DISSERTATION

zur Erlangung des Grades eines
Doktors der Wirtschaftswissenschaft
(Dr. rer. pol.)

eingereicht an der
Fakultät für Wirtschaftswissenschaften
der Universität Regensburg

Automatisierte Erschließung der Erkenntnispotentiale digitaler Daten

Konzeption und Entwicklung intelligenter Analysemethoden und Werkzeuge zur Operationalisierung nutzergenerierter Daten mittels Data Science und generativer Künstlicher Intelligenz

vorgelegt von
Janik Wörner (M. Sc. Wirtschaftsinformatik)

Prüfer/in
– Prof. Dr. Susanne Leist –
– Prof. Dr. Gregor Zellner –

Regensburg, 02.12.2025

Inhaltsverzeichnis

Genderhinweis

Zur besseren Lesbarkeit wird in dieser Arbeit bei Personenbezeichnungen in der Regel die grammatikalisch männliche Form (generisches Maskulinum) verwendet. Diese Entscheidung dient ausschließlich der sprachlichen Vereinfachung und impliziert keinerlei inhaltliche Gewichtung. Selbstverständlich sind damit stets alle Geschlechter gleichermaßen adressiert – Frauen, Männer sowie nicht-binäre Personen. Die gewählte Sprachform soll ausdrücklich alle Leserinnen und Leser gleichermaßen ansprechen und einschließen.

Abbildungsverzeichnis

Tabellenverzeichnis

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AAC	***A**ugmentative and **A**lternative **C**ommunication*
AISeL	***AIS E**lectronic **L**ibrary*
BERT	***B**idirectional **E**ncoder **R**epresentations from **T**ransformers*
CNN	***C**onvolutional **N**eural **N**etwork*
DESRIST	*International Conference on **De**sign **S**cience **R**esearch in **I**nformation **S**ystems and **T**echnology*
DMR	***D**irichlet **M**ultinomial **R**egression*
DSR	***D**esign **S**cience **R**esearch*
ECIS	***E**uropean **C**onference on **I**nformation **S**ystems*
EM	***E**lectronic **M**arkets*
FF	***F**orschungs**f**rage*
GPT	***G**enerative **P**re-trained **T**ransformers*
GQM	***G**oal **Q**uestion **M**etric*
HICSS	***H**awaii **I**nternational **C**onference on **S**ystem **S**ciences*
ICIS	***I**nternational **C**onference on **I**nformation **S**ystems*
IS	***I**nformations**s**ysteme*
KI	***K**ünstliche **I**ntelligenz*
LDA	***L**atent **D**irichlet **A**llocation*
LLM	***L**arge **L**anguage **M**odel*
MANTRA	***MA**rketi**N**g **TR**end **A**nalysis*
ML	***M**achine **L**earning*
PAM	***P**achinko **A**llocation **M**odel*
SDT	***S**elf **D**etermination **T**heory*
SNA	***S**ocial **N**etwork **A**nalysis*
SUMI	***S**oftware **U**sability **M**easurement **I**nventory*
UGC	***U**ser **G**enerated **C**ontent*
V	***V**eröffentlichung*
VoC	***V**oice **o**f the **C**ustomer*
ZS	***Z**ielsetzung*

1 Einleitung

Die vorliegende Dissertation vereint zwei thematisch komplementäre Forschungsfelder, die die Konzeption und Umsetzung intelligenter Analysemethoden und Werkzeuge zur Erschließung der in digitalen, nutzergenerierten Daten vorliegenden Erkenntnisspotentiale adressieren. Beide Forschungsstränge eint das Ziel, aktuelle Herausforderungen im Umgang mit diesen Daten durch den gezielten Einsatz von Methoden des Data Science sowie der Künstlichen Intelligenz (KI) zu bewältigen. Dabei wird ein besonderer Fokus auf die methodische Fundierung sowie die praxisorientierte Umsetzung gelegt.

Im ersten Forschungsgebiet stehen ökonomische Fragestellungen im Vordergrund (vgl. *Kapitel 1.1.1*). Untersucht wird, wie Daten aus sozialen Medien für die Erschließung der darin enthaltenen Erkenntnisspotentiale zur Unterstützung marketingorientierter Entscheidungsprozesse operationalisiert werden können. Hierzu werden datengetriebene Analysemethoden und Werkzeuge konzipiert, implementiert und evaluiert, die es Unternehmen ermöglichen, aus den unstrukturierten Kommunikationsdaten sozialer Medien verwertbare Informationen abzuleiten.

Das zweite Forschungsgebiet erweitert die Perspektive um eine gesellschaftliche Dimension und befasst sich mit der Entwicklung KI-gestützter Assistenzsysteme (vgl. *Kapitel 1.1.2*). Im Zentrum steht die Gestaltung adaptiver, nutzerzentrierter digitaler Kommunikationsagenten in den Bereichen der *Augmentative and Alternative Communication* (AAC)-Systeme sowie der gesundheitsbezogenen Chatbots.

Beide Forschungsbereiche tragen somit zur methodischen (Weiter-)Entwicklung und praktischen Anwendung digitaler Methoden und Werkzeuge im Bereich der Informationssysteme (IS) sowie angrenzender Forschungsdisziplinen bei. Sie zeigen exemplarisch, wie datengetriebene Technologien sowohl unternehmerischen Mehrwert stiften als auch gesellschaftliche Inklusion und individuelle Autonomie fördern können.

1.1 Motivation, Problemstellungen und Zielsetzungen

Die nachfolgenden Unterkapitel bilden die inhaltliche Grundlage der beiden zentralen Forschungsgebiete dieser Dissertation. Hierzu werden zunächst die in der aktuellen Forschung und Praxis vorherrschenden Problemstellungen systematisch aufgearbeitet, die bislang nur unzureichend adressiert wurden. Ausgehend von diesen Herausforderungen werden klar abgegrenzte Zielsetzungen formuliert, welche die im Rahmen dieser Dissertation verfolgten Untersuchungen strukturieren und inhaltlich rahmen.

1.1.1 Konzeption und Entwicklung kontextsensitiver Analyseverfahren

Soziale Medien haben sich in den letzten Jahren zu zentralen Infrastrukturen gesellschaftlicher Kommunikation entwickelt. Plattformen wie das soziale Kurznachrichten-Netzwerk X (ehemals Twitter), diskussionsbasierte Foren wie Reddit oder Bewertungsportale wie Yelp ermöglichen es Nutzern weltweit, Meinungen, Erfahrungen und Emotionen in Echtzeit öffentlich zu teilen (Hanna et al., 2011; Kietzmann et al., 2011). Daraus entstehen täglich enorme Mengen sogenannter nutzergenerierter Inhalte (*eng.: User-Generated Content*, UGC), die weit über alltägliche Kommunikation hinausreichen.

In der wissenschaftlichen Literatur wird der Begriff „Social Media" als Gruppe von internetbasierten Anwendungen definiert, welche auf den ideologischen und technologischen Grundlagen des Web 2.0 aufbauen und die Erstellung sowie den Austausch von UGC ermöglichen (Kaplan & Haenlein, 2010). Der hierdurch generierte UGC umfasst dabei bspw., wie Menschen Produkte bewerten (Mir & Rehman, 2013), politische Entscheidungen kommentieren (Inguanzo et al., 2021) oder auf gesellschaftliche Ereignisse reagieren (Wilson & Mihalcea, 2019) – und liefert dadurch aktuelle, kontextspezifische und umfassende Einblicke mit unmittelbarem Bezug zum sozialen Alltag und Leben der Menschen. In seiner Gesamtheit fungiert UGC somit nicht nur als Ausdruck individueller Perspektiven, sondern kann als Spiegel kollektiver Wahrnehmungen, geteilter Problemlagen und gesellschaftlicher Stimmungsbilder gesehen werden.

Allein auf X werden täglich hunderte Millionen Beiträge veröffentlicht (Radaideh & Dweiri, 2024); ein Datenvolumen, das traditionelle Instrumente wie Befragungen oder Interviews weder erfassen noch in Echtzeit auswerten können. Dieser dynamische und kontinuierlich wachsende Datenstrom und die damit einhergehenden Potenziale zur Analyse der darin enthaltenen kollektiven Meinungs- und Stimmungsbilder, markiert einen epistemischen Wandel (Neuberger et al., 2023): Wissen über gesellschaftliche Trends oder unternehmensbezogene Wahrnehmungen entsteht zunehmend nicht mehr nur durch klassische Erhebungsmethoden, sondern durch die Analyse digitaler Kommunikationsdaten – vorausgesetzt, es stehen geeignete methodische Zugänge zur Verfügung.

Genau hier zeigt sich die zentrale Herausforderung: Klassische Verfahren der Inhaltsanalyse, wie z.B. aus der Kommunikations- oder Sozialforschung (z. B. Mayring & Fenzl, 2019), stoßen bei der Verarbeitung großer Mengen unstrukturierter Textdaten schnell an ihre Grenzen. Sie sind weder für die Skalierung auf Massendaten noch für die Berücksichtigung kontextueller, semantischer und affektiver Dimensionen spezialisiert. Datenanalytische Verfahren aus dem Bereich des Data Science – insbesondere des maschinellen Lernens (ML)

– bieten hier neue Perspektiven. Sie ermöglichen es, große Mengen unstrukturierter Texte automatisiert zu strukturieren, die darin enthaltenen thematischen Muster zu identifizieren und so Bedeutungen im jeweiligen Kommunikationskontext zu erschließen (Shen et al., 2023; Yan et al., 2014). Auf diese Weise wird ein strukturierter Zugang zu den bislang oftmals verborgenen Informationen innerhalb UGC geschaffen und so ein tieferes Verständnis digitaler Kommunikationsdynamiken ermöglicht.

Gleichzeitig zeigt sich mit Blick auf die aktuelle Forschung und Praxis, dass trotz der Potenziale datenanalytischer Verfahren, deren tatsächliche Anwendbarkeit in vielen Anwendungsfeldern begrenzt bleibt. Dies gilt insbesondere für dynamische, datenintensive Domänen – wie z. B. dem Digitalen Marketing (Arora et al., 2022) –, in denen sich wettbewerbliche Rahmenbedingungen in hoher Geschwindigkeit verändern. Zwar existieren bereits zahlreiche Ansätze zur Analyse von UGC (vgl. z. B. Park et al., 2021, Bhor et al., 2018), doch sind diese in ihrer methodischen Ausgestaltung oftmals nicht spezifisch auf die Anforderungen der konkreten Nutzungsszenarien zugeschnitten (Reisenbichler & Reutterer, 2019). Hieraus resultiert die Problematik, dass viele Verfahren entweder zu generisch, zu voraussetzungsreich oder zu unflexibel sind, um die semantischen Dynamiken digitaler Kommunikationsdaten adäquat abzubilden.

In der Folge entsteht eine methodische Lücke zwischen dem analytischen Potenzial großer, unstrukturierter Datenströme an UGC und deren tatsächlicher domainspezifischer Nutzbarmachung. Vor diesem Hintergrund ergibt sich eine übergreifende Problemstellung, welche die Ausgangslage des ersten Forschungsgebiets dieser Dissertation markiert und das Promotionsprojekt visionär leitet:

> *Die hohe Vielfältigkeit datengetriebener Anwendungsbereiche führt dazu, dass für viele Problemstellungen keine hinreichend methodisch ausgereiften und praktisch umgesetzten datenanalytischen Verfahren und softwaregestützte Werkzeuge existieren, die die jeweiligen domänenspezifischen Anforderungen gezielt adressieren. Daraus resultiert die Notwendigkeit kontextsensitiver, skalierbarer und interpretierbarer Analyseansätze und Werkzeuge.*

Mit dieser Problemstellung verbindet sich der Anspruch der Dissertation, methodisch-technische (Weiter-)Entwicklungen datenanalytischer Verfahren und Werkzeuge zu leisten und diese entlang konkreter Anwendungsbedarfe neu zu denken, zu gestalten und zu evaluieren. Die Dissertation verfolgt im ersten Forschungsgebiet dafür insgesamt acht miteinander verzahnte Zielsetzungen (ZS1 bis ZS8), auf welche folgend detailliert eingegangen wird.

1. Forschungsgebiet: Herleitung der Zielsetzungen ZS1 bis ZS8.

Während sich die Interaktion zwischen Unternehmen und Kunden traditionell auf klassische Kommunikationskanäle wie Ladengeschäft, E-Mail oder Telefon stützte, eröffnen soziale Medien heute neue Möglichkeiten der unmittelbaren und öffentlichen Meinungsäußerung. In sozialen Netzwerken kommunizieren Nutzer ihre Erfahrungen, Erwartungen und Bewertungen zu Produkten und Dienstleistungen eines Unternehmens in Echtzeit und auf breiter öffentlicher Bühne. Diese spezielle Form der direkten, nutzergenerierten Kommunikation wird in der einschlägigen Literatur unter dem Konzept der *Voice of the Customer* (VoC) gefasst (Pande, 2014). Durch die Analyse der VoC wird es Unternehmen ermöglicht, differenzierte Einblicke in die Bedürfnisse, Ideen und Probleme aktueller sowie potenzieller Kundengruppen zu gewinnen (Hong et al., 2012; Lozano et al., 2017).

Gerade im digitalen Marketing ergeben sich durch die Analyse solcher VoC-Daten große Potenziale. So können Wahrnehmungen identifiziert, ihre zeitliche Entwicklung nachvollzogen und konkrete Rückschlüsse auf Produktbewertungen oder Markenwahrnehmungen gezogen werden (Cirqueira et al., 2017; Irawan et al., 2020; Endo et al., 2015). In einem zunehmend dynamischen und wettbewerbsintensiven Marktumfeld stellt dieser frühzeitige Erkenntnisgewinn einen zentralen strategischen Wettbewerbsfaktor dar (vgl. Zhong & Schweidel, 2020; Park et al., 2021).

Die Erschließung dieser Erkenntnisspotentiale ist jedoch mit erheblichen Herausforderungen verbunden. So liegen die Kommunikationsdaten in unstrukturierter Form vor, sind inhaltlich dynamisch und kontextsensitiv. In der Praxis kommen zur Analyse dieser VOC-Daten bislang vor allem schlagwortbasierte Verfahren zum Einsatz; z. B. im Rahmen gängiger Social Listening Tools wie *Brandwatch*, *Meltwater* oder *Symanto*. Diese Tools stützen sich auf die Frequenz vordefinierter Begriffe oder Hashtags (Lau et al., 2012), wodurch latente Kookkurrenzen verborgen bleiben und neue thematische Zusammenhänge nur unzureichend erfasst werden können (Khan et al., 2021; Reisenbichler & Reutterer, 2019).

Vor diesem Hintergrund rücken datenanalytische Verfahren zur automatisierten Inhaltsanalyse – insbesondere Topic Modelling – zunehmend in den Fokus wissenschaftlicher und praktischer Auseinandersetzungen. Topic Modelling erlaubt die probabilistische Identifikation thematischer Strukturen in Textdaten, ohne dass zuvor manuell Schlagwörter definiert werden müssen (Blei, 2012). Damit eignet sich dieser Ansatz insbesondere für die Analyse in explorativen Anwendungsszenarien mit hoher, heterogener Themenvielfalt und geringer

Strukturierung (Crain et al., 2012; Debortoli et al., 2016; Reisenbichler & Reutterer, 2019). Als Standardverfahren hat sich dabei insbesondere die Latent Dirichlet Allocation (LDA) etabliert (Blei et al., 2003; Blei, 2012; Eickhoff & Neuss, 2017).

LDA stößt jedoch im Rahmen marketingbezogener Anwendungsszenarien regelmäßig an methodische Grenzen. Das Modell berücksichtigt weder externe Kontextfaktoren (z. B. zeitliche oder geografische Gegebenheiten) noch bildet es dynamische Entwicklungen oder hierarchische Themenstrukturen adäquat ab (Lozano et al., 2017; Zhang et al., 2016). Gerade diese Anforderungen sind jedoch für viele domänenspezifische Anwendungsfälle im digitalen Marketing essenziell (Bae et al. 2018; Ding et al. 2020; Endo et al. 2015; Ha et al. 2017; Moro et al. 2020; Park et al. 2021). In der Folge wurde in der Forschung eine Vielzahl alternativer oder erweiterter Modellvarianten entwickelt.

Trotz der wachsenden Methodenvielfalt mangelt es bislang jedoch an systematischen, vergleichenden Analysen, welche diese Topic Modelling Verfahren im Hinblick auf die marketingspezifischen Anforderungen (z. B. der Trendanalyse) evaluieren (Vakansky & Kumar, 2020). Diese fehlende Transparenz über Leistungsfähigkeit, Einsatzgrenzen und kontextuelle Passung verhindert eine fundierte und zielgerichtete Auswahl geeigneter Verfahren. Hinzu kommt, dass ein Großteil der in der Forschung vorgeschlagenen Verfahren bislang auf konzeptioneller Ebene verbleiben. So existieren zwar zahlreiche theoretische Modellansätze des Topic Modelling, doch fehlt es häufig an praxistauglichen, implementierten Lösungen. Entweder stehen keine lauffähigen Implementierungen zur Verfügung oder ihre Anwendung setzt technisches Fachwissen voraus, das in vielen Marketingabteilungen nicht in ausreichendem Maße vorhanden ist (Vayansky & Kumar, 2020; Wörner et al., 2021). Diese Hürden erschweren die Nutzung von Topic Modelling und limitieren somit die methodische Weiterentwicklung durch die fehlende Erschließung anwendungsbezogener Erkenntnisspotentiale. Aus dieser Problemstellung ergibt sich die erste Zielsetzung dieser Dissertation:

Bezeichnung	**Zielsetzung 1 (ZS1)**
Beschreibung	Die systematische Evaluation bestehender Topic Modelling Verfahren im Hinblick auf ihre Eignung zur Analyse von UGC im digitalen Marketing. Ziel ist es, Transparenz über Leistungsfähigkeit, Einsatzgrenzen und kontextuelle Passung unterschiedlicher Topic Modelling Verfahren zu schaffen und deren Anwendbarkeit für konkrete Anwendungsszenarien (z. B. Trendanalyse) evidenzbasiert zu bewerten. Damit soll

	eine fundierte Entscheidungsgrundlage für die methodische Auswahl geeigneter Verfahren geschaffen werden.
Abhängigkeiten	Die Beantwortung dieser Zielsetzung informiert die Bearbeitung der Zielsetzung **ZS4** und **ZS6**, da die gewonnen Erkenntnisse mit in die Entwicklung der darin adressierten softwaregestützten Werkzeuge einfließen.

Tabelle 1. Beschreibung der 1. Zielsetzung.

Neben der methodischen Auswahl geeigneter Verfahren zur marketingorientierten Erschließung von VoC-Daten (vgl. ZS1) ergibt sich eine weiterführende Herausforderung: Die semantische Tiefe der in der VoC artikulierten Inhalte ist oftmals nicht allein durch inhaltsbezogene Analysen adäquat zu erfassen. (Digitale) Kommunikation entsteht nicht kontextfrei – sie ist stets eingebettet in zeitliche, geografische und kulturelle Rahmenbedingungen, welche die semantische Bedeutung einzelner Aussagen beeinflussen (vgl. Kegler et al., 2010; Green, 2015).

Gerade im digitalen Marketing reicht eine rein inhaltliche Analyse deshalb häufig nicht aus, um gewonnene Informationen aus VoC-Inhalten differenziert zu interpretieren. Wahrnehmungen und Bedürfnisse von Konsumenten im Hinblick auf Qualität, Preis-Leistungs-Verhältnisse, Serviceansprüche oder Nachhaltigkeitsorientierungen variieren erheblich in Abhängigkeit von regionalen, kulturellen und sozialen Faktoren (vgl. Cao et al., 2018; Cruz-Albrecht et al., 2017; Dodds et al., 2011; Lim et al., 2018; Lin, 2014; Plunz et al., 2019; Schwartz et al., 2019). Die Berücksichtigung solcher Kontextdimensionen ist daher essenziell, um aus VoC-Daten nicht nur thematische, sondern auch kontextuell verankerte Erkenntnisse für spezifische Zielgruppen, Zeiträume oder Kulturen zu gewinnen.

Obwohl digitale Plattformen kontextuelle Informationen (z. B. Zeitstempel, Standortdaten oder demografische Merkmale), mangelt es bislang an datenanalytischen Verfahren, die diese Kontextfaktoren automatisiert erfassen, strukturieren und in komplementäre Analyseprozesse überführen. Dies betrifft insbesondere dynamische Fragestellungen wie bspw. die Analyse thematischer Entwicklungen im Zeitverlauf oder über verschiedene Lokationen hinweg. Im Resultat bleiben kontextsensitive Erkenntnisspotentiale, welche z. B. zur zielgerichteten Kundensegmentierung im Rahmen differenzierter Marketingkampagnen benötigt werden, ungenutzt. Vor diesem Hintergrund adressiert die Dissertation eine zweite zentrale Zielsetzung:

Bezeichnung	**Zielsetzung 2 (ZS2)**
Beschreibung	Die Konzeption und Implementierung eines kontextsensitiven Analyseverfahrens, das es ermöglicht, zeitliche, geografische und demografische Kontextdimensionen systematisch in UGC zu identifizieren, zu extrahieren und in weiterführende Analyseverfahren zu integrieren. Ziel ist es, eine differenziertere Analyse der VoC durch die Integration externer Kontextfaktoren (z. B. Zeit, Geolokation) zu ermöglichen, die über eine rein inhaltlich-thematische Perspektive hinausgeht.
Abhängigkeiten	Die Beantwortung dieser Zielsetzung informiert die Bearbeitung der Zielsetzung **ZS4** und **ZS6**, da die gewonnen Erkenntnisse mit in die Entwicklung der darin fokussierten softwaregestützten Werkzeuge einfließen.

Tabelle 2. Beschreibung der 2. Zielsetzung.

Gerade mit Blick auf die Analyse von Trends und deren zeitliche oder geografische Entwicklung ist die Anwendung geeigneter Analyseverfahren, welche die Integration externer Kontextfaktoren berücksichtigen, von zentraler Bedeutung. Trends stellen komplexe, mehrdimensionale Phänomene dar, die sich im Zeitverlauf entwickeln, transformieren und weitreichende Auswirkungen auf gesellschaftliche und wirtschaftliche Strukturen entfalten können. Für Unternehmen ergibt sich daraus die strategische Notwendigkeit, solche Entwicklungen frühzeitig zu identifizieren, kontinuierlich zu beobachten und in Entscheidungen zu überführen (Endo et al., 2015; Lozano et al., 2017).

Soziale Medien bieten in diesem Zusammenhang einen besonders wertvollen Resonanzraum: Sie erlauben es ihren Nutzern, neue Entwicklungen (z. B. in Bereichen wie Technologie oder Lifestyle) offen zu diskutieren und ihre Wahrnehmungen öffentlich zu teilen (Jeong et al., 2019; Park et al., 2021; Yan et al., 2014). Dieser dynamische Diskurs ermöglicht es Unternehmen, relevante Trendthemen zu identifizieren, ihre potenziellen Entwicklungsrichtungen zu antizipieren und daraus Implikationen für die Gestaltung künftiger Leistungsangebote abzuleiten (Tucker & Kim, 2011; Zhong & Schweidel, 2020). Die am Markt verfügbaren Softwarelösungen zur Trendanalyse weisen in diesem Kontext jedoch deutliche methodische und funktionale Einschränkungen auf (vgl. Wörner et al., 2023). Insbesondere fehlt es an der Möglichkeit, geografische (z. B. zur Unterscheidung lokationsabhängiger Kundenwahrnehmungen), zeitliche (z. B. zur Berücksichtigung saisonaler Effekte) und demografische (z. B. zur Identifikation divergierender Anforderungen der Zielgruppen)

Kontextfaktoren flexibel und anwendungsbezogen in die Analyse zu integrieren. Dabei hat die Forschung gezeigt, dass sich die Wahrnehmungen von Kunden im Hinblick auf Produkte und Dienstleistungen teilweise erheblich über verschiedene Kontextfaktoren hinweg unterscheiden (vgl. Bae et al., 2018; Ding et al., 2020; Endo et al., 2015; Ha et al., 2017; Moro et al., 2020; Park et al., 2021). Im Resultat besteht ein klarer Bedarf an Trendanalysetools, die die systematische Einbindung und Nutzung der Kontextfaktoren ermöglichen.

Neben dieser Berücksichtigung kontextueller Einflussfaktoren müssen Trendanalysetools weitere Anforderungen erfüllen, damit Marketingabteilungen die aus der VoC extrahierten Informationen effektiv nutzen können (vgl. Wörner et al., 2022; Wörner et al., 2023). Eine besonders zentrale Anforderung betrifft die Analyse der Tonalitäten. Für die Bewertung von Trendthemen ist nicht nur entscheidend, *worüber* gesprochen wird, sondern auch, *wie* darüber gesprochen wird. Erst durch die Analyse der in den VoC zum Ausdruck gebrachten Stimmungen lassen sich Rückschlüsse auf die Wahrnehmung und Akzeptanz bestimmter Themen (z. B. spezifische Produkt- oder Servicemerkmale) ziehen. Verfahren der Sentimentanalyse ermöglichen es, Themen nach ihrer Tonalität zu klassifizieren – z. B. danach, ob sie positiv, neutral oder negativ diskutiert werden (Irawan et al., 2020; Jeong et al., 2019; Mastrogiacomo et al., 2021; Moro et al., 2020; Tucker & Kim, 2011). Diese Informationen sind von zentraler Bedeutung, um fundierte Einschätzungen zur Kundenwahrnehmung zu treffen und daraus strategisch relevante Maßnahmen abzuleiten.

Obwohl für einzelne Dimensionen der Trendanalyse, wie z. B. der Identifikation inhaltlich diskutierter Themen (z. B. *Topic Modelling*) oder der tonalitätsbasierten Identifikation der zugrundeliegenden Stimmungen (z. B. *Sentimentanalyse*), bereits spezialisierte Verfahren existieren, fehlt es bislang an integrierten Lösungen, die diese Analyseperspektiven in einem ganzheitlichen Rahmen integrieren. Die derzeit verfügbaren Analyseansätze zur Trendanalyse adressieren die relevanten Dimensionen lediglich isoliert – es existiert kein Tool, das eine thematisch-kontextuelle und tonalitätsbasierte Analyse von VoC-Daten in einem domänenspezifischen System vereint. Somit fehlt es bislang an einem Werkzeug zur (marketinggetrieben) Trendanalyse, welches die damit einhergehenden Anforderungen umfassend unterstützt. Daraus ergeben sich zwei aufeinander aufbauende Zielsetzungen, welche die methodische Fundierung, die technische Umsetzung und die anwendungsbezogene Evaluation eines datenanalytischen Werkzeugs zur marketingspezifischen Trendanalyse betreffen:

Bezeichnung	**Zielsetzung 3 (ZS3)**
Beschreibung	Ableitung von Anforderungen an ein Trendanalysewerkzeuge mit dem Ziel, ein marketingorientiertes, kontext- und tonalitätsbasiertes softwaregestütztes Werkzeug auf Basis der Entwicklung und Kombination moderner Data Science Verfahren (z. B. Topic Modelling, Sentimentanalyse, Kontextanalyse) zu entwickeln, welches die ganzheitliche Betrachtung der relevanten Trendanalyseperspektiven ermöglicht.
Abhängigkeiten	Die im Rahmen von ZS3 gewonnenen Erkenntnisse informieren die nachfolgende Zielsetzung **ZS4** durch die Bereitstellung eines Anforderungsprofils des softwaregestützten Trendanalysewerkzeugs.

Tabelle 3. Beschreibung der 3. Zielsetzung.

Bezeichnung	**Zielsetzung 4 (ZS4)**
Beschreibung	Konzeption, prototypische Umsetzung und technische Evaluation eines integrierten Trendanalysewerkzeugs, das thematische, kontextuelle und tonalitätsbasierte Analyseperspektiven integriert und dadurch eine ganzheitliche, marketingrelevante Trendanalyse unterstützt.
Abhängigkeiten	Die Beantwortung dieser Zielsetzung wird durch ZS1, ZS2 und ZS3 informiert: Die gewonnenen Erkenntnisse über die Leistungsfähigkeit der Topic Modelling Verfahren (**ZS1**), die Entwicklung des Analyseverfahrens zur Identifikation unterschiedlicher Kontextfaktoren (**ZS2**) sowie die Anforderungen an ein solches softwaregestütztes Werkzeug (**ZS3**) informieren die Entwicklung dieses Trendanalysewerkzeuges.

Tabelle 4. Beschreibung der 4. Zielsetzung.

Insbesondere im Bereich der Produktentwicklung – einer spezifischen Anwendungsperspektive der Trendanalyse – ist es von zentraler Bedeutung, die in der VoC enthaltenen Informationen nicht nur thematisch zu erfassen, sondern auf Ebene spezifischer Produkteigenschaften und Leistungsmerkmale differenziert zu analysieren. Nur durch diese Aspekt-basierte Betrachtung wird erkennbar, welche konkreten Eigenschaften eines Produkts oder einer Dienstleistung von Kunden wahrgenommen werden, welche Bedürfnisse mit ihnen verknüpft sind und mit welchem Stimmungsbild diese Aspekte behaftet sind (vgl. Mirtalaie et al., 2018; Tucker & Kim, 2011; Tuarob & Tucker, 2015; Vo et al., 2018).

Eine geeignete methodische Kombination stellt in diesem Kontext der kombinierte Einsatz von Topic Modelling und Aspekt-basierter Sentimentanalyse dar: Während das Topic Modelling die thematische Strukturierung und Extraktion relevanter Produktmerkmale ermöglicht, erlaubt die anschließende Aspekt-basierte Sentimentanalyse eine differenzierte Analyse des Stimmungsbilds dieser Merkmale. Damit kann beispielsweise erfasst werden, ob z. B. die Verarbeitungsqualität eines Produkts positiv, die Benutzerfreundlichkeit hingegen negativ bewertet wird.

Das Potenzial Aspekt-basierter Sentimentanalysen zur Identifikation von Kundenwahrnehmungen zu spezifischen Produkt- und Leistungsaspekten wurde in der Wissenschaft und Praxis bereits erkannt (Schouten & Frasincar, 2015; Tuarob & Tucker, 2015; Tucker & Kim, 2011) und hat zur Integration entsprechender Funktionen in am Markt verfügbaren Trendanalysetools geführt (vgl. Wörner et al., 2022). Allerdings zeigen sich bei diesen Tools nach wie vor erhebliche Defizite: Insbesondere fehlt es bislang an Verfahren, die Kontextfaktoren (z. B. Geolokationen, zeitliche Entwicklungen) flexibel und in ausreichender Granularität in die Analyse einbeziehen. Die daraus resultierenden produktbezogenen Trendanalysen bleiben dadurch häufig oberflächlich und kontextblind (Tuarob & Tucker, 2015).

Darüber hinaus fehlt es den bestehenden Tools an der Fähigkeit, bereits vorliegendes, domänenspezifisches Wissen wie z. B. Produktinformationen in Form von Ontologien in die Analyse zu integrieren. Gerade dieses strukturierte Expertenwissen bietet jedoch das Potenzial, besonders relevante Aspekte gezielt analytisch zu adressieren und semantische Zusammenhänge damit zu verbinden, wodurch die analytische Tiefe und Relevanz der Trendanalyse gesteigert werden kann (Tuarob & Tucker, 2015; Tucker & Kim, 2011; Vo et al., 2018). Vor diesem Hintergrund ist die Entwicklung eines Werkzeugs erforderlich, das sowohl inhaltliche, kontextuelle als auch domänenspezifische Analyseperspektiven (z. B. Analyse der Tonalität) integriert und eine ganzheitliche, Aspekt-basierte Erschließung der VoC-basierten Erkenntnispotentiale für die Produktentwicklung ermöglicht.

Bezeichnung	**Zielsetzung 5 (ZS5)**
Beschreibung	Ableitung von Anforderungen an ein Trendanalysewerkzeuge mit Blick auf die spezifische Anwendungsperspektive der Produktentwicklung. Ziel ist es, ein Aspekt-basiertes und kontext- bzw. tonalitätssensitives Trendanalysewerkzeug durch die Entwicklung und Kombination moderner Data Science Verfahren zu konzipieren, welches die Betrachtung spezifischer Produkt- und Leistungsaspekte ermöglicht.

Abhängigkeiten	Die Beantwortung der Zielsetzung wird durch ZS3 und ZS4 informiert: Die gewonnenen Erkenntnisse über die Anforderungen (genereller) Trendanalysetools mit Marketingorientierung (**ZS3**) sowie die Konzeption und Entwicklung des softwaregestützten Trendanalysewerkzeugs (**ZS4**) informieren die Konzeption dieses Trendanalysewerkzeuge vor dem Hintergrund der Produktentwicklung. Die in der Bearbeitung der ZS5 gewonnenen Erkenntnisse informieren die nachfolgende Zielsetzung **ZS6** durch die Bereitstellung eines Anforderungsprofils des umzusetzenden Trendanalysewerkzeugs.

Tabelle 5. Beschreibung der 5. Zielsetzung.

Bezeichnung	**Zielsetzung 6 (ZS6)**
Beschreibung	Konzeption, prototypische Entwicklung und Evaluation eines Trendanalysewerkzeugs im Anwendungskontext der Produktentwicklung. Ziel ist es, durch die Kombination moderner Data Science Verfahren (z. B. Topic Modelling, Aspekt-basierte Sentimentanalyse, Word Embeddings) eine differenzierte Analyse der Kundenstimmungen gegenüber konkreten Produkt- und Leistungsaspekte – unter Berücksichtigung geografischer und zeitlicher Kontextfaktoren – zu ermöglichen. Auf diese Weise soll eine ganzheitliche Analyse relevanter Merkmale von Produkten und Dienstleistungen unterstützt werden.
Abhängigkeiten	Die Beantwortung der Zielsetzung wird durch ZS1 bis ZS5 informiert: Die gewonnenen Erkenntnisse über die Leistungsfähigkeit der Topic Modelling Verfahren (**ZS1**) und die Entwicklung des Analyseansatzes zur Identifikation unterschiedlicher Kontextfaktoren (**ZS2**) informieren die Entwicklung dieses Trendanalysewerkzeugs. Zudem liefert die Bearbeitung von **ZS3** und **ZS4** design- und entwicklungsbezogenes Wissen, welches die Entwicklung dieses Trendanalysewerkzeuges zusätzlich informiert. Zuletzt liefern die Anforderungen an ein solches softwaregestütztes Werkzeug (**ZS5**) anwendungsbezogene Informationen, um das Analysewerkzeug zu konzipieren und technisch zu realisieren.

Tabelle 6. Beschreibung der 6. Zielsetzung.

Neben der Analyse der VoC zur Identifikation von Trendthemen, rückt mit Blick auf die Produktentwicklung eine weitere Anwendungsperspektive in den Fokus der Erschließung datenanalytischer Erkenntnisspotentiale: die Identifikation relevanter Nutzergruppen, die häufig diskutierte (Trend-)Themen frühzeitig erkennen, mitgestalten und verbreiten. Diese Nutzer sind nicht nur Beobachter, sondern aktive Impulsgeber im digitalen Diskurs – mit potenziell erheblichem Einfluss auf Innovationsprozesse im Rahmen der Produktentwicklung (Al-Zu'bi & Tsinopoulos, 2012; Brem et al., 2018; Ye & Kankanhalli, 2018).

Im Zentrum steht hierbei das Konzept der Lead User. Diese zeichnen sich durch eine vorausschauende Bedürfniswahrnehmung, eine ausgeprägte Problemlösungskompetenz sowie einem breiten Wissen über innovative Themenfeldern aus (von Hippel, 1986; Hienerth & Lettl, 2017). Ihre frühzeitige Identifikation eröffnet Unternehmen die Möglichkeit, externe Wissensressourcen strategisch in interne Entwicklungs- und Innovationsprozesse zu integrieren – mit dem Ziel, die Marktorientierung neuer Produkte und Dienstleistungen zu steigern (Schaarschmidt et al., 2019; Al-Zu'bi & Tsinopoulos, 2012; Ye & Kankanhalli, 2018).

Die Herausforderung besteht jedoch darin, diese Nutzer in großen, heterogenen digitalen Netzwerken zu identifizieren. Vor allem im Kontext sozialer Medien – geprägt durch ein hohes Kommunikationsvolumen – ist die Identifikation von Lead Usern einer der aufwändigsten und zugleich kritischsten Schritte innerhalb der Lead User Methode (Brem & Bilgram, 2015). In der Forschung wurden daher verschiedene Ansätze zur Identifikation von Lead Usern entwickelt. Diese decken jedoch nur einzelne Aspekte des Problems ab. So fokussieren sich viele Verfahren ausschließlich auf einzelne Lead User Merkmale wie die hohe Aktivität (Martínez-Torres, 2014) oder können aufgrund des manuellen Aufwands nur auf eine begrenzte Stichprobengrößen angewandt werden (Hau & Kang, 2016). Andere basieren auf aufwändigen Umfragen, Beobachtungsdaten oder Selbstauskünften (Brandtzaeg et al., 2016; Hung et al., 2011; Tuunanen et al., 2011), was zu geringer empirischer Belastbarkeit durch subjektiv-verzerrte Einschätzungen sowie hohem methodischen Aufwand führt (Hienerth & Lettl, 2017). Solche manuellen Ansätze widersprechen dem kurzfristigen, trendabhängigen Charakter der Lead User – was eine dynamische, datenbasierte Identifikation erforderlich macht. Im Resultat erfordert es einen Ansatz, der große Mengen UGC vor dem Hintergrund des Innovationsprozesses automatisiert analysieren und dabei alle in der Forschungsliteratur beschriebenen Merkmale von Lead Usern systematisch berücksichtigen kann. Ein solcher Identifikationsansatz fehlt bislang in der aktuellen Forschung. Die methodische Konzeption und die technische Realisierung dieses automatisierten Identifikationsansatzes begründen die zwei abschließenden Zielsetzungen des ersten Forschungsgebiets:

Bezeichnung	**Zielsetzung 7 (ZS7)**
Beschreibung	Ableitung und methodische Konzeption eines Anforderungsprofils für die datenbasierte Identifikation von Lead Usern. Hierzu sollen zentrale Lead User Merkmale (z. B. thematische Aktivität, Trendsetter-Eigenschaft) in der Literatur identifiziert sowie systematisch modelliert und softwaregestützt operationalisiert werden. Ziel ist es, diese Merkmale in digitalen Kommunikationsdaten automatisiert zu erfassen und für die softwaregestützte Lead Usern Identifikation nutzbar zu machen.
Abhängigkeiten	Die im Rahmen der Bearbeitung der ZS7 gewonnenen Erkenntnisse informieren die nachfolgende Zielsetzung **ZS8** durch die Bereitstellung eines methodischen Analyseansatzes, durch welchen die relevanten Lead User Merkmale innerhalb des zu entwickelnden automatisierten Analyseverfahrens technisch realisiert und operationalisiert werden.

Tabelle 7. Beschreibung der 7. Zielsetzung.

Bezeichnung	**Zielsetzung 8 (ZS8)**
Beschreibung	Entwicklung und Evaluation des konzipierten Analyseverfahrens zur automatisierten Identifikation von Lead Usern auf Basis von UGC-Daten. Der Ansatz soll multimethodische Analyseverfahren zu einem umfassenden Analysewerkzeug integrieren und dabei etablierte Data Science Verfahren (z. B. Topic Modelling zur Trendanalyse, Sentimentanalyse zur tonalitätsbasierten Bewertung sowie Social Network Analysis (SNA) zur Abbildung der Netzwerkstrukturen) zur automatisierten, merkmalbasierten Identifikation der Lead User kombinieren.
Abhängigkeiten	Die Zielsetzung wird durch ZS4, ZS6 und ZS7 informiert: Die gewonnenen Erkenntnisse im Rahmen der Bearbeitung von **ZS4** und **ZS6** liefern design- und entwicklungsbezogenes Wissen, welches die Entwicklung dieses softwaregestützten Werkzeuges informiert. Darüber hinaus liefern die spezifischen Anforderungen des Identifikationsverfahren (**ZS7**) relevante Informationen, um das Analysewerkzeug zur automatisierten Identifikation von Lead Usern zu konzipieren, technisch zu realisieren und schlussendlich zu evaluieren.

Tabelle 8. Beschreibung der 8. Zielsetzung.

1.1.2 Konzeption und Entwicklung intelligenter Assistenzsysteme

Während das erste Forschungsgebiet dieser Dissertation einen ökonomisch geprägten Kontext betrachtet, erweitert das zweite Forschungsgebiet die Entwicklung datengetriebener Analyseverfahren um eine gesellschaftliche Perspektive. Im Zentrum steht die Entwicklung nutzerzentrierter KI-gestützter digitaler Assistenzsysteme, welche neue Formen der adaptiven und personalisierten Interaktion ermöglichen sollen.

KI zählt zu den bahnbrechendsten Themen des 21. Jahrhunderts und hat sowohl in der wissenschaftlichen Literatur (vgl. Lu et al., 2018; Russell et al., 2015) als auch in der gesellschaftlichen Diskussion (vgl. Fast & Horvitz, 2017) erhebliche Aufmerksamkeit erfahren. So hat sich KI mittlerweile zu einem hochaktuellen Forschungsgegenstand entwickelt, der die Art und Weise, wie wir Technologie nutzen, domänenübergreifend revolutioniert hat. Laut einer aktuellen Studie des Pew Research Centers sind Experten insbesondere von den Chancen überzeugt, die KI in Bereichen wie dem Gesundheitswesen, der Automatisierung und der Kommunikation bietet (Pew Research Center, 2025).

Trotz dieser Chancen sehen sich bestehende intelligente Assistenzsysteme noch immer vor großen Herausforderungen hinsichtlich der Nutzerzentrierung. Dies betrifft insbesondere Anwendungen im Bereich der AAC und der digitalen Gesundheitsberatung. Trotz des KI-basierten technologischen Fortschritts bestehen in vielen dieser Systemen noch immer Einschränkungen hinsichtlich der individuellen Integration der Nutzer, worunter deren soziale Anschlussfähigkeit und letztendlich deren Inklusionsfähigkeit geschmälert wird. Damit bleibt das Potenzial bislang ungenutzt, generative KI nicht nur als funktionales Analysewerkzeug, sondern als individuelles, adaptives Kommunikationsmedium zu gestalten. Vor diesem Hintergrund ergibt sich die übergreifende Problemstellung, die das zweite Forschungsgebiet dieser Dissertation motiviert:

> *Trotz des Fortschritts generativer KI fehlt es an fundierten Konzepten und funktionsfähigen Implementierungen, die es ermöglichen, nutzerbezogene Informationen systematisch in personalisierte Assistenzsysteme zu überführen. Daraus resultiert die Notwendigkeit, intelligente Assistenzsysteme als adaptive, nutzerzentrierte Werkzeuge zur Inklusions- und Gesundheitsförderung neu zu denken und diese zu konzipieren und technisch unter Zuhilfenahme generativer KI zu realisieren.*

Die Dissertation verfolgt vor diesem Hintergrund vier inhaltlich aufeinander abgestimmte Zielsetzungen (ZS9 bis ZS12), die im Folgenden detailliert erläutert werden.

2. Forschungsgebiet: Herleitung der Zielsetzungen ZS9 bis ZS12.

Im Kontext menschlicher Kommunikation beschreibt Hourcade et al. (2004) den Austausch von Gedanken, Bedürfnissen, Wünschen und Gefühlen, als eine grundlegende Eigenschaft des Menschseins (vgl. Hourcade et al., 2004). Lautsprachliche Kommunikation wird dabei als zentrales Instrument eines Individuums verstanden, um mit der Umwelt zu interagieren (vgl. Fritzell, 1996; Kane et al., 2017; Pollak & Gallagher, 1989). Personen, die unter schweren körperlichen Behinderungen (z. B. Hirnverletzungen) oder Sprachstörungen wie Aphasie oder Autismus leiden, sind jedoch oft nicht in der Lage, ihre Muskulatur für die Lautbildung in ausreichendem Maße zu kontrollieren (Allen, 2005; Vanderheiden, 1983). Um diesen Menschen dennoch die Fähigkeit zur lautsprachlichen Kommunikation zu ermöglichen, sind alternative Kommunikationsformen notwendig, welche verschiedene Strategien und Techniken zur Förderung effektiver Kommunikation kombinieren (American Speech Language Hearing Association, 1991).

Solche digitalen AAC-Systeme wurden entwickelt, um Menschen mit Sprachbehinderungen die Möglichkeit zu bieten, lautsprachlich zu kommunizieren. Diese Softwareanwendungen übersetzen den Kommunikationsinhalt einer Person (generiert z.B. durch die Wahl entsprechender Piktogramm-Tasten oder der Tastaturfunktion [Mamlekar et al., 2023]) in Sprache und verleihen den Betroffenen auf diese Weise eine Stimme, um ihre Gedanken, Wünsche und Bedürfnisse lautsprachlich zu artikulieren (vgl. Bradshaw, 2013). Der positive Einfluss solcher AAC-Systeme auf das Kommunikationsverhalten von Menschen mit Sprachbehinderungen wurde bereits in mehreren Studien bestätigt (z. B. Brady, 2000; Desai et al., 2014; Schepis et al., 1996). Dennoch weisen die bestehenden Systeme signifikante Probleme auf, die den Nutzern in alltäglichen Gesprächssituationen Herausforderungen bereiten. Dazu gehören langsame Gesprächsraten, ein hoher Aufwand bei der Erstellung von Beiträgen und die unzureichende Anpassung an die individuellen Bedürfnisse und Persönlichkeiten der Nutzer (Dempster et al., 2010; Laxmidas et al., 2021; Obiorah et al., 2021).

Ein zentrales Problem besteht in der fehlenden Kontextualisierung gegenüber den unterschiedlichen Gesprächsphasen und -kontexten (Konadl et al., 2023). Derzeit ermöglichen es AAC-Systeme ihren Nutzern nicht, ihre Aussagen oder ihren Kommunikationsstil flexibel an den Kontext eines Gesprächs (z. B. in einer formellen oder informellen Situationen) oder an die unterschiedlichen Phasen eines Gesprächs (z. B. den Beginn oder das Ende einer Unterhaltung) anzupassen. Diese Inflexibilität führt zu einer mangelnden Personalisierung der Kommunikation, da die Nutzer nicht in der Lage sind, ihre Kommunikation nach ihren individuellen Vorlieben an die jeweiligen Gesprächssituationen anzupassen. So kann es für eine Person mit Sprachstörung bspw. bedeutend sein, sich in einem formellen Kontext besonders präzise oder eloquent zu artikulieren, während in einem informellen, legeren Umfeld eine alltagsnahe und entspannendere Sprache bevorzugt werden könnte (vgl. Bedrosian et al. 2003; Budde 2012; Elspaß 2018; Hoag et al. 2004; Todman & Alm 1997). Ohne die Möglichkeit, diese Nuancen in der Kommunikation zu berücksichtigen, bleibt die Ausdruckskraft des Nutzers eingeschränkt, was langfristig negative Auswirkungen auf seine Alltagsinklusion haben kann.

Der Einsatz von KI verspricht vielversprechende Lösungen, um diese Herausforderungen zu adressieren. Obwohl bereits zahlreiche wissenschaftliche Arbeiten den Einsatz von KI im Bereich der AAC-Systeme untersucht haben, fehlt es an einem strukturierten Überblick über die realisierten Potenziale und die noch unerforschten Möglichkeiten, KI zur Berücksichtigung differenzierter Gesprächsphasen und -Kontexte zu nutzen. Ein solcher Überblick kann helfen, Forschungslücken zu identifizieren und neue Lösungsansätze zu entwickeln, die AAC-Systeme konsequenter an den individuellen Kommunikationsbedürfnissen sprachbehinderter Menschen ausrichten und so eine persönlichere und authentischere Kommunikation fördern. Die hieraus resultierende Zielsetzung ZS9 ist wie folgt definiert:

Bezeichnung	**Zielsetzung 9 (ZS9)**
Beschreibung	Systematische Analyse der Potenziale KI-basierter Ansätze im Kontext digitaler AAC-Systeme sowie deren bisherige Realisierung im Hinblick auf Gesprächsphasen und -Kontexte. Ziel ist es, Einschränkungen bestehender Systeme hinsichtlich der Anpassungsfähigkeit an unterschiedliche kommunikative Kontexte (z. B. formelle vs. informelle) sowie Phasen eines Gesprächs (z. B. Anfang, Schluss) zu identifizieren und einzuordnen. Auf dieser Grundlage sollen bestehende Forschungslücken offengelegt und zentrale Anforderungen für die kontextsensitive, personalisierte Entwicklung intelligenter Kommunikationsagenten im Bereich der AAC abgeleitet werden.
Abhängigkeiten	Die im Rahmen der Bearbeitung der ZS9 gewonnenen Erkenntnisse über KI-Potenziale informieren die nachfolgenden Zielsetzung **ZS10** und **ZS11.**

Tabelle 9. Beschreibung der 9. Zielsetzung.

Die zuvor skizzierten Einschränkungen bestehender AAC-Systeme betreffen nicht nur die mangelnde Adaptionsfähigkeit gegenüber verschiedener Gesprächsphasen und -Kontexte hinweg, sondern offenbaren eine grundsätzliche Schwäche im Hinblick auf die Personalisierung von AAC-Systemen. In diesem Kontext bleibt auch die stilistische Anpassungsfähigkeit an die individuelle Persönlichkeit eines Nutzers in bestehenden AAC-Systemen bislang unberücksichtigt. Jedoch wirken standardisierte Ausgaben häufig unpersönlich, stilistisch starr und emotional entkoppelt (Elsahar et al., 2019; Light et al., 2021; Ascari et al., 2019). Nutzer können sich somit zwar generell verständlich machen, jedoch nicht in einer Weise, die ihrer individuellen Persönlichkeit und ihrer damit einhergehenden Ausdrucksweise entspricht. Dabei ist Kommunikation weit mehr als Informationsübertragung – sie dient der Beziehungsgestaltung und der persönlichen Selbstrepräsentation (Greenaway et al., 2015; Kane et al., 2017). Gerade für sprecheingeschränkte Menschen stellt diese fehlende Personalisierung eine Barriere für authentische Interaktion und gesellschaftliche Teilhabe dar.

Vor diesem Hintergrund gewinnt das Konzept der Digital Persona an Relevanz. Eine Digital Persona ist eine datenbasierte, virtuelle Nutzerrepräsentation, die dessen individuellen Eigenschaften abbildet und systemseitig modelliert (Clarke, 1994; Roosendaal, 2010). Im AAC-Kontext könnten so individuelle Persönlichkeitsmerkmale eines Nutzers modelliert und systemseitig im Rahmen der Textgenerierung operationalisiert werden. So könnten generierte Aussagen nicht nur funktional, sondern auch stilistisch kongruent zur Nutzeridentität formuliert und so bspw. in Bezug auf Tonalität oder Ausdrucksweise an die individuellen Nuancen des Nutzers angepasst werden. Während das Konzepte der Digital Persona bislang vorrangig im Bereich virtueller Assistenten oder Avatare zum Einsatz kommt (Belk, 2013; Sims, 2017), fehlt im Bereich der AAC-Systeme bislang eine systematische Integration.

Gleichzeitig eröffnet der technologische Fortschritt generativer KI – insbesondere durch sogenannte Large Language Models (LLMs) wie GPT-4 – neue Möglichkeiten: Diese Modelle sind in der Lage, nicht nur kontextuell kohärente, sondern auch stilistisch und emotional variantenreiche Textvorschläge automatisiert zu generieren (Brown et al., 2020; Jiang et al., 2023; Sorokovikova et al., 2024). Studien belegen, dass LLMs auf Basis von Nutzerprofilen personalisierte Kommunikationsstile imitieren und dynamisch anpassen können (Serapio-Garcia et al., 2023). Damit entsteht erstmals das Potenzial, automatisiert authentisch-personalisierte Äußerungen zu erzeugen, die sowohl in Inhalt als auch Form zur Persönlichkeit des jeweiligen Nutzers passen. Trotz dieser Potenziale fehlt es bislang an einem AAC-System, welches die individuelle Persönlichkeit eines Nutzers modelliert und in die Generierung personalisierter Kommunikationsbeiträge überführt.

Vor diesem Hintergrund setzt sich die vorliegende Dissertation das Ziel, das Potenzial generativer KI für die Entwicklung personalisierter AAC-Systeme systematisch zu untersuchen. Dabei sollen bestehende technische, kommunikative und gestalterische Anforderungen identifiziert und in ein integriertes AAC-Systemdesign überführt werden. Auf dieser Basis wird ein generatives AAC-System technisch umgesetzt und evaluiert, das individuelle, persönlichkeitsbasierte Kommunikationsstile berücksichtigen und somit eine stilistisch authentische Kommunikation ermöglichen soll. Die hieraus resultierenden Zielsetzungen ZS10 und ZS11 sind wie folgt definiert:

Bezeichnung	**Zielsetzung 10 (ZS10)**
Beschreibung	Konzeption eines Persönlichkeits-getriebenen AAC-Systems, das individuelle kommunikative Ausdrucksweisen anhand der Persönlichkeitsmerkmale der Nutzer systematisch berücksichtigt. Ziel ist es, auf Basis des Konzepts der Digital Persona ein integratives AAC-Systemdesign zu entwickeln, das LLMs zur automatisierten Generierung stilistisch authentischer und persönlichkeitskongruenter Kommunikationsbeiträge nutzt. Das System soll dabei in der Lage sein, sprachliche Stilpräferenzen sowie individuelle Ausdrucksformen adaptiv abzubilden und dynamisch an unterschiedliche Persönlichkeitsmerkmale anzupassen.
Abhängigkeiten	Die im Rahmen der Bearbeitung der ZS10 gewonnenen Erkenntnisse über die konzeptuellen Anforderungen des zu entwickelnden AAC-Systems informieren die nachfolgende Zielsetzung **ZS11.**

Tabelle 10. Beschreibung der 10. Zielsetzung.

Bezeichnung	**Zielsetzung 11 (ZS11)**
Beschreibung	Entwicklung und Evaluation eines generativen KI-gestützten AAC-Prototyps. Der zu entwickelnde Prototyp soll die Anforderungen (ZS10) des Digital Persona Ansatzes in einem funktionalen AAC-System operationalisieren, das dadurch in der Lage ist, nutzerspezifische Persönlichkeitsmerkmale in der Textgenerierung adaptiv abzubilden.
Abhängigkeiten	Die im Rahmen der Bearbeitung der **ZS10** gewonnenen Erkenntnisse über die konzeptuellen Anforderungen des zu entwickelnden AAC-Systems informieren diese Zielsetzung.

Tabelle 11. Beschreibung der 11. Zielsetzung.

Während sich ZS9 bis ZS11 auf die persönlichkeitsgetreue Generierung KI-basierter Kommunikation im AAC-Kontext fokussieren, eröffnet die Anwendung von generativer KI ebenfalls im gesundheitsbezogenen Kontext große Potenziale. Insbesondere kann das Problem der unzureichenden Personalisierung bestehender digitaler Gesundheitsanwendungen wie im Bereich der Darmgesundheit adressiert werden. Aktuelle Softwarelösungen, insbesondere regelbasierte Chatbots oder Apps mit vordefinierten Antwortmustern, sind häufig nicht in der Lage, individuelle Symptome, Verhaltensweisen oder Kontexte angemessen zu berücksichtigen. Dies erschwert eine zielgerichtete und wirksame Empfehlungserstellung, insbesondere in Bereichen, in denen physiologische, psychologische und lebensstilbezogene Einflussfaktoren eng miteinander verwoben sind.

Vor dem Hintergrund der zunehmenden Bedeutung der Darmgesundheit – insbesondere im Zusammenhang mit der sogenannten „Darm-Hirn-Achse“ (vgl. Avetisyan et al., 2015) – rückt die individuelle Unterstützung in diesem Gesundheitsbereich verstärkt in den Fokus. Der Zugang zu medizinischen Experten ist aufgrund des geringen Arzt-Patienten-Verhältnisses häufig eingeschränkt (vgl. Houwen et al., 2019), wodurch sich ein wachsender Bedarf an personalisierten Technologien ergibt, die individuelle Gesundheitsparameter erfassen, interpretieren und nutzerspezifische Empfehlungen ableiten können. Gleichzeitig mangelt es an systematischen Ansätzen zur Entwicklung solcher personalisierten Gesundheitsagenten. Insbesondere fehlt es an Designwissen darüber, wie die hierfür notwendigen digitalen Nutzerrepräsentationen ausgestaltet sein müssen, um als Grundlage für eine effektive, gesundheitsbezogene KI-gestützte Analyse dienen zu können.

Vor diesem Hintergrund zielt die letzte Zielsetzung (ZS12) darauf ab, die Potenziale generativer KI für gesundheitsbezogene, personalisierte Kommunikationsagenten im Bereich der Darmgesundheit systematisch zu untersuchen. Aufbauend auf einer Literaturrecherche, einer Marktanalyse bestehender Gesundheits-Apps sowie User Stories soll ein konzeptionelles Systemdesign eines solchen Agenten entwickelt werden. Im Zentrum steht dabei die Frage, welche funktionalen Anforderungen durch diesen gesundheitsbezogenen Kommunikationsagenten erfüllt werden müssen, um durch eine maßgeschneiderte Digital Persona nutzerzentrierte Empfehlungen für die Darmgesundheit zu ermöglichen.

Die Zielsetzung umfasst dabei sowohl die Ableitung konkreter Designanforderungen als auch die Entwicklung eines initialen Systemdesigns. Dieses soll als Ausgangspunkt für die iterative Umsetzung eines funktionsfähigen Prototyps dienen, der perspektivisch unter Einbezug potenzieller Endnutzer in realen Anwendungskontexten evaluiert werden kann. Mit dieser Zielstellung wird der Forschungsgegenstand dieser Dissertation abschließend um eine gesundheitsbezogene Perspektive erweitert, die das Potenzial generativer KI für die Förderung individueller Gesundheitskompetenz aufzeigt.

Bezeichnung	**Zielsetzung 12 (ZS12)**
Beschreibung	Konzeption eines Kommunikationsagenten auf Basis generativer KI zur personalisierten Unterstützung der Darmgesundheit. Ziel ist es, ein Systemdesign für einen nutzerzentrierten Kommunikationsagenten zu entwickeln, der mithilfe einer Digital Persona individuelle Gesundheitsinformationen kontextsensitiv verarbeiten. Aufbauend auf einer theoretisch-fundierten Literaturanalyse, einer explorativen Marktanalyse sowie auf User Stories sollen zentrale Anforderungen an ein solches System identifiziert und in ein initiales Systemkonzept überführt werden.
Abhängigkeiten	Die Beantwortung dieser Zielsetzung wird durch **ZS10 und ZS11** informiert: So fließen die gewonnenen Erkenntnisse zum Konzept sowie der technischen Umsetzung der Digital Personas aus ZS10 und ZS11 ergänzend in die Gestaltung dieses Systemdesigns ein. Zudem soll das konzeptionelle Systemdesign als Grundlage zukünftiger Forschungsarbeiten für die iterative Umsetzung eines KI-gestützten Gesundheitsagenten dienen, der präventive Gesundheitskompetenz fördert und den Zugang zu personalisierter Gesundheitskommunikation erleichtert.

Tabelle 12. Beschreibung der 12. Zielsetzung.

1.1.3 Inhaltlich-strukturelle Wirkungszusammenhänge der Zielsetzungen

Die Zielsetzungen der Dissertation sind nicht isoliert zu betrachten, sondern folgen einer konzeptionellen, inhaltlichen Struktur. So informieren Zielsetzungen, durch die im Rahmen der Bearbeitung gewonnenen methodischen und anwendungsbezogenen Erkenntnisse, die Grundlage zur Bearbeitung nachfolgender Zielsetzungen und bilden damit ein inhaltlich-strukturelles Wirkungsverhältnis. Dieses Wirkungsverhältnis wird systematisch in Tabelle 13 dargestellt.

j \ i	ZS1	ZS2	ZS3	ZS4	ZS5	ZS6	ZS7	ZS8	ZS9	ZS10	ZS11	ZS12
ZS1	-											
ZS2	-	-										
ZS3	-	-	-									
ZS4	→	→	→	-								
ZS5	-	-	→	→	-							
ZS6	→	→	→	→	→	-						
ZS7	-	-	-	-	-	-	-					
ZS8	-	-	-	→	-	→	→	-				
ZS9									-			
ZS10									→	-		
ZS11									→	→	-	
ZS12									-	→	→	-

Legende	
→	Die Bearbeitung der Zielsetzung *i* liefert wesentliche Erkenntnisse, welche die Bearbeitung von Zielsetzung *j* inhaltlich fundieren bzw. methodisch unterstützen.
-	Keine Beziehung.
	1. Forschungsgebiet
	2. Forschungsgebiet

Tabelle 13. Inhaltlich-strukturelle Wirkungszusammenhänge der Zielsetzungen.

1. Forschungsgebiet: Inhaltlich-strukturelle Verzahnung der Zielstellungen ZS1 bis ZS8 im Bereich der Entwicklung kontextsensitiver Analyseverfahren.

Die acht Zielstellungen des ersten Forschungsgebiets bauen inhaltlich aufeinander auf und tragen sukzessive zur Konzeption, Implementierung und Evaluation softwaregestützter Analyseverfahren und Werkzeuge zur Trend- bzw. Nutzeranalyse bei.

1. Zielsetzung. ZS1 bildet den methodischen Auftakt und adressiert die systematische Evaluation bestehender Topic Modelling Verfahren hinsichtlich ihrer Eignung zur Analyse von UGC im digitalen Marketing. Die hierbei gewonnenen Erkenntnisse über die Leistungsfähigkeit und kontextuelle Passung der Verfahren stellen eine essenzielle Entscheidungsgrundlage für alle nachfolgenden Zielstellungen dar, in denen Topic Modelling integraler Bestandteil softwaregestützter Analysetools ist. Insbesondere ZS4 und ZS6 profitieren direkt von der fundierten Methodeneinschätzung durch ZS1, da sie diese in die Entwicklung der dort fokussierten Werkzeuge integrieren.

2. Zielsetzung. In engem inhaltlichen Zusammenhang dazu steht ZS2, in der ein Analyseverfahren zur systematischen Erfassung von Kontextfaktoren in UGC konzipiert und implementiert wird. Die dort erarbeiteten Analyseverfahren zur Kontextmodellierung erweitern die in ZS1 betrachtete, rein inhaltlich-themenorientierte Perspektive der UGC-Analyse um eine zusätzliche semantische Tiefendimension und liefern einen entscheidenden Baustein für die Entwicklung ganzheitlicher Analyseverfahren in ZS4 und ZS6.

3. Zielsetzung. ZS3 erweitert diese methodischen Grundlagen und fokussiert die Ableitung von spezifischen Anforderungen an ein softwaregestütztes Trendanalysewerkzeug im Marketing. Diese Erkenntnisse (z. B. die Notwendigkeit einer Analyse der Tonalität) bilden das konzeptionelle Pflichtenheft für die anschließende technische Umsetzung des Analysewerkzeugs (ZS4).

4. Zielsetzung. ZS4 realisiert als direkte Fortsetzung von ZS3 die Konzeption, prototypische Umsetzung und Evaluation eines integrierten Trendanalysewerkzeugs und profitiert beim Systemdesign stark von den methodischen Grundlagen aus ZS1, ZS2 sowie ZS3. So fließen in das Systemdesign die aus ZS1 gewonnenen Erkenntnisse hinsichtlich der Leistungsfähigkeiten verschiedener Topic Modelling Verfahren, das in ZS2 entwickelte Kontextanalyseverfahren sowie die in ZS3 identifizierten marketingspezifischen Anforderungen an das Trendanalysewerkzeug ein.

5. Zielsetzung. Darauf aufbauend schließt ZS5 mit einer fokussierten Analyseperspektive an. Aufbauend auf den Erkenntnissen hinsichtlich der Entwicklung solcher automatisierten Trendanalysewerkzeuge aus ZS3 und ZS4, zielt ZS5 auf die Anforderungsanalyse eines Werkzeugs zur Aspekt-basierten Trendanalyse im Kontext der Produktentwicklung ab. Die gewonnenen Erkenntnisse bilden das konzeptionelle Pflichtenheft für die anschließende technische Umsetzung des Aspekt-basierten Analysewerkzeugs und ermöglichen die Entwicklung eines ganzheitlichen Trendanalysewerkzeugs, welches spezifisch an die Anforderungen der Produktentwicklung ausgerichtet ist. (ZS6).

6. Zielsetzung. Die in ZS5 erarbeiteten Anforderungen hinsichtlich des Aspekt-basierten Trendanalyseverfahrens fließen wiederum in ZS6 ein, welche die technische Umsetzung und die Evaluation des konzipierten Werkzeugs adressiert. ZS6 vereint damit das in ZS1 bis ZS5 aufgebaute Wissen zur Kontext-, Trend- und Sentimentanalyse zur Entwicklung des softwaregestützten Trendanalysewerkzeuges im Rahmen der Produktentwicklung und markiert damit einen integrativen Meilenstein der Arbeit.

7. Zielsetzung. Parallel hierzu zielt ZS7 auf die systematische Modellierung und Operationalisierung relevanter Merkmale von Lead Usern ab, um eine datenbasierten Identifikation der Lead User auf Basis objektivierbarer Eigenschaften zu ermöglichen. Die dort entwickelten methodischen Konzepte zur Operationalisierung der unterschiedlichen Lead User Merkmale dienen als methodische Grundlage für ZS8.

8. Zielsetzung. ZS8 schließt das erste Forschungsgebiet durch die Entwicklung und Evaluation eines skalierbaren, multimethodischen Analyseverfahrens zur automatisierten Lead User Identifikation. Neben den konzipierten Merkmalsmodellen aus ZS7 fließen auch die in ZS4 und ZS6 erarbeiteten Design- und Entwicklungserkenntnisse in ZS8 ein. Gerade mit Blick auf die technische Umsetzung der Lead User Merkmale (z. B. Trendführerschaft) bieten die gewonnen Einblicke hinsichtlich der entwickelten Trendanalyseverfahren einen großen Mehrwert und informieren die Operationalisierung dieser Merkmale.

> **2. Forschungsgebiet:** Inhaltlich-strukturelle Verzahnung der Zielstellungen ZS9 bis ZS12 im Bereich KI-gestützter digitaler Assistenzsysteme.

Die vier Zielstellungen des zweiten Forschungsgebiets folgen einer inhaltlich abgestimmten Forschungslogik, die von einer systematischen Analyse bestehender Defizite ausgeht und über die konzeptionelle Modellierung digitaler Persönlichkeitsrepräsentationen bis hin zur prototypischen Implementierung und Evaluation intelligenter Assistenzsysteme reicht. Durch die Erweiterung um eine gesundheitsbezogene Perspektive wird die Relevanz generativer KI zudem nicht nur im Kontext der AAC, sondern auch im Hinblick auf personalisierte digitale Gesundheitsberatung adressiert.

9. Zielsetzung. ZS9 bildet den analytischen Ausgangspunkt dieses Forschungspfads. Im Mittelpunkt steht die systematische Untersuchung bestehender AAC-Systeme hinsichtlich des darin realisierten Nutzungspotenzial von KI. Die Analyse fokussiert insbesondere die Nutzungspotenziale im Hinblick auf die Anpassungsfähigkeit der AAC-Systeme an unterschiedliche Gesprächsphasen und -Kontexte. Diese Ergebnisse liefern zentrale Erkenntnisse über bestehende Forschungslücken und ermöglichen so die Ableitung technischer und methodischer Nutzungspotenziale von KI zur personalisierten Kommunikationsunterstützung. Damit legt ZS9 die konzeptionelle und empirische Grundlage für die anschließenden Zielstellungen ZS10 und ZS11.

10. Zielsetzung. ZS10 greift diese Erkenntnisspotentiale auf und erweitert den Wissensstand gezielt um die konzeptionellen Anforderungen an ein KI-basiertes AAC-System, das individuelle Persönlichkeitsmerkmale des Nutzers systematisch in der stilistischen Anpassung der generierten Kommunikationsbeiträge berücksichtigt. Durch die Konzeption einer *Digital AAC Persona* wird ein integrativer Systemansatz entwickelt, der generative LLMs zur stilistisch authentischen und persönlichkeitsgetreuen Textgenerierung befähigt. Die in ZS10 entwickelten Anforderungen bilden wiederum die Grundlage für die nachfolgende Umsetzung in ZS11.

11. Zielsetzung. Mit ZS11 mündet der Forschungsstrang in die praktische Umsetzung und Evaluation des entwickelten Konzepts aus ZS10. Ziel ist die Entwicklung eines funktionsfähigen Prototyps, der die zuvor definierten Anforderungen technisch umsetzt. Der Fokus liegt hierbei auf der Berücksichtigung nutzerspezifischer Persönlichkeitsmerkmale bei der automatisierten Textgenerierung. Damit soll ein AAC-System entstehen, das nicht nur sprachliche Teilhabe ermöglicht, sondern auch authentische, individuell zugeschnittene Kommunikation fördert.

12. Zielsetzung. Die abschließende Zielsetzung ZS12 erweitert den zuvor skizzierten Forschungspfad um eine gesundheitsbezogene Anwendungsperspektive und untersucht das Potenzial der Digital Personas im Kontext der digitalen Gesundheitsförderung. Im Mittelpunkt steht die Konzeption eines personalisierten Kommunikationsagenten zur Unterstützung der Darmgesundheit, welcher mittels generativer KI personalisierte, kontextsensitive Empfehlungen formulieren soll. Die in ZS10 erarbeiteten Erkenntnisse über die konzeptionelle Ausgestaltung des personalisierten Analyseansatzes der Digital AAC Persona bilden dabei eine zentrale theoretische Grundlage. Zugleich fließen die in ZS11 gewonnenen Erkenntnisse zur system-technischen Operationalisierung der Digital AAC Persona in die Konzeption dieser gesundheitsspezifischen Digital Persona mit ein. Damit wird der in ZS10 und ZS11 etablierte Gestaltungsrahmen intelligenter, personalisierter Kommunikationssysteme in einen neuen Anwendungskontext übertragen, der die Digital Persona durch besondere Anforderungen im Hinblick auf die personalisierte Gesundheitsberatung beleuchtet.

Insgesamt ergibt sich somit eine logisch vernetzte und iterativ aufgebaute Forschungsarchitektur, bei der die Zielsetzungen systematisch auf den Ergebnissen vorangehender Zielsetzungen aufbauen und diese weiterführen – sei es durch methodische Verfeinerung, kontextuelle Spezialisierung oder softwaretechnische Umsetzung. So leistet die Dissertation durch die Konzeption und Umsetzung intelligenter Analysemethoden und Werkzeuge zur Operationalisierung digitaler nutzergenerierter Daten einen schrittweisen Beitrag zur Erschließung der darin enthaltenen Erkenntnispotentiale.

1.2 Ableitung der Forschungsfragen

Die Forschungsfragen FF1 bis FF13 sind als forschungsleitende Verdichtung zentraler Problemstellungen zu sehen, die im Rahmen der Zielstellungen systematisch bearbeitet werden. Ihre Formulierungen orientieren sich an den unterschiedlichen Anforderungsebenen der jeweiligen Zielsetzungen – von methodischer Grundlagenarbeit über anwendungsbezogene Systementwicklung bis hin zur Evaluation im Kontext konkreter Anwendungsszenarien. Sie gewährleisten damit sowohl inhaltliche Tiefe als auch empirische Anschlussfähigkeit und bilden die argumentative Brücke zwischen der problemzentrierten Zielstruktur (*vgl. Kapitel 1.1*) und den Erkenntnisinteressen dieser Dissertation.

> **1. Forschungsgebiet:** Konkretisierung der Zielstellungen ZS1 bis ZS8 in den Forschungsfragen FF1 bis FF8.

Im ersten Forschungsgebiet werden acht Zielsetzungen (ZS1 bis ZS8) verfolgt. Diese ergeben sich aus der übergeordneten Problemstellung, dass bislang keine hinreichend kontextsensitiven und skalierbaren Analysewerkzeuge existieren, um das Informationspotenzial in UGC systematisch vor dem Hintergrund marketingspezifischer Anwendungsszenarien zu erschließen. Die bestehenden Ansätze sind häufig entweder methodisch zu generisch, nicht auf konkrete Anwendungsfelder zugeschnitten oder technisch schwer zugänglich. Daraus ergibt sich die Notwendigkeit, datenanalytische Verfahren und Werkzeuge zu entwickeln, die den domänenspezifischen Anforderungen (z. B. zur Trendanalyse oder Lead User Identifikation) gerecht werden (vgl. *Kapitel 1.1.1*). Die acht Zielsetzungen greifen diese

Problematik in einer strukturierten Weise auf und münden jeweils in forschungsleitende Fragestellungen, deren Herleitung im Folgenden beschrieben wird.

1. und 2. Forschungsfrage. ZS1 adressiert die fehlende Vergleichbarkeit bestehender Topic Modelling Techniken im Hinblick auf ihre Eignung zur Analyse von UGC im digitalen Marketing. Während zahlreiche methodische Varianten existieren, fehlt es bislang an einer systematischen und praxisnahen Vergleichsstudie, die deren Einsatzgrenzen und Potenziale im Kontext marketingorientierter Nutzungsszenarien (Trendanalyse, Themenextraktion oder Inhaltsstrukturierung) bewertet. Diese Problematik führt zur Formulierung der ersten beiden Forschungsfragen. Die erste Frage ***FF1*** zielt auf die Notwendigkeit ab, überhaupt erst belastbare Vergleichsmaßstäbe zu entwickeln, die es erlauben, verschiedene Verfahren methodisch und domänenspezifisch sinnvoll zu evaluieren. Die zweite Frage ***FF2*** ergibt sich aus dem praktischen Bedürfnis, auf Basis dieses Evaluationsrahmens fundierte Empfehlungen für konkrete marketingspezifische Anwendungen abzuleiten. Die Differenzierung nach den drei Use Cases im Marketingbereich wurde dabei bewusst gewählt, da sie jeweils unterschiedliche Anforderungen an die Modelllogik und Ergebnisinterpretation stellen.

1. Forschungsfrage (FF1)
Nach welchen Kriterien lassen sich unterschiedliche Topic Modelling Verfahren systematisch vergleichen?

2. Forschungsfrage (FF2)
Welches Topic Modelling Verfahren eignet sich für die marketingbezogenen Anwendungsfälle (1) Trendanalyse, (2) Themenextraktion und (3) Inhaltsstrukturierung?

3. Forschungsfrage. ZS2 beschäftigt sich mit der bisher unzureichenden Berücksichtigung von Kontextfaktoren in der Analyse von UGC. Obwohl Plattformen wie X Metadaten zu Zeit, Ort oder Endgeräten bereitstellen, werden diese bislang kaum analytisch genutzt – insbesondere nicht in kombinierter Form. Die damit verbundenen Erkenntnispotenziale für die kontextsensitive Analyse von UGC bleiben folglich weitgehend ungenutzt. Die daraus abgeleitete Forschungsfrage ***FF3*** greift diese doppelte Forschungslücke auf. Zum einen wird ein automatisiertes Analyseverfahren entwickelt und explizit angewandt, um zu bewerten, ob eine kombinierte Analyse solcher Kontexte durch das entwickelte Verfahren die erhofften kontextualisierten Einsichten ermöglicht. Zum anderen wird untersucht, inwiefern technologische und geographische Kontextfaktoren das Stimmungsbild in UGC beeinflussen. Die

Frage ist also sowohl methodisch (Machbarkeit der Kontextanalyse) als auch deskriptiv (Wirkung der Kontexte) motiviert.

3. Forschungsfrage (FF3)
Welchen Einfluss haben die Kontextfaktoren Betriebssysteme mobiler Endgeräte und der geographische Kontext (Kontinente) auf die in sozialen Medien geäußerten Stimmungen, und wie lassen sich diese Kontextfaktoren in kombinierter Form analysieren?

4. und 5. Forschungsfrage. ZS3 und ZS4 befassen sich mit der Entwicklung eines softwaregestützten Trendanalysewerkzeugs, das in der Lage ist, inhaltliche, kontextuelle und tonalitätsbasierte Perspektiven ganzheitlich zu integrieren. Während ZS3 die Ableitung eines strukturierten Anforderungsprofils adressiert, stellt ZS4 die technische Umsetzung und Evaluation des Werkzeugs ins Zentrum. Vor diesem Hintergrund behandelt die vierte Forschungsfrage ***FF4***, welche funktionalen und gestalterischen Anforderungen ein solches Analysewerkzeug erfüllen muss. Die fünfte, darauf aufbauende Forschungsfrage ***FF5*** fragt nach der konkreten Gestaltung und Umsetzung eines entsprechenden Analysewerkzeugs. Dadurch wird geprüft, ob auf Basis der Anforderungen tatsächlich ein funktionales, softwaregestütztes Analysewerkzeug hervorgehen kann – ein Anspruch, den viele konzeptionelle Arbeiten bislang nicht berücksichtigen.

4. Forschungsfrage (FF4)
Welche Designanforderungen muss ein Topic-Modelling-basiertes Trendanalysewerkzeug erfüllen, das die marketingbezogenen Nutzungsszenarien Produktentwicklung, Kundenverhaltensanalyse und Markt-/Markenmonitoring unterstützt?

5. Forschungsfrage (FF5)
Wie lässt sich ein solches Trendanalysewerkzeug technisch gestalten und implementieren, und welche konkreten wissenschaftlichen sowie praktischen Beiträge können daraus abgeleitet werden?

6. Forschungsfrage. Mit ZS5 und ZS6 verschiebt sich der Fokus auf die Anwendungsperspektive der Produktentwicklung. Im Zentrum steht die Frage, wie sich die Entwicklung relevanter Produkt- und Serviceaspekte differenziert analysieren lässt, insbesondere im Hinblick auf die dazugehörigen Kundenwahrnehmungen. Die Herausforderung besteht darin, die Topic Modelling basierte Trendanalyse mit der Aspekt-basierten Sentimentanalyse zu kombinieren und dabei die kontextsensitiven Anforderungen der Produktentwicklung an die Analyse zu berücksichtigen. Die daraus abgeleitete Forschungsfrage ***FF6*** greift diese methodische Integration auf und vereint darin die beiden Zielsetzungen ZS5 und ZS6 durch die methodische Konzeption und technische Umsetzung des integrativen Analyseansatzes.

6. Forschungsfrage (FF6)
Wie könnte ein softwaregestütztes Analysewerkzeug zur Aspekt-basierten Sentimentanalyse gestaltet sein, das die Trendanalyse im Rahmen der Produktentwicklung unterstützt, und welche Anforderungen muss ein solches Werkzeug erfüllen?

7. und 8. Forschungsfrage. ZS7 und ZS8 behandeln die automatisierte Identifikation von Lead Usern, die durch ihre Innovationsnähe und Netzwerkposition als besonders wertvolle Informationsquelle für Innovationsprozesse gelten. Die Schwierigkeit liegt darin, diese Nutzer automatisiert und skaliert aus großen Mengen an UGC herauszufiltern – ein Problem, das bisherige Verfahren meist nur partiell lösen (vgl. Martínez-Torres, 2014; Hau & Kang, 2016). Vor diesem Hintergrund ergibt sich die siebte Forschungsfrage ***FF7*** dieser Dissertation. Diese Frage zielt auf die grundlegende Konzeptionalisierung und empirische Modellierung der relevanten Lead User Eigenschaften ab. Sie ist notwendig, um überhaupt ein systematisch fundiertes und automatisiertes Identifikationsverfahren entwickeln zu können. Die achte Forschungsfrage ***FF8*** erweitert diese Perspektive um die technische Umsetzung. Sie fragt danach, ob und wie ein solches Verfahren in einem prototypischen System abgebildet werden kann, welches die Identifikation der in FF7 erhobenen Merkmale eines Lead Users automatisiert in unstrukturierten UGC-Daten ermöglicht.

7. Forschungsfrage (FF7)
Welche charakteristischen Merkmale weisen Lead User in Online-Communities auf, und wie lassen sich diese Merkmale datenbasiert operationalisieren?

8. Forschungsfrage (FF8)
Wie kann die Identifikation von Lead Usern durch ein softwaregestütztes Analyseverfahren unterstützt werden, und welche technischen Verfahren sind hierfür geeignet?

In ihrer Gesamtheit adressieren die Forschungsfragen FF1 bis FF8 zentrale methodische und anwendungsbezogene Herausforderungen bei der Analyse von UGC. Sie zielen darauf ab, geeignete Verfahren zur kontextuellen Analyse methodisch zu fundieren und in softwaregestützten Werkzeugen umzusetzen. Damit leisten die Dissertation einen Beitrag zur Erschließung der Erkenntnisspotentiale digitaler UGC in marktorientierten Anwendungsfeldern wie der Trendanalyse, Produktentwicklung und Lead User Identifikation.

> **2. Forschungsgebiets:** Konkretisierung der Zielstellungen ZS9 bis ZS12 in den Forschungsfragen FF9 bis FF13.

Die Forschungsfragen des zweiten Forschungsgebiets ergeben sich aus einer schrittweisen Auseinandersetzung mit den methodischen und technologischen Herausforderungen, die mit dem Einsatz generativer KI im Kontext personalisierter Kommunikation in digitalen Assistenzsystemen verbunden sind. Im Zentrum steht dabei zunächst die lautsprachliche Kommunikation sprachbeeinträchtigter Menschen über digitale AAC-Systeme (ZS9–ZS11), bevor in einem letzten Schritt (ZS12) eine übertragene Anwendungsperspektive im Bereich der gesundheitsbezogenen Kommunikation eröffnet wird.

9. und 10. Forschungsfrage. Die beiden Forschungsfragen (FF9 und FF10) resultieren aus der Analyse bestehender AAC-Systeme und der in der wissenschaftlichen Literatur diskutierten Einsatzpotenziale von KI (ZS9). Ausgehend von der Beobachtung, dass aktuelle Systeme oft keine hinreichende Kontextsensitivität entlang unterschiedlicher Gesprächsphasen (z. B. Einstieg, Hauptteil, Abschluss) und -Kontexte (z. B. formell vs. informell) aufweisen, zielt ***FF9*** darauf ab, zu untersuchen, inwiefern die in der Literatur beschriebenen KI-Technologien diesen kommunikativen Anforderungen gerecht werden können. Aufbauend darauf fokussiert ***FF10*** die Ableitung konzeptioneller und gestalterischer Implikationen, die sich aus dieser Analyse der KI-Potenziale für zukünftige KI-basierte AAC-Systeme ergeben. Beide Forschungsfragen leiten sich somit direkt aus der in ZS9 vorgenommenen systematischen Untersuchung bestehender AAC-Lösungen vor dem Hintergrund ihrer KI-Nutzungspotenziale ab und bilden eine theoretische Grundlage für das nachfolgende Systemdesign in FF11 und FF12.

9. Forschungsfrage (FF9)
Decken die in der wissenschaftlichen Literatur vorgeschlagenen KI-Anwendungen für AAC-Systeme die Gesprächsphasen ‚Beginn', ‚Mitte ' und ‚Abschluss' sowie die Gesprächskontexte ‚Formell' und ‚Informell' in angemessener Weise ab?

10. Forschungsfrage (FF10)
Welche Implikationen ergeben sich für den Einsatz von KI im Kontext von AAC-Systemen hinsichtlich der Berücksichtigung von Gesprächsphasen und -kontexten?

11. und 12. Forschungsfrage. Die Forschungsfragen FF11 und FF12 adressieren die Entwicklung einer intelligenten, persönlichkeitsgetriebenen Digital Persona im AAC-Kontext

(ZS10 und ZS11). Während ***FF11*** darauf abzielt, Design Prinzipien für eine solche Digital Persona zu identifizieren, um stilistisch authentische und persönlichkeitskongruente Äußerungen zu erzeugen, beschäftigt sich ***FF12*** mit der technischen Umsetzung dieser Design Prinzipien innerhalb eines softwaregestützten Prototyps. Beide Fragen reflektieren zentrale Herausforderungen bei der Übertragung der theoretischen Anforderungen aus ZS10 in ein operatives, real nutzbares System gemäß ZS11. Sie zielen darauf ab, sowohl die Konzeption als auch die technische Realisierung intelligenter, personalisierter Kommunikationssysteme systematisch zu untersuchen.

11. Forschungsfrage (FF11)
Wie kann eine generative KI-basierte Digital Persona gestaltet werden, um persönlichkeitsgetriebene Äußerungen innerhalb eines AAC-Systems zu erzeugen?

12. Forschungsfrage (FF12)
Wie kann die generative KI-basierte Digital Persona implementiert und technisch realisiert werden?

13. Forschungsfrage. Die letzte Forschungsfrage FF13 erschließt sich aus ZS12 und erweitert den thematischen Fokus um eine gesundheitsbezogene Anwendungsperspektive. Vor dem Hintergrund der zunehmenden Relevanz personalisierter digitaler Gesundheitsberatung zielt ***FF13*** darauf ab, konkrete Design Anforderungen für einen Konversationsagenten zu identifizieren, der durch eine maßgeschneiderte Digital Persona personalisierte Empfehlungen zur Darmgesundheit generieren kann. Die Forschungsfrage ergibt sich aus der zuvor identifizierten Problematik unzureichend personalisierter Gesundheitsanwendungen sowie dem in ZS12 formulierten Ziel, systematische Grundlagen für die Gestaltung solcher Systeme zu schaffen. Im Fokus steht dabei sowohl die Ableitung kontextsensitiver Design Anforderungen als auch deren konzeptuelle Umsetzung in einem initialen Design des digitalen Kommunikationsagenten.

13. Forschungsfrage (FF13)
Welche funktionalen Anforderungen sollte ein digitaler Konversationsagent erfüllen, der mithilfe einer Digital Persona personalisierte Empfehlungen zur Darmgesundheit auf Basis generativer KI bereitstellt, und wie kann der Agent sowie die Persona entsprechend dieser Anforderungen gestaltet werden?

In ihrer Gesamtheit spiegeln die Forschungsfragen FF9 bis FF13 den interdisziplinären Anspruch des zweiten Forschungsgebiets wider: Sie verbinden sozial-kommunikative und technische Perspektiven und adressieren zentrale Herausforderungen bei der Entwicklung und Anwendung generativer KI zur Förderung digitaler, personalisierter Kommunikation – sowohl im Bereich der barrierefreien Inklusion als auch im gesundheitsorientierten Alltag.

1.3 Aufbau der Dissertation

Um das Thema der Dissertation *„Automatisierte Erschließung der Erkenntnispotenziale digitaler Daten – Konzeption und Umsetzung intelligenter Analysemethoden und Werkzeuge zur Operationalisierung digitaler nutzergenerierter Daten durch den Einsatz von Data Science Methoden und generativer Künstlicher Intelligenz*“ zu adressieren und die darin entwickelten Zielsetzungen und Forschungsfragen systematisch zu behandeln, ist die Dissertation wie folgt aufgebaut:

In **Kapitel 1 „Einleitung“** werden die forschungsleitenden Themen der Dissertation motiviert und in diesem Rahmen die zugrundeliegenden Problemstellungen dargelegt. Zudem werden darauf aufbauende Zielsetzungen und dazugehörige Forschungsfragen abgeleitet. Das erste Forschungsgebiet adressiert dabei die Analyse digitaler Kommunikationsdaten (UGC) zur Unterstützung ökonomisch geprägter Entscheidungen; das zweite Forschungsgebiet fokussiert die Entwicklung intelligenter Assistenzsysteme zur Steigerung der Inklusion und Gesundheitsförderung. Über beide Forschungsgebiet werden insgesamt die Zielstellungen ZS1 bis ZS12 formuliert und die Forschungsfragen FF1 bis FF13 abgeleitet.

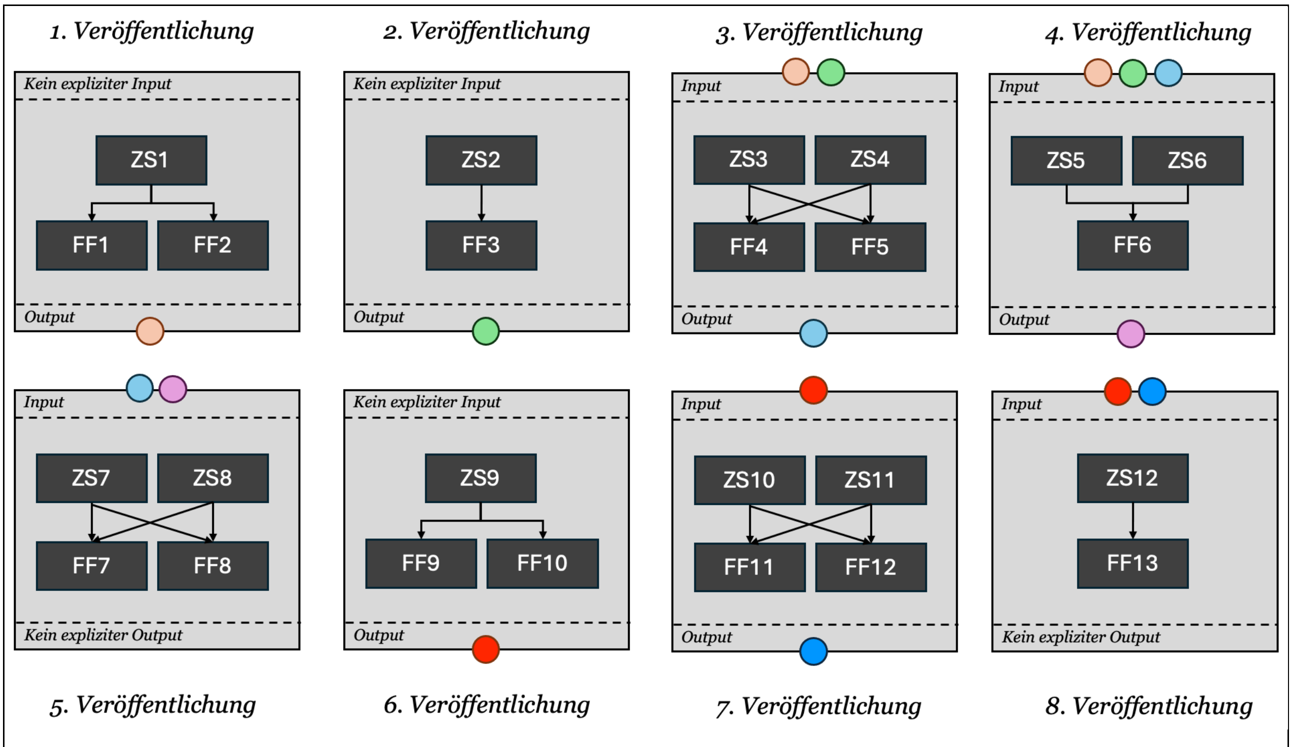
In **Kapitel 2 „Wissenschaftliche Beiträge“** werden die zentralen Ergebnisse der kumulativen Dissertation dargestellt. Das Kapitel umfasst insgesamt acht wissenschaftliche Veröffentlichungen, die jeweils mindestens einer Zielsetzungen zugeordnet sind. Diese Beiträge sind entsprechend ihrer thematischen Zugehörigkeit in die zwei Forschungsfelder gegliedert:

- Veröffentlichung ***2.1*** bis ***2.5*** bilden das ***erste Forschungsgebiet***.
- Veröffentlichung **2.6** bis **2.8** bilden das ***zweite Forschungsgebiet***.

In **Kapitel 3 „Schlussbetrachtung und Fazit“** werden die zentralen Erkenntnisse der Dissertation zusammengeführt. Dabei werden die geleisteten Beiträge der entstandenen Veröffentlichungen zu den (weiter-)entwickelten Technologien, der wissenschaftlichen Theoriebildung sowie zu den praxisrelevanten Implikationen reflektiert. Darüber hinaus werden Limitationen der Arbeit diskutiert und abschließend weiterführende Forschungsfelder skizziert.

2 Wissenschaftliche Veröffentlichungen

In diesem Kapitel werden die wissenschaftlichen Veröffentlichungen der kumulativen Dissertation vorgestellt. Zur besseren Orientierung wird zunächst ein Überblick über die relevante Informationen jeder Veröffentlichungen (z. B. Status, Ranking, Art der Veröffentlichung) gegeben (vgl. Tabelle 14). Anschließend werden die einzelnen Beiträge in den Kapiteln 2.1 bis 2.8 systematisch dargestellt. Mittels Fact Sheet wird für jede Veröffentlichung jeweils der inhaltliche Schwerpunkt erläutert, die zugrundeliegende Zielsetzung verortet, die Forschungsmethodik aufgeführt und die Kernergebnisse der Veröffentlichung zusammengefasst. Abbildung 1 veranschaulicht den strukturellen Zusammenhang der acht wissenschaftlichen Veröffentlichungen und zeigt auf, welche Beiträge auf methodischen, technischen oder inhaltlichen Erkenntnissen vorheriger Arbeiten aufbauen.



Legende	
	Farblich codiertes, im Rahmen der Veröffentlichung erzeugtes Wissen.
Input-Layer	Zeigt die Integration zuvor erschlossenen Wissens an.
Output-Layer	Zeigt die Erzeugung erschlossenen Wissens an.

Abbildung 1. Strukturübersicht und Erkenntnistransfers im Forschungsdesign.

Dabei werden sowohl explizite als auch implizite Formen des Erkenntnistransfers berücksichtigt: Während explizite Erkenntnistransfers konkrete Wissensinhalte (z. B. technische Anforderungen an ein zu entwickelndes Analysewerkzeug) markieren, bestehen ebenfalls implizite Verbindungen, die die schrittweise Entwicklung von Forschungskompetenzen (z.

B. im Bereich der gestaltungsorientierten Prototypenentwicklung) abbilden. Auf diese Weise wird sichtbar, wie sich im Verlauf des kumulativen Dissertationsprojekts ein wachsender Wissens- und Erfahrungsraum entwickelt, der die zielgerichtete Bearbeitung nachfolgender Forschungsbeiträge informiert.

Hinweis zur Seitennummerierung in Kapitel 2:

In den Kapiteln 2.1 bis 2.8 werden die wissenschaftlichen Beiträge im Originalformat eingebunden. Dabei erscheint zusätzlich zur regulären Seitenzahl der Dissertation eine zweite, leicht eingerückte und höher platzierte Seitenzahl, welche der Originalpaginierung des jeweiligen Beitrags entspricht. Die doppelte Seitennummerierung liegt bewusst vor und soll einerseits der Orientierung innerhalb des jeweiligen Beitrags sowie andererseits im Gesamtkomplex der Dissertation dienen.

	Kapitel	Titel	Autoren	Status	Ranking[1]	Autorenanteil
1. Forschungsgebiet	2.1	Comparison of Topic Modelling Techniques in Marketing – Results From an Analysis of Distinctive Use Cases	**Janik Wörner** Daniel Konadl Isabel Schmid Susanne Leist	**Veröffentlicht** in den Proceedings der 29. European Conference on Information Systems. (*ECIS*)	A	40 %
	2.2	Identifying Sentiment Influences Provoked by Context Factors – Results from a Data Analytics Procedure Performed on Tweets	Daniel Konadl **Janik Wörner** Susanne Leist	**Veröffentlicht** in den Proceedings der 54. Hawaii International Conference on System Sciences. (*HICSS*)	B	30 %
	2.3	MANTRA: A Topic Modeling-Based Tool to Support Automated Trend Analysis on Unstructured Social Media Data	**Janik Wörner** Daniel Konadl Isabel Schmid Susanne Leist	**Veröffentlicht** in den Proceedings der 44. International Conference on Information Systems, Hyderabad, Indien. (*ICIS*)	A	30 %
	2.4	Supporting Product Development by a Trend Analysis Tool Applying Aspect-Based Sentiment Detection	**Janik Wörner** Daniel Konadl Isabel Schmid Susanne Leist	**Veröffentlicht** in den Proceedings der 17. International Conference on Design Science Research in Information Systems and Technology, St. Petersburg, Florida, USA. (*DESRIST*)	B	30 %
	2.5	Automated identification of different lead users regarding the innovation process	**Janik Wörner** Isabel Schmid Janik Wörner Susanne Leist	**Veröffentlicht** im Electronic Markets Journal. (*EM*)	B	45 %
2. Forschungsgebiet	2.6	Artificial Intelligence in Augmentative and Alternative Communication Systems – A Literature-Based Assessment and Implications of Different Conversation Phases and Contexts	Daniel Konadl **Janik Wörner** Lucas Luttner Susanne Leist	**Veröffentlicht** in den Proceedings der 31. European Conference on Information Systems, Kristiansand, Norwegen. (*ECIS*)	A	30 %
	2.7	Personality is all that matters – Enhancing AAC Systems with Digital Personas Utilizing Generative Artificial Intelligence	**Janik Wörner**	**Unter Begutachtung** auf der 46. International Conference on Information Systems, Nashville, Tennessee, USA. (*ICIS*)	A	100 %
	2.8	Designing a Conversational Agent for User- Centric Gut Health Support Through Generative AI-Based Digital Personas	Daniel Konadl Florian Johannsen Susanne Leist **Janik Wörner** Isabel Reuter	**Unter Begutachtung** auf der 46. International Conference on Information Systems, Nashville, Tennessee, USA. (*ICIS*)	A	15 %
						320 %

Tabelle 14. Titel, Autoren, Status und Ranking der Veröffentlichungen.

