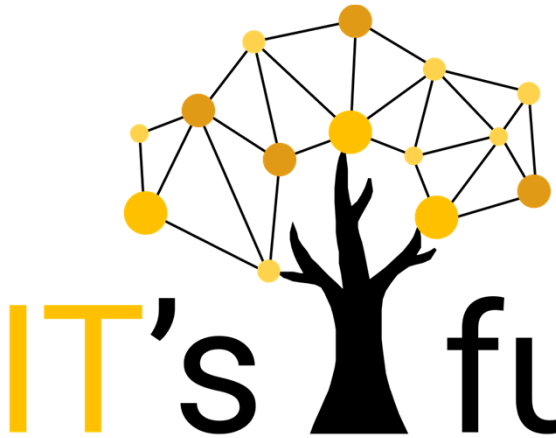
Informatics Inside



Informatics Inside
Frühling 2022

Tagungsband

Herausgeber*innen:
Uwe Kloos, Hochschule Reutlingen
Natividad Martinez, Hochschule Reutlingen
Gabriela Tullius, Hochschule Reutlingen

Impressum

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
Website: 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. Elma Bajraktari
B. Sc. Ronell Bresler
B. Sc. Philip Engler
B. A. Moritz Häcker
B. Sc. Kora Helm
B. Sc. Elisaveta Just
B. Sc. Elena Kirsch
B. Sc. Tobias Holger Lauxmann
B. Sc. Demis Moysidis
B. Sc. Michael Pietschmann
B. Sc. Anna Samira Praetorius
B. Eng. Michael Stadelmaier
B. Sc. Leila Taraman
B. Sc. Bruno Teixeira
B. Sc. Marcel Andre van Look
B. Sc. Julia Wagner
B. Eng. Clara Werner

Verlag: Hochschule Reutlingen
ISBN 978-3-00-072060-4

©2022 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,

die Informatics Inside ist seit über 13 Jahren ein fester Bestandteil des akademischen Jahres an der Fakultät für Informatik der Hochschule Reutlingen. Die Konferenz wird gegenwärtig von Studierenden des Masterstudiengangs Human-Centered Computing selbstständig organisiert und bildet einen wichtigen Teil der wissenschaftlichen Ausbildung. Die Studierenden haben ihre Themen selbständig gewählt und nicht selten sind es Fragen, die sie bereits durch das ganze Studium begleiten. Sie bereiten diese im Format einer wissenschaftlichen Ausarbeitung auf wobei Inhalt, Vollständigkeit und Nachvollziehbarkeit entscheidende Faktoren sind. Die Ergebnisse dieser vertieften Auseinandersetzung mit relevanten Anwendungsthemen der Informatik, können Sie in diesem Tagungsband nachlesen. Die Anwendungsdomänen reichen von der Medizin über Wirtschaft bis zu den Medien. Dabei werden aktuelle Fragestellungen des menschzentrierten Einsatzes von künstlicher Intelligenz, Softwaretechnik, Datenanalyse und Kommunikation sowie der digitalen Transformation behandelt. Es wird deutlich, dass der Nutzen von IT Lösungen für den Menschen im Mittelpunkt der Veranstaltung steht. In dieser Auflage illustriert die Auswahl der Themen daher, wie ich finde in eindrucksvoller Weise, auch die große Bandbreite und den starken Praxisbezug der Hochschulen für angewandte Wissenschaften.

Das Motto der Veranstaltung „IT's Future" ist Programm und macht die Relevanz der Informatik für alle Lebensbereiche sowie die zukünftige Innovations- und Wettbewerbsfähigkeit von Industrie und Forschung deutlich. „IT's Future" kann vor dem Hintergrund des erheblichen Mangels an kompetentem Nachwuchs von Informatikerinnen und Informatikern auch als Mahnung verstanden werden. Die rückläufigen Zahlen bei den Zulassungen in den Informatikstudiengängen in Deutschland bei gleichzeitig fortschreitender Digitalisierung sind insgesamt ein erhebliches Risiko für die Zukunftsfähigkeit unseres Standorts. Hier ist eine entsprechende neue gesellschaftliche Gewichtung der IT erforderlich. Umso erfreulicher ist es, dass die Studierenden mit Ihren Ausarbeitungen zeigen, dass sie eine am Menschen orientierte Informatik in Zukunft gestalten wollen und dazu auch in der Lage sind. Im vorliegenden Tagungsband können Sie sich von diesem Ansatz überzeugen und ich hoffe, dass sie sich von der Begeisterung der Autoren für die Möglichkeiten der IT anstecken lassen.

Ich wünsche Ihnen eine informative Tagung, mit spannenden Vorträgen, inspirierenden Diskussionen und nicht zuletzt der interessanten Lektüre dieses Tagungsbandes.

Reutlingen im Mai 2022

Prof. Dr. Christian Thies

Inhaltsverzeichnis

Vorwort III
Konzeption einer On-Premises Lernumgebung für Programmieraufgaben
Bresler, R. 1
Taxonomie- und Entwicklungsvorschlag für Serious Games
Lauxmann, T. 14
Pre-Study: Adaptation of 3D sign language avatars to the speakers appearance and its influence on user experience
Krautmacher, L. 26
Untersuchung der Verbesserung der User Experience durch maschinelles Lernen für Bewerbende auf Bewerbungsplattformen
Kruck, H. 37
Exploding Gradients bei Conv-LSTM-Netzwerken unter Verwendung des moving-MNIST Datensatzes
Müller, K. 48
Einsatz von KI im RE: Ein systematisches Literaturreview
Bajraktari, E. 60
Affective Computing zur Diagnose und Therapie von Erkrankungen
Uslu, I. 74
Sportdatenvisualisierung mittels Daten des Fußballsimulationsspiels FIFA
Moysidis, D. 89
Visualisierungskonzept zur Vermittlung der Blockchain Technologie für Laien
Wickert, F. 100
Methoden zur Optimierung von Path-Tracing in Echtzeitanwendungen
Grendatsch, D. 116
State-of-the-Art Analyse von 3D-Scans primitiver Formen
Stadelmaier, M. 128
BIM Collaboration Format - Theorie und Praxis
Häcker, M. 139
Autorenverzeichnis 148

Konzeption einer On-Premises Lernumgebung für Programmieraufgaben

Ronell Bresler
Hochschule Reutlingen
Reutlingen, Deutschland
Ronell.Bresler@Student.Reutlingen-University.de

Abstract

[Motivation/Kontext] Eine Lernumgebung für Programmieraufgaben ist im Kontext der Hochschule Reutlingen ein essenzieller Bestandteil für die Lehre. Derzeit bringen Studierende ihr eigenes Endgerät mit in die Informatiklabore und den Hörsaal. Dies wird auch als *Bring your own device* (BYOD) bezeichnet. Somit besitzen die Endgeräte unterschiedliche Hardware und Betriebssysteme. **[Problemstellung]** Durch die Verwendung verschiedener Plattformen kommt es zu Problemen bei der Installation der benötigen Software auf den Endgeräten. Aus diesem Grund stellt sich die Frage, wie eine einheitliche Lernumgebung für Programmieraufgaben zur Verfügung gestellt werden kann, die in Informatiklaboren und im Hörsaal funktioniert sowie BYOD unterstützt. Diese Lösung soll ebenfalls Hilfestellungen durch DozentInnen reduzieren. **[Lösungsidee/Ergebnisse]** Im Rahmen dieses Artikels wird ein Konzept für eine Lernumgebung für Programmieraufgaben entwickelt. Dabei werden ähnliche bestehende Systeme verglichen sowie Anforderungen an das System erhoben. **[Beitrag]** Mit dem Konzept wird eine Lernumgebung für Programmieraufgaben bereitgestellt, welche auf unterschiedlichen Plattformen ausführbar ist. Die Lernumgebung ermöglicht es, einen Buildprozess für Quellcode auszuführen, zu Debuggen und zu Versionieren. Die Evaluation zeigt, dass das Konzept auf Windows- und Linux Plattformen anwendbar ist. Durch eine Prüfung des Prototyps durch Stakeholder wurden Verbesserungsvorschläge identifiziert, welche in die nächste Iteration des Prototyps einfließen.

CCS Concepts

• Applied computing~Education~Interactive learning environments

Keywords

Software Engineering, BYOD, Learning environment, Requirements Engineering

Betreuer Hochschule: Prof. Dr. rer. nat. Kücherer
Hochschule Reutlingen
Christian.Kuecherer@Reutlingen-University.de
Informatics Inside Frühjahr 2022
11. Mai 2022, Hochschule Reutlingen
Copyright 2022 Ronell Bresler

1 Einleitung

Lernumgebungen für Programmieraufgaben müssen nicht immer kompliziert von den NutzerInnen auf dem eigenen Endgerät installiert werden. Dieses Vorgehen wird aktuell im Informatik3-Praktikum des Studienganges Medien- und Kommunikationsinformatik der Hochschule Reutlingen verwendet. Es besteht die Möglichkeit, die Lernumgebung bwLehrpool zu verwenden,

welche in Kapitel 3.2 genauer beschrieben wird. Nach dem Modulhandbuch[1] soll das Informatik3-Praktikum den Studierenden wichtige Kenntnisse im Bereich der Objektorientierung, Programmiersprachen, Pointer und Nebenläufigkeiten vermitteln. Ein Lernmedium, welches die Studierenden während der Informatik-Lehrveranstaltung und des Praktikums verwenden, sind ihre digitalen Endgeräte. Die DozentInnen vor Ort können Hilfestellung in Präsenzform bieten. Nach Ehlers [4] ist diese Form des Lernens das *Blended Learning*, welches eine Mischform des E-Learnings und der Präsenzlehre ist. Nach Pilotto [9] wird unter dem Begriff E-Learning eine Lehrmethode bezeichnet, welche über digitale Medien erfolgt und die DozentInnen und Studierenden räumlich voneinander trennen.

Für die Durchführung des Informatik Praktikum wird folgende Software benötigt: eine integrierte Entwicklungsumgebung (engl. Integrated Development Environment, abgekürzt IDE) und Git zur Versionskontrolle. Des Weiteren werden die Interpreter und Compiler der Programmiersprachen C, C++, C#, Java und Python benötigt. Durch die Verwendung unterschiedlicher Betriebssysteme und Hardware wird für die Installation der Software das erste Praktikum der Lehrveranstaltung benötigt und bei Installationsproblemen ist individuelle Hilfestellung notwendig. Aus diesem Grund soll ein Konzept einer Laborumgebung für Programmieraufgaben entwickelt werden, um die Installation der benötigten Software zu vereinfachen und Plattformunabhängig bereitzustellen.

1.1 Forschungsfragen

In diesem Kapitel werden die Forschungsfragen und die dazugehörigen Suchterme dargestellt. Die Forschungsfragen werden in Tabelle 1 aufgelistet.

Tabelle 1: Auflistung der Forschungsfragen

ID	Forschungsfrage
RQ1	Welche Arten von Softwarebereitstellung gibt es für Lernumgebungen im Software Engineering?
RQ2	Welche Funktionsbausteine besitzen Lernumgebungen im Software Engineering?

Der Suchterm für RQ1 lautet: *("deployment") AND ("learning environment") AND ("software engineering")*

Der Suchterm für RQ2 lautet: *("computer science education") AND ("programming") AND ("teaching") AND ("software")*

1.2 Ziel und Kontext der Arbeit

Das Ziel dieser Arbeit ist es, ein Konzept für eine Lernumgebung für Programmieraufgaben auszuarbeiten. Hierbei werden Anforderungen durch Interviews, Brainstorming und Dokumentenanalysen erhoben, die anschließend zur Konzeption einer einheitlichen Lernumgebung für Programmieraufgaben dienen. Eine genauere Beschreibung dieser Methoden aus dem Requirements Engineering (RE) befindet sich in Kapitel 2.1. Diese Systemumgebung soll für das Informatik-Praktikum, für Informatik Lehrveranstaltungen und auf den BYODs eingesetzt werden. Die

[1] https://www.inf.reutlingen-university.de/fileadmin/user_upload/Fakultaet_INF/Downloads/Studium/Bachelor/MKI/hsrt_informatik_MKIB-modulhandbuch.pdf

Studierenden können mit dieser Lernumgebung Programmieraufgaben lösen und abgeben. Hierzu soll die Lernumgebung folgende Plattformen unterstützen, welche nach dem Microsoft[2] Runtime-ID-Katalog (RID) in Tabelle 2 dargestellt werden.

Tabelle 2: Unterstützende Runtime Identifier der Systemumgebung

RID	Betriebssystem
win-x64	Windows
linux-x64	Ubuntu Desktop
osx-x64	MacOS

1.3 Literaturrecherche

Für diesen Artikel wird eine systematische Literaturrecherche durchgeführt. Bei der Literaturrecherche wird gezielt nach Fachliteratur, Konferenzen und Studien gesucht, die sich mit Lernumgebungen im Bereich des Software Engineerings (SWE) befassen. Dabei wurden folgende Datenbanken verwendet: Science Direct[3], IEEE-Xplore[4], ACM[5] und Springer Link[6]. Als erstes wird in den genannten Datenbanken nach den Suchtermen gesucht. Hierfür wurden die Einschlusskriterien aus Tabelle 3 verwendet. Danach wird das Abstract, die Zusammenfassung und falls vorhanden die Conclusion durchgelesen und beurteilt. Am Ende wird überprüft, ob die Literatur zur Beantwortung der Fragestellungen beiträgt.

Tabelle 3: Einschlusskriterien

ID	Einschlusskriterium
IC1	Der Artikel wurde zwischen 2010-2022 veröffentlicht und muss einen Beitrag zum Thema leisten
IC2	Literatur aus IEEE-Xplore, Science Direct, ACM und Springer Link und muss einen Beitrag zum Thema leisten
IC3	Die Literatur muss auf deutscher oder englischer Sprache veröffentlicht sein und muss einen Beitrag zum Thema leisten

1.4 Struktur der Arbeit

Um einen Überblick über das Themengebiet zu verschaffen, wird in Kapitel 2 der benötigte Grundlagenteil vermittelt. Kapitel 3 beschreibt den aktuellen Stand der Technik und somit welche bestehenden ähnliche Lernumgebungen für Programmieraufgaben bereits existieren. In Kapitel 4 wird das Konzept ausgearbeitet, indem Anforderungen erhoben und eine Ist-Analyse durchgeführt wird. In Kapitel 5 erfolgt die Evaluation des Konzepts anhand eines Prototyps, welcher durch ProbandInnen getestet wurde. Im letzten Kapitel erfolgt ein Fazit und die Diskussion.

2 Grundlagen

In diesem Kapitel werden die Begriffe des Requirements Engineering (RE), der Softwarebereitstellungsmethoden sowie der Unterschied zwischen virtuellen Maschinen und Containerdiensten genauer erläutert.

2.1 Requirements Engineering

Pohl und Rupp erklären in [10], dass das Requirements Engineering (RE) dazu dient, Problemstellungen zu erkennen und zu verstehen. Aus diesen Problemstellungen können dann Anforderungen an das System abgeleitet werden. Eine Anforderung beschreibt eine Fähigkeit oder Eigenschaft an

[2] https://docs.microsoft.com/de-de/dotnet/core/rid-catalog

[3] https://www.sciencedirect.com/

[4] https://ieeexplore.ieee.org/Xplore/home.jsp

[5] https://dl.acm.org/

[6] https://link.springer.com/

das System. In diesem Artikel werden Anforderungen in funktionale (FA) und Nicht-Funktionale Anforderungen (NFA) eingeteilt. FA sind Anforderungen, die das System erfüllen soll. NFA zeigen hingegen Qualitätsmerkmale, die das System aufweisen muss. Für die Ermittlung von Anforderungen werden drei Verfahren aus dem RE angewandt. Die Methode des Interviews, Brainstorming und die Dokumentenanalyse. Die Erhebungstechniken dienen dazu Anforderungen durch Gespräche mit Stakeholdern herauszufinden oder aus bestehenden ähnlichen Systemen abzuleiten [10]. Für die Durchführung eines Interviews zur Erhebung von Anforderungen werden Stakeholder benötigt. Die Erwartungen der Stakeholder an das System werden dabei dokumentiert [8]. Für diesen Artikel wurden Vorgespräche mit zwei Professoren der Hochschule Reutlingen geführt. Die daraus resultierenden Ergebnisse sind in Form von Anforderungen dokumentiert, welche in Tabelle 6 und 7 dargestellt sind. Um Anforderungen durch Brainstorming zu erheben, werden verschiedene Lösungsansätze für das zu lösende Problem analysiert. Hierfür wird der Ist-Zustand analysiert, welcher in Kapitel 4.1 genauer beschrieben wird. Die daraus entstehenden Anforderungen werden anschließend dokumentiert. Für das Verfahren der Dokumentenanalyse werden bestehende Dokumente analysiert, welche sich mit der Thematik des zu lösenden Problems beschäftigen [8]. Hierfür wurden die bestehenden Systeme in Kapitel 3 analysiert.

2.2 Softwarebereitstellungsmodelle

In diesem Kapitel werden die verschiedenen Softwarebereitstellungsmodelle beschrieben. Nach Braun et al. [1] wird der Begriff *Cloud* als ein Dienst bezeichnet, welcher Rechenleistung über das Internet oder Intranet bereitstellt. Reinheimer erklärt in [11], dass bei der Public Cloud die verfügbaren Ressourcen individuell an die NutzerInnen der Infrastruktur verteilt werden. Die Anzahl der NutzerInnen ist somit limitiert. Mit der Private Cloud besitzen definierte AnwenderInnen den Zugriff und die Kontrolle auf eine Infrastruktur. Der Standort des Servers hängt von einem externen Dienstleister ab. Die Hybrid Cloud ist eine Kombination der Private und Public Cloud, in der Sicherheitsunkritische Prozesse an die Public Cloud ausgelagert werden.

Für eine lokale vor Ort Lösung kommen *On-Premises* Lösungen zum Einsatz. Der Begriff On-Premises definiert eine Infrastruktur, welche selbst gehostet und vor Ort verwendet wird. Dieser Lösungsansatz funktioniert wie die Private Cloud, jedoch wird dieses Bereitstellungsmodell in Eigenverantwortung betrieben. Die gesamte Infrastruktur sowie benötigte Software wird über die On-Premises Lösung bereitgestellt und selbst gehostet [7, 11].

2.3 Virtuelle Maschinen und Containerdienste

Goniwada erklärt in [5], dass bei der Virtualisierung die bereits bestehende physische Hardware verwendet wird. Eine virtuelle Maschine (VM) benötigt einen Hypervisor, welcher die physische Hardware als einen Ressourcenpool verwendet. Wie Abbildung 1 darstellt, können die auf der physischen Hardware ausgeführten VMs eigene Betriebssysteme hosten und Anwendungen isoliert ausführen. Cloudanwendungen sind aus Sicht der Datenverarbeitung Containerdienste. Ein Container enthält alle Binärdaten und Abhängigkeiten einer Anwendung, welches sich in ein Containerimage packen lässt. Dieses Containerimage kann hardwareunabhängig ausgeführt werden. Vorteilhaft gegenüber einem VM-Image ist die Paketgröße, da keine Kopie eines Betriebssystems benötigt wird. Somit fällt die Paketgröße kleiner aus als ein VM-Image und die Startzeit ist schneller, da keine Initialisierung des Betriebssystems benötigt wird. Die Containeranwendung wird auf der bereits verwendeten Hardware ausgeführt. Wie Abbildung 1 darstellt, teilen sich die Container

ebenso wie die VMs ihre Host Hardware. Laut Goniwada [5] ist die Verwendung von Containern die nächste Evolutionsstufe der Serveranwendungen. Das Forschungsunternehmen Gartner prognostiziert für 2022, das 75% der weltweiten Unternehmen Containeranwendungen einsetzen werden. Als Containerisierungstechnologie ist die Open-Source Software Docker[7] weit verbreitet. Durch Docker können Containerimages verwaltet und ausgeführt werden [5].

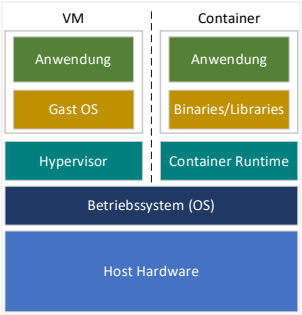


Abbildung 1: Virtuelle Maschine und Container. Eigene Darstellung nach [5]

3 Stand der Technik

In diesem Kapitel wird unter Berücksichtigung der Einschlusskriterien aus Tabelle 3 der aktuelle Stand der Technik dargestellt.

3.1 Webbasierte Programmierumgebungen

In dem Artikel von Tran et al. [14] wird eine webbasierte IDE im Lehrumfeld beschrieben. Diese webbasierte IDE ermöglicht es, jedem Studierenden die gleiche Lernumgebung zur Verfügung zu stellen. Die Lernumgebung wurde als *Software as a Service* (SaaS) implementiert und in einem privaten Netzwerk gehostet. SaaS bietet eine Bereitstellung der Dienste über das Internet an. Die AnwenderInnen dieser webbasierten IDE müssen somit keine Installation der Anwendung auf ihrer lokalen Maschine durchführen, sondern greifen über das Internet mit dem Browser auf die IDE zu. Das System bietet folgende Funktionsbausteine: ein Wiki, eine IDE, ein Kursmanagementsystem, ein Testingtool und ein Softwaredesigntool für Diagramme. Der Fokus in diesem Artikel liegt auf dem kollaborativen Arbeiten, um in einer Gruppe voneinander getrennten Studierenden Programmieraufgaben lösen zu können. Des Weiteren bietet dieses System eine Chatfunktion an, um sich mit den Studierenden innerhalb der Gruppe austauschen zu können. Der größte Vorteil des webbasierten Systems ist die Homogenität, da alle Studierenden das gleiche System nutzen. Die Studierenden bemängelten jedoch Verbindungsabbrüche zum Server.

In dem Artikel von Ding und Cao [3] wurde ein experimentelles E-Learning System mit dem Schwerpunkt auf dem kollaborativen Lernen entwickelt. Das System bietet folgende Funktionsbausteine an: eine webbasierte IDE, ein Forum, einen Chat und eine Filesharing Möglichkeit. Diese Implementierung richtet sich speziell an Personen, welche keine oder sehr wenig Programmierkenntnisse haben, um diese zu verbessern und Erfahrung beim Debugging zu sammeln. Meist wird in Gruppen von drei bis vier Personen gearbeitet, um Programmieraufgaben zu lösen. Über eine integrierte Audio- und Videoübertragung kann sich ein/e TutorIn remote bei benötigter Hilfe dazuschalten. Das Feedback der AnwenderInnen an diesem System zeigt, dass sowohl Studierende als auch Lehrende dieses System positiv wahrnehmen.

3.2 VM basierte Lernumgebungen für Programmieraufgaben

Suchodoletz et al. [13] beschreibt das VM basierte System bwLehrpool. Dieses System bietet eine digitale Infrastruktur für PC Räume, Labore und E-Klausuren. Das Ziel

[7] https://www.docker.com/

von bwLehrpool ist es, eine zentrale Infrastruktur aufzubauen, um virtuelle Laborumgebungen in Form von VMs bereitzustellen. Die VMs lassen sich auch außerhalb der Hochschule über das Internet durch den Browser[8] verwenden. Die Verwendung dieser VM basierten Umgebung soll den administrativen Aufwand in Laborumgebungen reduzieren. Die Computer in den PC-Räumen starten ihr Betriebssystem über die Netzwerkinfrastruktur mittels Ethernet[9]. Beim Start wird über Ethernet ein linuxbasiertes *Preboot Execution Environment* (PXE) durch einen Satellitserver bereitgestellt, welcher auf einer Benutzeroberfläche die Auswahl verfügbarer VM Images anzeigt. Der Satellitserver sowie die Infrastruktur von bwLehrpool ist in Abbildung 2 zu sehen. PXE hat sich nach Suchodoletz et al. [13] als Standard für das Booten über Ethernet bewiesen.

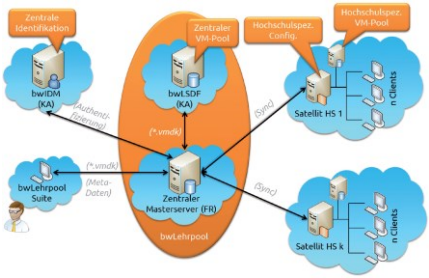


Abbildung 2: Infrastruktur von bwLehrpool aus [13].

Diese VM Images können für die gewünschte Arbeitsumgebung vorher konfiguriert und bereitgestellt werden. Nach der Auswahl des VM Images wird das zugehörige Betriebssystem gestartet [13]. Abbildung 3 stellt die verfügbaren Windows VM Images der Hochschule Reutlingen dar.

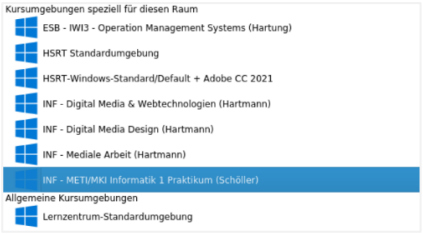


Abbildung 3: Verfügbare bwLehrpool Windows VM Images der Hochschule Reutlingen (Stand 13.03.2022), Bildschirmaufnahme[8]

Derzeit können Windows und Linux VMs verwendet werden. Die Verwendung von VM Images im Kontext von bwLehrpool benötigt keine lokale Installation auf dem Datenträger des Clientsystems im Informatiklabor. Für den Einsatz in Bildungseinrichtungen wird empfohlen, bwLehrpool auf einem Zentralserver auszuführen, welcher einen Pool an VMs bereitstellt. Der Vorteil einer VM basierten Umgebung ist die Hardwareunabhängigkeit. Bei Hardwareproblemen kann das Clientsystem ausgetauscht werden und beeinflusst nicht den Zugang zu bwLehrpool [13].

3.3 Übersicht über den Stand der Technik

Die Daten für die VM basierte Laborumgebung werden anhand eines ausgewählten VM Images von bwLehrpool erhoben. Das markierte VM Image in Abbildung 3 wird gestartet und untersucht, da sonst keine Funktionalität einer Programmierumgebung gegeben ist. Die folgende Tabelle 4 und 5 stellt den Funktionsumfang der webbasierten und der VM basierten Laborumgebung bwLehrpool aus Kapitel 3.2 zusammengefasst dar.

[8] https://rmz-bwl-guacamole.reutlingen-university.de/#/

[9] https://www.ieee802.org/3/

Tabelle 4: Zusammenfassung der Funktionsbausteine der webbasierten Laborumgebung

Feature	Web	VM
Wiki	✓	✗
Forum	✓	✗
Projektmanagement Tool	✓	✗
Lokale Installation auf Datenträger (Client)	✗	✗
Kursmanagementsystem	✓	✗
Audioübertragung	✓	✗

Tabelle 5: Zusammenfassung der Funktionsbausteine der VM basierten Laborumgebung

Feature	Web	VM
IDE	✓	✓
Versionskontrolle	✗	✗
Codeanalyse	✓	✓
File Sharing	✓	✓
Testing Tool	✓	✗
Softwaredesigntool	✓	✗
Durchführen eines Build-prozesses	✓	✓
Debugging	✓	✓

4 Konzeption einer On-Premises Lösung

In diesem Kapitel wird das Konzept einer Lernumgebung für Programmieraufgaben erarbeitet.

4.1 Ist-Analyse

Derzeit kann für das Informatik3-Praktikum des Studienganges Medien- und Kommunikationsinformatik die Laborumgebung bwLehrpool aus Kapitel 3.2 verwendet werden. Für die Durchführung des Praktikums können die Studierenden die Compiler und Interpreter der Programmiersprachen C, C++, C#, Python und Java auf einer bwLehrpool VM installieren. Die Voraussetzung für die Verwendung einer bwLehrpool Laborumgebung im Informatiklabor ist die Verfügbarkeit der PC-Poolräume. Für die Programmiersprachen C, C++ und C# wird aktuell die IDE Visual Studio[10] von Microsoft empfohlen, für Python PyCharm[11] und für Java die IDE Eclipse[12] empfohlen. Ebenso wird für die Versionskontrolle Git[13] als Client installiert. Für die Abgabe der Programmierprojekte besitzt die Hochschule Reutlingen bereits ein eigenes GitLab[14] Repository. Studierende können ihren eigenen Laptop mitbringen (BYOD), um die benötigte Software darauf zu installieren. Die Installation funktioniert nicht immer reibungslos, sodass individuelle Hilfestellung durch die DozentInnen notwendig ist. Hierfür wird meist das erste Praktikum benötigt. Diese Methode wird durch die Studierenden bevorzugt verwendet, da sie die Entwicklungsumgebung auch außerhalb der bwLehrpool Laborumgebung verwenden können. Um einen Zugriff auf die bwLehrpool VMs außerhalb des Hochschulnetzes zu ermöglichen, bietet bwLehrpool eine Remotefunktion an. Um die Remotefunktion zu verwenden, muss lediglich die Internetadresse[15] in den Browser eingegeben werden. Danach erfolgt eine Authentifizierung mit dem Hochschulkonto der

[10] https://visualstudio.microsoft.com/de/

[11] https://www.jetbrains.com/de-de/pycharm/

[12] https://www.eclipse.org/downloads/

[13] https://git-scm.com/

[14] https://gitlab.reutlingen-university.de/

[15] https://rmz-bwl-guacamole.reutlingen-university.de/#/

AnwenderIn. Ist die Authentifizierung erfolgt, kann ein verfügbares VM-Image wie bereits in Abbildung 3 dargestellt, ausgewählt werden. Abbildung 4 zeigt die Auswahl der bwLehrpool Laborumgebung über den Browser.

Abbildung 4: Auswahl einer verfügbaren bwLehrpool Lernumgebung, Bildschirmaufnahme[16]

4.2 *Systemidee und Ziele*

Die Systemidee hinter diesem Konzept ist es, den Studierenden eine homogene Lernumgebung für Programmieraufgaben hardwareunabhängig bereitzustellen. Das System soll für das Informatik Praktikum, für Informatik Lehrveranstaltungen und auf BYOD zum Einsatz kommen in denen bwLehrpool nicht zu Verfügung steht. Eine Abgabe der Programmieraufgaben erfolgt über Git.

4.3 *Identifizierung der Stakeholder*

Stakeholder sind Personen oder Organisationen, die ein Interesse an dem System haben [10]. Die identifizierten Stakeholder sind die Studierenden und die DozentInnen der Fakultät Informatik.

4.4 *Anforderungen*

Für die Formulierung der Anforderungen wurde die Satzschablone von Pohl und Rupp [10] wie Abbildung 5 darstellt, verwendet. Satzschablonen dienen zur einheitlichen Formulierung der Anforderungen und bieten ein leichteres Verständnis.

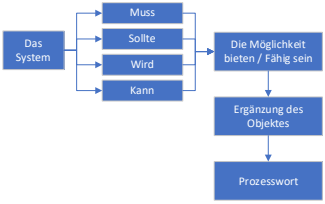


Abbildung 5: Satzschablone der Anforderungen in Anlehnung an [10]

Die erhobenen Anforderungen sind in Tabelle 6 und 7 aufgelistet.

Tabelle 6: Funktionalen Anforderungen

ID	Funktionale Anforderung
FA1	Das System muss fähig sein, Programmiersprachen für C, C++, C#, Java und Python zu unterstützen.
FA2	Das System muss fähig sein, einen Buildprozess durchführen zu können.
FA3	Das System muss über eine IDE verfügen.
FA4	Das System muss fähig sein, Programmierübungen mit Git abgeben zu können.
FA5	Das System muss fähig sein, ein Git Repository zu klonen.
FA6	Das System muss fähig sein, debuggen von Code zu ermöglichen.

Tabelle 7: Nicht-Funktionalen Anforderungen

ID	Nicht-Funktionale Anforderung
NFA1	Das System muss die Möglichkeit bieten, über den

[16] https://rmz-bwl-guacamole.reutlingen-university.de/#/

	Webbrowser erreichbar zu sein.
NFA2	Das System muss mit Windows, MacOS und Ubuntu Desktop kompatibel sein.
NFA3	Das System sollte einen geringen Installationsaufwand besitzen.
NFA4	Das System sollte intuitiv bedienbar sein.
NFA5	Das System sollte fähig sein, hardwareunabhängig zu funktionieren.
NFA6	Das System muss auf der lokalen Maschine der AnwenderInnen installierbar sein.
NFA7	Das System muss fähig sein, Dateien persistent zu speichern.

4.5 Konzept

Durch die Anforderungserhebung mit den genannten RE Methoden aus Kapitel 2.1 konnten Anforderungen formuliert und in Tabelle 6 und 7 dokumentiert werden. Die Designentscheidung der Lernumgebung für Programmieraufgaben fiel auf eine On-Premises Lösung, da die Lernumgebung auf den BYOD der Studierenden und DozentInnen ausführbar sein soll. Auf der lokalen Maschine kommt eine Containerlösung zum Einsatz, da wie in Kapitel 2.3 beschrieben, der Einsatz von Containern die bessere Lösung für das Bereitstellen von einzelnen Anwendungen ist. In dem Artikel von Combe et al. [2] werden verschiedene Containerlösungen betrachtet. Die Docker Engine[17] eignet sich besonders für das Bereitstellen von einzelnen Anwendungen, so wie es für dieses Konzept benötigt wird. Für Windows, MacOS und Ubuntu Desktop ist die Docker Engine verfügbar und deckt die benötigten Betriebssysteme von Tabelle 2 ab. Die Lernumgebung für Programmieraufgaben wird als Docker Containerimage bereitgestellt und auf der lokalen Maschine der AnwenderInnen ausgeführt. Die Nutzung von Docker ist eine Voraussetzung für die Verwendung dieses Konzepts.

Linuxserver.io[18] bietet die weitverbreitete Open Source IDE *vscode*[19] von Microsoft als Containerimage an. Diese IDE unterstützt die benötigten Programmiersprachen aus Tabelle 6 FA1. Aus diesem Grund wird dieses Containerimage verwendet, da es kontinuierlich Aktualisierungen erhält und den eigenen Wartungsaufwand durch das Aktualisieren neuer Softwareversionen verringert. Ebenfalls kann die IDE durch Erweiterungen erweitert werden, da sie einen Zugriff auf den OpenVSX[20] Marketplace besitzt. Dieser Zugriff ist notwendig, um die Debugger der benötigten Programmiersprachen aus Tabelle 6 zu installieren. Ein Git Client ist bereits in vscode integriert. Nach Singh [12] wird für die Ausführung von Containerimages unter Windows, die Aktivierung von *Windows Subsystem for Linux version 2* (WSL2) empfohlen, welches sich in den Windows-Features oder über die Kommandozeile aktivieren lässt. WSL2 bietet eine 2-3-Mal schnellere Input-Output (IO) Performance als die Hypervisor-basierte (Hyper-V) Virtualisierungslösung unter Windows. Ebenfalls empfiehlt Docker[21] die Verwendung von WSL2.

Aus diesem Grund ist das WSL2 Feature unter Windows eine Voraussetzung für die Verwendung für dieses Konzept. Abbildung

[17] https://www.docker.com/products/docker-desktop/

[18] https://docs.linuxserver.io/images/docker-code-server

[19] https://github.com/microsoft/vscode

[20] https://open-vsx.org/

[21] https://docs.docker.com/desktop/windows/wsl/

6 stellt das Deployment Diagramm der Lernumgebung für Programmieraufgaben dar.

Die AnwenderInnen starten die Anwendung mittels dem Docker Befehl *docker-compose up* auf ihrer lokalen Maschine und installieren das Containerimage. Der gestartete Docker Deamon dient dazu, das Containerimage auszuführen. Nachdem der Container gestartet ist, erfolgt der Zugriff auf die Programmierumgebung über den Browser. Des Weiteren besitzt der Container Zugriff auf das lokale Dateisystem [6]. Nach Xu et al. [16] nutzt das Containerimage einen eigenen abgegrenzten Bereich, welcher als *Docker Data Volume* bezeichnet wird. Daten, die auf diesem Containerspeicher gespeichert werden, bleiben auch nach dem Neustart des Containers bestehen. Für dieses Konzept ist es notwendig, alle Abhängigkeiten der benötigten Programmiersprachen sowie lokale Programmierprojekte innerhalb dieses Containerspeichers abzuspeichern, um die Daten zu persistieren.

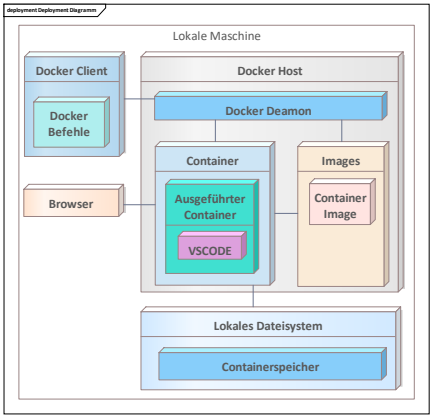


Abbildung 6: Deployment Diagramm der Lernumgebung für Programmieraufgaben, Erweiterung von [6]

5 Evaluation

Für die Evaluation der Anforderungen wurde nach Pohl und Rupp [10] die Prüfung durch einen Prototyp gewählt. Bei dem Prototyp handelt es sich um einen evolutionären Prototyp, welcher das Ziel hat, verbessert zu werden. Für die Durchführung wurden zwei Stakeholder akquiriert, welche das Betriebssystem Windows 11 und Ubuntu Desktop 20.04 verwenden, um die Hardwareunabhängigkeit zu überprüfen. Nach einer Unterweisung wurden die Stakeholder mit den nötigen Informationen versorgt und auf die Grenzen des Prototyps hingewiesen. Aktuell unterstützt der Prototyp nur die Programmiersprache C#. Für die Prüfung wurden folgende Prüfszenarien gewählt, welche in Tabelle 8 aufgelistet sind. Anschließen haben die Stakeholder die Möglichkeit, den Prototyp experimentell zu testen und Feedback zu geben.

Tabelle 8: Auflistung der Prüfszenarien

ID	Prüfszenario
P1	Erstellen Sie ein C# Projekt mit einem Beispielprogramm
P2	Führen Sie mit der erstellten Datei einen Buildprozess durch
P3	Setzen Sie einen Breakpoint und nutzen den Debugger
P4	Klonen Sie ein Repository aus GitLab
P5	Ändern Sie den Quellcode und nutzen Sie die Git Commit Funktion. Laden Sie anschließend den neuen Quellcode in das Repository

Tullis und Albert [15] empfehlen die Bewertung von Prüfszenarien in Binärform. Ein positives Ergebnis erhält den Wert 1 und ein negatives 0. Anschließend wird aus den Ergebnissen die *Task Success Rate* (TSR) berechnet. Für die Berechnung der TSR wird die Anzahl der erfolgreichen Prüfszenarien durch die Gesamtzahl der Prüfszenarien dividiert. Das Ergebnis wird dann mit 100 multipliziert.

Die Testdauer beträgt 40-60 Minuten. Nach Pohl und Rupp [10] soll ein Prüfszenario eigenständig gelöst werden. Das Prüfszenario

gilt somit als nicht bestanden, wenn Hilfe durch den Prüfer oder der Prüferin erforderlich ist. Die Ergebnisse der Prüfszenarien sind in Tabelle 9 aufgelistet.

Tabelle 9: Ergebnisse der Prüfszenarien

	P1	P2	P3	P4	P5
ProbandIn 1	0	1	1	1	1
ProbandIn 2	1	1	1	1	1

Die TSR aus Tabelle 9 ergibt 90% und zeigt, dass der Großteil der gestellten Aufgaben zu lösen ist. Das Feedback der ProbandInnen ergab folgende Verbesserungsvorschläge:

- Bereitstellung von Anleitungen und Tipps mit dem Umgang von vscode und Docker
- Verbesserung des Installationsprozesses durch fertige Skripte, um ein händisches Ausführen über die Kommandozeile zu vermeiden
- Ein Tool zum Modellieren von Unified Modeling Language (UML) Diagrammen wäre von Vorteil

Die Benutzeroberfläche wurde von den ProbandInnen als positiv und intuitiv empfunden. Ein Ausschnitt der Benutzeroberfläche zeigt in Abbildung 7 den Dateiexplorer sowie weitere Interaktionsmöglichkeiten an.

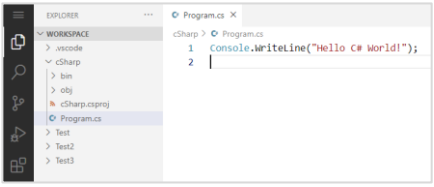


Abbildung 7: Dateiexplorer des Prototyps

Da die ProbandInnen Windows 11 und Ubuntu Desktop 20.04 als Betriebssystem verwenden, konnte sichergestellt werden, dass der Prototyp hardwareunabhängig ausführbar ist. Ebenso ist ein Buildprozess sowie das Debuggen von Quellcode in C# durchführbar. Der integrierte Git Client ermöglicht die Abgabe von Programmieraufgaben in ein GitLab Repository. Die Evaluation zeigt, dass die Anforderungen Erfüllbar sind. Da es sich um einen evolutionären Prototyp handelt, muss in der nächsten Iteration die Unterstützung der Programmiersprachen C, C++, Python und Java gewährleistet und das Feedback der ProbandInnen eingearbeitet werden. Das in Kapitel 4.5 beschriebene Docker Data Volume muss für eine vereinfachte Installation exportierbar und importierbar sein, um Konfigurationsdateien für die AnwenderInnen bereitzustellen. Durch den Import eines Docker Data Volumes in das installierte Containerimage, wird der Installationsaufwand deutlich verringert.

6 Diskussion und Fazit

Die Literaturrecherche zeigt ähnliche bestehende Systeme mit unterschiedlichen Softwarebereitstellungsmethoden und Funktionsbausteinen. Durch die angewandten Methoden des RE wie Interviews, Brainstorming und Dokumentanalysen wurden Anforderungen erhoben, um das Problem der fehlenden einheitlichen Lernumgebung für Programmieraufgaben zu lösen, welche BYOD unterstützt. Bei der Suche nach ähnlichen verwendeten Lernumgebungen für Programmieraufgaben innerhalb der Lehre hat sich herausgestellt, dass webbasierte und VM basierte Systeme zum Einsatz kommen. Die Ist-Analyse zeigt, dass die Studierenden ihre Software auf dem eigenen BYOD bevorzugt verwenden. Somit eignet sich eine On-Premises Lösung, um das Konzept der Lernumgebung für Programmieraufgaben bereitzustellen. Um eine On-Premises Lösung hardwareunabhängig einsetzen zu können, eignen sich Containerisierungstechnologien. Aus diesem Grund wurde für das Konzept eine Containeranwendung gewählt, welche als Containerdienst die Software Docker verwendet. Der Container wird auf dem Hostsystem ausgeführt und ermöglicht einen Zugriff auf die Lernumgebung für Program-

mieraufgaben über den Browser. Die Evaluation durch einen Prototyp zeigt, dass ein Buildprozess sowie das Debuggen von Quellcode und eine Abgabe der Programmieraufgabe mittels Git durchführbar ist. Die gestellten Prüfszenarien konnten zu 90% erfüllt werden. Der Prototyp wurde auf einem Windows- und Linux basiertem Betriebssystem erfolgreich ausgeführt und beweist die Hardwareunabhängigkeit für diese Plattformen. Das Feedback der ProbandInnen zeigt neue Erkenntnisse, die für die nächste Iteration des Prototyps einen Mehrwert bieten. Demnach werden Anleitungen zur Verwendung des Prototyps und Docker benötigt. Ebenso soll der Installationsprozess automatisiert werden, indem fertige Skripte ausgeführt werden sollen.

7 Literaturverzeichnis

[1] Baun, C., Kunze, M., und Ludwig, T. 2009. Servervirtualisierung. *Informatik Spektrum* 32, 3, 197–205. DOI: 10.1007/s00287-008-0321-6.

[2] Combe, T., Martin, A., und Di Pietro, R. 2016. To Docker or Not to Docker: A Security Perspective. *IEEE Cloud Comput.* 3, 5, 54–62. DOI: 10.1109/MCC.2016.100.

[3] Ding, Q. und Cao, S. 2017. RECT: A Cloud-Based Learning Tool for Graduate Software Engineering Practice Courses With Remote Tutor Support. *IEEE Access* 5, 2262–2271. DOI: 10.1109/ACCESS.2017.2664070.

[4] Ehlers, U.-D. 2011. *Qualität im E-Learning aus Lernersicht*. Medienbildung und Gesellschaft 15. VS Verlag für Sozialwissenschaften, Wiesbaden. 978-3-531-17589-8.

[5] Goniwada, S. R. 2022. *Cloud Native Architecture and Design. A Handbook for Modern Day Architecture and Design with Enterprise-Grade Examples*. Springer eBook Collection. Apress; Imprint Apress, Berkeley, CA. 978-1-4842-7225-1. DOI: 10.1007/978-1-4842-7226-8.

[6] Jaramillo, D., Nguyen, D. V., und Smart, R. 2016. Leveraging microservices architecture by using Docker technology. In *SoutheastCon 2016*. IEEE, 1–5. 978-1-5090-2246-5. DOI: 10.1109/SECON.2016.7506647.

[7] Mc Evoy, G. V., Schulze, B., und Garcia, E. L. M. 2009. Performance and deployment evaluation of a parallel application in an on-premises Cloud environment. In *Proceedings of the 7th International Workshop on Middleware for Grids, Clouds and e-Science - MGC '09*. ACM Press, New York, New York, USA, 1–6. 9781605588476. DOI: 10.1145/1657120.1657122.

[8] Niebisch, T. 2013. *Anforderungsmanagement in sieben Tagen*. Springer Berlin Heidelberg, Berlin, Heidelberg. 978-3-642-34856-3. DOI: 10.1007/978-3-642-34857-0.

[9] Pilotto, L. M. 2021. *Blended Learning*. Springer Fachmedien Wiesbaden, Wiesbaden. 978-3-658-31234-3. DOI: 10.1007/978-3-658-31235-0.

[10] Pohl, K. und Rupp, C. 2021. *Basiswissen Requirements Engineering. Aus- und Weiterbildung nach IREB-Standard zum Certified Professional for Requirements Engineering Foundation Level*. dpunkt.verlag, Heidelberg. 978-3-8649-0814-9.

[11] Reinheimer, S., Ed. 2018. *Cloud Computing. Die Infrastruktur der Digitalisierung*. Edition HMD. Springer Vieweg, Wiesbaden. 978-3-658-20966-7. DOI: 10.1007/978-3-658-20967-4.

[12] Singh, P. 2020. Exploring WSL2. In *Learn Windows Subsystem for Linux*, P. Singh, Ed. Apress, Berkeley, CA, 75–98. 978-1-4842-6037-1. DOI: 10.1007/978-1-4842-6038-8_5#DOI.

[13] Suchodoletz, D. von, Münchenberg, J., Honigberger, W., Rettberg, S., Wilson, M., und Lohr, M. 2014. bwLehrpool – ein landesweiter Dienst für die Bereitstellung von PC-Pools in virtualisierter Umgebung für Lehre und Forschung. *PIK - Praxis der Informationsverarbeitung und Kommunikation* 37, 1. DOI: 10.1515/pik-2013-0046.

[14] Tran, H. T., Dang, H. H., Do, K. N., Tran, T. D., und Nguyen, V. 2013. An interactive Web-based IDE towards teaching and learning in programming courses. In *Proceedings of 2013 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)*. IEEE, 439–444. 978-1-4673-6355-6. DOI: 10.1109/TALE.2013.6654478.

[15] Tullis, T. und Albert, B. 2013. *Measuring the user experience*. *Collecting, analyzing, and presenting usability metrics*. Elsevier/MK Morgan Kaufmann, Amsterdam, Boston, Heidelberg. 978-0-1241-5781-1.

[16] Xu, Q., Awasthi, M., Malladi, K. T., Bhimani, J., Yang, J., und Annavaram, M. 2017. Docker characterization on high performance SSDs. In *ISPASS 2017. IEEE International Symposium on Performance Analysis of Systems and Software : April 24-25, 2017, Santa Rosa, California*. IEEE, Piscataway, NJ, 133–134. 978-1-5386-3890-3. DOI: 10.1109/ISPASS.2017.7975282.

©2022 Ronell Bresler. 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/

Taxonomie- und Entwicklungsvorschlag für Serious Games

Tobias Lauxmann
Hochschule Reutlingen
Reutlingen, Deutschland
Tobias_Holger.Lauxmann@Student.Reutlingen-University.de

Zusammenfassung

Serious Games unterhalten und erfüllen pädagogische Ziele gleichzeitig. Um dies langfristig mit Erfolg umzusetzen ist ein Verständnis in Verhaltenspsychologie in Spielen, dem Lernen und eine ausführliche Planung essenziell. Dieses Grundverständnis wird in dieser Arbeit vermittelt. Dafür wurde eine umfangreiche Literaturrecherche und Analyse durchgeführt. Daraus resultierend wird eine Definition und die Eigenschaften eines Serious Games vorgestellt. Auf die grundlegenden Eigenschaften folgen Klassifizierungsmöglichkeiten, welche die Kategorisierung von unterschiedlichen Serious Games erleichtert. Abschließend wird durch ein entwickeltes Vorgehensmodell, eine detaillierte Handlungsempfehlung für das Design und die Entwicklung eines Serious Games vorgeschlagen.

CCS Concepts

• **Applied computing** → Interactive learning environments.

Keywords

serious games, educational games, learning in games, game design

1 Einleitung

Der pädagogische Ansatz *lernen durch Spielen* besitzt viel Potenzial, um Wissen zu vermitteln. Vor allem in Zeiten von wachsender Multimedianutzung bei Kindern, kann ein solcher Ansatz in digitaler Form äußerst attraktiv sein. Leider scheitert der gewünschte Lernerfolg bereits oft an der Umsetzung eines Spiels. Der Hauptzweck eines Spiels ist es seine Spieler zu unterhalten. Ein Spiel das einem anderen Zweck, zum Beispiel einem Lernzweck verfolgt wird Serious Game (SG) genannt [1][12, S.3]. Die Schwierigkeit bei der Umsetzung liegt darin, eine sinnvolle Balance zwischen der Unterhaltung und dem eigentlichen Ziel zu erschaffen [7]. Macht das Spiel zu viel Spaß, dann ignoriert der Spieler das gewünschte Lernziel. Besitzt das Spiel zu viele Lernelemente und wenig Unterhaltung, dann wird es schnell langweilig. Die Erkenntnis eines Ungleichgewichts im Spiel sollte nicht erst nach der Entwicklung auftreten. Es muss sich bereits im Spieldesign damit beschäftigt werden. Bevor in dieser Arbeit eine solche Vorgehensweisen präsentiert wird, werden die Kernelemente eines SG, in Form einer Definition und weiterer Eigenschaften erarbeitet.

2 Definiton

Die Definition für ein SG wurde bereits im Jahre 1968 durch das Buch Serious Games von Clark C. Abt geprägt [1]. In diesem sagt er, dass SGs Spiele sind, die einen ausdrücklichen und sorgfältig durchdachten Bildungszweck haben und nicht in erster Linie zum Vergnügen gespielt werden sollen. Diese noch

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

Informatics Inside Frühjahr 2022
11. Mai 2022, Hochschule Reutlingen
Copyright 2022 Tobias Lauxmann

vor der Ära der Informationstechnologie entstandene Definition, die sich ursprünglich auf Brett und Kartenspiele bezog, hat sich bis heute nur wenig verändert. Die Grundidee ist geblieben, doch eine Verschiebung in Richtung digitaler Spiele hat stattgefunden. Einige aktuellere Definitionen beziehen sich deswegen ausschließlich auf digitale Spiele [12, S.3]. Unabhängig von der Art der Spiele, herrscht Einigkeit darüber, dass bei einem SG neben einem unterhaltenden Faktor, eine Lern- oder verhaltensändernde Komponente im Vordergrund steht. In [12, S.3] wird ein SG als ein digitales Spiel definiert, das mit der Absicht entwickelt wird, zu unterhalten und mindestens ein zusätzliches charakterisierendes Ziel zu erreichen. Dieses Ziel hängt oft mit dem Anwendungsbereich des SG zusammen. So ermutigen beispielsweise Exergames aus dem Bereich Gesundheit, die Menschen, körperlich aktiv zu werden. SGs aus dem Bereich Bildung haben meistens das Ziel, dem Benutzer bestimmten Lernstoff zu vermitteln [12, S.4]. Serious Games zielen darauf ab, dieses charakteristische Ziel zu erreichen, ohne die Spielerfahrung zu beeinträchtigen [7]. Der Spieler sollte also immer das Gefühl des Spielens wahrnehmen. Dabei soll er dazu motiviert werden, das charakteristische Ziel durch geeignete Methoden, ein ansprechendes Spielerlebnis und durch den Einsatz geeigneter Interaktionstechnologie zu erreichen [7]. Je nach Anwendungsgebiet und Ziel des SG, können die Interaktionsmöglichkeiten sehr unterschiedlich ausfallen. Zum Beispiel kann körperliche Aktivität mit einem Gerät zur Bewegungserfassung effektiv erreicht und gemessen werden. Dabei muss darauf geachtet werden, dass genügend Raum zur Verfügung steht. Die Umgebungsvoraussetzung ist dementsprechend eine andere, wie von einem SG, das mit Maus und Tastatur bedient wird. Ein SG muss also unter Beachtung vieler Faktoren sorgfältig geplant und umgesetzt werden.

Im Zusammenhang mit SG treten oft die Begriffe Gamification und Game-based learning (GBL) auf. Diese sind jedoch von SGs abzugrenzen. Gamification kann verstanden werden, als die Verwendung von spielerischen Elementen in einem nicht spielerischen Kontext [10]. Die Idee dabei ist es die Motivation in einem Bereich anzuregen, der von der Großzahl der Benutzter als langweilig oder monoton empfunden wird. Eine weitere Definition beschreibt Gamification als Prozess, in dem ein Produkt oder Dienst mit spielerischen Erfahrungen angereichert wird, um die allgemeine Wertschöpfung des Nutzers zu unterstützen [14]. Die Definitionen machen deutlich, das Gamification eine extra Schicht darstellt, die in meist bereits bestehenden Produkten eingesetzt wird.
Der Begriff von GBL tritt oft im Zusammenhang mit SGs auf. Er beschreibt den Prozess des Lernens mittels eines Spiels [13]. Dabei muss dies nicht zwingend ein digitales Spiel sein. GBL kann auch auf Spiele angewendet werden, die ursprünglich nicht als Lernspiel entstanden sind, wie beispielsweise Schach. SG basieren auf dem Prozess des GBL, werden aber speziell für den Lern- oder pädagogischen Zweck entwickelt.

3 Taxonomie

Die Eigenschaften von SGs sind sehr vielfältig. Benutzer aber auch Entwickler von SGs benötigen ein gemeinsames Format, um SGs effektiv klassifizieren zu können. Es gibt bereits einige Vorschläge für die Klassifizierung von SGs mittels einer Taxonomie. Beispielweise stellen Laamarti et al. in An Overview of Serious Games folgende Oberkategorien vor: Anwendungsbereich, Aktivität, Modalität, Interaktionsart und Umgebung [16]. Diese Einteilung erscheint sinnvoll, da der Anwendungsbereich und die Aktivität grundlegende Eigenschaften eines SG beschreiben. Anzumerken ist jedoch, dass eine Information über den Hauptzweck des SG fehlt und

diese nur über die Oberkategorien interpretiert werden kann. Auch die Zielgruppe bleibt unbekannt. Ein Ansatz, der dies genauer beschreibt, ist das Standarddokument DIN SPEC 91380 [11]. Darüber können Informationen über ein SG systematisch und umfangreich beschrieben werden. Insgesamt umfasst das sogenannte Serious Games Metadata Format sechs obligatorische und drei optionale Oberkategorien. Die Oberkategorien und deren Beschreibung sind in Tabelle 1 abgebildet. Die Oberkategorien sind mit den Unterkategorien über eine Baumstruktur verknüpft und bilden zusammen den Kern (Level 1) und die Details (Level 2) einer Beschreibung. Außerdem existiert eine weitere Stufe für spezifische Anwendungsbereiche. Als Beispiel werden erweiterte Beschreibungen und Attributlisten für Lernspiele und Gesundheitsspiele präsentiert.
Die in der DIN SPEC 91380 beschriebene Struktur bietet die Möglichkeit ein SG standardisiert zu beschreiben. Dies ist für Anbieter, als Beschreibung und für Anwender als Suche äußerst hilfreich, weil Spiele mit speziellen Eigenschaften dadurch schnell zu finden sind. Da die Kern- und Detailkategorien zusammen jedoch sehr umfangreich sind, können für die Beschreibung eines SG auch ausschließlich die Kernelemente verwendet werden. Das wohl wichtigste Kernelement ist das charakterisierende Ziel. Dieses kann durch ein Attribut aus einer vordefinierten Liste gewählt werden. Außerdem kann dieselbe Attributliste genutzt werden, um zusätzliche Ziele zu ergänzen. Jede Zielsetzung kann nochmals mit Hilfe einer erweiterten Attributliste spezifiziert werden, die aus Subjekten, Adverbien und Verben besteht. Ein Beispiel für eine Zielsetzung ist: Trainings- und Simulationsumgebung/-spiel. Diese kann durch die Beschreibung: Fachkompetenzen des Spielers gezielt erproben, ergänzt werden. Dadurch wird unmissverständlich klar, dass es sich um ein SG handelt, das bereits einen fortgeschrittenen Wissensstand des Spielers erfordert.
Die grundlegendsten Elemente sind in Tabelle 2 zusammengefasst. Dabei sind die Elemente aus der Kernkategorie der DIN SPEC 91380 entnommen und wurden um den Punkt Interaktionsstil ergänzt. Es ist wichtig festzuhalten, dass es bisher noch keine allgemeingültige Norm zur Klassifizierung für SG gibt. Dabei ist es äußerst sinnvoll einen gemeinsamen Standard in diesem Bereich zu haben, um den zukünftigen Einsatz von SG, in bildungs- oder therapeutischen Einrichtungen zu unterstützen. Die dargestellte Taxonomie in Tabelle 2 liefern eine sehr komprimierte Zusammenfassung, für einen schnellen Überblick. Die Taxonomie, der DIN SPEC 91380 bietet hingegen umfangreiche Lösung, um SGs zu klassifizieren.

3.1 Weitere Eigenschaften

Für die Entwicklung eines SG gibt es grundlegende Überlegungen. Zunächst werden die Eigenschaften von exogenen und endogenen Spielen nach Thomas W. Malone und Mark R. Lepper unterschieden [24]. Exogene Spiele sind bereits existierende Spiele, in deren regelbasiertes System Lerninhalte eingebaut werden. Das bedeutet die Spielmechaniken existieren bereits. Ein einfaches Beispiel ist die Verwendung des Spiels hangman in Kombination mit Vokabeln einer zu lernenden Sprache. Der Lerninhalt beschränkt sich auf Fakten, die durch Wiederholung erlernt werden.
Im Gegensatz dazu verbinden endogene Spiele das Spieldesign mit dem Lerninhalt, damit dieser in der Struktur des Spiels integriert ist. Es kann sich an der Spielidee von existierenden Spielen orientiert werden, die Integration des Lerninhalts erfordert aber eine komplett neue Entwicklung.
Eine weitere grundlegende Eigenschaft, die den Lerninhalt betrifft ist das Einbinden des charakterisierenden Ziels in dynamischer oder

Beschreibungselement	Erklärung
SpielInfo	Allgemeine Informationen zum Spiel wie Titel und Zielsetzung
NutzungInfo	Nutzunginformationen wie Anwendungsbereich und Zielgruppe
DistributionInfo	Distributionskanäle inklusive Verkaufspreis
TechnikInfo	Verfügbarkeit auf Plattformen und technische Anforderungen
BeteiligteInfo	Bei der Entstehung und Verbreitung des Serious Game beteiligte Einrichtungen
MetadatenInfo	Angaben zu den Metadaten
QualitätsInfo	Verliehene Preise oder Auszeichnungen, Rezessionen, wissenschaftliche Studien
WeitereInfo	Optionale Informationen wie verwandte Serious Games oder Entwicklungskosten
Gesamtnote	Optionales Qualitätsmerkmal

Tabelle 1: Oberkategorien des Serious Games Metadata Format [11]

statischer Form [18]. Bei einer statischen Entwicklung wird der Lerninhalt in die Spielwelt integriert und beispielsweise mit der Erzählung des Spiels verbunden. Der Lerninhalt ist nicht mehr durch einen anderen Inhalt ersetzbar.
Eine dynamische Implementierung eines SG besitzt im Vergleich dazu, mehr Spielraum. Nach der Entwicklung kann der Lerninhalt des Spiels durch einen Anderen ausgetauscht werden. Dies ist nur möglich, wenn die Spielwelt getrennt vom Inhalt ist. Eine Gefahr dabei ist, dass das Spiel unharmonisch wirkt, da der Lerninhalt wenig bis gar keinen Bezug zum Spiel selbst hat.

4 Design- und Entwicklungsvorschlag

In diesem Kapitel wird zunächst eine klassische Methode für die Entwicklung von Spielen mittels des Mechanics, Dynamics und Aesthetics (MDA) Modells vorgestellt. Da dieses seinen Fokus auf die Unterhaltung und nicht auf ein charakterisierendes Ziel legt, wird für das Design und die Entwicklung von SGs ein anderes Verfahren benötigt. Das Design, Play, Experience (DPE) Modell beschreibt einen solchen Ansatz und wird aus diesem Grund ebenfalls kurz beschrieben. Daraus abgewandelt wird eine eigene Handlungsempfehlung für das Design und die Entwicklung von SGs vorgestellt.

4.1 MDA Modell

Für die Entwicklung von digitalen Spielen gibt es bereits mehrere Konzepte und Vorgehensweisen. Diese fokussieren sich auf den Spielfaktor und versuchen demnach möglichst viel Spielspaß zu generieren. Ein klassisches Beispiel für die Entwicklung von Spielen, ist das MDA Modell von Hunicke et al. [21]. Dieses beschreibt einen Spielprozess als aufgestellte Regeln, welche in einem System abgebildet werden, das gespielt zu Spaß führt. Parallel zu den Elementen: Spielprozess, Regeln und Spaß stehen die Komponenten Mechanics, Dynamics und Aesthetics, die den Designprozess beschreiben. Mechanics, sind

Titel	Beschreibung
Titel	Der Titel des SG.
Charakterisierendes Ziel	Eine kurze Beschreibung des Hauptziels des SG. Welche Lern-, Trainingseffekte oder therapeutischen Zwecke werden angestrebt?
Zielbeschreibung	Eine Beschreibung wie das Hauptziel erreicht wird.
Anwendungsbereich	Der oder die Anwendungsbereiche wie z.B. Medizin, Wirtschaft oder Bildung.
Zielgruppe	Hier können sowohl Spielvermittler wie Eltern, Lehrer, Therapeuten, als auch Spielergruppen, wie Schüler, Studenten genannt werden.
Spielzeit	Die durchschnittliche Zeit, die benötigt wird, um das Ende des Spiels zu erreichen. Optional bietet sich eine geschätzte Spielzeit zum Erreichen eines bestimmten Lern- oder Trainingseffekt an.
Spielmodus	Verschiedene Spielmodi wie: Einzelspielermodus, Lokaler Multiplayermodus, Online-Multiplayermodus; Unterscheidung bei Multiplayer zwischen kooperativ und kompetitiv.
Genre	Das Spielgenre wie beispielsweise Abenteuer, Sportsimulation, Logik- und Denkspiel.
Sprache	Die Sprache(n) in denen das SG gespielt werden kann.
Plattform	Nennt die Plattform(en) auf denen das SG verfügbar ist.
Interaktionsstil	Das Medium, mit dem welchen mit dem SG interagiert wird. Beispiel: Tastatur und Maus, Playstation-Controller, Eye Tracker.

Tabelle 2: Wesentliche Klassifizierungsmerkmale eines Serious Game in Anlehnung an [11]

Grundelemente auf Daten- und Algorithmenebene. Dynamics beschreibt deren Ausführung im Zusammenhang mit der Spielinteraktion. Aesthetics sind die erwünschten emotionalen Reaktionen, die während des Spiels hervorgerufen werden. Kleine Änderungen in einer Ebene, können große Auswirkungen in einer anderen erzeugen. Deswegen achtet das MDA Modell darauf, dass ein Spiel stets von der Design- als auch von der Spielerperspektive betrachtet wird. Vor allem letzteres ist ein sinnvoller Ausgangspunkt, um ein Spiel zu analysieren. Um die erwünschte Spielererfahrung beschreiben zu können wird eine beispielhafte Taxonomie vorgeschlagen. Die darin beschriebenen Begriffe können genutzt werden, um gewünschte Emotionen in Aestetics zu beschreiben. Daraus können gezielt Spielprinzipien im Teil Dynamics und folglich deren Umsetzung im Teil Mechanics entwickelt werden. Ein iterativer Prozess ist notwendig, um zu überprüfen, ob die umgesetzten Mechanics, die gewünschte Wirkung erzeugen.

