DLR Projektträger

Projektträger für das
Bundesministerium
für Wirtschaft
und Energie

INDUSTRIELLE GEMEINSCHAFTSFORSCHUNG

# Schlussbericht vom 31.01.2026

zum IGF-Vorhaben **22648 N**

## Thema

Flexibilisierung von Roboteroperationen in der Montage durch eine Ontologie- und KI-basierte Methodik mit integriertem Expertenwissen (ROOKIE)

## Berichtszeitraum

01.02.2023 bis 31.07.2025

## Forschungsvereinigung

Forschungsvereinigung Programmiersprachen für Fertigungseinrichtungen (FVP) e.V.
Theaterstraße 12
52062 Aachen

## Forschungseinrichtung(en)

Forschungsvereinigung Programmiersprachen für Fertigungseinrichtungen (FVP) e.V.
Theaterstraße 12
52062 Aachen

Werkzeugmaschinenlabor (WZL) der RWTH Aachen
Lehrstuhl für Werkzeugmaschinen
Steinbachstraße 19
52074 Aachen

[[FVP]]

INDUSTRIELLE GEMEINSCHAFTSFORSCHUNG

# Inhaltsverzeichnis


**1 Einleitung und Motivation ..... 3**

**2 Durchgeführte Arbeiten und Ergebnisse ..... 8**

2.1 Arbeitspaket 1: Bedarfs- und Anforderungsanalyse ..... 8

2.2 Arbeitspaket 2: Konzeption des Frameworks f. KI-basierte Roboteroperationen i. d. Montage.. 11

2.3 Arbeitspaket 3: Entwicklung der ontologiebasierten Systemmodellierung ..... 13

2.4 Arbeitspaket 4: Aufbau der virtuellen Module zur synthetischen Datenerzeugung ..... 15

2.5 Arbeitspaket 5: Benutzerschnittstellen zur Integration des Expertenwissens ..... 19

2.6 Arbeitspaket 6: Entwicklung des KI-Moduls „Lokalisieren“ ..... 22

2.7 Arbeitspaket 7: Entwicklung des KI-Moduls „Prüfen“ ..... 23

2.8 Arbeitspaket 8: Entwicklung des KI-Moduls „Greifen“ ..... 25

2.9 Arbeitspaket 9: Entwicklung des KI-Moduls „Fügen“ ..... 26

2.10 Arbeitspaket 10: Aufbau des Demonstrators ..... 28

2.11 Arbeitspaket 11: Validierung und Evaluierung ..... 29

2.12 Arbeitspaket 12: Projektmanagement, Dokumentation und Dissemination ..... 33

**3 Verwendung der Zuwendung ..... 34**

**4 Notwendigkeit und Angemessenheit der geleisteten Arbeit ..... 35**

**5 Darstellung des wissenschaftlich-technischen und wirtschaftlichen Nutzens der erzielten Ergebnisse insbesondere für KMU sowie ihres innovativen Beitrags und ihrer industriellen Anwendungsmöglichkeiten ..... 36**

**6 Wissenstransfer in die Wirtschaft ..... 37**

**7 Plan zum Ergebnistransfer in die Wirtschaft ..... 39**

7.1 Durchgeführte Transfermaßnahmen (vom Projektstart bis zum Projektende) ..... 39

7.2 Geplante Transfermaßnahmen (nach Projektende) ..... 41

# 1 Einleitung und Motivation

Entgegen den Argumenten für eine Automatisierung von Fertigungsprozessen, wie beispielsweise steigende Lohnkosten und Qualitätsanforderungen, wird in der Praxis in kleinen und mittleren Unternehmen (KMU) vor allem in der Montage meist auf manuelle Prozesse zurückgegriffen. Dies liegt vorrangig an einer häufig hohen Variantenvielfalt, kurzen Produktlebenszyklen sowie geringen Losgrößen. Diese Randbedingungen wirken sich negativ auf eine wirtschaftliche Umsetzung einer Automatisierung aus.

Eine Technologiekombination, welche das Potential hat, dieser Problematik entgegenzuwirken, ist die KI-basierte Robotik. Roboterbasierte Applikationen können dadurch deutlich flexibler, robuster und autonomer implementiert werden. Somit lässt sich eine Automatisierung auch bei geringerer Stückzahl rechtfertigen und die Gefahr von Fehlinvestitionen reduzieren. Um diese Potentiale für KMUs zugänglich zu machen, müssen allerdings noch einige Einstiegshürden überwunden werden.

Eine Herausforderung ist etwa die erforderliche Datenqualität und -quantität, welche gerade bei geringen Stückzahlen häufig schwer zu erreichen ist. Hier sind beispielsweise der Einsatz von synthetischen Datensätzen, die Anreicherung kleinerer Datensätze durch bekannte Engineering-Informationen oder die Integration von Expertenwissen mögliche Lösungsansätze. Eine weitere Herausforderung ist die Planung und Inbetriebnahme eines solchen Systems. Dazu wird in der Regel neben dem Montageplaner ein interdisziplinäres Fachpersonal aus den Bereichen Data Science, KI und Robotik benötigt. Diese Kompetenz ist gerade bei KMU in der Regel nicht vor-handen, weshalb Methoden entwickelt werden müssen, welche die Planung und Einführung von KI-basierten Montageprozessen autonomer und intuitiver gestalten.

Im Rahmen des beantragten Forschungsvorhabens „ROOKIE“ wird daher eine über-greifende Methodik mit ontologiebasiertem Modellierungsansatz und synthetischer Datenerzeugung entwickelt, welche den Einsatz von KI-basierten Roboteroperationen in der industriellen Montage ermöglicht. Dies soll die Flexibilität, Autonomie und Robustheit der Planung, Inbetriebnahme und Durchführung von roboterbasierten Montageoperationen steigern und somit die KI-basierte Robotik in der Montage für KMU zugänglich machen.

In Abbildung 1 werden die Teilaspekte dieser Methodik aufgezeigt, dessen Entwicklung primär die ganzheitliche Betrachtung folgender Forschungsschwerpunkte erfordert:

A. Synthetische Datenerzeugung für KI-basierte Roboteroperationen

B. Ontologiebasiertes Framework für KI-basierte Robotersysteme

C. Machine Learning Verfahren in der roboterbasierten Montage

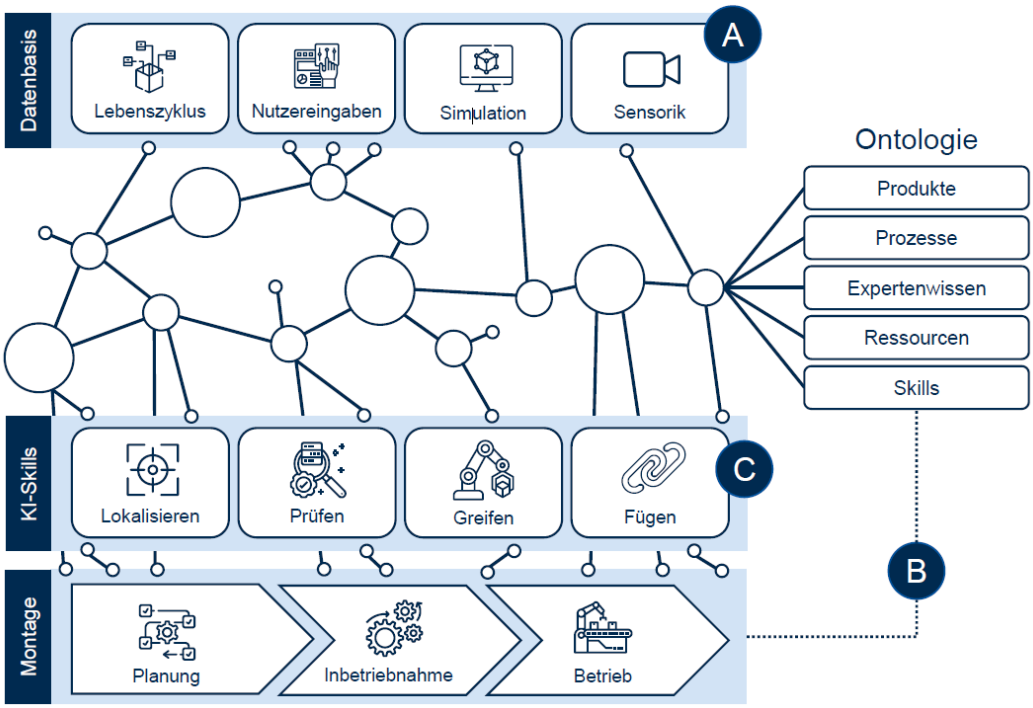


**Abbildung 1: KI-basierte Methodik für eine flexible roboterbasierte Montage**

Nachfolgend werden die Forschungsschwerpunkte näher erläutert.

## A. Synthetische Datenerzeugung für KI-basierte Roboteroperationen

Im Rahmen der umgesetzten Methodik werden vorrangig synthetische Daten für die Trainingsprozesse der KI-Module verwendet, bei dessen Erzeugung Expertenwissen integriert wird, um eine hohe Aussagekraft zu erreichen. Die zugrundeliegenden Argumente werden in diesem Kapitel erläutert. Dennoch sollen ergänzend heterogene Datentypen zum Einsatz kommen. Zunächst sind dabei Lebenszyklusdaten zu nennen, welche explizit in Form von Produkt-, Prozess- und Ressourceninformationen vorliegen. Dabei handelt es sich etwa um 3D-Modelle, die Montagereihenfolge oder Positionsdaten. Um die Industrietauglichkeit und die Aussagekraft dieser Daten für die Prozessplanung, das Training der KI-Module und die Durchführung der Montage zu erhöhen, sollen auch diese Daten durch das Fachwissen von Montage-Expert:innen angereichert, optimiert und kontrolliert werden.

Moderne Deep-Learning-Verfahren benötigen große Mengen an Trainingsdaten. Dies gilt vor allem für komplexe Aufgabenstellungen aus der Robotik oder dem maschinellen Sehen. Das Erstellen großer Datensätze ist allerdings ein hoher Kostenfaktor, da es meist in manuell durch Expert:innen durchgeführt wird. Dies stellt durch mangelnde Fachexpertise und fehlende Kapazität eine Hürde für KMUs dar und verhindert die Realisierung von Deep-Learning-Projekte.

Ein Ansatz mit großem Potential zur Überwindung dieses Problems sind sogenannte „synthetische Daten“, welche in einer Simulationsumgebung erzeugt werden. Die manuelle Arbeit beschränkt sich hierbei

hauptsächlich auf die einmalige Erstellung von 3D-Modellen und der virtuellen Umgebung, auf deren Basis beliebig große Datensätze autonom und günstig generiert werden können.

**Abbildung 2: Heterogene Datenbasis der in ROOKIE-gewählten Methodik**

Darüber hinaus bieten synthetische Daten noch weitere Vorteile. Roboter können flexibler eingesetzt werden, weil bei veränderten Prozessen oder Umgebungsparametern günstig neue Daten generiert werden können. Durch die Simulation ist zudem die Qualität der Datenbeschriftung höher und es lassen sich Daten für spezielle Aufgaben, welche in der Realität nicht ohne großen Aufwand möglich sind, günstig erzeugen. Beispielsweise die Erkennung der 6D-Pose von Werkstücken anhand von 2D-Bildern oder die Simulation von in der Realität zu vermeidenden Extremzuständen.

Im Bereich des Reinforcement Learning (RL) finden auch interaktive Simulationen Anwendung, in denen Robotersteuerungen trainiert werden. Reale Roboter können weiter für die Produktion eingesetzt werden, sodass durch das Trainieren kein Stillstand mehr entsteht. Auch wird das Training durch die parallele Simulation zahlreicher Roboter beschleunigt und skaliert somit stark mit der genutzten Rechenleistung.

Solche synthetischen Datensätze sollen daher die primäre Datenbasis für das Training der ROOKIE KI-Skills sein. Dazu werden einerseits virtuelle Umgebungen zum Rendering von visuellen 2D- und 3D-Datensätzen und andererseits physikalische Simulationen für das RL von Roboteroperationen aufgebaut. Aufgrund der Abweichung von Simulationen zur Realität lassen sich auf Basis von synthetischen Daten trainierte Algorithmen nicht immer störungsfrei in die Realität übertragen. Zur Entgegenwirkung dieser als „Sim2Real Gap“ bekannten Problematik existiert eine Vielzahl an Ansätzen, welche bei der Verwendung von synthetischen Daten zu beachten sind und die in ROOKIE zum Einsatz kommen sollen.

Abschließend kommen Sensordaten vor allem im Betrieb und der Evaluation zum Ein-satz. Beispielsweise werden dabei Kameradaten oder die Kraft-Momenten-Sensorik des sensitiven Roboters genutzt. So können

synthetisch trainierte KI-Module durch Sensordaten weiter optimiert werden oder einfach mit ihnen gespeist werden, um die Prozessdurchführung zu überwachen oder zu steuern.

### B. Ontologiebasiertes Framework für KI-basierte Robotersysteme

20% der Herstellungskosten und 50% der Produktionszeit entfallen auf die Montage, weshalb eine effektive Prozessplanung und -vorbereitung hohes Potential verspricht, um Kosten und Zeit einzusparen. Neben der Automatisierung der Montageoperationen durch KI-Skills soll daher auch die Prozessplanung und -vorbereitung unter-stützt werden. Durch die voll manuelle Durchführung dieser Arbeitsschritte steigt der Aufwand zur Inbetriebnahme neuer Montageprozesse. Ein informationstechnisches System zur teilautonomen Gestaltung der Prozessplanung, -vorbereitung und -durch-führung kommt damit einer prozesskritischen Rolle zu.

Als Grundlage für diese Funktionalität wird eine detaillierte Modellierung von Daten, Ressourcen, Zustände, Skills, Prozessparametern und Operationen, sowie die Modellierung der Beziehungen zwischen diesen Elementen benötigt. Montageprozesse würden bei einem solchen System nicht mehr als starre Abfolge von Handlungsoperationen definiert werden, sondern als Transition eines Ausgangszustands in einen Zielzustand. Die Beziehungen zwischen den Zustandsänderungen implizieren die notwendigen Montageoperationen, deren Anordnung und den sich daraus ergebenden Daten-, Ressourcen- und Fähigkeitsbedarf. Ontologien eignen sich dazu besonders, da sie es ermöglichen Wissen, Konzepte und Beziehungen zu modellieren. Insbesondere die W3C OWL 2 Web Ontology Language ist für die Entwicklung eines übergeordneten Modells zur Prozessplanung, -vorbereitung und -durchführung besonders gut geeignet. OWL 2 ist weitverbreitet, lässt eine flexible Erweiterung zu und ermöglicht es komplexes Wissen zu modellieren.

### C. Machine Learning Verfahren in der roboterbasierten Montage

Ein möglicher Ansatzpunkt zur Steigerung der Flexibilität und Robustheit von Montageoperationen ist der Einsatz von Machine Learning (ML) basierten Roboter-Skills, welche eine flexible oder automatisierte Durchführung von Montageoperationen er-möglichen. Diese werden nicht starr programmiert, sondern intelligent und kontext-adaptiv ausgeführt. Daher benötigen diese Methoden keine vorhergehende Programmierung von gut ausgebildeten Fachkräften, wodurch sich Kostenvorteile gegenüber der konventionellen Roboterprogrammierung ergeben. Die Skills werden in einer vo-rangehenden Trainingsphase erlernt und können etwa für das Lokalisieren von Objekten, die Prüfung des Montagefortschritts, das Greifen von unbekannten Objekten, und das sensitive Montieren von Bauteilen genutzt werden.

Die automatisierte Lokalisierung der Objekte und das anschließende KI-basierte Greifen erlaubt es, unbekannte Objekte auch in undefinierten Lagen zu lokalisieren und diese im Anschluss zu greifen. Dadurch benötigen solche Systeme weder Positioniersysteme, deren Anschaffung bei kleinen Losgrößen häufig unwirtschaftlich sind, noch bedarf es einer starren Programmierung der Greifpose, was die Flexibilität weiter erhöht. Darüber hinaus erlauben KI-Anwendung auch eine Verbesserung der Qualität. So liegt laut, trotz

Schulungen für Mitarbeiter:innen für die Qualitätskontrolle, die Gesamteffizienz der Qualitätsprüfung bei weniger als 80%. (ca. 20% - 30% der Fehler werden nicht gefunden) beschreibt die vollständige Abhängigkeit vom Menschen als den größten Nachteil der traditionellen Qualitätskontrolle. Neue KI-basierte An-sätze können die Werker:innen bei der Kontrolle sensorisch unterstützen. Eine weitere Verbesserung der Qualität und der Flexibilität kann beim KI-basierten, sensitiven Fügen erreicht werden. RL-basierte Ansätze ermöglichen es Montageoperationen in der Simulation zu erlernen, sodass kein starres programmieren der Roboter nötig ist. Zudem erlaubt der Einsatz von Kraftsensoren präzise Fügeoperationen mit geringen Toleranzen auszuführen.

