# Technologiegestützte, prozessparallele Erfassung von Erfahrungswissen in manuellen Produktionsprozessen von KMU

# Technology-Supported, Process-Parallel Capture of Experiential Knowledge in Manual Production Processes of SMEs

Von der Fakultät für Maschinenwesen
der Rheinisch-Westfälischen Technischen Hochschule Aachen
zur Erlangung des akademischen Grades eines
DOKTORS DER INGENIEURWISSENSCHAFTEN
genehmigte Dissertation

vorgelegt von

Yannick Dassen

Berichter/in: Univ.-Prof. Dr.-Ing. Christian Brecher

Univ.-Prof. Dr.-Ing. Verena Nitsch

Tag der Mündlichen Prüfung: 03.09.2026

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek online verfügbar

.

# Vorwort

Die vorliegende Dissertation entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Werkzeugmaschinenlabor WZL der Rheinisch-Westfälischen Technischen Hochschule Aachen.

Zunächst bedanke ich mich bei Prof. Dr.-Ing. Christian Brecher für die Möglichkeit zur Promotion an seinem Lehrstuhl sowie die damit verbundene Betreuung. Prof. Dr.-Ing. Verena Nitsch, Inhaberin des Lehrstuhls für Arbeitswissenschaften, danke ich für die Übernahme des Koreferats. Weiterhin möchte ich mich beim Inhaber des Lehrstuhls für Lasertechnik, Prof. Dr. rer. nat. Constantin Leon Häfner, für die Übernahme des Vorsitzes in der Promotionsprüfung bedanken.

Den Kolleginnen und Kollegen des Lehrstuhls für Werkzeugmaschinen sowie insbesondere der Abteilung Automatisierung, Robotik und Steuerungstechnik danke ich für die angenehme, inspirierende und freundschaftliche Zusammenarbeit. Meinen Oberingenieuren Dr.-Ing. Simon Storms und Dr.-Ing. Oliver Petrovic danke ich für das entgegengebrachte Vertrauen, die stete Unterstützung und die gewährten Freiräume. Stefan Wein, Katrin Schilling, Philipp Blanke, Josefine Monnet, Christoph Susen und Petar Tesic danke ich für die fachlichen Diskussionen und Denkanstöße bei der Themenfindung und Ausarbeitung meiner Dissertation. Mein Dank gilt ebenfalls den Studentinnen und Studenten, die mich im Zuge ihrer Abschlussarbeiten und Tätigkeit als studentische Hilfskräfte bei der Erarbeitung meiner Forschungsergebnisse unterstützt haben. Für zahlreiche Diskussionen in meinem Themenfeld und wertvolle Einblicke in die konkreten Probleme der Industrie danke ich allen Projektpartnern des Projektes „AKzentE4.0“.

Mein besonderer Dank gilt meinen Eltern Silvia Detro-Dassen und Gerd Dassen, die mir diesen Lebensweg ermöglicht haben. Euer Rückhalt und Vertrauen in mich waren das Fundament für die Entstehung dieser Arbeit.

Den größten Dank möchte ich meiner Frau Sonja aussprechen, die mir bedingungslos den Rücken während der Ausarbeitung der Dissertation und insbesondere während der Vorbereitung auf die Prüfung freigehalten hat. Du und Jonah wart der Lichtblick an langen Tagen am Schreibtisch und habt mir die Kraft gegeben, diese Arbeit zu leisten.

Aachen, September 2026 Yannick Dassen

# Kurzzusammenfassung

Der gezielte Einsatz von Wissensmanagement zur Sicherung und Verstetigung impliziten Wissens ist ein wesentlicher Faktor für die Wettbewerbsfähigkeit deutscher Unternehmen. Besonders kleine und mittlere Unternehmen (KMU) stehen vor der Herausforderung, dass relevantes Erfahrungswissen meist nicht systematisch dokumentiert ist, sondern in den Köpfen der Mitarbeitenden verbleibt. Beim Ausscheiden erfahrener Fachkräfte geht dieses Wissen häufig verloren. Bestehende kommerzielle Wissensmanagementsysteme sind für den Einsatz auf dem Shopfloor produzierender KMU nur eingeschränkt geeignet, da sie primär auf die Darstellung von Wissen fokussiert sind und eine intuitive, effiziente Erfassung kaum unterstützen. Daraus ergibt sich der Bedarf an neuen, auf KMU zugeschnittenen methodischen und technologischen Ansätzen.

Ziel der vorliegenden Arbeit ist daher die Untersuchung und Entwicklung innovativer Konzepte zur Integration von Wissensmanagement auf dem Shopfloor von KMU. Mitarbeitenden soll ermöglicht werden, ihr Erfahrungswissen prozessnah, intuitiv und mit geringem Zusatzaufwand digital zu erfassen, zu kontextualisieren und unternehmensweit verfügbar zu machen. Methodisch basiert die Arbeit auf einer literaturgestützten Analyse des Stands der Technik, einer empirischen Anforderungsanalyse, der Entwicklung und Implementierung neuer Konzepte sowie deren technischer Validierung und Evaluation.

Die Anforderungsanalyse zeigt, dass die größte Herausforderung in der Erfassung und strukturierten Dokumentation impliziten Wissens liegt. Aufbauend darauf wird das Prinzip der prozessparallelen Erfassung von Erfahrungswissen eingeführt, bei dem Wissensmanagement direkt in bestehende Arbeitsabläufe integriert wird. Es werden drei Konzepte entwickelt: aktiver, semi-aktiver und passiver Wissenstransfer. Der aktive Ansatz nutzt einen Chatbot zur gezielten Erfassung und Abfrage von Wissen, der semi-aktive Ansatz setzt auf Audio- und Videoaufzeichnungen erfahrener Mitarbeitender zur automatisierten Erstellung von Schulungsinhalten, während der passive Ansatz kamerabasierte Handbewegungsanalysen und Process Mining zur Ableitung von Prozesswissen verwendet. Alle Konzepte werden als On-Premise-Lösungen implementiert und im Hinblick auf ihre grundsätzliche Eignung für den Einsatz in KMU untersucht. Dabei erfolgen eine technische Betrachtung sowie eine Einordnung der Ansätze in Bezug auf die zuvor identifizierten Anforderungen. Abschließend werden die Konzepte vergleichend gegenübergestellt und im Kontext der Arbeit diskutiert. Obwohl die Ansätze die Grundanforderungen erfüllen, erfordert der Praxiseinsatz noch die Behebung technischer Mängel sowie eine höhere Benutzerfreundlichkeit.

# Brief Summary

The targeted use of knowledge management to secure and perpetuate implicit knowledge is a key factor in the competitiveness of German companies. Small and medium-sized enterprises (SMEs) in particular face the challenge that relevant experiential knowledge is usually not systematically documented but remains in the minds of employees. When experienced specialists leave the company, this knowledge is often lost. Existing commercial knowledge management systems are only of limited suitability for use on the shop floor of manufacturing SMEs, as they focus primarily on the presentation of knowledge and provide little support for intuitive, efficient recording. This results in a need for new methodological and technological approaches tailored to SMEs.

The aim of this thesis is therefore to investigate and develop innovative concepts for integrating knowledge management on the shop floor of SMEs. Employees should be enabled to digitally capture and contextualize their experiential knowledge in a process-oriented, intuitive manner with little additional effort and make it available company-wide. Methodologically, the work is based on a literature-based analysis of the state of the art, an empirical requirements analysis, the development and implementation of new concepts and their technical validation and evaluation.

The requirements analysis shows that the greatest challenge lies in the capture and structured documentation of implicit knowledge. Building on this, the principle of process-parallel capture of experiential knowledge is introduced, in which knowledge management is integrated directly into existing workflows. Three concepts are developed: active, semi-active and passive knowledge transfer. The active approach uses a chatbot for targeted knowledge capture and retrieval, the semi-active approach relies on audio and video recordings of experienced employees for the automated creation of training content, while the passive approach uses camera-based hand movement analysis and process mining to derive process knowledge. All concepts are implemented as on-premise solutions and examined with regard to their fundamental suitability for use in SMEs. This involves a technical assessment and classification of the approaches in relation to the previously identified requirements. Finally, the concepts are compared and discussed in the context of the thesis. Although the approaches meet the basic requirements, practical implementation still requires the resolution of technical shortcomings and improved user-friendliness.

# I. Inhaltsverzeichnis

*Table of Contents*

**I. Inhaltsverzeichnis ..... i**

**II. Formelzeichen und Abkürzungsverzeichnis ..... v**

**III. Abbildungsverzeichnis ..... ix**

**IV. Tabellenverzeichnis ..... xv**

**1 Einleitung ..... 1**

**2 Stand der Technik in Forschung und Industrie ..... 5**

2.1 Grundlagen des Wissensmanagements ..... 5
2.1.1 Der Wissensbegriff ..... 5
2.1.2 Wissensmanagement ..... 7
2.1.3 Wissenstransfer ..... 9
2.1.4 Wissensmodellierung und -speicherung ..... 10
2.2 Technische Grundlagen ..... 16
2.2.1 Large Language Models ..... 16
2.2.2 Retrieval Augmented Generation ..... 18
2.2.3 Rasa ..... 19
2.2.4 Sentence Transformer ..... 20
2.2.5 Computer Vision - YOLO ..... 21
2.2.6 Automatische Sprachverarbeitung ..... 22
2.2.7 Hand-Tracking und Gestenerkennung ..... 23
2.2.8 Process Mining ..... 25
2.3 Aktuelle Ansätze für Wissensmanagementsysteme ..... 26
2.3.1 Kommerzielle Systeme ..... 26
2.3.2 Experimentelle Systeme aus der Forschung ..... 32
2.4 Fazit zum Stand der Technik und Forschung ..... 37

**3 Zielsetzung und Vorgehensweise ..... 39**

3.1 Zielsetzung ..... 39
3.2 Vorgehensweise ..... 40

**4 Anforderungsanalyse ..... 43**

4.1 Ausrichtung, Methodik und Randbedingungen der Studie ..... 43
4.2 Ergebnisse der Studie ..... 46
4.2.1 Unternehmensspezifische Informationen ..... 47
4.2.2 Wissensmanagement ..... 49
4.2.3 Gestaltungsvorstellungen ..... 52

4.3 Anforderungen an Wissensmanagement auf dem Shopfloor .....53
4.3.1 Ableitung von Anforderungen auf Basis der Studienergebnisse ....53
4.3.2 Validierung der identifizierten Anforderungen .....56
4.4 Zusammenfassung und Bewertung der Anforderungsanalyse .....57

**5 Konzept für Wissensmanagement auf dem Shopfloor .....59**

5.1 Aktiver Wissenstransfer .....62
5.1.1 Chatbot-basiertes System für eine effiziente und intuitive Erfassung von Wissen auf dem Shopfloor .....64
5.1.2 Chatbot-basiertes System für einen effizienten und intuitiven Abruf von Wissen auf dem Shopfloor .....70
5.2 Semi-aktiver Wissenstransfer .....72
5.2.1 Automatisierte Analyse und Segmentierung von Prozessvideos zur Effizienzsteigerung beim Wissenstransfer auf dem Shopfloor .....74
5.3 Passiver Wissenstransfer .....82
5.3.1 Integration von Hand-Tracking und Process Mining zur Erfassung von Best Practices bei manuellen Montagen .....84
5.3.2 Einsatz von Hand-Tracking zur Überführung manueller Bearbeitungsverfahren von Bauteilen in roboterbasierte Prozesse und virtuelle Lernumgebungen .....94

**6 Referenzimplementierung der Konzepte .....99**

6.1 Umsetzung des aktiven Wissenstransfers .....99
6.1.1 Implementierung der Funktionen des Chatbots basierend auf dem Rasa-Framework .....100
6.1.2 Implementierung der Funktionen des Chatbots basierend auf Large Language Modellen .....104
6.2 Umsetzung des semi-aktiven Wissenstransfers .....109
6.3 Umsetzung des passiven Wissenstransfers .....118
6.3.1 Implementierung der Integration von Hand-Tracking und Process Mining zur Erfassung von Best Practices bei manuellen Montagen .....118
6.3.2 Implementierung von Hand-Tracking zur Überführung manueller Bearbeitungsverfahren von Bauteilen in roboterbasierte Prozesse und virtuelle Lernumgebungen .....131

**7 Validierung und Evaluation .....135**

7.1 Validierung der Funktionalitäten .....135
7.1.1 Validierung der Referenzimplementierung des aktiven Wissenstransfers .....135
7.1.2 Validierung der Referenzimplementierung des semi-aktiven Wissenstransfers .....140
7.1.3 Validierung der Referenzimplementierung des passiven Wissenstransfers .....146

7.2 Evaluation im Hinblick auf die Anforderungen von KMU .....151
7.3 Vergleich und Einordnung der Konzepte.....153
7.3.1 Beantwortung der Teilforschungsfragen .....155

**8 Zusammenfassung und Ausblick.....157**

8.1 Zusammenfassung.....157
8.2 Ausblick.....159

**V. Literaturverzeichnis.....163**

**VI. Anhang .....183**

# II. Formelzeichen und Abkürzungsverzeichnis

*Formula Symbols and Abbreviations*

| Formelzeichen | Einheit | Beschreibung |
|---|---|---|
| $a$ | | Vektor a |
| $b$ | | Vektor b |
| $c$ | | Confidence-Wert |
| $i$ | | Index zur Nummerierung |
| $j$ | | Index zur Nummerierung |
| $n$ | | Anzahl Tokens |
| $t$ | s | Zeitpunkt eines Montageschritts |
| $\Delta t_i$ | s | Dauer eines Montageschritts |
| $v$ | | Vektorrepräsentanz eines Fingergliedes |
| $w_G$ | | Gewichtungsfaktor der geringen Kräfte |
| $w_M$ | | Gewichtungsfaktor der mittleren Kräfte |
| $w_H$ | | Gewichtungsfaktor der hohen Kräfte |
| $x_{Pixel}$ | Pixel | Pixelbasierte X-Koordinate |
| $x_{Normalisiert}$ | | Auf Bildbreite normalisierte X-Koordinate |
| $y_{Pixel}$ | Pixel | Pixelbasierte Y-Koordinate |
| $y_{Normalisiert}$ | | Auf Bildbreite normalisierte Y-Koordinate |
| $B$ | | Belastung auf den Finger-Hand-Bereich |
| $B_{Total}$ | | Gesamtbelastung einer Montagetätigkeit auf den Finger-Hand-Bereich |
| $B_{Pro\ Sekunde}$ | | Durchschnittliche Belastung pro Sekunde auf den Finger-Hand-Bereich |
| $N$ | | Anzahl Sequenzen eines Videos |
| $P_{2D}$ | | Ortsvektor bzw. Koordinatenpunkt im zweidimensionalen Raum |
| $P_{3D}$ | | Ortsvektor bzw. Koordinatenpunkt im dreidimensionalen Raum |
| $S$ | | Ein Satz einer Transkription |
| $T_B$ | s | Zeitpunkt der Aufnahme eines Bildes in Sekunden |
| $T_{Total}$ | s | Gesamtprozesszeit einer Montage |
| $\theta$ | Rad | Winkel zwischen zwei Vektoren (Bogenmaß) |

| Abkürzung | Beschreibung |
|---|---|
| API | Applicational Programming Interface, Programmierschnittstelle |
| AR | Augmented Reality |
| BERT | Bidirectional Encoder Representation from Transformers |
| BPMN | Business Process and Notation |
| CNN | Convolutional Neural Network |
| CRUD | Create, Read, Update, Delete |
| CSV | Comma-seperated Values |
| DMS | Dokumentenmanagementsysteme |
| DNN | Deep Neural Network |
| FAISS | Facebook AI Similarity Search |
| GAN | Generative Adversarial Network |
| GPU | Graphics Processor Unit, Grafikprozessor |
| GUI | Graphical User Interface, Grafische Benutzerschnittstelle |
| HMD | Head-Mounted-Display |
| IoU | Intersection over Union |
| JSON | JavaScript Object Notation |
| KI | Künstliche Intelligenz |
| KMU | Kleine und mittlere Unternehmen |
| KMS | Knowledge Management Systems, Wissensmanagementsysteme |
| LLM | Large Language Model |
| NLP | Natural Language Processing |
| NLU | Natural Language Understanding |
| NoSQL | Nichtrelationale Datenbanken |
| OWL | Web Ontology Language |
| POV | Point of View |
| RAG | Retrieval-Augmented Generation |
| RDF | Resource Description Framework |
| RNN | Recurrent Neural Network |
| SAL | Smart Automation Lab, Laborumgebung des Werkzeug-maschinenlabors WZL der RWTH Aachen University |
| SBERT | Sentence-BERT |

| | |
|---|---|
| SQL | Structured Query Language, relationale Datenbank |
| STT | Speech to Text |
| SVM | Support Vector Machines |
| TAB | Das Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag |
| TOS | Time of Flight |
| TTS | Text to Speech |
| UX | User Experience, Benutzererfahrung |
| VUI | Voice User Interface |
| W3C | World Wide Web Consortium |
| WER | Word Error Rate |
| XML | Extensible Markup Language |

# III. Abbildungsverzeichnis

*List of Figures*

Abbildung 2.1: Auszugsweise Darstellung der Wissenstreppe nach [NORT21] .....6

Abbildung 2.2: SECI-Modell nach [NONA95].....8

Abbildung 2.3: Grundsätzlicher Prozessablauf von RAG angelehnt an [OMRA24] ...19

Abbildung 2.4: Unterscheidung von Klassifizierung, Lokalisierung, Segmentierung und Objektdetektion nach [DIWA23] .....21

Abbildung 3.1: Adaption der Phasen des Design Thinking für den Aufbau dieser Dissertation .....40

Abbildung 3.2: Detaillierte Vorgehensweise .....41

Abbildung 4.1: Kategorien der Sektion *Unternehmensspezifische Informationen* .....47

Abbildung 4.2: Kategorien der Sektion *Wissensmanagement*.....49

Abbildung 4.3: Kategorien der Sektion *Gestaltungsvorstellungen*.....52

Abbildung 4.4: Vergleich der identifizierten Anforderungen mit den in der VDI 5610 definierten Zielen von Wissensmanagement.....56

Abbildung 4.5: Zuordnung der identifizierten Anforderungen an Wissensmanagementsysteme zu den Teilforschungsfragen .....58

Abbildung 5.1: Systemarchitektur zur prozessparallelen Erfassung von Erfahrungswissen [DASS21] .....60

Abbildung 5.2: Vergleich der drei Konzepte für Wissensmanagement auf dem Shopfloor.....62

Abbildung 5.3: Systemarchitektur des aktiven Wissenstransfers.....63

Abbildung 5.4: Konzept der Chatbot-basierten Erfassung und Speicherung von Wissen .....64

Abbildung 5.5: Konzept zur Realisierung der Funktionen des Chatbots basierend auf dem Rasa-Framework.....67

Abbildung 5.6: Konzept zur Realisierung der Funktionen des Chatbots basierend auf Prompt Chaining für LLM .....69

Abbildung 5.7: Ablauf der Erfassung und des Abrufs von Wissen per Chatbot .....71

Abbildung 5.8: Systemarchitektur des semi-aktiven Wissenstransfers .....73

Abbildung 5.9: Grundsätzlicher Ablauf für die videobasierte Erfassung von Best Practices und Überführung in Lerneinheiten .....74

Abbildung 5.10: Neuartiger Ablauf für die videobasierte Erfassung von Best Practices mit automatisierter Aufbereitung .....75

Abbildung 5.11: Konzept der Analyse des Videos mittels Objekterkennung zur Detektion der Zeitpunkte von Werkzeugwechseln .....76

Abbildung 5.12: Vereinfachte Darstellung des Vorgangs der Transkription der Tonspur .....77

Abbildung 5.13: Konzept der Analyse der Tonspur mittels Schlüsselworterkennung 79

Abbildung 5.14: Konzept der Analyse der Tonspur zur Identifikation semantischer Ähnlichkeit .....79

Abbildung 5.15: Konzept des automatisierten Schneidens eines Videos auf Basis der per Objekterkennung, Schlüsselworterkennung oder durch die Identifikation von semantischer Ähnlichkeit identifizierten Zeitpunkte .....80

Abbildung 5.16: Ablauf der automatisierten Anreicherung .....81

Abbildung 5.17: Systemarchitektur des *passiven Wissenstransfers* .....83

Abbildung 5.18: Konzept zur Integration von Hand-Tracking und Process Mining zur Erfassung von Best Practices bei manuellen Montagen .....85

Abbildung 5.19: Modellhafte Darstellung des erfassten Arbeitsraums .....86

Abbildung 5.20: Schematische Darstellung der Positionen der Kameras und ihrer zugehörigen Koordinatensysteme .....87

Abbildung 5.21: Konzept der Ableitung von Koordinatenpunkten für die charakteristischen Punkte einer Hand auf Basis eines RGB-Bildes .88

Abbildung 5.22: Kombination der beiden zweidimensionalen Datensätze mittels Koordinatentransformation zur Bildung der dreidimensionalen Koordinatenpunkte ..... 89

Abbildung 5.23: Umwandlung der dreidimensionalen Koordinatenpunkte in die Vektorrepräsentanz einer Hand ..... 89

Abbildung 5.24: Analyse der Griffkombinationen zur Bestimmung der durchgeführten Montageschritte zu bestimmten Zeitpunkten ..... 91

Abbildung 5.25: Vergleich der individuellen Montageabläufe auf Basis von Petri-Netzen ..... 92

Abbildung 5.26: Konzept zum Einsatz von Hand-Tracking zur Überführung manueller Bearbeitungsverfahren von Bauteilen in roboterbasierte Prozesse und virtuelle Lernumgebungen ..... 95

Abbildung 5.27: Konzept der Visualisierung eines per Hand-Tracking erfassten Bewegungsablaufes in Form einer Trajektorie ..... 96

Abbildung 5.28: Visualisierung der Ghost Hands aus der Sicht eines lernenden Mitarbeitenden ..... 98

Abbildung 6.1: Systemarchitektur zur Implementierung der Erfassung und Ausgabe von Erfahrungswissen unter Verwendung des Rasa-Frameworks. 100

Abbildung 6.2: Grafisches Benutzerinterface der Web-Applikation des Rasa-basierten Chatbots ..... 101

Abbildung 6.3: Systemarchitektur zur Implementierung der automatisierten Analyse der Eingaben und Ableitung von Aktionen sowie der Wissensablage unter Verwendung des Rasa-Frameworks ..... 102

Abbildung 6.4: Auszug aus der zentralen Config-Datei des Rasa NLUs ..... 103

Abbildung 6.5: Auszug aus der Stories-Datei des Rasa Cores ..... 104

Abbildung 6.6: Systemarchitektur zur Implementierung der Chatbot-basierten Erfassung und Speicherung von Wissen unter Verwendung von LLM ..... 105

Abbildung 6.7: Grafisches Benutzerinterface der Web-Applikation des LLM-basierten Chatbots ..... 106

Abbildung 6.8: Systemarchitektur zur Implementierung der videobasierten Erfassung von Best Practices mit automatisierter Aufbereitung ..... 109

Abbildung 6.9: Darstellung der HoloLens 2 ..... 110

Abbildung 6.10: Exemplarische Darstellung der Werkzeugdetektion in einem Bild . 111

Abbildung 6.11: Ergebnis der Transkription einer Tonspur mittels Transformer-Modell ..... 113

Abbildung 6.12: Verwendete JSON-Struktur der Schlüsselwortliste ..... 113

Abbildung 6.13: Tab *Videoanalyse* der prototypischen GUI ..... 116

Abbildung 6.14: Tab *Analyse-Info* der prototypischen GUI ..... 117

Abbildung 6.15: Tab *Videoplayer* der prototypischen GUI ..... 117

Abbildung 6.16: Rekonfigurierbarer Arbeitstisch zur Reproduzierung realitätsnaher Arbeitsprozesse ..... 118

Abbildung 6.17: Bei der Referenzimplementierung betrachtete Baugruppe ..... 119

Abbildung 6.18: Systemarchitektur zur Implementierung der Erfassung der Aktivitäten bei manuellen Montagen ..... 120

Abbildung 6.19: Beispielhafte Darstellung der Erfassung der Hand-Landmarks und der zugehörigen Koordinaten ..... 121

Abbildung 6.20: Dreidimensionale Darstellung der Vektorrepräsentanz einer Hand ..... 123

Abbildung 6.21: Beispielhafte Darstellung des Winkels zwischen zwei aufeinanderfolgenden Verbindungsvektoren im zweidimensionalen Raum ..... 124

Abbildung 6.22: Definierte Griffklassen für die Montage der Referenzbaugruppe ... 124

Abbildung 6.23: Systemarchitektur zur Implementierung der Analyse der Aktivitäten bei manuellen Montagen ..... 127

Abbildung 6.24: Petri-Netz eines per Hand-Tracking erfassten Montageprozesses 129

Abbildung 6.25: Systemarchitektur zur Implementierung der Erfassung der Aktivitäten bei manuellen Bearbeitungsprozessen ..... 132

Abbildung 6.26: Visualisierung eines Bewegungsablaufs einer rechten Hand am Beispiel der Daumenspitze.....133

Abbildung 6.27: Visualisierung eines exemplarischen Geschwindigkeitsprofils einer rechten Hand am Beispiel der Daumenspitze .....134

Abbildung 7.1: Bildausschnitt der für die Validierung im KMU installierten Kamera inklusive des zu bearbeitenden Bauteils .....149

Abbildung 7.2: Vierfeldermatrix zur Einordnung der drei entwickelten Konzepte.....154

Abbildung D.1: Für die Referenzimplementierung des aktiven Wissenstransfer (Rasa) genutzte Struktur der SQL-Datenbank .....191

# IV. Tabellenverzeichnis

*List of Tables*

Tabelle 4.1: Teilnehmende der Studie ..... 45

Tabelle 4.2: Anforderungen aus der Sektion *Unternehmensspezifische Informationen* ..... 54

Tabelle 4.3: Kategorisierung der abgeleiteten Anforderungen und Klassifizierung in funktional und nicht-funktional ..... 55

Tabelle 5.1: Semantische Kategorien der Schlüsselworte für die Textanalyse ..... 78

Tabelle 6.1: Hardwarekonfiguration der für die Umsetzung verwendeten Workstation ..... 99

Tabelle 6.2: In der Vorstudie ermittelte Winkelbereiche des Daumens pro Griffklasse ..... 125

Tabelle 6.3: Zuweisung von Griffkombinationen zu Montageschritten für die Montage der Referenzbaugruppe ..... 128

Tabelle 6.4: Zuweisung der Montageschritte zu den bei der Ausführung auftretenden Kräften auf den Finger-Hand-Bereich ..... 130

Tabelle 7.1: Maßnahmen der Textnormalisierung zur Optimierung des Transkripts 136

Tabelle 7.2: Vergleich der von den drei evaluierten LLM gegebenen Antworten auf Fragen bezüglich eines hinterlegten Informationstextes ..... 139

Tabelle 7.3: Ergebnisse der für die Transkription verwendeten Whisper-Modelle bei der Validierung im SAL ..... 141

Tabelle 7.4: Ergebnisse der Werkzeugdetektion bei der Validierung im SAL ..... 142

Tabelle 7.5: Ergebnisse der für die Transkription verwendeten Whisper-Modelle bei der Validierung im KMU ..... 144

Tabelle 7.6: Ergebnisse der Werkzeugdetektion bei der Validierung im KMU ..... 145

Tabelle 7.7: Genauigkeit der Griffklassifizierung der verschiedenen Konfigurationen des Hand-Tracking-Systems ..... 147

Tabelle 7.8: Qualitative Erfüllungsgrade der Anforderungen .....151

Tabelle 7.9: Bewertung der Erfüllung der Anforderungen für die Konzepte des aktiven (A), semi-aktiven (SA) und passiven (P) Wissenstransfers .....152

Tabelle A.1: Übersicht bekannter Datenbankanbieter .....183

Tabelle A.2: Die fünf am häufigsten genutzten ECM-Systeme in KMU nach [BITK17] inklusive ihrer beworbenen Funktionalitäten .....184

Tabelle A.3: Übersicht aktuell verfügbarer Anbieter von Wikis und ihrer beworbenen Funktionalitäten .....185

Tabelle A.4: Aufstellung der von Bertram und Birtel betrachteten kommerziellen kognitiven Assistenzsysteme inklusive ihrer Funktionalitäten [BERT18].....186

Tabelle B.5: Abgeleitete Anforderungen auf Basis der Kernaussagen der Studienteilnehmenden in der Sektion Wissensmanagement .....187

Tabelle B.6: Abgeleitete Anforderungen auf Basis der Kernaussagen der Studienteilnehmenden in der Sektion Gestaltungsvorstellungen ...188

Tabelle E.7: Ergebnisse der manuellen, auf Schlüsselwörtern basierenden und semantischen Sequenzierung des Videos aus dem SAL .....192

Tabelle E.8: Zuordnung der mittels Schlüsselwort automatisiert erstellten Sequenzen zu den manuell erstellten Sequenzen des SAL-Prozesses und Berechnung von Overlap, Union und IoU.....192

Tabelle E.9: Zuordnung der mittels Sentence Transformer automatisiert erstellten Sequenzen zu den manuell erstellten Sequenzen des SAL-Prozesses und Berechnung von Overlap, Union und IoU.....193

Tabelle E.10: Ergebnisse der manuellen, auf Schlüsselwörtern basierenden und semantischen Sequenzierung des Videos aus dem KMU.....194

Tabelle E.11: Zuordnung der mittels Schlüsselwort automatisiert erstellten Sequenzen zu den manuell erstellten Sequenzen des KMU-Prozesses und Berechnung von Overlap, Union und IoU.....195

Tabelle E.12: Zuordnung der mittels Sentence Transformer automatisiert erstellten Sequenzen zu den manuell erstellten Sequenzen des KMU-Prozesses und Berechnung von Overlap, Union und IoU .....196

# 1 Einleitung

*Introduction*

Produzierende Unternehmen stehen kontinuierlich technischen und gesellschaftlichen Herausforderungen gegenüber, deren Ausprägungen einem steten Wandel unterliegen. Diese Dynamik resultiert aus gesellschaftlichen Entwicklungen sowie aus Vorfällen oder Veränderungen des Weltgeschehens. Noch vor wenigen Jahren galten der Bedarf an verkürzten Innovationszyklen, der Balanceakt zwischen zunehmend individualisierten Produkten und einer gleichzeitig effizienten Produktion sowie die aufgrund der Globalisierung zunehmende Konkurrenz als zentrale Herausforderungen [PAWL06; BREC11]. Insbesondere die im internationalen Vergleich niedrigeren Produktionskosten in Niedriglohnländern führten vielfach zur Verlagerung von Produktionsstätten in diese Regionen und beeinträchtigten die Wettbewerbsfähigkeit von Unternehmen in Hochlohnländern wie Deutschland [BREC11]. In den letzten Jahren haben tiefgreifende globale Veränderungen die Forderung nach nachhaltigem Wirtschaften erheblich verstärkt. Die produzierende Industrie steht zunehmend in der Verantwortung, ihre Produkte und Prozesse unter wirtschaftlichen, ökologischen und sozialen Gesichtspunkten ganzheitlich zu optimieren [BREC21; BREC23]. Die jüngst stark gestiegenen Energiekosten in Deutschland haben der Fokussierung auf Nachhaltigkeit und der Reduzierung der Produktionskosten zusätzliche Dringlichkeit verliehen, während der Wettbewerbsdruck aus dem Ausland, hauptsächlich aus China, durch kostengünstige und qualitativ hochwertige Produkte stetig zunimmt [SAUE24].

Neben diesen vorrangig extern bedingten Herausforderungen stellt der demografische Wandel, insbesondere die anstehende Verrentungswelle, für zahlreiche deutsche Unternehmen eine erhebliche Herausforderung dar [VDI09; HERT13; SAUE24]. Dies liegt vor allem an der Tatsache, dass die Ressource „Wissen“ in vielen Unternehmen ein essenzieller Wettbewerbsfaktor ist [VDI09; VANI12; HERT13; NORT18; SCHI23]. Der demografische Wandel birgt jedoch das Risiko eines signifikanten Wissensverlustes [VANI12; HERT13; SCHI23]. Kleine und mittlere Unternehmen (KMU) sind von dieser Entwicklung in gleichem Maße betroffen wie große Unternehmen [PAWL06; VANI12]. Da KMU ihre Produktion aufgrund der hohen erforderlichen Investitionskosten in der Regel nicht ins Ausland verlagern können, sind sie vermehrt auf technologieintensive und kundenspezifische Lösungen angewiesen [VANI12]. Dabei übernehmen die Mitarbeitenden oftmals zentrale Aufgaben der Steuerung und Qualitätsüberwachung von Prozessen, wodurch ihre langjährig aufgebaute Expertise von entscheidender Bedeutung ist [HOER23].

Aufgrund dessen ist ein gezieltes Wissensmanagement unerlässlich, um kritisches Wissen innerhalb des Unternehmens zu bewahren und die langfristige Wettbewerbsfähigkeit zu sichern [PAWL06; VDI09; HERT13; SCHU19]. In der Praxis wird Wissensmanagement jedoch häufig auf die Erstellung von Dokumentationen reduziert, die oft unstrukturiert und dezentral gespeichert sind, sodass relevante Informationen nicht ef-

fizient identifiziert und abgerufen werden können [ACKE18]. Viele der kommerziell verfügbaren Wissensmanagementsysteme sind aufgrund ihrer Struktur und Funktionalitäten zudem nur eingeschränkt für den Einsatz auf dem Shopfloor geeignet [ZAPP13].

Die aktuellen Entwicklungen im Bereich der Künstlichen Intelligenz (KI) bieten vielversprechende Potenziale, um die Generierung, Speicherung und Verteilung von Wissen effizienter zu gestalten [JARR23]. Entsprechend geeignete KI-Algorithmen sind in der Lage, Muster zu erkennen, Wissen automatisch zu kategorisieren und relevante Informationen gezielt bereitzustellen [TAHE23]. Dabei können die KI-Systeme in Form von intelligenten Assistenzsystemen oder Chatbots gestaltet werden, um eine personalisierte, intuitive und kontextabhängige Wissensbereitstellung zu ermöglichen [JARR23]. Für den Shopfloor von KMU besteht dennoch weiterhin erheblicher Forschungsbedarf hinsichtlich der praktischen Umsetzung solcher Systeme. Insbesondere die Erfassung von Wissen stellt einen Bereich des Wissensmanagements dar, der bislang nicht ausreichend adressiert wird.

Vor diesem Hintergrund ist das Ziel der vorliegenden Arbeit die Entwicklung neuer Ansätze für Wissensmanagementsysteme, mittels derer neu gewonnenes Erfahrungswissen unmittelbar, intuitiv und effizient gespeichert und bedarfsgerecht abgerufen werden kann. Dabei werden die spezifischen Anforderungen von KMU sowie deren teils unstrukturierte Arbeitsumgebungen gezielt berücksichtigt. Der Fokus der Untersuchungen liegt auf der Erfassung von Wissen und der damit einhergehenden Behandlung anwendungsnaher, ingenieurwissenschaftlicher Fragestellungen im Kontext der Konzeptionierung und Implementierung geeigneter Systeme. Methodisch folgt die Arbeit einem empirischen Ansatz, sodass die auf theoretischen Grundlagen entwickelten Konzepte prototypisch umgesetzt und praxisnah evaluiert werden. Dabei werden drei Varianten explorativ erarbeitet, implementiert und hinsichtlich ihrer Eignung zur Zielerreichung beurteilt.

Der Aufbau der Arbeit orientiert sich an den *sechs Phasen der Design-Thinking-Methode*, die ursprünglich als kundenorientierter Innovationsansatz zur Entwicklung kreativer Lösungen konzipiert wurde [MÜLL18]. Im Stand der Technik wird der Untersuchungsbereich definiert, ein fundiertes *Problemverständnis* aufgebaut und Defizite bestehender Lösungen herausgestellt. Basierend auf dem identifizierten Forschungsbedarf wird in Kapitel 3 die *Zielstellung* präzisiert sowie die daraus abgeleiteten Teilforschungsfragen formuliert. Kapitel 4 beinhaltet eine umfassende *Anforderungs- und Zielgruppenanalyse,* die auf einer empirischen, interviewgestützten Studie basiert. Die Generierung innovativer Lösungsansätze erfolgt in Kapitel 5, wobei drei *Konzepte* für den Wissenstransfer auf dem Shopfloor abgeleitet werden. In Kapitel 6 werden diese Konzepte *prototypisch implementiert*, bevor in Kapitel 7 eine *technische Validierung sowie eine Evaluation* hinsichtlich der definierten Anforderungen durchgeführt wird. Zudem werden die Ergebnisse im Kontext der Forschungsfragen kritisch diskutiert. Abschließend fasst Kapitel 8 das methodische Vorgehen sowie die zentralen Erkenntnisse zusammen und gibt einen Ausblick auf zukünftige Forschungsaktivitäten.

# Introduction

Manufacturing companies are constantly faced with technical and social challenges that are subject to constant change. This dynamic results from social developments as well as incidents or changes in world events. Just a few years ago, the need for shorter innovation cycles, the balancing act between increasingly individualized products and efficient production, and the growing competition resulting from globalization were considered key challenges [PAWL06; BREC11]. In particular, the lower production costs in low-wage countries compared to other countries often led to the relocation of production facilities to these regions and impaired the competitiveness of companies in high-wage countries such as Germany [BREC11]. In recent years, profound global changes have significantly increased the demand for sustainable business practices. The manufacturing industry is increasingly responsible for holistically optimizing its products and processes from an economic, ecological, and social perspective [BREC21; BREC23]. The recent sharp rise in energy costs in Germany has given additional urgency to the focus on sustainability and the reduction of production costs, while competitive pressure from abroad, mainly from China, is steadily increasing due to low-cost, high-quality products [SAUE24].

In addition to these primarily external challenges, demographic change, in particular the upcoming wave of retirements, poses a considerable challenge for many German companies [VDI09; HERT13; SAUE24]. This is mainly due to the fact that the resource “knowledge” is an essential competitive factor in many companies [VDI09; VANI12; HERT13; NORT18; SCHI23]. However, demographic change carries the risk of a significant loss of knowledge [VANI12; HERT13; SCHI23]. Small and medium-sized enterprises (SMEs) are affected by this development to the same extent as large companies [PAWL06; VANI12]. Since SMEs are generally unable to relocate their production abroad due to the high investment costs involved, they are increasingly reliant on technology-intensive and customer-specific solutions [VANI12]. Employees often take on key tasks in the control and quality monitoring of processes, which means that their expertise, built up over many years, is of crucial importance [HOER23].

For this reason, targeted knowledge management is essential in order to preserve critical knowledge within the company and ensure long-term competitiveness [PAWL06; VDI09; HERT13; SCHU19]. In practice, however, knowledge management is often reduced to the creation of documentation, which is in many cases unstructured and stored in a decentralized manner, meaning that relevant information cannot be efficiently identified and retrieved [ACKE18]. Due to their structure and functionalities, many of the commercially available knowledge management systems are also only suitable for use on the shop floor to a limited extent [ZAPP13].

Current developments in the field of artificial intelligence (AI) offer promising potential for making the generation, storage and distribution of knowledge more efficient [JARR23]. Appropriate AI algorithms are capable of recognizing patterns, automatically categorizing knowledge and providing relevant information in a targeted manner [TAHE23]. AI systems can be designed in the form of intelligent assistance systems or

chatbots to enable personalized, intuitive and context-dependent knowledge provision [JARR23]. Nevertheless, there is still a considerable need for research into the practical implementation of such systems for the shop floor of SMEs. In particular, the acquisition of knowledge is an area of knowledge management that has not yet been adequately addressed.

Against this background, the aim of this thesis is to develop new approaches for knowledge management systems that enable newly acquired experiential knowledge to be stored immediately, intuitively and efficiently and retrieved as needed. The specific requirements of SMEs and their sometimes unstructured working environments are taken into account in a targeted manner. The research focuses on the acquisition of knowledge and the associated treatment of application-oriented engineering questions in the context of the design and implementation of suitable systems. Methodologically, the work follows an empirical approach, so that the concepts developed on theoretical foundations are implemented as prototypes and evaluated in a practical manner. Three variants are being developed exploratively, implemented, and assessed in terms of their suitability for achieving the objectives.

The structure of this thesis is based on the six phases of the design thinking method, which was originally conceived as a customer-oriented innovation approach for developing creative solutions [MÜLL18]. The state of the art defines the area of investigation, establishes a sound understanding of the problem and highlights the shortcomings of existing solutions. Based on the identified research needs, Chapter 3 specifies the objectives and formulates the sub-research questions derived from them. Chapter 4 contains a comprehensive analysis of requirements and target groups based on an empirical, interview-supported study. Innovative solutions are generated in Chapter 5, where three concepts for knowledge transfer on the shop floor are derived. In Chapter 6, these concepts are implemented as prototypes before technical validation and evaluation with regard to the defined requirements are carried out in Chapter 7. In addition, the results are critically discussed in the context of the research questions. Finally, Chapter 8 summarizes the methodological approach and key findings and provides an outlook on future research activities.

# 2 Stand der Technik in Forschung und Industrie

*State of the Art in Research and Industry*

Wissensmanagement ist seit vielen Jahren Gegenstand der Forschung, findet jedoch auf dem Shopfloor von KMU bislang kaum oder lediglich in Ansätzen Anwendung. Um den Begriff des Wissensmanagements sowie die Hindernisse zu verstehen, die einer strukturierten Implementierung auf dem Shopfloor entgegenstehen, werden in Kapitel 2.1 zunächst die grundlegenden Konzepte und etablierten Methoden des Wissensmanagements, des Wissenstransfers sowie der Wissensmodellierung und -speicherung erläutert. Kapitel 2.2 widmet sich der Darstellung der technischen Grundlagen, welche die Basis für die in dieser Arbeit entwickelten Konzepte und Implementierungen bilden. Zur Abgrenzung der vorliegenden Arbeit von bestehenden Ansätzen wird in Kapitel 2.3 eine Übersicht über relevante kommerzielle und experimentelle Systeme und Methoden aus Industrie und Wissenschaft für die Umsetzung von Wissensmanagement auf dem Shopfloor gegeben. Anschließend wird in Kapitel 2.4 ein Fazit gezogen, in dem die Forschungslücken identifiziert und die daraus resultierenden Beiträge dieser Arbeit eingeordnet werden.

## 2.1 Grundlagen des Wissensmanagements

*Basics of Knowledge Management*

### 2.1.1 Der Wissensbegriff

*The Concept of Knowledge*

Für eine fundierte Auseinandersetzung mit dem Thema Wissensmanagement ist es zunächst erforderlich, den Begriff „Wissen“ zu definieren und ihn von den Konzepten „Daten“ und „Information“ abzugrenzen. DAVENPORT UND PRUSAK [DAVE98] definieren diese Begriffe wie folgt:

- **Daten:** Ein Satz diskreter, objektiver Fakten über Ereignisse, die moderne Organisationen üblicherweise in technologischen Systemen speichern. Diese Daten werden aus verschiedenen Quellen eingespeist.
- **Information:** Eine Nachricht, die in Form eines Dokuments oder einer hör- oder sichtbaren Mitteilung übermittelt wird. Informationen haben stets einen Sender und einen Empfänger und dienen dazu, die Wahrnehmung, das Urteil oder das Verhalten des Empfängers zu beeinflussen. Ob eine Nachricht informativ ist, entscheidet der Empfänger.
- **Wissen:** Eine fließende Mischung aus Erfahrungen, Werten, kontextbezogenen Informationen und Expertenwissen, die einen Rahmen für die Bewertung neuer Erfahrungen und Informationen bietet. Wissen entsteht im Kopf der Wissenden und wird dort angewandt. In Organisationen ist Wissen nicht nur in Dokumenten oder Ablagen eingebettet, sondern auch in organisatorischen Routinen, Prozessen, Praktiken und Normen.

Die Übergänge zwischen Daten, Informationen und Wissen sind fließend und lassen sich nicht eindeutig bestimmen. Laut DAVENPORT UND PRUSAK dient diese Definition als

Grundlage für ein einheitliches Verständnis, erhebt jedoch keinen Anspruch auf Vollständigkeit oder Endgültigkeit. [DAVE98]

Die schrittweise Transformation von Daten über Informationen zu Wissen und schließlich zur Handlungs- und Wettbewerbsfähigkeit eines Unternehmens wird durch die Wissenstreppe von NORTH [NORT21] anschaulich illustriert. Abbildung 2.1 zeigt die ersten Stufen der Wissenstreppe und verdeutlicht diese mithilfe eines Beispiels aus dem Produktionsumfeld.

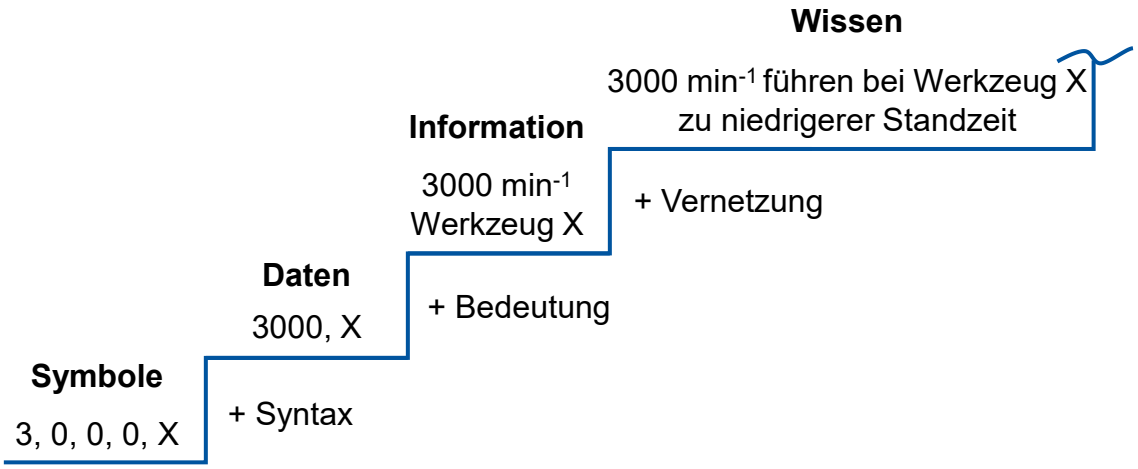



**Abbildung 2.1: Auszugsweise Darstellung der Wissenstreppe nach [NORT21]**

*Excerpt from the Knowledge Staircase According to [NORT21]*

Der erste Schritt auf der Wissenstreppe besteht in der Organisation von Symbolen, wie beispielsweise Buchstaben oder Ziffern, durch Syntax oder Codes zu Daten [NORT18; NORT21]. Daten werden gemäß dem DEUTSCHEN INSTITUT FÜR NORMUNG als Fakten über ein Objekt definiert [DEUT15; DEUT21]. Ein Objekt wird dabei als Einheit, Gegenstand oder etwas Wahrnehmbares oder Vorstellbares beschrieben und kann sowohl materiell (z.B. ein Motor, ein Werkzeug oder ein Blatt Papier), immateriell (z.B. ein Projektplan oder ein Konversionsverhältnis) als auch imaginär (z.B. eine hypothetische Prozessstruktur oder ein Design) sein. Ohne weiteren Zusammenhang oder Hintergrundinformationen bleiben Daten jedoch bedeutungslos und können isoliert nicht interpretiert werden [NORT21]. Aus diesem Grund werden Daten auch als „Rohmaterial“ bezeichnet [VDI09].

Im nächsten Schritt der Wissenstreppe wird den Daten Bedeutung verliehen, wodurch Informationen generiert werden [DEUT15; DEUT21; NORT18]. Informationen sind Daten, die durch das Hinzufügen von Relevanz und Zweck angereichert wurden und daher kontextualisiert, kategorisiert oder korrigiert werden können [FREY14; VDI09]. Ohne die Vernetzung mit anderen aktuellen oder gespeicherten Informationen bleiben isolierte Informationen für den Betrachter jedoch oft wenig aussagekräftig [NORT21].

Schließlich wird Wissen als Ergebnis der bewussten oder unbewussten Vernetzung von Informationen erzeugt [NORT18], wobei es das Anstellen von Vergleichen, das Herstellen von Verknüpfungen sowie das Treffen von Entscheidungen ermöglicht

[FREY14; VDI09]. Wissen ist kontextspezifisch, häufig personengebunden und wird durch individuelle Erfahrungen geprägt [NORT21].

Eine weitere in der Literatur häufig diskutierte Perspektive auf Wissen basiert auf der Unterscheidung zwischen **explizitem** und **implizitem** Wissen. NONAKA [NONA07] definiert diese beiden Arten von Wissen wie folgt:

- **Explizites Wissen** ist formalisiert und systematisch strukturiert, wodurch es leicht kommuniziert und geteilt werden kann.
- **Implizites Wissen** hingegen ist schwer zu artikulieren und nicht ohne Weiteres übertragbar, da es sich häufig in erworbenen Fähigkeiten manifestiert, die durch Erfahrung erworben werden.

Implizites Wissen wird oft durch langjährige praktische Tätigkeit entwickelt und kann beispielsweise im Handwerksbereich als „Fingerspitzengefühl“ beschrieben werden. Ein erfahrener Handwerker weiß intuitiv, worauf er achten muss, ist jedoch nicht immer in der Lage, die wissenschaftlichen oder technischen Prinzipien hinter seinen Handlungen explizit zu formulieren. [NONA07]

### 2.1.2 Wissensmanagement

*Knowledge Management*

Der Begriff *Wissensmanagement* (engl.: Knowledge Management) bezeichnet die systematische Gestaltung und Steuerung von Prozessen in Unternehmen, mit dem Ziel, Wissen als strategische Ressource nutzbar zu machen, um die Innovations- und Wettbewerbsfähigkeit der Unternehmen zu sichern und zu fördern [VDI09; NORT21]. Obwohl Wissensmanagement oft als modernes Konzept wahrgenommen wird, existiert es in informeller Form bereits seit Jahrhunderten. Schon in traditionellen Organisationen, Unternehmen oder Familien wurde Wissen in Form von Erfahrungen weitergegeben [GRAM20]. Grundlegend lassen sich dabei drei Kernschritte unterscheiden: Die *Erfassung*, *Speicherung* und *Bereitstellung* von Wissen [PREE01].

Aktuelle Methoden des Wissensmanagements zielen darauf ab, den Transfer insbesondere von implizitem Wissen innerhalb des Unternehmens zu fördern, um dieses Wissen im Unternehmen zu integrieren und dadurch Wettbewerbsvorteile zu erzeugen und zu sichern. Aus diesem Grund sollen individuelle Erfahrungen weitergereicht, Problemlösungen abstrahiert und die Verteilung von Kompetenzen in Organisationen ersichtlich gemacht werden. [CHEN10]

Eine der am häufigsten zitierte Methode des Wissensmanagements ist das SECI-Modell von NONAKA UND TAKEUCHI (siehe Abbildung 2.2), welches vier Phasen zur Schaffung von Wissen in Unternehmen beinhaltet [NONA95]:

- **Sozialisation:** Implizites Wissen wird über direkte Interaktion zweier Personen durch Beobachtung, Nachahmung und Übung weitergegeben. Da jedoch keine systematische Erfassung des impliziten Wissens erfolgt, findet keine Umwandlung in explizites Wissen statt. Deshalb kann das Wissen nicht ohne Weiteres vom gesamten Unternehmen genutzt werden.

- **Externalisierung:** Durch eine geeignete Formulierung und Dokumentation des impliziten Wissens wird es in explizites Wissen umgewandelt und kann anschließend systematisch mit allen Mitarbeitenden geteilt werden.
- **Kombination:** Durch die Kombination von bestehendem explizitem Wissen aus unterschiedlichen Quellen wird neues explizites Wissen generiert.
- **Internalisierung:** In der Phase der Internalisierung wird das neue explizite Wissen im Unternehmen weitergegeben und durch die Integration in die tägliche Arbeit von den Mitarbeitenden Stück für Stück verinnerlicht. Damit wird das Wissen wieder zu implizitem Wissen, welches die Mitarbeitenden dabei unterstützt, ihr Wissen zu erweitern, zu vertiefen und damit innovative Dinge zu entwickeln.

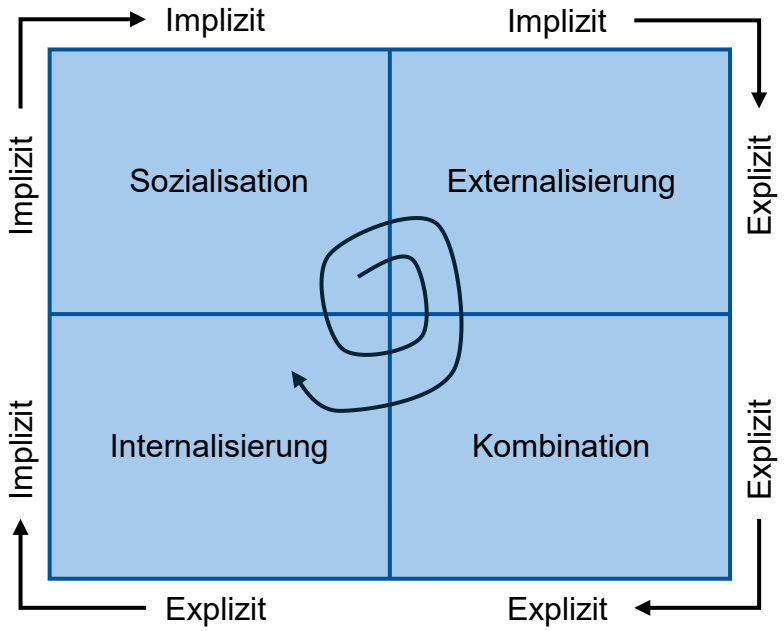



**Abbildung 2.2: SECI-Modell nach [NONA95]**

*SECI Model According to [NONA95]*

Die vier Phasen des SECI-Modells stehen miteinander in einer dynamischen Wechselwirkung. Sobald alle vier Phasen durchlaufen sind, beginnt diese Wissensspirale erneut, diesmal jedoch auf einer höheren Ebene. Damit wird die Wissensbasis eines Unternehmens mit jedem Zyklus sukzessive erweitert. Die Externalisierung und Internalisierung bilden die kritischen Phasen des Modells, da in beiden Fällen ein gewisses Maß an Engagement von den Mitarbeitenden notwendig ist. [NONA07]

Die technische Realisierung von Wissensmanagementsystemen ist heute weitgehend möglich. Die größere Herausforderung besteht jedoch in der Integration solcher Systeme in betriebliche Abläufe. Dies erfordert die aktive Einbindung der Mitarbeitenden, um Akzeptanz und Motivation zu schaffen. [DICK02]

Wissensmanagement hat sowohl eine stabilisierende als auch eine dynamisierende Wirkung auf Unternehmen. Die systematische Dokumentation von Prozessen, Erfahrungen und Best Practices verhindert den Verlust kritischen Wissens und schafft eine

stabile Basis. Gleichzeitig ermöglicht diese Basis dynamisches Handeln, indem Unternehmen schneller auf Veränderungen reagieren und innovative Problemlösungen sowie Next Practices entwickeln können. [ACKE18]

Die Implementierung von Wissensmanagement kann grundsätzlich durch *zentrale* oder *dezentrale* Ansätze erfolgen. Beim *zentralen* Ansatz werden alle Aufgaben von spezialisierten Expertinnen und Experten bearbeitet, die Methoden für die Identifikation und das Management von Wissen entwickeln. Dieser Ansatz erfordert eine permanente Analyse des Wissensbedarfs, um essenzielle Wissensressourcen zu identifizieren. Im Gegensatz dazu wird beim *dezentralen* Ansatz Wissensmanagement direkt von den Mitarbeitenden durchgeführt, was die Identifikation kritischen Wissens erleichtert. Hierbei sind jedoch geeignete Entscheidungshilfen notwendig, um Mitarbeitende bei der Auswahl passender Wissensmanagementmethoden zu unterstützen. [DOMB13]

In der Praxis wird der Einsatz solcher Entscheidungshilfen jedoch oft durch mangelnde Akzeptanz und Motivation behindert. Eine Erhebung des FIR e.V. an der RWTH Aachen zeigt, dass hohe Arbeitsbelastung und ein fehlender persönlicher Nutzen von Mitarbeitenden häufig genannte Gründe für die geringe Verbreitung von Wissensmanagementsystemen sind [BIRT18].

Stattdessen wird Wissensmanagement oftmals auf die Erstellung von Dokumentationen reduziert. Mitarbeitende verfassen Protokolle, erstellen Präsentationen oder legen Handbücher an. Dieser Ansatz führt jedoch zu Problemen bei der Auffindbarkeit relevanter Informationen, beispielsweise während der Einarbeitung neuer Mitarbeitender. Der Mangel an Übersichtlichkeit und die fehlende Fähigkeit, Informationen in wichtige und unwichtige Inhalte zu differenzieren, behindern die Effektivität. [ACKE18]

### 2.1.3 Wissenstransfer

*Knowledge Transfer*

Im vorherigen Kapitel wird erläutert, dass verschiedene Methoden des Wissensmanagements (z.B. das SECI-Modell) sich insbesondere mit der Weitergabe von implizitem Wissen befassen. Dieser Prozess, bei dem eine Entität (Individuum, Gruppe, Abteilung etc.) von den Erfahrungen einer anderen profitiert, wird als Wissenstransfer (engl.: Knowledge Transfer) bezeichnet [ARGO00].

Damit ein Wissenstransfer erfolgreich stattfinden kann, müssen spezifische Voraussetzungen erfüllt sein: Der Besitzer des Wissens muss bereit sein, sein Wissen zu teilen und gleichzeitig muss der potenzielle Empfänger bereit sein, dieses Wissen zu erhalten. Nicht alle Mitarbeitenden besitzen jedoch die notwendigen Vorkenntnisse, um spezifisches Wissen zu erhalten und unmittelbar zu verstehen und anzuwenden. Aus diesem Grund erfordert der Prozess des Wissenstransfers gezielte Maßnahmen, wie Schulungen und Trainings, um die Effizienz des Transfers durch eine zielgruppengerechte Aufbereitung des Wissens zu steigern und den Zeit- und Ressourcenaufwand zu minimieren. [SCHA16]

Vor diesem Hintergrund führten NAKANO ET AL. eine Studie in einem Unternehmen zur Glasherstellung durch, die sich auf den Wissenstransfer auf dem Shopfloor konzentrierte [NAKA13]. Im Rahmen der Untersuchung wurden die Mitarbeitenden über mehrere Monate hinweg systematisch befragt und ihr Verhalten beobachtet. Die Ergebnisse der Studie zeigen, dass die sogenannten "Blue-Collars" (Arbeiter*innen auf dem Shopfloor) ihr Wissen und ihre Erfahrungen primär durch direkten Austausch miteinander teilen. NAKANO ET AL. identifizierten auf Basis ihrer Ergebnisse bestimmte Bedingungen, die für einen erfolgreichen Wissenstransfer dieser Art erfüllt sein müssen:

- Die Mitarbeitenden müssen in einer offenen, kollegialen und vertrauensvollen Atmosphäre interagieren.
- Die Meinungen und Perspektiven der Mitarbeitenden müssen berücksichtigt werden, um ein Verantwortungsbewusstsein zu fördern.
- Die Mitarbeitenden müssen über einen vergleichbaren grundlegenden Wissensstand sowie eine gemeinsame Fachsprache verfügen.

Diese Faktoren sind insbesondere in Arbeitsumgebungen mit wenig strukturierten Prozessen von Bedeutung, in denen Tätigkeiten wie Qualitätskontrolle und Prozessanpassung im Vordergrund stehen und den Mitarbeitenden ein gewisses Maß an Freiheit bei der Ausführung ihrer Aufgaben eingeräumt wird. [NAKA13]

Im Gegensatz dazu steht der Wissenstransfer auf einem Shoopfloor mit klar definierten, regelmäßig wiederkehrenden Abläufen, wie sie unter anderem in der Automobilbranche häufig anzutreffen sind. Die dort oftmals eingesetzten produktionsbegleitenden Systeme (z.B. Lean Management Systems) definieren Prozesse, etablieren eine einheitliche Fachsprache und vernetzen zudem die verschiedenen Abteilungen und Bereiche des Unternehmens. Die Mitarbeitenden folgen dabei standardisierten Abläufen und können bereits geringste Abweichungen von der Norm direkt dokumentieren und bei Bedarf zusätzliche Informationen oder Hilfe durch Kollegen anfordern. Solche Systeme ermöglichen somit gleichzeitig die strukturierte Ablage und den effizienten Abruf von Wissen und tragen dadurch wesentlich zu einem erfolgreichen Wissenstransfer bei. [MIYA07]

### 2.1.4 Wissensmodellierung und -speicherung

*Knowledge Modelling and Storage*

Für die Nutzung von technischen Systemen zur Umsetzung von systematischem Wissenstransfer spielt die Modellierung und Speicherung von Wissen eine entscheidende Rolle. In der Literatur und der Praxis finden sich dazu verschiedene Ansätze, die sich in der Art und Weise unterscheiden, wie Wissen dargestellt, strukturiert, dokumentiert und kontextualisiert wird. Im Folgenden wird ein Überblick darüber gegeben.

**Modellierung und Organisation von Wissen**

Ein zentrales Ziel von Wissensmanagement besteht darin, vorhandenes Wissen sichtbar zu machen, um dessen Abrufbarkeit, Diskussion, Bewertung und Verwaltung zu erleichtern [SPAR98].

In diesem Zusammenhang kommen häufig Methoden der Wissensvisualisierung zum Einsatz, welche die Nutzung visueller Darstellungen zur Unterstützung des Wissensmanagements fokussieren. Diese Methoden umfassen grafische Mittel zur Entwicklung, Bewertung, Vermittlung oder Anwendung von Wissen und zielen darauf ab, Mitarbeitende zu befähigen, vorhandene Erfahrungen anderer zu identifizieren, zu rekonstruieren und korrekt anzuwenden. Beispiele hierfür sind heuristische Skizzen (z.B. ad hoc erstellte Zeichnungen in Meetings), konzeptionelle Diagramme (z.B. das Porters-Five-Forces-Diagramm) und Wissenslandkarten. Diese Formate enthalten nicht nur beschreibende Fakten, sondern auch präskriptive und prognostische Erkenntnisse und fördern Sinnstiftungsprozesse. [EPPL07]

Wissenslandkarten dienen dabei als Instrument zur Unterstützung der Navigation innerhalb vorhandener Wissensbestände, indem sie die Verteilung von Wissen, den Abgleich zwischen vorhandenem und benötigtem Wissen, die Identifikation von Stärken und Potenzialen sowie die Planung von Maßnahmen visualisieren. Abhängig vom darzustellenden Inhalt können Wissenslandkarten verschiedene Schwerpunkte setzen: Sie visualisieren beispielsweise Experten und deren Wissensgebiete (Wissensträgerkarten), die Speicherorte von Wissen (Wissensbestandskarten), Prozessabläufe (Wissensanwendungskarten) oder die Struktur des vorhandenen Wissens, um Zusammenhänge und Abhängigkeiten aufzuzeigen (Wissensstrukturkarten). [OTT15]

Eine spezialisierte Form der Wissensstrukturkarten bilden die sogenannten „Concept Maps“, die Beziehungen zwischen Konzepten (Begriffen) visualisieren [RITT24]. Ihre Struktur ähnelt der von Mindmaps, unterscheidet sich jedoch durch die Verwendung von Knoten und *gerichteten* Kanten mit Beschriftungen, welche die Art der Beziehung zwischen Konzepten spezifizieren [RITT24]. Concept Maps werden typischerweise hierarchisch aufgebaut und von oben nach unten gelesen, wobei häufig Querverweise zwischen Konzepten in unterschiedlichen Segmenten integriert werden [NOVA08]. Laut NOVAK UND CANAS fördert die Kombination aus der hierarchischen Struktur und den Querverweisen den Prozess des kreativen Denkens [NOVA08]. Ausgehend von der Situation, dass die Mitarbeitenden keine Möglichkeit zum direkten, persönlichen Austausch haben, kann die Verwendung von Concept Maps dazu beitragen, implizites Wissen explizit zu dokumentieren [SCHA16]. Gleichzeitig wird der Abruf von Wissen vereinfacht, da die Lesenden einer solchen Concept Map direkt erkennen können, wann und warum eine Information relevant und wie sie mit anderen Informationen verknüpft ist [SCHA16].

Eine weitere Form der strukturierten Wissensorganisation bilden Taxonomien, welche hierarchische Klassifizierungs- oder Kategorisierungssysteme darstellen [HEDD16]. *HEDDEN* ist der Ansicht, dass Taxonomien grundlegend drei Funktionen übernehmen können [HEDD16]:

1. **Indizierung**: Eine Taxonomie ermöglicht die Definition eines kontrollierten Vokabulars, das eine konsistente Indizierung und Katalogisierung von Wissen durch verschiedene Mitarbeitende sicherstellt.
2. **Abruf**: Eine Taxonomie unterstützt den Wissensabruf durch die kontextuelle Verknüpfung von Begriffen sowie die systematische Abbildung von Synonymen und verwandten Begriffen.
3. **Organisation und Navigation**: Eine Taxonomie dient der Klassifikation und Kategorisierung von Wissen, indem sie eine hierarchische Organisation der vorhandenen Informationen ermöglicht. Dadurch wird die Navigation der Mitarbeitenden innerhalb der Wissensbasis signifikant erleichtert.

Taxonomien für die Organisation und Navigation werden beispielsweise in der Informationsarchitektur von Websites, in der Strukturierung von Intranet-Inhalten sowie in Content-Management-Systemen eingesetzt [HEDD16].

In der Literatur wird die Erstellung eines kontrollierten Vokabulars sowie die systematische Verknüpfung von Begriffen und Synonymen häufig mit dem Begriff des Thesaurus bezeichnet [AITC00; HEDD16]. AITCHISON differenziert zwischen den Begriffen, indem er Taxonomien primär als Werkzeuge zur Klassifizierung oder Kategorisierung versteht, während Thesauri auf der semantischen Ebene operieren und die Bedeutung von Begriffen sowie deren Beziehungen zueinander fokussieren [AITC00].

Ein dritter Begriff, welcher immer wieder im Zusammenhang mit Taxonomien und Thesauri genannt wird, sind Ontologien. Ontologien werden von *GRUBER* als explizite Spezifikationen einer Konzeptualisierung beschrieben, die die Menge an Entitäten und deren Beziehungen eines Fachgebiets definieren und menschenlesbar beschreiben [GRUB92]. Ontologien sind somit formale Modelle einer Anwendungsdomäne, die dazu dienen, den Austausch und das Teilen von Wissen zwischen menschlichen und maschinellen Akteuren zu erleichtern [STAA02].

Im Kontext von Industrie 4.0 werden Ontologien in der Literatur häufig mit dem *Semantic Web* in Verbindung gebracht. Dieses erweitert das traditionelle Internet durch eine semantische Schicht, die Daten in einem maschinenlesbaren Format beschreibt, um den Zugriff auf Informationen zu optimieren und fortschrittliche Datenanalysen zu ermöglichen [SABO14]. Zur Realisierungen von Ontologien im Kontext des Semantic Web hat das World Wide Web Consortium (W3C) die Sprachen Resource Description Framework (RDF) und darauf aufbauend die Web Ontology Language (OWL) entwickelt, welche RDF um eine komplexere Semantik ergänzt [ANTO09]. Beide Sprachen modellieren Beziehungen zwischen Entitäten als Tripel, bestehend aus Subjekt, Prädikat und Objekt, beispielsweise: <MaschineX, isA, Maschine> [SABO14].

CHEN UND FU betrachten die Erstellung von Ontologien nicht nur als Instrument zur Organisation von Wissen, sondern auch als Methode, um die Transformation von implizitem in explizites Wissen zu unterstützen. Dies wird durch die Externalisierung (siehe SECI-Modell, Abbildung 2.2) während der Erstellung der Ontologie ermöglicht. [CHEN10]

### Speicherung von Wissen

Die Digitalisierung hat tiefgreifende Veränderungen in der Art und Weise bewirkt, wie Wissen generiert, abgerufen und geteilt wird. Unternehmen verfügen inzwischen über eine Vielzahl digitaler Speichersysteme, die die strukturierte Ablage und den bedarfsgerechten Abruf von Daten und Informationen ermöglichen. [NURN23]

Zu den am häufigsten genutzten digitalen Speichersystemen in Unternehmen zählen Datenbanken, die eine strukturierte Speicherung von Daten ermöglichen, deren Beständigkeit und Integrität sicherstellen und diese für berechtigte Nutzer zugänglich machen. Datenbanken lassen sich grundsätzlich in *relationale* und *nichtrelationale* Systeme unterscheiden. Relationale Datenbanken speichern und organisieren Daten in Tabellen, deren Struktur und wechselseitige Beziehungen durch ein Schema definiert werden. Jede Tabelle besteht aus Spalten (Attributen), die spezifische Datentypen und Formate aufweisen, sowie Zeilen (Tupeln), welche die konkreten Werte enthalten. Die gespeicherten Werte sind atomar, das heißt, sie können nicht weiter unterteilt werden. Für das Speichern, Abrufen und Manipulieren der Daten wird die standardisierte Structured Query Language (SQL) verwendet. [OLIV22]

Nichtrelationale Datenbanken, auch NoSQL-Datenbanken genannt, benötigen im Gegensatz zu relationalen Datenbanken kein festes Schema, wodurch sie sich besonders für die Speicherung semi-strukturierter Datenformate wie JavaScript Object Notation (JSON) oder Extensible Markup Language (XML) eignen. Die Interaktion mit NoSQL-Datenbanken erfolgt nicht über eine deklarative Abfragesprache wie SQL, sondern über Programmierschnittstellen (APIs) für CRUD-Operationen (Create, Read, Update, Delete), die speziell auf programmatische Zugriffe ausgelegt sind. Einige NoSQL-Datenbanken bieten zudem die Möglichkeit, mehrere Versionen von Daten mit Zeitstempeln zu speichern, um Änderungen nachverfolgen zu können. [ELMA16]

Abhängig von der Art der Modellierung der Daten bei der Speicherung werden NoSQL-Datenbanken in vier Kategorien differenziert [ELMA16; OLIV22]:

1. **Schlüsselwert-Datenbanken:** Schlüsselwert-Datenbanken speichern Daten als Paare aus einem eindeutigen Schlüssel und einem zugeordneten Wert. Dieser Wert kann in verschiedenen Formaten vorliegen, von einfachen Zeichenketten über JSON- oder XML-Dokumente bis hin zu noch komplexeren Strukturen. Die Daten sind aggregatorientiert, was bedeutet, dass über den Schlüssel stets nur der gesamte Wert abgerufen werden kann, ohne Zugriff auf einzelne Teile des Wertes. Schlüsselwert-Datenbanken sind besonders effizient für Anwendungen, die schnelle Lese- und Schreiboperationen benötigen.
2. **Dokumentenorientierte Datenbanken:** Dokumentenorientierte Datenbanken speichern Daten in dokumentbasierten Formaten wie JSON, XML oder PDF. Im Gegensatz zu Schlüsselwert-Datenbanken erlauben sie den Zugriff auf spezifische Inhalte innerhalb eines Dokuments, da die Daten selbstbeschreibend sind und Indizes auf bestimmte Attribute erstellt werden können. Dokumentenorientierte Datenbanken eignen sich damit ideal für Anwendungen, bei denen die

Daten eine komplexere Struktur aufweisen, z.B. E-Commerce-Systeme oder Content-Management-Systeme.

3. **Spaltenorientierte Datenbanken:** Spaltenorientierte Datenbanken organisieren Daten in Tabellen, wobei jede Zeile durch eine eindeutige ID identifiziert wird und eine variable Anzahl von Spalten enthalten kann. Die Spalten werden inhaltlich gruppiert und zu sogenannten Spaltenfamilien zusammengefasst. Verschiedenen Zeilen können unterschiedliche Spaltenfamilien oder nur Teile davon beinhalten. Bei Abfragen werden Spaltenfamilien typischerweise als Einheit abgerufen, wodurch die Effizienz bei der Verarbeitung großer Datenmengen mit variabler Struktur erhöht wird. Spaltenorientierte Datenbanken eignen sich deshalb besonders für Anwendungsfälle wie Echtzeitanalysen und Data Warehousing.
4. **Graph-Datenbanken:** Graph-Datenbanken speichern Daten als Netzwerke von Knoten und Kanten. Dabei repräsentieren die Knoten Entitäten, während die Kanten die Beziehungen zwischen diesen beschreiben. Sowohl Knoten als auch Kanten können mit Labels und Attributen versehen werden, die zusätzliche Informationen enthalten und damit die eigentlichen Daten repräsentieren. Das Netzwerk aus Knoten und Kanten benötigt kein festes Schema und ermöglicht sowohl die gezielte Abfrage einzelner Informationen zu Entitäten als auch die Analyse von Beziehungen und die Identifikation von Mustern. Graph-Datenbanken zeichnen sich durch ihre besondere Leistungsfähigkeit in Szenarien aus, in denen die Modellierung und Analyse von Beziehungen zwischen Daten eine zentrale Rolle spielt, wie etwa in sozialen Netzwerken oder Empfehlungssystemen.

Die Struktur sowie die Semantik der Beziehungen zwischen den Entitäten eines in einer Graph-Datenbank gespeicherten Graphen können sowohl durch definierte Regeln als auch durch Ontologien spezifiziert werden [HOGA22].

Die Entwicklung und breite Anwendung innovativer Modelle aus dem Gebiet der KI, wie Convolutional Neural Networks (CNNs) und Transformer-Modellen, ermöglicht die Transformation unstrukturierter, multimodaler Daten in mathematische Vektordarstellungen. Diese hochdimensionalen Vektoren, auch Embeddings genannt, codieren semantische Merkmale der ursprünglichen Daten. Durch die Berechnung der Abstände zwischen den Vektoren können Ähnlichkeiten zwischen verschiedenen Datensätzen ermittelt werden, was insbesondere die effiziente Suche in umfangreichen Datenbeständen erleichtert. [JING24]

Zur Speicherung und Verwaltung der Embeddings wurden spezialisierte Datenbanksysteme, sogenannte *Vector Stores*, entwickelt, welche eine effiziente Speicherung, Indizierung und den Abruf der Vektoren ermöglichen. Neben den Vektoren selbst werden im Vector Store in der Regel ein eindeutiger Identifier für jeden Vektor sowie die Daten, die der Vektor repräsentiert, gespeichert. Der Identifier verknüpft die Vektoren eindeutig mit ihren zugehörigen Daten. [TAIP24]

Die Kombination aus der Transformation unstrukturierter Daten in Embeddings und deren Verwaltung in Vector Stores stellt eine leistungsfähige Methode dar, um unstrukturierte Daten effizient zu speichern und zugänglich zu machen [JING24; TAIP24]. Diese Methode der Datenspeicherung und des Datenabrufs findet bereits Anwendung in verschiedenen Bereichen, wie beispielsweise der umgekehrten Bildsuche, in Chatbots sowie in Empfehlungssystemen und gewinnt zunehmend an Verbreitung [LI23].

**Kombinierte Systeme**

In der Praxis nutzen Unternehmen oftmals IT-Systeme, mit denen sie Daten, Informationen und Wissen in individuellen, firmenspezifischen Strukturen organisieren und speichern. Diese zum Teil historisch gewachsenen Systeme kombinieren in der Regel verschiedene der zuvor beschriebenen Ansätze zur Organisation, Modellierung und Speicherung von Wissen in unterschiedlicher Tiefe und Komplexität. Dabei lassen sich insbesondere zwei Typen von Systemen differenzieren: *Dokumentenmanagementsysteme* und *Wikis*.

Dokumentenmanagementsysteme (DMS) unterstützen Organisationen bei der Erfassung, Speicherung und Bereitstellung von explizitem Wissen, indem sie Funktionen für die Erstellung, Verwaltung und den Abruf von Dokumenten bereitstellen. Sie ermöglichen damit die Verwaltung großer Informationsmengen und werden oft als webbasierte Anwendung implementiert, die Zugriffsmöglichkeiten über Intranet oder Internet bieten. Inhalte können in verschiedenen Formaten (z.B. Word, Excel, PDF, Audio, Video) dargestellt werden, um Informationen zu einem Thema umfassend abzubilden. DMS basieren häufig nur auf einem zentralen Repository, damit alle wichtigen Informationen an einem Ort gespeichert und effizient verwaltet und abgerufen werden können. Darüber hinaus bieten DMS typischerweise Funktionen zur Speicherung von Metadaten, Versionierung von Dokumenten, Gewährleistung der Daten- und Informationssicherheit sowie zur Implementierung effizienter Such- und Abrufmöglichkeiten. [KHAN15]

Im Gegensatz zu traditionellen *zentralen* Wissensmanagementsystemen wie DMS können bei der Verwendung eines Wikis nicht nur ausgewählte Mitarbeitende Wissen in das System eingeben. Stattdessen wird Wissen *dezentral* von allen Beteiligten gemeinschaftliche generiert und gepflegt, indem jede Person neue Inhalte erstellen sowie bestehende abrufen und bearbeiten kann. Dieser offene Zugang birgt zwar das Risiko, dass fehlerhafte oder irreführende Informationen in das System eingebracht werden, bietet gleichzeitig jedoch die Chance, durch die kollektive Verantwortung der Beteiligten eine präzise und konsistente Dokumentation des gesammelten Wissens zu ermöglichen. Alle Mitarbeitenden tragen gemeinschaftlich die Verantwortung, potenzielle Fehler zu identifizieren und zu korrigieren, wodurch die Qualität der Wissensbasis langfristig gesichert wird. Durch diesen partizipativen Ansatz fördern Wikis die Nutzung kollektiver Intelligenz und Kreativität. Die verfügbaren Werkzeuge erleichtern es den Nutzern, ihr Wissen zu formulieren, zu teilen und effizient wiederzufinden. [WAGN04]

Laut *HESTER* kann die freiwillige Bereitschaft der Nutzer, Wissen beizutragen, durch die Sorge beeinträchtigt werden, dabei ihre Wissensvorteile oder -macht zu verlieren.

Zudem zögern viele Nutzer, bestehende Inhalte zu bearbeiten, da sie befürchten, dadurch Kollegen oder Vorgesetzte zu verärgern. Diese sozialen Herausforderungen können dazu führen, dass Wikis nicht vollständig von den Nutzern akzeptiert werden, was ihre Effektivität als Wissensmanagementsystem einschränken kann. [HEST11]

*WAGNER* unterstreicht diese Problematik und gibt zudem zu bedenken, dass ein entscheidender Faktor für den Erfolg von Wikis in Organisationen die Unternehmenskultur sei. In hierarchisch strukturierten Organisationen, die eine Top-Down-Kommunikation und klare Informationspolitik verfolgen, haben Wikis oft geringe Chancen auf Akzeptanz, da sie einen demokratischeren Ansatz zur Wissensschaffung und -verteilung fördern. Ohne eine entsprechende Unternehmenskultur kann die Einführung von Wikis demnach nur begrenzt erfolgreich sein. [WAGN04]

Sogenannte *semantische Wikis* erweitern die Funktionalitäten herkömmlicher Wikis durch die Integration eines semantischen Modells für die eingegebenen Daten. Mithilfe von Sprachen wie RDF und OWL werden Ontologien erstellt, die eine strukturierte und maschinenlesbare Repräsentation der Informationen ermöglichen. Die Benutzer sind dabei nicht darauf angewiesen, die Ontologien eigenständig zu entwickeln. Stattdessen unterstützt das Benutzerinterface des semantischen Wikis den Prozess, indem es die Erstellung und Pflege der Ontologien intuitiv gestaltet. Die hinterlegten Ontologien erweitern die Funktionalität des Wikis um die bereits erläuterten strukturellen und semantischen Vorteile. [SCHA05; SCHA07]

## 2.2 Technische Grundlagen

*Technical Principles*

In diesem Kapitel werden die im Rahmen dieser Arbeit eingesetzten Technologien zur Entwicklung, Implementierung und Validierung der Konzepte für Wissensmanagement auf dem Shopfloor vorgestellt. Die einzelnen Unterkapitel behandeln sowohl die theoretischen Grundlagen der jeweiligen Technologien als auch tiefergehende Aspekte, die für das Verständnis und die wissenschaftliche Einordnung dieser Arbeit von zentraler Bedeutung sind.

### 2.2.1 Large Language Models

*Large Language Models*

Large Language Models (LLM) sind eine Klasse von neuronalen Netzen, die häufig mehr als hundert Milliarden Parameter umfassen und sich durch die Fähigkeit auszeichnen, natürliche Sprache zu verstehen und zu generieren, weshalb sie eine wichtige Rolle im Bereich des Natural Language Processing (NLP) spielen [ZHAO24a]. Die Entwicklung solcher Modelle erfolgt in einem zweistufigen Prozess: Zunächst werden sie in einem Pre-Training auf äußerst umfangreichen Textkorpora trainiert, um grundlegende Sprachstrukturen und Muster zu erlernen [PART24]. Anschließend kann ein Fine-Tuning auf einem domänenspezifischen Datensatz erfolgen, um das Modell an dedizierte Aufgaben anzupassen [PART24]. Die Skalierung der Modellgröße (gemessen an der Anzahl der Parameter) und des Umfangs der Trainingsdaten führt zu einer gesteigerten Leistungsfähigkeit und erweitert ihrer Funktionalitäten [WEI22]. Dadurch

sind große LLM in der Lage, komplexe Aufgaben wie Programmierung und logisches Schlussfolgern (Reasoning) effizient zu bewältigen [WEI22; PLAA24].

Moderne LLM basieren in der Regel auf der Transformer-Architektur, die speziell für neuronale Netze zur Verarbeitung von Texten entwickelt wurde [ZHAO24a]. Im Gegensatz zu sequenziellen Modellen arbeitet diese Architektur nicht schrittweise, sondern betrachtet alle Wörter eines Textes gleichzeitig, um sprachliche Zusammenhänge besser erkennen zu können. Die ursprüngliche Transformer-Architektur besteht aus einem Encoder-Decoder-System, wobei der Encoder die Eingabe in eine numerische Repräsentation in Form von Vektoren (Embeddings) überführt, die die Bedeutung des gesamten Textes in einer dichten, mathematischen Form speichert. Der Decoder nutzt diese Embeddings, um Wort für Wort eine sinnvolle Ausgabe zu erzeugen. Zentrales Prinzip dieser Architektur ist der Self-Attention-Mechanismus, der den Kontext eines Wortes bestimmt, indem er dessen Beziehungen zu anderen Wörtern innerhalb eines Satzes gewichtet. Durch die darauf aufbauende Multi-Head-Attention werden mehrere parallele Aufmerksamkeitsebenen genutzt, um verschiedene semantische Beziehungen zwischen Wörtern zu erfassen. Diese Mechanismen ermöglichen eine präzisere Modellierung sprachlicher Zusammenhänge und verbessern die Erkennung komplexer Bedeutungsnuancen erheblich. [VASW17]

LLM werden mittlerweile der Klasse der sogenannten *Foundation Models* zugeordnet. Diese Art von Modellen zeichnet sich durch ein Training auf umfangreichen Datensätzen aus und ist in der Lage, durch gezielte Anpassungen für spezifische Anwendungsfälle optimiert zu werden. Seit etwa 2019 gab es eine rasante Entwicklung im Bereich der Foundation Models, die auf mehrere Faktoren zurückzuführen ist: Die kontinuierliche Verbesserung leistungsfähiger Hardware für KI-Anwendungen, insbesondere Grafikprozessoren (GPU), die Verfügbarkeit zunehmend umfangreicher Trainingsdatensätze sowie die Einführung und Weiterentwicklung der Transformer-Architektur ermöglichten eine massive Skalierung und Weiterentwicklung dieser Modelle. [BOMM22]

Insbesondere die Veröffentlichung von Modellen wie LLaMA [TOUV23], PaLM [CHOW23] und GPT-4 [OPEN24] sowie die Möglichkeit, diese entweder kostenfrei auf eigener Hardware oder über die von den Entwicklern bereitgestellten Cloud-Dienste zu nutzen, hat die Zugänglichkeit und Anwendung dieser Modelle für die breite Öffentlichkeit erheblich vereinfacht [BOMM22]. Diese Entwicklungen eröffnen zudem die Möglichkeit, die Nutzbarkeit der derzeit verfügbaren LLM für spezifische Anwendungsfälle, wie das Wissensmanagement auf dem Shopfloor von KMU, zu untersuchen.

Zur Kommunikation mit einem LLM werden sogenannte Prompts („Aufforderungen“) genutzt, welche das LLM analysiert und eine entsprechende Antwort generiert [KADD23]. Der genaue Wortlaut des Outputs eines LLM ist nicht exakt deterministisch vorhersagbar [RENZ24]. Dies liegt unter anderem an der stochastischen Natur des Modells, den Einstellungen der Hyperparameter sowie der Vielfalt der zugrunde liegenden Trainingsdaten [RENZ24]. Selbst bei identischer, mehrfacher Ausführung desselben Prompts können diese Faktoren zu leicht variierenden Antworten führen.

Zur präzisen Steuerung des Verhaltens eines LLM wird Prompt Engineering eingesetzt [CHEN24]. Diese Methode umfasst die gezielte Gestaltung und Optimierung von Eingaben, um die Genauigkeit, Relevanz und Kohärenz der generierten Ergebnisse zu verbessern [CHEN24]. Prompt Engineering ist ein essenzieller Prozess, um das volle Potenzial von LLM auszuschöpfen und ihre Anwendbarkeit in verschiedenen Domänen zu optimieren, ohne dass ein zusätzliches Fine-Tuning des Modells erforderlich ist [SAHO24]. Laut SAHOO ET AL. können die in der Literatur existierenden Techniken des Prompt Engineering dabei in 29 verschiedene Klassen eingeteilt werden, deren Anwendung unterschiedliche Leistungssteigerungen eines LLM in Bezug auf die Antwortqualität, Präzision und Anpassungsfähigkeit hervorrufen kann [SAHO24].

Im Kontext der vorliegenden Arbeit wird insbesondere die Methodik des Prompt Chaining angewendet, bei der mehrere Prompts sequenziell an das LLM gesendet werden. Das Ergebnis eines Prompts dient dabei als Eingabe für den nachfolgenden Prompt, wodurch eine schrittweise Verarbeitung komplexer Aufgaben ermöglicht wird. Dies führt zu einer signifikanten Leistungssteigerung des Models. [SUN24]

### 2.2.2 Retrieval Augmented Generation

*Retrieval Augmented Generation*

Obwohl moderne LLM über fortschrittliche Fähigkeiten in der Verarbeitung natürlicher Sprache verfügen, bleibt das Problem der sogenannten Halluzinationen bestehen. Dieser Begriff bezeichnet die Generierung von Inhalten, die von den Benutzereingaben abweichen, dem zuvor erzeugten Kontext widersprechen oder nicht mit etablierten Fakten übereinstimmen [ZHAN25]. Zur Verbesserung der Antwortqualität und zur Minimierung der Erzeugung von Falschinformationen wurde die Retrieval-Augmented Generation (RAG) entwickelt [GAO24]. Diese Methodik erweitert die Leistungsfähigkeit von LLM, indem sie relevante Dokumentenausschnitte aus externen, digitalen Wissensquellen abruft und in den Generierungsprozess einbindet [GAO24; ZHAO24b]. Durch die Integration externer Informationen, anstelle der ausschließlichen Verwendung des in das Modell eintrainierten Wissens, verringert RAG signifikant die Wahrscheinlichkeit fehlerhafter Antworten und gilt daher als Schlüsseltechnologie für die Weiterentwicklung von Chatbots [GAO24].