4.2 DPE Modell

Designvorschläge für SGs versuchen den Hauptfokus auf den pädagogischen Zweck, anstatt auf den Unterhaltungsfaktor zu legen. Ein Modell, das in diesem Zusammenhang

entstanden ist, ist das Design, Play, und Experience Modell, von Brian M. Winn [6]. Neben der Zielsetzung eine Balance zwischen Unterhaltung und pädagogischen Zweck zu schaffen, kritisiert Winn am MDA Modell, dass keine Aspekte von Game Design berücksichtigt werden, die sich über das Gameplay hinaus beschäftigen. Darunter fallen Aspekte wie Storytelling, User Experience und der Einfluss der Technologie auf das Design.
Analog zum MDA Modell, existiert im DPE Modell das gleiche Verhältnis von Spieldesigner zu Spieler. Dies geschieht über die Komponenten Design, Play und Experience. Im Design werden Ziele festgelegt und umgesetzt, welche während der Spielphase (Play) beim Spieler für Erfahrungen (Experience) sorgen. Der Designprozess soll iterativ stattfinden, um die erwünschten Erfahrungen zu testen, aber auch um das Ziel zu prüfen, das im Design festgelegt wird. Es wird außerdem darauf hingewiesen, dass die Zielgruppe eine bedeutende Rolle spielt. Der Grund dafür ist, dass die Spielerfahrung, die während dem Spielen entsteht, neben dem Design, auch stark von den kognitiven, sozialen, kulturellen und den persönlichen Erfahrungen des Spielers abhängig ist.
Das DPE Modell besteht weiterhin aus vier aufeinanderliegenden Schichten, welche alle in die Design, Play und Experience Phasen gegliedert sind. Zusätzlich existiert eine Technologieschicht, auf dem das Design aufbaut ist. Die aufeinanderliegenden Schichten heißen von oben nach unten: Learning, Storytelling, Gameplay und User Experience. Der Autor empfiehlt in der obersten Schicht zu beginnen und sich dann nach unten zu arbeiten. In der Learning-Schicht legt der Designer die Ziele fest, um anschließend den gesammelten Lern- oder pädagogischen Inhalt so zu gestaltet, damit die Ziele erreicht werden.
In der zweiten Schicht Storytellings entwirft der Designer den Kontext der Spielwelt, das Charakterdesign und die Erzählung. Der Spieler erfährt je nach Interaktion mit dem Spiel, seine eigene Geschichte. Die dritte Schicht, beschreibt den Spielverlauf und orientiert sich dabei am MDA Modell. Mechanics definieren die Spielregeln, die Herausforderungen und die Ziele, unter denen ein Spieler handeln kann. Das Handeln und die Interaktion über Zeit, wird über die Dynamics beschrieben. In Affects werden explizit die durch die Handlungen des Spielers hervorgerufenen Emotionen angesprochen. Der Designer muss sich also zunächst Emotionen überlegen, die er bei dem Spieler ansprechen möchte und darauf die Mechanics aufbauen, um sein Ziel über die Dynamics zu erreichen. Das Überprüfen der erwünschten Emotionen kann wiederum nur über Spieltests, bereits im Prototypstadium ermittelt werden. Die Schicht User Experience ist am meisten für den Spieler sichtbar. Diese soll dem Benutzer das Spiel zugänglich machen und durch Interkation erlauben, die vom Designer festgelegten Lernziele zu erreichen. Dabei sollte das User Interface intuitiv zu steuern sein und den Spieler nicht vom eigentlichen Spielfluss ablenken. Die letzte Schicht ist die Technologie, welche aufgrund des großen Einflusses auf die Entwicklung stets beachtet ist. Der Designvorschlag kann nach Winn für die Analyse von existierenden SG, als auch zum Erstellen von neuen SG verwendet werden. Der Fokus der Vorgehensweise liegt auf der Integration des charakterisierenden Ziels in das SG und die gegenseitige Wechselwirkung zwischen Spieldesigner und Spieler.

5 GIRP Modell

Im Zuge dieser Arbeit ist eine eigene Empfehlung für die Entwicklung eines SG entstanden. Dieser Vorschlag basiert auf der Grundidee des DPE Modells und somit auch auf dem MDA Modell. Auch wenn das DPE Modell bereits die einzelnen Schichten und den Einsatz von darin enthaltenen Spielfaktoren beschreibt, erscheint es sinnvoll die Schichten durch weitere Kriterien zu ergänzen. Aus den fünf Schichten des DPE Modells sind

Kriterium	Beschreibung
Charakterisierendes Ziel definieren	Den Hauptzweck des SG definieren. Dabei können die Attributlisten aus der DIN SPEC 91380 helfen. Die Pädagogik hat Vorrang gegenüber dem Spieldesign [17].
Anwendungsbereich	Den oder die Anwendungsbereiche des SG festlegen
Zielgruppe	Das Festlegen der Zielgruppe ist neben dem Anwendungsbereich eine weitere Randbedingung [18].
Art des SG festlegen	Endogen vs. Exogen.
	Dynamischer vs. statischer Inhalt.
Problemstellung aufstellen	Das Aufstellen einer Problemstellung, aus dem charakterisierenden Ziel und den Randbedingen bietet eine klare Orientierung für den Design und Entwicklungsprozess [18].
	Neben Randbedingungen wie Anwendungsbereiche und Zielgruppe können auch äußere Faktoren wie das Entwicklungsbudget eine Rolle spielen.
Lernziele festlegen	Falls Wissen vermittelt wird, sollten Lernziele über eine Taxonomie festgelegt werden. Ein Beispiel dafür ist Bloom's Taxonomie [4] und überarbeitete Versionen wie [15]. Je nach Anwendungsgebiet bieten es sich an Taxonomien aus dementsprechendem Fachgebiet zu wählen.
Domänenexperten einbinden	Domänenexperten sollten vor allem während des Designprozesses involviert sein [18]. Die Richtigkeit des Lernstoffs muss garantiert werden. Begrifflichkeiten aus der Domäne sollten richtig angewendet werden [7].

Tabelle 3: Hauptziel

die vier Phasen *Hauptziel* (goal), *Spielidee* (idea), *Konzeption und Umsetzung* (realization) und *Wahrnehmung* (perception) entstanden. Durch das Zusammensetzen der englischen Begriffe ergibt sich das GIRP Modell. Handlungsempfehlungen zu den einzelnen Schichten wurden tabellarisch festgehalten. Abbildung 1 beschreibt das Vorgehen des Designs und der Entwicklung anhand der vier Phasen. Das Aufstellen eines charakterisierenden Ziels ist die wichtigste Komponente der Phase Hauptziel. Dessen Wichtigkeit wird in [6][7][17][19][18] bestätigt. Das Aufstellen des Ziels wird in dieser Phase durch die Erarbeitung weiterer Kriterien unterstützt. Diese Kriterien werden in Tabelle 3 vorgelegt. Das Hauptziel sollte unter Einbindung von

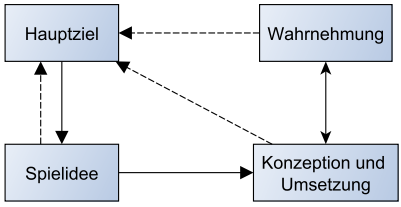


Abbildung 1: Entwicklungsempfehlung. Ausgefüllter Pfeil: Durchführung. Gestrichelter Pfeil: Evaluation

Kriterium	Beschreibung
Thema der Spielwelt	Der Kontext der Spielwelt muss entwickelt werden. Dazu zählt beispielsweise auch das Design von Charakteren. Alle weiteren Implementierungen von Spielfaktoren unterliegen der Spielfantasie und müssen zu dieser passen, ansonsten wirkt das Spiel unharmonisch [22].
Erzählung	Eine Geschichte kann einem stringenten Pfad folgen oder durch Entscheidungen des Spielers Verzweigungen erhalten, was zu komplexeren Geschichten mit mehreren Wegen führt [18]. Eine gute Erzählung leitet den Spieler durch das Spiel. Durch Verknüpfen der Erzählung mit Lerninhalt kann dabei Wissen unterbewusst vermittelt werden. Vermitteltes Wissen sollte kein Nebenprodukt sein, welches den Spieler vom Spielfluss ablenkt [20].
Soziale Interaktion	Soziale Interaktion im kollaborativen oder kompetitiven Stil sind ein wichtiger Bestandteil für ein gutes Spiel [5] [20]. Eine Umsetzung eines solchen Spielmodus kann sich auf die Handlung des Spiels auswirken und sollte deswegen früh in den Designprozess miteinbezogen werden [22]. Eine Auswahlmöglichkeit zwischen verschiedener Spielmodi macht das Spiel für eine größere Masse an Spielern zugänglich und für verschieden Spielertypen unterhaltsam [7].

Tabelle 4: Spielidee

Domänenexperten iterativ erarbeitet werden. Die Phase Spielidee beschreibt Kriterien die auf dem Hauptziel basierend entwickelt werden. Das Erarbeiten einer grundsätzlichen Spielidee mit einer Spielwelt und einer Erzählung ist ein iterativer Kreativprozess. Die festgelegten Ziele aus der ersten Phase sollten dabei mit einfließen. Durch das Durchlaufen dieses Prozesses, können sich bereits weitere Eigenschaften für das Spiel, wie beispielsweise das Genre bilden. Die zu beachtenden Kriterien dieser Phase sind in Tabelle 4 dargestellt.

Nachdem die grundlegende Spielidee des SG feststeht, kann eine Umsetzung dafür konzipiert werden. Dies geschieht in der nächsten Phase in der vor allem erarbeitete Randbedingungen zu beachten sind. Die Phase Konzept und Umsetzung benötigt eine Zusammenarbeit von Gamedesignern und Programmierern, um die Idee des SG in Spielmechaniken zu verwandeln.

Die Schwierigkeit bei dieser Umsetzung ist es die Balance zwischen ernsthaft und Spiel zu finden. Die Kriterien, welche für die Konzeption und Umsetzung eine große Rolle spielen, sind in Tabelle 5 abgebildet.

Die letzte Phase beschreibt die Wahrnehmung des Spielers auf audiovisueller Ebene. Sie ist ergänzend zur dritten Phase und sollte deswegen parallel zu dieser geschehen. Kriterien dieser Phase sind in Tabelle 6 abgebildet.

Der zurückführende Pfeil auf das Hauptziel in Abbildung 1 soll verdeutlichen, dass die

Kriterium	Beschreibung
Technologiewahl	Die Technologiewahl bildet den Grundstein für die Implementierung des Spiels und sollte deswegen gut begründet sein. Durch sie entstehen neue Möglichkeiten, aber auch Einschränkungen [6].
Erzeugung von Spaß	Nach der Flow Theorie von Czikszentmihalyi muss ein Spiel herausfordern, aber darf dabei nicht überfordern [8]. Es muss an die Fähigkeiten des Spielers angepasst sein. Sweetser und Wyeth stellen eine Möglichkeit vor Vergnügen mit Hilfe von Flow und anderen Spielfaktoren zu erreichen [23]. Für Spaß erzeugende Spielfaktoren existieren weitere Taxonomien wie von Hunicke et al. [21]. Heuristiken zur Spielbarkeit wie von Desurvire und Wiberg bieten vor allem in den Anfängen der Entwicklung eine gute Orientierung [9]. Es ist zu beachten, dass es verschiedene Spielertypen gibt, die unterschiedlich viel Spaß an bestimmten Spielfaktoren empfinden [3]. Ein breiteres Spektrum von Spielfaktoren, kann deswegen durchaus sinnvoll sein. Die Umsetzung der Spielfaktoren gilt es im Verlauf der Entwicklung durch Spieltests mit der Zielgruppe zu überprüfen.
Spielmechaniken	Neben der Entwicklung von grundlegenden Spielmechaniken wie beispielsweise der Spielsteuerung, werden Mechaniken benötigt, welche Spielfaktoren und das charakterisierende Ziel umsetzen. Während dieses Prozesses muss sich die Auswirkung auf den Spieler bewusst gemacht werden. Für die Vermittlung von pädagogischen Werten können beispielsweise Herausforderungen genutzt werden [22]. Eine weitere Möglichkeit ist das Aufteilen der Lernziele auf mehrere Level. Dadurch kann gezielt Wissen vermittelt oder abgefragt werden. Gleichzeitig ist der Einsatz von verschiedenen Spielmechanismen, beispielsweise in Form von Minigames möglich [2]. Dabei gilt es ein einheitliches Spielerlebnis zu erhalten.
Überwachung des Spielfortschritts	Eine Überwachung des Spielfortschritts, wie beispielsweise durch Spielstatistiken dienen nicht nur dem Spieler als Feedback, sondern können auch dazu verwendet werden den Spielprozess zu beobachten. Dies ist vor allem für die Auswertung des SG nützlich. Darüber können Schlüsse über das Erreichen des charakterisierenden Ziels geschlossen werden [7].

Tabelle 5: Konzeption und Umsetzung

Richtigkeit des Hauptziels in der Entwicklungsphase stets überprüft werden muss. Hierzu können Personen der Zielgruppe eingesetzt werden, die bereits erste funktionsfähige Prototypen testen. Involvierte Personen des gesamten Prozesses sind in Abbildung 2 dargestellt.

Kriterium	Beschreibung
Feedback	Um auf das charakterisierende Ziel hinarbeiten zu können, sollte der Spieler eine Rückmeldung über den Spielfortschritt erhalten. Beispielsweise in Form von Fortschrittsbalken[7]. Dies kann gezielt auf die Lerninhalte oder auch den Spielfortschritt angewendet werden. Das Sammeln von Punkten kann neben der Rückmeldung als kompetitiver Faktor eingesetzt werden
Belohnungen	Für eine positive Bestätigung und dadurch erzeugende Motivation, sollten Belohnungen wie Upgrades oder Skins in gezielter Form einsetzt werden [18].
Ästhetik	Die Grafiken des Spiels müssen an das Anwendungsgebiet und die Zielgruppe angepasst sein und sollten für diese ansprechend wirken [7]. Simulationen sollten versuchen ihre Umgebung möglichst realistisch darzustellen. Neben Grafiken, ist auch der Einsatz von passenden Audioeffekten und Musik hilfreich um die Aufmerksamkeit des Spielers aufrecht zu erhalten [20].
UserInterface	Es sollte zwischen zwei Benutzeroberflächen unterschieden werden [18]. Die Erste besteht aus einer Auswahl von Spielmenüs. Über sie gelangt der Spieler in das eigentliche Spiel und kann Einstellungen konfigurieren, welche das Spiel betreffen. Die zweite Benutzeroberfläche betrifft das Spiel selbst und lässt Interaktionsmöglichkeiten mit dem Spiel selbst zu. Die Benutzeroberflächen, sollten intuitiv gestaltet werden, damit sich der Spieler auf das Spielen konzentrieren kann [6]. Neben Designelementen wie natural mapping für die Steuerung, sollte sich an weiteren grundlegenden Designregeln orientiert werden.

Tabelle 6: Wahrnehmung

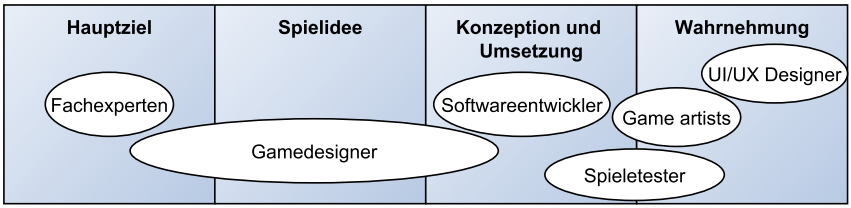


Abbildung 2: Involvierte Personengruppen in Anlehnung an [18]

6 Fazit

In diesem Artikel ist eine Definition, eine Taxonomie und eine Entwicklungsempfehlung für SGs vorgestellt worden. Nach der Untersuchung mehrerer Definitionen wird deutlich, dass sich ein SG dadurch definiert,

dass es neben einem unterhaltenen Faktor, in erster Linie ein pädagogisches Ziel verfolgt. Dieses auch charakterisierende Ziel genannt, dient besonders zur Klassifizierung von SGs, weil es von dessen Anwendungsbereich und der Zielgruppe abhängig ist. Da es bisher noch keinen offiziellen Standard für die Klassifizierung eines SG gibt, ist es wichtig, dass sich innerhalb der Branche auf einen Standard geeinigt wird. Die DIN SPEC 91380 ist gemeinsam durch Experten und Branchenmitgliedern innerhalb von Deutschland erstellt worden. Sie bietet eine Struktur für eine genaue Klassifizierung eines SG. Die Sinnhaftigkeit dieser Struktur wird, durch die in diesem Artikel untersuchte Literatur bestätigt. Eine kompakte Taxonomie wurde vorgeschlagen, die für eine schnelle Klassifizierung sinnvoll ist. Abschließend wurde die eigene Entwicklungsempfehlung des GIRP Modells vorgeschlagen, die von dem Schichtenmodell DPE [6] abgewandelt ist. Es wurden die vier Phasen Hauptziel, Spielidee, Konzeption und Umsetzung und Wahrnehmung definiert, die bei einer Entwicklung nacheinander, aber jeweils iterativ durchlaufen werden. Jede Phase wird durch bewerte Vorgehensweisen und Qualitätsattributen aus der Literatur gestützt. In der Zukunft wird darauf hingearbeitet, dass die Entwicklungsempfehlung auf ein reales Projekt angewendet wird, damit weitere Kriterien erarbeitet und Vorhandene verfeinert werden können.

Literatur

[1] Clark C. Abt. *Serious games.* Repr. d. Ausg. New York, 1970. Lanham: Univ. Press of America, 1987.

[2] André F. S. Barbosa u. a. „A New Methodology of Design and Development of Serious Games“. In: *International Journal of Computer Games Technology* 2014 (2014), S. 1–8. ISSN: 1687-7047. DOI: 10.1155/2014/817167.

[3] Richard Bartle. „Hearts, clubs, diamonds, spades: Players who suit MUDs“. In: *Journal of MUD research 1.1.*

[4] Benjamin Samuel Bloom. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook 1: Cognitive domain.* New York: David McKay Co Inc., 1956.

[5] Matthew Bond und Russell Beale. „What makes a good game? Using reviews to inform design“. In: Electronic Workshops in Computing. BCS Learning & Development, 2009. DOI: 10.14236/ewic/HCI2009.52.

[6] Brian M. Winn. „The Design, Play, and Experience Framework“. In: *Handbook of Research on Effective Electronic Gaming in Education.* IGI Global, 2009, S. 1010–1024. DOI: 10.4018/978-1-59904-808-6.ch058. URL: https://www.igi-global.com/chapter/design-play-experience-framework/20133.

[7] Polona Caserman u. a. „Quality Criteria for Serious Games: Serious Part, Game Part, and Balance“. In: *JMIR Serious Games* 8.3 (2020), e19037. ISSN: 2291-9279. DOI: 10.2196/19037.

[8] Czikszentmihalyi M. *Flow: The psychology of optimal experience.* 1990. URL: https://mktgsensei.com/amae/consumer%20behavior/flow_the_psychology_of_optimal_experience.pdf.

[9] Heather Desurvire und Charlotte Wiberg. „Game Usability Heuristics (PLAY) for Evaluating and Designing Better Games: The Next Iteration“. In: *Online communities and social computing.* Hrsg. von A. Ant Ozok und Panayiotis Zaphiris. Lecture notes in computer science. Berlin: Springer, 2009, S. 557–566. ISBN: 978-3-642-02774-1.

[10] Deterding, S., Dixon, D., Khaled, R., Nacke, L. „From game design elements to gamefulness: defining gamification“. In: (2011), S. 9–15.

[11] *DIN SPEC 91380:2018-06, Serious Games Metadata Format.* Berlin. DOI: 10.31030/2853739.

[12] Ralf Dörner u. a., Hrsg. *Serious games: Foundations, concepts and practice*. Switzerland: Springer, 2016. ISBN: 9783319406114. DOI: 10.1007/978-3-319-40612-1.

[13] Jan L. Plass, Bruce D. Homer, Charles K. Kinzer. „Foundations of Game-Based Learning“. In: *Educational Psychologist* (), S. 258–283. URL: https://doi.org/10.1080/00461520.2015.1122533.

[14] Juho Hamari Kai Huotari. *Defining gamification: a service marketing perspective*. 2012.

[15] David R. Krathwohl. „A Revision of Bloom's Taxonomy: An Overview“. In: *Theory Into Practice* 41.4 (2002), S. 212–218. ISSN: 0040-5841. DOI: 10.1207/s15430421tip4104_2.

[16] Fedwa Laamarti, Mohamad Eid und Abdulmotaleb El Saddik. „An Overview of Serious Games“. In: *International Journal of Computer Games Technology* 2014 (2014), S. 1–15. ISSN: 1687-7047. DOI: 10.1155/2014/358152.

[17] Stephanie B. Linek u. a. „When Playing Meets Learning: Methodological Framework for Designing Educational Games“. In: Springer, Berlin, Heidelberg, 2009, S. 73–85. DOI: 10.1007/978-3-642-12436-5_6. URL: https://link.springer.com/chapter/10.1007/978-3-642-12436-5_6.

[18] Philip Mildner und Florian 'Floyd' Mueller. „Design of Serious Games“. In: *Serious Games*. Springer, Cham, 2016, S. 57–82. DOI: 10.1007/978-3-319-40612-1_3. URL: https://link.springer.com/chapter/10.1007/978-3-319-40612-1_3.

[19] Aaron Pendleton und James Okolica. „Creating Serious Games with the Game Design Matrix“. In: Springer, Cham, 2019, S. 530–539. DOI: 10.1007/978-3-030-34350-7_51. URL: https://link.springer.com/chapter/10.1007/978-3-030-34350-7_51.

[20] Werner Siegfried Ravyse u. a. „Success factors for serious games to enhance learning: a systematic review“. In: *Virtual Reality* 21.1 (2017), S. 31–58. ISSN: 1434-9957. DOI: 10.1007/s10055-016-0298-4. URL: https://link.springer.com/article/10.1007/s10055-016-0298-4.

[21] Robin Hunicke, Marc LeBlanc, Robert Zubek. *MDA: A formal approach to game design and game research*. 2004. URL: https://www.aaai.org/papers/workshops/2004/ws-04-04/ws04-04-001.pdf.

[22] Yen-Ru Shi und Ju-Ling Shih. „Game Factors and Game-Based Learning Design Model“. In: *International Journal of Computer Games Technology* 2015 (2015), S. 1–11. ISSN: 1687-7047. DOI: 10.1155/2015/549684.

[23] Penelope Sweetser und Peta Wyeth. „GameFlow“. In: *Computers in Entertainment* 3.3 (2005), S. 3. DOI: 10.1145/1077246.1077253.

[24] Thomas W. Malone und Mark R. Lepper. „Making Learning Fun : A Taxonomy of Intrinsic Motivations for Learning“. In: *Aptitude, Learning, and Instruction*. Routledge, 2021, S. 223–254. ISBN: 9781003163244. DOI: 10.4324/9781003163244-10. URL: https://www.taylorfrancis.com/chapters/edit/10.4324/9781003163244-10/making-learning-fun-thomas-malone-mark-lepper.

©2022 Tobias Lauxmann. 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

Pre-Study: Adaptation of 3D sign language avatars to the speakers appearance and its influence on user experience

Lara Krautmacher
Hochschule Reutlingen
Reutlingen, Deutschland
Lara.Krautmacher@Student.Reutlingen-University.de

Abstract

For Deaf people sign language translations are essential for full accessibility, as spoken and written languages are foreign languages. The translation with interpreters is a time-consuming and costly process. Therefore other translation possibilities like machine translation with 3D sign language avatars were examined in the last years. Thereby the focus often laid on the technical implementation. This paper whereas questions whether 3D-avatars for online-video translation could have a positive effect on the users experience if avatars were adapted to the speakers appearance, which is not possible with human interpreters. A quantitative online survey with avatar prototype videos showed there could be positive effects, even if it was not possible to statistically prove it, as the sign language animations were not comprehensible enough due to a lack of sign language skills of the researcher.

CCS Concepts

• **Human-centered computing** → **Accessibility systems and tools**; *Visualization design and evaluation methods*; • **Applied computing** → *Language translation.*

Keywords

Sign Language Avatars, User Experience, 3D-Avatars, Online-Video Translation

1 Introduction

According to the World Health Organization (WHO) 5 % of the world's population have a disabling hearing loss[1]. To be able to communicate completely without hearing there are many fully-fledged sign languages all over the world deaf[2] people often use as their primary language. In history scientists have shown that deaf people have significant difficulties with learning to read and write [3]. Therefore it is important to make information also accessible for Deaf people not only in written language, but also in sign language.

Online-Videos for example are already subtitled most of the time but could be more accessible for the Deaf community if they were translated to sign language. But still this is rarely done as it is a time-consuming and costly process and there are not enough professional sign language interpreters.

[1] https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss (visited on 11/21/2021)

[2] In this paper the term deaf is used for people with a disabling hearing loss and the term Deaf is used to write about members of the Deaf community using sign language as their preferred language [4]

Betreuer/-in Hochschule: Prof. Dr. rer. nat. Gabriela Tullius
Hochschule Reutlingen
Lara.Krautmacher@Reutlingen-University.de
Informatics Inside Frühjahr 2022
11. Mai 2022, Hochschule Reutlingen
Copyright 2022 Lara Krautmacher

Additionally when translating videos to sign language there often is a discrepancy between the appearance of speaker and interpreter and multiple persons are translated by one interpreter. This also comes from the lack of interpreters and therefore lack of choice to match the appearance.
In the past years opportunities of automatic sign language translation through machine translation and 3D avatar animation were examined in detail. Also the acceptance of sign language avatars (SLA) was investigated, but there is no research about the different appearance of interpreter and speaker and the opportunities avatars provide in this sector. Therefore in this paper the question is addressed whether a discrepancy in the appearance has a negative influence on the user experience (UX) of translated online-videos. In addition it is examined if the UX could be improved by using SLAs, with a similar appearance as the speakers.
To answer these questions two video prototypes with SLAs adapted to the appearance of the speakers were created. Subsequently a quantitative online-survey was conducted, where the UX of the original videos is compared to the prototype videos by using a shortened version of the MeCue questionnaire.

2 Deaf culture

When conducting research for the deaf people it is important to understand that the Deaf community has its own language and also its own culture. For successful research this culture needs to be understood and integrated into the research design. Hence this chapter presents some background knowledge about the Deaf culture and research within this community.
The culture of the Deaf is influenced a lot by its sign languages. All over the world there are fully developed languages, with an own grammar and structure, based on body language, signs and facial expressions. With this languages the Deaf community also developed their own form of art like signed lyrics and theater.
In Germany, where this study is conducted, Deaf people use the German sign language "Deutsche Gebärdensprache" (DGS) and most of the Hearing people use spoken and written German (GD). These are the languages the paper refers to hereafter.
For the Deaf community it is important that information is translated for them as it is done for spoken foreign languages. As for many years sign languages had not been accepted and in many countries they are still not acknowledged [9], there often are no translations provided. In Germany this acceptance of DGS as a language happened in the year 1981[3].
As DGS differs a lot from GD, GD is a foreign language for people, who use DGS as their primary language. The language therefore needs to be learned, which is difficult for Deaf people, due to their hearing condition and the fact that there is no written form of DGS. Hence the illiteracy among Deaf people is higher than the average and also speaking skills are often lowered [4].
These lower speaking and writing skills the Deaf community resulted in a long history of discrimination and misunderstanding [4]. In history there have been "negative societal attitudes about deaf people and disability fostered by ignorance, neglect, superstition and fear" [4, p. 267]. And still there are a lot of limitations as not many hearing people are able to use DGS. These limitations also exist in scientific research, as papers usually have to be submitted in a written language. Hence a lot of research is made by Hearing persons

[3] https://www.taubwissen.de/content/index.php/geschichte/geschichte-der-deutschen-gebaerdensprache/kampf-um-anerkennung-der-dgs/579-kampfumdgsteil1.html (visited on 11/21/2021)

for the Deaf community and not by Deaf people themselves [9]. This can result in research not fitting for the Deaf themselves, especially if Deaf persons are not well integrated into the research planning and the process [9]. Therefore in this work Deaf people are used as consultants when planning the study.

3 Sign language avatars

SLAs have been a subject, which was often addressed over the last years. This paragraph displays relevant research projects concerning automated sign language translation with avatars and the acceptance of avatars.

3.1 Related work

Big parts of the research about SLAs focus on automated translations of written language to sign language with SLAs as form of presentation. Concerning the translation of written language to sign language existing research projects explore different machine translation approaches to translate the linguistic structure of written languages to the belonging sign language [2]. In those projects initially the grammar is translated and then separate avatar animations from a library are combined to a fluent translation. The project OpenSigns developed an overall solution to translate several languages with one application [2]. The application Hand Talk and the software SiMAX are thereby projects, which are already available for use [4] [5].

Additionally there are research projects examining sign language tracking methods for translations and automated generation of sign language animations. To recognise sign language Lang et al. use depth and image information from a Kinect and Hidden-Markov-Models [5]. Also sensor based recognition with glove wearable tracking fingermovement was explored [7]. As analysing Kinect data and finger movements with sensor-gloves can not deliver the needed level of detail, which includes facial tracking, image-processing and machine learning methods are examined in the newest research projects. Ye et al. and Li et al. use convolutional neural networks for sign language recognition from continuous video material [10][6]. In their experiments Li et al. were able to achieve a 95.52 % recognition rate [6]. Therefore a basis of sign language translation with avatars through machine translations is already set. Even though there a the need for further research, as there is still a lack of displaying the multi-channel information of sign language, like facial expressions and subtle body movements influencing the interpretation of signed sentences.

3.2 Acceptance of avatars

For avatar translations to be successful it is important to created animations and avatars, which are accepted by the Deaf community. Recent studies showed there is a general possibility that avatars find acceptance in situations, which are not face-to-face [4][8]. But this acceptance is influenced by factors like, naturalness, comprehensibility, attitude and impressions [8]. Studies showed that computer generated avatar animations are less accepted at the moment as they do not provide the necessary amount of naturalness and they were considered robotic, stiff and emotionless [4][8]. Handmade animations and motion-capture animations with more precise facial expressions and body movements receive a higher acceptance [4][8]. Therefore it is important to provide natural movements and facial expressions. Especially for people, who learned sign language earlier in life, as little details in the animations are very important for them [8]. Also the avatars style influence the acceptance of avatars, Adamo-Villani and Anasingaraju showed that a comic like representation is better than a realistic representation due to

[4] https://simax.media/wie-simax-funktioniert/ (visited on 03/23/2022)

[5] https://www.handtalk.me/en/app/ (visited on 03/23/2022)

an avoidance of the uncanny-valley effect [1]. Yet there has not been any research concerning the adaptation of avatars to the speakers appearance and its influence on the acceptance and experience. Therefore this paper gives first insights into this area.

4 Study design

This section explains the aim and the target group of this research. Additionally the study design is presented and it is addressed how the needs of Deaf people were met. Concluding it is illustrated how the prototype animations for the survey were created.

4.1 Aim of the study

The aim of this study is creating insights whether SLAs could be used to improve the UX for Deaf people when watching online-videos by matching the SLA's appearance to the speaker's appearance. As a basis it is also determined whether a difference in the appearance of speaker and translator or a translation of several people by one interpreter, is considered as negative by the Deaf community.
In detail in this paper, the following research questions are addressed:

RQ1: Does a discrepancy in the appearance of speaker and interpreter in sign language translations of online videos have a negative influence on the user's experience?

RQ2: Does the translation of several speakers by one interpreter have a negative influence on the user's experience?

RQ3: Can 3D-signing avatars adapted to the speaker's appearance improve the user experience of online-videos with sign language translation?

The results of this study should give a first impression if an appearance adaptation of avatars could help with further improvement of sign language translations. If the results display a positive impact, it can be used to decide on further research projects concerning the adaptation of avatars in automated translation. It can also build a decision basis for avatar design when transferring interpreters movements on avatars to adapt their appearance or make the interpreters invisible.

4.2 Questionnaire design

To answer the described research questions it was decided to conduct a quantitative accessible online survey. In this way a broad range of people in the Deaf community all over the country can be reached. To be able to gain insights about the UX of SLA videos, prototype videos have been created and compared to the original interpreter videos. For this purpose a shortened and adapted version of the MeCue questionnaire[6] was applied. The following section describes structure and accessibility adaptions in detail.

4.2.1 Structure

The online survey was structured into four parts. The first and the third part were constructed to answer RQ3, the second part tackles RQ1 and RQ2 and in the last part demographic questions were asked to gain further insights.

To be able to compare the UX of interpreter translated online videos and SLA translated videos the questionnaire MeCue was used on both types of video to get a comparable UX score and answer RQ3. MeCue was chosen as it asks for several aspects of UX, like usability, usefulness, aesthetics, product loyalty, positive and negative emotions. With a broad range of UX testing MeCue is a good fit, as for videos it is difficult to decide, which UX aspects are the most important. The UX of the two video types was measured in two sessions, one for each kind. In these sessions at first two example videos were presented and afterwards MeCue was queried. To

[6] http://mecue.de (visited on 03/23/2022)

separate the two video-evaluations as much as possible to avoid priming effects, they were split to the first (human interpreter) and the third part (avatar). Questions concerning RQ1 and RQ2 were asked in between.

To answer RQ1 and RQ2 the following questions were asked in the second section:

1. What do you think about sign language interpreter often looking different than the speaking people?

2. What do you think about multiple people being translated by one sign language interpreter?

As response option for both questions a semantic differential with five steps ("Doesn't bother me at all", "Doesn't bother me", "Neutral", "Bothers me", "Bothers me a lot") and the additional option "This never attracted my attention before" were used.
Additional prior experiences with SLAs were queried. Therefore three images of SLAs were presented as example. If experiences existed, it was asked whether the avatars were liked or not. These queries were intended to measure the influence of prior experiences with avatars on the results.

Afterwards part three of the questionnaire followed as described before and in the end a fourth part with demographic questions was presented. In this section the participants were asked about their age and gender identity to get further insights about the test group structure. It was also asked in which age the test persons learned sign language and whether they perceive themselves as hearing, hearing impaired or deaf. Finally the survey was closed with thanking words and an opportunity to give additional qualitative answers.

4.2.2 Accessibility

In order to make the survey accessible for the test group of people using sign language as their primary language, it was planned to translate the survey into sign language videos and present them equally to the texts. Therefore many sign language interpreters were contacted by e-mail and phone and asked for their help with the translation. Unfortunately even if more than 50 sign language interpreters were contacted it was not possible to find an interpreter to translate the questionnaire. As a result it was only possible to present the survey in a textual form.
To even tough make it easier to fill out the form for sign language speakers the questionnaire was adapted to their needs as much as possible. Following the instructions of a Deaf sign language interpreter the questionnaire was shortened radically, as long and complicated questions were not suited for the Deaf culture. Therefore especially the MeCue questionnaire was adapted, as the original version with a seven step semantic differential and 33 items is very long. When shortening MeCue the researcher was aware of the fact that the adaptions result in the UX score of the videos not being comparable to other MeCue results anymore. But as the video scores still can be compared to each other, as the same shortened version is used for the evaluation of both videos, this limitation was accepted.
MeCue was shortened in two ways, the semantic differential was reduced from seven to five steps and the number of items was truncated. To reduce the amount of items, the most important categories of MeCue were chosen. MeCue asks questions about usability (U), usefulness (N), aesthetics (A), status, bond, positive emotions (PE), negative emotions (NE), usage intention, product loyalty (L) and asks for an overall score. For this evaluation the categories U and N were considered important as a translation needs to be useful and usable as basis for a good UX. Additional the categories A, PE, NE and L were considered important to ask for. A is interesting as the evaluation focuses on the style of the translation and L is a direct indicator which translation is preferred. By asking

about PE and NE the experience can be measured. For each chosen category one items was selected, by searching for the most fitting one concerning the evaluation of videos instead of a physical product. For PE and NE each three items were selected to get a more precise insight about the actual experience. In this way the MeCue questionnaire was shortened to ten items with each five answer options, leading to a more fitting solution for an evaluation with sign language speakers.

4.3 Prototype development

To be able to compare the UX of interpreter translated online-videos with a differing appearance of interpreter and speaker to the UX of videos translated by SLAs with an adapted appearance, prototype videos with SLAs have been created. To achieve a maximum of comparability the prototype videos were created based on existing videos, which were available as well with translation and without. In this way it was possible to present the same video with avatar and with interpreter translation.

In order to show different kind of videos prototype videos for two genres were created. One video shows a politician, in a neutral setting, talking about news. With this video a neutral one person situation in a news context is tackled. To cover the entertaining video genre and the translation of multiple persons, an entertaining video with four people talking about cooking an egg in a volcano was translated.

To create the prototype videos, 3D-avatars were manually animated, based on the existing translation and subsequently added to the video without translator. For the animations avatars with a similar appearance of the speakers were needed and creating avatars from scratch was not possible with the timeline and resources of this study. Therefore avatars of the platform ReadyPlayerMe were used. ReadyPlayerMe is an online platform, where avatars can be individualized and downloaded as .glb files. To achieve a good match of avatar and speaker at first avatars were created as matching as possible on ReadyPlayerMe and downloaded. Afterwards they were further modified in Blender with a special focus on facial features and overall textures.

These avatars were then animated in Blender to present the video-translation. Therefore as well facial expressions as hand gestures were animated step by step to match the movements of the original sign language interpreter. As the person animating was not capable of using sign language, Deaf people were asked for feedback whether the animation was understandable. Based on this feedback animations were adapted in an iterative process. When it was perceived comprehensible the rendered animation was inserted with After Effects into the original video without translation. Figure 1 shows a screenshot of the resulting video prototype for the category entertaining videos.

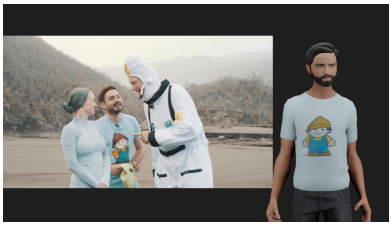


Figure 1: SLA prototype video screenshot

4.4 Test persons

For the conducted online-survey the aspired test group was a wide range of people considering themselves Deaf and using sign language as their primary language. To distribute the survey to those people, multiple channels were used. Sign language interpreters were used as multipliers spreading the online-survey to their clients. In this way a wide and heterogeneous group was reached. Another

way to spread the survey was searching for Deaf people on social media and asking them in direct messages to participate in the survey. By doing so, a very personal contact with a high chance of completion was created. Additionally snowballing was used, by asking all people to share the survey with their Deaf contacts.
In this way 49 people were recruited for the online survey. 30 of these participants have completed the survey fully and 19 cancelled in between. One reason for this poor completion rate might be the missing sign language translation of the questions asked. In the following explanations the 30 completed data-sets were used to analyze the results.

5 Results

To analyse the results of the survey at first the data of the semantic differentials were translated into numeric values to be able to statistically analyse them based on scores. The five steps were numbered from one to five, where "Doesn't bother me at all" got the number one and "Bothers me a lot" got number five. The additional answer opportunity "Never recognised before" was marked as zero an only indirectly integrated into the scoring. Based on these numeric translation the three research questions were analysed.

RQ1

To answer this research question the mean value of the answers about the peoples opinion about the different look of interpreter and speaker was calculated. With a mean value of 1,96 the study comes to the result that people do not recognise the difference as negative. 7 of the 30 test persons did not recognise it before.

RQ2

To answer this research question the mean value of the answers about the peoples opinion about the translation of several speakers by one interpreter was calculated. With a mean value of 2,09 the study comes to the result that people do not recognise it as negative. In this case only 2 of 30 test persons did not recognise it before. This leads to the assumption that multiple people translated by one interpreter is more noticeable than a difference in the appearance of interpreter and speaker.

RQ3

To answer this research question the mean value for each MeCue category was calculated for the interpreter and the avatar videos. The difference between the calculated results were statistically verified with t-tests. The calculated mean values and p-values of each category are shown in table 1. It can be seen that the mean values of N, U and A are higher than four for the interpreter videos, and lower than two for the avatar videos. This leads to the conclusion, that the avatar videos have not been useful, usable and aesthetic and therefore assumed not comprehensible enough. In comparison interpreter videos were very useful, usable and aesthetic. The difference between the two values of each category is significant as the p-values are 0,00000. It can be assumed that the lacking comprehensibility of the avatars has a high influence on the last three categories as well and these results would differ if the animations were better comprehensible. In the category PE interpreters have a higher value (3,77) than avatars (2,33). Therefore interpreter translations lead to an over average and avatars to an under average positive emotional response with a significant difference between the values. For NE interpreter videos have a result of 2,17, which is significantly lower than the score 3,63 of the avatar videos. Therefore the test persons had a under average negative load of emotions with the interpreters and an over average negative amount of emotions

Table 1: Results RQ3

Category	Interpreter	Avatar	t-test
U	4,07	1	0,00000
N	4,30	1,5	0,00000
A	4,30	1	0,00000
L	2,43	3	0,06278
PE	3,77	2,33	0,00000
NE	2,17	3,63	0,00007

with the avatar version. The only category in which the difference between the videos had not been significant is L. Interpreters gained a under average score with 2,43 and avatars an average score with 3 and a p-value of 0,06278. The higher value of the avatar videos in this category might be a clue for the avatars possibly being a good form of translation if the comprehensibility and the overall usefulness was improved even if it is still average and not significantly different with an alpha-value of 0,05. These results overall show that the UX of the avatar videos with adapted appearance more negative than the UX of the interpreter videos.

The examination of subgroups showed that the mean values of N, U and A were significantly higher for people, who learned sign language as children than for people who learned it later in their lives. The mean values of N, U, A and L of test persons who think the different appearance of interpreter and speaker is disturbing were significantly higher than for the rest of the test group. Additional a difference between people who prefer subtitles and people who prefer interpreter translations could be seen. The interpreter group showed higher values over all categories, but as the group was very small no statistical statement can be made.

Additional results

In addition to the research question results the data was analysed on correlations. Thereby correlations with scores between 0 and 0,4/-0,4 were considered weak, scores between 0,4/-0,4 and 0,7/-0,7 were considered moderate and a score higher/lower than 0,7/-0,7 were considered strong. In this way no strong correlations, but three moderate correlations were detected. There was a moderate correlation score of -0.45 between the age and the positive emotions towards the avatar. That shows the older the test person the less positive emotions did the avatar video awake. Also a moderate correlation score of -0,41 was detected between age and usefulness, that displays the older the test person the less useful the avatar video seemed. Another moderate correlation score of 0,49 was revealed between the age when sign language was learned and the score how much the difference in appearance of interpreter and speaker bothered the test person. Therefore it can be said, the later sign language was learned the more did the difference upset the test person.

Complementary to the numeric values and scores also five qualitative answers were collected. Three of them mentioned that the translation with the avatar was not natural and understandable enough. One of them mentioned, that the avatars could be good because of the neutrality, if they were created less exhausting. Another one described that the idea is nice, but it needs to be implemented in a better way. One person also described, that it is not really necessary that the speaker and the interpreter look the same, especially for topics like news. Even though the test person thinks it can be nice for some videos to use such kind of avatars, especially when translating several people.

These qualitative answers lead to the assumption that the results of the study may look different if the animation of the avatars looked more natural and was more understandable.

The answer, which told, that it was not necessary but might even though be nice for some videos, shows that the avatars might not be necessary for a good usability, but if well done could lead to a better experience.

6 Discussion

In this paper it was shown that the influence of an appearance adaption of SLAs to the speakers is an interesting research topic. It is not only interesting for computer generated avatar animations, but also for motion capture generated, or handmade animations. When talking to interpreters to get help with the questionnaire translation it became clear that there are not enough capacities to care for everything that needs to be translated. Even though the impression evoked that avatars were seen as a threat by the interpreters. Additionally translating online-videos is considered a problem by many researchers, as most of the interpreters do not want to have videos of themselves online, or request extra fees for it. Therefore avatars made with motion capture technology by interpreters could be a possibility to be able to make more translations, protect the interpreters privacy and counteract the perception of avatars as a threat. To evaluate whether avatars adapted to the speakers appearance can have a positive influence on UX prototype videos were created with handmade sign language animations. These were compared to the original interpreter videos in an online-survey with an adapted MeCue questionnaire. The results showed that the avatar videos had significant worse UX scores than the originals, which could issue from the bad scores for usability indicating that the animations were not comprehensible enough. Therefore the results might look completely different if the animations were better, which is also indicated by the qualitative answers participants gave. Additionally the study showed that users are mainly not bothered by the different appearance of speaker and interpreter and in average also do not experience the translation of multiple speakers by one person as annoying. Even though adapted avatars could be a possibility to enhance a neutral experience to a more positive one, with natural and understandable avatars.

6.1 Limitations

This research projects has some limitations, that should be mentioned at this point. The researchers themselves were not part of the Deaf community and also not capable of using sign language. As there was also no possibility to find an interpreter to translate the questionnaire this influences the study results. The study was therefore conducted in GD and not perfectly accessible for people with low skills in this language. In this way the main target group and part of the Deaf community benefiting from sign language translations and also preferring this style of dubbing was unrepresented in this study, which influencing the results. Additionally the animations were made by a person without sign language skills, just following the original video as detailed as possible, which can still not provide a perfect translation with the avatar. This might be the main reason for the negative usability scores and is therefore a big limitation for this study. The results could look different if the animations were made by someone with good sign language and animation skills.

6.2 Future Work

In the future research about adapted avatars could be conducted in a project with more resources. With the help of professional interpreters and animation artists, videos with animations, which are better comprehensible, could be created. In a survey with sign language translation the results could be reviewed and new findings could be generated. If this project shows that indeed the

UX can be improved with adapted avatars, when the animations were appealing and understandable, this builds a basis for future guidelines for avatar design. This can not only be used for computer generated animations, but could also be applied for motion capture translations. A motion capture system for interpreters could be developed, where they can perform their translations and they are transformed into animations with fitting characters. In this way the privacy of interpreters can be protected and new possibilities of one interpreter translating multiple people with fitting avatars or multiple interpreters translating one speaker with one avatar can be established.

References

[1] Nicoletta Adamo and Saikiran Anasingaraju. "Toward the Ideal Signing Avatar". In: *EAI Endorsed Transactions on e-Learning* 3 (June 2016), p. 151446. DOI: 10.4108/eai.15-6-2016.151446.

[2] Rostand Costa et al. "Towards an open platform for machine translation of spoken languages into sign languages". In: *Machine Translation* 33 (Dec. 2019). DOI: 10.1007/s10590-019-09238-5.

[3] Hannah Dostal and Kimberly Wolbers. "Developing Language and Writing Skills of Deaf and Hard of Hearing Students: A Simultaneous Approach". In: *Literacy Research and Instruction* 53 (May 2014), pp. 245–268. DOI: 10.1080/19388071.2014.907382.

[4] Michael Kipp et al. "Assessing the Deaf User Perspective on Sign Language Avatars". In: *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility*. ASSETS '11. Dundee, Scotland, UK: Association for Computing Machinery, 2011, pp. 107–114. DOI: 10.1145/2049536.2049557.

[5] Simon Lang, Marco Block, and Raúl Rojas. "Sign Language Recognition Using Kinect". In: *Artificial Intelligence and Soft Computing*. Ed. by Leszek Rutkowski et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 394–402.

[6] Wanbo Li, Hang Pu, and Ruijuan Wang. "Sign Language Recognition Based on Computer Vision". In: *2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*. 2021, pp. 919–922. DOI: 10.1109/ICAICA52286.2021.9498024.

[7] Nayan Mehta, Suraj Pai, and Sanjay Singh. "Automated 3D sign language caption generation for video". In: *Universal Access in the Information Society* 19 (Nov. 2020). DOI: 10.1007/s10209-019-00668-9.

[8] Lorna C. Quandt et al. "Attitudes Toward Signing Avatars Vary Depending on Hearing Status, Age of Signed Language Acquisition, and Avatar Type". In: *Frontiers in Psychology* 13 (2022). DOI: 10.3389/fpsyg.2022.730917.

[9] Amy T. Wilson and Rowena E. Winiarczyk. "Mixed Methods Research Strategies With Deaf People: Linguistic and Cultural Challenges Addressed". In: *Journal of Mixed Methods Research* 8.3 (2014), pp. 266–277. DOI: 10.1177/1558689814527943.

[10] Yuancheng Ye et al. "Recognizing American Sign Language Gestures from Within Continuous Videos". In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2018, pp. 2145–214509. DOI: 10.1109/CVPRW.2018.00280.

©2022 Lara Krautmacher. Licensee Reutlingen University, Germany. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license. http://creativecommons.org/licenses/by/4.0

Correction to Krautmacher (2022)

In the article “Pre-Study: Adaptation of 3D sign language avatars to the speakers appearance and its influence on user experience“ by Lara Krautmacher, there was an error in Table 1 and the belonging text references in the results section for RQ3.

In the third column of Table 1 named “Avatar”, median values were inserted instead of mean values. With mean values the table should look like this:

Category	Interpreter	Avatar	t-test
U	4,07	1,9	0,00000
N	4,30	1,77	0,00000
A	4,30	1,83	0,00000
L	2,43	2,97	0,06278
PE	3,77	2,14	0,00000
NE	2,17	3,37	0,00007

The paragraph RQ3 of the “Results” section reflected the error in Table 1. The belonging part of this paragraphs should read as follows:

In the category PE interpreters have a higher value (3,77) than avatars (2,14). Therefore interpreter translations lead to an over average and avatars to an under average positive emotional response with a significant difference between the values. For NE interpreter videos have a result of 2,17, which is significantly lower than the score 3,37 of the avatar videos. Therefore the test persons had a under average negative load of emotions with the interpreters and an over average negative amount of emotions with the avatar version. The only category in which the difference between the videos had not been significant is L. Interpreters gained a under average score with 2,43 and avatars an only a little under average score

Untersuchung der Verbesserung der User Experience durch maschinelles Lernen für Bewerbende auf Bewerbungsplattformen

Hannah Kruck
Hochschule Reutlingen
Reutlingen, Deutschland
Hannah.Kruck@Student.Reutlingen-University.de

Zusammenfassung

Bewerbungsplattformen und soziale Netzwerke wie LinkedIn werden immer häufiger für die Jobsuche verwendet und der größte Anteil des Recruiting-Budgets wird von Unternehmen für Stellenanzeigen im Internet verwendet. Es werden bereits einige Techniken des maschinellen Lernens eingesetzt, um den Bewerbungsprozess von Recruitern und Bewerbenden zu vereinfachen. Um die Prozesse der Bewerbenden zu vereinfachen werden bereits das Parsing von Lebensläufen oder One-Klick Bewerbungen eingesetzt. Ziel der Arbeit ist die Analyse der Methoden des Machine Learnings, die zur Verbesserung der User Experience von Bewerbungsplattformen beitragen sollen. Hierbei wird vor allem die Sicht der Bewerbenden betrachtet. Aktuelle und neue Ansätze werden untersucht, um daraus Handlungsempfehlungen abzuleiten und diese durch eine Umfrage mit Studierenden zu validieren. Die Untersuchungen haben gezeigt, dass Bewerbenden häufig die Transparenz auf Bewerbungsplattformen fehlt.

CCS Concepts

• **Human-centered computing** → *HCI design and evaluation methods*; • **Information systems** → **Users and interactive retrieval**.

Keywords

Machine Learning, User Experience, Online Recruiting, Candidate Experience

1 Einleitung

Durch den Ausbruch des SARS-CoV-2 Virus ist deutlich geworden wie wichtig digitale Prozesse mittlerweile für die Gesellschaft sind. So auch der Bewerbungsprozess. Mittlerweile gibt es unzählige Bewerbungsplattformen, ob firmenspezifisch oder für allgemeine Stellenausschreibungen unterschiedlicher Unternehmen. Auch soziale Netzwerke, die sich auf Beruf und Karriere spezialisieren, werden häufig für Bewerbungen genutzt. So ist es keine Überraschung, dass der größte Anteil des Recruiting-Budgets auf Stellenanzeigen im Internet entfällt [13]. Sobald ein digitales System entwickelt wird, muss auch ein positives Benutzererlebnis (engl. User Experience, abgekürzt mit UX) angestrebt werden. Im Zusammenhang mit Bewerbungen entstand hierzu der Begriff Candidate Experience (CX). Dieser Begriff bezeichnet „den Gesamteindruck, den ein potenzieller Bewerber im Rahmen der Prozesse des Recruitings und vom potenziellen Arbeitgeber erhält." [15, S. 12]. Beim Recruiting gibt es unterschiedliche Dimensionen zu berücksichtigen. Zum einen gibt es die Dimension von ArbeitgeberInnen hin zu den Bewerbenden. Diese beschäftigt

Betreuer Hochschule: Prof. Dr. Gabriela Tullius
Hochschule Reutlingen
Gabriela.Tullius@Reutlingen-University.de
Informatics Inside Frühjahr 2022
11. Mai 2022, Hochschule Reutlingen
Copyright 2022 Hannah Kruck

sich damit, wie passende KandidatInnen erreicht werden können. Diese umfassen Optimierungen des Recruitings. Der Fokus dieser Arbeit liegt jedoch auf den Prozessen der BewerberInnen, wie sie beispielsweise ihren Lebenslauf systemgestützt optimieren können. Es umfasst alle Maßnahmen, die das Ziel haben, die Prozesse von BewerberInnen in Richtung potenzieller ArbeitgeberInnen zu optimieren [17, S. 121].

1.1 Ziel der Arbeit

In dieser Arbeit sollen Vorschläge erarbeitet werden, wie die UX von Bewerbungsplattformen mithilfe von Machine Learning (ML) Methoden verbessert werden kann. Hierbei soll vor allem die Sicht der Bewerbenden vom Schreiben von Lebensläufen bis hin zum Feedback zu der gesendeten Bewerbung betrachtet werden. Folgende zentrale Leitfragen sollen unterstützend beantwortet werden, um passende Handlungsempfehlungen zu geben:

(1) Was macht eine gute UX beim Bewerbungsprozess aus?
(2) Wie kann eine Arbeitssuchende Person passende Vorschläge erhalten?
(3) Wie kann eine Arbeitssuchende Person motiviert werden, sich auf vorgeschlagene Jobs zu bewerben?

1.2 Methodik

In einer Literaturrecherche werden aktuell verwendete Ansätze des ML, die bereits im Zusammenhang mit der CX angewandt werden untersucht, aber auch vom Anwendungsfall losgelöste Methodiken des ML betrachtet. In einer Umfrage mit Studierenden wird die Wahrnehmung von ML im Bewerbungsprozess abgefragt. Zusätzlich werden die beliebtesten Bewerbungsplattformen explorativ untersucht. Aus der eigenen Untersuchung, den Ergebnissen der Umfrage und den Ergebnissen anderer Evaluationen von Bewerbungsplattformen werden Handlungsempfehlungen zum Einsatz des ML abgeleitet.

1.3 Aufbau der Arbeit

Zunächst werden einige Grundlagen im Bereich Recruiting, UX und ML zum Verständnis der Arbeit in Kapitel 2 erläutert. In Kapitel 3 wird auf die UX Probleme eingegangen, die bei den gängigen Bewerbungsplattformen auftreten. Weiterhin wird in Kapitel 4 auf die aktuell verwendeten Methoden des ML im Bewerbungsprozess eingegangen und in Kapitel 5 neue Ansätze vorgestellt. Anschließend werden in Kapitel 6 die Ergebnisse aus der eigenen Umfrage dargestellt. Aus den Untersuchungen und den neuen Ansätzen werden anschließend in Kapitel 7 Handlungsempfehlungen aus den vorigen Kapiteln abgeleitet. Zuletzt werden die Ergebnisse in Kapitel 8 kritisch betrachtet und ein Fazit gezogen.

2 Grundlagen

Der Begriff *Machine Learning* beschreibt Algorithmen, die sich durch Erfahrungswerte automatisch anpassen und verbessern. Die Anwendungsbereiche sind vielfältig und reichen von Datamining-Programmen bis hin zur Filterung von Informationen [9]. So findet es bereits den Einsatz in Bewerbungsplattformen wie in Kapitel 4 näher beschrieben wird.

Das *Online-Recruiting* beschreibt die Personalbeschaffung im Internet. Laut einer Statistik von 2019 wird die Mehrheit der MitarbeiterInnen bei großen deutschen Unternehmen über Stellenbörsen im Internet oder deren Unternehmenswebsite rekrutiert [13]. Beim Recruiting kann laut Verhoefen zwischen unterschiedlichen Dimensionen unterschieden werden [17]. Zum einen gibt es die Prozesse und Maßnahmen von Arbeitgebenden zu Bewerbenden, welche die Streuung von Stellenanzeigen, Werbemaßnahmen und KI-basierte Lösungen beinhalten. Interne Prozesse der Arbeitgebenden beschreiben beispielsweise automatisierte Matchingsysteme oder Rating-Lösungen. Zum anderen gibt es

die Prozesse und Maßnahmen von der Seite der Bewerbenden. Von den Bewerbenden zu Arbeitgebenden sind das Prozesse wie beispielsweise die Optimierung der Suche durch Chatbots oder das Matching auf Arbeitgeberseiten. Zu guter Letzt beschreiben die externen Prozesse der Bewerbenden diejenigen, die sich auf die Bewerbenden selbst beziehen. Dies wäre zum Beispiel die Optimierung des eigenen Lebenslaufes durch ein System [16, S. 121]. In dieser Arbeit liegt der Fokus auf den externen Prozessen der Bewerbenden und den Prozessen von Bewerbenden hin zu Arbeitgebenden. Im Zusammenhang mit diesen Prozessen entstand der Begriff *Candidate Experience* (CX) durch Verhoeven [15, S. 12], abgeleitet und erweitert von der Definition durch Kootz [7].

3 UX Probleme und Faktoren

Um zu klären, welche Faktoren die UX eines Bewerbungprozesses zu einer guten oder einer schlechten Erfahrung machen, muss zunächst geklärt werden, was eine gute UX einer Bewerbungsplattform ausmacht. Die Herausforderungen die Bewerbende bei der Nutzung von Bewerbungsplattformen haben, müssen identifiziert werden. Hierzu werden Studien hinzugezogen, die das Thema UX im Bewerbungsprozess adressieren. Außerdem werden die Grundsätze der Dialoggestaltung einbezogen und ein Modell betrachtet welches die Faktoren des individuellen Suchverhaltens aufzeigen soll.

3.1 Faktoren einer guten UX

In einer Studie von Jeff Sauro [11] aus dem Jahr 2019 wurden 213 Personen zu ihrer Erfahrung mit Bewerbungsplattformen befragt. Im Nachgang wurden 38 Personen näher zu ihren Erfahrungen und ihrem Erfolg bei der Jobsuche mit den Plattformen befragt. Aus den Ergebnissen der Befragungen wurden folgende wichtige Faktoren abgeleitet:

Ein wichtiger Faktor für eine gute UX ist laut Sauro das *Vertrauen* in die angezeigten Angebote. Einige Teilnehmende beklagten unechte oder unseriös wirkende Jobangebote und Spam E-Mails nach der Anmeldung auf einer der Plattformen. Der größte Faktor, über den die Teilnehmenden unzufrieden waren, sind irrelevante Job-Vorschläge aufgrund fehlender Filtermöglichkeiten. Somit empfanden die TeilnehmerInnen die *Zuverlässigkeit* der Plattform als nicht ausreichend. Außerdem werden unpassende Jobs angezeigt, die aufgrund von in den Stellenbeschreibungen verwendeten Begriffen von der Software als passend erachtet werden. Ein Nutzer merkte an, dass die Website möglicherweise mit Hilfe von KI besser erkennen könne, welche Jobs relevant sind, obwohl einige Schlüsselwörter nicht genau übereinstimmen. Viele NutzerInnen wünschten sich mehr *Transparenz*. Konkret wünschten sie sich mehr Information zum Gehalt, wobei hier angemerkt sei, dass es hier auch kulturelle Unterschiede in verschiedenen Ländern geben kann. Insgesamt ist bei der Integration von ML darauf zu achten, den NutzerInnen genügen Auskunft zu geben. Sauro hebt hervor, dass KI Methoden dazu neigen, als Blackboxen zu agieren und dem Benutzer nicht genügend Transparenz, Kontrolle und Interaktionsmöglichkeiten zu bieten, die jedoch als Hauptziele des nutzerzentrierten Designs erachtet werden [11].
Außerdem sollen NutzerInnen stets den Grund für eine erforderliche Aktion ersichtlich sein. In [14] fiel Venkatesh auf, dass in der Evaluation einer Bewerbungsplattform mit NutzerInnen eine der größten Frustrationen bei der Nutzung der Plattform der langwierige Registrierungsprozess ist, bevor überhaupt nach Stellenangeboten gesucht werden kann.

3.2 Usability

Zur Messung der Usability können zudem die sieben Grundsätze der Dialoggestaltung [4] hinzugezogen werden, da eine gut gestaltete

Benutzeroberfläche ebenfalls die Erfahrung bei der Nutzung beeinflussen kann. Bezogen auf Bewerbungsplattformen lassen sich den unterschiedlichen Grundsätzen positive Eigenschaften zuordnen, die erfüllt sein müssen. Diese wurden aus den eigenen Erfahrungen beim explorativen Untersuchen unterschiedlicher Plattformen erhoben. Damit die *Aufgabenangemessenheit* erfüllt ist, sollten beispielsweise bei Auswahlmöglichkeiten auch passende Optionen verfügbar sein. Sinnvolle Standardwerte können zum Beispiel in Dropdown-Feldern beim Eingeben vorgeschlagen werden um Zeit zu sparen, jedoch sollten aber auch weitere individuelle Begriffe möglich sein. Die Plattform muss *Selbstbeschreibungsfähig* sein. Dafür ist es für NutzerInnen hilfreich, wenn immer ersichtlich ist, bei welchem Schritt sie sich gerade im Bewerbungsprozess befinden. Dies gilt nicht nur für das Korrigieren des Ergebnisses des CV Parsings, sondern auch noch nach dem Absenden der Bewerbung. So ist zum Beispiel eine Benachrichtigung hilfreich, dass die Bewerbung die zuständigen Personen erreicht hat und die Bewerbung geprüft wird.
Um die *Erwartungskonformität* zu erfüllen können zur Registrierung auf der Plattform die gängigen Möglichkeiten angeboten werden. Zu diesen gehört eine Registrierung mit Google oder E-Mail. Damit die Plattform *fehlertolerant* ist, sollten Fehler, die NutzerInnen machen, markiert werden, um darauf aufmerksam zu machen. *Steuerbar* ist eine Bewerbungsplattform dann, wenn der Zwischenstand einer Bewerbung jederzeit speicherbar ist um an einem anderen Zeitpunkt damit fortzufahren. Außerdem müssen NutzerInnen jederzeit zu unterschiedlichen Schritten zurückkehren können. Die *Individualisierbarkeit* ist dann gegeben, wenn NutzerInnen die Möglichkeit dazu haben, spezifische Angaben leicht selbst tätigen können. Als Beispiel eignet sich hier die Angabe über das gewünschte Berufsfeld, da hier neue Berufsfelder der letzten Jahre oft nicht aufgeführt sind. Oft müssen NutzerInnen zudem Kompromisse schließen und einen Studiengang angeben, den sie gar nicht abgeschlossen haben, sondern nur einen, der dem begrifflich ähnelt. Singh merkt in [12] ebenfalls an, dass Lebensläufe teilweise selbst durch Dropdownfelder ausgefüllt werden müssen. Häufig gibt es hier nicht genügend Auswahl, sodass unpassende Alternativen ausgewählt werden [12]. Dies bietet NutzerInnen keine Kontrolle und wenig Freiheit. Eine Plattform ist dann *lernförderlich*, wenn Formularfelder zum Ausfüllen erklärt werden und Tipps gegeben werden, damit Bewerbende auch bei zukünftigen Bewerbungen dazu lernen und ihre Erfolgschancen somit erhöhen.

3.3 Faktoren des Suchverhaltens

In [6] untersuchen Kleemann et al. das Verhalten von NutzerInnen bei der Verwendung von Filtermechanismen, Recommender Systemen und Beratungsfunktionen in einem Online Shop. Im Beitrag wird anschließend ein Modell abgeleitet, welches Auskunft darüber geben soll, welche Methode den NutzerInnen am besten bei Entscheidungen hilft. Sie berücksichtigen hierbei Faktoren wie Charaktereigenschaften, demographische Daten und Situationsfaktoren [6, S. 500]. Die untenstehende Abbildung 1 wurde aus dem Modell der Autoren abgeleitet und in den Bereich des Online-Recruitings übertragen.
Der Einfluss der *Persönlichkeit* äußert sich darin, dass sie unter anderem die Akzeptanz der Vorschläge stark beeinflusst. Zudem sehr starken Einfluss hat der Wille, ein passendes Suchergebnis zu bekommen. Die Autoren beschreiben diese Charakteristik als *Maximization* [6]. Sogenannte *Maximizers* investieren viel Zeit um das bestmögliche Ergebnis

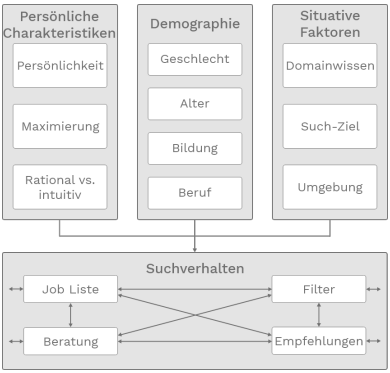


Abbildung 1: Faktoren des Suchverhaltens und Unterstützungsfunktionen, abgeleitet aus [6, S. 500]

zu finden. Im Gegensatz dazu geben sich *Sacrificers* eher schneller mit einem Ergebnis zufrieden, das zumindest teilweise ihren Ansprüchen genügt. Bei der Entscheidungsfindung unterscheiden die Autoren in rational oder intuitiv. Intuitiv gesteuerte Menschen lassen sich eher von den Suchergebnissen beeinflussen während rationale NutzerInnen eher analytischen Suchmethoden und ihrer ursprünglichen Intention folgen.
Ebenso wie die persönlichen Faktoren haben auch die *demographischen Daten* einen Einfluss auf das Suchverhalten. Laut den Autoren beeinflusst das Geschlecht das Suchverhalten insofern, dass sich Frauen beispielsweise mehr von der Darstellung eines Produktes beeinflussen lassen. Ältere Menschen und Männer setzen sich eher mit automatischen Vorschlägen bei der Suche auseinander. Der Bildungsgrad und der Beruf haben ebenso einen Einfluss [6]. Außerdem fiel in einer Evaluierung einer Bewerbungsplattform durch Venkatesh in [14] auf, dass erfahrene Arbeitssuchende gerne vorab durch die Ergebnisse stöbern um sich einen schnellen Überblick über das Angebot auf dem Arbeitsmarkt zu verschaffen.