Vor diesem Hintergrund sollen in ROOKIE diverse ML-Verfahren zu einer holistischen Methodik kombiniert werden, welche die Flexibilität, Autonomie und Robustheit der Planung, Inbetriebnahme und Durchführung von roboterbasierten Montageoperationen steigert. Aus dieser Motivation lässt sich folgende Hauptforschungs-frage ableiten:

*Inwiefern kann die Kombination aktueller Machine Learning Technologien mit einem ontologiebasierten Modellierungsansatz und einer synthetischen Datenbasis die Flexibilität und Robustheit von roboterbasierten Montageprozessen steigern?*

Zur Beantwortung dieser Forschungsfrage werden die im Folgenden beschriebenen Teilziele bearbeitet. Zunächst wird im Zuge der Planungsphase von ROOKIE ein über-greifendes Systemkonzept für die Methodik entwickelt. Dazu werden aus bestehenden Automationshürden und Bedarfen Anforderungen abgeleitet, welche die angestrebte Methodik erfüllen muss. Diese Anforderungen werden in Teilkonzepte der einzelnen Systembestandteile und das übergreifende Gesamtkonzept überführt. Im nächsten Schritt wird die Ontologie zur Modellierung des Systems entwickelt, welche das zu montierende Produkt, den Prozess, dessen Ressourcen und Skills sowie die Trai-ningsdaten abbilden kann. Zur Erzeugung der synthetischen Trainingsdaten werden virtuelle Simulationsmodule aufgebaut und die Daten in die Modellierung eingebettet. Anschließend wird die Benutzerschnittstelle konzipiert und umgesetzt, welche die Integration von Expertenwissen erlaubt. Im Zentrum der Methodik stehen die vier ML-Module „Lokalisieren“, „Prüfen“, „Greifen“ und „Fügen“ welche implementiert werden und mit der Kombination aus synthetischen Daten und integriertem Expertenwissen antrainiert werden. Abschließend wird die Methodik in einen am WZL bestehenden Demonstrator integriert und in industrienahen UseCases validiert und evaluiert.

# 2 Durchgeführte Arbeiten und Ergebnisse

Um die Forschungsfrage zu beantworten, wurden zwölf Arbeitspakete (AP) zu definiert und bearbeitet.

- AP1: Bedarfs- und Anforderungsanalyse
- AP2: Konzeption des Frameworks für KI-basierte Roboteroperationen in der Montage
- AP3: Entwicklung der ontologie-basierten Systemmodellierung
- AP4: Aufbau der virtuellen Module zur synthetischen Datenerzeugung
- AP5: Benutzerschnittstellen zur Integration des Expertenwissens
- AP6: Entwicklung des KI-Moduls „Lokalisieren“
- AP7: Entwicklung des KI-Moduls „Prüfen“
- AP8: Entwicklung des KI-Moduls „Greifen“
- AP9: Entwicklung des KI-Moduls „Fügen“
- AP10: Aufbau des Demonstrators
- AP11: Validierung und Evaluation
- AP12: Projektmanagement, Dokumentation und Dissemination

Nachfolgend werden Inhalt, Vorgehensweise und erzielte Ergebnisse für jedes Arbeitspaket dargelegt.

## 2.1 Arbeitspaket 1: Bedarfs- und Anforderungsanalyse

Basierend auf bestehenden Automatisierungshürden realer Montageprozesse galt es in AP 1 die Bedarfe und Anforderungen an die zu entwickelnde Methodik zu ermitteln. Dazu wurden aktuelle, überwiegend manuell durchgeführte Prozesse betrachtet und dabei auftretende Barrieren identifiziert, zu archetypischen Hürden geclustert und dessen kausale Zusammenhänge analysiert. Hier konnte auf Anwendungsbeispiele aus dem Projektkonsortium, welche in Workshops analysiert werden, sowie aus Industrie- und Forschungsprojekten des WZL zurückgegriffen werden. Aus den ermittelten Hürden wurde eine Anforderungsliste an die zu entwickelnde Methodik abgeleitet. Diese wurde hierarchisch in globale Anforderungen und Anforderungen an die Teilmodule gegliedert und um eine Bewertungsmetrik für die Evaluation in AP11 ergänzt.

**Abbildung 3: Zusammenspiel zwischen Produktion und Industrial Metaverse**

Das methodische Vorgehen in diesem Arbeitspaket zielte darauf ab, eine wissenschaftlich fundierte Basis für die darauffolgenden Entwicklungsphasen zu schaffen. Abbildung 3 zeigt das Zusammenspiel von Produktionsumgebung und dem Industrial Metaverse.

Die Schritte und die daraus resultierenden Ergebnisse stellen sich wie folgt dar:

- **Quantitative Prozess- und Kostenanalyse:** Initial wurde eine detaillierte Untersuchung industrieller Bildverarbeitungsprojekte durchgeführt. Dabei wurde als kritisches Ergebnis verifiziert, dass die Akquise und manuelle Annotation von Trainingsdaten über 80 % der Gesamtkosten verursacht. Pro Bild wurde ein Zeitaufwand von mehreren Minuten für das manuelle Labeling identifiziert, was die Skalierbarkeit herkömmlicher Deep-Learning-Ansätze in der Kleinserienfertigung verhindert.

- **Identifikation technischer Barrieren:** Methodisch wurden bestehende Toolchains zur synthetischen Datengenerierung auf ihre Autonomiegrade untersucht. Als Ergebnis wurden der „Programmierer-Flaschenhals“ (Notwendigkeit eines Programmierers für jeden neuen Datensatz) und die „Domain Gap“ (beschreibt in der künstlichen Intelligenz die Leistungsabnahme eines Modells, wenn es auf Daten angewendet wird, die sich in ihrer Verteilung oder Struktur von den Trainingsdaten unterscheiden) als technologische Hürden identifiziert.

- **Definition und technische Spezifikation der Use-Cases:** Aus der Analyse wurden drei primäre Anwendungsszenarien zur Unterstützung der Montage durch Cobots abgeleitet:
  - Kognitive Montage einer Getriebegruppe: Das Handling und die Montage einer Getriebegruppe als Referenz Use Case.
  - Oberflächenprüfung: Definition von Defekttypen (Kratzer) auf metallischen Oberflächen. Als technisches Ergebnis wurde festgelegt, die Defekte sowohl über Shader-Maps (Normal/Roughness) als auch über geometrische Mesh-Modifikationen (<0,01 mm Extrusion) abzubilden. Hierzu soll dem Cobot über ein intuitives Interface niederschwellig neue Checkpoints eingeteacht werden können.

- **Ableitung der hierarchischen Anforderungsliste:** Zu der gegliederten Anforderungsliste gehören:
  - Globale Anforderungen: Entkopplung von Programmieraufwand und Wissenseingabe zur Reduktion der benötigten Expertenrollen. Sicherstellung der Sim2Real-Robustheit durch Domain Randomization und Physically Based Rendering (PBR).

- Funktionale Modulanforderungen: Notwendigkeit einer webbasierten Nutzerschnittstelle für asynchrone Wissenserhebung, automatisierte Ground-Truth-Extraktion (Bounding Boxes, Keypoints, Segmentierungsmasken) sowie Unterstützung von Curriculum Learning zur Bewältigung komplexer Fügepfade.

Zur Optimierung und Flexibilisierung der Produktion kann das Industrial Metaverse dazu beitragen, indem es eine (oder mehrere) der folgenden Funktionen ausführt:

1. Erzeugung von fotorealistischen Renderings, welche später in KI-Trainingsprozessen genutzt werden
2. Physikalische Simulationen um bspw. KI-Anwendungen wie verstärkendes Lernen (englisch: Reinforcement Learning) zu ermöglichen
3. Konzeptionierung und Simulation von kollaborativen Workflows
4. Nutzung als digitaler Zwilling

Abbildung 4 zeigt exemplarische Anwendungsbereiche des Industrial Metaverse im Produktionsumfeld.

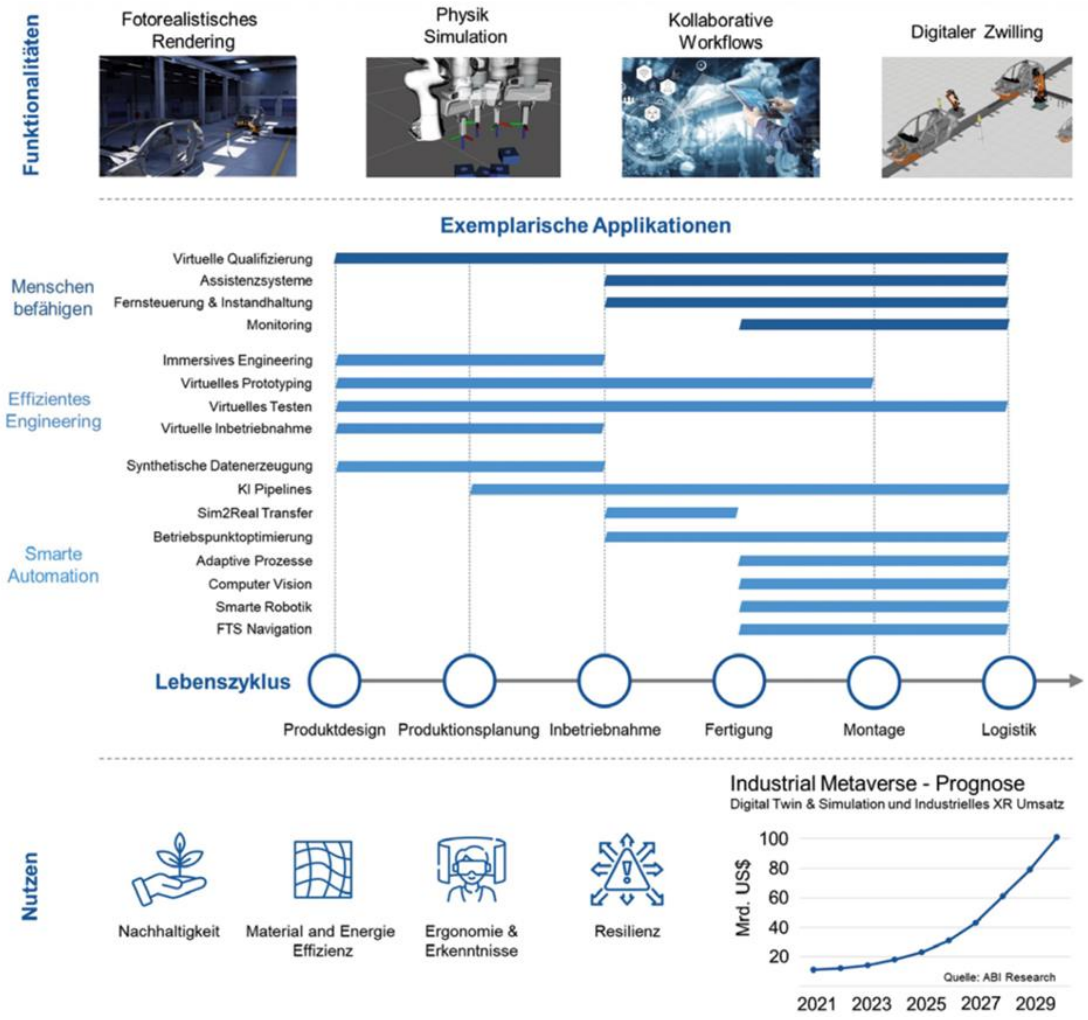


**Abbildung 4: Anwendungsbereiche des Industrial Metaverse**

Das Ergebnis von Arbeitspaket 1 bildet ein Katalog an bestehenden Automationshürden in der industriellen Montage sowie eine vollständige Anforderungsliste an die zu entwickelnde Methodik und den zugehörigen Demonstrator inklusive Bewertungsmetrik zur abschließenden Evaluation.

## 2.2 Arbeitspaket 2: Konzeption des Frameworks für KI-basierte Roboteroperationen in der Montage

Auf Basis der in AP1 erarbeiteten Anforderungsliste wurde in diesem AP ein übergreifendes Systemkonzept für die Methodik hergeleitet, welches in den folgenden AP in Form von detaillierten Teilkonzepten spezifiziert und erprobt werden kann. Dabei mussten vorrangig folgende Teilmodule des Gesamtsystems betrachtet werden: Die übergreifende Systemmodellierung (AP3), die synthetische Datenerzeugung (AP4), die Be-nutzerschnittstellen zur Integration von Expertenwissen (AP5), die KI-Module für die roboterbasiert Montage (AP6-AP9) und die physische Umsetzung in Form des Demonstrator (AP10). Diese Teilkonzepte galt es zu dem übergreifenden Systemkonzept zu kombinieren.

Die Konzeption des Frameworks für KI-basierte Roboteroperationen in der Montage basiert auf einem modularen Sim2Real-Ansatz, der darauf abzielt, die physische Produktion und eine hochreale Simulationsumgebung in einer gekoppelten Toolchain zu vereinen.

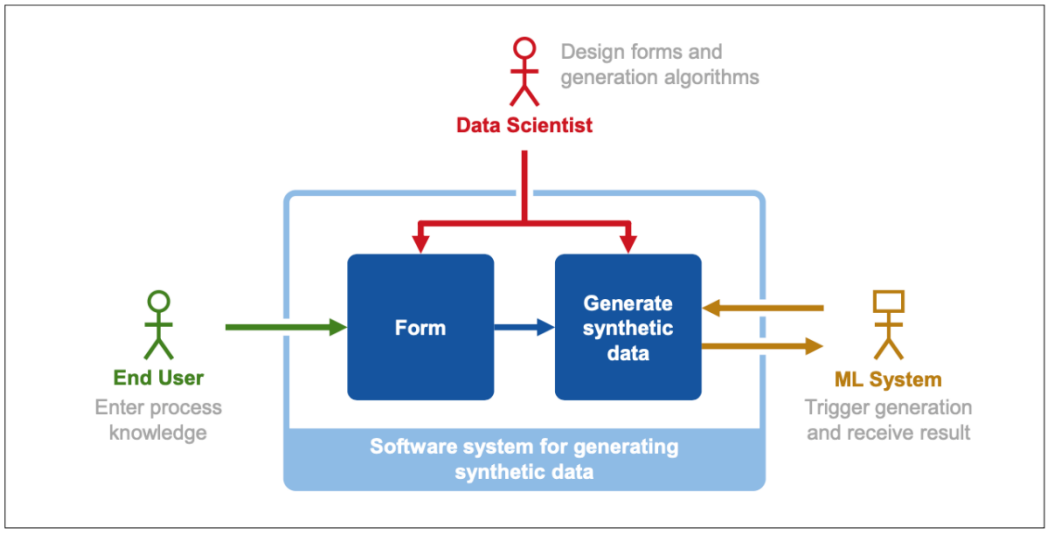


**Abbildung 5: Entkopplung der Experten für die synthetische Datenerzeugung**

Der gesamte Workflow der Datengenerierung ist in asynchrone Phasen unterteilt, um den „Programmierer-Flaschenhals“ zu eliminieren und die Autonomie des Gesamtsystems zu steigern, siehe Abbildung 5. In der Konfigurationsphase werden durch Data Scientists parametrisierte Skripte und die logischen Strukturen der Nutzerschnittstellen bereitgestellt. Die Wissenserhebungsphase ermöglicht es Domänenexperten, prozess-relevantes Wissen über intuitive Web-Formulare direkt in das System einzupflegen. Schließlich erlaubt die Datengenerierungsphase dem ML-System, die Erzeugung neuer Trainingsdatensätze on-demand über die API

auszulösen, wobei lediglich neue Parameter wie CAD-Modelle übergeben werden müssen. Dabei werden die Rollen des Data Scientists, des Endanwenders und des ML-Systems technologisch so entkoppelt, dass Aktualisierungen am System ohne tiefgreifende IT-Kenntnisse durchgeführt werden können.

Kernstück der Softwarearchitektur ist ein dreischichtiger Aufbau, der eine strikte Trennung zwischen der logischen Steuerung und der technischen Datengenerierung ermöglicht: Der Autonomy Layer verwaltet die asynchronen Prozessphasen und stellt die REST-API zur Verfügung, während der Ontology Layer als Abstraktionsschicht fungiert, um renderer-agnostische Szenenbeschreibungen zu definieren. Die unterste Ebene, der Rendering Layer, dient als spezifische Schnittstelle zur Ausführung der Simulation in Engines wie Blender oder NVIDIA Isaac Sim, siehe Abbildung 6.

