

# 1 Einleitung und Motivation

## 1.1 Ausgangssituation

Die von Unternehmen angebotenen Dienstleistungen, Services und Produkte sind in den letzten Jahren in vielen Bereichen austauschbar bzw. ersetzbar geworden. Um sich dennoch von der Konkurrenz abzuheben und sich einen strategischen Wettbewerbsvorteil zu sichern, setzen Unternehmen vermehrt auf flexiblen und adaptiven Kundenservice, Schnelligkeit und Termintreue zu günstigen Preisen [WASSERMANN 2013, S. 1 f.]. Verkürzte Produktlebenszyklen, schwankendes Kundenverhalten und die Notwendigkeit, sofort auf Marktschwankungen zu reagieren, sind nur einige der Herausforderungen, die in diesem Zusammenhang beherrscht werden müssen. Durch die steigende Dynamisierung und Globalisierung der Märkte wird es jedoch zunehmend schwieriger, Bedarfe und deren Schwankungen innerhalb der Lieferketten zu prognostizieren. KUHN UND HELLINGRATH formulieren diese Situation der Wirtschaft wie folgt:

*„Wir haben es heute mit einem globalem Wettbewerb zu tun, einer Internationalisierung der Technikentwicklung und Forschung, immer kürzeren Produktlebenszyklen, (stark) schwankender differenzierter Nachfrage, kurzen Reaktionszeiten, kürzesten Durchlaufzeiten, kleinsten Lagerbeständen, um die Nachfrage zu decken, geringsten Lieferstückkosten und Preisen, die die Kunden zufrieden stellen.“* [HELLINGRATH & KUHN 2013, S. 4]

Diese Unsicherheiten stellen die Bedarfsplanung innerhalb der Supply Chain (SC) vor große Herausforderungen, was sich auf die nachgelagerten Prozesse auswirkt, da diese als taktische und strategische Planungsgröße einen entscheidenden Einfluss auf die Effizienz und den Erfolg eines Unternehmens hat und andere Planungen (Distributionsplanung, Produktionsplanung, Beschaffungsplanung und Netzwerkplanung) beeinflusst [vgl. HELLINGRATH & KUHN 2013, S. 144 ff.]. Diese anhaltende Entwicklung ist branchenübergreifend zu beobachten und wird besonders im Bereich der Konsumgüterindustrie festgestellt [vgl. NEMTAJELA & MBOHWA 2017, S. 699 f.]. Durch die enge Verbindung von Konsumgüterherstellern und Einzelhandel verfügen die Hersteller nur über eine geringe Anzahl an Kunden, sodass sich Bedarfsschwankungen der einzelnen Kunden nicht über die Gesamtmenge der Nachfrage glätten. Zusätzlich erschweren Faktoren, wie z. B. ein Mindesthaltbarkeitsdatum (MHD), des Produkts eine Vorproduktion, um Bedarfsschwankungen mithilfe eines aufgebauten Bestands zu regulieren [vgl. BARBOSA et al. 2015, S. 7137].

Um kurzfristige Änderungen und umweltbedingte Schwankungen in der Lieferkette zu reduzieren bzw. zu berücksichtigen, ist es entscheidend, effiziente Prognosemodelle zu implementieren. Diese ermöglichen es den Unternehmen, sich auf zukünftige Situationen frühzeitig vorzubereiten und entsprechende Planungen anzupassen. Sowohl quantitative als auch qualitative Prognosemodelle sind jedoch aufgrund der zunehmenden Marktschwankungen nicht bzw. nur bedingt geeignet, um eine Prognose in ausreichender Qualität zu erstellen. Nach ERBEN UND ROMEIKE fördern unter anderem die zunehmende Datenmenge (Informationsvielfalt erschwert die genaue Identifikation von Kausalketten) und ein immer komplexeres Geschäftsumfeld die Unsicherheiten im Bereich der Prognosen. Dies führt über die einzelnen Planungsschritte zu

