

Michael Schulz ▪ Uwe Neuhaus ▪ Stephan Kühnel ▪ Heiko Rohde ▪ Sayed Hoseini ▪ René Theuerkauf  
(Hrsg.)

# DASC-PM v1.1

# Fallstudien



Dieses Werk ist lizenziert unter einer Creative Commons  
Namensnennung 4.0 Internationale Lizenz.  
<https://creativecommons.org/licenses/by/4.0/>

Elmshorn 2023

[info@dasc-pm.org](mailto:info@dasc-pm.org)

Herausgeber:

NORDAKADEMIE gAG Hochschule der Wirtschaft  
Köllner Chaussee 11  
25337 Elmshorn

Gefördert durch die  
NORDAKADEMIE-Stiftung

Die aktuelle Version des DASC-PM ist erschienen als

Schulz, Michael; Neuhaus, Uwe; Kaufmann, Jens; Kühnel, Stephan; Alekozai, Emal M.; Rohde, Heiko; Hoseini, Sayed; Theuerkauf, René; Badura, Daniel; Kerzel, Ulrich; Lanquillon, Carsten; Daurer, Stphan; Günther, Maik; Huber, Lukas; Thiée, Lukas-Walter; zur Heiden, Philipp; Passlick, Jens; Dieckmann, Jonas; Schwade, Florian; Seyffarth, Tobias; Badewitz, Wolfgang; Rissler, Raphael; Sackmann, Stefan; Gölzer, Philipp; Welter, Felix; Röth, Jochen; Seidelmann, Julian; Haneke, Uwe: **"DASC-PM v1.1 - Ein Vorgehensmodell für Data-Science-Projekte"**, NORDAKADEMIE gAG Hochschule der Wirtschaft, Hamburg 2022, ISBN: 978-3-9824465-0-9, DOI: 10.25673/85296.

# Inhaltsverzeichnis

<b>Vorwort</b>	<b>5</b>
<b>1 Wissenssicherung mit Wissensgraphen</b>	<b>6</b>
Carsten Lanquillon, Sigurd Schacht	
<b>2 Kollaborativer Empfehlungsdienst für stark korrelierende Dialogeingaben</b>	<b>16</b>
Julian Seidelmann	
<b>3 Entwicklung eines Machine-Learning-Modells zur Warenplanung in der Supply Chain</b>	<b>21</b>
Jonas Dieckmann, Daniel Badura	
<b>4 Projekt FLEMING – Predictive Maintenance von zentralen Komponenten des Mittelspannungsnetzes</b>	<b>29</b>
Philipp zur Heiden	
<b>5 Der Weg zum Projektauftrag</b>	<b>39</b>
Florian Schwade, Heiko Rohde	
<b>6 Identifikation von Gesichtsmasken</b>	<b>49</b>
René Theuerkauf, Tony Franke	

# Vorwort

Die Entwicklung des DASC-PM (Data Science Process Model) basiert auf dem Wissen einer großen Arbeitsgruppe, die sich aus Expertinnen und Experten der Data Science zusammensetzt. Wir sind der Meinung, dass die im DASC-PM beschriebene Vorgehensweise in datengetriebenen Projekten nutzbringend eingesetzt werden kann und eine Unterstützung bei der Strukturierung komplexer Vorhaben bietet. In der Version 1.1 des DASC-PM, die im März 2022 in deutscher Sprache und im Juni 2022 in englischer Sprache erschienen ist, haben wir die ursprüngliche Dokumentation auf Basis der Hinweise von Leserinnen und Lesern nochmals deutlich überarbeitet. Wir hoffen, dass das umfangreiche Dokument nun eingängiger und strukturierter aufbereitet ist und somit die Einstiegshürden bei der erstmaligen Nutzung des Vorgehensmodells reduziert.

Mit dem vorliegenden Dokument setzen wir einen weiteren Vorschlag der Leserinnen und Leser um: eine Sammlung von Fallstudien, bei der das DASC-PM angewendet wurde. Die Dokumentation des DASC-PM hat auf der einen Seite den Anspruch, das Vorgehen in einem Data-Science-Projekt möglichst detailliert aufzuzeigen. Auf der anderen Seite soll es aber auch eine Allgemeingültigkeit gewährleisten, die sich in einem durchaus breiten Anwendungsspektrum niederschlägt. Das beispielhafte Aufzeigen der direkten Anwendung und auch der Nutzbarmachung des DASC-PM für bestimmte Anwendungsdomänen stellt einen weiteren Aspekt dar, der die oben genannten Einstiegshürden reduzieren soll.

In dem vorliegenden Sammelband sind Fallstudien enthalten, die den Leserinnen und Lesern eine Hilfestellung für den Einsatz des DASC-PM bieten sollen. Dieser Einsatz wird dabei aus verschiedenen Perspektiven und auf ganz unterschiedlichen Abstraktionsebenen dargestellt. Auch werden sowohl Projekte aus der Praxis als auch aus der Wissenschaft betrachtet. Wir möchten allen Autoren der Fallstudien herzlich für Ihr Engagement in der Arbeitsgruppe danken und freuen uns sehr über die spannenden Beiträge, die daraus entstanden sind.

Diese Sammlung von Fallstudien soll allerdings nicht als abgeschlossen angesehen werden. Wir freuen uns jederzeit über Ihr Feedback zum Einsatz des DASC-PM in praktischen und wissenschaftlichen Projekten. Gerne nehmen wir auch Ihre Fallstudie in unsere Sammlung auf.

Wenn Sie Interesse an der Mitarbeit in der Arbeitsgruppe haben oder von uns über aktuelle Entwicklungen rund um das Modell informiert werden möchten, melden Sie sich gerne unter der unten angegebenen Kontaktadresse.

Elmshorn, Flensburg, Halle (Saale), Hamburg und Krefeld im Februar 2023

DASC-PM-Kernteam

Kontakt: [info@dasc-pm.org](mailto:info@dasc-pm.org)

# 1 Wissenssicherung mit Wissensgraphen

Carsten Lanquillon, Sigurd Schacht

## 1.1 Einleitung

Der Verlust von Erfahrungswissen ausscheidender Mitarbeiterinnen und Mitarbeiter kann gerade für kleinere und mittlere Unternehmen existenzbedrohend sein. Aber auch ein einfacher Zugriff auf bestehendes Erfahrungswissen in einem Unternehmen zur Förderung des Austausches und der Zusammenarbeit stellt einen großen Mehrwert bei der täglichen Arbeit dar. Insbesondere die Sicherung und der Transfer des bestehenden Erfahrungswissens sind zentrale Bausteine eines umfassenden Wissensmanagements (Probst, 2013).

Für einen erfolgreichen Einsatz muss das oft nur implizit in Dokumenten und in den Köpfen der Menschen verfügbare oder in Dialogen geäußerte Wissen in meist mühevoller Arbeit erschlossen werden. So kann durch strukturierte Interviews und Abschlussbesprechungen (Debriefings) beim Ausscheiden einer Mitarbeiterin oder eines Mitarbeiters oder zum Ende eines Projekts relevantes Erfahrungswissen dokumentiert und weitergegeben werden (Müller und Kaiser, 2006). Die manuelle Form der Wissenserfassung ist ein Engpass, der die Sicherung und Weitergabe des Erfahrungswissens bremst und eine erfolgreiche Umsetzung eines umfassenden Wissensmanagements gefährdet.

Ziel dieses Data-Science-Projekts war daher die Entwicklung und Anwendung KI-gestützter Methoden für eine stärkere Automatisierung des Wissenserfassungsprozesses. Mit Hilfe von Machine-Learning-Methoden und Methoden des Natural Language Processings (NLP) wurden Wissensfragmente aus Projekt-Artefakten wie Berichten, Dokumentationen oder der Kommunikation zwischen beteiligten Personen extrahiert und in einem Wissensgraphen abgelegt. Da diese das relevante Erfahrungswissen oft nur unvollständig und mit gewisser Unsicherheit abbilden, wurden sie anschließend durch die Anwendung klassischer Methoden der Wissenserfassung verfeinert und komplementiert, wie in Abbildung 1.1 dargestellt.

Eine geeignete Aufbereitung und Darstellung der bestehenden Wissensgraphen etwa in Form von Wissenslandkarten kann den Wissensmanagement-Expertinnen und -Experten als Vorbereitung für ihre Abschluss-Interviews und Debriefings dienen, so dass der Erfassungsprozess erheblich beschleunigt und die Qualität der erfassten Wissensfragmente gesteigert wird. Außerdem stellen Wissensgraphen auch eine wichtige Basis für weitere Wissensmanagement-Anwendungen wie Chatbots zur Beantwortung von Fragen oder anderer KI-basierter Assistenzsysteme dar.

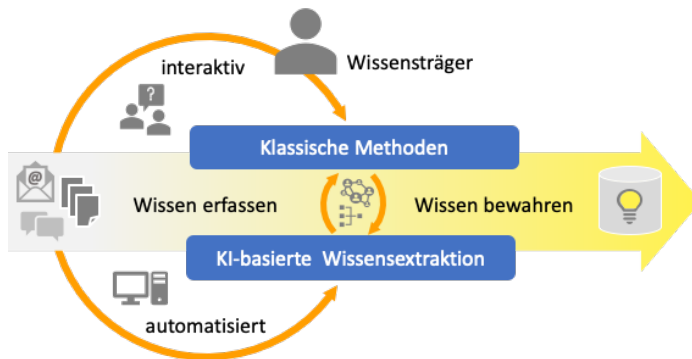


Abbildung 1.1: Wissen soll aus unterschiedlichen Quellen sowohl automatisch (KI-basiert) und interaktiv erfasst und gespeichert werden, so dass eine Nutzung bei neuen Aufgaben oder eine Weitergabe ermöglicht wird.

## 1.2 Vorgehensweise

### Schlüsselbereich „Wissenschaftlichkeit“

Diese Fallstudie beruht auf Ergebnissen und Erfahrungen von Forschungsprojekten im Bereich des KI-basierten Wissensmanagements. Sämtliche Phasen der Fallstudie wurden im Rahmen der Projekte von umfassenden Literaturanalysen zur Erfassung des aktuellen Stands der Forschung in den jeweiligen Teilbereichen begleitet, um geeignete Vorgehensweisen und Methoden auszuwählen. Aufgrund des Querschnitt-Charakters dieses Schlüsselbereichs wird in den folgenden Bereichen nicht mehr explizit darauf hingewiesen. Da bei dieser Fallstudie der Fokus auf der Beschreibung der Vorgehensweise liegt, wird auf eine ausführliche Darstellung der jeweiligen Erkenntnisse aus den Literaturanalysen sowie der zur Anwendung kommenden Analyseverfahren selbst verzichtet.

### Phase „Projektauftrag“

Der Gegenstand dieses Projekts war das innerbetriebliche Wissensmanagement. Es handelt sich um eine querschnittliche Disziplin, die in beliebigen fachlichen Domänen zum Einsatz kommen kann. Hintergrundwissen aus fachlichen Domänen ist bei der Festlegung, Erfassung und Bewertung relevanter Wissensfragmente zwar sehr wichtig, die Erarbeitung und Darstellung der Vorgehensweise in dieser Fallstudie werden davon jedoch nicht berührt. Daher wird im Rahmen dieser Fallstudie keine fachliche Domäne vorgestellt. Es wird lediglich angenommen, dass es um projektbezogene Tätigkeiten geht, für die geeignete textbasierte Artefakte vorliegen. Diese textuellen Artefakte werden auch kurz als Dokumente bezeichnet.

#### Use Case

Übergeordnetes Ziel des Projekts war der Aufbau eines Wissensgraphen mit allen relevanten Zusammenhängen (Wissensfragmenten) in einer definierten fachlichen Domäne als Baustein der innerbetrieblichen Wissenssicherung. Die Konstruktion des Wissensgraphen stellt aus der Perspektive des Wissensmanagements lediglich die Datenaufbereitung dar, die aus unstrukturierten Dokumenten (Projekt-Artefakte) einen strukturierten Datenbestand (Wissensgraph) erzeugt. Dieser wiederum entspricht im Rahmen der Verwendung in Wissensmanagement-Anwendungen und anderen KI-basierten Assistenzsystemen der analytische Datenquelle.

Während sich die Lösung einfacher kanonischer Data-Science-Aufgaben wie die Klassifikation oder numerische Prognose oft direkt und ganzheitlich in einem DASC-PM-Zyklus abbilden und beschreiben lassen, erfordern viele komplexe Aufgaben wie die Wissenssicherung mit Wissensgraphen basierend auf Dokumenten eine Zerlegung der Aufgabe in Teilprobleme.

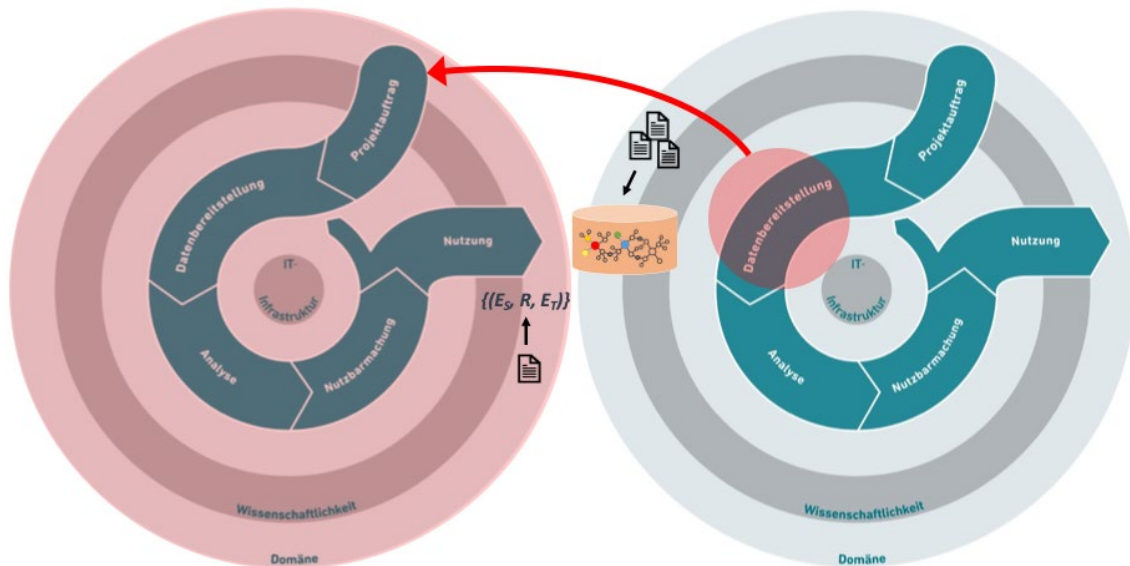


Abbildung 1.2: Die Anwendung von DASC-PM mit der Konstruktion eines Wissensgraphen aus Texten als Datenbereitstellung im Rahmen der Wissenssicherung auf der Makroebene (rechts) und als umfassende Wissensextraktion aus Texten auf der Mikroebene (links).

Derartig geschachtelte Lösungsmuster sind typisch für komplexe Problemstellungen in realen Anwendungsdomänen. Bei der Umsetzung ist es besonders wichtig, dass ein Vorgehensmodell zum Einsatz kommt, das eine einheitliche Anwendung sowohl auf der Makroebene des gesamten Projekts als auch auf der Mikroebene der einzelnen Schritte und Teilaufgaben erlaubt. In dieser Fallstudie wird beschrieben, wie DASC-PM mit seinen Schlüsselbereichen bzw. Phasen auf allen Granularitätsebenen angewendet wird.

Als Wissensfragmente betrachten wir hier sogenannte SPO-Tripel oder nur kurz Tripel  $(E_s, R, E_t)$  bestehend aus Subjekt  $E_s$ , Prädikat  $R$  und Objekt  $E_t$ . Das Prädikat definiert eine Beziehung (Relation) zwischen der Start-Entität  $E_s$  und der Ziel-Entität (Target-Entität)  $E_t$ , die in einem Graphen als gerichtete Kante zwischen den Entitäten als Knoten repräsentiert wird. Die Menge aller Tripel stellt den Wissensgraphen dar. Es handelt sich dabei wie oben dargestellt um eine komplexe Aufgabe, die aber mit Hilfe geeigneter Methoden des maschinellen Lernens und des Natural Language Processing (NLP) umgesetzt werden kann. Abbildung 1.2 zeigt die dafür vorgesehene geschachtelte Anwendung von DASC-PM.

### Projektskizze

In der Literatur werden zwei Extreme zum Aufbau von Wissensgraphen beschrieben (Zhao et al., 2018): Beim Top-Down-Ansatz wird vorab eine Ontologie als Menge von Schemata mit allen geforderten Entitätstypen und Beziehungstypen definiert und anschließend mit geeigneten Entitäten und Beziehungen zwischen diesen Entitäten befüllt, die aus den verfügbaren Dokumenten extrahiert werden. Beim Bottom-Up-Ansatz werden zunächst die Entitäten und Beziehungen als isolierte und in der Regel nicht typisierte Wissensfragmente extrahiert. Eine Typisierung und Zusammenfassung zu sinnvollen Strukturen erfolgen dann im Nachgang.



Zur Klärung der Umsetzbarkeit hat eine Erprobung der beiden Extreme gezeigt, dass beide gravierende Nachteile haben, denn sowohl das Modellieren einer vollständigen Ontologie vorab als auch das nachträgliche Strukturieren sind sehr zeitaufwendig. Beim Top-Down-Ansatz kommt erschwerend hinzu, dass eine vollständige Festlegung der Schemata im Vorfeld nicht praktikabel ist, da zu Projektbeginn oftmals noch nicht umfassend bekannt ist, welche Typen an Entitäten und Beziehungen in einer Anwendung benötigt werden.

Es sind hierbei eindeutig Parallelen zum Data Warehousing erkennbar. Denn auch dort ist eine ganzheitliche Modellierung im Vorfeld gemäß dem reinen Top-Down Ansatz nach Inmon zwar erstrebenswert, aber in der Praxis oftmals nicht realisierbar, während sich beim Bottom-Up-Ansatz nach Kimball insbesondere Probleme bei der Integration der Daten ergeben können. In der Praxis hat sich daher eine Mischform etabliert, die dem bekannten Leitbild „*think big, start small, grow step by step*“ folgt (SCN, 2013).

Dieses Prinzip liegt auch dem hier gewählten Vorgehen bei der Konstruktion von Wissensgraphen zugrunde. Die Vision ist ein vollständig typisierter Wissensgraph mit entsprechenden Schemata. Diese werden aber nicht wie im reinen Top-Down-Ansatz vollständig vorab modelliert und bereitgestellt, sondern iterativ. Wie in Abbildung 1.3 dargestellt, wird der Wissensgraph sukzessive für ausgewählte Fragestellungen um neue Wissensfragmente erweitert. Jede Iteration durch die Schritte der Wissensmodellierung leitet die Bearbeitung einer separaten Data-Science-Aufgabe ein, die dann auf der Mikroebene ebenfalls auf Basis von DASC-PM bearbeitet wird. Dabei entspricht die in jeder Iteration wiederkehrende Festlegung einer relevanten Fragestellung (Schritt 1) einem internen Projektauftrag zur Erstellung eines geeigneten Modells für die Wissensextraktion als Analyse-Artefakt und umfasst mit seiner Einbettung in die Anwendungsdomäne den Schlüsselbereich Domäne. Mit entsprechendem Fachwissen aus der Anwendung müssen die für die Beantwortung notwendigen und relevanten Entitäts- und Beziehungstypen identifiziert und als Schemata modelliert werden (Schritt 2). Dieser Schritt und die Bereitstellung von Trainingsdaten (Schritt 3) als analytische Datenquelle lässt sich der Phase Datenbereitstellung zuordnen. In Schritt 4 erfolgt die Erstellung der Modelle für die Wissensextraktion als Hauptaufgabe der Phase Analyse. Die Nutzbarmachung und Nutzung (Schritte 5 und 6) der Modelle zur Extraktion der Wissensfragmente aus allen relevanten Dokumenten wird hier explizit um die Möglichkeit der Integration von Feedback der Anwenderinnen und Anwender ergänzt (Schritt 7). Das teilweise auch aktiv eingeforderte Feedback wird auch im Wissensgraphen abgelegt und zur Evaluierung und Verbesserung der eingesetzten Modelle verwendet. Diese Vorgehensweise hat den Vorteil, dass die Wissensfragmente bereits in typisierter Form vorliegen und der Aufwand für das nachträgliche Typisieren und Strukturieren begrenzt ist. Zudem können erste Ergebnisse der Wissensmodellierung schon frühzeitig in Anwendungen genutzt und somit auch validiert werden.