[1]Die Angabe des Rankings erfolgte jeweils nach dem VHB JQ 4 Ranking von 2024.

2.1 Comparison of Topic Modelling Techniques in Marketing – Results From an Analysis of Distinctive Use Cases

<table>
<tr><td>Adressierte Zielsetzung</td><td>Systematische Entwicklung eines Evaluationsrahmens zur Analyse verschiedener Topic Modelling Verfahren hinsichtlich ihrer Eignung zur marketingspezifischen Analyse von UGC. Durch die Untersuchung soll eine empirische Vergleichsgrundlage geschaffen werden, anhand welcher die Verfahren vor dem Hintergrund ihrer Eignung im Rahmen marketingspezifischer Anwendungsszenarien bewertet und somit letztendlich Empfehlungen für die Wahl eines geeigneten, anwendungsspezifischen Verfahrens formuliert werden können.</td><td>ZS1</td></tr>
<tr><td rowspan="2">Abgeleitete Forschungsfragen</td><td>Nach welchen Kriterien lassen sich unterschiedliche Topic Modelling Verfahren systematisch vergleichen?</td><td>FF1</td></tr>
<tr><td>Welches Topic-Modelling-Verfahren eignet sich für die marketingbezogenen Anwendungsfälle (1) Themenextraktion, (2) Trendanalyse und (3) Inhaltsstrukturierung?</td><td>FF2</td></tr>
<tr><td>Genutzte Forschungsmethode</td><td colspan="2">Im Rahmen der Studie wurden die drei Topic Modelling Techniken Latent Dirichlet Allocation (LDA), Dirichlet Multinomial Regression (DMR) und Pachinko Allocation Model (PAM) anhand eines realen Facebook-Datensatzes mit über 4 Millionen Beiträgen auf ihre Eignung zur Analyse von UGC im Rahmen der drei marketingspezifischen Anwendungsszenarien untersucht. Die Untersuchung folgt dem Goal Question Metric (GQM) Ansatz nach Basili et al. (1994):
Goal: Ziel war es, quantitative und qualitative Vergleichsparameter abzuleiten, wodurch die marketingspezifischen Anforderungen in objektivierte Evaluationsmetriken übersetzt und die drei Topic Modelling Techniken systematisch miteinander verglichen werden können.
Question: Es wurde untersucht, inwieweit die drei Techniken LDA, PAM und DMR empirisch verglichen werden können und wie sie die Anforderungen der drei marketingbezogenen Anwendungsfälle der Themenextraktion, Trendanalyse und Inhaltsstrukturierung erfüllen.
Metric: Zur Objektivierung der Anforderungen wurden geeignete quantitative und qualitative Evaluationsmetriken aus der Forschungsliteratur abgeleitet. Diese dienen dazu, die Leistungsfähigkeit der Modelle vor dem Hintergrund der jeweiligen Anwendungsszenarien zu erfassen und vergleichbar zu machen.
Konkret wurden die quantitativen Metriken Log-Likelihood, Coherence und Build Time sowie die qualitativen Metriken Word- und Topic Intrusion abgeleitet, um sowohl die statistische Modellgüte als auch die semantische Interpretierbarkeit der Ergebnisse bewerten zu können.</td></tr>
<tr><td>Kernergebnisse (Überblick)</td><td colspan="2">Im Rahmen der Studie wurden insgesamt elf marketingspezifische Anforderungen an Topic Modelling Techniken identifiziert und in den fünf Vergleichsmetriken (Log-Likelihood, Coherence, Build Time, Word- und Topic Intrusion) operationalisiert. Die Anforderungen und Metriken wurden darüber hinaus den drei marketingbezogenen Anwendungsfällen (a) Themenextraktion, (b) Trendanalyse und (c) Inhaltsstrukturierung zugeordnet.</td></tr>
</table>

	Anhand eines realen Facebook-Datensatzes mit ca. 4,1 Millionen Beiträgen erfolgte ein systematischer Vergleich der drei Techniken LDA, DMR und PAM unter Anwendung der genannten Metriken. Die Analyse zeigte, dass keine der Techniken universell überlegen ist, sondern jeweils kontext- und anwendungsfallabhängige Stärken besitzt. Daraus lassen sich differenzierte Anwendungsempfehlungen ableiten: • Für die Themenextraktion empfiehlt sich **LDA**, insbesondere bei niedriger Themenanzahl; oder **DMR**, wenn eine größere thematische Bandbreite abgebildet werden soll. • Für die Trendanalyse überzeugt das semi-supervisierte Verfahren **DMR** aufgrund seiner Fähigkeit zur Einbindung von externen Dokumentenmetadaten (z. B. Zeitstempel) und seiner insgesamt höheren semantischen Kohärenz. • Für die Inhaltsstrukturierung ist **PAM** am besten geeignet, da es die hierarchische Strukturierung von Themen erlaubt und semantisch zusammenhängende Subthemen identifizieren kann. Diese Erkenntnisse ermöglichen eine fundierte Auswahl geeigneter Topic Modelling Verfahren für unterschiedliche analytische Zielsetzungen im Marketing.
Publikationsort	Das Paper wurde 2021 auf der European Conference on Information Systems (*ECIS*) veröffentlicht und ist in der AIS Electronic Library (*AISeL*) als peer-reviewtes Konferenzpapier gelistet.
Ranking VHB JQ 4	**A**
Autoren und Anteile	**Janik Wörner** **40%** Daniel Konadl 25% Isabel Schmid 25% Susanne Leist 10%
DOI	http://doi.org/10.5283/epub.46039

Tabelle 15. **Fact Sheet zur 1. Veröffentlichung.**

6-14-2021

COMPARISON OF TOPIC MODELLING TECHNIQUES IN MARKETING - RESULTS FROM AN ANALYSIS OF DISTINCTIVE USE CASES

Janik Wörner
University of Regensburg, janik.woerner@ur.de

Daniel Konadl
University of Regensburg, daniel.konadl@wiwi.uni-regensburg.de

Isabel Maria Schmid
University of Regensburg, isabel.schmid@wiwi.uni-regensburg.de

Susanne Leist
University of Regensburg, susanne.leist@ur.de

COMPARISON OF TOPIC MODELLING TECHNIQUES IN MARKETING - RESULTS FROM AN ANALYSIS OF DISTINCTIVE USE CASES

Research paper

Janik Wörner, University of Regensburg, Regensburg, Germany, Janik.Woerner@ur.de
Daniel Konadl, University of Regensburg, Regensburg, Germany, Daniel.Konadl@ur.de
Isabel Schmid, University of Regensburg, Regensburg, Germany, Isabel.Schmid@ur.de
Susanne Leist, University of Regensburg, Regensburg, Germany, Susanne.Leist@ur.de

Abstract

Currently, topic modelling is an effective analytical tool for the automated investigation of text data. However, identifying the underlying topics is still a challenging task that is dependent on the selection of the proper technique. Moreover, due to the considerable number of topic modelling techniques reported in the literature, uncertainty about the application of the techniques arises for both researchers and practitioners. Therefore, we conducted a comparison of three different topic modelling techniques (LDA, PAM, DMR) to give recommendations for three use cases identified in the literature: content extraction, trend analysis and content structuring. For each of them, we identified several requirements and by conducting the method 'Goal Question Metric', we derived several comparison metrics. We applied these metrics to a real-world Facebook data set (4,155,992 posts) to highlight the differences between the three topic modelling techniques and to give recommendations for our defined use cases.

Keywords: topic modelling, social media analysis, text analysis, marketing use cases

1 Introduction

Topic modelling is a prevalent kind of probabilistic generative model for extracting latent variables from large unstructured data sets (Liu et al., 2016). This can be applied to analyse different data such as bioinformatics data (Coelho et al., 2010), environmental data (Girdhar et al., 2013), and text data (Vayansky and Kumar, 2020). Thus, topic modelling has been studied in different disciplines and is also prevailing in information systems (IS) research, mainly focusing on Latent Dirichlet Allocation (LDA) (Blei et al., 2003) because of its simple applicability and good analysis results (Debortoli et al., 2016, Eickhoff and Neuss, 2017). With the continuous growth of social media and the consequential transformation of the way individuals interact with each other, increasing amounts of written data are created that can be analysed i.a. to support marketing related decision-making (Ghosh and Guha, 2013). Thus, social media data are increasingly used to enrich marketing tasks, such as complaint management (Einwiller and Steilen, 2015, Grégoire et al., 2015), innovation management (Mount and Martinez, 2014, Piller et al., 2012), or sales (Guesalaga, 2016, Marshall et al., 2012). However, the huge amount of written social media contents (Statista, 2020) complicates manual content analysis.

To solve this problem, automated topic discovery techniques and – in particular – topic modelling have been widely investigated (Chinnov et al., 2015, Eickhoff and Neuss, 2017, Hong and Davison, 2010). Topic modelling enables the analysis of a large amount of written social media data to extract embedded topics. Therefore, topic modelling has facilitated addressing marketing related questions and problems that have exceeded the feasibility of in-depth qualitative analysis (Eickhoff and Neuss, 2017). Thus, marketing related problems that refer to (1) content extraction, (2) trend analysis and (3) content structuring have often been discussed in the literature. Companies are required to base their products and services on customer requirements. Therefore, (1) content extraction with topic modelling is an appropriate application to extract customers' praise and criticism for product planning purposes (Irawan et

al., 2020, Rathore et al., 2018). To be aware of evolving trends concerning their own products and services, marketing departments conduct (2) trend analysis. Tracking evolving and changing requirements of customers is imperative to fulfil customers' wishes (Hong et al., 2012, Lozano et al., 2017). Moreover, (3) content structuring can help marketing departments to gain deeper insights into topics and their interrelatedness (Anoop et al., 2015, Srijith et al., 2017). Extracted hierarchical structures can reveal relationships between topics (e.g. price and product quality) and support more coordinated and sounder decision-making. Nevertheless, identifying the underlying topics of these documents is still a challenging task as the reasonable extraction of significant statistics and features from a dataset is dependent on the selection of the proper technique (Vayansky and Kumar, 2020).
As mentioned above, a growing number of IS-related investigations are currently using LDA. However, the basic LDA cannot represent all use cases (e.g. mapping hierarchies) for marketing related tasks so that extensions of LDA are essential. Therefore, not only do the numerous existing techniques for topic modelling hinder practical applications, but also the necessity of advanced technique-related knowledge. Liu et al. (2016) have divided various extensions of LDA into three areas: (I) extension of topic attributes (II) extension of document attributes and (III) extension of word attributes. Although numerous techniques are presented in these three areas, such as the Partially Labelled LDA (PLLDA) (Ramage et al., 2011), the Dirichlet Multinomial Regression (DMR) (Mimno and McCallum, 2008), or the Pachinko Allocation Model (PAM) (Li and McCallum, 2006), these extensions are scarcely applied. Due to the large number of topic modelling techniques in the current research literature, uncertainty about the selection of the right technique can arise. Vakansky and Kumar (2020) addressed this problem by conducting a theoretical comparison of different topic modelling techniques based on a structured literature review. Although this serves as a good overview of various topic modelling approaches and as a starting point for selecting a technique, differences only become obvious when applying them to a real-world data set. Furthermore, the results do not give clear suggestions which problem should be addressed with which technique. We addressed these problems by conducting a comparison between three different topic modelling techniques to give recommendations regarding the three use cases of (1) content extraction, (2) trend analysis, and (3) content structuring. We contribute to close this identified gap by comparing the practical application of the three techniques. This leads to the following research questions:

RQ1: Which criteria can be used to compare the different topic modelling techniques with each other?
RQ2: Which topic modelling technique can be recommended for the marketing related use cases (1) content extraction, (2) trend analysis, and (3) content structuring?

Topic modelling, both in general and especially regarding the analysis of companies´ social media posts, represents an important area for IS research. Accordingly, against the background of marketing we uncover various corporate use cases in the context of social media. By applying and comparing the three topic modelling techniques LDA, DMR, and PAM and by using build time, log-likelihood, coherence, word and topic intrusion as evaluation measures, we want to show differences between these techniques, identify advantages, disadvantages, and various potentials to enhance topic modelling techniques. Hence, we apply LDA, DMR, and PAM to a real-world data set. The remainder of this paper is as follows: section 2 provides a theoretical background. Then, we refer to the derivation of the three use cases and the respective requirements from literature. The transformation of them into topic modelling related metrics is also described here. Next, the procedure of the research 'Goal Question Metric' (cf. Basili, 1994) is described in section 3. Section 4 deals with the selection of the topic modelling techniques that are used for our comparison. The data analysis in section 5 achieves this comparison and further explains the data collection, the preparation of the data and the data analysis. Afterwards, in section 6, we present and discuss our results. Finally, section 7 draws an overall conclusion.

2 Theoretical Background

2.1 Social media

Social media can be defined as '*a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated*

Content (UGC)' (Kaplan and Haenlein, 2010, p.61). Social media connects people with the same interests, activities, backgrounds, or friendships (AlFalahi et al., 2014, Schneider et al., 2009). The active utilisation of social media enjoys particularly great popularity in private use. However, companies have also adopted social media to support value-creation (Hanna et al., 2011, McDonald and Aron, 2012). In particular, many companies apply these Internet-based applications such as content communities (e.g. YouTube), blogs, or social networks (e.g. Facebook or Twitter) to enable communication mainly with external stakeholders (Kietzmann et al., 2011). Thus, companies adopt social media to achieve different business objectives such as branding, including advertising, marketing, and content delivery for creating brand awareness (Culnan et al., 2010, Di Gangi et al., 2010, Kietzmann et al., 2011). Therefore, social media such as Facebook or Twitter serve as an important interface between companies and customers. This interface generates large amounts of data that need to be analysed and interpreted, as a company can strongly benefit from these data. In addition to structured social media data (e.g. timestamps, like counts, etc.), especially unstructured text data contain interesting contents for companies. Posts and comments often include an user´s major wishes, ideas, and expectations towards products, services, or a company in general (Hienerth et al., 2011, Sigala, 2012a, Sigala, 2012b). The so called 'Voice of the Customers' can be used to adjust marketing campaigns, to identify and support the position in the market and to adjust product features on customers´ favourability. However, to uncover this useful information from the large amount of data requires considerable effort (Dahal et al., 2019, Kumar and George, 2007, Womack and Jones, 1996). To avoid this problem, automated analysis of social media data such as social network analysis, sentiment analysis, and topic modelling can be conducted. Especially through the latter one, valuable information for companies can be extracted, as this technique is able to identify various (discussion) topics, perceptions, and opinions (Dahal et al., 2019, Lozano et al., 2017). However, automated analysis such as topic modelling are often complicated, as many companies are not familiar with the applied techniques, its implementation, and its purpose (Dai et al., 2011).

2.2 Topic modelling

Topic modelling aims to determine content structures in underlying document collections. Hereby, topic modelling refers to the use of generative probability models for determining latent relationships within a corpus of text data. The dataset under investigation is to be seen as a mixture of individual documents, where each document affects several corpus-wide topics, that in turn consist of frequently occurring words within the dataset (Blei, 2012). LDA can be considered as one of the most fundamental works in the topic discovery research area, wherefore a growing number of investigations currently uses this technique proposed by Blei et al. (2003). The authors describe their probabilistic model as *'a three-level hierarchical Bayesian model, in which each item of a collection is modeled as a finite mixture over an underlying set of topics'* (p. 993). This means that LDA is a generative model that is based on the assumption that documents are represented by a collection of different, latent topics. Each topic will be represented as a probability distribution over all words of a corpus (Krestel et al., 2009). LDA is frequently used in marketing to identify important issues for the adaptation of marketing campaigns or to identify product and service features currently being discussed (Chae and Park, 2018, Gao et al., 2012, Jeong et al., 2019, Ko et al., 2017, Luo et al., 2015, Xu and Xiong, 2020, Yu et al., 2019). However, the LDA as proposed by Blei et al. (2003) cannot represent all use cases (e.g. mapping hierarchies).
Thus, there are adaptations and extensions that are based on the probabilistic model of Blei et al. (2003). Generally, it can be differentiated between (I) topic-based extensions, (II) document-based extensions and (III) word-based extensions (cf. Liu et al., 2016). (I) Topic-based extensions derive structures and dependencies within the latent topics of a document (cf. Rathore et al., 2018, Rathore and Ilavarasan, 2017, Tuarob and Tucker, 2015). (II) Document-based extensions have the ability to incorporate an additional parameter into the model building (cf. Cheng et al., 2020, Lozano et al., 2017, Zhang et al., 2016). (III) The word-based extensions compute n-grams instead of Bag of Words (BoW) to incorporate the order of words in a document within the model generation procedure (cf. Wallach, 2006).
As can be concluded from previous research, the LDA approach is the predominantly applied technique (cf. Eickhoff and Neuss, 2017). Vakansky and Kumar (2020) provide a good overview of existing topic modelling techniques and accordingly develop a decision tree model supporting the selection of a

technique. These authors also include adaptations and extensions of the basic LDA for recommending an appropriate technique. However, their recommendations are theoretical in nature and not derived from practical applications of the techniques. There are numerous topic modelling techniques in the literature, but the number of papers comparing them is scare. Further, it is not clear if the theoretical overview and the decision tree of Vakansky and Kumar (2020) can withstand empirical investigations. The authors only conditionally contribute to the operational application of topic modelling as recommendations should ideally be deduced from the analysis of real-world social media data.

2.3 Corporate use cases and requirements of marketing

Within the prevailing literature, we identified papers that deal with utilising topic modelling techniques in marketing. These applications of topic modelling described in the identified papers cover a wide range of marketing tasks. In particular, three main use cases could be identified: (1) content extraction is concerned about consolidating insights of topics discussed in written social media data. By investigating brand-related content from social media, (1) content extraction enables marketing representatives to develop an understanding of topics and themes, e.g. sustainability or product feature favourability, that are discussed by customers and parties of interest. Thereby, companies can improve their external presentation by, e.g. emphasising the ecological superiority of the own products in brand communications (Chae and Park, 2018), and guide future product planning initiatives by putting more focus on product features that are appreciated by customers (Cirqueira et al., 2017, Irawan et al., 2020, Ko et al., 2017, Rathore et al., 2018). Providing a current overview of customers' major wishes, ideas, and thoughts is therefore a central requirement for these techniques as the ever-increasing amount of social media data along with the breadth of the user base may hinder marketing departments to focus on the essential aspects (cf. Cirqueira et al., 2017, Ko et al., 2017, Lee et al., 2016, Liu et al., 2017, Rathore et al., 2018, Rathore and Ilavarasan, 2017). Hereby, (1) content extraction is of a retrospective nature and is less concerned with speed. Because of the abundance of information provided by social media and the necessity to cover as much useful information as possible, the main aim is to cover the relevant and most frequent topics embedded in social media texts (cf. Gao et al., 2012, Ibrahim and Wang, 2019, Irawan et al., 2020, Wang et al., 2016, Yang et al., 2016). Therefore, we assume training time of the topic modelling techniques as secondary, because there are the two central quality dimensions, relevance and dominance of topics, as the basis for well-funded decisions. Furthermore, techniques for (1) content extraction need to support comparisons with competitors, assessments of a company´s position in the market (Aiello et al., 2013) and effectively support product and service opportunities generation (Ko et al., 2017, Rathore et al., 2018, Rathore and Ilavarasan, 2017).

The second use case we could identify is (2) trend analysis. It has in common with (1) content extraction that it deals with extracting topics from large amounts of written social media data. However, (2) trend analysis focuses on keeping track of emerging trends and their development, while the results of (1) content extraction, are instead point-in-time snapshots of the contents that do not illustrate the dynamic courses of the topics. Marketing representatives applying (2) trend analysis related techniques strive to track how certain topics (e.g., product favourability or customer satisfaction) evolve geographically and temporally (Hong et al., 2012, Jeong et al., 2019, Lozano et al., 2017, Zhang et al., 2017). Thereby, marketing departments can enhance the effectiveness of brand message placement and the allocation of appropriate resources to marketing campaigns depending on geographical and temporal developments. Furthermore, linking topics and contents to groups of interested parties and customers enables companies to adapt brand messages to meet the respective target groups´ expectations and attitudes (Zhang et al., 2016). Therefore, topic modelling techniques for (2) trend analysis need to flexibly incorporate different parameters like authors (Zhang et al., 2016), locations (Cheng et al., 2020, Hong et al., 2012, Lozano et al., 2017, Wang et al., 2007), or time (Cheng et al., 2020, Wang et al., 2012, Zhang et al., 2017) into the model building procedure. Ideally, it should even be possible to include implicit parameters such as places or times mentioned in the texts (Lozano et al., 2017). Since trends describe current issues that influence customers' decision-making, companies want to respond to the resulting customer demands to adapt e.g. marketing campaigns (cf. Luo et al., 2015, Rathore et al., 2018, Zhang et al., 2015, Zhong and Schweidel, 2020). Compared to (1) content extraction, especially with fast sequenced social

media data and topics of interests changing quickly, it is necessary to keep track of trends and topic transitions (cf. Wang et al., 2012, Zhang et al., 2017). The requirement to be able to react as quickly as possible is further intensified by reduced product lifecycles and globalised business environments that have made customer needs more dynamic (Jeong et al., 2019). For this reason, the applied topic modelling techniques must provide short training times (cf. Cheng et al., 2020, Wang et al., 2012, Zhang et al., 2017) and as well support quick comprehension of the extracted topics (cf. Jeong et al., 2019, Lozano et al., 2017) in order to keep up with the speed of the trends.
Compared to (3) content structuring, (2) trend analysis captures the dynamic courses of topics and do not collect hierarchical relationships and correlations. Topic modelling techniques for (3) content structuring enable deeper insights from textual social media by extracting not only the topics (cf. (1) content extraction) but also relationships and connections between them. In this way, (3) content structuring supports decisions that need to connect different aspects with each other (e.g. identifying the influence of different product features on customers´ favourability) (Rathore et al., 2018). Mining the inter-relatedness of individual topics can help to detect subtopics (Anoop et al., 2015, Nolasco and Oliveira, 2019, Park et al., 2015, Rathore and Ilavarasan, 2017, Shahbaznezhad, 2016, Srijith et al., 2017) that can be investigated more closely as driving or inhibiting factors. Hereby, it is also possible to identify niche topics at a finer level of granularity of the topical structure. In line with that, recognising the properties of subevents can enrich the understanding of the main event and to create a powerful knowledge about the scenario (Nolasco and Oliveira, 2019, Srijith et al., 2017). In general, corresponding techniques need to extract topics and at the same time establish connections between them. In line with that, the relationships identified between the topics should be understandable and applicable.
In the next step, the identified requirements (cf. tab. 1) are transformed into corresponding metrics that are necessary for evaluating and comparing the topic modelling techniques in section 5.3.

<table>
<tr><th></th><th>Requirements</th><th>Sources</th><th>Metrics</th></tr>
<tr><td rowspan="3">(1) Content Extraction</td><td>(a) Cover all relevant and the most frequent topics embedded in textual social media data</td><td>Gao et al. 2012, Ibrahim and Wang, 2019, Irawan et al. 2020, Lee et al., 2016, Liu et al., 2017, Wang et al. 2016, Yang et al. 2016</td><td rowspan="3">log-likelihood</td></tr>
<tr><td>(b) Provide a current overview of events and insights about customers' wishes and complaints</td><td>Aiello et al., 2013, Chae and Park, 2018, Cirqueira et al., 2017, Ibrahim and Wang, 2019, Ko et al., 2017, Lee et al., 2016, Liu et al., 2017, Rathore et al., 2018, Rathore and Ilavarasan, 2017</td></tr>
<tr><td>(c) Support comparisons with competitors and assessments of one's position in the market</td><td>Aiello et al., 2013, Ko et al., 2017, Rathore et al., 2018</td></tr>
<tr><td rowspan="5">(2) Trend Analysis</td><td>(d) Contextualise the extracted topics with additional parameters</td><td>Cheng et al., 2020, Hong et al., 2012, Lozano et al., 2017, Luo et al., 2015, Wang et al., 2007, Wang et al., 2012, Zhang et al., 2016, Zhang et al., 2017, Zhong and Schweidel, 2020</td><td rowspan="5">build time, coherence, word intrusion</td></tr>
<tr><td>(e) Support a flexible inclusion of different parameters (e.g. authors, locations or time)</td><td>Cheng et al., 2020, Lozano et al., 2017</td></tr>
<tr><td>(f) Support quick information provision</td><td>Cheng et al., 2020, Wang et al., 2012, Zhang et al., 2017</td></tr>
<tr><td>(g) Support quick comprehension of contents</td><td>Jeong et al., 2019, Lozano et al., 2017</td></tr>
<tr><td>(h) Support continuous tracking of trends and developments</td><td>Wang et al., 2012, Zhang et al., 2016, Zhang et al., 2017, Zhong and Schweidel, 2020</td></tr>
<tr><td rowspan="3">(3) Content Structuring</td><td>(i) Identify niche topics at a finer level of granularity of the topical structure</td><td rowspan="3">Anoop et al., 2015, Nolasco and Oliveira, 2019, Park et al., 2015, Rathore et al., 2018, Rathore and Ilavarasan, 2017, Shahbaznezhad, 2016, Srijith et al., 2017</td><td rowspan="3">topic intrusion</td></tr>
<tr><td>(j) Identify meaningful relationships and the inter-relatedness of topics</td></tr>
<tr><td>(k) Cover all aspects that are semantically related to the extracted topics</td></tr>
</table>

Table 1. Identified corporate use cases and their requirements

Comparison of topic modelling techniques

When evaluating a model with respect to (1) content extraction, the ability of the respective technique to (a) cover all relevant and the most frequent topics embedded in textual social media data is required. An excessive number of topics leads to the generation of not only relevant but also irrelevant topics. If the number of topics is too small, however, the given overview of topics will lack relevant content (cf. Liu et al., 2017, Yang et al., 2016). Therefore, the researcher must ensure that all relevant topics within an underlying dataset are considered within the analysis and thus integrated within the resulting extraction, which (b) provides a comprehensive and current overview of customers' wishes and complaints. This results in a comprehensive, decision driven base of information which supports (c) comparisons with competitors and assessments of one's position in the market. In order to evaluate this descriptive ability, the metric of log-likelihood is used. Using this, it is possible to quantify how accurately a model can represent the underlying data and thus models all relevant information (cf. Daud et al., 2010, Wallach et al., 2009). Furthermore, the evaluation has to consider different circumstances regarding the number of topics to be identified. Therefore, the evaluation of each technique takes place multiple times, with a continuously increasing number of topics. Thus, the strengths and weaknesses in modelling low (high) numbers of topics and thus lowly (highly) differentiated contextual information can be assessed.
When evaluating a model with regard to the described use case of (2) trend analysis, (d) the ability to contextualise the extracted topics and (e) hereby flexibly include different additional parameters (e.g. authors, location or time) is mentioned. Furthermore, corresponding techniques need to enable time-critical reactions to emerging circumstances so that (f) the provision of the topics should be as quick as possible. Therefore, the analysis of the different techniques is twofold. On the one hand, authors refer to the coherence measure (Dahal et al., 2019, Paul and Girju, 2009, Wang et al., 2007), which describes the property of the respective technique to generate topics that correlate well with the human understanding of semantically coherent topics. This results in a semantically meaningful and sound analysis output that (g) supports time-critical decisions and does not need further investigations to be applicable. On the other hand, to validate the calculated reasonability of the respective analysis output, we incorporate word intrusion (Chang et al., 2009). Thus, multiple subjects evaluate the consistency of the extracted topics by analysing the associated words. The objective is to identify the so-called 'intruder' within the topic, which is represented by a single word without contextual relevancy regarding the intruded topic. If the subjects are able to identify the respective intruder, the evaluated topic is consistent with the human understanding of a meaningful and sound topic. Besides the unhindered applicability of the analysis results, we also classified the build time as a time-critical evaluation metric. So, a short build time results in a timely output supporting faster decision-making. Beyond the aforementioned time-critical requirements, the ability to contextualise the extracted topics and (h) to continuously track their development is also required. To assess the possibility of accounting for further contextual information such as geological or time-based data, we qualify the ability to incorporate external information.
Concerning an evaluation of a model with regard to the described use case of (3) content structuring, the ability of the techniques to reveal the underlying structure within the data is focused. Thus, (i) to be able to extract hidden niche topics to identify relationships between the discussed topics, the ability of the respective technique to identify relationships within the data is qualified. Besides the pure ability to identify relationships, the assessment of the (j) meaningfulness of the identified relations is also required. In this regard, we opted to use the topic intrusion approach, which measures how well a topic model's decomposition of a document as a mixture of topics agrees with human associations of topics related within a document (Chang et al., 2009). Using these techniques, the (k) coverage of all aspects within the extracted topics as well as their interrelationships in the relevant document are analysed. The procedure of the analysis is similar to that of word intrusion. Specifically, the subjects are presented the document title alongside a short extract thereof. In addition to the document information, the subject receives four topics, of which three are the most probable topics assigned to the document and the remaining topic embodies the intruder topic to be identified. If the subjects are able to identify the erroneously listed topic, the topics and their contextual relationship to one another are meaningful and sound. A more detailed insight into the tasks of word and topic intrusion can be found in section 5.2.
As our review shows, different use cases and corresponding requirements have been reported in line with the literature that provide insights into the application of topic modelling techniques in marketing. However, the used data sets and the applied topic modelling techniques vary across the different papers,

so that recommendations made on this basis may not be sufficiently reliable. Therefore, within the research at hand, three different topic modelling techniques are applied to the same data set to give recommendations that are not only theory-driven but also based on the results of a data analysis.

3 Procedure of the Research

Our investigation follows the '*Goal Question Metric*' (GQM) approach outlined by Basili et al. (1994) for a systematic development of metrics for conducting a comparison of topic modelling techniques. The GQM approach is based on the idea that measurements in an organisational context depend on a thoroughly defined goal firstly operationalised by relevant enterprise data, which are then interpreted regarding the goal (Basili et al., 1994). Thus, this approach focuses on which informational needs a company exhibit in order to quantify them and consequently examine if the quantified information meets the goals or not. Especially for our investigation, the GQM approach is well qualified as it assures a systematic research procedure in which reproducible results are achieved. According to Basili et al. (1994) the GQM approach is divided into three different levels:

1. Conceptual level (**GOAL**): We set our goal as the development of means to compare different topic modelling techniques with each other regarding different use cases. Therefore, with this investigation we seek to highlight differences in the organisational application between the three different topic modelling techniques LDA, DMR, and PAM.
2. Operational level (**QUESTION**): In order to characterise how the assessment of our goal is performed, we formulate a question. In our study, the requirements of the different use cases for the techniques identified in the literature form the basis for our questions of the GQM approach: 'To what extent do the three topic modelling techniques (LDA, DMR, PAM) meet the requirements of the different use cases (a-k) (cf. tab. 1)?'
3. Quantitative level (**METRIC**): To answer this question, we conducted a quantitative analysis of the topic modelling techniques, that helps us to evaluate to what extent a selected topic modelling technique can meet the requirements. Subsequently, in order to meet these requirements, that we have already identified in a previous step, we now need proper metrics. Therefore, we consulted the research literature and established specific metrics for the requirements to enable a comprehensible comparison. Consequentially, we will be able to formulate recommendations for the allocation of the topic modelling techniques to the use cases.

4 Selection of the Topic Modelling Techniques

As mentioned above, the field of topic modelling has many different techniques, which all try to identify specific topics within large sets of text data by reducing the dimensionality and attaching different weights to the specific data set (Crain et al., 2012). In order to optimally meet the identified use cases, the selection of the techniques to be used is critical to success. Besides the aforementioned LDA with its extensions, a variety of different categories of topic modelling techniques like Latent Semantic Analysis, Probabilistic Latent Semantic Analysis, Correlated Topic Models, Dynamic Topic Models, or Topic Evolution Model exist (cf. Alghamdi and Alfalqi, 2015). Although they potentially could offer benefits in terms of different applications, most approaches lack a ready-to-use implementation or require an advanced technique-related knowledge and therefore suitable applicability for companies is not given. Differently LDA, where a clear dominance in the use has become apparent, as it offers simple applicability and good analysis results (Eickhoff and Neuss, 2017). Due to the multifaceted challenges to be mastered in the analysis of text data, different extensions of the basic LDA procedure have been developed over time which are suitable for the solution of different scenarios depending on their extending characteristics. Generally, a distinction is made between three expanding properties: (I) extension of topic attributes, (II) extension of document attributes, and (III) extension of word attributes (cf. Liu et al., 2016). In order to answer **RQ2**, the identification of topic modelling approaches for processing certain corporate use cases, we decided to compare LDA as well as selected extensions with respect to the identified use cases. Therefore, to achieve optimal coverage of different techniques, we selected one specific technique for each extension class.

The (I) topic-based extensions deal with the mapping of relations within the latent topics of a corpus. In this context, a variety of techniques is highlighted in literature (cf. Griffiths et al., 2004, Liu et al., 2016). The focus is on the identification of relationships between the inferred topics allowing a hierarchical representation of them. Mimno et al. (2007) compared the ability of their PAM algorithm to represent a hierarchical data structure and to predict a topic distribution for new data not included in the training set with a variety of techniques of the same extension class (cf. Mimno et al., 2007). Since PAM was characterised by better evaluation results, we decided to choose PAM as the topic-related extension. PAM represents the relationships among the topics as directed graphs, which allows the representation of a hierarchical structure within the topics.
The extension based on documents (II) enables the consideration of document-specific meta-information such as authors, document titles, points in time, or geographical information (cf. Liu et al., 2016). In this context, approaches such as the author-topic model (cf. Rosen-Zvi et al., 2012), Topics over Time (cf. Wang and McCallum, 2006), embedded topic model (Dieng et al., 2020) as well as DMR (cf. Mimno and McCallum, 2008) are highlighted in the current research literature. Because the technique presented by Mimno and McCallum (2008) is more flexible with respect to the incorporation of additional information as well as performs remarkable in terms of information quality, DMR is chosen for the implementation of (II) extension of document attributes. DMR is an upstream topic model with a particularly attractive technique for integrating any document features. Instead of defining specific random variables in the graphical model for each new document feature, DMR treats the document annotations as features in a log-linear model. The log-linear model parameterises the Dirichlet before the document's topic distribution, making the Dirichlet’s hyperparameter document-specific. Since no assumptions are made about the model structure of new random variables, DMR is flexible to include various types of features, resulting in a flexible use of DMR (Benton and Dredze, 2018).
The above-mentioned topic modelling techniques, which are all based on the BoW approach, do not consider the order of words within a document. This resulted in extension (III), attempting to eliminate the interchangeability of words. Therefore, Wallach (2006) argued that the consideration of word orders in the form of bi-grams can lead to improved results when using a topic modelling approach. Since the consideration of word orders in the form of bi-grams did not show any difference with regard to the generated topics and the underlying topic quality compared to LDA, it will be equated with the use of the basic technique in the following. By choosing these techniques, a selection was made which considers each extension class of the basic approach, whereby a broad spectrum of different techniques is compared with regard to their applicability against diverse use cases.

5 Data Analysis

5.1 Data collection

To identify the potential of the different techniques with regard to the applicability to different use cases, an existing data set of Facebook posts was used (cf. Martinchek, 2017). This comprises 4,155,992 documents from the 15 most popular news services in the United States of America for the period from 2012 to 2016. The raw text dataset contains information such as the respective picture URL or the like count, which are not relevant for the application of the implemented topic modelling approaches. Therefore, to reduce the dimensions of the data, a custom converter was developed and applied to the data. The resulting data set contains three parameters after conversion: the ID of the respective document, the respective year – which serves as feature input to determine the topic relevance at different points in time within the analysis via DMR – and the description – which reflects the actual text of the contribution. The following excerpt from the training data set gives an insight into the data (cf. tab. 2).

ID	Year	Description
52921	2016	Dow Drops More Than 300 Points Following Market Rout...
21049	2014	How to Greet People During Flu Season: Handshake, ...

Table 2. Structure of training data

5.2 Data cleansing and analysing

In order to compare the topic modelling techniques empirically on the basis of their analysis results, the data must first be prepared. With respect to this, we applied tokenisation, stopwords removal and case folding as proposed by Boyd-Graber et al. (2014). As the use of stemming procedures does not improve the interpretability of the results, but can potentially even deteriorate the topic stability (Schofield and Mimno, 2016), we did not incorporate stemming.
By conducting this comparison between LDA, DMR, and PAM based on the mentioned evaluation measures, we aim to reveal the strengths and weaknesses of the different techniques in terms of identifying embedded topics within written social media data. As it is necessary to provide similar conditions for a comparison to be valid, all techniques were configured with their default parameters and trained with iteratively increasing numbers of topics. Accordingly, the selected topic range includes 10, 30, 50, 100, and 300 topics (*k*). Furthermore, all evaluation metrics were validated by cross-validation to eliminate the choice of a potentially non-representative test dataset (Bramer, 2007). The evaluation approach is further distinguishing between intrinsic and extrinsic measures. Intrinsic evaluations measure the performance of a component on its defined subtask, usually against a defined standard in a reproducible laboratory setting. Extrinsic evaluations focus on the component's contribution to the performance of a complete application, which often involves the participation of a human in the loop (Resnik et al., 2006).

#	log-likelihood measurements ($\cdot 10^6$)			coherence measurements			build time (min)		
k	LDA	DMR	PAM	LDA	DMR	PAM	LDA	DMR	PAM
10	- 4.86	- 5.03	- 4.85	- 197.96	- 198.50	- 208.34	342	363	832
30	- 4.49	- 4.61	- 4.41	- 232.84	- 220.75	- 236.81	571	634	1,264
50	- 4.29	- 4.40	- 4.19	- 244.29	- 217.90	- 251.17	912	1,083	1,992
100	- 4.06	- 4.10	- 3.88	- 241.74	- 213.03	- 254.92	1,407	1,732	3,481
300	- 3.91	- 3.62	- 3.67	- 215.02	- 186.95	- 241.52	3,180	3,821	6,984

Table 3. Intrinsic evaluation measurements

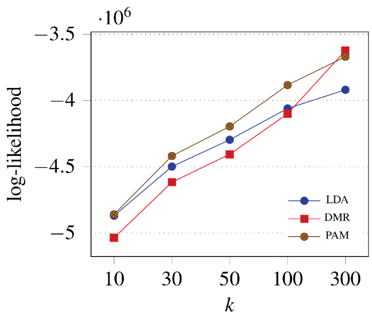


Figure 1. log-likelihood

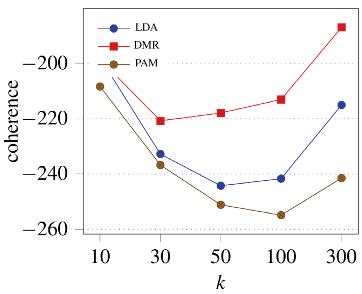


Figure 2. coherence

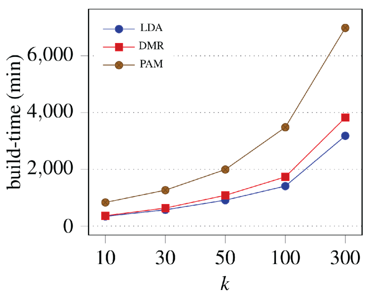


Figure 3. build time

In the first step of the evaluation, the models are compared according to their ability to represent the underlying data. Therefore, the metric of log-likelihood was applied (cf. tab. 3, fig. 1). A higher value represents a better model with regard to the ability to adequately represent the underlying data. The best model here is DMR, with k = 300 (-3.62). LDA performs better than DMR (e.g. k = 30 ≙ -4.86 > -5.03) and worse than PAM (e.g. k = 30 ≙ -4.49 < -4.41) until the number of topics exceeds 100. PAM and LDA have in common that, as the number of topics increases, the log-likelihood improves iteratively (cf. fig. 1). The increase in the number of topics for DMR results in a continuous improvement of the log-likelihood as well. Nevertheless, DMR always outperforms PAM and LDA when k >~ 300. This result implies that DMR is the technique that can be recommended when tasks require many different topics to be identified. However, within tasks that require a lower number of topics to be generated, our results indicate that PAM and LDA may be superior to DMR.
Besides the general ability of the models to adequately represent the underlying data, the semantic quality of the generated topics was evaluated. For that reason, the coherence measure was applied (cf. fig.

2). Therefore, we conclude that DMR generates topics, that are more coherent in general. LDA has a slightly better coherence for k = 10. However, any modelling of k > 10 is outperformed by DMR. Further, PAM performed worst in terms of coherence. Here, all different circumstances of topics are dominated by the other techniques. Nevertheless, PAM generated its most relatable topics for k = 10, which leads to the conclusion, that PAM, similar to LDA, exhibits its strength in modelling low topic dimensions.
To react in a time-critical manner to changing circumstances, it is also necessary to acquire supporting information as quickly as possible. Therefore, to measure the time an approach needs to extract the required information (cf. fig. 3), the build time of each technique is tracked. The respective build time includes data preprocessing, the actual model building and cross-validation. A lower build time indicates a quicker model training, which results in faster information provision. In this regard, LDA outperforms the other two techniques in terms of information extraction time. The difference between LDA and DMR in the lower range of topics is negligible (21 min), but the higher the value of topics, the larger the difference. The difference in the context of k = 300 amounts to 641 minutes. Because of the hierarchies to be modelled by PAM, it generally takes much longer to extract the respective information in comparison to LDA and DMR. Here, PAM needs at least twice as much time under almost all circumstances.
Besides the intrinsic measurements of log-likelihood, coherence and build time, the topic modelling techniques were further assessed by humans to evaluate the semantic quality of the generated topics as well as their interrelationships. Therefore, word and topic intrusion procedures were performed. The respective survey was undertaken by two researchers and administered to 18 participants, all of whom evaluated the semantic coherence of three randomly selected topics (word intrusion) and the decomposition of a single document into its topics and the corresponding relationships (topic intrusion). Each survey ranged between 38 and 51 minutes. The topics as well as the respective excerpts of a document were extracted randomly for each trained model. Further, to account the inter-rater reliability of the results, all participants evaluated the same set of topics or documents respectively for each trained model. Regarding the word intrusion task, the subjects had to identify the intruder within the topics that did not cohere to the semantics of the other presented words. The corresponding results for word intrusion are calculated as the sum of the correctly classified intruders by the test subjects in relation to the total number of tests per model. By analysing three topics per trained model, a total of 810 individual observations were carried out. Regarding the topic intrusion task, the participants had to identify the intruding topic by reading a document title alongside a short extract thereof. The respective document was randomly extracted for each trained model. All participants evaluated the decomposition of the same documents. By this, a total of 270 individual observations are accomplished. The respective results of the topic intrusion task are calculated as the amount of correctly classified intruder topics in relation to the total number of observations per model.

	Word intrusion (%)			Topic intrusion (%)		
k	LDA	DMR	PAM	LDA	DMR	PAM
10	74.0	85.1	53.7	83.3	72.2	83.3
30	72.2	79.6	57.4	66.6	61.1	72.2
50	62.9	75.9	51.8	72.2	72.2	72.2
100	64.8	70.3	48.1	55.5	66.6	55.5
300	55.5	77.7	46.2	55.5	72.2	44.4

Table 4. Extrinsic evaluation measurements

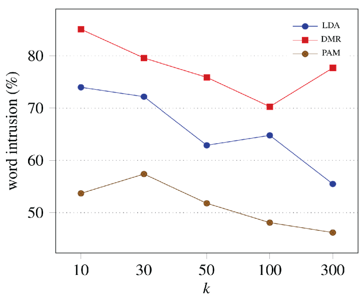


Figure 4. Word intrusion

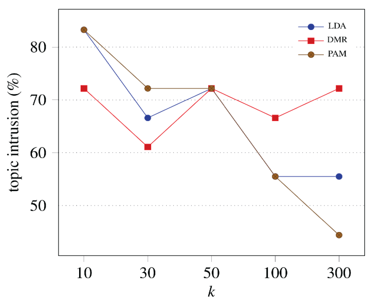


Figure 5. Topic intrusion

Comparison of topic modelling techniques

Regarding the word intrusion task, it becomes apparent that the previously determined values of coherence are in line with human understanding so DMR generates the most semantically coherent topics. Specifically, the best topics are generated by DMR with k = 10 (cf. fig. 4). The achieved model precision is 85.1% for k = 300. The evaluated minimum precision of DMR is 70.3%. As the results for word intrusion do not show any distinct extremes for different values of k, every model of DMR seems to produce a coherent word-topic distribution. Besides that, LDA also proves to generate the most coherent topics for k = 10 (74.0%). The higher the number of topics to be generated, the more difficult it becomes for the subjects to determine the intruder. This leads to the assumption, that the topics generated by LDA will be increasingly difficult to interpret with a rising number of k, thereby losing their semantic coherence and therefore their meaningfulness. The minimal precision of LDA is recorded for k = 300. Here, only 30 of 54 measurements are classified correctly, resulting in a model precision of 55.5%. There is a clear discrepancy in the quality of topics generated by PAM. The semantically most coherent topics are determined by the model with k = 30. The model precision achieved is 57.4%. The minimum semantic coherence of the topics is with k = 300, with a precision of 46.2%. Thus, it can be concluded that PAM, similar to LDA, has its strength in modelling a smaller number of topics but is clearly inferior to DMR for larger numbers. Besides the evaluation of the topic quality, their decomposition and interrelationships were evaluated. Therefore, topic intrusion tasks were conducted, where the best results were obtained with 83.3% for LDA and PAM with k = 10. Further, it is remarkable, that all techniques achieve the same score with k = 50 (cf. fig. 5), which leads to the conclusion, that all techniques generate consistent topic decompositions for a medium number of topics. The higher the number of topics to be generated, the worse LDA and PAM perform. The worst result is achieved by PAM with k = 300. That decrease of PAM for larger numbers of topics to be generated can be traced back to the formation of the many hierarchical levels, since a high number of k also means that a correspondingly large number of relationships between the different topics must be inferred.

6 Discussion

By evaluating the different techniques, the strengths and weaknesses (cf. tab. 5) of them were identified, that provide information about the applicability of the techniques related to the identified corporate use cases (1) content extraction, (2) trend analysis as well as (3) content structuring settled in marketing.

	Requirements	**LDA**	**DMR**	**PAM**
(1) Content Extraction	(a) Cover all relevant and the most frequent topics embedded in textual social media data	●	●	◗
	(b) Provide a current overview of events and insights about customers' wishes and complaints	●	●	◗
	(c) Support comparisons with competitors and assessments of one's position in the market	●	●	◗
(2) Trend Analysis	(d) Contextualise the extracted topics with additional parameters	○	●	○
	(e) Support a flexible inclusion of different parameters (e.g. authors, locations or time)	○	●	○
	(f) Support quick information provision	●	●	○
	(g) Support quick comprehension of contents	●	●	○
	(h) Support continuous tracking of trends and developments	◗	●	○
(3) Content Structuring	(i) Identify niche topics at a finer level of granularity of the topical structure	○	◗	●
	(j) Identify meaningful relationships and the inter-relatedness of topics	○	○	●
	(k) Cover all aspects that are semantically related to the extracted topics	○	◗	●

Table 5. Results of the comparison ●: applies fully ◗: applies partly ○: applies not

Comparison of topic modelling techniques

For the first use case – (1) content extraction – the models´ ability to provide an adequate overview of customers' major wishes, ideas, and thoughts to cover all relevant topics embedded in a collection of written social media data was evaluated. Therefore, we applied the metric of log-likelihood, which describes the ability of a model to represent the underlying data as appropriately as possible. This guarantees that the generated output contains all necessary information and that no relevant topics are missing. The standard procedure of LDA as well as that of PAM present their strength compared to DMR in the extraction of topics in the lower range of generated topics. DMR, in contrast, has its strength in the representation of large numbers of topics. This leads to the conclusion that the selection of the best fitting technique depends on the needs to be fulfilled within the extraction task:

- If the content must be very specific, e.g. to support a comparison with competitors and assessments of one's position in the market, a high number of topics is required. So, DMR should be considered.
- If the task requires an extraction on a more abstract level, a low number of topics will mostly be satisfying. If so, the usage of LDA and PAM could be considered. Due to the higher semantic coherence within the topics generated by LDA (cf. tab. 3, tab. 4, fig. 2, fig. 4), the use of LDA is recommended with respect to (1) content extraction for small dimensions.
- The use of PAM is considered as partly applicable, but not recommended, as it is outperformed by LDA and DMR for the criteria being evaluated.

Regarding the second use case (2) trend analysis, the ability of the techniques to generate immediately meaningful and sound output was analysed, which results e.g., in a quick decision supporting information base. This is indispensable regarding the need of time-critical actions within the volatile characteristic of trends. Hereby, DMR shows an advantage leading to the conclusion, that it generates the most reasonable topics (cf. tab. 3, fig. 2). To validate the collected intrinsic evaluation results, the techniques were further evaluated by humans regarding their semantic quality and soundness in an extrinsic way. Here, DMR could be confirmed to generate the most comprehensible output (cf. tab. 4, fig. 4). Further, LDA and PAM show their strength for modelling low numbers of topics as both techniques achieve their best results in the range of 10 (LDA) and 30 (PAM) topics. In addition to direct applicability, the time required for a technique to provide the necessary information was also accounted for. In this respect, LDA provides the fastest output, followed by DMR. The measured discrepancy between these two techniques is negligible for a small number of topics ($k = 10 \triangleq 21$ min.), but the larger the number of topics, the larger the gap ($k = 300 \triangleq 641$ min.). Since the focus of trend analysis is on identifying individual trends and tracking their development, the number of topics will not be that high. Therefore, LDA and DMR are considered capable of reacting to rapidly changing circumstances. PAM, however, requires at least twice as much time for each condition and is therefore not suitable. The strengths and weaknesses of the techniques in the context of (2) trend analysis are represented as follows:

- DMR provides the most reasonable and meaningful topics and is further able to provide them quickly. Additionally, DMR has the advantage of taking external parameters into account. This allows, e.g., tracking the development of topics over a certain period of time or based on geolocation data. Therefore, DMR should be considered regarding (2) trend analysis.
- If the contextualisation of topics does not apply, LDA can also be used for tracking trends and their development, as it can be used to quickly identify meaningful and sound analysis results. Therefore, LDA can be seen as partly applicable regarding trend analysis.
- PAM is not suitable due to the amount of time required and the lack of ability to contextualise topics.

Regarding the last identified use case – (3) content structuring - the ability of the techniques to reveal meaningful relationships and the inter-relatedness of topics was evaluated. Therefore, the extrinsic evaluation metric of topic intrusion was applied. By doing so, it could be guaranteed that all semantically related aspects were extracted. Here, PAM shows a slightly better result than the two remaining techniques for a small number of topics to be generated (cf. tab. 4, fig. 5). Further, DMR underlined its strength in the representation of a high number of topics. A large advantage of using PAM is represented by its ability to model hierarchical structures within the topics themselves. Thus, it is possible to extract general topics as well as their respective subtopics, whereby the topics can be divided into different, thematically consistent groupings that can support the identification of niche topics at a finer level of granularity. The elicited strengths and weaknesses of the techniques are listed in the following:

- As the best evaluation results regarding the topic intrusion task were achieved by PAM, we recommend this technique as an approach to support (3) content structuring.
- Besides that, DMR can also be applied if the number of topics to be generated reaches a comparatively high number. Thus, the applicability is considered partial.
- Since LDA is surpassed by the two techniques here, the application is not considered suitable.

In summary, all the investigated techniques have different strengths and weaknesses in their applicability to the identified use cases. Nevertheless, LDA showed its strength in modelling low-dimensional topics. In comparison, DMR showed to be superior in representing high-dimensional topics. Regarding trend analysis, DMR showed its strength within the generation of semantically meaningful results. For content structuring tasks, PAM showed superior results in extracting meaningful relationships compared to LDA and DMR.

7 Conclusion and Outlook

Analysing written social media data with automated techniques has massively gained in importance as being aware of customers' wishes is no longer sustainable with manual analysis due to the sheer volume of available posts. Topic modelling has shown to be an adequate instrument to support these tasks by extracting the topics discussed within documents (e.g. Eickhoff and Neuss, 2017, Vayansky and Kumar, 2020). However, it can be observed that especially LDA has been given particular attention for marketing-specific applications. Furthermore, since LDA cannot cover all fields of the identified use cases, marketing tasks may only be supported to a limited extent by automated approaches.

Within this work, the use cases (1) - (3) were identified from the literature due to the frequency that the identified papers related to them. For each of these use cases, corresponding requirements were identified and assigned to different metrics (log-likelihood, coherence, build time, word and topic intrusion) for evaluating these topic modelling techniques (cf. **RQ1**). Thus, LDA, DMR and PAM were applied to a real-world data set, evaluated and compared with each other. Thereafter, this work gives recommendations regarding which topic modelling technique could be applied for which use case (cf. **RQ2**).

Through our comparison of topic modelling techniques, practitioners are given means to select a technique that can best support their daily business activities. Decision makers in marketing can classify their concrete task into one of the three identified use cases and derive a recommendation for a suitable technique. Tasks in marketing, which can be enriched by topic modelling, can thus be supported more optimally and thus the performance of this division, which is so important for companies, can be increased. Beyond creating value for practitioners, theoretical contributions in the research area of IS are also provided. First, based on the use cases we derived several requirements for topic modelling techniques and assigned several evaluation criteria to each of them. Second, in order to provide recommendations, we compared LDA, DMR, and PAM with each other regarding five different evaluation metrics by analysing a real-world data set. This comparison is based on the GQM approach and assures therefore a systematic research procedure in which reproducible results are achieved. Based on this comparison potentials for further enhancements (e.g. considering a faster build time for PAM when $k > 50$) could be evolved. Third, we further deduced strengths (such as DMR should be considered regarding trend analysis, cf. section 6) and weaknesses (such as PAM is not suitable for trend analysis due to the amount of time required and the lack of ability to contextualise topics) of the three topic modelling techniques which is valuable for both researchers and practitioners.

There are some limitations to this study: first, the number of papers we incorporated in identifying use cases and related requirements for topic modelling within marketing is limited. Nevertheless, these requirements enabled the central metrics of topic modelling techniques to be assigned and generally valid recommendations for appropriate procedures to be derived. Second, the number of topic modelling techniques being compared is limited to three. Although the extensions proposed by Liu et al. (2016) could thus be covered to a large extent, we plan to include further techniques in future analysis for each extension and thereby further refine our recommendations.

References

Aiello, L. M., Petkos, G., Martin, C., Corney, D., Papadopoulos, S., Skraba, R., Göker, A., Kompatsiaris, I. and Jaimes, A. (2013). Sensing trending topics in Twitter. *IEEE Transactions on Multimedia,* 15 (6), 1268-1282.

AlFalahi, K., Atif, Y. and Abraham, A. (2014). Models of influence in online social networks. *International journal of intelligent systems,* 29 (2), 161-183.

Alghamdi, R. and Alfalqi, K. (2015). A survey of topic modeling in text mining. *Int. J. Adv. Comput. Sci. Appl.(IJACSA)*, 6 (1).

Anoop, V., Asharaf, S. and Alessandro, Z. (2015). Generating and visualizing topic hierarchies from microblogs: An iterative latent dirichlet allocation approach. *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 824-828.

Basili, G., Caldiera, V. and Rombach, H. D. (1994). The goal question metric approach. *Encyclopedia of software engineerin*g, 528-532.

Benton, A. and Dredze, M. (2018). Deep Dirichlet multinomial regression. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1, 365-374.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM,* 55 (4), 77-84.

Blei, D. M., Ng, A. Y. and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 993-1022.

Boyd-Graber, J., Mimno, D. and Newman, D. (2014). Care and feeding of topic models: Problems, diagnostics, and improvements. *Handbook of mixed membership models and their applications*, 225255.

Bramer, M. (2007). Principles of data mining. *Springer*, ISBN: 978-1-84628-765-7.

Chae, B. K. and Park, E. O. (2018). Corporate social responsibility (CSR): A survey of topics and trends using Twitter data and topic modeling. *Sustainability,* 10 (7), 2231.

Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J. L. and Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. *Advances in neural information processing systems*, 288-296.

Cheng, L., Li, J., Candan, K. S. and Liu, H. (2020). Tracking Disaster Footprints with Social Streaming Data. *Proceedings of the AAAI Conference on Artificial Intelligence*, 370-377.

Chinnov, A., Kerschke, P., Meske, C., Stieglitz, S. and Trautmann, H. (2015). An Overview of Topic Discovery in Twitter Communication through Social Media Analytics. *Americas Conference On Information Systems (AMCIS) 2015.*

Cirqueira, D., Pinheiro, M., Braga, T., Jacob Jr, A., Reinhold, O., Alt, R. and Santana, Á. (2017). Improving relationship management in universities with sentiment analysis and topic modeling of social media channels: learnings from ufpa. *Proceedings of the International Conference on Web Intelligence*, 998-1005.

Coelho, L. P., Peng, T. and Murphy, R. F. (2010). Quantifying the distribution of probes between subcellular locations using unsupervised pattern unmixing. *Bioinformatics,* 26 (12), 7-12.

Crain, S. P., Zhou, K., Yang, S.-H. and Zha, H. (2012). Dimensionality reduction and topic modeling: From latent semantic indexing to latent dirichlet allocation and beyond. *Mining text data*, Springer, 129-161.

Culnan, M. J., McHugh, P. J. and Zubillaga, J. I. (2010). How large US companies can use Twitter and other social media to gain business value. *MIS Quarterly Executive,* 9 (4).

Dahal, B., Kumar, S. A. and Li, Z. (2019). Topic modeling and sentiment analysis of global climate change tweets. *Social Network Analysis and Mining*, 9 (1), 1-20.

Dai, Y., Kakkonen, T. and Sutinen, E. (2011). MinEDec: a decision-support model that combines text-mining technologies with two competitive intelligence analysis methods. *International Journal of Computer Information Systems and Industrial Management Applications,* 3 (3), 165-173.

Daud, A., Li, J., Zhou, L. and Muhammad, F. (2010). Knowledge discovery through directed probabilistic topic models: a survey. *Frontiers of computer science in China,* 4 (2), 280-301.

Debortoli, S., Müller, O., Junglas, I. and vom Brocke, J. (2016). Text mining for information systems researchers: An annotated topic modeling tutorial. *Communications of the Association for Information Systems,* 39 (1), 7.

Di Gangi, P. M., Wasko, M. M. and Hooker, R. E. (2010). GETTING CUSTOMERS' IDEAS TO WORK FOR YOU: LEARNING FROM DELL HOW TO SUCCEED WITH ONLINE USER INNOVATION COMMUNITIES. *MIS Quarterly Executive,* 9 (4).

Dieng, A. B., Ruiz, F. J. and Blei, D. M. (2020). Topic modeling in embedding spaces. *Transactions of the Association for Computational Linguistics*, 8, 439-453.

Eickhoff, M. and Neuss, N. (2017). Topic modelling methodology: its use in information systems and other managerial disciplines. *European Conference On Information Systems (ECIS) 2017.*

Einwiller, S. A. and Steilen, S. (2015). Handling complaints on social network sites–An analysis of complaints and complaint responses on Facebook and Twitter pages of large US companies. *Public Relations Review,* 41 (2), 195-204.

Gao, W., Li, P. and Darwish, K. (2012). Joint topic modeling for event summarization across news and social media streams. *Proceedings of the 21st ACM international conference on Information and knowledge management*, 1173-1182.

Ghosh, D. and Guha, R. (2013). What are we 'tweeting'about obesity? Mapping tweets with topic modeling and Geographic Information System. *Cartography and geographic information science,* 40 (2), 90-102.

Girdhar, Y., Giguere, P. and Dudek, G. (2013). Autonomous adaptive underwater exploration using online topic modeling. *Experimental Robotics*, 789-802.

Grégoire, Y., Salle, A. and Tripp, T. M. (2015). Managing social media crises with your customers: The good, the bad, and the ugly. *Business Horizons,* 58 (2), 173-182.

Griffiths, T. L., Jordan, M. I., Tenenbaum, J. B. and Blei, D. M. (2004). Hierarchical topic models and the nested chinese restaurant process. *Advances in neural information processing systems*, 17-24.

Guesalaga, R. (2016). The use of social media in sales: Individual and organizational antecedents, and the role of customer engagement in social media. *Industrial Marketing Management,* 54, 71-79.

Hanna, R., Rohm, A. and Crittenden, V. L. (2011). We're all connected: The power of the social media ecosystem. *Business Horizons,* 54 (3), 265-273.

Hienerth, C., Keinz, P. and Lettl, C. (2011). Exploring the nature and implementation process of user-centric business models. *Long Range Planning,* 44 (5-6), 344-374.

Hong, L., Ahmed, A., Gurumurthy, S., Smola, A. J. and Tsioutsiouliklis, K. (2012). Discovering geographical topics in the twitter stream. *Proceedings of the 21st international conference on World Wide Web*, 769-778.

Hong, L. and Davison, B. D. (2010). Empirical study of topic modeling in twitter. *Proceedings of the first workshop on social media analytics*, 80-88.

Ibrahim, N. F. and Wang, X. (2019). A text analytics approach for online retailing service improvement: Evidence from Twitter. *Decision Support Systems,* 121, 37-50.

Irawan, M., Wijayanto, R., Shahab, M., Hidayat, N. and Rukmi, A. (2020). Implementation of social media mining for decision making in product planning based on topic modeling and sentiment analysis. *Journal of Physics: Conference Series*, 012068.

Jeong, B., Yoon, J. and Lee, J.-M. (2019). Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis. *International Journal of Information Management,* 48, 280-290.

Kaplan, A. M. and Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons,* 53 (1), 59-68.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P. and Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons,* 54 (3), 241-251.

Ko, N., Jeong, B., Choi, S. and Yoon, J. (2017). Identifying product opportunities using social media mining: application of topic modeling and chance discovery theory. *IEEE Access,* 6, 1680-1693.

Krestel, R., Fankhauser, P. and Nejdl, W. (2009). Latent dirichlet allocation for tag recommendation. *Proceedings of the third ACM conference on Recommender systems*, 61-68.

Kumar, V. and George, M. (2007). Measuring and maximizing customer equity: a critical analysis. *Journal of the Academy of Marketing Science,* 35 (2), 157-171.

Lee, T.-H., Sung, W.-K. and Kim, H.-W. (2016). A text mining approach to the analysis of information security awareness: Korea, United States, and China. *Pacific Asia Conference On Information Systems (PACIS) 2016.*

Li, W. and McCallum, A. (2006). Pachinko allocation: DAG-structured mixture models of topic correlations. *Proceedings of the 23rd international conference on Machine learning*, 577-584.

Liu, L., Tang, L., Dong, W., Yao, S. and Zhou, W. (2016). An overview of topic modeling and its current applications in bioinformatics. *SpringerPlus,* 5 (1), 1608.

Liu, X., Burns, A. C. and Hou, Y. (2017). An investigation of brand-related user-generated content on Twitter. *Journal of Advertising,* 46 (2), 236-247.

Lozano, M. G., Schreiber, J. and Brynielsson, J. (2017). Tracking geographical locations using a geo-aware topic model for analyzing social media data. *Decision Support Systems,* 99, 18-29.

Luo, J., Pan, X. and Zhu, X. (2015). Identifying digital traces for business marketing through topic probabilistic model. *Technology Analysis & Strategic Management,* 27 (10), 1176-1192.

Marshall, G. W., Moncrief, W. C., Rudd, J. M. and Lee, N. (2012). Revolution in sales: The impact of social media and related technology on the selling environment. *Journal of Personal Selling & Sales Management,* 32 (3), 349-363.

Martinchek, P. (2016). 2012-2016 Facebook Posts. URL: *https://data.world/martinchek/2012-2016-facebook posts*.

McDonald, M. and Aron, D. (2012). Amplifying the enterprise: the 2012 CIO agenda. *Gartner Executive Programs*, Gartner Inc.

Mimno, D., Li, W. and McCallum, A. (2007). Mixtures of hierarchical topics with pachinko allocation. *Proceedings of the 24th international conference on Machine learning*, 633-640.

Mimno, D. M. and McCallum, A. (2008). Topic models conditioned on arbitrary features with Dirichlet-multinomial regression. *UAI*, 411-418.

Mount, M. and Martinez, M. G. (2014). Social media: A tool for open innovation. *California Management Review,* 56 (4), 124-143.

Nolasco, D. and Oliveira, J. (2019). Subevents detection through topic modeling in social media posts. *Future Generation Computer Systems*, 93, 290-303.

Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., Ungar, L. H. and Seligman, M. E. (2015). Automatic personality assessment through social media language. *Journal of personality and social psychology,* 108 (6), 934-952.

Paul, M. and Girju, R. (2009). Cross-cultural analysis of blogs and forums with mixed-collection topic models. *Proceedings of the 2009 conference on empirical methods in natural language processing*, 1408-1417.

Piller, F. T., Vossen, A. and Ihl, C. (2012). From social media to social product development: the impact of social media on co-creation of innovation. *Die Unternehmung,* 65 (1).

Ramage, D., Manning, C. D. and Dumais, S. (2011). Partially labeled topic models for interpretable text mining. *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, 457-465.

Rathore, A. K., Das, S. and Ilavarasan, P. V. (2018). Social media data inputs in product design: Case of a smartphone. *Global Journal of Flexible Systems Management,* 19 (3), 255-272.

Rathore, A. K. and Ilavarasan, P. V. (2017). Social media analytics for new product development: case of a pizza. *International Conference on Advances in Mechanical, Industrial, Automation and Management Systems (AMIAMS)*, 213-219.

Resnik, P., Niv, M., Nossal, M., Schnitzer, G., Stoner, J., Kapit, A. and Toren, R. (2006). Using intrinsic and extrinsic metrics to evaluate accuracy and facilitation in computer-assisted coding. *Perspectives in Health Information Management Computer Assisted Coding Conference Proceedings*, 2006.

Rosen-Zvi, M., Griffiths, T., Steyvers, M. and Smyth, P. (2012). The author-topic model for authors and documents. *arXiv preprint,* 4169.

Schneider, F., Feldmann, A., Krishnamurthy, B. and Willinger, W. (2009). Understanding online social network usage from a network perspective. *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement*, 35-48.

Schofield, A. and Mimno, D. (2016). Comparing apples to apple: The effects of stemmers on topic models. *Transactions of the Association for Computational Linguistics,* 4, 287-300.

Shahbaznezhad, H. (2016). Fan page management via content generation and feedback strategies. *SSRN*, 2769428.

Sigala, M. (2012a). Exploiting web 2.0 for new service development: Findings and implications from the Greek tourism industry. *International Journal of Tourism Research*, 14 (6), 551-566.

Sigala, M. (2012b). Social networks and customer involvement in new service development (NSD). *International Journal of Contemporary Hospitality Management*.

Srijith, P., Hepple, M., Bontcheva, K. and Preotiuc-Pietro, D. (2017). Sub-story detection in Twitter with hierarchical Dirichlet processes. *Information Processing & Management,* 53 (4), 989-1003.

Statista (2020b). Number of social network users worldwide from 2017 to 2025. URL: *https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/* (visted on 25.10.2020).

Tuarob, S. and Tucker, C. S. (2015). Quantifying product favorability and extracting notable product features using large scale social media data. *Journal of Computing and Information Science in Engineering*, 15 (3).

Vayansky, I. and Kumar, S. A. (2020). A review of topic modeling methods. *Information Systems*, 94, 101582.

Wallach, H. M. (2006). Topic modeling: beyond bag-of-words. *Proceedings of the 23rd international conference on Machine learning*, 977-984.

Wallach, H. M., Murray, I., Salakhutdinov, R. and Mimno, D. (2009). Evaluation methods for topic models. *Proceedings of the 26th annual international conference on machine learning*, 1105-1112.

Wang, C., Wang, J., Xie, X. and Ma, W.-Y. (2007). Mining geographic knowledge using location aware topic model. *Proceedings of the 4th ACM workshop on Geographical information retrieval*, 65-70.

Wang, S., Chen, Z., Fei, G., Liu, B. and Emery, S. (2016). Targeted topic modeling for focused analysis. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 1235-1244.

Wang, X. and McCallum, A. (2006). Topics over time: a non-Markov continuous-time model of topical trends. *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, 424-433.

Wang, Y., Agichtein, E. and Benzi, M. (2012). TM-LDA: efficient online modeling of latent topic transitions in social media. *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, 123-131.

Womack, J. P. and Jones, D. T. (1996). Beyond Toyota: how to root out waste and pursue perfection. *Harvard business review,* 74 (5), 140-172.

Xu, S. and Xiong, Y. (2020). Setting socially mediated engagement parameters: A topic modeling and text analytic approach to examining polarized discourses on Gillette's campaign. *Public Relations Review,* 46 (5), 101959.

Yang, L., Lin, H., Lin, Y. and Liu, S. (2016). Detection and extraction of hot topics on chinese microblogs. *Cognitive Computation,* 8 (4), 577-586.

Yu, D., Xu, D., Wang, D. and Ni, Z. (2019). Hierarchical topic modeling of Twitter data for online analytical processing. *IEEE Access*, 7, 12373-12385.

Zhang, H., Kim, G. and Xing, E. P. (2015). Dynamic topic modeling for monitoring market competition from online text and image data. *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, 1425-1434.

Zhang, P., Gu, H., Gartrell, M., Lu, T., Yang, D., Ding, X. and Gu, N. (2016). Group-based Latent Dirichlet Allocation (Group-LDA): Effective audience detection for books in online social media. *Knowledge-Based Systems,* 105, 134-146.

Zhang, X., Zhao, L., Chen, Z., Boedihardjo, A. P., Dai, J. and Lu, C.-T. (2017). Trendi: Tracking stories in news and microblogs via emerging, evolving and fading topics. *2017 IEEE International Conference on Big Data*, 1590-1599.

Zhong, N. and Schweidel, D. A. (2020). Capturing changes in social media content: a multiple latent changepoint topic model. *Marketing Science,* 39 (4), 827-846.

2.2 Identifying Sentiment Influences Provoked by Context Factors – Results from a Data Analytics Procedure Performed on Tweets

<table>
<tr><td>Adressierte Zielsetzung</td><td>Konzeption und Evaluation eines kontextsensitiven Analyseverfahrens zur Integration von Kontextfaktoren in der Analyse von in UGC. Ziel war es, ein kontextsensitives Verfahren zu entwickeln und mit der Sentimentanalyse zu kombinieren, um systematisch zu untersuchen, wie technische (z. B. Betriebssysteme) und geographische (z. B. Kontinente) Kontextfaktoren die in sozialen Medien artikulierte Stimmung beeinflussen.</td><td>ZS2</td></tr>
<tr><td>Abgeleitete Forschungsfragen</td><td>Welchen Einfluss haben die Kontextfaktoren Betriebssysteme mobiler Endgeräte und der geographische Kontext (Kontinente) auf die in sozialen Medien geäußerten Stimmungen, und wie lassen sich diese Kontextfaktoren in kombinierter Form analysieren?</td><td>FF3</td></tr>
<tr><td>Genutzte Forschungsmethode</td><td colspan="2">Zur Beantwortung von FF2, inwiefern technische und geographische Kontextfaktoren das in UGC artikulierte Sentiment systematisch beeinflussen, wurde ein eigenes Analyseverfahren zur kontextsensitiven Sentimentanalyse konzipiert, umgesetzt und auf einen umfangreichen Social Media Datensatz (> 350 Mio. Tweets) angewendet.

Das Verfahren basiert auf der methodischen Grundlage zur Entwicklung eines Text Mining Verfahrens nach Aggarwal et al. (2012) und verfolgt das Ziel, Kontextinformationen automatisiert in großen Mengen an UGC zu identifizieren und mit den Sentimentausprägungen der Inhalte zu kombinieren. Das mehrstufige Analyseverfahren umfasst folgende Komponenten:

1. Datenextraktion und Preprocessing: Sammlung und Vorverarbeitung eines umfassenden Realweltdatensatzes mit 358.923.210 englischsprachigen Tweets.

2. Automatisierte Kontextanreicherung:
• Geographische Dimension: Ableitung der Kontinente durch Geolocation-Mapping.
• Technologische Dimension: Extraktion des mobilen Betriebssystems (Android vs. Apple) aus Tweet-Metadaten.
• Zeitliche Dimension: Abbildung der zeitlichen Entwicklung durch Zeitreihenanalyse.

3. Sentimentbewertung: Einsatz der lexikonbasierten VADER-Sentimentanalyse (Hutto & Gilbert, 2014) zur Bewertung der Stimmungslage pro Tweet auf einer Skala von −1 (negativ) bis +1 (positiv).
• Nach Ausschluss neutraler Beiträge verblieben 245.077.312 stimmungstragende Tweets für die Analyse der Wirkung bestimmter Kontextfaktoren auf die in sozialen Medien zum Ausdruck gebrachte Stimmung.

Das entwickelte Verfahren erlaubt es, Kontextfaktoren innerhalb großer Mengen UGC datengetrieben und skalierbar zu identifizieren und mit der Sentimentanalyse zu kombinieren – und liefert damit die methodische Grundlage zur Beantwortung der aufgestellten Forschungsfrage.</td></tr>
<tr><td>Kernergebnisse (Überblick)</td><td colspan="2">Im Rahmen der Studie wurde ein Analyseverfahren zur automatisierten, kontextsensitives Identifikation von Sentimentunterschieden in UGC durch die Anwendung des durch Aggarwal et al. (2012) formulierten Text Mininig Framework</td></tr>
</table>

	entwickelt und erfolgreich angewandt. Das Verfahren integriert Schritte der Datenextraktion, des Preprocessings, der Kontextanreicherung und der lexikonbasierten Sentimentanalyse, um den Einfluss technischer (z. B. *Betriebssysteme*) und geographischer (z. B. *Kontinente*) Faktoren auf online-artikulierte Stimmungsbilder datenbasiert sichtbar zu machen. Die Anwendung des Verfahrens auf über 245 Millionen Tweets ergab: • Signifikante Sentimentunterschiede sowohl auf Ebene einzelner Kontextfaktoren als auch bei deren Kombination, z. B: o Android-Nutzer artikulieren durchschnittlich positivere Sentiments als Apple-Nutzer. o Tweets aus Nordamerika weisen im Vergleich die negativsten durchschnittlichen Sentimentwerte auf, während Tweets aus Asien und Südamerika deutlich positivere Tendenzen zeigen. • Die Kombination beider Kontextdimensionen verstärkt diese Unterschiede: o Android & Asien führt zu den positivsten, o Apple & Nordamerika zu den negativsten durchschnittlichen Sentiments. Das Verfahren ermöglicht somit eine skalierbare Identifikation von Kontextfaktoren und wurde anhand der Anwendung auf ~245 Mio. Tweets demonstriert. Darüber hinaus liefert die Anwendung praktische Erkenntnisse für eine zielgerichtete, kontextsensitive Markt- und Kommunikationsstrategie – z. B. durch Identifikation von Risikokombinationen für negatives *Word-of-Mouth* oder positive Resonanzräume für die strategische Kommunikation.
Publikationsort	Das Paper wurde 2021 auf der 54th Hawaii International Conference on System Sciences (*HICSS*) veröffentlicht und ist in der AISeL als peer-reviewtes Konferenzpapier gelistet.
Ranking VHB JQ 4	**B**
Autoren und Anteile	Daniel Konadl 60% **Janik Wörner 30%** Susanne Leist 10%
URI	http://hdl.handle.net/10125/70922

Tabelle 16. Fact Sheet zur 2. Veröffentlichung

Proceedings of the 54th Hawaii International Conference on System Sciences | 2021

Identifying Sentiment Influences Provoked by Context Factors – Results from a Data Analytics Procedure Performed on Tweets

Daniel Konadl
University of Regensburg
daniel.konadl@ur.de

Janik Wörner
University of Regensburg
janik.woerner@ur.de

Susanne Leist
University of Regensburg
susanne.leist@ur.de

Abstract

Context factors have lasting impacts on people's sentiments. Exploring impacts that different contexts have on sentiments can be crucial for managing the increasing number of communications companies nowadays maintain with customers via social media channels. To help companies prevent impacts of negative word of mouth, we provide an overview about sentiment-influential contexts for tweets as one kind of social media texts previously discussed within the literature. We collected an overall amount of 358.923.210 tweets and performed analysis to uncover the effects of continents, mobile devices' operating systems (OS) and the combination of both on sentiments expressed within tweets. Our results show remarkable differences for tweets originating from North America and Apple devices, which turned out to be the tweets with the lowest sentiments compared to the other continents and the mobile OS Android.

1. Motivation

Over the last decade, social media have reached an immense widespread [1, 2]. In private settings, social media are used for connecting with friends, for communicating with each other, presenting oneself to other users, sharing personal experiences or achieving social standing and reputation within communities [3]. Textual social media (e.g. tweets on Twitter or posts in Facebook) have proven to be a channel for complaint articulation for customers, e.g. about inadequate behaviors of companies resulting from negative experiences with products and services [4, 5]. Negative perceptions and experiences can prompt consumers to articulate negative sentiments online in written social media [4-7]. As social media texts are widely visible within social media channels, negative sentiments articulated in social media texts can potentially infect huge masses of social media users and provoke negative sentiments [5-7]. Additionally, social media texts containing negative sentiments tend to spread more quickly [8] so that companies have to react fast to prevent damages. Companies reacting adequately and quickly to customers' concerns can turn negative experiences into positive perceptions [5, 9, 10]. Otherwise, negative perceptions can provoke other users to contribute their own negative experiences [5]. This so-called negative word of mouth (nWoM) can lead to the far-reaching dissemination of negative perceptions towards a company in the online as well as in the offline sphere [5, 7]. Beyond keeping potential customers away from buying companies' products and consuming their services, nWoM can furthermore lead to the termination of existing customer relationships [11-13]. Eventually, churning customers and the absence of new customers lead to decreasing sales [11-13], high costs for acquiring new customers [14] and thereafter even to companies' existence being threatened.

For monitoring the opinions expressed in social media texts, sentiment analysis offers a solution to automatically identify opinion polarities from huge volumes of textual data (cf. [15]). The results of sentiment analysis can inform companies about negative sentiments expressed in social media channels, to make companies try to meet and overcome customers' concerns. While sentiment analysis enables companies to identify negative customers' perceptions identified from the texts' contents, companies require knowledge about the circumstances provoking certain sentiments. In addition to the contents of the texts captured by sentiment analysis tools, there are other factors that influence the sentiments of people.

For example, the days during a week have been shown to predominantly provoking negative sentiments compared to weekends (cf. [16-18]). Furthermore, places related to spare-time activities (e.g. parks and green spaces [19-21]) are associated with more positive sentiments than workplaces where sentiments are on average less positive (e.g. [22]). Companies that are aware of contexts as sentiment-influential factors can tailor their customer communication according to the effects of contexts. For strengthening the effects of positive messages propa-

URI: https://hdl.handle.net/10125/70922
978-0-9981331-4-0

gated, companies can also rely on contexts provoking positive sentiments. Companies aiming at positive Word of Mouth (pWoM) spread (cf. [4, 11, 23]) for evoking positive perceptions towards their products, services or the company itself, should rather communicate related messages e.g. on weekends.

Depending on the locations (e.g. country or continent), customers may be more or less likely expressing negative sentiments within social media posts. To counter the sentiment tendencies prevalent in certain locations (e.g. countries or continents), companies could thereafter adjust their customer communication. We address this research gap and perform analysis to identify the influences mobile devices' operating systems (OS) have on sentiments expressed within tweets as one kind of social media posts. As a second context, we include continents. To address the lack of investigations considering at least two contexts, we perform analysis combining mobile devices' OS and continents as contextual factors.

The rest of the paper is organized as follows: In the next chapter, conceptional basics of social media, word-of-mouth, sentiment analysis and contexts are introduced. Afterwards, within the chapter related work, we give an overview of sentiment-influential context factors and corresponding effects. Subsequently, we describe the steps performed within our investigation and report the results of our analysis. After interpreting and discussing the said results, we draw on implications of our findings. The paper is rounded off with a conclusion including limitations and an outlook on future research.

2. Conceptional basics

In literature, the term "social media" is often described as "*a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content (UGC)*" ([24], p. 61). UGCs represent "*the sum of all ways in which people make use of social media*" ([24], p. 61). Social media tools provide users with functionalities to connect with friends, presenting themselves to other users and communicating with each other [3]. Social media posts are suitable to conduct word-of-mouth propaganda and to lead communication that includes personal experiences and opinions about a product, service or promotion [4, 5] with consumers, friends, colleagues or other acquaintances [23, 25]. Within these communications, messages as well as therein expressed sentiments are spread [23, 25] and widely noticeable by other users within a social media channel [4-7]. To help companies take notice and control the sentiments expressed within the ever-increasing amount of social media posts, sentiment analysis proposes algorithmic approaches to identify the polarity of texts [15]. In terms of sentiment analysis, there are different approaches, amongst others dictionary- and sentence-based sentiment analysis [15]. When performing dictionary-based sentiment analysis, the sentiment of each entity (e.g. each word) from a text is classified into a positive or negative class using the dictionaries. These dictionaries annotate opinion carrying words, and the sentiment of the whole sentence is determined by considering the sum of the combined scores of all its entities [26].

However, not only the content or the formulation of messages or opinions of customers on products, services or the company itself, but also contexts play an important role in provoking sentiments. Depending on the manifestations of contexts, people experience different sentiments [27]. Context can be defined as "*any information that can be used to characterize the situation of an entity*" ([28], p. 5) whereby "*an entity is a person, place, or object that is considered relevant*" ([28], p. 5). Contexts such as time, location or mobile OS as well as combinations of contexts can be investigated as sentiment-influential factors in the offline world. While time describes the temporal contexts (e.g. time of day, day of week, month of year) in which a tweet is posted, location relates to the spatial properties (e.g. county, country, continent) the user is surrounded by, when tweeting. Mobile devices' OS (e.g. Google Android or Apple iOS) responsible for operating essential system functions on mobile devices, are associated with different personality traits [29] and can thus be seen as another sentiment-influential context. Additionally, contexts are as well experienced, while being in the situation of writing social media posts (e.g. tweets). Therefore, contexts also act as sentiment-influential factors within social media spheres (e.g. Twitter).

3. Related Work

The idea of observing contextual factors and their influence on sentiments apparent within tweets is not a new one. Efforts have been made to uncover the effects of temporal factors such as time of day [16, 18, 30] and day of week [16-18, 31] on sentiments within tweets. Tuesdays, Wednesdays and Thursdays have been identified as days characterized by negative emotions [18], with Wednesdays having even been identified as being the most negative days of the week [16]. Throughout the week and towards weekends, the sentiments become more positive, with Fridays and Saturdays being the most positive days [16, 17]. Sundays are associated with peaks in positive sentiment [18]. However, literature also reports declining sentiments on Sundays [26]. Regarding weekends, the results suggest that both positive and negative sentiments are more present as opposed to the working week [31]. Regarding the times of a day, there are contradictory findings within the literature. While [30] state that the time between 5 a.m. and 6 a.m. is the happiest hour of a day, [18] identified the most negative sentiments for this phase and the most positive sentiments in the evenings. [16] showed that there is a maximum of positive sentiments from 11 a.m. to 1 p.m. as well as from 5 p.m. to 8 p.m.

As for spatial influences, efforts have been made to analyze the effects of different kinds of locations and the impact of their properties on sentiments [16, 17, 19-21, 32-37]. Staying in green spaces such as parks has been identified as inducing positive sentiments within tweets [19-21]. In line with that, [16] identified that how a location is used, influences the sentiments expressed in tweets. Farmland and places associated with public transportation or industry bring up more negative than positive sentiments. Public places are almost equally likely to provoke positive and negative sentiments. Commercial areas tend to generate more positive than negative sentiments [16]. Places where time is shared with friends and family also induce predominantly positive sentiments [17]. The weather prevailing at a location is another sentiment-influential context [33-37]. People are happier when temperatures drop slowly rather than rapidly [35]. Extremely hot and cold temperatures [34], cloudy weather [34], high humidity [34], excessive snowfall [35], hailstorms [35], and extreme weather events such as hurricanes [33, 36] or earthquakes [37] have negative influences on the sentiment expressed within tweets.

There is further context information that rely on characteristics of individuals. People with high incomes articulate themselves more positively within tweets as people living in poorer neighborhoods [32, 38, 39]. Older people express more fear compared to middle-aged people whose tweets contain more joyful terms [38]. Neighborhoods with higher proportions of White, Asian and Hispanic populations also share predominantly more joy-related expressions [38]. However, as [39] showed, Hispanic residents can also be associated with less positive and sadder emotions. People with an African background tend to use more dimness-related terms in their tweets [38]. Higher degrees of education shape and higher-earning populations share happier and more positive emotions [39]. Tweets that are sent from mobile devices tend to be more negative in terms of expressed sentiments than those that are sent from desktops [40, 41]. As for mobile devices' OS, more positive posts are more likely to stem from blackberry devices [42].

Within the literature, there are already approaches relating to the effects of contexts on sentiments expressed within tweets. There are many investigations concerning the identification of temporal, spatial and person-related contexts as well as corresponding effects provoking positive and negative sentiments. Nevertheless, related work mostly focuses on one context solely or considers them isolated from each other. [16, 17, 19-21, 32-37] have dealt with spatial factors and focused on certain countries (e.g. USA (cf. [16, 17, 19, 21, 32-36]), Haiti (cf. [37]) or Australia (cf. [20])) instead of whole continents. [42] analyzed the influence of mobile OS on sentiments. However, the authors do not match these influences with continents (cf. [42]) as we did within our paper. To address this research gap, we performed analysis that combine mobile devices' OS and continents.

4. Methodology

To identify the effects of mobile devices' OS, continents and the combination of both on sentiments expressed within tweets, we aligned our approach to the steps proposed within the text-mining procedure of [43]. As we aim at identifying the effects of contexts on sentiments, we also describe how we proceeded this task. We split our approach (*cf. figure 1*) into **(1)** preparing the analysis, **(2)** conducting the analysis and **(3)** reporting the results.

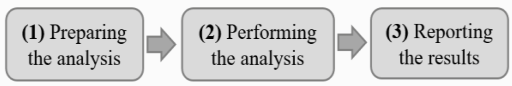


Figure 1: Steps applied within the approach

Within the **(1)** preparatory steps, we first familiarized with the particularities of tweets. Then, we extracted the tweets and performed exploratory analysis followed by conducting data reduction. Because mobile devices' OS and continents as the context factors to be investigated were not directly provided with the extracted tweets, we had to perform further steps to transform the provided data attribute values to obtain the said context factors. Then, we appended the continents and mobile devices' OS to the corresponding tweets and **(2)** determined sentiment values for the extracted tweets by applying an existing sentiment analysis approach. As with the continents and mobile devices' OS, sentiment values were appended to the corresponding tweets. Within the next step, summarization techniques are applied to the sentiment values for tweets of the investigated context factors. Hereby, the mean sentiment values for each context factor and the share of tweets at a certain sentiment level are determined. The last step **(3)** is about reporting the results. This includes the results to be presented and interpreted. The obtained results may contradict previous findings so that resolving these conflicts can be necessary. In the next chapters, we describe in more detail how we approached within these three steps.

5. Preparing the analysis

5.1. Particularities of tweets

As tweets are the research objects within our investigation, we firstly familiarized ourselves with the particularities of Twitter and tweets as the corresponding social media texts. Hereby, Twitter is a microblogging application "*stand[ing] halfway between traditional blogs and social networking sites, and are characterized by a high degree of self-presentation/self-disclosure and a medium to low degree of social presence/media richness*" ([44], p. 106). As a specific type of social media, it allows users to "*exchange small elements of content such as short sentences, individual images or video links*" ([44], p. 106). Tweets are amongst others characterized by their shortness (e.g. [31, 36, 42, 44, 45]). Since November 2017, the maximum number of characters to be used within a tweet is set to 280 (cf. [46, 47]). This shortness in text length must be taken into consideration when identifying sentiments expressed within tweets [36, 44]. Furthermore, our analysis has to cope with the huge number of users around the world that post tweets (cf. [31, 39, 42, 44, 48]). Therefore, we assume that the methods for processing the tweets need to be fast and performant and must deal with the shortness of tweets to obtain accurate and reliable results [45].

5.2. Data extraction, data preprocessing and data reduction

Using Twitter's sampling Application Programmable Interface (API) in the "*Spritzer*" version, we sampled approximately uniformly from all messages being posted via Twitter in 2019. The data collection procedure resulted in 358.923.210 English tweets in the time range from January 1st, 2019 to December 31st, 2019. For our investigation, we only collected tweets written in English. To that purpose, we filtered the provided language field to determine only tweets with the value "EN", which indicates that a tweet is written in English language. By this means, we were able to omit the complications of multiple languages (cf. [49]). Further restrictions beyond the language restriction as sampling by only incorporating tweets containing certain hashtags have not been applied.

Fluctuations in collective public emotions and sentiments can occur due to a multitude of competing effects (cf. [50]) and can influence the results of this investigation. By examining a whole year of tweets, seasonal influences and corresponding distortions resulting from deviating levels of sentiments expressed in certain months, such as higher temperature and more positive sentiments in summer months than in winter months, can be omitted. Therefore, because all twelve months of a year are included in our investigation, the influences of the incorporated contexts regarding sentiments are not distorted by a month being not included. As our data collection comprises 358.923.210 tweets, we assume that influences of external events and competing effects are smoothed to a high degree. We further regard the sample as a representative collection of tweets appropriate for our investigations because Twitter's "*Spritzer*" API provides 1 % of all tweets posted with a maximum margin of error of 0.06 at a confidence level of 99 % [51]. Each tweet delivered by Twitter comes as a JSON (Java Script Object Notation) object that contains tweet text and meta data characterizing both tweets and the situation in which it has been posted. These meta data, such as timestamp, language, source, geolocation and the device used for tweeting (cf. [19]), are logged and provided by Twitter.

Subsequently, we performed exploratory analysis to identify noise and outliers within the obtained data. By this means, we noticed that not every tweet is provided with location or device information. Thus, we agreed on performing analysis concerning influences of location and mobile devices' OS contexts only with tweets containing this context information. Filling up missing values by applying any of the commonly proposed strategies (cf. [52]) could have distorting effects as there were many data instances with missing attributes. To ensure the dataset to be without any redundant tweets, we additionally applied redundancy detection using tweet text and the creation date as the properties of a redundancy. The following steps of data reduction were carried out as part of the transformation of the JSON objects into CSV (Comma Separated Values) files. In this step, we excluded entries that don't match the goals of our investigation. Many of the provided meta data are not necessary for our analysis so that we only included the device field, the self-reported location field and the tweet text to identify sentiments expressed within the tweets. Only a small fraction of the tweets has geolocation coordinates that can be mapped directly to locations in terms of latitude and longitude (cf. [48]). Therefore, we relied on parsing the free-response location field that accompanies a tweet. By applying the methods from the Pycountry library (*https://pypi.org/project/pycountry/*), we mapped the self-reported location information to continents and appended the continent to the corresponding tweet within the CSV files.

5.3. Selection of methods to be applied

Identifying the influences of contexts on sentiments expressed within tweets requires a method to assign sentiment values to tweets. Considering the findings from *subchapter 5.1.*, we decided to apply the freely accessible "*Valence Aware Dictionary for sEntiment Reasoning*" (VADER) approach [45]. VADER was specifically developed for sentiment analysis in microblog-like texts and has achieved remarkable results compared to other sentiment analysis approaches [45]. The approach is also fast enough and can deal with huge numbers of tweets as required for our investigation. Because VADER identifies sentiment values using a built-in sentiment dictionary [45], there is no need for labeled training data. VADER provides a compound sentiment score that combines positive and negative sentiments into one single value. This score can take values in the range of -1 to +1 [45] and is calculated separately for each tweet contained in the data set. Furthermore, we applied summarization measures and decided to calculate mean values in terms of sentiments for all contexts being investigated. As the calculation of mean sentiment values consolidates several sentiment values to one single value, we additionally decided to determine the number of tweets whose sentiments are at certain intervals. This included counting the number of tweets that have certain sentiment values regarding the investigated context factors. In doing so, determining the sentiment intervals is independent from calculating the mean sentiment values.

To test the statistical significance of our findings, we additionally carried out t-tests (cf. [53]). In our case, there are independent samples since one sample selected from one population is not related in any way to the sample from another population [53]. This is because the assignment of tweets and the respective sentiment to a continent and a mobile devices' OS is exclusive and does not consider an assignment to more than one continent or mobile devices' OS. To be applicable, the tests require the dependent variable to be at least interval scaled. We see this requirement fulfilled as the dependent variable (sentiment value) of each tweet is numeric. Additionally, the independent variables (contextual factors) are at least nominal-scaled [53]. In the following chapters, every time the term "significant" or "significantly" is used, the differences in mean sentiment values or proportions within sentiment intervals showed to be significant by the pairwise calculated t-tests.

6. Performing the analysis

We observed a high proportion of neutral tweets by filtering for tweets with a sentiment value of "0" within the appended sentiment value field. As this huge number of neutral tweets can have remarkable influences on the results, we decided to exclude neutral tweets from further analysis. Therefore, we proceeded with the remaining 245.077.312 tweets being either positive or negative. Then, we filtered the tweets (*cf. table 1*) along with the corresponding sentiment values. For the analysis regarding one context, either continents or devices, we applied one filter criterion. E.g. by filtering the tweets with the value (a) "Europe" or (b) "Android", all tweets originating from (a) Europe *(cf. table 1 - IDs 1, 2 and 4)* or tweets sent from an Android OS *(cf. table 1 - IDs 1 and 4)* and their sentiment values are retrieved. When combining the contexts of continents and mobile devices' OS, we simultaneously set two filter criteria. We retrieve e.g. all tweets sent from Android powered devices originating from Europe together with the corresponding sentiment values by filtering with "Android" and "Europe" *(cf. table 1 - IDs 1 and 4)*.

Thereafter, the mean values of the sentiment values regarding the selected singular and combinatorial contexts were calculated and the numbers of occurrences of sentiment values in the respective sentiment intervals were counted. The results of determining the mean sentiment values and the sentiment intervals are more closely described in the next chapter.

Table 1: Example Tweets and Results

IDs	***Example Tweets***		
1	*I pray that your August will be full of good news, positivity and blessings.*		
2	*Hope you're having a great week.*		
3	*I wish you all the best :-)*		
4	*I have been on hold with you for 40 minutes and then the call hangs up. Poor customer service!!!*		
…	**…**		
IDs	***Results***		
	Continent	***Mobile OS***	***Sentiment***
1	Europe	Android	0.8481
2	Europe	Apple	0.7732
3	Asia	Apple	0.8481
4	Europe	Android	-0.4767
…	**…**	**…**	**…**

7. Reporting the results

7.1. Results of univariate analysis

Users of Apple devices turned out to be less positive compared to users of Android devices (*cf. table 2*). Tweets posted by Android devices have on average higher sentiments than tweets sent by Apple devices. Although the differences for mean sentiment values between Apple and Android users seem to be comparably low in our analysis, they showed to be significant. We additionally determined the number of tweets whose sentiments are at certain intervals (*cf. table 4*). Regarding the mobile devices' OS, the distributions reflect that 12.89 % of the tweets sent by Apple devices are strongly negative (within [-1; -0.66[), compared to 9.32 % for tweets sent by Android devices. As regards the strongly positive tweets (within [+0.66 to +1]), the proportion of Android tweets is about 5 % higher than the proportion of tweets sent by Apple devices.