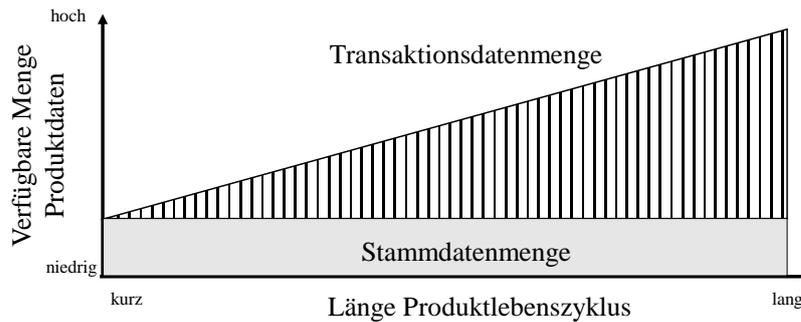
ineffizienten Lieferketten und zum Anstieg des Lagerbestands zur Beherrschung der Unsicherheiten. [vgl. ERBEN & ROMEIKE 2003, S. 43 ff.]

Die Unsicherheiten sind zum Teil umsatzbedingt bzw. marktbedingt und resultieren oft aus mangelndem Wissen oder falschen Informationen. Weiterhin sind häufig wechselnde Sortimente, neu auftretende Trends und schwankende Nachfrage aufgrund verschiedener störender Einflussfaktoren (Saisonalität, Preis- und Sortimentspolitik etc.) nur einige von vielen Punkten, die den Einsatz herkömmlicher statistischer Prognosemodelle erschweren, da komplexe Wechselbeziehungen den Bedarf beeinflussen. Das hat zur Folge, dass der tatsächliche Zusammenhang zwischen den Einflussfaktoren und dem Bedarfsverlauf für den Menschen oft nur schwer oder gar nicht erkennbar ist, was zu einem verstärkten Wunsch nach und steigendem Einsatz von Methoden der künstlichen Intelligenz (KI) in der Bedarfsplanung führt [vgl. DU et al. 2020, S. 1477 ff.].

Um trotz hoher Marktanforderungen und wechselnder Umfeldeinflüsse eine präzise Prognose zu erhalten, werden zunehmend Methoden des maschinellen Lernens (ML) oder des Deep Learning (DL) eingesetzt. Beide Methodentypen können als Teilbereiche bzw. Unterkategorie der künstlichen Intelligenz definiert werden. Vorteile von KI-basierten Methoden sind die Beherrschung der hohen Datenvielfalt und das Finden von Datenmustern und Abhängigkeiten in den Eingangsdaten, um diese für die anschließende Prognose des Zielwerts zu nutzen. [vgl. AMPAZIS 2015, S. 57 f.] Wie bei den statistischen Methoden gibt es auch bei KI-basierten Methoden keine allgemeingültige Methode, die für jede Situation oder jeden Anwendungsfall eine verbesserte Prognose generiert. Vielmehr können mit jeder Methode je nach Rahmenbedingung unterschiedliche Qualitäten erreicht werden [vgl. CARBONNEAU et al. 2008, S. 1153].

## **1.2 Problemstellung**

Um Machine-Learning-(ML-)Verfahren effizient in den Bedarfsplanungsprozess zu integrieren ist es notwendig, auf eine ausreichende Datenbasis zurückzugreifen. Die notwendige Ausprägung der Datenbasis ist individuell in Abhängigkeit vom betrachteten Anwendungsfall zu prüfen. Durch den Trend zu kürzeren Produktlebenszyklen gibt es aber oftmals gar nicht mehr die Möglichkeit, eine ausreichende Datenbasis zu generieren, um Machine-Learning-Verfahren erfolgreich für Prognoseaufgaben einzusetzen. Es herrscht somit eine gegenläufige Entwicklung (siehe Abbildung 1) zwischen zur Verfügung stehender Datenbasis (aufgeteilt in Stamm- und Transaktionsdaten) und Länge des Produktlebenszyklus. [Vgl. AMERI & DUTTA 2005, S. 584.]