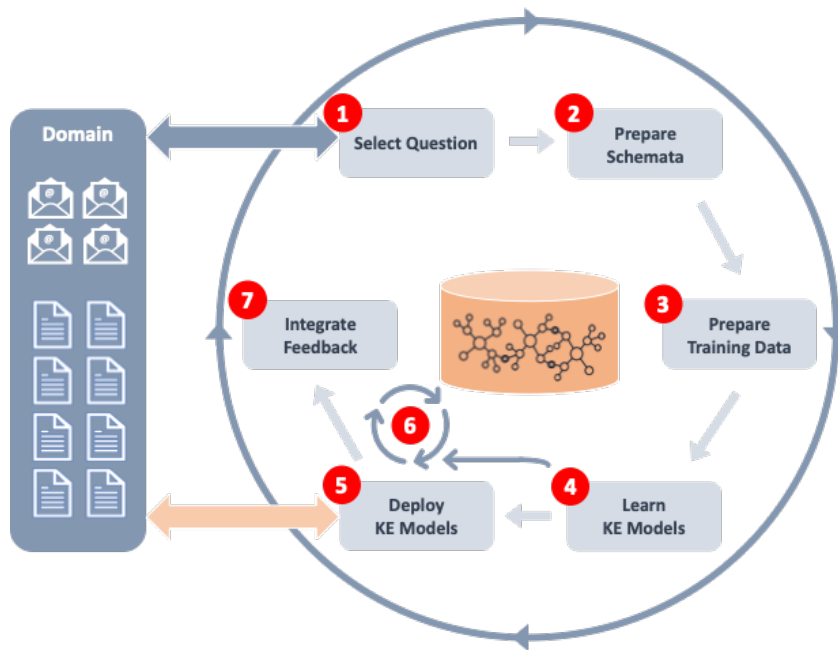


Abbildung 1.3: Die iterative Konstruktion eines Wissensgraphen auf Basis von Fragestellungen, die von den Anwenderinnen und Anwendern in der Domäne ausgewählt und vorgegeben werden, erfordert die Erstellung individueller Modelle für die Wissensextraktion (KE Models). Die Wissensmodellierung folgt dabei wohldefinierten Schritten, die sich nahtlos in DASC-PM einfügen.

## Phase „Datenbereitstellung“

Nach der Definition der Problemstellung spielt die Erstellung der analytischen Datenquelle eine zentrale Rolle für den Erfolg eines Data-Science-Projektes. Da in dieser Fallstudie primär unstrukturierte textuelle Daten die Grundlage für die Lernprozesse bilden, unterscheiden sich die durchzuführenden Arbeitsschritte und die geeigneten Methoden im Detail erheblich von Projekten mit strukturierten Daten. Die zentralen Aspekte der DASC-PM-Phasen behalten dabei dennoch ihre Gültigkeit und sind zur Strukturierung des Projekts sowohl auf der Makroebene also auch auf den Mikroebenen anwendbar.

### Ursprungsdatenquellen

Die Datengrundlage für diese Fallstudie ist eine Menge von textbasierten Projekt-Artefakten wie Berichte und Dokumentationen sowie elektronisch verfügbare Korrespondenz zwischen Personen, die an den Projekten beteiligt sind. Diese textbasierten Artefakte werden kurz als Dokumente und die gesamte Menge als Kollektion bezeichnet.

### Datenmanagement

Es wird vereinfacht davon ausgegangen, dass alle Artefakte, die für die Wissensextraktion verwendet werden sollen und dürfen, in speziell dafür vorgesehenen Verzeichnissen liegen und in einem Format vorliegen, aus dem der Text direkt entnommen werden kann. Für einen operativen Einsatz muss das Datenmanagement angepasst und erweitert werden. Dabei ist nicht nur ein automatisierter Zugriff auf relevante Projekt-Artefakte notwendig, sondern auch die Sicherstellung des Schutzes persönlicher oder vertraulicher Dokumente.

### Explorative Datenanalyse

Da die vorliegenden Daten unstrukturiert sind, unterscheidet sich die explorative Datenanalyse zum Aufbau eines Datenverständnisses sehr stark von Projekten mit strukturierten Daten. Nichtsdestotrotz ist ein Betrachten der Dokumente in ihrer ursprünglichen Form wichtig. So muss insbesondere

geprüft werden, ob die Dokumente in derselben angenommenen Sprache vorliegen oder möglicherweise mehrsprachig sind, was dann eine besondere Behandlung erforderte. Ebenso kann die sprachliche Qualität der Dokumente eine entscheidende Rolle spielen. Gerade schnell formulierte Nachrichten enthalten oft viele Abkürzungen und mehr sprachliche Fehler als sorgfältig bearbeitete Texte. In diesen Fällen kann eine vorgeschaltete Ersetzung von Abkürzungen oder eine automatisierte Korrektur die Qualität der folgenden Verarbeitungsschritte verbessern. Im Rahmen dieser Fallstudie gehen wir allerdings davon aus, dass einsprachige Dokumente in akzeptabler Qualität vorliegen.

#### Datenaufbereitung: Von der Ursprungsdatenquelle zur analytischen Datenquelle

Auf der Makroebene des Projekts entspricht die Extraktion relevanter Wissensfragmente aus den bereitgestellten Dokumenten der Datenaufbereitung. Der erstellte Wissensgraph ist in diesem Fall die analytische Datenquelle, auf die folgende Analyseverfahren aufsetzen. In den Phasen Nutzbarmachung und Nutzung wird die Thematik der regelmäßigen Wissensextraktion für eine Berücksichtigung neuer Dokumente adressiert.

Auf der Mikroebene des Projekts wird die Wissensextraktion durch Erlernen und Anwenden geeigneter Verfahren und Modelle realisiert. Da die meisten Verfahren nicht direkt mit Text als Eingabe umgehen können, stellt die Textrepräsentation einen entscheidenden Aspekt an dieser Stelle dar. Da die einzelnen Teilaufgaben innerhalb der Wissensextraktion oft unterschiedliche Anforderungen hinsichtlich der Datenaufbereitung haben, lässt sich eine allgemeingültige Repräsentation nicht festlegen. Abbildung 1.4 zeigt einige klassische Schritte der Datenaufbereitung für Text.

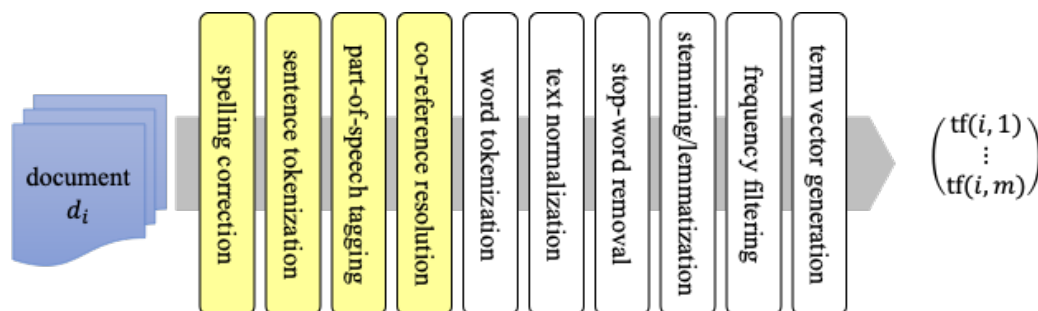


Abbildung 1.4: Klassische Schritte der Datenaufbereitung für die Überführung von Text in eine numerische Vektorrepräsentation, die für eine algorithmische Verarbeitung leichter zugänglich ist. Weniger typische Schritte sind hier gelb und häufiger verwendete Schritte weiß markiert.

Bei der dargestellten klassischen Aufbereitung von Text kommen oft viele verschiedene Schritte zum Einsatz, bis ein Dokument in eine numerische Vektorrepräsentation überführt wurde. Zu beachten ist, dass diese einfache Form der Aufbereitung Dokumente als Einheit auffasst und die Reihenfolge der Wörter ignoriert. Einige Aufgaben im Rahmen der Wissensextraktion erfordern aber eine Analyse auf Ebene der Wörter unter Berücksichtigung des Kontexts und somit der Reihenfolge der Wörter. So bleiben zwar einzelne Teilschritte relevant, deren Anwendung muss jedoch sehr sorgfältig gewählt werden. Teilweise führen die Ergebnisse der Teilschritte nicht mehr zu einer Transformation der ursprünglichen Daten, sondern zu einer Anreicherung der Repräsentation wie etwa um Tags für bestimmte Wortarten. Angewendet auf einzelne Wörter resultiert diese Vektorrepräsentation in einem sogenannten One-Hot-Encoding, bei dem das bloße Vorhandensein an der Stelle eine 1 steht, deren Index der Position des entsprechenden Wortes in einer definierten Wortliste (Vokabular) entspricht, und alle anderen Werte 0 sind. Da bei dieser Darstellung die Semantik der Wörter nicht berücksichtigt wird, so dass verschiedene Wörter alle gleichermaßen unähnlich sind, verwenden moderne NLP-Ansätze überwiegend globale Word Embeddings, die jedem Wort gemäß eines

Sprachmodells eine statische numerische Repräsentation zuordnen, oder Encoder-Repräsentationen, die auch den konkreten Kontext eines Wortes bei der Bestimmung einer geeigneten Repräsentation berücksichtigen. Bei diesen modernen Ansätzen wird ein signifikanter Teil der Datenaufbereitung automatisiert durchgeführt. Komponenten für die Erzeugung geeigneter Textrepräsentationen sind dann Teil einer umfassenden Modell-Architektur und werden entweder unverändert aus vortrainierten Modellen übernommen oder in einem Fine-Tuning-Schritt explizit an die Besonderheiten einer Domäne angepasst, so dass elementare Aufbereitungsschritte in dieser Phase entfallen.

Das Vorgehen für die Anreicherung der analytischen Datenquelle um mögliche Label für die Zielgrößen für die Anwendung überwachter Lernverfahren wird bei der Betrachtung der einzelnen NLP-Aufgaben im Schlüsselbereich Analyseverfahren diskutiert.

### Phase „Analyse“

Die Extraktion von Wissensfragmenten (Tripeln) umfasst mehrere NLP-Aufgaben wie in Abbildung 1.5 als Pipeline dargestellt. Während der erste Schritt (PREP) als klassische Datenaufbereitung der Phase Datenbereitstellung zuzuordnen ist, werden die folgenden Schritte durch spezielle Modelle realisiert, die entweder jeweils isoliert oder in Kombination aus Daten gelernt wurden (Phase Analyse). Die aus einem Dokument extrahierten Wissensfragmente werden in einem Wissensgraphen abgelegt. Der Extraktionsprozess wird auf alle bereitgestellten Dokumente angewendet, bei Verbesserung der einzelnen Modelle nach Bereitstellung von Feedback auch mehrfach (Phase Nutzung).

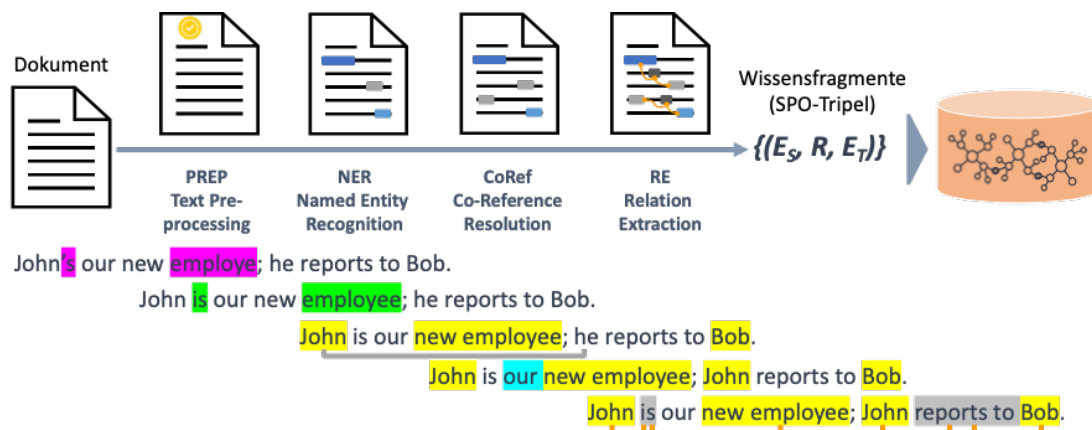


Abbildung 1.5: NLP-Aufgaben einer Verarbeitungs-Pipeline für die Extraktion von Wissensfragmenten aus einem Dokument.

Zum leichteren Verständnis und zur Vereinfachung der prozessbegleitenden Kontrolle und Evaluierung betrachten wir die Schritte zunächst isoliert. Zu beachten ist, dass für jede der NLP-Aufgaben eine separate Durchführung nach DASC-PM erforderlich ist (Mikroebene). Dabei ist insbesondere auf die Notwendigkeit einer ausführlichen Recherche des aktuellen Stands der Forschung hinzuweisen, um diese im Projekt adäquat berücksichtigen zu können. Im Folgenden werden lediglich die Ziele der einzelnen Phasen mit aktuellen Lösungsvarianten kurz skizziert.

#### NLP-Aufgaben für die Extraktion von Wissensfragmenten

Ziel der *Named Entity Recognition* (NER) ist die Identifikation von Entitäten bestimmter vorgegebener semantischer Typen wie etwa Personen, Orte oder Organisationen anhand ihnen zugeordneter fester Bezeichnungsausdrücke. Für diese generischen Typen existieren zahlreiche vortrainierten Modelle, die in allgemeinen Texten oft schon sehr gute Ergebnisse erzielen. Domänenspezifische Anwendungen erfordern oft nicht nur eine Anpassung hinsichtlich der vorkommenden Entitäten dieser Typen, sondern auch eine Erweiterung auf zusätzliche domänenspezifische Typen wie die Identifikation von Data-Science-Fachbegriffen in Forschungsartikeln. Die NER-Aufgabe wird zurzeit oft

mit Deep-Learning-Ansätzen wie bidirektionalen LSTM-Architekturen erfolgreich gelöst. Abhängigkeiten zwischen Wörtern etwa bei Entitäten, die aus mehreren Wörtern bestehen, werden durch eine abschließende Conditional-Random-Field-Schicht sehr gut berücksichtigt (Goyal et al., 2018; Panchendrarajan & Amaresan, 2018). Ergänzt durch Entity Linking werden Entitäten zusammengeführt, die sich auf dasselbe Objekt der realen Welt beziehen. Verfahren zum Auflösen von Koreferenzen (CoRef) ersetzen Bezugswörter in Dokumenten durch die Entitäten, auf die sie sich im Text beziehen. Nur dann können Wissensfragmente sinnvoll aus isolierten Satzteilen unabhängig vom weiteren Kontext extrahiert werden. Eine sehr bekannte Klasse von Verfahren dafür sind die Mention-Pair-Modelle (Ng, 2017). Die Relation Extraction (RE) soll relevante Beziehungen zwischen je zwei Entitäten erkennen, bei denen mindestens eine der relevanten Entitäten beteiligt ist. Auch für diese Aufgabe werden vermehrt rekurrente Deep-Learning-Ansätze eingesetzt (Wang et al., 2021).

### Bereitstellung der Trainingsdaten für überwachte Lernverfahren

Für die meisten NLP-Aufgaben erreichen überwachte Deep-Learning-Ansätze aktuell sehr gute Leistungen. Durch Verwendung vortrainierter Modelle mit anschließendem Fine-Tuning auf die jeweilige Domäne wird die Menge der für die Modellerstellung benötigten gelabelten Daten reduziert. Ein zweistufiger Ansatz vereinfacht die Bereitstellung der Trainingsdaten: Regelbasiert werden jeweils mit Listen bekannter Entitäten und Beziehungen als Seed-Werte erste Trainingsdaten erzeugt. Mit diesen wird im zweiten Schritt das Fine-Tuning der Deep-Learning-Ansätze durchgeführt. Der Lernprozess wird zusätzlich durch aktives Lernen auf Basis neuer Wissensfragmente unterstützt, die über die Feedback-Komponente gezielt erfragt werden können (Shen et al., 2018).

### Organisation der Lernprozesse

Die in jedem einzelnen Schritt erzielten Ergebnisse werden in der Nutzungsphase für die Wissensextraktion zu einer Verarbeitungs-Pipeline kombiniert. Zu bedenken ist, dass durch dieses Vorgehen etwaige Fehler jedes einzelnen Schritts sich bis zum Ende der Pipeline verstärken können und kumuliert werden. Dies ließe sich durch ein kombiniertes Lernen eines Gesamtmodells (Joint Modeling) für die Extraktion der Wissensfragmente vermeiden (Singh et al., 2013; Geng et al., 2021). Die Bereitstellung der Trainingsdaten dafür ist aber meist komplexer und aufwendiger und die Wiederverwendbarkeit zur Extraktion von Wissensfragmenten, die sich bestimmte Entitätstypen oder Beziehungstypen teilen, ist erschwert. Deshalb wird an dieser Stelle das Joint Modeling noch nicht verwendet.

## Phasen Nutzbarmachung und Nutzung

Auf der Mikroebene werden für jede einzelne Fragestellung Pipelines mit Modellen für die verschiedenen NLP-Aufgaben entwickelt, die im Zusammenspiel die Extraktion dedizierter Wissensfragmente ermöglichen. Diese müssen in die Phase Nutzbarmachung derart zur Anwendung gebracht werden, dass der Wissensgraph initial auf Basis aller verfügbaren Dokumente befüllt wird.

Über die vorgesehenen Feedback-Komponenten wird die Qualität der Extraktion während der Nutzung beobachtet. Die dabei teilweise auch aktiv erfragten neuen Wissensfragmente werden zur Erweiterung der analytischen Datenquellen verwendet. Sobald neue Dokumente zur Verfügung stehen oder Modelle der Extraktions-Pipeline auf Basis von Feedback verbessert wurden, wird der Wissensgraph entsprechend erweitert. Dieses Vorgehen ist vergleichbar mit der Anwendung und Pflege von ETL-Strecken beim Data Warehousing.

Bezogen auf das Gesamtprojekt (Makroebene) dient die Nutzung aller Extraktions-Pipelines zum Aufbau des Wissensgraphen als analytische Datenquelle als wesentliche Komponente der Phase Datenbereitstellung. Die analytische Datenquelle wird für weiterführende Wissensmanagement-An-

wendungen wie der Erstellung von Wissenslandkarten zur Unterstützung von Interviews und Debriefings oder der Entwicklung von KI-basierten Assistenzsystemen genutzt. Auf der Makroebene basieren diese Anwendungen auf Analyse-Artefakten, die in der Phase Analyse entwickelt werden und deren Nutzbarmachung und Nutzung explizit geplant werden muss.

### 1.3 Fazit

Diese Fallstudie zeigt, dass DASC-PM als Vorgehensmodell einen idealen Rahmen auch für Projekte mit komplexeren Problemstellungen bietet. Abgeleitet aus dem Use Case, in dieser Fallstudie eine KI-basierte Wissenssicherung, wurde über die Phase Projektauftrag das Projekt in kleinere Aufgaben heruntergebrochen. Dies lieferte einen klaren Rahmen für das Aufsetzen der Makroebene im Projekt. Die Lösung und Umsetzung jeder Teilaufgabe erfolgten ebenso über DASC-PM als Vorgehensmodell, um auch dort die einzelnen Aktivitäten wie etwa die Problemstellung, Erfassung und Aufbereitung der Daten und Durchführung der Analyse strukturiert bearbeiten zu können. In den Phasen Nutzbarmachung und Nutzung erfolgte im vorliegenden Anwendungsfall die Zusammenführung der Ergebnisse. Es hat sich somit gezeigt, dass DASC-PM flexibel genug ist, um komplexe Probleme mit mehrschichtigen Lösungskomponenten zu entwickeln. Schließlich ist der Schlüsselbereich Wissenschaftlichkeit hervorzuheben, da kaum ein KI-Projekt durchgeführt werden kann, ohne nicht auch den aktuellen Stand der Forschung mit einer Literaturanalyse zu erfassen und somit die Auswahl geeigneter Methoden auf allen Ebenen der Umsetzung auswählen zu können. Auch dafür bietet DASC-PM den notwendigen Rahmen.

## Literaturverzeichnis

- Geng, Z., Zhang, Y. & Han, Y. (2021). Joint entity and relation extraction model based on rich semantics. *Neurocomputing*, vol. 429, S. 132–140.
- Goyal, A., Gupta, V. & Kumar, M. (2018). Recent named entity recognition and classification techniques: A systematic review. *Computer Science Review*, Vol. 29, No. 1, S. 21–43.
- Müller, M. & Kaiser, R. (2006). Wissensbewahrung bei der Stadt Erlangen - Dokumentation und Kommunikation der Erfahrungen ausscheidender Wissensträger. In *Tagungsband KnowTech 2006*, S. 425–432.
- Ng, V. (2017). Machine Learning for Entity Coreference Resolution: A Retrospective Look at Two Decades of Research. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*, S. 4877–4884
- Panchendrarajan, R. & Amaresan, A. (2018). Bidirectional LSTM-CRF for Named Entity Recognition. In *Proceedings of PACLIC 2018*, S. 531–540.
- Probst, G., Raub, S. & Romhardt, K. (2013). *Wissen managen: Wie Unternehmen ihre wertvollste Ressource optimal nutzen*. 7. Auflage, SpringerGabler.
- Schulz, M., Neuhaus, U., Kaufmann, J., Kühnel, S., Alekozai, E., Rohde, H., Hoseini, S., Theuerkauf, R., Badura, D., Kerzel, U., Lanquillon, C., Daurer, S., Günther, M., Huber, L., Thiee, L., zur Heiden, P., Passlick, J., Dieckmann, J., Schwade, F., Seyffarth, T., Gölzer, P., Welter, F., Röth, J., Seidelmann, J. & Haneke, U. (2022). DASC-PM v1.1 - Ein Vorgehensmodell für Data-Science-Projekte. NORDAKA-DEMIE gAG.
- SCN Education B.C. (2013). *Data Warehousing: The Ultimate Guide to Building Corporate Business Intelligence*. Vieweg+Teubner Verlag.
- Shen, Y., Yun, H., Lipton, Z. C., Kronrod Y. & Anandkumar, A. (2018). Deep Active Learning for Named Entity Recognition. In *Proceedings of the 2<sup>nd</sup> Workshop on Representation Learning for NLP*, S. 252–256.
- Singh, S., Riedel, S., Martin, B., Zheng, J. & McCallum, A. (2013). Joint inference of entities, relations, and coreference. In *Proceedings of the 2013 Workshop on Automating Knowledge Base Construction (AKBC 2013)*. Co-located with CIKM 2013, S. 1–6.
- Wang, H., Lu, G., Yin, J. & Qin, K. (2021). Relation Extraction: A Brief Survey on Deep Neural Network Based Methods. In *Proceedings of the 4th International Conference on Software Engineering and Information Management (ICSIM 2021)*, S. 220–228.
- Zhao Z., Sung-Kook H. & So I.-M. (2018). Architecture of Knowledge Graph Construction Techniques. *International Journal of Pure and Applied Mathematics* 118(19): S. 1869–1883.