In terms of location as the second context, we obtained results for six continents, namely *Africa (AF), Asia (AS), Europe (EU), North America (NA), Oceania (OC)* and *South America (SA)*. Tweets originating from AS showed to be the most positive, followed by tweets from SA, EU and AF (*cf. table 2*). Interestingly, the average sentiment values for tweets from OC and NA show significant differences. In addition, we notice that tweets in the range of +0.66 to +1.0 originate most frequently from SA, followed by AS, EU and OC (*cf. table 4*). Considering the mean sentiment values, it was assumable for NA having a low proportion of tweets within this interval. Based on the mean sentiment value of AF, it is remarkable that tweets are strongly positive (within [+1; +0.66]) with a proportion of only 23.19 %. However, tweets with mean sentiment values between +0.33 and +0.66 occur most often for AF, while the other continents have at least 7.85 % fewer proportions in this interval. When investigating the intervals for negative sentiments, it is significant that NA consistently has the highest proportions, followed by OC. Asia having the highest mean sentiment value, interestingly shows as well comparably high proportions in the strongly negative and the negative intervals (within [-1; -0.66[and [-0.66; -0.33[).

Table 2: Sentiments of singular contexts

Singular Contexts		*Mean Sentiment Values*
Continents	***Asia***	0.21452
	South America	0.19716
	Europe	0.19327
	Africa	0.18906
	Oceania	0.16195
	North America	0.12869
Mobile OS	***Android***	0.20179
	Apple	0.13223

7.2. Results of bivariate analysis

The observations of the univariate analysis (*cf. tables 2 and 4*) are partly reflected within the bivariate analysis (*cf. table 3 and 5*). Combining each continent with Android gives on average always more positive sentiment values as if the same continents are combined with Apple. Hereby, the highest difference can be observed for Asian tweets, where Android achieves on average significantly higher sentiment values compared to Apple (Δ 0.10366). South American and African tweets from Android devices are also significantly more positive. Differences observed for NA, EU and OC are however comparably marginal. Tweets posted from AF, EU and OC using Apple devices have consistently higher proportions within the negative sentiment intervals and consistently lower proportions for the positive sentiment intervals (*cf. table 5*). South American tweets having a sentiment value in the range of]0; +0.33[occur slightly more often for Apple devices than Android. In Asia, there are also comparably many positive tweets that are sent from Apple devices compared to Android. However, strongly positive tweets (within [+0.66; +1.0]) originate significantly more often from Android devices (Δ 9.33 %). The observations for North America are also remarkable, as one would expect to find a higher number of positive tweets from Apple devices because these devices originate from NA. However, NA does not show the expected higher sentiment values due to a possible connection of this continent to the brand Apple that is based there. But Android powered devices predominantly originating from AS provoke more positive sentiment scores for this continent.

Table 3: Sentiments of combined contexts

Combined Contexts		*Mean Sentiment Values*
Android	***Asia***	0.25694
	South America	0.22781
	Africa	0.21481
	Europe	0.19987
	Oceania	0.16436
	North America	0.13537
Apple	***Europe***	0.17091
	Asia	0.15328
	Africa	0.15323
	South America	0.14674
	Oceania	0.14323
	North America	0.10668

Table 4: Proportions of sentiment intervals for continents and mobile devices' OS

Singular Contexts	[-1; -0.66[	[-0.66; -0.33[	[-0.33; 0[	]0; 0.33[	[0.33; 0.66[	[0.66; 1]
Africa	8.46 %	13.05 %	8.94 %	10.02 %	36.34 %	23.19 %
Asia	9.32 %	14.36 %	9.65 %	11.15 %	28.11 %	27.41 %
North America	12.03 %	16.92 %	10.77 %	11.41 %	26.49 %	22.38 %
Europe	9.02 %	15.12 %	10.02 %	11.63 %	28.49 %	25.72 %
Oceania	10.31 %	15.93 %	10.61 %	11.66 %	27.74 %	23.75 %
South America	8.12 %	13.89 %	8.86 %	10.28 %	26.33 %	32.52 %
Android	9.32 %	14.55 %	10.01 %	10.89 %	27.57 %	27.66 %
Apple	12.89 %	17.21 %	10.89 %	11.19 %	25.80 %	22.02 %

Table 5: Proportions sentiment intervals for the combinations of continents and mobile devices' OS

Combined Contexts	*Mobile OS*											
	Apple						*Android*					
Continents	[-1; -0.66[	[-0.66; -0.33[	[-0.33; 0[	]0; 0.33[	[0.33; 0.66[	[0.66; 1]	[-1; -0.66[	[-0.66; -0.33[	[-0.33; 0[	]0; 0.33[	[0.33; 0.66[	[0.66; 1]
Africa	10.73 %	16.19 %	10.50 %	11.55 %	27.08 %	23.95 %	8.34 %	14.04 %	10.19 %	11.64 %	28.13 %	27.66 %
Asia	10.16 %	16.55 %	10.68 %	11.13 %	29.43 %	22.05 %	7.77 %	13.21 %	8.80 %	11.21 %	27.63 %	31.38 %
North America	12.96 %	19.38 %	11.13 %	10.81 %	25.94 %	19.78 %	10.71 %	17.27 %	10.87 %	11.01 %	26.40 %	23.74 %
Europe	9.42 %	16.44 %	10.35 %	11.42 %	28.25 %	24.12 %	7.98 %	15.66 %	10.30 %	11.64 %	27.84 %	26.58 %
Oceania	8.26 %	18.15 %	11.75 %	12.50 %	27.53 %	21.81 %	9.46 %	16.45 %	10.58 %	11.96 %	27.57 %	23.98 %
South America	10.26 %	16.82 %	10.76 %	12.35 %	28.90 %	20.91 %	7.79 %	14.29 %	9.37 %	11.34 %	29.14 %	28.07 %

7.3. Interpretation and implications

Actions, decisions and sentiments expressed within social media texts are influenced by the contexts in which people act. Certain manifestations of contexts thereby provoke more likely positive or negative sentiments. While reacting is only possible after certain conditions have already occurred, including contexts into decision-making enables companies to perform preventive actions. They can benefit by adapting the way they communicate within certain markets by aligning to the sentiments provoked by contexts. E.g. as customers from certain countries or continents may be more or less likely expressing negative sentiments within social media posts, companies could thereafter adjust their customer communication to counter sentiment tendencies prevalent there. Regarding the results of our analysis, companies should especially adapt messages for customers originating from NA, and those North Americans that send tweets from their Apple devices. Although the contexts NA and Apple provoke negative sentiments, the combination of NA and Apple devices turned out to provoke the most negative sentiments according to our results (*cf. tables 2 and 3*). Therefore, we suppose that there is the most potential for this combination when companies want to include contexts into decision-making within the activities and tasks of customer communication. By providing tailored contents and messages for these customers, the negative attitudes caused by the corresponding context factors of NA (continent) and Apple (mobile devices' OS) could be countered. Instead, when companies intentionally aim at provoking pWoM, they should better concentrate on AS and Android users.

Our findings retrieved for the mobile devices' OS are supported by the findings of [29] who report their results from a psychological investigation. Users of Apple devices are associated with more negative traits. They are perceived as less honest, less humble and are considered to manipulate others more often to gain personal advantages [29]. Hereby, the results of [42] are contradicting the findings of [29] and our results. In this work, tweets sent from Blackberry devices are associated with more positive sentiments compared to tweets sent from Apple or Android devices (cf. [42]). As the data set of [42] comprises the time range of May 1st, 2012 to April 30th, 2014, we assume that this data no longer reflect the current situation of mobile devices being used. The observed differences of Android users being more positive within our results could therefore be justified by Blackberry users that switched to Android devices. The market share of Blackberry has continuously dropped so that there are nowadays predominantly two major mobile OS, Android and Apple's iOS, that dominate the mobile devices market [54].

8. Conclusion

Our paper provides an overview about sentiment-influential contexts within tweets, which is followed by the identification of the influences of continents, mobile devices' OS and the combination of both. Our approach is structured into preparatory steps, the execution of the analysis and the reporting of the corresponding results. The results of our analysis have implications for the management of customer communications within social media channels because companies strongly build on social media to foster the external communication with customers (e.g. [55, 56]). NWoM expressed by disappointed or angry customers within social media channels (e.g. Twitter or Facebook) has the potential to negatively impact the perception of (potential) customers towards a company (cf. [5-7, 11-13]). As customers reporting negative experiences within Twitter await replies within one to three hours, companies have to react quickly [10] or even better take preventive actions. Hereby, the contexts inducing negative sentiments can support the corresponding decision-making. In addition, we see contributions of contexts for companies intentionally aiming at provoking pWoM. Companies that use Twitter for customer communication benefit most from our results and findings about sentiment-inducing contexts due to the required higher reaction speed [10]. Our results can support this task as our investigation is based on tweets as research objects. Companies apply Facebook as well as a customer communication channel where customers also articulate negative perceptions and experiences (cf. [5, 6, 10]). Therefore, we recommend as a possible step for future research to give an overview of sentiment-influential contexts and corresponding effects within Facebook posts.

The paper on hand is however not free of limitations. First, the analysis performed in our investigation was only performed with the VADER sentiment approach. Secondly, our results can be seen rather as initial findings derived from a data analytics procedure. Therefore, it is a necessary step in future research to deduce more sound recommendations from our results. Thirdly, the results of our work are based on the tweets of the year 2019. For this reason, we propose to apply the described analysis procedure to the data of previous years and to figure out observed similarities and differences regarding the influences of context factors.

9. References

[1] Dickey, I.J., and Lewis, W.F., "The evolution (revolution) of social media and social networking as a necessary topic in the marketing curriculum: a case for integrating social media into marketing classes", Advances in Marketing: Embracing Challenges and Change-A Global Perspective. 2010.

[2] Chaffey, D., "Global social media research summary 2016", Smart Insights: Social Media Marketing. 2016.

[3] Kietzmann, J.H., Hermkens, K., McCarthy, I.P., and Silvestre, B.S., "Social media? Get serious! Understanding the functional building blocks of social media", Business horizons. 54(3), 2011, pp. 241-251.

[4] Jansen, B.J., Zhang, M., Sobel, K., and Chowdury, A., "Twitter power: Tweets as electronic word of mouth", Journal of the American society for information science and technology. 60(11), 2009, pp. 2169-2188.

[5] Grégoire, Y., Salle, A., and Tripp, T.M., "Managing social media crises with your customers: The good, the bad, and the ugly", Business Horizons. 58(2), 2015, pp. 173-182.

[6] Einwiller, S.A., and Steilen, S., "Handling complaints on social network sites–An analysis of complaints and complaint responses on Facebook and Twitter pages of large US companies", Public Relations Review. 41(2), 2015, pp. 195-204.

[7] Pfeffer, J., Zorbach, T., and Carley, K.M., "Understanding online firestorms: Negative word-of-mouth dynamics in social media networks", Journal of Marketing Communications. 20(1-2), 2014, pp. 117-128.

[8] Naveed, N., Gottron, T., Kunegis, J., and Alhadi, A.C., "Bad news travel fast: A content-based analysis of interestingness on twitter", in Proceedings of the 3rd international web science conference, 2011.

[9] Joireman, J., Grégoire, Y., Devezer, B., and Tripp, T.M., "When do customers offer firms a “second chance” following a double deviation? The impact of inferred firm motives on customer revenge and reconciliation", Journal of Retailing. 89(3), 2013, pp. 315-337.

[10] Istanbulluoglu, D., "Complaint handling on social media: The impact of multiple response times on consumer satisfaction", Computers in Human Behavior. 74, 2017, pp. 72-82.

[11] Chevalier, J.A., and Mayzlin, D., "The effect of word of mouth on sales: Online book reviews", Journal of marketing research. 43(3), 2006, pp. 345-354.

[12] Mizerski, R.W., "An attribution explanation of the disproportionate influence of unfavorable information", Journal of Consumer Research. 9(3), 1982, pp. 301-310.

[13] Kim, S.J., Wang, R.J.-H., Maslowska, E., and Malthouse, E.C., "“Understanding a fury in your words”: The effects of posting and viewing electronic negative word-of-mouth on purchase behaviors", Computers in Human Behavior. 54, 2016, pp. 511-521.

[14] Greenberg, P., "CRM at the speed of light: Capturing and keeping customers in Internet real time": Elsevier. 2001.

[15] Liu, B., "Sentiment analysis and opinion mining": Morgan & Claypool Publishers. 2012.

[16] Cao, X., MacNaughton, P., Deng, Z., Yin, J., Zhang, X., and Allen, J.G., "Using twitter to better understand the spatiotemporal patterns of public sentiment: a case study in Massachusetts, USA", International journal of environmental research and public health. 15(2), 2018.

[17] Hu, T., She, B., Duan, L., Yue, H., and Clunis, J., "A Systematic Spatial and Temporal Sentiment Analysis on Geo-Tweets", IEEE Access. 8, 2019, pp. 8658-8667.

[18] Golder, S.A., and Macy, M.W., "Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures", Science. 333(6051), 2011, pp. 1878-1881.

[19] Plunz, R.A., Zhou, Y., Vintimilla, M.I.C., Mckeown, K., Yu, T., Uguccioni, L., and Sutto, M.P., "Twitter sentiment in New York City parks as measure of well-being", Landscape and urban planning. 189, 2019, pp. 235-246.

[20] Lim, K.H., Lee, K.E., Kendal, D., Rashidi, L., Naghizade, E., Winter, S., and Vasardani, M., "The grass is greener on the other side: Understanding the effects of green spaces on Twitter user sentiments", in Companion Proceedings of the The Web Conference 2018.

[21] Schwartz, A.J., Dodds, P.S., O'Neil-Dunne, J.P., Danforth, C.M., and Ricketts, T.H., "Visitors to urban greenspace have higher sentiment and lower negativity on Twitter", People and Nature. 1(4), 2019, pp. 476-485.

[22] Sandstrom, G.M., Lathia, N., Mascolo, C., and Rentfrow, P.J., "Putting mood in context: Using smartphones to examine how people feel in different locations", Journal of Research in Personality. 69, 2017, pp. 96-101.

[23] Kaplan, A.M., and Haenlein, M., "Two hearts in three-quarter time: How to waltz the social media/viral marketing dance", Business horizons. 54(3), 2011, pp. 253-263.

[24] Kaplan, A.M., and Haenlein, M., "Users of the world, unite! The challenges and opportunities of Social Media", Business horizons. 53(1), 2010, pp. 59-68.

[25] Hennig-Thurau, T., Wiertz, C., and Feldhaus, F., "Does Twitter matter? The impact of microblogging word of mouth on consumers’ adoption of new movies", Journal of the Academy of Marketing Science. 43(3), 2015, pp. 375-394.

[26] Kundi, F.M., Ahmad, S., Khan, A., and Asghar, M.Z., "Detection and scoring of internet slangs for sentiment analysis using SentiWordNet", Life Science Journal. 11(9), 2014, pp. 66-72.

[27] Mehrabian, A., " An approach to environmental psychology", 1974, Cambridge, MA: MIT Press, 1974.

[28] Dey, A.K., "Understanding and using context", Personal and ubiquitous computing. 5(1), 2001, pp. 4-7.

[29] Shaw, H., Ellis, D.A., Kendrick, L.-R., Ziegler, F., and Wiseman, R., "Predicting smartphone operating system from personality and individual differences",

Cyberpsychology, Behavior, and Social Networking. 19(12), 2016, pp. 727-732.

[30] Dodds, P.S., Harris, K.D., Kloumann, I.M., Bliss, C.A., and Danforth, C.M., "Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter", PloS one. 6(12), 2011.

[31] Wang, W., Hernandez, I., Newman, D.A., He, J., and Bian, J., "Twitter analysis: Studying US weekly trends in work stress and emotion", Applied Psychology. 65(2), 2016, pp. 355-378.

[32] Mitchell, L., Frank, M.R., Harris, K.D., Dodds, P.S., and Danforth, C.M., "The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place", PloS one. 8(5), 2013.

[33] Mandel, B., Culotta, A., Boulahanis, J., Stark, D., Lewis, B., and Rodrigue, J., "A demographic analysis of online sentiment during hurricane irene", in Proceedings of the second workshop on language in social media, 2012: Association for Computational Linguistics.

[34] Baylis, P., Obradovich, N., Kryvasheyeu, Y., Chen, H., Coviello, L., Moro, E., Cebrian, M., and Fowler, J.H., "Weather impacts expressed sentiment", PloS one. 13(4), 2018.

[35] Li, J., Wang, X., and Hovy, E., "What a nasty day: Exploring mood-weather relationship from twitter", in proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, 2014.

[36] Neppalli, V.K., Caragea, C., Squicciarini, A., Tapia, A., and Stehle, S., "Sentiment analysis during Hurricane Sandy in emergency response", International journal of disaster risk reduction. 21, 2017, pp. 213-222.

[37] Oh, O., Kwon, K.H., and Rao, H.R., "An Exploration of Social Media in Extreme Events: Rumor Theory and Twitter during the Haiti Earthquake 2010", in ICIS, 2010.

[38] Lin, Y.-R., "Assessing sentiment segregation in urban communities", in Proceedings of the 2014 international conference on Social Computing, 2014.

[39] Lerman, K., Arora, M., Gallegos, L., Kumaraguru, P., and Garcia, D., "Emotions, demographics and sociability in Twitter interactions", in Tenth International AAAI Conference on Web and Social Media, 2016.

[40] Murthy, D., Bowman, S., Gross, A.J., and McGarry, M., "Do we tweet differently from our mobile devices? A study of language differences on mobile and web-based Twitter platforms", Journal of Communication. 65(5), 2015, pp. 816-837.

[41] Groshek, J., and Cutino, C., "Meaner on mobile: Incivility and impoliteness in communicating contentious politics on sociotechnical networks", Social Media+ Society. 2(4), 2016.

[42] Cruz-Albrecht, L., Xu, J., Ni, K.-Y., and Lu, T.-C., "Characterizing Regional and Behavioral Device Variations Across the Twitter Timeline: A Longitudinal Study", in Proceedings of the 2017 ACM on Web Science Conference, 2017.

[43] Aggarwal, C.C., and Zhai, C., "Mining text data": Springer Science & Business Media. 2012.

[44] Kaplan, A.M., and Haenlein, M., "The early bird catches the news: Nine things you should know about micro-blogging", Business horizons. 54(2), 2011, pp. 105-113.

[45] Hutto, C.J., and Gilbert, E., "Vader: A parsimonious rule-based model for sentiment analysis of social media text", in Eighth international AAAI conference on weblogs and social media, 2014.

[46] Newton, C., "Twitter just doubled the character limit for tweets to 280", The Verge. 2017.

[47] Busby, M., "Twitter to introduce expanded 280-character tweets for all its users", The Guardian. 2017.

[48] Schwartz, H.A., Eichstaedt, J.C., Kern, M.L., Dziurzynski, L., Lucas, R.E., Agrawal, M., Park, G.J., Lakshmikanth, S.K., Jha, S., and Seligman, M.E., "Characterizing geographic variation in well-being using tweets", in Seventh International AAAI Conference on Weblogs and Social Media, 2013.

[49] Thelwall, M., Buckley, K., and Paltoglou, G., "Sentiment in Twitter events", Journal of the American Society for Information Science and Technology. 62(2), 2011, pp. 406-418.

[50] Lansdall-Welfare, T., Dzogang, F., and Cristianini, N., "Change-point analysis of the public mood in UK Twitter during the Brexit referendum", in 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), 2016: IEEE.

[51] Gerlitz, C., and Rieder, B., "Mining one percent of Twitter: Collections, baselines, sampling", M/C Journal. 16(2), 2013.

[52] Bramer, M., "Principles of data mining", Vol. 180: Springer. 2007.

[53] Sakai, T., "Two sample t-tests for ir evaluation: Student or welch?", in Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval, 2016.

[54] Statista. "Marktanteile der führenden mobilen Betriebssysteme an der Internetnutzung mit Mobiltelefonen weltweit von September 2009 bis Mai 2020", visited on 06.07.2020, available from: https://de.statista.com/statistik/daten/studie/184335/umfrage/marktanteil-der-mobilen-betriebssysteme-weltweit-seit-2009/.

[55] Stobbe, A., "Enterprise 2.0–Wie Unternehmen das Web 2.0 für sich nutzen. Deutsche Bank Research", Economics. 78, 2010, pp. 1-16.

[56] Heidemann, J., Klier, M., and Probst, F., "Online social networks: A survey of a global phenomenon", Computer networks. 56(18), 2012, pp. 3866-3878.

2.3 MANTRA: A Topic Modeling-Based Tool to Support Automated Trend Analysis on Unstructured Social Media Data

<table>
<tr><td rowspan="2">Adressierte Zielsetzungen</td><td>Identifikation und Systematisierung von Anforderungen an ein Topic Modelling basiertes Trendanalysetool für die drei zentralen marketingbezogenen Anwendungsfelder der Produktentwicklung, der Kundenverhaltensanalyse und des Markt-/Markenmonitorings.</td><td>ZS3</td></tr>
<tr><td>Konzeption, Entwicklung und Evaluation eines softwaregestützten Analysewerkzeugs („MANTRA“) auf Basis von semi-supervisiertem Topic Modelling, das die identifizierten Anforderungen (ZS3) integriert und automatisierte Trendanalysen auf UGC-Daten ermöglicht.</td><td>ZS4</td></tr>
<tr><td rowspan="2">Abgeleitete Forschungsfragen</td><td>Welche Design Anforderungen muss ein Topic Modelling basiertes Trendanalysewerkzeug erfüllen, das die marketingbezogenen Anwendungsszenarien der Produktentwicklung, der Kundenverhaltensanalyse und des Markt-/Markenmonitorings unterstützt?</td><td>FF4</td></tr>
<tr><td>Wie lässt sich ein solches Trendanalysewerkzeug technisch gestalten und implementieren, und welche konkreten wissenschaftlichen sowie praktischen Beiträge können daraus abgeleitet werden?</td><td>FF5</td></tr>
<tr><td>Genutzte Forschungsmethode</td><td colspan="2">Die Studie folgt einem Design Science Research (DSR) Ansatz zur theoriebasierten Entwicklung und Evaluation eines datengetriebenen Softwareartefakts für die automatisierte Trendanalyse auf Basis von UGC.

Konkret wurde der DSR-Prozess nach Peffers et al. (2007) umgesetzt, der basierend auf einer fundierten Problemstrukturierung und der initialen Konzeption eines Lösungsansatzes, die Schritte Entwicklung, Demonstration und Evaluation eines funktionsfähigen Trendanalysetools umfasst. MANTRA basiert auf semi-supervisiertem Topic Modelling und kombiniert verschiedene Social Media Analysetechniken (z. B. Sentimentanalyse und Kontextanalyse [z. B. zeitlich, geografisch]) in einem modularen System.

Die Forschung orientiert sich zudem explizit an der drei-Zyklen-Logik von Hevner et al. (2004):
Design Cycle: Konzeption, technische Umsetzung und Anwendung des Artefakts sowie iterative Verfeinerung anhand der Tool-Demonstration.
Relevance Cycle: Orientierung an praxisrelevante Anforderungen aus den marketingbezogenen Anwendungsfeldern Produktentwicklung, Kundenverhaltensanalyse und Marken-/Marktmonitoring.
Rigor Cycle: Fundierung des Artefakts auf Basis bestehender Theorien wie der Customer Focus Theorie, Innovationstheorie und Social Media Analytics Theorie sowie Beitrag zur Weiterentwicklung der Kernel bzw. Nascent Design Theorien im Sinne von Gregor und Hevner (2013).
Die Evaluation erfolgte durch die Anwendung des entwickelten Analysetools auf einen Datensatz mit ca. 1,03 Mio. Yelp Reviews, wobei exemplarisch zwei Trends (Vegan Cuisine, Global Cuisine) hinsichtlich ihres zeitlichen Verlaufs, ihrer Sentimententwicklung und ihrer geographischen Ausprägungen untersucht wurden.</td></tr>
</table>

Kernergebnisse (Überblick)	Durch die Konzeption und Entwicklung des Analysewerkzeugs wurden mit Blick auf die FF4 und FF5 mehreren Kernergebnisse geschaffen: • Ableitung von 14 Design Anforderungen für trendbasierte Analysesysteme aus Literatur und relevanten Kerneltheorien. • Entwicklung von 6 Design Prinzipien (z. B. Kontextsensitivität, Wiederholbarkeit, Visualisierung, Sentimentintegration) für den Aufbau eines kontextsensitiven, modularen Trendanalysetools. • Implementierung des Tools *MANTRA* und Anwendung auf ca. 1,03 Mio. Yelp Reviews zur Demonstration und Evaluation des entwickelten Artefakts. • Empirische Analyse zweier Trends (*Vegan Cuisine*, *Global Cuisine*) mit differenzierter Auswertung entlang des Zeitverlauf, der geographischen Verteilung (*Massachusetts*, *Texas*, *Oregon*) und des Sentimentverlaufs. • Aufzeigen, dass reine Häufigkeitstrends ohne Kontext (z. B. negativ konnotiertes Wachstum eines Trends) zu Fehlschlüssen führen können – und dass Sentiment und Kontextinformationen essenziell für die korrekte Interpretation von Trends sind. • Beitrag durch die Ableitung von Design Prinzipien im Kontext automatisierter Trendanalysen sowie zur Weiterentwicklung der Social Media Analytics Theorie durch die Demonstration der Design Prinzipien und deren dadurch geschaffenen implizierte Fundierung (Heinrich und Schwabe, 2014).
Publikationsort	Das Paper wurde 2023 auf der 44. International Conference on Information Systems, Hyderabad, Indien (*ICIS*) veröffentlicht und ist in der AISeL als peer-reviewtes Konferenzpapier gelistet.
Ranking VHB JQ 4	**A**
Autoren und Anteile	**Janik Wörner** **35%** Daniel Konadl 35% Isabel Schmid 20% Susanne Leist 10%
DOI	https://aisel.aisnet.org/icis2023/dab_sc/dab_sc/15

Tabelle 17. Fact Sheet zur 3. Veröffentlichung.

ICIS 2023 Proceedings

International Conference on Information Systems (ICIS)

December 2023

MANTRA: A Topic Modeling-Based Tool to Support Automated Trend Analysis on Unstructured Social Media Data

Janik Wörner
University of Regensburg, janik.woerner@ur.de

Daniel Konadl
University of Regensburg, daniel.konadl@wiwi.uni-regensburg.de

Isabel Maria Schmid
University of Regensburg, isabel.schmid@wiwi.uni-regensburg.de

Susanne Leist
University of Regensburg, susanne.leist@ur.de

MANTRA: A Topic Modeling-Based Tool to Support Automated Trend Analysis on Unstructured Social Media Data

Completed Research Paper

Janik Wörner
University of Regensburg
Universitätsstraße 31,
93053 Regensburg, Germany
Janik.Woerner@ur.de

Daniel Konadl
University of Regensburg
Universitätsstraße 31,
93053 Regensburg, Germany
Daniel.Konadl@ur.de

Isabel Schmid
Regensburg, Germany
Isabel.m.schmid@gmail.com

Susanne Leist
University of Regensburg
Universitätsstraße 31,
93053 Regensburg, Germany
Susanne.Leist@ur.de

Abstract

The early identification of new and auspicious ideas leads to competitive advantages for companies. Thereby, topic modeling can serve as an effective analytical approach for the automated investigation of trends from unstructured social media data. However, existing trend analysis tools do not meet the requirements regarding (a) Product Development, (b) Customer Behavior Analysis, and (c) Market-/Brand-Monitoring as reflected within extant literature. Thus, based on the requirements for each of these common marketing-related use cases, we derived design principles following design science research and instantiated the artifact "MANTRA" (MArketiNg TRend Analysis). We demonstrated MANTRA on a real-world data set (~1.03 million Yelp reviews) and hereby could confirm remarkable trends of vegan and global cuisine. In particular, the importance of meeting all specific requirements of the respective use cases and especially flexibly incorporating several external parameters into the trend analysis is exemplified.

Keywords: Social Media Analytics, Trend Analysis, Topic Modeling-Based Tool, Design Science, Marketing

Introduction

In the fast-paced and competitive world of business today, it is crucial for companies to keep track of trends (e.g., Endo et al. 2015; Lozano et al. 2017). Trends are complex, multidimensional phenomena that evolve over time and that can exert a significant impact on society in case they prevail (cf. Burmeister et al. 2002). In this context, social media plays a pivotal role as it allows trendsetters to openly discuss emerging trends across various industries (e.g., technology, fashion, or food trends) (e.g., Jeong et al. 2019; Park et al. 2021) and enables customers to express their unfiltered opinions (Yan et al. 2014). Thus, based thereon, companies can get an impression of what customers discuss as trends and what they think that future products and services should look like. In this respect, there are three focal points in the extant marketing research literature that deal with identifying and implementing trends: Developing a sound comprehension of feasible development paths of products and services (Puranam et al. 2017), customers' cultural norms and behaviors (Humphreys and Wang 2018), and market structures (Netzer et al. 2012). Our research considers these focal points by three different use cases and aims to support the early exposure of trends.

To identify trending topics from social media, marketing departments can apply different automated analysis techniques (cf. Batrinca and Treleaven 2015; Fan and Gordon 2014; Stieglitz et al. 2018). There are quantitative approaches that identify trends based on the frequency by which certain keywords appear within social media posts over time. However, this is insufficient, as trends, whose constituent elements are not necessarily known in advance, cannot be found in this way. Supervised learning could also be applied to trend analysis, but this requires ex-ante knowledge about trend characteristics and trend impulses. Unsupervised learning techniques can autonomously extract disruptive and previously unknown trends from large text volumes. Trending topics can thereby include combinations of two or more words that are not unusually topical or frequently mentioned together (Khan et al. 2021; Reisenbichler and Reutterer 2019). In addition, identifying meaningful trends can benefit from previously gained domain knowledge (cf. Burmeister et al. 2002). Therefore, semi-supervised topic modeling turns out to be a suitable automated trend analysis technique as it determines latent relationships between the individual components of a trend (Crain et al. 2012; Hong and Davison 2010) and allows its users to provide seed information as semantic anchors to guide the topics to be discovered (Jagarlamudi et al. 2012).

The potential that automated trend analysis techniques propose has already been recognized in practice and science. Nevertheless, available solutions show remarkable drawbacks, as they do not cover important Design Requirements (DRs) that are deemed crucial for a trend analysis concerning the common marketing-related use cases (e.g., incorporation of different geolocation data and customer characteristics as external parameters) that we could derive. Hence, the capabilities for comprehensively assessing trends from different perspectives (i.a., products, customers, markets) are restricted. Companies need to operate different trend analysis approaches to adequately support these common marketing-related use cases, leading, i.a., to the burden of tediously compiling individual results to gain a holistic view of trends. Moreover, the available tools proposed for automated trend analysis of marketing-related use cases do not offer marketing representatives to incorporate prior domain knowledge. As a result, the extracted trends are not necessarily semantically tailored to the company and its fields of activity, which severely impedes the integration of the gained insights into decision-making processes within marketing departments.

We address these drawbacks of existing solutions and propose the design and a concrete instantiation of a software tool as an automated artifact for trend analysis. By including a topic modeling technique that particularly enables the incorporation of previously gained domain knowledge, we show the proficiency of this technique in automated trend analysis. Regarding the DRs that a corresponding automated trend analysis tool needs to meet for common marketing-related use cases, we focus especially on the combination of different machine learning techniques, which leads to the trend analysis being purposeful. Against this background, the following two Research Questions (RQs) are posed:

- ***RQ1:*** *Which DRs should a topic modeling-based trend analysis tool meet that supports the marketing-related use cases Product Development, Customer Behavior Analysis, and Market-/Brand-Monitoring?*
- ***RQ2:*** *How can such a topic modeling-based trend analysis tool be designed and implemented, and which contributions can be derived?*

The remainder of this paper is structured as follows: the next section provides the conceptual foundations and the related work. Subsequently, the research procedure, following the Design Science (DS) approach (Hevner et al. 2004; Peffers et al. 2007), is described. The next section deals with the compilation of the DPs that we have derived from the DRs. Then, we report the technical realization of the tool, demonstrate its application on about 1.03 million Online Customer Reviews (OCRs), and present the resulting outcomes. The paper closes with concluding remarks, potential limitations, and future research ideas.

Conceptual Foundations and Related Work

Conceptual Foundations

Social media entails *"a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0 and that allow the creation of User Generated Content"* (Kaplan and Haenlein 2010, p.61). Online social platforms (e.g., Yelp) as a form of social media thereby serve as an important medium for the exchange between companies and customers by unstructured and opinionated content (e.g., textual social media posts). Trendsetters (e.g., domain luminaries or pioneers in certain fields) can share their thoughts on trending topics and discuss what future products and services should look like (e.g., Jeong et al. 2019; Park et al. 2021; Tucker and Kim 2011). In this respect, OCRs constitute *"peer-generated product evaluations posted on company or third party websites"* (Mudambi and Schuff, 2010, p.186), enabling users to evaluate companies' current products and services by disseminating their experiences. Thus, OCRs are also beneficial for companies as they include customers' expectations and opinions concerning products, services, or the company in general (cf. Hicks et al. 2012; Sigala 2012). As OCRs need to be online, social media platforms have shown effective for their dissemination in practice. Therefore, companies can harness OCRs to gain insights about central marketing-related dimensions unfiltered and in real-time (Yan et al. 2014): i.a., opinions about products and services (Ko et al. 2018; Moro et al. 2020), customers' behaviors (Humphreys and Wang 2018), and market structure and competition (Netzer et al. 2012; Tirunillai and Tellis 2014). Keeping track of the evolution of topics and these marketing-related dimensions is central to trend analysis. In this way, marketing departments can enhance the effectiveness of brand message placement and the allocation of appropriate resources to marketing campaigns depending on geographical and temporal developments (Hong et al. 2012; Lozano et al. 2017).

To reveal trending topics, marketing representatives can apply topic modeling, which refers to the use of generative probability models for determining latent relationships within a corpus of text data. Hereby, semi-supervised topic modeling allows the user to guide the topic discovery process by providing seed information as semantic anchors (Jagarlamudi et al. 2012). In this way, different contextual information (e.g., geolocation, time, user characteristics, a company's field of activity) can be incorporated to tailor the trends to be derived. Despite these potentials, semi-supervised topic modeling has yet not comprehensively been harnessed for trend analysis within marketing use cases.

Distinctive Use Cases of Topic Modeling-Based Trend Analysis

To identify common and distinctive marketing-related use cases of topic modeling-based trend analysis, we have comprehensively searched and consolidated the extant research literature. Thereby, we have aligned with the procedure for conducting a systematic literature review as proposed by Vom Brocke et al. (2015). In the **first step**, we applied a keyword-based search on *ACM Digital Library, AIS Electronic Library, EBSCOhost, Emerald Insight, Google Scholar, IEEE Xplore, SAGE Publications, SpringerLink, and the Wiley Online Library (keywords: "topic model", "topic modeling", "trend analysis", "trend tracking", "marketing", "social media", and several combinations thereof)*. Within the keyword-based search, we conducted iterative keyword refinements and could thereby derive the following further keywords: *"online customer review", "product development", "customer behavior", "customer behavior change", "market surveillance", "market monitoring", and "brand monitoring"*. We only kept papers dealing with topic modeling within the domain of marketing for automated trend analysis on unstructured social media data. Based thereon, the three involved researchers independently analyzed the title, abstract, and thematic focus of the initially derived *682 papers*. It turned out that there are papers that mention topic modeling but do not apply it to trend analysis. Some papers used social media and marketing-related terms, but these concepts were not reflected in the application of topic modeling *(remaining: 52)*. In the cases where assessments of the researchers showed ambiguities, the relevance and fit of the respective paper to our investigation were determined by clearing up the ambiguities through discussions and justifications among the involved researchers *(remaining: 38)*. Duplicates were eliminated, only papers written in English and only those that have passed a peer-review assessment were kept *(remaining: 26)*. In the **second step**, we conducted a forward and a backward search, resulting in another *3 relevant papers* that meet the defined inclusion criteria. Therefore, in total *29 papers* could be derived.

Subsequently, we have investigated the contents of the derived papers. Thereby, the wording of the statements in the examined sources often differed considerably, and the logical consolidation of the individual statements offered room for interpretation. Using qualitative content analysis and inductive category development, we could carry out an abstraction from the basic data, while, at the same time, deriving a coherent image of these data (cf. Mayring 2000; Mayring 2002). In this way, in particular, three distinctive marketing-related use cases could be identified for automated topic modeling-based trend analysis: (a) **Product Development** (e.g., Bae et al. 2018; Irawan et al. 2020; Jeong et al. 2019), (b) **Customer Behavior Analysis** (e.g., Bhor et al. 2018; Chen et al. 2012; Zhang et al. 2016) and (c) **Market-/Brand-Monitoring** (e.g., Lozano et al. 2017; Qu et al. 2015; Zhao et al. 2019).

Design Requirements for a Topic Modeling-Based Trend Analysis Tool

After having derived the distinctive marketing-related use cases, we next aimed to derive DRs for the respective use cases. Because different papers pointed to the same requirements but in (slightly) different wordings, we applied the qualitative content analysis (cf. Mayring 2000; Mayring 2002) here as well. In this way, we could systematically condense the available data and consolidate the contents that relate to the respective DRs. In total, we could derive 14 DRs that we will describe in more detail in the following:

By applying topic modeling for (a) **Product Development**, marketing representatives can develop an understanding of how customers perceive their products, services, and corresponding features. Topics identified from OCRs may, e.g., indicate that customers perceive the product packaging as unecological. Harnessing OCRs may bring up the idea of replacing the existing plastic packaging with paper packaging so that customers' expectations are subsequently met (cf. Park et al. 2021). As the prevailing literature shows for the use case (a), customers' product and service expectations can differ across geographical markets. To be successful and competitive in a targeted market, marketing representatives need to know which features to design in which way to meet the global and local customers' expectations. Therefore, topic modeling needs to **(DR1)** incorporate different geolocation data as an external parameter (e.g., Bae et al. 2018; Ding et al. 2020; Ha et al. 2017) that match the generated topics with the corresponding geographical locations. Customers' self-reported opinions of products and services also play an important role. Marketing representatives aim at retaining features evoking positive customer perceptions, while features evoking negative sentiments need to be improved. Thus, a trend analysis tool needs to **(DR2)** incorporate the sentiment of social media posts (e.g., Irawan et al. 2020; Jeong et al. 2019). Thereby, numerous opportunities for improving products and services can be determined, and certain features are given greater importance than others. Yet, it is not feasible to address all identified issues because of limited resources in marketing departments. Thus, it is necessary to support marketing representatives **(DR3)** to assess the product and service feature favorability (e.g., Ding et al. 2020; Moro et al. 2020; Tucker and Kim 2011). Regarding use case (a), e.g., Production Theory and Innovation Theory can serve as theoretical foundations. Innovation Theory proposes harnessing external and internal sources for identifying new and auspicious ideas of what future products may look like. These ideas are assessed and refined so that the most promising ideas are integrated as input factors into the product refinement and development process (cf. Axtell et al. 2000; Ye and Kankanhalli 2018). So, we argue that it is imperative for a trend analysis tool to support the refinement of the identified trends and to enable marketing representatives to **(DR4)** link the anticipated customer preferences with new features (e.g., Jeong et al. 2019; Ko et al. 2018; Tucker and Kim 2011).

In contrast to (a), the focus in (b) **Customer Behavior Analysis** is on the authors of the social media posts. Marketing representatives responsible for customer targeting must know how customer behaviors and associated trends differ for certain customer characteristics like age or gender (e.g., Bhor et al. 2018; Chen et al. 2012). Customer Behavior Analysis can for example reveal that older people perceive comfort as an essential product property. To increase favorability among older customers, a company may communicate its superiority regarding comfort to this respective target group (cf. Duncan and Moriarty 1998). For certain customer characteristics, there are as well different customer behaviors. As such, the needs for online shopping do distinguish between males (e.g., the accuracy of product descriptions) and females (e.g., fair pricing) (cf. Ulbrich et al. 2011). In this vein, Customer Focus Theory can be harnessed as it proposes companies to align their actions to customers' characteristics and expectations (Gulati and Oldroyd 2005; Tseng and Piller 2003). To establish the topic-person connections, topic modeling techniques, therefore, need to **(DR5)** incorporate different user-related information and customer characteristics (e.g., age, gender, group-related attitudes, and values) as external parameters (e.g., Bhor et al. 2018; Shi et al. 2018; Zhang et al. 2016). Based thereon, marketing representatives can tailor marketing messages with proposals that the individual customer groups appreciate most. As (b) is about customer behavior, **(DR6)** a functionality to filter posts is essential to exclude company posts before generating customer-related topics (Iwata et al. 2009; Saha and Sindhwani 2012). Fast reactions **(DR7)** are important regarding topic modeling techniques for this use case because customers' perceptions can change rapidly (Bhor et al. 2018; Iwata et al. 2009; Sasaki et al. 2014) and, while, identified concerns about products or services require a certain time for refactoring, customers need to be rapidly contacted and appeased.

The third use case (c) **Market-/Brand-Monitoring** is about investigating how brands and companies act and which strategies they deploy. Therefore, (c) is concerned with the companies and brands instead of (a), their products and services. Applying trend analysis to the social media posts of competitors could reveal that competing brands include reporting about their sustainable acting within public communications. To support decisions on the market-/brand level, the topics must be generated for the respective companies or brands. Beyond generating topics according to geographical markets, the topic modeling technique **(DR8)** must also be capable of incorporating brands as external parameters (e.g., Endo et al. 2015; Lozano et al. 2017; Qu et al. 2015). Based thereon, a company may, e.g., adapt its market strategy to focus on niche markets that other brands have yet paid little attention to. Innovation Theory (cf. Sundbo 1995) thereby serves as a foundation as it recommends how to successfully place new innovative products and services within a market. Hereby, the Strategic Innovation Paradigm points out that companies need to assess how competitors shape their strategic communications concerning innovations. Thus, a trend analysis tool **(DR9)** must support the identification of competitors and their market orientations (e.g., Lozano et al. 2017; Park et al. 2021; Zhao et al. 2019). Moreover, marketers need not only support in identifying emerging competitors but also in understanding the behaviors of current ones (e.g., communication patterns and discount strategies). Therefore, **(DR10)** filter criteria like time periods, geolocations, and brands should be proposed to narrow down the dataset (e.g., Sohn et al. 2019; Valdez et al. 2018; Zhao et al. 2019).

Beyond use case-specific DRs, we could identify DRs that apply to all three use cases. One theory posing fundamental implications across all use cases is Social Media Analytics Theory (e.g., Fan and Gordon 2014; Stieglitz et al. 2018). When applying social media analytics for trend analysis on unstructured social media data, a central necessity is **(DR11)** dealing with a huge number of social media posts because of the ever-increasing speed of social media posts emergence (e.g., Jeong et al. 2019; Lozano et al. 2017; Stieglitz et al. 2018). Further, it is essential to provide means to include **(DR12)** historical data so that past developments of trends can be considered when assessing the current state of a trending topic (e.g., Bhor et al. 2018; Lozano et al. 2017). Because trends constitute dynamic developments and not solely static points in time, the topic modeling technique needs to integrate different temporal parameters into the topic generation. Marketers require **(DR13)** support in assessing the course of topics over time at different levels of granularity and in identifying emerging and declining trends (e.g., Ha et al. 2017; Zhong and Schweidel 2020). In general, applying the results of topic modeling requires advanced technique-related knowledge. Thus, **(DR14)** the results need to be visualized (by charts and plots) so that marketing representatives are supported in making sense of the derived topics (Bae et al. 2018; Ha et al. 2017; Yang et al. 2017).

Use Cases	Design Requirements	Sources
(a) Product Develop-ment	**DR1.** Incorporation of different geolocation data as external parameters	Bae et al. 2018; Ding et al. 2020; Endo et al. 2015; Ha et al. 2017; Moro et al. 2020; Park et al. 2021
	DR2. Consideration of sentiment	Irawan et al. 2020; Jeong et al. 2019; Ko et al. 2018; Moro et al. 2020
	DR3. Providing means to assess the importance of a product or service feature	Bai et al. 2023; Ding et al. 2020; Jeong et al. 2019; Mastrogiacomo et al. 2021; Moro et al. 2020; Tucker and Kim 2011
	DR4. Linking of newly discovered preferences with new features	Axtell et al. 2000; Bai et al. 2023; Ding et al. 2020; Jeong et al. 2019; Ko et al. 2018; Puranam et al. 2017; Tucker and Kim 2011; Ye and Kankanhalli 2018
(b) Customer Behavior Analysis	**DR5.** Incorporation of different user-related information and customer characteristics as external parameters	Bhor et al. 2018; Chen et al. 2012; Gulati and Oldroyd 2005; Puranam et al. 2017; Qu et al. 2015; Shi et al. 2018; Tseng and Piller 2003; Zhang et al. 2016
	DR6. Provide means to filter out social media posts being posted by the company itself	Gulati and Oldroyd 2005; Iwata et al. 2009; Saha and Sindhwani 2012; Tseng and Piller 2003
	DR7. Support fast reactions to changing customer perceptions	Bhor et al. 2018; Iwata et al. 2009; Sasaki et al. 2014
(c) Market-/ Brand-Monitoring	**DR8.** Incorporation of different brands and geolocation data as external parameters	Endo et al. 2015; Lozano et al. 2017; Qu et al. 2015; Sohn et al. 2019; Zhao et al. 2019
	DR9. Identify differences of how enterprises communicate and place themselves in markets	Park et al. 2021; Sundbo 1995; Tirunillai and Tellis 2014; Valdez et al. 2018; Zhao et al. 2019
	DR10. Apply time periods, geolocations, and brands as filter criteria	Sohn et al. 2019; Sundbo 1995; Valdez et al. 2018; Zhao et al. 2019
Use Case Independent Design Requirements	**DR11.** Deal with large numbers of social media posts	Fan and Gordon 2014; Jeong et al. 2019; Lozano et al. 2017; Puranam et al. 2017; Stieglitz et al. 2018; Yan et al. 2015
	DR12. Capability of analyzing historical data	Bai et al. 2023; Bhor et al. 2018; Ha et al. 2017; Lozano et al. 2017; Puranam et al. 2017; Sohn et al. 2019; Zhong and Schweidel 2020
	DR13. Identification of trends over time (topic evolution) at different levels of granularity (e.g., phases of a day, days of a week, seasons)	Ha et al. 2017; Lozano et al. 2017; Puranam et al. 2017; Tucker and Kim 2011; Zhang et al. 2015; Zhong and Schweidel 2020
	DR14. Visualization of the results	Bae et al. 2018; Ha et al. 2017; Yang et al. 2017

Table 1. Identified Use Cases and Corresponding Design Requirements

Assessment of Tools for Automated Trend Analysis on Social Media

To assess the capabilities of automated trend analysis tools for unstructured social media data, we first identified corresponding tool providers. The Google search engine and social media analytics groups within online social networks (i.a., LinkedIn) were drawn upon for this purpose. Tools identifying trends based on ex-ante knowledge (i.a., approaches that assess trends by the frequency by which certain keywords appear) were disregarded, as this tremendously hinders the extraction of disruptive and previously unknown trends.

By three researchers conducting the identification of tool providers independently, the possibility of overlooking established tool providers could be reduced. The identified tool providers were then consolidated and potential disagreements about the tools were resolved. In this way, we could cover a broad spectrum of providers and take an up-close look at the most established tools for automated trend analysis *(Brandwatch, Hootsuite, IBM, Keyhole, Meltwater, MineMyText, NetBase Quid, Pulsar, Q-Markets, Symanto, Synthesio, Trend-Sonar).* Based on the DRs that we could derive *(see Table 1)*, a tool survey was designed to assess whether the available trend analysis tools cover the posed requirements. The tool survey was validated by the involved researchers and subsequently applied to the data collection procedure. For this reason, we interviewed sales representatives from these tool providers on the specific features and additionally installed demo versions of the tools. Based on the interview results and by independently testing the functionalities of the trend analysis tools against the posed DRs, we could gain detailed and comprehensive insights into the capabilities as well as into the drawbacks of those trend analysis tools.

Solution Prov.	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
# Employees	~ 10	~ 30	~ 60	~ 70	~ 70	~ 120	~ 300	~ 1.3k	~ 1.4k	~ 2.2k	~2.5k	~ 290k
General Approach	NLP	NLP	AI	AI	NLP	NLP	NLP	AI	AI	NLP	NLP	AI
DR1	-	-	○	-	-	-	-	-	-	-	-	○
DR2	-	●	-	●	-	-	●	●	●	-	●	●
DR3	●	●	○	●	●	●	●	●	●	●	●	●
DR4	-	○	-	-	●	○	○	-	○	-	-	-
DR5	-	-	-	-	-	-	-	-	-	-	-	-
DR6	●	○	●	●	-	●	-	●	●	●	●	●
DR7	*DR7 was **not** included in the comparison because a fast execution is highly dependent on the underlying operating hardware.*											
DR8	-	-	○	-	-	-	-	-	-	-	-	-
DR9	-	-	-	○	○	○	●	-	-	-	-	●
DR10	●	●	●	●	●	●	●	●	●	●	●	●
DR11	●	●	●	●	●	●	●	●	●	●	●	●
DR12	●	○	●	●	●	●	●	●	●	●	●	●
DR13	○	○	○	○	○	○	○	○	○	○	○	○
DR14	●	●	●	●	●	●	●	●	●	●	●	●

Table 2. Results of Assessing Solutions for Automated Trend Analysis
AI: Artificial Intelligence (i.a., neural networks, support vector machines), **NLP:** Natural Language Processing (i.a., unsuperv. topic modeling)

● = applicable
○ = partly applicable

As our assessment could reveal, there is no automated tool that offers marketing representatives to unveil the latent relationships between the individual components of a trend and that enables them at the same time to incorporate their previously gained domain knowledge. The tools lack to meet various specific DRs of the marketing-related use cases (a)-(c), none of the tools thereby manages to cover all of these DRs, and, in particular, they cannot flexibly integrate the various external parameters at different granularity. Nonetheless, the perceptions of products and services differ across continents, countries, and federal states because of globally and locally differing customer expectations. Customer behaviors do differ for certain customer characteristics (e.g., age, gender, group-related attitudes, and values). Without the ability to flexibly integrate different external parameters like time (e.g., days of a week, phases of a day, seasons), geolocations, and user-related information at different granularity, an automated trend analysis tool does not sufficiently support trend analysis and the marketing-related use cases (a)-(c). Beyond the revealed drawbacks of the available tools, extant literature clearly points out the need of such a comprehensive topic modeling-based tool that supports all three use cases. With this research, we aim to close this gap.

Research Procedure

To accomplish the development of a tool for automated trend analysis regarding the common marketing-related use cases that we could derive from the extant research literature, we applied Design Science (DS) research. A widely recognized suggestion on how to conduct DS projects was introduced by March and Smith (1995) and Peffers et al. (2007). In this respect, DSR represents a synthesis of the activities "build/development" and "justify/evaluate" with the main goal of designing an IT artifact to address an organizational problem *(see Figure 1)* (cf. Gregor and Hevner 2013; Hevner et al. 2004; Peffers et al. 2007).

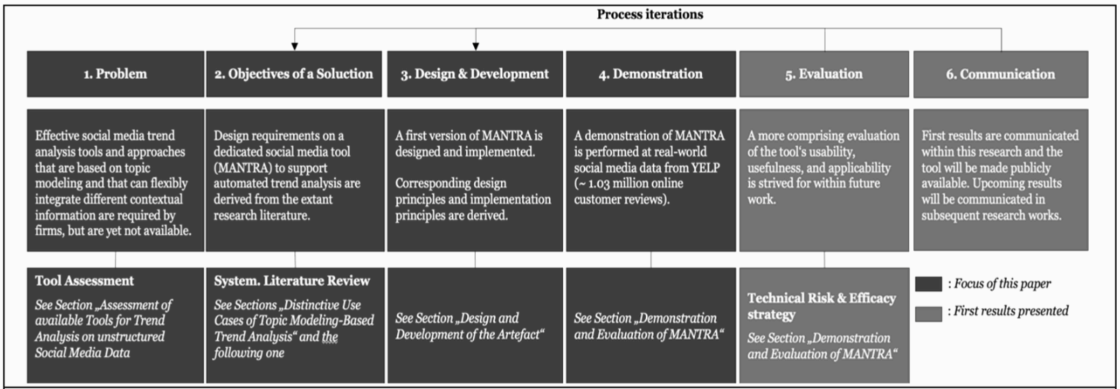


Figure 1. Procedure of the Research, own Representation Based on Peffers et al. (2007)

As a first step, **(1) corresponding problems** of existing software tools for trend analysis were identified *(see Section "Introduction" and Table 2)*. Automated tools for trend analysis neither do meet the requirements that are indispensable to the successful application of trend analysis nor do provide flexibly incorporating external parameters (e.g., by means of semi-supervised topic modeling). Thus, our **(2) objective** is to provide and combine a set of machine learning techniques, based on the DRs, to facilitate trend analysis regarding common marketing-related use cases based on semi-supervised topic modeling.

The third step contains the **(3) design and development** *(see Section "Design and Development of the Artifact")* of the solution. By conducting a systematic literature review (cf. Vom Brocke et al. 2015), we could identify DRs an automated trend analysis tool needs to meet to be comprehensive for common marketing-related use cases *(see Section "Distinctive Use Cases of Topic Modeling-Based Trend Analysis")*. To fill the gaps identified within step (1), we focus on the design of the technical realization of the tool by establishing DPs that are founded on the identified DRs. Thus, our approach was established to support a comprehensive trend analysis for marketing-related use cases and to eliminate the drawbacks of existing solutions.

Steps **(4)** and **(5)** deal with the **demonstration** and **evaluation** of the implemented artifact. To define a proper setting for rigorously demonstrating and evaluating the utility of the artifact, we related to Venable et al. (2016) who have defined four evaluation strategies for DSR projects. We aligned our DSR-project-specific evaluation strategy to the "Technical Risk & Efficacy strategy" because the focal aim of our investigation was to rigorously demonstrate that the evaluand "MANTRA" (MArketiNg TRend Analysis) achieves its purpose (i.e., identify meaningful and sound trends from unstructured social media data) and that the incorporation of semi-supervised topic modeling and meeting several DRs are thereby imperative (cf. March and Smith 1995; Peffers et al. 2012). Thus, we decided to apply one single summative evaluation episode (cf. Venable et al. 2016) in which trends that have already occurred are reproduced retrospectively. Thereby, the results produced by MANTRA are verified and confirmed with two trend reports. Beyond, we plan to investigate the tool within a controlled environment through a laboratory experiment where participants (i.a., experts from the food industry) identify the soundness of the trends and provide feedback on how to improve the tool before we will conduct a more extensive evaluation in a natural setting (e.g., within a field study). In step **(6)**, we aim to **document** and **communicate** the results of our research.

With this procedure, we align our research also with the guideline of Hevner et al. (2004). Regarding the design cycle, we present our artifact as the result that has gone through the process of demonstration (application to a Yelp dataset) and evaluation by a survey of an expert panel from the National Restaurant Association Board (NRAB) *(see Section "Demonstration and Evaluation of MANTRA")*. Regarding the relevance cycle, we identified several DRs from research literature and theory *(see Table 1)* that guided the design of the artifact, and so the practical application of our artifact brought up several contributions for practice *(see Section "Discussion and Contributions")*. In view of the rigor cycle, we used several techniques (i.a., semi-supervised topic modeling, sentiment analysis) to rigorously construct and evaluate our artifact. We have derived initial findings as contributions to theory, both kernel theory (e.g., Innovation Theory, Social Media Analytics Theory) and design theory. To contribute to a rather general and abstract knowledge base – "Nascent design theory" (Gregor and Hevner 2013) – and to design a purposeful artifact in a comprehensible way, we deduced Meta Requirements (MRs) and DRs (Gregor and Jones 2007; Heinrich and Schwabe 2014) for a trend analysis tool grounded in seminal works, resulting in the DPs.

Design and Development of the Artifact

Design Principles for a Trend Analysis Tool

The composition of MRs that describe "*what the system is for*" (Gregor and Jones 2007, p.325) is based on the purpose and scope of the trend analysis tool as discussed in the introduction. Thus, we define the solution objectives based on the problems of the investigation that our paper addresses. These MRs to be suitable for a class of artifacts are based on the current research literature (Gregor and Jones 2007; Heinrich and Schwabe 2014; Walls et al. 1992). Subsequently, we have synthesized DPs based on the DRs we have derived from the related work (i.e., extant literature) as well as based on DRs that are inspired by theory (cf. Möller et al. 2020; Purao et al. 2020). These DPs are defined as prescriptive statements, showing how to do something to achieve a certain goal (Gregor et al. 2020). The DPs we dispose fall into the category of action and materiality-oriented DPs that describe what an artifact should enable users to do and how the artifact should be built in order to do so (Chandra et al. 2015). For companies (= users) who are interested in conducting trend analysis based on unstructured social media data (= boundary conditions), and keeping our DRs for our artifact in mind, we derive six DPs for trend analysis tools:

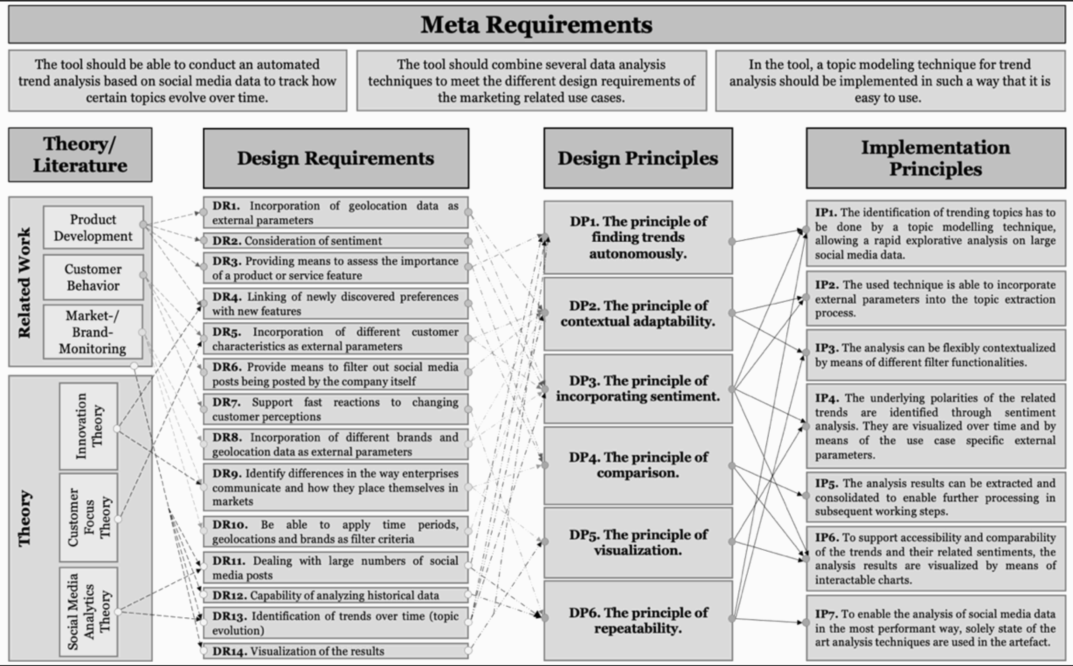


Figure 2. Design of an Automated Trend Analysis Tool for Unstructured Social Media Data

- **DP1: The principle of finding trends autonomously.** To track customers' evolving and changing requirements, it is necessary to retrace the evolution of topics. The tool should be able to find trends in large amounts of unstructured social media data and provide its users to apply autonomous trend analysis by including domain knowledge (i.a., seed words) to refine the identification of trending topics.
- **DP2: The principle of contextual adaptability.** Since external parameters (e.g., geolocations or customer characteristics) directly influence how a trend is pronounced, they must be included in the trend analysis tool. Therefore, the tool needs to provide the possibility to flexibly incorporate different external parameters and filters, so that the users can tailor the trend identification to their peculiarities.
- **DP3: The principle of incorporating sentiment.** To meet customers' expectations, it is necessary to identify issues evoking positive customer perceptions to retain them and to identify issues evoking negative customer perceptions to be improved. Thus, the tool needs to assign polarities to the individual trends and differentiate them over time and for the corresponding use case-specific parameters.
- **DP4: The principle of comparison.** To be competitive in the future, comparisons are necessary to identify the extent to which existing products and services meet customers' expectations (internal

comparison) and to compare a company's actions with those of competitors (external comparison). Thus, the tool should provide the retrieved analysis results to enable both kinds of comparisons.
- **DP5: The principle of visualization.** To obtain added value from the results and to benefit from the derived trends, they must be refined approachable. In doing so, the tool should be able to visualize the trends and sentiment values so that users are aided in interpreting and applying the derived results.
- **DP6: The principle of repeatability.** As a trend can change over time, the identification process should be executable often, and in a resource-saving way. Therefore, the tool should allow repetition at any time (including both historical and current data) to react quickly to changing circumstances.

These DPs are deduced from the DRs that are based on current research literature and theory. Gregor and Jones (2007) state that this reference to theory and literature discloses "*an explanation of why an artifact is constructed as it is and why it works*" (p.328). Based on related work and theory, we derive DRs our tool should meet. These DRs offer guidance by designing the artifact and advising the DPs (Böckle et al. 2021; Gregor and Jones 2007). These DPs refer to at least one DR and serve as an abstract "blueprint" of our artifact (Böckle et al. 2021; Gregor and Jones 2007; Heinrich and Schwabe 2014). By establishing these DPs, we made sure that they follow the value grounding (reference to the DRs) and the explanatory grounding (DPs are based on current literature and kernel theories) (Heinrich and Schwabe 2014). Furthermore, based on the instantiation of the DRs and DPs within our artifact MANTRA, we have defined Implementation Principles (IPs). In this way, we support "*the implementation in practice of an abstract, generic design method or development approach*" (Gregor and Jones 2007, p.329) in view of such an artifact for automated (topic modeling-based) trend analysis on unstructured social media data.

Technical Realization

To address the drawbacks of prevailing research and existing trend analysis tools, we have designed and developed the artifact *MANTRA* in the programming language *Python*. In the following, the instantiation of the DRs is described as they depict the required features of an automated topic modeling-based trend analysis tool in the most detailed way. Since the DPs and IPs depict a generic and prescriptive statement of how something should be done, they capture design-related knowledge and support the development of further IS (design) theories and new artifacts focusing on automated trend analysis. To reveal the association of the derived DPs/IPs to the DRs, their relations are pointed out as well.

Regarding the autonomous identification of trends **(DP1)**, the selection of the topic modeling technique to be applied is critical to success. As trends are time-dependent constructs, a rapid and explorative analysis must be provided. With this respect, unsupervised topic modeling techniques (e.g., LDA) are conceivable to enable an explorative analysis. However, given **DP2**, the incorporation of external parameters **(DR1, DR5, DR8)** is not feasible using unsupervised topic modeling. Beyond that, as stated by Chang et al. (2009) the potential of unsupervised topic modeling is stymied by their purely unsupervised nature, which often leads to topics that are neither meaningful nor effective at extrinsic tasks such as conducting a marketing campaign. Semi-supervised topic modeling techniques retain the flexibility of unsupervised techniques and can identify trends in an autonomous and performant way (Jagarlamudi et al. 2012). In doing so, they can incorporate external parameters such as geolocation information **(DR1, DR8)**, brands or products, and user-related external information **(DR5)**. As a semi-supervised topic modeling technique, *GuidedLDA* was chosen as it has achieved convincing analysis results within previous research (e.g., Ahadh et al. 2021; Chandrasekaran et al. 2020; Toubia et al. 2019). By providing seed information, such as a specific product/brand name or words related to an area such as the global food industry (e.g., burger sauce or burger patty), the topics generated by *GuidedLDA* will converge to gravitate into the contextual direction of those seed words, fostering a successful integration of identified trends into marketing decision-making.

Concerning use case (a) **Product Development**, the identification of trends and their matching with corresponding geolocations are required. Besides the incorporation of geolocation information through respective seed words, we, therefore, enabled MANTRA to pre-define different geolocations by applying a filter function, gaining maximum flexibility regarding the incorporation of external geolocation information **(DR1, DP2, IP3)**. This use case stipulates further the inclusion of sentiments within the analysis to determine the positive or negative tonality of the social media posts **(DR2, DP3, IP4)**. Therefore, we related to the "Valence Aware Dictionary for sEntiment Reasoning" (VADER) (Hutto and Gilbert 2014) technique. *VADER* is a lexicon- and rule-based sentiment analysis technique that is specifically attuned to sentiments expressed in social media and has achieved remarkable results (Hutto and Gilbert 2014). It can extract positive, neutral, and negative inflections in the range of -1 to +1. Additionally, the importance of a respective product or service feature has to be taken into account **(DR3, DP1, IP1)**. As probabilistic topic modeling techniques such as *GuidedLDA* infer their resulting topics based on various probabilistic distributions, depicting the relevance of the underlying topic words and thus the resulting topics (Crain et al. 2012), the assessment of the importance of the specific topic is met through the nature of topic modeling itself. Moreover, the linking of newly discovered preferences with new features **(DR4, DP4, IP5)** is compulsory. To develop an artifact that is suitable for various applications in marketing, including the processing of heterogeneous datasets, an automated linking of newly discovered preferences will not be feasible. Nevertheless, to take **DR4** into account, a function was implemented to extract the respective results into a Microsoft Excel file. In this way, the results are consolidated and can be used to manually map the identified preferences to new product or service features. Use case (b) **Customer Behavior Analysis** deals, i.a., with identifying behavior changes regarding different kinds of customers. Besides considering customer-related information by biasing the generated topics by means of seed words, users can pre-define customer groups based on attributes such as gender or age **(DR5, DP2, IP2)**. Further, a function to filter out posts created by the company itself to inhibit a bias in the analysis results **(DR6, DP2, IP3)** is required. Thus, MANTRA provides the possibility to dynamically define an identifier that is used to filter out associated data entries. **DR7 (DP6, IP7)** reflects a fast execution time when analyzing data concerning customer behavior. Here, MANTRA implements multi-threading, resulting in the parallel processing of independent tasks. Besides incorporating different brands and geolocation data with respective seed words **(DR8, DP2, IP2)**, (c) **Market-/Brand-Monitoring** requires identifying differences in the way companies communicate and place themselves within markets **(DR9, DP4, IP6)** by applying filter criteria such as time periods, geolocations, or specific brands **(DR10, DP2, IP3)**. Therefore, a filter function was implemented, handling the generation of different subsets of data concerning a specific time period or brand names. This allows comparisons based on diverse dimensions such as topics of interest or specific products and services. To enable the processing of large amounts of unstructured data **(DR11, DP6, IP7)**, only performant analysis techniques settled in the field of text mining are considered. Regarding the capability to view historical data **(DR12, DP6, IP7)**, flexible incorporation of different datasets including different time periods is feasible. The identification of trends over time **(DR13, DP2, IP2)** is considered by the possibility to determine specific time periods, resulting in a flexible identification of time-specific trends and their evolution over time. Finally, the trends, their use case-specific peculiarities, and all customizabilities are visualized **(DR14, DP5, IP6)** by a developed graphical user interface (GUI).

Demonstration and Evaluation of MANTRA

To assess MANTRA's capabilities, we followed our defined evaluation strategy *(see Steps "demonstration" and "evaluation" in Section "Research Procedure")*. In general, MANTRA is designed as an artifact to identify trends from various unstructured social media data. We decided to conduct the demonstration and evaluation on OCRs of the academic Yelp dataset, as it is a publicly available real-world dataset and has proven appropriate for verifying the capabilities of machine learning techniques (e.g., Binder et al. 2019; Heinrich et al. 2022). The academic Yelp dataset (Yelp 2021) comprises about 8.6 million OCRs of 160,585 businesses in different fields (i.a., restaurants, cultural sites, sports facilities) and contains different context information (i.a., creation time, federal state, author information) concerning these OCRs. Because this dataset has already been subject to trend analysis within the food and restaurant industry (e.g., Heinrich et al. 2022; Wörner et al. 2022), we decided to narrow our demonstration to this domain, resulting in a dataset of about 1.03 million OCRs. According to a 2018 survey by a panel of experts from the NRAB, *vegan cuisine* and *global cuisine* are likely to gain popularity in the US until 2030 (Statista 2019). By means of explorative analysis, we have next identified that restaurants with these two orientations and corresponding food offerings are actually included in the dataset. Together with the results from the previous trend analysis for the food and restaurant domain, this ensured that the possibility of MANTRA producing so-called false positives is considerably reduced (Pries-Heje et al. 2007). Consequently, MANTRA identifies topics for the two trends *vegan cuisine* and *global cuisine* because they are actually contained within the Yelp dataset.

The configuration *(see Figure 3)* represents the view when starting the tool and can be used to customize the trend analysis. The layout is aligned with the GUI design suggestions of Garrett (2010). *Section (i)* presents an excerpt of the dataset used, including the actual reviews and the accompanying business data (e.g., federal state or business name). To consider a representative analysis period concerning our demonstration, the analysis uses data from 2015 to 2021 *(ii)*, as the related NRAB survey conducted in 2018.

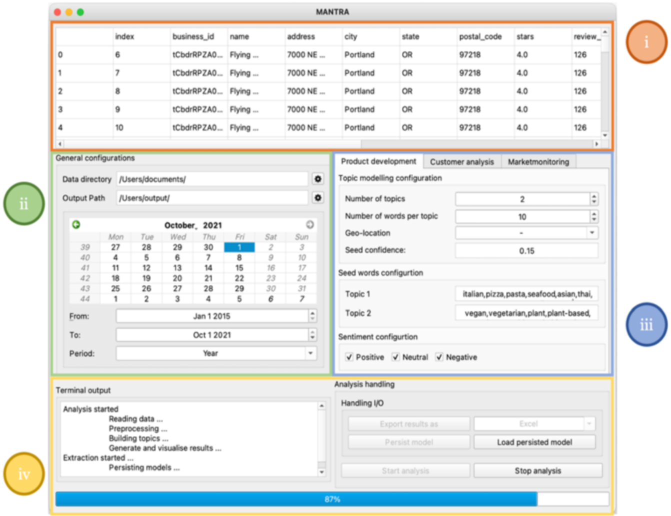


Figure 3. MANTRA – Configuration View

To incorporate the identified use cases, they were implemented in *section (iii)* using modular, dynamic tabs to enable a distinctive configuration. Here, regarding the use case (a) Product Development, the number of topics (*2*), words per topic (*10*), the respective external parameters of geolocation (*none*) and contextualized seed words (e.g., topic 1: Italian, Asian; topic 2: vegan, vegetarian) were defined. Moreover, all sentiment levels (*pos, neg, neu*) were considered. The start of the analysis, the handling of input and output (I/O) operations, and the monitoring take place in *section (iv)*. Once the process has been finished, the results are consolidated and displayed by responsive, interactive charts *(see Figures 4, 5, and 6)*.

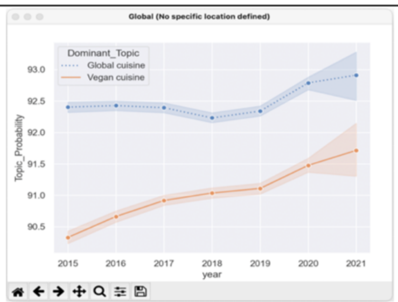


Figure 4. MANTRA – Identified Trends

Figure 5. Corresponding Sentiments (Vegan)

Figure 4 represents the determined topic probabilities with respect to the years, resulting in the tracking of the topic evolution based on their relevance. Here, both identified trends are continuously increasing in terms of their relevance. The trend *global cuisine* appears to slightly decrease from 2015 to 2018, but then continuously increases. The trend *vegan cuisine* increases steadily within the analyzed period. These results provide convincing evidence that both trends are on an upward trajectory and exhibit consistency with the findings described in the 2018 NRAB forecast. Moreover, the 2020 and 2022 Food & Health surveys by the International Food Information Council (IFIC 2020, 2022) show a significant percentage of US residents turned to a vegetarian/vegan diet relative to the previous years. As our tool detects a significant increase compared 2019 and 2020, convincing evidence is provided that our tool can identify even small deviations within the evolution of a trend. In addition, *Figure 5* comprises the trend-specific sentiments, by which a trend and the customer demands can be further deconstructed. All in all, our results confirm the general feasibility of MANTRA, as the elicited trends are consistent with the results of the two aforesaid surveys.

Trend topics	Products/Features (noun)	Descriptive characteristics (adj., verb)
vegan cuisine	food, restaurant, vegan, service, gluten, ...	good, great, delicious, best, unfriendly, free, ...

Table 3. Excerpt of the Resulting Excel

The analysis also comprises an Excel file to link the elicited results to new products/features. *Table 3* shows an excerpt of the resulting Excel file, containing the trend-related words and their classification into the part of speech for the topic *vegan cuisine*. This breaks down the trends into constituent parts, enabling the identification of specific products/features and related characteristics. As MANTRA depicts co-occurrences within the elicited trends by topic modeling, identifying sub-trends is feasible. Here, our analysis could, e.g., reveal that customers eating vegan also refer to gluten-free food in a vast frequency *(see Table 3)*.

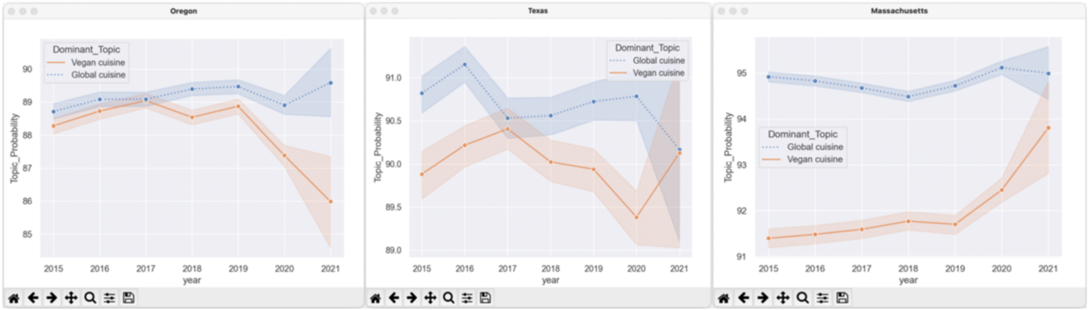


Figure 6. MANTRA - Identified Trends (From Left to Right: OR, TX, MA)

Since (a) **Product Development** refers to examining geolocation-based differences regarding the needs and expectations of customers, we further tested our artifact on sub-datasets referring to three different states in the US: *Massachusetts (MA), Texas (TX), and Oregon (OR)*. The results provide convincing evidence that MANTRA is able to identify geolocation-based discrepancies within the evolution of trends *(see Figure 6)*. It can be clearly seen that the topic *vegan cuisine* differs across the locations, as it consistently increases in MA but remains steady or even decreases in TX and OR. *Global cuisine* has declined since 2015 in TX, while slightly increasing in MA and OR. The sentiment distributions further reveal that – although the relevance of the vegan trend is vigorously increasing in TX from 2020 to 2021 – the respective reviews are negatively connotated *(pos: 0.27, neg: 0.61, neu: 0.12)*. In MA, the trend also increases sharply from 2019 to 2021 but is predominantly characterized by positive reviews *(pos: 0.56, neg: 0.3, neu: 0.14)*.

Generally, the development of our tool was based on the DRs, which are all met and technically realized *(see Figure 2 and the DPs and IPs)*. The 2018 and 2019/2020 trends could be verified by applying our tool to a representative real-world dataset, validating its functionality and therefore its practical applicability.

Discussion and Contributions

The implementation of the DRs and DPs has provided interesting results. As we have included the sentiment in our trend analysis tool, we could identify positive and negative inflections about the trends and how the sentiment values evolve over time. *Figure 5* shows that users spoke positively about the vegan trend in 2016. This positive resonance decreased until 2019, but in 2020 and 2021 the positive comments predominate again. Thus, our demonstration reveals that, although the vegan trend is consistently rising (in terms of topic probability) *(see Figure 4)*, the sentiment evolves distinctively. Without including the sentiment, one would not know that especially in 2019 ambivalent opinions were prevalent about the vegan trend and opening a vegan restaurant would not have yielded the success implicated by the upward trend.

To be successful and competitive in a targeted market, marketing representatives need to know which product and service features to design in which way to meet the global and local customers' expectations (Bae et al. 2018; Endo et al. 2015; Ha et al. 2017). This becomes particularly evident when comparing the three states of OR, TX, and MA *(see Figure 6)*. While both trends in MA have been growing steadily since 2015, the situation in TX and OR is different. In TX, the vegan trend declined between 2017 and 2020. However, in 2020, it started to gain importance. This is in contrast to the global cuisine trend which, peaked in 2016 but has dropped sharply since 2020. OR exhibits a completely different picture of the two trends. While the vegan trend initially remains stable and dropped sharply in 2019, the global cuisine trend remains stable over the years. The comparison of the three figures clearly indicates that flexibly including geolocation information at different levels of granularity (such as federal states) can have a considerable influence on whether and how to harness the trends. While aligning a restaurant to the vegan trend is (based on these analyses) advisable in MA, this is not the case in OR. To better assess these trends in TX and MA, we have combined the two alignments – geolocation and sentiment – because, as mentioned above, disclosing an upcoming trend does not directly mean that the trend exhibits necessarily positive connotations. Whereas the vegan trend in TX has gained in importance since 2020, the sentiment is predominantly characterized by negative contributions *(neg: 0.61)*. This indicates that either the vegan food offered so far does not meet the customers' expectations, or that vegan food is not well received in TX in general. In contrast to TX, the vegan trend in MA is predominantly characterized by positive reviews on average *(pos: 0.56)*. Without the inclusion of the geolocation data and the sentiment values, these aforementioned discrepancies in the results would not have been revealed.

Our investigation contributes to research and practice alike. As a contribution to practice, first companies can benefit from our comprehensive and modular artifact. MANTRA meets several DRs and, in contrast to the available solutions, does additionally provide a semi-supervised topic modeling technique *(GuidedLDA)*. Thus, MANTRA is capable of integrating several external parameters, which is deemed important in the extant research literature (Bae et al. 2018; Endo et al. 2015; Ha et al. 2017). Beyond that, MANTRA allows marketing representatives to integrate the use case-specific external parameters even at different granularity (e.g., federal states as geolocation) and therefore, i.a., to identify fine-grained discrepancies in customer perceptions. Moreover, we have implemented a functionality to link future customer preferences with new features (cf. DR4). By extracting the results of the trend analysis into an Excel file and classifying the trend-related words into the part of speech, trends and their constituent parts can be analyzed in more detail. For example, a company that has analyzed the trends that emerged in Section "*Demonstration and Evaluation of MANTRA*" could recognize that customers eating vegan also refer to gluten-free food with a vast frequency. As our results indicate, this functionality of a trend analysis tool is eminent as it supports companies, on the one hand, to identify fine-grained aspects of the elicited trends and, on the other hand, to support that the gained insights can be seamlessly included into the steps that are following the actual trend analysis (i.a., trend assessment and integration into the product and service portfolio of a company). In general, MANTRA has been developed in a modular way, enabling its users to define various settings through the corresponding GUI elements. For example, the time period for a specific trend can be set to conduct the analysis resource-efficiently. Several use case-specific DRs were implemented by offering modular and dynamic tabs to enable quick and almost effortless processing of common marketing-related use cases. As we have combined different machine learning approaches and designed our tool modularly, companies can adapt the analysis to their individual circumstances.

Furthermore, as an outcome of our DSR project, we achieved theoretical contributions to research that go beyond the technical contribution (i.e., the artifact). As we have designed a tool that is able to identify the voice of the customer, even about unknown customer demands and design features, our results also contribute to research theories such as Innovation Theory. Here, our approach can be included, e.g., in the stage-gate model of the innovation process that describes a conceptual and operational model for moving new product projects from idea to launch (Cooper 1996). Thereby, our results show that it is important to distinguish trends based on external parameters (e.g., geolocation), as customer perceptions may differ across global and local sites. By incorporating specific trend-related information within the different stages of the stage-gate model, the rigid specifications of stages and gates can be loosened up. Since social media enables customers to express their unfiltered opinions (Yan et al. 2014), thereon identified trends reflect customer perceptions, which can ensure that customers' needs are considered. Thus, by integrating the trend information in this way, the assessments at the go/kill checkpoints (i.e., gates) become less elaborate as the alignment with external requirements (e.g., customer wishes and demands) is maintained constantly. Moreover, Social Media Analytics Theory (especially concerning the research of automated trend analysis) states that the automated analysis of social media content still holds various challenges when it comes to its actual application (Fan and Gordon 2014; Stieglitz et al. 2018). With the design and development of our automated trend analysis tool, we purposefully combined various machine learning techniques (i.a., semi-supervised topic modeling, and sentiment analysis) and ensured that the DRs and the corresponding marketing-related use cases are technically realized. In this way, we contribute to Social Media Analytics Theory by proposing a need-fitted automated trend analysis for unstructured social media data.

The discussed results have shown the advantages of our tool "MANTRA" compared to existing ones, placing our DSR project in the group of improvements (development of new solutions for known problems) in terms of the DSR knowledge contribution framework of Gregor and Hevner (2013). DSR improvement projects contribute to both prescriptive theory, i.e., design theory (Gregor 2006), and descriptive theory, i.e., kernel theories such as Innovation- or Social Media Analytics Theory as described above (Gregor and Hevner 2013). By including GuidedLDA as a semi-supervised topic modeling technique within MANTRA that can flexibly integrate different external parameters as deemed crucial by our marketing-related use cases, our research as well affects the quadrant of exaptation (cf. Gregor and Hevner 2013). We have shown the proficiency of this technique and in this way contribute to IS and DSR research alike, as we propose semi-supervised topic modeling as a suitable technique for automated trend analysis (tools) to tailor the derived trends towards context factors and semantically towards the company and its fields of activity.

Based on the DRs derived from literature and kernel theories, DPs were formulated and proposed. By applying them during the design and development of the artifact followed by the demonstration and evaluation, an implicit empirical grounding of the DPs was achieved here (Heinrich and Schwabe 2014). Our DPs capture design-related knowledge and can thus support the development of further IS (design) theories and new artifacts. By considering e.g., DP2 Contextual adaptability, the importance of the contextualization (e.g., location) is highlighted. Since context directly impacts trend analysis results, the alignment with context leads to more meaningful insights. With the DPs, we made a first step towards contributing to design theory in terms of theory for design and action (Gregor 2006) as we comply with the conditions of March and Smith (1995) and Hevner et al. (2004), who pointed out under which a contribution to knowledge in DS has occurred: utility to a community of users, the novelty of the artifact and the persuasiveness of claims that it is effective.

Conclusion, Limitation, and Future Research

The early identification of new and auspicious ideas as well as trending topics regarding the development of products and services, the analysis of customers' behaviors, and the monitoring of markets and competing brands lead to competitive advantages for companies. However, prior literature and existing solutions for automated trend analysis on unstructured social media data do not incorporate semi-supervised topic modeling and do not sufficiently cover the different external use case-specific parameters as well as the specific requirements of the common marketing-related use cases (a) Product Development, (b) Customer Behavior Analysis, and (c) Market-/Brand-Monitoring.

Therefore, we identify several DRs and derive DPs. To technically realize the DPs, we combine several machine learning techniques and transfer these in an intuitive GUI to close the revealed research gaps *(see RQ1 and RQ2)*. Especially with the demonstration to about 1.03 million OCRs, we show that covering all DRs is essential for a target-oriented and feasible trend analysis. MANTRA also supports unsupervised trend analysis when no seed words are provided. Hence, also marketing representatives who have not gained prior knowledge of underlying trends can also conduct an explorative trend analysis using MANTRA.

Our investigation proposes contributions to practice and research alike *(see RQ2)*. Companies can benefit from our comprehensive and modular artifact, with which large amounts of unstructured social media data can be analyzed in a way best suited to the company's individual circumstances. We have also highlighted how our investigation made a first step towards contributing to design theory and kernel theories. Regarding Innovation Theory, the rigid sequence of stages and gates in the stage-gate model (Cooper 1996) can be loosened up by examining external requirements constantly. In this vein and according to the demonstration of MANTRA, it is exemplified that the stages should incorporate activities to assess the influences of the different external parameters (e.g., continents, phases of a day, personal characteristics; in the demonstration: federal states) on the perception of trends. By incorporating the DRs of a trend analysis tool, we enrich Social Media Analytics Theory with a need-fitted automated approach for analyzing social media content regarding the automated identification of trends.

It is worth mentioning, that the initial assessments on the practical applicability of MANTRA regarding (b) Customer Behavior Analysis and (c) Market-/Brand-Monitoring revealed promising results. Therefore, we will investigate these two use cases in more detail in the future and examine in more detail the potential of combining the three identified use cases with each other. In addition, to take the next steps toward a more mature design theory, we will first evaluate our artifact in a formative and artificial environment (e.g., a laboratory experiment). This allows us to improve our tool (whereby our DPs can be confirmed or adapted) before conducting a more elaborate evaluation in a more natural setting as a further part of the design cycle. There are also limitations to this research: Although we included a large set of investigations, we could identify probably even more use cases in further literature. Additionally, researchers from other fields could identify other use cases. Nevertheless, the identified use cases are undoubtedly important for marketing.

References

Ahadh, A., Binish, G., and Srinivasan, R. 2021. "Text mining of accident reports using semi-supervised keyword extraction and topic modeling," *Process Safety and Environmental Protection* (155).

Axtell, C., Holman, D., Unsworth, K., Wall, T., Waterson, P., and Harrington, E. 2000. "Shopfloor innovation: Facilitating the suggestion and implementation of ideas," *J. Occup. Psychol.* (73:3).

Bae, J., Havsol, J., Karpefors, M., Karlsson, A., and Mathiason, G. 2018. "Short Text Topic Modeling to Identify Trends on Wearable Bio-Sensors in Different Media Type," in *Proceedings of the 6th International Symposium on Computational and Business Intelligence*.

Bai, S., He, H., Han, C., Yang, M., Yu, D., Bi, X., and Panigrahi, P. 2023. "Exploring thematic influences on theme park visitors' satisfaction: An empirical study on Disneyland China," *J. Consum. Behav.* (22:2).

Batrinca, B., and Treleaven, P. 2015. "Social media analytics: a survey of techniques, tools and platforms," *AI & Society* (30), pp. 89-116.

Bhor, H., Koul, T., Malviya, R., and Mundra, K. 2018. "Digital Media Marketing Using Trend Analysis on Social Media," in *Proceedings of the 2nd International Conference on Inventive Systems and Control*.

Binder, M., Heinrich, B., Klier, M., Obermeier, A., and Schiller, A. 2019. "Explaining the stars: Aspect-based sentiment analysis of online customer reviews," in *Proceedings of the ECIS 2019*.

Böckle, M., Bick, M., and Novak, J. 2021. "Toward a Design Theory of User-Centered Score Mechanics for Gamified Competency Development," *Information Systems Management* (40:1), pp. 1-27.

Burmeister, K., Neef, A., Albert, B., and Glockner, H. 2002. "Zukunftsforschung und Unternehmen," *Praxis, Methoden, Perspektiven*. Essen.

Chandra, L., Seidel, S., and Gregor, S. 2015. "Prescriptive Knowledge in IS Research: Conceptualizing Design Principles in Terms of Materiality, Action, and Boundary Conditions," in *Proceedings of the 48th Hawaii International Conference on System Sciences*.

Chandrasekaran, R., Mehta, V., Valkunde, T., and Moustakas, E. 2020. "Topics, trends, and sentiments of tweets about the COVID-19 pandemic," *Journal of Medical Internet Research* (22:10).

Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J., and Blei, D. 2009. "Reading Tea Leaves: How Humans Interpret Topic Models," *Advances in Neural Information Processing Systems*, pp. 288-296.

Chen, J., Yu, J., and Shen, Y. 2012. "Towards Topic Trend Prediction on a Topic Evolution Model with Social Connection," in *Proceedings of the International Conferences on WI-IAT*.

Cooper, R. 1996. "Overhauling the new product process," *Industrial Marketing Management* (25:6).

Crain, S., Zhou, K., Yang, S., and Zha, H. 2012. "Dimensionality Reduction and Topic Modeling: From Latent Semantic Indexing to Latent Dirichlet Allocation and Beyond," *Mining Text Data*, pp. 129-161.

Ding, K., Choo, W., Ng, K., Ng, S. 2020. "Employing structural topic modelling to explore perceived service quality attributes in Airbnb accommodation," *International Journal of Hospitality Management* (91).

Duncan, T., and Moriarty, S. 1998. "A Communication-Based Marketing Model for Managing Relationships," *Journal of Marketing* (62:2), pp. 1-13.

Endo, Y., Toda, H., and Koike, Y. 2015. "What's Hot in the Theme: Query Dependent Emerging Topic Extraction from Social Streams," in *Proceedings of the 24th International Conference on the WWW*.

Fan, W., and Gordon, M. 2014. "The power of social media analytics," *Communications of the ACM* (57:6).

Garrett, J. 2010. "The Elements of User Experience: User-Centered Design for the Web and Beyond".

Gregor, S. 2006. "The nature of theory in information systems," *MIS Quarterly* (30:3), pp. 611-642.

Gregor, S., Chandra Kruse, L., and Seidel, S. 2020. "Research Perspectives: The Anatomy of a Design Principle," *Journal of the AIS* (21:6).

Gregor, S., and Hevner, A. 2013. "Positioning and Presenting Design Science Research for Maximum Impact," *MIS Quarterly* (37:2), pp. 337-355.

Gregor, S., and Jones, D. 2007. "The Anatomy of a Design Theory," *Journal of the AIS* (8:5).

Gulati, R., and Oldroyd, J. 2005. "The Quest for Customer Focus," *Harvard Business Review* (83:4).

Ha, T., Beijnon, B., Kim, S., Lee, S., and Kim, J. 2017. "Examining User Perceptions of Smartwatch through Dynamic Topic Modeling," *Telematics and Informatics* (34:7), pp. 1262-1273.

Heinrich, B., Hollnberger, T., Hopf, M., and Schiller, A. 2022. "Long-term Sequential and Temporal Dynamics in Online Consumer Ratings," in *Proceedings of the ECIS 2022*.

Heinrich, P., and Schwabe, G. 2014. "Communicating Nascent Design Theories on Innovative Information Systems through Multi-Grounded Design Principles," in *Proceedings of DESRIST 2014*.

Hevner, A., March, S., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *MIS Quarterly* (28:1), pp. 75–105.

Hicks, A., Comp, S., Horovitz, J., Hovarter, M., Miki, M., and Bevan, J. 2012. "Why People Use Yelp.com: An Exploration of Uses and Gratifications," *Computers in Human Behavior* (28:6), pp. 2274-2279.
Hong, L., Ahmed, A., Gurumurthy, S., Smola, A., and Tsioutsiouliklis, K. 2012. "Discovering Geographical Topics in the Twitter Stream," in *Proceedings of the 21st international conference on WWW*.
Hong, L., and Davison, B. D. 2010. "Empirical Study of Topic Modeling in Twitter," in *Proceedings of the 1st Workshop on Social Media Analytics*.
Humphreys, A., and Wang, J.-H. R. 2018. "Automated text analysis for consumer research," *Journal of Consumer Research* (44:6), pp. 1274-1306.
Hutto, C., and Gilbert, E. 2014. "Vader: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," in *Proceedings of the International AAAI Conference on Web and Social Media*.
IFIC (2020). "2020 Food & Health Survey. 9th June 2020," *https://foodinsight.org/2020-food-and-health-survey/*. Lastly accessed on 6th April 2023.
IFIC (2022). "2022 Food & Health Survey. 18th May 2022," *https://foodinsight.org/2022-food-health-survey/*. Lastly accessed on 6th April 2023.
Irawan, M., Wijayanto, R., Shahab, M., Hidayat, N., and Rukmi, A. 2020. "Implementation of Social Media Mining for Decision Making in Product Planning Based on Topic Modeling and Sentiment Analysis," *Journal of Physics: Conference Series*.
Iwata, T., Watanabe, S., Yamada, T., and Ueda, N. 2009. "Topic tracking model for analyzing consumer purchase behavior," in *Proceedings of the 21st International Joint Conference on AI*.
Jagarlamudi, J., Daumé III, H., and Udupa, R. 2012. "Incorporating Lexical Priors into Topic Models," in *Proceedings of the 13th Conference of the European Chapter of the ACL*.
Jeong, B., Yoon, J., and Lee, J.-M. 2019. "Social Media Mining for Product Planning: A Product Opportunity Mining Approach Based on Topic Modeling and Sentiment Analysis," *Int. J. Inf. Manag.* (48).
Kaplan, A., and Haenlein, M. 2010. "Users of the World, Unite! The Challenges and Opportunities of Social Media," *Business Horizons* (53:1), pp. 59-68.
Khan, H. U., Nasir, S., Nasim, K., Shabbir, D., and Mahmood, A. 2021. "Twitter trends: a ranking algorithm analysis on real time data," *Expert Systems with Applications* (164).
Ko, N., Jeong, B., Choi, S., and Yoon, J. 2018. "Identifying product opportunities using social media mining: application of topic modeling and chance discovery theory," *IEEE Access* (6), pp. 1680-1693.
Lozano, M., Schreiber, J., and Brynielsson, J. 2017. "Tracking Geographical Locations Using a Geo-Aware Topic Model for Analyzing Social Media Data," *Decision Support Systems* (99), pp. 18-29.
March, S., and Smith, G. 1995. "Design and Natural Science Research on Information Technology," *Decision Support Systems* (15:4), pp. 251-266.
Mastrogiacomo, L., Barravecchia, F., Franceschini, F., and Marimon, F. 2021. "Mining quality determinants of product-service systems from user-generated contents," *Quality Engineering* (33:3), pp. 425-442.
Mayring, P. 2000. "Qualitative Content Analysis," *Forum Qualitative Social Research* (1:2).
Mayring, P. 2002. "Introduction to qualitative social research," *Weinheim: Beltz*.
Möller, F., Guggenberger, T. M., and Otto, B. 2020. "Towards a method for design principle development in information systems," in *Proceedings of DESRIST 2020*.
Moro, S., Lopes, R., Esmerado, J., and Botelho, M. 2020. "Service quality in airport hotel chains through the lens of online reviewers," *Journal of Retailing and Consumer Services* (56).
Mudambi, S. M., and Schuff, D. 2010. "Research note: What makes a helpful online review? A study of customer reviews on Amazon.com," *MIS Quarterly* (34:1), pp. 185-200.
Netzer, O., Feldman, R., Goldenberg, J., and Fresko, M. 2012. "Mine your own business: Market-structure surveillance through text mining," *Marketing Science* (31:3), pp. 521-543.
Park, E., Kwon, J., and Kim, S. 2021. "Green marketing strategies on online platforms: a mixed approach of experiment design and topic modeling," *Sustainability* (13:8).
Peffers, K., Tuunanen, T., Rothenberger, M., Chatterjee, S. 2007. "A Design Science Research Methodology for Information Systems Research," *Journal of Management Information Systems* (24:3).
Peffers, K., Rothenberger, M., Tuunanen, T., and Vaezi, R. 2012. "Design science research evaluation," in *Proceedings of DESRIST 2012*.
Pries-Heje, J., Baskerville, R., and Venable, J. 2007. "Soft design science research: Extending the boundaries of evaluation in design science research," in *Proceedings of DESRIST 2007*, pp. 18-38.
Puranam, D., Narayan, V., and Kadiyali, V. 2017. "The effect of calorie posting regulation on consumer opinion: A flexible latent Dirichlet allocation model with informative priors," *Marketing Science* (36:5).
Purao, S., Kruse, L. C., and Maedche, A. 2020. "The Origins of Design Principles: Where do... they all come from?, " in *Proceedings of DESRIST 2020*, pp. 183-194.

Qu, Q., Chen, C., Jensen, C., and Skovsgaard, A. 2015. "Space-Time Aware Behavioral Topic Modeling for Microblog Posts," *IEEE Data Engineering Bulletin* (38:2), pp. 58-67.

Reisenbichler, M., and Reutterer, T. 2019. "Topic modeling in marketing: recent advances and research opportunities," *Journal of Business Economics* (89), pp. 327-356.

Saha, A., and Sindhwani, V. 2012. "Learning Evolving and Emerging Topics in Social Media: A Dynamic NMF Approach with Temporal Regularization," in *Proceedings of the Conference on WSDM 2012*.

Sasaki, K., Yoshikawa, T., and Furuhashi, T. 2014. "Online Topic Model for Twitter Considering Dynamics of User Interests and Topic Trends," in *Proceedings of the Conference on Empirical Methods in NLP*.

Shi, L., Wu, Y., Liu, L., Sun, X., and Jiang, L. J. 2018. "Event Detection and Identification of Influential Spreaders in Social Media Data Streams," *Big Data Mining and Analytics* (1:1), pp. 34-46.

Sigala, M. 2012. "Social Networks and Customer Involvement in New Service Development (Nsd): The Case of www.Mystarbucksidea.com," *International Journal of Contemporary Hospitality* (24:7).