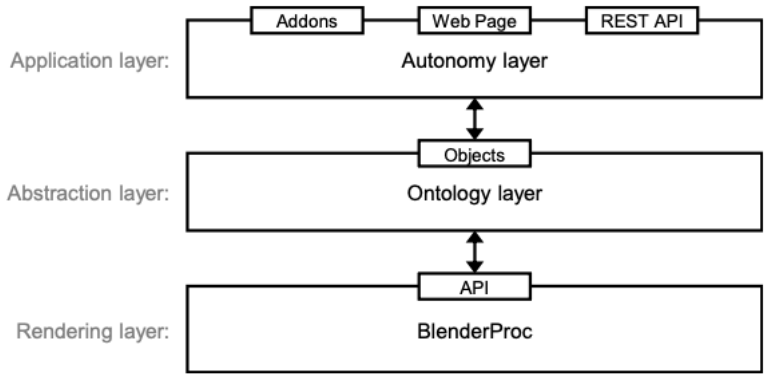


**Abbildung 6: Drei-Schichten Architektur zur Modularisierung der synthetischen Datenerzeugung**

Die funktionale Toolchain gliedert sich in drei spezialisierte Module: das Vision-Modul, das RL-Modul und das Control-Modul, siehe Abbildung 7. Das Vision-Modul realisiert die 6D-Posenschätzung der Bauteile durch eine Kombination aus YOLO-basierter Keypoint-Detektion und Perspective-n-Point (PnP) Algorithmen, wodurch eine Lokalisierungsgenauigkeit von unter einem Millimeter im Arbeitsraum erreicht wird. Das RL-Modul nutzt massiv parallelisierte, GPU-beschleunigte Simulationsumgebungen (Vectorized Environments), um komplexe Füge-Skills durch Reinforcement Learning zu trainieren. Hierbei werden physikalische Parameter wie Reibungskoeffizienten, Dämpfungskonstanten und viskoelastisches Materialverhalten präzise abgebildet, um eine hohe Transferierbarkeit der gelernten Policys zu gewährleisten. Das Control-Modul fungiert als zentrale Orchestrierungseinheit und steuert die Montageabläufe über ein „Decider-Network“, das als gerichteter azyklischer Graph (DAG) implementiert ist. Dieses Netzwerk ermöglicht eine fähigkeitsbasierte Auswahl zwischen KI-basierten Skills und klassischen Roboteroperationen. Die Verbindung zur realen Hardware wird über eine bidirektionale ROS-Schnittstelle hergestellt, die eine kontinuierliche Synchronisation zwischen dem Digitalen Zwilling und der Robotersteuerung ermöglicht. Um die Prozessresilienz bei kontaktintensiven Fügevorgängen zu erhöhen, integriert das Framework zudem eine kartesische Impedanzregelung zum Ausgleich mechanischer Spannungen sowie Strategien des Curriculum Learning, bei denen die geometrischen Toleranzen während des Trainings sukzessive reduziert werden.

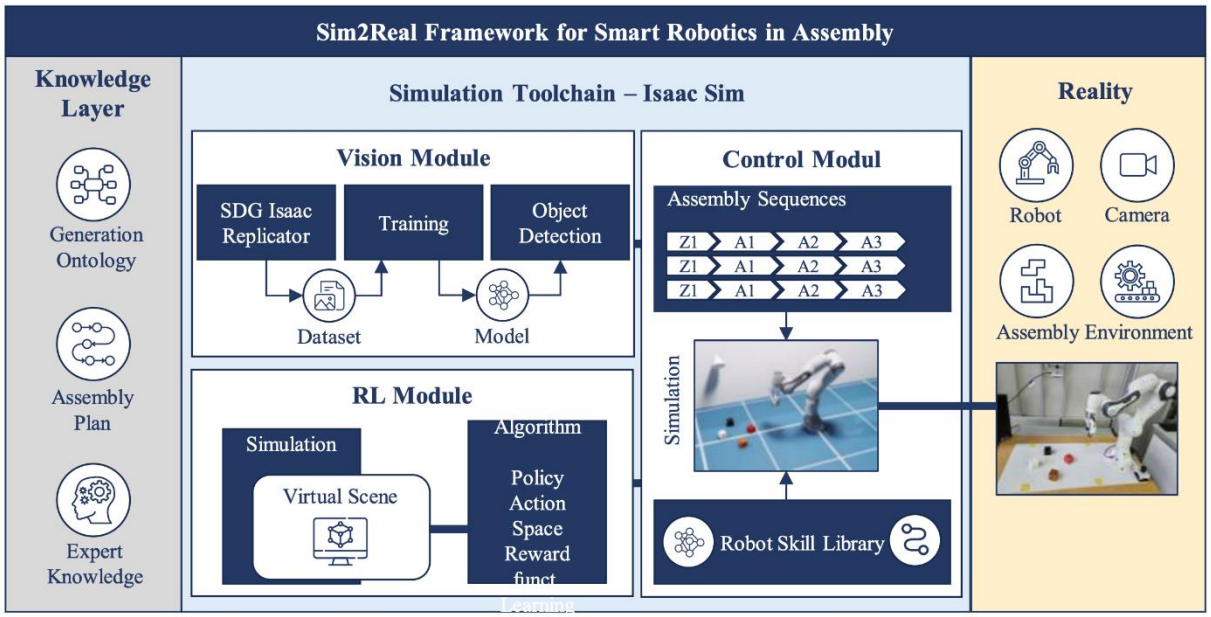


**Abbildung 7: Framework zur Flexibilisierung der kollaborativen Montage durch KI-Skills und einer Simulationsumgebung**

Das Ergebnis von Arbeitspaket 2 bildet ein umfassendes Konzept der Methodik als Grundlage für die darauffolgende Entwicklungsphase.

## 2.3 Arbeitspaket 3: Entwicklung der ontologiebasierten Systemmodellierung

Ein ontologiebasierter Modellierungsansatz sollte die Grundlage der entwickelten Methodik darstellen. Durch die semantische Beschreibung von Bauteilen, Montageabfolgen, Prozessressourcen und den Skills des Systems, lässt sich die Prozessplanung und -inbetriebnahme teilautomatisieren. Außerdem dient die Ontologie der logischen Verknüpfung von Expertenwissen und Trainingsdaten sowie der Beschreibung von Systemzuständen. Zum Aufbau der Ontologie wird die die Spezifikation Web Ontology Language (OWL) verwendet.

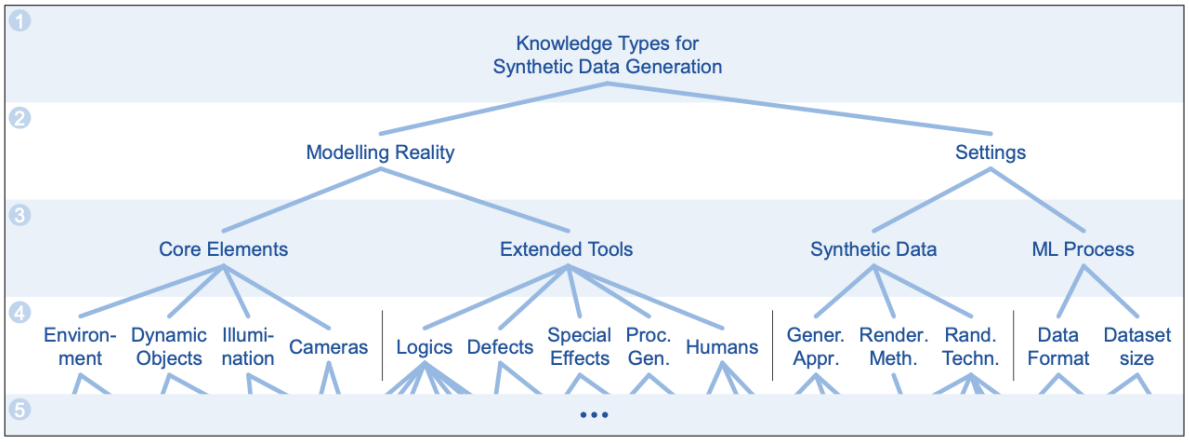


**Abbildung 8: Ontologie-basierte Systemmodellierung für die Generierung synthetischer Daten**

Das Vorgehen in Arbeitspaket 3 konzentrierte sich auf die systematische Formalisierung von Produktions- und Simulationswissen in einer maschinenlesbaren Struktur, um eine wissensbasierte Steuerung der Datengenerierung zu ermöglichen. Methodisch wurden zunächst relevante Wissenstypen aus wissenschaftlichen Publikationen, industriellen Computer-Vision-Projekten und bestehenden Rendering-Systemen gesammelt und konsolidiert. Durch ein hierarchisches Clustering dieser Informationen wurde eine umfassende Taxonomie mit 298 Knoten über neun Ebenen entwickelt, die als konzeptionelle Grundlage für die Ontologie dient. Wie in Abbildung 8 dargestellt, gliedert sich diese Struktur in zwei Hauptzweige: „Modelling Reality“, welcher die physische Szene (Kernkomponenten wie Objekte und Beleuchtung sowie erweiterte Werkzeuge für Defekte oder Logiken) beschreibt, und „Settings“, der technische Parameter für den ML-Prozess und das Rendering umfasst. Die gesamte Taxonomie ist einsehbar unter: https://zenodo.org/records/7270630. Die Modellierung zeigt die softwaretechnische Trennung zwischen „Properties“ und „Selection Methods“. Dies erlaubt es, komplexe Rendering-Attribute über benutzerfreundliche Schnittstellen wie vordefinierte Materialklassen (z. B. „Holz“, „Stahl“) anzusteuern, wobei die spezifischen PBR-Werte im Hintergrund automatisch gesetzt werden (siehe Abbildung 9).

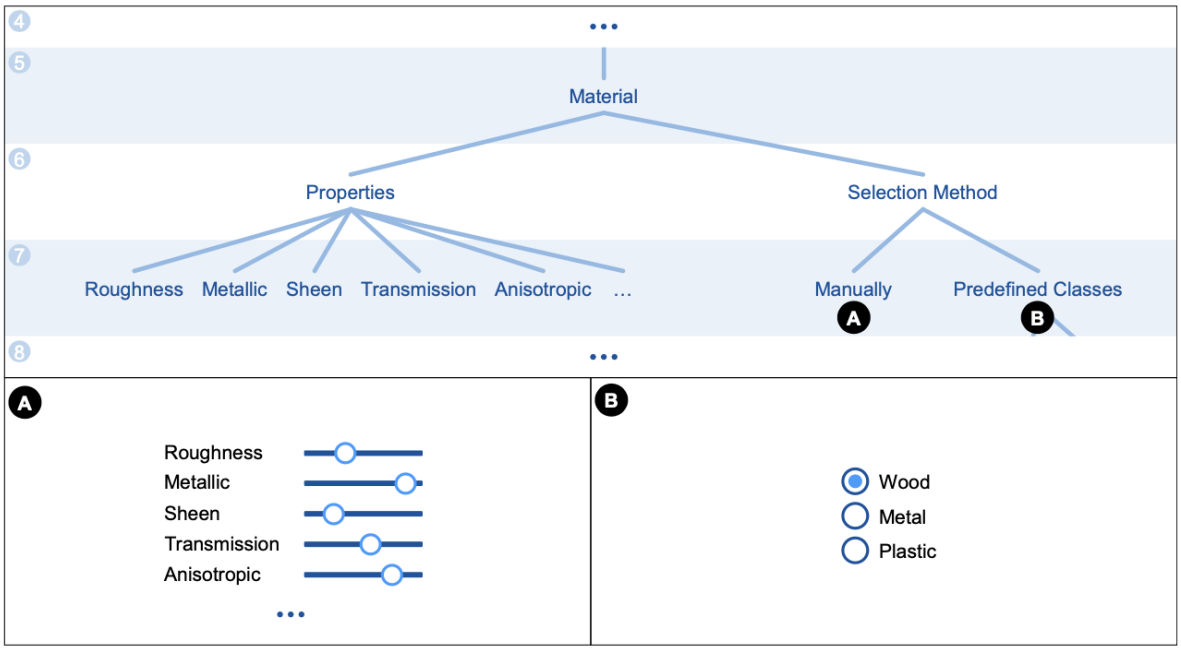


**Abbildung 9: Oben: Ausschnitt der Taxonomie. Unten: Beispielhaftes Interface für die beiden „Selection Methods“. Wenn eine Kategorie in (B) ausgewählt wird, werden die Entsprechenden Eigenschaftsausprägungen in (A) gesetzt.**

Zur Implementierung wurde eine „Ontology Layer“ als Abstraktionsebene im Framework realisiert, die Low-level Simulationsbefehle in allgemeine Klassen überführt. Innerhalb dieser Schicht fungiert die Klasse GenerationRoot als zentraler Startpunkt, an dem alle Szenenobjekte gebunden werden. Dabei wird Vererbung genutzt, um spezialisierte Strategien wie LookAtVolumeRotation für die automatisierte Objektplatzierung bereitzustellen. Die Definition von Belohnungsachsen auf Basis von Keypoints innerhalb der Ontologie

ermöglichen es, Positions- und Orientierungsfehler mathematisch in einem Belohnungswert für das Reinforcement Learning zusammenzufassen und dabei Symmetrieeigenschaften der Bauteile automatisch zu berücksichtigen. Das resultierende Modell bildet somit die semantische Brücke zwischen dem Expertenwissen und der physikalisch basierten Simulationsumgebung.

Das Ergebnis des Arbeitspakets bildet eine Ontologie zur semantischen Modellierung des Montageprozesses sowie sämtlicher Systembestandteile, deren Skills und Zuständen.

## 2.4 Arbeitspaket 4: Aufbau der virtuellen Module zur synthetischen Datenerzeugung

Die KI-Module in ROOKIE sollten vorrangig mit synthetischen Daten antrainiert werden. Zur Datenerzeugung mussten dementsprechend virtuelle Umgebungen aufgebaut werden. Im Zuge dieses AP werden zwei Module zur Datenerzeugung entwickelt. Ein Modul dient der Erzeugung von 2D- und 3D-Bilddaten aus den CAD-Modellen der Bauteile. Es wurden diverse Ansätze zum Rendern von fotorealistischen Bildern mit den definierten Anforderungen aus AP1 verglichen und ein ausgewähltes implementiert. Dabei werden zahlreiche Parameter, wie etwa Lichtverhältnisse, Verdeckungsgrad und Verzerrungen eingestellt, welche in definierten Bereichen randomisiert werden, um eine möglichst breite Datenbasis zu erzeugen. Damit antrainierte KI-Module (AP6, AP7 und AP8) können so robuster Er-kenntnisse auf Basis realer Kameradaten generieren.

Das zweite Modul ist eine physikalische Simulationsumgebung. Diese wird vorrangig für das RL von Roboteroperationen genutzt welches vor allem in AP9 beim Training der kraftbasierten Fügeprozesse und gegebenenfalls in AP8 zum Einsatz kommt. Dazu wird die von dem Roboter durchzuführende Operation simulativ auf Basis von Belohnungsfunktionen erlernt. Auch hier wird eine gewisse Streuung der Parameter bewusst eingeführt, um eine hohe Robustheit bei der Übertragung auf die Realität zu erreichen. Wichtig ist dabei auch, dass in der Simulation die gleiche Robotersteuerung wie in der Realität verwendet wird, daher wird das System ROS-basiert (Robot Ope-rating System) implementiert.

Im Arbeitspaket 4 erfolgte die technische Umsetzung der virtuellen Module zur automatisierten Generierung synthetischer Trainingsdaten, wobei der Fokus auf einer modularen und skalierbaren Softwarearchitektur lag. Die Pipeline zur synthetischen Datenerzeugung ist in Abbildung 10 dargestellt.

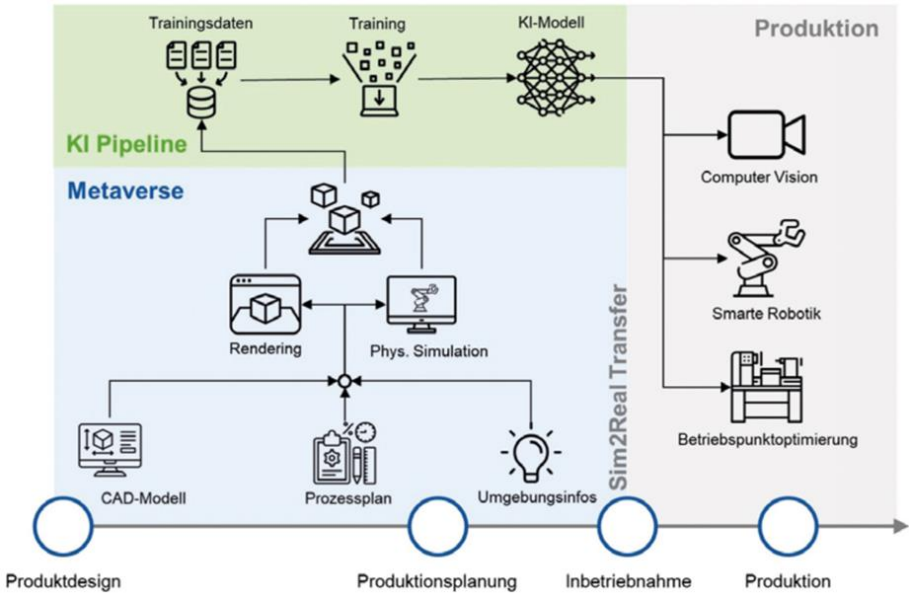


**Abbildung 10: Pipeline zur synthetischen Datenerzeugung**

Abbildung 11 zeigt den gesamten Aufbau der Software zur synthetischen Datenerzeugung.