Ein wichtiger Faktor der *situativen Faktoren* ist das Domänenwissen, welches NutzerInnen stark bei der Entscheidung und der Geschwindigkeit der Entscheidung beeinflusst [6]. Das Such-Ziel beeinflusst das Suchverhalten enorm. Es kann hier zwischen explorativer- und zielgerichteter Suche unterschieden werden. Mit der Umgebung ist das User Interface (UI) gemeint. Ein passend gestaltetes UI unterstützt NutzerInnen bei der Suche und kann die aufgewendete Zeit reduzieren. Die Autoren in [6] nannten als Argument, dass das Kaufverhalten stark von der dargestellten Information und der genutzten Technologie abhängig ist.
Die Liste der Jobs wird beeinflusst durch manuelle Filter, welche die NutzerInnen anwenden (siehe 1 im unteren Bereich). Je nachdem ob die angebotenen Stellenausschreibungen den Ansprüchen entsprechen, muss ein Filter angewandt werden. Die Beratung wird dargestellt durch einen Chatbot, der durch Konversation mit NutzerInnen die Angebote beeinflusst. Die Empfehlungen werden automatisiert gegeben. Die Schlussfolgerung der Untersuchung der Autoren in [6] ist, dass nicht jede Unterstützungsfunktion allen NutzerInnen in jeder Situation gleich unterstützen.

4 State-of-the-Art: Machine Learning im Bewerbungsprozess

Im Folgenden werden Methoden vorgestellt, die bereits Anwendung auf Bewerbungsplattformen finden.

4.1 CV Parsing

Das Curriculum Vitae (CV) Parsing ist eine ML Methode die bereits häufig eingesetzt wird und basiert auf natürlichen Sprachverarbeitungstechniken, um Muster in Lebensläufen zu erkennen [15, S. 96]. Dadurch kann solch ein System zum Beispiel deutsche

Adressen sehr einfach erkennen, da sie üblicherweise immer dieselbe Struktur besitzen [15, S. 100].
In [1] beschreiben Cernian et al. ein CV Parsing Modul, welches PDF Lebensläufe von BewerberInnen zunächst in Textdateien umwandelt, und dann den Inhalt in verschiedene Abschnitte wie persönliche Daten oder Arbeitserfahrung gliedert. Dann werden die Wörter in den Abschnitten klassifiziert. Nach der Klassifizierung werden die Eigenschaften in Bezug auf Kompetenzen, Fähigkeiten und Arbeitserfahrung der KandidatInnen analysiert. Anschließend wird ein Wert in Relation zu den Anforderungen der Stellenausschreibung für jeden Lebenslauf ermittelt, anhand dessen die Lebensläufe sortiert, und Recruitern präsentiert werden [1, S. 2-5].
In einem eigenen Versuch wurde das CV Parsing zweier bekannter Plattformen getestet. Auf der Plattform Indeed werden NutzerInnen auf unterschiedliche Plattformen weitergeleitet, über welche die Bewerbung abgeschlossen wird. Im Falle des Versuchs fand das CV Parsing über die Plattform *Successfactors* statt. Hier fiel auf, dass ausführliche Angaben nicht korrekt vom System erkannt werden. Im konkreten Fall konnte das System die unterschiedlichen Stichpunkte aus der Beschreibung des Bachelorstudiengangs nicht gut unterscheiden.
Über die Plattform *Monster* wurden beim CV Parsing sehr viele Qualifikationen erkannt. Diese konnten im Anschluss nach Relevanz geordnet werden und zu jeder Qualifikation konnte das Kenntnissniveau eingestellt werden. Negativ aufgefallen ist, dass zu viele Qualifikationen aus den Stichpunkten und Beschreibungen im Lebenslauf erkannt wurden. Es wurden Begriffe wie *CMS* oder *UX* als Qualifikation zugeordnet, aber auch Begriffe wie *Bereich* und *Unterstützung*. Hier muss im Anschluss an das Parsing einiges verbessert werden. Ursache für die fehlerhaften Parsingergebnisse könnte sein, dass das System nicht auf zwei- oder mehrspaltige Dokumente trainiert wurde.

4.2 One-Klick Bewerbungen

Ebenfalls in Bewerbungsplattformen vertreten, häufig in sozialen Netzwerken wie LinkedIn, ist das Prinzip der sogenannten One-Click Bewerbung. Die Bewerbung kann mit dem eigenen Lebenslauf oder dem Profil auf sozialen Netzwerken in bis zu drei Schritten abgeschlossen werden [15, S. 100]. Voraussetzung hierfür ist ein Profil im jeweiligen sozialen Netzwerk und die Aktualität der Angaben des Lebenslaufs und der Fähigkeiten.

5 Neue Machine Learning Ansätze zur Verbesserung der UX

Im folgenden Kapitel werden verschiedene Ansätze aus der Literatur vorgestellt, die ML im Bereich des Recruitings anwenden. Es gibt zahlreiche ML Methoden im Recruiting die bereits verwendet werden. Allerdings werden hier nur diejenigen vorgestellt, die auch für die Bewerbenden relevant sind. Zusätzlich werden eigene Überlegungen aus anderen Domänen vorgestellt, die zur Verbesserung der UX beitragen können.

5.1 Klassifikation der Fähigkeiten

Skill2Vec wurde von Le et al. [8] entwickelt und ist ein Ansatz der Natural Language Processing (NLP) verwendet, um die Suchstrategie nach KandidatInnen zu verbessern, indem die Fähigkeiten analysiert werden. Die Autoren haben ein Netzwerk von Beziehungen zwischen Fähigkeiten unter Verwendung eines neuronalen Netzes entwickelt. Laut den Autoren ist es möglich, qualitativ hochwertige Wortvektoren aus einem großen Datensatz zu berechnen.
Dieser Ansatz behandelt also die Herausforderungen, die Recruiter bei den Prozessen

der Personalbeschaffung haben, wie in Kapitel 2 beschrieben, und nicht die der Arbeitssuchenden. Dennoch kann der Ansatz auch in der anderen Richtung Anwendung finden, indem die Fähigkeiten aus der Bewerbung vorab geprüft werden (möglicherweise durch zusätzliche Prüfung und Verbesserung der erkannten Fähigkeiten durch die Bewerbenden selbst) und dann auf der Plattform passende Jobvorschläge angezeigt werden.

5.2 Klassifikation der Stellenbezeichnungen

Das Job Title Benchmarking (JTB) von Zhang et al. [18] zielt darauf ab, Stellenbezeichnungen mit ähnlichen Kompetenzen in verschiedenen Unternehmen abzugleichen. Das JTB könnte sowohl für die Recruiter als auch für Bewerbende bei der Vorhersage von Positionen und Gehältern eine Orientierung bieten. Die Daten zum Jobwechsel sind laut den Autoren unübersichtlich, da viele Stellen unterschiedliche mögliche Jobtitel besitzen. Als Beispiel nennen die Autoren den Beruf des Software Engineers. Hierfür werden in unterschiedlichen Unternehmen ähnliche Titel wie Software Development Engineer oder Computer Programmer verwendet [18, S. 1-2]. Dies erschwert die Suche für BewerberInnen. Ein weiteres Problem stellt die Datenmenge in sozialen Netzwerken wie LinkedIn dar, die die Anwendung von Data Mining Algorithmen erschwert. Als dritte Herausforderung wird gesehen, dass eine Person nur nach einer beschränkten Anzahl an Stellenbezeichnungen sucht, im Verhältnis zu den ausgeschriebenen Stellen.
Einen weiteren Ansatz zur Klassifikation der Stellenbezeichnung stellen Javed et al. in [5] vor. Die Job Title Classification ist eine Entwicklung eines auf maschinellem Lernen basierenden, semi-supervised Klassifikationssystems für das Online-Recruiting, bei dem die Titel der Stellenbeschreibungen genauer klassifiziert werden können.

5.3 Klassifikation mehrerer Einheiten

Die Google Cloud Jobs API unterscheidet laut Posse in Jobtitel, Beschreibungen, Fähigkeiten und Präferenzen und gleicht die Präferenzen von Arbeitssuchenden mit relevanten Stellenangeboten auf der Grundlage von Klassifizierungen und relationaler Modelle ab [10]. Das maschinelle Lernen realisiert die Erkennung und Extraktion von Berufen in Anfragen von Arbeitssuchenden und ordnet sie den Knoten in der Berufsontologie zu, ebenso wie die Bezeichnungen der Stellenausschreibungen. Außerdem werden Fähigkeiten in Anfragen von Arbeitssuchenden und Stellenausschreibungen erkannt und ebensfalls zu den Knoten der Fähigkeiten zugeordnet. Bevor dies geschieht, müssen die Titel der Stellenanzeigen standardisiert werden. Zunächst werden die Titel von allen irrelevanten Formulierungen bereinigt wie Firmenname oder Referenz-IDs. Danach wird die Vektorisierung der bereinigten Jobtitel berechnet. Zuletzt wird ein Jobtitel kodiert: Mehrere Klassifikatoren versuchen, den Titel auf null, einen oder mehrere Knoten abzubilden. So lässt sich beispielsweise die Berufsbezeichnung *Einzelhandelskaufmann* der umfassenden Kategorie *Verkauf und Einzelhandel* zuordnen.

5.4 Vorherige Lebenslauf Prüfung

In [2] wurden Menschen von Dillahunt et al., die arbeitssuchend sind und schlechte Aussichten bei der Arbeitssuche haben, befragt welche Probleme bei Bewerbungen bestehen. In [3] wurde daraufhin das Problem adressiert, dass häufig nicht die Arbeitssuche das Problem sei, sondern das fehlende Feedback zu den Lebensläufen. Daraufhin wurde eine Anwendung implementiert, in der Freiwillige Feedback zu Lebensläufen geben können. Diese Anwendung könnte mithilfe

von ML Ansätzen weiter digitalisiert werden. Lebensläufe müssten nicht durch Menschen analysiert werden, sondern durch ein System, das beispielsweise auf Rechtschreibung, Lücken im Lebenslauf oder fehlende Informationen hinweisen. Dies könnte mithilfe des CV Parsings realisiert werden und eine Rechtschreibkorrektur könnte beispielsweise mittels Deep Learning realisiert werden. So können Arbeitssuchende unterstützt werden, ihre Chancen auf eine Einladung zum Bewerbungsgespräch zu erhöhen.

5.5 Chatbots

Eine weitere Verbesserung der UX wäre es, wenn Bewerbende anstelle oder vor der Eingabe einer Suchanfrage durch Konversation mit einem Chatbot die Suche eingrenzen kann. So könnten beispielsweise mehrere ähnliche Jobtitel aus den vorhandenen Jobtiteln auf der Plattform abgefragt werden, ob sie passend wären oder nicht, oder die Wünsche der Bewerbenden abgefragt werden. So kann ein System individuell lernen was der Bewerbende wünscht, und was nicht. Zusätzlich könnte in der Liste der vorgeschlagenen Stellen manuell angegeben werden können, ob mehr Jobs dieser Art vorgeschlagen werden sollen, oder nicht. Automatisiert könnte das ebenso durch die gesendeten Bewerbungen einer Person geprüft werden. Sendet die Person auf eine Stellenanzeige hin eine Bewerbung, könnte das System lernen, dass dieser Job scheinbar relevant für die Person ist.

6 Umfrage

In der Umfrage sollen Informationen darüber gesammelt werden, welche Erfahrungen von Absolvierenden bereits mit Bewerbungsplattformen gesammelt wurden.

6.1 Fragebogendesign

Die Gruppe der Teilnehmenden wurde bewusst auf Absolvierende eingegrenzt, um vergleichbarere Ergebnisse zu erzielen. Denn wie durch das Modell zum Suchverhalten von [6] verdeutlicht wurde, fällt das Suchverhalten und die Wahrnehmung bei unterschiedlichen demographischen Faktoren unterschiedlich aus. In der Umfrage werden Fragen zu den wichtigen Faktoren gestellt, die die Erfahrung mit Bewerbungsplattformen beeinflussen (siehe Kapitel 3 [11]). Genauer sind das Vertrauen, Transparenz und Zuverlässigkeit. Um die Ergebnisse besser einordnen zu können wurde zudem am Ende des Fragebogens der Studienhintergrund und das Alter der Person abgefragt. Es wurde hier die Annahme getroffen, dass Studierende mit Informatikhintergrund den Einsatz von ML Methoden anders einschätzen und mehr vertrauen in die Systeme haben, als Studierende anderer Fakultäten.

6.2 Auswertung

Insgesamt nahmen 20 Studierende an der Umfrage teil. Zunächst wurden Fragen zu den bisherigen Erfahrungen mit Bewerbungsplattformen gestellt. 70 % der Teilnehmenden gaben an, dass sie im Allgemeinen zufrieden mit dem Bewerbungsprozess waren. Unter den Plattformen, mittels welcher sich die Studierenden bewarben, wurden am häufigsten LinkedIn, Xing und Stepstone genannt. 60 % der Teilnehmenden nannten zudem Firmenspezifische Plattformen wie zum Beispiel Smartrecruiters von Bosch. In einem weiteren Fragenblock ging es um die Transparenz von Machine-Learning Methoden. Knapp die Hälfte der Teilnehmenden gaben an, dass sie nicht wissen, ob bei ihrer Bewerbung Machine-Learning Methoden zum Einsatz kamen. Teilnehmede, die wussten, dass solche Methoden Anwendung fanden, nannten One-Klick Bewerbungen und CV-Parsing. Im letzten inhaltlichen Abschnitt wurden Fragen gestellt, die auf die Akzeptanz von bisher unbekannten Methoden abzielte. 70 % der Teilnehmenden würden den Einsatz der vorgeschlagenen Methoden (siehe Kapitel 5)

begrüßen, haben größtenteils großes Vertrauen in ML und würden die Funktionen nutzen.

6.3 Ergebnisse

Auffällig war, dass häufig das CV Parsing als hilfreich erachtet wurde. Auch die Schnelligkeit bei One-Klick Bewerbungen auf LinkedIn wurde betont. Als Problem stellte sich heraus, dass die Suche erschwert wird durch unterschiedliche Bezeichnungen desselben Berufs. Insgesamt war das Vertrauen in ML sehr groß und unbekannte Methoden trafen auf Akzeptanz. Gewünscht wurde aber auch, dass transparent aufgezeigt werden muss, wie Ergebnisse zustandekommen.

7 Handlungsempfehlungen

Generell sollte darauf geachtet werden, dass für eine gute UX auch eine angemessen gestaltetes UI nötig ist. Um die Frage zu beantworten, was eine gute UX beim Bewerbungsprozess ausmacht ist eine konkrete Handlungsempfehlung, stets die NutzerInnen zu informieren wofür sie ihre Daten preisgeben. Wie in Kapitel 3 beschrieben, verstanden NutzerInnen einer Evaluation aus [14] nicht, warum der Registrierungsprozess langwierig ist, bevor überhaupt nach Stellenangeboten gesucht werden kann. Hier sollte den NutzerInnen in angemessener Form erklärt werden, warum die Angabe der Daten wichtig ist und zu ihrem Erfolg beitragen kann. Bei Auswahlfeldern sollte bei der Berufsauswahl für den gesuchten Job gängige Beschreibungen angeboten werden. Diese können aus den Daten bisheriger Nutzerregistrierungen entnommen werden und durch Klassifizierungssysteme korrekt eingeordnet werden, damit auch neue Berufsfelder in der Auswahl möglich sind und NutzerInnen keine Kompromisse eingehen müssen. So können auch bereits für Recruiter eingesetzte Matchingsysteme für Bewerbende angewandt werden, um ihre Erfolgschancen zu erhöhen und passendere Vorschläge erhalten. Die Frage, wie Arbeitssuchende motiviert werden, sich auf vorgeschlagene Jobs zu bewerben, konnte durch die Umfrage und die betrachteten Paper beantwortet werden. Durch die Umfrageergebnisse konnte bestätigt werden, dass NutzerInnen motiviert werden, sich zu bewerben, wenn sichtbar ist, dass die eigenen Angaben aus der Bewerbung mit den Angaben aus der Stellenanzeige eine hohe Übereinstimmung haben. Dies kann auch durch einen Fragebogen bei der Registrierung oder einen Chatbot abgefragt werden. Die Suchergebnisse und die Vorschläge durch einen Chatbot können zudem analysiert werden, ob NutzerInnen sich auf die Vorschläge hin bewerben und ob die Bewerbung erfolgreich war. So könnte die Erfolgsrate von Bewerbungen durch ständige Anpassung der Vorschläge durch ML verbessert werden.

Bereits erfolgreich eingesetzte Methoden wie das CV-Parsing müssen stets verbessert werden, um die Zuverlässigkeit zu verbessern. Wie gezeigt wurde, ist diese noch nicht ausreichend. Wichtig ist dabei, dass NutzerInnen im Falle einer Fehlfunktion angemessen und effizient in die Korrektur ihrer Daten eingreifen können.

Auch One-Klick Bewerbungen können erfolgreich sein und finden großen Anklang, wenn ausreichende Angaben durch NutzerInnen erbracht werden. Auch hier ist es wichtig, die NutzerInnen zu informieren, dass ihre Daten auf den sozialen Netzwerken aktuell und vollständig sind.

Bei ML Methoden muss darauf geachtet werden, die Trainingsdaten zu prüfen. Werden zum Beispiel Daten von zuvor erfolgreich eingestellten Personen verwendet, könnte dies je nach Recruiter variieren und den Datensatz möglicherweise verzerren. Wenn also explizit auf Diversität geachtet werden soll, muss auch der Datensatz divers sein.

8 Fazit & Ausblick

Es wurden Untersuchungen zu der Erfahrung von Bewerbenden angestellt, um diese durch ML Ansätze zu verbessern. Die Untersuchung zeigte, dass eine gute UX beim Bewerbungsprozess unter anderem unter fehlender Transparenz, fehlendem Vertrauen und Zuverlässigkeit leidet. Durch den Einsatz von ML können diese Faktoren verbessert werden. Konkreter bedeutet dies den Einsatz von CV Parsing, One-Klick-Bewerbungen und Klassifizierungssystemen zu fördern. Klassifizierungssysteme können die Titel der Stellenbezeichnungen aber auch den Inhalt betreffen, um das Matching zwischen den Angaben der Bewerbenden und den Stellenausschreibungen zu verbessern. Voraussetzung für die zuverlässige Funktion dieser Systeme sind die ausreichenden Angaben der Daten von NutzerInnen der Plattformen. Über den Einsatz der ML Methoden und den Nutzen müssen Bewerbende ausreichend informiert werden. Die gesammelten Daten erfolgreich rekrutierter KandidatInnen müssen geprüft werden, um die Diversität sicherzustellen. Die Umfrage und die Erkenntnisse aus der Literatur lassen vermuten, dass es möglicherweise zudem eine Abhängigkeit zwischen dem Job-Angebot auf den Plattformen, dem akademischen Grad und dem Berufsfeld gibt. Dies sollte weiter untersucht werden. Das vorgestellte Modell zum Suchverhalten auf Online-Shopping Plattformen wurde als Vorschlag vorgestellt, um Faktoren zu identifizieren, die die Erfahrung mit Bewerbungsplattformen verbessern könnten. Ob das Modell tatsächlich übertragbar ist, muss jedoch geprüft werden. In der Umfrage dieser Arbeit wurde nur eine sehr kleine Stichprobengröße befragt, was die Ergebnisse eingrenzt. Daher muss zusätzlich geprüft werden, ob die Umfrageergebnisse auch übertragbar ist. Bei der Recherche ist aufgefallen, dass in der Literatur ML Ansätze hauptsächlich für Recruiter verwendet werden und weniger für die Prozesse der Bewerbenden. Die Nutzung von ML Methoden könnte nicht nur dazu führen, dass Bewerbende mehr Erfolg und Motivation bei Bewerbungen haben, sondern auch, dass nur die Fähigkeiten der KandidatInnen betrachtet werden, und nicht die Vorlieben der Recruiter. Die Ergebnisse dieser Arbeit zeigen erste Tendenzen, dass neue ML Ansätze gewünscht und auch unter der Annahme ausreichender Transparenz akzeptiert werden.

Literatur

[1] Alexandra Cernian, Dorin Carstoiu und Bogdan Martin. „Semi-automatic Tool for Parsing CVs and Identifying Candidates' Abilities and Competencies". In: *DEStech Transactions on Social Science, Education and Human Science* emass (2017). DOI: 10.12783/dtssehs/emass2016/6802.

[2] Tawanna R. Dillahunt und Amelia R. Malone. „The Promise of the Sharing Economy among Disadvantaged Communities". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Hrsg. von Bo Begole u. a. New York, NY, USA: ACM, 2015, S. 2285–2294. ISBN: 9781450331456. DOI: 10.1145/2702123.2702189.

[3] Tawanna R. Dillahunt u. a. „Designing for Disadvantaged Job Seekers". In: *DIS 2016*. Hrsg. von Marcus Foth. New York, NY: ACM, 2016, S. 905–910. ISBN: 9781450340311. DOI: 10.1145/2901790.2901865.

[4] *DIN EN ISO 9241-110:2020-10, Ergonomie der Mensch-System-Interaktion_-Teil_110: Interaktionsprinzipien (ISO_9241-110:2020); Deutsche Fassung EN_ISO_9241-110:2020*. Berlin. DOI: 10.31030/3147467.

[5] Faizan Javed u. a. *Towards a Job Title Classification System*. 2016. URL: https://arxiv.org/pdf/1606.00917.

[6] Timm Kleemann u. a. „Modeling User Interaction at the Convergence of Filtering Mechanisms, Recommender Algorithms and Advisory Components“. In: *Mensch und Computer 2021*. Hrsg. von Stefan Schneegass, Bastian Pfleging und Dagmar Kern. New York, NY, USA: ACM, 2021, S. 499–511. ISBN: 9781450386456. DOI: 10.1145/3473856.3473859.

[7] Jochen Kootz. „Kundenorientiertes Personalrecruiting. Eine empirische Untersuchung unter besonderer Berücksichtigung von Customer Experience Management“. Doctoralthesis. Universität Trier, 2014. DOI: 10.25353/ubtr-xxxx-d960-6a1f.

[8] Van-Duyet Le, Minh-Quan Vo und Dang Quang-An. *Skill2vec: Machine Learning Approaches for Determining the Relevant Skill from Job Description*. 2017.

[9] Tom M. Mitchell. *Machine learning*. International ed., [Reprint.] McGraw-Hill series in computer science. New York, NY: McGraw-Hill, 2010. ISBN: 0070428077.

[10] Christian Posse. *Cloud Jobs API: machine learning goes to work on job search and discovery*. 2016. URL: https://cloud.google.com/blog/products/gcp/cloud-jobs-api-machine-learning-goes-to-work-on-job-search-and-discovery.

[11] Jeff Sauro. *The UX of Online Job Searching Websites & Apps*. 2019. URL: https://measuringu.com/ux-jobsearching/.

[12] Yogesh Singh. *How better UX for job portals can bring better job applications?* 2019. URL: https://medium.com/@cyogesh56/how-better-ux-for-job-portals-can-bring-better-job-applications-2155d7e4b3b5.

[13] Statista Research Department. *Statistiken zum Thema Online-Recruiting*. 2019. URL: https://de.statista.com/themen/710/online-recruiting/.

[14] Harini Venkatesh. *Search made easy with the JobStreet app — a UX case study*. 2018. URL: https://uxdesign.cc/jobstreet-com-app-revamp-91523cd7eccb.

[15] Tim Verhoeven. *Candidate Experience: Ansätze für eine positiv erlebte Arbeitgebermarke im Bewerbungsprozess und darüber hinaus*. 1. Aufl. s.l.: Springer Gabler, 2015. ISBN: 978-3-658-08895-8. DOI: 10.1007/978-3-658-08896-5.

[16] Tim Verhoeven, Hrsg. *Digitalisierung im Recruiting: Wie sich Recruiting durch künstliche Intelligenz, Algorithmen und Bots verändert*. Wiesbaden: Springer Fachmedien Wiesbaden GmbH und Springer Gabler, 2020. ISBN: 978-3-658-25884-9. DOI: 10.1007/978-3-658-25885-6.

[17] Tim Verhoeven. „Künstliche Intelligenz im Recruiting: Die möglicherweise größte Veränderung des Recruitings steht erst noch bevor und wartet darauf, gestaltet zu werden“. In: *Digitalisierung im Recruiting*. Hrsg. von Tim Verhoeven. Wiesbaden: Springer Fachmedien Wiesbaden GmbH und Springer Gabler, 2020, S. 113–127. ISBN: 978-3-658-25884-9.

[18] Denghui Zhang u. a. „Job2Vec: Job Title Benchmarking with Collective Multi-View Representation Learning“. In: *CIKM'19*. Hrsg. von Wenwu Zhu, Dacheng Tao und Xueqi Cheng. ACM Digital Library. New York: Association for Computing Machinery, 2019, S. 2763–2771. ISBN: 9781450369763. DOI: 10.1145/3357384.3357825.

©2022 Hannah Kruck. 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

Exploding Gradients bei Conv-LSTM-Netzwerken unter Verwendung des moving-MNIST Datensatzes

Kai Müller
Hochschule Reutlingen
Reutlingen, Deutschland
Kai.Mueller@Student.Reutlingen-University.de

Zusammenfassung

Long short-term memory Netze (LSTMs) sind für das Erlernen zeitlicher Zusammenhänge in Daten eine erfolgreich verwendete Architektur. LSTMs sind vom Problem explodierender Gradienten betroffen. Dabei wachsen die Gradienten der Netzwerkgewichte während des Trainings exponentiell. Conv-LSTMs sind eine Architekturvariation von LSTMs, in denen Faltungsoperationen zur besseren Erfassung von Bildinformationen genutzt werden. In dieser Arbeit wird untersucht, inwieweit Conv-LSTM ebenfalls vom Problem explodierender Gradienten betroffen sind. Um diese Frage zu erörtern, wurde ein Conv-LSTM auf Basis des moving-MNIST Datensatzes[1] trainiert. Beim durchgeführten Experiment kam es nicht zu explodierenden Gradienten. Die Arbeit zeigt, dass es für die verwendeten Trainingsgegebenheiten nicht zum Problem explodierender Gradienten kommen muss. Dadurch wird eine Basis für die weitere Untersuchung von Gradienten bei Conv-LSTMs gelegt und das Problem dokumentiert und analysiert. Außerdem wird das komplexe Problem explodierender Gradienten anschaulich dargestellt und mögliche Gründe werden aufgezeigt.

CCS Concepts

• **Computing methodologies → Artificial intelligence**; Supervised learning.

Keywords

Conv-LSTM, explodierende Gradienten, recurrent neural networks, Sequenz Modellierung

[1] https://www.cs.toronto.edu/ nitish/unsupervised_video/

1 Einleitung

Das Verständnis zeitlicher Abläufe ist wichtig für die Lösung verschiedener Probleme im Bereich des maschinellen Lernens. Grundlage hierfür sind sequenzielle Daten, also Daten, die einen zeitlichen Bezug zueinander aufweisen. So basiert in einem Video immer ein Bild auf dem Vorigen. Recurrent neural networks (RNNs), die die long short-term memory (LSTM) Architektur verwenden, wurden erfolgreich für verschiedene überwachte sequenzielle Lernaufgaben eingesetzt, zum Beispiel für die Spracherkennung [8], für die automatisierte Generierung von Untertiteln [2], oder für die Aktionserkennung in Videos [4]. Um aus einer Sequenz an Bildern eine

Betreuer/-in Hochschule	Prof. Dr.-Ing. Cristóbal Curio Hochschule Reutlingen cristobal.curio@reutlingen-university.de
Betreuer/-in Extern	Dr. Tobias Windisch Bosch GmbH Tobias.Windisch@de.bosch.com

Informatics Inside Frühling 2022
11. Mai 2022, Hochschule Reutlingen
Copyright 2022 Kai Müller

darauf Folgende vorhersagen zu können (Sequenz zu Sequenz), kann ein Conv-LSTM eingesetzt werden. Conv-LSTMs vereinen die in der Bildverarbeitung genutzte Methode der Faltung (engl. convolution) mit der LSTM-Architektur. Beim Training von RNN Architekturen kann es zum Problem der explodierenden Gradienten kommen, welches in Kapitel 7.1 genau beschrieben wird. Beim Problem explodierender Gradienten wachsen die Gradienten der Gewichte in einem Netz exponentiell, was zum Beispiel zur Folge haben kann, dass das Netz seine Vorhersagen nicht weiter verbessern kann. Das Problem existiert ebenfalls für besonders tiefe neuronale Netze und LSTM Architekturen [17]. In den folgenden Kapiteln wird zunächst der Stand der Forschung zu explodierenden Gradienten im Bezug auf Conv-LSTM Architekturen beschrieben. Anschließend folgt eine Beschreibung der Conv-LSTM Architektur und der Architekturen die ihr zu Grunde liegen (RNN, GRNN und LSTM). Im weiteren Verlauf wird das durchgeführte Experiment zur Untersuchung explodierender Gradienten bei Conv-LSTMs beschrieben und dessen Ergebnisse präsentiert. Nachfolgend wird ein Fazit aus den Ausführungen und dem durchgeführten Experiment gezogen.

2 Stand der Forschung

Bisher existiert keine Literatur, die einfache Conv-LSTM Architekturen unter Verwendung des moving-MNIST Datensatzes im speziellen auf das Verhalten der Gradienten untersucht. Luo et.al. [11] trainieren erfolgreich eine Variation der Conv-LSTM Architektur auf Basis des moving-MNIST Datensatzes. Elsayed et. al. [6] untersuchen Conv-LSTMs mit verschiedene Aktivierungsfunktionen auf dem moving-MNIST Datensatz. Es werden allerdings keine Aussagen über den Einfluss der Gradienten getroffen. Pascanu et. al. [15] beschreiben in ihrer Arbeit, dass LSTMs generell nicht von Problem verschwindender Gradienten betroffen sind, jedoch vom Problem der explodierenden Gradienten. Andere Autoren beziehen die Aussagen von Pascanu et. al. [15] als Basis für Ihre Arbeiten mit ein [13] [12] [9]. RNNs bilden die Grundlage für LSTMs und Conv-LSTMs. Sie werden im folgenden Kapitel spezifisch erörtert.

3 Recurrent neural networks

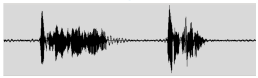
Ein recurrent neural network (RNN) ist ein in den 80er und 90er Jahren entwickeltes neuronales Netzmodell zur Modellierung von Zeitreihen bzw. von sequenziellen Daten [15]. Die Eigenschaft der zeitlichen Abhängigkeit, also die Abhängigkeit eines Datenpunktes zum Vorherigen, wird durch die Architektur von RNNs widergespiegelt. Sie bilden die Basis für spätere Entwicklungen wie Gated Recurrent Neural Networks (GRNNs) oder die zu GRNNs gehörigen LSTMs [16].

Die Struktur von RNNs sind aneinandergereihte Neuronen, welche Informationen aus allen vorigen Schritten mit einbeziehen. Durch die Verwendung von Informationen aus vorigen Schritten, kann das Modell Informationen über die Vergangenheit nutzen, sodass es zeitliche Korrelationen zwischen Ereignissen entdecken kann, die in den Daten voneinander entfernt sind [3].

Die Einsatzgebiete von RNNs sind zum Beispiel im Natural Language Processing, Übersetzung oder videobasierte Aktivitätserkennung oder das Erkennen grammatischer Strukturen [5]. Beispielfälle sind in Tabelle 1 dargestellt.

Während das rekurrente Netz im Prinzip ein einfaches und leistungsfähiges Modell ist, ergeben sich aus der Rekurrenz des Netzes im Trainingsprozess Schwierigkeiten [16]. Einer der Hauptgründe für die Schwerfälligkeit dieses Modells ist das in Bengio et al. (1994)

Tabelle 1: Sequenzielle Daten

Anwendungsfall	sequenzielle Daten	RNN Ausgabe
Spracherkennung		„Hallo Kathrin, ich komme heute vorbei“
Gefühlserkennung	„Es hat mir gar nicht gefallen"	
Maschinelle Übersetzung	„Do you like this book?”	Magst du dieses Buch?
Videobasierte Aktivitätserkennung		Fußball spielen

[15] beschriebenen Probleme des verschwindenden und explodierenden Gradienten. Alle großen machine learning frameworks bieten effiziente, produktionsreife Implementierungen einer Reihe von RNN-Architekturen an [16]. In Gleichung 1 wird die Berechnung des hidden state h_t einer RNN Zelle zum Zeitpunkt t definiert.

$$h_t = f(h_{t-1}, X_t, W) \tag{1}$$

Dabei ist f die Funktion, welche den sogenannten hidden state h_t auf Basis des vorigen hidden state h_{t-1} und dem Input zum Zeitpunkt t X_t berechnet [3]. W beschreibt die im Training zu optimierenden Gewichte. Um die Rekurrenz bzw. den Informationsfluss im RNN zu beschrieben, kann die obige Funktion, wie in Gleichung 2 dargestellt, in der Zeit aufgefaltet werden. Die vorigen hidden states h_{t-n} und die Informationen aus vorigen Zeitschritten die auf Basis der vorigen Inputs X_{t-n} berechnet wurden.

$$\begin{aligned} h_t &= f(h_{t-1}, X_t, W) \\ &= f(f(h_{t-2}, X_{t-1}, W), X_t, W) \\ &= f(f(f(h_{t-3}, X_{t-2}, W)X_{t-1}, W), X_t, W) \\ &= f(f(f(f(...))), X_t, W) \end{aligned} \tag{2}$$

In Abbildung 1 wird eine einzelne Zelle eines RNN Netzwerks dargestellt. Dabei ist $\hat{y}$ der errechnete Output zum Zeitpunkt t. Die verwendeten Gewichte W sind in einem forward pass immer die gleichen. Eine weitere

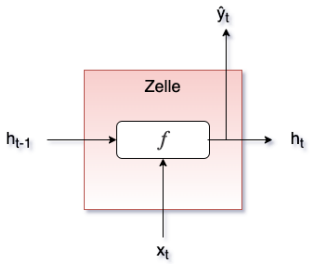


Abbildung 1: RNN Zelle

Eigenschaft von RNNs ist, dass der Einfluss von h_{t-n} auf h_t kleiner ist je größer n [3]. Generell gesehen beeinflussen nur vorherige In- und Outputs h_t, eine Ausnahme hierzu sind Bidirektionale RNNs. Diese verarbeiten auch Informationen die sequenziell vor dem Output liegen [16].

3.1 *Architektur Beispiele*

Die Architektur von RNNs kann verschiedene Formen annehmen. In Abbildung 2 werden Beispiele für verschiedene RNN Architekturen gegeben. Ein sequenzieller Input bei

einem einzigen Output, wie in Tabelle 1 dargestellt, kann im Anwendungsfall von Gefühlerkennung oder der videobasierten Aktivitätserkennung gegeben sein. Im Beispiel der Gefühlserkennung von Tabelle 1 könnte jedes Wort der Filmrezession ein Input x_t sein. Eine Weitere Architekturvariante sind Sequenz zu Sequenz Modelle. Dabei wird aus einer Sequenz eine weitere Sequenz als Output berechnet. Die Länge einer Sequenz wird mit T bezeichnet. So kann zum Beispiel in der maschinellen Übersetzung $T_x \neq T_y$ sein, da ein Satz in verschiedenen Sprachen verschiedene Anzahl an Wörtern haben kann.

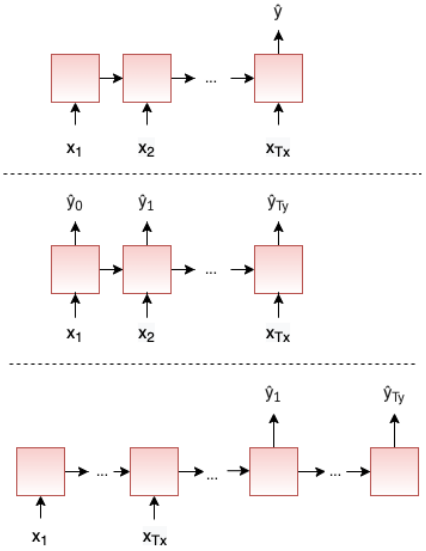


Abbildung 2: Beispiele für RNN Architekturen

4 Gated Recurrent Neural Networks

Die Gated Recurrent Unit (GRU) ist die Architekturbesonderheit der Gated Recurrent Neural Networks (GRNN). Sie wurde erstmals 2014 von Cho. et.al. [2] vorgestellt. Der Unterschied von GRNNs zu RNNs ist die Verwendung von gates [3].

$$\begin{aligned} \widetilde{h_t} &= tanh(W_c[G_r * h_{t-1}, x_t] + b_c) \\ G_u &= \sigma(W_u[h_{t-1}, x_t] + b_u) \\ G_r &= \sigma(W_r[h_{t-1}, x_t] + b_r) \\ h_t &= G_u * \widetilde{h_t} + (1 - G_u) * h_{t-1}) \end{aligned} \tag{3}$$

Dabei ist h_t der hidden state der GRU. $\widetilde{h_t}$ ist ein Platzhalterwert welcher variabel zum hidden state hinzugefügt werden kann. W sind die Gewichte. h_{t-1} ist der vorige hidden state und gleichzeitig der Input. b sind die Bias. Und G sind die gates, welche Einfluss darauf nehmen, ob Informationen vergessen werden und in wie weit $\widetilde{h_t}$ und h_{t-1} zu h_t beitragen. GRNNs haben auf verschiedenen Datensätzen gute Erfolge zeigen können [3]. Durch dem verringerten Rechenaufwand im Vergleich zu LSTMs sind sie leichter skalierbar.

5 Long Short Term Memory Netze

Für die allgemeine Sequenzmodellierung haben sich LSTMs als spezielle GRNN-Struktur in verschiedenen früheren Studien als stabil und leistungsfähig für die Modellierung weitreichender Abhängigkeiten erwiesen [17]. Eine LSTM Zelle lässt sich wie folgt definieren (adaptiert von [3].

$$\begin{aligned} \widetilde{c_t} &= tanh(W_c[h_{t-1}, x_t] + b_c) \\ G_u &= \sigma(W_u[h_{t-1}, x_t] + b_u) \\ G_f &= \sigma(W_f[h_{t-1}, x_t] + b_f) \\ G_o &= \sigma(W_o[h_{t-1}, x_t] + b_o) \\ c_t &= G_u * \widetilde{c_t} + G_f * c_{t-1} \\ h_t &= G_o * tanh(c_t) \end{aligned} \tag{4}$$

Die wichtigste Neuerung von LSTMs ist, dass die Zelle im Unterschied zum RNN nicht nur über einen hidden state h_t, sonder auch über einen cellstate c_t, sowie über drei parametrisierte gates verfügt, welche den Informationsfluss steuern. Kommt ein neuer Input h_{t-1},

x_t, c_t wird dieser, auf Basis des Zustandes des Input-gates G_u, mit dem hidden state verrechnet. Der vergangene cellstate c_{t-1} wird durch das forget-gate G_f reguliert, kann also "vergessen" werden. Das dritte gate G_o (Output-gate) reguliert das Einfließen des aktuellen cellstate c_t in den hidden state der Zelle h_t. Die gates sind eine sigmoidale Schicht, geben also Zahlen zwischen Null und Eins aus. Ein Vorteil der Verwendung des cellstate in Verbindung mit dem forget-gate ist, dass der Gradient gehindert wird, zu schnell zu verschwinden, was ein Problem von RNNs ist [17].

5.1 *Architektur Bezeichnungen*

In der Literatur gibt es unterschiedliche Auslegungen der Begriffe CNN-LSTMs, Conv.-LSTMs, Conv-LSTMs bzw. convolutional LSTM. CNN Architekturen sind meist Architekturen, die ein Convolutional Neural Network mit einem LSTM verbinden. Verschiedene Autoren wie zum Beispiel Behera et. al. [1] oder Kahn und Karim [10] bezeichnen CNN-LSTMs als Conv.-LSTMs (Behera et. al. [1] verwendet Co-LSTM als Abkürzung). Beide Autorengruppen verwenden eine Kombination aus einem Convolutional Neural Network (CNN) und einem LSTM. Dabei bekommt das LSTM als Input die Outputs des CNN. Der Unterschied zur üblichen Definition von Conv.-LSTM ist, dass bei den oben genannten Beispielen die beiden Netze CNN und LSTMs als alleinstehende Einheit definiert sind. Für diese Arbeit ist ein Conv-LSTM eine Architektur in der die Faltung in der eigentlichen LSTM Zelle stattfindet, die beiden Architekturen also verschmolzen werden. Die meisten Veröffentlichungen folgen dieser Definition. Es gibt noch weitere Architekturvariationen. Die unterschiedlichen Auslegungen der Begrifflichkeiten sind bei der Recherche zu beachten.

6 Conv-LSTM

Das Conv-LSTM ersetzt die Matrixmultiplikationen des LSTMs mit Faltungsoperationen. Während LSTMs zeitliche Vorhersagen zulassen, haben Conv-LSTMs zusätzlich die Eigenschaft räumliche/ bildliche Informationen mit zu verwenden [17].
Eine Conv-LSTM Zelle lässt sich wie folgt beschreiben (adaptiert nach [18]):

$$
\begin{aligned}
\widetilde{c_t} &= tanh(W_c[W_k[h_{t-1}, x_t]] + b_c) \\
G_u &= \sigma(W_u[W_k[h_{t-1}, x_t]] + b_u) \\
G_f &= \sigma(W_f[W_k[h_{t-1}, x_t]] + b_f) \\
G_o &= \sigma(W_o[W_k[h_{t-1}, x_t]] + b_o) \\
c_t &= G_u * \widetilde{c_t} + G_f * c_{t-1} \\
h_t &= G_o * tanh(c_t)
\end{aligned}
\tag{5}
$$

Die Berechnungen bleiben im Vergleich zum LSTM gleich, außer der Verwendung von W_k. W_k beschreibt die Faltungsoperation von $h_t - 1$ und x_t. Es gibt noch weiter Architektur Variationen, in welchen zum Beispiel für x_t und $h_t - 1$ verschiedene Faltungsoperationen im Trainingsprozess erlernt werden. Der Informationsfluss ist in Abbildung 3 dargestellt.

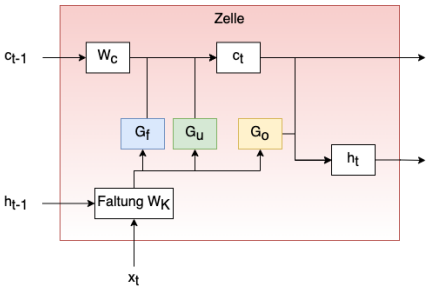


Abbildung 3: Conv-LSTM Zelle

Betrachtet man die cell- und hidden states als verborgene Repräsentationen sich bewegender Objekte (im Fall einer Video Sequenz als Input), so kann ein Conv-LSTM mit einem größeren Kernel für die Faltungsoperationen

schnellere Bewegungen erfassen, während ein Modell mit einem kleineren Kernel langsamere Bewegungen erfassen kann [17]. Der Kernel ist eine Filtermaske die mehrere Pixelwerte in einem vereint. Er wird in der Faltungsoperation über das Bild geschoben wird. Ein einfacher Kernel könnte zum Beispiel als Input neun Pixelwerte verwenden, diese addieren und durch neun teilen und als Output diesen Wert für ein neues Pixel festlegen. Faltungsoperationen schneiden durch die Kernel Informationen vom Bildrand ab, deshalb muss nach Anwendung der Faltungsoperation eine Auffüllung (padding) erfolgen, sodass die nächste Zelle die gleichen Input Dimensionen wie die vorige bekommt. Die Faltungsoperation in Conv-LSTMs wird in Abbildung 4 dargestellt.

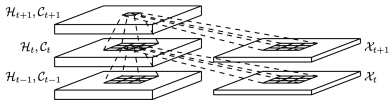


Abbildung 4: Faltungsoperationen im Conv-LSTM [17]

7 Explodierende Gradienten

Im Folgenden wird das Problem explodierender Gradienten im allgemeinen beschrieben. Der generelle Prozess beim Training eines Neuronalen Netzes läuft wie folgt ab: 1. Präsentation von Daten 2. Errechnung einer Vorhersage 3. Veränderung der Gewichte, sodass die nächste Vorhersage besser ist. Um ein Gewicht richtig zu verändern, muss errechnet werden in wie weit dieses Gewicht zur errechneten Vorhersage beigetragen hat. Dafür wird die Lossfunktion bezüglich jedes Gewichts abgeleitet. Eine Funktion zur Berechnung der Qualität einer Vorhersage. Der Gradient ist das Ergebnis dieser partiellen Ableitung. Eine genauere Beschreibung findet in Kapitel 7.1 statt. Vergrößert sich dieser Wert schlagartig spricht man von einem explodierenden Gradienten [15]. Das Problem verschwindender Gradienten bezeichnet den gegenteiligen Effekt, also kleiner werdende Gradienten [15]. Das hat zur Folge, dass sich die Gewichte über das Training nicht mehr ändern, das Lernen also stoppt. Explodierende Gradienten können sich zum Beispiel dadurch zeigen, dass sich die Vorhersagen des Netzes nicht verbessern oder dass es große, sprungartige Veränderungen des Losswerts gibt. Durch Gradient Clipping kann das Problem explodierender Gradienten vermieden werden. Dabei werden die Gradienten ab einer bestimmten Größe herunter skaliert [15].

7.1 Explodierende Gradienten in RNNs

Im Folgenden wird das Problem explodierender Gradienten in RNNs spezifisch dargestellt. Für die Ausführungen wird ein einfaches RNN angenommen welches eine Input Sequenz der Länge zwei erhält. Folglich gibt es zwei hidden states die uns auch als Output dienen.

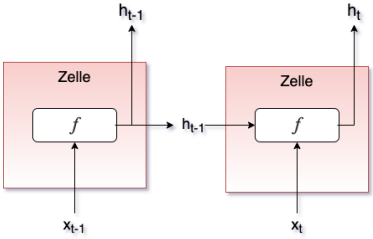


Abbildung 5: RNN mit einer Input Sequenz von zwei. Der Output ist gleich der hidden state

Auf Basis des Outputs h_t wird, mit Hilfe der Lossfunktion, der Loss $L(h_t)$ errechnet. Um die Gewichte anzupassen, muss die Lossfunktion partiell zu einem Gewicht abgeleitet werden [7]. Allgemein gesprochen, inwieweit ändert sich ein Loss L bei der Änderung eines Gewichtes W. Das Ergebnis der Ableitung ist

der Gradient.

$$\frac{\partial L}{\partial W} \tag{6}$$

Der Gradient, das Gewicht W und die Lernrate η werden anschließend genutzt um das neue Gewicht W_u wie folgt zu berechnen.

$$W_u = W - \eta \frac{\partial L}{\partial W} \tag{7}$$

Möchte man den Gradienten für das Gewicht W errechnen, ergibt sich unter Verwendung der Kettenregel Gleichung 8 (adaptiert von [7]):

$$\begin{aligned} \frac{\partial L}{\partial W} &= \frac{1}{T} \sum_{t=1}^{T} \frac{\partial l(h_t)}{\partial W} \\ &= \frac{1}{T} \sum_{t=1}^{T} \frac{\partial l(h_t)}{\partial h_t} \frac{\partial h_t}{\partial W} \end{aligned} \tag{8}$$

Dabei ist L der (gesamte) zu errechnende Loss, der sich zusammensetzt aus der Summe des errechneten Loss l für jeden Output h_t zu allen Zeitschritten t. T ist in unserem Beispiel zwei. Durch die mehrfache Multiplikation von Werten kann es leicht zu sehr großen oder sehr kleinen Werten kommen. Bei der Berechnung von $\frac{\partial h_t}{\partial W}$ ist zu beachten, dass auch vorige Zustände von W beeinflusst werden/ recurrent sind. So basiert h_t sowohl auf h_{t-1} als auch auf W. Wobei die Berechnung von h_{t-1} wiederum ebenfalls auf W beruht. Unter Verwendung der Kettenregel ergibt sich für die Berechnung von $\frac{\partial h_t}{\partial W}$ Gleichung 9 (adaptiert von [7]).

$$\frac{\partial h_t}{\partial W} = \frac{\partial (h_t)}{\partial W} + \frac{\partial (h_t)}{\partial h_{t-1}} \frac{\partial (h_{t-1})}{\partial W} \tag{9}$$

Daraus folgt, je tiefer das Netz bzw. je länger die Input Sequenz, desto länger wird die Gleichung zur Berechnung von $\frac{\partial h_t}{\partial W}$. Folglich erhöht sich je nach Länge des Inputs die Instabilität des Trainings und die wahrscheinlich von explodierenden Gradienten.
Um das Problem, das aus der wiederholten Anwendung einer Gleichung auf sich selbst entsteht, visuell darzustellen, wird auf Gleichung 2 zurückgegriffen. In dieser Gleichung wird gezeigt wie die Funktion f einer Zelle, wie in Abbildung 1 dargestellt, im Trainingsprozess immer wieder auf sich selbst angewandt wird. Um die entstehende Instabilität, bei der wiederholten Anwendung einer Funktion auf sich selbst und damit auch die Instabilität der Ableitungen zu verdeutlichen, wird hier Beispielhaft die Sinusfunktion $sin(4x)$ mehrfach auf sich selbst angewandt, siehe Abbildung 6. Der Output der inneren Sinusfunktion ist der Input für die jeweils Nächste. So wird die Steigung der entstehenden Funktionen immer extremer und damit auch die Ableitung. Kleine Änderungen im Wert x führen zu großen Veränderungen des Ergebnisses. Im folgenden Kapitel wird das Experiment zur Untersuchung explodierender Gradienten in LSTMs beschrieben.

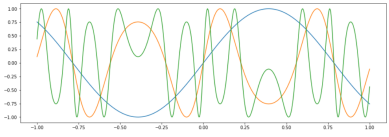


Abbildung 6: Blau: sin(4x), Gelb: sin(sin(4x)), Grün: sin(sin(sin(4x)))

8 Experiment

Um LSTMs auf das Problem von explodierenden Gradienten zu untersuchen, wurde ein LSTM mit einer Input Sequenz von 10 Bildern des moving-MNIST Datensatzes trainiert. Das Netz hat die Aufgabe, die nächsten zehn Bilder der Sequenz vorherzusagen. Es handelt sich also um eine Sequenz zu Sequenz Architektur (siehe Abbildung 2). Der Versuch wurde mit einem 3x3, einem 5x5 und einem 7x7 Kernel für die Faltungsoperationen durchgeführt. Das Training lief etwa zwei Tage pro Kernel. Es wurde eine Nvidia GTX 1080 verwendet. Während des Trainings wurden neben Hyperparametern, die Gewichte, die Gradienten der Gewichte, die Bias und die Gradienten der Bias, für beide

Tabelle 2: Netzwerkarchitektur [14]

Schichtname	Output Bedeutung	Output Dimensionen
Inputlayer	(batch size, time step, n channel, Breite, Höhe)	(12, 10, 1, 64, 64)
Conv-LSTM encoder 1	(batch size, time step, nf, Breite, Höhe)	(12, 10, 64, 64, 64)
Conv-LSTM encoder 2	(batch size, time step, nf, Breite, Höhe)	(12, 10, 64, 64, 64)
Conv-LSTM decoder 1	(batch size, time step, nf, Breite, Höhe)	(12, 10, 64, 64, 64)
Conv-LSTM decoder 2	(batch size, time step, nf, Breite, Höhe)	(12, 10, 64, 64, 64)
3D-CNN decoder	(batch size, time step, n channel, Breite, Höhe)	(12, 10, 1, 64, 64)

de- und encoder Zellen sowie für die CNN Schicht dokumentiert.

8.1 Netzwerkarchitektur

Die verwendete Netzwerkarchitektur ist in Tabelle 2 dargestellt. Die Architektur wurde von [14] übernommen, da diese auf dem moving-MNIST Datensatz gute Ergebnisse erzielen konnte. Das Netz besteht aus zwei encoder und zwei decoder Zellen. Die finale decoder Zelle gibt 64 featuremaps der Dimension 64x64 für jedes vorhergesagte Bild aus, siehe Tabelle 2. Die Anzahl der featuremaps wird mit nf bezeichnet. Da im Wesentlichen eine Regression (Vorhersage von Pixelwerten) durchgeführt wird, müssen die Feature in tatsächliche Vorhersagen umgewandelt werden, ähnlich wie bei einer pixelgenauen Bildklassifizierung. Um dies zu erreichen, ist eine 3D-CNN-Schicht implementiert, welche den Output der finalen decoder Zelle (64, Breite=64, Höhe=64) als Input bekommt. Als Ausgabe wird ein (1, 64, 64) Feature generiert. Als Optimizer wurde Adam verwendet. Die Lernrate betrug 0,0001. Es fand kein Gradient Clipping statt.

8.2 Daten

Die zum Training verwendeten Daten basieren auf dem moving-MNIST Datensatz. Dieser enthält 20 Bilder lange Videosequenzen sich gleichmäßig bewegender Zahlen. Die Bilder haben eine Dimension von (1, 64, 64).

9 Ergebnisse

Im Folgenden werden die Ergebnisse des Experiments dargestellt. Die gesamten Ergebnisse mit Histrogrammen und Verteilungen sind für den 3x3 Kernel[2], den 5x5 Kernel[3] und den 7x7 Kernel[4] öffentlich einsehbar.

9.1 Gradienten

Die Gradientenwerte werden mithilfe von Verteilungen visualisiert. Die Verteilungen sind durch die Linien in Sektionen unterteilt, wobei die Aufteilung von oben nach unten [*Maximum*, 93%, 84%, 69%, 50%, 31%, 16%, 7%, *Minimum*] sind. Die Percentile können auch als Markierungen für die Standardabweichung σ gesehen werden. Deshalb

[2] 3x3 Kernel: https://tensorboard.dev/experiment/gAJuMXQgR9u3g9hliemYmA
[3] 5x5 Kernel: https://tensorboard.dev/experiment/WdV1J1F2R3uYrVvl5D3Aww
[4] 7x7 Kernel: https://tensorboard.dev/experiment/rEwq3WJKRLSBoHqRCkAqOw

können die Linien von oben nach unten auch wie folgt interpretiert werden $[Maximum, \mu + 1.5\sigma, \mu + \sigma, \mu + 0.5\sigma, \mu, \mu - 0.5\sigma, \mu - \sigma, \mu - 1.5\sigma, Minimum]$. Die drei eingeschlossenen Zonen haben die Größe $[\sigma, 2\sigma, 3\sigma]$.

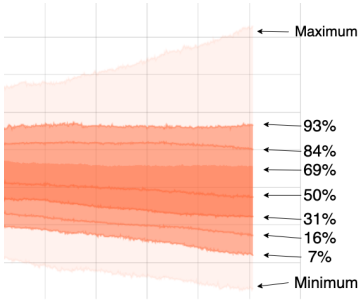


Abbildung 7: Histogramm Legende

Die Gradienten über den Verlauf des Trainings sind in Abbildung 8 dargestellt. Auf die Visualisierung der Gewichte, des Loss und der Bias wird in dieser Arbeit verzichtet. In keinem der Trainings kam es zu explodierenden Gradienten. Es zeigt sich, dass die Gradienten der CNN Schicht größer sind als die der decoder Schicht, welche wiederum größer sind als die der encoder Schicht. Der Grund dafür liegt wahrscheinlich in der Netzwerkarchitektur, da der Einfluss der Gewichte auf den Output umso größer ist, je näher sie im Informationsfluss am Output liegen (Input - decoder - encoder - CNN - Output). Außerdem fällt auf, dass die Streuung der Gradienten der CNN Schicht deutlich größer ist. Der Loss über den Verlauf des Trainings nahm mit allen Kerneln auch nach Beendigung des Trainings weiter ab. Das zeigt sich auch an den Vorhersagen des Netzes, die über die Epochen hinweg besser wurden. Grund für die Stabilität der Gradienten könnte die Input Sequenzlänge sein. Diese liegt bei zehn In- und Output Bildern . Es werden damit in der Berechnung der Gradienten maximal 20 Schritte zurück iteriert (siehe Kapitel 7.1). Je nach Einsatzgebiet, können längere Input Sequenzen verwendet werden, zum Beispiel in der Sprachverarbeitung oder in der Wettervorhersage. Außerdem könnte die Verwendung einer höheren Lernrate dazu beitragen das Problem auftreten zu lassen. Je höher die Lernrate, desto größer die Veränderung der Gewichte (siehe Gleichung 7).

10 Fazit

Unter keiner der verschiedenen Kernelgrößen kam es zu exponentiell wachsenden Gradienten. So konnte gezeigt werden, dass es unter Verwendung der gegebenen Trainingsparameter nicht zu exponentiell wachsenden Gradienten kommen muss. Die Ergebnisse bilden die Basis für weitere Untersuchungen zukünftiger Studien zur Problematik explodierender Gradienten. Außerdem wurde durch die Arbeit das komplexe Problem explodierender Gradienten anschaulich dargestellt und ein Bezug zur Architektur von LSTMs hergestellt. Die Architektur von Conv-LSTMs wird dabei aus der Architektur von RNNs, GRNNs und LSTMs abgeleitet. Die Arbeit dient dazu, das allgemeine Problem verständlich darzustellen und mögliche Faktoren, die das Trainieren der Netzwerke beeinflussen, aufzuzeigen. Das ist insofern besonders wichtig, da es für das Problem verschiedenste, fallindividuelle Faktoren geben kann und somit ein Verständnis des Problems zentral ist. Es kann keine allgemeine Lösung gefunden werden, die für jede Anwendung explodierenden Gradienten vermeidet. Zukünftige Arbeiten können Lernrate, Input Sequenzlänge, Netzwerkarchitektur oder auch Datensatz variieren, um so die gewonnenen Erkenntnisse noch zu erweitern.

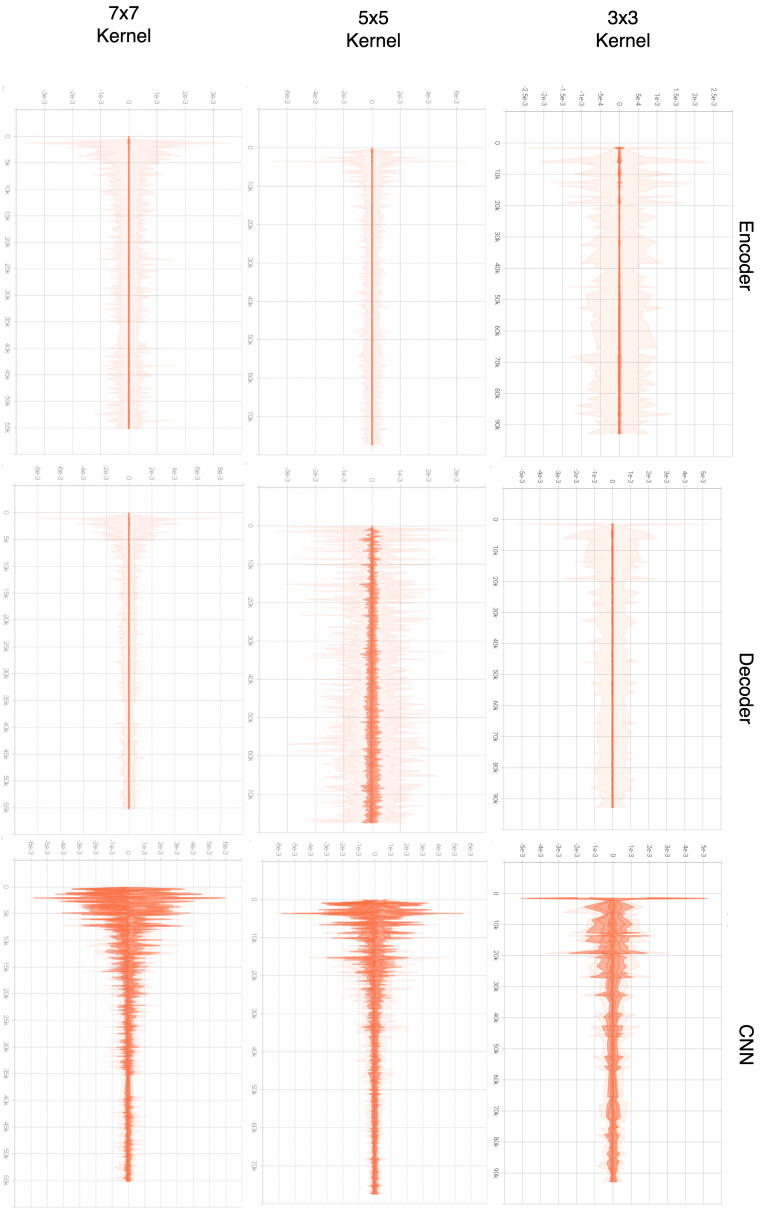


Abbildung 8: Gradienten im Vergleich

Literatur

[1] Ranjan Kumar Behera u. a. „Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data“. In: *Information Processing &amp Management* 58.1 (Jan. 2021), S. 102435. DOI: 10.1016/j.ipm.2020.102435. URL: https://doi.org/10.1016/j.ipm.2020.102435.

[2] KyungHyun Cho u. a. *On the Properties of Neural Machine Translation: Encoder-Decoder Approaches*. URL: http://arxiv.org/pdf/1409.1259v2.

[3] Junyoung Chung u. a. *Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling*. URL: http://arxiv.org/pdf/1412.3555v1.

[4] Jeff Donahue u. a. „DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition“. In: *Proceedings of the 31st International Conference on Machine Learning*. Hrsg. von Eric P. Xing und Tony Jebara. Bd. 32. Proceedings of Machine Learning Research 1. Bejing, China: PMLR, 22–24 Jun 2014, S. 647–655. URL: https://proceedings.mlr.press/v32/donahue14.html.

[5] Jeffrey L. Elman. „Distributed representations, simple recurrent networks, and grammatical structure“. In: *Machine Learning* 7.2-3 (Sep. 1991), S. 195–225. DOI: 10.1007/bf00114844. URL: https://doi.org/10.1007/bf00114844.

[6] Nelly Elsayed, Anthony S. Maida und Magdy Bayoumi. „Empirical Activation Function Effects on Unsupervised Convolutional LSTM Learning“. In: *2018 IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI)*. 2018, S. 336–343. DOI: 10.1109/ICTAI.2018.00060.

[7] Aston f u. a. „Dive into Deep Learning“. In: *arXiv preprint arXiv:2106.11342* (2021). Kapitel 8.7.3.

[8] Alex Graves und Navdeep Jaitly. „Towards End-To-End Speech Recognition with Recurrent Neural Networks“. In: *Proceedings of the 31st International Conference on Machine Learning*. Hrsg. von Eric P. Xing und Tony Jebara. Bd. 32. Proceedings of Machine Learning Research 2. Bejing, China: PMLR, 22–24 Jun 2014, S. 1764–1772. URL: https://proceedings.mlr.press/v32/graves14.html.

[9] Andrej Karpathy, Justin Johnson und Li Fei-Fei. *Visualizing and Understanding Recurrent Networks*. 2015. DOI: 10.48550/ARXIV.1506.02078. URL: https://arxiv.org/abs/1506.02078.

[10] Muhammad Khan, Md. Karim und Yangwoo Kim. „A Scalable and Hybrid Intrusion Detection System Based on the Convolutional-LSTM Network“. In: *Symmetry* 11.4 (Apr. 2019), S. 583. DOI: 10.3390/sym11040583. URL: https://doi.org/10.3390/sym11040583.

[11] Weixin Luo, Wen Liu und Shenghua Gao. „Remembering history with convolutional LSTM for anomaly detection“. In: *2017 IEEE International Conference on Multimedia and Expo (ICME)*. 2017, S. 439–444. DOI: 10.1109/ICME.2017.8019325.

[12] Yixuan Ma, Zhenji Zhang und Alexander Ihler. *Multi-Lane Short-Term Traffic Forecasting With Convolutional LSTM Network*. https://ieeexplore.ieee.org/abstract/document/9000828. Feb. 2020.

[13] Roni Mittelman. *Time-series modeling with undecimated fully convolutional neural networks*. 2015. DOI: 10.48550/ARXIV.1508.00317. URL: https://arxiv.org/abs/1508.00317.

[14] Andreas Nielsen. Juli 2021. URL: https://towardsdatascience.com/video-prediction-using-convlstm-with-pytorch-lightning-27b195fd21a2.

[15] Razvan Pascanu, Tomas Mikolov und Yoshua Bengio. „On the difficulty of training recurrent neural networks“. In: *Proceedings of the 30th International Conference on Machine Learning*. Hrsg. von Sanjoy Dasgupta und David McAllester. Bd. 28. Proceedings of Machine Learning Research 3. Atlanta, Georgia, USA: PMLR, 17–19 Jun 2013, S. 1310–1318. URL: https://proceedings.mlr.press/v28/pascanu13.html.

[16] Alex Sherstinsky. „Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network“. In: *Physica D: Nonlinear Phenomena* 404.8 (2020), S. 132306. ISSN: 01672789. DOI: 10.1016/j.physd.2019.132306. URL: http://arxiv.org/pdf/1808.03314v9.

