

Informatics Inside  
Herbst 2021

# Tagungsband

Herausgeber:

Uwe Kloos, Hochschule Reutlingen

Natividad Martinez, Hochschule Reutlingen

Gabriela Tullius, Hochschule Reutlingen

# Impressum



Hochschule Reutlingen  
Reutlingen University

## Anschrift:

Hochschule Reutlingen  
Reutlingen University  
Fakultät Informatik  
Human-Centered Computing  
Alteburgstraße 150  
D-72762 Reutlingen



Telefon: +49 7121 / 271-4002  
Telefax: +49 7121 / 271-4042

E-Mail: [infoinside@reutlingen-university.de](mailto:infoinside@reutlingen-university.de)  
Website: [infoinside.reutlingen-university.de](http://infoinside.reutlingen-university.de)

## Organisationskomitee:

Prof. Dr. rer. nat. Uwe Kloos, Hochschule Reutlingen  
Prof. Dr.-Ing habil. Natividad Martinez, Hochschule Reutlingen  
Prof. Dr. rer. nat. Gabriela Tullius, Hochschule Reutlingen

B. Sc. Dennis Grendatsch  
B. Sc. Felix Welker  
B. Sc. Irem Uslu  
B. Sc. Jana Ickenroth  
B. Sc. Hannah Kruck  
B. Sc. Manuel Haynes  
B. Sc. Maximilian Rabus  
B. A. Lara Krautmacher  
B. A. Pauline Weimann  
B. Sc. Patrick Beyersdorffer  
B. Sc. Philip Storz  
B. Sc. Stephan Bruhn

Verlag: Hochschule Reutlingen  
ISBN 978-3-00-070521-2



©2021 bei den Autoren. Lizenznehmer Hochschule Reutlingen, Deutschland. Dieser Artikel ist ein Open-Access-Artikel unter den Bedingungen und Konditionen der Creative Commons Attribution (CC BY)-Lizenz. <http://creativecommons.org/licenses/by/4.0/>

# Vorwort

Liebe Leserinnen und Leser,

seit über 12 Jahren findet nun die Informatics Inside als Informatikkonferenz an der Hochschule Reutlingen statt, in diesem Jahr zum zweiten Mal in einem halbjährigen Rhythmus, d.h. auch im Herbst. Diese Wissenschaftliche Konferenz des Masterstudiengangs Human-Centered Computing wird von den Studierenden selbst organisiert und durchgeführt. Sie erhalten während ihres Masterstudiums die Gelegenheit sich in einem selbstgewählten Fachthema zu vertiefen. Dies kann an der Hochschule, in einem Unternehmen, einem Forschungsinstitut oder im Ausland durchgeführt werden. Gerade diese flexible Ausgestaltung des Moduls „Wissenschaftliche Vertiefung“ führt zu einem sehr breiten Themenspektrum, das von den Studierenden bearbeitet wird. Neben der eigentlichen fachlichen Vertiefung spielt auch die Präsentation und Verteidigung von wissenschaftlichen Ergebnissen eine wichtige Rolle und dies weit über das Studium hinaus. Ein gewähltes Fachgebiet so allgemeinverständlich aufzubereiten und zu vermitteln, dass es auch für Nicht-Spezialisten verständlich wird, stellt immer wieder eine besondere Herausforderung dar. Dieser Herausforderung stellen sich die Studierenden im Rahmen der Herbstkonferenz zur Wissenschaftlichen Vertiefung am 24. November 2021. Bereits zum vierten Mal wird die Veranstaltung in einem online-Modus stattfinden, einschließlich eines virtuellen Begleitprogramms.

Das Themenspektrum der diesjährigen Herbstkonferenz ist wieder sehr vielfältig und breit gefächert. So erwarten Sie u.a. Beiträge aus dem Gesundheitssektor, dem Maschinellen Lernen, der KI und VR sowie dem Marketing und E-Learning. Allen gemein ist ein sehr starker Bezug zu innovativen Informatikansätzen, was sich auch in dem Wortspiel und Motto „RockIT Science“ der Konferenz widerspiegelt. Die Informatik durchdringt fast alle beruflichen und privaten Anwendungsbereiche und hat zunehmend größeren Einfluss auf unser tägliches Leben. Dies kann einerseits Besorgnis und andererseits Begeisterung auslösen. Gerade letzteres wollen die Studierenden mit Ihren Beiträgen erreichen und es auch mal im Informatiksektor „rocken“ lassen.

Mit diesem Tagungsband halten Sie die schriftliche Ausarbeitung der Fachbeiträge in der Hand und können sich selber ein Bild von den Arbeiten der Studierenden machen. Ich wünsche Ihnen eine spannende Lektüre, anregende Diskussionen und interessante Gespräche sowohl bei der Konferenz als auch beim anschließende virtuellen Treffen.

Reutlingen im November 2021

Prof. Dr. Uwe Kloos  
Studiendekan Human-Centered Computing

# Inhaltsverzeichnis

Steigerung der intrinsischen Motivation in E-Learning-Systemen durch psychologische Trigger . . . . .	1
Requirements for an Interactive Marketing Dashboard . . . . .	16
Dynamisches Matching für Peer to Peer Mitfahrgelegenheitssysteme . . . . .	29
Status Quo von künstlicher Intelligenz in Extended Reality-Anwendungen . . . . .	44
Überblick über maschinelle Lernverfahren für Predictive Maintenance . . . . .	56
Vorhersagen von Aktienkursen basierend auf Regressionsmodellen und Sentiment-Analysen . . . . .	68
Vergleich von maschinellen Lernalgorithmen zur Klassifizierung von wissenschaftlichen Abstracts . . . . .	80
A Scientometric Review about Current and Upcoming Trends of Artificial Intelligence in the Healthcare Sector . . . . .	98
Digitale Gesundheitslösungen zur Unterstützung von Patienten mit chronisch obstruktiver Lungenerkrankung: Eine systematische Literaturrecherche . . . . .	110
Anforderungen an ein Computer-Vision-System zur automatischen Leistungsdocumentation in der ambulanten Pflege . . . . .	124
Analyse des Krebsgenomik-Workflow-Management-Systems One Touch Pipeline . . . . .	137
Autorenverzeichnis . . . . .	150

# Steigerung der intrinsischen Motivation in E-Learning-Systemen durch psychologische Trigger

Manuel Haynes  
Hochschule Reutlingen

Manuel\_Samuel.Haynes@Student.Reutlingen-University.de

## Abstract

In Folge des menschlichen Wissensdurstes und der angestrebten Digitalisierung werden Lernsysteme auf ihre Fähigkeit der effizienten und motivierenden Lernerfahrung hin entwickelt. Ein durch soziale Netzwerke geprägter Ansatz für eine vermehrte und langanhaltende Nutzung kann der Einsatz von *psychologischen Triggern* darstellen. Diese Arbeit zeigt eine zusammengefasste Darstellung möglicher Ansätze und Anforderungen für die erfolgreiche Integration von *psychologischen Triggern* zur Erhöhung der *intrinsischen Motivation*. Hierbei wird ein Modell zur Anforderungsanalyse eines triggerbasierten E-Learning-System vorgestellt und der Gebrauch von Triggern anhand von bestehenden Applikationen aufgezeigt.

## CCS Concepts

•Applied computing → Education → E-learning; •Applied computing → Education → Interactive learning environments;

---

Betreuerin Hochschule: Prof. Dr. Gabriela Tullius  
Hochschule Reutlingen  
Gabriela.Tullius@  
Reutlingen-University.de

Informatics Inside Herbst 2021  
24. November 2021, Hochschule Reutlingen  
Copyright ©2021 Manuel Haynes

## Keywords

E-Learning, psychologische Trigger, Gamification, Motivation, triggerbasiertes E-Learning, E-Learning-Systemstruktur

## 1 Einleitung

Lernplattformen können den Prozess der Wissensseinprägung durch die Vielfalt der Darstellungsmöglichkeiten und dem Widerspiegeln realer Bedingungen nachweislich fördern. Studien in diesem Gebiet belegen eine positive Korrelation zwischen der Lernplattformnutzung und der daraus resultierenden Benotung [1]. Dieser Zusammenhang wird von Miller [2] bestätigt und zeigt eine Abhängigkeit des Interaktionsgrades mit dem zu lernenden Material und der tatsächlichen Wissensseinprägung auf. Der Prozess, eine Person für eine solche Plattform langfristig zu motivieren und für eine wiederkehrende und effektive Lernerfahrung zu sorgen, stellt Systembetreiber bei der Umsetzung jedoch vor eine große Hürde [1,3]. In den Bachelorarbeiten von Krebs [4] zur *Untersuchung psychologischer Trigger erfolgreicher Social-Media- und Streaming-Plattformen* und der darauf aufbauenden Evaluation von Aubele [5] zu Netflix und Instagram geht es um das Erkennen und Verstehen genutzter Trigger. Anhand von solchen Triggern werden bestimmte Reaktion bei Menschen, die das System nutzen, hervorgerufen und dienen somit als Auslöser für gewollte Interaktionen [4]. Aus Alblwi et al. [6] geht hervor, dass sich durch diese Methodik nicht nur die Verweildauer auf den

Plattformen positiv beeinflussen lässt, sondern auch direkte Interaktionen mit dem System schneller vollzogen werden. Hingegen beschreiben Krebes und Aubele [4,5] wie Menschen durch solche Trigger stärker zur Prokrastination neigen. Prokrastination beschreibt den Zustand des vor sich Herschiebens von Aufgaben, um dafür etwas anderes, meist Erfreulicheres, zu machen. Im Gegensatz dazu können die gleichen *psychologischen Trigger* auch für die vermehrte Nutzung eines E-Learning-Systems genutzt werden.

## 1.1 Ziel der Arbeit

Das Ziel dieser Arbeit ist es, anhand der erarbeiteten Grundlagen von Krebes und Aubele [4,5] in Kombination mit weiteren Recherchen, eine ganzheitliche Betrachtung über die Motivationssteigerung auf Lernplattformen durch den Einsatz von *psychologischen Triggern* aufzuzeigen. Auf Basis dessen ergeben sich folgende Fragestellungen, die im Laufe dieser Arbeit beantwortet werden:

- Welche Faktoren gibt es, die den Einsatz von *psychologischen Triggern* zur Motivationssteigerung mindern oder gar erfolglos machen?
- Welche Schritte müssen bei der Implementierung *psychologischer Triggern* in E-Learning-Systemen beachtet werden?
- Gibt es bereits existierende Umsetzungen, die Erfolge durch den Einsatz von Triggern aufweisen?

Für eine ganzheitliche Betrachtung wird eine Recherche in zwei Teilbereiche untergliedert. Der erste Bereich befasst sich mit dem Status quo der Wissenschaft und bildet Studien und Theorien zu den Themen E-Learning, Motivation, *psychologische Trigger* und Lernen ab. Dabei wird der Fokus auf Studien gelegt, die einen Nachweis des effektiven Triggereinsatzes und daraus resultierende

Steigerung der Motivation aufzeigen. Als Grundlage dienen, wie in 1.1 beschrieben, die Werke von Krebes und Aubele [4,5], die durch ihre Arbeiten den effektiven Einsatz *psychologischen Triggern* in Social Media Plattformen nachweisen konnten. Als Suchterme wurden die folgenden Wortkombinationen verwendet: Trigger, psychologische Trigger, Engagement Trigger, E-Learning, Motivation und Motivation in E-Learning. Gesucht wurde hierbei auf den Plattformen Google Scholar, Web of Science und Springer.

Der zweite Bereich der Recherche befasst sich mit bereits bestehenden Systemen, die auf den Trigger-Einsatz untersucht werden sollen. Hierbei wird die Methodik des *Cognitive Walktroughs* eingesetzt, bei dem eine Person mit Expertenwissen die Rolle eines Lernenden einnimmt und das interaktive System anhand von vordefinierten Kriterien durchläuft [7]. Anhand dessen können dann Aussagen und Erkenntnisse der Systeme verglichen werden. Da diese Methodik der Usability-Evaluierung entspringt, muss für die Analyse der Fokus des *Walktroughs* an die Trigger-Evaluierung angepasst werden.

## 1.2 Aufbau der Arbeit

Die weiteren Kapitel dieses Artikels beschäftigen sich im nächsten Schritt mit der *Motivationspsychologie* und dem darauf aufbauenden Trigger-Einsatz. Hierbei werden die einzelnen Arten von Triggern und der Prozess zum Aufbau einer Gewohnheit beleuchtet. Darauf aufbauend handelt das nächste Kapitel von den Grundstrukturen im *E-Learning* und der Spezifikation hin zu einem motivationsfördernden Aufbau. Kapitel 4 zeigt im Anschluss die Erkenntnisse aus den in 1.2 beschriebenen Rechercheteilen. Untergliedert wird dieses Kapitel in *Lernapplikations-Analyse, Leitung des Nutzerflusses mithilfe von Triggern* und *Einsatz der Trigger in E-Learning-Systemen*. Nach

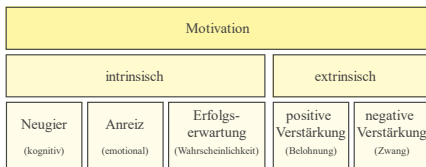
der Analyse werden im darauffolgenden Schritt die Ergebnisse in den Kontext der Übertragbarkeit und Implementierungsart gestellt. Zum Schluss werden die Erkenntnisse und die beantworteten Fragestellungen nochmals als *Fazit und Ausblick* dargestellt.

## 2 Motivationspsychologie

Das von einem Menschen erwartete Endergebnis, bei dem Energie für die Erreichung eingesetzt werden muss, wird als Motivation bezeichnet. Unterscheiden lässt sich dieser innere Motivationsprozess in den folgenden drei Determinanten [8]:

- *Motiv*: Das Motiv zielt auf die angeborenen Bedürfnisse oder auf die gelernten Verhaltensweisen ab.
- *Anreiz*: Durch die emotionale Bedeutung einer Sache wird ein Anreiz erzeugt.
- *Kognitive Prozesse*: Diese bilden Entscheidungen und Pläne auf Basis von Erwartungen ab.

Die gängigere Unterscheidungsart von Motivation bildet die Einwirkung aus einem Selbst heraus, die *intrinsische Motivation*, und eine äußere Einwirkung, die *extrinsische Motivation*. Auf Basis dieser Unterscheidung und der vorangegangenen Determinanten zeigt Edelman [8] in seiner Arbeit eine vereinfachte Übersicht von Motivation, die in Abbildung 1 zu sehen ist.



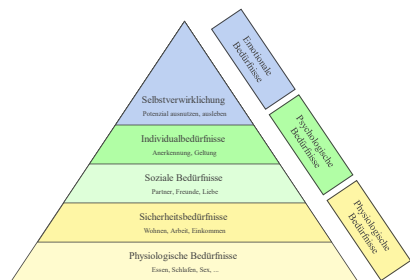
**Abbildung 1: Intrinsische und extrinsische Motivation im Überblick nach [8]**

Die Entscheidungen, die wir bewusst oder unbewusst treffen und die uns antreiben,

können durch die Triebtheorie von Taormina und Gao [9] erklärt werden. Bei dieser sind Bedürfnisse als Trieb in unserem Unterbewusstsein fest verankert. Der Mensch agiert aus einem Defizit seiner Bedürfnisse heraus, um diese zu stillen. Somit zielen alle Motivationsarten und Determinanten auf die Befriedigung der individuellen Bedürfnisse ab. Gegliedert werden diese nach Deci und Ryan [10] in drei Teilbereiche, die für den Menschen gleichermaßen wichtig sind. Die Erfüllung dieser Bereiche wird gleichermaßen angestrebt. Zu diesen zählen die emotionalen, physiologischen und psychologischen Bedürfnisse. Diese spiegeln sich ebenfalls in der Bedürfnispyramide nach Maslow wieder, der seine Aufgliederung von den Grundbedürfnissen hin zu immer abstrakteren Bedürfnissen wie der Selbstverwirklichung vornimmt [11], siehe Abbildung 2.

Ergänzend nennt Deci und Ryan [10] den Drang des Menschen, Motivation und Wohlbefinden zu erzeugen. Dies gilt für sich selbst und andere.

Für jede Art von Motivation kann ein Auslöser benannt werden, der zur Handlung anregt. Dieser innere Prozess wird im nächsten Unterkapitel im Detail beleuchtet und zeigt ebenfalls, wie Gewohnheiten aus diesem Prozess entstehen können. Im darauffolgenden Kapitel werden die Auslöser des Prozesses und ihre Kategorisierung beschrieben.



**Abbildung 2: Bedürfnispyramide nach Maslow [11]**

## 2.1 Langzeitmotivation

Um eine Langzeitmotivation aufzubauen, benötigt es einen Übergang von der durchgeführten Handlung hin zu einer regelmäßigen Verhaltensweise, die auch als Gewohnheit bezeichnet wird. Um zu verstehen wie dieser Übergang funktioniert, kann das Hakenmodell nach Eyal [12] verwendet werden, siehe Abbildung 3.

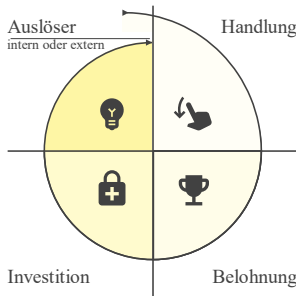


Abbildung 3: Hakenmodell nach Eyal [12]

Dieses untergliedert sich in vier Phasen und startet mit dem Auslöser beziehungsweise dem Trigger der Handlung. Der Mensch wird in diesem Schritt mit einem internen oder externen Reiz konfrontiert. Der interne Reiz entsteht durch einen inneren iterativen Prozess, bei dem eine Verknüpfung von Handlungen und Emotionen entsteht. Im Gegensatz zum internen Reiz taucht ein externer Reiz im Wahrnehmungsspektrum der Person auf.

Die zweite Phase beinhaltet die Handlung der Person, die ein spezifisches Ziel oder eine spezifische Motivation innehat. Eyal [12] merkt den Zusammenhang zwischen Usability, also der Benutzbarkeit, und der tatsächlichen Handlungsdurchführung an. Er beschreibt eine klare Korrelation zwischen der Einfachheit der Interaktion und der Bereitschaft die Umsetzung einer Handlung durchzuführen.

Nach der Handlung folgt die Phase der Belohnung bei der positive Gefühle durch das Bilden von dem Neurotransmitter Dopamin in verschiedenen Teilregionen des Gehirns

ausgelöst werden [13]. Durch das Stillen der Bedürfnisse und dem Ausschütten von Dopamin wird das wiederholte Handeln stark angeregt [4]. Doch auch das Erwarten einer Belohnung schüttet bereits Dopamin aus. So können zusätzliche Belohnungen für die Handlung miteingebracht werden, um den Drang des vermehrten Durchführens zu steigern [12].

Die Endphase ist das zu erbringende Investment, um den Zyklus neu zu starten. Hierbei muss, um eine dauerhafte Belohnung erhalten zu können, ein Gegenwert von Zeit, Aufwand, Daten oder Geld erbracht werden. Aus dieser Investition können dann weitere Auslöser hervorgehen, die den iterativen Prozess fördern [12]. Dieser Vorgang lässt sich somit als Spirale beschreiben, die den Menschen, je öfter die Wiederholungen stattfindet, immer stärker zu einer regelmäßigen Handlung führt.

## 2.2 Psychologische Trigger

Um sich näher mit den Auslösern zu beschäftigen, kann die Einordnung nach Alblwi et al. [6] in *soziale*, *Social-Network-Sites* (SNS) und *persönliche Trigger* betrachtet werden. Die Kategorie der *sozialen Trigger* befasst sich mit den Auslösern für das angestrebte Verbundenheitsgefühl der Menschen. Hierbei tritt der Mensch in Kommunikation mit anderen Menschen oder wird am Leben des anderen durch digitale Funktionen beteiligt [4]. Doch nicht nur das Bedürfnis der Verbundenheit spielt in dieser Kategorie eine Rolle, auch das Vermeiden einer Isolation kann diesen Trigger auslösen. Hinter dieser Trigger-Art stehen somit psychologische Bedürfnisse, die bei Maslow als Bedürfnis nach Zugehörigkeit näher spezifiziert sind [10]. Zu dieser Gruppe gehören die *Trigger Fear of Missing Out* (FoMO), *sozialer Einfluss* und *soziale Angst* [6].

Anders als die *sozialen Trigger* kommen die *Social-Network-Sites-Trigger* fast ausschließlich auf sozialen Netzwerken vor und basieren auf den Grundkonzepten des gegenseitigen Interagierens mittels Profilen. Untergliedert wird hier in *Benachrichtigungen*,



*immersives Design, Identität, Präsenzfunktion* und *Dialogfunktion* [6]. Je nach Trigger-Typ wird hier das Bedürfnis nach Zugehörigkeit oder Wertschätzung ausgelöst.

Die Kategorie mit der meisten Individualität stellen die *persönlichen Trigger* dar. Diese sprechen die höchste Stufe der Bedürfnispyramide nach Maslow an, die Selbstverwirklichung. Diese können auf verschiedene Art und Weise wirken und unterscheiden sich von Individuum zu Individuum sehr stark. Ein Beispiel hierfür ist das Ausleben von Talenten oder das Weiterbilden in Bereichen von eigenem Interesse. Unterteilt werden sie in *persönliche Situationen, geringes Selbstwertgefühl* und *geringe Selbstkontrolle* [6].

### 3 E-Learning

Das Lernen mit sämtlichen Arten mediengestützter Inhalten wird nach Baumgartner et al. als E-Learning bezeichnet [14]. Das Spektrum solcher Systeme reicht somit von der lokalinstallierten Softwareanwendung bis hin zu Apps oder Live Streams. Durch die Wiederverwendbarkeit der Inhalte gelten sie als kostenträgere Alternative zu herkömmlichem Lehrpersonal. Zusätzlich wird die losgelöste Nutzung von Zeit und Ort als einer der Präferenzgründe angegeben [3,14,15]. Anhand von soziopsychologischen Synergien, die sich bei dem kollaborativen Lernen ergeben und sich positiv auf die Wissensaneignung auswirken, wird der Einsatz von Systemen, die eine Interaktion der Lernenden miteinander zulässt, als effizienter beschrieben [2,16]. Dies kann durch Mischformen der Präsenz und Onlineblöcken umgesetzt werden oder durch Funktionalitäten, die das kollaborative Lernen ermöglichen.

In dieser Arbeit liegt der Fokus auf Systemen, die das selbstständige Lernen ermöglichen und die Integration von *psychologischen Triggern* ermöglichen. Hinzu kommt das Bestreben nach der Motivationssteigerung, die sich durch gewisse Aufbauten der E-Learning-Systeme begünstigen lassen. Auf Basis dessen werden in den folgenden Unterkapiteln gängige Aufbauten von E-

Learning-Systemen auf Basis ihrer *Lerntheorie* sowie ihrer Struktur und ihrem Inhalt betrachtet.

#### 3.1 Lerntheorien im E-Learning

In der heutigen Praxis dienen zunehmend Lerntheorien als Grundpfeiler für E-Learning-Konzepte. Durch den Einsatz solcher Theorien erhält der Konzeptionsprozess eine klare Ausrichtung und kann so effizienter erarbeitet werden. Baumgartner [17] nennt die folgenden drei Lernsysteme als gängigste: Behaviorismus, Kognitivismus und Konstruktivismus. Entscheidend für diese Arbeit ist der Grad der Selbstbestimmtheit beim Lernen.

Den geringsten Grad besitzt der Behaviorismus, der in deutschen Schulen angewandt wird. Als Basis dient die lehrende Person mit dem Wissen auf Korrektheit und den Erfahrungen, um ihr Wissen an Lernende weiterzugeben [17]. Daenecke [16] nennt als häufigsten Einsatz das Erlernen von Faktenwissen wie beispielsweise das Erlernen von Sprachen in Vokabeln.

Die Eigenverantwortung steigt im Kognitivismus mit dem Erhalt von vereinfachten Problemstellungen ohne weitere Aufgabenstellungen. Die lehrende Person agiert dabei als Tutor, die den Lösungsprozess begleitet und für Hilfestellungen zurate gezogen werden kann [17]. Um dies als System umsetzbar zu machen, muss dem Lernenden ein variabler Ablauf zur Hand gegeben werden, da die Lösungsansätze so individuell sein können wie die Personen, die das System nutzen. Zusätzlich wird die Berücksichtigung der persönlichen Bedürfnisse in den Fokus gerückt [16].

Im Konstruktivismus liegt dem Lernenden eine komplexe Aufgabenstellung vor, die durch seine persönlichen Erfahrungen in Aufgabenabfolgen und schlussendlich zu einem Lösungsweg führen sollen. Als Beobachter steht dem Lernenden die lehrende Person zur Verfügung, die als Coach nicht

alles wissen muss, sondern nur beratend interagiert [17].

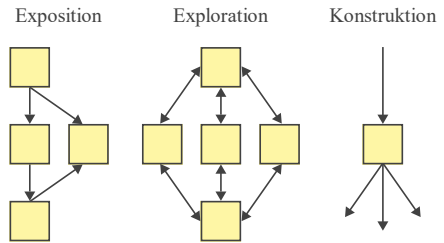
Die Wahl der Lerntheorie kann anhand folgender Faktoren betrachtet werden: Inhalt, Motivationssteigerung, Technik und Art des Lernens. Für den gezielten Einsatz der Lerntheorien muss anhand des zu erlernenden abgewägt werden, welche der drei Theorien die passende ist [17]. Als Beispiel hierfür kann das Erlernen von Vokabeln nur schwer mit der Lerntheorie des Konstruktivismus umgesetzt werden. Für die Steigerung der intrinsische Motivation sieht Daenecke [16] die Erhöhung der Selbstbestimmtheit als einen der Faktoren, weshalb der Konstruktivismus demnach den höchsten Motivationsgrad erzielen würde. Doch auch die technische Umsetzung spielt eine Rolle. Diese stellt einen negativen Zusammenhang zwischen der steigenden Selbstbestimmtheit und der programmatischen Komplexität dar. Ein weiterer Faktor bildet die Art des Lernens, die unterstützt werden soll. Hierbei unterscheidet die Literatur in formelles und informelles Lernen. Rohs [18] beschreibt in seiner Arbeit die verschwimmende Abgrenzung zwischen beiden Lernarten und wie sie als Kombination im Berufsleben Anklang finden. Hierbei ordnet er das informelle, erfahrungsgeleitete Lernen am ehesten dem Konstruktivismus zu. Der Behaviorismus stellt für ihn den Inbegriff des formalen und organisierten Lernens dar.

Als Fazit muss eine Abwägung im individuellen Fall vorgenommen werden, jedoch können auch Mischformen angestrebt werden. So empfiehlt es sich für ein motivationssteigerndes Konzept, welches eine mittlere programmatische Komplexität aufweisen soll, einen Mix aus Kognitivismus und Konstruktivismus anzustreben.

### 3.2 Struktur und Inhalte von E-Learning-Systemen

Wurde sich für eine Lerntheorie entschieden, muss ein didaktisches Struktursystem ausgewählt werden, das die jeweilige Theorie begünstigt. Schüpbach [15] beschreibt in seiner

Arbeit die drei gängigsten Systeme, die in Abbildung 4 zu sehen sind.



**Abbildung 4: Didaktische Struktur von E-Learning-Software [15]**

Die erste Struktur, die Exposition, bildet eine sequenzielle Abfolge ab und wird laut Schüpbach [15] am häufigsten eingesetzt. Durch den geführten Nutzerfluss und dem klar definierten Start- und Endpunkt kann dieses System dem Behaviorismus zugeordnet werden. Durch ein Netz aus miteinander verknüpften Blöcken kann der Lernende bei dem System der Exploration frei wählen. Ebenfalls bietet die Konstruktion eine freie Entscheidung der Blöcke an und stellt einen sequenziellen Aufbau dar [15]. Ähnlich wie bei einem Baumdiagramm können so universelle Lernwege wiedergegeben werden. Aus diesen zwei Struktursystemen ist es möglich, den Kognitivismus und Konstruktivismus systematisch nachzustellen. Hierbei können ebenfalls Kombinationen aus beiden Systemen verwendet werden, um ein möglichst reales Szenario mit ergänzenden Wissensmaterialien abzubilden.

In der granularen Betrachtung des Aufbaus besteht ein E-Learning-System aus Lernobjekten, die durch Medien für die Präsentation von Wissen verwendet werden [19]. Zu diesen Medien gehören Videos, Bilder, Texte, Grafiken oder Animationen. Für eine motivierende Lernerfahrung muss der Inhalt, trotz der sich immer wiederholenden Wissensblöcken so aufbereitet sein, dass es sich trotzdem immer individuell und neu anfühlt.

Um auch auf didaktischer Ebene Abwechslung zu bieten, kann die Betrachtung der

Lernobjekte mit ihrer jeweiligen Wissensart betrachtet werden. Anderson und Krathwohl [20] untergliedern diese in Faktenwissen, Zusammenhangswissen und Methoden- und Handlungswissen. Faktenwissen bilden Lernobjekte, die reine Fakten oder Szenarien beinhalten. Laut Daenecke wird hier meist mit Texten gearbeitet, die durch Bilder, Grafiken, Tabellen und Animationen ergänzt werden. Doch auch Videos seien für diese Wissensart nicht unüblich. Beim Zusammenhangswissen werden Argumente und Fakten in Korrelation gestellt und mit ähnlichen Medien wie dem Faktenwissen dargestellt. Die Häufigkeit der Videoverwendung sinkt in dieser Wissensart ab. Methoden- und Handlungswissen hilft den Lernenden sich themenspezifisches Wissen über Methoden anzueignen und in den jeweiligen Kontexten zu adaptieren. Bei diesem Wissen werden häufig Animationen mit Text verwendet, die oftmals auch interaktiv genutzt werden können. Je nach Themengebiet wird eine andere Zusammensetzung dieser drei Wissensarten verwendet [20]. Trotz einer Übergewichtung des Faktenwissens sollte dennoch eine durchmischte Abfolge angestrebt und durch Interaktionen und Aufgaben erweitert werden.

Bei der Gliederung der Lernobjekte in thematische Bereiche muss darauf geachtet werden, den Lernenden dauerhaft das Ziel des aktuellen Bereichs vor Augen zu halten. Auch muss der Stand des Lernenden stets ersichtlich sein. Abhängig vom didaktischen Konzept muss des Weiteren darauf geachtet werden, ob die Lernobjekte in Abhängigkeit stehen und ob der Lernende jederzeit das benötigte Basiswissen erlernt oder angezeigt bekommt.

## 4 Trigger in E-Learning

In diesem Kapitel werden zuerst die Ergebnisse der Analyse von bestehenden Lernapplikationen betrachtet, die im Rahmen dieser Arbeit durchgeführt wurde. Anhand dessen sowie der Social Media Analysen von Krebses und Aubele [4,5] werden dann in den weiteren Unterkapiteln die

Adaption und Integration der von Alblwi et al. [6] definierten Trigger betrachtet.

### 4.1 Lernapplikations-Analyse

Für die Arbeit wurden 50 Applikationen auf einem Smartphone mit Betriebssystem iOS analysiert. Die Wahl der Kriterien fand in Anlehnung an Krebses [4] statt und ergab folgende Punkte für die Betrachtung:

1. Welche Trigger bietet das interaktive System?
2. Welche Rolle spielt das Design der verwendeten Trigger?
3. Welche Trigger motivieren den Lernenden das interaktive System möglichst lange zu nutzen?
4. Was lösen die Trigger des interaktiven Systems aus?

Anhand dieser Kriterien können die einzelnen Systeme und ihre Trigger verglichen werden. Durch eine Betrachtung der Themengebiete der einzelnen Applikationen bildet sich folgende Aufteilung:

- 11x Sprachenlernen
- 3x Jura
- 2x Programmierung
- 29x Finanzen
- 5x Sonstige

Bei sieben der Applikationen wurden die E-Learning-Elemente ergänzend für die Hauptfunktionalitäten verwendet. Somit boten diese für ihren speziellen Fall Lernobjekte an. Diese Objekte wurden genutzt, um Grundwissen für die Interaktion mit dem System zu vermitteln oder den Gebrauch der Applikation zu optimieren. Die Auswahl fand anhand der App Store Charts in den Themenfeldern Bildung und Finanzen statt.

Kriterium eins, also der Einsatz von Triggern in den Systemen wird in Tabelle 1 auf Basis der Häufigkeit aufgelistet. Des Weiteren wird in der Tabelle betrachtet, welche der Trigger die Lernenden zur App-Nutzung

motivieren. Hierbei steht fünf für eine starke Motivation und eins für gar keine Motivation. Festgelegt wurde dieser Wert an dem dahinterstehenden Bedürfnis und der objektiven Wahrnehmung beim Cognitive Walkthrough.

**Tabelle 1. Übersicht der genutzten Trigger mit ihrem Durchschnittswert von geschaffener Motivation**

Trigger	Anzahl der Apps	Durchschnittswert der Motivationssteigerung
Benachrichtigungen	12	4
Sozialer Einfluss	9	1
Persönliche Situation	8	4
Soziale Angst	7	2
Identität	4	3
FoMO	3	5
Dialogfunktion	3	2
Präsenzfunktion	2	1
Immersives Design	1	1
Geringes Selbstwertgefühl	(nicht bewertbar)	
Geringe Selbstkontrolle	(nicht bewertbar)	

Die erkennbare Nutzung von Triggern beläuft sich bei den geprüften Applikationen auf vier Apps (Duolingo, Quizduell, Knowunity und Jurafuchs), die vermehrt Trigger aufweisen, Acht, die den Einsatz integriert hatten und 38 Applikationen ohne integrierte Trigger. Zu den acht Apps gehören: Babel, Fastic, 7Mind, Memrise, Blinkist, Simple Club, Busuu und Tandem. Somit beläuft sich die Anzahl der in Tabelle 1 betrachteten Applikationen auf zwölf Stück. Im Kontext der Themenfelder ist der Bereich des interaktiven Sprachenlernens mit 41,67% am stärksten vertreten. Darauf folgen unterschiedliche Themenfelder, die sich nicht gruppieren lassen. Die Analyse der Trigger *geringes Selbstwertgefühl* und *ge-*

*ringe Selbstkontrolle* konnten bei der Analyse nicht eindeutig festgestellt werden und benötigen weitere Untersuchungen. Bei der Motivationsbewertung muss in Funktionalitäten, die den Lernenden unterstützen oder seine Neugier befriedigen, unterschieden werden. Auf Basis dessen wurden *FoMO*, *Benachrichtigung* und *persönliche Situation* als potenziell stärkster Antrieb für den Motivationsgewinn angesehen.

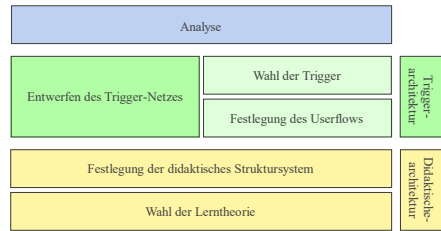
Die Analyse zeigt eine Tendenz zu den *sozialen Triggern*. Hierbei findet der Trigger des sozialen Einflusses am häufigsten Einsatz, um neue Personen anzuwerben. Der Trigger der Benachrichtigung wird in allen Applikationen verwendet und in fünf Applikationen zusätzlich mit E-Mail-Benachrichtigungen erweitert. Die Trigger selbst werden vom Design durch Größenunterschiede oder farbliche Hervorhebung prägnanter angezeigt. Auch Animationen, wie das Aufblinken von Buttons kommen in sechs Systemen vor. Um somit die Analyse von Kriterium zwei zusammenzufassen wird beim Design darauf geachtet, den Trigger optisch möglichst hervorzuheben.

Die höchste Motivation zur Appnutzung, welche bei Kriterium drei betrachtet wurde, kam bei *FoMO*, *Benachrichtigung* und *persönliche Situation* vor. Viele der Trigger zielen auf die soziale Interaktion zwischen den Lernenden ab, je bekannter der Gegenüber desto motivierter findet der Austausch oder Konkurrenzkampf im Lernen statt. Diese These muss anhand von weiteren Untersuchungen bekräftigt werden. Durch die Trigger wird versucht, die Interaktion zwischen den Lernenden zu verstärken oder ein Lernobjekt mehr zu konsumieren als beim Eintritt in die Applikation. Doch auch das Anwerben von neuen Lernenden wird durch Trigger angeregt. Somit unterteilt sich die Analyse von Kriterium 4 in:

- Konsum des Inhalts
- Interaktion mit anderen
- Anwerben von neuen Lernenden

## 4.2 Leitung des Nutzerflusses mithilfe von Triggern

Um ein E-Learning-System mit *psychologischen Triggern* zu erweitern muss die didaktische Architektur des Systems betrachtet werden. Hierbei liegt das Hauptaugenmerk zu Beginn auf der Lerntheorie und wird durch das daraus erstellte didaktische Struktursystem erweitert, welches in Kapitel 2.1 und 2.2 beleuchtet wird. Auf Basis dessen können erste Skizzen des Systems angefertigt werden, die den Fluss des Lernenden abbilden. Je nach System können hier unterschiedliche Architekturen verwendet werden. Der gängigste Aufbau eines E-Learning-Systems besteht im Kern aus einer Startseite mit dem aktuellen Lernstand des Systems und Lernobjekt-Gruppen, die entsprechend der Vorkenntnisse des Lernenden aufgebaut sind. Alle weiteren Funktionalitäten dienen als eine Erweiterung der Hauptfunktionalität und sollten deshalb nicht im Fokus stehen. Für einen effektiven Trigger Einsatz gilt es dann im Einzelnen abzuwägen, was das Ziel des Triggers sein soll und wie dieser dann an Funktionalität oder Aussehen umgesetzt werden soll. Ist dieser Prozess abgeschlossen, erhält man ein Trigger-Netz, welches sich über das E-Learning-System legt und den Lernenden verstärkt in die Hauptfunktionalitäten des Systems leitet. Die Herangehensweise für einer solchen Analyse findet sich in Abbildung 5 wieder. In den Arbeiten von Krebs und Aubele [4,5] wird ein solches Netz von aufeinander aufbauenden und verstärkenden Triggern aufgezeigt. Somit kann von einer gezielten Verkettung bei Social Media Plattformen ausgegangen werden, die es auf den Bereich des E-Learnings zu übertragen gilt.

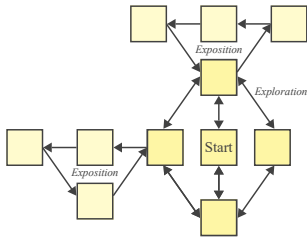


**Abbildung 5: Modell zur Analyse eines triggerbasierten E-Learning-Systems**

## 4.3 Einsatz von Triggern in E-Learning-Systemen

Bei dem Einsatz von Triggern in E-Learning-Systemen kann das in 4.2 beschriebene Modell verwendet werden. Als Grundlage des Modells wird im Vorhinein festgelegt, welchen Inhalt das System hat und was das Ziel des Systems ist. Auf Basis dessen kann dann die Lerntheorie festgelegt werden. Wie in 3.1 beschrieben muss hierbei zwischen dem Freiheitsgrad zur Entfaltung des eigenständigen Lernens und der programmatischen Komplexität abgewägt werden. Da es in dieser Arbeit um die Motivationssteigerung geht, wird der weitere Verlauf mit dem Kognitivismus beschrieben. Somit bildet das System einen virtuellen Tutor ab, der bei der Exploration der Lernobjekte unterstützen soll.

Um die didaktische Architektur zu vollenden wird nun ein didaktisches Struktursystem für die Lernobjekte festgelegt. Für den digitalen Tutor kann eine Kombination aus der Exploration und Exposition genutzt werden. Anhand dieser zwei kann eine Basis von Lernobjektgruppen erzeugt werden, die inhaltlich koexistieren und nicht aufeinander aufbauen. Zugleich können sie aber durch die Exposition je nach Bedarf des Lernenden in tiefergehende Verzweigungen münden. Abbildung 6 zeigt einen solchen Aufbau.



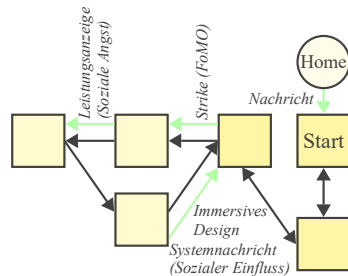
**Abbildung 6: Beispielhafte didaktische Struktur eines E-Learning-Systems**

Durch die Pfeile in der Abbildung werden die möglichen Pfade der Lernenden aufgezeigt. Beim Aufbau dieser Pfade ist darauf zu achten, dass keine Endpunkte entstehen, sodass der Lernende immer weiß, was er als nächstes zu tun hat und nicht aus dem Lernfluss kommt. Solche Übergänge der Endpunkte sorgen für den Trigger *immersives Design*. Um die Lernenden zu motivieren, das System nicht nach einem Lernobjekt zu verlassen, kann die Strike Funktion aus der Sprachenlernapp Duolingo verwendet werden. Diese belohnt die Lernenden, wenn sie drei Lernobjekte in Folge beenden. Hierfür erhalten sie extra Punkte. Die Funktion basiert auf *FOMO* und *soziale Angst*, da durch die Punkte der Rang zu anderen Lernenden verändert wird. Durch die erhaltenen Punkte wird auch der Trigger der persönlichen Situation angeregt. Durch die positive Bestätigung des Systems vor anderen Lernenden können positive Gefühle beim Lernenden aufkommen. Eine weitere Möglichkeit bieten Meldungen über das System. Hierbei wird bei Beendigung von definierten Lernobjekten eine Glückwunschnachricht mit Namen des jeweiligen Lernenden über das gesamte System versendet. Diese Nachricht löst den Trigger des sozialen Einflusses aus, da die anderen Lernenden von den Nachrichten in ihrem Handeln beeinflusst werden.

Für eine fokussierte Durcharbeitung eines Lernobjekts kann eine Anzeige am Ende des Objekts sorgen, die dem Lernenden seine eigene und die durchschnittliche Leistung anzeigt. Anhand dessen kann der Trigger des

*sozialen Einflusses*, der *sozialen Angst* und der *persönliche Situation* ausgelöst werden. Zudem kann so auch ein Anreiz gesetzt werden, um noch ein Lernobjekt mit einer besseren Leistung zu beenden.

Für den Einsatz dieser Funktionalitäten muss die Applikation zuerst gestartet werden. Dies wird in den untersuchten Anwendungen durch den Trigger der Benachrichtigung versucht. In diesem Fall den Lernenden News, neue Inhalte oder einfache Erinnerungen zu gespielt. Hierbei lassen sich viele Ähnlichkeiten zu Instagram und Netflix finden. Im Home-Menü des Smartphones werden Neuigkeiten in der Applikation durch rote Hervorhebungen am Appicon symbolisiert. Auch *Benachrichtigung* über neue Funktionen oder Inhalte auf dem Homescreen oder in der App sind bereits in den gängigeren E-Learning-Systemen vertreten. Als Ergänzung zu den in Social Media festgestellten *Benachrichtigungen* wird in E-Learning-Systemen vermehrt bei der lernenden Person nach sogenannten Lernzeitpunkten gefragt. Dieser wird dann in einem festgelegten Zyklus verwendet, um den Lernenden an das Lernen zu erinnern.



**Abbildung 7: Beispielhafte didaktische Struktur mit integrierten Triggern**

Durch die eben genannten Beispielen lassen sich die Wege der Lernenden in der Applikation ergänzen und verstärken den Drang den Pfad weiterzugehen, siehe Abbildung 7. Hierbei kann frei an Triggern und Einsatzorten variiert werden.

Eine weitere Möglichkeit für den Triggereinsatz in E-Learning-Systemen bietet das Bestreben, Lernziele zu erreichen oder eine Community an Lernenden aufzubauen. Die weiteren Unterkapitel beschreiben weitere Möglichkeiten auf Basis von bestehenden E-Learning Systemen oder Übertragungen aus Sozialen Plattformen. Um die Übersichtlichkeit zu gewährleisten, gliedern sich die Kapitel in die Trigger Kategorien.

### 4.3.1 Soziale Trigger

Der weitere Einsatz des *FoMO* im Kontext eines E-Learning-Systems kann anhand des Beispiels der Applikationen Quizduell und Duolingo aufgezeigt werden. In der Allgemeinwissensapplikation Quizduell tritt die lernende Person gegen andere Lerner in einem rundenbasierten Quiz an. Ist eine Person am Zug, muss sie die Runde innerhalb von 24 Stunden beenden, andernfalls wird die Partie abgebrochen. Um sein Gegenüber nicht zu enttäuschen, wird je länger die Partie vorangeschritten ist, versucht diese 24 Stunden nicht ablaufen zu lassen. Die Konsequenz daraus ist die vermehrte App Nutzung. Das zweite Beispiel stammt aus Duolingo. Der Lernende erhält für jeden Tag an dem er die App nutzt weitere Erfahrung. Tut er das nicht, so erhält er keine Erfahrungspunkte und fällt im Punktstand im Vergleich zu seinen Mitlernenden zurück. Zusammenfassend kann *FoMO* als Druck zwischen zwei Lernenden verwendet werden oder, um Belohnungen reizvoller zu gestalten, wie es auch in der Arbeit von Maderthaler [21] beschrieben wird. Aber auch das temporäre Teilen von Lernerfolgen könnte einen ähnlich erfolgreichen Einsatz des *FoMO* mit sich bringen. Die Funktionalität des temporären Teilens von Inhalten wird im Kontext von Social Media als sehr effektiv beschrieben [4,5]. Dies ist einer der Gründe weshalb die Funktionalität von vielen sozialen Plattformen nun integriert wurde. Alle drei Beispiele sorgen für einen weiteren Trigger, der die Handlung des Applikationsöffnens anregt. Die Belohnungen hierfür sind Stillung der Neugier, Anerkennung oder

eine negative Verstärkung im Falle des Quizduells.

Um den Trigger des sozialen Einflusses zu nutzen, wird auf sozialen Plattformen vermehrt die Funktion des Teilens angeboten, bei dem Inhalte an Freunde weitergeleitet werden können [5]. In den E-Learning-Applikationen kommt es zu ähnlichen Funktionalitäten. Freunde können in die App eingeladen werden, was vermehrt mit Vergünstigungen als Belohnung beworben wird und den Trigger somit verstärkt. Auch das Teilen von einzelnen Textpassagen als Zitat wird in einigen Applikationen bereits angeboten. Beide Funktionen werden in Blinkist genutzt, welches Sachbücher in Kernaussagen zusammenfasst. Der Zweck der Funktionalitäten und damit des Triggers kann dem Communityaufbau zugeschrieben werden. Hervorzuheben ist der positive Einfluss den befreundete Lernende aufeinander haben können.

Der letzte *sozialen Trigger* wird nur indirekt in bestehenden E-Learning-Applikationen verwendet. Durch eine öffentliche Liste mit Erfahrungsstand oder Freundeslisten mit einsehbaren Profilen kann *soziale Angst* ausgelöst werden. Systeme, die diese Funktionalität nutzen sind: Duolingo, Quizduell und Jurafuchs. Als Basis des Triggers findet keine direkte Negativbewertung statt, jedoch kann eine Negativassoziation mit einem niedrigeren Rang oder Stufe einhergehen. Um dies zu verhindern bieten manche Systeme, ähnlich wie in den sozialen Plattformen, an, Profile anonym und damit nicht dem Namen zugeordnet oder sogar nicht für andere sichtbar zu schalten. Der Trigger *soziale Angst* wird somit in bestehenden Systemen so verwendet, dass die Individualbedürfnisse durch eine vermehrte Nutzung befriedigt werden.

### 4.3.2 Social-Media-Site-Trigger

Auch im Bereich der SNS-Trigger finden sich viele Funktionalitäten in den E-Learning-Systemen wieder. In 4.3. wurde bereits der Einsatz von *Benachrichtigung* aufgezeigt. Weshalb sich dieses Kapitel auf den

Einsatz der Trigger *immersives Design* und *Identität* fokussiert.

*Immersives Design* entsteht in den sozialen Medien durch infinite Scrolling, Pull-to-refresh (Ptr) und Pagnation [4]. Aufgrund der häufig genutzten linearen Struktur und der zu geringen Anzahl von Lernobjekten, können solche Navigationen oft nicht genutzt werden. Um trotz des Aufbaus infinite Scrolling zu ermöglichen, muss Content auch von den Personen, die das System nutzen, kommen. So kann ein zeitlicher Verlauf etabliert werden, der Spielraum für Ptr und Infinite Scrolling zulässt. Als Beispiel kann hierfür die Applikation Knowunity betrachtet werden. Diese kann von den Lernenden als Plattform für den Austausch von Lernzetteln und Präsentationen genutzt werden. Ein weiteres Beispiel für die Menüstrukturierung bietet die Meditationsapplikation 7Mind. In ihr werden die Lernobjekte ähnlich wie in Netflix in Kategorien untergliedert. In den Kategorien selbst wird seitliches Infinity Scrolling genutzt. Für die klassische Pagnation, also dem Auflisten von Daten in unterschiedlichen Seiten mit Seitenangaben, können Nachschlagewerke integriert werden, bei denen die lernende Person die Inhalte entweder durch Suche oder durchs willkürliche Stöbern durchgeht. Um den Ptr zu etablieren kann eine Box eingeblendet werden, die Tipps oder kleinere Wissensinhalte anzeigt. Diese kann sich jedes Mal neu laden, wenn Ptr aktiviert wird. Gerade bei dem Trigger *immersives Design* muss das Ziel des Systems im Vordergrund stehen. Es muss abgewägt werden, inwieweit ein Stöbern gefördert werden sollte, da es auch zur Prokrastination führen kann.

Der Trigger *Identität* wird im Bereich des Profils aufgebaut. Keines der getesteten Appprofile hat ausreichend Individualisierungsmöglichkeiten, um den Trigger in seiner Gänze zu nutzen. In den betrachteten Applikationen konnte nur ein Profilbild, Beschreibung und Nutzernamen angelegt werden. Welche Auswirkungen die Anzeige des Wissensstandes im Zuge des Profils auf den Trigger hat, muss in einer zukünftigen Arbeit

noch erforscht werden. Auch kann das Selbsterstellen eigener Lernobjekten mit diesem Trigger gefördert werden.

Der Trigger *Dialogfunktion* wird in den meisten Fällen mit dem Trigger *Präsenzfunktion* verstärkt.

Die *Präsenzfunktion* kann so verwendet werden, dass befreundete Lernende eine Nachricht bekommen (*sozialer Einfluss*), wenn der jeweils andere am Lernen ist. Eine weitere Möglichkeit ist das Einblenden des Onlinestatus in der Bestenliste. Wird der Trigger der *Präsenzfunktion* betrachtet muss im gleichen Zug *Dialogfunktion* betrachtet werden. Dieser wird meist durch den Einsatz von *Präsenzfunktionen* verstärkt. Beispiele für den Einsatz bilden hierfür FAQs, Foren und Chats. Bei den Chats zwischen Lernenden können Awareness Patterns eingesetzt werden, die der Person zeigen, dass ihr gegenüber gerade schreibt oder etwas aufnimmt [4]. Eine visuelle Verstärkung kann durch ein Icon vorgenommen werden, welches die Präsenz des Gegenübers auf dem Chat anzeigt. Der Trigger *Präsenzfunktion* findet somit seinen Einsatzort bei allen Berührungspunkten zwischen Lernenden.

### 4.3.3 Persönliche Trigger

Die persönlichen Trigger können im Gegensatz zu den beiden anderen Kategorien nur indirekt zu Funktionen zugeordnet werden. Sie ergeben sich in der Nutzung des Systems und wirken sich individuell auf den Lernenden aus. Um einen Trigger in dieser Kategorie zu verwenden kann immer die Überlegung angestellt werden, wie die Funktion das Bedürfnis nach Selbstverwirklichung oder die Individualbedürfnisse befriedigt.

Im Kontext eines Triggers *persönliche Situation* muss die App die persönliche Entwicklung und das Ausleben möglich machen. Dies passiert im Bereich von Social Media innerhalb des Profils und kann im E-Learning-System auf Lernprofile übertragen werden. In diesen von anderen einsehbaren Profilen wird das Wissen des Lernenden und seine Persönlichkeit wiedergespiegelt.



Der Trigger *geringes Selbstwertgefühl* kann durch das Erreichen von Lernerfolgen oder dem Zusammenschließen von Gleichgesinnten Lernern verbessert werden.

Der letzte Trigger *geringe Selbstkontrolle* kann durch die produktive Nutzung des Systems abgebaut werden. Der Trigger selbst sollte zudem in einem E-Learning-System vermieden werden, da dieser entgegen des Zwecks eines solchen Systems verläuft. Dieser lautet, den Lernenden die Lernobjekte effektiv zu vermitteln.

## 5 Fazit und Ausblick

In der Arbeit konnten viele unterschiedliche Einsatzmöglichkeiten von Triggern aufgezeigt werden. Wie die Arbeit von Aubele [5] aufzeigt, können Trigger allerdings unterschiedlich wahrgenommen werden, weshalb auch nur von positiven Tendenzen bei der Motivationssteigerung der Lernenden gesprochen werden sollte. Eine eindeutige Minderung beim Einsatz von Triggern konnte in der Arbeit nicht festgestellt werden. Nur die Wahl der Funktionen, die den Trigger auslösen, können eine negative Auswirkung auf den Lernerfolg eines E-Learning-Systems haben und sollten sich deshalb nach dem E-Learning-System richten.

Um also eine erfolgreiche Implementierung von Triggern sicherzustellen, sollte die erste Analyse des E-Learning-Systems mit dem Modell, welches in 4.2 beschrieben und in 4.3 an einem kurzen Beispiel erklärt wurde, vorgenommen werden. Auf Basis des Modells werden Grundbedingungen aufgezeigt, die die Motivationssteigerung beim Aufbau eines solchen Systems begünstigen. Durch die in Didaktische- und Trigger-Architektur gegliederte Analyse kann eine strukturierte Planung sichergestellt werden. Des Weiteren können Synergien zwischen dem bestehenden Nutzerfluss und den Triggern hergestellt werden, um zum einen eine lückenlose Integration zu bewerkstelligen und zum anderen die Lernenden dazu zu motivieren, mehr zu lernen. Dieses Modell kann als Grundlage für andere Arbeiten dienen und könnte hierbei durch weitere Ebenen wie dem Einsatz

von Gamification oder adaptives Lernen ergänzt werden.

Anhand der Analyse von Applikationen und Literatur konnten verschiedene existierende Umsetzungen für den erfolgreichen Einsatz von Triggern nachgewiesen werden. Ebenfalls konnten dadurch wichtige Erkenntnisse über den Einsatz und ihre Anforderungen an das System festgehalten werden. An dieser Stelle muss jedoch auch die Limitation in der Vollständigkeit der Arbeit beschrieben werden. Hierbei konnte im Rahmen des Projekts nur eine Grundlage an Literatur und Applikationen betrachtet werden. Auch beschränkte sich die Analyse nur auf mobile Applikationen. Es ist davon auszugehen, dass es Unterschiede zur Nutzung auf Desktopanwendungen gibt. Des Weiteren ist der Cognitive Walkthrough nur für ein Überblick gedacht. Um die Auswirkungen von Triggern näher zu untersuchen, kann eine ähnliche Vorgehensweise wie in der Bachelorarbeit von Aubele [5] verwendet werden. Dabei wurde eine behavioristische Studie auf Basis des DECIDE Frameworks durchgeführt. Dieses Framework hilft dabei, eine strukturierte Planung und Durchführung bei einer Evaluation zu bewerkstelligen.

Bei der Betrachtung der Adaptierbarkeit *psychologischer Trigger* der Social Media Plattformen wird eine Annäherung der Funktionalitäten vorgenommen, sodass man von einem Social-E-Learning-System sprechen kann. Dieses kombiniert die Interaktion der Menschen, die das System nutzen, mit den Inhalten und Kontexten eines E-Learning-Systems. Durch das Bilden einer Lern-Community können soziale Bedürfnisse während des Lernprozesses befriedigt werden. Auf Basis dieses Grundgedankens muss hinterfragt werden, inwieweit die Vermengung der Funktionalitäten mit dem Zweck eines E-Learning-Systems übereinstimmen. Durch soziale und kollaborative Funktionalitäten kann die Motivation der Nutzung zwar gesteigert werden, doch können im gleichen Zuge dieselben Funktionen den Lernenden von den Lernobjekten abhalten. Dieser Grad an Vermengung bietet Potenzial für weitere

Untersuchungen, die eine Limitierung der Community-Funktionen in einem solchen System analysieren und abgrenzen kann.

## Literaturverzeichnis

- [1] Sclater, N. e. A. 2016. *Learning Analytics in Higher Education. A review of UK and international practice*. Jisk.
- [2] Miller, G. E. 1990. The assessment of clinical skills/competence/performance. *Academic medicine: journal of the Association of American Medical Colleges* 65, 9 Suppl, S63-7.
- [3] Rose, J., Dahm, M., Czarnetzki, M., and Köhler, M. 2020. *Steigerung der Motivation und des Lernerfolgs von Programmieranfängern mithilfe von Gamification*. DOI=10.18420/muc2020-ws103-267.
- [4] Krebes, E. 2021. *Untersuchung psychologischer Trigger-Mechanismen erfolgreicher Social-Media- und Streaming-Plattformen*. Bachelorarbeit, Hochschule Reutlingen.
- [5] Aubele, A. 2021. *Studie zur Untersuchung der Wahrnehmung und Wirkung psychologischer Trigger-Mechanismen am Beispiel von Instagram und Netflix*. Bachelorarbeit, Hochschule Reutlingen.
- [6] Alblwi, A., Stefanidis, A., Phalp, K., and Ali, R. 2019. Procrastination on Social Networks: Types and Triggers. In *2019 6th International Conference on Behavioral, Economic and Socio-Cultural Computing (BESC)*. IEEE, 1–7. DOI=10.1109/BESC48373.2019.8963036.
- [7] Lewis, C., Wharton C. 1997. Cognitive Walkthroughs. In *Handbook of Human-Computer Interaction*. Elsevier Science & Technology, Amsterdam, 333–347.
- [8] Edelmann, W. Intrinsische und extrinsische Motivation. In *Grundschule: Ideen, Erfahrungen, Konzepte*, 30–32.
- [9] Taormina, R. J. and Gao, J. H. 2013. Maslow and the motivation hierarchy: measuring satisfaction of the needs. *The American journal of psychology* 126, 2, 155–177.
- [10] Deci, E. L. and Ryan, R. M. 1993. Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39, 223–238.
- [11] Maslow, A. H. 1993. *The farther reaches of human nature*. Arkana, New York, N.Y., U.S.A.
- [12] Eyal, N. 2019. *Hooked. Wie Sie Produkte erschaffen, die süchtig machen*. REDLINE Verlag, München.
- [13] Makalesi, A., et al. 2019. A Research On Social Media Addiction and Dopamine Driven Feedback. *Journal of Mehmet Akif Ersoy University Economics and Administrative Sciences Faculty*, 882–897.
- [14] Baumgartner, P., Häfele, H., and Maier-Häfele, K. 2002. E-Learning: Didaktische und technische Grundlagen. In *CD Austria*, 4–29.
- [15] Schüpbach, E. et al. 2003. *Didaktischer Leitfaden für E-Learning*. Pädagogik. h.e.p.-Verl., Bern.
- [16] Daenecke, R. 2008. *Aufbau des E-Learning-Systems - Informationssysteme und ihre Vernetzung in der Milcherzeugung - bei besonderer Beachtung einer entwicklungsbegleitenden Evaluation*. Dissertation, Martin-Luther-Universität Halle-Wittenberg.
- [17] Baumgartner, P. 2003. E-Learning: Lerntheorien und Lernwerkzeuge. In *Österreichische Zeitschrift für Berufsbildung (ÖZB)*, 3–6.
- [18] Rohs, M. 2007. *Zur Theorie formellen und informellen Lernens in der IT-Weiterbildung*. Dissertation, Universität der Bundeswehr Hamburg.

- [19] Walther, P. e. A. 2005. Standardisierung im E-Learning – Entwicklungsstand und Anwendungsbeispiel aus dem Agrarbereich. In *Zeitschrift für Agrarinformatik*. Landwirtschaftsverlag. Münster, 27–31.
- [20] Anderson, L. W. et al. 2009. *A taxonomy for learning, teaching, and assessing. A revision of Bloom's taxonomy of educational objectives*. Longman, New York.
- [21] Maderthaner, R. 2017. *Psychologie*. utb basics 2772. facultas; UTB, Wien, Stuttgart.



©2021 Manuel Haynes. Lizenznehmer Hochschule Reutlingen, Deutschland. Dieser Artikel ist ein Open-Access-Artikel unter den Bedingungen und Konditionen der Creative Commons Attribution (CC BY)-Lizenz. <http://creativecommons.org/licenses/by/4.0/>

# Requirements for an Interactive Marketing Dashboard

Jana Maria Ickenroth

Reutlingen University

Jana\_Maria.Ickenroth@Student.Reutlingen-University.de

## Abstract

The growing number of heterogeneous data sources and the ever increasing amount of available data makes it more and more difficult for marketers to efficiently analyze and target their marketing activities. Many platforms such as Google Analytics, LinkedIn, and Facebook offer their own analytics capabilities that measure the performance of activities on the respective channel. However, in order to evaluate a campaign holistically, the analytics data from the various sources must be combined. In this thesis, functional and non-functional requirements are identified for an interactive marketing dashboard that allows the central analysis of marketing data from different sources. The requirements elicitation is carried out by examining the state of the art as well as conducting and evaluating expert interviews.

## CCS Concepts

• **Information systems** → **Decision support systems**; • **Human-centered computing** → **Information visualization**;

## Keywords

Data-driven Marketing, Dashboard, Business Intelligence, Requirements Analysis, Scrum, Data Integration, Data Analysis

## 1 Introduction

In 2018, Germans over the age of 14 spent on average 196 minutes online per day. That is an increase of 240 % compared to just ten years earlier and more than eleven times the amount in 2000 [2]. Consequently,

public media consumption behavior and the appropriate customer approach have also changed. While TV and print advertising used to dominate, modern approaches focus on contacting customers via a wide variety of (online) channels [4]. Social networks such as Facebook, Instagram, and LinkedIn have their own analytics sections where performance metrics can be viewed. Additionally, there are web analytics tools that evaluate user behavior on websites and provide it in a third-party interface. But how can these key figures from different sources be combined to ultimately evaluate whether a marketing campaign was successful or not? Although Bosch already utilizes tools that can be used to centrally plan, control, and evaluate campaigns, reporting is still time-consuming. For this reason, an interactive marketing dashboard is to be designed and prototyped which combines key figures from various data sources such as social media channels, analytics services, and content management tools, and allows for central evaluation.

The overall goal of the work is to elicit functional and non-functional requirements

---

Supervisor	University	Prof. Dr. Uwe Kloos Hochschule Reutlingen Uwe.Kloos@Reutlingen- University.DE
Supervisor	Organization	Sebastian Weiss Bosch.IO GmbH Sebastian.Weiss@bosch.io

Informatics Inside  
24. November 2021 Hochschule Reutlingen  
Copyright 2021 Jana Maria Ickenroth

for an interactive marketing dashboard. The resulting requirements catalog serves as the basis for the subsequent creation of the dashboard, which is not part of this article.

In order to establish context, chapter 2 first provides relevant fundamentals on data-driven marketing, the starting point of this approach within many companies, and the use of business intelligence tools as a solutions approach. Then, in chapter 3, the state of the art of interactive dashboards is examined in order to gain insights and derive initial requirements. Chapter 4 explains the chosen methodology for conducting the requirements analysis which is followed by the requirements elicitation based on expert interviews, documentation, and analysis in chapter 5. The concluding chapter 6 summarizes the gained insights and includes an outlook on future work.

## 2 Background

The following chapter provides background information to ensure a better understanding in the course of the paper.

### 2.1 Data-driven Marketing

The additional communication and analytics channels described in chapter 1 represent new data sources that can be used in marketing to create value and serve as a basis for decision-making. In order to make a data-driven decision, it's crucial to extract the important information and combine relevant data sources in an intelligent way, because „data alone does not provide insights“ [4]. To move from Big Data to Smart Data, data analytics technologies and capabilities are needed to process not only formatted data but also semi-structured and unstructured data [4]. Effective analysis not only involves finding meaningful patterns in data, but also linking those insights to the business context [22]. In their "Maturity Model for Data Analytics", the market research company Gartner

shows the relationship between difficulty of the data analysis techniques to be used and the added value to be gained [8] (Figure 1). Accordingly, there are four types of analytics that organizations can apply: Descriptive analytics, diagnostic analytics, predictive analytics, and prescriptive analytics. Descriptive analytics are the most basic form and rely primarily on past-related structured data to uncover historical problems as well as potentials. „General examples [...] include sums, averages, and percentage changes“ [22]. This type of data analysis is often used in a corporate context and in connection with business intelligence systems. Diagnostic analytics go one step further and analyze the why. Causal analyses or structure-discovering procedures are used to clarify causes, effects, and interactions. The next step is predictive analytics where statistical methods or machine learning methods are used to make predictions about future events and its possible outcomes based on past data. This type of analytics is helpful for goal setting and planning [22]. The final step according to Gartner is prescriptive analytics which, for example, uses simulations and if-then analyses to show the effects of different courses of action and thus enable optimized decision-making [8][22][4].

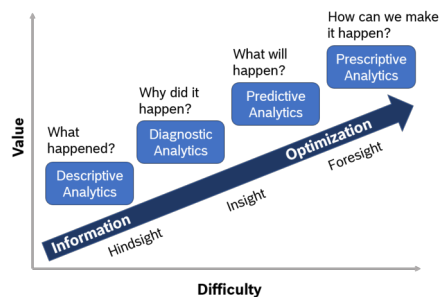


Figure 1: Maturity Model for Data Analytics [8]

## 2.2 Initial Situation in Organizations

Nearly all of the participants in the study "Future Ready" [9] conducted by Wunderman in 2018 agree that data is a crucial factor for marketing success. However, implementation is often still a problem. Although three quarters of companies have already deployed a marketing technology stack, only one third of them manage to design their marketing measures according to the data-based insights. The people interviewed see outdated structures within the companies as a major hurdle to overcome towards a data-driven approach.

