



IIP-Ecosphere

Schlussbericht

Zuwendungsempfänger: Leibniz Universität Hannover, Marposs Monitoring Solutions GmbH, Lenze SE, Univ. Hildesheim, slashwhy GmbH & Co. KG, Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutsche Messe Technology Academy, Sennheiser electronic GmbH & Co. KG, Univ. Koblenz-Landau, Gildemeister Drehmaschinen GmbH, Siemens AG, Fraunhofer ISST, RapidMiner GmbH, Verein Deutscher Werkzeugmaschinenfabriken e.V., KIProtect GmbH, Volkswagen AG, Bitmotec GmbH	Förderkennzeichen: 01MK20006 A-Q
Vorhabenbezeichnung: Verbundprojekt: IIP-Ecosphere – Next Level Ecosphere for Intelligent Industrial Production	
Laufzeit des Vorhabens: Vom 01.01.2020 bis zum 30.09.2023	

Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages

Autoren (max 2 pro Partner)

Dr. Claudia Niederée, Dr. Holger Eichelberger, Per Schreiber, Amir Shayan Ahmadian, Dr. Alexander Broos, Dr. Andreas Dewes, Marco Ehl, André Heinke, Dr. Petra Hildebrandt, Nils Jahnke, Julius Kirschbaum, Ralf Klinkenberg, Prof. Dr.-Ing. Dirk Lange, Dr. Kai Litwinski, Dr. Gregory Palmer, Daniela Quardt, Jeff Reimer, Michael Schick-tanz, Hans-Dieter Schmees, Geeske Sieckman, Andrea Suckro, Dr. Heiko Stichweh, Jürgen Urdich, Matthäus Wilga, Rainer Wilkens, Peter Wißbrock;

Weitere Beitragende

Ahmad Alamoush, Daniel Arnold, Jonas Becker, Götz Görisch, Dr. Nils Haldenwang, Klaas Heide, Thomas Hildebrandt, Alexander Inglessi, Prof. Dr. Jan Jürjens, Daniel Kemp, Jens Königsmann, Sofiane Laridi, Prof. Dr. Angela Roth, Christian Severin Sauer, Annika Schumann, Timon Sengewald, Monika Staciwa, Tobias Stiehl, Alexander Weber, Sören Wilmsmeier, Martin Winkler

1 Kurze Darstellung des Projektinhalts.....	4
1.1 Aufgabenstellung.....	4
1.2 Voraussetzungen des Vorhabens.....	4
1.3 Planung und Ablauf.....	5
1.4 Wissenschaftlicher und technischer Stand bei Antragstellung	7
1.5 Zusammenarbeit mit anderen Stellen	8
2 Eingehende Darstellung.....	10
2.1 Erzielte Ergebnisse	10
2.1.1 Initialisierung und Anforderungen (AP 1).....	10
2.1.2 Think Tank „Plattformen“ (AP 2).....	10
2.1.3 Think Tank „Daten“ (AP 3).....	34
2.1.4 Think Tank „KI & Produktion“ (AP 4)	47
2.1.5 Think Tank „Geschäftsmodelle“ (AP 5).....	63
2.1.6 KI-Accelerator (AP 6)	74
2.1.7 Demonstrator „Werkübergreifende Optimierung der Taktzeiten“ (AP 7).....	91
2.1.8 Demonstrator „KI-basierte Assistenzsysteme zur Zyklenuptimierung“ (AP 8) ..	98
2.1.9 Demonstrator „Autonome Roboterzellen zur Endprüfung in der Leiterplattenfertigung“ (AP 9).....	105
2.1.10 Demonstrator „KI-basierte Prozess-Stabilität in der Glasformung“ (AP 10)....	143
2.1.11 Exploration und Best Practice (AP 11).....	150
2.1.12 Innovation und Stakeholder-Einbindung (AP 12)	155
2.1.13 Dissemination/Sichtbarkeit (AP 13)	157
2.1.14 Verwertung (AP 14).....	162
2.1.15 Projektmanagement/Governance (AP 15)	164
2.2 Voraussichtlicher Nutzen und Verwertbarkeit	165
2.3 Relevante Fortschritte/Ergebnisse Dritter	167
2.4 Veröffentlichung der Ergebnisse.....	167
3 Literaturverzeichnis.....	176

1 Kurze Darstellung des Projektinhalts

1.1 Aufgabenstellung

Die Vision von IIP-Ecosphere ist ein Innovationssprung im Bereich der Selbstoptimierung der Produktion auf Basis vernetzter, intelligenter, autonomer Systeme zur Steigerung der Produktivität, Flexibilität, Robustheit und Effizienz. Ziel der Umsetzungsphase ist es daher, ein neuartiges Ökosystem – die Next Level Ecosphere for Intelligent Industrial Production (IIP-Ecosphere) – aufzubauen, das eine „nächste Ebene“ in der intelligenten Produktion ermöglicht. Zur Umsetzung dieser Vision sieht IIP-Ecosphere Aktivitäten vor, die die Anwendbarkeit von KI-Methoden in der intelligenten Produktion erleichtern und in realen Anwendungsszenarien demonstrieren. Die Aktivitäten zielen auf die Beseitigung der unten genannten Hemmnisse und fördern die nachhaltige Einbeziehung neuer Stakeholder. Hierbei sollen insbesondere KMU und Startups befähigt werden, KI-Methoden zur intelligenten Produktion selbst erfolgreich anzuwenden und weiterzuentwickeln.