Sohn, B., Lim, H., and Choi, J. 2019. "The Prediction of Sales Volume and WoM Effect through Network Structure and Topic Modeling on Social Media," in *Proceedings of the PACIS 2019*.

Statista. 2019. "Ranking der prognostizierten Restaurant-Trends im Bereich 'Food & Menu'," *https://de.statista.com/statistik/daten/studie/1125284*. Lastly Accessed on 6th April 2023.

Stieglitz, S., Mirbabaie, M., Ross, B., and Neuberger, C. 2018. "Social Media Analytics – Challenges in Topic Discovery, Data Collection, and Data Preparation," *Int. J. Inf. Manag.* (39), pp. 156-168.

Sundbo, J. 1995. "Three paradigms in innovation theory," *Science and Public Policy* (22:6), pp. 399-410.

Tirunillai, S., and Tellis, G. 2014. "Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation," *Journal of Marketing Research* (51:4).

Tseng, M., and Piller, F. 2003. "The customer centric enterprise," *Springer Science & Business Media*.

Toubia, O., Iyengar, G., Bunnell, R., and Lemaire, A. 2019. "Extracting features of entertainment products: A guided latent dirichlet allocation approach informed by the psychology of media consumption," *Journal of Marketing Research* (56:1), pp. 18-36.

Tucker, C., and Kim, H. 2011. "Predicting Emerging Product Design Trend by Mining Publicly Available Customer Review Data," in *Proceedings of the 18th International Conference on Engineering Design*.

Ulbrich, F., Christensen, T., and Stankus, L. 2011. "Gender-Specific on-Line Shopping Preferences," *Electronic Commerce Research* (11:2), pp. 181-199.

Valdez, D., Padon, A., Barry, A., and Russell, A. 2018. "Alcohol Advertising on Twitter - a Topic Model," *American Journal of Health Education* (49:4), pp. 256-263.

Venable, J., Pries-Heje, J., and Baskerville, R. 2016. "FEDS: a framework for evaluation in design science research," *European Journal of Information Systems* (25:1), pp. 77-89.

Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., and Cleven, A. 2015. "Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research," *Communications of the AIS* (37:1), pp. 205-224.

Walls, J., Widmeyer, G., and El Sawy, O. 1992. "Building an Information System Design Theory for Vigilant EIS," *Information Systems Research* (3:1), pp. 36-59.

Wörner, J., Konadl, D., Schmid, I., and Leist, S. 2022. "Supporting Product Development by a Trend Analysis Tool Applying Aspect-Based Sentiment Detection," in *Proceedings of DESRIST 2022*.

Yan, X., Guo, J., Lan, Y., Xu, J., and Cheng, X. 2015. "A Probabilistic Model for Bursty Topic Discovery in Microblogs," in *Proceedings of the 29th AAAI Conference on Artificial Intelligence*.

Yan, Z., Xing, M., Zhang, D., Ma, B., and Wang, T. 2014. "A Context-Dependent Sentiment Analysis of Online Product Reviews Based on Dependency Relationships," in *Proceedings of the ICIS 2014*.

Yang, Y., Yao, Q., and Qu, H. 2017. "Vistopic: A Visual Analytics System for Making Sense of Large Document Collections Using Hierarchical Topic Modeling," *Visual Informatics* (1:1), pp. 40-47.

Ye, H. J., and Kankanhalli, A. 2018. "User service innovation on mobile phone platforms: Investigating impacts of lead userness, toolkit support, and design autonomy," *MIS Quarterly* (42:1), pp. 165-188.

Yelp. 2021. "Yelp Open Dataset," *https://www.yelp.com/dataset*. Accessed in April 2022.

Zhang, H., Kim, G., and Xing, E. P. 2015. "Dynamic Topic Modeling for Monitoring Market Competition from Online Text and Image Data," in *Proceedings of the 21th KDD Conference*.

Zhang, P., Gu, H., Gartrell, M., Yang, D., Ding, X., Gu, N. 2016. "Group-based Latent Dirichlet Allocation (Group-LDA): Effective audience detection for books in online social media," *Knowl. Based Syst.* (105).

Zhao, K., Cong, G., Chin, J.-Y., and Wen, R. 2019. "Exploring Market Competition over Topics in Spatio-Temporal Document Collections," *The VDLB Journal* (28:1), pp. 123-145.

Zhong, N., and Schweidel, D. 2020. "Capturing Changes in Social Media Content: A Multiple Latent Changepoint Topic Model," *Marketing Science* (39:4), pp. 827-846.

2.4 Supporting Product Development by a Trend Analysis Tool Applying Aspect-Based Sentiment Detection

<table>
<tr><td rowspan="2">Adressierte Zielsetzungen</td><td>Ableitung der Anforderungen an ein aspekt-, kontext- und tonalitätsbasiertes Trendanalysewerkzeug zur Unterstützung der Produktentwicklung. Ziel ist die Konzeption eines datengetriebenen Analyseansatzes, der durch die Kombination moderner Verfahren der Data Science (z. B. Topic Modelling, Aspekt-basierte Sentimentanalyse, Word Embeddings) eine feingranulare Auswertung relevanter Produkt- und Serviceaspekte in UGC ermöglicht.</td><td>ZS5</td></tr>
<tr><td>Prototypische Entwicklung und Demonstration des konzipierten Trendanalysewerkzeugs zur differenzierten Beobachtung von Kundenstimmungen hinsichtlich relevanter Produkt- und Serviceaspekte. Durch die Integration kontextsensitiver Data Science Methoden soll eine zeit- und ortsbezogene Analyse relevanter Produkt- und Dienstleistungsmerkmale ermöglicht werden, um eine ganzheitliche Entscheidungsunterstützung im Innovationsprozess zu fördern.</td><td>ZS6</td></tr>
<tr><td>Abgeleitete Forschungsfragen</td><td>Wie könnte ein softwaregestütztes Analysewerkzeug zur Aspekt-basierten Sentimentanalyse gestaltet sein, das die Trendanalyse im Rahmen der Produktentwicklung unterstützt, und welche Anforderungen muss ein solches Werkzeug erfüllen?</td><td>FF6</td></tr>
<tr><td>Genutzte Forschungsmethode</td><td colspan="2">Umsetzung nach dem DSR-Prozessmodell von Peffers et al. (2007). Orientierung an den drei DSR-Zyklen nach Hevner et al. (2004):
Design Cycle: Konzeption, Umsetzung, Demonstration und iterative Verfeinerung des kontextsensitiven, Aspekt-basierten Trendanalysewerkzeugs.
Relevance Cycle: Orientierung an praxisrelevante Anforderungen aus dem Anwendungsfeld der Produktentwicklung.
Rigor Cycle: Theoretische Fundierung des Artefakts durch die Integration der Social Media Analytics Theory sowie durch die Leistung eines Beitrags zur Weiterentwicklung der Nascent Design Theorie im Sinne von Gregor und Hevner (2013).
Die technische Umsetzung basiert auf einer Kombination differenzierter Data Science Verfahren:
Topic Modelling zur automatisierten Aspektextraktion.
Convolutional Neural Networks (CNN) zur aspektgesteuerten Analyse auf Basis domainspezifischer Word Embeddings.
Sentimentanalyse zur Bewertung der Tonalität.</td></tr>
<tr><td>Kernergebnisse (Überblick)</td><td colspan="2">Durch die Konzeption und Entwicklung des Analysewerkzeugs wurden mit Blick auf FF6 mehreren Kernergebnisse geschaffen:
Ableitung und Operationalisierung von zehn Design Anforderungen sowie Entwicklung entsprechender Design Prinzipien für die Gestaltung Aspekt-basierter Analysewerkzeuge.
Kombination unsupervisierter (Topic Modelling) und supervisierter (CNN und Word Embeddings) Data Science Verfahren, wodurch flexible Aspektextraktion – mit und ohne Vorwissen – ermöglicht wird.</td></tr>
</table>

	• Demonstration des Tools anhand von 37.638 Yelp-Reviews eines US-Restaurantkonzerns. o Identifikation relevanter Produktaspekte (z. B. *Fleisch*, *Sauce*) des Produkts *Burger* durch die Analyse der VoC. o Analyse der Aspekt-basierten Sentimentverläufe über die Zeit zur Erkennung produktbezogener Trendentwicklungen. o Berücksichtigung regionaler Unterschiede (Massachusetts vs. Texas) zur Identifikation ortsspezifischer Wahrnehmungsdifferenzen. • Erkenntnis: Das entwickelte Tool ermöglich eine fein-granulare und kontext-sensitive Sentimentanalyse und liefert somit differenziertere Entscheidungsgrundlagen für die Produktentwicklung als herkömmliche Analysen. • Beitrag zur *Nascent Design Theorie*: Durch die Ableitung von Design Prinzipien im Kontext automatisierter Trendanalysen wurde die Nascent Design Theorie durch die implizierte Fundierung der Design Prinzipien (Heinrich und Schwabe, 2014) erweitert.
Publikationsort	Das Paper wurde 2022 auf der 17th International Conference on Design Science Research in Information Systems and Technology (*DESRIST*), St. Petersburg, Florida (USA) veröffentlicht und ist als peer-reviewtes Konferenzpapier gelistet.
Ranking VHB JQ 4	**B**
Autoren und Anteile	**Janik Wörner** **30%** Daniel Konadl 30% Isabel Schmid 30% Susanne Leist 10%
DOI	http://doi.org/10.5283/epub.52213

Tabelle 18. **Fact Sheet zur 4. Veröffentlichung.**

Supporting Product Development by a Trend Analysis Tool applying Aspect-Based Sentiment Detection

Abstract. Incorporating product trends into innovation processes is imperative for companies to meet customers' expectations and to stay competitive in fiercely opposing markets. Currently, aspect-based sentiment analysis has proven an effective approach for investigating and tracking towards products and corresponding features from social media. However, existing trend analysis tools on the market that offer aspect-based sentiment analysis capabilities, do not meet the requirements regarding the use case Product Development. Therefore, based on these requirements, we implemented an automated artifact by following the design science research. We applied our tool to real-world social media data (37,638 Yelp reviews) from one major fast-food restaurant in the US, and thereby demonstrated that our tool is capable of identifying remarkable and fine-grained product trends.

Keywords: trend analysis tool, aspect-based sentiment, product development.

1 Motivation

Social media such as Yelp or Twitter have evolved rapidly over the last years. These platforms have become increasingly important for interaction in both private and business contexts [1, 2]. As social media is a channel for the exchange of user-generated content and unfiltered voices about products, services and the company in general, social media data contain the so-called "Voice of the Customer" (VoC). Thus, the VoC provides deep insights into customers' current expectations. To meet customers' expectations, marketing representatives need to identify and continuously track trending topics regarding product and service features and incorporate the VoC into product innovation processes. For example, identified product features and correspondingly mentioned opinions may indicate shortcomings (e.g., low battery capacity of a smartphone) and which improvements to be made to meet customers' requirements (e.g., [3]). One possibility to identify these shortcomings in an automated way from social media texts is to conduct aspect-based sentiment analysis [4].

The potential of aspect-based sentiment analysis for tracking fine-grained trends over time has already been recognized in practice and in theory (e.g., [3, 5-9]). This has led to the emergence of trend analysis tools that include aspect-based sentiment analysis functionalities. However, trend analysis tools available on the market have remarkable drawbacks as they do not cover the comprehensive requirements that are deemed essential within the extant literature for the use case Product Development (e.g., [3, 5-9]).

With this work at hand, we make practical as well as theoretical contributions. We address drawbacks of existing software tools by suggesting a comprehensive artifact for automated trend analysis that allows marketing representatives to conduct aspect-based sentiment analysis. To meet several use case-specific requirements, we focus especially on the combination of different data analysis methods regarding the particular requirements, leading to a constructive trend analysis. By this, we aim to propose an automated solution for identifying ideas as the basis of (incremental) product innovation. Summing up, the research at hand is guided by the following research question:

What could an aspect-based sentiment analysis tool that supports trend analysis for Product Development purposes look like, and which requirements should such a tool meet?

The remainder of this paper is structured as follows: In the next section, we provide conceptual basics and related work. Following on this, we turn to the Design Requirements (DRs) and Design Principles (DPs) for implementing our tool, and to the shortcomings of trend analysis tools on the market. Next, we show the research methodology. After a description of the tool's design and development as well as its demonstration, the paper concludes with a discussion and its contributions to theory and practice.

2 Foundations and Related Work

2.1 Conceptual Background

Social media serves as an important interface between companies and customers. In content communities, users can evaluate e.g., products by disseminating their opinions in form of online customer reviews (OCR). In doing so, customers not only rate products as a whole but express their opinions and attitudes towards different features of the rated items (e.g., service quality in a Yelp restaurant review). In this way, OCR not only help customers to make informed decisions but are also beneficial for driving innovations of products within companies. As OCR include customers' experiences and expectations of product features [10], unfiltered and in real-time [11], they can serve as a valuable resource for product innovations. Thus, OCR can be harnessed to identify ideas, to either develop new value propositions (i.e., disruptive innovation) or to improve the performance of existing products (i.e., incremental innovation) [12, 13].

To identify ideas for product improvements as well as product development and therefore to drive incremental product innovations, marketing representatives can conduct aspect-based sentiment analysis. The first step of an aspect-based sentiment analysis deals with extracting aspects from OCR. For this purpose, unsupervised as well as supervised techniques can be applied. Topic modeling techniques (e.g., LDA [14]) suggest a possibility to identify aspects without prior knowledge (i.e., unsupervised) [14]. Compared to that, supervised techniques (e.g., artificial neural networks) need first to be trained on training data (e.g., ontologies) to extract the proper aspects (cf. [4]). Subsequently, the expressed tonalities can be identified for each of the aspects [4] by means of automated sentiment analysis techniques [4, 15]. Therefore, aspect-based sentiment analysis offers benefits in terms of Product Development. For marketing representatives that lack the ability to implement aspect-based sentiment analysis themselves, trend analysis tools on the market offer this functionality in a ready-to-use way. However, these tools show remarkable drawbacks as they do not cover the comprehensive requirements that are deemed essential within the extant literature for the use case of Product Development.

2.2 Design Requirements and Available Tools on the Market

In a first step, we have comprehensively searched and consolidated literature (cf. [16]) to identify DRs of a trend analysis tool that applies aspect-based sentiment analysis for Product Development purposes. Based on the attention and importance received, we could derive several DRs (DR1)-(DR10) (see fig. 2) from the extant literature.

Concerning the identification of product features, the tool (DR1) should be capable of extracting the aspects autonomously from social media posts (e.g., [6-8, 17]). However, if marketing representatives have already knowledge documented about a domain problem (e.g., domain ontologies or product trees), the tool (DR2) should provide the option to include this prior knowledge into the automated identification of aspects (cf. [6, 7, 17, 18]). Beyond that, customers' self-reported opinions of product features play an important role for Product Development. Marketing representatives aim to retain the features that evoke positive perceptions, while features evoking negative perceptions need to be improved. Thus, (DR3) determining the polarity as well as the intensity of the opinions expressed about respective aspects is mandatory [3, 6, 8, 19, 20]. Subsequently, the product features can be adapted so that customers' requirements are met (e.g., increasing smartphone screen size). However, adapting features may influence perceptions of the features customers currently appreciate (e.g., high battery capacity of the smartphone). Therefore, to support informed improvement decisions, the tool (DR4) needs to identify the dependencies between product features [7, 9, 17]. Furthermore, to be successful and competitive in a targeted market, marketing representatives must decide which product features to propose in which way to meet customers' expectations within geographical markets. Product trends do also converge over time as they are dynamic developments and not solely static points in time. It is therefore essential that the tool (DR5) can flexibly match aspect-sentiment relations to different geographical (e.g., continents, countries, federal states) and temporal (e.g., days of a week, phases of a day) parameters [3, 5, 17-20]. The huge volume of available social media posts requires the incorporated techniques (DR6) to deal with vast amounts of textual data [8, 9, 18-20]. As past developments of trends are essential for assessing the current state of trending topics, the tool (DR7) needs to allow the user to consider historical data [3, 17-19]. To support Product Development in prioritizing product improvement decisions, the tool (DR8) needs to output aggregated sentiment values for the identified aspects [3, 6, 8, 19, 20] and (DR9) illustrate the frequency of the identified aspects [3, 8, 20, 21]. To immediately identify the most important aspect-sentiment relations, the tool (DR10) should provide means to rank the results in either descending or ascending order [3, 21].

In the second step, we searched the market for available trend analysis tools. We took an up-close look at the most popular tools (e.g., Brandwatch, Meltwater, Symanto) that offer trend analysis by means of aspect-based sentiments. We analyzed the functionality of these tools by installing demo versions. To verify the drawbacks regarding the DRs for each of the tools, we also turned to sales representatives from these companies to confirm our observations. As it turned out, there is indeed a lack of a software tool that meets all the specific requirements for the use case Product Development. Firstly, these tools lack the ability to flexibly match different temporal and geographical parameters to aspect-sentiment relations. Perceptions of product features differ across geolocations (e.g., continents, federal states), trends are temporal developments (e.g., days of a week, phases of a day), and both require different and flexible contexutalizations. Secondly, existing trend analysis tools that apply aspect-based sentiment analysis extract aspects, either with or without incorporating prior knowledge. However, we could not observe a trend analysis tool that includes both possibilities. Literature unveils the need for a comprehensive trend analysis tool that meets all the requirements for the use case Product Development. With this research, we aim to close this gap.

3 Research Procedure

In order to develop a systematic artifact for the automated trend analysis in marketing, we followed the Design Science (DS) approach [22, 23] and aligned our research activities with the DS procedure as proposed by [23] (see Fig. 1).

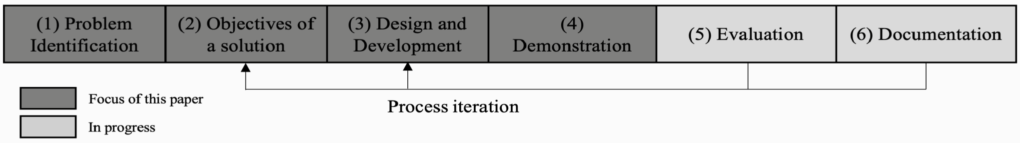


Fig 1. Design Science Research (DSR) Procedure

As a first step, (1) corresponding problems and drawbacks of previously existing approaches regarding the automated trend analysis using aspect-based sentiment analysis were identified (see sections 1 and 2.2). Hence, the revised tools supporting trend analysis by aspect-based sentiment detection do not meet the requirements that are indispensable to the successful application. Consequently, our (2) objective is to address drawbacks of existing software tools by suggesting a comprehensive artifact for automated trend analysis that allows marketing representatives to conduct aspect-based sentiment analysis (see sections 2.2 and 4). The third step of our DS process model contains the (3) design and development (see section 4) of an artifact. To fill the gaps identified within phase (1), we focus on the design of the technical realization of the tool by combining different machine learning techniques, following our derived DPs. Thus, our approach was established to support the trend analysis and to eliminate the existing disadvantages. By (4) demonstrating our artifact (see section 5.1), we highlight the application of our tool on 37,638 Yelp reviews [24]. Thus, we showed the implementation of the requirements identified in literature. In Step 5 the usefulness, applicability and usability of the tool are to be analyzed in a larger field study. Finally, the tool will be further enhanced before it is provided to marketing departments of large companies (6).

The orientation towards the procedure by [23] also makes it possible to align our research with the guidelines of [22] or [25], respectively. According to the design cycle, we present our artifact as the result that has gone through the process of demonstration (application of our approach to a Yelp dataset). In view of the relevance cycle, we identified several DRs from current research literature that guided the design of the artifact, and so the practical application of our artifact brought up several contributions for practice. In view of the rigor cycle, we used several methods to rigorously construct our artifact (e.g., topic modeling, sentiment analysis, neural networks) and derived initial findings as contributions to (nascent) design theory.

4 Design and Development

First, the composition of Meta Requirements (MRs) that describe "*what the system is for*" ([26], p. 325) is based on the purpose and scope of the tool that was discussed in the motivation. Thus, we define the solution objectives based on the investigations´ problems and present them in Figure 2. Besides the MRs, the Design Principles (DPs) are synthesized in a next step. DPs are defined as prescriptive statements that show how to do something to achieve a goal [27]. These DPs are deduced from the Design Requirements (DRs) that are based on current research literature. The DPs we derive from our results fall into the category of "action and materiality-oriented design principles", describing what an artifact should enable users to do and how the artifact should be built to do so [28]. The development of the DPs follows the guidelines of [28] and [27].

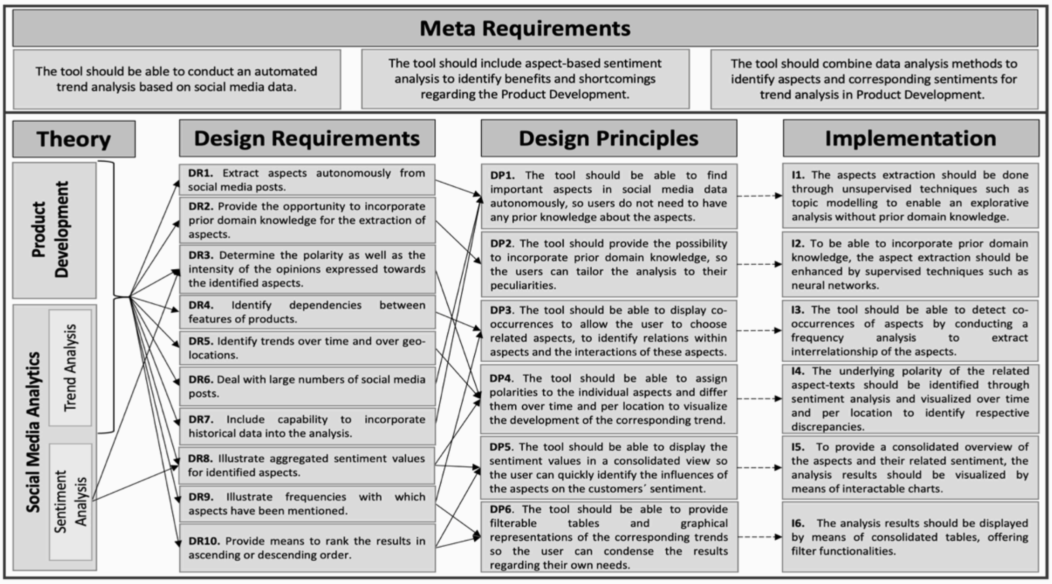


Fig 2. Design of the Artifact

Since no particular machine learning technique is capable of accurately representing all DRs, a combination of techniques was essential. Regarding *DP1*, the autonomous extraction of aspects, unsupervised techniques (e.g., topic modeling) are required to enable an explorative analysis without prior domain knowledge. However, as stated by [29] the potential of totally unsupervised techniques is stymied by their purely unsupervised nature. Thus, semi-supervised techniques have arisen in the past, facilitating an effective way to guide the analysis specific to a user by manipulating the analysis process even without structured prior domain knowledge [30]. Therefore, to take advantages of semi-supervised techniques while maintaining the flexibility of unsupervised ones, the known semi-supervised topic modeling technique GuidedLDA found application as it achieves convincing analysis results [30]. Besides the explorative analysis, the artifact must provide the ability to incorporate prior domain knowledge (*DP2*). Thus, the artifact provides a supervised aspect extraction using deep learning. Specifically, it applies a convolutional neural network (CNN) as proposed by [31], using two types of pre-trained embeddings for the aspect extraction: a general-purpose embedding and a domain-specific embedding, containing domain related information used by the CNN to learn the specific domain peculiarities. Consequently, the analysis can easily be adapted and tailored to the users' own circumstances by changing the underlying domain-embedding, resulting in a highly generic and customizable artifact. With respect to *DP3*, the artifact must depict co-occurrences of the related aspects to detect their interrelationship. As probabilistic topic modeling techniques such as GuidedLDA infer the resulting topics based on various probabilistic distributions, depicting the relations of the underlying topic words (and thus the resulting aspects) [32], the identification of their interrelationships is met through the nature of topic modeling itself. Considering the use of the CNN, the artifact depicts the co-occurrences of the aspects by conducting a frequency analysis. Here, sub-aspects are identified for each extracted (main-)aspect by analyzing their respective occurrence in the context of the related main-aspect, resulting in an n-dimensional occurrence-tree. To determine the tonality of each aspect (*DP4*), the "Valence Aware Dictionary for sEntiment Reasoning" (VADER) [33] technique (a lexicon and rule-based sentiment analysis technique specifically attuned to sentiments expressed in social media) has been implemented. To further consolidate and visualize (*DP5*) the underlying aspects' sentiment as well as its evolution over time, corresponding line-charts are implemented using the well-known python library matplotlib [34]. Finally, to ensure an adequate illustration of the analysis results (*DP6*), the extracted aspects, the corresponding sentiment values and their means, but also specific references to the extracted aspects are displayed using filterable tables and lists as demonstrated in the following section.

5 Demonstration and Discussion of the Artifact

5.1 Demonstration of the Artifact

To examine the tool's ability to identify meaningful and sound trends (including related aspects and their sentiment), we applied it to a real-world dataset. Therefore, we consulted the academic Yelp dataset [24], represented by a subset of real-world reviews and businesses in the US from one of the most popular online communities for crowd-sourced reviews. The dataset, which comprises a total of ~ 8.6 million OCR concerning 160,585 businesses in different fields such as restaurants, cultural sites and sports facilities spans the period from October 13th, 2004 to January 28th, 2021. To demonstrate our tool, we narrowed the data to the multi-faceted field of the food and restaurant industry. Specifically, we narrowed the analysis to the reviews of a fast-food restaurant with multiple franchises in various locations to extract the relevant aspects and associated customer perceptions, resulting in 37,638 reviews. The evaluation of the analysis and the tool's usability will be carried out, as described in section 6, within future work.

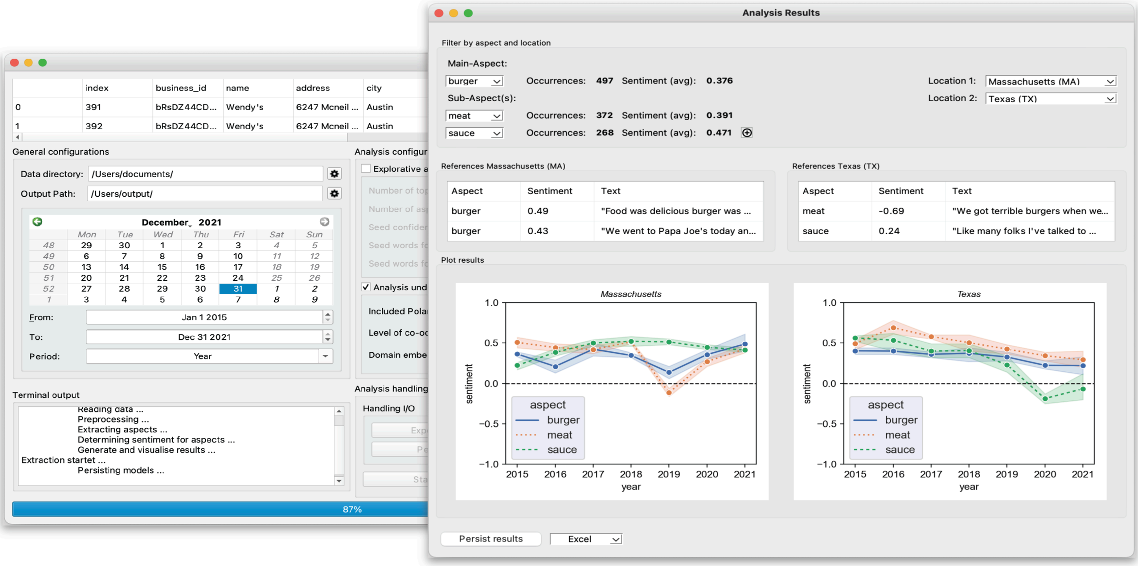


Fig 3. Configuration and Result View of the Artifact

Figure 3 represents the tool's configuration view (left) and results view (right). By the configuration view, the underlying analysis can be customized to the one's own needs. Here, in case of not all periods of time included within the data deemed necessary, a specific period can be individually defined. For the purposes of our demonstration, the analysis was conducted using data spanning the years 2015 to 2021. To further account for the two different analysis settings (with/without the incorporation of prior domain knowledge), both were implemented using dynamic Graphical User Interface (GUI) elements to enable a distinctive configuration. Concerning the demonstration, the analysis under consideration of prior knowledge has been conducted. Therefore, all sentiment levels (*positive, neutral, negative*) are considered. Moreover, the level of co-occurrences has been set to two, resulting in a two-dimensional occurrence-tree. The domain knowledge used was extracted from the renowned dataset of the 2016 SemEval task [35]. By using this appropriate word embedding tailored to the use case of restaurant reviews, the CNN is trained in the domain of the food and restaurant industry.

The result view represents the sentiments of the extracted aspects with respect to the years, resulting in the monitoring of their evolution based on their customer perceptions. Here, both locations refer to the same aspects (main-aspect: *burger*; sub-aspects: *meat, sauce*), facilitating a comparison of their temporal and location-based discrepancies. In Massachusetts (MA), the main-aspect *burger* exhibits slight oscillation across the years but generally remains stable. The worst average customer perception occurs in 2019 and is represented by a slightly positive sentiment score of 0.131. Furthermore, the two sub-aspects *meat* and *sauce* differ strongly in their course in 2019. Here, the course of *meat* collapses drastically (-0.109), while the course of *sauce* remains nearly constant (0.503) compared to the previous year, leading to the assumption that the negative reflections of the *meat* aspect may influence the main-aspect *burger*. Moreover, this assumption is supported by the development of the respective aspects. Here it becomes apparent that the significant improvement of the aspect *meat* also potentially causes an improvement of the main aspect *burger*, supporting the conclusion that our tool is able to identify meaningful and sound trends based on the corresponding aspects and their customer perception. The results provide further convincing evidence that our tool can identify location-based discrepancies in the evolution of trends. Thus, it can be seen that the customer perceptions of the aspects differ across the respective locations, as they overall remain quite constant in MA but consistently decreases in Texas (TX).

Generally, the development of our tool was based on the DPs as shown in Figure 2, which were all implemented as described in section 4. Accordingly, all expectations posed in regard to our artifact were technically realized. In addition, several trends and their evolution could be identified by applying our tool to a representative real-world dataset, validating its functionality. To subsequently evaluate its practical applicability by means of an identification of meaningful and sound trends, an evaluation will be conducted in future work.

5.2 Discussion of the Demonstration

The implementation of both the MRs and the DPs enabled us to design and develop a tool which has provided interesting results. As we establish the opportunity to extract the aspects in either a supervised or an unsupervised way (*DP1, DP2*), we can identify in the first instance aspects and/or product features which are discussed in the social media data under consideration. This allows us to identify those aspects which are most important from the customer´s point of view and, above all, which must be considered in Product Development [3, 8, 20, 21]. Figure 4 shows that the most frequently named aspects in our data are *burger, meat* and *sauce* (497, 372 and 268 occurrences). This means that when reviews about the restaurant are written, customers address mainly these three aspects. In the case of Product Development, the restaurant can start screening these aspects as they are particularly important for the customers. The extraction of aspects is possible on the one hand via a CNN including word embedding, providing the restaurant to incorporate prior domain knowledge (*DP2*) such as aspects about food, drinks, and processes in the restaurant. Therefore, the extraction of the aspects is tailored precisely to the company. But, on the other hand if the restaurant wants to extract aspects without exerting any influence, the tool can also identify them autonomously.

Moreover, the demonstration of the artifact has also shown that the customers have spoken differently about the three aspects. Here, our results show that users spoke about *sauce* more positively (0.471) over the years than about the aspects *burger* (0.376) and *meat* (0.391). However, the results become even more interesting when the related sentiment values are observed over time. Thus, numerous opportunities for improving products can be identified, and certain features can be given greater importance [5, 19]. Figure 4 shows that for MA the values of the individual aspects can change considerably: While in 2015-2018 the sentiment values of *meat* range between 0.40 and 0.52, in 2019 it slipped down into the negative range with -0.12. After this decline, the opinion about *meat* then improved again in 2020-2021. At the same time as the sentiment value of *meat* has fallen, that of the *burger* has fallen too. In this context, it is therefore possible that the negative sentiment about the *sauce* had also influenced the customers' opinion about *burger*. If a company did not have this fine-grained information gained through the aspect-based sentiment analysis and wanted to adjust the product in 2019 based on the negative reviews, it is possible that it would have changed product features which were actually rated positively. Also, in TX, the consideration of the aspects is inevitable as the results show that the *sauce* needs to be changed so that the customer's opinion about it and therefore the opinion of the whole burger can be improved again. Already existing trend analysis tools often include sentiment analysis to show the general tonality about the company or the product over time – without considering that different aspects influence the product's evaluation. However, this leads to a distortion of the results and the benefit for companies is no longer a given.

Furthermore, including geolocations and therefore considering location-based discrepancies can also create significant benefits. Customers' product requirements can differ across different geographical markets. To be competitive in a targeted market, companies need to know which features should be designed in which way to meet the local customers' expectations [17, 36]. This becomes particularly evident when comparing MA and TX (see fig. 4). While we can see in terms of the restaurant in MA that the *burger* exhibits a positively connotated trend over the years, the sentiment over time in TX shows worse values. Especially in 2020 and 2021, the sentiment values of the *burger* differ immensely (MA: 0.35, 0.45; TX: 0.28, 0.22). In both locations, there are different reasons for the burger's better (MA) or worse (TX) rating. While in TX the *sauce* must be adjusted to the taste of the people, this is not necessary in MA due to the steady positive values. Without the inclusion of the geolocations and also the sentiment values, these discrepancies in the results would not have been revealed, which clearly is a benefit in comparison to other existing tools.

6 Conclusion, Contribution and Outlook

Assessing and identifying people's opinions about a particular aspect and its future impact (thus a social media trend), is difficult, especially given the vast amount of social media data. Thus, techniques for analyzing textual social media data, e.g., aspect-based sentiment analysis, topic modeling and neural networks, have gained in importance, as companies need to be aware of customers' expectations regarding products. However, prior literature and existing tools do not incorporate them to identify future trends, do not include external parameters (such as geolocation) and do not cover specific requirements (simultaneous identification of trends with/without prior knowledge) crucial in the field of Product Development (see sections 2.2), although the early identification of new and auspicious ideas and trends means a competitive advantage for companies [19]. Thus, we provide a comprehensive tool by combining several machine learning approaches and transfer these in a highly responsive and platform-independent GUI. Especially with the demonstration of our tool on 37,638 OCR from Yelp we have shown that considering all identified requirements is necessary to analyze trends.

Besides creating value for practitioners, theoretical contributions in the research area of IS are also provided. To acknowledge the importance of different DSR perspectives, we have related our DSR contribution to the category "design artifacts" according to [37], including both: the demonstration of the artifacts practical benefit and design theory contributions [38]. Therefore, by providing a tool for automated trend analysis that can identify aspects that are discussed within social media, we enable companies to gain deep insights into customers' current opinions and future expectations to tailor their products. Hence, to meet these customers' needs, a company has to identify and continuously track product features by incorporating the VoC into internal Product Development processes. Thus, as tracking evolving and changing customer requirements is imperative to meet customers' wishes [36, 39], companies can respond to them quickly and with minimal effort as huge amounts of social media data can be processed with our tool. Compared to already existing trend analysis tools, our tool meets all the specific requirements set out within the extant literature regarding Product Development. In particular, our tool can flexibly match different temporal and geographical parameters to identify aspect-sentiment relations and it provides users the possibility to extract the aspects either with or without prior domain knowledge. Thus, customer perceptions for specific periods or geolocations can be displayed to track fine-grained variabilities. On the one hand, this makes it possible to visualize influences, affecting the sentiment. On the other hand, downward trends in sentiment can be counteracted and upward trends can be strengthened. Hence, this combined with the integration of geolocations can especially support large companies with multiple branches in their efforts to easily perceive location-specific sentiment changes and explicitly react to them. As we have combined different machine learning approaches and designed our tool modularly, companies can adapt the analysis to their specific needs. A further contribution of implementing aspect-based sentiment analysis is that the tool can be supportive in identifying the aspects of the products to be changed to meet the customer's expectations (remove existing aspects of the product, others need to be enhanced). With the realization of (*DP4*)-(*DP6*), we created a tool with which companies can track the overall customer perceptions. In summary, companies can benefit from our comprehensive and modular artifact by analyzing large amounts of data in a way best suited to their circumstances, aiming to analyze trends regarding their aspect-based sentiment values.

Besides our technical contribution (i.e., the artifact), we achieved prescriptive theoretical contributions as a further outcome of our DSR project. Therefore, we formulated and proposed DPs based on the DRs (see section 2.2) derived from current research literature. By applying them in the course of the design and development of the artifact followed by the demonstration, an implicit empirical grounding of the DPs was achieved. Our DPs capture design-related knowledge and can therefore support the development of further IS (design) theories and new artifacts. For designing further (trend analysis) tools in related areas, our DPs can be applied as we have formulated them generally by describing what the artifact should enable users to do and how it should be built. For example, by DP4, the importance of including time, geolocation and/or further external parameters (e.g., customers' characteristics) in a trend analysis tool is highlighted. As these external parameters have a direct impact on the customers' sentiment and therefore on the analysis results (cf. [40]), the alignment to them will lead to a more targeted trend analysis tool. Thus, for researchers that intend to design a trend analysis tool, we suggest considering the influences that are evoked by external factors. So, with the compilation of the DPs, we made a first step towards contributing to nascent design theory. To take a next step towards a more mature design theory, we intend to evaluate our DPs by further evaluating our tool. Therefore, we will first evaluate our tool in a formative and artificial environment (i.e., a laboratory experiment). Here, participants will use the tool to identify relevant aspects and associated customer perceptions in OCR. Subsequently, they will complete a questionnaire to indicate their perceptions of the analysis quality and tool usability. This allows us to identify difficulties and improve our tool (whereby our DPs can be confirmed or adapted) before conducting a more elaborate evaluation in a more natural setting as a further part of the design cycle [41]. This second evaluation will be a field study with a large restaurant chain that plans to integrate a software artifact to support its marketing departments.

However, there are also some limitations to this research: Although we included a large set of investigations, we could identify probably even more requirements the tool should meet in further literature. Nevertheless, the identified DRs are undoubtedly important for Product Development in other areas and other DRs could be identified.

References

1. Kaplan, A.M., Haenlein, M.: Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons 53(1), 59-68 (2010).
2. Kim, Y.K., Lee, D., Lee, J., Lee, J.-H., Straub, D.W.: Influential users in social network services: the contingent value of connecting user status and brokerage. ACM SIGMIS: The Database for Advances in Information Systems 49(1), 13-32 (2018).
3. Tuarob, S., Tucker, C.S.: Quantifying product favorability and extracting notable product features using large scale social media data. JCISE 15(3) (2015).
4. Schouten, K., Frasincar, F.: Survey on aspect-level sentiment analysis. IEEE Transactions on Knowledge and Data Engineering 28(3), 813-830 (2015).
5. Tucker, C.S., Kim, H.: Predicting emerging product design trend by mining publicly available customer review data. In: 18th International Conference on Engineering Design (2011).
6. Mirtalaie, M.A., Hussain, O.K., Chang, E., Hussain, F.K.: Sentiment analysis of specific product's features using product tree for application in new product development. In: International Conference on Intelligent Networking and Collaborative Systems, pp. 82-95 (2017).
7. Mirtalaie, M.A., Hussain, O.K., Chang, E., Hussain, F.K.: Extracting sentiment knowledge from pros/cons product reviews: Discovering features along with the polarity strength of their associated opinions. Expert Systems with Applications 114, 267-288 (2018).
8. Mirtalaie, M.A., Hussain, O.K.: Sentiment aggregation of targeted features by capturing their dependencies: Making sense from customer reviews. IJIM 53 (2020).
9. Vo, A.-D., Nguyen, Q.-P., Ock, C.-Y.: Opinion–aspect relations in cognizing customer feelings via reviews. IEEE Access 6, 5415-5426 (2018).
10. Hicks, A., Comp, S., Horovitz, J., Hovarter, M., Miki, M., Bevan, J.L.: Why people use Yelp.com: An exploration of uses and gratifications. Computers in Human Behavior 28(6), 2274-2279 (2012).
11. Yan, Z., Xing, M., Zhang, D., Ma, B., Wang, T.: A context-dependent sentiment analysis of online product reviews based on dependency relationships. In: 35th ICIS (2014).
12. Ye, H.J., Kankanhalli, A.: User service innovation on mobile phone platforms: Investigating impacts of lead userness, toolkit support, and design autonomy. MIS Quarterly 42(1), 165-188 (2018).
13. Axtell, C.M., Holman, D.J., Unsworth, K.L., Wall, T.D., Waterson, P.E., Harrington, E.: Shopfloor innovation: Facilitating the suggestion and implementation of ideas. Journal of Occupational and Organizational Psychology 73(3), 265-285 (2000).
14. Blei, D.M., Ng, A.Y., Jordan, M.: Latent dirichlet allocation. JMLR 3, 993-1022 (2003).
15. Liu, B.: Sentiment analysis and opinion mining. Synth. Lect. Hum. Lang. Technol. 5(1) (2012).

16. Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., Cleven, A.: Standing on the shoulders of giants. Communications of the AIS 37(1) (2015).
17. Han, Y., Moghaddam, M.: Analysis of sentiment expressions for user-centered design. Expert Systems with Applications 171 (2021).
18. Cheng, L.-C., Chen, K., Lee, M.-C., Li, K.-M.: User-Defined SWOT analysis – A change mining perspective on user-generated content. Information Processing & Management 58(5) (2021).
19. Jeong, B., Yoon, J., Lee, J.-M.: Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis. IJIM 48, 280-290 (2019).
20. Zhang, J., Zhang, A., Liu, D., Bian, Y.: Customer preferences extraction for air purifiers based on fine-grained sentiment analysis of online reviews. Knowledge-Based Systems 228 (2021).
21. Yu, J., Zha, Z.-J., Wang, M., Chua, T.-S.: Aspect ranking: identifying important product aspects from online consumer reviews. In: 49th annual meeting of the ACL, pp. 1496-1505 (2011).
22. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Quarterly, 75-105 (2004).
23. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for information systems research. JMIS 24, 45-77 (2007).
24. Yelp open dataset, https://www.yelp.com/dataset, Accessed on 15.12.2021.
25. Hevner, A.: A three cycle view of design science research. SJIS 19(2) (2007).
26. Gregor, S., Jones, D.: The anatomy of a design theory. Journal of the AIS 8(5) (2007).
27. Gregor, S., Chandra, L., Seidel, S.: Research perspectives: The anatomy of a design principle. Journal of the AIS 21(6) (2020).
28. Chandra, L., Seidel, S., Gregor, S.: Prescriptive knowledge in IS research: Conceptualizing design principles in terms of materiality, action, and boundary conditions. In: 48th HICSS (2015).
29. Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J.L., Blei, D.M.: Reading tea leaves: How humans interpret topic models. In: Advances in neural information processing systems, pp. 288-296 (2009).
30. Jagarlamudi, J., Daumé III, H., Udupa, R.: Incorporating lexical priors into topic models. In: 13th Conference of the European Chapter of the ACL, pp. 204-213 (2012).
31. Xu, H., Liu, B., Shu, L., Philip, S.: Double Embeddings and CNN-based Sequence Labeling for Aspect Extraction. In: 56th Annual Meeting of the ACL, pp. 592-598 (2018).
32. Crain, S.P., Zhou, K., Yang, S.-H., Zha, H.: Dimensionality reduction and topic modeling: From latent semantic indexing to latent dirichlet allocation and beyond. Mining text data, 129-161 (2012).
33. Hutto, C., Gilbert, E.: Vader: A parsimonious rule-based model for sentiment analysis of social media text. In: International AAAI Conference on Web and Social Media (2014).
34. Hunter, J.: Matplotlib: A 2D graphics environment. Comput. Sci. Eng 9(3), 90-95 (2007).
35. Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., De Clercq, O.: Semeval-2016 task 5: Aspect based sentiment analysis. In: International workshop on semantic evaluation, pp. 19-30 (2016).
36. Lozano, M.G., Schreiber, J., Brynielsson, J.: Tracking geographical locations using a geo-aware topic model for analyzing social media data. Decision Support Systems 99, 18-29 (2017).
37. Baskerville, R., Baiyere, A., Gregor, S., Hevner, A., Rossi, M.: Design science research contributions: Finding a balance between artifact and theory. Journal of the AIS 19(5) (2018).
38. Gregor, S., Hevner, A.: Positioning and presenting design science research for maximum impact. MIS Quarterly, 337-355 (2013).
39. Hong, L., Ahmed, A., Gurumurthy, S., Smola, A.J., Tsioutsiouliklis, K.: Discovering geographical topics in the twitter stream. In: 21st International Conference on WWW, pp. 769-778 (2012).
40. Konadl, D., Wörner, J., Leist, S.: Identifying Sentiment Influences Provoked by Context Factors– Results from a Data Analytics Procedure Performed on Tweets. In: 54th HICSS (2021).
41. Venable, J., Pries-Heje, J., Baskerville, R.: FEDS: a Framework for Evaluation in Design Science Research. European Journal of Information Systems 25, 77-89 (2016).

2.5 Automated Identification of Different Lead Users Regarding the Innovation Process

Adressierte Zielsetzungen	Ableitung eines Anforderungsprofils an die methodische Konzeption eines datenbasierten Identifikationsverfahrens von Lead Usern. Hierzu werden zentrale Lead User Eigenschaften (z. B. Trendsetter, Netzwerkposition, produktbezogenes Wissen) in der Literatur identifiziert, systematisch modelliert und softwaregestützt operationalisiert. Ziel ist es, ein automatisiertes Analyseverfahren zu konzipieren, welches diese Lead User Merkmale in digitalen Kommunikationsdaten automatisiert erfasst und so für die softwaregestützte Lead Usern Identifikation nutzbar machen kann.	***ZS7***
	Entwicklung, Demonstration und empirische Evaluation eines skalierbaren Analysewerkzeugs zur automatisierten Identifikation von Lead Usern auf Basis des konzipierten Analyseverfahrens. Dabei soll ein multimethodisches Analysewerkzeug durch die Integration etablierte Data Science Verfahren (z. B. Topic Modelling zur Trendanalyse, Sentimentanalyse zur Stimmungsanalyse, SNA zur Abbildung der Netzwerkstrukturen) entwickelt werden, welches eine automatisierte, merkmalbasierte Identifikation von Lead User ermöglicht.	***ZS8***
Abgeleitete Forschungsfragen	Welche charakteristischen Merkmale weisen Lead User in Online-Communities auf, und wie lassen sich diese Merkmale datenbasiert operationalisieren?	***FF7***
	Wie kann die Identifikation von Lead Usern durch ein softwaregestütztes Analyseverfahren unterstützt werden, und welche technischen Lösungsansätze sind hierfür geeignet?	***FF8***
Genutzte Forschungsmethode	Umsetzung gemäß dem DSR-Prozess nach Peffers et al. (2007) mit Fokus auf die Entwicklung, Demonstration und Evaluation eines funktionalen Analyseartefakts zur Lead-User-Identifikation. Orientierung an den drei Zyklen von Hevner et al. (2004): • **Design Cycle**: Konzeption, technische Umsetzung und Evaluation des Analyseansatzes mit Fokus auf die datenbasierte Operationalisierung relevanter Lead User Merkmale. Aufbauend auf der ersten Evaluationsepisode mit den identifizierten Lead Usern sowie einem Innovationsexperten wurde das Artefakt verfeinert. • **Relevance Cycle**: Orientierung an praxisrelevanten Anforderungen aus den Anwendungsfeldern des Produktentwicklung und der Innovationsprozesse sowie Rückkopplung des entwickelten Artefakts durch dessen Evaluation mit Experten aus der Praxis. • **Rigor Cycle**: Theoretische Fundierung des Artefakts anhand der Berücksichtigung bestehender Theorien wie der Lead User Theorie sowie der Innovationstheorie. Das entwickelte Analyseverfahren nutzt ein modulares Methodenset aus Sentimentanalyse, SNA, Frequenzanalyse und Topic Modelling, um typische Lead User Merkmale systematisch und automatisiert zu identifizieren.	

Kernergebnisse (Überblick)	Durch die Konzeption und Entwicklung des Analysewerkzeugs wurden mit Blick auf FF7 und FF8 mehreren Kernergebnisse geschaffen: • Identifikation und datenbasierte Operationalisierung von sechs zentralen Lead User Merkmalen: *Trend Leadership*, *Dissatisfaction*, *Enjoyment*, *High Activity Level, High Product-Related Knowledge*, *Opinion Leadership* • Differenzierte Betrachtung des Lead User Konzepts entlang des Innovations-prozesses: Es wurden zwei unterschiedliche Typen von Lead Usern für die Phasen *Ideengenerierung* und *Produktentwicklung* abgeleitet. • Entwicklung eines automatisierten Lead User Identifikationsverfahrens, das alle sechs Lead User Merkmale automatisiert in digitalen UGC analysieren kann. • Demonstration des entwickelten Analyseverfahrens anhand realer Daten aus einer innovationsgetriebenen Online-Community: Durch die Anwendung des Verfahren konnte die Segmentierung unterschiedlicher Lead User Typen je nach Innovationsphase demonstriert werden. • Zweigleisige Evaluation des Analyseverfahrens: Einerseits wurden die Ergebnisse durch asynchrone Nutzerinterviews mit identifizierten Lead Usern sowie andererseits mit einem Experteninterview mit einem Innovationsmanager evaluiert. • Es werden unterschiedliche Beiträge zur datengetriebenen Entscheidungsunterstützung im Innovationsmanagement und zur Weiterentwicklung der Lead User Theorie geleistet.
Publikationsort	Das Paper wurde 2022 im **Electronic Markets Journal** (*EM*) veröffentlicht und ist als peer-reviewtes Journalpapier gelistet.
Ranking VHB JQ 4	**A**
Autoren und Anteile	Isabel Schmid 45% **Janik Wörner** **45%** Susanne Leist 10%
DOI	https://doi.org/10.1007/s12525-022-00568-6

Tabelle 19. Fact Sheet zur 5. Veröffentlichung.

Electronic Markets
https://doi.org/10.1007/s12525-022-00568-6

RESEARCH PAPER

Automated identification of different lead users regarding the innovation process

Isabel Schmid[1] · Janik Wörner[1] · Susanne Leist[1]

Received: 12 May 2021 / Accepted: 18 May 2022

Abstract
Lead users are often established in an organizational innovation process to attenuate the difficulties a company faces, such as high costs or the obscurity of customers' needs. But to benefit from these lead users a major challenge is to characterize and identify them especially in the fast-moving world of social media. Therefore, we aim to design a tool to identify lead users automatically for the two innovation phases ("Idea generation" and "Development") by combining different approaches such as social network analysis, topic modeling and sentiment analysis. Thus, we consulted the design science approach and applied our artifact to 11,481 contributions of an online digital platform. The technical realization of the six different characteristics and their respective weighting according to the different phases of the innovation process resulted in different lead users and showed the necessity of distinguishing between them. Our results were evaluated and confirmed by the identified lead users and an expert. Hence, our investigation contributes to both practice and theory (kernel theories and design theory) alike.

Keywords Lead user · Innovation process · Digital platforms

Introduction

In today's dynamic environment, innovation is central for the competitiveness of companies. To create competitive advantages a profound understanding of the sources of innovation is necessary (Von Hippel, 2007). New and innovative ideas can be released by both internal and external sources. According to Innovation Theory generating a new idea by consulting the corporate research and development (R&D) department can be seen as the internal way of innovation creation to advance a company's technology (Freeman & Soete, 1997; Marxt & Hacklin, 2005). The external way of innovation creation, however, consults innovation ideas initialized from parties outside the company such as customers, suppliers, universities or individuals (West & Bogers, 2014). Integrating an individual in the innovation process can take place in different ways e.g. in terms of co-creation (Ramaswamy, 2010), crowdsourcing (Poetz & Schreier, 2012) or open innovation (Martínez-Torres, 2014).

In current research literature, there are a lot of documented examples of successful collaboration with external parties in the innovation process. General Electrics, for example, has banded together with a number of venture capital companies to arrange the "Ecomagination Challenge," a $200 million fund for identifying and investing in innovative ideas and business models regarding renewable energy, grid efficiency and energy consumption. They created a platform where different external stakeholders submitted their ideas and in total they attracted more than 5000 ideas (King & Lakhani, 2013). Moreover, Lilien et al. (2002) have shown that in the company 3 M the inclusion of external individuals in the innovation process results in ideas that have greater commercial potential than ideas without the inclusion of external persons. Whereas TopCoder arranged a two-sided innovation platform to bring software programmers and companies together in order to fix IT-related problems (Lakhani et al., 2010).

Responsible Editor: Markus Bick

Isabel Schmid
Isabel.Schmid@ur.de

Janik Wörner
Janik.Woerner@ur.de

Susanne Leist
Susanne.Leist@ur.de

[1] Business Informatics, University of Regensburg, Universitätsstraße 31, 93053 Regensburg, Germany

Published online: 27 July 2022

Therefore, new ways of communication such as social media or online communities, a form of digital platforms, provide companies the possibility to access a huge number of users for a new way to innovate (Brem & Bilgram, 2015; Gawer & Cusumano, 2014). Hence, a company can use social media like a magnet to capture customer feedback, improve market research and facilitate innovation (Gallaugher & Ransbotham, 2010). Thus, on the one hand the company benefits from the collaboration with a user as it may result in ideas for extending product varieties, in entirely new products and/or in modifications to existing ones (Al-Zu'bi and Tsinopoulos, 2012). On the other hand, the user also benefits strongly from the innovative products as these are tailored to their own needs (Tuarob & Tucker, 2014; Von Hippel, 1986). In order to realize these benefits, some approaches consult the opinions and suggestions of a crowd of people (cf. open innovation). Although a company thereby receives a lot of input, the ideas are often futile, as they are either not innovative, not feasible or are formulated too superficially. Furthermore, the processing and evaluation of the ideas is very time-consuming as the example of Fiat Mio shows (Saldanha & Pozzebon, 2015). The Fiat Mio team aimed to create a concept car by composing a collaborative website where they received 21,000 ideas and 45,000 comments. The whole process took 15 months and a lot of resources – both human and capital – to screen all the posted ideas and suggestions. To avoid such an intricate and expensive process it is more constructive to concentrate on single persons who are able, due to their individual characteristics, to support a company's innovation process – so called lead users. Hence, a lead user is a user who identifies needs and trends in the market months or years before other people do and who benefit significantly by obtaining a solution to those needs according to the Lead User Theory (Hienerth & Lettl, 2017; Schaarschmidt et al., 2019; Von Hippel, 1986).

These lead users can attenuate the difficulties a company faces during the innovation process, such as high costs or the unsteadiness of customers' acceptance of a company's innovation (Ye & Kankanhalli, 2018). Therefore, a lead user is often established at the beginning and at the end of an innovation process. In the early phases of this process lead users formulate their needs which can result in new ideas. At the end of the process, a lead user can be incorporated to test the product's functionality and durability (Al-Zu'bi and Tsinopoulos, 2012). But in order to benefit from lead users, one major challenge in both research and practice is to characterize and identify them (Ernst et al., 2013). Amongst other factors, the tremendous amount of online community data is responsible for the fact that the identification of lead users is the most difficult and time-consuming aspect within the lead user method (Brem & Bilgram, 2015). In current research literature there are a lot of different lead user identification approaches, but these investigations only covered a limited point of view as they either focus on only one lead user characteristic such as the high level of activity (Martínez-Torres, 2014) or include a very small amount of data (Hau & Kang, 2016). Moreover, various investigations base their approach on observations or online questionnaires (Brandtzaeg et al., 2016; Hung et al., 2011; Tuunanen et al., 2011), resulting in rather low sample efficiency and high costs. Additionally, lead user characteristics are derived based on self-assessments, which may bias the results due to subjective assessments (Hienerth & Lettl, 2017). Finally, most of the aforementioned identification methods are time-consuming, which contradicts to the trend specific short-term construct of lead users (Hienerth & Lettl, 2017).

With this work at hand, we address these aforementioned problems by following the Design Science (DS) approach suggesting an automated and effective approach for the lead user identification. Therefore, this research seeks to answer the following research questions informed by the Lead User- and Innovation Theory and is thus built on a descriptive knowledge base:

RQ1: What different characteristics does a lead user in an online community exhibit?
RQ2: How can the identification of a lead user be supported by a software artifact?

With this investigation we seek to cover the major characteristics of a lead user found in the literature. We aim to identify this user type automatically by means of combining different analysis approaches such as social network analysis (SNA), topic modeling and sentiment analysis. By screening the current research literature, it became obvious that different lead users can be identified in different phases of the innovation process. Thus, we further aim to show the differences between the lead users in these two innovation phases and therefore the legitimacy of the differentiation with respect to the phase-specific characteristics. To cope with this, our goal is to develop a software tool for the automated lead user identification, enabling e.g. the identification of different lead users regarding the two phases of the innovation process, the mapping of all prior identified characteristics in an automated manner and the processing of large amounts of online community data containing relevant information regarding the characteristics of a lead user. To show the applicability of the designed identification approach we apply our artifact to real-world data of an online community for kitesurfing. Kitesurfing is a water sport in which the athlete surfs through the water by pulling a large, controllable kite while standing on a special board. It is a popular example for lead user innovation as this sport was initiated by surfers who – driven by the desire to jump higher and further – experimented with the combination of a surfboard and sails from hang gliding. Moreover, this area of application is further suitable as these individuals in this area are quite active as innovators and kitesurfing is comprised of a young community, essentially all serious participants are active members in some kind of online community (Franke et al., 2006; Von Hippel, 2005; Wagner & Piller, 2011).

In addition to the instantiated artifact and the results obtained from the demonstration we want to highlight contributions to both practice and theory. Thereby, we want to acknowledge both perspectives of contribution in a design science research project – the artifact school of thought (cf. Hevner et al., 2004) and the design school of thought (cf. Gregor & Jones, 2007). Furthermore, different knowledge contributions will be taken into account to contribute to theory—the descriptive knowledge base in the course of kernel theories and the prescriptive knowledge base in the course of design theory (by deriving the design principles and evaluating them in the course of applying the artifact). To achieve this goal, our investigation follows the third research question:

> **RQ3:** What different contributions for theory and practice can be derived from our Design Science project?

The remainder of this paper is as follows: the following section "Conceptual basics" provides a theoretical background by introducing important definitions and related work regarding lead users and their characteristics. Next, the procedure of the research following the DS approach (Hevner et al., 2004; Peffers et al., 2007) is described in the subsequent section. The section "Design and development" particularly deals with the technical realization and derivation of the design principles to enable the automated identification of lead users regarding the different phases of the innovation process. The following section "Demonstration, evaluation and discussion" shows the application of the demonstrated approach on approximately 12,000 online community data and presents as well as discuss the resulting outcomes, which are additionally evaluated by an interview with our cooperating partner and interviews with the identified lead users. The paper concludes with the contribution for practice and theory and a conclusion.

Conceptual basics

Online communities

Social media are defined as internet-based applications that offer opportunities for interactive and dynamic communication, collaboration and participation (Kaplan & Haenlein, 2010; Obar & Wildman, 2015). Thus, different types of social media can be identified: whereas social network sites (SNS) especially enable users to connect with other people by creating personal profiles, online communities, as a further hyponym of social media, focus on sharing content between users (Kaplan & Haenlein, 2010). Therefore, online communities can be defined as internet-based platforms for communicating and exchanging content among users who are interested in a given product or technology (Autio et al., 2013; Breitsohl et al., 2018; Preece & Maloney-Krichmar, 2003). Online communities as digital, multisided platforms benefit mainly from so called "network effects": the more users the platform access, the more valuable the platform becomes for both users and companies (de Reuver et al., 2018; Gawer & Cusumano, 2014).

For online communities, which have become increasingly popular due to the rise of social media, various characteristics were defined early and are still relevant today: such as (1) users follow a shared goal, interest or need (Breitsohl et al., 2018; Preece & Maloney-Krichmar, 2003; Tuunanen et al., 2011) (2) users participate actively, interact with each other and build up ties (Dahlander & Frederiksen, 2012; Fisher, 2019; Füller et al., 2007; Preece & Maloney-Krichmar, 2003) and (3) users have access to shared resources (like knowledge or information) (Breitsohl et al., 2015, 2018; Preece & Maloney-Krichmar, 2003). The communication in online communities is established around discussion threads. This means that users initialize new threats in order to start a new discussion, issue or call for advice (Autio et al., 2013). Thus, as online communities often cover one main topic (e.g. mountain biking or kitesurfing), this subtype of social media focus more on connecting people with the same interests than SNS.

Moreover, companies can also benefit from the broad dissemination of digital, multisided platforms in terms of online communities because of social media's reach (via social media a lot of people can be reached) and richness (social media platforms provide various types of information) (Shang et al., 2017). This kind of communication medium gives a company the opportunity to communicate and engage with (potential) customer communities (Fisher, 2019). Thus, as users discuss their experiences, news, improvements or ideas, companies become aware of the customers' needs (Autio et al., 2013; Kaplan & Haenlein, 2010; Tuunanen et al., 2011). Especially in brand communities, excited users group together and share brand-related content (Breitsohl et al., 2015, 2018). However, a company can not only benefit from online communities in terms of nurturing brand commitment and the awareness of customers' needs but also from the fact that these digital platforms can serve as a source of innovation (Dahlander & Frederiksen, 2012; Fisher, 2019). In terms of the discussions taking place in online communities, users also provide new ideas, offer solutions for problems, work out details and test new product ideas (Füller et al., 2007). Thus, these platform-based new product developments can be consulted to increase product variety, meet diverse customer requirements and business needs (Gawer & Cusumano, 2014).

All in all, online communities allow communication and interaction between users and companies in different ways. Gallaugher and Ransbotham (2010) take this on in their 3-M framework including the three different customer dialog approaches: Megaphone (firm-initiated dialog), Monitor (customer-to-customer dialog) and Magnet (customer-initiated dialog). From this follows that a company can use digital platforms in terms of online communities not only as a megaphone (in terms of spreading marketing messages) but also especially as a monitor to get to know customers' needs. Thus, by monitoring customer-to-customer dialogs companies can gain insights in customers' opinions or market intelligence (Gallaugher & Ransbotham, 2010). Furthermore, a company can also use online communities as a magnet, the customer-initiated dialog, to capture customer feedback, improve market research and facilitate innovation (Dahlander & Frederiksen, 2012; Fisher, 2019; Gallaugher & Ransbotham, 2010).

Lead user innovation

Innovation is a central construct for organizational competitiveness and effectiveness (Wolfe, 1994). It can be seen as an essential process for driving economic growth (Chen et al., 2018). In general, innovation can be defined as a process that includes the generation, adoption and implementation of new ideas, practices, or artifacts in organizations (Axtell et al., 2000; Ye & Kankanhalli, 2018). So, innovation is a complex issue that comprises many theories, each with a different focus (Wolfe, 1994). In addition, there are also many innovation process models that describe how innovations can be implemented step by step. Cooper (1996), for example, established the so-called stage-gate model and divided the innovation process into the following five different phases (stages): 1. preliminary investigation, 2. detailed investigation, 3. development, 4. testing and validation, 5. full production and market launch. The stage-gate model describes a conceptual and operational model for moving new product projects from idea to launch. Other widely spread innovation process models (cf. Crawford, 1994; Herstatt, 1999; Ulrich & Eppinger, 1995) are similar to the process of Cooper's (1996) approach.

By scanning further research literature about innovation and keeping the process models in mind, exemplarily the stage-gate model, it became apparent that most innovation approaches identify two comprising key phases: (A) the idea generation which means the "awareness" of an innovation and incorporates therefore the preliminary and detailed investigation phases of the stage-gate model, and (B) the development of an innovation incorporating the development as well as the testing and validation phases of the stage model (Amabile, 1988; Axtell et al., 2000; Unsworth et al., 2000). We follow this approach and concentrate on the two phases "Idea generation" and "Development". Consequently, we excluded in our investigation e.g. the step "market launch" as here another user type – the influencer – can be applied to support this phase optimally (Schmid, 2020). According to the definition of an influencer, this user type is applied by companies for disseminating information, for spreading marketing messages, and for changing the opinions and even the purchase decisions of people in its direct and indirect environments (Schmid, 2020). If an influencer would be involved in upstream value creation stages respectively innovation related phases (such as "Idea Generation" or "Development"), the user would feature characteristics of lead users (e.g., ahead of trends, etc.) and would therefore be – in addition of being an influencer – also a lead user. So, as our overall goal of this paper is to identify users who can support a company in their innovation process, we focus the characteristics of a lead user who can also appear in other phases as other type of user.

In the last decades, it has become conventional that consumers or users themselves support one or even both phases of the illustrated innovation process. Hence, this user innovation can be conveyed from the shift of traditional firm-centered innovation to user-centered innovation (Von Hippel, 2005). Prior research highlights that users, rather than firms, are frequently the ones who initiate new product ideas and product developments (Dong & Wu, 2015). Thus, user innovation can be defined as innovative activities undertaken by users who are the source of innovative ideas and who engage actively in developing and modifying products also to meet their own needs (Zheng & Zhou, 2017). These users can invent, prototype, and test new innovations (Roy, 2018). The advantages of user innovations can be mainly attributed above all to the nature of digital innovation platforms.

From a company's point of view, the most important driver for user innovation is to overcome information stickiness. Innovation requires both information about the problem and problem-solving information or, in other words, need-related and solution-related knowledge (Von Hippel, 1994, 2005). Often the information about user's needs and the information used in problem solving is costly to acquire and therefore "sticky" (Piller, 2006; Von Hippel, 1994). Hence, the acquisition as well as transferring costs of the information that is decisive for initiating innovation have tremendous influence on where innovation is created (Idota, 2019). Therefore, as users with highly sticky information can create innovation, organizations should include them in their innovation process to get to know the user's needs, to solve (product) problems and to reduce R&D costs. Thus, User Innovation Theory postulates i.a. that "*innovation among users tends to be concentrated on lead users (people with* *high lead userness) of those products or services*" (Ye & Kankanhalli, 2018). This means that those users who carry out user innovation are so-called lead users (Von Hippel, 1986).

In current research literature there is no consistent definition of a lead user, but the Lead User Theory of Von Hippel (1986) is often used as a starting point for defining them: "*Lead users face needs that will be general in a marketplace – but face them months or years before the bulk of that marketplace encounters them, and – Lead users are positioned to benefit significantly by obtaining a solution to those needs*." (Von Hippel, 1986, p.796). Thus Lead User Theory states that lead users can be used as a source of innovative and commercially attractive ideas about products and services and are characterized by two distinct characteristics: ahead of trend and high benefits from innovation (Hau & Kang, 2016; Von Hippel, 1986, 2005). Hence, lead users are able to invent, prototype and field test innovations (Roy, 2018). Therefore, they can either be applied for the entire innovation process (cf. Ye & Kankanhalli, 2018) or can be applied for only one part of the innovation – either need or solution related tasks (cf. Von Hippel & Katz, 2002). As lead users can constrict the challenges a company faces during the innovation process and as they are simultaneously able to disclose new ideas, lead users can be seen as a valuable resource for companies in terms of different phases of the innovation process (Al-Zu'bi and Tsinopoulos, 2012; Ye & Kankanhalli, 2018). Several studies have shown that their involvement in this process, especially in the early and late phases, can enhance product success (Brem et al., 2018; Schreier et al., 2007). Hence, an intensive collaboration with lead users can increase the product variety as well as the rapidness of a new product development process (Al-Zu'bi and Tsinopoulos, 2012).

Furthermore, as the lead user is the only user type who can be applied in terms of user innovation and therefore support a company in their innovation process, we focus on this type of user. To benefit from a lead user, one major challenge in both research and practice is to characterize and identify him/her (Ernst et al., 2013) – the second step in Von Hippel (1986) 4-step process of utilizing lead users (1. identification of the trend, 2. identification of a lead user, 3. analyze lead user need data, 4. project lead user data onto the general market). Here, the identification of adequate lead users is mostly accompanied by horrendous monetary, time and human resources (e.g. Brandtzaeg et al., 2016; Hung et al., 2011; Tuunanen et al., 2011). Therefore, to reduce the devoted resources regarding the characterization and identification of lead users, we aim to design a software tool, enabling the automated identification of lead users based on their descriptive characteristics. This tool is intended to automate the identification process which is described in current research literature as the most difficult and time-consuming aspect within the Lead User Theory (Brem & Bilgram, 2015). However, to automate the identification process we need to characterize the lead user in detail first.

Characterization of lead users in online communities

In order to characterize lead users in online communities, we conducted an extensive literature search. This resulted in 18 investigations (see Table 1) that focus on the characterization of lead users in terms of online communities. A minority of the 18 investigations (3 out of 18) examines lead users within SNS rather than explicitly in online communities. Nevertheless, since these investigations specify SNS with the same characteristics as online communities and since the authors of these three investigations also base their research primarily on identifying lead users in online communities, these investigations are also included here. Numerous research papers that are not related to the online area were excluded as well as those that do not focus on the identification or characterization process.

However, the description of lead users by Von Hippel (1986) in the course of the description of the Lead User Theory (see section: "Lead user innovation"), were used as a starting point for the characterization of lead user as almost every investigation mentioned in the following relate to these two major lead user characteristics: (1) trend leadership/being ahead of trend and (2) the high expected benefit from innovative solutions, meaning that lead users benefit strongly from adopting new products tailored to their needs (Brem et al., 2018). Prior research about the identification and characterization of lead users in the social media sphere have shown that these two characteristics of the basic model of the Lead User Theory remain valid (Pajo et al., 2017; Schaarschmidt et al., 2019; Tuarob & Tucker, 2014; Ye & Kankanhalli, 2018).

(1) Trend leadership incorporates the degree to which a user can be seen as a leading edge with respect to a certain trend (Franke & von Hippel, 2003). That means lead users have prevailing information and expertise about major trends of products and services as well as future demands for them in the market (Hau & Kang, 2016; Tuarob & Tucker, 2014). Hence, a lead user is a consumer of a product that identifies problems and unmet needs that will later be experienced by the public. This means that the innovations lead users strive for often do not exist on the market (Franke & von Hippel, 2003). Therefore, as lead users recognize what the mass desires months or years before others do, they are ahead of trends (Brandtzaeg et al., 2016; Pajo et al.,

Table 1 Prior research on lead users

		Trend leadership	Sentiment		High Activity Level	High product related knowledge	Opinion leadership
			Dissatis-faction	Enjoy-ment			
Idea generation	(Belz and Baumbach, 2010)	X	X		X	X	X
	(Chen et al., 2019)	X		X	X	X	
	(Miao & Zhang, 2017)				X		
	(Marchi et al., 2011)	X			X	X	
	(Martínez-Torres, 2014)				X		
	(Hau & Kang, 2016)	X	X			X	
	(Hung et al., 2011)	X	X		X		
	(Pajo et al., 2017)	X	X		X	X	X
	(Tuarob & Tucker, 2014)	X				X	
Development	(Brandtzaeg et al., 2016)	X					
	(Chen et al., 2019)	X		X	X	X	
	(Franke and Hippel, 2003)	X				X	
	(Franke et al., 2006)	X				X	
	(Li & Tang, 2016)	X				X	
	(Miao and Zhang, 2017)				X	X	
	(Mahr & Lievens, 2012)	X				X	
	(Pajo et al., 2014)	X	X		X	X	X
	(Pajo et al., 2017)	X	X		X	X	X
	(Schaarschmidt et al., 2019)	X	X			X	
	(Schreier et al., 2007)	X				X	X
	(Tuunanen et al., 2011)	X			X		

2017; Pajo et al., 2014; Tuarob & Tucker, 2014; Ye & Kankanhalli, 2018).

(2) In addition, the characteristic high expected benefit is broken down into further sub-characteristics in the current research literature to make this characteristic more tangible and (especially against the background of the large amount of social media data) more measurable (Ye & Kankanhalli, 2018). We agree with this approach and focus on these sub-characteristics (e.g. dissatisfaction) when defining and characterizing a lead user in the following section. Thus, a lead user does not only come up with attractive innovations to help others but they also benefit strongly from the adoption of new or improved products (cf. high expected benefit) (Schreier et al., 2007; Von Hippel, 1986). Often it is not the financial benefit that motivates a lead user to innovate, but e.g. the chance to execute their sports more effectively. By undertaking their sports, users become aware of the mismatch of expected and experienced performance of the products which can lead to dissatisfaction (Lüthje, 2004). Therefore, the discrepancy between the users' needs and the solutions available on the market leads to dissatisfaction. Given the nature of the kitesurfing or mountain bike community and their genesis, it was through the dissatisfaction of the athletes that the sport emerged, which leads to this proxy measure for users' expected benefit (Belz & Baumbach, 2010; Pajo et al., 2014, 2017; Schaarschmidt et al., 2019). The unmet needs and the relating dissatisfaction of a user lead to the expectation to benefit significantly from an innovative solution (Pajo et al., 2017; Ye & Kankanhalli, 2018). Although this characteristic is prevalent in the current research literature, a discrepancy can be determined. Chen et al. (2019) e.g. introduce a new model (ITF model) for determining a user's index of innovativeness including the three dimensions of involvement, thinking and feeling. The last dimension "feeling" is related to the extent of a user's enjoyment, exploration and creativity, which in turn enables the users to make full use of their potential innovativeness. Therefore, the authors refer to the emotional attachement and the preference for the product by users and therefore state that a lead user exhibits positive sentiment rather than negative sentiment such as dissatisfaction (Chen et al., 2019).

Additionally, with regard to the topic of lead users in online communities, multiple other characteristics, beside the abovementioned, can be identified e.g. the high level of activity with regards to the involvement. According to various investigations (Martínez-Torres, 2014; Miao & Zhang, 2017; Pajo et al., 2017) lead users are more active in a community than the rest of their members. Moreover, the authors Hung et al. (2011) emphazise the lead user's creative and active participation in order to facilitate effective innovations and to encourage innovation communication. The more a lead user's participation level is, the more they get involved in the community. High involvement of users usually implies that there will be high effort made by the users in interacting with the product (Chen et al., 2019). This active involvement is necessary to disclose the sticky information that resides in a lead user. This information can only be valuable in terms of innovations when a user exhibits a high product related knowledge (Franke et al., 2006; Li & Tang, 2016). According to Schaarschmidt et al. (2019) a lead user differs most "*from `typical` consumers as they also have considerable levels of solution knowledge*" (Schaarschmidt et al., 2019, S. 4). This kind of product-related knowledge contains expertise about the product architecture, the used materials and the technologies as this is the basis for creating new ideas (Franke et al., 2006; Schreier et al., 2007). Only by having high product related knowledge, a lead user is able to formulate the needs into concrete innovation ideas and/or concrete specifications of new products (Chen et al., 2019; Marchi et al., 2011; Pajo et al., 2017; Pajo et al., 2014; Tuarob & Tucker, 2014).

As lead users not only have ideas for realizing innovation but also diffusing them, a lead user can also be described by the characteristic "opinion leadership". Opinion leadership is the ability to enable the flow of information and especially to diffuse it. Strong social relationships and a high level of engagement are premises for a functioning exchange of ideas and innovation (Pajo et al., 2014, 2017).

However, lead users can be defined not only in terms of these different characteristics but also – as already mentioned in the section "Lead user innovation" – in terms of the different phases of the innovation process where a lead user can be applied. Therefore, to support the identification of lead users regarding these different innovation phases, we further allocate the aforementioned characteristics to the respective innovation phase.

- Lead users can be applied in the phase "Idea generation" of the innovation process and are therefore more problem-oriented (Belz & Baumbach, 2010; Miao & Zhang, 2017). Lead users in this phase of the innovation process describe problems and unmet needs with the already existing products (cf. dissatisfaction) (Belz & Baumbach, 2010; Hau & Kang, 2016). Furthermore they bring forward new ideas which might help to fix the problem described before (cf. trend leadership). These ideas tend to be unique and can possibly be useful for the development of the next generation (Tuarob & Tucker, 2014). In online communities lead users can share their innovative ideas and other community members can comment and evaluate these ideas. The users offer suggestions on the one hand about modifications and adaptions regarding product attributes, positioning, etc.. On the other hand, lead users formulate innovative ideas about completely new products which can be realized afterwards by a company's R&D team (Marchi et al., 2011; Martínez-Torres, 2014). Therefore, lead users are incorporated in a very early phase in the innovation process (Hung et al., 2011). This phase "Idea generation" can be seen as a venue for brainstorming to make the free exchange of ideas possible (Muller et al., 2012; Paulus et al., 2002). When brainstorming, people are encouraged to generate as many ideas as possible and therefore a high participation as well as a high activity level is necessary here (Chen et al., 2019; Hung et al., 2011; Miao & Zhang, 2017))
- Lead users are not only able to provide new ideas but can also be integrated into the "Development" phase of the innovation process. Because of their high product related knowledge and their vast experience lead users are able to suggest concrete solutions instead of describing problems or stating customer needs (Mahr & Lievens, 2012). Hence, scientific articles which characterize and identify lead users in terms of the "Development" phase focus on users e.g. who have already done security-related modifications to a web server software (Franke & von Hippel, 2003) or who have already developed applications for different plattforms (Schaarschmidt et al., 2019). Hence, Mahr and Lievens (2012) summarize it and state that lead users are best suited for improvements pertaining to functionality. Thus, lead users in this second phase of the innovation process are able to support companies in order to develop new products and solutions with the aim of meeting rapidly changing consumer needs and to stay competitive (Pajo et al., 2014). This can diminish failure rates of new product introduction. So utilizing this high-product related knowledge combined with the high level of trend leadership, a lead user can be conducive in strengthening a company's innovation competitive advantage (Li & Tang, 2016).

The assignment of the characteristics to the different phases in the innovation process and thus the difference made by the lead users in the two innovation phases can be detected in Table 1.

Related work

The review of the research literature has shown that, in addition to the two characteristics from the basic model of the Lead User Theory, there are many different characteristics to describe and characterize lead users in the online environment (see Table 1), whereby different approaches such as screening (Brandtzaeg et al., 2016; Hung et al., 2011; Tuunanen et al., 2011), pyramiding (Von Hippel et al., 2009), SNA (Martínez-Torres, 2014), or netnography (Belz & Baumbach, 2010; Mahr & Lievens, 2012) have been used. However, these studies on lead user identification covered only a limited point of view as they either focus only on one characteristic of a lead user, like the high level of activity (Martínez-Torres, 2014), or they include a very small amount of data (Hau & Kang, 2016). Furthermore, investigations are based on observations or (online) questionnaires (Brandtzaeg et al., 2016; Hung et al., 2011; Tuunanen et al., 2011) which results in a low sample efficiency and high costs. In addition, lead user characteristics are thereby based on the self-assessment of respondents, which means that the results can be affected by subjective assessments (Hienerth & Lettl, 2017). Another problem within the current research literature is the aspect of time (Brem et al., 2018; Hienerth & Lettl, 2017). Most of the aforementioned identification methods are time-consuming. This contrasts with the fact that the concept of "lead userness" is not a long-term construct but it is trend specific and can change over time. A lead user today may or may not be a lead user in distant future (Hienerth & Lettl, 2017).

Furthermore, as already mentioned, lead users can be identified in both innovation phases: the "Idea generation" and "Development" (Füller et al., 2007). However, in current research literature it is common practice to identify a lead user for one of the two phases (Marchi et al., 2011; Martínez-Torres, 2014; Schaarschmidt et al., 2019). Moreover, only a minority of the investigations incorporates an innovation process but identifies a lead user for all of its phases (Miao & Zhang, 2017; Pajo et al., 2017). Consequently, the current research literature is incomplete here as there is no approach that identifies different lead users for every phase of the innovation process and so an overall approach is missing. Thus, we aim to show the differences between the lead users in the two innovation phases and the legitimacy of the different identification processes, although we have seen in our summary Table 1 that there are no large differences between the two phases.

All in all, in order to avoid these negative aspects of the prior research literature, we have come to the conclusion that a tool for automated lead user identification is needed. This tool should meet the following design requirements and should therefore be able to:

- map all prior identified characteristics,
- process a large amount of online community data,
- apply objective identification methods,
- repeat the identification process for lead users at any time as lead users are trend specific, and
- identify different lead users regarding the two phases of the innovation process.

Procedure of the research

In order to make the development of a systematic approach for the automated identification of lead user comprehensible, we applied Design Science (DS) research. Research projects that follow the DS paradigm are concerned with the design, development, implementation, use, and evaluation of socio-technical systems in organizational contexts. Design scientists produce and apply knowledge of tasks or situations to create effective artifacts (March & Smith, 1995). These artifacts are delineated in different structured forms such as software, formal logic, and rigorous mathematics to informal natural language descriptions (Hevner et al., 2004).

An important step in DS research is to prove the utility, quality, and efficacy of the artifact via well-executed evaluation methods. Since the artifact's performance is related to the environment in which it is used, an incomplete understanding of the environment can induce inappropriately designed artifacts (March & Smith, 1995). Therefore, Hevner's "design cycle" (Hevner, 2007) substantiates the importance of constructing and evaluating the artifact, and suggests balancing the efforts spent on both activities, which must additionally be convincingly based in relevance and rigor (Hevner, 2007). Consequently, DS research is based on and contributes to scientific knowledge by performing the research process rigorously (e.g., by reflecting the construction or/and evaluation of the artifact) which is represented by Hevner's "rigor cycle". DS research additionally uses practical knowledge and leads to several practical contributions which constitutes Hevner's "relevance cycle" and which can be seen as self-evident objectives of a DS research project (Hevner, 2007).

We followed the DS research paradigm (Gregor & Hevner, 2013; Hevner et al., 2004) and aligned our research activities with the procedure as proposed by Peffers et al. (2007) (see Fig. 1). This procedure provides a commonly accepted framework for conducting research based on DS principles. In addition, Peffers et al. (2007) designed the procedure as a result of a consensus-building approach, which comprises well-agreed process elements (Peffers et al., 2007). As a first step, (1) corresponding problems and drawbacks of already existing approaches to identify lead users in online communities were identified (see sections "Introduction" and "Conceptual basics"). Hence, in

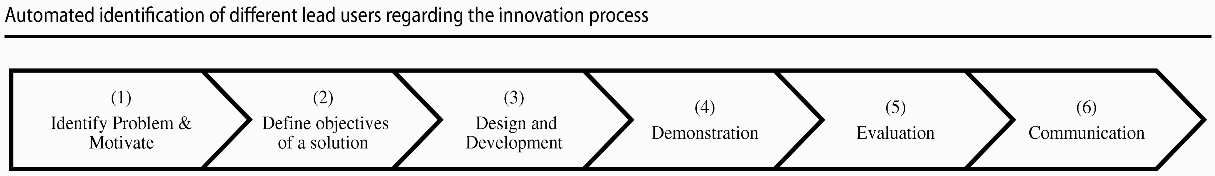


Fig. 1 Design science process

current research literature there are a lot of different lead user identification approaches, but these investigations only cover a limited point of view as they either focus only on one lead user characteristic, they include a very small amount of data or their approach is reliant on the self-assessment of users. Consequently, our (2) objective is to provide and combine a set of methods, based on the characteristics of a lead user, in order to identify this type of user automatically in an online community (see sections "Conceptual basics" and "Design and development"). The third step of the DS process model contains the (3) design and development (see section "Design and development") of a solution or an artifact, respectively. Such artifacts can be constructs, models, methods or instantiations. In order to fill the gaps identified within phase (1), we focus on the design of the technical realization of the tool by means of the combination of different methods such as SNA, topic modeling and sentiment analysis. Thus, our approach was established to support and simplify the lead user identification process and to eliminate the existing disadvantages. In the next step, the (4) demonstration, we show the application of the demonstrated approach on approximately a data set of about 12,000 contributions from an online community about kitesurfing. Kitesurfing is a suitable area of application as individuals in this area are quite active as innovators. Furthermore, kitesurfing comprises a young community, essentially all serious participants are active members in some kind of online communities (Franke et al., 2006; Von Hippel, 2005; Wagner & Piller, 2011). The overall results of the application of the analysis are shown in this chapter, consolidated in a summary table and discussed in detail. These results are additionally evaluated (5) by conducting both interviews with lead users and an in-depth interview with an expert (head of marketing of our cooperating partner) in the field of kitesurfing. In terms of these interviews, we evaluated our artifact and showed that our approach provides an added value. We have further discussed the results of the evaluation as well. Finally, the results are then (6) communicated.

The orientation towards the procedure by Peffers et al. (2007) also makes it possible to align our research with the guidelines of Hevner et al. (2004) or Hevner (2007), respectively. According to the design cycle, we present our artifact as the result that has gone through the process of demonstration (application of our approach to an online community about kitesurfing and evaluation with several interviews (see section "Demonstration, evaluation and discussion")). In view of the relevance cycle, we identified several design requirements (from literature including several case studies (see section "Conceptual basics")) that guided the design of the artifact, and so the practical application of our artifact brought up several contributions for practice (e.g. identifying relevant users for innovation/trends (see section "Discussion of the results of demonstration")). In view of the rigor cycle, we used several methods and techniques to rigorously construct and evaluate our artifact (e.g. topic modeling, SNA, frequency analysis) and derived initial findings as contributions to theory, both kernel theory (Lead User- and Innovation Theory) and Design Theory (see section "Contribution for practice and research"). Thus to contribute to a rather general and abstract knowledge base – "nascent design theory" (Gregor & Hevner, 2013) – and in order to design a purposeful artifact in a comprehensible way, we first established both, a set of meta-requirements and design principles (Gregor & Jones, 2007; Heinrich & Schwabe, 2014). Thus, the design of the lead user identification tool is grounded on design requirements retrieved from seminal works on Lead User- and Innovation Theory. In a next step, we then describe our prototypical implementation that demonstrates the feasibility of the design principles and meta-requirements in the tool.

Design and development

Design principles for a lead user identification tool

First, the composition of meta-requirements (MRs) that describe "*what the system is for*" (Gregor & Jones, 2007, p.325) is based on the purpose and scope of the identification tool that has been discussed in the introduction. Thus, we define the solution objectives based on the class of problems our paper addresses and present them in Fig. 2. These MRs established to be suitable for a class of artifacts and are based on the current research literature (Gregor & Jones, 2007; Heinrich & Schwabe, 2014; Walls et al., 1992). Besides the MRs, the design principles are synthesized in a next step. Design principles are defined as prescriptive statements that show how to do something to achieve a goal (Gregor et al., 2020). The design principles that we dispose fall into the category of action and materiality-oriented design principles that describe what an artifact should enable users to do and how the artifact should be built in order to

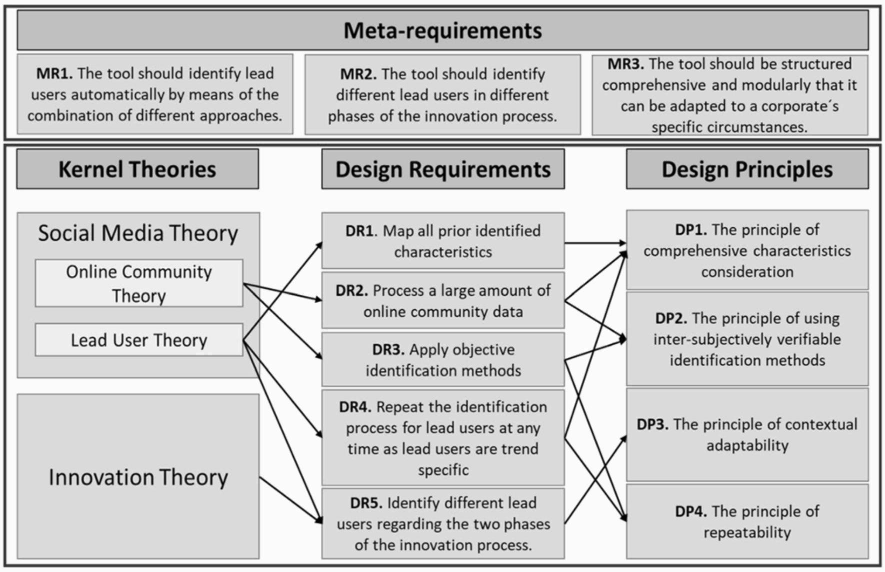


Fig. 2 Overview of the design phase

do so (Chandra et al., 2015). Regarding companies (= users) who are interested in identifying lead users in online communities (= boundary conditions) and keeping our design requirements for our artifact in mind, we derive four design principles for (lead) user identification tools:

1. The principle of comprehensive characteristics consideration. In order to identify specific user types in online communities, e.g. a lead user, it is necessary to precisely define and describe their characteristics. Thus, the automated identification of a lead user requires a technical implementation of its characteristics that we have derived from the current research literature. Therefore, the tool should be able to incorporate and technically realize all relevant lead user characteristics (trend leadership, sentiment, high activity level, high product related knowledge and opinion leadership) to obtain precisely targeted results.
2. The principle of using inter-subjectively verifiable identification methods. In order to counteract the subjective self-assessment of respondents of (online) questionnaires different inter-subjective methods should be consulted and combined to identify a lead user. Therefore, the tool should use comprehensible and inter-subjectively verifiable identification methods to make the identification process traceable.
3. The principle of contextual adaptability. Since lead users are applied in terms of innovations in a company, the identification of a lead user must also take into account the different phases of the innovation process the user supports. Therefore, the tool should be able to adapt the weights of the characteristics according to the different circumstances of the companies and their aim to apply lead user in different stages of the innovation process.
4. The principle of repeatability. As lead userness is a short-term construct, which means that lead users can change over time, the identification process should be executable often and in a resource-saving way. Therefore, the tool should allow repetition of the identification process at any time to react quickly to changing circumstances such as trends.

These design principles are deduced from the design requirements that are further based on kernel theories and prior research literature. Gregor and Jones (2007) state that these kernel theories disclose "*an explanation of why an artifact is constructed as it is and why it works*" (p.328). So, these kernel theories include the body of knowledge that is necessary to meet the design requirements (Böckle et al., 2021). Hence, based on the discussion of the kernel theories and thus the related work, we derive design requirements our tool should meet. These design requirements offer guidance by designing the artifact and advise the design principles (Böckle, et al., 2021; Gregor & Jones, 2007). These principles refer to at least one requirement and serve as an abstract "blueprint" of our artifact (Böckle, et al., 2021; Gregor & Jones, 2007; Heinrich & Schwabe, 2014). By establishing these design principles, we made sure that they follow the value grounding (reference to the requirement) and the explanatory grounding (design principles are based on the current literature and thus on kernel theories) (Heinrich & Schwabe, 2014).

For each of the design principles, its instantiation in the artifact is described in the following sections.

Weighting of the according lead user characterizations

To address the shortcomings of the prevailing research and therefore consider the derived design principle *1. The principle of comprehensive characteristics consideration*, we aim to compose an automatic identification approach including all characteristics identified in literature. Furthermore, to account for the different circumstances of the two innovation phases (see section "Lead user innovation"), we also distinguish between lead users associated to the phase “Idea generation” and lead users associated to the phase “Development” of the innovation process. Therefore, to consider the different relevancies related to the identified characteristics with respect to each innovation phases, the characteristics are weighted accordingly (see Table 2) based on their occurrences within the current research literature (see Table 1). In addition, with respect to the derived design principle *3. The principle of contextual adaptability*, companies are enabled to adapt the respective weights to their circumstances and thus to apply lead users in different stages of the innovation process.

Table 2 summarizes the characteristics' accompanied relevancies in the context of the respective innovation phases. Here, the weights illustrate that the respective focus of the innovation process within the activity of the users (Idea generation) as well as their product-specific knowledge (Development) differs considerably. Additionally, users also differ in the mood they exhibit. Here, the characteristic dissatisfaction is given greater meaning in the “Idea generation” phase since users express their unmet needs of a product or service within negatively afflicted communication. These differentiations enable to adequately consider the circumstances of the two innovation phases, resulting in the determination of precisely fitting and goal-oriented lead users.

Table 2 Weighting the relevance of each characteristic for both innovation phases

Characteristics	Weight by occurrence in literature	
	Idea generation	Development
Trend leadership	5	5
Dissatisfaction	3	2
Enjoyment	1	1
High activity level	5	3
High product related knowledge	4	4
Opinion leadership	2	2

Technical realization

To enable the automated identification of lead users based on the above determined relevancies, the previously identified characteristics (see Table 1) must be mapped in an automatic manner. Therefore, we have implemented each identified characteristic in the programming language *Python*. As the underlying data (e.g. online community posts, network interactions, etc.) are mainly represented in a textual way, we focused on finding computer-based procedures from the research field of text mining to map the identified characteristics. Text mining enables an automatic identification of hidden structures or patterns within a corpus of textual data (Feldman & Sanger, 2007; Heyer et al., 2006). In addition, we conducted the SNA, which best suits the identification of users within a potentially high influence to be solved, as SNA enables us to show the relations in a structured network via nodes and ties to state quantitative characteristics of users. Furthermore, due to the different nature of each characteristic (see Table 3), the values must be normalized to make them comparable. Therefore, we have conducted the Min–Max normalization (Han et al., 2006) to rescale each characteristic into a value range between [0;1]. In the course of normalization, the specific values of all users were related to each other. Thus, the higher the respective value, the more the respective user exhibits the specific characteristic. By this, all values are located at the same scale and therefore can be weighted by their allocated relevance (see Table 2). To give an overview of the characteristics and their technical realization, they are further summarized in Table 3. Here, to consider the derived design principle *2. The principle of using inter-subjectively verifiable identification methods* and therefore ascertain an adequate analysis process,

Table 3 Technical conception of the identified characteristics

Characteristics	Technical realization	Implementation	Outcome
Trend leadership	Latent Dirichlet Allocation (LDA)	Gensim; Mallet (McCallum, 2002)	Probability
Dissatisfaction	Sentiment Analysis	VADER (Hutto & Gilbert, 2014)	Classification
Enjoyment	Sentiment Analysis	VADER (Hutto & Gilbert, 2014)	Classification
High activity level	Frequency Analysis	Self-created	Frequency
High product related knowledge	Frequency Analysis	Self-created	Frequency
Opinion leadership	Closeness, Betweeness and Degree centrality	NetworkX (Hagberg et al., 2020)	Centrality measure

all technical realizations are based on broadly known and prevalent quantitative and qualitative content analysis methods. To further meet the particular needs associated to the respective characteristic, the methods used have further been adapted as described in the following.

With respect to the characteristic of trend leadership, the aim is the identification of users who talk about trends before they became general, community-wide discussed topics. To meet these requirements, we mainly had to split the automation into two separate sub-sequences: (1) identify trends (e.g. frequently discussed product issues or service properties), based on UGC; (2) identify users who talked about one or multiple of these previously identified trends, before they became discussed community-wide. With regard to these two identified sub-sequences, we focused on the use of text mining methods, enabling the automatic processing of unstructured, unlabeled data such as online community posts. More specifically, as trends represent frequently emerging topics as well as the advantage of topic modeling compared to other text mining techniques to operate directly on the textual data instead of solely comparing their underlying structure (Aggarwal & Zhai, 2012), we have chosen topic modeling for the automatic identification of trends. Topic modeling can project the textual corpus of contributions into a topical space by reducing the dimensionality and attaching different weights, which results in semantically coherent groups of words (topics), which represents our trends (Crain et al., 2012; Xie & Xing, 2013). Specifically, because of LDA's simple applicability but also its satisfactory analysis results within the topic modeling (Eickhoff & Neuss, 2017), the choice was made for LDA. For the implementation of LDA within the automated identification approach, the established python library Gensim was used in combination with Mallet (see Table 3). In order to achieve the highest quality of results possible, we further automatically prepared the data for the analysis by applying tokenization, stop word removal and case folding (cf. Boyd-Graber et al., 2014). Furthermore, in order to take the characteristics of trends into account (1), we adapted LDA to only consider contributions of the last eight weeks to extract the trending topics. By this adaption, the identification of those user who were already talking about these trending topics within their contributions at an earlier point in time than eight weeks ago (2) is feasible. The identification takes place through statistical inference and reflects the cumulated probability with which a user talks about one of the identified trending topics.