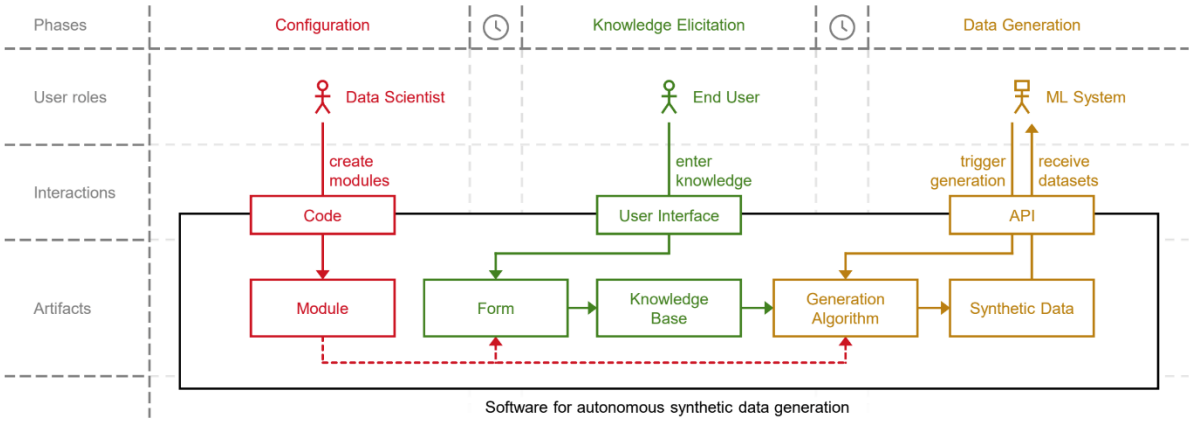


**Abbildung 11: Software zur synthetischen Datenerzeugung**

Ein zentrales technisches Ergebnis stellt die Implementierung der Mapping-Logik innerhalb der Funktion json_to_onto dar, welche es ermöglicht, domänenspezifisches Expertenwissen aus JSON-Konfigurationsdateien dynamisch in Instanzen der Ontologie-Klassen zu übersetzen. Dieser Prozess erlaubt die Kombination von flexiblen Nutzereingaben, wie etwa variierenden Arbeitsraumdimensionen, mit fest hinterlegten systemtechnischen Parametern, siehe Abbildung 12 für ein Beispiel.

```python
def json_to_onto(onto_classes, end_user_data, ml_system_data):
    volume = onto_classes.SimpleVolume(
        Has_XCoordinate = [-end_user_data["area_length_x"]/2],
        Has_YCoordinate = [-end_user_data["area_length_y"]/2],
        Has_ZCoordinate = [500.0],
        Has_XLength = [end_user_data["area_length_x"]],
        Has_YLength = [end_user_data["area_length_y"]]
    )
    # ...
    root = onto_classes.GenerationRoot(
        Has_NumberOfImagesToRender = [ml_system_data["count"]],
        Has_Volume = [volume],
        Has_Object = obj,
        Has_Camera = [camera],
        Has_Label = [label]
    )
```

a

b

**Abbildung 12: Beispiel json_to_function. Der Großteil des Wissens ist vom Data Scientist hard-gecode). Für die Tiefe und Breite des Volumens werden jedoch Werte verwendet, die ein Endnutzer über das Formular eingibt (siehe a), während für die Datensatzgröße ein vom ML-System übermittelter Parameter genutzt wird (siehe b). Ein Root-Objekt wird erstellt, um alle Komponenten miteinander zu verknüpfen.**

Zur Erzeugung der Daten wurden verschiedene Simulations-Backends integriert, um sowohl hohe Anforderungen an den visuellen Realismus als auch an die physikalische Rechengeschwindigkeit zu erfüllen. Während BlenderProc für physikalisch basiertes Path Tracing zur Erstellung fotorealistischer RGB-Bilddatensätze genutzt wurde, kam NVIDIA Isaac Replicator für die echtzeitfähige Generierung visueller und physikalischer Ground-Truth-Daten zum Einsatz. In Bezug auf die automatisierte Qualitätsprüfung wurden spezialisierte Module zur Modellierung von Oberflächenfehlern entwickelt, die Defekte wie Kratzer entweder über PBR-Shader-Maps (Normal, Roughness) oder durch geometrische Mesh-Modifikationen mit Extrusionen im Mikrometerbereich (<0,01 mm) abbilden. Der hierfür entwickelte Workflow zur Defektgenerierung ist in Abbildung 13 dargestellt.

Für die Simulation komplexer Montagevorgänge, wie etwa Schnappverbindungen (Snap-fits), wurde ein Modellierungsansatz gewählt, der elastisches Materialverhalten stabil in starren Physik-Engines abbildet. Bauteile mit einem Young-Modul von bspw. 2000 MPa werden hierbei als kinematische Ketten aus starren Teilkörpern modelliert, die über Gelenke mit spezifisch kalibrierten Steifigkeiten verbunden sind. Diese methodische Darstellung ermöglicht eine effiziente Berechnung der Verformungen unter Fügekräften. Um den Lernprozess für Reinforcement-Learning-Agenten zu beschleunigen, wurde auf das Isaac Gym Factory Framework zurückgegriffen, welches massiv parallelisierte Simulationen von bis zu 128 Instanzen gleichzeitig auf der GPU berechnet. Parallel dazu extrahieren die Module während des Renderings automatisch präzise Ground-Truth-Informationen, darunter 2D/3D-Bounding-Boxes, Segmentierungsmasken sowie die exakten Koordinaten von Keypoints für die 6D-Posenschätzung.

**Algorithm 1**: Rendering of synthetic images

```
set axis rotation and camera location
FOR each camera position and orientation
    determine randomly the number of scratches
    FOR each scratch
        choose randomly between two scratch models straight/curved
        determine randomly the rotation angle around the z axis
        scale randomly in x direction
    ENDFOR
    FOR each metal plate
        get dimension and position
        FOR each scratch
            generate scratch coordinates within plate boundaries
        ENDFOR
    ENDFOR
    update camera position and rotation
    randomize lighting intensity in interval of [int_min, int_max]
    render the scene and save the generated image
    FOR each scratch
        compute scratch bounding box coordinates in camera frame
            RETURN (x_min, y_min), (x_max, y_max)
    ENDFOR
    save coordinates and delete generated objects
ENDFOR
```

**Abbildung 13: Workflow zur Generierung synthetischer Daten für die Aufbringung von Defekten auf metallische Oberflächen**

Ergebnis: Zwei Module zur Erzeugung synthetischer Trainingsdaten. In einem werden realistische 2D- und 3D-Bilddaten gerendert und das andere bietet eine physikalische Simulationsumgebung zum Training von Roboteroperationen an.

## 2.5 Arbeitspaket 5: Benutzerschnittstellen zur Integration des Expertenwissens

Zur Gewährleistung einer hohen Industrietauglichkeit und Robustheit der trainierten Module sowie der erstellten Montagepläne soll eine Benutzerschnittstelle entwickelt werden, über die ein Montageexperte mit dem System interagieren kann. Dieser kann durch Integration seines Fachwissens die erzeugte Datenbasis anreichern, optimieren und kontrollieren. Über die Schnittstelle sollen etwa die Montageaufgaben eingepflegt, die Parameter zur Erzeugung der Trainingsdaten eingestellt und der durchzuführende Montageprozess kontrolliert und angepasst werden. Das dadurch aufgenommen Expertenwissen muss in das Systemmodell semantisch integriert und den entsprechen-den Modulen zur Verfügung gestellt werden.

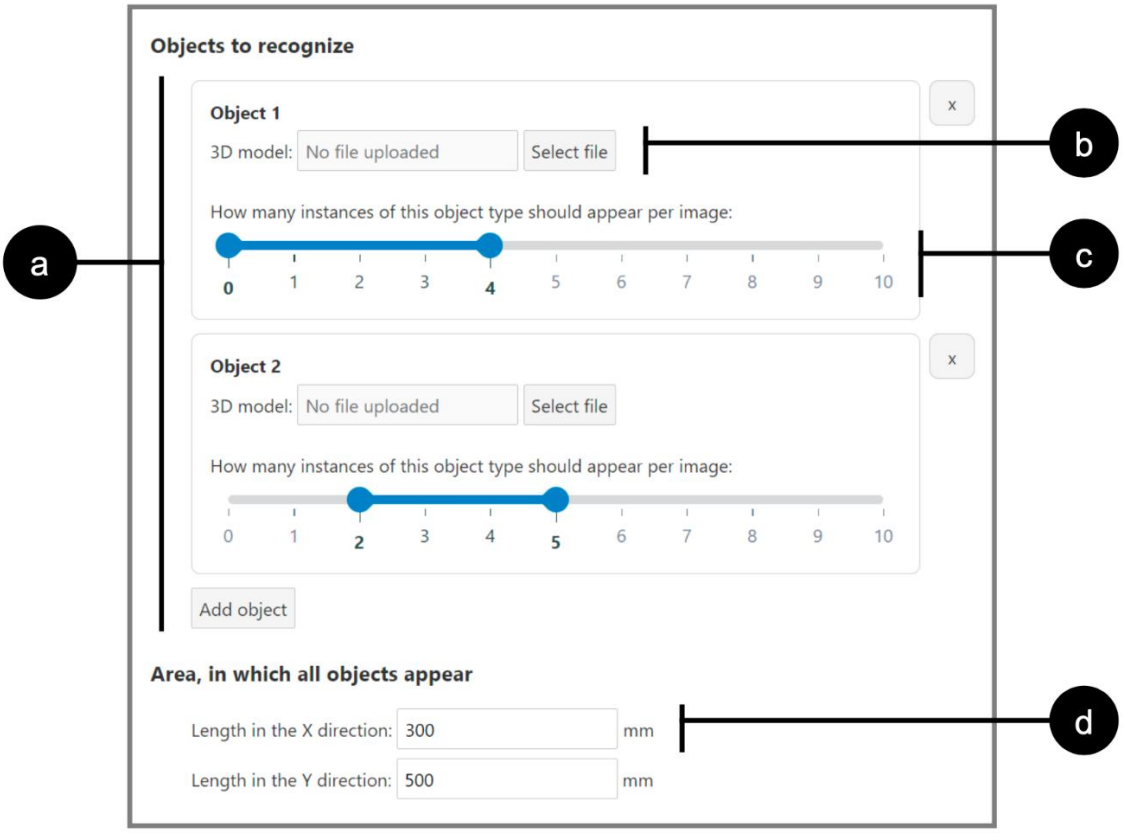


**Abbildung 14: Beispielinterface zur Einstellung von Parametern zur Objekterkennung. a) Dynamische Liste. b)3DModellEingabe, c)BereichEingabe d)NummerEingabe**

Das Arbeitspaket 5 fokussierte sich auf die Konzeption und Realisierung von Schnittstellen, die es Domänenexperten ermöglichen, prozessrelevantes Wissen ohne Programmierkenntnisse in die automatisierte Toolchain zu integrieren. Methodisch wurde hierzu ein rollenbasiertes Interaktionsmodell entwickelt, welches strikt zwischen dem „Data Scientist“, dem „End User“ (Domänenexperten) und dem „ML-System“ unterscheidet, um den Programmieraufwand von der Wissenseingabe zu entkoppeln. Ein zentrales Ergebnis ist die Entwicklung einer webbasierten Nutzerschnittstelle, die auf modular aufgebauten Formularkomponenten basiert. Diese Komponenten, wie etwa Schieberegler, Dropdown-Menüs oder dynamische Listen (z. B.

zur Definition einer variablen Anzahl von Zielobjekten), ermöglichen eine intuitive Konfiguration der Simulationsszenarien. Die so erfassten Daten werden in einem standardisierten JSON-Format gespeichert, was eine hohe Wiederverwendbarkeit und die automatisierte Verarbeitung durch nachgelagerte Module sicherstellt. Ein wesentlicher technologischer Meilenstein war die Umsetzung des Konzepts „Modelling Reality" innerhalb der Benutzeroberflächen. Hierbei wird die Komplexität der technischen Simulationsparameter abstrahiert: Anstatt numerische Reflexionswerte oder komplexe Shader-Parameter direkt zu manipulieren, interagieren Experten mit realitätsnahen Kategorien wie Materialklassen (z. B. „gebürstetes Aluminium", „Polymer") oder physikalischen Objektbeziehungen, siehe Abbildung 14. Die zugrundeliegende Ontologie-Schicht übersetzt diese Eingaben im Hintergrund automatisch in die für den Renderer erforderlichen technischen Spezifikationen.

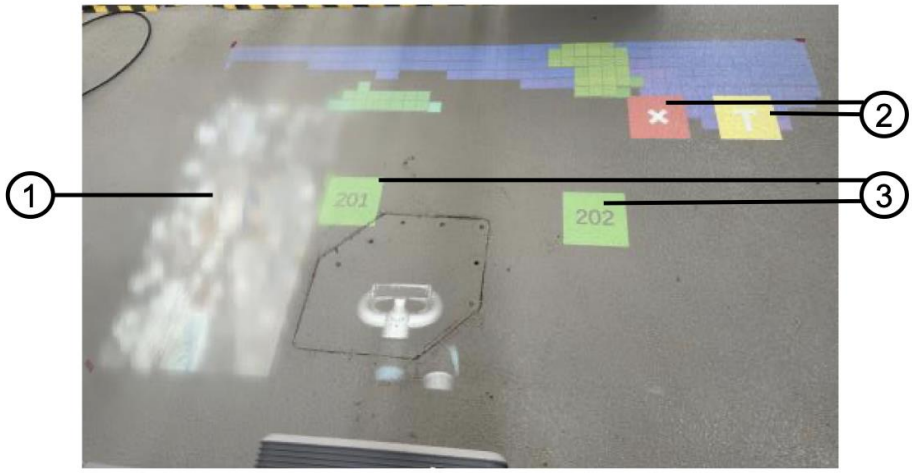


**Abbildung 15 Interface während der Teach-In Phase. 1) Kamerabild; 2) Knöpfe zur Bestätigung/Abbruch des letzten Checkpoints; 3) Check-points**

Darüber hinaus wurden Konzepte für multimodale Schnittstellen zur Erfassung impliziten Expertenwissens untersucht. Hierbei kommen Computer-Vision-Systeme und KI-Applikationen zum Einsatz, um manuelle Handhabungsvorgänge auf dem Shopfloor – beispielsweise bei Schweiß- oder Montageprozessen – exakt in Form von Koordinaten oder Punktwolken zu digitalisieren. Diese digitalisierten Bewegungsabläufe dienen als Basis für die Erstellung virtueller Trainingsszenarien, in denen unerfahrene Mitarbeiter über Augmented Reality (AR) oder Virtual Reality (VR) die Abläufe visuell nachvollziehen und durch systemseitiges Feedback optimieren können. Dies ermöglicht eine Konvergenz von physischer Produktion und virtuellem Wissensmanagement und eine hochgradig skalierbare Qualifizierung von Fachkräften. Das Expertenwissen wird somit direkt aus der Wissensschicht abgerufen, um sowohl die synthetische Datenerzeugung als auch die Planung adaptiver Roboterpfade im Control-Modul zu steuern.

Um die Interaktion zwischen Mensch und Cobot zu optimieren, wurde mittels AR ein User-Interface entwickelt, welches diverse Buttons und Informationen in den Arbeitsbereich des Roboters projiziert. Zur Optimierung der Interaktion zwischen Mensch und Maschine wurde zwischen der Phase des „Teach-Ins" und dem

Produktionsmodus unterschieden. Abbildung 15 zeigt das Interface während des Teach-Ins. Abbildung 16 zeigt das Interface während der Produktionsphase.

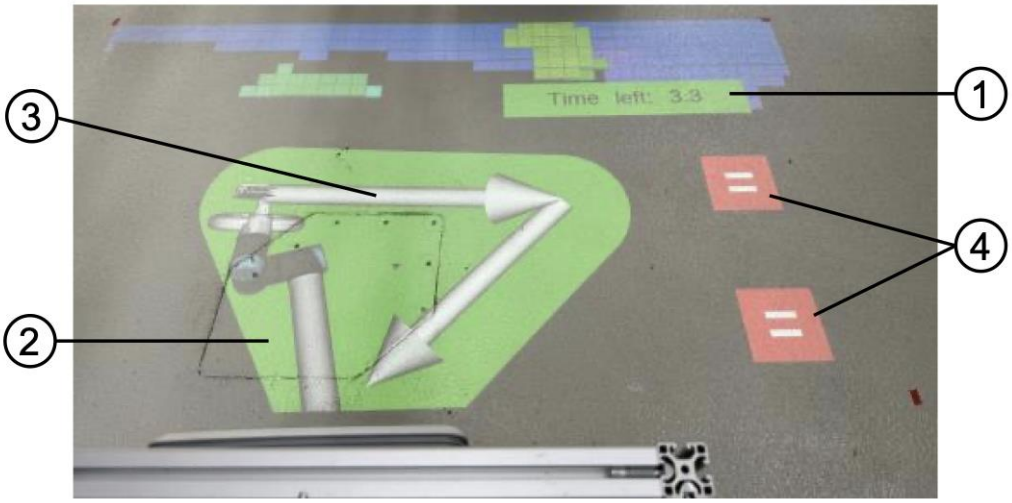


**Abbildung 16: Produktionsmodus. 1) Verbleibende des Prüfzyklus; 2) Aktuelle Cobot Position; 3) Arbeitsbereich des Cobots; 4) Stop**

Die Implementierte Architektur der Nutzerschnittstelle mittels ROS2 ist in Abbildung 17 dargestellt. Hierbei wird eine virtuelle Kamera genutzt, um das virtuelle Bild aufzunehmen und in die Realität zu projizieren.