1.2 Voraussetzungen des Vorhabens

Die Digitalisierung der Industrie sowie die zunehmende Vernetzung erhöhen die Leistungsfähigkeit und Komplexität technischer Systeme und der zugehörigen Prozesse. Im industriellen Umfeld wird der intelligenten Produktion (Industrie 4.0) ein disruptives Innovationspotenzial eingeräumt, wobei künstliche Intelligenz (KI) als Schlüsseltechnologie gilt. Studien prognostizieren durch den Einsatz von KI eine Steigerung der Produktivität um bis zu 50 % [BAU16]. Allein für den Maschinen- und Anlagenbau wird ein zusätzliches Marktpotenzial in Höhe von 99,8 Milliarden Euro gesehen [BIT19].

Der Ausbau der intelligenten Produktion erfordert den Einsatz von KI auf allen Ebenen des Produktentstehungsprozesses und über die Grenzen einzelner Unternehmen hinweg. Die Wissenschaftliche Gesellschaft für Produktionstechnik (WGP) sieht in diesem Kontext u. a. folgende Forschungsschwerpunkte: Übertragung cyberphysischer Systemarchitekturen in die Fertigung, Methoden zur Selbstoptimierung von Produktionsprozessen und Prozessparametern, KI-basierte Produktionssteuerung und -planung, Methoden zur dezentralen Analyse und Optimierung sowie zur automatisierten Entscheidung über die Informationsweitergabe [BAU16]. In 65 % der deutschen Unternehmen mit über 100 Mitarbeitenden kommen bislang (2019) allerdings noch keine KI-Methoden zum Einsatz [BMW19]. Besonders kleine und mittelständische Unternehmen (KMU) stehen vor Hemmnissen, die den Einsatz von KI-Methoden verhindern können:

- Fehlende technische Voraussetzungen (bspw. Sensorik, Cloud-Computing) [GAB18]
- Uneinheitliche Schnittstellen und Plattformen für unterschiedliche Maschinen, Steuerungen, Sensoren und Bearbeitungsprozesse [BIT19]
- Bedenken bezüglich Datenhoheit und Geheimhaltung [BMW19, GAB18, LIC19]
- Unpassendes Geschäftsmodell oder unklare Gewinnaussichten [MÜL16]
- Fehlende interne Kompetenzen im Umgang mit Big Data und KI [BIT19, GAB18, LIC18]

Daher ist es zwingend erforderlich, KMU den Zugang zu leistungsfähigen Methoden des Maschinellen Lernens (ML) zu erleichtern sowie geeignete Geschäftsmodelle aufzuzeigen und diese zu ermöglichen.

Dies und die Integration KI-basierter Komponenten in reale Prozessketten sowie deren Weiterentwicklung ist nur möglich, wenn die Randbedingungen in Form eines geeigneten Ökosystems geschaffen werden. Diese Randbedingungen sind beispielsweise einfache Verwendbarkeit, übergreifende Standardisierung sowie die Bereitstellung von Mehrwertdiensten mit geringerer Einstiegsschwelle.