This goes along with Roth's statement [1] that many companies are still not taking a holistic approach to performance management. If they have an integrated management reporting, the data and information often stems from completely different systems. Combined with different methodical approaches, Roth states that this can lead to delayed management decisions or even wrong decisions, because many queries and clarifications are necessary to obtain transparency. Another obstacle on the way to holistic performance management, explains Roth, is that many companies continue to use tools such as Excel and PowerPoint for reporting. This results in „unnecessary coordination and validation steps as well as manual adaptations of numbers [which] require valuable reaction time“ [1]. Also, numerous interfaces, manual process steps, and room for interpretation „imply risks for the security, reliability [and] compliance with deadlines and quality of information“ [1]. Instead of a bottom-up approach, Roth suggests a top-down view for target-oriented performance management, meaning that operative activities should be oriented towards the company's objectives and strategies [1].

## 2.3 Business Intelligence Tools

Business Intelligence (BI) tools such as Microsoft Power BI, Tableau, or Google Data Studio make it possible to connect heterogeneous data sources to gain new insights and thus provide users with a central overview of all relevant information. This enables targeted work, facilitates decision-making, and creates transparency [22][3]. These tools allow the creation of reports and dashboards in a quick way and are relatively easy to use. BI dashboards primarily consist of context-relevant metrics, Key-Performance-Indicators (KPIs), and textual information [24].

Nevertheless, BI solutions often suffer from low utilization rates, supported by the fact that „only 20% of users actually make use of [deployed] BI technologies“ [3]. Beelders et al. conducted an eye-tracking usability study [3] to determine possible causes of low usage rates. They found that the problem might be caused by the user interfaces: The participants sometimes struggled to relate legends to visualizations, which might be due to the high concentration of information displayed in the tested dashboards. In addition, fixation duration as well as fixation number increased with ongoing exposure, suggesting that the cognitive load of working with dashboards increased within the study. For these reasons, the authors recommend that the development of dashboards should be done in cooperation with end users, thereby increasing usage rates.

## 3 State of the Art

In this chapter, three dashboards from different sectors are presented to show the current state of the art. First, however, it should be clarified what is meant by the term dashboard in this context, as the definition is a matter of interpretation. As Sariaka et al.

state, „the dashboard concept has evolved from single-view reporting screens to include interactive interfaces with multiple views and purposes [...]“ [17]. This is especially due to the increasing amount of data as well as the availability of visualization techniques to everyone. This paper uses the definition by Wexler et al. [23] who describe dashboards as „visual display[s] of data used to monitor conditions and/or facilitate understanding“ [23]. Regarding interactivity, there are three forms according to [17]: 1) Options to allow users to (re-)design the user interface (e.g. by moving or adding/deleting elements), 2) filtering or slicing data, and 3) modifying the state of the data, e.g. by adding what-if-analyses.

### 3.1 Management Cockpit

The Management Cockpit at Reutlingen University (Figure 2) is a work environment for decision makers, in order to establish holistic performance management. Thus, it not only provides them with business-relevant key figures but also enables information analysis, visualization, decision-making and communication. Managers are presented with all available information to get an overview of the situation and possible outcomes and thus be able to make well-informed decisions [16][11].



**Figure 2: Management Cockpit at Reutlingen University [11]**

The Management Cockpit is used for various use cases: For instance, management simulation games as part of lectures, social media

monitoring, and real-time data analysis. Regarding structuring and categorization of information, Roth suggests to divide the cockpit into several walls, each of which fulfills its own functionality, e.g. one to present the most important KPIs, one to monitor competitors, and one to provide insights of internal processes and resources. [16]. According to Roth, this structure strengthens the recognition factor and helps management to keep a clear overview. Regarding visualization he says: „It is essential that management can capture the core message simply and quickly“ [16].

In order to achieve strategic objectives with operative activities, it's important to „identify the cause-and-effect relations between the different measures and objectives [...]“[16][14]. Thus, the cockpit allows managers to simulate different scenarios by using if-then analyses to weigh up different courses of action against each other. As a result, the Management Cockpit can be classified at the highest level in the maturity model for data analytics according to Gartner (Figure 1) [8].

### 3.2 COVID-19 Dashboard

In January 2020, the Center for Systems Science and Engineering (CSSE) at John Hopkins University first shared an online interactive dashboard that was able to visualize and track reported cases of COVID-19 in real time. Their goal was to make „accurate and robust epidemiological, clinical, and laboratory data [available] early in an epidemic“ [5]. In November 2020, TIME recognized the John Hopkins Coronavirus Resource Center as the „go-to data source“ for COVID-19 [21]. Data is collected from various sources<sup>1</sup>, such as the World Health Organization (WHO) or

<sup>1</sup>The aggregated datasets, as well as the original data sources, are available for public access at: <https://github.com/CSSEGISandData/COVID-19>



**Figure 3: COVID-19 Dashboard by CSSE [12]**

the European Centre for Disease Prevention and Control (ECDC). The dashboard (shown in Figure 3) is divided into four sections. The top section shows the main KPIs, namely the total number of reported cases, deaths, and vaccine doses administered. On the left side, there is a list of countries affected by the virus with respective reported cases. The panel on the right contains interactive timelines for daily cases, daily deaths, and total vaccine doses. In the center is a world map showing the number of cases, deaths, and vaccine doses per country. All panels, except the map, are interconnected, so if the user changes the timelines on the right, for example, the numbers on the left and at the top are updated accordingly. What was noticed during testing as well were long loading times. It is not clear whether this was due to the VPN connection used, a heavy server load, or other factors. Nevertheless, responsiveness is an important point to keep in mind when eliciting requirements for dashboard development.

### 3.3 Voder

Voder is a system developed by Srinivasan et al. in 2018 [20] that uses interactive data facts to help users interpret visualizations, explore the underlying data, and communicate their findings. Data facts are defined as „a textual description of the result of one or more statistical functions applied to the data used to create a visualization“ [20]. To generate these potentially useful insights, Voder

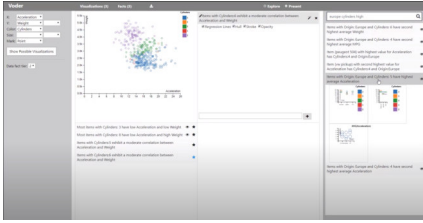
uses a collection of predefined heuristics. Depending on the attribute combination (numerical/categorical/combinations of both), the following tasks are applied: *Find anomalies*, *correlation*, *characterize distribution*, *find extremum*. Also, they cover *derived value* and *filter* tasks in conjunction with the tasks listed before. The generated data facts alongside visualizations allow people to verify assumptions they made about a chart, help them discover information they might have missed otherwise, and assist in interpreting complex visualizations. However, it can be challenging to mentally map the system-generated sentences to the accompanying visualization. Therefore, the authors specified that their system should visually link data facts to visualizations by using embellishments. To support data exploration, the system suggests related facts that are more detailed or more aggregated. Also, the user is supported by suggestions of alternative visualizations or embellishments to pair with a data fact. Another important aspect of Voder is communication: Visualizations and data facts can be saved and later organized in a slide-like layout or a dashboard layout. When finalized, all insights can be shared by exporting them as an interactive HTML page. A system video<sup>2</sup> shows Voder’s user interface and its capabilities in detail.

### 3.4 Insights from the State of the Art

The Management Cockpit has been developed since 2008 and therefore represents the current state of the art well. Thus, important insights can be drawn from this work: (1) The dashboard should provide a steering direction to align operational

<sup>2</sup>[https://drive.google.com/file/d/1HdHf3qH\\_jzhnP16sfa3kT8388PGTSv0/view](https://drive.google.com/file/d/1HdHf3qH_jzhnP16sfa3kT8388PGTSv0/view)





**Figure 4: Voder User Interface (Screenshot from System Video – Link in Footnote)**

marketing activities with the overall strategic goal. (2) Structuring the information according to internal corporate structures and content-related contexts leads to better orientation and supports comprehension. (3) Data and information have to be visualized in a way that is quick and easy to understand.

Regarding the COVID-19 dashboard, the following findings emerge for the present work: (1) Cross-visual filtering helps understanding relationships between data. (2) The loading times of the dashboard should be as short as possible to facilitate work and not cause frustration to the user.

Last, Voder [20] is a good example of interactive dashboards to support data exploration. Also, the paper contains extensive and very insightful dashboard development information. As a result, some valuable conclusions can be drawn from this: (1) Textual descriptions alongside visualizations facilitate interpretation, help discover information that might have been missed otherwise, and enable easier communication. (2) Data facts should be visually linked to the corresponding charts. (3) The dashboard should suggest alternative visualization possibilities. (4) Related visualizations and data facts from different abstraction levels help data exploration.

## 4 Methodical Approach

The following chapter explains how the requirements analysis is carried out methodically, how the interview guide is designed and evaluated, and how the requirements documentation is done.

### 4.1 Requirements Analysis

The overall process is carried out according to Ebert [7] and is divided into ten phases, eight of which are relevant in the present project: (1) Clarify relevant claimants (stakeholders, target group), (2) specify goals and product vision, (3) use different investigation techniques, (4) document and structure requirements, (5) analyze and review requirements, (6) solve conflicts, (7) prioritize requirements, and (8) decide on requirements. Since the dashboard deliverable is based on a project with a lot of ambiguity in the beginning, it would be risky to define all requirements and specifications in too much detail. For this reason, Scrum was used for the individual process steps, as it is a lightweight framework that takes these initial uncertainties into account. Furthermore, Scrum has the great advantage that after each sprint a project status is reached, which can be demonstrated to customers in order to obtain constant feedback [6].

### 4.2 Interview Guide

The interview guide was designed based on [10]. Accordingly, a four-part structure is implemented, consisting of an information phase, a warm-up, a main phase, and a concluding phase. It is important to establish a pleasant atmosphere in which the interviewee can talk freely [10]. Furthermore, in the context under consideration, it was important not to formulate hypotheses in advance, but to understand the initial situation, including positive aspects and pain points, through open questioning. The questions for the interview guide were first derived based on the

studied literature. Subsequently, several revisions were made in consultation with a user researcher as well as with the technical supervisor. The actual structure of the interview guide and thus the interviews themselves can be obtained from section 5.

### 4.3 Interview Evaluation

The evaluation of the expert interviews corresponds to the methodology of [13]. First, a selective transcript or paraphrase is made. A paraphrase is a condensation of the statements, whereby the content of the message remains unchanged. The second step of data preparation is to categorize the statements made (Figure 5). In the case of a guided interview, it is advisable to use the topics of the guide as a basis.

Case	Sub-Topic (T 1-5)						
C 1	T2	T5	T2	T1	T3	T4	T3
C 2	T1	T4	T2	T5	T4	T3	T2
C 3	T5	T1	T5	T1	T4	T3	T2
C 4	T5	T2	T1	T3	T5	T4	T2

Figure 5: Categorization of Statements (According to [13])

The final step is to remove the sequentiality of the individual interviews – instead, the statements are clustered by topic (Figure 6). This also shows which topics include the most statements. Subsequently, this can be taken into account during the prioritization process.

Case	Sub-Topic (T 1-5)				
C 1	T1	T2	T3	T4	T5
C 2	T1	T2	T3	T4	T5
C 3	T1	T2	T3	T4	T5
C 4	T1	T2	T3	T4	T5

Figure 6: Clustering of Statements (According to [13])

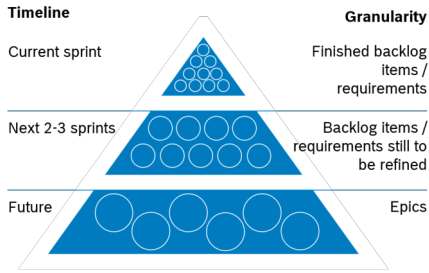
## 4.4 Requirements Documentation

Requirements are divided into functional and non-functional requirements and documented as user stories. These are concise descriptions of product features that provide added value for a user or buyer of the product. They follow a fixed template: As a <user role>, I want <functionality> so that <benefit>. In order to determine when a requirement is considered fulfilled, acceptance criteria are formulated for each user story. These are written in the form of a checklist. In Scrum, requirements should fulfill the INVEST criteria, meaning they should be Independent, Negotiable, Valuable, Estimatable, Sized appropriately, and Testable. The granularity with which requirements are described is primarily based on when they will be implemented. With ongoing iterations, the requirements are described more precisely [6]<sup>3</sup>. This is illustrated in Figure 7. Nevertheless, it became clear during the requirements analysis described in chapter 5 that there are also requirements that can be defined very precisely from the start, such as filters and their options, although these may not yet need to be implemented in the first sprints. To determine an initial order of the requirements, the MuSCoW prioritization is used, which stands for "Must have, Should have, Could have, Won't have" [15].

## 5 Requirements for the Marketing Dashboard

The following chapter contains the target group, the template that was used as a base to formulate the product vision, and requirements that were elicited from expert interviews and literature research.

<sup>3</sup>Note: Within Scrum, epics are usually higher-level software packages to which user stories can be assigned. In the book by Dräther et al., however, epics are seen as very rough requirements that function more as reminders for team members.



**Figure 7: Requirements of Different Granularity Depending on Timeline (Own Figure According To [6])**

## 5.1 Target Group and Product Vision

The target group for the dashboard are marketing associates, who have to manage multiple tasks at once and do not have enough time for a regular and holistic marketing reporting. As [7] states, the product vision describes the purpose of a product and functions as the guiding principle of the project. It goes beyond the individual requirements and thus helps to decide which requirements fit together and are prioritized. Furthermore, a product vision should help in motivating the team, the stakeholders, and users. The product vision for the marketing dashboard was created by following the vision statement template proposed by Schuurman [18]: For *target customer* who *statement of need or opportunity* the *product name* is a *product category* that *key benefit, reason to buy*. Unlike *primary competitive alternative* our *product statement of primary differentiation*. Schuurman also advises to share the product vision often and adapt it to changes in user or market needs, if necessary.

## 5.2 Expert Discussions

The goal of the expert discussions was to understand the work environments, experiences, knowledge level of the subjects as well as positive aspects and pain points in the current process. The following chapter describes the process starting from the design of the interview guide to deriving the requirements.

**5.2.1 Interview Structure and Guide** The resulting interview guide and thus the interviews were divided into four parts:

- (1) The introduction aimed at opening the conversation and making the participant feel comfortable. Here, the project was introduced, the interview process was explained and details were given on how the information collected was further processed.
- (2) The second part of the interview was considered the warm-up phase: the goal here was to get to know the interviewees by asking questions about their job profile and work environment. In addition, the transition to the main part of the interview was created by asking a number of questions about their experience with analytics.
- (3) In the main phase of the interview, initial questions were asked about the general workflows in content marketing. This included questions about the participants tasks, how content is planned, whether and how the performance of campaign activities from different channels is compared, what exemplary campaign goals look like, and how it is monitored whether set goals are achieved. After that, questions about the tools used in marketing activities were asked to identify possible interfaces that would need to be incorporated in the dashboard. This was followed by a change in interview technique: In order to understand how reporting is done across all tools to ultimately assess how a campaign has

performed, a user journey map was filled in together with the interviewees. It consisted of seven levels: (i) The whole process on a meta level, (ii) the exact activities per process step, (iii) person responsible for each process step, (iv) time required for each process step, (v) suitable smiley for each process step (happy, neutral, sad), (vi) what works well?, (vii) what doesn't work well? After the user journey map, the interviewees were asked additional questions about the reporting to see which analyses were especially interesting to them and how they decided between relevant and irrelevant data. The next part of the interview involved questions about data visualization and collaboration to see which charts are often used, who the reports are shared with, and which abstraction level is necessary. To finish the main part of the interview, the interviewees were asked which data sources they would like to combine and why.

(4) In the closing phase of the interview, possible open questions were clarified, feedback was requested and the interview was concluded.

**5.2.2 Conduction of the Expert Discussions**  
A total of four interviews with five individuals were conducted. Three people were marketing employees from different Bosch business units and two people had a digital business consulting role within Bosch.IO. This leads to two different perspectives on the topic: While the consultants have more of an overarching view and are familiar with the problems of many different customers, the marketing employees reported on their very specific activities and problems. Both perspectives were valuable. Each interview was scheduled for 90 minutes and was conducted via Microsoft Teams. The user journey map was prepared in Microsoft Whiteboard. Also, all interviews were recorded in order to enable an easier follow-up. The evaluation was done as described in chapter 4.2.

### 5.3 Derivation of Requirements

A total of 18 functional and 6 non-functional requirements were identified from the conducted interviews and the examined literature. The initial prioritization resulted in the following distribution: Eight functional and five non-functional requirements rated as "Must", two functional and one non-functional requirements prioritized as "Should", and eight functional requirements prioritized as "Could". Since the implementation of the dashboard will not take place until the subsequent master thesis, and because the requirements should be applicable to different settings also outside of Bosch, the granularity of some of the requirements regarding larger features have been kept rather general at this point and can be compared with the backlog items from Figure 7. However, to illustrate what refined requirements may look like, FReq1.1 is added as an example.

### 5.4 Functional Requirements

A functional requirement describes a function that is to be provided by the system or a system component. Furthermore, they describe system or function behaviors that are to be triggered when certain conditions are met [7]. In the following section, five functional requirements are shown, which were evaluated with the priority "Must" within the scope of internal company processes. The associated acceptance criteria are shown as an example in FReq1 and FReq1.1. For the other requirements, only the user stories are listed, but all acceptance criteria are documented internally.

**FReq1: Management summary – Metrics**

*User Story:* As a user, I want to have a

management summary with the three most important KPIs of my use case to get a quick overview of the current metrics.

*Acceptance Criteria:*

- Each metric receives one column in the management summary.
- Each metric is shown as a number and as a chart.
- All three columns have the same width and height.
- All information is displayed within the desktop viewport, without the users having to scroll down.

When referring to the performance monitoring of social media campaign activities, the requirement can be refined:

#### **FReq1.1: Campaign management summary – Metrics**

*User Story:* As a campaign manager, I want to have a management summary that shows the current performance metrics of my campaigns to evaluate my activities quickly and enable fast reporting.

*Acceptance Criteria:*

- There are KPI columns for the number of conversions, the average conversion rate across social platforms, and the average engagement rate across social platforms.
- The respective charts show the number of conversions per platform, the conversion rate per platform, and the engagement rate per platform.
- All three columns have the same width and height.
- All information is displayed within the desktop viewport, without the users having to scroll down.

#### **FReq2: Time frame filters**

*User Story:* As a user, I want to be able to filter by specific time frames to analyze data from a certain time period.

#### **FReq3: Relative performance changes**

*User Story:* As a user, I want to see the percentage change of the current period compared to the previous period to know if my performance is increasing or decreasing over time.

#### **FReq4: Integrating marketing data from different data sources**

*User Story:* As a user, I want content marketing data, web analytics data, and social media analytics data available in one central dashboard so I don't have to switch between different tools to get a holistic overview of my marketing activities' performance.

#### **FReq5: Cross visual filtering**

*User Story:* As a user, I want the areas in different charts referring to the same data to be connected visually in order to understand relationships between data better.

## **5.5 Non-Functional Requirements**

Non-functional requirements, also called quality requirements, describe qualitative properties that the system or system components must fulfill. Examples are reliability, availability, modifiability, safety, security and also usability [7]. The non-functional requirements rated as "Must" are listed below:

#### **NFReq1: User rights for dashboard modifications**

*User Story:* As a user, I want to have sufficient user rights to make modifications to my dashboards to adopt it to my own needs.

#### **NFReq2: Easy-to-understand reports**

*User Story:* As a user, I want to be able to draw conclusions from the displayed data confidently to make data-driven decisions.

*Acceptance Criteria:*

- In a usability test, an error-free rate (critical and non-critical errors) of 80% is the goal for each task.

### **NFReq3: Data access limited to Business Unit**

*User Story:* As an admin, I need to make sure that customers only have access to the data belonging to their Business Unit to keep data privacy.

### **NFReq4: Loading times**

*User Story:* As a user, I want the dashboard to load fast, to ensure a good user experience and not cause frustration when interacting with the dashboard.

*Acceptance Criteria:*

- The maximum initial loading time of the dashboard is 10 seconds.
- The maximum loading time after changed filter settings is 5 seconds.

### **NFReq5: Intuitive user interface**

*User Story:* As a user, I want the navigation, filter options, and interactive elements to be intuitive enough to use with only short initial training.

*Acceptance Criteria:*

- In a usability test, the initial training does not take more than 10 minutes.
- In a usability test, the completion rate for each task is 95%.

## **5.6 Reflection on Derived Requirements**

In 1996, Ben Shneiderman stated that „exploring information collection becomes increasingly difficult as the volume grows. A page of information is easy to explore, but when the information becomes the size of a book [...], it may be difficult [...] to gain an overview“ [19]. Although this was 25 years ago, it was still reflected in the user interviews today.

The solution Shneiderman suggests as a starting point is called the Information Seeking Mantra: *Overview first, zoom and filter, then details-on-demand*. This can also be recognized in FReq1 regarding the management summary. Users would find it helpful to have a self-updating overview with the most important KPIs from their use cases which they can access at any time to get a quick overview of their current online performance. Also, several filter options were named during the interviews, allowing for a drill-down of the data. Another statement that was made multiple times during the expert interviews was that „less is more“. This was also confirmed by Prof. Dr. Armin Roth – creator of the management cockpit – who explained in an expert interview that a dashboard that is too complex will not be used, also pointing in the direction that details should only be displayed on demand.

## **6 Conclusion**

Nowadays, the challenge is not gathering data but to take action based on it – We need to transform Big Data into Smart Data so that we can learn from it. Although many inspiring data dashboards already exist and most people understand the importance of data-driven decisions, many companies still struggle to implement a solution that fits their needs. This is often due to a lack of time and/or a rather high-level understanding of data analysis. Within this paper, the state of the art was examined to derive insights from existing interactive dashboard. Also, expert interviews assessed the wishes and needs regarding a holistic and automated marketing reporting solution. The derived requirements can be applied to different use cases and function as a starting point to create an interactive dashboard for marketing associates.

## References

- [1] A. Roth. “Conception of a holistic and long-term performance management: five part disciplines as systematizing framework”. In: *Quarterly Review of Business Disciplines* (2014). URL: <https://publikationen.reutlingen-university.de/frontdoor/deliver/index/docId/104/file/104.pdf>.
- [2] ARD and ZDF. *Entwicklung der durchschnittlichen täglichen Nutzungsdauer des Internets in Deutschland in den Jahren 2000 bis 2018 (in Minuten)*. Last accessed on 10.11.2021. URL: <https://de.statista.com/statistik/daten/studie/1388/umfrage/taegliche-nutzung-des-inter%5C-nets-in-minuten/>.
- [3] Tanya Beelders and Eduan Kotzé. “Augmenting the Business Intelligence Lifecycle Model with Usability: Using eye-Tracking to Discover the why of Usability Problems”. In: 23 (2020), pp. 96–111. DOI: 10.34190/EJISE.20.23.1.007.
- [4] Silvia Boßow-Thies, Christina Hofmann-Stölting, and Heike Jochims. *Data-driven Marketing Insights aus Wissenschaft und Praxis: Insights aus Wissenschaft und Praxis*. Jan. 2020. ISBN: 978-3-658-29994-1. DOI: 10.1007/978-3-658-29995-8.
- [5] Ensheng Dong, Hongru Du, and Lauren Gardner. “An interactive web-based dashboard to track COVID-19 in real time”. In: *The Lancet Infectious Diseases* 20 (2020). DOI: 10.1016/S1473-3099(20)30120-1.
- [6] Rolf Dräther, Holger Koschek, and Carsten Sahling. *Scrum-kurz & gut*. O’Reilly, 2019. ISBN: 9783960090946.
- [7] Christof Ebert. *Systematisches Requirements Engineering: Anforderungen ermitteln, dokumentieren, analysieren und verwalten*. dpunkt. verlag, 2019. ISBN: 978-3-86490-562-9.
- [8] Gartner. *Gartner’s Data Analytic’s Maturity Model*. Last accessed on 10.11.2021. URL: <https://towardsdatascience.com/how-analy%5C-tics-maturity-models-are-stunting-data-science-teams-962e3c62d749>.
- [9] David Hein. *Wunderman-Studie: Zwei Drittel aller Unternehmen können aus Daten keine Marketing-Maßnahmen ableiten*. Last accessed on 10.11.2021. URL: <https://www.horizont.net/marketing/nach%5C-richten/Wunderman-Studie-Zwei-Drittel-aller-Unternehmen-koennen-aus-Daten-keine-Marketing-Mass%5C-nahmen-ableiten-165526>.
- [10] Cornelia Helfferich. “Qualitative Interviews”. German. In: *Soziologische Revue* 41.3 (2018), pp. 508–512. URL: <https://search.proquest.com/docview/2493133943>.
- [11] Hochschule Reutlingen. *Management Cockpit*. Last accessed on 10.11.2021. n.d. URL: <https://mc.reutlingen-university.de/>.
- [12] John Hopkins University & Medicine. *COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE)*. Last accessed on 10.11.2021. n.d. URL: <https://coronavirus.jhu.edu/map.html>.
- [13] Renate Liebold and Rainer Trinczek. “Experteninterview”. In: *Handbuch Methoden der Organisationsforschung* (2009), pp. 32–56. DOI: 10.1007/978-3-531-91570-8\_3.
- [14] Koen Pauwels et al. “Dashboards as a Service : Why, What, How, and What Research Is Needed?” In: *Journal of Service Research* 12 (2009), pp. 175–189. DOI: 10.1177/1094670509344213.
- [15] Roman Pichler. *Scrum: agiles Projektmanagement erfolgreich einsetzen*. dpunkt. verlag, 2013. ISBN: 9783898644785.

- [16] Armin Roth. "Management Cockpit as a Layer of Integration for a Holistic Performance Management". In: *Quarterly Review of Business Disciplines* 2.2 (2015), pp. 165–175. issn: 2329-5163. URL: <https://faculty.utrgv.edu/louis.falk/qrbd/QRBDaug15.pdf#page=91>.
- [17] Alper Sarikaya et al. "What Do We Talk About When We Talk About Dashboards?" In: *IEEE Transactions on Visualization and Computer Graphics* 25.1 (2019), pp. 682–692. DOI: 10.1109/TVCG.2018.2864903.
- [18] Robbin Schuurman. *10 Tips for Product Owners on the Product Vision*. Last accessed on 10.11.2021. 29.11. 2017. URL: <https://www.scrum.org/resources/blog/10-tips-product-owners-product-vision>.
- [19] B. Shneiderman. "The eyes have it: a task by data type taxonomy for information visualizations". In: *Proceedings 1996 IEEE Symposium on Visual Languages*. 1996, pp. 336–343. DOI: 10.1109/VL.1996.545307.
- [20] Arjun Srinivasan et al. "Augmenting Visualizations with Interactive Data Facts to Facilitate Interpretation and Communication". English. In: *IEEE transactions on visualization and computer graphics* 25.1 (2019), pp. 672–681.
- [21] TIME. *2020's Go-To Data Source: John Hopkins Coronavirus Resource Center*. Last accessed on 10.11.2021. Nov. 2020. URL: <https://time.com/collection/best-inventions-2020/5911434/johnsh-hopkins-coronavirus-resource-center/>.
- [22] Prakash Ukhalkar et al. "Business Intelligence and Analytics: Challenges and Opportunities". In: *International Journal of Advanced Science and Technology* 29.12s (2020), pp. 2669–2676.
- [23] Steve Wexler, Jeffrey Shaffer, and Andy Cotgreave. *The big book of dashboards: visualizing your data using real-world business scenarios*. John Wiley & Sons, 2017. ISBN: 978-1-119-28271-6.
- [24] Jack Zheng. "Data Visualization for Business Intelligence". In: 2017, pp. 67–82. ISBN: 1138203688. DOI: 10.4324/9781315471136-6.





# Dynamisches Matching für Peer to Peer Mitfahrgelegenheitssysteme

Marius Matthäus Bauer

Hochschule Reutlingen

marius\_matthaeus.bauer@student.reutlingen-university.de

## Abstract

**[Kontext und Motivation]** Die zunehmende Sorge um die Umwelt, eine immer größer werdende Anzahl an Kraftfahrzeugen (KFZ) und die damit einhergehende Überlastung der Straßen sorgt für einen Bedarf an alternativen Transportmöglichkeiten. Peer-to-peer (P2P) Mitfahrgelegenheiten bieten eine solche Alternative. Sie beschreiben Fahrgemeinschaften zwischen Privatpersonen und stellen eine kostengünstige und komfortable Transportmöglichkeit dar. Studenten der Hochschule Reutlingen arbeiten an einem solchen P2P Mitfahrgelegenheitssystem, namens *URide*, das die Route eines KFZ automatisch aufzeichnet und so automatisch Mitfahrgelegenheiten für die Teilnehmer des Systems vorschlagen kann. Der Kern eines solchen Mitfahrgelegenheitssystem besteht daraus, Nutzern des Systems potentielle Mitfahrgelegenheiten, die ähnliche Reiserouten besitzen, vorzuschlagen. Um dieses Optimierungsproblem zu lösen wurden in der Vergangenheit viele verschiedene Ansätze und Algorithmen veröffentlicht. **[Fragestellung und Methodik]** In dieser Arbeit wird daher, mithilfe einer systematischen Literaturrecherche, untersucht, welcher Ansatz für den Einsatz in *URide* geeignet ist. Hierfür werden Arbeiten, die in den letzten 10 Jahren veröffentlicht wurden, betrachtet. Um einen möglichst breiten Stand der Technik abzubilden wurden die drei größten Datenbanken IEEE, ScienceDirect und ACM verwendet. **[Ergebnisse]** Es stellt sich heraus, dass die untersuchten Arbeiten nicht quantitativ vergleichbar sind da diese keinen einheitlichen und direkt vergleichbaren Test durchführen.

Nichtsdestotrotz wurden die untersuchten Arbeiten in Hinblick auf Komplexität und Dynamik eingeordnet und für eine erste Entscheidungsgrundlage gegenübergestellt. Der Ansatz von Schreieck et al. [1] erwies sich am passendsten für den gegebenen Anwendungsfall.

## CCS Concepts

• **Applied computing** → *Transportation*;

## Keywords

Dynamic ridesharing, matching algorithms, ride matching, optimization algorithms, smart transportation system

## 1 Einführung

Durch die immer größer werdende Anzahl von Kraftfahrzeugen (KFZ) werden Verkehrsnetze, vor allem in Ballungsräumen wie bspw. Großstädte, immer häufiger durch Verkehrsstaus blockiert. Diese Verkehrsstaus wirken sich nicht nur direkt auf die Reisezeit und den Kraftstoffverbrauch aus, sondern verursachen auch weitere indirekte Kosten wie bspw. erhöhte Emissionen, die sich nachteilig auf die menschliche Gesundheit und den Klimawandel auswirken können [2]. Ein Ansatz um diesem Problem entgegen zu wirken

---

Betreuer/-in Hochschule

Prof. Dr.-Ing. Marcus Schöller  
Hochschule Reutlingen  
marcus.schoeller@reutlingen-university.de

Informatics Inside Herbst 2021  
24. November 2021 Hochschule Reutlingen  
Copyright 2021 Marius Matthäus Bauer

sind Mitfahrgelegenheiten oder Fahrgemeinschaften. Diese beschreiben eine Transportart, bei der sich einzelne Reisende ein Fahrzeug für eine Reise teilen und Reisekosten wie bspw. Benzin, Maut und Parkgebühren unter den Mitfahrenden aufgeteilt werden [3]. Sie unterscheiden sich im Vergleich zu den öffentlichen Verkehrsmitteln darin, dass sie keine festen Routen und Zeitpläne besitzen und können daher wesentlich flexibler eingesetzt werden. Inzwischen haben sich bereits verschiedene Arten von Mitfahrgelegenheitssystemen etabliert.

*Peer-to-Peer (P2P)* Mitfahrgelegenheitssysteme ermöglichen es bspw. Nutzern des Systems andere Nutzer vorzuschlagen, die eine ähnliche Route, sowie Zeitplan besitzen. Beispiele solcher Systeme sind *Zimride*<sup>1</sup>, welches Mitfahrgelegenheiten für Firmen und Universitäten bereitstellt und *BlaBlaCar*<sup>2</sup>, das sich auf die gemeinsame Nutzung von Langstreckenfahrten zwischen größeren Städten konzentriert. Im Gegensatz zu Ridehailingsystemen haben die Fahrer bei P2P Mitfahrgelegenheiten ihre eigenen Ziele und fahren nicht nur, um die Anfragen der Fahrer zu bewältigen [4, 5].

Bei Ridehailingsystemen hingegen sind Fahrer, ähnlich wie bei Taxiunternehmen angestellt, um Nutzer der Anwendung an Ihren Zielort zu bringen. Diese Systeme sind auch unter dem Begriff *Ride-Sourcing* bekannt und werden oftmals in Form von Mobilanwendungen angeboten. Beispiele für solche Mobilanwendungen sind bspw. Uber<sup>3</sup> oder Lyft<sup>4</sup> [5]. Ridehailing-Mitfahrgelegenheiten bergen allerdings den Nachteil, dass weitere Leerkilometer dem Verkehrsnetz hinzugefügt werden können. Fahrer müssen bspw. alleine von

einem Reiseziel zum nächsten Kunden fahren oder fahren umher während sie auf die Zuordnung eines neuen Kunden warten [5]. Außerdem unterscheiden sich P2P Mitfahrgelegenheiten zu gewöhnlichen Fahrgemeinschaften insofern, dass keine langfristige Vereinbarungen und Wiederholungen nötig sind [4].

## 1.1 Kontext

In dieser Arbeit soll ein passender Matching-Algorithmus für ein P2P Mitfahrgelegenheitssystem namens *URide*, das an der Hochschule Reutlingen entwickelt wird, gefunden werden. *URide* soll unter anderem Autofahrer, die ähnliche Routen, sowie Zeitpläne besitzen, anhand von GPS Daten, erkennen. So könnten Autofahrern automatisch vorgeschlagen werden, dass eine Mitfahrergemeinschaft mit einer oder mehreren Personen möglich ist. Hierfür müssen GPS Daten der Fahrer entsprechend verglichen und zugeordnet werden. Das Ziel dieser Arbeit ist es, herauszufinden welche State of the Art Ansätze es gibt und welche am besten für den beschriebenen Anwendungsfall geeignet sind.

## 1.2 Problembeschreibung

Das P2P-Ride-Matching Problem ist eine Generalisierung des *Dial-a-ride Problem* (DARP). DARP wurde ursprünglich entworfen um bspw. medizinische Pflegedienste, bei denen Kunden an der Haustür abgeholt werden und zu einem Ziel gebracht werden, zu modellieren [2, 4]. Das System um das es in dieser Arbeit geht besteht aus einer Menge von Teilnehmern  $T$ . Ein Teilnehmer  $t \in T$  teilt mit dem System automatisch, mithilfe eines GPS und internetfähigen Mobilgerät, Herkunft und Ziel von gefahrenen Routen  $r \in R$ , bezeichnet als  $h_{t,r}$  und  $z_{t,r}$ . Außerdem wird das Zeitfenster der Route, bezeichnet als  $f_{t,r} = [z_{t,r}^{start}, z_{t,r}^{ende}]$  automatisch erkannt. Das System versucht dann, übereinstimmende Paare von Routen, die sowohl ähnliche Herkunft und Zielpunkte als auch Zeitfenster haben, zwischen Teilnehmern des Systems zu finden

<sup>1</sup><https://zimride.com/>

<sup>2</sup><https://www.blablacar.de/>

<sup>3</sup><https://www.uber.com/>

<sup>4</sup><https://www.lyft.com/>

und versucht so Mitfahrmöglichkeiten den Teilnehmern vorzuschlagen.

## 2 Methodik

Um den aktuellen Stand der Technik zu ermitteln wurde eine systematische Literaturrecherche durchgeführt. Hierbei wurden die drei Literaturdatenbanken IEEE<sup>5</sup>, ScienceDirect<sup>6</sup> und ACM<sup>7</sup> verwendet. Um die relevante Literatur zu identifizieren sind zuerst Ein- und Ausschlusskriterien sowie Schlüsselwörter definiert worden.

Die Schlüsselwörter wurden wie folgt definiert:

- Ridesharing AND (Matching OR Matching Algorithm)
- Ridepooling AND (Matching OR Matching Algorithm)
- Peer-to-peer AND (Ridesharing OR Ridepooling)

Es wurden lediglich Arbeiten betrachtet die folgende Einschlusskriterien erfüllen:

- Veröffentlichungsdatum in den letzten 10 Jahren (2011 - 2021)
- Peer-to-Peer Mitfahrgelegenheitssysteme
- Matching Algorithmen
- Reviews

Folgender Typ von Arbeiten wurden nicht betrachtet und somit als Ausschlusskriterium definiert:

- Studien
- Ridehailing Systeme und Algorithmen
- White Paper

Um weitere relevante Literatur zu identifizieren, wurde anschließend eine *Backward Search*, anhand der bereits gefundenen Arbeiten, durchgeführt. Die gefundenen Arbeiten wurden zuerst anhand des Abstracts auf die

Eignung für diese Arbeit überprüft. Im zweiten Schritt wurden die Arbeiten in eine Vergleichstabelle (Tabelle 1) eingesetzt und versucht, so detailliert wie möglich, die Arbeit zu verstehen. Im Rahmen dieses Prozesses wurden Eigenschaften wie Typ, Jahr, Matchingansatz, Komplexität, Fokus und Dynamik der Arbeiten in die Tabelle eingetragen. Die Komplexität der Arbeit wurde anhand der subjektiven Wahrnehmung des Aufwands um den Ansatz zu verstehen bestimmt. Da der Matching-Algorithmus in Zukunft von Studenten mit ähnlichem Hintergrund und Wissensstand implementiert wird, ist dies eine valide Methodik um die Komplexität und somit den Aufwand der späteren Implementierung zu bestimmen.

## 3 Kategorisierung von Mitfahrgelegenheitssystemen

Matching Algorithmen für Mitfahrgelegenheitssysteme lassen sich nach Tafreshian et al. [4] anhand der Anzahl an Stops für das Ein- und Aussteigen von Mitfahrern und der Anzahl an Umstiegen zwischen Fahrzeugen in drei verschiedene Kategorien einteilen: *one-to-one*, *one-to-many*, and *many-to-many*. Das *One-to-One* Matching beschreibt eine Mitfahrgelegenheit bei der ein Mitfahrer, oder mehrere, sofern diese gleiche Herkunft und Ziel teilen, von einem Fahrer mitgenommen werden.

*One-to-Many* Mitfahrgelegenheiten ermöglichen hingegen die Mitnahme von mehreren Personen die auch unterschiedliche Startpositionen und Ziele haben können. Ein Transfer der Mitfahrer in eine weitere Mitfahrgelegenheit ist bei beiden Formen nicht erlaubt.

Das *Many-to-Many* Matching bietet abschließend den größten Grad an Flexibilität, da mehrere Mitfahrer aufgenommen werden können und diese auch zwischen Fahrzeugen von

<sup>5</sup><https://ieeexplore.ieee.org/Xplore/home.jsp>

<sup>6</sup><https://www.sciencedirect.com/>

<sup>7</sup><https://dl.acm.org/>

verschiedenen Mitfahrgelegenheiten umsteigen können. Diese Art der Mitfahrgelegenheit könnte in einem multimodalen Transportsystem verwendet werden, das bspw. öffentliche Verkehrsmittel oder Taxis integriert. Der Nachteil dieses Matching-Typs ist allerdings eine hohe Berechnungskomplexität.

## 4 Ansätze

In diesem Kapitel werden die verschiedenen Ansätze der untersuchten Literatur aufgeführt. Gegliedert werden die Arbeiten in die in Kapitel 3 vorgestellte Kategorisierung von Mitfahrgelegenheitssystemen.

### 4.1 One-to-One

Der wohl primitivste Ansatz des *One-to-One* Matching besteht daraus, die Distanz zwischen korrespondierenden Start- und Zielpunkten zweier Teilnehmer zu vergleichen. Diese Aufgabe lässt sich für alle Teilnehmer des Systems, laut Masoud et al. [2], in Polynomialzeit berechnen. Allerdings birgt dieser Ansatz das Problem, dass Mitfahrer, deren Start- und Endziel entlang der vom Fahrer angegebenen Route liegen, nicht gematched werden [2]. Pelzer et al. [6] verfolgen unter anderem aus diesem Grund einen Zerlegungsansatz bei dem das Straßennetz in Partitionen aufgeteilt wird. Die Route eines Fahrers kann so als eine Reihenfolge von Partitionen dargestellt werden. Der Matching-Algorithmus kann dadurch auch Übereinstimmungen von Start- und Zielorten finden, die entlang der Route des Fahrers liegen. Die Hauptverkehrsstraßen werden hierbei von einer Art Kontur umhüllt. Straßen die sich überlappen oder kreuzen werden zu einer Kontur vereinigt. Flächen die von diesen Konturen eingeschlossen werden, bilden, wie in Abb. 1 zu sehen, die Partitionen. Die Breite der Kontur wird durch einen Parameter  $\alpha$  festgelegt. Die Partitionen können außerdem durch einen Parameter  $\beta$  skaliert werden. Sobald  $\beta > \alpha$  gewählt wird, überlappen sich angrenzende Partitionen. Das hat zur Folge, dass Reisende die sich in diesem Grenzgebiet



**Abbildung 1:** Kontur die die Hauptverkehrsstraße in Pelzer et al. [6] umhüllt. Die Fläche, die durch die Kontur eingeschlossen wird, repräsentiert eine Partition.

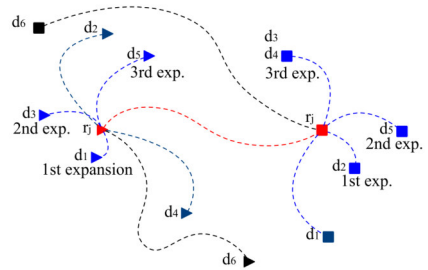
aufhalten zu Reisenden beider Partitionen gematched werden können. Dieser Ansatz des Matchings ermöglicht eine Reduzierung des Suchraums und kann so für Systeme mit einer großen Anzahl an Teilnehmern verwendet werden. Dadurch benötigt dieser Algorithmus im Vergleich zu einer Greedy Heuristik 30% weniger Zeit [6].

Schreieck et al. [1] verwenden hingegen einen Ansatz, bei dem die schnellste Route einer angebotenen Mitfahrgelegenheit vom Dijkstra Algorithmus berechnet wird. Die berechnete Route wird dann als eine Menge von geographischen Knoten in einer Datenstruktur mit invertiertem Index gespeichert. Solche Datenstrukturen werden unter anderem von Suchmaschinen eingesetzt um Suchwörter mit Dokumenten zu verknüpfen und können so Suchanfragen beschleunigen. In diesem Fall werden sie eingesetzt um von einzelnen Knoten auf die berechneten Routen zu schließen. Die Datenstruktur speichert also alle dem System bekannten Knoten und deren zugehörige Mitfahrangebote die diesen Knoten in ihrer Route haben. Der eigentliche Matching-Prozess besteht daraus, naheliegende Knoten, in einem Radius  $r$ , von

Start- und Zielort eines Teilnehmers der eine Mitfahrgelegenheit sucht (Mitfahrer) zu finden. Wenn Knoten einer Mitfahrgelegenheit sowohl im Radius des Start- als auch des Zielorts gefunden wurden, kann diese Mitfahrgelegenheit von diesem Mitfahrer verwendet werden. Sofern mehrere Routen von potentiellen Mitfahrgelegenheiten durch die beiden Radien führen, wird für jede Route der nächste Knoten zum Start- bzw. Zielpunkt berechnet. Anschließend wird der Durchschnitt zwischen der Distanz des nächsten Knotenpunkts zur Startposition und Zielposition berechnet. So kann festgestellt werden, welche Route einen kleineren Umweg für den Fahrer bedeutet. Diese Mitfahrgelegenheit wird dem Mitfahrer dann vorgeschlagen.

Eine weitere Arbeit von Ta et al. [7] versucht, ähnlich wie die Arbeit zuvor, den Prozentsatz der gemeinsamen Strecke von Fahrer und Mitfahrer (eng. shared route percentage, SRP) zu maximieren und die  $k$  besten Mitfahrgelegenheiten mit absteigender SRP, für einen Mitfahrenden zu finden. Der Ansatz basiert auf einem erweiterungsbasiertem Ansatz. Hierbei werden zuerst zwei verschiedene Mengen mithilfe des Dijkstra Algorithmus berechnet. Eine Menge  $D_{asc}^{start}$ , bildet alle Fahrer nach aufsteigender Distanz vom Startpunkt bis zum Startpunkt des Mitfahrers ab. Eine zweite Menge  $D_{asc}^{ziel}$  bildet alle Fahrer nach aufsteigender Distanz vom Zielpunkt bis zum Zielpunkt des Mitfahrers ab. Dann wird eine Grenze mithilfe des Fahrers  $D_{asc}^{start}[0]$ , der den geringsten Abstand zwischen den beiden Startpositionen hat und dem Fahrer  $D_{asc}^{ziel}[0]$  der den geringsten Abstand zwischen beiden Zielpositionen hat, berechnet. Anschließend wird die SRP dieser beiden Fahrer und dem Mitfahrer berechnet. Die Ergebnisse werden in einer Prioritätswarteschlange  $Q$  gespeichert. In der nächsten Iteration wird die SRP der beiden nächsten Fahrer  $D_{asc}^{start}[1]$  und  $D_{asc}^{ziel}[1]$  berechnet.

Falls diese eine höhere SRP als bereits gespeicherte Mitfahrgelegenheiten besitzen werden sie der Prioritätswarteschlange hinzugefügt. Außerdem wird die Grenze mithilfe dieser Fahrer aktualisiert. Bei jeder neuen Iteration wird außerdem geprüft, ob die definierte Grenze kleiner als die der SRP des  $k$ 'ten Fahrers in der Prioritätswarteschlange ist. Falls dies zutrifft, terminiert der Algorithmus und es wurden alle  $k$  Fahrer mit der höchsten SRP für einen Mitfahrer gefunden. Abb.2 zeigt ein Beispiel bei dem Dreiecke Startpunkte und Vierecke Zielpunkte darstellen. Da dieser Ansatz aber voraussetzt, dass viele Distanzen mithilfe des Dijkstra Algorithmus berechnet werden müssen, kann er hohe Berechnungskosten mit sich bringen.



**Abbildung 2: Beispiel des erweiterungsbasierten Ansatz von Ta et al. [7]**

Das *Go!Caronas* Projekt, das von Cruz et al. [8] veröffentlicht wurde, beschreibt ein Echtzeit Mitfahrgelegenheitssystem mit Social Network Funktionalitäten. Um mögliche Mitfahrgelegenheiten den Nutzern des Systems vorzuschlagen werden die Trajektorien und Zeitpläne der Nutzer verglichen und geclustert. Trajektorien bezeichnen Wege oder Pfade in der Physik. Der hierfür verwendete Ansatz kann in drei Schritte unterteilt werden: (i) Diskretisierung der Trajektorie (ii) zeitliche Filterung und (iii) Clustering. Im ersten Schritt wird die Anzahl der Wegpunkte einer Trajektorie reduziert, indem eine Teilmenge,

bestehend aus charakteristischen Wegpunkten, gebildet wird. Vorallem bei Trajektorien, die durch GPS aufgezeichnet wurden, kann dieser Schritt hilfreich sein, um redundante Punkte und die Menge der Daten zu verringern [9]. Diese Wegpunkte werden auch *Points of Interest (POI)* genannt. Für diesen Schritt wurde ein Algorithmus namens *RotaFacil* benutzt, der in [10, 11] veröffentlicht wurde. Im nächsten Schritt werden die Nutzer gefiltert, so dass nur Nutzer, die eine ähnliche Abfahrts- und Ankunftszeit besitzen für eine weitere Verarbeitung infrage kommen. Im letzten Schritt werden die Nutzer anhand der Trajektorien geclustert. Hierbei wird die Ähnlichkeitsfunktion, die in [9] veröffentlicht wurde, verwendet. Diese Ähnlichkeitsfunktion basiert auf einer Distanzfunktion die die Wölbung der Erde mit einberechnet. Um einem Mitfahrer mit einem Fahrer zu matchen, werden die Distanzen der POI's eines Fahrers zu den Start und Zielpunkten des Mitfahrers berechnet. Wenn eine berechnete Distanz kleiner als ein Schwellenwert  $\epsilon$  ist, besteht ein Match. Der Parameter  $\epsilon$  beschreibt also die maximale Entfernung eines Fahrers von seiner Route, die er in kauf nehmen muss.

Cao et al. [12] veröffentlichen mit *SHAREK* einen skalierbaren und effizienten Mitfahrgelegenheitsservice. *SHAREK* ermöglicht es Nutzern, die auf der Suche nach einer Mitfahrgelegenheit sind, einen maximalen Preis, sowie maximale Wartezeit anzugeben. Der Preis einer Mitfahrgelegenheit hängt hierbei von der Länge der Strecke und dem erzeugten Umweg für den Fahrer ab. Wie viele Algorithmen zuvor, erreicht *SHAREK* seine Skalierbarkeit durch die Minimierung von rechenaufwendigen Berechnungen des kürzesten Wegs. Wichtig zu erwähnen ist, dass in dieser Arbeit Distanz, Zeit und Preis miteinander verglichen werden. Die Autoren Argumentieren, dass die Werte proportional zueinander

sind und lediglich ein Faktor miteinander berechnet werden sollte um bspw. den letztendlichen Preis für den Mitfahrer zu berechnen. Im Detail besteht der Matching-Algorithmus aus drei verschiedenen Phasen: (i) *Euclidian Temporal Pruning*, (ii) *Euclidean Cost Pruning* und (iii) *Semi-Euclidean Skyline-aware Pruning*. In der ersten Phase werden mithilfe der Euklidischen Distanz Fahrer aussortiert, die bei einer optimalen Route (Luftlinie) nicht in der vorgegebenen Wartezeit beim Mitfahrer ankommen würden. Da die reale Route der Fahrer aber meist länger als die Distanz zwischen Fahrer und Mitfahrer ist, können unter den restlichen Fahrern noch immer false-positives sein, die nicht rechtzeitig beim Mitfahrer ankommen würden. In der zweiten Phase werden daher weitere Fahrer, die folgende Formel nicht erfüllen, aussortiert.

$$PickupED(d, r) + 2 * Trip(r) + ReturnED(d, r) - Trip(d) < r_{max\_price} \quad (1)$$

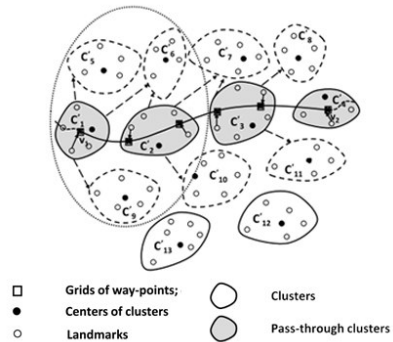
Wobei *PickupED* die Euklidische Distanz zwischen der aktuellen Position des Fahrers und der Startposition des Mitfahrers ist und *ReturnED* die Euklidische Distanz zwischen den beiden Zielpunkten. *Trip* beschreibt den kürzesten Weg vom Start zum Ziel von entweder Fahrer ( $r$ ) oder Mitfahrer ( $d$ ). Diese muss lediglich einmal berechnet werden für jeden Fahrer und Mitfahrer.  $r_{max\_price}$  beschreibt den maximalen Preis den der Mitfahrer bereit ist zu zahlen. In der dritten und letzten Phase wird die echte Entfernung, von der aktuellen Position des Fahrers bis zur Startposition des Mitfahrers, über das Straßennetz, berechnet.

$$Pickup(d, r) + 2 * Trip(r) + ReturnED(d, r) - Trip(d) < r_{max\_price} \quad (2)$$

Diese Distanz wird dann benutzt um die Semi-Euklidischen Kosten mithilfe Formel 2 zu berechnen. Mithilfe eines inkrementellen Straßennetz nearest-neighbor Algorithmus werden dann die übrig gebliebenen Fahrer anhand der echten Entfernung über das Straßennetz zum Mitfahrer sortiert. Falls der erste und somit nächste Fahrer es nicht in der angegebenen Wartezeit ( $r_{max\_time}$ ) bis zum Mitfahrer schafft, können alle Fahrer verworfen werden, da die nachfolgenden Fahrer eine noch längere Strecke bis zum Mitfahrer zurück legen müssten. Falls der Fahrer aber in angegebener Wartezeit beim Mitfahrer ankommen könnte und Formel 2 erfüllt muss zuletzt die Strecke vom Zielort des Mitfahrers bis zum Ziel des Fahrers über das Straßennetz berechnet werden. Somit wurde für diesen Fahrer die komplette reale Strecke berechnet und falls diese kleiner als der vom Mitfahrer angegebene maximale Preis ist, kann dieser in die Ergebnismenge aufgenommen werden. Um nur noch weitere Fahrer zu finden die besser sind, wird außerdem  $r_{max\_price}$  auf die Kosten des gefundenen Fahrers gesetzt. Sobald ein Fahrer gefunden wurde, der es nicht in der angegebenen Zeit bis zum Mitfahrer schafft oder alle Fahrer iteriert wurden, kann der Algorithmus terminiert und die Ergebnismenge zurückgegeben werden.

Thangaraj et al. [13] veröffentlichen ein weiteres System, namens *Xhare-a-ride*, das skalierbar, effizient und hochgradig suchoptimiert ist, sodass für jede Anfrage mehrere potentielle Fahrer gefunden werden können. Das System nutzt eine Vorverarbeitung, bei der geographische Regionen hierarchisch in drei verschiedene Ebenen aufgeteilt werden. Die erste Ebene besteht aus einem *Grid*, das aus  $100\text{ m}^2$  großen Quadraten besteht. Die nächste Ebene besteht aus sogenannten Landmarken, die interessante Punkte wie bspw. Bushaltestellen, Einkaufszentren oder bedeutende Gebäude darstellen. Jedes Quadrat des Grids bekommt eine Landmarke zugewiesen

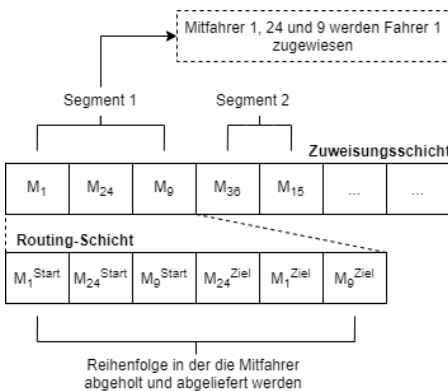
die, am nächsten zum Mittelpunkt des Quadrats liegt. Hierbei wird sowohl die Fahrtdistanz über das Straßennetz als auch die Distanz zu Fuß berechnet. Die dritte Ebene beschreibt *Cluster* aus Landmarken, die nicht weiter als eine definierte Distanz voneinander entfernt sind. Jede Landmarke wird einem einzigen Cluster zugewiesen und die Anzahl der Cluster wird minimiert. Da dieses Problem der Clusterbildung NP-Schwer ist, veröffentlichten die Autoren außerdem einen Greedy Algorithmus, der eine Annäherung der Lösung berechnen kann. Ein Unterschied zu den bereits vorgestellten Systemen ist, dass der Nutzer bei einer Anfrage angeben kann, wie weit er maximal Laufen möchte, um von der Mitfahrgelegenheit abgeholt zu werden. Das System speichert für jede angelegte Mitfahrgelegenheit alle Cluster die durchquert werden und die innerhalb des Umweglimits, das vom Fahrer angegeben wird, erreicht werden können. Die Idee des Matchings ist, dass jede Anfrage die aus solch einem Cluster kommt zur korrespondierenden Mitfahrgelegenheit gematched werden kann. Abb. 3 zeigt ein Beispiel einer Route die durch verschiedene Cluster führt.



**Abbildung 3: Beispiel des Matchingansatz von Thangaraj et al. [13] bei dem die Route durch Cluster von Landmarken führt.**

## 4.2 One-to-Many

Weitere Arbeiten von Huang et al. [14] und Javidi et al. [15], die im Rahmen dieser Arbeit untersucht wurden, nutzen Evolutionäre Algorithmen. Diese Art der Algorithmen orientieren sich an den in der Natur vorkommenden Evolutionsprinzipien, um eine Lösung für ein Optimierungsproblem zu finden. Evolutionäre Algorithmen sind metaheuristische Verfahren, die unter anderem für NP-Vollständige Probleme verwendet werden können. Da das Matchingproblem bei Mitfahrgelegenheitssystemen NP-Schwer ist, bietet sich also die Verwendung dieser Algorithmen an [15]. Huang et al. verwenden einen Genetischen Algorithmus, der eine spezielle Form eines Evolutionären Algorithmus darstellt. Die grundlegende Vorgehensweise dieser Art von Algorithmen besteht daraus, die Eingabeparameter in Form von sogenannten Chromosomen darzustellen. In [14] bestehen Chromosomen bspw. aus zwei verschiedenen Schichten. Die Zuweisungsschicht speichert wie in Abb. 4 zu sehen die Zuweisungen von Fahrern zu Mitfahrern. Hierbei werden Mitfahrer chronologisch in Segmente aufgeteilt, die einzelnen Fahrern zugewiesen werden. Die einzelnen Felder der Seg-



**Abbildung 4: Chromosomenrepräsentation die von Huang et al. [14] verwendet wird.**

mente, die Mitfahrerindizes beinhalten, werden als Gene bezeichnet. Die zweite Schicht des Chromosoms wird als Routing-Schicht bezeichnet und beschreibt die Reihenfolge, in der die Mitfahrer aufgenommen und abgesetzt werden sollen. Die einzelnen Segmente jedes Fahrers werden dann, mithilfe einer Distanzmetrik initialisiert. Dieser Initialisierungsalgorithmus liefert eine Population von Chromosomen, die einer festgelegten Größe entspricht. Anschließend werden die Chromosomen bzw. Segmente, mithilfe einer sogenannten Fitnessfunktion bewertet. Während dieses Prozesses wird die schnellste Route jedes Segments berechnet und in die Routing-Schicht des Chromosoms eingetragen. Im nächsten Schritt eines Evolutionären Algorithmus findet eine Mutation bzw. Evolution der Chromosomen statt. Huang et al. kreuzen bspw. Chromosomen mit schlechter Bewertung und guter Fitnessbewertung. Hierbei werden die Segmente der beiden Chromosomen verglichen und die Segmente mit besserer Fitness beibehalten. Während dieses Vorgangs kann es vorkommen, dass ein Mitfahrer zu mehr als einem Fahrer zugewiesen wurde. Um diesen Fehler zu reparieren werden an diesen Stellen Mitfahrer eingesetzt, die noch keinen Fahrer zugewiesen bekommen haben. Dieser Vorgang, des Bewerten der Population und der anschließenden Mutation wird solange fortgeführt bis ein geschätztes Optimum erreicht ist. Das Ergebnis ist eine annähernd optimale Lösung des Problems.

Javidi et al. [15] verwenden einen ähnlichen evolutionsbasierten Ansatz, der die Biogeografie-basierte Optimierung (BBO) [16] nutzt. BBO beschreibt ein Verfahren, bei dem mögliche Lösungen für ein Optimierungsproblem als Inseln dargestellt werden. Diese Inseln tauschen verschiedene Merkmale miteinander aus, um so zu besseren Lebensräumen zu werden. Dieser Austausch findet solange statt, bis ein akzeptabler Lebensraum gefunden



wurde, der die Lösung für das Optimierungsproblem darstellt. Der an das Mitfahrgelegenheitsproblem angepasste Algorithmus erzeugt in einem ersten Schritt eine festgelegte Anzahl an Populationen, die aus verschiedenen Konfigurationen von Fahrer und Mitfahrer bestehen. Da jede Population einzigartig sein sollte wird ein Teil der Matches zufällig generiert, der restliche Teil wird ähnlich, wie in [14] mithilfe eines Greedy Algorithmus berechnet. Die einzelnen Konfigurationen von Fahrer und Mitfahrer, also eine Instanz einer Population, entspricht einer Insel des BBO Algorithmus. Die Merkmale einer Insel bestehen aus den Fahrern der Matches. Diese Merkmale werden dann durch einen Migrierungsalgorithmus zwischen den Inseln ausgetauscht, sodass immer bessere Lösungen entstehen. Die Güte einer Insel wird hierbei mithilfe einer Kostenfunktion berechnet, die sowohl die Maximierung der Matchingrate, als auch die Minimierung des Umwegs für die Fahrer beinhaltet.

Escalona et al. veröffentlichen in [17] eine modifizierte Version des suchbasierten Algorithmus von Ta et al. [7], der bis zu drei Personen einer Mitfahrgelegenheit zuweisen kann und somit in die Kategorie One-to-Many fällt. Hierfür muss der zugrundeliegende Algorithmus lediglich drei mal wiederholt werden. Um die Performance des Algorithmus zu verbessern werden außerdem nicht alle SRP's von Fahrer-Mitfahrer Paaren berechnet, sondern nur die, deren Start und Zielpunkt nah beieinander liegen. Diese Änderung basiert auf der Beobachtung, dass Fahrer-Mitfahrer Paare die eine hohe SRP besitzen auch nahe gelegene Start- und Zielpunkte besitzen.

Fan et al. veröffentlicht in [18] einen effizienten Algorithmus um Mitfahrgelegenheiten für viele Nutzer zu bilden. Der Ansatz verfolgt eine ähnliche Strategie wie [12] und versucht den Suchraum zu verkleinern indem potentielle Fahrer für einen Mitfahrer in drei

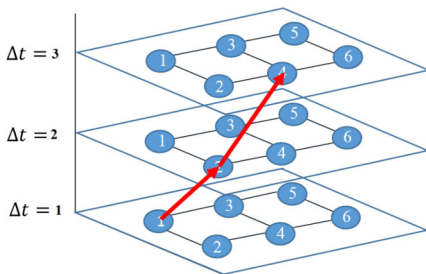
verschiedenen Schritten gefiltert werden. Im ersten Schritt werden Fahrer aussortiert, die eine frühere Abfahrtszeit haben als die späteste Zeit, die vom Fahrer angegeben wurde, um abgeholt zu werden. Im zweiten Schritt werden weitere Fahrer aussortiert, indem die euklidische Distanz zum Mitfahrer berechnet wird. Fahrer, die mit durchschnittlicher Geschwindigkeit diese Distanz bis zum spätesten Abholzeitpunkt des Mitfahrers nicht zurücklegen können, werden aussortiert. Des Weiteren wird in diesem Schritt überprüft, ob die Kosten eines Fahrers nicht die maximalen angegebenen Kosten des Mitfahrers übersteigen. Falls dies der Fall sein sollte, wird dieser Fahrer ebenfalls aussortiert. Im dritten und letzten Schritt des Algorithmus werden die Distanzen über das Straßennetz berechnet und der optimale Fahrplan für einen Fahrer ermittelt. Hierfür wird für jeden Kandidatenfahrer eines Mitfahrers die Start- und Zielposition in den Fahrplan eines Fahrers eingesetzt. Durch einen Brute-force Algorithmus werden alle Konfigurationen eines Fahrplans ausprobiert und Beschränkungen von bereits zugewiesenen Mitfahrern überprüft. So kann für jeden Fahrer die reale Umwegdistanz berechnet werden und der Fahrer mit dem niedrigsten Umweg dem Mitfahrer zugewiesen werden. Der Algorithmus findet, laut Fan et al. [18], aus mehr als 100.000 Fahrern, innerhalb von durchschnittlich einer halben Sekunde, den optimalen Fahrer für einen Mitfahrer.

### 4.3 Many-to-Many

Masoud und Jayakrishnan veröffentlichen in [19] einen Algorithmus um das Many-to-Many Matchingproblem zu lösen. Der Ansatz verwendet vordefinierte Stationen, die von den Teilnehmern benutzt werden können und vordefiniert sind. Die Stationen können von den Teilnehmern verwendet werden um Mitfahrgelegenheiten zu starten oder zu beenden oder zwischen Fahrzeugen oder

Transportarten umzusteigen. Die Autoren argumentieren, dass diese Herangehensweise mehrere Vorteile mit sich bringt. Zum einen kann so sichergestellt werden, dass die Fahrer diese Stationen problemlos finden können und Parkplätze vorhanden sind. Zum anderen kann so der Wohnort der Mitfahrer geheim gehalten werden. Außerdem kann so laut Stiglic et al. [20] die Rate der erfolgreichen Matches erhöht werden. Das System diskretisiert zudem die Zeit, indem Graphen der vordefinierten Stationen für Zeitintervalle mit geringer Dauer,  $\Delta t$ , erstellt werden. Wenn ein Teilnehmer des Systems von einer Station zu einer anderen Station reist, wird dies mit einem Link  $(n_i, n_j) = (t_i, s_i, t_j, s_j)$  dargestellt, wobei  $t_i$  den Zeitintervall angibt, bei dem der Reisende die Station verlassen muss um während des Zeitintervall  $t_j$  anzukommen.  $s_i$  und  $s_j$  repräsentieren Indizes von vordefinierten Stationen. Das Konzept der Diskretisierung von Raum und Zeit und die Verwendung von Links wird in Abb. 5 anschaulich dargestellt. Da dem System Teil-

Minute gesetzt werden. In einem ersten Vorverarbeitungsschritt des Matchings werden die Anzahl der für jeden Teilnehmer zugänglichen Links begrenzt indem Fahrer gefiltert werden, die aufgrund mangelnder räumlicher und zeitlicher Kompatibilität zwischen ihren Fahrten nicht Teil der Route eines Mitfahrers sein können. Hierfür werden alle Links der Teilnehmer berechnet die entlang der geplanten Route von Start- zur Zielstation, innerhalb der definierten maximalen Reisezeit, benutzt werden könnten. Diese Links können anschließend verwendet werden um die Menge an Fahrern und Mitfahrern zu verringern. Um als Mitfahrer eine Mitfahrgelegenheit zu finden muss mindestens ein Fahrer sowohl an der Startstation als auch der Zielstation vorhanden sein. Falls dies nicht der Fall ist kann der Mitfahrer aussortiert werden. Außerdem werden Fahrer aussortiert die keine Schnittmenge an Links mit anderen Mitfahrern des Systems bilden können. Anschließend wird eine Menge von Tupeln, aus Fahrern und Mitfahrer, deren Links eine Schnittmenge haben, gebildet. Anschließend wird iterativ versucht eine Konfiguration aus einem oder mehreren Fahrern für jeden Mitfahrer zu finden. Die grundlegende Idee hierbei ist es in jeder Iteration eine Anzahl von Teilproblemen zu lösen die das gesamte System darstellen können. Falls die Lösung dieser Teilprobleme keine Konflikte aufweisen wird der Algorithmus beendet und das globale Optimum des Optimierungsproblems wurde gefunden.