Der grundsätzliche Prozessablauf von RAG ist in Abbildung 2.3 angelehnt an [OMRA24] dargestellt und besteht aus drei aufeinanderfolgenden Schritten: Zunächst werden relevante Informationen aus externen Wissensquellen abgerufen (*Retrieval*). Dazu wird die Benutzereingabe (Prompt) mittels eines Embedding Models vektorisiert und anschließend mit den in einem Vectore Store gespeicherten Informationen durch semantische Ähnlichkeitsberechnungen abgeglichen (siehe auch Kapitel 2.4.4). Die als relevant identifizierten Informationen werden im nächsten Schritt, der *Augmentation*, mit der ursprünglichen Eingabe kombiniert und an das LLM übergeben. Im letzten Schritt, der *Generation*, erzeugt das LLM unter Berücksichtigung der Benutzereingabe und der ergänzenden Informationen aus dem Vector Store eine kohärente Antwort. [WU24; ZHAO24b]

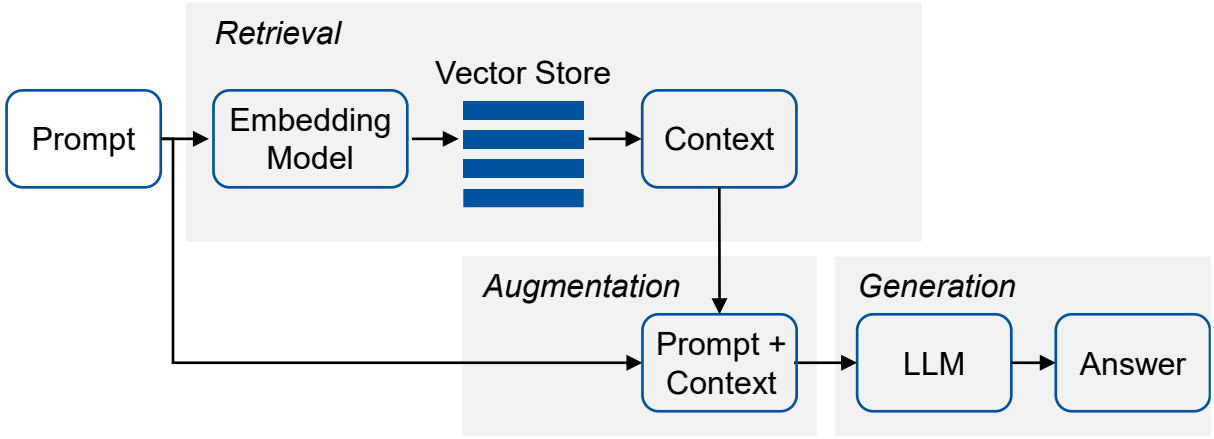



**Abbildung 2.3: Grundsätzlicher Prozessablauf von RAG angelehnt an [OMRA24]**

*Basic Process Flow of RAG Based on [OMRA24]*

Aktuelle Herausforderungen bei der Anwendung von RAG betreffen insbesondere die hohen Rechenanforderungen, die durch den Abruf externer Daten entstehen und mit zunehmendem Datenvolumen weiter ansteigen. Zudem schränkt die derzeit noch unzureichende Verarbeitung mehrdeutiger Informationen und domänenspezifischen Kontexts die breite Anwendbarkeit von RAG ein. [GUPT24]

### 2.2.3 Rasa

*Rasa*

Rasa ist ein Open-Source-Framework zur Entwicklung kommunikativer KI-Anwendungen, das umfassende Werkzeuge für die natürliche Sprachverarbeitung, die Dialogverwaltung und die Gestaltung interaktiver Konversationen mit Chatbots und virtuellen Assistenten bereitstellt [RAJP23; SHAR20]. Die Funktionalitäten von Rasa lassen sich in die beiden Hauptkomponenten Rasa Natural Language Understanding (NLU) und Rasa Core unterteilen. Rasa NLU ist für die Analyse und Verarbeitung der Spracheingaben der Nutzer verantwortlich, während Rasa Core mittels Machine-Learning-gestütztem Dialogmanagement die entsprechenden Reaktionen und Aktionen des Chatbots steuert [BOCK17].

Bei der Gestaltung eines Chatbots unter Verwendung des Rasa-Frameworks folgt die Verarbeitung der Benutzereingaben einem strukturierten Ablauf: Zunächst analysiert Rasa NLU die Intention der Benutzereingabe mithilfe sogenannter *Intents*, welche typische Benutzerabsichten repräsentieren. Damit der Chatbot unterschiedliche Intents differenzieren kann, müssen diese vorab definiert und mit exemplarischen Texteingaben annotiert werden. Diese Beispiele dienen als Trainingsdaten für das zugrunde liegende NLP-Modell, welches auf Basis dieser Daten die Erkennung der Intents erlernt und zur Generalisierung befähigt wird, sodass auch unbekannte, aber semantisch ähnliche Formulierungen zugeordnet werden können. [DINE21; RASA24a]

Neben der Erkennung von Intentionen ermöglicht Rasa die Extraktion spezifischer Informationen aus den Benutzereingaben, die für die korrekte Ausführung einer zugehörigen Aktion erforderlich sind. Diese sogenannten *Entities* umfassen Namen, Orte oder

andere strukturierte Daten. Zur Identifikation und Verarbeitung von Entities stellt Rasa verschiedene Methoden bereit: Eine Möglichkeit ist die Hinterlegung von *Synonymen* in den Trainingsdaten, wodurch das NLP-Modell semantisch gleichwertige Begriffe erkennen und korrekt zuordnen kann. Für Entities mit festen strukturellen Mustern, wie Nummern oder Codes, können *Regular Expressions* verwendet werden, welche die zugrunde liegende Zeichen- oder Zahlenlogik definieren. Darüber hinaus bietet Rasa die Option, *Lookup-Tabellen* zu nutzen, in denen vordefinierte Begriffslisten hinterlegt werden. Diese Methode eignet sich insbesondere für die Identifikation von Eigennamen oder anderen klar abgegrenzten Entitätsklassen, indem sie dem Modell eine feste Auswahl möglicher Begriffe oder Werte bereitstellt. [DINE21; RASA24a]

Sobald die Absicht der Benutzereingabe erkannt und die relevanten Entities extrahiert wurden, entscheidet der Rasa Core, welche Aktion ausgeführt werden muss [DINE21]. Zur Festlegung eines deterministischen Verhaltens des Chatbots im Umgang mit Benutzereingaben können im Rasa Core *Stories* definiert werden, die den Dialogfluss zwischen Benutzer und Chatbot festlegen, indem sie Sequenzen aus Intents und den daraus resultierenden Aktionen beschreiben [DINE21; RASA24b]. Auf Grundlage dieser vordefinierten Abläufe kann der Chatbot die gewünschte Reaktion auf bestimmte Abfolgen von Benutzerabsichten erlernen [DINE21; RASA24b]. Die vom Chatbot ausführbaren Aktionen umfassen sowohl Antworten und Rückfragen an den Benutzer als auch den Aufruf von *Costum Actions* [RAJP23]. Custom Actions ermöglichen die Ausführung beliebigen Programmcodes, wodurch unter anderem API-Aufrufe sowie Datenbankabfragen realisiert werden können [RASA24c].

### 2.2.4 Sentence Transformer

*Sentence Transformer*

Sentence Transformer bilden eine Klasse von Transformer-Modellen, die zur effizienten Berechnung numerischer Repräsentationen von Sätzen in Form hochdimensionaler Vektoren entwickelt wurden [AARS20]. Die meisten Sentence Transformer basieren auf einer Variante von BERT (Bidirectional Encoder Representation from Transformers) [AARS20], einem vortrainierten Transformer-Modell, welches kontextabhängige Vektordarstellungen (Embeddings) von Wörtern und Sätzen erzeugt [DEVL19]. Diese Embeddings können für diverse Aufgaben im Bereich des NLP wie Textklassifizierung, Fragenbeantwortung und Übersetzung verwendet werden [DEVL19].

Während BERT bei vielen NLP-Aufgaben hervorragende Ergebnisse erzielt, ist es für den semantischen Vergleich vieler Sätze ineffizient, da es Sätze stets paarweise verarbeitet. Die beiden zu vergleichenden Sätze werden als gemeinsame Eingabe in das Modell gegeben, wodurch ein semantischer Ähnlichkeitswert für dieses Paar ausgegeben wird. Diese Methode skaliert jedoch schlecht für umfangreiche Mengen an Sätzen. So erfordert beispielsweise die Identifikation des semantisch ähnlichsten Satzpaares in einer Sammlung von $n = 10000$ Sätzen insgesamt $\frac{n*(n-1)}{2} = 49.995.000$ paarweise Vergleiche. Diese enorme Anzahl an Modellaufrufen führt auf modernen GPUs zu einer Rechenzeit von etwa 65 Stunden. Zur Lösung dieses Skalierungsprob-

lems entwickelten REIMERS UND GUREVYCH 2019 das erste Sentence Transformer-Modell mit dem Namen Sentence-BERT (SBERT). SBERT nutzt eine sogenannte siamesische Architektur, bei der zwei BERT-Modelle mit identischen Parametern jeweils einen Satz $i$ unabhängig voneinander in einen Vektor $v_i$ überführen, der anschließend gespeichert wird. Da jeder Satz bei dieser Methodik nur einmal durch das Modell laufen muss, wird die Anzahl der Modellaufrufe auf nur 10000 und die resultierende Rechenzeit auf etwa 5 Sekunden reduziert. Die semantische Ähnlichkeit zwischen den mit SBERT erzeugten Embeddings zweier Sätze kann darauf aufbauend durch den mathematischen Vergleich ihrer Vektoren bestimmt werden. [REIM19]

Hierfür wird häufig die Cosinus-Ähnlichkeit berechnet, welche die relative Ausrichtung der Vektoren im Raum ermittelt [THAK21; MUEN23]. Die Cosinus-Ähnlichkeit wird definiert als das Verhältnis des Skalarprodukts zweier Vektoren zu dem Produkt ihrer euklidischen Normen [MANN09]:

$$\cos(\theta) = \frac{a \cdot b}{||a||\ ||b||} = \frac{a \cdot b}{\sqrt{a^2 \cdot b^2}} = \frac{\sum_{i=1}^{n} a_i \cdot b_i}{\sqrt{\sum_{i=1}^{n}(a_i)^2} \cdot \sqrt{\sum_{i=1}^{n}(b_i)^2}} \quad (1)$$

### 2.2.5 Computer Vision - YOLO

*Computer Vision - YOLO*

Computer Vision bezeichnet den Einsatz von technischen Systemen zur Erfassung und Analyse visueller Informationen und wird in der Produktionstechnik insbesondere für die automatisierte Qualitätskontrolle und Prozessüberwachung eingesetzt [SHAR23]. Die Grundlage moderner Computer Vision-Systeme bilden oftmals CNN, welche basierend auf ihrem Training Merkmale in Bilddateien klassifizieren, lokalisieren und segmentieren können [PAAS20]. In Abbildung 2.4 ist die Unterscheidung dieser Mechanismen dargestellt.




**Abbildung 2.4: Unterscheidung von Klassifizierung, Lokalisierung, Segmentierung und Objektdetektion nach [DIWA23]**

*Differentiation of Classification, Localization, Segmentation and Object Detection According to [DIWA23]*

Während bei der Klassifizierung ein Bild basierend auf seinem Inhalt einer vordefinierten Klasse zugeordnet wird, bestimmt die Lokalisierung die Position eines Objektes mithilfe einer Bounding Box [DIWA23; PAAS20]. Die Segmentierung steigert die Genauigkeit der Lokalisierung, indem die exakten Pixelbereiche der Kanten des Objektes bestimmt und damit die Konturen des Objektes im Bild ermittelt werden [DIWA23; PAAS20] . Gemäß DIWAN ET AL. beschreibt der Begriff *Objektdetektion* (oder auch Objekterkennung) darüber hinaus die Kombination aus Klassifizierung und Lokalisierung oder Segmentierung zur Differenzierung verschiedener Objekte in einem Bild [DIWA23].

YOLO (You Only Look Once) zählt derzeit zu den meistgenutzten Modellen im Bereich der Objektdetektion und zeichnet sich durch eine optimale Balance zwischen Geschwindigkeit und Genauigkeit aus [TERV23]. Diese Eigenschaft resultiert aus dem Umstand, dass YOLO im Gegensatz zu anderen Ansätzen das *gesamte* Bild genau *einmal* betrachtet und dabei direkt eine Objektdetektion als Kombination von Klassifizierung und Lokalisierung durchführt [REDM16]. Seit der Veröffentlichung von YOLOv1 im Jahr 2016 wurde die Modellarchitektur kontinuierlich weiterentwickelt und verbessert [HUSS23; HUSS24]. Mit der Publikation von YOLOv10 durch Wang et al. im Mai 2024 [WANG24] wurde das bis dahin hinsichtlich Genauigkeit und Leistungsfähigkeit als effizientestes Modell geltende YOLOv8 übertroffen [HUSS24].

### 2.2.6 Automatische Sprachverarbeitung

*Automated Speech Processing*

Die Realisierung einer intuitiven Kommunikation zwischen Mensch und Maschine über natürliche Sprache stellt seit langem ein zentrales Ziel in der Erforschung und Entwicklung von Benutzerschnittstellen dar [SUN23]. In den vergangenen Jahren konnte die Qualität der Sprachverarbeitungssysteme durch den exponentiellen Anstieg verfügbarer, annotierter Daten sowie den umfassenden Einsatz von Deep-Learning-Techniken und der gesteigerten Rechenleistung moderner GPU signifikant verbessert werden [ROGE22]. Allgemein bekannte Beispiele für die Anwendung dieser Technologie sind virtuelle Assistenten wie Amazons Alexa, Apples Siri und Google Assistant, die natürliche Sprache mittels Speech-to-Text (STT)-Technologien erfassen [REDD23], den Inhalt analysieren und anschließend eine Antwort über ein Text-to-Speech (TTS)-Interface generieren [AHMA24].

Moderne STT-Systeme basieren auf Deep Neural Networks (DNN), die direkt die Wellenform eines Audiosignals, also die zeitlichen Schwankungen des Schalldrucks, analysieren und in Wörter transkribieren [GEOR21]. Ein bahnbrechendes Modell in diesem Bereich war *DeepSpeech*, eines der ersten erfolgreichen End-to-End-STT-Systeme, das auf einem Recurrent Neural Network (RNN) trainiert wurde und direkt auf rohen Audiodaten operierte [HANN14]. Mittlerweile wurde dieser Ansatz durch Transformer-basierte Modelle abgelöst, die signifikante Leistungssteigerungen aufweisen. Ein prominentes Beispiel hierfür ist *Whisper* von OpenAI, das auf einem umfangreichen Datensatz von 680.000 Stunden trainiert wurde und dadurch eine robuste

Spracherkennung selbst bei der Verarbeitung von Audiosignalen mit schwierigen Bedingungen wie Akzenten, Hintergrundgeräuschen oder domänenspezifischen Fachterminologien bietet [REDD24]. Whisper nutzt eine Transformer-Architektur, um die Tonspur der eingegebenen Audiodatei basierend auf Stille, Pausen und Intonationsmustern in kleinere Passagen zu segmentieren und anschließend eine zeitlich ausgerichtete Transkription durchzuführen [RADF23].

Die Synthese von Text-to-Speech (TTS), bei der ein in natürlicher Sprache formulierter Text in ein entsprechendes Audiosignal übersetzt wird, erfolgt üblicherweise in einem zweistufigen Prozess [SHEN18; AHMA24]: Zunächst wandelt ein akustisches Modell, der Feature Generator, den Eingabetext in ein Spektrogramm um. In der zweiten Stufe wird dieses Spektrogramm von einem Vocoder, einem spezialisierten neuronalen Netzwerk, in ein kontinuierliches Audiosignal im Zeitbereich konvertiert. Abhängig vom Anwendungsfall können in beiden Stufen unterschiedliche Modellarchitekturen integriert werden. Ein aktueller Trend zeigt, dass zunehmend Transformer-Modelle für die erste Stufe verwendet werden, während für die zweite Stufe verstärkt Generative Adversarial Networks (GAN) zum Einsatz kommen [AHMA24].

Die Open-Source-Bibliothek *CoquiTTS* bietet eine umfangreiche Sammlung verschiedenster vortrainierter Modelle für beide Stufen des TTS-Prozesses und erleichtert durch ihre Funktionen sowohl die Implementierung als auch das individuelle, bedarfsgerechte Training und Fine-Tuning der Modelle [COQU23].

### 2.2.7 Hand-Tracking und Gestenerkennung

*Hand-Tracking and Gesture Recognition*

Die jüngsten Fortschritte im Bereich der KI haben die Erfassung, Identifikation und Differenzierung menschlicher Aktivitäten durch kamerabasierte Systeme maßgeblich verbessert, indem sie eine präzise Analyse komplexer Muster und Merkmale in visuellen Daten ermöglichen [GUPT22]. Dabei steht insbesondere das Tracking von Handbewegungen im Fokus, da diese eine zentrale Rolle in der menschlichen Interaktion mit der Umwelt spielen, sei es durch Berührung, Gestikulation, die Bedienung von Steuerungselementen oder die Nutzung von Werkzeugen [SIMO17]. Aus technologischer Sicht werden gegenwärtig überwiegend drei Ansätze zur Umsetzung von kamerabasiertem Hand-Tracking unterschieden: *Stereo-Kameras, Tiefensensoren* und *RGB-Kameras*.

Die Verwendung von *Stereo-Kameras* basiert auf dem Prinzip der Stereo-Vision, das die Bestimmung der dreidimensionalen Koordinaten eines Objektes anhand von zwei simultan aufgenommenen Bildern aus leicht versetzten Perspektiven ermöglicht [HART03]. Die Tiefe eines Punktes wird berechnet, indem das Produkt aus der Brennweite der Kamera und dem Abstand der beiden Kameralinsen (Base-line) durch die Disparität geteilt wird, welche die Differenz der horizontalen Bildkoordinaten desselben Objektes in beiden Bildern beschreibt [HART03]. Zur Minimierung optischer Verzerrung und Rekonstruktionsfehler ist eine präzise Kalibrierung der Kameras bei diesem Ansatz unerlässlich [HART03]. Da diese klassische Berechnung der Disparität fehler-

anfällig ist und bereits kleinste Abweichungen große Auswirkungen auf die Tiefenberechnung haben, werden in aktuelleren Ansätzen neuronale Netze eingesetzt, um die dreidimensionalen Koordinaten mit einer höheren Genauigkeit zu ermitteln [LI22; ZHAN16]. Dabei besteht sowohl die Möglichkeit, die Disparitätsberechnung zu optimieren [ZHAN16] als auch gänzlich auf eine explizite Berechnung der Disparität zu verzichten und stattdessen direkt aus den Stereo-Bildern die 3D-Koordinaten zu ermitteln [LI22].

Kommerzielle *Tiefensensoren* basieren üblicherweise entweder auf dem Prinzip der Time-of-Flight (ToF) oder dem Einsatz von strukturiertem Licht. Bei ToF wird ein Lichtimpuls von einer Lichtquelle ausgestrahlt, trifft auf ein Objekt und wird auf einen Sensor reflektiert [HANS13]. Durch die Messung der Zeit, die das Licht für diesen Weg benötigt, wird die Entfernung des Objektes zum Sensor berechnet [HANS13]. Alternativ kann ein moduliertes Lichtsignal ausgesendet werden, um die Phasenverschiebung zwischen gesendetem und reflektiertem Signal zur Ermittlung der Entfernung heranzuziehen [HANS13]. Der Nutzung von strukturiertem Licht liegt ein vergleichbares Konzept wie bei der Stereo-Vision zugrunde: Ein Projektor projiziert ein definiertes Lichtmuster (z.B. ein Gitter) auf das betrachtete Objekt, dessen Verzerrung durch das Objekt von einer Kamera erfasst wird [ZANU16]. Anhand der Disparität des Lichtmusters kann die Entfernung von der Kamera zum Objekt bestimmt werden [ZANU16]. Beide Prinzipien werden häufig in Kombination mit Methoden des maschinellen Lernens zur Gestenerkennung eingesetzt [LIU14; MARI16; DING22].

RGB-basierte Verfahren zur Gestenerkennung stellen eine leistungsfähige und kosteneffiziente Alternative zum Einsatz von Stereo-Kameras und Tiefensensoren dar, da sie ohne spezialisierte Hardware auskommen und in einer Vielzahl von Anwendungsbereichen sowie bestehenden Systemen integrierbar sind [SCHN19; SANT20]. Ein bekanntes Beispiel für die Nutzung von RGB-Daten zum Hand-Tracking und der Erkennung von Gesten ist MediaPipe Hands von Google [GOOG24]. In dieser Pipeline werden zwei neuronale Netze kombiniert: Zunächst wird das betrachtete Bild mit dem Palm Detection Model auf das Vorhandensein von Händen untersucht [ZHAN20]. Sofern eine Hand detektiert wird, identifiziert das Hand-Landmark-Model pro Hand 21 charakteristische Punkte, sogenannte Hand-Landmarks, deren jeweilige X- und Y-Koordinaten in Bezug auf ihre Position im Bild erfasst werden [ZHAN20]. Die so gewonnenen 2D-Koordinaten werden durch eine zusätzliche Schätzung der relativen Tiefe zu einer 2.5D-Repräsentation erweitert. Hierbei bestimmt das Hand-Landmark-Model die räumliche Entfernung jedes Hand-Landmarks relativ zum Handgelenks-Landmark [ZHAN20]. Während diese Schätzung eine Differenzierung von Gesten ermöglicht, stellt sie keine absolute Tiefeninformation bereit. Die Implementierung von MediaPipe Hands kann in Verbindung mit der Python-Bibliothek OpenCV erfolgen, die eine Vielzahl von Funktionen zur Bildverarbeitung bietet [BRAD08; LAGA17].

Unabhängig von der für das Hand-Tracking verwendeten Technologie wird bei der anschließenden Analyse der erfassten Gesten zwischen *statischen* und *dynamischen* Gesten unterschieden [SANT20; ZANU16]. Statische Gesten werden durch eine defi-

nierte Haltung der Hand zu einem bestimmten Zeitpunkt charakterisiert, während dynamische Gesten durch eine kontinuierliche Bewegung der Hand entlang einer Trajektorie beschrieben werden [ZANU16]. Zur Differenzierung statischer Gesten können basierend auf der Stellung der Finger und ihrer Positionen relativ zueinander unterschiedliche Griffe definiert werden [FEIX16].

### 2.2.8 Process Mining

*Process Mining*

Process Mining ist eine Technik zur Analyse von Geschäftsprozessen auf Basis von Ereignisprotokollen (Event-Logs), bei deren Anwendung reale Prozesse entdeckt, überwacht und verbessert werden können, indem Wissen aus den Ereignisprotokollen extrahiert wird [VAN 12; VAN 16a; VAN 16b]. Damit ermöglicht Process Mining Unternehmen ihre Geschäftsprozesse zu verstehen und zu managen [PETE19]. Grundlage für die Anwendung von Process Mining ist die Erstellung eines Event-Logs, welcher den Ablauf des betrachteten Prozess als Auflistung einzelner Aktivitäten (Events) beschreibt [VAN 16a]. Jeder Eintrag für ein Event enthält relevante Informationen über den zugehörigen Prozessschritt (z.B. Namen, Daten, Parameter) und ist über eine Case-ID eindeutig der betrachteten Prozessinstanz zugeordnet [VAN 12; VAN 16b]. VAN DER AALST unterscheidet drei grundlegende Arten von Process Mining [VAN 12; VAN 16a; VAN 16b]:

- **Entdeckung (Discovery):** Diese Methode zielt darauf ab, ein Prozessmodell ausschließlich auf Basis eines Event-Logs zu generieren, ohne auf zusätzliche Informationen zurückzugreifen. Sie stellt die prominenteste Technik im Bereich des Process Mining dar.
- **Übereinstimmung (Conformance):** Hierbei erfolgt ein Abgleich eines bestehenden Prozessmodells mit einem Event-Log desselben Prozesses. Ziel ist es, zu überprüfen, inwieweit das im Event-Log dokumentierte tatsächliche Prozessverhalten mit dem Modell übereinstimmt und umgekehrt.
- **Erweiterung (Enhancement):** Diese Form des Process Mining dient der Optimierung oder Erweiterung eines bestehenden Prozessmodells, indem Informationen über den tatsächlichen Prozess verwendet werden, die im Event-Log aufgezeichnet sind.

Zur Verwendung von Process Mining können je nach Anwendungsfall verschiedene Algorithmen wie Alpha-Algorithmen, heuristische Algorithmen oder induktive Miner eingesetzt werden [VAN 16b]. Insbesondere induktive Miner sind in der Lage, aus umfangreichen und komplexen Event-Logs formal korrekte Prozessmodelle abzuleiten, wo andere Algorithmen bereits an ihre Grenzen stoßen [VAN 16b]. Die von induktiven Minern ausgegeben Ergebnisse können in unterschiedlichen Visualisierungsformen wie Petri-Netzen, Business Process and Notation (BPMN)-Modellen und Prozessbäumen dargestellt werden [VAN 16b].

## 2.3 Aktuelle Ansätze für Wissensmanagementsysteme

*Current Approaches for Knowledge Management Systems*

Auf Basis der zuvor darstellten Grundlagen des Wissensmanagements, insbesondere der Methoden sowie technischen Ansätze zur Modellierung, Organisation und Speicherung von Daten, Informationen und Wissen, haben sich im Zuge der fortschreitenden Digitalisierung sogenannte Wissensmanagementsysteme (engl.: Knowledge Management Systems, KMS) als zentrale Werkzeuge etabliert. KMS sind eine Klasse von Systemen, die darauf abzielen, die Prozesse der Wissensgenerierung, -speicherung, -übertragung und -anwendung innerhalb von Unternehmen zu unterstützen und zu optimieren. Moderne KMS verwenden verschiedene IT-Technologien, um unter anderem relevante Informationsquellen zu identifizieren und zugänglich zu machen, den Austausch von Wissen in Teams zu ermöglichen sowie auf Historiendaten vergangener Projekte zuzugreifen, um aktuelle Problemstellungen effizienter verstehen und lösen zu können [ALAV01]. Diese Definition von KMS durch ALAVI UND LEIDNER verdeutlicht, dass der Begriff in der Literatur hinsichtlich der erforderlichen Funktionalitäten relativ breit gefasst wird. In Übereinstimmung damit vertreten ARIS UND SENSUSE die Ansicht, dass diverse technische Systeme mit unterschiedlichen Schwerpunkten in den Bereichen Identifikation, Erfassung, Speicherung, Verteilung und Anwendung von Wissen als KMS betrachtet werden können [ARIS20].

Im Folgenden wird ein Überblick über sowohl kommerziell erhältliche KMS als auch forschungsrelevante Systeme und Ansätze gegeben. Dabei werden überwiegend KMS betrachtet, die einen Bezug zur produzierenden Industrie aufweisen und eine potenzielle Anwendung auf dem Shopfloor unterstützen. Diese Fokussierung ist aufgrund der Vielzahl an existierenden Ansätzen für unterschiedlichste Domänen notwendig und stellt sicher, dass bei der Analyse des Stands der Technik und Forschung die spezifischen Anforderungen der Produktionstechnik berücksichtigt werden.

### 2.3.1 Kommerzielle Systeme

*Commercial Systems*

Die im Kontext dieser Arbeit als relevant identifizierten, kommerziellen KMS werden in den fünf Kategorien *Datenbanken*, *Enterprise-Content-Management-/Dokumentenmanagementsysteme, Wikis*, *Wissensdatenbanken* und *kognitive Assistenzsysteme* vorgestellt und diskutiert.

#### Datenbanken

Datenbanken bilden in Unternehmen häufig die unterste Ebene der Wissensbasis, da sie, wie der Name impliziert, primär der Speicherung von Daten dienen. Sie allein erfüllen nicht die Anforderungen eines KMS, bilden jedoch häufig die technologische Grundlage für die Wissensspeicherung und werden deshalb der Vollständigkeit halber hier benannt. Während die Datenbanktypen SQL und NoSQL im Kontext von Industrie 4.0 bereits in der produzierenden Industrie etabliert sind [OLIV22], ist der Einsatz der spezialisierteren Formen der Graph-Datenbanken und Vectore Stores noch überwiegend Gegenstand der Forschung.

Tabelle A.1 im Anhang gibt eine Übersicht bekannter Anbieter kommerzieller Versionen und Open-Source-Varianten der verschiedenen Datenbanktypen.

**Enterprise-Content-Management – Erweiterung des Dokumentenmanagements**

Der Begriff des Dokumentenmanagementsystems und die damit einhergehenden Funktionen (siehe Kapitel 2.1.4) werden zunehmend als Bestandteile von Enterprise-Content-Management-Systemen (ECM-Systeme) integriert. ECM-Systeme umfassen dabei nicht nur Funktionen für das Erstellen, Verwalten, Archivieren und den Abruf von Dokumenten, Web-Inhalten und anderen digitalen Assets, sondern darüber hinaus auch für die Definition von Prozessen und Workflows [ROSM20].

Gemäß einer repräsentativen Studie des BITKOM E.V. aus dem Jahr 2019 setzen im Mittelstand 47% der Unternehmen auf den Einsatz von ECM-Systemen für das Dokumentenmanagement, bei dem vor allem Office-Dokumente (Word, Excel, PowerPoint) und PDF-Dateien verarbeitet werden. Viele Unternehmen gaben an, im Zuge ihrer Digitalisierungsprozesse Papierdokumente einzuscannen und als PDF-Dokumente in das digitale Dokumentenmanagement einzupflegen. Die Mehrheit der befragten KMU (48%) setzt dabei auf Software in Form privater Cloud-Anwendungen, welche in externen Rechenzentren gehostet werden. [BITK19]

Aus einer ähnlichen Studie des BITKOM E.V. aus dem Jahr 2017 geht hervor, dass die befragten KMU eine Vielzahl unterschiedlicher Lösungen von insgesamt 43 verschiedenen Herstellern verwenden [BITK17]. In Tabelle A.2 im Anhang sind die fünf am häufigsten genannten Anbieter in absteigender Reihenfolge ihrer Nennung aufgeführt. Ergänzend dazu werden die vom jeweiligen Anbieter auf den offiziellen Internetseiten beworbenen Funktionen der ECM-Systeme dargestellt.

Diese fünf ECM-Systeme bieten grundsätzlich vergleichbare Funktionalitäten. Dabei ist auffällig, dass ausschließlich Dokumente und Prozesse im Bereich der Office-Workflows adressiert werden. Konkrete Funktionalitäten, die das Management von Wissen im Bereich technischer Dokumentationen oder Produktionsprozesse unterstützen, werden von den Anbietern nicht bereitgestellt. Dies wird auch durch die Studie des BITKOM E.V. von 2017 unterstrichen, aus der sich entnehmen lässt, dass nur 15,3% der Unternehmen ECM-Systeme in der Produktion einsetzen [BITK17]. Da die Studie keinen Zusammenhang zwischen den eingesetzten Systemen und ihrem jeweiligen Verwendungszweck herstellt, kann über die Gründe für diese Nutzungseinschränkungen an dieser Stelle lediglich spekuliert werden. Allerdings liegt der Schluss nahe, dass aufgrund fehlender spezifischer Funktionalitäten für den Shopfloor-Bereich die ECM-Systeme dort nur in begrenztem Umfang eingesetzt werden. Ihre Nutzung beschränkt sich dabei vermutlich überwiegend auf den Kontakt mit Office-Workflows, beispielsweise im Rahmen der Auftragsabwicklung.

**Wikis**

Im Rahmen der Recherche zur Nutzung von Wikis im Wissensmanagement wurde deutlich, dass der Großteil der wissenschaftlichen Publikationen und Studien zu diesem Thema vor mehr als einem Jahrzehnt entstanden ist. Diese Beobachtung wird

von KNEBEL UND LÄMMEL bestätigt, die präzisieren, dass Wikis in der Forschung vorwiegend im Zeitraum von 1995 bis 2010 thematisiert wurden [KNEB17]. Folglich liefert die bestehende Literatur nur begrenzte Erkenntnisse über die aktuellen Funktionen kommerzieller Wikis und deren Anwendung in produzierenden Unternehmen. Einige Autoren liefern dennoch grundlegende Bewertungen zum Einsatz von Wikis, die auch heute noch als relevant angesehen werden können:

STOCKER UND TOCHTERMANN präsentieren in ihrer Mehrfachfallstudie detaillierte Einblicke in die Verwendung von Wikis in zehn deutschsprachigen Unternehmen. Die Ergebnisse dieser Untersuchung verdeutlichen, dass Wikis überwiegend zur Unterstützung der Suche und des Auffindens relevanter Informationen genutzt werden, um die individuelle Arbeitsweise zu erleichtern. Für eine effiziente Nutzung ist es jedoch entscheidend, Wikis in klar definierte Business-Cases zu integrieren und sie somit als festen Bestandteil von Geschäftsprozessen zu etablieren. Ein reines Bereitstellen von Wikis, das auf freiwilliger Nutzung basiert, erweist sich hingegen als unzureichend, um ein effektives Wissensmanagement zu gewährleisten. [STOC10]

Die vergleichbare Studie von HACKERMEIER verfolgt ebenfalls einen empirischen Ansatz, um die psychologischen Aspekte der Akzeptanz von Wikis zu analysieren. HACKERMEIER kommt dabei zu ähnlichen Ergebnissen wie STOCKER UND TOCHTERMANN und stellt fest, dass wahrnehmbare Vorteile für die eigene Arbeit den stärksten Einfluss auf die Akzeptanz technischer Wissensmanagementsysteme haben. Die Untersuchung zeigt, dass Wikis zwar einen entscheidenden Beitrag zum Wissensmanagement leisten, jedoch den persönlichen Austausch nicht vollständig ersetzen können. [HACK12]

In den Jahren 2009 bis 2011 führte das Enterprise 2.0 Fallstudien-Netzwerk, bestehend aus Mitgliedern der Universitäten St. Gallen und Koblenz sowie der Universität der Bundeswehr München und der European Business School, insgesamt vierzehn Fallstudien in deutschen Unternehmen verschiedenster Branchen durch, bei denen die Einführung von Wissensmanagementsystemen begleitet wurde [DIGI11]. Aus den Studien lässt sich analysieren, dass die Unternehmen sehr unterschiedlich an die Thematik herangingen: Während Rheinmetall auf eine Kombination von kommerziell erhältlichen IBM Services setzte, um unter anderem Wiki-Funktionalitäten in ihr System einzubringen [KOCH11], wurde bei der ABB AG das Angebot von Microsoft SharePoint als Grundlage genutzt [STEI10]. Bei Siemens hingegen entschied man sich, eine nicht näher benannte, in Java entwickelte Open-Source-Software zu verwenden [MÖRL11]. Die Auswahl der jeweils verwendeten Software basierte dabei auf den individuellen Anforderungen an die Funktionalitäten der Systeme und die Einbettung in die bereits bestehende Softwarelandschaft. Sofern kommerziell angebotene Software nicht die erforderlichen Funktionalitäten bietet, tendieren insbesondere große Unternehmen mit umfangreichen IT-Abteilungen dazu, ihre eigenen Systeme zu implementieren.

In diesem Zusammenhang analysiert ZAPP, dass am Markt verfügbare IT-Werkzeuge für das Wissensmanagement aufgrund mangelnder Flexibilität und zu zeitintensivem Nutzungsaufwand in der Produktion nur begrenzt verbreitet sind [ZAPP14]. Aufgrund

dessen stellt ZAPP ein semantisches Wiki vor, welches ein Wissensmodell nutzt, um das von Mitarbeitenden dokumentierte Wissen mit Daten aus dem Shopfloor zu verknüpfen [ZAPP13; ZAPP14]. Zudem postuliert ZAPP, dass Wissensmanagementsysteme für den Shopfloor die Erfassung unstrukturierter, multimodaler Inhalte erlauben und intuitiv in die Produktionsumgebung und -prozesse eingebettet sein müssen [ZAPP14].

Im Jahr 2013 führten FIGURA UND GROSS einen Vergleich von sieben damals weit verbreiteten Wiki-Anbietern durch (BlueSpice, Confluence, Drupal Wiki, TikiWiki, TWiki, SharePoint) und gelangten zu dem Schluss, dass die Systeme unterschiedliche Stärken und Schwächen aufweisen. Sie betonen, dass die Auswahl des optimalen Wikis für ein Unternehmen stark von den individuellen Anforderungen abhängt. Zudem argumentieren sie, dass Wikis insbesondere für kleinere Unternehmen geeignet sind, da der Verwaltungsaufwand bei einer größeren Wissensbasis erheblich zunimmt und die Suchqualität der analysierten kommerziellen Wikis als unzureichend bewertet wurde. Als Alternative empfehlen FIGURA UND GROß die Nutzung von Wissensdatenbanken. [FIGU13]

Neben diesen älteren wissenschaftlichen Arbeiten existieren vereinzelt auch neuere Untersuchungen zum Thema Wikis: So analysieren CLEMENTE UND DOMINGUES den Einsatz von kommerziell erhältlichen Projektmanagement-Tools im Wissensmanagement [CLEM23]. Zu den untersuchten Tools zählt auch Confluence von Atlassian, das allgemein als Wiki eingestuft wird. Sie gelangen zu dem Ergebnis, dass Confluence eine zentrale Speicherung und Bearbeitung von Projektdokumentationen ermöglicht, den Wissensaustausch innerhalb von Teams fördert und eine nachvollziehbare Historie für zukünftige Projekte bereitstellt. Aufgrund dieser Eigenschaften wird Confluence als insgesamt gut geeignet für den Einsatz im Wissensmanagement bewertet [CLEM23]. Dabei liegt der Schwerpunkt des Tools insbesondere auf der Wissenserfassung und -speicherung, weniger jedoch auf der praktischen Anwendung des Wissens [CLEM23]. Diese Einschätzung zu Confluence als geeignetem Werkzeug für Wissensmanagement teilt auch SANUGOMMULA in ihrer Arbeit [SANU23]. Zusätzlich stellt sie mit Notion, Microsoft SharePoint, Google Workspace und Slite vier aktuelle Alternativen vor, mit denen ähnliche Funktionalitäten wie die von Confluence realisiert werden können [SANU23].

In Tabelle A.3 im Anhang sind die heute noch relevanten Wikis aus den Untersuchungen von FIGURA UND GROSS sowie die von SANUGOMMULA analysierten Tools dargestellt. Ergänzend werden Anbieter aufgeführt, die im Rahmen dieser Arbeit durch eine eigenständige Marktrecherche zusätzlich identifiziert wurden. Die Tabelle gibt eine Übersicht über die jeweiligen Systeme und ihre beworbenen Funktionalitäten. Bei der Analyse der in Tabelle A.3 benannten Funktionalitäten fällt auf, dass die deutschen Anbieter BlueSpice und QWiki explizit die Unterstützung ISO-konformer Prozesse hervorheben. Diese Eigenschaft kann für viele deutsche Unternehmen ein Kaufargument darstellen, da sie zur Aufrechterhaltung bestehender Zertifizierungen beitragen kann. Abgesehen davon weisen alle betrachteten Systeme ähnliche Funktionalitäten auf, wobei die größeren Anbieter wie Atlassian, Google und Microsoft verstärkt den Einsatz

von KI-gestützten Features betonen. Analog zu den ECM-Systemen finden sich auch bei den Wikis keine Anbieter, die ausdrücklich auf die Unterstützung produktionsnaher Dateien und Prozesse ausgerichtet sind. Allerdings bieten Open-Source-Plattformen wie BlueSpice und Wiki.js aufgrund ihrer offenen Architektur das Potenzial, spezialisierte Funktionen für diesen Anwendungsbereich individuell zu implementieren.

**Wissensdatenbanken**

Am Beispiel von Microsoft SharePoint, welches in der Literatur sowohl im Bereich der ECM-Systeme als auch der Wikis genannt wird, zeigt sich, wie fließend die Übergänge zwischen den Definitionen dieser Systemkategorien sind. Der Begriff *Wissensdatenbank* erweitert den Kreis der Terminologien, die Software zur Speicherung, Verwaltung und zum Abruf von Wissen beschreiben. SCHWARZ nennt in einem Blog-Beitrag Confluence, Notion, SharePoint, Guru und Bloomfire als Beispiele für Werkzeuge zur Realisierung von Wissensdatenbanken [SCHW24]. Diese Nennung von Confluence, Notion und SharePoint verdeutlicht die Überlappungen und fließenden Grenzen zwischen den Begrifflichkeiten. Guru und Bloomfire werben beide mit der KI-basierten Erstellung einer zentralen Wissensdatenbank, die effizient durchsucht werden kann [BLOO25; GURU25]. Dieser Fokus auf Funktionalitäten, die auch bei vielen der in Tabelle A.3 aufgeführten Tools zu finden sind, verdeutlicht erneut die fehlende klare Trennung zwischen den verschiedenen Systemkategorien.

**Kognitive Assistenzsysteme**

Neben den bereits erörterten kommerziellen Systemen für strukturiertes Wissensmanagement, die primär auf den systematischen Umgang mit Daten und Dokumenten ausgerichtet sind, werden an dieser Stelle kognitive Assistenzsysteme zur Unterstützung der Mitarbeitenden auf dem Shopfloor betrachtet.

DAS BÜRO FÜR TECHNIKFOLGEN-ABSCHÄTZUNG BEIM DEUTSCHEN BUNDESTAG (TAB) definiert kognitive Assistenzsysteme als Technologien, die Mitarbeitende durch bedarfsgerechte Informationsbereitstellung, Hinweise auf korrekte Arbeitsabläufe und die Überprüfung ausgeführter Arbeitsschritte unterstützen. Sie bieten das Potenzial, die Produktivität, Qualität und Kosteneffizienz durch die enge Interaktion von Mensch und Technik zu steigern. Ein recht weit verbreitetes Beispiel sind Systeme zur Werkerführung, die insbesondere in der Montage oder Kommissionierung eingesetzt werden. Diese Systeme können idealerweise individuell auf die Fähigkeiten der Mitarbeitenden angepasst werden, um Fähigkeitslücken durch Wissensvermittlung sowie Lernerfolgskontrolle zu schließen. Damit dienen sie auch als Werkzeuge zur Weiterbildung im Arbeitsprozess. [BOVE20]

Das TAB gibt an, dass Pick/Place-To-Light-Systeme heute zu den am häufigsten eingesetzten kognitiven Assistenzsystemen in der manuellen Montage zählen und hebt Systeme der *LAP GmbH Laser Applikationen* und der *MiniTec GmbH* als positive Beispiele hervor [BOVE20]. Während die LAP GmbH Laser Applikation Laserprojektionssysteme zur Anzeige von Montagepositionen auf Montageplatten vertreibt [LAP 25], hat die MiniTec GmbH den MiniTec SmartAssistant entwickelt, der die Mitarbeitenden

durch einen Montageprozess leitet und die korrekte Ausführung kontrollieren kann [MINI24]. In Kombination mit diesem digitalen Assistenzsystem bietet MiniTec GmbH komplette Arbeitsplätze ausgestattet mit Laserprojektor, Pick-to-Light-System, Bilderkennung, Touchdisplay und Handscanner an, um den vollen Funktionsumfang des SmartAssistant nutzbar zu machen [MINI24]. Die Erstellung der Arbeitsanleitungen erfolgt dabei von erfahrenen Mitarbeitenden über einen entsprechenden Editor [MINI24].

BERTRAM UND BIRTEL geben einen Überblick kommerzieller kognitiver Assistenzsysteme und vergleichen deren Funktionsumfang mit prototypischen Entwicklungen aus Forschungsprojekten. Die in der Analyse berücksichtigten kommerziellen Systeme sind im Anhang in Tabelle A.4 zusammen mit ihren jeweiligen Funktionalitäten aufgelistet. Laut BERTRAM UND BIRTEL verfügen alle vier Systeme über Funktionen zur Werkerführung bei Arbeitsaufgaben, zur Erfassung von Werkzeugen und Komponenten sowie zur Verfolgung des Montagefortschritts. Allerdings bietet keines der Systeme eine automatisierte Erstellung von Programmabläufen oder Anweisungen für die Werker, wodurch diese manuell generiert werden müssen. [BERT18]

Zusätzlich zu den in Tabelle A.4 dargestellten Anbietern wurden im Rahmen der Literaturrecherche für die vorliegende Arbeit weitere Anbieter identifiziert, die ähnliche Systeme zur Werkerführung durch Visualisierungen auf Bildschirmen, Projektionen oder Pick-to-Light-Technologien sowie Kamera-basierte Fortschritts- und Qualitätskontrollen anbieten. Die Funktionalitäten und technischen Ansätze dieser Systeme scheinen jedoch, basierend auf den verfügbaren Informationen, denen der in der Analyse beschriebenen Systeme weitgehend zu entsprechen. Ein Anbieter, der mit einem neuartigen technischen Ansatz hervorsticht, ist das Start-up Anticipate GmbH. Das von Anticipate entwickelte System zur Werkerführung und Montagefortschrittskontrolle integriert eine KI-basierte visuelle Qualitätskontrolle, die eine Kombination aus Hand-Tracking und Bilderkennung nutzt, um die Aktivitäten der Mitarbeitenden bei der Montage zu erfassen [ANTI25].

KLAPPER ET AL. führten 2019 eine empirische Untersuchung durch, um die Potenziale digitaler Assistenzsysteme zu analysieren. Hierbei wurden 144 Personen aus produzierenden Unternehmen zu ihrer Einschätzung verschiedener Technologien befragt. Die Studie ergab, dass Unternehmen neben den bereits etablierten Technologien wie der Visualisierung von Prozessschritten und der Nutzung von Bilderkennung für Qualitäts- und Prozessfortschrittskontrollen insbesondere auch großes Potenzial im Einsatz von Augmented Reality in den Bereichen der Informationssammlung und -aufbereitung, der Entscheidungsunterstützung sowie des Erlernens von Abläufen sehen [KLAP19].

Der Begriff Augmented Reality (AR) bezeichnet eine physische Realität, die durch die Integration virtueller Elemente ergänzt wird. Ziel von AR ist es, die Wahrnehmung eines Nutzers zu erweitern, indem die reale Umgebung mit der Darstellung digital generierter Informationen (Zahlen, Symbole, Audio, Video oder Grafiken) angereichert wird. Dazu kommen in der Regel sogenannte Head-Mounted-Displays (HMD) zum Einsatz,

die über Bildschirme in oder vor Brillengläsern verfügen, mittels derer die digitalen Inhalte im Sichtfeld des Nutzers erscheinen. [AREN22]

Zur Bewertung der Eignung von AR-Anwendungen zur Unterstützung von Mitarbeitenden bei der manuellen Montage nahmen ROBERTO ET AL. eine vergleichende Analyse von neun kommerziellen AR-Systemen vor: Fieldbit, Hololight, LightGuide, Plutonmen, TeamViewer Frontline, TechSee, VSight, vuforia und Zumoko CAD DT. Die Untersuchung wurde in Zusammenarbeit mit Experten aus den Bereichen Maschinenbau, User Experience (UX) und AR-Technologien durchgeführt, um verschiedene Perspektiven und fachliche Sichtweisen einzubeziehen. Die Ergebnisse zeigen, dass die evaluierten AR-Anwendungen ähnliche Zielsetzungen, Funktionen und Benutzerschnittstellen aufweisen. Der Fokus liegt dabei primär auf der Remote-Unterstützung bei der Aufgabenausführung sowie der Befähigung zur Durchführung von Montageprozessen. In diesen Anwendungsfeldern tragen die durch AR bereitgestellten Informationen zur Steigerung der Effizienz der Nutzenden bei. Gleichzeitig offenbarten die Analysen jedoch Schwächen, insbesondere in Bezug auf die Robustheit der Anwendungen, den Umfang der verfügbaren Individualisierungsmöglichkeiten sowie die Benutzerfreundlichkeit der vorhandenen Autorentools zur Erstellung von AR-Inhalten und Arbeitsanweisungen. [ROBE24]

Die Vuforia-Plattform von PTC bietet neben den bereits benannten Funktionen die Komponente *Expert Capture*, welche die Erfassung und Verbreitung von Expertenwissen durch AR-gestützte Arbeitsanweisungen unterstützt. Erfahrene Mitarbeitende nutzen die *vuforia capure* App, um Prozessabläufe in Form von Schritt-für-Schritt-Anweisungen per Augmented-Reality-Geräten (z.B. AR-Brille) aufzuzeichnen. Anschließend können diese Aufzeichnungen im *vuforia editor* strukturiert aufbereitet und unerfahrenen Mitarbeitenden per *vuforia vantage* als AR-Inhalte dargestellt werden. [PTC21]

### 2.3.2 Experimentelle Systeme aus der Forschung

*Experimental Systems from Research*

Das Ziel der Literaturrecherche des Stands der Forschung war es, sowohl praktische als auch theoretische Ansätze zu identifizieren, die Innovationen in Bezug auf Methodik oder eingesetzte Technologien für die Implementierung von Wissensmanagement auf dem Shopfloor produzierender Unternehmen aufzeigen. Darüber hinaus wurden auch Technologien und Methoden aus anderen Domänen berücksichtigt, deren Anwendung Potenzial für den Einsatz in KMS bietet. Die als relevant bewerteten experimentellen Systeme und Methoden werden in Anlehnung an die Schritte des strukturierten Wissensmanagements in drei Kategorien eingeteilt: Wissenserfassung, Wissensspeicherung und Wissensbereitstellung. Entlang dieser Systematik wird im Folgenden der aktuelle Stand der Forschung dargestellt und diskutiert.

### Wissenserfassung

Das Thema der Wissenserfassung ist in der aktuellen Forschung Gegenstand verschiedener wissenschaftlicher Disziplinen.

PRATIWI ET AL. stellen einen Ansatz aus qualitativen und quantitativen Methoden vor, bei dem Experteninterviews, Fokusgruppen, Umfragen und Fragebögen als Werkzeuge für eine effiziente Erfassung von Wissen genutzt werden [PRAT24]. Durch die Kombination der verschiedenen Erhebungsmethoden werden sowohl individuelle Erfahrungen und Expertise erfasst als auch Möglichkeiten zum interdisziplinären Austausch geschaffen [PRAT24]. Die Autoren kommen zu dem Schluss, dass Unternehmen durch die systematische Anwendung dieser Ansätze ein effektives Wissensmanagement etablieren können [PRAT24]. Eine kritische Auseinandersetzung mit potenziellen Herausforderungen dieses arbeitswissenschaftlichen Vorgehens, insbesondere dem zusätzlichen Aufwand für nicht unmittelbar wertschöpfende Tätigkeiten, bleibt jedoch aus.

Einen deutlich technischeren Ansatz verfolgen KHAN ET AL. in ihrer Arbeit, in der sie den Einsatz eines DMS für die Erfassung von explizitem Wissen validieren [KHAN15]. Dazu implementieren sie ein eigenes DMS basierend auf der Oracle Developer Suite und der Oracle Database, statten es mit den bekannten Funktionen eines DMS (siehe Kapitel 2.1.4) aus und validieren seine Funktionalität anhand von Personaldaten aus dem Bereich der Human Ressources [KHAN15]. Obwohl die Autoren ihr System als *Explicit Knowledge Management System* bezeichnen, weist es keine signifikanten Neuerungen gegenüber kommerziell verfügbaren DMS-Lösungen auf. Zudem bleibt die Herausforderung der Erfassung von implizitem Wissen unberücksichtigt.

Im Gegensatz zu den zuvor behandelten Ansätzen fokussiert sich HEZLER explizit auf Prozesse auf dem Shopfloor produzierender Unternehmen. Er stellt die Implementierung einer Wissensmanagementsoftware bei einem Automobilzulieferer vor, die zur systematischen Erfassung von Störungen, deren Ursachen sowie effektiven Gegenmaßnahmen dient. Das System ermöglicht es Mitarbeitenden jederzeit neue Einträge bestehend aus Text und Dateien (Bilder, Videos etc.) zu erstellen und bereits hinterlegte Störungen über ein Kategoriensystem zu identifizieren. Eingetragene Informationen werden unmittelbar für alle Nutzer zugänglich gemacht, wobei eine nachträgliche Validierung und Anpassung durch Experten optional vorgesehen ist. Die durch das System ermöglichte Wiederverwendbarkeit von Erfahrungswissen trägt laut HEZLER nachweislich zur Reduktion von Anlagenstillständen bei und steigert die Produktivität in der Fertigung signifikant. [HEZL20]

Während die in [PRAT24], [KHAN15] und [HEZL20] dargestellten Untersuchungen methodisch eng an den Prinzipien des Wissensmanagements ausgerichtet sind, präsentieren andere Autoren innovativere Ansätze. ANDRIANAKOS ET AL. untersuchen die automatisierte Überwachung manueller Montageprozesse unter Einsatz von Computer Vision und Machine Learning, um Abweichungen in der Montagereihenfolge in Echtzeit zu erkennen. Dazu wird ein Kamerasystem eingesetzt, welches den Montageablauf kontinuierlich verfolgt und mithilfe von neuronalen Netzen sowohl die Bauteile als auch

die Hände der Mitarbeitenden im Video detektiert. Zur Bestimmung der Interaktion zwischen Hand und Bauteil berechnet das System die Überlappung der Bounding-Boxen, wobei eine Überschreitung eines definierten Schwellenwerts als Indikator für das Greifen eines Bauteils dient. Durch die Kombination von Objektdetektion und Hand-Tracking kann das System die korrekte Montagereihenfolge überprüfen und bei Abweichungen in der Ausführung unmittelbar eine visuelle Rückmeldung an die Mitarbeitenden über ein Bildschirminterface ausgeben. [ANDR19]

Einen ähnlichen Ansatz zur feingranularen Erkennung von Arbeitsabläufen und darauf basierender Fehleridentifikation präsentieren KOBAYASHI ET AL., indem sie die Erkennung und Klassifikation der Position beider Hände per Machine Learning als zentrale Methode zur Differenzierung einzelner Montageschritte heranziehen. Dabei wird jede Hand durch einen Punkt im Bild in zweidimensionalen Koordinaten repräsentiert. Diese Positionsinformationen werden mit Bildmerkmalen aus der Umgebung der Hände kombiniert, um unterschiedliche Montageschritte zu klassifizieren. [KOBA19]

FULLEN ET AL. fokussieren sich auf die automatische Erkennung des aktuellen Montageschritts, um die entsprechenden Montageanleitungen bedarfsgerecht bereitzustellen. Sie verwenden eine Tiefenkamera zur Erfassung der Handbewegungen, wobei jede Hand als ein dreidimensionaler Punkt im Raum repräsentiert wird [FULL19]. Die erfassten Bewegungsdaten werden als Zeitreihe verarbeitet und mittels Machine Learning klassifiziert, um den aktuellen Montageschritt automatisch zu bestimmen [FULL19]. Dieses Ziel wird auch von LEHMANN ET AL. verfolgt, die jedoch anstelle optischer Systeme Datenhandschuhe nutzen, um die ausgeführten Griffe zu erfassen und darauf basierend den aktuellen Montageschritt zu identifizieren [LEHM22].

FACCIO ET AL. gehen einen Schritt weiter und bedienen sich eines Systems von im Raum verteilten Tiefenkameras, um die Bewegungen des ganzen Körpers der Mitarbeitenden vollständig zu erfassen und daraus ein digitales 3D-Modell der Arbeitsabläufe zu generieren. Dieses Modell ermöglicht eine detaillierte Analyse von Arbeitswegen, der Handhabung von Bauteilen sowie der Nutzung von Werkzeugen und dient der präzisen Zeiterfassung einzelner Tätigkeiten, wodurch eine objektive Bewertung und Optimierung manueller Arbeitsprozesse ermöglicht wird. [FACC19]

In [ANDR19], [KOBA19], [FULL19] und [FACC19] wird somit die Anwendung von Computer Vision und Machine Learning zur Erfassung menschlicher Bewegungen in industriellen Prozessen mit dem Ziel untersucht, die Effizienz der Arbeiten zu steigern und Fehler zu minimieren. Die vorgestellten Systeme besitzen jedoch darüber hinaus das Potenzial für die Erfassung von Wissen im Sinne der Evaluation unterschiedlicher Arbeitsstrategien zur Identifikation optimaler Vorgehensweisen eingesetzt zu werden. Aktuell wird dies nur von FACCIO ET AL. im Ausblick ihrer Arbeit in Betracht gezogen [FACC19].

Andere Autoren haben dieses Potenzial jedoch bereits erkannt. So stellen BURGGRÄFE ET AL. ein Konzept zur automatischen Generierung von Montageanleitungen vor, um den hohen manuellen Aufwand bei deren Erstellung zu reduzieren und gleichzeitig

implizites Wissen von erfahrenen Mitarbeitenden zu erfassen. Dabei werden die Montageprozesse per Kamera aufgezeichnet und mittels Deep-Learning-Algorithmen analysiert, wodurch Bauteile, Werkzeuge und Handbewegungen erkannt werden. Die gewonnenen Informationen werden im nächsten Schritt zu einer strukturierten Montageanleitung zusammengeführt, die eine schrittweise Beschreibung der Tätigkeiten enthält. Diese Anleitungen können unerfahrenen Mitarbeitenden per User-Interface bereitgestellt werden. Herausforderungen bestehen aktuell noch in der Anpassung des Systems an verschiedene Montageprozesse und der dafür notwendigen Erstellung qualitativ hochwertiger Trainingsdatensätze. [BURG24]

Neben der Verwendung von Computer Vision und Machine Learning-Algorithmen wird in der Forschung zunehmend auch der Einsatz von Process Mining zur Untersuchung menschlichen Verhaltens diskutiert. THEIS UND DARABI untersuchen den Einsatz von Process Mining zur Analyse von Mensch-Computer-Interaktionen, um ineffiziente Nutzerverhaltensweisen zu identifizieren und personalisierte Empfehlungen zur Optimierung der Interaktion abzuleiten [THEI19]. Die Erfassung der Nutzeraktivitäten erfolgt dabei durch das Tracking von Eingabegeräten wie Mausklicks und Tastatureingaben [THEI19]. MANNHARDT ET AL. setzen sich mit den datenschutzrechtlichen Herausforderungen des Process Mining in industriellen Umgebungen auseinander, wobei die Aktivitäten der Mitarbeitenden über Wearables erfasst werden [MANN18]. Alternativ setzen sowohl DOGAN ET AL. als auch MA'ARIF auf unterschiedliche Sensortechnologien zur Erfassung menschlichen Verhaltens. Während DOGAN ET AL. Infrarotsensoren zur Bewegungserkennung in Innenräumen verwenden [DOGA19], nutzt MA'ARIF Sensordaten aus Smart-Home-Systemen, um Verhaltensmuster zu analysieren und zu modellieren [MAAR17].

**Wissensspeicherung**

Wie bereits in Kapitel 2.1.4 dargestellt, können zur Speicherung und Modellierung von Wissen unter anderem Knowledge Graphs und Ontologien eingesetzt werden. Knowledge Graphs werden laut einer Analyse von BUCHGEHER ET AL. aktuell verstärkt für die Aggregation und Kombination („Wissensfusion") von Informationen und Daten aus unterschiedlichen Quellen verwendet und bilden damit die Grundlage für die Erstellung digitaler Zwillinge und die automatische Prozessintegration [BUCH21]. Gleichzeitig bietet diese Wissensfusion auch Vorteile für den Einsatz im Wissensmanagement, um eine ganzheitliche Wissensbasis zu schaffen.

ZHANG führt einen Vergleich von relationalen Datenbanken und Graph Datenbanken durch und kommt zu dem Schluss, dass sich die Knowledge Graphs in Kombination mit Graph Datenbanken für das Wissensmanagement in Unternehmen besonders eignen, da sie durch ihre Struktur komplexe Beziehungen exakt und gleichzeitig einfach durchsuchbar modellieren können [ZHAN17].

NORDSIECK ET AL. untersuchen die Anwendung von Knowledge Graphs zur Erfassung von implizitem Wissen im Kontext der Parametrisierung von Produktionsprozessen. Über eine speziell entwickelte grafische Benutzeroberfläche können Mitarbeitende

während der Parametrisierung gezielt jene Einflussfaktoren markieren, die ihre Entscheidungsfindung bei der Einstellung der Parameter maßgeblich beeinflusst haben [NORD21]. Diese Einflussgrößen sowie die vorgenommenen Parametrisierungen werden in einem semantischen Wissensgraphen gespeichert und bilden ein Netzwerk aus Entscheidungsregeln, das für zukünftige Produktionsprozesse als Unterstützung genutzt werden kann [NORD21].

**Wissensbereitstellung**

Im Kontext des Wissensabrufs durch Mitarbeitende aus digitalen Systemen werden derzeit verschiedene technologische Ansätze untersucht. Die jüngsten Fortschritte in der KI-gestützten Verarbeitung natürlicher Sprache, insbesondere der Einsatz von LLM, eröffnen vielversprechende Potenziale für die Implementierung eines effizienten Wissensmanagements.

Im Zuge einer Studie in einem produzierenden Unternehmen untersuchten FREIRE ET AL. den Einsatz von LLM in Kombination mit RAG, um den effizienten Abruf von Wissen aus technischen Dokumentationen (Maschinenhandbücher, Prüfprotokolle, Arbeitsanweisungen etc.) und schriftlich festgehaltenem Expertenwissen in Form von Berichten zu ermöglichen. Dabei evaluieren sie unterschiedliche LLM und bewerten deren Eignung in ihrem System anhand der Korrektheit und Vollständigkeit der generierten Antworten. Die Ergebnisse zeigen, dass GPT-4 unter den getesteten Modellen die höchste Leistungsfähigkeit für diesen Anwendungsfall aufweist. Insgesamt trägt das System dazu bei, die Informationssuche erheblich zu beschleunigen und somit die Effizienz bei der Problemlösung signifikant zu steigern. [FREI24]

BARRON ET AL. verwenden ein vergleichbares System aus LLM und RAG, ergänzen dieses jedoch zusätzlich um einen Knowledge Graph. Während im Vector Store verschiedene Dokumente gespeichert sind, stellt der Knowledge Graph Zusammenhänge und Beziehungen von Wissen dar (z.B. „Technik A nutzt Methode B“). Bei einer Abfrage an das System wird zunächst der RAG-Prozess durchlaufen. Liefert der kein eindeutiges Ergebnis, so wird im nächsten Schritt der Knowledge Graph durchsucht. Durch die Kombination der Technologien wird gemäß der Autoren eine höhere Genauigkeit und Qualität der vom LLM erzeugten Antworten erreicht. [BARR24]

Anstelle der Anwendung der RAG-Methodik verfolgen LEE ET AL. einen Ansatz, der auf das spezifische Fine-Tuning eines LLM zur Extraktion von Wissen aus technischen Dokumenten fokussiert ist. Im Rahmen ihrer Untersuchung vergleichen sie die Leistungsfähigkeit zweier auf identische Weise feinjustierter Modelle, GPT-3 und LLaMA2. Die Ergebnisse zeigen, dass LLaMA2 signifikant weniger Halluzinationen erzeugt und somit qualitativ hochwertigere Antworten liefert. [LEE24]

Die in [FREI24], [BARR24] und [LEE24] vorgestellten Systeme bieten vielversprechende Ansätze für die Implementierung eines effizienten Wissensabrufs. Allerdings spezifizieren lediglich LEE ET AL. die für die Validierung verwendete Hardware, wodurch die Bewertung der Antwortqualität der Systeme von FREIRE ET AL. und BARRON ET AL. nur eingeschränkt aussagekräftig ist. Alle drei Arbeiten gehen zudem davon aus, dass

Expertenwissen bereits schriftlich dokumentiert vorliegt und lassen damit den Prozess der Formulierung von implizitem Wissen unbeachtet.

Neben der Kombination aus LLM und RAG können jedoch noch andere Technologien für den Wissensabruf in Betracht gezogen werden. MÜLLER ET AL. untersuchen den Einsatz von Recommender Engines zur Identifikation und zum Abruf bereits dokumentierter Probleme und zugehöriger Lösungsansätze im Produktionsumfeld. Die zugrunde liegende Technologie der Recommender Engines basiert auf textbasierten Suchmechanismen, die vorhandene Dokumente nach Wortübereinstimmungen mit der Nutzeranfrage durchsuchen und relevante Inhalte extrahieren. Dabei erfolgt die Verarbeitung der Texte durch sogenannte TF-IDF-basierte Vektorisierung, wodurch sowohl die hinterlegten Problembeschreibungen als auch die Suchanfragen in numerische Repräsentanzen umgewandelt und mittels Cosinus-Ähnlichkeit verglichen werden. [MÜLL20]

WANG entwickelt ein KMS auf Basis von Support Vector Machines (SVM), das die Klassifizierung, Strukturierung und Extraktion von Wissen aus Texten ermöglichen soll. Auch in diesem Ansatz werden sowohl die gespeicherten Texte als auch die Anfragen numerisch repräsentiert, indem sie in Vektorräume transformiert werden. Auf Grundlage dieser numerischen Repräsentationen entscheidet die SVM, welche der hinterlegten Informationen den Mitarbeitenden auf Basis ihrer Anfragen zurückgegeben werden. Die eingesetzte SVM lernt durch vorheriges Training, Entscheidungsgrenzen für die verschiedenen Kategorien des hinterlegten Wissens zu definieren, um eingehende Anfragen gezielt den relevantesten Kategorien zuzuordnen. [WANG23]

Die Ansätze in [MÜLL20] und [WANG23] weisen strukturelle Ähnlichkeiten zur RAG-Methodik auf, da sie auf der numerischen Repräsentation von Wissen zur Beantwortung von Suchanfragen basieren, bieten jedoch keine semantische Analyse der Inhalte. Aufgrund dieser Einschränkung kann angenommen werden, dass Recommender Engines und SVM-basierte Klassifikationssysteme für den Wissensabruf technologisch überholt sind und durch moderne Ansätze basierend auf LLM und RAG ersetzt werden könnten.

## 2.4 Fazit zum Stand der Technik und Forschung

*Conclusion on the State of the Art and Research*

Die Analyse des aktuellen Stands der Technik verdeutlicht, dass die Integration effizienter Wissensmanagementprozesse in betriebliche Abläufe eine zentrale Herausforderung darstellt, deren Bewältigung maßgeblich von der Akzeptanz und Motivation der Mitarbeitenden abhängt [DICK02; STOC10]. Die bloße Dokumentation vorhandenen Wissens erweist sich als unzureichend; vielmehr ist eine digitale Speicherung erforderlich, die einen bedarfsgerechten, kontextsensitiven Abruf ermöglicht [ACKE18]. Insbesondere auf dem Shopfloor besteht ein erhöhter Bedarf an Systemen, die die Erfassung unstrukturierter, multimodaler Inhalte unterstützen und zugleich intuitiv in die Produktionsumgebung und -prozesse integriert sind [ZAPP14].

Die identifizierten, kommerziell erhältlichen ECM-Systeme und Wikis bieten keine spezifischen Funktionalitäten zur Unterstützung wissensintensiver Prozesse im Shopfloor-Bereich und sind deshalb nur bedingt für die Handhabung von produktionsnahem Wissen geeignet. Während moderne Assistenzsysteme zur Werkerführung teilweise über fortschrittliche Funktionen zur Unterstützung manueller Tätigkeiten verfügen, fehlt es den untersuchten Systemen an Möglichkeiten zur automatisierten Generierung von Arbeitsanweisungen oder Programmabläufen, weshalb diese aufwendig manuell erzeugt werden müssen [BERT18]. Gleichzeitig offenbart die Analyse verfügbarer AR-Tools ähnliche Schwächen, insbesondere in Bezug auf die Robustheit der Anwendungen, den Umfang der Individualisierungsmöglichkeiten sowie die Benutzerfreundlichkeit der vorhandenen Autorentools [ROBE24]. Insgesamt decken kommerzielle Systeme grundsätzlich die Erfassung, Speicherung und Bereitstellung von Wissen ab, adressieren jedoch unzureichend die Erfassung impliziten Wissens, die kontextabhängige Bereitstellung relevanter Informationen sowie die Automatisierung von Prozessen zur Wissensdokumentation.

Die in der aktuellen Forschung diskutierten experimentellen Methoden und Systeme liefern vielversprechende Ansätze, weisen jedoch weiterhin Defizite auf. Qualitative und quantitative Methoden zur Wissenserfassung, beispielsweise durch Interviews und Umfragen, wie in [PRAT24] vorgeschlagen, verursachen einen signifikanten zusätzlichen Aufwand für nicht wertschöpfende Tätigkeiten. Experimentelle DMS wie in [KHAN15] erleichtern zwar die Erfassung expliziten Wissens, bieten jedoch keine adäquaten Mechanismen zur Identifikation und Speicherung von implizitem Wissen. Andere Systeme, beispielsweise [HEZL20], sind primär für strukturierte Produktionsumgebungen konzipiert und lassen sich nur schwierig auf die variablen Rahmenbedingungen von KMU übertragen. Arbeiten wie [ANDR19], [KOBA19], [FULL19] und [FACC19] zeigen vielversprechende Ansätze zur Steigerung der Prozesseffizienz und Fehlervermeidung in Arbeitsabläufen, schöpfen jedoch das Potenzial für ein umfassendes Wissensmanagement, insbesondere hinsichtlich der Erfassung und Bewahrung von Best Practices, nicht aus. Auch RAG-Systeme wie in [FREI24], [BARR24] und [LEE24] bieten zwar fortschrittliche Mechanismen für einen effizienten Wissensabruf, adressieren jedoch nicht die initial notwendige Erfassung des Wissens.

Zusammenfassend lässt sich feststellen, dass nach Kenntnis des Autors derzeit kein ganzheitlicher Ansatz zur systematischen Bewahrung und optimalen Nutzung von implizitem Wissen in KMU existiert. Während sowohl kommerzielle als auch experimentelle Systeme einzelne Aspekte adressieren, bleiben insbesondere die Erfassung von implizitem Wissen, die Automatisierung von Wissensmanagementprozessen und die intuitive Einbettung in die betrieblichen Abläufe auf dem Shopfloor weitgehend unzureichend berücksichtigt. Dies unterstreicht die Notwendigkeit der Entwicklung neuer methodischer und technologischer Ansätze für Wissensmanagement, die gezielt auf die spezifischen Anforderungen von KMU und die Gegebenheiten in deren Produktionsumgebungen abgestimmt sind.

# 3 Zielsetzung und Vorgehensweise

*Objective and Procedure*

Bislang wird Wissensmanagement auf dem Shopfloor von KMU kaum bis gar nicht strukturiert durchgeführt. Die verfügbaren kommerziellen Methoden und Systeme konzentrieren sich überwiegend auf eine intuitive Darstellung von Informationen für die Mitarbeitenden, erfordern jedoch häufig einen erheblichen manuellen Aufwand zur initialen Dokumentation und Speicherung von Wissen und Erfahrung. Der dadurch entstehende Zeitaufwand, kombiniert mit teilweise mangelnder Benutzerfreundlichkeit und Problemen bei der kontextualisierten und zielgruppengerechten Formulierung von Abläufen, hemmen aktuell die Einführung eines prozessbegleitenden, intuitiven Wissensmanagements.

## 3.1 Zielsetzung

*Objective*

In dieser Arbeit werden Möglichkeiten aufgezeigt, den dargestellten Herausforderungen im Wissensmanagement auf dem Shopfloor zu begegnen. Auf Basis des herausgearbeiteten Handlungsbedarfs wird die folgende Forschungsfrage definiert:

**Wie kann Wissensmanagement auf dem Shopfloor von KMU derart in die Produktionsprozesse eingebettet werden, sodass neu gewonnenes Erfahrungswissen unmittelbar, intuitiv und effizient gespeichert und bei Bedarf abgerufen werden kann?**

Neben der dafür grundsätzlich notwendigen Konzeption geeigneter Methoden erfordern insbesondere die Integration des Wissensmanagements in die üblichen Arbeitsabläufe und die strukturierte Speicherung des Wissens die Beantwortung weiterer Teilfragen:

1. **Wie können Mitarbeitende ihr Wissen an bestimmten Punkten ihrer Tätigkeit intuitiv und mit möglichst minimalem Mehraufwand *aktiv* digital speichern bzw. weitergeben?**
2. **Wie kann Wissen in Form von „Best Practices“ *passiv* erfasst bzw. „getrackt“ werden, ohne die Mitarbeitenden bei der Ausführung ihrer Tätigkeiten zu behindern?**
3. **Wie kann eine kontextualisierte, strukturierte Speicherung des Wissens realisiert werden, um eine bedarfsgerechte Bereitstellung zu gewährleisten?**

Zur Beantwortung dieser Fragestellungen wird die Lösungshypothese aufgestellt, dass sich durch die geschickte Kombination

- von Technologien zur Mensch-Maschine-Interaktion,
- Methoden aus dem Bereich des Natural Language Processing (NLP),
- Methoden aus dem Bereich der Computervision
- und Methoden der Prozessbeschreibung und -analyse

ein systematisches, prozessparalleles Erfassen von Erfahrungswissen auf dem Shopfloor realisieren lässt. Vor diesem Hintergrund ist das Ziel dieser Arbeit die explorative Entwicklung und Erforschung von innovativen Konzepten des Wissenstransfers zur Realisierung einer zeit-, orts- und personenunabhängigen Verfügbarkeit von Expertenwissen in KMU.

## 3.2 Vorgehensweise

*Approach*

Der Aufbau der vorliegenden Dissertation orientiert sich an den sechs Phasen des Design Thinking-Prozesses (*Verstehen, Beobachten, Definieren, Ideen finden, Prototyp entwickeln* und *Testen*), wie sie von MÜLLER-ROTHENBERG in [MÜLL18] beschrieben sind. Die Zielstellungen der einzelnen Phasen werden inhaltlich adaptiert und in eine geeignete Kapitelstruktur überführt. Abbildung 3.1 zeigt die Zuordnung der Phasen zu den Kapiteln der Dissertation.

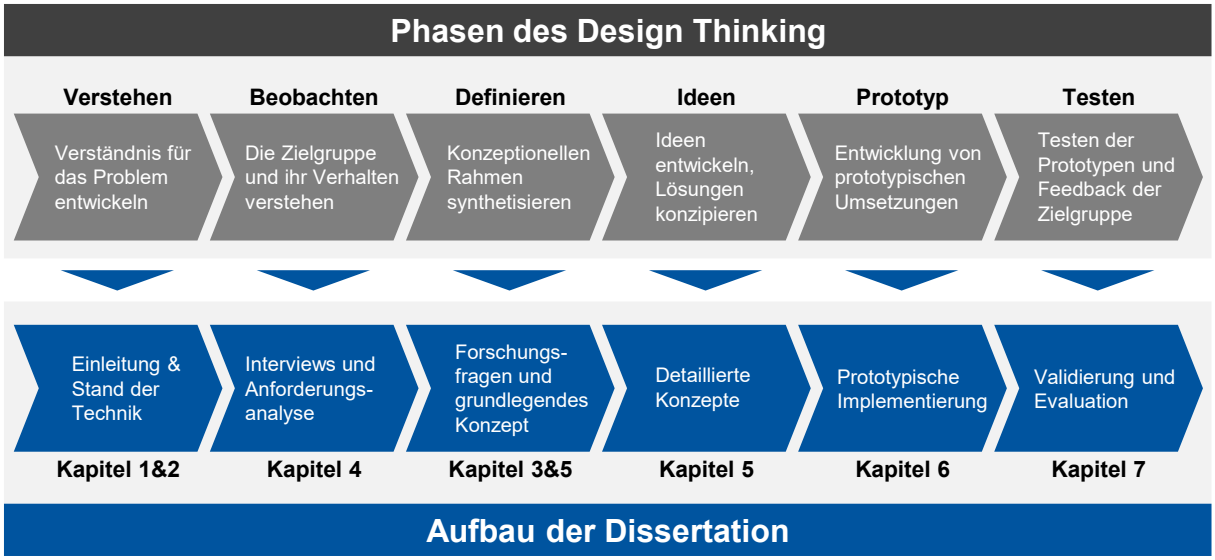



**Abbildung 3.1: Adaption der Phasen des Design Thinking für den Aufbau dieser Dissertation**

*Adaptation of the Phases of Design Thinking for the Structure of this Dissertation*

In Kapitel 2 erfolgte zunächst eine Darstellung und anschließende Analyse des aktuellen Stands der Forschung und Technik im Hinblick auf strukturiertes Wissensmanagement auf dem Shopfloor. Dabei wurden die Begriffe Wissensmanagement und Wissenstransfer eingeführt und aktuelle Methoden sowie für diese Arbeit relevante Technologien erörtert. Anschließend wurden sowohl kommerziell erhältliche Lösungen als auch prototypische Ansätze aus der Forschung diskutiert. Auf das abschließende Fazit zum Stand der Technik und Forschung folgte in Kapitel 3, unter Berücksichtigung der Defizite der bisherigen Ansätze, die Herleitung der Forschungsfrage und der damit einhergehenden Zielsetzung. Die weitere Vorgehensweise dieser Arbeit ist in Abbildung 3.2 zusammenfassend dargestellt.

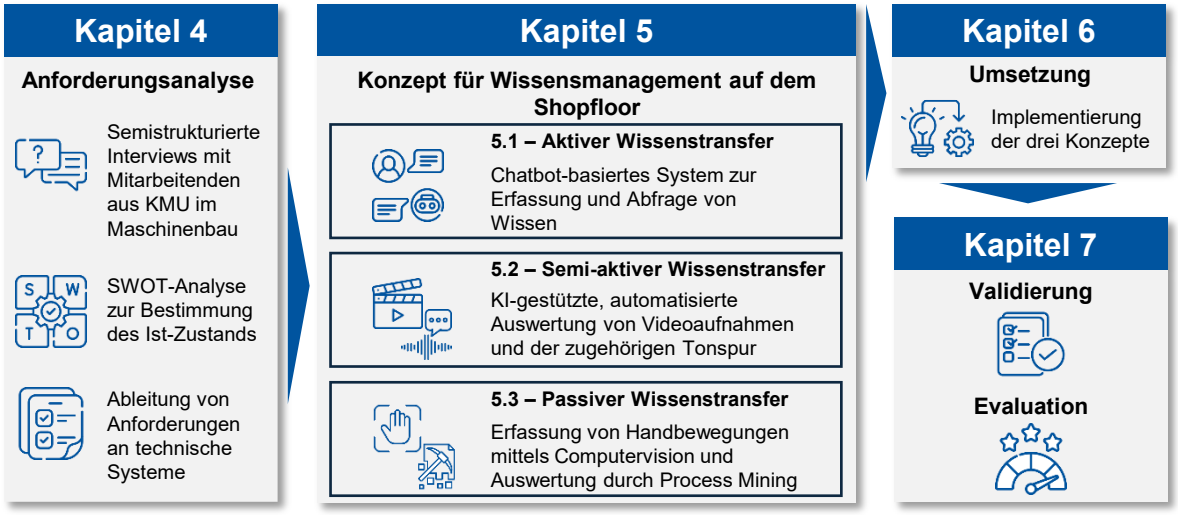



**Abbildung 3.2: Detaillierte Vorgehensweise**

*Detailed Approach*

Entsprechend der in Kapitel 3 definierten Zielsetzung werden in Kapitel 4 die spezifischen Anforderungen an das Wissensmanagement auf dem Shopfloor von KMU analysiert. Hierzu wird eine Studie vorgestellt, bei der Mitarbeitende hinsichtlich ihres aktuellen Vorgehens im Wissensmanagement, ihres Fortschritts in der Digitalisierung sowie ihrer generellen Vorstellungen und Anforderungen in diesem Bereich interviewt wurden. Im Zuge dieser semistrukturierten Interviews wird pro Unternehmen eine SWOT-Analyse zur Bestimmung des Ist-Zustandes durchgeführt. Ausgehend von den konsolidierten Ergebnissen dieser Studie werden Anforderungen an neu zu entwickelnde Methoden und Systeme abgeleitet, die im weiteren Verlauf der Arbeit als Referenz für die Forschungs- und Entwicklungsaktivitäten dienen.

Basierend auf der Anforderungsanalyse werden in Kapitel 5 drei Konzepte vorgestellt, welche unterschiedliche Ansätze zur Implementierung von Wissensmanagement auf dem Shopfloor beschreiben. Die drei Ansätze unterscheiden sich insbesondere durch die Art und den Umfang der Aktivitäten, die von den Mitarbeitenden im Prozess des Wissenstransfers geleistet werden müssen. Im ersten Unterkapitel wird ein Konzept für einen *aktiven* Wissenstransfer vorgestellt, das auf dem Prinzip eines Chatbots basiert. Dazu wird eine Systemarchitektur konzipiert, welche den Chatbot mit Datenspeichern für produktionsspezifische Informationsarten und Dokumententypen verknüpft und somit den Mitarbeitenden eine intuitive Möglichkeit zum Speichern und Abrufen von Wissen per natürlichsprachlichem Dialog ermöglicht. In den darauffolgenden beiden Unterkapiteln wird jeweils ein Konzept für einen *semi-aktiven* und einen *passiven* Wissenstransfer präsentiert. Das *semi-aktive* Konzept thematisiert die automatisierte Auswertung von Videoaufnahmen und der zugehörigen Tonspur mittels Methoden der Computer Vision und des NLP. Dabei werden die Videos entsprechend der im Bild und Ton enthaltenen Informationen derart sequenziert, dass einzelne Tätigkeiten separiert werden können. Das *passive* Konzept befasst sich mit den Möglichkeiten der auf Computer Vision basierten Erfassung der individuellen Handbewegungen von Mitarbeiten-

den in Kombination mit Methoden des Process Mining zur automatisierten Dokumentation und Bewertung bewährter Vorgehensweisen („Best Practices“) bei manuellen Produktionsprozessen.

Die prototypische Umsetzung der entwickelten Konzepte in Form von Referenzimplementierungen wird in Kapitel 6 beschrieben. Aufgrund der dynamischen technologischen Fortschritte im Bereich der Chatbots während der Durchführung dieser Arbeit werden für das Konzept des *aktiven* Wissenstransfers zwei unterschiedliche Implementierungen vorgestellt, die jeweils auf verschiedenen Technologien aus dem Bereich des NLP basieren.

Die Validierung der technischen Funktionalitäten der drei Referenzimplementierungen erfolgt zu Beginn von Kapitel 7. Anschließend werden die drei Konzepte sowie deren jeweilige Umsetzung hinsichtlich der zuvor in Kapitel 4 definierten Anforderungen systematisch evaluiert. Abschließend erfolgt eine vergleichende Analyse der Konzepte, bei der ihr jeweiliger Beitrag zur Beantwortung der in diesem Kapitel formulierten Teilforschungsfragen kritisch untersucht wird.

Die Arbeit schließt in Kapitel 8 mit einer Zusammenfassung der wesentlichen Ergebnisse und einem Ausblick auf mögliche weiterführende Untersuchungen im Zusammenhang mit den Forschungsfragen. Die in dieser Dissertation beschriebenen Entwicklungen und Untersuchungen entstanden im Kontext des öffentlich geförderten Projekts *AKzentE4.0*[1] unter Einbeziehung industrieller Anwendungsunternehmen. Durch einen intensiven Austausch, insbesondere innerhalb des im Projekt initiierten Arbeitskreises „Wissensmanagement“, konnten praxisnahe Anforderungen identifiziert, potenzielle Lösungsansätze kontinuierlich diskutiert und anhand realer Problemstellungen evaluiert werden.

[1] AKzentE4.0 – Arbeitswissenschaftliches Kompetenzzentrum für die Erwerbsarbeit in der Industrie 4.0. Dieses Forschungs- und Entwicklungsprojekt wird durch das Bundesministerium für Bildung und Forschung (BMBF) im Programm „Zukunft der Wertschöpfung – Forschung zu Produktion, Dienstleistung und Arbeit“ (Förderkennzeichen: 02L19C400) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei dem Autor.

# 4 Anforderungsanalyse

*Requirements Analysis*

In diesem Kapitel werden die Anforderungen von KMU an technische Systeme und Methoden zur Realisierung von Wissensmanagement auf dem Shopfloor erfasst und anschließend analysiert, um Richtlinien für die in Kapitel 5 erfolgende Konzeptentwicklung abzuleiten. Dazu wird zunächst eine Studie vorgestellt, in der KMU hinsichtlich ihres aktuellen Stands und ihrer Vorstellungen in Bezug auf Wissensmanagement befragt wurden. Anschließend werden die Ergebnisse der Studie systematisch ausgewertet und in funktionale sowie nicht-funktionale Anforderungen überführt. Abschließend erfolgt eine detaillierte Bewertung der identifizierten Anforderungen. Das Vorgehen und die Ergebnisse der Studie wurden bereits in [DASS23] diskutiert und werden im Folgenden ausführlicher vorgestellt und vertieft analysiert.

## 4.1 Ausrichtung, Methodik und Randbedingungen der Studie

*Orientation, Methodology and Boundary Conditions of the Study*

Als Grundlage für die Entwicklung eigener Konzepte für die Realisierung von Wissensmanagement auf dem Shopfloor ist zunächst die Schaffung eines umfassenden Überblicks der aktuellen Situation in deutschen KMU notwendig. Insbesondere die herrschenden Probleme bei der Realisierung von Wissensmanagement und die Wünsche der Mitarbeitenden, welche die Systeme später unmittelbar nutzen, sollen systematisch erfasst werden. Die der Studie vorangegangene Literaturrecherche ergab, dass bisherige Untersuchungen, wie z.B. [ORTH08], [HOLG15] und [MARC21], sich häufig auf Business Prozesse fokussieren und die besonderen Randbedingungen des Shopfloors vernachlässigen. Gleichzeitig wurde deutlich, dass der Einsatz von Informationstechnologien eine zentrale Rolle für effizientes Wissensmanagement spielen kann, indem durch diese das Erfassen, Speichern und Abrufen von Wissen strukturiert und beschleunigt wird [ABUB19; PETA20; SÁNC22]. Laut [ZIMM18] treiben KMU die Digitalisierung nicht erheblich voran und investieren weniger als größere Unternehmen. Als Gründe dafür werden u.a. fehlendes IT-Fachwissen, Probleme bei der Anpassung der unternehmensinternen Abläufe und Strukturen sowie schlechte Internetqualität genannt. Dennoch schätzen in einer Studie von 2016 etwa die Hälfte der befragten Unternehmen ihren Digitalisierungsgrad im Vergleich zu ihren Wettbewerbern als normal entwickelt ein [BLEY16]. Da diese Aussagen zum Teil konträr sind, wurde der Aspekt Digitalisierung dem Untersuchungsrahmen der Studie hinzugefügt.

Das Ziel dieser Studie besteht darin, die Funktionen der derzeit in den Unternehmen eingesetzten Wissensmanagementsysteme und die gewünschten Funktionen und Merkmale zu erfassen. Zur besseren Einordnung der Aussagen wird zudem der aktuelle Stand der digitalen Transformation in der Produktion dokumentiert. Vor diesem Hintergrund ergeben sich die folgenden vier Forschungsfragen (RQ) für die Studie:

- **RQ1**: Wie ist der Status Quo von KMU in Bezug auf Digitalisierung und Realisierung der Konzepte der Industrie 4.0?

Von besonderem Interesse sind dabei die aktuell genutzten Implementierungen zur digitalen Erfassung und Speicherung von Daten und Informationen, die zur Durchführung von Prozessen benötigt werden. Mit der zweiten Forschungsfrage sollen aktuell verwendete Systeme und Methoden identifiziert werden:

- **RQ2**: Wie ist der Status Quo von KMU in Bezug auf Wissensmanagement auf dem Shopfloor?

Die dritte Forschungsfrage knüpft inhaltlich unmittelbar an die zweite an und soll zur kritischen Auseinandersetzung mit der aktuellen Situation dienen:

- **RQ3**: Was sind die Stärken und Schwächen der derzeit für Wissensmanagement eingesetzten Systeme und Methoden?

Abschließend folgt die vierte und im Kontext der angestrebten Entwicklungen in dieser Arbeit wichtigste Frage:

- **RQ4**: Welche Anforderungen und Wünsche haben die KMU an neu zu entwickelnde Wissensmanagementsysteme und -methoden?

Die Analyse der mit dieser vierten Frage erfassten Gestaltungsvorstellungen der Unternehmen soll vor dem Hintergrund der Antworten auf die Fragen RQ1-3 eine fundierte Ausgangsbasis für die Entwicklung neuer, für den Shopfloor von KMU geeigneter Wissensmanagementsysteme liefern.

Zur Beantwortung der vorgestellten Forschungsfragen wurden leitfadengeführte, semistrukturierte Interviews mit Mitarbeitenden von KMU aus Nordrhein-Westfalen durchgeführt. Trotz ihres zeitintensiven Charakters ermöglicht diese qualitative Methodik die Erstellung eines detaillierten Überblicks der fokussierten Thematik und bietet gleichzeitig die Möglichkeit, bei relevanten Aspekten tiefgreifende Nachfragen zu stellen, um die individuellen, ehrlichen Meinungen der Mitarbeitenden zu erfassen [ADAM15]. Zur möglichst vollständigen Abbildung des Status Quo wurden sowohl Mitarbeitende des Shopfloors als auch Personen aus dem Management interviewt. Aufgrund der hohen Auftragslage bei gleichzeitigem Personalmangel konnten nicht in jedem Unternehmen jeweils Mitarbeitende des Shopfloors und des Managements für das Interview bereitgestellt werden. Grundsätzlich wurde pro Unternehmen nur ein Interview durchgeführt, sodass bis zu drei Personen gleichzeitig interviewt wurden. In Tabelle 4.1 sind die Teilnehmer der Studie anonymisiert mit Informationen zu ihrem jeweiligen Geschlecht, Alter und dem Bereich des Unternehmens, in dem sie tätig sind, dargestellt. Aufgrund zeitlicher Verfügbarkeit konnten vorrangig Personen aus dem Management von KMU für die Interviews gewonnen werden. Dies ist jedoch keine Einschränkung hinsichtlich der Fokussierung auf den Shopfloor, da die interviewten Personen aufgrund flacher Hierarchien sowie ihrer operativen Nähe zur Produktion und den Mitarbeitenden über umfassende Kenntnisse der Bedürfnisse und Herausforderungen verfügen. Insgesamt wurden sieben Interviews per Videokonferenz in deutscher Sprache durchgeführt. Die durchschnittliche Dauer der Interviews betrug 46 Minuten. Zur Auswertung wurden die Interviews mit Einverständnis der Teilnehmenden aufgezeichnet.