## 2 Kollaborativer Empfehlungsdienst für stark korrelierende Dialogeingaben

Julian Seidelmann

### 2.1 Motivation und Projektauftrag

Im Jahre 2020 wurden Waren im Wert von über 1.000 Milliarden Euro nach Deutschland importiert. Über die importierten Waren kann erst dann verfügt werden, wenn diese zu einem Zollverfahren angemeldet und von einer Zollstelle überlassen worden sind. Für die Anmeldung der Waren ist eine Zollanmeldung von Nöten. Diese Anmeldung ist ein amtliches Dokument, in dem eingeführte oder auch ausgeführte Waren aufgelistet und Angaben zu diesen aufgeführt werden.

Zollanmeldungen werden heutzutage typischerweise elektronisch über Softwaredienste und das Informationssystem des deutschen Zolls, ATLAS, abgegeben. Die Erfassung von Zollanmeldungen ist ein sowohl zeitaufwändiger als auch komplexer Prozess und verlangt vertiefende Fachkenntnisse des Zollwesens und der zollrechtlichen Bestimmungen. Zusätzlich zu der Komplexität und dem zu investierenden Zeitaufwand unterliegt die manuelle Erfassung von Daten immer der Gefahr des menschlichen Versagens. Die Digitalisierung von Anträgen und Formularen kann als Chance genutzt werden den Nutzer bei der Erfassung zu unterstützen.

Im Rahmen eines 75 Personentage umfassenden Projektes soll untersucht werden, ob Nutzer bei der Erfassung von Formularen durch Techniken der Data Science unterstützt werden können und damit den Nutzer sowohl bei der fachlichen Komplexität als auch beim zeitlichen Aufwand zu entlasten. Dieser Proof of Concept soll als Grundlage und Eignungsprüfung für Folgeprojekte und eine mögliche Integration in das bestehende System dienen. Das Risiko dieses Projektes ist als niedrig einzustufen. Zur Durchführung des Projekts wird das Vorgehensmodell DASC-PM genutzt. Für die Projektbeteiligten, mit geringer Data-Science-Projekten Erfahrungen, ist die durch das Vorgehensmodell vorgegebene Strukturierung besonders hilfreich.

Ein höheres Maß an Digitalisierung im Zollprozess kann nicht nur für effizientere Abläufe, sondern unter anderem auch neue Wege für bessere Compliance, mehr Sichtbarkeit in der Lieferkette und im Endeffekt einen Wettbewerbsvorteil bedeuten. Ein Ansatz, um Anwender bei der Nutzung von Dialogen zu unterstützen, ist die Empfehlung von möglichen zu tätigen Eingaben in Dialogen. Eine Art von Empfehlungsdiensten sind die kollaborativen Empfehlungsdienste. Der deutsche Zoll definiert über seine Implementierungshandbücher diverse Vorgaben und Validierungen, die eine starke Korrelation der Daten bedingen. Auch die unterschiedlichen unternehmensspezifischen Prozesse lassen auf eine starke Korrelation der Daten schließen. Nur diese verschiedenen Korrelationen können genutzt werden, um aus der Masse an Zollanmeldungen auf die Eingaben einer einzelnen Zollanmeldung zu schließen.



## 2.2 Datenbereitstellung

Zollanmeldungen als steuerliche Nachweise unterliegen einer gesetzlichen Aufbewahrungsfrist von zehn Jahren. Die Datengrundlage für dieses Forschungsprojekt erstreckt sich somit über mehrere Jahre.

Es wurde sich dazu entschieden eine Auswahl von zehn Unternehmen zu untersuchen und den Zeitraum auf die letzten fünf Monate zu begrenzen. Die zehn Unternehmen wurden ausgewählt, um eine Bandbreite an Variationen der Anzahl an Zollanmeldungen abzubilden. Die Anzahl der Datensätze pro Monat variiert hierbei von 200 bis 5.000 Zollanmeldungen. Die Ursprungquelle der Daten ist eine relationale Datenbank, in der alle benötigten Daten strukturiert und intern vorliegen. Eine Besonderheit dieses Projektes ist, dass nur eine Datenquelle genutzt wird. Es ist nicht nötig diese Daten durch externe Quellen anzureichern. Die Korrelation der Daten aus den Zollanmeldungen reicht aus, um diesen Empfehlungsdienst zu trainieren. Eine explorative Datenanalyse der Daten und eine Visualisierung als Heatmap (siehe Abbildung 2.1) kann die bereits vermutete starke Korrelation bestätigen aber auch eine verschieden verteilte Korrelation innerhalb der Dialogfelder als auch zwischen den Unternehmen offenlegen.

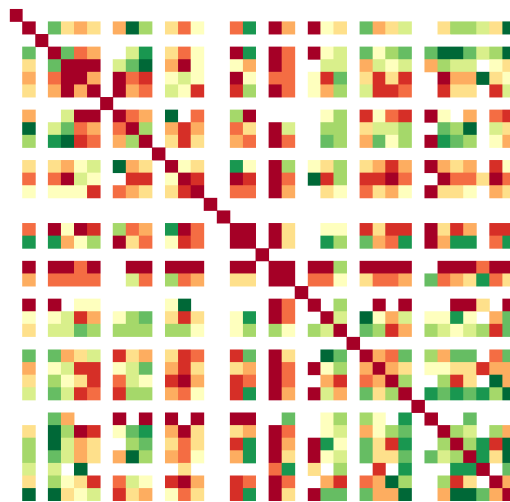


Abbildung 2.1: Beispielhafte Heatmap des Korrelationskoeffizienten Cramér's  $V$  aller Dialogfelder eines Unternehmens.

Um die durch Freitextfelder entstehende hohe Dimensionalität der Daten zu reduzieren, wurden Einträge mit geringem Aufkommen als Ausreißer angesehen und damit als 'andere' zusammengefasst. Damit der Informationsgehalt und Detailgrad der Daten nicht weiter reduziert wird, wurde auf weitere Anpassungen verzichtet. Die Einträge aus Freitextfeldern werden nun als kategorial angesehen.

Da nur binäre und kategoriale Daten vorliegen kann auf eine Normierung der Daten verzichtet werden. Auch auf eine umfangreiche Validierung der Daten ist unnötig, da nur positive durch den Zoll bestätigte Zollanmeldungen selektiert wurden und die Qualität der Daten bereits sichergestellt ist. Fehlende Daten müssen nicht aufgefüllt werden. Dieses spezifische Domänenmerkmal reduziert somit die in Data-Science-Projekten meist aufwändige Datenaufbereitung enorm.

## 2.3 Analyse

### Identifikation geeigneter Analyseverfahren

Kollaborative Empfehlungsdienste beruhen typischerweise auf Distanzen, bspw. k-Nächsten-Nachbarn oder auf Wahrscheinlichkeitsverteilungen, wie bei Bayesschen Netzen. Auch das maschinelle Lernen bietet Ansätze für kollaborative Empfehlungsdienste. Besonders Autoencoder, eine Art der neuronalen Netze, eignen sich, um Strukturen aus einer Menge von Daten zu extrahieren und diese erlernten Informationen als Empfehlungsgrundlage zu nutzen.

Bei auf Wahrscheinlichkeiten basierenden Modellen sind die erlernten Repräsentationen immer mit den A-posteriori-Verteilungen verbunden. Diese Verteilungen können bei Modellen mit mehreren Schichten sehr kompliziert und auch unlösbar werden. Die gesuchten numerischen Werte werden schlussendlich aus der Wahrscheinlichkeitsverteilung abgeleitet. Autoencoder als Alternative berechnet die deterministischen numerischen Werte direkt und effizient und wurden als geeignetes Analyseverfahren identifiziert.

Autoencoder sind eine besondere Form des unüberwachten Lernens und versuchen die gegebenen Eingaben möglichst exakt an der Ausgabe zu rekonstruieren. Ein Autoencoder besteht dabei aus zwei Teilen: Dem Encoder und dem Decoder. Der Encoder projiziert die gegebenen Daten auf eine latente Repräsentation dieser Daten und der Decoder rekonstruiert aus dieser Repräsentation wiederum die eigentlichen Daten. Damit der Autoencoder nicht nur die Identitätsfunktion erlernt, kann dieser durch diverse Techniken eingeschränkt werden, um relevante Abhängigkeiten zu erlernen.

Die typischste Einschränkung ist, die Dimension der latenten Repräsentation zu reduzieren. So ist der Autoencoder gezwungen, nur relevante Daten in dieser Schicht abzubilden. Eine weitere wichtige Form ist der Denoising Autoencoder. Diese Variante fügt den gegebenen Daten Rauschen hinzu und versucht diese teilweise zerstörten Daten wiederherzustellen.

### Anwendung von Analyseverfahren

Das sukzessive Ausfüllen von Dialogen kann als Analogie zur Entfernung von Rauschen aus Daten angesehen werden. Das initiale Autoencoder-Modell orientiert sich an der Arbeit von Wu et al. (2016). Diese Arbeit nutzt ebenso implizites Feedback und stützt sich auf binäre Daten.

Um eine konsistente Vergleichbarkeit zwischen Testläufen zu schaffen, wurden die Daten in drei Mengen aufgeteilt: Training, Test und Validierung. Mit dem Training- und Testdatensatz wurde das Autoencoder-Modell trainiert und mit diesem trainierten Modell wurden die Metriken Precision und Recall aus dem Validierungsdatensatz berechnet. Zusätzlich wurden die Metriken mehrfach berechnet und gemittelt, um statistische Ungenauigkeiten zu eliminieren. Mit diesem Vorgehen ist stets eine Vergleichbarkeit und Reproduzierbarkeit gegeben.

Mit Hilfe der Metriken Precision und Recall wird das initiale Modell von Wu et al. (2016) kontinuierlich bewertet und durchläuft mehrfach die Phase der Evaluation. Das Modell wird somit mehrfach überprüft und erweitert.

Dieses initiale rudimentäre in Python implementierte Autoencoder Modell liefert bereits mit einer Precision von 77% und einem Recall von 69% ordentliche Ergebnisse und übertrifft Metriken, welche mit typischerweise genutzten Datensätzen wie dem des MovieLens erreicht werden.

Tabelle 2-1: Empfehlungsgüte mit verschiedenen Aktivierungsfunktionen und Optimierungsalgorithmen.

	Sigmoid			ReLu		
	Epochen	Precision	Recall	Epochen	Precision	Recall
<b>SGD</b>	4999	0.777	0.688	4999	0.787	0.692
<b>AdaGrad</b>	3000	0.862	0.834	700	0.871	0.833
<b>Adam</b>	360	0.875	0.846	170	0.878	0.844

Durch das Ersetzen des SGD Optimierungsalgorithmus durch den AdaGrad können diese Metriken noch auf 86% Precision und 83% Recall verbessert werden. Wird statt des AdaGrad der Adam Optimizer genutzt können vergleichbare Werte in einem Zehntel der Epochen und damit weit effizienter erreicht werden.

Wut et al. (2016) nutzen die Sigmoid Funktion als Aktivierungsfunktion sowohl für den Encoder als auch den Decoder. Ersetzt man die Sigmoid Funktion des Encoders durch die bekannte ReLu Funktion kann nach 170 Epochen eine Precision von 87% und ein Recall von 84% errechnet werden.

Ein Einfluss auf die Precision oder den Recall durch das Ersetzen der Sigmoid Funktion des Decoders kann nicht beobachtet werden. Weitere beeinflussende Parameter sind die Anzahl Neuronen des Hidden Layers oder die genutzte Batchgröße. Eine engmaschige Untersuchung dieser beiden Parameter konnte indessen keinen Einfluss auf die Empfehlungsgüte des Systems offenbaren.

Die vorherigen Beobachtungen haben gezeigt, dass trotz der ordentlichen Ergebnisse einige Validierungsdatensätze Ausreißer darstellen. Die Vermutung ist, dass durch eine künstliche Erhöhung der mit Rauschen belegten Daten durch Kopieren die Varianz verringert wird. Es kann gezeigt werden, dass bereits durch vier zusätzliche Datensätze der Recall auf 87% bei einer Precision von 91% gesteigert werden kann. Der Grad des Rauschens als Analogie zu der Anzahl der Felder, die ein Nutzer bereits getätigt hat, ist auch als Parameter zu nennen. Eine prozentuale Gleichverteilung des Rauschens führt auch zu einer gleichmäßigen Güte der vorgeschlagenen Empfehlungen bei unterschiedlich vielen getätigten Eingabe.

Vergleicht man die Validierungsergebnisse aller zehn ausgewählten Unternehmen, dann stellt man fest, dass neben der Anzahl zur Verfügung stehender Daten die Korrelation zwischen den Dialogeingaben als maßgebende Eigenschaft in Bezug auf die gewählten Metriken hervorzuheben ist. Die Spärlichkeit der Daten, welche oft als beschreibende Eigenschaft genutzt wird, hat in diesem Kontext keine Relevanz.

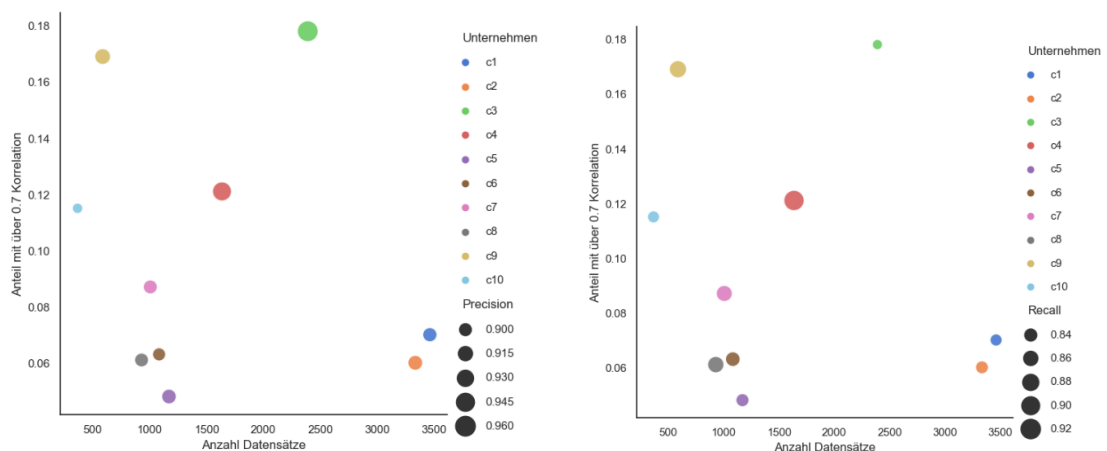


Abbildung 2.2 Abhängigkeit der Precision/Recall von der Anzahl Daten und Dialogeingaben mit einem Cramérs V von über 0.7.

Neben der mittleren Korrelation über alle Eingaben ist auch die Verteilung der Korrelation wichtig für die Qualität der Empfehlungen. Die gezielte Untersuchung von wenig getätigten Eingaben hat gezeigt, dass eine minimale Menge von Dialogfeldern existiert, die den restlichen Inhalt der Zollanmeldung mit einer Precision von 96% und einem Recall von 93% bestimmt. Die Heatmap der Korrelation kann hier als erklärendes Element hinzugezogen werden.

Die gezielte Eingabe der am stärksten korrelierenden Felder kann somit die Auswirkungen des Kalt-Start Problems reduzieren und den Eingabeprozess beschleunigen. Eine Untersuchung über mehrere Monate zeigte zwar eine leichte Abnahme der Metriken, aber nicht in einer Stärke, die auf ein mögliches Overfitting hinweisen würde. Die messbare Korrelation zwischen Dialogfeldern kann ebenfalls als Indikator für die Erklärbarkeit von Empfehlungen dienen und die Transparenz des Empfehlungssystems erhöhen.

## 2.4 Nutzbarmachung

Im Rahmen des Forschungsprojektes konnte somit durch iterative und strukturierte Erweiterung einer bereits existierenden Forschungsgrundlage ein Empfehlungsdienst entwickelt werden, welcher in der Lage ist, mit hoher Genauigkeit Empfehlungen für stark korrelierende Dialogfelder, wie die der Zollanmeldung zu generieren. Dieses Projekt war somit ein Erfolg und wird in Folgeprojekten in die Nutzbarmachung überführt und in die Systemlandschaft integriert.

## Literaturverzeichnis

Wu, Y., DuBois, C., Zheng, A. X. & Ester, M. (2016). Collaborative Denoising Auto-Encoders for Top-N Recommender Systems. In Proceedings of the Ninth ACM International Conference on Web Search and Data Mining (WSDM '16). Association for Computing Machinery, New York, NY, USA, 153–162

# 3 Entwicklung eines Machine-Learning-Modells zur Warenplanung in der Supply Chain

Jonas Dieckmann, Daniel Badura

## 3.1 Domäne & Projektauftrag

Die SCHRAMME AG ist führender Anbieter von Verbänden, Pflastern und Bandagen. Das Management sieht qualitatives Optimierungspotential sowie Einsparmöglichkeiten innerhalb ihrer Warenplanung und den resultierenden Produktionsprozessen. Die Geschäftsführung beauftragt einen internen Projektleiter mit der Aufgabe, ein Machine-Learning basiertes Modell für die Waren- und Bedarfsplanung innerhalb der Supply Chain zu entwickeln. Aufgrund negativer Erfahrungen aus vorherigen Data Science Projekten wird vorgeschlagen, das bevorstehende Projekt erstmals mithilfe eines Vorgehensmodells zu entwickeln. Die Wahl ist dabei auf das DASC-PM gefallen, um sowohl ein strukturiertes als auch ein wissenschaftliches Vorgehen für das Projektmanagement sicherzustellen. Um sich einen Überblick über den Projektauftrag zu verschaffen, arbeitet der Projektleiter zunächst verschiedene Use-Cases heraus, die anschließend auf Eignung und Umsetzbarkeit geprüft werden. Aus den geeigneten Use-Cases wird dann die konkrete Problemstellung sowie die Projektausgestaltung herausgearbeitet. Diese Ausgestaltung sollte abschließend noch einmal auf Eignung und Umsetzbarkeit geprüft werden.

### Auslöser und Use-Case-Entwicklung

Aktuell werden die über 2.500 verschiedenen Produkte des Unternehmens manuell geplant und anschließend produziert. In den letzten Quartalen kam es immer häufiger vor, dass einige Produktreihen häufig zu wenig Bestand vorgewiesen haben, andere Einzelprodukte die Lagerkapazitäten wiederum deutlich überstiegen. Während das Controlling die steigenden Lagerkosten durch ungenaue Planung bemängelt, beklagen die Warenplaner\*innen hingegen mangelnde Zeit für die Planung. Die Leiterin der Supply Chain kritisiert schon seit einiger Zeit die Tatsache, dass die Planungsaktivität ausschließlich manuell stattfindet und die Chancen der Digitalisierung scheinbar ungenutzt bleiben.

Als Ziel des Projektes wird demnach die Entwicklung eines Machine-Learning-Modells herausgearbeitet, wodurch ein Großteil des Produktbedarfs in Zukunft anhand verschiedener Einflussfaktoren automatisch geplant werden sollen. Die Warenplaner\*innen sollen sich somit vermehrt auf die Planung wichtiger Produktgruppen und die Werbeplanung kümmern. Das System soll Saisonalitäten, Trends und Marktentwicklungen berücksichtigen, sowie eine Planungsgenauigkeit von 75% erreichen. Das bedeutet, dass die Mengenprognosen jedes einzelnen Produktes maximal 25% von dem tatsächlichen Bedarf abweichen sollen. Als potenzielle Datenquellen kommen Bestellhistorien, Bestands- und Abverkaufszahlen der Kunden und interne Werbepläne in Frage. Neben der Beteiligung des Fachbereichs Supply Chain wird vor allem auch eine enge Zusammenarbeit mit dem Vertrieb und der IT erwartet.

Das Planungsteam innerhalb der Supply Chain besteht derzeit aus einem globalen Market-Demand-Planning-Team, welches sich basierend auf Marktentwicklungen, Produktlebenszyklen und strategischen Ausrichtungen um die langfristige Warenplanung (6 - 18 Monate) kümmert. In den einzelnen Märkten existieren lokale Customer-Demand-Planning-Teams, die für die entsprechenden Vertriebskanäle die kurzfristige Waren- und Werbeplanung (0 - 6 Monate) für den Handel umsetzen. Das zu entwickelnde Data-Science-Modell soll die monatlichen Planungszyklen unterstützen und dabei sowohl den kurzfristigen als auch den langfristigen Warenbedarf quantifizieren. Die Projektion wird anschließend in die interne Planungssoftware geladen und soll dort analysiert und bei Bedarf ergänzt oder korrigiert werden. Die endgültige Planmenge wird schlussendlich von den Fabriken für die Produktionsplanung verwendet. Zur Berücksichtigung von kunden- und produktspezifischem Fachwissen, besonderen Saisonalitäten und Erfahrungen der Vergangenheit sollen einzelne Teammitglieder der Planungsteams in das Projekt in bis zu 20% ihrer Arbeitszeit mit einbezogen werden.