**Abbildung 5: Darstellung zwei verschiedener Links durch diskretisierten Raum und Zeit**

nehmer nach und nach beitreten wird ein sogenanntes *rolling horizon framework* eingesetzt, das eine erneute Optimierung des System, nach einem festgelegten Zeitintervall, vornimmt. Um einem Echtzeitsystem gerecht zu werden, könnte dieser Zeitintervall, laut den Autoren der Arbeit, bspw. auf eine

Masoud et al. [19] veröffentlichte noch im gleichen Jahr eine weiterführende Arbeit, welche die Verwendung des zuvor vorgestellten Algorithmus in einem Echtzeit Mitfahrgelegenheitssystem erlaubt. Hierzu wurde das System so verändert, sodass für jeden Mitfahrer, in einem first-come first-serve (FCFS) Schema, ein passender Fahrer gesucht wird. Jedoch werden die Fahrstrecken der zugewiesenen Fahrer nicht endgültig festgelegt.

Autoren	Jahr	Typ	Matching-ansatz	Komplexität	Fokus	Dynamisch
Ta et al. [7]	2018	one-to-one	Erweiterungsbasiert	● ● ● ●	Effizienz	X
Pelzer et al. [6]	2015	one-to-one	Zerlegungsalgorithmus	● ● ● ● ●	Skalierbarkeit	✓
Cruz et al. [8]	2016	one-to-one	Distanzfunktion	● ● ● ● ●	Anwendung	X
Thangaraj et al. [13]	2017	one-to-one	Clustering + Greedy Algorithmus	● ● ● ● ●	Suchoptimierung	✓
Cao et al. [12]	2015	one-to-one	Filterung + Skyline Algorithmus	● ● ● ● ●	Effizienz	✓
Schreieck et al. [1]	2016	one-to-many	Intelligente Datenstruktur	● ● ● ● ●	Effizienz	✓
Huang et al. [14]	2015	one-to-many	Evolutionärer Algorithmus (Genetisch)	● ● ● ● ●	Effizienz	X
Javidi et al. [15]	2021	one-to-many	Evolutionärer Algorithmus (Biogeographie-basiert)	● ● ● ● ●	Mehrzieloptimierung	✓
Escalona et al. [17]	2020	one-to-many	Modifikation von [7]	● ● ● ● ●	Erweiterung	X
Fan et al. [18]	2018	one-to-many	Filterung durch Abfahrtszeit + Euklidische Distanz und Straßennetz Distanz	● ● ● ● ●	Effizienz	✓
Masoud et al. [2]	2017	many-to-many	Vorverarbeitung + Zerlegungsalgorithmus	● ● ● ● ●	Problemlösung	✓
Masoud et al. [19]	2017b	many-to-many	Vorverarbeitung + Zerlegungsalgorithmus	● ● ● ● ●	Effizienz	✓

**Tabelle 1: Übersicht der vorgestellten Matchingalgorithmen**

Falls ein neuer Mitfahrer lediglich von einem Fahrer, der zuvor schon einem anderen Mitfahrer zugewiesen wurde, mitgenommen werden kann, können diese beiden Mitfahrer eine Verhandlung starten und einen Austausch der Fahrer veranlassen.

## 5 Diskussion

Im vorherigen Kapitel wurden einige Arbeiten aus dem Stand der Technik vorgestellt. In diesem Kapitel soll diskutiert werden, welche der vorgestellten Ansätze oder welche Teile davon für den vorgestellten Anwendungsfall des *URide* Mitfahrgelegenheitssystem verwendet werden könnten. Grundsätzlich könnten fast alle vorgestellten Arbeiten für das Matching in *URide* verwendet werden. Allerdings gibt es große Differenzen in der Komplexität der Algorithmen. Tabelle 1 liefert eine Übersicht der vorgestellten Arbeiten und deren Komplexität. Die Komplexität der Ansätze korreliert mit dem Matchingtyp. Das bedeutet die Ansätze werden aufsteigend von *One-to-One* zu *Many-to-Many* immer komplexer, da die Problemstellung auch schwieriger wird. *Many-to-Many* Ansätze bieten zwar Vorteile wie bspw. eine theoretisch höhere Matchingrate als andere Matchingtypen, allerdings kommen sie für *URide* nicht infrage, da die Teilnehmer, die vorerst auf die Studenten der Hochschule Reutlingen beschränkt sind, immer den gleichen Zielort bzw. Startort haben. Ein Transfer von Fahrzeug zu Fahrzeug bringt somit keinen Vorteil. Viele der vorgestellten *One-to-One* und *One-to-Many* Arbeiten [1, 7, 12, 14, 18, 19] fokussieren sich darauf, die Effizienz und Leistung zu verbessern. Im Falle der *URide* Anwendung spielt dieser Faktor allerdings eine untergeordnete Rolle, da die Zuordnung von Mitfahrern zu Fahrern nicht in Echtzeit ablaufen muss, sondern auch über einen Zeitraum von mehreren Minuten oder Stunden durchgeführt werden kann. Zudem behandeln viele der vorgestellten Arbeiten ein dynamisches

Mitfahrgelegenheitssystem indem Teilnehmer, nach und nach, dem System beitreten und eine nahezu sofortige Antwort erhalten wollen. *URide* ist insofern dynamisch, dass neue Teilnehmer dem System hinzugefügt werden können, jedoch wird die in den meisten Arbeiten beschriebene Dynamik nicht benötigt. Falls *URide* für weitere Teilnehmer veröffentlicht werden würde, wäre natürlich auch die Verwendung eines *Many-to-Many* Ansatzes möglich. Die Verwendung eines solchen Ansatzes würde es außerdem ermöglichen, weitere Transportmittel wie bspw. Bus und Bahn einzubinden, um ein multimodales Mitfahrgelegenheitssystem zu bilden. *One-to-One* Ansätze bieten die geringste Komplexität, besitzen allerdings eine geringere Matchingrate als *One-to-Many* Ansätze, da mehrere Personen von einem Fahrzeug transportiert werden können. Ein Nachteil dieser Ansätze ist allerdings, dass das Transportieren von mehreren Gästen, einen potentiell höheren Umweg für den Fahrenden bedeutet. Da Fahrer aber letztendlich, abhängig von der Länge der Strecke, für das transportieren von Mitfahrern belohnt werden müssen, ist dieser Aspekt zu vernachlässigen. Falls ein Fahrer diesen Potentiell größeren Umweg nicht in Kauf nehmen möchte, kann er außerdem auswählen wie viel Personen er maximal mitnehmen möchte. Aus diesen Gründen ist ein *One-to-Many* Matchingalgorithmus am besten für die Verwendung in *URide* geeignet. Die Entscheidung, welcher der vorgestellten *One-to-Many* Algorithmen am besten für den *URide* Anwendungsfall passt, ist allerdings nicht trivial. Die vorgestellten Arbeiten können nämlich nicht ohne weiteres quantitativ verglichen werden, da die in den Arbeiten erhobenen Metriken wie bspw. Matchingrate, durchschnittliche Wartezeit oder Kilometer Einsparungen durch unterschiedliche Simulationen und Konfigurationen dieser Simulationen entstanden sind. Außerdem besteht keine Konsistenz, der erhobenen Metriken, bei den untersuchten Arbeiten.

Der Algorithmus, der von Schreieck et al. [1] veröffentlicht wurde sticht mit seiner simplen Herangehensweise heraus und bietet sogar eine ausreichende Performance für Echtzeitanwendungen. Die geringe Komplexität bedeutet auch, dass eine relativ einfache Implementierung des Algorithmus möglich ist. In einer von Schreieck et al. [1] durchgeführten Simulation mit 10000 Mitfahrgelegenheiten braucht der Algorithmus 0,4 Sekunden um den besten Fahrer für einen Mitfahrer zu finden. Des weiteren könnte die Diskretisierung von GPS Aufzeichnung aus [8, 9] in *URide* Anwendung finden. Wie bereits erwähnt können mit dem von Cruz et al. [8] verwendeten *RotaFacil* Algorithmus die oftmals redundant vorhandenen GPS Positionsdaten in POI's diskretisiert und somit die Datenmenge der Routen verringert werden. Das hätte außerdem zur Folge, dass die Performance des Matching Algorithmus von Schreieck et al. [1] verbessert werden würde.

## 6 Fazit

Ziel dieser Arbeit war es, ein passenden Algorithmus für den Einsatz im *URide* Mitfahrgelegenheitssystem zu finden. Mithilfe einer systematischen Literaturrecherche wurden verschiedene Ansätze des Stand der Technik ermittelt, um Teilnehmer eines Mitfahrgelegenheitssystem mit ähnlichen Reiserouten zu finden. Es stellte sich heraus, dass viele der untersuchten Arbeiten unterschiedliche Evaluationsmethoden verwenden und teilweise auch unterschiedliche Problemstellungen behandeln. Daher sind diese Arbeiten nicht quantitativ vergleichbar. Um den besten Matchingalgorithmus zu finden müsste also ein einheitlicher Benchmark erstellt werden mit dem diese Algorithmen verglichen werden könnten. Nichtsdestotrotz wurde ein passender Algorithmus für den *URide* Anwendungsfall, anhand des Matchingtyps, der Komplexität und der Performance identifiziert. Dieser

von Schreieck et al. [1] veröffentlichte Ansatz kann in Zukunft in das *URide* System eingebaut werden. Hierfür wird, aufgrund der niedrigen Komplexität, ein geringer Aufwand erwartet. Laut den Autoren wäre es außerdem denkbar, vor und nach der Mitfahrgelegenheit, Vorschläge für öffentliche Verkehrsmittel in das System zu integrieren, um so eine Multimodalität des Systems zu erreichen.

## Literatur

- [1] M. Schreieck, H. Safetli, S. A. Siddiqui, C. Pflügler, M. Wiesche und H. Krcmar, „A Matching Algorithm for Dynamic Ridesharing,“ *Transportation Research Procedia*, Jg. 19, Nr. 4, S. 272–285, 2016, ISSN: 23521465.
- [2] N. Masoud und R. Jayakrishnan, „A real-time algorithm to solve the peer-to-peer ride-matching problem in a flexible ridesharing system,“ *Transportation Research Part B: Methodological*, Jg. 106, Nr. 9, S. 218–236, 2017, ISSN: 01912615.
- [3] M. Furuhata, M. Dessouky, F. Ordóñez, M.-E. Brunet, X. Wang und S. Koenig, „Ridesharing: The state-of-the-art and future directions,“ *Transportation Research Part B: Methodological*, Jg. 57, Nr. 9, S. 28–46, 2013, ISSN: 01912615.
- [4] A. Tafreshian, N. Masoud und Y. Yin, „Frontiers in Service Science: Ride Matching for Peer-to-Peer Ride Sharing: A Review and Future Directions,“ *Service Science*, Jg. 12, Nr. 2-3, S. 44–60, 2020, ISSN: 2164-3962.
- [5] N. V. Bozdog, M. X. Makkes, A. van Halteren und H. Bal, „RideMatcher: Peer-to-Peer Matching of Passengers for Efficient Ridesharing,“ in *18th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, Piscataway, NJ: IEEE, 2018, S. 263–272, ISBN: 978-1-5386-5815-4. DOI: 10.1109/CCGRID.2018.00041.

- [6] D. Pelzer, J. Xiao, D. Zehe, M. H. Lees, A. C. Knoll und H. Aydt, „A Partition-Based Match Making Algorithm for Dynamic Ridesharing,“ *IEEE Transactions on Intelligent Transportation Systems*, Jg. 16, Nr. 5, S. 2587–2598, 2015, ISSN: 1524-9050.
- [7] N. Ta, G. Li, T. Zhao, J. Feng, H. Ma und Z. Gong, „An Efficient Ride-Sharing Framework for Maximizing Shared Route,“ *IEEE Transactions on Knowledge and Data Engineering*, Jg. 30, Nr. 2, S. 219–233, 2018, ISSN: 1558-2191.
- [8] M. Cruz, H. Macedo, E. Mendonca und A. Guimaraes, „GO!Caronas: Fostering ridesharing with online social network, candidates clustering and ride matching,“ in *2016 8th Euro American Conference on Telematics and Information Systems (EATIS)*, Y. A. Rodríguez, Hrsg., Piscataway, NJ: IEEE, 2016, S. 1–8, ISBN: 978-1-5090-2436-0. DOI: 10.1109/EATIS.2016.7520120.
- [9] M. O. Cruz, H. Macedo und A. Guimaraes, „Grouping Similar Trajectories for Carpooling Purposes,“ in *2015 Brazilian Conference on Intelligent Systems*, Piscataway, NJ: IEEE, 2015, S. 234–239, ISBN: 978-1-5090-0016-6. DOI: 10.1109/BRACIS.2015.36.
- [10] R. Teles, B. Barroso, A. Guimaraes und H. Macedo, „Automatic Generation of Human-like Route Descriptions: A Corpus-driven Approach,“ *Journal of Emerging Technologies in Web Intelligence*, Jg. 5, Nr. 4, 2013, ISSN: 1798-0461. DOI: 10.4304/jetwi.5.4.413-423.
- [11] R. Telles, A. Guimarães und H. Macedo, „Automated feeding of POI base for the generation of route descriptions,“ in *6th Euro American Conference on Telematics and Information Systems (EATIS), 2012*, J. J. S. Zapater, Hrsg., Piscataway, NJ: IEEE, 2012, S. 253, ISBN: 9781450310123. DOI: 10.1145/2261605.2261643.
- [12] B. Cao, L. Alarabi, M. F. Mokbel und A. Basalamah, „SHAREK: A Scalable Dynamic Ride Sharing System,“ in *2015 16th IEEE International Conference on Mobile Data Management (MDM)*, Piscataway, NJ: IEEE, 2015, S. 4–13, ISBN: 978-1-4799-9972-9. DOI: 10.1109/MDM.2015.12.
- [13] R. S. Thangaraj, K. Mukherjee, G. Raravi, A. Metrewar, N. Annamaneni und K. Chattopadhyay, „Xhare-a-Ride: A Search Optimized Dynamic Ride Sharing System with Approximation Guarantee,“ in *ICDE 2017*, Piscataway, NJ: IEEE, 2017, S. 1117–1128, ISBN: 978-1-5090-6543-1. DOI: 10.1109/ICDE.2017.156.
- [14] S.-C. Huang, M.-K. Jiau und C.-H. Lin, „A Genetic-Algorithm-Based Approach to Solve Carpool Service Problems in Cloud Computing,“ *IEEE Transactions on Intelligent Transportation Systems*, Jg. 16, Nr. 1, S. 352–364, 2015, ISSN: 1524-9050.
- [15] H. Javidi, D. Simon, L. Zhu und Y. Wang, „A multi-objective optimization framework for online ridesharing systems,“ in *2021 IEEE International Conference on Big Data and Smart Computing (BigComp)*, IEEE, 1/17/2021 - 1/20/2021, S. 252–259, ISBN: 978-1-7281-8924-6. DOI: 10.1109/BigComp51126.2021.00054.
- [16] D. Simon, „Biogeography-Based Optimization,“ *IEEE Transactions on Evolutionary Computation*, Jg. 12, Nr. 6, S. 702–713, 2008, ISSN: 1089-778X. DOI: 10.1109/TEVC.2008.919004.
- [17] J. Angelo Escalona, B. Manalo, W. J. R. Limjoco und C. C. Dizon Electrical, „A Ride Sharing System based on An Expansive Search-Based Algorithm,“ in *2020 IEEE REGION 10 CONFERENCE (TENCON)*, IEEE, 11/16/2020 - 11/19/2020, S. 870–874, ISBN: 978-1-7281-8455-5. DOI: 10.1109/TENCON50793.2020.9293943.

- [18] J. Fan, J. Xu, C. Hou, B. Cao, T. Dong und S. Cheng, „URoad: An Efficient Algorithm for Large-Scale Dynamic Ridesharing Service,“ in *2018 IEEE International Conference on Web Services*, Piscataway, NJ: IEEE, 2018, S. 9–16, ISBN: 978-1-5386-7247-1. DOI: 10.1109/ICWS.2018.00009.
- [19] N. Masoud und R. Jayakrishnan, „A decomposition algorithm to solve the multi-hop Peer-to-Peer ride-matching problem,“ *Transportation Research Part B: Methodological*, Jg. 99, S. 1–29, 2017, ISSN: 01912615. DOI: 10.1016/j.trb.2017.01.004.
- [20] M. Stiglic, N. Agatz, M. Savelsbergh und M. Gradisar, „The Benefits of Meeting Points in Ride-Sharing Systems,“ *SSRN Electronic Journal*, 2015. DOI: 10.2139/ssrn.2567274.



©2021 Marius Matthäus Bauer. Lizenznehmer Hochschule Reutlingen, Deutschland. Dieser Artikel ist ein Open-Access-Artikel unter den Bedingungen und Konditionen der Creative Commons Attribution (CC BY)-Lizenz. <http://creativecommons.org/licenses/by/4.0>

# Status Quo von künstlicher Intelligenz in Extended Reality-Anwendungen

Alexander Brückner

Hochschule Reutlingen

alexander.brueckner@Student.Reutlingen-University.de

## Abstract

Diese Arbeit setzt sich zum Ziel, einen Status Quo bezüglich der Anwendung von KI-Verfahren im Bereich von Extended Reality zu schaffen. Die grundlegenden Verfahren und Techniken werden eingeführt und eine umfangreiche Literaturrecherche dient als Mittel der Wahl, um bestehende Arbeiten zu sondieren und daraus einen Konsens zu schaffen. Die immense Zahl relevanter Arbeiten wird auf eine Auswahl der prägnantesten reduziert und die entsprechenden Arbeiten werden genauer betrachtet und diskutiert. Abschließend ist ein Ausblick in die Zukunft möglich - so können KI-basierte Systeme durchaus als Alternative zu traditionellen Verfahren im Bereich XR verstanden werden oder lediglich unterstützende Rollen einnehmen. Umgekehrt kann Extended Reality jedoch als vielversprechendes Mittel in der Vermittlung von KI-relevanten Inhalten gesehen werden.

## CCS Concepts

• **Human-centered computing** → **Virtual reality**; Mixed / augmented reality; • **Computing methodologies** → **Machine learning approaches**;

## Keywords

Virtual Reality, Augmented Reality, Mixed Reality, Artificial Intelligence

## 1 Einleitung

KI-Anwendungen werden immer häufiger als zuverlässige Lösungen für Problemstellungen in Bereich des Autonomen Fahrens oder der Bildverarbeitung eingesetzt und sind ein

aktives Feld der Forschung. Im Forschungsfeld von Extended Reality existieren aktuelle Herausforderungen wie die Vorhersage von VR Sickness [1], [2] und Verfahren wie Redirected Walking [3] [4], die ebenfalls immer öfter durch Einsatz von künstlicher Intelligenz gelöst werden können. Diese Arbeit setzt sich zum Ziel, entsprechende Schnittpunkte zwischen Herausforderungen im Bereich XR und etablierten Verfahren der künstlichen Intelligenz zu finden und zu analysieren. Der Status Quo soll durch die Betrachtung mehrerer KI-basierter Lösungsansätze hervorgehoben werden.

## 2 Methodik

Dieses Kapitel bietet Aufschluss über die allgemeine Vorgehensweise, mit der ein Status Quo etabliert werden soll. Der genaue Ablauf einer Literaturrecherche und deren Ergebnisse werden erörtert und diskutiert.

### 2.1 Literaturrecherche

Die in der Arbeit vorgestellten Erkenntnisse sind im Rahmen einer Literaturrecherche entstanden, die auf einer Keywordsuche basiert. Da Extended Reality als Überbegriff von VR, AR und MR gesehen werden kann [5], wurden die entsprechenden Begriffe einzeln in

---

Betreuer/-in Hochschule

Prof. Dr. Uwe Kloos  
Hochschule Reutlingen  
uwe.kloos@Reutlingen-  
University.de

Informatics Inside Herbst 2021  
24. November 2021 Hochschule Reutlingen  
Copyright 2021 Alexander Brückner



Verbindung Artificial Intelligence und Machine Learning als Keywords für die Suche verwendet. Die erste Keywordsuche wurde in den drei Datenbanken IEEExplore von IEEE <sup>1</sup>, ACM Digital Library <sup>2</sup> und Google Scholar<sup>3</sup> durchgeführt. Dabei wurden lediglich Arbeiten zwischen einschließlich 2017 und 2021 betrachtet. Die genauen Suchanfragen an die Datenbanken setzen sich wie folgt zusammen:

- Virtual Reality & Artificial Intelligence (VR AI)
- Virtual Reality & Machine Learning (VR ML)
- Augmented Reality & Artificial Intelligence (AR AI)
- Augmented Reality & Machine Learning (AR ML)
- Mixed Reality & Artificial Intelligence (MR AI)
- Mixed Reality & Machine Learning (MR ML)

Im ersten Schritt wurden lediglich die Anzahl an Suchergebnissen in den jeweiligen Datenbanken festgehalten. Da eine Analyse von einer so großen Zahl von Arbeiten den Rahmen dieser Arbeit eklatant übersteigen würde, werden lediglich Ergebnisse von IEEExplore im weiteren Verlauf berücksichtigt. Als Rechtfertigung kann die Tatsache gesehen werden, dass IEEE selbst eine jährliche Konferenz im Bereich Virtual Reality abhält <sup>4</sup>, in welcher aktuelle Entwicklungen im entsprechenden Bereich vorgestellt werden, der selbst wieder unter die Kategorie Extended Reality fällt.

<sup>1</sup><https://ieeexplore.ieee.org/Xplore/home.jsp>

<sup>2</sup><https://dl.acm.org/>

<sup>3</sup><https://scholar.google.com/>

<sup>4</sup><https://ieeevr.org/2021/>

	ACM	Google Scholar	IEEE
VR AI	24257	23300	1698
VR ML	24496	71600	911
AR AI	8148	18900	984
AR ML	23586	33000	396
MR AI	8293	21400	315
MR ML	23772	5700	96

**Tabelle 1: Ergebnisse der ersten Keyword-basierten Literaturrecherche**

Um die Ergebnisse auf IEEExplore weiter einzugrenzen, werden primär Arbeiten im Bereich von Virtual Reality anstelle von Augmented Reality und Mixed Reality betrachtet. Weiterhin werden die Ergebnisse im VR Bereich wie folgt kategorisiert:

- Sickness
- Rendering
- Content
- Locomotion

Eine letzte Keyword-basierte Suche vereint die vier Kategorien mit den Begriffen Virtual Reality und Artificial Intelligence bzw. Virtual Reality und Machine Learning.

VR AI Sickness	23
VR ML Sickness	15
VR AI Rendering	115
VR ML Rendering	32
VR AI Content	154
VR ML Content	15
VR AI Locomotion	34
VR ML Locomotion	87

**Tabelle 2: Anzahl der Ergebnisse von Recherche mit Bezug auf Kategorien**

Die deutlich eingegrenzte Quantität an relevanten Arbeiten wurde dann letztmalig auf Arbeiten eingegrenzt, die folgenden Kriterien entsprechen, um die aussagekräftigsten Ergebnisse zu erhalten:

- Virtual Reality als Zielsystem
- Anwendung eines Machine Learning/KI Ansatzes
- Durchführung einer Nutzerstudie

Ergebnisse dieser Suchanfrage wurden stichprobenartig näher betrachtet, dabei wurde der Abstract gelesen und besonders auffällige Arbeiten wurden in Gänze gelesen. Auffällig war die vergleichsweise große Zahl von Arbeiten, die sich zwar im Bereich Extended Reality bewegen und laut Keywords auch im Bereich Artificial Intelligence bzw. Machine Learning tätig sind, letzteres bei näherer Inspektion aber nicht der Fall ist. So wird in mehreren Fällen zwar auf entsprechende Literatur verwiesen, aber weder näher erörtert oder aufbereitet. Daraus folgt eine im Vergleich zur ursprünglich gefundenen Menge an Arbeiten vehement verkleinerte Auswahl von tatsächlich brauchbarer Literatur.

Aus den daraus resultierenden Ergebnissen werden die prägnantesten Arbeiten weiter vorgestellt, um ein Bewusstsein für den State of the Art von KI-unterstützten Extended Reality Systemen zu schaffen und einen Blick in die Zukunft möglich zu machen.

### 3 Extended Reality

Dieses Kapitel soll den Begriff Extended Reality aufbereiten und eine Einführung in die aktuellen Entwicklungen und Herausforderungen bieten. Neben der genauen Definition von Extended Reality und deren Sparten wird der State of the Art der Forschung diskutiert und aufbereitet.

#### 3.1 Definition Extended Reality

Extended Reality (XR) selbst kann als Überbegriff für die Technologien Virtual Reality (VR), Augmented Reality (AR) und Mixed

Reality (MR) verstanden werden [5]. Die wesentlichen Unterschiede zwischen VR und AR liegen in der Art und Weise, wie den Benutzenden Personen "künstliche Informationen" präsentiert werden [5]. Während die Benutzenden Personen einer Virtual Reality Applikation mit einer insgesamt virtuellen Welt ohne Bezug zur realen Welt interagieren und sich bewegen, in der sowohl die Umgebung als auch dessen Inhalt als künstlich zu sehen sind, wird den Benutzenden Personen einer Augmented Reality-Applikation lediglich künstliche Informationen auf eine natürliche Umgebung projiziert [5].

#### 3.2 State of the Art Extended Reality

Aktuelle Entwicklungen in der Welt von XR weisen eine Vielzahl von Ansätzen auf, welche das Ziel besitzen, die Immersion und Präsenz solcher Systeme zu verbessern. Eine Auswahl von Arbeiten aus diesem Bereich soll den State of the Art nahelegen. Ein Projekt der Tampere University in Finnland, genannt „HUMOR“ [6] befasst sich mit den Problemen im Bereich der Wahrnehmung innerhalb von XR-Anwendungen. Multimodale Interaktion, also Interaktionsmethoden, die sich nicht nur auf eine einzige Eingabe- und Ausgabemöglichkeit wie beispielsweise Handcontroller oder Treadmills beschränken, sondern mehrere Sinne der Benutzenden Personen gleichzeitig ausnutzen, um eine Interaktion möglich zu machen, sind eines der aktiven Forschungsthemen [6]. Multimodale Interaktion selbst kann auf mehrere Ebenen gruppiert werden, darunter parallele und synergistische Verfahren - als Beispiel einer parallelen Interaktion kann ein System verstanden werden, welches mittels interpretierter Handgesten eine Aktion ausführt, während gleichzeitig mittels Sprachbefehlen eine weitere, unabhängige Aktion ausgeführt werden kann. Eine synergistische Interaktion würde dieselben Interaktionen in eine Einzige

zusammenfassen, so könnte wie bei Rakkolainen et.al [6] beschrieben ein graphisches Objekt mittels einer Geste ausgewählt und zugleich mit einem Sprachbefehl verändert werden. Unter Multimodalität können also Ansätze verstanden werden, die sowohl eine erhöhte Usability als auch eine effizientere Interaktion mit dem System ermöglichen [6]. Neben Arbeiten im Bereich der Interaktion finden sich immer mehr Anwendungen mit einem Fokus auf Kollaboration. Danielsen et.al. beschreiben ein kollaboratives VR System, in dem Szenarien für Notfallmanagement-Situationen trainiert werden können [7]. Dafür wurden entsprechende Notfallszenarien als Trainingsfall entwickelt, welche dann von mehreren Benutzenden Personen innerhalb des Virtual Reality Systems durchschritten werden [7]. Die verfassenden Personen beschreiben eine mehrschichtige Kommunikationshierarchie, welche von den Benutzenden Personen als Virtual Reality Avatare belegt wird und den Rollen zugehörige Interaktionen zur Verfügung stellt - so kann ein Feuerwehrmann etwa mit Objekten wie einer Löschausrüstung interagieren und Medizinproufessional eine Triage durchführen [7]. Ergebnisse einer Evaluation des Systems zeigen, dass kollaborative Systeme im Bereich von XR Vorteile im Bereich des Lernens bieten - so berichten die verfassenden Personen von positivem Feedback in Bezug auf die Lernerfahrung in einer ihrer Usergruppen [7].

## 4 Künstliche Intelligenz/Machine Learning

Unter die Bezeichnung Künstliche Intelligenz und Machine Learning fallen Verfahren, welche Computersysteme gestatten, sich menschenähnlich zu verhalten [8]. Alan Turing definiert KI als eine Maschine, die von Menschen ebenfalls als menschlich wahrgenommen werden wird. Grundsätzlich werden aber auch Verfahren, die enorme Datenmengen in kürzester Zeit so verarbeiten können,

dass für Menschen leicht verständliche Ergebnisse zu Tage treten, verstanden [8]. Machine Learning und Deep Learning können als Unterbegriffe von künstlicher Intelligenz verstanden werden [8]. Im Kontext dieser Arbeit soll KI als Überbegriff solcher Verfahren etabliert werden und zur besseren Lesbarkeit gleichbedeutend sein. So kann ein Machine Learning Ansatz auch als KI Ansatz verstanden werden. Die im Rahmen dieses Beitrags relevanten Methodiken wie Reinforcement Learning und Decision Trees sollen in diesem Kapitel eingeführt werden.

### 4.1 Reinforcement Learning

Ansätze auf Basis von Reinforcement Learning verfolgen das Ziel, einem System durch Belohnungsfunktionen nahe zu legen, wann ein Sachverhalt als günstig und wann als weniger günstig zu betrachten ist [8]. Hierbei wird anders als bei Supervised Learning kein extern definierter Datensatz eingesetzt, welcher einem Agenten Aufschluss über korrektes Verhalten gibt. Umgekehrt fällt Reinforcement Learning aber auch nicht unter die Kategorie von Unsupervised Learning - die verfassenden Personen von **rlbasics** beschreiben es als ein Weiteres Machine Learning Paradigma. Die Anwendungsgebiete sind weitläufig, so gibt es Lösungsansätze auf Basis von RL in den Bereichen von Robotersteuerung, Simulationen, als Lösung für Optimierungprobleme und weiteren Use Cases [9]. Als verstärkende Metrik, die ein RL-basiertes System zu Ergebnissen führt, gilt eine Belohnungsfunktion. So kann in einem Kontext von Robotern die zurückgelegte Distanz als Belohnung gelten. Steigt diese an, gilt der verfolgte Ansatz als günstig und wird vom System verstärkt - im Falle einer Verminderung wird die vom Agenten verwendete Lösungsstrategie angepasst [9]. Auf diese Weise ist ein RL-basiertes System in der Lage, sich sukzessiv im Hinblick auf Resultate zu verbessern [9].

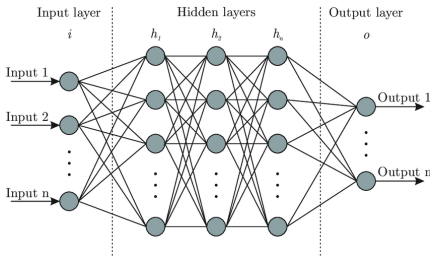
## 4.2 Decision Trees

Entscheidungsbäume (engl. Decision Trees) gelten als Datenstruktur mit dem Ziel zur Klassifikation von Daten und deren Vorhersage [10]. Strukturell können sie als absteigendes Geflecht aus Knoten betrachtet werden, wobei jeder Knoten selbst eine Überprüfung auf Attribute verstanden werden kann [10]. Blätter im Baum (Knoten ohne Kindsknoten) gelten als finales Element in der Entscheidungsreihe und zählen somit als Endergebnis des Entscheidungsbaums [10]. Die Bäume selbst werden mit einem Trainingsdatensatz trainiert, welcher dem System dann gestattet, die jeweiligen Entscheidungen bezüglich der Attribute der Eingabemenge zu treffen und entsprechende Elemente zu klassifizieren [10]. Beispielhaft können mittels eines Entscheidungsbaumes und einem passenden Trainingsdatensatz Vorhersagen getroffen werden, so konnte Zhang 2021 ein System auf Basis von Entscheidungsbäumen entwickeln, welches Entscheidungen in einem Tennisspiel klassifizieren und Aufschluss über deren Beitrag zum Erfolg liefern konnte [10]. Ein ähnlicher Ansatz wird von Padmanaban et. al. [1] verfolgt, um das Auftreten von Krankheitssymptomen während der Benutzung von 360° Medieninhalten vorherzusagen zu können.

## 4.3 Neuronale Netze

Künstliche Neuronale Netze werden als beliebtes Werkzeug für Herausforderungen im Machine Learning Bereich eingesetzt und replizieren den Lernprozess in biologischen Wesen, speziell anhand des menschlichen Nervensystems [11]. Als Neuronen bezeichnete Nervenzellen innerhalb des menschlichen Nervensystems sind durch sogenannte Axonen und Dendriten miteinander verbunden - diesen Verbindungspunkt bezeichnet man auch als Synapsen [11]. Ein externer Stimulus verändert die Stärke dieser synaptischen Verbindungen, was die Basis des menschlichen Lernprozesses abbildet und in

einem künstlichen Neuronalen Netz repliziert werden soll, um einen möglichst menschenähnlichen Lernprozess zu gewährleisten [11]. Als künstliche Synapsen (also die Verbindungen zwischen den Neuronen) versteht man im Kontext von künstlichen neuronalen Netzen eine Gewichtung - ändert sich diese Gewichtung, wird dadurch auch die Ausgabe des jeweiligen Neurons angepasst [11]. Die dafür im biologischen Prozess notwendigen, externen Stimuli werden bei einem künstlichen neuronalen Netz durch Trainingsdaten repliziert [11]. So können Bilddaten in eine Eingabeschicht aus Pixeln und eine Ausgabeschicht bestehend aus deren Merkmalen umgewandelt werden. Dabei können vom neuronalen Netz generierte Fehler als künstliche Form eines negativen Lernfeedbacks verstanden werden, welche das Netz wiederum so beeinflussen, dass die Ausgabe sich verändert [11]. Dies kann durch eine Veränderung der Gewichtungen anhand von fehlerhaften Vorhersagen geschehen [11]. Somit kann ein neuronales Netz wie im von Aggerwal genannten Beispiel trainiert werden, bestimmte Objekte zu erkennen, wie etwa Bananen. Wird das neuronale Netz mit einer möglichst großen Zahl an verschiedenen Bildern von dem zu identifizierenden Objekt versorgt, ist es nach einer gewissen Zeit in der Lage, die Banane auch in Bildern zu erkennen, die das Netz selbst noch nicht gesehen hat. Der unterliegende Prozess wird dabei als model generalization (dt. Generalisierung) verstanden [11].



**Abbildung 1: Schichten eines künstlichen neuronalen Netzwerks. Sichtbar die Eingabeschicht, die versteckte als computational layer [11] bezeichnete Mittelschicht, welche die Klassifikation bzw. Erkennung durchführt und die Ausgabeschicht, welche die Ergebnisse zeigt [12]**

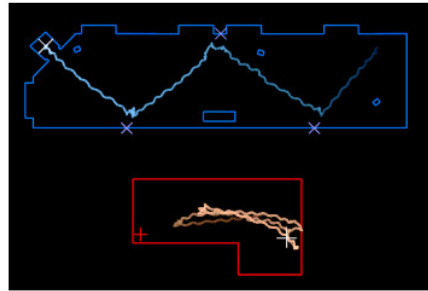
## 5 Einsatz von KI in XR

Die Recherche ergab eine Vielzahl von Anwendungsfällen von KI im Bereich Extended Reality. Eine Betrachtung aller recherchierten Artikel wäre zu zeitintensiv und umfangreich, stattdessen soll das Potenzial von KI-basierten Lösungen anhand von einigen wenigen, aber umso prägnanteren Beispielen erfolgen. So konnte die ursprünglich von Razaque et. al. [13] beschriebene Technik Redirected Walking bereits mit sichtbaren Fortschritten durch Reinforcement Learning angewendet werden.

### 5.1 Redirected Walking - Reinforcement Learning

Unter Redirected Walking versteht man Verfahren, welche die Wahrnehmung von Benutzenden Personen eines Virtual Reality Systems so verändern, dass eine Bewegung in einem großen virtuellen Raum auch in einem kleineren physikalischen Raum erfolgen können [13]. In einem Virtual Reality-System liegt das Verhältnis zwischen physikalischer Bewegung und virtueller Bewegung in der Regel bei eins zu eins - soll heißen, eine Bewegung über eine diskrete Distanz in der realen Welt entspricht der gleichen Distanz in der

virtuellen Welt [13]. Hierbei entsteht die Problematik, dass virtuelle Welten nicht zwangsweise an Grenzen gebunden sind, sie können beliebig groß gestaltet sein. Dennoch existieren für Benutzende physikalische Limitierungen wie die Größe des Raumes, in dem sie sich befinden. Durch Redirected Walking kann eine Schwäche der menschlichen Wahrnehmung ausgenutzt werden, um den Benutzenden Personen im Kreis"laufen zu lassen, ohne dafür ein Bewusstsein in der virtuellen Welt zu schaffen [13].



**Abbildung 2: Steering durch Redirected Walking. Blau zeigt den virtuell zurückgelegten Weg, Rot den tatsächlich im physikalischen Raum zurückgelegten Weg [13].**

Dies kann entweder durch vorgefertigte Wegpunkte, an denen die Umgebung dann geplant verändert wird (predictive redirected walking) geschehen oder durch reaktive Verfahren - genannt Steering Algorithms - umgesetzt werden [4]. Beispielhaft kann ein unendlicher Korridor gesehen werden - laufen Benutzende diesen Korridor entlang, wird ein sogenannter Steering Algorithmus die virtuelle Welt so anpassen, dass eine subtile Seitwärtsbewegung bzw. Drehung in der realen Welt erfolgt. Somit geht man virtuell zwar gradeaus, in der Realität aber im Kreis - ohne dabei an die Grenzen des physikalisch verfügbaren Raumes zu geraten [13]. Sollte dies doch der Fall sein, spricht man von einem

Hard Reset, in dem die Applikation pausiert wird und die Benutzenden Personen dazu gebeten werden, eine sichere Position einzunehmen, um dann fortzufahren, was jedoch massive Einschränkungen der Immersion bedeutet [4].

Abbildung 1 illustriert beispielhaft die Funktionsweise eines Steering Algorithmus anhand des Steer to Center Verfahrens: Der tatsächlich zurückgelegte Pfad der Benutzenden Personen in der wirklichen Welt, dargestellt durch rot, weicht deutlich von der virtuell zurückgelegten Strecke ab [13]. Mehrere solcher Algorithmen existieren, darunter Steer-to-Center, welcher die Benutzenden Personen stets zum Zentrum des physikalischen Raumes lenkt und Steer-to-Orbit, welcher eine Kreisbahn um das Zentrum herum anstrebt [4]. Darüber hinaus werden auch Ansätze mit Abstoßfeldern verfolgt, welche die Umgebung anhand deren Nähe zu den Benutzenden Personen anpassen [4]. Strauss et al. [4] beschreiben in ihrer Arbeit einen Ansatz auf Basis von Reinforcement Learning, in welchem die Manipulation der Umgebung durch ein neuronales Netz erfolgt.

Das neuronale Netz selbst wird mittels Reinforcement Learning darauf trainiert, wann Benutzende mit welcher Stärke umgelenkt werden, hierfür wird auf die Methodik von Proximal Policy Optimization zurückgegriffen, welche die vom Netz erlernte Policy vor erheblichen Abweichungen durch Aktualisierungen der Lernparameter schützt [4]. Das System selbst definiert eine Belohnungsfunktion, welche große Abstände zu den Grenzen des physikalischen Raumes als gut und kleine als schlecht bewertet, um somit Gains zu definieren, die schlussendlich Intensität und Dauer des eigentlichen Umlenkens festlegen. Auf diese Weise kann das System trainiert werden, die Benutzenden Personen möglichst lange innerhalb der freien Zone des physikalischen Raumes zu bewegen und die begehbare Distanz, bis ein Hard Reset notwendig ist optimieren [4].

Als Trainingsdaten kommen generierte, simulierte Userpfade zum Einsatz, die in ihrer Ursprungsrotation und Position variieren, um eine möglichst große Vielfalt von Szenarien zu gewährleisten. Evaluiert wird das System mit Probanden und Probandinnen und Probandinnen, welche eine virtuelle Umgebung erkunden und dabei von dem eingelernten Neuronalen Netz umgelenkt werden. Als Metrik, mit der die Effektivität ihres Ansatzes mit der von klassischen Steering Algorithmen verglichen wird, wird die gesamte zurückgelegte Distanz bis zu einem Hard Reset gemessen. Sind Benutzende also in der Lage, mit einem Reinforcement Learning Ansatz 400 Meter zurückzulegen, bevor sie wie bei Steer-to-Center nach nur 20 Metern einen Hard-Reset durchführen müssen, gilt der Reinforcement Learning Ansatz als effektiver [4].

Der Reinforcement Learning basierte Ansatz erwies sich als kaum effektiver als der Steer-to-Center Algorithmus, jedoch wesentlich effektiver als ein rein auf Hard-Resets basierender Ansatz, die verfassenden Personen berichten dabei jedoch von dem wesentlichen Unterschied, dass deren RL-Ansatz grundsätzlich immer eine Umleitung vornimmt, während Steer-to-Center nur dann umleitet, wenn sich Benutzende nicht direkt in Richtung des Zentrums des Raumes bewegen [4]. Als wichtigen Vorteil geben sie jedoch die Möglichkeit an, den Reinforcement Learning-Ansatz auf bestimmte Umgebungen beziehungsweise Pfade trainieren zu können, was in weniger allgemeinen Szenarien wie vorgefertigten Räumlichkeiten oder unkonventionell geformten Umgebungen hilfreicher ist als eine weniger spezifizierte Lösungsstrategie wie etwa Steer-to-Center [4]. Abschließend wird die Möglichkeit diskutiert, Reinforcement Learning nicht als alleinigen Ansatz für Redirected Walking einzusetzen, sondern es viel mehr als Unterstützung von Steering-Algorithmen anzuwenden, um deren Umleitungsverhalten zu optimieren [4].

## 5.2 Virtual Reality Induced Sickness - Vorhersage

Neben Verfahren, die direkt in die User Experience eingreifen, existieren auch Machine Learning-gestützte Verfahren zur Vorhersage von Virtual Reality Induced Sickness (VR Sickness), um Benutzenden Personen vorab Informationen über potenzielle Einschränkungen der User Experience zur Verfügung zu stellen bzw. diese zu reduzieren. Padmanaban et. al. beschreiben einen Machine Learning-basierten Ansatz zur Vorhersage von VR Sickness in stereoskopischen 360° Videos [1].

Unter VR Sickness werden Symptome wie Übelkeit, Schwindel, Kopfschmerzen und weitere verstanden, welche durch die Nutzung eines VR Systems verursacht werden, der Ursprung liegt dabei in einem Konflikt der Wahrnehmungsorgane - speziell im Bereich der visuellen Wahrnehmung können etwa die Augen dem Gehirn melden, dass man sich bewegt, während andere Sinnesorgane das Gegenteil vermitteln, der daraus entstehende Konflikt im Gehirn erzeugt dann entsprechende Symptome [1]. Die Abgrenzung zur Motion Sickness besteht darin, dass Motion Sickness durch tatsächliche Bewegungen wie etwa das Wippen eines Schiffes verursacht wird, VR Sickness jedoch auch bei statischer Lage eintritt [1].

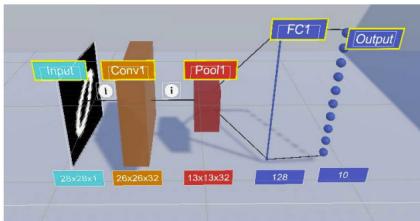
Um die Art und die Schwere der Symptome zu erfassen, werden entsprechende Fragebögen wie der von Kim et. al. entworfene Virtual Reality Sickness Questionnaire (VRSQ) [14] verwendet. Ziel war die Entwicklung eines allgemeingültigen Fragebogens zur Feststellung von eventuell auftretenden Symptomen bei Virtual Reality Systemen [14]. Die Vorhersage von VR Sickness wird von Padmanaban et. al. mit Hilfe von Decision Trees getroffen [1]. Die hierfür notwendigen

Features werden anhand des optischen Flusses ermittelt. Die von den Decision Trees getroffenen Vorhersagen wurden mit der erhobenen Ground Truth verglichen und zeigten sich grundsätzlich als zuverlässig, mit geringen Abweichungen. In zwei Beispielen wurde hingegen eine enorme Abweichung festgestellt, was sich als Fehler im Trainings Datensatz herausstellt hat [1]. Grundsätzlich lassen die Schlüsse von Padmanaban et. al. zu, dass Machine Learning Verfahren als denkbare Vorgehensweise in Bezug auf die Vorhersage von VR Sickness in entsprechenden 360° Videos in Frage kommt.

Ein ähnlicher Ansatz auf Basis von Deep Learning wird von Hell und Agyriou [2] verfolgt. Eine VR-Applikation, in der Benutzende Personen eine Achterbahnfahrt erleben dient als Grundlage für ein System, in welchem das Auftreten und die Intensität von VR Sickness durch Kamerabewegungen quantifiziert werden soll. Die verfassenden Personen berichten, dass der in Form einer Achterbahn vergleichsweise spezielle Use-Case auch als allgemeingültiger Ansatz in Frage kommen kann. Probanden und Probandinnen bewerten eine Achterbahn durch Eigenschaften wie Schwere ihrer Symptome und wie viel Spaß sie dabei erlebt haben. Kombiniert werden die erhobenen Daten durch die physikalischen Eigenschaften der Achterbahn selbst, etwa Beschleunigungen und Höchstgeschwindigkeiten oder zu überwindende Höhen. Durch den daraus entstehenden Datensatz werden neuronale Netze trainiert, die schlussendlich in der Lage sind, Aussagen über neue, von Benutzenden Personen eingereichte Achterbahnen im Bezug auf die zu erwartende VR Sickness zu treffen. Das Neuronale Netz war mit geringen Abweichungen in der Lage vorherzusagen, ob eine Achterbahn entsprechende Symptome hervorrufen kann [2].

### 5.3 Neuronale Netze - Visualisierung in VR

Über die bisher vorgestellten KI-Ansätze hinaus, welche überwiegend als Unterstützung für Virtual Reality selbst gesehen werden können, existieren auch Ansätze, die Virtual Reality nutzen, um KI-Verfahren zu visualisieren, speziell im Kontext der Lehre. So beschreiben Meissler et. al. [15] ein System, welches Convolutional Neural Networks in einer immersiven Virtual Reality Umgebung interaktiv erklären soll. Hierfür werden CNNs im Kontext der Bildverarbeitung mit dem Ziel, Bilder zu klassifizieren eingesetzt.



**Abbildung 3: Ausschnitt der VR Umgebung aus dem in [15] beschriebenen System. Sichtbar die interaktiven Schichten des visualisierten CNNs.**

Neben den individuellen Schichten der CNNs werden auch dessen Zwischenergebnisse und Pooling Layer visualisiert und den Anwendern mit 3D-Objekten präsentiert. Dabei wird sich auch die Tiefenwahrnehmung, welche durch Virtual Reality gewährleistet wird, zu Nutze gemacht, um den verfügbaren Raum voll auszunutzen. Eine durchgeführte Nutzerstudie kommt zu dem Ergebnis, dass Virtual Reality als brauchbares Werkzeug zur Vermittlung von Lehrinhalten mit dem Bezug zu Convolutional Neural Networks eingesetzt werden kann. So berichteten die Mehrheit der Probanden und Probandinnen, welche überwiegend Vorkenntnisse zu Virtual Reality und Convolutional Neural Networks besitzen, von positiven Eindrücken und deutlich

vermittelten Sachverhalten durch das System [15].

### 5.4 Neuronale Netze - KI-basiertes Rendering mit DeepFocus

Unter dem Begriff KI-Upscaling werden Vorgehensweisen verstanden, die ein Medium von einer geringen Auflösung auf eine höhere Auflösung umrechnen, um dadurch die Anzeigequalität zu erhöhen [16]. Hierfür werden oftmals Convolutional Neural Networks (CNN) eingesetzt [16].

Die Firma Oculus VR Inc., fortan abgekürzt Oculus, ist Entwickler der VR-Systeme Oculus Rift und Oculus Quest stellt ein neues VR-Headset vor, welches ein internes KI-System in Kombination mit neuartigen optischen Linsen verwendet, um die Darstellungsqualität zu erhöhen<sup>5</sup>. Das als HalfDome bezeichnete Produkt nutzt dabei die intern entwickelte OpenSource-Lösung DeepFocus<sup>6</sup>. In der 2018 von Xiao et. al. [17] veröffentlichten Arbeit wird DeepFocus als System beschrieben, welches mehrere Typen von VR-Headsets unterstützen soll. Bislang verfügbare VR-Headsets besitzen einen statischen Fokus, welcher durch neuartige Systeme wie varifocal bzw. multifocal headsets verbessert werden kann [17]. Unter einem multifocal display kann eine Anzeige mit verschiedenen Fokusebenen verstanden werden, während ein varifocal display eine Anzeige mit einer variablen Fokustiefe darstellt [17]. Das Ziel von DeepFocus ist hierbei die softwareseitige Erzeugung von realistischer Unschärfe am Rand des menschlichen Sichtfeldes [17]. Aktuelle VR-Systeme nutzen eine künstlich erzeugte Unschärfe, um die Tiefenwahrnehmung in Virtual Reality zu ermöglichen [17]. Mit DeepFocus und neuartigen Linsen sollen

<sup>5</sup><https://www.oculus.com/blog/introducing-deepfocus-the-ai-rendering-system-powering-half-dome/>

<sup>6</sup><https://www.oculus.com/blog/introducing-deepfocus-the-ai-rendering-system-powering-half-dome/>



Systeme wie HalfDome eine natürliche Unschärfe durch die vom menschlichen Auge durchgeführte Fokusanpassung ermöglichen, um die Tiefenwahrnehmung und generelle Anzeigequalität zu verbessern [17]. Multifocal displays benötigen eine Reihe von Bildern mit verschiedenen Fokustiefen, die einer rechnerisch aufwendigen, iterativen Zerlegung bedarf - DeepFocus realisiert diese Zerlegung mit Hilfe eines neuronalen Netzes [17]. Die Qualität der produzierten Bilddaten wird mit ähnlichen Ansätzen im Hinblick auf Höchstes Signal-Rausch-Verhältnis (eng. peak signal-to-noise ratio, PSNR), struktureller Ähnlichkeit (eng. structural similarity, SSIM) und der Laufzeit in Millisekunden verglichen.

	Akeley et al.	Narain et al.	Mercier et al.	DeepFocus
PSNR	28.81	35.20	27.56	43.45
SSIM	0.8680	0.9791	0.9296	0.9963

**Abbildung 4: Vergleich von DeepFocus mit anderen Ansätzen der Fokus-Dekomposition bezüglich Signal-Rausch-Verhältnis und struktureller Ähnlichkeit [17]**

Resolution	Narain et al.	Mercier et al.	DeepFocus
512 <sup>2</sup>	13.3	7.5	0.0154
1024 <sup>2</sup>	62.0	25.0	0.0503

**Abbildung 5: Laufzeitvergleich von DeepFocus und anderen Ansätzen mit zwei verschiedenen Auflösungen [17]**

Die von den verfassenden Personen veröffentlichten Ergebnisse lassen den Schluss zu, dass DeepFocus im Vergleich mit anderen Methoden zur Fokusdekomposition nicht nur präzisere Ergebnisse liefern kann, jedoch auch die Möglichkeit besitzt, diese mit einer wesentlich höheren Geschwindigkeit zu erzeugen, als andere existierende Lösungen.

## 6 Fazit

Die Arbeit führt die Begriffe Extended Reality und Künstliche Intelligenz im Kontext der

Zielsetzung ein und liefert einen Einblick in deren State of the Art. Die durchgeführte Literaturrecherche mit dem Ziel, einen Konsens zum Status Quo von KI in XR Anwendungen zu schaffen, lieferte eine enorme Zahl an potenziellen Arbeiten, wovon sich nur ein sehr kleiner Anteil hinsichtlich der verwendeten Suchbegriffe und Anforderungen als brauchbar erwiesen hat. So weist eine hohe Zahl der ursprünglich betrachteten Arbeiten entweder keinen Bezug zur Thematik auf, oder die Arbeiten wurden nicht in englischer beziehungsweise deutscher Sprache verfasst. Die vorgestellten Arbeiten zu Redirected Walking, Sickness Prediction und Visualisierung zeigen jedoch, dass KI-Verfahren ihren Weg in die Welt von Extended Reality gefunden haben und bereits erste Erfolge erzielen konnten. So zeigt die Arbeit von Strauss et. al. [4], dass ML-Verfahren wie hier Reinforcement Learning Ergebnisse zu Tage fördern, die bestehende Verfahren sogar übertreffen. Die dort gewonnenen Erkenntnisse lassen aber auch den Schluss zu, dass ML-Verfahren sich eher einer unterstützenden Rolle von bestehenden Ansätzen annehmen, anstatt sie gänzlich zu ersetzen. So kann zukünftig ein Redirected Walking Ansatz auf Basis von konventionellen Steering Algorithmen, die durch einen Reinforcement Learning Ansatz in Hinblick auf Präzision und Immersion augmentiert werden, denkbar sein. Über Redirected Walking hinaus kann durch die von [2] und [1] gewonnenen Erkenntnisse bezüglich der Vorhersage vom Auftreten der VR Sickness ein automatisches, KI-gestütztes Screening von Virtual Reality Inhalten vorgenommen werden, welches interessierte Benutzende vor dem potenziellen Auftreten und der Stärke von Symptomen in Kenntnis setzen kann. Durch die Arbeit von Meissen et. al. [15] wird jedoch auch ersichtlich, dass der umgekehrte Fall denkbar ist - Extended Reality wird in ihrer Arbeit eingesetzt, um Machine Learning Verfahren wie Convolutional Neural Networks interaktiv zu vermitteln. So

könnte sich Virtual Reality als legitimes Mittel zum Vermitteln von Inhalten in Bezug auf Künstliche Intelligenz etablieren. Im Hinblick auf das Rendering von Bilddaten bei Virtual Reality Systemen bietet das in Kapitel 5.4 vorgestellte System DeepFocus einen interessanten Einblick in die aktuelle Rolle von Neuronalen Netzen. Die von den verfassenden Personen vorgestellten Ergebnisse lassen den Schluss zu, dass ein KI-basierter Rendering Ansatz bald Einzug in kommerziell verfügbare Virtual Reality Systeme finden kann, um sowohl die Anzeigequalität zu erhöhen als auch die von Menschen empfundene Tiefenwahrnehmung zu steigern. In Kombination mit neuartigen Linsen mit mehreren beziehungsweise variablen Fokustiefen kann hier eine durchweg höhere Darstellungsqualität erwartet werden. Bei aufbauenden Arbeiten wäre eine präzisere Auswahl von Suchbegriffen bzw. eine weitere Eingrenzung eventuell hilfreich. So ist es schwierig, einen allgemein gültigen Status Quo auf die Gesamtheit von Extended Reality Anwendungen mit KI-Ansätzen zu definieren. Viel mehr wäre eine Konkretisierung denkbar, welche sich einzelner Ansätze aus der Welt des Machine Learnings, beispielsweise Neuronale Netze, Clustering-Verfahren oder Reinforcement Learning bedient, um einen zwar begrenzten, aber wesentlich präziseren Konsens zu schaffen. Die vorgestellten Arbeiten können hier einen Startpunkt für weitere, tiefgreifendere Recherchen in den jeweiligen Anwendungsgebieten darstellen.

## Literatur

- [1] N. Padmanaban, T. Ruban, V. Sitzmann, A. M. Norcia und G. Wetzstein, „Towards a Machine-Learning Approach for Sickness Prediction in 360° Stereoscopic Videos,“ *IEEE Transactions on Visualization and Computer Graphics*, Jg. 24, Nr. 4, S. 1594–1603, 2018. doi: 10.1109/TVCG.2018.2793560.
- [2] S. Hell und V. Argyriou, „Machine Learning Architectures to Predict Motion Sickness Using a Virtual Reality Rollercoaster Simulation Tool,“ in *2018 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, 2018, S. 153–156. doi: 10.1109/AIVR.2018.00032.
- [3] T. Nescher, Y.-Y. Huang und A. Kunz, „Planning redirection techniques for optimal free walking experience using model predictive control,“ in *2014 IEEE Symposium on 3D User Interfaces (3DUI)*, 2014, S. 111–118. doi: 10.1109/3DUI.2014.6798851.
- [4] R. R. Strauss, R. Ramanujan, A. Becker und T. C. Peck, „A Steering Algorithm for Redirected Walking Using Reinforcement Learning,“ *IEEE Transactions on Visualization and Computer Graphics*, Jg. 26, Nr. 5, S. 1955–1963, 2020. doi: 10.1109/TVCG.2020.2973060.
- [5] S. Martirosov und P. Kopeček, „Virtual Reality and its Influence on Training and Education - Literature Review,“ in Jan. 2017, S. 0708–0717, ISBN: 9783902734112. doi: 10.2507/28th.daaam.proceedings.100.
- [6] I. Rakkolainen, A. Farooq, J. Kangas u. a., „State of the Art in Extended Reality - Multimodal Interaction, Technical Report,“ in Jan. 2021, ISBN: 9783902734112. doi: 10.2507/28th.daaam.proceedings.100.
- [7] J. Molka-Danielsen, E. Prasolova-Førland, M. Fominykh und K. Lamb, „Use of a Collaborative Virtual Reality Simulation for Multi-Professional Training in Emergency Management Communications,“ in *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, 2018, S. 408–415. doi: 10.1109/TALE.2018.8615147.

- [8] A. Joshi, Machine Learning and Artificial Intelligence. Springer International Publishing, 2019, isbn: 9783030266226. Adresse: <https://books.google.de/books?id=ybyxDwAAQBAJ>.
- [9] P. Shengguang, „Overview of Meta-Reinforcement Learning Research,“ in 2020 2nd International Conference on Information Technology and Computer Application (ITCA), 2020, S. 54–57. doi: 10.1109/ITCA52113.2020.00019.
- [10] B. Zhang, „Tactical Decision System of Table Tennis Match based on C4.5 Decision Tree,“ in 2021 13th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), 2021, S. 632–635. doi: 10.1109/ICMTMA52658.2021.00146.
- [11] C. C. Aggarwal, „An Introduction to Neural Networks,“ in Neural Networks and Deep Learning: A Textbook. Cham: Springer International Publishing, 2018, S. 1–52, isbn: 978-3-319-94463-0. doi: 10.1007/978-3-319-94463-0\_1. Adresse: [https://doi.org/10.1007/978-3-319-94463-0\\_1](https://doi.org/10.1007/978-3-319-94463-0_1).
- [12] F. Bre, J. Gimenez und V. Fachinoti, „Prediction of wind pressure coefficients on building surfaces using Artificial Neural Networks,“ Energy and Buildings, Jg. 158, Nov. 2017. doi: 10.1016/j.enbuild.2017.11.045.
- [13] E. Jonathan, C. Roberts, S. Presentati-on, S. Razaque, Z. Kohn und M. Whit-ton, „Redirected Walking,“ Proceedings of Eurographics, Aug. 2001.
- [14] H. K. Kim, J. Park, Y. Choi und M. Choe, „Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment,“ Applied ergonomics, Jg. 69, S. 66–73, Mai 2018. doi: 10.1016/j.apergo.2017.12.016.
- [15] N. Meissler, A. Wohlan, N. Hochgeschwender und A. Schreiber, „Using Visualization of Convolutional Neural Networks in Virtual Reality for Machine Learning Newcomers,“ in 2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR), 2019, S. 152–1526. doi: 10.1109/AIVR46125.2019.00031.
- [16] W. Shi, J. Caballero, F. Huszár u. a., „Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network,“ in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, S. 1874–1883. doi: 10.1109/CVPR.2016.207.
- [17] L.Xiao, A.Kaplanyan, A.Fix, M.Chapman und D. Lanman, „DeepFocus: Learned Image Synthesis for Computational Displays,“ ACM Trans. Graph., Jg. 37, Nr. 6, Dez. 2018, issn: 0730-0301. doi: 10.1145/3272127.3275032. Adresse: <https://doi.org/10.1145/3272127.3275032>.



# Überblick über maschinelle Lernverfahren für Predictive Maintenance

Giuseppe Ferrera  
Hochschule Reutlingen

Giuseppe.Ferrera@Student.Reutlingen-University.de

## Abstract

Mit der Einführung und der Umsetzung der Industrie 4.0 Grundkonzepte entstanden in der Vergangenheit weitere Verfahren wie Predictive Maintenance. Dank dieser digitalen Transformation ist es möglich große Mengen an Daten zu sammeln und diese Daten für eine automatisierte Diagnose und Fehlererkennung zu nutzen und damit Ausfallzeiten von Maschinen zu prognostizieren und die Restnutzungsdauer einzelner Bauteile zu berechnen. Im Rahmen von Industrie 4.0 haben sich maschinelle Lernverfahren als ein vielversprechendes Werkzeug für die Umsetzung von Predictive Maintenance erwiesen. Im Rahmen dieser Arbeit wird ein Überblick über die eingesetzten maschinellen Lernverfahren gegeben und es wird dabei betrachtet, welche Aspekte zur Datenbeschaffenheit, dem Umfeld und dem erfolgreichen Betrieb zu beachten sind.

## CCS Concepts

- **Computing methodologies ~ Machine learning ~ Machine learning approaches**
- Computer systems organization ~ Embedded and cyber-physical systems ~ Sensors and actuators

## Keywords

Predictive Maintenance; Machine Learning; Industry 4.0; Remaining Useful Life; Defect Detection

## 1 Einleitung

Mit Industrie 4.0 wird die vierte industrielle Revolution bezeichnet. Der Begriff setzt dabei voraus, dass die vorherigen drei industriellen Revolutionen bereits abgeschlossen sind. Dies ist auch der Fall, denn die erste Revolution war der Übergang dazu die Handarbeit durch maschinelle Produktion zu ersetzen. Die zweite nutzte dies als Grundlage und basierte hauptsächlich auf die Etablierung von Arbeitsteilung, Standardisierung, Präzisionsfertigung und Fließbandarbeit. Die dritte Evolution basierte auf der Entwicklung der IT – Infrastruktur [17]. Die vierte industrielle Revolution hat nun das Ziel, digitale Fabriken durch die Verwendung von digitalen Repräsentationen der physikalischen Operationen zu schaffen. Oft werden diese Repräsentationen Cyber Physical Models (dt. Cyber-Physische-Modelle) oder Digital Twins (dt. Digitale Zwillinge)

---

Betreuer Hochschule: Prof. Dr. Cristobal Curio  
Hochschule Reutlingen  
Cristobal.Curio@Reutlingen-  
University.de

Informatics Inside Herbst 2021  
24. November 2021, Hochschule Reutlingen

genannt [4]. Die dadurch entstehende intelligente Fabrik bietet die Möglichkeit, den Kunden neue Dienstleistungen und Produkte anzubieten, deren Effizienz, Qualitätsstandards und Zuverlässigkeit höher sind als bisher. Unter den vielen neuen Möglichkeiten, die dadurch entstehen ist Predictive Maintenance (dt. vorausschauende Wartung, PM) die bekannteste. PM nutzt meist Sensordaten, um Anomalien (d. h. Abweichungen von den normalen Betriebsbedingungen) in Produktionsprozessen, Fertigungsanlagen und Produkten zu erkennen und zu diagnostizieren. Dies wird erreicht, indem das Auftreten eines anomalen Zustands charakterisiert wird um zukünftige Ausfälle zu prognostizieren und, indem die zukünftige Entwicklung des anomalen Zustands bis hin zum Ausfall vorhergesagt wird [3]. PM hat sich von seiner einfachsten Methode, der visuellen Inspektion, zu automatisierten Methoden entwickelt, bei denen Signalverarbeitungstechniken auf der Grundlage von Mustererkennung und maschinellen Lernverfahren eingesetzt werden. Dadurch können unter anderem der unnötige Austausch von Teilen vermieden sowie Produktionsstillstände durch geplante Wartungsintervalle reduziert werden [13]. Es gibt unterschiedliche maschinelle Lernverfahren, die für die Lösung des PM Problems eingesetzt werden. Je nach gewähltem Verfahren können die Leistung und die Vorteile unterschiedlich sein. Folglich zielt diese Arbeit darauf ab, die Forschungsbereiche der letzten fünf Jahre in Bezug auf PM und der eingesetzten maschinellen Lernverfahren zu untersuchen. Im Folgenden wird beschrieben, wie diese Arbeit aufgebaut ist: Zunächst werden in den Subkapiteln das Ziel sowie die angewandte Methodik dieser Arbeit beschrieben. Im zweiten Kapitel wird ein kurzer Überblick über PM und den zugrunde liegenden Instandhaltungsverfahren gegeben. Im dritten Kapitel werden Publikationen, die nach der beschriebenen Methodik im ersten Kapitel gefunden wurden, vorgestellt und Aspekte wie Einsatzgebiete, Datengrundlage und die Algorithmen näher betrachtet. Abschließend werden eine Diskussion und ein Ausblick auf

der Grundlage der in der Literatur durchgeführten Analyse von den Algorithmen für PM geführt.

### 1.1 Motivation und Ziele

Ziel dieser Arbeit ist es, einen systematischen Überblick über den aktuellen Stand der Technik von PM der letzten fünf Jahre zu geben. Dabei wird untersucht, in welcher Anwendungsdomäne das Forschungsfeld von PM den Fokus gelegt hat. Ein weiterer Punkt, der untersucht werden soll, ist die Datengrundlage, mit der die erzeugten Modelle trainiert werden. Außerdem soll ein Überblick über die eingesetzten Modelle gegeben werden. Zur Beantwortung dieser Fragen wurden in Tabelle 1 die Forschungsfragen definiert, welche durch die folgende Recherche beantwortet werden sollen.

**Tabelle 1: Forschungsfragen**

FF1	In welchen Bereichen wird PM am häufigsten eingesetzt?
FF2	Welche Art von Daten werden für die PM-Algorithmen verwendet?
FF3	Welche Algorithmen werden für PM am häufigsten verwendet?