[17] Xingjian SHI u. a. „Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting“. In: *Advances in Neural Information Processing Systems*. Hrsg. von C. Cortes u. a. Bd. 28. Curran Associates, Inc., 2015. URL: https://proceedings.neurips.cc/paper/2015/file/07563a3fe3bbe7e3ba84431ad9d055af-Paper.pdf.

[18] Liang Zhang u. a. „Attention in Convolutional LSTM for Gesture Recognition“. In: *Advances in Neural Information Processing Systems*. Hrsg. von S. Bengio u. a. Bd. 31. Curran Associates, Inc., 2018. URL: https://proceedings.neurips.cc/paper/2018/file/287e03db1d99e0ec2edb90d079e142f3-Paper.pdf.

©2022 Kai Müller. 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

Einsatz von KI im RE: Ein systematisches Literaturreview

Elma Bajraktari
Hochschule Reutlingen
Reutlingen, Deutschland
Elma.Bajraktari@Student.Reutlingen-University.de

Abstract

[Motivation/Kontext] Der Einsatz vom Requirements Engineering (RE) ist für die Entwicklung eines softwarebasierten Produkts notwendig, um die Bedürfnisse der NutzerInnen zu unterstützen. Dabei werden im RE alle Anforderungen bei StakeholderInnen erhoben, die dann dokumentiert, verifiziert sowie validiert und verwaltet werden. Durch den Einsatz von Formen der künstlichen Intelligenz (KI) können diese vier RE-Aktivitäten unterstützt werden. **[Problemstellung]** Dabei fehlt den Requirements Engineers der Überblick darüber, welche Formen der KI aktuell in welcher der oben genannten RE-Aktivitäten als Unterstützung eingesetzt werden können. **[Lösung/Ergebnis]** Im Rahmen dieses systematischen Literaturreviews wird Literatur bezüglich des Einsatzes von KI im RE analysiert. Die Ergebnisse zeigen, welche Form der KI in welcher RE-Aktivität eingesetzt werden kann und welche Ergebnisse die verwendete Form der KI liefert. Es wird aufgezeigt, dass nicht alle RE-Aktivitäten gleichermaßen anhand von KI unterstützt werden und somit bei der Verwaltung und Verifikation sowie Validierung von Anforderungen weitere Forschungsarbeit erforderlich ist. **[Beitrag]** Durch den vorliegenden Artikel wird ein Überblick über den Einsatz von KI im RE verschafft.

CCS Concepts

• Software and its engineering~Software creation and management~Designing software~Requirements analysis • Computing methodologies~Artificial intelligence

Keywords

Künstliche Intelligenz (KI), Requirements Engineering (RE), Systematisches Literaturreview (SLR)

1 Einleitung

Bei der Entwicklung von softwarebasierten Produkten sind laut Jantzer et al. [20] mehrere StakeholderInnen mit verschiedenen Vorstellungen beteiligt, die ein gemeinsames Ziel erreichen wollen. Der Fokus liegt hierbei darauf, dass das entwickelte Produkt auf dem Markt erfolgreich wird. Daher ist laut Gilz [15] bei der Entwicklung von Produkten wichtig, bereits zu Beginn zu ermitteln, was genau entwickelt werden soll. Hierfür kommt das Requirements Engineering (RE) zum Einsatz. Unter RE wird laut Gilz [15] ein Prozess verstanden, in dem Anforderungen für ein Produkt erhoben, dokumentiert, verifiziert sowie validiert und verwaltet werden. Da das RE komplex sein kann, können die Requirements Engineers bei der Ausführung der RE-Aktivitäten laut Palomares et al.

Betreuer Hochschule: Prof. Dr. Christian Kücherer
Hochschule Reutlingen
Christian.Kuecherer@Reutlingen-University.de

Informatics Inside Frühjahr 2022
11. Mai 2022, Hochschule Reutlingen
Copyright 2022 Elma Bajraktari

[29] durch künstliche Intelligenz (KI) unterstützt werden, indem die Aufgaben in den einzelnen RE-Aktivitäten automatisiert werden. Ziel dieses Artikels ist anhand von einem systematischen Literaturreview (SLR) ein Überblick über den Einsatz von KI im RE zu verschaffen. Ein SLR ist laut Kitchenham und Charters [22] eine wissenschaftliche Vorgehensweise, womit die verfügbare Literatur zu einer Forschungsfrage und zugleich zu einem Themenbereich eingesammelt, bewertet und interpretiert wird. Dabei handelt es sich um eine Übersicht des aktuellen Wissens zu einem Themenbereich. Eine solche Art von SLR wurde bereits von Batarseh et al. [5] und Breuninger et al. [6] durchgeführt. Beide Artikel unterscheiden sich hinsichtlich folgender Aspekte zu diesem SLR:

Batarseh et al. [5] betrachten den Einsatz von KI in allen Software Engineering (SE) Phasen, wozu auch das RE gehört. Dabei wurden Artikel aus den Jahren von 1975 bis 2017 einbezogen, die auf den Datenbanken IEEE, Association for the Advancement of Artificial Intelligence, Springer, ACM Digital Library und Elsevier enthalten waren. Die RE-Aktivitäten wurden nicht spezifisch definiert, es wurde generell nach dem Einsatz von KI im SE gesucht. Aus dem Artikel wird nicht ersichtlich, welcher Suchterm für die Literaturrecherche verwendet wurde. Als Ergebnis wurde ermittelt, dass Forscher sich einig sind, dass KI-Methoden in den SE Phasen eingesetzt werden können aber auf diesem Gebiet weitere Forschungsarbeiten erforderlich sind.

Breuninger et al. [6] betrachten den Einsatz von KI im RE. Dabei wurden Artikel aus den Jahren von 2010 bis 2020 einbezogen, die auf den Datenbanken EBSCOhost, ScienceDirect und ACM Digital Library enthalten waren. Die RE-Aktivitäten wurden als Anforderungserhebung, -analyse, -spezifikation und -validierung definiert. Der für die Literaturrecherche verwendete Suchterm ist grob, da es nur die UND-Verknüpfung von den Begriffen „artificial intelligence“ und „requirements engineering“ beinhaltet. Als Ergebnis wurde ermittelt, dass KI in allen definierten RE-Aktivitäten zum Einsatz kommt, aber bei der Validierung von Anforderungen weitere Forschungsarbeit erforderlich ist.

In diesem vorliegenden SLR wird der Einsatz von KI im RE betrachtet. Da die Artikel aus den Jahren von 1975 bis 2017 auf den Datenbanken SpringerLink, ScienceDirect, IEEE Xplore und ACM Digital Library bereits von Batarseh et al. [5] betrachtet wurden, werden in diesem SLR nur Artikel aus den Jahren von 2018 bis 2022 einbezogen. Dabei werden die RE-Aktivitäten, anders als bei Breuninger et al. [6], wie nach Gilz [15] als Anforderungserhebung, -dokumentation, -verifikation sowie -validierung und -verwaltung definiert. Im Gegensatz zu Breuninger et al. [6] werden in diesem SLR zusätzlich Artikel aus den Datenbanken IEEExplore und SpringerLink einbezogen und der Suchterm wird spezifischer definiert, wie in Kapitel 3.2 beschrieben wird.

2 Aufbau des Artikels

In Kapitel 3 wird die Methodik beschrieben, die für dieses SLR angewandt wird. Kapitel 4 befasst sich mit den Grundlagen zu den Themenbereichen RE und KI. In Kapitel 5 werden die Ergebnisse aufgelistet. Dabei wird dargestellt, welche Formen der KI in welchen RE-Aktivitäten eingesetzt werden können und welche Ergebnisse die verwendeten Formen der KI liefern. In Kapitel 6 wird die Diskussion und das Fazit dargestellt.

3 Methodik

In diesem Artikel wird ein SLR nach den Prinzipien von Kitchenham und Charters [22] durchgeführt, wobei folgende Aspekte betrachtet werden: Forschungsfragen, Suchstrategie, Ein- und Ausschlusskriterien, Datenextraktion und Datensynthese. Zudem wird die Auswahl der einbezogenen Literatur anhand von einer Abbildung dargestellt.

3.1 Forschungsfragen

Die Forschungsfragen von diesem Artikel lauten folgendermaßen:

RQ1: Welche Formen der KI werden für das RE bei den RE-Aktivitäten nach Gilz eingesetzt?

RQ2: Welche Ergebnisse liefert die verwendete Form der KI?

3.2 Suchstrategie

RQ1 und RQ2 werden mit der gleichen Literatur beantwortet, weshalb nur ein Suchterm aufgestellt wird. Der Suchterm lautet folgendermaßen: *("requirements engineering" OR "requirements elicitation" OR "requirements specification" OR "requirements verification" OR "requirements validation" OR "requirements management") AND "artificial intelligence"*. Dieser Suchterm wird auf den Datenbanken SpringerLink[1], ScienceDirect[2], IEEE Xplore[3] und ACM Digital Library[4] eingegeben.

3.3 Ein- und Ausschlusskriterien

Für die Auswahl der relevanten Literatur werden Ein- und Ausschlusskriterien festgelegt, die der Tabelle 1 und der Tabelle 2 zu entnehmen sind. Dadurch wird irrelevante Literatur aussortiert.

Tabelle 1: Einschlusskriterien

ID	Einschlusskriterium
EK1	Englischsprachige Literatur
EK2	Literatur aus den Jahren von 2018 bis 2022
EK3	Literatur mit kostenlosem Zugriff über die Hochschule Reutlingen
EK4	Literatur, die Ergebnisse zu beiden Forschungsfragen liefert

Tabelle 2: Ausschlusskriterien

ID	Ausschlusskriterium
AK1	SLR, das bereits in diesem Bereich durchgeführt wurde
AK2	Literatur ohne Peer-Review
AK3	Literatur, in der keine Evaluation der KI stattfand
AK4	Literatur, in der durch die Evaluation negative oder unschlüssige Ergebnisse festgestellt wurden
AK5	Duplikat (gleiche Literatur auf verschiedenen Datenbanken)

3.4 Auswahl der einbezogenen Literatur

Bei der Aussortierung der relevanten Literatur werden fünf Phasen durchlaufen, in denen Teile der Literatur analysiert sowie die Ein- und Ausschlusskriterien angewandt werden. Der Aussortierungsprozess der Literatur ist in Abbildung 1 dargestellt.

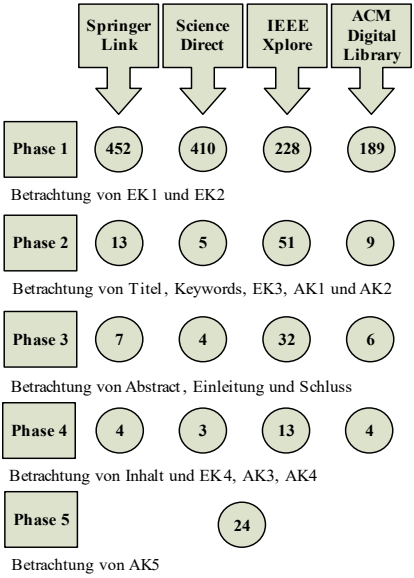


Abbildung 1: Aussortierungsprozess

[1] https://link.springer.com/

[2] https://www.sciencedirect.com/

[3] https://ieeexplore.ieee.org/Xplore/home.jsp

[4] https://dl.acm.org/

Zu Beginn werden 1279 Artikel ermittelt, von denen nach dem Aussortierungsprozess noch 24 übrigbleiben. Diese werden für die Darstellung der Ergebnisse verwendet, die in Kapitel 5 aufzufinden sind. Dadurch, dass keine Duplikate vorhanden sind, ist die Anzahl der Artikel in der Phase 4 und in der Phase 5 identisch.

3.5 Datenextraktion

Bei der Datenextraktion wird der Inhalt der Literatur betrachtet und es werden folgende Daten extrahiert und in einem Word-Dokument festgehalten: Name des Ansatzes, die betroffene RE-Aktivität, die eingesetzte Form der KI in der RE-Aktivität und die durch die Evaluation der verwendeten Formen der KI erzielten Ergebnisse.

3.6 Datensynthese

In der Datensynthese werden die relevanten Inhalte der 24 Artikeln zusammengefasst. Diese Zusammenfassung ist in Kapitel 5 einsehbar.

4 Grundlagen

In diesem Kapitel werden die Begriffe KI und RE erläutert.

4.1 Requirements Engineering

Der RE-Prozess besteht laut Gilz [15] aus vier RE-Aktivitäten, die in Abbildung 2 dargestellt sind.

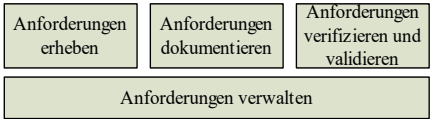


Abbildung 2: RE-Aktivitäten in Anlehnung an [15]

Unter einer Anforderung wird eine Eigenschaft eines Produktes verstanden, die erfüllt werden muss, um ein bestimmtes Ziel erreichen zu können. Dabei wird zwischen funktionalen und nicht-funktionalen Anforderungen unterschieden. Funktionale Anforderungen beschreiben, was das Produkt leisten und nicht-funktionale Anforderungen beschreiben Qualitätsmerkmale, die ein Produkt besitzen muss [10]. Für die Anforderungserhebung werden laut Gilz [15] StakeholderInnen bezüglich Eigenschaften befragt, die das Produkt erfüllen muss oder es werden Anforderungen in bereits existierenden Dokumenten erhoben. Zudem müssen die erhobenen Anforderungen so dokumentiert werden, dass sie adäquat, vollständig, widerspruchsfrei, verständlich, eindeutig und prüfbar sind. Die Anforderungen können beispielsweise anhand von Satzschablonen textuell dokumentiert oder anhand von Diagrammen wie Anwendungsfalldiagrammen und Aktivitätsdiagrammen bildlich modelliert werden. Bei der Verifikation und Validierung der Anforderungen werden die Qualität und der Inhalt der Anforderungen sichergestellt und überprüft, um Fehler rechtzeitig zu entdecken und zu beheben [15]. Dies kann beispielsweise anhand von einer Inspektion stattfinden [34]. Die Verwaltung von Anforderungen ist laut Gilz [15] auch unter dem Begriff Requirements Management (RM) bekannt. Im RM werden die Anforderungen mit ihren dazugehörigen Informationen und ihren Relationen zueinander verwaltet. Diese Aktivität ist relevant, da Anforderungen sich ändern, neu hinzukommen oder gelöscht werden können.

4.2 Künstliche Intelligenz

Unter KI wird laut Kreutzer und Sirrenberg [23] zum einen die Untersuchung des intelligenten Problemlösungsverhaltens und zum anderen die darauf basierte Erstellung von Systemen verstanden, die intelligente Lösungen automatisiert generieren. Die KI beinhaltet Teilbereiche wie neuronale Netze, maschinelles Lernen, Deep Learning, Natural Language Processing (NLP) und Knowledge Representation [14][23]. Diese Teilbereiche und ihre dazugehörigen Techniken stellen Formen der KI dar, die im RE als Unterstützung eingesetzt werden können, wie aus Kapitel 5 entnommen werden kann.

4.2.1 Neuronale Netze

Laut Kreutzer und Sirrenberg [23] stellt ein neuronales Netz im Kontext der KI ein System aus Hardware und Software dar, was dem Aufbau eines biologischen Gehirns nahekommt. Dieses Netz besteht aus einer Reihe von parallel arbeitenden und in mehreren Sichten angeordneten Prozessoren. Innerhalb der Schichten erfolgt die Weiterleitung von Informationen, wodurch das System bestenfalls dazu lernt. Dabei gibt es laut Paaß und Hecker [28] unterschiedliche Arten wie Convolutional Neural Network oder Recurrent Neural Network.

4.2.2 Maschinelles Lernen

Beim maschinellen Lernen werden laut Paaß und Hecker [28] die zur Verfügung stehenden Daten von einem System analysiert, wodurch sich das System selbstständig verbessert, um eine Aufgabenstellung besser ausführen zu können. Dabei können die drei Lernarten überwachtes Lernen, unüberwachtes Lernen und verstärktes Lernen zum Einsatz kommen [23][28]. Das überwachte Lernen lernt anhand von gekennzeichneten Daten, das unüberwachte Lernen lernt anhand von ungekennzeichneten Daten und das verstärkte Lernen, lernt aus einem Trial-and-Error-Prozess dazu [23]. Zum überwachten Lernen gehören Algorithmen wie Random Forest [2] und zum unüberwachten Lernen Algorithmen wie der k-Means-Algorithmus [23] dazu.

4.2.3 Deep Learning

Deep Learning ist eine Methode des maschinellen Lernens und verwendet neuronale Netze. Dabei kann diese Methode eine große Menge an Daten über mehrere Schichten des Netzwerkes verarbeiten, wobei weniger Hilfe von Personen bei der Datenvorverarbeitung benötigt wird. Das System lernt selbstständig durch die Verarbeitung der Daten dazu und kann den Zusammenhang zwischen neuen Eingabedaten und alten Eingabedaten herstellen [23].

4.2.4 Natural Language Processing

Unter NLP wird die Verarbeitung natürlicher Sprache verstanden [28]. Dieser Teilbereich der KI kann beispielsweise bei der Verarbeitung von textuellen Anforderungen verwendet werden, wodurch der manuelle Aufwand dieser Aktivität vermieden wird. Dabei werden unterschiedliche Schritte wie Text Tokenization oder Part-of-Speech Tagging ausgeführt [24]. Bei Text Tokenization wird der Text in Tokens zerlegt. Beim Part-of-Speech Tagging werden jedem Token Wortarten wie Nomen oder Verb zugewiesen [33].

4.2.5 Knowledge Representation

Knowledge Representation ist ein Teilbereich der KI, wobei es darum geht, Wissen in Computersystemen abzubilden. Dafür kann beispielsweise die formale Ontologie eingesetzt werden, womit ein bestimmter Bereich anhand von Entitäten und deren dazugehörigen Beziehungen formal dargestellt wird [14].

5 Ergebnisse zum Einsatz von KI im RE

In diesem Kapitel werden die Ergebnisse aufgelistet, die durch die Betrachtung der 24 relevanten Artikel erstellt werden. Dabei werden die Artikel den RE-Aktivitäten zugeordnet, die durch eine Form der KI unterstützt werden. Der Einsatz der Formen der KI kann anhand von Performanz-Metriken wie Precision, Recall und F1-Score evaluiert werden, weshalb in einigen Ergebnisdarstellungen auf diese eingegangen wird [32].

5.1 Anforderungserhebung

Zhang et al. [38] stellen die CCM-TFIDFE Methode vor, die auf Ontology Learning basiert. Durch den Einsatz dieser Methode werden alle Domänenkonzepte aus Dokumenten der Benutzeranforderungen extrahiert. Diese Konzepte werden für die automatisierte oder halbautomatisierte Erstellung der Domänenontologie verwendet. Anhand dieser Domänenontologie können StakeholderInnen

dazu gebracht werden die Beziehungen zwischen den Konzepten zu verstehen und genauere Anforderungen zu stellen. Die Evaluation zeigt, dass durch den Einsatz dieser Methode die Domänenkonzepte mit hoher Genauigkeit extrahiert werden können. Die Anzahl der extrahierten Domänenkonzepte ist höher als die Anzahl, die durch ein/eine SoftwareentwicklerIn erreicht wurde.

Falkner et al. [13] stellen das OpenReq Classification Service (ORCS) vor, das Bestandteil vom OpenReq-Ansatz ist [29]. Anhand vom ORCS können Anforderungen aus einem Request for Proposal extrahiert und irrelevante Texte aussortiert werden. Hierfür verwendet ORCS Naïve Bayes Klassifikatoren, die zu den überwachten Algorithmen des maschinellen Lernens gehören. Die während der Evaluation ermittelten Ergebnisse anhand der Performanz-Metriken sind gut, es muss aber trotzdem der Recall erhöht werden. Der Recall repräsentiert in diesem Fall die Vollständigkeitsquote der Anforderungen, die in Zukunft beispielsweise durch den Einsatz von NLP erhöht werden soll.

Gupta und Gupta [17] stellen den halbautomatisierten und abhängigkeitsbasierten Ansatz CDBR vor, der zur kollaborativen Priorisierung von Anforderungen eingesetzt werden kann. Durch CDBR wird eine Prioritätenreihenfolge von Anforderungen unter Berücksichtigung der Prioritätenzuweisung von StakeholderInnen sowie EntwicklerInnen und der Abhängigkeit zwischen den Anforderungen erzeugt. CDBR verwendet Particle Swarm Optimization, um die Meinungsverschiedenheiten der Prioritätenzuweisungen zwischen StakeholderInnen und EntwicklerInnen zu minimieren und die endgültigen Prioritäten für die Anforderungen zu erzeugen. Die Evaluation zeigt, dass CDBR eine akzeptable Prioritätenliste erstellen kann und das sogar auch bei einer großen Anzahl von Anforderungen.

Onyeka et al. [27] stellen das KI-Tool COTIR vor, anhand dem implizierte Anforderungen in den Anforderungsspezifikationen identifiziert werden können. Dabei handelt es sich um Anforderungen, die nicht spezifiziert werden, aber indirekt in expliziten Anforderungen enthalten sind. Für diese Identifizierung verwendet COTIR Commonsense Knowledge, Ontologie und Text Mining. Commonsense Knowledge und die Ontologie dienen als Wissensrepräsentation. Text Mining besteht aus einem NL-Prozessor, der NLP verwendet, einem Feature-Extraktor und einem heuristischen Klassifikator. Die Evaluation zeigt, dass durch den Einsatz von COTIR der manuelle Aufwand zur Identifizierung der Quellen von impliziten Anforderungen reduziert und die Softwarequalität durchschnittlich um circa 20 % verbessert wird. Dieses Tool soll in Zukunft weiterhin optimiert werden.

Abualhaija et al. [2] stellen einen Ansatz vor, der auf überwachtem maschinellem Lernen basiert. Dieser grenzt in einer textuellen Anforderungsspezifikation Anforderungen und irrelevante Informationen voneinander ab. Dafür wird die Anforderungsspezifikation anhand von NLP verarbeitet, woraus eine Anforderungsspezifikation mit Annotationen resultiert. Danach findet anhand von einem vorab trainierten Klassifizierungsmodell die binäre Klassifizierung statt. Dabei weist sich Random Forest mit kostensensitivem Lernen nach dem Vergleich mit anderen Algorithmen als der geeignetste Klassifikationsalgorithmus für diesen Ansatz aus. Das resultierende Ergebnis ist eine Anforderungsspezifikation, in der gekennzeichnet wurde, was eine und was keine Anforderung ist. Die Evaluation zeigt anhand der Performanz-Metriken, dass dieser Ansatz im Gegensatz zu anderen Modellen bessere Ergebnisse liefert. Der manuelle Aufwand der Abgrenzung von Anforderungen und irrelevanten Informationen kann aber nicht vollständig aufgelöst werden.

Chatterjee et al. [7] stellen einen Ansatz vor, durch den funktionale Anforderungen aus einer Software Requirements Spezifikation automatisch identifiziert und klassifiziert werden, die Einfluss auf die Architektur des

Softwaresystems haben. Dieser Ansatz basiert auf dem Recurrent Neural Network Bi-LSTM Attention. Bi-LSTM ist ein Deep-Learning-Ansatz. Die Evaluation zeigt, dass dieser Ansatz bei der Identifizierung und Klassifizierung von funktionalen Anforderungen bessere Ergebnisse liefert als ein anderer Ansatz, der Naïve Bayes für die Identifizierung und RAndom K-labEL Naïve Bayes für die Klassifizierung nutzt.

Ibrahim et al. [19] stellen ein Modell vor, das auf maschinellem Lernen basiert und die Auswahl der Erhebungstechnik von Anforderungen automatisieren kann. Zu den Erhebungstechniken gehören Techniken wie Interviews und Brainstorming dazu. Die Auswahl solcher Erhebungstechniken wird normalerweise von Requirements Engineers getroffen. Für die Automatisierung wird der k-Nearest-Neighbors Algorithmus und die Maschinensprache verwendet. Die Evaluation zeigt anhand von einem Vergleich der Ergebnisse, die durch dieses Modell und der Ergebnisse, die durch Experten und Expertinnen erzielt wurden, dass das auf maschinellem Lernen basierte Modell positive Ergebnisse liefert und die Abhängigkeit von Expertenwissen verringert. Dennoch kann bei diesem Modell nicht komplett auf die Rollen AnalytikerIn oder Experte und Expertin verzichtet werden. Um die positiven Ergebnisse bestätigen zu können, muss dieses Modell erneut getestet werden.

Quba et al. [32] stellen einen Ansatz vor, bei dem maschinelles Lernen verwendet wird, um Softwareanforderungen automatisch in funktionale und nicht-funktionale Anforderungen zu klassifizieren. Dabei wird die Bag-of-Words Technik in Kombination mit den zwei Algorithmen des maschinellen Lernens Support Vector Machine und k-Nearest-Neighbor verglichen. Bag-of-Words ist eine Technik, mit der Texte vektorisiert werden können. Die Evaluation zeigt anhand von Precision, Recall und F1-Score, dass die Bag-of-Words Technik in Kombination mit der Support Vector Machine bessere Ergebnisse bei der Klassifizierung von Anforderungen erzielt. Dabei stellt der Recall die Vollständigkeitsquote und die Precision die Genauigkeitsquote der klassifizierten Anforderungen dar. F1-Score ist die Kombination von Precision und Recall. In Zukunft soll die Precision durch Anpassung des Ansatzes erhöht werden.

Baker et al. [3] stellen einen Ansatz vor, der ein neuronales Netzwerk verwendet, um nicht-funktionale Anforderungen automatisch in die Kategorien Wartbarkeit, Bedienbarkeit, Performanz, Sicherheit und Gebrauchstauglichkeit zu klassifizieren. Dabei wird der Ansatz anhand von einem Artificial Neural Network und einem Convolutional Neural Network verglichen. Das Convolutional Neural Network erzielt anhand von Precision, Recall und F1-Score bei der Klassifizierung der nicht-funktionalen Anforderungen bessere Ergebnisse. Dieser Ansatz kann nicht-funktionale Anforderungen effektiv klassifizieren und den manuellen Aufwand der Klassifizierung reduzieren. Trotzdem soll dieser Ansatz in Zukunft weiterhin ausgebaut werden, um noch bessere Ergebnisse zu erzielen.

5.2 Anforderungsdokumentation

Nasiri et al. [26] stellen einen Ansatz vor, durch den User Stories automatisch in ein UML Klassendiagramm umgewandelt werden können. Für die Umwandlung müssen Elemente wie Klassen, Attribute, Beziehungen und Operationen von Klassen aus den User Stories extrahiert werden. Dafür werden die User Stories anhand von dem NLP Tool Standford CoreNLP verarbeitet. Die Evaluation zeigt anhand von einem Vergleich der manuellen und automatisierten Extraktion der Elemente, dass die automatische Extraktion gute Ergebnisse liefert.

Elallaoui et al. [11] stellen ein Plugin vor, durch das User Stories automatisch in UML Use Case Diagramme umgewandelt werden. Dafür werden die User Stories anhand von dem NLP Tool TreeTagger verarbeitet.

Durch dieses Tool wird für alle User Stories jeweils ein Syntaxbaum erstellt, der die Extraktion von Akteuren, Anwendungsfällen und deren Beziehungen erleichtert. Die Evaluation zeigt anhand von einem Vergleich der manuellen und automatisierten Modellierung, dass die automatisierte Modellierung positive Ergebnisse liefert. Die Ergebnisse indizieren, dass der Einsatz des Plugins Zeit bei der Modellierung einspart. Das Plugin muss aber weiterhin optimiert werden, damit beispielsweise auch komplizierte Sätze mit Inclusion- und Exclusion-Beziehungen unterstützt werden.

Abdelnabi et al. [1] stellen einen Ansatz vor, durch den aus englischen Anforderungen UML Klassen-, Anwendungsfall-, Aktivitäts-, Sequenz- und Kollaborationsdiagramme generiert werden können. Dafür werden NLP und heuristische Regeln verwendet. Für NLP wird das Tool Standford CoreNPL eingesetzt. Die Evaluation zeigt anhand von einem Vergleich von diesem Ansatz mit anderen Ansätzen, dass dieser Ansatz in der Lage ist, fünf UML Diagrammarten zu erstellen und Analysten bei der Erstellung von Diagrammen eine Unterstützung ist. Dadurch können Kosten, Zeit und Aufwand reduziert werden. Dieser Ansatz ist eine Erweiterung zum Ansatz von Maatuk und Abdelnabi [24], weshalb auf [24] in diesem Kapitel nicht näher eingegangen wird.

Jebril et al. [21] stellen den manuellen und halbautomatisierten Ansatz AAEAA vor, durch den Akteure sowie Akteurinnen und Anwendungsfälle aus textuellen funktionalen Anforderungen extrahiert werden können. Für die Extraktion der Akteure sowie der Akteurinnen und Anwendungsfälle wird NLP eingesetzt. Genauer wird dafür die Semantic Role Labeling Demo Software verwendet. Die Erzeugung der Anwendungsfalldiagramme erfolgt über eine weitere Software. Die Evaluation zeigt, dass der Einsatz von AAEAA Analysten beim Erstellen von Anwendungsfalldiagrammen unterstützen und somit den Aufwand reduzieren kann, der bei der manuellen Erstellung entsteht. In Zukunft soll dieser Ansatz erweitert werden, sodass in einer Entwicklungsumgebung sowohl die Extraktion als auch die Generierung der Anwendungsfalldiagramme stattfindet.

Bashir et al. [4] stellen READ vor, ein Tool, durch das UML Klassendiagramme anhand von Anforderungen generiert werden können. Dafür verwendet das System NLP und Domain Ontologie. READ besteht aus einem NLP Modul, einem READ Knowledge Extraction Modul und einem Verfeinerungsmodul. Das NLP Modul vorverarbeitet die Anforderungen und im READ Knowledge Extraction Modul findet die Extraktion von Klassen, Attributen, Methoden und Beziehungen anhand von heuristischen Regeln statt. Zum Schluss kommt die Domain Ontologie im Verfeinerungsmodul zum Einsatz, wobei irrelevante Elemente entfernt werden. Die Evaluation zeigt, dass READ im Vergleich zu anderen Tools einen höheren Recall erzielt, die Precision aber in Zukunft verbessert werden muss.

5.3 Anforderungserhebung und -dokumentation

Do et al. [9] stellen ein Framework vor, das frei verfügbare Anforderungen anderer vergleichbarer Software nutzt und anhand dieser automatisch kreative Anforderungen generiert. Dafür setzt das Framework NLP und maschinelles Lernen ein. Die gefundenen Anforderungen durchlaufen einen Clustering-Prozess. In diesem Prozess werden die maschinellen Lerntechniken doc2vec und BIRCH eingesetzt. Durch NLP werden Prozesse, Objekte und zusätzliche Details aus den Anforderungen extrahiert. Die extrahierten Attribute werden in einer vereinfachten Form der Satzschablonenvorlage von Chris Rupp [31] eingefügt, um die neuen Anforderungen zu generieren. Die Evaluation zeigt, dass das vorgestellte Framework fähig ist, kreative Anforderungen zu generieren und innovatives Denken bei den EntwicklerInnen, die verschiedene Erfahrungen besitzen, zu fördern. Das Framework soll in Zukunft

weiterhin optimiert werden.

Panichella und Ruiz [30] stellen Requirements-Collector vor, ein Tool, dass zum einen das Analysieren des Benutzerfeedbacks und zum anderen die Dokumentation von Anforderungen aus RE-Meetings automatisiert. Das Benutzerfeedback und die Anforderungen werden anhand von maschinellem Lernen und Deep-Learning klassifiziert. Die Evaluation zeigt, dass die Klassifizierung des Benutzerfeedbacks und der Anforderungen aus den RE-Meetings sowie die generelle Extraktion der Anforderungen zuverlässig ist.

Yanuarifiani et al. [37] stellen ein regelbasiertes Ontologie-Framework (ROF) vor, durch das Anforderungsspezifikationen automatisch generiert werden. Dafür wird zu Beginn eine Anforderungsontologie erstellt, die alle durch Gap-Analyse, Interviews und Brainstorming ermittelten Anforderungen beinhaltet. Diese Anforderungsontologie wird automatisch erstellt, indem die Grammatik anhand von Part-of-Speech-Funktionen durch eine englische Bibliotheksdatenbank identifiziert wird. Danach wird durch die Verwendung dieser Anforderungsontologie und anhand von regelbasierten Methoden automatisch eine Anforderungsspezifikation erstellt, wozu ein Software Requirements Specification Dokument und ein Business Process Model and Notation Diagramm dazu gehört. Die Evaluation anhand von einem Informationssysteme Projekt zeigt, dass das Framework den manuellen Aufwand für die Erstellung der Anforderungsontologie sowie der Anforderungsspezifikation reduziert.

5.4 Anforderungsverifikation und -validierung

Dalpiaz et al. [8] stellen einen Ansatz vor, der sich aus Techniken der Informationsvisualisierung und einem NLP gestützten Algorithmus zusammen setzt und Analysten dabei hilft, Unvollständigkeit und Mehrdeutigkeit in User Stories zu entdecken. Der eingesetzte NLP gestützte Algorithmus erkennt Synonymie zwischen Begriffen, die Mehrdeutigkeiten auslösen können. Dafür integriert der NLP gestützte Algorithmus den Cortical.io Algorithmus, welches auf der Semantic Folding Theory basiert, um semantische Ähnlichkeiten zwischen Begriffen zu berechnen. Die Evaluation zeigt, dass dieser Ansatz im Vergleich zur manuellen Inspektion die Vollständigkeit beim Auffinden von Mehrdeutigkeiten und fehlenden Anforderungen erhöhen kann.

Ezzini et al. [12] stellen einen Ansatz vor, der NLP nutzt, um Mehrdeutigkeiten in Anforderungen zu erkennen sowie zu interpretieren. Zudem basiert dieser Ansatz auf der automatisierten Generierung eines domänenspezifischen Korpus aus Wikipedia. Die Evaluation zeigt, dass das Einbeziehen von Domänenwissen zu einer signifikanten Verbesserung der Genauigkeit bei der Erkennung und Interpretation von Mehrdeutigkeit führt. Durch diesen Ansatz können nicht alle Mehrdeutigkeiten gefunden werden und daher ist ein wenig manueller Aufwand von Analysten noch notwendig. Trotzdem hat dieser Ansatz Potenzial in der Praxis eine Unterstützung zu sein.

5.5 Anforderungserhebung, -verifikation und -validierung

Sonbol et al. [36] stellen ReqVec vor, eine Methode, wodurch aus dem Anforderungstext ein ReqVec erzeugt wird. Anhand des ReqVec wird eine Anforderung als semantische Vektordarstellung dargestellt. Diese Vektordarstellung kann für die Erkennung semantischer Anforderungsbeziehungen und für die Anforderungsklassifikation eingesetzt werden. Für diese Methode werden NLP in Form vom Natural Language Toolkit, Heuristiken und Word2vec verwendet. Durch Word2vec wird jedes Wort anhand von einem Vektor dargestellt. Die Evaluation zeigt, dass ReqVec positive Ergebnisse bei der Erkennung semantischer Anforderungsbeziehungen und der Klassifizierung

von Anforderungen erzielt.

5.6 Anforderungsverwaltung

Mezghani et al. [25] stellen einen Ansatz vor, durch den die Traceability-Matrix [16] von zwei Anforderungsspezifikationen automatisch generiert werden kann. Dabei wird der k-Means-Algorithmus eingesetzt, der zu den unüberwachten maschinellen Lernverfahren gehört. Durch diesen Algorithmus werden ähnliche Anforderungen aus den zwei Anforderungsspezifikationen gruppiert. Die Evaluation zeigt, dass dieser Ansatz für Anforderungen ohne Fachbegriffe und mit jeweils einer Verbindung zwischen Anforderungen gute Ergebnisse erzielt. Anforderungen, die Fachbegriffe enthalten, müssen vor der Anwendung des k-Means-Algorithmus erst durch NLP vorverarbeitet werden, um gute Ergebnisse zu erzielen. Generell eignet sich dieser Ansatz nur für ein bis zwei Verbindungen zwischen Anforderungen. Durch diesen Ansatz kann eine Traceability-Matrix generiert werden, um den manuellen Aufwand zu reduzieren. Dieser Ansatz soll weiterhin verbessert werden.

Samer et al. [35] stellen einen Ansatz von einem content-based Recommender-System vor, durch welches automatisch Abhängigkeiten vom Typ requires zwischen Anforderungen erkannt und empfohlen werden können. Dabei werden zwei Ansätze von content-based Recommender-Systemen in unterschiedlichen Kombinationen verglichen. Die Evaluation zeigt, dass der Ansatz des content-based Recommender-Systems basierend auf Random Forest, das Probabilistic Features verwendet, die besten Ergebnisse liefert. Durch Probabilistic Features werden die gemeinsam auftretenden Wörter in zwei Anforderungen gemessen, was wichtige Informationen für den Klassifikator darstellen. Durch diesen Ansatz kann die Ermittlung von Abhängigkeiten unterstützt werden.

5.7 Anforderungserhebung und -verwaltung

Haider et al. [18] stellen ein Framework vor, das KI für die Priorisierung und Traceability von Anforderungen nutzt. Dabei wird das RE in der globalen Softwareentwicklung betrachtet. Für die Priorisierung und Traceability der Anforderungen werden die maschinellen Lerntechniken Case-Based Reasoning und j48 verwendet. Dabei ist der j48-Algorithmus im Weka-Tool vorhanden. Die Evaluation zeigt, dass der Einsatz des Frameworks gute Ergebnisse erzielt. Durch das Framework kann die Priorisierung und Traceability von Anforderungen bei einer globalen Softwareentwicklung verbessert werden.

Die Ergebnisse werden als Übersicht in der Tabelle 3 dargestellt.

Tabelle 3: Übersicht der Ergebnisse

Quelle	KI	RE-Aktivität
[38]	Ontology Learning	E
[13]	Naïve Bayes Klassifikatoren	E
[17]	Particle Swarm Optimization	E
[27]	Commonsense Knowledge, Ontologie und Text Mining	E
[2]	NLP und Random Forest mit kostensensitivem Lernen	E
[7]	Recurrent Neural Network (Bi-LSTM Attention)	E
[19]	k-Nearest-Neighbors Algorithmus	E
[32]	Bag-of-Words in Kombination mit Support Vector Machine	E
[3]	Convolutional Neural Network	E
[26]	NLP (Standford CoreNLP)	D
[11]	NLP (TreeTagger)	D

Quelle	KI	RE-Aktivität
[1] [24]	NLP und heuristische Regeln	D
[21]	NLP (Semantic Role Labeling)	D
[4]	NLP und Domain Ontologie	D
[9]	NLP und Techniken des maschinellen Lernens (doc2vec + BIRCH)	E und D
[30]	Maschinelles Lernen und Deep-Learning	E und D
[37]	Part-of-Speech Tagging und Ontologie	E und D
[8]	NLP gestützter Algorithmus, der Cortical.io integriert	VV
[12]	NLP	VV
[36]	NLP, Heuristiken und Word2vec	E und VV
[25]	k-Means-Algorithmus	V
[35]	Random Forest, das Probabilistic Features nutzt	V
[18]	Case-Based Reasoning und j48 (Weka-Tool)	E und V

Legende: E = Erhebung; D = Dokumentation; VV = Verifikation und Validierung; V= Verwaltung

6 Diskussion und Fazit

Dieses SLR verschafft einen Überblick über den Einsatz von KI im RE. Dabei werden die Formen der KI aus 24 Artikel den RE-Aktivitäten nach Gilz [15] zugeordnet, in denen sie Unterstützung leisten. Beim Vergleich der Ergebnisse dieses SLRs mit den Ergebnissen von Breuninger et al. [6] wird deutlich, dass auch die Datenbanken SpringerLink und IEEExplore relevante Artikel zum Einsatz von KI im RE liefern. Es werden in diesem SLR 20 weitere Artikel behandelt, die im SLR von Breuninger et al. [6] nicht abgedeckt werden. Daher ist dieses SLR eine Erweiterung zu den bereits existierenden SLRs. Anhand der Ergebnisse in Kapitel 5 wird deutlich, dass der Einsatz von Formen der KI nicht den vollständigen manuell erbrachten Aufwand von Requirements Engineers vermeiden kann. Durch die vorgestellten Ansätze werden beispielsweise nicht alle Anforderungen gefunden, die aus Dokumenten extrahiert werden sollen. Trotzdem kann der Einsatz von Formen der KI den manuellen Aufwand reduzieren, den die Requirements Engineers erbringen müssen. Anhand der Tabelle 4 wird ersichtlich, dass bei der generellen Betrachtung des REs vor allem NLP und maschinelles Lernen eingesetzt wird. Bei der spezifischeren Betrachtung der einzelnen RE-Aktivitäten wird anhand der Tabelle 4 ersichtlich, dass bei der Anforderungserhebung und -verwaltung vor allem maschinelles Lernen und bei der Anforderungsdokumentation, sowie der Anforderungsverifikation und -validierung vor allem NLP als Unterstützung eingesetzt wird. Aus Kapitel 5 ist zudem ersichtlich, dass nicht

Tabelle 4: Prozentuale Auswertung vom Einsatz der KI im RE

RE-Aktivität / KI	Generell RE	E	D	VV	V
NLP	43,75 %	≈ 28,571 %	≈ 61,538 %	75 %	0 %
Maschinelles Lernen	34,375 %	≈ 42,857 %	≈ 15,385 %	25 %	100 %
Knowledge Representation	12,5 %	≈ 14,286 %	≈ 15,385 %	0 %	0 %
Deep Learning	3,125 %	≈ 4,762 %	≈ 7,692 %	0 %	0 %
Neuronale Netze	6,25 %	≈ 9,524 %	0 %	0 %	0%

Legende: E = Erhebung; D = Dokumentation; VV = Verifikation und Validierung; V= Verwaltung

alle RE-Aktivitäten gleichermaßen von KI unterstützt werden. In den 24 Artikeln wird die Anforderungserhebung zu circa 48,276 %, die Anforderungsdokumentation zu circa 31,034 %, die Anforderungsverifikation und -validierung zu circa 10,345 % und die Anforderungsverwaltung zu circa 10,345 % durch Formen der KI unterstützt. Dadurch wird deutlich, dass viel mehr Forschungsarbeiten zu den RE-Aktivitäten Anforderungserhebung und -dokumentation als zu den RE-Aktivitäten Anforderungsverifikation sowie -validierung und -verwaltung existieren. Daher sollte der Einsatz von KI in diesen RE-Aktivitäten in Zukunft weiter untersucht werden.

Literaturverzeichnis

[1] A. Abdelnabi, E., M. Maatuk, A., und M. Abdelaziz, T. 2021. An Algorithmic Approach for Generating Behavioral UML Models Using Natural Language Processing. In *The 7th International Conference on Engineering & MIS 2021 (ICEMIS'21)*. ACM, New York, NY, USA, Artikel 36, 1–6. DOI: 10.1145/3492547.3492612.

[2] Abualhaija, S., Arora, C., Sabetzadeh, M., Briand, L. C., und Vaz, E. 2019. A Machine Learning-Based Approach for Demarcating Requirements in Textual Specifications. In *2019 IEEE 27th International Requirements Engineering Conference (RE)*. IEEE, 51–62. DOI: 10.1109/RE.2019.00017.

[3] Baker, C., Deng, L., Chakraborty, S., und Dehlinger, J. 2019. Automatic Multi-class Non-Functional Software Requirements Classification Using Neural Networks. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*. IEEE, 610–615. DOI: 10.1109/COMPSAC.2019.10275.

[4] Bashir, N., Bilal, M., Liaqat, M., Marjani, M., Malik, N., und Ali, M. 2021. Modeling Class Diagram using NLP in Object-Oriented Designing. In *2021 National Computing Colleges Conference (NCCC)*. IEEE, 1–6. DOI: 10.1109/NCCC49330.2021.9428817.

[5] Batarseh, F. A., Mohod, R., Kumar, A., und Bui, J. 2020. The application of artificial intelligence in software engineering: a review challenging conventional wisdom. In *Data Democracy*. Elsevier, 179–232. DOI: 10.1016/B978-0-12-818366-3.00010-1.

[6] Breuninger, J., Kücher, F., und Misic, N. 2021. Künstliche Intelligenz im Requirements Engineering. Eine Systematische Literaturrecherche. In *SKILL 2021 - Studierendenkonferenz Informatik - Komplettband. (Hrsg.), SKILL 2021. Gesellschaft für Informatik, Bonn*, 63–74. ISBN: 978-3-88579-751-7.

[7] Chatterjee, R., Ahmed, A., und Anish, P. R. 2020. Identification and Classification of Architecturally Significant Functional Requirements. In *2020 IEEE Seventh International Workshop on Artificial Intelligence for Requirements Engineering (AIRE)*. IEEE, 9–17. DOI: 10.1109/AIRE51212.2020.00008.

[8] Dalpiaz, F., van der Schalk, I., und Lucassen, G. 2018. Pinpointing Ambiguity and Incompleteness in Requirements Engineering via Information Visualization and NLP. In *Requirements Engineering: Foundation for Software Quality (REFSQ)*, E. Kamsties, J. Horkoff und F. Dalpiaz, Hrsg. Lecture Notes in Computer Science. Springer International Publishing AG, Cham, 119–135. DOI: 10.1007/978-3-319-77243-1_8.

[9] Do, Q. A., Chekuri, S. R., und Bhowmik, T. 2019. Automated Support to Capture Creative Requirements via Requirements Reuse. In *Reuse in the Big Data Era (ICSR)*, X. Peng, A. Ampatzoglou und T. Bhowmik, Hrsg. Lecture Notes in Computer Science. Springer Nature Switzerland AG, Cham, 47–63. DOI: 10.1007/978-3-030-22888-0_4.

[10] Ebert, C. 2019. *Systematisches Requirements Engineering. Anforderungen ermitteln, dokumentieren, analysieren und verwalten* (6. überarb. und erw. Aufl.). dpunkt.verlag GmbH, Heidelberg. ISBN: 978-3-86490-562-9.

[11] Elallaoui, M., Nafil, K., und Touahni, R. 2018. Automatic Transformation of User Stories into UML Use Case Diagrams using

NLP Techniques. *Procedia Computer Science* 130, 3, 42–49. DOI: 10.1016/j.procs.2018.04.010.

[12] Ezzini, S., Abualhaija, S., Arora, C., Sabetzadeh, M., und Briand, L. C. 2021. Using Domain-Specific Corpora for Improved Handling of Ambiguity in Requirements. In *2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE)*. IEEE, 1485–1497. DOI: 10.1109/ICSE43902.2021.00133.

[13] Falkner, A., Palomares, C., Franch, X., Schenner, G., Aznar, P., und Schoerghuber, A. 2019. Identifying Requirements in Requests for Proposal: A Research Preview. In *Requirements Engineering: Foundation for Software Quality (REFSQ)*, E. Knauss und M. Goedicke, Hrsg. Lecture Notes in Computer Science. Springer Nature Switzerland AG, Cham, 176–182. DOI: 10.1007/978-3-030-15538-4_13.

[14] Galton, A. 2009. Spatial and temporal knowledge representation. *Earth Sci Inform* 2, 3, 169–187. DOI: 10.1007/s12145-009-0027-6.

[15] Gilz, T. 2014. Requirements Engineering und Requirements Management. In *Modellbasierte virtuelle Produktentwicklung*, M. Eigner, D. Roubanov und R. Zafirov, Hrsg. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, 53–75. DOI: 10.1007/978-3-662-43816-9_3.

[16] Gotel, O., Cleland-Huang, J., Hayes, J. H., Zisman, A., Egyed, A., Grünbacher, P., Dekhtyar, A., Antoniol, G., Maletic, J., und Mäder, P. 2012. Traceability Fundamentals. In *Software and Systems Traceability*, J. Cleland-Huang, O. Gotel und A. Zisman, Hrsg. Springer-Verlag London, London, 3–22. DOI: 10.1007/978-1-4471-2239-5_1.

[17] Gupta, A. und Gupta, C. 2022. CDBR: A semi-automated collaborative execute-before-after dependency-based requirement prioritization approach. *Journal of King Saud University - Computer and Information Sciences (JKSUCI)* 34, 2, 421–432. DOI: 10.1016/j.jksuci.2018.10.004.

[18] Haider, W., Hafeez, Y., Ali, S., Jawad, M., Ahmad, F. B., und Rafi, M. N. 2019. Improving Requirement Prioritization and Traceability using Artificial Intelligence Technique for Global Software Development. In *2019 22nd International Multitopic Conference (INMIC)*. IEEE, 1–8. DOI: 10.1109/INMIC48123.2019.9022775.

[19] Ibrahim, H. M. E., Ahmad, N., Rehman, M. B., Ahmad, I., und khan, R. 2019. Implementing and Automating Elicitation Technique Selection using Machine Learning. In *2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*. IEEE, 564–569. DOI: 10.1109/ICCIKE47802.2019.9004398.

[20] Jantzer, M., Nentwig, G., Deininger, C., und Michl, T. 2019. *Die Kunst, eine Produktentwicklung zu führen. Erfolgreiche Konzepte aus der Unternehmenspraxis*. Springer-Verlag GmbH Deutschland, Berlin, Heidelberg. DOI: 10.1007/978-3-662-57899-5.

[21] Jebril, E. M., Imam, A. T., und Al-Fayuomi, M. 2018. An algorithmic approach to extract actions and actors (AAEAA). In *Proceedings of the International Conference on Geoinformatics and Data Analysis (ICGDA '18)*. ACM, New York, NY, USA, 13–17. DOI: 10.1145/3220228.3220247.

[22] Kitchenham, B.A. und Charters, S. Charters. 2007. *Guidelines for performing Systematic Literature Reviews in Software Engineering*. Version 2.3. EBSE Technical Report EBSE 2007-01. Keele University und Durham University.

[23] Kreutzer, R. T. und Sirrenberg, M. 2019. *Künstliche Intelligenz verstehen. Grundlagen – Use-Cases – Grundlagen – Use-Cases – unternehmenseigene KI-Journey*. Springer Fachmedien Wiesbaden GmbH. DOI: 10.1007/978-3-658-25561-9.

[24] M. Maatuk, A. und A. Abdelnabi, E. 2021. Generating UML Use Case and Activity Diagrams Using NLP Techniques and Heuristics Rules. In *International Conference on Data Science, E-learning and Information Systems 2021 (DATA'21)*. ACM, New York, NY, USA, 271–277. DOI: 10.1145/3460620.3460768.

[25] Mezghani, M., Kang, J., Kang, E.-B., und Sedes, F. 2019. Clustering for Traceability Managing in System Specifications. In *2019*

IEEE 27th International Requirements Engineering Conference (RE). IEEE, 257–264. DOI: 10.1109/RE.2019.00035.

[26] Nasiri, S., Rhazali, Y., Lahmer, M., und Chenfour, N. 2020. Towards a Generation of Class Diagram from User Stories in Agile Methods. *Procedia Computer Science* 170, 831–837. DOI: 10.1016/j.procs.2020.03.148.

[27] Onyeka, E., Varde, A. S., Anu, V., Tandon, N., und Daramola, O. 2020. Using Commonsense Knowledge and Text Mining for Implicit Requirements Localization. In *2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE, 935–940. DOI: 10.1109/ICTAI50040.2020.00146.

[28] Paaß, G. und Hecker, D. 2020. *Künstliche Intelligenz. Was steckt hinter der Technologie der Zukunft?* Springer Fachmedien Wiesbaden GmbH, Wiesbaden. DOI: 10.1007/978-3-658-30211-5.

[29] Palomares, C., Franch, X., und Fucci, D. 2018. Personal Recommendations in Requirements Engineering: The OpenReq Approach. In *Requirements Engineering: Foundation for Software Quality (REFSQ)*, E. Kamsties, J. Horkoff und F. Dalpiaz, Hrsg. Lecture Notes in Computer Science. Springer International Publishing AG, Cham, 297–304. DOI: 10.1007/978-3-319-77243-1_19.

[30] Panichella, S. und Ruiz, M. 2020. Requirements-Collector: Automating Requirements Specification from Elicitation Sessions and User Feedback. In *2020 IEEE 28th International Requirements Engineering Conference (RE)*. IEEE, 404–407. DOI: 10.1109/RE48521.2020.00057.

[31] Pohl, K. 2010. *Requirements Engineering. Fundamentals, Principles, and Techniques*. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg. ISBN: 978-3-662-51888-5.

[32] Quba, G. Y., Al Qaisi, H., Althunibat, A., und AlZu'bi, S. 2021. Software Requirements Classification using Machine Learning algorithm's. In *2021 International Conference on Information Technology (ICIT)*. IEEE, 685–690. DOI: 10.1109/ICIT52682.2021.9491688.

[33] Roelen-Blasberg, T. 2019. *Automatisierte Präferenzmessung. Extraktion und Evaluation von Produktattributen auf Basis von Online-Rezensionen*. Springer Fachmedien Wiesbaden, Wiesbaden. DOI: 10.1007/978-3-658-23831-5.

[34] Rupp, C. 2021. *Requirements-Engineering und -Management. Das Handbuch für Anforderungen in jeder Situation* (7. aktual. und erw. Aufl.). Hanser eLibrary. Carl Hanser Verlag München, München. ISBN: 978-3-446-45587-0.

[35] Samer, R., Stettinger, M., Atas, M., Felfernig, A., Ruhe, G., und Deshpande, G. 2019. New Approaches to the Identification of Dependencies between Requirements. In *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE, 1265–1270. DOI: 10.1109/ICTAI.2019.00-91.

[36] Sonbol, R., Rebdawi, G., und Ghneim, N. 2020. Towards a Semantic Representation for Functional Software Requirements. In *2020 IEEE Seventh International Workshop on Artificial Intelligence for Requirements Engineering (AIRE)*. IEEE, 1–8. DOI: 10.1109/AIRE51212.2020.00007.

[37] Yanuarifiani, A. P., Chua, F.-F., und Chan, G.-Y. 2020. Feasibility Analysis of a Rule-Based Ontology Framework (ROF) for Auto-Generation of Requirements Specification. In *2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*. IEEE, 1–6. DOI: 10.1109/IICAIET49801.2020.9257838.

[38] Zhang, J., Yuan, M., und Huang, Z. 2019. Software Requirements Elicitation Based on Ontology Learning. In *Software Engineering and Methodology for Emerging Domains*, Z. Li, H. Jiang, G. Li, M. Zhou und M. Li, Hrsg. Communications in Computer and Information Science. Springer Nature Singapore Pte Ltd, Singapore, 83–98. DOI: 10.1007/978-981-15-0310-8_6.

©2022 Elma Bajraktari. 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/

Affective Computing zur Diagnose und Therapie von Erkrankungen

Irem Uslu
Hochschule Reutlingen
Reutlingen, Deutschland
Irem.Uslu@Student.Reutlingen-University.de

Abstract
Die Forschung nach besseren Diagnose- und Therapiemöglichkeiten bei Erkrankungen wird immer weiter fortgeführt. Vor allem psychische Erkrankungen oder Verhaltensstörungen können nicht immer direkt diagnostiziert werden. Auch bei der Therapie stoßen Arzneimittel an ihre Grenzen. Durch die rasante Entwicklung der künstlichen Intelligenz könnten in Zukunft weitere Diagnose- und Therapiemöglichkeiten geboten werden. Dabei ist Affective Computing ein Bereich der intelligenten Computersysteme, bei dem Schnittstellen geschaffen werden, die die Gefühlslage eine:r Anwender:in mithilfe von neuen Technologien identifizieren und auf diese folglich eine Reaktion geben. Das Ziel dieser Arbeit ist es, mithilfe einer systematischen Literaturrecherche den aktuellen Forschungsstand zum möglichen Einsatz von Affective Computing zur Diagnostik und Therapie von Erkrankungen zusammenzufassen. Dabei werden die eingesetzten Methoden zur Emotionserkennung der Computersysteme und die Ergebnisse dieser Methoden aufgezeigt. In der Untersuchung hat sich ergeben, dass es derzeit Forschungen und Studien im Bereich der Diagnostik und Therapie gibt, diese jedoch noch keinen verbreiteten medizinischen Einsatz finden.

CCS Concepts
• **Computing methodologies** → *Artificial Intelligence* → Philosophical/theoretical foundations of artificial intelligence → **Cognitive Science**;

Keywords
Emotion AI, Affective Computing, Diagnosis, Therapy, Affective Technology

1 Einleitung
In unterschiedlichen Medizinbranchen wie z. B. der Psychologie wird stetig nach Möglichkeiten gesucht, die Diagnose und Therapie von Erkrankungen mithilfe neu eingeführter Technologien zu erweitern [9]. Eine neuere Technologie stellt mit dem Erkennen und Reagieren auf Emotionen das Affective Computing dar. Im MIT Media Lab der MIT University werden bspw. derzeit unterschiedliche Projekte aus unterschiedlichen Bereichen mit Affective Computing durchgeführt. Unter anderem gibt es Projekte im medizinischen Bereich wie das aSpire, das ein mobiles pneumatisch-haptisches Gerät darstellt, welches an den Bund der Hose befestigt werden kann. Dieses Gerät ermöglicht es Patient:innen mit Atemproblemen bei einer zu schnellen Atemfrequenz, welche durch Stressverhalten ausgelöst werden kann, die Reduzierung der Atemfrequenz durch eine durch leichte Berührung induziertes Feedback zu reduzieren [10]. In einem weiteren Projekt wird das Stressverhalten bei Multiple Sklerose Patient:innen untersucht.

Betreuer Hochschule: Prof. Dr. Uwe Kloos
Hochschule Reutlingen
Uwe.Kloos@Reutlingen-University.de

Informatics Inside Frühjahr 2022
11. Mai 2022, Hochschule Reutlingen
Copyright 2022 Irem Uslu

Dabei soll auch hier durch Sensoren, die am Körper befestigt werden können oder transportierbar sind, das Stresslevel dieser Patient:innen gemessen werden, um ihren Krankheitsverlauf zu verlangsamen, da Stress zu einer Verschlechterung des Krankheitsbildes führt [20]. Solche und weitere Projekte werden von unterschiedlichen Forschungsinstituten behandelt. Hierbei kann Feedback beim Affective Computing nicht nur durch physiologische Komponenten wie Sensoren erreicht werden, sondern auch durch die Erkennung von menschlicher Mimik und Gestik durch Aufnahmen oder durch die Erkennung von Emotionen aus natürlicher Sprache [13].

1.1 Ziel der Arbeit

Diese Arbeit verfolgt das Ziel, durch ein systematisches Literaturreview die wichtigsten Forschungsergebnisse zum Einsatz von Affective Computing zur Diagnose und Therapie von Erkrankungen zusammenzufassen. Der Fokus liegt auf der Beantwortung folgender Forschungsfragen:

FF1: Wie kann Affective Computing zur Therapie von Erkrankungen eingesetzt werden?

FF2: Wie kann Affective Computing zur Diagnose von Erkrankungen eingesetzt werden?

1.2 Strukturierung der Arbeit

Die Anwendung der systematischen Vorgehensweise und Auswahl der Literatur werden in Kapitel 2 vorgestellt. In Kapitel 3 werden Grundbegriffe wie Affective Computing, Emotionen und deren Erkennungsmöglichkeiten behandelt. In Kapitel 4 werden die Forschungsergebnisse aus den selektierten Literaturen der systematischen Literaturrecherche zu den beiden Forschungsfragen FF1 und FF2 zusammengefasst. Eine Diskussion und ein Vergleich zwischen den ausgesuchten Literaturen werden in Kapitel 5 durchgeführt. In Kapitel 6 wird ein Fazit zur Arbeit und ein Ausblick zum Affective Computing bei Erkrankungen und zum erzielten Literaturreview gegeben.

2 Methodik

Die Methodik dieser Arbeit stützt sich auf einem systematischen Literaturreview. Mithilfe eines systematischen Literaturreviews werden alle zugänglichen Literaturen aus den ausgewählten Datenbanken, die für die vordefinierten Forschungsfragen von großer Bedeutung sind, analysiert und zusammengetragen. Hierdurch wird ein Überblick über die bestehenden Forschungsarbeiten zum Thema Affective Computing zur Diagnose und Therapie von Erkrankungen in einer Arbeit resümiert. Um ein systematisches Literaturreview durchzuführen, werden mehrere Schritte benötigt. In dieser Arbeit wird die Beschreibung und Vorgehensweise von Kitchenham [19] und der praktische Leitfaden für eine systematische Literaturrecherche von Neiva und da Silva [22] verwendet. Dabei werden zunächst die zwei Forschungsfragen FF1 und FF2 aus Kapitel 1.1 gewählt, um die Diagnose von Erkrankungen mittels Affective Computing von der Therapie zu trennen. Anschließend werden die wissenschaftlichen Datenbanken zur Suche nach den Literaturen festgelegt. Für die Untersuchung und Selektion der Literaturen dieser Arbeit werden die Suchdatenbanken *SpringerLink*, *IEEE Xplore*, *ACM Digital Library* und *ScienceDirect* verwendet. Um in den genannten Datenbanken nach geeigneter Literatur suchen zu können, die die Forschungsfragen beantworten, wird wie in Tabelle 1 zu sehen pro Forschungsfrage ein Suchterm erstellt. Dabei ist der Suchterm S1 für FF1 und der Suchterm 2 für FF2. Mithilfe dieser Suchterme wird in den Datenbanken gesucht. Innerhalb der Suchterme werden mehrere Synonyme verwendet, um die Ergebnisquote zu erhöhen. Die Literaturen, die nach der Eingabe der Suchterme resultieren, werden mithilfe von Ein- und Ausschlusskriterien, die in Tabelle 2 und 3 aufgelistet sind, gefiltert. Die Einschlusskriterien definieren, welche Informationen oder Anforderungen ein Paper enthalten muss. Die Ausschlusskriterien sind Gründe, weshalb eine bestimmte Literatur nicht verwendet wird.

Tabelle 1: Suchterme zu den Forschungsfragen

ID	Suchterm
S1	(“affective computing“ OR “emotion ai“ OR “emotion recognition” OR “affective therapy“ OR "affective technology" OR “emotion measurement”) AND (“therapy” OR "treatment")
S2	(“affective computing“ OR “emotion ai“ OR “affective therapy“ OR "affective technology" OR “emotion measurement”) AND (“diagnose” OR "diagnosis")

Diese Ein- und Ausschlusskriterien helfen zur Eingrenzung der Literaturergebnisse, welche nach ihrem Inhalt analysiert und je nach Eignung zur Beantwortung der Forschungsfrage erneut aussortiert werden. Die einzelnen Schritte, wie bei der Selektion vorgegangen wird, können aus der Tabelle 4 entnommen werden. Hier wird die Anzahl der Suchergebnisse in den Datenbanken nach Eingabe der Kriterien wie Suchterme, Ein- und Ausschlusskriterien, Titel und Keywords schrittweise angegeben. In den Schritten wird verdeutlicht, welche dieser Kriterien in welchem Schritt betrachtet werden.

Tabelle 2: Einschlusskriterien bei der Literaturauswahl

ID	Einschlusskriterien
EK1	Literatur zwischen den Jahren 2017 und 2021
EK2	Literatur, die zur Beantwortung und Ergebnisführung der Forschungsfrage beiträgt
EK3	Literatur, die öffentlich oder durch die Hochschule Reutlingen zugänglich ist
EK4	Der Inhalt muss ein Forschungsartikel sein

Tabelle 3: Ausschlusskriterien bei der Literaturauswahl

ID	Ausschlusskriterien
AK1	Literaturen die keinem Peer-Review unterzogen wurden
AK2	Literaturen, die in den Datenbanken doppelt vorkommen
AK3	SLR zu diesem Thema

Zum Schluss resultieren für die FF1 13 Ergebnisse und für die Forschungsfrage FF2 10 Ergebnisse, die für die Beantwortung der Forschungsfragen relevant sind. Die Inhalte der Paper werden in Kapitel 4 näher vorgestellt.

Tabelle 4: Anzahl der Ergebnisse

ID	Eingabe	Anzahl der Suchergebnisse				
		Springer Link	IEEE Xplore	ACM Digital Library	Science Direct	Gesamt
1	ST1	511	400	637	516	2064
	ST2	322	132	128	228	810
2	EK 1, EK4 und AK1 für ST1	93	246	279	262	880
	EK 1, EK4 und AK1 für ST2	45	85	87	129	346
3	AK2 & EK3 für ST1	65	246	224	77	592
	AK2 & EK3 für ST2	45	82	68	38	235
4	AK3, Titel & Keywords für ST1	11	39	41	24	123
	AK3, Titel & Keywords für ST2	9	15	10	14	48
5	Abstract für ST1	5	17	11	7	40
	Abstract für ST2	4	8	5	7	33
6	EK2 für ST1	1	6	4	2	**13**
	EK2 FÜR ST2	0	3	3	2	**8**

3 Grundlagen

Menschen werden in ihren Handlungsweisen oder Reaktionen in bestimmten Situationen durch Emotionen oder Stimmungen angetrieben. Emotionen sind Gefühle, die den Menschen ermöglichen eine Beurteilung einer Sachlage durchzuführen und auf diese dementsprechend zu reagieren. Diese Emotionen oder Stimmungen können zusammengefasst werden als Affekte. Affekte sind dabei kurz anhaltende Gefühle oder Gefühlsänderungen, die durch bestimmte äußere Umwelteinflüsse oder Handlungen hervorgerufen werden. Diese Änderungen können in Mimik, Gestik, Sprache oder dem physiologischen Zustand der Menschen erkannt werden. Unter Affective Computing wird hierbei die Erkennung der Veränderung in der Mimik, Stimme, Gestik oder inneren Physiologie verstanden, um die Emotionen des Menschen detektieren und demententsprechend darauf reagieren zu können [2].