**Tabelle 4.1: Teilnehmende der Studie**

*Interview Respondants*

| Pseudonym | Geschlecht | Alter | Domäne | Nummer |
|---|---|---|---|---|
| 1_M_m_30 | Männlich | 30 | Management | 1 |
| 2_M_m_57 | Männlich | 57 | Management | 2 |
| 3_S_m_44 | Männlich | 44 | Shopfloor | 3 |
| 4_M_m_30 | Männlich | 30 | Management | 4 |
| 5_M_m_59 | Männlich | 59 | Management | 5 |
| 6_M_w_27 | Weiblich | 27 | Management | 6 |
| 6_S_m_57 | Männlich | 57 | Management | 6 |
| 7_M_m_53 | Männlich | 53 | Management | 7 |
| 7_S_m_56 | Männlich | 56 | Shopfloor | 7 |
| 7_S_m_x | Männlich | - | Shopfloor | 7 |

Die Fragen des zur Strukturierung der Interviews verwendeten Leitfadens sind in fünf thematische Sektionen gegliedert:

**Einleitung:** In der Einleitung werden die Teilnehmenden ausführlich über das Vorgehen und das Thema der Studie informiert. Anschließend wird ihr Einverständnis zur Aufzeichnung der Interviews eingeholt und dabei auf die Tatsache hingewiesen, dass die Teilnahme freiwillig erfolgt und sie jederzeit das Interview abbrechen können.

**Persönliche Informationen:** In diesem zweiten Abschnitt werden der Vollständigkeit halber Alter und Geschlecht der Teilnehmenden erfasst. Daraufhin wird der jeweilige schulische und berufliche Werdegang besprochen. Dabei werden insbesondere Ausbildungen und Erfahrungen mit Bezug zu Digitalisierung und Wissensmanagement vermerkt.

**Unternehmensspezifische Informationen:** Die Teilnehmenden werden gebeten, ihr derzeitiges Unternehmen kurz zu beschreiben. Dies umfasst die Unternehmensgröße (Anzahl der Mitarbeitenden und Standorte), die Kernkompetenz des Unternehmens sowie die technische Ausstattung (Maschinen, digitale Infrastruktur etc.). Zusätzlich wird auf die Art der Inhalte eingegangen, mittels derer Informationen aktuell gespeichert und transportiert werden (Texte, Zeichnungen, 3D-Modelle, Bilder, Videos etc.). Darüber hinaus werden die Teilnehmenden nach ihrer Wahrnehmung der Nutzung und Akzeptanz neuer Technologien im Unternehmen befragt. Abschließend wird der jeweilige Onboarding-Prozess der Befragten thematisiert, wobei besondere Aufmerksamkeit auf die Methoden und technischen Hilfsmittel beim Lernprozess gelegt wird.

**Wissensmanagement:** Zur Einleitung dieses Abschnittes werden die Teilnehmenden gebeten, ein Beispiel für besonders gutes oder schlechtes Wissensmanagement in ihrem Unternehmen zu nennen. Dieses Beispiel dient im weiteren Verlauf als Referenz. Anhand des Beispiels sollen die Befragten die Prozesse der Speicherung und Übertragung von Wissen im Unternehmen detailliert erläutern, indem sie verwendete

Methoden und Praktiken darstellen. Abschließend wird eine SWOT-Analyse durchgeführt, um potenzielle Stärken, Schwächen, Chancen und Risiken in Bezug auf ihr aktuelles Wissensmanagement zu identifizieren.

**Gestaltungsvorstellungen:** Im letzten Abschnitt sollen die Teilnehmenden ihre Vorstellungen und gewünschten Spezifikationen für ein optimales Wissensmanagementsystem festlegen. Zur Präzisierung der Fragestellung werden die Teilnehmenden dazu angehalten, anzugeben, welche Art von Unterstützung sie zur Implementierung eines solchen Systems benötigen würden, wie viel Zeit die Mitarbeitenden wöchentlich (in Stunden) für die Pflege des Systems aufwenden dürften und welche Schritte erforderlich wären, um ein optimales System zu realisieren. Dabei sollen die Befragten ihre Erwartungen möglichst detailliert beschreiben.

Zur Auswertung der Interviews wird die Methodik der qualitativen Datenanalyse angewandt, wie sie von MILES ET AL. in [MILE14] beschrieben wird. Dabei werden die Aufnahmen der Interviews zunächst manuell wörtlich transkribiert, um eine vollständige und genaue Datengrundlage zu erhalten. Die Aussagen in den Transkripten werden anschließend codiert, indem die enthaltenen Sätze oder Satzgruppen auf ihre Kernaspekte heruntergebrochen und mit einem Überbegriff (dem Code) versehen werden. Die so erstellten Codes werden im nächsten Schritt über alle Interviews hinweg gesammelt und zu Kategorien zusammengefasst. Abschließend werden die Kategorien analysiert, um Muster und Schlüsselaussagen zu erkennen und auf Basis dessen fundierte Schlussfolgerungen zu ziehen.

Die Anzahl der teilnehmenden Unternehmen wurde im Vorfeld nicht festgelegt, da die Bereitschaft zur Teilnahme an der zeitaufwendigen Befragung im Vorhinein nicht eingeschätzt werden konnte. Die letztlich befragte Anzahl von sieben Unternehmen wurde im Kontext dieser Arbeit als ausreichend für das weitere Vorgehen bewertet, da sich bereits in diesen Gesprächen ein hohes Maß an inhaltlicher Wiederholung und thematischer Übereinstimmung zeigte. Eine wesentliche Erweiterung oder Veränderung der Erkenntnisse durch weitere Interviews mit KMU ähnlicher Branche, Größe und Standortstruktur war nicht zu erwarten. Dieser Zustand wird in der Literatur als *theoretische Sättigung* bezeichnet und beschreibt den Punkt, an dem zusätzliche Datenerhebung keine neuen relevanten Informationen mehr hervorbringt [FLIC95]. Dies entspricht den Beobachtungen von GUEST ET AL., dass die theoretische Sättigung häufig bereits bei einer Interviewanzahl zwischen sechs und zwölf eintritt [GUES06].

## 4.2 Ergebnisse der Studie

*Results of the Study*

Im Folgenden werden die Ergebnisse der Studie vorgestellt, die auf der in Kapitel 4.1 beschriebenen Methodik zur Analyse der Interviews basieren. Dazu werden die gefundenen Kategorien und zugehörigen Codes für die Sektionen *Unternehmensspezifische Informationen*, *Wissensmanagement* und *Gestaltungsvorstellungen* präsentiert. Die Sektionen *Einleitung* und *Persönliche Informationen* ergaben keine für diese Arbeit relevanten Ergebnisse und werden deshalb vernachlässigt. Anhand repräsentativer

Zitate der interviewten Personen werden die Kategorien der drei relevanten Sektionen exemplarisch veranschaulicht und die Ergebnisse je Sektion diskutiert.

### 4.2.1 Unternehmensspezifische Informationen

*Company-Specific Information*

Die Aussagen der Teilnehmenden der Studie in Bezug auf die Fragen zur Sammlung der *Unternehmensspezifischen Informationen* lassen sich in vier Kategorien gliedern, welche inklusive der genutzten Codes in Abbildung 4.1 dargestellt sind.

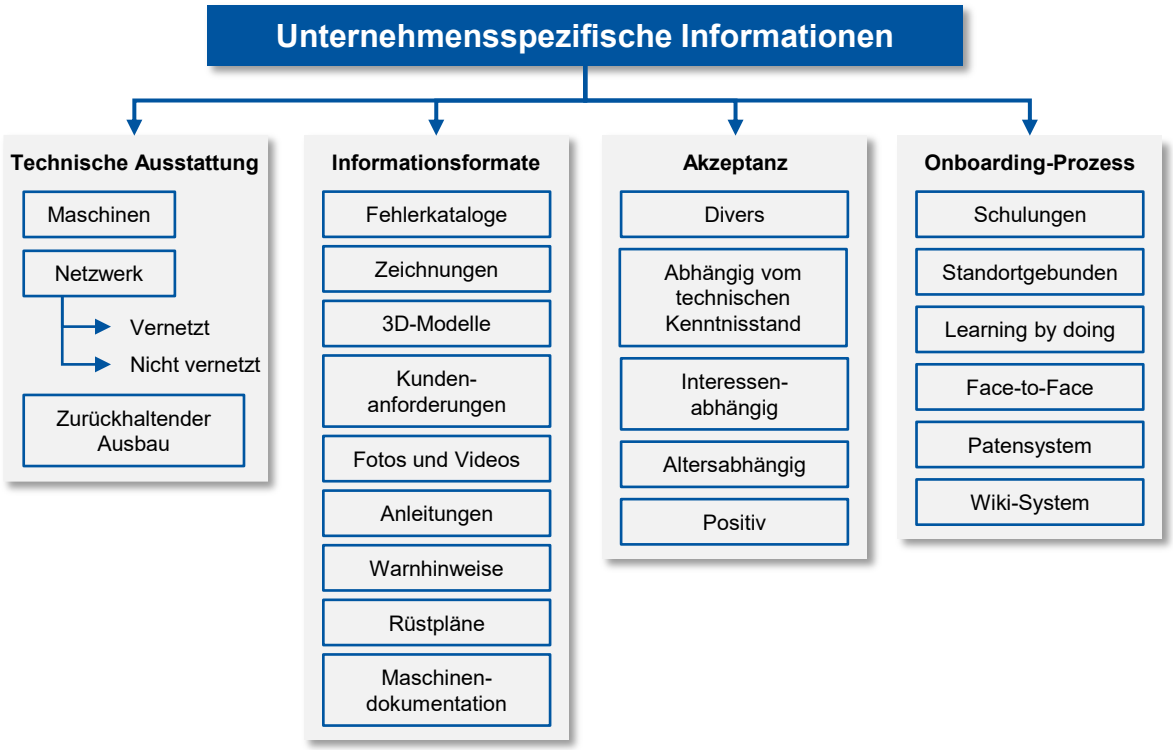



**Abbildung 4.1: Kategorien der Sektion *Unternehmensspezifische Informationen***

*Categories of the Company-specific Information Section*

Da die Fragen des Leitfadens in dieser Sektion direkt auf die technische Ausstattung, die vorhandenen Informationsformate, die Akzeptanz der Mitarbeitenden gegenüber neuen Technologien und den Ablauf der Onboarding-Prozesse zielen, können die gegebenen Antworten entsprechend unter diese Begriffe gruppiert werden. Grundsätzlich räumten die meisten Befragten ein, dass Wissensmanagement ein wichtiger Bestandteil der zukünftigen Entwicklung ihres Unternehmens ist. Sie nannten jedoch auch Herausforderungen, die es zu bewältigen gilt, um effizientes Wissensmanagement in den täglichen Betrieb zu integrieren. Angesprochen auf die technische Ausstattung im Sinne von Industrie 4.0 gaben einige der Befragten an, dass die vorhandenen Maschinen kaum vernetzt seien und somit keine automatische Datenerfassung erfolgt. Dieser Mangel an Vernetzung ist nicht immer auf technologische Beschränkungen zurückzuführen, sondern wird zum Teil auch aus Sicherheitsgründen bewusst beibehalten. Ein Befragter (1_M_m_30) sagte beispielsweise: „Nein, sie [die Maschinen] sind komplett getrennt, damit kein Zugriff oder keine Manipulation von außen

stattfinden kann.“ Andere Teilnehmende erklärten, dass sie in ihrem Unternehmen Daten und Programme der Maschinen speichern und nach Möglichkeit für zukünftige Projekte wiederverwenden.

Auf die Frage nach den Arten von Wissen, die generiert werden, nannten die Teilnehmenden eine Vielzahl von Informationsformaten, in denen Wissen dokumentiert ist. Neben klassischen, produktionstypischen Formaten, wie Zeichnungen, 3D-Modellen, Rüstpläne und Maschinendokumentationen wurden auch Videos, Fotos und selbstentwickelte Anleitungen für Prozesse genannt, mittels derer bereits versucht wird, Prozessabläufe und Fehlerbehandlungen anschaulich zu dokumentieren. Die für eine Tätigkeit relevanten Dokumente sind dabei oftmals über verschiedene digitale Datenspeicher verteilt, sodass das Finden aller relevanten Unterlagen viel Aufwand bedeutet.

Die Akzeptanz von neuen Technologien wird laut den Teilnehmenden der Studie in ihren Unternehmen durch Faktoren wie den individuellen technischen Kenntnisstand, das jeweilige Alter der Person und allgemeines Interesse an der Technologie beeinflusst. Auch der technologische Fortschritt im Unternehmen kann einen Einfluss auf das Interesse haben, wie eine Befragte (6_M_w_27) erklärte: „Es gibt ein paar Mitarbeiter, die sich besonders dafür [für neue Technologien] interessieren. Andere haben es schon aufgegeben, weil sie merken, dass sich nichts tut. Man hat ein breites Feld an unterschiedlichen Charakteren […].“ Grundsätzlich kann das Ausmaß an Akzeptanz in den befragten KMU als sehr divers bezeichnet werden, wobei zumindest der Eindruck der Befragten in Summe für ihr jeweiliges Unternehmen eher positiv war.

Der letzte Teil dieser Sektion des Interviews adressierte den Onboarding-Prozess und den dabei erfolgenden Transfer von Wissen auf neue Mitarbeitende. Die Teilnehmenden nannten verschiedene Methoden, darunter allgemeine Schulungen, die standortgebunden zu festen Zeiten durchgeführt werden, Learning by doing („Ausprobieren“) und Patensysteme, bei denen erfahrene Mitarbeitende die neuen KollegInnen („Face-to-Face“) als Mentor begleiten und sie anlernen. Einige Befragte nannten auch ein vorhandenes Wiki-System, in welchem Prozesse, Verantwortlichkeiten und Lerneinheiten abgelegt seien. Auf Nachfrage stellte sich jedoch heraus, dass diese Systeme aufgrund mangelnder Pflege in der Praxis nur begrenzt hilfreich sind.

In Bezug auf **RQ1** lässt sich zusammenfassen, dass die teilnehmenden Unternehmen der Studie grundsätzlich ein Interesse an Digitalisierung haben, insbesondere um von der papierbasierten Fertigung wegzukommen. Bisher werden die digitalen Dokumente jedoch überwiegend unstrukturiert abgelegt und es existiert keine einheitliche Datenbasis. Alle Befragten nutzen den Desktop-PC als einzige Hardware für den Wissenszugriff, während innovative Technologien wie AR bisher nicht zum Einsatz kommen. Die auf dem Shopfloor genutzten Technologien und Methoden für Wissensmanagement entsprechen im Wesentlichen denen im Büro. Die im Sinne von Industrie 4.0 angestrebte Vernetzung wird unterschiedlich gehandhabt. Während einige Unternehmen Daten aus ihren Maschinen nutzen, um bereits bekannte Probleme schneller zu lösen oder Programme wiederzuverwenden, gibt es bei anderen Bedenken hinsichtlich

der Sicherheit. Einige Unternehmen schließen ihre Maschinen gar nicht an ein Netzwerk an und verhindern so, dass die Daten für eine intelligentere Produktionssteuerung genutzt werden können.

### 4.2.2 Wissensmanagement

*Knowledge Management*

Für die Sektion *Wissensmanagement* der Interviews lassen sich die Antworten der Teilnehmenden in die zwei Kategorien „Aktuelle Methoden des Wissensmanagement“ und „Aktuelle Wissensspeicher“ einteilen (vergleiche Abbildung 4.2). Auf die Frage nach aktuell verwendeten Methoden des Wissensmanagements wurde von den Befragten am häufigsten der persönliche Wissenstransfer („Face-to-Face“) genannt (N=4). Ein Befragter (6_S_m_57) merkte dabei an, dass bei dieser Methode Wissen verloren geht und nie wieder aufzufinden ist, wenn ein Mitarbeiter das Unternehmen verlässt. Zur Verstetigung von Wissen im Unternehmen nutzen die Befragten deshalb Dokumentationsmethoden wie Protokolle, die oftmals auf digitalen Wissensspeichern wie internen Servern abgelegt werden (N=5). Bei manchen Teilnehmenden wird (zumindest teilweise) noch analog mit Stift und Papier dokumentiert (N=3). Positiv lässt sich dabei anmerken, dass eine generelle Bereitschaft zum Teilen von Wissen festgestellt werden konnte. Dies wird unter anderem durch eine Aussage von 3_S_m_44 deutlich: „Wissen wird verbal weitergegeben. Jeder teilt sein Wissen. Es gibt niemanden, der sagt: ‚Ich behalte mein Wissen lieber für mich selbst‘.“ Dedizierte digitale Wissensmanagementsysteme wurden von keinem Teilnehmenden benannt.

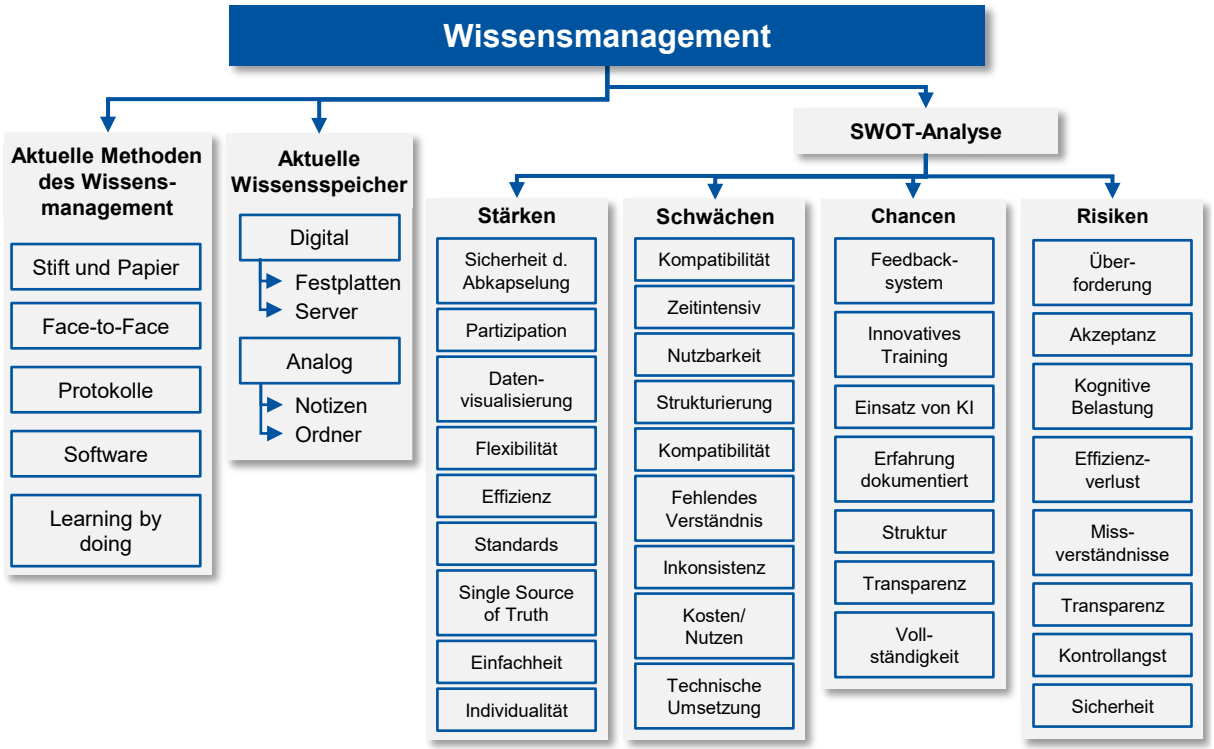



**Abbildung 4.2: Kategorien der Sektion *Wissensmanagement***

*Categories of the Knowledge Management Section*

Die Ergebnisse der SWOT-Analyse sind entsprechend dem Vorgehen der Methode in Stärken, Schwächen, Chancen und Risiken gegliedert und werden im Folgenden entlang dieser Kategorien vorgestellt.

**Stärken**

Generell berichteten die Befragten, dass ihrer Meinung nach Wissensmanagement durch die Etablierung einer zentralen Wissensquelle („Single Source of Truth“) in Kombination mit uneingeschränktem Wissenstransfer, intuitiver Datenvisualisierung und der Anwendung von standardisierten Prozessen zu mehr Effizienz, Produktivität und Flexibilität führen kann. So gab ein Befragter an, dass Aufträge schneller abgewickelt werden könnten, da das notwendige Wissen dann für die Mitarbeiter sofort und direkt zugänglich ist, ohne auf persönliche Kommunikation angewiesen zu sein. Ein gut eingeführtes Wissensmanagementsystem kann daher zu einer besseren Reproduzierbarkeit führen, was dem Unternehmen einen Wettbewerbsvorteil in Bezug auf Schnelligkeit und Anpassungsfähigkeit verschafft ("Man muss es nicht immer wieder machen. So ist man natürlich schneller und hat einen gewissen Wettbewerbsvorteil gegenüber anderen, die es zum ersten Mal machen müssen", 1_M_m_30). Als ein weiterer Vorteil von Wissensmanagementsystemen wurde die Möglichkeit genannt, strukturierter und zielgerichteter auf Probleme reagieren zu können: „Die Sammlung von Zahlen, Daten und Fakten gibt uns auch die Möglichkeit, strukturierter an mögliche Problemlösungen heranzugehen.“, 7_S_m_x. Als eine der wenigen genannten Stärken der aktuellen Methoden zum Wissensmanagement wurde angemerkt, dass man aufgrund der manuellen, analogen Vorgänge eine hohe Sicherheit gegenüber Industriespionage habe.

**Schwächen**

Als größten Schwachpunkt der aktuellen Methoden und Systeme für Wissensmanagement nannten die Befragten (N=6) den hohen Zeitaufwand, der in keinem ausgewogenen Verhältnis zum Nutzen steht. Dies wurde von einem Befragten (1_M_m_30) folgendermaßen erklärt: „Der Punkt ist, dass es einfach anstrengend ist, all dieses Wissen aufzuschreiben. Ich hatte einmal einen erfahrenen Kollegen, der seinen ganzen Tag damit verbrachte, alles aufzuschreiben, was er wusste. Das war ein Fulltime-Job für mehrere Jahre. Und dann begann er sich zu fragen, ob es das wirklich wert ist." Diese momentan herrschende Ineffizienz wurde von den Befragten mit mangelnden Strukturen und inkonsequenter Umsetzung begründet. Die vorhandenen heterogenen Maschinen, Programme und Abläufe ließen sich aktuell nicht vollständig zu einem System integrieren. Die Befragten gaben an, dass das meiste Wissen noch sehr stark an Mitarbeitende gebunden und nicht systematisch dokumentiert ist: „Bis heute ist das Wissen an die Menschen gebunden und sobald sie das Unternehmen verlassen, verlässt natürlich auch das Wissen das Unternehmen.", 1_M_m_30.

**Chancen**

Die Teilnehmenden gaben an, dass Wissensmanagement ihrer Meinung nach das Potenzial hat, die Transparenz innerhalb des Unternehmens signifikant zu erhöhen. Ein Befragter sagte, dass die Mitarbeitenden dadurch motiviert werden könnten, dass sie

Einblick in die vom Wissensmanagementsystem generierten Daten erhalten und Auffälligkeiten erkennen („Ich meine, die Leute sollten eigentlich dadurch motiviert werden, dass sie die Zahlen sehen und sich sagen: ‚Wir müssen etwas tun!' ", 7_M_m_53). Außerdem sehen die Befragten eine Chance für bessere Schulungsmethoden und Onboarding-Prozesse, da neue Mitarbeitende auf einen strukturierten, fundierten Wissens- und Erfahrungsschatz zurückgreifen könnten („Wie eine Schatzkiste voll Erfahrung!", 5_M_m_59).

**Risiken**

Das von den Befragten am häufigsten genannte Risiko war die Angst vor einer übermäßigen Digitalisierung, aus der eine technologische Abhängigkeit entstehen könnte. Die Möglichkeit, dass Systeme das Ziel von Cyberkriminalität werden oder fehlerhaft sein könnten und dadurch die Produktion zum Stillstand käme, würde die Akzeptanz deutlich schmälern. Außerdem fürchten die Teilnehmenden, dass die Teamarbeit durch Missverständnisse beeinträchtigt werden könnte, wenn ein Großteil des Wissens über digitale Kanale gespeichert und transferiert werden würde. Damit einher geht die Befürchtung, dass die Mitarbeitenden ihre Fähigkeit zum kritischen Denken verlieren könnten. Dies wird von einem Befragten (1_C_m_30) wie folgt formuliert: „Das Risiko besteht natürlich darin, dass man sich zu sehr auf das verlässt, was geschrieben steht. Das heißt, wenn es ein bestimmtes Programm gibt, dem man vertraut, aber vielleicht hat jemand einen Fehler gemacht oder irgendwie ist ein Fehler reingekommen, dass man sich darauf verlässt und sich nicht mehr hinterfragt, oder? Dass man sein Gehirn nicht mehr benutzt." Darüber hinaus sahen die Befragten ein Risiko darin, dass die Mitarbeitenden mit dem zusätzlichen Aufwand, der für die Pflege des Wissensmanagementsystems notwendig wäre, überfordert sein könnten. Sie erwähnten auch die große Menge an Wissen, die tagtäglich erzeugt werde und ins System eingebracht werden müsste. Dies wiederum könne zu einem Effizienzverlust führen, falls mehr Zeit für die Verwaltung von Wissen aufgewendet würde als für die eigentliche Schaffung von Wissen und die Herstellung von Produkten.

Auf Basis dieser Informationen können die Forschungsfragen **RQ2** und **RQ3** folgendermaßen beantwortet werden: Die Teilnehmer sind sich bewusst, wie schnell Wissen verloren geht und welche Probleme damit einhergehen. Die Ergebnisse der Studie zeigen aber auch deutlich, dass die derzeit verwendeten Wissensmanagementmethoden noch sehr traditionell sind. Vieles geschieht durch direkten Austausch zwischen den Mitarbeitenden. Learning-by-doing, Mentoring und persönliche Gespräche sind immer noch weit verbreitet. Ob Wissen digital gespeichert wird, hängt vor allem von der Verfügbarkeit der notwendigen Soft- und Hardware sowie dem Engagement der Mitarbeitenden ab.

Die Teilnehmenden der Studie hatten Probleme, die Bereiche Stärken und Chancen der SWOT-Analyse zu differenzieren, sodass sie häufig bereits bei den Stärken potenziell positive Auswirkungen von strukturiertem Wissensmanagement im Unternehmen beschrieben und sich weniger auf die Stärken des aktuellen Vorgehens bezogen. Als

einzig konkrete Stärke der aktuellen Methoden wurde die Sicherheit vor Industriespionage genannt. Bei den Schwächen wurden insbesondere der aktuell hohe Zeitaufwand und die Ineffizienz beim Dokumentieren und Finden von Wissen hervorgehoben. Hinzu kommt die Tatsache, dass viel Wissen momentan noch ausschließlich in den Köpfen der Mitarbeitenden existiert und somit nicht jederzeit von allen Personen abgerufen werden kann.

### 4.2.3 Gestaltungsvorstellungen

*Design Ideas*

In der abschließenden Sektion der Interviews wurden die Teilnehmenden gebeten, ihre Gestaltungsvorstellungen und Wünsche an zukünftige Wissensmanagementsysteme zu beschreiben. Zusätzlich sollten sie erklären, welche Unterstützung sie für die Implementierung im Unternehmen benötigen würden. Entsprechend gliedern sich die Antworten der Befragten in die beiden Kategorien „Notwendige Unterstützung“ und „Erwartungen“, wie in Abbildung 4.3 dargestellt.

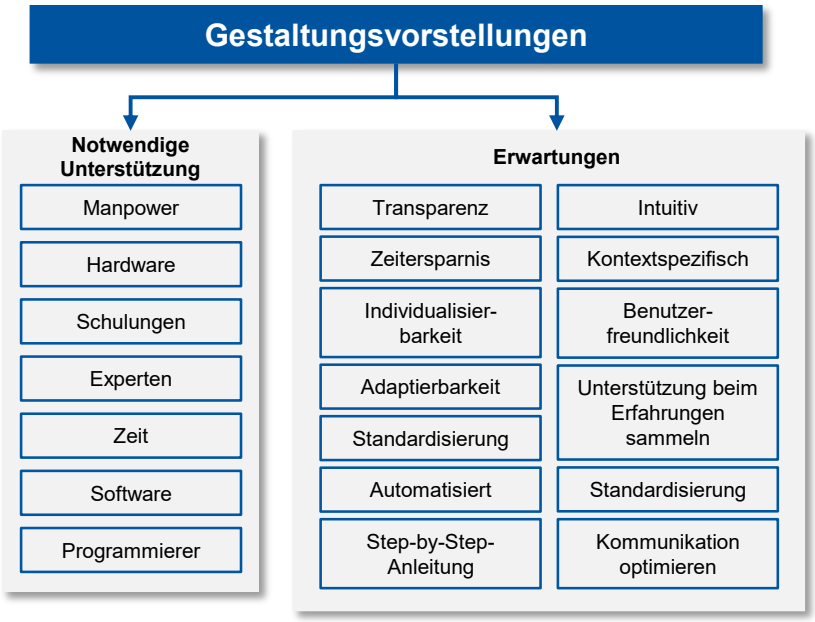



**Abbildung 4.3: Kategorien der Sektion *Gestaltungsvorstellungen***

*Categories of the Design Ideas*

In Bezug auf die benötigte Unterstützung nannten die Befragten im Wesentlichen drei Hauptpunkte: Erstens benötigen sie Zeit und Personal zur Einrichtung und Wartung der Systeme. Zweitens stellten die Befragten fest, dass die erforderliche Hard- und Software in den meisten KMU fehlt und von entsprechendem Fachpersonal eingerichtet werden müsse, welches aktuell in den Unternehmen ebenfalls nicht zu finden sei. Drittens bedarf es zielgerichteter Schulungen, um die Mitarbeitenden an die Arbeit mit den neuen Systemen heranzuführen. Dabei wurde mehrfach erwähnt, dass ein Experte für Wissensmanagement im Unternehmen hilfreich wäre.

Wie in Abbildung 4.3 zu sehen ist, äußerten die Teilnehmenden eine Vielzahl von Erwartungen, die sich jedoch grundsätzlich in jedem der KMU ähnelten: Zeitersparnisse

beim Suchen von Wissen, automatisierte Dokumentation von Abläufen und die Generierung von Schritt-für-Schritt-Anleitungen („Step-by-Step“) wurden dabei am häufigsten für die Prozesse auf dem Shopfloor genannt. Die Vorstellungen der Teilnehmenden begrenzten sich jedoch nicht nur auf den Shopfloor, sondern bezogen sich häufig auf das gesamte Unternehmen. Das System müsse transparent und individuell erweiterbar sein, um die spezifischen Anforderungen der verschiedenen Abteilungen (z.B. Controlling, Fertigung, Finanzen, Personalwesen) abzudecken. Nur so könne die notwendige Akzeptanz der Mitarbeitenden zur Nutzung erreicht werden. Dabei sei es wichtig, dass die Systeme von allen Mitarbeitenden unabhängig von ihrer Vorerfahrung und Ausbildung intuitiv genutzt werden können.

Entsprechend dieser Erwartungen kann zur Beantwortung der Forschungsfrage **RQ4** zusammengefasst werden, dass die Teilnehmenden sich grundsätzlich ein vollumfassendes Wissensmanagement für alle Bereiche des Unternehmens wünschen, welches intuitiv, adaptierbar und transparent sein soll. Für das Wissensmanagement auf dem Shopfloor wünschen sich die Befragten vor allem Verbesserungen in Bezug auf die Zeitersparnisse beim Suchen von Wissen, die automatisierte Dokumentation von Abläufen und die Generierung von Schritt-für-Schritt-Anleitungen.

## 4.3 Anforderungen an Wissensmanagement auf dem Shopfloor

*Requirements for Knowledge Management on the Shopfloor*

Die zuvor dargestellten Ergebnisse der Studie werden nachfolgend ausgewertet und in funktionale sowie nicht-funktionale Anforderungen an Wissensmanagement auf dem Shopfloor überführt. Aufgrund der begrenzten Stichprobengröße der Studie erfolgt zur Gewährleistung der Relevanz der definierten Anforderungen anschließend ein Abgleich mit den in der VDI 5610 formulierten Zielen und Nutzen von Wissensmanagement (siehe Anhang Teil C).

### 4.3.1 Ableitung von Anforderungen auf Basis der Studienergebnisse

*Derivation of Requirements Based on the Study Results*

Im ersten Schritt werden die Kernaussagen der Teilnehmenden der Studie in den drei Sektionen *Unternehmensspezifische Informationen*, *Wissensmanagement* und *Gestaltungsvorstellungen* zusammengefasst und anschließend in spezifische Anforderungen überführt. In Tabelle 4.2 ist dieses Vorgehen exemplarisch für die Sektion *Unternehmensspezifische Informationen* dargestellt. Die Ergebnisse der beiden übrigen Sektionen sind im Anhang in Tabelle B.5 und Tabelle B.6 zu finden.

Bei genauer Betrachtung der abgeleiteten Anforderungen lässt sich eine gewisse Redundanz über die Sektionen hinweg feststellen. Aufgrund dessen werden die Anforderungen zur besseren Übersichtlichkeit auf die wesentlichen Aspekte reduziert und inhaltlich kategorisiert. Die resultierenden Kategorien sind in Tabelle 4.3 aufgeführt. Zur Nachvollziehbarkeit des Ursprungs der Anforderungen ist jeweils die laufende Nummer aus den Tabellen 4.2, B.5 und B.6 angegeben.

**Tabelle 4.2: Anforderungen aus der Sektion *Unternehmensspezifische Informationen***

*Requirements from the Company-Specific Information Section*

| Sektion | Kernaussage | Anforderung | Nr. |
|---|---|---|---|
| **Unternehmensspezifische Informationen** | Die vorhandenen Maschinen sind zum Teil aufgrund von Sicherheitsbedenken nicht vernetzt. | Gewährleistung von **Sicherheit gegenüber Cyberkriminalität** und damit Schutz vor Industriespionage und Sabotage. | 1 |
| | Daten und Programme von Maschinen werden gespeichert und für zukünftige Aufträge wiederverwendet. | Gewährleistung von Möglichkeiten zur **Wiederverwendbarkeit von vorhandenem Wissen** zur Vermeidung von Redundanz. | 2 |
| | Wissen wird in vielfältigen Formaten (teilweise noch analog) dokumentiert und in verschiedenen Datenspeichern angelegt, sodass das Finden des notwendigen Wissens sehr zeitaufwendig ist. | Realisierung eines **zentralen, digitalen Datenspeichers** für alle relevanten Informationsformate und Möglichkeit zur **kontextabhängigen, schnellen Suche von Wissen**. | 3 |
| | Die Akzeptanz von neuen Methoden und Systemen hängt von individuellen Faktoren (u.a. Alter, Vorkenntnisse, Interesse) ab. | Realisierung von Möglichkeiten zur **Individualisierbarkeit der Benutzerschnittstellen** des Systems. | 4 |
| | Onboarding-Prozesse werden analog durch den direkten Austausch von Wissen zwischen Mitarbeitenden vorgenommen. Die Qualität des Wissenstransfers ist personenabhängig und damit örtlich und zeitlich eingeschränkt. | Unterstützung der Onboarding-Prozesse durch eine **zentrale Wissensquelle**, welche **orts- und zeitunabhängig** von Mitarbeitenden genutzt werden kann. | 5 |

Zur weiteren Strukturierung werden die identifizierten Anforderungen nach ROBERTSON in *funktional* und *nicht-funktional* klassifiziert. Eine *funktionale* Anforderung beschreibt eine Aktion, die das Produkt ausführen muss, damit es für einen Benutzer nützlich ist. Diese Aktion resultiert aus der Arbeit, die die Nutzer erledigen müssen. Fast jede Aktion (meist durch ein aktives Verb ausgedrückt) kann dabei Grundlage für eine funktionale Anforderung sein. *Nicht-funktionale* Anforderungen sind Eigenschaften oder Qualitäten, die das Produkt haben muss, damit es von seinen Nutzern akzeptiert wird. In einigen Fällen sind nicht-funktionale Anforderungen entscheidend für den Erfolg des Produktes, da sie eine Mindestanforderung zur Benutzung darstellen. Eigenschaften wie Leistung, Aussehen, Handhabung, Benutzerfreundlichkeit und Sicherheit fallen in diese Klasse von Anforderungen. [ROBE13]

Das Ergebnis der auf Basis dieser Definitionen durchgeführten Klassifizierung ist in Tabelle 4.3 in der Spalte K dargestellt. Die *funktionalen* Anforderungen sind durch „**f**“ und die *nicht-funktionalen* Anforderungen durch „**nf**“ markiert. Grundsätzlich ist die Unterscheidung zwischen den beiden Klassen nicht immer eindeutig möglich, da *nicht-funktionale* Anforderungen meistens eine *funktionale* Anforderung näher erläutern. Wurde diese zugehörige *funktionale* Anforderung jedoch vorher nicht explizit genannt,

steckt sie oftmals implizit in der *nicht-funktionalen* Anforderung. Dies ist zum Beispiel in Tabelle 4.3 in der Kategorie „User Experience“ bei der Anforderung „Intuitive Datenvisualisierung“ der Fall. Entsprechend der oben genannten Definition kann diese Anforderung als *nicht-funktional* klassifiziert werden, da sie mit dem Adjektiv „intuitiv“ eine Eigenschaft beschreibt. Sie setzt jedoch voraus, dass eine Datenvisualisierung überhaupt besteht, welche nicht explizit als eigene funktionale Anforderung genannt wird. Dies trifft auf alle Anforderungen der Kategorie „User Experience“ zu. Da die dort vorausgesetzten Funktionen unverzichtbare Bestandteile eines technischen Wissensmanagementsystems sind, ist die für diese Arbeit relevante Information in der Beschreibung der *nicht-funktionalen* Anforderung enthalten, sodass diese Anforderungen mit „nf“ klassifiziert werden.

**Tabelle 4.3: Kategorisierung der abgeleiteten Anforderungen und Klassifizierung in funktional und nicht-funktional**

*Categorization of the Derived Requirements and Classification in Functional and Non-functional*

| Kategorie | Anforderungen | Nr. | K |
|---|---|---|---|
| **Allgemeine Anforderungen** | Wiederverwendbarkeit von Wissen | 2 | f |
| | Transfer von individuellem Wissen in Unternehmenswissen | 6 | f |
| | Möglichkeit zur Validierung des gespeicherten Wissens | 12 | f |
| **Sicherheit** | Sicherheit gegenüber Cyberkriminalität | 1 | nf |
| **User Experience** | Individualisierbarkeit der Benutzerschnittstellen | 4 | nf |
| | Intuitive Datenvisualisierung | 8 | nf |
| | Intuitive Implementierung und Nutzung | 14 | nf |
| | Intuitive Benutzerschnittstellen | 18 | nf |
| | Offenes, erweiterbares, transparentes System | 19 | nf |
| **Wissens-dokumentation** | Intuitive, prozessparallele, digitale Dokumentation | 7 | f |
| | Reduktion des Zeitaufwands bei der Dokumentation | 9 | nf |
| | Systematische Dokumentation | 10 | f |
| | Strukturierte Dokumentation | 11 | f |
| | Effizientes Wissensmanagement mit geringem Aufwand | 13 | nf |
| | Automatisierte Dokumentation von Prozessabläufen und Best Practices | 16 | f |
| | Effiziente Erstellung von Schritt-für-Schritt-Anleitungen | 17 | f |
| **Wissens-speicher** | Zentraler, digitaler Datenspeicher für alle relevanten Informationsformate | 3 | f |
| | Zentrale Wissensquelle, zentrale Speicherung | 8, 11 | f |
| **Wissensabruf** | Kontextabhängige, schnelle Suche von Wissen | 3 | f |
| | Orts- und zeitunabhängige Nutzung der Wissensquelle | 5 | f |
| | Uneingeschränkter Zugriff | 8 | nf |
| | Schnelle Suche und Identifikation von relevantem Wissen | 15 | nf |

In der Kategorie „Wissensspeicher“ bei der Anforderung „Zentraler, digitaler Datenspeicher für alle relevanten Informationsformate“ erfolgt die Auslegung der Uneindeutigkeit der Klassifizierung in die andere Richtung. In diesem Fall wird das Vorhandensein eines zentralen, digitalen Datenspeichers als *funktionale* Anforderung an das System interpretiert und damit als „f“ markiert. Auf diese Weise werden alle Anforderungen vor dem Hintergrund der Fragestellung dieser Arbeit klassifiziert.

### 4.3.2 Validierung der identifizierten Anforderungen

*Validation of the Identified Requirements*

Zur Validierung der identifizierten Anforderungen an Wissensmanagementsysteme für den Shopfloor sowie zur Überprüfung ihrer Gültigkeit über die befragten Unternehmen hinaus, wird ein Abgleich mit den in der VDI 5610 (siehe Anhang C) definierten Zielen von Wissensmanagement durchgeführt. Dabei wird überprüft, ob jede der Anforderungen mindestens einem Ziel zugeordnet werden kann. Das Ergebnis dieses Vergleichs ist in Abbildung 4.4 in Form einer Matrix dargestellt, in der die Anforderungen gemäß Tabelle 4.3 auf der X-Achse und die in der VDI 5610 definierten Ziele auf der Y-Achse aufgetragen sind. Aufgrund des begrenzten Platzes sind nur Spalten für jene Ziele vorhanden, bei denen eine Entsprechung gefunden wurde.

| | VDI 5610 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Perspektive des Unternehmens | | | | | | | | | Perspektive der Mitarbeitenden | |
| | Kosten einsparen durch… | | | Produkt- und Qualitätsverbesserungen durch… | | | | | | | |
| **Anforderungen** | Zeiteinsparung bei der Suche | Vermeidung von Doppelarbeit | Reduzierung von Wissensverlust | Transparente Organisation | Transparente Prozesse | Standardisierte Dokumentation | Standardisierte IT | Systematische Aufbereitung von Wissen | Zugriff auf aktuelle Informationen | Steigerung der Zufriedenheit | Leichtere Einarbeitung |
| Wiederverwendbarkeit von Wissen | | X | | | | | | | | | |
| Transfer von individuellem Wissen in Unternehmenswissen | | | X | | | | | | | | |
| Möglichkeit zur Validierung des gespeicherten Wissens | | | | | | | | X | | | |
| Sicherheit gegenüber Cyberkriminalität | | | | | | | | | | | |
| Individualisierbarkeit der Benutzerschnittstellen | | | | | | | | | | X | |
| Intuitive Datenvisualisierung | | | | | | | | | | X | |
| Intuitive Implementierung und Nutzung | | | | | | | | | | X | |
| Intuitive Benutzerschnittstellen | | | | | | | | | | X | |
| Offenes, erweiterbares, transparentes System | | | | X | X | | | | | | |
| Intuitive, prozessparallele, digitale Dokumentation | | | | | | X | X | | | | |
| Reduktion des Zeitaufwands bei der Dokumentation | | | | | | X | X | | | | |
| Systematische Dokumentation | | | | | | X | X | | | | |
| Strukturierte Dokumentation | | | | | | X | X | | | | |
| Effizientes Wissensmanagement mit geringem Aufwand | | | | | | X | X | | | | |
| Automatisierte Dokumentation von Prozessabläufen und Best Practices | | | | | | X | X | | | | X |
| Effiziente Erstellung von Schritt-für-Schritt-Anleitungen | | | | | | X | | | | | X |
| Zentraler, digitaler Datenspeicher für alle relevanten Informationsformate | | | | | | | X | | | | |
| Zentrale Wissensquelle, zentrale Speicherung | | | | | | | X | | | | |
| Kontextabhängige, schnelle Suche von Wissen | X | | | | | | | | X | | |
| Orts- und zeitunabhängige Nutzung der Wissensquelle | X | | | | | | | | | | |
| Uneingeschränkter Zugriff | X | | | | | | | | | | |
| Schnelle Suche und Identifikation von relevantem Wissen | X | | | | | | | | | | |

**Abbildung 4.4: Vergleich der identifizierten Anforderungen mit den in der VDI 5610 definierten Zielen von Wissensmanagement**

*Comparison of the Identified Requirements with the Objectives of Knowledge Management Defined in VDI 5610*

Nahezu jede der abgeleiteten Anforderungen lässt sich mindestens einem Ziel des Wissensmanagements gemäß der VDI 5610 zuordnen, wodurch sichergestellt ist, dass die Anforderungen eine ausreichende Relevanz für die weitere Untersuchung

bieten. Lediglich für die Anforderung hinsichtlich der *Sicherheit* von Wissensmanagementsystemen gegenüber Cyberkriminalität kann keine passende Einordnung vorgenommen werden. Da sich diese Anforderung jedoch allgemein auf IT-Systeme bezieht, wird sie im weiteren Verlauf dieser Arbeit nicht explizit betrachtet.

## 4.4 Zusammenfassung und Bewertung der Anforderungsanalyse

*Summary and Evaluation of the Requirements Analysis*

In den vorangegangenen Unterkapiteln wird eine Studie vorgestellt, mittels derer Anforderungen von Mitarbeitenden in KMU an Wissensmanagement(-systeme) für den Shopfloor erhoben wurden. Die so identifizierten Anforderungen werden in funktional und nicht-funktional kategorisiert (vgl. Tabelle 4.3) und anschließend zur Validierung mit den in der VDI 5610 beschriebenen Zielen von Wissensmanagement verglichen (vgl. Abbildung 4.4). Durch dieses Vorgehen wird die Menge der erfassten Anforderungen strukturiert, reduziert und auf ausreichende Relevanz für die weitere Untersuchung überprüft. Grundsätzlich sind die dargestellten Anforderungen im Kontext von Wissensmanagement sinnvoll, verständlich und konsistent, sodass sie sich alle für die weitere Untersuchung eignen. Da der Schwerpunkt dieser Arbeit auf der prototypengestützten Erforschung technischer Wissensmanagementsysteme für den Shopfloor von KMU liegt, dienen insbesondere die identifizierten funktionalen Anforderungen als Richtlinien für die in den folgenden Kapiteln dargestellte Entwicklung, Implementierung und Validierung innovativer Konzepte.

Insbesondere die in der Kategorie User Experience enthaltenen Anforderungen in Bezug auf Intuitivität, Individualisierbarkeit und Erweiterbarkeit werden als nicht-funktionale Anforderungen erst relevant, sobald die notwendigen, grundlegenden Funktionen vorhanden sind. Außerdem lässt sich die User Experience eines Systems über die Gestaltung des Frontends problemlos nach der Implementierung der zugrundeliegenden Funktionalitäten einstellen. Deshalb werden die Anforderungen der Kategorie User Experience in dieser Arbeit nur indirekt berücksichtigt. Der Fokus liegt zunächst auf der Exploration neuartiger Ansätze und der Untersuchung deren Realisierbarkeit mit aktuell verfügbaren Technologien. Die in Tabelle 4.3 in den weiteren Kategorien aufgeführten nicht-funktionalen Anforderungen beziehen sich auf grundlegende Funktionen, die bei der Implementierung im Backend realisiert werden. Aufgrund ihrer Relevanz werden sie, neben den funktionalen Anforderungen, als primäre Anforderungen betrachtet. Insgesamt können die identifizierten Anforderungen zur Spezifizierung des jeweiligen Untersuchungsraums der in Kapitel 3.1 vorgestellten Teilforschungsfragen genutzt werden. In Abbildung 4.5 ist die Zuordnung der Anforderungen zu den Teilforschungsfragen dargestellt.

Dabei lässt sich erkennen, dass die *Allgemeinen Anforderungen* hinsichtlich Wiederverwendbarkeit und Validierungsmöglichkeiten von Wissen die dritte Teilforschungsfrage adressieren, während der Transfer von individuellem Wissen in Unternehmenswissen sowohl für die erste als auch für die zweite Teilforschungsfrage von Bedeutung ist. Die Anforderungen an die *User Experience* sind für jede Interaktion zwischen Mit-

arbeitenden und Wissensmanagementsystem relevant und detaillieren damit den Untersuchungsraum aller drei Forschungsfragen. Die Kategorie *Wissensdokumentation* bezieht sich sowohl auf die erste als auch auf die zweite Forschungsfrage. Die Anforderungen an *Wissensspeicher* und *Wissensabruf* korrespondieren mit der dritten Forschungsfrage. Entsprechend dieser Zuordnung werden die Anforderungen im weiteren Verlauf der vorliegenden Arbeit für die Untersuchung der Teilforschungsfragen genutzt.

Da die Analyse des Stands der Technik insbesondere ein Defizit bei den technischen Möglichkeiten zur strukturierten Erfassung von implizitem Wissen, der Automatisierung von Wissensmanagementprozessen und der intuitiven Einbettung in die betrieblichen Abläufe aufgedeckt hat, werden in der vorliegenden Arbeit die Hauptkategorien *Wissensdokumentation, Wissensspeicher* und *Wissensabruf* besonders fokussiert. Diese drei Kategorien repräsentieren gemeinsam den grundlegenden Prozessablauf des Wissensmanagements: Erfassung, Speicherung und Bereitstellung. Die Anforderungen aus den Kategorien *Allgemeine Anforderungen* und *User Experience* lassen sich vollständig auf diesen Prozessablauf beziehen und werden deshalb nicht separat, sondern im Kontext der drei Hauptkategorien betrachtet.

| | | Teilforschungsfragen | | |
|---|---|---|---|---|
| **Kategorie** | **Anforderungen** | 1. Wie können Mitarbeitende ihr Wissen an bestimmten Punkten ihrer Tätigkeit intuitiv und mit möglichst minimalem Mehraufwand aktiv digital speichern bzw. weitergeben? | 2. Wie kann Wissen in Form von „Best Practices“ passiv erfasst bzw. „getrackt“ werden, ohne die Mitarbeitenden bei der Ausführung ihrer Tätigkeiten zu behindern? | 3. Wie kann eine kontextualisierte, strukturierte Speicherung und bedarfs-gerechte Bereitstellung des Wissens realisiert werden? |
| **Allgemeine Anforderungen** | Wiederverwendbarkeit von Wissen | | | X |
| | Transfer von individuellem Wissen in Unternehmenswissen | X | X | |
| | Möglichkeit zur Validierung des gespeicherten Wissens | | | X |
| **User Experience** | Individualisierbarkeit der Benutzerschnittstellen | X | | X |
| | Intuitive Datenvisualisierung | X | | X |
| | Intuitive Implementierung und Nutzung | X | X | X |
| | Intuitive Benutzerschnittstellen | X | X | X |
| | Offenes, erweiterbares, transparentes System | X | X | X |
| **Wissens-dokumentation** | Intuitive, prozessparallele, digitale Dokumentation | X | X | |
| | Reduktion des Zeitaufwands bei der Dokumentation | X | X | |
| | Systematische Dokumentation | X | X | |
| | Strukturierte Dokumentation | X | X | |
| | Effizientes Wissensmanagement mit geringem Aufwand | X | X | X |
| | Automatisierte Dokumentation von Prozessabläufen und Best Practices | | X | |
| | Effiziente Erstellung von Schritt-für-Schritt-Anleitungen | X | X | |
| **Wissens-speicher** | Zentraler, digitaler Datenspeicher für alle relevanten Informationsformate | | | X |
| | Zentrale Wissensquelle, zentrale Speicherung | | | X |
| **Wissensabruf** | Kontextabhängige, schnelle Suche von Wissen | | | X |
| | Orts- und zeitunabhängige Nutzung der Wissensquelle | | | X |
| | Uneingeschränkter Zugriff | | | X |
| | Schnelle Suche und Identifikation von relevantem Wissen | | | X |

**Abbildung 4.5: Zuordnung der identifizierten Anforderungen an Wissensmanagementsysteme zu den Teilforschungsfragen**

*Assignment of the Identified Requirements for Knowledge Management Systems to the Sub-Research Questions*

# 5 Konzept für Wissensmanagement auf dem Shopfloor

*Concept for Knowledge Management on the Shopfloor*

Aufbauend auf der Analyse des Stands der Technik (siehe Kapitel 2) sowie der vorangegangenen Anforderungsanalyse (siehe Kapitel 4) liegt der Fokus im weiteren Verlauf dieser Dissertation auf der Entwicklung technologiebasierter Ansätze zur Erfassung, Speicherung und zum Abruf von implizitem Wissen bei der Durchführung von Tätigkeiten auf dem Shopfloor von produzierenden KMU. Für die Konzeptentwicklung wird angenommen, dass in der betrachteten Zielgruppe von Unternehmen bislang weder dedizierte Wissensmanagementsysteme noch produktionsspezifische, arbeitsbegleitende IT-Lösungen zum Einsatz kommen. Wissen wird überwiegend in digitaler Form über text- und bildbasierte Dokumentation gespeichert und weitergegeben. Ziel ist es daher, vollständig eigenständige Systeme zu konzipieren, die eine arbeitsnahe und intuitive Erfassung sowie den Abruf von Wissen ermöglichen, ohne auf bestehende IT-Systeme angewiesen zu sein. Eine optionale Anbindung an vorhandene Systeme wird vorgesehen, ist jedoch nicht Voraussetzung für den Betrieb. Im Vordergrund steht die Untersuchung der technischen Umsetzbarkeit. Wirtschaftliche, arbeitsorganisatorische und rechtliche Rahmenbedingungen werden nicht vertieft behandelt, jedoch bei der Konzeptentwicklung in grundlegender Form berücksichtigt, um realitätsferne oder praktisch nicht umsetzbare Lösungen zu vermeiden.

Grundlage für die in diesem Kapitel erfolgende Entwicklung von Konzepten für Wissensmanagement auf dem Shopfloor ist die bereits in [DASS21] vorgestellte Idee einer *prozessparallelen Erfassung von Erfahrungswissen*. Kern dieser Idee ist die Vision, dass Mitarbeitende auf dem Shopfloor jederzeit in der Lage sein sollen, die bei der Ausführung ihrer Tätigkeiten gewonnenen Erfahrungen systematisch, digital und mit möglichst wenig Aufwand zu speichern. Das auf diese Weise erfasste Wissen soll anschließend orts- und zeitunabhängig von anderen Mitarbeitenden abrufbar sein, um den Erfahrungsschatz auch dann zugänglich zu machen, wenn die ursprünglichen Wissensgeber nicht vor Ort sind oder keine Zeit haben. Im Gegensatz zu einer nachträglichen Dokumentation, die erst nach Abschluss der Tätigkeit erfolgt, bietet die prozessparallele Erfassung den Vorteil, dass die gemachten Erfahrungen noch „frisch“ und unverfälscht sind. Besonders in Fällen, in denen Mitarbeitende ihre Erfahrungen nur sporadisch nachträglich dokumentieren, besteht die Gefahr, dass wichtige Sachverhalte und Details in Vergessenheit geraten oder verfälscht werden. Gerade diese Details können jedoch für die Rekapitulation und Abstraktion eines auftretenden Problems oder einer Lösungsstrategie von entscheidender Bedeutung sein.

Zur Veranschaulichung dieser grundlegenden Idee ist in Abbildung 5.1 eine Systemarchitektur zur Realisierung der prozessparallelen Erfassung von Erfahrungswissen dargestellt, die aus drei Schichten besteht. Die unterste Schicht, der sogenannte *Process Layer*, umfasst die Prozesse und Arbeitsabläufe, denen die Mitarbeitenden auf dem

Shopfloor nachgehen. Darüber liegt der *Interface Layer*, welcher die Interaktion zwischen Menschen und dem Wissensmanagementsystem im darüber liegenden *Storage Layer* ermöglicht. Der Interaktionsprozess zwischen Mitarbeitenden und Wissensmanagementsystem verläuft dabei wie folgt: Stoßen Mitarbeitende während ihrer Arbeit auf ein Problem, können sie das Auftreten dieses Problems und gegebenenfalls die von ihnen entwickelte Lösung unmittelbar nach dem Auftreten mithilfe eines geeigneten Gerätes an das Wissensmanagementsystem übermitteln. In diesem System wird die Erfahrung erfasst, gespeichert und mit weiteren relevanten Informationen verknüpft und somit kontextualisiert. Tritt ein vergleichbares Problem erneut bei anderen Mitarbeitenden auf, können diese das hinterlegte Erfahrungswissen abrufen und zur effizienteren Lösung des Problems nutzen.

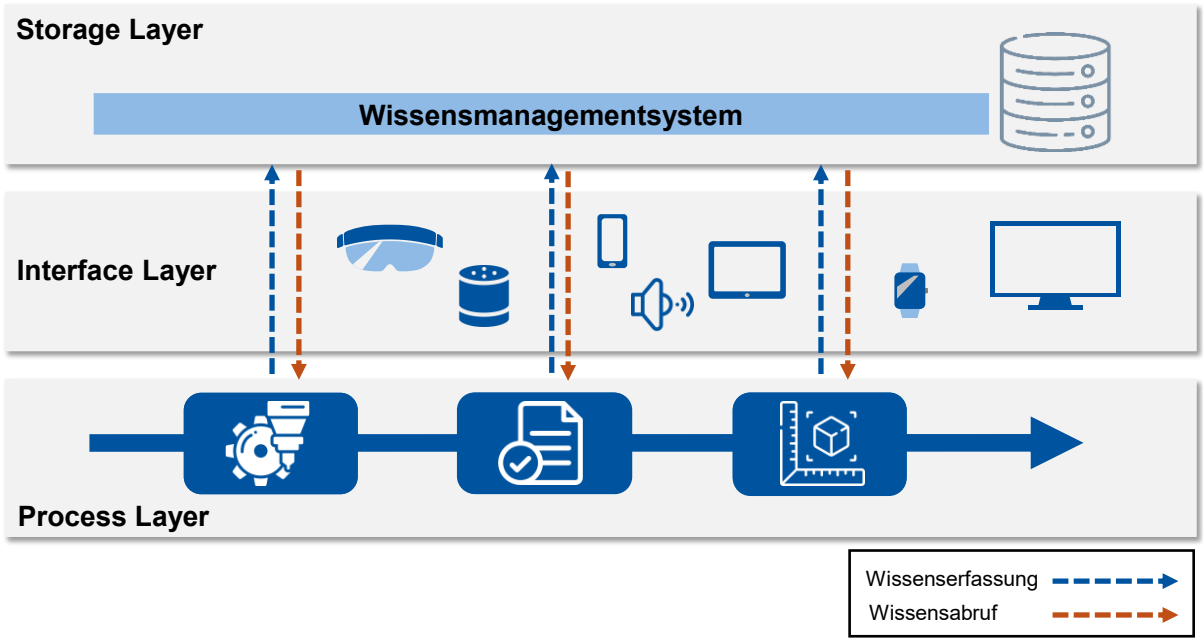



**Abbildung 5.1: Systemarchitektur zur prozessparallelen Erfassung von Erfahrungswissen [DASS21]**

*System Architecture for Process-Parallel Recording of Empirical Knowledge [DASS21]*

Etablierte Architekturen der Produktionstechnik, wie die Automatisierungspyramide [MEUD17] und das Referenzarchitekturmodell Industrie 4.0 [PLAT18], bilden primär (informations-)technische Hierarchien und Strukturen in Unternehmen ab. Diese basieren üblicherweise auf der Modellierung technisch eindeutig identifizierbarer und übertragbarer Signale, beispielsweise von Sensoren, Steuergeräten und Maschinen. Die Systemarchitektur zur prozessparallelen Erfassung von Erfahrungswissen fokussiert jedoch die Erfassung und Nutzung von Wissen als Produkt menschlicher Tätigkeiten. Da ein unmittelbarer, kontinuierlicher Datenfluss aus dem menschlichen Gehirn nach heutigem Stand der Technik nicht praktikabel realisierbar ist, erfolgt die Erfassung von Wissen über die Beobachtung von Handlungen sowie über verbale und nonverbale Kommunikation zwischen Mensch und Wissensmanagementsystem.

Die vorgestellte Systemarchitektur basiert auf der Annahme, dass sämtliche Prozesse im Process Layer in klar definierbare Blöcke einzelner Tätigkeiten untergliedert werden

können. Dies gilt im produktionstechnischen Kontext insbesondere für Prozesse, die nach standardisierten Abläufen erfolgen, wie die Montage von Bauteilen oder das Einrichten einer Maschine. In solchen strukturierten Prozessen ist den Mitarbeitenden stets bewusst, in welchem Schritt bzw. Tätigkeitsblock sie sich befinden. Diese exakte Bestimmung des Schritts innerhalb einer Prozesskette ist sowohl für die Erfassung als auch für den Abruf von Wissen von entscheidender Bedeutung. Durch die präzise Zuordnung von Wissen zu spezifischen Tätigkeiten kann eine bedarfsgerechte Bereitstellung von Informationen ermöglicht werden. Wie bereits im Stand der Technik dargelegt (siehe Kapitel 2.1.3), ist die Verwendung einer einheitlichen Fachsprache hierfür unerlässlich. Sollten Mitarbeitende unter denselben Begrifflichkeiten unterschiedliche Tätigkeiten, Objekte oder Werkzeuge verstehen oder damit assoziieren, führt dies unweigerlich zu Komplikationen und Missverständnissen, wodurch die Effizienz des Wissenstransfers und die operative Zusammenarbeit erheblich beeinträchtigt wird.

Zur Gewährleistung eines möglichst intuitiven Wissenstransfers müssen geeignete Geräte für den Interface Layer ausgewählt werden. Die konkreten Anforderungen an diese Geräte ergeben sich aus den spezifischen Umgebungsbedingungen der Arbeitsumgebung, wie Beleuchtung, Lärmbelastung und Verschmutzungsgrad sowie aus den durch die Ausführung der Tätigkeit bedingten Einschränkungen. Die manuelle Montage einer Baugruppe erfordert typischerweise den Einsatz beider Hände, wodurch eine prozessparallele Weitergabe von Erfahrungswissen per Tastatureingaben ausgeschlossen ist. Zwar bietet sich in solchen Fällen die Nutzung eines VUI an, doch können laute Umgebungsgeräusche die Effektivität dieser Technologie erheblich beeinträchtigen. Darüber hinaus sollten die individuellen Präferenzen der Mitarbeitenden in Bezug auf die Benutzererfahrung berücksichtigt werden, um mangelnder Akzeptanz vorzubeugen. Einige Menschen empfinden Unbehagen bei der Verwendung von z.B. AR-Brillen oder anderen HMDs und sollten daher die Möglichkeit haben, alternative Geräte zu nutzen, um zusätzliche Belastungen am Arbeitsplatz zu vermeiden.

Um die angestrebte Effizienzsteigerung durch den prozessparallelen Wissenstransfer zu realisieren, muss die Speicherung von Erfahrungswissen im Storage Layer systematisch und kontextbezogen erfolgen. Tritt dasselbe Problem unter identischen Bedingungen auf, ist eine präzise Zuordnung der erfassten Erfahrungen zum spezifischen Problem ausreichend. Komplexer gestaltet sich der Wissensabruf, wenn ein Problem nicht unter exakt gleichen Bedingungen auftritt. Es ist beispielsweise möglich, dass ein ähnliches Problem an unterschiedlichen Maschinen vorkommt. Wenn die gespeicherten Erfahrungen im Wissensmanagementsystem des Storage Layers ausschließlich mit Maschine A verknüpft sind, haben Mitarbeitende, die nach Informationen zu Fehlern an Maschine B suchen, keinen Zugriff auf die Lösung. Dies erfordert eine gewisse Flexibilität beim Abruf des Wissens, um verwandte Problemlösungen zu identifizieren.

Zur Entwicklung innovativer Ansätze für Wissensmanagement auf dem Shopfloor von KMU werden basierend auf den in Kapitel 4 analysierten Anforderungen und der

grundlegenden Idee der *prozessparallelen Erfassung von Erfahrungswissen* im Folgenden drei Konzepte vorgestellt, welche sich insbesondere in der Art und dem Umfang der Aktivitäten unterscheiden, die von den Mitarbeitenden im Prozess des Wissenstransfers zu leisten sind. Alle drei Konzepte lassen sich in die in Abbildung 5.1 gezeigte Systemarchitektur einordnen, bedienen sich jedoch unterschiedlicher Technologien und Methoden in den drei Layern zur Erfassung, Speicherung und Bereitstellung von Wissen (siehe Abbildung 5.2).

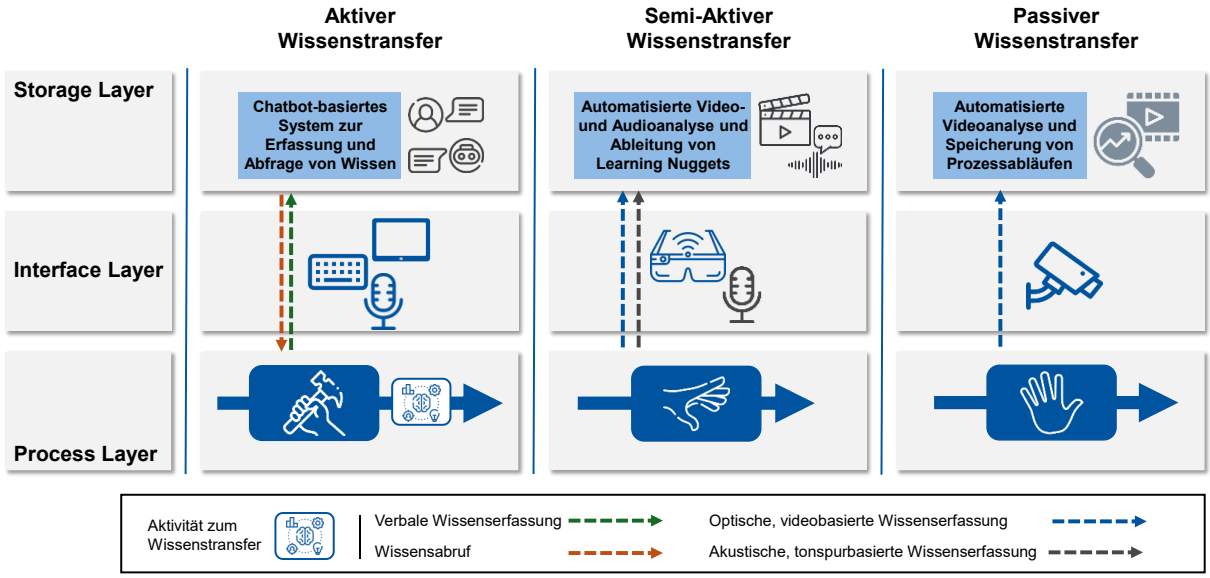



**Abbildung 5.2: Vergleich der drei Konzepte für Wissensmanagement auf dem Shopfloor**

*Comparison of the Three Concepts for Knowledge Management on the Shopfloor*

## 5.1 Aktiver Wissenstransfer

*Active Knowledge Transfer*

Das in dieser Dissertation entwickelte Konzept des *aktiven Wissenstransfers* leitet sich unmittelbar aus dem Grundgedanken der *prozessparallelen Erfassung von Erfahrungswissen* ab und bedient sich damit der Annahme, dass sich alle Tätigkeiten im Process Layer in einzelne Aktivitäten segmentieren lassen. Mitarbeitende sollen demnach eine technische Möglichkeit erhalten, ihre Erfahrungen systematisch nach Abschluss einer Aktivität und vor Beginn der nächsten zu dokumentieren. Der *aktive* Ansatz des Wissenstransfers sieht dabei vor, dass die Mitarbeitenden selbst dafür verantwortlich sind, ihre neugewonnenen Erkenntnisse zu formulieren und diese mithilfe eines geeigneten Interfaces entweder verbal oder schriftlich digital zu speichern. Dieser Ansatz kann in Form der in Abbildung 5.3 visualisierten Systemarchitektur verdeutlicht werden.

Im Process Layer haben Mitarbeitende die Möglichkeit, zusätzliche Aktivitäten zum aktiven Wissenstransfer zwischen den wertschöpfungsrelevanten Tätigkeiten einzufügen. Dabei nutzen sie im Interface Layer geeignete Eingabegeräte, wie Tastaturen und Mikrofone, um ihr Wissen an das Wissensmanagementsystem im Storage Layer weiterzugeben. Die Interaktion mit dem Wissensmanagementsystem erfolgt über eine

Chatbot-Architektur, welche die Eingaben der Mitarbeitenden verarbeitet und diese entsprechend in geeignete digitale Wissensspeicher einordnet und abspeichert. Neben der Erfassung der Erfahrungen der Mitarbeitenden kann analog auch ein Wissensabruf erfolgen: Die Mitarbeitenden stellen dabei Anfragen an das Chatbot-basierte System und erhalten relevante Informationen, die im Dialog mit dem Chatbot individuell aufbereitet und bei Bedarf ergänzt werden können.

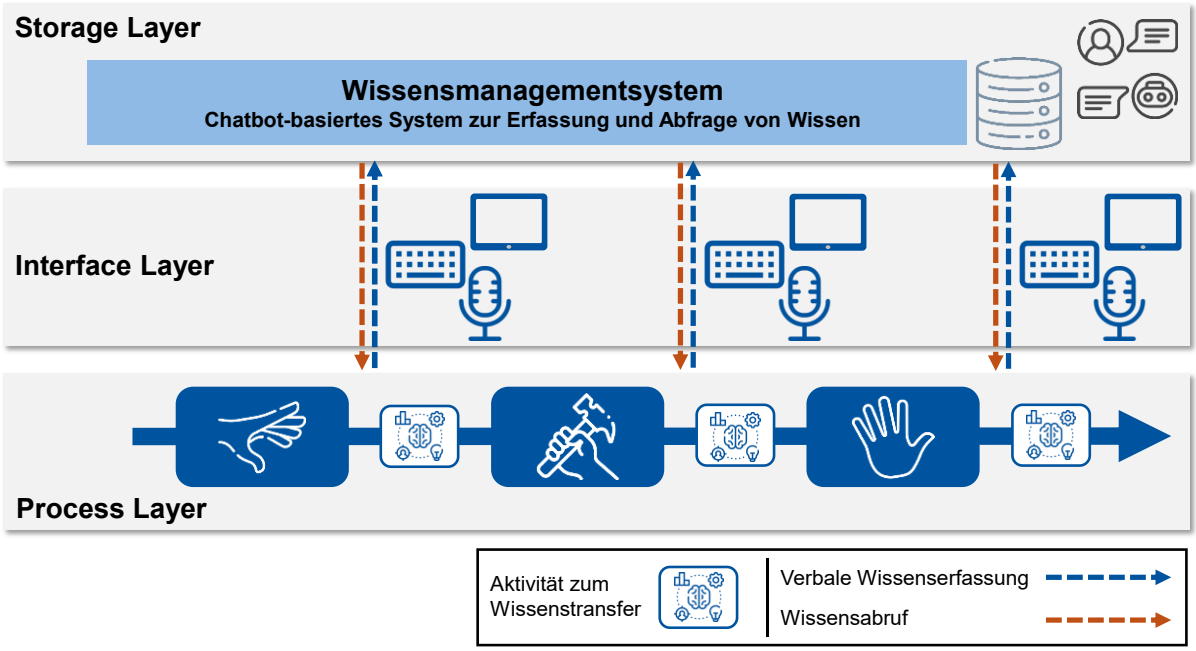



**Abbildung 5.3: Systemarchitektur des aktiven Wissenstransfers**

*System Architecture of the Active Knowledge Transfer*

Die Realisierung des zentralen Chatbots kann durch die Integration verschiedener aktueller Technologien aus den Bereichen der KI und des NLP erfolgen. Die konkrete Umsetzung hängt maßgeblich von der Art und dem Umfang der zu erfassenden oder abzurufenden Informationen sowie von den verwendeten Technologien zur Datenspeicherung ab. Im Rahmen dieser Arbeit liegt der Fokus auf textbasierten Informationen, wobei das Konzept modular gestaltet wird, sodass das Erfassen und Abrufen multimodaler Informationen zu einem späteren Zeitpunkt integriert werden kann. Da der Abruf und die Darstellung von Informationen bereits umfassend durch den Stand der Technik abgedeckt sind, liegt der Schwerpunkt auf der Erfassung und kontextualisierten Speicherung von Erfahrungswissen zur Ergänzung der bestehenden digitalen Wissensbasis.

In den folgenden Unterkapiteln wird das Konzept zur Realisierung des Chatbot-basierten *aktiven Wissenstransfers* anwendungsorientiert und detaillierter ausgearbeitet. Dabei werden insbesondere die für KMU relevanten textbasierten Informationsträger wie Arbeitsanweisungen, Anleitungen, Fehlermeldungslisten sowie Zustandsdaten von Produktionssystemen und -maschinen im Kontext der Prozesse der Wissenserfassung und des Wissensabrufs berücksichtigt.

### 5.1.1 Chatbot-basiertes System für eine effiziente und intuitive Erfassung von Wissen auf dem Shopfloor

*Chatbot-based System for Efficient and Intuitive Knowledge Capturing on the Shop Floor*

Aktuell für den Shopfloor verfügbare Wissensmanagementsysteme, wie beispielsweise Wikis oder DMS, setzen üblicherweise auf Formulare oder Eingabemasken, die den Mitarbeitenden eine starre Struktur für die Erfassung von Wissen vorgeben (siehe Kapitel 2.3.1). Im Gegensatz dazu erfordert der angestrebte Chatbot aufgrund seiner semantischen Analysefähigkeiten keine Einhaltung vordefinierter Strukturen, weshalb für den aktiven Wissenstransfer eine grundlegende sprachliche Verständlichkeit der Eingaben ausreichend ist. Diese innovative Flexibilität bei der Eingabe verbalisierter Informationen stellt eine Weiterentwicklung im Vergleich zum Stand der Technik dar und ermöglicht es, die im Process Layer notwendige zusätzliche Aktivität des Wissenstransfers möglichst intuitiv und effizient zu gestalten. Gleichzeitig werden dadurch potenzielle Hemmschwellen der Mitarbeitenden reduziert und die Bereitschaft gefördert, eigenes Wissen zu dokumentieren. Das Ziel besteht darin, dass die Mitarbeitenden ihr Wissen in einer Weise verbalisieren können, die dem direkten Austausch mit einer neuen Person im Unternehmen während einer Einarbeitung ähnelt. Der Chatbot soll darüber hinaus in der Lage sein, sowohl bei der Wissensdokumentation als auch beim Wissensabruf eine gewisse Abstraktion zu leisten. Dadurch wird sichergestellt, dass Informationen kontextualisiert gespeichert und entsprechend abrufbar gemacht werden können. Das aus diesen Überlegungen abgeleitete Konzept für die systematische Erfassung von Wissen per Chatbot ist in Abbildung 5.4 dargestellt. Die einzelnen Komponenten des Konzepts werden nachfolgend ausführlich beschrieben.

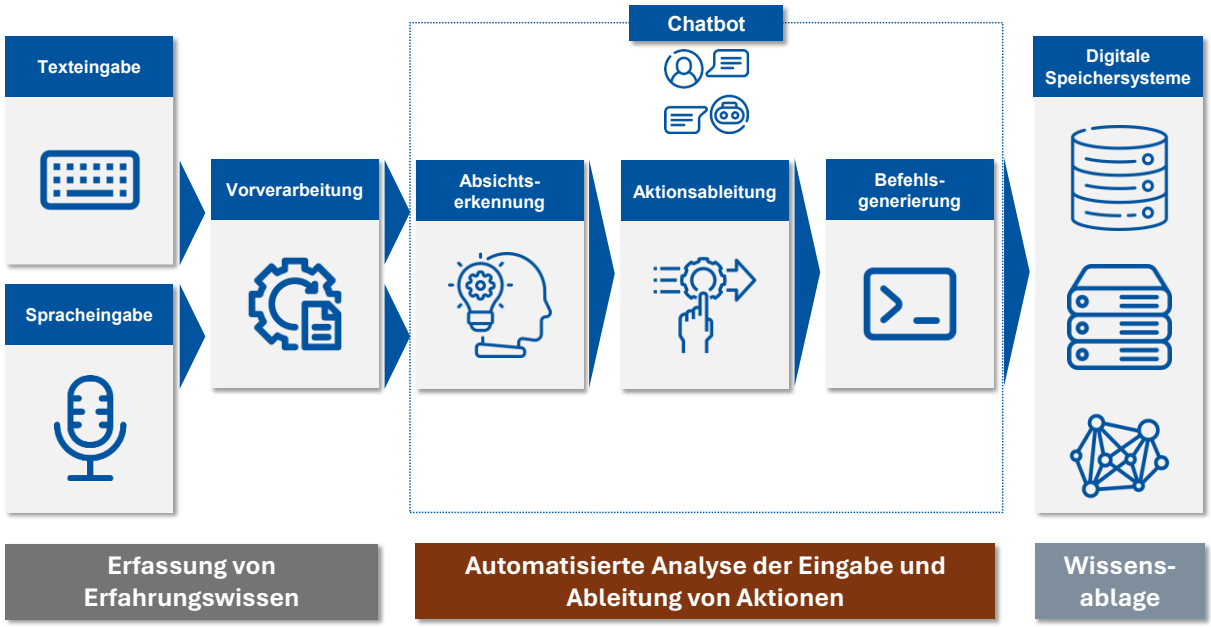



**Abbildung 5.4: Konzept der Chatbot-basierten Erfassung und Speicherung von Wissen**

*Concept of Chatbot-Based Capture and Storage of Knowledge*

### Erfassung von Erfahrungswissen – *Text- und Spracheingabe*

Entsprechend des allgemeinen Ansatzes für den aktiven Wissenstransfer sollen Mitarbeitende die Möglichkeit erhalten, ihre gemachten Erfahrungen unmittelbar nach Abschluss einer Aktivität zu formulieren und an ein digitales Speichersystem zu übermitteln. Dies kann in der Praxis entweder schriftlich als Texteingabe über eine Tastatur oder verbal durch eine Spracheingabe mittels eines geeigneten Mikrofons erfolgen. Im Rahmen dieser Arbeit liegt der Schwerpunkt auf der schriftlichen Eingabe per Tastatur, da an den Arbeitsplätzen der im Projekt AKzentE4.0 beteiligten KMU in vielen Fällen bereits Tastaturen und Zugänge zu Computersystemen vorhanden sind. Der Fokus auf die schriftliche Eingabe ermöglicht damit die Nutzung bestehender Geräte, wodurch eine spätere Implementierung in der Praxis erheblich erleichtert wird. Gleichzeitig stellt die Tastatur ein Interface dar, das den Mitarbeitenden sowohl aus ihrer beruflichen Tätigkeit als auch aus ihrer Freizeit vertraut ist, was Vorteile hinsichtlich einer späteren Akzeptanz des Systems bietet.

Die Eingabe per Tastatur erfolgt in einem Freitextfeld, das keinerlei Vorgaben bezüglich der Struktur oder Form der eingegebenen Texte enthält. Der Chatverlauf wird oberhalb des Eingabefeldes in einem Chatfenster dargestellt. Dieses Konzept für das User-Interface orientiert sich bewusst an den aus dem Freizeitbereich bekannten Messenger-Apps wie beispielsweise WhatsApp, um durch ein vertrautes Bedienungskonzept die Einstiegshürden zu senken und die intuitive Nutzung zu fördern. Die Mitarbeitenden haben durch das Freitextfeld die Möglichkeit, ihre Erfahrungen in ihrem individuellen Stil und Sprachgebrauch zu formulieren. Im Rahmen der Validierung wird untersucht, inwieweit der Chatbot in der Lage ist, die durch diese Gestaltungsfreiheit möglicherweise entstehende Ungenauigkeit und Uneinheitlichkeit der eingegebenen Informationen zu kompensieren und dennoch eine strukturierte sowie kontextualisierte Speicherung des Wissens zu gewährleisten.

Da die Spracheingabe im Konzept des semi-aktiven Wissenstransfers (siehe Kapitel 5.2) eine größere Bedeutung hat als im Konzept des aktiven Wissenstransfers, werden die Erfassung von Wissen mittels Tonspur sowie der Ablauf der für die weitere Verarbeitung erforderlichen Transkription in den Kapiteln 5.2 und 5.2.1 ausführlich erläutert. Das dort beschriebene Vorgehen kann auch für den aktiven Wissenstransfer genutzt werden.

#### *Vorverarbeitung*

Neben der Transkription können weitere Vorverarbeitungsschritte erforderlich sein, um den eingegebenen oder transkribierten Text für die nachfolgende Analyse durch den Chatbot aufzubereiten. Dazu zählt die Bereinigung der Textdaten, bei der Rechtschreibfehler korrigiert und die Groß- und Kleinschreibung vereinheitlicht werden, um die anschließende Verarbeitung zu erleichtern. Zusätzlich können je nach praktischen Anforderungen weitere Schritte wie die Sprachidentifikation und Übersetzung notwendig sein, um mehrsprachige Eingaben zu ermöglichen. Dies trägt zur Überwindung von Sprachbarrieren bei der Verbalisierung von Erfahrungswissen bei und erweitert die

Nutzbarkeit des Systems erheblich. Ergänzend zu diesen Vorgängen zur Textaufbereitung kann die Vorverarbeitung eine Benutzer- und Rechteverwaltung umfassen, die durch Authentifizierung den aktuellen Benutzer identifiziert und dessen Zugriffsberechtigungen überprüft. Dies ist insbesondere bei sensiblen oder kritischen Prozessinformationen essenziell, um den Zugang zu geschützten Daten und Funktionen angemessen zu regulieren.

In der vorliegenden Arbeit liegt der Fokus auf der Untersuchung der Chatbot-basierten automatisierten Analyse der Eingabe, weshalb die Betrachtung der Vorverarbeitung auf ein Minimum reduziert wird. Aufgrund der Modularität des Konzepts können jedoch bei Bedarf zusätzliche Vorverarbeitungskomponenten problemlos integriert werden.

**Automatisierte Analyse der Eingaben und Ableitung von Aktionen**

Die zentrale Logik des Systems für den aktiven Wissenstransfer wird durch einen Chatbot umgesetzt, der die Eingaben der Mitarbeitenden analysiert, darauf basierend entsprechende Aktionen ableitet und diese durch geeignete Befehle an die digitalen Speichersysteme weitergibt. Die Kernfunktionen des Chatbots – *Absichtserkennung*, *Aktionsableitung* und *Befehlsgenerierung* – können mithilfe unterschiedlicher technologischer Ansätze umgesetzt werden. In dieser Dissertation wurden zwei Ansätze untersucht, die sich hinsichtlich der Implementierungsweise der Kernfunktionen und des Grades der Generalisierung unterscheiden. Zunächst wurde das Konzept mittels klassischer NLP-Methoden in Kombination mit Machine-Learning-Ansätzen unter Verwendung des Rasa-Frameworks (siehe Kapitel 2.2.1) entwickelt und umgesetzt. Aufgrund der rasanten technologischen Entwicklung im Bereich der Foundation Models während des Bearbeitungszeitraums der Dissertation (2021-2024) (siehe Kapitel 2.2.1) wurde das Konzept zusätzlich auf die Verwendung eines LLM als Kern des Chatbots angepasst. Beide Ansätze werden in den folgenden Abschnitten detailliert vorgestellt. Dabei werden ihre Unterschiede sowie die jeweiligen Vor- und Nachteile für den aktiven Wissenstransfer benannt.

***Absichtserkennung, Aktionsableitung und Befehlsgenerierung – Rasa***

Im ersten Schritt führt der Chatbot eine semantische Analyse der vorverarbeiteten Texte durch, um die Intention der Eingabe zu erkennen. Bei der Gestaltung des Chatbots unter Verwendung des Rasa-Frameworks (siehe Kapitel 2.2.3) wird dazu die Funktion der *Intents* verwendet. Bezogen auf den aktiven Wissenstransfer in KMU könnte beispielsweise die Benutzereingabe „Lege bitte einen neuen Lösungsansatz zur Behebung von Fehler Nummer XB103 an Werkzeugmaschine 7 an“ dem Intent *create_solution_approach_for_error_at_MT* zugeordnet werden. Zur Erfassung weiterer notwendiger Informationen, die für die Ausführung eines Intents relevant sind, werden zudem spezifische *Entities* aus der Benutzereingabe extrahiert. Im zuvor genannten Beispiel der Eingabe eines neuen Lösungsvorschlags für einen Maschinenfehler könnten „Werkzeugmaschine 7“ als Entity der Kategorie *machineTool* und „Fehler Nummer XB103“ als Entity der Kategorie *errorNumber* identifiziert werden. Die Extraktion von Entities kann dabei durch alle drei in Kapitel 2.2.3 vorgestellten Methoden des Rasa-Frameworks erfolgen:

- **Synonyme**: Begriffe wie „WZM“, „Maschine“, „MachineTool“ oder „MT“ können als Synonyme für die Entity *machineTool* definiert werden, sodass unterschiedliche Benennungen derselben Bedeutung korrekt zugeordnet werden.
- **Regular Expressions**: Strukturen wie „XB103“ können mittels regulärer Ausdrücke identifiziert werden, indem entsprechend spezifische Textstrukturen hinterlegt werden.
- **Lookup Tables**: Eine Liste von Namen oder Nummerierungen, beispielsweise aller Werkzeugmaschinen, kann genutzt werden, um die Entity *machineTool* präzise zu erkennen.

Zur Festlegung eines deterministischen Verhaltens des Chatbots im Umgang mit Benutzereingaben werden darüber hinaus passende *Stories* definiert. Bezogen auf das oben genannte Beispiel könnte der Chatbot als Reaktion auf die Eingabe beispielsweise dem Benutzer die erfolgreiche Identifikation des Intents bestätigen und ihn zur Formulierung des Lösungsansatzes auffordern. Eine mögliche Rückmeldung des Chatbots könnte lauten: „Verstanden! Ich kann gerne einen neuen Lösungsansatz zur Behebung von Fehler Nummer XB103 an Werkzeugmaschine 7 anlegen. Bitte beschreibe mir deinen Lösungsansatz.“ Die daraufhin vom Benutzer eingegebene Beschreibung wird im nächsten Schritt in der Fehlerdatenbank dem genannten Fehler und der entsprechenden Maschine zugeordnet und gespeichert. Dazu wird in der Story eine Custom Action hinterlegt, die beim Eintreten des zugehörigen Intents ausgeführt wird. Das mit den Möglichkeiten des Rasa-Frameworks entwickelte Konzept zur Realisierung der Funktionen des Chatbots zur automatisierten Analyse der Nutzereingaben und Ableitung von entsprechenden Aktionen ist in Abbildung 5.5 visualisiert.