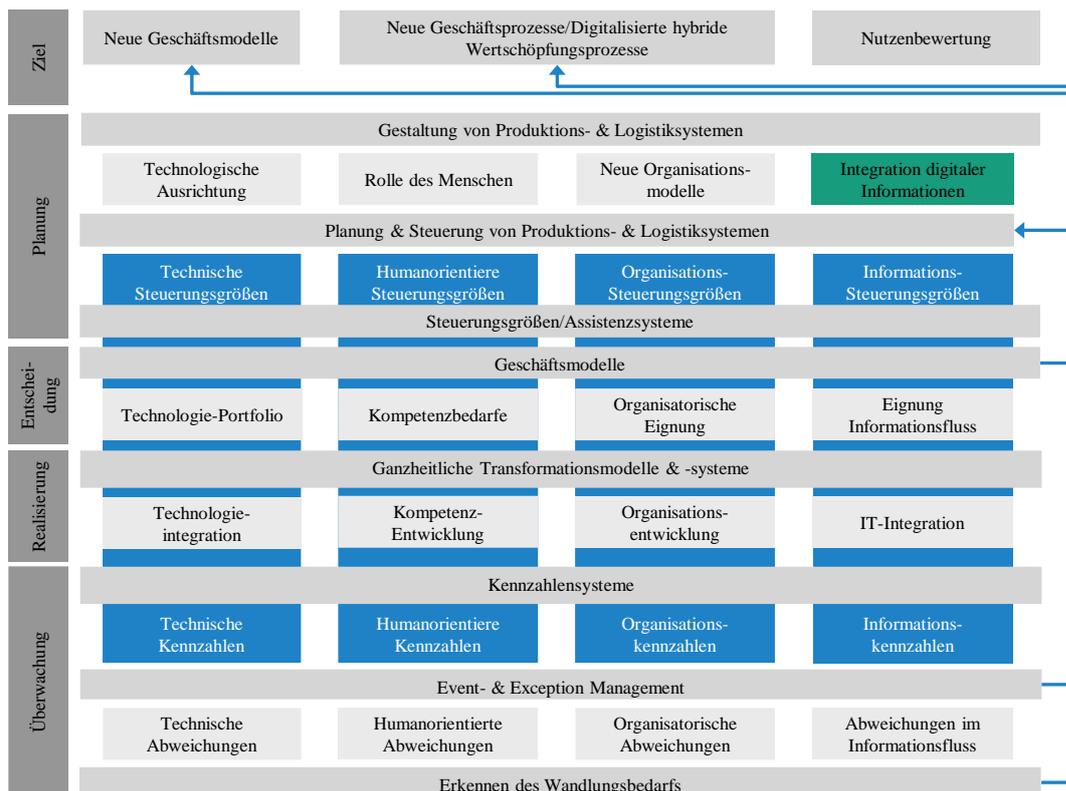


**Abbildung 1: Gegenläufige Entwicklung zwischen verfügbarer Datengrundlage und Länge des Produktlebenszyklus (eigene Darstellung)**

Die ML-basierten Prognosemodelle zur Beherrschung der Marktdynamisierung brauchen eine valide Datenbasis, und die immer kürzer werdenden Produktlebenszyklen erschweren die Bereitstellung der notwendigen Daten. Zusätzlich korrelieren Bedarfsverläufe häufig mit weiteren Einflussfaktoren wie einem saisonalen Faktor, sodass sich die Bedarfe innerhalb eines Jahreszyklus anpassen und variieren können. Produkte mit einem Produktlebenszyklus, der kleiner als ein Jahr ist, haben somit keine Möglichkeiten, die über das Jahr verteilten zeitlichen Schwankungen als Bedarfsinformationen zu interpretieren und aufzunehmen. Zusätzlich zu den produkteigenen Informationen müssen für die Prognose des Anlaufverhaltens auch marktspezifische Einflüsse mitberücksichtigt werden, die das Anlaufverhalten eines Produkts erheblich beeinflussen können. Machine-Learning-Verfahren können solche Trends berücksichtigen und als zusätzliche Dateninformation (Daten-Feature) in das Prognosemodell integrieren. Grundvoraussetzung für den Einsatz von Machine-Learning-Verfahren ist jedoch eine ausreichende Datengrundlage, deren Bereitstellung für den einzelnen Artikel in Zukunft immer herausfordernder wird.