Für die Emotionserkennung ist es wichtig, Emotionen zunächst in Kategorien zu unterteilen. Es gibt hierbei mehrere Emotionen, die unterschiedlich definiert werden. Laut Plutchik wird zwischen den Emotionen Wut, Angst, Trauer, Ekel, Überraschung, Neugier, Akzeptanz und Freude unterschieden. Ekman jedoch lässt die Emotionen Neugier und Akzeptanz raus. Die Ekman Betrachtung ist hierbei verbreiteter, da diese Emotionen unabhängig von äußeren subjektiven Faktoren sind [37][13]. Im diagnostischen und therapeutischen Kontext betrachtet, können Emotionen und Affekte Einsicht über die Gefühlslage von Patient:innen geben. Die Gefühlslage von Patient:innen kann verdeutlichen, ob diese unter Stress stehen, Schmerz erleiden oder wie sie sich im Allgemeinen fühlen. Durch diese Erkenntnis können eventuelle Erkrankungen diagnostiziert oder auch Therapiemöglichkeiten für diese Patient:innen entwickelt werden. Die Diagnose ist dabei die Feststellung einer Erkrankung und die Therapie, die Schritte, die eingeleitet werden, um Patient:innen heilen zu können. Die Bestimmung der Emotionen von Patient:innen kann hierbei anhand unterschiedlicher Faktoren oder gegebenenfalls mittels äußerer Sensoren erfolgen [4][13].

3.1 Emotionserkennung

In diesem Kapitel wird die Erkennung von Emotionen in der Sprache, in der Mimik, in der Gestik und durch physiologische Zustände näher erläutert.

3.1.1 Sprache

Die Sprache ist ein Kommunikationsweg, den Menschen täglich nutzen. Durch diese können viele sachliche, aber auch emotionale Informationen preisgegeben werden. Alles, was ein Mensch sagt, kann emotional interpretiert werden. Diese Informationsübertragung kann zum einen durch den Primärkanal erfolgen. Diese ist die direkte Sprache, wie der Mensch die Dinge ausspricht. Der andere Weg ist der Sekundärkanal, der die Lautstärke oder die Körpersprache, die beim Sprechen ausgeübt wird, ausmacht [31][13]. Zudem kann die Erkennung von Gefühlsänderungen durch die Sprache auch textuell erfolgen [2]. Die Messung dieser Worte und das Erkennen von Emotionen aus diesen Worten kann durch Natural Language Processing Methoden geschaffen werden. Dabei wird die gesprochene Sprache durch verschiedene Datenvorverarbeitungsschritte in wichtige Informationen gefiltert. Anschließend können neuronale Netze verwendet werden, um diese Emotionen anhand von Machine Learning zu erlernen und bei erneuter Kommunikation eine Reaktion auf eine Situation zu geben.

3.1.2 Mimik

Durch die Mimik, auch Gesichtsausdrücke genannt, können Emotionen indirekt geäußert werden. Hierbei werden die verändernden Gesichtszüge eines Menschen in einer Situation betrachtet. Eine Situation kann etwas sein, das ein Mensch hört, sieht oder fühlt. Die veränderten Gesichtszüge können die Änderung der Lippen, der Nase, des Mundes oder der Gesichtsmuskeln sein. Diese Veränderung verdeutlichen die Emoti-

onen des Menschen in bestimmten Situationen. In dieser Art von Emotionserkennung werden durch unterschiedliche Schnittstellen die entscheidenden Gesichtspunkte für den Ausdruck von Emotionen durch Video oder Kameraaufnahmen getrackt und kategorisiert [15]. Dabei gibt es unterschiedliche Tools, die angeboten werden, um die Gesichtspunkte und die damit verbundenen Emotionen zu messen. Diese stellen Garcia et al. [15] in ihrer Arbeit vor. Dabei fokussieren sich die meisten Tools auf die von Ekman definierten Emotionen. Lernverfahren zum Klassifizieren der Gesichtspunkte in die Emotionen können hier ebenfalls eingesetzt werden. In der Arbeit von Atanassov et al. [6] hat sich ergeben, dass z. B. der Einsatz von Convolutional Neural Networks zum Erlernen von Gesichtsdaten gute Ergebnisse liefern.

3.1.3 Gestik

Durch die Gestik eines Menschen können wie auch bei der Mimik viele emotionale Informationen erkannt werden. Die Körperhaltung eines Menschen ändert sich bei jeder Emotion. Diese kann z. B. bei der Haltung vom Kopf deutlich werden, wenn er herunterhängt, dass es eine negative Emotion wie Trauer darstellt. In der aktuellen Forschung sind nicht alle Körperteile in ihren Bewegungen in Bezug auf Emotionen erforscht. Dabei gibt es mehr Forschungen zum Thema Emotionserkennung aus Hand- oder Kopfbewegungen als aus Emotionserkennungen vom Torso. Die gängigste Methode, Emotionen aus Körperbewegungen zu gewinnen, ist die Verwendung eines kinematischen Modells am Körper eines Menschen. Ein kinematisches Modell stellt dabei eine Kollektion von angegliederten Gelenken mit festgelegten Freiheitsgraden dar, die analog zum menschlichen Skelett sind. Nachdem die Körperpunkte durch ein kinematisches Modell festgelegt werden, werden Datenvorverarbeitungsschritte durchgeführt, um alle Störfaktoren aus den aufgenommenen Bewegungsdaten zu beseitigen. Diese gefilterten Daten können dann Emotionen zugeordnet werden, die im Vorfeld definiert werden. Auch hier kann die Zuordnung erlernt und durch maschinelle Lernverfahren angewendet werden [23]. So wie die Erkennung der Emotionen aus der Mimik als auch der Gestik wurde in der Arbeit von Atanas V. Atanassov et al. [6] eingeführt. Dabei hat sich ergeben, dass für das Erlernen und Erkennen von Emotionen aus Körpersprache Deconvolutional Neural Networks gute Ergebnisse zeigten. Eine weitere Forschungsarbeit zu dem Thema, wie herangegangen werden kann, ist die Vorgehensweise von Bernhardt und Robinson [7] bei dem verschiedene Formen der Gestik in Klassen unterteilt werden. Diese Klassen wurden an unterschiedliche Emotionen zugewiesen und durch maschinelle Lernverfahren erlernt. Als Klassifikationsmethode für die Einteilung der Körpergesten in Emotionen wurde ein Support-Vector-Machine-basierter Klassifikator verwendet und zuletzt mithilfe einiger Tests evaluiert, die positive Ergebnisse erzielt haben.

3.1.4 Physiologische Körperüberwachung

Physiologische Zustände des Körpers können die Gefühlslage eines Menschen verdeutlichen. Die Änderung des emotionalen Zustandes kann durch deutliche Symptome wie die Änderung der Herzfrequenzaktivität, der Funktion des Blutkreislaufes oder durch Reaktionen auf der Haut erkannt werden. Wenn der Blutkreislauf nicht richtig funktioniert, kann dies z. B. an der Farbe der Haut des Menschen erkannt werden. Stressreaktionen können bspw. mit Symptomen wie Schwitzen oder Zappeln assoziiert werden. In dieser Art von Messung können Emotionen nicht genau definiert werden und sollten anhand von mehrdimensionalen Klassifikationsmethoden kategorisiert werden. Dabei werden die Emotionen als Status einer körperlichen Reaktion wie z. B. Fieber definiert [24][25]. Diese Methode kann oft bei der Diagnose oder Therapie von Erkrankungen behilflich sein, da der physiologische Zustand eines Menschen die Anzeichen für eine Erkrankung verdeutlichen kann.

4 Ergebnisse

In diesem Kapitel werden die Literaturergebnisse aus der systematischen Literaturrecherche vorgestellt. Dabei werden die Inhalte der Paper in einer kompakten Form beschrieben, indem das Thema, die Erkrankung, die Art und Weise der Emotionserkennung- und Kommunikation beim Affective Computing und das Ergebnis des Papers, falls vorhanden, dargelegt werden. Anschließend wird in Tabelle 5 eine Übersicht über die Literaturen und der darin beinhalteten Informationen gegeben.

Die von Aranha et al. [5] vorgestellte Arbeit handelt von einer Studie zur automatischen Anpassung von Spielen an die Anwender:innen bei Rehabilitationsaufgaben nach einer körperlichen Behinderung. Das Ziel ist die Steigerung der Motivation und somit auch der Rehabilitationsergebnisse von Patient:innen. Das verwendete Framework EasyAffecta besteht aus der Emotionserkennung, dem Controller-Modul und dem Anpassungsmodul. Die Emotionserkennung basiert auf einer Gesichtserkennungstechnik mithilfe einer Kamera. Mithilfe der Software Affectiva SDK wird das Kodiersystem FACS eingesetzt, das die relevanten Gesichtspunkte misst und einer Emotion zuordnet. Das Controller-Modul wird eingesetzt, damit die behandelnden Therapeut:innen Gruppen von Anwenderprofilen erstellen kann. So können im nächsten Schritt durch das Adaptionsmodul passende Adaptionsmöglichkeiten für die Anwender:innen ausgewählt werden. Die Evaluation der Ergebnisse der Studie zeigen, dass die Psychotherapeut:innen mit den Adaptionsvorschlägen des Systems übereinstimmten und einen positiven Einfluss auf die Rehabilitation hatten.

In der Arbeit von Siegmund et al. [30] wird eine Designstudie über ein tragbares Gerät namens Talis für die Behandlung von Depressionspatient:innen durchgeführt. Talis ist ein kleines Gerät, das am Hals z. B. als Halskette getragen werden kann. Es funktioniert, in dem es Affective Computing mithilfe von Sprache nutzt. Dabei hört Talis den Nutzer:innen und dessen Umgebung zu und nimmt diese gesammelten Daten auf. Diese wertet er aus und schlägt den Patient:innen einen Audiosound vor, um deren Laune zu erhöhen. Die Nutzer:innen haben auch die Möglichkeit, selbst Audioaufnahmen anzufordern, anzuhören oder aufzunehmen. Talis kann dementsprechende Audios vergeben, die die Laune der Patient:innen erhöhen können. In diesem Gerät wird das Emo-Voice Framework verwendet, dass Emotionen nach ihrer Frequenz, nach ihrer Lautstärke und des Zusammenhangs zwischen der Lautstärke und der Störfaktoren klassifiziert. Für die Klassifikation werden Support Vector Machine (SVM) Klassifikatoren verwendet. Das Gerät wurde in Form von Nutzerstudien getestet, welche ergaben, dass die Form und die Funktion von Talis zur Therapie von Depressionen weiterhelfen könnte.

In der Arbeit von Kim et al. [18] wird das Ziel verfolgt, einen Klassifikator in einen Roboter einzubauen, der Emotionen mithilfe der Sprache erkennt und klassifiziert. Dabei liegt der Fokus auf der therapeutischen Behandlung von Kindern mit Autismus-Spektrum-Störungen, für die eine Therapiemöglichkeit geboten wird. Hier wird ein interaktiver Roboter entwickelt, der aus zwei Systemen besteht, wobei eines zur Verdeutlichung der Emotion des Roboters in Form von Gestik und das andere zur Verdeutlichung von Emotionen mithilfe von Gesichtsausdrücken und der Sprache gegenüber den Kindern, dient. Hierfür wurde der Datensatz IEMOCAP für die Auswahl von Merkmalen zur Erkennung von Emotionen erlernt. Für die Klassifikation der Sprache wurde das openSMILE toolkit verwendet, das auf einem SVM-Klassifikator basiert und die Stimme, Lautstärke und weitere Merkmale der Stimme aussortiert, um diese zu bewerten. Diese Forschung basiert auf einer Entwicklung, die noch nicht getestet wurde, aber das Ziel verfolgt, die Bereitschaft des Kindes korrekt einzuordnen und die Reaktion so zu ändern, dass die Kommunikation zwischen dem Kind und dem Roboter verbessert wird und das Kind mit dem Roboter lernen kann.

Um die richtige schrittweise Atmung zu regulieren und eine gesunde Atemfrequenz zu kontrollieren, haben Zafar et al. [34] in ihrer Arbeit drei Spiele entwickelt, die durch Biofeedback das Erlernen von kontrollierter Atmung ermöglichen sollen. Biofeedback ist dabei die Therapiemethode, bei der die Kontrolle von nicht bewussten Abläufen im Körper erlernt werden soll. Die Messung dieser Körperfunktion wird durch Biofeedbackgeräte, die am Körper tragbar sind, ermöglicht. Das Pacman Zen ist eine modifizierte Form von Pacman, bei der Spieler:innen bei jeder neuen Runde eine Atemfrequenz von 7bpm in 60 Sekunden erreichen müssen. Schaffen sie dies nicht, verlieren sie ein Leben. Bei erfolgreicher Durchführung kriegen sie Belohnungen. Die erweiterte Version von Dodging Stress hat zusätzlich die Funktion, dass sich die Hindernisse im Spiel jedes Mal erhöhen, sobald die Atemfrequenz der Patient:innen nicht zwischen 7 und 14bpm liegt. Im dritten Spiel Chill Out wird ebenfalls mithilfe der Atemfrequenz von 7bpm die Atmung kontrolliert. Die Spiele wurden an einer Stichprobe von Menschen getestet. Die Ergebnisse zeigten, dass sich die Regulation der Atemfrequenz ins Positive geändert hat und durch die kontrollierten Atmungen zu einer besseren Spielleistung führte.

Eine Methode der Messung von Emotionen mithilfe von physiologischen Zuständen stellt das von R. W. Picard et al. [10] eingeführte aSpire dar. aSpire ist ein haptisches Gerät, das am Körper getragen werden kann und die Patient:innen dabei unterstützen soll, eigene Kontrolle über die Atemfrequenz übernehmen zu können. Dabei gibt das Gerät bei uneinheitlicher Atemfrequenz leichte Vibrationen ab, die den Patient:innen ein Feedback darüber geben, dass ihre Atemfrequenz unebenmäßig ist. Die Stärke des fühlbaren Musters können Patient:innen selbst in einer hierfür entwickelten App auf dem Smartphone, dass mit dem aSpire verbunden ist, auswählen. Die Implementierung besteht auf mehreren Simulationsmodulen, die fühlbare Muster in eine Dimension richten. Zusätzlich besteht es aus weiteren Modulen, die den Antrieb und den Druck reglementieren. Die Ergebnisse der Teststudie zeigten, dass aSpire bei der Kontrolle und Reduktion der Atemfrequenz der Menschen unterstützt und das Leistungslevel dadurch aufrechterhält.

In der Arbeit von Doyran et al. [12] wird eine Spieltherapie angewandt, um die Affekte von Kindern mithilfe von Video und Sprache zu erkennen, die Qualität der Spieltherapie zu testen und ggf. zu verbessern. Die Gesichtszüge der Kinder werden durch statische Kameras während der Spieltherapie aufgenommen. Das Tool OpenPose wurde für die Erkennung der Gesichter verwendet. Die Erkennung der Emotionen und Affekte der Kinder wurde durch zwei tiefe neuronale Netze (NN) realisiert. Dabei wurde der AffectNet Datensatz trainiert, um die Mimiken zu erkennen. Mit dem zweiten NN wurde die Konnotation für die Gesichter errechnet, um erkennen zu können, ob das Kind sich wohlfühlt oder ruhig ist. Auch die Körpergestik wurde mithilfe von OpenCV ausgewertet. Die Sprache des Kindes wurde in Text umgewandelt und ausgewertet. Die Auswertung des Textes basiert auf einer Sentimentanalyse, bei der die Stimmung eines Kindes erkannt wird. Hierfür wird das Tool SentiTurkNet verwendet, das die Polaritätswerte ausgerechnet. Das Ergebnis zeigt, dass die Emotionsanalyse mithilfe von Mimikanalysen funktioniert, aber nicht genügt. Das Zusammenspiel mit der Spracherkennung führt zu einem besseren Ergebnis und ist für die Auswertung der Emotionsdaten entscheidend. Durch diese Methode können in Zukunft bei verbesserter ML-Verfahren sowohl eine Diagnose als auch eine Therapie von Autismus erkrankten Kindern ermöglicht werden.

Eine weitere Art der Therapie wird im Artikel von Zygadlo [36] beschrieben. Hier wird ein therapeutischer Chatbot implementiert, der mit den Nutzer:innen textuell kommunizieren kann. Der Chatbot wird hierbei mithilfe von Rasa, einem Framework zur Erstellung von Chatbots, aufgebaut. Innerhalb von Rasa sind vorgefertigte Datenbanken zur natürlichen Verarbeitung von Sprache eingebaut. Da kein Korpus auf polnischer Sprache

existiert, wird ein eigener mit den entsprechenden zugeordneten Emotionen zur Emotionserkennung erstellt. Die Einteilung der Emotionen wird auf Basis des englischen Datensatzes DailyDialog und EmpathicDialog realisiert. Für das Trainieren der Modelle wird in einem Durchlauf der Text mittels der Klassifikationsverfahren Naive Bayes und SVM trainiert. In einem weiteren Durchlauf wird das Prinzip der bigramme zur Kategorisierung der Texte verwendet. Im letzteren wurde der BERT-basierte Klassifikator verwendet, der von den drei vorgestellten Verfahren, die besten Ergebnisse erzielte. Dieses Modell der Klassifikation wird im weiteren Schritt in Rasa integriert, so dass der Chatbot mit den Nutzer:innen kommunizieren kann. Diese Kommunikation wird in einer Gruppe von Testnutzer:innen getestet, um den therapeutischen Ansatz zu ermöglichen.

Ein Chatbot wird auch in der Arbeit von Shah et al. [29] vorgestellt. Dieser Chatbot EDRA soll für Menschen mit mentalen Problemen eine Therapieumgebung aufbauen. Dabei werden die Emotionen mithilfe von Emotionserkennungssystemen aus den Texteingaben erkannt und durch tiefe neuronale Netze eine Antwort produziert. Für die Kommunikation wird zunächst ein Korpus sowohl für die Analyse der Stimmung als auch für die Antworten des Chatbots erstellt. Für die Einteilung der Stimmung aus den Emotionen wird ein SVM Klassifikator eingesetzt. Um die Sprache verarbeiten zu können NLP-Prozesse eingesetzt und diese anschließend mithilfe von DNN, die auf Feed Forward Neural Networks basieren, in Ihre Auswertungen kategorisiert. Je nach Resultat erstellt der Chatbot ein Protokoll, das die mentale Verfassung der Patient:innen zusammenfasst und Empfehlungen ausspricht. Die Ergebnisse haben ergeben, dass die Genauigkeit der richtigen Emotionsbestimmung bei etwa 93 % liegt.

In der Arbeit von Amanda Funabashi et al. [14] wird ein Spielansatz vorgestellt, der eine der Rehabiliationstherapie für Schlaganfallpatient:innen mithilfe von virtueller Realität bieten soll. Die Aufgabe des Spieles ist es, zwei zueinander passende Bilder miteinander zu verbinden, in dem die Bilder zueinander platziert werden. Dabei werden die Emotionen der Patient:innen mithilfe von Affectiva SDK erfasst. Diese erkennt Gesichtsmerkmale und klassifiziert sie mit Machine Learning Verfahren in Emotionen. Die Bewegungen beim Spielen werden mithilfe einer Kinect erfasst. Es wird ein visuelles Feedback gegeben, bei dem angezeigt wird, ob ein Error oder ein Punkt erzielt wurde. Das Ziel hierbei ist die Senkung der Fehler in jedem Spieldurchlauf. Die Evaluation der Sensordaten der Kinect, der Gesichtsdaten und der Spielerfolge von Patient:innen, deren Schlaganfall über 6 Monate her ist, haben ergeben, dass bei verschiedenen Patient:innen andere Ergebnisse resultieren und deshalb keine allgemeine Schlussfolgerung gezogen werden kann, jedoch bei Öfteren Durchläufen ein höheres Engagement erzielt wurde.

Ein Einsatz von physiologischen Köperüberwachungen bei der Therapie von Kindern mit Autismus wird in der Arbeit von Washington et al. [33] vorgestellt. Dabei kommt Google Glass zum Einsatz und wird hier SuperpowerGlass genannt. Diese Brille besitzt eine Kamera, die die Sicht und das Gesicht der Träger:innen aufnimmt, womit die Emotionen der Kinder erkannt werden können. Die Reaktionen auf die Emotionen können sowohl auf dem Heads-Up Display als auch als Audio gehört werden. Die Gesichtserkennung wurde mithilfe einer logistischen Regression realisiert. Mithilfe von unterschiedlichen Aktivitäten wie *Erkenne das Lächeln*, *Errate die Emotionen* oder auch als *unstrukturierte Aktivität*, indem es dem Kind ein Feedback gibt, wenn es eine bestimmte Emotion aufnimmt ohne sich in einem Spiel zu befinden. Die Eltern der Kinder können die Aktivitäten per einer mit der Brille verbundenen App ansehen und erneut starten. Nach einer dreimonatigen Studie hat sich ergeben, dass die tragbaren Geräte wie die Brille bei kontinuierlichem Einsatz helfen können, aber die Effektivität auch von der Schwergrad der Erkrankung abhängt.

In der Arbeit von Biswal [8] wird ein therapeutisches Musikempfehlungssystem vorgestellt, das Lieder bei einer entsprechenden Emotion abspielt. Die Emotionserkennung findet dabei anhand eines intelligenten Algorithmus statt, der die Emotionsveränderungen von Menschen in bestimmten Aktivitäten erkennt, um dementsprechend ein Lied vorzuschlagen. Das Ziel dabei ist die Verbesserung der emotionalen Stabilität und der Leistung in einer bestimmten Situation. Das System wird in Form von Mobilgeräten, Infotainmentsystemen für Fahrzeuge oder tragebaren Geräten zur Verfügung stehen und in Form von der Sprache realisiert. Um diese Musik kategorisieren zu können werden die Metadaten der Lieder in Abschnitte klassifiziert wie Text, Rhythmus usw. Darauffolgend werden diese Metadaten Emotionen zugeordnet, die dann je nach Emotion der Nutzer:innen ein passendes Lied abspielen lassen können. Diese Anwendung kann hierbei zu selbsttherapeutischen Zwecken genutzt werden, um die Stimmung eines bspw. depressiven Menschen zu verbessern. Dieses Modell wird in der Zukunft noch implementiert und getestet werden.

In der Arbeit von Tironi et al. [32] wird eine virtuelle Trainerin Hannah vorgestellt, die die Therapie von Patient:innen mit psychischen Erkrankungen oder auch physischen Erkrankungen, die Patient:innen psychisch belasten, verbessern soll, indem sie eine emotionale Beziehung zu den Patient:innen aufbaut. Dabei werden die Emotionen der Patient:innen mithilfe der Microsoft Kinect und dem Wii Board für die Mimik und Gestik gemessen. Zudem kann auch sprachlich mit Hannah kommuniziert werden. Hierbei werden die vier Emotionen Freude, Trauer, Wut und Überraschung von Paul Ekman für die Reaktion von Hannah ausgewählt. Für die Verarbeitung der Erkennungen werden Text-To-Speech für die Sprache, ein Animator für die Gestik und ein Memory-Speicher, um die Daten den Patient:innen anzupassen, verwendet. Die Verarbeitung und Einteilung erfolgt nach dem Fuzzy-Logic Prinzip. Die Ausgaben des Avatars werden durch Paul Ekmans FACS für die Mimik und durch Mixamo für die Gestik realisiert. Dabei führt Hannah durchgehend durch eine virtuelle Welt, interagiert und führt Spiele durch. Das Spiel wurde mit einer Patientin, die an Multiple Sklerose leidet, getestet. Hannah hat ihr während dem Spiel Hinweise gegeben, Ihre Haltung zu verbessern, was präventiv gegen die Verschlechterung der Krankheit helfen kann. Es hat sich gezeigt, dass sich während dem Spiel die Haltung der Patientin gebessert hat.

In der Arbeit von Akinloye et al. [3] wird ein e-Gesundheits-System für Kinder mit Autismus-Spektrum-Störungen (ASD) in afrikanischen Ländern entwickelt. Dabei wurde hier ein physiologisches Wearable eingesetzt, das elektrodermale Aktivitäts- und Spracherkennungssensoren verwendet, um physiologische Daten von autistischen Kindern zu erhalten. Außerdem wird mit diesem Sensor auch Spracherkennung für die Erkennung der Emotionen eingesetzt. Die eingesetzten Sensoren basieren aus einfachen Arduinosystemen. Aus diesen Sensoren werden Daten analysiert und mithilfe von Fuzzy-Logic in sieben Emotionen klassifiziert. Es hat sich ergeben, dass die Klassifikationsrate sehr genau war. Nach der Erkennung der Emotion wurde auch ein Empfehlungssystem zur Musiktherapie angeboten, welches ebenfalls gute Ergebnisse geliefert hat.

Zur Diagnostik von Erkrankungen mittels Affective Computing haben Muzammel et al. [21] die Erkennung und Bewertung von Depressionsmerkmalen untersucht. Dabei wird zur Erkennung die Sprache als Mittel genutzt. Die Depressionsmerkmale werden aus den sechs akustischen Merkmalen für die Depressionserkennung herausgenommen. Aus diesen Kategorisierungen kann auch erkannt werden, welchen Schweregrad die Depression hat. Die Sprache an sich wird vorverarbeitet und die wichtigen Merkmale aus der Sprache extrahiert. Dazu werden zum einen Vokale und zum anderen Konsonanten bei der Sprache analysiert. Anschließend werden Convolutional Neural Networks eingesetzt, um die Merkmale so zu erlernen,

dass sie im Nachhinein klassifiziert oder bewertet werden können. In dieser Arbeit werden vor allem die Phonem-Einheiten für die Erkennung der Erkrankung mittels Machine Learning untersucht. Hierbei werden drei Arten von neuronalen Netzen zusammengebracht, um eine bessere Performanz zu erreichen. Die durchgeführten Studien zeigen, dass die Merkmale, die auf Konsonanten basieren, erkenntlichere Ergebnisse bei der Emotionserkennung geliefert haben als die Vokalbasierten.

Auch in der Arbeit von He und Cao [16] liegt der Fokus auf der Sprache. Hier werden ebenfalls Informationen aus der Sprache verwendet, um eine Diagnose einer Depression feststellen zu können. Zusätzlich werden hier auch Deep-Learning Algorithmen eingesetzt. Dabei wird eine Methode eingesetzt, um mit Deep Convolutional Neural Networks Sprachinformationen zu erlernen und anschließend anhand von manuellen Extraktionen weitere Informationen zu gewinnen. Diese Erfassungen werden zuletzt kombiniert, um Depressionen besser erkennen zu können. Diese Vorgehensweise wurde auf die zwei Datensätze von AVEC2013/2014 verwendet und haben gezeigt, dass dieser Ansatz performantere Leistungen gezeigt hat als andere Erkennungsansätze für Depressionen.

Im Jahr 2018 [27] und 2019 [28] fand der Audio/Visual Emotion Challenge and Workshop statt, bei dem 2018 die Erkennung bipolarer Störungen und 2019 die Erkennung von Depressionen mittels künstlicher Intelligenz behandelt wurden. Bei der Erkennung von bipolaren Störungen wurde zuerst ein Korpus mit Audioaufnahmen aus Interviews mit Menschen, die an bipolarer Störung leider erstellt. Für die Auswertung wurde zum einen die Datenbank SEWA für spontane Audiodaten und zum anderen RECOLA, in der sich das soziale Verhalten bei kollaborativem Arbeiten befindet. Für die Erkennung der Sprache werden drei KI-Methoden eingesetzt: überwachtes, halbüberwachtes und unüberwachtes Lernen. Die Klassifikation der Daten wird mithilfe eines SVM realisiert, jedoch hat sich ein LSTM-RNN für die Audiodaten performanter gezeigt. Durch diese Klassifikation kann das Vorhandensein einer bipolaren Störung festgestellt werden. Die Ergebnisse im Jahr 2018 zeigen, dass der Einsatz von unüberwachtem Lernen bei der Erkennung der Audiodaten mit 58,2 % die besten Ergebnisse erzielt. 2019 lag der Schwerpunkt auf der Erkennung von Depressionen, bei dem drei Unterwettbewerbe veranstaltet wurden. Einmal SoMS, bei dem die Stimmung aus positiven und negativen persönlichen Geschichten vorhergesagt werden soll. Zum zweiten DDS, bei dem die Depression aus Interviews vorhergesagt werden soll und zum dritten CES, bei dem die affektiven Veränderungen in einem kulturübergreifenden Paradigma abgeleitet werden sollen. Bei der Einhaltung angegebener Anforderungen waren 2019 die Ergebnisse auf einem guten Stand.

In der Arbeit von Qian et al. [26] wurde ebenfalls die Diagnose von Depressionen untersucht. Dabei werden auch hier maschinelle Lernverfahren eingesetzt, um eine automatisierte Erkennung von Depressionen in Menschen ermöglichen zu können. Hier werden mittels eines physiologischen Wearables, eine Uhr, die Aktivitäten von Patient:innen getrackt und auf Depressionsanzeichen zugeordnet. Diese werden mithilfe eines Bag-of-Words Ansatzes und einem Funktionsansatzes extrahiert. Nach der Extraktion der Daten werden diese Sensordaten mittels eines SVM-Klassifikators eingeteilt. Nach Durchführung eines Experimentes haben die Ergebnisse gezeigt, dass bei einer erhöhten Abtastrate der Funktions- und bei keiner erhöhten Abstastrate der BoBW-Ansatz bessere Leistungen liefert.

In der Arbeit von Abdaei und al Osman [1] werden die Zustände, in denen sich Patient:innen mit einer bipolaren Störung in einer Phase befinden, in der sie eine extreme Emotion erleben, gemessen. Diese Zustände lassen sich in Manie, Hypomanie und Rückbildung unterteilen. So kann die Krankheit bestimmt und eine mögliche Therapie hieraus gefunden werden. Dabei werden Emo-

Tabelle 5: Literaturergebnisse

Literatur	Jahr	Art der Emotionserkennung	Einsatz	Erkrankung
[5]	2017	Mimik	Therapie	Körperliche Behinderung
[30]	2017	Physiologische Zustand und Sprache	Therapie	Depression
[18]	2017	Mimik, Sprache, Gestik	Therapie	Autismus-Spektrum Störung
[34]	2019	Physiologischer Zustand	Therapie	Atemprobleme
[10]	2021	Physiologischer Zustand	Therapie	Atemprobleme
[12]	2019	Mimik, Sprache, Gestik	Therapie	Autismus
[36]	2021	Text (Sprache)	Therapie	Psychische Störung
[29]	2021	Text (Sprache)	Therapie	Psychische Krankheit
[14]	2017	Physiologischer Zustand und Mimik	Therapie	Schlaganfall
[33]	2017	Mimik	Therapie	Autismus
[8]	2021	Sprache	Therapie	Psychische Gesundheit
[32]	2019	Gestik, Mimik und Sprache	Therapie	Psychische Krankheit, Multiple Sklerose
[3]	2020	Physiologischer Zustand und Sprache	Therapie	Autismus-Spektrum Störung
[21]	2020	Sprache	Diagnose	Depression
[16]	2018	Sprache	Diagnose	Depression
[27]	2018	Sprache	Diagnose	Bipolare Störung
[28]	2019	Mimik und Sprache	Diagnose	Depression
[26]	2019	Physiologischer Zustand	Diagnose	Depression
[1]	2020	Mimik	Diagnose	Bipolare Störung
[11]	2018	Mimik	Diagnose	Depression
[35]	2020	Physiologischer Zustand	Diagnose	Depression

tionen aus Gesichtsmerkmalen herausgefiltert und in ein Convolutional Neural Network getan. Die Gesichter werden durch Videoaufnahmen aufgenommen und zur Weiterverarbeitung korrigiert. Dann werden die Patient:innen in Ihre Zustände mithilfe von LSTM Klassifikatoren eingeteilt. Die Methode wurde mit existierenden Modellen verglichen und es hat sich ergeben, dass das vorgestellte Modell eine bessere Genauigkeitsrate von 63,32% als die anderen Modelle hat.

Die Arbeit von Cholet und Paugam-Moisy [11] basiert auf der Diagnose von depressiven Stimmungen und dessen automatische Klassifikation. Hierfür wird ein Tool vorgestellt, dass depressive Stimmungen durch Gesichtsmerkmale und Landmarken auf dem Gesicht, erkennt. Mit diesen Gesichtsinformationen können Emotionen herausgelesen werden. Die Gesichter werden anhand von Videos aufgenommen. Für die Einteilung der Emotionen wird die AVEC2014 Datenbank verwendet, bei der depressive Gesichter definiert wurden. Für die Klassifikation der Stimmungen wurde ein inkrementeller Klassifikator verwendet. In einer Studie mit depressiven und gesunden Menschen lag der Erfolg der Einteilungsrate bei 90,73%. Zwischen den Klassifikatoren hat sich ergeben, dass mit 94,87% der Random Forest Algorithmus zur Bestimmung vorne liegt.

Eine weitere Untersuchung zur Erkennung von Depressionen stellt Zhu et al. [35] mit einem verbesserten Klassifikationsmodell mittels EEG und Eye-Tracking Daten dar. Mit Mithilfe von EEG können Funktion im Hirn gemessen und interpretiert werden. Dabei werden die Daten depressiver Patient:innen mithilfe neuronaler Netze eingeteilt und erreichten eine Genauigkeit von 88,6 %. Durch die Messung konnten depressive Hinweise gemessen werden. Mithilfe vom Eye-Tracking werden Sichtdaten bei bestimmten Aufgaben gemessen. Sowohl die Eye-Tracking- als auch die EEG-Daten haben nach einem durchgeführten Experiment ergeben, dass das verbesserte Klassifikationsmodel CBEM bessere Ergebnisse erzielt hat als die traditionellen Klassifikationsmethoden.

5 Diskussion und Fazit

Die Ergebnisse des systematischen Literaturreviews in diesem Artikels bestehen zu 61,9 % aus Literaturen, die eine mögliche Therapie von Erkrankungen mittels Affective Computing beschreiben und zu 38,1 % aus Literaturen, die eine Möglichkeit zur Diagnose von Erkrankungen bieten. An den Ergebnissen ist zu sehen, dass es einige Forschungen gibt, die sich mit dem Einsatz von therapeutischen oder diagnostischen Zielen im Zusammenhang mit Affective Computing beschäftigen. Hierbei werden unterschiedliche Mittel zur Emotionserkennung verwendet, wie die Sprache bei Talis, dem tragbaren Gerät für die Erkennung und Reaktion auf menschliche Emotionen oder auch die Roboter, die durch ein mobiles Gerät die Sprache der Kinder verarbeiten und dementsprechend reagieren. Auch für die Autismus-Spieltherapie wurden Mittel zur Spracherkennung verwendet. Die Erkennung der Mimik wurde bei der Rehabilitationsverbesserung wie auch für die Autismus-Therapie sowohl bei den Robotern als auch bei der Spieltherapie eingesetzt. Zudem hat die Spieltherapie auch die Gestik betrachtet. Die Erkennung der Emotionen aus physiologischen Zuständen können auch gemessen werden, wie z. B. beim Biofeedback mit drei Spielen, die zur Regulation der Atemfrequenzen eingesetzt wurde und beim aSpire, welches ein Feedback gibt, wenn der Körper eine unregelmäßige Atemfrequenz aufweist. Hierbei kann für die Forschungsfragen zusammengefasst werden, dass die Diagnose oder Therapie von verschiedenen Erkrankungen mittels Affective Computing mithilfe der Emotionserkennung und -reaktion in der Sprache, der Mimik, der Gestik oder durch physiologische Zustände erfolgen kann. Dabei können diese Mittel für unterschiedlichen Anwendungsfälle je nach Erkrankung eingesetzt werden. Aus den in Kapitel 4 vorgestellten Literaturen wird erkannt, dass die Sprache und die Mimik ein häufig eingesetztes Mittel beim Affective Computing für die Therapie oder Diagnose von Erkrankungen

ist. Dabei wird die Sprache in dieser Recherche in 57,1 % und die Mimik in 42,9 % der Fälle eingesetzt. Die physiologischen Zustände bieten auch viel Potenzial durch die Körperveränderungen bei veränderten Emotionen und werden in 33,3% der Papers eingesetzt, jedoch sollten weitere Entwicklungen in dem Bereich durchgeführt werden, da diese vor allem bei der Erkennung von Erkrankungen eine aufschlussreiche Stütze bieten können. Für Affective Computing bei Erkrankungen wird die Gestik derzeit nicht so häufig wie die anderen Mittel eingesetzt, da diese viel komplexere Analysen benötigt [23]. In dieser Arbeit wird die Gestik in nur 14,3 % der Paper eingesetzt. Zusätzlich wird erkenntlich, dass Affective Computing oft bei psychischen Erkrankungen oder Verhaltensstörungen eingesetzt wird und weniger bei körperlichen Erkrankungen. Innerhalb der Literaturen sind allein 52,4 % für psychische Erkrankungen und 28,6% für Verhaltensstörungen, was eine große Menge der Literaturen einnimmt. Körperliche Erkrankungen oder Behinderungen werden in 23,8 % der Literaturen behandelt. Dies kann daran liegen, dass psychische Erkrankungen oder Verhaltensstörungen eher mit den Emotionen des Menschen korrelieren. Bei der Recherche wurde zusätzlich erkenntlich, dass es mehr Artikel zur Stressreduktion oder zur Verbesserung des Wohlbefindens gab als für die Diagnostik oder Therapie von Erkrankungen. Die Evaluationen der Forschungen haben ergeben, dass der Einsatz von Affective Computing für diese Anwendungsfälle positive Ergebnisse liefern. Dabei wurde in z. B. Doyran et al. [12] und vielen weiteren Paper erwähnt, dass die Forschung der künstlichen Intelligenz (KI) weiter voranschreiten muss, um solche Systeme endgültig einsetzen zu können. Außerdem ist ersichtlich, dass es sich meist nur um Forschungsarbeiten handelt, die sich derzeit nicht im medizinischen Umfeld reguliert einsetzen lassen. Dies kann an der unzureichenden Forschung liegen und muss weiter ausgebaut werden, um dem medizinischen Umfeld mehr Sicherheit bieten zu können.

Literaturverzeichnis

[1] Abaei, N. and Osman, H.A. 2020. A Hybrid Model for Bipolar Disorder Classification from Visual Information. *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (Barcelona, Spain, May 2020), 4107–4111. DOI: 10.1109/ICASSP40776.2020.9054648

[2] Ahmad, K. ed. 2011. *Affective Computing and Sentiment Analysis*. Springer Netherlands. DOI: 10.1007/978-94-007-1757-2

[3] Akinloye, F.O. et al. 2020. Development of an affective-based e-healthcare system for autistic children. *Scientific African*. 9, (Sep. 2020), e00514. DOI: 10.1016/j.sciaf.2020.e00514

[4] Alarcao, S.M. 2017. Reminiscence therapy improvement using emotional information. 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII) (San Antonio, TX, Oct. 2017), 561–565. DOI: 10.1109/ACII.2017.8273655

[5] Aranha, R.V. et al. 2017. Using Affective Computing to Automatically Adapt Serious Games for Rehabilitation. *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)* (Thessaloniki, Jun. 2017), 55–60. DOI: 10.1109/CBMS.2017.89

[6] Atanassov, A.V. et al. 2021. Hybrid System for Emotion Recognition Based on Facial Expressions and Body Gesture Recognition. *2021 International Conference Automatics and Informatics (ICAI)* (Varna, Bulgaria, Sep. 2021), 135–140. DOI: 10.1109/ICAI52893.2021.9639829

[7] Bernhardt, D. and Robinson, P. 2007. Detecting emotions from everyday body movements. (Cambridge, 2007).

[8] Biswal, S. 2021. I Can Feel You: A Self-Trained Therapeutic Music Recommender System. *2021 11th International Conference on Advanced Computer Information Technologies (ACIT)* (Deggendorf, Germany, Sep. 2021), 753–756. DOI: 10.1109/ACIT52158.2021.9548471

[9] Calvo, R. et al. 2015. Cyberpsychology and Affective Computing. *The Oxford Handbook of Affective Computing*. R. Calvo et al., eds. Oxford University Press. DOI: 10.1093/oxfordhb/9780199942237.013.017

[10] Choi, K.Y. et al. 2021. aSpire: Clippable, Mobile Pneumatic-Haptic Device for Breathing Rate Regulation via Personalizable Tactile Feedback. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama Japan, May 2021), 1–8. DOI: 10.1145/3411763.3451602

[11] Cholet, S. and Paugam-Moisy, H. 2018. Prototype-based Classifier for Automatic Diagnosis of Depressive Mood. *2018 International Conference on Biomedical Engineering and Applications (ICBEA)* (Funchal, Jul. 2018), 1–6. DOI: 10.1109/ICBEA.2018.8471736

[12] Doyran, M. et al. 2019. Video and Text-Based Affect Analysis of Children in Play Therapy. *2019 International Conference on Multimodal Interaction* (Suzhou China, Oct. 2019), 26–34. DOI: 10.1145/3340555.3353760

[13] Ekman, P. 1994. *The nature of emotion: Fundamental questions.* Oxford University Press.

[14] Funabashi, A.M.M. et al. 2017. AGaR: A VR Serious Game to Support the Recovery of Post-Stroke Patients. *2017 19th Symposium on Virtual and Augmented Reality (SVR)* (Curitiba, Nov. 2017), 279–288. DOI: 10.1109/SVR.2017.15

[15] Garcia-Garcia, J.M. et al. 2017. Emotion detection: a technology review. *Proceedings of the XVIII International Conference on Human Computer Interaction* (Cancun Mexico, Sep. 2017), 1–8. DOI: 10.1145/3123818.3123852

[16] He, L. and Cao, C. 2018. Automated depression analysis using convolutional neural networks from speech. *Journal of Biomedical Informatics*. 83, (Jul. 2018), 103–111. DOI: 10.1016/j.jbi.2018.05.007

[17] Hervas, R. et al. 2017. A Learning System to Support Social and Empathy Disorders Diagnosis through Affective Avatars. *15th International Conference on Ubiquitous Computing and Communications and 2017 International Symposium on Cyberspace and Security (IUCC-CSS)* (Granada, Dec. 2017), 93–100. DOI: 10.1109/IUCC-CSS.2016.021

[18] Kim, J.C. et al. 2017. Audio-based emotion estimation for interactive robotic therapy for children with autism spectrum disorder. *2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)* (Jeju, Jun. 2017), 39–44. DOI: 10.1109/URAI.2017.7992881

[19] Kitchenham, B. 2007. Guidelines for performing Systematic Literature Reviews in Software Engineering.

[20] Lopez, D.M. and Picard, R. 2016. Wearable Technologies for Multiple Sclerosis: The Future Role Wearable Stress Measurement in Improving Quality of Life. *In the proceedings of the Second International Conference on Smart Portable, Wearable, Implantable and Disability-Oriented Devices and Systems (SPWID'16)* (May 2016).

[21] Muzammel, M. et al. 2020. AudVowelConsNet: A phoneme-level based deep CNN architecture for clinical depression diagnosis. *Machine Learning with Applications*. 2, (Dec. 2020), 100005. DOI: 10.1016/j.mlwa.2020.100005

[22] Neiva, F.W. and de Souza da Silva, R.L. 2016. Systematic Literature Review in Computer Science - A Practical Guide. Federal University of Juiz de Fora Technical Report of Computer Science Department. DOI: 10.13140/RG.2.2.35453.87524

[23] Noroozi, F. et al. 2021. Survey on Emotional Body Gesture Recognition. *IEEE Transactions on Affective Computing*. 12, 2 (Apr. 2021), 505–523. DOI: 10.1109/TAFFC.2018.2874986

[24] Picard, R.W. 2009. Future affective technology for autism and emotion communication. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 364, 1535 (Dec. 2009), 3575–3584. DOI: 10.1098/rstb.2009.0143

[25] Qi, H. et al. 2017. Video-based human heart rate measurement using joint blind source separation. *Biomedical Signal Processing and Control*. 31, (Jan. 2017), 309–320. DOI: 10.1016/j.bspc.2016.08.020

[26] Qian, K. et al. 2019. Automatic Detection of Major Depressive Disorder via a Bag-of-Behaviour-Words Approach. *Proceedings of the Third International Symposium on Image Computing and Digital Medicine - ISICDM 2019* (Xi'an, China, 2019), 71–75. DOI: 10.1145/3364836.3364851

[27] Ringeval, F. et al. 2018. AVEC 2018 Workshop and Challenge: Bipolar Disorder and Cross-Cultural Affect Recognition. *Proceedings of the 2018 on Audio/Visual Emotion Challenge and Workshop* (Seoul Republic of Korea, Oct. 2018), 3–13. DOI: 10.1145/3266302.3266316

[28] Ringeval, F. et al. 2019. AVEC 2019 Workshop and Challenge: State-of-Mind, Detecting Depression with AI, and Cross-Cultural Affect Recognition. *Proceedings of the 9th International on Audio/Visual Emotion Challenge and Workshop - AVEC '19* (Nice, France, 2019), 3–12. DOI: 10.1145/3347320.3357688

[29] Shah, H. et al. 2021. EDRA – An Emotional Health Detection And Recognition Assistant. *2021 Asian Conference on Innovation in Technology (ASIANCON)* (PUNE, India, Aug. 2021), 1–6. DOI: 10.1109/ASIANCON51346.2021.9544973

[30] Siegmund, D. et al. 2017. Talis — A Design Study for a Wearable Device to Assist People with Depression. *2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)* (Turin, Italy, Jul. 2017), 543–548. DOI: 10.1109/COMPSAC.2017.228

[31] Sloman, A. 1999. Review of Affective Computing. *AI Magazine*. 20, 1 (1999), 127–133. DOI: 10.1609/aimag.v20i1.1448

[32] Tironi, A. et al. 2019. An Empathic Virtual Caregiver for assistance in exergame-based rehabilitation therapies. *2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH)* (Kyoto, Japan, Aug. 2019), 1–6. DOI: 10.1109/SeGAH.2019.8882477

[33] Washington, P. et al. 2017. SuperpowerGlass: A Wearable Aid for the At-Home Therapy of Children with Autism. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 1, 3 (Sep. 2017), 1–22. DOI: 10.1145/3130977

[34] Zafar, M.A. et al. 2020. Gaming Away Stress: Using Biofeedback Games to Learn Paced Breathing. *IEEE Transactions on Affective Computing*. 11, 3 (Jul. 2020), 519–531. DOI: 10.1109/TAFFC.2018.2816945

[35] Zhu, J. et al. 2020. An Improved Classification Model for Depression Detection Using EEG and Eye Tracking Data. *IEEE Transactions on NanoBioscience*. 19, 3 (Jul. 2020), 527–537. DOI: 10.1109/TNB.2020.2990690

[36] Zygadlo, A. 2021. A Therapeutic Dialogue Agent for Polish Language. *2021 9th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW) (*Oct. 2021), 1-5. DOI: 10.1109/ACIIW52867.2021.9666281

©2022 <Irem><Uslu>. 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/

Sportdatenvisualisierung mittels Daten des Fußballsimulationsspiels FIFA

Demis Moysidis
Hochschule Reutlingen
Reutlingen, Deutschland
Demis.Moysidis@Student.Reutlingen-University.de

Abstract

Eines der beliebtesten Spiele im elektronischen Sport ist das Spiel FIFA, welches sich zur Aufgabe gemacht hat den Fußball zu simulieren. Die Fähigkeiten der realen Fußballprofis werden anhand von definierten Kriterien eingeschätzt und somit im Spiel simuliert. Die Fähigkeitsbewertung unterscheidet sich jährlich und ist an die reale Leistung des Spielers orientiert. In diesem Paper wird das selbstentwickelte Spieleranalyse-Tool Tiko vorgestellt, welches die Bewertungsdaten der FIFA-Spieler von FIFA 15 bis FIFA 22 zur genaueren Betrachtung bereitstellt. Die verwendeten Visualisierungstechniken sind aus einer Literaturrecherche extrahiert. Der erste Prototyp von Tiko konnte erfolgreich als Web-Anwendung umgesetzt werden. Die Fähigkeitswerte werden anhand Netzdiagrammen, modifizierten Kreisdiagrammen und Balkendiagrammen dargestellt. Der Entwicklungsverlauf der Gesamtbewertung der Spieler wird anhand eines Liniendiagramms dargestellt.

Betreuer Hochschule: Prof. Dr. Uwe Kloos
Hochschule Reutlingen
Uwe.Kloos@Reutlingen-University.de
Informatics Inside Frühjahr 2022
11. Mai 2022, Hochschule Reutlingen
Copyright 2022 Demis Moysidis

CCS Concepts

•Human-centered computing~Visualization~Visualization techniques

•Human-centered computing~Visualization~Visualization systems and tools

Keywords

Datenvisualisierung (engl. *data visualization*), FIFA-Spielerdaten (engl. *FIFA player data*), Analyseanwendung (engl. *analyze tool*)

1 Einleitung

Heutzutage spielen virtuelle Spiele für die junge Generation eine wichtige Rolle, die schon sehr früh in diese Welt eintaucht, indem sie auf Streaming-Websites wie Twitch spielt oder zuschaut [6]. Das Konzept des elektronischen Sports (eSports) wächst jedes Jahr als ein weltweites Phänomen, bei dem professionelle Spieler sowie auch Gelegenheitsspieler, in der Branche auch bezeichnet als *casual gamer*, mit Hilfe von elektronischen Geräten, d. h. Computern, Konsolen und mobilen Endgeräten, an Turnieren teilnehmen [9]. Der ständige Zuwachs an Spielern und Zuschauern trägt zur Popularität von eSports bei. Dabei ist FIFA eines der beliebtesten Spiele. Es fokussiert sich auf die Simulierung des Fußballs. Jährlich erscheint ein neuer FIFA-Teil, in dem die fußballerischen Fähigkeiten der realen Spieler anhand ihrer Leistung im vergangenen Jahr bewertet werden.

1.1 Motivation und Zielsetzung

Im professionellen eSports-Kontext ist es wie im traditionellen Sport von entscheidender Bedeutung, dass Trainer und Analysten die Entwicklung von Spielern verstehen, da dies zur Erleichterung des Individualtrainings und zur Verbesserung der Fähigkeiten des Spielers genutzt werden kann [12]. Der Einsatz von Visualisierungstechniken zur Analyse von Spieler- und Mannschaftsleistungen anhand verschiedener Metriken, wie z.B. erzielte Punkte oder Anzahl erfolgreicher Pässe, wurde bereits in vielen Sportarten (z.B. Fußball oder Basketball) eingeführt und häufig von Profisportlern und Sportinstitutionen genutzt [5]. Eine relevante Forschungsherausforderung besteht darin, Visualisierungstechniken zu erforschen, um Wissen aus diesen Datensätzen zu extrahieren, das Spielern und Trainern bei ihren Analysezielen und Analyseaufgaben hilft [7].

Diese Arbeit zielt darauf ab, diese Herausforderung anzugehen, indem Visualisierungstechniken zur Untersuchung der Spielerfähigkeiten von Fußballspielern vorgeschlagen werden. Als Datenbasis wird ein Datensatz mit allen Bewertungen der FIFA-Spieler von FIFA 15 bis FIFA 22 genutzt. FIFA zeichnet sich dadurch aus, dass es die Fähigkeiten der realen Spieler simuliert und somit eine möglichst nachvollziehbare Bewertung der Leistungsfähigkeit der Spieler bereitstellt. Es werden Fähigkeiten wie z.B. die Sprintgeschwindigkeit oder die Passqualität bewertet. Die Bewertungsskala liegt zwischen 0 und 99. Je höher der Wert, desto besser beherrscht der Spieler die Fähigkeit. Dabei unterscheiden sich die Schwerpunkte je nach Position des Spielers. Das heißt ein Abwehrspieler hat einen besseren Defensiv-Wert als ein Stürmer, dafür hat aber ein Stürmer einen besseren Schuss als ein Verteidiger. Feldspieler und Torhüter werden anhand unterschiedlicher Fähigkeiten bewertet. Bei Feldspielern werden die Fähigkeiten Tempo (engl. *pace*), Schießen (engl. *shooting*), Passen (engl. *passing*), Dribbling (engl. *dribbling*), Defensive (engl. *defending*) und Physis (engl. *physical*), welche wiederrum mittels weiterer untergeordneter Fähigkeiten bestimmt werden (siehe Kapitel 4.1), abgebildet. Bei Torhütern werden die Fähigkeiten Tempo (engl. *speed*), Positionsspiel (engl. *positioning*), Hechten (engl. *diving*), Fangsicherheit (engl. *handling*), Abschlag (engl. *kicking*) und Reflexe (engl. *reflexes*) abgebildet.

Um dieses Ziel zu erreichen, wird das Spieleranalyse-Tool Tiko entwickelt. Es soll ein erster Prototyp von Tiko entwickelt werden, welcher Visualisierungstechniken für die Analyse eines einzelnen Spielers besitzt. Der Name Tiko ist selbst festgelegt und wurde gewählt, weil er kurz und leicht einprägsam ist.

1.2 Struktur der Arbeit

Diese Arbeit ist wie folgt aufgebaut. Kapitel 2 stellt die in dieser Arbeit angewendete Methodik näher vor. In Kapitel 3 werden verwandte Arbeiten im Bereich der Visualisierung von Fähigkeitsdaten von Fußballspielern sowie der Sportdatenvisualisierung vorgestellt. In Kapitel 4 wird das entwickelte Spieleranalyse-Tool Tiko vorgestellt. Erst wird der verwendete Datensatz näher erläutert und anschließend wird auf die Struktur und die eingesetzten Visualisierungstechniken eingegangen. In Kapitel 5 wird ein Ausblick für zukünftige Arbeiten gegeben. Abschließen wird in Kapitel 6 der gesamte Inhalt dieser Arbeit zusammengefasst.

2 Methodik

Die Vorgehensweise zur Umsetzung des Projekts setzt sich aus zwei Tätigkeiten zusammen. Zum einen wurde eine Literaturrecherche durchgeführt, um entsprechende Visualisierungen für das zu entwickelnde Tool ausfindig zu machen und zum anderen fand die Implementierung des ersten Prototyps von Tiko statt. Weiterführung werden diese zwei Tätigkeiten näher ausgeführt.

2.1 Literaturrecherche

Die Literaturrecherche dient als Informationsquelle, um geeignete Visualisierungstechniken für die zu entwickelnde Anwendung zu bestimmen. Dabei wurden die Fachdatenbanken *IEEE*[1], *ACM*[2], *Science Direct*[3] und *Web of Science*[4] nach geeigneter Literatur untersucht. Gesucht wurden Paper die sich mit der Visualisierung von Sport- bzw. Fähigkeitsdaten auseinandersetzen. Der Veröffentlichungszeitraum der Fachartikel wurde auf den Zeitraum von 2015 bis 2022 eingeschränkt. Der erste Suchterm (Suchterm 1) wurde wie folgt definiert:

visualization* AND data AND (sport OR soccer)

Der erste und zweite Teil von Suchterm 1(visualization* AND data) beschränkt sich auf die Übereinstimmung mit dem Titel des Fachartikels. Der letzte Teil beschränkt sich auf eine Übereinstimmung im Gesamten Text des Fachartikels. Der zweite Suchterm (Suchterm 2) wurde wie folgt definiert:

visual* AND analytic* AND (game OR soccer OR player OR sport)

Der gesamte Suchterm 2 beschränkt sich auf die Übereinstimmung mit dem Titel des Fachartikels. Die in Suchterm 1 und Suchterm 2 verwendeten Schlagwörter wurden so gewählt, dass zum einen thematisch passende Fachartikel gefunden werden und zum anderen die Ergebnismenge der Fachartikel eingegrenzt wird. In Tabelle 1 sind die Ergebnismengen abhängig der Suchterme dargestellt.

Aus den Ergebnissen der Fachdatenbanken ACM und ScienceDirect konnten keine für diese Arbeit brauchbaren Artikel entnommen werden. Zudem blieb der Zugriff auf die meisten Fachartikel über Web of Science verwehrt.

Tabelle 1: Ergebnismengen der Suchterme

Datenbank	Ergebnismenge	
	Suchterm 1	*Suchterm 2*
IEEE	10	14
ACM	4	11
ScienceDirect	3	2
Web of Science	37	23

Neben den Fachartikeln dienen die in Kapitel 3 genannten verwandten Anwendungen als Quelle. Sie beinhalten teilweise Visualisierungstechniken, die in ähnlicher Form in Tiko umgesetzt werden können. In Kapitel 4.3 werden die aus den genutzten Fachartikeln und verwandet Anwendungen extrahierten Visualisierungstechniken vorgestellt.

2.2 Technische Umsetzung von Tiko

Anhand den ermittelten Visualisierungstechniken wurde anschließend das Spieleranalyse-Tool Tiko implementiert.

Der Datensatz mit allen Informationen und Fähigkeitsbewertungen der FIFA-Spieler ist aus kaggle[5] entnommen. Im ersten Schritt wird mittels dem Programm XAMP[6] eine MySQL Datenbank aufgesetzt. XAMP ist ein Programmpaket, dass über einen Klick das aufsetzten eines Apache Webservers verbunden mit einer MySQL Datenbank ermöglicht. Für jede FIFA-Serie wird eine separate Tabelle angefertigt. Über die insert-Funktion des Admin-Bereichs werden die Daten in die

[1] https://ieeexplore.ieee.org/

[2] https://dl.acm.org/

[3] https://www.sciencedirect.com/

[4] https://www.webofscience.com/

[5] https://www.kaggle.com/

[6] https://www.apachefriends.org/de/index.html

jeweilige Tabelle vollautomatisch eingepflegt. Lediglich zu Beginn musste die Form der Tabelle mit den entsprechenden Attributbezeichnungen angelegt werden. Da der Datensatz gut aufbereitet ist, musste keine eigenständige Datenvorverarbeitung am Datensatz durchgeführt werden.

Die Visualisierung der Daten wird in Form einer Web-Anwendung realisiert. Mittels JavaScript (Front- und Back-End), HTML, und CSS (Front-End) sowie den Frameworks Chart.js[7], Require.js[8] und JQuery.js[9] wird das Projekt implementiert. Chart.js dient zur Generierung der Graphen wie z.B. den Balkendiagrammen und Netzdiagrammen. Für das Hosting der Webseiten sowie die Übermittlung der Daten ist ein Node.js[10] Server (Back-End) zuständig. Die Datenübermittlung findet über eine im Node.js Server implementierte API-Schnittstelle statt. Beides läuft über denselben Port.

3 Verwandte Arbeiten

Die Herausforderungen, die mit der Datenanalyse von Sportdaten bzw. Fähigkeiten von mehreren Einzelspielern verbunden sind, liegen in der geeigneten Darstellung und dem Umgang der Datenmenge [3]. Aus diesen Gründen werden in den folgenden Abschnitten einige relevante Arbeiten in den Bereichen Datenvisualisierung und Visualisierung der Fähigkeitsdaten von FIFA-Spielern beschrieben.

Die Videospiel- sowie die Sportindustrie setzten schon seit langem Visualisierungstechniken ein [1][2][8]. Mehrere Spieleentwickler und unabhängige Gruppen bieten bereits Visualisierungstools und APIs an, die den Zugriff auf Videospieldaten ermöglichen. Zu den gängigen Tools, die zur Analyse der FIFA-Spieler Fähigkeitswerte verwendet werden, gehören Anwendungen von Drittanbietern wie FutBin[11], Sofifa[12] oder FutHead[13]. Diese Tools liefern allgemeine Spielerdaten sowie die Fähigkeitsbewertungen, die sich auf die Leistung der Spieler im Spiel beziehen. Die Fähigkeitsdaten werden nur teilweise mittels Visualisierungen dargestellt. An vielen Stellen werden nur die Zahlenwerte angezeigt, wodurch Unterschiede in den einzelnen Fähigkeitswerten nicht auf den ersten Blick wahrgenommen werden können und somit eine nähere Betrachtung nötig ist. FutBin und FutHead beinhalten neben den Standardfähigkeitswerten, die den Datensatz dieser Arbeit darstellen, auch noch die Fähigkeitsdaten von Spezialkarten der Spieler, die in einem separaten FIFA Spielmodus genutzt werden, der sich Ultimate Team nennt. Sofifa beinhaltet die Fähigkeitsdaten eines weiteren FIFA Spielmodus, in dem die Fähigkeitswerte für einen Onlinespielmodus wöchentlich aktualisiert werde.

Diese speziellen Anpassungen der Fähigkeitswerte der Spieler werden in diesem Paper nicht berücksichtigt. Der Datensatz in dieser Arbeit beinhaltet lediglich die erste Momentaufnahme der Fähigkeitswerte jeder FIFA-Serie. Zudem mangelt es diesen Tools in der Regel an Visualisierungen von Übersichten, welche die Entwicklung der Spieler über einen längeren Zeitraum bzw. über mehrere FIFA-Teile hinweg, wiedergeben. Außerdem mangelt es diesen Tools an geeigneten Filterfunktionen, um entsprechende Spieler auch aus vergangenen FIFA-Teilen genauer untersuchen zu können. Diese Arbeit zielt darauf ab, einige dieser Herausforderungen anzugehen, indem geeignete Visualisierungen in den ersten Prototypen von

[7] https://www.chartjs.org/

[8] https://requirejs.org/

[9] https://jquery.com/

[10] https://nodejs.org/

[11] https://www.futbin.com/

[12] https://sofifa.com/

[13] https://www.futhead.com/

Tiko implementiert werden, um eine Analyse der Entwicklung der Spieler über einen längeren Zeitraum zu ermöglichen.

Krishnamurthy und Nanda stellen in [8] eine Anwendung vor, die sich ebenfalls mit der Visualisierung von Fähigkeitswerten von Fußballspielern auseinandersetzt. Laut der Beschreibung des Datensatzes nutzt ihre Anwendung denselben Datensatz wie das hier entwickelte Spieleranalyse-Tool Tiko. Der Datensatz wurde ebenso aus kaggle entnommen, weist eine Historie von acht Jahren nach und besitzt die gleichen Fähigkeitsbeschreibungen mit derselben Bewertungsskala von 0 bis 99 hat aber die Bezeichnung *European Soccer Database*. Daraus lässt sich schließen, dass in ihrem Datensatz nur Spieler aus Europa aufgelistet sind, während der Datensatz in diesem Paper weltweite Spielerdaten enthält.

Du und Yuan stellen in [4] eine Taxonomie der Sportdatenvisualisierung auf und fassen dessen Stand der Forschung zusammen. Darüber hinaus werden Visualisierungstechniken für die Sportdaten auf der Grundlage von Datenmerkmalen in fünf Kategorien vorgeschlagen. Es wird unter den Kategorien *hochdimensionaler Datenvisualisierung*, *Zeitreihenvisualisierung*, *Graphen- bzw. Netzwerkvisualisierung*, *glyphenbasierter Visualisierung* und *anderen Visualisierungen* klassifiziert. Ziel des Papers ist es geeignete Visualisierungstechniken für verschiedene Datentypen und unterschiedliche Aufgaben zu empfehlen. Das Paper soll zudem auch als Richtlinie für zukünftige Forschungsarbeiten im Bereich der Sportdatenvisualisierung dienen.

4 Spieleranalyse-Tool Tiko

Dieser Abschnitt beschreibt Tiko, das selbstentwickelte Visualisierungstool zur Analyse der Fähigkeitsdaten der Fußballspieler. Es werden die Fähigkeitsdaten über die acht vergangenen FIFA-Teile von FIFA 15 bis FIFA 22 bereitgestellt. Zu Beginn wird der hier genutzte Datensatz etwas genauer erläutert. Anschließend wird der erste Prototyp von Tiko vorgestellt. Dabei wird auf den Aufbau und die Funktionen von Tiko näher eingegangen. Abschließend werden die umgesetzten Visualisierungstechniken in Tiko näher erläutert.

4.1 Datensatz

Bei dem FIFA-Datensatz handelt es sich um einen multivariaten Datensatz mit 110 Attributen. Er enthält personenbezogene Informationen wie z.B. den Namen, das Alter, die Größe der Spieler, vereinsbezoge Informationen wie z.B. den Club-Namen, die Nation des Clubs und die Fähigkeitswerte der Spieler wie z.B. die Springgeschwindigkeit oder die Passqualität. Die Bewertung eines Feldspielers wird anhand den Fähigkeitskategorien (Oberkategorie) *pace*, *shooting*, *passing*, *dribbling*, *defending* und *physical* bestimmt. Jeder Oberkategorie sind weitere Fähigkeitskategorien (Unterkategorie) untergeordnet, aus denen wiederum die Bewertung der Oberkategorie bestimmt wird. Der Oberkategorie *pace* sind z.B. die Unterkategorien *acceleration* und *sprint speed* untergeordnet. Der maximale Wert den Fähigkeiten annehmen können liegt bei 99. In Tabelle 2 sind Ober- und Unterkategorien der Fähigkeiten von Feldspielern dargestellt. Torhüter besitzen die Fähigkeitskategorien (Oberkategorie) *diving*, *handling*, *kicking*, *reflexes*, *speed* und *positioning*. Diese werden aber nicht anhand weiterer Unterkategorien beschrieben, bis auf *speed*. Diese Oberkategorie besitzt die Unterkategorien *acceleration* und *sprint speed*. Zudem wird die Qualität eines Spielers anhand einer Gesamtbewertung wiedergegeben. Diese kann ebenfalls maximal den Wert 99 annehmen und gibt unabhängig der Betrachtung einzelner Schwächen und Stärken des Spielers ein Gesamtbild dessen fußballerischer Qualität wieder. Die Gesamtbewertung eines Spielers wird anhand der Bewertung der Oberkategorien bestimmt.