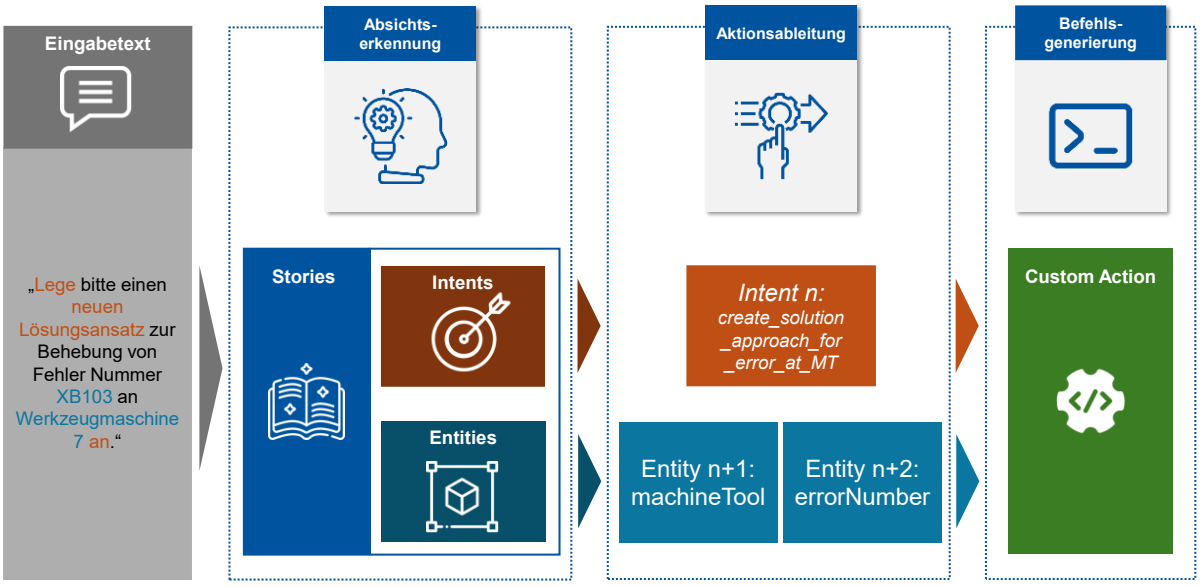



**Abbildung 5.5: Konzept zur Realisierung der Funktionen des Chatbots basierend auf dem Rasa-Framework**

*Concept of the Realization of the Functions of the Chatbot Based on the Rasa Framework*

Wie anhand des Beispiels beschrieben, erfolgt die Absichtserkennung durch die Kombination von Stories, Intents und Entities. Nach der Identifikation eines zur Benutzereingabe passenden Intents und der zugehörigen Entities wird die entsprechende Aktion abgeleitet und eine Custom Action ausgeführt. Diese nutzt die extrahierten Informationen der Entities, um spezifische Befehle zur Kommunikation mit digitalen Speichersystemen zu generieren. In dieser Custom Action müssen die erforderlichen Befehle zur Kommunikation mit den digitalen Speichersystemen so vorstrukturiert sein, dass die Informationen aus den Entities direkt eingefügt werden können, wodurch eine fehlerfreie Ausführung der Befehle gewährleistet wird.

Diese vorgegebene Struktur des Rasa-Frameworks bietet die Möglichkeit, die Verhaltensweisen eines Chatbots präzise zu definieren und bei Bedarf beliebig zu erweitern. Die dabei genutzte Kombination aus vordefinierten Regeln und Machine-Learning-Ansätzen zeichnet sich durch einen vergleichsweise geringen Ressourcenbedarf hinsichtlich der erforderlichen Rechenkapazität aus. Im Kontext des aktiven Wissensmanagements besteht jedoch die Herausforderung, alle potenziellen Absichten der Mitarbeitenden bei der Eingabe vorab zu antizipieren, um die notwendigen Intents, Entities und Stories zu entwickeln. Dies erfordert einen erheblichen manuellen Aufwand bei der Implementierung und schränkt die Generalisierbarkeit des Systems ein. Darüber hinaus besteht das Risiko, das tatsächliche Eingabeverhalten der Nutzenden nicht präzise genug abzubilden, was potenziell zu Missverständnissen führen kann.

***Absichtserkennung, Aktionsableitung und Befehlsgenerierung – LLM***

Im Gegensatz zu dem zuvor erläuterten Ansatz, der auf dem Rasa-Framework basiert und eine umfassende Determinierung des Verhaltens des Chatbots voraussetzt, wird bei der Nutzung eines LLMs auf eine derart konkrete Vorstrukturierung verzichtet. Statt das Benutzerverhalten vorab zu antizipieren und das Modell auf spezifische Szenarien zu trainieren, wird das LLM durch gezieltes Prompt Engineering so konfiguriert, dass es die für die Systemfunktionalität erforderlichen Aktionen ausführt. Eine für diese Anwendung besonders vielversprechende Methode des Prompt Engineerings ist das Prompt Chaining (siehe Kapitel 2.2.1), bei dem mehrere Prompts sequenziell an das LLM übermittelt werden. Das Ergebnis eines Prompts wird als Eingabe für den nachfolgenden Prompt verwendet, wodurch eine schrittweise Bearbeitung komplexer Aufgaben ermöglicht wird. Auf Basis dieses Vorgehens wurde das in Abbildung 5.6 dargestellte Konzept entwickelt.

Für die drei Kernfunktionen des Chatbots – Absichtserkennung, Aktionsableitung und Befehlsgenerierung – wird jeweils ein spezifisch abgestimmter Prompt formuliert. Die Ergebnisse eines Prompts werden anschließend als Eingabe für die nachfolgende Funktionalität genutzt. Innerhalb der Prompts wird dem LLM explizit mitgeteilt, welche Aufgabe auszuführen ist und welche Rahmenbedingungen dabei zu berücksichtigen sind. Die in Abbildung 5.6 dargestellten Prompts dienen als generisches Beispiel und veranschaulichen den grundsätzlichen Aufbau solcher Eingaben. Für die praktische Umsetzung des Systems müssen die Prompts jedoch mit detaillierten, domänenspe-

zifischen Informationen angereichert werden, wie beispielsweise der Struktur der angebundenen Datenbanken und der zur Kommunikation mit diesen erforderlichen Befehlen.[2] Da der Output eines LLMs, wie in Kapitel 2.2.1 beschrieben, nicht exakt deterministisch ist, werden die Ergebnisse eines Prompts bewusst als „*Erwartetes Ergebnis*“ bezeichnet. Dies berücksichtigt die inhärente Variabilität in der Generierung natürlicher Sprache durch das Modell.

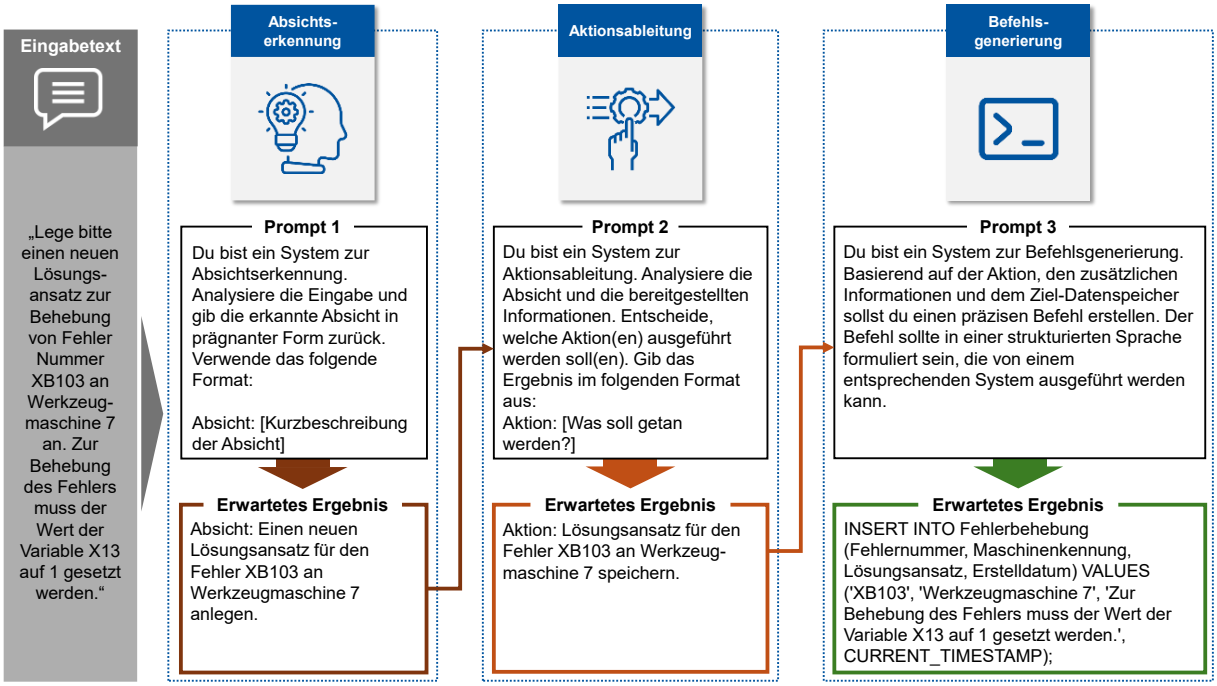



**Abbildung 5.6: Konzept zur Realisierung der Funktionen des Chatbots basierend auf Prompt Chaining für LLM**

*Concept for Realizing the Functions of the Chatbot Based on Prompt Chaining for LLM*

Je nach spezifischem Anwendungsfall im Kontext des aktiven Wissenstransfers kann der Einsatz einer Kombination mehrerer LLMs mit unterschiedlichen Stärken eine sinnvolle Strategie darstellen. Für die Analyse der Absichten der Nutzereingaben sowie die Ableitung geeigneter Aktionen sind insbesondere semantische und logische Fähigkeiten erforderlich. Diese Aufgaben sollten daher bevorzugt von einem LLM übernommen werden, welches ein ausgeprägtes semantisches Textverständnis aufweist. Die anschließende Generierung von Befehlen kann hingegen durch ein spezialisiertes LLM erfolgen, das über besondere Kompetenzen in der Erstellung spezifischer Befehlsformate (SQL-Queries, SPARQL-Abfragen oder Programmcode) verfügt. Hierbei bleibt das Prinzip des Prompt Chaining erhalten: Der Output eines LLM wird in den Prompt für die nachfolgende Aufgabe integriert. In diesem Fall wird der zweite Prompt jedoch an ein anderes, spezialisiertes LLM weitergegeben, um die jeweilige Aufgabe optimal zu erfüllen.

[2] Eine Vorstellung der konkret in dieser Arbeit verwendeten Prompts erfolgt in Kapitel 6.1.2 im Zuge der Erläuterung der Umsetzung des Konzepts.

Grundsätzlich besteht zudem die Möglichkeit, ein LLM durch Fine-Tuning an individuelle Anforderungen anzupassen. Im Kontext dieser Arbeit wird das Fine-Tuning jedoch nicht näher betrachtet, da der Fokus auf der Entwicklung von Ansätzen liegt, die ein leicht integrierbares Wissensmanagement für den Shopfloor in KMU ermöglichen. Fine-Tuning erfordert umfassende Kenntnisse über den Umgang mit LLM und deren Trainingsprozesse. Im Vergleich zur Nutzung verfügbarer, vortrainierter Modelle stellt dies eine erhebliche Herausforderung dar und erschwert die praktische Umsetzbarkeit sowie die Akzeptanz in KMU.

Insgesamt bietet die Nutzung von LLM zur Realisierung der Funktionen des Chatbots im Vergleich zur Verwendung des Rasa-Frameworks eine höhere Flexibilität bei gleichzeitig reduziertem Implementierungsaufwand. Während bei dem Ansatz mit Rasa das Risiko besteht, dass das Benutzerverhalten nicht hinreichend antizipiert wird, ist dieses Problem bei dem Vorgehen mit LLM weniger ausgeprägt. Stattdessen hängt die Effektivität maßgeblich von der Auswahl eines geeigneten LLM für die Textanalyse und Aufgabenbearbeitung ab.

Ein wesentlicher Nachteil des LLM-basierten Ansatzes ist dessen geringere Deterministik, was potenziell zu Inkonsistenzen bei der Erfassung und Speicherung von Wissen führen kann. Zudem fungieren LLM häufig als Blackbox-Systeme, bei denen der Prozess von der Eingabe bis zur Ausgabe weniger transparent ist als bei regelbasierten Ansätzen wie Rasa. Weiterhin können die benötigten Hardwareressourcen, abhängig von der Größe und Komplexität des LLMs, erheblich variieren, was die Anwendbarkeit des Ansatzes insbesondere für KMU einschränkt. Wie zu Beginn dargelegt, werden im Rahmen der vorliegenden Arbeit beide Ansätze implementiert und validiert, um abschließend eine fundierte Bewertung vorzunehmen und eine begründete Schlussfolgerung hinsichtlich des vorteilhafteren Ansatzes zu ziehen.

### 5.1.2 Chatbot-basiertes System für einen effizienten und intuitiven Abruf von Wissen auf dem Shopfloor

*Chatbot-based System for Efficient and Intuitive Knowledge Retrieval on the Shop Floor*

Das anhand von Abbildung 5.4 vorgestellte Konzept für den Chatbot-basierten Wissenstransfer sowie die beiden detaillierteren Konzepte zur Umsetzung der Kernfunktionen des Chatbots mittels Rasa (siehe Abbildung 5.5) und LLM (siehe Abbildung 5.6) beziehen sich in der bisherigen Beschreibung und Darstellung ausschließlich auf das *Erfassen* von Erfahrungswissen von Mitarbeitenden. Dieser Prozess repräsentiert im Wesentlichen den *Upstream* des aktiven Wissenstransfers, bei dem Wissen vom Nutzer über den Chatbot in die digitalen Speichersysteme übertragen wird. Abhängig vom Informationsgehalt der Nutzereingaben erfolgt dabei eine dialogbasierte Interaktion zwischen Nutzer und Chatbot, die auf einer klassischen Request-Response-Struktur basiert. Der Nutzer initiiert den Prozess durch eine Anfrage (Request) an den Chatbot, der diese verarbeitet, eine entsprechende Aktion ableitet und gegebenenfalls zusätzliche Informationen anfordert, falls diese für die Ausführung der Aktion erforderlich sind. Diese Rückfrage erfolgt über eine generierte Antwort (Response), in der gezielt

nach weiteren Details gefragt wird. Sobald alle notwendigen Informationen vorliegen, führt der Chatbot die Aktion aus und das Wissen des Nutzers wird in das entsprechende digitale Speichersystem übertragen. Für den *Abruf* von bereits gespeichertem Wissen können die vorgestellten Konzepte sowie der dialogbasierte Austausch mit dem Chatbot analog verwendet werden. Dieser Prozess repräsentiert den *Downstream* des aktiven Wissenstransfers, bei dem Wissen aus den digitalen Speichersystemen über den Chatbot zum Nutzer übertragen wird. Dabei werden zunächst die gleichen Schritte wie bei der Erfassung von Wissen durchlaufen (siehe Abbildung 5.4). Nach der Kommunikation mit den digitalen Speichersystemen erfolgt zusätzlich eine Antwort des Chatbots, der das angeforderte Wissen in geeigneter Form an den Nutzer transferiert und somit den Abrufprozess abschließt. Der sich daraus ergebende Ablauf der Erfassung sowie des Abrufs von Wissen ist in Abbildung 5.7 schematisch dargestellt.

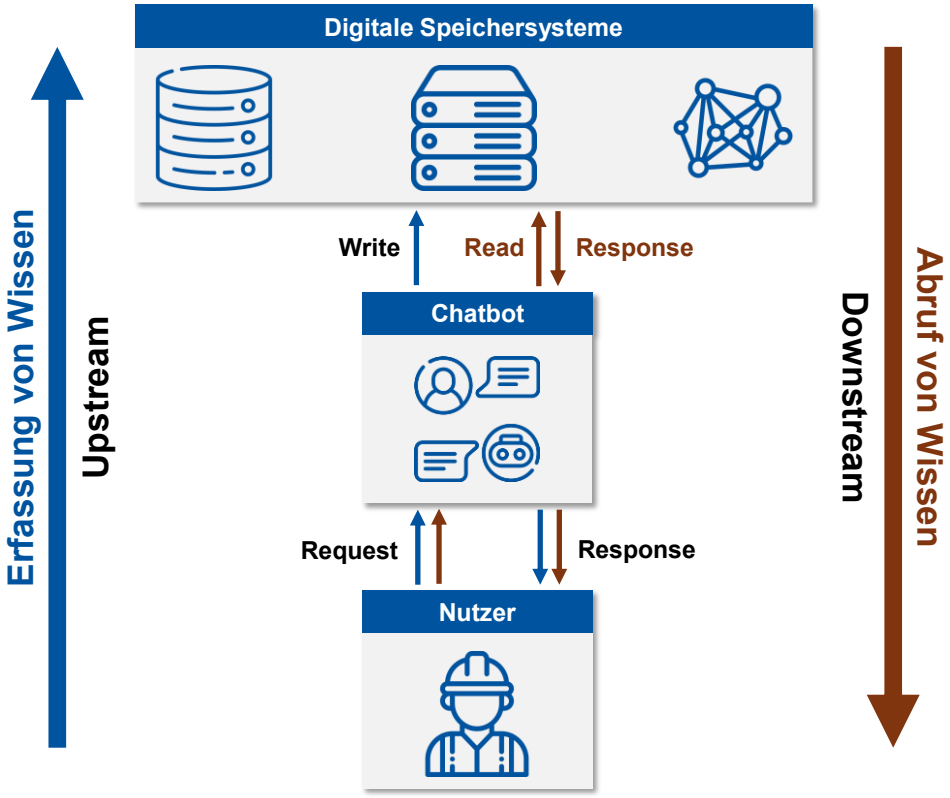



**Abbildung 5.7: Ablauf der Erfassung und des Abrufs von Wissen per Chatbot**

*Process of Capturing and Retrieving Knowledge via Chatbot*

Der wesentliche Unterschied zwischen diesen beiden Prozessen liegt in der Art der Kommunikation zwischen dem Chatbot und den digitalen Speichersystemen. Während bei der Wissensspeicherung Informationen mittels eines Write-Befehls in die digitalen Speichersysteme eingetragen werden, erfolgt beim Wissensabruf eine Anforderung durch einen Read-Befehl. Das adressierte Speichersystem liefert daraufhin eine Response, die die angeforderten Informationen enthält. Diese Informationen werden anschließend durch den Chatbot in einen für den Benutzer verständlichen Text aufbereitet. Für diesen Schritt können entweder die im Rasa-Framework vorgesehenen

Funktionen wie Stories, Intents und Entities oder entsprechend geeignete Prompts für ein LLM verwendet werden.

Die Funktionen *Wissensabruf* und *Wissenserfassung* können im User Interface des Wissensmanagementsystems grundsätzlich getrennt realisiert werden, sodass die Mitarbeitenden zu Beginn zwischen den beiden Prozessen auswählen müssen. In diesem Fall werden für den Upstream (Erfassung) und den Downstream (Abruf) separate Systeme mit jeweils eigenständigen Stories, Intents und Entities (bei Nutzung des Rasa-Frameworks) oder speziell angepassten Prompts (bei Nutzung eines LLM) erstellt. Alternativ können beide Funktionen in einem einzigen System integriert werden. In diesem Szenario übernimmt die Absichtserkennung des Chatbots die initiale Differenzierung, ob der Nutzer eine Erfassung oder einen Abruf von Wissen vornehmen möchte. Diese Kernfunktion erhält damit eine mehrstufige Struktur, die bei der Konfiguration im Rasa-Framework oder bei der Gestaltung des Prompt Chainings für ein LLM entsprechend berücksichtigt werden muss.

## 5.2 Semi-aktiver Wissenstransfer

*Semi-active Knowledge Transfer*

Im vorangegangenen Unterkapitel wird ein Konzept für Wissenstransfer vorgestellt, bei dem Mitarbeitende ihre Erfahrungen zur Problemlösung und Aufgabenbewältigung aktiv verbal oder schriftlich formulieren können. Dabei obliegt es ihnen, selbstständig den Umfang und Detailgrad der weitergegebenen Informationen festzulegen. Daraus resultiert jedoch die Gefahr, dass Mitarbeitende ihre Erfahrungen auf Grundlage ihres eigenen Kenntnisstands dokumentieren und dabei bestimmte Informationen auslassen, die für sie selbstverständlich sind, für weniger erfahrene Kolleginnen und Kollegen jedoch von grundlegender Bedeutung sein könnten. Zudem ist eine rein textuelle Beschreibung in vielen Fällen nicht ausreichend, um eine Arbeitsaufgabe hinreichend präzise darzustellen. Die technologische Unterstützung des vorgestellten Chatbots ermöglicht zwar eine strukturierte Erfassung des eingegebenen Wissens und dessen Anreicherung mit zusätzlichen Informationen in verschiedenen Formaten (z.B. Bilder, Videos, CAD-Dateien), setzt jedoch voraus, dass die referenzierten Informationen entsprechend verfügbar sind.

An dieser Stelle knüpft das Konzept des *semi-aktiven Wissenstransfers* an. Dieser Ansatz beruht auf der Überlegung, die Aktivitäten der Mitarbeitenden aus ihrer Perspektive – dem sogenannten „Point-of-View“ (PoV) – mittels optischer und akustischer Erfassungssysteme aufzuzeichnen und digital zu speichern. Die erfassten Daten werden anschließend automatisiert analysiert, um die dokumentierten Aktivitäten auf Basis einzelner Prozessschritte in kleinere Untereinheiten zu differenzieren und daraus kompakte, leicht konsumierbare Wissenseinheiten, sogenannte Learning Nuggets [BAIL06], zu generieren. Diese Wissenseinheiten sollen unerfahrene Mitarbeitende beim Erlernen von neuen Tätigkeiten und dem Umgang mit ihnen unbekannten Situationen unterstützen. Entsprechend dem semi-aktiven Ansatz erfolgt der Wissenstransfer dabei sowohl passiv über die visuelle Dokumentation als auch aktiv über verbale

Beschreibungen relevanter Aspekte der durchgeführten Aktivitäten durch die erfahrenen Mitarbeitenden. Die Mitarbeitenden führen ihre Tätigkeiten dabei wie gewohnt aus und sind lediglich gefordert, eine geringe zusätzliche kognitive Leistung in Form der Erläuterungen zu erbringen. Dieses Vorgehen kann gemäß dem zugrunde liegenden Prinzip der *prozessparallelen Erfassung von Erfahrungswissen* anhand der in Abbildung 5.8 dargestellten Systemarchitektur veranschaulicht werden.

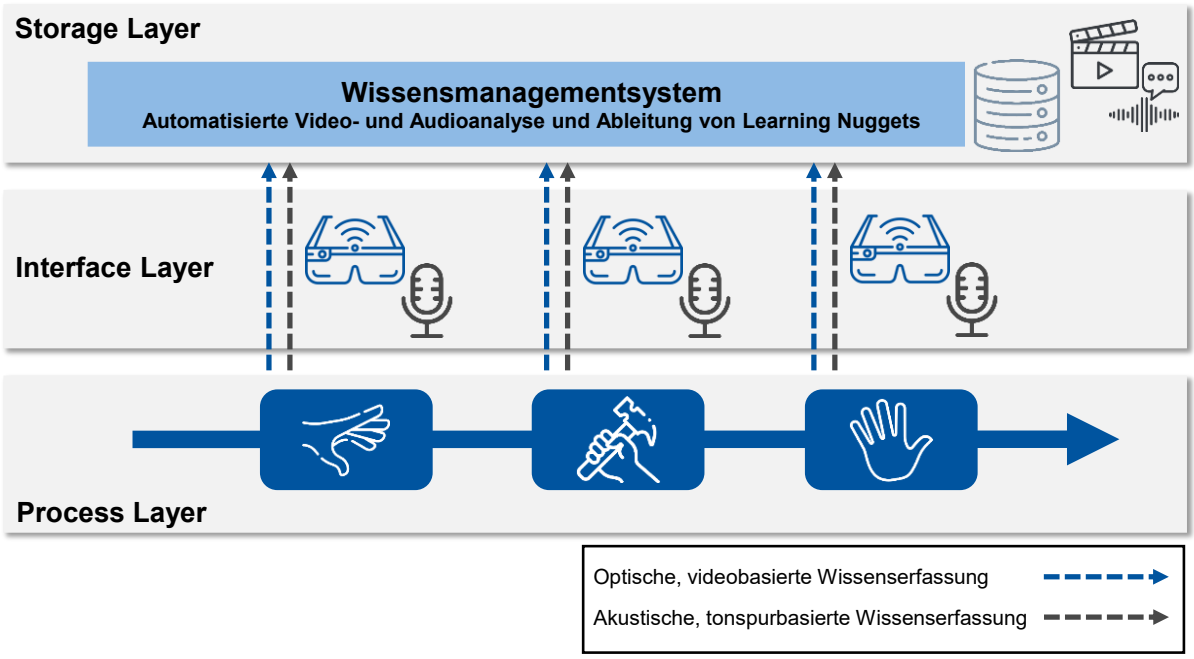



**Abbildung 5.8: Systemarchitektur des semi-aktiven Wissenstransfers**

*System Architecture of the Semi Active Knowledge Transfer*

Die im Process Layer stattfindenden Aktivitäten der Mitarbeitenden werden prozessparallel mithilfe einer Videokamera und eines Mikrofons erfasst und im Wissensmanagementsystem im Storage Layer automatisiert ausgewertet. Die zur Erfassung eingesetzte Kamera und das Mikrofon werden dabei von den erfahrenen Mitarbeitenden mitgeführt. Um eine umfassende Erfassung der Handlungen aus der PoV-Perspektive zu gewährleisten, wird die Kamera am Kopf positioniert und entsprechend der Blickrichtung ausgerichtet. Dies lässt sich beispielsweise durch den Einsatz von Helmkameras oder Smart Glasses realisieren. Für eine optimale Klangqualität sollte das eingesetzte Mikrofon in unmittelbarer Nähe des Mundes der Mitarbeitenden positioniert werden und kann in die Kamera integriert sein. Durch die Kombination aus videobasierter Dokumentation und ergänzendem Audiokommentar werden sowohl die ausgeführten Tätigkeiten und deren konkrete Durchführung als auch die Begründung, warum bestimmte Handlungen auf eine spezifische Weise ausgeführt werden müssen, erfasst.

Zur Ableitung der Learning Nuggets, also zur inhaltlichen Segmentierung der aufgenommenen Videosequenzen, können verschiedene Technologien genutzt werden. Unter Verwendung spezialisierter Methoden des maschinellen Lernens können sowohl einzelne Objekte im Video detektiert als auch die verbalen Beschreibungen auf der Tonspur semantisch analysiert werden. Basierend auf diesen Auswertungen können

die Video- und Audiodaten anschließend in prozessschrittspezifische Segmente unterteilt werden. Auf diese Weise werden Best Practices digital dokumentiert und können unmittelbar in Lernmaterialien für einen effizienten Wissenstransfer integriert werden.

Im folgenden Unterkapitel wird ein praxisorientiertes Konzept zur Umsetzung des semi-aktiven Wissenstransfers erläutert, welches die im Stand der Technik identifizierten Defizite derzeit verfügbarer kommerzieller Systeme zur videobasierten Erfassung von Prozessabläufen und deren Überführung in Lerneinheiten adressiert und eine Optimierung ermöglicht.

### 5.2.1 Automatisierte Analyse und Segmentierung von Prozessvideos zur Effizienzsteigerung beim Wissenstransfer auf dem Shopfloor

*Automated Analysis and Segmentation of Process Videos to Increase Efficiency in Knowledge Transfer on the Shop Floor*

In Kapitel 2.3.1 des Stands der Technik werden verschiedene AR-Anwendungen benannt, die eine grundlegende Möglichkeit bieten, Best Practices bei der Prozessausführung auf dem Shopfloor videobasiert zu erfassen und in Schritt-für-Schritt-Anleitungen zu überführen. Die auf diese Weise erstellten Instruktionen können anschließend von weniger erfahrenen Mitarbeitenden über ein geeignetes Endgerät, wie z.B. eine AR-Brille, zur Aneignung der Vorgehensweisen genutzt werden. Abbildung 5.9 veranschaulicht den grundlegenden Ablauf dieses Prozesses, der bei den meisten kommerziell verfügbaren Lösungen ähnlich strukturiert ist: Nach der Erfassung der Aktivitäten per Video und Ton erfolgt eine manuelle Inhaltsaufbereitung, bei der das Videomaterial händisch entsprechend der Prozessschritte sequenziert und mit weiteren relevanten Informationen (wie etwa Bauteil- oder Werkzeuginformationen, technische Zeichnungen, Schaltpläne) verknüpft wird.




**Abbildung 5.9: Grundsätzlicher Ablauf für die videobasierte Erfassung von Best Practices und Überführung in Lerneinheiten**

*Basic Procedure for the Video-Based Recording of Best Practices and Transfer into Learning Units*

Einige kommerzielle Anwendungen (z.B. PTC Vuforia) bieten zudem die Möglichkeit, einen Teil der manuellen Aufbereitung vorwegzunehmen, indem eine feste Struktur für den Erfassungsprozess definiert wird. Die erfahrenen Mitarbeitenden werden dabei

schrittweise durch den Erfassungsprozess geführt und geben selbst an, wann ein Prozessschritt abgeschlossen ist und der nächste beginnt. Grundsätzlich erfordern kommerziell verfügbare Lösungen jedoch immer einen manuellen Aufwand bei der Aufbereitung, der sinnvollerweise nur von Mitarbeitenden durchgeführt werden kann, die über ein hinreichendes Fachwissen in der jeweiligen Thematik verfügen. Dies ist erforderlich, um eine präzise Sequenzierung des Videomaterials vorzunehmen und die zugehörigen ergänzenden Inhalte passend auszuwählen und zuzuordnen.

Im Sinne einer effizienten Erfassung von Best Practices und deren Überführung in Lerneinheiten zielt das folgende Konzept auf die Automatisierung des bislang manuellen Aufbereitungsprozesses ab. Entsprechend des semi-aktiven Wissenstransfers wird dabei sichergestellt, dass die erfahrenen Mitarbeitenden bei der Dokumentation ihrer Aktivitäten nicht durch vorgegebene Aufzeichnungsabläufe eingeschränkt werden. Abbildung 5.10 zeigt das Konzept der videobasierten Erfassung von Best Practices mit automatisierter Aufbereitung in schematischer Form.

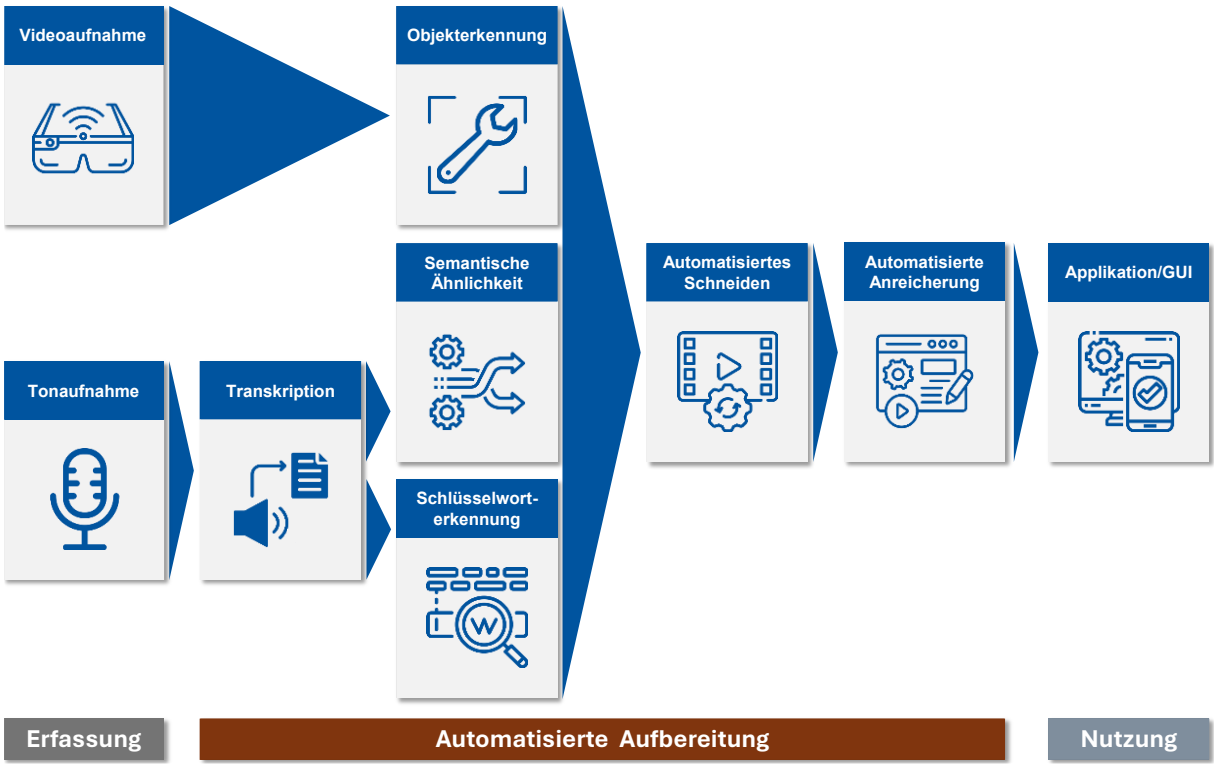



**Abbildung 5.10: Neuartiger Ablauf für die videobasierte Erfassung von Best Practices mit automatisierter Aufbereitung**

*Innovative Process for the Video-Based Recording of Best Practices with Automated Processing*

Während die Erfassung der Aktivitäten analog zu den kommerziell erhältlichen Anwendungen geschieht, liegt die Innovation in der Entwicklung der *automatisierten Aufbereitung*. Diese umfasst mehrere konsekutive Schritte, welche die üblicherweise manuell ausgeführten Tätigkeiten substituieren. Die einzelnen Bestandteile des Verfahrens der automatisierten Aufbereitung sowie die darauf aufbauende, angepasste Nutzung der resultierenden Inhalte werden in den folgenden Abschnitten detailliert beschrieben.

### Erfassung der Aktivitäten – *Video- und Tonaufnahme*

Die Aufnahme von Video und Ton aus der PoV-Perspektive erfolgt über eine am Kopf des Mitarbeitenden befestigte Kamera und ein Mikrofon. Die Kamera muss so positioniert werden, dass sie in Blickrichtung ausgerichtet ist und den visuellen Fokus des Blickfeldes des Mitarbeitenden abbildet. Für eine hinreichend scharfe, flüssige und detailreiche Erfassung des Blickfelds sollte die Kamera mindestens eine Auflösung von 1920 x 1080 Pixeln (Full HD) und eine Aufzeichnungsrate von 30 Bildern pro Sekunde unterstützen. Zur exakten Erfassung der Stimme des Mitarbeitenden ist ein gerichtetes Mikrofon erforderlich, das auf den Frequenzbereich und Dynamikumfang menschlicher Sprachaufnahmen abgestimmt ist. Idealerweise sollte die Empfindlichkeit des Mikrofons an die Umgebungsbedingungen angepasst werden können. Die resultierenden Video- und Audiodaten werden im weiteren Verlauf getrennt analysiert.

### Automatisierte Aufbereitung – *Objekterkennung*

Für die spätere Sequenzierung der Videoaufnahme in handlungsbezogene Learning Nuggets wird der Inhalt des Videos im Hinblick auf die von dem erfahrenen Mitarbeitenden ausgeführten Prozessschritte analysiert. Dazu wird angenommen, dass ein Werkzeugwechsel den Beginn eines neuen Prozessschritts markiert. Zur Identifizierung von Werkzeugwechseln sowie zur Erfassung der verwendeten Werkzeuge wird eine Objekterkennung eingesetzt, die mithilfe eines geeigneten neuronalen Netzwerks jedes Bild des Videos analysiert und die erkannten Werkzeuge detektiert. In Abbildung 5.11 ist das Konzept der Analyse des Videos mittels Objekterkennung zur Detektion der Zeitpunkte von Werkzeugwechseln schematisch dargestellt. Bei erfolgreicher Detektion eines Werkzeugs wird im analysierten Bild eine Bounding Box um jedes erkannte Werkzeug gezogen (siehe Kapitel 2.2.5).

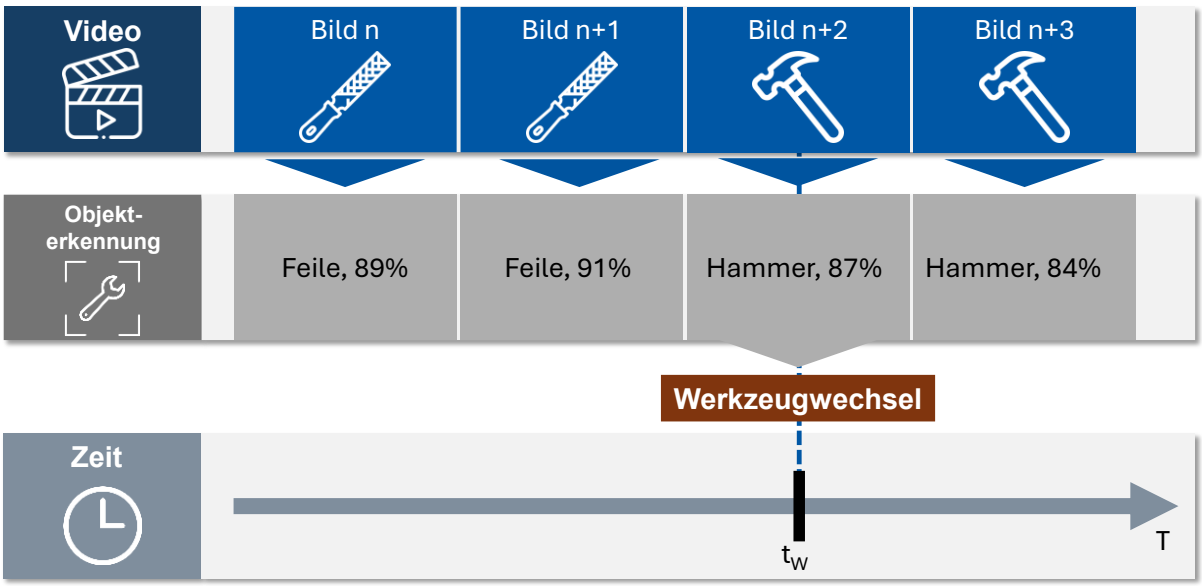



**Abbildung 5.11: Konzept der Analyse des Videos mittels Objekterkennung zur Detektion der Zeitpunkte von Werkzeugwechseln**

*Concept of Analyzing the Video Using Object Recognition to Detect the Time of Tool Changes*

Zusätzlich werden die zugehörige Klassifizierung sowie das Konfidenzniveau, also die Wahrscheinlichkeit, mit der das Werkzeug korrekt erkannt wurde, dargestellt. Die auf diese Weise überarbeiteten Einzelbilder werden nach Abschluss der Analyse aller Bilder des ursprünglichen Videos zu einem neuen Video zusammengefügt, das dem Betrachter die identifizierten Informationen zu den erkannten Werkzeugen visualisiert.

### Transkription

Parallel zur Analyse der Bilddaten wird auch die zugehörige Audiodatei automatisiert ausgewertet, um Informationen über den Prozessablauf zu extrahieren. Dazu wird die Tonspur im ersten Schritt transkribiert, indem ein spezialisiertes neuronales Netzwerk aus dem Bereich des NLP angewendet wird. Das Ergebnis der Transkription enthält die transkribierten Textsegmente, jeweils versehen mit Start- und Endzeitpunkten in Sekunden sowie einem Konfidenzniveau, welches die Wahrscheinlichkeit der Korrektheit der Transkription angibt. Dieses Vorgehen ist in Abbildung 5.12 vereinfacht visualisiert.

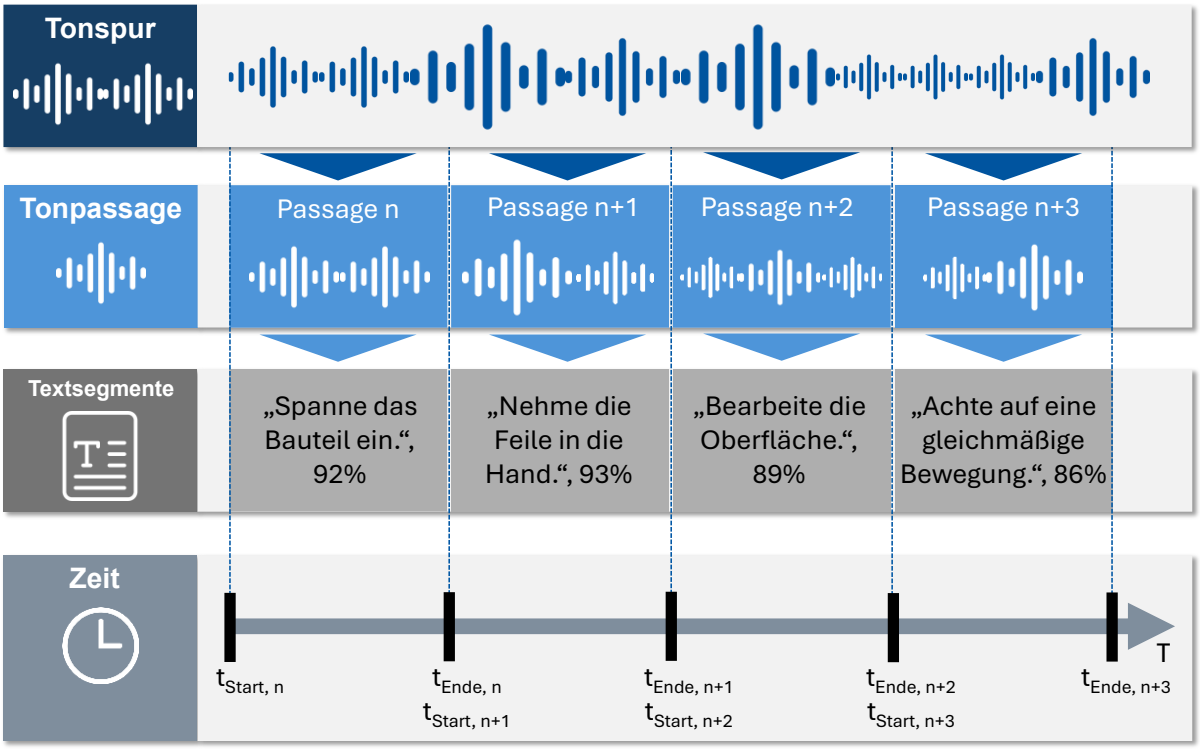



**Abbildung 5.12: Vereinfachte Darstellung des Vorgangs der Transkription der Tonspur**

*Simplified Representation of the Process of Transcribing the Audio Track*

Die von dem erfahrenen Mitarbeitenden verbal ausgedrückten Informationen liegen nun digital in Textform vor und können im nächsten Schritt inhaltlich analysiert werden. Für diese Analyse werden zwei unterschiedliche Methoden angewendet, deren Ergebnisse in der späteren Validierung (siehe Kapitel 7.1.2) miteinander verglichen werden. Die erste Methode untersucht die Texte anhand vordefinierter Schlüsselwörter, während die zweite Methode ein Transformer-Modell verwendet, um die Texte semantisch zu analysieren.

***Schlüsselworterkennung***

Der Inhalt der transkribierten Texte wird zunächst analysiert, indem die Texte auf das Vorhandensein von vorab definierten Schlüsselworten geprüft werden. Diese Schlüsselworte lassen sich inhaltlich differenzieren, indem sie in semantische Kategorien eingeteilt werden. Im Rahmen dieser Arbeit werden fünf solcher Kategorien unterschieden, die in Tabelle 5.1 definiert sind. Abhängig von der betrachteten Tätigkeit und dem zugehörigen Produktionsbereich können die Kategorien *Aktionen*, *Objekte* und *Werkzeuge* spezifisch befüllt werden. Dazu werden Listen mit Begriffen erstellt, die auf die jeweilige Tätigkeit abgestimmt sind. Im Gegensatz dazu lassen sich die Kategorien *Sequenzindikatoren* und *Warnhinweise* allgemeiner anwenden, da die darin enthaltenen Begriffe unabhängig von einer spezifischen Tätigkeit oder einem bestimmten Produktionsbereich im Sprachgebrauch verwendet werden. Die Sequenzindikatoren liefern eindeutige Hinweise auf den Beginn eines neuen Prozessschritts und können deshalb als verlässliche Grundlage der Sequenzierung der transkribierten Texte genutzt werden. Ähnliches gilt auch für die Warnhinweise, da es im Sinne einer sicheren Arbeitsumgebung wichtig ist, dass unerfahrene Mitarbeitende beim Erlernen neuer Tätigkeiten gezielt auf sicherheitsrelevante Aspekte hingewiesen werden. Sobald ein Warnhinweis auftritt, sollte der nachfolgende Inhalt für die spätere Weiterverarbeitung separiert werden.

**Tabelle 5.1: Semantische Kategorien der Schlüsselworte für die Textanalyse**

*Semantic Categories of Keywords for Text Analysis*

| Kategorie | Definition | Beispiele |
|---|---|---|
| **Aktionen** | Verben und Tätigkeitsbeschreibungen | feilen, kleben, messen, anbringen, schrauben |
| **Objekte** | Physische Teile, die für eine Aufgabe verwendet werden | Schraube, Mutter, Lager, Welle, Zahnrad, Dichtung |
| **Sequenzindikatoren** | Begriffe, die die Reihenfolge der Arbeitsanweisungen andeuten | zuerst, danach, anschließend, darauf aufbauend |
| **Warnhinweise** | Begriffe, die auf potenzielle Gefahren oder Risiken hinweisen | Achtung, riskant, gefährlich, problematisch, Vorsicht |
| **Werkzeuge** | Für eine Aufgabe verwendete Werkzeuge und Hilfsmittel | Schraubenschlüssel, Hammer, Feile, Zange |

Abhängig von den sprachlichen Gewohnheiten des erfahrenen Mitarbeitenden kann eine Sequenzierung anhand der beschriebenen Aktionen oder der verwendeten Objekte und Werkzeuge hilfreich sein, da Sequenzindikatoren nicht immer explizit genutzt werden. In diesem Kontext wird erneut angenommen, dass ein Werkzeugwechsel den Übergang zu einem neuen Prozessschritt signalisiert. Die Fokussierung auf Objekte ist insbesondere bei Montagetätigkeiten sinnvoll, da Montageanleitungen typischerweise der Reihenfolge der zu montierenden Bauteile folgen. Auf Basis dieses Vorgehens kann die Sequenzierung der transkribierten Texte, einschließlich der zugehörigen Zeitstempel, durch eine Kombination regelbasierter Algorithmen und String-basierter

Operationen zur strukturierten Textbearbeitung erfolgen. Das Konzept der Analyse der Tonspur auf Grundlage der Schlüsselworterkennung ist in Abbildung 5.13 dargestellt.

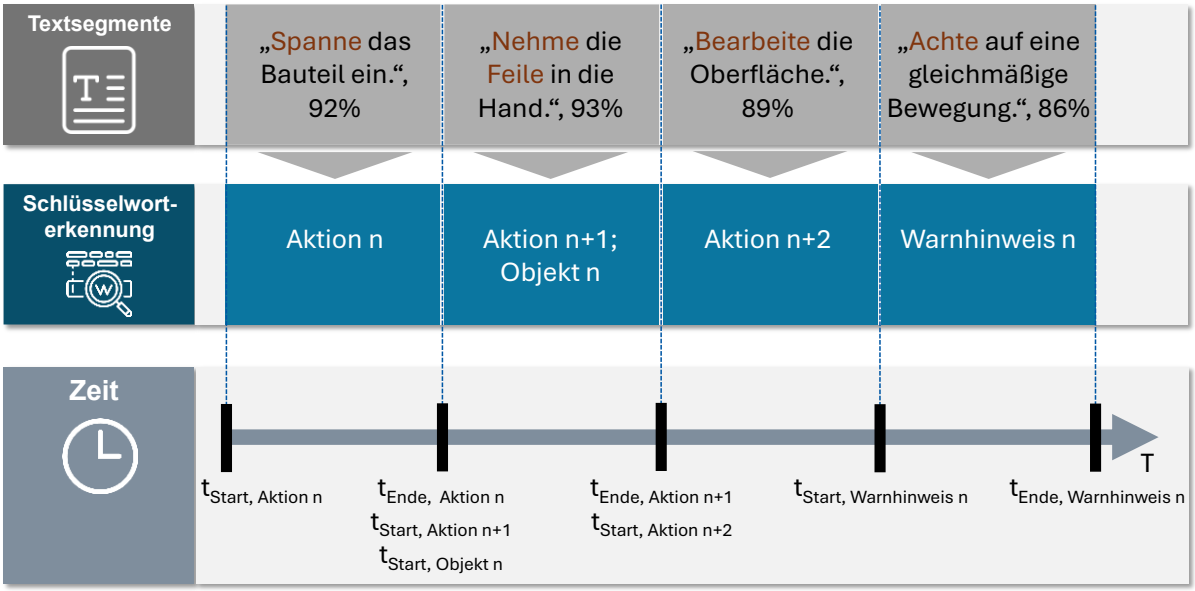



**Abbildung 5.13: Konzept der Analyse der Tonspur mittels Schlüsselworterkennung**

*Concept of Analyzing the Audio Track Using Keyword Recognition*

### *Semantische Ähnlichkeit*

Die zweite Methode zur Analyse der transkribierten Texte basiert auf der semantischen Ähnlichkeitsbewertung aufeinanderfolgender Textsegmente. Abbildung 5.14 zeigt das Konzept der semantischen Analyse der Tonspur in schematischer Form. Im ersten Schritt wird der Text in einzelne Sätze segmentiert und anschließend mithilfe geeigne-

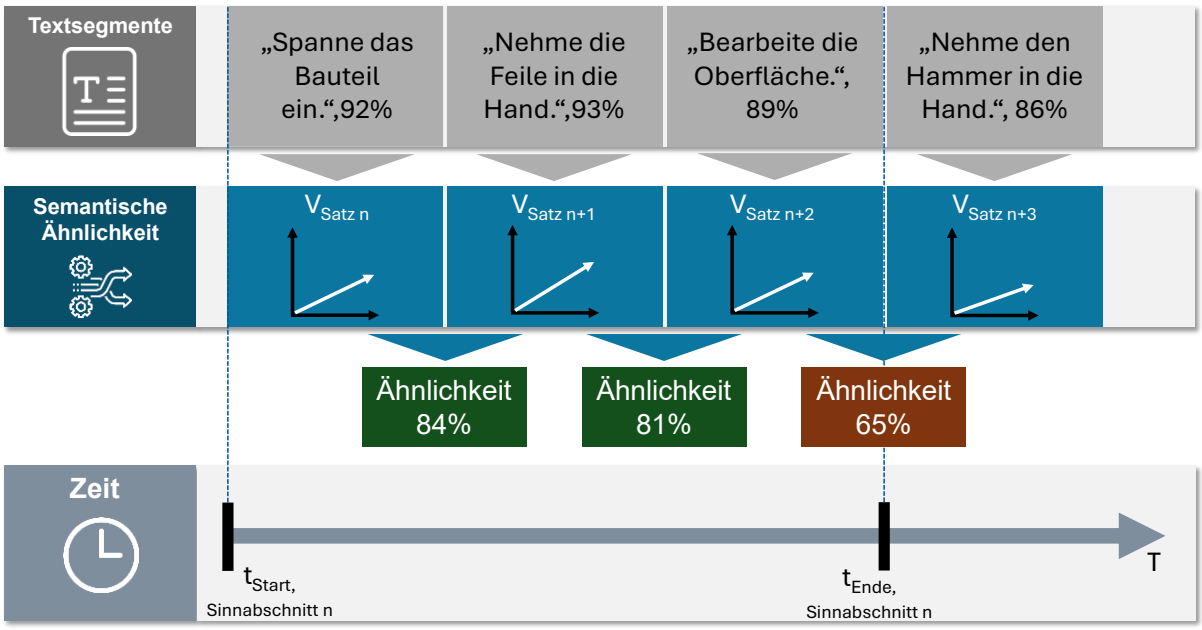



**Abbildung 5.14: Konzept der Analyse der Tonspur zur Identifikation semantischer Ähnlichkeit**

*Concept of Analyzing the Audio Track to Identify Semantic Similarity*

ter Sentence Transformer (siehe Kapitel 2.2.4) in Embeddings überführt. Zur Identifikation zusammenhängender Sinnabschnitte wird ein Schwellenwert für die Ähnlichkeit der Embeddings der Sätze definiert. Solange der Ähnlichkeitswert zwischen den betrachteten Textsegmenten den Schwellenwert nicht unterschreitet, werden die Segmente als inhaltlich kohärent und zusammengehörig interpretiert. Wird der Schwellenwert jedoch unterschritten, markiert dies den Beginn eines neuen Sinnabschnitts, für den die Prüfung der Zusammengehörigkeit der nachfolgenden Textsegmente nach demselben Verfahren erfolgt. Dieses iterative Vorgehen wird fortgeführt, bis die Analyse des gesamten Textes abgeschlossen ist.

### *Automatisiertes Schneiden*

Sowohl die Analyse der Bilddateien des Videos per Objekterkennung als auch die semantische Auswertung der Tonspur mittels Schlüsselworterkennung und Ähnlichkeitsberechnung liefern wertvolle Informationen über die Zeitpunkte, an denen im aufgezeichneten Arbeitsablauf neue Prozessschritte beginnen. Diese Informationen bilden die Grundlage für das automatisierte Schneiden des Videos in didaktisch sinnvoll strukturierte Learning Nuggets. Die ermittelten Zeitpunkte werden an geeignete programmiertechnische Routinen aus dem Bereich des Video Processings übergeben, um das Video an den entsprechenden Stellen zu unterteilen. Dieses Vorgehen ist in Abbildung 5.15 exemplarisch dargestellt, wobei im gezeigten Fall der Zeitpunkt eines durch Objekterkennung detektierten Werkzeugwechsels als Schnittstelle dient.

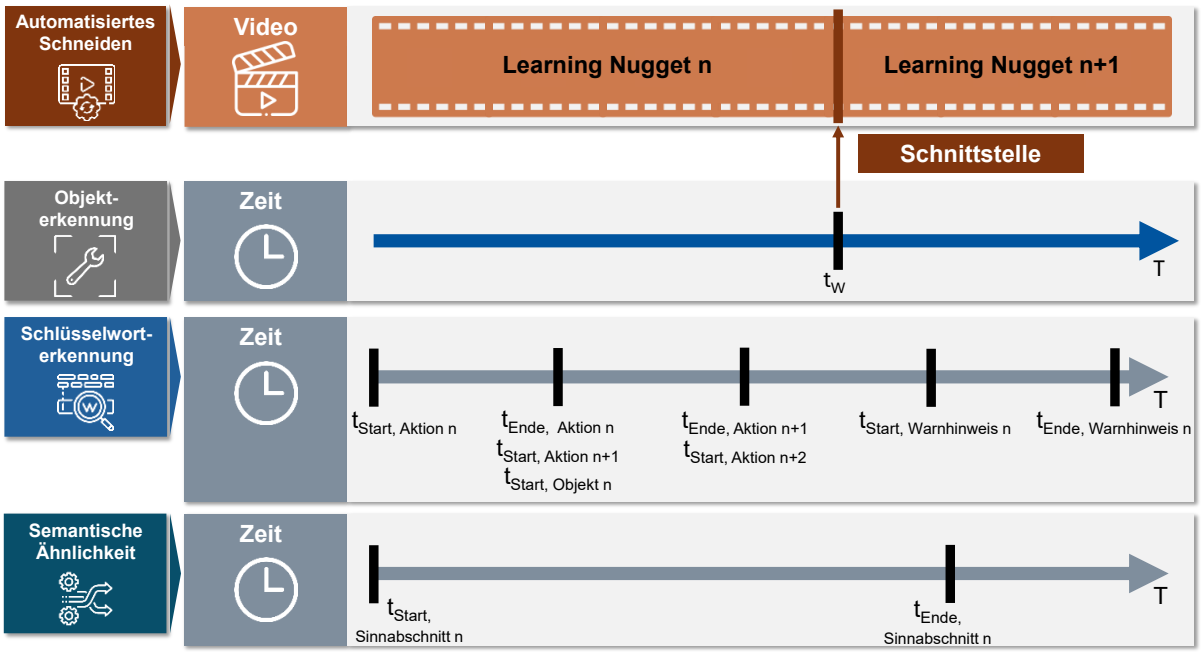



**Abbildung 5.15: Konzept des automatisierten Schneidens eines Videos auf Basis der per Objekterkennung, Schlüsselworterkennung oder durch die Identifikation von semantischer Ähnlichkeit identifizierten Zeitpunkte**

*Concept of Automated Cutting of a Video Based on the Points in Time Identified by Object Recognition, Keyword Recognition or by Identification of Semantic Similarity*

Im Rahmen dieser Arbeit werden die durch Objekterkennung, Schlüsselworterkennung und semantische Ähnlichkeit ermittelten Zeitpunkte jeweils separat für die Sequenzierung des Videos verwendet, um die Korrektheit der daraus resultierenden Prozessschritte im Anschluss evaluieren zu können (siehe Kapitel 7.1.2). Grundsätzlich ist auch eine Kombination der drei verschiedenen Informationsquellen möglich, wodurch eine potenziell robustere und präzisere Sequenzierung der Videos erzielt werden kann. Beispielsweise könnte die Schnittmenge der Zeitpunkte bestimmt werden, sodass ausschließlich die von allen drei Methoden als Übergänge zwischen zwei Prozessschritten identifizierten Zeitpunkte für die Sequenzierung herangezogen werden. Die derart sequenzierten Learning Nuggets werden mit den Informationen aus dem genutzten Verfahren zur Identifikation der Zeitpunkte der Schnittstellen angereichert. Im Beispiel aus Abbildung 5.15 werden die Learning Nuggets n und n+1 entsprechend mit den Werkzeugen getaggt, welche in ihren jeweiligen Videosegmenten vorkommen. Analog können die erkannten Schlüsselwörter oder semantisch ausgewerteten Inhalte dafür genutzt werden. Die so generierten Metadaten können nicht nur zur Erzeugung eines Titels für das jeweilige Learning Nugget genutzt werden, sondern bilden auch die Basis für weiterführende Funktionen im Bereich der automatisierten Anreicherung und Nutzung der Inhalte.

### *Automatisierte Anreicherung*

Die automatisierte Anreicherung (siehe Abbildung 5.16) ermöglicht die Ergänzung der Learning Nuggets durch bereits digital hinterlegte Informationen, die eine inhaltliche Vertiefung oder Erweiterung des jeweiligen Nuggets darstellen. Grundlage hierfür sind die in den Metadaten enthaltenen Informationen über die zugehörigen Prozessschritte. Ein Mapping zwischen den identifizierten Werkzeugen, Objekten oder Begriffen und den in einer Datenbank hinterlegten Ressourcen (z.B. technischen Zeichnungen, Handbüchern) gewährleistet eine zielgerichtete Zuordnung relevanter Zusatzinformationen.

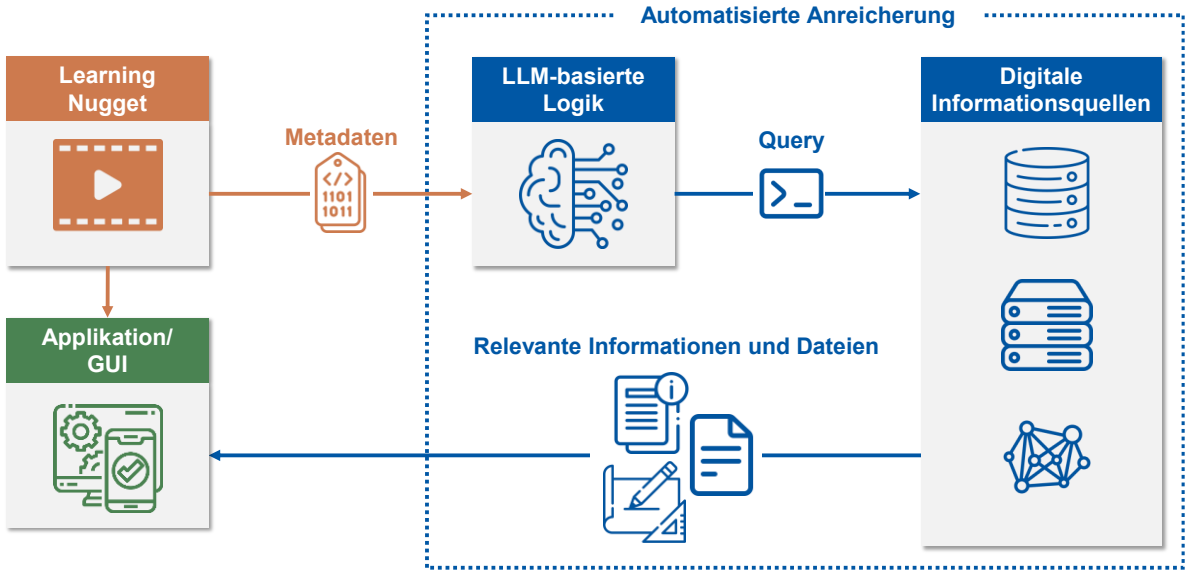



**Abbildung 5.16: Ablauf der automatisierten Anreicherung**

*Procedure for Automated Enrichment*

So wird beispielsweise bei der Identifikation eines Dremels in einem Learning Nugget die dazugehörige Betriebsanweisung automatisiert abgerufen und dem Nugget zugeordnet. Dieses Mapping kann durch die Nutzung eines LLMs realisiert werden, welches die in den Metadaten enthaltenen Informationen semantisch interpretiert und darauf basierend zielgerichtete Anfragen an Datenbanken, Fileserver und Ontologien formuliert. Auf diese Weise können relevante Informationen und Dateien automatisiert identifiziert und abgerufen werden. Das Ergebnis dieser Anreicherung wird im User Interface visualisiert und bietet unerfahrenen Mitarbeitenden zusätzliche Unterstützung im Lernprozess.

**Nutzung – *Applikation/GUI***

Für die Nutzung des zuvor beschriebenen Konzepts ist im letzten Schritt die Entwicklung einer entsprechenden Applikation mit einer geeigneten grafischen Benutzerschnittstelle (GUI) notwendig. Diese GUI soll die Erstellung und Nutzung der Learning Nuggets ermöglichen. Dazu sollen Funktionen vorgesehen werden, welche das Laden eines Videos sowie anschließend die Auswahl des Vorgehens zur Segmentierung in Learning Nuggets umfassen. Für das effektive Erlernen neuer Tätigkeiten bedarf es zudem einer übersichtlichen Darstellung der vorhandenen Learning Nuggets. Bei der Auswahl eines Learning Nuggets sollen die zusätzlich relevanten Informationen und Dokumente strukturiert und übersichtlich dargestellt werden.

Für die Anwendung in der Praxis ist eine funktionale Trennung zwischen der Erstellung der Learning Nuggets durch erfahrene Mitarbeitende und deren anschließender Nutzung im Lernprozess durch unerfahrene Mitarbeitende sinnvoll. Diese Trennung kann entweder durch den Einsatz zweier separater Systeme oder durch rollenspezifische Ansichten innerhalb einer gemeinsamen Applikation realisiert werden. Je nach Anwendungsbereich und vorhandenen digitalen Assistenzsystemen kann es zudem sinnvoll sein, insbesondere die Elemente zur Nutzung der Learning Nuggets in bestehende Systeme zu integrieren. Dies reduziert Redundanzen und minimiert die kognitive Belastung der Mitarbeitenden, die andernfalls im Lernprozess mit der Nutzung mehrerer Systeme konfrontiert wären.

## 5.3 Passiver Wissenstransfer

*Passive Knowledge Transfer*

In den vorherigen Unterkapiteln werden zwei unterschiedliche Ansätze des Wissenstransfers vorgestellt, bei denen die Mitarbeitenden eine mehr oder weniger aktive Rolle übernehmen. Dabei sind sie in unterschiedlichem Maße gefordert, ihr Wissen und ihre Erfahrungen unter Nutzung technologischer Hilfsmittel aktiv zu dokumentieren und zu erklären. Dies erfordert jedoch, dass die Mitarbeitenden dazu in der Lage sind. Wie im Stand der Technik erklärt (vergleiche Kapitel 2.1 und 2.3), stellt die Umwandlung von implizitem in explizites Wissen eine komplexe Herausforderung dar, weil Mitarbeitende die Prinzipien und Kausalitäten, die ihren Handlungen zugrunde liegen, oft nicht präzise genug formulieren können. Darüber hinaus besteht die Gefahr, dass wesentliche

Informationen ausgelassen werden, da diese für erfahrene Mitarbeitende so selbstverständlich sind, dass sie nicht an eine explizite Erwähnung oder Dokumentation denken.

Vor diesem Hintergrund wurde das Konzept des *passiven Wissenstransfers* entwickelt. Dieses Konzept basiert auf der Idee, dass die Aktivitäten der Mitarbeitenden durch optische Systeme erfasst, digital dokumentiert und anschließend automatisiert analysiert werden, um Best Practices zu identifizieren, mit allen Mitarbeitenden zu teilen und somit das Wissen im Unternehmen langfristig zu sichern. Gemäß dem passiven Ansatz erfolgt der Wissenstransfer vollständig ohne aktive Beteiligung der Mitarbeitenden. Sie führen ihre Tätigkeiten wie gewohnt aus, ohne dass zusätzliche Handlungen von ihnen erforderlich sind. Entsprechend des grundlegenden Prinzips der *prozessparallelen Erfassung von Erfahrungswissen* ergibt sich die in Abbildung 5.17 dargestellte Systemarchitektur.

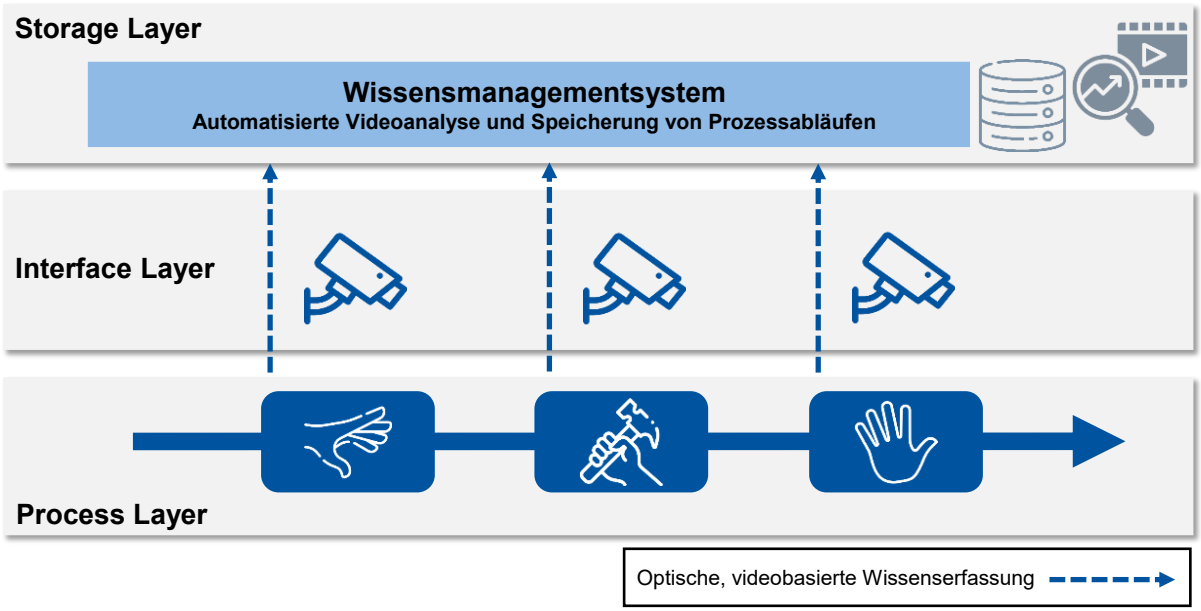



**Abbildung 5.17: Systemarchitektur des *passiven Wissenstransfers***

*System Architecture of the Passive Knowledge Transfer*

Die im Process Layer ablaufenden Aktivitäten der Mitarbeitenden werden prozessparallel über den Interface Layer mittels Videokameras erfasst und im Wissensmanagementsystem des Storage Layers automatisiert ausgewertet. Im Gegensatz zum *semiaktiven Wissenstransfer* (siehe Kapitel 5.2) wird die Kamera in diesem Ansatz nicht von den Mitarbeitenden geführt, sondern ist an einer fest installierten Position im Arbeitsraum platziert. Die Position ist so zu wählen, dass die Handlungen der Mitarbeitenden umfassend erfasst werden können. Abhängig von den betrachteten Prozessabläufen können unterschiedliche Technologien zur Analyse der erfassten Daten eingesetzt werden: Spezialisierte Methoden des maschinellen Lernens ermöglichen die Erkennung sowohl einzelner Objekte als auch von Kollisionen zwischen verschiedenen Objekten. Darüber hinaus können die Bewegungen von Menschen erfasst und in Form von Datenpunkten gespeichert werden.

In dieser Arbeit wird der Schwerpunkt auf die Erfassung von Handbewegungen der Mitarbeitenden gelegt, da diese trotz des zunehmenden Automatisierungsgrades weiterhin eine zentrale Rolle in zahlreichen Produktionsprozessen von KMU einnehmen.[3] Im Laufe der Zeit entwickeln Mitarbeitende dabei häufig individuelle Routinen, die effizienter oder ergonomischer durchführbar sind als die vom Unternehmen ursprünglich per Anleitung vorgegebenen Herangehensweisen. Diese Best Practices werden jedoch in der Regel nicht systematisch dokumentiert, was dazu führt, dass Unternehmen nicht in der Lage sind, die bestehenden Vorgaben entsprechend anzupassen und dieses wertvolle Wissen an alle Mitarbeitenden weiterzugeben. Gleichzeitig behindert der aktuelle Mangel an dokumentierten Best Practices die gezielte Automatisierung manueller Prozesse, da das notwendige Wissen über die optimalen Abläufe fehlt, um die gewohnte Qualität der Produkte und die Effizienz der automatisierten Prozesse sicherzustellen.

In den folgenden beiden Unterkapiteln werden zwei praxisnahe Konzepte zur Realisierung des *passiven Wissenstransfers* vorgestellt, die auf unterschiedliche Anwendungsfälle ausgerichtet sind: Das erste Konzept konzentriert sich auf manuelle Montagetätigkeiten und dient zur Erfassung der Best Practice Routinen zum Wissenstransfer zwischen Mitarbeitenden. Das zweite Konzept fokussiert die manuelle Bearbeitung von Bauteilen und adressiert den Transfer von implizitem Wissen in automatisierte Prozessabläufe und virtuelle Lernumgebungen.

### 5.3.1 Integration von Hand-Tracking und Process Mining zur Erfassung von Best Practices bei manuellen Montagen

*Integration of Hand Tracking and Process Mining to Capture Best Practices in Manual Assembly*

Im Gegensatz zu den im Stand der Technik vorgestellten Arbeiten im Bereich der Hand- und Gestenerkennung in der Montage (siehe Kapitel 2.3.2), die sich primär auf die Sicherstellung korrekter Montagereihenfolgen oder die einmalige Ableitung von Anweisungen für Montageprozesse konzentrieren, zielt das vorliegende Konzept auf die kontinuierliche Erfassung von Montageprozessen sowie die gleichzeitige Verbesserung der Arbeitsabläufe der Mitarbeitenden ab. Dabei werden die auf Erfahrung und Routine basierenden, individuell entwickelten Montageabläufe systematisch erfasst und hinsichtlich ihrer Effizienz und Ergonomie analysiert. Das Ziel besteht darin, bewährte Techniken, die bisher nur einem begrenzten Kreis von Mitarbeitenden bekannt sind, automatisiert zu identifizieren und diese im Rahmen eines unternehmensinternen Wissenstransfers allen Mitarbeitenden zugänglich zu machen.

Dazu wird erneut die Annahme zugrunde gelegt, dass sich die Aktivitäten der Mitarbeitenden im Process Layer in einzelne Aktionen differenzieren lassen. Im Kontext von Montagetätigkeiten kann eine naheliegende Differenzierung anhand der Bauteile der

[3] Diese Erkenntnis basiert auf Beobachtungen und Gesprächen, die im Rahmen des Forschungsprojekts *AKzentE4.0*, auf mehreren Messen sowie in diversen Industrieprojekten gemacht und geführt wurden.

Montagegruppe vorgenommen werden, sodass die Montage jedes Bauteils einen Prozessschritt darstellt. Dies entspricht der gängigen Praxis in vielen Montageanleitungen. In dieser Arbeit wird jedoch ein alternativer Differenzierungsansatz gewählt: Anstelle einer Differenzierung entlang der Bauteile erfolgt die Unterscheidung der Prozessschritte anhand der Handbewegungen, genauer gesagt der verwendeten Griffe der Mitarbeitenden. Diese präzise Erfassung der Aktionen ist entscheidend für die Identifikation und Bewertung individueller Routinen und stellt aufgrund ihres innovativen Charakters eine Abgrenzung zum Stand der Forschung dar. Die Idee besteht darin, unterschiedliche Griffe den verschiedenen Tätigkeiten im Montageprozess, wie dem Aufnehmen der Bauteile, der Werkzeuge oder der Verbindungselemente sowie den Fügeprozessen, zuzuordnen. Durch die Betrachtung der Griffe wird nicht nur eine Bewertung der Effizienz ermöglicht, sondern auch eine ergonomische Analyse individueller Routinen. Als weiterer innovativer Ansatz wird in dieser Arbeit untersucht, ob sich Methoden des Process Minings (siehe Kapitel 2.2.8) auf die Analyse der Aktivitäten der Mitarbeitenden anwenden lassen. Die Eigenschaften des Process Minings zur retrospektiven Analyse von Geschäftsprozessen durch die Auswertung von Event-Logs werden dazu im Folgenden adaptiert und auf die Analyse von Montageprozessen übertragen. Dieses Konzept sowie die zugehörige Umsetzung und Validierung (siehe Kapitel 6.3.1 und 7.1.3) wurden in reduzierter Form bereits in [DASS24] vorgestellt.

In Abbildung 5.18 ist das Konzept zur Integration von Hand-Tracking und Process Mining zur Erfassung von Best Practices bei manuellen Montagetätigkeiten schematisch visualisiert. Dabei sind sowohl der Ablauf der Erfassung der Aktivitäten der Mitarbeitenden als auch die daran anschließenden Analyseschritte dargestellt, deren jeweilige Funktionsweise im Folgenden näher erläutert wird.

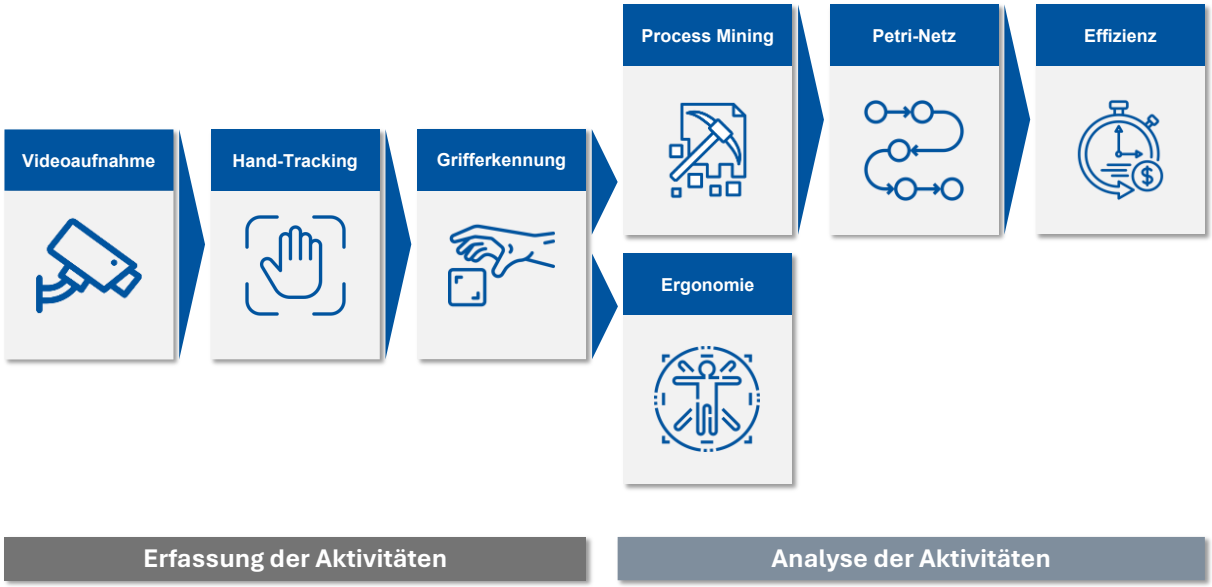



**Abbildung 5.18: Konzept zur Integration von Hand-Tracking und Process Mining zur Erfassung von Best Practices bei manuellen Montagen**

*Concept for Integrating Hand Tracking and Process Mining to Capture Best Practices in Manual Assembly*

### Erfassung der Aktivitäten – *Videoaufnahme*

Zur digitalen Erfassung der manuellen Arbeitsabläufe wird im ersten Schritt eine Videoaufnahme des Montageprozesses durchgeführt. Gegenstand der Aufnahme bildet dabei der Arbeitsraum, welcher den Bereich umfasst, der einer Person zur Ausführung der erforderlichen Tätigkeiten zum Erreichen eines vorgesehenen Arbeitsergebnisses zur Verfügung steht [DEUT16]. In dieser Arbeit wird die Systemgrenze des Arbeitsraumes so definiert, dass sie den betrachteten Arbeitsplatz (Montagetisch), alle im Montageprozess genutzte Arbeitsmittel (Werkzeuge, Verbindungselemente, Bauteile etc.) sowie die jeweils zur Aufbewahrung genutzten Ablageflächen, Behälter und Halterungen beinhaltet. Für eine ganzheitliche Erfassung der Handbewegungen der Mitarbeitenden wird die Kamera senkrecht oberhalb des Arbeitsraums positioniert, um eine durchgehende Sicht auf die Hände zu gewährleisten und visuelle Hindernisse in Form von Gegenständen oder Körperteilen zu minimieren. Dazu wird die Kamera so positioniert, dass der Rand des sichtbaren Bildausschnitts exakt der definierten Systemgrenze entspricht. In Abbildung 5.19 ist der auf diese Weise erfassbare Arbeitsraum am Beispiel eines manuellen Montageprozesses modellhaft dargestellt.

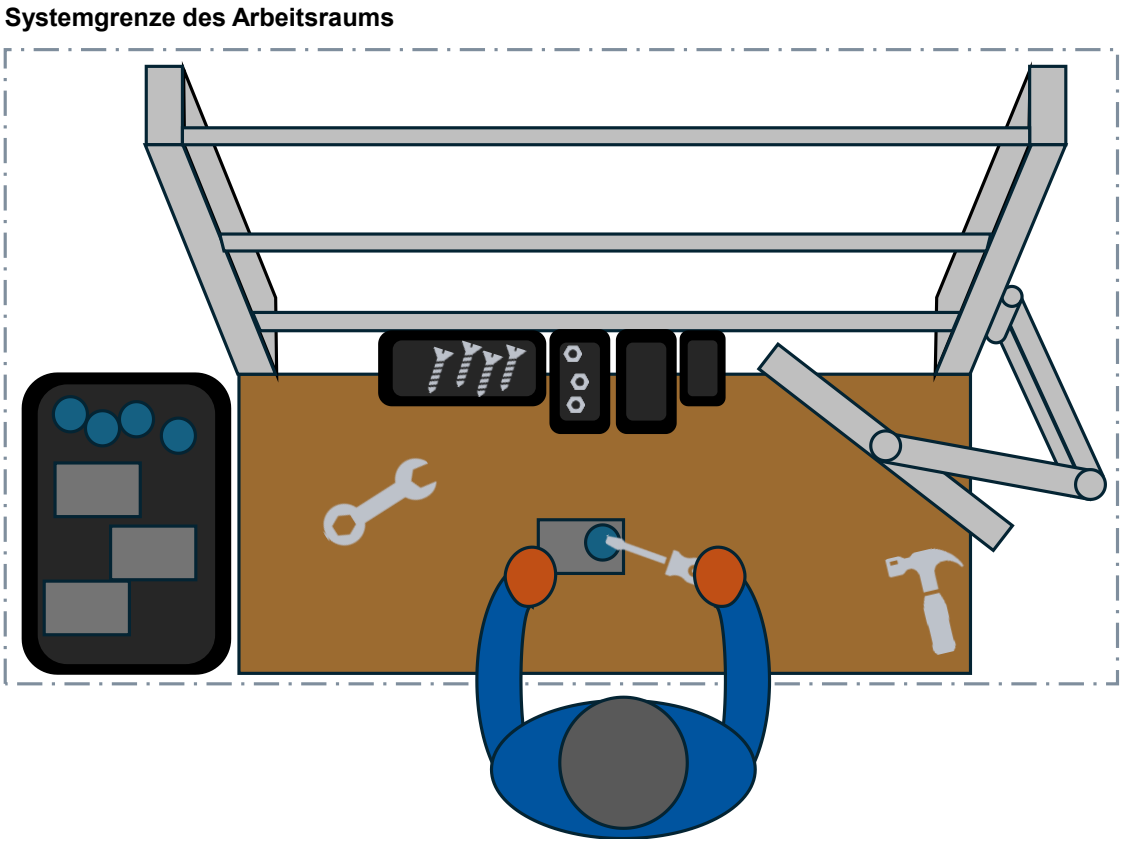



**Abbildung 5.19: Modellhafte Darstellung des erfassten Arbeitsraums**

*Exemplary Representation of the Captured Workspace*

Neben der Wahl des Bildausschnitts sind die Beleuchtung des Arbeitsraumes sowie die Auflösung und Bildrate der Kamera entscheidende Faktoren für die nachfolgende Auswertung der Videoaufnahme. Die Beleuchtung sollte ausreichend und gleichmäßig sein, um Schatten zu minimieren und alle Prozessdetails klar erkennbar zu machen.

Die Auflösung muss so gewählt werden, dass die Szene scharf und detailliert erfasst wird. Eine Standardauflösung von 1920 x 1080 Pixeln (Full HD) ist hierbei in der Regel ausreichend. Aufgrund der mitunter schnellen Wechsel zwischen verschiedenen Handbewegungen und Griffen wird eine Bildrate von mindestens 30 Bildern pro Sekunde angestrebt, um flüssige Bewegungsabläufe ausreichend detailliert aufzeichnen zu können.

Die bisher beschriebene Erfassung der Aktivitäten der Mitarbeitenden mittels einer Kamera birgt das Risiko, dass bestimmte Handbewegungen aufgrund von Verdeckungen nicht ausreichend präzise oder im schlimmsten Fall gar nicht erfasst werden können. Um diesem potenziellen Mangel entgegenzuwirken, wird das Konzept im Folgenden um die Möglichkeit der Nutzung einer zweiten Kamera erweitert. Die Verwendung von zwei Kameras ermöglicht grundsätzlich die Ermittlung von 3D-Koordinaten von Objekten per Stereo-Vision (siehe Kapitel 2.2.7). Vor dem Hintergrund dieser Arbeit, innovative Wissensmanagementsysteme und -methoden für den Shopfloor von KMU zu entwickeln, wurde die Notwendigkeit einer aufwendigen Kalibrierung für den Einsatz klassischer Stereo-Vision-Ansätze jedoch als impraktikabel für die meisten KMU bewertet. Deshalb wird ein abweichendes Vorgehen genutzt: Anstatt dem Prinzip der Stereo-Vision zu folgen, werden die beiden Kameras rechtwinklig in derselben Ebene zueinander positioniert, sodass sich ihre optischen Achsen im Mittelpunkt des Arbeitsraumes kreuzen. In Abbildung 5.20 sind die Positionen der beiden Kameras relativ zum betrachteten Arbeitsraum sowie die jeweils daraus resultierenden Koordinatensysteme schematisch dargestellt. Während die erste Kamera senkrecht über der Arbeitsplatte positioniert ist, wird die zweite Kamera aus der Perspektive der Mitarbeitenden hinter der Arbeitsfläche angebracht. Die in Abbildung 5.19 skizzierte Aufnahme des Arbeitsraumes wird damit um eine zweite Aufnahme ergänzt, welche die Mitarbeitenden und ihre Aktivitäten frontal erfasst.

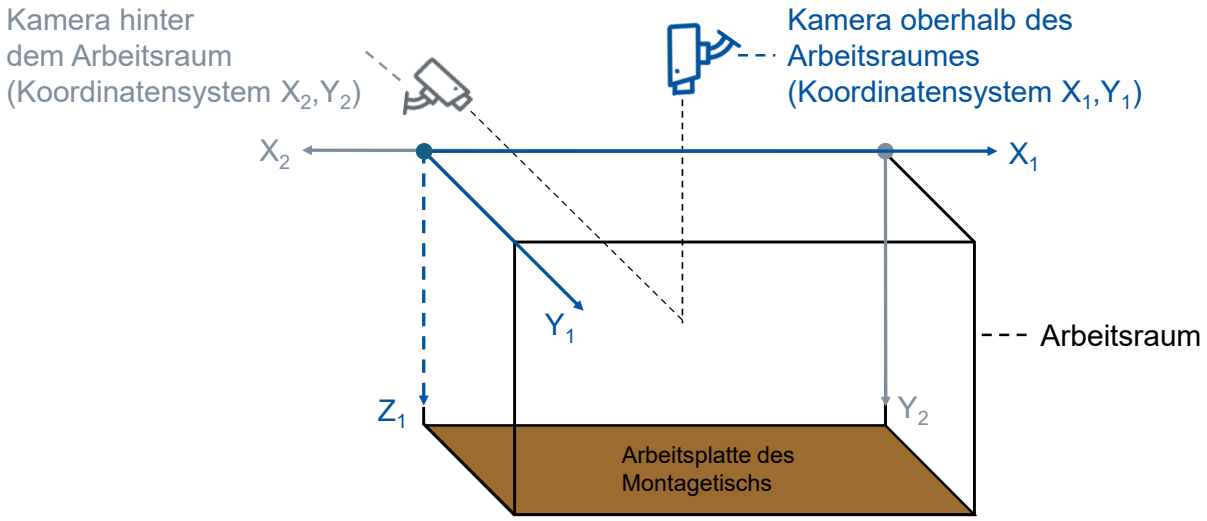



**Abbildung 5.20: Schematische Darstellung der Positionen der Kameras und ihrer zugehörigen Koordinatensysteme**

*Schematic Representation of the Positions of the Cameras and their Associated Coordinate Systems*

Zur Definition der beiden zweidimensionalen Koordinatensysteme wird der Ursprung des jeweiligen Koordinatensystems in der Bildebene der Kamera in die obere linke Ecke positioniert. Dabei ist erkennbar, dass die X-Achsen beider Kamera-Koordinatensysteme zwar kollinear verlaufen, jedoch aufgrund der unterschiedlichen Koordinatenursprünge entgegengesetzt orientiert sind. Zudem ist ersichtlich, dass die für eine dreidimensionale Betrachtung notwendige $Z_1$-Achse parallel zur $Y_2$-Achse verläuft. Ausgehend von dem Koordinatensystem der Kamera oberhalb des Arbeitsraumes können somit dreidimensionale Koordinaten berechnet werden, indem die $Y_2$-Koordinate als $Z_1$-Koordinate interpretiert wird. Die eingesetzten Kameras müssen hinsichtlich Auflösung und Bildrate übereinstimmen und die Aufnahmeprozesse müssen zeitlich synchronisiert erfolgen, um eine eindeutige und zuverlässige Erfassung der Koordinatenpunkte sicherzustellen.

Obwohl einige, vergleichsweise kostengünstige, kommerziell verfügbare Kamerasysteme über Tiefensensoren verfügen, wird in dieser Arbeit bewusst ausschließlich auf RGB-Bilder zurückgegriffen. Diese Entscheidung zielt darauf ab, die Übertragbarkeit der entwickelten Konzepte und Methoden auf ein breiteres Spektrum an Kamerasystemen zu gewährleisten und somit deren Eignung für eine spätere praktische Anwendung zu erhöhen. Im Zuge der Validierung der Umsetzung des Konzeptes in Kapitel 7.1.3 wird ein Vergleich zwischen den beiden Ansätzen (Eine Kamera vs. Zwei Kameras) durchgeführt, um den Mehrwert einer zusätzlichen Kamera zu bewerten.

***Hand-Tracking***

Die aufgezeichneten Videoaufnahmen des manuellen Prozessablaufs werden anschließend im Hinblick auf statische Gesten (siehe Kapitel 2.2.7) analysiert. Hierbei wird in jedem Bild der Videos die Position und Lage der einzelnen Finger beider Hände durch die Bestimmung der Hand-Landmarks detektiert. Dieses Konzept ist in Abbildung 5.21 exemplarisch anhand eines Bildes einer Videoaufnahme von Handbewegungen dargestellt.