Vergleicht man verschiedene Umfragen und Veröffentlichungen zu den Herausforderungen bei der Implementierung von KI- und Big-Data-Technologien, so lassen sich unterschiedliche Gemeinsamkeiten feststellen. Es existieren Kernherausforderungen aus den Bereichen der strategie- und unternehmensbezogenen, der technologie- und datenbezogenen sowie der menschenbezogenen Ebene. Primär geht es innerhalb dieser Arbeit um die technologische und die datenbezogene Ebene und somit um Herausforderungen wie z. B. Komplexität der Daten, unzureichende Datenqualität, unzureichender Datenzugang/Datenaustausch zwischen Unternehmen, begrenzte technologische Fähigkeiten oder Bereitschaft der IT-Infrastruktur und Mangel an verfügbaren Tools. Der Schwerpunkt ist dabei besonders die Bereitstellung einer ausreichenden Datengrundlage für den Einsatz einer KI-Methode. [Vgl. MOKTADIR et al. 2019, S. 5 f.; vgl. EAGER et al. 2020, S. 26 f.; vgl. ALSHEIBANI et al. 2019, S. 5 ff.]

Die formulierte Problemstellung ist auch im Dortmunder Management Modell der Industrie 4.0 wiederzufinden. Grundsätzlich wurde das Modell entwickelt, um Unternehmen bei der digitalen Transformation zu unterstützen. Es untergliedert sich in die Bereiche Überwachung, Realisierung, Entscheidung, Planung und Ziel. Die Nutzbarmachung von Machine-Learning-Modellen in der Bedarfsprognose lässt sich der Kategorie Planung zuordnen und beschäftigt sich mit dem Baustein der Integration digitaler Informationen (grün markiert in Abbildung 2). [Vgl. MICHALIK et al. 2019, S. 2327.]



**Abbildung 2: Einordnung der Problemstellung in das Dortmunder Management Modell**  
[vgl. HENKE et al. 2018, S. 561]

Nach der Beschreibung der führenden Problemstellung der Arbeit wird im nächsten Abschnitt genauer auf die konkrete Zielsetzung eingegangen.

### 1.3 Zielsetzung

Mit dem Fokus auf die beschriebene Problemstellung besteht das wesentliche Ziel der vorliegenden Arbeit darin, ein systematisches Vorgehensmodell zur Generierung einer Datenbasis für die Befähigung von Machine-Learning-Ansätzen in der Bedarfsprognose von Anlaufprodukten zu entwickeln:

- **Systematisches Vorgehensmodell** beschreibt das Zielbild des zu entwickelnden Prozesses zur Bereitstellung einer Datenbasis für die Trainingsphase eines Machine-Learning-Algorithmus.
- **Generierung einer Datenbasis** drückt aus, dass auf Basis der gesammelten Informationen innerhalb des systematischen Vorgehensmodells ein eigener Datensatz für das zu betrachtende Anlaufprodukt erstellt wird.
- **Befähigung** beschreibt den Zustand, dass Machine-Learning-Modelle für die Bedarfsprognose eingesetzt werden können und ein ähnlich gutes Ergebnis erzielen wie Machine-Learning-Modelle bzw. konventionelle Prognoseverfahren, die auf realen Daten in ausreichender Menge trainiert worden sind.
- **Anlaufprodukte** fasst die zu betrachtende Produktart zusammen. Es geht primär um Produkte mit unzureichender bzw. ohne Datenbasis, die besonders in der Produkteinführungsphase zu finden sind.

Das systematische Vorgehensmodell soll Unternehmen dazu befähigen, Produkte mit unzureichender Datenbasis in der Produktanlaufphase mithilfe von KI-Ansätzen zu prognostizieren und somit statistische Verfahren in der Prognose abzulösen. Um die beschriebenen Forschungsziele zu erreichen, wird innerhalb dieser Arbeit folgende forschungsleitende Fragestellung analysiert und das erarbeitete Vorgehensmodell anhand zweier Anwendungsfälle evaluiert:

*Wie können Machine-Learning-Ansätze in der Bedarfsprognose von Anlaufprodukten vor dem Hintergrund ihrer spezifischen Anforderungen an die Datenverfügbarkeit und -qualität bei kürzer werdenden Produktlebenszyklen im industriellen Kontext zur Anwendung gebracht werden?*

Zur Detaillierung und systematischen Aufarbeitung wird die zentrale Forschungsfrage in weitere Teilforschungsfragen (F1 bis F4) untergliedert, die in der folgenden Tabelle vorgestellt werden.