Tabelle 2: Ober- und Unterkategorien der Fähigkeiten eines Feldspielers

Oberkategorie					
pace	shooting	passing	dribbling	defending	physical
Unterkategorie					
acceleration	positioning	vision	agility	interception	jumping
sprint speed	finishing	crossing	balance	hd. accur.[14]	stamina
	shot power	fk. accur.[15]	reactions	def. awar.[16]	strength
	long shots	short pass	ball control	st. tackle[17]	aggression
	volleys	long pass	dribbling	sl. tackle[18]	
	penalties	curve	composure		

Die Dateigröße des Datensatzes beläuft sich auf etwas mehr als 94 Megabyte und besitzt 142.079 Datensätze. Somit enthält der Datensatz etwas mehr als 142.000 Spielerinformationen mit den entsprechenden Fähigkeitsdaten. Da sich aber Spieler in den FIFA-Teilen wiederholen besitzt jeder Datensatz einer FIFA-Serie zwischen 16.000 und 19.000 Datensätze bzw. Spielerdaten. Da sich die FIFA Fähigkeitsdaten auf die reale Leistung der Spieler beziehen, werden diese in dieser Arbeit als Datenbasis im Sinne der Sportdatenvisualisierung verwendet.

4.2 Erster Prototyp

Die Umsetzung der Visualisierung wird in Form eines ersten Prototyps des Spieleranalyse-Tools Tiko realisiert. Auf der Startseite (Home) kann jeder Spieler von FIFA 15 bis FIFA 22 ausgewählt werden. Somit stellt sie eine Datenbank mit Filterfunktionen dar, über die der zu analysierende Spieler ausgewählt werden kann. Den Filtereinstellungen entsprechend werden alle passenden Spieler zur Wahl angezeigt. Nach Wahl eines Spielers erscheint dieser in einem Auswahlbereich, wo dann dessen Fähigkeitswerte der Oberkategorie zu sehen sind. In Abbildung 2 ist das Auswahlfenster der Spieler dargestellt.

Zudem ist das Auswahlfenster mit Filterfunktionen ausgestattet, sodass die Suche erleichtert ist und gezielt nach bestimmten Merkmalen gesucht werden kann. Über ein Suchfeld kann ein Spieler anhand seines Namens gefunden werden. Zudem bietet eine Filterfunktion bezüglich der Ligen die Eingrenzung der angezeigten Spieler auf nur eine bestimmte Liga. Auch das Filtern nach Spielern auf nur bestimmten Positionen ist möglich, wie in Abbildung 3 dargestellt.

Große Dimensionen von Datensätzen können für die meisten Visualisierungstechniken schwerwiegende Probleme verursachen. Sie erhöhen nämlich die Generierungszeit zur Darstellung aller nötigen Informationen,

[14] heading accuracy

[15] free kick accuracy

[16] defending awareness

[17] standing tackle

[18] sliding tackle

was wiederum die Informationswahrnehmung und die Analyse der Daten behindert [12]. Dieses Bereitstellungsproblem von vielen Daten gleichzeitig ist ein gängiges Visualisierungsproblem [10]. Im Fall von Tiko trifft es zu, da bei der Spielerwahl Informationen aus mehreren Spielen und von mehreren Spielern berücksichtigt werden müssen. Ein gängiger Ansatz zur Lösung dieser Probleme ist die Verringerung der Datenmenge durch Techniken, wie z. B. dem Clustering oder der Filterung [11]. Um diesem Problem entgegenzuwirken, wurden die oben beschriebene Liga-Filterfunktion konzipiert. Sie ermöglicht das *clustern* der Spieler anhand ihrer Liga. Somit werden nur Spieler einer bestimmen bzw. ausgewählten Liga zur Wahl angezeigt, wodurch die Spielermenge von 1.000 Spielern nie überschritten wird. Folglich konnte die Generierungszeit der Spielerwahl in einem angemessenen Rahmen gehalten werden.

Auf der nächsten Seite (Detail View) sind die detaillierten Fähigkeitswerte des ausgewählten Spielers in Form von Visualisierungen dargestellt. Zudem kann bei Bedarf zwischen den unterschiedlichen FIFA-Teilen, solange der Spieler in diesen vertreten ist, getauscht werden. Das ermöglicht Änderungen der Fähigkeitswerte des Spielers über die FIFA-Serien zu erkennen. Die *detail view* beinhaltet Informationen wie den Namen, das Geburtsdatum, das Gewicht, die Größe und weitere personenbezogene Daten des Spielers. In Abbildung 1 wird diese Ansicht dargestellt.

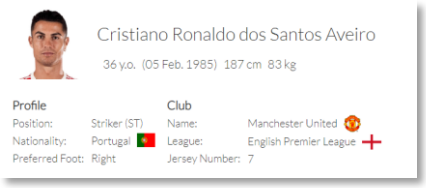


Abbildung 1: Spielerinformationen

Mittels eines Netzdiagramms werden die sechs Fähigkeiten (Oberkategorie) der Spieler abgebildet und veranschaulichen somit übersichtlich die Schwächen und Stärken des Spielers auf einen Blick. Anhand modifizierter Kreisdiagramme, auch bekannt als *donut chart*, werden die sechs Fähigkeitsattribute (Oberkategorien) erneut im Vergleich zueinander dargestellt. Die detaillierteren Fähig-

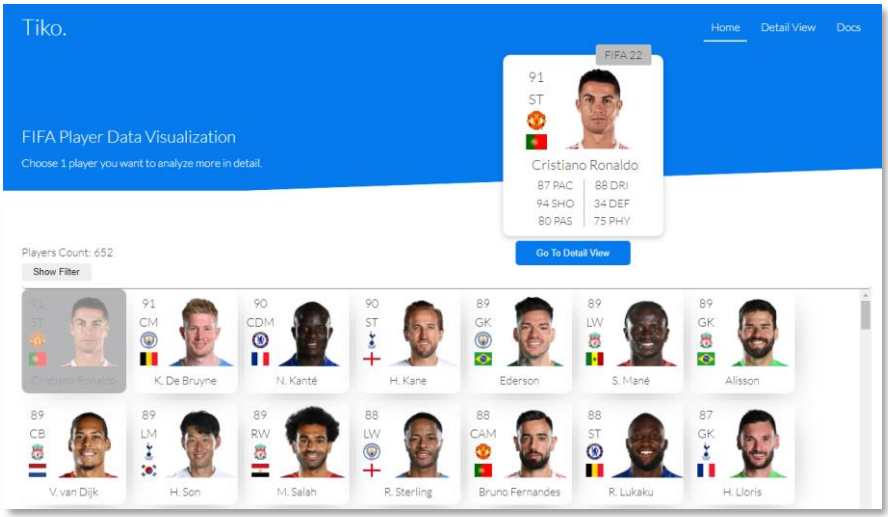


Abbildung 2: Spielerwahl

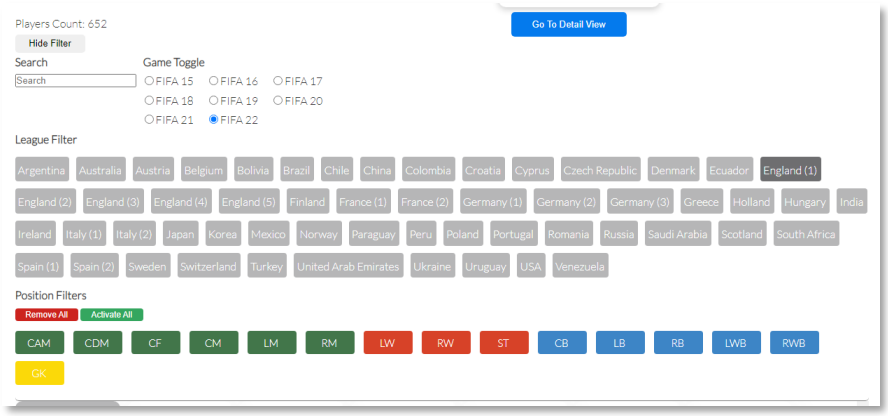


Abbildung 3: Filterfunktionen der Spielerwahl

keitswerte (Unterkategorien) werden in Form von Balkendiagrammen dargestellt. Über ein Liniendiagramm wird zudem der Entwicklungsverlauf des Spielers über die acht vergangenen FIFA-Teile dargestellt. Somit kann der ausgewählte Spieler bezüglich seiner Fähigkeitswerte sowie auch seinem Entwicklungsverlauf analysiert werden. In Kapitel 4.3 wird näher auf die Visualisierungen eingegangen.

4.3 Visualisierungstechniken in Tiko

In diesem Kapitel werden die in Tiko umgesetzten Visualisierungen näher erläutert. Ihre Wahl und Nützlichkeit werden anhand der verwendeten Quellen begründet.

4.3.1 Farben

Zur besseren Veranschaulichung der Fähigkeitswerte wurde die Fähigkeitsskala (0 bis 99) in Farbstufen aufgeteilt. Es wird unter fünf unterschiedlichen Fähigkeitsstufen unterschieden. Jeder Fähigkeitsstufe ist eine Farbe zugeordnet. Die Fähigkeitswerte von 0 bis 25, von 16 bis 45 und von 46 bis 65 werden in roter, orangener und gelber Farbe dargestellt, da Bewertungen in diesem Bereich als *schwach* eingestuft werden. Die Fähigkeitswerte in den Bereichen von 66 bis 85 und von 86 bis 99 werden in hell-grüner und dunkel-grüner Farbe dargestellt, da diese in einem Bewertungsbereich liegen der als *stark* eingestuft wird. Die Bereiche der Fähigkeitsstufen sind selbst gewählt. Die Farbwahl der Fähigkeitsstufen wurde durch Sofifa inspiriert. Die Farben werden im *Netzdiagramm*, den *Balkendiagrammen* und den *modifizierten Kreisdiagrammen* eingesetzt.

4.3.2 Netzdiagramm

Zur Veranschaulichung der Fähigkeiten wird wie in [4] und [8] vorgeschlagen ein Netzdiagramm genutzt. Dieses gibt einen guten Überblick über alle in FIFA bestimmten Fähigkeiten (Oberkategorie) eines Spielers wieder, wodurch Stärken und Schwächen des Spielers schnell ersichtlich werden. Beispielhaft werden in Abbildung 4 die Fähigkeiten eines Spielers mittels eines Netzdiagramms aus Tiko dargestellt. Durch die unterschiedlichen Stufen mit den entsprechenden Farben wird zudem verdeutlicht, ob gewisse Stärken oder Schwächen besonders ausgeprägt sich. Außerdem wird besser klar, welche Fähigkeiten im selben Bewertungsbereich liegen. Die Farbskala der Bewertungsstufen ist zudem immer unterhalb des Netzdiagramm zu sehen.

Abbildung 4: Netzdiagramm der Fähigkeiten eines Spielers (Oberkategorien)

4.3.3 Modifiziertes Kreisdiagramm

Um die Fähigkeitswerte der Oberkategorie von denen der Unterkategorie trennen zu können werden diese, wie bei FutHead, anhand veränderter Kreisdiagramme, auch bekannt als *donut chart*, dargestellt. Sie veranschaulichen, inwieweit ein Spieler diese Fähigkeit beherrscht. In Farbe wird der Wert abgebildet, während die restliche Fläche grau dargestellt wird. Abbildung 5 veranschaulicht beispielhaft ein solches *donut chart*. Je nach dem in welchem Stufenbereich sich der Fähigkeitswert befindet, passt sich die Farbe (siehe Kapitel 4.3.1) entsprechend an. Oberhalb des *donut charts* wird die Beschreibung der Fähigkeit angegeben. Unterhalb wird der Fähigkeitswert als Ziffer wiedergegeben.

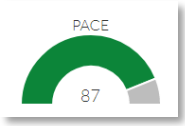


Abbildung 5: Modifiziertes Kreisdiagramm einer Fähigkeit (Oberkategorie)

4.3.4 Balkendiagramm

Die Fähigkeitsbewertungen der Unterkategorien der Spieler werden, wie bei FutBin und in [1] vorgeschlagen, in Form von Balkendiagrammen dargestellt. Abbildung 6 veranschaulicht die Visualisierung der Unterkategorien der Fähigkeit *pace*. Auch hier wird je nach Bewertungswert der Fähigkeit die Farbe des Balkens entsprechend der Stufenfarbe (siehe Kapitel 4.3.1) angepasst. Oberhalb wird die Bezeichnung der Unterkategorie der Fähigkeit und der Zahlenwert angezeigt. Die Balken besitzen eine Wertskala, wodurch die Werte selbst und Unterschiede zwischen den Fähigkeiten leichter wahrgenommen werden können.

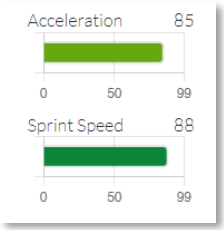


Abbildung 6: Balkendiagramm der Fähigkeiten (Unterkategorien)

4.3.5 Liniendiagramm

Der Entwicklungsverlauf der Spieler über die acht FIFA-Teile wird durch ein Liniendiagramm, wie in [4] in der aufgestellten Kategorie *Zeitreihenvisualisierung* vorgeschlagen, dargestellt. Es eignet sich gut für zeitlich-bedingte Verläufe. Da FIFA jährlich einen neuen FIFA-Teil veröffentlicht kann dieser als zeitlicher Verlauf in einem Liniendiagramm dargestellt werden. Positive sowie negative Entwicklungen eines Spielers werden verdeutlicht. Das Liniendiagramm veranschaulicht lediglich die Entwicklung der Gesamtbewertung eines Spielers. In Abbildung 7 wird beispielhaft ein Liniendiagramm aus Tiko dargestellt.

5 Ausblick

Der erste Prototyp von Tiko stellt die Basis für ein noch ausbaufähiges Spieleranalyse-Tool dar. Durch die bereits aufgebaute MySQL-Datenbank und der umgesetzten API-Schnittstelle können weitere Visualisierungstechniken integriert oder Verbesserungen an den umgesetzten Visualisierungen

vorgenommen werden. Zudem können in Zukunft weitere Datensätze von neuen FIFA-Teilen in die Datenbank eingepflegt werden. Die Nützlichkeit und Usability von Tiko sollte durch Nutzertests untersucht werden. Die daraus generierte Rückmeldung kann verwendet werden, um Tiko in Bezug auf die Visualisierung nützlicher und auch generell Benutzerfreundlicher zu gestallten.

Außerdem gilt in weiteren Arbeiten zu untersuchen, inwieweit die implementierten Filter geeignet sind. Denkbar wären Erweiterung der Filterfunktionen auf beispielsweise die Suche nach Spielern hinsichtlich ihrer Nation. Die Einzelspieleranalyse (Detail View) könnte um Visualisierungen erweitert werden, die neben dem Entwicklungsverlauf der Gesamtbewertung auch den Entwicklungsverlauf der Ober- und Unterkategorien ermöglichen. Außerdem könnte Tiko neben der Analyse eines einzelnen Spielers die Analyse von direkten Spielervergleichen ermöglichen. Auch die Analyse von Spielergruppen wäre denkbar. Um z.B. zu untersuchen, ob Spieler einer bestimmten Nation einen besseren *passing*-Wert vorweisen als Spieler anderer Nationen. Durch Ergebnisse in Evaluierungsstudien kann dann entsprechend ein zweiter bzw. verbesserter Prototyp implementiert werden.

6 Zusammenfassung

Dieses Paper beschäftigt sich mit der Visualisierung von Fähigkeiten von Fußballspielern. Aus einer Literaturrecherche wurden geeignete Visualisierungstechniken extrahiert und in einem selbstentwickelten Tool, das den Namen Tiko trägt, praktisch umgesetzt. Tiko wurde als Web-Anwendung mittels HTML, CSS, JavaScript und weiteren JavaScript-Frameworks implementiert. Als Datenbasis wird ein FIFA-Spieler Datensatz genutzt. Dieser beinhaltet personen- und vereinsbezogene Informationen der Spieler sowie auch die fußballerischen Fähigkeitswerte, die im Simulationsspiel FIFA die Leistungsfähigkeit der realen Fußballspieler darstellen soll. Der Datensatz enthält alle im Spiel vertretenen Spieler von FIFA 15 bis FIFA 22.

Die Spielerwahl, für den zu analysierenden Spieler, erfolgt durch ein separates Fenster und besitzt geeignete Filter. So kann vereinfacht nach gewünschten Spielern gesucht

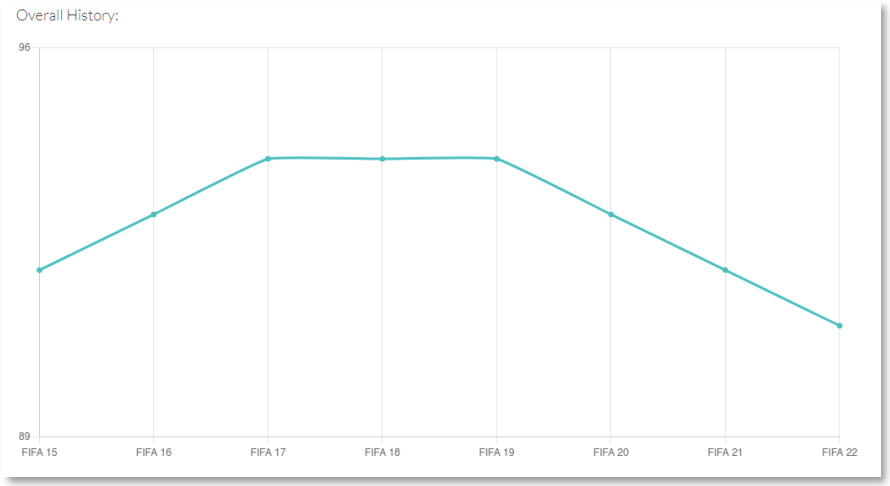


Abbildung 7: Liniendiagramm der Entwicklung des Spielers über die acht FIFA-Teile

werden. Die Daten werden in einer MySQL-Datenbank verwaltet. Über eine integrierte API-Schnittstelle greift das Front-End auf die Daten zu. Durch eine Literaturrecherche konnten geeignete Visualisierungen ermittelt und in Tiko realisiert werden. Die Fähigkeitswerte der Fußballspieler werden mittels Netzdiagrammen, modifizierten Kreisdiagrammen und Balkendiagrammen dargestellt. Die Entwicklung der Gesamtbewertung der Spieler über die acht FIFA-Teile wird in Form eines Liniendiagramms dargestellt.

In weiteren Arbeiten gilt es Tiko auf dessen Mehrwert zu untersuchen und mittels des Feedbacks einen verbesserten (zweiten) Prototypen zu entwickeln. So kann Tiko um weitere hilfreiche und nützliche Visualisierungen erweitert werden.

Literaturverzeichnis

[1] Afonso, A. P., Carmo, M. B., & Afonso, R. „VisuaLeague: Visual Analytics of Multiple Games“. In *2021 25th International Conference Information Visualisation (IV)*. pp. 54-62. IEEE. 2021.

[2] Bowman, B., Elmqvist, N., & Jankun-Kelly, T. J. „Toward visualization for games: Theory, design space, and patterns“. *IEEE transactions on visualization and computer graphics*, 18(11), pp. 1956-1968. 2012.

[3] Drachen, A., & Schubert, M. „Spatial game analytics and visualization“. In *2013 IEEE Conference on Computational Inteligence in Games (CIG)*. pp. 1-8. IEEE. 2013.

[4] Du, M., & Yuan, X. „A survey of competitive sports data visualization and visual analysis“. *Journal of Visualization*, 24(1). pp. 47-67. 2021.

[5] Gudmundsson, J., & Horton, M. „Spatio-temporal analysis of team sports“. *ACM Computing Surveys (CSUR)*, 50(2), pp. 1-34. 2017.

[6] Hamari, J., & Sjöblom, M. „What is e-Sports and why do people watch it?“. *Internet research*. 2017.

[7] He, J., Chen, H., Chen, Y., Tang, X., & Zou, Y. „Variable-based spatiotemporal trajectory data visualization illustrated“. *IEEE Access*, 7, 143646-143672. 2019.

[8] Krishnamurthy, P., & Nanda, N. „Dynamic Data Visualization for Soccer Performance Metrics“. In *2021 6th International Conference for Convergence in Technology (I2CT)*. pp. 1-11. IEEE. 2021.

[9] Olsen, A. H. „The evolution of e-Sports: An analysis of its origin and a look at its prospective future growth as enhanced by Information Technology Management tools“. *arXiv preprint arXiv:1509.08795*. 2015.

[10] Schertler, R., Kriglstein, S., & Wallner, G. „User guided movement analysis in games using semantic trajectories“. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. pp. 613-623. 2019.

[11] Tobler, W. R. „Experiments in migration mapping by computer“. *The American Cartographer*, 14(2). pp. 155-163. 1987.

[12] Wallner, G., Halabi, N., & Mirza-Babaei, P. „Aggregated visualization of playtesting data“. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. pp. 1-12. 2019.

©2022 Demis Moysidis. 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/

Visualisierungskonzept zur Vermittlung der Blockchain Technologie für Laien

Felix Wickert
Hochschule Reutlingen
Reutlingen, Deutschland
Felix_Johannes.Wickert@Student.Reutlingen-University.de

Abstract
Die hinter Kryptowährungen wie Bitcoin liegende Blockchain (BC) Technologie ist für viele potentielle Anwendende unverständlich, da sie sich durch ihre Funktionsweise stark von gewohnten Zahlungssystemen und den zugrundeliegenden Gedankenmustern unterscheidet. In der vorliegenden Arbeit werden anhand einer Literaturrecherche Herausforderungen in der Wahrnehmung von BC Laien identifiziert, um bereits bestehende Visualisierungen auf Ihren Aussagegehalt bewerten zu können. Das Ziel ist einen Anforderungskatalog für ein Visualisierungskonzept zu entwerfen, dass Laien die Eigenschaften und Funktionsweise der BC begreifbar macht und dieses prototypisch umzusetzen.

CCS Concepts
•Human-centered computing~Visualization~Visualization application domains~Information visualization

Keywords
Blockchain, Bitcoin, Kryptowährungen, Laien, Verständnis, Adoption, Human Centered Computing, Figma, Visualisierung

1 Einleitung
Laut [1] ist Bitcoin (BTC) eine aufstrebende Technologie, die noch nicht vollständig erforscht ist und daher daran scheitert, potentiellen Nutzenden ein vollständiges Bild ihrer Funktionsweise zu gewähren. In [10] wurde festgestellt, dass es für den regulären Internetnutzer allgemein schwierig ist, sich in das Themenfeld von Kryptowährungen (KW) und der zugrundeliegenden Blockchain (BC) Technologie einzuarbeiten. Dabei ist BTC eine von vielen KW. Die Schwierigkeit resultiert laut [10] daher, dass sich die meisten vorhandenen Informationen auf Nutzer beziehen, die eine höhere Technologieaffinität besitzen, wie Software Entwickler. Für einen Laien sei die Frustration und die benötigte Zeit, um sich in das Thema einzuarbeiten solch ein kritischer Faktor, dass das Interesse an dem Thema vorher verloren geht. Es sei bedeutsam, dass potentielle Nutzer in jeder Anwendungsform von BTC Zugang zu einer Art Hilfestellung erhalten würden, damit die Funktionen und Merkmale des Systems besser verstanden werden können. Das Ziel sei laut [10], einen Leitfaden oder ein Tool für Laien zu erstellen, dass das Wissen über die Technologie, sowie ihrer Vor- und Nachteile näherbringen kann. Zusätzlich sei

Betreuer Hochschule: Prof. Dr. Uwe Kloos
Hochschule Reutlingen
Uwe.Kloos@Reutlingen-University.de

Informatics Inside Frühjahr 2022
11. Mai 2022, Hochschule Reutlingen
Copyright 2022 Felix Wickert

es von höchster Bedeutung, dass technologische Vorteile, wie eine erhöhte Sicherheit der personenbezogenen Daten, stärker hervorgehoben würden. In [1] wird zusammenfassend ausgesagt, dass die Forschung zu KW trotz ihres kontroversen Austausches in Politik und Gesellschaft noch wenig vorhanden sei und die Schulung von potentiellen Anwendern eine neue Denkweise erfordere. Daher sollte ein Modell implementiert werden, dass die Grundlagen der KW für Laien verständlich veranschaulicht.

1.1 Adoption von BC

Für das BC Netzwerk, auf dem KW wie BTC aufbauen, ist es aufgrund der Funktionsweise der zugrundeliegenden Sicherheitsfunktion unerlässlich, die Adoption voranzutreiben. Bereits [2, 16] konstatierten, dass sich die Sicherheit des BC Netzwerkes erst mit kritischer Masse von Teilnehmern manifestiere. Auch [4] sehen die Notwendigkeit, Menschen über die Vorteile der Verwendung von BTC aufzuklären. Sie gehen noch weiter und empfehlen politischen Entscheidungsträgern sowie Händlern, sich auf die steigende Adoption von BTC vorzubereiten. In El Salvador wurde BTC als gesetzliches Zahlungsmittel anerkannt [3] und in Europa hat man sich gegen ein Verbot resultierend aus dem Stromverbrauch der Technologie entschieden[1]. Dies bestätigt, dass KW wie BTC zunehmend Einzug in den Alltag erhalten. Kritikpunkte wie ein hoher Stromverbrauch stellen Hemmnisse für die Adoption der BC Technologie dar, weswegen ein Bedarf für die Vermittlung der Vorteile der zugrunde liegenden Technologie von Bedeutung ist. Dabei muss zwischen BTC, der ursprünglichen KW, sowie deren Derivaten unterschieden werden. Gemeinsam haben KW allesamt, dass hinter ihnen eine Technologie steht, die als BC bezeichnet wird. Diese Technologien können in Bezug auf ihre Adoption gewissen Phasen zugeordnet und psychologisch nachvollzogen werden. Ähnlich wie die Adoptorkategorien[2], die die Verwendung einer Innovation von Individuen misst, visualisiert der Gartner Hype Cycle [8, 13] anhand verschiedener Emotionaler Zustände die Gefühlslagen, die Menschen beim Erscheinen einer innovativen Technologie durchlaufen. Es ist also bekannt, dass neue Technologien im ersten Schritt zu einem sogenannten Hype, einem Gipfel überzogener Erwartungen führen, der eine neue Technologie stark in den Fokus des Interesses rückt. Betrachtet man die Gartner Analyse für BC [8] erkennt man, dass sich diese in einem Pfad der Erleuchtung befinden. Das bedeutet, dass laut Gartner für BC eine Phase der steigenden Produktivität folgen wird. Hemmnisse für die BC Technologie resultieren einmal aus negativer Berichterstattung, wie dem Prozess des Minings, also dem Herstellen von BTC. Mining ist für seinen hohen Strombedarf bekannt, was in Europa[1] beinahe zu einem Verbot geführt hätte [9]. Ein weiteres Hemmnis ist die bereits in der Einleitung kritisierte Usability. Tatsächlich wird laut [4] das BC System hinter KW unter anderem aufgrund seiner komplexen Kryptographie von Nutzern und Händlern als zu schwierig zu verstehen angesehen, um es zu nutzen. Derartige Kritikpunkte stellen Hemmnisse für die Adoption der BC Technologie dar, weswegen ein Bedarf für die Vermittlung der Vorteile der zugrunde liegenden Technologie von Bedeutung ist.

1.2 Zielgruppe

Als Zielgruppe sind wie in [1, 6, 7] all jene zu verstehen, die bislang weder KW besessen, noch sich mit diesen auseinandergesetzt

[1] Davies, P. 2022. EU votes down proposal limiting PoW cryptos but sets new draft rules: *https://www.euronews.com/next/2022/03/14/europe-to-vote-on-limiting-pow-crypto-mining-used-by-bitcoin-and-ethereum.* Accessed: 04.05.2022.

[2] Haas, H. Definition: Adoptorkategorien: *https://wirtschaftslexikon.gabler.de/definition/adoptorkategorien-30845*, Accessed: 04.05.2022

haben. Diese Personen sich noch nicht mit den Grundlagen der BC vertraut. Momentan scheinen sich noch vornehmlich männliche Personen für die BC Technologe zu interessieren, ein Wandel ist aber im Gange. Während es 2013 laut [3] nur etwa 5 % weibliche BTC Nutzer gab, spricht [4] von inzwischen etwa 30 % Anteil an Nutzerinnen. Überlegt man, welche Zielgruppe mit einer Visualisierung angesprochen werden soll, die die grundlegenden Eigenschaften einer BC erläutern soll, schließt sich die Frage an, wer heutzutage welche Medien konsumiert. In [5] wurden Jugendliche und ihre Gedankenmuster untersucht. Es wurde herausgefunden, dass die Mediennutzung innerhalb von Familien immer selbstverständlicher wird und Eltern inzwischen selbst versierte Onliner seien [5]. Der Anspruch an eine sinnstiftende Visualisierung, um nicht nur Jugendliche anzusprechen, sollte also derart aufgebaut werden, als dass auch Menschen weit über 60 von den Erkenntnissen profitieren können. Dies insbesondere, da die Verbreitung des Internets gerne als Maßstab für die Adoption von KW verwendet wird[5]. Es kann geschlussfolgert werden, dass Nutzer aller Altersgruppen das Internet benutzen. Betrachtet man auch über 80-jährige als Zielgruppen, ist [24] relevant. Hier wurde eine Umfrage zum Thema Onlinebanking durchgeführt und herausgefunden, dass gerade ältere Nutzer neueren oder unbekannten Bezahlmethoden gegenüber skeptisch sind. Aus diesem Grund fokussiert sich diese Ausarbeitung auf allen Personen, die auch aufgrund von Unwissenheit oder Skepsis noch keine Berührungspunkte mit der BC Technologie hatten. Für die Definition von Laien werden wie in [1, 6, 7] alle Nutzer bezeichnet, die bislang nicht mit KW zu tun hatten oder diese besaßen.

2 Ziel der Arbeit

Ziel dieser Ausarbeitung soll sein, einen Beitrag zur Adoption der BC zu leisten, indem eine Visualisierung entworfen wird, die von unterschiedlichsten Altersgruppen gleichermaßen nachvollzogen werden kann. Um das Verständnis für die Funktionsweise der BC für Laien voranzubringen, sind die einer Visualisierung zu extrahierenden Informationen von Bedeutung. Aufgrund dessen wurden bereits vorhandene Visualisierungen untersucht. Anhand dieser Erkenntnisse ist es wichtig, Laien mit Hilfe von Informationen zu der BC über deren Funktionsweise und Sinnhaftigkeit aufzuklären. In dieser Ausarbeitung wird deshalb untersucht, welche Hemmnisse einer breiten Adoptierung entgegenstehen, indem verschiedene Lektüren gesichtet und wichtige Kriterien extrahiert werden sollen, um die hinter den KW liegende Technik der BC der Allgemeinheit besser zugänglich zu machen. Dabei wird die Forschungsfrage gestellt,

Wie kann die Funktionsweise der BC einfach und verständlich für Einsteiger in das Thema als Visualisierung aufbereitet werden?

3 Grundlagen der BC Technologie

Um ein tieferes Verständnis für die BC Technik zu erhalten und die essentiellen, der Zielgruppe zu vermittelnden Grundlagen zu identifizieren, wird im Folgenden der Hintergrund der Technologie betrachtet.

---

[3] Simulacrom 2013. The Demographics of Bitcoin. *https://simulacrum.cc/en/the-demographics-of-bitcoin/* Accessed: 04.05.2022.

[4] Konish, L. 2022: Why U.S. minority communities may turn to cryptocurrencies to pay their bills: *https://www.cnbc.com/2022/02/08/-research-shows-cryptocurrency-adoption-among-hispanics-is-high.html*. Accessed: 04.05.2022.

[5] Wan, S. 2021. Internet vs. crypto adoption: *https://cryptoslate.com/internet-vs-crypto-adoption-chart-predicts-1-billion-users-by-2027/*. Accessed: 04.05.2022

3.1 Hintergrund

Die Funktionsweise der BC wurde 2008 als White Paper veröffentlicht und das Netzwerk 2009 vollumfänglich zugänglich gemacht [2, 15, 18]. Die BC enthält verschiedene Techniken wie Kryptographie, die in ihren einzelnen Funktionsweisen bereits zuvor veröffentlicht, aber bis zu diesem Zeitpunkt nie gemeinsam nutzbar waren [2, 21]. Als Ausgeber des White Papers und der Veröffentlichung steht das Pseudonym Satoshi Nakamoto, dessen Anteil an BTC bis zum Datum der Ausarbeitung seit Jahren nicht bewegt worden ist [15, 18]. Im Gegensatz zu anderen KW existiert beim BTC Netzwerk keine zentrale Instanz wie ein Management, während Entwickler gemeinsam an einem störungsfreien Betrieb arbeiten [2, 15].

3.2 Funktionsweise des BC Netzwerks

Die vornehmliche Aufgabe des BC Netzwerkes ist es, Transaktionen wie auf einem verteilten Hauptbuch eines Kontos abzubilden und gegen Veränderung zu schützen. Das Hauptbuch fasst alle Transaktionsvorgänge wie Ein- und Ausgänge zusammen. Dabei kommunizieren ausschließlich verschiedene Peers miteinander, es existiert keine zentrale Instanz wie eine Bank als Intermediär. Es wird über Konsensverfahren wie dem Proof of Work entschieden, welche Transaktionen, die zuvor im Mempool vorgehalten wurden, in den nächsten Block integriert werden. Das Konsensverfahren ist wie ein Protokoll zu verstehen, dass die Bedingungen definiert, unter denen Werte ausgetauscht werden dürfen. Der Mempool umfasst alle Transaktionen, die noch nicht in der BC erfasst und dauerhaft gespeichert wurden. Im Kapitel 3.4 wird nochmal auf die Elemente der BC eingegangen. Im Folgenden werden zunächst die Eigenschaften der BC beschrieben, die sich von der Funktionsweise des BC Netzwerkes ableiten lassen.

3.3 Eigenschaften der BC

Für das Verständnis der Funktionsweise des BC Netzwerkes ist es nicht erforderlich, den Quellcode zu analysieren. Durch verschiedene Literatur wie [11, 12, 19, 20] wurden bereits Kriterien abgeleitet, die als Eigenschaften des Implementierten Codes verstanden werden können und vom Autor auf Verständnis innerhalb der Generation Z geprüft wurden [26]. Die identifizierten Eigenschaften, die die BC ausmachen sind *Dezentralität, Unveränderbarkeit durch Konsens, Vertrauen, Redundanz, Offenheit, Anonymität, Transparenz und Stromverbrauch.* Sie werden im Folgenden erläutert. Die BC des BTC wird in [6] und [7] als einzige KW vom Vorstand der US-Börsenaufsicht SEC aufgrund seiner Dezentralität als Ressource bezeichnet. Das bedeutet, dass er steuerlich Gold gleichgestellt wird. Die restlichen KW werden eher Aktienanteilen zugeordnet und somit strenger reguliert. Die Dezentralität bedeutet, dass im Gegenzug zu einer Bank kein Intermediär wie ein Bänker nötig ist, dem der Nutzer vertrauen muss. Das Vertrauen in die Kompetenz des Bänkers, Angriffsflächen wie Double Spending Attacken zu vereiteln und eine funktionale Infrastruktur bereitzustellen, wird durch Regeln in Form von Algorithmen ersetzt. Diese prüfen angestoßene Transaktionen selbstständig und führen sie mit Hilfe des Mining Prozesses durch [11, 15, 20]. Als Resultat der Dezentralität entstehen weitere Eigenschaften der BC. Um einen Intermediär auszuschalten, muss das Haupt-

[6] Bloomberg SEC's Meme Stock Response Coming Next Week, Gensler Says: *https://www.bloomberg.com/news/videos/2022-02-03/sec-s-meme-stock-response-coming-next-week-gensler-says-video.* Accessed: 04.05.2022.

[7] SEC gov Remarks Before the Aspen Security Forum: *https://www.sec.gov/news/public-statement/gensler-aspen-security-forum-2021-08-03?utm_medium=email&utm_source=govdelivery.* Accessed: 04.05.2022.

buch mit den Informationen zu den Kontentransaktionen redundant an möglichst vielen verschiedenen Orten aufbewahrt werden, wodurch auch eine Unveränderbarkeit erreicht wird. Dies kommt daher, da mindestens 51% der Netzwerkteilnehmer in Form der Konsensbildung einer Veränderung der Historie zustimmen müssten, um Daten der BC zu verfälschen [2, 15]. Zusätzlich ist derselbe Aufwand an Strombedarf nötig, den es für das Herstellen der zu verändernden Blöcke benötigt hat [2, 15]. Es existiert ein hohes Maß an Transparenz, ohne dass die Verfolgbarkeit von Zahlungen innerhalb des Netzwerkes nicht möglich wäre. Hierzu können Webseiten wie der BC Explorer[8] verwendet werden, um Transaktionen und Beträge abzufragen, die sich auf Wallets befinden. Eine weitere Eigenschaft ist die Offenheit für jeden, der dem Netzwerk beitreten möchte sowie die Anonymität, die sich aus der Repräsentanz der Teilnehmer in Form von Hashwerten erschließt. Werden die BC Vorgänge Mithilfe des BC Explorers[8] betrachtet, werden keine Spitznamen oder Avatare von Personen dargestellt, sondern nur deren Wallet Adressen. Diese Anonymität hat aber Grenzen und wird auch als Pseudo-Anonymität bezeichnet, da mit Hilfe von Aus- und Einzahlungen auf Konten von Personen geschlossen werden kann[9]. Auch wurden Betrüger bereits erfolgreich durch digitale Forensik und das einfache Verfolgen der Transaktionsverläufe durch den BC Explorer[8] identifiziert[9]. Im Folgenden werden die Elemente der BC dargestellt, die durch Explorer[8] und Quellcode nachvollzogen werden können [2, 15].

3.4 Elemente der BC

Die höchste Granularität hat die Transaktion, die auch als TX bezeichnet wird. Zu ihr gehört auch eine Gebühr. Transaktionsdatensätze werden in einem Mempool vorgehalten, bis eine neue Datenstruktur namens Block von einem Miner gefunden wird. Für das Minen muss eine Kopie der Historie des Hauptbuches gespeichert und mit Hilfe einer Recheneinheit wie einer Grafikkarte oder einem Prozessor ein anspruchsvolles Rätsel gelöst werden. Dies bedingt einen hohen Stromverbrauch. Die gefundenen Blöcke aggregieren dann eine gewisse Anzahl an Transaktionen und erhalten eine Verschlüsselung, die den Status aller in einem Block enthaltenen Transaktionen miteinschließt. Der nächste Block erhält diese Kennung und gibt sie weiter, sodass alle Blöcke kryptografisch verbunden sind. Es entsteht eine verbundene Blockkette. Hierdurch wurde ein sich selbst kontrollierendes System geschaffen, dass bemerkt, wenn einzelne Transaktionen innerhalb eines Blockes unerlaubt verändert werden sollen. Da sich so der von den in einem Block befindlichen Transaktionen abgeleitete Hashwert in einen ungültigen Zustand ändern würde, lehnt der folgende Block die Manipulation ab. Sogenannte Knotenpunkte agieren ähnlich wie Miner. Anstatt neuer Blöcke zu finden benötigen sie weniger Strombedarf, weil sie nur die gesamte Historie des Hauptbuches verwalten und im Netzwerk verteilen. Sie synchronisieren also die BC, während die Miner dazu beitragen, dass neue Blöcke entstehen und sich die Historie der gespeicherten Daten vergrößert. Ein BC Nutzer besitzt ein Wallet, dass sozusagen seine digitale Repräsentanz auf der BC darstellt und Transaktionen empfangen und versenden kann. Nutzer können so viele Wallets anlegen, wie sie möchten. Es wird empfohlen, für jeden Empfang eine eigene Walletadresse zu erstellen [15]. Denn durch den Transaktionsverlauf können Rückschlüsse auf die zu einem Wallet und

[8] Blockchain Explorer 2020. *https://www.blockchain.com/explorer*. Accessed: 04.05.2022.

[9] Biselli, A. 2019. Blockchain-Forensik: Wer steckt hinter einer Bitcoin-Zahlung? netzpolitik.org: *https://netzpolitik.org/2019/blockchain-forensik-wer-steckt-hinter-einer-bitcoin-zahlung/*. Accessed: 04.05.2022.

ihren Transaktionen zugehörigen realen Besitzern geschlossen werden. Reguläres Währungsgeld wird als FIAT bezeichnet, was bedeutet, dass es von einer Regierung ausgegeben wird und nicht mit Gold gedeckt ist [2]. FIAT kann auf KW Börsen gegen KW getauscht werden. Da die BC des BTC nur wenige Transaktionen gleichzeitig verarbeiten kann, wird an weiteren Ebenen gearbeitet. Diese Ebenen werden auch Layer genannt. Lightning ist ein Netzwerk, das auf der BC von BTC aufsetzt und dass die Transaktionen mit Hilfe einer weiteren Ebene um ein Vielfaches erhöhen soll [17].

4 Stand der Forschung

Im Folgenden werden die Grundlagen der Informationsvermittlung skizziert, sowie unterschiedliche BC Visualisierungen auf ihre Eigenschaften und deren Aussagekraft für Laien hin betrachtet. Diese Betrachtungen dienen als Basis, um ein zielgruppengerechtes Visualisierungskonzept der BC zu entwickeln.

4.1 Informationsvermittlung

Während das Wissen zu der BC Technologie noch vornehmlich per Sprache, Schriften und YouTube Videos weitergegeben wird, gab es bislang wenig Versuche, BC Interessenten Informationen vereinfacht zukommen zu lassen. Dabei wurde schon früh realisiert, dass sich Informationen gepaart mit Bildern leichter im Gedächtnis speichern lassen [23]. Der Paradigmenwechsel der Wissenschaft in Form des Narratives kognitiver Arbeit des menschlichen Gehirns lässt sich bspw. durch Colin Ware nachvollziehen, der sein Buch *Visual Thinking* aus dem Jahr 2008 just 2022 mit einem neuen Kapitel aus der kognitiven Wissenschaft ergänzt hat [25]. Es wird hinterfragt, dass der Mensch sein visuelles Denken nicht wie ein Speicher eines Computers verwendet, der die Vergangenheit fokussiert. Stattdessen sind wir auf unser zukünftiges Handeln spezialisiert, prädiktive Kognition genannt [25]. Der Mensch lebt also von Natur aus in einem Zusammenspiel aus zukunftsorientiertem Denken, gepaart mit dem Bedürfnis, sich nicht zu weit aus seiner Komfortzone entfernen zu müssen [14]. Für eine Visualisierung ist eine wichtige Erkenntnis, dass der Mensch Gelerntes verwendet, um aus dem Gelernten Prognosen für die Zukunft zu erstellen. Das bedeutet, dass Menschen womöglich durch eine geeignete Visualisierung der Grundlagen der BTC dem Thema auch mit mehr Vorwissen und Sicherheit anstatt Skepsis begegnen.

4.2 Visualisierungen im Kontext von BC

Da sich diese Ausarbeitung an Laien wendet, wurden insbesondere Visualisierungen betrachtet, die eine gewisse Aussagekraft zur Funktionsweise enthalten. Bei der Recherche nach BTC Visualisierungen wurden verschiedene Ansätze betrachtet. Hervorzuheben ist [22], wo bereits 76 Visualisierungen gefunden und nach Zielgruppe und Fokus kategorisiert wurden. Für diese Ausarbeitung von Interesse sind vornehmlich 16 Visualisierungen, die sich aufgrund der niedrigen Analyseebene an Laien richten sollen. Diese anfängergeeigneten Visualisierungen sollen zwanglos über die BC informieren. Beispielsweise visualisieren aus [22] Bitbonkers, BTC Monitor, BTC Rain, Bitcoin Transaction Visualization, BTC City, Bitforce5, Bitlisten, TxStreet, chainFlyer, Realtime Bitcoin Globe vornehmlich das grundlegende Element der BC, nämlich die unwiderrufliche und vollumfängliche historische Abbildung von Transaktionen. Blockstream[10] visualisiert in einem Video die Anzahl an Transaktionen, die von unterschiedlichen Systemen wie Lightning pro Sekunde ausgeführt werden können. Hier besteht ein gewisses Lernpotential, da mit Visa

[10] Blockstream 2022. With an estimated limit of 500 TPS. Blockstream@Twitter: *https://twitter.com/Blockstream/status/1502115599874359297*. Accessed: 04.05.2022.

und Paypal Relationen zu bekannten Zahlungsdiensten vorliegen und zumindest abstrakt verstanden werden kann, dass sich diese Systeme in irgendeiner Form von KW abgrenzen. Auch bei Monospace[11] liegt der Fokus bei der Abbildung der Transaktionen auf der BC, sodass die bisher nicht bestätigten Transaktionen im unteren Bereich des Bildschirms als Blöcke dargestellt werden. Das Ziel der Visualisierung ist, den Übergang der Transaktionen in einen durch einen Miner geschürften Block darzustellen. Zur Interaktion kann der Nutzer die Maus über die Blöcke legen, um Informationen zu den einzelnen Transaktionen zu erhalten. Auf der VIS2020, einer IEEE Visualisierungskonferenz wurde SilkViser[12] vorgestellt [28]. Das Tool wurde mit vier verschiedenen Views angelegt, die eine Transaktion im BC Netzwerk hierarchisch als einzelne Komponenten darstellen. Es soll in erster Linie die Frage beantworten, welcher Zusammenhang zwischen der BC selbst, den Blöcken und Transaktionen, bis hin zu den einzelnen Adressen herrscht, hinter denen sich die BC Nutzer befinden. Noch umfangreicher ist TradeBlock[13] aufgebaut. Hier erhalten Nutzer in Graphen alle Informationen aufbereitet, die sonst auf KW Börsen für das Handeln vorhanden sind. Es werden Kursverläufe gezeigt, aber auch Informationen über den Mempool oder KW Bestände von Börsen sind visualisiert. In [22] wurden auch anspruchsvolle Projekte wie [14] und [15] vorgestellt. Bei Plantoids haben verschiedene Künstler und Programmierer zu einer, laut ihrer Aussage neuen Lebensform beigetragen, die zwischen Elektronik, Metall, Digitalisierung und BC existiert und durch eine digitale wie auch physische Repräsentanz in all diesen Ebenen existieren und interagieren soll. Auch OnBrink[15] visualisiert die Vorgänge auf der BC physisch, indem auf einem Band die entstehenden Blöcke in Realtime als Erdhaufen mit der Größe passend zu der Anzahl an Transaktionen repräsentiert werden. The Symphony of Blockchains aus [22] verfolgt einen astronomisch künstlerischen Ansatz. Die Blöcke stellen Planetenlinien dar, die um eine Zusammenhäufung aus Sternen zirkulieren, die den Mempool visualisieren.

5 Diskussion und Ergebnisse

Für die Ableitung einer für Laien geeigneten BC Visualisierung wird die Aussagekraft aktueller Visualisierungen bewertet, um Potentiale für ein aussagekräftiges Konzept zu identifizieren. Es werden allgemeine Herausforderungen von Laien einbezogen. Im Anschluss wird das Konzept vorgestellt.

5.1 Bewertung vorhandener BC Visualisierungen

Die in Kapitel 3.3 beschriebenen Visualisierungen verfolgen verschiedene Ziele der Informationsvermittlung und sind meist zweckgetrieben. Da KW Transaktionen abbilden, wird der Fokus der Darstellung auf Themen reduziert, die dem Visualisieren von Transaktionen und einem historischen Verlauf nachgehen. Da die Funktionsweise der

[11] Bitfeed: *https://bits.monospace.live/*. Accessed: 04.05.2022.

[12] Zhong, Z. et al. 2020. SilkViser: A Visual Explorer of Blockchain-based Transaction Data: *https://virtual.ieeevis.org/year/2020/session_f-papers-finance-blockchain.html*. Accessed: 04.05.2022

[13] TradeBlock: *https://tradeblock.com*. Accessed: 04.05.2022.

[14] Plantoids 2015: Blockchain-based life forms: *http://plantoid.org/*. Accessed: 04.05.2022.

[15] Stupp, D. 2018. ON BRINK: Live Physicalization of the Bitcoin Blockchain: *http://dataphys.org/list/on-brink-live-physicalization-of-the-bitcoin-blockchain/*. Accessed: 04.05.2022.

BC vornehmlich auf den in Kapitel 3.3 beschriebenen Eigenschaften der BC beruht, ist fraglich, ob Visualisierungen der Transaktionsvorgänge helfen, dem in [7] beschriebenen, fremden Gefühl der Nutzer entgegenzuwirken. Im Gegensatz zu den erwähnten Visualisierungen enthält die Darstellung Monospace[11] bereits mehr Aussagekraft über die Funktionsweise des BC Netzwerkes des BTC, indem sie den Mempool visualisiert. Es existieren bislang vergleichsweise wenige Anwendungen wie Silkviser oder auch TradeBlock, die zumindest das Potential bieten, Laien über die BC aufzuklären. Hierbei besteht dasselbe Risiko, das in [7] beschrieben wurde: Ein Laie ohne grundlegendes Verständnis wird anhand der Visualisierung nur interpretieren können, wie das zugrundeliegende technische Konstrukt arbeitet und sich somit aufgrund der Komplexität von dem Konzept distanzieren, was im Umkehrschluss zu einer niedrigeren Akzeptanz führt. Während diese Visualisierungen hilfreich für das Verständnis sein können, müssen grundlegende Elemente und Funktionsweisen der BC vorher verstanden werden. Besonders die Visualisierungen mit physikalischer Repräsentanz lassen die Annahme zu, dass Menschen nach Metaphern suchen, um die Funktionsweise der BC zu ergründen und diese außenstehenden künstlerisch nahezubringen. Es kann ein Bedarf unterstellt werden, für das Begreifbar machen von BC Anwendungen. Für das Verständnis der der BC zugrundeliegenden Komplexität können diese Visualisierungen als Metapher dienen, aber ein tieferes Verständnis vermitteln sie aufgrund des abstrakten Levels genauso wenig, wie die zu pragmatischen Live Abbildungen von Transaktionen. Für eine höhere Akzeptanz muss der Laie bei seinem Stand des Wissens abgeholt werden. Selbst das pure Abbilden von Transaktionsvorgängen setzt voraus, dass dem Zuschauer bewusst ist, dass diese Transaktionen live abgebildet werden und tatsächlich Transaktionen in einem vollständig dezentralisierten System sind.

5.2 Anforderungen für BC Visualisierungen

Es gibt mehrere Studien wie [1], in denen Herausforderungen für Laien beim Anlegen eines Accounts bei Börsen beobachtet, und Defizite der Benutzung auf Webseiten oder beim Anlegen eines eigenen Wallets identifiziert wurden. Das vereinfachte Darstellen des User Interfaces oder das Anbieten einer zweiten, für Laien geeigneten Darstellung auf der Webseite einer Börse kann hierbei ein erster Ansatz sein und wird auch in [1] vorgeschlagen. Dem gegenübergestellt werden kann das Ziel dieser Ausarbeitung, dem Nutzer das grundlegende Verständnis zu den aus Kapitel 3.3 erwähnten Eigenschaften der BC zu vermitteln. Den geschilderten Erkenntnissen der Studie [7] wurde entnommen, dass die Technologie der BC den Probanden technisch anspruchsvoll, schwer zu begreifen und fremd vorkam. Um diesem Umstand vorzubeugen ist ein Ansatz, die in Kapitel 3.3 beschriebenen Eigenschaften der BC anhand eines FIAT Transfers zu visualisieren. Eine Anforderung an eine Visualisierung stellt sich in der Form dar, als dass sie den Laien Berührungsangst nehmen soll, indem sie beispielsweise in Börsen, Schulungen oder Webseiten eingebettet werden kann und somit überall zur Verfügung steht. Sie soll Laien von einer Nichtnutzung abhalten, die von der Angst resultiert, eine Technik zu verwenden, die man nicht vollständig durchdrungen hat. Da wie in [7] festgestellt, auch die Konzepte des Tauschmittels Gold oder die einer Kreditkarte zugrundeliegende Funktionalität nicht von allen Nutzern durchdrungen wurden, sind die Ansprüche an die Visualisierung nicht in die Tiefe gehend. Zudem ist anzunehmen, dass eine schriftliche oder videobasierte Erläuterung einer BC nicht den erwünschten Lerneffekt hätte. Diese Erkenntnis lassen [7, 24] in Bezug zu Gold und Kreditkarten zu. Auch würde das bloße umschreiben des Codes nicht mit den in [7, 20] dargestellten Vorurteilen aufräumen. Es soll eine visuell unterstützte, datenbasierte Möglichkeit gefunden

werden, die umfangreichen Eigenschaften der BC auch an technisch nicht versierte Menschen zu vermitteln. Das Darstellen von Transaktionen ist genauso wenig ein Ziel, wie eine neue Übersicht zu der Relation zwischen Transaktionen und Blöcken. Es gilt vielmehr, grundlegende Informationen zu vermitteln, damit sich die Technik weniger fremd [7] anfühlt. Für das Verständnis ist nicht relevant, wie genau ein Hashwert einer Wallet Adresse zuzuordnen ist und wie Private Keys mit Public Keys eingesetzt werden. Sondern dass es Aufgabe der Kryptographie ist, die persönlichen Informationen von Nutzern zu verschlüsseln. Und dass es im Kontrast zur Anonymität Aufgabe der Transparenz ist, die BC durch Webseiten wie den BC Explorer mit seiner vollständigen Historie trotzdem vollständig einsehbar zu machen. Eine Annahme ist, dass neben dem Anbieten eines für Laien geeigneten User Interfaces die Nutzer einer KW Börse mehr Verständnis für die Komplexität hätten, wenn Sie über die Funktionsweise und somit die Vorteile bei der Nutzung einer BC besser informiert wären. Gerade der Vorgang des Minings ist den Nutzern nicht geläufig und differenziert sich stark von regulären Zahlungssystemen. Diese Eigenschaften sollen Einzug in das Visualisierungskonzept erhalten.

5.3 Visualisierungskonzept für Laien

Bereits in [24] wurde herausgefunden, dass ältere (mind. 70 Jahre) Nutzer neueren oder unbekannten Bezahlmethoden gegenüber skeptisch sind. Dabei wurde eine Umfrage zum Thema Onlinebanking durchgeführt. Um das Problem des fehlenden Vertrauens zu lösen, wurde in [24] das System Cheque Mate ins Leben gerufen, bei dem Schecks mit einem digitalen Stift ausgefüllt wurden, um die handschriftlichen Daten mit diesem Transfer zu digitalisieren. Nimmt man dieses Vorgehen als Beispiel, gilt es zunächst ein Medium zu finden, das von unterschiedlichen Altersgruppen verwendet wird. Es soll nach Möglichkeit keine Personengruppe mit anspruchsvollen Installationsroutinen oder vorauszusetzendem Fachwissen ausgeschlossen werden. Im Kontext von KW steht das Abbilden von Finanzen und Transaktionen im Fokus, weswegen eine Verbindung zu bisherigen Zahlungssystemen, ähnlich wie in [24] umgesetzt, denkbar ist. Sucht man wie in [24] den kleinsten gemeinsamen Nenner, ist für die Arbeit und das Mieten/Vermieten einer Wohnung unerlässlich, ein Konto zu führen. Laut [16] besitzen 99 % der Erwachsenen in Deutschland Konten. Das Verwenden von Transaktionen über einen Überweisungsträger oder per Onlinebanking Möglichkeit sollte somit für potentielle KW Nutzer gewohnt sein. Zum Ausführen und Nutzen eines Kontos gehören Kontoauszüge, die die Benutzer direkt in der Bank ausdrucken, online abrufen oder die sie per Post oder E-Mail zugesandt bekommen. Es kann also konstatiert werden, dass geschäftsfähige, erwachsene Personen Erfahrungen mit Überweisungen und Kontenauszügen gemacht haben. Daher soll metaphorisch ein Dokument aus einer Bank so graphisch manipuliert werden, dass die Eigenschaften aus Kapitel 3.3 anhand kleiner Veränderungen am Dokument, aber auch durch das Vermitteln einer Geschichte nachvollzogen werden können. Die Geschichte soll erzählt werden, um auch Menschen anzusprechen, die keine Erfahrungen mit Konten haben und sich mehr auf diese Darstellungsform konzentrieren möchten. Um den Transfer zwischen FIAT Geld und KW zu veranschaulichen, soll der Nutzer anhand verschiedener Themenfelder Punkte auswählen können, die für sich gesehen die einzelnen Eigenschaften der KW darstellen. Sie sollen

[16] Brandt, M. 2019: Wo ein Bankkonto nicht selbstverständlich ist: *https://de.statista.com/infografik/18482/top-10-laender-nach-anteil-der-erwachsenen-die-kein-bankkonto-haben/. Accessed*: 04.05.2022.

aber in einem vorher definierten Ablauf dargestellt werden, damit die einzelnen Eigenschaften, wie im Folgenden beschrieben, ineinandergreifen und aufeinander aufbauen können. Es wird quasi eine Hierarchie erstellt, die sich auch dem Zweck der BC selbst entnehmen lässt. Während die wichtigste Eigenschaft die Dezentralität ist, ergeben sich daraus wie in Kapitel 3.3 beschrieben, die weiteren Eigenschaften der BC.

Der im Folgenden beschriebene Figma Prototyp kann unter [27] betrachtet werden und ist im Anhang beispielhaft dargestellt. Um die Möglichkeiten der Mobilität zu visualisieren, wurde ähnlich zu den Bussen aus TxStreet in [22] ein Boot verwendet. Die Bedeutung von internationalen Gewässern soll allgemein verständlich und eine Metapher für die Dezentralität sein. Sie soll für einen unregulierten Raum zwischen den Ländern stehen. Der Prototyp kann mit den Buttons oben links linear bedient werden, um vorwärts oder rückwärts zu blättern. Betätigt der Nutzer den Button Dezentralität, verschwindet der Bankenname und eine Metapher für das Peer to Peer Netzwerk wird angezeigt (vgl. Anhang Darstellung 2). Wird die Metapher eines Bootes bemüht, ist es sinnstiftend, noch vor der Dezentralität die *Offenheit* des Netzwerkes zu visualisieren. Hierzu können Menschen verschiedenster Farben, denen man ansieht, dass sowohl Reiche als auch Arme unter ihnen sind, zu Beginn in das Boot steigen (vgl. Anhang Darstellung 1). Um den Punkt der *Transparenz* zu visualisieren können die Finanzbeträge, die die Bootsinsassen besitzen, angezeigt werde. Das Boot selbst hat Informationen darüber, wie viele BTC es transportiert, wie viele es jemals geben wird und wie viele momentan existieren (vgl. Anhang Darstellung 3). Wird der Punkt *Anonymität* betrachtet, wird der Name aus der Überweisung mit einem Wallet überschrieben und die Bootsfahrer erhalten Masken (vgl. Anhang Darstellung 4). Wird der Button Mining aktiviert, werden die Transaktionen auf dem Überweisungsträger hervorgehoben, um darauf hinzuweisen, dass das Mining mit Transaktionen zu tun hat. Die Bootsinsassen zeigen gleichzeitig Beträge an, die sie anschließend wechseln möchten und die im nächsten Schritt an Land von einer Instanz in einem Block gespeichert werden steigen (vgl. Anhang Darstellung 5). Die hierzu nötige Arbeit und der Ressourcenverbrauch werden durch Aktivieren des nächsten Buttons, *Stromverbrauch* dargestellt. Während Blitze den Stromverbrauch und die Spitzhacke die Arbeit visualisieren, zeigen die Insassen an, wer welchen Betrag tauschen möchte steigen (vgl. Anhang Darstellung 6). Die Informationen werden alle in den Block geschrieben, den die Instanz an Land gefunden hat. Erst wenn der Block gefunden wurde, werden die finalen Beträge angezeigt. Um die *Redundanz* darzustellen, erhalten alle Personen Informationen über den Inhalt des Blockes steigen (vgl. Anhang Darstellung 7). Dann wird der neue Block von der Instanz an Land mit der BC verbunden. Einer der Mitfahrer will dann die BC manipulieren und fordert etwas zurück. Da die BC *unveränderlich* und gegen äußere Angriffe geschützt ist, zeigen die folgenden Blöcke an, dass sie den Eingriff ablehnen und somit schlägt der Manipulationsversuch fehl steigen (vgl. Anhang Darstellung 8). [27]

6 Fazit

Bei der Literaturrecherche konnte ein langjähriger Bedarf zur Aufklärung der Funktionsweise der BC Technologie festgestellt werden. Es wird kritisiert, dass das Wissen über die BC größtenteils Entwicklern und Experten vorbehalten ist. Hier ist klar ein Forschungsbedarf zu identifizieren, für den diese Ausarbeitung einen Beitrag leisten soll. Denn obwohl es bereits viele unterschiedliche Herangehensweisen an eine Visualisierung gibt, fehlen noch einsteigerfreundliche Darstellungen, die potentiellen Nutzern die Grundlagen begreifbar machen. Dabei ist es eine Herausforderung, den komplexen BC

Aufbau strukturiert und verständlich zu vermitteln. Die in dieser Ausarbeitung gewonnenen Erkenntnisse und die Visualisierung sollen dabei helfen, BC Laien einen ersten Einblick in die Funktionsweise der BC zu gewähren. Der Prototyp für die Visualisierung kann als erste Iteration für eine Evaluation dienen, um seine Aussagekraft für Laien zu prüfen. Er soll anhand seiner Ausarbeitung dem Anspruch genügen, trotz seiner trivialen Darstellung ein tieferes Verständnis zu erzeugen. Hiermit soll ein Beitrag geschaffen werden, Menschen über die Funktionsweise der BC aufzuklären und die Adoption und das Verständnis für KW voranzutreiben.

Literaturverzeichnis

[1] Alshamsi, A. and Andras, Prof.P. 2019. *User perception of Bitcoin usability and security across novice users.* International Journal of Human-Computer Studies. DOI: https://doi.org/10.1016/j.ijhcs.2019.02.004

[2] Ammous, S. 2019. *Der Bitcoin-Standard: die dezentrale Alternative zum Zentralbankensystem.* Aprycčt. ISBN: 978-1-119-47386-2.

[3] Arslanian, H. et al. 2021. El Salvador's law: a meaningful test for Bitcoin. pwc.com: *https://www.pwc.com/gx/en/financial-services/pdf/el-salvadors-law-a-meaningful-test-for-bitcoin.pdf.*

[4] Baur, A. et al. 2015. *Cryptocurrencies as a Disruption? Empirical Findings on User Adoption and Future Potential of Bitcoin and Co*. https://www.doi.org/10.1007/978-3-319-25013-7_6

[5] Calmbach, M. et al. 2016. *Wie ticken Jugendliche 2016? Lebenswelten von Jugendlichen im Alter von 14 bis 17 Jahren in Deutschland*. Springer. https://doi.org/10.1007/978-3-658-12533-2.

[6] Froehlich, M. et al. 2021. *Don't Stop Me Now! Exploring Challenges Of First-Time Cryptocurrency Users*. Designing Interactive Systems Conference DOI: https://doi.org/10.1145/3461778.3462071.

[7] Gao, X. et al. 2016. *Of Two Minds, Multiple Addresses, and One Ledger: Characterizing Opinions, Knowledge, and Perceptions.* DOI: https://doi.org/10.1145/2858036.2858049.

[8] Gartner 2019 Hype Cycle for Blockchain: *https://www.gartner.com/en/newsroom/press-releases/2019-09-12-gartner-2019-hype-cycle-for-blockchain-business-shows*. Accessed: 04.05.2022.

[9] Ghosh, E. and Das, B. 2020. A Study on the Issue of Blockchain's Energy Consumption. *Proceedings of International Ethical Hacking Conference 2019*. Springer Singapore. DOI: https://doi.org/10.1007/978-981-15-0361-0_5.

[10] Glomann, L. et al. 2019. *Improving the Blockchain User Experience - An Approach to Address Blockchain Mass Adoption Issues from a Human-Centred Perspective*. DOI: https://doi.org/10.1007/978-3-030-20454-9_60

[11] Henke, M. 2017. *Understanding Blockchain - Gegenwart und Zukunft der Technologie hinter Bitcoin*. Universität der Künste Berlin.

[12] Hosp, J. 2018. *Blockchain 2.0 – einfach erklärt.* FinanzBuch Verlag. ISBN 978-3-95972-159-2.

[13] Hype Cycle Forschungsmethodik: *https://www.gartner.de/de/methoden/hype-cycle*. Accessed: 04.05.2022.

[14] Kahneman, D. 2017. *Schnelles Denken, langsames Denken*. Penguin Verlag. ISBN: 978-3-88680-886-1.

[15] Nakamoto, S. 2008. Bitcoin: Ein elektronisches Peer-to-Peer- Bezahlsystem: *https://bitcoin.org/de/bitcoin-paper*. Accessed: 04.05.2022

[16] Palka, S. and Wittpahl, V. 2018. *Vertrauen und Transparenz – Blockchain-Technologie als digitaler Vertrauenskatalysator*. ISBN: 978-3-89750-193–5.

[17] Poon, J. and Dryja, T. The Bitcoin Lightning Network: www.lightning.network/lightning-network-paper.pdf. Accessed: 04.05.2022.

[18] Satoshi Nakamoto: Wer steckt hinter dem geheimnisvollen Bitcoin-Erfinder? *https://www.focus.de/finanzen/boerse/kryptowaehrungen/krypto-mythos-satoshi-nakamoto-wer-steckt-hinter-dem-geheimnisvollen-bitcoin-erfinder_id_24468460.html*. Accessed: 04.05.2022.