Considering the characteristic of dissatisfaction or enjoyment aims to identify users with either negative or positive mood. Therefore, the global mood of each user within their contributions has to be identified. The automatic identification of moods within textual data is summarized under the term “sentiment analysis”. Through this, for instance, it is possible to identify users who have unfulfilled expectations and thus show a significant potential for improvement of a product or service (Pajo et al., 2017). To determine the mood of each post by a user, we implemented the "Valene Aware Dictionary for sEntiment Reasoning" (VADER) (Hutto and Gilbert, 2014) technique. VADER is a lexicon and rule-based sentiment analysis technique that is specifically attuned to sentiments expressed in social media and has achieved remarkable results compared to other prominent sentiment analysis techniques (Hutto and Gilbert, 2014). To determine the sentiment value, VADER uses a labelled dictionary adapted to the contextual characteristics of social media data. Hereby, VADER is able to combine the positive and negative inflections and generates a single sentiment score within the range of -1 to + 1. In order to determine the global sentiment value of each user, we further adjusted the technique to build a consolidated sentiment score for each user reflecting its global mood by setting the individual scores of each contribution into relation to the total amount of contributions of a user. This results in the mean value of all mood-bearing contributions of a single user, which reflects their average mood.

To measure the activity level of a user, we further determined the amount of user interactions within the community. For this purpose, the number of posts and transacted comments per user within the analyzed period was identified to attain information about the activity level of a user (Miao & Zhang, 2017).

In the case of high product related knowledge, the aim is to identify users who have an immense knowledge of product specific information. To accomplish this, we considered splitting the determination of the characteristic into two parts. In the first step, a dictionary of product-specific terms was extracted from product and service descriptions e.g. product brochures. Secondly, following the generation of the product-specific dictionary, the occurrence of the extracted product-specific words in the contributions were determined. Therefore, matching word candidates from the contributions are identified and reflected against the product-related dictionary. If an entity matches with a product-specific word, the total sum of the user's usage of product-related words will be increased. After all contributions of the related user have been analyzed, the number of product-specific words is divided by the total number of all words used by the specific user. The resulting value reflects the average use of product-related words by a user and allows conclusions to be drawn about the product knowledge of a user.

With regards to the determination of the user's ability to enable the flow of information and especially diffuse it, which are prerequisites for opinion leadership, we have considered several centrality measurements which best suit the identification of strong social relationships within a social network (Pajo et al., 2017). These measures are those of *degree, closeness* and *betweenness centrality*, and are fundamentally related to the concept of social influence in terms of the structural effects of different connections within a network of users (Aggarwal, 2011). Degree centrality σ_D is used to determine the number of direct interactions of a participant within the network, which represents an indicator of quality for the member's interconnectedness. Through the use of an adjacency matrix $A=(a_{ij})$, the degree centrality can be formulized as follows:

$$\sigma_D(x) = \sum a_{ix}. \tag{1}$$

As a consequence, the higher the centrality score $\sigma_D(x)$ is, the more contacts a node x has. Thus, by implementing the degree centrality, we are able to identify those users who have the most interactions with other network participants (Aggarwal, 2011). The closeness centrality σ_C is based on the idea that nodes with a short distance to other nodes can disseminate information very productively in the network. To calculate $\sigma_C(x)$ of a node x, the distances between node x and all other nodes in the network are summed up. By using the reciprocal value, the closeness increases when the distance to another node decreases, i.e., when the integration into the network is improved. The closeness centrality can be formulized as follows:

$$\sigma_C(x) = \frac{1}{\sum_{i=1}^{n} d_G(x,i)} \tag{2}$$

In this respect, through the implementation of the closeness centrality, we are able to identify those users who distribute information among other network participants as efficiently as possible (Latora & Marchiori, 2007). In case of the third centrality measure—the betweenness centrality σ_B—a network member is well connected if it is located on as many of the shortest paths as possible between pairs of other nodes. The underlying assumption of this centrality measure is that the interaction between two non-directly connected nodes x and y depends on the nodes between x and y. The betweenness centrality for a node x can therefore be formulized as

$$\sigma_B(x) = \sum_{i=1, i\neq x}^{n} \sum_{j=1, j<i, j\neq x}^{n} \frac{g_{ij}(x)}{g_{ij}} \tag{3}$$

with g_{ij} representing the number of shortest paths from node i to node j, and $g_{ij}(x)$ denoting the number of these paths which pass through the node x. Through this, we are able to identify those situated on the shortest path distance between various actors, showing that a user has fast access to and control over network flows (AlFalahi et al., 2014; Freeman & Soete, 1997). By these centrality measures, we are able to subdivide the users on the basis of their network characteristics. Regarding the calculation of the respective centrality measurements, the well-known and widely used python library *NetworkX* found application (see Table 3). Besides the plain calculation of the centrality measures of each user, we further adapted the technique to normalize the calculated values into the range of [0;1]. Based on this normalization, it is possible to consolidate the different centrality measures into a single value by calculating their mean. By this, the respective user's position in the network and therefore their ability to enable the flow of information is being represented.

Following the calculation of the individual metrics, the automatic identification of the lead users per phases in the innovation process takes place. Therefore, the result per metric is multiplied by the corresponding weight of the respective phases (see Table 2) and summed up for each specific user. Finally, the calculated sum is divided by the maximum number of points to be achieved (see (4)). Thus, two resulting scores for each user, one each for the two phases in the innovation process, will be generated. These two resulting scores represent the cumulative relevance of a user with respect to the phases in the innovation process. The higher the resulting score for a respective user is, the more highly the user is defined by his characteristics as a lead user for one of the respective phases: "Idea generation" or "Development".

$$\text{score}_i = \frac{\sum \left(x_i * w_i\right)}{\sum w} | w = weight; \; x = metric \tag{4}$$

An identification of lead users according to the described procedure enables the determination of users who show particular strength in terms of relevant characteristics such as their influence on other participants within the community, their product related knowledge or their level of activity. In combination with an individual weighting of these characteristics, the two identified phases of "Idea generation" and "Development" are also considered. In addition, with respect to the derived design principle *4. The principle of repeatability*, the artifact is designed in a modular and generic way. Thus, the underlying data and the respective characteristics' weighting can be easily adapted, allowing the identification process to be carried out at any time without further restrictions to e.g., react quickly to changing circumstances such as trends. Finally, as our design principles follow our purpose and scope and found consideration within the designed artifact as described above, the derived meta requirements (see Fig. 2) can be seen as successfully met since they are representing our underlying purpose and scope.

Demonstration, evaluation and discussion

To demonstrate the applicability of the developed artifact – including the identification approach and the corresponding software tool – we have conducted several steps. In order to verify the consideration of the design principles, the underlying design requirements are examined for their met using our specific use case (see Table 4). Subsequently, the artifact was applied on a real-world kitesurfing dataset to ensure its usability for practice. Further, we conducted interviews with our identified lead users and with an expert from our cooperating partner, a market leader in kite- and watersports to evaluate both the usability and the generated added value for practice.

Review of the identified requirements

In order to verify the derived design principles, we further review whether and how the elicited design requirements of our artifact (see Fig. 2) were met. Therefore, we specify them in more detail in Table 4.

Demonstration of the artifact

In order to facilitate the accessibility concerning the use of the developed artifact, including the designed identification process, all its customizabilities, as well as the monitoring of the analysis, a graphical user interface (GUI) was developed. To ensure the development of a highly responsive, performant and platform-independent interface, the GUI was developed using the standardized and well-known Python library PyQt5. Figures 3 and 4 show the two main interfaces – namely the configuration and the result table view – of the developed GUI.

The configuration (see Fig. 3) represents the initial view when starting the tool and can be used to customize the underlying analysis approach to one's own needs. The layout was designed based on three sections (i)—(iii), following an adaptation of the design principles of Garrett (2010).

Fig. 3 GUI—Configuration view

In section (i), the user can flexibly specify the data to be analysed as well as the output path for storing the analysis results by selecting the appropriate directories within the native filesystem. Further, in case of not all elicited

Table 4 Met of the previously identified requirements

Design Requirements	*Requirements met*
Map all prior identified characteristics	As all identified characteristics of lead users are considered within the artifact (see Table 3), we see the requirement as met.
Process a large amount of social media data	The artifact was applied successfully to a real-world dataset of 11,481 contributions (see Table 5). Furthermore, by choosing adequate analysis techniques settled in the field of text mining or SNA respectively (see Table 3), the analysis of larger data sets is easily possible.
Apply objective identification methods	The implemented characteristics were either implemented by well-known text mining or SNA techniques or by self-developed techniques that build on patterns of well-known analysis methods from text mining (see section "Technical realization").
Repeat the identification process for lead users at any time as lead users are trend specific	As the developed artifact is implemented in a generic way, the underlying data and the respective weighting of the characteristics can be easily adapted, allowing the identification to be carried out at any time without further restrictions (see section "Technical realization") to e.g., react quickly to changing circumstances such as trends.
	Additionally, the identification of lead users within the real-world application with 11,481 contributions took 1317 s, which is why a quick identification based on a new dataset is also easily possible.
Identify different lead users regarding the two phases of the innovation process	As the proposed artifact differs between lead users regarding the phase "Idea generation" and lead users regarding the phase "Development", we see the requirement as met.

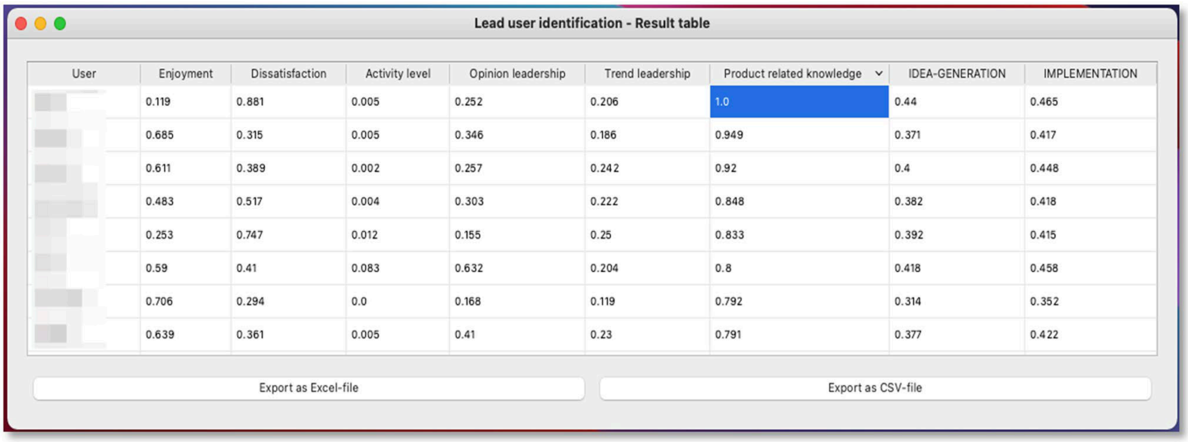


Fig. 4 GUI—Resulting table view

characteristics being deemed necessary, a subset of them can be individually defined, comprising all specific characteristics relevant to the current circumstances. This ensures that only favoured characteristics are considered in the analysis, resulting in a resource-efficient identification of lead users who show particular strength in terms of significant characteristics such as their influence on other participants within the community (opinion leadership) or their level of activity. In addition, to incorporate the two identified phases "Idea Generation" and "Development", they were implemented modularly using dynamic tabs to enable a distinctive configuration (see Fig. 3, ii). Here, the weightings for each respective phase elicited from literature (see section "Weighting of the according lead user characterizations") are defined as default within the phase's configuration. Nevertheless, to provide maximum flexibility and to be able to react quickly and almost effortlessly to changing circumstances, the pre-defined weighting can also be individualized per phase through the corresponding text fields highlighted in section (ii). The start of the analysis process as well as its monitoring takes place in section (iii). As soon as the process is initiated, all relevant information concerning the process such as the current state or occurring errors will be monitored and logged within the designated text area (see Fig. 3, iii).

Once the process has finished, the results will be consolidated and displayed by a responsive, sortable table (see Fig. 4). Here, the results are subdivided into each characteristic as well as the two calculated phase-specific scores, featuring the dispensation of the users to each characteristic or innovation phase, respectively. To facilitate the selection of relevant lead users, it is further possible to filter the identified users based on each calculated value (see Fig. 4, "Product related knowledge"). This allows companies to select users in an intuitive way based on either a specific characteristic such as product-related knowledge or the overall scores.

To preserve the obtained results for later usage, two functions were implemented to handle the extraction by use of either a Microsoft Excel or a Comma-separated value (CSV) file. These file formats enable a platform-independent presentation of the results for e.g., marketing campaigns (Excel) as well as the automated processing by a proprietary third-party system, such as importing the generated information into the company's active directory (CSV).

To demonstrate the practical applicability of our developed tool, a representative real-world dataset was needed. Thus, we extracted data from one of the most popular online communities for kiteboards (https://www.seabreeze.com.au/), which comprises a total of 11,481 contributions of 945 users. The data were extracted using the ParseHub extraction tool and span the period from January 1st, 2018 to April 10th, 2020.

Based on these data, the analysis was undertaken to identify the respective lead users. Table 5 presents the top five identified users per phases of the innovation process. The values represent the previously identified characteristics by a normalization within a scale of [0;1]. Thus, a high value implies the respective characteristic is strongly distinctive. The identified users are differentiated regarding the two innovation phases. Accordingly, the weights of the characteristics were adapted to the respective needs of the phase (see section "Weighting of the according lead user characterizations"). The "Overall score" represents the affiliation of the respective user in each phase and is determined through the weighting of the characteristics.

A cursory glance at Table 5 reveals that lead users can be identified in both phases of the innovation process. Thus, the two identified lead users: user #1 and user #2 are identified as lead users exhibiting the highest values compared

Table 5 Top five identified lead users for the specific innovation phases

			Enjoyment	*Dissatisfaction*	*High product related knowledge*	*Trend Leadership*	*High Activity Level*	*Opinion leadership*	*Overall Score*
	User	Weight	1	3	4	5	5	2	
Idea Generation	user #1		0.347	0.653	0.367	0.489	1.0	1.0	***0.662***
	user #2		0.395	0.605	0.438	0.374	0.84	0.925	***0.602***
	user #3		0.346	0.654	0.506	0.304	0.353	0.81	***0.472***
	user #4		0.295	0.705	0.291	0.291	0.497	0.743	***0.46***
	user #5		0.364	0.636	0.36	0.3	0.381	0.79	***0.442***
	User	Weight	1	2	4	5	3	2	
Development	user #1		0.347	0.653	0.367	0.489	1.0	1.0	***0.6***
	user #2		0.395	0.605	0.438	0.374	0.84	0.925	***0.557***
	user #6		0.119	0.881	1.0	0.204	0.005	0.252	***0.51***
	user #7		0.611	0.389	0.92	0.24	0.002	0.257	***0.489***
	user #8		0.59	0.41	0.8	0.203	0.083	0.632	***0.478***

to all users of the innovation process. We assume that the identification of users present in both innovation phases is a rarely occurring exception resulting from extremely conspicuous users. Here, the two identified users have a significantly higher activity level than the lead users identified for a specific phase, which supports this conclusion. User #1 even has the highest activity level (1.0) among all 945 users. In addition to the identification of users who are present in both phases, lead users, who differ significantly in their descriptive characteristics, were further identified for each innovation phase. User #6 e.g., shows an activity level way below average (0.002), but exhibits remarkable results regarding the presence of product related knowledge (0.92). Therefore, the user is considered as lead user regarding the second innovation phase "Development".

To be able to identify lead users scaled to the different circumstances of enterprises in a resource-optimized way, a high degree of generalizability was considered in the implementation of the artifact. Therefore, to be able to adequately react to specific circumstances, the weighting of the respective characteristics can be individualized at the beginning of the analysis process. Thus, the identification approach can be specifically geared to users who dominate a single criterion or a combination of criteria and can therefore be easily adapted to different conditions.

Finally worth to mention, the related lead user characteristics (and therefore the lead users themselves) are validated in an intrinsic way by incorporating different evaluation metrics (e.g., topic coherence) during the identification process. In this way, a high information quality is ensured, supporting the practical applicability of both, the identification process and the retrieved lead users. In this regard, by applying our tool, we revealed promising lead users for the specific innovation phases based on their remarkable characteristics exhibited (see Table 5, e.g. user #1). However, as the intrinsic evaluation of probabilistic models such as topic modeling (trend leadership) poses various challenges and drawbacks (Chang et al., 2009), it is not sufficient to verify the elicited results. Thus, we evaluate the identified lead users and their characteristics in an extrinsic way by verifying the identified lead users through an interview with an expert of our cooperating partner (a market leader in kite- and watersports) and respective lead users (see section "Evaluation of the artifact"). Therefore, we will evaluate the derived lead users as well as the identification process by applying them to our specific use case, revealing their meaningfulness and potential regarding their practical applicability in a first step.

Discussion of the results of demonstration

Our results have shown that our identification approach and the corresponding software tool are working immaculately. The implementation of the design principles was thus feasible, resulting in the identification of lead users for both innovation phases (see Table 5). Prior research literature is inchoate here as only a minority of the investigations incorporates an innovation process but identifies a lead user for all of the phases in an innovation process (Miao & Zhang, 2017; Pajo et al., 2017). Thus, we provide a new approach that identifies different lead users for every phase of the innovation process. Most of the identified lead users are better suited for one of the two phases but there are also lead users who exhibit very high values in both innovation phases. We have shown with our results that a clear differentiation of the two phases as well as the separated identification and consideration of lead users is necessary as they have different competencies, characteristics and application areas.

The lead users #3 or #4 are according to our results an adequate choice when searching for a lead user in terms of the innovation phase "Idea generation". User #3 e.g. features a high value in the dimensions "trend leadership" (0.304) and "high activity level" (0.353). This means that this lead user can be seen as an active member in the kitesurfing-lifestyle scene. His/her creative and active participation in the online discussions shows that this lead user is highly involved in the kite community. Their active participation and involvement additionally lead to the awareness of unmet needs about existing solutions in the kiteboarding scene (see dissatisfaction: 0.654). Because of his/her high value in the dimensions "trend leadership" (0.304), we can assume that this lead user is able to "translate" his/her dissatisfaction into concrete ideas. Against the background of the fact that a company requires many initial ideas from lead users (as only a few of them can be realized anyway) especially the requirement "repeat the identification process for lead users at any time" is important in this first innovation phase. Approaches that are established and discussed in prior research literature such as screening (Brandtzaeg et al., 2016; Hung et al., 2011; Tuunanen et al., 2011), pyramiding (Von Hippel et al., 2009) and other lead user identification procedures are often based on interviews or (online) questionnaires (Brandtzaeg et al., 2016; Hung et al., 2011; Tuunanen et al., 2011) which makes it almost impossible to repeat the identification process for lead users at any time. However, most business-to-consumer industries are fast-moving and therefore identifying innovative lead users and their ideas repeatedly with little expenditure of time must be focused on. Thus, with our artifact a company is able, on the one hand, to identify lead users who are currently ahead of trends and, on the other hand, to react to changing circumstances such as trends. Furthermore, with this procedure we also counteract the low sample efficiency and the high costs that results from conducting interviews and online questionnaires.

By examining the detailed results of the phase "Development" it is especially interesting that user #6 and user #7 are determined as lead users although they exhibit very low values in the dimension "high activity level" (user #6: 0.005; user #7: 0.002). This shows that due to the medium weight of the dimension "high activity level" in the phase "Development" a lead user does not necessarily exhibit a high active usage behavior which contradicts the results of Martínez-Torres (2014). This means that in our case the lead users #6 and #7 posted only few contents and therefore they do not participate a lot in these online community discussions. But if they did submit a post, it contained very valuable and detailed content including high product-related knowledge (user #6:1.0 user #7: 0.92). These users suggested concrete solutions based on their broad expertise about the products, the components and how they mesh with each other.

The combination of this "high product related knowledge" with the relatively high value for "trend leadership" (0.204) and the simultaneously low value for the "high activity level" (0.005) led to the assignment of user#6 to the phase "Development". The results have shown that this combination of characteristics is more decisive than, for example, the dimension "dissatisfaction" as user #6 is the lead user (compared to all other identified users) with the highest value in the dimension "dissatisfaction" (0.881). This high value of "dissatisfaction" would actually speak for being assigned to the "Idea generation" phase as it is weighted higher (Idea generation: 3, Development: 2) here. However, the high level of product related knowledge and the associated ability to suggest concrete solutions for new products or their improvements is the main factor for a promising cooperation in terms of the "Development" phase. Therefore, observing both the individual results of the characteristics and the overall score has shown that not only is the weighting and selection of the characteristics plausible, but that our software tool is also capable of finding them.

Moreover, we have identified lead users for both phases as they exhibit the highest overall scores: user #1 and user #2. Both users exhibited extraordinary results, especially for the dimensions "high activity level" (user #1: 1.0; user #2: 0.84) and "opinion leadership" (user #1: 1.0; user#2: 0.925). So, their high level of involvement in the kitesurfing scene is characterized by their active participation as well as by their central position in the network. Displaying both strong social relationships and high levels of engagement enable the users to spread information in the online community. Consequently, these users know the overall sentiment and can also identify unmet needs that will later be experienced by the public (see trend leadership and dissatisfaction). Based on this, they formulate and disseminate ideas for new products as well as suggest detailed solutions for the prior identified needs. Thus, they facilitate effective innovation and encourage innovation communication. Only in some dimensions such as "high product related knowledge" other users (e.g. users: #6, #7 and #8) exhibit better results. Nevertheless, the users' overall scores in both phases show their outstanding position as lead users, which we assume, however, that this can be seen rather as an exception. Furthermore, we are convinced, that it makes sense not only to focus on lead users who are suitable for both innovation phases but rather consider for what purpose a lead user should be engaged and to adjust the weighting accordingly. If a lead user is only active in one phase, then s/he can focus on either the objective generating many good and innovative ideas (see Idea generation) or on the objective developing explicit solutions for unmet needs (see Development) and applies his/her strengths accordingly. In other words, if a lead user would act in both phases and thus focuses on both objectives simultaneously, e.g. knowledge about the realization and development could have a negative impact on the creativity in the idea generation phase. Concentrating on lead users for both phases simultaneously also means that lead users who have extraordinary new ideas but only less product-related knowledge would be excluded. This would lead to a loss of new and potentially successful ideas.

In summary, observing both the individual results of the characteristics and the overall score has shown that not only is the weighting and selection of the characteristics plausible, but that our software tool is also capable of finding them. Our results follow from combining, weighting and considering all relevant lead user characteristics. Previous research literature (e.g., Miao & Zhang, 2017; Martínez-Torres, 2014; Tuarob & Tucker, 2014) has often focused on only few characteristics for identifying lead users and thus resulted in many and ambiguous lead users.

Evaluation of the artifact

To evaluate the generated artifact and our results, we conducted both an in-depth interview with the head of marketing of our cooperating partner and interviews with some of our identified lead users.

The in-depth interview was undertaken by two researchers, recorded, transcribed, and finally condensed to the most important statements. It lasted approximately two hours and we aimed to investigate the artifact's applicability and its generated added value. In the course of the interview, we presented the expert both an excerpt of our results and randomly selected posts from the identified lead users. Thus, we wanted to find out if he can benefit from the lead users identified by the tool and if the expert agrees with the differentiation of the user types. Accordingly, by analyzing the selected posts as well as the excerpt of our results, the interviewee stated first that by means of the software tool he is now able to analyze a huge amount of social media data. Previously, his team only analyzed social media data by hand, which cost a lot of resources (e.g. time, human resources, etc.) and often led to incorrect results.

Second, the expert highlighted the distinction to be beneficial as it allows him to address users for different stages in the innovation process. Ultimately, he was aware of the two lead users that were identified by the tool. As he knew them in advance, he has already incorporated them successfully into the company's innovation process. However, there are also lead users, and therefore also innovative ideas and content, that were unknown to him so far. To reveal what the users are talking about we discussed randomly selected posts with him. Hence, the aim was to detect whether both the analysis of posts in the context of the innovation process and the differentiation of the lead users made sense from the practitioner's point of view. Consequently, by analyzing and discussing the provided posts, our interviewee already detected some ideas and suggestions for new products or for variations of existing ones. He contemplated the involvement of the users in the company's internal brainstorming/idea-finding process to talk about ideas for new products or about drawbacks of pre-existing ones. Therefore, he is committed to include the dissatisfaction in the identification approach here. To get a better understanding of how the expert came to this decision, we included a short excerpt of a selected post as an example:

> *"[…] The male velcro is facing the wrong way, which means its going to chaff like 60 grit sandpaper if you don't wear a thick rashie or wetsuit. […]"*

According to the interviewee, from this short excerpt, it can already be recognized that the user is pointing to a certain problem (*"male velcro is facing the wrong way"*) and therefore identify an unmet need. Accordingly, the expert identifies here a starting point for improving a specific product. The expert added that the selected posts also show that these identified lead users seem to be "*passionate individuals who are true ambassadors of the kitesurfing scene*" and can therefore be auspicious, prospective partners for the company's internal idea-finding process with the aim of uncovering existing problems, unmet needs and new product ideas.

Furthermore, after showing the interviewee the posts that our tool assigned to the phase "Development", he was enthusiastic about the high level of product-related knowledge. According to our interviewee this and the positiveness of user #7 could, for example, support the engagement of a promising cooperation regarding the development of new products. Including a lead user means the incorporation of the user's vast experience and knowledge. Thus, this cooperation could potentially lead to decreasing failure rates in product innovations. For demonstration purposes we also included here a selected post of a lead user. To exclude potential influence, mentioned competitors were removed here.

> "*Yes it is[,] if you like to ride powered, the early edition was a fave of mine[.] The next imho was completely different, lost all the flex that you want in choppy conditions and a lot of feel by going heavier on build. The newbronq crb 4 is back to the original, been riding the ts from […], great board but […] struggle [to] make a non spray board especially in chop conditions, in flat water it isn't an issue, the monk will cover most riders as mentioned. Demo I'd say"*

Phrases such as "*The newbronq crb 4 is back to the original*" led the expert to the assumption that this user exhibits a high product-related knowledge and thus is able to formulate precise solutions, which are both indicators for assigning this post/user to the "Development" phase. From the analysis of the posts the expert drew the conclusion that these lead users identified here, represent "*progressive riders who are continuously pushing their riding and the sport to new levels*". Therefore, the expert referred to the fact that these users propose to apply new materials and technologies to create constantly better performing products. Moreover, the interviewee stated that not only the identification of promising lead users will be supported with the tool but also analyzing the posts and contents of the respective users, which represents another added value for him.

He further noted that because of these different application areas, it is very constructive to identify lead users regarding the two innovation phases and therefore differentiate between them. The interviewee had already involved some kitesurfers in the company's (innovation) processes and therefore drew the conclusion that he would definitely involve the two lead users who exhibit the highest overall scores in both innovation phases (user#1 and user#2) but he also highlighted that it would be reluctant to focus only on them. He reasoned as follows: First, it is advisable for a company to include more than just two lead users so that multiple perspectives can be included in the company. Second, according to his experience, when a user is part of each innovation phase, the generation of ideas is inhibited if the user always keeps the development and its boundaries in mind. This would limit the venue for brainstorming that should be ensured in the "Idea generation" phase. Moreover, a user can have interesting new ideas, but s/he has too little product related knowledge to implement and develop them. This would also lead to a loss of new and potentially successful ideas. Furthermore, our interviewee noted that the lead users #1 and #2 exhibited good overall scores (and thus he would definitely include them), but he tended to prefer to incorporate the user with the highest level of product-related knowledge (user #6) in the "Development" phase. Overall, the expert confirmed that a clear differentiation between the two phases as well as a separated identification and consideration of lead users is necessary.

In addition to the evaluation with the expert, we discussed our results with three of the identified lead users. To strengthen the results of our tool and to make sure that the identified users are the appropriate lead users for the particular innovation phase we also examine the lead user perspective by conducting short interviews. This should also confirm and, where appropriate, extend the characteristics that we have found in literature. In doing so, the interviewees all confirmed being lead users in this online community and they postulated that they are all aware of making major contributions to idea generation and/or product development in line with our results. In addition, in surprising harmony they all mentioned similar characteristics (enthusiasm for the sport, high activity level in the online community and experience in the field) as essential for lead users. Only on the time needed to be an experienced kitesurfer there was no agreement. Two of the respondent lead users stated a minimum of 3 and 5 years to be an experienced kitesurfer. The remaining one quantified the respective time by the kitesurfing sessions performed and different kitesurfing locations visited.

With these statements our lead users supported the results of our research and in consequence validated the applicability of our approach. Furthermore, the essential characteristics fit to those we have found in literature and thus confirmed the characteristics we have included in our tool. The characteristic "enthusiasm for the sport " is implemented with "opinion leadership" in our approach. According to the current research literature "opinion leadership" is the ability to enable the flow of information and especially to diffuse it. Strong social relationships and a high level of engagement are premises for a functioning exchange of ideas and innovation (Pajo et al., 2014, 2017). Accordingly, a user who is motivated to build relationships in the community and thus exhibits high centrality scores is highly enthusiastic about the sport. The high activity level in the online community, calculated in our tool by the number of posts and transacted comments per user within the analyzed period, represents the second characteristic the lead users have mentioned. The "Experience in the field " can be partly covered by our characteristic "high product-related knowledge". However, the number of training hours, e.g., could also be included here.

Discussion of the results of evaluation

The evaluation has revealed not only the applicability but also the added value of our artifact in a practical environment. Thus, the in-depth interview with the head of marketing of our cooperating partner has highlighted that he is enthusiastic about the results as he benefits from them in many ways. First, the interviewee was able to assign the innovation potential to many posts by recognizing trends that were talked about in the posts, months before their realization. Moreover, the content of the comments has already made him aware of ideas on how to improve certain products in the company. Second, our expert has also confirmed that the high level of product related knowledge, vast experience and the associated ability to suggest concrete solutions for new products or their improvements is the main factor for a promising cooperation with a lead user in terms of the "Development" phase. The lead users identified with our tool have suggested concrete solutions based on their broad expertise about the products, the components and how they mesh with each other. As our expert has confirmed, this high product related knowledge can lead to decreasing failure rates in new product introductions or improvements because these users are aware of very specific facts such as every tiny change to a kite's profile can have enormous impact on its flight characteristics. Third, our expert also stated that he would be reluctant to focus only on the lead users who exhibit the highest overall scores in both innovation phases. He further highlighted that he would prefer in the "Development" phase the user with the highest level of product-related knowledge (user #6) and not user #1 or #2 who have higher overall scores. All in all, he confirmed that a clear differentiation of the two phases as well as a separated identification and consideration of lead users are necessary.

Finally, the interviews with the lead user confirm our approach and implicated further interesting perspectives and provided indications on how our approach can be further refined. In future research, these and other possible aspects and characteristics mentioned by the users have to be evaluated additionally.

Contribution for practice and research

Our investigation contributes to research and practice alike. As a contribution to practice, first companies can benefit from our comprehensive and modular approach. By applying our approach companies can resource-efficiently identify lead users which is an important process as the acquisition as well as the transferring costs of the information that are decisive for initiating innovation have tremendous influence on where innovation is created (Idota, 2019). Therefore, as lead users feature highly sticky information and are able to create innovations, organizations benefit from including them in their innovation process in order to overcome their information stickiness and so get to know the user's needs to solve (product) problems and reduce R&D costs. Thus, we stand out against other approaches that follow more resource intensive approaches (e.g. Brandtzaeg et al., 2016; Hung et al., 2011; Tuunanen et al., 2011).

Second, we created an artifact, respectively a tool, that is able to process a large number of social media data which can be repeated at any time as lead users are trend specific. This counteracts i.a. weaknesses of previous approaches that include only a small amount of data in the identification (Hau & Kang, 2016). By means of our tool, companies are able to start and monitor the current state of the identification process, display the analysis results by an intuitive, sortable table to easily enable either the selection of the overall lead users by the respective overall-scores or specific lead users by their identified results of an explicit characteristic and extract and persist the elicited results to various file formats (Excel, CSV) for later usage.

Third, a high degree of generalizability was taken into account to identify a lead user by considering several characteristics depending on the different circumstances of different companies. Thus, they are able to customize the identification process to their own needs by uploading their own dataset and applying all or a selected set of characteristics either following our pre-defined weights for each of the two innovation phases or individualize them as well. Hence, the weight of the respective characteristic is determined in the beginning of the analysis process. The identification process therefore can be specifically geared to users who dominate a single criterion or a combination of criteria. When a company, for instance, wants to focus more on lead users who express a sentiment of enjoyment in the innovation process, then the company can set the weights for dissatisfaction very low (or even to zero) and for enjoyment very high. Thus, we created an extensive, flexible, and resource-saving approach which can be easily applied by companies and which is based on objective traceable characteristics (different to other approaches that include self-assessment of respondents (Hienerth & Lettl, 2017)).

Fourth, the evaluation of our results has shown their contribution for our cooperating partner and therefore for practice. As the expert has highlighted, the tools enable him to turn away from analyzing social media data by hand, which costed a lot of resources (e.g. time, human resources, etc.) and often led to incorrect results. Further he identified lead users and therefore also innovative ideas and content, that had been unknown to him so far. So, our tool also allows to analyze posts and contents of the respective users and is thus able to detect new ideas and suggestions for new products or for variations of existing ones. For practice this can mean decreasing failure rates in product innovations.

In summary, companies aiming to identify different lead users for the particular phases in the innovation process can benefit from our comprehensive and modular artifact, since they are enabled to autonomously analyze large amounts of data and therefore automatically identify respective lead users adapted to the corporate's specific circumstances. Thus, we automated the lead user identification process, the most difficult and time-consuming aspect within the lead user method (Brem & Bilgram, 2015).

Furthermore, as outcomes of our DS research project we achieved theoretical contributions to research that go beyond the technical contribution (i.e., the artifact) and which are explained in more detail in the following. In terms of the DSR knowledge contribution framework of Gregor and Hevner (2013) the enhancements of our artifact over existing ones in the literature can be classified in the group of improvement (development of new solutions for known problems). DSR improvement projects make contributions to both prescriptive theory i.e. Design Theory (Gregor, 2006) and descriptive theory i.e. kernel theories (Gregor & Hevner, 2013). Thus, first, in terms of prescriptive theory our artifact contributes to a rather general and abstract knowledge base – “nascent design theory” (Gregor & Hevner, 2013). Therefore, design principles based on kernel theories and resulting design requirements were formulated and proposed in the section “Design principles for a lead user identification tool”. By applying them in the course of the design and development of the artifact followed by the demonstration and evaluation, an implicit empirical grounding of the design principles was achieved here (Heinrich & Schwabe, 2014). Our design principles capture design-related knowledge and can therefore support the development of further IS (design) theories and new artifacts. For designing further (identification) tools in related areas our design principles can be applied as we have formulated them generally by describing what the artifact should enable users to do and how the artifact should be built. By considering e.g. the design principle 3. Contextual adaptability, the importance of the context is highlighted in which the respective tool should be created. Since the context has a direct impact on the definition and implementation of the requirements, the alignment with the context also will lead to a more targeted identification tool. So, with the compilation of the design principles, we made a first step towards contributing to Design Theory in terms of theory for design and action (Gregor, 2006) as we comply with conditions as March and Smith (1995) and Hevner et al. (2004) pointed out under which a contribution to knowledge in DS has occurred: utility to a community of users, the novelty of the artifact and the persuasiveness of claims that it is effective. To take a next step towards mature Design Theory, according to Gregor and Jones (2007), a total of eight components are necessary. We have shown the “Purpose and scope” by means of the meta-requirements and the “Principles of form and function” by means of the design principles (see both Fig. 2). Furthermore, the latter is based on kernel theories (Lead User Theory and Innovation Theory) which entails the inclusion of a further component – “Justificatory knowledge”. Also, the “Constructs” that are described as the most basic levels of the theory are involved with the alignment to the characteristics of a lead user and the two phases of the innovation process. These components resulted in the “Expository instantiation”, i.e. in the application of the designed tool in a real world setting. However, the inclusion of the “artifact mutability”, the “testable propositions” and the “principles of implementation” explicitly and aligning the investigation on these eight components in general, as for example Böckle et al. (2021) have done, would need to be undertaken as a next step towards a mature contribution to Design Theory.

Beside that, our results also contribute to the kernel theory knowledge base regarding the social media theory as well as the innovation related theory. Moreover, our results contribute to different kernel theories by providing the following useful implications which previous investigations have barely considered until now. First, this study sheds a new light on the lead user construct itself – the core of Lead User Theory – as our investigation has shown that it is meaningful to differentiate lead users according to the different innovation phases as they have different competencies, characteristics and application areas. Until now, no distinction has been made in defining and characterizing lead users in terms of the innovation process. The basic model of Lead User Theory (Von Hippel, 1986) indicates indeed a distinction between lead users, but only against the background of whether the product innovation supported by the lead user is a novelty or one that requires commercially viable modifications and enhancements (Von Hippel, 1986). Our results highlighted that a separated consideration implicates a more targeted identification. If a lead user is active in one phase, then s/he can focus on either the objective generating many good and innovative ideas (see Idea generation) or on the objective developing explicit solutions for unmet needs (see Development) and applies his/her strengths accordingly. When lead users are examined and identified separately in the two phases, the generation of ideas is not inhibited by keeping the development and its boundaries in mind. Additionally, our approach takes also into account who have extraordinary new ideas but only less product-related knowledge and would therefore be excluded from prior identification approaches. Thus, our approach contracts a loss of new and potentially successful ideas. So, our study has revealed a new point by defining a lead user against the background of the purpose of its use (based on the innovation process) whereby we introduce a new dimension to the Lead User Theory. This can constitute an important new implication which includes that the definition of a lead user should not only focus on Von Hippels’ characteristics but also on the purpose of its use (Von Hippel, 1986).

Second our investigation contributes to the process of utilizing lead users included in the Lead User Theory. Von Hippel (1986) introduced a 4-step process – which has often been taken up in other studies (cf. Hung et al., 2011) –, including (1) the identification of an important market or technical trend, (2) the identification of a lead user leading that trend, (3) analyzing the lead user need data and (4) project lead user data onto the general market (Von Hippel, 1986). Our approach and results have shown that the (1) identification of a trend before (2) identifying a lead user for that respective trend is no longer deemed necessary as the identification of trend(s) can be included within the identification of corresponding lead users. Thus, the initial step (1) identification of trends is no longer considered as a necessary sequential premise for the (2) lead user identification, since the emerging trend is identified and considered simultaneously, resulting in a more flexible and easier-to-use process. Moreover, our approach provides the opportunity to consider multiple trends simultaneously, rather than being limited to one previously identified trend (Von Hippel, 1986). Therefore, multiple trends reflected in the underlying data can be dynamically considered when identifying lead users, enabling the identification of target-oriented lead users associated with each trend. Thus, the 4-step process can be enhanced in terms of its applicability and ease of use by enabling the automated identification of underlying trends when identifying accompanying lead users, as well as in terms of its functional scope by including multiple trends instead of solely considering the previously, manually identified trend.

Third, this study sheds another light to Lead User Theory and contributes to the automated identification of lead users in online communities in more specific (and thus a further contribution to the identification process). With this work at hand, we provide initially a comprehensive and structured overview of lead user characteristics based on the current research literature. Beyond that, we further technically realized these characteristics by means of an adaption of several machine learning methods (see section "Technical realization") and enriched the related Lead User Theory by establishing synergies of these research areas. Thus, future research in Lead User Theory will benefit from the advantages of automated analysis techniques and will therefore be supported by our concrete techniques for the identification of lead user characteristics. In addition, we distinguish ourselves from investigations that define and identify lead users by including only one or two characteristics (cf. Miao & Zhang, 2017; Tuarob & Tucker, 2014; Tuunanen et al., 2011), as our identification process enables an identification of lead users considering all identified characteristic. This enables the consideration of each relevant characteristic, allowing lead users to be identified in a more target-oriented and fine-grained manner. Moreover, to take a step further in the identification of respective lead users and in order to account the differentiation of them in the innovation process, we have adapted the identification process to incorporate priorities (weights) regarding the characteristics with respect to the different innovation phases. Consequently, contrary to the current research literature which treats all characteristics equally, we assign different weights to different lead user characteristics in the course of the identification process to make this process even more targeted.

Finally, for innovation theories our research identified relevant characteristics of users who can contribute to the different stages of the innovation process. Our results have shown that it is important to consider for what purpose a lead user should be engaged and to adjust the weighting of the identified characteristics accordingly. This has implications for the theories dealing with the process of innovation, e.g. the stage-gate model. By applying specific lead users within the stages preliminary and detailed investigations as well as in development, testing and validation, the rigid sequence of stages and gates can be broken up. By integrating the user's and therefore the external point of view the assessments at the go/kill checkpoints (i.e. gates) become less elaborate as the alignment with the external requirements is maintained constantly. This results in a more agile and target group-oriented approach. Based on this, the innovation process must be specified more concretely in terms of interactive value creation, especially the open innovation. Thus, including different lead users adds new tasks for companies in the innovation process. These different lead user types can be taken into account by introducing process variants.

Conclusion

In the existing literature there are a lot of different lead user identification approaches, but these investigations only covered a limited point of view as they either focus on only a few lead user characteristics (Martínez-Torres, 2014), include a very small amount of data (Hau & Kang, 2016) or base their approach on the self-assessment of users (Hienerth & Lettl, 2017). This problem is further compounded by the tremendous amount of online community data which makes it even more difficult, costly and time-consuming to identify lead users. We approached this research gap by introducing an automated and – according to our interviewed expert – effective method for identifying lead users. After consulting the research literature, we focused on two main phases of the user innovation process (A) the "Idea generation" of an innovation and (B) the "Development" of an innovation. In both cases (A and B), a lead user is a valuable resource for companies. Furthermore, we have demonstrated that six different characteristics (trend leadership, dissatisfaction, enjoyment, high level of activity, product related knowledge, opinion leadership) are prevalent in existing research literature regarding lead user identification in online communities (see RQ1). Based on this, we further designed and implemented a tool that, on the one hand, combines all of the aforementioned characteristics and, on the other hand, considers the fact that lead users can be applied in different phases of the innovation process (see RQ2). To demonstrate the applicability of our artifact we applied it to 11,481 contributions of 945 users from a popular online forum for kiteboarding. After identifying the lead users, we evaluated our results by interviewing the respective lead users as well as an expert. In conclusion, following the DS research, we derived numerous contributions for both, theory (kernel theories: Innovation- and Lead User Theory; Design Theory: Design Principles) and practice (e.g., the artifact) (see RQ3).

In the previous section (see "Contribution for practice and research") we have shown that companies can benefit from our comprehensive and modular artifact, with which large amounts of data can be analyzed adapted to the corporate's specific circumstances with the aim of identifying different lead users for the particular phases in the innovation process. Thus, we automated the lead user identification process, the most difficult and time-consuming aspect within the lead user method. Furthermore, we have highlighted how our investigation made a first step towards contributing to Design Theory (theory for design and action (Gregor, 2006)) by formulating four design principles. These design principles (comprehensive characteristics consideration, using inter-subjectively verifiable identification methods, contextual adaptability and repeatability) can support the design of further user identification tools. Beside that we also highlighted our contribution to the kernel theories: Our study has revealed a new point by defining a lead user against the background of the purpose of its use (based on the innovation process) whereby we have introduced a new dimension to the Lead User Theory. Moreover, we enhanced Von Hippel's 4-step lead user utilization process in terms of its applicability and ease of use by enabling the automated identification of underlying trends when identifying accompanying lead users, as well as in terms of its functional scope by including multiple trends instead of solely considering the previously, manually identified trend. Finally with respect to the current Lead User Theory which treats all characteristics equally, we have assigned different weights to different lead user characteristics in the course of the identification process to make it even more targeted. Regarding the Innovation Theory the rigid sequence of stages and gates can be broken up and further parallelized by applying specific lead users within the different stages. However, our research is not without limitations. We have identified the characteristics that are decisive for a lead user in the current research literature. It is possible that there are further characteristics distinctive for a lead user that we have not considered. During our research we came upon areas of further research. In terms of a further evaluation of our results, we are intent on carrying out a study to assess the completeness and usefulness of our approach with other cooperating partners. Further, as the users noted in the interviews that experience in the respective field of application is important and we only partially cover this with the characteristic "product-related knowledge", the question "At what point can a lead user be seen as experienced?" may be subject of future work.

Funding Open Access funding enabled and organized by Projekt DEAL.

References

Aggarwal, C. C. (2011). An introduction to social network data analytics. In C. Aggarwal (Ed.), *Social network data analytics* (pp. 1–15). Boston. https://doi.org/10.1007/978-1-4419-8462-3_1

Aggarwal, C. C., & Zhai, C. (2012). A survey of text classification algorithms. In *Mining text data* (pp. 163–222). Boston. https://doi.org/10.1007/978-1-4614-3223-4_6

Al-Zu'bi, Z. B. M., & Tsinopoulos, C. (2012). Suppliers versus lead users: Examining their relative impact on product variety. *Journal of Product Innovation Management, 29*(4), 667–680. https://doi.org/10.1111/j.1540-5885.2012.00932

AlFalahi, K., Atif, Y., & Abraham, A. (2014). Models of Influence in Online Social Networks. *International Journal of Intelligent Systems, 29*(2), 161–183. https://doi.org/10.1002/int.21631

Amabile, T. M. (1988). A model of creativity and innovation in organizations. *Research in Organizational Behavior, 10*(1), 123–167.

Autio, E., Dahlander, L., & Frederiksen, L. (2013). Information exposure, opportunity evaluation, and entrepreneurial action: An investigation of an online user community. *Academy of Management Journal, 56*(5), 1348–1371. https://doi.org/10.5465/amj.2010.0328

Axtell, C. M., Holman, D. J., Unsworth, K. L., Wall, T. D., Waterson, P. E., & Harrington, E. (2000). Shopfloor innovation: Facilitating the suggestion and implementation of ideas. *Journal of Occupational and Organizational Psychology, 73*(3), 265–285. https://doi.org/10.1348/096317900167029

Belz, F. M., & Baumbach, W. (2010). Netnography as a method of lead user identification. *Creativity and Innovation Management, 19*(3), 304–313. https://doi.org/10.1111/j.1467-8691.2010.00571.x

Böckle, M., Bick, M., & Novak, J. (2021). Toward a design theory of user-centered score mechanics for gamified competency development. *Information Systems Management,* 1–27. https://doi.org/10.1080/10580530.2021.1975852

Boyd-Graber, J., Mimno, D., & Newman, D. (2014). Care and feeding of topic models: Problems, diagnostics, and improvements. In E. M. Airoldi, D. Blei, E. A. Erosheva, & S. E. Fienberg (Eds.), *Handbook of Mixed Membership Models and Their Applications* (pp. 3–34). CRC Press.

Brandtzaeg, P. B., Haugstveit, I. M., Lüders, M., & Følstad, A. (2016). How should organizations adapt to youth civic engagement in social media? A lead user approach. *Interacting with Computers, 28*(5), 664–679. https://doi.org/10.1093/iwc/iwv041

Breitsohl, J., Kunz, W. H., & Dowell, D. (2015). Does the host match the content? A taxonomical update on online consumption communities. *Journal of Marketing Management, 31*(9–10), 1040–1064. https://doi.org/10.1080/0267257X.2015.1036102

Breitsohl, J., Roschk, H., & Feyertag, C. (2018). *Consumer brand bullying behaviour in online communities of service firms. In: Service business development*. Springer, 289–312. https://doi.org/10.1007/978-3-658-22424-0_13

Brem, A., & Bilgram, V. (2015). The search for innovative partners in co-creation: Identifying lead users in social media through netnography and crowdsourcing. *Journal of Engineering and Technology Management, 37*, 40–51. https://doi.org/10.1016/j.jengtecman.2015.08.004

Brem, A., Bilgram, V., & Gutstein, A. (2018). Involving lead users in innovation: A structured summary of research on the lead user method. *International Journal of Innovation and Technology Management, 15*(03). https://doi.org/10.1142/S0219877018500220

Chandra, L., Seidel, S., & Gregor, S. (2015). Prescriptive knowledge in IS research: Conceptualizing design principles in terms of materiality, action, and boundary conditions. In *48th Hawaii International Conference on System Sciences* (4039–4048). https://doi.org/10.1109/HICSS.2015.485

Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J., & Blei, D. (2009). Reading tea leaves: How humans interpret topic models. *Advances in Neural Information Processing Systems, 22*, 288–296.

Chen, J., Yin, X., & Mei, L. (2018). Holistic innovation: An emerging innovation paradigm. *International Journal of Innovation Studies, 2*(1), 1–13. https://doi.org/10.1016/j.ijis.2018.02.001

Chen, X., Hu, X., Wang, Y., & Tao, D. (2019). Extending Lead Users to Average User Innovation: A Novel Segmentation Framework Based on Users' Innovativeness. *Proceedings of The IEEE International Conference on Industrial Cyber Physical Systems (ICPS)*. https://doi.org/10.1109/ICPHYS.2019.8780274

Cooper, R. G. (1996). Overhauling the new product process. *Industrial Marketing Management, 6*(25), 465–482. https://doi.org/10.1016/S0019-8501(96)00062-4

Crain, S. P., Zhou, K., Yang, S.-H., & Zha, H. (2012). *Dimensionality reduction and topic modeling: From latent semantic indexing to latent dirichlet allocation and beyond. In: Mining text data*. Springer. 129–161. https://doi.org/10.1007/978-1-4614-3223-4_5

Crawford, C. M. (1994). *New products management*. Irwin.

Dahlander, L., & Frederiksen, L. (2012). The core and cosmopolitans: A relational view of innovation in user communities. *Organization Science, 23*(4), 988–1007. https://doi.org/10.1287/orsc.1110.0673

de Reuver, M., Sørensen, C., & Basole, R. C. (2018). The digital platform: A research agenda. *Journal of Information Technology, 33*(2), 124–135. https://doi.org/10.1057/s41265-016-0033-3

Dong, J. Q., & Wu, W. (2015). Business value of social media technologies: Evidence from online user innovation communities. *The Journal of Strategic Information Systems, 24*(2), 113–127. https://doi.org/10.1016/j.jsis.2015.04.003

Eickhoff, M., & Neuss, N. (2017). Topic modelling methodology: its use in information systems and other managerial disciplines. *Proceedings of the European Conference of Information Systems (ECIS)*. 1327–1347.

Ernst, M., Brem, A., & Voigt, K.-I. (2013). Innovation Management, Lead Users and Social Media: Introduction of a Conceptual Framework for Integrating Social Media Tools in Lead User Management. *Social media in strategic management, Advanced Series in Management, 11,* 169–195. https://doi.org/10.1108/S1877-6361(2013)0000011013.

Feldman, R., & Sanger, J. (2007). *The text mining handbook: Advanced approaches in analyzing unstructured data*. Cambridge University Press.

Fisher, G. (2019). Online communities and firm advantages. *Academy of Management Review, 44*(2), 279–298. https://doi.org/10.5465/amr.2015.0290

Franke, N., Von Hippel, E., & Schreier, M. (2006). Finding commercially attractive user innovations: A test of lead-user theory. *Journal of Product Innovation Management, 23*(4), 301–315. https://doi.org/10.1111/j.1540-5885.2006.00203.x.

Franke, N., & von Hippel, E. A. (2003). *Finding Commercially Attractive User Innovations: A Performance Evaluation of the "lead User Construct"*. https://doi.org/10.2139/ssrn.367140

Freeman, C., & Soete, L. (1997). *The economics of industrial innovation*. Psychology Press.

Füller, J., Jawecki, G., & Mühlbacher, H. (2007). Innovation creation by online basketball communities. *Journal of Business Research, 60*(1), 60–71. https://doi.org/10.1016/j.jbusres.2006.09.019

Gallaugher, J., & Ransbotham, S. (2010). Social media and customer dialog management at Starbucks. *MIS Quarterly Executive, 9*(4), 197–212.

Garrett, J. J. (2010). *The elements of user experience: user-centered design for the web and beyond*. Pearson Education.

Gawer, A., & Cusumano, M. A. (2014). Industry Platforms and Ecosystem Innovation. *Journal of Product Innovation Management, 31*(3), 417–433. https://doi.org/10.1111/jpim.12105

Gregor, S. (2006). The nature of theory in information systems. *MIS quarterly*, *30*(3), 611–642. https://doi.org/10.2307/25148742

Gregor, S., & Jones, D. (2007). The anatomy of a design theory. *Journal of the Association for Information Systems, 8*(5), 313–335.

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly, 37*(2), 337–355. https://doi.org/10.25300/MISQ/2013/37.2.01

Gregor, S., Chandra Kruse, L., & Seidel, S. (2020). Research perspectives: The anatomy of a design principle. *Journal of the Association for Information Systems*, 21(6). https://doi.org/10.17705/1jais.00649

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function using NetworkX. In G. Varoquaux, T. Vaught, & J. Millman (Eds.), *Proceedings of the 7th Python in Science Conference* (SciPy2008) (pp. 11–15).

Han, J., Pei, J., & Kamber, M. (2006). *Data Mining*. Elsevier Science.

Hau, Y. S., & Kang, M. (2016). Extending lead user theory to users' innovation-related knowledge sharing in the online user community: The mediating roles of social capital and perceived behavioral control. *International Journal of Information Management, 36*(4), 520–530. https://doi.org/10.1016/j.ijinfomgt.2016.02.008

Herstatt, C. (1999). Theorie und Praxis der frühen Phasen des Innovationsprozesses. *Management Zeitschrift Industrielle Organisation, 68*(10), 80–91.

Heinrich, P., Schwabe, G. (2014). Communicating nascent design theories on innovative information systems through multi-grounded design principles. In M. C. Tremblay, D. VanderMeer, M. Rothenberger, A. Gupta, V. Yoon (eds.), *Advancing the impact of design science: Moving from theory to practice* (vol. 8463, pp. 148–163). DESRIST 2014. Lecture Notes in Computer Science. Cham: Springer. https://doi.org/10.1007/978-3-319-06701-8_10

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems, 19*(2), 87–92.

Hevner, A. R., Salvatore, M. T., Jinsoo, P., & Sudha, R. (2004). Design science in information systems research. *MIS Quarterly, 28*(1), 75–105. https://doi.org/10.2307/25148625

Heyer, G., Quasthof, U., & Wittig, T. (2006). *Text mining: Wissensrohstoff tex*. W3l, Herdecke, 18

Hienerth, C., & Lettl, C. (2017). Perspective: Understanding the nature and measurement of the lead user construct. *Journal of Product Innovation Management, 34*(1), 3–12. https://doi.org/10.1111/jpim.12318

Hung, C.-L., Chou, J.C.-L., & Dong, T.-P. (2011). Innovations and communication through innovative users: An exploratory mechanism of social networking website. *International Journal of Information Management, 31*(4), 317–326. https://doi.org/10.1016/j.ijinfomgt.2010.12.003

Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Paper presented at the Proceedings of the International AAAI Conference on Web and Social Media.*

Idota, H. (2019). Empirical Study on Consumer Innovation by Using Social Media in Japan. In *Paper presented at the UK Academy for Information Systems Conference Proceedings 2019.*

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons, 53*(1), 59–68. https://doi.org/10.1016/j.bushor.2009.09.003

King, A., & Lakhani, K. R. (2013). Using open innovation to identify the best ideas. *MIT Sloan Management Review, 55*(1), 41–48.

Lakhani, K. R., Garvin, D. A., & Lonstein, E. (2010). Topcoder (a): *Developing software through crowdsourcing* (January 15, 2010). Harvard Business School General Management Unit Case No. 610-032, Available at SSRN:https://ssrn.com/abstract=2002884

Latora, V., & Marchiori, M. (2007). A measure of centrality based on network efficiency. *New Journal of Physics, 9*(6), 188–199.

Li, Z., & Tang, H. (2016). Identifying Lead User in Mass Collaborative Innovation Community: Based on Knowledge Supernetwork. In *Paper presented at the International Symposium on Knowledge and Systems Sciences.*

Lilien, G. L., Morrison, P. D., Searls, K., Sonnack, M., & Hippel, E. V. (2002). Performance assessment of the lead user idea-generation process for new product development. *Management Science, 48*(8), 1042–1059.

Lüthje, C. (2004). Characteristics of innovating users in a consumer goods field: An empirical study of sport-related product consumers. *Technovation, 24,* 683–695. https://doi.org/10.1016/S0166-4972(02)00150-5.

Mahr, D., & Lievens, A. (2012). Virtual lead user communities: Drivers of knowledge creation for innovation. *Research Policy, 41*(1), 167–177. https://doi.org/10.1016/j.respol.2011.08.006

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems, 15*(4), 251–266. https://doi.org/10.1016/0167-9236(94)00041-2

Marchi, G., Giachetti, C., & De Gennaro, P. (2011). Extending lead-user theory to online brand communities: The case of the community Ducati. *Technovation, 31*(8), 350–361. https://doi.org/10.1016/j.technovation.2011.04.005.

Martínez-Torres, M. R. (2014). Analysis of open innovation communities from the perspective of social network analysis. *Technology Analysis & Strategic Management, 26*(4), 435–451. https://doi.org/10.1080/09537325.2013.851378

Marxt, C., & Hacklin, F. (2005). Design, product development, innovation: all the same in the end? A short discussion on terminology. *Journal of Engineering Design, 16*(4), 413–421.

McCallum, A. K. (2002). A machine learning for language toolkit. *MALLET, 15*(2), 131–136.

Miao, Y., & Zhang, H. (2017). A biclustering-based lead user identification methodology applied to xiaomi. *Proceedings of the Fourth International Forum on Decision Sciences*. https://doi.org/10.1007/978-981-10-2920-2_80

Muller, M., Ehrlich, K., Matthews, T., Perer, A., Ronen, I., & Guy, I. (2012). Diversity among enterprise online communities: Collaborating, teaming, and innovating through social media. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2815–2824*,. https://doi.org/10.1145/2207676.2208685

Obar, J. A., & Wildman, S. S. (2015). Social media definition and the governance challenge-an introduction to the special issue. *Telecommunications Policy, 39*(9), 745–750. https://doi.org/10.2139/ssrn.2663153

Pajo, S., Vandevenne, D., & Duflou, J. R. (2017). Automated feature extraction from social media for systematic lead user identification. *Technology Analysis & Strategic Management, 29*(6), 642–654. https://doi.org/10.1080/09537325.2016.1220517.

Pajo, S., Verhaegen, P.-A., Vandevenne, D., & Duflou, J. R. (2014). Lead User Identification through Twitter: Case Study for Camera Lens Products. In *Proceedings of Nord Design 2014.*

Paulus, P. B., Putman, V. L., Dugosh, K. L., Dzindolet, M. T., & Coskun, H. (2002). Social and cognitive influences in group brainstorming: Predicting production gains and losses. *European Review of Social Psychology, 12*(1), 299–325. https://doi.org/10.1080/14792772143000094.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Piller, F. T. (2006). User Innovation: Der Kunde kann's besser. In O. Drossou, S. Krempl & A. Poltermann (Eds.), *Die wunderbare Wissensvermehrung: Wie Open Innovation unsere Welt revolutioniert* (pp. 85–97). Heise.

Poetz, M. K., & Schreier, M. (2012). The value of crowdsourcing: can users really compete with professionals in generating new product ideas? *Journal of Product Innovation Management, 29*(2), 245–256. https://doi.org/10.1111/j.1540-5885.2011.00893.x

Preece, J., & Maloney-Krichmar, D. (2003). Online communities. In J. Jacko, & A. Sears (Eds.), *Handbook of human-computer interaction* (pp. 596–620). Lawrence Erlbaum Associates Inc.

Ramaswamy, V. (2010). Competing through co-creation: Innovation at two companies. *Strategy & Leadership, 37*(2), 32–37.

Roy, R. (2018). Role of relevant lead users of mainstream product in the emergence of disruptive innovation. *Technological Forecasting and Social Change, 129*, 314–322. https://doi.org/10.1016/j.techfore.2017.09.036

Saldanha, F. P., & Pozzebon, M. (2015). *Fiat Mio: The project that embraced open innovation, crowdsourcing and creative commons in the automotive industry* (Vol. 13, no. 1). Harvard Business School Press.

Schaarschmidt, M., Stol, K.-J., Walsh, G., & Bertram, M. (2019). Lead Users' Innovative Work Behavior in Digital Platform Ecosystems: A Large Scale Study of App Developers. *Proceedings of International Conference on Information Systems.*

Schmid, I. M. (2020). INFLUENTIAL USERS IN SOCIAL MEDIA NETWORKS: A LITERATURE REVIEW, *Proceedings of the European Conference on Information Systems (ECIS).*

Schofield, A., & Mimno, D. (2016). Comparing apples to apple: The effects of stemmers on topic models. *Transactions of the Association for Computational Linguistics, 4*, 287–300. https://doi.org/10.1162/tacl_a_00099

Shang, S. S., Wu, Y. L., & Li, E. Y. (2017). Field effects of social media platforms on information-sharing continuance: Do reach and richness matter? *Information & Management, 54*(2), 241–255. https://doi.org/10.1016/j.im.2016.06.008

Schreier, M., Oberhauser, S., & Prügl, R. (2007). Lead users and the adoption and diffusion of new products: Insights from two extreme sports communities. *Marketing Letters, 18*(1–2), 15–30. https://doi.org/10.1007/s11002-006-9009-3

Tuarob, S., & Tucker, C. S. (2014). Discovering next generation product innovations by identifying lead user preferences expressed through large scale social media data. *Paper Presented at the International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. https://doi.org/10.1115/DETC2014-34767

Tuunanen, T., Bragge, J., Haivala, J., Hui, W., & Virtanen, V. (2011). A method for recruitment of lead users from virtual communities to innovate it enabled services for consumers in global markets. *Pacific Asia Journal of the Association for Information Systems, 3*(2), 31–56. https://doi.org/10.17705/1pais.03202

Ulrich, K. T., & Eppinger, S. D. (1995). *Product Design and Development*. McCraw-Hill. Inc.

Unsworth, K. L., Brown, H., & McGuire, L. (2000). Employee innovation: The roles of idea generation and idea implementation. *Proceedings in SIOP Conference.*

Von Hippel, E. (1986). Lead users: A source of novel product concepts. *Management Science, 32*(7), 791–805. https://doi.org/10.1287/mnsc.32.7.791

Von Hippel, E. (1994). "Sticky information" and the locus of problem solving: Implications for innovation. *Management Science, 40*(4), 429–439. https://doi.org/10.1287/mnsc.40.4.429

Von Hippel, E. (2005). *The democratization of innovation*. Mass.

von Hippel, E. (2007). The Sources of Innovation. In C. Boersch, & R. Elschen (Eds), *Das Summa Summarum des Management* (pp. 111–120). Gabler. https://doi.org/10.1007/978-3-8349-9320-5_10

Von Hippel, E., Franke, N., & Prügl, R. (2009). Pyramiding: Efficient search for rare subjects. *Research Policy, 38*(9), 1397–1406. https://doi.org/10.1016/j.respol.2009.07.005

Von Hippel, E., & Katz, R. (2002). Shifting innovation to users via toolkits. *Management Science, 48*(7), 821–833.

Wagner, P., & Piller, F. T. (2011). *Mit der Lead-User-Methode zum Innovationserfolg: ein Leitfaden zur praktischen Umsetzung* (1–28). CLIC.

Walls, J. G., Widmeyer, G. R., & El Sawy, O. A. (1992). Building an information system design theory for vigilant EIS. *Information Systems Research, 3*(1), 36–59. https://doi.org/10.1287/isre.3.1.36

West, J., & Bogers, M. (2014). Leveraging external sources of innovation: A review of research on open innovation. *Journal of Product Innovation Management, 31*(4), 814–831. https://doi.org/10.1111/jpim.12125

Wolfe, R. A. (1994). Organizational innovation: Review, critique and suggested research directions. *Journal of management studies, 31*(3), 405–431. https://doi.org/10.1111/j.1467-6486.1994.tb00624.x.

Xie, P., & Xing, E. P. (2013). Integrating document clustering and topic modeling. *arXiv preprint* arXiv:1309.6874.

Ye, H., & Kankanhalli, A. (2018). User Service Innovation on Mobile Phone Platforms: Investigating Impacts of Lead Userness, Toolkit Support, and Design Autonomy. *MIS quarterly, 42*(1). 165–187. https://doi.org/10.25300/MISQ/2018/12361

Zheng, T.-T., & Zhou, Q.-J. (2017). Analysis of the Motive Mechanism of User Innovation Based on System Dynamics. *Proceedings of the Annual International Conference on Management, Economics and Social Development, 21*, 17–23.

2.6 Artificial Intelligence in Augmentative and Alternative Communication Systems – A Literature-Based Assessment and Implications of Different Conversation Phases and Contexts

Adressierte Zielsetzungen	Systematische Analyse der Potenziale KI-basierter Ansätze zur Unterstützung digitaler AAC-Systeme im Hinblick auf die Berücksichtigung spezifischer Gesprächsphasen und -Kontexte. Ziel ist es, Einschränkungen bestehender Systeme hinsichtlich der Anpassungsfähigkeit an unterschiedliche kommunikative Kontexte (z. B. formelle vs. informelle) sowie Phasen eines Gesprächs (z. B. Anfang, Schluss) zu identifizieren und einzuordnen. Auf dieser Grundlage werden bestehende Forschungslücken offengelegt und zentrale Anforderungen für die kontextsensitive, personalisierte Entwicklung intelligenter Kommunikationsagenten im Bereich der AAC abgeleitet.	***ZS9***
Abgeleitete Forschungsfragen	Decken die in der wissenschaftlichen Literatur vorgeschlagenen KI-Anwendungen für AAC-Systeme die Gesprächsphasen ‚Beginn', ‚Mitte ' und ‚Abschluss' sowie die Gesprächskontexte ‚Formell' und ‚Informell' in angemessener Weise ab?	***FF9***
	Welche Implikationen ergeben sich für den Einsatz von KI im Kontext von AAC-Systemen hinsichtlich der Berücksichtigung von Gesprächsphasen und -kontexten?	***FF10***
Genutzte Forschungsmethode	**Systematischer Literaturüberblick (Systematic Literature Review)** Die Studie folgt methodisch den Empfehlungen etablierter Frameworks zur Literaturauswertung, insbesondere von Cooper (1988), Cooper & Hedges (1994), Vom Brocke et al. (2015) sowie Webster & Watson (2002). Die Literaturanalyse erfolgte in mehreren Phasen: • Systematische Suche nach einschlägigen wissenschaftlichen Arbeiten. • Reduktions- und Auswahlverfahren zur Eingrenzung relevanter Beiträge. • Inhaltsanalytische Auswertung auf Basis der qualitativen Inhaltsanalyse nach Mayring (2004). Die Analyse zielte auf drei Ebenen ab: • Ableitung zentraler KI-Potenziale für AAC-Systeme und Skizzierung KI-basierter Lösungen zur Realisierung dieser Potenziale. • Zuordnung dieser Potenziale zu spezifischen Gesprächsphasen (Eröffnung, Mitte, Abschluss) und Gesprächskontexten (formell/informell). • Identifikation von Forschungslücken durch systematische Kontrastierung in einer Konzeptmatrix.	
Kernergebnisse (Überblick)	Durch die Literaturanalyse wurden vier zentrale Kategorien generativer KI-Potenziale in AAC-Systemen identifiziert: 1. *Aufrechterhaltung des Gesprächsverlaufs*, 2. *Kontextsensitivität* (z. B. Ort, Thema), 3. *Adaptivität gegenüber dem Gesprächspartner*, 4. *Personalisierung gegenüber dem AAC-Nutzer*.	

<table>
<tr><td></td><td>Die Potenziale wurden systematisch in einer Konzeptmatrix den Gesprächsphasen (Beginn, Mitte, Abschluss) sowie Gesprächskontexten (formell vs. informell) zugeordnet. Die Analyse zeigt:
• Der Großteil bestehender KI-basierter AAC-Systeme fokussiert sich auf einzelne Phasen (v. a. Gesprächsbeginn).
• Informelle Gesprächskontexte werden besser abgedeckt als formale.
Daraus ergibt sich ein klarer Forschungsbedarf hinsichtlich kontextübergreifender, adaptiver Kommunikationsunterstützung.
Die Studie diskutiert abschließend konkrete Ansatzpunkte zur Schließung dieser Lücken, insbesondere durch gezielte Nutzung und Erweiterung generativer KI-Ansätze um kontextuelle und personenspezifische Steuerungselemente.</td></tr>
<tr><td>Publikationsort</td><td>Das Paper wurde 2023 auf der 31. European Conference on Information Systems, Kristiansand, Norwegen (ECIS) veröffentlicht und ist in der AISeL als peer-reviewtes Konferenzpapier gelistet.</td></tr>
<tr><td>Ranking VHB JQ 4</td><td>A</td></tr>
<tr><td>Autoren und Anteile</td><td>Daniel Konadl 50%
Janik Wörner 30%
Lucas Luttner 10%
Susanne Leist 10%</td></tr>
<tr><td>DOI</td><td>http://doi.org/10.5283/epub.54202</td></tr>
</table>

Tabelle 20. Fact Sheet zur 6. Veröffentlichung.

5-11-2023

ARTIFICIAL INTELLIGENCE IN AUGMENTATIVE AND ALTERNATIVE COMMUNICATION SYSTEMS - A LITERATURE-BASED ASSESSMENT AND IMPLICATIONS OF DIFFERENT CONVERSATION PHASES AND CONTEXTS

Daniel Konadl
University of Regensburg, daniel.konadl@wiwi.uni-regensburg.de

Janik Wörner
University of Regensburg, janik.woerner@ur.de

Lucas Luttner
University of Regensburg, Lucas.Luttner@ur.de

Susanne Leist
University of Regensburg, susanne.leist@ur.de

ARTIFICIAL INTELLIGENCE IN AUGMENTATIVE AND ALTERNATIVE COMMUNICATION SYSTEMS - A LITERATURE-BASED ASSESSMENT AND IMPLICATIONS OF DIFFERENT CONVERSATION PHASES AND CONTEXTS

Research Paper

Daniel Konadl, University of Regensburg, Regensburg, Germany, Daniel.Konadl@ur.de

Janik Wörner, University of Regensburg, Regensburg, Germany, Janik.Woerner@ur.de

Lucas Luttner, University of Regensburg, Regensburg, Germany, Lucas.Luttner@ur.de

Susanne Leist, University of Regensburg, Regensburg, Germany, Susanne.Leist@ur.de

Abstract

Even though AAC systems and corresponding AI approaches have been investigated within the extant research, they still show remarkable drawbacks, resulting in a low prevalence among speech-impaired individuals. As the suggestions and adaptions proposed by AI within AAC systems may show insufficiencies in certain situations (e.g., unreliable suggestions, low conversational rates, unauthentic adaptions towards users), we take an up-close look, especially at the conversation phases and contexts in which the supporting AI is applied. Therefore, we have conducted a systematic literature review as well as a literature analysis. Thereby, we could reveal that there are indeed several gaps within the extant research on AI regarding the coverage of the conversational context "informativeness" and the conversation phases "beginning" and "closing". To dismantle the existing communication barriers that speech-impaired individuals suffer from, several implications for investigating AI in the context of AAC systems are derived and proposed for future (IS) research.

Keywords: Augmentative and Alternative Communication Systems, Artificial Intelligence, Conversation Phases, Conversational Contexts, Systematic Literature Review.

1 Motivation

"Perhaps the single quality most central to humanness is the ability to exchange thoughts, ideas, and feelings with others" (Hourcade et al., 2004, p. 235). In this context, speech is seen as an individual's most important instrument to be in contact with its human surroundings (cf. Fritzell, 1996; Kane et al., 2017; Pollak and Gallagher, 1989). However, individuals that suffer from severe physical disabilities or brain injury are not able to control their oral-respiratory musculature sufficiently for speech (Allen, 2005; Vanderheiden, 1983). This restriction of verbal communication, and hence, the "separation from the mainstream of society" (Hourcade et al., 2004, p. 235) holds true for people with damages to the vocal tract or other impairments (e.g., aphasia, autism) affecting speech as well (Allen, 2005).

To get in exchange with their surroundings, speech-impaired individuals need to apply Augmentative and Alternative Communication (AAC) which combines strategies, symbols, and techniques to promote goal-directed communication (American Speech Language Hearing Association, 1991). AAC comprises different analogue forms of communication (i.a., sign language, picture boards). In the context of sign language, expressions are formed using mimics and gestures. But these expressions can easily be overlooked or misinterpreted when the interlocutors have little or no experience in sign language (Boyes-Braem et al., 1994). Compared to vocal language, sign language lacks expressiveness because it provides only a few prepositions and conjunctions. Furthermore, it contains no genus markers to differentiate between masculinum and femininum for nouns (Louis-Nouvertné, 2001).

To give individuals with speech impairments a voice and enable them to express themselves using vocal language, digital AAC solutions have emerged that enable the translation of an individual's intended meaning into speech and subsequent voice outputs (cf. Bradshaw, 2013). The positive influence of these voice-generating AAC systems on the interaction behaviour of speech-impaired individuals has already been proven in numerous studies (e.g., Brady, 2000; Desai et al., 2014; Schepis et al., 1996). Despite the positive influence and although these AAC systems manage to meet several design requirements *(see Section 2.2)*, they still show drawbacks that research has attempted to address by various Artificial Intelligence (AI) approaches (e.g., Dempster et al., 2010; Neamtu et al., 2019; Obiorah et al., 2021). While it is undeniable that these AI approaches could increase the practicability of AAC systems (e.g., Klauer et al., 2021; Laxmidas et al., 2021; Obiorah et al., 2021), these systems nevertheless have only a low prevalence among the target group as they still face different problems.

For example, AAC systems propose only low conversation rates as the suggestions generated, and the utterances composed by AI often do not fit the concrete situation in which the AAC user takes part. As a consequence of unreliable suggestions, speech-impaired individuals have to exert effort to produce an utterance or have to adjust the suggestions to generate an appropriate contribution to the conversation (e.g., Dempster et al., 2010; Laxmidas et al., 2021; Obiorah et al., 2021). The possibilities AI approaches propose concerning personalisation towards the user of the AAC system show limitations as well (e.g., within situation-depending adaptions of the speech output tonality). The adaptions proposed by the AI are often not perceived as authentic by the conversation participants given a certain situation because certain situational factors (e.g., time progression and changes of involvement and attitude towards a conversational topic) are not incorporated into the adaption (e.g., Ascari et al., 2020; Mills et al., 2014; Shen et al., 2022). Because the suggestions and adaptions proposed by AI approaches within AAC systems show insufficiencies in certain situations, this suggests that the AI approaches in particular need to be investigated and adapted concerning the conversations in which they are applied. Indeed, conversations are performed in different conversation phases ("beginning", "middle" and "closing" of a conversation) and within different contexts (e.g., formal conversations within meetings at the workplace vs. informal conversations with friends within spare-time activities).

So, while the beginning of a conversation with an acquaintance in an informal context does not require the assessment of the conversation partner (i.a., personality, attitudes, conversational behaviour), the conversational beginning with a business customer in a formal negotiation within a meeting requires the assessment of him or her. In the same vein, while the beginning of a conversation is comparably standardised (cf. Henne and Rehbock, 2012; Spiegel and Spranz-Fogasy, 2001), the middle of conversations may be more complex as the involved conversation participants may aim at diverging conversational goals, requiring mutual and coordinated conversational acting until a targeted and mutually accepted state of the conversation is reached. As can be concluded, the different conversation phases and conversational contexts consequently lead to different demands and requirements of a supporting AI approach that need to be met to comprehensively and optimally support a speech-impaired individual.

With the work at hand, we address these issues by conducting a systematic literature review (Cooper, 1988; Vom Brocke et al., 2015; Webster and Watson, 2002) to establish an initial basis for the underlying research topic and thus develop a deeper understanding of the research field of AI in the context of AAC systems as well as to identify existing research problems and gaps and justify the relevance of addressing them (Vom Brocke et al., 2015). In this context, we identify AI-based potentials in view of AAC systems and assign them to the phases "beginning", "middle" and "closing" (cf. Henne and Rehbock, 2012) of a conversation as well as to the two conversational contexts "socialness" and "informativeness" (cf. Bedrosian et al., 2003; Hoag et al., 2004; Todman and Alm, 1997). Against this background, the following two research questions (RQs) are posed: ***(RQ1)*** *Do the applications of Artificial Intelligence that are proposed within the extant research literature for AAC systems cover the conversation phases "beginning", "middle" and "closing" and the conversational contexts "socialness" and "informativeness" appropriately?* ***(RQ2)*** *What are possible implications for AI in the context of AAC systems regarding the consideration of conversation phases and conversational contexts?*

We contribute to the inclusion of speech-impaired individuals into societal life and accelerate the corresponding Information Systems (IS) research so that these individuals are proposed comprehensive and individualised AI-supported AAC. The remainder of this paper is structured as follows: Section 2 provides the theoretical background and the related work. In Section 3, we describe the procedure of the literature review and literature analysis. Afterwards, in Section 4, we present our results and discuss the reasonable implications of the research in Section 5. Finally, Section 6 draws an overall conclusion.

2 Conceptional Basics and Related Work

2.1 Conceptional Basics

To investigate the problems of current AI-based approaches within the context of AAC systems in more detail, we take an up-close look at the conversations in which AI is applied. In this context, extant communication research has investigated structures and components (i.a., communication activities, contents, sequences) within daily conversations. According to Henne and Rehbock (2012), a conversation can be structured into an opening, a middle and a closing part. Each of these conversation phases has specific goals and comprises individual communication activities the conversation participants need to perform (cf. Spiegel and Spranz-Fogasy, 2001). The opening part concerns, i.a., the recognition of incoming communication requests, the identification, assessment and greeting of the conversation partner as well as the introduction of the conversation topics (cf. Henne and Rehbock, 2012). The closing part deals with closing the conversation, thanking the conversation partners and farewelling them (cf. Schegloff and Sacks, 1973; Spiegel and Spranz-Fogasy, 2001). In comparison to the opening and closing part, the middle part aims to fulfil specific communication activities that are related to the conversation goal and the purpose framed within the opening part (Henne and Rehbock, 2012).

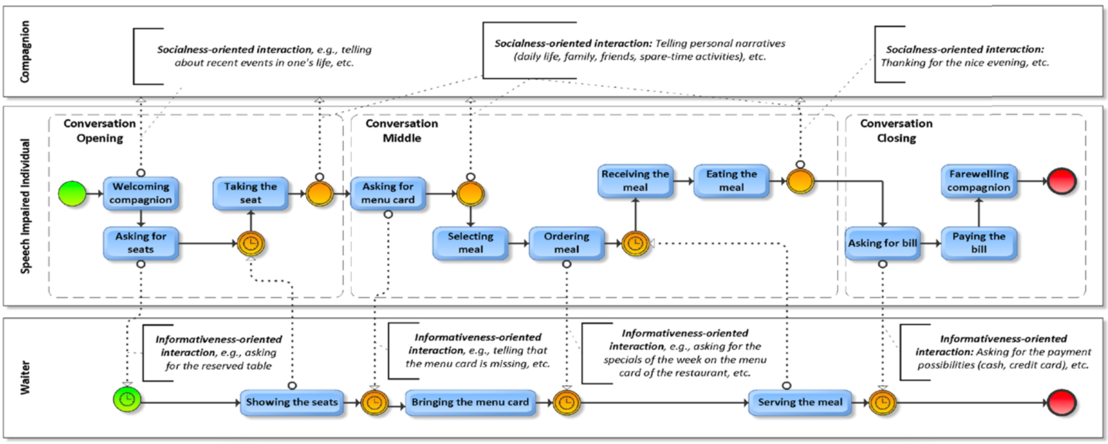


Figure 1. Illustration of different types of interactions (in particular "socialness-" and "informativeness"-oriented) in the example of visiting a restaurant with an acquaintance, own illustration related to Ehlich and Rehbein (1972) and Schindler (2013).

The conversation activities being performed within the respective conversation phases are further determined by different factors, i.a., the number of conversation participants involved (one-to-one conversations vs. group conversations) (Ehlich and Rehbein, 1972), the relationships that exist between the conversation partners (acquaintances vs. strangers) (Spiegel and Spranz-Fogasy, 2001) and the conversational context ("socialness"-oriented vs. "informativeness"-oriented conversations (cf. Bedrosian et al., 2003; Hoag et al., 2004; Todman and Alm, 1997)). Accordingly, group conversations require more complex activities to synchronise the conversation flow and the conversation utterances being exchanged. Further, interactions with friends or acquaintances, e.g., within a restaurant visit, are rather "socialness"-oriented and informal while the interactions with the waiter at the restaurant are rather "informativeness"-oriented and formal (cf. Ehlich and Rehbein, 1972; Schindler, 2013). As a consequence, humans communicate differently and apply different articulations depending on whom they interact with *(see Figure 1)*. Nonetheless, speech-impaired individuals lack the fundamental capability to apply vocal language and therefore need to apply Augmentative and Alternative Communication.

Augmentative and Alternative Communication encapsulates strategies, techniques, and tools that support individuals with speech impairments in expressing their thoughts, desires, feelings, and ideas (Hourcade et al., 2004). In this context, AAC systems can be described as VOCAs (Voice Output Communication Aids) or SGDs (Speech Generating Devices) providing speech-impaired individuals with the ability to express themselves using vocal language. Currently available high-tech AAC systems encompass the use of electronic devices (e.g., smartphones, tablet PCs) and the devices' built-in peripherals (e.g., camera, microphone) (e.g., Baldassarri et al., 2014; Laxmidas et al., 2021). The board within AAC systems contains the vocabulary the users can employ. To contribute utterances to conversations, users of AAC systems can either relate to pre-stored sentences (e.g., Alm and Higginbotham, 2008), can compose single words within the so-called speech line (Klauer et al., 2021) or can make use of keyboard capabilities and a combination of letters to express words and sentences (e.g., Kristensson et al., 2020). The utterance formed and to be spoken towards the conversation partner is then processed by the speech synthesising module that generates the voice output that is subsequently emitted towards the conversation partner by the loudspeaker of the device. This voice output can in turn be interpreted and reciprocated by the conversation partner so that the conversation can progress using mutual and coordinated utterances until a targeted state of the conversation is reached.

2.2 Design Requirements of AAC Systems

To be purposeful for speech-impaired individuals, research has investigated several design requirements for AAC systems. There have been considered general requirements, such as the support of multiple languages (e.g., Guasch et al., 2019; Kane et al., 2017), availability for different operating systems (e.g., iPad OS, Android) and devices (e.g., smartphones, tablet PCs) (e.g., Baldassarri et al., 2014; Dattilo et al., 2008), reliability and confidentiality of the AAC system concerning generated communication data (e.g., Dattilo et al., 2008; McNaughton and Bryen, 2007). Beyond that, an AAC system needs to propose a core vocabulary that enables users to exchange needs, wants and thoughts within various daily life situations (cf. McNaughton and Bryen, 2007; Van Tilborg and Deckers, 2016). In this connection, the AAC systems must offer a dynamic expansion of the vocabulary (e.g., Van Tilborg and Deckers, 2016; Wang et al., 2018). To open up the variability of speech, the storage of synonymous outputs for respective the same utterance needs to be enabled (Alm and Higginbotham, 2008). For individuals hindered to edit the vocabulary on their own because of severe cognitive disorders accompanying their speech impairment, family members, caregivers, and nursing staff need to be given access to the AAC system (Baldassarri et al., 2014). Cognitive disorders also require assigning pictograms to the respective utterances to support the understanding of the concept related to an utterance (e.g., Guasch et al., 2019; Klauer et al., 2021). To support efficiency in selecting vocabulary, it is necessary to propose means to structure the utterances, e.g., with topic- and concept-specific folders or grids (e.g., Klauer et al., 2021; Wang et al., 2018). Further, there are design requirements for the individualisation of the AAC system towards the users' personalities. To customise the way an utterance is outputted, AAC systems propose the user means to adjust the speed and the volume of the voice output (e.g., via sliders in the GUI) (e.g., Arnott and Alm, 2013; Kane et al., 2017). In this vein, maintaining individual user profiles that include user-related preferences enables the customisation of the speech output (e.g., gendering the voice output, storing the user's personal life history) (Alm and Higginbotham, 2008; Guasch et al., 2019).

Although these several design requirements have been investigated within the extant literature, AAC systems still show drawbacks, i.a., low communication rates (cf. Arnott and Alm, 2013; Kane et al., 2017; Klauer et al., 2021) and low user- and situation-specific individualisation (e.g., Arnott and Alm, 2013; Kane et al., 2017; Murphy, 2004; Van Tilborg and Deckers, 2016). Therefore, research has started to investigate the potentials of AI approaches within AAC systems.

2.3 AI Approaches and AI-based Potentials in the Context of AAC Systems

Artificial intelligence relates to the ability of a machine to apply characteristics of human intelligence to solve specific problems and adapt to changing environments (Boden, 2018). The capabilities of AI have recently been incorporated within AAC systems, leading to advances in the support of speech-impaired individuals. Corresponding AI approaches consider for example the user's past selection data, including the object of interest, time of use or, related GPS locations, to predict what a user wants to express in a specific context (e.g., Klauer et al., 2021; Neamtu et al., 2019; Obiorah et al., 2021). To further enhance the users' communication competency, approaches to model conversational topics that could be of interest in a conversation take advantage of Natural Language Generation (NLG) (e.g., Dempster et al., 2010; Vertanen, 2017). AI-based speech recognition and parsing of the interlocutors' verbal expressions propose need-fitting information that could improve both rate and relevancy of the utterances (e.g., Neamtu et al., 2019; Wisenburn and Higginbotham, 2009). Besides the consideration of the interlocutors' voice, facial features can be useful when retrieving relevant information regarding the conversation (e.g., Ascari et al., 2020; Kane et al., 2012).

As can be derived, there are AI approaches that can support speech-impaired individuals within AAC systems. Nevertheless, it remains unclear whether the conversation phases and conversational contexts have so far been sufficiently addressed. However, it is necessary to cover several conversation phases and conversational contexts to effectively reduce the communication gap speech-impaired individuals suffer from. Thus, within the research at hand, we conduct a differentiated investigation of the extant AI-based potentials and thereby aim to contribute to closing this communication gap.

3 Research Approach

3.1 Collection and Evaluation of the Literature

Within our investigation, we performed two systematic literature reviews (Cooper, 1988; Vom Brocke et al., 2015; Webster and Watson, 2002). With the **first** literature review, we aimed at identifying design requirements posed for AAC systems within the extant literature. Thereby, it could be revealed that, although several design requirements are being investigated, there are still drawbacks these AAC systems show, i.a., low communication rates (e.g., Arnott and Alm, 2013; Kane et al., 2017) and low user- and situation-specific individualisation (e.g., Arnott and Alm, 2013; Murphy, 2004; Van Tilborg and Deckers, 2016). To address these drawbacks, the prevailing research has proposed various AI approaches. Because these AAC systems nevertheless show only a low prevalence among speech-impaired individuals, with the **second** literature review, we focused on the AI approaches and AI-based potentials for AAC systems that are proposed within the extant research literature and that we subsequently aim to investigate in more detail regarding the conversation in which the AI is applied.

Considering the taxonomy of Cooper (1988), the focus of this literature review was to investigate the potentials of AI in the context of AAC systems. Thereby, we could identify that applying AI for supporting individuals suffering from speech impairments by means of AAC systems across different phases of a conversation and within different conversational contexts is yet an under-researched topic *(see Section 2.3)*. To address this gap, firstly we identified current AI approaches and corresponding potentials for AAC systems and, secondly within the analysis of the literature, assigned the identified AI-based potentials to the three conversation phases "beginning", "middle" and "closing" as well as to the two conversational context "informativeness" and "socialness". As the differentiated investigation of AI approaches and their potentials in the context of AAC systems has so far not been addressed sufficiently within the extant research, it is nonetheless essential in order to design and develop more suitable AI approaches for AAC systems and their users.

The literature review followed a five-step procedure (Cooper and Hedges, 1994): **(I)** Problem Statement, **(II)** Collection of the Data, **(III)** Data Evaluation, **(IV)** Analysis and Interpretation of the Data, and **(V)** Presentation of the Results. The problem statement has been introduced above in *Sections 1, 2.2 and 2.3*, while the remaining aspects **(II)** - **(V)** are described in detail in the upcoming sections. At first, a topic-based search was performed to query scholarly databases (Vom Brocke et al., 2009; Webster and Watson, 2002). The following databases were analysed: ACM Digital Library, AISeL, Emerald Insight, Google Scholar, IEEE Xplore, SpringerLink, Science Direct and Web of Science. These databases cover leading journals of the IS discipline, proceedings of highly renowned IS conferences, as well as technical reports and dissertations. By means of these selected databases, we not only cover the extant IS research but also research communities that explicitly address the investigation of AAC systems and their support potentials for speech-impaired individuals (e.g., the "Augmentative and Alternative Communication" Journal within ScienceDirect). Next, by following an iterative refinement process (cf. Kitchenham et al., 2009; Wohlin et al., 2012), we derived keywords for our topic-based search which focuses on AI applications and AI-based potentials to support speech-impaired individuals by means of AAC systems. In general, there are numerous techniques that can potentially be applied in the realm of AI (e.g., for artificial neural networks: convolutional neural network or backpropagation neural network). Since it is difficult to cover all the specific AI techniques within the literature search, we took an initial look at the extant literature on AI in the context of AAC. Thereby, we could reveal that these papers also relate to the categories of AI techniques (e.g., "machine learning", "deep learning", "supervised learning", "unsupervised learning"). Thus, we applied these and additionally "AI" and "artificial intelligence" as keywords to cover the AI aspect within the literature collection. In the context of VOCAs and SGDs that we focused on as solutions for supporting people with speech impairments, it could be revealed that the relevant literature persistently referred to those in connection with the term "augmentative and alternative communication" or the abbreviation "AAC". Overall, the following keywords, and combinations of them, were therefore applied: "augmentative and alternative communication", "AAC", "artificial intelligence", "AI", "machine learning", "supervised learning", "unsupervised learning" and "deep learning". As a result of this literature search procedure, **107 papers** could be derived.

To assess the appropriateness of these 107 publications, we first scanned the titles of the papers. If the title hinted at the application of AI in the context of AAC to support speech-impaired individuals, in the second step, we read the abstract and the conclusion of the respective paper to verify its appropriateness. Regarding the purpose of our research, we applied the following preconditions within the literature assessment: ***(1)*** *Does the publication deal with the application of AI in the context of AAC?* ***(2)*** *Does the publication reveal how these AI applications propose support for individuals suffering from a speech impairment?* ***(3)*** *Does the publication limit its application to the end device of the user and does not include further external devices, hindering the portability of the AAC systems within daily life situations, e.g., EEG sensors, sensors for acquisition of nasal pressure?* (cf. Elsahar et al., 2019).

After applying these inclusion criteria to the identified literature, 25 papers remained. For these **25 papers**, we additionally and iteratively applied backward and forward searches (cf. Webster and Watson, 2002). This resulted in **three** additional references that could meet the above-mentioned inclusion criteria. Regarding the distribution of the finally resulting **28 papers** used for the upcoming analysis across the searched databases, we found the following (with duplicates and papers not in the English language already sorted out): **14** references were extracted via Google Scholar, **three** via SpringerLink, **three** via Web of Science**, two** via Science Direct, **two** via the ACM Digital Library, **two** via IEEE Xplore, **one** via Emerald Insight and **one** via the AISeL.

3.2 Literature Analysis

After evaluating the literature collection, we analysed the relevant publications. Here, qualitative content analysis (cf. Mayring, 2004) was applied, which helps to filter out those fragments of information from larger texts that are suitable for answering the RQs.

3.2.1 Deriving the AI-based Potentials in the context of AAC Systems

The AI-based potentials being harnessed in the context of AAC systems were derived as follows: First, the publications were screened regarding any text statement that hinted at a realisable benefit (potential) of AI approaches to support speech-impaired individuals in the course of a conversation by means of an AAC system. On the one hand, some papers directly related to the potentials of applying the described

AI approaches, i.a., within the problem statements or the application and evaluation of the AI approaches (e.g., increase the speed of response or the individualisation of conversation contributions). On the other hand, in cases when there was no AI-based potential immediately observable, we related to the notion of key performance indicators in the context of AAC systems (e.g., number of clicks to produce a conversation contribution, flexibility of the utterances) and to the already derived potentials. In this way, we could derive a broad range of potentials for AI approaches for AAC systems. For the subsequent assignment to conversation phases and conversational contexts, we aimed to derive a manageable number of categories. Therefore, an abstraction of the AI-based potentials into coherent categories was carried out (cf. Mayring, 2004). Based on the attention and importance received within the literature, and led by the entities and structures that constitute a conversation (cf. Henne and Rehbock, 2012; Schank and Schoenthal, 2016), in particular, four categories could be derived *(see Section 4)*.

3.2.2 Assigning AI-based Potentials to Conversation Phases and Conversational Contexts

Subsequently, we aimed at assigning the identified AI-based potentials to the phases “beginning”, “middle” and “closing” as well as to the conversational contexts “informativeness” and “socialness”. The following assignment criteria could be derived from the extant literature and were applied: The publications were screened regarding any text statement that hinted at the phases “beginning”, “middle” and “closing” of a conversation or the conversational context “informativeness” and “socialness”. Hereby, the text statements were observed whether communication activities that are commonly performed within the respective phases are described within the paper (e.g., recognition of incoming communication requests, the identification, assessment and greeting of the conversation partner as well as the introduction of the topics of the conversation in the “beginning”, recognition of the fulfilment of the conversational goals, closing the conversation, thanking the involved conversation partners and farewelling them in the “closing”). Regarding “informativeness” and “socialness”, the assignments related to the description of the situation in which the conversations in the respective papers were performed. Regarding the criteria for an assignment, “informativeness”-oriented interactions are rather formal and aim at conciseness, clarity and relevance of the contents (cf. Bedrosian et al., 2003; Hoag et al., 2004). In comparison, “socialness”-oriented interactions are rather informal, aim at keeping pace with the speed of the conversational flow and are rather observed in social contexts (cf. Todman and Alm, 1997). In those cases when none of the conversation phases or conversational contexts could be observed within the respective paper, an assignment to “Not specified” was applied. If any assignment faced ambiguities, these were cleared within discussions and justifications among the involved researchers.

4 Results of the Literature Analysis

To effectively synthesise the extant research literature, we derived four coherent categories, representing the potentials of AI in the context of AAC systems, by applying inductive category development (Mayring, 2004). These categories represent the potentials of AI for AAC systems, distinguished by four dimensions: *1. Formal Course of a Conversation, 2. Natural Context Factors, 3. Interlocutor, and 4. Speech-Impaired Individual (i.e., the user of the AAC system).* In the following, exemplary potentials of the derived categories are presented.

Maintaining the Formal Course of a Conversation

The first category deals with the potential of AI in terms of the formal course of a conversation. The main objective is to ensure the grammatical correctness of the utterances, supporting their comprehensiveness and thus maintaining the overall conversation course. From the perspective of Natural Language Processing (NLP), speech errors from speech-impaired individuals can be interpreted as grammatical errors at the morphological and syntactic levels. Therefore, Park et al. (2021) applied neural Grammatical Error Correction (GEC) to ensure the formal correctness of the utterances. More precisely, they applied a deep learning-based speech-to-text approach, taking the voice of the user as input and converting it into text. Subsequently, the GEC algorithm revises various linguistic errors to replace missing or wrong words and to enable the users to practice the vocal language. To support users with autism, Hervás et al. (2020) proposed an application to automatically compose messages based on pictograms that include AI predictive capabilities. Since predictive systems and orthographic correctors can sometimes persistently provide incorrect or undesirable options, the authors implemented a functionality whereby users can correct messages using NLP-based part-of-speech decomposition, thereby ensuring the semantic and formal correctness of the message.

Incorporating Natural Context Factors

To further support AAC systems through the potential of AI, researchers attempted to incorporate natural context factors to improve the proposed user responses by recommending situationally appropriate utterances. In this regard, numerous scholars have recognised the importance of environmental-related (e.g., GPS coordinates) or time-related (e.g., daytime) information (e.g., Klauer et al., 2021; Neamtu et al., 2019; Reddington and Tintarev, 2011; Waller, 2019).