**Tabelle 1: Detaillierung der leitenden Forschungsfrage**

F1	Welche charakteristischen Eigenschaften muss ein Datensatz aufweisen, um ihn in der KI-basierten Bedarfsprognose einsetzen zu können?
F2	Welche Möglichkeiten gibt es, Anlaufprodukte zu klassifizieren, um Informationsquellen als Basis für das Anlaufverhalten eines Produkts systematisch aufzunehmen?
F3	Wie kann eine Zeitreihe systematisch erzeugt werden, um als Grundlage für das Training eines KI-basierten Prognosemodells eingesetzt zu werden?
F4	Welchen Mehrwert kann ein systematisch generierter Trainingsdatensatz für den frühzeitigen Einsatz einer ML-basierten Bedarfsprognose in der Anlaufphase eines Produkts schaffen?

Um eine Evaluierung des entwickelten Vorgehens zur Beantwortung der Forschungsfrage F4 zu ermöglichen, wird das systematische Vorgehensmodell beispielhaft anhand zweier Unternehmen aus dem Bereich der Konsumgüterindustrie durchgeführt und getestet. Die Erkenntnisse anhand der realen Daten fließen anschließend in die Beantwortung der vierten Hauptforschungsfrage.

## 1.4 Definition des Untersuchungsbereichs

Basierend auf der beschriebenen Ausgangssituation, Problemstellung und Zielsetzung wird im Folgenden eine erste Definition des Untersuchungsbereichs mithilfe eines morphologischen Kastens (Methode, um komplexe Problembereiche vollständig zu erfassen und alle möglichen Lösungen systematisch aufzunehmen) durchgeführt. Bei der Methodik wird mithilfe der Definition von Merkmalen und entsprechenden Ausprägungen systematisch eine Matrix aufgebaut, auf deren Basis unterschiedliche Kombinationsmöglichkeiten charakterisiert werden können [vgl. ZWICKY 1949, 4 ff.].

1. Definition der Merkmale (z. B. Planungsebene)
2. Mögliche Ausprägungen der Merkmale (z. B. exekutiv, operativ, taktisch, strategisch)
3. Auswahl mindestens einer Ausprägung jedes Merkmals (z. B. Planungsebene = operativ)

Der Untersuchungsraum wird mithilfe der drei Kernthemen (Bedarfsplanung, Künstliche Intelligenz und Produktbetrachtung) der Forschungsarbeit charakterisiert (siehe Abbildung 3).

Bei der Bedarfsplanung liegt der Fokus auf der verbrauchsgebundenen Bedarfsermittlung mit operativem/taktischen Planungshorizont. Die zu untersuchende KI-Ausprägung für den Einsatz im Prognosemodell wird an dieser Stelle der Arbeit noch nicht weiter spezifiziert, um auf Basis des folgenden Kapitels die Vor- und Nachteile der jeweiligen Verfahren in Bezug auf die zu lösende Problemstellung der Arbeit zu beurteilen und gegenüberzustellen. Bei dem betrachteten Produktspektrum wird es sich in dieser Arbeit um Anlaufprodukte aus dem Bereich der Verbrauchsgüter im Business-to-Business-(B2B)-Bereich handeln. Das Anwendungsfeld Business-to-Customer (B2C) wird innerhalb dieser Arbeit nicht weiter berücksichtigt. Konkret soll innerhalb der Forschungsarbeit das entwickelte Vorgehensmodell exemplarisch anhand von zwei Fallbeispielen aus dem Konsumgüterbereich evaluiert werden. Generell soll eine allgemeine Anwendbarkeit geschaffen werden, sodass branchenunabhängig Dateninformationen für die KI-basierte Prognose von Anlaufprodukten bzw. Produkten mit unzureichender Datenlage bereitgestellt werden können.

	Bereich/Charakteristik	Ausprägungen			
Bedarfsplanung	Planungsebene	exekutiv	operativ	taktisch	strategisch
	Bedarfsermittlung	programmgebunden		verbrauchgebunden	
KI	KI Ausprägung	Machine Learning		Deep Learning	
Produktbetrachtung	Produktlebenszyklus	Anlaufprodukt	Hochprodukt	Auslauf	etc.
	Klassifizierungsmerkmal Produktart	Gebrauchsgüter		Verbrauchsgüter	Dienstleistungen
	Kundenspektrum	Privatkunden (B2C)	Gemischter Kundenstamm (B2C u. B2B)		Geschäftskunden (B2B)

	wird nicht betrachtet
	wird betrachtet

**Abbildung 3: Erste Definition des Untersuchungsbereichs (eigene Darstellung)**

Nach der Kurzeinführung in den Untersuchungsbereich der Arbeit sollen im folgenden Abschnitt noch der Aufbau und die gewählte Forschungsmethodik der Arbeit spezifiziert werden.