Ein wichtiger Teilaspekt während der Use-Case-Auswahl ist die Eignungsprüfung. Der Projektmanager versucht daraufhin zu untersuchen, ob das Projekt grundsätzlich als durchführbar eingestuft werden kann und ob die Anforderungen mit den gegebenen Ressourcen durchgeführt werden können. Mittels Experteninterviews hat sich gezeigt, dass sich die Problemstellung im Allgemeinen sehr gut für den Einsatz von Data Science eignet und entsprechende Projekte bereits extern durchgeführt und auch publiziert wurden. Das Data Science Team hat währenddessen bestätigt, dass ausreichend potenziell geeignete Methoden für dieses Projekt existieren und erforderliche Datenquellen zur Verfügung stehen.

Abschließend nimmt der Projektmanager eine Analyse der Umsetzbarkeit vor. Neben Abstimmungen mit der IT zur Prüfung der vorhandenen Infrastruktur muss auch die Expertise der beteiligten Mitarbeiter geprüft werden. Durch die vorhandene Cloud-Infrastruktur von Microsoft und der Erfahrung des Data Science Teams mit der Software Databricks scheint das Projekt grundsätzlich realisierbar. Das Projektrisiko wird generell als moderat eingestuft, da die Planer insbesondere in der Implementierungsphase eine starke Kontrollfunktion einnehmen und die Ergebnisse prüfen werden.

## **Projektausgestaltung**

Basierend auf der Problemstellung und den Domänenspezifika obliegt es nun dem Projektmanager gemeinsam mit der Leiterin Supply Chain sowie einem Data Scientist das Projekt formal zu ausgestalten. Dabei wird als Projektziel die Verbesserung der Planungsgenauigkeit sowie die Reduktion der manuellen Vorgänge angenommen und mit der Projektzielsetzung verknüpft, ein entsprechendes Modell für das Vorhaben zu entwickeln. Der Kostenrahmen beträgt nach anfänglicher Schätzung 650.000 Euro. Als Zeitrahmen für die Entwicklung wurde eine Dauer von sechs Monaten veranschlagt, wobei anschließend weitere sechs Monate für die Prozessintegration vorgesehen werden.

Da eine vollständige Planung und Beschreibung des Ablaufes von Projekten im Data-Science-Kontext im Gegensatz zu vielen anderen Projekten in der Regel nicht möglich ist, erstellt der Projektmanager für dieses Vorgehen lediglich eine Projektskizze mit grundsätzlichen Eckpfeilern, die bereits in den vorherigen Abschnitten genannt wurden. Das Budget umfasst finanzielle Ressourcen für 1 VE Projektmanager, 2 VE Data Scientist und 0,5 VE Data Engineer. Wie bereits genannt, sollen die Warenplaner\*innen etwa 20% ihrer Arbeitszeit bereitstellen, um Fachwissen und Erfahrungen zu teilen.

Insgesamt soll das Projekt in agiler Arbeitsweise und anhand der Phasen des DASC-PM in Scrum-Methodik erarbeitet werden. Dabei wird sich iterativ durch die Bereiche Datenbereitstellung, Analyse, Nutzbarmachung und Nutzung gearbeitet, wobei aus jeder Phase heraus sowohl die vorherige als auch die nachfolgende Phase in den Fokus gerückt werden kann. Die Rückschritte sind insbesondere dann wichtig, wenn in Schlüsselbereichen Lücken oder Probleme festgestellt werden, die nur durch eine erneute Durcharbeitung der vorherigen Phase lösbar werden. Die Projektskizze wird visuell aufbereitet und in den Räumlichkeiten der SCHRAMME AG für alle Beteiligten gut sichtbar platziert. Abschließend wird der gesamte Projektauftrag noch einmal auf Eignung und Umsetzbarkeit überprüft, bevor das Vorgehen in die nächste Phase übergeht.

## 3.2 Datenbereitstellung

### Datenaufbereitung

Die SCHRAMME AG verfügt über eine Vielzahl an Datenquellen, welche in die automatisierte Planung einfließen können. Neben historischen Verkaufsdaten aus dem ERP-System kommen auch Bestellhistorien und Kundendaten aus dem CRM-System in Frage, sowie Lagerbestände und Marketingmaßnahmen. Mithilfe von Azure Data Factory wird eine cloudbasierte Pipeline erstellt, welche die Daten aus den verschiedenen Quellsystemen lädt, transformiert und integriert. Als Grundlage für die automatischen Forecasts sollen in erster Linie die Bestellhistorien dienen. Die restlichen Daten werden entweder als Hintergrundinformationen für die Planungsteams genutzt oder um bei Bedarf im Vorfeld Clusteranalysen durchzuführen. In der Anfangsphase des Projektes weisen die einzelnen Datenquellen noch große Unterschiede in Hinblick auf Qualität und Struktur auf. Deshalb werden in Zusammenarbeit mit IT und Fachbereichen Anpassungen vorgenommen, um die Forecasts später auf einer soliden Grundlage durchführen zu können.

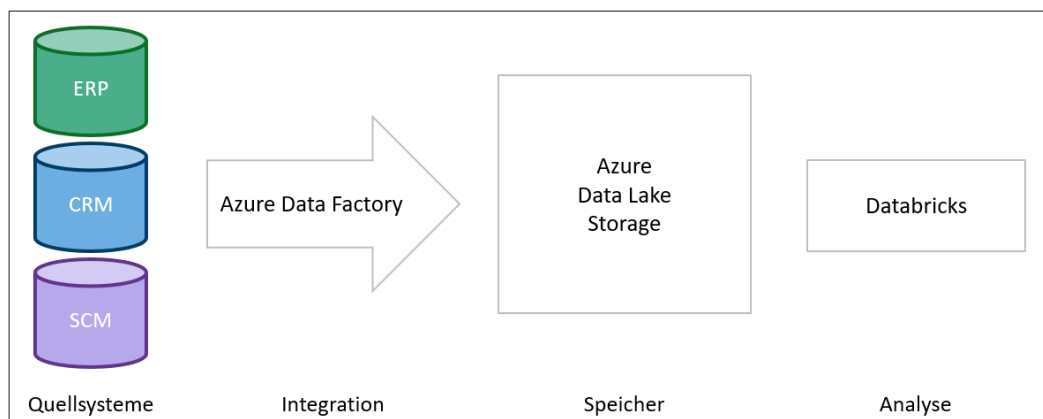


Abbildung 3.1: ELT Prozess von Datenaufbereitung zur Analyse.

### Datenmanagement

Der Prozess des Datenmanagements wird von den Data Engineers automatisiert und nach einem täglichen Zeitplan ausgeführt, um immer auf dem neuesten Stand zu bleiben. Um die Komplexität in einem annehmbaren Rahmen zu halten, werden zunächst die vielversprechendsten Datenquellen verarbeitet und die Pipeline dann mit Continuous Integration / Continuous Deployment (CI/CD) inkrementell ausgebaut. Nach dem Deployment werden die verarbeiteten Daten in Azure Data Lake Storage gespeichert, wo sie zur weiteren Analyse mit Azure Databricks verwendet werden können.

Im Data Lake werden zudem Backups der aufbereiteten Daten und Analyseergebnisse sowie weitere Daten wie Protokolle, Qualitätsmetriken und Berechtigungsstrukturen gespeichert. Über Schreib- und Leseberechtigungen sowie Planungsversionen wird zudem sichergestellt, dass immer nur die aktuelle Planungsperiode bearbeitet werden kann, damit Werte aus der Vergangenheit sich nicht mehr verändern.

## Explorative Datenanalyse

Ein wichtiger Schritt in der Datenaufbereitung stellt die explorative Datenanalyse (EDA) dar, bei der zunächst verschiedene Statistiken und Visualisierungen erzeugt werden. Es entsteht ein Überblick über Verteilungen, Ausreißer und Korrelationen in den Daten. Die Ergebnisse der EDA liefern Aufschluss darüber, welche Merkmale potenziell für die nächste Phase der Analyse in Betracht zu ziehen sind. Im zweiten Schritt werden mithilfe von Feature Selection und Feature Engineering relevante Merkmale ausgewählt oder neue Features erzeugt. Bei Daten mit hoher Dimensionalität wird dazu eine Methode der Dimensionsreduktion angewandt, wie z. B. eine Hauptkomponentenanalyse. Die EDA gibt darüber hinaus Aufschluss über die existierenden Bedarfshistorien der SCHRAMME AG.

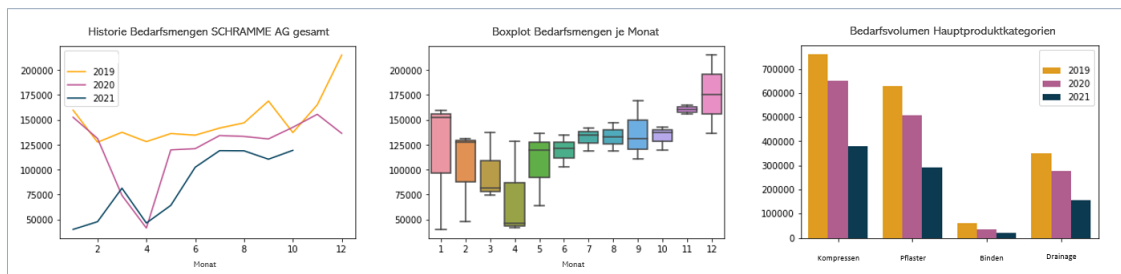


Abbildung 3.2: Exemplarische Ergebnisse der explorativen Datenanalyse.

## 3.3 Analyse

### Identifikation geeigneter Analyseverfahren

Basierend auf der Prüfung der Umsetzbarkeit zu Beginn des Projektes wurde deutlich, dass das vorliegende Projekt in der Tat mit Data-Science-Methoden gelöst werden kann und soll. Die beiden involvierten Data-Science-Mitarbeiter verschaffen sich daher zunächst einen Überblick über existierende Verfahren, die gut zu dem zugrundeliegenden Problem passen können.

Die vorliegende Problemstellung gehört grundsätzlich in die Problemklasse Regression innerhalb der Supervised-Learning-Algorithmen. Grundlegend handelt es sich hier um eine Art Zeitreihenanalyse, die sich durch zusätzliche Faktoren oder Multiple Regression erweitern lässt. In Verbindung des Schlüsselbereiches Wissenschaftlichkeit, wurde der aktuelle Forschungsstand zu vergleichbaren Problemstellungen durchgeführt. Diese hat ergeben, dass XGBoost, ARIMA, Facebook Prophet und LightGBM häufig genannte Methoden für die Problemklasse sind. Ein Data Scientist dokumentiert die entsprechenden Vor- und Nachteile der jeweiligen Methode und sortieren sie anhand ihrer Komplexität und Rechenintensität. Um erste Indikationen über die Modellierbarkeit der Produkte der SCHRAMME AG zu erhalten, werden zunächst einfachere Modelle selektiert, woraufhin das Projektteam mit klassischem exponentiellem Glätten und der Modellfamilie ARIMA beginnen wird.



## Anwendung von Analyseverfahren

Da in dem vorliegenden Projekt mehrere Anwender am Analyseverfahren beteiligt sind, setzt das Team zunächst eine geeignete Notebook basierte Entwicklungsumgebung in Databricks auf. Entlang des typischen Machine Learning Workflows wird zunächst der Code für den Import und das Data Cleaning implementiert. Zur Sicherstellung der Validität wird das zugrundeliegende Datenset mittels Cross Validation schließlich in Trainings-, Validierungs- und Testdaten unterteilt. Trainings- und Validierungsdatensätze werden anschließend auf die ausgewählten Verfahren angewandt, um das Modell zu optimieren. In dem Zuge wird auch wiederholend versucht, die Parameter der Verfahren zu optimieren, sowie ggf. die Anzahl der vorliegenden Dimensionen sinnvoll zu reduzieren.

Die Data Scientists der SCHRAMME AG dokumentieren die Durchführung und die Validierungsergebnisse der einzelnen Durchläufe. Es zeigt sich, dass die Modelle der ARIMA Familie grundsätzlich eine bessere Performance gegenüber dem exponentiellen Glätten aufweisen, wenn auch die Zielgenauigkeit von 75% mit einem aktuell erreichten Wert von 62,4% noch nicht erreicht werden kann. Die Metriken RMSE und MAPE zeigen ebenfalls noch Optimierungspotential.

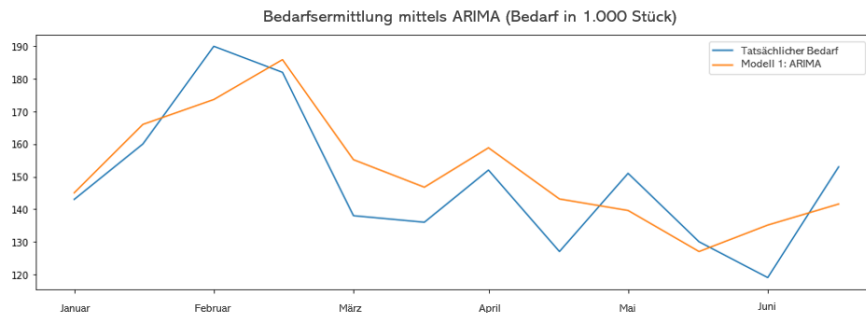


Abbildung 3.3: Vergleich des ARIMA Forecasts gegenüber tatsächlichem Bedarf.

Die Parameterkonfiguration und die Begründung der Auswahl des finalen Modells der ersten Anwendungsiteration werden festgehalten und für den Projektleiter sowie die Leiterin Supply Chain fachlich verständlich aufbereitet. Insbesondere wird aufgezeigt, dass einige Produktgruppen sehr ungewöhnliche Saisonalitäten aufweisen und bestimmte Produkte generell recht schlecht vorhersehbar sind. Wenn auch das Produktportfolio der SCHRAMME AG eher geringfügig durch die temporären Schließungen (Lockdowns) im Rahmen der Corona-Pandemie betroffen ist, zeigt sich ein leichter Rückgang des Bedarfs an Verbandsmitteln. Es wird angenommen, dass durch weniger Aktivität und Transport auch die Anzahl an Unfällen und Verletzungen zurückgegangen ist. Dieser Trend lässt sich in den verwendeten Analyseverfahren aber recht gut modellieren.

Um die Zielgenauigkeit zu verbessern, werden in einem weiteren Experiment nun technisch aufwändigere Verfahren eingesetzt, welche sich im Rahmen der Identifikation geeigneter Methoden als relevant und einsetzbar erwiesen haben. Nach einigen Iterationen der Parameteroptimierung und Kreuzvalidierung haben die Verfahren Prophet und XGBoost mit 73,4% und 65,8% die höchsten Validierungsergebnisse hervorgebracht. Die Data Scientists stellen Prophet als geeignetste Methode unter den bisher angewandten Verfahren heraus und ermitteln die Planungsgenauigkeit gegenüber der Testzeitreihe. Wenn auch die Genauigkeit mit 73,4% leicht unter dem Zielwert liegt, ist eine signifikante Verbesserung der Planungsgenauigkeit gegeben. Der MAPE liegt bei 16,64% und der RMSE bei 8.130, was im Vergleich zum RMSE der XGBoost Methode (10.134) eine geringere absolute Ab-

weichung impliziert. Ähnlich dem ersten Experiment zeigen sich allerdings Produktgruppen, die insgesamt sehr schwach vorgesagt werden (37,2%) und die kumulative Genauigkeit somit negativ beeinflussen.

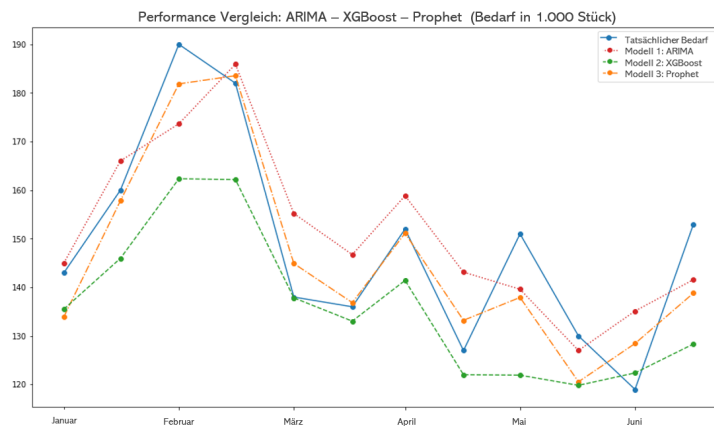


Abbildung 3.4: Performance-Gegenüberstellung der verschiedenen Verfahren.

### Evaluation

Basierend auf den Ergebnissen der Analysen erfolgt anschließend eine nachvollziehbare Bewertung sowie Einordnung durch die Leiterin der Supply Chain und den Analysten, welche durch den Projektmanager organisiert und moderiert wird. Als Metriken zur Evaluation wird die bereits im Vorweg festgelegte kumulative Planungsgenauigkeit aller Produkte in Kombination mit den gängigen Metriken RMSE und MAPE angenommen. Für den Fachbereich ist es hier von großer Wichtigkeit, eine realistische, nachvollziehbare und zuverlässige Bedarfsermittlung auf Produktebene zu erhalten.

Tabelle 3-1: Evaluation der drei besten Modelle.

	Genauigkeit	RMSE	MAPE
ARIMA	62,4 %	11.437,6	21,8%
XGBoost	65,8 %	10.134,2	19,1%
Prophet	73,4 %	8.130,8	16,6%

Tabelle 3-2: Evaluation des besten Modells, verteilt nach Produktgruppen

Prophet	Genauigkeit	RMSE	MAPE
<b>Total</b>	<b>73,4 %</b>	<b>8.130,8</b>	<b>16,6%</b>
Kompressen	81,9 %	6.923,3	11,3 %
Verbandmittel	86,3 %	6.027,9	9,3 %
Drainagen	37,2 %	21.738,2	62,5 %
Binden	70,7 %	8.345,4	18,4 %
Pflaster	92,8 %	2.173,3	6,3%

Als Benchmark der Planungsgenauigkeit kann die aktuell (manuell geplante) mittlere Genauigkeit der letzten zwei Jahre von 58% angenommen werden. In der Ergebnisevaluation zeigt sich, dass viele Produktgruppen insgesamt eine sehr hohe Planungsgenauigkeit mithilfe des Data Science Modells aufweisen und den Benchmark weit übersteigen. Es gibt allerdings auch Produktgruppen, die eine ähnliche Genauigkeit gegenüber der manuellen Planung zeigen. Diskutiert werden muss vor allem der Produktbereich *Drainagen*, der mithilfe des Modells deutlich schlechter als der händische Plan abschneidet und unter den bisher verwendeten Methoden eher ungeeignet für eine statistische Bedarfsermittlung erscheint.

Aus fachlicher Sicht ergibt es laut der Leiterin Supply Chain wenig Sinn, solche Produktgruppen statistisch zu planen, die aufgrund ihrer saisonalen oder trendbedingten Spezifika nur eine geringe Planungsgenauigkeit aufweisen. Sie schlägt vor, einen Fehlerschwellenwert auf Produktbasis einzuführen, wodurch festgelegt wird, welche Produkte mithilfe des Modells vorhergesagt werden sollen und welche Produktgruppen aus der Modellierung herausgenommen und entsprechend weiterhin manuell geplant werden. Als geeigneter Schwellenwert wird ein Bereich leicht unter dem aktuellen Benchmark gesehen, da ähnlich gute Genauigkeit bei weniger manuellem Aufwand aus Sicht des Fachbereiches immer noch eine Verbesserung zur Erreichung des Projektziels darstellt. Der Projektleiter dokumentiert die Ergebnisse der Evaluation mit den getroffenen Entscheidungen und Maßnahmen.

Als Analyseergebnis lassen sich nach der ersten realen Modellierung die Bedarfsmengen aller ausgewählten Produkte für die nächsten 18 Monate festhalten. Diese gilt es nun nutzbar zu machen und in den Planungsprozess der Teams zu integrieren.

## **3.4 Nutzbarmachung**

Für die Integration geht das Team nun in die Phase der Nutzbarmachung des DASC-PM über.

### **Technisch-methodische Bereitstellung**

Bei der Nutzbarmachung kann auf die bestehende Infrastruktur gesetzt werden. Die Forecasts werden in die Planungssoftware IBM Planning Analytics geladen, wo sie geprüft und weiterbearbeitet werden. Zur Automatisierung der Ladeprozesse wird der sogenannte TurboIntegrator genutzt, der einen zentralen Bestandteil von IBM Planning Analytics bildet. Über die OLAP-Struktur von Planning Analytics können flexible Ansichten erstellt werden, in denen die Anwender ihren Kontext (Zeitbezug, Produktgruppen etc.) selbst wählen können und Berechnungen sich in Echtzeit anpassen. Für tieferegehende Analysen wird zudem die Reporting-Software QlikSense angebunden. Hier können zum einen die Komponenten der Zeitreihen (Trends, Saisonalität, Rauschen) visualisiert und zum anderen weitere Informationen wie Ausreißer und Mittelwerte angezeigt werden. Die fertigen Pläne werden im Anschluss an die Bearbeitung durch die Planungsteams wieder in den Data Lake geladen, um zukünftig abrufbar zu sein.