In this context, Emil et al. (2020) developed an AI approach facilitating communication by analysing users' past selection data, including the object of interest, time of use, touch duration, GPS location, and pictogram location on the screen. Similar to this, Laxmidas et al. (2021) proposed an approach including a recommendation engine based on nearest-neighbour clustering. Each time an utterance is selected, the engine updates its internal ranking and takes into account the given time of the day and the number of selections. Subsequently, these features are partitioned into ideal clusters so that an optimal collection of relevant utterances is available for any given time of day. Also, Neamatu et al. (2019) turned their attention to the use of AI to make context-based recommendations for utterances. By analysing historical usage data (e.g., the object of interest, time of usage, GPS data, touch time) the developed machine learning algorithm predicts what a person with a speech impairment may want to express in a specific context. Further, an approach to model conversational topics that would be of interest in social conversations is presented by Dempster et al. (2010), which takes advantage of NLG to automatically create and adopt a user-specific ontology (cf. Karakatsiotis et al., 2008). To do so, several contextual information (location, time) are turned into useful conversational utterances through a template-driven utterance generation system. In addition to looking at contextual factors in general, several investigations took a step further and tried to determine how contextual factors can support specific daily activities. Obiorah et al. (2021) focused on supporting a specific leisure activity (dining in a certain restaurant) by facilitating dynamic language productions by incorporating location-related web content to support the comprehension of novel vocabulary. Similarly, Tintarev et al. (2016) focused on supporting children with complex communication needs in school to facilitate interactive narratives about personal experiences. To achieve this, the authors developed an AI-based clustering algorithm considering several locations, time and voice recording data to create concrete utterances, reflecting holistic personal narratives. To personalise the respective utterances, the approach identifies different event boundaries in terms of location changes and events specified in the children's timetables.

Tailoring towards the Interlocutor

Apart from the consideration of natural context factors such as geolocation and time of day, consideration of the interlocutor has also gained attention in the research field of AAC. In this regard, the approaches to integrate the interlocutor can be divided into two areas - speech recognition and image recognition - based on the technology used.

Wisenburn and Higginbotham (2009) demonstrated a computerised communication program utilising an NLP strategy consisting of speech recognition and a parser that proposes utterances incorporating the noun phrases spoken by the interlocutor as modular speech keys. In addition, Neamtu et al. (2019) enable speech-impaired individuals to respond more quickly by integrating a natural language conversation feature, which continuously listens to the surroundings and searches for specific trigger words and activates appropriate utterances. Therefore, the obstacle to conversation initiation and thus the communication gap are reduced. Further, the authors implemented a natural language sentence classifier to examine the raised question of the interlocutor and to classify whether the question is dichotomous. If so, the AAC system will solely present the answers of "Yes" and "No" to further enhance the communication exchange. Similarly, Emil et al. (2020) could reduce the time and effort required to communicate using a text-to-speech module in conjunction with an NLP-based classifier. The algorithm detects specific questions and presents the most relevant utterances to users based on the interlocutors' statements. In contrast, Vertanen (2017) developed an approach to initially extract the mentioned conversation topics and subsequently generate holistic utterances by using a neural network specifically tuned to the conversation topic. In this way, the cohesion between individual parts of the conversation can be increased, resulting in a reduction of the communication gap.

Besides the consideration of the interlocutors' voice, also facial features or gestures can be useful when retrieving relevant information regarding the course of the conversation. Accordingly, Kane et al. (2012) developed an approach to generate person-associated utterances based on facial image recognition of the interlocutor, resulting in a more personalised and convenient system. Ascari et al. (2020) further applied AI in the form of a gesture recognition classification algorithm to identify non-verbal expressions of the interlocutor and transform them into associated emotions. Subsequently, they will be included in the generation of the users' responses, allowing them to respond to the mood or emotions of the interlocutor and thus strengthen their ability to communicate.

Tailoring towards the Speech-Impaired Individual (AAC System User)

In addition, numerous researchers have harnessed the potential of AI to better tailor the conversation towards the user (e.g., characteristics, attitudes and preferences of the user). Shen et al. (2022) proposed a multi-turn dialogue system that applies a language model with bag-of-keywords and personality-related information to generate high-quality responses based on keyword entries. By the generation of meaningful utterances with very little user input, additional keystrokes were saved, leading to a significant reduction in the communication gap. Further, Mancilla et al. (2015) developed an ontology personalisation approach for an AAC system, which automatically extends a given domain ontology into a user-related one, including user-specific information. By including personal information in the communication process, the message quality improves as it becomes more complete and personal. According to Hernández et al. (2014), user behaviour is essential to establish personalised recommendations. Thus, the authors implemented a memory-based recommender system that tracks users by user archetypes, representing their behaviour patterns, goals and needs. Subsequently, the archetypes are used in conjunction with AI and statistical methods to predict and suggest need-fitted utterances. Further, Heo and Kang (2019) proposed a system to incorporate user-related information by applying a slot-filling method. By employing decision trees, it automatically learns slot priority strategies to enable the efficient selection of information for the user and thus an efficient way to prevent unnecessary inputs and minimise the number of keystrokes. Another method for integrating user-related information by generating context-related vocabulary from photographs of personally relevant events is presented by De Vargas and Moffatt (2021), who aimed at supporting people with speech impairments in recounting their past experiences. Here, NLP-based photo-to-story vocabulary generation reduced irritation and navigation errors and generated improved lexical retrieval during activity retelling. Johnson et al. (2020) demonstrated the usage of various sensors for measuring the voice and video recordings of non-verbal users. By using speech and image recognition, the user's moods, emotions and sounds can be measured and assigned to different categories (e.g., laughter, crying, or protest), thereby clarifying the underlying intention and supporting the user's expressiveness. Finally, as stated by Mills et al. (2014), several individuals suffering from speech impairment use the same synthetic voice, diminishing their expressiveness. To enable a more personalised voice, the authors developed an approach that extracts prosodic properties from the target talker's voice and applies these features to a surrogate talker's database, generating a synthetic voice with the vocal identity of the target talker and the clarity of the surrogate talker.

5 Discussion and Future Research

To identify existing gaps in the coverage of AI-based potentials concerning the conversational phases and contexts and thus deduce related research prospects, the AI potentials were further assigned to the appropriate conversation phases and contexts as described in Section 3.2.2. The respective assignments are shown in Table 1.

Authors / Publications	AI-based Potentials of AAC Systems				Conversation Contexts			Conversation Phases			
	Maintaining the Formal Course of a Conversation	Incorporat-ing Natural Context Factors	Tailoring towards the Inter-locutor	Tailoring to-wards the Speech Im-paired	So-cialn.	In-form.	Not Spec.	Beg.	Mid.	Clos.	Not Spec.
Ascari et al. (2020)			X				X				X
Dempster et al. (2010)		X	X	X	X				X		
De Vargas and Moffatt (2021)		X		X	X				X		
Emil et al. (2020)		X	X	X			X		X		
Heo and Kang (2019)				X			X				X
Hernández et al. (2014)				X	X						X
Hervás et al. (2020)	X			X		X					X
Hossain et al. (2018)		X					X				X
Johnson et al. (2020)				X	X				X		
Kane et al. (2012)		X			X						X
Klauer et al. (2021)		X					X				X
Kristensson et al. (2020)		X				X					X
Laxmidas et al. (2021)		X		X			X				X
Li et al. (2022)				X			X				X
Mancilla et al. (2015)		X		X	X				X		
Mills et al. (2014)				X			X				X
Neamtu et al. (2019)		X	X	X	X			X	X		
Obiorah et al. (2021)		X	X			X		X	X		
Park et al. (2021)	X	X		X			X		X		
Reddington and Tintarev (2011)		X			X				X		
Roy et al. (2021)	X		X	X			X	X	X		
Shen et al. (2022)		X		X	X				X		
Tintarev et al. (2016)		X		X	X						X
Vertanen (2017)			X				X		X		
Waller et al. (2013)				X	X				X		
Waller (2019)		X	X		X				X		
Wandmacher et al. (2007)		X					X				X
Wisenburn and Higginbotham (2009)			X		X				X		

Explanation of the Abbreviations: *Socialn.: Socialness, Inform.: Informativeness; Beg.: Beginning; Mid.: Middle; Clos.: Closing; Not Spec.: Not Specified.*

Table 1. Assignment of AI-related potentials to conversation phases and conversational contexts.

The general picture that emerges from the analysis is that some studies have already addressed specific contexts and phases of conversations. However, significant gaps in the literature could be revealed, which indicate shortcomings in the research on the examined subject of AAC and AI. To guide research into the future, we derived several implications for research below.

5.1 Implications for Future Research

The findings deduced from the assignment of the AI-based potentials to the respective conversation phases and conversational contexts are intended to help researchers and practitioners to make sense of the accumulated knowledge in this emerging research field. Based thereon, we discuss implications concerning the observations throughout this investigation and outline novel directions for future research that may yield promising potentials in the field of AAC.

First, as a cursory glance at Table 1 reveals, there is no prevailing research that leverages the potentials of all derived categories to support speech-impaired individuals. However, to support communication holistically using an AAC system, the consideration of all conversation-related dimensions (categories) and thus the use of their associated potentials is indispensable. Thus, studies considering all relevant dimensions of a conversation simultaneously could be promising avenues for further research.

- The prevalent literature on the first category is sparse, with only three articles specifically addressing the use of AI to maintain a formal course within a conversation. In comparison, 17 articles deal with the integration of natural context factors as well as nine and 17 articles deal with the consideration of the conversation partner and the speech-impaired individual, respectively. This leads to the assumption, that there is a necessity for research concerning maintaining the formal course of the conversation by means of tailored AI approaches. In this respect, additionally to neural GEC, word prediction engines could be leveraged to influence the choices of spell checkers or to automatically correct minor spelling errors. Moreover, prediction engines that use syntactic or semantic components could similarly be used for grammar correction and disambiguation of meaning. For example, Widgit's Communicate software (Widgit, 2022) uses syntactic disambiguation to provide the appropriate symbol when an ambiguous word such as "can" occurs in a sentence. In general, probability-based prediction engines can be used in many situations where textual ambiguity exists, and support both conversational comprehension and users' communicative competence.

- Moreover, AI holds several potentials for text simplification and summarisation and thus supporting the maintenance of the formal course of conversation. Text simplification converts the associated text message into simpler sentence structures and vocabulary while retaining the meaning of the original (Al-Thanyyan and Azmi, 2021). This could also support the formal course of the conversation by breaking down more complex issues into simpler elements of communication, and thus reducing the likelihood of errors.
- Similar to text simplification, text summarisation aims to provide a brief overview of what is contained within large texts. Using an automatic summarisation approach, particularly important sentences in a document can be identified and included in the summary (Adhikari, 2020). Although comparable technologies are occasionally used for AAC systems (e.g., Obiorah et al., 2021), the potential of automatic text summarisation has yet to be exhausted. In this way, complex text messages could be simplified for individuals with cognitive-linguistic challenges. Newspaper or magazine articles, website content, or emails could be automatically processed and inserted into the AAC system to transform them into modules for expressive communication. With a simplified and appropriately segmented set of text modules, the user could select from these to discuss current events or personal narratives with grammatically and orthographically correct sentences.

Second, the consideration of conversational contexts is worth mentioning. As previously indicated, considering specific conversational contexts is one way to address the associated requirements (e.g., formal vs. informal) to specifically support the conversational process. However, as Table 1 reveals, solely three articles refer to the conversational context of "informativeness", while 13 articles take into account the conversational context of "socialness". Since the differentiation between these conversational contexts is indispensable to holistically support AAC-based communication *(see Sections 1 and 2.1)* and thus achieve a high level of communication competence in speech-impaired individuals, future research will need to shed light on how the conversational contexts can be supported by appropriate AI methods. Especially concerning the context of "informativeness", our results implicate research gaps, since the prevailing literature to date mainly focuses on the context of "socialness".

- Since "informativeness"-oriented communication is more concerned with the formal exchange of relevant information that focuses on conciseness, clarity, and relevance of content rather than the speed of exchange as in "socialness"-oriented communication (cf. Bedrosian et al., 2003; Hoag et al., 2004), "informativeness"-oriented communication aims at a holistic view of a topic, including the coverage of all topic-specific relationships. In contrast to "socialness"-oriented conversations, the focus is thus mainly on the meticulous elaboration of a specific topic and less on the superficial but more flexible coverage of a broad spectrum of differentiated topics.
- One conceivable approach for the holistic mapping of a discourse domain is the AI-based generation of an ontology that can be applied to map all topic-related concepts and their relationships. Ontologies model particular knowledge through a representative vocabulary determined by the formal definition of a set of concepts associated with a particular domain and the meta-relationships between them (Mancilla et al., 2015). In this way, criteria of "informativeness"-oriented communication such as clarity, coherence, and extensibility are further met (Gruber, 1995). To enable the holistic coverage of a conversation topic with all associated concepts, deep learning approaches (e.g., neural networks) can be applied in conjunction to dynamically identify and integrate topic-relevant concepts. Once a specific topic has been identified, probabilistic neural networks can allocate (potentially)

relevant concepts and dynamically extend the ontology with the topic-specific concepts. Therefore, semantically related concepts in various degrees of detail can be identified, resulting in a holistic coverage of the discussed topic and thus enabling, e.g., the formation of logical conversation chains supporting the formal exchange of information. As an example, when ordering a dish in a restaurant *(see Figure 1, Section 2.1)*, the user would also like to receive information about certain aspects that might be relevant to ordering the dish. If the user wants to ask about the origin of certain ingredients or possible allergens, the relevant topic-specific concepts must be available within the ontology. To achieve the dynamic extension of the given ontology, the associated concepts (e.g., the origin of specific ingredients) could be determined by a neural network and dynamically integrated into the ontology. By dynamically integrating topic-specific concepts, all relevant facets of a conversation can be taken into account, ensuring conciseness as well as the relevance of the respective content on the one hand and supporting clarity through a logical structure of the topics on the other.

- In this vein, studies additionally have to examine which event boundaries can be defined to dynamically switch between the conversational contexts to adapt the proposed utterances in a target-oriented way. Automatic speech recognition could be utilised to analyse the spoken words of the interlocutor regarding specific keywords to determine whether a communication shifts from "socialness" to “informativeness”. If a shift in the communication context is identified, the generation of utterances can be specifically geared to the requirements of the currently identified context, thereby increasing both the prospect of conversational success and the conversational competence of the user.

Third, in view of the conversation phases, each phase has to meet specific goals and therefore comprise individual communication activities the conversation participants need to perform *(see Section 2.1)*. Taking into account the assignment of the AI-based potentials to the respective communication phases, it becomes apparent, that the middle part is densely covered (15 articles), while the opening part (three articles) and the closing part (null articles) are not sufficiently considered yet. Nonetheless, since each conversation phase is associated with specific activities such as recognising incoming communication requests or closing the conversation by thanking the interlocutors, the consideration of each conversation phase and its associated requirements is essential to support a conversation holistically.

- In view of the conversation initiation, image recognition techniques and deep learning-based computer vision propose promising potentials. In this context, action and event recognition could identify an emerging conversation beginning by analysing, from a sequence of images, different actions that a human performs (Leo et al., 2018). Common examples of actions are, e.g., “answering the phone” or “shaking hands”, but also facial actions like “smiling” or “crying”. By identifying certain actions that indicate an upcoming conversation situation (e.g., a greeting gesture such as waving), a conversation situation could be initiated, whereupon the corresponding utterances would be suggested automatically. For illustration purposes, if a speech-impaired user enters a restaurant and is guided to his seat by a waiter, the AAC system could automatically identify, based on the waiter's gesture, that the user is guided to take a seat. Accordingly, the system can generate appropriate utterances that support the respective conversation situation and thus enable the generation of targeted utterances. Likewise, the facial expression of the conversation partner could be included in the automated adaptation of the vocabulary. Thus, in the case of an angry interlocutor, positively connotated vocabulary could be used to appease in.
- In the context of conversation closing, the focus is on farewelling the interlocutor (cf. Schegloff and Sacks, 1973; Spiegel and Spranz-Fogasy, 2001). Thereby, deep learning-based computer vision approaches could be used to enable automated recognition of the closing of the conversation by identifying gestures of farewell and adapting the proposed utterances accordingly. Further potential in the conversation phase could be represented by the automated evaluation of the conversation, the verification of the achievement of conversation goals, and the extraction and evaluation of discussed conversation topics. In this context, machine learning approaches such as automatic topic extraction (cf. Jelodar et al., 2019) could be used to extract the discussed topics and subsequently have them evaluated by the user. If the user rates a topic positively, it can be prioritised in the automated identification of future conversation topics concerning upcoming conversations.

Overall, our results emphasise that AI techniques are ubiquitous in current AAC technologies and propose great potentials for development in the future. Therefore, we encourage research to address the consideration of phase- and context-specific requirements of an AAC conversation to develop appropriate techniques to support speech-impaired individuals and thus better integrate them into social life.

5.2 Limitations

The investigation at hand has several limitations. Although we conducted a broad and structured database search covering major outlets, other researchers might have uncovered additional search terms and additional relevant papers. Yet, this structured literature review allows for a transparent, replicable, and comprehensive overview of AI-based potentials in the context of AAC systems. The assignments of the derived AI-based potentials to conversation phases and conversational contexts bear to a certain extent subjectivity. However, several assignments were performed by three researchers individually. In case of disagreements, the article in question was analysed by a fourth researcher, followed by a discussion until a consensus was reached. Despite these limitations, we are confident that our findings help to get a first overview and a better understanding of the current AI approaches in the context of AAC systems and the insufficiencies of the AI decisions that can be related to the underlying conversations.

6 Conclusion

Digital systems of Augmentative and Alternative Communication for supporting speech-impaired individuals have gained tremendous importance as a means of the inclusion of speech-impaired individuals into societal life. Although the extant research has investigated several design requirements and AI approaches, there are certain circumstances where AAC systems deduce recommendations and apply adaptions that are insufficient given the specific situation in which they are applied. This is due to the divergent requirements imposed by the conversation context (e.g., formal vs. informal conversations) and its structure (e.g., conversation phases). Thus, as the AI approaches may need to be adapted concerning the conversations in which they are applied, we took an up-close look thereon within our investigation.

To establish a first basis for the underlying research topic and to develop a deeper understanding of the research field of AI in the context of AAC systems, we have therefore conducted a systematic literature review *(see Section 3)*. In this way, we have derived corresponding AI approaches and deduced related AI-based potentials. For their subsequent assignment to conversation phases and conversational contexts, we could derive four categories that are based on the attention and importance they have received within the extant literature and are additionally led by the entities and structures that constitute a conversation (e.g., conversational partners, the course of the conversation). Subsequently, based on the determinants of the phases and contexts that are described in the corresponding literature, we have assigned the categories of AI-based potentials to the conversation phases "beginning", "middle", "closing" and to the two conversational contexts "informativeness" and "socialness". As this assignment revealed, there are indeed several gaps within the extant research on AI regarding the coverage of the conversational context "informativeness" and the conversation phases "beginning" and "closing" *(see RQ1)*.

To this extent, this paper proposes several contributions. First, based on the extant literature, it provides a comprehensive and structured overview of AI-based potentials in the context of AAC systems. Moreover, we have consolidated these AI-based potentials into an abstraction of four coherent categories *(see Section 4)* and could establish synergies of these research areas. Second, our results propose possibilities for AI research to investigate the identified AI-based potentials within conversation phases and conversational contexts that have received less attention yet. Based on the contributions of the work at hand, AI approaches that are established within other fields (e.g., within voice assistants for voice commerce, e.g., Reinkemeier et al., 2022) and current generative AI-based language models could be investigated in specific for the case of AAC systems. To guide future research, we contribute several implications for investigating AI in the context of AAC systems, especially regarding the conversation in which the supporting AI is applied *(see RQ2)*. We hope future IS research strives to address the revealed gaps to dismantle communication barriers speech-impaired individuals suffer from.

References

Adhikari, S. (2020). NLP-based machine learning approaches for text summarization. *Proceedings of the 4th International Conference on Computing Methodologies and Communication*, 535-538.

Allen, J. (2005). Designing desirability in an augmentative and alternative communication device. *Universal Access in the Information Society*, 4(2), 135-145.

Alm, N., and Higginbotham, J. (2008). Whole utterance approaches in AAC. *Augmentative and Alternative Communication*, 24(3), 235-254.

Al-Thanyyan, S. S., and Azmi, A. M. (2021). Automated text simplification: a survey. *ACM Computing Surveys*, 54(2), 1-36.

American Speech Language Hearing Association (1991). Report: augmentative and alternative communication. *ASHA*, 33, 9–12.

Arnott, J. L., and Alm, N. (2013). Towards the improvement of Augmentative and Alternative Communication through the modelling of conversation. *Computer Speech & Language*, 27(6), 1194-1211.

Ascari, R. E., Pereira, R., and Silva, L. (2020). Computer vision-based methodology to improve interaction for people with motor and speech impairment. *ACM Transactions on Accessible Computing*, 13(4), 1-33.

Baldassarri, S., Marco, J., Cerezo, E., and Moreno, L. (2014). Accessibility evaluation of an alternative and augmentative communication (AAC) tool. *Proceedings of the 2014 International Conference on Universal Access in Human-Computer Interaction*, 529-540. Springer, Cham.

Bedrosian, J. L., Hoag, L. A., and McCoy, K. F. (2003). Relevance and speed of message delivery trade-offs in augmentative and alternative communication. *Journal of Speech, Language, and Hearing Research*, 46, 800–817.

Boden, M. A. (2018). Artificial Intelligence: A very short introduction. *Oxford, England: Oxford University Press*.

Boyes-Braem, P., Erting, C. J., Johnson, R. C., Smith, D. L., and Snider, B. D. (1994). An overview of current sign language projects in Switzerland. *The Deaf Way: Perspectives from the International Conference on Deaf Culture*, 382-387. Washington, DC: Gallaudet University Press.

Bradshaw, J. (2013). The use of augmentative and alternative communication apps for the iPad, iPod and iPhone: an overview of recent developments. *Tizard Learning Disability Review*, 18(1), 31-37.

Brady, N. (2000). Improved comprehension of object names following voice output communication aid use: Two case studies. *Augmentative and Alternative Communication*, 16(3), 197-204.

Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society*, 1(1), 104-126.

Cooper, H. M., and Hedges, L. V. (1994). Research synthesis as a scientific enterprise. *The Handbook of Research Synthesis*, 285-299.

Dattilo, J., Estrella, G., Estrella, L. J., Light, J., McNaughton, D., and Seabury, M. (2008). "I have chosen to live life abundantly": Perceptions of leisure by adults who use augmentative and alternative communication. *Augmentative and Alternative Communication*, 24(1), 16-28.

Dempster, M., Alm, N., and Reiter, E. (2010). Automatic generation of conversational utterances and narrative for Augmentative and Alternative Communication: a prototype system. *Proceedings of the NAACL HLT 2010 Workshop on Speech and Language Processing for Assistive Technologies*, 10-18.

Desai, T., Chow, K., Mumford, L., Hotze, F., and Chau, T. (2014). Implementing an iPad-based alternative communication device for a student with cerebral palsy and autism in the classroom via an access technology delivery protocol. *Computers & Education*, 79, 148-158.

De Vargas, M. F., and Moffatt, K. (2021). Automated Generation of Storytelling Vocabulary from Photographs for use in AAC. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics,* 1353-1364.

Ehlich, K., and Rehbein, J. (1972). Zur Konstitution pragmatischer Einheiten in einer Institution: Das Speiserestaurant. *Linguistische Pragmatik*, 209-254.

Elsahar, Y., Hu, S., Bouazza-Marouf, K., Kerr, D., and Mansor, A. (2019). Augmentative and alternative communication (AAC) advances: A review of configurations for individuals with a speech disability. *Sensors*, 19(8).

Emil, Z., Robbertz, A., Valente, R., and Winsor, C. (2020). Towards a more inclusive world: Enhanced augmentative and alternative communication for people with disabilities using AI and NLP. *Major Qualifying Projects, Digital Worcester Polytechnic Institute*.

Fritzell, B. (1996). Voice disorders and occupations. *Logopedics Phoniatrics Vocology*, 21(1), 7-12.

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing?. *International Journal of Human-Computer Studies*, 43(5-6), 907-928.

Guasch, D., Martín-Escalona, I., Macías, J. A., Moreno, L., Hervás, R., and Bautista, S. (2019). Expert-based assessment of an augmentative and alternative communication tool. *Proceedings of the 20th International Conference on Human Computer Interaction*, 1-8.

Henne, H., and Rehbock, H. (2012). Einführung in die Gesprächsanalyse. *Einführung in die Gesprächsanalyse*. de Gruyter.

Heo, Y., and Kang, S. (2019). Augmentative and Alternative Communication System Using Information Priority and Retrieval. *Proceedings of the 2019 IEEE International Conference on Big Data and Smart Computing*, 1-4.

Hernández, S. S., Mancilla, D., Medina, J. M., and Iregui, M. (2014). User-centric Recommendation Model for AAC based on Multi-criteria Planning. *Proceedings of the 2014 International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies and Services*.

Hervás, R., Bautista, S., Méndez, G., Gavlán, P., and Gervás, P. (2020). Predictive composition of pictogram messages for users with autism. *Journal of Ambient Intelligence and Humanized Computing*, 11(11).

Hoag, L. A., Bedrosian, J. L., McCoy, K. F., and Johnson, D. E. (2004). Trade-offs between informativeness and speed of message delivery in augmentative and alternative communication. *Journal of Speech, Language, and Hearing Research*, 47, 1270–1285.

Hossain, M. S., Takanokura, M., Sakai, H., and Katagiri, H. (2018). Using context history and location in context-aware AAC systems for speech-language impairments. *Proceedings of the International Multiconference of Engineers and Computer Scientists*, 1.

Hourcade, J., Everhart Pilotte, T., West, E., and Parette, P. (2004). A history of augmentative and alternative communication for individuals with severe and profound disabilities. *Focus on Autism and other Developmental Disabilities*, 19(4), 235-244.

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., and Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78(11), 15169-15211.

Johnson, K. T., Narain, J., Ferguson, C., Picard, R., and Maes, P. (2020). The ECHOS Platform to Enhance Communication for Nonverbal Children with Autism: A Case Study. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-8.

Kane, S. K., Linam-Church, B., Althoff, K., and McCall, D. (2012). What we talk about: designing a context-aware communication tool for people with aphasia. *Proceedings of the 14th International ACM Conference on Computers and Accessibility*, 49-56.

Kane, S. K., Morris, M. R., Paradiso, A., and Campbell, J. (2017). "At times avuncular and cantankerous, with the reflexes of a mongoose" Understanding Self-Expression through Augmentative and Alternative Communication Devices. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1166-1179.

Karakatsiotis, G., Galanis, D., Lampouras, G., and Androutsopoulos, I. (2008). NaturalOWL: Generating texts from OWL ontologies in Protégé and in Second Life. *Proceedings of the 18th European Conference on Artificial Intelligence*.

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., and Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. *Information and Software Technology*, 51(1), 7-15.

Klauer, A., Johannsen, F., and Leist, S. (2021). "BeMyVoice"–An iPad-Based Helper for Speech Impaired Persons. *Proceedings of the International Conference on Design Science Research in Information Systems and Technology*, 40-51. Springer, Cham.

Kristensson, P. O., Lilley, J., Black, R., and Waller, A. (2020). A design engineering approach for quantitatively exploring context-aware sentence retrieval for nonspeaking individuals with motor disabilities. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-11.

Laxmidas, K., Avra, C., Wilcoxen, C., Wallace, M., Spivey, R., Ray, W., Polsley, S., Kohli, P., Thompson, J., and Hammond, T. (2021). CommBo: Modernizing Augmentative and Alternative Communication. *International Journal of Human-Computer Studies*, 145.

Leo, M., Furnari, A., Medioni, G. G., Trivedi, M., and Farinella, G. (2018). Deep learning for assistive computer vision. *Proceedings of the European Conference on Computer Vision Workshops*.

Li, W., Qiu, X., Li, Y., Ji, J., Liu, X., and Li, S. (2022). Towards a novel machine learning approach to support augmentative and alternative communication (AAC). *International Journal of Speech Technology*, 25(2), 331-341.

Louis-Nouvertné, U. (2001). Was sind Gebärdensprachen?: Eine Einführung in die wichtigsten Ergebnisse der linguistischen Gebärdensprachforschung. *Sprache und Literatur*, 32(2), 3-20.

Mancilla, D., Sastoque, S., and Iregui, M. (2015). Towards ontology personalization to enrich social conversations on AAC systems. *Proceedings of the 10th International Symposium on Medical Information Processing and Analysis*, 9287, 212-217.

Mayring, P. (2004). Qualitative content analysis. *A Companion to Qualitative Research*, 1(2), 159-176.

McNaughton, D., and Bryen, D. N. (2007). AAC technologies to enhance participation and access to meaningful societal roles for adolescents and adults with developmental disabilities who require AAC. *Augmentative and Alternative Communication*, 23(3), 217-229.

Mills, T., Bunnell, H. T., and Patel, R. (2014). Towards personalized speech synthesis for augmentative and alternative communication. *Augmentative and Alternative Communication*, 30(3), 226-236.

Murphy, J. (2004). "I Prefer Contact This Close": Perceptions of AAC by People with Motor Neurone Disease and their Communication Partners. *Augmentative and Alternative Communication*, 20(4), 259-271.

Neamtu, R., Camara, A., Pereira, C., and Ferreira, R. (2019). Using artificial intelligence for augmentative alternative communication for children with disabilities. *Proceedings of the 2019 IFIP Conference on Human-Computer Interaction*, 234-243. Springer, Cham.

Obiorah, M. G., Piper, A. M., and Horn, M. (2021). Designing AACs for people with aphasia dining in restaurants. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1-14.

Park, C., Jang, Y., Lee, S., Seo, J., Yang, K., and Lim, H. (2021). PicTalky: Augmentative and Alternative Communication Software for Language Developmental Disabilities.

Pollak, V. A., and Gallagher, B. (1989). A fast communication aid for non-verbal subjects with severe motor handicaps. *Journal of Medical Engineering & Technology*, 13(1-2), 23-27.

Reddington, J., and Tintarev, N. (2011). Automatically generating stories from sensor data. *Proceedings of the 16th International Conference on Intelligent User Interfaces*, 407-410.

Reinkemeier, F., Gnewuch, U., and Toporowski, W. (2022). Can Humanizing Voice Assistants Unleash the Potential of Voice Commerce?. *Proceedings of the 43rd International Conference on Information Systems*.

Roy, Q., Berlioux, S., Casiez, G., and Vogel, D. (2021). Typing Efficiency and Suggestion Accuracy Influence the Benefits and Adoption of Word Suggestions. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1-13.

Schank, G., and Schoenthal, G. (2016). Gesprochene Sprache. *Gesprochene Sprache*. Max Niemeyer Verlag.

Schegloff, E. A., and Sacks, H. (1973). Opening up closings. De Gruyter.

Schepis, M. M., Reid, D. H., and Behrman, M. M. (1996). Acquisition and functional use of voice output communication by persons with profound multiple disabilities. *Behavior Modification*, 20(4), 451-468.

Schindler, E. (2013). Kommunikative Ritualität im Gastronomiebereich. *Freiburger Arbeitspapiere zur germanistischen Linguistik*, 11. Freiburg: Germanistische Linguistik, Universität Freiburg.

Shen, J., Yang, B., Dudley, J. J., and Kristensson, P. O. (2022). Kwickchat: A multi-turn dialogue system for aac using context-aware sentence generation by bag-of-keywords. *Proceedings of the 27th International Conference on Intelligent User Interfaces*, 853-867.

Spiegel, C., and Spranz-Fogasy, T. (2001). Aufbau und Abfolge von Gesprächsphasen. *Text-und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischer Forschung*. 2. Halbband, 1241-1252. De Gruyter.

Tintarev, N. Reiter, E., Black, R., Waller, A., and Reddington, J. (2016). Personal storytelling: Using Natural Language Generation for children with complex communication needs, in the wild …. *International Journal of Human-Computer Studies*, 92, 1-16.

Todman, J., and Alm, N. (1997). Pragmatics and AAC approaches to conversational goals. In Natural language processing for communication aids. *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics (ACL)*.

Vanderheiden, G. C. (1983). Non-Conversational Communication Technology Needs of Individuals with Handicaps. *Rehabilitation World*, 7(2), 8-12.

Van Tilborg, A., and Deckers, S. R. (2016). Vocabulary selection in AAC: Application of core vocabulary in atypical populations. *Perspectives of the ASHA Special Interest Groups*, 1(12), 125-138.

Vertanen, K. (2017). Towards improving predictive aac using crowdsourced dialogues and partner context. *Proceedings of the 19th International ACM Conference on Computers and Accessibility*, 347-348.

Vom Brocke, J. Simons, A., Niehaves, B., Reimer, K., Plattfaut, R., and Cleven, A. (2009). Reconstructing the giant: On the importance of rigour in documenting the literature search process. *Proceedings of the 17th European Conference on Information Systems*.

Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., and Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the Association for Information Systems*, 37(1).

Waller, A., Menzies, R., Herron, D., Prior, S., Black, R., and Kroll, T. (2013). Chronicles: Supporting conversational narrative in alternative and augmentative communication. *Proceedings of the 2013 IFIP Conference on Human-Computer Interaction*, 364-371. Springer, Berlin, Heidelberg.

Waller, A. (2019). Telling tales: unlocking the potential of AAC technologies. *International Journal of Language & Communication Disorders*, 54(2), 159-169.

Wandmacher, T., Antoine, J. Y., and Poirier, F. (2007). SIBYLLE: a system for alternative communication adapting to the context and its user. *Proceedings of the 9th International ACM Conference on Computers and accessibility*, 203-210.

Wang, E. H., Zhou, L., Chen, S. H., Hill, K., and Parmanto, B. (2018). Development and evaluation of a mobile AAC: a virtual therapist and speech assistant for people with communication disabilities. *Disability and Rehabilitation: Assistive Technology*, 13(8), 731-739.

Webster, J., and Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, xiii-xxiii.

Widgit (2022). SymWriter 2 – Easy symbol writing for everyone. URL: https://www.widgit.com/products/symwriter/index.htm (visited on November 01, 2022).

Wisenburn, B., and Higginbotham, J. (2009). Participant evaluations of rate and communication efficacy of an AAC application using natural language processing. *Augmentative and Alternative Communication*, 25(2), 78-89.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., and Wesslén, A. (2012). Systematic literature reviews. *Experimentation in Software Engineering*. Springer, Berlin-Heidelberg. 45-54.

2.7 Personality is all that matters – Enhancing AAC Systems with Digital Personas Utilizing Generative Artificial Intelligence

<table>
<tr><td rowspan="2">Adressierte Zielsetzungen</td><td>Konzeption einer intelligenten, personalisierten Digital AAC Persona unter Zuhilfenahme von generativer KI, durch die nutzerspezifische Persönlichkeitsmerkmale adaptiv bei der Generierung der Kommunikationsbeiträge berücksichtigt werden können. Ziel ist die Entwicklung eines persönlichkeitsgetriebenen AAC-Systems zur stilistisch authentischen und nuancierten Kommunikation.</td><td>ZS10</td></tr>
<tr><td>Entwicklung und Evaluation des AAC-Prototyps zur Integration des konzipierten Digital Persona Ansatzes. Das System soll identifizierte Anforderungen an die Personalisierung eines AAC-Systems technisch und gestalterisch umsetzen und die persönlichkeitsgetriebene Generierung von Kommunikationsbeiträge durch die Digital Persona ermöglichen.</td><td>ZS11</td></tr>
<tr><td rowspan="2">Abgeleitete Forschungsfragen</td><td>Wie kann eine generative KI-basierte Digital Persona gestaltet werden, um persönlichkeitsgetriebene Äußerungen innerhalb eines AAC-Systems zu erzeugen?</td><td>FF11</td></tr>
<tr><td>Wie kann die generative KI-basierte Digital Persona implementiert und technisch realisiert werden?</td><td>FF12</td></tr>
<tr><td>Genutzte Forschungsmethode</td><td colspan="2">Die Studie folgt dem DSR-Prozess nach Peffers et al. (2007) und beinhaltet die systematische Konzeption, Entwicklung, Demonstration und Evaluation eines prototypischen AAC-Systems zur KI-gestützten, persönlichkeitsadaptiven Kommunikation.

Die Forschung orientiert sich an den drei Zyklen nach Hevner et al. (2004):
Design Cycle: Konzeption, technische Umsetzung und Evaluation des Prototyps zur Generierung personalisierter Kommunikationsbeiträge unter Berücksichtigung der identifizierten Anforderungen.
Relevance Cycle: Identifikation eines initialen, anwendungsorientierten Anforderungsprofils zur Umsetzung des AAC-Prototyps durch die Literatur. Aufbauend auf dieser initialen Anforderungsanalyse wurden die Anforderungen mit betroffenen Personen verifiziert und um die Anforderungssicht der Nutzer ergänzte.
Rigor Cycle: Theoretische Fundierung des Artefakts durch die Berücksichtigung der Theorie zur Digital Persona sowie der Integration des Big Five Persönlichkeitsmodells. Zudem wurde zur Weiterentwicklung der Nascent Design Theorie im Sinne von Gregor & Hevner (2013) ein Beitrag geleistet.
Methodisch wurde ein mehrstufiger Entwicklungsprozess durchgeführt, bestehend aus:
Ableitung von Designanforderungen aus Theorie und Nutzerbefragungen
Technische Umsetzung einer Digital Persona mittels GPT-gestütztem Sprachmodell (GPT-4), dem Big Five Persönlichkeitsmodell sowie ML-basierter Analyseverfahren (z. B. Memory Network)
Evaluation des AAC-Prototyps anhand von Experteninterviews hinsichtlich der Fähigkeit, persönlichkeitsgetreue Kommunikationsbeiträge zu erzeugen.</td></tr>
</table>

Kernergebnisse (Überblick)	Es wurde ein innovativer AAC-Prototyp entwickelt, der eine generative KI-basierte Digital Persona integriert und stilistisch-persönlichkeitsgetriebene Kommunikationsbeiträge erzeugen kann. Dieser Prototyp integriert hierfür ein LLM-gestütztes Sprachmodell (*GPT-4*) mit dem Big Five Persönlichkeitsmodell sowie ML-basierter Analyseverfahren, um nutzerspezifische, persönlichkeitsbasierte Ausdrucksweisen authentisch abzubilden. Im Rahmen der Evaluation wurde gezeigt: • Die durch die *Digital AAC Persona* generierte Sprache ist aus Perspektive der psychologischen Experten stilistisch konsistent, glaubwürdig und individuell nachvollziehbar. • Besonders im Kontext informeller Kommunikation kann die Digital Persona zur Steigerung der kommunikativen Autonomie beitragen. Neben den geleisteten Beiträgen zu den Kerneltheorien (z. B. Big Five) bildet und evaluiert die Studie zudem designbezogenes Wissen in Form von Design Prinzipien für die weitere Gestaltung KI-gestützter AAC-Systemen und liefert hierdurch einen Beitrag zur (Nascent) Design Theory (Gregor, 2006).
Publikationsort	Das Paper wurde 2025 auf der International Conference on Information Systems (ICIS), Nashville, Tennessee (USA) zur Begutachtung eingereicht.
Ranking VHB JQ 4	**A**
Autoren und Anteile	**Janik Wörner 100%**
DOI	-

Tabelle 21. Fact Sheet zur 7. Veröffentlichung.

Personality is all that matters – Enhancing AAC Systems with Digital Personas Utilizing Generative Artificial Intelligence

Completed Research Paper

Introduction

Effective communication is a cornerstone of human interaction, enabling individuals to express their thoughts, feelings, and ideas and facilitating their integration into society (Lewar et al., 2023). It is not only essential for personal expression but also critical for maintaining social relationships and a sense of belonging (Fritzell, 1996; Greenway, 2015). For individuals with speech impairments, however, the ability to communicate is severely compromised, often resulting in social isolation and a diminished quality of life (Baxter et al., 2012, Kane et al., 2017). Augmentative and Alternative Communication (AAC) systems have been developed to assist speech-impaired individuals by providing tools that enable them to express themselves by vocal language (Light et al., 2014). While AAC systems have made significant strides, their standardized outputs often fail to capture the uniqueness of the individual user, leaving communication impersonal and ineffective (Elsahar et al., 2019; Light et al., 2021). This limitation highlights the need for more intelligent, adaptive, and personalized solutions that align with the user's personality.

Personality influences not only what is communicated but also how it is conveyed, affecting the tone, style, and nuances of the interaction (Bakker-Pieper and Vries, 2013). Research has shown that personality traits such as extraversion, agreeableness, and conscientiousness significantly affect communication effectiveness across various contexts, including social, educational, and professional settings (Frederickx and Hofmans, 2014; Hassan et al., 2019). For example, teachers high in extraversion and conscientiousness often adopt expressive and precise communication styles, which resonate with their audience and improve overall communication effectiveness (Dhillon et al., 2023).

Despite the central role of personality in communication, existing AAC systems generate utterances that fail to reflect the user's unique personality. Consequently, AAC-generated messages are often perceived as rigid and lacking emotional nuance (Ascari et al., 2019). Communication is not merely about conveying information, it is a powerful tool for building relationships and expressing identity (Greenaway et al., 2015). Thus, the drawback of current AAC systems to integrate an individual's personality contributes to a sense of emotional and social disconnection for users, ultimately undermining their autonomy and social integration (Kane et al., 2017; Rayman et al., 2023).

The integration of a user's personality into AAC systems offers a promising avenue for addressing these challenges. A so-called Digital Persona represents a virtual user representation, designed to reflect their unique personality (Clarke et al., 1994; Roosendaal, 2010). By incorporating such personality-driven concepts, AAC systems can generate utterances that reflect the user's individual personality (traits), leading to more authentic and personalized interactions (Sims, 2017). Generative Artificial Intelligence (AI), particularly Generative Pre-trained Transformers (GPT), offers a powerful tool for facilitating the development of Digital Personas (Sorokovikova et al., 2024). Large Language Models (LLMs) like GPT-4.0 have demonstrated remarkable abilities in generating contextually appropriate, human-like text, interpreting, and incorporating context, tone, and personality (Brown et al., 2020; Jiang et al., 2023). LLMs are especially well-suited for developing Digital Personas, as studies indicate that advancements have enabled the measurement and manipulation of synthetic personas in automatically generated text (Serapio-Garcia et al., 2023).

The potential impact of integrating Digital Personas into AAC systems is profound. By aligning the system's output with the user's personality, not only can the authenticity and effectiveness of communication be enhanced, but also deeper emotional connections and greater social inclusion, and improved quality of life for people with speech impairments can be fostered. Specifically, this research addresses the following Research Questions (RQ):

- ***RQ1:*** *How can a Generative AI-based Digital Persona be designed to generate personality-driven utterances within an AAC system?*
- ***RQ2:*** *How can the Generative AI-based Digital Persona be instantiated and technically realized?*

The paper unfolds as follows: Next, the research background is discussed. Then, the research methodology, based on the DSR approach (Hevner et al. 2004; Peffers et al. 2007), is outlined. The following section presents the compilation of Design Principles (DPs) derived from related Design Requirements (DRs). Afterwards, the technical implementation of the AI-based Digital Persona is described, including its demonstration in a daily social life setting (doctor's visit). The paper then addresses the key contributions, potential limitations, and suggestions for future research. Finally, concluding remarks summarize the findings.

Research Background

Effective communication is deeply influenced by the individual characteristics and personality traits of the interlocutors (Greenaway et al., 2015). Studies revealed that personality significantly shapes communication behaviors, impacting how people initiate, participate in, and interpret conversations (Frederickx and Hofmans, 2014; Hassan et al., 2019). The widely known Five-Factor Personality Model (Costa and McCrae, 1992) – proven particularly effective in assessing personality traits in the field of software engineering (Jia et al., 2015) – categorizes personality into five core traits: openness, conscientiousness, extraversion, agreeableness, and neuroticism. These traits significantly impact how individuals engage in dialogues, respond to social cues, and adapt their communication styles to different contexts (Dhillon et al., 2023; Sims, 2017). However, despite the well-established link between personality and communication, research on supporting speech-impaired individuals has yet to explore how these insights can be applied within AAC systems.

AAC systems are designed to assist individuals with complex communication needs. They include strategies and techniques (e.g., symbol-based communication boards, text-to-speech technology) to facilitate communication for speech-impaired individuals, enabling them to express their thoughts, ideas, and emotions (Mamlekar et al., 2023). Users can generate speech by selecting pre-recorded phrases, composing sentences from individual words or pictograms, or using keyboard input for personalized message creation (Elsahar et al., 2019). The resulting utterance is then transformed into spoken language via a speech synthesis module, enabling voice-based communication that allows speech-impaired individuals to more authentically express their individual personalities (Kane et al., 2017).

Regardless of their crucial role in enabling communication, existing AAC systems primarily address basic communication needs without accounting for individual user preferences (Ascari et al., 2019; Mills, 2014; Light, 2021). This lack of personalization leads to communication experiences that are perceived as artificial or disconnected, constraining the effectiveness of AAC systems (Rayman et al., 2023). Although progress has been made in personalizing AAC-based communication (cf. Kocaballi et al., 2019), research indicates a need for deeper integration of user-specific characteristics (Light et al., 2021; Reichle et al., 2019).

A promising solution to these limitations lies in integrating a Digital Persona into AAC systems. Digital Personas refer to detailed and personalized virtual user representations designed to reflect their unique personality (Clarke, 1994; Roosendaal, 2010). Successfully applied in virtual assistants, Digital Personas enhance user engagement by creating more personalized and relatable experiences (cf. Belk, 2013). In AAC systems, this concept could transform standardized outputs into utterances reflecting the user's individual personality, thus enhancing the authenticity of communication (Sims, 2017), increasing user engagement and satisfaction (Light et al., 2021), and facilitating more natural communication that improves social inclusion and relationships (Moorcroft et al., 2018). Recent advancements in Generative AI, particularly in the LLM GPT-4.0 (OpenAI, 2025), have opened new avenues for implementing Digital Personas in AAC systems. Generative AI refers to computational models capable of producing human-like text that is contextually appropriate and coherent (Brown et al., 2020; Radford et al., 2019), adjusting outputs to reflect specific personality traits by fine-tuning models with personality-aligned data (Serapio-Garcia et al., 2023), and continuously learning and adapting to the user's communication style over time (Jiang et al., 2023). Further, recent studies in user modeling indicate that LLMs facilitate the accurate representation, measurement, and dynamic manipulation of synthetic user representations in automatically generated text (Serapio-Garcia et al., 2023; Jiang et al., 2023; Sorokovikova et al., 2024). Despite these advancements, there is no extant research that has integrated the concept of Digital Personas with AAC systems using Generative AI to create a comprehensive, personality-driven approach for utterance generation.

Research Procedure

To accomplish the development and integration of a Digital Persona within an AAC system that enables speech-impaired individuals to generate personalized utterances that reflect their personality (traits), DSR was applied (cf. Gregor and Hevner, 2013; Hevner et al., 2004). The research process adheres to the framework established by Peffers et al. (2007) (see Figure 1). As a first step, **(1) corresponding problems** of existing AAC tools were identified (*see Sections "Introduction" and "Research Procedure"*). Available AAC systems are not capable of generating personalized utterances that reflect an individual user's personality (traits), leading to a lack of user-specific articulation. However, integrating personality-driven utterances within AAC systems is indispensable for enabling speech-impaired individuals to express themselves authentically and for fostering more natural interactions. Consequently, the **(2) objective** is to develop, based on the identified DRs, a Generative AI-based Digital Persona in the context of AAC. Thereby, ChatGPT-4.0 is purposefully leveraged to facilitate the synthetic representation of a user (Digital Persona) that not only generates contextually and semantically correct utterances but also reflects the user's unique personality (traits), allowing for more personalized communication. The third step contains the **(3) design** and **development** of the solution (*see Section "Design and Development of the Artifact"*). Drawing on established theories (e.g., the Five Factor Personality Model (Costa and McCrae, 1992)), a systematic literature review (cf. Vom Brocke et al., 2015), and semi-structured interviews with target users, DRs for the Generative AI-based Digital AAC Persona were identified. Founded on the DRs, DPs were derived and instantiated to ensure the Digital Persona produces utterances tailored to the user's personality traits.

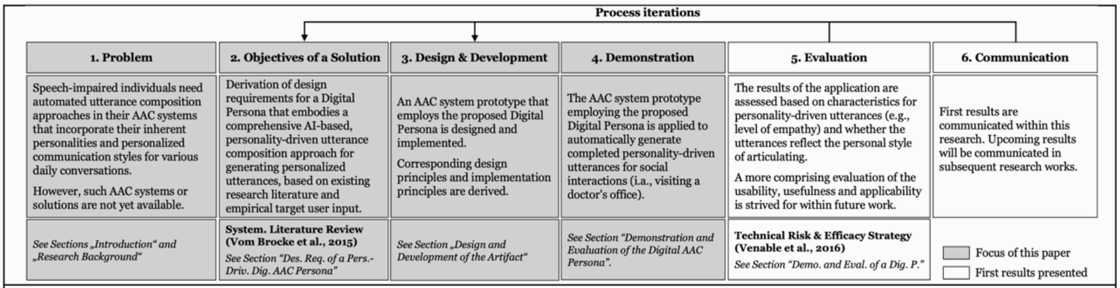


Figure 1. Procedure of this Research (own Representation, based on Peffers et al. 2007).

Steps **(4)** and **(5)** involve the **demonstration** and **evaluation** of the artifact. A rigorous testing framework was defined to assess the artifact's effectiveness. In alignment with Venable et al. (2016), an evaluation strategy based on the "Technical Risk & Efficacy Strategy" was employed to verify the artifact's ability to meet its intended purpose: generating personality-driven utterances, authentically reflecting a user's personality (traits) and their inherent communication style. The evaluation further aimed to confirm that meeting all the identified DRs and effectively leveraging GPT-4.0 is critical to achieve this outcome (cf. March and Smith, 1995). A summative evaluation was conducted, where the Digital Persona was applied to five communication activities based on a typical daily life scenario (visiting a doctor), evaluating the artifact for its accuracy in reflecting the users' personality nuances. To further ensure the quality of the derived DPs, their evaluation was guided by Iivari et al. (2021), who emphasize that effective DPs should be internally valid, externally applicable, and useful for future design efforts. These criteria were operationalized into three dimensions – internal coherence, applicability, and perceived utility – and assessed through structured evaluator feedback alongside the artifact's functional evaluation. Future evaluations will employ the "Human Risk & Effectiveness Strategy" (cf. Venable et al., 2016) in a controlled setting with target users as part of a larger empirical study. Step (6) involves the **communication** of the research findings.

Design and Development of the Artifact

Design Requirements of a Personality-Driven Digital AAC Persona

To derive DRs for a comprehensive approach to integrate a user's personality into an AAC system by employing a Digital Persona, a two-folded approach was applied: First, an extensive systematic literature review (SLR) was conducted. Second, semi-structured interviews with a focus group of six speech-impaired individuals were carried out to complement the findings from the SLR and ensure that the resulting DRs reflect the needs and expectations of actual target users.

The SLR was aligned with the guidelines as proposed by Vom Brocke et al. (2015). Initially, a keyword search was performed across the academic databases ACM Digital Library, AIS Electronic Library, EBSCOhost, Emerald Insight, Google Scholar, IEEE Xplore, SAGE Publications, and SpringerLink. The initial search terms included "augmentative and alternative communication," "personalization", "generative artificial intelligence", and various combinations and abbreviations; further, they were aligned with the wording and concepts of the research questions as proposed by Snyder (2019). Beyond that, iterative keyword refinement was employed to ensure comprehensive coverage (e.g., "large language model", "user modeling", "digital persona"). As a result, 621 articles were initially identified, focusing on different aspects of the personalization of AAC systems. Titles, abstracts, and thematic foci were thoroughly reviewed, with inclusion criteria emphasizing studies that dealt specifically with personalization in AAC systems. After applying these criteria and removing duplicates and non-English articles, a total of **33** relevant studies were retained. To assess the coding, inter-rater reliability between two independent reviewers was calculated using Cohen's Kappa (Cohen, 1960). The rating procedure yielded a Cohen's Kappa value of 0.86, indicating a high level of agreement on the inclusion of relevant articles. In the few cases of disagreement, the coding was discussed to reach a consensus about including or excluding the respective article. This rigorous approach ensured a robust and reliable selection of studies, forming a solid foundation for the derivation of DRs.

To complement the findings from the SLR by a practical, user-grounded perspective, semi-structured interviews with a focus group of six speech-impaired individuals were conducted. Participants were selected to represent a diverse range of communication impairments (e.g., congenital speech impairments, impairments resulting from neurological events such as stroke). The interviews explored participants' perspectives on personalization, communication preferences, and expectations toward a digital persona in AAC systems.

Finally, the contents of the selected articles and the transcripts of the semi-structured interviews were consolidated using qualitative content analysis and inductive category development to abstract the data and derive a coherent image of the findings (cf. Mayring, 2004). This approach facilitated the inductive derivation of DRs by applying the concept of a DR as any specific condition, capability, or functionality that the intended artifact must implement to achieve its goals and purposes (cf. Möller et al., 2020; Walls et al., 1992). Both the literature and the interview data were systematically screened for any statements that suggested a DR. Statements indicating a DR were consolidated across both sources, with the user input serving to validate, refine, or extend the DRs initially identified through the literature. Beyond that, relevant theories (e.g., Five Factor Personality Theory, Cognitive Load Theory) were analyzed to derive additional implications for the development of the personality-driven Digital AAC Persona. Based on the combined attention, validation, and practical relevance that the respective DRs received within literature and user feedback, nine DRs were formulated and organized according to three deduced categories I–III (see Table 1).

Design Requirements	**Sources**
(I) Dynamic Personalization and Adaption	
DR1. Recording and Understanding a User's Personality – The Digital Persona must record and understand a user's personality traits to adapt communication styles accordingly.	Black et al., 2012; Costa and McCrae, 1992; Fontana De Vargas et al., 2024; Franco et al., 2018; Shen et al., 2022; Sorokovikova et al., 2024);
	User-Informed
DR2. Dynamic Persona Updates – The Digital Persona must continuously update, using both real-time and historical user communication data, to adapt to evolving personality traits and needs.	Black et al., 2012; Hornero et al., 2015; Lang et al., 2023; Pereira et al., 2024; Tett and Burnett, 2003; Valencia et al., 2023; Zeng et al., 2024
DR3. AI-driven Communication Efficiency – The Digital Persona must predict and suggest personality-driven communication options based on Generative AI to improve communication efficiency.	Bhattacharjee et al., 2024; Choi et al., 2024; Cai et al., 2024 Shen et al., 2022; Sweller, 1988; Valencia et al., 2023;
DR4. Continuous Feedback – The Digital Persona must refine its personalization by user feedback, ensuring continuous adaption.	Ascari et al., 2021; Lang et al., 2023; Laxmidas et al., 2021; Monalisa et al., 2023; Wang et al., 2022
	User-Informed
(II) Contextual and Situational Awareness	
DR5. Context-Awareness for Relevant Suggestions – The Digital Persona's communication suggestions must consider personality, social, and environmental contexts.	Cai et al., 2024; Shen et al., 2022; Farzana et al., 2021; Konadl et al., 2023; Laxmidas et al., 2021; Obiorah et al., 2021; Roy et al., 2022; Zeng et al., 2024
	User-Informed
DR6. Recognition of Emotional States – The Digital Persona must adapt communication to the user's emotional state.	Na et al., 2016; Valencia et al., 2023;
	User-Informed
(III) User Inclusivity and Security	
DR7. Scalability – The Digital Persona must scale to support different users with diverse personalities, languages, and backgrounds.	Elsahar et al., 2019; Mancilla et al., 2015
DR8. Multimodal Input Support – The Digital Persona must be accessible by multiple input methods (e.g., text, voice).	Ascari et al., 2021; Ghatkamble et al., 2014

DR9. Data Privacy and Security – The Digital Persona must protect user data, ensuring privacy and security of sensitive data.	Hornero et al., 2015; Mills et al., 2014; O'Leary et al., 2012 User-Informed

Table 1. Design Requirements of a Digital AAC Persona and Related References.

(I). Concerning *Dynamic Personalization and Adaption*, **DR1** emphasizes the need for the Digital Persona to record and assess a user's personality traits to adapt communication accordingly. By integrating a personality model (c.f. Costa and McCrae, 1992), the Persona can build a detailed synthetic representation of the users' personality. This synthetic representation allows the Digital Persona to dynamically adapt its utterance generation, ensuring that it aligns with the user's personality traits and communication preferences (e.g., Franco et al., 2018; Fontana De Vargas et al., 2024; Sorokovikova et al., 2024). Furthermore, **DR2** states that the Digital Persona must dynamically update the user's synthetic representation based on real-time interactions and historical conversation data to ensure continuous personality-adaption (e.g., Lang et al., 2023; Pereira et al., 2024; Zeng et al., 2024). This presupposes the monitoring of how personality traits affect communication styles and adjusting the Digital Persona's behavior accordingly. Trait Activation Theory suggests that personality traits are activated by specific situations or cues (e.g., public speaking) (Tett and Burnett, 2003), and therefore, the Digital Persona should recognize these triggers and adjust its communication style to match the trait being expressed. **DR3** emphasizes enhancing the communication efficiency of the Digital Persona using AI (e.g., Bhattacharjee et al., 2024; Choi et al., 2024; Valencia et al., 2023). By leveraging Generative AI (e.g., GPT-4.0), the Digital Persona must predict and suggest phrases that match the user's style, tone, and preferences, which can significantly reduce cognitive load while ensuring that communication remains consistent with the user's natural expression. Cognitive Load Theory (Sweller, 1988) suggests that by reducing the amount of mental effort required to select appropriate words or phrases, the system minimizes extraneous cognitive load, allowing the user to focus more on the contents of the communication rather than the mechanics of constructing them. This ensures that the user can communicate more fluidly, with less mental strain, while maintaining personality-aligned interactions. **DR4** emphasizes the importance of continuous user feedback to refine the Digital Persona's personalization capabilities. Thus, the Digital Persona should be able to incorporate user feedback on its ability to reflect the user's personality traits, consequently ensuring that it adapts and evolves with the user's preferences over time (e.g., Ascari et al., 2021; Monalisa et al., 2023; Wang et al., 2022).

(II). With respect to *Contextual and Situational Awareness*, **DR5** highlights the importance of context awareness in providing appropriate utterances. The Digital Persona should consider the user's personality traits alongside environmental, situational, and social contexts to deliver communication options that align with the user's personality in various settings (e.g., Farzana et al., 2021; Konadl et al., 2023; Zeng et al., 2024). Further, **DR6** stresses the Digital Persona's ability to recognize emotional states and adapt its responses accordingly (Na et al., 2016; Valencia et al., 2023).

(III). To *User Inclusivity and Security*, **DR7** emphasizes the scalability of the Digital Persona, ensuring that it can accommodate a wide variety of users with, e.g., different personality traits and languages. Scalability is key to enabling the Digital Persona to remain flexible and adaptable across diverse user groups (Elsahar et al., 2019; Mancilla et al., 2015). Furthermore, **DR8** focuses on the need for multimodal input support when accessing the Digital Persona, which enables users to interact in ways that align with their personality traits and physical abilities (Ascari et al., 2021; Ghatkamble et al., 2014). Finally, **DR9** underlines the critical importance of data privacy and security, particularly when handling sensitive user data. Thus, the Digital Persona must protect user personality data (e.g., by anonymization) and ensure that all data is processed securely, maintaining user trust while providing highly personalized communication support (Hornero et al., 2015; Mills et al., 2014; O'Leary et al., 2012).

Design of a Personality-Driven Digital AAC Persona

Meta-requirements (MRs) define the overarching purpose of a system, indicating "what the system is for" (Gregor and Jones, 2007, p. 325) and shaping the solution's objectives based on the specific problems the research addresses. For instance, an MR for the Digital Persona is to empower speech-impaired individuals to express themselves through voice communication. These meta-requirements are applicable to a class of artifacts, are derived from the research literature (Gregor and Jones, 2007; Heinrich and Schwabe, 2014; Walls et al., 1992) and are presented in Figure 2.

Design Principles (DPs) are synthesized from DRs, which are informed by both existing literature and theory (Möller et al., 2020; Purao et al., 2020). DRs specify the necessary functionalities, conditions, and capabilities that an artifact must implement to support users in completing specific tasks within a defined problem domain (Möller et al., 2020; Walls et al., 1992). After carefully examining the DRs, those with overlapping meanings were consolidated into unified DPs (Möller et al., 2020; Purao et al., 2020). For example, DP2 describes the Digital Persona's ability to generate personality-driven responses, covering requirements such as personality reflection (DR1), responsiveness to dynamic user feedback (DR6), and communication efficiency (DR3). Defined as generic and prescriptive statements, the DPs foster the transfer and application of design-related knowledge to similar problem classes beyond the specific problem class of Generative AI-based Digital AAC Personas (Gregor et al., 2020).

Boundary conditions (BCs) define the constraints or limitations (e.g., technical, legal, or user-specific) within which the artifact operates, ensuring the artifact remains feasible, usable, and relevant within certain operational contexts (Chandra et al., 2015). A boundary condition for the Digital Persona is, e.g., that the system requires the user to have a sufficient cognitive level to express their personality traits and compose utterances independently. Keeping the DRs of the artifact in mind, there was derived the following design regarding the proposed AI-based Digital AAC Persona including **six DPs**:

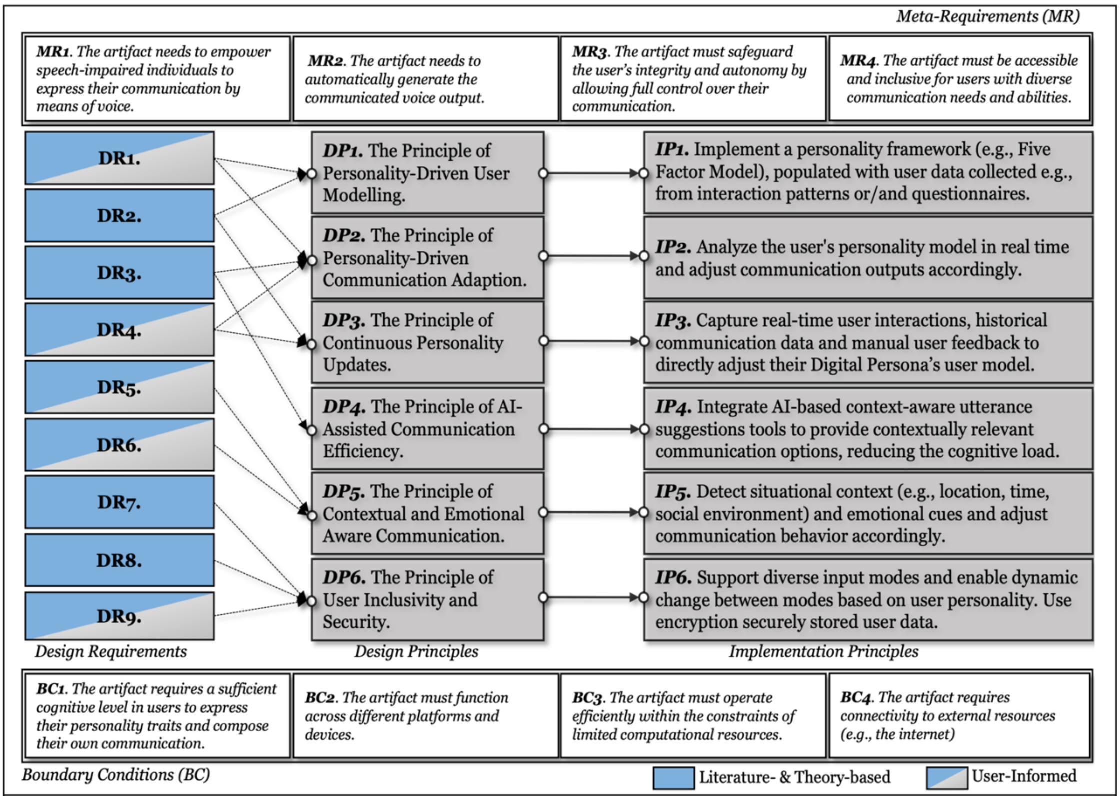


Figure 2. Design of a Personality-Driven Digital AAC Persona.

- **DP1. The Principle of Personality-Driven User Modeling**: The Digital Persona should construct and maintain a dynamic model of the user's personality grounded in established psychological frameworks. This user model should reflect the user's stable characteristics and preferences over time, ensuring accurate mirroring of the user's core personality (traits).
- **DP2. The Principle of Personality-Driven Communication Adaption**: The Digital Persona should dynamically adapt its communication style based on the user's personality, ensuring all interactions reflect the user's unique traits, communication preferences, and natural way of expressing. This ensures that communication feels personalized, natural, and relevant to the user.
- **DP3. The Principle of Continuous Personality Updates**: The Digital Persona should update the user's personality model and communication preferences using real-time interactions, historical data, and user feedback. This ensures that the Digital Persona evolves with the user's changing traits and communication needs, maintaining ongoing accuracy in reflecting the user's personality.

- **DP4. The Principle of AI-Assisted Communication Efficiency**: The Digital Persona should use AI-based models to predict, suggest, and refine communication options tailored to the user's personality. This reduces cognitive load for the user while improving communication efficiency by offering personalized and contextually relevant communication options.
- **DP5. The Principle of Contextual and Emotional Aware Communication**: The Digital Persona should adapt its communication based on the user's personality alongside the situational context and emotional states. This ensures that responses are contextually appropriate, emotionally aligned, and relevant to the user's current circumstances.
- **DP6. The Principle of User Inclusivity and Security**: The Digital Persona should support multimodal interaction methods (e.g., text, voice) and accommodate diverse languages and personalities while ensuring that user data is protected through robust security measures. This fosters accessibility and trust, ensuring the system is usable by a broad audience while safeguarding user information.

Figure 2 visually illustrates how the MRs, DPs, and BCs interrelate within the design of the proposed artifact, providing a comprehensive overview of how the system's overarching objectives, core functionalities, and operational constraints are integrated. The proposed DPs are derived from DRs grounded in current research literature, established theoretical frameworks, and empirical insights from target users. Gregor and Jones (2007) emphasize that theory-backed references provide "an explanation of why an artifact is constructed as it is and why it works" (p. 328). The DRs outline the core requirements that a Digital AAC Persona should meet, drawing from relevant research and theory. These DRs guide the design of the artifact and shape the DPs (Böckle et al., 2021; Gregor and Jones, 2007). Each DP corresponds to at least one DR, serving as an abstract "blueprint" for the artifact, thereby ensuring value and explanatory grounding (as the DPs are based on current literature and insights from target users) (Heinrich and Schwabe, 2014). Implementation Principles (IPs) emerge from applying the DRs and DPs in the artifact, supporting "the implementation in practice of an abstract, generic design method or development approach" (Gregor and Jones, 2007, p. 329), specifically for AI-driven Digital Persona-based utterance composition in AAC systems.

Technical Realization of a Personality-Driven Digital AAC Persona

To overcome the limitations of existing AAC systems, a Digital Persona artifact has been designed and developed in Python, chosen for its versatility and suitability for AI-based applications. This artifact has been integrated into a more comprehensive AAC system prototype, enabling application and user interaction with the Digital AAC Persona. The following section details the technical instantiation (IP) of each DP by elaborating on the technical realization of each associated DR, as they represent the essential features of the proposed AI-based approach.

DP1. In alignment with ***DR1***, the Digital Persona employs a custom questionnaire based on the well-established Big Five Inventory (Rammstedt et al., 2007) to assess the user's personality traits. Responses are measured on a Likert scale from "strongly disagree" (0) to "strongly agree" (5), encapsulating user preferences across the five personality dimensions. This personality information, along with general user data (e.g., pseudonym, age, vocal preferences), is stored as JSON in an AES-encrypted foundation document, which serves as the core representation of the user (Holzinger et al., 2023).

DP2. Building upon ***DR1***'s foundation, the data stored in the foundation document directly informs the adaptive communication capability of the Digital Persona. The AAC system, serving as a middleware, dynamically adjusts the prompts provided to GPT-4.0 based on the user's personality traits, ensuring that each interaction aligns with the user's personality-driven communication preferences. To further refine the personalized communication experience, few-shot learning has been leveraged by incorporating a curated set of conversation examples that reflect the user's preferred communication style (***DR2***).

DP3. To achieve continuous personalization in line with ***DR4***, the Digital Persona adapts to evolving personality traits through both implicit and explicit user feedback. Users can provide manual feedback on generated utterances, facilitating immediate refinement. Additionally, the system autonomously infers feedback from user interactions; once a modified utterance is finally approved by the user, the system stores the approved utterances in a Memory Network, which serves as a foundational dataset for few-shot learning. This allows the model to draw on a history of approved interactions to produce responses that align with the user's evolving communication style and preferences. By selectively retaining utterances that reflect current personality traits and preferences, the Digital Persona supports continuous adaptation, ensuring that the responses remain dynamically aligned with the user's evolution over time.

DP4. As the core of the Digital Persona, GPT-4.0 generates responses that are contextually relevant and personalized based on the user's personality data, fulfilling ***DR3***. Additionally, the Memory Network's static and dynamic components utilize an adaptive pruning mechanism, which effectively manages memory constraints by retaining essential interactions while removing irrelevant ones (Shao and Shin, 2022). Leveraging this set of AI-based techniques ensures efficient data retrieval, supports continuity in communication, and reduces cognitive load, thereby enhancing user engagement.

DP5. To deliver contextually and emotionally aware communication following ***DR5***, the Digital Persona adapts its responses based on situational factors and the user's emotional state. Contextual information is derived from sensor data (e.g., microphone inputs). For example, the device's microphone detects ambient noise levels, allowing the system to distinguish between a quiet one-on-one setting and a louder group environment. This enables the Digital Persona to adjust responses accordingly, offering concise replies in noisy settings and more detailed ones in quieter situations. Emotional awareness is supported through dynamic emotion selection and sentiment analysis. Users can manually indicate their current emotional state via on-screen emotion buttons, facilitating real-time tone modulation. Additionally, sentiment analysis on refined utterances allows the system to autonomously detect mood shifts, aligning new responses with the user's emotional state (***DR6***).

DP6. To ensure versatility and security, the Digital Persona is designed to accommodate diverse user needs across multiple dimensions, meeting the standards set by ***DR7***. Built on a modular, cloud-based infrastructure utilizing GPT-4.0, the system is scalable and can support different personalities, languages, and cultural backgrounds (Radford, 2019). It enables multimodal input by integrating both text and voice options (***DR8***); voice input utilizes automatic speech recognition to facilitate seamless interaction for users who prefer or require vocal communication. AES encryption and anonymization mechanisms (e.g., pseudonymization) safeguard sensitive user information, with strict access controls to maintain data privacy, ensuring that user data remains secure throughout interactions (***DR9***).

Lifecycle of a Personality-Driven Digital AAC Persona

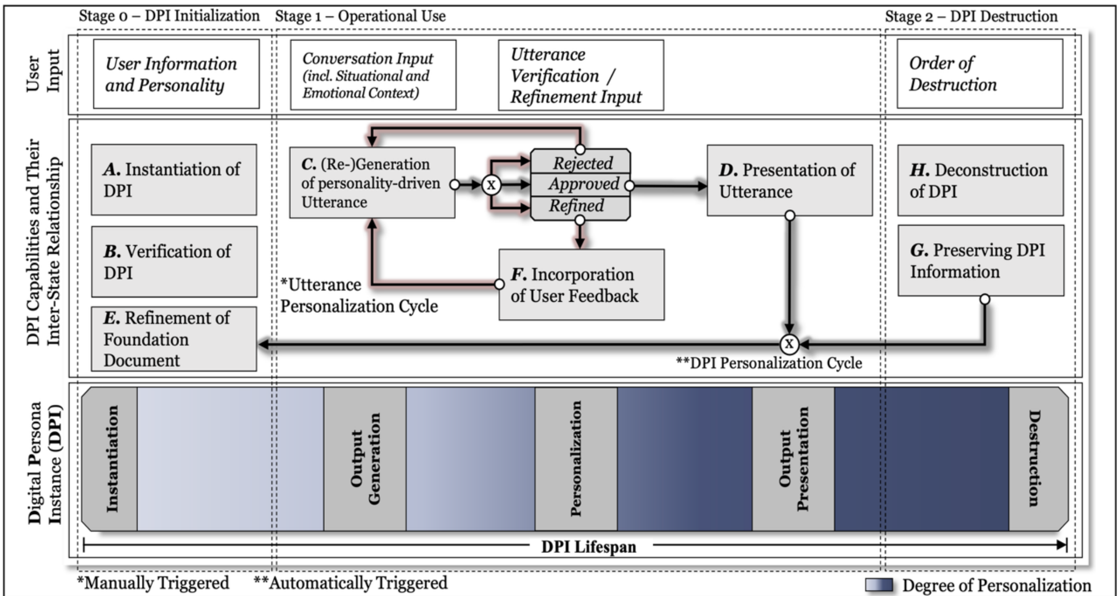


Figure 3. State-based Lifecycle Model of a Personality-Driven DPI for AAC.

Within the AAC system, the Digital Persona Instance (DPI) operates to produce customized utterances that resonate with the user's unique communicative style. Figure 3 delineates the dynamic lifecycle of the DPI, organized into five key states: Instantiation, Output Generation, Personalization, Output Presentation, and Destruction. These states are consolidated into three stages: Stage 0 (DPI Initialization), Stage 1 (Operational Use), and Stage 2 (DPI Destruction), each representing critical stages in DPI lifecycle management. The model exemplifies the DPI's functionalities and inter-state connections across these stages, illustrating how user input (e.g., provision of conversation input, approval of utterances) prompts particular transitions within the state sequence. The sequence flow, denoted by the unidirectional arrows, highlights two primary cycles within the DPI lifecycle: the utterance refinement cycle and the DPI personalization cycle.

The utterance refinement cycle addresses the iterative process of adjusting proposed utterances based on manual user feedback, thereby enabling each output to reflect the user's preferences more accurately. Further, upon user approval of an utterance, the personalization cycle initiates automatically. With each approval, the DPI captures personalized communication knowledge (e.g., preferred vocabulary) reflective of the user's personality, subsequently integrating this knowledge into the DPI's foundation information. This automatic personalization mechanism ensures that the Digital Persona continues to evolve seamlessly without requiring additional manual input from the user.

Stage 0 – For the DPI Instantiation, the user's personality traits are assessed by the AAC system using the Five Factor Personality Model. This assessment, along with additional user information (e.g., pseudonym, age), is stored in a foundation document, serving as a central repository for all relevant user data during the instantiation of a Digital Persona (Holzinger et al., 2023). The foundation document provides the basis for personalization by capturing the user's general information and personality traits, thereby enabling the Digital Persona to dynamically adapt its responses to reflect the user's unique personality. Based on the foundation document, the Digital Persona is instantiated (***A***) within the LLM GPT-4.0, creating a DPI that encapsulates the user's personality traits and communication preferences. Once instantiated, the DPI undergoes technical verification to ensure a stable connection, data access, and reliable functionality within the system architecture (***B***).

Stage 1 – During Operational Use, the user provides conversation input (e.g., voice, text) via the AAC interface. This input is transformed into a contextualized prompt that incorporates the user's real-time situational cues (e.g., social setting) and emotional states (e.g., anxiety). The DPI then processes the prompt and generates a contextually appropriate, personality-driven utterance that mirrors the user's personality (***C***). The generated utterance is then presented to the user for verification. If the user approves it, the DPI presents the final utterance to the interlocutor via the AAC system's speech module (***D***). Further, the foundation document is updated with the successful utterance, strengthening the basis for few-shot learning that informs future DPIs by leveraging the accumulated communication knowledge (***E***). If the user refines the output, the DPI incorporates the user feedback (***F***) and updates the generated utterance accordingly (***C***). If the user rejects the output, the DPI regenerates the utterance by resampling the set of few-shot learning examples, ensuring a different communication style (***C***).

Stage 2 – When the interaction is closed, the DPI is destructed. Prior to preserving the gained knowledge for future instantiations of Digital Personas, the conversation and associated personalization data are consolidated, condensed, extracted, and integrated into the foundation document (***G***). This process ensures that, upon re-instantiation in Phase 0, the foundation document (and thus the new DPI) includes the prior conversation and personalization knowledge, enabling continuous personalization. Finally, once the system preserved the gained knowledge, the DPI is destructed (***H***).

Demonstration and Evaluation of the Digital AAC Persona

The demonstration and evaluation align with the outlined evaluation strategy (*see Section "Research Procedure"*) and focus on the context of a doctor's visit, a critical interaction scenario for many speech-impaired individuals (Morris et al., 2013). According to Venable et al. (2016), the first essential step when evaluating an artifact is to assess its capabilities and functionality by conducting a "Technical Risk & Efficacy" evaluation phase, primarily focusing on quantitative metrics to gauge the artifact's applicability. Thus, this evaluation initially assesses the Digital AAC Persona's ability to generate personality-driven, context-sensitive communication outputs. Future evaluation episodes dedicated to qualitative feedback will include a "Human Risk & Effectiveness" evaluation strategy (Venable et al., 2016) (*see Section "Research Procedure"*), incorporating explicit feedback from the target user group. This section includes (a) a comparative overview of two fictive personas (see Table 2) to illustrate the Digital Persona approach, (b) a demonstration of the handling of the integrated Digital Persona within the AAC system prototype, and (c) an evaluation assessing the Digital Persona's ability for generating personalized, context-aware utterances.

(***a.***) Two fictive personas, Anna and Ben, were developed to illustrate the Digital Persona approach in an AAC context. Their distinct personality traits serve as benchmarks for evaluating the prototype's ability to generate personality-driven utterances. Table 2 provides a comparative overview of Anna and Ben's personality traits. The exemplary effects on the expression of these personality traits are derived from existing literature (e.g., Dhillon et al., 2023; Hassan et al., 2019; Bakker-Pieper and Vries, 2013). Anna, characterized by high openness and extraversion, approaches conversations with curiosity and warmth, aiming to create an engaging atmosphere. In contrast, Ben demonstrates high conscientiousness and introversion, preferring structured and precise communication to ensure clarity and professionalism.

Trait	Anna	Exemplary Effects	Ben	Exemplary Effects
Openness	4/5	Curious, imaginative, enjoys exploring new ideas.	1/5	Prefers familiar routines and concrete, practical discussions.
Conscientiousness	3/5	Flexible and adaptable, prefers spontaneous approach.	5/5	Organized, methodical, values structure and accuracy.
Extraversion	5/5	Outgoing, energetic, enjoys leading conversations.	1/5	Reserved, prefers listening, speaks only, when necessary, in a calm way.
Agreeableness	5/5	Warm, empathetic, aims to make others feel comfortable.	2/5	Polite and respectful but maintains a professional tone.
Neuroticism	1/5	Resilient, stays calm and optimistic, even in stress.	4/5	Cautious in unfamiliar situations, seeks reassurance by precise communication.

Table 2. Comparative Overview of Fictive Personas Anna and Ben.

(***b.***) Figure 4 illustrates an example of the prototype leveraging Anna's Digital Persona. It comprises three primary graphical user interfaces. The (*1*) profile screen enables customization of user information and personality traits, assessed through validated questionnaires based on the Five Factor Model. Scores from these assessments are consolidated to form a comprehensive trait profile that defines the Digital Persona's characteristics. The (*2*) main interface provides input options (i.a., pictograms, keyboard interface) for generating communication outputs. When generating a personalized phrase, the Digital Persona processes the user's input, conversational context, and emotional state to produce a personality-driven utterance. The (*3*) context manager supports situational adaptability by allowing to define contextual and emotional information. Through integration with a personal calendar, the Digital Persona incorporates historical and planned events, ensuring responses are contextually relevant. Moreover, users can specify their current emotional state and social setting (e.g., group vs. one-on-one conversation).

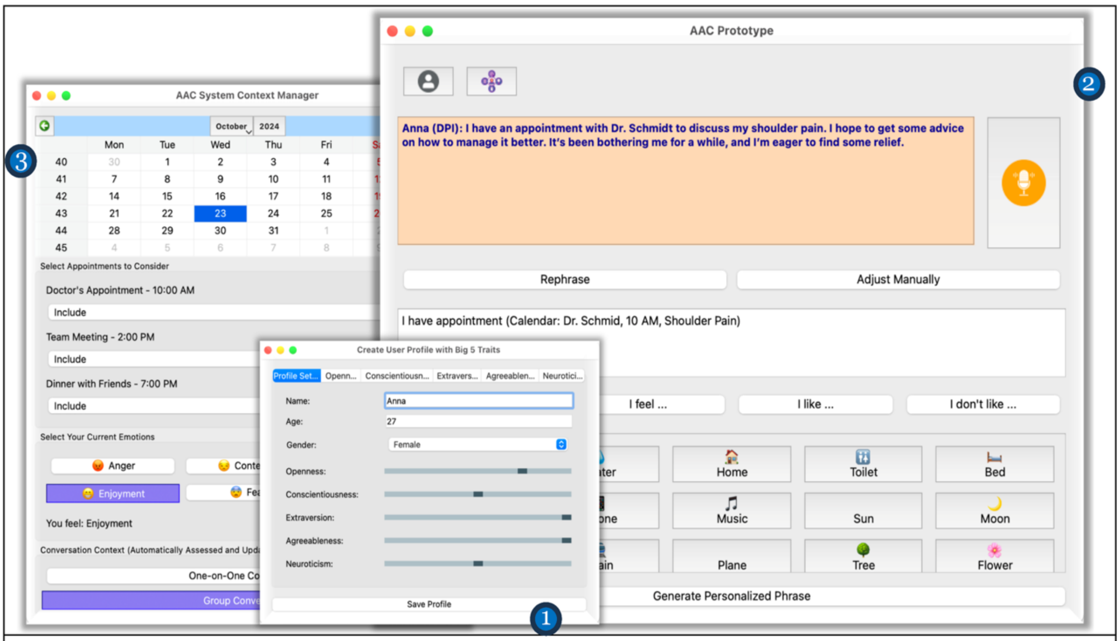


Figure 4. Exemplary usage of the AAC Prototype utilizing Anna's DPI.

(***c.***) To evaluate the DPI's efficacy, five standardized conversational exchanges were defined based on common interactions when visiting a doctor's office. This evaluation aimed to assess the system's ability to generate personality-driven responses tailored to the user's contextual and emotional nuances. For each of the five interactions, ten responses were generated for each Digital Persona, yielding a total of 100 responses. Seven independent evaluators rated each response based on three dimensions: personality alignment (***P***), situational context (***S***), and emotional appropriateness (***E***). The evaluators were briefed on both, the Five-

Factor Personality traits (as outlined in Table 2) as well as the specific trait levels and exemplary effects of each persona to ensure consistent interpretation and reduce subjective variability in assessments. The evaluator panel comprised four psychological experts and three technically literate laypersons, totaling seven independent evaluators. The psychological experts contributed their specialist knowledge of personality assessment, thereby ensuring that the Digital Persona's reflective capacity was in line with the Five Factor Personality traits. Meanwhile, the lay evaluators, although not formally trained in psychological research, offered valuable insights based on their innate human ability to recognize and interpret personality traits. This heterogeneous mix of evaluators provided a well-rounded evaluation, using both academic rigor and everyday intuition to judge the artifact's effectiveness in promoting a more personalized communication experience. Each response was analyzed regarding its ability to convey the persona's specific communication style (e.g., Anna's warmth and openness, or Ben's reserved and precise tone) in conjunction with situational context (e.g., conversation start) and emotional sensitivity (e.g., enjoyment). Accuracy measures for personality alignment (***A(P)***), situational context (***A(S)***), and emotional appropriateness (***A(E)***) were calculated by assessing the frequency with which evaluators perceived each respective dimension as correctly reflected (see Table 3). In addition, the standard deviation (***SD***) was determined to evaluate the consistency of perceptions across evaluators.

Comm. Activity / User's Input	Exemplary Response (**Anna's DPI**)						Exemplary Response (**Ben's DPI**)					
	A(P)	*SD*	*A(S)*	*SD*	*A(E)*	*SD*	*A(P)*	*SD*	*A(S)*	*SD*	*A(E)*	*SD*
Conversation Start	*Receptionist: "Hello, how can I help you today?"*											
"I have appointment (Calendar: Dr. Schmid, Shoulder Pain)."	Hi there! I have an appointment with Dr. Schmidt to discuss my shoulder pain. I hope to get some advice on how to manage it better. It's been bothering me for a while, and I'm eager to find some relief.						Good morning. I have an appointment with Dr. Schmid at 10:30. Could you confirm if my records are ready?					
	0.84	0.37	0.97	0.17	0.77	0.42	0.9	0.3	0.97	0.17	0.54	0.5
Waiting Room Inquiry	*Receptionist: "The doctor will be with you shortly. Do you need anything else?"*											
"No. I wait."	Thank you! I'll just be here with my magazine, but if anything changes, feel free to let me know!						Thank you. Could you let me know if there are any updates on the expected wait time?					
	0.93	0.26	0.9	0.3	0.79	0.41	0.83	0.38	0.86	0.35	0.43	0.5
Doctor's Conversation	*Doctor: "Hello, what brings you in today?"*											
"Hello. I shoulder pain."	Hi, Dr. Schmid! I've had this shoulder pain that won't seem to go away. I'd love to get your take on what might be going on.						Hello, Dr. Schmid. I've been experiencing some recurring symptoms that I'd like to discuss in detail. Specifically, I'm concerned about consistent shoulder pain.					
	0.86	0.35	0.91	0.28	0.73	0.45	0.97	0.17	0.94	0.23	0.81	0.39
Treatment Clarification	*Doctor: "We'll start with medication and see how it goes. Does that sound good?"*											
"Yes. Side effects?"	Absolutely, thank you! I'm happy to give it a try. Any side effects I should know about?						Could you clarify the dosage and expected side effects? I'd like to be sure I understand the regimen fully.					
	0.96	0.2	0.9	0.3	0.96	0.2	0.91	0.28	0.9	0.3	0.77	0.42
Conversation End	*Doctor: "That should be all for today. We'll follow up in two weeks."*											
„Thank you."	Thank you so much, Dr. Schmid! I really appreciate your help, and I'll be sure to keep you updated.						Thank you, Dr. Schmid. I appreciate your time and guidance. I'll be sure to follow the plan and monitor my symptoms.					
	0.91	0.28	0.83	0.38	0.83	0.38	0.86	0.35	0.83	0.38	0.53	0.5
Mean Score	***A(P)***	***SD***	***A(S)***	***SD***	***A(E)***	***SD***	***A(P)***	***SD***	***A(S)***	***SD***	***A(E)***	***SD***
	0.9	**0.29**	**0.9**	**0.29**	**0.81**	**0.37**	**0.89**	**0.3**	**0.9**	**0.29**	**0.61**	**0.46**

Table 3. Accuracy in Reflecting Personality-driven Differences in DPI Utterance Generation.

The evaluation reveals strong performance in generating responses aligned with predefined personality traits and situational contexts, as reflected in high mean scores for personality alignment and situational context across both personas ($A(P_{Anna})$ and $A(S_{Anna})$ with 0.9; $A(P_{Ben})$ with 0.89 and $A(S_{Ben})$ with 0.90), accompanied by low standard deviations (SD) ranging from 0.29 to 0.30, indicating high inter-rater agreement. However, a notable variation emerged in the emotional awareness, where Anna achieved a higher average (0.81, SD = 0.37) compared to Ben (0.61, SD = 0.46), highlighting differential performance in the Digital Persona's ability to infuse emotional sensitivity into responses. Anna's DPI consistently demonstrated her open, extroverted characteristics with friendly and engaging responses, thus scoring highly in emotional tone alignment and contextual appropriateness. In contrast, Ben's responses, while precise and professionally aligned, revealed a narrower emotional scope, particularly in settings where empathy or a softer tone may be expected. Ben's DPI maintained a more factual tone, aligning well with his conscientious persona but occasionally missing the emotional nuance. These findings imply that, while the AAC prototype stands out in replicating distinct personality traits and contextual relevance, there is room for enhancing responsiveness to emotional cues in reserved personas.

Finally, to assess the DPI's ability to preserve continuity in aligning to the user's personality while retaining conversation context, the Digital Personas were both re-instantiated in a follow-up session to evaluate their awareness of the prior doctor's visit. Both Anna's and Ben's DPIs demonstrated a memory of past interactions, referencing prior discussions on shoulder pain and treatment steps, which facilitated a seamless continuation of the conversation. For example, in the identically-recreated Communication Activity "Doctor's Conversation" (see Table 3), Anna's DPI responded:

> "*Hi Dr. Schmid! It's wonderful to see you again. I wanted to chat because even after taking the medications you prescribed last week, I'm still experiencing some shoulder pain. I'm really interested in exploring other treatment options with you.*".

This retention of conversational history supports not only continuity but also the nuanced evolution of each persona's characteristics, allowing the DPIs to build on previous exchanges in a way that reflects their personality-driven perspectives. This capacity for context retention enhances the AAC system in creating a coherent, evolving communication experience, ensuring that each interaction aligns with the user's ongoing personality traits and fosters a more personalized, naturalistic dialogue over time.

Evaluation of the Derived Design Principles

Beyond evaluating the artifact's technical functionality, the quality of the derived DPs is evaluated, ensuring that the DPs are not only effective within the immediate artifact but also reusable and prescriptive for future AAC system designs. Following the evaluation framework proposed by Iivari et al. (2021), which identifies internal validity, external validity, and prescriptive usefulness as essential evaluation criteria, this study operationalized these criteria into three evaluation dimensions: internal coherence, applicability, and perceived utility. Internal coherence was defined as the logical consistency and clarity of the DP's formulation. Applicability referred to the ease with which a DP could be instantiated into concrete features within an AAC system. Perceived utility captured the evaluators' judgment on whether the DP would assist other researchers in developing similar systems. Each evaluator assessed each of the six DPs using a 5-point Likert scale, with each dimension assessed through a corresponding set of four targeted statements (e.g., 1. Statement for Int. Coherence: "The Design Principle is formulated clearly and unambiguously."). To evaluate the DPs, the same panel of evaluators – comprising four psychological experts and three technically literate laypersons – was engaged. The results are presented in Table 4.

Design Principle	Int. Coherence (M | SD)		Applicability (M | SD)		Utility (M | SD)	
DP1	4.71	0.49	4.57	0.53	4.29	0.49
DP2	4.43	0.53	4.14	0.38	3.86	0.69
DP3	4.86	0.38	4.71	0.49	4.57	0.53
DP4	4.29	0.49	3.86	0.38	3.71	0.49
DP5	4.57	0.53	4.57	0.53	4.14	0.38
DP6	4.00	0.58	3.71	0.49	3.71	0.49

Table 4. Evaluation Results of the DPs - Mean scores (M) and Standard deviation (SD) of evaluator ratings (n = 7) for each DP, assessed on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree).

As shown in Table 4, the evaluation of the six DPs yielded consistently high ratings for internal coherence, with mean values ranging from 4.00 (DP6) to 4.86 (DP3) and standard deviations between 0.38 and 0.58. These results indicate that the DPs were perceived as logically sound, clearly articulated, and largely consistent across evaluators. Similarly, applicability was rated favorably across all principles, with DP3 receiving the highest rating (M = 4.71, SD = 0.49) and DP6 the lowest (M = 3.71, SD = 0.49). The narrow range and low variability across the applicability suggest that the DPs can be translated effectively into artifact features. Perceived utility, while exhibiting slightly greater variance, also demonstrated generally positive outcomes. DP3 (M = 4.57, SD = 0.53) and DP1 (M = 4.29, SD = 0.49) received the highest evaluations in terms of their potential to support future AAC system design. Even the lowest utility scores, assigned to DP4 and DP6 (M = 3.71, SD = 0.49), remain above the scales' neutral midpoint, indicating that all DPs were regarded as relevant and potentially reusable. Taken together, these results reflect a high degree of conceptual robustness and practical relevance across all proposed DPs, supporting their utility as prescriptive knowledge for guiding the design of personality-driven AAC systems.

Discussion and Contributions

The implementation of the DRs and DPs demonstrated notable success in enhancing personalized communication within AAC systems. By utilizing GPT-4.0 in combination with various machine learning techniques (e.g., automated speech recognition, reinforcement learning), a Digital Persona capable of reflecting users' personality traits and generating contextualized and personality-driven utterances has been developed (*see Section "Demonstration and Evaluation of the Digital AAC Persona*"). This supports speech-impaired users by alleviating the cognitive demands associated with producing personality-driven and grammatically accurate utterances. By correctly depicting the users' unique personality traits, consistently maintaining conversational context, and adapting to the users' emotional cues, the generated utterances collectively foster a more tailored and immersive interaction experience, enhancing users' autonomy and engagement. As the conducted evaluation suggests, the Digital Persona can also retain conversational history, allowing for adaptive responses that evolve with ongoing interactions. This not only promotes effective communication for speech-impaired users but also strengthens their independence and capacity for continuous integration in daily life. In terms of the DSR knowledge contribution framework of Gregor and Hevner (2013), this DSR-project is positioned in the group of improvements (new solutions for known problems). DSR improvement projects contribute to both prescriptive, i.e., design theory (Gregor, 2006), and descriptive theory, i.e., kernel theories (e.g., Five Factor Personality Model) (cf. Gregor and Hevner, 2013). In this way, contributions to IS, Sociopsychology, and DSR research are made alike.

From a practical perspective, the Digital Persona approach contributes by tailoring automated utterance generation to reflect the user's personality and thus by enabling AAC users to communicate in ways that authentically convey their unique communication styles. The ability to produce responses that are contextually and emotionally appropriate and aligned with the user's personality underscores the system's effectiveness, validated by high accuracy scores in personality alignment and situational relevance (see Table 3). Notably, the system's adaptive, contextual awareness allows it to differentiate responses across varied environments – whether casual (e.g., responding to a waiting room inquiry) or formal (communicating with the doctor) – adding a layer of responsiveness, ultimately enhancing user versatility and satisfaction. Despite these advancements, evaluation results also indicate areas for growth, particularly in enhancing emotional sensitivity for reserved personalities, where responses may occasionally lack the warmth needed in sensitive settings. This could be addressed by explicit fine-tuning on emotional expressions that modulate responses based on subtle affective markers, especially for interactions demanding warmth or empathy.

Further, the investigation achieved theoretical contributions that go beyond the technical contribution (i.e., the artifact) by applying the Five Factor Personality Model (Costa and McCrae, 1992) within the AAC field. The ability to capture and reflect personality traits (e.g., extraversion, agreeableness) demonstrates how AAC technology can leverage these traits to produce communication outputs that reflect a user's social and emotional communication preferences. This provides a new pathway for applying personality theory in practical, adaptive technologies, thus reinforcing the value of the Five Factor Personality Model as a framework for personalized digital interactions. Furthermore, the study advances Cognitive Load Theory (Sweller, 1988) by illustrating how Generative AI can reduce cognitive strain for AAC users. The system's AI-driven language generation lowers the mental load associated with constructing utterances, enabling users to focus on the content of communication rather than its mechanics. This reduction in cognitive demand is particularly valuable for speech-impaired individuals, who often face high cognitive loads in crafting coherent, socially attuned messages (Valencia et al., 2023).

Finally, DPs were synthesized based on the DRs derived from the literature. By applying them during the design and development of the artifact, followed by the demonstration and evaluation, an implicit empirical grounding of the DPs was achieved (Heinrich and Schwabe 2014). In addition, a structured evaluation was conducted to assess the DPs in terms of their internal coherence, applicability, and perceived utility (see Table 4), thereby providing an explicit empirical grounding for their validity and practical relevance. The evaluation yielded consistently promising results, confirming the conceptual clarity and design utility of the DPs. These findings support their role in guiding the development of future IS (design) theories and new artifacts and thus contribute to design theory, specifically as a theory for design and action (Gregor, 2006).

Conclusion

To facilitate the social integration of speech-impaired individuals, it is essential to empower them to communicate their genuine thoughts and feelings in a manner that reflects their unique communication style. However, existing AAC solutions and prior research have not yet considered the incorporation of a user's personality, restricting their ability to communicate in ways aligned with their true self. This study addresses this gap by proposing specific DRs and DPs for a Generative AI-based approach that integrates an AAC user's personality into the aided communication process. Generative AI-based techniques (e.g., GPT-4.0) are purposefully combined with machine learning (e.g., automated speech recognition, reinforcement learning), resulting in an automated, personality-driven, and context-aware utterance generation approach (Digital Persona). This approach aims to close the identified gaps in existing research as outlined in RQ1 and RQ2. The application and evaluation of the Digital AAC Persona in a replicated real-life scenario (doctor visits) underscored the effectiveness of the approach, thus highlighting the importance of meeting all derived DRs to achieve a comprehensive and effective approach.

This research advances practice and theory alike: The developed artifact empowers speech-impaired individuals to express their true selves by generating personality-driven utterances, effectively reducing communication barriers and thus fostering more authentic, personalized, and effective communication. Further, the study contributes to theory by, e.g., applying the Five Factor Personality Model to AAC and consequently extending it from theoretical psychology to adaptive communication technology.