## 1.5 Aufbau und Forschungsmethodik der Arbeit

Zur Beantwortung der in Abschnitt 1.3 skizzierten Fragestellungen orientieren sich der Aufbau und die Forschungsmethodik der Forschungsarbeit am Forschungsprozess der angewandten Wissenschaften (Angewandte Forschung) nach ULRICH aus dem Jahre 1981. Die vorliegende Arbeit ist dem Bereich der konzeptionellen Forschung zur Entwicklung eines praxisrelevanten Modells bzw. Konzepts zuzuordnen. Während bei der Grundlagenforschung (explorative bzw. theoretische Forschung) die Fragestellungen eng mit dem theoriebeeinflussten Bereich der Wissenschaft verknüpft sind, leiten sich die Fragestellungen in der angewandten Wissenschaft aus real existierenden Problemstellungen ab (hier: die fehlenden Datengrundlagen für den Einsatz von Machine Learning Methoden in der Produkthanlaufphase) [vgl. ULRICH 1981, S. 19 ff.]. Der Forschungsschwerpunkt der Arbeit ist somit die lösungsorientierte Kombination bzw. Adaption von vorhandenem Wissen und Methodiken.

Da das Vorgehensmodell auf getroffenen Theorien und Thesen aufbaut, die anhand von einzelnen Experimenten geprüft werden, hat die vorliegende Arbeit einen experimentellen deduktiven Forschungsaufbau. Im Gegensatz zu einem induktiven Forschungsaufbau, bei dem man zielgerichtet und basierend auf Daten oder Beobachtungen Theorien ableiten würde, stützt sich die vorliegende Arbeit auf getätigte Hypothesen, um diese anschließend mit Experimentiererergebnissen zu bestärken und einen allgemeinen Erkenntnisgewinn für den systematischen Aufbau einer Datengrundlage zu ermöglichen [vgl. REINDERS & DITTON 2011, S. 44 ff.]. Zum Aufbau der wissenschaftlichen Grundlagen wird zu Beginn die Methode der systematischen Literaturrecherche eingesetzt, um darauf aufbauend ein systematisches Vorgehensmodell zu entwickeln und die aufgestellten Thesen anhand von unterschiedlichen Anwendungsfällen zu prüfen und zu validieren.

Innerhalb der Einleitung wurde das zu untersuchende praxisrelevante Problem der Datenbereitstellung für eine KI-basierte Bedarfsprognose in der frühen Phase des Produktlebenszyklus bereits erfasst und beschrieben und die Zielsetzung der vorliegenden Arbeit wurde abgeleitet. Darauf aufbauend wurden die Forschungs- bzw. Teilforschungsfragen formuliert und eine erste Definition des Untersuchungsbereichs durchgeführt.

Im Kapitel 2 werden zunächst wesentliche theoretische und praktische Grundlagen aus dem Bereich der Bedarfsplanung, der KI und der Produktklassifizierung erläutert, relevante Begriffe definiert und der Untersuchungsbereich weiter konkretisiert. Ziel des Kapitels ist die Durchführung der grundlegenden Informationsbeschaffung als ein entscheidender Bestandteil des von ULRICH beschriebenen Forschungsprozesses [vgl. ULRICH 1981, S. 19]. Das Kapitel 3 beschreibt den aktuellen Forschungsstand, der mithilfe einer detaillierten systematischen Literaturrecherche erfasst wird. Es werden Anforderungen bzgl. der Ausprägung des Datensatzes formuliert und Ansätze für die systematische Produktklassifizierung und bereits existierende Verfahren für die Datenbereitstellung vorgestellt. Aufbauend auf den gewonnenen Erkenntnissen wird der sich ergebende Forschungsbedarf vorgestellt und als Kernelement für das zu entwickelnde Lösungskonzept identifiziert. Im weiteren Verlauf der Arbeit wird der identifizierte Forschungsbedarf als Rahmenwerk genutzt, um gezielt die ausgewählten Forschungsfragen zu beantworten.