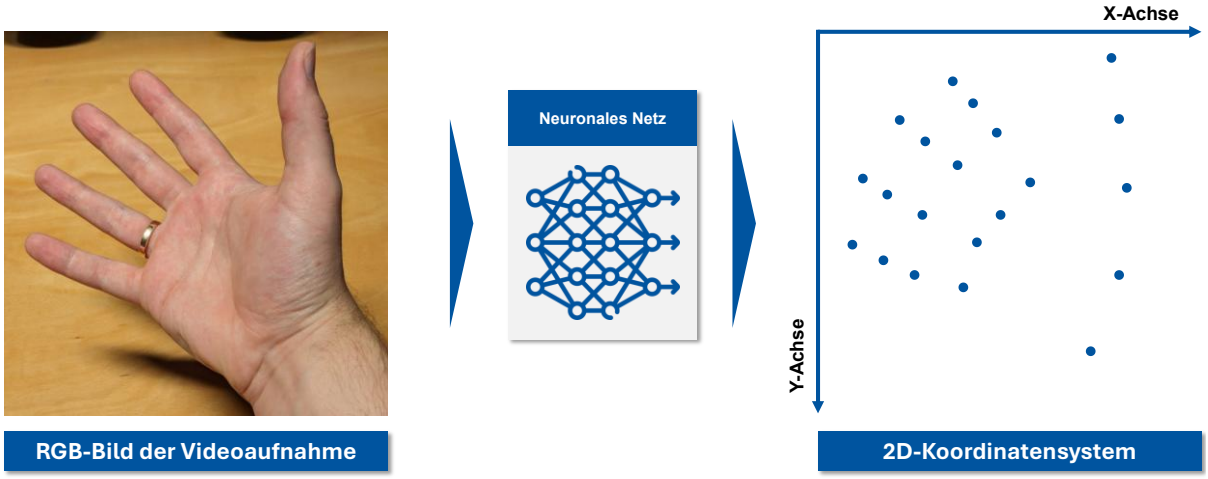



**Abbildung 5.21: Konzept der Ableitung von Koordinatenpunkten für die charakteristischen Punkte einer Hand auf Basis eines RGB-Bildes**

*Concept of Deriving Coordinate Points for the Characteristic Points of a Hand Based on an RGB Image*

Auf Grundlage der in Abbildung 5.20 gezeigten Beziehung zwischen den beiden Kamera-Koordinatensystemen können im nächsten Schritt die beiden Datensätze mit zweidimensionalen Koordinatenpunkten durch eine geeignete Koordinatentransformation zu einem Datensatz mit dreidimensionalen Koordinatenpunkten kombiniert werden. Abbildung 5.22 visualisiert diesen Vorgang in schematischer Form.

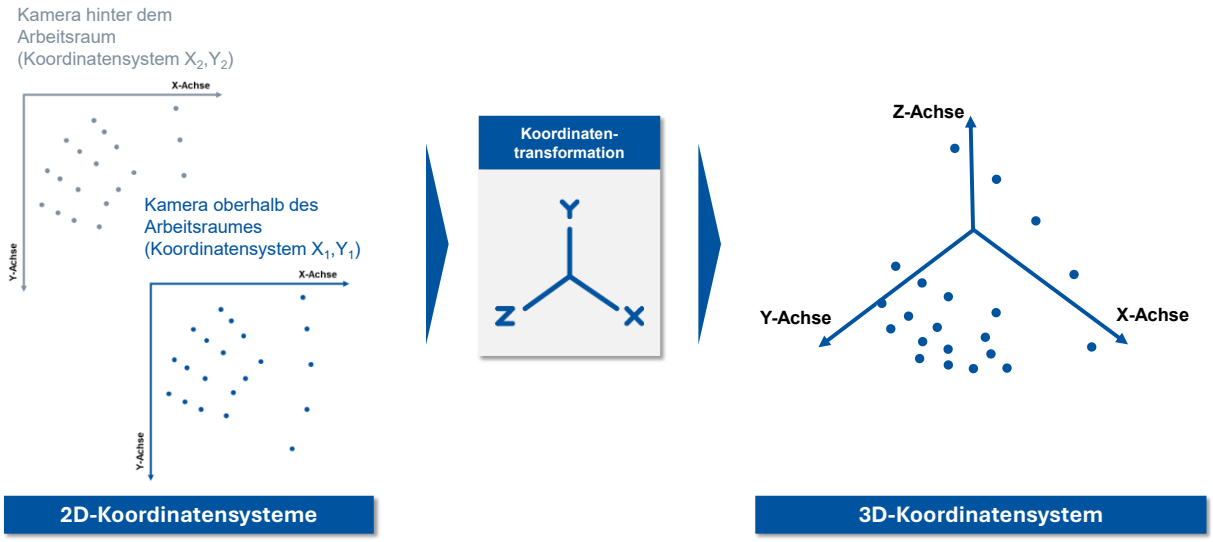



**Abbildung 5.22: Kombination der beiden zweidimensionalen Datensätze mittels Koordinatentransformation zur Bildung der dreidimensionalen Koordinatenpunkte**

*Combination of the Two Two-Dimensional Data Sets by Means of Coordinate Transformation to Form the Three-Dimensional Coordinate Points*

### *Grifferkennung*

Zur Differenzierung der Hand- und Fingerstellungen in unterschiedliche Griffarten werden die abgeleiteten dreidimensionalen Koordinatenpunkte der Hände genutzt, um durch die Anwendung mathematischer Beziehungen der linearen Algebra eine Vektordarstellung der einzelnen Fingersegmente zu erzeugen (siehe Abbildung 5.23, für die konkrete Umsetzung siehe Kapitel 6.3.1).

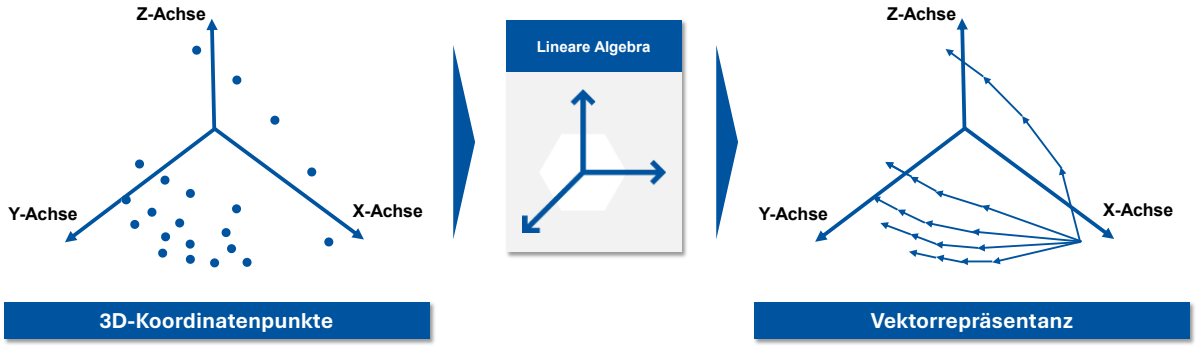



**Abbildung 5.23: Umwandlung der dreidimensionalen Koordinatenpunkte in die Vektorrepräsentanz einer Hand**

*Conversion of the Three-Dimensional Coordinate Points into the Vector Representation of a Hand*

Jeder der so erzeugten Vektoren beschreibt exakt die Position und Orientierung des jeweiligen Fingersegments. Auf Basis dieser Vektorrepräsentanz wird anschließend untersucht, welche Griffart vorliegt. Dazu wird ein analytisches Verfahren angewandt, bei dem die Winkelbeziehungen zwischen den Vektoren betrachtet werden. Für die Differenzierung der verschiedenen Griffarten werden zwei unterschiedliche Ansätze verfolgt: Beim ersten Ansatz werden Winkelbereiche vorab definiert, welche die Winkelbeziehungen der Vektoren für die unterschiedlichen Griffarten beschreiben. Die Vektorrepräsentanz wird anschließend per *regelbasiertem Ansatz* auf die Zugehörigkeit in einen der Winkelbereiche geprüft und der passenden Griffart zugeordnet. Der zweite Ansatz stützt sich ebenfalls auf die Winkelbeziehungen zwischen den Fingergliedern. Anstatt jedoch die ermittelten Werte in vordefinierte Winkelbereiche einzuordnen, wird die Machine-Learning-Methode der *Support Vector Machines* (SVM) verwendet. SVMs sind besonders gut geeignet, zwei Klassen von Merkmalen zu unterscheiden und die Eingabedaten einer dieser Klassen zuzuordnen [SURI12]. Da im betrachteten Fall zwischen mehreren Griffarten differenziert werden muss, wird die SVM an eine Mehrklassen-Spezifikation angepasst. Hierzu wird der sogenannte One-vs-All-Ansatz [DUAN07] verwendet: Für jede Klasse, in diesem Fall für jede Griffart, wird eine separate SVM trainiert, die als Klassifikator fungiert. Der auf Basis eines neuen Bildes berechnete Winkeldatensatz wird von allen Klassifikatoren ausgewertet und jeder Klassifikator vergibt eine Punktzahl, die angibt, wie gut der erfasste Griff in die jeweilige Klasse passt. Die höchste Punktzahl wird schließlich zur Zuweisung der wahrscheinlichsten Klasse (Griffart) verwendet.

Bei beiden Ansätzen muss berücksichtigt werden, dass eine Griffart nicht immer identisch ausgeführt wird. Diese Variabilität ergibt sich sowohl aus der Form und Beschaffenheit des gegriffenen Objekts als auch durch die individuellen physischen Merkmale der Hände der Mitarbeitenden. Zur Sicherstellung einer geeigneten Generalisierung, müssen die vordefinierten Winkelbereiche und die Trainingsdatensätze für die SVM diesen Aspekt adäquat abdecken.

Vor der Anwendung des hier beschriebenen Konzeptes muss eine Vorstudie durchgeführt werden, in der der Montageprozess der betrachteten Baugruppe per Video dokumentiert und anschließend einer teilautomatisierten Analyse unterzogen wird. Durch die manuelle Betrachtung des Videos werden die verwendeten Griffe identifiziert. Gleichzeitig werden die Winkelbeziehungen zwischen den Vektoren der Fingerglieder erfasst, um eine Ausgangslage für die Festlegung der zu den Griffen passenden Winkelbereiche zu schaffen. Nachdem dies erfolgt ist, werden Trainingsdaten für die SVM erstellt, indem Bilder von den zuvor passend zur betrachteten Montage definierten Griffklassen aufgenommen werden. Dazu werden die einzelnen Griffklassen gezielt mit leichten Variationen sowie in unterschiedlichen Positionierungen ausgeführt und jeweils einzeln per Foto oder Video erfasst.

Mit der Erkennung der verschiedenen Griffklassen ist der Prozess der Erfassung der Aktivitäten abgeschlossen, sodass im Anschluss die Analyse dieser erfolgt (vergleiche Abbildung 5.18). Im Folgenden wird zunächst die Analyse der Effizienz und anschließend die Analyse der Ergonomie dargestellt.

### Analyse der Aktivitäten – *Process Mining*

Das in dieser Arbeit entwickelte Konzept zur Effizienzanalyse adaptiert Methoden des Process Minings für die retrospektive Analyse von Montageprozessen. Dazu müssen zunächst Event Logs auf Grundlage der erkannten Griffe generiert werden. Hierzu wird der Montageprozess als eine Sequenz der von den Mitarbeitenden ausgeführten Griffbewegungen modelliert. Zur Erstellung eines geeigneten Formats für die Event Logs werden die erfassten Informationen wie folgt verarbeitet: Die Aktivitäten, d.h. die Griffe der linken und rechten Hand, werden separat und kontinuierlich aufgezeichnet. Bei der Verwendung einer Kamera mit einer Bildfrequenz von 30 Bildern pro Sekunde resultiert dies in der Detektion von 30 Griffen pro Hand und Sekunde. Da diese Erfassungsrate die tatsächliche Geschwindigkeit menschlicher Griffbewegungen übersteigt, wird die Datenmenge durch eine Aggregation reduziert. Dabei wird pro Sekunde der am häufigsten erfasste Griff extrahiert, um den dominanten Griff innerhalb des betrachteten Zeitintervalls zu definieren. Auf diese Weise kann aus den Videoaufnahmen bzw. Bildserien der Montageprozesse eine Sequenz der ausgeführten Griffe rekonstruiert werden, die jeweils mit Zeitinformationen versehen sind. Durch die Kombination der zum betrachteten Zeitpunkt erfassten Griffe der linken und rechten Hand wird der in diesem Moment durchgeführte Montageschritt identifiziert. Dieses Vorgehen ist in Abbildung 5.24 schematisch dargestellt.

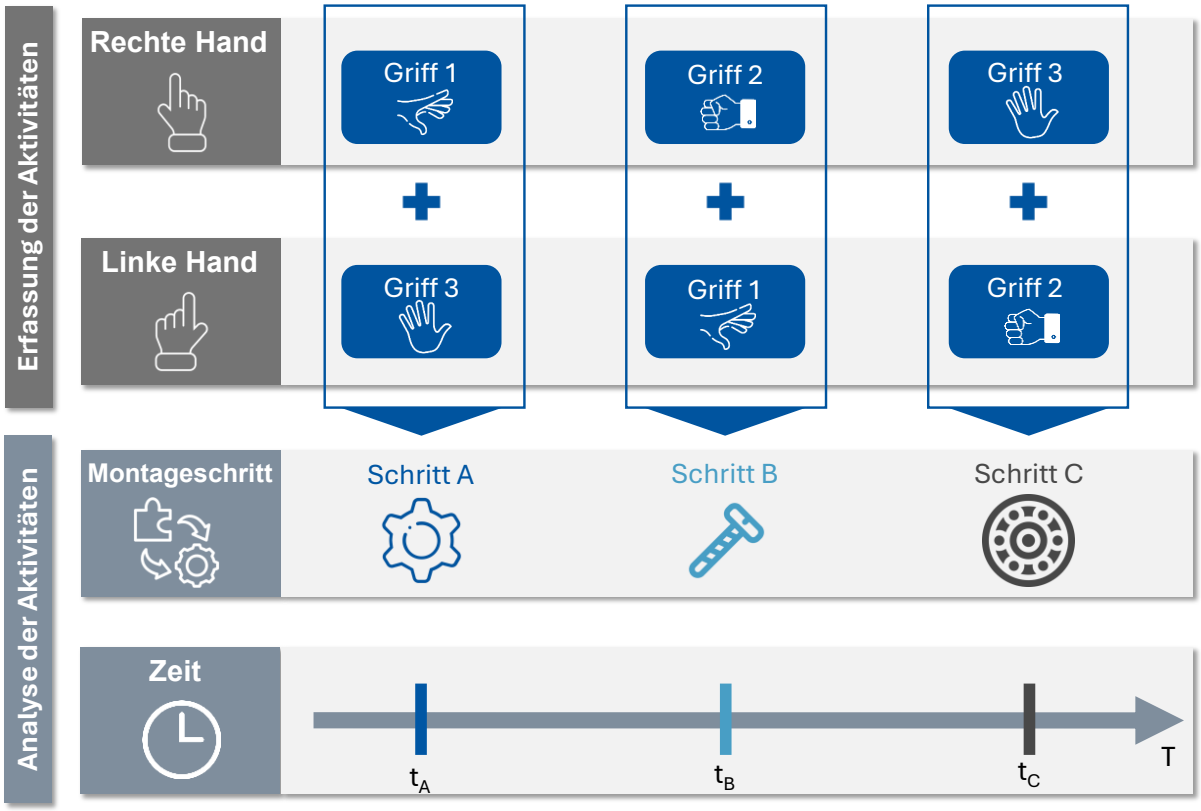



**Abbildung 5.24: Analyse der Griffkombinationen zur Bestimmung der durchgeführten Montageschritte zu bestimmten Zeitpunkten**

*Analysis of Grip Combinations to Determine the Assembly Steps Performed at Specific Points in Time*

Die Identifikation des zu einem bestimmten Zeitpunkt durchgeführten Montageschritts auf Grundlage der erfassten Griffkombinationen erfordert eine vorgelagerte Referenzaufnahme, in der sowohl die verwendeten Griffe als auch die damit gehandhabten Komponenten und die durchgeführten Handhabe- und Fügeprozesse explizit erfasst werden. Dies kann ebenfalls in der oben erläuterten Vorstudie erfolgen. Durch diese initiale Analyse wird die grundlegende Zuordnung von Griffkombinationen und Montageschritten geschaffen, anhand derer im späteren Prozess die beschriebene Hand-Tracking-basierte Identifikation der jeweiligen Montageschritte erfolgt. Die so erfassten Informationen – bestehend aus der Griffkombination, dem zugehörigen Montageschritt sowie dem Zeitpunkt der Aktivität und einem eindeutigen Identifikator – werden als Event betrachtet und im Event Log gespeichert.

***Petri-Netz und Effizienz***

Die erzeugten Event Logs des Montageprozesses können im weiteren Verlauf mit Techniken des Process Minings visualisiert und im Hinblick auf Effizienz der individuellen Routinen und Prozessabläufe analysiert werden. In dieser Arbeit wird die Analyse der Event Logs unter Einsatz eines Induktiven Miners (siehe Kapitel 2.2.8) durchgeführt, um eine Prozessentdeckung zu ermöglichen und den Prozess in einem blockstrukturierten Prozessmodell zu visualisieren. Die daraus resultierenden Petri-Netze sind geeignet, um die Reihenfolge der von den Mitarbeitenden ausgeführten Prozessschritte, die Dauer jedes einzelnen Schritts sowie die Gesamtdauer des Prozesses zu veranschaulichen und dadurch die Identifikation besonders effizienter Vorgehensweisen zu ermöglichen. Das Prinzip dieses Vorgehens ist in Abbildung 5.25 visualisiert.

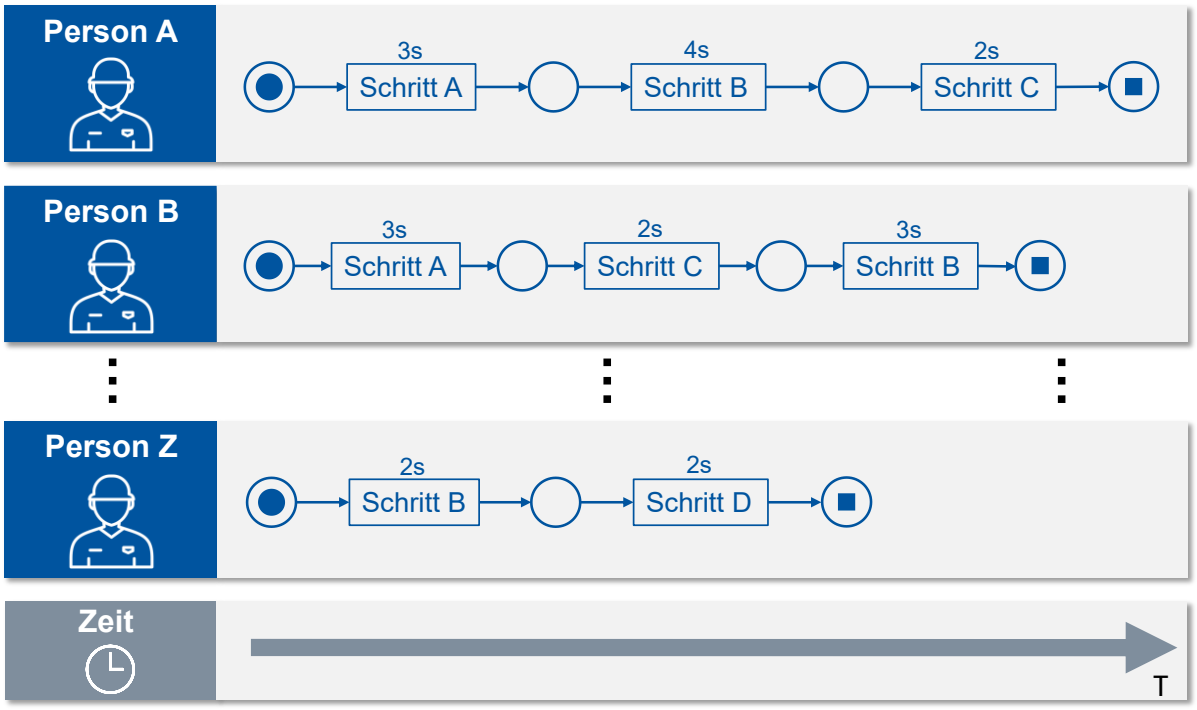



**Abbildung 5.25: Vergleich der individuellen Montageabläufe auf Basis von Petri-Netzen**

*Comparison of the Individual Assembly Sequences Based on Petri Nets*

Die einzelnen Aktionen der Montageabläufe der verschiedenen Mitarbeitenden (Person A, B, Z) werden als Transitionen eines Petri-Netzes dargestellt. Etwaige Unterschiede in den Vorgehensweisen (Vertauschen von Schritten, Weglassen von Schritten, Hinzufügen neuer Schritte) werden somit als Veränderungen in der Reihenfolge der Transitionen sichtbar. Jeder Montageschritt bzw. jede Transition ist zusätzlich mit der dafür benötigten Bearbeitungszeit versehen, wodurch eine Analyse der ggf. durch veränderte Abläufe erzielten Zeiteinsparungen ermöglicht wird. Zur Bewertung der Gesamteffizienz der Montageprozesse können die Bearbeitungszeiten der einzelnen Schritte zu einer Gesamtprozesszeit aggregiert werden.

***Ergonomie***

Parallel zur Analyse der Effizienz der Montageabläufe kann auch eine ergonomische Bewertung erfolgen: Mithilfe von Gefährdungsbeurteilungen für manuelle Tätigkeiten, wie sie beispielsweise gemäß dem Leitfaden der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin [KLUß22] durchgeführt werden können, lassen sich die einzelnen Montageschritte hinsichtlich der erforderlichen Kräfte für die Ausführung der Griffkombinationen bewerten. Für jeden individuellen Montageprozess kann anschließend ermittelt werden, über welchen Zeitraum welche Belastungen auf Hände und Finger wirken. Die Sensibilisierung und Verbreitung von Wissen über ergonomische Prozesse fördern die Motivation der Mitarbeitenden und tragen zur Reduzierung krankheitsbedingter Ausfallzeiten bei, was letztlich sowohl den Mitarbeitenden als auch dem Unternehmen zugutekommt. Belastungen der Hände durch manuelle Tätigkeiten sind in diesem Zusammenhang ein ernstzunehmendes Problem. In einer repräsentativen Studie der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin im Jahr 2018 gaben rund 17% der befragten Arbeitnehmenden an, Schmerzen in den Händen zu haben [LÜCK19].

Beide Aspekte können auch kombiniert analysiert werden, um beispielsweise einen Kompromiss zwischen Effizienz und Ergonomie zu finden. Hierzu kann eine geeignete Gewichtung der relevanten Parameter (wie Anzahl der Montageschritte, Gesamtprozesszeit und Kategorien der aufgewendeten Kräfte) vorgenommen werden, um eine differenzierte Gesamtbewertung zu ermöglichen. Nach der erfolgreichen Durchführung des so beschriebenen Konzepts zur Integration von Hand-Tracking und Process Mining zur Erfassung von Best Practices bei manuellen Montagen können die vorhandenen Arbeitsanweisungen entsprechend den neugewonnenen Erkenntnissen angepasst werden, um das Wissen und die Erfahrung einzelner Personen dem gesamten Unternehmen bereitzustellen.

### 5.3.2 Einsatz von Hand-Tracking zur Überführung manueller Bearbeitungsverfahren von Bauteilen in roboterbasierte Prozesse und virtuelle Lernumgebungen

*Use of Hand Tracking to Transfer Manual Processing Methods of Components to Robot-based Processes and Virtual Learning Environments*

Neben der manuellen Montage werden auf dem Shopfloor von KMU auch verschiedene manuelle Bearbeitungsverfahren an Bauteilen durchgeführt, bei denen die individuelle Erfahrung der Mitarbeitenden eine entscheidende Rolle für die Effizienz der Prozessdurchführung sowie die Qualität der Bauteile spielt. Insbesondere in den Fertigungshauptgruppen Fügen, Trennen und Beschichten gibt es Verfahren, bei denen das Erreichen der angestrebten Qualitätsstandards eine spezifische Bewegungsausführung erfordert. Bei Prozessen wie Schleifen, Bürsten oder Lackieren werden häufig spezifische optische Oberflächeneigenschaften (z.B. Schliffbilder oder Muster im Lack) angestrebt, die sich nur durch die präzise Ausführung der jeweiligen Techniken beim Auf- bzw. Abtragen des Materials erreichen lassen. Das Erlernen und Beherrschen dieser Techniken erfordert in der Regel intensive Übung, um die optimalen Bewegungsabläufe zu perfektionieren. Selbst für erfahrene Mitarbeitende gestaltet es sich jedoch oft schwierig, die durch Training erworbenen Bewegungsmuster sowie die dabei zu beachtenden Besonderheiten zu verbalisieren, um einen traditionellen Wissenstransfer zu ermöglichen. So sind Mitarbeitende beispielsweise häufig nicht in der Lage, den Radius kreisender Bewegungen präzise zu beschreiben. Gleichzeitig erschwert dies auch die Automatisierung solcher Bearbeitungsverfahren, da für die Übertragung menschlicher Bewegungsabläufe auf Roboter eine exakte Quantifizierung der Bewegungen erforderlich ist.

Vor diesem Hintergrund wird die im vorherigen Unterkapitel vorgestellte Methodik zur videobasierten Erfassung der Aktivitäten von Mitarbeitenden bei der manuellen Montage adaptiert, um die Bewegungsabläufe von Mitarbeitenden bei der Ausführung manueller Bearbeitungsverfahren zu quantifizieren. Die dabei entstehenden mathematischen Beschreibungen der Bewegungsabläufe können anschließend als Input für die roboterbasierte Automatisierung sowie zur Erstellung von virtuellen Lernumgebungen genutzt werden. Das Konzept der Kombination von Computer-Vision-Systemen und KI-Applikationen zur Erfassung von Aktivitäten von Mitarbeitenden und Nutzung dieser Informationen zur Entwicklung innovativer Trainingskonzepte in virtuellen Lernumgebungen wurde bereits in [PETR23] grundlegend thematisiert und wird im Folgenden ausführlicher im Kontext dieser Arbeit beschrieben.

In Abbildung 5.26 ist das Konzept zur Erfassung von Bewegungsabläufen für die Übertragung in automatisierte Prozesse und virtuelle Lernumgebungen visualisiert. Der Ablauf der Erfassung der Aktivitäten der Mitarbeitenden und die daran anschließende Verwertung werden im Folgenden schrittweise vorgestellt.

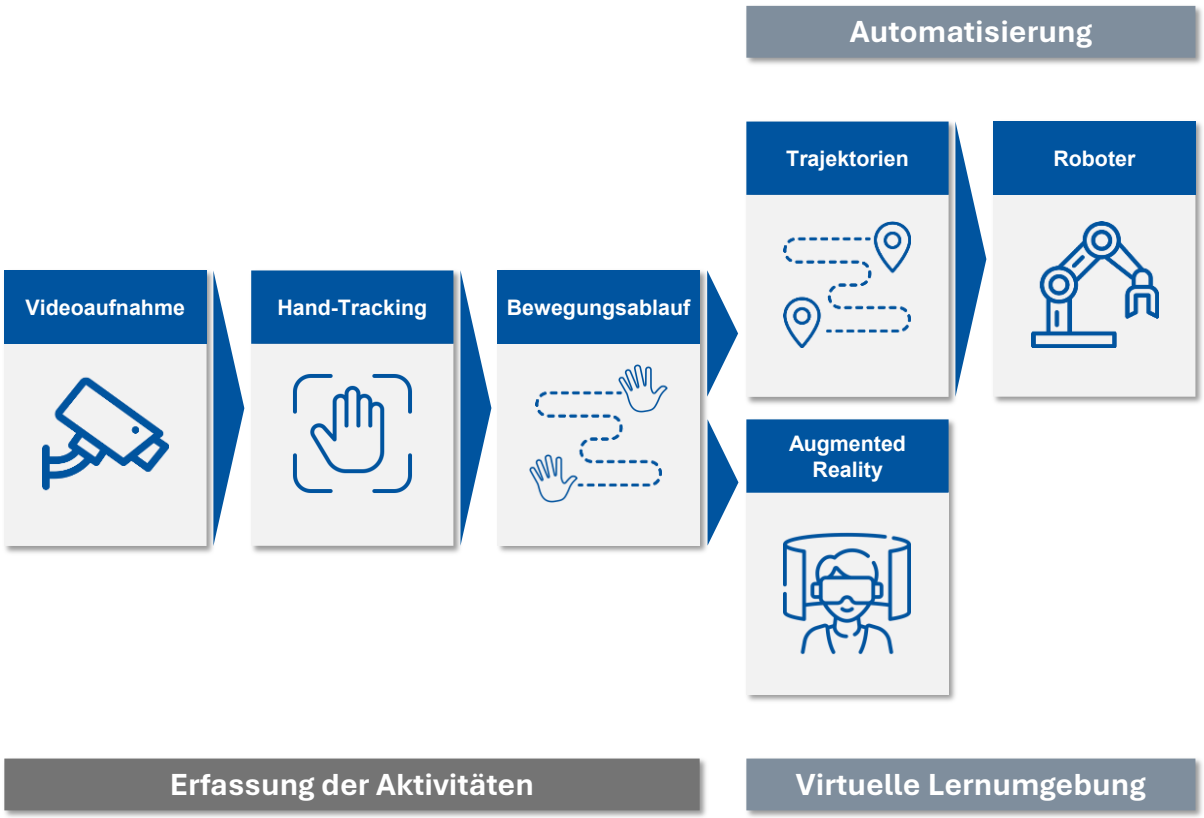



**Abbildung 5.26: Konzept zum Einsatz von Hand-Tracking zur Überführung manueller Bearbeitungsverfahren von Bauteilen in roboterbasierte Prozesse und virtuelle Lernumgebungen**

*Concept for Using Hand Tracking to Transfer Manual Processing Methods of Components into Robot-Based Processes and Virtual Learning Environments*

### Erfassung der Aktivitäten – *Videoaufnahme und Hand-Tracking*

Die Erstellung der Videoaufnahmen des Arbeitsraums, in dem die manuellen Bearbeitungsverfahren ausgeführt werden, sowie die anschließende Auswertung dieser Aufnahmen mittels Hand-Tracking erfolgen analog zu dem im vorherigen Unterkapitel 5.3.1 erläuterten Vorgehen.

Im weiteren Verlauf dieser Arbeit wird ein manueller Schleifprozesses zur Erzeugung eines definierten Schliffbildes auf einer Funktionsfläche als Referenz für manuelle Bearbeitungsprozesse genutzt.[4] Bei diesem Anwendungsfall wird das Bauteil so auf dem Arbeitstisch positioniert, dass die zu bearbeitende Oberfläche senkrecht nach oben zeigt. Die auszuführenden Schleifarbeiten werden dementsprechend in einer Ebene parallel zur Tischplatte vorgenommen. Da in diesem Prozess keine vertikale Bewegung relativ zur Oberfläche erforderlich ist, reicht die Erfassung der zweidimensionalen

[4] Im Rahmen des Forschungsprojekts *AKzentE4.0* wurde der Anwendungsfall des manuellen Schleifens von Oberflächen zum Aufbringen eines definierten Schliffbildes auf eine Funktionsfläche gemeinsam mit einem Partnerunternehmen aus dem Projektkonsortium als relevant eingestuft. Eine nähere Erläuterung des Anwendungsfalls erfolgt im Verlauf der Validierung der Referenzimplementierung des Konzeptes in Kapitel 7.1.3.

Hand-Landmarks in der Bearbeitungsebene aus, um alle relevanten Bewegungen präzise zu dokumentieren. Daher wird in diesem Konzept ausschließlich die oberhalb des Arbeitsraums installierte Kamera zur Erfassung der zweidimensionalen Koordinaten verwendet.

Während bei der Analyse manueller Montagen statische Gesten (Griffe) erfasst werden, liegt der Fokus in diesem Fall auf der Erfassung dynamischer Gesten (siehe Kapitel 2.2.7). Das Ausführen der *Videoaufnahme* und der Analyse der resultierenden Bilder mittels *Hand-Tracking* liefert bei einer Kamera mit einer Bildfrequenz von 30 Bildern pro Sekunde auch in diesem Fall 30 Koordinatenpunkte pro Hand-Landmark und pro Hand innerhalb einer Sekunde als Ergebnis. Anders als bei den statischen Gesten ist eine Aggregation der Daten auf Sekundenebene in diesem Fall nicht zweckmäßig, da der Mensch selbst innerhalb von Bruchteilen einer Sekunde signifikante Handbewegungen im Bereich mehrerer Zentimeter ausführen kann.

***Bewegungsablauf***

Zur Analyse des Bewegungsablaufes der Hände auf Basis der erfassten Koordinatenpunkte der Hand-Landmarks können zwei Aspekte betrachtet werden: Zunächst lässt sich eine Trajektorie bilden, welche die Bewegungsbahn einer Hand bzw. der einzelnen markanten Punkte auf einer Hand darstellt. Dazu werden die über die Zeit erfassten Koordinatenpunkte der Hand-Landmarks in ein Koordinatensystem eingetragen und entsprechend ihrer zeitlichen Reihenfolge miteinander verbunden. Durch die Bildung dieser Trajektorie werden Bewegungsmuster sichtbar und können anschließend verglichen und analysiert werden. Dieses Konzept ist in Abbildung 5.27 anhand des Bewegungsablaufes der Spitze eines Zeigefingers schematisch visualisiert.

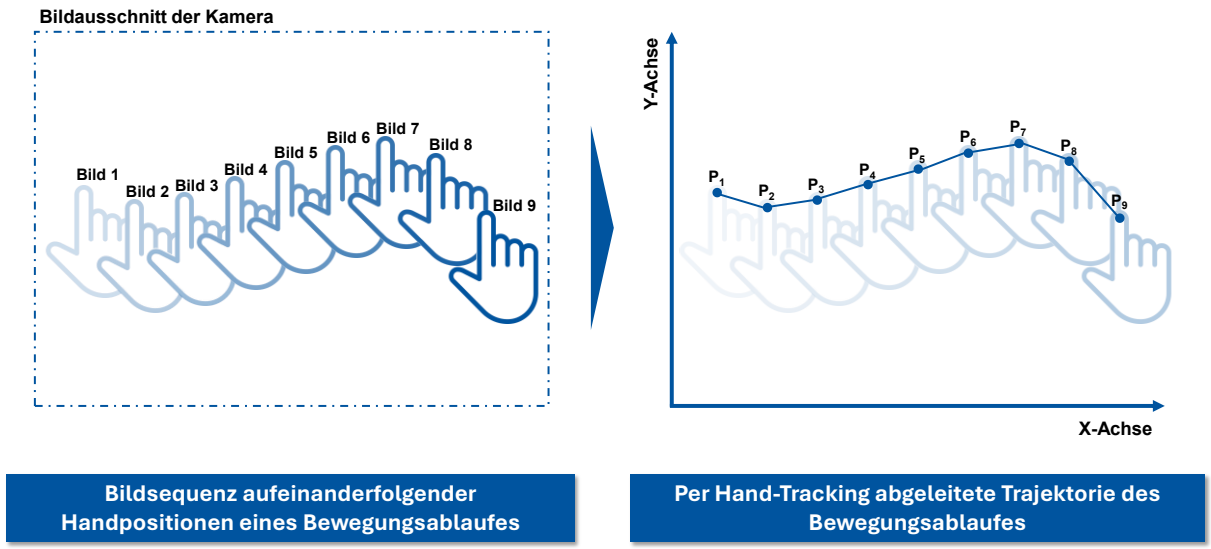



**Abbildung 5.27: Konzept der Visualisierung eines per Hand-Tracking erfassten Bewegungsablaufes in Form einer Trajektorie**

*Concept of Visualizing a Movement Sequence Recorded by Hand Tracking in the Form of a Trajectory*

Die bei der Aufnahme des Bewegungsablaufs verwendete Kamera muss konstant in einer fest definierten Position verbleiben, um sicherzustellen, dass die Bewegung der Hand relativ zum Kamerakoordinatensystem erfasst wird. Die Erhöhung der Bildfrequenz der verwendeten Kamera führt zu einer präziseren Annäherung der erzeugten Trajektorie an die tatsächlich ausgeführte Bewegung.

Neben der Visualisierung des Bewegungsablaufs kann auch der Verlauf der Bewegungsgeschwindigkeit bestimmt werden. Hierzu wird die zurückgelegte Strecke zwischen zwei in aufeinanderfolgenden Bildern detektierten Koordinatenpunkten einer Hand-Landmark durch die Zeitdifferenz zwischen den Bildern dividiert. Die auf diese Weise berechneten Geschwindigkeitswerte können anschließend als Funktion der Zeit in ein geeignetes Koordinatensystem eingetragen werden, um das Geschwindigkeitsprofil des betrachteten Bewegungsablaufs zu erhalten. Die kombinierte Analyse und Visualisierung der Bewegungsabläufe und der zugehörigen Geschwindigkeitsprofile liefert wertvolle Erkenntnisse über spezifische Bewegungseigenschaften wie etwa Richtungsänderungen, zurückgelegte Weglängen sowie Bewegungsmuster.

**Automatisierung – *Trajektorien und Roboter***

Manuelle Bearbeitungsverfahren zur Erreichung spezifischer optischer Oberflächeneigenschaften sind häufig schwer zu automatisieren, da sowohl die präzise Erfassung der zu bearbeitenden Bereiche (einschließlich Position, Orientierung und ggf. Art der Oberflächenfehler) als auch die Auswahl einer geeigneten Bearbeitungsstrategie (z.B. Werkzeugauswahl und -führung sowie Anpassung der Prozessparameter) umfangreiches prozesstechnisches Wissen erfordern [ALT24]. Vor diesem Hintergrund kann das zuvor beschriebene Konzept zur Erfassung und Analyse manueller Bewegungsabläufe die roboterbasierte Automatisierung gezielt unterstützen: Durch die quantitative Erfassung der Bewegungen erfahrener Mitarbeitender bei der Ausführung des manuellen Prozesses wird eine solide Grundlage für die Planung der Trajektorien eines roboterbasierten Prozesses geschaffen. Die erfassten Koordinatenpunkte müssen dabei in das Koordinatensystem des Roboters transformiert werden, wofür eine Kalibrierung zwischen dem Hand-Tracking-Koordinatensystem (d.h. dem Kamera-Koordinatensystem) und dem Roboterkoordinatensystem erforderlich ist. Abhängig von der Kontinuität und Genauigkeit der aufgezeichneten Bewegungen und der dazugehörigen Geschwindigkeitsprofile kann eine mathematische Glättung notwendig sein, um die Trajektorie des Roboters präzise und gleichmäßig zu gestalten. Zudem müssen bei der Transformation die Gelenkstellungen und Freiheitsgrade des Roboters berücksichtigt werden.

Die Transformation manueller Bearbeitungsprozesse in roboterbasierte, automatisierte Abläufe wird in dieser Arbeit nicht durch eine Umsetzung und Validierung vertieft, da sie im Fokus paralleler Forschungsarbeiten der am Lehrstuhl für Werkzeugmaschinen des WZLs der RWTH Aachen University liegt. Diese Arbeiten werden ebenfalls im Projekt AKzentE4.0 durchgeführt und voraussichtlich 2026 erstmals veröffentlicht.

**Virtuelle Lernumgebungen**

Aufgrund der evolutionären Entwicklung des Menschen wird schon seit Urzeiten das Prinzip des *Imitation Learning* angewandt, bei dem Menschen durch die Beobachtung anderer deren Aktivitäten verstehen und nachahmen [BAND71; RIZZ04]. Für das Erlernen und die Perfektionierung von Bewegungsabläufen ist jedoch insbesondere das wiederholte Üben von zentraler Bedeutung [STE-12]. Um diesen Aspekt des wiederholten Übens auf Basis beobachteter Aktivitäten in Form einer virtuellen Lernumgebung zu ermöglichen, kann das aus Rennspielen bekannte Konzept des *Ghost Racers* adaptiert werden. Dieses System beschreibt eine Methode zum Training von Fahrern in einer simulierten Umgebung, bei der die Aktivitäten eines vorausgegangenen Fahrers für den Lernenden visualisiert werden [MONC96]. In vielen Rennspielen wird dies durch transparente, „geisterhafte“ Darstellungen der Fahrzeuge realisiert, sodass der lernende Fahrer die Bewegungen anderer beobachten, vergleichen und nachahmen kann. Die digitale Erfassung der Aktivitäten erfahrener Mitarbeitender per Hand-Tracking ermöglicht es, die Bewegungsabläufe auf ein virtuelles Modell von Händen zu übertragen. Die aufgezeichneten Bewegungen und zugehörigen Geschwindigkeitsprofile dienen als Input für die Modellierung des virtuellen Modells. Mithilfe eines geeigneten AR-Devices können Lernende die Bewegungsabläufe der Experten in der realen Umgebung visualisiert betrachten. Entsprechend des Prinzips des *Ghost Racers* werden die Hände als *Ghost Hands* visualisiert. So wird eine direkte Vergleichbarkeit der eigenen Handlungen mit der Idealbewegung ermöglicht, wodurch Abweichungen von der optimalen Bewegungslinie unmittelbar erkannt und angepasst werden können. In Abbildung 5.28 ist das Ergebnis dieses Konzepts beispielhaft dargestellt.

Die Idee der *Ghost Hands* entstand erst zu einem späten Zeitpunkt während der Bearbeitung dieser Dissertation. Aufgrund der mit der Umsetzung verbundenen Komplexität konnte deshalb keine vollständige Implementierung erfolgen.

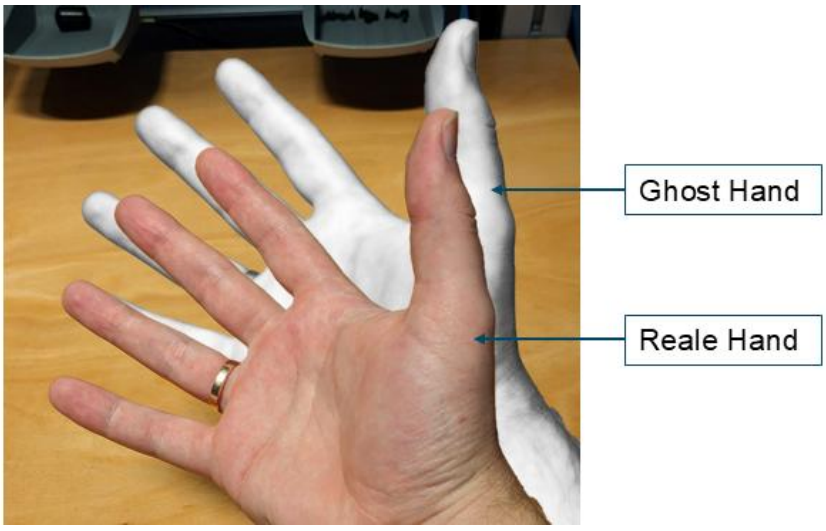



**Abbildung 5.28: Visualisierung der Ghost Hands aus der Sicht eines lernenden Mitarbeitenden**

*Visualization of Ghost Hands from the Perspective of a Learning Employee*

# 6 Referenzimplementierung der Konzepte

*Reference Implementation of the Concepts*

In diesem Kapitel werden die Referenzimplementierungen der drei in Kapitel 5 beschriebenen Konzepte für den Wissenstransfer detailliert vorgestellt. Sie bilden die Grundlage für die anschließende technische Validierung und Evaluation, die in Kapitel 7 im Hinblick auf die in Kapitel 4 abgeleiteten Anforderungen erfolgt.

Für die Umsetzung der Referenzimplementierungen wurde eine On-Premise-Lösung gewählt. Diese Entscheidung erfolgte aus dem wissenschaftlichen Anspruch, eine konsistente und kontrollierte Umgebung zu schaffen, um die notwendige Reproduzierbarkeit der Untersuchungsergebnisse sicherzustellen. Die Spezifikationen der hierfür eingesetzten Workstation sind in Tabelle 6.1 aufgeführt.

**Tabelle 6.1: Hardwarekonfiguration der für die Umsetzung verwendeten Workstation**

*Hardware Configuration of the Workstation Used for Realization*

| Komponente | Details |
|---|---|
| **CPU** | Intel Xeon W-2245 (8 Kerne, 16 Threads, 3.9 GHz Basis, 4.7 GHz Turbo, 16.5 MB Cache, 155 W TDP, DDR4-2933, ECC-Unterstützung) |
| **GPU** | 2x NVIDIA RTX A4000 (je 16 GB GDDR6 VRAM, insgesamt 32 GB VRAM) |
| **RAM** | 256 GB DDR4-2933 (8x 32 GB RDIMM ECC) |
| **Speicher** | 1x 512 MB NVMe SSD (PCIe 4.0, High-Speed-Storage)<br>3x 1 TB NVMe SSD (PCIe 4.0, High-Speed-Storage) |
| **Betriebssystem** | Windows 10 Pro, Version 22H2 |

Die Hardwarekonfiguration der eingesetzten Workstation stellt in gewissem Maße einen limitierenden Faktor bei der Umsetzung der Konzepte dar, da einige KI-Applikationen sehr ressourcenintensiv sind. Sehr rechenintensive Anwendungen und Modelle können aufgrund der resultierenden hohen Rechenzeiten für den praktischen Einsatz nicht berücksichtigt werden. Diese Einschränkung spiegelt jedoch die Realität vieler KMU wider und stellt im Kontext der Forschungsfragen dieser Arbeit die methodisch sinnvollste Herangehensweise dar. Der potenzielle Einfluss der Hardwarekonfiguration auf die jeweilige Referenzimplementierung wird, sofern relevant, in den folgenden Unterkapiteln entsprechend erläutert.

## 6.1 Umsetzung des aktiven Wissenstransfers

*Implementation of the Active Knowledge Transfer*

In den folgenden Kapiteln werden die Referenzimplementierungen der beiden in Kapitel 5.1 vorgestellten Ansätze zur Chatbot-basierten Gestaltung des aktiven Wissenstransfers präsentiert.

### 6.1.1 Implementierung der Funktionen des Chatbots basierend auf dem Rasa-Framework

*Implementation of the Chatbot Functions Based on the Rasa framework*

Die Referenzimplementierung des Chatbots auf Basis des Rasa-Frameworks erfolgte entlang der konkreten Anforderung aus der Industrie, die Interaktion mit Fehlerdatenbanken von Maschinen zu optimieren, um deren Abfrage und Befüllung durch Mitarbeitende intuitiver und effizienter zu gestalten. Insgesamt ergaben sich für diesen Anwendungsfall die folgenden Randbedingungen:

- Die Interaktion mit dem Chatbot soll über ein Sprachinterface erfolgen.
- Die Maschinenfehler werden in einer SQL-Datenbank als Text gespeichert.
- Der Chatbot soll sowohl die Abfrage von Maschinenfehlern, deren Ursachen und Lösungsvorschläge als auch das Hinzufügen neuer Einträge ermöglichen.

Aufbauend auf dem in Kapitel 5.1 vorgestellten Konzept (vgl. Abbildung 5.4 und Abbildung 5.7) visualisiert Abbildung 6.1 die detaillierte Ausgestaltung der einzelnen Systembestandteile zur Erfassung und Ausgabe von Erfahrungswissen per Sprachinterface. Die einzelnen Module der Implementierung werden nachfolgend vorgestellt.

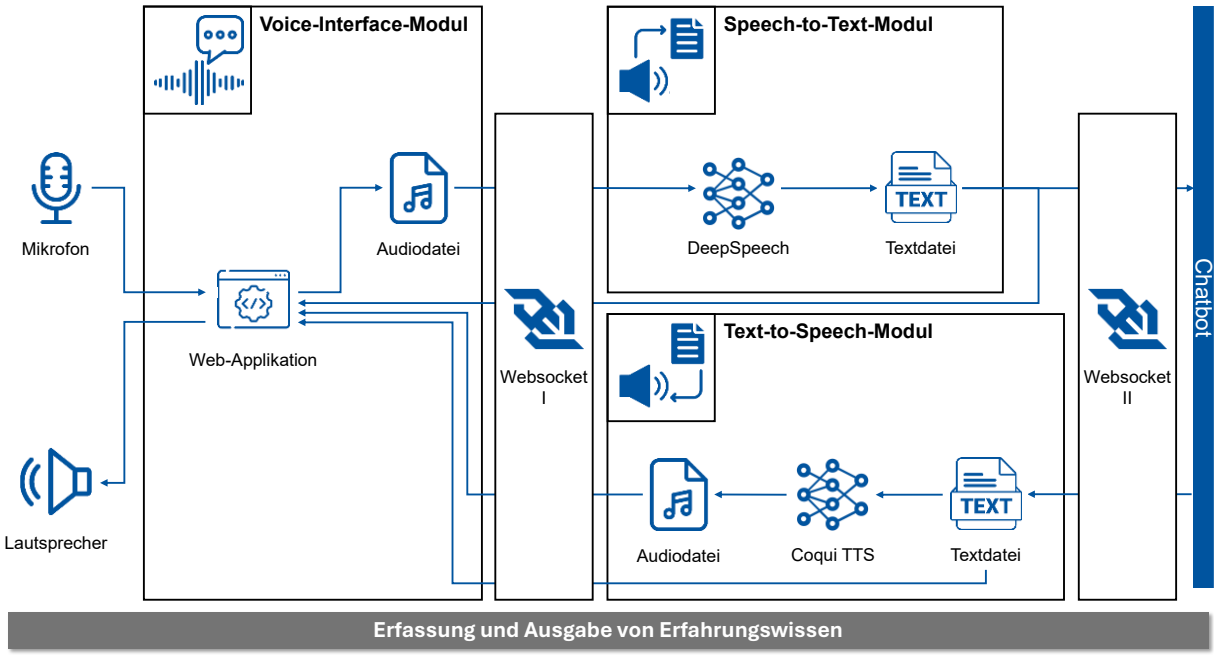



**Abbildung 6.1: Systemarchitektur zur Implementierung der Erfassung und Ausgabe von Erfahrungswissen unter Verwendung des Rasa-Frameworks**

*System Architecture for Implementing the Capture and Output of Experiential Knowledge when Using the Rasa Framework*

Die Interaktion des Benutzers mit dem Chatbot erfolgt über das ***Voice-Interface-Modul***, das durch einen Python-basierten HTTP-Server realisiert wird. Dieser hostet unter Port 8065 eine lokale Web-Applikation, welche zur Aufnahme und Ausgabe von Audiodateien dient. Der Aufbau der Web-Applikation ist in Abbildung 6.2 zu sehen.

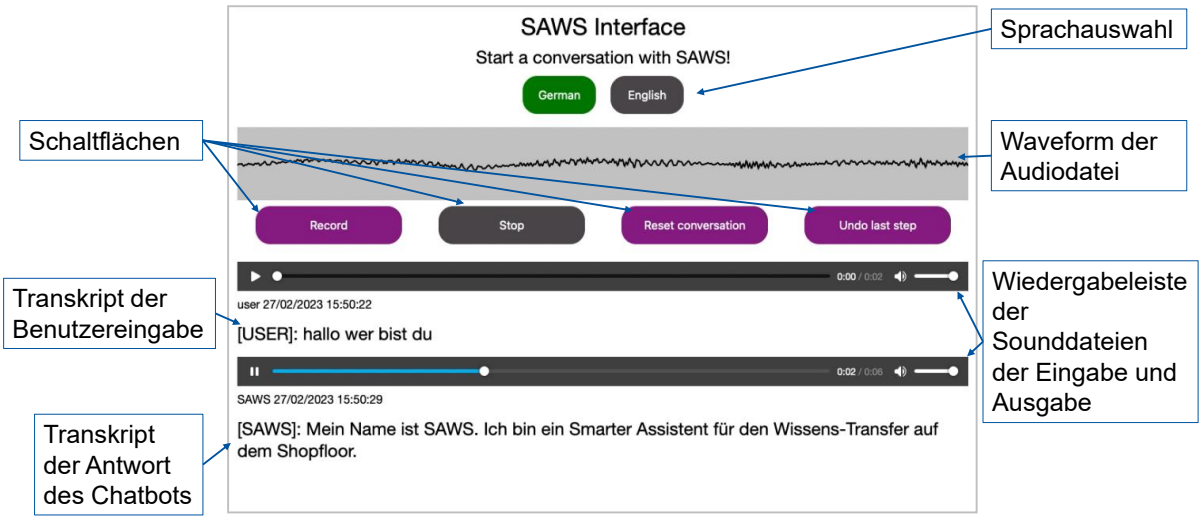



**Abbildung 6.2: Grafisches Benutzerinterface der Web-Applikation des Rasa-basierten Chatbots**

*Graphical User Interface of the Web Application of the Rasa-Based Chatbot*

Zur Kommunikation mit dem Chatbot kann der Benutzer zunächst zwischen den Sprachen Deutsch (German) und Englisch (English) wählen. Mit der Schaltfläche *Record* wird die Aufnahme einer Audiodatei gestartet und durch Betätigen der Schaltfläche *Stop* beendet. Über die Schaltflächen *Reset Conversation* und *Undo last Step* können der gesamte Chatverlauf oder die letzte Spracheingabe zurückgesetzt werden. Unterhalb dieser Bedienelemente wird der Chatverlauf visualisiert, wobei sowohl die Benutzereingaben als auch die Antworten des Chatbots als Transkripte dargestellt werden. Jedes Transkript wird zusätzlich durch eine Wiedergabeleiste ergänzt, die die zugehörige Audiodatei abspielt. Die Antworten des Chatbots werden initial automatisch wiedergegeben und können anschließend über die Wiedergabeleiste beliebig oft erneut abgespielt werden. Ergänzend dazu wird die aktuell wiedergegebene Audiodatei in einer Waveform-Darstellung visualisiert, die die Amplitude des Audiosignals im Zeitverlauf zeigt. Für die Aufnahme der Anfragen sowie zur Ausgabe der Antworten des Chatbots wurde ein Jabra Evolve 40 MS Stereo Headset verwendet.

Die über die Web-Applikation aufgezeichnete Audiodatei der Benutzereingabe wird vom Voice-Interface-Modul über eine Websocket-Verbindung (Websocket I) zur Vorverarbeitung an das ***Speech-to-Text-Modul*** übertragen. Die Verwendung einer Websocket-Verbindung ist für die Implementierung der Funktionen eines Chatbots besonders geeignet, da sie eine bidirektionale Kommunikation ermöglicht. Durch die permanente Verbindung zwischen dem Websocket-Server und den angebundenen Clients wird ein nahezu verzögerungsfreier Datenaustausch gewährleistet.

Um die Benutzereingabe effizient weiterzuverarbeiten, wird die Audiodatei, die im WAV-Format mit einer Abtastrate von 16 kHz vorliegt, innerhalb des Voice-Interface-Moduls in Base64 kodiert und in ein JSON-Objekt eingebettet. Dieses JSON-Objekt wird anschließend über den Websocket I an das Speech-to-Text-Modul weitergeleitet. Dort wird der Base64-String aus dem JSON-Objekt extrahiert und in das ursprüngliche WAV-Format zurückkonvertiert. Die WAV-Datei wird mithilfe von DeepSpeech (siehe

Kapitel 2.2.6) transkribiert. Ist im Benutzerinterface Englisch als Eingabesprache ausgewählt, wird der transkribierte Text zusätzlich ins Deutsche übersetzt, da die nachfolgenden Rasa-Funktionen im Rahmen der Referenzimplementierung ausschließlich auf Deutsch trainiert wurden. Der transkribierte und vorverarbeitete Text wird anschließend über eine zweite Websocket-Verbindung (Websocket II) an den Chatbot weitergeleitet. Parallel dazu wird der transkribierte Text über Websocket I zurück an das Voice-Interface-Modul gesendet, wo er im Benutzerinterface angezeigt wird.

Die vom Chatbot über Websocket II übermittelten Antworten werden im ***Text-to-Speech-Modul*** für die Darstellung in der Web-Applikation aufbereitet. Falls die Sprache Englisch für den Dialog mit dem Chatbot ausgewählt wurde, erfolgt zunächst eine Übersetzung des Antworttextes ins Englische. Anschließend wird der Text mit Funktionen der Open-Source-Bibliothek *CoquiTTS* in eine Audiodatei im WAV-Format umgewandelt. Über Websocket I werden sowohl der Antworttext vom Chatbot als auch die generierte Audiodatei an das Voice-Interface-Modul zurückgesendet und entsprechend im Benutzerinterface im Chatverlauf dargestellt (siehe Abbildung 6.2).

**Implementierung der automatisierten Analyse der Eingabe und Ableitung von Aktionen sowie der Wissensablage**

Die Prozessschritte des in Kapitel 5.1 vorgestellten Konzeptes zur automatisierten Analyse der Eingaben und Ableitung entsprechender Aktionen sowie für die Erstellung der erforderlichen Wissensablage (vgl. Abbildung 5.4 und Abbildung 5.5) werden in Form der in Abbildung 6.3 visualisierten Systemarchitektur implementiert.

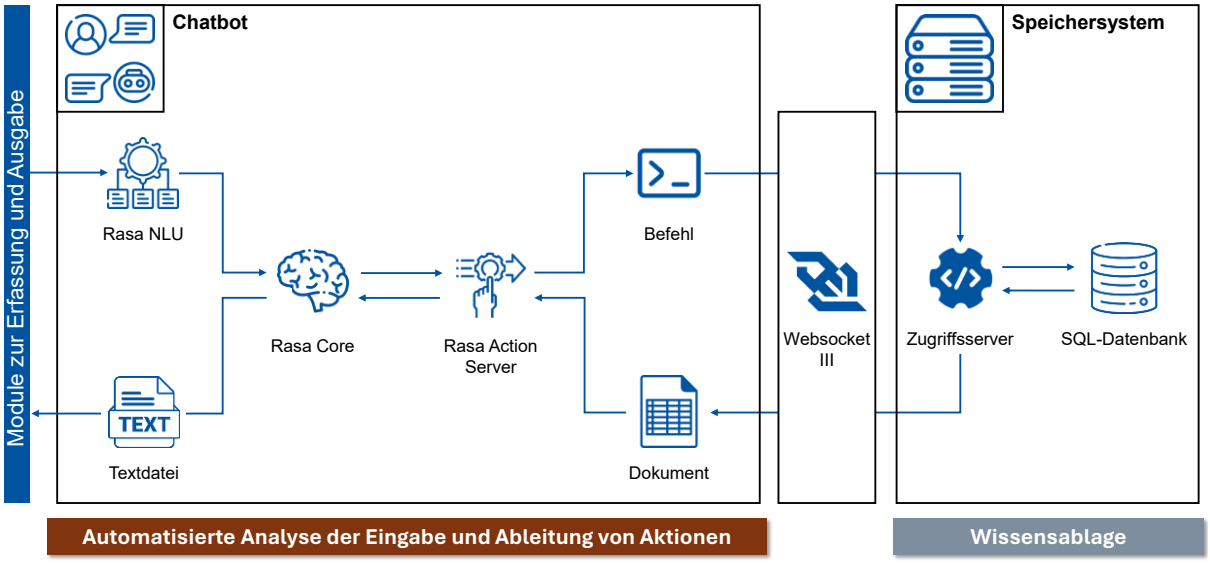



**Abbildung 6.3: Systemarchitektur zur Implementierung der automatisierten Analyse der Eingaben und Ableitung von Aktionen sowie der Wissensablage unter Verwendung des Rasa-Frameworks**

*System Architecture for Implementing the Automated Analysis of Inputs and Derivation of Actions as well as Knowledge Storage when Using the Rasa Framework*

Dabei werden die Strukturen Rasa NLU, Rasa Core und Rasa Action Server für die Realisierung des Chatbots genutzt. Die vom Speech-To-Text-Modul über Websocket II (siehe Abbildung 6.1) an den *Chatbot* übermittelte Textdatei wird zunächst vom Rasa NLU analysiert, um die Absicht des Benutzers sowie relevante Informationen im Text zu identifizieren. Hierzu werden in der zentralen Konfigurationsdatei „nlu.yml“ des Rasa NLUs spezifische Intents definiert, die mögliche Benutzereingaben im Kontext der Interaktion mit Fehlerdatenbanken von Maschinen beschreiben. Anhand dieser Beispiele wird das Modell trainiert, um Benutzereingaben korrekt den entsprechenden Intents zuzuordnen. In Abbildung 6.4 ist ein Auszug aus der verwendeten Config-Datei zu sehen.

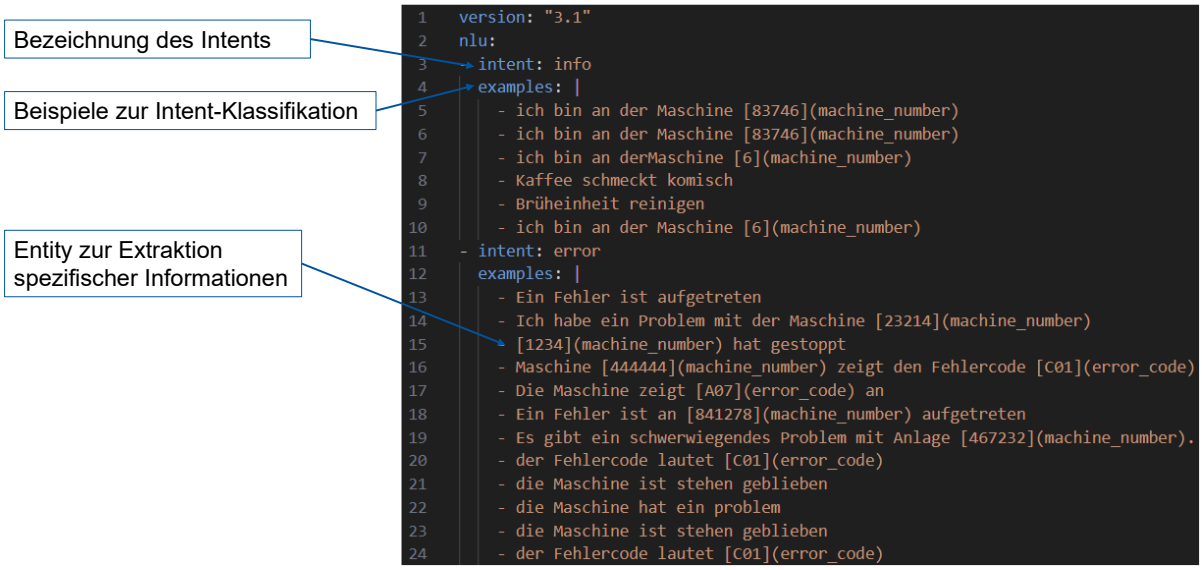



**Abbildung 6.4: Auszug aus der zentralen Config-Datei des Rasa NLUs**

*Extract from the Central Config File of the Rasa NLUs*

Der erkannte Intent sowie gegebenenfalls extrahierte Entities werden an den Rasa Core übergeben, der diese Informationen verwendet, um basierend auf den vordefinierten Stories den nächsten Dialogschritt mit dem Benutzer zu bestimmen. Dazu wird die Datei „stories.yml“ (siehe Abbildung 6.5) nach einer passenden Story durchsucht, die sowohl den erkannten Intent als auch den bisherigen Dialogverlauf berücksichtigt.

Die Aktionen des Chatbots werden durch die in der jeweiligen Story definierte Abfolge von Intents und Aktionen (Actions) gesteuert. Dabei gibt der Chatbot entweder eine entsprechende Antwort auf die Anfrage des Benutzers oder löst eine Custom Action aus. Für die Ausführung einer Custom Action sendet der Rasa Core eine Anforderung an den Rasa Action Server. Die Custom Actions sind als Python-Funktionen in der Datei „actions.py“ hinterlegt. In der Referenzimplementierung beinhalten die Custom Actions unterschiedliche Methoden zur Generierung und Verarbeitung von SQL-Abfragen, die sowohl das Lesen als auch das Schreiben von Daten in einer SQL-Datenbank ermöglichen.

Die aus den erkannten Intents und Entities extrahierten Informationen, beispielsweise zur angefragten Maschine oder zum aufgetretenen Fehler, werden innerhalb der entsprechenden Custom Action in eine SQL-Abfrage umgewandelt. Diese SQL-Query

wird über Websocket III an den Java-basierten Zugriffsserver des Speichersystems übermittelt. Der Zugriffsserver verarbeitet die Abfrage, indem er entweder die angeforderten Informationen aus der angebundenen SQL-Datenbank abruft oder ein entsprechendes Update in der Datenbank durchführt. Die Struktur der entwickelten SQL-Datenbank ist in Abbildung D.1 im Anhang visualisiert.

Das Ergebnis der SQL-Query wird anschließend als String-basiertes Dokument zurück an den Rasa Action Server gegeben, der innerhalb der Ausführung der Custom Action auf eine Antwort des Zugriffsservers wartet. Die empfangene Antwort wird an den Rasa Core übergeben, der den Dialogzustand basierend auf der definierten Story aktualisiert und entweder die nächste Custom Action ausführt oder eine Antwort an den Benutzer formuliert. Im Falle einer Antwort an den Benutzer wird diese als Text formuliert und über Websocket II sowie das Text-to-Speech-Modul an das Voice-Interface-Modul zurückgegeben. Dort wird die Antwort sowohl in der Web-Applikation visualisiert als auch als Audio ausgegeben (siehe Abbildung 6.1). Dieser Ablauf schließt den Interaktionszyklus mit dem Rasa-basierten Chatbot ab.




**Abbildung 6.5: Auszug aus der Stories-Datei des Rasa Cores**

*Extract of the Stories File of the Rasa Cores*

### 6.1.2 Implementierung der Funktionen des Chatbots basierend auf Large Language Modellen

*Implementing of the Chatbot Functions Based on Large Language Models*

Die Implementierung des Chatbots unter Verwendung von LLM geschah zu einem späteren Zeitpunkt während der Bearbeitung der vorliegenden Dissertation. Im Vergleich zur Variante mit Rasa ergaben sich deshalb veränderte Anforderungen aus der Praxis: In den am Projekt AKzentE4.0 beteiligten KMU besteht die Herausforderung, dass umfangreiches Unternehmenswissen in Form unstrukturierter Textdokumente digital vorliegt. Die schiere Menge an Informationen erschwert es jedoch erheblich, relevante Inhalte für spezifische Aufgabenstellungen bei Bedarf manuell zu identifizieren. Daher lag der Fokus bei der Implementierung mittels LLM nicht auf der Bereitstellung

einer Möglichkeit zur Interaktion mit strukturierten Datenbanken, sondern auf der Entwicklung einer effizienten Lösung zur semantischen Suche innerhalb einer textbasierten, diffusen Wissensbasis. Um diesen Anforderungen gerecht zu werden, wurde anstelle einer SQL-Datenbank eine Kombination aus einem Vector Store (siehe Kapitel 2.2.2) und einer Knowledge Base auf Basis von JSON-Objekten verwendet. Darüber hinaus bestand in diesem Fall kein Wunsch nach einem Sprachinterface, sodass die Benutzerschnittstelle zur Ein- und Ausgabe von Wissen ausschließlich über Tastatur und Bildschirm realisiert wurde.

Analog zum Vorgehen bei der Implementierung mit Rasa wird auch für die Realisierung des Chatbots unter Verwendung von LLM das in Kapitel 5.1 vorgestellte Konzept (vgl. Abbildung 5.4 und Abbildung 5.7) in eine Systemarchitektur überführt. Diese ist in Abbildung 6.6 dargestellt. Die einzelnen Bestandteile werden im weiteren Verlauf erläutert.

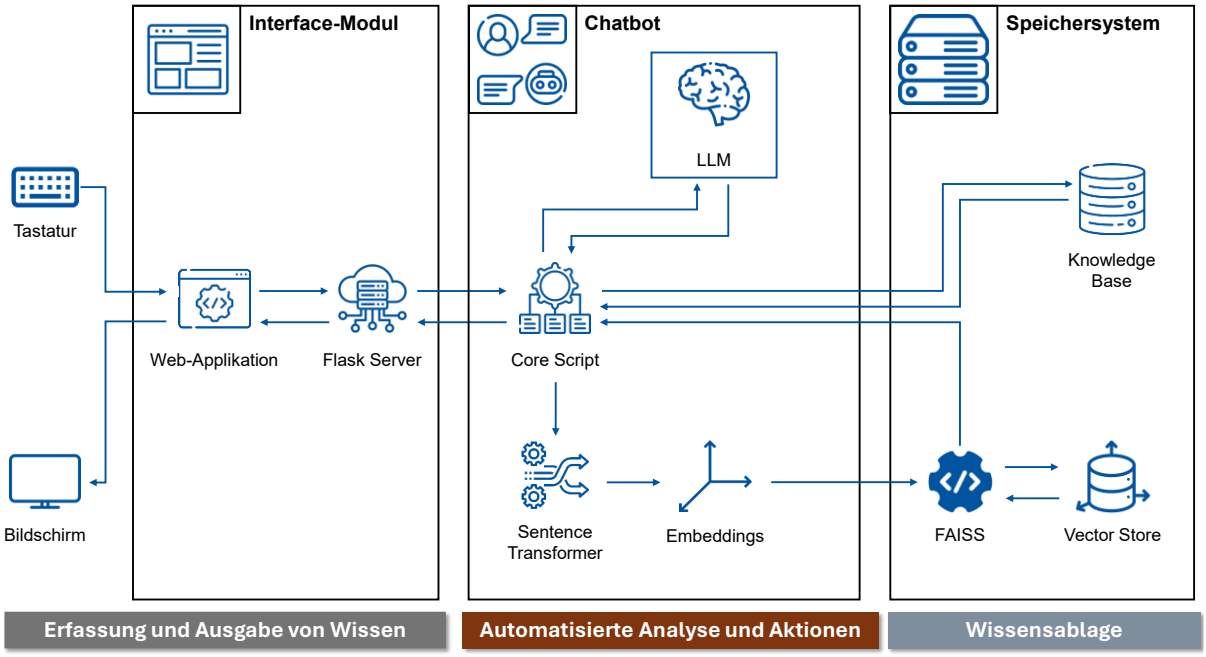



**Abbildung 6.6: Systemarchitektur zur Implementierung der Chatbot-basierten Erfassung und Speicherung von Wissen unter Verwendung von LLM**

*System Architecture for the Implementation of Chatbot-Based Knowledge Capture and Storage Using LLM*

### Implementierung der Erfassung und Ausgabe von Wissen – *Interface-Modul*

Die Interaktion zwischen Benutzer und Chatbot läuft textbasiert über eine Web-Applikation ab, die mithilfe von HTML und JavaScript entwickelt wurde und über die URL „http://localhost:8000“ im Browser lokal auf der Workstation aufgerufen werden kann. Der Aufbau der Web-Applikation ist in Abbildung 6.7 gezeigt. Das Benutzerinterface umfasst ein Chatfenster mit einem zugehörigen Eingabefeld zur textbasierten Kommunikation mit dem Chatbot sowie eine Funktionsfläche zur Auswahl und zum Upload von PDF-Dokumenten. Die Upload-Funktion dient der initialen Generierung einer Da-

tenbasis, auf die anschließend über den Chatbot zugegriffen werden kann. Diese Datenbasis simuliert in grundlegender Weise die textbasierte, unstrukturierte Wissensbasis, wie sie in den befragten KMU vorzufinden ist.

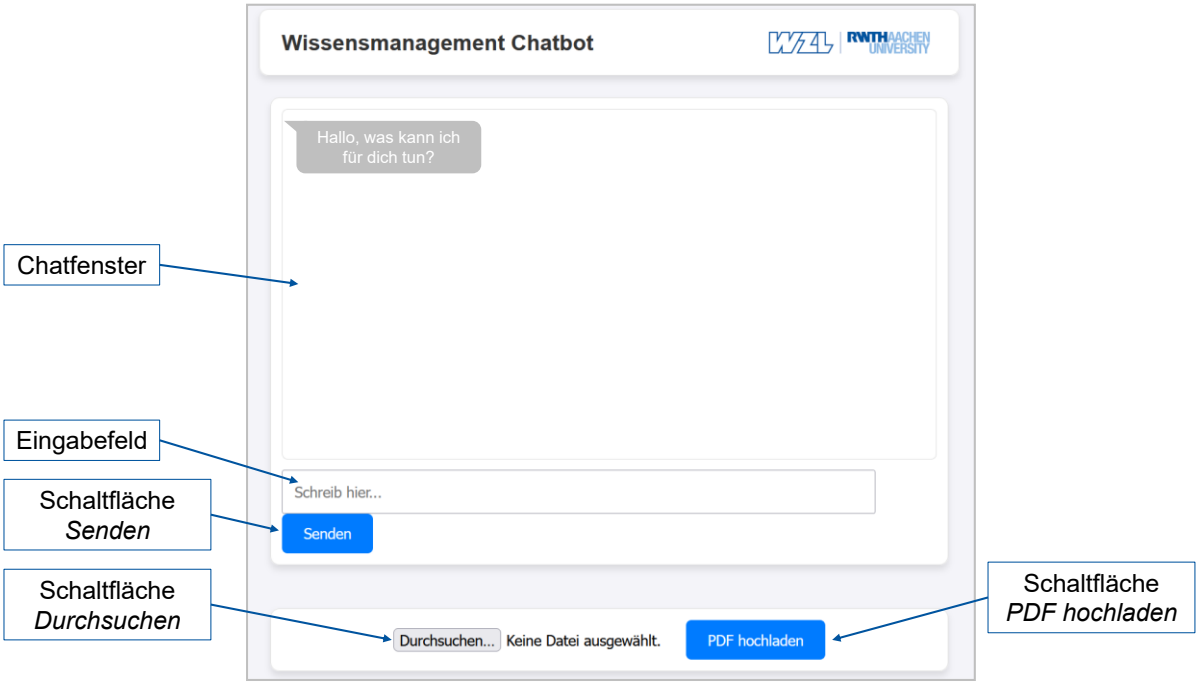



**Abbildung 6.7: Grafisches Benutzerinterface der Web-Applikation des LLM-basierten Chatbots**

*Graphical User Interface of the Web Application of the LLM-Based Chatbot*

Die Web-Applikation wird über einen HTTP-Server bereitgestellt, der auf Funktionen des für Python entwickelten Web-Frameworks Flask basiert. Dieser Server übernimmt die Verarbeitung der Benutzereingaben sowie der hochgeladenen PDF-Dokumente. Während die Benutzereingaben in Form von JSON-Objekten an das Backend übermittelt werden, erfolgt die Übertragung der PDF-Dokumente als JavaScript-spezifisches FormData-Objekt.

### Implementierung der automatisierten Analyse, der Ableitung von Aktionen und der Wissensablage

Der Flask-Server dient als Schnittstelle zwischen der Web-Applikation und dem Core Script, indem er das Core Script als Modul lädt und die entsprechenden Funktionen für die Verarbeitung von Benutzereingaben und hochgeladenen PDF-Dokumenten aufruft.

#### *Speicherung von PDF-Dokumenten*

Die hochgeladenen PDF-Dokumente werden durch Funktionen der Python-Bibliothek *pdfminer* geöffnet und der darin enthaltene Text wird extrahiert. Anschließend wird der extrahierte Text in Abschnitte, sogenannte Chunks, aufgeteilt. Für die Referenzimplementierung wurde die Länge der Chunks auf 256 Tokens festgelegt, da das verwendete Sentence-Transformer-Modell eine maximale Token-Länge in diesem Bereich er-

fordert, um eine korrekte Verarbeitung zu gewährleisten. Der Chunking-Prozess beginnt mit der Zerlegung des Textes in einzelne Sätze, wobei Satzgrenzen anhand von Schlusspunkten erkannt werden. Anschließend werden die Sätze mithilfe eines speziellen Algorithmus, des sogenannten Tokenizers, in Tokens überführt (siehe Kapitel 2.2.4). Diese Tokens werden dann schrittweise zu Chunks der festgelegten Größe zusammengefügt. Vor dem Hinzufügen der Tokens eines Satzes wird geprüft, ob die Token-Längenbegrenzung überschritten werden würde. Falls dies der Fall ist, wird der aktuelle Chunk abgeschlossen und ein neuer Chunk begonnen. Auf diese Weise endet jeder Chunk stets mit einem vollständigen Satz, um die semantische Kohärenz der Chunks zu gewährleisten.

Die erzeugten Chunks werden im nächsten Schritt durch das Sentence Transformer Modell in Embeddings, also multidimensionale Vektoren, überführt. Für die Referenzimplementierung wurde das Sentence Transformer Modell „all-MiniLM-L6-v2“ verwendet, das Vektoren mit einer Dimension von 384 erzeugt. Diese Modellwahl stellt einen Kompromiss aus kompakten und effizienten Embeddings dar, die wenig Speicherplatz benötigen und gleichzeitig eine gute semantische Repräsentanz bieten.

Die Embeddings werden anschließend an eine Funktion der Facebook AI Similarity Search (FAISS) übergeben, wodurch sie im Vectore Store, dem FAISS-Index, gespeichert werden. Parallel dazu werden die Chunks in Textform zusammen mit Metadaten (Timestamps) in einem JSON-Dokument, der Knowledge Base, gespeichert. Die Embeddings im FAISS-Index und die Einträge im JSON-Dokument erhalten dabei die gleiche Index-ID, wodurch eine eindeutige Zuordnung gewährleistet wird. Nach Abschluss des Vorgangs sind die Inhalte des hochgeladenen PDFs im Speichersystem hinterlegt und können durch den Benutzer über semantische Abfragen abgerufen werden.

***Automatisierte Analyse der Benutzereingaben und Ableitung von Aktionen***

Die Verarbeitung der über das Chatfenster der Web-Applikation erfassten Benutzereingaben unterscheidet sich in einigen Aspekten von der Verarbeitung hochgeladener PDF-Dokumente. Der Text der Benutzereingabe wird zunächst zur Absichtserkennung an das LLM[5] übergeben. Mithilfe eines geeigneten Prompts wird das LLM aufgefordert, die Absicht der Eingabe zu klassifizieren. In der Referenzimplementierung wurden verschiedene Varianten des folgenden Prompts getestet:

„*Bestimme die Absicht des folgenden Textes. Handelt es sich dabei um eine Frage oder eine Anweisung? Antworte nur mit ‚Frage‘ oder ‚Anweisung‘.“*

Basierend auf der Antwort des LLMs (‚*Frage*‘ oder ‚*Anweisung*‘) wird im Anschluss die entsprechende Verarbeitung initiiert. Dazu erfolgt eine Überprüfung der Antwort durch eine bedingte Anweisungsstruktur (If-Else-Konstrukt).

[5] Im Rahmen der Validierung der beschriebenen Referenzimplementierung wurden verschiedene Varianten der drei LLM GPT-4, Bard-2 und Llama-2 zur Realisierung der beschriebenen Funktionen des LLMs eingesetzt. Zur besseren Lesbarkeit wird im Folgenden einheitlich der Begriff „das LLM“ verwendet, obwohl alle drei Modelle die Funktionen erfüllen.

Handelt es sich bei der Benutzereingabe um eine *Frage*, wird der zugehörige Text dem Sentence Transformer übergeben. Analog zum zuvor beschriebenen Vorgehen bei der Verarbeitung von PDF-Dokumenten wird der Text sequenziert, tokenisiert, zu Chunks zusammengefügt und in Embeddings überführt. Das erzeugte Embedding wird anschließend an FAISS übergeben, wo eine semantische Ähnlichkeitssuche durchgeführt wird: Die Embeddings im Vector Store werden mit dem Embedding der Frage auf Basis der euklidischen Distanz verglichen. Dabei berechnet FAISS die Abstände zwischen dem Anfrage-Embedding und allen gespeicherten Embeddings im Vektorraum. Als Ergebnis liefert FAISS die fünf ähnlichsten Embeddings zurück, die die geringsten euklidischen Distanzen aufweisen. Über die zugehörigen Index-IDs werden die entsprechend zugehörigen Inhalte in Textform aus der Knowledge Base (JSON-Dokument) abgerufen. Die ursprüngliche Benutzerfrage sowie die fünf ähnlichsten Inhalte aus der Knowledge Base werden daraufhin an das LLM übergeben. Dieses wird aufgefordert, die Frage des Benutzers basierend auf den bereitgestellten Informationen zu beantworten. Der verwendete Prompt dafür lautet wie folgt:

*„Beantworte basierend auf den folgenden Informationen diese Frage: {user_message} \n \n Wissen: \n {combined_context}."*

In die Variablen *{user_message}* und *{combined_context}* werden die Frage und die identifizierten, passenden Einträge der Knowledge Base eingefügt. Das Ergebnis des LLMs wird anschließend über den Flask-Server an die Web-Applikation übergeben und dort als Antwort des Chatbots im Chatfenster dargestellt. Dieses Vorgehen entspricht der im Stand der Technik vorgestellten Methode der RAG (siehe Kapitel 2.2.2).

Wenn die Benutzereingabe vom LLM als *Anweisung* klassifiziert wird, erfolgt die Ableitung einer entsprechenden Aktion. Im Rahmen der Referenzimplementierung wurde ausschließlich die Möglichkeit berücksichtigt, dass sich die Anweisung auf das Speichern von Wissen im Vector Store bezieht. Die resultierende Aktion umfasst die Verarbeitung der Eingabe durch denselben Workflow wie bei einem hochgeladenen PDF-Dokument. Dabei wird der Text der Eingabe in Embeddings umgewandelt und im Vector Store gespeichert, während der ursprüngliche Text gemeinsam mit zugehörigen Metadaten in der Knowledge Base abgelegt wird.

Grundsätzlich besteht jedoch die Möglichkeit, an dieser Stelle verschiedene Intentionen zu unterscheiden und darauf basierende Aktionen zu implementieren (siehe auch Kapitel 5.1.1 und Abbildung 5.6). Nach der Identifikation einer Anweisung könnte beispielsweise analysiert werden, welche Informationen der Benutzer eingegeben hat. Im Falle der Verfügbarkeit unterschiedlicher Speichersysteme kann das LLM zudem prüfen, in welches Speichersystem (z.B. SQL-Datenbank, Ontologie, JSON-Dokument) die Informationen am besten eingeordnet werden können. Zu diesem Zweck wird dem LLM durch einen gezielten Prompt mitgeteilt, welche Speichersysteme zur Verfügung stehen und welche Art von Informationen in diesen jeweils enthalten sind. Darüber hinaus können exemplarische Anweisungen bereitgestellt werden, um die Wahrscheinlichkeit einer korrekten Klassifizierung durch das LLM zu erhöhen.

## 6.2 Umsetzung des semi-aktiven Wissenstransfers

*Implementation of the Semi Active Knowledge Transfer*

In diesem Kapitel wird die Referenzimplementierung des semi-aktiven Wissenstransfers vorgestellt. Entsprechend des in Kapitel 5.2.1 anhand von Abbildung 5.10 konzeptionell erläuterten Ablaufs zur Erfassung und automatisierten Aufbereitung von Best Practices wird die in Abbildung 6.8 visualisierte Systemarchitektur abgeleitet.

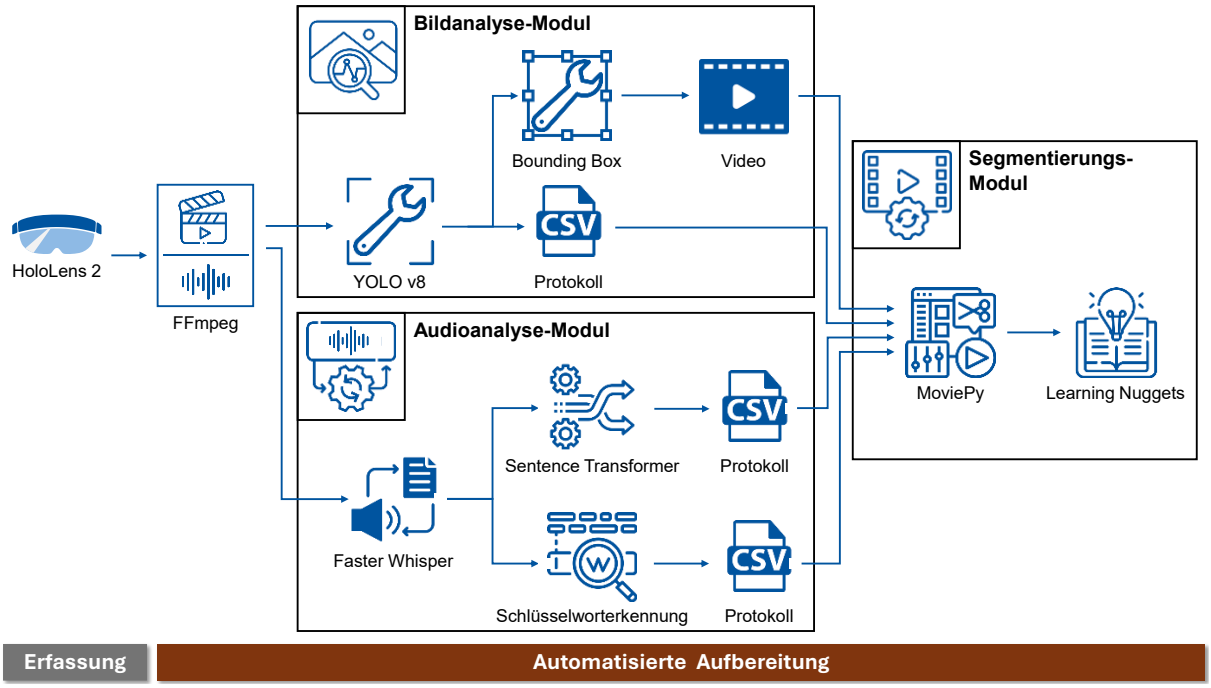



**Abbildung 6.8: Systemarchitektur zur Implementierung der videobasierten Erfassung von Best Practices mit automatisierter Aufbereitung**

*System Architecture for Implementing the Video-Based Recording of Best Practices with Automated Processing*

Gemäß dem Konzept des semi-aktiven Wissenstransfers (siehe Kapitel 5.2) erfolgt die ***Erfassung der Aktivitäten der Mitarbeitenden*** aus deren Perspektive mittels Video- und Tonaufnahmen, um ihr Erfahrungswissen digital zu dokumentieren. Dafür kommt eine Microsoft HoloLens 2 zum Einsatz, die in Abbildung 6.9 zu sehen ist. Dieses HMD bietet eine Vielzahl an Funktionalitäten, die typischerweise für die Entwicklung von AR-Anwendungen genutzt werden. Im Rahmen der vorliegenden Referenzimplementierung beschränkt sich die Nutzung des Geräts jedoch auf die eingebauten Mikrofone und die RGB-Kamera.

Die RGB-Kamera der HoloLens 2 zeichnet Videos mit einer Auflösung von 1920 x 1080 Pixeln und einer Bildfrequenz von 30 Bildern pro Sekunde (FPS) auf. Sie ist mittig an der Vorderseite des Geräts angebracht, sodass die Videoaufnahme aus der Perspektive der Stirn des Trägers erfolgt. Zusätzlich ist die HoloLens 2 mit fünf Mikrofonen ausgestattet, die jeweils separate Audiokanäle liefern. Zwei der Mikrofone befinden sich an der Unterseite des Visiers und dienen primär der Aufzeichnung der

Stimme des Nutzers. Die drei Mikrofone an der Oberseite erfassen gezielt Hintergrundgeräusche, die anschließend aus der kombinierten Tonspur der unteren Mikrofone herausgefiltert werden, um die Audioqualität der erfassten Stimme zu optimieren. [MICR23a] Das von der HoloLens 2 aufgezeichnete Video, einschließlich der zugehörigen Tonspur, wird zur weiteren Verarbeitung vom internen Speicher der HoloLens 2 auf die Workstation übertragen, auf der die einzelnen Module zur Realisierung der automatisierten Aufbereitung implementiert wurden.