1.3 Planung und Ablauf

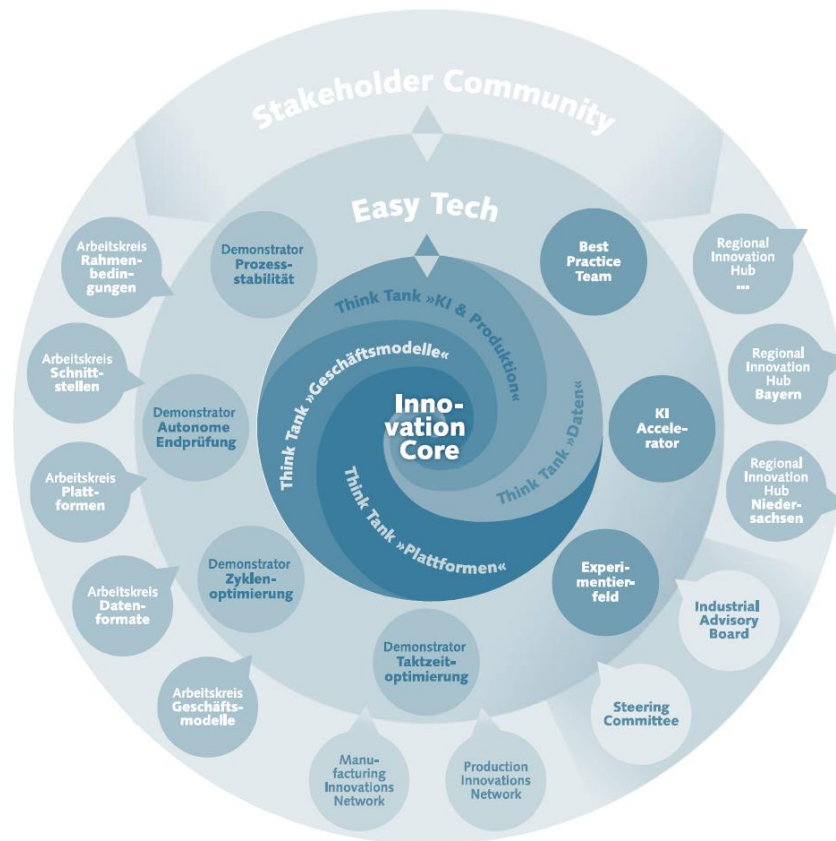


Abbildung 1: Umsetzungskomponenten der IIP-Ecosphere

Um die genannten Ziele zu erreichen, setzt sich IIP-Ecosphere aus komplementären Bausteinen zusammen, die in drei Schichten angeordnet sind (siehe Abbildung 1). Den Kern des Ökosystems bildet der **InnovationCore**, in dem die relevanten anwendungsorientierten Forschungsaktivitäten in vier **Think Tanks** (TT) gebündelt sind. Die zweite Schicht bilden die **EasyTech**-Komponenten, die die Aktivitäten und Angebote des Ökosystems für einen vereinfachten Zugang von KMU und anderen Unternehmen zu KI-Technologien bündeln. Hierzu gehören als wichtige Bausteine die **IIP-Demonstratoren** und der **KI-Accelerator**, der einen systematischen Zugang zu KI-Projekten in der intelligenten Produktion fördert. Ebenfalls Teil der EasyTech-Schicht sind das **Experimentierfeld**, welches niedrigschwellig das Experimentieren neuer KI-Technologien erlaubt, und das **Best Practice Team**. Die äußere Schicht der Ecosphere (**Stakeholder Community**) bilden Aktivitäten, die die Interaktion mit den unterschiedlichen Stakeholder-Gruppen und insbesondere auch mit Akteuren außerhalb des IIP-Ecosphere Konsortiums fördern. Hierzu zählen zum einen die **Arbeitskreise** (AK) zur Diskussion und Gestaltung relevanter Themen für KI in der Produktion mit relevanten Akteuren und Multiplikatoren. Zum anderen sind **Regional Innovation Hubs** (RIH) geplant, welche regionale Akteure der Innovationsförderung zum

Thema „KI für die intelligente Produktion“ vernetzen, um Technologie-Transfer, Innovation und Gründungsaktivitäten in diesem Umfeld durch gemeinsame Aktivitäten voranzubringen.

Die Arbeit im Projekt ist in vier Phasen gegliedert: In der **Initialisierungsphase** (3 Monate) werden die Anforderungen präzisiert, die Prozesse für die Zusammenarbeit und die Kommunikationskanäle im Projekt festgelegt und erste Veranstaltungen zur Interaktion mit den Stakeholdern organisiert. In der zweiten Phase – der **Aufbauphase** – wird in allen Bereichen des Projekts der Aufbau der Ecosphere gezielt und koordiniert vorangetrieben. Dabei werden unter anderem erste Forschungs- und Entwicklungsergebnisse aus den Think Tanks erzielt, der Lösungskatalog im KI-Accelerator gefüllt, die Demonstratoren weiterentwickelt sowie mit der Umsetzung der virtuellen Plattform begonnen. Zugleich werden verstärkt Aktivitäten für das Wachstum der IIP-Ecosphere gestartet und die Arbeit in den Arbeitskreisen und im Experimentierfeld aufgenommen.

Als Mittel der Erfolgskontrolle und der anwendungsorientierten Steuerung des Projekts sind **jährliche Reviews** mit dem Industrial Advisory Board (IAB) eingeplant, in denen die im jeweiligen Jahr erreichten Ergebnisse der Umsetzungsphase auf den Prüfstand gestellt werden. Das erste Review bildet zugleich den Startschuss zur dritten Phase, der **Erweiterungs- und Konsolidierungsphase**. Diese Phase widmet sich der Konsolidierung der bisherigen Ergebnisse, insbesondere auch auf der Basis der Review-Ergebnisse und der systematischen Erweiterung. Dabei wird neben der Erarbeitung weiterer Forschungs- und Entwicklungsergebnissen insbesondere die Nutzung der Ergebnisse in weiteren Kontexten vorbereitet, z. B. im KI-Accelerator durch Stabilisierung, Umsetzung in Form von Standardkomponenten und geeignete Parametrisierung. Die Ergebnisse werden in die Demonstratoren zurückgespiegelt und deren Anwendung im Kontext der virtuellen Plattform erprobt und validiert. Verstärkt werden auch Best Practices aus den Erfahrungen im Projekt und aus Interaktion mit der wachsenden Anzahl von Akteuren in der Ecosphere gesammelt und dokumentiert.

In der letzten Phase des Projekts - der **Mehrwertphase** - wird besonderer Wert auf die Vorbereitung der Ecosphere auf den langfristigen nachhaltigen Betrieb gelegt. In dieser Phase wird insbesondere der Mehrwert durch Dienste und Angebote, den die Ecosphere für ihre Zielgruppen anbietet, weiter ausgebaut. Dies dient dazu, die Attraktivität der Ecosphere weiter zu steigern und einen Übergang in Verstärkung vorzubereiten.