In Kapitel 4 wird der konzeptionelle Rahmen des Lösungsansatzes vorgestellt und auf die einzelnen Bausteine innerhalb des Vorgehensmodells zur Beantwortung der Forschungsfragen eingegangen. Die darauffolgenden drei Kapitel befassen sich mit der Entwicklung, der Implementierung und der Evaluierung des Vorgehensmodells und gehen somit auf die Forschungsfragen F2 bis F4 ein. Abbildung 4 fasst das Vorgehen innerhalb der vorliegenden Arbeit zusammen und konkretisiert die Ausgestaltung der Rückkopplungsschleifen zwischen den einzelnen Kapiteln.

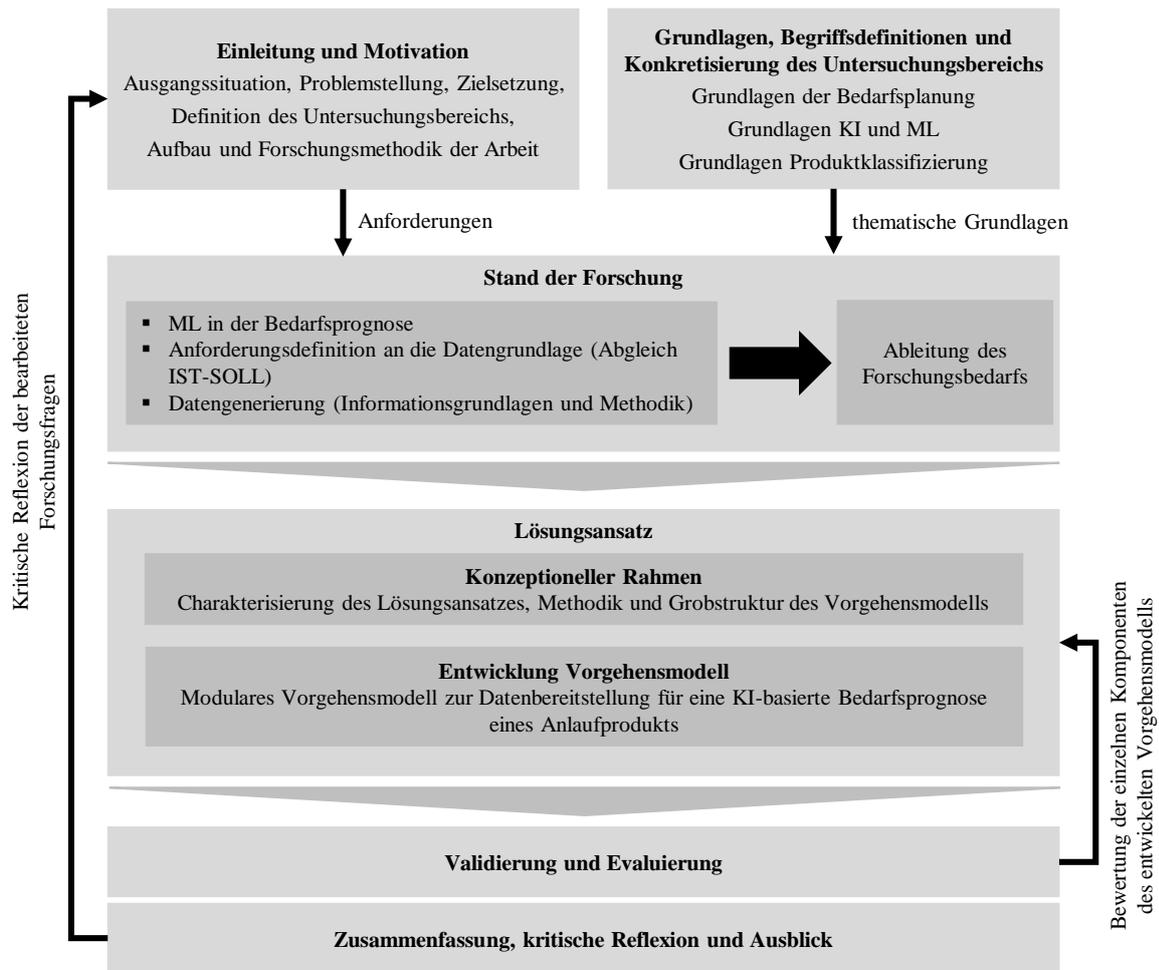


Abbildung 4: Struktureller Aufbau der Arbeit (eigene Darstellung)