### 1.2 Methoden

Zur Beantwortung der zuvor definierten Forschungsfragen wurde eine systematische Literaturrecherche durchgeführt. Für die Suchstrategie wurden wissenschaftliche Datenbanken sowie Schlüsselwörter definiert. In Anbetracht des technischen Rahmens dieser Arbeit wurden die Datenbanken der Association for Computing Machinery (ACM), dem Institute of Electrical and Electronics Engineers (IEEE) und Science Direct als die vielversprechendsten Datenquellen ermittelt und für diese Untersuchung ausgewählt. Zur Ermittlung von Publikationen wurden durch die Kombination von definierten Schlüsselwörtern mehrere Suchterme erzeugt und in den Datenbanken zur Suche genutzt. Die Suchterme wurden wie folgt erzeugt:

(predictive maintenance oder smart manufacturing oder intelligent device) und industry 4.0 und (iot) und (wear oder failure) und prediction und (algorithm oder model)

Außerdem wurden Einschlusskriterien definiert, die in Tabelle 2 gelistet sind. Nach dem Filtern der Einschlusskriterien wurde außerdem das Abstract der Publikationen durchgelesen. Durch das Abstract musste ersichtlich sein, dass es sich in der Arbeit um die maschinelle Vorhersage oder der Erkennung eines Defektes handelt. Nach Filterung der beschriebenen Kriterien wurde die Publikation vertieft untersucht, wie im Vorgehen in Abbildung 1 veranschaulicht.

**Tabelle 2: Einschlusskriterien**

Nr.	Beschreibung
K1	Es werden Publikationen zwischen 2016 – 2021 berücksichtigt
K2	Publikationen müssen auf Englisch oder Deutsch sein
K3	Es werden Fachartikel oder Konferenzmaterialien berücksichtigt
K4	Publikationen müssen sich mit der verbleibende Nutzungsdauer eines Bauteils oder der Erkennung eines Defektes einer Maschine beschäftigen

Nach den beschriebenen Kriterien wurden 10 Publikationen gefunden, welche in dieser Arbeit näher betrachtet werden.

## 2 Predictive Maintenance

PM ist eine Sammlung von Aktivitäten, welche Veränderungen im physischen Zustand der Ausrüstung erkennen und anhand dieser Veränderungen darauf schließen können, ob

Wartungsarbeiten durchzuführen sind. Ziel ist es, die Lebensdauer von Maschinen und Maschinenteilen zu maximieren, ohne, dass das Risiko eines Ausfalls erhöht wird [16]. Es sind Aktivitäten, die früher in ihrer rudimentärsten Version von erfahrenen Technikern ausgeführt wurden, indem Sie ihre Sinne nutzen, wie sehen, hören oder fühlen, um dadurch auf eine Veränderung des physischen Zustandes zu schließen. Was sich geändert hat, ist, dass nun unter PM der Einsatz von großen Datenmengen verstanden wird, die z. B. durch Sensoren erfasst werden und diese Daten die Aufgabe des Sehens, Riechens und des Tastens übernehmen. Die Einschätzung des Zustandes wird durch Algorithmen übernommen [14]. Diese großen Datenmengen stammen aus einem der Grundsätze der Industrie 4.0, die die traditionelle Fertigung in intelligente, mit Sensoren ausgestattete Fabriken umwandelt, in denen Technologie omnipräsent ist [4]. PM lässt sich in zwei Kategorien aufteilen, die wie folgt lauten [8]:

- **PM auf Basis von statistischer Auswertung:** Daten von Ausfällen und die daraus gewonnen Informationen werden zur Entwicklung statischer Modelle herangezogen und ermöglichen so die Entwicklung einer vorbeugenden Wartungspolitik.
- **PM auf Basis von Zuständen:** Kontinuierliche oder periodische Überwachung des Zustands der Ausrüstung, um Anzeichen eines Ausfalls zu erkennen und Wartungsentscheidungen zu treffen.

Die kostengünstigste Instandhaltungstechnik ist die zustandsorientierte Instandhaltung. Die zustandsorientierte Instandhaltung bietet die Möglichkeit, die Verfügbarkeit von Ma-



**Abbildung 1: Vorgehen Literaturrecherche**

schinen zu optimieren und die Instandhaltungskosten deutlich zu senken [12]. Zustandsorientierte Instandhaltungstechniken bieten eine Bewertung des Systemzustands auf der Grundlage von Daten, die durch kontinuierliche Systemüberwachung gesammelt wurden. Ziel ist es, das notwendige Instandhaltungsprogramm zu bestimmen, bevor ein erwarteter Ausfall eintritt. Daher zielen Instandhaltungsstrategien darauf ab, die Kosten zu minimieren, indem sie die Betriebszuverlässigkeit verbessern und die Schwere und Anzahl von Systemausfällen verringern [8].

Es gibt drei unterschiedliche Ansätze, wie man sich dem Problem von PM nähern kann die wie folgt lauten [13]:

- **Physikalisch modellbasierter Ansatz:** Bei diesem Ansatz werden die physikalischen Modelle des Bauteils für die Schätzung seines gesunden Zustands und die Vorhersage seiner Verschlechterung verwendet. Derzeit bilden diese Modelle die Grundlage für die Entwicklung von digitalen Zwillingen. Die Entwicklung physikalisch basierter Modelle ist jedoch nicht immer praktikabel, da sie sehr teuer sind. Darüber hinaus berücksichtigen diese Modelle die Auswirkungen externer Bedingungen oft nicht vollständig und beruhen auf Parametern, die schwer zu schätzen sind [3].
- **Datengetriebener Ansatz:** Dieser Ansatz verwendet historische Betriebsdaten, um ein Modell zu trainieren. Bei Mangel an historischen Daten können diese in Experimenten vor Ort gesammelt werden. Dieser Ansatz ist auch als Data-Mining oder als Ansatz des maschinellen Lernens bekannt [13].
- **Hybrider Ansatz:** Kombination von Merkmalen des datengesteuerten Ansatzes und des modellbasierten Ansatzes [11].

## 2.1 Instandhaltungspolitik

In der Instandhaltung gibt es drei Kategorien für Maßnahmen, und PM kann als die neueste und am meisten gewünschte Strategie in der Entwicklung des Instandhaltungsmanagements angesehen werden. Im Vergleich zu den traditionellen und immer noch angewandten Instandhaltungsstrategien bietet sie in jeder Hinsicht die beste Leistung, ist aber dafür schwer zu implementieren, da es eine sorgfältige Planung von Hardware, Software und Personal erfordert [14].

Die einfachste Instandhaltungsstrategie ist die ungeplante Instandhaltung (auch „run to failure“ genannt), bei der keine Instandhaltung durchgeführt wird, bis ein Maschinenausfall eintritt. In dieser Situation kann die Lebensdauer einer Maschinenkomponente bis zum Ausfall gestreckt werden, aber ungeplante Ausfallzeiten sind unvermeidlich. Zwar werden damit Kosten gespart, indem noch funktionstüchtige Teile nicht ausgetauscht werden, sondern weiter Verwendung finden jedoch steht dies im Kontrast mit einem unerwarteten und teuren Ausfall von Produktionen [10].

Bei der vorbeugenden Wartung, die in der Industrie am weitesten verbreitet ist, werden die Komponenten in regelmäßigen Abständen inspiziert und gewartet, um unerwartete Maschinenausfälle zu vermeiden. Komponenten werden erfahrungsgemäß in gleichen Zeitabständen ausgetauscht. Dadurch kann es vorkommen, dass Teile trotz Austausch noch viel länger eingesetzt werden könnten. Es wird also die maximale Lebensdauer des Bauteils nicht effektiv genutzt. Es kommt zu regelmäßigen Ausfallzeiten und mehr Teile müssen gelagert werden [10].

Während die ungeplante und die vorbeugende Instandhaltung einen Kompromiss darstellen, ist PM eine vielversprechende Methode, die diesen Kompromiss beseitigen und gleichzeitig die Lebensdauer und Betriebszeit der Komponenten maximieren kann.

## 2.2 Benefits

Die Anwendung von PM in Produktionsumgebungen bringt einige Vorteile mit sich. Auf der höchsten Ebene wird die Produktivität durch die Verringerung von Systemausfällen erhöht, was auch die Profitabilität erhöht [18]. Dies wird erreicht, indem man fehlerhafte Geräte oder Bauteile erkennt und die Fehler behebt. Außerdem wird die Lebenszeit von Bauteilen voll ausgeschöpft, was Lagerkosten und Kosten, die durch Ausfälle anfallen, reduziert. Neben den finanziellen Vorteilen liefert es für das Instandhaltungsmanagement mehr Transparenz, da sie einen besseren Überblick über den Gesundheitszustand ihrer Produktionsanlagen bekommen [4]. Die Beziehung zwischen den Instandhaltungsmethoden und den Benefits sowie mögliche Nachteile werden in Abbildung 2 dargestellt.

## 3 Stand der Technik

Ausgehend von den im vorigen Kapitel definierten Suchkriterien wurden 10 Publikationen gefunden, die mit den Kriterien übereinstimmen. In diesem Kapitel und in den folgenden Kapiteln werden diese Publikationen im Detail vorgestellt und Aspekte wie die Einsatzgebiete, die Datengrundlage und die verwendeten Algorithmen beleuchtet. Eine Übersicht hierzu bietet Tabelle 3, welche die soeben genannten Punkte kurz zusammenfasst.

Wie bereits beschrieben gibt es drei Ansätze, um PM umzusetzen. Nach der vorangegan-

genen Recherche wurden allerdings nur datengetriebene Ansätze und ein hybrider Ansatz gefunden.

Bei der Verwendung datengesteuerter Modelle können die Daten entweder durch bereits vorhandene und in eine Maschine eingebaute Sensoren oder durch die Nachrüstung von geeigneten Sensoren erfasst werden. Letzteres wurde durch die Arbeit von Hesser und Markert [5] erreicht, in der sie einen Ansatz zur Nachrüstung einer alten CNC-Fräsmaschine vorstellen. Die Maschine war mehr als 15 Jahre in Betrieb und wurde mit einer Sensorplattform als eingebettetes programmierbares Sensorsystem nachgerüstet, um die Datenerfassung und kabellose Kommunikation für die nachgerüstete Maschine zu ermöglichen. Der Defekt von Interesse ist der Verschleiß einer Fräse, die zum Fräsen der Werkstücke verwendet wird. Dazu wurden Datensätze erstellt, indem die Fräse bis zum Ausfall betrieben und die Daten anschließend gelabelt wurden. Als Sensor dient ein Beschleunigungssensor, der Vibrationsdaten liefert. Die Zeitreihe wurde in drei Merkmale umgewandelt, nämlich in die Standardabweichung, den quadratischen Mittelwert und die Fläche des Signals. Dies ist eine gängige Praxis, da die Anzahl der Abtastungen und die Komplexität der Rohsignale zu hoch sind und das Training und die Validierung der Modelle beeinträchtigen. Anschließend wurde ein neuronales Netz mit diesen drei Features angelernt, das zwischen

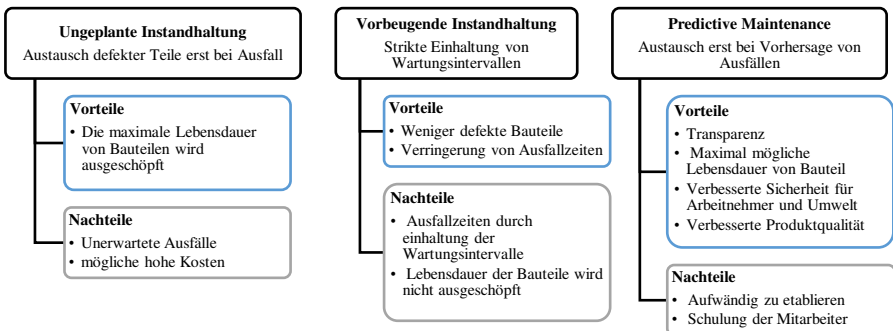


Abbildung 2: Benefits



neuer Fräse und verschlissener Fräse unterscheiden kann. Ein weiterer Beitrag in diesem Bereich liefern Jain et al. [6] mit ihrer Arbeit. Sie stellen ebenfalls einen datengetriebenen Ansatz, der neben einem Beschleunigungssensor noch einen Kraftsensor und ein Mikrofon für die akustische Emission als Datenquellen verwendet. Hier wird ebenfalls die Abnutzung der Fräse gemessen, indem durch Tests die Fräse bis zum endgültigen Defekt beansprucht wird. Die hochdimensionalen Signale wurden in unterschiedliche statistische Features umgewandelt und die Daten mit drei unterschiedlichen Stufen von Verschleiß gelabelt. Anhand dieser Daten wurden unterschiedliche Klassifikatoren angelehrt und die Autoren stellen fest, dass sich in Abhängigkeit der Datenquellen, unterschiedliche Klassifikatoren eignen. So liefert zum Beispiel der Naive Bayes (NB) Klassifikator die besten Ergebnisse für die Daten des Kraftsensors, die Logistische Regression (LR) die besten Ergebnisse für die Vibrationsdaten und der Random Forest (RF) Klassifikator die besten Ergebnisse für die akustischen Daten. Abhängig vom Ergebnis dieser Klassifikatoren werden die Daten weiter an ein Long short-term memory (LSTM) Netzwerk weitergeleitet, das für die Prognose der verbleibenden Nutzungsdauer zuständig ist. Dieses Netzwerk wurde mit den gleichen Daten antrainiert. Lee et al. [10] stellen eine weitere Publikation im Anwendungsfall von CNC-Fräsmaschinen vor. Ähnlich wie in der vorherigen wird der Verschleiß und die verbleibende Lebensdauer vorhergesagt. In ihrer Methode werden Daten zu Spannung, Vibration und Ton in Experimenten gesammelt. Im Unterschied zur vorherigen Publikation möchten Sie auch eine Funktionsstörung eines Kugellagers erkennen. Für beide Bauteile werden die Daten in drei Klassen gelabelt (Normal, Warnung, Fehler). Für die Fräse wird eine Support Vector Maschine (SVM) antrainiert. Für das Kugellager ein Convolutional Neural Network (CNN) und ein Recurrent Neural Network (RNN). Die verbleibende Lebensdauer leiten die Autoren von den drei Klassen ab. Somit ergibt die Klassifikation des Zustands „Normal“ eine

Funktionstauglichkeit des Bauteils zwischen 100 %-66 %. Die Anwendung der Algorithmen auf die aus den Versuchsdaten extrahierten Merkmale zeigt, dass die Bedingungen effektiv überwacht werden, um die Verschlechterung der Werkzeuge zu bewerten. Traini et al. [15] arbeiten in ihrer Arbeit mit den gleichen Sensoren und stellen eine weitere Methode vor, um den Verschleiß von Fräsen zu erkennen. Ihre Zielvariablen sind den Verschleiß in die Klassen sicher und abgenutzt zu klassifizieren und die verbleibende Nutzungsdauer in Form von verbleibenden Durchläufen zu schätzen. Für die Klassifikationen erreichen die Autoren mit einem Boosted Decision Tree (BDT) Modell die besten Ergebnisse und mit einem Neural Network Regression (NNR) Modell die besten Ergebnisse für die Schätzung der verbleibenden Nutzungsdauer.

Die letzte Methode, die den Anwendungsfall von CNC-Fräsmaschinen behandelt, ist ein hybrider Ansatz, der von Luo et al. [11] vorgestellt wird. In diesem Ansatz liefert ein Kraft- und ein Beschleunigungssensor sowie ein Mikrofon die notwendigen Daten für das datengetriebene RF Modell. Von der Fräse gibt es ein mathematisches Modell, welches durch Eigenschaften der physikalischen Maschine Daten zum Systemstatus liefert. Die Beobachtungsdaten der Sensoren sowie die Daten des Systemmodells werden durch ein Partikelfilteralgorithmus zusammengeführt. Durch die Nutzung beider Datenquellen erzielen die Autoren bessere Ergebnisse für die Schätzung der verbleibenden Lebensdauer der Fräse, als wenn nur Simulationsdaten oder Beobachtungsdaten verwendet werden [11].

Methoden für den Einsatz von PM in der Halbleiterfertigung werden in den Arbeiten [2] sowie [1] vorgestellt. Der vorgeschlagene Ansatz von Cao et al. basiert auf dem kombinierten Einsatz von Fuzzy-Clustering und Semantik, wobei Fuzzy-Clustering-Techniken eingesetzt werden, um die Gefährlichkeit von Fehlern auf der Grundlage historischer Maschinendaten zu erlernen und regelbasierte Ontologien die Ergebnisse des

**Tabelle 3: Übersicht der Algorithmen**

Quelle	Zusammenfassung	Kategorie	Einsatzgebiet	Variablen	Bester Algorithmus
[2]	<ul style="list-style-type: none"> <li>Erkennung von Fehlersequenzen</li> <li>Klassifikation von drei kritischen Fehlerklassen</li> <li>Klassifikation Fehler und vorhersage RUL</li> </ul>	Datengetrieben	Halbleiterherstellung	Numerische Sensordaten, nicht weiter bekannt	Sequential Pattern Mining, Fuzzy C Means Clustering
[5]	<ul style="list-style-type: none"> <li>Nachrüstung alter Maschine</li> <li>Erzeugung von Daten</li> <li>Klassifikation von Abnutzung der Fräse</li> </ul>	Datengetrieben	CNC-Fräsmaschine	Vibration	Artificial Neural Network
[6]	<ul style="list-style-type: none"> <li>Erzeugung von Daten durch „Run-To-Failure“ tests</li> <li>Umwandlung von Signalen in statistische Features</li> <li>Klassifikation der Abnutzung in drei Klassen</li> <li>Trainieren des LSTM zur Berechnung der RUL</li> </ul>	Datengetrieben	CNC-Fräsmaschine	Kraft, Vibration, Schall	Logistische Regression, Naiver Bayes Klassifikator, Random Forest Klassifikator, Long-short-term-memory
[10]	<ul style="list-style-type: none"> <li>Klassifikation Abnutzungszustand Fräse sowie Kugellager</li> <li>Erzeugung der Daten durch „Run-To-Failure“ tests</li> <li>Umwandlung von Signalen in statistische Features</li> </ul>	Datengetrieben	CNC-Fräsmaschine	Spannung, Vibration, Schall	Support-Vector-Maschine, Recurrent-Neural-Netzwerk
[15]	<ul style="list-style-type: none"> <li>Klassifikation von Abnutzung in zwei Klassen</li> <li>Schätzung der RUL</li> </ul>	Datengetrieben	CNC-Fräsmaschine	Spannung, Vibration, Schall	Boosted Decision Trees, Neural Network Regression
[11]	<ul style="list-style-type: none"> <li>Datengetriebener Ansatz zur Schätzung der RUL</li> <li>Modellbasierter Ansatz zur Schätzung der RUL</li> </ul>	Hybrid	CNC-Fräsmaschine	Kraft, Vibration, Schall	Random Forest
[7]	<ul style="list-style-type: none"> <li>Cloudbasiertes System zur Überwachung Zustand Motor</li> <li>Feature Selektion durch PCA</li> <li>Klassifikation durch Konfidenzniveaus</li> </ul>	Datengetrieben	Motor	Vibration	Principal Component Analysis, Hotelling T2
[1]	<ul style="list-style-type: none"> <li>Berechnung der Time To Failure</li> <li>Klassifikation der Fehlerursachen</li> </ul>	Datengetrieben	Halbleiterfertigung	Domänen Spezifische Daten (Fehlerprotokoll, Zeitstempel)	Support-Vector-Maschine, Bayessches Netz
[9]	<ul style="list-style-type: none"> <li>Vorhersage von RUL anhand von ERP Daten</li> <li>Klassifikation der Fehlerbereiche</li> </ul>	Datengetrieben	Verpackungsroboter	ERP – Daten	Multilayer perceptron
[19]	<ul style="list-style-type: none"> <li>Mapping von Fehlerevents mit verfügbaren Daten</li> <li>Erweiterung Datensatz mit Zeit-Bis-Ausfall Zeiten</li> <li>Einteilung der Zeit-Bis-Ausfall Zeiten in diskrete Klassen</li> <li>Klassifikation Zeit-Bis-Ausfall Zeiten</li> </ul>	Datengetrieben	Hydraulische Presse	Sensordaten, Fehlerprotokoll, Daten nicht näher beschrieben	Feed-Forward-Neural-Network

Fuzzy-Clustering nutzen, um den Zeitpunkt von Fehlern und deren Gefährlichkeit vorherzusagen. In [1] setzen die Autoren SVM

ein, um die verbleibende Lebensdauer zu berechnen und BN um die Ursache des Defekts anhand der Daten zu klassifizieren. Die

Daten für ihre Algorithmen stammen aus dem Betrieb der Anlage über mehrere Jahre und sind angereichert durch Fehlerprotokolle. Die Modalität der Daten wird in ihrer Arbeit nicht näher beschrieben.

In der Arbeit von Joung et al. [7] wird ein cloudbasiertes System zur Überwachung des Motorzustands implementiert, um abnormale Betriebszustände von Motoren zu erkennen. Als Datenquelle werden die Daten eines Beschleunigungssensors genutzt. Die T2-Statistik von Hostelling und die Statistik des quadrierten Vorhersagefehlers (SPE) werden dann angewandt, um Kriterien für abnormale Betriebszustände anhand der gesammelten Vibrationsdaten zu ermitteln.

Im Hinblick auf die Datengrundlage unterscheidet sich die Arbeit von Koca et al. [9] darin, dass statt Sensordaten Enterprise Resource Planning (ERP)-Daten genutzt wurden, um ein MLP Modell zu trainieren und dadurch Anlagenausfälle vorherzusagen. Die Daten bestehen im Wesentlichen aus den

von der Produktionslinie gesammelten Fehlermeldungen. Außerdem sind sie mit ihrer Methode in der Lage, Vorhersagen über den Ausfallort in Ihrer Anlage zu treffen.

### 3.1 Einsatzgebiete

Unter den zehn gefundenen Publikationen befassten sich fünf Publikationen mit der Vorhersage von Defekten beim Betrieb von CNC-Fräsmaschinen. Beim Fräsen trägt ein rotierende Fräse Material von einem Werkstück ab, um die gewünschte Form zu erhalten. Durch den Kontakt der Fräse mit dem Werkstück verschleißt die Fräse durch die Erzeugung von Wärme und Spannung während des Prozesses. Folglich verschlechtert sich die Leistung des Fräsaufsatzes, was sich auf die Qualität des Werkstücks auswirkt [10]. Alle fünf haben sich dabei mit der Abnutzung des Fräsaufsatzes beschäftigt und eine Publikation befasst sich zusätzlich mit dem Defekt eines eingesetzten Kugellagers [10].

Das Einsatzgebiet der Halbleiterfertigung war in der Recherche zwei Mal vertreten. In

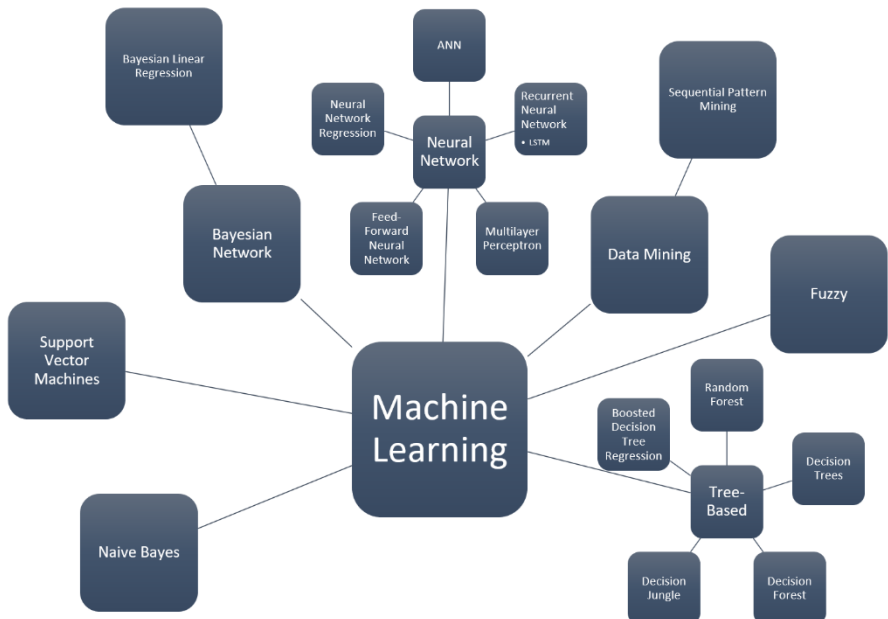


Abbildung 3: Taxonomie der Algorithmen

der Halbleiterherstellung geht es darum, vordefinierte Muster auf das Substrat zu übertragen, um bestimmte physikalische Eigenschaften des späteren elektrischen Bauteils zu erhalten. Moderne Halbleiterfertigungsanlagen zeichnen sich durch einen hohen Automatisierungsgrad aus, der es ihnen ermöglicht, ganze Chargen von Wafern mit unterschiedlichen Rezepten zu verarbeiten. Aufgrund dieses hohen Automatisierungsgrads und des Bedarfs an Reinraumwartung, einschließlich der Notwendigkeit, die Kammer zu reinigen und nach jedem Vorgang Testzyklen durchzuführen, bestehen hohe Anforderungen an die Wartungsplanung, um Zeit und Ressourcen zu sparen [1].

Weitere Einsatzgebiete, in denen nach den gefundenen Publikationen kein Schwerpunkt finden lässt, waren Verpackungsroboter, Motoren und hydraulische Pressen.

### 3.2 *Datengrundlage*

Zu den Daten ist festzustellen, dass in keiner der vorgestellten Veröffentlichungen Techniken wie Datenaugmentierung oder Simulationen verwendet werden, um Daten für ihre datengetriebenen Ansätze zu generieren. Die Daten sind entweder bereits in der realen Umgebung vorhanden [1, 2, 7, 9, 11, 19], werden durch "run to failure"-Experimente erzeugt [5, 6, 15] oder die Methoden werden an bereits verfügbaren Datensätzen für die Forschung zum maschinellen Lernen bewertet [10]. Oft werden Sensoren eingesetzt, um messbare Eigenschaften für den Betrieb der Maschinen zu erfassen. Aus der Recherche geht hervor, dass Beschleunigungssensoren für die Messung von Vibration, Kraftsensoren für die Messung von mechanischer Kraft sowie Mikrofone für die Messung von Schall bzw. der akustischen Emission genutzt werden. Diese Sensoren können entweder bereits in der Maschine verbaut sein oder nachgerüstet werden. Die Arbeit von Koca et al. [9] hat auch gezeigt, dass nicht nur Sensordaten für PM geeignet sind, sondern auch prozessuale Daten, die zum Beispiel aus einem ERP- System stammen.

### 3.3 *Algorithmen*

Um einen Überblick über die Algorithmen-Landschaft zu erhalten, wurde in Abbildung 3 eine Taxonomie erstellt, welche die Verbindung zwischen verschiedenen Klassen von Algorithmen aufzeigt. PM basierend auf maschinellen Lernverfahren kann in zwei Hauptklassen eingeteilt werden: (1) überwachtes Lernen und (2) unüberwachtes Lernen. Algorithmen des überwachten Lernens werden verwendet, wenn Informationen über das Auftreten von Ausfällen vorhanden sind. Unüberwachte Lernverfahren werden eingesetzt, wenn Letzteres nicht vorhanden ist, jedoch Logistik- und/oder Prozessinformationen. Wenn die Datenlage es ermöglicht, sind überwachte Lösungen vorzuziehen. Je nach Ausgang des Datensatzes sind zwei Klassen von überwachten Problemen möglich: (1) Regressionsprobleme (wenn der Ausgang kontinuierliche Werte annimmt) und (2) Klassifikationsprobleme (wenn der Ausgang kategorische Werte annimmt) [13]. Für die Klassifikation und Regression wurden in den gefundenen Veröffentlichungen ausschließlich überwachte Verfahren verwendet.

### 4 *Diskussion und Ausblick*

In Anbetracht der Tatsache, dass die meisten der gefundenen Ergebnisse auf Daten beruhen, die in Experimenten mit "run-to-failure"-Experimente erzeugt wurden, kann festgestellt werden, dass es immer noch an annotierten Daten zur Lösung des Problems mangelt. Durch die Erstellung solcher Daten kann zwar bewiesen werden, dass PM möglich ist, aber die erstellten Lösungen sind bei der Erkennung von anomalem Verhalten sehr begrenzt. Das liegt vor allem daran, dass die Modelle für ein bestimmtes Fehlermuster trainiert wurden, z. B. für ein fehlerhaftes Fräswerkzeug. In der Praxis beruhen die auf PM untersuchten Maschinen nicht nur auf einem einzigen Teil, sondern setzen sich aus anderen mechanischen und elektrischen Teilen zusammen. In den Studien fehlte eine Bewertung, die zeigt, wie die Algorithmen funktionieren, wenn die aufgezeichneten

Signale nicht nur den besten Fall aufzeichnen, in dem ein Teil ausfällt, sondern auch Signale von Teilen, die fehlerhaft sind und ebenfalls zur Maschine gehören. Um das Problem der fehlenden Daten zu überwinden, sollten Wartungsstrategien in Betracht gezogen werden, die erworbenes Wissen durch Fehler erhalten. Dieses Wissen sollte so erhalten werden, dass es mit Sensordaten verbunden werden kann. Wenn diese Praxis im täglichen Betrieb angewendet wird, können auf lange Sicht sehr vertrauliche Daten erzeugt werden, für die nur sehr wenig Feature Engineering erforderlich ist. Es sollte auch in Betracht gezogen werden, sich nicht nur auf Sensordaten zu verlassen. Die Arbeit von Koca et al. [9] hat gezeigt, dass allein durch die Integration von Domänenwissen ERP-Daten zur Vorhersage von Fehlern genutzt werden können. Die Kombination beider Strategien könnte zu einem robusteren System führen, das auch auf Veränderungen in der Umgebung reagieren kann.

## 5 Fazit

In dieser Arbeit wurde versucht, den aktuellen Stand der Technik von PM zusammenzustellen. Hierfür wurde zuerst PM und die unterschiedlichen zugrunde liegenden Instandhaltungsstrategien sowie die Vor- und Nachteile dieser Strategien zusammengefasst. Die durch die Suchkriterien gefundenen Publikationen wurden anschließend näher betrachtet und im Hinblick auf die Einsatzgebiete, Datengrundlage und die Algorithmen analysiert. Hierdurch wurden die Forschungsfragen FF1-FF3 beantwortet. In Anbetracht der Ergebnisse dieser Überprüfung werden beim aktuellen Stand der Technik Algorithmen des maschinellen Lernens für das PM-Problem eingesetzt. Angesichts der Anzahl der gefundenen Veröffentlichungen über Fräsmaschinen gibt es einen klaren Konsens über die Datenquellen, die zur Erkennung von Defekten und der verbleibenden Nutzungsdauer benötigt werden. Die in der Literatur verwendeten Algorithmen sind meist unterschiedlich für das Problem. Für die Vorhersage von Defekten werden Klassifizierungsalgorithmen eingesetzt, für die Ermittlung

der Restnutzungsdauer werden überwiegend Regressionsalgorithmen verwendet.

## Literaturverzeichnis

- [1] Biebl, F., Glawar, R., Jalali, A., Ansari, F., Haslhofer, B., Boer, P. de, and Sih, W. A conceptual model to enable prescriptive maintenance for etching equipment in semiconductor manufacturing. In *Procedia CIRP*, 64–69. DOI=10.1016/j.procir.2020.05.012.
- [2] Cao, Q., Samet, A., Zanni-Merk, C., Beuvron, F. B. de, and Reich, C. An Ontology-based Approach for Failure Classification in Predictive Maintenance Using Fuzzy C-means and SWRL Rules. In *Procedia Computer Science*, 630–639. DOI=10.1016/j.procs.2019.09.218.
- [3] Compare, M., Baraldi, P., and Zio, E. Challenges to IoT-Enabled Predictive Maintenance for Industry 4.0. In *IEEE Internet of Things Journal*, 4585–4597. DOI=10.1109/JIOT.2019.2957029.
- [4] Dalzochio, J., Kunst, R., Pignaton, E., Binotto, A., Sanyal, S., Favilla, J., and Barbosa, J. Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges. In *Computers in Industry*, 103298. DOI=10.1016/j.com-pind.2020.103298.
- [5] Hesser, D. F. and Markert, B. Tool wear monitoring of a retrofitted CNC milling machine using artificial neural networks. In *Manufacturing letters*, 1–4. DOI=10.1016/j.mfglet.2018.11.001.
- [6] Jain, A. K., Chouksey, P., Parlikad, A. K., and Lad, B. K. Distributed Diagnostics, Prognostics and Maintenance Planning: Realizing Industry 4.0. In *IFAC-PapersOnLine*, 354–359. DOI=10.1016/j.ifa-col.2020.11.057.
- [7] Joung, B. G., Lee, W. J., Huang, A., and Sutherland, J. W. Development

- and Application of a Method for Real Time Motor Fault Detection. In *Procedia Manufacturing*, 94–98. DOI=10.1016/j.promfg.2020.07.002.
- [8] Kiangala, K. S. and Wang, Z. Initiating predictive maintenance for a conveyor motor in a bottling plant using industry 4.0 concepts. In *The International Journal of Advanced Manufacturing Technology*, 3251–3271. DOI=10.1007/s00170-018-2093-8.
- [9] Koca, O., Kaymakci, O. T., and Mercimek, M. 2020. Advanced Predictive Maintenance with Machine Learning Failure Estimation in Industrial Packaging Robots. In *2020 International Conference on Development and Application Systems (DAS). 15th edition, May 21-23, 2020, Suceava, Romania : conference proceedings*. IEEE, Piscataway, NJ, 1–6. DOI=10.1109/DAS49615.2020.9108913.
- [10] Lee, W. J., Wu, H., Yun, H., Kim, H., Jun, M. B., and Sutherland, J. W. Predictive Maintenance of Machine Tool Systems Using Artificial Intelligence Techniques Applied to Machine Condition Data. In *Procedia CIRP*, 506–511. DOI=10.1016/j.procir.2018.12.019.
- [11] Luo, W., Hu, T., Ye, Y., Zhang, C., and Wei, Y. A hybrid predictive maintenance approach for CNC machine tool driven by Digital Twin. In *Robotics and Computer-Integrated Manufacturing*, 101974. DOI=10.1016/j.rcim.2020.101974.
- [12] Mustakerov, I. and Borissova, D. 2013. An intelligent approach to optimal predictive maintenance strategy defining. In *2013 IEEE I-NISTA*, 1–5.
- [13] Paolanti, M., Romeo, L., Felicetti, A., Mancini, A., Frontoni, E., and Loncarski, J. 2018. Machine learning approach for predictive maintenance in industry 4.0. In *2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, 1–6. DOI=10.1109/MESA.2018.8449150.
- [14] Selcuk, S. Predictive maintenance, its implementation and latest trends. In *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 1670–1679. DOI=10.1177/0954405415601640.
- [15] Traini, E., Bruno, G., D’Antonio, G., and Lombardi, F. Machine Learning Framework for Predictive Maintenance in Milling. In *IFAC-PapersOnLine*, 177–182. DOI=10.1016/j.ifacol.2019.11.172.
- [16] Wang, K. Intelligent predictive maintenance (IPdM) system-Industry 4.0 scenario. In *WIT Transactions on Engineering Sciences*, 259–268.
- [17] Wang, S., Wan, J., Di Li, and Zhang, C. Implementing Smart Factory of Industrie 4.0: An Outlook. In *International Journal of Distributed Sensor Networks*, 3159805. DOI=10.1155/2016/3159805.
- [18] Zhai, S., Gehring, B., and Reinhart, G. Enabling predictive maintenance integrated production scheduling by operation-specific health prognostics with generative deep learning. In *Journal of Manufacturing Systems*. DOI=10.1016/j.jmsy.2021.02.006.
- [19] Züfle, M., Agne, J., Grohmann, J., Dörtoluk, I., and Kounev, S. 2021. A Predictive Maintenance Methodology: Predicting the Time-to-Failure of Machines in Industry 4.0. In *2021 IEEE 19th International Conference on Industrial Informatics (INDIN)*, 1–8. DOI=10.1109/INDIN45523.2021.9557387.



©2021 Giuseppe Ferrera. Lizenznehmer Hochschule Reutlingen, Deutschland. Dieser Artikel ist ein Open-Access-Artikel unter den Bedingungen und Konditionen der Creative Commons Attribution (CC BY)-Lizenz. <http://creativecommons.org/licenses/by/4.0/>

# Vorhersagen von Aktienkursen basierend auf Regressionsmodellen und Sentiment-Analysen

Maximilian Rabus  
Hochschule Reutlingen

Maximilian.Rabus@Student.Reutlingen-University.de

## Abstract

Über die letzten Jahrzehnte gab es vor allem durch die immer größere Verfügbarkeit von Tools und Rechenleistung für Machine-Learning-Modelle einen großen Anstieg von Ansätzen, die sich mit dem Vorhersagen von Aktienkursen und -märkten beschäftigen. Im Rahmen dieser Arbeit sollen auf Basis dieser Ansätze aus historischen Daten des Internetforums Reddit verschiedenste Kennzahlen identifiziert werden, die mit der Entwicklung eines Aktienpreises korrelieren. Anschließend werden diese Daten und deren Kennzahlen genutzt, um ein Regressionsmodell zu trainieren und eine Aussage über dessen Genauigkeit zu treffen. Die Ergebnisse indizieren durch Kennzahlen wie einer Genauigkeit von 74,9%, dass durch die Auswahl geeigneter Kennzahlen in einem passenden Umfeld Vorhersagen über die Entwicklung eines Aktienkurses getroffen werden können.

## CCS Concepts

Computing methodologies → Modeling and simulation → Model development and analysis

---

Betreuer Hochschule: Prof. Dr. Benjamin Himpel  
Hochschule Reutlingen  
Benjamin.Himpel@Reutlingen-University.de

Informatics Inside Herbst 2021  
24. November 2021, Hochschule Reutlingen  
Copyright 2021 Maximilian Rabus

## Keywords

Sentiment Analysis, NLP, Regression modeling

## 1 Einleitung

Im Folgenden wird die Ausgangslage, der gewählte Ansatz und das daraus abgeleitete Vorgehen genauer beschrieben.

### 1.1 Ausgangslage

Eine der bekanntesten Theorien, die sich gegen die Vorhersage von Aktienkursen und -märkten ausspricht, ist die sogenannte Random Walk Theorie [1]. Sie besagt unter anderem, dass die Ereignisse, Trends und Kurse der Vergangenheit nicht genutzt werden können, um daraus eine Kursentwicklung für die Zukunft vorhersagen zu können. Eine der kontroverseren Aussagen ist, dass eine zufällige Auswahl von Aktien in der Regel nicht besser oder schlechter als die eines von Experten ausgewählten Portfolios abschneidet – eine Unterstreichung der Zufallsnatur eines Aktienmarktes.

Diese Theorie hat seit 1973, dem ursprünglichen Geburtsjahr der Theorie, eine große Anzahl von Veröffentlichungen nicht davon abgehalten eigene Ansätze für Vorhersagen von Aktienmärkten zu entwickeln. Konzepte, die dabei eine Rolle spielen, sind häufig aus dem Bereich des Machine-Learning [2], aber auch andere Ansätze aus dem Bereich des NLP (Natural Language Processing) und damit zusammenhängend Sentiment-Analysen finden hier Anwendung [3].



## 1.2 *Ansatz*

Der grundlegende Ansatz dieses Papers ist es, durch die Analyse von Posts in dem Internetforum Reddit verschiedenste Kennzahlen aus den Bereichen der Sentiment-Analyse [9] und der deskriptiven Statistik zu identifizieren, die für die Entwicklung von Aktienkursen relevant sind. Diese Identifikation findet über das Auswerten von Korrelationen zu der Entwicklung eines Aktienkurses statt. Mit Hilfe dieser Kennzahlen sollen anschließend Regressionsmodelle trainiert und getestet werden.

Der Fokus soll in dieser Arbeit hauptsächlich auf der kurzfristigen Vorhersage von sogenannten Pennystocks liegen. Ganz allgemein gilt eine Aktie als ein Pennystock, die oft nur wenige Cents, bis einige Dollar als Wert pro Anteil vorweisen und denen grundsätzlich eine hohe Volatilität zugeschrieben werden kann. Diese Volatilität leitet sich aus den großen Diskrepanzen zwischen Potenzial und Realität der finanziellen Potenz einer Firma oder eine Geschäftsidee ab.

Die Idee dieses Papers ist es die Ereignisse und Kennzahlen dieser hohen Volatilität zu erkennen und sich dabei der „Schwarmintelligenz“ eines großen Forums wie Reddit zu bedienen, um diese Trends vorherzusehen.

## 1.3 *Vorgehen*

In Kapitel 2 werden zunächst historische Daten gesammelt, die als Grundlage für die Analyse dienen sollen. Dort werden auch erste potenzielle Kennzahlen aus den Metadaten der gesammelten Daten festgehalten.

Kapitel 3 enthält eine genauere Beschreibung wie Sentiment-Analysen in dieser Arbeit ablaufen und vergleicht zwei verschiedene Ansätze.

Zusätzlich zu den potenziellen Kennzahlen der Metadaten, werden in Kapitel 4 eine

Reihe von weiteren Kennzahlen definiert, wo ein Zusammenhang zu der Entwicklung eines Aktienpreises vermutet wird.

Alle potenziellen Kennzahlen werden in Kapitel 5 getestet, indem die Korrelation zwischen einer Kennzahl und der Entwicklung des Aktienpreises festgestellt wird. Diese Entwicklung wird in Kapitel 5 ebenfalls als feste Zielgröße definiert.

Basierend auf diesen Tests wird in Kapitel 6 ein Modell entwickelt und definiert. Zu diesem Zweck werden eine Reihe von Regressionsmodellen getestet, verglichen und ein Sieger identifiziert. Das finale Modell wird anschließend mit den vorhandenen Daten trainiert, getestet und anhand einer Confusion Matrix ausgewertet.

Kapitel 7 fasst die gewonnenen Erkenntnisse zusammen und zieht ein Fazit.

Abschließend wird in Kapitel 8 ein Ausblick auf die weitere Arbeit und Ansatzpunkte dieses Papers skizziert.

Die Umsetzung findet in Python statt. Eine genauere Übersicht der verwendeten Bibliotheken und der Code befindet sich in Google Drive<sup>1</sup>.

## 2 **Sammeln und Aufbereiten der Daten**

Für das Trainieren eines passenden Modells, welches die bereits skizzierten Anforderungen erfüllt, muss eine entsprechende Datengrundlage geschaffen werden. Dieses Kapitel beschäftigt sich mit den einzelnen Schritten, die auf dem Weg zu dieser Datengrundlage durchgeführt wurden.

### 2.1 *Reddit und Pennystocks*

Reddit bezeichnet sich selbst als „the frontpage of the internet“ und ist im Kern eine

---

<sup>1</sup> <https://colab.research.google.com/drive/1MEuY8jfuZ79TJyz-2KN9otScn0kc1Fjx?usp=sharing>

Sammlung verschiedenster Sub-Foren zu jedem erdenklichen Thema [4].

Zu den Themen mit denen sich in den einzelnen sogenannten Subreddits auseinandergesetzt wird, gehören auch ganz allgemein Finanzen oder in spezifischeren Subreddits Aktien. Die Mitgliederanzahl dieser Subreddits belaufen sich auf mehrere Millionen und setzen in den einzelnen Unterforen unterschiedliche Schwerpunkte.

Spezifisch für Pennystocks gibt es einen Subreddit mit demselben Namen, der annähernd 1,8 Millionen Mitglieder vorweist und in dem täglich tausende von Pennystocks vorgeschlagen und diskutiert werden<sup>1</sup>.

Aufgrund der großen thematischen Überschneidung zum Thema dieser Arbeit wird der Pennystocks Subreddit im Folgenden als Grundlage für die Sammlung der historischen Daten verwendet.

## 2.2 Sammeln der Rohdaten

Für das Sammeln der historischen Datengrundlage, die für den Zweck dieser Ausarbeitung unabdinglich ist, wurde die Pushshift-API verwendet. Pushshift ist ein Datensatz, der seit 2015 regelmäßig aktualisiert wird und historische (Meta-)Daten von Posts und Kommentaren auf Reddit seit dessen Existenz beinhaltet [5].

Für eine ausreichend große Datengrundlage wurden über die Pushshift-API alle Posts im Subreddit „Pennystocks“<sup>2</sup> vom 1. November 2020 bis zum 29. Juni 2021 – also etwa über 8 Monate – gesammelt.

Jeder dieser Posts beinhaltet eine große Anzahl von Meta-Daten: Der Titel und Text des Posts, die Anzahl der Kommentare, das Datum der Veröffentlichung und etwa 100 weitere Attribute, die nicht alle für den Zweck dieser Ausarbeitung dienlich sind.

## 2.3 Relevante Attribute

Um die Komplexität des Datensatzes zu reduzieren und einen Fokus auf relevante Attribute zu ermöglichen wurden aus den etwa 100 existierenden Attributen jedes Posts 7 ausgewählt, die für die weitere Verarbeitung potenziell relevant sind:

1. **created\_utc**: Enthält den Zeitpunkt der Veröffentlichung
2. **id**: Identifier
3. **title**: Der Titel
4. **selftext**: Der tatsächliche Text des Posts
5. **score**: Die Upvotes minus die Downvotes
6. **upvote\_ratio**: Das Verhältnis von Up- zu Downvotes
7. **num\_comments**: Die Anzahl der Kommentare
8. **num\_crossposts**: Wie oft ein Post an einer anderen Stelle erneut gepostet wurde

Als potenziell relevant wurden hierbei Attribute bezeichnet und ausgewählt, wo eine Korrelation zwischen dem Attribut und der Entwicklung eines Aktienpreises vermutet wurde. Diese Auswahl basierend auf Vermutungen ist nötig, da eine Auswertung aller hundert Attribute zu zeitaufwändig wäre.

Konzentriert wurde sich also vor allem auf Attribute, die das Auslesen eines Sentiments ermöglichen oder einen andersartigen Hinweis auf das Sentiment des Autors oder der Reaktionen auf diesen Post liefern.

Attribute wie die „id“ oder „created\_utc“ wurden für den wichtigen zeitlichen Kontext und die Möglichkeit weitere Informationen über diesen Post später auslesen zu können ausgewählt.

---

<sup>2</sup> <https://www.reddit.com/r/pennystocks/>

Eine genauere Auswertung inwiefern diese Attribute auch Kennzahlen mit einer Aussagekraft für die Entwicklung eines Aktienkurses haben, findet in Kapitel 5 statt.

Attribute wie „created\_utc“ und die „id“ sind selbstverständlich keine potenziellen Kennzahlen, sondern werden für die Erstellung eigener Kennzahlen in Kapitel 4 verwendet.

## 2.4 Bezug zu einer Aktie

Für die Suche nach Korrelationen zwischen einzelnen Attributen und der Entwicklung eines Aktienpreises ist es notwendig einen Post eindeutig einer Aktie zuweisen zu können.

Im „Pennystocks“-Subreddit ganz speziell, aber allgemein in entsprechenden finanzbezogenen Foren werden für die Zuordnung von einer Zeichenkette zu einer Aktie die sogenannten Symbols verwendet, welche seit jeher als eindeutige Identifier einer Aktie dienen [6].

Um zu erkennen, ob ein Post sich eindeutig einer Aktie zuordnen lässt, wurden die Titel der einzelnen Posts mit einer Liste von Symbolen abgeglichen, die einen Großteil der existierenden Aktien an den Märkten des NASDAQ, NYSE und AMEX abdeckt<sup>3</sup>.

Nach diesem Abgleich blieben 12359 Posts aus dem in 3.1 erwähnten Zeitraum übrig, die sich eindeutig 1349 verschiedenen Aktien zuordnen ließen.

## 3 Sentiment-Analyse

Die im nächsten Kapitel definierten Kennzahlen basieren unter anderem auf den Sentimenten verschiedenster Daten. Im Folgenden wird genauer beschrieben welche Konzepte und Bibliotheken verwendet werden, um diese Sentimente zu erheben.

### 3.1 VADER

Für das Ziel der Sentiment-Analyse gibt es eine Reihe von Ansätzen und entsprechenden Bibliotheken. Arbeiten im Umfeld von

Social Media wie z. B. auf Twitter oder auch im spezifischen auf Reddit setzen hierbei in der Regel auf VADER [7, 8], eine Python-Bibliothek für die Analyse von Sentimenten.

Grundsätzlich gibt VADER basierend auf einem Input-String eine Ausgabe zurück, die die sogenannten „Polarity Scores“ enthält. Diese beziffern für die Sentimente „Negativ“, „Neutral“ und „Positiv“ jeweils einen Wert von 0 bis 1 und beschreiben in ihrer Gesamtheit das festgestellte Sentiment des Eingabetextes.

### 3.2 Emojis

Gerade im Umfeld von Social Media und allgemein bei Textnachrichten aller Art sind Emojis als zusätzlicher Indikator für ein Sentiment ein essentialer Faktor, der für die Analyse ebenfalls betrachtet werden soll.

Standardmäßig ist VADER bereits in der Lage Emojis in Texten zu erkennen und deren Intention in die Auswertung mit einfließen zu lassen. Tests mit verschiedensten Emojis legen hier jedoch offen, dass VADER eine Großzahl der Emojis in dem Kontext des Pennystocks-Subreddits entweder nicht erkennt oder ein zu neutrales Sentiment zuordnet.

Auf der Suche nach einer Alternative für die Sentiment-Analyse von Emojis wurde sich letztendlich für die Python-Bibliothek „emosent-py“ entschieden. „emosent-py“ liefert als Rückgabe für einen Emoji unter anderem einen sogenannten „sentiment\_score“, der das festgestellte Sentiment von -1 bis 1 beinhaltet.

Erste Tests mit dieser Bibliothek zeigten, dass „emosent-py“ eher das tatsächliche Sentiment eines Emojis widerspiegelte. Aufgrund der ersten positiven Resultate wurde sich für einen größer angelegten Test zweier verschiedener Ansätze für die Analyse von

<sup>3</sup> <https://www.nasdaq.com/market-activity/stocks/screener>

Sentimenten entschieden, um eine klare Aussage über den passenderen Ansatz für den Zweck dieser Arbeit liefern zu können.

### 3.3 Hybrid aus VADER und emosent-py

Da „emosent-py“ keine Aussage über das Sentiment eines reinen Textes treffen kann, ist ein Hybrid-Ansatz nötig, um das tatsächliche Sentiment eines Textes zu erfassen.

Für diesen Ansatz wurden zunächst die Emojis aus einem Input-String extrahiert und separat die Emojis von „emosent-py“ und der reine Text von VADER bewertet.

Die zwei Werte werden anschließend aufsummiert. Hier wurde sich bewusst gegen ein Mitteln der Werte entschieden um der Aussagekraft von Emojis einen höheren Stellenwert zuzuordnen und allgemein die Unterscheidung zwischen positiv und negativ klarer zu machen und die Sentiment-Aussage zu schärfen.

### 3.4 Vergleich der beiden Ansätze

Im Folgenden soll der Ansatz aus 3.3 mit einer reinen Vorhersage von VADER verglichen werden.

Zu diesem Zweck wurden zufällig 1000 Einträge aus dem Datensatz von 2.4 ausgewählt und deren Sentiment manuell eingetragen.

Die Titel der einzelnen Posts aus diesem Datensatz wurden anschließend von beiden Ansätzen bewertet und mit dem gewünschten Ergebnis verglichen.

**Tabelle 1. Vergleich von „korrekten“ Sentiment-Analysen zwischen VADER und einem Hybrid-Ansatz**

	Treffer	Trefferquote
VADER	567	56.7%
Hybrid-Ansatz	721	72.1%

In der Auswertung ist klar zu erkennen, dass der Hybrid-Ansatz aus den 1000 Posts eine deutlich höhere Trefferquote vorweist und

aus diesem Grund auch für weitere Sentiment-Analysen dieser Arbeit verwendet wird.

## 4 Definieren eigener Kennzahlen

Zusätzlich zu den in Kapitel 2.3 ausgewählten Kennzahlen aus den bereits vorhandenen Meta-Daten der einzelnen Posts, sollen im Folgenden zehn weitere potenzielle Kennzahlen definiert werden, welche in Kapitel 5 ausgewertet und getestet werden. Für eine genauere Unterscheidung der einzelnen Kennzahlen wurden diese in 3 Kategorien aufgeteilt.

### 4.1 Meta-Kennzahlen

Die folgende Auflistung enthält alle selbst definierten Kennzahlen, welche sich aus der Analyse der ursprünglichen (Meta-)Daten oder deren direkt zusammenhängenden Informationen ableiten lassen:

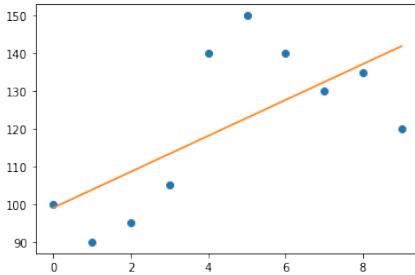
- **Titel-Sentiment:** Das Sentiment des Titels des Posts.
- **Body-Sentiment:** Das Sentiment des tatsächlichen Inhalt des Posts.
- **Kommentar-Sentiment:** Das gemittelte Sentiment aller Kommentare des Posts.
- **Body-Länge:** Die Anzahl der Zeichen im Body des Posts.

### 4.2 Trend-Kennzahlen

Durch die große Anzahl von Datenpunkten, die analysiert werden, sind diese isoliert betrachtet nicht ausreichend, um ein vollumfassendes Bild von der potenziellen Entwicklung einer Aktie zu erhalten. In dieser Arbeit werden Trend-Analysen verwendet, um den Entwicklungen und Trends einzelner Zahlen eine der Realität zutreffendere Kennzahl definieren zu können.

Für alle Trend-Kennzahlen wurden einzelne Datenpunkte einer Zeitachse zugeordnet und ein Polynom 1. Ordnung als Regressionsgerade berechnet. Die Steigung dieser Gerade

ist der Wert, der den einzelnen Trend-Kennzahlen zugeordnet wurde.



**Abbildung 1: Lineare Regression mit Beispiels-Datenpunkten**

Abbildung 1 zeigt eine solche Regressionsgerade mit Beispiel-Datenpunkten. Gut zu erkennen ist hierbei, wie die Steigung der Gerade eine gute Beschreibung für den generellen Trend eines Wertes ist.

Für das Verständnis der folgenden Kennzahlen ist das Konzept des Aktiensentiments relevant, was wie folgt definiert wird: Das Aktiensentiment ist das wahrgenommene Sentiment auf Reddit für eine gewisse Aktie in einem gewissen Zeitraum. Im Rahmen dieser Arbeit werden hierfür die Sentimente aller Post-Titel auf ganz Reddit die eine bestimmte Aktie erwähnen in einem Zwei-Tages-Zeitraum ausgelesen und anschließend der Durchschnitt als finaler Wert des Aktiensentiments ausgerechnet.

Für die folgenden Trendbetrachtungen werden vier Datenpunkte des Aktiensentiments, also die letzten acht Tage, erhoben:

- **Trend Aktiensentiment:** Trend, der das Aktiensentiment der letzten acht Tage beschreibt
- **Aktie Posts-Trend:** Trend, der die Anzahl der Posts zu einer bestimmten Aktie beschreibt.
- **SD-Trend Aktiensentiment:** Trend, der die Standardabweichung des Aktiensentiments der letzten acht Tage beschreibt

- **Aktie Kurs-Trend:** Trend, der die Entwicklung des Kurspreises der letzten acht Tage beschreibt.

Die Standardabweichung definiert sich wie folgt:

$$SD = \frac{\sum(x_i - \bar{x})^2}{n - 1}$$

Die Standardabweichung beschreibt die Streuung von Datenpunkten rund um einen Mittelwert. Für die Kennzahl SD-Trend Aktiensentiment kann dies helfen zu verstehen wie „einig“ die Meinungen zu einer bestimmten Aktie ausfallen.

### 4.3 Durchschnitts-Kennzahlen

Die folgende Auflistung enthält alle selbst definierten Kennzahlen, welche einen Durchschnitt von Daten beschreiben, der sich nicht aus den ursprünglichen (Meta)Daten ableiten lässt:

- **Aktie SD-Sentiment-Durchschnitt:** Der Durchschnitt der Standardabweichung des Sentiments einer Aktie, die einem Post zugeordnet wurde.
- **Aktie Sentiment-Durchschnitt:** Der Durchschnitt des Sentiments einer Aktie, die einem Post zugeordnet wurde.

### 4.4 Anreichern des Datensatzes

Nach der Definition der eigenen Kennzahlen wurde der ursprüngliche Datensatz aus 2.4 mit diesen angereichert.

Für die Bestimmung der eigenen Kennzahlen wurde sich neben der Pushshift-API auch bei der eigentlichen Reddit-API und der Yahoo Finance-API bedient.

Die Reddit-API war zusätzlich zu Pushshift nötig, da die (Meta-)Daten von Pushshift zu einem einzelnen Post nicht immer alle Daten vollständig enthalten. Gerade Attribute wie der Text des Bodys oder dessen Länge waren

hierbei oft entweder leer oder enthielten falsche Werte. Aus diesem Grund wurden diese separat durch die „id“ eines Posts von der Reddit-API abgefragt.

Die Yahoo-Finance API wurde immer dann verwendet, wenn historische oder aktuelle Preise einer Aktie abgefragt werden mussten.

Sollten keine Daten zu einem bestimmten Post gefunden werden, so wurde dieser aus dem Datensatz gestrichen, um die Aussagekraft des Ergebnisses nicht durch zu viele Null-Ergebnisse zu verschleiern. Häufigstes Beispiel dafür war eine Nichtauffindung von Aktienpreisen für ein identifiziertes Symbol. Da es im Bereich der Pennystocks häufig auch um eher unbekannte Aktien geht, sind die Aufzeichnungen der historischen Aktienpreise oft unvollständig.

Nach der Aussortierung dieser Daten blieb ein finaler Datensatz mit insgesamt 3571 Einträgen und 257 eindeutigen Aktien übrig, der in den folgenden Kapiteln für die weiteren Auswertungen und Berechnungen verwendet wird.

## 5 Auswertung der Kennzahlen

Nachdem in Kapitel 2 und 4 eine Reihe von potenziellen Kennzahlen identifiziert wurden, sollen diese im Folgenden getestet und ausgewertet werden.

Ziel ist es ein finales Set an Kennzahlen zu definieren, mit welchem ein Regressionsmodell trainiert werden kann.

### 5.1 Zielgröße

Da das Testen und die Auswertung der Kennzahlen basierend auf Korrelationen erfolgen soll, muss eine Zielgröße festgelegt werden, zu der die Korrelationen aufgezeigt werden können.

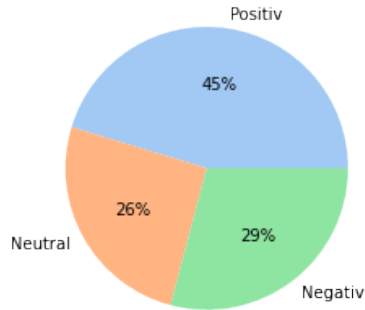


Abbildung 2: Anteile der verschiedenen Sentimente in finalen Datensatz

Aufgrund der hohen Volatilität und der überwiegend positiven Sentimente (Abbildung 2) innerhalb des Pennystock Subreddits, wurde eine Zielgröße gewählt, die den höchsten Anstieg innerhalb der nächsten vier Wochen widerspiegelt. Für diesen Zweck werden die Aktienpreise der nächsten vier Wochen ab dem Zeitpunkt der Veröffentlichung des Posts analysiert und der höchste „Close“-Price ausgewählt. Für den „Close“-Preis wurde sich entschieden, da dies den letzten Preis darstellt, zu welchem ein tatsächlicher Kauf noch stattfinden kann. „High“- „Low“- und „Open“-Preise sind oft anfällig für kurzfristige Anstiege und Abfälle des Preises, die nicht wirklich den tatsächlichen Wert repräsentieren. Dieser höchste „Close“-Preis wird anschließend ins Verhältnis zum Preis der entsprechenden Aktie zum Zeitpunkt der Veröffentlichung des Posts gesetzt und der prozentuale Anstieg/Abfall festgestellt.

Der prozentuale Anstieg/Abfall ist als Zielgröße aus mehreren Gründen noch nicht geeignet. Egal wie stark eine Korrelation mit dieser Zielgröße ausfällt oder wie gut ein Modell ausfällt was diese prozentualen Anstiege vorhersagt, eine Zufallskomponente wird immer bleiben, die die Aussagen eines Modells verfälschen kann. Aus diesem Grund wurde sich für eine Vereinfachung entschieden.

Anstatt den prozentualen Anstieg/Abfall zu verwenden wird eine 0 für einen maximalen Anstieg von kleiner 20 Prozent und eine 1 für einen Anstieg von größer oder gleich 20 Prozent eingetragen. Anstatt also den genauen prozentualen Anstieg vorherzusagen, geht es darum allgemein eine Empfehlung für oder gegen einen Kauf auszusprechen. Dies hat den Vorteil, dass Schwankungen durch zufällige Events durch eine solche Kategorisierung eher abgefangen werden. Auch die 20 Prozent haben den Hintergrund dies auszugleichen und bei einer Kaufempfehlung eine gewisse „Sicherheit“ vorweisen zu können.

Die Zielgröße wird im Folgenden „Kaufempfehlung“ genannt.

### 5.2 Korrelation der Kennzahlen mit Zielgröße

Für die Auswertung der Korrelation zur Zielgröße Kaufempfehlung wird eine Pearson-Korrelation verwendet, die als Rückgabe einen Wert zwischen -1 und 1 zurückgibt. Eine 1 steht für eine perfekte positive und eine -1 für eine perfekte negative Beziehung der beiden Werte.

Tabelle 2 enthält die Korrelationen zwischen allen in den vorangegangenen Kapiteln aufgestellten Kennzahlen und der Zielgröße Kaufempfehlung, sortiert von positiv nach negativ.

**Tabelle 2. Korrelationen der einzelnen Kennzahlen mit der Zielgröße „Kaufempfehlung“**

Attribut	Korrelation
Aktie Sentiment-Durchschnitt	0.085
Kommentar-Sentiment	0.029
Body-Sentiment	0.012
Titel-Sentiment	0.003
num_crossposts	-0.000
Trend Aktiensentiment	-0.010

score	-0.012
num_comments	-0.012
Aktie Posts-Trend	-0.023
upvote_ratio	-0.024
Aktie SD Sentiment-Trend	-0.031
Body-Länge	-0.047
Aktie SD-Sentiment-Durchschnitt	-0.107
Aktie Kurs Trend	-0.171

### 5.3 Auswertung der Korrelationen

Ganz allgemein lässt sich auf den ersten Blick erkennen, dass selbst die größten positiven und negativen Korrelationen mit der Zielgröße Kaufempfehlung eher klein ausfallen. Dies ist zu erwarten, da eine gewisse Zufallsnatur dem Aktienmarkt nicht abzusprechen ist. Eine perfekte Vorhersage wird also nie möglich sein. Ziel ist es eine möglichst hohe Trefferquote über 50 Prozent zu erzielen, die ermöglicht über die Menge der Käufe und Verkäufe trotzdem einen Gewinn zu erzielen.

Zunächst lässt sich festhalten, dass das Sentiment des Autors selbst, sowohl des Titels wie auch des Bodys, keine tatsächliche Korrelation mit der Zielgröße vorweist. Auch andere Kennzahlen, die direkt mit dem Post zusammenhängen wie die Anzahl der Kommentare, die Anzahl der Crossposts oder der eigentliche Score weisen eine Korrelation vor, die fast 0 ist.

Die stärkste Korrelation ist eine negative: Aktie Kurs Trend. Die negative Korrelation weist daraufhin, dass vergangene negative Entwicklungen des Aktienpreises in der Vergangenheit einen positiven Trend in der Zukunft begünstigen können.

Bis auf die Kennzahl „Trend Aktiensentiment“ weisen auch die meisten Sentiment-

Kennzahlen eine gewisse positive oder negative Korrelation vor.

Herauszustellen ist hierbei noch die Kennzahl „Aktie SD-Sentiment-Durchschnitt“. Die negative Korrelation impliziert, dass eine geringere Streuung eines vorherrschenden Sentiments - was bedeutet, dass sich die Menge an Investoren auf Reddit „einig“ sind - einen positiven Aktientrend zur Folge haben kann.

Ziel dieses Unterkapitels ist es auch ein finales Set an Kennzahlen festzulegen, welches im nächsten Kapitel für das Trainieren eines Modells verwendet werden kann. Aus diesem Grund werden die folgenden Kennzahlen aufgrund von zu kleinen Korrelationen entfernt:

- Body-Sentiment
- Titel-Sentiment
- num\_crossposts
- Trend Aktiensentiment
- score
- num\_comments

Übrig bleiben also acht Kennzahlen, die in Abbildung 3 noch einmal in einer Heatmap dargestellt werden.

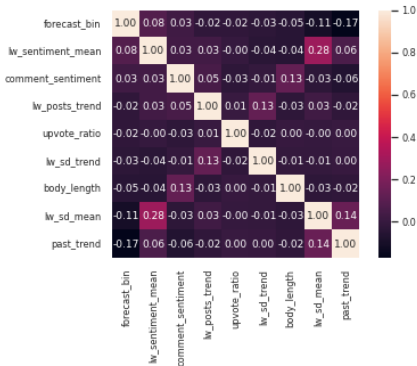


Abbildung 3: Korrelationen der final ausgewählten Kennzahlen

## 6 Trainieren und Testen des Modells

Nachdem in Kapitel 5 ein finales Set an Kennzahlen definiert wurde, kann in diesem Kapitel ein Modell ausgewählt und trainiert werden.

### 6.1 Wahl des Regressionsmodells

Die „sklearn“-Bibliothek für Python stellt eine Reihe von Regressionsmodellen bereit, aus denen folgende für einen Test ausgewählt wurden:

- LinearRegression
- Lasso
- ElasticNet
- KNeighborsRegressor
- DecisionTreeRegressor
- GradientBoostingRegressor

Ein finaler Datensatz wurde aus den angereicherten Daten aus 4.4 und den Kennzahlen aus 5.3 erstellt und in ein Trainings- und Test-Set unterteilt.