[19] Schlatt, V. and Schweizer, A. 2016. Blockchain: Grundlagen, Anwendungen und Potentiale. Fraunhofer: *https://publica.fraunhofer.de/entities/publication/dc9322f0-3063-4792-b69a-b2006982d857/details*. Accessed: 04.05.2022

[20] Sixt, E. 2017. *Bitcoins und andere dezentrale Transaktionssysteme: Blockchains als Basis einer Kryptoökonomie*. Springer Fachmedien Wiesbaden. DOI: https://doi.org/10.1007/978-3-658-02844-2.

[21] The Cypherpunks: 2019. *https://nakamoto.com/the-cypherpunks/*. Accessed: 04.05.2022.

[22] Tovanich, N. et al. 2019. Visualization of Blockchain Data: A Systematic Review. *IEEE Transactions on Visualization and Computer Graphics*. PP, (Dec. 2019). DOI: https://doi.org/10.1109/TVCG.2019.2963018.

[23] Tufte, E.R. 2013. *The visual display of quantitative information*. Graphics Press. ISBN 978-1-930824-13-3.

[24] Vines, J. et al. 2012. The joy of cheques: trust, paper and eighty somethings. *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* https://doi.org/10.1145/2145204.2145229.

[25] Ware, C. 2008. *Visual thinking for design: active vision, attention visual queries, gist, visual skills, color, narrative, design*. Morgan Kaufmann Publishers, an imprint of Elsevier. ISBN: 978-0-12-370896-0.

[26] Wickert, F. 2020. *Evaluation zur Wahrnehmung von Kryptowährungen innerhalb der Generation Z mit Prognose für die Akzeptanz in der Zukunft*. BA Thesis: https://hdms.bsz-bw.de/frontdoor/index/index/docId/6732. Hochschule der Medien. Stuttgart. Accessed: 04.05.2022

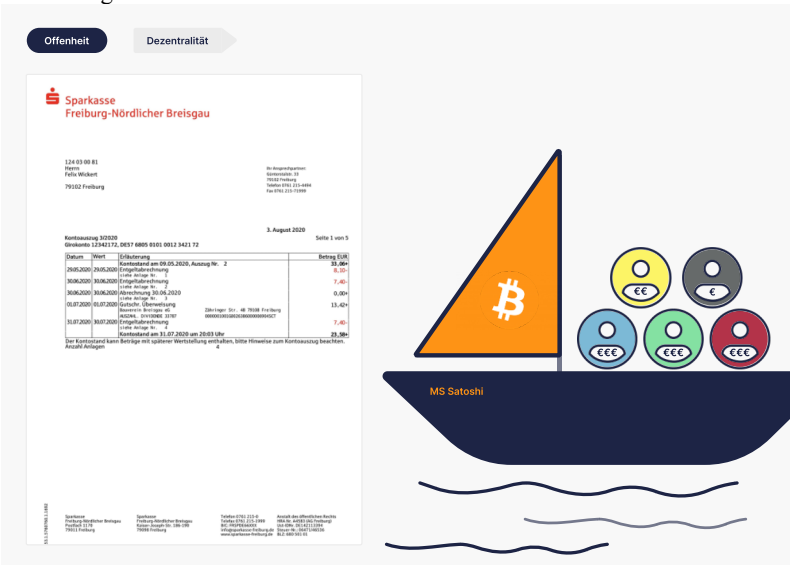
[27] Wickert, F. 2022. Blockchain Visualisierung. Figma: *https://www.figma.com/proto/ztrmvqoLsKd3bmbiQm4V76/WVK_Wickert_Blockchain_Vis?scaling=scale-down&page-id=0%3A1&starting-point-node-id=11%3A922&node-id=11%3A598*. Accessed: 04.05.2022.

[28] Zhong, Z. et al. 2020. SilkViser: *A Visual Explorer of Blockchain-based Cryptocurrency Transaction Data*. DOI:https://doi.org/10.1109/VAST50239.2020.00014

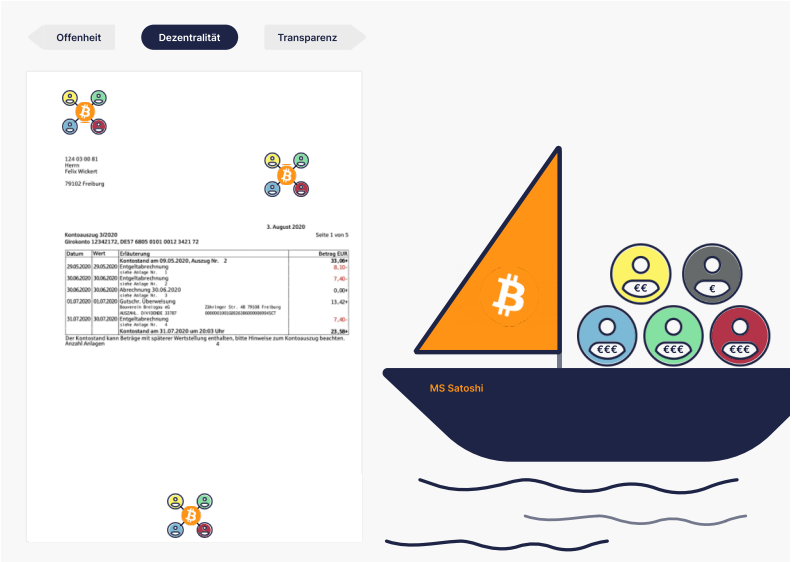
©2022 Felix Wickert. Lizenznehmer Hochschule Reutlingen, Deutschland. Dieser Artikel ist ein Open-Access-Artikel unter den Bedingungen und Konditionen der Creative Commons Attribution (CC BY)-Lizenz. http://creativecommons.org/licenses/by/4.0/

Anhang

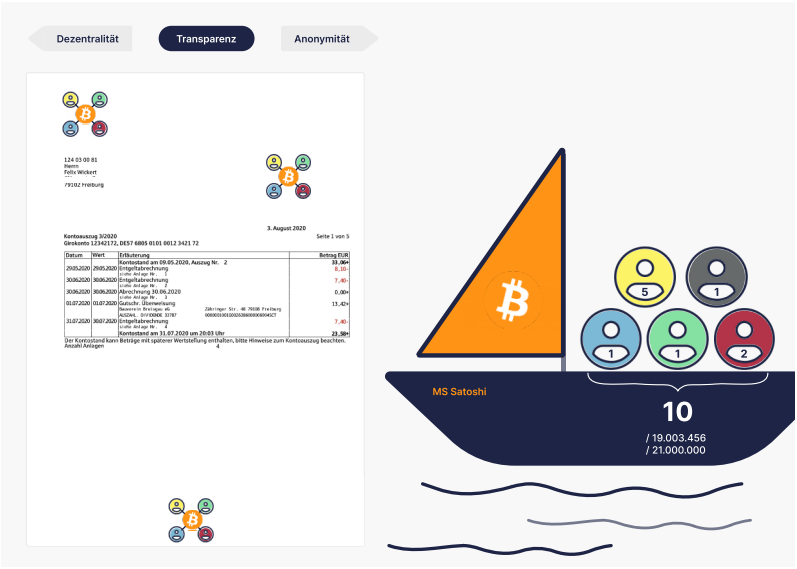
Darstellung 1: Offenheit



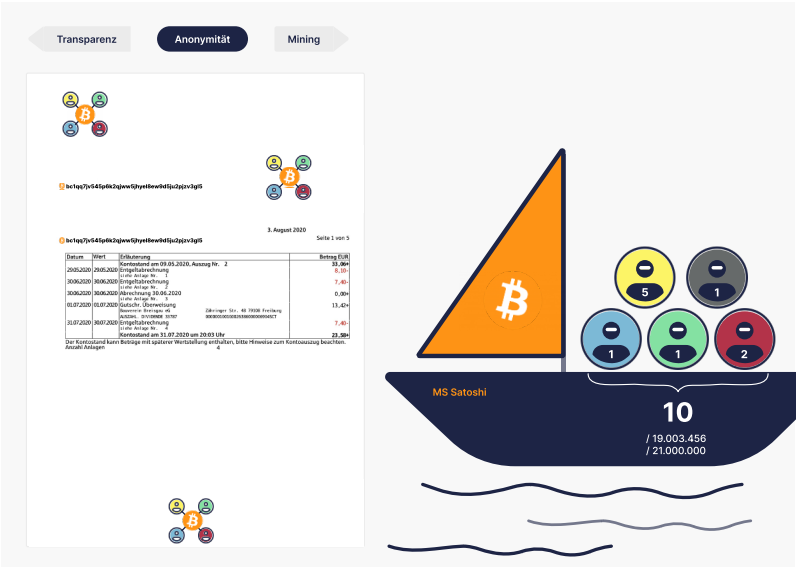
Darstellung 2: Dezentralität



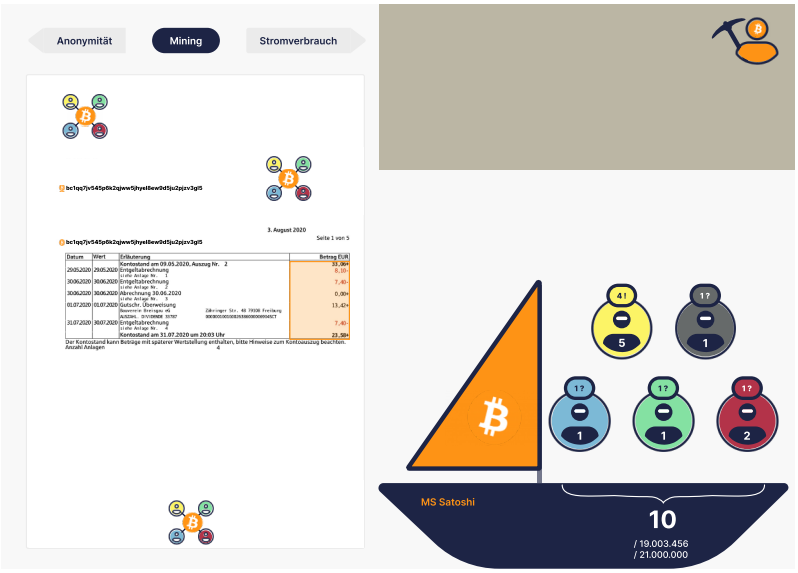
Darstellung 3: Transparenz



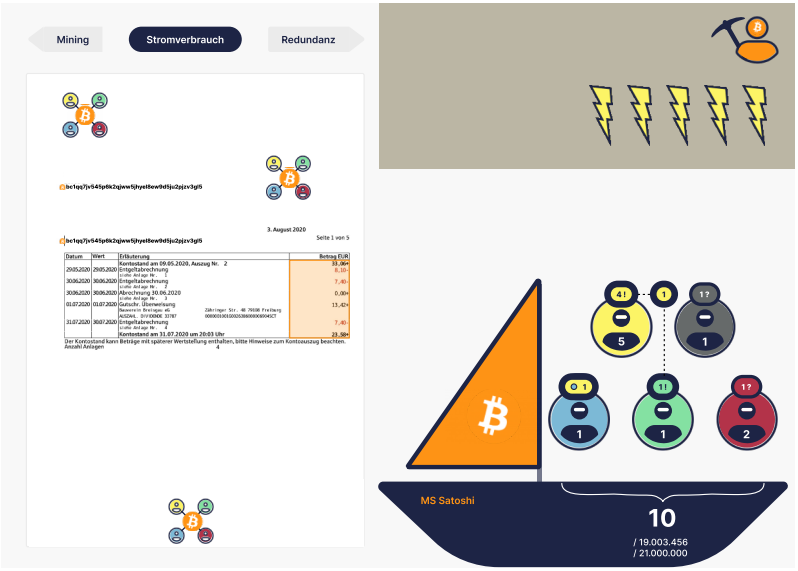
Darstellung 4: Anonymität



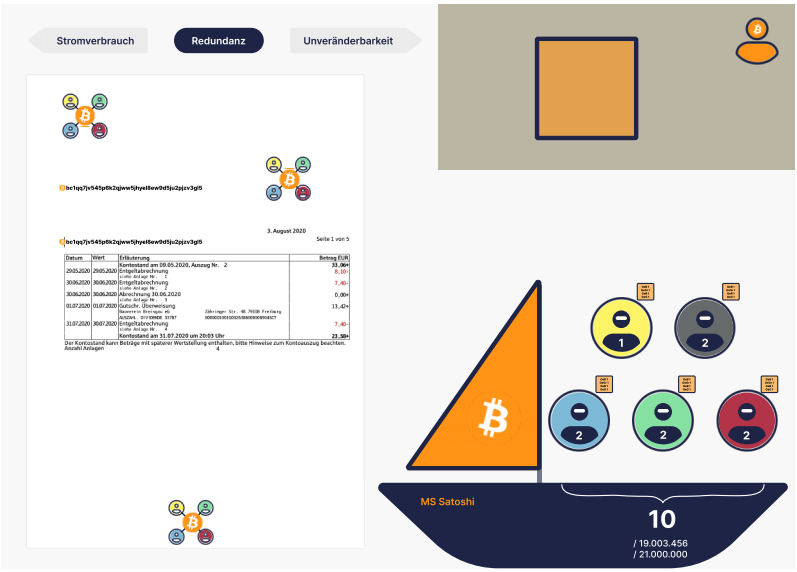
Darstellung 5: Mining



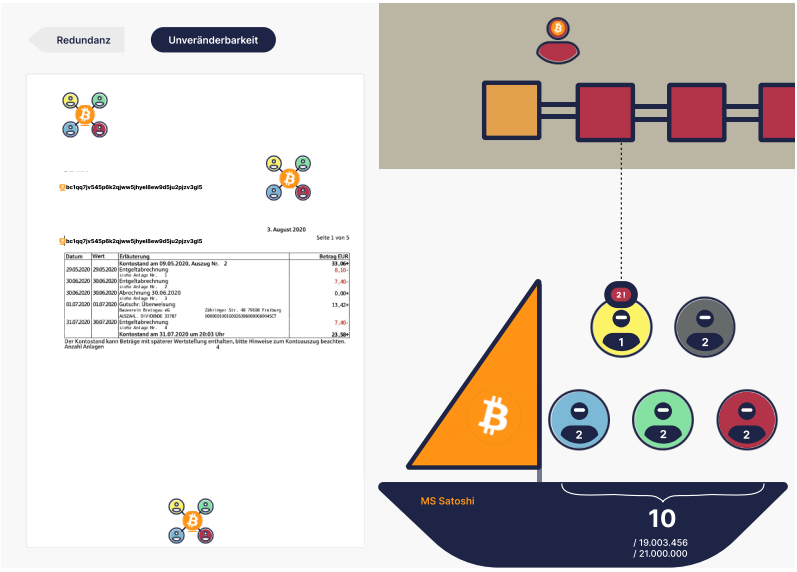
Darstellung 6: Stromverbrauch



Darstellung 7: Redundanz



Darstellung 8: Unveränderbarkeit



Methoden zur Optimierung von Path-Tracing in Echtzeitanwendungen

Dennis Grendatsch
Hochschule Reutlingen
Reutlingen, Deutschland
Dennis.Grendatsch@Student.Reutlingen-University.de

Zusammenfassung

Ziel der Arbeit ist aufzuzeigen, welche Methoden eingesetzt werden, um Path-Tracing für Echtzeitanwendungen (z.B. Videospiele) zu beschleunigen. Der Fokus dieser Arbeit liegt hierbei auf der Simulation von indirektem diffusen Licht. Als Erstes werden die Grundlagen von Path-Tracing vermittelt. Hierbei soll die Problematik hinter Path-Tracing verdeutlicht werden und welche Methoden seitdem entwickelt wurden, um dieser entgegenzuwirken. Als Nächstes wird, mithilfe einer Literaturreche, ein Überblick zu den aktuellen Methoden aufgezeigt, die Path-Tracing hinsichtlich der Geschwindigkeit und/oder der Varianz optimieren. Im letzten Schritt werden die Methoden untersucht, welche in der Praxis eingesetzt werden. Die Untersuchung zeigt, dass Methoden wie u.a. SVGF, Shading Level of Detail etc. Path-Tracing beschleunigen, wodurch eine Ausführung in Echtzeit unter gewissen Umständen durchaus möglich ist.

CCS Concepts

• **Computing methodologies → Ray-Tracing**.

Keywords

Path-Tracing, Ray-Tracing, Optimierung, Echtzeit, indirektes diffuses Licht

1 Einleitung

In diesem Kapitel wird als Erstes die Problematik hinter Echtzeit Ray-Tracing erläutert und es wird erklärt, weshalb Ray-Tracing für physikalisch plausible Ergebnisse so wichtig ist. Außerdem wird das Ziel dieser Ausarbeitung besprochen, um die genannte Problematik anzugehen.

1.1 Problemstellung

Heutige Videospiele sehen zunehmend realistischer aus. Mittlerweile kommen Videospiele an die Qualität von computeranimierten Kinofilmen heran. Jedoch existieren diverse Effekte, die in Videospielen bzw. interaktiven 3D-Anwendungen schwer darzustellen sind. Das liegt daran, dass für Videospiele seit jeher eine Technik namens Rasterisierung für das Rendering verwendet wird, [9, S. 5] welche Vor- und Nachteile hat. Ein Vorteil ist bspw., dass der Grafikprozessor (GPU) für diese Aufgabe prädestiniert ist, [8, S. 22] und daher schnelle Ergebnisse erzeugen kann. Ein Nachteil ist allerdings, dass die Primitiven (z.B. *triangles*) individuell erzeugt werden und somit kein Zusammenhang zwischen diesen hergestellt wird [2, S. 438]. Daher kann in der Rasterisierung nur direktes Licht (s. Abbildung 1, Abschnitt (1)) berücksichtigt werden, da indirektes Licht (s. Abschnitt (2)) die Informationen der umliegenden Objekte benötigt. Für die Simulation von spiegelnden Oberflächen (z.B. das Lack eines

Betreuer/-in Hochschule: Prof. Dr. Uwe Kloos, Hochschule Reutlingen, Uwe.Kloos@Reutlingen-University.de

Informatics Inside Frühjahr 2022
11. Mai 2022, Hochschule Reutlingen
Copyright 2022 Dennis Grendatsch

Autos) stellt dies ein großes Problem dar. Mit einigen Techniken, wie z.B. *screen space reflections*, können zwar Reflexionen in der Rasterisierung dennoch simuliert werden, aber mit Bildartefakten. [9, S. XVI] Ray-Tracing ist in der Lage, neben direktem Licht, auch indirektes Licht zu simulieren, wodurch eine globale Beleuchtung (engl. *global illumination*) möglich ist (s. Abschnitt (3)) [12, S. 9]. Andererseits benötigt der Rendering-Algorithmus eine hohe Rechenleistung, wodurch eine Ausführung in Echtzeit sehr anspruchsvoll sein kann. Die Veröffentlichung der Demo „Star Wars: Reflections“ ab Anfang August 2018 seitens Nvidia hat jedoch gezeigt, dass Ray-Tracing in Echtzeit ausgeführt werden kann[1].

1.2 Begründung der Relevanz

Die globale Beleuchtung kann einen erheblichen Beitrag zum Realismus eines gerenderten Bildes leisten [9, S. 519]. Wie bereits erwähnt, kann Ray-Tracing globale Effekte erzielen. Allerdings ist der Algorithmus sehr anspruchsvoll, weshalb Kompromisse eingegangen werden müssen, um die Berechnungen zu beschleunigen. Ein Kompromiss ist bspw. die Reduzierung der Abtastrate je Pixel (engl. *samples per pixel* (SPP)). Jedoch verstärkt eine Reduzierung der SPP das Rauschen bzw. die Varianz in dem endgültigen Ergebnis. [2, S. 445] Daher werden Techniken benötigt, um die Varianz zu verringern.

1.3 Vorgehen

Ziel ist die Ermittlung von Methoden, die Path-Tracing beschleunigen, um eine Ausführung in Echtzeit-Anwendungen (z.B. in Videospiele) zu ermöglichen. Der Fokus dieser Arbeit wird hierbei auf der Simulation von indirektem diffusem Licht liegen.

Als Erstes (Kapitel 2) werden die Fortschritte von Ray-Tracing chronologisch von 1968 bis heute erläutert. Als nächstes (Kapitel 3) werden die Grundlagen von Path-Tracing erklärt, um ein Grundverständnis bzgl. der Varianz Problematik und der Leistungsanforderung zu entwickeln. Darauffolgend (Kapitel 4) wird der aktuellste Stand der Forschung präsentiert, um Path-Tracing hinsichtlich der Geschwindigkeit und/oder der Varianz zu optimieren. Anschließend (Kapitel 5) werden ausgewählte Methoden besprochen, die in Minecraft RTX genutzt werden. Am Ende (Kapitel 6) erfolgt ein abschließendes Fazit.

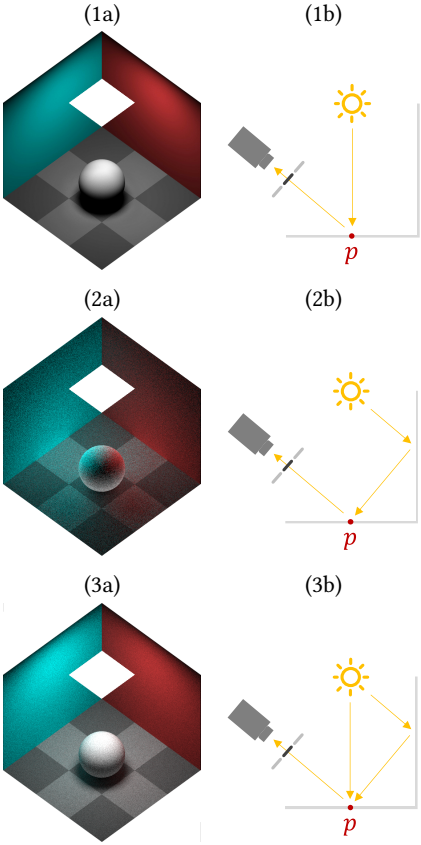


Abbildung 1: (1) direktes Licht, (2) indirektes Licht, (3) globale Beleuchtung.

[1] Nvidia Blog: https://blogs.nvidia.com/blog/2020/08/25/rtx-real-time-ray-tracing/

2 Geschichte hinter Ray-Tracing

Ray-Tracing ist ein Rendering Algorithmus, der heutzutage fotorealistische Bilder generieren kann, die kaum von der Realität unterscheidbar sind (s. Abbildung 2, Abschnitt (d)) [16, S. 7]. In diesem Kapitel werden die Fortschritte in Ray-Tracing chronologisch von 1968 (s. Abschnitt (a)) bis heute (s. Abschnitt (d)) erläutert.

2.1 Ray-Tracing (1968)

Im Jahr 1968 wurde der erste Ray-Tracing Algorithmus von Arthur Appel [3] vorgestellt [7, S. 530]. Ziel ist es, die 3D Szene auf einem 2D Bild abzubilden. Dazu wird von der Kamera aus durch jeden Bildpixel ein Strahl (engl. *ray*) geschossen. Dieser Strahl wird in der Fachliteratur als Primärstrahl (engl. *primary ray*) bezeichnet. Der Strahl ermittelt dann, mit welchen Objekten ein Schnitt- bzw. Treffpunkt p vorliegt. Der Treffpunkt, der am nächsten zur Kamera liegt, wird ausgewählt. Als nächstes wird die Objektfarbe, anhand von dem Treffpunkt, ermittelt. [7, S. 530] Neben der Objektfarbe, kann zusätzlich ermittelt werden, ob der Treffpunkt p von einer Lichtquelle beleuchtet wird. Dazu wird ein Schattenstrahl (engl. *shadow ray*) von dem Schnittpunkt aus in Richtung Lichtquelle geschickt (s. Abbildung 2, Abschnitt (a)). Falls der Strahl, währenddessen etwas schneidet, liegt der Oberflächenpunkt p im Schatten, daher der Name Schattenstrahl. [9, S. 9]

2.2 Classical Ray-Tracing

Im Jahr 1979 hat Turner Whitted [23] am Konzept von Appel weitergearbeitet und diese erweitert. Die angepasste Variante konnte zusätzliche Effekte, wie Reflexionen und Brechungen (z.B. durch Glas) simulieren (s. Abbildung 2, Abschnitt (b)). [7, S. 531] Dazu werden weitere Sekundärstrahlen[2] am jeweiligen Schnittpunkt erzeugt. Die Anzahl an Sekundärstrahlen ist von dem Material abhängig. Bei einer spiegelnden Oberfläche sind es bspw. zwei Sekundärstrahlen, jeweils für die Schattierung und Reflexion. Die Sekundärstrahlen erzeugen neue Schnittpunkte, die bei der Ermittlung der endgültigen Pixel Farbe mitberücksichtigt werden.

2.3 Stochastic/Distributed Ray-Tracing

Bisher konnten Effekte wie Schatten, Reflexionen und Brechungen mit Ray-Tracing simuliert werden. Diese Effekte waren bisher gestochen scharf (s. Abbildung 2, Abschnitt (b)), jedoch können diese Effekte in der realen Welt auch unscharf sein. Eine Reflexion wirkt bspw. weicher, je rauer die Oberfläche (z.B. gebürsteter Stahl) ist. Um diese Probleme zu lösen, hat Cook et al. [5] im Jahr 1984 die Methode *stochastic/distributed ray-tracing* präsentiert [9, S. 8]. Dafür werden am Schnittpunkt mehrere Sekundärstrahlen, anstelle von einem Sekundärstrahl, erzeugt (s. Abschnitt (c)).

2.4 Path-Tracing

Im Jahr 1986 hat Jim Kajiya [10] die *rendering equation* und damit Path-Tracing vorgestellt. Mit Path-Tracing können Effekte wie indirektes diffuses Licht, Kaustik etc. simuliert werden, da Path-Tracing das komplexe Verhalten von Licht simuliert. [16, S. 47] Verwirrung kann der Begriff Path-Tracing stiften, da in den vorherigen Kapiteln stets von Ray-Tracing die Rede war. Bisher wurde an jedem Schnitt- bzw. Auftreffpunkt mind. ein Sekundärstrahl erzeugt. In Path-Tracing ist es je Schnittpunkt höchstens ein Sekundärstrahl. Mit dieser Besonderheit entsteht

[2] Im weiteren Verlauf dieser Arbeit wird der neu erzeugte Strahl am Treffpunkt p als Sekundärstrahl (engl. *secondary ray*) verallgemeinert [7, S. 531].

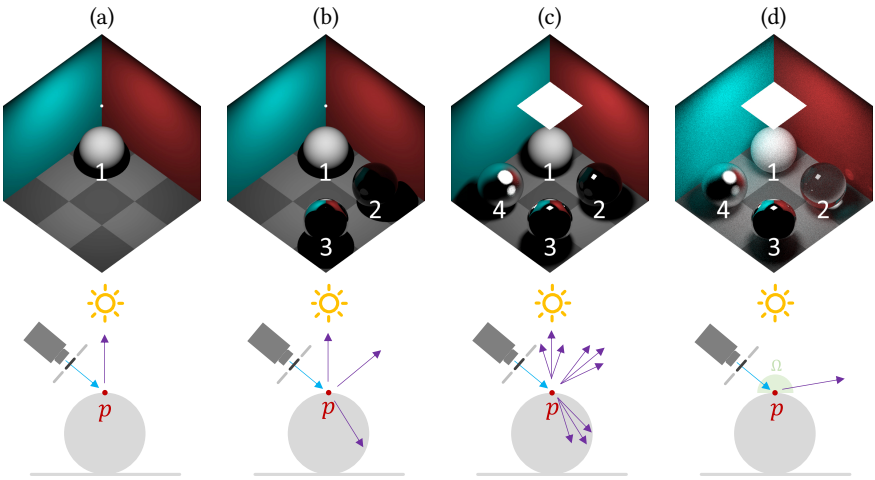


Abbildung 2: Darstellung verschiedener globaler Effekte auf Kugeln. (1) diffus, (2) gläsern, (3) metallisch, (4) diffus-metallisch.

am Ende ein einziger Lichtpfad. [10] Der Lichtpfad wird mit der Rendering-Gleichung (s. Formel 1) erzeugt. Eine ausführliche Erklärung zum Path-Tracing Algorithmus erfolgt in Kapitel 3.

3 Path-Tracing Grundlagen

In diesem Kapitel wird der Path-Tracing Algorithmus detailliert erläutert, indem u.a. die Rendering-Gleichung und die dazugehörige BRDF (*bidirectional reflectance distribution function*) erklärt wird, die zur Ermittlung der Streuung benötigt wird. Außerdem wird die Monte-Carlo-Integration behandelt, die zur Lösung der Rendering-Gleichung benötigt wird. Des Weiteren wird die Funktionsweise von dem Path-Tracing Algorithmus veranschaulicht sowie das größte Problem von Path-Tracing angesprochen, die Varianz. Zuletzt werden bisherige Optimierungen besprochen, die Path-Tracing hinsichtlich der Geschwindigkeit und/oder der Varianz verbessern.

3.1 Rendering-Gleichung

Mit der Rendering-Gleichung (engl. *rendering equation*) kann das komplexe Verhalten von Licht simuliert werden [16, S. 47]. Eine Variante der Rendering-Gleichung [2, S. 437] von Kajiya [10] ist in Formel 1 abgebildet und wird im Folgenden ausführlich erklärt.

$$L_0(p,v) = L_e(p,v) + \int_{l\in\Omega} f(l,v) L_0(r(p,l),-l)(n\cdot l)^+ dl \quad (1)$$

Mit der Kamera wird der Oberflächenpunkt p über den Vektor $-v$ betrachtet (s. Abbildung 3). Mit der Rendering-Gleichung bzw. $L_0(p,v)$ kann ermittelt werden, wie hoch die Beleuchtungsstärke (engl. *irradiance*) von p ist. $L_e(p,v)$ sagt aus, wie viel Licht p emittiert [2, S. 437]. Eine Lampe ist bspw. ein emittierendes Objekt. Falls p Licht emittiert, entfällt die Berechnung von dem Integral. Als nächstes werden die Komponenten in dem Integral erläutert. $f(l,v)$ ist die BRDF und wird in Kapitel 3.2 erläutert. In $L_0(r(p,l),-l)$ wird die Energie von dem Lichtstrahl $-l$ berechnet,

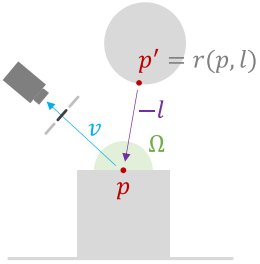


Abbildung 3: Visualisierung der Gleichung $L_0(r(p,l),-l)$, die Bestandteil der Rendering-Gleichung (s. Formel 1) ist. Quelle: In Anlehnung an [2, S. 438]

der von p' auf p trifft (s. Abbildung 3) [2, S. 437f.]. In $(n \cdot l)^+ dl$ wird berechnet, in welchem Winkel der Lichtstrahl $-l$ auf p trifft. Dazu wird der Winkel zwischen den Vektoren n (ist der Normalenvektor von p) und l berechnet. Der Punkt p wird bspw. stärker beleuchtet, wenn der Lichtstrahl senkrecht auf die Oberfläche trifft. [6, S. 384] Das Integral berücksichtigt jeden Lichtstrahl $-l$, der auf p trifft. Die Hemisphäre Ω repräsentiert die Richtungen, aus denen die Lichtstrahlen einfallen können (s. Abbildung 3). Das gesamte einfallende Licht $-l$ wird mit dem Integral aufsummiert.

3.2 Bidirectional Reflectance Distribution Function

Die BRDF (*bidirectional reflectance distribution function*) gibt an, wie eingehende Lichtstrahlen, die auf Punkt p treffen, reflektiert bzw. gestreut werden [17, Kap. 8].

Die BRDF $f(l,v)$ in der Rendering-Gleichung nimmt als Parameter den ausgehenden Lichtstrahl l und den ausgehenden Blickstrahl v. Der eingehende Lichtstrahl $-l$ wird, je nach Einfallswinkel, perfekt reflektiert. Der reflektierte ausgehende Lichtstrahl wird als l' bezeichnet (s. Abbildung 4, Abschnitt (1)). Wenn v mit l' übereinstimmt, liefert die BRDF den Wert 1 (s. Abschnitt (1a)), anderenfalls den Wert 0 (s. Abschnitt (1b)) [6, Kap. BRDF]. In Abschnitt (2a) liefert die BRDF z.B. den Wert 0.6, da die Intensität von dem mittleren Strahl (von l') aufgrund der Streuung geringer ist.

Für eine diffuse Reflexion kann bspw. die Lambertian BRDF genutzt werden, da das Licht in allen möglichen Richtungen innerhalb der Hemisphäre gestreut wird [2, S. 497]. Für eine spiegelnde Reflexion kann z.B. die Cook-Torrance BRDF genutzt werden. Weitere Informationen zur Cook-Torrance BRDF können in Vries 2020 [6], unter dem Kapitel 'BRDF' nachgelesen werden.

3.3 Monte-Carlo-Integration

Die Rendering-Gleichung (s. Formel 1) kann in ihrer jetzigen Form nicht gelöst werden. Es können nicht alle Lichtstrahlen der Hemisphäre Ω berücksichtigt werden, die auf Punkt p auftreffen, da dafür unendlich viele Möglichkeiten existieren. Aus diesem Grund wird die Monte-Carlo-Integration verwendet, die anhand einer Stichprobe N eine gute Annäherung zum Endresultat liefert. Es ist

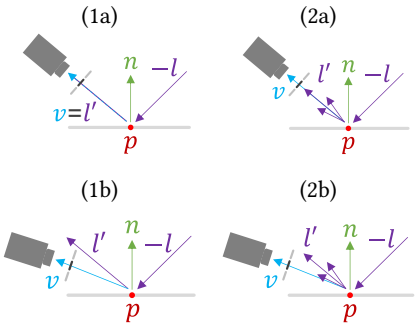


Abbildung 4: Je nach Oberflächenbeschaffenheit von p wird der eingehende Lichtstrahl unterschiedlich gestreut. In Abschnitt (1) ist die Oberfläche spiegelnd und in Abschnitt (2) glänzend (engl. *glossy*).

wichtig anzumerken, dass die Stichprobenauswahl in der Monte-Carlo-Integration zufällig erfolgt. [17, Kap. 13] Das heißt, dass bei einer Stichprobenmenge von bspw. $N = 20$ die Beleuchtungsstärke von p anhand 20 zufälliger einfallender Lichtstrahlen $-l$ ermittelt wird. Je höher die Stichprobenmenge N ist, desto besser ist die Annäherung an das Integral der Rendering-Gleichung. Die Formel zur Monte-Carlo-Integration [17, Kap. 13.2] ist in Formel 2 abgebildet.

$$F_N = \frac{1}{N} \sum_{i=1}^{N} \frac{f(X_i)}{p(X_i)} \tag{2}$$

In dem Integrand $f(X_i)$ wird $f(l, v)L_0(r(p, l), -l)(n \cdot l)^+ dl$ von der Rendering-Gleichung ausgerechnet [12, S. 12]. Mit der *probability distributed function* (PDF) $p(X_i)$ kann der Abtastbereich der Hemisphäre Ω beeinflusst werden, [17, Kap. 13.2] da einige einfallende Lichtstrahlen mehr zur Beleuchtung von p beitragen, als andere Lichstrahlen. Daher ist es wichtig, dass $p(X_i)$ dem Integrand $f(X_i)$ möglichst nahe kommt, da eine unpassende PDF das Ergebnis der Monte-Carlo-Integration sogar verschlechtern kann. [17, Kap. 13.10] Die einfachste PDF ist *uniform sampling*. Für diffuse Oberflächen ist *uniform sampling* bzw. eine einheitliche Abtastung der Hemisphäre durchaus sinnvoll, aber nicht für spiegelnde Oberflächen, die eingehende Lichtstrahlen in eine bestimmte Richtung lenken. [12, S. 15] Das heißt, dass mehr Stichproben N benötigt werden, um zufriedenstellende Reflexionen zu erhalten. Eine weitere PDF-Variante ist *light sampling*, in der gezielt die Lichtquelle abgetastet wird [21, Kap. 9].

3.4 Algorithmus Ablauf

Die Folgende Erklärung bezieht sich auf Abbildung 5. Zur Berechnung von einem Bildpixel, wird ein Primärstrahl von der Kamera aus durch diesen zu berechnenden Pixel geschossen. Bei einem Schnittpunkt wird gemäß der PDF *uniform sampling* ein zufälliger Bereich der Hemisphäre abgetastet. Um die Lichtenergie aus der abgetasteten Richtung zu bestimmen, wird ein Sekundärstrahl von p in die abgetastete Richtung geschossen. Punkt p' ist der Schnittpunkt mit dem Sekundärstrahl. Um die Energie in Punkt p‘ zu berechnen, wird die Rendering-Gleichung erneut angewandt. Dieser Prozess wird so lange wiederholt, bis entweder die maximale Anzahl an Lichtsprüngen (engl. *bounces*) erreicht wurde, oder ein emittierendes Objekt (z.B. eine Lampe) getroffen wurde. Die maximale Anzahl an Lichtsprüngen ist dabei frei wählbar. Eine hohe Zahl erhöht zwar die Wahrscheinlichkeit eine Lichtquelle zu treffen, verlangsamt jedoch die Berechnung der Beleuchtungsstärke von p. Falls kein emittierendes Objekt nach der maximalen Anzahl an Lichtsprüngen erreicht wurde, liefert der Lichtpfad keine Energie bzw. den Farbwert (Rot: 0, Grün: 0, Blau: 0).

Dieser Algorithmus wird für jeden Bildpixel separat ausgeführt. Die Stichprobenmenge N beeinflusst dabei die Anzahl an Ausführungen je Pixel. Bei einer Stichprobenmenge von $N = 1$ entsteht ein 1 SPP Rendering, bei $N = 100$ ein 100 SPP Rendering usw. Die Komplexität einer Szene beeinflusst die Anzahl an Schnittpunkt Überprüfungen je Strahl [20, Kap. 3]. Folglich beeinflusst die Stichprobenmenge N, die Bildauflösung und/oder die Komplexität einer Szene die endgültige Renderzeit [9, S. 23].

3.5 Varianz

In Kapitel 3.3 wurde erwähnt, dass mehr Stichproben N zu einer besseren Annäherung der Rendering-Gleichung (s. Formel 1) führen. Bei einer geringen Anzahl an Stichproben hat das Endresultat eine hohe Varianz [2, S. 445] und wirkt stark verrauscht. Der Grund ist, dass jeder Pixel gesondert berechnet wird und manche Lichtstrahlen, wie in Kapitel 3.4 erläutert wurde, am Ende kein emittierendes

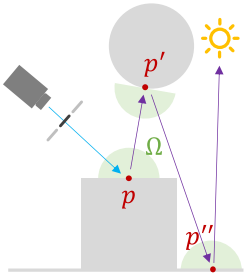


Abbildung 5: Ein beispielhafter Lichtpfad zur Berechnung der Beleuchtungsstärke (engl. *irradiance*) von p.

Objekt treffen und somit keine Lichtenergie zum Pixel transportieren. Eine Verringerung der Anzahl der Stichproben führt zwar zu schnelleren Ergebnissen, verursacht jedoch eine hohe Varianz.

3.6 Bisherige Optimierungen

In diesem Kapitel wird das Verfahren *importance sampling* besprochen, um die Varianz insgesamt zu reduzieren. Es existieren weitere Methoden wie bspw. *multiple importance sampling*, *russian roulette*, *bounding volume hierarchy*, *bidirectional path-tracing* etc., allerdings werden diese Techniken in dieser Ausarbeitung nicht behandelt. Auf Hardwarebeschleunigtes Ray-Tracing wird ebenfalls nicht näher eingegangen.

Importance sampling verringert die Varianz in dem Rendering. Dafür wird die *probability distributed function* (PDF) benötigt, die in Kapitel 3.3 erwähnt wurde. Für diffuse Oberflächen ist bspw. *uniform sampling* geeignet, jedoch könnte eine andere PDF die Abtastung zusätzlich optimieren, da senkrecht auftreffende Lichtstrahlen mehr zur Beleuchtung von p beitragen, als schräg auftreffende Lichtstrahlen [13, S. 315]. Diese Herangehensweise wird als *cos-weighted sampling* bezeichnet. [2, S. 451] Der visuelle Unterschied ist zwar sehr marginal, aber im Detail dennoch erkennbar. Bei spiegelnden Oberflächen ist der Unterschied zwischen *uniform sampling* und einer dafür angepassten PDF größer. [12, S. 15]

4 Aktueller Stand der Forschung

In diesem Kapitel wird der aktuellste Stand der Forschung präsentiert, um Path-Tracing hinsichtlich der Geschwindigkeit und/oder der Varianz zu optimieren. Im Folgenden sind wissenschaftliche Publikationen aufgelistet, die ihren Beitrag dazu leisten.

Müller et al. (2021) [15] fokussieren sich darauf, ein rauschärmeres Bild zu erzeugen und das bei gleichbleibender Anzahl an SPP. Dazu werden die Strahlen (engl. *radiances*) mit Hilfe eines neuronalen Netzes zwischen gespeichert, wodurch spätere Berechnungen wegfallen sollen.

Stengel et al. (2021) [22] verteilen den Rechenaufwand auf zwei Systeme. Der Server berechnet das indirekte Licht und das lokale System die Sichtbarkeit von dem primären Strahl. Damit soll Ray-Tracing auch auf schwächerer Hardware ausgeführt werden können.

Schied et al. (2018) [19] schlagen eine Methode vor, die Artefakte, wie bspw. *ghosting*, verhindern soll, falls zeitliche Informationen genutzt werden, um das Rauschen zu entfernen. Dazu sollen frühzeitige Veränderungen auf darauffolgenden Bildern detektieren werden.

Koskela et al. (2019) [11] verfolgen das Ziel mit Regression basierten Rausch Filterung, das zuvor nur im Offline Bereich aufgrund der Rechenleistung verwendet wurde, 1 SPP Bilder zu entrauschen. Dabei soll diese Methode u.a., verglichen mit vorherigen Path-Tracing Rekonstruktionsmethoden, bis zu 1.8x schneller sein.

Meister et al. (2020) [14] möchten mit der Sortierung von Strahlen die Kohärenz erhöhen, um die Auffindung von Schnittpunkten im

Trace Kernel zu beschleunigen. Path-Tracing erzeugt zunehmend inkohärente Strahlen, aufgrund der Streuung von diffusen Oberflächen. Grund ist, dass die Verfolgung von inkohärenten Strahlen mehr Ressourcen als die Verfolgung von kohärenten Strahlen benötigen.

Yang et al. (2020) [24] schlagen vor die Menge an erzeugten Strahlen, abhängig von der Szene, variabel zu halten und notfalls zu minimieren. Dadurch soll eine FPS Erhöhung um 30% ermöglicht werden. Jedoch kann diese Variante zu einem minimalen Qualitätsverlust führen.

5 Methoden in der Praxis

Jakub Boksansky und Oli Wright haben in der Game Developers Conference (GDC) 2020 Einblicke in die Path Tracer Implementierung von Minecraft RTX ermöglicht [4]. Die Präsentation besteht aus zwei Teilen und deckt folgende Themen ab: Der erste Teil liefert eine Übersicht zum Renderer, Details zu den Materialien (diffus, spiegelnd etc.) und Informationen zum Denoising Prozess. Teil zwei beschäftigt sich mit der diffusen Beleuchtung, *shading level of detail* und dem *irradiance cache*.

In diesem Kapitel wird *irradiance cache* erläutert, um die Berechnungen der Oberflächenpunkte zu beschleunigen. Im Anschluss wird *shading level of detail* besprochen, um den Berechnungsaufwand bei der Schattierung anzupassen. Als letztes wird der SVGF (*spatiotemporal variance-guided filtering*) Algorithmus erklärt, der 1 SPP Renderings in Echtzeit entrauscht. Die anderen Themen der Präsentation werden im Rahmen dieser Arbeit nicht behandelt.

5.1 Irradiance Cache

Die Methode *irradiance cache* beschleunigt die Berechnung von Oberflächenpunkten p_i, die eine diffuse Oberfläche besitzen. Die Beschleunigung erfolgt dadurch, dass nur ein Teil der sichtbaren Oberflächenpunkte berechnet wird. Die jeweilige berechnete Beleuchtungsstärke (engl. *irradiance*) wird als *record* zwischengespeichert. Die fehlenden Oberflächenpunkte werden anhand der zuvor berechneten Oberflächenpunkte interpoliert. [12, S. 24ff.] Das interpolieren fehlender Werte funktioniert, da diffuses Licht kaum Details, im Gegensatz zu spiegelnde Oberflächen, besitzt [12, S. 17]. Für spiegelnde Oberflächen ist hingegen der *radiance cache* geeignet [12, S. 62].

Als nächstes wird die Funktionsweiße von dem *irradiance cache* erläutert. Falls in der Nähe von dem Oberflächenpunkt p nicht mindestens ein *record* vorliegt, wird die Beleuchtungsstärke von p mit Hilfe der Monte-Carlo-Integration ermittelt und am Ende ebenfalls als *record* zwischengespeichert [12, S. 25]. Um zu bestimmen, ob p interpoliert oder mit der Monte-Carlo-Integration berechnet wird, hängt von dem Fehler ab, die eine Interpolation verursachen könnte. Ziel ist es, einen

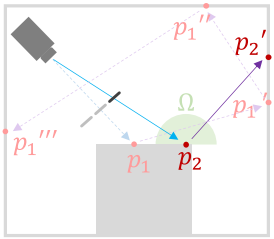


Abbildung 6: Beispielhafte Darstellung des Prinzips hinter *recursive radiance caching*. Ziel ist es die Beleuchtungsstärke von p_2 zu ermitteln. Das Prime-Zeichen repräsentiert die jeweilige Rekursionstiefe. In p_2 wird die Monte-Carlo-Integration durchgeführt, da der Interpolierungsfehler mit p_1 zu hoch ist. Bei p_2‘ hingegen wird der Wert anhand von dem *record* p_1‘ interpoliert, da der zulässige Fehler in der zweiten Rekursionstiefe höher ist.

konstanten Fehler in der gesamten Szene zu erreichen. [12, S. 26] Die Formel zur Ermittlung des Interpolationsfehlers wird in dieser Arbeit nicht behandelt. Weitere Informationen sind in dem Werk von Krivanek, Gautron (2009) [12] unter dem Kapitel 2.2.1 ‚Interpolation' zu finden.

Es existiert auch eine rekursive Variante der *irradiane cache* Methode, die als *recursive irradiance cache* bezeichnet wird. Diese Variante optimiert die Berechnung der sekundären Lichtstrahlen, die für die Simulation von indirektem Licht benötigt werden. [12, S. 64] Es wird angenommen, dass die rekursive Variante in Minecraft RTX verwendet wird. In dem nächsten Absatz erfolgt eine Erklärung zu *recursive irradiance cache*.

Mit zunehmender Rekursionstiefe verringert sich der Beitrag zur Beleuchtung von dem Oberflächenpunkt p. Anders gesagt, trägt die Rekursionstiefe 1 mehr zur indirekten Beleuchtung von p bei, als die Rekursionstiefe 3. Diese Tatsache nutzt *recursive irradiance cache*, indem der konstante Fehler bei jeder Rekursionstiefe angehoben wird und zudem die Menge an neu erzeugten Sekundärstrahlen am jeweiligen Treffpunkt verringert wird. Jede Rekursionstiefe besitzt dabei einen eigenen *irradiance cache*, in denen die *records* gespeichert werden. [12, S. 64] In Abbildung 6 wird das Prinzip veranschaulicht.

5.2 Shading Level of Detail

Mit der Technik *shading level of detail* werden die Punkte p_i, die von der Kamera aus zu sehen sind, entweder mit einem hohen oder mit einem niedrigen Detailgrad schattiert. Der Detailgrad beeinflusst dabei die Menge an Berechnungen, die für die Schattierung notwendig sind. [4] In Kapitel 5.2.1 wird erläutert, wie der Schattierungslevel von einem Punkt p ermittelt wird. In Kapitel 5.2.2 und 5.2.3 werden die Unterschiede beider Schattierungslevels erläutert.

5.2.1 Schattierungslevel bestimmen. Mit einem *ray cone* (dt. Strahlenkegel) kann der Schattierungslevel für einen Oberflächenpunkt p ermittelt werden. Die Oberflächenbeschaffenheit von p beeinflusst dabei den *ray cone* Öffnungswinkel. [4] In diesem Kapitel wird angenommen, dass p eine diffuse Oberfläche ist.

Im Folgenden wird die Funktionsweise erläutert. Der Primärstrahl wird als *ray cone* mit einem konstanten Öffnungswinkel in Richtung p geschossen. Am Treffpunkt bzw. p wird ein Sekundärstrahl erzeugt, das ebenfalls ein *ray cone* ist und ebenfalls einen Treffpunkt liefert. Dieser Treffpunkt hinterlässt einen *footprint* (dt. Fußabdruck), womit der Schattierungslevel bestimmt werden kann. Je größer die Fläche von dem *footprint* ist, desto wahrscheinlicher ist ein niedriger Schattierungslevel. Der Grund dafür ist, dass p seine Beleuchtungsstärke (engl. *irradiance*) von vielen Quellen bzw. eingehenden Lichtstrahlen erhält (s. die *footprint* Fläche) und daher diese in p gemittelt werden. Bei einem kleineren *footprint* muss hingegen ein höherer Detailgrad verwendet werden. Anstelle von dem *footprint* kann auch die Länge von dem Strahlenpfad (von der Kamera bis zum endgültigen Treffpunkt) verwendet werden, da

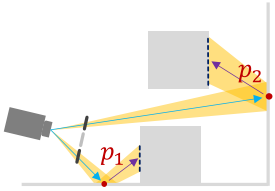


Abbildung 7: Die Ermittlung von dem Schattierungslevel. Bei p_2 wird aufgrund der Entfernung oder der footprint Größe ein *low detail shader* verwendet. Bei p_1 hingegen ein *high detail shader*. Quelle: In Anlehnung an: [1]

die Fläche von dem *footprint* mit zunehmender Entfernung steigt. [4, 1] In Abbildung 7 ist ein dazugehöriges Beispiel.

Als nächstes erfolgt eine Erklärung der beiden Schattierungslevel *low-* und *high detail* in Kapitel 5.2.2 und 5.2.3.

5.2.2 High Detail Shading. Die *high detail shading* Variante liefert die beste Qualität, ist allerdings aufwändiger zu berechnen [4]. Zur Berechnung der Beleuchtungsstärke von p wird folgendermaßen vorgegangen: Die indirekten Lichtstrahlen werden mit der *irradiance cache* Methode ermittelt. Die *records* befinden sich in den Vertices der Objekte und beinhalten die *irradiance* von eintreffendem Licht (s. $L_0(r(p,l),-l)$ in Formel 1). Das direkte Licht wird mit einem Schattenstrahl ermittelt. Falls das Objekt Licht emittiert, wird dies auch berücksichtigt. [4]

5.2.3 Low Detail Shading. Die low detail shading Variante liefert schnellere Resultate, allerdings mit Verlusten in der Qualität. Die indirekten Lichtstrahlen werden, wie bei dem *high detail shading*, mit der *irradiance cache* Methode ermittelt, die *records* werden jedoch in den *faces* abgespeichert. Die *records* beinhalten die *irradiance* von ausgehendem Licht (s. $L_0(p,v)$ in Formel 1). [4]

5.3 Spatiotemporal Variance-Guided Filtering

SVGF (*spatiotemporal variance-guided filtering*) ist eine Lösung, um verrauschte Ray-Tracing bzw. Path-Tracing Bilder zu entrauschen. Der Algorithmus ist auf das Entrauschen von 1 SPP Bilder spezialisiert. [18]

Der *path tracer* in dem SVGF-Algorithmus generiert jeweils ein 1 SPP Rendering für direktes- und indirektes Licht. Das jeweilige Rendering wird in einem dafür spezialisiertem *reconstruction filter* entrauscht. Das Bild mit direktem Licht wird dabei anders entrauscht, als das Bild mit indirektem Licht, um bei beiden Renderings ein optimales Ergebnis zu erreichen. In dem *reconstruction filter* finden zwei wichtige Prozesse statt. Ein zeitlicher Filter (engl. *temporal filter*) und ein räumlicher Filter (engl. *spatial filter*). Der zeitliche Filter speichert alle generierten 1 SPP Renderings in einem Puffer ab. Mit jedem neu hinzugefügtem Rendering steigt die Abtastrate (SPP) in dem Puffer. Als Beispiel: Aus einhundert 1 SPP Renderings kann ein 100 SPP Rendering generiert werden. Der räumliche Filter entfernt die hochfrequenten Details in dem Bild, um das Rauschen möglichst zu entfernen. Jedoch wird zuvor die Albedo Map entfernt, da ansonsten der räumliche Filter ebenfalls die hochfrequenten Details in der Textur entfernt. Nachdem die Bilder entrauscht wurden, werden sie zusammengefügt, womit ein Rendering entsteht, das direktes- sowie indirektes Licht enthält. Die zuvor entfernte Albedo Textur wird wieder hinzugefügt. [18] Die letzten beiden Schritte *tone mapping* und *temporal antialiasing* werden in dieser Ausarbeitung nicht erläutert. Diese Informationen können in der dazugehörigen wissenschaftlichen Ausarbeitung von Schied et al. (2017) [18] entnommen werden.

6 Fazit

Ziel war es herauszufinden, mit welchen Methoden der Path-Tracing Algorithmus beschleunigt werden kann, um eine Echtzeit Ausführung zu ermöglichen.

Der SVGF-Algorithmus ist eine wichtige Methode, um eine Echtzeit-Ausführung zu realisieren, denn die 1 SPP Renderings können relativ, je nach Komplexität der Szene und der Rendering-Auflösung, schnell generiert werden. Mit *irradiance cache* können viele aufwendige Berechnungen wegfallen, die für diffuse indirekte Lichtstrahlen notwendig sind. Das bringt einen deutlichen Performance Verbesserung und das bei fast gleichbleibender

Qualität. Mit *shading level of detail* kann ebenfalls viel Leistung eingespart werden, da viele Oberflächenpunkte keine hochdetaillierte Schattierung benötigen.

Die Ergebnisse haben gezeigt, dass viele Optimierungsmöglichkeiten bzgl. Path-Tracing existieren, um eine Echtzeitausführung zu ermöglichen. In Zukunft sind weitere Fortschritte auf diesem Gebiet zu erwarten, die die Varianz und/oder die Performance bzgl. Path-Tracing verbessern.

Literatur

[1] Tomas Akenine-Möller, Cyril Crassin, Jakub Boksansky, Laurent Belcour, Alexey Panteleev, and Oli Wright. 2021. Improved Shader and Texture Level of Detail Using Ray Cones. *Journal of Computer Graphics Techniques (JCGT)* 10, 1 (2021), 1–24. http://jcgt.org/published/0010/01/01/

[2] Tomas Akenine-Möller, Eric Haines, Naty Hoffman, Angelo Pesce, Michał Iwanicki, and Sébastien Hillaire. 2018. *Real-time rendering* (fourth edition ed.). CRC Press Taylor & Francis Group, Boca Raton and London and New York.

[3] Arthur Appel. 1968. Some techniques for shading machine renderings of solids. In *Proceedings of the April 30–May 2, 1968, spring joint computer conference on - AFIPS '68 (Spring)*, Unknown (Ed.). ACM Press, New York, New York, USA, 37. https://doi.org/10.1145/1468075.1468082

[4] Jakub Boksansky and Oli Wright. 2020. Minecraft with RTX: Crafting a Real-Time Path-Tracer for Gaming. https://www.gdcvault.com/play/1026716/Minecraft-with-RTX-Crafting-a

[5] Robert L. Cook, Thomas Porter, and Loren Carpenter. 1984. Distributed ray tracing. *ACM SIGGRAPH Computer Graphics* 18, 3 (1984), 137–145. https://doi.org/10.1145/964965.808590

[6] Joey de Vries. 2020. *Learn OpenGL - Graphics programming: Learn modern OpenGL graphics programming in a step-by-step fashion.* Kendall & Welling, o.O.

[7] Borivoje Furht. 2009. *Handbook of Multimedia for Digital Entertainment and Arts.* Springer, Dordrecht. https://doi.org/10.1007/978-0-387-89024-1

[8] Fabio Ganovelli, Massimiliano Corsini, Sumanta Pattanaik, and Marco Di Benedetto. 2015. *Introduction to Computer Graphics: A Practical Learning Approach.* CRC Press, Hoboken. http://gbv.eblib.com/patron/FullRecord.aspx?p=1648266

[9] Eric Haines and Tomas Akenine-Möller (Eds.). 2019. *Ray Tracing Gems: High-Quality and Real-Time Rendering with DXR and Other APIs.* Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-4427-2

[10] James T. Kajiya. 1986. The rendering equation. *ACM SIGGRAPH Computer Graphics* 20, 4 (1986), 143–150. https://doi.org/10.1145/15886.15902

[11] Matias Koskela, Kalle Immonen, Markku Mäkitalo, Alessandro Foi, Timo Viitanen, Pekka Jääskeläinen, Heikki Kultala, and Jarmo Takala. 2019. Blockwise Multi-Order Feature Regression for Real-Time Path-Tracing Reconstruction. *ACM Transactions on Graphics* 38, 5 (2019), 1–14. https://doi.org/10.1145/3269978

[12] Jaroslav Krivanek and Pascal Gautron. 2009. *Practical global illumination with irradiance caching.* Synthesis lectures on computer graphics and animation, Vol. 10. Morgan & Claypool, S.l.

[13] Boaz Livny. 2008. *Mental ray for Maya, 3ds max, and XSI: A 3d artist's guide to rendering.* John Wiley, Indianapolis, Ind. http://site.ebrary.com/lib/academiccompletetitles/home.action

[14] Daniel Meister, Jakub Boksansky, Michael Guthe, and Jiri Bittner. 2020. On Ray Reordering Techniques for Faster

GPU Ray Tracing. In *Symposium on Interactive 3D Graphics and Games*. ACM, New York, NY, USA, 1–9. https://doi.org/10.1145/3384382.3384534

[15] Thomas Müller, Fabrice Rousselle, Jan Novák, and Alexander Keller. 2021. Real-time neural radiance caching for path tracing. *ACM Transactions on Graphics* 40, 4 (2021), 1–16. https://doi.org/10.1145/3450626.3459812

[16] Jon Peddie. 2019. *Ray Tracing: A Tool for All*. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-17490-3

[17] Matt Pharr, Wenzel Jakob, and Greg Humphreys. 2017. *Physically Based Rendering: From Theory to Implementation* (third edition ed.). Elsevier Morgan Kaufmann Publishers, Amsterdam and Boston and Heidelberg and London and New York and Oxford and Paris and San Diego and San Francisco and Singapore and Sydney and Tokyo.

[18] Christoph Schied, Anton Kaplanyan, Chris Wyman, Anjul Patney, Chakravarty R. Alla Chaitanya, John Burgess, Shiqiu Liu, Carsten Dachsbacher, Aaron Lefohn, and Marco Salvi. 2017. Spatiotemporal variance-guided filtering. In *Proceedings of High Performance Graphics (ACM Digital Library)*, Morgan McGuire (Ed.). ACM, New York, NY, 1–12. https://doi.org/10.1145/3105762.3105770

[19] Christoph Schied, Christoph Peters, and Carsten Dachsbacher. 2018. Gradient Estimation for Real-time Adaptive Temporal Filtering. *Proceedings of the ACM on Computer Graphics and Interactive Techniques* 1, 2 (2018), 1–16. https://doi.org/10.1145/3233301

[20] Peter Shirley. 2020. *Ray Tracing: The Next Week* (v3.2.3 ed.). https://raytracing.github.io/books/RayTracingTheNextWeek.html

[21] Peter Shirley. 2020. *Ray Tracing: The Rest of Your Life* (v3.2.3 ed.). https://raytracing.github.io/books/RayTracingTheRestOfYourLife.html

[22] Michael Stengel, Zander Majercik, Benjamin Boudaoud, and Morgan McGuire. 2021. A distributed, decoupled system for losslessly streaming dynamic light probes to thin clients. In *Proceedings of the 12th ACM Multimedia Systems Conference*, Özgü Alay, Cheng-Hsin Hsu, and Ali C. Begen (Eds.). ACM, New York, NY, USA, 159–172. https://doi.org/10.1145/3458305.3463379

[23] Turner Whitted. 1980. An improved illumination model for shaded display. *Commun. ACM* 23, 6 (1980), 343–349. https://doi.org/10.1145/358876.358882

[24] Jinyuan Yang, Xiaoli Li, and Abraham G. Campbell. 2020. Variable Rate Ray Tracing for Virtual Reality. In *SIGGRAPH Asia 2020 Posters*. ACM, New York, NY, USA, 1–2. https://doi.org/10.1145/3415264.3425451

©2022 Dennis Grendatsch. 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

State-of-the-Art Analyse von 3D-Scans primitiver Formen

Michael Stadelmaier
Hochschule Reutlingen
Reutlingen, Deutschland
Michael.Stadelmaier@Student.Reutlingen-University.de

Zusammenfassung

Die Automatisierung von industriellen Fertigungsanlagen durch Roboter schreitet immer weiter voran. Des Weiteren werden Lasermessverfahren und Tiefenkameras immer genauer und günstiger. Die Punktwolken dieser Kameras und Lasermessverfahren werden dadurch vermehrt in Roboterapplikationen verwendet, um die stetig wachsenden Aufgaben durch die dreidimensionalen Informationen präziser und flexibler zu erledigen. Der State-of-the-Art der Robot-Vision-Applikationen sowie Arbeiten mit künstlicher Intelligenz in Zusammenhang mit Punktwolken werden in dieser Arbeit dargestellt. Diese Arbeit liefert einen Überblick über die möglichen Einsatzgebiete und Grenzen der Lasermessverfahren und Tiefenkameras und der anschließenden Verarbeitung von Punktwolken.

CCS Concepts

• **Hardware** → *Scanners*; *Sensors and actuators*; • **Computer systems organization** → *Robotic autonomy*; • **Computing methodologies** → *Computer vision problems.*

Keywords

Punktwolken, Objekterkennung, 3D-Scans

1 Einleitung

Um Maschinen mit Rohmaterial zu beladen, werden zunehmend automatisierte Anlagen mit Robotern verwendet. Damit Roboter die Bauteile greifen können, müssen diese bisher auf einer exakten Position liegen. Außerdem müssen die Rohmaterialien dieselben Dimensionen aufweisen, damit sie im Greifer festgehalten werden können. Eine falsche Positionierung der Bauteile durch Verrutschen oder Unachtsamkeit der bedienenden Person kann zu Störungen durch eine Kollision des Roboters führen. Dies kann – unter bestimmten Umständen – auch zu einem Schaden am Roboter, Greifer oder auch der Rohmaterialunterlage führen. Solch vermeidbaren Schäden verursachen Kosten.

Außerdem müssen für eine effiziente Beladung der Automation die Träger der Werkstücke getauscht werden. So etwas führt, vor allem bei neuen Bauteilen, zu langen Rüst- und Einrichtungszeiten. Während solcher Zeiträume steht die Anlage, und somit auch die Produktion still.

In dieser Arbeit wird eine State-of-the-Art Analyse vom Stand der aktuellen Technik durchgeführt um herauszufinden, ob und mit welchen Systemen die unterschiedlichen Körper erfasst werden und welche Möglichkeiten hierbei bestehen. Dabei werden ebenfalls Methoden zur Erkennung komplexer Formen, sowie ähnliche Verfahren berücksichtigt, um ein möglichst breites Spektrum

Betreuer/-in Hochschule	Prof. Dr.-Ing. Cristóbal Curio Hochschule Reutlingen Cristobal.Curio@Reutlingen-University.de
Betreuer/-in Extern	Dr. Holger Baur CMO-SYS GmbH, Göppingen h.baur@cmo-sys.com

Informatics Inside Frühjahr 2022
11. Mai 2022, Hochschule Reutlingen
Copyright 2022 Michael Stadelmaier

abbilden zu können. Das Hauptaugenmerk wird dabei auf der Erkennung und Einsetzbarkeit von primitiven Formen und der Berechnung der Greifposition des Roboters bei unterschiedlicher Lage der Bauteile liegen.

Im weiteren Verlauf dieser Arbeit werden in Abschnitt 2 die Grundlagen zu den ausgewählten Scanverfahren dargestellt, mit welchen dreidimensionale Objekte digitalisiert werden können. Anschließend folgt in Abschnitt 3 der Stand der Wissenschaft mit den Ergebnissen der Literaturrecherche zur Erkennung von Objekten – mit und ohne Einsatz von künstlicher Intelligenz – sowie eine kleine Auswahl an Software zur Verarbeitung der digitalisierten dreidimensionalen Datenstrukturen. In Abschnitt 4 folgt die Diskussion mit einer Bewertung der Laserscanverfahren, sowie eine Zusammenfassung und Einsetzbarkeit der vorgestellten Arbeiten. Die Arbeit endet mit einer Schlussfolgerung und einem kurzen Ausblick in Abschnitt 5.

2 Grundlagen

Nach aktuellem Forschungsstand sind unterschiedliche Systeme im Einsatz, mit welchen dreidimensionale Objekte digitalisiert werden. Hierzu zählen Time-Of-Flight (TOF) Sensoren, welche einen definierten Lichtimpuls aussenden und die Dauer bis zur eintreffenden Reflexion im Sensor messen. Über diese zeitliche Verzögerung kann ein Rückschluss auf die Distanz gezogen werden. Weitere Sensoren werden in dem Leitfaden zur industriellen Bildverarbeitung von der Frauenhofergesellschaft weitere mögliche Verfahren beschrieben (vgl. [13]). In dieser Arbeit werden in den Grundlagen jedoch nur die laserbasierten Verfahren beschrieben, da diese überwiegend in der Literaturrecherche aufgetreten sind. Hierzu zählt das Triangulationsverfahren, Lichtschnittverfahren, sowie das Streifenprojektionsverfahren. In allen drei Verfahren wird mithilfe von Laserlicht und dessen Reflexion auf den Oberflächen gearbeitet.