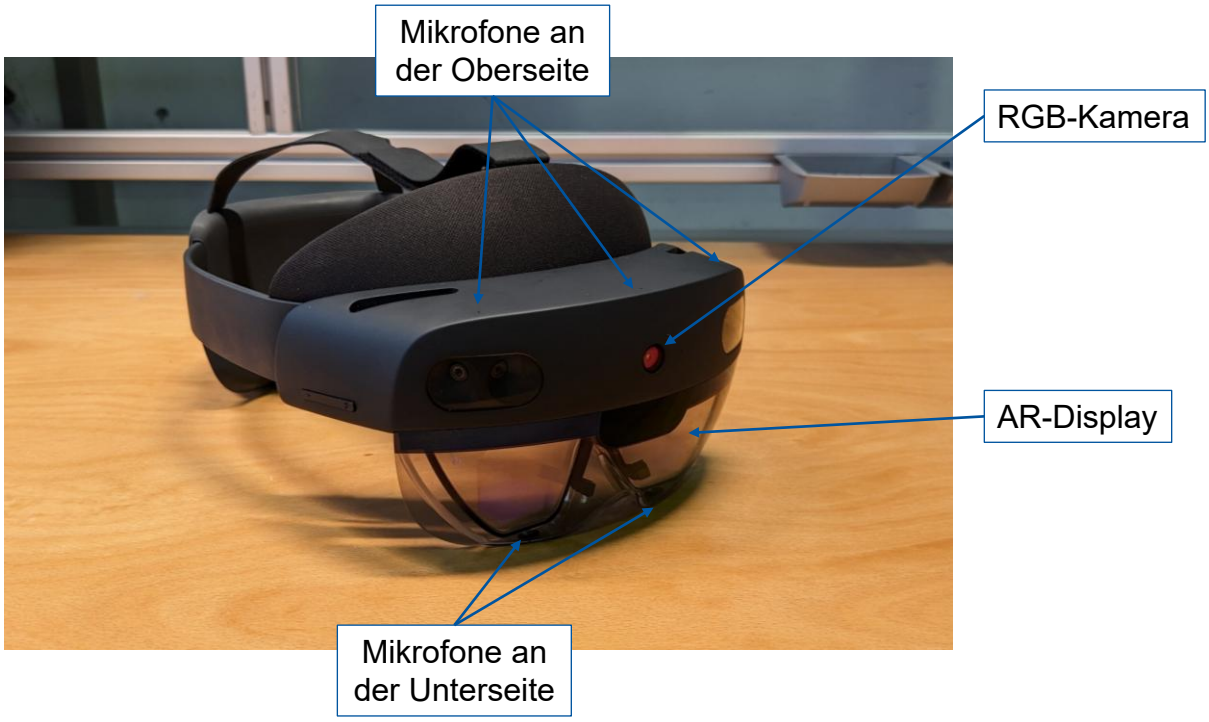



**Abbildung 6.9: Darstellung der HoloLens 2**

*Presentation of HoloLens 2*

Im ersten Schritt der ***automatisierten Aufbereitung*** wird mithilfe des Multimedia-Toolkits *FFmpeg* das Video in die einzelnen 30 Bilder pro Sekunde zerlegt und die Tonspur aus dem aufgezeichneten Video extrahiert. Dabei können spezifische Parameter wie der Audio-Codec, die Abtastrate sowie die Anzahl der Audiokanäle festgelegt werden. Für die Referenzimplementierung werden die folgenden Konfigurationen verwendet:

- Audio-Codec: pcm_s16le (Linear Pulse Code Modulation), 16-Bit, Little Endian (zur unkomprimierten Speicherung der Audiodaten)
- Abtastrate: 44.100 Hz (Standard vieler Audioanwendungen wie Audio-CDs)
- Anzahl der Audiokanäle: Zwei (Stereo), um eine räumliche Klangwiedergabe zu gewährleisten

Die extrahierte Tonspur wird anschließend im WAV-Format ausgegeben, welches eine verlustfreie Weiterverarbeitung der Audiodaten ermöglicht. Insgesamt gewährleisten die gewählten Einstellungen eine für die weitere Verarbeitung und Analyse geeignete

Audioqualität. Im nächsten Schritt werden die resultierenden Bilder des Videos im Bildanalyse-Modul und die als isolierte Datei vorliegende Tonspur im Audioanalyse-Modul parallel verarbeitet.

Zur Analyse der durch die Videoaufnahme erfassten, visuellen Informationen über den betrachteten Prozess wird im ***Bildanalyse-Modul*** ein YOLOv8-Modell (siehe Kapitel 2.2.5) eingesetzt, um die von den Mitarbeitenden bei der Ausführung des Prozesses verwendeten Werkzeuge zu detektieren. Dafür wurde das YOLO-Modell zunächst auf das Erkennen verschiedener Werkzeugtypen trainiert, wobei der frei verfügbare METU-ALET-Datensatz (Automated Labeling of Equiptment and Tools) [KURN19] als Trainingsgrundlage diente. Der Datensatz umfasst 49 Werkzeugklassen aus unterschiedlichen Tätigkeitsbereichen, darunter Werkstätten, Büros, Gärtnereien, Schreinereien und Steinmetzbetriebe. Insgesamt enthält der Datensatz 2788 Bilder mit über 200 annotierten Instanzen pro Werkzeugklasse, die aus einer Mischung von realen und synthetisch erzeugten Aufnahmen bestehen.

Das auf die Erkennung von Werkzeugen trainierte YOLO-Modell untersucht jedes Einzelbild des aufgezeichneten Videos auf das Vorhandensein eines oder mehrerer Objekte aus den erlernten Werkzeugklassen. Für jedes erkannte Werkzeug werden zwei Ausgaben generiert: Zunächst wird ein Ausgabebild des analysierten Eingabebilds erzeugt, auf dem um jedes erkannte Werkzeug eine Bounding Box gezogen wird, an der die zugehörige Klassifizierung (Name des Werkzeugs) sowie das Konfidenzniveau dargestellt werden. Dies ist exemplarisch in Abbildung 6.10 visualisiert.

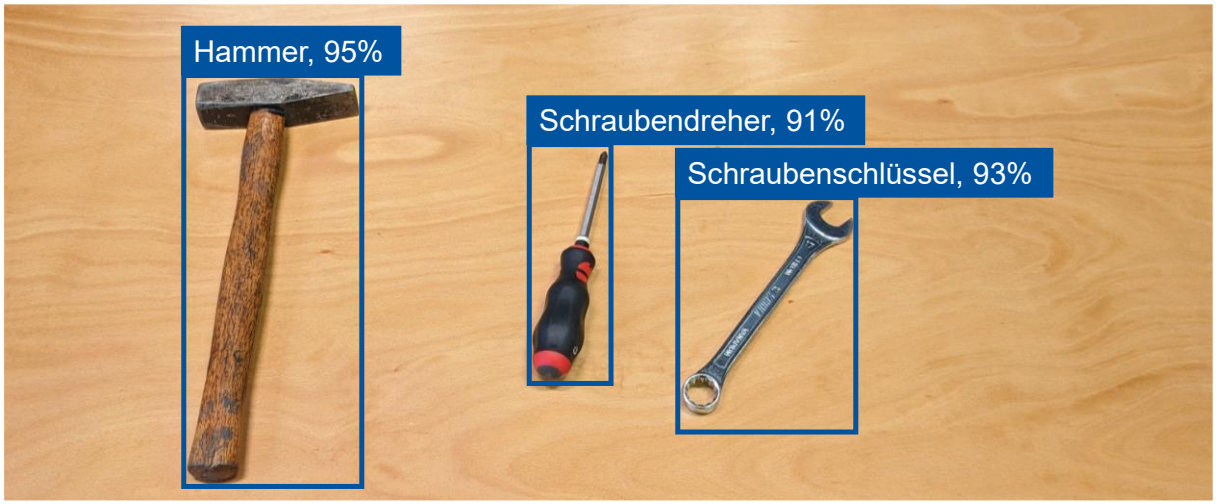



**Abbildung 6.10: Exemplarische Darstellung der Werkzeugdetektion in einem Bild**

*Exemplary Representation of Tool Detection in an Image*

Die auf diese Weise erzeugten Ausgabebilder werden nach Abschluss der Analyse aller Einzelbilder des ursprünglichen Videos zu einem neuen Video zusammengefügt, welches dem Betrachter später die identifizierten Informationen zu den erkannten Werkzeugen visualisiert.

Gleichzeitig wird eine CSV-Datei erzeugt, in der für jedes Einzelbild der Zeitpunkt der Aufnahme ab Beginn des Videos sowie die darin detektierten Werkzeugklassen und ihre Konfidenzniveaus dokumentiert werden.

Der Zeitpunkt $T_{B,i}$ eines Einzelbildes $i$ im Video wird als Quotienten aus der Nummer des Bildes und der Bildfrequenz des Videos bestimmt:

$$T_{B,i} = \frac{Nummer\ des\ Bildes}{Bildfrequenz} \tag{2}$$

Bei einer Bildfrequenz von 30 Bildern pro Sekunde berechnet sich beispielsweise der Zeitpunkt $T_{B,1500}$ des Bildes 1500 wie folgt:

$$T_{B,1500} = \frac{1500}{30} = 50\ Sekunden \tag{3}$$

Mit der Erstellung des mit Bounding Boxes annotierten Videos sowie der Generierung der CSV-Datei ist die Verarbeitung durch das Bildanalyse-Modul abgeschlossen.

Die isolierte Tonspur wird parallel zur Bildanalyse mithilfe des ***Audioanalyse-Moduls*** untersucht. Dazu erfolgt zunächst die Transkription der Audiodatei mittels eines Faster Whisper-Modells (siehe Kapitel 2.2.6). Die von Faster Whisper transkribierten Segmente werden mit einem geeigneten Python-Skript zu einem zusammenhängenden Text vereint. Anschließend werden die einzelnen Sätze anhand der Satzzeichen differenziert. Dieses Vorgehen ist notwendig, da Faster Whisper die Transkription segmentbasiert durchführt und Satzgrenzen nicht explizit identifiziert, sodass ein Segment mitten im Satz zu Ende sein kann. Nach der Differenzierung des transkribierten Textes in einzelne Sätze werden die Start- und Endzeitpunkte jedes Satzes bestimmt. Hierfür werden die von Faster Whisper bereitgestellten Tokens, einschließlich ihrer Start- und Endzeitpunkte, mit den ermittelten Sätzen abgeglichen. Die Startzeit eines Satzes wird durch den Startzeitpunkt des ersten zugehörigen Tokens definiert, während die Endzeit durch den Endzeitpunkt des letzten Tokens bestimmt wird. Darüber hinaus wird für jeden Satz $S_i$ die durchschnittliche $Confidence_{S,i}$ berechnet, um die Erkennungszuverlässigkeit zu quantifizieren. Die Confidence wird als arithmetisches Mittel der von Faster Whisper bei der Transkription bestimmten Confidence-Werte $c_j$ aller $n$ Tokens des jeweils betrachteten Satzes berechnet:

$$Confidence_{S,i} = \frac{\sum_{j=1}^{n} c_j}{n} \tag{4}$$

Die ermittelten Sätze, einschließlich ihrer Start- und Endzeitpunkte in Sekunden sowie der durchschnittlichen Confidence-Werte, werden in einer JSON-Datei gespeichert. Ein Beispiel für die daraus resultierende JSON-Struktur ist in Abbildung 6.11 zu sehen.

Im nächsten Prozessschritt des Audioanalyse-Moduls werden die von Faster Whisper erzeugte Token-Liste der Tonspur $T_{Tonspur}$ und die in der JSON-Datei gespeicherten, transkribierten Sätze separat analysiert. Die Token-Liste $T_{Tonspur}$ wird systematisch auf das Vorkommen spezifischer Schlüsselwörter untersucht, während die transkribierten Sätze im Hinblick auf semantische Zusammenhänge analysiert werden.

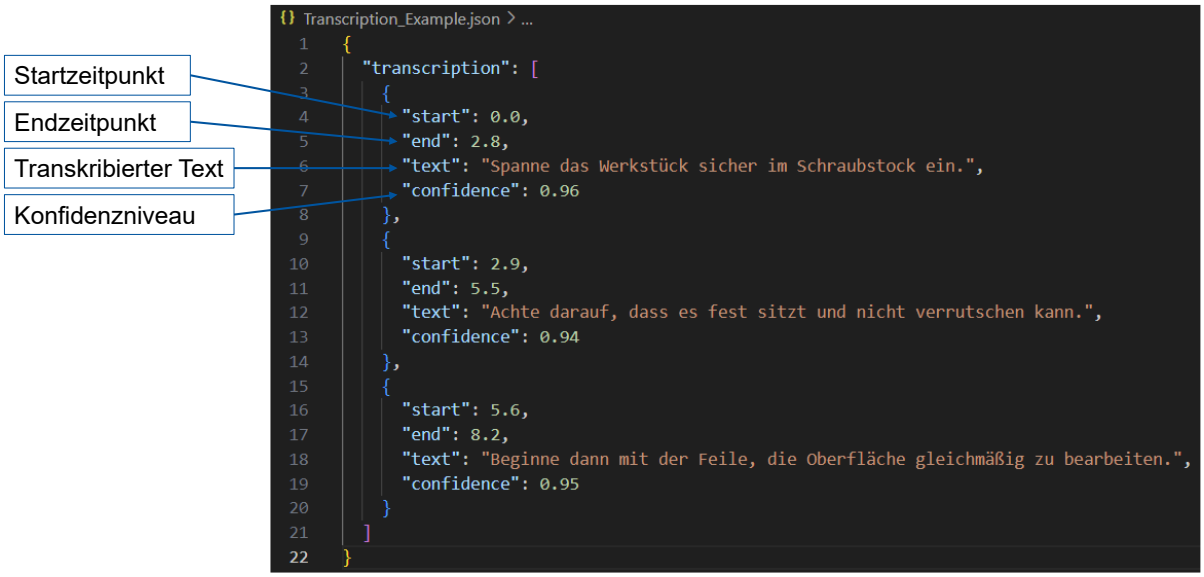



**Abbildung 6.11: Ergebnis der Transkription einer Tonspur mittels Transformer-Modell**

*Result of Transcribing an Audio Track Using a Transformer Model*

Basierend auf den im Konzept des semi-aktiven Wissenstransfers in Tabelle 5.1 definierten semantischen Kategorien von ***Schlüsselwörtern*** (*Aktionen, Objekte, Sequenzindikatoren, Warnhinweise, Werkzeuge*) wird eine ergänzende JSON-Datei erstellt, die insgesamt etwa 500 Schlüsselwörter enthält. Die gewählten Schlüsselwörter stehen alle inhaltlich im Zusammenhang mit Tätigkeiten auf dem Shopfloor produzierender Unternehmen. In Abbildung 6.12 ist die Struktur dieser Wortliste mit einem Schlüsselwort pro Kategorie exemplarisch dargestellt.




**Abbildung 6.12: Verwendete JSON-Struktur der Schlüsselwortliste**

*JSON Structure Used for the Keyword List*

Die Token-Liste $T_{Tonspur}$ wird durch eine Kombination aus For- und If-Schleifen systematisch auf das Vorkommen der hinterlegten Schlüsselwörter untersucht. Da ein einzelnes Token nicht zwangsläufig einem vollständigen Wort entspricht, werden die Schlüsselwörter vorab mit dem gleichen Tokenizer, der auch von Faster Whisper verwendet wird, in ihre Tokenrepräsentanz konvertiert und in einer separaten Token-Liste $T_{Schlüsselwörter}$ gespeichert. Anschließend erfolgt die Iteration durch $T_{Tonspur}$, bei der geprüft wird, ob eine Übereinstimmung mit einem Eintrag in $T_{Schlüsselwörter}$ vorliegt. Für Schlüsselwörter, die in mehrere Tokens segmentiert werden, erfolgt eine sequenzielle Prüfung aufeinanderfolgender Tokens in $T_{Tonspur}$, wobei die Anzahl der Tokens des jeweiligen Schlüsselworts berücksichtigt wird. Wurde ein Schlüsselwort beispielsweise in drei Tokens aufgeteilt, so werden jeweils drei aufeinanderfolgende Tokens in $T_{Tonspur}$ geprüft. Bei einem Treffer werden das erkannte Schlüsselwort sowie die zugehörigen Start- und Endzeitpunkte in einem Protokoll erfasst und in einer weiteren CSV-Datei dokumentiert.

Die mittels Faster Whisper transkribierten Sätze, welche in der in Abbildung 6.11 dargestellten JSON-Struktur vorliegen, werden mithilfe des in Kapitel 2.2.4 vorgestellten ***Sentence Transformer***-Frameworks analysiert, um Sinnabschnitte zu bilden. Gemäß dem in Kapitel 5.2 vorgestellten Konzept wird überprüft, ob zwei aufeinanderfolgende Sätze $S_i$ und $S_{i+1}$ eine ausreichende semantische Ähnlichkeit aufweisen. Für diese Prüfung werden die Sätze zunächst in Embeddings überführt.

Zum Zeitpunkt der hier vorgestellten Erstellung der Referenzimplementierung des Konzepts des semi-aktiven Wissenstransfer stand eine Vielzahl für diese Aufgabe geeigneter Sentence Transformer-Modelle zur Verfügung, bei denen es sich um spezialisierte Varianten des Basismodells BERT handelt (siehe Kapitel 2.2.1). Im Zuge der Implementierung wurde eine gezielte Recherche durchgeführt, um das für die Aufgabe der semantischen Ähnlichkeitsbewertung am besten geeignete Modell zu identifizieren. Die Recherche berücksichtigte die spezifische Aufgabenstellung (Ähnlichkeitsvergleich), die zugrundeliegende Datenstruktur (domänenspezifische, deutsche Texte), die Performance-Kriterien (Genauigkeit der Ähnlichkeitsanalyse) sowie die Hardware-Ressourcen der für diese Dissertation verfügbaren Workstation (siehe Tabelle 6.1). Als Resultat dieser Recherche wurden sechs Sentence Transformer Modelle identifiziert, welche den Anforderungen entsprachen und kostenfrei verfügbar sind: *Roberta-Base*, *Bert-based-german-dbmz-uncased*, *German-Bert-Large-paraphrase-cosine*, *German-BERT-large-paraphrase-euclidean*, *German-Zeroshot* und *German-semantic-base.* Zur Auswahl des optimalen Modells wurden Vorversuche mit domänenspezifischen deutschen Texten durchgeführt, die gezielt mit Sinnbrüchen versehen wurden. Ziel dieser Vorversuche war es, das Modell zu identifizieren, das die Sinnbrüche am besten erkennt, indem es die semantische Ähnlichkeit zwischen dem letzten Satz eines Sinnabschnitts und dem ersten Satz des darauffolgenden Abschnitts möglichst gering bewertet. Aus diesen Versuchen ging *German-Bert-Large-paraphrase-cosine* als das geeignetste Modell hervor und wurde für die Referenzimplementierung ausgewählt.

Mit dem gewählten Sentence Transformer werden die transkribierten Sätze in einem 1024-dimensionalen Vektorraum abgebildet und ihre semantische Ähnlichkeit mithilfe der Cosinus-Ähnlichkeit mathematisch verglichen (siehe Kapitel 2.2.4). Der Schwellenwert für die sinnhafte Zusammengehörigkeit von zwei Sätzen wurde für die Referenzimplementierung auf 0,667 festgelegt.

Die Zeitpunkte der Sinnabschnitte werden auf Grundlage der in der JSON-Datei gespeicherten Start- und Endzeitpunkte der Sätze berechnet. Der Startzeitpunkt eines Sinnabschnitts wird durch den Startzeitpunkt des ersten Satzes innerhalb des Abschnitts definiert, während der Endzeitpunkt durch den Endzeitpunkt des letzten Satzes bestimmt wird. Die als zusammengehörig identifizierten Sätze werden zu einem kohärenten Text kombiniert und gemeinsam mit den ermittelten Start- und Endzeitpunkten in einer dritten CSV-Datei dokumentiert. Mit der Erstellung der CSV-Dateien, die die erkannten Schlüsselwörter sowie die semantisch verknüpften Sätze enthalten, ist die Betrachtung der Tonspur durch das Audioanalyse-Modul abgeschlossen.

Die in den drei CSV-Dateien dokumentierten Zeitpunkte von Werkzeugdetektionen, Schlüsselwortnennungen und semantischen Sinnabschnitten dienen als Eingabedaten für das ***Segmentierungs-Modul*** und bilden die Grundlage für das automatisierte Schneiden des Videos in Learning-Nuggets (siehe Abbildung 6.8). Hierbei wird das im Bildanalyse-Modul annotierte Video, welches Bounding Boxen um die detektierten Werkzeuge enthält, sowie die zu Beginn mit FFmpeg isolierte Tonspur genutzt.

Unter Verwendung der Python-Bibliothek *MoviePy* werden sowohl das annotierte Video als auch die Tonspur entsprechend der ermittelten Zeitpunkte zugeschnitten. Die zugeschnittenen Videoclips und die korrespondierenden Tonspursegmente werden anschließend wieder zu einer synchronisierten Datei zusammengeführt und in einem lokalen Verzeichnis gespeichert. Abhängig vom spezifischen Anwendungsfall sowie der Qualität und dem Inhalt des erfassten Videos und der zugehörigen Tonspur können die für die Segmentierung heranzuziehenden Zeitpunkte flexibel gewählt werden. Dabei besteht die Möglichkeit, zwischen den Zeitpunkten der Werkzeugdetektionen, Schlüsselwortnennungen oder semantischen Sinnabschnitten zu wählen. Die Segmentierung wird durch den gezielten Aufruf der jeweils implementierten Methoden durchgeführt.

Die Erstellung der resultierenden Learning Nuggets bildet den abschließenden Schritt der im Zuge der Referenzimplementierung umgesetzten automatisierten Aufbereitung. Der im Konzeptkapitel des semi-aktiven Wissenstransfers (siehe Kapitel 5.2) in Abbildung 5.10 vorgestellte, neu entwickelte Ablauf für die videobasierte Erfassung von Best Practices umfasst innerhalb der automatisierten Aufbereitung eigentlich auch den Schritt der *Automatisierten Anreicherung*. Das Vorgehen der automatisierten Anreicherung entspricht jedoch dem in Kapitel 5.1.2 im Kontext des *Aktiven Wissenstransfers* beschriebenen Konzepts des Abrufs von Wissen aus digitalen Informationsquellen. Da der Fokus dieser Arbeit auf der Erforschung neuer technologiegestützter Ansätze zur Erfassung von Wissen liegt, würde eine doppelte Implementierung desselben Konzepts keinen zusätzlichen wissenschaftlichen Mehrwert bieten. Aus diesem

Grund wurde die Implementierung dieses Vorgehens ausschließlich bei der Umsetzung der Referenzimplementierung des *Aktiven Wissenstransfers* realisiert und wird entsprechend in Kapitel 6.1 detailliert beschrieben.

Im Anschluss an die Erfassung der Aktivitäten der Mitarbeitenden und die automatisierte Aufbereitung erfolgt gemäß Abbildung 5.10 die ***Nutzung*** der aufbereiteten Inhalte über eine geeignete Applikation oder eine entsprechende GUI. Für die in Kapitel 7 folgende Validierung der Funktionalitäten der Referenzimplementierung wurde eine rudimentäre GUI entwickelt, deren primärer Zweck in der einfachen Ausführung der zuvor beschriebenen automatisierten Aufbereitung besteht. Die GUI wurde unter Einsatz von *PyQt 6* entwickelt. *PyQt 6* ist eine Python-basierte Open-Source-Variante des plattformübergreifenden Qt-Frameworks, das speziell für die Erstellung von Anwendungen mit grafischen Benutzeroberflächen konzipiert wurde [QT G25]. Die GUI umfasst drei Tabs: *Videoanalyse*, *Analyse-Info* und *Videoplayer*.

Im Tab Videoanalyse (siehe Abbildung 6.13) können Benutzer das erfasste Video über eine entsprechende Schaltfläche hochladen oder öffnen. Ein Drop-Down-Menü ermöglicht die Auswahl der Analysemethode, wobei zwischen Objektdetektion, Schlüsselworterkennung und semantischer Analyse gewählt werden kann. Mit der Schaltfläche „Videoanalyse starten“ wird die automatisierte Analyse der Videodatei und der zugehörigen Tonspur ausgeführt.

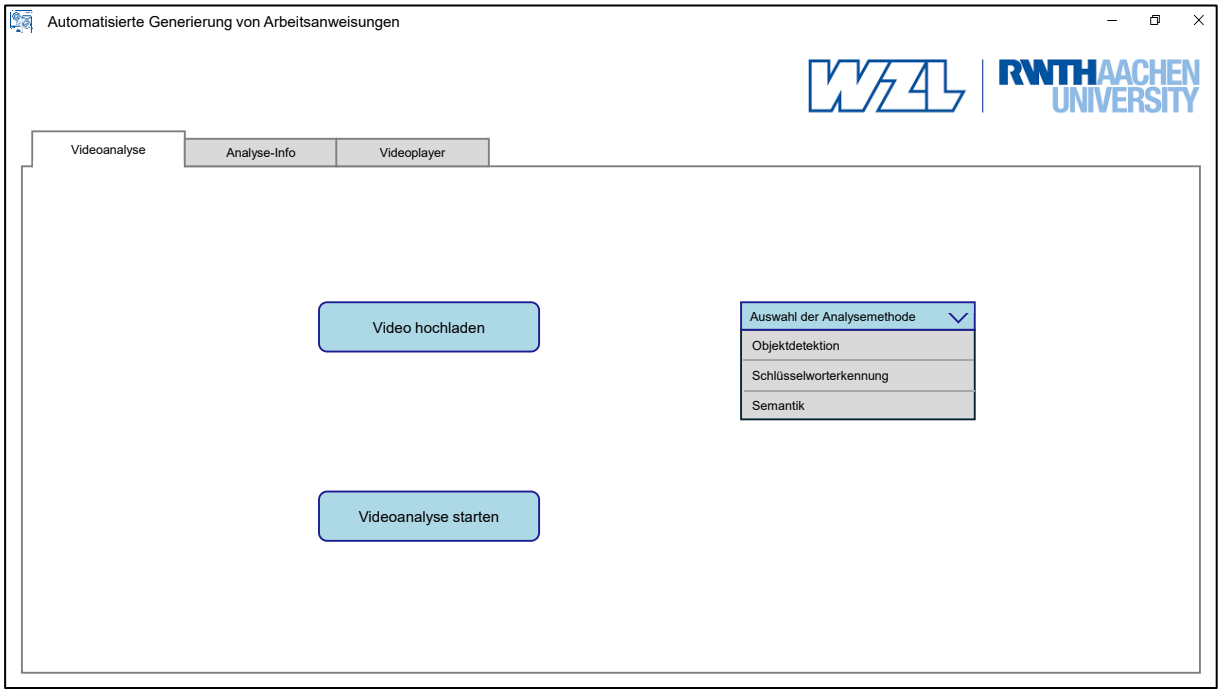



**Abbildung 6.13: Tab *Videoanalyse* der prototypischen GUI**

*Tab Videoanalyse of the Prototypical GUI*

Im Tab *Analyse-Info* (siehe Abbildung 6.14) werden nach erfolgreich durchgeführter Analyse des Videos die transkribierten Sätze inklusive Start- und Endzeitpunkte tabellarisch dargestellt. Diese Darstellung ermöglicht es, die Qualität der Transkription zu evaluieren und eventuelle Abweichungen zu identifizieren.

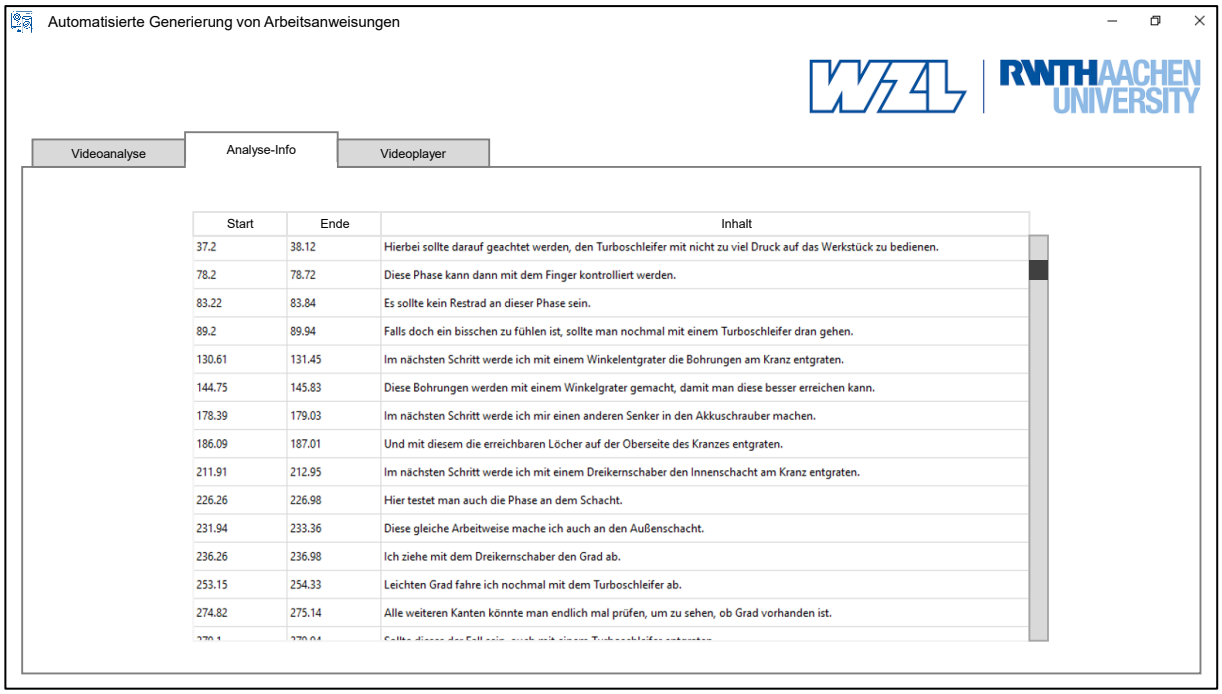



**Abbildung 6.14: Tab *Analyse-Info* der prototypischen GUI**

*Tab Analyse-Info of the Prototypical GUI*

Der Tab *Videoplayer* (siehe Abbildung 6.15) listet die resultierenden Learning-Nuggets in einer geordneten Übersicht chronologisch untereinander auf. Jedes Learning-Nugget wird dabei durch ein Thumbnail repräsentiert. Durch Anklicken eines Thumbnails wird der entsprechende Videoclip mit der synchronisierten Tonspur im zentralen Bereich der Benutzeroberfläche auf der rechten Seite wiedergegeben.




**Abbildung 6.15: Tab *Videoplayer* der prototypischen GUI**

*Tab Videoplayer of the Prototypical GUI*

## 6.3 Umsetzung des passiven Wissenstransfers

*Implementation of the passive Knowledge Transfer*

Das in Kapitel 5.3 vorgestellte Konzept des passiven Wissenstransfers besteht aus zwei Ansätzen. Die ersten beiden Schritte beider Varianten sind konzeptionell identisch (vgl. Abbildung 5.18 und Abbildung 5.26) und basieren daher in der Umsetzung auf einer gemeinsamen Implementierung. Die jeweils nachfolgenden Funktionen wurden jedoch spezifisch auf Basis der jeweiligen Anforderungen entwickelt und separat umgesetzt. Im Folgenden werden die Referenzimplementierungen beider Ansätze vorgestellt.

### 6.3.1 Implementierung der Integration von Hand-Tracking und Process Mining zur Erfassung von Best Practices bei manuellen Montagen

Die Referenzimplementierung der Integration von Hand-Tracking und Process Mining zur Erfassung von Best Practices bei manuellen Montagen wurde an einem rekonfigurierbaren Arbeitstisch im Smart Automation Lab (SAL) des Werkzeugmaschinenlabors WZL der RWTH Aachen University durchgeführt. Die für diese Versuche genutzte Konfiguration des Arbeitstisches ist in Abbildung 6.16 dargestellt.

Der Arbeitstisch besteht aus einem modularen Auf- und Unterbau, der mit Profilen der Firma item konstruiert wurde. Diese Bauweise ermöglicht eine flexible und bedarfsgerechte Anpassung an unterschiedliche Produktionsszenarien. Zur Untersuchung des Hand-Tracking-Ansatzes für manuelle Montagetätigkeiten wurde der Tisch mit einer Werkzeughalterung sowie mehreren Greifboxen ausgestattet, die als Stauraum für Verbindungselemente wie Schrauben und Muttern dienen. Dieser Aufbau bildet typische manuelle Montagearbeitsplätze nach und schafft dadurch die Grundlage für die Untersuchung des Konzepts unter realitätsnahen Bedingungen.




**Abbildung 6.16: Rekonfigurierbarer Arbeitstisch zur Reproduzierung realitätsnaher Arbeitsprozesse**

*Reconfigurable Worktable for Reproducing Realistic Work Processes*

Zusätzlich zum Arbeitsplatz wird eine Referenzbaugruppe benötigt, um das Konzept zu validieren. Dazu wurde eine 3D-gedruckte Lampenhalterung gewählt, die in Abbildung 6.17 gezeigt ist. Diese Baugruppe besteht aus vier Einzelkomponenten, die mithilfe von drei Schraubverbindungen montiert werden. Dabei kommen jeweils eine M5-Zylinderschraube mit Innensechskant, eine M5-Mutter und zwei Unterlegscheiben zum Einsatz.




**Abbildung 6.17: Bei der Referenzimplementierung betrachtete Baugruppe**

*Assembly Group Considered in the Reference Implementation*

Basierend auf dieser Hardware wurden sämtliche erforderlichen Funktionalitäten zur Erfassung und anschließenden Analyse von Handbewegungen gemäß des in Kapitel 5.3.1 vorgestellten Konzeptes in der Programmiersprache Python implementiert. Aufbauend auf dem in Abbildung 5.18 schematisch dargestellten Prozessablauf illustriert Abbildung 6.18 die systematische Verortung der einzelnen Komponenten zur Erfassung der Aktivitäten der Mitarbeitenden bei manuellen Montagen in einer Systemarchitektur. Der Prozess der Erfassung wird durch die Implementierung eines Hand-Tracking-Moduls und eines Grifferkennungs-Moduls realisiert, welche die mittels Kamera aufgezeichneten Montagetätigkeiten verarbeiten.

**Implementierung der Erfassung der Aktivitäten**

Bevor die Analyse durchgeführt werden kann, müssen die Montageabläufe zunächst videobasiert aufgezeichnet werden. Hierzu wurden zwei Kameras an dem rekonfigurierbaren Arbeitstisch installiert. Aufgrund der Verfügbarkeit kamen eine XBOX Kinect V2 und eine Intel RealSense D435 zum Einsatz (siehe Abbildung 6.16). Beide Kameramodelle arbeiten im RGB-Farbraum mit einer Auflösung von 1920 x 1080 Pixeln und einer Bildwiederholungsrate von 30 Bildern pro Sekunde.

Entsprechend des in Abbildung 5.20 gezeigten Konzeptes zur vereinfachten Kombination der 2D-Informationen zweier Kameras, wurde die XBOX Kinect senkrecht über der Arbeitsplatte des Tisches angebracht, während die Intel RealSense hinter dem Tisch montiert wurde.

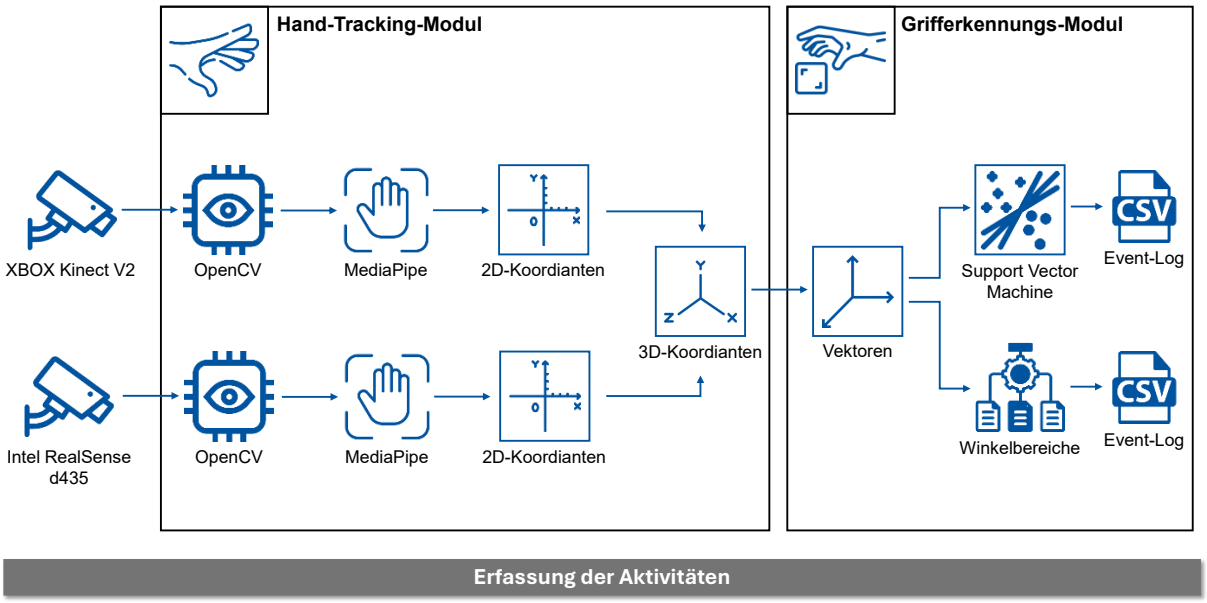


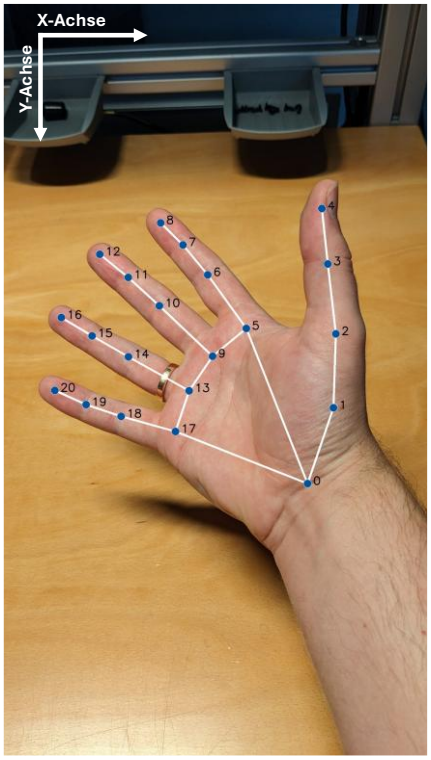

**Abbildung 6.18: Systemarchitektur zur Implementierung der Erfassung der Aktivitäten bei manuellen Montagen**

*System Architecture for Implementing the Recording of Activities of Manual Assembly*

Die Linsen beider Kameras sind dabei in einer Ebene angeordnet, die senkrecht zur Arbeitsplatte verläuft. Der Abstand der Linsen beträgt sowohl vertikal als auch horizontal 70 cm. In Übereinstimmung mit dem in Abbildung 5.19 dargestellten Konzept für die Erfassung des relevanten Arbeitsraumes wurde die Systemgrenze dabei so definiert, dass die gesamte Arbeitsfläche des Tisches mit den Abmessungen von 180 cm x 75 cm aus der Top-Down-Perspektive von der XBOX Kinect vollständig erfasst wird.

Die Videoaufnahmen der beiden Kameras werden im ersten Schritt des ***Hand-Tracking-Moduls*** mittels Methoden der Python-Bibliothek *OpenCV* parallel aus einem lokalen Verzeichnis geladen. Anschließend werden jeweils die einzelnen Bilder der Videos extrahiert. Da *OpenCV* bei der Extraktion der Bilder diese automatisch im BGR-Farbraum speichert, wird daraufhin noch eine Rücktransformation in den RGB-Farbraum vorgenommen, um die weitere Verarbeitung zu ermöglichen.

Die extrahierten Bilder werden im nächsten Schritt mittels der im MediaPipe-Framework enthaltenen Bibliothek zur Detektion von Hand-Landmarks (siehe Kapitel 2.2.7) weiterverarbeitet. Pro Bild wird die maximale Anzahl der erkennbaren Hände auf zwei begrenzt. Entsprechend werden maximal zwei Datensätze mit den zugehörigen Koordinatenpunkten generiert. In Abbildung 6.19 ist die daraus resultierende Erfassung der Hand-Landmarks exemplarisch für das Bild einer Hand visualisiert. Es ist erkennbar, dass pro Finger vier Landmarks identifiziert werden, deren Positionen kombiniert die Lage des jeweiligen Fingers beschreiben. Der Ursprung des Koordinatensystems (0, 0) befindet sich in der oberen linken Ecke des Bildes, wobei die X-Koordinate horizontal nach rechts und die Y-Koordinate vertikal nach unten ansteigt.




| Landmark | $x_{Normalisiert}$ | $y_{Normalisiert}$ | $x_{Pixel}$ | $y_{Pixel}$ |
|---|---|---|---|---|
| 0 | 0.716490 | 0.638145 | 1625 | 2573 |
| 1 | 0.777337 | 0.537450 | 1763 | 2167 |
| 2 | 0.783069 | 0.439236 | 1776 | 1771 |
| 3 | 0.764109 | 0.346478 | 1733 | 1397 |
| 4 | 0.750000 | 0.272817 | 1701 | 1100 |
| 5 | 0.572310 | 0.432044 | 1298 | 1742 |
| 6 | 0.481481 | 0.360863 | 1092 | 1455 |
| 7 | 0.423280 | 0.321181 | 960 | 1295 |
| 8 | 0.370811 | 0.291915 | 841 | 1177 |
| 9 | 0.492945 | 0.468998 | 1118 | 1891 |
| 10 | 0.367725 | 0.402034 | 834 | 1621 |
| 11 | 0.294533 | 0.364335 | 668 | 1469 |
| 12 | 0.227513 | 0.333581 | 516 | 1345 |
| 13 | 0.437831 | 0.514881 | 993 | 2076 |
| 14 | 0.294974 | 0.469742 | 669 | 1894 |
| 15 | 0.209436 | 0.441964 | 475 | 1782 |
| 16 | 0.137125 | 0.417659 | 311 | 1684 |
| 17 | 0.405644 | 0.568948 | 920 | 2294 |
| 18 | 0.277778 | 0.548859 | 630 | 2213 |
| 19 | 0.195326 | 0.533234 | 443 | 2150 |
| 20 | 0.121693 | 0.514385 | 276 | 2074 |

**Abbildung 6.19: Beispielhafte Darstellung der Erfassung der Hand-Landmarks und der zugehörigen Koordinaten**

*Exemplary representation of the recording of the hand landmarks and the associated coordinates*

Die neben dem Bild dargestellte Tabelle zeigt pro Landmark sowohl die normalisierten als auch die pixelbasierten X- und Y-Koordinaten. Die Normalisierung erfolgt relativ zur Größe des Bildes und liegt im Bereich von 0 bis 1. Ein X-Koordinatenwert von 0 entspricht der linken Bildkante, während ein Wert von 1 die rechte Bildkante repräsentiert. Analog steht ein Y-Koordinatenwert von 0 für die obere Bildkante und ein Wert von 1 für die untere Bildkante. Die Berechnung der normalisierten Koordinaten erfolgt durch Division der pixelbasierten Koordinatenwerte $x_{Pixel}$ und $y_{Pixel}$ durch die jeweilige Bildbreite bzw. Bildhöhe, wobei die Bildbreite und Bildhöhe die maximale Anzahl von Pixeln in der horizontalen bzw. vertikalen Ausrichtung des Bildes repräsentieren:

$$x_{Normalisiert} = \frac{x_{Pixel}}{Bildbreite} \tag{5}$$

$$y_{Normalisiert} = \frac{y_{Pixel}}{Bildhöhe} \tag{6}$$

Das hier verwendete Beispielbild weist eine Auflösung von 4032 Pixeln in der Höhe (Y-Achse) und 2268 Pixeln in der Breite (X-Achse) auf. Für Landmark 0 ergibt sich beispielhaft folgende Berechnung:

$$x_{Normalisiert} = \frac{1625}{2268} = 0.716490 \tag{7}$$

$$y_{Normalisiert} = \frac{2573}{4032} = 0.638145 \tag{8}$$

Mit diesem Vorgehen erhält man bei einer Bildrate von 30 Bildern pro Sekunde 30 2D-Koordinatenpunkte pro Landmark $i$ pro Sekunde, die wie folgt beschrieben werden können:

$$P_{2D,i}(x_i, y_i) = \begin{pmatrix} x_i \\ y_i \end{pmatrix} \tag{9}$$

Durch die in Abbildung 6.16 gezeigte Positionierung der beiden Kameras am rekonfigurierbaren Arbeitstisch können die zweidimensionalen Koordinatenpunkte beider Kameras mittels der in Abbildung 5.20 vorgestellten Koordinatentransformation zu dreidimensionalen Positionsdaten zusammengeführt werden. Die resultierenden Punkte bzw. deren Ortsvektoren im Koordinatensystem im dreidimensionalen Raum lauten demnach wie folgt:

$$P_{3D,i}(x_{1,i}, y_{1,i}, z_{1,i}) = \begin{pmatrix} x_{1,i} \\ y_{1,i} \\ z_{1,i} \end{pmatrix} = \begin{pmatrix} x_{1,i} \\ y_{1,i} \\ y_{2,i} \end{pmatrix} \tag{10}$$

Dabei entsprechen $x_1, y_1, z_1$ den dreidimensionalen Koordinaten des Punktes $P_{3D,i}$ ausgehend vom Koordinatensystem der oberhalb des Arbeitsraumes angebrachten Kamera (XBOX Kinect V2).

Die somit für jedes Bild des Videos erfassten dreidimensionalen Koordinatenpunkte einer Hand werden zur späteren Analyse der verwendeten Griffe zunächst im ***Grifferkennungs-Modul*** in eine Vektordarstellung überführt. Bezogen auf die verwendeten 21 Hand-Landmarks lässt sich jeder Finger durch drei Verbindungsvektoren darstellen, die jeweils durch die Subtraktion zweier Ortsvektoren von aufeinanderfolgenden Hand-Landmarks berechnet werden:

$$v_{i,i+1} = P_{i+1} - P_i \tag{11}$$

Dabei ist „i“ der Index für die Hand-Landmarks entsprechend der Nummerierung in Abbildung 6.19. Beispielsweise ergibt sich so der folgende Vektor, der die Verbindung zwischen dem proximalen Gelenk (erstes Gelenk von der Spitze aus betrachtet) und der Spitze des Daumens darstellt:

$$v_{3,4} = P_4 - P_3 \tag{12}$$

Jeder Verbindungsvektor beschreibt die Position und Ausrichtung des jeweiligen Fingersegments, das zwischen den betrachteten Hand-Landmarks liegt. Die in (11) und (12) genutzte Notation ist jedoch ausschließlich innerhalb der Nummerierung der einzelnen Finger anwendbar. Ergänzend dazu werden die Verbindungsvektoren zwischen dem Hand-Landmark 0 (Handgelenk) und den Startpunkten der Finger (Hand-Landmarks 1, 5, 9, 13 und 17) beschrieben. Diese Vektoren lassen sich zwar nicht direkt mit der oben genannten Notation darstellen, folgen jedoch dem gleichen Prinzip der Subtraktion der Ortsvektoren. Die Methodik ist sowohl auf die zweidimensionalen als auch auf die dreidimensionalen Koordinaten der Hand-Landmarks anwendbar. Zur Veranschaulichung dieses Prozessschritts ist in Abbildung 6.20 ein Beispiel für die so entstehende Vektorrepräsentanz einer Hand im dreidimensionalen Raum dargestellt.

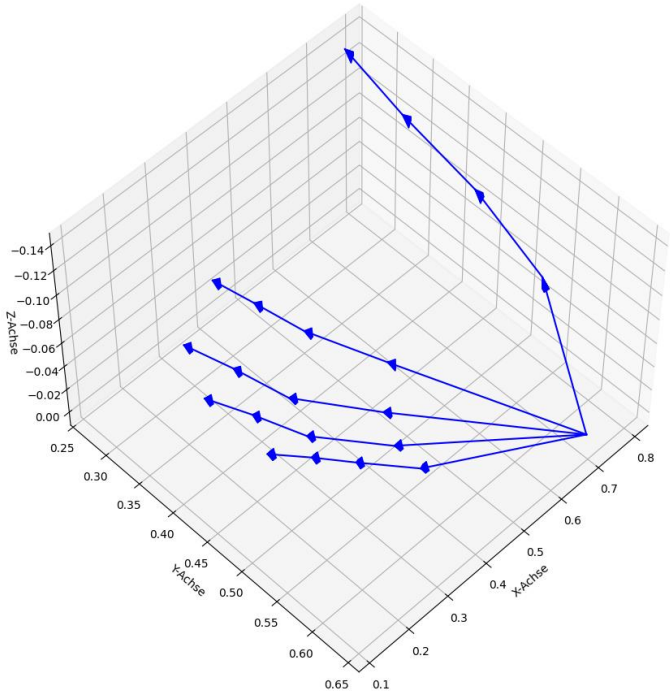



**Abbildung 6.20: Dreidimensionale Darstellung der Vektorrepräsentanz einer Hand**

*Three-Dimensional Illustration of the Vector Representation of a Hand*

Die abgebildeten Vektoren basieren auf dem in Abbildung 6.19 gezeigten Foto einer rechten Hand. Die Achsen des Diagramms stellen die Positionen in normalisierten Koordinaten dar. Die Werte auf der Z-Achse sind negativ, da die Achse in positiver Richtung senkrecht nach unten verläuft (vergleiche Abbildung 5.20). Die Beugung und Orientierung der Finger können daraufhin durch die Berechnung des Winkels θ zwischen den aufeinanderfolgenden Verbindungsvektoren bestimmt werden. Hierzu wird die Gleichung für den Cosinus des Winkels formuliert, die dem Verhältnis des Skalarprodukts zweier Vektoren zu dem Produkt der Beträge dieser Vektoren entspricht. Anschließend wird die Gleichung nach dem Winkel θ umgeformt:

$$\cos(\theta_{v_{i,i+1},v_{i+1,i+2}}) = \frac{v_{i,i+1} \cdot v_{i+1,i+2}}{|v_{i,i+1}||v_{i+1,i+2}|} \tag{13}$$

$$\theta_{v_{i,i+1},v_{i+1,i+2}} = \arccos\left(\frac{v_{i,i+1} \cdot v_{i+1,i+2}}{|v_{i,i+1}||v_{i+1,i+2}|}\right) \tag{14}$$

Abbildung 6.21 visualisiert den zuvor beschriebenen Prozess der Vektorenbildung basierend auf den Hand-Landmarks sowie die Position des Winkels θ zwischen zwei aufeinanderfolgenden Verbindungsvektoren. Zur Erhöhung der Anschaulichkeit wurde dabei die Darstellung im zweidimensionalen Raum gewählt. Auf der linken Seite sind die resultierenden Vektoren direkt auf der Hand dargestellt, während sie auf der rechten Seite in das zugehörige, zweidimensionale, normalisierte Koordinatensystem übertragen werden. Darin ist der Winkel $\theta_{v_{1,2},v_{2,3}}$ zwischen den Vektoren $v_{1,2}$ und $v_{2,3}$ exemplarisch eingezeichnet. Die Lage der übrigen Winkel verhält sich analog und das beschriebene Verfahren lässt sich gleichermaßen auf den dreidimensionalen Raum übertragen.

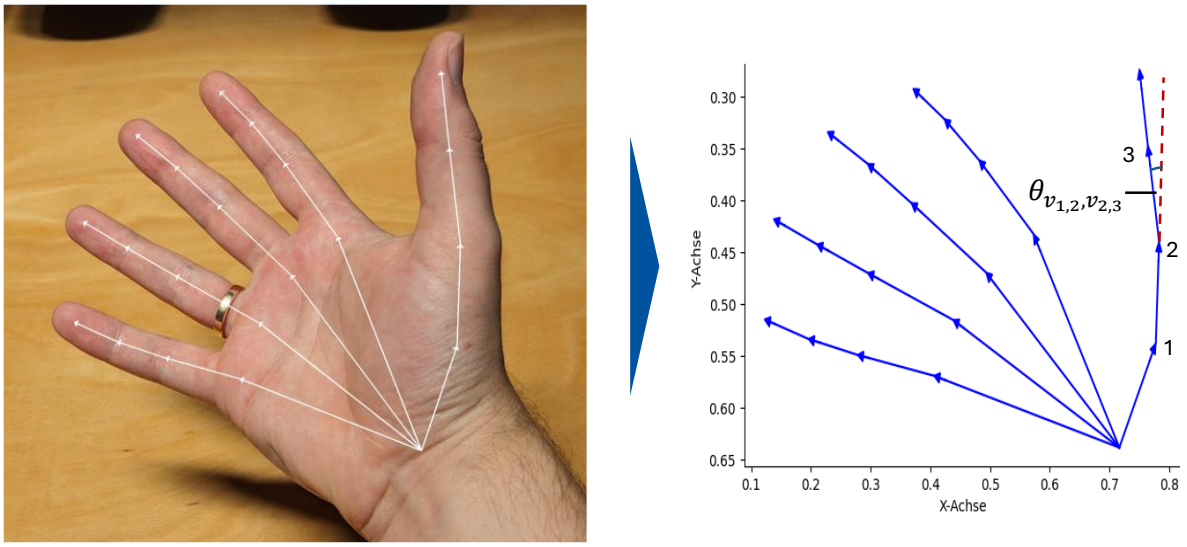



**Abbildung 6.21: Beispielhafte Darstellung des Winkels zwischen zwei aufeinanderfolgenden Verbindungsvektoren im zweidimensionalen Raum**

*Exemplary Representation of the Angles Between Two Consecutive Connecting Vectors in Two-Dimensional Space*

Auf Grundlage der berechneten Winkel zwischen den Vektoren der Fingerglieder wird die Differenzierung der während des betrachteten Montageprozesses verwendeten Griffarten vorgenommen. Dieser Schritt bedarf jedoch (wie in Kapitel 5.3.1 erklärt) einer Vorstudie zur Ermittlung der Griffarten und der zugehörigen Winkelbereiche.

In der ***Vorstudie*** zur Referenzimplementierung wurde die Montage der Referenzbaugruppe (siehe Abbildung 6.17) von fünf unterschiedlichen Personen jeweils zweimal durchgeführt und per Video dokumentiert. Die Aufnahmen wurden anschließend manuell analysiert, um die verwendeten Griffarten, die jeweils verwendete Hand (links oder rechts) sowie die Zeiträume, in denen sie im Video ausgeführt wurden, zu identifizieren. Auf Basis dieser Analyse wurden fünf Griffklassen definiert, die bei der Montage der 3D-gedruckten Lampe regelmäßig zum Einsatz kamen. Diese Griffklassen sind in Abbildung 6.22 dargestellt und umfassen den *Umschließenden Griff* (Greifen einer Komponente mit allen Fingern), den *Pinzettengriff* (Greifen einer Komponente mit Daumen und Zeigefinger), den *Drei-Punkt-Griff* (Greifen der Komponente mit Daumen, Zeigefinger und Mittelfinger), den *Werkzeuggriff* und die *Faust.*




**Abbildung 6.22: Definierte Griffklassen für die Montage der Referenzbaugruppe**

*Defined Grip Classes for Mounting the Reference Assembly*

Der *Umschließende Griff* kam insbesondere beim Handling der größeren 3D-gedruckten Komponenten zum Einsatz, während *Pinzettengriff* und *Drei-Punkt-Griff* vorrangig für die Handhabung und das Fügen der kleineren Elemente wie Schrauben, Muttern und Unterlegscheiben verwendet wurden. Die Griffklasse *Faust* wurde hinzugefügt, um in der weiteren Untersuchung als Indikator für den Abschluss des Montagevorgangs der Baugruppe zu dienen.

Zur Bestimmung der charakteristischen Winkelbereiche jeder Griffklasse wurden das Hand-Tracking-Modul sowie eine modifizierte Version des Grifferkennungs-Moduls eingesetzt. Die Modifikation ermöglichte es, die zuvor manuell erfassten Zeitintervalle der dokumentierten Griffausführungen sowie die jeweils verwendete Hand im System zu hinterlegen. In diesen Zeitabschnitten wurden die Winkel zwischen den Vektordarstellungen der Fingerglieder berechnet und den entsprechenden Griffklassen zugeordnet. Die zugeordneten Winkelwerte wurden über alle Personen hinweg aggregiert und zur Eliminierung von Ausreißern wurden das 5. und 95. Perzentil der Winkel als untere bzw. obere Grenze der gültigen Winkelbereiche je Griffklasse festgelegt. In Tabelle 6.2 sind die resultierenden Winkelbereiche für die einzelnen Griffklassen exemplarisch für den Daumen dargestellt.

**Tabelle 6.2: In der Vorstudie ermittelte Winkelbereiche des Daumens pro Griffklasse**

*Angle Ranges of the Thumb per Grip Class Determined in the Preliminary Study*

| Griffklasse | Segment 1 | Segment 2 | 5. Perzentil [Grad] | 95. Perzentil [Grad] |
|---|---|---|---|---|
| **Umschließender Griff** | 0-1 | 1-2 | 11,62185975 | 47,42242424 |
| **Umschließender Griff** | 1-2 | 2-3 | 11,08072117 | 38,1112991 |
| **Umschließender Griff** | 2-3 | 3-4 | 5,063831834 | 31,6209382 |
| **Pinzettengriff** | 0-1 | 1-2 | 7,709546335 | 59,12788518 |
| **Pinzettengriff** | 1-2 | 2-3 | 8,680121932 | 36,57540013 |
| **Pinzettengriff** | 2-3 | 3-4 | 4,064700826 | 33,12930861 |
| **Drei-Punkt-Griff** | 0-1 | 1-2 | 13,6479378 | 45,91543168 |
| **Drei-Punkt-Griff** | 1-2 | 2-3 | 13,57562391 | 38,38205843 |
| **Drei-Punkt-Griff** | 2-3 | 3-4 | 6,330988039 | 29,03307195 |
| **Werkzeuggriff** | 0-1 | 1-2 | 25,96264766 | 64,13798482 |
| **Werkzeuggriff** | 1-2 | 2-3 | 11,80080572 | 30,62056325 |
| **Werkzeuggriff** | 2-3 | 3-4 | 5,432967583 | 24,71933579 |
| **Faust** | 0-1 | 1-2 | 30,90949093 | 32,20628341 |
| **Faust** | 1-2 | 2-3 | 37,94506988 | 44,76602781 |
| **Faust** | 2-3 | 3-4 | 30,77197053 | 34,32577464 |

Die dargestellten Bezeichnungen „Segment 1“ und „Segment 2“ beziehen sich auf die anatomischen Abschnitte des Daumens, die durch jeweils zwei benachbarte Hand-

Landmarks definiert und als Vektoren repräsentiert werden. Die verwendeten Zahlenkodierungen entsprechen den Indizes der Hand-Landmarks gemäß dem MediaPipe-Handmodell. Die dargestellten Winkelwerte wurden auf Basis der berechneten Vektoren ermittelt und in Grad angegeben.

Auf Basis der in der Vorstudie bestimmten Winkelbereiche werden sowohl ein regelbasierter als auch ein Machine-Learning-basierter Ansatz zur Klassifikation der verwendeten Griffe eingesetzt (siehe Abbildung 6.18):

Der *regelbasierte Ansatz* nutzt eine deterministische Entscheidungslogik zur Klassifikation, die als Kombination aus For- und If-Schleifen implementiert wird. Beim Durchlaufen dieser Schleifen werden die Winkel zwischen den Vektoren der Fingersegmente berechnet und anschließend wird überprüft, ob diese innerhalb der vordefinierten Winkelbereiche der Griffklassen liegen. Sofern eine Übereinstimmung vorliegt, wird die zugehörige Griffklasse als Ergebnis ausgegeben.

Im *Machine-Learning-basierten Ansatz* wird eine SVM verwendet, um die Griffarten automatisiert zu klassifizieren. Zur Erstellung des für die SVM notwendigen Trainingsdatensatzes wurden für jede Griffart separate Videosequenzen mit einer Dauer von 60 Sekunden aufgenommen, in denen die jeweilige Griffart mit leichten Variationen und aus unterschiedlichen Perspektiven demonstriert wurde. Die Aufzeichnung erfolgte mit der XBOX Kinect (siehe Abbildung 6.16) mit einer Bildfrequenz von 30 Bildern pro Sekunde, sodass für jede Griffart 1800 Bilder entstanden. Aus diesen Bildern extrahierte das modifizierte Grifferkennungs-Modul die relevanten Winkel zwischen den Vektoren der Fingersegmente. Die ermittelten Merkmale wurden zusammen mit der zugehörigen Griffklasse in eine CSV-Datei überführt, wobei die Griffklassen numerisch codiert wurden (0: Umschließender Griff, 1: Pinzettengriff, 2: Drei-Punkt-Griff, 3: Werkzeuggriff, 4: Faust). Für das Training der SVM wurden 80 % der Datensätze (1440 Bilder) verwendet, während die verbleibenden 20 % (360 Bilder) zur Validierung dienten. Da im betrachteten Fall eine Differenzierung zwischen fünf Griffklassen (Nummer 0-4) erforderlich ist, erfolgte eine Anpassung an eine Mehrklassen-Spezifikation. Hierzu wurde der sogenannte One-vs-All-Ansatz angewendet, bei dem für jede Griffklasse eine separate binäre SVM trainiert wird, die jeweils die Zugehörigkeit zu ihrer Klasse gegen alle anderen differenziert. Beim Einsatz bewertet jede dieser fünf SVM den Vektordatensatz eines neu erfassten Bildes separat und gibt eine Bewertung aus, die das Maß der Übereinstimmung mit der jeweiligen Griffklasse beschreibt. Diejenige Klasse mit der höchsten Bewertung wird als Ergebnis der Klassifikation ausgewählt.

Der regelbasierte und der SVM-basierte Ansatz wurden als alternative Methoden zur Identifizierung der verwendeten Griffklassen implementiert. Beide Ansätze werden in der Validierung (siehe Kapitel 7.1) hinsichtlich ihrer Klassifikationsgenauigkeit verglichen. Unabhängig vom verwendeten Ansatz generiert das Grifferkennungsmodul als Ergebnis einen Event-Log in Form einer CSV-Datei, in der für jedes Bild des untersuchten Videos die jeweils erkannte Griffklasse pro Hand, einschließlich des zugehörigen Zeitstempels, protokolliert wird. Mit der Ausgabe dieses Event-Logs ist die Erfassung der Aktivitäten abgeschlossen (siehe Abbildung 6.18).

### Implementierung der Analyse der Aktivitäten

Entsprechend dem in Abbildung 5.18 konzeptionell vorgestellten Prozessablauf erfolgt nach Erfassung die Analyse der Aktivitäten. Die dafür in der Implementierung verwendete Systemarchitektur ist in Abbildung 6.23 dargestellt und wird im Folgenden erläutert.

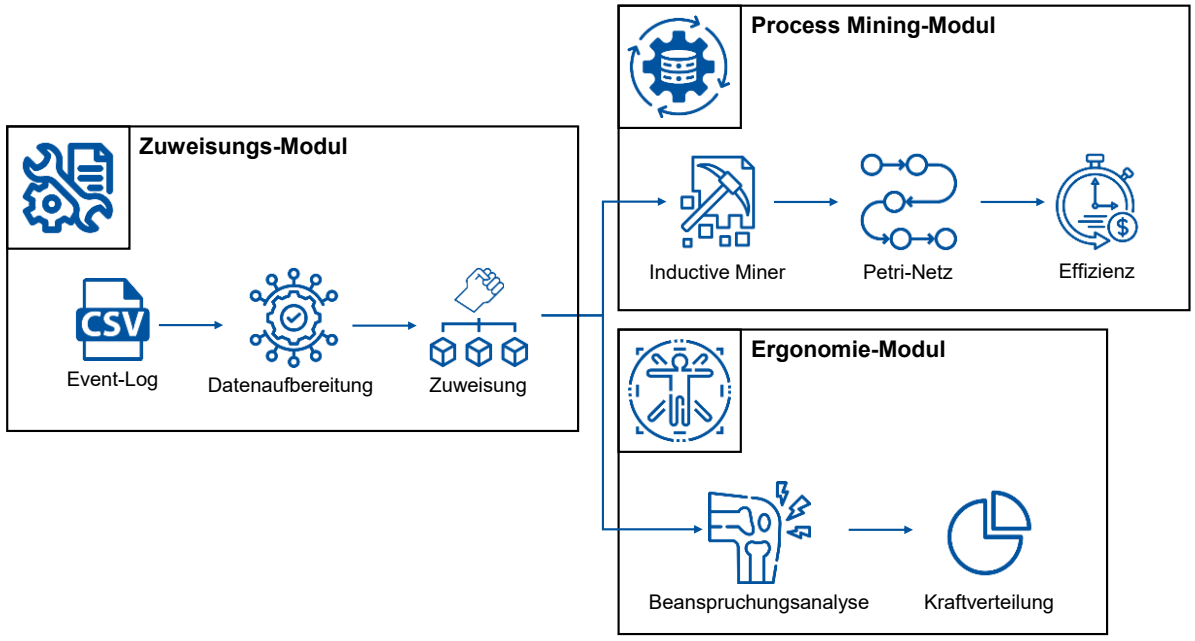



**Abbildung 6.23: Systemarchitektur zur Implementierung der Analyse der Aktivitäten bei manuellen Montagen**

*System Architecture for Implementing the Analysis of Activities of Manual Assembly*

Der zuvor erzeugte Event-Log wird im ersten Schritt der Analyse einer strukturierten Datenaufbereitung durch das ***Zuweisungs-Modul*** unterzogen. Dabei erfolgt eine Bereinigung des Datensatzes, bei der alle Zeilen ohne relevante Informationen entfernt werden. Dazu zählen Leerzeilen, die entstehen, wenn das System aufgrund fehlender oder unvollständiger Hände in einem Bild keine erfolgreiche Handerkennung durchführen kann. Zusätzlich werden Zeilen entfernt, bei denen „None“ als Ergebnis der erkannten Griffklasse dokumentiert ist. Dieser Eintrag drückt aus, dass zwar eine Hand detektiert wurde, dieser jedoch keine Griffklasse zugewiesen werden konnte.

Daran anschließend wird eine Datenreduktion durchgeführt. Da die hier beschriebene Umsetzung mit einer Kamera mit einer Bildfrequenz von 30 Bildern pro Sekunde erfolgte, enthält der Event-Log maximal 30 Griffklassen pro Hand und Sekunde. Diese Datenmenge wird durch eine Aggregation auf ein Dreißigstel seiner ursprünglichen Größe reduziert, indem pro Sekunde der Modus der Daten, also die häufigste erfasste Griffklasse, für jede Hand bestimmt und als repräsentativ für die betrachtete Sekunde angenommen wird. Anschließend werden aufeinanderfolgende Zeilen im Event-Log, in denen identische Griffklassen pro Hand erkannt wurden, zusammengeführt. Die entsprechenden Zeitintervalle werden dabei zu einem einzigen Zeitraum aggregiert.

Abschließend erfolgt eine Zuweisung von Kombinationen von gleichzeitig mit der linken und rechten Hand ausgeführter Griffklassen zu Komponenten der montierten Baugruppe. Die Grundlage für diese Zuweisung bildet erneut die zuvor beschriebene Vorstudie, in der neben der Erfassung der Winkelbereiche für jede Griffklasse auch dokumentiert wurde, welche Kombinationen von Griffklassen die Probanden zur Durchführung der einzelnen Montageschritte zur Montage der Referenzbaugruppe verwendeten. Auf Grundlage dieser Beobachtungen wurden die in Tabelle 6.3 zu sehenden Zuweisungen durchgeführt.

**Tabelle 6.3: Zuweisung von Griffkombinationen zu Montageschritten für die Montage der Referenzbaugruppe**

*Assignment of Grip Combinations to Assembly Steps for the Assembly of the Reference Assembly*

| Griffklasse linke Hand | Griffklasse rechte Hand | Montageschritt |
|---|---|---|
| Umschließender Griff | Umschließender Griff | Handhabung eines Bauteils |
| Pinzettengriff | Pinzettengriff | Kombination von Schraube und Unterlegscheibe |
| Drei-Punkt-Griff | Pinzettengriff | |
| Drei-Punkt-Griff | Umschließender Griff | Fügen der Schraube in die Bauteile |
| Pinzettengriff | Umschließender Griff | Aufschrauben der Mutter |
| Werkzeuggriff | Umschließender Griff | Anziehen der Schraube |
| Faust | Faust | Beendigung der Montage |

Der verkleinerte Event-Log enthält somit nicht mehr ausschließlich Informationen über die Zeitintervalle, in denen bestimmte Griffklassen ausgeführt wurden, sondern liefert darüber hinaus eine Aussage dazu, welcher Montageschritt in welchem Zeitraum stattgefunden hat. Der aufbereitete Event-Log wird im nächsten Schritt parallel weiterverarbeitet.

Zur Bewertung der Effizienz des betrachteten Montageprozesses wird der aufbereitete Event-Log mittels Methoden des Process Minings im ***Process Mining-Modul*** weitergehend analysiert. Dazu wird mithilfe der Python-Bibliothek *pm4py* ein Induktive Miner eingesetzt. Zunächst wird der Event-Log in ein für *pm4py* lesbares Format gebracht, welches aus drei Spalten besteht:

1. Case_ID: Eindeutiger Identifikator des betrachteten Montageprozesses
2. Activity: Bezeichnung der Montageschritte, die durch die Kombination der Griffklassen der linken und rechten Hand bestimmt wurden (siehe Tabelle 6.3)
3. Time: Zeitstempel, der den Zeitpunkt der Ausführung des jeweiligen Montageschritts dokumentiert

Auf Basis dieses neustrukturierten Event Logs wird durch den Einsatz des Induktiven Miners ein Prozessbaum für die betrachtete Montage erstellt. Da ein Prozessbaum eine eher abstrakte Darstellung der Prozesslogik bietet, erfolgt zur besseren menschlichen Lesbarkeit eine Konvertierung des Prozessbaums in ein Petri-Netz. Ein Beispiel für diese Darstellungsform ist in Abbildung 6.24 visualisiert.

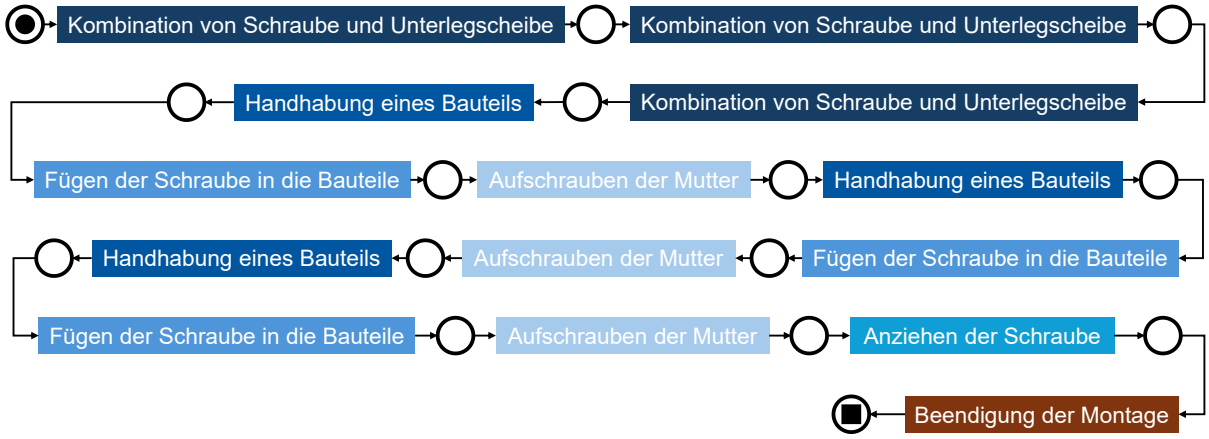



**Abbildung 6.24: Petri-Netz eines per Hand-Tracking erfassten Montageprozesses**

*Petri Net of an Assembly Process Recorded by Hand-Tracking*

Die Analyse der Petri-Netze ermöglicht einen systematischen Vergleich der Montagereihenfolgen unterschiedlicher Mitarbeitender. Dabei werden individuelle Vorgehensweisen in Form von Abweichungen in den Prozessabläufen unmittelbar sichtbar, wodurch Optimierungspotenziale identifiziert werden können.

Ergänzend dazu werden die im Event-Log erfassten Zeitstempel verwendet, um die Effizienz einzelner Montageschritte sowie des Gesamtprozesses zu bewerten. Dazu wird die Dauer pro Montageschritt $\Delta t_i$ bestimmt, indem die Differenz zwischen dem Zeitstempel des nachfolgenden Schritts $t_{i+1}$ und dem Zeitstempel des aktuellen Schritts $t_1$ gebildet wird:

$$\Delta t_i = t_{i+1} - t_i \tag{15}$$

Die Gesamtzeit des Montageprozesses $T_{total}$ ergibt sich als Summe der einzelnen Dauern aller Montageschritte:

$$T_{total} = \sum_{i=1}^{n} \Delta t_i \tag{16}$$

Hierbei bezeichnet $n$ die Anzahl der Montageschritte innerhalb des betrachteten Prozesses. Diese zeitbasierten Analysen ermöglichen Rückschlüsse darauf, wie die Reihenfolge der Montageschritte die Gesamtdauer des Montageprozesses sowie die Dauer der einzelnen Prozessschritte beeinflusst. Die auf diese Weise identifizierten effizientesten Vorgehensweisen können durch die Anpassung der bestehenden Montageanleitungen an alle Mitarbeitenden kommuniziert und als standardisiertes Vorgehen etabliert werden. Damit wird das Erfahrungswissen der Mitarbeitenden systematisch erfasst und nachhaltig im Unternehmen gesichert.

Durch den Einsatz des ***Ergonomie-Moduls*** erfolgt eine Bewertung der physischen Beanspruchung auf den Finger-Hand-Bereich der montierenden Person während der Ausführung des betrachteten Montageprozesses. Hierzu wird der vom Zuweisungs-Modul generierte und aufbereitete Event-Log parallel zur Analyse der Effizienz einer zusätzlichen ergonomischen Untersuchung unterzogen (vergleiche Abbildung 6.23). In Anlehnung an die Leitmerkmalmethode zur Gefährdungsbeurteilung manueller Arbeitsprozesse [KLUß22] wurde im Ergonomie-Modul ein Dictionary implementiert, welches den einzelnen Montageschritten spezifische Kraftkategorien zuordnet. Die resultierenden Zuordnungen für die Referenzbaugruppe sind in Tabelle 6.4 abgebildet.

**Tabelle 6.4: Zuweisung der Montageschritte zu den bei der Ausführung auftretenden Kräften auf den Finger-Hand-Bereich**

*Assignment of the Assembly Steps to the Force Categories Occurring on the Finger-Hand Area During Execution*

| Montageschritt | Kraftkategorie |
|---|---|
| Handhabung eines Bauteils | Mittlere Kräfte |
| Kombination von Schraube und Unterlegscheibe | Mittlere Kräfte |
| Fügen der Schraube in die Bauteile | Geringe Kräfte |
| Aufschrauben der Mutter | Geringe Kräfte |
| Anziehen der Schraube | Hohe Kräfte |
| Beendigung der Montage | Keine Kräfte |

Unter Verwendung des Dictionarys wird jedem Eintrag im Event-Log die entsprechende Kraftkategorie zugeordnet. Zur quantitativen Bewertung der Gesamtbelastung des Prozesses werden die Kraftkategorien mit Gewichtungsfaktoren versehen, um die unterschiedliche Beanspruchung zu berücksichtigen. Die zugeordneten Gewichtungsfaktoren lauten wie folgt:

- Geringe Kräfte: $w_G = 1$
- Mittlere Kräfte: $w_M = 2$
- Hohe Kräfte: $w_H = 3$

Für jeden Montageschritt $i$ kann darauf aufbauend der Belastungswert $B_i$ bestimmt werden:

$$B_i = \Delta t_i \cdot w_i \tag{17}$$

Dabei ist $\Delta t_i$ die Dauer des Montageschritts $i$ in Sekunden, die sich aus den Zeitstempeln der Einträge im Event-Log ergibt. Zur Ermittlung der Gesamtbelastung $B_{total}$ des gesamten Montageprozesses werden die Belastungswerte aller einzelnen Montageschritte aufsummiert. Dies lässt sich wie folgt ermitteln:

$$B_{total} = \sum_{i=1}^{n} B_i = \sum_{i=1}^{n} \Delta t_i \cdot w_i \quad (18)$$

Hierbei bezeichnet $n$ die Gesamtzahl der Montageschritte. Die Gesamtbelastung $B_{total}$ stellt eine quantitative Metrik der physischen Belastung auf den Finger-Hand-Bereich der montierenden Person dar. Ein höherer Wert von $B_{total}$ signalisiert eine höhere körperliche Beanspruchung.

Für eine bessere Vergleichbarkeit der Belastungen unterschiedlicher Montageprozesse wird zusätzlich die durchschnittliche Belastung pro Sekunde berechnet. Dazu wird die Gesamtbelastung $B_{total}$ durch die Gesamtzeit des Prozesses $T_{total}$ geteilt:

$$B_{pro\ Sekunde} = \frac{B_{total}}{T_{total}} \quad (19)$$

Dieses Vorgehen liefert eine ergonomische Bewertung des betrachteten Montageprozesses, welche genutzt werden kann, um physiologisch günstige Arbeitsweisen zu identifizieren und das Risiko von körperlichen Beschwerden und Verletzungen der Mitarbeitenden zu minimieren.

### 6.3.2 Implementierung von Hand-Tracking zur Überführung manueller Bearbeitungsverfahren von Bauteilen in roboterbasierte Prozesse und virtuelle Lernumgebungen

Die in dieser Arbeit umgesetzte Referenzimplementierung des Einsatzes von Hand-Tracking zur Überführung manueller Bearbeitungsverfahren in roboterbasierte Prozesse und virtuelle Lernumgebungen baut zu wesentlichen Teilen auf der bereits in Kapitel 6.3.1 beschriebenen Implementierung zur Erfassung der Aktivitäten der Mitarbeitenden auf. Konkret betrifft dies das Hand-Tracking-Modul der in Abbildung 6.18 vorgestellten Systemarchitektur. Wie in Kapitel 5.3.2 erläutert, erfolgt die in diesem Unterkapitel beschriebene Implementierung unter der Berücksichtigung des spezifischen Anwendungsfall eines manuellen Schleifprozesses zur Erzeugung eines definierten Schliffbildes auf einer Oberfläche. Dieses Beispiel stellt einen Prozess in der Herstellung von Bauteilen für Sondermaschinen im Bereich der Vakuumtechnik dar. In diesem Fall ist insbesondere die präzise Führung des Werkzeugs entscheidend, da die Funktionsflächen so bearbeitet werden müssen, dass die Schleifriefen ausschließlich längs zur Funktionsfläche verlaufen, um potenzielle Vakuumverluste zu verhindern. Aufgrund dieser Fokussierung wird die Erfassung der Aktivitäten der Mitarbeitenden nur im zweidimensionalen Raum durchgeführt. Im Gegensatz zur Erfassung der manuellen Montageprozesse erfolgt in diesem Fall keine Detektion der angewandten Griffklassen, sondern der ausgeführten Bewegungsabläufe. Für die Umsetzung des in Abbildung 5.26 gezeigten Konzeptes wird die in Abbildung 6.25 visualisierte Systemarchitektur verwendet.

Die Referenzimplementierung wurde erneut an dem rekonfigurierbaren Arbeitstisch (siehe Abbildung 6.16) umgesetzt. Analog zu dem beschriebenen Vorgehen in Kapitel 6.3.1 werden die Aktivitäten der Mitarbeitenden mit der XBOX Kinect aufgenommen und durch das Hand-Tracking-Modul analysiert.

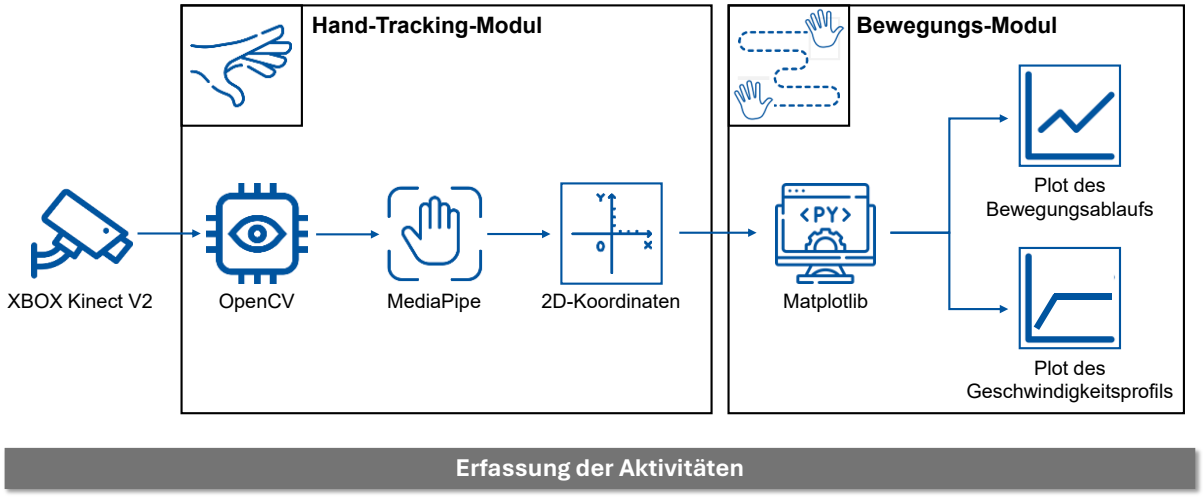



**Abbildung 6.25: Systemarchitektur zur Implementierung der Erfassung der Aktivitäten bei manuellen Bearbeitungsprozessen**

*System Architecture for Implementing the Analysis of Activities of Manual Processing Operations*

Die über *OpenCV* und *Mediapipe* erzeugten zweidimensionalen Koordinatenpunkte der Hand-Landmarks werden anschließend an das Bewegungs-Modul übergeben. Unter Verwendung der Python-Bibliotheken *Math* und *Matplotlib* werden die Koordinatenpunkte der Hand-Landmarks im ersten Schritt innerhalb des ***Bewegungs-Moduls*** in ein geeignetes Koordinatensystem eingetragen. Anschließend werden die jeweils aufeinanderfolgenden Koordinatenpunkte miteinander verbunden, wodurch die Bewegung der Hand als Trajektorie visualisiert wird. In Abbildung 6.26 ist die Visualisierung des Bewegungsablaufs einer rechten Hand während eines manuellen Schleifprozesses dargestellt.

Zur besseren Übersicht ist dabei ausschließlich die Trajektorie der Spitze des Daumens (Hand-Landmark 4) abgebildet. Beide Achsen der Abbildung 6.26 stellen die normalisierten Kamerakoordinaten dar. Die abgebildete Trajektorie kann nun analysiert werden: Der aufgezeichnete Bewegungsablauf zeigt eine kreisförmige Bewegung, die sich translational in X-Richtung fortsetzt. Aus der Darstellung wird ersichtlich, dass bei dem erfassten Bewegungsmuster kein konstanter Radius eingehalten wurde und der Mittelpunkt der Kreisbewegung in Y-Richtung variiert.

Zur Bestimmung der realen Abweichungen der Radien und Mittelpunkte von den Referenzkreisen müssen die normierten Pixelwerte mithilfe eines Maßstabs in reale Einheiten (z.B. Zentimeter) umgerechnet werden. Dazu wird ein Kalibrierungsobjekt mit bekannten Abmessungen als Referenz im betrachteten Arbeitsraum positioniert und mit der Kamera erfasst. Durch eine geeignete Bildanalyse wird die Länge des Kalibrierungsobjektes in Pixeln ermittelt, sodass der Maßstab zur Umrechnung in reale Einheiten bestimmt werden kann:

$$Maßstab = \frac{Reale\ Länge\ des\ Objektes\ [cm]}{Länge\ des\ Objektes\ in\ Pixeln} \tag{20}$$

Alternativ zur Bestimmung des Maßstabs auf Basis der Pixelkoordinaten kann die Berechnung auch analog auf Grundlage der normalisierten Koordinaten erfolgen. Die aus der Darstellung gewonnenen Informationen dienen nicht nur der präzisen Erfassung des Bewegungsablaufs in Form einer mathematischen Beschreibung, sondern liefern auch wertvolle Erkenntnisse zur Optimierung manueller Arbeitsabläufe der Mitarbeitenden. Statt Kreisbewegungen können auch beliebige andere Bewegungsmuster und Abfolgen visualisiert, vermessen und damit quantifiziert werden.

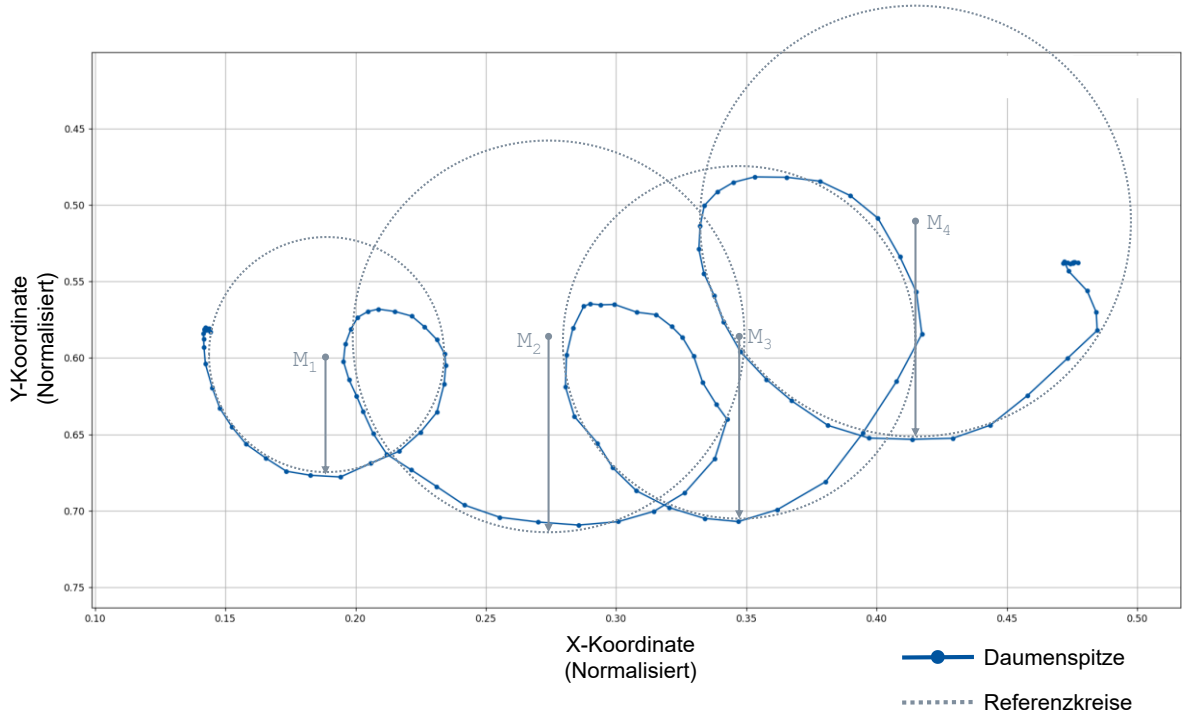



**Abbildung 6.26: Visualisierung eines Bewegungsablaufs einer rechten Hand am Beispiel der Daumenspitze**

*Visualization of an Exemplary Movement Sequence of a Right Hand Using the Tip of the Thumb as an Example*

Im zweiten Schritt innerhalb des Bewegungs-Moduls erfolgt die Bestimmung der Bewegungsgeschwindigkeiten, mit denen der Bewegungsablauf ausgeführt wird. Zur Berechnung der Geschwindigkeit wird die zurückgelegte Strecke zwischen den Koordinatenpunkten $P_{2D,i}(x_{k,i}, y_{k,i})$ einer Hand-Landmark $i$ in zwei aufeinanderfolgenden Bildern $k$ und $k+1$ ermittelt und durch die Zeitdifferenz $\Delta t$ zwischen den Bildern dividiert:

$$v = \frac{\sqrt{(x_{k+1,i} - x_{k,i})^2 + (y_{k+1,i} - y_{k,i})^2}}{\Delta t} \tag{21}$$

Die Zeitdifferenz $\Delta t$ kann aus der Bildfrequenz der Videoaufnahme (Frames pro Sekunde) abgeleitet werden. Sie entspricht dem Kehrwert der Bildfrequenz:

$$\Delta t = \frac{1}{Bildfrequenz} \tag{22}$$

Bei einer Bildfrequenz von 30 Bildern pro Sekunde beträgt $\Delta t$ demnach $\frac{1}{30}$ Sekunde. Für den in Abbildung 6.26 dargestellten Bewegungsablauf lassen sich die zugehörigen Geschwindigkeiten der Handbewegungen zwischen den einzelnen Bildern entsprechend berechnen. In Abbildung 6.27 sind die ermittelten Geschwindigkeiten in normalisierten Einheiten pro Sekunde über der Zeit in Sekunden aufgetragen. Durch die Verbindung der einzelnen Geschwindigkeitswerte entsteht ein Geschwindigkeitsprofil, welches zur weiteren Analyse der Charakteristiken der betrachteten Bewegung herangezogen werden kann. So lassen sich in der dargestellten Kurve beispielsweise vier Segmente ($s_1$, $s_2$, $s_3$, $s_4$) identifizieren, die jeweils einen ähnlichen Verlauf, jedoch mit verschobenen lokalen Maxima und Minima, aufweisen.