Um eine ungefähre Normalverteilung der einzelnen Werte der Spalten – also der Kennzahlen – zu gewährleisten, wurden alle Werte standardisiert<sup>4</sup>

Durch eine „Cross-Validation“ des Test-Sets mit den einzelnen Regressionsmodellen konnte der „mean squared error“ – im Folgenden MSE abgekürzt - und dessen Standardabweichung festgestellt werden<sup>5</sup>. Diese

<sup>4</sup> <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html>

<sup>5</sup> [https://scikit-learn.org/stable/modules/generated/sklearn.model\\_selection.cross\\_val\\_score.html#sklearn.model\\_selection.cross\\_val\\_score](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.cross_val_score.html#sklearn.model_selection.cross_val_score)



werden zur Bewertung der einzelnen Modelle genutzt, um ein finales auszuwählen.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Obige Formel zeigt die Berechnung des MSE. Allgemein beschreibt der MSE die Genauigkeit zwischen einem vorhergesagten und beobachteten Wert – in dem Fall dieses Vergleichs also die Vorhersagen der Modelle gegenüber der tatsächlichen Werte - als Summe der quadrierten Werte der einzelnen Differenzen und ist damit ein passendes Bewertungskriterium für die Bewertung der einzelnen Modelle.

Tabelle 3 zeigt die einzelnen Modelle und deren Ergebnisse der „Cross-Validation“ der Trainings-Daten.

**Tabelle 3. Ergebnisse der „Cross-Validation“ der einzelnen Regressionsmodelle**

	MSE	STD
<b>LinearRegression</b>	-0.211429	0.012392
<b>Lasso</b>	-0.221321	0.009791
<b>ElasticNet</b>	-0.221321	0.009791
<b>KNeighborsRegressor</b>	-0.210006	0.014742
<b>DecisionTreeRegressor</b>	-0.266103	0.031002
<b>GradientBoostingRegressor</b>	-0.173251	0.009610

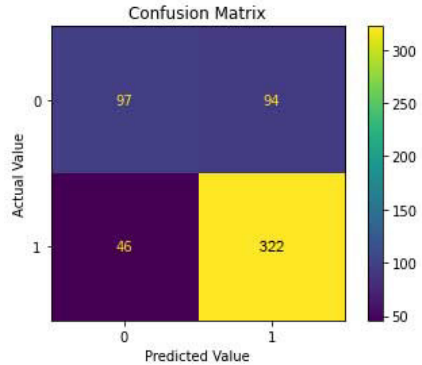
Basierend auf diesen Ergebnissen wurde der „GradientBoostingRegressor“ als Modell für das weitere Verfahren ausgewählt, da er die geringste durchschnittliche Abweichung und - auch wenn dies ein sekundäres Bewertungskriterium war – die geringste Standardabweichung vorweisen kann.

## 6.2 Trainieren und Testen des Regressionsmodells

Nach dem Testen durch die „Cross-Validation“ kann das Modell des „GradientBoostingRegressors“ jetzt final mit den Trainingsdaten trainiert werden.

Das trainierte Modell wurde anschließend genutzt, um Vorhersagen mit den Test-Daten zu treffen. Die Vorhersagen des Modells sind nicht exakt 0 und 1, sondern bewegen sich immer zwischen diesen beiden Werten.

Um eine Auswertung in Form einer Confusion Matrix vornehmen zu können, wurden die Vorhersagen in 0 - alle Vorhersagen kleiner 0,5 - und in 1 - alle Vorhersagen größer oder gleich 0,5 - umgewandelt.



**Abbildung 4: Confusion Matrix des finalen Modells**

Abbildung 4 zeigt von links oben nach rechts unten die True Negatives (TN), False Positives (FP), False Negatives (FN) und die True Positives (TP) des Modells.

Aus der Confusion Matrix lassen sich jetzt weitere Kennzahlen ableiten.

$$Recall = \frac{TP}{TP + FN} = 0.875$$

Der Recall beschreibt wie viele der tatsächlichen positiven Werte korrekt vorhergesagt wurden.

$$Precision = \frac{TP}{TP + FP} \approx 0.774$$

Die Precision beschreibt wie viele der vorhergesagten positiven Werte auch tatsächliche

$$Accuracy = \frac{TP + TN}{n} \approx 0.749$$

Die Accuracy beschreibt für die Menge der Vorhersagen  $n$  den Prozentsatz der korrekt vorhergesagten Werte.

## 7 Fazit

Grundsätzlich sprechen die Ergebnisse dieser Arbeit für die These, dass es möglich ist durch die Analyse vergangener (Sentiment-)Daten Vorhersagen zu treffen, die bessere Ergebnisse als ein zufälliger Trading-Algorithmus liefern. Indizien dafür sind die guten Ergebnisse der Kennzahlen Recall, Precision und Accuracy.

Die Meinungen der Individuen spielen – wie man an den Korrelationen der Sentimente des eigentlichen Posts sehen kann – dabei kaum eine Rolle. Die Reaktionen auf diesen Post und das generelle Sentiment gegenüber einer Aktie zu einem gewissen Zeitpunkt besitzen eine deutlich höhere Aussagekraft.

Die Trendbetrachtungen über die Steigung der Regressionskurven hat sich ebenfalls als valides Indiz für Kursentwicklungen sowohl für Sentimente wie auch für preisliche Kursentwicklungen bewiesen.

Trotzdem sind diese Ergebnisse mit Vorsicht zu genießen. Diese Arbeit kann und muss eher als Grundlagenarbeit verstanden werden, an der es noch einige Ansatzpunkte gibt. Grund für diese Vorsicht ist unter anderem der verhältnismäßig kleine Datensatz mit dem die Regressionsmodelle getestet und trainiert wurden. Aufgrund der vielen APIs, die für die Berechnung der Kennzahlen abgefragt werden müssen, sind die Rechenzeiten ein Hindernis für umfangreicheres Testen.

Auch die Natur der Zielgröße Kaufempfehlung und das Trainieren des Modells auf überwiegend positive Vorhersagen durch die 20%-Grenze können Ursachen für die große Anzahl an positiven Werten und Treffern der Confusion Matrix sein und die Aussagekraft der guten Werte für Recall, Precision und Accuracy reduzieren.

Zukünftige Arbeiten müssen hier durch umfangreicheres Testen die Ansätze dieser Arbeit validieren.

## 8 Ausblick

Auch wenn das grundsätzliche Ergebnis dieser Arbeit ein Indiz für die Richtigkeit der verwendeten Ansätze ist, so gibt es doch eine Reihe von Punkten, an denen man diese vertiefen kann.

Zunächst sei die grundsätzliche Idee erwähnt, die mit ihrem Fokus auf die volatile Natur der Pennystocks einen nur kleinen Teil der Welt der Aktienmärkte abdeckt. Die Definition von Kennzahlen mit längerfristigen Implikationen würde hier eine spannende Herausforderung darstellen.

Die Datengrundlage bietet einen weiteren Ansatzpunkt. Die Größe des Datensatzes könnte noch einmal etwas größer gewählt werden um Einflussfaktoren wie die Schwankungen der Weltmärkte insgesamt und statistische Anomalien besser ausgleichen zu können.

Auch die Betrachtung von Reddit bzw. des Pennystocks Subreddit ist eine zwar bewusst gewählte aber allgemein nicht notwendige Einschränkung. Die Daten, Sentimente und Analysen weiterer finanzbezogener Foren könnten hier hinzugezogen werden.

Die gewählte Zielgröße Kaufempfehlung hat sich in dieser Arbeit zwar bewährt, jedoch besteht hier großes Potenzial zum Beispiel durch Trendbetrachtungen andere Zielgrößen zu definieren und zu testen.

Auch die Regressionsmodelle bilden nur einen kleinen Teil der Landschaft in diesem Bereich ab. Hier einen breiter angelegten Test durchzuführen und das finale Modell zu verbessern wäre ein interessanter Ansatz für die Zukunft.

Langfristig wäre eine tatsächliche Portfolio-simulation mit dem in dieser Arbeit definierten Modell interessant, die regelmäßig aktuelle Posts im Pennystocks Subreddit analysiert und darauf basierend Kaufempfehlungen ausspricht. Besonders spannend wäre dies gekoppelt mit „echten“ Trading-Strate-

gien, die die Vorhersagen des Modells ergänzen um so einen allgemeinen, robusten Trading-Algorithmus zu schaffen.

## Literaturverzeichnis

- [1] Malkiel, Burton Gordon. 2020. A random walk down Wall Street. The time-tested strategy for successful investing. New York: W.W. Norton & Company.
- [2] Pang, X., Zhou, Y. and Wang, P. 2020. An innovative neural network approach for stock market prediction. *J Supercomput* 76, 2098–2118 (2020). <https://doi.org/10.1007/s11227-017-2228-y>
- [3] Mittal A., Goel A. 2011. Stock prediction using twitter sentiment analysis 2011; Stanford University, CS229.
- [4] Anderson, K.E. 2015. Ask me anything: what is Reddit? *Library Hi Tech News*, Vol. 32 No. 5, pp. 8-11. <https://doi.org/10.1108/LHTN-03-2015-0018>
- [5] Baumgartner, J., Zannettou, S., Keegan, B., Squire, M. and Blackburn, J. 2020. The Pushshift Reddit Dataset. *Proceedings of the International AAAI Conference on Web and Social Media*. 14, 1 (May 2020), 830-839.
- [6] Anderson, Alyssa G. and Larkin, Yelena. 2019. Does Noninformative Text Affect Investor Behavior? In *Financial Management* 48 (1), pp. 257–289. DOI: 10.1111/fima.12225.
- [7] A. Shelar and C. -Y. Huang. 2018. Sentiment Analysis of Twitter Data. *International Conference on Computational Science and Computational Intelligence (CSCI)*, 2018, pp. 1301-1302, doi: 10.1109/CSCI46756.2018.00252.
- [8] Davis, Jenny L. and Graham, Timothy. 2021. Emotional consequences and attention rewards: the social effects of ratings on Reddit. In *Information, Communication & Society* 24 (5), pp. 649–666. DOI: 10.1080/1369118X.2021.1874476.
- [9] Ronen Feldman. 2013. Techniques and applications for sentiment analysis. *Commun. ACM* 56, 4 (April 2013), 82–89. DOI:<https://doi.org/10.1145/2436256.2436274>



# Vergleich von maschinellen Lernalgorithmen zur Klassifizierung von wissenschaftlichen Abstracts

Philip Storz

Hochschule Reutlingen

Philip.Storz@Student.Reutlingen-University.de

## Abstract

Die automatisierte Klassifizierung von Dokumenten findet in vielen Gebieten Anwendung und kann eine schnellere Verarbeitung und einen geordneten Überblick über große unstrukturierten Datenmengen bieten. In dieser Arbeit wird ein Überblick über Lernalgorithmen zur Textklassifikation gegeben. Zudem wird neben einem State-of-the-Art Vergleich ein eigener Vergleich von drei Lernalgorithmen vorgestellt. Dabei wird die Performance von einer Support-Vektor-Maschine (SVM) mit Word2Vec, einer SVM mit *Term Frequency-Inverse Document Frequency* (TF-IDF) und einem rekurrenten neuronalen Netz (RNN) mit *Long Short-Term Memory* (LSTM) bei der Klassifikation von Abstracts verglichen. Als Datengrundlage wurde mit der Elsevier API ein Korpus bestehend aus 27329 Abstracts aus 26 Themengebieten zusammengestellt und annotiert. Bei der Klassifikation von fünf Themengebieten mit jeweils mehr als 1000 Abstracts hatte die SVM mit TF-IDF die beste Performance mit einer Genauigkeit von 86 % und einem F1-Wert von 86 %. Das RNN hat mit 82 % eine ähnliche Genauigkeit wie die SVM, hatte aber aufgrund eines niedrigen Präzisionswertes von 57 % nur einen F1-Wert von 70 %. Die SVM mit Word2Vec hat mit einer Genauigkeit von 43 % und einem F1-Wert von 43 % die schlechteste Performance. Es wurde bewiesen, dass eine Klassifikation von Abstracts in ihre Themengebiete möglich ist. Es muss hierfür jedoch eine bessere Datengrundlage aus annotierten frei verfügbaren Korpora geschaffen werden.

## CCS Concepts

• **Computing methodologies** → **Information extraction**; *Support vector machines*; *Neural networks*;

## Keywords

Natural Language Processing, Machine Learning, LSTM, RNN, SVM, Abstract, Evaluation

## 1 Einleitung

Technischer Fortschritt sorgt für immer mehr Daten, welche sich aus unterschiedlichen Themengebieten zusammensetzen. Die Aufmerksamkeitsspanne der Menschen geht mit den Jahren immer mehr zurück, was unter anderem auch an der kurzlebigen Interaktion von sozialen Medien hängt [18]. Lange Texte werden weniger gelesen als kurze. Twitter hat dieses Konzept bereits sehr früh umgesetzt und bietet nur Kurztexte mit 280 Zeichen an. Dieses Prinzip kann ebenfalls auf wissenschaftliche Veröffentlichungen übertragen werden. Abstracts bieten einen schnellen Überblick über die Methodik, Ergebnisse und Diskussion zu unterschiedlichen Fachgebieten. Sie spielen eine wichtige Rolle für die

---

Betreuer/-in Hochschule

Prof. Dr. rer. medic. Christian Thies  
Hochschule Reutlingen  
Christian.Thies@Reutlingen-University.de

Informatics Inside  
24. November 2021 Hochschule Reutlingen  
Copyright 2021 Philip Storz

wissenschaftliche Arbeit, da sie durch diesen kurzen Überblick die Wissensfindung verbessern. Daten sind für den Menschen die Grundlage seiner Entscheidungsfindung, weshalb diese in einer strukturierten Form vorhanden sein müssen, um korrekt interpretiert zu werden [19].

Daten können in den unterschiedlichsten Formen auftreten. Eine der am häufigsten verwendeten Formen in der menschlichen Kommunikation ist die natürliche Sprache. Diese Sprache hat sich auf eine natürliche Weise entwickelt und ist nicht aus einer Planung oder Modellierung heraus entstanden [3]. Menschen verstehen diese Art der Kommunikation und der Daten instinktiv und können diese ebenfalls in textlicher oder gesprochener Form verwenden. Eine Kommunikation zwischen Menschen und Maschine basierend auf natürlicher Sprache ist jedoch mit mehreren Schwierigkeiten behaftet. Die Bedeutung der einzelnen Sätze, die Struktur und Formulierung sowie der regionale Unterschied der Sprache sind ein paar Beispiele hierfür. Um diese Schwierigkeiten zu umgehen kommt das Natural Language Processing (NLP) zum Einsatz, eine Technik mit welcher Maschinen natürliche Sprache verstehen und verarbeiten können [18].

Bei NLP handelt es sich um ein Teilgebiet der Informatik, welches erstmals in den 1960er Jahren erforscht wurde [1]. Die Technik zielt darauf ab, natürliche Sprachen zu analysieren, um ihre Bedeutung zu verstehen und zu extrahieren [3]. Die Analyse kann morphologisch, syntaktisch oder semantisch erfolgen. Eine morphologische Analyse bezieht sich hauptsächlich auf die Struktur der Wörter, während die syntaktische und die semantische Analyse die Beziehung und Bedeutung der Wörter im Bezug auf ihren Syntax und ihre Semantik betrachten [3]. NLP ist ein breit erforschtes Gebiet, welches den Hauptfokus auf die englische Sprache gelegt hat [6]. Bei der Verarbeitung einzelner oder mehrerer Sprachen wird in den meisten Fällen eine

Grundsprache definiert, in diese vor der Auswertung übersetzt wird [10].

Für NLP gibt es ein breites Anwendungsgebiet, welches über die Jahre entstanden ist. Allem voran können Übersetzer oder Spracherkennungssysteme dabei helfen Texte aus anderen Sprachen besser zu verstehen. Zur besseren Übersicht und Organisation von vielen Informationen können Texte kategorisiert, zusammengefasst oder analysiert werden. Zur Kategorisierung zählt unter anderem auch die Spam-Filterung von E-Mails. [18, 19]

In dieser Arbeit werden die folgenden Forschungsfragen untersucht:

- RQ1:** Welche NLP-Methoden eignen sich für die Textklassifizierung?
- RQ2:** Kann ein Klassifikator anhand des Abstracts einer Veröffentlichung das Themengebiet vorhersagen?

In einer anfänglichen Literaturrecherche wird die Forschungsfragen RQ1 untersucht. Daraufhin wird eine eigene Implementierung einen Vergleich von unterschiedlichen Techniken aufzeigen, um RQ2 zu beantworten. Dabei werden die Techniken verwendet, um Abstracts von wissenschaftlichen Veröffentlichungen zu kategorisieren. Diese Arbeit ist in die folgenden Kapitel unterteilt: In Kapitel 2 wird die Methodik vorgestellt, mit welcher die Ergebnisse für die Literaturrecherche und der eigenen Implementierung erreicht wurden. In den Ergebnissen der Literaturrecherche wird eine Textverarbeitungs-Pipeline (Kapitel 3.1) und die Extraktion von Merkmalen (Kapitel 3.2) erklärt. Anschließend werden unterschiedliche Techniken zur Erkennung von Thematiken aus Texten erklärt (Kapitel 3.3) und ein Vergleich von unterschiedlichen State-of-the-Art Veröffentlichungen gezogen (Kapitel 3.4). Nach den Ergebnissen der eigenen Implementierung (Kapitel 3.5) folgen eine Diskussion (Kapitel 4) und ein abschließender Ausblick (Kapitel 5).

## 2 Methoden

Die Literaturrecherche wurde in den drei Suchportalen *EDDI*, *Web of Science* und *Google Scholar* durchgeführt. Dabei wurden unterschiedliche Kombinationen der nachfolgenden Suchbegriffe verwendet.

- Natural Language Processing OR NLP
- Machine Learning
- Evaluation OR Comparison
- Classification OR Categorization

Ziel war es, nicht nur State-of-the-Art-Systeme zu finden, sondern ebenfalls einen Überblick über die Forschung und die Verarbeitung natürlicher Sprachen zu erhalten. Kriterien für die Quellen sind, dass es sich um eine wissenschaftliche Veröffentlichung handelt, diese sich mit NLP beschäftigt und optional einen Vergleich zwischen mehreren Methoden oder eine Evaluation einer Methode beinhaltet. Die Quellen werden in einem ersten Suchlauf mit den Suchbegriffen gesammelt und anhand des Titels und der Schlagwörter vorausgewählt. Nach der Vorauswahl werden die Abstracts jeder Quelle betrachtet und anhand dieser weitere Quellen aussortiert. Alle übrigen Quellen werden ausgearbeitet und gegebenenfalls anhand des Inhaltes aussortiert. Bei der eigenen Implementierung werden Methoden zur Erkennung der Thematik in englischen Abstracts implementiert und anhand von Metriken miteinander verglichen. Die verwendeten Metriken und verglichenen Algorithmen hängen von den Ergebnissen der Literaturrecherche ab.

## 3 Ergebnisse

Nachfolgend werden die Ergebnisse der Literaturrecherche und der eigenen Implementierung gezeigt. Es wurden in der Literaturrecherche 113 Quellen gefunden, welche das erste Auswahlkriterium erfüllen. Davon wurden nach der zweiten Auswahl 65 Quellen aussortiert, da diese zu spezielle Themen betrachteten

oder keine Evaluation von Methoden beinhalten. Von den 48 Quellen wurden abschließend 26 aufgrund ihres Inhaltes ausgewertet, wobei 14 dieser Quellen State-of-the-Art-Systeme zeigen und Methoden miteinander verglichen.

### 3.1 Textvorverarbeitung

Eine große Herausforderung von NLP ist die Mehrdeutigkeit von unstrukturierten Texten natürlicher Sprache [3]. Die englische Sprache besitzt viele unterschiedliche Wortarten [18]. Wörter können in unterschiedlichen Kombinationen auftreten, was die Interpretation des Textes durch eine Maschine erschwert. Die Textvorverarbeitung hilft dabei Teile eines Rohtextes zu bestimmen und zu entfernen, um das Rauschen der Daten zu reduzieren [9]. Dadurch werden uninformative Elemente bei der Interpretation und dem Lernprozess ausgelassen, um eine bessere Auswertung zu erzielen.

Je nach Art der Daten können unterschiedliche Vorverarbeitungsmethoden angewandt werden. Ist der Text nicht in der gewünschten Sprache vorhanden, so muss dieser zuerst übersetzt werden. Zur automatischen Erkennung und Übersetzung von Sprachen können entweder die Anzahl der typischen Wörter einer Sprache gezählt oder die Häufigkeit der typischen Buchstabenfolgen analysiert werden. Bei Ersterem werden Wortarten wie Artikel, Präpositionen, Pronomen oder Adverbien betrachtet. Letzteres nutzt n-Gramme und analysiert die Häufigkeit wie oft ein n-Gramm auftritt. Ein n-Gramm ist eine Folge von n aufeinanderfolgenden Buchstaben, wobei ein Monogramm einen Buchstaben, ein Bigramm zwei Buchstaben und ein Trigramm drei Buchstaben enthält. Die Häufigkeitsverteilung der n-Gramme ist für jede Sprache typisch, weshalb diese Methode dafür geeignet ist, um Sprachen zu erkennen. [6] In einem weiteren Schritt können Elemente des Textes entfernt werden, welche der

Strukturierung dienen. Hierzu zählen HTML-Elemente, Überschriften und Kopf- oder Fußzeilen. Diese Elemente bieten der Auswertung des Textes keinen Mehrwert. [16]

Ein sehr wichtiger Schritt der Textvorverarbeitung ist die Tokenisierung. Die Textdaten werden dabei in kleinere aussagekräftige Komponenten zerlegt [19]. Dies kann mit der Hilfe von regulären Ausdrücken (engl. *regular expressions* - RegEx) durchgeführt werden [9]. Man unterscheidet dabei zwischen der Tokenisierung von Sätzen und von Wörtern. Bei der Aufteilung in Sätze wird das gesamte Textdokument unterteilt. Dabei werden auf Trennzeichen wie Punktation, Ausrufe- und Fragezeichen, Zeilenumbrüche und Semikola geachtet [18]. Diese Sätze werden dann weiter in einzelne Wörter unterteilt. Dabei werden Leerzeichen verwendet und Kommata entfernt [18].

Der tokenisierte Text muss im nächsten Schritt standardisiert werden. Dazu zählt die Umwandlung aller Zeichen in ASCII-Zeichen [18]. Dadurch wird aus einem È ein E oder aus einem Ä eine Ae. Diese Umwandlung ist jedoch nicht in jeder Sprache möglich und nötig. In normalen Sätzen werden häufig Abkürzungen von Wörtern verwendet, um diese Texte kompakter zu gestalten. Bei der Analyse des Textes sind diese Abkürzungen jedoch hinderlich, weshalb sie entweder entfernt oder expandiert werden müssen. Bei Letzterem werden gesammelte Listen genutzt, welche für eine Sprache typische Abkürzungen enthalten, und damit iterativ alle Abkürzungen im Text ersetzt [18]. Eine weitere Anwendung von RegEx ist das Entfernen von Punktationen und spezielle Zeichen [23]. Hierunter fallen unter anderem @, €, % oder &. Zur Verbesserung der Interpretierbarkeit kann der Text ebenfalls entweder in Groß- oder Kleinbuchstaben umgewandelt werden [7].

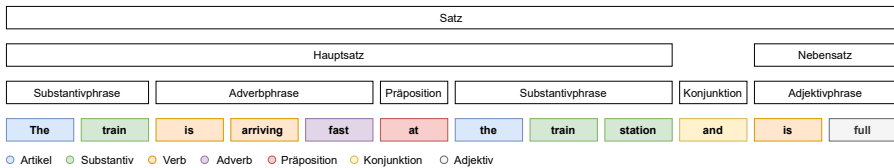
Abhängig von der Quelle der Textdaten kann es vorkommen, dass eine bestimmte Schreibweise von Wörtern gegeben ist und der Text

Schreibfehler enthält. Texte aus sozialen Medien besitzen häufig langgezogene Wörter, bei welchen ein Buchstabe in einem Wort mehrmals wiederholt wird, um dem Wort mehr Ausdruck zu verleihen. Diese Wörter müssen zurück in die grammatikalisch korrekte Form gebracht werden. Dafür wird kontinuierlich ein wiederholender Buchstabe entfernt und geprüft, ob das Wort korrekt ist. Bei Schreibfehlern werden mit Hilfe von Datenbanken Kandidatenwörter gesucht, welche die höchste Übereinstimmung mit dem geschriebenen Wort besitzen. Das Wort mit der wahrscheinlichsten Übereinstimmung anhand von Buchstabenfolgen, Anzahl an Zeichen oder n-Grammen wird als Korrektur eingetragen. [18]

Ein Wort besitzt mehrere Wortformen, welche unterschiedliche Bedeutungen besitzen. Für die Analyse des Textes ist es sinnvoll das Wort auf seinen Wortstamm zurückzuführen. Eine häufig verwendete Variante des Stemming (Stammformreduktion) entfernt Affixe vom Wortstamm [11]. Ein Affix ist ein unselbstständiger Wortbestandteil, welcher vor (Präfix) oder nach (Suffix) einem Wort angebracht werden kann, um dessen Bedeutung zu ändern [18]. Beim Stemming wird beispielsweise aus *jumping* nach Entfernung des Suffixes *-ing* der Wortstamm *jump*.

Eine dem Stemming ähnliche Methode ist die Lemmatisierung. Dabei werden alle Wörter in ihre Grundform (Lemma) umgewandelt [9]. Dabei werden Wörter zusammengeführt, welche die gleiche grundsätzliche Bedeutung besitzen. Diese unterscheiden sich häufig in der Mehrzahl eines Wortes. Ebenfalls werden Synonymen und Homonymen aufgelöst, um doppelten Entitäten zu entfernen [3].

Nachdem alle Wörter in eine einheitliche Form gebracht wurden, können nun unnötige Token und Stoppwörter entfernt werden. Diese haben meistens keine Bedeutung für den Satz und können anhand von Listen entfernt werden [16]. Dabei handelt es sich zu meist um Artikel, Adverbien, Konjunktionen,



**Abbildung 1: Englischer Satzbau mit Part of Speech Tagging**

Pronomen, Präpositionen oder Hilfsverben [6].

Unterschiedliche Wortarten werden auch *Parts of Speech* genannt und können in einem Text mit einem Part of Speech Tagger gekennzeichnet werden [1]. Dies ist ein Prozess, der als Klassifizierung verwendet wird, um Wörter nach ihrer Wortart zu filtern und ihre Wichtigkeit darzustellen [18]. Beim Tagging müssen zusätzlich benachbarte Begriffe beachtet werden, damit die eindeutige Bedeutung des Wortes erkannt wird [6]. Ein Beispiel ist in Abbildung 1 zu sehen.

Das *Shallow Parsing/Chunking* ist eine Technik, mit welcher die Struktur eines Satzes analysiert wird und einzelne Bestandteile in übergeordnete Phrasen zusammengefasst werden. Das Hauptziel ist es dabei semantisch sinnvolle Phrasen zu erhalten, um die Beziehung zwischen diesen zu beobachten. Dies hilft dabei Subjekt-Verb-Objekt-Reihenfolgen genauer zu erkennen. [18]

### 3.2 Feature-Engineering

Die vorverarbeiteten Rohdaten müssen nun nach wichtigen Merkmalen untersucht werden. Der Prozess des Feature-Engineering bietet Techniken, um diese Rohdaten in Merkmale (engl. features) umzuwandeln. Herausforderungen von unstrukturierten Daten wie unvorhersehbare Inhalte oder Syntax, sowie das Format des Textes machen das Erstellen von Merkmalen aus Rohdaten nötig. Dadurch ist auch eine bessere Darstellung der Daten, mehr Flexibilität und ein besseres Modell beim Lernprozess möglich. [19]

Merkmale sind eindeutig messbare Eigenschaften jedes Datenpunktes in einem Datensatz. Von Maschinen verwendbare Daten befinden sich immer in numerischer Form, da diese interpretierbar sind. Viele Feature-Engineering-Modelle beruhen auf dem Prinzip textuelle Informationen in numerische Vektorraum-Modelle umzuwandeln. Ein Vektorraum besitzt eine Anzahl an Wörtern, welche sich über unterschiedliche Dokumente erstreckt. Ein Dokument besteht dementsprechend aus  $n$  gewichteten Wörtern. Das Gewicht kann unter anderem die Frequenz des Auftretens sein. [19]

Feature-Engineering-Modelle können in traditionelle und fortgeschrittene Modelle unterteilt werden. Traditionelle Modelle verwenden herkömmliche zählungsbasierte Strategien, welche mit algebraische und statistischen Methoden erstellt werden [18]. Ein grundlegendes Vektorraum-Modell ist das Bag-of-Words-Modell (BoW). Dabei sind alle Dokumente als numerischer Vektor dargestellt [2]. Jede Dimension des Vektors ist ein bestimmtes Wort aus dem Textkorpus. Ein Korpus ist eine Sammlung von Textdokumenten mit unterschiedlichen Themen und Subjekten [18]. Der im Vektor angegebene Wert kann die Häufigkeit, das Vorkommen oder einen gewichteten Wert des Wortes annehmen. Die Reihenfolge des Auftretens von Wörtern oder die zugrundeliegende Grammatik und Syntax werden beim BoW-Modell nicht beachtet, sondern es wird lediglich das Verhältnis der einzelnen Worte zu den unterschiedlichen Dokumenten festgestellt [18].



Das BoW-Modell kann zu einem Bag-of-N-Grams-Modell erweitert werden. In diesem Fall ist ein N-Gramm eine Sammlung von Wort-Token, welche in einem Textdokument zusammenhängend auftreten. Ein Bi-Gramm repräsentiert in diesem Fall die Reihenfolge zweier Wörter. Häufig miteinander auftretende Wortkombinationen werden bei dieser Methode in einzelnen Beuteln gesammelt und analysiert. [18] Bei sehr großen Textkorporas haben BoW-Modelle Probleme, um kleinere Abhängigkeiten zu erkennen. Die Merkmalsvektoren basieren auf der Häufigkeit eines Begriffes, wodurch weniger häufig auftretende Wörter, welche für bestimmte Kategorien wichtig sind, überschattet werden [18].

*Term frequency-inverse document frequency* (TF-IDF)-Modelle kombinieren Matrizen aus der Wortfrequenz und der inverse Dokumentenfrequenz, um dieses Problem zu lösen [17]. Die Wortfrequenz ist dabei der Wert, welcher durch das BoW-Modell berechnet wird. Die Dokumentenfrequenz berechnet sich durch eine logarithmische Skalierung für jeden Begriff und dessen Häufigkeit in allen Dokumenten des Korpus [17]. Die Normalisierung der TF-IDF-Matrix kann durch die L2-Norm erfolgen [18].

Aufbauend auf dem BoW- und TF-IDF-Modell kann die Ähnlichkeit zweier Dokumente festgestellt werden. Dabei werden ähnlichkeitsbasierte Metriken wie der Kosinus-Abstand oder der euklidische Abstand verwendet, um den Abstand oder Winkel zwischen den Dokumentenvektoren zu bestimmen. Je kleiner diese Messung ist, desto ähnlicher sind die beiden Dokumente. [18]

Neuere und fortgeschrittene Feature-Engineering-Modelle versuchen Nachteile des BoW-Modells zu umgehen, indem sie Wortvektoren in den kontinuierlichen Vektorraum anhand von semantischen und kontextuellen Ähnlichkeiten einbetten [18]. Diese Worteinbettungen werden häufig für maschinelle Lernmethoden verwendet, da es

ihnen ermöglicht rohe Textdaten zu kontextualisieren [23]. Die Einbettungsmethoden beziehen sich dabei fast ausschließlich auf neuronale Netzen [1].

Zu den beiden gebräuchlichsten Worteinbettungsmethoden zählen Word2Vec und GloVe [2]. Word2Vec betrachtet die Struktur und die Platzierung von Wörtern in ihrem Korpus [17]. Es basiert auf Deep Learning und berechnet verteilte und kontinuierlich dichte Vektordarstellungen von Wörtern. Durch die geringe Dimensionalität im Vergleich zu traditionellen BoW-Modellen kann eine bessere Erfassung der kontextuellen und semantischen Ähnlichkeit erfolgen [18]. Word2Vec kann dabei zwei unterschiedliche Modellarchitekturen nutzen. Ein *Continuous BoW-Modell* sagt ein Zielwort anhand der umgebenen Wörter vorher. In einem neuronalen Netz sind dabei die Kontextwörter die Eingaben und das Zielwort die Ausgaben [17]. Das *Skip-Gram-Modell* hingegen nutzt das Zielwort, um die umliegenden Wörter vorherzusagen. Dabei werden Zielwort-Kontextwörter-Paare in einzelne Zielwort-Kontextwort-Paare umgewandelt, von welchen die Relevanz errechnet wird [17].

Das *Global Vectors* (GloVe)-Modell ist ein unüberwachtes Lernmodell, mit welchem dichte Wortvektoren ermittelt werden. Es wird dazu verwendet, um Nachteile anderer Modelle, wie der schlechten Verwendung von Wortanalogien und statistischen Informationen, zu umgehen. Das GloVe-Modell bildet eine Wort-Kontext-Cocurrence-Matrix, welche aus einzelnen Wort-Kontext-Paaren besteht. Jedes Element in der Matrix stellt dar, wie oft ein Wort in einem Kontext vorkommt. Die Matrix wird durch eine Matrix-Faktorisierung approximiert und setzt sich aus einer Wort-Merkmal-Matrix und Merkmal-Kontext-Matrix zusammen. [18]

### 3.3 Erkennung von Themen-gebieten

Für die Herausarbeitung von Themen aus einem Text können bei NLP unterschiedliche Ansätze gewählt werden. Bei der Klassifizierung werden Texte abhängig von ihrem Inhalt in bestimmte Kategorien eingeordnet. Klassifikationsmodelle nutzen in der Regel überwachtes Lernen, da bereits vorgegebene Klassen benötigt werden, um Daten zuzuweisen. Es gibt jedoch auch unüberwachte Methoden wie das Clustering, mit welchem gleiche Themen gruppiert werden können. Neben der Klassifikation von Texten kann auch die Zusammenfassung der Inhalte als eine Methode angewendet werden. Dabei werden Schlüsselsätze und -wörter aus den Dokumenten extrahiert und so eine Übersicht über das Thema erhalten. [18]

Bei der überwachten Klassifikation wird zuvor manuell ein Trainingsdatensatz erstellt. Auf diesem wird ein Modell durch einen Klassifikationsalgorithmus gelernt, welches danach auf einen neuen Testdatensatz angewendet werden kann. Bei der Aufteilung des Trainings- und Testdatensatzes muss darauf geachtet werden, dass der Trainingsdatensatz groß genug ist, um eine Überanpassung des Modells zu verhindern. Bei einer Überanpassung ist der Testdatensatz zu spezifisch

und das trainierte Modell kann nicht auf viele Szenarien angewendet werden. Klassifizierungen können binär, mehrklassig und mehrkategorisch mit mehreren Klassen pro Element sein. Eine genauere Übersicht über die überwachte Klassifikation ist in Abbildung 3 zu sehen. [18]

Anhand des Testdatensatzes kann das Modell evaluiert werden. Für die Evaluation gibt es unterschiedliche Metriken, wobei die meisten auf den Einträgen einer Konfusionsmatrix basieren. Eine erweiterte Konfusionsmatrix ist in Abbildung 2 zu sehen. Die Konfusionsmatrix basiert auf dem Verhältnis von vorhergesagten und tatsächlichen positiven oder negativen Zuständen. Aus den vier Kombinationen richtig positiv, falsch negativ, falsch positiv und richtig negativ können weitere Metriken errechnet werden. Nachfolgend werden die in der Literatur am häufigsten verwendeten Metriken genannt.

Für die Evaluation von Modellen wird häufig die Sensitivität und Spezifität berechnet [2, 7, 8, 10, 22]. Die Sensitivität gibt an, zu wieviel Prozent positive Zustände richtig erkannt wurden. Die Spezifität hingegen misst, wieviel Prozent der falschen Zustände richtig zugeordnet wurden. Die positiven und negativen Vorhersagewerte werden häufig verwendet, um den Anteil der korrekt als

<sup>1</sup>[https://en.wikipedia.org/wiki/Confusion\\_matrix](https://en.wikipedia.org/wiki/Confusion_matrix)

		Vorhersage		Informiertheit	Prävalenzschwelle
		Positiv	Negativ		
Resultat	Alle Elemente				
	Positiv	Richtig positiv	Falsch negativ	Sensitivität	Fehlerrate
Negativ	Falsch positiv	Richtig negativ	Ausfall	Spezifität	
Prävalenz	Positiver vorhergesagter Wert	Falsche Auslassungsrate	Positives Wahrscheinlichkeitsverhältnis	Negatives Wahrscheinlichkeitsverhältnis	
Gemauigkeit	Falsche Erkennungsrate	Negativer vorhergesagter Wert	Markertheit	Diagnosewahrscheinlichkeitsverhältnis	
Angepasste Genauigkeit	F1-Score	Fowkes-Mallows Index	Matthews Korrelationskoeffizient	Jaccard-Index	

Abbildung 2: Erweiterte Konfusionsmatrix mit berechenbaren Metriken<sup>1</sup>

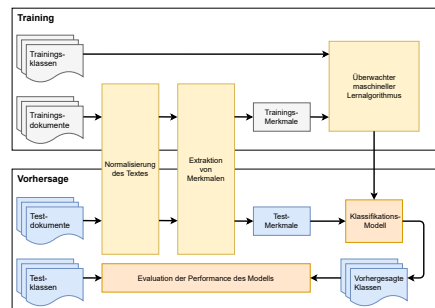


Abbildung 3: Klassifikation mit einem überwachtem Lernalgorithmus nach [18]

positiv oder negativen klassifizierten Ergebnisse an der Gesamtheit der als positiv bzw. negativen klassifizierten Ergebnisse zu erkennen [2, 7, 8, 10, 22]. Die Genauigkeit gibt an, wieviel Prozent aller tatsächlichen Zustände auch richtig zugeordneten wurden [7, 8]. Der F1-Score ist eine kombinierte Metrik aus dem positiven Vorhersagewert und der Sensitivität. Er gibt an, wie gut die Kombination seiner beiden Einzelteile ist [7, 22].

Basierend auf den Evaluationsmetriken können die Hyperparameter der Klassifikationsalgorithmen angepasst werden. Jedes Modell ist eine mathematische Funktion mit Parametern, welche die Komplexität, das Lernen und weitere Eigenschaften beeinflussen. Durch das Tuning dieser Parameter kann ein besseres Klassifikationsergebnis erhalten werden. [18].

Für die Klassifikation können unterschiedliche Algorithmen verwendet werden. Ein *naiver Bayes*-Klassifikator nutzt das Bayes-Theorem mit dem naiven Ansatz, dass jedes Merkmal unabhängig voneinander ist. Es handelt sich hierbei um einen generativen Klassifikator, welcher die Zusammensetzung der einzelnen Klassen analysiert und jedem Dokument eine Klasse zuordnet, indem er die Wahrscheinlichkeit der einzelnen Klassen berechnet [17]. Der naive Bayes-Klassifikator kann multinomial sein und dadurch mehr als nur eine Klasse in seiner Vorhersage verwenden [18].

Die logistische Regression ist ein Klassifikator, welcher die logistische Sigmoidfunktion der gewichteten Summe der Merkmale und einen Bias-Term nutzt, um die Klassifizierung zu bestimmen [4]. Der diskriminierende Klassifikator bestimmt, welche Merkmale einzelne Klassen am stärksten unterscheiden und schätzt dadurch die Wahrscheinlichkeit der Klassen [17].

Eine *Support-Vektor-Maschine* ist ein Klassifizierungsalgorithmus, der Stichproben als Punkte im Raum darstellt [4]. Die zusammengehörigen Punktecluster werden durch

eine Hyperebene voneinander getrennt, welche den maximalen Abstand der beiden Cluster angibt [17]. Neue Datenpunkte werden den einzelnen Klassen anhand der Seite der Hyperebene zugeordnet [18]. Die äußersten Punkte in jeder Klasse werden Supportvektoren genannt [17].

Ensemble-Modelle sind zusammengesetzt aus unterschiedlichen kleineren Modellen. Die einzelnen Teile sind dabei in der Regel nicht in der Lage alleine genaue Vorhersagen zu treffen [18]. Eine Kombination aus mehreren Teilen, wie bei einem *Random Forest* aus vielen einzelnen Entscheidungsbäumen, kann zu verlässlichen Vorhersagen führen [4]. Es wird hierbei zwischen *Bagging* und *Boosting* unterschieden. Bagging teilt die Daten in einzelne Teile auf und nutzt diese in einem parallelen Training [18]. Jedes Modell erzeugt dann eine Vorhersage und der Mittelwert aller Vorhersagen ergibt die endgültige Vorhersage. Beim Boosting werden die Modelle sequentiell mit den Daten des vorherigen Modells versorgt [18]. Die Modelle führen daraufhin eine Fehlerkorrektur der Vorhersage des vorherigen Modells durch. Ein Beispiel hierfür ist die *Gradient Boosting Machine*, welche ein additives Modell als Kombination aus Modellen und Verlustfunktionen bildet [18].

### 3.4 State-of-the-Art-Systeme

Von den 14 ausgewerteten Quellen mit State-of-the-Art-Systemen ist das hauptsächliche Anwendungsgebiet die Medizin mit sieben von 14 Quellen. Bei den implementierten Methoden ist ein breiteres Feld aufgetreten. Drei Quellen verwenden Entscheidungsbäume und drei Quellen Worteinbettungsalgorithmen. Jeweils zwei Quellen implementieren SVM, Bayes, RNN/CNN, Random Forests oder LSTMs. Eine Übersicht der Quellen ist in Tabelle 1 zu sehen.

Quelle	Themengebiet	Evaluierte/Verglichene Methode(n)
Barash et al. [2]	Medizin	Bag-of-Words-Modell Word2Vec-Modell LSTM-Modell Aufmerksamkeitsmodell
Carrodeguas et al. [4]	Medizin	SVM Random Forest Logistische Regression LSTM-RNN
Gagnon et al. [7]	Medizin	Varianten von CNNs
Gavriellov-Yusim et al. [8]	Soziale Medien	Co-Occurrence-Methode Algorithmus zur Erkennung benannter Entitäten
Horecki et al. [9]	Automatisierte Vorhersage	Max-Entropy-Klassifikator Naïver Bayes-Klassifikator Entscheidungsbaum MLP-Klassifikator
Jungnickel et al. [10]	Sprachübergreifende Textverarbeitung	TF-IDF Repräsentativer überwachter Keyphrase-Extraktionsalgorithmus Graphenbasierter Keyphrase-Extraktionsalgorithmus
Lansley et al. [13]	Social Engineering	Entscheidungsbäume Random Forest MLP
Lewis et al. [14]	Informations-extraktion	Entscheidungsbaum Bayesian-Klassifikator
Richard et al. [17]	Medizin	SVM mit TF-IDF Wortembeddings mit TF-IDF
Shree et al. [20]	Natural Language Processing	Varianten von Wortsinndisambiguierungen
Solti et al. [21]	Radiologie	Regelbasierter Algorithmus Maximum Entropy
Taggart et al. [22]	Medizin	Regelbasierter Algorithmus SVM Extra-Trees-Klassifikator CNN
Wang et al. [23]	Biomedizin	Word2Vec-Modell mit einer Skip-Gram-Architektur
Yin et al. [24]	Natural Language Processing	CNN LSTM-RNN GRU-RNN

**Tabelle 1: State-of-the-Art-Systeme**

Lewis [14] vergleicht Entscheidungsbäume mit einem Bayes Klassifikator. Die Algorithmen filtern aus Nachrichtenartikeln Informationen heraus und kategorisieren diese. Entscheidungsbäume erzielen sowohl bei dem positiven Vorhersagewert als auch bei der Sensitivität ein besseres Ergebnis. Auffällig ist, dass beim Bayes-Klassifikator nach zehn Merkmalen bereits ein Abfall der Performance zu sehen ist. Entscheidungsbäume hingegen können über 90 Merkmale verwenden. Ein Word2Vec-Modell mit einer Skip-Gram-Architektur wird von Wang et al. [23] eingesetzt, um aus klinischen Notizen, biomedizinischen Publikationen, Wikipedia und Nachrichten zu lernen. Das Hinzufügen von Worteinbettungen liefert verbesserte Werte, eine Auswahl zwischen wissenschaftlichen und öffentlichen Quellen erzeugt jedoch keinen Unterschied beim Training.

In [21] vergleichen Solti et al. einen regelbasierten Algorithmus mit maschinellen Lernalgorithmen, um Röntgenreports zu klassifizieren. Der auf Schlüsselwörtern und einfachen Heuristiken basierte Algorithmus schneidet im Vergleich zur Maximum Entropy und 6-Gramm-Systemen sowohl beim F1-Score als auch bei einzelnen Werten schlechter ab. Zudem wird hervorgehoben, dass regelbasierte Systeme weniger flexibel sind bei der Anwendung auf unterschiedliche Gebiete.

Verschiedene Arten von neuronalen Netzen werden von Yin et al. [24] verglichen. Dabei schneiden das LSTM-RNN und ein *Gated Recurrent Unit*-RNN in einem breiten Spektrum besser als das CNN. Bei sowohl CNNs als auch RNNs sind jedoch die versteckte Größe und die Stapelgröße entscheidend für eine gute Performance.

NLP-Techniken können für die Erkennung von Blutungen in medizinischen Notizen eingesetzt werden. Taggart et al. [22] vergleichen einen regelbasierten Ansatz mit drei unterschiedlichen maschinellen Lernalgorithmen. Dabei werden die Datensätze einmal unbearbeitet und einmal herunterskaliert

verwendet, um den Effekt von einer eventuellen Überanpassung zu beobachten. Der herunterskalierte Extra-Trees-Algorithmus erreicht zusammen mit dem regelbasierten Modell die beste Sensitivität, wobei dieses zusätzlich den besten positiven Vorhersagewert besitzt. Die SVM erhält bei beiden Datensätzen die beste Spezifität. Trotzdem erzielt das regelbasierte Modell allgemein eine bessere Performance als die maschinellen Lernalgorithmen. Taggart et al. begründen dies mit den eindeutig definierten Begriffen, welche auf Blutungen hinweisen, wodurch maschinelle Algorithmen zu umfangreich wären. [22]

Gagnon et al. [7] vergleichen drei unterschiedliche Architekturen eines CNNs für die automatische Erkennung von mentalen Statusveränderungen in Notizen von Ärzten in Notaufnahmen. Jedes Modell nutzt Worteinbettungen mithilfe des Word2Vec-Algorithmus. Das mehrkanalige CNN ist in den meisten gemessenen Metriken, einschließlich Genauigkeit und Verlust besser als die beiden anderen Modelle. Das sequentielle CNN ist in allen Kategorien besser als das aufmerksame CNN.

Ein weiteres System, welches in der Medizin angewendet wird, wurde von Barash et al. [2] entwickelt. Dieses klassifiziert anhand eines NLP-Algorithmus CT-Berichte des Kopfs. Untersucht wurden ein LSTM-Algorithmus und ein DL-Algorithmus mit BoW. Das LSTM wurde mit einer Aufmerksamkeitsarchitektur und Worteinbettungen erweitert. Diese Kombination schlägt oder gleicht dadurch auch den DL-Algorithmus in fünf von sechs Metriken aus.

Zur Identifizierung von unerwünschten Ereignissen und Berichterstattungen von patientengenerierten sozialen Medieninhalten vergleichen Gavrielov-Yusim et al. [8] eine merkmalsbasierte statistische Lernmethode mit einer Ko-Ereignis-Methode. Die Gesamtgenauigkeit der NLP-Methode übersteigt die der Ko-Ereignis-Methode mit 61,5 % zu 52,2%.

Dies liegt an der besseren Spezifität und dem positiven Vorhersagewert.

NLP-Algorithmen arbeiten meistens in einer Grundsprache, um das beste Ergebnis zu erhalten. Jungnickel et al. [10] untersuchen in diesem Kontext wie maschinelle Übersetzungssysteme und andere Techniken zur Verarbeitung natürlicher Sprache kombiniert werden können, um das beste Ergebnis zu erzielen. Der Multipartite-Rank-Algorithmus erzielt die besten Ergebnisse, wobei der Schlüsselsatzextraktions-Algorithmus die TF-IDF-Methode leicht übertrifft. Die Ergebnisse sind jedoch sehr abhängig von dem gewählten Übersetzer und der Reihenfolge von Übersetzung und Verarbeitung. Zudem ist die Leistung der Algorithmen zwischen ein- und mehrsprachigen Daten stark unterschiedlich.

Horecki et al. [9] finden bei ihrer Arbeit heraus, dass der naive Bayes-Klassifikator sich durch seine hohe Trainingsgeschwindigkeit, Möglichkeit auf geringen Datensätzen zu arbeiten und hohen Klassifikationsgenauigkeit am besten für die Klassifikation eignet. Zusätzlich legen die Autoren einen Merkmalsfilter und eine große Menge an Merkmalen nahe, wenn eine gute Klassifikation erreicht werden soll.

Wortsinnaufschlüsselungen versuchen die korrekte Semantik von Wörtern und Satzteilen zu erfassen. Shree et al. [20] vergleichen unterschiedliche wissensbasierte, auf maschinellem Lernen basierende und hybride Ansätze für WSD miteinander. Wissensbasierte Ansätze sind auf große Wissensressourcen angewiesen, welche es nicht in jeder Sprache gibt. Zusätzlich sind diese Ansätze leistungsschwächer als andere. Überwachte Ansätze hingegen werden aufgrund ihrer genauen Ergebnisse häufig verwendet, jedoch ist hierbei die Wissensbeschaffung schwieriger. Shree et al. kommen zu dem Schluss, dass hybride Systeme eine optimale Leistung besitzen und dadurch sich am besten eignen.

Lansley et al. [13] haben eine Methode zur Erkennung von Social-Engineering-Angriffen entwickelt. Diese basiert auf natürlicher Sprachverarbeitung und künstlichen neuronalen Netzen. Verglichen wurden Entscheidungsbäume, Random Forests und ein Multi-Layer-Perceptron (MLP), welches die höchste Genauigkeit bei zwei Datensätzen erzielt. Ein MLP ist ein vollständig verbundenes künstliches neuronales Netz mit einer Eingabe-, einer Ausgabe- und mindestens einer verborgenen Schicht.

Richard et al. [17] kombinieren die logistische Regression, einen naiven Bayes-Klassifikator und eine SVM mit Worteinbettungen über Word2Vec und der TF-IDF. Dabei wird eine Kategorisierung von Protokollabweichungen in verschiedenen Therapiebereichen miteinander verglichen. Die SVM erzielt mit dabei die höchste Genauigkeit.

Zur Erkennung von Nachfolgeempfehlungen in Radiologieberichten verglichen Carrodeguas et al. [4] traditionelle maschinelle Lernalgorithmen mit einem DL-Modell. Für das LSTM-RNN wurde eine optimale Architektur von 100 Knoten erfasst. Im Vergleich der Performance erzielt jedoch die SVM mit einer Trainingszeit von unter einer Sekunde und durchschnittlich besseren Werten bei dem positiven Vorhersagewert, der Sensitivität und dem F1-Score bessere Ergebnisse als die logistische Regression, der Random Forest und das LSTM-RNN.

Weitere Quellen mit systematischen Literaturrecherchen zum Vergleich von Methoden wurden ebenfalls betrachtet. Casey et al. [5] betrachten 164 Veröffentlichungen und untersuchen die Anwendungsgebiete sowie häufig verwendete Methoden. Das häufigste Anwendungsgebiet ist die Medizin und umfasst die automatische Auswertung von radiologischen oder medizinischen Berichten. Bei den NLP-Methoden ist die SVM mit 34 von 74 die meistverwendete Methode gefolgt von der logistischen Regression, den Random Forests und den naiven Bayes-Klassifikatoren.

Casey et al. heben hervor, dass herkömmliche NLP-Methoden häufiger als DL-Methoden verwendet werden, da diese oft auch als Benchmarktest dienen.

Le Glaz et al. [1] führen eine Zusammenfassung und Charakterisierung von Studien, in denen maschinelles Lernen und NLP-Techniken für die psychische Gesundheit eingesetzt wurden durch. Dabei stellte sich heraus, dass die SVM mit Abstand zu den häufigsten Methoden von maschinellem Lernen gehört. Gefolgt wird sie von der logistischen Regression, Random Forests und einzelnen Entscheidungsbäumen. Die am häufigsten gewählte Plattformen ist Python.

Eine systematischen Literaturrecherche, um bestehende klinische Systeme zur Verarbeitung natürlicher Sprache (NLP) zu identifizieren, die strukturierte Informationen aus unstrukturiertem Freitext generieren wird von Kreimeyer et al. [12] durchgeführt. Als Ergebnis wird der Trend erkannt, dass mehr regelbasierte Ansätze als maschinelle oder hybride Systeme verwendet wurden.

Lopez-Martinez et al. [15] untersuchen ebenfalls Trends und den Stand der NLP-Forschung. Sie finden heraus, dass das Forschungsinteresse sein 2011 um fast das fünf-fache gestiegen ist. Dabei werden immer neue Forschungsgebiete miteinbezogen und erschlossen.

### 3.5 Eigener Vergleich

Basierend auf den ausgewerteten Quellen wurde sich bei der eigenen Entwicklung dafür entschieden eine SVM mit Worteinbettungen, eine SVM mit TF-IDF und ein LSTM-RNN zu vergleichen. Eine zentrale Herausforderung der Umsetzung war die Beschaffung eines geeigneten Korpus von wissenschaftlichen Veröffentlichungen. Die meisten Datensätze sind entweder nicht frei verfügbar, decken nur wenige Themengebiete ab, oder sind nicht ausreichend für die Klassifikation geeignet. Für eine anfängliche Entwicklung der NLP-Pipeline wurde der OA-STM

Korpus<sup>2</sup> von Elsevier mit 110 wissenschaftlichen Artikeln aus zehn Themengebieten verwendet. Mit elf Abstracts pro Kategorie, eignet sich dieser Korpus nicht, um einen umfangreichen Klassifikator zu trainieren. Er diente hauptsächlich des ersten Entwurfs der NLP-Pipeline.

Die Veröffentlichungen des Korpus wurden über den *PlainTextReader* der Natural-Language-Toolkit-Bibliothek<sup>3</sup> eingelesen und für jedes wurde das Abstract extrahiert. In der Vorverarbeitung wurden alle Wörter in Kleinbuchstaben umgewandelt und auf ASCII-Zeichen beschränkt. Daraufhin wurden alle HTML-Tags, Emails, URLs und weitere Sonderzeichen entfernt. Zahlen wurden ebenfalls entfernt. Daraufhin wurden die Abstracts in einzelne Wort-Tokens umgewandelt. Die einzelnen Wörter wurden lemmatisiert und nach Stopwörtern gefiltert. Für die Worteinbettungen wurde Word2Vec verwendet. Das Model wurde auf allen vorverarbeiteten Abstracts trainiert. Es erstellt einen 50-dimensionalen Vektor für jedes Wort. Das trainierte Word2Vec-Modell wird daraufhin verwendet, um die einzelnen Wörter in Worteinbettungssequenzen umzuwandeln. Bei der SVM mit TF-IDF wurden durch den *TfidfVectorizer* der sklearn-Bibliothek<sup>4</sup> die einzelnen Abstracts enkodiert. Für das RNN wurden die einzelnen Wörter aller Abstracts in Zahlen umgewandelt und jedes Abstract in eine einheitliche Länge umgewandelt. Zusätzlich wurde für die Labels ein *One-Hot-Encoding* durchgeführt.

Die vorverarbeiteten Daten wurden in einen Trainings- und Testdatensatz aufgeteilt. 80 % der Daten werden für das Training verwendet und 20 % für das Testen. Die Verteilung der einzelnen Themengebiete ist gleichmäßig.

<sup>2</sup><https://researchcollaborations.elsevier.com/en/datasets/open-access-corpus-of-scientific-technical-and-medical-content>

<sup>3</sup><https://www.nltk.org/>

<sup>4</sup><https://scikit-learn.org/stable>

Für die SVMs wurde jeweils ein linearer Kernel verwendet. Bei dem RNN wurde ein sequentielles Model verwendet, welches aus einer Embedding-Schicht, einer LSTM-Schicht mit 100 LSTM-Zellen und einer Dense-Schicht mit einer Sigmoid-Aktivierung besteht. Als Verlustfunktion wurde die kategorische Kreuzentropie und der Adam-Optimierer verwendet. Zusätzlich wurde ein Early-Stopping-Callback hinzugefügt, welcher nach drei Epochen ohne Verbesserung das Training abbricht.

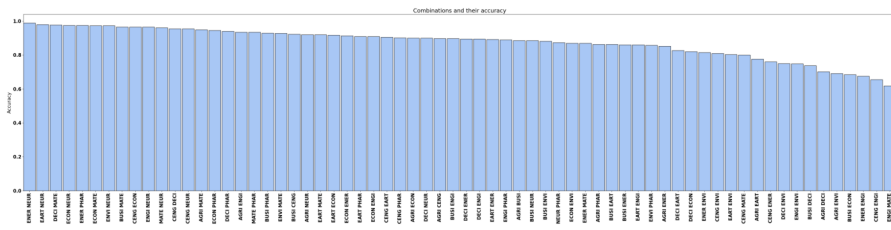
Nachdem die erste Implementierung der NLP-Pipeline abgeschlossen war, wurde ein eigener wissenschaftlicher Korpus erstellt. In Tabelle 2 ist eine Übersicht über den erstellten Korpus zu sehen. Für die Erstellung wurde die Elsevier API<sup>5</sup> verwendet. Der Standard API-Schlüssel von Elsevier bietet die Möglichkeit auf 5000 Veröffentlichungen in der Scopus-Datenbank zuzugreifen. Da es sich bei Scopus jedoch um eine Referenzdatenbank handelt, liegt nicht für jede Veröffentlichung ein Volltext vor. Nach Volltexten wurde mit dem folgenden Suchstring gesucht: SUBJAREA(<topic>) AND OPENACCESS(1) AND LANGUAGE(english) AND (DOCTYPE(ar) OR DOCTYPE(bk) OR DOCTYPE(ch) OR DOCTYPE(cp)). Der Platzhalter <topic> wird dabei von den jeweiligen Themengebieten von Elsevier ersetzt. Eine Übersicht dieser ist ebenfalls in Tabelle 2 zu finden. Des Weiteren wurden nur englische Veröffentlichungen verwendet und Dokumente, welche entweder ein Artikel, ein Buch, ein Buchkapitel oder eine Veröffentlichung auf einer Konferenz sind. Für jedes Themengebiet wurde eine .json-Datei angelegt, welche eine Liste von Abstracts mit dem zugehörigen Titel der Veröffentlichung und einem Label enthält. Der neu erstellte Korpus wurde daraufhin für die Optimierung der Klassifikation verwendet. Die Klassifikation wurde auf alle Themengebiete im Korpus mit 1000 Abstracts limitiert, wodurch eine Klassifikationsproblem

<sup>5</sup><https://dev.elsevier.com/>

Abk.	Themengebiet	Anzahl
ENVI	Environmental Science	2920
ENGI	Engineering	2607
DECI	Decision Sciences	2021
MATE	Material Sciences	1663
ENER	Energy	1630
NEUR	Neuroscience	1295
BUSI	Business, Management and Accounting	1236
CENG	Chemical Engineering	1226
ECON	Economics, Econometrics and Finance	1168
EART	Earth and Planetary Sciences	1061
AGRI	Agricultural and Biological Sciences	1058
PHAR	Pharmacology, Toxicology and Pharmaceutics	1058
PSYC	Psychology	976
SOCI	Social Science	895
COMP	Computer Science	821
NURS	Nursing	765
IMMU	Immunology and Microbiology	674
DENT	Dentistry	656
MATH	Mathematics	644
VETE	Veterinary	621
ARTS	Arts and Humanities	611
MEDI	Medicine	463
HEAL	Health Professions	437
CHEM	Chemistry	344
BIOC	Biochemistry, Genetics and Molecular Biology	268
PHYS	Physics and Astronomy	211

**Tabelle 2: Abstract-Verteilung des erstellten Korpus**





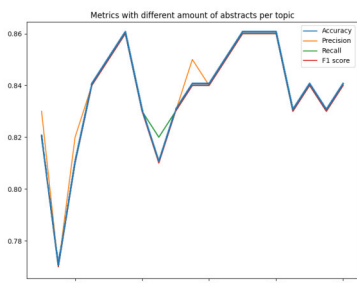
**Abbildung 4: Binäre Klassifikationsgenauigkeit aller Themengebiete mit über 1000 Abstracts**

mit zwölf Klassen entsteht. Die Genauigkeit der Klassifikationen für alle drei Algorithmen war zu Beginn der Optimierungen für TF-IDF bei 53,00 %, für Word2Vec bei 22,42 % und bei dem RNN bei 37,52 %.

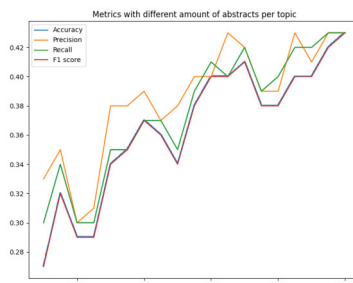
Weitere Optimierungen änderten die Ergebnisse jedoch nur geringfügig. Ein Grund hierfür, sind die breitgefächerten Kategorien, welche von Elsevier vorgegeben sind. Kategorien wie *Chemical Engineering* und *Engineering, Business, Management and Accounting* und *Economics, Econometrics and Finance* oder *Immunologie and Microbiology* und *Pharmacology, Toxicology and Pharmaceutics* haben viele ähnliche Thematiken, weshalb eine genaue Unterscheidung schwierig ist.

Um die Abgrenzung zwischen den Themengebieten besser zu erfassen, wurde eine binäre Klassifikation von jedem Themengebiet mit mehr als 1000 Abstracts durchgeführt. Dabei wurde die Genauigkeit der Klassifizierung als Evaluationsmaß genommen. In

Abbildung 4 ist eine Gegenüberstellung der Genauigkeiten zu sehen. Die höchste Genauigkeit erhielt die binäre Klassifikation von ENER und NEUR mit 99 %. Themengebiete, welche sich ebenfalls sehr gut von anderen abgrenzen lassen, sind ECON, MATE und EART. Diese besitzen die höchste gemeinsame Genauigkeit von 97,4 %. Die Klassifikation wurde daraufhin mit diesen fünf Kategorien für alle drei Algorithmen weiter optimiert. Die Anzahl an benötigten Abstracts wurde für jeden Algorithmus untersucht. Die Klassifikation wurde für jeden Algorithmus mit 100–1000 Abstracts durchgeführt. Dabei wurde in jedem Schritt die verwendete Anzahl um 50 Abstracts erhöht. Durch dieses Vorgehen soll die Anzahl an Abstracts herausgefunden werden, bei welchen die Performance der Lernalgorithmen am besten ist. Die Ergebnisse werden in den Abbildungen 5, 6 und 7 gezeigt. Eine detailliertere Darstellung zeigt Tabelle 3.



**Abbildung 5: Evaluation der SVM mit TF-IDF**

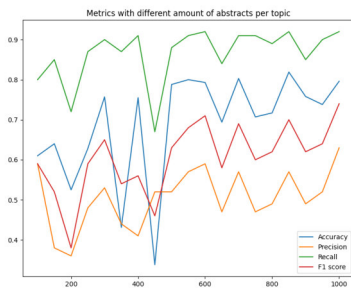


**Abbildung 6: Evaluation der SVM mit Word2Vec**

Algorithmus	Anzahl an Abstracts	Genauigkeit	Präzision	Sensitivität	F1-Wert
SVM mit TF-IDF	800	86 %	86 %	86 %	86 %
SVM mit Word2Vec	1000	43 %	43 %	43 %	43 %
LSTM-RNN	850	82 %	57 %	92 %	70 %

**Tabelle 3: Vergleich der besten Performance der Algorithmen**

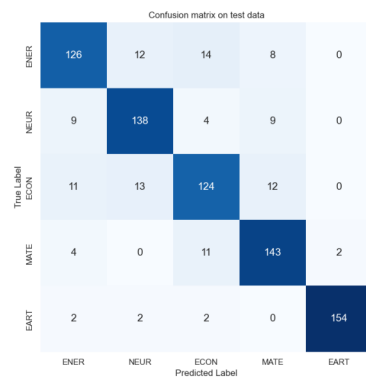
Das beste Gesamtergebnis konnte die SVM mit TF-IDF erhalten. Sie erzielte eine maximale Genauigkeit von 86 % und einen F1-Wert von 86 % bei 800 Abstracts. Die Konfusionsmatrix der SVM mit TF-IDF in Abbildung 8 zeigt, dass die meisten Fehlklassifikationen bei den Themengebieten ENER (21,25 %) und ECON (22,5 %) vorkommen. Bei EART ist die Fehlerrate die geringste mit 3,75 %. Die SVM mit Worteinbettungen hingegen ist bei der Klassifikation nicht über 43 % Genauigkeit gekommen. Dies ist bei 1000 Abstracts erreicht worden. Aus Abbildung 6 kann jedoch erkannt werden, dass sich mit einer höheren Anzahl an Abstracts die Performance der Klassifikation verbessert. Die Genauigkeit des RNNs beträgt 82 %, jedoch liegt die Wahrscheinlichkeit der richtig positiven Ergebnisse bei 57 %.