2.1 Triangulationsverfahren

Das grundlegendste Verfahren für die Erfassung von 3D-Objekten ist das Triangulationsverfahren. Bei diesem Prinzip wird ein Objekt mit einem Laserstrahl beleuchtet. Dieser punktförmige Laserstrahl beleuchtet einen einzelnen Punkt, welcher durch die Oberfläche des Objektes reflektiert und gestreut und dadurch von einem Detektor erfasst wird (vgl. Abbildung 1). Durch den bekannten Winkel aus der Kalibrierung zwischen Laserstrahl und Detektor, sowie der erkannten Position auf dem Detektor kann die Höhe des Punktes auf dem Objekt exakt erkannt werden. Die Lasertriangulation mit punktueller Abtastung ist nach Versuchen von Grote und Schwab sehr robust [3].
Die genaue Funktionsweise des Triangulationsverfahren ist in [3, S. 590] ausführlich beschrieben und kann dort nachgeschlagen werden.

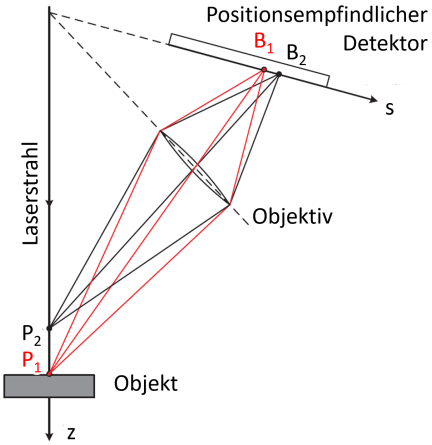


Abbildung 1: Abbildungsgeometrie Lasertriangulation. Abbildung aus [3][1].

[1] Die Grafik wurde leicht angepasst um eine bessere Lesbarkeit der Texte herzustellen.

2.2 Lichtschnittverfahren

Das Lichtschnittverfahren baut auf dem Prinzip des Triangulationsverfahren auf. Als Erweiterung wird – wie in Abbildung 2 zu sehen – von der Laserquelle eine ganze Linie und nicht nur ein einzelner Punkt ausgestrahlt. Der Bildsensor erfasst die Reflexion der Laserlinie und kann dadurch die Kontur des Messobjektes erkennen. Diese Verfahren sollte auch schneller, als das reine punktförmige Triangulationsverfahren sein, da hier eine ganze Linie ausgewertet werden kann und der Lichtstrahl nicht durch mechanischen Aufwand verschoben wird.

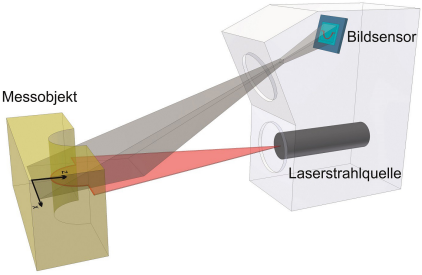


Abbildung 2: Messprinzip eines Lichtschnittsensors aus [9].

Ein Lichtschnittsensor besteht hierbei aus einem Bildsensor, Linienlaser und der benötigten Abbildungsoptiken [9]. Dieses Messverfahren ist in der optischen Formerfassung weit verbreitet. Der Bildsensor erfasst die reflektierenden Strahlen der Laserlinie. Wie in Abbildung 2 dargestellt ist der Verlauf der Linie je nach Objekthöhe und Form unterschiedlich gekrümmt. Für die anschließende Berechnung der Objekthöhen wird das Triangulationsverfahren verwendet. [9]

2.3 Streifenprojektionsverfahren

Das Musterverfahren ist eine Weiterentwicklung aus dem Lichtschnittverfahren. Dabei wird die Szene mit einem oder auch mehreren Mustern aus Laserlicht beleuchtet. Über eine Abbildungsoptik werden die projizierten Streifenmuster erfasst. Durch die Objekte in der Szene erscheinen die Projektionsmuster nach der Aufnahme mit einer gewollten Verzerrung.[15] In dem Leitfaden der Fraunhofer Vision zur optischen 3D-Vermessung [10] wird außerdem beschrieben, dass durch die Streifenprojektion die räumlichen Koordinaten nur an den Streifengrenzen oder in Streifenmitte exakt berechnet werden können. Die Koordinaten dazwischenliegender Punkte müssen interpoliert werden. Durch die Phasenverschiebung (Helligkeitsmodulation) wird dieses Problem jedoch vermieden. Die Informationen und Bilder der Phasenverschiebung wird mit dem Grauwertprofil in einer Phasenkarte zusammengefügt. Über ein unwrapping-Verfahren wird die kontinuierliche Phasenkarte erstellt und daraus können die Höheninformationen abgeleitet, sowie die Punktwolke erstellt werden [15]. In Abbildung 3 ist das Vorgehen unter Verwendung von verschiedenen Mustern und einer Phasenverschiebung der Muster dargestellt. Die Kinect©, welche von Microsoft entwickelt wurde, ist eine bewegliche, genaue, günstige und schnelle Tiefenkamera zum Erfassen von dreidimensionalen Objekten, welche dieses Verfahren verwendet. Die Genauigkeit der Kamera liegt unter 2 mm. [14] Die Kinect©-Kamera arbeitet mit einer RGB-Kamera, einer IR-Kamera und einem Projektor. Um die Tiefe zu ermitteln wird die Szene durch den Projektor mit einem Projektionsmuster beleuchtet und mittels Triangulation die Entfernungen zu den Objekten in der Szene berechnet. [12]

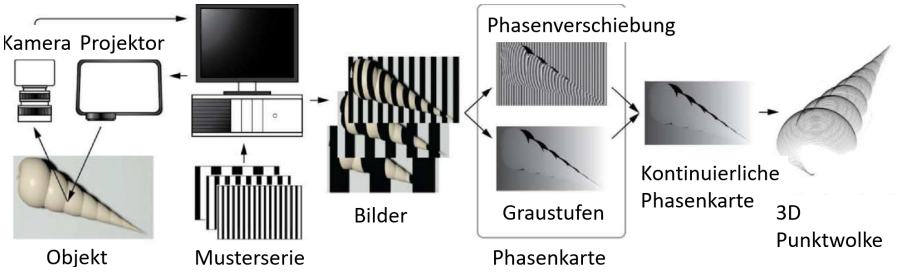


Abbildung 3: Arbeitsprozess des Streifenprojektionsverfahren. Abbildung übersetzt aus [15]

3 Stand der Wissenschaft

Dieser Abschnitt ist in die drei Bereiche „Erkennung von Objekten", „Künstliche Intelligenz und Punktwolken" und „Entwicklungsumgebungen zur Verarbeitung von Punktwolken" untergliedert. Im Abschnitt 3.1 werden Arbeiten vorgestellt, welche die Verfahren aus den Grundlagen verwenden um Objekte zu erkennen und ggfs. mit einem Roboter zu manipulieren. Anschließend folgen in Abschnitt 3.2 Methoden auf der Basis von künstlicher Intelligenz zur Verarbeitung von Punktwolken. Zuletzt folgen in Abschnitt 3.3 einige Softwarelösungen mit welchen die Punktwolken verarbeitet werden können.

3.1 Erkennung von Objekten

In diesem Abschnitt werden unterschiedliche Arbeiten dargestellt um einen Überblick über die verschiedenen Einsatzgebiete von Objekterkennungen mit Punktwolken darzustellen. Dabei handelt es sich um eine kleine Auswahl an Anwendungsmöglichkeiten.

3.1.1 Depalettierung

Ein mögliches Einsatzgebiet für die Einbindung eines Lichtschnittsensor in eine bestehende Automation wurde von Grote und Schwab [3] beschrieben. In dieser Arbeit wurde in einem bestehenden Prozess eine Robot-Vision-Applikation erfolgreich integriert. Konkret geht es in dieser Veröffentlichung darum, einen Depalettierungsprozess von Rohren zu automatisieren. Bei diesem Vorgang liegen die Rohre nicht immer so sauber geordnet in dem Materialrahmen wie in Abbildung 4.

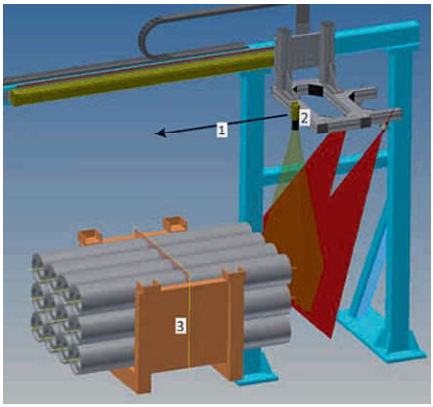


Abbildung 4: Aufbau der Scaneinheit. Abbildung aus [3]. Aufbau besteht aus einem Lichtschnittsensor (Kamera + zwei Linienprojektoren) (2), welcher mittels Linearachse den Weg (1) über die Palette mit den Rohren (3) fährt um die Bauteile zu scannen.

Bei einer so geordneten Lage des Materials – wie in Abbildung 4 – werden nach der Entnahme des ersten Rohrs, die anderen Rohre in die Mulden der Lage darunter rutschen womit sich die Lage der Rohre beliebig ändern kann. Durch die Verbindung des Lichtschnittsensor an einer Linearachse und die Koppelung mit einem inkrementellen Weggeber

kann die Punktwolke in einen korrekten metrischen Zusammenhang gebracht werden. Der Prismenmagnetgreifer greift anschließend eines der Rohre, an der von der Computer-Vision berechneten mittigen Position. Das Bildverarbeitungssystem segmentiert hierfür die einzelnen Bauteile durch den gegebenen Rohrdurchmesser des Prozessleitsystems. Durch die Z-Koordinaten der berechneten Greifpunkte kann das oberste Bauteil gefunden und somit die Auswahl des als nächstes zu greifende Rohr getroffen werden. Um Kollisionen durch die Bewegung der Bauteile nach einer Entnahme gegenzusteuern, werden die Rohre vor jedem Greifvorgang erneut gescannt und wieder das oberste Bauteil gegriffen. Durch dieses Verfahren kann das letzte Bauteil erkannt und diese Information an das Prozessleitsystem weitergeleitet werden.

3.1.2 Prozessoptimierung und Greifpunktberechnung

Die Arbeit von Wang und Yu [15] verwendet das Lichtschnittverfahren um Achsträger von Kraftfahrzeugen mit einem Roboter zu greifen. Die Bauteile laufen auf einem Fließband von der Schmiede zur Entgratung und werden dabei von dem Lichtschnittsensor vermessen. Anhand der vom Bildverarbeitungssystem berechneten Greifkoordinaten werden die Bauteile von einem Roboter gegriffen und der Entgratung zugeführt. Zusätzlich kann der entstandene Grat, sowie die Bauteilgeometrie direkt vermessen werden.

3.1.3 Maßschneiderung

Ebenfalls wird in der Arbeit von Wang und Yu [15] auf die Verwendung im Bekleidungsbereich hingewiesen, um die Maßschneiderungen schneller und präziser zu gestalten. Außerdem wäre es für den Schneider möglich, alle Maße jederzeit neu einsehen zu können. Ein vergessenes Maß wäre dann kein Problem mehr. Ein Menschlicher Körper müsste zu diesem Zweck mehrmals aufgenommen werden, da es hier viele Schattierungen geben kann. Diese einzelnen Punktewolken müssen im Nachgang noch zusammengesetzt werden.

3.1.4 Erkennung anatomischer Strukturen

Im medizinischen Kontext ist die Erfassung von 3D-Volumen ebenfalls sinnvoll, um die anatomischen Strukturen eines Patienten zu bestimmen, um dadurch navigiert operieren zu können. In der Arbeit von Scarpin und Wahrburg [11] wird ein Lichtschnittsensor entwickelt und dieser an einem Roboter über dem Patienten geführt. Durch die Führung des Lichtschnittsensors am Roboter wird die aufgenommene Punktewolke homogen und die Position des Sensors kann durch die Robotergeometrie berechnet werden. Mithilfe der linearen Bewegung des Roboters kann nun der zu erfassende Bereich – im Gegensatz zur Sensorführung von Hand – zitter- und bewegungsfrei aufgenommen werden von Scarpin und Wahrburg aus Kosten- und Gewichtsgründen selbst entwickelt [11]. Dadurch lässt sich laut Scarpin und Wahrburg das 3D Bild, im Anschluss an die Datenaufnahme erheblich einfacher rekonstruieren. Um die Position des Roboters in dem Koordinatensystem des Bildverarbeitungssystems zu berechnen, ist eine Kalibrierung notwendig. Der eigens entwickelte Lichtschnittsensor wurde mit einer Genauigkeit von kleiner 1mm bei einer Entfernung von 600mm angegeben.

3.2 Künstliche Intelligenz und Punktwolken

In diesem Abschnitt werden zwei unterschiedliche neuronale Netze vorgestellt, auf welche im Rahmen der Literaturrecherche häufig gestoßen wurde.

3.2.1 Point-View Network

You et. al. [17] schlagen zur Objekterkennung von 3D-Punktwolken das Point-View Network (PVNet) vor. Das PVNet arbeitet mit der 3D-Punktwolke und mehreren 2D-Bildern der Punktwolke aus verschiedenen Ansichten. Hierbei handelt es sich um eine reine Objekterkennung wie bspw. einem Flugzeug. Die Autoren vergleichen in ihrer Arbeit das vorgestellte PVNet mit anderen neuronalen Netzen und erreichen durch die zusätzliche Verwendung von 2D-Bildern eine Gesamtgenauigkeit von 93, 2% in der Objekterkennung und dadurch eine ein bis vier Prozent höhere Genauigkeit im Vergleich zu neuronalen Netzen [4, 6, 7, 16], welche rein auf Punktewolken basieren. Für die Verwendung des PVNet ist für die bessere Erkennung ein Mehraufwand durch die Erstellung der 2D-Bilder nötig.

3.2.2 PointNet++

In der Arbeit von Zhang [18] werden Objekte mit der Kinect© 2-Kamera aufgenommen. Die aufgenommene Punktewolke wird anschließend dem *PointNet++* [7] übergeben. Dieses tiefe neuronale Netzwerk von Qi et al. [7] dient zur Segmentierung und Klassifizierung von 3D-Punktwolken. Dabei extrahiert das Netzwerk hierarchische Merkmalsstrukturen aus der Punktwolke. Dieses tiefe neuronale Netz wird in der Arbeit von Zhang [18] als Grundlage verwendet um Merkmale aus Punktwolken zu extrahieren. Anschließend werden die extrahierten Merkmale weiter verarbeitet um die Greifposition für einen Zwei-Finger-Greifer zu berechnen. In der Arbeit konnte der Roboter etwa 80.3% der Bauteile durch die errechnete Greifposition greifen.

3.3 Entwicklungsumgebungen zur Verarbeitung von Punktwolken

In diesem Abschnitt folgen Beispiele für in der Industrie und Forschung verwendete Software zur Verarbeitung von Punktwolken. Hierbei handelt es sich um eine kleine Auswahl.

3.3.1 HALCON

"HALCON ist die umfassende Standardsoftware für die industrielle Bildverarbeitung (Machine Vision) mit integrierter Entwicklungsumgebung (IDE), die weltweit zum Einsatz kommt."[5] Mit diesen Worten beginnt die Münchner Firma MVTec ihre Broschüre zur Vorstellung von HALCON. Die Software ist auf Windows, Linux und macOS verwendbar und kann durch Programmiersprachen wie C, C++ sowie über das .Net-Framework angesprochen werden. Des Weiteren garantiert MVTec eine Hardwareunabhängigkeit durch Schnittstellen zu verschiedenen Industrie-Kameras und Framegrabbern. Ebenfalls kann die Software auf Embedded-Vision-Plattformen verwendet werden. Im Automobilbau und der Robotik bewirbt MVTec HALCON mit der Positionserkennung von 3D-Objekten, der Extraktion von 3D-Daten zur Kommissionierung, sowie die Roboterbahnplanung. Auch die neuen Trends zu maschinellen Lernverfahren unterstützt HALCON mit der Funktion, verschiedene Deep-Learning-Methoden verwenden zu können. Diese maschinellen Lernverfahren können zum Beispiel zur Qualitätskontrolle oder auch zur Objekterkennung dienen. Ebenfalls können in HALCON, mit Kombination einer 3D-Kamera, Kanten detektiert sowie Konturen analysiert werden.
HALCON ist eine lizenzpflichtige Software und Programmierbibliothek. Die Kosten für die Lizenzen hängen hierbei von Art und Anzahl an benötigter Lizenzen ab.

3.3.2 MATLAB

Neben HALCON ist Matlab ebenfalls eine Standardsoftware in der Industrie sowie in der Forschung und Lehre. Diese Software von dem US-Unternehmen MathWorks ist eine Plattform für die Programmierung, numerischen Berechnungen, Datenanalysen, Algorithmenentwicklung oder auch für die Erstellung von Modellen. Die Software ist eine kombinierte Desktop-Umgebung mit eigener Programmiersprache, sowie einem Editor zum Programmieren und Ausgeben der visualisierten Daten. Dabei bietet der Editor zusätzlich die Verwendung von Toolboxen und interaktiven Apps. Neben dem eigenen Editor können die Matlab-Bibliotheken ebenfalls mit Programmiersprachen wie Python, C/C++ oder Java genutzt werden. Ebenfalls lassen sich mit diesen Bibliotheken Aufgaben der Bildverarbeitung und Computer Vision, maschinellem Lernen, Robotik, Signalverarbeitung, uvm.[2] lösen. Die grundlegenden Datentypen von MATLAB, Vektoren und Matrizen, eignen sich hervorragend für Probleme in der Robotik und der Computer Vision [2]. Im Bereich der Robotik ist bei Matlab ebenfalls die Verarbeitung und Visualisierung von Sensordaten aus Tiefenkameras, welche Punktwolken der aufgenommenen Umgebung ausgeben, zu finden.

3.3.3 Point-Cloud-Library

Im Gegensatz zu HALCON und MATLAB ist die Point-Cloud-Library (PCL) ein Open-Source-Projekt unter der BSD-Lizenz und steht damit jedem kostenfrei zur Verfügung und kann kommerziell genutzt werden. PCL wurde 2011 von Rusu und Cousins [8] vorgestellt. In dem Paper zur Vorstellung von PCL wurde beschrieben, dass die Bibliothek in der Sprache C++ programmiert wurde und für die Plattformen Windows, Linux und macOS verfügbar ist. Eine Unterstützung für Android wurde zum damaligen Zeitpunkt in Aussicht gestellt und sei in Arbeit.

[2] siehe https://de.mathworks.com/solutions.html

PCL enthält eine Vielzahl von 3D Algorithmen für die Verarbeitung von Punktwolken. Hierzu gehören unter anderem die Filterung, Merkmalsschätzung, Segmentierung und die Registrierung. Die einzelnen Algorithmen sind in der Bibliothek jeweils als eine Basisklasse mit allen gemeinsamen Funktionen implementiert. Dadurch kann die Implementierung der Algorithmen kompakt und sauber gehalten werden. Eine grundlegende Verarbeitungspipeline für Punktwolken in PCL besteht aus:

- Erstellung des Verarbeitungsobjekte (z.B. Filter).
- Übergabe der Punktwolke an das Verarbeitungsmodul.
- Setzen von Parametern.
- Ausgabe.

Damit im bestehendem Code, welcher auf PCL aufbaut, durch Änderungen der Autoren am Quellcode keine Fehler entstehen und der Code nicht mehr verwendbar ist, wird der Quellcode regelmäßigen automatisierten Unit- und Regressionstests unterzogen. Für die Programmierbibliothek wurde außerdem eine Wiki-Seite mit einer großen Anzahl an Quellcodebeispielen und Schritt-für-Schritt Anleitungen bereitgestellt. Außerdem können Robotikanwendungen durch die volle Integration des Robot Operation System (ROS)[3] direkt in PCL umgesetzt werden.

4 Diskussion

Das Triangulationsverfahren ist im Vergleich zum Lichtschnittverfahren und Streifenprojektionsverfahren, bedingt durch die punktuelle Abtastung der Szene, das langsamste Verfahren zur Digitalisierung von dreidimensionalen Szenen. Das Triangulationsverfahren hat sich dennoch aufgrund seiner Robustheit und Genauigkeit bereits in vielen Anwendungen bewährt. [1] Durch das Lichtschnittverfahren kann eine Szene schneller als beim Triangulationsverfahren erfasst werden, da hier eine ganze Linie gescannt wird.

[3] siehe https://www.ros.org

Das Lichtschnittverfahren würde sich Beispielsweise gut eignen, um Bauteile auf einem Laufband oder besonders lange Werkstücke zu erfassen, wie Rohre in der Arbeit von Grote und Schwab [3].
Beim Streifenprojektionsverfahren wird die Szene vollständig von der Kamera unter einer Musterbeleuchtung – Anordnung von unterschiedlichen Streifenmuster (vgl. Abbildung 3) – aufgenommen und kann dadurch eine große Szene beinahe in Echtzeit digitalisieren. Die Genauigkeiten dieser Systeme liegen bspw. bei der Microsoft Kinect© bei unter 2mm [14], bei der Intel RealSense L515 liegt die Tiefengenauigkeit bei etwa 14mm auf einer Distanz von 9m (vgl. Intel[4]).
Bei allen dieser drei Verfahren liegen die physikalischen Grenzen der Messunsicherheit bei der Blendenöffnung der Kamera, der Oberflächenbeschaffenheit sowie dem Triangulationswinkel (Winkel zwischen Laserquelle und Kamera). Die Messunsicherheit wird durch einen großen Triangulationswinkel verringert, jedoch wird dadurch die Abschattung – besonders an steilen Kanten – vergrößert. Wodurch das Messen von Bohrungen und die Erfassung von Hinterschneidungen problematisch sind. Kritisch sind bei diesen Messverfahren ebenfalls spiegelnde Oberflächen sowie ausgeprägte Riefenstrukturen. Diese rauen Oberflächen führen zu einem Rauschen durch die unterschiedlichen Lichtreflexionen der Berge und Täler. Ebenfalls führen Lichtdurchlässige Materialien wie Glas oder bestimmte Kunststoffe zu großen Messfehlern, da diese das Laserlicht nicht an der Oberfläche reflektieren.
Bei dem Triangulationsverfahren und Lichtschnittverfahren ist es notwendig diese an einer definierten Achse mit einem Wegmesssystem anzubringen, um einen korrekten metrischen Zusammenhang zu erzielen, wenn diese Sensoren über die Szene bewegt werden. Wenn Bauteile auf einem Fließband darunter durchfahren, sollte die Strecke bzw. die Geschwindigkeit der Bandes erfasst werden, um die Punktwolke in einen korrekten metrischen Zusammenhang zu bringen.

In den vorgestellten Arbeiten in Abschnitt 3.1 wurden keine primitiven Formen anhand der geometrischen Gegebenheiten aus den Punktwolken erkannt, vermessen und mit dem Roboter gegriffen. Jedoch lassen sich diese Arbeiten gut zusammenführen. In der Arbeit von Grote und Schwab [3] wird die Lage der Rohre im Palettensystem erfasst und dadurch ein Greifpunkt berechnet. Der Durchmesser der Rohre wird in dieser Implementierung allerdings nicht durch die Punktwolke berechnet, stattdessen wird der hinterlegte Durchmesser im Prozessleitsystem hierfür verwendet. Durch dieses definierte Maß werden die einzelnen Rohre der Punktwolke voneinander getrennt und anschließend die Greifposition des Magnetprismengreifer für ein einzelnes Rohr berechnet. In der Arbeit von Wang und Yu [15] wird das Lichtschnittverfahren verwendet, um ebenfalls die Greifposition des Roboters zu berechnen. Bei den Achsträgern in dieser Arbeit handelt es sich nicht um primitive Körper, welche direkt erkannt werden können. Jedoch werden hier Vermessungen anhand der Punktewolke durchgeführt. In den beiden weiteren Arbeiten in Abschnitt 3.1 handelt es sich nicht um die Greifpunktberechnung eines Roboters sondern um die Erkennung anatomischer Strukturen sowie das Ausmessen von Personen für eine Maßschneiderung. Diese Arbeiten zeigen ebenfalls das Potential auf, welches diese Verfahren für die Vermessung von Objekten und Körpern besitzen. Zusätzlich zeigen die Arbeiten auf, dass eine Vermessung von Objekten mit anschließender Greifpunktberechnung für einen Roboter möglich sind.
Neuronale Netze, wie das PointNet++ von Qi et al. [7], können die aufgenommenen Punktwolken zur Segmentierung und Klassifizierung verarbeiten. In der Arbeit von Zhang [18] wurde das PointNet++ zur Extrahierung

[4] https://www.intelrealsense.com/lidar-camera-l515/

von Merkmalen aus der Punktwolke verwendet. Anhand jener Merkmale wurde anschließend die Greifposition eines Zwei-Finger-Greifers berechnet. Dies zeigt, dass ebenfalls der Einsatz von künstlicher Intelligenz für die Greifpunktberechnung aus Punktwolkendaten unterstützen kann.

5 Schlussfolgerung

Durch die in Abschnitt 3 vorgestellten Arbeiten ist zu sehen, dass bereits viele Anwendungsgebiete für die Erfassung von dreidimensionalen Objekten und Szenen existieren. Mithilfe der dadurch erzeugten Punktwolken können Messungen zur Qualitätssicherung und Greifpunktberechnung für Roboter durchgeführt werden. Bei der Segmentierung von Bauteilen muss darauf geachtet werden, dass ein gewisser Abstand zwischen den Bauteilen vorhanden ist. Ohne diesen Abstand ist es schwer die Bauteile voneinander zu unterscheiden. Zwei sehr dicht aneinander stehende Würfel würden in diesem Fall – voraussichtlich – immer als Quader erkannt und folglich mit dem Roboter falsch gegriffen werden. Dies bestätigt auch das Vorgehen von Grote und Schwab [3], weshalb bei der Implementierung der Durchmesser der Rohre gegeben war, um die Segmentierung zu ermöglichen. Eventuell wäre es hier möglich mit Hilfe von maschinellen Lernverfahren die Segmentierungsprobleme zu umgehen.
In der Masterarbeit, welche an diese State-of-the-Art Analyse anknüpft, soll mit den hieraus gelernten Erkenntnissen ein Algorithmus entwickelt werden, um unterschiedliche Körper zu erfassen. Aus der Punktwolke sollen die Objekte und ihre Grundformen segmentiert sowie vermessen werden. Zusätzlich soll die Lage und Ausrichtung auf der Grundfläche sowie mögliche Greifpunkte für einen Roboter berechnet werden. Hierfür soll ebenfalls ein Schwellenwert hinterlegt werden können, ab welcher Höhe bzw. Größe von Körpern diese Berechnungen durchzuführen sind. Durch diesen Schwellenwert sollen kleine Körper – wie Schrauben, Unterlegscheiben o.ä. – herausgefiltert werden können. Ein solches Vorgehen soll mit – und ohne – den Einsatz von künstlicher Intelligenz erprobt werden. Als Programmierbibliothek soll in diesem Projekt auf das Open-Source-Projekt PCL zurück gegriffen werden, da es sich um eine kostenfreie Bibliothek handelt und die Ansteuerung von Robotern durch ROS durchgeführt werden kann.

Literatur

[1] Dirk Berndt und Norbert Bauer, Hrsg. *Leitfaden zu Grundlagen und Anwendungen der optischen 3-D-Messtechnik*. ger. 1. Aufl. Bd. 6. Vision. Erlangen: Fraunhofer-Allianz Vision, 2003. 40 S. ISBN: 3816762972. URL: http://www.vision.fraunhofer.de/de/publikationen/leitfaeden/band-6.html.

[2] Peter Corke. „MATLAB toolboxes: robotics and vision for students and teachers". In: *IEEE Robotics Automation Magazine* 14.4 (Dez. 2007), S. 16–17. ISSN: 1558-223X. DOI: 10.1109/MRA.2007.912004.

[3] Alexander Grote und Erwin Schwab. „Realisierung einer Robot-Vision-Applikation mit Lichtschnittverfahren für einen bestehenden Prozess". In: *teme* 81.12 (2014), S. 585–594. ISSN: 0171-8096. DOI: 10.1515/teme-2014-1056.

[4] Yangyan Li u. a. *PointCNN: Convolution On X-Transformed Points*. To be published in NIPS 2018, code available at github.com/yangyanli/PointCNN. 2018. URL: http://arxiv.org/pdf/1801.07791v5.

[5] MVTec. *HALCON Broschüre*. Hrsg. von MVTec. Version halcon-20.11. URL: https://go.mvtec.com/acton/attachment/43208/f-929b0bb1-e00f-464a-aab9-1a9968cd8bd9/1/-/-/-/-/halcon_20.11_steady_brochure_de.pdf (besucht am 25. 03. 2022).

[6] Charles R. Qi u. a. *PointNet: Deep Learning on Point Sets for 3D Classification*

and Segmentation. CVPR 2017. 2016. URL: https://arxiv.org/pdf/1612.00593.

[7] Charles R. Qi u. a. *PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space*. 2017. URL: https://arxiv.org/pdf/1706.02413.

[8] Radu Bogdan Rusu und Steve Cousins. „3D is here: Point Cloud Library (PCL)". In: *2011 IEEE International Conference on Robotics and Automation. (ICRA 2011) ; Shanghai, China, 9 - 13 May 2011*. 2011 IEEE International Conference on Robotics and Automation (ICRA) (Shanghai, China, 5. Sep. 2021). Hrsg. von Antonio Bicchi. IEEE Robotics and Automation Society. Piscataway, NJ: IEEE, 2011, S. 1–4. ISBN: 978-1-61284-386-5. DOI: 10.1109/ICRA.2011.5980567.

[9] Michael Sackewitz, Hrsg. *Handbuch zur industriellen Bildverarbeitung. Qualitätssicherung in der Praxis*. ger. 3., vollständig überarbeitete und aktualisierte Auflage. Sackewitz, Michael (HerausgeberIn). Stuttgart: Fraunhofer Verlag, 2017. 447 S. ISBN: 3839612268. URL: http://www.vision.fraunhofer.de/de/publikationen/leitfaeden/band-17.html.

[10] Michael Sackewitz, Hrsg. *Leitfaden zur optischen 3D-Messtechnik*. ger. Bd. 14. Vision-Leitfaden. Sackewitz, Michael (Hrsg.) Stuttgart: Fraunhofer-Verl., 2014. 109 S. ISBN: 3839607612.

[11] D Scarpin und J Wahrburg. „Entwicklung eines robotergeführten Lichtschnittsensors für die berührungslose Erfassung anatomischer Strukturen". In: URL: http://ceur-ws.org/vol-1475/proceedings_curac_2010_paper_46.pdf.

[12] Jan Smisek, Michal Jancosek und Tomas Pajdla. „3D with Kinect". In: *Consumer depth cameras for computer vision. Research topics and applications*. Hrsg. von Andrea Fossati u. a. Advances in computer vision and pattern recognition. London: Springer, 2012, S. 3–25. ISBN: 978-1-4471-4640-7. DOI: 10.1007/978-1-4471-4640-7_1.

[13] Klaus Spinnler. *Leitfaden zur industriellen Bildverarbeitung*. ger. Unter Mitarb. von Michael Sackewitz. Bd. 13. Vision Leitfaden. Sackewitz, Michael (Red.) Erlangen und Stuttgart: Fraunhofer-Allianz Vision und Fraunhofer-Verl., 2012. 88 S. ISBN: 3839604478. URL: http://shop.vision.fraunhofer.de/artikel/leitfaden-13-industrielle-bildverarbeitung.

[14] Yan Wan u. a. „Edge Voxel Erosion for Noise Removal in 3D Point Clouds Collected by Kinect©". In: *Proceedings of the 2020 2nd International Conference on Image, Video and Signal Processing*. IVSP '20: 2020 2nd International Conference on Image, Video and Signal Processing (Singapore Singapore, 20. März 2020). ACM Digital Library. New York, NY, United States: Association for Computing Machinery, 2020, S. 59–63. ISBN: 9781450376952. DOI: 10.1145/3388818.3388821.

[15] Kesheng Wang und Quan Yu. „Accurate 3D object measurement and inspection using structured light systems". In: *Proceedings of the 12th International Conference on Computer Systems and Technologies*. the 12th International Conference (Vienna, Austria, 16. Juni 2011). Hrsg. von Boris Rachev. ACM Other conferences. Association for Computing Machinery. New York, NY, United States: ACM, 2011, S. 221. ISBN: 9781450309172. DOI: 10.1145/2023607.2023646.

[16] Yue Wang u. a. *Dynamic Graph CNN for Learning on Point Clouds*. 2018. URL: https://arxiv.org/pdf/1801.07829.

[17] Haoxuan You u. a. „PVNet: A Joint Convolutional Network of Point Cloud and Multi-View for 3D Shape Recognition“. In: *Proceedings of the 26th ACM international conference on Multimedia*. MM '18: ACM Multimedia Conference (Seoul Republic of Korea, 22. Okt. 2018). Hrsg. von Susanne Boll. ACM Conferences. Association for Computing Machinery-Digital Library und ACM Special Interest Group on Multimedia. New York, NY: ACM, 2018, S. 1310–1318. ISBN: 9781450356657. DOI: 10.1145/3240508.3240702.

[18] Tianze Zhang. „An optimal grasping point identification method based on deep learning and point cloud processing“. In: *2021 International Conference on Aviation Safety and Information Technology*. ICASIT 2021: 2021 International Conference on Aviation Safety and Information Technology (Changsha China, 18. Dez. 2021). New York, NY, United States: ACM, 2021, S. 667–671. ISBN: 9781450390422. DOI: 10.1145/3510858.3511355.

©2022 Michael Stadelmaier. 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

BIM Collaboration Format - Theorie und Praxis

Moritz Häcker
Hochschule Reutlingen
Reutlingen, Deutschland
Moritz.Häcker@Student.Reutlingen-University.de

Abstract
Das Bim Collaboration Format ist ein Standard für die Kommunikation von Planern im BIM-Prozess. In einer Analyse wird untersucht, wie das Format aufgebaut ist und welche Anwendungsmöglichkeiten es bietet. Darauf aufbauend wird durch eine Expertenumfrage dargelegt, wie das Format im Projektumfeld funktioniert und welche Schwachstellen und Schwierigkeiten dabei auftreten. Daraus werden Handlungsempfehlungen für zukünftige BCF-Softwareentwicklungen abgeleitet.

CCS Concepts
• **Software and its engineering → Software creation and management → Software verification and validation → Empirical software validation**

Keywords
Building Information Modelling, Building Collaboration Format

Betreuer Hochschule: Prof. Dr. Gabriela Tullius
Hochschule Reutlingen
Gabriela.Tullius@Reutlingen-University.de
Betreuer Extern: Matthias Aust
Fraunhofer IAO
Matthias.Aust@iao.Fraunhofer.de
Informatics Inside Frühjahr 2022
11. Mai 2022, Hochschule Reutlingen
Copyright 2022 Moritz Häcker

1 Einleitung
Die Nutzung von Building Information Modelling (kurz BIM) als Planungsstandard für Bauprojekte erfährt immer mehr Verbreitung. Während zu Beginn der 2000er Jahre nur wenige Bauprojekte mit BIM durchgeführt wurden, wird heute bei ca. 70% aller Bauprojekte mindestens BIM Level 0 genutzt. Seit 2020 ist BIM bei allen öffentlichen Bauprojekten, Bundesbauprojekten und Infrastrukturprojekten verpflichtend.[1]

Bei BIM-Projekten wird während der Planung, für den Austausch der 3D-Modelle zwischen den verschiedenen Planern, das Industry Foundation Classes (kurz IFC) Schema verwendet. Der IFC-Standard wird von *buildingSMART* seit 1997 entwickelt und ist ein herstellerneutrales, offenes Dateiformat zum Austauschen von 3D Modellen zwischen verschiedenen Planern.[2] 2013 ist die aktuelle IFC-Version 4 erschienen, welche heute der Standard ist.[3]

Die Grundlage des IFC-Schemas bildet das Standard for the Exchange of Product model data (kurz STEP- Dateiformat) in Kombina-

[1] https://www.planradar.com/de/bim-europavergleich. Besucht am 04.04.2022.

[2] https://technical.buildingsmart.org/standards/bcf/. Besucht am 04.04.2022

[3] https://www.baunetzwissen.de/bim/fachwissen/standardisierung/ifc-der-offene-standard-fuer-bim-modelle-5288161. Besucht am 04.04.2022

tion mit der objektorientierten Datenmodellierungssprache *Express*. Express enthält eine Vielzahl von Klassen, die Attribute und Beziehungen zu anderen Klassen haben können. In Express werden Klassen, *Entitytypen* genannt. Jeder Entitytyp abstrahiert ein Objekt der realen Welt, bspw. ein Fenster, eine Wand oder ein anderes Bauteil. So ist ein BIM-Modell ein digitales Abbild des sich in Planung oder Bau befindenden Gebäudes, welches alle Bauteile als IFC-Modelle mit ihren verschiedenen Parametern beinhaltet. Jedes dieser Bauteile, welche als IFC-Modelle in einem Projekt bearbeitet werden, hat neben seinen Klassen und Attributen einen festen Global Unique Identifier (kurz GUID), welcher projektweit zur Identifizierung und Nachverfolgung genutzt wird. Dieser GUID ist für die Kommunikation mit dem Bim Collaboration Format (kurz BCF) wichtig, da BCF-Topics von ihm einem oder mehreren IFC-Modellen zugewiesen werden können [1]. Wie eine BCF-Datei aufgebaut ist und welche Attribute ein Topic besitzt, wird in Kapitel 2 erklärt.

Alle beteiligten Planer, die mit BIM-Standards arbeiten, können ein einheitliches Format zur Kommunikation verwenden. Die gesamte Kommunikation, welche sich um spezielle Bauteile oder sonstige im 3D-Modell verortbare Inhalte dreht, kann über das BCF abgewickelt werden.

1.1 Ziele der Arbeit

Ziel der Arbeit ist es, den BCF-Standard bezüglich seines Aufbaus und seiner Nutzung im realen Projektumfeld zu analysieren, um daraus Handlungsempfehlungen und Mehrwerte für die Nutzung des Standards in Virtual Reality (kurz VR) abzuleiten. Neben der Analyse wird im weiteren Schritt eine Umfrage mit Experten zu ihren Erfahrungen mit der Kommunikation mittels BCF durchgeführt.
Dabei soll der Fokus auf die Limitierungen und Schwächen des Standards gelegt werden. Die befragten Planer, also die Nutzer des BCF-Standards, werden in den Interviews ebenfalls gebeten, Wünsche bezüglich noch zu entwickelnder Funktionen zu äußern. Mit den Aussagen aus den Experteninterviews sowie der Literaturrecherche sollen dann Handlungsempfehlungen abgeleitet werden. Diese dienen als Grundlage für die darauf aufbauende Masterthesis, welche sich damit beschäftigen wird, wie das BCF in VR genutzt werden kann. Im Rahmen der Masterthesis soll folgende Fragestellung untersucht werden: Welche Funktionen und die dazugehörigen Interaktionsmethoden braucht es, um effizient in VR mit dem BCF zu arbeiten?

2 Analyse

Während der Analyse hat sich gezeigt, dass die Diskussion über Arbeitsweisen und Details rund um das BCF aktuell weniger auf wissenschaftlichen Konferenzen und in Veröffentlichungen stattfindet. Stattdessen finden sich eher Artikel in Fachmagazinen sowie Beiträge auf Online Blogs von Softwareherstellern. Dieser Umstand hat sicherlich damit zu tun, dass das BCF ein offener Standard ist. Aus diesem Grund sind für die Analyse des Standarts neben wissenschaftlichen Artikeln auch einige Internetquellen referenziert.

2.1 BIM Collaboration Format

In diesem Kapitel wird geklärt, was das BIM Collaboration Format ist und wie es funktioniert. Das BIM Collaboration Format ist ein Kommunikations-Austauschformat für die BIM-basierte Gebäudeplanung. Der Standard wurde von *Solibri* und *Tekla* in Zusammenarbeit mit dem *Institut für angewandte Bauinformatik der Hochschule München* im Jahr 2009 konzipiert. Später entwickelte *buildingSMART* daraus einen internationalen openBIM Standard. Ziel des Dateiformates

ist der offene, herstellerneutrale Kommunikationsaustausch zwischen IFC-basierter Software.[4]

Mit dem BCF lässt sich ein Großteil der Planungskommunikation abwickeln. Dies ist möglich, weil das Format herstellerneutral ist und von verschiedensten Fachplanern direkt in ihren Planungswerkzeugen eingesetzt werden kann. Somit ist weder das Schreiben von E-Mails noch Telefonabsprachen nötig. Ebenfalls entfällt das aufwendige Erstellen und Bearbeiten von Änderungstabellen [3]. Ein weiterer Vorteil ist, dass Planer keine 3D-Geometriedaten zum Zwecke der Kommunikation versenden müssen.[5] Es reicht, wenn die BCF-Datei verschickt wird. Die Planer bekommen für ihre jeweilige Planung ein Referenzmodell im IFC-Format. Dieses Referenzmodell enthält alle Modelle, welche sich mittels ihres GUID nachverfolgen lassen. Durch diese GUIDs können BCF-Topics gezielt mit Bauteilen verknüpft werden. Mit einem BCF-Topic lassen sich z. B. Probleme, Aufgaben, Fragen oder Änderungswünsche standardisiert abbilden [3]. In einer Untersuchung von 18000 Topics fand der Softwarehersteller *BIMTrack* heraus, dass Kommentar (19%), Anfrage (19%), Lösung (19%), Issue (17%), Frage (3%), Andere (4%), Kein Typ angegeben (19%) die am meisten genutzten Typen von Topics im BCF sind.[6]

Durch die Standardisierung lassen sich BCF-Topics sehr gut automatisiert erstellen. Dies geschieht bspw. bei einer Kollisionserkennung zweier 3D-Modelle. Bei einer Kollisionserkennung werden verschiedene 3D- oder Teilmodelle auf Kollisionen überprüft. Tritt eine Kollision auf, wird automatisch ein BCF-Topic erstellt, welches alle nötigen Informationen über die Problematik enthält [6]. Wie eine BCF-Datei aufgebaut ist, wird im Folgenden erklärt.

Es gibt zwei verschiedene Herangehensweisen, wie BCF-Dateien ausgetauscht werden können: mittels Datei-basiertem Austausch oder mit Hilfe eines Webservers via Application Programming Interface (kurz API). Bei dem Datei-basierten Austausch wird die BCF-Datei (.bcfzip oder .bcf) zwischen Planern hin und her geschickt. BCF-Dateien werden in der Planungssoftware eines Planers erstellt, exportiert sowie entsprechend der Zuständigkeit weitergegeben. Ein anderer Planer kann diese Datei dann in seine fachspezifische Planungssoftware laden, dort auf die Inhalte zugreifen und die nötigen Änderungen vornehmen. Wenn Änderungen vorgenommen wurden, muss die Datei wieder zurückgeschickt werden. Dabei ist darauf zu achten, dass nicht mehrere Kopien einer BCF-Datei mit unterschiedlichen Planungs- und Kommunikationsständen im Umlauf sind. Um dieses Problem zu umgehen, gibt es einen Arbeitsablauf, bei dem auf das Verschicken der BCF-Datei verzichtet werden kann. Dafür wird eine Web-basierte (RESTful) API Schnittstelle implementiert. Dort werden die Inhalte der BCF-Kommunikation zentral gespeichert. Somit kann sichergestellt werden, dass alle Planer mit dem gleichen Planungs- bzw. Kommunikationsstand arbeiten [6].

2.2 Einsatz in Planungswerkzeugen

Damit in allen Planungswerkzeugen die BCF-Topics gleich dargestellt werden ohne das wichtige Information zwischen Export aus der einen Software und Import bei einer

[4] https://technical.buildingsmart.org/standards/bcf. Besucht am 04.04.2022

[5] https://bimcorner.com/bim-collaboration-format-whatsapp-of-bim. Besucht am 04.04.2022

[6] https://bimtrack.co/blog/blog-posts/why-issue-tracking-is-about-more-than-clashes. Besucht am 04.04.2022

anderen Software, verloren gehen, ist es erforderlich, dass Softwarehersteller den Standard richtig implementieren. Dies untersuchten Treldal et. al [4] unter anderem in ihrer Studie bei sieben gängigen Softwarelösungen und Plug-Ins, welche das BCF-Standard einlesen und schreiben können. Bei keinem der von ihnen untersuchten Software-Werkzeugen war die exportierte BCF-Datei von der Spezifikation her komplett korrekt. In einigen Fällen fehlten sogar die sog. nicht-optionalen Attribute und Elemente in den BCF-Dateien. Die optionalen Attribute wurden bei den wenigsten Softwareherstellern umgesetzt. Fehlen etwa gewisse Parameter oder Attribute in den XML-Dateien des BCF, führt dies zum Verlust von Information auf dem Weg von einer Software zur anderen. Im schlechtesten Fall kann dann die BCF-Datei nicht eingelesen werden.

2.3 Aufbau des BCF-Dateiformats

Im nächsten Schritt soll geklärt werden, aus welchen Komponenten das BCF-Dateiformat besteht. Eine BCF-Datei besteht im Grunde aus mehreren Extensible Markup Language (kurz XML-)Dateien, welche in einer vorgegebenen Ordnerstruktur abgelegt sind. Der oberste Ordner, auch *root Ordner* genannt, ist im ZIP-Format mit der Dateiendung .bcf oder .bcfzip verpackt. Wird das ZIP-Archiv entpackt, befinden sich dort mehrere XML-Dateien. Die Datei *bcf.version* gibt Aufschluss über die verwendete BCF-Versionsnummer. In der Datei *project.bcfp* sind alle Metainformationen zum BCF-Projekt gespeichert. Die *extensions.xml* enthält alle extra hinzugefügten Schemata wie Farbschemata oder zusätzliche Typen. Auf derselben Ebene im ZIP-Archiv können sich noch beliebig viele Ordner mit einem GUID als Namen befinden. Jedes dieser Verzeichnisse ist ein BCF-Topic (Problem, Frage, Kommentar oder Ähnliches), der Verzeichnisname entspricht der Topic-ID. In einem Topic-Verzeichnis befinden sich in der Regel drei XML-Dateien: eine *Markup.bcf*, eine *Viewpoint.bcfv* und eine PNG-Datei mit dem Namen *Snapshot*.

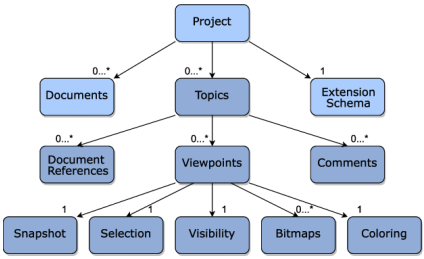


Abbildung 1: Struktur des BCF

Quelle: Schulz, O., Oraskari, J., & Beetz, J. (2021). *bcfOWL: A BIM collaboration ontology*.

In der *Markup.bcf* sind die grundlegenden Informationen über das Topic festgeschrieben. Dort finden sich etwa Name, Beschreibung, Datum und Name des Autors sowie der Name des zugewiesenen Verantwortlichen oder die Priorität des Topics. *Viewpoint.bcfv* enthält alle Informationen über die Bauteile, welche mit dem Topic verknüpft sind, sowie Information darüber, ob die Bauteile ein- oder ausgeblendet werden sollen. Ebenso ist die exakte Position der Kamera, von der aus der Snapshot erstellt wurde, in der *Viewpoint.bcfv* gespeichert. Die Kameraposition ermöglicht es, dass Nutzer in anderer Software genau an die Position im Modell bewegt werden können, von der aus es möglich ist, die Problematik des Topics nachzuvollziehen. Der Snapshot dient als Vorschaubild, damit Topics schneller wiedergefunden werden können [1, 2].

2.4 Einsatz des BCF

Es stellt sich nun die Frage, in welchem Kontext das BCF zum Einsatz kommen sollte. Laut der *buildingSMART*-Dokumentation könnte das BCF nicht nur in allen Phasen des Planens und Bauens, sondern auch nach der Fertigstellung zum Einsatz kommen. Es könnte zunächst in der Design Phase, bspw. zur Dokumentation der Qualitätssicherung

von BIM-Modellen, zur Kollisionserkennung und zum Annotieren von Design-Optionen eingesetzt werden. Während der Beschaffungsphase empfiehlt sich seine Nutzung wiederum zur Ausschreibungskoordination und Planung von Kosten- und Lieferanteninformationen für spezielle Baugruppen oder Systeme. In der Bauphase hingegen kann es bei Qualitätskontrollen nutzen, um den Planungsstand mit dem Baustand abzugleichen. Im Betrieb könnte das BCF außerdem genutzt werden, um Veränderungen am Gebäude zu dokumentieren sowie um Umbaumaßnahmen zu planen.[7]

3 Einsatz im Projektumfeld

Nachdem die Recherche das BCF aus theoretischer Sicht behandelte, wird zur Entwicklung von Handlungsempfehlungen der Einsatz des BCF-Standards im Projektumfeld analysiert. Dazu wird gemäß dem DECIDE-Framework [5] eine Expertenumfrage konzipiert und durchgeführt.

3.1 Konzeption und Durchführung der Interviews

Es wurden acht Personen in verschiedenen Planungsbüros angeschrieben, wovon fünf einer Teilnahme zustimmten. Ein Interview-Partner verweigerte die Unterzeichnung der Einverständniserklärung, weshalb das Interview nicht aufgezeichnet werden konnte. Einige generelle Aussagen konnten aber trotzdem in die Auswertung einfließen. Da die Teilnehmerzahl sehr klein ist, wurden bewusst Experten aus verschiedenen Bereichen befragt, um trotzdem ein möglichst breites Bild mit verschiedenen Blickwinkeln auf die Thematik zu ermöglichen. Zu den Interviewteilnehmern gehören Architekten, Tragwerksplaner, Fachplaner, Baufirmen und Immobiliendienstleister. Aus Gründen des Datenschutzes wird nicht näher auf die Interviewteilnehmer eingegangen. Nach den Interviews wurden die Aufzeichnungen transkribiert und in tabellarischer Form zur Auswertung gebracht.

3.2 Fragen

Um das Ziel der Umfrage – herauszufinden, wie Planer das BCF in der Praxis nutzen und welche Probleme ihnen dabei begegnen – erreichen zu können, wurde ein Fragenkatalog formuliert. Die Fragen dienten als Leitfaden für jeden Interviewpartner. Bei den meisten Interviews entwickelte sich nach den ersten Fragen ein interessantes, offenes Gespräch über die BCF-Thematik.

Zu Beginn wurde nach formellen Angaben gefragt:

- Wer sind Sie?
- Was tut die Firma, für die Sie arbeiten?
- Was tun Sie bei der Firma, für die Sie arbeiten?

Danach wurde direkt und spezifisch nach dem BCF-Standard gefragt:

- Wird das BCF-Format in BIM-Projekten in der Firma, in der Sie arbeiten, genutzt?
- Wie sieht ein typischer Arbeitsablauf aus, bei dem BCF-Topics bearbeitet werden?
- Welche Softwarewerkzeuge werden für die Arbeitsabläufe verwendet?
- Für was genau werden BCF-Topics genutzt?
- Wie wird mit einer sehr großen Zahl von Topics umgegangen (bspw. nach Kollisionserkennung)?

[7] https://technical.buildingsmart.org/standards/bcf. Besucht am 04.04.2022

- Welche Schwierigkeiten fallen beim Arbeiten mit BCF auf? Was fällt positiv auf?
- Wie zeitaufwendig ist das Abarbeiten der BCF-Topics?

Neben den Fragen zur alltäglichen Arbeit mit BCF wurden außerdem Fragen zu Erfahrungen mit Virtual-Reality-Geräten gestellt. Ebenso, ob sich die Teilnehmer vorstellen könnten, gewisse Tätigkeiten beim Arbeiten mit BCF-Topics in einer Virtual Reality-Umgebung mit einem Head-Mounted-Display (kurz HMD) auszuführen.

- Haben Sie bereits Erfahrungen mit VR gemacht?
- Könnten Sie sich vorstellen, die Bearbeitung von BCF-Topics in VR zu machen?

3.3 Ablauf

Die Teilnehmer wurden per E-Mail kontaktiert und um einen Interview-Termin gebeten. Bei einer Zusage wurde vorab eine Einverständniserklärung unterzeichnet. Das Interview wurde für Transkriptionszwecke mit Bild und Ton aufgezeichnet. Aufgrund der Corona-Pandemie und langen Reisewegen wurden die Interviews nicht persönlich, sondern per Microsoft Teams durchgeführt. Um eventuellen Audioproblemen vorzubeugen und die spätere Transkription zu erleichtern, wurden den Interviewteilnehmern die Fragen mittels Präsentationsfolien per Bildschirmübertragung präsentiert. Um nicht zu viel Arbeitszeit der Interviewteilnehmer in Anspruch zu nehmen und damit die Bereitschaft einer Teilnahme zu erhöhen, wurde die Dauer des Interviews auf 30 Minuten festgelegt.

4 Ergebnisse

Da sich während den Interviews meist ein Gespräch entwickelte und Fragen dabei nicht zwangsweise in der vorgesehenen Reihenfolge beantwortet wurden, werden die Aussagen der Teilnehmer im Folgenden als Hauptthemen der Fragen zusammengefasst.

4.1 Teilnehmer und Firmen

Die Teilnehmer sind Mitarbeiter in einem Architektenbüro, Ingenieurbüro, Technischer Gebäudeausstattung, Immobilienentwicklung sowie ein Bauunternehmen im Hoch-Tiefbau. Sie alle haben Erfahrungen mit BIM und nutzen bei ihrer Tätigkeit aktiv das Austauschformat BCF.

4.2 Wann wird BCF verwendet?

Planer aus der Architektur und Tragwerksplanung sagten, dass „[…] es vor Leistungsphase 3 nach HOAI[8] kein Sinn macht [mit BCF-Topics zu arbeiten]“. Der Grund hierfür sei, dass sich die Dinge in den Phasen davor noch so häufig und grundlegend ändern, dass BCF-Topics nicht sinnvoll nutzbar seien. Außerdem würden sich die IFC-Modelle und damit verbunden auch der GUID häufig komplett austauschen, wodurch keine sinnvolle Nachverfolgung möglich sei. Ab Leistungsphase 3 nach HOAI sei es laut den Befragten sinnvoll, BCF-Topics zur Kommunikation einzusetzen. Leistungsphase 3 beschreibt die Entwurfsplanung, welche nach der Grundlagenermittlung und Vorplanung kommt. Die Teilnehmer verwendeten die gängigen Softwarewerkzeuge, die auf dem Markt für das entsprechende Gewerk erhältlich sind, wie bspw. *Revit*, *Navisworks*, *AutoCAD*, *Solibri*, *Rhino*, *Grasshopper*, *Sophistik*. Als reiner Modellbetrachter wurde *BIMCollabZoom* von mehreren Teilnehmern genannt.

[8] Honorarordnung für Architekten und Ingenieure

Alle Teilnehmer merkten an, dass projektabhängig, je nach Bauherrn, verschiedene Kollaborationsplattformen und Common-Data-Environment (CDE)-Plattformen verwendet würden. Dazu gehören z. B. *BimSync*, *revizto* oder *Bim360*. Über diese Plattformen lässt sich die Kommunikation für alle Projektpartner einfach an einem Ort zu bündeln, was erheblich Zeit spart. Das Erlernen der vielen verschiedenen Plattformen wurde von den Teilnehmern als herausfordernd und zeitintensiv empfunden.

4.3 Arbeitsablauf mit BCF

Die Teilnehmer gaben an, dass sie BCF-Topics an ihrem Desktop-Computer mit der Software des entsprechenden Gewerkes bearbeiten. Durch das „AssignedTo“-Feld eines BCF-Topics sei sehr schnell klar, was wer zu tun hat, sagt eine Interviewteilnehmerin.

Die reine Bearbeitungszeit, ohne die evtl. erforderliche inhaltliche Abstimmung mit Kollegen oder anderen Planern, wurde von allen Teilnehmern auf weniger als 5 Minuten geschätzt. Mit etwas Übung brauche ein erfahrener Planer sogar weniger als 2 Minuten pro BCF-Topic, um dieses zu bearbeiten, kommentieren oder zu lösen. Alle Teilnehmer gaben an, dass der Austausch mit BCF deutlich schneller sei, als das Verfassen einer E-Mail. Die Kommentarfunktion im BCF wird von allen Teilnehmern genutzt und es finden dort auch Diskussionen statt. Mehrere Teilnehmer merkten aber auch an, dass der persönliche Austausch sehr wichtig sei und man auch „[…] einfach mal den Hörer in die Hand nehmen und anrufen muss“, um bei manchen Problemen schnell zu einer Lösung zu kommen.

Wenn es nach einer Kollisionskontrolle eine sehr große Anzahl an BCF-Topics gibt, weil sich bspw. mehrere Bauteile überschneiden, haben die Planer verschiedene Strategien, die große Anzahl an Topics abzuarbeiten. Als erstes werde damit begonnen, die BCF-Topics in verschiedene Kategorien zu ordnen. Diese Kategorien seien hierarchisch gegliedert, je nachdem wie kritisch die Problematik für die weitere Planung oder Ausführung ist. Kategorien, die genannt wurden, sind: blockierend, dringend, normal, unkritisch, pausiert. Danach werden denjenigen Mitarbeitern, die für einen gewissen Bereich im Gebäude zuständig sind, die Aufgaben zugewiesen. Die verschiedenen Planungsbüros haben im Detail verschiedene Vorgehen. Die Unterteilung in verschiedene Kategorien und die Priorisierung lassen sich aber häufiger wiederfinden.

4.4 Schwierigkeiten und fehlende Funktionen

Eine Teilnehmerin wies darauf hin, dass es sowohl für einen 2D-Grundriss als auch für ein 3D-Modell hilfreich wäre, die Verortung von BCF-Topics zu sehen

Mehrere Planer äußerten sich unzufrieden mit den Möglichkeiten, BCF-Topics zu kategorisieren, um sie später sinnvoll filtern und wiederfinden zu können. Laut den Teilnehmern sollte es neue Schemata zum Verknüpfen von BCF-Topics mit Gebäudeteilen, Stockwerken oder Räumen geben. Denn bei großen Bauprojekten mit mehreren Gebäuden sind mehrere tausend BCF-Topics keine Seltenheit. Deshalb braucht es bessere Filter, Kategorien und Label, damit sich Topics schnell wieder finden lassen. Planer nannten in diesem Zusammenhang auch die fehlende Funktionalität des Zusammenfassens von BCF-Topics. Oftmals würden viele BCF-Topics an einem übergeordneten Problem hängen. Werde dieses gelöst, lösen sich auch die anderen. Die Möglichkeit, diese Abhängigkeit darzustellen, ist aktuell nicht vorhanden.

4.5 Erfahrungen mit Virtual Reality

Alle Teilnehmer gaben an, erste Erfahrungen mit Virtual Reality gemacht zu haben. Diese beschränken sich allerdings auf kurzes Testen und Ausprobieren, bspw. an Messeständen oder sonstigen Demonstratoren. Drei

Teilnehmer merkten an, VR schon zu Testzwecken und kleineren Demonstrationen für Kunden verwendet zu haben. Im alltäglichen Arbeitsablauf und speziell bei der Arbeit mit BCF kann sich kein Teilnehmer vorstellen, dass VR das Potenzial hat, einen Vorteil im Arbeitsalltag zu bieten.

Der Grund hierfür sei, dass das Abarbeiten der BCF-Topics am Desktop-Monitor schnell funktioniere und die Werkzeuge der verschiedenen BIM-Plattformen und Software Tools (BIM Viewer, CAD-Software) ausreichend Funktionen hätten, mit denen Gebäudeteile ausreichend übersichtlich dargestellt werden können. Für das Einfügen von Schnittebenen oder das Einfärben von Bauteilen seien „[…] für die Besprechung kurzer Zwischenstände (sind) Modellviewer besser geeignet [als VR Geräte]“. Im Zusammenspiel mit einer Multi-User-Plattform kann sich ein Teilnehmer vorstellen, dass VR einen großen Vorteil bieten könnte, da verschiedene Planer gleichzeitig im 3D-Modell über Probleme abstimmen und diskutieren können. Dies wäre eine Neuerung und könnte einige Meetings ersetzen bzw. die Effizienz von Absprachen erhöhen. Dafür müssten die Anforderungen an den Aufbau und die Inbetriebnahme der VR-Geräte jedoch noch sinken, da Mitarbeiter in Planungsbüros in der Regel keine Erfahrung damit haben.

Die Teilnehmer gaben – mit einer Ausnahme – an, dass sie keinen konkreten Nutzen im VR-Bereich zur Bearbeitung von BCF-Topics sehen. Diese generelle Aussage könnte damit zusammenhängen, dass die Befragten nur sehr wenig bis gar keine Erfahrung mit VR haben. Denn ohne einen gewissen Erfahrungswert über die Darstellungs- und Interaktionsmöglichkeiten ist es schwer, sich neue Arbeitsweisen und Wege im Zusammenhang mit BCF vorzustellen.

5 Handlungsempfehlungen

Wie durch die Auswertung der Interviews mit Experten verschiedener Branchen gezeigt werden konnte, ist das BCF ein hilfreicher Standard für die Kommunikation in BIM-Bauprojekten. Durch die Experteninterviews konnte ein Bedarf an Weiterentwicklungen identifiziert werden.

Erstens kann generell – aufgrund mehrerer Aussagen – festgehalten werden, dass es Bedarf für die Entwicklung von zusätzlichen Darstellungen der BCF-Topics gibt. Dies gilt sowohl für die Darstellung als 2D-Liste als auch für die Darstellung im 3D-Modell, wo Topics räumlich verortet werden könnten. Aktuell gibt es nur die Möglichkeit, die Topics als 2D-Liste in den Software-Tools zu sehen.

Zweitens wären zusätzliche Filter-Werkzeuge zur schnellen und effizienten Topic-Findung hilfreich. Denn hat ein Projekt sehr viele Topics, reichen die vorhandenen Filter-Werkzeuge nicht aus. Hier wären zusätzliche Kategorien und Parameter sinnvoll, um besser filtern zu können.

Des Weiteren bestätigen die Interviewergebnisse die Ergebnisse von Treldal et al. [4]: Softwarehersteller setzen die BCF-Spezifikationen nur teilweise und nicht ausreichend um. Dadurch kann es zu einem Informationsverlust bei der Übertragung von einer Planungs-Software zur anderen kommen. Deshalb sollten Softwarehersteller in Zukunft noch mehr darauf achten, dass die Spezifikationen vollständig umgesetzt sind, damit der Standard mit all seinen Eigenschaften Herstellerübergreifend genutzt werden kann.

Zuletzt soll auf den Bedarf der Entwicklung von kollaborativen VR-Anwendungen, die das Bearbeiten von BCF-Topics ermöglichen, hingewiesen werden. Denn derzeit sind solche kaum vorhanden. Die Auswertung der Umfrageergebnisse hat – trotz weniger Erfahrung mit VR seitens der Interviewten – gezeigt, dass kollaborative Anwendungen wünschenswert sind.

Der Mehrwert eines – um VR erweiterten – Planungswerkzeugs für BCF sollte daher in weiteren Arbeiten untersucht werden, da diese noch viel Potential für zukünftige Entwicklungen bieten.

6 Literaturverzeichnis

[1] André Borrmann, Markus König, Christian Koch, Jakob Beetz. 2015. *Building Information Modeling. Technologische Grundlagen und industrielle Praxis*. VDI-Buch. Springer Vieweg, Wiesbaden. ISBN: 978-3-658-05605-6.

[2] Lindhard, K. and Steinmann, R. 2015. BIM-collaboration processes - from fuzziness to practical implementation. In *eWork and eBusiness in architecture, engineering and construction. Proceedings of the 10th European Conference on Product and Process Modelling (ECPPM 2014), Vienna, Austria, 17-19 September 2014*. A Balkema book. CRC Press, Boca Raton, Fla., 27–31.

[3] Marian Behaneck. 2021. BIM mit Fachplanern Teil 2: Austauschformat BCF. *DAB Deutsches Architekten Blatt*, 11/2021. https://www.dabonline.de/2021/11/17/bim-servermodell-fachplaner-teil2-austauschformat-bcf-cloud/. Aufgerufen 04.04.2022.

[4] Niels Treldal, Hussain Parsianfar, Jan Karlshøj. 2016. Using BCF as a mediator for task management in building design. In *Proceedings of the International RILEM Conference Materials, Systems and Structures in Civil Engineering 2016. BIM in Civil Engineering – Open Data Standards in Civil Engineering*, 48–59.

[5] Preece, J., Rogers, Y., and Sharp, H. 2002. *Interaction design. Beyond human-computer interaction*. Wiley, New York, NY. 978-0471492788.

[6] Tobias Döring. 2019. BCF – endlich verständlich. Datenformat BCF. In *Build-Ing 1/2019*, Huss-Medien, Ed., Berlin, 48–54.

©2022 Moritz Häcker. 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/

Autorenverzeichnis

B

Bajraktari, E. 60

Bresler, R. 1

G

Grendatsch, D. 116

H

Häcker, M. 139

K

Krautmacher, L. 26

Kruck, H. 37

L

Lauxmann, T. 14

M

Moysidis, D. 89

Müller, K. 48

S

Stadelmaier, M. 128

U

Uslu, I. 74

W

Wickert, F. 100

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
Website: infoinside.reutlingen-university.de

ISBN
978-3-00-072060-4