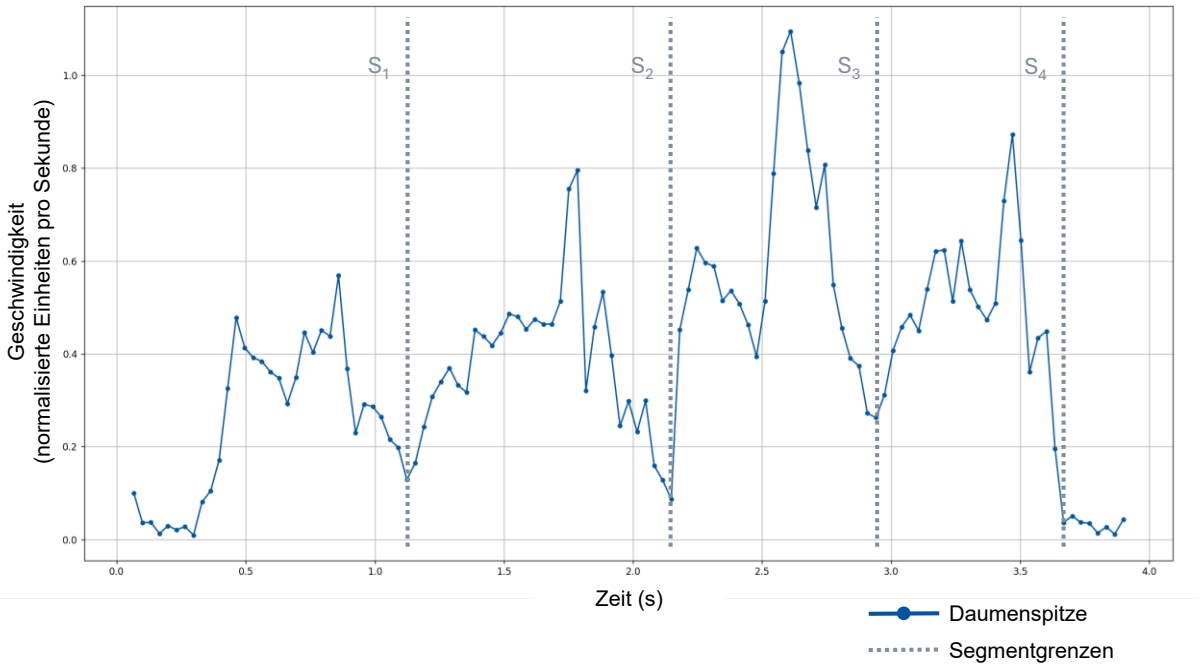



**Abbildung 6.27: Visualisierung eines exemplarischen Geschwindigkeitsprofils einer rechten Hand am Beispiel der Daumenspitze**

*Visualization of an Exemplary Velocity Curve of a Right Hand Using the Tip of the Thumb as an Example*

Die Informationen aus der Visualisierung des Bewegungsablaufs und dem zugehörigen Geschwindigkeitsprofil ermöglichen die Identifikation relevanter Merkmale manueller Bewegungen. Ergänzend zur Analyse der Visualisierungen können die Bewegungsdaten durch den Einsatz mathematischer Methoden weitergehend untersucht werden. Abhängig vom Ziel der Analyse lassen sich dabei Beschleunigungsprofile erstellen oder Bewegungsmuster klassifizieren. In Kombination können die gewonnenen Erkenntnisse gezielt genutzt werden, um die Bewegungscharakteristik in roboterbasierten Prozessen oder virtuellen Lernumgebungen realitätsnah und anwendungsgerecht abzubilden.

# 7 Validierung und Evaluation

*Validation and Evaluation*

Dieses Kapitel widmet sich der Untersuchung und Bewertung der in dieser Arbeit entwickelten Ansätze zur Unterstützung des Wissenstransfers auf dem Shopfloor. In Kapitel 7.1 werden die Ergebnisse der technischen Validierung der in Kapitel 6 beschriebenen Referenzimplementierungen des *aktiven, semi-aktiven und passiven Wissenstransfers* präsentiert. Darauf aufbauend erfolgt in Kapitel 7.2 eine Evaluation der Konzepte im Hinblick auf die in Kapitel 4 definierten Anforderungen. Abschließend wird in Kapitel 7.3 eine vergleichende Bewertung der Ansätze vorgenommen, wobei diese sowohl in Bezug auf die übergeordnete Zielsetzung der Arbeit als auch hinsichtlich ihrer praktischen Anwendbarkeit und Relevanz eingeordnet werden. Dabei wird aufgezeigt, in welchem Umfang die vorgestellten Konzepte und deren Umsetzung zur Beantwortung der zentralen Forschungsfragen dieser Dissertation beitragen.

## 7.1 Validierung der Funktionalitäten

*Validation of Functionalities*

Zur technischen Validierung der drei Referenzimplementierungen wurden primär Funktionstests auf Systemebene durchgeführt, um die korrekte Interaktion der Softwaremodule und die fehlerfreie Verarbeitung von Benutzereingaben zu prüfen. Ergänzend erfolgten gezielte Validierungen einzelner Module sowie Vergleiche der Leistungsfähigkeit der unterschiedlichen Implementierungsvarianten je Konzept, um die geeignetsten Technologiekombinationen und Varianten zu identifizieren. Entsprechend des zugrunde liegenden Design-Thinking-Ansatzes wurden die Tests bewusst als formative Evaluation mit wenigen, aber repräsentativen Durchläufen konzipiert. Ziel war es, frühzeitig praxisnahe Rückmeldungen zur Funktionsweise zu gewinnen, zentrale Fehlermuster zu identifizieren und Optimierungspotenziale aufzuzeigen. Eine statistisch abgesicherte Generalisierbarkeit wurde dabei nicht angestrebt.

### 7.1.1 Validierung der Referenzimplementierung des aktiven Wissenstransfers

*Validation of the Reference Implementation of Active Knowledge Transfer*

Das Konzept des aktiven Wissenstransfers wurde in zwei Varianten umgesetzt (siehe Kapitel 6.1), deren Konzeptionierung und Implementierung mit einem zeitlichen Versatz von etwa einem Jahr durchgeführt wurden. Aufgrund dessen erfolgte die Validierung initial jeweils separat und wurde retrospektiv verglichen.

Die Validierung der Funktionalität des auf dem Rasa-Framework basierenden Chatbots zielte darauf ab, die Korrektheit und Zuverlässigkeit der Interaktion zwischen Benutzer, Chatbot und SQL-Datenbank sicherzustellen. Als Methodik wurde zunächst ein Black-Box-Testansatz gewählt, bei dem funktionale Tests durchgeführt wurden, die sich auf die Verifizierung des Ein- und Ausgabeverhaltens des in Kapitel 6.1.1 vorgestellten Systems konzentrierten, ohne die internen Abläufe einzelner Module direkt zu analysieren. Die vom Chatbot generierten Antworten wurden systematisch mit den tat-

sächlich in der SQL-Datenbank gespeicherten Informationen abgeglichen, um mögliche Abweichungen zu identifizieren. Zusätzlich wurde geprüft, ob neu eingegebene Informationen korrekt in der SQL-Datenbank hinterlegt wurden, um eine konsistente Datenintegrität sicherzustellen.

Während der ersten Tests zeigte sich, dass das Systemverhalten in mehreren Fällen von den erwarteten Ergebnissen abwich. Der Chatbot reagierte häufig mit der Rückmeldung, die Benutzereingabe nicht verstanden zu haben oder gab bei Anfragen fälschlicherweise an, dass keine entsprechenden Daten in der SQL-Datenbank vorlägen. Mithilfe der grafischen Benutzeroberfläche der Web-Applikation, in der die Transkripte der über das Sprachinterface eingegebenen Benutzereingaben dargestellt werden (siehe Abbildung 6.2), wurde festgestellt, dass die Benutzereingaben teilweise nicht korrekt erfasst wurden. Infolgedessen wurde eine detaillierte Analyse des Voice-Interface-Moduls sowie des Speech-to-Text-Moduls durchgeführt. Dabei stellte sich heraus, dass die Verarbeitung von Fehlercodes durch das Speech-to-Text-Modul unzureichend implementiert war, was zu Fehlinterpretationen und Missverständnissen führte. Um diese Probleme zu beheben, wurde eine Textnormalisierung auf das vom Speech-to-Text-Modul erzeugte Transkript angewandt. Diese umfasste die Konvertierung ausgeschriebener Zahlen in numerische Darstellungen sowie die Korrektur der Groß- und Kleinschreibung bei Zeichenfolgen. Das überarbeitete Vorgehen wird in Tabelle 7.1 anhand eines Beispiels verdeutlicht.

**Tabelle 7.1: Maßnahmen der Textnormalisierung zur Optimierung des Transkripts**

*Text normalization measures to optimize the transcript*

| Transkript ohne Textnormalisierung | Transkript mit Textnormalisierung |
|---|---|
| eins sechs achtzehn | 1618 |
| a | A |
| a eins sechs achtzehn | A1618 |

Durch diese Maßnahme konnte die Anzahl der Missverständnisse des Chatbots signifikant reduziert werden. Dennoch blieb eine gewisse Unschärfe bei der Transkription der erfassten Spracheingaben bestehen. Insbesondere Maschinenbezeichnungen und Fehlercodes wurden wiederholt fälschlicherweise als Wörter interpretiert und dadurch fehlerhaft transkribiert. Dieses Verhalten ist auf das verwendete Speech-to-Text-Modell *DeepSpeech* zurückzuführen, das nicht speziell für die Erkennung alphanumerischer Fehlercodes trainiert wurde. Die Leistung des Modells hängt maßgeblich von der Qualität und der spezifischen Relevanz der zugrunde liegenden Trainingsdaten ab. Da alphanumerische Fehlercodes sowie Eigennamen, beispielsweise von Werkzeugmaschinenherstellern, in den Trainingsdaten des Modells nicht ausreichend repräsentiert sind, treten bei der Transkription entsprechende Probleme auf.

Zur Vermeidung der zuvor beschriebenen Probleme wurden zusätzlich Tests ohne Verwendung des Sprachinterfaces durchgeführt. Die Anfragen und Eingaben durch den Benutzer wurden stattdessen direkt als Text übermittelt. Wie erwartet, traten die

zuvor beobachteten Missverständnisse hierbei nicht auf. Insgesamt wurden einhundert Abfragen und Eingaben an das System gestellt und die Ergebnisse systematisch überprüft. Sofern keine Fehler bei der Texteingabe (z.B. Zahlendreher, Tippfehler bei Buchstaben) auftraten, verhielt sich das System wie erwartet und führte sowohl Abfragen als auch Eingaben korrekt aus. Im Verlauf der Tests zeigte sich jedoch, dass das System erhebliche Schwankungen in den Antwortzeiten aufwies. Diese variierten zwischen fünf Sekunden und einer Minute. Ein eindeutiger Kausalzusammenhang zwischen den Benutzereingaben und den Schwankungen in der Antwortzeit konnte nicht festgestellt werden. Selbst identische Benutzereingaben führten zu unterschiedlich langen Antwortzeiten. Dabei lagen die Abweichungen im Bereich von ein bis zehn Sekunden.

Die angestrebte Funktionalität des auf dem Rasa-Framework basierenden Chatbots konnte insgesamt erfolgreich implementiert werden. Jedoch ist die vorliegende Referenzimplementierung in ihrer aktuellen Form nur eingeschränkt für den praktischen Einsatz geeignet, da die teilweise langen Antwortzeiten die Benutzererfahrung erheblich beeinträchtigen und somit die Akzeptanz des Systems negativ beeinflussen.

Die in Kapitel 6.1.2 vorgestellte Systemarchitektur zur Realisierung der Funktionen des Chatbots auf Basis von LLM umfasst zwei zentrale Komponenten, die maßgeblich das Systemverhalten und die Qualität der Nutzerinteraktion beeinflussen (siehe Abbildung 6.6): Das Sentence Transformer Modell und das LLM selbst. Das Sentence Transformer Modell bildet die Grundlage für den semantischen Vergleich durch FAISS, indem es Embeddings sowohl der hochgeladenen PDF-Dokumente als auch der Benutzereingaben erstellt. Die Qualität dieser Embeddings ist entscheidend für die Leistungsfähigkeit der FAISS-Ähnlichkeitssuche: Je differenzierter die semantischen Zusammenhänge durch die Embeddings repräsentiert werden, desto präziser können relevante Inhalte identifiziert werden. Das LLM übernimmt die Aufgaben der Klassifizierung der Benutzereingaben sowie der Generierung systemseitiger Antworten auf Benutzerfragen. Die Leistungsfähigkeit des LLMs hängt insbesondere von der Fähigkeit ab, komplexe Texte semantisch zu analysieren und relevante Informationen zielgerichtet zu extrahieren.

Während der Entwicklung des Systems wurden verschiedene Ansätze mit unterschiedlichen LLM und Alternativen zu Sentence Transformers untersucht, bevor schließlich die in Abbildung 6.7 dargestellte Systemarchitektur entstand. Dieser Prozess erforderte mehrere Iterationen von Systemtests und Anpassungen, um die optimale Konfiguration zu erreichen. Zu Beginn der Implementierungsphase wurde erwogen, die Klassifizierung der Benutzereingaben, die Generierung der Antworten sowie die Erstellung der Embeddings für PDF-Dokumente und Benutzereingaben vollständig durch ein zentrales LLM auszuführen. Dieser Gedanke führte zum Vergleich der drei Modelle *Llama-2-70b-chat*, *PaLM-2* und *GPT-4*, welche laut dem damaligen Stand der Technik als leistungsstarke Modelle galten. Ziel des Vergleichs war es, die Eignung der Modelle für den Einsatz als zentrales Element des Chatbots zu bewerten. Diese Untersuchung wurde in einer Vorversuchsreihe losgelöst von der später entwickelten Systemarchitektur durchgeführt.

Das LLaMA-Modell wurde auf der vorhandenen Workstation (siehe Tabelle 6.1) implementiert, während PaLM-2 und GPT-4 aufgrund fehlender Open-Source-Versionen über die von Google und OpenAI bereitgestellten Chatbots BARD und ChatGPT getestet wurden. Für das lokal implementierte LLaMA-Modell wurde zur Bereitstellung des PDFs als Wissensgrundlage der Ansatz der RAG (siehe Kapitel 2.2.2) adaptiert und lokal mithilfe des GPT4ALL-Python-SDKs sowie des Vector Store Chroma umgesetzt. Zum Testen von PaLM-2 und GPT-4 wurden bereits vorhandene Funktionen der Chatbots genutzt. BARD ermöglichte den Zugriff auf ein PDF-Dokument über die „Google Workspace“-Erweiterung aus einem Google Drive-Verzeichnis, während bei ChatGPT das Plugin „AskYourPDF“ zum Einsatz kam.

Alle drei Testsysteme wurden nach einem einheitlichen Verfahren evaluiert: Den LLM wurden per Prompt Fragen gestellt, die sie ausschließlich auf Basis der Informationen aus einem hinterlegten PDF beantworten sollten. Neben der Beantwortung der Fragen wurden die LLM aufgefordert, die entsprechende Quelle in Form des Textabschnitts aus dem PDF zu nennen, der zur Beantwortung herangezogen wurde. Als Testdokument diente das Skript der ersten Vorlesung des Fachs „Werkzeugmaschinen“, das im Sommersemester 2022 in der Lehre der RWTH Aachen University eingesetzt wurde. In Tabelle 7.2 sind fünf der gestellten Fragen, die Antworten (A) der drei Modelle sowie die Bewertung, ob die korrekte Quelle bzw. Textstelle aus dem PDF (Q) verwendet wurde, exemplarisch dargestellt. Richtig beantwortete Fragen und identifizierte Quellen sind in grüner, falsche in roter Schrift hervorgehoben. Die Einordnung der Antworten als „richtig“ oder „falsch“ erfolgte durch eine manuelle Prüfung der faktischen Genauigkeit. Dabei wurden bereits geringfügige Abweichungen, etwa Rundungsdifferenzen bei Zahlenangaben, als fehlerhaft klassifiziert, um eine konsistente Bewertung sicherzustellen. Die ergänzende Analyse der jeweils genutzten Textstellen hatte zum Ziel, Halluzinationen seitens der LLM zu identifizieren und Zufallstreffer ohne inhaltliche Grundlage auszuschließen.

Die Übersicht zeigt, dass das Testsystem mit *GPT-4* alle fünf Fragen korrekt beantwortete und dabei stets die relevante Textstelle zur Beantwortung heranzog. Im Gegensatz dazu lieferten die Systeme mit *PaLM-2* und *LLaMA-2-70B-Chat* mehrere fehlerhafte Antworten, was darauf zurückzuführen war, dass die relevanten Textpassagen im hinterlegten Dokument nicht korrekt identifiziert wurden.

Aus diesen Ergebnissen wurden zwei zentrale Erkenntnisse abgeleitet: Alle drei LLM sind grundsätzlich in der Lage, Fragen zu einem Text korrekt zu beantworten, sofern sich die Frage eindeutig auf die Inhalte des vorliegenden Texts bezieht. Die korrekte Identifikation der für die Beantwortung relevanten Textpassagen spielt beim Hinterlegen textbasierter Dokumente wie PDFs eine entscheidende Rolle. Das LLM kann nur dann eine korrekte Antwort liefern, wenn es die notwendigen Informationen aus den richtigen Textpassagen erhält.

**Tabelle 7.2: Vergleich der von den drei evaluierten LLM gegebenen Antworten auf Fragen bezüglich eines hinterlegten Informationstextes**

*Comparison of the Answers Given by the Three Evaluated LLM to Questions Relating to a Stored Information Text*

| Fragen | GPT-4 (ChatGPT) | PaLM-2 (BARD) | Llama-2-70b-chat |
|---|---|---|---|
| Pressengestelle von Reibspindelpressen werden häufig über was vorgespannt? | A: Zuganker<br>Q: Richtig | A: Zuganker<br>Q: Richtig | A: Zuganker<br>Q: Richtig |
| Wie hoch war der Produktionswert von Werkzeugmaschinen in Deutschland im Jahr 2009? | A: 7,7 Milliarden Euro<br>Q: Richtig | A: 2,56 Milliarden Euro<br>Q: Falsch | A: 46,7 Milliarden Euro<br>Q: Falsch |
| Was ist das Funktionsprinzip von Hämmern? Energiegebunden, weggebunden oder kraftgebunden? | A: Energiegebunden<br>Q: Richtig | A: Energiegebunden<br>Q: Falsch | A: Energiegebunden<br>Q: Richtig |
| In welchem Bereich liegen die Betriebsdrücke von hydraulischen Pressen? | A: 200 bis 300 Bar<br>Q: Richtig | A: 50 bis 300 Bar<br>Q: Falsch | A: 50 bis 200 Bar<br>Q: Falsch |
| Was kommt als Überlastschutz an Umformmaschinen zum Einsatz? | A: Mechanische Sollbruchstellen und hydraulische Systeme<br>Q: Richtig | A: Mechanische Sollbruchstellen und hydraulische Systeme<br>Q: Richtig | A: Mechanische, elektrische und hybride Schutzsysteme<br>Q: Falsch |

Auf Grundlage dieser Erkenntnisse wurde in der nächsten Iteration der Implementierungsphase der Fokus auf die Qualität der Erstellung der Embeddings sowie die Leistungsfähigkeit des verwendeten Vector Stores gelegt. Nach gezielter Recherche wurden die Sentence Transformer-Modelle und FAISS als geeignet für das angestrebte System identifiziert, woraufhin die in Abbildung 6.6 visualisierte Systemarchitektur entwickelt wurde. Zur Validierung dieser finalen Systemarchitektur wurden zusätzliche Versuchsreihen durchgeführt. Analog zur Validierung des Chatbots auf Basis des Rasa-Frameworks wurde ein Black-Box-Testansatz gewählt, um das Ein- und Ausgabeverhalten des Systems zu prüfen. Als LLM wurde dabei aufgrund der verfügbaren Open-Source-Version *Llama-2-70b-chat* angebunden. Dem System wurde erneut das Skript der ersten Vorlesung des Fachs „Werkzeugmaschinen“ als PDF übergeben und die in Tabelle 7.2 aufgeführten Fragen gestellt. In dieser Konfiguration wurden alle Fragen korrekt beantwortet und die jeweils relevanten Textpassagen präzise identifiziert.

Zusätzlich wurde das Verhalten des Systems in Bezug auf die zunehmende Größe der hinterlegten Wissensbasis untersucht. Zu diesem Zweck wurde zunächst das Buch

„Introduction to Mechanical Engineering“ von J.Paulo Davim mit einem Umfang von 372 Seiten und anschließend das Buch „Werkzeugmaschinen Fertigungssysteme 3“ von Christian Brecher und Manfred Weck mit 718 Seiten als Wissensbasis hinterlegt. Analog zum Vorgehen mit dem Skript wurden auch hier Fragen zu den in den Büchern enthaltenen Informationen gestellt. Dabei zeigte sich, dass die Antwortzeiten des Systems mit der Größe der Wissensbasis anstiegen. Insbesondere bei den Tests mit dem 718-seitigen Buch wurden Antwortzeiten von bis zu 20 Sekunden gemessen. Die generierten Antworten blieben korrekt, sofern sich die gestellten Fragen auf Textpassagen aus dem Buch und nicht auf Inhalte in Abbildungen bezogen.

Die Eingabe neuen Wissens verlief in den Tests problemlos: Das LLM erkannte die Intention zur Wissensspeicherung zuverlässig und die Informationen wurden wie vorgesehen im Vector Store sowie in der Knowledge Base abgelegt. Bei entsprechenden Abfragen nutzte der Chatbot die neuen Einträge zur Beantwortung. Im Fall widersprüchlicher Informationen zeigte das System jedoch ein inkonsistentes Verhalten: Es nannte entweder eine der möglichen Antworten oder verwies auf fehlende Eindeutigkeit. Eine Kausalität für die Wahl des Verhaltens konnte nicht identifiziert werden, was auf die nicht deterministische Natur des LLMs (siehe Kapitel 2.2.1) zurückzuführen ist.

Die angestrebte Funktionalität des LLM-basierten Chatbots konnte damit erfolgreich umgesetzt werden. Für den praktischen Einsatz in KMU sind jedoch weiterführende Tests unter Berücksichtigung spezifischer Rahmenbedingungen und eingesetzter Speichersysteme erforderlich. Die in dieser Arbeit gewonnenen Erkenntnisse deuten darauf hin, dass sowohl die Größe der Wissensbasis als auch die Qualität und Korrektheit der Informationen maßgeblich die Effektivität und Nutzbarkeit des Chatbots beeinflussen. Im Vergleich zum Rasa-basierten System zeigte der LLM-gestützte Chatbot eine deutlich höhere Leistungsfähigkeit, insbesondere in Form kürzerer Antwortzeiten und einer als natürlicher empfundenen Interaktion. Aufgrund dieser Ergebnisse wird der LLM-basierte Chatbot als die geeignetere Lösung zur Unterstützung des aktiven Wissenstransfers bewertet.

### 7.1.2 Validierung der Referenzimplementierung des semi-aktiven Wissenstransfers

*Validation of the Reference Implementation of Semi Active Knowledge Transfer*

Die Validierung der Referenzimplementierung des semi-aktiven Wissenstransfers (siehe Kapitel 6.2) erfolgte in zwei Phasen. Zunächst wurde das System unter kontrollierten Bedingungen im SAL des WZLs getestet, um die grundlegenden Funktionen zu verifizieren. Anschließend wurde die Validierung anhand eines Prozesses in einem realen Unternehmensumfeld durchgeführt, um die Anwendbarkeit und Effektivität des Systems unter praxisnahen Bedingungen zu bewerten. In beiden Phasen der Validierung wurden drei unterschiedliche Whisper-Modelle zur Transkription der Tonspur eingesetzt und die resultierenden Texte auf Fehler geprüft. Für die Bildanalyse wurden verschiedene Konfidenzniveaus des YOLOv8-Modells bei der Objektdetektion getestet, um den Einfluss auf mögliche Fehldetektionen zu untersuchen. Abschließend wurden die Ergebnisse der automatisierten Video-Segmentierung mit einer manuellen

Segmentierung durch einen Experten verglichen, um die Leistungsfähigkeit und Präzision der automatisierten Methoden zu beurteilen.

In der ersten Phase der Validierung im SAL wurde die Referenzimplementierung zur Erfassung und Analyse der Endmontage eines Tisches bestehend aus item-Komponenten angewendet. Während der Montage wurde an einem bereits vormontierten Gestell eine zusätzliche Strebe zur Versteifung angebracht und anschließend die Tischplatte verschraubt. Dabei kamen ein Schraubendreher, ein Akkuschrauber und ein Schonhammer zum Einsatz. Zusätzlich wurde eine Zange im Hintergrund platziert, um die Robustheit der Objekterkennung gezielt zu testen. Wie im Konzept in Kapitel 5.2.1 beschrieben, erklärte der Monteur bei der Aufnahme sein Vorgehen. Die Dauer der Aufnahme betrug 5 Minuten, 49 Sekunden und 77 Millisekunden, wobei die mündlichen Erklärungen des Monteurs insgesamt 150 Wörter umfassten. Diese Wortanzahl wurde manuell aus dem Video extrahiert und diente als Referenz. Zum Vergleich der Leistungsfähigkeit der drei unterschiedlichen Whisper-Modelle wurde jeweils die Word Error Rate (WER) über den Quotienten der Summe aller falsch erkannten Wörter (Substitutionen und Insertionen) sowie der nicht erkannten Wörter (Deletionen) und der Gesamtzahl der tatsächlich gesprochenen Wörter (150) berechnet:

$$WER = \frac{Substitutionen + Deletionen + Insertionen}{Gesamtzahl\ gesprochener\ Worte} * 100 \qquad (23)$$

Die Ergebnisse der Berechnung und die Transkriptionszeiten sind in Tabelle 7.3 dargestellt.

**Tabelle 7.3: Ergebnisse der für die Transkription verwendeten Whisper-Modelle bei der Validierung im SAL**

*Results of the Whisper Models Used for the Transcription in the Validation in SAL*

| Modell | Anzahl erkannter Worte | Fehlerhaft transkribierte Worte | WER [%] | Transkriptionszeit [s] |
|---|---|---|---|---|
| Whisper Medium | 148 | 3 Substitutionen | 2,03 | 30,86 |
| Whisper LargeV2 | 146 | 1 Substitution | 0,67 | 28,21 |
| Whisper LargeV3 | 148 | 3 Substitutionen | 2,03 | 28,93 |

Die drei getesteten Whisper-Modelle weisen in allen Kategorien sehr ähnliche Leistungen auf und erreichen eine sehr gute WER von unter 5 %. Auffällig ist jedoch, dass keines der Modelle, alle gesprochenen Wörter fehlerfrei erkannte. Die nicht erkannten Wörter variierten zwischen den Modellen und zeigten keine systematischen Gemeinsamkeiten, sodass kein kausaler Zusammenhang identifiziert werden konnte. Bei den fehlerhaft transkribierten Wörtern liegt stets eine Substitution durch phonetisch ähnliche Begriffe vor. Ein Beispiel hierfür ist die Transkription von „Endkappen“ als „Handkappen“. Dieses Verhalten lässt darauf schließen, dass die Modelle Wörter, die sie nicht explizit gelernt haben, durch phonetisch ähnliche Alternativen ersetzen.

Im Rahmen der Bildanalyse mit dem YOLOv8-Modell wurde das Video in 10.493 einzelne Bilder zerlegt, die auf das Vorhandensein der erlernten Werkzeugklassen untersucht wurden. Die festgelegte Schwelle für die Confidence des YOLOv8-Modells wurde während der Validierung variiert, um deren Einfluss auf die Detektion zu analysieren. Zur Bewertung der Leistungsfähigkeit der Bildanalyse werden die Metriken *Precision* und *Recall* herangezogen:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} * 100 \tag{24}$$

$$Recall = \frac{True\ Positives}{Ground\ Truth} * 100 \tag{25}$$

Dabei bezeichnet True Positives die korrekt erkannten Werkzeugklassen, False Positives die fälschlicherweise erkannten Werkzeugklassen und die Ground Truth die Anzahl der tatsächlich im Video enthaltenen Werkzeugklassen (in diesem Fall vier). Die Ergebnisse der Werkzeugdetektion sind in Tabelle 7.4 aufgeführt.

**Tabelle 7.4: Ergebnisse der Werkzeugdetektion bei der Validierung im SAL**

*Results of Tool Detection During Validation in the SAL*

| Confidence | True Positives | False Positives | Precision [%] | Recall [%] |
|---|---|---|---|---|
| 0,667 | 4 | 20 | 16,67 | 100 |
| 0,75 | 4 | 10 | 28,57 | 100 |
| 0,85 | 4 | 10 | 28,57 | 100 |

Mit zunehmender Confidence verringerte sich zwar die Anzahl der False Positives, unabhängig vom Confidence-Niveau wurden jedoch stets False Positives ermittelt. Die True Positives wurden dabei stets vollständig und korrekt identifiziert. Die hohe Zahl der False Positives und die daraus resultierende geringe Precision ist auf die Vielfalt der im Trainingsdatensatz (METU-ALET-Datensatz [KURN19], siehe auch Kapitel 6.2) enthaltenen Werkzeuge zurückzuführen. Viele dieser Werkzeuge weisen eine hohe visuelle Ähnlichkeit auf, wodurch das Modell insbesondere bei variierenden Blickwinkeln Schwierigkeiten bei der korrekten Klassifikation hatte. Eine gezielte Erweiterung des Trainingsdatensatzes um repräsentative Bilder der für den spezifischen Anwendungsfall relevanten Werkzeugklassen, die eine hohe Varianz hinsichtlich Blickwinkel, Beleuchtungsbedingungen sowie unterschiedlicher Ausprägungen der Werkzeugeigenschaften (z.B. Farbe und Form der Griffe) abdecken, sollte die Differenzierung ähnlicher Klassen signifikant verbessern und somit die Erkennungsgenauigkeit steigern.

Grundsätzlich kann eine zuverlässige Bildanalyse die Zeitpunkte von Werkzeugwechseln identifizieren und damit potenzielle Schnittmarken für die Videosegmentierung liefern. In den meisten Fällen folgen die Werkzeugwechsel allerdings den standardisierten Arbeitsanweisungen und reflektieren somit keine individuellen Erfahrungswerte. Zwar kann ein unkonventioneller Werkzeugwechsel auf das Erfahrungswissen eines

Mitarbeitenden hindeuten, jedoch ist davon auszugehen, dass ein solcher Schritt auch verbal begleitet und durch explizite Erläuterungen begründet wird. Die verbalen Äußerungen enthalten darüber hinaus oft weiterführende, prozessbezogene Informationen, die weit über das visuell Erfassbare hinausgehen. Aufgrund der geringen Precision der implementierten Bildanalyse sowie der inhaltlich größeren Relevanz der Tonspur wurde die Sequenzierung des Videos im weiteren Verlauf ausschließlich auf Basis der Transkription durchgeführt.

Entsprechend der in Abbildung 6.8 dargestellten Systemarchitektur kamen hierfür die semantische Analyse mittels der Schlüsselworterkennung sowie des Sentence Transformer-Modells zum Einsatz. Das Video wurde zudem vom Monteur manuell sequenziert, um eine Referenz (Ground Truth) zu schaffen. Die Ergebnisse aller drei Sequenzierungen sind im Anhang in Tabelle E.7 aufgeführt. Zur Bewertung der Leistungsfähigkeit beider automatisierter Sequenzierungsmethoden wird der durchschnittliche Grad der Übereinstimmung der automatisiert erstellten Sequenzen mit den manuell erzeugten Sequenzen bestimmt. Dies geschieht über die Berechnung der Metrik „Intersection of Union“ (IoU), welche die zeitliche Überlappung zweier Sequenzen ins Verhältnis zu deren vereinter Gesamtlänge setzt. Die IoU berechnet sich für ein Sequenzpaar nach der folgenden Formel:

$$IoU = \frac{Overlap}{Union} \tag{26}$$

mit

$$Overlap = \max(0, \min(p_{Ende}, q_{Ende}) - \max(p_{Start}, q_{Start})) \tag{27}$$

und

$$Union = (p_{Ende} - p_{Start}) + (q_{Ende} - q_{Start}) - Overlap \tag{28}$$

Dabei gilt:

$$Segment_{Manuell} = [\, p_{Anfang}, p_{Ende} \,] = [Startzeit, Endzeit] \tag{29}$$

$$Segment_{Auto} = [\, q_{Anfang}, q_{Ende} \,] = [Startzeit, Endzeit] \tag{30}$$

Um die gesamte Sequenzierung zu bewerten, wird anschließend der Mittelwert der berechneten IoU-Werte über alle Sequenzpaare gebildet:

$$Mean\ IoU = \frac{1}{N} \sum_{i=1}^{N} IoU_i \tag{31}$$

Da die manuelle Sequenzierung und die beiden automatisierten Sequenzierungen jeweils eine unterschiedliche Anzahl an Sequenzen erzeugen, wird für jede manuelle Sequenz zunächst die am besten passende automatisch erzeugte Sequenz identifiziert. Anschließend werden für jedes so gebildete Sequenzpaar gemäß den Formeln (26) bis (30) die Metriken Overlap, Union und IoU berechnet. Die Ergebnisse sind im Anhang in Tabelle E.8 (Schlüsselwort-basierte Sequenzierung) und Tabelle E.9 (Sequenzierung mittels Sentence Transformer) dargestellt. Unter Verwendung von Formel (31) resultiert aus diesen Werten ein Mean IoU von 0,605 für die Schlüsselwort-

basierte Sequenzierung und ein Mean IoU von 0,374 für die Sequenzierung mittels Sentence Transformer. Der Vergleich dieser Werte zeigt, dass die Schlüsselwort-basierte Sequenzierung eine deutlich höhere Übereinstimmung mit der manuellen Sequenzierung aufweist als die auf Sentence Transformers basierende Variante. Ein Mean IoU von 0,605 stellt jedoch nur eine moderate Übereinstimmung zur manuellen Referenz dar und bietet noch Verbesserungspotenzial.

Die zweite Phase der Validierung wurde in Zusammenarbeit mit einem KMU durchgeführt, das als Projektpartner im Forschungsprojekt AKzentE4.0 beteiligt ist. Dazu erhielt ein Mitarbeiter des Unternehmens eine kurze Einweisung in die Nutzung der Microsoft HoloLens 2 (siehe Abbildung 6.9) und bekam anschließend die Instruktion, seinen Arbeitsprozess wie gewohnt durchzuführen und dabei besonders relevante Aktivitäten durch Hinweise und zu beachtende Aspekte zu erläutern. Der untersuchte Prozess umfasste das Entgraten der Kanten und Bohrungen eines Kranzes, der später Bestandteil einer Beschichtungsmaschine ist.[6] Während des Arbeitsablaufs kamen ein Turboschleifer, ein Akkuschrauber mit Kegelsenker sowie ein Dreikantschaber zum Einsatz. Zusätzlich sind auf einigen Passagen des Videos eine Zange und ein Schraubenschlüssel im Hintergrund zu sehen, welche jedoch keine Relevanz für den betrachteten Prozess haben. Das aufgezeichnete Video hat eine Länge von 13 Minuten, 6 Sekunden und 23 Millisekunden, während die zugehörige Tonspur insgesamt 331 Wörter umfasst. Tabelle 7.5 stellt die Ergebnisse der Transkription dar.

**Tabelle 7.5: Ergebnisse der für die Transkription verwendeten Whisper-Modelle bei der Validierung im KMU**

*Results of the Whisper Models Used for the Transcription in the Validation in SME*

| Modell | Anzahl erkannter Worte | Fehlerhaft transkribierte Worte | WER [%] | Transkriptionszeit [s] |
|---|---|---|---|---|
| Whisper Medium | 339 | 8 Insertionen, 9 Substitutionen | 4,78 | 42,69 |
| Whisper LargeV2 | 331 | 7 Substitutionen | 2,07 | 28,21 |
| Whisper LargeV3 | 331 | 7 Substitutionen | 2,07 | 28,93 |

Trotz der Durchführung des semi-aktiven Wissenstransfers in einer produktiven Werkstattumgebung mit zahlreichen Störgeräuschen (z.B. Radiolärm und Geräusche anderer Tätigkeiten) transkribierten die beiden Large Whisper-Modelle ausschließlich die Erklärungen des Mitarbeiters, der den betrachteten Prozess ausführte. Im Gegensatz dazu erkannte das Medium-Modell fälschlicherweise acht Wörter aus dem Hintergrund und fügte diese in die Transkription ein. Insgesamt erreichen alle Modelle erneut eine sehr gute WER von unter 5 %. Die fehlerhaft transkribierten Wörter resultierten erneut aus phonetischen Ähnlichkeiten. Beispielsweise wurde das vom Mitarbeiter verwendete Wort „Fase“ von allen drei Modellen als „Phase“ interpretiert. Hinsichtlich der

[6] Zur Wahrung von Betriebsgeheimnissen dürfen keine Bilder dieses Bauteils gezeigt werden.

Transkriptionszeiten zeigte lediglich das Medium-Modell einen leichten Leistungsabfall, während die beiden Large-Modelle im Vergleich zum unter Laborbedingungen im SAL betrachteten Prozess konstante Zeiten aufwiesen (vergleiche Tabelle 7.3).

Die Ergebnisse der Bildanalyse des in 23.587 Bilder zerlegten Videos sind in Tabelle 7.6 zusammengefasst. Die Metriken wurden analog zum Vorgehen für Tabelle 7.4 mittels der Formeln (24) und (25) berechnet, wobei die Ground Truth in diesem Fall fünf betrug. Von den drei prozessrelevanten Werkzeugen konnte nur der Akkuschrauber korrekt detektiert werden. Zusätzlich beinhalten die drei True Positives die im Hintergrund befindlichen, für die Auswertung irrelevanten Werkzeuge Zange und Schraubenschlüssel. Grundsätzlich zeigt sich, dass mit steigender Confidence die Anzahl False Positives abnimmt, was gleichzeitig zu einer Steigerung der Precision führt. Insgesamt ist die erreichte Precision jedoch unzureichend.

**Tabelle 7.6: Ergebnisse der Werkzeugdetektion bei der Validierung im KMU**

*Results of Tool Detection During Validation in SME*

| Confidence | True Positives | False Positives | Precision [%] | Recall [%] |
|---|---|---|---|---|
| 0,667 | 3 | 31 | 8,82 | 60,00 |
| 0,75 | 3 | 27 | 10,00 | 60,00 |
| 0,85 | 3 | 16 | 15,79 | 60,00 |

Da die Werkzeugerkennung auch in dieser zweiten Phase der Validierung keine ausreichende Leistungsfähigkeit liefert, wird die Sequenzierung des Videos erneut ausschließlich auf Basis der Schlüsselworterkennung und des Sentence Transformer-Modells durchgeführt und mit der vom Mitarbeiter vorgenommenen, manuellen Sequenzierung verglichen. Die Ergebnisse der Sequenzierungen sind im Anhang in Tabelle E.10 dargestellt. Analog zum Vorgehen in der ersten Phase der Validierung im SAL wird erneut die Zuordnung der automatisiert erstellten Sequenzen zu den manuell erstellten Sequenzen vorgenommen und anschließend Overlap, Union und IoU berechnet. Die Ergebnisse dessen sind im Anhang in der Tabelle E.11 (Schlüsselwort-basierte Sequenzierung) und Tabelle E.12 (Sequenzierung mittels Sentence Transformer) aufgeführt. Daraus ergibt sich unter Verwendung von Formel (31) ein Mean IoU von 0,499 für die Schlüsselwort-basierte Sequenzierung und ein Mean IoU von 0,406 für die Sequenzierung mittels Sentence Transformer. Damit zeigt die Sequenzierung mittels Schlüsselworterkennung erneut eine höhere Leistungsfähigkeit, erreicht jedoch auch in diesem Fall nur eine moderate Übereinstimmung mit der manuellen Sequenzierung.

Die Referenzimplementierung des semi-aktiven Wissenstransfers erfüllt insgesamt die angestrebten Funktionalitäten. Während die Transkription der Tonspur äußerst zuverlässig funktioniert, weist die Werkzeugerkennung noch Optimierungspotenzial auf. Dies ist primär auf das Training des YOLOv8-Modells (siehe Kapitel 6.2) zurückzuführen. Der verwendete Datensatz deckt nicht alle während der Validierung eingesetzten Werkzeuge ab, was zu fehlerhaften Werkzeugklassifikationen führt. Zudem tragen

Werkzeuge, die im Hintergrund des Videos sichtbar sind, zu falschen Detektionen bei, wodurch die Erkennung des tatsächlich genutzten Werkzeugs beeinträchtigt wird. Die automatisierte Sequenzierung des Videos ist daher ausschließlich auf Basis der Transkription sinnvoll, wobei die Verwendung der Schlüsselworterkennung die besten Ergebnisse erzielt. Allerdings ist die erreichte Übereinstimmung der automatisiert erstellten Sequenzen des Videos mit den manuell erstellten noch ausbaufähig. Die Sensibilisierung der erfahrenen Mitarbeitenden zur Verwendung von Schlüsselworten zur verbalen Signalisierung von Prozessschrittübergängen sowie die Erweiterung der hinterlegten Liste an Schlüsselwörtern in der Implementierung um individuell verwendete Signalwörter bieten Möglichkeiten, um die Leistungsfähigkeit der automatisierten Sequenzierung zu steigern.

Zusätzlich zu den Erkenntnissen über die Leistungsfähigkeit der einzelnen Komponenten der Referenzimplementierung wurde in der zweiten Phase der Validierung festgestellt, dass der korrekte Umgang mit der HoloLens 2 zur Aufnahme der eigenen Tätigkeit gezielt geübt werden muss. Obwohl die Kamera zentral positioniert ist und den Bereich vor der Person aufzeichnet, erfordert sie eine präzise Ausrichtung durch die Bewegung des Kopfes. Insbesondere bei Mitarbeitenden, die eher über die Augenbewegung ihren Fokuspunkt verändern, besteht die Gefahr, dass relevante Teile der Umgebung und des Prozesses nicht erfasst werden. Diese Problematik könnte durch eine fest in der Umgebung positionierte Kamera behoben werden. Allerdings müsste die Position entsprechend gewählt werden, dass alle relevanten Arbeiten jederzeit ohne Verdeckung sichtbar sind.

Die Ergebnisse der Validierung wurden gemeinsam mit dem beteiligten KMU analysiert. Die Referenzimplementierung wurde dabei insgesamt als eine vielversprechende Alternative zu kommerziell erhältlichen Systemen bewertet. Für den dauerhaften Einsatz im Arbeitsalltag eines KMU sind jedoch insbesondere die Inbetriebnahme, Installation sowie die grafische Benutzerschnittstelle weiterzuentwickeln und zu vereinfachen, um eine reibungslose und benutzerfreundliche Implementierung und Anwendung zu gewährleisten. Der aktuelle Implementierungsstand ist für das KMU aufgrund mangelnder Programmierkenntnisse weder selbstständig installier- noch erweiterbar.

### 7.1.3 Validierung der Referenzimplementierung des passiven Wissenstransfers

*Validation of the Reference Implementation of Passive Knowledge Transfer*

Das Konzept des passiven Wissenstransfers wurde in dieser Arbeit anhand von zwei unterschiedlichen Anwendungsfällen ausdetailliert und umgesetzt (siehe Kapitel 6.3). Beide dabei entwickelten Referenzimplementierungen wurden jeweils anhand eines geeigneten praktischen Anwendungsfalls validiert.

Die Validierung der Referenzimplementierung der Integration von Hand-Tracking und Process Mining zur Erfassung von Best Practices bei manuellen Montagen wurde im SAL des WZLs durchgeführt. Hierbei kamen der rekonfigurierbare Arbeitstisch (siehe Abbildung 6.16) und die Referenzbaugruppe (siehe Abbildung 6.17) zum Einsatz. Die

Ergebnisse wurden bereits teilweise in [DASS24] vorgestellt und werden im Folgenden detaillierter analysiert und diskutiert.

Vor der Durchführung eines umfassenden Systemtests der Referenzimplementierung wurde eine separate Validierung des Hand-Tracking-Moduls und des Grifferkennungs-Moduls (siehe Abbildung 6.18) mittels Konfigurationstests vorgenommen. Dabei wurde untersucht, welche Kombination aus der Verwendung von 2D-Koordinaten versus 3D-Koordinaten und dem regelbasierten Ansatz versus SVM die höchste Klassifikationsgenauigkeit bei der Erkennung der Griffklassen erzielt. Zur Bestimmung der optimalen Konfiguration wurden alle vier möglichen Kombinationen der beiden Bestandteile evaluiert. Anstelle der Erfassung vollständiger Montageprozesse erfolgte dafür die gezielte Aufzeichnung sequenzieller Griffabfolgen, um die Erkennungsrate der korrekten Griffklassen für jede Systemkonfiguration zu quantifizieren. Dazu wurden zehn Videos mit einer Dauer von jeweils einer Minute erstellt, in denen wechselnde Grifffolgen entsprechend der in Abbildung 6.22 vorgestellten Griffklassen demonstriert wurden. Die Ausführung der Griffe geschah in der Luft, ohne ein Objekt wirklich zu greifen. Die Videos wurden anschließend durch das System analysiert, wobei die Ergebnisse mit einer parallel durchgeführten manuellen Videoanalyse abgeglichen wurden, um die Erkennungsraten zu bestimmen. Die Resultate dieser Untersuchung sind in Tabelle 7.7 dargestellt.

**Tabelle 7.7: Genauigkeit der Griffklassifizierung der verschiedenen Konfigurationen des Hand-Tracking-Systems**

*Accuracy of the Grip Classification of the Different Hand Tracking System Configurations*

| | Winkelbereiche | Support Vector Machine |
|---|---|---|
| **2D-Koordinaten** | 42 % | 74 % |
| **3D-Koordinaten** | 54 % | 84 % |

Die Verwendung der Aufnahmen beider Kameras zur Generierung der 3D-Koordinaten der Hand-Landmarks in Kombination mit dem Einsatz der SVM zur Klassifikation der eingesetzten Griffe erzielte das höchste Ergebnis. Aufgrund dessen wurde diese Konfiguration für die nachfolgende Validierungsphase ausgewählt, in der Systemtests zur Überprüfung der Funktionalität des Gesamtsystems erfolgten. Dazu wurden Videoaufnahmen der Montage der Referenzbaugruppe erstellt und durch das System analysiert. Die Aufnahmen wurden von vier Probanden angefertigt, die jeweils einmal die Montage durchführten. Im Rahmen der Tests wurde die Korrektheit der durch das System analysierten Aktivitäten überprüft. Dies umfasste das erzeugte Petri-Netz des Prozesses, die ermittelten Zeitwerte pro Montageschritt sowie für den Gesamtprozess und die Ergebnisse der Belastungsanalyse.

Die automatisiert generierten Petri-Netze der aufgezeichneten Montageprozesse zeigten, dass die Erfassung der Sequenz der ausgeführten Montageschritte auf Basis der

klassifizierten Griffklassen nicht die gewünschte Präzision erreicht. Keines der vier Petri-Netze bildet den jeweiligen Montageprozess vollständig korrekt ab. Stattdessen werden bestimmte Montageschritte häufiger als tatsächlich durchgeführt und andere seltener oder gar nicht korrekt erkannt. Dieses Ergebnis ist auf eine reduzierte Genauigkeit bei der Klassifizierung der Griffklassen zurückzuführen, die durch wiederholte Verdeckungen der Finger während des Montageprozesses durch die gegriffenen Bauteile oder Werkzeuge entsteht. Dadurch ist das MediaPipe-Modell nicht in der Lage, die Position der verdeckten Finger korrekt zu detektieren, sondern trifft Schätzungen, die nicht ausreichend präzise sind. Dieses Verhalten trat bei der Ermittlung der in Tabelle 7.7 dargestellten Genauigkeiten nicht auf, da in diesem Szenario keine Bauteile gegriffen und somit keine problematischen Verdeckungen verursacht wurden.

Trotz dieser Abweichungen bei der erfassten Montagereihenfolge konnte die Gesamtprozesszeit für alle vier Montagen korrekt bestimmt werden. Die Detektion der Ausführung einer Faust mit beiden Händen funktioniert problemlos, da in diesem Fall keine Verdeckung durch Bauteile oder Werkzeuge auftritt. Auch die Erfassung der Zeitwerte pro Montageschritt bei den korrekt erfassten Montageschritten stimmen mit den real benötigten Zeiten überein. Die Qualität der Belastungsanalyse steht in linearer Abhängigkeit von der Genauigkeit der Erfassung der Montagereihenfolge, da die Belastungsanalyse durch die direkte Zuordnung der Montageschritte zu den während der Ausführung auftretenden Kräften im Finger-Hand-Bereich (siehe Tabelle 6.4) erfolgt. Dementsprechend liefert die Validierung der Referenzimplementierung auch diesem Zusammenhang keine zufriedenstellenden Resultate.

Die Ergebnisse dieser Validierung zeigen, dass die entwickelte Referenzimplementierung die erwartete Leistungsfähigkeit nicht erreicht. Insbesondere die Erfassung der Montagereihenfolge auf Grundlage der Klassifizierung der ausgeführten Griffarten weist unter realen Bedingungen signifikante Defizite auf, die die Qualität nachgelagerter Analysen erheblich beeinträchtigen. Für die Entwicklung einer praxistauglichen Systemversion sind daher substanzielle Verbesserungen der Genauigkeit der Griffklassifizierung erforderlich, insbesondere im Hinblick auf die Robustheit gegenüber auftretenden Verdeckungen.

Eine Möglichkeit zur Steigerung der Präzision des Hand-Trackings besteht in einem kombinierten Ansatz aus der Festlegung physikalischer Einschränkungen und der Anwendung von temporaler Glättung. Dabei werden die festen Längenverhältnisse und maximal möglichen Winkel der Handknochen zueinander als Grenzwerte definiert, um unrealistische Positionen der Hand-Landmarks zu erkennen und zu korrigieren. Gleichzeitig werden die Positionen zu vorherigen Zeitpunkten herangezogen, um die Plausibilität aktueller Positionen zu prüfen und damit abrupte Sprünge zu glätten. Durch diese Kombination sollten sich Fehler und Inkonsistenzen reduzieren, wodurch ein robusteres und plausibleres Hand-Tracking erreicht wird.

Die Validierung der Referenzimplementierung des Hand-Trackings zur Überführung manueller Bearbeitungsverfahren wurde in Zusammenarbeit mit demselben KMU

durchgeführt, auf dessen Shopfloor bereits die zweite Phase der Validierung der Referenzimplementierung des semi-aktiven Wissenstransfers stattgefunden hat. Im Fokus der Untersuchung stand dabei der manuelle Schleifprozess zur Erzeugung eines definierten Schliffbildes auf einer Oberfläche, der bereits in der Konzeptentwicklung als Inspiration diente (siehe Kapitel 5.3.2).

Zur Erfassung des Prozesses wurde eine XBOX Kinect V2 über dem Arbeitstisch installiert, an dem die Mitarbeitenden den Schleifprozess ausführen. Abbildung 7.1 zeigt den Bildausschnitt der Kamera, einschließlich des zu bearbeitenden Bauteils.

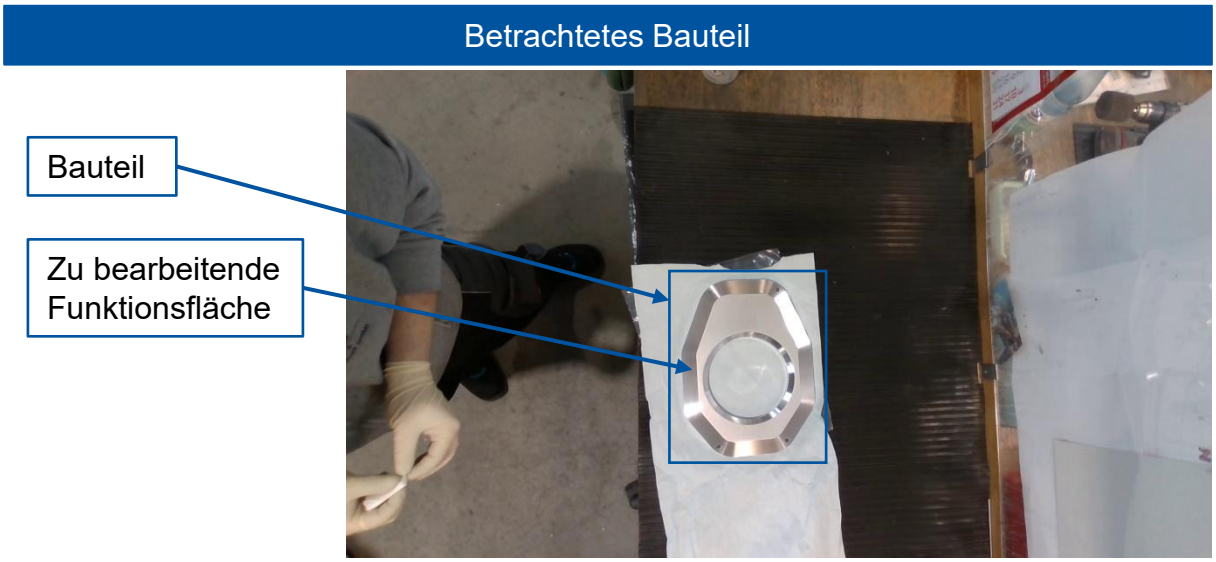



**Abbildung 7.1: Bildausschnitt der für die Validierung im KMU installierten Kamera inklusive des zu bearbeitenden Bauteils**

*Image Section of the Camera Installed in the SME for Validation, Including the Component to be Processed*

Die Mitarbeitenden führen das Schleifpapier während des Prozesses mit dem Daumen über die Funktionsfläche und bewegen es parallel zur Oberfläche. Ziel dieses Vorgehens ist die Entfernung von Querkratzern auf der Funktionsfläche, um die Integrität des in der fertigen Maschine erzeugten Vakuums zu gewährleisten. Die Bearbeitung erfolgt in zwei Schritten. Zunächst wird die dem Körper abgewandte Hälfte der umlaufenden Funktionsfläche geschliffen. Danach rotieren die Mitarbeitenden das Bauteil, um die verbleibende Hälfte analog zu bearbeiten. Jede Hälfte wird mehrfach geschliffen, wobei die Anzahl der Durchgänge von der visuellen Begutachtung der Fläche und dem Vorhandensein sichtbarer Querkratzer abhängig ist. Im Rahmen der Validierung wurden die Bearbeitungsprozesse von insgesamt vier Bauteilen aufgezeichnet, die alle von demselben Mitarbeiter ausgeführt wurden. Die erfassten Bilder wurden anschließend vom System automatisiert analysiert, wobei sowohl der Bewegungsablauf als auch das zugehörige Geschwindigkeitsprofil visualisiert wurden. In Abbildung 7.2 und Abbildung 7.3 sind beide Visualisierungen exemplarisch für eine Iteration der Bearbeitung einer Seite des Bauteils dargestellt.

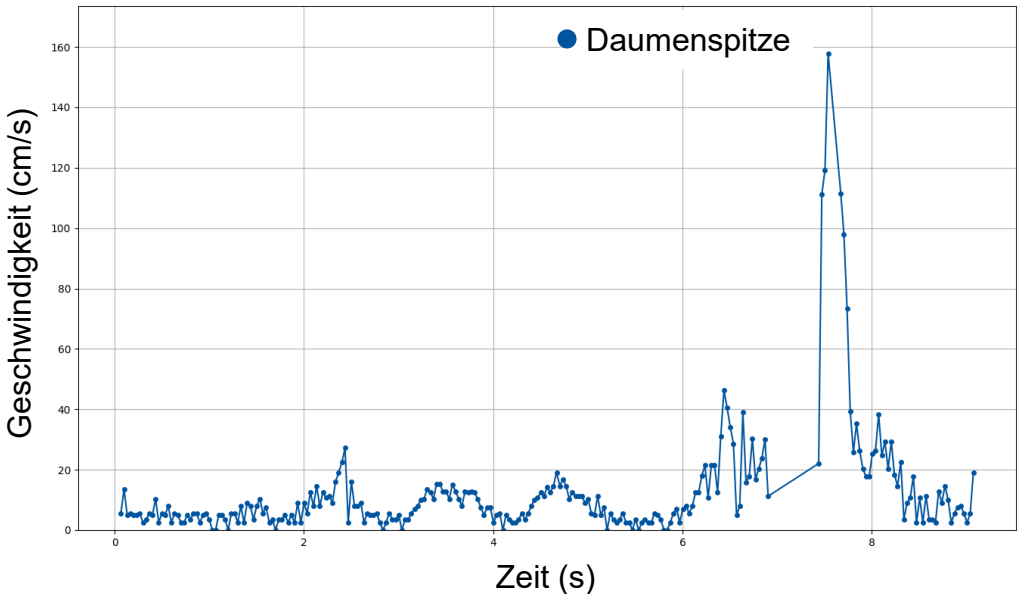



**Abbildung 7.2: Exemplarische Visualisierung des Geschwindigkeitsprofils des Bewegungsablaufs zur Bearbeitung der Funktionsfläche**

*Exemplary Visualization of the Speed Profile of the Motion Sequence for Processing the Functional Surface*

Zur Berechnung des Maßstabs diente das in Abbildung 7.1 gezeigte Bauteil als Referenz. Basierend auf der abgebildeten Perspektive beträgt die Höhe des Bauteils 395 Pixel, was der realen Höhe von 33,10 cm entspricht. Die Höhe in Pixeln konnte mithilfe einer automatisierten Kantendetektion bestimmt werden. Dazu kam ein Python-Script zum Einsatz, in welchem mithilfe des Canny-Algorithmus die Objektgrenzen im Graustufenbild ermittelt, darauf aufbauend Konturen und Bounding Box des Objekts extrahiert und schließlich die Abmessungen des Objekts in Pixeln bestimmt werden.

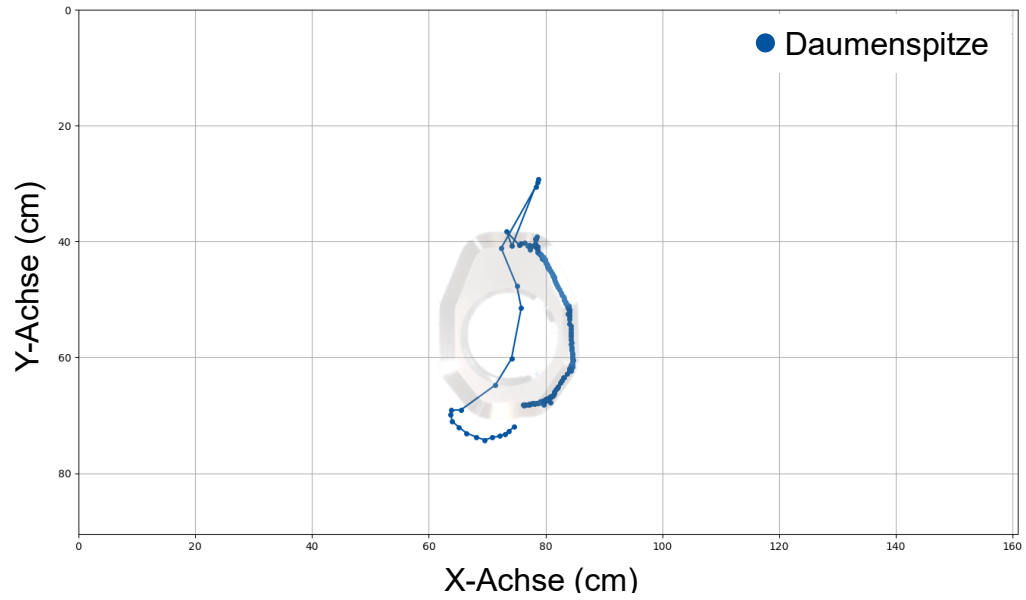



**Abbildung 7.3: Exemplarische Visualisierung des Bewegungsablaufs zur Bearbeitung der Funktionsfläche**

*Exemplary Visualization of the Movement Sequence for Processing the Functional Area*

Wie in Abbildung 7.3 erkennbar, wird die Bewegung der Daumenspitze entlang der Funktionsfläche des Bauteils überwiegend zuverlässig erfasst. Allerdings treten auch bei dieser Anwendung des Hand-Trackings immer wieder signifikante Abweichungen in der Detektion der Position der Hand-Landmarks auf, welche sich im gezeigten Fall in Form von Ausreißern oberhalb des Bauteils manifestieren. Diese Problematik ist erneut auf Verdeckungen zurückzuführen, die durch die Bewegungen des Mitarbeitenden relativ zur Kamera entstehen. Während des Prozesses beugt sich der Mitarbeitende wiederholt über das Bauteil, sodass die Sichtlinie der Kamera auf seine Hand blockiert wird. Die Punkte unterhalb des Bauteils entstanden im Zuge des Neuansetzens des Daumens für den nächsten Durchlauf des Schleifprozesses. Der im Geschwindigkeitsprofil in Abbildung 7.2 sichtbare Sprung korreliert mit den Zeitpunkten, an denen das MediaPipe-Modell eine fehlerhafte Detektion der Position der Daumenspitze vornimmt. Ähnlich wie bei der manuellen Montage könnte auch in diesem Fall der Einsatz von physikalischen Einschränkungen in Kombination mit der Anwendung von temporaler Glättung zu einer Steigerung der Präzision des Hand-Trackings führen. Zusätzlich wäre die Verwendung einer Objekterkennung denkbar, die die Kontur des zu bearbeitenden Bauteils bestimmt. Positionen des Daumens, die außerhalb dieser Kontur liegen, könnten so identifiziert und ab einer definierten Distanz als Ausreißer klassifiziert und für die Analyse entfernt werden.

Die Ergebnisse der Validierungen beider Referenzimplementierungen des passiven Wissenstransfers zeigen, dass die entwickelten Systeme grundsätzlich die vorgesehenen Funktionalitäten bieten. Unter realen Bedingungen erweist sich die Handerkennung mittels MediaPipe jedoch aufgrund von Verdeckungen als unzureichend präzise. Diese Einschränkung führt dazu, dass die Systeme aktuell nicht die für den praktischen Einsatz erforderliche Genauigkeit erreichen und daher entsprechende Verbesserungen erforderlich sind.

## 7.2 Evaluation im Hinblick auf die Anforderungen von KMU

*Evaluation Regarding the Requirements of SMEs*

Basierend auf den Ergebnissen der Validierung der Referenzimplementierungen wird im Folgenden eine Evaluation der Systeme hinsichtlich der in Kapitel 4 definierten Anforderungen durchgeführt. Dazu wird für jede Anforderung analysiert, in welchem Umfang sie durch das jeweilige Konzept erfüllt wird. Diese Bewertung erfolgt anhand von drei Erfüllungsgraden, deren Codierung in Tabelle 7.8 definiert ist. Die Ergebnisse der Evaluation sind in Tabelle 7.9 abgebildet.

**Tabelle 7.8: Qualitative Erfüllungsgrade der Anforderungen**

*Qualitative Degrees of Fulfillment of the Requirements*

| | |
|---|---|
| **Nicht berücksichtigt:** Außerhalb des Untersuchungsrahmens | **NB** |
| **Nicht erfüllt**: Kein Teil der Anforderung wurde erfüllt. | **NE** |
| **Teilweise erfüllt**: Die Anforderung wurde etwa zur Hälfte erfüllt. | **TE** |
| **Vollständig erfüllt**: Die Anforderung wurde in vollem Umfang erfüllt. | **VE** |

**Tabelle 7.9: Bewertung der Erfüllung der Anforderungen für die Konzepte des aktiven (A), semi-aktiven (SA) und passiven (P) Wissenstransfers**

*Evaluation of the fulfillment of the requirements for the concepts of active (A), semi-active (SA) and passive (P) knowledge transfer*

| Kategorie | Anforderungen | A | SA | P |
|---|---|---|---|---|
| **Allgemeine Anforderungen** | Wiederverwendbarkeit von Wissen | VE | VE | VE |
| | Transfer von individuellem Wissen in Unternehmenswissen | VE | VE | TE |
| | Möglichkeit zur Validierung des gespeicherten Wissens | NB | NB | NB |
| **User Experience** | Individualisierbarkeit der Benutzerschnittstellen | NB | NB | NB |
| | Intuitive Datenvisualisierung | NB | NB | NB |
| | Intuitive Implementierung und Nutzung | TE | TE | TE |
| | Intuitive Benutzerschnittstellen | TE | TE | VE |
| | Offenes, erweiterbares, transparentes System | WE | WE | WE |
| **Wissens-dokumentation** | Intuitive, prozessparallele, digitale Dokumentation | TE | VE | VE |
| | Reduktion des Zeitaufwands bei der Dokumentation | TE | VE | VE |
| | Systematische Dokumentation | TE | TE | TE |
| | Strukturierte Dokumentation | TE | TE | TE |
| | Effizientes Wissensmanagement mit geringem Aufwand | TE | VE | VE |
| | Automatisierte Dokumentation von Best Practices | TE | TE | VE |
| | Effiziente Erstellung von Schritt-für-Schritt-Anleitungen | TE | VE | NE |
| **Wissens-speicher** | Zentraler Datenspeicher für alle relevanten Informations-formate | VE | TE | TE |
| | Zentrale Wissensquelle, zentrale Speicherung | VE | TE | TE |
| **Wissens-abruf** | Kontextabhängige, schnelle Suche von Wissen | VE | NE | NE |
| | Orts- und zeitunabhängige Nutzung der Wissensquelle | VE | VE | NE |
| | Uneingeschränkter Zugriff | VE | VE | NE |
| | Schnelle Suche und Identifikation von relevantem Wissen | VE | NE | NE |

Die Anforderungen an die *Wiederverwendbarkeit* sowie den *Transfer von individuellem Wissen in Unternehmenswissen* werden von allen drei Konzepten überwiegend erfüllt. Einschränkungen bestehen jedoch beim passiven Wissenstransfer, da zwar die Handlungen der Mitarbeitenden erfasst, jedoch keine Beweggründe dokumentiert werden können. Die *Validierung des gespeicherten Wissens* wird in keinem der Konzepte unterstützt, da dieser Aspekt nicht primär Gegenstand der vorliegenden Arbeit war (vergleiche Kapitel 4.4).

Auch die Anforderungen an die *Individualisierbarkeit der Benutzerschnittstellen* und die *Visualisierung von Daten* bleiben unberücksichtigt, da sie außerhalb des Fokus dieser Arbeit lagen (vergleiche Kapitel 4.4). Aufgrund des prototypischen Entwicklungsstands sind die *Implementierungen* derzeit nicht benutzerfreundlich, was sich negativ auf die *Nutzung* auswirkt. Dennoch sind die entwickelten *Benutzerschnittstellen* grundsätzlich intuitiv gestaltet. Positiv hervorzuheben ist, dass der passive Wissenstransfer keinerlei Einschränkungen für die Mitarbeitenden verursacht. Alle Systeme zeichnen sich durch *Offenheit und Erweiterbarkeit* aus, weisen jedoch eine eingeschränkte *Transparenz* auf, da keine Informationen über interne Abläufe der Systeme bereitgestellt werden.

Alle drei Konzepte bieten eine *intuitive, prozessbegleitende digitale Dokumentation*, die den *Zeitaufwand für die Dokumentation* reduziert. Allerdings erfolgt der aktive Wissenstransfer prozessintegriert, nicht prozessparallel. In Bezug auf die *Struktur der Wissensdokumentation* bestehen Schwächen: Während beim aktiven Wissenstransfer die Speicherung in Vector Stores für Nutzende unstrukturiert erscheint, erfolgt sie im semi-aktiven und passiven Ansatz über lokale Verzeichnisse. Trotz dieser Einschränkungen bieten alle Konzepte eine höhere *Effizienz* gegenüber dem Status Quo in KMU. Der *Automatisierungsgrad* steigt vom aktiven über den semi-aktiven bis zum passiven Wissenstransfer an. Das semi-aktive Konzept adressiert die *Erstellung von Schritt-für-Schritt-Anleitungen*, während dies beim aktiven Wissenstransfer nur indirekt durch Dialoge mit dem Chatbot möglich ist. Der passive Wissenstransfer ist hierfür ungeeignet.

Hinsichtlich der *Wissensspeicherung* erfüllt das aktive Konzept die Anforderungen vollständig, da es den Zugriff auf externe Datenbanken und Wissensspeicher erlaubt. Die semi-aktiven und passiven Konzepte bieten derzeit keine solche Anbindung. Auch der *Wissensabruf* ist im aktiven Wissenstransfer umfassend umgesetzt. Das semi-aktive Konzept erlaubt *orts- und zeitunabhängigen Zugriff* auf erfasstes Wissen, jedoch ohne integrierte *Suchfunktion*. Der passive Wissenstransfer beschränkt sich auf Dokumentation, Analyse und Speicherung und ermöglicht keinen gezielten Wissensabruf.

## 7.3 Vergleich und Einordnung der Konzepte

*Comparison and Classification of the Concepts*

Zur vergleichenden Analyse der Konzepte des *aktiven*, *semi-aktiven* und *passiven Wissenstransfers* werden diese in der in Abbildung 7.2 dargestellten Matrix eingeord-

net. Die X-Achse repräsentiert den Grad der Automatisierung. Ein niedriger Automatisierungsgrad weist auf Konzepte hin, die stark auf die aktive Interaktion der Nutzenden angewiesen sind, während ein hoher Automatisierungsgrad Konzepte beschreibt, die ohne direkte Interaktionen der Nutzenden auskommen. Die Y-Achse unterscheidet zwischen explizitem und implizitem Wissen. Explizites Wissen wird durch direkte Nutzereingaben erfasst, während implizites Wissen hingegen durch die Beobachtung und Analyse von Handlungen der Nutzenden abgeleitet wird.

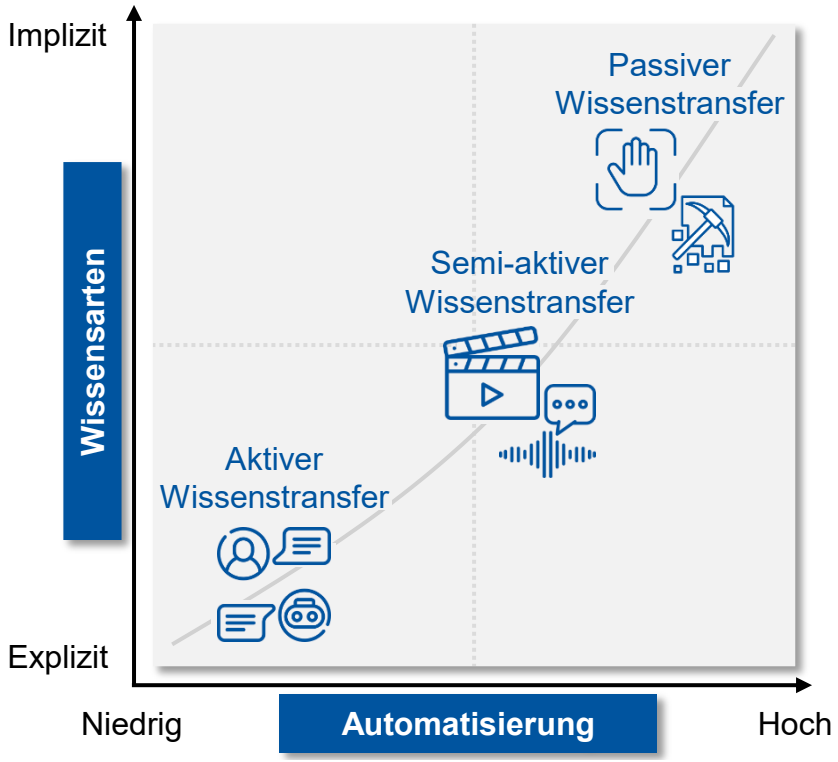



**Abbildung 7.2: Vierfeldermatrix zur Einordnung der drei entwickelten Konzepte**

*Four-field Matrix for Classifying the Three Concepts Developed*

Die Wissenserfassung und -abfrage des aktiven Wissenstransfers über einen Chatbot ist stark interaktiv und erfordert von den Mitarbeitenden eine präzise Formulierung ihrer Erfahrungen und Anfragen. Aufgrund dieser aktiven Beteiligung der Nutzenden ist der Automatisierungsgrad gering. Das Konzept ermöglicht zwar die Erfassung und den Abruf von implizitem Wissen, setzt jedoch voraus, dass die Mitarbeitenden in der Lage sind, ihr Wissen in einer klaren und für andere verständlichen Weise auszudrücken.

Das Konzept des semi-aktiven Wissenstransfers weist einen deutlich höheren Automatisierungsgrad auf. Die Nutzenden müssen lediglich ihre Tätigkeiten kommentieren, wobei der zusätzliche Aufwand gering ist. Die Prozessschritte bleiben unverändert und die visuelle sowie auditive Erfassung erlaubt eine systematische Dokumentation sowohl ausgeführter Handlungen als auch subjektiv relevanter Informationen. Ein entscheidender Vorteil liegt in der Erfassung von Aktivitäten, die den Ausführenden möglicherweise nicht bewusst bedeutsam erscheinen, für neue Mitarbeitende jedoch relevant sein können. Dadurch wird mehr implizites Wissen als beim aktiven Wissenstransfer erfasst.

Der passive Wissenstransfer wird durch videobasiertes Hand-Tracking vollständig automatisiert umgesetzt. Da keine aktive Mitwirkung erforderlich ist, erreicht dieses Konzept den höchsten Automatisierungsgrad. Die uneingeschränkte Ausführung der Tätigkeiten gewährleistet eine authentische, ungefilterte Erfassung individueller Vorgehensweisen. Aufgrund dessen hat dieses Konzept hohes Potenzial zur Erfassung von implizitem Wissen. Allerdings bleiben die Beweggründe für bestimmte Handlungen unklar und müssen durch ergänzende Methoden nachträglich erschlossen werden.

### 7.3.1 Beantwortung der Teilforschungsfragen

*Answering the Sub-Research Questions*

Insgesamt können die in Kapitel 3 aufgestellten Teilforschungsfragen auf Basis der gewonnenen Erkenntnisse wie folgt beantwortet werden:

**Wie können Mitarbeitende ihr Wissen an bestimmten Punkten ihrer Tätigkeit intuitiv und mit möglichst minimalem Mehraufwand *aktiv* digital speichern bzw. weitergeben?**

Für Prozesse mit diskreten Schritten eignet sich der Einsatz eines Chatbot-basierten Systems, das Mitarbeitenden die Möglichkeit bietet, ihr Wissen zwischen den Prozessschritten intuitiv und effizient *aktiv* zu dokumentieren. Alternativ ermöglicht der Einsatz eines HMDs in Kombination mit einer automatisierten Analyse von Video- und Tonspuren eine prozessparallele, *semi-aktive* Weitergabe von Wissen. Dies bietet den Vorteil, dass die Mitarbeitenden Erklärungen über ihre Tätigkeiten hinzufügen können, wodurch Gründe für bestimmte Vorgehensweisen erfasst werden.

**Wie kann Wissen in Form von „Best Practices“ *passiv* erfasst bzw. „getrackt“ werden, ohne die Mitarbeitenden bei der Ausführung ihrer Tätigkeiten zu behindern?**

Bei stationären Arbeitsplätzen kann ein videobasiertes Hand-Tracking eingesetzt werden, um die Aktivitäten der Mitarbeitenden zu erfassen. Auf diese Weise können individuelle Routinen und Best Practices *passiv* dokumentiert werden, ohne die Ausführung der Tätigkeiten zu beeinträchtigen. Für Tätigkeiten mit wechselnden Arbeitsplätzen kann ein HMD in Kombination mit einer automatisierten Analyse von Video- und Tonspuren verwendet werden. Falls die Erfassung vollständig ohne Aktionen der Mitarbeitenden erfolgen soll, kann ausschließlich das aufgezeichnete Video zur Analyse herangezogen werden, während auf die zusätzliche Tonspur verzichtet wird. Dieser Ansatz erlaubt ebenfalls eine uneingeschränkte Ausführung der Tätigkeiten, während implizite Best Practices *passiv* oder *semi-aktiv* dokumentiert werden.

**Wie kann eine kontextualisierte, strukturierte Speicherung des Wissens realisiert werden, um eine bedarfsgerechte Bereitstellung zu gewährleisten?**

Die in Verbindung mit dem Chatbot-basierten System verwendeten Vector Stores bieten eine flexible Möglichkeit, Wissen bedarfsgerecht abzurufen, ohne auf vorab definierte Speicherstrukturen angewiesen zu sein. Die Transformation des Wissens in Embeddings ermöglicht eine semantische Suche, die kontextbezogene Inhalte liefert. Al-

ternativ können in Videos detektierte Objekte und in Audiodaten erkannte Schlüsselwörter als Identifikatoren dienen, um Wissen kontextbezogen zuzuordnen. In Kombination mit strukturierten Speichersystemen (z. B. SQL-Datenbanken oder Ontologien) können diese Identifikatoren für die automatische Speicherung und das situationsgerechte Abrufen des Wissens verwendet werden.

# 8 Zusammenfassung und Ausblick

*Summary and Outlook*

## 8.1 Zusammenfassung

*Summary*

Der gezielte Einsatz von Wissensmanagement zur Verstetigung von individuellem, implizitem Wissen stellt einen zentralen Faktor zu Sicherung der Wettbewerbsfähigkeit deutscher Unternehmen dar. Insbesondere KMU stehen vor der Herausforderung, dass Wissen häufig nur unzureichend dokumentiert ist und primär als Erfahrungswissen in den Köpfen der Mitarbeitenden existiert. Ein durch das Ausscheiden langjähriger Mitarbeitender hervorgerufener Wissensverlust kann daher kaum kompensiert werden. Die aktuell kommerziell erhältlichen Systeme zur Unterstützung von Wissensmanagement bieten nur unzureichende Funktionalitäten zur effizienten, intuitiven Nutzung im Umfeld produzierender Unternehmen und fokussieren überwiegend die Darstellung von Wissen, nicht dessen Erfassung. Aus dieser Situation resultiert die Notwendigkeit, innovative methodische und technologische Ansätze für Wissensmanagement zu entwickeln, die gezielt auf die spezifischen Anforderungen von KMU und deren Produktionsumgebungen abgestimmt sind.

Die vorliegende Arbeit verfolgte daher das Ziel, die Integration von Wissensmanagement auf dem Shopfloor von KMU zu untersuchen und neue Ansätze zu entwickeln, welche den Mitarbeitenden ermöglichen, ihr Erfahrungswissen unmittelbar, intuitiv und effizient digital zu speichern, wodurch das Wissen kontextualisiert und unternehmensweit abrufbar gemacht wird.

Das methodische Vorgehen dieser Arbeit basiert auf sechs Schritten:

1. Literaturgestützte Analyse des aktuellen Stands der Technik zur Erarbeitung eines fundierten Verständnisses der aktuellen Situation bezüglich des Einsatzes von Wissensmanagement auf dem Shopfloor von KMU
2. Analyse der spezifischen Anforderungen von KMU an Wissensmanagementsysteme und -methoden für den Einsatz auf dem Shopfloor anhand einer Studie
3. Entwicklung innovativer Wissensmanagementkonzepte zur gezielten Adressierung der identifizierten Anforderungen
4. Implementierung der Konzepte unter Berücksichtigung des aktuellen Digitalisierungsstands sowie der technischen Rahmenbedingungen in KMU
5. Technische Validierung der entwickelten Referenzimplementierungen in Laborumgebungen sowie unter realen Bedingungen
6. Evaluation der Referenzimplementierungen anhand der identifizierten Anforderungen

Die Anforderungsanalyse zeigt, dass das größte Problem der Unternehmen tatsächlich in der Erfassung und strukturierten Dokumentation von vorhandenem, implizitem Wissen besteht. Die systematische Speicherung und Bereitstellung dieses Wissens

werden als essenziell erachtet, um die langfristige Kompetenz des Unternehmens sicherzustellen.

Im Rahmen der Konzeptentwicklung wird die Grundidee der prozessparallelen Erfassung von Erfahrungswissen eingeführt, die eine Integration von Wissensmanagementprozessen in bestehende betriebliche Abläufe vorsieht. Darauf aufbauend werden drei unterschiedliche Konzepte für innovative Wissensmanagementsysteme für den Shopfloor vorgestellt, die sich hinsichtlich des erforderlichen Aufwands für die Mitarbeitenden und der Art des Wissenstransfers unterscheiden.

Das Konzept des *aktiven* Wissenstransfers beschreibt ein Chatbot-basiertes System, welches den Mitarbeitenden die simultane Erfassung und Abfrage von Wissen über ein Chatfenster ermöglicht. Bei der Eingabe von Wissen müssen die Mitarbeitenden selbstständig und aktiv die relevanten Inhalte auswählen und verständlich formulieren, sodass andere Mitarbeitende ihr hinterlegtes Wissen bei Bedarf über den Chatbot abrufen und verstehen können. Der Chatbot kann während der Arbeitsabläufe, insbesondere zwischen einzelnen Prozessschritten, genutzt werden. Zur technischen Realisierung werden zwei Ansätze untersucht: das Rasa-Framework und der Einsatz von LLM.

Beim *semi-aktiven* Wissenstransfer werden erfahrene Mitarbeitende mit einem Mikrofon und einer am Kopf befestigten Kamera ausgestattet, welche die Arbeitsumgebung und ihre Aktivitäten aus ihrer Perspektive aufzeichnen. Ergänzend sollen die Mitarbeitenden bei der Ausübung ihrer Tätigkeit relevante Aspekte verbal erläutern. Die aufgezeichneten Audio- und Videodaten werden mittels KI-Methoden analysiert und entsprechend den Prozessschritten sequenziert und in Learning-Nuggets überführt, die als Schulungsmaterial für neue Mitarbeitende dienen. Dieser Ansatz verringert den zusätzlichen Arbeitsaufwand für die Wissensdokumentation erheblich, da lediglich die verbale Erläuterung eine geringe kognitive Zusatzbelastung darstellt.

Das Konzept des *passiven* Wissenstransfers sieht vor, dass im Arbeitsraum des betrachteten Arbeitsplatzes zwei Kameras platziert werden, welche die Aktivitäten der Hände der Mitarbeitenden aufzeichnen. Diese Handbewegungen werden anschließend mittels Machine Learning-Methoden ausgewertet, um verschiedene Griffklassen zu differenzieren. Anhand der zu einem spezifischen Zeitpunkt mit beiden Händen von den Mitarbeitenden ausgeführten Griffklassen wird automatisiert bestimmt, welche Tätigkeit zu diesem Zeitpunkt ausgeführt wird. Die resultierenden Prozessdaten werden in Form eines Event-Logs gespeichert und mittels Process Mining analysiert und visualisiert, um Arbeitsabläufe hinsichtlich Effizienz und Ergonomie zu optimieren. Dieser Ansatz erfordert keine zusätzlichen Interaktionen der Mitarbeitenden über ihre regulären Arbeitsschritte hinaus.

Alle drei Konzepte wurden als On-Premise-Lösungen auf KMU-tauglicher Hardware realisiert und technisch validiert. Die Validierung zeigt, dass beim aktiven Wissenstransfer der Einsatz von LLM zur Realisierung des Chatbots die bessere Performance liefert, jedoch stark von der Leistungsfähigkeit der verwendeten Hardware abhängt. Beim semi-aktiven Wissenstransfer liefert die KI-gestützte Analyse der Tonspur überzeugende Ergebnisse, während die Objekterkennung im Videomaterial mittels

YOLOv8 ein intensiveres Training erfordert, um die gewünschte Genauigkeit bei der Werkzeugerkennung zu erreichen. Zudem erweist sich die Nutzung einer am Kopf montierten Kamera als gewöhnungsbedürftig und bedarf Übung, bis der relevante Bereich der Arbeitsumgebung sicher erfasst wird. Die Validierungsergebnisse des passiven Wissenstransfers zeigen, dass das Konzept zwar viel Potenzial hat, die Umsetzung jedoch bei der präzisen Erfassung der Hände noch zu fehleranfällig ist.

Ein anschließender Abgleich der implementierten Konzepte mit den initial identifizierten Anforderungen bestätigt, dass die entwickelten Ansätze die grundlegenden Anforderungen der Unternehmen erfüllen. Für einen Einsatz in der Praxis müssen die benannten, technischen Qualitätsmängel jedoch beseitigt und die Benutzerfreundlichkeit bei der Implementierung gesteigert werden.

## 8.2 Ausblick

*Outlook*

Die vorliegende Arbeit leistet einen wesentlichen Beitrag zur Etablierung eines effektiven Wissensmanagement auf dem Shopfloor von KMU. Die Ergebnisse verdeutlichen jedoch, dass die praxisgerechte, robuste und benutzerfreundliche Implementierung der vorgestellten Wissensmanagementsysteme eine Herausforderung darstellt, die nicht allein durch erhöhte Implementierungsaufwände zu bewältigen ist. Vielmehr erfordert sie eine gezielte Analyse der spezifischen Anforderungen der Mitarbeitenden hinsichtlich des Designs und der Funktionalität der Benutzerschnittstellen, um eine hohe Akzeptanz und nachhaltige Nutzung sicherzustellen.

Die drei separat entwickelten Konzepte besitzen das Potenzial, in ein integriertes Gesamtsystem überführt zu werden, das eine zentrale Wissensbasis verwaltet und zugleich verschiedene Methoden zur Wissensdokumentation und -abfrage bereitstellt. Dies ermöglicht eine an individuelle Präferenzen und arbeitsplatzspezifische Rahmenbedingungen angepasste Nutzung. In diesem Kontext sollte auch die Verknüpfung mit bestehenden Systemen wie PLM, MES und ERP sowie mit Wikis, Fileservern und Office-Anwendungen untersucht werden. Eine verbesserte Interoperabilität mit bestehenden Infrastrukturen erhöht die zugängliche Wissensbasis und steigert somit die Anwendungsrelevanz des Systems.

Während die vorliegende Arbeit vorrangig die Erfassung von Wissen fokussierte und die Bereitstellung bzw. den Abruf nur am Rande behandelte, birgt letzterer Aspekt ein erhebliches Potenzial für weiterführende wissenschaftliche Untersuchungen. Insbesondere Virtual Reality und Augmented Reality bieten vielfältige Möglichkeiten zur Gestaltung innovativer Lehr- und Lernkonzepte, bei denen neue Tätigkeiten in immersiven, risikofreien Umgebungen realitätsnah simuliert und beliebig oft trainiert werden können. Eine systematische Validierung dieser Technologien im Vergleich zu konventionellen Schulungsformaten stellt ein vielversprechendes Forschungsfeld dar und kann als Erweiterung des passiven Wissenstransfers betrachtet werden.

# Summary and Outlook

**Summary**

The targeted use of knowledge management to consolidate individual, implicit knowledge is a key factor in ensuring the competitiveness of German companies. SMEs in particular face the challenge that knowledge is often inadequately documented and exists primarily as experiential knowledge in the minds of employees. The loss of knowledge caused by the departure of long-standing employees can therefore hardly be compensated for. The systems currently available on the market to support knowledge management offer insufficient functionality for efficient, intuitive use in the manufacturing environment and focus primarily on the presentation of knowledge rather than its capture. This situation has resulted in the need to develop innovative methodological and technological approaches to knowledge management that are specifically tailored to the specific requirements of SMEs and their production environments.

The aim of this thesis was therefore to investigate the integration of knowledge management on the shop floor of SMEs and to develop new approaches that enable employees to store their experiential knowledge digitally in an immediate, intuitive and efficient manner, thereby contextualizing the knowledge and making it accessible throughout the company.

The methodological approach of this work is based on six steps:

1. Literature-based analysis of the current state of the art to develop a sound understanding of the current situation regarding the use of knowledge management on the shop floor of SMEs
2. Analysis of the specific requirements of SMEs for knowledge management systems and methods for use on the shop floor based on a study
3. Development of innovative knowledge management concepts to specifically address the identified requirements
4. Implementation of the concepts, taking into account the current state of digitization and the technical framework conditions in SMEs
5. Technical validation of the developed reference implementations in laboratory environments and under real conditions
6. Evaluation of the reference implementations based on the identified requirements

The requirements analysis shows that the biggest problem for companies is actually the collection and structured documentation of existing, implicit knowledge. The systematic storage and provision of this knowledge is considered essential to ensure the long-term competence of the company.

As part of the concept development, the basic idea of process-parallel recording of experiential knowledge is introduced, which envisages the integration of knowledge management processes into existing operational procedures. Building on this, three different concepts for innovative knowledge management systems for the shop floor

are presented, which differ in terms of the effort required by employees and the type of knowledge transfer.

The concept of active knowledge transfer describes a chatbot-based system that enables employees to simultaneously capture and query knowledge via a chat window. When entering knowledge, employees must independently and actively select the relevant content and formulate it in a comprehensible manner so that other employees can retrieve and understand their stored knowledge via the chatbot when needed. The chatbot can be used during work processes, especially between individual process steps. Two approaches are being investigated for technical implementation: the Rasa framework and the use of LLM.