### **Sicherstellung technischer Umsetzbarkeit**

Die Forecasts selbst werden zu Monatsbeginn automatisch neu generiert. Die Planer können ihre Korrekturen während der ersten vier Arbeitstage des Monats vornehmen und die Resultate in Echtzeit im Planungssystem sehen. Da die Algorithmen in einer Cloudumgebung arbeiten, kann bei Bedarf die Rechenleistung skaliert werden. Damit alle Prozesse automatisiert laufen können, sollen Veränderungen bei den Datenquellen minimiert werden. Gibt es Anpassungsbedarf, wird der Data Engineer darüber informiert und das Schnittstellendokument aktualisiert, in dem alle Informationen zu Datenquellen und -verbindungen festgehalten werden. Das Planungs- und Forecastingsystem ist somit eine Mischung aus Cloud (Microsoft Azure) und On-Premise (Planning Analytics), wobei die Planer nur aktiv auf die On-Premise-Strukturen zugreifen. Hier werden Berechtigungen vergeben, wodurch die lokalen Planer nur Zugriff auf ihre Bereiche haben, wogegen die globalen Planer Einblick in alle Themen bekommen. Nach Abschluss der Entwicklungsphase wird der Support hauptsächlich von der IT übernommen. Bei komplexen Problemen werden auch Data Scientist oder Data Engineer eingeschaltet.

## Anwendbarkeitssicherstellung

Adressaten der Lösung sind die lokalen und globalen Planungsteams. Da die Mitglieder der Teams eher fachlich als technisch orientiert sind, werden Schulungen veranstaltet, die ihnen dabei helfen sollen, die Forecasts zu interpretieren und ihre Qualität einzuordnen. Auch die Benutzeroberfläche wird mit einem Fokus auf Übersichtlichkeit und Verständlichkeit gestaltet. Es werden einfache Lini- und Balkendiagramme für Verläufe und Benchmarks sowie auf das Wesentliche reduzierten Tabellen eingesetzt. Die Anwender werden von Beginn an in die Entwicklung mit einbezogen, um die fachliche Richtigkeit und Relevanz sicherzustellen und um vor Ende der Entwicklungsphase bereits für Vertrautheit mit der Lösung zu sorgen. Zusätzlich wird eine umfassende Dokumentation erstellt. Der technische Teil der Dokumentation baut dabei im Wesentlichen auf dem Schnittstellendokument auf, in dem Datenstrukturen und -verbindungen festgehalten sind, während der inhaltliche Teil gemeinsam mit den Anwendern erarbeitet wird.

## Fachliche Bereitstellung

Um sicherzugehen, dass die neue Lösung nicht nach wenigen Monaten an Relevanz oder Qualität verliert, wird auch nach Abschluss der ersten Entwicklungsphase weiterhin an Verbesserungen gearbeitet, wenn auch mit deutlich reduziertem Zeitaufwand. Der wichtigste Aspekt der laufenden Verbesserung ist die ständige Anpassung des Vorhersagemodells an neue Daten, die automatisiert abläuft. Andere Bestandteile des Systems, die zu Beginn noch manuellen Aufwand erfordern, werden im Laufe der Zeit ebenfalls automatisiert. Die Änderung verschiedener Parameter, wie Forecast-Horizont oder Schwellwerte für die Genauigkeit der Vorhersagen, können von den Planern selbst in Planning Analytics vorgenommen werden, wodurch das Modell flexibel bleibt. Probleme, die nach dem Release der ersten Version auftreten, werden über das IT-Ticketsystem erfasst und dem Data-Science-Bereich zugeordnet. In regelmäßigen Abständen wird zudem überprüft, ob das Modell noch die Ansprüche des Unternehmens erfüllt, oder Änderungen nötig werden.

## 3.5 Nutzung und Fazit

Mit dem Übergang in die Nutzung des entwickelten Modells geht das Data-Science-Vorgehensmodell DASC-PM in seine letzte Phase. Insgesamt konnte die SCHRAMME AG mithilfe des strukturierten und ganzheitlichen Ansatzes die selbstgesteckten Ziele innerhalb des Bereichs Supply Chain erreichen. Von der Nutzung ausgehend lassen sich nun weitergehende oder neue Projektaufträge ableiten. Die Planungsvorgänge wurden weitestgehend automatisiert und mittels Machine-Learning-Algorithmen unterstützt. Dabei ist ein hoher Zufriedenheitsgrad bei den relevanten Stakeholdern aus Management, Finance und Supply Chain entstanden. Das Planungsteam selbst ist nach anfänglicher Skepsis nun ebenfalls von der Entlastung und möglichen Priorisierung überzeugt. Es ist aber auch denkbar, dass während der Nutzung Schwachstellen auftreten und eine weitere Iteration in vorherigen Phasen erfordern.

Insgesamt hat die Case Study gezeigt, dass insbesondere nichtlineare Vorgehensmodelle für den Bereich der Data Science vorteilhaft sind. Das DASC-PM stellt dabei ein geeignetes und neuartiges Vorgehen dar, welches sich ebenfalls auf zahlreiche andere Domänen und Problemstellungen transferieren lässt.

# 4 Projekt FLEMING – Predictive Maintenance von zentralen Komponenten des Mittelspannungsnetzes

Philipp zur Heiden

## 4.1 Einleitung

In der Politik und Gesellschaft sind die Energie- und Mobilitätswende zwei der aktuell kontrovers diskutierten Themen. Die Energiewende beschreibt dabei die Umstellung von fossilen Energieträgern (vor allem Kohle und Gas) zu regenerativen Alternativen im deutschen Stromnetz (Schiffer 2019; Renn und Marshall 2016). Dazu zählt auch, dass das Energienetz sich von einem unidirektionalen Netzwerk zu einem bidirektionalen, dezentralen Netzwerk mit schwankend generierten Energiemengen entwickeln wird (Bundesnetzagentur 2019). Unter der Mobilitätswende wird der Umstieg von Fahrzeugen, die durch fossile Energieträger angetrieben werden, hin zu alternativen, emissionslosen Antrieben verstanden, welche einen großen Bedarf an elektrischer Energie hervorrufen. Eine zentrale Herausforderung für das Verteilnetz ergibt sich daher bei der Beanspruchung von zentralen Komponenten, bspw. Lasttrennschaltanlagen in Mittelspannungsnetzen. Hier bedarf es daher einer adaptierten Instandhaltungsstrategie, damit möglichst wenige Ausfälle im Mittelspannungsnetz entstehen (Hoffmann et al. 2020).

In Theorie und Praxis sind drei verschiedene Instandhaltungsstrategien verbreitet (vgl. im Folgenden Mobley 2002): Reaktive Instandhaltung, präventive Instandhaltung und prädiktive Instandhaltung (engl.: Predictive Maintenance). Bei der reaktiven Instandhaltung werden relevante Komponenten nicht gewartet oder instandgehalten, solange sie noch funktionieren. Erst zum Zeitpunkt des Ausfalls oder einem Defekt der Komponente wird diese ausgetauscht oder repariert. Während die reaktive Instandhaltung die Lebensdauer von Komponenten maximal ausreizt, entstehen Ausfälle, die vor allem im Energienetz eine Unterbrechung der Stromzufuhr bedeuten und demnach schwerwiegende Folgen für z. B. Krankhäuser oder die Lebensmittelindustrie haben können. Die präventive Instandhaltung setzt mit einer regelmäßigen Wartungsstrategie an diesem Punkt an, indem, basierend auf Zeiträumen oder Nutzungszahlen (z. B. Anzahl der Schaltungen), die Komponenten überprüft werden. So können zwar einige Ausfälle verhindert werden, allerdings sind Ausfälle weiterhin möglich und die Kosten deutlich höher als bei einer reaktiven Instandhaltungsstrategie. Predictive Maintenance kombiniert die Vorteile der beiden geschilderten Strategien. Durch permanente Zustandsüberwachung und Data-Science-Algorithmen kann der Status von Komponenten präzise abgebildet und vorhergesagt werden, sodass Wartungen erst dann eingeplant werden, wenn die maximale Lebenszeit von Komponenten erreicht ist, allerdings noch bevor Defekte auftreten und Ausfälle ausgelöst werden.

Ziel des in dieser Fallstudie vorgestellten Projektes ist die Entwicklung eines umfassenden Predictive Maintenance Systems für zentrale Komponenten des Mittelspannungsnetzes. Das Forschungsprojekt *FLEMING* (Flexible Monitoring- und Regelsysteme für die Energie- und Mobilitätswende im Verteilnetz durch Einsatz von Künstlicher Intelligenz) besteht dabei aus einem Konsortium von Projektpartnern, die diverse Bereiche für die erfolgreiche Durchführung eines Data-Science-Projektes ermöglichen, u.a. Projektmanager, Domänenexperten, Data Scientists und Data Engineers. Dazu zählen die Forschungseinrichtungen des Karlsruher Institut für Technologie (KIT), des FIR e.V. an der RWTH Aachen, der Software Innovation Campus Paderborn (SICP) sowie die Industriepartner ABB Forschungszentrum in Ladenburg und Heimann Sensor GmbH.

Das gesamte *FLEMING*-Projekt kann als ein Data-Science-Projekt verstanden werden, da Predictive Maintenance als Strategie der permanenten Zustandsüberwachung und -vorhersage verschiedenste Daten und ausgereifte Analysemethoden voraussetzt. In dieser Fallstudie wird der Projektauftrag von *FLEMING* nach dem DASC-PM beschrieben, wobei ein besonderer Fokus auf den Schlüsselbereich *Domäne* gelegt wird. Mit dem Mittelspannungsnetz spielen sich weite Teile des *FLEMING*-Projektes in einem besonderen Anwendungsgebiet für Data-Science-Projekte ab, welches gesondert betrachtet werden sollte. Die folgenden Abschnitte stellen daher jeweils ein Unterthema des Schlüsselbereichs *Domäne* aus dem DASC-PM im Detail dar: Problemstellung und Ziele, Beteiligte und Stakeholder, Projektorganisation und Ressourcen (kombiniert in einem Abschnitt) und Vorfahrungen. Dabei wird vor allem auf die Beantwortung der Fragen zu den wesentlichen Projektmerkmalen im Projektauftrag, die im Rahmen des DASC-PM als Fragenkatalog bereitgestellt werden, eingegangen. Für Data-Science-Projekte ist dieser Fragenkatalog nicht zwingend in der hier beschriebenen Detailtiefe auszufüllen, allerdings empfiehlt sich das gerade bei aufwändigen und großen Data-Science-Projekten. Final wird ein Ausblick auf die Projektergebnisse und das weitere Vorgehen im *FLEMING*-Projekt gegeben.

## 4.2 Problemstellung und Ziele

*Welche Problemstellung soll mit dem Projekt gelöst werden?*

Durch die Energie- und Mobilitätswende wird sich das Elektrizitätsnetz zeitnah von einem unidirektionalen Netz zu einem bidirektionalen, dezentralen Netz mit schwankenden Energiemengen wandeln (Bundesnetzagentur 2019). Daraus resultiert, dass die zentralen Komponenten des Mittelspannungsnetzes zukünftig deutlich stärker beansprucht werden. Beispielsweise kann dies ein Problem für Lasttrennschaltanlagen im Mittelspannungsnetz sein, welche in der Regel auf eine Laufzeit von bis zu 40 Jahren ausgelegt sind. Das *FLEMING*-Projekt soll daher das Problem lösen, dass der aktuelle Zustand der Lasttrennschaltanlagen unbekannt ist, d.h. die Funktionsfähigkeit und die Wahrscheinlichkeit eines Ausfalls oder einer Störung, und die Entwicklung des Zustands nicht eingeschätzt werden kann, da diverse externe und teilweise unvorhersehbare Faktoren den Zustand beeinflussen, z. B. Temperatur, Luftqualität, Luftfeuchtigkeit, Tiere und Pflanzen. Ausfälle in der Stromversorgung sind daher nicht abzusehen.

*Welche Ziele werden mit dem Projekt verfolgt?*

Die Ziele, die mit diesem Projekt verfolgt werden, lassen sich in mehrere Kategorien einteilen. Zunächst soll die beschriebene Problemstellung gelöst werden, nach welcher der aktuelle Zustand von Lasttrennschaltanlagen im Mittelspannungsnetz unklar und nicht vorhersehbar ist. Zudem soll für die Betreiber der Schaltanlagen der Aufwand und die Kosten für die Instandhaltung gesenkt werden, indem nicht mehr auf Ausfälle bzw. fixe Wartungszyklen, sondern auf Predictive Maintenance als Instandhaltungsstrategie gesetzt wird. Für die Hersteller der Schaltanlagen könnte dies auch eine Erschließung von neuen Geschäftsfeldern und Zielgruppen bedeuten.

*Welche Ergebnisse werden erwartet?*

Es wird erwartet, dass das Konsortium aus Forschungs- und Anwendungspartnern einen Prototyp für ein ganzheitliches Predictive-Maintenance-System entwickelt. Dazu zählt die Generierung von Daten zur Analyse des aktuellen Zustands und zur Vorhersage zukünftiger Zustände basierend auf innovativer Sensorik, die Entwicklung von Data-Science-Modellen, die Evaluation der Ergebnisse und die Veröffentlichung von Beiträgen auf wissenschaftlichen Konferenzen und in Fachzeitschriften.

*Wie wird Erfolg gemessen?*

Aufgrund des vornehmlich explorativen Ziels des Forschungsprojektes ist der Erfolg nicht durch KPIs oder vergleichbare Metriken zu messen. Stattdessen wird der Erfolg im Projekt durch die fortschreitenden Ergebnisse der einzelnen Projektpartner und die gewonnenen Erkenntnisse im Rahmen der gemeinsamen Treffen zur Projektabstimmung bestimmt, bspw. durch die prototypische Umsetzung und den Demonstrator.

*Worin besteht die Motivation, die Problemstellung durch ein Data-Science-Projekt zu adressieren?*

Bei Lasttrennschaltanlagen ist es möglich, dass auftretende Probleme schon vor dem Ausfall bemerkt werden, da verschiedene interne und externe Einflüsse Ausfälle oder Störungen herbeiführen können (bspw. erhöhte Temperaturen, durch Ruß verschmutzte Luft und Tiere in der Anlage). Dadurch könnten Ausfälle vorzeitig erkannt und im Idealfall verhindert werden. Die Zustandsüberwachung ist allerdings aufgrund der räumlich breit gefächerten Streuung der Anlagen eines Betreibers im Mittelspannungsnetz sehr kostenaufwändig und kann bei manueller Betrachtung fehleranfällig sein. Daher wird im Rahmen des Projektes neue Sensorik entwickelt, die die Schaltanlagen und deren Zustand zu überwachen erlaubt. Mithilfe dieser Zustandsdaten kombiniert mit weiteren Daten (bspw. von benachbarten Komponenten, geografisch zusammenhängenden Daten wie Luftverschmutzung und Wetter) ist davon auszugehen, dass Ansätze der Data Science zielführende Ergebnisse liefern werden.

*Welche projektnahen Ziele sollen definitiv nicht verfolgt werden?*

Das Projekt vertritt den Charakter eines vorwettbewerblichen Forschungsprojektes und verfolgt demnach nicht das Ziel, dass eine am Markt erbringbare und im Vergleich zu aktuell am Markt befindlichen Leistungen vergleichbare Leistung aus dem Projekt resultiert. Es soll somit definitiv keine vollwertige, komplett implementierte Lösung entstehen, sondern es soll im Rahmen von Prototypen ein Mehrwert für den Stand der Forschung in den verschiedenen Disziplinen, die am Projekt beteiligt sind, erreicht werden. Dazu zählt die Forschung zu Machine-Learning-Algorithmen, Sensorik, Lasttrennschaltanlagen und dem Design für Monitoring- und Instandhaltungssysteme.



## 4.3 Domänenspezifika

Um die Besonderheiten des *FLEMING*-Projekts zu verstehen, ist eine rudimentäre Einordnung des Projekts und der zugrundeliegenden Problemstellung in den Kontext der Mittelspannungsnetze notwendig. Das deutsche Stromnetz kann grundlegend in vier verschiedene Spannungsebenen eingeteilt werden: Höchstspannung (220-380 kV), Hochspannung (60-110 kV), Mittelspannung (6-30 kV) und Niederspannung (230-400 V) (BMW 2021; Kamper 2010). Während in den Ebenen der Höchst- und Hochspannung Kraftwerke und Erzeugungsanlagen Energie produzieren und nur wenige industrielle Großkunden Energie abnehmen, sind auf der Mittelspannungsebene durch Windkraft- und Photovoltaik-Anlagen sowie Industrie- und Bürogebäude eine Vielzahl verschiedener Produzenten und Konsumenten von Energie angesiedelt. Auch in der Niederspannung wird Energie in das Stromnetz eingespeist (bspw. durch Solardächer) und von Privatkunden konsumiert.

Die Energie muss nach der Produktion in Höchst- bzw. Hochspannung zunächst durch Umspannwerke auf die niedrigere Spannungsebene transformiert werden. Die Mittelspannung endet schließlich in der Netzstation (Siemens 2018). Unter Netzstation werden die Gebäude zusammengefasst, die in etwa der Größe einer Garage entsprechen und die Energie von der Mittelspannung auf mehrere Niederspannungsphasen transformiert. Innerhalb dieser Netzstationen befinden sich die zentralen Komponenten, die durch die zunehmende Energie- und Mobilitätswende besonders beansprucht werden (Hofstätter et al. 2012). Dazu zählt auch die Lasttrennschaltanlage in der Mittelspannung.

Kommt es aufgrund der steigenden Belastung der Komponenten zu einem Ausfall in einer Ortsnetzstation, kann dies auch einen Ausfall für die dahinterliegenden Energiekonsumenten der Niederspannung bedeuten, sodass schließlich die Stromversorgung kleinerer Unternehmen und Privathaushalte unterbrochen werden würde. Bei einer voranschreitenden Elektrifizierung der Mobilität wird die durchgehende Versorgung mit elektrischer Energie damit zukünftig noch wichtiger als sie derzeit schon ist.

## 4.4 Beteiligte und Stakeholder

*Welche Organisationseinheiten sind konkret involviert?*

Als Konsortialführer leitet das ABB Forschungszentrum Deutschland das *FLEMING*-Projekt. An dem Projekt sind als Forschungseinrichtungen das Institut für Elektroenergiesysteme und Hochspannungstechnik am KIT, das FIR e.V. an der RWTH Aachen sowie der Lehrstuhl für Wirtschaftsinformatik, insb. Betriebliche Informationssysteme und die Gruppe Intelligente Systeme am Software Innovation Campus Paderborn (SICP) beteiligt. Zudem ist die Heimann Sensor GmbH als Entwickler und Hersteller von Sensorik an dem Projekt konkret involviert. Die Organisationseinheiten lassen sich somit nicht eindeutig in Management, Fachabteilung, IT, Data-Science-Team und Externe einsortieren, vielmehr verschwimmen diese Grenzen im Rahmen dieses Forschungsprojekts.

*Welche Organisationseinheiten sind konkret verantwortlich?*

Die verschiedenen Arbeitsbereiche des Forschungsprojekts, die in Abbildung 4.1 dargestellt sind, werden von verschiedenen Projektpartnern bzw. einer Kombination der Projektpartner geleitet und verantwortet. Es ist dabei jeweils sichergestellt, dass der Partner mit der größten Expertise für



den jeweiligen Arbeitsbereich diesen auch inhaltlich leitet und verantwortet. Zudem ist sichergestellt, dass jeder Projektpartner mindestens einen Arbeitsbereich verantwortet.

*Wer hat das Projekt beauftragt?*

Das Projekt wurde vom Bundesministerium für Wirtschaft und Klimaschutz (BMWK, vorher Bundesministerium für Wirtschaft und Energie, BMWi) beauftragt.

*Welche Anspruchsgruppen dienen zusätzlich zu den Projektbeteiligten als Input-Geber zu fachlichen Aspekten?*

Im Rahmen des Projektes sind drei Unternehmen als zusätzliche Input-Geber für fachliche Aspekte zu nennen. Das Versorgungsunternehmen Städtische Werke Überlandwerke Coburg GmbH (SÜC) betreibt Mittel- und Niederspannungsnetze in Stadt und Landkreis Coburg, sodass es für die Analyse der Anforderungen, die Evaluierung und die prototypische Umsetzung des Projektes genutzt werden kann. Als zusätzlicher assoziierter Partner fungiert die Westfalen Weser Netz GmbH, welche regionale Mittelspannungsnetze im Raum Westfalen anbietet und die WestfalenWIND Gruppe, welche erneuerbare Energien (besonders Windparks) in Westfalen betreibt.

*Wer unterstützt / fördert das Projekt?*

Das Projekt wird gefördert vom BMWK entsprechend dem Förderaufruf „Betriebsmittel und Komponenten in Stromnetzen“ im Rahmen des 7. Energieforschungsprogramms der Bundesregierung.

*Gibt es mögliche Störer für das Projekt?*

Mögliche Störer sind im Projekt nicht bekannt.

*Welches sind die Aufgabenfelder eines externen Dienstleisters?*

Im Rahmen des Forschungsprojektes werden keine Aufgaben an externe Dienstleister vergeben.