**Abbildung 7: Evaluation des LSTM-RNNs**

## 4 Diskussion

Die implementierte SVM mit TF-IDF ist in diesem Vergleich am besten abgeschnitten. SVMs werden häufig bei der Textklassifizierung verwendet, da diese bereits von sich aus eine hohe Genauigkeit liefern [1, 5]. Die SVM benötigt zudem eine geringere Trainingszeit als das LSTM-RNN. Beim Durchlauf der Klassifikationen mit 100-1000 Abstracts mit den fünf Themengebieten hat die SVM ca. 25 Minuten benötigt, das RNN ca. 75 Minuten. Eine grundlegende Limitation dieser Veröffentlichung ist das Fehlen von frei verfügbaren Korpora, welche wissenschaftliche Veröffentlichungen beinhalten. Der wichtigste



**Abbildung 8: Konfusionsmatrix der SVM mit TF-IDF bei 800 Abstracts**

Bestandteil einer jeden Klassifikation sind geeignete Daten. Ohne eine gute Datengrundlage kann kein breitgefächertes Wissen aufgebaut werden, welches bei der Klassifikation zu einem besseren Ergebnis beiträgt. Korpora, wie der *Academic Corpus*<sup>6</sup> der Universität Wellington oder der *Reading Academic Text Corpus*<sup>7</sup> der Universität Reading bieten einen englischen Korpus mit vielen wissenschaftlichen Veröffentlichungen aus unterschiedlichen Themengebieten an. Beide Korpora sind jedoch nur für Angehörige der jeweiligen Universitäten verfügbar.

Der eigene erstellte Korpus ist aus dem Mangel an verfügbaren Korpora entstanden. Dieser hat jedoch selbst einige Limitationen. Die breit gefassten Themengebiete erschweren eine genauere Klassifikation. Eine manuelle Überarbeitung des Korpus wäre notwendig, um die Themengebiete weiter zu unterteilen. Eine andere Möglichkeit wäre, mehr Daten in den Korpus einzupflegen. Dies kann entweder durch den Zugriff auf die erweiterte Elsevier API oder die Verwendung einer anderen API-Schnittstelle ermöglicht werden. Weitere Daten könnten im Falle der SVM mit Worteinbettungen ein besseres Ergebnis liefern. Unklar ist jedoch, ob sich der Trend, welcher in Abbildung 6 zu sehen ist, sich tatsächlich so weiterentwickelt. Zu den weiteren APIs zählen unter anderen die Springer API<sup>8</sup> und die IEEE Xplore API<sup>9</sup>. Aufgrund des limitierten Zeitraums der Bearbeitung dieser Veröffentlichung, war die Erstellung weiterer Textkorpora nicht möglich.

Nazemi et al. [16] stellen bei ihrem Vergleich von Volltexten und Abstracts fest, dass Abstracts häufig zu wenig Informationen liefern, um eine genaue Klassifikation zu gewährleisten. In Abstracts werden zudem Ergebnisse anhand von Zahlen dargestellt, was ein

größeres Rauschen bei den geringen vorhandenen Informationen hervorbringt. In dem erstellten Korpus ist die durchschnittliche Länge der Abstracts 217 Wörter, mit einem Minimum von 21 und einem Maximum von 872 Wörtern. Ein Vorteil der Abstracts ist jedoch, dass in den meisten Fällen immer die Methoden und spezifische Fachwörter des Themengebietes genannt werden. Dadurch kann sehr gut eine Übereinstimmung gefunden werden bei ähnlichen Themengebieten.

## 5 Ausblick

In dieser Veröffentlichung wurde ein Vergleich verschiedener maschineller Lerntechniken bei der Klassifizierung von Abstracts angestrebt. In einer vorangegangenen Literaturrecherche wurden unterschiedliche Techniken untersucht und bereits bestehende Vergleiche dargestellt. Basierend auf den Ergebnissen wurden drei Algorithmen implementiert und verglichen.

Trotz der genannten Limitationen hat diese Veröffentlichung demonstriert, dass sowohl eine SVM mit TF-IDF, als auch ein LSTM-RNN dazu verwendet werden kann Abstracts nach ihren Themengebieten zu klassifizieren. Die Ergebnisse zeigen ebenfalls, dass es nötig ist mehr öffentlich verfügbare Korpora anzubieten, um die Forschung weiter voranzutreiben.

## Literatur

- [1] Le Glaz Aziliz u. a. „Machine Learning and Natural Language Processing in Mental Health: Systematic Review“. In: *Journal of Medical Internet Research* (Juli 2021). DOI: 10.2196/preprints.15708.
- [2] Y. Barash u. a. „Comparison of deep learning models for natural language processing-based classification of non-English head CT reports“. In: *Neuroradiology* 62 (2020), S. 1247–1256. DOI: 10.1007/s00234-020-02420-0.

<sup>6</sup><https://www.wgtn.ac.nz/lals/resources/academicwordlist/information/corpus>

<sup>7</sup><http://www.reading.ac.uk/internal/appling/corpus.htm>

<sup>8</sup><https://dev.springernature.com/>

<sup>9</sup><https://developer.ieee.org/>

- [3] Ana Bordignon u. a. „Natural Language Processing in Business Process Identification and Modeling: A Systematic Literature Review“. In: Juni 2018. doi: 10.1145/3229345.3229373.
- [4] Emmanuel Carrodegua u. a. „Use of Machine Learning to Identify Follow-Up Recommendations in Radiology Reports“. In: *Journal of the American College of Radiology* 16 (Dez. 2018). doi: 10.1016/j.jacr.2018.10.020.
- [5] Arlene Casey u. a. „A systematic review of natural language processing applied to radiology reports“. In: *BMC Medical Informatics and Decision Making* 21.1 (Juni 2021). doi: 10.1186/s12911-021-01533-7.
- [6] Stefano Ferilli. *Automatic Digital Document Processing and Management*. Jan. 2011. doi: 10.1007/978-0-85729-198-1.
- [7] Kevin Gagnon, Tami Crawford und Jihad Obeid. „Comparison of Convolutional Neural Network Architectures and their Influence on Patient Classification Tasks Relating to Altered Mental Status“. In: Dez. 2020, S. 2752–2756. doi: 10.1109/BIBM49941.2020.9313156.
- [8] Natalie Gavriellov-Yusim u. a. „Comparison of text processing methods in social media-based signal detection“. In: *Pharmacoepidemiology and Drug Safety* 28.10 (2019), S. 1309–1317. doi: 10.1002/pds.4857.
- [9] Krystian Horecki und Jacek Mazurkiewicz. „Natural Language Processing Methods Used for Automatic Prediction Mechanism of Related Phenomenon“. In: *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)* 9120 (Juni 2015), S. 13–24. doi: 10.1007/978-3-319-19369-4\_2.
- [10] Robert Jungnickel u. a. „Evaluation and Comparison of Cross-lingual Text Processing Pipelines“. In: *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*. 2019, S. 417–425. doi: 10.1109/SSCI44817.2019.9002921.
- [11] Tarek Kanan u. a. „A Review of Natural Language Processing and Machine Learning Tools Used to Analyze Arabic Social Media“. In: *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*. 2019, S. 622–628. doi: 10.1109/JEEIT.2019.8717369.
- [12] Kory Kreimeyer u. a. „Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review“. In: *Journal of Biomedical Informatics* 73 (Sep. 2017), S. 14–29. doi: 10.1016/j.jbi.2017.07.012.
- [13] Merton Lansley, Nikolaos Polatidis und Stelios Kapetanakis. „SEADER: A Social Engineering Attack Detection Method Based on Natural Language Processing and Artificial Neural Networks“. In: Aug. 2019, S. 686–696. doi: 10.1007/978-3-030-28377-3\_57.
- [14] David Lewis u. a. „A Comparison of Two Learning Algorithms for Text Categorization“. In: *Third Annual Symposium on Document Analysis and Information Retrieval* (Apr. 1994), S. 14.
- [15] Roberto Lopez-Martinez und Gerardo Sierra. „Research Trends in the International Literature on Natural Language Processing, 2000-2019 — A Bibliometric Study“. In: *Journal of Scientometric Research* 9 (Dez. 2020), S. 310–318. doi: 10.5530/jscires.9.3.38.
- [16] Kawa Nazemi u. a. „Comparison of Full-text Articles and Abstracts for Visual Trend Analytics through Natural Language Processing“. In: *2020 24th International Conference Information Visualisation (IV)*. 2020, S. 360–367. doi: 10.1109/IV51561.2020.00065.

- [17] Emma Richard und Bhargava Reddy. „Text Classification for Clinical Trial Operations: Evaluation and Comparison of Natural Language Processing Techniques“. In: *Therapeutic Innovation & Regulatory Science* 55 (2021), S. 447–453. DOI: 10.1007/s43441-020-00236-x.
- [18] Dipanjan Sarkar. *Text Analytics with Python: A Practitioner’s Guide to Natural Language Processing*. Apress, 2019, S. 688. DOI: 10.1007/978-1-4842-4354-1.
- [19] Dipanjan Sarkar, Raghav Bali und Tuskar Sharma. *Practical Machine Learning with Python A Problem-Solver’s Guide to Building Real-World Intelligent Systems*. Apress, 2018, S. 545. DOI: 10.1007/978-1-4842-3207-1.
- [20] M. Rajani Shree und Shambhavi B.R. „Performance comparison of Word sense disambiguation approaches for Indian languages“. In: *2015 IEEE International Advance Computing Conference (IACC)*. 2015, S. 166–169. DOI: 10.1109/IADCC.2015.7154691.
- [21] Imre Solti u. a. „Automated Classification of Radiology Reports for Acute Lung Injury: Comparison of Keyword and Machine Learning Based Natural Language Processing Approaches“. In: *IEEE International Conference on Bioinformatics and Biomedicine*. Bd. 2009. IEEE, 2009. DOI: 10.1109/BIBMW.2009.5332081.
- [22] Maxwell Taggart u. a. „Comparison of 2 Natural Language Processing Methods for Identification of Bleeding Among Critically Ill Patients“. In: *JAMA Network Open* 1.6 (Okt. 2018), e183451–e183451. DOI: 10.1001/jamanetworkopen.2018.3451.
- [23] Yanshan Wang u. a. „A comparison of word embeddings for the biomedical natural language processing“. In: *Journal of Biomedical Informatics* 87 (2018), S. 12–20. DOI: <https://doi.org/10.1016/j.jbi.2018.09.008>.
- [24] Wenpeng Yin u. a. *Comparative Study of CNN and RNN for Natural Language Processing*. 2017.



# A Scientometric Review about Current and Upcoming Trends of Artificial Intelligence in the Healthcare Sector

Felix Welker

Hochschule Reutlingen

Felix.Welker@Student.Reutlingen-University.de

## Abstract

Artificial intelligence and healthcare are in relation to each other still a very new but increasingly important field of research. However, a visual representation of this connection with a focus on current trends based on scientometrics does not yet exist. The results of this paper showed that there are many different research hotspots like "cancer", "covid-19", "diagnosis", "gene", and "protein". The research focus evolved mainly from basic human understanding of "genes" and "proteins", then to more complex methods like "diagnosing diseases" to "cancer research". In 2020 the research focus shifted to the ongoing pandemic and its "coronavirus". Publications between the beginning of 2011 and the end of 2020 were taken from Web of Science and used as a source for this paper. These hotspots could be identified and visualized using different visualization software like CiteSpaceV, Carrot<sup>2</sup> and ScientoPy. This paper's aim is to promote the development of related artificial intelligence theories

in medicine and to identify relevant research directions for future work.

## CCS Concepts

- Information systems~Information retrieval~Evaluation of retrieval results~Presentation of retrieval

## Keywords

Artificial Intelligence, Healthcare, Medical field, Visualization, Research hotspots, Emerging trends, Scientometrics

## 1 Introduction

The global health industry has grown in the last couple of years at a rapid pace. The healthcare sector was worth \$8.45 trillion in 2018 and had an annual growth rate of 7.3% since 2014. [1] It is expected that the industry grows even faster during the adoption of digitalization and the newly deployed technologies such as cloud computing or artificial intelligence (AI). Especially the Covid-19 pandemic is forcing the health care sector to innovate and adapt in a short period of time. New treatments, methods to diagnose and more is announced on a daily basis.

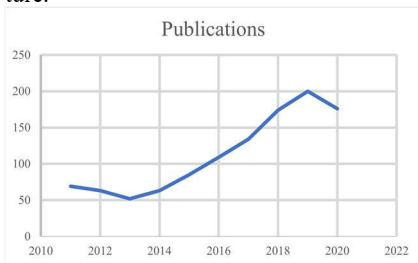
Artificial intelligence is becoming more sophisticated at doing what humans do, and is at the same time more efficiently, faster and does its job for less money. Especially more complex tasks which required a lot of time in the past like protein folding can now be done in fraction by using advanced machine learning techniques.

---

Betreuer Hochschule: Prof. Dr. Sven Steddin  
Hochschule Reutlingen  
Sven.Steddin@Reutlingen-  
University.de

Informatics Inside Herbst 2021  
24. November 2021, Hochschule Reutlingen  
Copyright 2021 Felix Welker

Yet, both topics, the health sector and artificial intelligence, are still emerging topics. An investigation has shown that the annual number of published papers from Web of Science increased drastically over the past ten years. Figure 1 shows all published articles of the last ten years starting at 2011 and ending in 2020. The small decline in 2020 can be explained by the coronavirus crisis which made research temporarily more difficult. The publications of 2021 are not included in the study, because 2021 is not over yet and therefore wouldn't show its true picture.



**Figure 1 – Published articles of the last years**

Figure 2 shows the regions where current research of these topics is conducted. The heatmap has been generated by using a plugin for the vue web framework<sup>1</sup>. Google Maps serves as a basis for the world map. It shows that there are hotspots all over the world. The most papers published are from North America. However, the European countries aren't this far behind and Asia as well. It shows that China is also very interested in the topic and therefore heavily researching it. In Figure 3 you see exactly which countries are leading in this field. The United States do clearly lead with its elite universities like MIT, Stanford and Harvard. However, China is gaining momentum very quickly by nearly doubling its publications in the last 1½ years. The United Kingdom mainly with Oxford makes the third place in the list.

Some of the most focused topics are cancer research, gene research, diagnosis methods, protein research and now a lot of covid-19 research. Research shows that especially processes like protein folding are very hard to compute and therefore AI can help speed the process up by a significant amount of time. [2] Covid-19, which is obviously a newer research topic, also has major challenges that can be solved by using artificial intelligence [3].

Scientometrics is no new field of study and therefore multiple analyses about different topics have been made. In 2014 Yao et. al did a study about current scientometrics trends of global health systems. They came to the conclusion that the sector has increased rapidly over the last 20 years. The focus was mainly on public, environmental and occupational health care. [4] Another study from Soleimani-Roozbahani et. al. has been made about healthcare and big data. The results showed that big data was a research topic which was rapidly rising but rather neglected the health care aspect. [5]

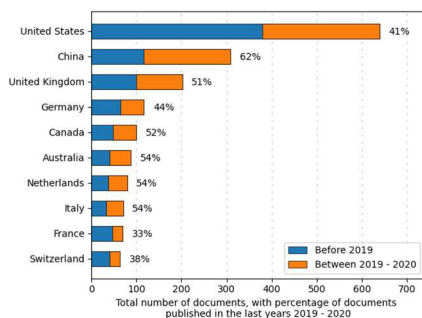
When checking if scientometrics have been applied with artificial intelligence, many papers were found. However, they were all in conjunction with topics like agriculture or engineering. No recent studies could be found about artificial intelligence in relation to healthcare. This paper should try to close that gap and identify relevant research directions for future work.

The rest of the paper is organized in the following manner. Chapter two introduces the methodology adopted by this paper. This includes the source of the data and the tools used. Different graphs like the co-cited references network, co-occurrence keywords network, burst keywords and the timeline view have been used as a visual evaluation

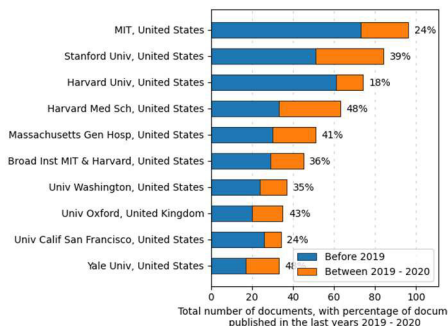
<sup>1</sup> <https://github.com/bobby-white/vue-google-heatmap>



**Figure 2 – A heatmap of research hotspots worldwide. Red dots represent hotbeds of scientific activity.**



(a)



(b)

**Figure 3 – Total number of documents, with percentage of documents in the last years 2019 – 2020: (a) → Countries with publications (b) → Institutions by countries**

method to find all the trends. These visualizations were able to highlight current and upcoming trends of AI and healthcare in conjunction. Finally, in the last sector, the conclusion is reached and presented.

## 2 Methodology

This chapter shows the methods used for data acquisition and which software was used to visualize the hotspots.

### 2.1 Data collection

The dataset used in this paper is exclusively from Web of Science, which is a website

with access to multiple databases that provide access to an extensive citation index of different academic publications. In order to collect the needed data for this paper the following advanced search query has been conducted: (ALL=(Artificial Intelligence OR Machine Learning)) AND ALL=(Medical OR Health OR treatment OR diagnosis OR diagnostic).

The language was set to English, and the timeframe was set between 2011 up until the end of 2020 (the time of retrieval was August



2021). In order to only get the most important papers only highly cited papers were requested. This query led to 1113 records as a result. Each record contains the title, all authors involved, keywords and the abstract, as some more metadata like the year of publication and all cited papers within the article. Web of Science offers the possibility to download the data in different file formats. For our use case the file format “*tab delimited*” has been chosen and the data has been downloaded. This gave the possibility to import the data in scientometrics software and analyze it.

## 2.2 *Scientometrics software analysis*

Scientific publications have been growing rapidly in the last year. Especially during the coronavirus pandemic scientists and clinicians have hastened to comprehend and reduce the threat, sharing what they've learned with other researchers. [6] To better understand this overwhelming data, researchers developed different technologies to better visualize this information and therefore make it easier to spot and find hidden trends and patterns in the huge amount of data. [7] One of the software used in this paper is CiteSpace<sup>2</sup>. CiteSpace is a freely accessible computational tool that uses interactive visual explorations of patterns and trends provided by the body of scientific literature to find solutions to these queries. CiteSpace can assist in better understanding the scientific literature on a certain issue. [8] As a result, the most recent version of CiteSpace (5.8.R3c) was chosen for this paper's research.

CiteSpace is used to create the co-cited references network, the co-occurrence keywords network, and the time zone view. Depending on the importance of a hotspot CiteSpace creates a cluster in a fitting size. The more important the subject, the bigger

the cluster. This demonstrates in a co-cited references network which are the main subjects' researchers are currently working on. By evaluating relationships between citations in the literature, a co-citation network focuses on studying the structure of scientific communication. [9] Meanwhile, the co-cited references network's burst references, which are a lot of citations in a short period of time, and the co-occurrence keywords network's burst keywords, which are a lot of occurrences in a short timeframe, help to assess whether the references or keywords are related with each other. The examination of these networks will show the current research trends and upcoming trends of AI in relation to healthcare in this paper.

Apart from CiteSpace, ScientoPy<sup>3</sup> was also used to analyze the publications. ScientoPy is an Open-Source project written in Python which allows an import of the Web of Science export. It processes the exports and can display various different graphs like top authors, countries, or institutions based on first authors or all documents' authors. Various different visualizations can then be displayed like bar charts, bar trends, timelines and more.

As a last tool Carrot<sup>24</sup> has been used. Carrot<sup>2</sup> is an Open-Source Search Results Clustering Engine. It can naturally put together little assortments of archives into topical classifications. You can select different clustering algorithms and cluster various different resources like keywords, authors and more.

These tools are all used in conjunction to make the analysis on Ai and health in relation to each other more comprehensively.

---

<sup>2</sup> <http://cluster.cis.drexel.edu/~cchen/citespace/>

<sup>3</sup> <https://www.scientopy.com/en/>

<sup>4</sup> <https://search.carrot2.org/#/search/web>

### 3 Current and Upcoming Trends of AI in the healthcare sector

This chapter focuses on the main evaluation of the different visualizations in order to find current and upcoming trends. For this, different visualization techniques and different software has been used.

#### 3.1 Evaluating the references network

When two papers are cited in the same document at the same time, it is referred to as a co-cited references network in CiteSpace. These two documents are linked by a co-citation. These co-citation relationships can be used to create maps that indicate a domain's research focus and current knowledge about it. [8, 10]

To generate the co-cited references network of AI and Healthcare, all the 1113 document records were imported into CiteSpace. The parameters in CiteSpace are configured as follows. Because the application of artificial intelligence has mainly appeared in recent years only the literature as a study object of the past 10 years have been selected. Therefore, the parameter “*Time Slicing*” has been set to 2011–2020 and “*Years per slice*” to 1 year. The parameter “*Top N*” has been set to 50, which means the top 50 levels of most cited or occurred nodes will be selected from each slice. The node type is set to “*Reference*”.

Between 2011 and the end of 2020, the co-cited references network produced 64,855 cited papers (from 1113 citing documents) and a total of 707 nodes, 2960 links, and 14 visibly labeled clusters. Figure 4 shows the rendered co-citation map. The modularity  $Q = 0.8043$  in the figure implies that the network map is highly modular and that the network spectrum clustering findings are excellent. Everything higher than 0.7 is a good score. [8]

The biggest and most representative tags of these clusters have a bold and red font, and

their names have been highlighted in Figure 4. The nodes in the diagram represent co-cited references, and the bigger they are the more often have they been cited. This means, the bigger a node, the more valuable the research is. Co-citation links are lines that connect nodes, with different colors revealing two nodes connected for the first time. The year is shown in the time bar at the top of the network.

The entire knowledge map displays the cluster knowledge structure's spatial and temporal evolution, as well as the most trending topics in AI and healthcare studies. The diagram shows that AI and healthcare research has resulted in a huge variety of research subjects, with some connections between them. It demonstrates that AI and healthcare data research has evolved into a system rather than a single entity. Only one cluster is separated from the others.

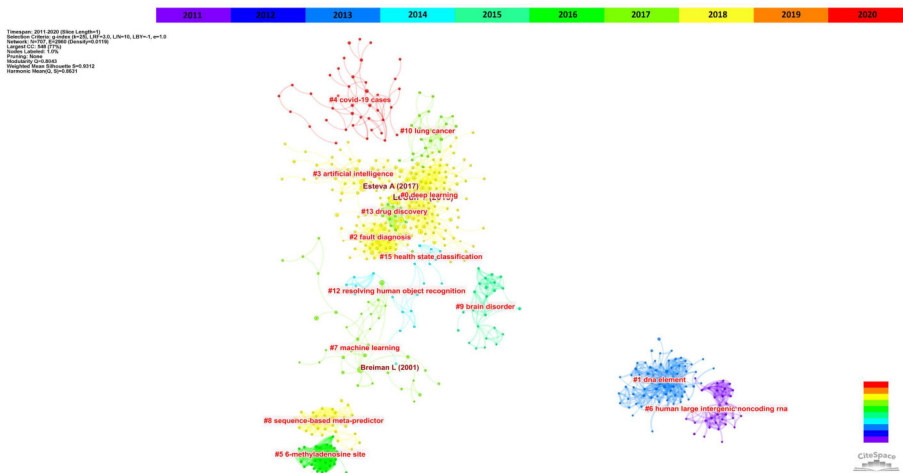
The results show that there are strong links between “#0 *deep learning*”, “#13 *drug discovery*” and “#2 *fault diagnosis*” which implies that the research of new drugs and the fault diagnosis of diseases strongly relate to each other but also connect very strongly with deep learning and therefore basically with applied artificial intelligence techniques. The topic “#4 *covid-19 cases*” appears at first in 2020, however has no visible connections to other clusters yet. Its location is however close to “#3 *artificial intelligence*” which shows that this topic is trending in that direction and may become a hot research topic along with AI in the future. Also “#10 *lung cancer*” is strongly interconnected with “#0 *deep learning*” and has been mainly focused in 2017 and 2018. Furthermore, the majority of the nodes are green and orange (2016 – 2019), indicating that AI and healthcare research hotspots have only recently emerged, and the topic is a relatively new research field. But not to be forgotten are “#1 *dna element*” and “#6 *human large intergenic noncoding rna*”, because they are the earliest clusters of 2011 and 2012. These are the topics were the researchers started with.

The co-cited references network only shows clusters of cited literature. In order to analyze the literature itself more extensively the literature has been analyzed with Carrot<sup>2</sup> in Figure 5. Carrot<sup>2</sup> is an Open Source search results clustering engine. [11] It can group small collections of documents, such as search results or document abstracts, into relevant groups. Carrot<sup>2</sup> offers different clustering algorithms, however for this analysis the Lingo algorithm was chosen. The Lingo algorithm is a new technique for grouping search results that prioritizes the quality of cluster descriptions. [12] Foam Tree is the name of the visualization created by Lingo in Carrot<sup>2</sup>. The importance of a cluster is represented by the size of a foam. The cluster is more relevant if the foam bubble is bigger. The colors

of the clusters in the foam tree have no meaning.

The results of the foam tree show some similarities with the co-cited references network. “Deep Learning”, “Machine Learning”, “Artificial Intelligence” and “Neural Networks” are by far the largest clusters. All of these except for “Neural Networks” exist in the co-cited references network. Other bigger clusters like “Fault Diagnosis”, “Prediction by Machine Learning”, “Diagnosis using Deep”, “Predictive Modeling” are in co-cited references network in a modified manner.

Some new trends like “Big Data”, “Medical Image” and “Alzheimer’s Disease” in smaller foams could be identified as well.



**Figure 4 – Co-cited references network generated using CiteSpace. The clusters are interconnected co-cited references. The lines that connect these clusters are co-citation links. The colors of the time bar at the top indicate the year of a cluster.**

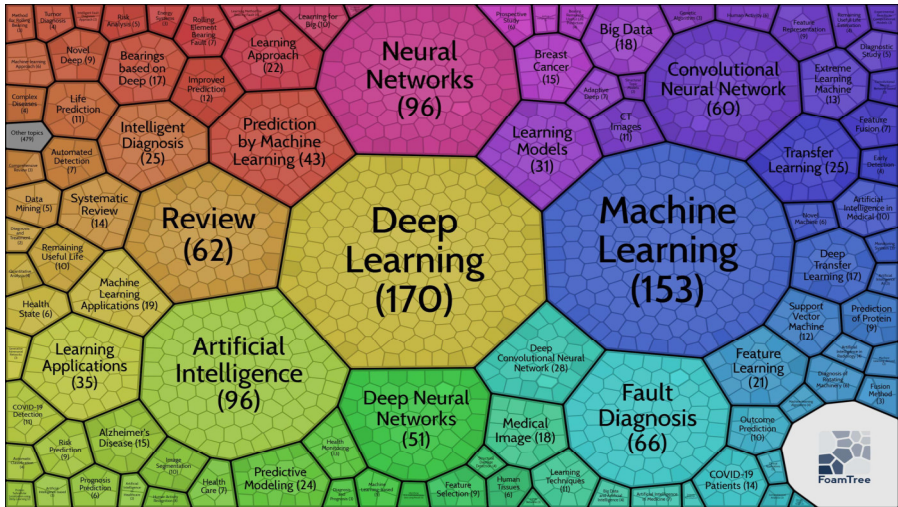


Figure 5 – The Foam Tree of AI and healthcare: generated by using Lingo clustering method from Carrot<sup>2</sup>

### 3.2 Evaluating burst references

Figure 6 shows a burst detection analysis of referenced papers in order to investigate the literature that has made important contributions to AI and healthcare. Burst references are those that have a large number of citations in a short period of time, indicating the fields of interest to the scientific community, i.e. current trends. [7] We may quickly understand the main study fields of the scientific community at any given time by watching the temporal trends of burst references. The graph in figure 6 shows the top 15 citations with a strong citation burst. The citation burst's duration is indicated by the horizontal red bar. The dark blue bar shows the time after its publication. The burst references exist practically every year from 2011 to 2018, with the exception of 2016, 2019 and 2020. Despite the fact that the strength and duration of these bursts varies, and their study content continues to evolve, the chart clearly illustrates that AI and healthcare has been a constant research hotspot since 2011. That there are no very recent burst references can

be explained by the coronavirus crisis since research and its focus changed temporarily. The two most notable years are 2011 and 2017 when a huge number of publications were published. This indicates the years where the level of interest in AI and healthcare research came mainstream. There was a good start with some important research topics in 2011 which continued for a long time. Right when most of them stopped new topics were started in 2017. The interest basically never stifled.

In order to assess the data more representatively the strongest citation bursts of the years with the most publications have been examined more closely to uncover new research trends.

In the year 2011 the research mainly focused on human genome studies. Some of the strongest citation bursts covered these in particular Ernst et. Al. [13] In this specific paper no artificial intelligence approach was used however in other weaker bursts first approaches with support vector machines have been made.

References	Year	Strength	Begin	End	2011 - 2020
Emst J, 2011, NATURE, V473, P43, DOI 10.1038/nature09906	2011	6.86	2011	2015	
Hindorf LA, 2009, P NATL ACAD SCI USA, V106, P9362, DOI 10.1073/pnas.0903103106	2009	5.14	2011	2015	
Guttman M, 2010, NAT BIOTECHNOL, V28, P503, DOI 10.1038/nbt.1633	2010	5.07	2011	2013	
Ashburner M, 2000, NAT GENET, V25, P25, DOI 10.1038/75556	2000	4.98	2011	2016	
Tibshirani R, 1996, J ROY STAT SOC B MET, V58, P267, DOI 10.1111/j.2517-6161.1996.tb0280.x	1996	4.53	2011	2017	
Lindblad-Toh K, 2011, NATURE, V478, P476, DOI 10.1038/nature10530	2011	5.38	2012	2015	
Breiman L, 2001, MACH LEARN, V45, P5, DOI 10.1023/A:1010933404324	2001	5.24	2013	2017	
Dunham I, 2012, NATURE, V489, P57, DOI 10.1038/nature11247	2012	7.26	2014	2016	
Chang CC, 2011, ACM T INTEL SYST TEC, V2, P0, DOI 10.1145/1961189.1961199	2011	6.51	2015	2016	
Hinton GE, 2006, SCIENCE, V313, P504, DOI 10.1126/science.1127647	2006	6.32	2015	2016	
BENJAMINI Y, 1995, J R STAT SOC B, V57, P289, DOI 10.1111/j.2517-6161.1995.tb02031.x	1995	4.62	2015	2017	
Ciresan DC, 2013, LECT NOTES COMPUT SC, V8150, P411, DOI 10.1007/978-3-642-40763-5_51	2013	4.83	2017	2018	
Szegedy C, 2015, PROC CVPR IEEE, V0, P1, DOI 10.1109/CVPR.2015.7298594	2015	4.8	2017	2018	
He KM, 2016, PROC CVPR IEEE, V0, P770, DOI 10.1109/CVPR.2016.90	2016	6.09	2018	2020	
Esteva A, 2017, NATURE, V542, P115, DOI 10.1038/nature21056	2017	5.19	2018	2020	

**Figure 6 – Top 15 citations with a strong citation burst**

Some of the stronger bursts overall appeared in 2017 and 2018 with its main focus on classification of different cancer types and image recognition. [14, 15] The surge started in 2018 and lasted through 2019, implying that during this time, experts paid particular attention to the difficulties with cancer and image detection. It can only be speculated that the coronavirus pandemic changed the research focus temporarily and that the cancer research area will come back into focus in the future. Evaluating co-occurrence keywords networks and burst keywords

When two keywords appear in a document at the same time, they are usually considered content relevant. A co-occurrence keywords network is just a collection of such keywords that indicates the development and progression of key terms in a certain field. As a result, evaluating the co-occurrence keywords network map might show hotspots and developing trends in a certain subject. [7] The parameters of the network have been set as follows: "Node Types" was set as "Keyword" and the parameter for "Top N" has been set to 25, which means the top 25 levels of most cited or occurred nodes will be selected from each slice in CiteSpace. The data was not pruned. The view has been set to a timeline view of co-occurrence. Figure 7 shows the generated network. The network

consists of 673 keywords and 2875 links between 2011 and end-2020. The nodes in the network are the keywords. In order to highlight and display the important nodes, small nodes have been removed from the periphery. The modularity  $Q = 0.8044$  in the figure implies that the network map is highly modular and that the network spectrum clustering findings are excellent. The nodes in the diagram represent co-cited references, and the bigger they are, the more often they have been cited. This means, the bigger a node, the more valuable the research is. Co-occurrence links are lines that connect nodes, with different colors indicating when a connection was created for the first time. The year is shown on the time bar at the top of the network. The thickness of the lines indicates the co-occurrence strength.

The network indicates that "diagnosis", "prediction", "identification", "support vector machine", "algorithm", "gene" are research trends of the years 2011-2013.

It also indicates that "neural network", "cancer", "segmentation", "disease" were research hotspots of 2014-2015

The trends of 2016-2018 have been "convolutional neural network", "image", "computer aided detection".

In the years 2019-2020 “health”, “coronavirus”, “pneumonia”, “lung” and “artificial intelligence” were research trends.

Obviously, “neural network” is one of the core issues of AI and healthcare research, which is why it’s the biggest node of all. But also “prediction”, “diagnosis”, “cancer” and “identification” are very clearly displayed. This indicates that the topics AI and healthcare have received a lot of attention in relation together in the recent years.

Keywords are highly refined and distilled essential content from publications. Burst keywords can indicate the rapid emergence of a specific keyword. As a result, burst keywords can be used as a tool to locate hotspots and rising trends. [7]

Figure 8 lists the top 15 keywords including the time of burst and usage. The result shows that all keywords already appeared in 2011. The earliest burst keywords also already appeared in 2011, including “gene”, “identification” and “discovery”. Whereas more recent burst keywords were “cancer”, “dna” and “generation” indication the focus change from gene research to cancer research. This goes on par with the findings in the previous sections.

To analyze not only the cited keywords, the literature itself has also been analyzed. For this the tool ScientoPy has been used. The data has been imported as tab delimited file and the criterion “authorKeywords” has been set. The timespan has been set from 2011 to 2020, the window is two years. In figure 9a you can see the top 3 keywords are “Machine learning”, “Deep learning” and “Artificial intelligence”. All these three keywords got especially in the last two years a lot of traction. “COVID-19” made a respectable fourth place even though it was never mentioned before 2020. Another interesting keyword is “big data”. It makes sense that big data is also important, because in order to train models for machine learning, a lot of data has to be collected and used. For figure 9b the same parameters as before have been used however the graph type is a timeline. It clearly shows how the trend is going. Artificial intelligence related topics and also “fault diagnosis” and “COVID-19” will be definitely a future research trend.

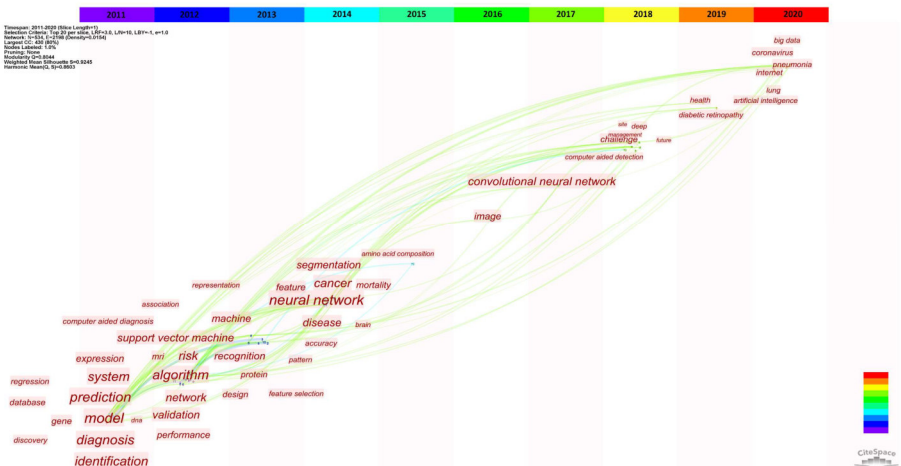


Figure 7 –Time zone map of co-occurrence network

Keywords	Year	Strength	Begin	End	2011 - 2020
gene	2011	4.19	2011	2016	
gene expression	2011	3.59	2011	2016	
expression	2011	3.58	2011	2013	
identification	2011	3.35	2011	2012	
discovery	2011	3.28	2011	2015	
transcription factor binding	2011	3.22	2011	2012	
chromatin	2011	3.14	2011	2013	
binding	2011	3.21	2012	2015	
brain	2011	3.4	2014	2015	
database	2011	3.72	2015	2016	
organization	2011	3.42	2015	2016	
cancer	2011	3.72	2018	2019	
site	2011	3.05	2018	2019	
dna	2011	2.77	2018	2019	
generation	2011	2.77	2018	2019	

Figure 8 – Top 15 Keywords with the strongest citation bursts

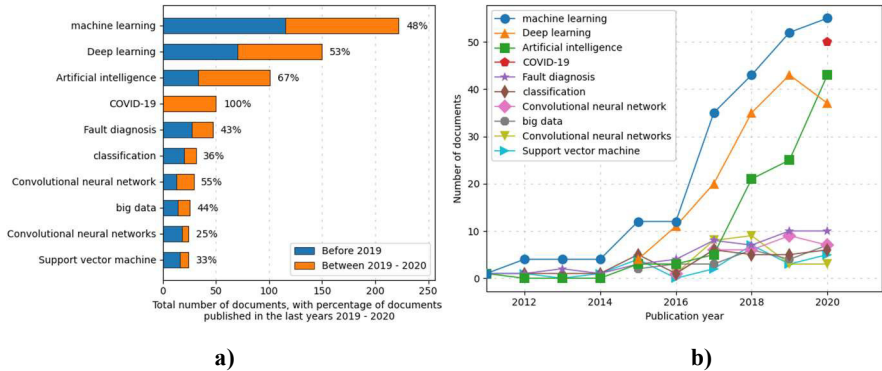


Figure 9 – Top keywords of the literature by publication year and by percentage of documents published

## 4 Summary and discussion

Based on material gathered from Web of Science for the years 2011–2020, the knowledge map of AI and healthcare was created using information visualization technologies CiteSpace, ScientoPy, and Carrot<sup>2</sup>. The hotspots and developing trends of AI and healthcare have been found in a multicriteria and thorough research method using a co-cited references network, a co-

current keywords network, burst references, burst keywords and the timezone view. In addition, the Vue Google Maps tool was used to produce the area distribution of AI and healthcare study data around the world, which is displayed in section 1. To compensate for CiteSpace's shortcoming in this part, Carrot<sup>2</sup> and ScientoPy was used to conduct a citation literature analysis.

## 4.1 Results

According to Web of Science data, the annual number of papers about AI and healthcare has gradually increased over the last ten years. The generated heatmap also showed that the topic is being researched all over the globe. The United States, China, European countries and Korea have been very active in this field. Clearly, more and more academics are becoming interested in AI and healthcare as a topic to study.

Distinct evaluations yielded different hotspots and developing trends when it comes to AI and healthcare. However, the following conclusions can be made after integrating a range of findings. The research trends of AI and healthcare mainly include “deep learning”, “machine learning”, “artificial intelligence”, “neural networks”, “fault diagnosis”, “prediction”, “predictive modeling”, “drug discovery”, “covid-19”, “pneumonia”, “cancer”, “dna”, “gene”, “image recognition”, “identification”, “discovery”, etc. Additionally, the evaluations showed that in the beginning of the recent 10 years, the focus lied mainly on gene and DNA research, which is more about understanding difficult concepts with a new type of technology. Later, the focus shifted to drug discovery and especially research on diseases like cancer where no treatments are available yet. In the end the coronavirus pandemic hit and the whole search shifted mainly to studying Covid-19. This seems to be still ongoing, but cancer research and drug discovery will likely make a comeback since these are still unsolved problems. Upcoming trends are therefore “covid-19”, “cancer research” and “drug discovery”.

## 4.2 Outlook

In recent years, the development of AI and healthcare has shown many new trends, such as cancer research and the diagnosing of diseases. Covid-19 was also a big use case for AI approaches e.g., trying to predict the spread of it. These use cases of AI are perfect and therefore can be foreseen that these will

see a more widespread use in the future. This paper outlined these findings and can be used to identify relevant research directions for future work.

## 5 References

- [1] Smiljanic Stasha, *The State of Healthcare Industry – Statistics for 2021*. [Online]. Available: <https://policyadvice.net/insurance/insights/healthcare-statistics/> (accessed: Sep. 14 2021).
- [2] A. Senior, J. Jumper, D. Hassabis, and P. Kohli, *AlphaFold: Using AI for scientific discovery*. [Online]. Available: <https://deepmind.com/blog/article/AlphaFold-Using-AI-for-scientific-discovery> (accessed: Sep. 14 2021).
- [3] F. Al-Turjman, Ed., *Artificial Intelligence and Machine Learning for COVID-19*. Cham, Switzerland: Springer, 2021.
- [4] Q. Yao *et al.*, “Scientometric trends and knowledge maps of global health systems research,” *Health Res Policy Sys*, vol. 12, no. 1, p. 26, 2014, doi: 10.1186/1478-4505-12-26.
- [5] F. Soleimani-Roozbahani, A. Rajabzadeh Ghatari, and R. Radfar, “Knowledge discovery from a more than a decade studies on healthcare Big Data systems: a scientometrics study,” *J Big Data*, vol. 6, no. 1, 2019, doi: 10.1186/s40537-018-0167-y.
- [6] Claire Jarvis, *Journals, Peer Reviewers Cope with Surge in COVID-19 Publications*. [Online]. Available: <https://www.the-scientist.com/news-opinion/journals-peer-reviewers-cope-with-surge-in-covid-19-publications-67279> (accessed: Sep. 14 2021).
- [7] C. Chen, R. Dubin, and T. Schultz, “Science Mapping,” in *Encyclopedia of information science and technology*, M. Khosrowpour, Ed., 3rd ed., Hershey, Pa.: Information Science Reference, 2015, pp. 4171–4184.



- [8] Chaomei Chen, *How to Use CiteSpace*. British Columbia, Canada: Lean Publishing, 2016. [Online]. Available: [https://www.researchgate.net/publication/320699133\\_How\\_to\\_Use\\_CiteSpace\\_c\\_2015-2020](https://www.researchgate.net/publication/320699133_How_to_Use_CiteSpace_c_2015-2020)
- [9] S. Radhakrishnan, S. Erbis, J. A. Isaacs, and S. Kamarthi, “Novel keyword co-occurrence network-based methods to foster systematic reviews of scientific literature,” *PLoS ONE*, vol. 12, no. 3, e0172778, 2017, doi: 10.1371/journal.pone.0172778.
- [10] H. J. Kim, Y. K. Jeong, and M. Song, “Content- and proximity-based author co-citation analysis using citation sentences,” *Journal of Informetrics*, vol. 10, no. 4, pp. 954–966, 2016, doi: 10.1016/j.joi.2016.07.007.
- [11] Open Source, *Carrot2*. [Online]. Available: <https://github.com/carrot2/carrot2> (accessed: Sep. 24 2021).
- [12] S. Osiński, J. Stefanowski, and D. Weiss, “Lingo: Search Results Clustering Algorithm Based on Singular Value Decomposition,” in *Proceedings of the International 2S: 2PWM'04 Conference held in Zakopane, Poland, May 17-20, 2004*, J. Kacprzyk, Ed., 1st ed., Berlin/Heidelberg: Springer Berlin Heidelberg, 2000, pp. 359–368.
- [13] J. Ernst *et al.*, “Mapping and analysis of chromatin state dynamics in nine human cell types,” *Nature*, vol. 473, no. 7345, pp. 43–49, 2011, doi: 10.1038/nature09906.
- [14] A. Esteva *et al.*, “Dermatologist-level classification of skin cancer with deep neural networks,” *Nature*, vol. 542, no. 7639, pp. 115–118, 2017, doi: 10.1038/nature21056.
- [15] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *29th IEEE Conference on Computer Vision and Pattern Recognition: CVPR 2016 : proceedings : 26 June-1 July 2016, Las Vegas, Nevada, Las Vegas, NV, USA, 2016*, pp. 770–778.



# Digitale Gesundheitslösungen zur Unterstützung von Patienten mit chronisch obstruktiver Lungenerkrankung: Eine systematische Literaturrecherche

Yailda Aini

Hochschule Reutlingen

Yailda.Aini@student.reutlingen-university.de

## Abstract

Die chronische obstruktive Lungenerkrankung (COPD) stellt die weltweit dritthäufigste Todesursache dar. Sie geht mit einer Verengung und einer dauerhaften Entzündung der Atemwege einher. Dies bedingt eine immer weiter abnehmende Luftversorgung der Lunge, wodurch das Atmen erschwert wird. Zu den klinischen Anzeichen gehören Husten, Schleimauswurf und Atemnot. Eine Exazerbation (Verschlechterung) der Symptome schränkt die Lebensqualität des Betroffenen immens ein und kann sogar lebensbedrohlich werden. COPD-Patienten können durch effektive Behandlung und aktivem Selbstmanagement positiven Einfluss auf die Erkrankung nehmen und so eine Exazerbation verhindern. Nicht selten kommt es vor, dass sich der Zustand des Patienten bereits vor dem nächsten Nachsorgetermin verschlechtert hat. Durch moderne Informations- und Kommunikationstechnologien wurden in den letzten Jahrzehnten diverse digitale Gesundheitslösungen entwickelt, welche eine Überwachung des Ges-

undheitszustands und des Selbstmanagements von Patienten ermöglichen sollen. Um den Stand der Forschung zu ermitteln, wird eine systematische Literaturrecherche durchgeführt, mit dem Ziel Aufschluss darüber zu geben, inwiefern digitale Gesundheitslösungen COPD-Patienten unterstützen und damit Exazerbationen verringern können.

## CCS Concepts

- **Applied computing** → Health informatics
- **Information systems** → Information systems applications
- **Human-centered computing** → *User studies*

## Keywords

Digital Health, Remote Monitoring, Chronic Obstructive Pulmonary Disease, COPD, Systematic Review, Self-management, Exacerbation

## 1 Einleitung

Die *chronisch obstruktive Lungenerkrankung* (engl. Chronic Obstructive Pulmonary Disease, COPD) stellt weltweit die dritthäufigste Ursache für Morbidität und Mortalität dar [1]. Laut der *World Health Organization* starben im Jahr 2019 weltweit 3,23 Millionen Menschen an einer COPD [1]. Es handelt sich hierbei um eine chronisch fortschreitende Atemwegserkrankung, deren Krankheitsverlauf irreversibel ist. Als Hauptursache für die Erkrankung sind das aktive oder passive Rauchen zu nennen [2].

---

Betreuer: Prof. Dr. Christian Thies  
Hochschule: Hochschule Reutlingen  
Christian.Thies@reutlingen-university.de

Informatics Inside Herbst 2021  
24. November 2021 Hochschule Reutlingen  
Copyright 2021 Yailda Aini

Mithilfe eines speziellen Messgeräts, dem Spirometer, kann die Lungenfunktion überprüft werden. Der Spirometer gibt Aufschluss über die Leistungsfähigkeit der Lunge und kann so Auskunft über das Vorliegen einer COPD geben. Ein wichtiger Messwert stellt das *forcierte expiratorische Volumen* (FEV<sub>1</sub>) dar. Bei diesem Test wird das Volumen gemessen, welches sich bei maximal beschleunigter Ausatmung innerhalb einer Sekunde messen lässt. Anhand des FEV<sub>1</sub> wird das COPD-Stadium bestimmt [3]. Zu den klinischen Symptomen der COPD gehören Atemnot, Husten und Auswurf. Im Laufe der Zeit, können Exazerbationen auftreten. Findet eine Exazerbation über mehrere Tage statt, dann spricht man von einer „akuten Exazerbation“. Bei einer akuten Exazerbation werden zusätzliche Behandlungen benötigt, die wiederum mit stationären Krankenhausaufenthalten verbunden sind, sowie zu einer beschleunigten irreversiblen Verschlechterung der Lungenfunktion oder zum Tod führen können [2], [4]. Exazerbationen können vermieden werden, jedoch wird in der Literatur beschrieben, dass es bei vielen COPD-Patienten an Selbstmanagement mangelt und die Adhärenz (Befolgung der verordneten Therapie) nur suboptimal ist [4]. Folglich ist es von zentraler Bedeutung mit Hilfe von modernen Informations- und Kommunikationstechnologien (IKT) das Selbstmanagement zu stärken und den Gesundheitszustand der betroffenen Personen enghesiger zu beobachten. Um die Wirksamkeit digitaler Gesundheitslösungen und die Akzeptanz der COPD-Patienten hinsichtlich der digitalen Technologien zu untersuchen, wird eine systematische Literaturrecherche durchgeführt.

## 2 Methodik

Um einen Überblick über die Thematik zu erhalten, erfolgte eine erste Recherche über *Google Scholar*. Keywords der Suche waren unter anderem: *digital health, digital health technology, chronic obstructive pulmonary disease, COPD*. Daraufhin wurden die Forschungsfragen definiert (s. Kapitel 2.1) und

anschließend erfolgte der Aufbau der Analyse. Nachfolgend fand die Sichtung der Artikel hinsichtlich der Qualität statt. Dies umfasste die Datenextraktion und -synthese. Abschließend wurden die Erkenntnisse der Studien wiedergegeben und diskutiert. In den folgenden Unterkapiteln wird die Vorgehensweise der systematischen Literaturrecherche näher erläutert.

### 2.1 Forschungsfragen

Die erste Recherche ergab, dass insbesondere Themen wie „Fernüberwachung“ (engl. Remote Monitoring, RM) und „COPD-Selbstmanagement“ eine zentrale Rolle einnehmen, um Exazerbationen entgegenzuwirken. Diesbezüglich wurden Forschungsfragen festgelegt, welche in Tabelle 1 dargestellt sind.

**Tabelle 1. Forschungsfragen**

<b>RQ1</b>	Kann das Selbstmanagement von COPD-Patienten durch digitale Gesundheitslösungen verbessert und damit verbunden Exazerbationen reduziert werden?
<b>RQ2</b>	Können durch die Überwachung der Symptome und Vitalparameter Exazerbationen frühzeitig erkannt werden?
<b>RQ3</b>	Wie ist die Akzeptanz der COPD-Patienten bei der Nutzung digitaler Gesundheitslösungen?

### 2.2 Suchstrategie und Artikelauswahl

Nach der Festlegung der Forschungsfragen, wurden für die Recherche nach relevanten wissenschaftlichen Artikeln folgende Datenbanken festgelegt: *Web of Science, PubMed, ScienceDirect* und *IEEE*. Um einen umfassenden Einblick in die digitalen Interventionen für Personen mit COPD zu erhalten, wurden folgende Suchbegriffe festgelegt:

- *digital health,*
- *digital health technology,*
- *telemedicine,*
- *monitoring,*
- *chronic obstructive pulmonary disease,*
- *COPD.*

Damit zusammenhängende Arbeiten gefunden werden konnten, wurden diese Begriffe mit Hilfe von Operatoren wie *AND* und *OR* verknüpft und für die Suche verwendet: „(digital health OR digital health technology OR telemedicine OR monitoring) AND (COPD OR chronic obstructive pulmonary disease)“. Anhand dieses Suchterms wurde in den zuvor genannten Datenbanken nach Literatur gesucht. Nachfolgend werden die Ein- und Ausschlusskriterien für die Auswahl der Studien sowie die Datensynthese näher erläutert.

### 2.2.1 Kriterien für die Studienauswahl

Für die Auswahl der Studien wurden Ein- und Ausschlusskriterien festgelegt (s. Tabelle 2 und Tabelle 3). Um einen aktuellen Überblick über den Forschungsstand zu erhalten, wurden Studien, die im Jahr 2020 und 2021 publiziert wurden, berücksichtigt. In der Abbildung 1 wird anhand eines Flussdiagramms ein Überblick über den Suchprozess und die Auswahl der inkludierten Studien für die Arbeit dargestellt.

**Tabelle 2. Einschlusskriterien**

<b>E1</b>	Studien, die im Jahr 2020 und 2021 veröffentlicht wurden
<b>E2</b>	E1 + auf frühere Studien verweisen, die für das Verständnis benötigt werden
<b>E3</b>	Journals, Reviews und Konferenzbeiträge
<b>E4</b>	Studien mit dem Fokus auf COPD
<b>E5</b>	Studien, welche das COPD-Selbstmanagement unterstützen
<b>E6</b>	Studien, im Zusammenhang mit der Nutzung von IKT und RM

**Tabelle 3. Ausschlusskriterien**

<b>A1</b>	Studien mit größerem Fokus auf andere Atemwegserkrankungen
<b>A2</b>	Studien lediglich mit Fokus auf die medizinische Forschung
<b>A3</b>	Studien mit Fokus auf RM und medizinischem Personal

### 2.2.2 Datensynthese

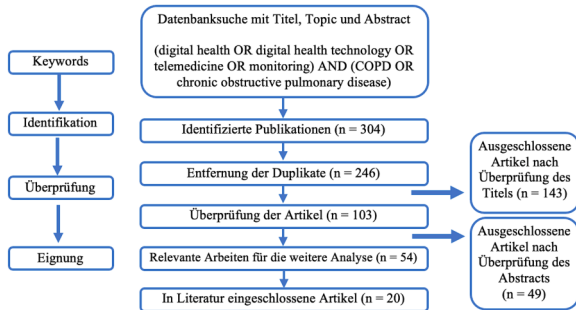
Die Suche ergab mit den in Kapitel 2.2 definierten Suchbegriffen insgesamt 304 Publikationen. Die Literaturergebnisse wurden in das Literaturverwaltungsprogramm *Zotero* überführt. Mit Hilfe von *Zotero* ließen sich die Duplikate (n=58) entfernen. Anschließend erfolgte die Sichtung der Publikationstitel nach den definierten Kriterien. Dabei wurden 143 Publikationen exkludiert. Nachdem wurden die Abstracts der Publikationen gesichtet und 49 der Publikationen ausgeschlossen. Es wurden 54 relevante Publikationen für die Analyse eingeschlossen und genauer untersucht. Nach der näheren Untersuchung erfolgte eine weitere Selektion hinsichtlich der Forschungsfragen. Für diese Arbeit wurden 20 Artikel eingeschlossen.

## 3 Erkenntnisse der Studien

Wie in 2.2.1 dargestellt, lag der Fokus auf Studien, die im Jahr 2020 und 2021 publiziert wurden. 30 % der inkludierten Studien wurden im Jahr 2020 publiziert und 45 % der Studien im Jahr 2021. Weitere Studien (25 %) die eingeschlossen wurden, sind zwischen 2016 und 2019 publiziert worden [12], [16], [17], [20], [24]. Angemerkt sei hier, dass einige Studien auf frühere Artikel verwiesen haben, die eine detailliertere Beschreibung der Interventionsmerkmale enthielten. Da diese für ein besseres Verständnis benötigt wurden, wurden diese Publikationen ebenfalls in Betracht gezogen. Im Folgenden wird auf das Studiendesign (3.1 Studienmerkmale, 3.2 Interventionsmerkmale) eingegangen und daraufhin werden die Studien näher beschrieben.

### 3.1 Studienmerkmale

Die Studien werden überwiegend durchgeführt, um die Auswirkung der IKT auf das COPD-Selbstmanagement und die Akzeptanz der Betroffenen zu untersuchen. Das Durchschnittsalter der in den Studien eingebundenen Personen lag bei  $\geq 60$  Jahren. StudienteilnehmerInnen wurden aus verschiedenen Orten wie Atemwegsambulanzen, Rehabilitationsprogrammen oder nach Krankenhauseinweisungen rekrutiert. Vier Studien



**Abbildung 1: Flussdiagramm – Prozess der Such- und Artikelauswahl**

schlossen Personen ab einem Alter von  $\geq 40$  Jahre ein [5-8]. Andere Studien nahmen Personen ab 18 Jahre auf oder wählten diese nach dem Schweregrad des COPD-Stadiums aus [9-13]. Die Studien aus Dänemark ( $n=281$ ) [12], USA ( $n=138$ ) [6] und Kanada ( $n=122$ ) [9] weisen die breitesten Stichprobengrößen auf. Angemerkt sei hier, dass einige Studien eine Stichprobengröße von mehr als 100 Personen hatten, die sich aus den Interventions- und Kontrollgruppen zusammensetzten. Drei Studien wurden als einarmige Studien mit 11 [7], 29 [10] und 90 [11] Personen durchgeführt. Die Studie [9] stellte eine dreiarmlige Studie dar. Die Studiendauer reichte von drei [8] bis 18 Monate [10], wobei die Intervention von Buul et al. [10] keinen Endzeitpunkt enthielt, da die Intervention als Teil der üblichen Versorgung im Bravis-Krankenhaus (Roosendaal, NLD) eingesetzt wird. Die durchgeführten Studien liefern Informationen darüber, ob eine Intervention sich positiv auf diverse Aspekte des COPD-SM ausgewirkt hat und ob dadurch Patienten Fähigkeiten erlangt haben, sodass diese mit ihrer Erkrankung im Alltag selbstsicher umgehen können. Um Aufschluss über die Wirksamkeit der Interventionen zu erhalten, werden zum Studienbeginn verschiedene Patientendaten erhoben, die dann anhand von Prä-Post-Vergleiche, statistischer Analysen oder maschinellen Lernverfahren mit den ermittelten Daten im späteren Studienverlauf verglichen werden. Die Daten umfassen unter anderem die Lebensqualität, Selbstmanagement, Selbstwirksamkeit,

Vitalparameter oder Anzahl der Exazerbationen. Daten wie die Lebensqualität, werden anhand von krankheitsspezifischen Messinstrumenten in Form von Fragebögen ermittelt [5]. Zur Ermittlung der Akzeptanz wird nach Studienende von dem Betroffenen eine Rückmeldung hinsichtlich der Intervention eingeholt. Die Ein- und Ausschlusskriterien der Studien fallen sehr heterogen aus und umfassen beispielsweise Daten wie FEV<sub>1</sub>, COPD-Stadium, Komorbiditäten, Raucherstatus, physische und psychische Einschränkungen, welche die Nutzung der Interventionen unausführbar machen, Mobilität oder Wohnort.

### 3.2 Interventionsmerkmale

Zur Fernüberwachung von COPD-Patienten und Unterstützung des SM, kommen in den Studien diverse Technologien zum Einsatz. In vier Studien werden die Interventionen über ein Smartphone durchgeführt [5], [6], [10], [13]. Fünf Studien führten die Interventionen über mobile Tablets aus [7-9], [11], [12]. In vier der Studien werden weitere Peripheriegeräte wie z.B. ein Spirometer, Pulsoximeter, Thermometer oder eine Gewichtswaage verwendet [7], [9], [11], [12]. Abhängig von der Zielsetzung der Studie werden weitere Technologien wie Smartwatches [7], [10], Sensoren für Inhalationsgeräte [6], [7] oder Hand-Free Überwachungssensoren zur Messung von Vitalparameter und der Aktivität [19] eingesetzt. Die Peripheriegeräte lassen sich per Bluetooth mit der jeweiligen App verbinden und liefern weitere relevante

Daten zur Ermittlung des Gesundheitszustands. In einigen Studien werden automatisierte Systeme zur Unterstützung der Entscheidungsfindung (engl. Decision Support Systems, DSS) eingesetzt. Beim Überschreiten von Schwellenwerten, aufgrund der Selbstauskunft zunehmender Symptome oder der erfassten Vitaldaten, wird vom DSS eine individuelle Handlungsanweisung, basierend auf den Aktionsplan, an den Betroffenen gesendet [7], [8], [10], [11]. Ein Aktionsplan dient zur frühzeitigen Erkennung von Symptomen und gibt Schritt-für-Schritt Anweisungen, die bei zunehmenden Symptomen durchgeführt werden sollten. In einer Studie wird bei der Überschreitung des Schwellenwerts das medizinische Gesundheitspersonal nicht direkt miteingeschlossen [10]. In allen anderen Studien wird im Falle einer Überschreitung des Schwellenwerts das betreuende Personal z.B. über eine SMS, E-Mail oder Dashboard kontaktiert. Drei der Studien geben an, dass die Intervention unter Mitwirkung von Betroffenen und fachärztliches Personal konzipiert und entwickelt wurden [8], [10], [11].

### 3.3 Ziele und Beschreibung der Studien

Das Ziel der Studien ist es die Wirksamkeit digitaler Selbstmanagement (SM) Interventionen bei COPD-Erkrankten zu untersuchen, das SM zu stärken und Exazerbationen frühzeitig zu erkennen. Zum COPD-SM zählt beispielsweise die richtige Verhaltensweise bei Atemnot, Adhärenz und die korrekte Ausführung der Inhalationsmedikation, das selbstständige Erkennen und Behandeln von Exazerbationen, die Verhaltensveränderung in Bezug auf das Rauchen sowie die körperliche Betätigung [9], [13]. Für viele Betroffene stellt jedoch die COPD eine große Herausforderung dar, sodass zuvor genannte Aspekte überwiegend suboptimal beziehungsweise nicht ausgeführt werden. Dies führt zu einer rapiden Verschlechterung des Gesundheitszustands und führt somit zu weiteren immensen Einschränkungen der Lebensqualität [13], [14]. Um das SM auf

langfristige Sicht zu verbessern, ist es laut Wang et al. [5] notwendig, dass digitale Interventionen Komponenten einsetzen, die das Wissen, die Motivation und die Fähigkeiten von Betroffenen unterstützen. Daher basiert die mobile SM-Intervention von Wang et al. auf ein *Informations-Motivations-Verhaltens-Modell*, welches drei Module zur Verfügung stellt. Die Interventionsgruppe (IG) erhielt zusätzlich zur Standardversorgung Zugriff auf die mobile Intervention. Ersteres vermittelt wichtige Informationen über die Erkrankung, mit dem Ziel das Wissen der Erkrankten zu verbessern, um eine bessere Gesundheitskompetenz zu erzielen [5]. Ferner werden visuelle Elemente (Bilder, Videos) genutzt, um Informationen oder Sport- und Atemtechnikübungen reizvoller zu gestalten. Ziel ist es, den Erkrankten Kompetenzen bezüglich der Krankheitskontrolle (z.B. Verhalten bei Atemnot) nahezubringen. Letzteres zielt auf die Motivation ab und bietet einen Peer-Support Chatroom an, in dem sich Betroffene austauschen und gegenseitig motivieren können. Zudem bietet es ein weiteres Portal für Expertenunterstützung an [5]. Dieser Ansatz wies in der IG im Vergleich zur Kontrollgruppe (KG) eine signifikante Verbesserung hinsichtlich des Selbstmanagementverhaltens (CSMS), der Lebensqualität (CAT), Anzahl sportlich aktiver TeilnehmerInnen und des Raucherverhaltens (s. Tabelle 4) auf. Ein höherer CSMS-Wert weist auf ein besseres SM und ein niedriger CAT-Wert auf eine bessere Lebensqualität hin. Der CAT umfasst Fragen bezüglich der Symptome, Aktivität, Selbstvertrauen und liefert ein Maß zum Schweregrad der COPD [5].

**Tabelle 4. Auswirkung der Intervention [5]**

Auswirkung(mean)	Gruppe	Studienbeginn	Nach 12 M.
CSMS	IG	127.73	150.74
	KG	124.52	126.80
CAT	IG	22.6	18.3
	KG	22.2	20.9
Aktivität	IG	14	32
Raucher	IG	11	2

CAT=COPD Assessment Test, CSMS=COPD Self-Management Scale [5]

Ein beständiges Problem, stellt die mangelnde Adhärenz der Medikation und die fehlerhafte Durchführung der Inhalationstechnik (IHT) dar [14], [15]. Dabei soll die Adhärenz zwischen 40 % und 57 % liegen [6]. Die fehlerhafte Durchführung der IHT hat zur Folge, dass das Medikament nicht bis in die Lunge und somit nicht am Wirkungsort angelangt. Damit wird das Risiko einer Exazerbation erhöht [15].

Um Abhilfe zu schaffen, wird in der Intervention von Sloots et al. [7] ein zusätzlicher Sensor von *Respiro*, das kompatibel mit dem Inhalator *Ellipta* ist, eingesetzt, um die Adhärenz und die Ausführung der IHT zu messen [7]. Der Sensor enthält ein Zählwerk, dass die Inhalationsanzahl erfasst [16]. Die korrekte Durchführung der IHT wird anhand der Dauer des Inhalationsflusses (Stärke des Atemzugs beim Einatmen), dem Spitzeninspirationsfluss (engl. Peak Inspiratory Flow, PIF), der Ausrichtung und anhand des korrekten Verschlusses der Kappe des Inhalators gemessen. Als Kriterium für die korrekte IHT wird die Dauer des Inhalationsflusses als  $> 1,25$  Sekunden definiert [7]. Grund hierfür ist, dass Inhalationen unter der Mindestschwelle von einer Sekunde gerätebedingt als suboptimal registriert werden [16]. Die Ermittlung der PIF gibt Aufschluss darüber, ob Betroffene in der Lage sind, aus einem Inhalator ausreichend zu inhalieren [17]. Der PIF wurde hier  $> 30\text{L}/\text{min}$  definiert, um eine effektive Behandlung zu ermöglichen. Durch Algorithmen des maschinellen Lernens, werden die Daten aus dem Sensor mit den festgelegten Schwellenwerten verglichen und bei fehlerhafter Anwendung werden Betroffene über die App darauf aufmerksam gemacht [7], [16]. Die Studie zeigte eine hohe Adhärenz bezüglich der Medikationseinnahme (98,4 %) und der täglichen Symptomeingaben (91 %). Der häufigste Inhalationsfehler war zu kurzes Einatmen (79,6 %) gefolgt von zu schwachem Einatmen mit 16,9 %. Trotz der hohen Adhärenz

zeigten die TeilnehmerInnen gemäß der gerätebedingten und von [7] festgelegten kritischen Kriterien eine schlechte IHT auf. Während der Nachbeobachtungszeit wurden von 752 durchgeführten Inhalationen, lediglich 390 (51,9 %) korrekt durchgeführt [7]. Angemerkt sei, dass es diverse Inhalationsgeräte gibt, die sich sowohl technisch als auch in der Ausführung voneinander unterscheiden. Demzufolge gibt es zur Ermittlung der IHT gerätebedingte Checklisten, welche festgelegte Definitionen von kritischen Fehlern enthalten und als Maß für eine gute oder schlechte IHT dienen [7], [8]. Die *Deutsche Atemwegsliga e.V.* stellt verschiedene Checklisten für die richtige IHT diverser Inhalationsgeräte zur Verfügung<sup>1</sup>. Erinnerungen an die Medikation erfolgten über audiovisuelle Zeichen des Sensors und der App. Zudem wurde ein *Emodied Conversational Agent* (ECA) eingesetzt. Es handelt sich dabei um ein intelligentes, interaktives Softwareprogramm mit einer Verkörperung, das Betroffene über die App Face-to-Face an Symptomeingaben erinnert, Rückmeldungen bezüglich der Adhärenz der Inhalationsmedikation und -technik gibt und über diverse gesundheitliche Themen unterrichtet [7].