Despite the comprehensive scope of this research, there are certain limitations. While the study identifies a robust set of DRs, further interdisciplinary research could uncover additional requirements. Though the research makes initial contributions to the Five Factor Personality Theory and Cognitive Load Theory, more extensive investigations are needed to fully mature these theoretical contributions. While the initial evaluation of the artifact yielded promising results, the absence of direct involvement of the target user group (i.e., speech-impaired individuals) limits the external validity of the findings. Thus, future research will address this limitation by incorporating end-user perspectives within the evaluation (*see Section "Research Procedure"*) to enhance practical relevance and support broader generalizability. Nevertheless, the contributions made are undoubtedly significant for research in both AAC and Sociopsychology.

References

Ascari, R., Silva, L., and Pereira, R. 2019. "Personalized interactive gesture recognition assistive technology," *Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems.*

Bakker-Pieper, A., and Vries, R. 2013. "The Incremental Validity of Communication Styles Over Personality Traits for Leader Outcomes," *Human Performance* (26:1).

Baxter, S., Enderby, P., Evans, P., and Judge, S. 2012. "Barriers and facilitators to the use of high-technology AAC devices: a systematic review and qualitative synthesis," *Int. J. Lang. Commun. Disord. (*47:2).

Belk, R. 2013. "Extended Self in a Digital World," *Journal of Consumer Research* (40:3).

Bhattacharjee, A., Zeng, Y, Xu, S., and Kulzhabayeva, D. 2024. "Understanding the Role of LLMs in Personalizing and Scaffolding Strategies to Combat Academic Procrastination," *CHI.*

Black, R., Waller, A., Turner, R., and Reiter, E. 2012. "Supporting Personal Narrative for Children with Complex Communication Needs," *ACM Trans. Comput. Hum. Interact.* (19:2).

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., and et al. 2020. "Language Models are Few-Shot Learners," *arXiv:2005.14165.*

Böckle, M., Bick, M., and Novak, J. 2021. "Toward a Design Theory of User-Centered Score Mechanics for Gamified Competency Development," *Information Systems Management* (40:1).

Cai, R., Janaka, N., Chen, Y., Wang, L., Zhao, S., and Liu, C. 2024. "PANDALens: Towards AI-Assisted In-Context Writing on OHMD During Travels," *CHI.*

Chandra, L., Seidel, S., and Gregor, S. 2015. "Prescriptive Knowledge in IS Research: Conceptualizing Design Principles in Terms of Materiality, Action, and Boundary Conditions," *HICSS.*

Choi, D., Lee, S., Kim, S.-I., and Lee, K. (2024). "Unlock Life with a Chat(GPT): Integrating Conversational AI with LLMs into Everyday Lives of Autistic Individuals," *CHI.*

Clarke, R. 1994. "The Digital Persona and Its Application to Data Surveillance," *Inf. Soc.* (10:2).

Cohen, J. 1960. "A coefficient of agreement for nominal scales," *EPM* (20:1).

Costa, P., and McCrae, R. 1992. "The Five-Factor Model of Personality and Its Relevance to Personality Disorders," *Journal of Personality Disorders* (6:4).

Dhillon, N., and Kaur, G. 2023. "Impact of Personality Traits on Communication Effectiveness of Teachers: Exploring the Mediating Role of Their Communication Style," *SAGE Open* (13:2).

Elsahar, Y., Hu, S., Bouazza-Marouf, K., Kerr, D., and Mansor, A. 2019. "AAC Advances: A Review of Configurations for Individuals with a Speech Disability," *Sensors* (19:8).

Farzana, W., Sarker, F., Chau, T., and Mamun, A. 2021. "Technological Evolvement in AAC Modalities to Foster Communications of Verbally Challenged ASD Children: A Systematic Review," *IEEE Access*.

Franco, N., Fidalgo, R., Cunha, P., Nóbrega, O., and Ramos, A. 2018. "Customizing Usability Heuristics for AAC Systems," *EATIS*.

Frederickx, S., and Hofmans, J. 2014. "The Role of Personality in the Initiation of Communication Situations," *Journal of Individual Differences* (35:1).

Fritzell, B. 1996. "Voice disorders and occupations," *Logopedics Phoniatrics Vocology* (21:1).

Fontana De Vargas, M., Yu, C., and Shane, H. 2024. "Co-Designing QuickPic: Automated Topic-Specific Communication Boards from Photographs for AAC-Based Language Instruction," *CHI*.

Ghatkamble, R., Son, J., and Park, D. 2014. "A Design and Implementation of Smartphone-Based AAC System," *JKIICE* (18:8).

Greenaway, K., Wright, R., Willingham, J., Reynolds, K., and Haslam, A. 2015. "Shared Identity Is Key to Effective Communication," *Personality and Social Psychology Bulletin* (41:2).

Gregor, S. 2006. "The nature of theory in information systems," *MIS Quarterly* (30:3).

Gregor, S., and Hevner, A. 2013. "Positioning and Presenting Design Science Research for Maximum Impact," *MIS Quarterly* (37:2).

Gregor, S., and Jones, D. 2007. "The Anatomy of a Design Theory," *J. Assoc. Inf. Syst* (8:5).

Gregor, S., Chandra Kruse, L., and Seidel, S. 2020. "Research Perspectives: The Anatomy of a Design Principle," *J. Assoc. Inf. Syst* (21:6).

Grice, H. 1975. "Logic and conversation," *Syntax and semantics* (3).

Hassan, N., Sumardi, N., and Aziz, R. 2019. "The Influence of Personality Traits on Communication Competence," *IJ-ARBSS* (9:13).

Heinrich, P., and Schwabe, G. 2014. "Communicating Nascent Design Theories on Innovative Information Systems through Multi-Grounded Design Principles," *DESRIST*.

Hevner, A., March, S., and Park, J. 2004. "Design Science in Information Systems Research," *MIS Q* (28:1).

Holzinger, A., Kargl, M., Kipperer, B., Regitnig, P., Plass, M., and Müller, H. 2023. "Personas for Artificial Intelligence (AI) An Open Source Toolbox," *IEEE Access* (10).

Hornero, G., Conde, D., Quilez, M., and Domingo, S. 2015. "A Wireless AAC System for People With Speech Disabilities," *IEEE Access* (3).

Jiang, H., Zhang, X., Cao, X., Kabbara, J., and Roy, D. 2023. "PersonaLLM: Investigating the Ability of GPT-3.5 to Express Personality Traits and Gender Differences," *arXiv:2305.02547*.

Kane, S. K., Morris, M. R., Paradiso, A., & Campbell, J. 2017. "Understanding Self-Expression through Augmentative and Alternative Communication Devices," CSCW.

Kocaballi, A., Berkovsky, S., Quiroz, J., and Laranjo, L. 2019. "The Personalization of Conversational Agents in Health Care: Systematic Review," *Journal of Medical Internet Research* (21:11).

Konadl, D., Wörner, J., Luttner, L., and Leist, S. 2023. "AI in AAC systems - a Literature-Based Assessment and Implications of Different Conversation Phases and Contexts," *ECIS*.

Lewar, E., Latifah, A., and Atoillah, T. 2023. "Effective Communication in Social Life," *JCEH* (6:1).

Jia, J., Zhang, P., and Zhang, R. 2015. "A comparative study of three personality assessment models in software engineering field," *ICSESS*.

Lang, R., McLay, L., and Rispoli, M. 2023. "Advanced Language Models: Potential to Improve AAC for Individuals with Intellectual and Developmental Disabilities," *Adv. Neurodev. Disord* (7:3).

Laxmidas, K., Avra, C., Wilcoxen, C., Wallace, M., and Spivey, R. 2021. "CommBo: Modernizing Augmentative and Alternative Communication," *Int. J. Hum. Comput.* (145).

Light, J., Barwise, A., Gardner, A., and Flynn, M. 2021. "Personalized Early AAC Intervention to Build Language and Literacy Skills," *Topics in Language Disorders* (41:3).

Light, J., and Mcnaughton, D. 2014. "Communicative Competence for Individuals who require AAC: A New Definition for a New Era of Communication?," *AAC* (30:1).

March, S., Smith, G. 1995. "Design and natural science research on information technology," *DSS,* (15:4).

Mamlekar, C., Dietz, A., Wendt, O. 2024. "History and Evolution of AAC," *Principles and Practices in AAC*.

Mancilla, V., Sastoque, H., and Iregui, G. 2015. "Towards ontology personalization to enrich social conversations on AAC systems," *SIPAIM*.

Mayring, P. 2004. "Qualitative Content Analysis," *A companion to qualitative research*.

Mills, T., Bunnell, T., and Patel, R. 2014. "Towards Personalized Speech Synthesis for AAC," *AAC* (30:3).
Monalisa, N. T., Himi, S., Parvin, K., and Uddin, M. 2023. "Shanketik: Revolutionizing Communication for the Deaf and Mute - A Reinforcement Learning-Enhanced Bangla AAC Application," *ICCIT*.
Moorcroft, A., Scarinci, N., and Meyer, C. 2018. "A systematic review of the barriers and facilitators to the provision and use of low-tech and unaided AAC systems for people with complex communication needs and their families," *Disability and Rehabilitation: Assistive Technology* (14:7).
Morris, M., Dudgeon, B., and Yorkston, K. 2013. "A qualitative study of adult AAC users' experiences communicating with medical providers," *Disability and Rehabilitation: Assistive Technology* (8:6).
Möller, F., Guggenberger, T., and Otto, B. 2020. "Towards a method for design principle development in information systems," *DESRIST*.
Na, J., Wilkinson, K., Karny, M., Blackstone, S., and Stifter, C. 2016. "A synthesis of relevant literature on the development of emotional competence: Implications for design of AAC systems," *AJSLP* (25:3).
Obiorah, M., and Piper, A. 2021. "Designing AAC for People with Aphasia Dining in Restaurants," *CHI*.
O'Leary, K., Delahunt, C., Dowden, P., Darmansya, I., Heng, J., Riskin, E., Ladner, R., and Wobbrock, J. 2012. "Design goals for a system for enhancing AAC with personalized video," *ASSETS*.
OpenAI. 2025. "ChatGPT-4.0," *Accessed on 15.04.2025 via this* Link.
Peffers, K., Tuunanen, T., Rothenberger, M., and Chatterjee, S. 2007. "A design science research methodology for information systems research," *JMIS* (24:3).
Purao, S., Kruse, L., and Maedche, A. 2020. "The Origins of Design Principles: Where do… they all come from?," *DESRIST*.
Pereira, J., Pereira, J., Zanchettin, C., and do Nascimento Fidalgo, R. 2024. "PrAACT: Predictive AAC with Transformers," *Expert Systems with Applications* (240).
Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., and Sutskever, I. 2019. Language models are unsupervised multitask learners," *OpenAI blog* (1:8).
Rammstedt, B., and John, O. 2007. "Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German," *Journal of Research in Personality* (41:1).
Rayman, A., Satchidanand, A., Higginbotham, J.2023. "Simulation of other-initiated repair using AAC," *AAC* (40:2).
Reichle, J., Simacek, J., Wattanawongwan, S., and Ganz, J. (2019). "Implementing Aided Augmentative Communication Systems With Persons Having Complex Communicative Needs," *BMO* (43:6).
Roosendaal, A. 2010. "Digital Personae and Profiles as Representations of Individuals," *Springer*.
Roy, A., Karim, E., Bin Farukee, M., and Makedon, F. 2022. "ChatGPT as an Assistive Technology: Enhancing Human-Computer Interaction for People with Speech Impairments," *PETRA*.
Serapio-Garcia, G., Safdari, M., Crepy, C., Fitz, S., Romero, P., Sun, L., Abdulhai, M., Faust, A., and Mataric, M. 2023. "Personality Traits in Large Language Models," *arXiv:2307.00184*.
Shao, T., and Shin, D. 2022. "Structured Pruning for Deep Convolutional Neural Networks via Adaptive Sparsity Regularization," *COMPSAC*.
Shen, J., Yang, B., Dudley, J., and Kristensson, P. 2022. "KWickChat: A Multi-Turn Dialogue System for AAC Using Context-Aware Sentence Generation by Bag-of-Keywords," *IUI*.
Sims, C. 2017. "Do the Big-Five Personality Traits Predict Empathic Listening and Assertive Communication?," *International Journal of Listening* (31:3).
Snyder, H. 2019. "Literature review as a research methodology: An overview and guidelines," *JBR* (104).
Sorokovikova, A., Rezagholi, S., Fedorova, N., and Yamshchikov, I. 2024. "LLMs Simulate Big5 Personality Traits: Further Evidence," *Proceedings of 1st Workshop on Personalization of Generative AI Systems*.
Sweller, J. 1988. "Cognitive load during problem solving: Effects on learning," *Cognitive science* (12:2).
Tett, R., and Burnett, D. 2003. "A personality trait-based interactionist model of job performance," *J. Appl. Psychol.* (88:3).
Vom Brocke, J., Simons, A., Riemer, K., and et al. 2015. "Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research," *CAIS* (37:1).
Venable, J., Pries-Heje, J., and Baskerville, R. 2016. "FEDS: a framework for evaluation in design science research," *EJIS* (25:1).
Walls, J., Widmeyer, G., and El Sawy, O. 1992. "Building an Information System Design Theory for Vigilant EIS," *Information systems research* (3:1).
Wang, M., Muthu, B., and Sivaparthipan, C. 2022. "Smart assistance to dyslexia students using artificial intelligence based AAC," *Int. J. Speech Technol.* (25:2).
Zeng, Y, Tang, Q., Chen, S., and Chen, X. 2024. "Integration of a LLM with AAC tool for oncological aphasia rehabilitation," *Asia-Pacific Journal of Oncology Nursing* (11:2).

2.8 Designing a Conversational Agent for User- Centric Gut Health Support Through Generative AI-Based Digital Personas

<table>
<tr><td>Adressierte Zielsetzungen</td><td>Konzeption eines intelligenten, nutzerzentrierten Kommunikationsagenten zur Unterstützung der Darmgesundheit. Ziel ist die Entwicklung eines KI-basierten Systemdesigns, das auf einer Digital Persona basiert, individuelle Gesundheitsinformationen kontextsensitiv verarbeitet und darauf aufbauend nutzerspezifische Empfehlungen zur Gesundheitsförderung formuliert.</td><td>ZS12</td></tr>
<tr><td>Abgeleitete Forschungsfragen</td><td>Welche funktionalen Anforderungen sollte ein digitaler Konversationsagent erfüllen, der mithilfe einer Digital Persona nutzerzentrierte Empfehlungen zur Darmgesundheit auf Basis generativer KI bereitstellt, und wie kann der Agent sowie die Persona entsprechend dieser Anforderungen gestaltet werden?</td><td>FF13</td></tr>
<tr><td>Genutzte Forschungsmethode</td><td colspan="2">Die Studie folgt dem DSR-Prozess nach Peffers et al. (2007) und kombiniert verschiedene Methoden zur Ableitung und Konsolidierung von Design Anforderungen zur Entwicklung des gesundheitsbezogenen Systemdesigns des Konversationsagenten.

Die Forschung orientiert sich an den drei DSR-Zyklen nach Hevner et al. (2004):
• Design Cycle: Entwicklung eines initialen Systemdesigns, bestehend aus Designanforderungen, Designprinzipien, Implementationsprinzipien und einem technischen Blueprint zur Realisierung des digitalen Agenten.
• Relevance Cycle: Ableitung praxisrelevanter Anforderungen durch eine systematische Literaturrecherche, eine explorative Marktanalyse und die Auswertung von User Stories (in Zusammenarbeit mit medizinischen Experten und Betroffenen).
• Rigor Cycle: Theoretische Fundierung des Artefakts durch die Berücksichtigung der theoretischen Konstrukte der Digital Persona, LLM-basierter Personalisierung und gesundheitsbezogener Chatbots.

Zur Ableitung der Anforderungen kamen qualitative Inhaltsanalysen (Mayring, 2004) mit mehrfacher Forscherkodierung und Interraterreliabilität-Überprüfung (Fleiss' Kappa: 0,87 bis 0,89) zum Einsatz.</td></tr>
<tr><td>Kernergebnisse (Überblick)</td><td colspan="2">Durch die Konzeption des Systemdesigns wurden mehrere Kernergebnisse erzielt:
• Ableitung von 11 spezifischen Designanforderungen an einen KI-gestützten Konversationsagenten zur Unterstützung der Darmgesundheit.
• Entwicklung eines integrierten Systemdesigns mit Meta- und Designanforderungen, Design- und Implementierungsprinzipien sowie Randbedingungen, welches die zentralen Anforderungen an die personalisierte Gesundheitsförderung innerhalb des Systemdesigns zusammenführen.
• Entwicklung eines Blueprints für die technische Realisierung des konzipierten Systemdesigns.
• Beitrag zur Design Theory durch die Entwicklung des Systemdesigns im Bereich digitaler Gesundheitssysteme, insbesondere zur theoriebasierten Integration von LLMs und Digital Personas in gesundheitsbezogene Kommunikationsagenten (Gregor, 2006).</td></tr>
</table>

Publikationsort	Das Paper wurde 2025 auf der International Conference on Information Systems (ICIS), Nashville, Tennessee (USA) zur Begutachtung eingereicht.
Ranking VHB JQ 4	**A**
Autoren und Anteile	Daniel Konadl 40% Florian Johannsen 20% Susanne Leist 15% **Janik Wörner 15%** Isabel Reuter 10%
DOI	-

Tabelle 22. Fact Sheet zur 8. Veröffentlichung.

Designing a Conversational Agent for User-Centric Gut Health Support Through Generative AI-Based Digital Personas

Short Paper

Introduction

In human medicine, the gut is frequently referred to as the "second brain" (Avetisyan et al., 2015; Gershon, 1999). This is because it has its own nervous system, consisting of around one hundred million nerve cells (Avetisyan et al., 2015; Schneider et al., 2019). By means of the so-called "enteric nervous system", the gut can regulate many digestive processes independently of the human brain (Schneider et al., 2019). The composition of the intestinal flora has extensive effects on the physiological and psychological condition of a human being (Zhou et al., 2020). It influences various gut-related symptoms and diseases such as abdominal cramps, diverticulitis, diarrhea, and gut cancer. Certain intestinal bacteria can promote inflammatory processes that lead to severe complications like vasoconstriction, cardiovascular concerns, and fat accumulation in the liver (Tang et al., 2017). Moreover, via the "brain-gut axis", gut concerns can even influence emotions, stress, and trigger severe mental illnesses such as depression (Labanski et al., 2020).

The far-reaching consequences of poor gut health, exacerbated by common harmful habits (e.g., poor nutrition, physical inactivity) (Inczefi et al., 2022) underline the urgent need for prevention and treatment of intestinal issues (Bhattarai & Janaswamy, 2022). However, access to medical experts in the context of gut health is limited due to the low doctor-patient and advisor-patient ratios (Houwen et al., 2019). Mobile apps can provide immediate support for gut health concerns. Yet, these apps lack the sense of human-like interactions that patients value in healthcare (Alqahtani et al., 2021). To bridge this gap, conversational agents dedicated to gut health have emerged. However, these conversational agents fall short of essential functionalities, particularly in providing personalized and user-centric recommendations – a prominent drawback existing in extant research on chatbots and conversational agents (e.g., Breiter et al., 2024; Göldi & Rietsche, 2024; Liu et al., 2022). As a consequence, the delivered recommendations are not necessarily tailored towards the specific health parameters of an individual, which can severely impede their meaningfulness and effectiveness in supporting gut health (cf. Samaan et al., 2025; Yang et al., 2024).

We address these drawbacks by proposing in a first step the design of a comprehensive conversational agent that can deliver user-centric recommendations for gut health. In this regard, we derived several Design Requirements (DRs) through a Systematic Literature Review (SLR) (Vom Brocke et al., 2015), a market study analyzing gut health mobile apps and conversational agents, and three user stories representing individuals affected by gut health problems (Schilling, 2016). To enable personalized recommendations, we incorporate digital personas (Clarke, 1994; Roosendaal, 2009) into the conversational agent and purposefully tailor them to the context of gut health. These gut health digital personas reflect the users' unique gut health profiles, develop tailored prompts, and leverage a Generative AI technique like ChatGPT-4o to generate the responses for the gut health agent. In this short paper, we answer the following research question:

> *Which DRs should a conversational agent providing user-centric gut health recommendations through a tailored digital persona meet, and how to design the agent and the persona based on these DRs?*

This research-in-progress unfolds as follows: The next section provides conceptual foundations and related work. Then, the research procedure following the Design Science Research (DSR) approach (Hevner et al., 2004; Peffers et al., 2007; Tuunanen et al., 2024) is described. The next section deals with the design of the gut health conversational agent and the gut health digital persona. Finally, we elaborate the implications of this research, the (expected) theoretical and practical contributions, and the next steps in our DSR-project.

Conceptual Foundations and Related Work

Generally, a conversational agent is a computer program that can hold a conversation with human beings using natural language (Abdul-Kader & Woods, 2015). The newly gained interest for text-based conversational agents these days is often explained by the spread of mobile applications and virtual assistants, which all influence people's way of communicating and connecting with one another (Dale, 2016).

As revealed by the analysis of existing literature conducted in this paper, there have also emerged conversational agents that address the specific domain of gut health (e.g., Hunt et al., 2021; Morato et al., 2023; Yang et al., 2024; Zhou et al., 2024). Most of them fall into the categories of task-oriented or health-coaching conversational agents, primarily focusing on prevention, education, and symptom monitoring (e.g., Hunt et al., 2021; Narang et al., 2024). However, a significant limitation is their lack providing user-centric recommendations, which not necessarily align with an individual's unique gut health condition. Moreover, they address only isolated symptoms like inflammatory bowel diseases or generalized dietary advice instead of proposing holistic gut health support (e.g., Samaan et al., 2025; Sciberras et al., 2024; Zand et al., 2020). Our market study analyzing conversational agents targeting gut health support has confirmed these drawbacks. These agents available in the market are rather rudimentary and rule-based, lacking adaptive and context-sensitive dialogue capabilities. Their functionalities are limited to single use cases, such as dietary suggestions, lifestyle tips, or probiotic guidance. As a result, there is a lack of a comprehensive gut health conversational agent capable of providing personalized recommendations for gut health issues.

In the context of personalization, digital personas offer a highly promising approach. These virtual representations can capture various aspects of a user's individual conditions and behaviors (Clarke, 1994; Roosendaal, 2009). Within a conversational agent focused on gut health, a digital persona can model various gut health-related dimensions (e.g., symptoms, dietary habits, physical activities, stress level) and track their interrelationships over time. Based on this information, the digital persona develops tailored prompts to leverage a Generative AI technique like ChatGPT-4o, analyzing users' unique health profiles and their specific intents (e.g., seeking gut-healthy dietary). In this way, these digital personas empower the gut health agent to generate effective and user-centric recommendations. However, neither academic research nor industry efforts have sufficiently addressed the need for such a comprehensive gut health agent. To foster holistic support in the prevention and treatment of intestinal issues, our DSR-project aims to design the comprehensive gut health conversational agent featuring a tailored gut health digital persona.

Research Procedure

This research-in-progress follows the DSR paradigm (cf. Hevner et al., 2004; Tuunanen et al., 2024) and is structured according to the steps of Peffers et al. (2007). We are currently in the first cycle of this iterative process. Therefore, the design of our solution will be revised in future cycles based on the insights and feedback gained. Figure 1 provides an overview of the procedure of our holistic DSR-project and highlights the steps dealt with in this short paper.

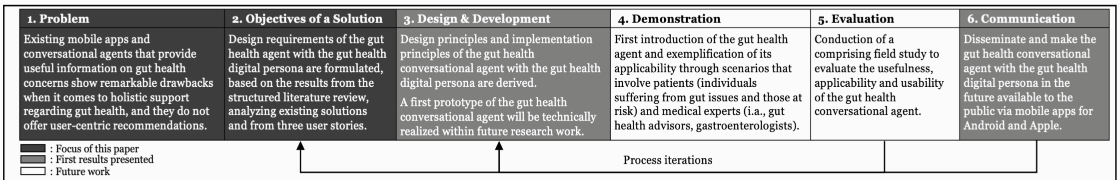


Figure 1. Procedure of this Research (own Representation, based on Peffers et al., 2007)

First, the problem statement was formulated (**Step 1 – "Problem"**; *see Section "Introduction"*). Next, we gathered DRs through a SLR and a market study analyzing existing gut health mobile apps and conversational agents. In addition, user stories were developed in collaboration with individuals affected by gut health issues and medical experts (**Step 2 – "Objectives of a Solution"**; *see Section "Design of the Gut Health Conversational Agent"*). The results were then consolidated and discussed among the five involved researchers to reach a consensus on the design of the gut health agent and the digital persona. In this way, a manageable set of DRs was derived for an initial prototype. Following this, Design Principles (DPs) and Implementation Principles (IPs) were deduced from the DRs (**Step 3 – "Design & Development"**; *see Section "Preliminary Design for the Gut Health Agent"*).

In the next step of our DSR-project, we will proceed with the technical realization of the gut health conversational agent and the gut health digital persona. Subsequently, the resulting prototype will undergo a series of demonstration and evaluation steps (**Step 4 – "Demonstration"** and **Step 5 – "Evaluation"**). This procedure allows to further revise and improve the prototype and several intermediate artifacts (e.g., DRs, DPs, IPs) in further DSR-cycles, before conducting more elaborate evaluations in different natural settings (cf. Venable et al., 2016). Finally, (**Step 6 – "Communication"**), we document the results of this research, while future insights will be communicated in subsequent research papers.

Design of the Gut Health Conversational Agent

Design Requirements of the Gut Health Conversational Agent

To gather DRs for the gut health conversational agent, we followed a three-step approach: **(I)** We defined three user stories representing individuals who have suffered from or have been at risk of developing gut health concerns. A first user story was developed in collaboration with a person suffering from gut problems (diverticulitis) who would see himself as a potential user of the gut health agent. A second (irritable bowel syndrome) was developed in collaboration with a doctor at a university hospital. A third was created collaboratively by a gut health advisor, a doctor specialized in nutrition medicine, and an individual experiencing frequent gut concerns (bloating and diarrhea). These user stories focus on providing knowledge to prevent gut concerns and offering solutions for providing immediate treatment of acute symptoms. In doing so, we adhered to the recommendations of Schilling (2016) about the pre-conception of mobile apps and requirements engineering (Rupp et al., 2009). Beyond the user stories, we conducted **(II)** a structured review and synthesis of extant literature, and **(III)** carried out a market study analyzing existing mobile apps and conversational agents focused on gut health.

The SLR **(II)** was aligned to the guidelines as proposed by Vom Brocke et al. (2015). In a **first step**, a keyword search was performed on ACM Digital Library, AIS Electronic Library, EBSCOhost, Google Scholar, IEEE Xplore, SAGE Publications, ScienceDirect, SpringerLink, and the Wiley Online Library. According to Snyder (2019), the search strings for conducting a literature review should refer to words and concepts directly linked to the research question. Thus, "conversational agent", "gut", and "bowel" were derived as initial search terms. To obtain even further search terms, iterative keyword refinement was performed (e.g., "intestinal concern", "chatbot", "gastroenterology", various gut symptoms like "irritable bowel syndrome"). As a result, 758 papers were initially derived. We only kept papers that focus on conversational agents providing support and recommendations regarding gut health issues to individuals by means of human-like interactions. According to that, four of the involved researchers independently analyzed the title, abstract and thematic focus of every paper. Several decisions about including or excluding the respective papers were cross-validated by all four researchers. To assess the coding, interrater reliability was calculated using Fleiss' Kappa. The rating procedure yielded a Fleiss' Kappa value of 0.89, indicating a high level of agreement between the raters (Fleiss et al., 2013). In the very few cases of disagreement, the coding was discussed to come to a consensus about including or excluding the respective paper. After applying this criterion, removing duplicates and non-English articles, 14 relevant papers remained. In the **second step**, forward and backward searches were performed according to Webster & Watson (2002), yielding two additional relevant papers. As follows, a total of 16 relevant papers could be derived from the SLR.

As part of the market study **(III)**, four of the involved researchers independently identified existing mobile apps and conversational agents dedicated to gut health support. Thereby, the Google search engine as well as the search engines of the Google Play Store and the Apple App Store were drawn upon, applying the same search terms and combinations as in the SLR. The results were cross-validated by all four researchers to ensure that no essential tool was overlooked. In this way, a total of 37 gut health apps and five conversational agents could be derived.

Subsequently, we investigated the contents derived from **(I)**, **(II)**, and **(III)**. Through qualitative content analysis and inductive category development, we could carry out a consolidation from these data to derive a coherent image. To inductively extract the DRs from the consolidated materials, qualitative content analysis was applied as well (cf. Mayring, 2000). In this process, a DR was considered any specific condition, capability, or functionality that the intended artifact needs to implement in order to achieve its intended goals and purposes (cf. Möller et al., 2020; Walls et al., 1992). Any text statement was screened against this criterion pointing out a DR. Moreover, the mobile apps and conversational agents were tested out to achieve further functionalities for the design of the gut health agent with the gut health digital persona. Then, the resulting DRs were cross-validated by all the five involved researchers. Thereby, the gained Fleiss' Kappa value of 0.87 indicates a high level of interrater agreement (Fleiss et al., 2013). In case of any disagreement, the respective DR was analyzed again, followed by discussions until reaching a consensus. Based on this procedure, the following **eleven DRs** could be derived:

Design Requirements	Descriptions The Gut Health Agent needs to ...	Sources
DR1. Tracking of Health Parameters	... record and track relevant health parameters (e.g., age, sleep, stress level, nutrition, physical activities, symptoms).	Narang et al., 2024 User Stories + Market Study

DR2. Personalized Gut-Related Recommendations	... analyze the provided health parameters to deliver personalized interactions and tailored recommendations regarding gut health.	Huo et al., 2024; Kashyap et al., 2023; Samaan et al., 2025; Yang et al., 2024 Market Study
DR3. Personalized Nutrition and Recipe Suggestions	... provide tailored recipe suggestions based on the user's dietary habits, intolerances, and gut health goals.	User Stories + Market Study
DR4. Adaptive Real-Time Responses	... dynamically adjust to real-time changes in user behavior and gut health data, ensuring context-sensitive and continuously relevant interactions.	Yang et al., 2024; Zhou et al., 2024 User Stories
DR5. Integration of Specialized Medical Knowledge	... incorporate medical data sources (rather than solely relying on large language models) for ensuring accurate, context-specific and medically-proven responses. Thereby, expert input should enhance the plausibility of the conversational agent's recommendations and responses.	Giuffrè et al., 2024; Gravina et al., 2024; Li, 2023; Pellegrino et al., 2024; Samaan et al., 2025; Yang et al., 2024; Zheng et al., 2023; Zhou et al., 2024 User Stories
DR6. Provision of Information about the Gut and Gut-Healthy Behaviors	... provide medically-proven background information on the gut, the digestive system, and habits as well as behaviors that support gut health.	Hunt et al., 2021; Zand et al., 2020 User Stories + Market Study
DR7. Provision of Connections to Experts	... enable users to connect with medical experts, nutritionists, or like-minded individuals to facilitate knowledge exchange, support, and treatment in case of severe gut issues.	User Stories + Market Study
DR8. Preparation and Visual Representation of Health Data	... generate highly understandable and appealing reports, including diagrams, statistics, and summaries. These representations should illustrate the development of gut health and overall well-being in an accessible and structured format.	Narang et al., 2024 User Stories + Market Study
DR9. Clear and User-Friendly Language	... use a language style that is simple, understandable, and accessible for the users. Thereby, the agent should also cite its sources to ensure that its statements are verifiable.	Gravina et al., 2024; Morato et al., 2023; Sciberras et al., 2024; Zhao et al., 2023; Zhou et al., 2024
DR10. High Usability and Intuitive Design	... offer an intuitive user interface to ensure ease of use and a smooth user experience.	Morato et al., 2023
DR11. Real-Time Data Processing	... process data quickly to provide real-time responses and seamless interactions.	Zand et al., 2020 Market Study

Table 1. Design Requirements of a Gut Health Conversational Agent

Preliminary Design for the Gut Health Agent

Similar to the process of deriving the DRs, several steps in developing the preliminary design for the gut health agent featuring the gut health digital persona were conducted and cross-validated by all five involved researchers. Beyond the DRs, **Meta Requirements (MRs)** were derived for a purposeful and comprehensible design (Gregor & Jones, 2007; Heinrich & Schwabe, 2014). MRs define the overarching purpose of the system, specifying "what the system is for" (Gregor & Jones, 2007, p. 325) and shaping the solution's objectives based on the specific problems that the research addresses. For instance, one MR for the gut health conversational agent is to uphold data ethics and ensure confidentiality within several operations performed with a user's health data. Grounded in research literature, the MRs apply to a broader class of related artifacts (Gregor & Jones, 2007; Heinrich & Schwabe, 2014; Walls et al., 1992).

The **Design Principles** are synthesized from the derived set of eleven DRs *(see Table 1)*. DRs that semantically correspond to each other indicate conceptual influences to be considered when proposing a DP. Therefore, the derived DRs were carefully examined and semantically-related DRs were consolidated into a respective DP (cf. Möller et al., 2020). Defined as generic and prescriptive statements, these DPs facilitate the transfer and application of the design-related knowledge to similar problem classes beyond the specific domain of a gut health conversational agent (Gregor et al., 2020). Based on the DRs and the DPs, we formulated preliminary **Implementation Principles**. These IPs are intended to guide and support the implementation in practice by proposing generic design methods (Gregor & Jones, 2007) in view of a gut health conversational agent that incorporates a gut health digital persona.

Boundary conditions (BCs) define the constraints and limitations (e.g., technical, legal, or user-specific) in which the artifact must operate, ensuring that it remains feasible, usable, and relevant within certain operational contexts (Chandra et al., 2015). One example of a BC for the gut health agent is its reliance on truthful user input (i.a., specifications about health parameters, symptoms, dietary habits and intended goals) to be capable of generating reasonable, user-specific, and effective recommendations for gut health support. Keeping in mind several DSR-related intermediate artifacts, there was derived the following blueprint for the design of the gut health conversational agent with the gut health digital persona:

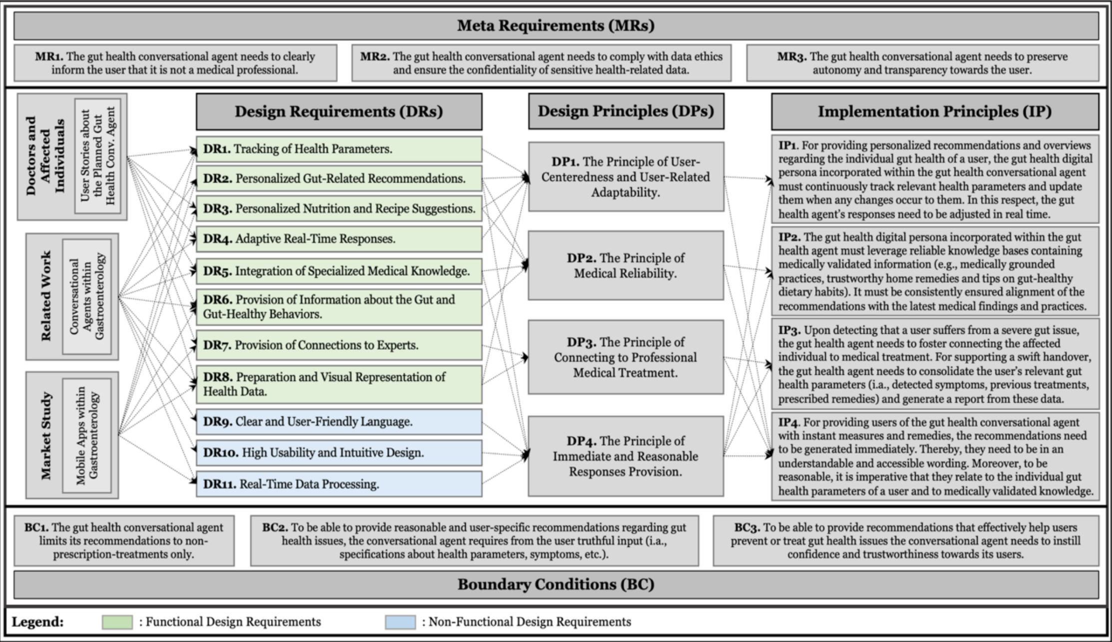


Figure 2. Derived Design of the Gut Health Agent with the Gut Health Digital Persona

- **DP1:** For providing user-centric gut health recommendations, the gut health digital persona within the gut health conversational agent needs to track and analyze the interrelationships of relevant gut health parameters over time. Any changes need to be immediately considered in the generated responses.
- **DP2:** For ensuring medical validity of the delivered gut health recommendations, the gut health digital persona needs to align the incorporated knowledge bases with the latest medical findings and practices.
- **DP3:** When detecting a potentially severe gut issue, the gut health digital persona needs to connect the affected individual with medical experts to facilitate immediate medical treatment. Thereby, it should also consolidate user-specific health data for a swift handover to the medical experts.
- **DP4:** For providing users effective remedies, the gut health agent must support immediate generation of understandable responses that consider the user-specific parameters and pertinent medical knowledge.

Blueprint for the Technical Realization of the Gut Health Conversational Agent

In the following, we outline how the gut health agent could be instantiated based on the derived design. When a conversation has been started, the agent prompts the user to enter its name and credential. It then verifies this information against the encrypted entries in the user database. Upon successful authentication, the agent loads the user's individual gut health digital persona, which is then enabled to leverage ChatGPT-4o via its API and via tailored prompts, enriched by user-specific parameters and the user's specific intents.

To be comprehensive, the gut health agent supports different gut health-related use cases (i.a., tracking and analyzing gut health parameters, nutrition counseling, challenges to improve gut health). If a user suffering from gut issues selects nutrition counseling, the agent prompts to specify the current issues and its intensity. To provide user-centric recommendations and help alleviate the issues (e.g., adjusted dietary habits based on potential intolerances like lactose or gluten), the agent assesses the symptoms and their severity using a curated medical knowledge base **(DR5)**. In case of any questions from the user, the agent **(DR6)** provides relevant background information (i.a., from the LLM) and justifies its assumptions (from the curated medical knowledge base). Simultaneously, the user-provided information is integrated into the user-specific gut health digital persona **(DR1)** to be considered for upcoming recommendations **(DR4)**. The agent also checks this digital persona whether the user has previously reported the same complaints. If a severe gut issue is detected (e.g., long-lasting diarrhea) or if previously provided recommendations **(DR2 / DR3)** have proven ineffective, the gut health agent recommends seeking advice from a medical expert

(DR7). To support the medical expert in understanding the user's background, the user can prompt the agent to prepare a report (e.g., Microsoft Excel sheets) that includes the detected gut issues, symptoms, and health history **(DR8)**. After the user visited the medical expert and re-authenticates with the gut health agent, it inquires about treatments, findings, prescribed remedies, and the healing process. This information is then integrated into the user's unique gut health digital persona to enhance future gut health recommendations.

To assess this blueprint for the technical realization, we engaged two individuals affected by gut issues and two medical experts who had contributed to defining the user stories. They expressed appreciation of the gut health digital persona to deliver user-centric gut health recommendations, as they emphasized gut concerns are inherently individual issues. They also valued the holistic approach to gut health support, ranging from identifying and assessing the gut issues to, when necessary, referring to professional medical care.

Implications and Discussion

The central role the human gut plays in both, the physical and mental health, compounded by the far-reaching consequences of poor gut health and common harmful habits (Inczefi et al., 2022), leads to an increasing demand in the prevention and treatment of intestinal issues (Bhattarai & Janaswamy, 2022). However, access to medical experts in the context of gut health is limited due to the low doctor-patient and advisor-patient ratios. Therefore, we generated foundational design-related knowledge for a comprehensive gut health conversational agent in a first iteration of the DSR-cycle. By means of tailored gut health digital personas, the gut health agent can provide user-centric recommendations that support individualized gut health needs. It is important to emphasize that, while this agent can manage gut issues with a low risk for an individual's health, severe gut problems must always be subject to the investigation by a licensed medical expert. By conducting this DSR-project, we aim to generate desirable results to both theory and practice.

First, we contribute to the ongoing discussions in IS research on the personalization of recommendations delivered by chatbots and conversational agents (e.g., Breiter et al., 2024; Göldi & Rietsche, 2024; Liu et al., 2022). The technical realization of the gut health digital persona will reveal how to incorporate and map different user-specific parameters for meaningful and user-centric recommendations. In this way, our work will make valuable contributions to related research areas like the personalization of chatbot responses in service encounters. **Second**, we contribute to "Nascent Design Theory" by designing the artifact in a purposeful and comprehensible way (Gregor & Hevner, 2013). For example, DP1 emphasizes that by accounting for an individual's unique gut health conditions (e.g., symptoms, habits, behaviors), this digital persona delivers recommendations that are user-centric and contextually appropriate. In this regard, we ensure the DPs follow the value grounding (reference to the DRs) and explanatory grounding (DPs are based on extant literature, established mobile apps and conversational agents in the market, user stories created by experienced medical professionals and affected individuals (Heinrich & Schwabe, 2014)). **Third**, further iterations within the DSR-cycle are expected to yield valuable implications for pertinent kernel theories. As one key theory in the context of gut health, the brain-gut axis theorem (e.g., Labanski et al., 2020) explains the bidirectional communication between the central and the enteric nervous system. This axis links emotional and cognitive centers of the brain with peripheral intestinal functions. The gut health agent can serve as a technological extension of the brain-gut axis by influencing how gut health information is processed. Our DSR-project can expand this theorem by investigating how different recommendation delivery methods (e.g., empathetic tone, solution-oriented communication) influence the user's perceptions and intestinal responses. An optimized recommendation style for the gut health agent based on the concrete transmission mechanisms of the brain-gut axis could alleviate gut issues rather than exacerbating them.

Outlook and Next Steps

Our next step will be the technical realization of the gut health agent, the gut health digital persona, and their seamless integration. Leveraging ChatGPT-4o may not be purposeful without extensive adaptations, which we will all realize in this phase. The derived MRs, DRs, DPs, and IPs will guide several design decisions, such as structuring the conversational flow and selecting a suitable chatbot framework (e.g., Rasa Open Source). Afterward, the prototype will undergo a series of demonstrations and assessments based on common everyday scenarios (e.g., nutrition planning, assessing food compatibility when visiting a restaurant). To prepare this step, we are currently building up a database by recruiting gut health experts and a broad group of affected individuals who are willing to contribute to our gut health agent.

After successfully passing the "Technical Risk & Efficacy" evaluation (Venable et al., 2016), and confirming or adapting the DRs, DPs, and IPs, the prototype will undergo a field study following the "Human Risk & Effectiveness" evaluation strategy (Venable et al., 2016). This will involve a broad sample of medical experts and patients with diverse gut health concerns. In this way, we can intensify the assessment of the gut health agent's and the gut health digital persona's usefulness, applicability, and usability, amongst others. These insights facilitate to further revise and improve the artifact and will support contributions to extant kernel theories (e.g., brain-gut axis theorem) and to design theory. Finally, we will make the artifact available to the public as a comprehensive gut health companion via mobile apps for Android and Apple devices.

References

Abdul-Kader, S., & Woods, J. (2015). Survey on chatbot design techniques in speech conversation systems. *International Journal of Advanced Computer Science and Applications, 6*(7).

Alqahtani, F., Winn, A., & Orji, R. (2021). Co-designing a mobile app to improve mental health and well-being: focus group study. *JMIR - Formative Research, 5*(2).

Avetisyan, M., Schill, E., & Heuckeroth, R. (2015). Building a second brain in the bowel. *The Journal of Clinical Investigation, 125*(3), 899-907.

Bhattarai, S., & Janaswamy, S. (2022). The nexus of gut microbiota, diet, and health. *FFS Online*, *2*(2).

Breiter, K., Schnaak, F., & Gimpel, H. (2024). Designing Conversational Agents to Promote Environmentally Friendly User Behavior. *Proceedings of ICIS 2024.*

Chandra, L., Seidel, S., & Gregor, S. (2015). Prescriptive Knowledge in IS Research: Conceptualizing Design Principles in Terms of Materiality, Action, and Boundary Conditions. *Proceedings of HICSS 2015.*

Clarke, R. (1994). The Digital Persona and Its Application to Data Surveillance. *Information Society, 10*(2).

Dale, R. (2016). The return of the chatbots. *Natural Language Engineering, 22*, 811-817.

Fleiss, J., Levin, B., & Paik, M. (2013). *Statistical methods for rates and proportions*. John Wiley & Sons.

Gershon, M. (1999). The enteric nervous system: a second brain. *Hospital Practice, 34*(7), 31-52.

Giuffrè, M., Kresevic, S., Pugliese, N., You, K., & Shung, D. (2024). Optimizing large language models in digestive disease: strategies and challenges to improve clinical outcomes. *Liver International, 44*(9).

Göldi, A., & Rietsche, R. (2024). Making Sense of Large Language Model-Based AI Agents. *Proceedings of ICIS 2024.*

Gravina, A., Pellegrino, R., Cipullo, M., Palladino, G., & Imperio, G. (2024). May ChatGPT be a tool producing medical information for common inflammatory bowel disease patients' questions? An evidence-controlled analysis. *World Journal of Gastroenterology, 30*, 17-33.

Gregor, S., & Jones, D. (2007). The Anatomy of a Design Theory. *Journal of the AIS, 8*(5), 312-335.

Gregor, S., Chandra, L., & Seidel, S. (2020). Research Perspectives: The Anatomy of a Design Principle. *Journal of the AIS, 21*(6), 1622-1652.

Gregor, S., & Hevner, A. (2013). Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly, 37*(2), 337-355.

Heinrich, P., & Schwabe, G. (2014). Communicating Nascent Design Theories on Innovative Information Systems through Multi-Grounded Design Principles. *Proceedings of DESRIST 2014.*

Hevner, A., March, S., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly, 28*(1), 75-105.

Houwen, J., Lucassen, P., Verwiel, A., Stappers, H., & Assendelft, S. (2019). Which difficulties do GPs experience in consultations with patients with unexplained symptoms: a qualitative study. *BMC Family Practice, 20*, 1-9.

Hunt, M., Miguez, S., Dukas, B., Onwude, O., & White, S. (2021). Efficacy of Zemedy, a Mobile Digital Therapeutic for the Self-management of Irritable Bowel Syndrome: Crossover Randomized Controlled Trial. *JMIR mHealth and uHealth, 9.*

Huo, B., Calabrese, E., Sylla, P., Kumar, S., & Ignacio, R. (2024). The performance of artificial intelligence large language model-linked chatbots in surgical decision-making for gastroesophageal reflux disease. *Surgical Endoscopy, 38*(5).

Inczefi, O., Bacsur, P., Resal, T., Keresztes, C., & Molnar, T. (2022). The influence of nutrition on intestinal permeability and the microbiome in health and disease. *Frontiers in Nutrition, 9.*

Kashyap, R., Roy, V., Patil, P., Manhar, A., & Roy, L. (2023). Deep Learning's Role in Advancing Gastroenterology and Digestive Health. *Proceedings of the ICTBIG.*

Labanski, A., Langhorst, J., Engler, H., & Elsenbruch, S. (2020). Stress and the brain-gut axis in functional and chronic-inflammatory gastrointestinal diseases. *Psychoneuroendocrinology, 111*.

Li, L. (2023). Role of chatbots on gastroenterology: Let's chat about the future. *Gastrointest. Endosc., 1*(3).

Liu, Y., Yan, W., Hu, B., Li, Z., & Lai, Y. (2022). Effects of personalization and source expertise on users' health beliefs and usage intention toward health chatbots. *Digital Health, 8*.

Mayring, P. (2000). Qualitative Content Analysis. *Forum Qualitative Social Research, 1*(2).

Möller, F., Guggenberger, T., & Otto, B. (2020). Towards a method for design principle development in information systems. *Proceedings of DESRIST 2020*.

Morato, J., do Nascimento, J., Roque, G., de Souza, R., & Santos, I. (2023). Development, Validation, and Usability of the Chatbot ESTOMABOT to Promote Self-care of People With Intestinal Ostomy. *Computers, Informatics, Nursing, 41*(12), 1037-1045.

Narang, G., Chen, Y., Wedel, N., Wu, M., & Luo, M. (2024). Development of a Digital Patient Assistant for the Management of Cyclic Vomiting Syndrome: Patient-Centric Design Study. *JMIR, 8*.

Peffers, K., Tuunanen, T., Rothenberger, M., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*(3), 45-77.

Pellegrino, R., Federico, A., & Gravina, A. (2024). Conversational LLM Chatbot ChatGPT-4 for Colonoscopy Boston Bowel Preparation Scoring: An Artificial Intelligence-to-Head Concordance Analysis. *Diagnostics, 14*(22).

Roosendaal, A. (2009). Digital Personae and Profiles as Representations of Individuals. *Proceedings of the IFIP PrimeLife International Summer School on Privacy and Identity Management for Life*.

Rupp, C., Simon, M., & Hocker, F. (2009). Requirements Engineering und Management. *HMD Praxis der Wirtschaftsinformatik, 46*(3), 94-103.

Samaan, J., Issokson, K., Feldman, E., Fasulo, C., & Rajeev, N. (2025). Examining the Accuracy and Reproducibility of Responses to Nutrition Questions Related to Inflammatory Bowel Disease by Generative Pre-trained Transformer-4. *Crohn's & Colitis*, *360*(7).

Schilling, K. (2016). *Apps machen - Der Kompaktkurs für Designer*. Hanser.

Schneider, S., Wright, C., & Heuckeroth, R. (2019). Unexpected roles for the second brain: enteric nervous system as master regulator of bowel function. *Annual Review of Physiology, 81*(1), 235-259.

Sciberras, M., Farrugia, Y., Gordon, H., Furfaro, F., & Allocca, M. (2024). Accuracy of Information given by ChatGPT for patients with Inflammatory Bowel Disease in relation to ECCO Guidelines. *Journal of Crohns & Colitis, 18*(8), 1215-1221.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333-339.

Tang, W., Kitai, T., Hazen, S. (2017). Gut microbiota in cardiovascular health and disease. *Circ. Res.*, *120*(7).

Tuunanen, T., Winter, R., & Vom Brocke, J. (2024). Dealing with complexity in design science research: A methodology using design echelons. *MIS Quarterly, 48*(2), 427-458.

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: a framework for evaluation in design science research. *European Journal of Information Systems, 25*(1), 77-89.

Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the AIS*, *37*(1), 205-224.

Walls, J., Widmeyer, G., & El Sawy, O. (1992). Building an Information System Design Theory for Vigilant EIS. *Information Systems Research*, *3*(1), 36-59.

Webster, J., & Watson, R. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, *26*(2), xiii-xxiii.

Yang, Z., Khatibi, E., Nagesh, N., & Azimi, I. (2024). ChatDiet: Empowering Personalized Nutrition-Oriented Food Recommender Chatbots through an LLM-Augmented Framework. *Smart Health, 32*.

Zand, A., Sharma, A., Stokes, Z., Reynolds, C., & Montilla, A. (2020). An Exploration Into the Use of a Chatbot for Patients With Inflammatory Bowel Diseases: Retrospective Cohort Study. *JMIR*, *22*(5).

Zhao, Y., Gao, C., Zhang, L., Gao, X., & Zhang, Z. (2023). The application of chatbot in gastroenterology nursing. *Gastroenterology & Endoscopy*, *1*(3), 136-138.

Zheng, D., Ren, G., Min, J., Yew, D., & Tan, C. (2023). ChatGPT for advice on common gastrointestinal endoscopic procedures - the promise and the peril. *iGIE, 2*(4), 547-553.

Zhou, B., Yuan, Y., Zhang, S., Guo, C., & Li, X. (2020). Intestinal flora and disease mutually shape the regional immune system in the intestinal tract. *Frontiers in Immunology, 11*.

Zhou, Q., Liu, C., Duan, Y., Sun, K., & Li, Y. (2024). GastroBot: a Chinese gastrointestinal disease chatbot based on the retrieval-augmented generation. *Frontiers in Medicine, 11*.

3 Schlussbetrachtung und Fazit

Die vorliegende Dissertation behandelt zwei thematisch komplementäre Forschungsfelder, die sich mit der Konzeption und Umsetzung intelligenter Analysemethoden und Werkzeuge zur Erschließung der in digitalen Daten liegenden Erkenntnispotenziale befassen (vgl. *Kapitel 1.1.1* und *1.1.2*). Beide Forschungsstränge eint dabei das Ziel, aktuellen Herausforderungen im Umgang mit nutzergenerierten Daten durch den gezielten Einsatz von Methoden der KI sowie des Data Science zu begegnen. Aufbauend auf insgesamt acht wissenschaftlichen Veröffentlichungen wurden in beiden Forschungsgebieten zentrale Problemstellungen identifiziert, theoretisch fundiert analysiert und durch methodisch reflektierte sowie empirisch validierte Beiträge adressiert (vgl. *Kapitel 2*).

Das abschließende Kapitel dieser Dissertation fasst im ersten Schritt die zentralen Forschungsergebnisse entlang der beiden adressierten Forschungsgebiete zusammen (vgl. *Kapitel 3.1*). Daran anschließend werden die geleisteten Wertbeiträge, die im Rahmen der Dissertation erzielt wurden, anhand der drei Perspektiven *technischer*, *theoretischer* und *praktischer* Wertbeiträge systematisiert (vgl. *Kapitel 3.2*). Abschließend werden bestehende Limitationen der Dissertation reflektiert und darauf aufbauend gezielte Forschungsrichtungen abgeleitet, die konkrete Ansatzpunkte für deren systematische Bearbeitung in zukünftigen Arbeiten bieten (vgl. *Kapitel 3.3*). Ziel des letzten Kapitels ist es, die Ergebnisse der Dissertation systematisch aufzuarbeiten, in ihrer Gesamtheit sichtbar zu machen und darauf aufbauend die Anschlussfähigkeit der Dissertation sicherzustellen.

3.1 Zusammenfassung der Forschungsergebnisse

1. Forschungsgebiet: Zusammenfassung der Forschungsergebnisse im Rahmen der Konzeption und Entwicklung kontextsensitiver Analyseverfahren**.**

1. Veröffentlichung. Unternehmen – und im speziellen Marketingabteilungen – können umfassend von den in UGC enthaltenen Erkenntnissen profitieren; insbesondere von den ungefilterten Meinungen, Gedanken und Bedürfnissen ihrer (potenziellen) Kunden, wie sie sich in Social Media Posts wiederfinden (VoC). Die Analyse dieser großen Mengen an textbasierten Daten erfordert dabei automatisierte Analyseverfahren: Topic Modelling hat sich hierbei als etabliertes Verfahren zur Extraktion kohärenter Themen und thematischer Strukturen bewährt und durch viele unterschiedliche Verfahrenstypen ausgezeichnet.

Im Rahmen der ersten Veröffentlichung wurde auf Basis der GQM-Methode (Basili et al., 1994) eine vergleichende Analyse von drei unterschiedlichen Topic Modelling Verfahren

(*LDA*, *DMR*, *PAM*) durchgeführt. Der Fokus lag dabei auf deren Eignung für die drei zentralen Anwendungsszenarien im Marketing: *Themenextraktion*, *Trendanalyse* und *Themenstrukturierung*. Die Evaluation von LDA, DMR und PAM erfolgte entlang der fünf abgeleiteten Vergleichskriterien – die drei quantifizierbaren Metriken *Perplexity*, *Coherence* und *Log-Likelihood* sowie die beiden qualitativen Metriken *Word Intrusion* und *Topic Intrusion* – und diente der Beantwortung von **FF1** und **FF2**.

Die Ergebnisse zeigen, dass der De-facto-Standard *LDA* nicht pauschal die beste Methode für alle Anwendungsfälle darstellt. Stattdessen erweisen sich semi-supervisierte Topic Modelling Verfahren wie *DMR* für die Analyse von Trends und hierarchische Topic Modelling Verfahren wie *PAM* für die strukturierende Analyse der diskutierten Themen als besonders geeignet. Damit liefert die Veröffentlichung konkrete, anwendungsbezogene Handlungsempfehlungen für die fundierte Auswahl geeigneter Topic Modelling Verfahren in der Marketingpraxis – in Abhängigkeit vom jeweiligen Nutzungskontext und Ziel der Analyse.

2. Veröffentlichung. Während sich klassische Analyseverfahren zur Analyse von UGC in sozialen Medien primär auf die inhaltliche Analyse fokussieren (z. B. zur Identifikation thematischer Trends), bleiben die kontextuellen Dimensionen nutzergenerierter Kommunikation oftmals unberücksichtigt; obwohl digitale Plattformen Metadaten zu Zeit und Ort bereitstellen. Dabei hat sich in der Forschung gezeigt, dass Konsumentenwahrnehmungen erheblich in Abhängigkeit solcher Kontextfaktoren variieren. Eine kontextsensitive Analyse sozialer Medien bietet so bspw. Potenziale für die Segmentierung von Kundengruppen zur differenzierten Marktansprache oder der zeitlichen Einordnung von Stimmungsdynamiken.

Die zweite Veröffentlichung der Dissertation widmet sich dieser Forschungslücke, indem ein Text Mining Verfahren entwickelt wurde, das dynamisch zeitliche, geographische und technologische Kontextinformationen automatisiert identifiziert, kombiniert und in die nachgelagerte Analyse des Stimmungsbilds (mittels Sentimentanalyse) integriert. Methodisch wurde sich hierzu an dem nach Aggarwal et al. (2012) vorgeschlagene Vorgehen zur Entwicklung eines solchen Text Mining Verfahrens orientiert. Das entwickelte Analyseverfahren geht dabei über die rein tonalitätsbasierte Sentimentanalyse hinaus, indem es externe Kontextfaktoren (z. B. Kontinent, verwendetes Endgerät) als Einflussgrößen bei der Stimmungsanalyse berücksichtigt und so kontextsensitive Analyseergebnisse erzielt (**FF3**).

Die Anwendung des Analyseansatzes auf ~358 Mio. Tweets und die dabei erzielten Analyseergebnisse zeigen signifikante Stimmungsunterschiede in Abhängigkeit vom geografischen und technologischen Kontext. So konnte beispielsweise festgestellt werden, dass

Nutzer in unterschiedlichen Regionen (z. B. Nordamerika bzw. Asien) oder Endgeräten (iOS bzw. Android) mit höherer positiver Emotionalität kommunizieren. Gleichzeitig zeigte sich, dass die Anwendung des entwickelten Analyseverfahrens kontextualisierte Ergebnisse liefert, wodurch die Erschließung kontextsensitiver Erkenntnisse im Rahmen der Analyse sozialer Medien unterstützt wird.

3. Veröffentlichung. Aufbauend auf den gewonnenen Erkenntnissen der ersten beiden Veröffentlichungen adressiert die dritte Veröffentlichung die Frage, wie diese datenanalytischen Verfahren zu einem ganzheitlichen Trendanalysewerkzeug integriert werden können. Insbesondere rückt die Entwicklung eines softwaregestützten Analysewerkzeugs ins Zentrum, das in der Lage ist, die inhaltliche Entwicklung häufig diskutierter Themen kontextsensitiv im Zeitverlauf zu analysieren.

Vor diesem Hintergrund wurde im Rahmen der dritten Veröffentlichung ein integriertes Analysewerkzeug zur kontextsensitiven Trendanalyse konzipiert und prototypisch umgesetzt. Zunächst wurden im Rahmen des durchgeführten DSR-Projekts zentrale Anforderungen an die marketingorientierte Trendanalyse (z. B. Berücksichtigung der trendspezifischen Stimmungsbilder) abgeleitet (**FF4**). Diese Anforderungen setzen unter anderem die Kombination von Topic Modelling, Kontextanalyse (z. B. Ort, Zeit) und Sentimentanalyse in einem integrierten Systemdesign voraus. Aufbauend darauf wurde ein Softwareprototyp entwickelt und evaluiert (**FF5**). Der entwickelte Prototyp – MANTRA – ermöglicht dabei die kontextualisierte Identifikation sowie stimmungsbezogene Bewertung von Trendthemen auf Basis unstrukturierter, nutzergenerierter Daten (UGC).

Die Demonstration anhand 1,03 Mio. Online Customer Reviews der Plattform *Yelp* zeigt, dass MANTRA die identifizierten Anforderungen an eine automatisierte Trendanalyse im Marketing umfassend erfüllt, indem inhaltliche, kontextuelle und tonalitätsbasierte Analyseebenen sinnvoll integriert werden. So wurden im Rahmen der Demonstration zwei exemplarischen Trends – *Vegan Cuisine* und *Global Cuisine* – untersucht. Die Analyse dieser Trends erfolgte dabei entlang deren thematischer Entwicklungen im Zeitverlauf, deren geografischer Verteilungen (*Massachusetts*, *Texas* und *Oregon*) sowie deren assoziierter Stimmungsbilder.

Die Analyseergebnisse zeigen dabei, dass eine rein frequenzbasierte Analyse von Trends (z. B. durch gängige Social Listening Tools) ohne die Berücksichtigung von Kontextinformationen zu Fehlschlüssen führen kann. So wurde durch die Analyse im Bundesstaat *Texas* ein signifikanter Anstieg des Trends *Vegan Cuisine* zwischen 2020 und 2021 festgestellt, der auf Basis seiner Häufigkeitsentwicklung im Zeitverlauf als positives Signal interpretiert werden könnte. Eine integrierte Betrachtung dieser Ergebnisse mit der Sentimentanalyse offenbart jedoch, dass der Trend überwiegend negativ konnotiert ist (negativ: 0,61; positiv: 0,27), was auf enttäuschte Erwartungen bzw. einer kritischeren Haltung gegenüber veganer Ernährung hinweisen könnte. In *Massachusetts* hingegen zeigt sich ein ebenfalls steigender Trendverlauf, jedoch eine überwiegend positive Bewertung (positiv: 0,56; negativ: 0,30) des assoziierten Trends. Diese Differenzierung wird erst durch die Einbeziehung kontextueller und tonalitätsbezogener Informationen möglich. Hierdurch wird gezeigt, dass die bloße Identifikation eines Trends keineswegs automatisch auf eine positive Marktdynamik hinweist, sondern erst durch die Berücksichtigung relevanter Kontextfaktoren angemessen interpretiert werden kann. Das entwickelte Tool MANTRA adressiert diese Herausforderung gezielt, indem es die kontextsensitive Analyse von Trendentwicklungen in einem integrierten Analysewerkzeug bereitstellt und somit eine differenzierte Trendanalyse ermöglicht.

4. Veröffentlichung. Aufbauend auf den entwickelten methodischen Grundlagen zur kontextuellen Trendanalyse in der dritten Veröffentlichung wird im vierten Beitrag der Fokus der Trendanalyse auf eine spezifische Anwendungsperspektive vertieft: die *Produktentwicklung*. Im Zentrum steht dabei die Frage, wie VoC-Daten so analysiert werden können, dass nicht nur allgemeine Kundenwahrnehmungen zu einem Produkt(-trend) identifiziert, sondern differenzierte Einblicke in die Wahrnehmung der Kunden hinsichtlich spezifischer Produkt- und Serviceaspekte ermöglicht werden.

Vor diesem Hintergrund wurde ein integriertes Analyseverfahren konzipiert, welches Topic Modelling zur Identifikation relevanter Produkttrends mit Aspekt-basierter Sentimentanalyse kombiniert. Das Analyseverfahren wurde dabei auf Basis einer dedizierten Anforderungserhebung um spezifische Anforderungen der Produktentwicklung (z. B. Integration domänenspezifischen Expertenwissens [bspw. Produktontologien]) erweitert. Basierend auf dieser Konzeption wurde ein softwaregestütztes Analysewerkzeug prototypisch umgesetzt und evaluiert. Hierbei hat sich gezeigt, dass die Kundenwahrnehmung in VoC-Daten zu einzelnen Produkteigenschaften automatisiert analysiert und deren Entwicklung im Zeitverlauf kontextsensitiv abgebildet werden kann (**FF6**).

Die Anwendung des Analysewerkzeuges auf 37.638 Yelp-Reviews eines US-Restaurantkonzerns zeigt dabei z. B., dass bestimmte Aspekte eines Produkts (*Fleisch* bzw. *Soße* eines Burgers) in urbanen Regionen zeitweise positiver bewertet werden als in ländlichen Gebieten. Das entwickelte Analysewerkzeug erlaubt damit nicht nur eine spezifische Verfolgung von

Wahrnehmungsverschiebungen einzelner Produkt- und Serviceeigenschaften, sondern auch die Identifikation von ortsspezifischen Innovationspotenzialen. Damit trägt die Veröffentlichung zur Weiterentwicklung datengetriebener Analysewerkzeuge bei und stärkt insbesondere die Anschlussfähigkeit der Trendanalysen an operative Entscheidungen in der Produktentwicklung.

5. Veröffentlichung. Im fünften Beitrag rückt eine weitere zentrale Perspektive der *Produktentwicklung* in den Fokus: die automatisierte Identifikation besonders innovationsrelevanter Nutzergruppen auf Basis der Analyse von UGC. Im Zentrum steht hierbei das Konzept der *Lead User*, die durch ihre perspektivische Bedürfniswahrnehmung, ihre aktive Rolle im digitalen Diskurs sowie ihrer Vernetzungspotenziale als besonders einflussreiche Akteure für Innovationsprozesse gelten.

Die Identifikation solcher Nutzer stellt jedoch eine methodische Herausforderung dar – insbesondere vor dem Hintergrund der unterschiedlichen Anforderungen, die je nach Phase des Innovationsprozesses (*Ideengenerierung* vs. *Produktentwicklung*) bestehen. Während klassische Methoden zur Identifikation von Lead Usern häufig auf subjektiven Einschätzungen (Hienerth & Lettl, 2017), oder kleinen Stichproben beruhen (Hau & Kang, 2016), eröffnet die Analyse großer Mengen UGC die Möglichkeit, Lead User automatisiert und datenbasiert zu identifizieren. Im Zentrum dieses Vorhabens stehen dabei zwei sich ergänzende Forschungsgegenstände: Einerseits die konzeptionelle Modellierung eines merkmalsbasierten Analyseverfahrens zur automatisierten Identifikation von Lead Usern für verschiedenen Innovationsphasen (**FF7**) sowie andererseits die technische Realisierung des konzipierten Analyseverfahrens (**FF8**). Der fünfte Beitrag adressiert diese Potenziale, indem ein integratives Identifikationsverfahren entwickelt wird, das zentrale Merkmalsdimensionen von Lead Usern (z. B. *thematische Aktivität*, *Netzwerkpositionierung*) technisch abgebildet und mit Hilfe etablierter Verfahren des Data Science (z. B. *Topic Modelling*, *Sentimentanalyse*, *SNA*) operationalisiert werden.

Um die Anwendbarkeit und den geleisteten praktischen Mehrwert des entwickelten Analysewerkzeugs zu demonstrieren, wurde es auf *11.481* Beiträge von *945* Nutzern eines Online-Forums für *Kiteboarding* angewendet. Die identifizierten Lead User wurden im Anschluss durch qualitative Befragungen der Community, der identifizierten Nutzer sowie eines innovationsspezifischen Experten evaluiert und in ihrer Rolle bestätigt. Durch die erfolgreiche Entwicklung und Evaluation des Analysewerkzeugs konnten mehrere Implikationen für Forschung und Praxis herausgearbeitet werden – sowohl im Sinne domänenspezifischer

Kerneltheorien (Innovations- und Lead User Theorie) als auch mit Blick auf die Ableitung gestaltungsorientierter Designprinzipien (vgl. *Kapitel 3.2.2*). Darüber hinaus ermöglicht diese automatisierte und phasenspezifische Identifikation von Lead Usern eine gezielte Einbindung innovationsrelevanter Nutzergruppen entlang des gesamten Innovationsprozesses, wodurch Unternehmen ihre Innovationsstrategien passgenauer bedienen und letztlich marktnähere sowie bedarfsgerechtere Lösungen entwickeln können.

2. Forschungsgebiet: Zusammenfassung der Forschungsergebnisse im Rahmen der Konzeption und Entwicklung intelligenter Assistenzsysteme.

Mit dem Wechsel zum zweiten Forschungsgebiet verschiebt sich der Fokus der Dissertation von ökonomischen Anwendungsfeldern hin zu gesellschaftlichen Kontexten. Während im ersten Forschungsgebiet die Entwicklung datengetriebener Analyseverfahren zur kontextsensitiven Analyse von UGC mit Rahmen des Marketings bzw. der Produktentwicklung im Zentrum standen, wird im zweiten Forschungsgebiet das Potenzial generativer KI zur Gestaltung intelligenter, personalisierter Assistenzsysteme adressiert.

6. Veröffentlichung. Ausgehend von der zunehmenden Relevanz generativer KI beantwortet die sechste Veröffentlichung die Frage, inwiefern KI zur Weiterentwicklung von AAC-Systemen beitragen kann, um sprachbeeinträchtigte Personen gezielter bei der Bewältigung komplexer Gesprächssituationen zu unterstützen. Im Zentrum steht dabei die Analyse, welche KI-Nutzungspotenziale bereits in bestehenden AAC-Systeme mit Blick auf unterschiedliche Gesprächskontexte (z. B. *formell* vs. *informell*) und Gesprächsphasen (z. B. *Gesprächsbeginn*, *Gesprächsabschluss*) realisiert wurden und welche Potenziale sich durch den gezielten Einsatz von KI noch ergeben können.

Vor diesem Hintergrund wurde eine umfassende Literaturanalyse durchgeführt, um bestehende KI-basierte Lösungsansätze im Rahmen der AAC systematisch zu erfassen, zu klassifizieren und im Hinblick auf bestehende Forschungslücken zu bewerten (**FF9**). Auf dieser Grundlage wurden vier zentrale Wirkdimensionen identifiziert, in denen KI zur Unterstützung von AAC beitragen kann: die *Strukturierung des Gesprächsverlaufs*, die *kontextuelle Anpassungsfähigkeit*, die *Berücksichtigung des Gesprächspartners* sowie die *Individualisierung gegenüber dem AAC-Nutzer*.

Die Analyse bestehender KI-Nutzungspotenziale hinsichtlich dieser Wirkdimensionen zeigt, dass insbesondere im formellen Gesprächskontext ein erheblicher Forschungsbedarf besteht. Darüber hinaus offenbart die Literatur Defizite in Bezug auf eine ganzheitliche

Gesprächsabdeckung: Während sich viele bestehende Ansätze auf die Gesprächsmitte konzentrieren, fehlen systematische Lösungen für die Gesprächseröffnung und den Gesprächsabschluss. Aufbauend auf diesen Forschungspotenzialen wurden abschließend konkrete Ansatzpunkte zur Schließung dieser methodischen Lücken durch den Einsatz von KI (z. B. Nutzung künstlicher neuronaler Netze zur Unterstützung des gesprächsübergreifenden Konversationsverlaufs) entwickelt (**FF10**).

Generell verdeutlichen die Ergebnisse der Literaturanalyse einen wachsenden Handlungsbedarf im Hinblick auf die *Personalisierbarkeit* von AAC-Systemen. Bestehende Lösungen orientieren sich bislang nur unzureichend an den individuellen Bedürfnissen der Nutzer, was deren kommunikativen Handlungsspielraum einschränkt. Die Möglichkeit, Systeme stärker auf persönliche Ausdrucksstile und Anforderungen zuzuschneiden, stellt daher eine zentrale Voraussetzung für wirksame, anschlussfähige Kommunikation dar und bleibt eine zentrale Herausforderung für die Entwicklung inklusiver AAC-Systeme.

7. Veröffentlichung. Anknüpfend an die zentrale Notwendigkeit, digitale Assistenzsysteme stärker an die Nutzer anzupassen, richtet sich der Fokus des siebten Beitrags auf stilistische Passung generierter Sprachinhalte an die individuelle Persönlichkeit des Nutzers. Trotz großer Fortschritte im Bereich der AAC bleiben aktuelle AAC-Systeme in ihrer Ausdrucksfähigkeit häufig stilistisch starr, emotional entkoppelt und somit unzureichend personalisiert. Dieser Umstand führt für sprachbehinderte Nutzer zu einer Einschränkung der authentischen Selbstrepräsentation und der sozialen Anschlussfähigkeit.

Im Zentrum des Beitrags steht daher die Entwicklung eines intelligenten AAC-Systems, das in der Lage ist, individuelle Persönlichkeitsmerkmale seiner Nutzer systematisch zu erfassen und in die Generierung von Gesprächsbeiträgen einzubeziehen. Grundlage hierfür bildet das Konzept der Digital Persona – eine datenbasierte Repräsentation charakteristischer Persönlichkeitsdimensionen (basierend auf dem *Big Five Modell*) eines Nutzers. Diese Merkmale werden mittels generativer KI entlang sprachlich-stilistischer Dimensionen operationalisiert und bei der automatisierten Textproduktion berücksichtigt. Ziel ist es, eine personalisierte und authentische Kommunikation zu ermöglichen, die sich stärker an den individuellen Ausdrucksweisen der Nutzer orientiert.

Zur Realisierung wurde ein DSR-Projekt durchgeführt, welches sich an den Phasen des DSR-Prozesses nach Peffers et al. (2007) orientiert. Zunächst wurde die fehlende stilistische Personalisierung der Gesprächsbeiträge bestehender AAC-Systeme als zentrale Schwachstelle identifiziert. Auf dieser Basis wurden die Anforderungen an eine *Digital AAC Persona* abgeleitet, welche die Persönlichkeitsdimensionen des Nutzers abbildet, die eine stilistisch authentische Sprachgenerierung beeinflussen (z. B. *Extraversion*, *Offenheit für Erfahrungen*, *Neurotizismus*). Diese Dimensionen wurden in ein konzeptionelles Modell einer Digital Persona überführt, welches als integratives Artefaktdesign – basierend auf Designanforderungen und modelliert durch Design- und Implementierungsprinzipien – dient (**FF11**). Darauf aufbauend wurde das konzipierte Modell technische innerhalb eines prototypischen AAC-Systems implementiert (**FF12**). Hierbei wurde das LLM *GPT-4* durch eine Kombination verschiedener ML-Verfahren sowie gezieltem Prompt Engineering so gesteuert, dass nutzerspezifische, persönlichkeitsgetriebene Stilpräferenzen in der Textausgabe abgebildet wurden.

Schließlich wurde das Artefakt im Rahmen einer Evaluation hinsichtlich seiner Fähigkeit zur stilistisch persönlichkeitskongruenten Kommunikation getestet. Die Evaluation mit psychologischen Experten bestätigte dabei, dass die generierten Äußerungen mit den intendierten Persönlichkeitsprofilen der Nutzer resonieren und über verschiedene Gesprächssituationen hinweg als kontextuell passend wahrgenommen wurden.

8. Veröffentlichung. Nach der Auseinandersetzung mit personalisierten Kommunikationssystemen im Bereich der AAC (FF9 bis FF12) erweitert die achte Veröffentlichung den thematischen Fokus um eine gesundheitsbezogene Anwendungsperspektive. Im Zentrum steht die Frage, wie generative KI – in Kombination mit individuell ausgestalteten Digital Personas – zur Entwicklung nutzerzentrierter Kommunikationsagenten im Gesundheitsbereich beitragen kann. Besonders adressiert wird dabei das Anwendungsfeld der *Darmgesundheit*, in dem stark kontextabhängige Symptome, psychosomatische Einflussfaktoren und ein begrenzter Zugang zu medizinischer Beratung einen hohen Bedarf an individualisierter Unterstützung erzeugen.

Ausgehend von einer Literaturrecherche, einer Marktanalyse aktueller Gesundheits-Apps sowie User Stories wurden im Rahmen der Untersuchung elf zentrale Anforderungen an einen solchen digitalen Gesundheitsagenten systematisch abgeleitet. Diese umfassen unter anderem das *kontinuierliche Tracking und die Integration persönlicher Gesundheitsdaten* zur Erstellung personalisierter Gesundheitsempfehlungen oder die *Bereitstellung nutzerspezifischer Ernährungstipps* zur Förderung einer individuellen Darmgesundheit (z.B. basierend auf Unverträglichkeiten). Basierend auf diesen Anforderungen wurde ein initiales Systemdesign eines gesundheitsbezogenen Kommunikationsagenten entwickelt, der die Einbettung einer Digital Persona erlaubt und generative KI gezielt zur nutzerspezifischen

Empfehlungserstellung einsetzt (**FF13**). Die Untersuchung orientierte sich methodisch am DSR-Prozessmodell nach Peffers et al. (2007), wodurch eine theoriegeleitete und zugleich anwendungsorientierte Konzeption des Designs ermöglicht wurde. Damit schafft das entwickelte Systemdesign eine anschlussfähige Grundlage für nachfolgende Forschungsarbeiten, die nahtlos an das entwickelte Systemdesign anschließen und dessen technische Realisierung sowie empirische Evaluation weiterführen können.

Die in dieser Dissertation vorgestellten acht Veröffentlichungen verdeutlichen exemplarisch das Potenzial daten- und KI-gestützter Analysemethoden und Werkzeuge in unterschiedlichen Anwendungskontexten. Während im ersten Forschungsgebiet datengetriebene Analyseverfahren zur kontextsensitiven Analyse von UGC konzipiert, technisch umgesetzt und empirisch validiert wurden, rückte im zweiten Forschungsgebiet die Entwicklung intelligenter, personalisierter Assistenzsysteme in den Fokus. Beide Forschungsstränge verbindet dabei das Ziel, aktuelle Herausforderungen im Umgang mit digitalen Daten durch innovative, anwendungsnahe Data Science und KI-Ansätze zu adressieren und dabei sowohl wissenschaftliche Erkenntnisgewinne als auch konkrete praktische Gestaltungsimpulse zu generieren. Im folgenden Kapitel werden die daraus hervorgehenden technischen, theoretischen und praktischen Beiträge systematisch herausgearbeitet und in ihren forschungsübergreifenden Implikationen reflektiert.

3.2 Technologische, theoretische und praktische Wertbeiträge

Die Wertbeiträge der einzelnen Veröffentlichungen dieser kumulativen Dissertation werden jeweils im Rahmen der zugehörigen Artikel (vgl. *2. Kapitel*) detailliert dargestellt und diskutiert. Die Seitenverweise auf die entsprechenden Unterkapitel jeder Veröffentlichungen können Tabelle 23 entnommen werden.

Veröffentlichung	*Titel des Kapitels*	*Seitenverweis*
V1.	Discussion	47
V2.	Interpretation and Implications	64
V3.	Discussion and Contributions	81
V4.	Conclusion, Contribution and Outlook	97
V5.	Contribution for practice and research	122
V6.	Implications for Future Research	141
V7.	Discussion and Contributions	164
V8.	Implications and Discussions	175

Tabelle 23. Seitenverweise auf die Kapitel der Wertbeiträge jeder Veröffentlichung.

Die nachfolgenden Kapitel 3.2.1 und 3.2.2 zielen darauf ab, die forschungsübergreifenden Wertbeiträge der Dissertation herauszuarbeiten, die sich aus der ganzheitlichen Betrachtung der einzelnen Forschungsgegenstände ergeben haben. Im Fokus steht dabei nicht nur die Summe der Einzelerkenntnisse, sondern insbesondere deren systematische Verdichtung entlang *technischer*, *theoretischer* und *praktischer* Perspektiven. Dabei wird strukturiert aufgezeigt, inwiefern die im Rahmen der Dissertation entwickelten Konzepte, Methoden und Werkzeuge über die jeweiligen Einzelbeiträge hinaus einen substanziellen Mehrwert für Forschung und Praxis leisten. Die Abbildungen 2 bis 4 zeigen für jede der drei Perspektiven auf, in welchem Umfang die einzelnen Veröffentlichungen (*V1* bis *V8*) Beiträge zu den jeweiligen Schwerpunkten geleistet haben. Für die spezifischen Wertbeiträge jeder einzelnen Veröffentlichungen ist auf die Kapitel in den Veröffentlichungen verwiesen (vgl. *Tabelle 23*).

3.2.1 Technologiebezogene Beiträge

Im Rahmen der in dieser Dissertation entwickelten Analyseverfahren und Werkzeuge wurden konkrete technologiebezogene Wertbeiträge zur Weiterentwicklung der adressierten Technologien geleistet. Dabei wurden sowohl bestehende Methoden und Techniken weiterentwickelt als auch neuartige Analyseverfahren konzipiert, implementiert und evaluiert. Abbildung 2 gibt einen systematischen Überblick über die in den einzelnen Veröffentlichungen geleisteten technologiebezogenen Beiträge und ordnet diese entlang der Technologien, die jeweils im Fokus standen, ein. Dabei wird zwischen methodischen und technischen Wertbeiträgen differenziert.

	Topic Modelling	Sentimentanalyse	Social Network Analyse	Large Language Models
V1	○	-	-	-
V2	-	○ ▣	-	-
V3	○ ▣	-	-	-
V4	○ ▣	○ ▣	-	-
V5	▣	▣	○ ▣	-
V6	-	-	-	○
V7	-	-	-	○ ▣
V8	-	-	-	○

Legende	
▣	Geleisteter Beitrag ist technischer Natur.
○	Geleisteter Beitrag ist methodischer Natur.
-	Kein nennenswerter Beitrag.

Abbildung 2. Einordnung der geleisteten technischen Wertbeiträge je Veröffentlichung.

1. Forschungsgebiet: Einordnung der technologiebezogenen Wertbeiträge durch die Konzeption und Entwicklung kontextsensitiver Analyseverfahren.

Im Rahmen des ersten Forschungsgebiets wurden mehrere technologiebezogene Beiträge zur Weiterentwicklung datengetriebener Analyseverfahren geleistet. Im Zentrum stand die Entwicklung neuer Data Science Methoden und Werkzeuge für die kontextsensitive Analyse von UGC. Ausgangspunkt war die Beobachtung, dass bestehende Data Science Verfahren wie *Topic Modelling*, *Sentimentanalyse* und *SNA* zwar vielfach erprobt sind, sie jedoch hinsichtlich ihrer Aussagekraft und domainspezifischer Anschlussfähigkeit häufig an konzeptionelle Grenzen stoßen. Ziel der durchgeführten Forschungsarbeiten war es daher, die methodischen Grundlagen dieser Technologien gezielt zu erweitern und dies technisch so zu adaptieren, dass sie den domänenspezifischen Anforderungen des Marketings sowie der Produktentwicklung gerecht werden. Im Ergebnis entstanden kontextsensitive Analyseverfahren und -Techniken, die einerseits theoretisch fundiert sowie andererseits praxistauglich erprobt wurden. Dieses Unterkapitel systematisiert die geleisteten technologischen Beiträge entlang der neu- bzw. weiterentwickelten Analysetechniken.

Ein zentraler technischer Wertbeitrag wurde im Bereich des *Topic Modelling* geleistet. Aufbauend auf einer empirischen Evaluation der Modellansätze *LDA*, *DMR* und *PAM* vor dem Hintergrund dreier marketingspezifischer Anwendungsszenarien wurde gezeigt, dass die Eignung einzelner Verfahren stark vom jeweiligen Anwendungskontext abhängt. Vor diesem Hintergrund wurde ein mehrdimensionaler Evaluationsrahmen entwickelt, der marketingspezifischen Anforderungen sowohl in quantifizierbare (z. B. *Topic Coherence*, *Log-Likelihood*) als auch qualitative (z. B. *Word-*, *Topic-Intrusion*) Evaluationsmetriken übersetzt (vgl. *Kapitel 2.1*). Dieser Evaluationsrahmen erweitert damit bestehende Ansätze zur Evaluation um eine praxisnahe, marketingspezifische Perspektive und erlaubt es, unterschiedliche Topic Modelling Technologien neben ihrer statistischen Güte, auch hinsichtlich ihrer funktionalen Eignung systematisch zu bewerten. Die Konzeption dieses Evaluationsschemas stellt damit einen zentralen methodisch-technologiebezogenen Beitrag zum Topic Modelling dar.

Daneben wurden mehrere technische Erweiterungen im Topic Modelling realisiert (vgl. *Abbildung 2*), die das klassische Topic Modelling um die systematische Integration von Kontextfaktoren erweitert. Indem bspw. im Rahmen der dritten Veröffentlichung (vgl. *Kapitel 2.3*) zeitliche und geografische Kontextfaktoren in die Modellbildung einbezogen wurden, konnte ein Topic Modelling Verfahren entwickelt werden, welches Themen kontextsensitiv modellieren und deren Dynamiken über verschiedene Kontextfaktoren (z. B. *Geolokation*, *Zeit*) hinweg abbilden kann. Diese Weiterentwicklung ermöglicht eine differenziertere Analyse der generierten Themen und erhöht nicht nur die Interpretierbarkeit der Modellergebnisse (was eine zentrale Herausforderung im Topic Modelling darstellt), sondern verbessert auch die Anschlussfähigkeit des Topic Modellings an kontextsensitive Anwendungsszenarien (wie z. B. im Rahmen der Innovationsanalyse [*vgl. Kapitel 3.2.2.1*]). Diese geleistete Weiterentwicklung transformiert das klassische Topic Modelling damit von einer rein inhaltlichen Analysetechnik zu einem kontextsensitiven Analyseinstrument und leistet damit einen wichtigen technologiebezogenen Beitrag.

Auch im Bereich der *Sentimentanalyse* wurden technologiebezogene Fortschritte erzielt; insbesondere durch die technische Integration Aspekt-basierter Sentimentanalysen mit Topic Modelling zur Analyse produktspezifischer Kundenwahrnehmungen. Während herkömmliche Sentimentanalysen die Stimmungsbilder in UGC häufig auf einer generischen Ebene (z. B. dokumentenbasiert) aggregieren, bilden gerade Aspekt-basierte Einblicke in die Kundenwahrnehmungen großes Erkenntnisspotential im Rahmen der Produktentwicklung. Vor diesem Hintergrund wurde ein Aspekt-basiertes Sentimentanalyseverfahren entwickelt, welches *Topic Modelling* zur Identifikation häufig diskutierte Produkt- und Servicetrends integriert und auf den Modellergebnissen des Topic Modellings Aspekt-basierte Stimmungsbilder identifiziert. Durch diese Erweiterung wurde die Sentimentanalyse gezielt kontextualisiert, wodurch die Stimmungsanalyse nicht mehr nur rein isoliert für generische, sondern flexibel und kontextuell eingebettet für Aspekt-basierte Analysen erfolgen kann. Dies stellt eine technologische Erweiterung dar, indem die Sentimentanalyse in ihrer interpretativen Anschlussfähigkeit an themenspezifische Fragestellungen erweitert wurde. Zudem wurde die Aspekt-basierte Sentimentanalyse dahingehend erweitert, Stimmungsbilder differenziert vor dem Hintergrund verschiedener Lokationen analysieren zu können. Diese kontextualisierte Erweiterung erlaubt es, geografische Unterschiede innerhalb der in UGC zum Ausdruck gebrachten Stimmungen zu identifizieren. Durch diese beiden Adaptionen konnte die Aspekt-basierte Sentimentanalyse methodisch und technisch erweitert werden, was sie von einer reinen Stimmungsanalyse zu einer kontextsensitiven, stimmungsbezogenen Analysetechnik erweitert und ihre Anschlussfähigkeit an kontextgetriebene Anwendungsszenarien wie der Produktentwicklung erhöht.

Ein weiterer technologischer Beitrag besteht in der konzeptionellen und technischen Erweiterung von Verfahren der SNA durch die Integration datenbasierter Analyseverfahren. Während sich klassische SNA-Ansätze primär auf die Analyse struktureller Netzwerkmetriken

(z. B. Zentralität) fokussieren, wurde im Rahmen der Dissertation ein ganzheitliches Analyseverfahren entwickelt, das diese Netzwerk-Strukturanalysen systematisch mit inhaltlichen Informationen aus der Analyse von UGC kombiniert. Ziel war es, ein Netzwerk-basiertes Analyseverfahren zur automatisierten Identifikation von Lead Usern zu entwickeln, welches deren Merkmale technisch operationalisiert und somit Lead User automatisiert und merkmalsbasiert identifiziert. Das entwickelte Analyseverfahren integriert hierfür verschiedene Verfahren des Data Science (z. B. *Topic Modelling*, *Sentimentanalyse*, *Frequenzanalyse*) und geht damit über ein klassisches SNA-Verfahren hinaus, indem es ausgehend von strukturellen auch inhaltsbezogene Dimensionen in die Netzwerkanalyse integriert. So entstand ein neuartiger, multidimensionaler Analyseansatz, welcher eine flexible Integration unterschiedlicher Netzwerk- und Inhaltsanalysen vereint und ihre Resultate harmonisiert (vgl. *Kapitel 2.5*). Diese integrative Erweiterung erlaubt eine bisher nicht realisierte analytische Tiefenschärfe bei der Analyse sozialer Netzwerkstrukturen und erweitert damit das technologische Feld der SNA.

Insgesamt zeigt sich, dass durch die methodisch-technischen Beiträge dieser Dissertation zentrale Verfahren des Data Science – insbesondere *Topic Modelling*, *Aspekt-basierte Sentimentanalyse* und *SNA* – nicht nur einzeln weiterentwickelt, sondern systematisch miteinander integriert und um kontextspezifische Steuerungsmechanismen ergänzt wurden. Diese entwickelten Analyseverfahren stellen die Grundlage für die Umsetzung der Analysewerkzeuge dar (vgl. *Kapitel 3.2.3*) und ermöglichen eine anschlussfähige Analyse digitaler UGC zur automatisierten Erschließung der darin enthaltenen Erkenntnisspotentiale.

2. Forschungsgebiet: Einordnung der technologiebezogenen Wertbeiträge durch die Konzeption und Entwicklung intelligenter Assistenzsysteme.

Im Rahmen des zweiten Forschungsgebiets wurden technologiebezogene Beiträge zum Feld der generativen KI geleistet. Der Fokus lag auf der systematischen Anwendung und Adaption von LLMs – insbesondere GPT-4 – für die AAC und die digitale Gesundheitsförderung. Die im Rahmen dieser Dissertation entwickelten Technologien adressieren dabei eine neue Perspektive: die gezielte Nutzung generativer KI in intelligenten und *personalisierten* Assistenzsystemen.

Ein erster technologiebezogener Beitrag zur KI wurde durch die umfassenden Analyse der KI-Einsatzpotenziale zur Gesprächsunterstützung im Rahmen der AAC – differenziert entlang konkreter Gesprächsphasen (*Beginn, Mitte, Schluss*) und -Kontexten (*formell vs.*

informell) – geleistet. Diese Analyse ermittelte, dass vor allem die Wahrung des formalen Konversationsverlaufs sowie die allgemeine Nutzerzentriertheit bestehender AAC-Systeme noch Forschungspotenziale bieten. Auf Basis dieser Erkenntnisse wurden systematisch KI-basierte Techniken skizziert, um die identifizierten Herausforderungen zu adressieren. Diese konzipierten Lösungsansätze erweitern das bisherige technische KI-Verständnis um domänenspezifische, technische Anwendungsimpulse und erhöhen damit die technische Anschlussfähigkeit (generativer) KI an den Anwendungsfall der AAC.

Vor dem Hintergrund der erkannten Notwendigkeit, AAC-Systeme stärker auf ihre Nutzer auszurichten, wurden im Rahmen der Dissertation intelligente, personalisierte Assistenzsysteme konzipiert, prototypisch realisiert und evaluiert. Als abschließenden technologischen Beitrag verschiebt die Dissertation damit den Fokus von LLMs von der generischen Sprachgenerierung hin zur anwendungsspezifischen Adaption für die personalisierte Kommunikation und erschließt damit neue Anwendungspotenziale generativer KI. Dabei wurden LLMs durch ein nutzerzentriertes Designparadigma (*Digital Persona*) erweitert und aufgezeigt, wie generative KI individuell ausgestaltet werden muss, um als inklusive und gesundheitsfördernde Assistenztechnologie wirksam zu sein. Die Dissertation verdeutlicht dabei, dass der Mehrwert generativer KI nicht nur in der generischen Textproduktion liegt, sondern in ihrer Fähigkeit zur *nutzerzentrierten*, *sozial anschlussfähigen* Kommunikation. Die hierdurch eröffneten neuen Perspektiven transformieren LLMs von einer Technologie zur generischen Textgenerierung hin zu einer personalisierten und nutzerzentrierten Kommunikationstechnologie und leisten einen Wertbeitrag zur Weiterentwicklung des Forschungsfelds der generativen KI – insbesondere im Bereich der LLMs. Um diese technischen Forschungspotenziale weiter zu entfalten, erfordert es an neuen Forschungsarbeiten, die Methoden und Techniken zur Steuerung dieser Fähigkeiten entwickeln und somit die Potenziale von LLMs umfassend erschließen.

3.2.2 Theoriegeleitete Beiträge

Die theoretischen Wertbeiträge dieser Dissertation lassen sich zwei übergeordneten Bereichen zuordnen: Zum einen leistet sie Beiträge zur theoretischen Fundierung des wissenschaftlichen Diskurs mehrerer IS-Forschungsbereiche, indem zentrale Methoden, Modelle und Kerneltheorien weiterentwickelt und empirisch fundiert werden. Zum anderen ergeben sich durch die gestaltungsorientierte Ausrichtung einiger Untersuchungen Beiträge zur Designtheorie im Sinne der DSR. Um dieser doppelten Perspektive gerecht zu werden, wird folgend zwischen theoretischen Beiträgen zu domänenspezifischen Forschungsbereichen –

einschließlich einschlägiger Kerneltheorien – und zur Designtheorie unterschieden. Abbildung 3 ordnet die Wertbeiträge je Veröffentlichung entlang der beiden theoretischen Betrachtungsfelder ein.

	Forschungsbereiche					Designtheorie
	Social Media Analytics	Innovation	Lead User	AAC	Digital Persona	
V1	△	-	-	-	-	-
V2	△	-	-	-	-	-
V3	△	△	-	-	-	◆
V4	△	△	-	-	-	◆
V5	△	△	△	-	-	◆
V6	-	-	-	△	-	-
V7	-	-	-	△	△	◆
V8	-	-	-	-	△	◆

Legende	
△	Geleisteter Beitrag ist theoretischer Natur.
◆	Geleisteter Beitrag ist gestaltungsorientierter Natur.
-	Kein nennenswerter Beitrag.

Abbildung 3. Einordnung der geleisteten theoretischen Wertbeiträge je Veröffentlichung.

3.2.2.1 Beiträge zu den domänenspezifischen Forschungsbereichen

Durch die Dissertation wurden mehrerer theoretische Erkenntnisse generiert, die zur Weiterentwicklung domänenspezifischer Forschungsbereiche und darin enthaltener Kerneltheorien beitragen. Die nachfolgenden Abschnitte systematisieren diese geleisteten Wertbeiträge entlang ihrer theoriegeleiteten Schwerpunkte: (1) *Social Media Analytics*, (2) *Innovation*, (3) *Lead User,* (4) *AAC und (5) Digital Persona.*

> **1. Forschungsgebiet:** Einordnung der geleisteten Wertbeiträge im Rahmen der Konzeption und Entwicklung kontextsensitiver Analyseverfahren.

(1) Beiträge zum wissenschaftlicher Diskurs der Social Media Analytics

Durch die Konzeption und Entwicklung der datengetriebenen Analyseverfahren dieser Dissertation konnten theoretische Wertbeiträge zum Forschungsbereich sozialer Medien geleistet werden. Ein besonders relevanter theoretischer Bezugspunkt stellt dabei die *Social Media Analytics* (Phillips, 2012; Stieglitz et al., 2014; Stieglitz et al., 2018) dar.

Das Forschungsfeld der *Social Media Analytics* adressiert die methodischen und technologischen Herausforderungen bei der Auswertung großer Mengen unstrukturierter Social

Media Daten (vgl. Fan & Gordon, 2014; Stieglitz et al., 2014). Eine zentrale Problemstellung besteht dabei in der fehlenden Kontextsensitivität der Analysen und der daraus resultierenden eingeschränkten Interpretierbarkeit der Analyseergebnisse vor dem Hintergrund domänenspezifischer Anwendungsszenarien (Phillips, 2012). Besonders die Erkennung und Nutzbarmachung von emergenten Diskussionsthemen stellt die Forschung vor Herausforderungen, da deren Interpretierbarkeit stark vom eingebetteten Kontext abhängt (Stieglitz et al., 2018). Dadurch lässt sich die Bedeutung der Themen im jeweiligen Anwendungskontext oftmals nur schwer erschließen, was deren Verwendung zur Beantwortung domänenspezifischer Fragestellungen (z. B. in der Produktentwicklung) oftmals unzugänglich macht.