Thematisch ist die Arbeit in 15 Arbeitspakete aufgeteilt. Jedes Arbeitspaket wird von einem Partner im Kernkonsortium geleitet, der den Arbeitspaketleiter stellt. Das **Arbeitspaket Initialisierung/Anforderungen** ist insbesondere in der Initialisierungsphase aktiv. In diesem Arbeitspaket führen alle Partner eine gemeinsame Konzeptionierung der durchzuführenden Arbeiten durch und leiten grundlegende Anforderungen für alle weiteren Arbeitspakete ab. Die vier Think Tanks des Bereichs Innovation Core, werden durch vier korrespondierende **Forschungs-Think-Tank-Arbeitspakete** repräsentiert. Der Bereich EasyTech wird durch mehrere Arbeitspakete abgebildet: Im **Arbeitspaket KI-Accelerator** arbeiten Industriepartner am KI-Lösungskatalog sowie an Standardkomponenten und unterstützen die Demonstratoren. Die vier Demonstratoren werden in vier zugehörigen **Demonstrator-Arbeitspaketen** entwickelt, die insbesondere in der Initialisierungs- sowie in der Erweiterungs- und Konsolidierungsphase aktiv sind. Weiterhin gehört das **Arbeitspaket Exploration und Best Practices** zum Bereich EasyTech. In diesem Arbeitspaket werden die dynamischen Demonstratoren durchgeführt, das Experimentierfeld aufgebaut und verwaltet sowie Best Practices aus den gewonnenen Erfahrungen abgeleitet

Aktivitäten rund um die regionalen Innovation Hubs, Arbeitskreise und die für IIP-Ecosphere geplanten öffentlichen Events werden vom **Arbeitspaket Innovation und Stakeholder-Einbindung** gebündelt. Neben den genannten inhaltlichen Arbeitspaketen sind noch die folgenden begleitenden Arbeitspakete geplant: Das **Arbeitspaket Dissemination/Sichtbarkeit** erstellt und verwaltet das Internetportal für die Ecosphere, stellt die Projektergebnisse der interessierten Öffentlichkeit vor und führt die Öffentlichkeitsarbeit von IIP-Ecosphere zum Auf- und Ausbau einer Community durch. Das **Arbeitspaket Verwertung** führt Aktivitäten zur Vorbereitung der Verwertung der Projektergebnisse durch, insbesondere in Form von Standardisierung Vorstudien zur Übernahme von Ergebnissen in Produkte sowie Erstellung von wissenschaftlichen und technischen Veröffentlichungen. Das **Arbeitspaket Projektmanagement/Governance** beinhaltet das Gesamtprojektmanagement inklusive Steering Committee und Industrial Advisory Board, exploriert Geschäftsmodelle zum nachhaltigen Bestand der Ecosphere und führt Aktivitäten zum eigenständigen Betrieb nach der Umsetzungsphase durch.

1.4 Wissenschaftlicher und technischer Stand bei Antragstellung

Basierend auf der kontinuierlichen Steigerung der Rechenleistung bei gleichzeitiger Kostenminderung und der Verfügbarkeit großer Mengen annotierter Daten wird verstärkt Forschung im Bereich der „Neuronale Netze“ und des „Deep Learning“ betrieben. Diese Verfahren bieten die Möglichkeit komplexe Entscheidungsfunktionen zu lernen [GOD09] und erzielen unter anderem in der Signalverarbeitung [WAN14, YU11] und Informationsbeschaffung [XU18] bereits hervorragende Ergebnisse, die menschliche Leistung immer öfter übertreffen [GEI18, WIC17]. Aktuelle Arbeiten zielen unter anderem auf für Menschen verständliche Erklärbarkeit [ANA18, WU18] der maschinellen Entscheidungen ab und steigern damit die Bedienbarkeit und Einsetzbarkeit der Modelle.

Im Rahmen des Megatrends Industrie 4.0 [BMW16] wird das Ziel verfolgt, physische Produktionssysteme mittels der Integration von digitalen Komponenten zu optimieren. Die objektbezogene Kopplung der physischen Komponenten einer Fertigung mit der softwarebasierten Entscheidungsfindung ist auch als cyber-physisches System (CPS) beziehungsweise cyber-physisches Produktionssystem (CPPS) bekannt [BAU16, MON16]. Diese Systeme sind durch eine hohe digitale Konnektivität und autonome Entscheidungskompetenz von physischen Komponenten innerhalb einer digitalisierten Fertigung gekennzeichnet [KAG13]. Durch die digitale Datenerfassung können große Datenmengen erfasst werden, die einerseits ein großes Optimierungspotenzial bieten. Zur Optimierung, Vorhersage oder Steuerung von Bearbeitungsprozessen wurden daher bereits viele auf KI-Prozessmodellen basierende Überwachungssysteme für die intelligente Produktion entwickelt und Datenanalysen oder Datenbereinigung vorgenommen [ABE10]. Herausfordernd ist dabei die Gewährleistung einer stetigen Aktualisierung der Trainingsdaten, um die Klassifikationsgüte bei einem zeitvarianten Verhalten von Maschinen zu erhalten [AMI19]. Ferner stellt die Datenqualität in Produktionsprozessen einen limitierenden Faktor hinsichtlich der Anwendung von KI-Modellen in der Produktion dar [SAH14]. Zusätzlich benötigen die datengetriebenen Ansätze eine große Anzahl an Stichproben, die im Bereich der Produktion, insbesondere für fehlerhafte Prozesse, häufig nicht zur Verfügung stehen. Um diesen Herausforderungen zu begegnen, muss ein verstärkter Datenaustausch sowie die Verfügbarkeit öffentlicher Datenquellen gefördert werden [SEV16].