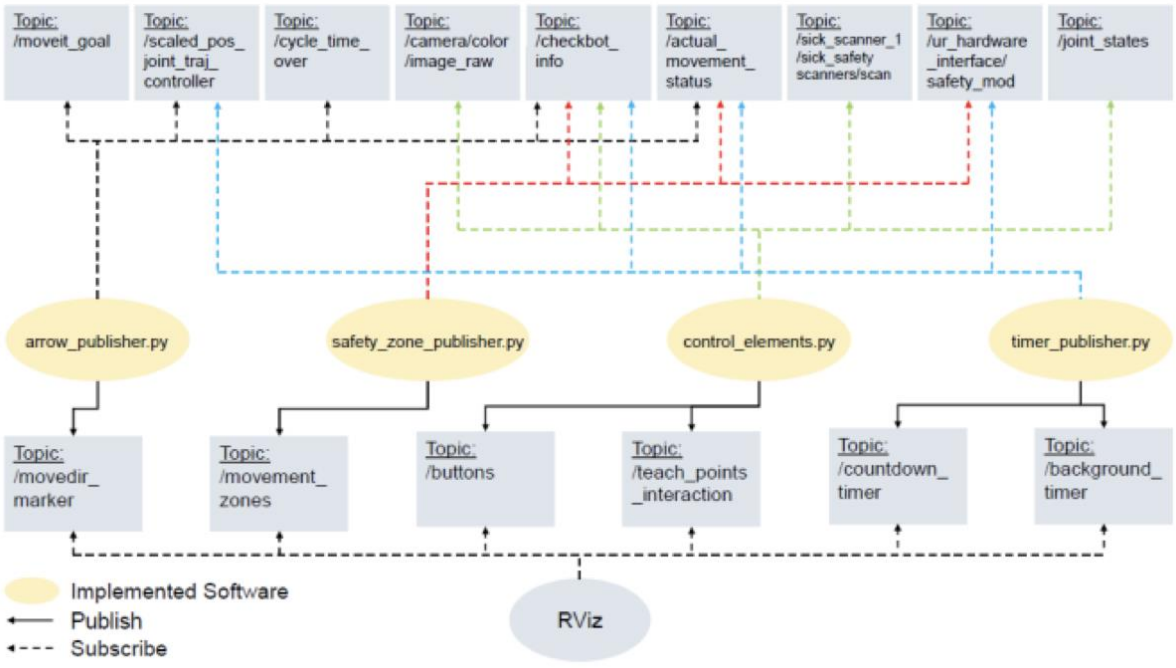


**Abbildung 17: Systemarchitektur der Nutzerschnittstelle in ROS2**

Ergebnis: Benutzerschnittstelle zur Interaktion mit dem Gesamtsystem und zur In-tegration von Experten-wissen in das Modell.

## 2.6 Arbeitspaket 6: Entwicklung des KI-Moduls „Lokalisieren“

In diesem Arbeitspaket wird das erste KI-Modul entwickelt, welches der kamerabasier-ten Lokalisierung von Bauteilen dient. Es bildet auch die direkte Grundlage für die KI-Module aus AP7 und AP8. Das Modul wird durch die erzeugten und beschrifteten Bild-daten aus AP4 trainiert. Dazu wird ein geeigneter KI-Algorithmus ausgewählt und die entsprechende Trainingsumgebung aufgebaut. Der Algorithmus soll in der Lage sein die 6D-Pose von Bauteilen zu ermitteln, um beispielsweise eine korrekte Positionie-rung zu prüfen oder auf dessen Basis Greifpunkte für den Roboter zu bestimmen. Außerdem bestimmt die Wahl des Algorithmus auch, ob 2D- oder 3D-Bilddaten bzw. eine Kombination verwendet werden.

Im Arbeitspaket 6 wurde das KI-Modul zur Lokalisierung von Montagebauteilen entwickelt, welches als zentrale Schnittstelle zwischen der virtuellen Planungsebene und der physischen Realität fungiert. Das methodische Vorgehen konzentrierte sich auf die Implementierung einer robusten Inferenz-Pipeline zur 6D-Posenschätzung, die Translation und Rotation der Objekte im dreidimensionalen Raum präzise bestimmt. Als technologische Basis wurde die Architektur YOLOv8 gewählt, die speziell für die Detektion von Bauteil-Keypoints konfiguriert wurde. Im Gegensatz zu konventionellen Ansätzen, die lediglich 2D-Begrenzungsrahmen (Bounding Boxes) liefern, lokalisiert dieses Modul spezifische geometrische Merkmale am Objekt, die zuvor in der Simulationsumgebung definiert wurden. Abbildung 18 zeigt die Realisierung der zweistufigen Lokalisierungslogik.

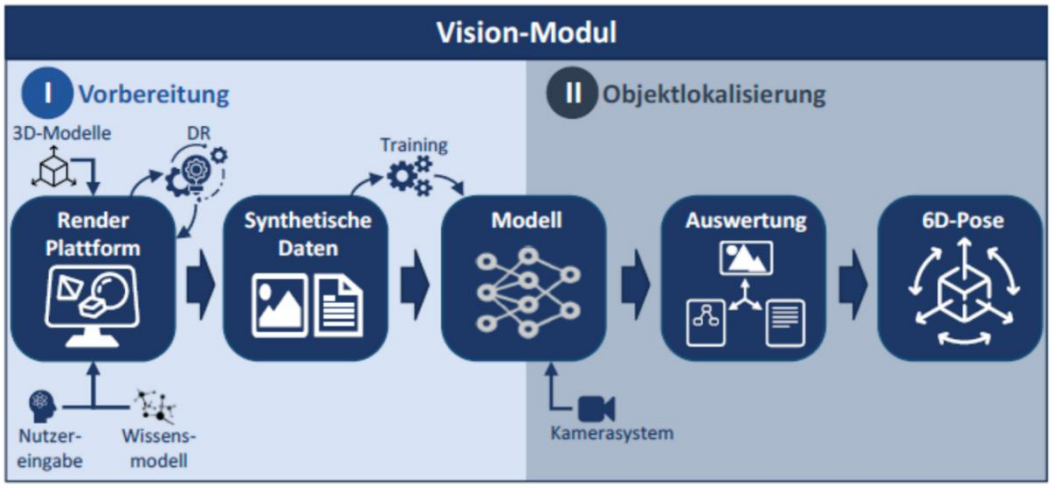


**Abbildung 18: Konzeption der Lokalisierungslogik**

Zunächst identifiziert das YOLOv8-Modell Keypoints in den zweidimensionalen Bilddaten der Kamera. Anschließend wird ein Perspective-n-Point (PnP) Algorithmus eingesetzt, um unter Einbeziehung der intrinsischen Kameraparameter und der bekannten 3D-Koordinaten der Keypoints im Objektkoordinatensystem die Transformation zum Kamerakoordinatensystem zu berechnen. Dieser Prozess ermöglicht eine hochgenaue Bestimmung der 6D-Pose, die für nachgelagerte Greifoperationen (AP8) essenziell ist.

Die Inferenz-Hardware und Software-Anbindung basiert auf einem Intel RealSense Kamerasystem, dessen Bildströme über einen spezialisierten ROS-Node (Robot Operating System) zur latenzarmen Synchronisation mit dem Steuerungssystem direkt in die Toolchain eingespeist werden. In Validierungstests erreichte das ausschließlich auf synthetischen Daten trainierte Modell beachtliche Performance-Metriken mit einem mAP_50 (mean Average Precision) von 0,856 für Keypoints sowie 0,975 für Bounding Boxes. Trotz der Verwendung kostengünstiger Sensorik konnte im Sim2Real-Transfer eine Positioniergenauigkeit von unter 1 mm im Arbeitsraum des Franka Emika Panda Roboters nachgewiesen werden. Durch gezieltes Training mit variierenden Lichtintensitäten, Texturrandomisierungen und virtuellen Verdeckungen weist das Modul zudem eine hohe Robustheit gegenüber industriellen Störfaktoren auf. Die berechneten Positionsdaten werden kontinuierlich an das Control-Modul übermittelt, um die virtuelle Repräsentation in NVIDIA Isaac Sim mit der realen Szenerie abzugleichen und so eine präzise, dynamische Pfadplanung zu realisieren.

Ergebnis: Auf synthetischen Bilddaten trainierbares Modul zur KI-basierten Erkennung von 6D-Posen von Bauteilen.

## 2.7 Arbeitspaket 7: Entwicklung des KI-Moduls „Prüfen“

Das in diesem Arbeitspaket entwickelte Modul soll den Montageablauf verfolgen und erkennen, ob dieser korrekt durchgeführt wird. Auch hier dienen Bilddaten der einzelnen Montageschritte in Kombination mit dem zuvor modellierten Prozessablauf als Basis der Erkenntnisgewinnung. Durch die Ermittlung des aktuellen Montagezustands kann die Montagereihenfolge sukzessive abgearbeitet werden bzw. ein potentieller Fehler im Montageprozess detektiert werden. Denkbar sind sowohl eine qualitative Ermittlung des Fortschritts als auch eine quantitative Prüfung der Bauteilpositionen.

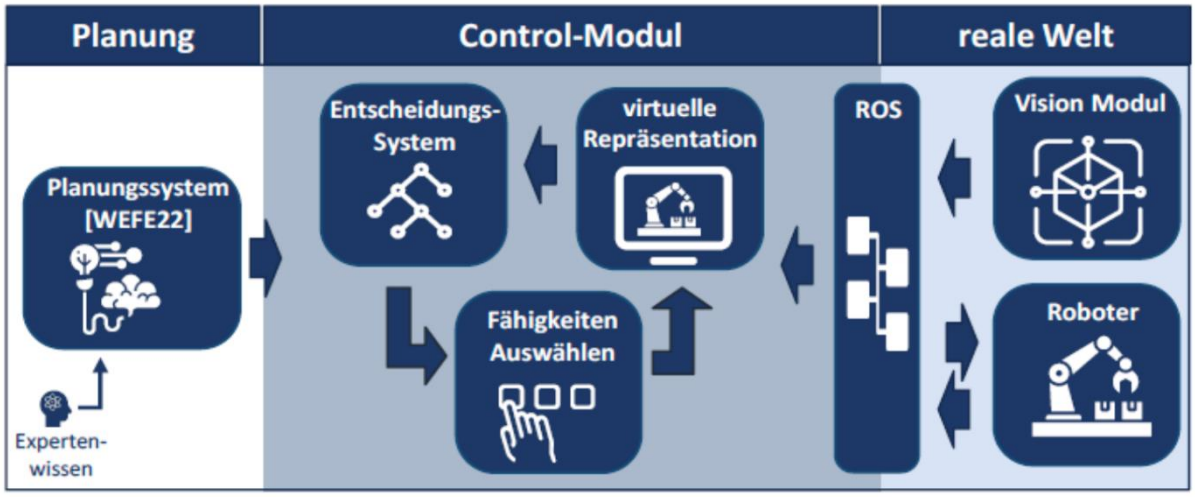


**Abbildung 19: Planung und Steuerung, beeinflusst durch den aktuellen Montagezustand, detektiert mittels Kamera**

Die Entwicklung des KI-Moduls „Prüfen“ konzentrierte sich methodisch auf die Etablierung einer automatisierten Pipeline, die sowohl die Überwachung des aktuellen Montagezustands als auch die präzise Detektion von Qualitätsmängeln auf Basis rein synthetischer Trainingsdaten ermöglicht. Abbildung 19 zeigt das Zusammenspiel der Planung und Steuerung, welche auf Basis von Kameradaten des aktuellen Montagezustandes beeinflusst wird.

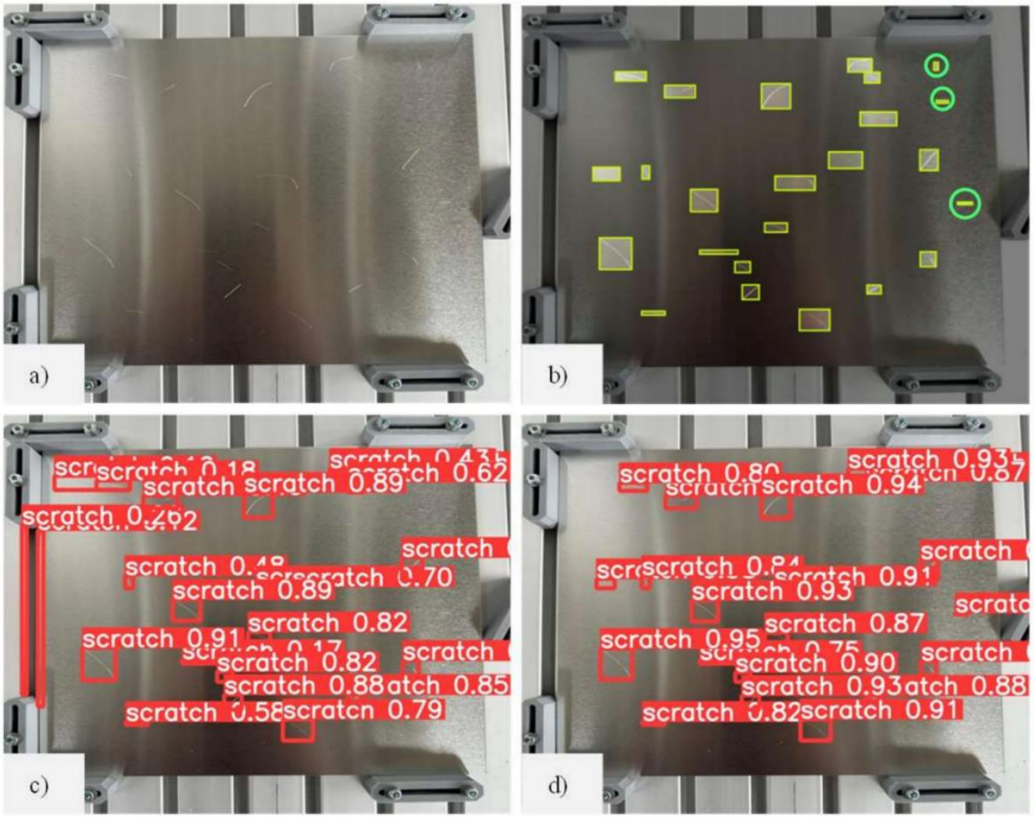


**Abbildung 20: Ergebnisse der Vorhersage auf Realdaten. a) Abbildung des Aluminiumblechs mit Kratzern. b) Annotierte Abbildung. c) Datensatz 1: TP: 19, FP: 3, FN: 5; Präzision: 79,2 %, Recall: 86,3 %. d) Datensatz 2: TP: 21, FP: 0, FN: 3; Präzision: 87,5 %, Recall: 100 %**

Die Prüfung auf Oberflächendefekten basiert auf einer dualen Strategie zur Datengenerierung, um die visuelle Komplexität industrieller Oberflächenfehler abzubilden. Hierbei wurden in der Simulationsumgebung zwei komplementäre Modellierungsansätze verfolgt: Einerseits die shader-basierte Erzeugung von Defekten mittels PBR, bei der Normal- und Roughness-Maps modifiziert werden, um optische Kratzer ohne Geometrieänderung darzustellen, und andererseits die geometrische Modifikation des Meshes durch Extrusionen im Mikrometerbereich (<0,01 mm), um realistische Schattenwürfe zu provozieren. Im Rahmen der Architekturauswahl wurde ein technischer Benchmark zwischen verschiedenen Deep-Learning-Modellen wie YOLOv8, YOLOv9 und DETR (Detection Transformer) durchgeführt. Als zentrales Ergebnis stellte sich die Überlegenheit von YOLOv9 heraus, da diese Architektur durch den Einsatz von Programmable Gradient Information und Generalized Efficient Layer Aggregation Networks (GELAN) einen deutlich verbesserten Informationsfluss

während des Trainings gewährleistet. Dies führte bei der Erkennung von feinen Kratzern auf metallischen Oberflächen zu einem mAP von 66,2 %. Es konnte festgestellt werden, dass die Objektdichte innerhalb der Trainingsbilder (Anzahl der Objekte pro Frame) einen signifikanteren Einfluss auf die Robustheit des Modells hat als die rein zufällige räumliche Platzierung der Objekte im Arbeitsraum.

Bei der Übertragung der Modelle in die reale Fertigungsumgebung (Sim2Real-Transfer) identifizierte das Modul zunächst Herausforderungen durch Fehlalarme an konstruktionsbedingten Kanten, wie etwa T-Nuten an Montageprofilen. Um diese „Domain Gap“ technisch zu schließen, wurde eine Strategie implementiert, bei der dem synthetischen Trainingssatz lediglich ca. 15 % ungelabelte reale Bilddaten beigemischt wurden, was die Fehlerrate massiv senkte, siehe Abbildung 20.