## 4.5 Projektorganisation

*Welche Projektmanagementmethode ist vorgesehen?*

Durch das breit aufgestellte und räumlich verteilte Konsortium an Forschungs- und Industriepartnern ist es notwendig, dass das Projekt reibungsarm koordiniert und gesteuert wird. Dazu zählen, neben den aus dem DASC-PM bekannten Aspekten für Data-Science-Projekte, die Aufgaben der zeitlichen und inhaltlichen Koordination der Arbeitsbereiche, die Überwachung des Gesamtprojektplans, die Koordination des Berichtwesens, der Ausgleich der Interessen unterschiedlicher Partner, die Organisation und Durchführung der Projekttreffen und die Sicherstellung des Informationsaustauschs zwischen den Partnern und Arbeitsbereichen. Für die Durchführung der genannten Aufgaben kann keine alleinstehende Projektmanagementmethode ausgewählt werden, sondern es sind verschiedene Methoden zu kombinieren. Hinzu kommt, dass die einzelnen Arbeitsbereiche (siehe Abbildung 4.1) von den jeweiligen Projektpartnern, teilweise auch in Kombination, durchgeführt und verwaltet werden müssen, sodass hier wiederum eigene Methoden angewendet werden können.

Welche Rollen sind am Projekt beteiligt?

Die Projektpartner fungieren als jeweils eigenständige Rollen basierend auf ihren Arbeitsbereichen im Projekt, welche in Abbildung 4.1 dargestellt sind. Die Rolle des Data Scientist wird von Mitarbeitern von ABB und SICP eingenommen. Sie fungieren ebenfalls als Data Engineer in Kombination mit den Mitarbeitern von ABB, KIT und Heimann Sensor für das thermische Monitoring (AB 2) bzw. Schalterantriebsmonitoring (AB 3), die Bereitstellung der Infrarotsensorik (AB 5) und der experimentellen Datenerhebung (AB 6). ABB als Hersteller von Lasttrennschaltanlagen sowie die SÜC als Verteilnetzbetreiber gelten als Domänenexperten für die Besonderheiten der im FLEMING-Projekt thematisierten Anwendungsdomäne. Zudem sind das FIR und der SICP im Bereich der Anforderungsanalyse und Betreibererwartung (AB 1) sowie der Erarbeitung von Instandhaltungsstrategien und Betriebskonzepten (AB 8) als Domänenexperten zu sehen. Die Rolle des Projektmanagers für das Gesamtprojekt (AB 9) übernimmt ABB.

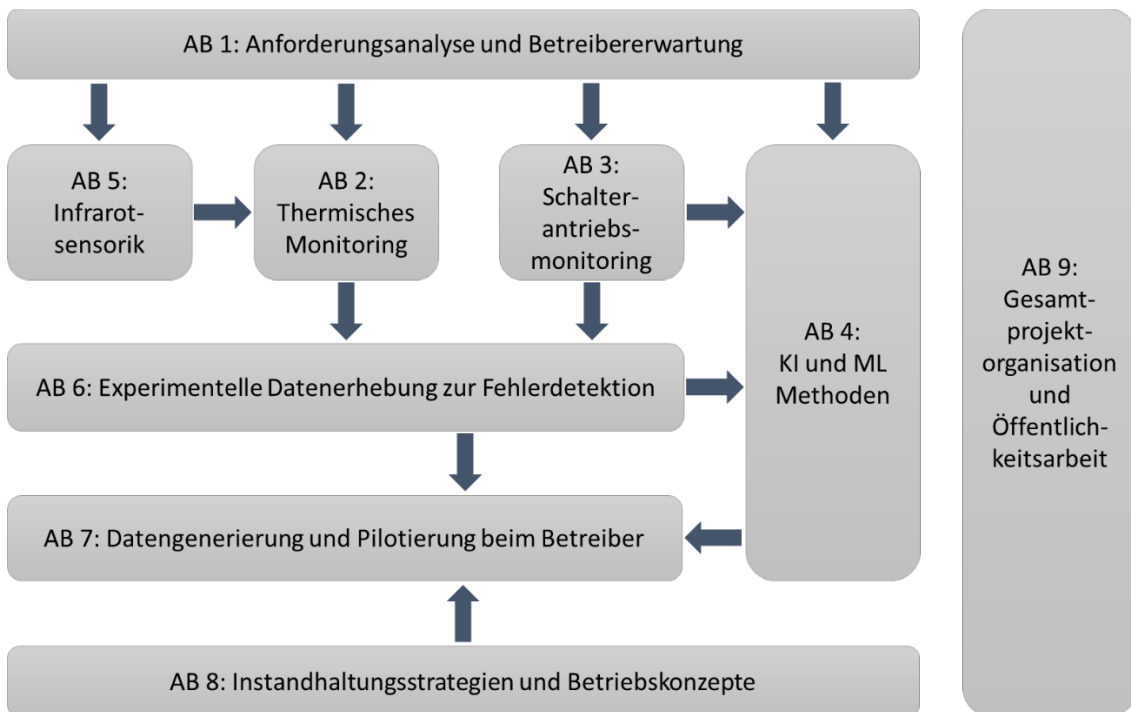


Abbildung 4.1: Übersicht der Projektorganisation des FLEMING-Projekts.

Wie sieht die Organisationsform des Projekts aus?

Die Organisation des FLEMING-Projekts ist schematisch in Abbildung 4.1 dargestellt. Die Abbildung zeigt die verschiedenen Arbeitsbereiche, die Verbindungen zwischen den Arbeitsbereichen und die Verantwortlichkeiten für die Arbeitsbereiche. Das bedeutet selbstverständlich nicht, dass bestimmte Arbeitsbereiche von einzelnen Partnern allein durchgeführt werden, sondern lediglich, dass die Verantwortung bei bestimmten Partnern liegt.

Welche zeitlichen Rahmenbedingungen existieren während der Projektdurchführung bis zur Vorlage des Ergebnisses?

Das Projekt wird über einen Zeitraum von drei Jahren durchgeführt. Dabei ist zu beachten, dass die jeweiligen Arbeitsbereiche (siehe Abbildung 4.1) mit ihren einzelnen Arbeitspaketen detailliertere

Zeitpläne aufweisen. Bspw. startet der AB 1 direkt zu Beginn, da die Ergebnisse für alle nachfolgenden Arbeitsbereiche starke Relevanz besitzen. So ist es vor allem wichtig zu wissen, welche Anforderungen an ein System für Predictive Maintenance von den Betreibern der Schaltanlagen zu erwarten sind. Die Instandhaltungsstrategien und Betreiberkonzepte (AB 8) sind dagegen erst zu erstellen, wenn die Anforderungen und Betreibererwartungen (AB 1) bekannt sind, da sie darauf aufbauen.

*Welche Kompetenzen haben die Projektmitglieder?*

Die Industrie- und Forschungspartner sind Experten in ihren jeweiligen Bereichen und bringen die notwendigen Kompetenzen mit, die zur Bearbeitung des Projektes notwendig sind. Dies ist bei den wissenschaftlichen Partnern aus den durchgeführten Forschungsprojekten und deren Publikationen ersichtlich, während die Industriepartner ihre Akzeptanz am Markt und ihr Wachstum überzeugen.

*Welche finanziellen Rahmenbedingungen gibt es?*

Das Projekt wird vom Bundesministerium für Wirtschaft und Klimaschutz (BMWK) gefördert. Dabei sind zu verwendende Gelder je Projektpartner vordefiniert und auf Personenmonate, IT-Infrastruktur, Reisemittel und weitere Posten aufgeteilt.

*Wieviel Vorlaufzeit existiert bis das Projekt beginnen muss?*

Das Projekt wurde im Juni 2019 vom BMWK bewilligt und die Bearbeitung hat im September 2019 gestartet. Die Vorlaufzeit von drei Monaten ist allerdings von geringer Relevanz, da in den meisten Arbeitsbereichen Aufwände für die Einarbeitung in die jeweiligen Themen eingeplant ist.

## 4.6 Vorerfahrungen

*Welche Lösungsansätze gibt es schon?*

Predictive Maintenance ist im Bereich der Energienetze keine neue Technologie, die im Rahmen dieses Projektes erstmals angewendet wird. In der Energiebranche wurde das mögliche Potenzial, welches die Predictive Maintenance als Instandhaltungsstrategie bieten kann, bereits erfolgreich eingesetzt. So existieren bereits Forschungsarbeiten zu dem Einsatz von Predictive Maintenance auf Windturbinen (Canizo et al. 2017) und in Kernkraftwerken (Hashemian 2011). Daher liegt in diesem Projekt der Fokus auch auf der Sichtung bestehender datengetriebener Methoden und der Auswahl von geeigneten Methoden für die Anwendungsdomäne von Lasttrennschaltanlagen in der Mittelspannung. Auch erste Lösungen für das Monitoring von Schalterantrieben und die Infrarotsensoren existieren bereits, allerdings sind auch diese nicht für den genannten Einsatzzweck optimiert.

*Welche Erfahrungen wurden durch vorherige ähnliche Projekte gesammelt?*

Die einzelnen Projektpartner haben in vorherigen Projekten, die teils öffentlich gefördert und teils intern durchgeführt wurden, bereits Erfahrungen zu ihren Fachbereichen erarbeitet. Im Rahmen des Konsortialverbunds ist bisher kein gemeinsames Projekt bearbeitet worden.

*Wo lagen bei vergangenen Projekten die Schwierigkeiten?*

Da es keine gemeinsamen Projekte gab, sind keine Schwierigkeiten bekannt, die das gesamte Projekt betreffen bzw. mehrere Partner übergreifen.

*Welche Organisationseinheiten haben Vorerfahrung mit Data Science?*

Das ABB Forschungszentrum in Ladenburg, besonders die in AB 4 involvierten Gruppen, hat bereits intensive Erfahrungen mit Data Science im Rahmen von internen und externen Projekten gesammelt. Der SICP kann besonders mit der Arbeitsgruppe Intelligente Systeme bisher deutliche und vielfältige Erfahrungen im Bereich Data Science in das Projekt einbringen. Das FIR ist Experte für datengetriebenen Geschäftsmodelle, sodass es bei den Data-Science-Aufgaben beratend zur Seite stehen kann, ebenso der Lehrstuhl für Wirtschaftsinformatik am SICP. Heimann Sensors hat durch Anwendungen, die auf der Sensorik basieren, bereits erste Erfahrungen in Data-Science-Ansätzen, bspw. durch Personendetektion und Hot-Spot-Detektion. Das KIT, welches am Projekt *FLEMING* beteiligt ist, kann im Vergleich zu den anderen Projektpartnern bisher wenig Vorerfahrung zu Data Science Aktivitäten vorweisen. Ebenso sieht es bei der SÜC als Betreiber von Verteilnetzen aus.

## 4.7 Fazit und Ausblick

Das Projekt *FLEMING* stellt ein großes, mehrere Industrieunternehmen und Forschungseinrichtungen umfassendes Data-Science-Projekt dar. An vielen Stellen sind die Grenzen zwischen den klassischen Rollen im DASC-PM nicht klar ersichtlich und sie überschneiden sich. Dennoch bietet das DASC-PM eine sinnvolle Vorgehensweise zur Durchführung und Strukturierung des gesamten Forschungsprojektes. Die im Verlauf dieser Fallstudie beantworteten Fragen zeigen die Herausforderungen, die im Rahmen des Projektauftrags identifiziert und in ein Projekt mit einem breit gefächerten Team eingegliedert wurden.

Wie bei anderen Data-Science-Projekten auch, kann es allerdings besondere Umstände geben, die die Ausführung des Projektes deutlich erschweren. Im Falle des *FLEMING*-Projektes ist dabei vor allem die COVID-19-Pandemie zu nennen, welche zu Beginn des Projekts nicht vorhersehbar war. Durch die Pandemie war die Beschaffung von Lasttrennschaltanlagen für die experimentelle Datengenerierung (AB 6) deutlich erschwert und führte dazu, dass die auf diese Daten angewiesenen Arbeitsbereiche ebenfalls nicht wie geplant Ergebnisse erzielen konnten. Da Data-Science-Projekte grundsätzlich ohne Daten nicht sinnvoll durchgeführt werden können (siehe auch den Schlüsselbereich *Daten*), ist hier besondere Umsicht geboten.

Die bisher im Projekt erzielten Ergebnisse reichen von der Analyse der Anforderungen der Betreiber (zur Heiden und Priefer 2021), die Weiterentwicklung von AutoML als eine Analysemethode zur Berechnung von Restlebenszeiten von Komponenten (Tornede et al. 2021), bis hin zu einer ganzheitlichen Gestaltung des Predictive-Maintenance-Systems (zur Heiden et al. 2022). Die Bearbeitungszeit ist zum Zeitpunkt des Verfassens dieser Fallstudie noch nicht ausgeschöpft, sodass die finalen Ergebnisse und eine Bewertung, ob das Projekt die zu Beginn aufgestellten Ziele erreicht hat, noch nicht abzuschätzen sind.

## 4.8 Förderhinweis

Dieser Beitrag beschreibt das Forschungsprojekt *FLEMING*, welches vom Bundesministerium für Wirtschaft und Klimaschutz (BMWK) unter dem Förderungskennzeichen 03E16012F gefördert wird. Wir danken dem BMWK sowie dem Projektträger Jülich (PtJ) für die Unterstützung..

## Literaturverzeichnis

ABB (2021): Luftisolierte Schaltanlagen. Online verfügbar unter <https://new.abb.com/medium-voltage/de/mittelspannungs-schaltanlagen/luftisolierte-mittelspannungsschaltanlagen>, zuletzt aktualisiert am 08.11.2021, zuletzt geprüft am 08.11.2021.

BMWi (2021): Das deutsche Stromnetz. Online verfügbar unter <https://www.bmwi.de/Redaktion/DE/Infografiken/Energie/abbildung-das-deutsche-stromnetz.html>, zuletzt geprüft am 26.08.2021.

Bundesnetzagentur (2019): Bericht - Zustand und Ausbau der Verteilernetze 2018. Bonn.

Canizo, Mikel; Onieva, Enrique; Conde, Angel; Charramendieta, Santiago & Trujillo, Salvador (2017): Real-time predictive maintenance for wind turbines using Big Data frameworks. In: 2017 IEEE international conference on prognostics and health management (icphm). IEEE, S. 70–77.

Hashemian, H. M. (2011): State-of-the-Art Predictive Maintenance Techniques. In: *IEEE Trans. Instrum. Meas.* 60 (1), S. 226–236. DOI: 10.1109/tim.2010.2047662.

Hoffmann, Martin W.; Wildermuth, Stephan; Gitzel, Ralf; Boyaci, Aydin; Gebhardt, Jörg & Kaul, Holger et al. (2020): Integration of novel sensors and machine learning for predictive maintenance in medium voltage switchgear to enable the energy and mobility revolutions. In: *Sensors* 20 (7), S. 2099.

Hofstätter, Frank; Weber, Thomas & Rabanus, Sebastian (2012): Intelligente Ortsnetzstationen als Alternative zum Netzausbau. In: *etz elektrotechnik & automation* (4).

Kamper, Andreas (2010): Dezentrales Lastmanagement zum Ausgleich kurzfristiger Abweichungen im Stromnetz: KIT Scientific Publishing.

Mobley, R. Keith (2002): An introduction to predictive maintenance: Elsevier.

Renn, Ortwin & Marshall, Jonathan Paul (2016): Coal, nuclear and renewable energy policies in Germany: From the 1950s to the “Energiewende”. In: *Energy Policy* 99, S. 224–232. DOI: 10.1016/j.enpol.2016.05.004.

Schiffer, Hans-Wilhelm (2019): Zielvorgaben und staatliche Strategien für eine nachhaltige Energieversorgung. In: *Wirtschaftsdienst* 99 (2), S. 141–147. DOI: 10.1007/s10273-019-2408-x.

Siemens, A. G. (2018): Planung der elektrischen Energieverteilung-Technische Grundlagen. Erlangen, Deutschland.

Tornede, Tanja; Tornede, Alexander; Wever, Marcel & Hüllermeier, Eyke (2021): Coevolution of remaining useful lifetime estimation pipelines for automated predictive maintenance. In: Francisco Chicano (Hg.): Proceedings of the Genetic and Evolutionary Computation Conference. GECCO '21: Genetic and Evolutionary Computation Conference. Lille France, 10 07 2021 14 07 2021. ACM Special Interest Group on Genetic and Evolutionary Computation. New York, NY, United States: Association for Computing Machinery (ACM Digital Library), S. 368–376.

zur Heiden, Philipp & Priefer, Jennifer (2021): Transitioning to Condition-Based Maintenance on the Distribution Grid: Deriving Design Principles from a Qualitative Study. In: Michael H. Breitner, Sebastian Lehnhoff, Astrid Nieße, Philipp Staudt, Christof Weinhardt und Oliver Werth (Hg.): Energy Informatics and Electro Mobility ICT. Community Workshop Proceedings, Pre-Conference 16. International Congress on Wirtschaftsinformatik Universität Duisburg-Essen: BIS-Verlag der Carl von Ossietzky Universität Oldenburg, S. 72–87. Online verfügbar unter <https://oops.uni-oldenburg.de/5084/>.

## Projekt FLEMING – Predictive Maintenance von zentralen Komponenten des Mittelspannungsnetzes

zur Heiden, Philipp; Priefer, Jennifer & Beverungen, Daniel (2022): Utilizing Geographic Information Systems for Condition-Based Maintenance on the Energy Distribution Grid. In: *Proceedings of the 55th Hawaii International Conference on System Sciences*.

# 5 Der Weg zum Projektauftrag

Florian Schwade, Heiko Rohde

## 5.1 Ausgangslage und Beschreibung des Fallstudienunternehmens

Die ACM Technology AG ist als Automobilzulieferer (A), Hersteller und Zulieferer von Consumer Electronics (C) und medizin-technischer Ausrüstung (M) tätig. Das Unternehmen zählt etwa 3.500 Mitarbeitende und produziert schwerpunktmäßig in den Kategorien *Luftfilter*, *Ventilatoren* und *Membranen*. Die Produkte werden bspw. benötigt in KFZ-Klimaanlagen (Automobilindustrie), Lüftern und Kühlern für Server, PCs und Laptops (Consumer Electronics) und medizinischen Beatmungsgeräten (Medizintechnik). Aufgrund des breiten Geschäftsfeldes ergeben sich für das Unternehmen Herausforderungen in der Zuweisung von Materialien und der Materialbedarfsplanung. Insbesondere aus dem Geschäftsbereich der Medizintechnik ergeben sich häufig Eilbestellungen, die bevorzugt bearbeitet werden müssen, sodass sich die Produktions- und Materialbedarfsplanung auch innerhalb eines Tages ändern kann. Eine stark schwankende Anzahl von Aufträgen aus der Automobilindustrie und deren häufige Stornierung erschweren die Planungen zusätzlich. Nicht zuletzt durch Lieferengpässe für wichtige Komponenten wie Chips und andere Rohstoffe, sieht die ACM Technology AG schon länger großen Bedarf darin, die Materialbedarfsplanung und die Allokation von (knappen) Materialien zu verbessern, um allen Geschäftsbereichen gerecht zu werden. Ende 2020 startete das Unternehmen ohne Vorerfahrungen oder Expertise Versuche Optimierungen mit Hilfe von Verfahren des Operations Research und Data-Science-Methoden hervorzurufen. Dabei wurden vor allem die Komplexität und das Management dieses Data Science Projektes unterschätzt, sodass gute Ansätze nicht erfolgreich umgesetzt werden konnten. Herausforderungen ergaben sich dabei hinsichtlich mangelnder Expertise des Projektteams, fehlendem Rückhalt innerhalb der Organisation und bei technologischen Aspekten. Dieses als fehlgeschlagen eingestufte Projekt wurde aufgearbeitet. Infolgedessen wurde ein neues und bewusster ausgewähltes Projektteam zusammengestellt.

Das Projekt soll nun strukturiert anhand des DASC-PM umgesetzt werden. Die beteiligten Personen haben sich darauf geeinigt, in einem ersten Schritt die Fragestellungen und die damit verbundenen Ziele sauber in Form eines Projektauftrages zu formulieren. Die Ausarbeitung des Projektauftrages ist Gegenstand dieser Fallstudie und orientiert sich an der Phase Projektauftrag des DASC-PM. Dies kann somit als Auslöser für dieses Data Science Projekt gesehen werden.

Um erfolgreicher in die Definition des Projektes zu starten, wurden als erste Maßnahme im Zuge der Zusammenstellung des Projektteams Kompetenzprofile (anhand der Dimensionen *Mathematik/Statistik*, *Informationstechnik*, *Anwendungsbereich*, *Kommunikation*, *Strategie* und *Management*) ermittelt, wodurch sichergestellt werden konnte, dass die notwendigen Expertisen und Rollen (v.a. Domänenexperte, Data Scientist, Data Engineer und Projektmanager) ausgewogen im Projektteam repräsentiert waren. In Konsequenz waren nicht nur Data Scientists am Projekt beteiligt, sondern auch Experten bzw. Product Owner der wichtigsten Produkte, die über Expertise und Domänenwissen der Geschäftsbereiche, Zielmärkte und Produktion verfügten. Dadurch, dass der CIO und der

COO direkte Projektsponsoren waren, wurde der Rückhalt im Management sichergestellt. Insgesamt bestand das Projekt aus neun Teilnehmern:

- Projektmanager
- SCRUM-Master
- Product Owner Atemschutzgeräte (Domänenexperte)
- Produktionsverantwortlicher 3D-Fertigung (Domänenexperte)
- Head of Sales (Domänenexperte)
- Head of Procurement (Domänenexperte)
- Data Scientist
- Head of IT (technischer Support)
- Compliance Officer (Compliance Support)

## 5.2 Entwicklung der Use Cases

Nach der Zusammenstellung des Projektteams wurden die Mitglieder zu einer gemeinsamen Fokusgruppe eingeladen. Die Fokusgruppe wurde von einer Mitarbeiterin geleitet, die zwar mit den Details des Projektes vertraut ist, aber nicht beteiligt war. Da diese Mitarbeiterin promoviert ist und Fokusgruppeninterviews bereits als Forschungsmethode durchgeführt hat, verfügte sie über entsprechende Expertise in deren Gestaltung und Durchführung. Die Entscheidung eine Fokusgruppe als Methode zur Detaillierung des Projektes zu wählen, wurde gefällt, da Fokusgruppen auf Diversität in Ideen und Meinungen abzielen und unterschiedliche Wissensstände oder Expertisen gleichermaßen berücksichtigen. Zudem waren die Teilnehmer für eine Fokusgruppe aufgrund der Diversität ihrer Rollen ideal geeignet und elf Teilnehmer gelten als ideale Größe einer Fokusgruppe.