Im industriellen Umfeld gibt es eine Reihe von Anwendungsfällen für KI-Methoden. Die prominentesten sind datenbasierte Methoden für die vorausschauende Instandhaltung [MOB02]. Durch sie können Betriebskosten der Produktion gesenkt und die Zuverlässigkeit von unternehmenskritischen Systemen erhöht werden [HAS11]. In diesem Zusammenhang spielt der Einsatz maschineller Lernverfahren eine entscheidende Rolle [ABD10, JUN17, SUS15]. Weitere wichtige KI-Anwendungsgebiete sind die Qualitätskontrolle und Qualitätsregelkreise. Aus der Verfügbarkeit zahlreicher Sensormessungen ergibt sich großes Potenzial für eine automatisierte Qualitätskontrolle [MIT15], wobei trotz der zunehmenden Verfügbarkeit automatisierter datengetriebener Methoden die Qualitätskontrolle nach wie vor häufig noch abhängig von der Expertise der Mitarbeiter ist [UHL17]. Im Bereich der optischen Qualitätskontrolle sind Deep Learning Ansätze sehr vielversprechend und auch bei hohen Stückzahlen einsetzbar [HUA15]. Auch im Bereich Arbeitsplanung und Fertigungssteuerung werden datengetriebenen Methoden bereits vielseitig eingesetzt, um z. B. Vorgabezeiten und Produktionskosten zu prognostizieren [EVE02, WES06, WIE09]. Methoden der Künstlichen Intelligenz sowohl prozessnah als auch übergeordnet auf planerischer Ebene angewendet. So erforschen aktuelle Arbeiten den Einsatz digitaler Schatten, die als prozessparallele Simulationen eingesetzt werden, um eine erweiterte Prozessüberwachung [ALT17] sowie eine systematische Prozessoptimierung mittels Künstlicher Intelligenz zu realisieren [DIT19].

Industrie 4.0 Produktionssysteme verwenden üblicherweise Software-Plattformen zur Integration von Daten und zur Vereinheitlichung der Schnittstellen für höhere Ebenen der Automatisierungspyramide (DIN EN 62264-1, IEC 62264, ISA-88.01, ISA-95). Die Verwendung von KI-Methoden in derartigen Produktionssystemen erfordert, dass die entsprechenden Software-Plattformen auf den verschiedenen Ebenen Integrationsmöglichkeiten bieten. Leider ist die Landschaft aktueller Industrie 4.0 Plattformen sehr divers. Arbeiten wie [ADA19, BMW19, DIN18, LIC19] beschreiben Hemmnisse wie Herstellerabhängigkeit (Vendor Lock-In), unterschiedliche Terminologien, unzureichende Offenheit, oft fehlende Verwendung offener Standards oder Komponenten, erzwungene Cloud-Nutzung, fehlende Sicherheitsmechanismen und zu komplexe aber auch nicht konsistente Konfigurationsvorgänge. Zudem sind Architekturen und Schnittstellen sehr heterogen [TRA17] und oft nicht an aktuelle Standards wie RAMI4.0 (DIN SPEC 91345, IEC PAS 63088) angelehnt. Wie eine systematische Untersuchung der für dieses Projekt relevantesten IIoT-Plattformen in der Wettbewerbsphase ergeben hat, bieten die meisten Plattformen zudem keine (hinreichenden) Schnittstellen und sind auch nicht auf dynamische Änderungen der Produktionsprozesse, insbesondere nicht auf autonome Selbstoptimierung vorbereitet.

1.5 Zusammenarbeit mit anderen Stellen

Für die Umsetzung hat IIP-Ecosphere ein Konsortium von Partnern aus Wirtschaft und Wissenschaft gewinnen können. Dies sind Volkswagen Nutzfahrzeuge (VWN), Siemens, Sennheiser, Gildemeister, Artis Marposs, Gerresheimer, RapidMiner, KIPROTECT, Lenze, slashwhy und Bitmotec, die mit erheblichem Eigenanteilen teilnehmen. Vervollständigt wird das Konsortium durch die Multiplikatoren Deutsche Messe Technology Academy (DMTA) und den Verein der deutschen Werkzeugmaschinenfabriken (VDW). Forschungsseitig nehmen die Institute und Arbeitsgruppen der Leibniz Universität Hannover (LUH, hier das L3S mit den Gruppen L3S-KBS, L3S-IRI und L3S-TNT und das Institut für Fertigungstechnik und Werkzeugmaschinen IFW), die Universität Hildesheim (UHi), die Friedrich-Alexander Universität Erlangen-Nürnberg (FAU), die Universität Koblenz-Landau (UKL) und das Fraunhofer Institut für Software- und Systemtechnik (ISST) teil.

Zudem sind 57 Partner (darunter 13 KMU, 6 Startups und 11 Multiplikatoren wie LNI, MIN) assoziiert. Im Bereich der Innovation Hubs wurde besonders mit den assoziierten Multiplikatoren kooperiert. Dazu gehörten beispielsweise die IHKS in Nürnberg und Hannover oder Hannover Impuls. Mit einer Reihe von assoziierten Partnern wurde eine intensive fachliche Zusammenarbeit initiiert. Beispielhaft sind hier die Fa. Phoenix Contact und das LNI 4.0 im Bereich der Plattformentwicklung zu nennen.

Im Rahmen eines Ideenwettbewerbes (siehe Abschnitt 1.6.11) wurde das Projekt um vier weitere „Dynamische Demonstratoren“ erweitert. Für diese KI-Kleinprojekte erfolgte eine Zusammenarbeit mit den Fa. Novoa, Koehn AI, Synthetic Images und MIP-Technology.