In semi-active knowledge transfer, experienced employees are equipped with a microphone and a camera attached to their head, which records the work environment and their activities from their perspective. In addition, employees are asked to verbally explain relevant aspects while performing their tasks. The recorded audio and video data is analysed using AI methods, sequenced according to the process steps and converted into learning nuggets that serve as training material for new employees. This approach significantly reduces the additional workload for knowledge documentation, as the verbal explanation alone represents a low additional cognitive burden.

The concept of passive knowledge transfer involves placing two cameras in the workspace of the workplace in question to record the activities of the employees' hands. These hand movements are then evaluated using machine learning methods to differentiate between different grip classes. Based on the grip classes performed by employees with both hands at a specific point in time, the system automatically determines which activity is being performed at that moment. The resulting process data is stored in the form of an event log and analysed and visualized using process mining to optimize workflows in terms of efficiency and ergonomics. This approach does not require any additional interaction from employees beyond their regular work steps.

All three concepts were implemented as on-premise solutions on SME-compatible hardware and technically validated. The validation shows that, in terms of active knowledge transfer, the use of LLM to implement the chatbot delivers better performance but is heavily dependent on the performance of the hardware used. In semi-active knowledge transfer, AI-supported analysis of the audio track delivers convincing results, while object recognition in video material using YOLOv8 requires more intensive training to achieve the desired accuracy in tool recognition. In addition, the use of a head-mounted camera takes some getting used to and requires practice until the relevant area of the working environment is reliably captured. The validation results of passive knowledge transfer show that although the concept has great potential, its implementation is still too prone to errors when it comes to the precise capture of hands.

A subsequent comparison of the implemented concepts with the initially identified requirements confirms that the developed approaches meet the basic requirements of the companies. However, for practical use, the identified technical quality deficiencies must be eliminated and user-friendliness during implementation must be increased.

**Outlook**

This work makes a significant contribution to establishing effective knowledge management on the shop floor of SMEs. However, the results make it clear that the practical, robust, and user-friendly implementation of the knowledge management systems presented is a challenge that cannot be overcome by increased implementation efforts alone. Rather, it requires a targeted analysis of the specific requirements of employees with regard to the design and functionality of the user interfaces in order to ensure a high level of acceptance and sustainable use.

The three separately developed concepts have the potential to be transferred into an integrated overall system that manages a central knowledge base and at the same time provides various methods for knowledge documentation and retrieval. This enables use tailored to individual preferences and workplace-specific conditions. In this context, the connection to existing systems such as PLM, MES and ERP, as well as to wikis, file servers and office applications, should also be investigated. Improved interoperability with existing infrastructures increases the accessible knowledge base and thus enhances the application relevance of the system.

While this work focused primarily on the acquisition of knowledge and only dealt with its provision and retrieval in passing, the latter aspect holds considerable potential for further scientific investigation. Virtual reality and augmented reality in particular offer a wide range of possibilities for designing innovative teaching and learning concepts in which new activities can be simulated realistically in immersive, risk-free environments and practiced as often as desired. Systematic validation of these technologies in comparison to conventional training formats represents a promising field of research and can be considered an extension of passive knowledge transfer.

# V. Literaturverzeichnis

*References*

[AARS20] Aarsen, T.: Sentence-Transformers: Embeddings, Retrieval, and Reranking. URL: https://github.com/UKPLab/sentence-transformers. [Stand: 11.02.2025]

[ABUB19] Abubakar, A. M.; Elrehail, H.; Alatailat, M. A.; Elçi, A.: Knowledge management, decision-making style and organizational performance. In: Journal of Innovation & Knowledge. 4. Jg., 2019, Nr. 2, S. 104–114

[ACKE18] Ackermann, B.; Krancher, O.; North, K.; Schildknecht, K.; Schorta, S.: Erfolgreicher Wissenstransfer in agilen Organisationen. Wiesbaden: Springer Gabler, 2018

[ADAM15] Adams, W. C.: Conducting semi-structured interviews. In: Handbook of practical program evaluation, 2015, S. 492–505

[AHMA24] Ahmad, H. A.; Rashid, T. A.: Planning the development of text-to-speech synthesis models and datasets with dynamic deep learning. In: Journal of King Saud University - Computer and Information Sciences. 36. Jg., 2024, Nr. 7, S. 102131

[AITC00] Aitchison, J.; Gilchrist, A.; Bawden, D.: Thesaurus Construction and Use. 4. Aufl. London: Aslib IMI, The Association for Information Management, 2000

[ALAV01] Alavi, M.; Leidner, D. E.: Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. In: MIS Quarterly. 25. Jg., 2001, Nr. 1, S. 107–136

[ALT24] Alt, B.; Stöckl, F.; Müller, S.; Braun, C.; Raible, J.; Alhasan, S.; Rettig, O.; Ringle, L.; Katic, D.; Jäkel, R.; Beetz, M.; Strand, M.; Huber, M. F.: RoboGrind: Intuitive and Interactive Surface Treatment with Industrial Robots. In: 2024 IEEE International Conference on Robotics and Automation (ICRA), 2024, S. 2140–2146

[ANDR19] Andrianakos, G.; Dimitropoulos, N.; Michalos, G.; Makris, S.: An approach for monitoring the execution of human based assembly operations using machine learning. In: Procedia CIRP. 86. Jg., 2019, S. 198–203

[ANTI25] Anticipate GmbH: Anticipate. URL: https://www.anticipate.ml/. [Stand: 28.01.2025]

[ANTO09] Antoniou, G.; van Harmelen, F.: Web Ontology Language: OWL. In: Staab, S.; Studer, R. (Hrsg.): Handbook on Ontologies. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, S. 91–110

[AREN22] Arena, F.; Collotta, M.; Pau, G.; Termine, F.: An Overview of Augmented Reality. In: Computers. 11. Jg., 2022, Nr. 2

[ARGO00] Argote, L.; Ingram, P.: Knowledge Transfer: A Basis for Competitive Advantage in Firms. In: Organizational Behavior and Human Decision Processes. 82. Jg., 2000, Nr. 1, S. 150–169

[ARIS20] Aris, R.; Sensuse, D.: Knowledge Management Systems Development and Implementation: A systematic Literature Review. In: International Conference on Creative Economics, Tourism & Information Management (ICCETIM) 2020, 2020

[ATLA25] Atlassian: Confluence. URL: https://www.atlassian.com/de/software/confluence. [Stand: 28.01.2028]

[BAIL06] Bailey, C.; Zalfan, M. T.; Davis, H. C.; Fill, K.; Conole, G.: Panning for Gold: Design Pedagogically-inspired Learning Nuggets. In: Educational Technology & Society. 9. Jg., 2006, Nr. 1, S. 133 - 122

[BAND71] Bandura, A.: Social Learning Theory. New York: General Learning Press, 1971

[BARR24] Barron, R. C.; Grantcharov, V.; Wanna, S.; Eren, M. E.; Bhattarai, M.; Solovyev, N.; Tompkins, G.; Nicholas, C.; Rasmussen, K. Ø.; Matuszek, C.; Alexandrov, B. S.: Domain-Specific Retrieval-Augmented Generation Using Vector Stores, Knowledge Graphs, and Tensor Factorization. In: 2024 International Conference on Machine Learning and Applications (ICMLA), 2024, S. 1669–1676

[BERT18] Bertram, P.; Birtel, M.; Quint, F.; Ruskowski, M.: Intelligent Manual Working Station through Assistive Systems. In: IFAC-PapersOnLine. 51. Jg., 2018, Nr. 11, S. 170–175

[BIRT18] Birtel, F.; Wötzel, A.: Branchenindikator Instandhaltung,FIR an der RWTH Aachen, 2018

[BITK17] Bitkom e.V.: ECM im Mittelstand. Berlin, 2017

[BITK19] Bitkom e.V.: Digital Office im Mittelstand 2019. Berlin, 2019

[BLEY16] Bley, K.; Leyh, C.: Status Quo der Digitalisierung deutscher Industrieunternehmen. In: Multikonferenz Wirtschaftsinformatik 2016, 2016, S. 1651–1662

[BLOO25] Bloomfire: Bloomfire. URL: https://bloomfire.com/. [Stand: 28.01.2025]

[BOCK17] Bocklisch, T.; Faulkner, J.; Pawlowski, N.; Nichol, A.: Rasa: Open Source Language Understanding and Dialogue Management. In: arXiv, 2017

[BOMM22] Bommasani, R.; Hudson, D. A.; Adeli, E.; Altman, R.; Arora, S.; Arx, S. v.; Bernstein, M. S.; Bohg, J.; Bosselut, A.; Brunskill, E.; Brynjolfsson, E.; Buch, S.; Card, D.; Castellon, R.; Chatterji, N.; Chen, A.; Creel, K.; Davis,

J. Q.; Demszky, D.; Donahue, C.; Doumbouya, M.; Durmus, E.; Ermon, S.; Etchemendy, J.; Ethayarajh, K.; Fei-Fei, L.; Finn, C.; Gale, T.; Gillespie, L.; Goel, K.; Goodman, N.; Grossman, S.; Guha, N.; Hashimoto, T.; Henderson, P.; Hewitt, J.; Ho, D. E.; Hong, J.; Hsu, K.; Huang, J.; Icard, T.; Jain, S.; Jurafsky, D.; Kalluri, P.; Karamcheti, S.; Keeling, G.; Khani, F.; Khattab, O.; Koh, P. W.; Krass, M.; Krishna, R.; Kuditipudi, R.; Kumar, A.; Ladhak, F.; Lee, M.; Lee, T.; Leskovec, J.; Levent, I.; Li, X. L.; Li, X.; Ma, T.; Malik, A.; Manning, C. D.; Mirchandani, S.; Mitchell, E.; Munyikwa, Z.; Nair, S.; Narayan, A.; Narayanan, D.; Newman, B.; Nie, A.; Niebles, J. C.; Nilforoshan, H.; Nyarko, J.; Ogut, G.; Orr, L.; Papadimitriou, I.; Park, J. S.; Piech, C.; Portelance, E.; Potts, C.; Raghunathan, A.; Reich, R.; Ren, H.; Rong, F.; Roohani, Y.; Ruiz, C.; Ryan, J.; Ré, C.; Sadigh, D.; Sagawa, S.; Santhanam, K.; Shih, A.; Srinivasan, K.; Tamkin, A.; Taori, R.; Thomas, A. W.; Tramèr, F.; Wang, R. E.; Wang, W.; Wu, B.; Wu, J.; Wu, Y.; Xie, S. M.; Yasunaga, M.; You, J.; Zaharia, M.; Zhang, M.; Zhang, T.; Zhang, X.; Zhang, Y.; Zheng, L.; Zhou, K.; Liang, P.: On the Opportunities and Risks of Foundation Models. In: arXiv, 2022

[BOVE20] Bovenschulte, M.: Kognitive Assistenzsysteme,Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag, 2020

[BRAD08] Bradski, G. R.; Kaehler, A.: Learning OpenCV. Sebastopol (Kalifornien, USA): O'Reilly Media, 2008

[BREC11] Brecher, C.: Integrative Produktionstechnik für Hochlohnländer. Berlin, Heidelberg: Springer, 2011

[BREC21] Brecher, C.; Biernat, B.; Fey, M.; Kehne, S.; Lohrmann, V.; Spierling, R.; Wiesch, M.: Data Science in Production. In: AWK21 - Internet of Production – Turning Data into Sustainability. Deutsche Ausgabe, 2021, S. 225–261

[BREC23] Brecher, C.; Petrovic, O.; Dassen, Y.; Blanke, P.; Trinh, M.; Wurm, S.: Nachhaltige Produktion-as-a-Service. In: AWK23 - Empower Green Production, 2023, S. 247–276

[BUCH21] Buchgeher, G.; Gabauer, D.; Martinez-Gil, J.; Ehrlinger, L.: Knowledge Graphs in Manufacturing and Production: A Systematic Literature Review. In: IEEE Access. 9. Jg., 2021, S. 55537–55554

[BURG24] Burggraef, P.; Adlon, T.; Steinberg, F.; Broehl, F.; Moriz, A.; Engeln, C.; Schütz, M.; Jaworek, F.: Automatic Generation of Assembly Instructions by Analyzing Process Recordings – A Concept Overview. In: Procedia CIRP. 126. Jg., 2024, S. 775–780

[CHEN10] Chen, C.; Fu, R.: A new frame of tacit knowledge transfer based on ontology. In: Seventh International Conference on Fuzzy Systems and Knowledge Discovery, 2010, S. 1756–1760

[CHEN24] Chen, B.; Zhang, Z.; Langrené, N.; Zhu, S.: Unleashing the potential of prompt engineering in Large Language Models: a comprehensive review. In: Patterns. 6. Jg., 2024, Nr. 6

[CHOW23] Chowdhery, A.; Narang, S.; Devlin, J.; Bosma, M.; Mishra, G.; Roberts, A.; Barham, P.; Chung, H. W.; Sutton, C.; Gehrmann, S.; Schuh, P.; Shi, K.; Tsvyashchenko, S.; Maynez, J.; Rao, A.; Barnes, P.; Tay, Y.; Shazeer, N.; Prabhakaran, V.; Reif, E.; Du Nan; Hutchinson, B.; Pope, R.; Bradbury, J.; Austin, J.; Isard, M.; Gur-Ari, G.; Yin, P.; Duke, T.; Levskaya, A.; Ghemawat, S.; Dev, S.; Michalewski, H.; Garcia, X.; Misra, V.; Robinson, K.; Fedus, L.; Zhou, D.; Ippolito, D.; Luan, D.; Lim, H.; Zoph, B.; Spiridonov, A.; Sepassi, R.; Dohan, D.; Agrawal, S.; Omernick, M.; Dai, A. M.; Pillai, T. S.; Pellat, M.; Lewkowycz, A.; Moreira, E.; Child, R.; Polozov, O.; Lee, K.; Zhou, Z.; Wang, X.; Saeta, B.; Diaz, M.; Firat, O.; Catasta, M.; Wei, J.; Meier-Hellstern, K.; Eck, D.; Dean, J.; Petrov, S.; Fiedel, N.: PaLM: Scaling Language Modeling with Pathways. In: J. Mach. Learn. Res. 24. Jg., 2023, Nr. 1

[CLEM23] Clemente, M.; Domingues, L.: Analysis of Project Management Tools to support Knowledge Management. In: Procedia Computer Science, 2023, Nr. 219, S. 1769–1776

[COQU23] Coqui-AI: Coqui TTS. URL: https://github.com/coqui-ai/TTS?tab=readme-ov-file. [Stand: 16.02.2025]

[DASS21] Dassen, Y.; Herfs, W.: Systematische Erfassung und Applikation der Erfahrungen von Mitarbeitern in der Produktion. In: Fortschritt-Berichte VDI, Reihe 02 Fertigungstechnik, Konvergenz von OT und IT, 2021, S. 57–68

[DASS23] Dassen, Y.; Liehner, L.; Brauner, P.; Storms, S.; Herfs, W.; Ziefle, M.: Current State and Future Direction of Digital Knowledge Management in Small and Medium Enterprises. In: Proceedings of the 24th European Conference on Knowledge Management ECKM 2023, 2023, Nr. 24, S. 284–292

[DASS24] Dassen, Y.; Zhao, Y.; Petrovic, O.; Herfs, W.: Integration of Hand Tracking and Process Mining to improve Knowledge Transfer on the Shopfloor. In: 2024 10th International Conference on Control, Automation and Robotics (ICCAR), 2024, S. 16–20

[DAVE98] Davenport, T. H.; Prusak, L.: Working Knowledge: How Organizations Manage What They Know. In: Ubiquity - An ACM IT Magazine and Forum, 1998

[DEUT15] (November 2015). DIN EN ISO 9000, Qualitätsmanagementsysteme - Grundlagen und Begriffe (ISO 9000:2015); Deutsche und Englische Fassung EN ISO 9000:2015

[DEUT16] (Dezember 2016). DIN EN ISO 6385, Grundsätze der Ergonomie für die Gestaltung von Arbeitssystemen (ISO 6385:2016); Deutsche Fassung EN ISO 6385:2016

[DEUT21] (August 2021). DIN SPEC 91443, Systematisches Wissensmanagement für KMU – Instrumente und Verfahren

[DEVL19] Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: Proceedings of the 2019 Conference of the North {A}merican Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 2019

[DICK02] Dick, M.; Wehner, T.: Wissensmanagement zur Einführung: Bedeutung, Definition, Konzepte. In: Wissensmanagement-Praxis. Einführung, Handlungsfelder und Fallbeispiele, 2002, S. 7–27

[DIGI11] Digitale Bibliothek der GEsellschaft für Informatik e.V.: Schriftreihen zu Enterprise 2.0-Fallstudien. URL: https://dl.gi.de/collections/a68efd58-6438-44a9-b41d-f45431c82356. [Stand: 28.01.2025]

[DINE21] Dinesh, T.; R, A. M.; Newton, T. T.; R, S. G.: AI Bot for Academic Schedules using Rasa. In: 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 2021, S. 1–6

[DING22] Ding, I.-J.; Zheng, N.-W.: RGB-D Depth-sensor-based Hand Gesture Recognition Using Deep Learning of Depth Images with Shadow Effect Removal for Smart Gesture Communication. In: Sensors and Materials. 34. Jg., 2022, Nr. 1, S. 203–216

[DIWA23] Diwan, T.; Anirudh, G.; Tembhurne, J. V.: Object detection using YOLO: challenges, architectural successors, datasets and applications. In: Multimedia tools and applications, 2023, Nr. 82, S. 9243–9275

[DOCU25] DocuWare GmbH: DocuWare. URL: https://start.docuware.com/de/. [Stand: 28.01.2025]

[DOGA19] Dogan, O.; Martinez-Millana, A.; Rojas, E.; Sepúlveda, M.; Munoz-Gama, J.; Traver, V.; Fernandez-Llatas, C.: Individual Behavior Modeling with Sensors Using Process Mining. In: Electronics. 8. Jg., 2019, Nr. 7, S. 766

[DOMB13] Dombrowski, U.; Mielke, T.; Anne Reimer: Dezentrales Wissensmanagement in Ganzheitlichen Produktionssystemen. In: Zeitschrift für wirtschaftlichen Fabrikbetrieb. 108. Jg., 2013, Nr. 6

[DUAN07] Duan, K.-B.; Rajapakse, J. C.; Nguyen, M. N.: One-Versus-One and One-Versus-All Multiclass SVM-RFE for Gene Selection in Cancer Classification. In: Marchiori, E., Moore, J.H., Rajapakse, J.C. (eds) Evolutionary

Computation,Machine Learning and Data Mining in Bioinformatics. Evo-BIO 2007. Lecture Notes in Computer Science, 2007, Nr. 4447

[EASY25] easy Software AG: easy Software. URL: https://easy-software.com/de/. [Stand: 28.01.2025]

[ELMA16] Elmasri, R.; Navathe, S.: Fundamentals of database systems. 7. Aufl. Hoboken NJ: Pearson Higher Education, 2016

[ELO 25] ELO Digital Office GmbH: elo. URL: https://www.elo.com/de-de/software/elo-ecm-suite/digitalisierungsplattform.html. [Stand: 28.01.2025]

[EPPL07] Eppler, M. J.; Burkhard, R. A.: Visual representations in knowledge management: framework and cases. In: Journal of Knowledge Management. 11. Jg., 2007, Nr. 4, S. 112–122

[FACC19] Faccio, M.; Ferrari, E.; Gamberi, M.; Pilati, F.: Human Factor Analyser for work measurement of manual manufacturing and assembly processes. In: The International Journal of Advanced Manufacturing Technology, 2019, Nr. 103, S. 861–877

[FEIX16] Feix, T.; Romero, J.; Schmiedmayer, H.-B.; Dollar, A. M.; Kragic, D.: The GRASP Taxonomy of Human Grasp Types. In: IEEE Transactions on Human-Machine Systems. 46. Jg., 2016, Nr. 1, S. 66–77

[FIGU13] Figura, M.; Gross, D.: Die Qual der Wiki-Wahl. In: Open Journal of Knowledge Management. 4. Jg., 2013

[FLIC95] Flick, U.; Kardorff, E. von; Keupp, H.; Rosenstiel, L. von; Wolff, S.: Handbuch qualitative Sozialforschung. 2. Aufl. Weinheim: Beltz Psychologie-Verl.-Union, 1995

[FREI24] Freire, S. K.; Wang, C.; Foosherian, M.; Wellsandt, S.; Ruiz-Arenas, S.; Niforatos, E.: Knowledge Sharing in Manufacturing using Large Language Models: User Evaluation and Model Benchmarking. In: arXiv, 2024

[FREY14] Frey-Luxemburger, M.: Wissensmanagement - Grundlagen und praktische Anwendung. (Reihe: SpringerLink Bücher). 2. Aufl. Wiesbaden: Springer Vieweg, 2014

[FULL19] Fullen, M.; Maier, A.; Nazarenko, A.; Aksu, V.; Jederny, S.; Röcker, C.: Machine learning for assistance systems: pattern-based approach to online step recognition. In: 2019 IEEE 17th International Conference on Industrial Informatics (INDIN), 2019, S. 296–302

[GAO24] Gao, Y.; Xiong, Y.; Gao, X.; Jia, K.; Pan, J.; Bi, Y.; Dai, Y.; Sun, J.; Wang, M.; Wang, H.: Retrieval-Augmented Generation for Large Language Models: A Survey. In: arXiv, 2024

[GEOR21] Georgescu, A.-L.; Pappalardo, A.; Cucu, H.; Blott, M.: Performance vs. hardware requirements in state-of-the-art automatic speech recognition.

In: EURASIP Journal on Audio, Speech, and Music Processing, 2021, Nr. 28

[GOOG24] Google: Hand landmarks detection guide. URL: https://ai.google.dev/edge/mediapipe/solutions/vision/hand_landmarker. [Stand: 12.11.2024]

[GOOG25] Google: Google Workspace. URL: https://workspace.google.com/intl/de/. [Stand: 28.01.2028]

[GRAM20] Grames, P. P.: Entwicklung eines Methodenkonzeptes des Wissensmanagements zur Sicherung von Erfahrungswissen, zur Vermeidung von Wissensverlust und zur Einarbeitung neuer Mitarbeiter. Hamburg, 2020

[GRUB92] Gruber, T. R.: A Translation Approach to Portable Ontology Specifications. In: Knowledge Acquisition. 5. Jg., 1992, Nr. 2, S. 199–220

[GUES06] Guest, G.; Bunce, A.; Johnson, L.: How Many Interviews Are Enough? In: Field Methods. 18. Jg., 2006, Nr. 1, S. 59–82

[GUPT22] Gupta, N.; Gupta, S. K.; Pathak, R. K.; Jain, V.; Rashidi, P.; Suri, J. S.: Human activity recognition in artificial intelligence framework: a narrative review. In: Artificial intelligence review, 2022, Nr. 55, S. 4755–4808

[GUPT24] Gupta, S.; Ranjan, R.; Singh, S. N.: A Comprehensive Survey of Retrieval-Augmented Generation (RAG): Evolution, Current Landscape and Future Directions. In: arXiv, 2024

[GURU25] Guru Technologies: Guru. URL: https://www.getguru.com/de. [Stand: 28.01.2025]

[HACK12] Hackermeier, I.: Wiki im Wissensmanagement: Determinanten der Akzeptanz eines Web 2.0 Projektes innerhalb eines internationalen Zulieferers der Automobilindustrie. München, 2012

[HALL25] Hallo Welt GmbH: BlueSpice. URL: https://bluespice.com/de/. [Stand: 28.01.2028]

[HANN14] Hannun, A.; Case, C.; Casper, J.; Catanzaro, B.; Diamos, G.; Elsen, E.; Prenger, R.; Satheesh, S.; Sengupta, S.; Coates, A.; Ng, A. Y.: Deep Speech: Scaling up end-to-end speech recognition. In: arXiv, 2014

[HANS13] Hansard, M.; Lee, S.; Choi, O.; Horaud, R.: Time-of-Flight Cameras. London: Springer, 2013

[HART03] Hartley, R.; Zisserman, A.: Multiple view geometry in computer vision. Cambridge UK: Cambridge University Press, 2003

[HEDD16] Hedden, H.: The accidental taxonomist. 2. Aufl. Medford New Jersey: Information Today Inc, 2016

[HERT13] Hertling, S.: Leitfaden - Wissen in Unternehmen halten und verteilen,RKW Rationalisierungs- und Innovationszentrum der Deutschen Wirtschaft e.V., 2013

[HEST11] Hester, A. J.: A comparative analysis of the usage and infusion of wiki and non-wiki-based knowledge management systems. In: Information Technology and Management, 2011, Nr. 12, S. 335–355

[HEZL20] Hezler, C.: Modernes Erfahrungs- und Wissensmanagement auf dem Shopfloorindustrieller Hightech-Produktionen, 2020

[HOER23] Hoerner, L.; Schamberger, M.; Bodendorf, F.: Using Tacit Expert Knowledge to Support Shop-floor Operators Through a Knowledge-based Assistance System. In: Computer Supported Cooperative Work (CSCW), 2023, Nr. 32, S. 55–91

[HOGA22] Hogan, A.; Blomqvist, E.; Cochez, M.; D'amato, C.; Melo, G. de; Gutierrez, C.; Kirrane, S.; Gayo, J. E. L.; Navigli, R.; Neumaier, S.; Ngomo, A.-C. N.; Polleres, A.; Rashid, S. M.; Rula, A.; Schmelzeisen, L.; Sequeda, J.; Staab, S.; Zimmermann, A.: Knowledge Graphs. In: ACM Computing Surveys. 54. Jg., 2022, Nr. 4, S. 1–37

[HOLG15] Holger Kohl, Ronald Orth and Erik Steinhöfel: A Practical Approach to Process-Oriented Knowledge Management. In: The Electronic Journal of Knowledge Management. 13. Jg., 2015, Nr. 1, S. 74–87

[HUSS23] Hussain, M.: YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection. In: Machines, 2023, Nr. 11, S. 677–701

[HUSS24] Hussain, M.: YOLOv5, YOLOv8 and YOLOv10: The Go-To Detectors for Real-time Vision. In: arXiv, 2024

[JARR23] Jarrahi, M. H.; Askay, D.; Eshraghi, A.; Smith, P.: Artificial intelligence and knowledge management: A partnership between human and AI. In: Business Horizons, 2023, Nr. 66, S. 87–99

[JING24] Jing, Z.; Su, Y.; Han, Y.: When Large Language Models Meet Vector Databases: A Survey. In: arXiv, 2024

[KADD23] Kaddour, J.; Harris, J.; Mozes, M.; Bradley, H.; Raileanu, R.; McHardy, R.: Challenges and Applications of Large Language Models. In: arXiv, 2023

[KHAN15] Khan, S.; Rani, U.; Prasad, B. V. N.; Srivastava, A. K.; Selvi, S.; Gautam, D. K.: Document Management System: An Explicit Knowledge Management System. In: 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom 2015), 2015, S. 402–405

[KLAP19] Klapper, J.; Gelec, E.; Pokorni, B.; Hämmerle, M.; Rothenberger, R.: Potenziale Digitaler Assistenzsysteme,Fraunhofer IAO, 2019

[KLUß22] Klußmann, A.; Schäfer, A.; Serafin, P.; Lang, K.-H.; Schust, M.; Liebers, F.: Manuelle Arbeitsprozesse,Bundesanstalt für Arbeitsschutz, 2022

[KNEB17] Knebel, F.; Lämmel, U.: Einsatz von Wiki-Systemen im Wissensmanagement. (Reihe: Wismarer Diskussionspapiere. Bd. Heft 2017, 02). Wismar: Hochschule Wismar Fakultät für Wirtschaftswissenschaften, 2017

[KOBA19] Kobayashi, T.; Aoki, Y.; Shimizu, S.; Kusano, K.; Okumura, S.: Fine-Grained Action Recognition in Assembly Work Scenes by Drawing Attention to the Hands. In: 2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), 2019, S. 440–446

[KOCH11] Koch, M.; Bentele, M.: Rheinmetall: Unified Communication and Collaboration (UC²). In: Schriftreihe zu Enterprise 2.0 Fallstudien, 2011

[KONT25] KontextWork: Drupal Wiki. URL: https://drupal-wiki.com/. [Stand: 28.01.2028]

[KURN19] Kurnaz, F. C.; Hocaoğlu, B.; Yılmaz, M. K.; Sülo, İ.; Kalkan, S.: ALET (Automated Labeling of Equipment and Tools): A Dataset, a Baseline and a Usecase for Tool Detection in the Wild. In: arXiv, 2019

[LAGA17] Laganiere, R.: OpenCV 3 Computer Vision Application Programming. 3. Aufl. Birmingham: Packt Publishing Ltd., 2017

[LAP 25] LAP GMBH LASER APPLIKATIONEN: Laserbasierte Werkerführung. URL: https://www.lap-laser.com/de/branchen/werkerfuehrung/. [Stand: 28.01.2025]

[LEE24] Lee, J.; Jung, W.; Baek, S.: In-House Knowledge Management Using a Large Language Model: Focusing on Technical Specification Documents Review. In: Applied Sciences, 2024, Nr. 24, S. 2096–2113

[LEHM22] Lehmann, M. A.; Porsch, R.; Mai, C.: Assembly Process Digitization Through Self-learning Assistance Systems in Production. In: Andersen, AL., et al. Towards Sustainable Customization: Bridging Smart Products and Manufacturing Systems. CARV MCPC 2021 2021. Lecture Notes in Mechanical Engineering, 2022, S. 216–223

[LI22] Li, Y.; Xue, Z.; Wang, Y.; Ge, Liuhao, Ren, Zhao, Rodriguez, Jonathan: End-to-End 3D Hand Pose Estimation from Stereo Cameras. In: arXiv, 2022

[LI23] Li, F.: Modernization of Databases in the Cloud Era: Building Databases that Run Like Legos. In: Proceedings of the VLDB Endowment. 16. Jg., 2023, Nr. 12, S. 4140–4151

[LIU14] Liu, K.; Chen, C.; Jafari, R.; Kehtarnavaz, N.: Fusion of Inertial and Depth Sensor Data for Robust Hand Gesture Recognition. In: IEEE Sensors Journal. 14. Jg., 2014, Nr. 6, S. 1898–1903

[LÜCK19] Lück, M.; Hünefeld, L.; Brenscheidt, S.; Bödefeld, M.; Hünefeld, A.: Grundauswertung der BIBB/BAuA-Erwerbstätigenbefragung 2018 (2. Auflage),Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA), 2019

[MAAR17] Maarif, M. R.: Revealing Daily Human Activity Pattern using Process Mining Approach. In: Proceeding of the Electrical Engineering Computer Science and Informatics. 4. Jg., 2017, S. 19–21

[MANN09] Manning, C.; Raghavan, P.; Schuetze, H.: Introduction to Information Retrieval. Cambridge: Cambridge University Press, 2009

[MANN18] Mannhardt, F.; Petersen, S. A.; Oliveira, M. F.: Privacy Challenges for Process Mining in Human-Centered Industrial Environments. In: 2018 International Conference on Intelligent Environments, 2018, S. 64–71

[MARC21] Marcus V. Goncalves: Knowledge management as a catalyst for business process digitalisation. In: Proceedings Applied Research International Conference 2021, 2021

[MARI16] Marin, G.; Dominio, F.; Zanuttigh, P.: Hand gesture recognition with jointly calibrated Leap Motion and depth sensor. In: Multimedia Tools and Applications. 75. Jg., 2016, Nr. 22, S. 14991–15015

[MEUD17] Meudt, T.; Pohl, M.; Metternich, J.: Die Automatisierungspyramide - Ein Literaturüberblick,TU Darmstadt, 2017

[MICR23a] Microsoft: Microsoft Hololense. URL: https://learn.microsoft.com/de-de/hololens/. [Stand: 24.12.2024]

[MICR23b] Microsoft: Enterprise Content Management. URL: https://learn.microsoft.com/de-de/sharepoint/dev/scenario-guidance/enterprise-content-management. [Stand: 28.012025]

[MICR25] Microsoft: Erstellen und Bearbeiten eines Wiki. URL: https://support.microsoft.com/de-de/office/erstellen-und-bearbeiten-eines-wikis-dc64f9c2-d1a2-44b5-ac59-b9d535551a32. [Stand: 28.01.2025]

[MILE14] Miles, M. B.; Huberman, A. M.; Saldaña, J.: Qualitative data analysis. 3. Aufl. Thousand Oaks Califorinia: SAGE Publications Inc, 2014

[MINI24] MiniTec GmbH: Smartes Lernen per Assistenzsystem, 2024

[MIYA07] Miyake, D. I.; Nakano, D.: Implementation of corporate production systems in the brazilian auto indutry: managing knowledge through practice. In: Int. J. Automotive Technology and Management. 7. Jg., 2007, S. 153–167

[MODE25] Modell Aachen GmbH: QWiki. URL: https://www.modell-aachen.de/de/qwiki. [Stand: 28.01.2025]

[MONC96] Moncrief, R. L.; Mott, S. J.; Behensky, M. L.; Margolin, J.: System and method for driver training with multiple driver competition. In: United States Patent, 1996

[MÖRL11] Mörl, S.; Heiss, M.; Richter, A.: Siemens: Wissensvernetzung mit Techno-Web2.0. In: Schriftreihe zu Enterprise 2.0-Fallstudien, 2011

[MUEN23] Muennighoff, N.; Tazi, N.; Magne, L.; Reimers, N.: MTEB: Massive Text Embedding Benchmark. In: Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics, 2023, S. 2014–2037

[MÜLL18] Müller-Roterberg, C.: Handbook of Desing Thinking, 2018

[MÜLL20] Müller, M.; Terziev, G.; Metternich, J.; Landmann, N.: Knowledge management on the shop floor through recommender engines. In: Procedia Manufacturing. 52. Jg., 2020, S. 344–349

[NAKA13] Nakano, D.; Muniz, J.; Dias Batista, E.: Engaging environments: tacit knowledge sharing on the shop floor. In: Journal of Knowledge Management. 17. Jg., 2013, Nr. 2, S. 290–306

[NONA07] Nonaka, I.: The knowledge-creating company. In: Harvard Business Review, 2007 (ursprgl. 1991)

[NONA95] Nonaka, I.; Takeuchi, H.: The Knowledge Creating Company: How Japanese Companies Create the Dynamics of Innovation. New York: Oxford University Press, 1995

[NORD21] Nordsieck, R.; Heider, M.; Winschel, A.; Hahner, J.: Knowledge Extraction via Decentralized Knowledge Graph Aggregation. In: 2021 IEEE 15th International Conference on Semantic Computing (ICSC), Laguna Hills, CA, USA, 2021, 2021, S. 92–99

[NORT18] North, K.; Maier, R.: Wissen 4.0 – Wissensmanagement im digitalen Wandel. In: HMD Praxis der Wirtschaftsinformatik. 55. Jg., 2018, Nr. 4, S. 665–681

[NORT21] North, K.: Wissensorientierte Unternehmensführung. (Reihe: Lehrbuch). 7. Aufl. Wiesbaden, Heidelberg: Springer Gabler, 2021

[NOTI25] Notion Labs: Notion. URL: https://www.notion.com/de. [Stand: 28.01.2025]

[NOVA08] Novak, J. D.; Canas, A. J.: The Theory Underlying Concept Maps and How to Construct Them. In: Technical Report IHMC CmapTools 2006-01 Rev 01-2008, Florida Institute for Human and Machine Cognition, 2008

[NURN23] Nurnaninsih, A.; Ahmad Muktamar, B.; Muthmainah, H. N.: Knowledge Management in the Digital Age: Harnessing Information and Innovation with Knowledge Management Systems. In: The Eastasouth Journal of Information System and Computer Science. 1. Jg., 2023, Nr. 1, S. 25–34

[OLIV22] Oliveira, V. F. de; Pessoa, M. A. d. O.; Junqueira, F.; Miyagi, P. E.: SQL and NoSQL Databases in the Context of Industry 4.0. In: Machines. 10. Jg., 2022, Nr. 1, S. 20

[OMRA24] Omrani, P.; Hosseini, A.; Hooshanfar, K.; Ebrahimian, Z.; Toosi, R.; Ali Akhaee, M.: Hybrid Retrieval-Augmented Generation Approach for LLMs Query Response Enhancement. In:: 2024 10th International Conference on Web Research (ICWR). Tagungsband zu 2024 10th International Conference on Web Research (ICWR). Tehran, Iran, Islamic Republic of, 24.04.2024 - 25.04.2024, S. 22–26

[OPEN24] OpenAI; Achiam, J.; Adler, S.; Agarwal, S.; Ahmad, L.; Akkaya, I.; Aleman, F. L.; Almeida, D.; Altenschmidt, J.; Altman, S.; Anadkat, S.; Avila, R.; Babuschkin, I.; Balaji, S.; Balcom, V.; Baltescu, P.; Bao, H.; Bavarian, M.; Belgum, J.; Bello, I.; Berdine, J.; Bernadett-Shapiro, G.; Berner, C.; Bogdonoff, L.; Boiko, O.; Boyd, M.; Brakman, A.-L.; Brockman, G.; Brooks, T.; Brundage, M.; Button, K.; Cai, T.; Campbell, R.; Cann, A.; Carey, B.; Carlson, C.; Carmichael, R.; Chan, B.; Chang, C.; Chantzis, F.; Chen, D.; Chen, S.; Chen, R.; Chen, J.; Chen, M.; Chess, B.; Cho, C.; Chu, C.; Chung, H. W.; Cummings, D.; Currier, J.; Dai, Y.; Decareaux, C.; Degry, T.; Deutsch, N.; Deville, D.; Dhar, A.; Dohan, D.; Dowling, S.; Dunning, S.; Ecoffet, A.; Eleti, A.; Eloundou, T.; Farhi, D.; Fedus, L.; Felix, N.; Fishman, S. P.; Forte, J.; Fulford, I.; Gao, L.; Georges, E.; Gibson, C.; Goel, V.; Gogineni, T.; Goh, G.; Gontijo-Lopes, R.; Gordon, J.; Grafstein, M.; Gray, S.; Greene, R.; Gross, J.; Gu, S. S.; Guo, Y.; Hallacy, C.; Han, J.; Harris, J.; He, Y.; Heaton, M.; Heidecke, J.; Hesse, C.; Hickey, A.; Hickey, W.; Hoeschele, P.; Houghton, B.; Hsu, K.; Hu, S.; Hu, X.; Huizinga, J.; Jain, S.; Jain, S.; Jang, J.; Jiang, A.; Jiang, R.; Jin, H.; Jin, D.; Jomoto, S.; Jonn, B.; Jun, H.; Kaftan, T.; Kaiser, Ł.; Kamali, A.; Kanitscheider, I.; Keskar, N. S.; Khan, T.; Kilpatrick, L.; Kim, J. W.; Kim, C.; Kim, Y.; Kirchner, J. H.; Kiros, J.; Knight, M.; Kokotajlo, D.; Kondraciuk, Ł.; Kondrich, A.; Konstantinidis, A.; Kosic, K.; Krueger, G.; Kuo, V.; Lampe, M.; Lan, I.; Lee, T.; Leike, J.; Leung, J.; Levy, D.; Li, C. M.; Lim, R.; Lin, M.; Lin, S.; Litwin, M.; Lopez, T.; Lowe, R.; Lue, P.; Makanju, A.; Malfacini, K.; Manning, S.; Markov, T.; Markovski, Y.; Martin, B.; Mayer, K.; Mayne, A.; McGrew, B.; McKinney, S. M.; McLeavey, C.; McMillan, P.; McNeil, J.; Medina, D.; Mehta, A.; Menick, J.; Metz, L.; Mishchenko, A.; Mishkin, P.; Monaco, V.; Morikawa, E.; Mossing, D.; Mu, T.; Murati, M.; Murk, O.; Mély, D.; Nair, A.; Nakano, R.; Nayak, R.; Neelakantan, A.; Ngo, R.; Noh, H.; Ouyang, L.; O'Keefe, C.; Pachocki, J.; Paino, A.; Palermo, J.; Pantuliano, A.; Parascandolo, G.; Parish, J.; Parparita, E.; Passos, A.; Pavlov, M.; Peng, A.; Perelman, A.; Peres, F. d. A. B.; Petrov, M.; Pinto, H. P. d. O.; Michael; Pokorny; Pokrass, M.; Pong, V. H.; Powell, T.; Power, A.; Power, B.; Proehl, E.; Puri, R.; Radford, A.; Rae, J.; Ramesh, A.; Raymond, C.; Real, F.; Rimbach, K.; Ross, C.; Rotsted, B.; Roussez, H.; Ryder, N.; Saltarelli, M.; Sanders, T.; Santurkar, S.; Sastry, G.; Schmidt, H.; Schnurr, D.; Schulman, J.; Selsam, D.; Sheppard, K.; Sherbakov, T.; Shieh, J.; Shoker, S.;

Shyam, P.; Sidor, S.; Sigler, E.; Simens, M.; Sitkin, J.; Slama, K.; Sohl, I.; Sokolowsky, B.; Song, Y.; Staudacher, N.; Such, F. P.; Summers, N.; Sutskever, I.; Tang, J.; Tezak, N.; Thompson, M. B.; Tillet, P.; Tootoonchian, A.; Tseng, E.; Tuggle, P.; Turley, N.; Tworek, J.; Uribe, J. F. C.; Vallone, A.; Vijayvergiya, A.; Voss, C.; Wainwright, C.; Wang, J. J.; Wang, A.; Wang, B.; Ward, J.; Wei, J.; Weinmann, C. J.; Welihinda, A.; Welinder, P.; Weng, J.; Weng, L.; Wiethoff, M.; Willner, D.; Winter, C.; Wolrich, S.; Wong, H.; Workman, L.; Wu, S.; Wu, J.; Wu, M.; Xiao, K.; Xu, T.; Yoo, S.; Yu, K.; Yuan, Q.; Zaremba, W.; Zellers, R.; Zhang, C.; Zhang, M.; Zhao, S.; Zheng, T.; Zhuang, J.; Zhuk, W.; Zoph, B.: GPT-4 Technical Report. In: arXiv, 2024

[OPEN25] Opentext Corporation: Opentext. URL: https://www.opentext.com/de-de. [Stand: 28.01.2025]

[ORTH08] Orth, R.; Voigt, S.: Wissen greifbar machen: Den Umgang mit Wissen beschreiben und bewerten,Fraunhofer IPK und Fraunhofer IFF, 2008

[OTT15] Ott, G.: Nutzung von Wissenslandkarten zur Verwaltung von Wissenskapital in Unternehmen. In: eBusinessLotse, 2015

[PAAS20] Paass, G.; Hecker, D.: Künstliche Intelligenz. Wiesbaden, Heidelberg: Springer, 2020

[PART24] Parthasarathy, V. B.; Zafar, A.; Khan, A.; Shahid, A.: The Ultimate Guide to Fine-Tuning LLMs from Basics to Breakthroughs: An Exhaustive Review of Technologies, Research, Best Practices, Applied Research Challenges and Opportunities. In: arXiv, 2024

[PAWL06] Pawlowsky, P.; Gerlach, L.; Hauptmann, S.; Puggel, A.: Wissen als Wettbewerbsvorteil in kleinen und mittelständischen Unternehmen,Forschungsstelle Organisationale Kompetenz und Strategie, 2006

[PETA20] Petana, G.; Rosa, C.: Digital Transformation and the Impact in Knowledge Management. In: International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management. 3. Jg., 2020, S. 180–187

[PETE19] Peters, R.; Nauroth, M.: Process-Mining. Wiesbaden: Springer Gabler, 2019

[PETR23] Petrovic, O.; Dassen, Y.; Brecher, C.: Potenziale und Anwendung des Industrial Metaverse. In: Industrie 4.0 Management. 39. Jg., 2023, Nr. 5, S. 27–32

[PLAA24] Plaat, A.; Wong, A.; Verberne, S.; Broekens, J.; van Stein, N.; Back, T.: Reasoning with Large Language Models, a Survey. In: arXiv, 2024

[PLAT18] Plattform Industrie 4.0: RAMI4.0, 22.11.2018

[PRAT24] Pratiwi, M. E.; Ramadhan, Y.; Sensuse, D. I.; Lusa, S.; Safitri, N.; Elisabeth, D.: Knowledge Capture in Agile Organizations: Methods and Strategies for Enhancing Effective and Efficient Process. In: International Journal of Artificial Intelligence Research. 8. Jg., 2024, Nr. 1, S. 65

[PREE01] Preece, A.; Flett, A.; Sleeman, D.; Curry, D.; Meany, N.; Perry, P.: Better knowledge management through knowledge engineering. In: IEEE Intelligent Systems. 16. Jg., 2001, Nr. 1, S. 36–43

[PTC21] PTC: vuforia expert capture. URL: https://www.tristar.com/wp-content/uploads/2022/03/Vuforia-Expert-Capture-Product-Brief.pdf. [Stand: 11.03.2025]

[QT G25] Qt Group: Qt Framework. URL: https://www.qt.io/product/framework. [Stand: 03.01.2025]

[RADF23] Radford, A.; Kim, J. W.; Xu, T.; Brockman, G.; McLeavey, C.; Sutskever, I.: Robust Speech Recognition via Large-Scale Weak Supervision. In: Proceedings of the 40th International Conference on Machine Learning, 2023

[RAJP23] Rajpurohit, R.: Conversational AI with RASA: A Comprehensive Guide. URL: https://medium.com/@rakeshrajpurohit/conversational-ai-with-rasa-a-comprehensive-guide-45073e830c03. [Stand: 28.01.2025]

[RASA24a] rasa: NLU Training Data. URL: https://rasa.com/docs/rasa/nlu-training-data/. [Stand: 26.11.2024]

[RASA24b] rasa: Stories. URL: https://rasa.com/docs/rasa/stories/. [Stand: 26.11.2024]

[RASA24c] rasa: Custom Actions. URL: https://rasa.com/docs/rasa/custom-actions. [Stand: 11.02.2025]

[REDD23] Reddy, V. M.; Vaishnavi, T.; Kumar, K. P.: Speech-to-Text and Text-to-Speech Recognition Using Deep Learning. In: 2023 2nd International Conference on Edge Computing and Applications (ICECAA), 2023, S. 657–666

[REDD24] Reddy Yerramreddy, D.; Marasani, J.; Venkata Gowtham, P. S.; Harshit, G.; Anjali: Speech Recognition Paradigms: A Comparative Evaluation of SpeechBrain, Whisper and Wav2Vec2 Models. In: 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), 2024, S. 1–6

[REDM16] Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A.: You Only Look Once: Unified, Real-Time Object Detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, S. 779–788

[REIM19] Reimers, N.; Gurevych, I.: Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In: arXiv, 2019

[RENZ24] Renze, M.; Guven, E.: The effect of sampling temperature on problem solving in large language models. In: Findings of the Association for Computational Linguistics EMNLP, 2024, S. 7346–7356

[REQU25] Requarks: Wiki.js. URL: https://js.wiki/. [Stand: 85.01.2025]

[RITT24] Ritter, R.: Concept Mapping als Methode zur Messung vernetzten Wissens? In: Wehner, A.; Masanek, N.; Hellmann, K.; Heinz, T.; Grospietsch, F. (Hrsg.): Vernetzung von Wissen bei Lehramtsstudierenden. Eine Black-Box für die Professionalisierungsforschung?: Verlag Julius Klinkhardt, 2024, S. 167–183

[RIZZ04] Rizzolatti, G.; Craighero, L.: The mirror-neuron system. In: Annual review of neuroscience. 27. Jg., 2004, S. 169–192

[ROBE13] Robertson, S.; Robertson, J.: Mastering the Requirements Process. 3. Aufl. New York, USA: Addison-Wesley, 2013

[ROBE24] Roberto, R.; Breyer, F.; Betts, J.; Newnes, L.; Shokrani, A.: A Study of Commercial Augmented Reality Applications in Manufacturing: A Subject Matter Expert Analysis. In: Procedia CIRP. 128. Jg., 2024, S. 186–191

[ROGE22] Roger, V.; Farinas, J.; Pinquier, J.: Deep neural networks for automatic speech processing: a survey from large corpora to limited data. In: EURASIP Journal on Audio, Speech, and Music Processing, 2022, Nr. 19

[ROSM20] Rosman, M. R. M.: The 5Ws of Enterprise Content Management (ECM) Research. In: Open Journal of Science and Technology. 3. Jg., 2020, Nr. 1, S. 46–70

[SABO14] Sabou, M.; Kovalenko, O.; Ekaputra, F.; Biffle, S.: Beiträge des Semantic Web zum Engineering für Industrie 4.0. In: Vogel-Heuser, B.; Bauernhansl, T.; Hompel, M. ten (Hrsg.): Handbuch Industrie 4.0 Bd.2. Berlin: Springer Vieweg, 2014, S. 293–314

[SAHO24] Sahoo, P.; Singh, A. K.; Saha, S.; Jain, V.; Mondal, S.; Chadha, A.: A Systematic Survey of Prompt Engineering in Large Language Models: Techniques and Applications. In: arXiv, 2024

[SÁNC22] Sánchez Ramírez, S.; Guadamillas Gómez, F.; González Ramos, M. I.; Grieva, O.: The Effect of Digitalization on Innovation Capabilities through the Lenses of the Knowledge Management Strategy. In: Administrative Sciences. 12. Jg., 2022, Nr. 4, S. 144

[SANT20] Santos, C. C. d.; Samatelo, J. L. A.; Vassallo, R. F.: Dynamic Gesture Recognition by Using CNNs and Star RGB: a Temporal Information Condensation. In: Neurocomputing. 400. Jg., 2020, S. 238–254

[SANU23] Sanugommula, H.: Exploring Confluence: Enhancing Collaboration and Knowledge Management in Modern Organizaitons. In: International Journal for Multidisciplinary Research. 5. Jg., 2023, Nr. 4

[SAUE24] Sauer, S.; Wohlrabe, K.: Wettbewerbsfähigkeit der deutschen Industrie im freien Fall? In: ifo Schnelldienst digital, 2024, Nr. 9

[SCHA05] Schaffert, S.; Gruber, A.; Westenthaler, R.: A Semantic Wiki for Collaborative Knowledge Formation. In: Proceedings of SEMANTICS 2005 Conference, 2005

[SCHA07] Schaffert, S.; Bry, F.; Baumeister, J.; Kiesel, M.: Semantic Wiki. In: Informatik-Spektrum. 30. Jg., 2007, Nr. 6, S. 434–439

[SCHA16] Schafermeyer, R. G.; Hoffman, R. R.: Using Knowledge Libraries to Transfer Expert Knowledge. In: IEEE Intelligent Systems. 31. Jg., 2016, Nr. 2, S. 89–93

[SCHI23] Schiedermair, I.; Kick, E.; Baumgartner, M.; Kopp, T.; Kinkel, S.: Wissensmanagement in KMU. In: Zeitschrift für wirtschaftlichen Fabrikbetrieb. 118. Jg., 2023, Nr. 6, S. 395–399

[SCHN19] Schneider, P.; Memmesheimer, R.; Kramer, I.; Paulus, D.: Gesture Recognition in RGB Videos UsingHuman Body Keypoints and Dynamic Time Warping. In: Chalup, S., Niemueller, T., Suthakorn, J., Williams, MA. (eds) RoboCup 2019: Robot World Cup XXIII. RoboCup 2019. Lecture Notes in Computer Science(), 2019

[SCHU19] Schuh, G.; Kelzenberg, C.; Wiese, J.; Ochel, T.: Data Structure of the Digital Shadow for Systematic Knowledge Management Systems in Single and Small Batch Production. In: Procedia CIRP. 84. Jg., 2019, S. 1094–1100

[SCHW24] Schwarz, D.: 5 Software-Lösungen für eine Wissensdatenbank. URL: https://www.bitrix24.de/articles/5-software-loesungen-fuer-eine-wissensdatenbank.php. [Stand: 28.01.2025]

[SHAR20] Sharma, R. K.; Joshi, M.: An Analytical Study and Review of open Source Chatbot framework, RASA. In: International Journal of Engineering Research & Technology (IJERT). 9. Jg., 2020, Nr. 06

[SHAR23] Sharma, H.; Kumar, H.; Gupta, A.; Shah, M. A.: Computer vision in manufacturing: a bibliometric analysis and future research propositions. In: The International Journal of Advanced Manufacturing Technology. 127. Jg., 2023, Nr. 11-12, S. 5691–5710

[SHEN18] Shen, J.; Pang, R.; Weiss, R. J.; Schuster, M.; Jaitly, N.; Yang, Z.; Chen, Z.; Zhang, Y.; Wang, Y.; Skerry-Ryan, R. J.; Saurous, R. A.; Agiomyrgiannakis, Y.; Wu, Y.: Natural TTS Synthesis by Conditioning WaveNet on Mel

Spectrogram Predictions. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018

[SIMO17] Simon, T.; Joo, H.; Matthews, I.; Sheikh, Y.: Hand Keypoint Detection in Single Images using Multiview Bootstrapping. In: arXiv, 2017

[SLIT25] Slite: Slite. URL: https://slite.com/. [Stand: 28.01.2025]

[SPAR98] Sparrow, J.: Knowledge in Organizations: Acess to Thinking at Work. Thousand Oaks, CA: Sage Publications, 1998

[STAA02] Staab, S.: Wissensmanagement mit Ontologien und Metadaten. In: Informatik-Spektrum. 25. Jg., 2002, Nr. 3, S. 194–209

[STE-12] Ste-Marie, D. M.; Law, B.; Rymal, A. M.; Jenny, O.; Hall, C.; McCullagh, P.: Observation interventions for motor skill learning and performance: an applied model for the use of observation. In: International Review of Sport and Exercise Psychology. 5. Jg., 2012, Nr. 2, S. 145–176

[STEI10] Steinhüser, M.; Räth, P.: ABB AG: Blog und Wiki in der Unternehmenskommunikation. In: Schriftreihe zu Enterprise 2.0 Fallstudien, 2010

[STOC10] Stocker, A.; Tochtermann, K.: Wissenstransfer mit Wikis und Weblogs: Fallstudien zum erfolgreichen Einsatz von Web 2.0 in Unternehmen: Gabler, 2010

[SUN23] Sun, L.; Kang, S.; Liu, X.; Meng, H.: Speech Recognition and Text-to-Speech Synthesis. In: Huang, C.-R.; Hsieh, S.-K.; Jin, P. (Hrsg.): Chinese language resources. (Reihe: Text, speech and language technology. Bd. volume 49). Cham: Springer, 2023

[SUN24] Sun, S.; Yuan, R.; Cao, Z.; Li, W.; Liu, P.: Prompt Chaining or Stepwise Prompt? Refinement in Text Summarization. In: Findings of the Association for Computational Linguistics: ACL 2024, 2024, S. 7551–7558

[SURI12] Suriya Prakash, J.; Annamalai Vignesh, K.; Ashok, C.; Adithyan, R.: Multi class Support Vector Machines classifier for machine vision application. In: 2012 International Conference on Machine Vision and Image Processing (MVIP), 2012, S. 197–199

[TAHE23] Taherdoost, H.; Madanchian, M.: Artificial Intelligence and Knowledge Management: Impacts, Benefits, and Implementation. In: Computers. 12. Jg., 2023, Nr. 4, S. 72–89

[TAIP24] Taipalus, T.: Vector database management systems: Fundamental concepts, use-cases, and current challenges. In: Cognitive Systems Research. 85. Jg., 2024, S. 101216

[TERV23] Terven, J.; Córdova-Esparza, D.-M.; Romero-González, J.-A.: A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1

to YOLOv8 and YOLO-NAS. In: Machine Learning and Knowledge Extraction. 5. Jg., 2023, Nr. 4, S. 1680–1716

[THAK21] Thakur, N.; Reimers, N.; Daxenberger, J.; Gurevych, I.: Augmented SBERT: Data Augmentation Method for Improving Bi-Encoders for Pairwise Sentence Scoring Tasks. In: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2021, S. 296–310

[THEI19] Theis, J.; Darabi, H.: Behavioral Petri Net Mining and Automated Analysis for Human-Computer Interaction Recommendations in Multi-Application Environments. In: Proceedings of the ACM on Human-Computer Interaction. 3. Jg., 2019, Nr. EICS

[TOUV23] Touvron, H.; Lavril, T.; Izacard, G.; Martinet, X.; Lachaux, M.-A.; Lacroix, T.; Rozière, B.; Goyal, N.; Hambro, E.; Azhar, F.; Rodriguez, A.; Joulin, A.; Grave, E.; Lample, G.: LLaMA: Open and Efficient Foundation Language Models. In: arXiv, 2023

[VAN 12] van der Aalst, W. M. P.: Process Mining: Overview and Opportunities. In: ACM Transactions on Management Information Systems, 2012, Nr. 99

[VAN 16a] van der Aalst, W. M. P.: Process Mining. In: Liu, L.; Özsu, M. T. (Hrsg.): Encyclopedia of Database Systems. New York, NY: Springer New York, 2016, S. 1–3

[VAN 16b] van der Aalst, W. M. P.: Process Mining. 2. Aufl. Berlin Heidelberg: Springer, 2016

[VANI12] Vanini, U.; Hauschildt, J.: Wissensmanagement in KMU – Erfahrungen und Implikationen aus dem Praxistest eines Wissensmanagement-Audits in schleswig-holsteinischen Unternehmen. In: Management Summary: Wissensmangement in KMU, 2012

[VASW17] Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; Polosukhin, I.: Attention Is All You Need. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017, S. 6000–6010

[VDI09] VDI: VDI 5610 - Wissensmangement im Ingenierwesen, 2009

[WAGN04] Wagner, C.: Wiki: A Technology for Conversational Knowledge Management and Group Collaboration. In: Communications of the Association for Information Systems. 13. Jg., 2004

[WANG23] Wang, Z.: Design of Knowledge Management Based on Support Vector Machine. In: Procedia Computer Science. 228. Jg., 2023, S. 1168–1176

[WANG24] Wang, A.; Chen, H.; Liu, L.; Chen, K.; Lin, Z.; Han, J.; Ding, G.: YOLOv10: Real-Time End-to-End Object Detection. In: Proceedings of the

38th International Conference on Neural Information Processing Systems, 2024

[WEI22] Wei, J.; Tay, Y.; Bommasani, R.; Raffel, C.; Zoph, B.; Borgeaud, S.; Yogatama, D.; Bosma, M.; Zhou, D.; Metzler, D.; Chi, E. H.; Hashimoto, T.; Vinyals, O.; Liang, P.; Dean, J.; Fedus, W.: Emergent Abilities of Large Language Models. In: Transactions on Machine Learning Research, 2022

[WU24] Wu, S.; Xiong, Y.; Cui, Y.; Wu, H.; Chen, C.; Yuan, Y.; Huang, L.; Liu, X.; Kuo, T.-W.; Guan, N.; Xue, C. J.: Retrieval-Augmented Generation for Natural Language Processing: A Survey. In: arXiv, 2024

[ZANU16] Zanuttigh, P.; Marin, G.; Dal Mutto, C.; Dominio, F.; Minto, L.; Cortelazzo, G. M.: Time-of-Flight and Structured Light Depth Cameras. 1. Aufl: Springer, 2016

[ZAPP13] Zapp, M.; Hoffmeister, M.; Verl, A.: Methodology to Apply Semantic Wikis as Lean Knowledge Management Systems on the Shop Floor. In: Procedia CIRP. 12. Jg., 2013, S. 444–449

[ZAPP14] Zapp, M.: Semantische Wiki-Systeme in der wandlungsfähigen Produktion. (Reihe: Stuttgarter Beiträge zur Produktionsforschung. Bd. Bd. 30). Stuttgart: Fraunhofer-Verl., 2014

[ZHAN16] Zhang, J.; Jiao, J.; Chen, M.; Qu, L.; Xu, X.; Yang, Q.: 3D Hand Pose Tracking and Estimation Using Stereo Matching. In: arXiv, 2016

[ZHAN17] Zhang, Z.: Graph Databases for Knowledge Management. In: IT Professional. 19. Jg., 2017, Nr. 6, S. 26–32

[ZHAN20] Zhang, F.; Bazarevsky, V.; Vakunov, A.; Tkachenka, A.; Sung, G.; Chang, C.-L.; Grundmann, M.: MediaPipe Hands: On-device Real-time Hand Tracking. In: arXiv, 2020

[ZHAN25] Zhang, Y.; Li, Y.; Cui, L.; Cai, D.; Liu, L.; Fu, T.; Huang, X.; Zhao, E.; Zhang, Y.; Chen, Y.; Wang, L.; Luu, A. T.; Bi, W.; Shi, F.; Shi, S.: Siren's Song in the AI Ocean: A Survey on Hallucination in Large Language Models. In: Computational Linguistics 2025. 54. Jg., 2025, Nr. 4, S. 1373–1418

[ZHAO24a] Zhao, W. X.; Zhou, K.; Li, J.; Tang, T.; Wang, X.; Hou, Y.; Min, Y.; Zhang, B.; Zhang, J.; Dong, Z.; Du Yifan; Yang, C.; Chen, Y.; Chen, Z.; Jiang, J.; Ren, R.; Li, Y.; Tang, X.; Liu, Z.; Liu, P.; Nie, J.-Y.; Wen, J.-R.: A Survey of Large Language Models. In: arXiv, 2024

[ZHAO24b] Zhao, P.; Zhang, H.; Yu, Q.; Wang, Z.; Geng, Y.; Fu, F.; Yang, L.; Zhang, W.; Jiang, J.; Cui, B.: Retrieval-Augmented Generation for AI-Generated Content: A Survey. In: arXiv, 2024

[ZIMM18] Zimmermann, V.: Digitalisation in German SMEs: state of implementation and investment,KfW Research, 2018

# VI. Anhang

*Appendix*

## A. Anhang zum Stand der Technik und Forschung

**Tabelle A.1: Übersicht bekannter Datenbankanbieter**

*Overview of known Database Providers*

| Datenbanktyp | Kommerziell | Open-Source |
|---|---|---|
| **SQL** | Oracle Database | MySQL |
| | Microsoft SQL Server | PostgreSQL |
| | Amazon RDS | MariaDB |
| **NoSQL** | Amazon DynamoDB | MongoDB |
| | Google Firestore | Apache Cassandra |
| | Azure Cosmos DB | RavenDB |
| **Graph-Datenbanken** | Neo4j | JanusGraph |
| | Amazon Neptune | ArangoDB |
| | TigerGraph | Dgraph |
| **Vector Stores** | Pinecone | FAISS |
| | Zilliz Cloud | Chroma |
| | Weaviate | Milvus |

**Tabelle A.2: Die fünf am häufigsten genutzten ECM-Systeme in KMU nach [BITK17] inklusive ihrer beworbenen Funktionalitäten**

*The Five Most Frequently Used ECM Systems in SMEs According to [BITK17] Including Their Advertised Functionalities*

| Anbieter | Funktionalitäten |
|---|---|
| **DocuWare** [DOCU25] | • Cloudbasiertes Dokumentenmanagement<br>• Workflow-Automatisierung in den Bereichen<br>  ○ Auftragsabwicklung<br>  ○ Rechnungsprozesse<br>  ○ Vertragsverwaltung<br>  ○ Qualitätsmanagement |
| **easy Software** [EASY25] | • Rechtskonforme Datenarchivierung<br>• Rechnungsprüfprozesse<br>• Vertragsverwaltung<br>• HR-Automation |
| **Microsoft SharePoint** [MICR23b] | • Kollaboratives Dokumentenmanagement<br>• Web-Content-Management<br>• Integration in Microsoft 365 (z.B. Teams, Outlook)<br>• KI-basierte Unterstützung für die Dokumentenerstellung |
| **ELO** [ELO 25] | • Digitales Dokumentenmanagement<br>• Workflow-Management und Automatisierung von Geschäftsprozessen<br>• Low-Code Prozessmodellierung<br>• Integration in unter anderem Microsoft 365, SAP, Salesforce |
| **OpenText** [OPEN25] | • Dokumentenverwaltung<br>• Prozessautomatisierung<br>• Datenerfassung und intelligente Verarbeitung von Dokumenten<br>• Integration in SAP, Microsoft 365, Salesforce |

**Tabelle A.3: Übersicht aktuell verfügbarer Anbieter von Wikis und ihrer beworbenen Funktionalitäten**

*Overview of Currently Available Wiki Providers and Their Advertised Functionalities*

| Anbieter | Funktionalitäten |
|---|---|
| **BlueSpice** [HALL25] | • Open-Source Wiki auf Basis von MediaWiki<br>• ISO-konformes Qualitäts-, Umwelt- und Risikomanagement<br>• IT- und technische Dokumentation |
| **Confluence** [ATLA25] | • Umfangreiche Tools für Wissensmanagement, Kollaboration und Dokumentation<br>• Flexible Templates und Dashboards<br>• KI-Funktionen zur Content-Erstellung |
| **Drupal Wiki** [KONT25] | • Flexibles Open-Source Wiki<br>• Dokumentenbasiertes Wissensmanagement mit Rechtemanagement und Suchfunktionalitäten<br>• Prozess- und Qualitätsmanagement |
| **Google Workspace** [GOOG25] | • Einfache Erstellung und Verwaltung von Dokumenten<br>• Grundlegende Wissensmanagement-Funktionen<br>• Nahtlose Integration in Google-Dienste<br>• KI-Funktionen zur Content-Erstellung |
| **Notion** [NOTI25] | • Vielseitige All-in-One-Plattform für Wissens- und Projektmanagement<br>• Dokumentenmanagement inklusiver leistungsstarker Suche<br>• Integration von Daten aus Projektmanagementtools |
| **QWiki** [MODE25] | • Speziell für Prozessdokumentation und Wissensmanagement<br>• Strukturierte Organisation<br>• ISO-konformes Qualitätsmanagement |
| **Microsoft SharePoint** [MICR25] | • Kollaboratives Dokumentenmanagement<br>• Web-Content-Management<br>• Integration in Microsoft 365 (z.B. Teams, Outlook)<br>• KI-basierte Unterstützung für die Dokumentenerstellung |
| **Slite** [SLIT25] | • Kollaborativer Dokumentenmanagement<br>• Templates für die Content-Erstellung<br>• Ki-Funktionen zur Verwaltung, Aktualisierung und Archivierung von Dokumenten<br>• Chatbot zur Abfrage von Inhalten<br>• Integration einer großen Palette von Office-Tools |
| **Wiki.js** [REQU25] | • Open-Source Wiki mit moderner Benutzeroberfläche und modularer Architektur<br>• Asset-Manager zur Verwaltung multimodaler Daten<br>• Leistungsstarke Suchfunktion |

**Tabelle A.4: Aufstellung der von Bertram und Birtel betrachteten kommerziellen kognitiven Assistenzsysteme inklusive ihrer Funktionalitäten [BERT18]**

*List of Commercial Cognitive Assistance Systems Considered by Bertram and Birtel, Including Their Functionalities [BERT18]*

| Anbieter | Name | Funktionalitäten |
|---|---|---|
| **Bosch Rexroth** | Active-Assist | • Konfigurierbare Webplattform mit kontextbezogener Informationsbereitstellung<br>• Standardisierte Schnittstellen für zusätzliche Systemkomponenten (z.B. Pick-to-Light, RFID)<br>• Integration in übergeordnete IT-Systeme wie MES und ERP möglich<br>• Flexible, erweiterbare Architektur mit intuitiver Benutzeroberfläche |
| **Optimum datamanagement solutions GmbH** | Schlauer Klaus | • Visualisierung von relevanten Daten aus MES und ERP<br>• Identifikation von Bauteilen und Anweisung der nächsten Schritte, z.B. Lagerplatzzuweisung bei Wareneingang<br>• Schrittweise Anleitung der Mitarbeitenden während des Montageprozesses; Überwachung und automatische Fortschrittsanzeige<br>• Unterstützung bei komplexen Prüfaufgaben oder der Kontrolle von Produkten<br>• Unterstützung bei der Platzierung von Endprodukten oder Bauteilen, z.B. für Transport oder Weiterverarbeitung |
| **Schnaithmann Maschinenbau GmbH** | cubu:S | • In-situ-Projektion zur Anzeige der Montageschritte auf der Arbeitsfläche<br>• Pick-by-Light-System zur Hervorhebung der benötigten Komponenten für den nächsten Arbeitsschritt<br>• Identifikation von Position und Identität der Bauteile im Arbeitsbereich<br>• Akustische oder visuelle Warnsignale bei falscher Auswahl von Komponenten |
| **Sight Proc GmbH** | Der Assistent | • All-in-one-System ausgestattet mit physischen Schnittstellen sowie integrierten Sensor- und Aktor-Schnittstellen<br>• Multi-Sensor-Kamera zur Überwachung des Arbeitsfortschritts und Projektor zur Anzeige von Videos, Bildern und Texten auf der Arbeitsfläche<br>• Softwareplattform „ARP“ als App-Store mit Unterstützung für externe Anwendungen und eigener App-Entwicklung |

# B. Anhang zur Anforderungsanalyse

**Tabelle B.5: Abgeleitete Anforderungen auf Basis der Kernaussagen der Studienteilnehmenden in der Sektion Wissensmanagement**

*Derived Requirements Based on the Core Statements of the Study Participants in the Knowledge Management Section*

| Sektion | Kernaussage | Anforderung | Nr. |
|---|---|---|---|
| **Wissensmanagement** | Durch den Abgang von Mitarbeitenden aus dem Unternehmen geht Wissen verloren. | Realisierung des **Transfers von individuellem Wissen in Unternehmenswissen.** | 6 |
| | Zur Dokumentation von Wissen werden Protokolle erstellt, die teilweise analog mit Stift und Papier entstehen. | Realisierung von Möglichkeiten zur **intuitiven, prozessparallelen, digitalen Dokumentation**. | 7 |
| | Eine zentrale Wissensquelle, kombiniert mit uneingeschränktem Zugriff und intuitiver Datenvisualisierung, kann zu einer schnelleren Abwicklung von Aufträgen führen. | Realisierung einer **zentralen Wissensquelle** mit **uneingeschränktem Zugriff** und **intuitiver Datenvisualisierung**. | 8 |
| | Mangelnde Strukturen bedingen einen hohen Aufwand beim Dokumentieren von Wissen. | **Reduktion des Zeitaufwands bei der Dokumentation** von Wissen. | **9** |
| | Der Großteil des Wissens ist noch stark an Mitarbeitende gebunden und nicht systematisch dokumentiert. | Realisierung von Möglichkeiten zur **systematischen Dokumentation** von Wissen. | 10 |
| | Durch Etablierung einer strukturierten Wissensbasis können Schulungsmaßnahmen und Onboarding-Prozesse optimiert werden. | Realisierung von Möglichkeiten zur **strukturierten Dokumentation** und **zentralen Speicherung** von Wissen. | 11 |
| | Werden Fehler bei der Speicherung von Wissen gemacht, kann dies weitreichende Auswirkungen haben. | Realisierung von Mechanismen zur **Validierung** des gespeicherten Wissens. | 12 |
| | Die Pflege von Wissensmanagementsystemen kann einen hohen Aufwand erfordern und dadurch ineffizient sein. | Realisierung von **effizientem Wissensmanagement** mit möglichst **geringem Aufwand**. | 13 |

**Tabelle B.6: Abgeleitete Anforderungen auf Basis der Kernaussagen der Studienteilnehmenden in der Sektion Gestaltungsvorstellungen**

*Derived Requirements Based on the Core Statements of the Study Participants in the Design Ideas Section*

| Sektion | Kernaussage | Anforderung | Nr. |
|---|---|---|---|
| **Gestaltungsvorstellungen** | Aktuell fehlt es an qualifiziertem Personal für die Einführung von neuen Technologien und Methoden. | Möglichst selbsterklärende, **intuitive Gestaltung des Implementierungsprozesses und der Nutzung** von neuen Systemen. | 14 |
| | Das Suchen von relevantem Wissen in den aktuell sehr heterogenen Datenspeichern kostet viel Zeit. | Realisierung einer Möglichkeit zur **schnellen Suche und Identifikation von relevantem Wissen**. | 15 |
| | Statt einer händischen wäre eine automatisierte Dokumentation wünschenswert. | Realisierung von Möglichkeiten zur **automatisierten Dokumentation von Prozessabläufen und Best Practices**. | 16 |
| | Schritt-für-Schritt-Anleitungen würden den Onboarding-Prozess signifikant verbessern. | Realisierung von Möglichkeiten zur **effizienten Erstellung von Schritt-für-Schritt-Anleitungen**. | 17 |
| | Neue Systeme müssen von allen Mitarbeitenden nutzbar sein, unabhängig von ihren Vorerfahrungen. | Realisierung von **intuitiven Benutzerschnittstellen**. | 18 |
| | Das ideale Wissensmanagementsystem soll Abteilungsübergreifend funktionieren und dabei transparent sein. | Realisierung eines **offenen, erweiterbaren** und **transparenten** Wissensmanagementsystems. | 19 |

## C. Auszug aus der VDI 5610

Die 2009 erschienene Richtlinie VDI 5610 des Vereins Deutscher Ingenieure bietet Hilfestellungen zur Einführung, Umsetzung und Etablierung von Wissensmanagement im Ingenieurbereich. Dabei werden praxisnahe Herausforderungen benannt und entsprechende Vorgehensweisen zur Bewältigung vorgestellt. Die Richtlinie beschränkt sich auf Informationen, die für eine erfolgreiche Umsetzung von Wissensmanagement relevant sind und ist in die vier Kapitel „Anwendungsbereich“, „Grundlagen des Wissensmanagements“, „Wesentliche Elemente des Wissensmanagements in Ingenieurbereichen“ und „Einführung und Anwendung“ unterteilt. Die Ziele von Wissensmanagement sind im Kapitel „Grundlagen des Wissensmanagements“ aufgeführt und werden aus drei Perspektiven betrachtet [VDI09]:

**Perspektive des Unternehmens**

Aus Sicht des Unternehmens ergeben sich folgende Ziele und Nutzen von Wissensmanagement:

- Kosten einsparen durch:
  - Zeiteinsparung bei Suchaktionen, Recherchen in Daten und Quellen sowie für interne Abstimmungen
  - Vermeidung von Redundanzen und Doppelarbeiten durch erhöhte Transparenz und verbesserte Organisation
  - Effizientere Nutzung vorhandener Ressourcen
  - Reduzierung von Verlusten von erfolgskritischem Wissen
- Zusätzliche Erlöse erzielen durch:
  - Vermarktung von Wissen als Dienstleistung an Dritte
  - Bessere Positionierung der eigenen Produkte im Markt durch schnellere Berücksichtigung der Kundenanforderungen
- Produkt-/Qualitätsverbesserungen durch:
  - Transparente Aufbau- und Ablauforganisation
  - Transparente Prozesse
  - Standardisierung bezüglich Anforderungen an Inhalte und Gestaltung von Ergebnissen, Papier-Ablagen, Erfahrungsberichten
  - Standardisierung bezüglich IT-Lösungen durch Vereinheitlichung von Zugriffskonzepten, Vereinheitlichung von Suchtechnologien, Ablagen auf Laufwerken in einheitlicher Ordnerstruktur, Datei- und Ordner-Namenskonventionen
  - Systematische Aufbereitung von Erfahrungswissen
  - Zugriff auf tagesaktuelle aufgaben- und bereichsrelevante Informationen
  - Systematisierte direkte Kommunikation mit den Kunden

**Perspektive der Mitarbeitenden**

Aus Sicht der Mitarbeitenden ergeben sich folgende Ziele und Nutzen:

- Steigerung der Zufriedenheit: Höhere Motivation durch kürzere Such- und Zugriffszeiten auf Informationen, bessere Vernetzung
- Mitarbeiterqualifikationen „on the job“: Veränderungs-, Lern- und Entwicklungsbedarfe und -möglichkeiten der Mitarbeitenden werden mobilisiert:
    - Kompetenzerweiterung, Bereicherung der eigenen Arbeit durch Kenntnis und Zusammenarbeit über Teamgrenzen hinaus
    - Kreativitäts- und Innovationsfähigkeit über Impulse von außen: Dialog mit „Andersdenkenden“, Erweiterung der Perspektive
- Überwindung von Denkblockaden
- Identifikation von Wissensträgern
- Bessere Anpassung an ein sich rasch veränderndes Unternehmensumfeld
- Leichtere Einarbeitung in neue Aufgaben und Projekte

**Perspektive der Kunden**

Alle Ziele und Nutzen aus der Perspektive des Unternehmens und der Mitarbeitenden unterstützen die Prozessabläufe im Unternehmen und haben damit auch positive Effekte auf:

- Effiziente Abwicklung von Kundenaufträgen
- Gestaltung der kollaborativen Prozesse mit Externen
- Verbesserung der Prozess-, Produkt- und Servicequalität
- Erhöhung der Kundenzufriedenheit
- Sicherstellung der langfristigen Kundenbindung

# D. Anhang zu den Referenzimplementierungen

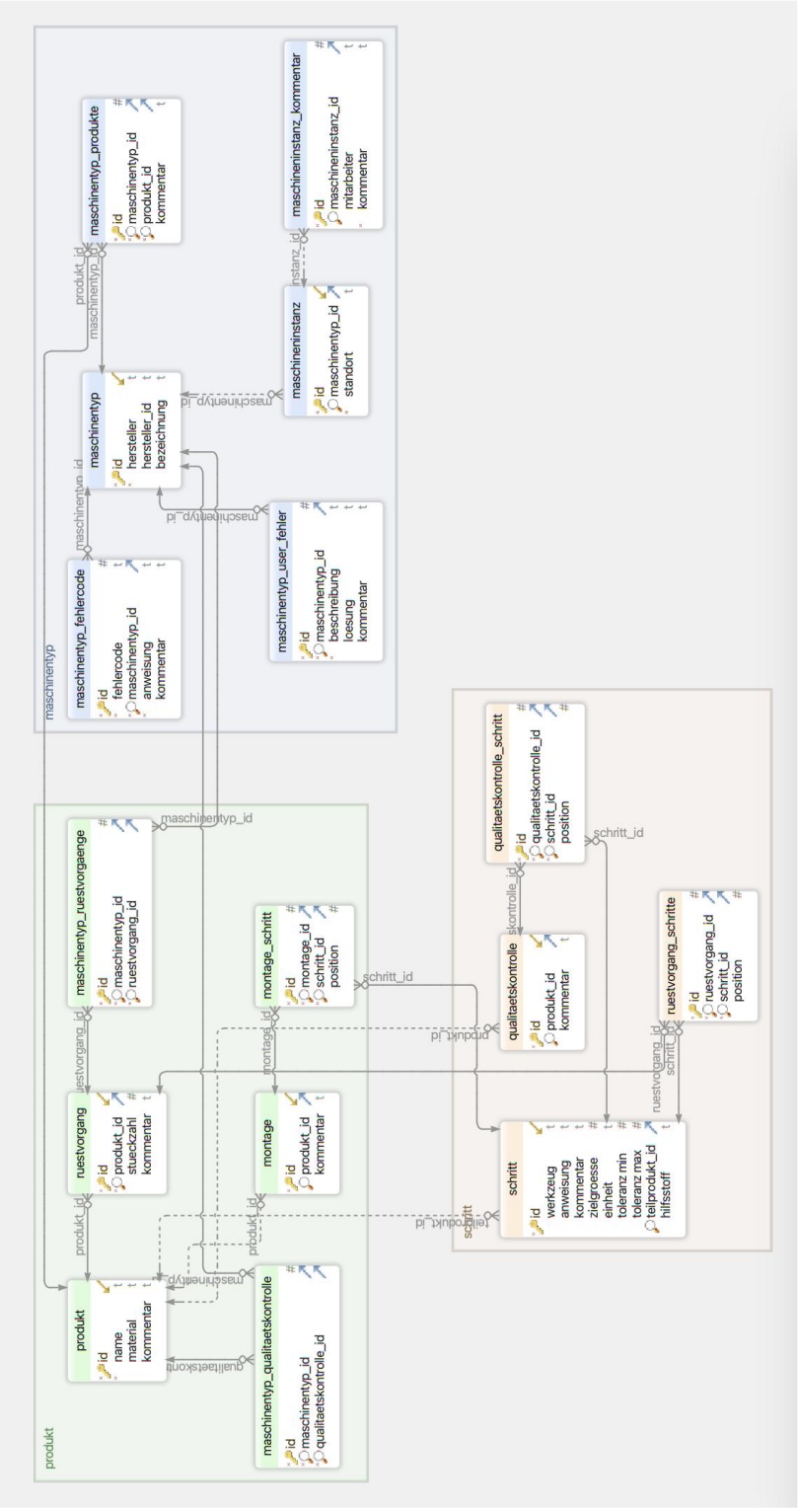



**Abbildung D.1: Für die Referenzimplementierung des aktiven Wissenstransfer (Rasa) genutzte Struktur der SQL-Datenbank**

*SQL Database Structure Used for the Reference Implementation of the active Knowledge Transfer (Rasa)*

# E. Anhang zur Validierung und Evaluation

**Tabelle E.7: Ergebnisse der manuellen, auf Schlüsselwörtern basierenden und semantischen Sequenzierung des Videos aus dem SAL**

*Results of Manual, Keyword-based and Semantic Sequencing of the Video from the SAL*

| | Manuell | | Schlüsselwörter | | Sentence Transformer | |
|---|---|---|---|---|---|---|
| Sequenz | Start | Ende | Start | Ende | Start | Ende |
| 1 | 00:00:00 | 00:00:24 | 00:00:00 | 00:00:24 | 00:00:00 | 00:03:10 |
| 2 | 00:00:24 | 00:00:58 | 00:00:24 | 00:01:00 | 00:03:10 | 00:03:25 |
| 3 | 00:00:58 | 00:03:13 | 00:01:00 | 00:03:25 | 00:03:25 | 00:03:57 |
| 4 | 00:03:13 | 00:03:33 | 00:03:25 | 00:03:58 | 00:03:57 | 00:04:03 |
| 5 | 00:03:33 | 00:03:50 | 00:03:58 | 00:04:34 | 00:04:03 | 00:05:01 |
| 6 | 00:03:50 | 00:04:20 | 00:04:34 | 00:05:01 | - | - |
| 7 | 00:04:20 | 00:04:50 | - | - | - | - |
| 8 | 00:04:50 | 00:05:01 | - | - | - | - |

**Tabelle E.8: Zuordnung der mittels Schlüsselwort automatisiert erstellten Sequenzen zu den manuell erstellten Sequenzen des SAL-Prozesses und Berechnung von Overlap, Union und IoU**

*Assignment of Sequences Automatically Generated by Keyword to Manually Generated Sequences of the SAL Process and Calculation of Overlap, Union and IoU*

| Manuell | Schlüsselwörter | Overlap | Union | IoU |
|---|---|---|---|---|
| Sequenz | Sequenz | [s] | [s] | [-] |
| 1 | 1 | 24 | 24 | 1 |
| 2 | 2 | 34 | 36 | 0,944 |
| 3 | 3 | 133 | 147 | 0,905 |
| 4 | 4 | 8 | 45 | 0,178 |
| 5 | 4 | 17 | 33 | 0,515 |
| 6 | 5 | 22 | 44 | 0,500 |
| 7 | 6 | 16 | 41 | 0,390 |
| 8 | 6 | 11 | 27 | 0,407 |

**Tabelle E.9: Zuordnung der mittels Sentence Transformer automatisiert erstellten Sequenzen zu den manuell erstellten Sequenzen des SAL-Prozesses und Berechnung von Overlap, Union und IoU**

*Assignment of Sequences Automatically Generated by Sentence Transformer to Manually Generated Sequences of the SAL Process and Calculation of Overlap, Union and IoU*

| Manuell | Transformer | Overlap | Union | IoU |
|---|---|---|---|---|
| Sequenz | Sequenz | [s] | [s] | [-] |
| 1 | 1 | 24 | 190 | 0,126 |
| 2 | 1 | 34 | 190 | 0,179 |
| 3 | 1 | 132 | 193 | 0,684 |
| 4 | 2 | 12 | 23 | 0,522 |
| 5 | 3 | 17 | 32 | 0,531 |
| 6 | 5 | 17 | 71 | 0,239 |
| 7 | 5 | 30 | 58 | 0,517 |
| 8 | 5 | 11 | 58 | 0,190 |

**Tabelle E.10: Ergebnisse der manuellen, auf Schlüsselwörtern basierenden und semantischen Sequenzierung des Videos aus dem KMU**

*Results of Manual, Keyword-based and Semantic Sequencing of the Video from the SME*

| | Manuell | | Schlüsselwörter | | Sentence Transformer | |
|---|---|---|---|---|---|---|
| Sequenz | Start | Ende | Start | Ende | Start | Ende |
| 1 | 00:00:00 | 00:01:37 | 00:00:00 | 00:01:48 | 00:00:00 | 00:00:56 |
| 2 | 00:01:37 | 00:02:25 | 00:01:48 | 00:02:43 | 00:00:56 | 00:01:01 |
| 3 | 00:02:25 | 00:02:58 | 00:02:43 | 00:03:09 | 00:01:01 | 00:01:07 |
| 4 | 00:02:58 | 00:03:43 | 00:03:09 | 00:05:04 | 00:01:07 | 00:01:18 |
| 5 | 00:03:43 | 00:04:57 | 00:05:04 | 00:07:27 | 00:01:18 | 00:03:24 |
| 6 | 00:04:57 | 00:07:10 | 00:07:27 | 00:11:31 | 00:03:24 | 00:03:34 |
| 7 | 00:07:10 | 00:08:10 | 00:11:31 | 00:11:47 | 00:03:34 | 00:03:51 |
| 8 | 00:08:10 | 00:09:45 | 00:11:47 | 00:13:06 | 00:03:51 | 00:04:12 |
| 9 | 00:09:45 | 00:10:28 | - | - | 00:04:12 | 00:04:17 |
| 10 | 00:10:28 | 00:11:20 | - | - | 00:04:17 | 00:05:04 |
| 11 | 00:11:20 | 00:11:45 | - | - | 00:05:04 | 00:07:27 |
| 12 | 00:11:45 | 00:13:06 | - | - | 00:07:27 | 00:08:58 |
| 13 | - | - | - | - | 00:08:58 | 00:10:09 |
| 14 | - | - | - | - | 00:10:09 | 00:10:46 |
| 15 | - | - | - | - | 00:10:46 | 00:13:01 |
| 16 | - | - | - | - | 00:13:01 | 00:13:06 |

**Tabelle E.11: Zuordnung der mittels Schlüsselwort automatisiert erstellten Sequenzen zu den manuell erstellten Sequenzen des KMU-Prozesses und Berechnung von Overlap, Union und IoU**

*Assignment of Sequences Automatically Generated by Keyword to Manually Generated Sequences of the SME Process and Calculation of Overlap, Union and IoU*

| Manuell | Schlüsselwörter | Overlap | Union | IoU |
|---|---|---|---|---|
| Sequenz | Sequenz | [s] | [s] | [-] |
| 1 | 1 | 97 | 108 | 0,898 |
| 2 | 5 | 37 | 66 | 0,561 |
| 3 | 5 | 15 | 44 | 0,341 |
| 4 | 6 | 34 | 126 | 0,270 |
| 5 | 10 | 74 | 115 | 0,643 |
| 6 | 11 | 126 | 150 | 0,840 |
| 7 | 12 | 43 | 261 | 0,165 |
| 8 | 12 | 95 | 244 | 0,389 |
| 9 | 14 | 43 | 244 | 0,176 |
| 10 | 14 | 52 | 244 | 0,213 |
| 11 | 15 | 14 | 27 | 0,519 |
| 12 | 15 | 79 | 81 | 0,975 |

**Tabelle E.12: Zuordnung der mittels Sentence Transformer automatisiert erstellten Sequenzen zu den manuell erstellten Sequenzen des KMU-Prozesses und Berechnung von Overlap, Union und IoU**

*Assignment of Sequences Automatically Generated by Sentence Transformer to Manually Generated Sequences of the SME Process and Calculation of Overlap, Union and IoU*

| Manuell | Transformer | Overlap | Union | IoU |
|---|---|---|---|---|
| Sequenz | Sequenz | [s] | [s] | [-] |
| 1 | 1 | 97 | 108 | 0,898 |
| 2 | 2 | 37 | 66 | 0,561 |
| 3 | 3 | 15 | 44 | 0,341 |
| 4 | 4 | 34 | 126 | 0,27 |
| 5 | 4 | 74 | 115 | 0,643 |
| 6 | 5 | 126 | 150 | 0,84 |
| 7 | 6 | 43 | 261 | 0,165 |
| 8 | 6 | 95 | 244 | 0,389 |
| 9 | 6 | 43 | 244 | 0,176 |
| 10 | 6 | 52 | 244 | 0,213 |
| 11 | 7 | 14 | 27 | 0,519 |
| 12 | 8 | 79 | 81 | 0,975 |