Das fertige Prüfmodul fungiert nun als integraler Bestandteil der Prozesssteuerung, indem es Detektionsergebnisse in Echtzeit an das „Decider-Network“ zurückspielen kann. Diese Logik ermöglicht es dem System, den Fortschritt kontinuierlich zu verifizieren und bei detektierten Fehlern oder unvollständigen Montageschritten autonom entsprechende Korrekturmaßnahmen einzuleiten.

Ergebnis: Auf synthetischen Bilddaten trainierbares Modul zur Erkennung und Verfolgung des aktuellen Montagezustands auf Basis von Kameradaten.

## 2.8 Arbeitspaket 8: Entwicklung des KI-Moduls „Greifen“

Auch dieses Arbeitspaket baut auf den Ergebnissen von AP6 auf. Um Bauteile die erfolgreich lokalisiert wurden auch greifen zu können muss ein robuster Greifpunkt am Bauteil erfasst werden und in eine Zielpose für den Roboter überführt werden. Um einen möglichst robusten Griff durchführen zu können, sollen Parameter wie beispielsweise die Lage des Schwerpunkts, die Position des Bauteils im Arbeitsraum sowie Randbedingungen wie etwa Montageflächen bei der Wahl geeigneter Greifpunkte berücksichtigt werden.

Auch hier wird die Wahl des implementierten Algorithmus auf Basis der ermittelten Anforderungen gefällt. Gegebenenfalls kann eine KI-basierte Greifpunktbestimmung durch nachgelagerte heuristische Ansätze optimiert werden.

Im Arbeitspaket 8 wurde das Modul „Greifen“ entwickelt, welches die Ergebnisse der 6D-Posenschätzung aus dem Vision-Modul (AP6) in physische Roboterbewegungen übersetzt. Das methodische Vorgehen basierte auf der Implementierung einer „Grasping Skill“ innerhalb einer modularen „Robot Skill Library“, die als Teil des übergeordneten Control-Moduls fungiert. Ein wesentlicher Bestandteil war dabei die Einbettung dieses Skills in ein „Decider-Network“, welches als gerichteter azyklischer Graph (DAG) implementiert wurde, um komplexe und adaptive Montagefolgen zu orchestrieren.

Eine automatisierte Transformationspipeline wurde implementiert, welche die vom Vision-Modul gelieferte 6D-Pose des Objekts vom Kamerakoordinatensystem in das Basiskoordinatensystem des Roboters überführt.

Hierbei wurde eine „Eye-in-Hand“-Konfiguration verwendet, bei der die extrinsische Kalibrierung der Kamera am Endeffektor eine präzise Berechnung der Zielkoordinaten für den Greifer ermöglicht. Somit bildet das Greif-Modul das Bindeglied zwischen der bildbasierten Perzeption (Vision) und der manipulationstechnischen Ausführung.

Die Realisierung des Greifvorgangs basiert auf einer Skill-basierten Architektur, bei der der Prozess als eigenständiger, parametrisierbarer Dienst gekapselt und über ein Decider-Network initiiert wird, was einen agilen Wechsel zwischen Greifstrategien je nach Bauteilgeometrie erlaubt. Zur virtuellen Absicherung wurde das System vor dem physischen Einsatz in NVIDIA Isaac Sim unter Nutzung von Isaac Cortex validiert, wodurch Singularitäten und Kollisionen bereits in der Planungsphase detektiert und vermieden werden konnten. Die operative Steuerung erfolgt über eine Echtzeit-Synchronisation mittels ROS-Nodes, wobei eine kontinuierliche Rückkopplung der Gelenkwinkel eine präzise Korrektur der Greiferpose sicherstellt. Ergänzend dazu nutzt das System eine wissensbasierte Pose-Optimierung auf Basis der in AP3 entwickelten Taxonomie, um optimale Greifpunkte direkt aus CAD-Metadaten zu extrahieren. Dies ermöglicht eine dynamische Anpassung der Anfahrbewegung an die jeweilige Objektorientierung.

In Kombination mit der hochpräzisen 6D-Lokalisierung gewährleistet dieses Modul die sichere Handhabung von Bauteilen selbst in unstrukturierten Ausgangslagen und bildet damit die prozesskritische Grundlage für die nachfolgenden Fügeoperationen.

Ergebnis: KI-Modul zur Bestimmung von Greifpunkten an lokalisierten Objekten.

## 2.9 Arbeitspaket 9: Entwicklung des KI-Moduls „Fügen“

Das vierte zu entwickelnde Modul betrifft kraftbasierte Fügeoperationen durch den Roboter. Hier sind vorrangig die Fügeoperationen Zusammensetzen sowie Steckverbindungen zu nennen. Es wird ein RL Modell implementiert, welches in der in AP4 entwickelten physikalischen Simulation das kraftbasierte Fügen von diversen Bauteilgeometrien erlernen kann. Dazu werden die entsprechenden Bauteilpaarungen in die Simulationsumgebung überführt und dessen Zusammensetzen in zahlreichen Epochen randomisiert erlernt. Somit können sensitive Fügeoperationen durchgeführt werden, ohne den Roboter explizit zu programmieren. Dies ist vor allem für die Robustheit von Montageprozessen relevant, da kontaktreiche Operationen in der Robotik häufig zu Störfällen und somit zu Produktivitätseinbußen führen können. Zusätzlich zu den Hyper- und Trainingsparametern muss bei einem RL-Modell auch die Belohnungsfunktion entwickelt werden.

Das Arbeitspaket 9 befasste sich mit der technischen Entwicklung des KI-Moduls „Fügen“, dessen primäres Ziel die Automatisierung hochkomplexer, kontaktbehafteter Montageprozesse wie der Montage von Schnappverbindungen (Snap-fits) war. Das methodische Vorgehen stützte sich maßgeblich auf den Einsatz von Deep Reinforcement Learning (DRL) in Kombination mit Strategien des Curriculum Learning (CL), um die

Herausforderungen enger geometrischer Toleranzen und nichtlinearer Materialverformungen zu bewältigen. Als Trainingsumgebung wurde NVIDIA Isaac Gym eingesetzt, welches durch seine end-to-end GPU-Beschleunigung die Simulation von bis zu 128 parallelisierten Umgebungen (Vectorized Environments) ermöglichte, was die Konvergenzrate des RL-Agenten im Vergleich zu CPU-basierten Simulationen massiv steigerte.

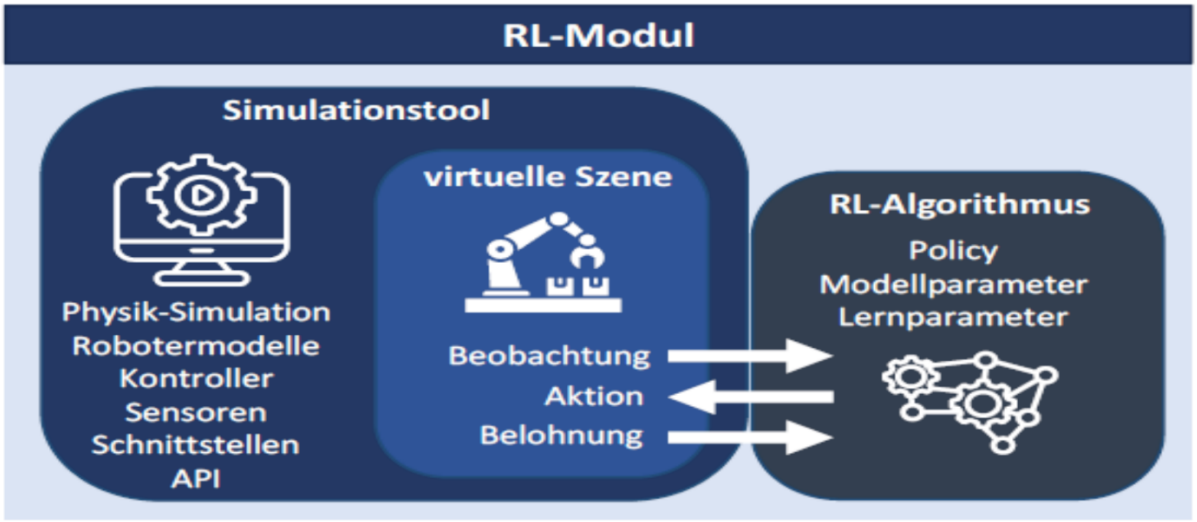


**Abbildung 21: Interaktion zwischen Reinforcement Learning und der virtuellen Szene zur Optimierung der Policy des Agenten**

Eine hybriden Belohnungsfunktion (Reward Function) wurde definiert, die auf der Minimierung des euklidischen Abstands zwischen den Bauteil-Keypoints und der Zielpose basiert. Um das elastische Verhalten der Schnappharken realitätsnah abzubilden, wurden diese als kinematische Ketten aus starren Teilkörpern modelliert, deren Gelenksteifigkeiten spezifisch kalibriert wurden. Diese physikalische Abstraktion erlaubte es dem RL-Agenten, die notwendigen Fügekräfte und Trajektorien autonom zu erlernen, ohne dass eine explizite Programmierung der Kontaktphysik erforderlich war. Abbildung 11 zeigt die Interaktion zwischen dem Reinforcement Learning Agenten und der Simulationsumgebung.

Um die Robustheit des Moduls zu erhöhen, wurde ein technisches Framework für Curriculum Learning integriert. Hierbei wurde der Schwierigkeitsgrad der Fügeaufgabe während des Trainings sukzessive gesteigert, indem die geometrischen Toleranzen (Clearance) zwischen den Bauteilen schrittweise reduziert wurden – beginnend bei einer hohen Fehlertoleranz bis hin zur Zieltoleranz von unter 0,1 mm. Dieser Ansatz führt zu einer signifikant höheren Erfolgsrate und einer schnelleren Exploration des Zustandsraums im Vergleich zu herkömmlichem Training. Als Algorithmus kam dabei bevorzugt Soft Actor-Critic (SAC) zum Einsatz.

Die Übertragung der gelernten Strategien auf den realen Franka Emika Panda Roboter erfolgte über das Control-Modul unter Nutzung einer kartesischen Impedanzregelung. Diese technische Implementierung ermöglichte es dem Roboter, während des Fügevorgangs eine definierte Nachgiebigkeit (Compliance) zu ermgöglichen, wodurch mechanische Spannungen bei minimalen Fehlstellungen der 6D-Pose kompensiert werden

konnten. Das Modul „Fügen“ bildet somit den technologischen Abschluss der Kette von der Perzeption über die Planung bis hin zur hochpräzisen physischen Manipulation, wobei die gelernten Skills über das Decider-Network bedarfsgerecht abgerufen werden.

Ergebnis: KI-Modul für das sensitive Fügen von Bauteilen

## 2.10 Arbeitspaket 10: Aufbau des Demonstrators

Die in AP11 geplante Validierung und Evaluation der im Forschungsprojekt erarbeite-ten Systemarchitekturen und Softwaremodule wird anhand eines Demonstrators pro-zessnah erfolgen. Dazu steht im Smart Automation Lab des WZL bereits ein Prüfstand aus dem Forschungsprojekt KoKoMo zur Verfügung, welcher vier Leichtbauroboter (Franka Emika Panda, Kuka LBR iiwa, Rethink Robotics Sawyer, Doosan M1013), eine Zellenumgebung und diverse Sensorik beinhalten. Diese werden im Zuge dieses AP um die zusätzlich benötigten Komponenten ergänzt. Zunächst wird dazu die für die in AP4 bis AP9 benötigte Rechenleistung, Sensorik und Greifmittel beschafft und inte-griert. Anschließend werden auch die entwickelten KI-Module und die Benutzerschnitt-stelle integriert. Wenn nötig, wird das bestehende System aus Werkzeugen und einem Werkzeugwechselsystem erweitert, um dem Roboter ein ausreichend breites Spekt-rum an Funktionalitäten zu ermöglichen.

Im Arbeitspaket 10 erfolgte die physische Realisierung und der Aufbau des Demonstrators, welcher als Validierungsplattform für das entwickelte Sim2Real-Framework und die in den vorangegangenen Arbeitspaketen erarbeiteten KI-Module dient. Das methodische Vorgehen konzentrierte sich auf die technologische Konvergenz zwischen dem digitalen Zwilling und der realen Fertigungszelle, um eine hochgradige Synchronisation im Sinne des „Industrial Metaverse“ zu erreichen. Als zentrale Hardware-Komponente wurde ein Franka Emika Panda Roboter mit sieben Freiheitsgraden (7-DOF) integriert, der aufgrund seiner feinfühligen Drehmomentsensorik in allen Gelenken die notwendigen Voraussetzungen für die Umsetzung der in AP9 entwickelten impedanzgeregelten Fügeoperationen bietet. Die visuelle Perzeption wird durch ein am Endeffektor montiertes Intel RealSense D435 Kamerasystem realisiert, dessen extrinsische Kalibrierung („Eye-in-Hand“) eine präzise Transformation der detektierten 6D-Posen in das Roboterkoordinatensystem erlaubt.

Der Aufbau basiert auf einer bidirektionalen Kommunikationsinfrastruktur auf Basis des Robot Operating System (ROS). Diese Schnittstelle ermöglicht einen kontinuierlichen Datenaustausch zwischen der realen Steuerung und der virtuellen Simulationsumgebung in NVIDIA Isaac Sim. Hierbei werden Gelenkzustände, Sensordaten und Transformationsmatrizen in Echtzeit abgeglichen, was die Grundlage für das konzipierte Control-Modul bildet. Der Demonstrator fungiert als physisches Äquivalent zur virtuellen Toolchain, wobei das „Decider-Network“ die Orchestrierung der Montageabläufe übernimmt.

Der Demonstrator wurde spezifisch für die im Projekt definierten Use-Cases ausgerüstet. Für die kognitive Montage wurden 3D-gedruckte Bauteile verwendet, die komplexe Schnappverbindungen aufweisen. Die Modellierung dieser Bauteile als kinematische Ketten in der Simulation konnte erfolgreich am physischen Aufbau validiert werden, indem der Roboter die gelernten Reinforcement-Learning-Strategien zur Kompensation von Lagefehlern autonom ausführte. Für die Oberflächenprüfung wurde eine Prüfstation in den Demonstrator integriert, die unter kontrollierten Lichtverhältnissen die Erkennung von Defekten auf metallischen Profilen ermöglicht.

Ergebnis: Demonstrators als Grundlage für die Validierung der in den anderen AP erstellten Konzepte und Module.

## 2.11 Arbeitspaket 11: Validierung und Evaluierung

Die entwickelten Methoden und Ansätze werden im Rahmen dieses AP an dem Demonstrator hinsichtlich ihrer Einsetzbarkeit und Funktion getestet und abschließend bewertet. Dazu stehen diverse, realitätsnahe Anwendungsfälle aus dem PA zur Verfügung. Diese Anwendungsfälle werden in Demonstrationsszenarien überführt anhand derer die entwickelte Methode überprüft werden kann. Dabei wird die Funktionsfähigkeit der Methode und der entwickelten KI-Module zur Erfüllung der in AP1 aufgestellten Anforderungen überprüft.

Das Arbeitspaket 11 bildete den abschließenden Meilenstein des Projekts und konzentrierte sich auf die umfassende technische Validierung und Evaluierung des Gesamtsystems hinsichtlich seiner Performanz, Robustheit und Effizienzsteigerung.

Die Evaluation des Vision-Moduls zeigte, dass bei der 6D-Posenschätzung eine Lokalisierungsgenauigkeit von unter 1 mm im realen Arbeitsraum des Franka Emika Panda Roboters erreicht wurde.

In Bezug auf die kognitive Montage wurde die Erfolgsrate der durch Reinforcement Learning (RL) trainierten Fügeskills evaluiert. Die Ergebnisse zeigen, dass der Einsatz von Curriculum Learning (CL) in Kombination mit dem Soft Actor-Critic (SAC) Algorithmus die Konvergenzgeschwindigkeit im Vergleich zu herkömmlichen RL-Ansätzen um den Faktor drei steigerte. Während der RL-Agent in der Simulation (NVIDIA Isaac Gym) eine nahezu 100-prozentige Erfolgsquote erreichte, konnte im realen Sim2Real-Transfer eine Erfolgsrate von über 92 % bei der Montage von-Schnappverbindungen nachgewiesen werden. Im Curriculum-Learning-Ansatz wurde der PPO-Algorithmus in vier Stufen trainiert, wobei in jeder Phase die Geometrie des Aufnahmeteils (female part) angepasst wurde. Nach jeweils 1.000 Epochen pro Stufe wurde das leistungsfähigste Modell als Basis für die nächste Stufe verwendet. Abbildung 22 illustriert den Verlauf der Erfolgsraten über den gesamten Trainingszeitraum. Diese beschreiben den Anteil der Umgebungen, in denen ein Fügevorgang innerhalb einer bestimmten Epoche erfolgreich abgeschlossen wurde. Die x-Achse jedes Graphen ist kumulativ

aufgetragen, wobei jede Stufe bei der Epoche beginnt, in der das beste Modell der vorangegangenen Stufe ausgewählt wurde. Bereits zu Beginn des Trainings in Stufe eins konnte ein signifikanter Lernfortschritt beobachtet werden, wobei die Erfolgsquoten schließlich 95 % erreichten. Dies bestätigt die Eignung der Trainingskonfiguration für eine präzise Positionierung. In einer anspruchsvolleren Phase mit verringertem Spiel (Stufe zwei) erbrachte der Agent bereits ab der ersten Epoche gute Leistungen. Dies deutet darauf hin, dass die in Stufe einer erlernten Positionierung bereits präzise genug für eine engere Passung war.