Darüber hinaus fanden zu dem Projekt zahlreiche Interaktionen mit Firmen und Organisationen außerhalb des Projekts statt. Diese umfassen die Darstellung des Projekts, die Diskussion möglicher Teilnahme im Projekt, und die Diskussion von Anforderungen zur Verfeinerung des Konzepts. Insbesondere wurde die Zusammenarbeit mit externen im Rahmen von Arbeitskreisen gebündelt, die wichtige Impulse für die Arbeit in den Projektteilen geliefert haben.

Zudem gaben die von IIP-Ecosphere organisierten Veranstaltungen (siehe auch Abschnitt 1.6.13) Gelegenheit für den Austausch und die Vernetzung mit weiteren Stakeholdern des Ökosystems. Aus diesem Austausch wurden wichtige Erkenntnisse für die Ausgestaltung der Umsetzungsphase gewonnen z. B. zu Fragen des Data Sharing und zur Erleichterung des Technologiezugangs. Weiterhin hat das IIP-Ecosphere Team an zahlreichen Veranstaltungen teilgenommen, wie etwa an der Hannover Messe 2022 und 2023 sowie an der Messe EMO 2023. Dabei konnte die Sichtbarkeit des Projekts erhöht werden, sowie neue Kontakte geknüpft und Projektideen diskutiert werden.

2 Eingehende Darstellung

2.1 Erzielte Ergebnisse

2.1.1 Initialisierung und Anforderungen (AP 1)

In AP 1, „Initialisierung /Anforderung“ wurden in Zusammenarbeit mit AP 15 die Strukturen und Kanäle im Projekt aufgesetzt. Details hierzu siehe AP 15. Zur Stärkung der gemeinsamen Vision wurde ein Kickoff-Meeting durchgeführt. In den einzelnen Projektbereichen wurden die Anforderungen geschärft. Anforderungen wurden im Rahmen eines agilen Vorgehens fortlaufend den Erfahrungen und sich weiterentwickelnden Erfordernissen angepasst.

2.1.2 Think Tank „Plattformen“ (AP 2)

Die Arbeitsaufteilung im Think Tank „Plattformen“ entsprach im Wesentlichen der Projektplanung. Zudem wurden Zusammenarbeiten mit anderen Projektteilen aufgebaut, insbesondere durch regelmäßige Treffen mit den KI-Accelerator (AP 6) und dem Demonstrator zur Leiterplatten-Endprüfung (AP 9). Durch diese enge Zusammenarbeit ergaben sich insbesondere gemeinsame Arbeiten mit Partnern aus AP6, die thematisch AP 2 bzw. AP 6 und teilweise auch beiden Arbeitspaketen zuzuordnen sind. Daher benennen wir insbesondere in diesen Fällen die involvierten Partner die laut Projektplanung ggf. nicht an AP 2 beteiligt waren.

Die Ergebnisse der Zwischenschritte wurden in regelmäßigen Treffen mit den am Think Tank beteiligten Projektpartner besprochen. Dabei waren oft nicht nur „offizielle“ AP 2 Partner zugegen, sondern im Regelfall auch interessierte Partner aus dem KI-Accelerator (AP 6) oder aus dem Demonstrator zur Leiterplatten-Endprüfung (AP 9).

State-of-the-Art Plattformen (UAP 2.1)

Arbeitsschritt 2.1.1: State of the Art Plattformen

Als Basis für die Plattform-Arbeiten wurde in diesem Arbeitsschritt die Hersteller-Dokumentation von 21 industriellen IIoT/Industrie4.0-Plattformen (inklusive Siemens Mindsphere, AWS IoT, GE Predics, PTC ThinkWorkx) in 16 Dimensionen analysiert, darunter angebotene Protokolle, KI-Integration, Edge-Nutzung, Nutzung digitaler Zwillinge, Offenheit, Sicherheit Konfigurierbarkeit. Die Dimensionen für diese Arbeit entstammen der Plattform-Visionsbildung, die in intensiven Diskussionen im KI-Accelerator (AP6) erarbeitet wurde. Die Durchführung der Analyse und Aufarbeitung der 21 industriellen Plattformen wurde gemeinsam geplant, im Wesentlichen von AP 2 durchgeführt und gemeinsam besprochen. Die Ergebnisse wurden im IIP-Ecosphere Whitepaper [SAU21] zusammengefasst, diskutiert und gegen die Vision der IIP-Ecosphere Plattform abgegrenzt.

Parallel wurden Arbeiten zu den KI-Architekturen der großen Plattformen wie AWS IoT, Google IoT Code und Siemens Mindsphere sowie ein wissenschaftliches Systematic Literature Review zum Thema Industrie 4.0 Plattformen (mit den gleichen Dimensionen wie [SAU21]) gestartet. Die zugehörigen Veröffentlichungsarbeiten werden nach dem Projektende weitergeführt.

Arbeitsschritt 2.1.2: State of the Art Sicherheit

In diesem Arbeitsschritt wurde, gemeinsam mit Arbeitsschritt 2.1.1, der Stand von Wissenschaft und Technik für Sicherheits- und Datenzugriffsmechanismen für Plattformen / deren Realisierung auch jenseits IIoT erarbeitet und analysiert. Die Ergebnisse wurden als Dimensionen Sicherheit und Datenschutz im IIP-Ecosphere Plattform-Whitepaper [SAU2021] diskutiert und veröffentlicht.