In Criner et al. [6] wird das *BreathMate* Monitoringgerät, zur Erfassung der Adhärenz sowie der Unter- und Überdosierung der Inhalationsmedikation eingesetzt. Es werden Datum und die ausgeführte Inhalation erfasst. In der IG wurde beobachtet, dass durch audio-visuelle Erinnerungen an die Medikation der Nicht-, Unter-, und Übergebrauch im Vergleich zur KG signifikant geringer war. Laut Criner et al. könnte hiermit der Medikamentenmissbrauch verringert werden. Hierfür sind jedoch weitere Studien notwendig [6]. Im Gegensatz zur Studie von Sloots et al. [7] wird in Criner et al. zusätzlich präventiv sowie beim Auslassen der Medikation über die App an die Wichtigkeit der Adhärenz erinnert [6]. In der Studie von North et al. [8] stand der IG Videos zur Inha-

<sup>1</sup> <https://www.atemwegsliga.de/checklisten.html>

lationstechnik zur Verfügung. Zum Studienende hatte sich die Anzahl der kritischen Fehler in der IG von 101 auf 20 kritische Fehler reduziert. Im Vergleich zur Standardversorgung, die einen papierbasierten Aktionsplan zur IHT erhielten, wurde eine Verbesserung von 100 auf 72 kritische Fehler festgestellt. Durch eine Längsschnittanalyse konnten nach drei Monaten klinische Verbesserungen des CAT-Werts beobachtet werden, was auf eine bessere Krankheitskontrolle deutete. Zudem wurde ein Rückgang der registrierten Exazerbationen festgestellt (s. Tabelle 5) [8].

**Tabelle 5. Auswirkung der Intervention [8]**

Auswirkung(mean)	Gruppe	Studienbeginn	Nach 3 M.
CAT	IG	26,0	20,7
	KG	28	25,1
Exazerbation	IG	2,9	1,06
	KG	3,2	1,88

CAT=COPD Assessment Test

Je nach COPD-Stadium können bei Betroffenen Atemnot aufgrund körperlicher Belastung (Belastungsdyspnoe) oder bereits in Ruhe (Ruhedyspnoe) eintreten. Viele Erkrankte vermeiden daher körperliche Aktivitäten. Diese Schonungshaltung führt zu einem Abbau der Muskulatur, einschließlich der Atemhilfsmuskulatur, die jedoch für COPD-PatientInnen von zentraler Bedeutung sind, da diese die Ein- und Ausatmung unterstützen. Es ist bekannt, dass Exazerbationen mit einer geringen körperlichen Aktivität einhergehen [18].

Daher wird in der retrospektiven Studie von Buul et al. [10] untersucht, ob zusätzlich zur Überwachung der subjektiven Symptomangaben, die körperliche Aktivität als prädiktiver Indikator zur frühzeitigen Erkennung von Exazerbationen genutzt werden kann. Die verwendete SM-Intervention nennt sich *EmmaCOPD*. Zur Überwachung der körperlichen Aktivität wird eine programmierbare Smartwatch *Pebble Time* genutzt, welche mit einem Beschleunigungssensor und

Gyrometer ausgestattet ist, um die Anzahl der Schritte der TeilnehmerInnen zu überwachen. Um die Aktivität eines Patienten vergleichen zu können, wurde in den ersten drei Wochen das Ausgangsniveau der Aktivität (Schrittzahl) ermittelt und anschließend eine individuelle Schrittzahl von einem Physiotherapeuten festgelegt [10]. Zur Vorhersage beziehungsweise Verhinderung von Exazerbationen wird ein digitaler Aktionsplan für Lungenentzündungen<sup>2</sup> verwendet, welcher Einblick in den Symptomverlauf ermöglicht. Je nach Symptomeingaben oder Abnahme der körperlichen Aktivität können sich TeilnehmerInnen in vier Zonen (grün, gelb, orange, rot) bewegen und erhalten dementsprechend Handlungsanweisungen durch das Decision Support System. Wird beispielsweise ein Rückgang der Schrittzahl um 40 % vom System erkannt, wird eine Nachricht an die betroffene Person und einer nahestehenden Person versendet, mit der Information, dass sich dieser in der „roten Zone“ befindet und ein Notruf getätigt werden soll. Durch die Intervention konnte ein signifikanter Rückgang der Krankenhaustage und der Gesamtzahl der Exazerbationen festgestellt werden. 12 Monate vor Beginn der Intervention betrug die Gesamtzahl der Exazerbationen 2,0 (median) und die Anzahl der Krankenhaustage 8,0 (median). 18 Monate nach Beginn der Intervention wurde ein Median von 1,0 bei der Gesamtzahl von Exazerbationen und 2,0 Krankenhaustage festgestellt. Daraus erschließt sich, dass durch die frühen Meldungen TeilnehmerInnen schneller bei Anzeichen einer Exazerbation selbstständig reagiert haben. Demnach kann eine Abnahme der körperlichen Aktivität ein erster Hinweis auf eine mögliche Exazerbation sein [10].

COPD führt häufig zu Atemnot. Als Reaktion darauf, folgt eine verstärkte Atmung, wodurch sich die Atemfrequenz (Respiratory Rate, RR) erhöht [19].

<sup>2</sup><https://bestellen.longfonds.nl/product/longaanval-actieplan-copd/>



Polsky et al. [19] stellt einen Anwendungsfall eines kommerziellen RPM-Systems *Spire Health* vor, welcher im klinischen Umfeld zur frühzeitigen Erkennung von Exazerbationen eingesetzt wird [19]. Der verwendete Sensor *Spire Health Tag* umfasst ein Photoplethysmogram (PPG-Sensor), ein drei-Achsen-Beschleunigungsmesser und einen speziellen Kraftsensor [19], [20]. Der Sensor kann an der Unterwäsche oder bei Frauen an der Innenseite des BH-Flügels angeklebt werden. Abbildung 2 stellt den Sensor dar. Ein wesentlicher Vorteil des Sensors ist, dass es nach Anbringen und einmaligem Verbinden durch Bluetooth mit dem jeweiligen Endgerät keine weitere Interaktion durch den Patienten mehr benötigt [20].



Abbildung 2: Spire Health Tag [20]

Anhand des PPG-Sensors lässt sich die Pulsfrequenz (Blutstrom in den Gefäßen) ermitteln. Hierbei wird eine hohe Lichtintensität auf eine Hautstelle gestrahlt. Die Reflektion des Lichts, das durch den Sensor erfasst wird, gibt Aufschluss über die Pulsfrequenz. Der Beschleunigungsmesser ermöglicht die Ermittlung der Aktivität. Damit die Atemfrequenz gemessen werden kann, werden die Bewegungen und Volumenänderungen, die sich beim Atmen durch den Brustkorb und das Abdomen gegen den Kraftsensor des Geräts ergeben, erfasst. Mit einer Abtastrate von 25 Hz, werden die vom Sensor erzeugten Atmungswellenformen, welche die Ein- und Ausatmung darstellen, abgetastet [20]. In Abbildung 3 wird der Anwendungsfall von [19] dargestellt. Der Sensor erkannte beim Patienten eine tägliche RR von  $\geq 10\%$  über dem täglichen Ausgangswert der RR und alarmierte das klinische Team. Trotz der Benachrichtigung des klinischen Teams vermutete der Patient die erhöhte RR und die leichten Symptome auf die körperliche Arbeit zu Hause und suchte keinen Arzt auf. Nach fünf Tagen ging ein erneuter Alarm aufgrund des

Anstieges der RR und der Pulsfrequenz ein und beim Patienten zeigten sich zunehmende Symptome auf. Es wurde eine Exazerbation diagnostiziert [19]. Zu erkennen ist, dass eine erhöhte RR bis zum Eintreten der Exazerbation einen gewissen Zeitraum einnehmen kann. Dieser Zeitraum wird als *Prodromalstadium* bezeichnet und stellt eine frühe Krankheitsphase dar, indem Erkrankte erst nach diesem Stadium die ersten Anzeichen der Exazerbationssymptome aufweisen. Folglich können präventive Maßnahmen frühzeitig eingeleitet und damit verbunden eine verbesserte Versorgung gewährleistet werden [11], [19]. Eine erhöhte RR kann somit als prädikativer Indikator gesehen werden, das auf frühzeitige Exazerbationen hinweisen kann [19].

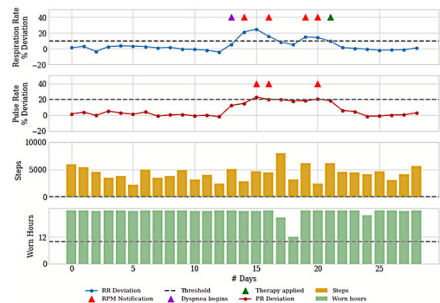


Abbildung 3: Erkannte Exazerbation durch Spire Health RM [19]

Ferner wurde in einer Zwischenanalyse von Spire Health festgestellt, dass durch Verfahren des maschinellen Lernens zur Neuheitserkennung, auffällige Atemzüge, die vom Health Tag Sensor erfasst und sich von der üblichen Patientenphysiologie differenzieren, identifiziert werden können, sodass Alarme noch vor einem Krankenhausaufenthalt ausgelöst werden. Die mittlere Benachrichtigungsdauer betrug 4,4 Tage bei 16 von 19 identifizierten Krankenhausaufenthalten [21].

In der Studie von Patel et al. [11] wird die Validität des SM-Fernüberwachungs-App *COPDPredict* untersucht. Folgende Punkte werden im Rahmen der Studie von Patel et al. evaluiert: Unterscheidung zwischen ver-

schiedenen Gesundheitsphasen (Stabil, Prodromal, Exazerbation), die Vorhersagbarkeit von Exazerbationen sowie die rechtzeitige Benachrichtigung der Betroffenen und des betreuenden Personals. Zur Bestimmung des Gesundheitszustands und Vorhersagbarkeit bevorstehender Exazerbationen wird ein Algorithmus, dass auf einem Entscheidungsbaum basiert, verwendet [11]. Der Entscheidungsbaum setzt sich aus den prozentualen Schwellenwerten der Veränderung des Gesundheitszustands zusammen. Diese basieren auf den subjektiven Angaben des Wohlbefindens eines Patienten, dem FEV<sub>1</sub> und das C-reaktive Protein (CRP). Das Wohlbefinden der TeilnehmerInnen wird anhand von fünf Fragen bezüglich Symptomatik und Aktivität ermittelt. Als Antwortskala wird eine fünfstufige Likert-Skala genutzt, dessen Gesamtwert anschließend automatisch ermittelt wird. Der CRP-Wert wird durch eine Blutprobe mittels Fingerstich-Test gemessen und gibt Aufschluss über Entzündungen im Körper [11]. Zur Messung des FEV<sub>1</sub> wird ein Bluetooth-fähiger Spirometer verwendet.

Um für jeden Patienten individuelle kritische Ausgangswerte festlegen zu können, wird das Wohlbefinden der PatientInnen in den ersten zwei Studienwochen täglich ermittelt. Ebenfalls wird zu Studienbeginn und am 14. Tag der CRP-Wert sowie an jedem dritten Tag der FEV<sub>1</sub>-Wert, ermittelt [11]. Dies stellte zudem die Lernphase des Vorhersagemodells (VM) für Exazerbationen dar.

Anschließend erfolgte bis zum Ende des Studienzeitraums die sequenzielle Aufteilung des Entscheidungsbaums durch die täglichen Angaben des Wohlbefindens und der wöchentlichen FEV<sub>1</sub>-Messung [11]. Im Falle einer Überschreitung des Schwellenwerts zuvor genannter Daten, werden weitere Tests (FEV<sub>1</sub>, CPR-Test) angefordert. Um Alarmschwellenwerte für die Veränderung des Gesundheitszustands - stabil und prodromal - festzulegen, wird der prozentuale Unterschied zwischen dem Median des Prodromal-Wert, dem Ausgangswert und die absoluten Werteänderungen festgelegt [11]. Die Alarmschwellenwerte für stabile Zustände

bis hin zur Exazerbation, werden anhand von Receiver-Operating-Curves (ROC) festgelegt. Um zu untersuchen, ob die vom VM ausgelösten Benachrichtigungen eine Frühwarnung vor einer bevorstehenden Exazerbation darstellt, wird für das Prodromalstadium einer Exazerbation ein Zeitraum von zwei Wochen vor dem Auftreten einer vom ärztlichen Fachperson definierten Exazerbation festgelegt. Durch einen Vergleich zwischen den vom Algorithmus definierten Exazerbationsmeldungen und den von den Ärzten definierten Exazerbationen konnte eine Kreuzanalyse durchgeführt werden, sodass die Genauigkeit, Sensitivität, Spezifität sowie der bevorstehende Zeitpunkt zur Erkennung von Exazerbationen des VM ermittelt werden konnten. Zur Validierung des Algorithmus wurden insgesamt 3257 Messwerte verwendet. Die Ergebnisse der Kreuztabellenanalyse ergaben, dass der Algorithmus bereits Exazerbationen 7 Tage vor einem vom Arzt definierte Exazerbation identifizieren konnte. Die Genauigkeit des Algorithmus betrug bezüglich der Vorhersagbarkeit einer Exazerbation 85,3 %. Der Algorithmus hat in 285 von 291 Messzeitpunkten auf eine bevorstehende Exazerbation hingewiesen. Dabei lag die Sensitivität bei 97,9 %, die Spezifität bei 84 % und der negative Vorhersagewert bei 99,8 %. Demnach ist COPD-Predict in der Lage Exazerbationen sehr genau zu erkennen und den Betroffenen sowie das ärztliche Fachpersonal frühzeitig zu benachrichtigen. Eine weitere Analyse mit lediglich Eingabe des Wohlbefindens und der FEV<sub>1</sub> zeigte, dass der Algorithmus bei der Vorhersage bevorstehender Exazerbationen eine Sensitivität von 77% und eine Spezifität von 65 % erzielte. Folglich können Biomarker wie CPR die Vorhersagbarkeit von Exazerbationen optimieren [11].

## 4 Diskussion

Nachdem diverse Interventionen zur Unterstützung des SM und zur Erkennung von Exazerbationen vorgestellt wurden, werden in diesem Kapitel die Erkenntnisse diskutiert und abschließend die Forschungsfragen (s. Kapitel 5) dieser Arbeit beantwortet.

Die Integration von Schulungssitzungen vor einer Intervention und das Anbieten von Aufklärungs- und Motivationskomponenten können sich, wie Wang et al. [5] nachgewiesen haben, positiv auf die Verhaltensänderung und das Selbstmanagement eines Patienten auswirken. Studien die zuvor genannte Komponenten einsetzten, betonten, dass diese ebenfalls starken Einfluss auf die Adhärenz haben können [7], [8], [11] und daher fester Bestandteil einer COPD SM-Intervention sein sollten [23]. Dies steht im Einklang mit den Erkenntnissen aus der Meta-Analyse von Hong et al. [24] die ferner bekräftigen, dass vor allem Betroffene mit mittel bis sehr schwerem COPD-Schweregrad von solchen Komponenten profitieren könnten. Laut Stamenova et al. [9] ist das der zentrale Grund gewesen, weshalb keine signifikanten Unterschiede in Bezug des SM, Wissen und Symptome in Ihrer Studie erzielt werden konnten.

Es muss berücksichtigt werden, dass COPD-Patienten häufig an mehreren Komorbiditäten, wie koronare Herzkrankheiten, Depressionen oder Angststörungen leiden. Laut Sloots et al. [7] sollen daher ebenfalls Symptome der Komorbiditäten erfasst und zudem Aktionspläne für multimorbide Exazerbationen in die Intervention integriert werden, um Betroffene eine sichere Versorgung zu ermöglichen. Daher wird in [7] z.B. der *NT-proBNP-Wert*, welches Aufschluss über Schädigungen am Herzen geben kann, ermittelt. Lediglich drei Studien berücksichtigen die Angst und Depressionen [7], [8], [13].

In der Studie von Sloots et al. wurde festgestellt, dass sechs generierte Anweisungen, die das Anrufen des zuständigen Personals beinhalteten, nicht durchgeführt wurden. Als Grund wird genannt, dass der Schwellenwert zu hoch festgelegt wurde, sodass Betroffene aus eigener Erfahrung in Bezug auf Exazerbationen, bereits früher einen Arzt aufgesucht oder sich selbst behandelt hatten [7]. Interventionen, die statische und manuell eingestellte Schwellenwerte nutzen, können

daher unvorteilhaft sein [7], [11]. Wie in Patel et al. vorgestellt wurde, kann der Einsatz intelligenter Algorithmen Abhilfe schaffen, indem automatisch patientenspezifische Ausgangsprofile (Schwellenwerte) erstellt werden [11].

Weiterhin werden verschiedene Vitalparameter der PatientInnen überwacht, um einen Überblick über den Gesundheitszustand zu erhalten und so Exazerbationen frühzeitig zu identifizieren. Welche der prädikativen Indikatoren (z.B. Aktivität, FEV<sub>1</sub>, Atemfrequenz) sich jedoch am besten dafür eignen, scheint noch nicht ganz klar zu sein. Die Intervention von Stamenova et al. wies das umfangsreichste RM-Geräteset auf, sodass mehrere Parameter ermittelt werden konnten. Dennoch konnten keine signifikanten Veränderungen des SM erzielt und somit kein Beitrag zur Vermeidung von Exazerbationen geleistet werden [9]. Wie Patel et al. jedoch nachweisen konnten, kann die Robustheit von Algorithmen durch Biomarker (z.B. CRP) erhöht werden, da keine subjektiven Einflüsse gegeben sind [11].

Trotz der positiven Ergebnisse betonen die Studien, dass weitere Forschungsarbeiten mit längeren Nachbeobachtungszeiten und größeren Patientenkohorten notwendig sind, um Langzeitnachweise der Interventionen zu erhalten und die Verallgemeinerbarkeit zu gewährleisten [6-8], [10]. Denn hinsichtlich des Langzeiteffekts der Interventionen liegen keine Ergebnisse vor, da die Dauer der einzelnen Studien nicht dem einer Langzeitstudie gleichen. Patel et al. [11] erwähnt, dass eine umfangreiche Studie<sup>3</sup> bereits in Planung ist.

Angemerkt sei, dass sich die Interventionen der Studien stark voneinander unterscheiden. Dies liegt größtenteils daran, dass es keine Standardisierung gibt. Die Unterschiede reichen von der Ausführung der Intervention auf verschiedenen Endgeräten und Betriebssystemen über diverse Peripheriegeräte,

<sup>3</sup> <https://clinicaltrials.gov/ct2/show/NCT04136418>

Häufigkeit der Messungen, Echtzeitübermittlungen und -überwachungen der Daten, verwendete Komponenten und Funktionalitäten, Visualisierung der Gesundheitsdaten sowie genutzte Kommunikationswege zwischen Erkrankten und betreuendem Personal. Diesbezüglich beschreiben Korpershoek et al. [23] einen umfassenden iterativen nutzentrierten Entwicklungsprozess, der Forschern und Entwicklern bei der Entwicklung von evidenzbasierten und benutzerfreundlichen mobilen Interventionen zur Verbesserung des Selbstmanagement als Leitlinie dienen kann [23].

## 5 Beantwortung der Forschungsfragen

Das Ziel der systematischen Literaturrecherche war es die Forschungsfragen aus Kapitel 2.1 zu beantworten. Durch die Erkenntnisse aus den Studien lassen sich RQ1 bis RQ3 positiv beantworten.

Die Studienergebnisse haben gezeigt, dass durch den Einsatz digitaler Interventionen diverse Aspekte des COPD-SM unterstützt und verbessert werden können [5-8], [10], [12]. Insbesondere die Adhärenz bezüglich der Medikamenteneinnahme kann durch Echtzeit-Erinnerungen über die App, Sensoren und ECA verbessert werden [6-8], [10]. In einigen Studien konnte durch die Verbesserung des SM ebenfalls eine Verbesserung der Lebensqualität [12] und Selbstwirksamkeit festgestellt werden [5], [7]. Digitale Gesundheitslösungen können demzufolge das COPD-SM verbessern, sodass eine bessere Krankheitskontrolle ermöglicht und folglich das Risiko von Exazerbationen gesenkt werden kann.

RQ2 lässt sich positiv durch die Ergebnisse von [10], [11], [19], [21] beantworten. Die Studien haben gezeigt, dass die Überwachung der subjektiven Symptomangaben in Kombination weiterer Indikatoren, wie die Aktivität, Atemfrequenz oder bio-physiologische Parameter, ermöglichen Exazerbationen frühzeitig zu erkennen. Dabei betonen einige Studien, dass zukünftig die Entwicklung von Algorithmen zur Früherkennung

von Exazerbationen für künftige RM-Systeme eine zentrale Rolle spielen werden [9], [11], [21].

Die Akzeptanz der Patienten bezüglich der Interventionen (RQ3) wurden anhand von Fragebögen nach Studienende erhoben. Betroffene die bereits Erfahrung mit Smartphones oder Tablets hatten beziehungsweise eine hohe Motivation aufwiesen, empfanden die Interventionen meist als einfach zu bedienen und zeigten eine höhere Adhärenz als diejenigen, die keinerlei Erfahrung mit diesen Technologien hatten [10], [13]. Grundsätzlich sahen die PatientInnen die smarten Inhalationsgeräte und die Erinnerungsfunktion zur Medikamenteneinnahme als sehr hilfreich [6], [7]. In der Tabelle 6 werden die absoluten und relativen Zahlen der Interventionsabbrüchen innerhalb einiger Studien veranschaulicht.

**Tabelle 6. Anzahl der Interventionsabbrüche (IA)**

Ref.	Anzahl Patienten in IG	IA Absolut	IA Relativ (%)
[5]	39	4	1,56
[6]	68	18	12,24
[8]	20	3	0,6
[9]	81	11	8,91
[10]	29	6	1,74
[12]	141	32	45,12
[13]	41	4	1,64

Ref. = Referenz, IG = Interventionsgruppe

Angemerkt sei, dass in der Studie von [12] 17 TeilnehmerInnen verstorben sind. Es wird nicht näher beschrieben, ob die TeilnehmerInnen zur IG oder KG gehören, daher kann nur eine Gesamtaussage gemacht werden. Gründe für Interventionsabbrüche sind z.B. gewesen: Intervention zu aufwendig, Technologie zu komplex [9], Technische Schwierigkeiten [6], [7], [9], [13], keine Kraft mehr weiterzumachen [12], kein Interesse mehr [5] oder verstorben. Erwähnt wird jedoch auch, dass Personen aufgrund von Bedenken hinsichtlich des Datenschutzes, Desinteresse [7], Angst vor den Technologien oder aufgrund der nicht Verfügbarkeit von Internetzugang und Endgeräte (n=65 in [13]), nicht an den Studien teilnehmen wollten oder konnten [8], [13].

## 6 Fazit

Zusammenfassend kann gesagt werden, dass digitale Gesundheitslösungen eine nützliche und effiziente Lösung sind, um COPD-Erkrankte im Alltag zu unterstützen, ihr Selbstmanagement zu optimieren und damit verbunden das Risiko von Exazerbationen zu reduzieren oder zu verhindern. Die kontinuierliche Fernüberwachung der subjektiven Symptomangaben, der körperlichen Aktivität und der Vitalparameter, die durch diverse Messinstrumente und Geräte ermittelt werden können, ermöglichen einerseits einen Einblick in den Gesundheitszustand beziehungsweise im Krankheitsverlauf der PatientInnen und andererseits lassen sich dadurch abnorme Veränderungen des Gesundheitszustands frühzeitig erkennen, so dass präventive Maßnahmen rechtzeitig veranlasst werden können.

Betrachtet man den demographischen Wandel, so nimmt die Zahl der alternden Menschen und somit auch die Zahl der Erkrankungen von Jahr zu Jahr zu. Zudem ergibt sich durch die COVID-19 Pandemie ein weiterer Risikofaktor für COPD-Patienten, wodurch die Relevanz digitaler Lösungen für diese Patientengruppe nochmals verdeutlicht wird. In der Studie von Boyce et al. [22] gab die Hälfte der Befragten an, dass Sie im Jahr 2020 mit der Nutzung von Telemedizinischen Interventionen begonnen haben (Videotelefonie, Nutzung von Apps zu Atemtechniken oder Online-Rehabilitationsprogramm). Deutlich wird aus dieser Umfrage, dass vor allem COPD-Erkrankte, die am stärksten gefährdet sind, telemedizinische Interventionen in Anspruch nehmen [22]. Die Innovation digitaler Gesundheitslösungen für COPD-Patienten ist somit vielversprechend und hat ein großes Potenzial Erkrankte eine individuelle, effiziente und verbesserte Versorgung und Therapiestaltung zu ermöglichen. Wie gezeigt werden konnte, können mit Hilfe von künstlich intelligenten Algorithmen Exazerbationen frühzeitig erkannt werden. Daher wäre eine intensivere Untersuchung des Forschungsstands in diesem Bereich interessant.

## Literaturverzeichnis

- [1] Statista. Häufigste Todesursachen weltweit im Jahr 2019. (2020). Online verfügbar: <https://de.statista.com/statistik/daten/studie/38995/umfrage/weltweite-odesfaelle-aufgrund-chronischer-krankheiten/> Letzter Zugriff: 14.11.2021.
- [2] 2021 GOLD Report. Global Initiative for Chronic Obstructive Lung Disease. (2021). *Published in Fontana, WI, USA*. Online verfügbar: [www.goldcopd.org](http://www.goldcopd.org). Letzter Zugriff: 14.11.2021.
- [3] Buess C., Schlegelmilch R. M., Kramme R. Spirometrie und erweiterte Lungenfunktionsdiagnostik. *Medizintechnik. Springer Reference Technik*. (2015). DOI: 10.1007/978-3-662-45538-8\_13-1
- [4] Jardim J. R., Nascimento O. A., The Importance of Inhaler Adherence to Prevent COPD Exacerbations. *Medical Sciences*. (2019). 7(4):54. <https://doi.org/10.3390/medsci7040054>
- [5] Wang L., Guo Y., Wang M., et al., A mobile health application to support self-management in patients with chronic obstructive pulmonary disease: a randomised controlled trial. *Clinical Rehabilitation*. (2021). S. 90-101. doi:10.1177/0269215520946931
- [6] Criner G. J., Cole T., et al., The Impact of Budesonide/Formoterol pMDI Medication Reminders on Adherence in Chronic Obstructive Pulmonary Disease (COPD) Patients: Results of a Randomized, Phase 4, Clinical Study. *Int J Chron Obstruct Pulmon Dis*. (2021). S. 563-577. doi: 10.2147/COPD.S253770
- [7] Sloots J., Bakker M., et al. Adherence to an eHealth Self-Management Intervention for Patients with Both COPD and Heart Failure: Results of a Pilot Study. *Int J Chron Obstruct Pulmon*

- Dis.* (2021). S. 2089–2103. Doi: 10.2147/COPD.S299598
- [8] North M., Bourne S., et al., A randomised controlled feasibility trial of E-health application supported care vs usual care after exacerbation of COPD: the RESCUE trial. *npj Digit. Med.* 3, 145. (2020). <https://doi.org/10.1038/s41746-020-00347-7>
- [9] Stamenova V., Liang K., et al., Technology-Enabled Self-Management of Chronic Obstructive Pulmonary Disease With or Without Asynchronous Remote Monitoring: Randomized Controlled Trial. *J Med Internet Res.* (2020). DOI: 10.2196/18598
- [10] Van Buul A., Derksen C., Hoedemaker O., et al., eHealth Program to Reduce Hospitalizations Due to Acute Exacerbation of Chronic Obstructive Pulmonary Disease: Retrospective Study. *JMIR Formative Research.* (2021). DOI: 10.2196/24726
- [11] Patel N., Kinmond K., et al., Validation of COPDPredict<sup>TM</sup>: Unique Combination of Remote Monitoring and Exacerbation Prediction to Support Preventative Management of COPD Exacerbations. *Int J Chron Obstruct Pulmon Dis.* (2021), S. 1887-1899. doi:10.2147/COPD.S309372
- [12] Tupper O. D., Gregersen T.L., et al., Effect of tele-health care on quality of life in patients with severe COPD: a randomized clinical trial. *Int J Chron Obstruct Pulmon Dis.* (2018). S. 2657-2662. DOI: 10.2147/COPD.S164121
- [13] Kooij L., Vos P., Dijkstra A., et al., Effectiveness of a Mobile Health and Self-Management App for High-Risk Patients With Chronic Obstructive Pulmonary Disease in Daily Clinical Practice: Mixed Methods Evaluation Study. *JMIR Mhealth Uhealth.* (2021). DOI: 10.2196/21977
- [14] Schrijver J. W., Effing T., et al., Predictors of patient adherence to COPD self-management exacerbation action plans. *Patient Education and Counseling.* (2021). Volume 104, S. 163-170. DOI: 10.1016/j.pec.2020.06.015
- [15] Elander A., Gustafsson M., Inhaler Technique and Self-reported Adherence to Medications Among Hospitalised People with Asthma and COPD. *Drugs - real world outcomes.* (2020). S. 317–323. DOI:10.1007/s40801-020-00210-x
- [16] Braido F., Paa F., et al., A new tool for inhalers 'use and adherence monitoring: the Amiko<sup>®</sup> validation trial. *Int J Eng Res Sci.* (2016). S. 159–166. Online verfügbar: <https://ijoer.com/Paper-October-2016/IJOER-OCT-2016-25.pdf> Letzter Zugriff: 14.11.2021.
- [17] Kreutz, I., Inhalator-Schulung ist bei COPD das A & O. CME (2018). <https://doi.org/10.1007/s11298-018-6707-0> Letzter Zugriff: 14.11.2021
- [18] Hurst J. R., Skolnik N., et al., Understanding the impact of chronic obstructive pulmonary disease exacerbations on patient health and quality of life. *Eur J Intern Med.* (2020). S. 1–6. <https://doi.org/10.1016/j.ejim.2019.12.014>
- [19] Polsky M. B., Moraveji N., Early identification and treatment of COPD exacerbation using remote respiratory monitoring. *Respiratory medicine case reports.* (2021). DOI: 10.1016/j.rmcr.2021.101475
- [20] Moraveji N., Golz P., et al., Long-term, Ambulatory Respiratory Monitoring of COPD Patients using Garment-Adhered Sensors. *IEEE Int Sym on Med Me and Applic.* (2019). DOI: 10.1109/MeMeA.2019.8802187
- [21] Murray R. K., Holt M., et al., Pre-Hospitalization Notification Using

Novelty Detection in Chronic Obstructive Pulmonary Disease. *American Thoracic Society*. (2020). A5062-A5062. Online verfügbar: [https://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2020.201.1\\_Meeting-Abstracts.A5062](https://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2020.201.1_Meeting-Abstracts.A5062) Letzter Zugriff: 14.11.2021

- [22] Boyce D. M., Thomashow B. M., et al., New Adopters of Telemedicine During the Coronavirus-19 Pandemic in Respondents to an Online Community Survey: The Case for Access to Remote Management Tools for Individuals with Chronic Obstructive Pulmonary Disease. *Int J Chron Obstruct Pulmon Dis*. (2021). S. 213–218. <https://doi.org/10.15326/jcopdf.2020.0181>
- [23] Korpershoek Y., Hermsen S., et al., User-Centered Design of a Mobile Health Intervention to Enhance Exacerbation-Related Self-Management in Patients With Chronic Obstructive Pulmonary Disease (Copilot): Mixed Methods Study. *J Med Internet Res*. (2020). DOI: 10.2196/15449
- [24] Hong Y., Lee S. H., Effectiveness of tele-monitoring by patient severity and intervention type in chronic obstructive pulmonary disease patients: A systematic review and meta-analysis. *Int J Nurs Stud*. (2019). DOI: 10.1016/j.ijnurstu.2018.12.006



# Anforderungen an ein Computer-Vision-System zur automatischen Leistungsdokumentation in der ambulanten Pflege

Stephan Frieder Bruhn  
Hochschule Reutlingen

Stephan\_Frieder.Bruhn@Student.Reutlingen-University.de

## Abstract

Die Pflege stellt unter verschiedenen Aspekten ein umfangreiches Gebiet der Gesellschaft dar, in dem soziale Interaktion und zeitlicher Druck zunehmend konkurrieren. Neben reinen Pflege Tätigkeiten wird die Dokumentation von erbrachten Leistungen trotz aller Bestrebungen, schlankere Methoden zu etablieren, immer zentraler und zeitintensiver. Hier könnte ein System der Computer-Vision die Pflegekräfte unterstützen, indem erbrachte Leistungen automatisch erkannt und dokumentiert werden. Die Voraussetzung dafür sind Recherchen über den Status Quo in der Pflege sowie über entsprechende Technologie, aus denen Anforderungen generiert werden können. Diese Recherchen wurden in Form einer Kombination aus Literaturrecherche und Interviews durchgeführt, deren Ergebnisse in dieser Arbeit zusammenfassend vorgestellt werden. Basierend auf den Resultaten der Recherchen wurden

Anforderungen ermittelt, die anschließend dargestellt werden. Dabei handelt es sich sowohl um funktionale als auch um nicht-funktionale Anforderungen an das Gesamtsystem. Außerdem wurden Aspekte ermittelt, die einer Einführung des Zielsystems aktuell noch im Wege stehen könnten - insbesondere im Bereich Datenschutz und Akzeptanz der Patienten. Die Ergebnisse dieser Arbeit bieten damit eine informative Grundlage als Unterstützung bei der Umsetzung eines solchen Systems.

## CCS Concepts

•Computing methodologies~Artificial intelligence~Computer vision~Image and video acquisition •Applied computing~Life and medical sciences •Machine learning~Machine learning approaches

## Keywords

Computer Vision, Action Detection, Requirements, Health Care, First Person Action Recognition, Egocentric View, Healthcare Documentation

## 1 Einleitung

Zu Beginn dieser Ausarbeitung wird dargestellt, aus welcher Problemstellung heraus die Erarbeitung der Thematik motiviert ist und wie vorgegangen wird, um eine entsprechende Lösung zu erarbeiten.

---

Betreuer / -in Hochschule Prof. Dr.-Ing. Cristóbal Curio  
Hochschule Reutlingen  
Cristobal.Curio@Reutlingen-University.de

Betreuer / -in Firma Dr. Michael Gebhart  
iteratec GmbH, Stuttgart  
Michael.Gebhart@iteratec.com

Informatics Inside Herbst 2021  
24. November 2021, Hochschule Reutlingen  
Copyright 2021 Stephan Frieder Bruhn



## 1.1 Problemstellung

In der Pflege, auch der mobilen Pflege, herrscht seit langem Fachkraftmangel vor [1]. Die Pflegekräfte sehen sich somit einem enormen Zeitdruck ausgesetzt [2], wobei rotierende Schichtpläne und Aushilfskräfte auch nicht dazu beitragen, ausreichend Zeit bei den Hausbesuchen zu schaffen. Da die Anzahl alter und pflegebedürftiger Menschen zunehmend steigt [3] und deren Pflege ein unerlässlicher und zwingender Teil der Gesellschaft ist, muss Pflegekräften vor Ort das Arbeiten erleichtert werden.

Pflegekräfte sind dazu verpflichtet, erbrachte Leistungen zu dokumentieren [4], was zu weiterem Aufwand führt und eine Fehlerquelle darstellt, gerade weil die Dokumentation teils im Nachgang erfolgt [5]. Hier soll ein System helfen, das die Leistungen automatisch erkennt und dokumentiert. Ziel ist, dass die Pflegekraft entlastet wird, was Stress reduziert und somit mehr Zeit für die zu pflegende Person aufgebracht werden kann.

Um dies umsetzen zu können, bedarf es vorangehender Recherchen, die einen Überblick verschaffen und eine Basis für Anforderungen an ein entsprechendes System bilden, die nachfolgend ausgearbeitet werden. Diese Arbeit beschäftigt sich mit folgender Frage:

*„Welche Anforderungen ergeben sich an ein System der Computer-Vision zur automatischen Dokumentation von Pflegeleistungen?“*

## 1.2 Vorgehen

Zur Erarbeitung der Fragestellung wird eingangs der State of the Art betrachtet, wobei neben der Aktionserkennung auch auf die Dokumentation in der Pflege und auf Aufzeichnungsmodalitäten aus der Perspektive der ersten Person eingegangen wird. Anschließend werden die Anforderungen vor dem Kontext der ambulanten Pflege und die sich dadurch zusätzlich ergebenden Herausforderungen erarbeitet. Dazu werden Er-

kenntnisse aus der Literatur mit den Ergebnissen der Interviews zusammengeführt, wodurch die Thematik aus unterschiedlichen Perspektiven beleuchtet wird. Dadurch können praktische Anforderungen vor technologischem Hintergrund erhoben werden. Dabei wurde insbesondere auf die Bereiche Funktional, Qualität, Usability, Datenschutz und Trainingsdaten eingegangen. Bevor die Arbeit mit einem Fazit abgeschlossen wird, erfolgt eine Diskussion. Dabei wird festgehalten, dass für einen realen Einsatz eines solchen Systems noch einige Hürden zu überwinden sind, wie beispielsweise das Fehlen von Trainingsdaten, rechtliche Aspekte oder die Wahrung der Würde der Patienten.

## 2 State of the Art

Für das System sind Einblicke in verschiedene Bereiche notwendig. So muss sowohl die Pflegedomäne mit den aktuellen Prozessen verstanden wie auch ein Überblick über die Technologie gewonnen werden. Dabei ist sowohl die Verarbeitung der Daten relevant wie auch die Aufnahmemodalitäten. Bei der Aufnahme wird sich auf interaktionsfreie Modalitäten beschränkt, da beispielsweise eine Erfassung per Spracheingabe weder vor Vergessen schützt noch die Dokumentation automatisch durchgeführt wird.

### 2.1 Pflegedokumentation

Will man in einen Prozess eingreifen und dabei einen Teilprozess technologisch ersetzen, ist es unabdingbar, sich vorab mit dem Bestandsprozess und dem Bedarf hinter diesem Prozess auseinanderzusetzen. Nur so kann der neue Teilprozess so ausgestaltet werden, dass er sich gut in das prozedurale Umfeld einfügt. Daher wurden neben einer Recherche auch Interviews mit Pflegekräften bzw. mit Angestellten aus einschlägigen Einrichtungen durchgeführt [16]. Obgleich die Pflegedokumentation teilweise als überflüssige Bürokratie angesehen wird, sind Pflegeeinrichtungen dazu verpflichtet, eine Dokumentation über alle für die Pflege relevanten Informationen durchzuführen [17]. Dazu gehören u. a. die Pflegeanamnese, die Pflegeplanung, der Pflegebericht, Angaben

über den Einsatz von Pflegehilfsmitteln sowie über durchgeführte Pflegeleistungen [17]. In dieser Arbeit liegt der Fokus dabei auf letzterem, nämlich der Dokumentation von Pflegedienstleistungen. Diese Leistungen werden in zwei Kategorien eingeteilt: die Grundpflege und die Behandlungspflege. Während gewöhnliche und wiederkehrende Tätigkeiten zur Verrichtung des Alltags (z. B. waschen, essen, umkleiden) in die Grundpflege fallen, zählen ärztlich verordnete Leistungen (z. B. Wundversorgung, Verbandswechsel, Medikamentenvergabe) zu der Behandlungspflege [18]. Manchmal wird zusätzlich von hauswirtschaftlicher Versorgung gesprochen, wobei sich die Leistungen der Grundpflege in Leistungen am Patienten und Leistungen im Haushalt unterteilen [19]. Im Rahmen dieser Arbeit wird jene weitere Unterteilung nicht vorgenommen und von nur zwei Kategorien ausgegangen.

Während Pflegekräfte Tätigkeiten der Grundpflege nach eigenem Ermessen durchführen, sind Leistungen der Behandlungspflege also vom Arzt vorgegeben [16]. Der zuständige Arzt stellt wöchentliche Listen zusammen, in denen diese medizinischen Leistungen aufgeführt werden [16]. Diese Listen sind inzwischen zumeist digital, jedoch findet auch die Papierform weiter Anwendung [17]. Sie dienen als Checkliste, auf denen die erbrachten Leistungen direkt abgehakt werden können [16].

Unabhängig davon, ob diese Listen der Pflegekraft in digitaler oder analoger Form vorliegen, bedarf es zusätzlichen Aufwands, diese zu führen, womit problematische Aspekte einhergehen [16]. Zum einen ist es Zeitaufwand, der von der Pflegezeit am Patienten abgeht. Zum anderen liegt dadurch nicht die volle Aufmerksamkeit auf der zu pflegenden Person, was dazu führen kann, dass sie sich weniger in einer sozialen Interaktion, sondern mehr als behandeltes Objekt sehen kann [16]. Dies ist auch der Tatsache geschuldet, dass die ambulante Pflege ein Umfeld ist, wo Arbeiten mit und an einem Menschen und entsprechende Interaktionen,

auch sozialer Natur, im Mittelpunkt stehen. Pflegebedürftige Menschen wünschen sich einen persönlichen Kontakt. Dies ergab u. a. eine Studie aus dem Jahr 2017, aus der hervorgeht, dass sich 59% der Gepflegten mehr Zeit für Gespräche wünschten [6]. Der Versuch, der pflegebedürftigen Person dieses Gefühl durch eine nachgelagerte Dokumentation zu nehmen, mündet jedoch häufig in Vergessen der Dokumentation einzelner Leistungen, was dazu führt, dass der Dokumentationspflicht nicht vollständig nachgekommen wird, wodurch die Pflegequalität gefährdet wird [17]. Selbst bei unkritischen Leistungen der Grundpflege können die Pflegeprozesse dabei nicht effektiv durchgeführt werden. Des Weiteren wurde erwähnt, dass, insbesondere im Bereich der Grundpflege, auch Sprachbarrieren die Dokumentation erschweren [16].

Seit Anfang 2017 läuft über *EinSTEP* eine 2013 vom Bundesministerium für Gesundheit (kurz BMG) beauftragte Initiative [31, 32], das Strukturmodell. Es beinhaltet eine sogenannte Strukturierte Informationssammlung (SIS®), worüber Dokumentationsprozesse vereinfacht werden sollen [31]. Dieses Modell gibt einen ergänzenden Rahmen für vereinfachte Dokumentation und wird bislang von etwa 80% der ambulanten und stationären Pflegeeinrichtungen in Deutschland umgesetzt [31]. Eine zentrale Einführung durch dieses Modell ist es, bei der Grundpflege nur Abweichungen vom Pflegeplan zu dokumentieren [32]. Für die Bestimmung der Abweichungen muss dabei weiterhin ein Überblick über den Pflegeplan sowie die durchgeführten Leistungen gegeben sein. Auch scheinen Verschlinkungen in den Prozessen durch weitere Auflagen und Anforderungen an die Dokumentation mehr als kompensiert zu werden, sodass auch nach Einführung und breiter Annahme des SIS® über zu viel Dokumentationsaufwand geklagt wird [36].

Da es sich auch bei der Dokumentation von Pflegeleistungen um personenbezogene Daten handelt, ist auch die Einsicht in die Dokumentation ein zu beachtendes Thema.

Grundsätzlich gilt, dass ohne (rechtlichen) Grund kein Anspruch auf Einsicht besteht [17]. Da die Einsicht in die Dokumentation für beteiligte Pflegekräfte erforderlich ist, um die vorgeschriebene Tätigkeit durchzuführen, ist hier ein triftiger Grund gegeben. Umstritten ist jedoch, inwieweit dies für Verwaltungsangestellte und Pflegekräfte anderer Teams gilt, da hier eine Einsicht der Dokumentation nicht für die Aufgabenverrichtung erforderlich ist [17]. Während bei der digitalen Daten-Verarbeitung derartige Zugriffe leicht über Berechtigungen gesteuert werden können, ist die Zugriffssteuerung bei dem nach wie vor realen Einsatz von analogen Dokumenten [16] schwieriger durchzuführen, da sie in diesem Falle physisch erfolgen muss. Unabhängig davon, ob digitale oder analoge Ansätze verfolgt werden, gilt der Grundsatz der Eigenhändigkeit, was bedeutet, dass die pflegerische Dokumentation nicht delegiert werden darf [30].

Abweichungen von gegebenen Zeitplänen sind in der Regel bewusst getroffen und strategisch im Patientenwohl begründet [16]. Die Umsetzung hierbei ist händisch leichter zu realisieren. Wie sich dies mit einer automatischen Dokumentation vereinbaren ließe, gilt es zu prüfen.

## 2.2 Erste-Person-Aufzeichnung

In der Grammatik gelten es als Definition für die Bestimmung sprachlicher Perspektiven die erste, die zweite und die dritte Person. Dionysios Thrax definierte diese Personen dadurch, dass die erste Person diejenige ist, von der die Rede ausgeht, die Rede an die zweite Person gerichtet ist und die dritte Person das Objekt der Rede ist [7]. Diese Perspektiven werden parallel auf die Computer-Vision übertragen, wobei es hier nicht um Reden geht, sondern um abbildende Repräsentationen der Wirklichkeit. Dadurch lässt sich die zweite Person als Empfänger der Rede nicht sinnvoll übertragen, die erste und die dritte Person allerdings doch.

In der Computer-Vision ist die erste Person diejenige, die als fokussierter Akteur im jeweiligen Kontext auftritt, die dritte Person ist eine beliebige beobachtende Person. In Bezug auf die Perspektiven ist die Perspektive der ersten Person die Wahrnehmung aus Sicht des Akteurs, die Perspektive der dritten Person beschreibt die Sicht von außen auf den fokussierten Akteur. Im Kontext dieses Projekts wurde die Aufzeichnung aus der Perspektive der ersten Person gewählt (Begründung siehe Kapitel 3, Anforderungen). Für die Aufzeichnung der relevanten Daten für die Aktionserkennung aus der Perspektive der ersten Person gibt es verschiedene Möglichkeiten, die kurz beleuchtet werden.

Eine häufig genutzte Variante zur Aufzeichnung stellt die Aufnahme mit einer RGB-Kamera dar. In den meisten Fällen wird sie dafür am Kopf befestigt [9], um durch Augennähe die Immersion der Perspektive der ersten Person zu verstärken [8]. Ein weiterer Grund für diese Position der Kamera ist, dass das Sichtfeld der Kamera dem tatsächlichen Sichtfeld des Akteurs sehr nahekommt, ungeachtet der Augenbewegung, und es somit wahrscheinlich ist, dass die Kamera durch die Kopfbewegungen größtenteils auf das wesentliche Geschehen ausgerichtet wird. Dieser Vorteil kommt jedoch auch mit neuen Herausforderungen einher. So ist die Kamera hierbei häufigen ruckartigen Bewegungen ausgesetzt, die sich im aufgezzeichneten Bild niederschlagen und bei der weiteren Bearbeitung berücksichtigt werden müssen [9]. Des Weiteren kann es bei Routinearbeiten zu Problemen führen. So zum Beispiel, wenn eine Aufgabe aufgrund geringer Komplexität oder sehr hoher Wiederholungsrate ohne das Hinschauen des Akteurs ausgeführt wird. In diesem Fall entziehen sich die Aktion oder relevante Objekte möglicher Weise dem Sichtfeld der Kamera.

Neben oder anstatt der Aufzeichnung von Videodaten können für die Aktionserkennung auch Tiefeninformationen erfasst werden [13], die anschließend beispielsweise in Deep Motion Maps übertragen werden kön-

nen [15]. Probleme bezüglich Kamerabewegung oder Sichtfeld lassen sich dadurch nicht umgehen, jedoch kann die Erfassung von zusätzlichen Tiefeninformationen für die Verarbeitung hilfreich sein [15], wie etwa zum Segmentieren relevanter Informationen.

Eine weitere, weniger verbreitete Möglichkeit zur Positionierung der Kamera ist es, die Kamera am Handgelenk anzubringen. Da sich Aktionen insbesondere durch Objekte und der Interaktion mit diesen beschreiben lassen, bringt das Anbringen der Kamera am Handgelenk diesbezüglich einen wesentlichen Vorteil: Die aktionsgegenständlichen Objekte sind größer abgebildet und das Sichtfeld der Kamera ist stets an der Hand und die durch sie manipulierten Objekte ausgerichtet – ohne viele harte Bewegungen zwischen Kamera und Beobachtungsbereich.

Diese Variante wurde in einer Arbeit von Onishi et al. in Hinblick auf die Aktionserkennung näher beleuchtet [10]. Dabei im Fokus stand u. a. ein Vergleich zu der Variante der am Kopf befestigten Kamera. Die dabei durchgeführten Tests, die im Kontext von Aktivitäten des täglichen Lebens durchgeführt wurden, ergaben dabei, dass sich vermutete Vorteile bestätigten und im Allgemeinen bessere Ergebnisse erzielt werden konnten. In manchen Aktionsklassen wurden jedoch schlechtere Ergebnisse erzielt, wodurch auch die Eignung dieser Variante kontextabhängig ist. Ein weiteres Problem ist, dass eine Hand nicht immer an allen Aktionen beteiligt ist, auch unabhängig von der Händigkeit des Akteurs. Dadurch kann eine Aktion, die mit der weniger bevorzugten Hand durchgeführt wurde, vollständig außerhalb des so aufgezeichneten Sichtfelds stattfinden.

Des Weiteren kann die Aufzeichnung von Aktivitäten in Form von Bewegungsdaten erfolgen, die anhand einer inertialen Messeinheit (kurz IMU, Inertial Measurement Unit) erfasst werden. Diese besteht aus einer Kombination von einem Bewegungssensor

und einem Gyroskop, mit deren Hilfe die Beschleunigung und die Winkelgeschwindigkeit gemessen werden [11]. Beide Sensoren haben typischer Weise eine 3-Achsen-Erfassung. Diese Sensoren können um einen Magnetfeldsensor ergänzt werden. In der Arbeit von Cristiano et al. [12] wurde eine Aufzeichnung von Bewegungsdaten zur Aktionserkennung durchgeführt, wobei eine IMU ohne Magnetfeldsensor verwendet wurde. Diese Messeinheit war dabei ebenfalls am Kopf befestigt. Die Auswertung der Bewegungsinformationen bei den durchgeführten Tests ergab signifikante Rückschlüsse auf die zugrundeliegenden Aktionen. Die Genauigkeit nahm jedoch rasant ab, wenn weniger bewegungsintensive Aktionen durchgeführt wurden [12]. Die Aufzeichnung und Auswertung von Bewegungsinformationen eignet sich daher primär für Aktivitäten, bei denen Bewegungen des Körpers des Akteurs eine zentrale Rolle spielen.

### 2.3 Aktionserkennung

Durch die rapide Entwicklung und Integration von Hard- und Software in Freizeit und Beruf besteht ein wachsender Bedarf darin, Interaktionshürden zwischen Menschen und Maschinen abzubauen. Ein großes und aktuelles Forschungsfeld ist dabei die Computer-Vision, da sie wesentlich dazu beiträgt, dass Maschinen ihr reales Umfeld wahrnehmen und einschätzen können. Ein Teil dieses Bereichs stellt dabei die menschliche Aktionserkennung dar [14], wo das korrekte Interpretieren der komplexen menschlichen Bewegungen in ihrer großen Unterschiedlichkeit eine zentrale Herausforderung darstellt. Ein Faktor für diese Vielfaltigkeit ist die menschliche Individualität, da Aktionen von unterschiedlichen Menschen unterschiedlich ausgeführt werden können, oder gar derselbe Mensch eine Wiederholung derselben Aktion wiederum etwas anders durchführen könnte. Im Bereich der ambulanten Pflege kommt darüber hinaus hinzu, dass sehr viele der Aktionen in irgendeiner Weise auf Interaktionen mit einem weiteren Menschen basieren, wodurch die Individualität in der

Ausführung noch größer ist als beispielsweise bei herkömmlichen Aktionen des täglichen Lebens (kurz ADL, Actions of Daily Living). Auch handelt es sich insbesondere bei Leistungen der Behandlungspflege überwiegend um komplexe Bewegungen und Aktionsabläufe, die sich untereinander teilweise nur wenig unterscheiden.

Somit stellt die Aktionserkennung ein vielfältiges und nach wie vor herausforderndes Themengebiet dar, in dem viele unterschiedliche Ansätze und Methoden entwickelt und erprobt werden. Im Allgemeinen wird dabei hauptsächlich zwischen zwei Bereichen unterschieden, nämlich zwischen der Klassifizierung und der Detektion [20]. Bei der Aktions-Klassifizierung geht es darum, ein Bild oder eine Bildsequenz anhand enthaltener Informationen zu klassifizieren. Diese Daten sind dabei in der Regel kurze Videoabschnitte, die jeweils eine Aktion enthalten, die bestimmt werden soll [20]. Je nach Anwendungsfall und Aktion können auch einzelne Bilder klassifiziert werden. Während also bei der Klassifizierung einer definierten Sequenz die enthaltene Aktion zugeordnet werden soll, ist es das Ziel der Aktions-Detektion, in einem unbestimmten Sequenzfenster bekannte Aktionen zu erkennen und zu lokalisieren, wozu meist das sogenannte Sliding Window Verfahren eingesetzt wird, was jedoch sehr rechenintensiv ist [20].

Da es sehr viele verschiedene Einsatzbereiche und Aufzeichnungsvarianten gibt, bedarf es auch vielerlei Methoden, um den verschiedenen Anforderungen gerecht zu werden. Während bei einer statischen Kamera beispielsweise eine Subtraktion der Bilder vorgenommen werden kann, um Bewegungen zu detektieren [20], funktioniert dies nicht bei einer bewegten Kamera. Auch die Perspektive spielt bei der Wahl einer geeigneten Methode eine wesentliche Rolle. Während aus der Perspektive der dritten Person die Posen des Akteurs analysiert werden können [21], sind diese aus der Perspektive der ersten Person in der Regel nicht im Sichtfeld enthalten, was die Aktionserkennung komplexer gestaltet [9] und einige Verfahren

ausschließt. Im Folgenden werden daher nur Methoden betrachtet, die für Aufzeichnungen aus der Perspektive der ersten Person geeignet sind.

Zentral für die Aktionserkennung ist es, Merkmale aus den Bilddaten zu generieren [20], wobei unter den verschiedenen Ansätzen unterschiedliche Merkmale als Grundlage generiert werden. Methoden, die allein auf personenbezogene Trajektorien [22] oder Posen des Akteurs basieren, schneiden bei der Perspektive der ersten Person nicht gut ab [9], da zumeist lediglich die Hände des Akteurs zu sehen sind. Jedoch bieten Trajektorien eine gute Basis, um Rückschlüsse auf die Bewegungen des Akteurs sowie auf interessante Bildausschnitte zu liefern [29]. Dieser Ansatz, Merkmale aus der räumlich-zeitlichen Verteilung von Punkten und daraus abgeleiteten Interessen-Hotspots zu gewinnen [28], ist sehr verbreitet. Trotz des Nachteils, dass Kamerabewegungen zu einem erhöhten Hintergrundrauschen führen, sind diese Merkmale dennoch für viele Bereiche der Aktionserkennung maßgeblich [10, 23].

In Ergänzung dazu spielen neben der Hand-Pose die Objekte der Interaktion eine wesentliche Rolle [9], wodurch die Objekt-Detektion [24] Teil der meisten Verfahren ist. Dabei gibt es unterschiedliche Ansätze, Objekte in die Analyse mit einzubeziehen. Ma et al. haben in ihrer Arbeit [25] eine Twin-Stream-Netzwerkarchitektur unter Einsatz von Convolutional Neural Networks (kurz CNN) vorgestellt, wobei ein Stream das manipulierte Objekt detektiert, während der zweite Stream die Aktion analysiert. Beide Streams werden schlussendlich wieder zusammengeführt und aus Objekt- und Aktionsinformationen die durchgeführte Tätigkeit aggregiert [25]. Auch Three-Stream-Architekturen finden Einsatz, wobei in der Arbeit von Singh et al. zusätzlich Streams mit Daten aus der Perspektive der dritten Person verwendet wurden [9]. Andere Methoden basieren darauf, dass Objekte anders aussehen, wenn sie aktiv, also Gegenstand der durch-

geführten Interaktion sind [23]. Dabei werden unterschiedliche Objekt-Detektoren trainiert, um zwischen aktiven und inaktiven Objekten zu unterscheiden. Diese Informationen werden um die Relation zu den Händen angereichert, um eine optimierte Aktionserkennung durchführen zu können [23, 26].

Die meisten dieser Ansätze zielen darauf ab, geschlossene Aktionen zu erkennen. Je länger sich Aktionen zeitlich erstrecken und je komplexer solch Tätigkeiten sind, desto komplizierter wird diese Aufgabe. Daher verfolgen Giannakeris et al. einen Ansatz, Aktivitäten in Mikro-Aktionen zu unterteilen, die charakteristisch für die übergeordnete Aktivität sind [27]. Zur Identifizierung der Mikro-Aktionen wird die Bewegung der Objekte betrachtet. Dabei können verschiedene Mikro-Aktionen entweder eine Abfolge darstellen, oder sich innerhalb der Aktivität fortlaufend wiederholen [27].

Alle diese Methoden basieren darauf, dass sie, oder entsprechende Teilsysteme, anhand von Trainingsdaten trainiert werden. Die Trainingsdaten müssen dabei späteren Produktivdaten so ähnlich sein, dass das System beim Trainieren lernen kann, relevante Merkmale in den Produktivdaten wiederzuerkennen und korrekt einzuordnen. Die Verfügbarkeit geeigneter Daten aus dem Bereich der Pflege und der Perspektiver der ersten Person ist aktuell nicht gegeben. So existieren zwar Videodatensätze aus der Perspektive der ersten Person, wie beispielsweise GTEA [33], UCF101 [34], oder auch ein Begleitdatensatz zu einem Paper [35]. Jedoch sind hier nur Aktionen des täglichen Lebens abgebildet, während der Pflegesektor nicht abgedeckt wird.

### 3 Anforderungen

Basierend auf den Erkenntnissen zu dem jeweiligen Status Quo der Pflegedokumentation, der Aufzeichnungsmodalitäten und der Aktionserkennung können konkrete Anforderungen erhoben werden, die grundsätzlich an ein solches System gestellt werden müssen. Dabei wird auf die Aspekte Funktional,

Qualität, Usability, Datenschutz und Trainingsdaten eingegangen, die gleichzeitig als Kategorien der Anforderungen verwendet werden. Die dadurch erhobenen Anforderungen stellen ein Ausgangsset dar und können in dieser allgemeinen Form nicht sinnvoll verfeinert werden. Weitere Spezifikationen, wie die der Anzeigeelemente oder Schwellenwerte, müssen in konkreten Umsetzungen für die jeweilige Applikation und das Einsatzumfeld definiert werden.

#### 3.1 Funktional

Das System muss pflegerische Leistungen erkennen können, die im Voraus definiert wurden (AF-1). Dies ist das wesentliche Feature des Systems und muss entsprechende Priorität haben.

Um die Aktionserkennung durchführen zu können, ist es erforderlich, dass das System alle von der jeweiligen Pflegeeinrichtung regelmäßig durchgeführten Leistungen kennt (AF-2), da nur bekannte Aktionen dokumentiert werden können. Dafür muss ein Leistungskatalog mit relevanten Daten erhoben und dem System hinterlegt werden.

Das System muss zudem hinterlegt haben, aus welchen Teilschritten die jeweiligen Leistungen bestehen (AF-3). Die begleitende soziale Interaktion während der Pflege, häufig in Form von Kommunikation und damit verbundenem Blickkontakt, sorgt dafür, dass die Videodaten keine kontinuierlichen Repräsentationen der durchgeführten Aktionen liefern, sondern lediglich einzelne Ausschnitte. Je nach Tätigkeit und Routine der behandelnden Pflegekraft, können dabei auch Teilschritte durchgeführt werden, die gar nicht im Sichtfeld der Kamera erscheinen. Des Weiteren kommt es vor, dass einige Aktionen nicht kontinuierlich durchgeführt werden und sich dadurch überlappen. So können Teilschritte erkannt werden, die nicht in Verbindung mit einer zuvor durchgeführten Teilaktion steht, was in folgendem einfachen Beispiel verdeutlicht wird.

*Beispiel:* Während des Besuchs bei einer zu pflegenden Person entfernt die Pflegekraft einen Verband. Als nächstes misst sie der

Person Fieber. Anschließend wird ein Verband am selben Arm angelegt, wo zuvor der Verband abgenommen wurde. Die drei einzelnen erkannten Aktionen (Verband abnehmen, Fieber messen, Verband anlegen) resultieren in zwei Leistungen, nämlich Verbandswechsel und Fieber messen. Die jeweiligen Teilaktionen können dabei noch granularer sein, als es in diesem Beispiel der Fall ist. Das System muss also trotz Unterbrechungen in Aktionsabläufen oder Abweichungen in der Durchführungsdauer der Aktionen eine sichere Erkennung gewährleisten können.

Trotz einer automatisierten Leistungserkennung handelt es sich insbesondere im Bereich der Behandlungspflege weiterhin um Listen, die abzuarbeiten sind. Daher braucht das System Kenntnis über den jeweiligen Pflegeplan (AF-4), um entsprechende Einträge abhaken zu können (AF-5). Dies erleichtert zudem das Klassifizierungsproblem erkannter Aktionen. So müssen in den Dokumenten mehr Details notiert werden, als die automatische Erfassung erstmal hergibt. Beispielsweise kann das System in diesem Stadium nicht erkennen, welches Medikament in welcher Dosierung gespritzt wurde. Sehr wohl erkennen muss es in diesem Beispiel aber, dass eine Spritze verabreicht wurde. Gibt der Pflegeplan vor, dass ein Medikament  $x$  in der Dosierung  $y$  verabreicht und der rechte Verband gewechselt werden soll, reicht es zum Abhaken aus, zu erkennen, dass eine Spritze verabreicht und ein Verband gewechselt wurde. Die Kenntnis über den Pflegeplan bietet somit nicht nur die Basis zum Abhaken, sondern ermöglicht zudem weitere Details.

Um den Überblick zu behalten, was bzw. ob etwas vom System erkannt wurde, soll das System der Pflegekraft ein Feedback über erkannte Aktionen geben (AF-6). Zusätzlich sollen erbrachte Leistungen in der Liste der zu erledigenden Pflegeleistungen entsprechend markiert sein.

Das System muss auch erkennen können, wie oft eine Aktion ausgeführt wurde (AF-

7). So sollten bei einer beispielhaften intravenösen Verabreichung zweier Medikamente erkannt werden, dass zwei Injektionen erfolgten, um beide Leistungen als erledigt markieren zu können.

### 3.2 *Qualität*

Die Erkennungsrate muss sehr hoch sein (AQ-1), da der Mehrwert des Systems schnell in einen zusätzlichen Aufwand umschlägt, wenn das System nicht zuverlässig arbeitet und die Dokumentation im Nachgang korrigiert werden muss. Obgleich eine Erkennungsrate von 100% hier sehr wünschenswert wäre, wird eine derart hohe Schwelle in der Praxis kaum erreichbar sein. Für die Definition eines entsprechenden Akzeptanzkriteriums muss dafür im konkreten Umsetzungsprojekt ein Schwellenwert festgelegt werden, der in Relation zur geplanten Einsatzreife stehen sollte.

Die Leistungserkennung soll dabei in nahezu Echtzeit erfolgen (AQ-2), wobei eine maximale Latenz im konkreten Projekt zu definieren ist; ein nachträgliches Auswerten aufgezeichneter Videodaten scheidet aus. Dies begründet sich darin, dass das Feedback an die Pflegekraft wichtig ist, um einen Überblick über die Erkennung zu haben, da im Nachhinein wieder das Problem des Vergessens gegeben wäre. Zudem müssten Bilddaten als solche für die spätere Auswertung gespeichert werden, was nicht sein soll.

Um eine effiziente und genaue Erkennung zu gewährleisten, muss das System in der Erkennung robust gegen folgende Einflüsse sein (AQ-3). Die Erkennung darf nicht durch wechselnde Lichtverhältnisse verhindert werden, da das jeweilige Umfeld bei der ambulanten Pflege stark variieren kann. Zudem muss es das System nichtdiskriminierend sein. Die Erkennung der Aktionen darf nicht ethnisch abhängig sein. Auch muss das System auch mit häufigen harten Kamerabewegungen weiterhin valide Ergebnisse liefern. Gleiches gilt für zeitlich begrenzte Verdeckungen oder Teilschritte außerhalb des Sichtfeldes.

### 3.3 Usability

Ein zentrales Ziel des Systems ist es, der Pflegekraft die tägliche Arbeit zu erleichtern und den Fokus mehr auf die pflegebedürftige Person zu lenken. Damit dieses Ziel erreicht werden kann, muss sichergestellt sein, dass die Verwendung des Systems nicht aufwändiger ist als die herkömmlichen Verfahren (AU-1). Da die aktuelle Leistungsdokumentation einen überschaubaren Aufwand bedeutet, muss die automatische Leistungserfassung sehr beiläufig zu handhaben sein. Vorteilhaft in Hinblick auf diesen Aspekt ist, dass es sich bei diesem System nur um ein Teilsystem handelt. Das Gesamtsystem, worin die Leistungsdokumentation eingebettet wird, steigert diesen Mehrwert deutlich. Dennoch sollte auch jedes Feature für sich einen Mehrwertsbeitrag leisten.

Eine Grundvoraussetzung dafür ist, dass die Videodaten aus der Perspektive der ersten Person erfasst werden, da ein Mitführen und Aufbauen gesonderter Kameras zu viel Aufwand bedeuten würden, um diesen Nutzen zu rechtfertigen. Zudem muss die Kamera leicht anzubringen sein und darf die Pflegekraft nicht in ihrer Ausführung behindern (AU-2).

Da trotz aller Maßnahmen damit gerechnet werden muss, dass vereinzelt Leistungen falsch oder nicht erkannt werden, muss das System eine Möglichkeit bieten, nachträglich manuelle Korrekturen vornehmen zu können. Dies muss ohne komplizierte Menüs oder Abläufe im Sinne der Usability/User Experience möglich sein (AU-3).

### 3.4 Datenschutz

Weil es sich bei den erfassten Daten um Patientendaten handelt, müssen aktuelle datenschutzrechtliche Aspekte berücksichtigt werden. Das System selbst und alle eventuell beteiligten Teilsysteme müssen daher DSGVO-konform sein (AD-1).