Das primäre Ziel der Fokusgruppe war es, Use Cases detaillierter herauszuarbeiten und die eigentlichen Problemstellungen und Herausforderungen zu identifizieren, sodass eine Annäherung an die Definition des Projektes stattfinden konnte. Hierbei sollten, da es sich um mehrere Teilprojekte handelt, die Merkmale der „Auslöser“ weiter detailliert und hinterfragt werden. Insbesondere standen *fachlicher Zweck, Ziele, Anwendungsrahmen* sowie die *Komplexität* im Vordergrund. Die Fragestellung nach den benötigten Daten wurde ebenfalls adressiert. Diese Fragestellung wird jedoch nicht im Rahmen dieser Case Study behandelt.

Nachfolgend werden die für die Phase *Projektauftrag* wichtigsten Ergebnisse der Fokusgruppe zusammengefasst. Durch die kreativen Elemente in der Fokusgruppe konnte eine Vielzahl möglicher Use Cases identifiziert werden. Im Rahmen dieser Fallstudie werden die drei Aspekte der Materialbedarfsplanung, Rüstzeitoptimierung und Logistik fokussiert. Die weiteren möglichen Use Cases werden außen vorgelassen.

Eine der größten Herausforderungen für die ACM Technology AG ist die Materialbedarfsplanung. Diese ergibt sich, wie in der Ausgangslage beschrieben, aus sich wandelnden Anforderungen der Auftraggeber und der derzeit größtenteils knappen Ressourcen, wie bspw. Microchips. Ressourcenknappheit führt oftmals dazu, dass nicht alle Aufträge sofort vollständig bedient werden können.



Folglich muss häufig eine Triage durchgeführt werden. Das Unternehmen muss also entscheiden, welche Aufträge priorisiert werden müssen und welche Aufträge zu welchem Volumen bedient werden können. Die Herausforderung wird durch die fortwährende COVID-19-Pandemie, aufgrund derer die kurzfristigen Eilbestellungen aus dem Bereich Medizintechnik, bspw. für Beatmungsgeräte, stark gestiegen sind, noch weiter verschärft. Durch solche Eilbestellungen muss kurzfristig, manchmal innerhalb weniger Stunden die Produktionsplanung für den Tag nahezu in Echtzeit angepasst werden. Dies wiederum hat als Konsequenz, dass die Produktion umpriorisiert werden muss, was in der Planung der Auslastung nicht vorgesehen war. Ein weiterer Bestandteil des Projektes ist damit die Planung, Berechnung, Vorhersage und Simulation von Kapazitäten (Mensch, Maschine, Rohstoffe) bei sich ändernden Parametern.

Als eines der zentralen und neuen Probleme der ACM Technology AG wurde die Rüstzeitoptimierung identifiziert. Das Unternehmen hat vor einigen Monaten weite Teile der Produktionslinien auf 3D-Druckfertigung umgestellt und noch wenige Erfahrungen mit der Optimierung der Produktionsreihenfolge gesammelt. Hier sind insbesondere der Wechsel der benötigten Filamente und teilweise notwendige Stillstandzeiten aufgrund von Reinigung herausfordernd. Durch Data Science erhofft man sich Verbesserungen in der Produktion vor allem hinsichtlich der Produktionsreihenfolge und Rüstzeitoptimierung, um Maschinen möglichst effizient nutzen und Stillstandzeiten minimieren zu können.

Schlussendlich wurde als weiteres mögliches Ziel die Optimierung der Logistik thematisiert. In diesem Bereich erhofften sich die Mitglieder der Fokusgruppe Verbesserungen bei der Logistik, insbesondere Kommissionierung von Waren. Der Verpackungsverbrauch steht hierbei eher im Hintergrund. Vielmehr ging es um die Entscheidung, wann Lieferungen verschickt werden, insbesondere bei Eilbestellungen. Zentrale Fragestellungen sind hier: Sollen einzelne Produkte aus Eilbestellungen sofort einzeln nach Fertigstellung oder konsolidiert als Teil- oder Gesamtlieferung versendet werden? Da die zuvor genannten Aspekte größere Herausforderungen sind, wurde beschlossen, die Logistikoptimierung in ein Folgeprojekt zu verlagern.

Nach einer Zusammenfassung der Ergebnisse der Fokusgruppe wurde vom Management gefordert, dass ein kleiner Teil des Projektes fokussiert und als Proof of Concept umgesetzt werden soll, da großes Potenzial für einen immensen Wettbewerbsvorteil durch Verbesserungen gesehen wurde. Durch den Proof of Concept sollte schnell und mit wenigen Ressourcen ein erster Prototyp entwickelt werden, der dann sukzessive zu einem unternehmensweit sichtbaren Leuchtturmprojekt entwickelt wurde. Nach einer abschließenden Diskussion über die möglichen Risiken wurde festgehalten einen sauberen Prototypen und keine "Quick & Dirty" Lösung zu entwickeln.

Im Zuge der Nachbereitung der Fokusgruppe wurde die Moderatorin ebenfalls in das Projektteam aufgenommen, um die Perspektive aus der Wissenschaft stärker im Team zu repräsentieren.

Aus diesem Grund sollte nun zunächst für die Problemstellungen erarbeitet werden, ob sie vorwiegend eine strategische Perspektive oder eine operative Perspektive haben, ob sie gegebenenfalls sogar nur rein datengetrieben sind und wie stark sie auf das Marktumfeld wirken bzw. eine Wirkung erzielen. Mit der Klassifizierung sollte die Frage geklärt werden, ob es sich grundsätzlich um Data-Science-Projekt handelt. Zwei Aspekte wurden hier besonders betrachtet und mit externen Data Scientists verifiziert – erstens die Datenverfügbarkeit von geeigneten Datensätzen im Unternehmen, bzw. aus angeschlossenen Datenquellen, sowie zweitens ob überhaupt der Einsatz von Data Science gegenüber klassischen Analytics Methoden überlegen erscheint. Zusätzlich sollte diesmal ein klares Ziel definiert werden, so dass auch für den Fall, dass eine Problemstellung nicht zu einem gewünschten Ergebnis führt, dennoch ein Ergebnis erzielt wird.

Der erste Use Case der Materialplanung wurde zunächst operativ gesehen, da durch kurzfristige Produktionsanpassungen ein starker operativer Einfluss erkennbar war. Allerdings haben die Domänenexperten in der Fokusgruppe schnell darlegen können, dass wenige Rohmaterialien für alle Produktionsprozesse die Grundlage bilden, so dass das Endmaterial hauptsächlich vom Herstellungsprozess mit seinen Rezepturen abhängig ist und weniger von speziellen Materialien. Es gibt hier zwar Additive, die aber in der Masse kaum eine Rolle spielen. Dadurch würde das Thema stärker in den Bereich „Marktumfeld und Strategie“ verschoben, denn die Fragestellungen beziehen sich hier auf die Lieferkette, die strategische Auswahl von Zulieferern und die langfristige Sicherstellung der Verfügbarkeit von Produktionsrohstoffen.

Der zweite Use Case der Produktionsoptimierung im 3D Druck geht auch klar in die strategische Richtung und hat einen sehr innovativen Charakter. Durch die Prozessbetrachtung sollen neben der Optimierung der Verfahren auch neue Geschäftsfelder erschlossen werden, wie der Fertigung beim Kunden.

Das dritte Thema, welches aber in ein eigenes Projekt ausgelagert werden soll, wird als operatives Thema gesehen. Die Optimierung in der Logistik betrifft zwar den Kunden und hat eine gewisse Marktrelevanz, allerdings setzen immer mehr Unternehmen auf Nachhaltigkeit und erwarten dies auch von ihren Zulieferern. Es wird kein großes Risiko gesehen, wenn Lieferungen zusammengefasst werden, solange noch die Lieferziele erreicht werden. Auch hier sehen die Domänenexperten weiteres Potential, was gegebenenfalls bis in die Produktionsplanung greifen könnte.

### 5.3 Eignungsprüfung und Sicherstellung der Umsetzbarkeit

Die begleitende *Eignungsprüfung* war für dieses Vorhaben sehr spannend. Gerade die Eignung der Problemstellungen wurde vom Management kritisch betrachtet, da das erste Projektvorhaben nicht erfolgreich gewesen ist und sich die Problemstellungen nur unwesentlich von denen des ersten Versuchs unterscheiden.

In der Vorstudie zum Projekt musste die Sicherstellung der Umsetzbarkeit der Use Cases durchgeführt werden. Um hier ein besseres Bild zu erhalten, sollten vergleichbare Herausforderungen, beschrieben in Publikationen und Online-Quellen, mit den eigenen Problemstellungen verglichen werden, um herauszufinden, ob die ACM Restriktionen unterliegt, die bei der Projektdurchführung eine Rolle spielen.

Die wohl wichtigsten domänenspezifischen Anforderungen korrelieren auch mit den Problemstellungen. Für die Problemstellungen aus dem Bereich Produktion für medizinische Produkte ist in hohem Maße fachliches Prozesswissen notwendig, da hier Chargenverfolgung, Materialanforderungen durch Regulierung und spezielle Reinheitsverfahren zur Anwendung kommen. So können Luftfilter für die Automobil-Sparte nicht einfach in Medizinprodukte umgewandelt werden, ohne auf die spezifischen Herstellungsparameter zu achten, auch, wenn die Produkte sehr ähnlich bis baugleich sind. Dieses Fachwissen ist allerdings durch jahrelange Erfahrung im Unternehmen vorhanden, so dass dieses in das Projekt direkt eingebracht werden kann. Das Risiko zu wenig Expertise für ein Projekt zu haben, kann hier ausgeschlossen werden.

Ein weiteres Domänenspezifikum wurde durch die COVID-19-Pandemie hervorgerufen: der politische Druck auf Unternehmen aus der Medizinbranche. Regelmäßig ist das Management von gesundheitspolitischer Seite eingeladen und um Hilfe ersucht worden, um gemeinsam mit anderen Herstellern schnell die notwendigen Gerätschaften, wie Luftfilter, zu liefern. Auch hier stand das Unternehmen ACM zwischen unterschiedlichen Interessengruppen, die es zu bedienen galt. Die Konsequenzen einer erneuten Verschärfung der Situation konnte nicht ausgeschlossen werden, allerdings kann das direkte Projektrisiko gering bewertet werden, beziehungsweise sieht das Management hier Data-Science als Chance.

Viel spannender als die spezielle Fertigung in der Medizintechnik ist allerdings generell das Herstellungsverfahren in der 3D-Drucktechnik. Dieses wurde als strategisch eingestuft mit der Hoffnung, ist die Produktion nachhaltig in allen Sparten zu optimieren. Aus dem Fachbereich gibt es sogar Überlegungen, dass die Fertigung an einen Kundenstandort ausgelagert werden könnte, so dass das Produkt vor Ort gedruckt wird.

### 5.4 Vorgehen zur Auswahl des Use Cases

Im nächsten Schritt soll die Bewertung der Use Cases vorgenommen werden, um die projektrelevanten Use Cases zu identifizieren und allgemein eine Eignung festzustellen. Das DASC-PM unterscheidet die beiden begleitenden Aufgaben *Eignungsprüfung* und *Sicherstellung der Umsetzbarkeit*, allerdings hat man sich hier darauf geeinigt für dieses Projekt eine Neun-Felder-Matrix zu nutzen, um alle Ergebnisse der Teilaufgaben in eine Bewertung einfließen zu lassen. Auf der x-Achse ist dabei die Einschätzung der Durchführbarkeit und auf der y-Achse die Relevanz für das Unternehmen abgebildet, jeweils in den Ausprägungen gering, mittelmäßig und hoch (siehe Abbildung 5.1).

Obwohl dieses Vorgehen keinen streng wissenschaftlichen Maßstäben folgt, so konnte durch die Beteiligten der Fokusgruppe aufgezeigt werden, welcher Use Case gegenüber einem anderen zu priorisieren war. Zudem bietet diese Matrix die Möglichkeit Aspekte wie IT-Infrastruktur, Kosten-Nutzen, die Einschätzung von Risiken oder ggf. die spätere Integration in Anwendungen im aktuellen Unternehmenskontext zu bewerten, was der begleitenden Aufgabe *Sicherstellung der Umsetzbarkeit* folgt. Auch die Akzeptanz im Unternehmen aus organisatorischer Sicht kann auf die Durchführbarkeit einzahlen. In diesem Kontext mussten durch die Fokusgruppe mit Experten zunächst Widerstände abgebaut werden. Mit Rückendeckung des Managements wurde der Fokus auf die technische Infrastruktur und die Datenverfügbarkeit gelegt.

Die Relevanz für das Unternehmen wurde über den Innovationsgrad einer Lösung, sowie der Chance für Umsatzsteigerung oder Kosteneinsparung definiert, da sich hier der stärkere Managementeinfluss bemerkbar machte, wodurch eine interne Priorisierung der Use-Cases geschaffen wurde.

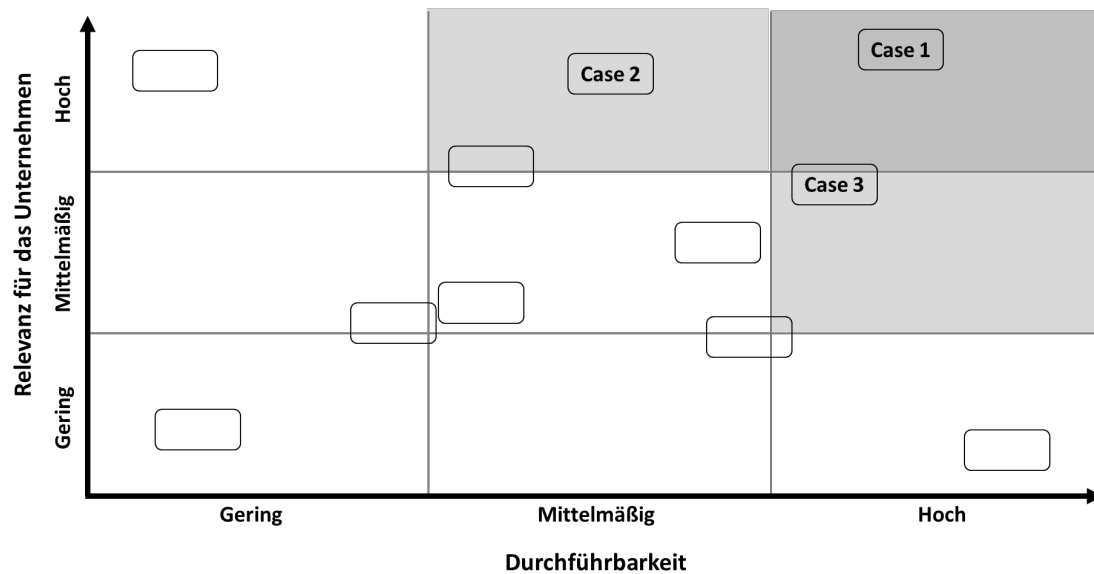


Abbildung 5.1: Einordnung der Use Cases nach Durchführbarkeit und Relevanz

Für dieses Projekt konnten so die Use Cases definiert werden und mögliche weitere für zukünftige Projekte. Eine einfache Matrix-Darstellung hätte durch weitere wissenschaftliche Methoden ergänzt werden können, falls eine triviale Auswahl nicht möglich gewesen wäre.

## 5.5 Projektausgestaltung

Nachdem die Use-Cases ausgewählt wurden und die Prüfungen stattgefunden haben, wurde zur internen Präsentation vor den beiden Projektspensoren (CIO & COO) folgender Projektsteckbrief zur Vorbereitung eingereicht, worin die Projektausgestaltung weiter beschrieben war

### Projektsteckbrief

#### Projekttitel:

Materialbedarfsplanung & Rüstzeitoptimierung durch Data Science

#### Motivation:

Das Unternehmen steht häufig vor der Herausforderung, Bestellungen priorisieren zu müssen. Dies gilt insbesondere für Eilbestellungen aus der Medizintechnik. In diesem Kontext sind auch verknappte Ressourcen, bspw. aufgrund des andauernden Chipmangels eine Herausforderung. Daher möchte das Unternehmen auch zukünftig für etwaige Ressourcenknappheit besser gewappnet sein. Eng verbunden mit diesen Herausforderungen ist die Rüstzeitoptimierung. Das Unternehmen ist neu in der 3D-Druckfertigung tätig und möchte die Rüstzeiten für die 3D-Drucker bedingt durch Filamentwechsel verbessern.

#### Ziele:

Die Ziele des Projektes sind somit zum einen die Verbesserung der Materialbedarfsplanung und die Verbesserung der Rüstzeiten der 3D-Drucker. Die Ziele können anhand der folgenden Kennzahlen gemessen werden:

- Menge der verwendeten Materialien im Verhältnis zum Produktionsoutput
- Menge des Ausschusses im Verhältnis zu den eingesetzten Ressourcen
- Auslastung der Produktionskapazitäten (Menschen und Maschine)
- Stillstandszeiten
- Veränderungen in der Produktion nach Eingang von Eilbestellungen

**Innovation:**

Die Innovation des Projektes besteht für die ACM Technology AG darin, durch die Anwendung von Data Science die Materialbedarfsplanung und die Rüstzeiten von Druckern und Maschinen in der Produktion zu optimieren. Insbesondere hinsichtlich der Rüstzeiten und den optimierten Wechsel der Filamente wird ein Wettbewerbsvorteil gegenüber der Konkurrenz erwartet. Ein weiterer Aspekt der Innovation ist die Erweiterung der Data Science Kompetenzen im Unternehmen, was für künftige Data Science Projekte zuträglich sein wird.

**Nutzungsmöglichkeiten:**

Die primären Nutzungsmöglichkeiten der Projektergebnisse liegen in den skizzierten Bereichen der Materialbedarfsplanung und Rüstzeitoptimierung. Die gewonnenen inhaltlichen, technischen, methodischen und Data-Science-Kompetenzen können nach Projektabschluss auf weitere Bereiche ausgeweitet werden. Es wird erwartet, die im Projekt entwickelten Analyseverfahren auch zur Verbesserung der Absatz- und Betriebsplanung (S&OP) anwenden zu können. Da insbesondere der Aspekt der Rüstzeitoptimierung für das Pilotprojekt zunächst speziell auf den 3D-Druck ausgerichtet ist, ist zu erwarten, dass insbesondere die Analysen und Erkenntnisse hierzu auf sämtliche Produktionsschritte und Standort der gesamten Fertigung ausgeweitet werden können.

**Meilensteine und Zeitplanung:**

Es ist vorgesehen, dass das Projekt bereits nach einem Monat einen "Quick Win" in Form eines Prototypen vorweisen kann. Mittels des Prototypen soll zum einen die technische Machbarkeit demonstriert werden und zum anderen erste rudimentäre Erkenntnisse gewonnen werden können. Die Gesamtdauer des Projektes ist für sechs Monate angesetzt. Um diese Ziele zu erreichen orientiert sich das Projekt am DASC-PM und bezieht weiter agile Methoden für die Durchführung und das Management des Projektes ein. Die Meilensteine ergeben sich entsprechend aus den Schlüsselbereichen des DASC-PM. Das Projekt soll zum 03.01.2022 starten. Die nachfolgend skizzierten Meilensteine konzentrieren sich auf die Erlangung des Proof of Concept nach einem Monat. Da im Projekt agil gearbeitet wird, kann auch nach Entwicklung des Prototypen iterativ an bereits abgeschlossenen Meilensteinen gearbeitet werden, bspw. falls die Erkenntnisse aus der Anwendung des Prototyps zeigen, dass neue Datenquellen einbezogen oder Daten weiter aufbereitet werden müssen.

Tabelle 5-1: Meilensteine im Projekt

Meilensteine	Fälligkeit
Definition des Projektes	31.12.2021
Identifikation und Selektion der Ursprungsdatenquellen	07.01.2022
Datenaufbereitung	24.01.2022
Identifikation und Evaluation geeigneter Analyseverfahren	07.01.2022
Entwicklung von Analyseverfahren	31.01.2022
Sicherstellung technischer Umsetzbarkeit	31.01.2022
Technisch-methodische Bereitstellung	31.05.2022
Fachliche Bereitstellung	30.06.2022

### Technische Voraussetzungen:

Die technischen Voraussetzungen für das Projekt können vorab nicht vollständig abgeschätzt werden, da diese teilweise von den ausgewählten Analyseverfahren und Datenmengen abhängen. Es wurde beschlossen, dass der Prototyp in Python implementiert wird, sodass eine Entwicklungsumgebung in Python vorhanden sein muss. Zudem benötigen die Mitglieder des Projektteams entsprechend leistungsfähige Laptops. Später könnten Cloud-Umgebungen eingesetzt werden, um eine bessere Verfügbarkeit und Einbindung in die IT-Landschaft zu gewährleisten. Hiervon wird aber zunächst abgesehen, da kein Cloud-Architekt zur Verfügung steht.