Arbeitsschritt 2.1.3: Technology Watch

Als Grundlage zur praktischen Umsetzung der IIP-Ecosphere Plattform wurden relevante Technologien identifiziert und Details zu diesen Technologien gesammelt. Dazu zählen insbesondere Projekte aus dem Eclipse IoT Ökosystem (Eclipse Milo, Eclipse Paho, BaSyx), Spring (Cloud Stream), Docker, Linux Container (LXC), RabbitMQ, HThingsBoard oder Prometheus. Für diese Technologien wurden regelmäßige Überwachungsaktivitäten gestartet um zu bewerten, ob Updates der jeweiligen Versionen im Laufe der Umsetzungsarbeiten möglich, sinnvoll oder sogar gefährlich für die Realisierung der IIP-Ecosphere Plattform sind. Dazu gehören ebenso Entwicklungen der zugehörigen Standards aus technischer Sicht bzw. aus dem Blickwinkel der technischen Integration, z. B. bei BaSyx die aktuellen Standardisierungs-Aktivitäten der IDTA für Industrie 4.0 Verwaltungsschalen.

Aus diesen Aktivitäten wurden letztlich mehrere Updates der Plattform-Integrationen von Eclipse BaSyx, Eclipse Milo and Eclipse Paho angestoßen und durchgeführt, während beispielsweise Updates des Stromprozessierungsframeworks Spring Cloud Stream aus Risiko- bzw. Kompatibilitätsgründen nicht durchgeführt wurden. Aus ähnlichen Erwägungen heraus wurde EASy-Producer, das zentrale Konfigurations-Modellierungs- und Instantiierungswerkzeug der Plattform einmal zu Beginn der Implementierungstätigkeiten auf den neuesten Stand gebracht und im Verlaufe von IIP-Ecosphere nur zur Ergänzung weiterer Funktionalität und zur Fehlerbeseitigung aktualisiert und jeweils mit der IIP-Ecosphere-Plattform als Release freigegeben.

Arbeitsschritt 2.1.4: Einarbeitung geänderter Anforderungen

Die Grundlagenarbeiten aus den Arbeitsschritten 2.1.1 und 2.1.2 bilden ein solides Fundament für die Planung, das Design und die Umsetzung der IIP-Ecosphere Plattform. Die Anpassungen, die durch die Technologieüberwachung in Arbeitsschritt 2.1.3 angestoßen wurden, führten zwar zu technologischen Anpassungen oder zur Änderung interner Schnittstellen aber zu keinen neuen Anforderungen.

Im Laufe der weiteren Diskussionen im Projekt, der Diskussionen mit assoziierten und externen Partnern sowie den Arbeiten an den IIP-Ecosphere Demonstratoren ergaben sich über die Projektlaufzeit einige Verschiebungen von Schwerpunkten sowie neue Ideen für Plattformfunktionalitäten. Ursprünglich war geplant, diese neuen Anforderungen in die initiale Anforderungsdokumentation [STI21] [EIC21a] einzuarbeiten. Da allerdings die verschiedenen Versionen des Plattformhandbuchs [EIC20, EIC21b, EIC22d, EIC23a, EIC23c] die Release-Versionen der Plattform dokumentieren, wurden geänderte Anforderungen, angepasste Konzepte sowie der Umsetzungsgrad der einzelnen Anforderungen aus pragmatischen Gründen dort dokumentiert.

Konzepte der virtuellen Plattform (UAP 2.2)

Die Konzeptentwicklung für die IIP-Ecosphere-Plattform fand in enger Zusammenarbeit mit dem KI-Accelerator (AP 6) statt. Die Grundideen dieser Konzepte entstammen der gemeinsamen Visions-Entwicklung mit dem KI-Accelerator (AP 6), bei denen offensichtlich wurde, dass eine rein virtuelle Plattform

nicht den Zukunftsinteressen der Industriepartner, die repräsentativ für zukünftige Industrie 4.0 Anwendungen argumentierten, entspricht. Damit wurde das container-virtualisierte Deployment von Künstlicher Intelligenz auf Edge Geräten als Kernanforderung für die Plattform (insbesondere von Lenze, Bitmotec und Phoenix Contact vorangetrieben) identifiziert. Im gleichen Kontext stellte sich auch die besondere Relevanz des von der Plattform Industrie 4.0 vorgeschlagenen (und inzwischen von der Industrial Digital Twin Association übernommenen) Standards der Industrie 4.0 Verwaltungsschale für die Plattform heraus, eine Anforderung, die insbesondere von Lenze eingebracht wurde.