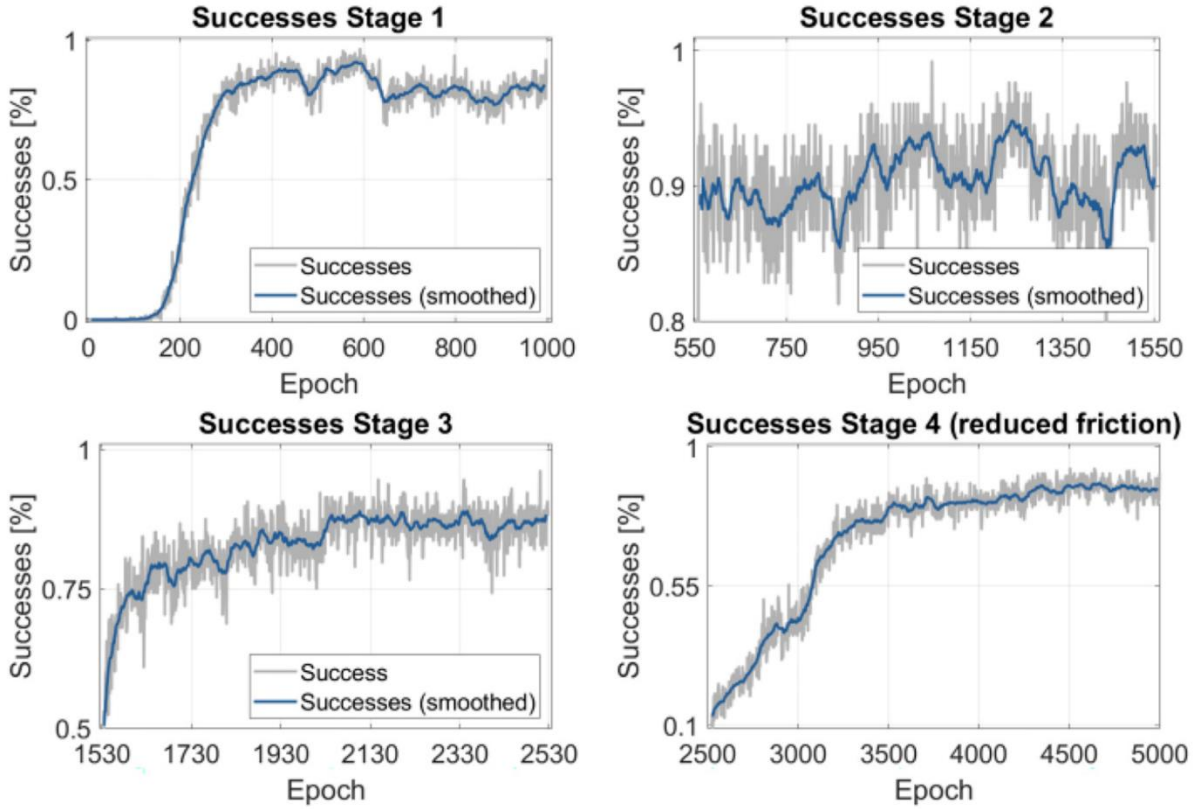


**Abbildung 22: Erfolgsrate des Trainings zum Fügen der Schnappverbindung**

In Stufe drei, in der die Aufgabe erstmals das Aufbringen einer Fügekraft erforderte, zeigte das Modell eine erhebliche Verbesserung: Die mittlere Erfolgsquote stieg von anfangs 50 % auf über 96 % an. Dies unterstreicht die Fähigkeit des Modells, sich an Aufgaben anzupassen, bei denen eine Kraftausübung notwendig ist. Mit der Einführung der finalen Geometrie des Aufnahmeteils (Stufe vier) wurden zunächst keine erfolgreichen Versuche verzeichnet. Während der Roboter die Checkpoint-Position zuverlässig erreichte, misslang das Aufbringen der erforderlichen Fügekraft für das finale Einsetzen. Trotz vorheriger Anpassungen der Gelenksteifigkeit für ein realistischeres Verhalten offenbaren sich hier erneut Einschränkungen von Isaac Sim bei der exakten Simulation der gewählten Steifigkeitsparameter sowie die Herausforderung, Werte zu finden, die das reale Systemverhalten originalgetreu widerspiegeln. Eine Reduzierung des Reibungskoeffizienten von 0,4 auf 0,05 verringerte die benötigte Fügekraft in der Simulation weiter, was zu einer schrittweisen Verbesserung der Erfolgsquote führte. Da die Kurve nach 1.000 Epochen noch nicht konvergierte, wurde das

Training um zusätzliche 1.500 Epochen verlängert, was schließlich zu einer Erfolgsquote von 92,3 % für die Zielaufgabe führte. Diese Ergebnisse belegen, dass der RL-Agent tatsächlich lernen kann, die Schnappverbindung inklusive der benötigten Fügekraft zu montieren – und das mit einer Reward-Funktion, die sich ausschließlich auf die Optimierung der Positionsgenauigkeit konzentriert.

| Construct | Model | Items |
|---|---|---|
| Performance Expectancy | UTAUT<br>UTAUT2 | PE1: I would find the system useful in my job.<br>PE2: Using the system enables me to accomplish tasks more quickly.<br>PE3: Using the system increases my productivity.<br>PE4: If I use the system, I will increase my chances of getting a raise. |
| Effort Expectancy | UTAUT<br>UTAUT2 | EE1: My interaction with the system would be clear and understandable.<br>EE2: It would be easy for me to become skillful at using the system.<br>EE3: I would find the system easy to use.<br>EE4: Learning to operate the system is easy for me. |
| Social Influence | UTAUT<br>UTAUT2 | SI1: People who influence my behavior think that I should use the system.<br>SI2: People who are important to me think that I should use the system.<br>SI3: The senior management of this business has been helpful in the use of the system.<br>SI4: In general, the organization has supported the use of the system. |
| Facilitating Conditions | UTAUT<br>UTAUT2 | FC1: I have the resources necessary to use the system.<br>FC2: I have the knowledge necessary to use the system.<br>FC3: The system is not compatible with other systems I use.<br>FC4: A specific person (or group) is available for assistance with system difficulties. |
| Hedonic Motivation | UTAUT2 | HM1: Using the system is fun.<br>HM2: Using the system is enjoyable.<br>HM3: Using the system is very entertaining. |
| Price Value | UTAUT2 | PV1: The system is reasonably priced.<br>PV2: The system is a good value for the money.<br>PV3: At the current price, the system provides a good value. |
| Habit | UTAUT2 | HT1: The use of the system has become a habit for me.<br>HT2: I am addicted to using the system.<br>HT3: I must use the system. |
| Behavioral Intention | UTAUT<br>UTAUT2 | BI1: I intend to continue using the system in the future.<br>BI2: I will always try to use the system in my daily life.<br>BI3: I plan to continue to use the system frequently. |

**Abbildung 23: Konstrukte und Items des UTAUT und UTAUT2 Fragebogens**

Die Evaluierung der übergeordneten Systemeffizienz bestätigte das Ziel der Reduktion des „Programmierer-Flaschenhalses“. Durch die Einführung der ontologiebasierten Nutzerschnittstelle konnte die Zeitspanne von der Aufgabenstellung (Einreichen eines neuen CAD-Modells) bis zum einsatzbereiten KI-Modell drastisch verkürzt werden. Im Vergleich zum konventionellen manuellen Labeling-Prozess, der bei industriellen Projekten oft über 80 % der Gesamtkosten beansprucht, ermöglichte das Framework eine deutliche Reduktion des

manuellen Aufwands, da der Domänenexperte lediglich funktionale Anforderungen und Geometriedaten definieren muss, während die technische Datengenerierung autonom erfolgt.

Die grafische Benutzerschnittstelle zum Teachen der Checkpoints des kollaborativen Roboters wurde in einer Nutzerstudie mit 23 Teilnehmer evaluiert. Die Hälfte der Teilnehmer arbeitete zuerst mit dem konservativen System (ausgestattet ohne die entwickelte, auf augmentierter Realität basierender Nutzerschnittstelle), die andere Hälfte zuerst mit dem innovativen System (ausgestattet mit der entwickelten Nutzerschnittstelle). Anschließend wurde ein Fragebogen mit Fragen der Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) beantwortet, welche in Abbildung 23 zu sehen sind.

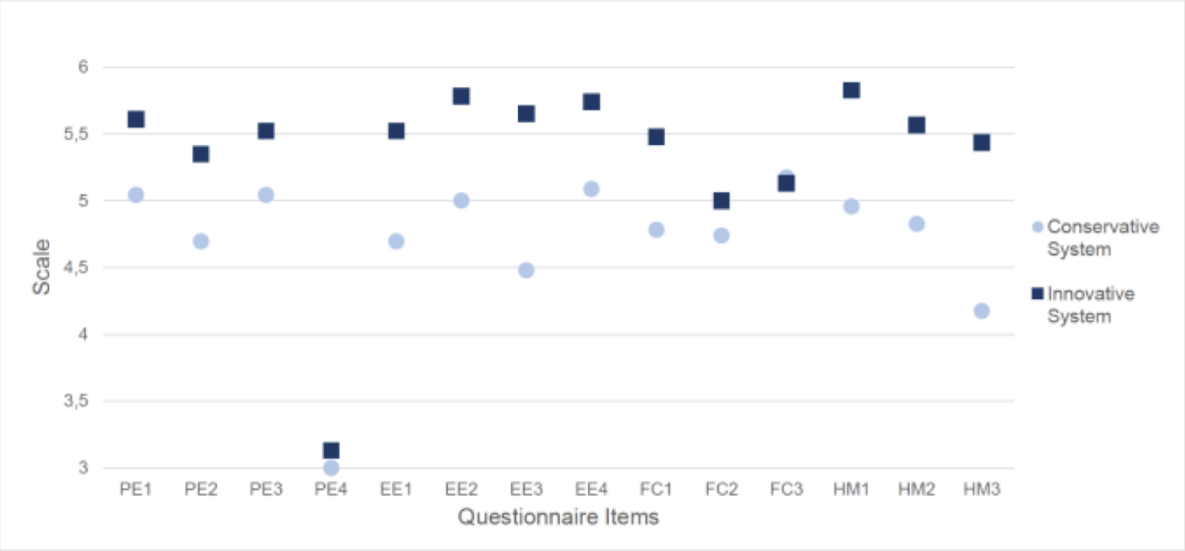


**Abbildung 24: Ergebnisse der Antworten der Teilnehmenden**

Die Antworten der Teilnehmer sind in Abbildung 24 dargestellt. Insgesamt lässt sich festhalten, dass die Nutzer, die auf augmentierter Realität basierende Nutzerschnittstelle positiv bewerteten.

Abschließend wurde die Resilienz des Gesamtsystems durch das „Decider-Network“ validiert. Das System demonstrierte die Fähigkeit, bei Detektion von Prozessfehlern durch das Prüfmodul (z. B. ein nicht korrekt eingerasteter Schnapphaken) autonom in einen Fehlerbehebungsmodus zu wechseln oder den Prozess sicher abzubrechen.

Ergebnis: Aussage über die Funktionsfähigkeit sowie die qualitative Flexibilitäts- und Robustheitssteigerung.

## 2.12 Arbeitspaket 12: Projektmanagement, Dokumentation und Dissemination

Die im Projekt entwickelten Methoden und Ansätze sowie erarbeiteten Inhalte sind zu dokumentieren. Dabei belegt die Dokumentation den Projektfortschritt und bildet die Grundlage für die projektbegleitende Dissemination. Neben der Zusammenarbeit mit dem PA über die Laufzeit hinweg wird bei Erreichen der gesetzten Meilensteine der jeweils aktuelle Stand der Dokumentation an alle Teilnehmer übermittelt. Detaillierte Transfermaßnahmen sind in Kapitel 5 beschrieben. Das Projektmanagement über-wacht dabei über die Projektlaufzeit hinweg den aktuellen Fortschritt im Projekt und spiegelt diesen am eingereichten Arbeitsplan wider. Weiter trägt das Projektmanagement die organisatorische Verantwortung für die Projekttreffen des PA. Die geleisteten und weiterhin geplanten Transfermaßnahmen lassen sich Kapitel 7 entnehmen.

# 3 Verwendung der Zuwendung

Für die beschriebenen wissenschaftlichen Tätigkeiten wurden die in Tabelle 1 aufgeführten Personalaufwände für wissenschaftlich-technisches Personal der Stufe HPA-A und HPA-B benötigt.

**Tabelle 1: Personaleinsatz**

| Forschungseinrichtung | Personal | Personenmonate |
|---|---|---|
| FVP | Wissenschaftlicher Angestellter HPA-A | 16,25 PM |
| FVP | Technischer Angestellter HPA-B | 7,25 PM |
| WZL | Wissenschaftlicher Angestellter HPA-A | 23,73 PM |

Die im Rahmen des Projektes beschäftigten studentischen Hilfskräfte waren projektbegleitend mit Recherche-, Dokumentations- und einfachen Implementierungsaufgaben betraut. Die Ausgaben für die Beschäftigung der Stufen HPA-C und HPA-F lassen sich dem beiliegenden zahlenmäßigen Nachweis (bZ) entnehmen.

Zur Bearbeitung des Vorhabens wurden die in Tabelle 2 aufgelisteten Geräte beschafft.

**Tabelle 2: Gerätebeschaffungen**

| Forschungseinrichtung | Bezeichnung | Kosten |
|---|---|---|
| WZL | Dell Precision 7920 Tower XCTO Base - KI Rechnersystem | 11.868,65 € |

Die Beschaffung des Rechnersystems diente der Umsetzung der geplanten Tätigkeiten in AP 6 bis AP 11. Für den Aufbau und die Evaluierung der KI-Basierten Montage sollen heute oft CPU-limitierte Berechnungsschritte mittels moderner GPU-Technologie parallelisiert werden. Hierdurch konnte die notwendige Berechnungsgeschwindigkeit, die für die industrielle Nutzung des Gesamtansatzes vorausgesetzt wird, erreicht werden. Das zu beschaffenden Geräte dient daher ausschließlich der Umsetzung der beschriebenen Konzepte und Demonstratoren. Der Demonstratoraufbau soll so den beteiligten Forschungsstellen und Unternehmen ermöglichen, die gewonnenen Erkenntnisse praktisch umzusetzen. Unternehmen des PA können an dem Demonstrator ihre Montageprozesse an Beispielbauteilen validieren (Endanwender). Der Aufbau am WZL wird darüber hinaus gerade KMU mit begrenzten finanziellen Mitteln die Chance bieten, die gewonnen Erkenntnisse zeit- und praxisnah umzusetzen (CAD/CAM-Systementwickler).

# 4 Notwendigkeit und Angemessenheit der geleisteten Arbeit

Insgesamt entsprechen die geleisteten Arbeiten dem begutachteten und bewilligten Antrag und waren daher für die Durchführung des Forschungsvorhabens notwendig und angemessen. Die ausgeführten Arbeiten wurden mit den Mitgliedern des PA besprochen und für das Erreichen des Projektziels gemeinsam für notwendig und angemessen befunden. Das Vorhaben wurde zur Laufzeit kostenneutral um sechs Personenmonate verlängert. Die Änderung begründet sich aus Beschaffungsverzögerungen durch lange Lieferzeiten erfolgskritischer Rechenleistung in der ersten Projekthälfte. Diese wird zur Simulation und dem Training von KI-Applikationen benötigt und stellt somit eine erfolgskritische Komponente zur Entwicklung und Demonstration der angestrebten Ergebnisse dar. Da der Rechner verzögert zur Verfügung stand, können die Arbeiten im Projekt nicht vollständig nach Plan durchgeführt und die bewilligten Mittel nicht vollständig abgerufen werden.

# 5 Darstellung des wissenschaftlich-technischen und wirtschaftlichen Nutzens der erzielten Ergebnisse insbesondere für KMU sowie ihres innovativen Beitrags und ihrer industriellen Anwendungsmöglichkeiten

KMU werden zurecht als Rückgrat der deutschen Wirtschaft bezeichnet – dies belegen wirtschaftliche Kennzahlen. So wird ein Anteil von 99,6 % der deutschen Unternehmen von KMU repräsentiert, in denen insgesamt 57,9 % der Arbeitnehmer beschäftigt sind. Auch im Bereich der sogenannten „Hidden Champions“ stellen deutsche KMU mit 1.307 von 2.734 rund 48 % der weltweit identifizierten. Dieser Erfolg lässt sich vorrangig auf die Tatsache zurückführen, dass KMU sich auf die Produktion von Produkten spezialisiert haben, die in ihrem jeweiligen Markt ein Alleinstellungsmerkmal darstellen. Dieses stellt dabei oft das Angebot komplexer, kundenindividueller Produkte dar, welche sich nicht durch die Massenproduktion in lohngünstigeren Ländern realisieren lassen. Die kundenindividuelle Produktion bringt allerdings in logischer Konsequenz auch Herausforderungen wie eine höhere Produktvarianz und kürzere Produktlebenszyklen mit sich. Dies gilt vor allem in der Montage, da die Variantenbildung meist in der Wertschöpfungskette möglichst weit nach hinten verlagert wird, um dessen Einfluss zu minimieren.