### Projektteam:

Das Projektteam besteht insgesamt aus neun Personen. Das Projekt setzt sich aus folgenden Teilnehmern und Rollen zusammen:

- Projektmanager
- SCRUM-Master
- Product Owner Atemschutzgeräte (Domänenexperte)
- Produktionsverantwortlicher 3D-Fertigung (Domänenexperte)
- Head of Sales (Domänenexperte)
- Head of Procurement (Domänenexperte)
- Data Scientist
- Head of IT (technischer Support)
- Compliance Officer (Compliance Support)

Die Auswahl der Teilnehmer und Zuordnung zu Rollen orientiert sich an den Empfehlungen des DASC-PM. Die Zusammenstellung des Projektteams zeigt, dass das Projekt nicht als IT-Projekt angelegt wurde. Es wurde insbesondere Wert daraufgelegt, dass die Domänenexpertise der betroffenen Fachbereiche genutzt werden kann, um so viel Wert wie möglich aus dem Projekt ziehen zu können. Trotz der anteilmäßig hohen Gewichtung der Rollen aus dem "Business" wird dem Data Scientist die Kernrolle im Projekt zugeschrieben.

### Kostenabschätzung:

Die Projektkosten belaufen lassen sich in Personal und Sachkosten gliedern. Für die Sachkosten ist die Anschaffung der Entwicklungsumgebung bzw. der notwendigen Hardware notwendig. Hier ist nicht von einem größeren Investitionsvolumen auszugehen. Außerdem sind externe Beratungskosten im Umfang von ca. 80 Projekttagen vorgesehen. Eine erste Kostenabschätzung beläuft sich hier auf 100.000 – 120.000€. Vor allem externes Data Science Know-how wird für das Projekt benötigt und muss extern eingekauft werden, bis die notwendigen internen Ressourcen trainiert sind.

Aus internen Ressourcen sind 150 Projekttag geplant. Diese teilen sich wie folgt auf:

Tabelle 5-2: Zusammensetzung des Projektteams inkl. Personentage

Projektteam	Tage
Projektleitung / Product Owner	30 Tage
Scrum Master	15 Tage
Domänenexperten (3 Personen, zusätzlich zum Product Owner)	40 Tage
Data Scientist (1 Person, ohne extern)	50 Tage
Supportrollen (IT, Compliance)	15 Tage

Als grobe Schätzung der internen Kosten können für die Projektzeit von 150 Tagen interne Kosten von 50.000€ angesetzt werden.

## Projektskizze

Für die Interne Projektsteuerung nach Freigabe des Projektantrags wurde Scrum als Methode gewählt, um in dem Rahmen des DASC-PM das Projektmanagement zu organisieren.

Hierzu wurde ein Sprint-Zyklus von anfänglich zwei Wochen in der POC Phase gewählt und dieser danach auf drei Wochen erweitert, um jeweils einmal das DASC-PM mit allen Phasen zu durchlaufen, um ein Increment, also einen fertigen, nutzbaren Case, zu implementieren. Die Scrum-Ereignisse (Events) wurden mit dem Ablauf des DASC-PM Verbunden. Im Rahmen einer Review sollte immer wieder erneut die Definition des Projekts überprüft und gegebenenfalls angepasst werden, so dass mit dem Sprint-Planning ein neuer Zyklus durch das DASC-PM starten kann.

Die agile Arbeitsweise soll ermöglichen, das größte Problem eines Data Science Projektes früh zu erkennen, nämlich die Möglichkeit, dass kein nutzbares Ergebnis nach der fachlichen Zieldefinition möglich ist. Die Gründe können hierfür vielfältig sein, aber durch kurze Zyklen können diese schneller erkannt werden.

Eine kleine Abweichung zum klassischen Scrum wurde jedoch implementiert. Die Definition of Done sollte zwar sagen, dass das Ergebnis, also das Increment, eines Sprints kompilierbar und lauffähig sein sollte, allerdings wurden für die Stories, bzw. Aufgaben, jeweils weitere „Definitions of Done“ angelegt, die sich auf die einzelnen Phasen des DASC-PM beziehen sollten.

So soll beispielweise für die Phase *Datenbereitstellung* gelten, dass die Daten aus der Datenquelle extrahiert, transformiert und qualitätsgesichert sind, in dem sie eine fachliche Validierung durchlaufen haben. Für die Phase *Analyse* sind zwei Definitionen erstellt worden, die in Abhängigkeit angewendet werden können. Wenn Testdaten zur Verfügung stehen, dann ist ein Test des Modells gegen ein Validierungsdatenset möglich. Alternativ wird das Analyseartefakt durch die Fachbereiche bzw. Stakeholder validiert. Für die Nutzbarmachung und die Nutzung wurden zunächst keine weiteren Definitionen entwickelt. Mit den Merkmalen der Phasen des DASC-PM könnten hierzu noch weitere entwickelt werden.

Für den Projektstart sind schon einige Backlog-Items generiert worden, die von der fachlichen Prozessdefinition, über die Datenextraktion und -transformation bis zu möglichen Analyseverfahren reicht. Da ein vollständiges Backlog hier zu lang wäre, anbei ein paar Beispiele für die Materialbedarfsplanung. Hier sollen die aktuellen Prozessaufträge aus dem ERP-System extrahiert werden, um diese mit den historischen Absätzen zu vergleichen. Außerdem soll geschaut werden, ob es Ereignisse in der Vergangenheit gegeben hat, die zu Umplanungen geführt und die unter Umständen eine Regelmäßigkeit haben, bzw. anhand von Merkmalen eine frühere Erkennung möglich machen. Die Domänenexperten sollen grundsätzlich die Absatzströme besser prozessual beschreiben und

Datenmerkmal benennen. Aus den Diskussionen im vorherigen Projekt konnte zum Beispiel festgestellt werden, dass sich Materialien je nach Auftragsart unterscheiden können. Ein langfristiger Kontrakt ist nicht so anfällig für Änderungen, wie eine ausgelöste Bestellung, wobei es auch hier materialspezifische Abweichungen geben kann, die in eine Optimierung einfließen könnten.

Zusammengefasst lässt sich das DASC-PM gut mit agilen Methoden kombinieren, wobei die Methoden aufeinander abgestimmt werden müssen. Das wichtigste für ein Data Science Projekt aber ist ein motiviertes Team, der Rückhalt in der Organisation, ein abgestimmtes und methodisches Vorgehen (z.B. DASC-PM) sowie ein realisierbarer Use Case, der im ersten Schritt nicht „nach den Sternen greift“.



# 6 Identifikation von Gesichtsmasken

René Theuerkauf, Tony Franke

## 6.1 Ausgangslage

Das Unternehmen BIG KAUF AG möchte zur effizienteren Gestaltung der Zutrittskontrolle für ihre Ladengeschäfte eine Software zur Identifikation des Tragens einer Mund-Nasen-Abdeckung in Auftrag geben. Die bisherige Praxis sieht vor, dass die Zutrittspunkte eines Gebäudes durch Mitarbeiter vor Ort während der Öffnungszeiten vollumfänglich überwacht werden. Durch die vorliegende Personalausstattung muss die Anzahl der Zutrittspunkte stark reduziert werden. Die softwareseitige Unterstützung soll es ermöglichen, die Anzahl der Zugänge bei Bedarf und bei gleichbleibendem Personaleinsatz, wieder erhöhen zu können. Bei der Umsetzung muss die Entscheidung aufgrund von Auflagen des zuständigen Gesundheitsamtes mit einer hinreichenden Verlässlichkeit erfolgen. Dem Unternehmen soll eine Applikation zur Verfügung gestellt werden, welche bei Verstößen automatisch eine Aktion ausführt/auslöst und ein Monitoring aus der Ferne ermöglicht. Unter Anwendung des Vorgehensmodells DASC-PM soll die Entwicklung der Software durchgeführt werden. Es wird ein Projekthorizont von drei Monaten vereinbart. Das Budget umfasst finanzielle Ressourcen für 0,25 VE Projektmanagern, 2 VE Data Scientist und 0,5 VE Data Engineer. Die Abnahme des Projektes erfolgt durch den Kunden.

## 6.2 Einleitung

Aufgrund der seit 2019 anhaltenden Corona-Pandemie und den damit verbundenen Hygienevorschriften gilt für Filialen mit Publikumsverkehr unter anderem die Pflicht der Kontrolle des Tragens eines Mund-Nasen-Schutzes bei ihren Kunden, bspw. in Form einer medizinischen Maske. Die entsprechende Überprüfung kann beim Zutritt in das Gebäude erfolgen. Dies stellt für Unternehmen mit größeren Filialen, welche mehrere Zugänge besitzen, ein aufwändiges Unterfangen dar. Hiermit kann ein deutlich erhöhter Personaleinsatz zur Umsetzung dieser Kontrollen einhergehen. Um dem entgegenzuwirken, kann mit Software-Lösungen aus dem Bereich der Data Science gearbeitet werden. In dieser Case Study wird die Erstellung eines Software-Prototyps anhand des Vorgehensmodells DASC-PM beschrieben, das als Forschungsprojekt von Studierenden im Rahmen eines Projektseminars der Martin-Luther-Universität Halle-Wittenberg umgesetzt wurde. Es wurden die fünf Phasen „Projektauftrag“, „Datenbereitstellung“, „Analyse“, „Nutzbarmachung“ und „Nutzung“ des Vorgehensmodells sukzessive geplant und bis einschließlich der Nutzbarmachung umgesetzt. Im Folgenden werden diese Phasen ausführlich beschrieben und die Anwendbarkeit des DASC-PM auf realwirtschaftliche Projekte demonstriert.

## 6.3 Projektauftrag

Das Unternehmen BIG KAUF AG möchte eine Softwarelösung für die automatisierte Maskenerkennung in Live-Video-Feeds in Auftrag geben. Diese soll im Eingangsbereich der Filialen des Unternehmens zum Einsatz kommen und Kunden bei unterlassener Nutzung einer medizinischen Maske automatisiert auf die Einhaltung der aktuellen Hygienevorschriften hinweisen. Das Budget umfasst finanzielle Ressourcen für 0,25 VE Projektmanagern, 2 VE Data Scientist und 0,5 VE Data Engineer für einen Dreimonatshorizont. Das Projektziel wurde von der BIG KAUF AG wie folgt definiert. Die erstellte Software soll mit einer Genauigkeit von 80% zwischen Kunden mit Maske und Kunden ohne Maske, resp. einer nicht korrekten Nutzung dieser, unterscheiden können. Ein visuelles und akustisches Signal soll bei Identifikation der falschen Nutzung den Kunden auf diesen Sachverhalt hinweisen.

## 6.4 Datenbereitstellung

Die Datenbereitstellung wurde im Rahmen des Projekts in zwei Schritten realisiert. Im ersten Schritt wurde nach einer kostenfreien Datenquelle bzw. Datenbank gesucht, welche Bilder beinhaltet, die bereits dahingehend gekennzeichnet sind, ob die Person auf dem Bild eine Maske trägt oder ob es sich um eine Person ohne Maske handelt. Da keine entsprechende Datengrundlage gefunden werden konnte, mussten die notwendigen Bilder eigenständig zusammengetragen werden. Die Bilder wurden mithilfe unterschiedlicher Suchmaschinen gesammelt. Hierbei wurden die Suchmaschinen Google, Bing und Pexels verwendet. Bei allen Suchmaschinen wurden die folgenden Suchbegriffe berücksichtigt: „person with medical mask“, „person with mask“, „covid masked face“, „face“, „person“. Aufgrund der Kostenrestriktionen im Rahmen des Projektes konnte nicht auf die von Google und Bing bereitgestellten Schnittstellen zurückgegriffen werden. Stattdessen wurde ein Web-Crawler mithilfe von Selenium realisiert, welcher automatisiert nach Bildern in den entsprechenden Suchmaschinen gesucht und heruntergeladen hat.

Für die Suchmaschine Pexels konnte hierfür hingegen die bereitgestellte Schnittstelle genutzt werden. Anhand der verwendeten Suchbegriffe wurden die gefundenen Bilder automatisiert in die jeweiligen Kategorien (Maske, keine Maske) eingeordnet. Diese Einordnung wurde im Anschluss manuell überprüft und gegebenenfalls angepasst.

Es wurden jeweils 500 Bilder gesammelt. Im Anschluss wurde eine Datenerweiterung durchgeführt, in welcher die Bilder mit zuvor definierten Anpassungen vervielfältigt wurden. Die Anpassungen umfassten das Spiegeln, Rotieren, Veränderungen von Sättigungen und das Hinzufügen von Gaußschem Rauschen auf die zuvor gesammelten Bilder. Durch dieses Vorgehen konnte die Datengrundlage besser an die sich wechselnden Umstände in Form von Lichtverhältnissen und Bewegungen in einem Ladengeschäft angepasst werden. Somit konnte ein Datensatz von 5.000 Bildern erstellt werden, welcher die Grundlage für das Training der im Folgenden beschriebenen Methode bildet.

## 6.5 Analyse

Im anschließenden Schritt, der Analyse, wurde ein Verfahren für die Klassifikation, ob eine Person eine Maske trägt oder nicht, der Daten aus dem Live-Video-Feed entwickelt.

### Identifikation geeigneter Analyseverfahren

Zu Beginn dieser Phase erfolgte eine Sichtung und Evaluation möglicher Verfahren und Methoden, welche für die Realisierung der Klassifizierung in Frage kommen könnten. Hierbei wurde zur Auswahl der zu evaluierenden Verfahren vor allem auf Erfahrungen aus vorherigen Data-Science-Projekten zurückgegriffen. Es wurde ebenfalls evaluiert, ob eine neue Methode entwickelt werden muss oder auf ein bereits vorhandenes Konzept zurückgegriffen werden kann.

### Anwendung von Analyseverfahren

Im Rahmen der Sichtung potenzieller Verfahren wurden zuerst die Support Vector Machine, der Random Forest, XGBoost, Naive Bayes, Convolutional Neural Networks (CNN) und Long Short-Term Memory Networks näher betrachtet und testweise umgesetzt. Hierbei wurden gängige Implementierungen mit Standardparametern genutzt. Die Evaluation erfolgte anhand der statistischen Gütekriterien Genauigkeit, Sensitivität und F1-Maß. Nach Auswertung der Tests wurde das CNN als vielversprechendster Kandidat ausgewählt. Im Anschluss erfolgte eine Sichtung vorhandener Netze für die Objekterkennung in Bildern. Im zweiten Schritt wurden fünf bereits trainierte Netze, welche von Google bereitgestellt werden, getestet. Diese sind das DenseNet, das SqueezeNet, das ResNet\_V2, das Mobilnet\_V1 und das Mobilnet\_V2. Alle Netze wurden auf ihre Zeit je Klassifikationsvorgang und die TOP5-Genauigkeit getestet. Tabelle 6-1 zeigt die Ergebnisse der Tests.

Tabelle 6-1: Evaluation Bildklassifizierungsmodelle

<b>CNN</b>	<b>Klassifikationsdauer</b>	<b>TOP5-Genauigkeit</b>
<i>DenseNet</i>	<i>195 ms</i>	<i>85%</i>
<i>SqueezeNet</i>	<i>36 ms</i>	<i>72%</i>
<i>ResNet_V2</i>	<i>526 ms</i>	<i>93%</i>
<i>Mobilenet_V1</i>	<i>1,7 ms</i>	<i>70,2%</i>
<i>Mobilenet_V2</i>	<i>17,5 ms</i>	<i>90,6%</i>

### Evaluation

Zur Ermittlung des geeigneten Netzes für das Projekt wurde anhand der TOP5-Genauigkeit eine erste Selektion vorgenommen. Dabei wurden die beiden Netze mit der höchsten Genauigkeit ausgewählt, diese sind das ResNet\_V2 (93%) und das Mobilnet\_V2 (90,6%). Anschließend wurde die Klassifikationsdauer in den beiden Netzen in Betracht gezogen. Diese ist relevant, da das Unternehmen die Maskenerkennung in Verbindung mit einem Live-Video-Feed nutzen möchte. ResNet\_V2 bietet zwar die höchste Genauigkeit, jedoch benötigt eine Klassifizierung ca. 0,5 Sek. Dies ist im Rahmen der Nutzung eines Live-Video-Feeds nicht praktikabel. Daher wurde das Mobilnet\_V2 als Grundlage für die Maskenerkennung genutzt. Es bietet eine leicht schlechtere Genauigkeit, jedoch beträgt die Klassifizierungsdauer nur 17,5 Millisekunden und ist damit für die Maskenerkennung in Video-Feeds in HD-Qualität geeignet. Das bestehende Mobilnet\_V2 wurde im Anschluss mit den Trainingsdaten des generierten Datensatzes erneut trainiert und durch weitere Layer auf den Anwendungsfall angepasst. Die Evaluation des Modells erfolgte auf Basis der Sensitivität und des F1-Maßes der Klassifikationsergebnisse auf den Testdaten.

## 6.6 Nutzbarmachung und Nutzung

Im Rahmen der Nutzbarmachung erfolgte die technisch-methodische Bereitstellung und die Sicherstellung der Anwendbarkeit. Um dies zu gewährleisten, wurden zwei zusätzliche Komponenten geschaffen. Zuerst wurde das Klassifikationsergebnis mit der gewünschten Fehlertoleranz auf dem Live-Video-Feed durch die Verwendung unterschiedlich gefärbter Boxen um das Gesicht der zu klassifizierenden Person kenntlich gemacht. Die Box um das Gesicht eines Menschen mit Maske wurde hierbei mit einer grünen Farbe hinterlegt. Die rote Farbe signalisierte, dass der Algorithmus zu der Entscheidung kam, dass keine Maske getragen wird. Zur Vereinfachung der Anwendung wurde für den Auftraggeber eine Applikation konzipiert und implementiert. Diese beinhaltet neben einer grafischen Benutzeroberfläche auch die Möglichkeit, den Schwellenwert der Klassifikationsentscheidung der entwickelten Methode manuell anzupassen. Im Rahmen des Prozessschritts der Nutzbarmachung werden Daten jeder Klassifikationsentscheidung in einer Datenbank gespeichert. Diese Daten sollen der Nachhaltung der Ergebnisse dienen und gegebenenfalls im weiteren Verlauf als neue oder zusätzliche Trainingsdaten zur kontinuierlichen Performancesteigerung der Methode verwendet werden können.

Anschließend wurde die Software-Lösung in die IT-Infrastruktur der BIG KAUF AG in Zusammenarbeit mit dem Domänenexperten eingebunden. In diesem Zusammenhang wurde die Verwendung der Klassifikationsergebnisse zur Ansteuerung der audio-visuellen Signalgebung umgesetzt. Hierfür wurde an den entsprechenden Eingangsbereichen des Unternehmens eine Kombination von Warnleuchte und -ton installiert und ein Monitor angebracht. Wenn eine Person, welche ein negatives Klassifikationsergebnis des Algorithmus ausgelöst hat, den zuvor definierten Bereich überschreitet, wird das audio-visuelle Signal ausgelöst und das aktuelle Kamerabild auf dem Monitor für kurze Zeit dargestellt. Dadurch soll die Person an die Einhaltung der aktuellen Hygienevorschriften automatisiert erinnert werden. Darüber hinaus wurde ein zentrales Monitoring der Applikation durch eine Kombination des Live-Video-Feeds mit abgespeicherten Informationen aus der Datenbank in Form eines Dashboards mit aggregierten Informationen bereitgestellt.

Perspektivisch ist ein Folgeauftrag zu der bisherigen Umsetzung denkbar. Als Projektziel könnte eine höhere Genauigkeit von beispielsweise 95% definiert werden. Für die Realisierung dieses Ziels könnten zwei Vorgehensweisen relevant sein. Zuerst könnte der Algorithmus mit den kontinuierlich weiter gesammelten Bildern, welche die Klassifikationsergebnisse enthalten, neu trainiert werden. Diese Daten müssen manuell auf die korrekte Durchführung der Klassifikation überprüft werden. Darüber hinaus ist es möglich, dass die Struktur des CNN um weitere Layer ergänzt werden muss. Des Weiteren könnte realisiert werden, dass nicht nur ein Signal ausgesendet, sondern auch eine Aktion ausgeführt wird. So könnte beispielsweise eine Erweiterung implementiert werden, welche die Ansteuerung einer Drehtür beinhaltet, sodass sich diese je nach Klassifikationsergebnis öffnet. Hiermit könnte sichergestellt werden, dass der Zutritt nur Personen, welche eine medizinische Maske tragen, gestattet wird. Dies ist bei der aktuellen Umsetzung ohne weiteren Personalaufwand nicht realisierbar. Durch einen anschließenden Projektauftrag könnte der Zutritt restriktiver umgesetzt werden. Ebenfalls ist es möglich, weitere mögliche Vorgaben im Rahmen von Hygieneverordnungen umzusetzen. Es ist etwa durch eine Anpassung der Datengrundlage möglich, zwischen der Verwendung von medizinischen Masken und FFP2-Masken zu differenzieren.