Die in der Dissertation entwickelten *kontextsensitiven* Analyseverfahren setzen genau an dieser Herausforderung an (vgl. *Kapitel 2.3, 2.4, 2.5*). So konnten bspw. zentrale Anforderungen an die kontextsensitive Trendanalyse auf Basis sozialer Medien fundiert abgeleitet und durch die methodische und technisch Entwicklung eines datengetriebenen Analyseverfahrens (*MANTRA*) realisiert werden. MANTRA ermöglicht es dabei, thematische Strukturen in großen Mengen unstrukturierter Social Media Daten automatisiert zu identifizieren und zugleich deren kontextuelle Einbettung (Entwicklung im Zeitverlauf, geografische Differenzierung) systematisch zu analysieren. So konnte gezeigt werden, wie zeitliche und geografische Kontexte in einem integrierten Analyseansatz vereint werden können, um eine kontextsensitive Erschließung emergenter Trendthemen in sozialen Medien zu ermöglichen.

Durch die in der Dissertation entwickelten Analyseverfahren wird der wissenschaftliche Diskurs zum Forschungsbereich der *Social Media Analytics* inhaltlich um domänenspezifische Analyseverfahren ergänzt und somit um domänenbezogene Tiefenschärfungen erweitert. Damit leistet die Dissertation einen Beitrag zur Weiterentwicklung des Forschungsfelds der *Social Media Analytics*, indem sie konkrete methodische und technische Lösungsimpulse für die bislang unzureichend adressierten Herausforderungen der kontextsensitiven Analyse sozialer Medien eröffnet.

(2) Beitrag zur Innovationstheorie

Das *Stage Gate Modell* nach Cooper (1996) stellt ein etabliertes theoretisches Konzept zur Steuerung von Innovationsprozessen dar. Es basiert auf der grundlegenden Annahmen der Innovationstheorie, dass Innovationen entlang strukturierter Phasen entwickelt und durch Entscheidungspunkte gesteuert werden können, um so Unsicherheiten in der Innovationsentwicklung zu reduzieren und die Marktfähigkeit neuer Produkte systematisch zu erhöhen.

Es stellt damit eine praxisorientierte Ausgestaltung der Innovationstheorie dar, indem es theoretische Konzepte der Innovationssteuerung in ein konkret anwendbares Prozessmodell überführt. Das Stage Gate Modell untergliedert den Innovationsprozess dafür in aufeinanderfolgende Phasen („Stages"), die durch sogenannte Go-/Kill-Entscheidungspunkte („Gates") strukturiert werden. An diesen Entscheidungspunkten wird auf Basis externer Marktanalysen (z. B. durch Kundenfeedback) bewertet, ob eine Produktinnovation weiterentwickelt oder verworfen wird. Die Einbindung dieser Marktsicht erfolgt dabei in erster Linie punktuell; zu Beginn des Innovationsprozesses im Rahmen der Ideengenerierung bzw. im Rahmen der *Innovationsbewertung* an den Entscheidungspunkten im Übergang zwischen den einzelnen Innovationsphasen (vgl. Cooper, 1996).

Die im Rahmen dieser Dissertation entwickelten Trendanalysewerkzeuge (vgl. *Kapitel 2.3* und *2.4*) ermöglichen es, (Produkt-)Trends automatisiert und flexibel durch die Analyse großer Mengen an digitalen Kundenstimmen zu identifizieren. Dies könnte die Möglichkeit eröffnen, relevante Marktinformationen (*Trends*) nicht mehr ausschließlich punktuell in der Innovationsbewertung (*Gates*), sondern konstant über alle Prozessphasen hinweg zu integrieren. Durch diese kontinuierliche Integration authentischer, ungefilterter Kundenstimmen in die phasenspezifischen Innovationsentwicklungen könnte sich eine Flexibilisierung des Innovationsprozesses ergeben. So könnten externe Marktinformationen nicht mehr nur als Evaluationsmaß zur Bewertung der Produktinnovation innerhalb der Gates dienen, sondern als ein konstitutiver Bestandteil zur Innovationssteuerung genutzt werden: Die Marktsicht würde damit zur ad hoc Justierung von Innovationen durch deren Ausrichtung an den realen Kundenbedürfnissen beitragen.

Vor dem Hintergrund dieser konstanten Marktausrichtung stellt sich die Frage, ob die bislang zentrale Prüf- und Bewertungsfunktion der Gates in ihrer bisherigen Form weiterhin notwendig wäre. Theoretisch könnte die Bewertung der Marktfähigkeit einer Produktinnovation – was zuvor Teil der Gate-Entscheidungspunkte war – reduziert oder gar obsolet werden, da auf Basis der kontinuierlichen Integration der Kundensicht die Marktfähigkeit konstant validiert und die punktuelle Überprüfung innerhalb der Gates redundant werden würde. Die Bewertungsfunktion der Gates würde damit in den laufenden Entwicklungsprozess integriert. Dies könnte strukturelle Flexibilisierungen des Stage Gate Modells ermöglichen, wodurch zentrale Strukturprinzipien des Modells neu gedacht werden könnten. So könnten Gates bspw. nicht mehr zwingend als obligatorische Übergangspunkte gesehen, sondern situativ aktiviert werden – z. B. bei kritischen Entscheidungen wie einer strategischen Neupositionierung einer Produktinnovation. Damit könnte eine neue Perspektive auf ein hybrides Innovationsmodell eröffnet werden, in dem die klassische Stage/Gate-Logik durch datenbasierte Steuerungsmechanismen ergänzt oder partiell abgelöst wird. In der Konsequenz könnte das die Grundlage für ein flexibleres und anpassungsfähigeres Innovationsmodell bilden, das die klassische Prozesslogik agiler gestaltet und eine höhere Reaktionsfähigkeit auf dynamische Innovationskontexte ermöglicht. Damit könnte auf die zunehmende Volatilität und Schnelllebigkeit von Innovationszyklen reagiert werden, welche eine zentrale Herausforderung moderner Innovationsprozesse darstellt (Gaglio et al., 2022).

Neben diesen strukturellen Implikationen könnte sich durch die phasenübergreifende Integration automatisierter Trendanalysen zudem das Potenzial konvergenter Innovationsentwicklungen erschließen lassen. Durch die fortlaufende, phasenweise Rückkopplung der Kundensicht an die entwickelten Produktinnovationen könnten sich die Innovation von Phase zu Phase enger an die tatsächlichen Bedürfnisse der Zielgruppe annähern. Die Produktinnovation könnte somit – über die Phasen hinweg – anhand tatsächlicher Kundenbedürfnisse iterativ verfeinert und spezifiziert werden. Aus theoretischer Perspektive ergibt sich hierdurch eine erweiterte Form der Innovationsleistungsfähigkeit, bei der nicht nur Risiken minimiert, sondern auch die Passung zwischen Innovation und Marktsicht systematisch erhöht werden könnte. Diese progressive Konvergenz zwischen realisierter Innovation und realer Kundenwahrnehmung könnte einen zentralen neuen Steuerungsmechanismus darstellen, um Marktpotenziale präziser zu erschließen und die Erfolgswahrscheinlichkeit von Innovationen zu erhöhen.

(3) Beitrag zur Lead User Theorie

Ein weiterer theoretischer Beitrag liegt in der Weiterentwicklung der Lead User Theorie. Durch die Konzeption eines modularen Analyseverfahrens zur automatisierten Identifikation von Lead Usern konnten zentrale Konzepte der Lead User Theorie differenziert aufgegriffen und methodisch erweitert werden.

So wurde das klassische Lead User Modell nach von Hippel (1986) um eine nutzungsorientierte Perspektive entlang der beiden Phasen des Innovationsprozesses (*Ideengenerierung* und *Produktentwicklung*) erweitert. Im Rahmen der Dissertation wurde gezeigt, dass sich Lead User nicht nur durch ihre definierenden Merkmale wie ihre Trend- oder Meinungsführerschaft charakterisieren lassen, sondern dass ihre innovative Wirksamkeit ebenfalls stark vom jeweiligen Einsatzzweck innerhalb des Innovationsprozesses abhängt. Eine differenzierte Betrachtung entlang der Innovationsphasen offenbarte dabei, dass unterschiedliche Typen von Lead Usern – z. B. ideengebende vs. anwendungserfahrene Lead User –

spezifische Stärken einbringen, die gezielt zur Steigerung der Innovationsleistung genutzt werden können. Damit konnte eine neue Dimension in das bestehende Lead User Modell eingeführt werden: Lead User sollten nicht nur anhand ihrer Merkmale klassifiziert, sondern auch im Hinblick auf ihren spezifischen Verwendungszweck innerhalb einzelner Innovationsphasen bewertet werden. Diese zweckorientierte Sichtweise ermöglicht eine zielgerichtetere Zuordnung gemäß der individuellen Stärken der Lead User mit Blick auf die jeweilige Innovationsphase und stärkt die konzeptionelle Grundlage für eine wirkungsorientierte Integration von Lead Usern in Innovationsprozesse.

Darüber hinaus könnte durch den entwickelten Analyseansatz, welcher eine automatisierte Trendanalyse in die Identifikation von Lead Usern integriert (vgl. *Kapitel 2.5*), eine methodische Flexibilisierung der Vier-Stufen-Methode zur Identifikation und Einbindung von Lead Usern nach von Hippel (1986) ermöglicht werden. Während die ursprüngliche Methode von einer strikt sequenziellen Abfolge der Trend- und Lead User Identifikation ausgeht – 1. Schritt Trendidentifikation, 2. Schritt Identifikation trendspezifischer Lead User – ermöglicht der vorgestellte Ansatz eine simultane Durchführung beider Schritte. Relevante Trends und dazugehörige Lead User werden dabei automatisiert in einem integrierten Verfahren identifiziert. Durch diese simultane Identifikation von Trends und zugehörigen Lead Usern auf Basis desselben Datenraums könnte einerseits deren kontextuelle Passung erhöht werden, da ein potenzieller Zeitverzug zwischen Trenddefinition und nachgelagerter Lead User Auswahl vermieden wird. Zum anderen stärkt die regelgeleitete, datenbasierte Identifikation die Objektivität und Replizierbarkeit des Prozesses und reduziert damit die subjektive Verzerrung, wie sie in explorativ geprägten Verfahren häufig auftritt. Dies könnte nicht nur die Effizienz des Identifikationsprozesses steigern, sondern auch die Anschlussfähigkeit der Methode an datenintensive Anwendungskontexte erhöhen.

Ein letzter Beitrag zur Lead User Theorie konnte durch die Entwicklung des mehrdimensionales Identifikationsverfahren geleistet werden, das die in der Literatur beschriebenen Lead User Merkmale technisch operationalisiert und in einer integrierten Analysemethode zusammenführt (vgl. *Kapitel 2.5*). Für jedes Merkmal wurde dabei eine technische Abbildung konzipiert, die dessen automatisierte Identifikation in unstrukturierten digitalen Daten ermöglicht. Diese technischen Merkmalsabbildungen wurde in einem gewichtbaren Identifikationsmechanismus vereint, der die relative Bedeutung einzelner Merkmale flexibel berücksichtigen kann. Der entwickelte Gewichtungsmechanismus ermöglicht damit eine graduelle Bewertung von Lead Usern entlang spezifischer Merkmalsausprägungen. Dadurch müssen Lead User nicht mehr ausschließlich auf Basis binärer Vorkommnisse – also ob ein

Merkmal vorhanden ist oder nicht – klassifiziert werden, sondern können stufenweise entlang eines Merkmalsprofils eingeordnet werden. Diese differenzierte Betrachtung eröffnet neue Perspektiven für die Kategorisierung und Auswahl von Lead Usern, da individuelle Nutzerprofile gezielter auf die Anforderungen unterschiedlicher Anwendungskontexte abgestimmt werden können. Aus theoretischer Perspektive erweitert das entwickelte Verfahren die Lead User Theorie nach von Hippel (1986) damit um eine datenbasierte, objektive Identifikationslogik, die eine zentrale Herausforderung der Lead User Methode adressiert: die zielgerichtete Identifikation geeigneter Lead User (vgl. Brem & Bilgram, 2015).

2. Forschungsgebiet: Einordnung der Wertbeiträge zu den Forschungsbereichen im Rahmen der Konzeption und Entwicklung intelligenter Assistenzsysteme.

(4) Beiträge zum wissenschaftlicher Diskurs der AAC und zur Big Five Persönlichkeitstheorie

Mit Blick auf den wissenschaftlichen Diskurs im Bereich der AAC leistet die Dissertation mehrere Beiträge zur theoretischen Fundierung. Ein theoretischer Erkenntnisgewinn zum Forschungsgebiet der AAC liegt in der systematischen Aufdeckung defizitärer Bereiche innerhalb aktueller AAC-Systeme. So zeigt die Untersuchung, dass insbesondere das Aufrechterhalten eines formalen Gesprächsverlaufs sowie die Unterstützung gesprächsübergreifender Konversationsverläufe bislang nur unzureichend durch bestehende Systeme unterstützt werden. Auf Basis dieser Defizite wurden konkrete KI-gestützte Forschungspotenziale erarbeitet. So konnte theoretisch skizziert werden, wie die Adaption künstlicher neuronaler Netze zur Wahrung gesprächsübergreifender Konversationskontexte genutzt werden könnten, wodurch eine theoretische Ausgangsbasis zur gezielten Adressierung bestehender Herausforderungen des Forschungsfelds der AAC geleistet wurde.

Darüber hinaus ergibt sich ein weiterer zentraler theoretischer Beitrag durch die theoretische Fokussetzung auf die bislang nur begrenzt untersuchte *Personalisierbarkeit* der AAC. So wurde im Rahmen der Dissertation verdeutlicht, dass bestehende AAC-Systeme die individuellen Ausdrucksbedürfnisse und stilistischen Präferenzen ihrer Nutzer oftmals nur unzureichend berücksichtigen. Dies schränkt die kommunikative Selbstbestimmung der Nutzer ein und verhindert deren authentische und anschlussfähige Selbstrepräsentation (Kane et al., 2017). Vor diesem Hintergrund wurde ein Ansatz zur personalisierten Sprachgenerierung entwickelt. Aufbauend auf der Big Five Persönlichkeitstheorie wurden zentrale Persönlichkeitsdimensionen technisch erfasst und in die sprachliche Ausdrucksweise generativer

KI-Systeme integriert. Durch die Abbildung der Persönlichkeitsdimensionen konnten persönlichkeitsabhängige Stilpräferenzen durch gezielte Manipulation des LLM *GPT-4* operationalisiert werden, wodurch sich die generierten Sprachbeiträge der Nutzer stilistisch an deren individuellen Persönlichkeit orientieren. Damit wird der theoretische Diskurs der AAC um eine methodisch-theoriegestützten Ansatz zur Integration der Nutzerpersönlichkeit erweitert und eröffnet eine neue Verbindung zwischen psychologischen Persönlichkeitsmodellen und generativer Sprachproduktion im Bereich der AAC. Die positiven Evaluationsergebnisse legen dabei nahe, dass zukünftige AAC-Forschungsarbeiten nicht nur auf die inhaltliche Verständlichkeit, sondern verstärkt auch auf die individuelle Persönlichkeitsrepräsentation ausgerichtet sein sollten, um eine authentische Selbstrepräsentation der Nutzer zu ermöglichen. Damit wurde ein erster Impuls gesetzt, der den theoretischen Fokus der AAC von der reinen Kommunikationsübertragung hin zur personalisierbaren Identitätsvermittlung verschiebt. Weitere Forschungsarbeiten können hier ansetzen und eine vertiefte Integration psychologischer und technischer Perspektiven adressieren. Die entwickelten Konzepte liefern hierfür erste theoretische und methodische Grundlagen und ebnen den Weg für ein neues, interdisziplinäres Forschungsparadigma für die Weiterentwicklung personalisierter AAC.

Angrenzend an diesen Kontext leistet die Dissertation darüber hinaus einen theoretischen Beitrag zur *Big Five Persönlichkeitstheorie (*Costa und McCrae, 1992*)*. Durch die Integration der *Big Five* Persönlichkeitsdimensionen in ein generative KI-gestütztes Kommunikationssystem zur Generierung persönlichkeitsgetreuer Kommunikation wurde das Persönlichkeitsmodell aus seiner klassischen, diagnostisch-deskriptiven Anwendungsperspektive in ein handlungsleitendes Steuerungskonzept für die automatisierte AAC-Textgenerierung überführt. Die Persönlichkeitsmerkmale stellen hierbei nicht länger nur deskriptive Kategorien dar, sondern steuern die sprachlich-stilistischen Gesprächselemente der generierten Gesprächsbeiträge. Eine erste empirische Evaluation mit psychologischen Experten unterstützt dabei die Annahme, dass sich die *Big Five* Dimensionen nicht nur zur Beschreibung menschlicher Persönlichkeitsunterschiede eignen, sondern auch als strukturgebende Grundlage für die synthetische Erzeugung personalisierter Kommunikation genutzt werden können. Auf diese Weise erweitert die Dissertation den Anwendungsrahmen der *Big Five Persönlichkeitstheorie* um eine digital operationalisierte Persönlichkeitsmodellierung zur persönlichkeitsgetreuen Sprachgenerierung.

(5) Beiträge zum wissenschaftlichen Diskurs der Digital Persona

Abschließend leistet die Dissertation durch die Konzeption und prototypische Umsetzung domänenspezifischer Digital Persona Modelle in den Anwendungsfeldern der AAC und der digitalen Gesundheitsförderung einen theoretischen Beitrag zur methodischen Weiterentwicklung dieses Forschungsfeldes. Im Kontext der AAC wurde ein Konzept für eine persönlichkeitsbasierte Digital AAC Persona entwickelt, das auf den Dimensionen des Big Five Persönlichkeitsmodells basiert und die automatisierte Generierung personalisierter Gesprächsbeiträge ermöglicht. Im Bereich der digitalen Gesundheitsförderung stand die Entwicklung einer gesundheitsbezogenen Digital Persona zur individuellen Unterstützung der Darmgesundheit im Fokus.

Die entwickelten Digital Persona Modelle repräsentieren dabei eine konzeptionelle Neupositionierung: Das Digital Persona Konzept wird dabei nicht länger nur als statische Nutzerrepräsentation verstanden (vgl. Roosendaal, 2010), sondern als interaktives, sich entwickelndes System, das individuelle Nutzerparameter kontinuierlich erfasst und in personalisierte Verhaltenssteuerung überführt. Durch diese Perspektivverschiebung wird das bestehende theoretische Verständnis der Digital Persona von einem passiven Abbild individueller Nutzereigenschaften um eine verhaltenssteuernde Dimension ergänzt. Damit wird ein Impuls zur theoretischen Weiterentwicklung des Konzepts der Digital Persona geliefert, der dessen Verständnis von einer statischen Nutzerrepräsentation hin zu einem dynamischen, verhaltenssteuernden Abbild eines Nutzers weiterentwickelt und als Grundlage für weiterführende Untersuchungen in diesem Bereich dienen kann.

Insgesamt zeigt die Dissertation mit ihren geleisteten theoretischen Beiträgen, wie die Entwicklung datenanalytischer Verfahren nicht nur technische Innovationen hervorbringen, sondern auch zur theoretischen Fundierung domänenspezifischer Forschungsfelder beitragen kann. Durch die Entwicklung mehrerer Analyseverfahren und deren theoriebasierter Rückbindung liefert die Dissertation substanzielle Beiträge zur Weiterentwicklung etablierter Kerneltheorien und schlägt damit eine Brücke zwischen datengetriebener Analysepraxis und theoriebasierter Systemgestaltung, wodurch ein integrativer Beitrag zur Weiterentwicklung der IS-Forschung geleistet wird.

3.2.2.2 Beiträge zur Designtheorie

Zusätzlich zu den Beiträgen zu den Kerneltheorien leistet die Dissertation einen substanziellen Beitrag zur Designtheorie im Sinne gestaltungsorientierter Forschung. Im Zentrum stehen hierbei die systematische Entwicklung und Evaluation domänenspezifischer Designprinzipien, die auf Basis explizit abgeleiteter Designanforderungen formuliert wurden. In mehreren DSR-Projekten – u. a. im Kontext der Trendanalyse in Social Media, der datenbasierten Lead User Identifikation sowie der Personalisierung digitaler Assistenzsysteme – wurden konkrete Designprinzipien entwickelt, technisch umgesetzt und im Rahmen der Artefaktentwicklung demonstriert. Dadurch konnte ein Beitrag zur „*nascent design theory*“ (Gregor & Hevner, 2013) geleistet werden, indem frühe theoriebasierte Erkenntnisse zur Gestaltung von IT-Artefakten systematisch strukturiert und dokumentiert werden.

Die Designprinzipien erfassen dabei theoriebasiertes und domänenspezifisches Gestaltungswissen, das die funktionale und strukturelle Gestaltung vergleichbarer Artefakte in verwandten Domänen unterstützen kann. Die empirische Fundierung dieser Design Prinzipien erfolgte einerseits durch die Evaluation der Artefakte in realitätsnahen Anwendungsszenarien, wodurch die Anwendbarkeit und Wirksamkeit der Prinzipien implizit validiert wurde (vgl. Heinrich & Schwabe, 2014). Andererseits wurden die entwickelten Designprinzipien im Rahmen der Entwicklung der *Digital AAC Persona* durch Experteninterviews evaluiert, wodurch neben der impliziten Fundierung durch die realitätsnahe Demonstration der Designprinzipien, ebenso eine explizite Fundierung durch Experten erreicht wurde.

Der geleistete Beitrag dieser Dissertation zur Designtheorie kann beispielhaft durch die Entwicklung des Analysewerkzeugs *MANTRA* verdeutlicht werden. Aus den zugrunde liegenden domänenspezifischen Anforderungen wurde ein spezifisches Systemdesign – inklusive der darin geltenden Designprinzipien – abgeleitet. So betont etwa das Prinzip der *kontextabhängigen Analysefähigkeit* die Notwendigkeit, den Trendanalyseansatz so zu konzipieren, dass geografische Unterschiede innerhalb der Analyse berücksichtigt werden können. Durch die Umsetzung und Evaluation dieses Prinzips konnte ein gestaltungsbezogener Erkenntnisgewinn erzielt werden, der über das konkrete Artefakt hinaus als Blaupause für die Entwicklung vergleichbarer Systeme dient.

Insgesamt leistet die Dissertation damit einen übergreifenden Beitrag zur Designtheorie, indem sie domänenspezifische, empirisch fundierte und theoriegeleitete Designprinzipien zur Gestaltung mehrerer softwarebasierter Analyse- und Unterstützungswerkzeuge bereitstellt.

3.2.3 Praxisorientierte Beiträge

Zudem wurden mehrere praxisorientierte Wertbeiträge in Form von Analysewerkzeugen geleistet. Die entwickelten Werkzeuge adressieren unterschiedliche Anwendungskontexte – von der kontextsensitiven Trendanalyse über die Lead User Identifikation bis hin zur personalisierten Kommunikationsunterstützung – und operationalisieren theoretisch fundierte Analyseansätze in nutzbaren Systemlösungen. Abbildung 4 gibt einen systematischen Überblick über die in den einzelnen Veröffentlichungen realisierten praktischen Beiträge und ordnet diese entlang der jeweils adressierten Domänen ein.

	Trendanalyse		Produkt- bzw. Innovationsanalyse		Lead User Analyse		Intelligente Assistenzsysteme	
V1	-		-		-		-	
V2	-		-		-		-	
V3	○	⬢	-		-		-	
V4	○	⬢	○	⬢	-		-	
V5	-		○	⬢	○	⬢	-	
V6	-		-		-		○	
V7	-		-		-		○	⬢
V8	-		-		-		○	

Legende	
⬢	Geleisteter Beitrag ist ein technisches Analyseverfahren bzw. Werkzeug.
○	Geleisteter Beitrag ist methodischer Natur.
-	Kein nennenswerter Beitrag.

Abbildung 4. Einordnung der geleisteten praktischen Wertbeiträge je Veröffentlichung.

1. Forschungsgebiet: Einordnung der praktischen Wertbeiträge durch die Konzeption und Entwicklung kontextsensitiver Analyseverfahren.

Ein praktischer Beitrag dieser Dissertation besteht in der Entwicklung von Analysewerkzeugen zur datengetriebenen *Trendanalyse* auf Basis von UGC aus sozialen Medien. Die entwickelten Werkzeuge ermöglichen es Unternehmen, zentrale Fragestellungen datenbasiert zu untersuchen und können somit eine fundierte Entscheidungsfindung unterstützen.

So wurden im Rahmen der Dissertation zwei softwaregestützte Analysewerkzeuge entwickelt, die den praktischen Einsatz datenanalytischer Verfahren in Unternehmen gezielt unterstützen. Das Tool MANTRA wurde dabei zur explorativen Trendanalyse konzipiert und erlaubt es, externe Kontextfaktoren wie zeitliche, geografische oder demografische Merkmale flexibel in die Analyse von Trends einzubeziehen. Im Unterschied zu herkömmlichen

Social Listening Tools, die auf der Häufigkeitsanalyse vordefinierter Schlagwörter beruhen, basiert MANTRA auf der Identifikation signifikanter Kookkurrenzen in großen Textmengen. Dadurch lassen sich auch bislang unbekannte Trends erkennen – ohne dass ein spezifisches Vorwissen (z. B. trendspezifische Schlagwörter) erforderlich ist. Unternehmen können so frühzeitig und explorativ aufkommende Entwicklungen im Markt identifizieren. Daran anknüpfend wurde mit Blick auf die Produktentwicklung ein weiterer praktischer Wertbeitrag durch das zweite Analysewerkzeug zur Aspekt-basierten Analyse von Kundenwahrnehmungen geleistet. Das entwickelte Analysewerkzeug erweitert die klassische Trendanalyse durch die Kombination von Topic Modelling und Aspekt-basierter Sentimentanalyse um die Möglichkeit, Kundenwahrnehmungen gezielt auf Ebene einzelner Produkt- und Serviceaspekte sowie über verschiedene Zeiträume und Standorte hinweg auszuwerten und deren Entwicklung zu verfolgen. Unternehmen, welche ihre Produkte bspw. parallel in mehreren regionalen Märkten anbieten, können so lokationsabhängige Unterschiede in der Kundenwahrnehmung analysieren und gezielte Maßnahmen zur Produkt- oder Serviceanpassung ableiten.

Darüber hinaus wurde ein softwaregestütztes Analysewerkzeug entwickelt, das die automatisierte Identifikation von Lead Usern ermöglicht. Dabei adressiert das Analysewerkzeug eine der größten Herausforderungen der klassischen Lead User Methode: den zeit- und ressourcenintensiven Identifikationsprozess geeigneter Nutzer (Brem & Bilgram, 2015). Das entwickelte Analysewerkzeug erlaubt es Unternehmen, große Mengen an Social Media Daten effizient zu analysieren und potenzielle Lead User automatisiert anhand definierter Charakteristika zu identifizieren. Das Analysewerkzeug bietet hierfür einen modularen Analyseansatz, der auf unternehmensspezifische Anforderungen zugeschnitten werden kann. So können eigene Datensätze zur datenbasierten Identifikation der darin enthaltenen Lead User hochgeladen, spezifische Lead User Charakteristika flexibel in die Analyse in- bzw. exkludiert und deren Gewichtung individuell angepasst werden. Auf diese Weise wird eine automatisierte, flexibel anpassbare und zugleich methodisch fundierte Identifikation von Lead Usern ermöglicht – differenziert nach den unterschiedlichen Phasen des Innovationsprozesses. Damit werden Unternehmen mit einem leistungsfähigen Werkzeug unterstützt, um strategisch relevante Nutzergruppen datenbasiert zu identifizieren und gezielt in Innovationsprozesse zu integrieren. Das entwickelte System leistet damit einen praxisrelevanten Beitrag zur Operationalisierung der Lead User Methode und unterstützt insbesondere innovationsgetriebene Unternehmen bei der effizienteren Gestaltung der einzelnen Innovationsphasen.

2. Forschungsgebiet: Einordnung der praktischen Wertbeiträge durch die Konzeption und Entwicklung intelligenter Assistenzsysteme.

Darüber hinaus leistet die Dissertation mehrere substanzielle Beiträge zur praktischen Nutzbarmachung generativer KI im Kontext intelligenter Assistenzsysteme. Ein erster praktischer Beitrag ergibt sich aus der durchgeführten systematischen Literatur- und Marktanalyse, die bestehende KI-Ansätze in AAC-Systemen strukturiert erfasst und entlang konkreter Gesprächsphasen und Gesprächskontexte systematisiert. Die hieraus hervorgehende Übersicht bildet einen praxisnahen Analyserahmen, der insbesondere Anbietern von AAC-Systemen als Grundlage für strategische Entwicklungsentscheidungen dienen kann. Die identifizierten Einsatzpotenziale ermöglichen dabei eine zielgerichtete Systementwicklung und eröffnen damit neue Perspektiven für die bedarfsgerechte Weiterentwicklung bestehender AAC-Systeme. Ein daran anknüpfender praktischer Beitrag für die Anbieter resultiert aus der Konzeption und prototypischen Umsetzung des intelligenten, persönlichkeitsgetriebenen AAC-Prototyps. Dieser Prototyp zeigt auf, dass generative KI nicht nur syntaktisch korrekte, sondern auch individuell stilisierte Gesprächsbeiträge erzeugen kann – vorausgesetzt, die Spracherzeugung wird durch geeignete Persönlichkeitsmodelle theoretisch fundiert. Durch die Implementierung des *Big Five* Persönlichkeitsmodells und dessen Integration in die automatisierte Textgenerierung wurde ein nutzerzentriertes AAC-System entwickelt, das personalisierte Kommunikation ermöglicht. Das dabei entwickelte Systemdesign und das Wissen hinsichtlich der Implementierung des AAC-Prototyps stellt für Systemanbieter eine anwendungsbezogene Referenz dar, die aufzeigt, wie die funktionalen Integration generativer KI für personalisierte AAC-Anwendungen konkret umgesetzt werden können.

Von dem persönlichkeitsgetriebenen AAC-Ansatz können jedoch nicht nur Systemanbieter profitieren, sondern vor allem die sprachbeeinträchtigten Nutzer selbst. Für Personen mit Sprachstörung leistet der entwickelte Prototyp eine wichtigen Impuls zur Erweiterung ihrer kommunikativen Ausdrucksfähigkeit. Bestehende AAC-Systeme erzeugen häufig unpersönliche, generische Sprachbeiträge, die nicht dem individuellen Artikulationsstil der Nutzer entsprechen und somit sozial distanzierend wirken können (Light et al., 2021). Die generierten Sprachbeiträge des Prototyps zeigen jedoch, dass eine stilistisch authentische Kommunikation der Nutzer ermöglicht werden kann, wodurch deren Selbstrepräsentation und das kommunikative Selbstbewusstsein gestärkt werden könnte. Durch den Prototyps konnte ein erster praktischer Impuls für die Entwicklung solch nutzerzentrierter und personalisierter AAC-Systeme geleistet werden.

Über den Bereich der AAC hinaus wird auch im Kontext der digitalen Gesundheitsförderung ein praktischer Beitrag geleistet. Die Konzeption eines nutzerzentrierten Gesundheitsagenten zur Förderung der Darmgesundheit zeigt exemplarisch auf, wie generative KI zur individuellen Gesundheitsförderung eingesetzt werden kann, um personalisierte Empfehlungen auf Basis nutzerspezifischer Merkmale zu generieren. Das im Rahmen der Dissertation entwickelte Konzept dieser *Gut Health Digital Persona* liefert abschließend erste Impulse zur gesundheitsfördernden Verhaltenssteuerung durch individuelle, nutzerzentrierte Adaption generativer KI. Auf Basis des entwickelten Konzepts können so konkrete Entwicklungsimpulse zur technische Umsetzung eines solchen Kommunikationsassistenten geliefert werden, wodurch Softwarehersteller bei der praktischen Umsetzung unterstützt werden können. Für Menschen mit chronischen Verdauungsproblemen oder eingeschränktem Zugang zu ärztlicher Beratung könnte so ein niedrigschwelliges, individuell adaptierbares Unterstützungsangebot im Alltag eröffnet werden.

3.3 Kritische Würdigung und Ausblick auf weitere Forschungsfelder

Die Veröffentlichungen (V1 bis V8) der Dissertation leisten mehrere Wertbeiträge (vgl. *Kapitel 3.2*). Dabei wurden zentrale technische, theoretische sowie praktische Herausforderungen adressiert und innovative Lösungsansätze entwickelt. Dennoch weisen die einzelnen Beiträge auch Limitationen auf, die sich aus methodischen, theoretischen bzw. praktischen Fokussetzungen ergeben. Diese Limitationen sind dabei nicht als Schwächen im Sinne einer unzureichenden Qualität zu verstehen; vielmehr verkörpern sie Grenzmarkierungen, an denen gezielt neue Forschungsvorhaben anknüpfen können. Zur abschließenden Systematisierung weiterer Forschungspotenziale der Dissertation werden in *Kapitel 3.3.1* zuerst die Limitationen der Veröffentlichungen betrachtet und die vier übergeordneten Limitationskategorien *(A) Anforderungserhebung, (B) Technikvielfalt bzw. -aktualität, (C) interdisziplinäre Fundierung* sowie *(D) realweltliche Evaluation* abgeleitet (vgl. *Tabelle 24*). Die strukturierte Betrachtung dieser Kategorien ermöglicht eine konsolidierte Reflexion darüber, in welchen Bereichen anschließende Forschungsarbeiten Potenziale zur Vertiefung der bestehenden Forschungsergebnisse bieten. *Kapitel 3.3.2* überführt diese Erkenntnisse in konkrete Perspektiven zukünftiger Forschung.

3.3.1 Limitationen der Dissertation

1. Veröffentlichung. In der ersten Veröffentlichung besteht eine Limitation durch die begrenzte Auswahl an Topic Modelling Verfahren, welche für den Vergleich herangezogen wurden. Mit LDA, DMR und PAM wurden zwar drei etablierte Methoden in den Vergleich einbezogen, die jeweils für eine unterschiedliche Erweiterungskategorie (z. B. DRM für semi-supervisiert; PAM für hierarchische Topic Modelling Verfahren) des klassischen Topic Modellings stehen (vgl. Liu et al., 2016), dennoch bilden sie nur einen Ausschnitt des verfügbaren technischen Spektrums ab. Insbesondere jüngere Verfahren wie z. B. neuronale Topic Modelling Verfahren (vgl. Wu et al., 2024), konnten bei der Durchführung des Vergleichs aufgrund fehlender methodischer und technischer Umsetzungen noch nicht berücksichtigt werden und könnten zu abweichenden oder noch präziseren Technologie-Empfehlungen führen.

Darüber hinaus orientieren sich die abgeleiteten Anforderungen, welche zur Bewertung der einzelnen Topic Modelling Techniken formuliert wurden, vorrangig an den in der Literatur formulierten praktischen Anforderungen des digitalen Marketings. Zwar erlaubt diese fokussierte Perspektive eine passgenaue Bewertung der Techniken mit Blick auf die domänenspezifischen Anwendungsszenarien (z. B. Trendanalyse), jedoch ist hierbei nicht auszuschließen, dass die Durchführung einer realen Marktstudie oder die Integration von Domänenexperten aus der Praxis zu abweichenden oder ergänzenden Anforderungen führen würden. Diese könnten das Evaluationsschema erweitern bzw. die bestehenden Bewertungen verschieben und so zu leicht veränderten Einschätzungen hinsichtlich der Eignung einzelner Verfahren führen.

2. Veröffentlichung. Die in der zweiten Veröffentlichung entwickelte Methode zur Identifikation von Kontextfaktoren und deren anschließende Anwendung im Rahmen der kontextsensitiven Sentimentanalyse stellt einen innovativen Beitrag zur kontextsensitiven Analyse von UGC in sozialen Medien dar. Eine zentrale Limitation besteht jedoch darin, dass der entwickelte Analyseansatz bislang keiner systematischen Evaluation unterzogen wurde. So wurde der entwickelte Analyseansatz bislang primär im Rahmen der Anwendung zur Stimmungsanalyse in sozialen Medien (V2) sowie im Rahmen der kontextualisierten Trendanalyse (V3 und V4) angewandt, wodurch ein impliziertes Grounding des Verfahrens erzielt werden konnte; eine explizite Evaluation, die die zielgerichtete Anwendung und Bewertung des Verfahrens vor dem Hintergrund der Extraktion dynamischer Kontextfaktoren untersucht, steht jedoch noch aus.

Darüber hinaus bezieht sich die durchgeführte Analyse auf einen zeitlich begrenzten Datenausschnitt, der ausschließlich Tweets aus dem Jahr 2019 umfasst. Auch wenn dieser Fokus eine konsistente Vergleichbarkeit innerhalb des Untersuchungsdesigns gewährleistet, ist nicht auszuschließen, dass sich einzelne Effekte in anderen Zeiträumen anders darstellen.

3. und 4. Veröffentlichung. In der dritten und vierten Veröffentlichung wurde jeweils softwarebasierte Analysewerkzeug zur datengetriebenen Trendanalyse entwickelt. Beide Artefakte leisten einen wesentlichen Beitrag zur Operationalisierung datenanalytischer Verfahren und eröffnen neue domänenspezifische Perspektiven für die Analyse von UGC. Dennoch weisen auch diese Untersuchungen Limitationen auf.

Einerseits besteht die Möglichkeit, dass durch eine Analyse interdisziplinärer Forschungsliteratur zusätzliche Anforderungen identifiziert werden könnten, die im bisherigen Systemdesign nicht berücksichtigt wurden. Andererseits erfolgte die Evaluation der entwickelten Werkzeuge bislang primär auf technischer Ebene. So wurde die funktionale Umsetzung der Anforderungen durch die entwickelten Artefakte erfolgreich technisch demonstriert und evaluiert, jedoch fehlen bislang systematische Untersuchungen zur Nutzungstauglichkeit aus Anwendersicht. Eine anschließende Evaluation der Usability, beispielsweise im Rahmen einer SUMI-Studie (Software Usability Measurement Inventory; Kirakowski & Corbett, 1993), könnte wertvolle Hinweise auf potenzielle Verbesserungsbedarfe liefern.

5. Veröffentlichung. Die in der fünften Veröffentlichung identifizierten Charakteristika zur automatisierten Identifikation potenzieller Lead User basieren auf einer systematischen Auswertung einschlägiger Literatur und bilden eine fundierte Grundlage für deren technischen Operationalisierung im Rahmen des entwickelten Analyseverfahren. Trotz einer umfassenden und systematischen Literaturanalyse ist jedoch nicht auszuschließen, dass in weiteren Untersuchungen zusätzlich relevante Charakteristika identifiziert werden würden, die im Rahmen der vorliegenden Analyse nicht berücksichtigt wurden. Dies könnte zu einer Erweiterung der bisher berücksichtigten Merkmale oder zu einer Anpassung der Merkmalsgewichtung führen, was im Resultat einen Einfluss auf die Identifikation der Lead User haben könnte.

Zum anderen wurde der Ansatz bislang primär im Rahmen eines domänenspezifischen Anwendungsfalls (*Identifikation von Lead Usern im Bereich des Kitesurfings*) erprobt. Durch die dynamische Integration der Lead User Merkmale sowie durch den entwickelten Gewichtungsmechanismus kann die Analyse flexibel an verschiedene Kontexte angepasst werden. Hierbei stellt sich die Frage, inwiefern die entwickelten Identifikationskriterien und der Gewichtungsmechanismus auf andere Domänen übertragbar sind.

6. Veröffentlichung. Die in der sechsten Veröffentlichung durchgeführte systematische Literaturanalyse strukturiert den aktuellen Forschungsstand zu den Potenzialen KI-gestützter Ansätze im Kontext der AAC. Wie bei Literaturanalysen üblich, handelt es sich jedoch um eine Momentaufnahme, die zeitlich bedingt nur jene Arbeiten erfassen kann, die zum Zeitpunkt der Recherche in einschlägigen wissenschaftlichen Datenbanken verfügbar waren. Trotz einer sorgfältigen Durchführung des Reviews – einschließlich der Einbindung mehrerer Forscher zur Sicherstellung der Interrater-Reliabilität – lässt sich nicht ausschließen, dass einzelne relevante Publikationen unberücksichtigt blieben. Dies könnte im Einzelfall zu geringfügigen Verschiebungen bei der Gewichtung der identifizierten Forschungslücken führen. Zudem basieren die abgeleiteten KI-Nutzungspotenziale zur Unterstützung von AAC-Systemen bislang ausschließlich auf Erkenntnissen aus der Forschungsliteratur. So könnte eine ergänzende Validierung durch eine systematische Marktanalyse aufzeigen, in welchen Bereichen die in der Forschung identifizierten Potenziale bereits konkrete in am Markt verfügbaren AAC-Systemen umgesetzt wurden und wo noch ungenutzte Einsatzpotenziale bestehen.

7. Veröffentlichung. Im Rahmen der siebten Veröffentlichung wurde ein AAC-Prototyp entwickelt, der auf einer generativen KI-basierten Digital Persona zur personalisierten Kommunikationsunterstützung basiert. Während die technische Funktionsfähigkeit des Systems im Sinne der von Venable et al. (2016) vorgeschlagenen „Technical Risk & Efficacy"-Evaluationsstrategie evaluiert werden konnte, steht eine weiterführende empirische Evaluation im realen Nutzungskontext noch aus. Damit bleibt die Frage offen, inwiefern die entwickelten Personalisierungsmechanismen unter realen Einsatzbedingungen durch die Betroffenen bspw. als hilfreich, intuitiv und angemessen wahrgenommen werden.

8. Veröffentlichung. Wie bei systematischen Literaturanalysen üblich, stellt auch der im Rahmen der achten Veröffentlichung durchgeführte SLR lediglich eine Momentaufnahme des aktuellen Forschungsstands dar. Trotz eines strukturierten Suchprozesses, der Nutzung einschlägiger Datenbanken sowie einer mehrstufigen Qualitätskontrolle durch vier unabhängige Forscher (*Interrater-Reliabilität nach Fleiss' Kappa = 0,89*), lässt sich nicht ausschließen, dass relevante Studien nicht identifiziert oder nicht berücksichtigt wurden. Dies betrifft insbesondere die Auswahl der Designanforderungen, auf deren Grundlage das Systemdesign entwickelt wurde. Vor diesem Hintergrund ist nicht auszuschließen, dass bei einer alternativen Studienauswahl andere Schwerpunkte bei den Anforderungen identifiziert und damit auch ein abweichendes Systemdesign resultieren könnten.

Die in den Veröffentlichungen identifizierten Limitationen lassen sich systematisch entlang vier übergeordneter Kategorien (***A*** bis ***D***) strukturieren:

(A) Anforderungserhebung
Die in mehreren Veröffentlichungen durchgeführte Anforderungsanalyse auf Basis systematischer Literaturauswertungen ermöglicht eine erste, theoriegeleitete Fundierung technischer Lösungen. Jedoch bleibt diese Perspektive in ihrer Reichweite begrenzt, da sie primär auf wissenschaftlich dokumentierte Anforderungen zurückgreift und potenzielle Bedarfe aus der realen Anwendungspraxis nur indirekt abbildet. Dadurch besteht das Risiko, dass bestimmte Nutzungskontexte oder Zielgruppenbedürfnisse unberücksichtigt bleiben und die entwickelten Systeme nicht in vollem Maße auf praktische Herausforderungen abgestimmt sind.
(B) Technikvielfalt bzw. -aktualität
Obwohl in den bisherigen Arbeiten zentrale und bewährte datenanalytische Techniken verwendet wurden, konnte die rasanten Weiterentwicklung dieser dynamischen Technologien bislang nur eingeschränkt berücksichtigt werden. Dies kann zur Folge haben, dass Potenziale neuere oder leistungsfähigere Techniken, welche zum Zeitpunkt der durchgeführten Forschungsarbeiten noch nicht verfügbar waren, mit höherer Anpassungsfähigkeit, Interpretierbarkeit oder Robustheit ungenutzt bleiben.
(C) Interdisziplinäre Fundierung
Die durchgeführten Arbeiten orientierten sich konsequent an etablierten Methoden und Modellen der IS, insbesondere im Kontext der gestaltungsorientierten Forschung. Dabei wurden bereits erste interdisziplinäre Anknüpfungspunkte aufgegriffen, z. B. durch die Integration psychologischer Persönlichkeitsmodelle. Allerdings blieb diese disziplinübergreifende Theoriebildung bislang punktuell und wurde nicht systematisch vertieft. Dadurch bleiben Potenziale ungenutzt, um die entwickelten Systeme theoretisch noch stärker zu fundieren und deren domänenspezifische Anschlussfähigkeit zu erhöhen.
(D) Realweltliche Evaluation
Die entwickelten Artefakte wurden alle erfolgreich technisch-funktional demonstriert und evaluiert. Damit konnten ihre grundsätzliche Funktionalität nachgewiesen und ein erster methodischer Beitrag zur Gestaltung dieser Systeme geleistet werden. Für einige der Artefakte steht jedoch eine umfassende Evaluation mit den intendierten Nutzergruppen im realen Anwendungskontext noch aus. Dadurch bleibt die Fragen offen, wie die Systeme tatsächlich angenommen, verstanden und genutzt werden.

Tabelle 24. Klassifikation der Limitationen.

Um die Limitationen systematisch einordnen und darauf aufbauend gezielt in zukünftigen Forschungsarbeiten adressieren zu können, wurden diese entlang des (abstrahierten) Forschungsvorgehens dieser Dissertation betrachten (vgl. Abbildung 5). Die durchgeführten Forschungsarbeiten folgten – trotz unterschiedlicher methodischer Ausrichtung – alle einem generischen Forschungsprozess, der sich grob in drei Phasen unterteilt lässt:

(1) die initiale Problem- und Anforderungsdefinition,

(2) die Konzeption und Entwicklung einer Problemlösung,

(3) die Evaluation der Zielerreichung.

Vor diesem Hintergrund lassen sich die Limitationskategorien (A bis D) den jeweiligen Phasen wie folgt zuordnen:

Kategorie A (*Anforderungserhebung*) betrifft dabei typischerweise die erste Phase des Forschungsprozesses, in der Problemstellungen analysiert und domänenspezifische Anforderungen an eine Lösung formuliert werden. Kategorien B (*Technikvielfalt/-aktualität*) und C (*Interdisziplinäre Fundierung*) lassen sich der zweiten Phase zuordnen, da sie zentrale Aspekte der Konzeptualisierung, Entwicklung und disziplinären Verankerung einer Problemlösung betreffen. Kategorie D (*Realweltliche Evaluation*) betrifft schließlich die abschließende Phase, in der erzielte Forschungsergebnisse vor dem Hintergrund ihrer Zielerreichung überprüft und bewertet werden.

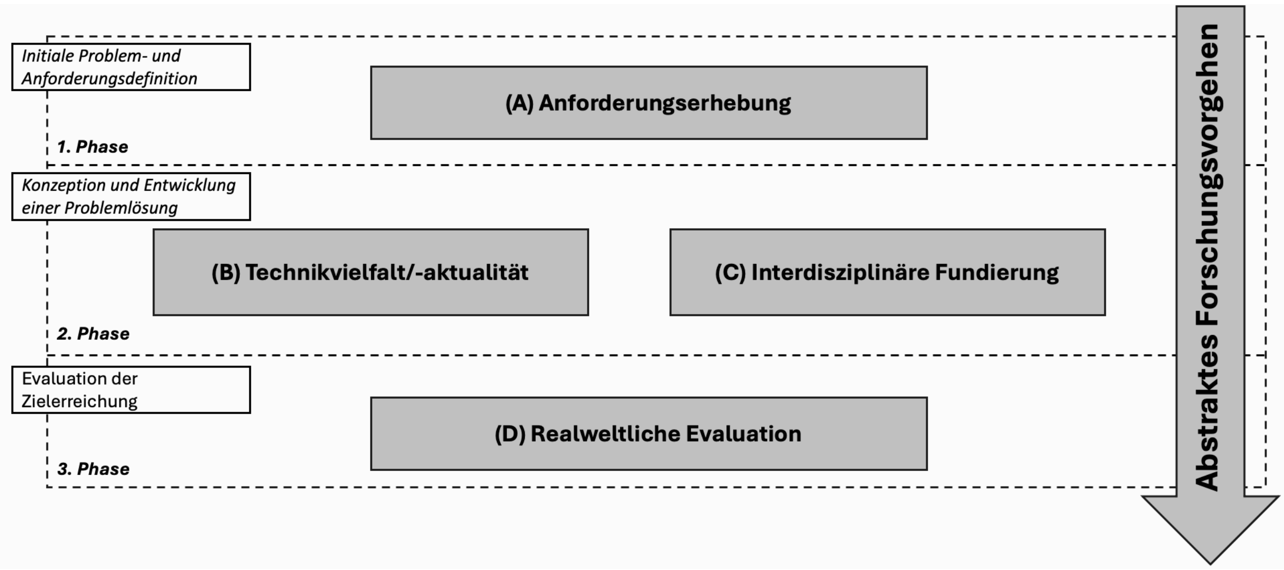


Abbildung 5. Verortung der Limitationen im Forschungsvorgehen der Dissertation.

Durch diese Verortung entlang des abstrahierten Forschungsvorgehens der Dissertation lassen sich gezielte Ansatzpunkte für zukünftige Forschungsarbeiten zur Adressierung dieser Limitationen ableiten (vgl. *Kapitel 3.3.2*).

3.3.2 Perspektivische Forschungsansätze zur Adressierung bestehender Limitationen

Aufbauend auf der in Kapitel 3.3.1 vorgenommenen Systematisierung zentraler Limitationen lassen sich gezielte Perspektiven für zukünftige Forschungsvorhaben ableiten (vgl. *Abbildung 6*). Diese Perspektiven stellen dabei eine Reaktion auf bestehende Limitationen dar und setzen direkt auf den bisherigen Erkenntnissen der Dissertation auf.

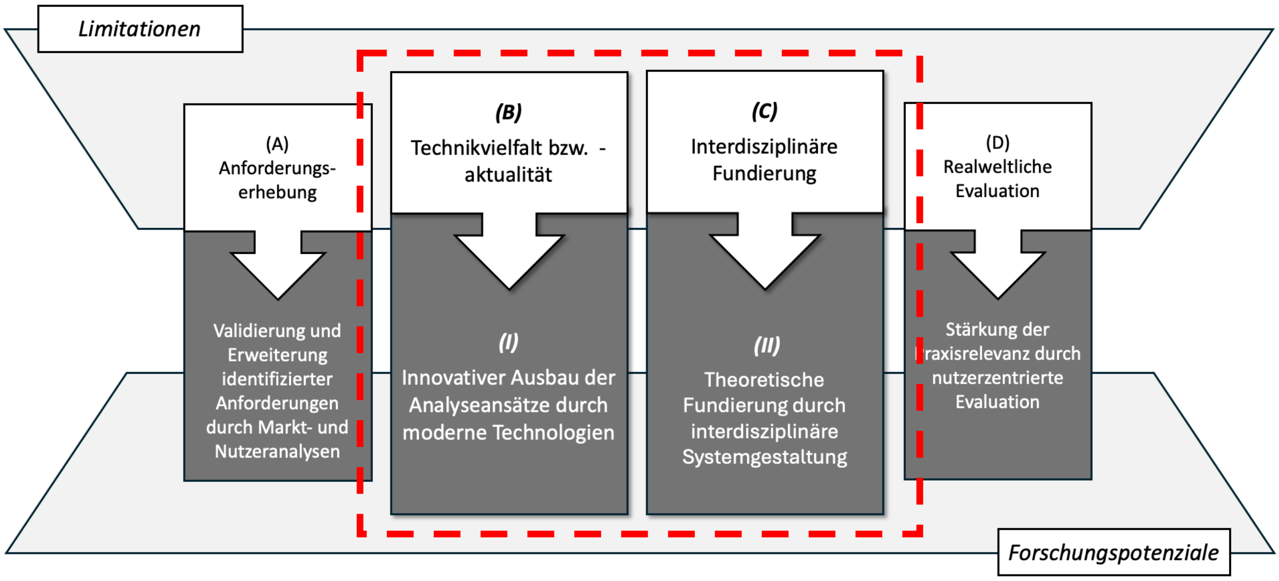


Abbildung 6. Systematisierung perspektivischer Forschungspotenziale.

Während die Limitationen der Kategorien ***B*** und ***C*** hauptsächlich die theoretische bzw. technische Ausgestaltung der zu entwickelnden Problemlösung betreffen, beziehen sich die Kategorien ***A*** und ***D*** auf die Einbindung externer Akteure in den Forschungsprozess. Die Einbindung externer Akteure ist in der Forschungsliteratur methodisch bereits gut dokumentiert, wodurch sich diese Limitationen durch etablierte Vorgehensweisen wie z. B. zur Durchführung semi-strukturierter Interviews (vgl. Adeoye-Olatunde & Olenik, 2021) oder einer nutzerzentrierten Evaluationen (vgl. Kirakowski & Corbett, 1993) direkt adressieren lassen. Demgegenüber besteht zur Adressierung der Limitationen B und C ein höherer theoretischer Forschungsbedarf, da sie nicht nur methodisch-operative Anpassungen, sondern neue theoretische und technische Perspektiven erfordern.

Aus diesem Grund liegt der Fokus der folgenden Betrachtung auf den beiden Limitationskategorien **B** und **C**. Um zukünftige Forschungsarbeiten zur Adressierung dieser Limitationen und zur damit einhergehenden Erschließung der darin liegenden Forschungspotenziale perspektivisch zu leiten, werden im Folgenden die zwei Forschungsperspektiven ***I*** und ***II*** skizziert. Ziel ist es, eine Orientierung für anschlussfähige Arbeiten zu bieten und damit die nachhaltige Wirkung der Dissertation im wissenschaftlichen Diskurs zu stärken.

(I) Innovativer Ausbau der Analyseansätze durch moderne Technologien

Die in dieser Dissertation entwickelten Analyseansätze basieren auf etablierten, vielfach erprobten Verfahren des Data Science wie z. B. semi-supervisiertem Topic Modelling, Aspekt-basierter Sentimentanalyse oder LLMs (vgl. *Kapitel 2*). Angesichts der dynamischen Entwicklungen dieser Technologiefelder, kann jedoch nicht ausgeschlossen werden, dass in der Zwischenzeit modernere Technologien entwickelt wurden, welche die identifizierten Herausforderungen (z. B. im Rahmen der kontextsensitiven Trendanalyse) noch besser adressieren könnten. Besonders im ersten Forschungsgebiet, worin klassische Topic Modelling Verfahren zur Trendmodellierung genutzt wurden, könnten sich vielversprechendes Potenziale durch die Integration modernerer Technologien ergeben.

1. Forschungspotenzial: Untersuchung des Potenzials neuronaler Topic Modelling Verfahren zur Weiterentwicklung der Trendanalyse.

Zukünftige Forschungsarbeiten könnten untersuchen, ob sich für die entwickelten Trendanalyseverfahren durch die Integration *neuronaler Topic Modelling* Techniken (vgl. Wu et al., 2024) konkrete Vorteile gegenüber der verwendeten Analysetechnik eröffnen könnten. Während die in der Dissertation entwickelten Verfahren auf klassischem Topic Modelling basieren, welches zur Themenmodellierung auf die statistische Analyse von Wortverteilungen setzt, nutzen neuronale Verfahren vortrainierte Sprachmodelle (z. B. BERT, GPT), um die semantische Bedeutung der Themen besser erfassen zu können. Um die damit einhergehenden (potenziellen) Vorteile zu untersuchen, werden Forschungspotenziale skizziert, die sich durch die Integration des neuronalen Topic Modelling Verfahrens *BERTopic* (Grootendorst, 2022) in das bestehende Trendanalyseverfahren ergeben könnten.

Ein zentraler Vorteil neuronaler Verfahren wie BERTopic gegenüber klassischen Topic Modelling Ansätzen besteht in der Möglichkeit, Themen nicht mehr nur inhaltlich, sondern auch semantisch zu erschließen. Während klassische Topic Modelling Verfahren Themen auf Basis häufigkeitsbasierter Schlagwörter identifiziert, generiert BERTopic kontextsensitive, sprachlich zusammenhängende Themenbeschreibungen (Grootendorst, 2022). Anstatt lediglich einzelne Schlüsselbegriffe pro Thema zu extrahieren, können durch BERTopic so vollständige, semantisch interpretierbare Themen formuliert werden, was die Interpretierbarkeit der Themen erhöhen kann. Ein klassisches Topic Modelling Verfahren könnte ein Thema rund um Kundenfeedback zu einem neuen Smartphone-Update anhand der Schlagwörter „*Update*", „*Fehler*", „*Akku*" identifizieren. Diese Begriffe geben dabei zwar Hinweise auf das Thema, erfordern jedoch eine manuelle Interpretation, um den inhaltlichen

Zusammenhang zu verstehen. BERTopic hingegen könnte dasselbe Thema als zusammenhängende Beschreibung modellieren, aus der ersichtlich wird, dass durch die Kunden *Kritik am neuen Update wegen schnellerem Akkuverbrauch* geäußert wurde. Diese semantisch angereicherte Themenformulierung erleichtert die direkte inhaltliche Einordnung und verringert einerseits den Interpretationsaufwand und andererseits die Subjektivität bei der Interpretation.

Daraus könnte sich ein vielversprechendes Forschungspotenzial für die kontextsensitive Trendanalyse ergeben. Dabei gilt es zu untersuchen, ob BERTopic – insbesondere unter Berücksichtigung der domänenspezifischen Anforderungen an kontextsensitive Trendanalysen im Marketing – in der Lage ist, valide und realitätsnahe Trendbeschreibungen zu erzeugen. Sollte sich dies empirisch bestätigen lassen (z. B. durch Experteninterviews), könnte durch die Integration von BERTopic in den bestehenden Analyseansatz die bisher notwendige nachgelagerte Interpretation der identifizierten (*schlagwortbasierten*) Trends durch semantische, kontextsensitive Trendbeschreibungen ergänzt werden. Entscheidungsträger hätten somit unmittelbar Zugriff auf interpretierbare, semantisch angereicherte Trendinformationen, was vor dem Hintergrund der Schnelllebigkeit von Trends einen essenziellen Vorteil darstellen könnte.

Ein weiteres Potenzial zur Weiterentwicklung der Trendanalyseverfahren könnte zudem in der Echtzeitfähigkeit von BERTopic liegen. Im Gegensatz zum klassischen Topic Modelling, bei dem neue Daten in der Regel einen erneuten Durchlauf des Modell-Trainingsprozesses erfordern, ermöglicht BERTopic durch die Nutzung vortrainierter Sprachmodelle die dynamische Integration neuer Inhalte, ohne dass das zugrunde liegende Modell erneut trainiert werden muss. Dadurch können neue Daten unmittelbar in das bestehende Modell eingespeist und ohne Verzögerung analysiert werden.

Zukünftige Forschungsarbeiten könnten daher untersuchen, ob diese Fähigkeit zur kontinuierlichen Datenintegration genutzt werden kann, um ein laufend aktualisiertes Trendmodell zu entwickeln, das neu aufkommende Trends zuverlässig und (nahezu) in Echtzeit identifizieren kann. Auf diese Weise ließe sich der bislang notwendige Zwischenschritt des erneuten Modelltrainings umgehen lassen, wodurch eine ressourcenschonendere Trendanalyse ermöglicht werden könnte. Zudem würde der manuelle Schritt der Erhebung bzw. Bereitstellung neuer Daten für das erneute Modelltraining entfallen, wodurch der Automatisierungsgrad der Trendanalyse verbessert werden könnte. Zur Flexibilisierung der Trendanalyse wäre es technisch somit denkbar, aktuelle Datenströme durch die Anbindung von

Social Media APIs in das neue Trendanalyseverfahren zu integrieren, um so neue Beiträge (nahezu) in Echtzeit in das Analyseverfahren einzuspeisen und somit automatisiert in die Trendanalyse zu integrieren. Dies würde sowohl die Reaktionsgeschwindigkeit der Trendanalyse steigern als auch das Automatisierungspotenzial der eingesetzten Werkzeuge erweitern – was einen wesentlichen Vorteil bei der frühzeitigen Erkennung und Bewertung dynamischer Trends darstellen würde.

Insgesamt eröffnen sich durch die Integration neuronaler Topic Modelling Verfahren wie BERTopic zwei zentrale Perspektiven für zukünftige Forschung: Einerseits könnte durch die semantisch fundiertere Themenerschließung die inhaltliche Interpretierbarkeit der identifizierten Trends verbessert werden; andererseits bietet die Echtzeitfähigkeit dieser Modelle das Potenzial, neue Inhalte automatisiert und ohne Verzögerung in die Analyse zu integrieren. Beide Potenziale adressieren damit zentrale Anforderungen kontextsensitiver Trendanalysen und markieren vielversprechende Ansatzpunkte für die Weiterentwicklung der datengetriebenen Analyseverfahren.

(II) Theoretische Fundierung durch interdisziplinäre Systemgestaltung

Die in dieser Dissertation entwickelten Analyseverfahren und Werkzeuge orientierten sich konsequent an etablierten Methoden und Modellen der IS, insbesondere im Kontext der gestaltungsorientierten Forschung. Dabei wurden bereits erste interdisziplinäre Anknüpfungspunkte aufgegriffen. Allerdings blieb diese disziplinübergreifende Theoriebildung bislang punktuell und wurde nicht systematisch vertieft. Dadurch bleiben Potenziale ungenutzt, um die entwickelten Systeme theoretisch noch stärker zu fundieren und deren domänenspezifische Anschlussfähigkeit zu erhöhen. Besonders im zweiten Forschungsbereich bietet die Integration interdisziplinärer Theorien vielversprechende Forschungspotenziale, um die bestehenden Untersuchungen zur AAC theoretischer zu fundieren und die identifizierten Limitationen in (C) zu adressieren.

2. Forschungspotenzial: Untersuchung des Potenzials der Self Determination Theory zur Erweiterung der nutzerzentrierten AAC-Systemgestaltung.

Im Rahmen des zweiten Forschungsgebiets der Dissertation wurde ein AAC-System entwickelt, das auf Basis der Big Five Persönlichkeitstheorie (Costa & McCrae, 1992) personalisierte Sprachäußerungen generiert. Die zugrundeliegende Idee bestand darin, die generierten Sprachbeiträge stilistisch stärker an die individuelle Persönlichkeit des Nutzers anzupassen, um so eine authentischere Kommunikation für den Nutzer zu ermöglichen. Ziel war es, durch diese personalisierte Anpassung einen Beitrag zur Stärkung der Nutzerzentrierung zu leisten und damit einen ersten Weg in Richtung inklusiverer AAC-Systeme zu ebnen. Während durch die Big Five eine etablierte Grundlage zur Erfassung der individuellen Persönlichkeiten der Nutzer geschaffen wurde, erscheint es vor dem Hintergrund der Stärkung der Nutzerzentrierung vielversprechend, ergänzend auch die *psychologischen Grundbedürfnisse* der Nutzer zu berücksichtigen.

Ein vielversprechendes Potenzial zur theoretisch-fundierten Berücksichtigung der individuellen psychologischen Nutzerbedürfnisse könnte sich durch die Integration der *Self Determination Theory* (SDT) von Deci & Ryan (1985) in das AAC-Systemdesign ergeben. Die SDT beschreibt drei universelle psychologische Grundbedürfnisse: *Autonomie*, *Kompetenz* und *soziale Eingebundenheit*. Die Theorie besagt, dass eine Befriedigung dieser drei Grundbedürfnisse zu einer höheren intrinsischen Motivation, einem besseren Wohlbefinden und einer positiven Persönlichkeitsentwicklung führt. Umgekehrt können mangelnde Befriedigung dieser Bedürfnisse zu Demotivation, Stress und negativen Folgen für die psychische Gesundheit führen. Die SDT könnte sich daher ideal als komplementäre Theorie zur Big Five Personalisierung eignen, da sie den motivationalen Unterbau personalisierter Kommunikation liefert. Während die Big Five beschreiben, *wie* Menschen sind, beschreibt die SDT, *was Menschen motiviert* – und was sie benötigen, um sich ernst genommen und eingebunden zu fühlen (Deci & Ryan, 1985). So könnte die Berücksichtigung dieser Dimensionen einen zentralen Beitrag für die Gestaltung nutzerzentrierter AAC-Systeme darstellen. Zukünftige Forschungsarbeiten könnten daher untersuchen, wie die drei psychologischen Grundbedürfnisse der SDT konkret in der Gestaltung von AAC-Systemen berücksichtigt und technologisch operationalisiert werden können.

So könnten zukünftige Forschungsarbeiten bspw. untersuchen, wie das psychologische Grundbedürfnis nach Autonomie im Design von AAC-Systemen gezielt berücksichtigt und technisch umgesetzt werden könnte. Autonomie im Sinne der SDT beschreibt das menschliche Bedürfnis, selbstbestimmte Entscheidungen zu treffen und Kontrolle über das eigene Handeln auszuüben (Deci & Ryan, 1985) – ein Aspekt, der gerade in der AAC eine große Rolle spielt. Denn Personen mit eingeschränkter Sprachfähigkeit erleben häufig einen Kontrollverlust über ihre Ausdrucksmöglichkeiten, was ihre Autonomie im kommunikativen Alltag erheblich einschränkt (Uchoa et al., 2021).

Zukünftige Forschungsarbeiten könnten daher untersuchen, wie AAC-Systeme so gestaltet werden können, dass sie den Nutzern ein höheres Maß an Kontrolle über die

Kommunikation ermöglichen. Dabei könnten zwei zentrale Richtungen einschlagen werden: Erstens ist es notwendig, systematisch zu untersuchen, welche Gestaltungsmerkmale von AAC-Systemen überhaupt geeignet sind, das psychologische Bedürfnis der Nutzer nach Autonomie gezielt zu fördern. Hierzu könnten z. B. qualitative Studien mit Betroffenen durchgeführt werden, um Autonomie-Barrieren und -Präferenzen im Hinblick auf selbstbestimmte Kommunikation zu identifizieren. Solche Erkenntnisse würden es ermöglichen, konkrete Gestaltungsvorgaben wie individuell konfigurierbare Steuerungsmechanismen zur personalisierten Kommunikationserzeugung abzuleiten, um das Gefühl von Kontrolle über die eigene Kommunikation zu stärken. Zweitens ließe sich auf diesen Erkenntnissen aufbauend die technologische Umsetzung und Evaluation entsprechender Systemfunktionalitäten realisieren – z. B. durch die prototypische Entwicklung adaptiver Steuerungsmechanismen, die in explorativen Nutzungsstudien auf ihre Wirkung im Hinblick auf wahrgenommene Autonomie und Nutzerakzeptanz geprüft werden können.

Insgesamt zeigt sich, dass durch die Integration interdisziplinärer Theorien wie der SDT ein Beitrag zur theoretischen Fundierung nutzerzentrierter AAC-Systemgestaltung geleistet werden könnte. Durch die systematische Berücksichtigung psychologischer Grundbedürfnisse wie Autonomie lassen sich neue Gestaltungsideen und Entwicklungspotenziale für AAC-Systeme ableiten, welche zukünftigen Forschungsarbeiten zur gezielten Adressierung zentraler Limitationen im Hinblick auf die theoretische Fundierung (vgl. Kategorie C) dienen können. Schlussendlich könnte damit zu einer stärker nutzerzentrierten Kommunikationsunterstützung beigetragen werden, wodurch die Inklusionsfähigkeit der AAC-Systeme gesteigert und ein Beitrag für unterstützungsbedürftige Menschen geleistet werden könnte.

Literaturverzeichnis

Hinweis: Die hier angegebene Literatur wird in den Kapiteln 1 und 3 referenziert. In Kapitel 2 wird je Forschungsbeitrag die referenzierte Literatur im Anschluss an den Beitrag gelistet.

Abbasi, A., & Chen, H. (2008). CyberGate: A design framework and system for text analysis of computer-mediated communication. *MIS Quarterly*, 32(4), 811–837. https://doi.org/10.2307/25148859

Adeoye-Olatunde, O. A., & Olenik, N. L. (2021). Research and scholarly methods: Semi-structured interviews. *Journal of the american college of clinical pharmacy*, 4(10), 1358-1367.

Aguilar-Moreno, J. A., Palos-Sánchez, P. R., & Pozo-Barajas, R. (2024). Sentiment analysis to support business decision-making: A bibliometric study. *AIMS Mathematics, 9*(2), 4337–4375. https://doi.org/10.3934/math.2024223

Allen, J. (2005). Designing desirability in an augmentative and alternative communication device. *Universal Access in the Information Society*, 4(2), S. 135– 145.

Al-Zu'bi, Z. B. M., & Tsinopoulos, C. (2012). Suppliers versus lead users: Examining their relative impact on product variety. *Journal of Product Innovation Management*, 29(4), 667–680. https://doi.org/10.1111/j.1540-5885.2012.00932

American Speech Language Hearing Association (1991). Report: augmentative and alternative communication. *ASHA*, 33, 9–12.

Arora, K., Alwiyah, A., & Faisal, M. (2022). the use of data science in digital marketing techniques: work programs, performance sequences and methods. *Startupreneur Business digital*, 1(2), 143-155. https://doi.org/10.34306/sabda.v1i2.110

Ascari, R., Silva, L., & Pereira, R. 2019. Personalized interactive gesture recognition assistive technology. *Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems*.

Atkin, D. J., Hunt, D. S., & Lin, C. A. (2018). Diffusion theory in the new media environment: Toward an integrated technology adoption model. In *Advances in foundational mass communication theories* (pp. 225-252). Routledge.

Avetisyan, M., Schill, E., & Heuckeroth, R. (2015). Building a second brain in the bowel. *The Journal of Clinical Investigation*, 125(3), 899-907.

Bae, J., Havsol, J., Karpefors, M., Karlsson, A., & Mathiason, G. (2018). Short Text Topic Modeling to Identify Trends on Wearable Bio-Sensors in Different Media Type. In 2018 6th *International Symposium on Computational and Business Intelligence* (pp. 89-93). IEEE.

Basili, G., Caldiera, V. & Rombach, H. D. (1994). The goal question metric approach. *Encyclopedia of software engineering*, 528-532.

Bedrosian, J., Hoag, L., & McCoy, K. (2003). Relevance and speed of message delivery trade-offs in augmentative and alternative communication. *Journal of Speech, Language, and Hearing Research*, 46, 800–817.

Belk, R. 2013. Extended Self in a Digital World. *Journal of Consumer Research* (40:3).

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing Artificial Intelligence. *MIS Quarterly, 45*(3), iii–ix. https://doi.org/10.25300/MISQ/2021/15385

Bhor, H., Koul, T., Malviya, R., & Mundra, K. 2018. Digital Media Marketing Using Trend Analysis on Social Media. In *Proceedings of the 2nd International Conference on Inventive Systems and Control.*

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55 (4), 77-84.

Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 993-1022.

Bradshaw, J. (2013). The use of augmentative and alternative communication apps for the iPad, iPod and iPhone: an overview of recent developments. *Tizard Learning Disability Review*, 18(1), 31-37.

Brady, N. (2000). Improved comprehension of object names following voice output communication aid use: Two case studies. *Augmentative and Alternative Communication*, 16(3), 197-204.

Brandtzaeg, P. B., Haugstveit, I. M., Lüders, M., & Følstad, A. (2016). How should organizations adapt to youth civic engagement in social media? A lead user approach. *Interacting with Computers*, 28(5), 664–679. https://doi.org/0.1093/iwc/iwv041

Brem, A., Bilgram, V., & Gutstein, A. (2018). Involving lead users in innovation: A structured summary of research on the lead user method. *International Journal of Innovation and Technology Management*, 15(03). https://doi.org/10.1142/S0219 87701 85002 20

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., & et al. 2020. Language Models are Few-Shot Learners. arXiv:2005.14165.

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in numbers: How does data-driven decision-making affect firm performance? *Proceedings of the 32nd International Conference on Information Systems*, Shanghai.

Budde, M. 2012. Über Sprache reflektieren. *Kassel University Press.*

Cao, X., MacNaughton, P., Deng, Z., Yin, J., Zhang, X., & Allen, J. (2018). Using twitter to better understand the spatiotemporal patterns of public sentiment: a case study in Massachusetts, USA. *International Journal of Environmental Research and Public Health*, 15(2).

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly, 36*(4), 1165–1188. https://doi.org/10.2307/41703503

Cherenkov, E., Benga, V., Lee, M., Nandwani, N., Raguin, K., Sueur, M. C., ... & Sun, G. (2024). from machine learning algorithms to superior customer experience: business implications of machine learning-driven data analytics in the hospitality industry. *Journal of Smart Tourism*, 4(2), 5-14. https://doi.org/10.52255/smarttourism.2024.4.2.2

Churchill, R., & Singh, L. (2022). The evolution of topic modeling. *ACM Computing Surveys*, 54(10s), 1-35. https://doi.org/10.1145/3507900

Cirqueira, D., Pinheiro, M., Braga, T., Jacob Jr, A., Reinhold, O., Alt, R. & Santana, Á. (2017). Improving relationship management in universities with sentiment analysis and topic modeling of social media channels: learnings from ufpa. *Proceedings of the International Conference on Web Intelligence*, 998-1005.

Clarke, R. 1994. The Digital Persona and Its Application to Data Surveillance. *The information society*, 10(2), 77-92.

Constantiou, I. D., & Kallinikos, J. (2015). New games, new rules: Big data and the changing context of strategy. *Journal of Information Technology, 30*(1), 44–57. https://doi.org/10.1057/jit.2014.17

Cooper, H. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society*, 1(1).

Cooper, H., & Hedges, L. (1994). Research synthesis as a scientific enterprise. *The Handbook of Research Synthesis*, 285-299.

Cooper, R. G. (1996). Overhauling the new product process. *Industrial Marketing Management*, 6(25), 465–482. https://doi.org/10.1016/S0019-8501(96)00062-4

Costa, P., & McCrae, R. 1992. The Five-Factor Model of Personality and Its Relevance to Personality Disorders. *Journal of Personality Disorders* (6:4).

Crain, S. P., Zhou, K., Yang, S.-H. and Zha, H. (2012). Dimensionality reduction and topic modeling: From latent semantic indexing to latent dirichlet allocation and beyond. Mining text data. *Springer*, 129-161.

Cruz-Albrecht, L., Xu, J., Ni, K.-Y., & Lu, T.-C. (2017). Characterizing Regional and Behavioral Device Variations Across the Twitter Timeline: A Longitudinal Study. *Proceedings of the ACM Conference on Web Science*.

Debortoli, S., Müller, O., Junglas, I., & vom Brocke, J. (2016). Text mining for information systems researchers: An annotated topic modeling tutorial. *Communications of the Association for Information Systems, 39*(1), 110–135. https://doi.org/10.17705/1CAIS.03907

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and selfdetermination in human behavior. *New York: Plenum.*

Demoulin, N. T. M., & Coussement, K. (2020). Acceptance of text-mining systems: The signaling role of information quality. *Information & Management, 57*(1), 103120. https://doi.org/10.1016/j.im.2018.05.002

Dempster, M., Alm, N., & Reiter, E. (2010). Automatic generation of conversational utterances and narrative for Augmentative and Alternative Communication: a prototype system. *Proceedings of the NAACL HLT 2010 Workshop on Speech and Language Processing for Assistive Technologies*, 10-18.

Desai, T., Chow, K., Mumford, L., Hotze, F., & Chau, T. (2014). Implementing an iPad-based alternative communication device for a student with cerebral palsy and autism in the classroom via an access technology delivery protocol. *Computers & Education*, 79, 148-158.

Ding, K., Choo, W., Ng, K., Ng, S. (2020). Employing structural topic modelling to explore perceived service quality attributes in Airbnb accommodation. International Journal of Hospitality Management, 91.

Dodds, P., Harris, K., Kloumann, I., Bliss, C., & Danforth, C. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLOS One*, 6(12).

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of big data – Evolution, challenges and research agenda. *International Journal of Information Management, 48*, 63–71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021

Dwivedi, Y. K., Ismagilova, E., Hughes, D. L., Carlson, J., Filieri, R., Jacobson, J., ... & Wang, Y. (2021). Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management, 59*, 102168. https://doi.org/10.1016/j.ijinfomgt.2020.102168

Eickhoff, M., & Neuss, N. (2017). Topic Modelling Methodology: Its Use in Information Systems and Other Managerial Disciplines. *Proceedings of the 25th European Conference on Information Systems*.

Elsahar, Y., Hu, S., Bouazza-Marouf, K., Kerr, D., & Mansor, A. 2019. AAC Advances: A Review of Configurations for Individuals with a Speech Disability. *Sensors* (19:8).

Elspaß, S. 2018. Sprachvariation und Sprachwandel. Sprache in sozialen Gruppen, pp. 87–107.

Endo, Y., Toda, H., & Koike, Y. (2015). What's Hot in the Theme: Query Dependent Emerging Topic Extraction from Social Streams. *Proceedings of the 24th International Conference on the World Wide Web*.

Fan, W., & Gordon, M. D. (2014). The power of social media analytics. *Communications of the ACM, 57*(6), 74–81. https://doi.org/10.1145/2602574

Fast, E., & Horvitz, E. (2017). Long-term trends in the public perception of artificial intelligence. *Proceedings of the 31st AAAI Conference on Artificial Intelligence.*

Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering, 66*(1), 111–126. https://doi.org/10.1007/s12599-023-00834-7

Fritzell, B. (1996). Voice disorders and occupations. *Logopedics Phoniatrics Vocology*, 21(1), 7-12.

Gaglio, C., Kraemer-Mbula, E., & Lorenz, E. (2022). The effects of digital transformation on innovation and productivity: Firm-level evidence of South African manufacturing micro and small enterprises. *Technological Forecasting and Social Change*, 182, 121785.

Germann, F., Lilien, G. L., & Rangaswamy, A. (2020). Performance implications of deploying marketing analytics. *International Journal of Research in Marketing, 37*(1), 20–35. https://doi.org/10.1016/j.ijresmar.2019.06.001

Green, T. (2015). Places of inequality, places of possibility: mapping “opportunity in geography” across urban school-communities. *The Urban Review*, 47(4), 717-741. https://doi.org/10.1007/s11256-015-0331-z

Greenaway, K., Wright, R., Willingham, J., Reynolds, K., & Haslam, A. 2015. Shared Identity Is Key to Effective Communication. *Personality and Social Psychology Bulletin* (41:2).

Gregor, S. 2006. The nature of theory in information systems. MIS Quarterly (30:3).

Gregor, S., & Hevner, A. (2013). Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly*, 37(2), 337-355.

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv preprint arXiv:2203.05794.

Ha, T., Beijnon, B., Kim, S., Lee, S., & Kim, J. (2017). Examining User Perceptions of Smartwatch through Dynamic Topic Modeling. *Telematics and Informatics*, 34(7), 1262-1273.

Han, L. (2024). The rise of digital media: Transforming communication, culture, and commerce. *Global Media Journal*, 22(72). https://doi.org/10.36648/1550-7521.22.72.472

Hanna, R., Rohm, A., & Crittenden, V. (2011). We‘re all connected: The power of the social media ecosystem. *Business Horizons*, 54(3), 265-273.

Hau, Y. S., & Kang, M. (2016). Extending lead user theory to users’ innovation-related knowledge sharing in the online user community: The mediating roles of social capital and perceived behavioral control. *International Journal of Information Management*, 36(4), 520–530. https://doi.org/10.1016/j.ijinfomgt.2016.02.008

Heinrich, P., & Schwabe, G. (2014, May). Communicating nascent design theories on innovative information systems through multi-grounded design principles. *In International Conference on Design Science Research in Information Systems* (pp. 148-163). Cham: Springer International Publishing.

Hienerth, C., & Lettl, C. (2017). Perspective: Understanding the nature and measurement of the lead user construct. *Journal of Product Innovation Management*, 34(1), 3–12. https://doi.org/10.1111/jpim.12318

Hoag, L., Bedrosian, J., McCoy, K., & Johnson, D. (2004). Trade-offs between informativeness and speed of message delivery in augmentative and alternative communication. *Journal of Speech, Language, and Hearing Research*, 47, 1270– 1285.

Hong, L., Ahmed, A., Gurumurthy, S., Smola, A., & Tsioutsiouliklis, K. (2012). Discovering Geographical Topics in the Twitter Stream. *Proceedings of the 21st International Conference on the World Wide Web.*

Hourcade, J., Everhart Pilotte, T., West, E., & Parette, P. (2004). A history of augmentative and alternative communication for individuals with severe and profound disabilities. *Focus on Autism and other Developmental Disabilities*, 19(4), 235-244.

Houwen, J., Lucassen, P., Verwiel, A., Stappers, H., & Assendelft, S. (2019). Which difficulties do GPs experience in consultations with patients with unexplained symptoms: a qualitative study. *BMC Family Practice*, 20, 1-9.

Hung, C.-L., Chou, J.C.-L., & Dong, T.-P. (2011). Innovations and communication through innovative users: An exploratory mechanism of social networking website. *International Journal of Information Management*, 31(4), 317–326. https://doi.org/10.1016/j.ijinfomgt.2010.12.003

Inguanzo, I., Zhang, B., Zhang, B., Gil de Zúñiga, H., Gil de Zúñiga, H., & Gil de Zúñiga, H. (2021). Online cultural backlash? sexism and political user-generated content. *Information, Communication & Society*, 24(14), 2133–2152. https://doi.org/10.1080/1369118X.2021.1962940

Irawan, M., Wijayanto, R., Shahab, M., Hidayat, N., & Rukmi, A. (2020). Implementation of Social Media Mining for Decision Making in Product Planning Based on Topic Modeling and Sentiment Analysis. *Journal of Physics: Conference Series.*

Jeong, B., Yoon, J., & Lee, J.-M. (2019). Social Media Mining for Product Planning: A Product Opportunity Mining Approach Based on Topic Modeling and Sentiment Analysis. *International Journal of Information Management*, 48, 280-290.

Jiang, H., Zhang, X., Cao, X., Kabbara, J., & Roy, D. 2023. PersonaLLM: Investigating the Ability of GPT-3.5 to Express Personality Traits and Gender Differences. arXiv:2305.02547.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons, 53*(1), 59–68. https://doi.org/10.1016/j.bushor.2009.09.003

Kane, S. K., Morris, M. R., Paradiso, A., & Campbell, J. 2017. Understanding Self-Expression through Augmentative and Alternative Communication Devices. *CSCW*.

Kegler, M. C., Rigler, J., & Honeycutt, S. (2010). How does community context influence coalitions in the formation stage? a multiple case study based on the community coalition action theory. *BMC Public Health*, 10(1). https://doi.org/10.1186/1471-2458-10-90

Khan, H., Nasir, S., Nasim, K., Shabbir, D., & Mahmood, A. (2021). Twitter trends: a ranking algorithm analysis on real time data. *Expert Systems with Applications*, 164.

Kietzmann, J., Hermkens, K., McCarthy, I., & Silvestre, B. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, 54(3), 241-251.

Kirakowski, J., & Corbett, M. (1993). SUMI: The Software Usability Measurement Inventory. *British Journal of Educational Technology*, 24(3), 210–212.

Konadl, D., Wörner, J., Luttner, L., & Leist, S. (2023). Artificial intelligence in Augmentative and Alternative communication systems – a literature-based assessment and implications of different conversation phases and contexts. *Proceedings of the 31st European Conference on Information Systems*.

Lau, J., Collier, N., & Baldwin, T. (2012). On-Line Trend Analysis with Topic Models:# Twitter Trends Detection Topic Model Online. *Proceedings of the 2012 International Conference on Computational Linguistics*.

Laxmidas, K., Avra, C., Wilcoxen, C., Wallace, M., Spivey, R., Ray, W., Polsley, S., Kohli, P., Thompson, J., & Hammond, T. (2021). CommBo: Modernizing Augmentative and Alternative Communication. *International Journal of Human-Computer Studies*, 145.

Li, S., Liu, F., Zhang, Y., Zhu, B., Zhu, H., & Yu, Z. (2022). Text mining of user-generated content (UGC) for business applications in e-commerce: A systematic review. *Mathematics, 10*(19), 3554. https://doi.org/10.3390/math10193554

Li, X., Tse, Y. K., & Fastoso, F. (2024). Unleashing the power of social media data in business decision making: An exploratory study. *Enterprise Information Systems, 18*(1), Article 2243603. https://doi.org/10.1080/17517575.2023.2243603

Light, J., Barwise, A., Gardner, A., & Flynn, M. 2021. Personalized Early AAC Intervention to Build Language and Literacy Skills. *Topics in Language Disorders* (41:3).

Lim, K., Lee, K., Kendal, D., Rashidi, L., Naghizade, E., Winter, S., & Vasardani, M. (2018). The grass is greener on the other side: Understanding the effects of green spaces on Twitter user sentiments. *Proceedings of the 'The Web' Conference*.

Lin, Y.-R. (2014). Assessing sentiment segregation in urban communities. *Proceedings of the 2014 International Conference on Social Computing*.

Lozano, M., Schreiber, J., & Brynielsson, J. (2017). Tracking Geographical Locations Using a Geo-Aware Topic Model for Analyzing Social Media Data. *Decision Support Systems*, 99, 18-29.

Lu, H., Li, Y., Chen, M., Kim, H., & Serikawa, S. (2018). Brain intelligence: go beyond artificial intelligence. *Mobile Networks and Applications*, 23(2), 368-375.

Ma, L., & Sun, B. (2020). Machine learning and AI in marketing – Connecting computing power to human insights. *International Journal of Research in Marketing, 37*(3), 481–504. https://doi.org/10.1016/j.ijresmar.2020.04.005

Malik, N., & Bilal, M. (2024). Natural language processing for analyzing online customer reviews: A survey, taxonomy, and open research challenges. *PeerJ Computer Science, 10*, e2203. https://doi.org/10.7717/peerj-cs.2203

Mamlekar, C., Dietz, A., Wendt, O. 2024. History and Evolution of AAC. *Principles and Practices in AAC*.

Martínez-Torres, M. R. (2013). Analysis of open innovation communities from the perspective of social network analysis. *Technology Analysis & Strategic Management*, 26(4), 435–451. https://doi.org/10.1080/09537325.2013.851378

Mastrogiacomo, L., Barravecchia, F., Franceschini, F., & Marimon, F. 2021. Mining quality determinants of product-service systems from user-generated contents. *Quality Engineering* (33:3), pp. 425-442.

Mayring, P. 2004. Qualitative Content Analysis. A companion to qualitative research.

Mayring, P., & Fenzl, T. (2019). Qualitative Inhaltsanalyse (pp. 633-648). *Springer Fachmedien Wiesbaden*.

Mir, I. A., & Ur Rehman, K. (2013). Factors affecting consumer attitudes and intentions toward user-generated product content on YouTube. *Management & Marketing*, 8(4).

Mirtalaie, M., Hussain, O., Chang, E., & Hussain, F. (2018). Extracting sentiment knowledge from pros/cons product reviews: Discovering features along with the polarity strength of their associated opinions. *Expert Systems with Applications*, 114, 267-288.

Moro, S., Lopes, R., Esmerado, J., & Botelho, M. (2020). Service quality in airport hotel chains through the lens of online reviewers. *Journal of Retailing and Consumer Services*, 56.

Neuberger, C., Bartsch, A.-M., Fröhlich, R., Hanitzsch, T., Reinemann, C., & Schindler, J. (2023). The digital transformation of knowledge order: a model for the analysis of the epistemic crisis. *Annals of the International Communication Association*, 47(2), 180–201. https://doi.org/10.1080/23808985.2023.2169950

Obiorah, M., Piper, A., & Horn, M. (2021). Designing AACs for people with aphasia dining in restaurants. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1-14.

Pande, P., Neuman, R., & Cavanagh, R. (2014). The Six Sigma Way: How to maximize the impact of your change and improvement efforts. *McGraw Hill Professional.*

Park, E., Kwon, J., & Kim, S. 2021. Green marketing strategies on online platforms: a mixed approach of experiment design and topic modeling. *Sustainability* (13:8).

Pew Research Center. (2025, April 3). Views of risks, opportunities, and regulation of AI. https://www.pewresearch.org/internet/2025/04/03/views-of-risks-opportunities-and-regulation-of-ai/

Phillips, J. (2012). Social media analytics. Win With Advanced Business Analytics, 247-269. https://doi.org/10.1002/9781119205371.ch12

Phillips-Wren, G., & Virvou, M. (2025). Issues and trends in generative AI technologies for decision making. *Intelligent Decision Technologies, 16*(1), 5–17. https://doi.org/10.1177/18724981251320551

Plunz, R., Zhou, Y., Vintimilla, M., Mckeown, K., Yu, T., Uguccioni, L., & Sutto, M. (2019). Twitter sentiment in New York City parks as measure of well-being. *Landscape and Urban Planning*, 189, 235-246.

Pollak, V., & Gallagher, B. (1989). A fast communication aid for non-verbal subjects with severe motor handicaps. *Journal of Medical Engineering & Technology*, 13(1-2), 23-27.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data, 1*(1), 51–59. https://doi.org/10.1089/big.2013.1508

Radaideh, A. and Dweiri, F. (2024). Use of ai applications to learn the sentiment polarity of public perceptions: a case study of the covid-19 vaccinations in the uae. *International Journal of Computing and Digital Systems*, 15(1), 1191-1199. https://doi.org/10.12785/ijcds/160187

Reisenbichler, M., and Reutterer, T. 2019. Topic modeling in marketing: recent advances and research opportunities. *Journal of Business Economics* (89), pp. 327-356.

Rogers, E. (2003). Diffusion of Innovations. Fifth edition. *Free Press: New York.*

Roosendaal, A. 2010. Digital Personae and Profiles as Representations of Individuals. *Springer.*

Russell, S., Dewey, D., & Tegmark, M. (2015). Research priorities for robust and beneficial artificial intelligence. *AI Magazine*, 36(4), 105-114.

Schaarschmidt, M., Stol, K.-J., Walsh, G., & Bertram, M. (2019). Lead Users' Innovative Work Behavior in Digital Platform Ecosystems: A Large Scale Study of App Developers. *Proceedings of International Conference on Information Systems*.

Schepis, M., Reid, D., & Behrman, M. (1996). Acquisition and functional use of voice output communication by persons with profound multiple disabilities. *Behavior Modification*, 20(4), 451- 468.

Schouten, K., & Frasincar, F. (2015). Survey on aspect-level sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering*, 28(3), 813-830.

Schurig, D., Riemer, K., & Schlagwein, D. (2022). Aspect-based sentiment analysis for the analysis of user-generated content: A framework and research agenda. *Information Systems Journal, 32*(5), 825–862. https://doi.org/10.1111/isj.12350

Schwartz, A., Dodds, P., O'Neil-Dunne, J., Danforth, C., & Ricketts, T. (2019). Visitors to urban greenspace have higher sentiment and lower negativity on Twitter. *People and Nature*, 1(4), 476-485.

Serapio-Garcia, G., Safdari, M., Crepy, C., Fitz, S., Romero, P., Sun, L., Abdulhai, M., Faust, A., & Mataric, M. 2023. Personality Traits in Large Language Models. arXiv:2307.00184.

Shen, H., Ju, Y., & Zhu, Z. (2023). Extracting useful emergency information from social media: a method integrating machine learning and rule-based classification. *International Journal of Environmental Research and Public Health*, 20(3), 1862. https://doi.org/10.3390/ijerph20031862

Sims, C. 2017. Do the Big-Five Personality Traits Predict Empathic Listening and Assertive Communication?. *International Journal of Listening* (31:3).

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research, 70*, 263–286. https://doi.org/10.1016/j.jbusres.2016.08.001

Sorokovikova, A., Rezagholi, S., Fedorova, N., & Yamshchikov, I. 2024. LLMs Simulate Big5 Personality Traits: Further Evidence. *Proceedings of 1st Workshop on Personalization of Generative AI Systems*.

Stieglitz, S., Dang-Xuan, L., Bruns, A., & Neuberger, C. (2014). Social media analytics: An interdisciplinary approach and its implications for information systems. *Business & Information Systems Engineering, 6*(2), 89–96. https://doi.org/10.1007/s12599-014-0315-7

Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics–Challenges in topic discovery, data collection, and data preparation. *International journal of information management*, 39, 156-168.

Todman, J., & Alm, N. (1997). Pragmatics and AAC approaches to conversational goals. *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics*.

Tuarob, S., & Tucker, C. (2015). Quantifying product favorability and extracting notable product features using large scale social media data. *Journal of Computing and Information Science in Engineering*, 15(3).

Tucker, C., & Kim, H. (2011). Predicting Emerging Product Design Trend by Mining Publicly Available Customer Review Data. *Proceedings of the 18th International Conference on Engineering Design.*

Tuunanen, T., Bragge, J., Haivala, J., Hui, W., & Virtanen, V. (2011). A method for recruitment of lead users from virtual communities to innovate it enabled services for consumers in global markets. Pacific Asia *Journal of the Association for Information Systems*, 3(2), 31–56. https://doi.org/10.17705/1pais.03202

Uchoa, J. P., Falcão, T. P., Nascimento, A. C., Miranda, P. B., & Mello, R. F. (2021, July). Fostering Autonomy through Augmentative and Alternative Communication. In *2021 International Conference on Advanced Learning Technologies*, (pp. 320-324).

Vayansky, I. & Kumar, S. A. (2020). A review of topic modeling methods. *Information Systems*, 94, 101582.

Vanderheiden, G. (1983). Non-Conversational Communication Technology Needs of Individuals with Handicaps. *Rehabilitation World*, 7(2), 8-12.

Verma, S., & Yadav, N. (2021). Past, present, and future of electronic word of mouth (eWOM). *Journal of Interactive Marketing, 53*, 111–128. https://doi.org/10.1016/j.intmar.2020.07.001

Vo, A., Nguyen, Q., & Ock, C. (2018). Opinion–aspect relations in cognizing customer feelings via reviews. IEEE Access, 6, 5415-5426.

Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the association for information systems*, 37(1), 9.

Von Hippel, E. (1986). Lead users: A source of novel product concepts. *Management Science*, 32(7), 791–805. https://doi.org/10.1287/mnsc.32.7.791

Von Hippel, E. (1994). “Sticky information” and the locus of problem solving: Implications for innovation. *Management science*, 40(4), 429-439

Webster, J., & Watson, R. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, 26(2), 13-23.

Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing, 80*(6), 97–121. https://doi.org/10.1509/jm.15.0413

Wilson, S. R., & Mihalcea, R. (2019). Predicting Human Activities from User-Generated Content. *Meeting of the Association for Computational Linguistics*, 2572–2582. https://doi.org/10.18653/V1/P19-1245

Wong, R. Y. M., & Ngai, E. W. T. (2024). From information to intelligence: Transforming social media big data into business decision-making. *Decision Support Systems, 175*, 114023. https://doi.org/10.1016/j.dss.2023.114023

Wu, X., Nguyen, T., & Luu, A. T. (2024). A survey on neural topic models: methods, applications, and challenges. *Artificial Intelligence Review*, 57(2), 18.

Wörner, J., Konadl, D., Schmid, I., & Leist, S. (2021). Comparison of Topic Modelling Techniques in Marketing - Results From an Analysis of Distinctive Use Cases. *Proceedings of the 29th European Conference on Information Systems*.

Wörner, J., Konadl, D., Schmid, I., & Leist, S. (2022). Supporting Product Development by a Trend Analysis Tool Applying Aspect-Based Sentiment Detection. *Proceedings of the 17th International Conference on Design Science Research in Information Systems and Technology*.

Wörner, J., Konadl, D., Leist, S., & Schmid, I. (2023). MANTRA: A Topic Modeling- Based Tool to Support Automated Trend Analysis on Unstructured Social Media Data. *Proceedings of the 44th International Conference on Information Systems*.

Yan, Z., Xing, M., Zhang, D., & Ma, B. (2014). A Context-Dependent Sentiment Analysis of Online Product Reviews Based on Dependency Relationships. *Proceedings of the 35th International Conference on Information Systems*.

Ye, H., & Kankanhalli, A. (2018). User Service Innovation on Mobile Phone Platforms: Investigating Impacts of Lead Userness, Toolkit Support, and Design Autonomy. *MIS quarterly*, 42(1). 165–187. https://doi.org/10. 25300/MISQ/2018/12361

Zhang, P., Gu, H., Gartrell, M., Yang, D., Ding, X., & Gu, N. (2016). Group-based Latent Dirichlet Allocation (Group-LDA): Effective audience detection for books in online social media. *Knowledge-Based Systems*, 105, 134-146.

Zhong, N., & Schweidel, D. (2020). Capturing Changes in Social Media Content: A Multiple Latent Changepoint Topic Model. *Marketing Science*, 39(4), 827-846.