Eine große Chance diese Komplexität zu beherrschen ist die Digitalisierung. Auch wenn die deutschen KMU die strategische Bedeutung der Digitalisierung erkannt haben, existiert noch ein großes Defizit in der Umsetzung. So erreichen KMU in 2018 im Schnitt nur 5 von 100 Punkten im Digital Index. Im Vergleich dazu erreichen die großen Unternehmen im Schnitt 24 Punkte. Laut einer Studie der Telekom AG hat die aktuelle Krise durch COVID-19 zwar die hohe Relevanz der Digitalisierung nochmals verdeutlicht, ist aber gleichzeitig ein starker Dämpfer für die Investitionsbereitschaft in digitale Technologien. Zwei der größten Strömungen der Digitalisierung sind die Robotik und die KI. Während bisher im Schnitt nur etwa 10 % der Produktionsaufgaben durch den Einsatz von Robotern automatisiert sind, wird ein Anstieg auf 25 % bis 45 % bis 2030 prognostiziert. Ähnlich sieht es bei KI aus, denn erst etwa 10% produzieren-der Unternehmen haben überhaupt eine Anwendung mit KI im Einsatz.

Die Einführung der in ROOKIE erarbeiteten, fortschrittlicher Robotik und KI kann dabei die Produktivität in vielen Branchen um 30 % steigern und gleichzeitig die Lohnkosten um 18-33 % senken. Gerade die Kombination der Technologien in fortschrittlichen Robotersystemen kann Einsparungen von bis zu 40% ermöglichen. Diese Effizienzsteigerung und Kostensenkung ist in vielen Fällen notwendiger Bestandteil, um unter Sicherung des Produktionsstandorts Deutschland konkurrenzfähig bleiben zu können. Darüber hinaus gilt es im Hinblick auf den demographischen Wandel und den drohenden Fachkräftemangel die hohe Abhängigkeit der Produktivität deutscher KMU ein Stück weit von der Expertise von Fachkräften aufzulösen.

Aufgrund dieser Entwicklungen und Argumente ist es von fundamentaler Bedeutung deutsche KMU dazu zu befähigen, KI-gestützte und roboterbasierte Automationslösungen einführen zu können. Vor allem Montageprozesse bieten hier aufgrund des noch sehr geringen Automationsgrads große unausgeschöpfte Potentiale. Dennoch existieren in der Industrie bisher kaum Montageapplikationen, bei denen künstliche Intelligenz mit Robotern kombiniert wird. Dies liegt an der fehlenden Kompetenz von KMU in der Roboterprogrammierung, fehlender Erfahrung und Strategien in der Einführung und Anwendung von künstlicher Intelligenz, mangelnder Planungsunterstützung und einer häufig

ungenügenden Datenqualität und -quantität. Abschließend gilt es Wege zu finde, Expertenwissen in KI-basierte Systeme zu integrieren um die Abhängigkeit von Fachkräften zu reduzieren.

Im Zuge von ROOKIE wurde auf die Herausforderungen reagiert und eine holistische, KI- und roboterbasierte Montage-Methodik entwickelt, mit deren Hilfe die genannten Einführungshürden reduziert werden können. Durch die Kombination aktueller KI-Verfahren mit einer ontologiebasierten Modellierung wurde die Flexibilität gegenüber der angesprochenen Produktvarianz von KMU sowie die Flexibilität, Autonomie und Robustheit der Planung, Inbetriebnahme und Durchführung von roboterbasierten Montageoperationen verbessert. Außerdem wurde über Benutzerschnittstellen Expertenwissen in das System integriert, was die Industrietauglichkeit weiter verbessert. Abschließend wurden die entwickelten KI-Skills vorrangig mit synthetischen Daten trainiert, was die Möglichkeit eröffnet trotz geringer Losgrößen die benötigten Datensätze zu erzeugen. Zusammenfassend werden KMU durch die Methodik in die Lage versetzt, von den Vorteilen der KI-basierten Robotik profitieren zu können, um den Fertigungsstandort Deutschland weiter zu festigen.

# 6 Wissenstransfer in die Wirtschaft

Das Ziel des Projekts, KMU den Zugang zu KI- und roboterbasierten Prozessen zu ermöglichen, wurde erreicht. Dazu gehört die Entwicklung neuer, KI-basierter Methoden und Strate-gien zur Senkung von Kompetenzhürden sowie die Steigerung der Flexibilität und Robustheit von Montageprozessen. Letztere soll durch eine teilautonome Planung und Programmierung sowie eine KI-gestützte Durchführung der Montageoperationen er-reicht werden. Mit Hilfe des Demonstrators werden die entwickelten Ansätze in bestehenden Anwendungsszenarien validiert und die Methoden und Werkzeuge aus den beschriebenen AP industrienah erprobt. Der Demonstrator soll die beteiligten Forschungsstellen und Unternehmen dazu befähigen, die gewonnenen Erkenntnisse praktisch umzusetzen. Unternehmen des PA können am Demonstrator sowohl die Montage von Beispielbauteilen validieren. Mit dem Demonstrator am WZL wird gerade KMU mit begrenzten finanziellen Mitteln die Chance geboten, die entwickelten Ergeb-nisse zeit- und praxisnah kennenzulernen.

Ergebnisse des Forschungsvorhabens können projektbegleitend und darüber hinaus genutzt werden und einen Beitrag zur Entwicklung von Leitfäden und Einführungsstrategien leisten. Das Entwickeln solcher Strategien für die wirtschaftliche, robuste und industrietaugliche Umsetzung von KI-basierten Prozessen ist im Hinblick auf die aktuellen Entwicklungen im industriellen Umfeld von besonderem Interesse. Auch Unternehmen, die nicht Teil des PA sind, können die erarbeiteten Ergebnisse nutzen und als Basis für weitere Entwicklungen heranziehen.

Für die Mitglieder im PA kann der Ergebnistransfer wie folgt skizziert werden:

- Unternehmenskategorie Endanwender: Hierbei handelt es sich um Firmen aus dem produzierenden Gewerbe, bei denen die manuelle, teilautomatisierte oder automatisierte Montage zum täglichen Geschäft gehört oder welche an der Ein-führung der entwickelten Technologie interessiert sind. Eine Varianz im zu fer-tigenden Produkt ist dabei sichergestellt. Die Anwender werden durch die im

Projekt entwickelten Methoden und Ansätze in die Lage versetzt, roboterbasierte Montageprozesse durch den Einsatz von KI schneller und robuster in Be-trieb zu nehmen und sie darüber hinaus sicherer und intuitiver zu betreiben

- Unternehmenskategorie Experten für Robotik und Automatisierung: Diese Gruppe von Unternehmen wird durch die Forschungsergebnisse in die Lage versetzt, das bestehende Portfolio an Automatisierungslösungen zu erweitern und Gesamtlösungen für KI- und roboterbasierte Montageprozesse anbieten zu können. Durch ihre Fachkompetenz in der Implementierung von Automations-lösungen können diese Unternehmen vor allem im Zuge der Systemkonzeption und der Umsetzung eines sicheren Gesamtsystems unterstützen und außer-dem die Ergebnisse von ROOKIE als Multiplikatoren über das Projektende hin-aus verbreiten
- Unternehmenskategorie Technologieanbieter und IT-Unternehmen: Um eine effiziente Einbindung von Hardware- und Softwarekomponenten in die entwickelte Methodik sicherzustellen, sind auch Technologie- und IT-Anbieter Teil des PA. Diese können ROOKIE durch ihre Expertise fachlich beraten und profitieren selbst durch das Erschließen neuer, innovativer Anwendungsfelder für ihre Produkte. Außerdem können diese Unternehmen die Einhaltung von gängigen Industriestandards im Zuge von ROOKIE beurteilen ebenfalls als Multiplkatoren fungieren.
- Unternehmenskategorie Technologie- und Kompetenzzentren: Diese Verbünde bieten zahlreiche domänenübergreifende Transferkanäle zu deutschen KMU. Diese Kanäle erlauben sowohl einen Informationsfluss in das Projekt als auch einen Transfer der Ergebnisse nach draußen.

# 7 Plan zum Ergebnistransfer in die Wirtschaft

In Kapitel 7.1 und Kapitel 7.2 werden die bereits zur Projektlaufzeit durchgeführten und weiterhin geplanten Transfermaßnahmen nach Projektende erläutert.

## 7.1 Durchgeführte Transfermaßnahmen (vom Projektstart bis zum Projektende)

Während der Projektlaufzeit sind vier wissenschaftliche Veröffentlichungen erfolgt:

- Josefine Monnet, Oliver Petrovic, Christian Brecher. Investigating the generation of synthetic data for surface defect detection: A comparative analysis. In: 57th CIRP Conference on Manufacturing Systems 2024, S. 767–773
- Josefine Monnet, Emma Heyen, Oliver Petrovic, Werner Herfs. Leveraging reinforcement and curriculum learning for flexible robot-based snap-fit assembly automation. In: At-Automatisierungstechnik 2025 73(5), S. 331-340.
- Oliver Petrovic, David Leander Dias Duarte, Werner Herfs. Generating Synthetic Data Using a Knowledge-based Framework for Autonomous Productions. In: 2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), S. 1086-1093.
- Oliver Petrovic, David Leander Dias Duarte, Simon Storms, Werner Herfs. Towards Knowledge-based Generation of Synthetic Data by Taxonomizing Expert Knowledge in Production. In: Intelligent Human Systems Integration (IHSI) Vol. 69, 2023, S. 1-8.
- Oliver Petrovic, Petar Tesic, Christian Brecher. Sim2Real Framework for Smart Robotics in Assembly.In: Annals of Scientific Society for Assembly, Handling and Industrial Robotics 2024 (In Press)
- Oliver Petrovic, Yannick Dassen, Christian Brecher. Potenziale und Anwendungen des Industrial Metaverse. In: Industrie 4, 2023, S. 27-32.
- Oliver Petrovic, Josefine Monnet, Petar Tesic, Yannick Dassen, Werner Herfs. Potentials and Applications of the industrial metaverse using the example of synthetic data generation. In: Prinz, W., Trauth, D. (eds) Tokenizing the Future. Springer Cham.,2025, S. 423-435.
- David Kötter, Gabriel Wiedon, Daniel Meierkord, Minh Trinh, Oliver Petrovic, Christian Brecher. Development of an Augmented Reality User Interface for Collaborative Robotics in Quality Inspection for Manufacturing. In: 2023 the 5th International Conference on Control and Robotics. S. 107-112.
- David Kötter, Denis Teknetas, Melanie Kolditz, Oliver Petrovic, Christian Brecher. Flexible voice control for collaborative robotics. In: at-Automatisierungstechnik 73(5), 2025, S.341-351.
- David Kötter, Fabian Nolte, Oliver Petrovic, Christian Brecher. Status recognition system for collaborative robotics. In: wt-Werkstattstechnik Online, Band 115, Nr. 7-8, 2025, S.632-636.

Darüber hinaus wurden die in Tabelle 3 dargelegten Transfermaßnahmen geleistet.

**Tabelle 3: Durchgeführte Transfermaßnahmen**

| Transfermaßnahme | Datum |
|---|---|
| Erstellung und Pflege einer Projekt-Website | 01.02.2023 bis 31.07.2025 |
| Treffen des projektbegleitenden Ausschusses | 16.02.2023 |
| Wöchentliche digitale Showroom Events in Aachen (FVP) | 01.02.2023 bis 31.12.2023 |
| Vorstellung der Projektergebnisse in den Vorlesungen „Mechatronik und Steuerungstechnik für Produktionsanlagen“, „NC-Programmierung von Werkzeugmaschinen“ und „Automatisierungstechnik für Produktionssysteme“ des WZL | 04/2023 bis 05/2025 |
| Treffen des projektbegleitenden Ausschusses | 10.05.2023 |
| Vorstellung des Projektes und der erzielten Ergebnisse auf dem General Meeting der EXAPT Systemtechnik GmbH in Wiesbaden | 03.05.2023 bis 04.05.2023 |
| Vorstellung des Projekts und der erzielten Projektergebnisse im Rahmen des Aachener Werkzeugmaschinenkolloquiums AWK 2023 in Aachen | 11.05.2023 bis 12.05.2023 |
| Treffen des projektbegleitenden Ausschusses | 13.09.2023 |
| Vorstellung der Projektergebnisse auf der Weltleitmesse für Produktionstechnologie EMO 2023 in Hannover | 18.09.2023 bis 23.09.2023 |
| Treffen des projektbegleitenden Ausschusses | 04.12.2023 |
| Vorstellung der Ergebnisse auf der Mitgliederversammlung FVP in Aachen | 21.02.2024 bis 22.02.2024 |
| Vorstellung des Projektes und der erzielten Ergebnisse auf der Hauptversammlung der EXAPT Systemtechnik GmbH in Aachen | 15.05.2024 |
| Treffen des projektbegleitenden Ausschusses | 12.06.2024 |
| Vorstellung der Ergebnisse auf der Internationalen Messe für Metallbearbeitung AMB 2024 in Stuttgart | 10.09.2024 bis 14.09.2024 |
| Treffen des projektbegleitenden Ausschusses | 09.10.2024 |
| Treffen des projektbegleitenden Ausschusses | 22.01.2024 |
| Vorstellung der Ergebnisse auf der Mitgliederversammlung FVP in Aachen | 08.04.2025 bis 09.04.2025 |
| Vorstellung des Projektes und der erzielten Ergebnisse auf der Hauptversammlung der EXAPT Systemtechnik GmbH in Aache | 14.05.2025 bis 15.05.2025 |
| Treffen des projektbegleitenden Ausschusses | 26.06.2025 |

## 7.2 Geplante Transfermaßnahmen (nach Projektende)

Es sollen in Zukunft sowohl öffentliche Förderformate als auch bilaterale Industrieprojekte an dieses Projekt anschließen. Der Demonstrator dient dabei als greifbares Transferobjekt, an dem die Funktionsweise des Frameworks sowie die Integration in industrielle Abläufe gezeigt werden kann. Aufbauend auf den im Projekt erzielten Ergebnissen werden aktiv Folgeprojekte angestrebt, um (i) die Integration in reale KMU-Umgebungen zu vertiefen, (ii) die Bedienbarkeit weiter zu, (iii) Safety-/Security-Aspekte ergänzen und (iv) Planungs- und Optimierungsfunktionen unter realen Datenbedingungen zu validieren.

Am 29. und 30. April 2027 findet das Aachener Werkzeugmaschinen-Kolloquium (AWK) statt. Unter dem Leitthema „Shaping Future Production“ werden einem breiten Publikum u. a. die Inhalte und Ergebnisse des ROOKIE-Projektes vorgestellt. Der enge fachliche Austausch und die Diskussion mit zahlreichen Anwendern sowie ausgewiesenen Experten anderer Forschungsinstitute werden dabei neue Impulse liefern. Darüber hinaus sind seitens WZL und FVP weitere Transfermaßnahmen nach Projektende geplant. Einen Überblick hierzu liefert Tabelle 4.

**Tabelle 4: Geplante Transfermaßnahmen**

| Transfermaßnahme | Datum |
|---|---|
| Vorstellung der Projektergebnisse auf der Weltleitmesse für Produktionstechnologie EMO 2025 in Hannover | 22.09.2025 bis 26.09.2025 |
| Vorstellung des Projektes und einer Ergebnisse im WZL-Jahresbericht 2026 | 01/2026 |
| Veröffentlichung des Schlussberichts auf den Websites von WZL und FVP | 03/2026 |
| Vorstellung der Projektergebnisse in den Vorlesungen „Mechatronik und Steuerungstechnik für Produktionsanlagen“, „NC-Programmierung von Werkzeugmaschinen“ und „Automatisierungstechnik für Produktionssysteme“ des WZL | 04/2026 bis 11/2026 |
| Vorstellung des Projekts und der erzielten Projektergebnisse im Rahmen des Aachener Werkzeugmaschinenkolloquiums AWK 2027 in Aachen | 29.04.2027 bis 30.04.2027 |
| Vorstellung der Projektergebnisse auf der Messe „Toolmanagement in der Praxis 2026“ in Böblingen | 04.03.2026 bis 05.03.2026 |
| Vorstellung der Ergebnisse auf der Mitgliederversammlung FVP in Aachen | 06.05.2026 bis 07.05.2026 |
| Vorstellung des Projektes und der Projektergebnisse auf der Hauptversammlung und technischen Tagung der EXAPT Systemtechnik GmbH in Aachen | 06.05.2026 bis 07.05.2026 |
| Vorstellung der Ergebnisse auf der Internationalen Messe für Metallbearbeitung AMB 2026 in Stuttgart | 15.09.2026 bis 19.09.2026 |