Insbesondere dürfen nur relevante Daten gespeichert werden. Videodaten sind dabei nicht relevant, sondern ausschließlich die aus ihnen extrahierten Merkmale. Videodaten

dürfen ausschließlich innerhalb der Erkennungspipeline zwischengespeichert werden und müssen nach Extraktion der Features mit gelöscht werden (AD-2).

### 3.5 Trainingsdaten

Bei den zuvor betrachteten Datensätzen sind, durch die Positionierung der Kamera ebenfalls am Kopf der aktionsausführenden Person, zwar harte Bewegungen in der Kameraführung gegeben, mit denen die Verfahren zur Aktionserkennung umgehen können müssen, jedoch sind die durchgeführten Aktionen während der Videoaufnahmen dabei meist durchgehend auf dem Bild repräsentiert und es sind nur einzelne Bereiche von relevanten Objekten verdeckt. Zudem sind die Aktionen in sich geschlossen. Beide Punkte können in der Pflege wie dargestellt nicht als gegeben angenommen werden und müssen, neben weiterhin durchgängigen Aktionen (AT-1), zusätzlich in den Trainings- und Testdaten repräsentiert sein (AT-2).

Um die Robustheit des Systems zu fördern, müssen die Testdaten zudem unterschiedliche Beleuchtungsszenarien, Hautfarben und Umfelder wiedergeben (AT-3). Dies kann auch durch Data Augmentation gefördert werden.

Außerdem müssen die Testdaten in Qualität und Erscheinungsbild den Bilddaten des Zielgeräts gleichen (AT-4).

### 3.6 Zusammenfassung

Um einen kompakten Überblick über die erhobenen Anforderungen zu bieten, werden diese nachfolgend in Tabellenform zusammengefasst.

Kategorie	Kürzel	Inhalt
Funktional	AF-1	Alle relevanten Leistungen erkennen
	AF-2	Relevante Leistungen kennen
	AF-3	Teilschritte der Aktionen kennen
	AF-4	Aktuellen Pflegeplan der relevanten Patienten kennen



	AF-5	Erkannte Leistungen im Pflegeplan markieren
	AF-6	Feedback über Erkennungsstatus
	AF-7	Durchführungen zählen
Qualität	AQ-1	Sehr hohe Erkennungsrate
	AQ-2	Leistungserkennung in nahezu Echtzeit
	AQ-3	Robuste Erkennung gewährleisten ( <i>Licht, Ethnie, Umfeld, Kameraführung</i> )
Usability	AU-1	Benutzungsaufwand minimal halten
	AU-2	Aufnahme per leicht an der Pflegekraft anzubringender Kamera
	AU-3	Effizientes Editieren der Dokumentation
Datenschutz	AD-1	DSGVO-Konformität unterstützen
	AD-2	Keine Bilddaten speichern, ausschließlich Merkmale
Trainingsdaten	AT-1	Durchgängige Aktionen enthalten
	AT-2	Gesplittete Aktionen enthalten
	AT-3	Unterschiedliche Beleuchtungsszenarien, Hautfarben und Umfelder
	AT-4	In Qualität und Erscheinung den Daten des Zielgeräts gleichen

**Tabelle 1: Zusammenfassung der Anforderungen nach Kategorien**

## 4 Diskussion

Im Einsatz des Systems müssen Bilddaten nicht gespeichert werden, da die erfassten Daten in nahezu Echtzeit analysiert und Merkmale aus ihnen generiert werden können. Dabei müssen lediglich die extrahierten Merkmale für die weitere Verarbeitung bis zur Dokumentation gespeichert werden, während die Bilddaten bereits wieder verworfen werden können. Anders sieht es allerdings beim Trainieren des Modells aus, da dem System die Zuordnung von Videorepräsentationen zu relevanten Merkmalen beigebracht werden muss. Somit müssen für dieses System sämtliche relevanten Aktionen und Situationen in mehrfacher und bildlicher

Form aufgezeichnet und gespeichert werden. Während sich bei einem Verbandswechsel voraussichtlich einige Probanden für derartige Aufzeichnungen bereiterklären werden, gilt dies sicher nicht für jeden Bereich, wie beispielsweise bei der Intimpflege.

Problematisch könnte es in diesen Bereichen jedoch auch sein, dass sich pflegebedürftige Personen neben dem generellen Unbehagen als Pflegebedürftige noch unwohler dabei fühlen, wenn eine Kamera auf sie gerichtet ist. Gegen dieses subjektive Unwohlsein dürfte es auch kaum helfen, dass das Videomaterial nicht dauerhaft gespeichert und nur in Echtzeit ausgewertet wird. Muss die Brille, bzw. die Kamera, bei einzelnen Leistungen jedoch abgelegt werden, reduziert sich der Mehrwert schnell auf ein Minimum. Damit der Einsatz dieses Systems hilfreich sein kann, muss also gegeben sein, dass die pflegebedürftige Person mit einem durchgehenden Verwenden einverstanden ist. Hierzu muss vorab eine Einverständniserklärung eingeholt werden.

Außerdem müsste rechtssicher in Erfahrung gebracht werden, wie sich der Grundsatz der Eigenhändigkeit (Delegationsverbot) [30] bereits heute mit dem automatischen Abhaken durch ein System vereinbaren lässt. Jedoch sollte jedem Anwender auch bewusst gemacht werden, dass es sich um keine Delegation handelt, sondern um eine Unterstützung. Die Verantwortung für die korrekte Dokumentation liegt weiter bei der Pflegekraft. Bereits vor diesem Gesichtspunkt muss ein Editieren der erfassten Angaben möglich sein.

Dass aktuell keine Trainings- und Testdaten existieren, erschwert die Umsetzung. Zwar können auch eigene Daten aufgezeichnet werden, jedoch ist es mit hohem Aufwand verbunden, einen umfassenden Datensatz zu erstellen, sodass davon auszugehen ist, dass dabei eine geringe Variabilität der Daten sowie ein reduzierter Umfang in Kauf genommen werden muss. Dadurch kann das System nicht so trainiert werden, dass sichergestellt ist, dass eine sehr hohe Erkennungsrate in

wechselnden Umfeldern und Akteuren erzielt werden kann. Da die Datenerhebung wie beschrieben bereits rechtlich nicht trivial ist, werden derartige Aufnahmen, die als Trainings- und Testdaten dienen können, jedoch auch nicht ohne Grund aufgenommen werden. Hier könnte bereits ein eingeschränkter Prototyp, basierend auf einem unzulänglichen Datensatz, einen möglichen Bedarf aufzeigen und somit dafür sorgen, dass entsprechende Daten in der notwendigen Vielfältigkeit aufgezeichnet werden, und künftige Systeme auf breit aufgestellte Videodaten zurückgreifen können.

## 5 Fazit

In der Pflege besteht Unterstützungsbedarf, wo auch die Informatik ihren Beitrag leisten kann, wenn es darum geht, die Pflegekräfte in ihrem täglichen Geschäft zu unterstützen. Eine Möglichkeit dazu bietet eine Automatisierung des Dokumentationsprozesses. Dies lässt sich am einfachsten über einer mobilen, an der Pflegekraft befestigten, Kamera realisieren, hier in Form einer AR-Brille. Die integrierte Kamera stellt Daten zur Analyse zur Verfügung. Für die Auswertung dieser Bilddaten gibt es eine Vielzahl an Ansätzen und Methoden, die für eine konkrete Umsetzung tiefer auf ihre Eignung für diesen speziellen Kontext geprüft bzw. angepasst werden müssen. Der Kontext der ambulanten Pflege stellt einige Herausforderungen, aus denen sich weitere Anforderungen ergeben. Insbesondere die Patientenorientierung, sehr unterschiedliche Umgebungen, Datenschutz und die Art von Leistungen kommen dabei mit Herausforderungen einher.

Somit bietet dieses Forschungsgebiet ein spannendes Thema, das bei einer erfolgreichen Umsetzung den Alltag von Pflegekräften erleichtern kann. Bis zu einer umfänglichen Lösung bedarf es allerdings noch einiger Forschung und weiterer Daten.

Trotz aller Einschränkungen, mit denen ein Prototyp momentan voraussichtlich noch einherkommen würde, kann dennoch ein wesentlicher Beitrag erzielt werden, indem künftiger Innovation der Weg bereitet wird.

## Literaturverzeichnis

- [1] Süddeutsche Zeitung, 2021. Online verfügbar unter <https://www.sueddeutsche.de/bayern/bayern-pflege-personalmangel-wiedereinsteiger-holet-schek-1.5331871>; zuletzt aufgerufen am 30.09.2021.
- [2] DERWESTEN, 2021. Online verfügbar unter <https://www.derwesten.de/politik/zuviel-stress-zu-wenig-zeit-ambulante-pflegedienste-schlagen-alarm-id7840454.html>; zuletzt aufgerufen am 30.09.2021.
- [3] DESTATIS, Statistisches Bundesamt, 2020. Online verfügbar unter <https://www.destatis.de/DE/The-men/Gesellschaft-Umwelt/Gesundheit/Pflege/Publicationen/Downloads-Pflege/pflege-deutschlandergebnisse-5224001199004.pdf>; zuletzt aufgerufen am 30.09.2021.
- [4] Rolf Höfert, Thomas Meißner, 2008. Von Fall zu Fall — Ambulante Pflege im Recht. DOI: 10.1007/978-3-540-75599-9\_25.
- [5] Monika Habermann, Heidrun Biedermann, 2007. Die Pflegevisite als Instrument der Qualitätssicherung in der ambulanten Pflege. Online verfügbar unter [https://www.researchgate.net/publication/265642687\\_Die\\_Pflegevisite\\_als\\_Instrument\\_der\\_Qualitatssicherung\\_in\\_der\\_ambulanten\\_Pflege](https://www.researchgate.net/publication/265642687_Die_Pflegevisite_als_Instrument_der_Qualitatssicherung_in_der_ambulanten_Pflege); zuletzt aufgerufen am 30.09.2021.
- [6] Ann-Kathrin Seemann, Harald Fischer, 2017. Was macht Freude im Arbeitsalltag, und was belastet?. DOI: 10.1007/s16024-017-0308-9.
- [7] Wilfried Kürschner, 1996. Die Lehre des Grammatikers Dionysios. ISBN: 90-6831-881-0/2-87723-307-3, S. 177-215.
- [8] Diego Monteiro, Hai-Ning Liang, Andrew Abel, Nilufar Baghaei, Rita de Cassia Monteiro, 2018. Evaluating En-

- agement of Virtual Reality Games based on First and Third Person Perspective using EEG and Subjective Metrics.  
DOI:10.1109/AIVR.2018.00015.
- [9] Suriya Singh, Chetan Arora, C. V. Jawahar, 2016. First Person Action Recognition Using Deep Learned Descriptors.  
DOI: 10.1109/CVPR.2016.287.
- [10] Katsunori Ohnishi, Atsushi Kanehira, Asako Kanezaki, Tatsuya Harada, 2016. Recognizing Activities of Daily Living with a Wrist-Mounted Camera.  
DOI: 10.1109/CVPR.2016.338.
- [11] Norhafizan Ahmad, Raja Ariffin Raja Ghazilla, and Nazirah M. Khairi, 2013. Reviews on Various Inertial Measurement Unit (IMU) Sensor Applications.  
DOI: 10.12720/ijsp.1.2.256-262.
- [12] Alessia Cristiano, Alberto Sanna, Diana Trojaniello, 2019. Daily Physical Activity Classification using a Head-mounted device.  
DOI: 10.1109/ICE.2019.8792621.
- [13] Jing Liu, Lei Wang, Xubo Yang, 2013. Research for Human Action Recognition based on Depth Information.  
DOI: 10.1109/ICCSNT.2013.6967335.
- [14] Chongkai Yu, 2021. Action Recognition Algorithm based on 2D Human Pose Estimation Method. DOI: 10.23919/CCC52363.2021.9550204.
- [15] Chen Chen, Kui Liu, Nasser Kehtarnavaz, 2013. Real-time human action recognition based on depth motion maps.  
DOI: 10.1007/s11554-013-0370-1.
- [16] Michelle Schlicher, iteratec GmbH, Notizen von Interviews mit verschiedenen Pflegekräften und Klinikangestellten. Zusammenfassung online verfügbar unter:  
<https://explore.iteratec.com/blog/medical-care-glasses-mit-augmented-reality-die-zukunft-der-pflege-gestalten; zuletzt abgerufen am 14.10.2021.>
- [17] Guido Steinke, Bundesinteressenvertretung für alte und pflegebetroffene Menschen (BIVA) e.V., 2016. Einsichtnahme in die Pflegedokumentation. Online verfügbar unter <https://www.biva.de/dokumente/broschueren/Pflegedokumentation-rechtliche-Grundlagen.pdf; zuletzt abgerufen am 14.10.2021.>
- [18] Alexandra Jorzig, 2017. Die Abgrenzung von Grund- und Behandlungspflege unter haftungsrechtlichen Gesichtspunkten. Online verfügbar unter <https://www.gesundheitskongresse.de/berlin/2017/dokumente/presentationen/Jorzig-Alexandra---Haftungsrecht-Grund-und-Behandlungspflege.pdf?m=1485256048&; zuletzt abgerufen am 14.10.2021.>
- [19] pflege.de, 2016. Behandlungspflege: Medizinische Versorgung im Rahmen der Häuslichen Krankenpflege. Online verfügbar unter <https://www.pflege.de/altenpflege/behandlungspflege/; zuletzt abgerufen am 14.10.2021.>
- [20] Hong-Bo Zhang, Yi-Xiang Zhang, Bineng Zhong, Qing Lei, Lijie Yang, Ji-Xiang Du, Duan-Sheng Chen, 2019. A Comprehensive Survey of Vision-Based Human Action Recognition Methods. DOI: 10.3390/s19051005.
- [21] Jinchun Wen, Jiannong Chi, Chengdong Wu, Xiaosheng Yu, 2021. Human Pose Estimation Based Pre-training Model and Efficient High-Resolution. Representation. DOI: 10.23919/CCC52363.2021.9549849.
- [22] D. Meyer, H. Niemann, 1997. Automatische Bestimmung der Trajektorien ohne Marker zur Bewegungserkennung von Personen. Online verfügbar unter: [http://www5.informatik.uni-erlangen.de/~meyer/papers/97\\_automatische\\_bestimmung\\_der\\_trajektorien\\_ohne\\_marker\\_zur\\_bewegungserkennung\\_von\\_personen.pdf; zuletzt abgerufen am 14.10.2021.](http://www5.informatik.uni-erlangen.de/~meyer/papers/97_automatische_bestimmung_der_trajektorien_ohne_marker_zur_bewegungserkennung_von_personen.pdf; zuletzt abgerufen am 14.10.2021.)

langen.de/Forschung/Publikationen/1997/Meyer97-ABD.pdf; zuletzt aufgerufen am 18.10.2021.

- [23] Hamed Pirsiavash, Deva Ramanan, 2012. Detecting Activities of Daily Living in First-person Camera Views. DOI: 10.1109/CVPR.2012.6248010.
- [24] Pedro F. Felzenszwalb, Ross B. Girshick, David McAllester, Deva Ramanan, 2009. Object Detection with Discriminatively Trained Part-Based Models. DOI: 10.1109/TPAMI.2009.167.
- [25] Minghuang Ma, Haoqi Fan, Kris M. Kitani, 2016. Going Deeper into First-Person Activity Recognition. DOI: 10.1109/CVPR.2016.209.
- [26] Yang Zhou, Bingbing Ni, Richang Hong, Xiaokang Yang, Qi Tian, 2016. Cascaded Interactional Targeting Network for Egocentric Video Analysis. DOI: 10.1109/CVPR.2016.210.
- [27] Panagiotis Giannakeris, Panagiotis C. Petrantonakis, Konstantinos Avgerinakis, Stefanos Vrochidis, Ioannis Kompatsiaris, 2020. First-person activity recognition from micro-action representations using convolutional neural networks and object flow histograms. DOI: 10.1007/s11042-020-09902-6.
- [28] Ivan Laptev, Marcin Marszałek, Cordelia Schmid, Benjamin Rozenfeld, 2008. Learning realistic human actions from movies. DOI: 10.1109/CVPR.2008.4587756.
- [29] Heng Wang, Dan Oneata, Jakob Verbeek, Cordelia Schmid, 2016. A Robust and Efficient Video Representation for Action Recognition. DOI: 10.1007/s11263-015-0846-5.
- [30] Heinz R. Sträßner, 2010. Sicher dokumentieren - Rechtliche Aspekte der Pflegedokumentation. Online verfügbar unter: [https://www.thieme.de/statics/dokumente/thieme/final/de/dokumente/tw\\_pflege/le4\\_110\\_1-schutz.pdf](https://www.thieme.de/statics/dokumente/thieme/final/de/dokumente/tw_pflege/le4_110_1-schutz.pdf); zuletzt aufgerufen am 20.10.2021.
- [31] Bundesministerium für Gesundheit, Entbürokratisierung in der Pflegedokumentation, 2021. Online verfügbar unter: <https://www.bundesgesundheitsministerium.de/themen/pflege/entbuerokratisierung.html>; zuletzt aufgerufen am 20.10.2021.
- [32] EinSTEP, über EinSTEP. Online verfügbar unter: <https://www.ein-step.de/ueber-einstep/>; zuletzt aufgerufen am 20.10.2021.
- [33] Georgia Tech Egocentric Activity Datasets. Online verfügbar unter: <http://cbs.ic.gatech.edu/fpv/>; zuletzt aufgerufen am 27.10.2021.
- [34] UCF101 – Action Recognition Data Set. Online verfügbar unter: <https://www.crcv.ucf.edu/research/data-sets/ucf101/>; zuletzt aufgerufen am 27.10.2021.
- [35] Guillermo Garcia-Hernando, Shanxin Yuan, Seungryul Baek, Tae-Kyun Kim, 2018. First-Person Hand Action Benchmark with RGB-D Videos and 3D Hand Pose Annotations. Online verfügbar unter: [https://docs.google.com/forms/d/e/1FAIpQLScoksYrmthDbJeAV0\\_y-sXJDMfvZmzMMsX0\\_Uhkb6H6DHYGBtg/viewform](https://docs.google.com/forms/d/e/1FAIpQLScoksYrmthDbJeAV0_y-sXJDMfvZmzMMsX0_Uhkb6H6DHYGBtg/viewform); zuletzt aufgerufen am 27.10.2021.
- [36] ÄrzteZeitung, Springer Medizin, 2020. Genug Pfleger, aber zu viel Dokumentation. Online verfügbar unter: <https://www.aerztezeitung.de/Politik/Genug-Pfleger-aber-zu-viel-Dokumentation-405451.html>; zuletzt aufgerufen am 04.11.2021.



©2021 Stephan Frieder Bruhn. Lizenznehmer Hochschule Reutlingen, Deutschland. Dieser Artikel ist ein Open-Access-Artikel unter den Bedingungen und Konditionen der Creative Commons Attribution (CC BY)-Lizenz. <http://creativecommons.org/licenses/by/4.0/>

# Analyse des Krebsgenomik-Workflow-Management-Systems One Touch Pipeline

Merve Emirzeoglu  
Hochschule Reutlingen

Merve.Emirzeoglu@Student.Reutlingen-University.de

## Abstract

**[Kontext und Motivation]** Die Präzisionsmedizin strebt danach, genomische Daten zu nutzen, um passende Behandlungsmethoden für individuelle Patienten bereitzustellen. Mit der Technologie Next-Generation Sequencing (NGS) kann ein gesamtes menschliches Genom sequenziert werden. Vor allem im Bereich der Onkologie bietet es sich an, Genome von Tumorproben zu sequenzieren, um Mutationen zu detektieren. Das Wissen über die individuellen Mutationen eines Tumors lässt Aussagen über die Dysregulation von Signalkaskaden zu und kann über das Auffinden von sogenannten zielgerichteten Läsionen Aussagen zur individualisierten Therapie und Krankheitsmanagement zulassen. Zur Analyse der komplexen Daten, die durch NGS erzeugt werden, sind bioinformatische Pipelines erforderlich. In der Bioinformatik wird eine Pipeline erstellt, die ein Programm ausführt, dann transformiert, dann ein anderes Programm wieder ausführt. Das deutsche Krebsforschungszentrum (DKFZ) in Heidelberg verwendet ein System mit dem Namen One Touch Pipeline (OTP), das eine Automatisierungsplattform ist, die NGS-Daten verwaltet und entsprechende Pipelines zur Bearbeitung dieser Daten aufruft. **[Fragestellung und Problem]** Es besteht ein hohes Maß an Vielfalt an der Art und Weise, wie Vertreter der Molekulargenetik und der Pathologie solche Pipelines entwickeln und validieren können. Um diesen Prozess am DKFZ normkonform zu halten, möchte die Forschungseinrichtung ein

Qualitätsmanagementsystem einsetzen. Um dies zu ermöglichen, wird eine Analyse des Aufbaus des vorhandenen Systems OTP benötigt. **[Lösungsidee]** In dieser Arbeit wird das bestehende System OTP genauer betrachtet. Es wird der Datenfluss analysiert und mit Systems Modeling Toolbox (SYSMOD) modelliert. Das Rahmenwerk Task- and Object-oriented Requirements Engineering (TORE) wird für die Ermittlung der Aufgaben im Datenfluss eingesetzt. **[Beitrag]** Das Ziel dieser Arbeit ist die Dokumentation des Systems OTP mit dem Fokus auf die Systemschnittstellen, welche für die Auswahl des Qualitätsmanagementsystems und den Datenfluss benötigt werden.

## CCS Concepts

• **Software and its engineering** → **Abstraction, modeling and modularity**; *Data flow architectures*; Object oriented architectures;

## Keywords

Oncology, DNA- and RNA- Sequencing, Next-Generation Sequencing, One-Touch Pipeline, SYSMOD, TORE

## 1 Einleitung

Die Onkologie ist einer der am stärksten von der Genomik beeinflussten Bereiche der Medizin. Die rasanten Fortschritte in der Genomik und im medizinischen Verständnis haben nicht nur zu Entwicklungen in der Behandlung und im Management beigetragen, sondern auch unzählige Möglichkeiten für die Forschung eröffnet, die es ermöglichen Krebskrankungen besser zu verstehen [1].

Die Forschung auf dem Gebiet der Genomik hat sich in den vergangenen Jahren massiv weiterentwickelt, während gleichzeitig durch die Kostensenkung die NGS-Technologie in den Krankenhausalltag eingeführt werden konnte [2]. NGS wird in der molekularen Onkologie zur Detektion von somatischen Varianten eingesetzt. Durch die Zunahme der Genomik sowie der Einführung von NGS, haben sowohl der Umfang als auch die Breite der im Gesundheitswesen und in der medizinischen Forschung erzeugten Datensätze zugenommen. Die Zunahme der erzeugten Datensätze stellt eine Herausforderung für die Kapazitäten der biowissenschaftlichen Datenzentren dar, da diese die Daten korrekt und effizient verarbeiten und speichern müssen [3]. Daher werden große Datenspeicher und Rechenkapazitäten, weitreichende Verarbeitungsplattformen und ein automatisiertes Datenhandling benötigt, um den Arbeitsaufwand des Personals zu reduzieren. Bei einer Automatisierung können sich BioinformatikerInnen auf die Entwicklung der Methoden zur Analyse und die Interpretation konzentrieren, statt die Daten im Voraus zu verarbeiten. Des Weiteren können Fehler durch die Automatisierung vermieden bzw. reduziert werden [4]. Das DKFZ, als einer

der größten biomedizinischen Forschungszentren in Europa, verwendet die Automatisierungsplattform OTP<sup>1</sup>. Diese Automatisierungsplattform dient zur Bereitstellung strukturierter Daten, dem Datenzugriffmanagement sowie der Verarbeitung von NGS-Daten [2].

### 1.1 Ziel der Arbeit

Das Ziel dieser Arbeit ist die Durchführung einer Ist-Analyse des bestehenden Systems OTP bzgl. der Verarbeitung der Daten sowie die Modellierung der Systemschnittstellen, der enthaltenen Module und der durch das System ausgeführten Workflows. Um das System zu modellieren, werden die Ansätze SYSMOD [5] und TORE [6] betrachtet und verwendet. Der Ansatz SYSMOD wird für die Modellierung von OTP, der Schnittstellen und der Module verwendet, während der Ansatz von TORE für die Ermittlung der Aufgaben aus dem Datenfluss eingesetzt wird.

### 1.2 Forschungsfragen

Aus dem vorherigen Kapitel ergeben sich folgende Forschungsfragen, die in dieser Arbeit beantwortet werden sollen:

*RQ1: Wie sieht der Ablauf der Datenverarbeitung bei der DNA-Sequenzierung und der RNA-Sequenzierung aus?*

*RQ2: Welche Variantenaufrufe werden bei DNA-Sequenzierung eingesetzt und welche bei der RNA?*

*RQ3: Wie sehen die Schnittstellen und die enthaltenen Module des Systems OTP aus?*

### 1.3 Aufbau

Kapitel 2 befasst sich mit der Erklärung einiger Begriffe die in der Präzisionsonkologie relevant sind. In Kapitel 3 wird die Methodik die in dieser Arbeit verwendet wurde beschrieben. Kapitel 4 befasst sich mit dem Stand der Technik. Anschließend wird in Kapitel 5 die Analyse des Systems OTP anhand der definierten Vorgehensweise präsentiert.

---

Betreuer/-in Hochschule	Prof. Dr. Christian Kücherer Hochschule Reutlingen Christian.Kuecherer@Reutlingen-University.de
Betreuer/-in Extern	Dr. rer. nat. Dr. med. Daniel Hübschmann Deutsches Krebsforschungszentrum Heidelberg D.Huebschmann@dkfz-heidelberg.de

Informatics Inside Herbst 2021  
24. November 2021, Hochschule Reutlingen  
Copyright 2021 Merve Emirzeoglu

<sup>1</sup><https://otp.dkfz.de/otp/>

Abschließend wird in Kapitel 6 das Vorgehen sowie die Ergebnisse zusammengefasst.

## 2 Hintergrund

Dieser Abschnitt beschäftigt sich mit Begrifflichkeiten aus der Präzisionsonkologie, die für das Grundverständnis des Systems OTP relevant sind.

### 2.1 DNA-Sequenzierung

Die DNA (*engl. deoxyribonucleic acid*) ist das Erbgut des Menschen und vieler anderer Organismen. Jede Zelle im Körper eines Menschen hat nahezu die identische DNA. Ein Großteil der DNA befindet sich im Zellkern, wobei ein geringer Anteil auch in den Mitochondrien zu finden ist. Mitochondrien sind Zellstrukturen, welche die Energie aus der Nahrung in eine für die Zelle nutzbare Form umwandeln. In der DNA ist die Information in Form eines Codes gespeichert, welcher aus vier chemischen Basen besteht. Diese Basen sind Adenin (A), Guanin (G), Cytosin (C) und Thymin (T). Diese Basen der DNA paaren sich, um Einheiten zu bilden, die als Basenpaare bezeichnet werden [7]. Die DNA-Sequenzierung ist das Verfahren zur Ermittlung der Reihenfolge von Nukleotide in der DNA. Nukleotide sind Bausteine aus der Nukleinsäure und bildet den Grundbaustein für DNA und RNA. Diese Abfolge von Basensequenz der DNA liefert die nötigen Daten, die eine Zelle benötigt, um Protein- und RNA-Moleküle zusammensetzen. Anhand dieser Sequenz kann ebenfalls festgestellt werden, welche Art von genetische Information in einem bestimmten Abschnitt der DNA enthalten ist sowie Veränderungen in einem Gen aufzeigen, die zu Krankheiten führen können [8].

### 2.2 RNA-Sequenzierung

Die RNA (*engl. ribonucleic acid*) ist eine Nukleinsäure, genannt auch Makromolekül, deren Struktur zwar der DNA ähnelt, sich aber in wesentlichen Aspekten, wie dass die RNA einzelsträngig ist, unterscheidet. RNA wird

von der Zelle für unterschiedliche Aufgaben verwendet. Einer dieser Aufgaben ist die mRNA. Hierbei handelt es sich um das Informationsmolekül der Nukleinsäure, welches Informationen aus dem Genom durch die Übersetzung in Proteine überträgt [9]. In der RNA-Sequenzierung kann mit Hilfe von NGS die Menge und die Sequenzen der RNA in einer Probe untersucht werden. Dabei wird das Transkriptom analysiert, das angibt, welche der in unserer DNA kodierten Gene aktiviert oder deaktiviert sind [10].

### 2.3 Next-Generation-Sequencing

NGS ist eine Technologie zur Bestimmung der DNA- oder RNA-Sequenz, um genetische Variationen im Zusammenhang mit Krankheiten oder anderen biologischen Erscheinungen zu untersuchen. Die Technologie NGS wurde bei der Einführung als Massively-Parallel Sequencing bezeichnet, da sie die Sequenzierung mehrerer DNA-Stränge gleichzeitig ermöglicht, anstatt einen nach dem anderen wie bei der Sanger Sequenzierung. Der Einsatz dieser Sequenzierungstechnologie hat in der genomischen Forschung stark zugenommen, vor allem im Bereich der Onkologie. NGS ermöglicht die gleichzeitige Untersuchung von Hunderten bis Tausenden von Genen in mehreren Proben sowie die Entdeckung und Analyse verschiedener Arten von genomischen Merkmalen in einem Sequenzierungslauf [4, 11]. Um dies zu ermöglichen gibt es Plattformen mit diversen Verfahren zur Sequenzierung der DNA wie die Illumina Sequenzierung [8]. Das Grundprinzip von NGS basiert auf vier Schritten. Der erste Schritt ist die Probenextraktion, die mit Proben, bei der DNA oder RNA anfällt, durchgeführt werden kann. Nach der Extraktion werden Menge und Qualität der DNA oder RNA bestimmt, um ein hochwertiges Ausgangsmaterial für die Sequenzierung zu erhalten. Falls die Ausgangssequenz RNA ist, wird ein weiterer Schritt benötigt, bei dem die RNA

durch reverse Transkription in cDNA umgewandelt wird. Im zweiten Schritt wird aus der gewonnenen Probe eine Sequenzierungsbibliothek erstellt. Diese Bibliotheken werden in der Regel durch Fragmentierung der DNA wie die PCR-Amplifikation und Hinzufügen von Sequenzier-Adaptoren, welche an die NGS-Plattform angebunden werden erstellt. Um Ressourcen zu sparen, können mehrere Bibliotheken in einem Pool zusammengefasst und im selben Lauf sequenziert werden. Anschließend werden die Adapterbindungen jeder Bibliothek mit eindeutigen Indexsequenzen oder Barcodes versehen, um bei der Datenanalyse zwischen den Bibliotheken unterscheiden zu können. Während dem Sequenzierungsschritt, die den dritten Schritt bildet, werden die Bibliotheken auf einen Flowcell geladen und auf den Sequenzierer gelegt. Die Cluster von DNA-Fragmenten werden in einem Prozess vervielfältigt und zu Millionen von Kopien einzelsträngiger DNA geführt. Durch die Aufteilung in Cluster können mehrere Sequenzierungsvorgänge gleichzeitig durchgeführt werden. Nach Abschluss der Sequenzierung wird eine Software eingesetzt, um die große Menge an Daten auszuwerten. Der letzte Schritt ist somit die Ausrichtung und Datenanalyse. Hierbei werden die Reads, die bei der Sequenzierung entstanden sind, nach Qualität, Amplikongröße und Übereinstimmung zwischen den gepaarten Enden gefiltert. Die Reads werden dann zusammengesetzt und an Referenzgenom angeglichen. Anschließend können die Reads mit einer Referenzsequenz oder mit Reads aus einer weiteren Proben verglichen werden, um bspw. Krankheitsvarianten zu identifizieren [8, 12].

## 2.4 Whole-Genome-Sequencing

Whole-genome-sequencing (WGS) ist ein Methode zur Analyse ganzer Genome. Die Durchführbarkeit der Analyse wird durch die NGS-Technologie unterstützt und beinhaltet

die Sequenzierung der gesamten DNA sowie der in den Mitochondrien enthaltenen DNA [13].

## 2.5 Whole-Exome-Sequencing

Die Whole-Exome-Sequencing (WES) ist eine gängige NGS-Methode, bei der es um die Sequenzierung der proteinkodierenden Regionen des Genoms bzw. der mRNA-kodierenden Region (Exons) geht. Laut Rabani et al. [14] entspricht das menschliche Exom weniger als 2% des Genoms, aber enthält 85% der bekannten krankheitsverursachenden Mutationen bei Mendelschen Störungen sowie für polygene Krankheiten wie Alzheimer.

## 2.6 Sequenz-Alignment

Im Bereich der Bioinformatik ist ein Sequenz-Alignment ein Verfahren zur Anordnung von DNA-, RNA- oder Proteinsequenzen, um Sequenzen zu vergleichen und Ähnlichkeiten zwischen diesen identifizieren lassen, welche auf funktionale, strukturelle oder evolutionäre Beziehungen zwischen den Sequenzen zurückzuführen sind. Diese Informationen sind für die Berechnung der Länge und die Planung von Mutationen erforderlich [15].

## 2.7 Variantenaufrufe

Die NGS-Technologien haben eine enorme Ausweitung der klinischen Gentests sowohl für Erbkrankheiten als auch für Krankheiten wie Krebs ermöglicht. Die genaue Bestimmung der Varianten in NGS-Daten ist ein wichtiger Schritt, auf den sich praktisch alle nachgelagerten Analyse- und Interpretationsprozesse stützen. Am DKFZ werden die Variantenaufrufe Single nucleotide variants (SNV), Indel, Structural variant (SV) und Copy number variant (CNV) für die DNA-Datenverarbeitung eingesetzt. Für die RNA-Datenverarbeitung werden die Transcript abundance und Fusion gene detection verwendet siehe Abbildung 1. Der linke Teil der Abbildung zeigt die Datenverarbeitung aus



der DNA-Sequenzierung, die für die Erkennung von Mutation verwendet werden, während der rechte Teil die Daten aus der RNA-Sequenzierung zeigt.

### 3 Methodik

In diesem Kapitel wird die Vorgehensweise der systematischen Literaturrecherche und die Methoden für die Analysen wiedergegeben.

#### 3.1 Literatur

Um den aktuellen Stand der Wissenschaft zu ermitteln wurde eine systematische Literaturrecherche durchgeführt. Hierbei wurden die Literaturdatenbanken IEEE<sup>2</sup>, PubMed<sup>3</sup> und ScienceDirect<sup>4</sup> benutzt. Um die relevante Literatur zu identifizieren wurden zuerst Ein- und Ausschlusskriterien, sowie Schlüsselwörter definiert. Die Schlüsselwörter wurden wie folgt definiert:

- (precision oncology) AND (genomics (OR cancer genomics (somatic mutation)))
- (data processing) AND (DNA (OR RNA)) AND (sequencing)
- (variant) OR (alignment) OR (algorithm) AND (DNA sequencing (OR RNA sequencing))
- (one touch pipeline) AND (data flow (OR structure (OR interfaces)))

Die Einschlusskriterien beinhalten Veröffentlichungen von 2011 bis 2021, englische und deutsche Veröffentlichungen sowie den Typ Buch, Fachartikel und Konferenzmaterial. Berücksichtigt wurden keine White Paper.

#### 3.2 Analyse

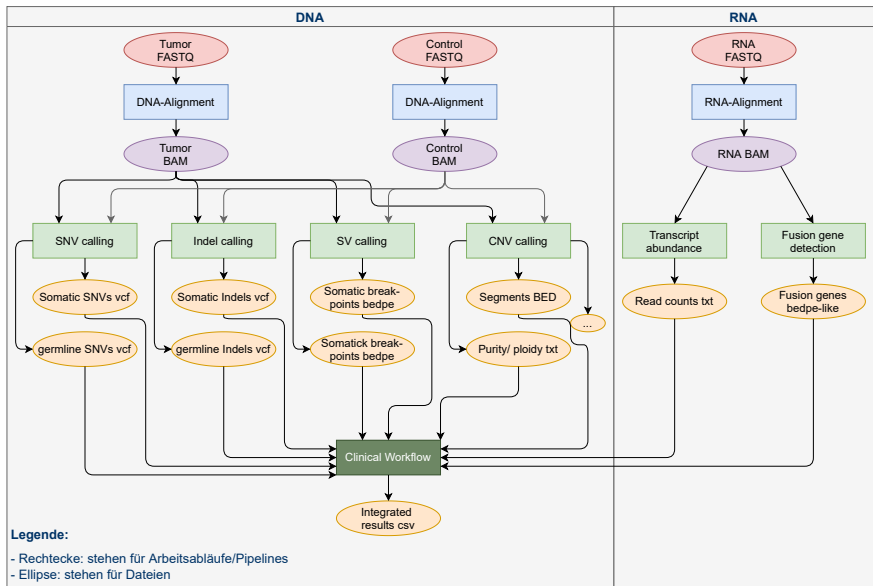
Für die Analyse von OTP werden aus dem Bereich des System Engineering der Ansatz SYSMOD sowie aus dem Bereich Requirements Engineering TORE herangezogen.

SYSMOD ist ein von Tim Weillkiens [5] entwickelter pragmatischer Ansatz für die Modellierung der Anforderungen und der Architektur eines Systems. SYSMOD nutzt zur Modellierung die Systems Modeling Language (SysML), kann aber ebenso mit semiformalen Notationen wie UML verwendet werden. SYSMOD deckt die Themenbereiche Anforderungsmanagement, Systemkontext, Anwendungsfälle ab und beinhaltet Leitfäden für die verschiedenen Aspekte der Modellierung. Dieser Ansatz wird für die Modellierung der Datenflüsse in OTP verwendet [5]. TORE ist eine grundlegende Methode zur Ermittlung von Anforderungen. Hierbei handelt es sich um eine aufgabenorientierte Ermittlung der Anforderungen, wobei die Anforderungsermittlung aus Sicht der Nutzer betrachtet werden. Beim Rahmenwerk TORE handelt es sich um ein Entscheidungsframework, welches 18 Entscheidungen auf vier Ebenen klassifiziert [6]. Auf der Ziel- und Aufgabenebene werden die Stakeholder sowie deren Ziele und Aufgaben ermittelt. Die Ermittlung der Stakeholder kann anhand von Benutzerbeschreibungen oder Rollenbeschreibungen dargestellt werden. Auf der Domänenebene findet die Analyse der aktuellen Situation und die Soll-Analyse statt. Außerdem werden die Systemverantwortlichen ermittelt und die Daten, die zur Domäne gehören behandelt. Die Interaktionsebene definiert die Interaktion zwischen System und dem Nutzer sowie die Funktionen des Systems. Die Systemebene gliedert sich in das Frontend und Backend auf. Das Frontend beinhaltet die Graphical User Interface (GUI) in der das Layout und Struktur sowie die Dialoge und Funktionen der GUI beschrieben werden. Das Backend wiederum befasst sich mit der Systemarchitektur und den internen Funktionen und Daten [16]. TORE wird für die Ermittlung der Aufgaben aus Sicht des Nutzers eingesetzt.

<sup>2</sup><https://ieeexplore.ieee.org/Xplore/home.jsp>

<sup>3</sup><https://pubmed.ncbi.nlm.nih.gov/>

<sup>4</sup><https://www.sciencedirect.com/>



**Abb. 1: Eine schematische Übersicht über die Datenverarbeitung für die Präzisionsonkologie Programme**

## 4 Stand der Technik

Für die Analyse des Systems OTP wurde die Arbeit von Reisinger et al. [2] herangezogen. Reisinger et al. beschreibt den ausführlichen Nutzen von OTP. Außerdem wird die Entwicklung des Systems beschrieben sowie der Vergleich zu anderen Plattformen wie HTS-Flow [4], WASP System [17], QuickNGS [18], Chipster [19] und Omics Pipe [20] gezogen. Beim Vergleich der Plattformen wurden Aspekte berücksichtigt wie die Speicherung der Metadaten in der Datenbank, die Benutzeroberfläche, die Zugriffskontrolle, der Automatisierungsprozess, die Reproduzierbarkeit und die Verfügbarkeit.

Die Plattform HTS-Flow, dass vom Italienischen Institut in Mailand entwickelt wurde, interagiert mit dem Labor-Informations-Management-System und automatisiert die NGS-Analyse über eine Benutzeroberfläche [4]. Das System WASP dient der Koordination der Probeneinreichung und der anschließenden Verarbeitung. Hierbei liegt der

Schwerpunkt auf der Open-Source-Strategie, wobei andere Institute aktiv zur Zusammenarbeit eingeladen werden [21]. QuickNGS beinhaltet eine Vielzahl von Werkzeugen zur Beschleunigung der am häufigsten verwendeten Verfahren für NGS-Daten [18]. Die Plattform Chipster bietet die Möglichkeit, Workflows in einem Desktop-Programm zu erstellen und zu speichern, wobei eine Schnittstelle zu einem lokalen Server besteht. Der Fokus bei dieser Plattform liegt auf der Integration vieler Tools und externer Datensätze [19]. Die von Fischer et al. [20] vorgestellte Plattform Omics Pipe bietet ein Python-Paket, das viele Tools integriert und Nutzern eine Batch-Verarbeitung ermöglicht.

## 5 Analyse

In diesem Abschnitt wird das System OTP analysiert bzw. modelliert. Als Modellierungssprache werden SysML und UML verwendet. Das Vorgehen basiert hierbei auf den

Ansätzen SYSMOD und TORE, die in Kapitel 3.2 beschrieben sind. Da es sich hierbei um Rahmenwerke handelt, werden nicht alle Aspekte der Prozesse verwendet, sondern nur die, die zum Kontext passen.

### 5.1 Beschreibung des Systems

Die One Touch Pipeline (OTP) ist eine Plattform, in der die Datenerfassung, die Ausführung von Bioinformatik Pipelines sowie die Verarbeitung von NGS-Daten und die daraus resultierenden Daten und Metadaten verwaltet werden [2]. Die Plattform entstand am DKFZ und verwaltet aktuell genomische Daten im Petabyte-Bereich. Der Code der Plattform ist OpenSource<sup>5</sup>.

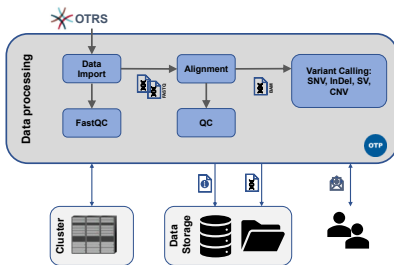


Abb. 2: Datenverarbeitung in OTP [2]

In Abbildung 2 ist der Ablauf des Datenflusses in OTP dargestellt. Das externe System Open Technology Real Service (OTRS), das in der Abbildung zu sehen ist, ist ein Ticket-system und wird für den automatischen Datenimport verwendet. Im ersten Schritt werden die die Rohdaten im Dateisystem und die Metadaten in einer Datenbank gespeichert. Im nächsten Schritt werden die Qualitätskontrollwerte für die FASTQ-Dateien mittels FASTQC berechnet <sup>6</sup>. Eine FASTQ-Datei ist ein ASCII-basierte Textdatei, die Sequenzdaten aus den Clustern beinhaltet,

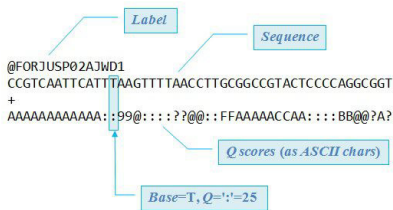
welche den Filter auf einer Flowcell durchlaufen haben. D.h. die Datei besteht aus einem Eintrag pro sequenziertem DNA-Fragment ("Reads". Jeder Eintrag besteht aus 4 Zeilen, von denen eine die Basenpaar-Reihfolge enthält, und eine weitere einen hexadezimal kodierten Qualitätswert für die Sequenzierung angibt [22]. Sofern das Referenzgenom des sequenzierten Organismus bekannt ist, müssen diese Reads auf das Referenzgenom abgebildet werden. Dieser Prozess des Mappings wird als Alignment bezeichnet. Das Ergebnis eines Alignments ist eine Datei des Typs Sequence Aglinment Map (SAM). Bei SAM-Dateien handelt es sich um Textdateien, die aus mehreren Einträgen bestehen und beinhalten Informationen darüber, wo der Read im Referenzgenom zugeordnet ist sowie Informationen über die Zuordnungsqualität. Durch die Komprimierung wird das Textdateiformat SAM in das Binärformat Binary Alignment Map (BAM) umgewandelt. Dieser Vorgang der Umwandlung findet statt, da die SAM-Dateien zu riesigen Textdateien werden können und somit sehr viel Speicherplatz auf des Festplatte beansprucht wird und der Zeitaufwand beim durchsuchen der Datei zu hoch ist. Anschließend wird die BAM-Datei sortiert und Duplikate markiert. Abschließend werden weitere Verarbeitungsschritte, die in Abbildung 1 abgebildet sind, wie das Varianten Calling, dass zur Identifizierung und Extraxtion von Mutationen definiert ist, durchgeführt [2]. Nach Abschluss der Analyse wird der Dateneinreicher und die im Projekt zugehörigen Mitglieder per E-Mail benachrichtigt [2].

### 5.2 Systemidee und -ziele

Die Sequenzierprogramme am DKFZ verarbeiten mehrere tausend Proben pro Jahr aus dem Bereich der Krebsgenomik mit mehreren Pipelines in hunderten von Projekten sowie in verschiedenen Kombinationen und mit unterschiedlichen Parametern. Dabei automatisiert OTP den kompletten digitalen Prozess

<sup>5</sup><https://gitlab.com/one-touch-pipeline/otp>

<sup>6</sup>Tool zur Analyse von FastQ-Dateien



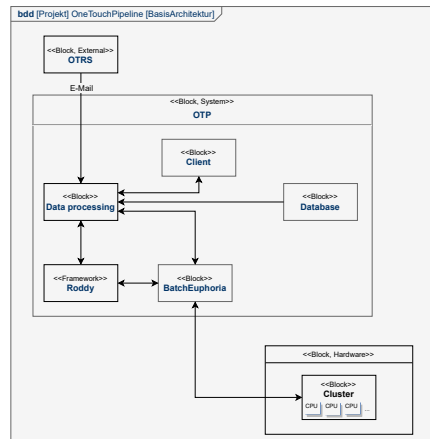
**Abb. 3: Beispiel für einen Eintrag in einer FASTQ-Datei**

vom Import der Rohsequenzen bis hin zum Alignment und die Erkennung von Mutationen. Dabei ergeben sich aus dieser Idee Systemziele wie die Senkung von Zeitaufwand bei der Verarbeitung der durch die Sequenzierung erzeugten Daten, die Automatisierung des Ablaufs und die Möglichkeit zu bieten in einem zentralen System zu arbeiten.

### 5.3 Basisarchitektur

Die Basisarchitektur ermöglicht einen ersten Überblick für die Umsetzung von Systemidee sowie den Systemzielen [5]. Hierbei wird die Architektur von OTP modelliert. Dies ist notwendig, um weitere Ebenen zu abstrahieren. Die Basisarchitektur ist in Abbildung 4 abgebildet. Dabei sind die grundlegenden Blöcke des Systems: ein Service-Management-System, eine Datenbank, ein Framework Roddy und die Bibliothek BatchEuphoria, welche die Kommunikation mit dem Cluster ermöglicht. Das **Service-Management-System** wird dazu verwendet, um den manuellen Aufwand beim Import zu minimieren und den Prozess des Datenimports zu automatisieren. Außerdem wird damit der Import in der Nacht und am Wochenende ermöglicht. Die **relationale Datenbank** wird für die Organisation, Strukturierung, Verarbeitung und Visualisierung der Daten verwendet. Der Cluster erhält ein Batch-Verarbeitungssystem. Dabei kann die Übermittlung der Daten direkt oder indirekt, mittels der Ausführung des Frameworks **Roddy**, verlaufen. Bei simplen Aufgaben wie die Datenverwaltung und

der Ausführungen von FastQC, übermittelt OTP diese direkt dem Cluster. Komplexere Abläufe hingegen werden auf dem indirekten Weg mit Roddy übermittelt. **BatchEuphoria** ist eine Bibliothek für Entwickler von Cluster-/Batch-Systemen, um Batch-Aufträge von Java aus zu erstellen [2]. Auf der Hardwarekomponente High performance computing (HPC) Cluster werden die Verarbeitungen der Aufgaben/Jobs von genomischen Daten, die rechenintensiv sind, bearbeitet.



**Abb. 4: Basisarchitektur**

### 5.4 Stakeholder

Als Stakeholder werden alle Personen oder Institutionen, die ein Interesse am Erfolg des Projektes haben, bezeichnet. Die Stakeholder können als potenzielle Quellen für Anforderungen gesehen werden, da diese die potentiellen Nutzer des Systems sein können [5]. In Tabelle 1 werden die Stakeholder und die verschiedenen Interessen aufgelistet.

### 5.5 Fluss der Aktivitäten

In Abbildung 5 ist der manuelle Import von Daten abgebildet. Die Daten werden anhand von vor definierten Eigenschaften überprüft, die für die Verarbeitung und Speicherung auf dem Dateisystem erforderlich sind. Bei einer

**Tab. 1: Stakeholder und ihr Interesse**

Stakeholder	Interesse
ProjektleiterInnen	Möchte die Projektplanung durchführen und eine Übersicht des Projektes erhalten.
BioinformatikerInnen	Möchte die Arbeitsabläufe und Workflows verwalten.
DatenmanagerInnen	Möchte die Daten für die Verarbeitung importieren.
ÄrzteInnen im molekularen Tumorboard	Möchte die gewonnenen Ergebnisse aus OTP analysieren.
Genomics and Proteomics Core Facility (GPCF)	Möchte die Sequenzdaten für OTP liefern.
Nidergelassene ÄrzteInnen	Möchte nach Abschluss des Prozesses Empfehlungen für eine Behandlung aussprechen.

erfolgreichen Validierung erfolgt der tatsächliche Importvorgang. Dabei werden die Metadaten in der Datenbank und die Sequenzierungsdateien auf dem Dateisystem installiert. Falls die Validierung aufgrund von fehlenden Informationen fehlschlägt, wird eine manuelle Interaktion benötigt, um die Metadaten zu vervollständigen. Neben dem manuellen Import kann ebenso ein automatischer gestartet werden. Dieser ist in Abbildung 6 dargestellt. Hierfür sendet der/die NutzerIn eine E-Mail mit Sequenzereinheiten an das Service Management System OTRS. Dieser Vorgang kann nur stattfinden, sobald die Sequenzierdaten, vom GPCF auf dem Speicher liegen. Aus der E-Mail wird ein Ticket in OTRS generiert sowie die URL auf dem Server aufgerufen. Anschließend wird der Inhalt der

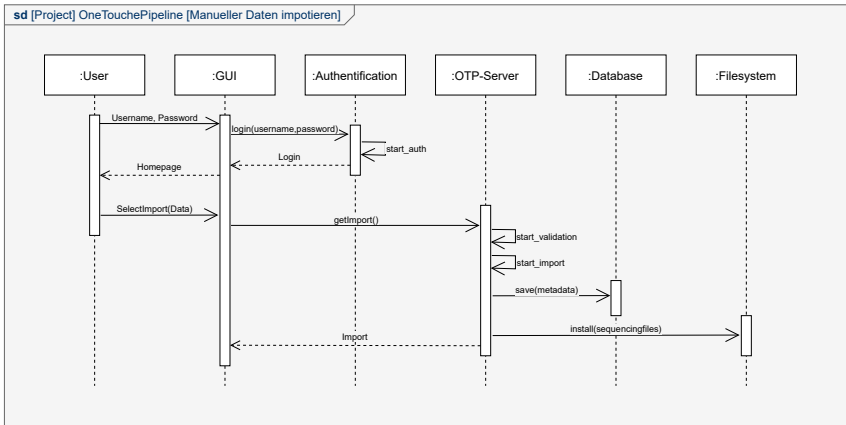
Sequenzierungsdatei geparkt und der Importvorgang eingeleitet. Die nächsten Schritte sind identisch zum manuellen Import.

Nach dem Importieren der Sequenzierdaten wird das Alignment und mehrere Verarbeitungsschritte wie in Kapitel 5.1 beschrieben, durchgeführt. Die Informationen nach dem die Verarbeitung der Daten stattfand, wird im clinical Workflow zusammengeführt. Die Ergebnisse der Pipelines werden auf dem Dateisystem und die QC-Werte, die nach dem Alignment erzeugt werden, sowie weitere Ergebnisse in der Datenbank gespeichert. Abschließend werden die meisten Ergebnisse über einen Webclient bereitgestellt und der Dateneinreicher sowie die Projektmitglieder per E-Mail benachrichtigt.

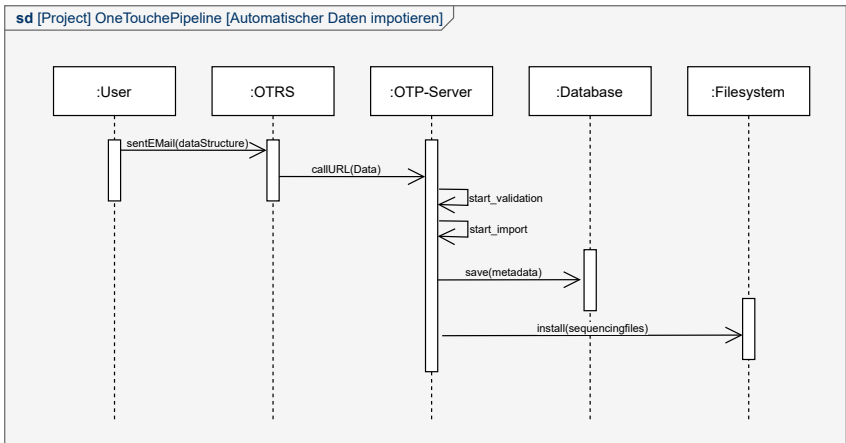
In Abbildung 7 wird in einem Sequenzdiagramm die Kommunikation der integrierten Komponenten in OTP während einer Verarbeitung der Daten dargestellt. Zuerst werden alle benötigten Daten bzgl. Eingabeinformation und Konfigurationsinformation aus der Datenbank gezogen und in OTP gesammelt. Diese Daten werden anschließend Roddy zur Verfügung gestellt, welcher nach dem Empfang die Analyse triggert und startet. Nach dem Start werden die Jobs an den Cluster über BatchEuphoria gesendet, der im Anschluss IDs erstellt und diese zurück an Roddy gibt. Roddy stellt diese IDs wiederum OTP zu Verfügung, um im Nachhinein die Status der Jobs über BatchEuphoria mit der ID auf dem Cluster abfragen zu können.

## 5.6 Systemkontext

Das Systemkontextdiagramm, das in Abbildung 8 zu sehen ist, repräsentiert diesen Systemkontext und bietet erste Informationen über den Kommunikationsfluss von und zum System. Die externen Interaktionspartner sind die Systemakteure. Das Systemkontextdiagramm ist kein definiertes SysML Diagramm, sondern Teil des SYSMOD Ansatzes. Daher wird in dieser Arbeit ein Blockdefinitionsdiagramm für die Modellierung des



**Abb. 5: Aktivitätsdiagramm - Manueller Datenimport**



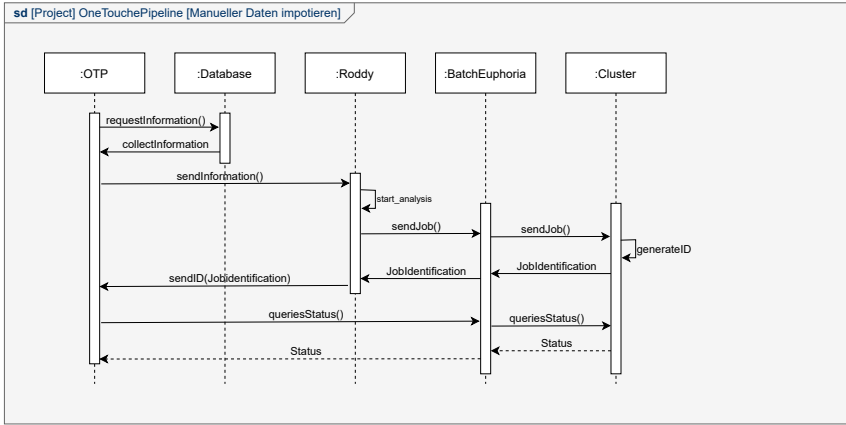
**Abb. 6: Sequenzdiagramm - Automatischer Datenimport**

Systemkontextes gewählt [5]. Beim Dateneinreicher handelt es sich um den Akteur, der den automatischen Import startet bzw. die E-Mail in OTRS verfasst. Das DKFZ befasst sich mit der Datenanalyse nach der Sequenzierung der Daten. Das GPCF ist für die Einlieferung der Sequenzierdaten zuständig. Die Ärzte sind die, die Patienten eine Behandlungsmethode empfehlen bzw. sie behandeln. Neben den Systemakteuren gibt es noch das externe System Open Technology Real Service (OTRS), das ein Ticketsystem ist. Die

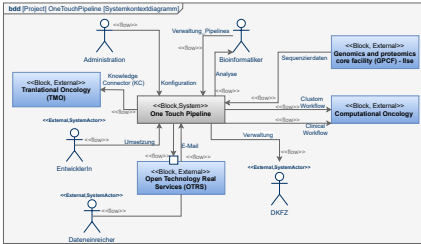
Administration nutzt das System, indem es das System konfiguriert und verwaltet. Die BioinformatikerInnen integrieren mit dem System, um Arbeitsabläufe zu verwalten und um die Analyse auszuwerten.

## 5.7 Anwendungsfälle

In diesem Kapitel werden die Anwendungsfälle beschrieben. Die Anwendungsfälle ergeben sich durch die Anforderungen und repräsentieren die Dienstleistungen des Systems

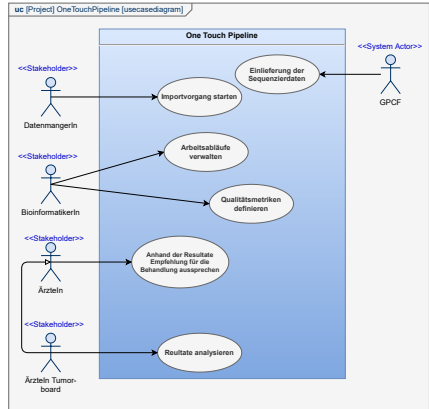


**Abb. 7: Sequenzdiagramm - Kommunikation der Komponenten bei der Datenverarbeitung**



**Abb. 8: Systemkontextdiagramm**

und die Nutzer welche diese Dienstleistungen in Anspruch nehmen [5]. Diese Anwendungsfälle werden mithilfe eines SysML Anwendungsfalldiagramms modelliert.



**Abb. 9: Anwendungsfalldiagramm**

## 6 Fazit

In dieser Arbeit wurde das System OTP modelliert, das Nutzern ermöglicht Genomdaten zu sequenzieren. Hierzu wurden die Basisarchitektur, der Ist-Stand der Datenverarbeitung, die System- und Actor-Interaktion, der Anwendungsfall sowie der Systemkontext mithilfe von SysML und UML modelliert. Die einzelnen Schritte des Ansatzes erscheinen sinnvoll. Allerdings sollte, wie auch der Herausgeber dieses Ansatzes, Tim Weilkens, es beschreibt [5], abgewogen werden welche

dieser Schritte für ein Projekt sinnvoll sind. Vor allem für Neueinsteiger dieses Ansatzes ist nicht immer klar welche Schritte benötigt werden und welche nicht. Die in dieser Arbeit modellierten Diagramme umfassen nicht alle Inhalte und vor allem nicht bis in die tiefste Ebene. Dies liegt daran dass die vollständige Modellierung von OTP den Rahmen dieser Arbeit sprengen würde und daher die

wesentlichen Bestandteile und Vorgänge dokumentiert wurden, um einen Überblick des Systems zu erhalten.

## Literatur

- [1] J.-D. Hoppe und P. Scriba, „Richtlinien zur prädiktiven genetischen Diagnostik“ 2003.
- [2] E. Reisinger u. a., „OTP: An automated system for managing and processing NGS data,“ *Journal of biotechnology*, Jg. 261, 2017. DOI: 10.1016/j.jbiotec.2017.08.006.
- [3] L. J. Jennings u. a., „Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists,“ *The Journal of molecular diagnostics : JMD*, Jg. 19, Nr. 3, 2017. DOI: 10.1016/j.jmol dx.2017.01.011.
- [4] V. Bianchi u. a., „Integrated Systems for NGS Data Management and Analysis: Open Issues and Available Solutions,“ *Frontiers in genetics*, Jg. 7, 2016, ISSN: 1664-8021. DOI: 10.3389/fgene.2016.00075.
- [5] T. Weilkens und R. M. Soley, *Systems Engineering mit SysML/UML: Anforderungen, Analyse, Architektur*, 3., überarb. und aktualisierte Aufl. Heidelberg: dpunkt.verl., 2014, ISBN: 3864900913.
- [6] B. Paech und K. Kohler, „Task-Driven Requirements in Object-Oriented Development,“ in *Perspectives on Software Requirements*, Ser. The Springer International Series in Engineering and Computer Science, J. C. S. Prado Leite und J. H. Doorn, Hrsg., Boston, MA: Springer, 2004. DOI: 10.1007/978-1-4615-0465-8\{textunderscore}3.
- [7] J. A. Shapiro und R. von Sternberg, „Why repetitive DNA is essential to genome function,“ *Biological Reviews*, Jg. 80, Nr. 2, 2005. DOI: 10.1017/S1464793104006657.
- [8] S. Behjati und P. S. Tarpey, „What is next generation sequencing?“ *Archives of disease in childhood. Education and practice edition*, Jg. 98, Nr. 6, 2013. DOI: 10.1136/archdischild-2013-304340.
- [9] D. E. Draper, „A guide to ions and RNA structure,“ *RNA*, Jg. 10, Nr. 3, 2004, ISSN: 1355-8382. DOI: 10.1261/rna.5205404.
- [10] Z. Wang, M. Gerstein und M. Snyder, „RNA-Seq: a revolutionary tool for transcriptomics,“ *Nature Reviews Genetics*, Jg. 10, Nr. 1, 2009. DOI: 10.1038/nrg2484.
- [11] R. Kamps u. a., „Next-Generation Sequencing in Oncology: Genetic Diagnosis, Risk Prediction and Cancer Classification,“ *International journal of molecular sciences*, Jg. 18, Nr. 2, 2017. DOI: 10.3390/ijms18020308.
- [12] S. Y. Moorcraft, D. Gonzalez und B. A. Walker, „Understanding next generation sequencing in oncology: A guide for oncologists,“ *Critical reviews in oncology/hematology*, Jg. 96, Nr. 3, 2015. DOI: 10.1016/j.critrevonc.2015.06.007.
- [13] C. G. van El u. a., „Whole-genome sequencing in health care: recommendations of the European Society of Human Genetics,“ *European Journal of Human Genetics*, Jg. 21, Nr. 6, 2013. DOI: 10.1038/ejhg.2013.46.
- [14] B. Rabbani, M. Tekin und N. Mahdieh, „The promise of whole-exome sequencing in medical genetics,“ *Journal of Human Genetics*, Jg. 59, Nr. 1, 2014, ISSN: 1435-232X. DOI: 10.1038/jhg.2013.114.



- [15] A. D. Prjibelski, A. I. Korobeynikov und A. L. Lapidus, „Sequence Analysis,“ in *Encyclopedia of bioinformatics and computational biology*, S. Ranganathan, M. Gribskov, K. Nakai und C. Schönbach, Hrsg., Amsterdam, Boston und Heidelberg: Elsevier, 2019, S. 292–322. DOI: 10.1016/B978-0-12-809633-8.20106-4.
- [16] S. Adam, J. Doerr, M. Eisenbarth und A. Gross, „Using Task-oriented Requirements Engineering in Different Domains &#150; Experiences with Application in Research and Industry,“ in *17th IEEE International Requirements Engineering Conference, 2009*, IEEE Computer Society, 2009. DOI: 10.1109/RE.2009.16.
- [17] A. S. McLellan u. a., „The Wasp System: an open source environment for managing and analyzing genomic data,“ *Genomics*, Jg. 100, Nr. 6, 2012. DOI: 10.1016/j.ygeno.2012.08.005.
- [18] P. Wagle, M. Nikolić und P. Frommolt, „QuickNGS elevates Next-Generation Sequencing data analysis to a new level of automation,“ *BMC genomics*, Jg. 16, 2015. DOI: 10.1186/s12864-015-1695-x.
- [19] M. A. Kallio u. a., „Chipster: user-friendly analysis software for microarray and other high-throughput data,“ *BMC genomics*, Jg. 12, 2011. DOI: 10.1186/1471-2164-12-507.
- [20] K. M. Fisch u. a., „Omics Pipe: a community-based framework for reproducible multi-omics data analysis,“ *Bioinformatics (Oxford, England)*, Jg. 31, Nr. 11, 2015. DOI: 10.1093/bioinformatics/btv061.
- [21] E. McPherson, „Genetic diagnosis and testing in clinical practice,“ *Clinical medicine & research*, Jg. 4, Nr. 2, 2006, ISSN: 1539-4182. DOI: 10.3121/cmr.4.2.123.
- [22] I. Gabdank u. a., „Prevention of data duplication for high throughput sequencing repositories,“ *Database : the journal of biological databases and curation*, Jg. 2018, 2018. DOI: 10.1093/database/bay008.



# Autorenverzeichnis

## A

Aini, Y. .... 110

## B

Bauer, M. .... 29

Bruhn, S. .... 124

Brückner, A. .... 44

## E

Emirzeoglu, M. .... 137

## F

Ferrera, G. .... 56

## H

Haynes, M. .... 1

## I

Ickenroth, J. .... 16

## R

Rabus, M. .... 68

## S

Storz, P. .... 80

## W

Welker, F. .... 98

---

Hochschule Reutlingen  
Reutlingen University  
Fakultät Informatik  
Human-Centered Computing  
Alteburgstraße 150  
D-72762 Reutlingen

Telefon: +49 7121 / 271-4002  
Telefax: +49 7121 / 271-4042

E-Mail: [infoinside@reutlingen-university.de](mailto:infoinside@reutlingen-university.de)  
Website: [infoinside.reutlingen-university.de](http://infoinside.reutlingen-university.de)

ISBN 978-3-00-070521-2