Dokumentiert wurden die Konzepte und deren initiale Detaillierung insbesondere in den zwei Anforderungsdokumenten, die als IIP-Ecosphere Whitepaper veröffentlicht wurden, dem Plattform Usage View [STI21] (von Partner Lenze angestoßen und vorangetrieben) in Form von 67 Benutzungsszenarien und dem den funktionalen Anforderungen bzw. Qualitätsanforderungen (insbesondere mit den Partnern L3S und aus dem KI-Accelerator Siemens sowie KIPROTECT) in [EIC21a] in Form von 141 Anforderungen mit insgesamt 181 Unteranforderungen in 14 Themenbereichen. Die initialen Konzepte wurden weiter konkretisiert, in ein Software-Design überführt und schließlich die Realisierung und der Anforderungserfüllungsgrad im Plattformhandbuch [EIC20, EIC21b, EIC22d, EIC23a, EIC23c] diskutiert. Die Anforderungsermittlung wurde virtuell und interaktiv in Workshops mit allen relevanten Projektpartnern durchgeführt und von den genannten Partnern auch für zukünftige Plattformentwicklungen dokumentiert. Die zugehörigen Ansätze und Erfahrungen wurden in [EIC22e][SAU23b] veröffentlicht.

Bereits im Konzeptstadium diskutierten SUH, L3S, die Deutschen Messe Technology Academy, Phoenix Contact und die Volkswagen Digital Academy einen Ausbildungs-Demonstrator auf dem Gebiet der Messe Hannover. Dabei stellte sich heraus, dass zwar die Automatisierungs- und Steuerungstechnik für die IIP-Ecosphere Plattform relevant und geeignet ist, die dort umsetzbaren Applikationen leider nicht hinreichende Mengen realistischer Daten für einen demonstrierbaren Einsatz von Methoden der Künstlichen Intelligenz aus AP4 ermöglichen.

Arbeitsschritt 2.2.1: Konzepte Plattformdienste

Die Plattformdienste für die IIP-Ecosphere Plattform umfassen insbesondere die Basisdienste, ohne die die Plattform nicht einsatzfähig. Im Gegensatz zu den meisten industriellen Plattformen in [SAU21] sollen diese Basisdienste auf lokalen Ressourcen, d.h. on premise, oder, bei Bedarf, auch in einer Cloud installiert und ausgeführt werden können. Zu diesen Diensten gehören die zentralen Plattform-Server-Dienste für Verwaltungsschalen (Registry, AAS Server), das zentrale Plattform-Monitoring, der zentrale Datenverteiler (im Regelfall ein MQTT oder AMQP Broker), die Verwaltungs-Benutzeroberfläche, das Kommandozeileninterface (für tiefere Administrationstätigkeiten) sowie der Status-Verteiler als Brücke zwischen der Benutzeroberfläche und anderen Plattformdiensten. Neben den zentralen Plattform-Diensten, sind insbesondere die verteilten Verwaltungsdienste erforderlich, nämlich die Geräte-Abstraktion (ECS-Runtime von Edge, Cloud, Server-Runtime) die Geräteinformationen in Form eines Verwaltungsschalen-Submodells anbietet und auf Anweisung der zentralen Plattformdienste Applikationscontainer startet. In jedem dieser Applikationscontainer läuft eine Instanz des Service-Managers der Plattform, der zur Geräte-Verwaltungsschale beiträgt und auf Anweisung der Plattform Applikationen und Dienste starten kann. Falls aus Ressourcengründen erforderlich können ECS-Runtime und Service-Manager in einen Betriebssystemprozess zusammengefasst werden. Abbildung 2 stellt die grundlegenden Plattformdienste grafisch dar.

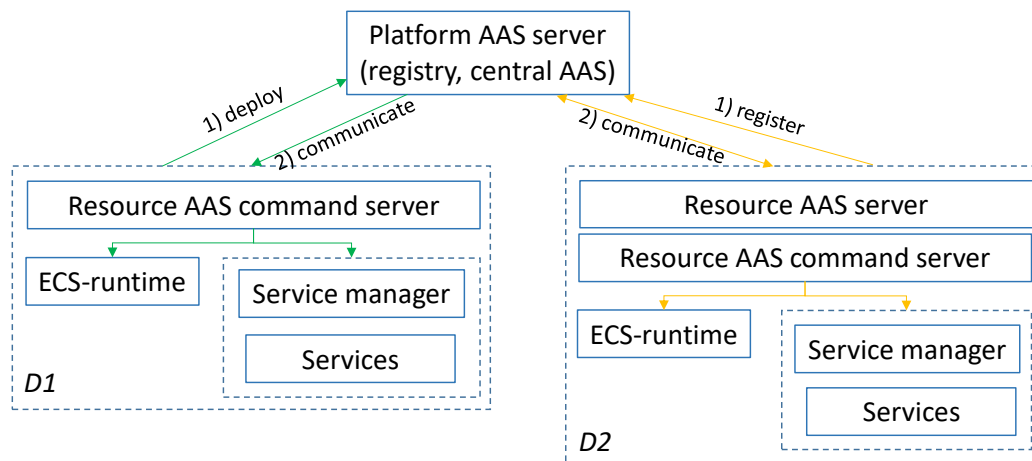


Abbildung 2: Grundlegende Plattformdienste

Insgesamt erweitern die Konzepte der IIP-Ecosphere Plattform aktuelle Plattformfähigkeiten [SAU21] in Bezug auf Interoperabilität/KI-Dienste, Standardnutzung (z. B. vertikale Verwaltungsschalenintegration), modell-getriebene Konfigurierbarkeit, low-code-basierte Code-Generierung zur Instantiierung der Plattform und der IIoT-Applikationen, dynamisches Deployment sowie Funktionalität zum Data Sharing und zur Sicherheit/IDS. Die relevanten Konzepte für diese Themen wurden in diesem UAP in 5 zugehörigen Arbeitsschritten aus den Ergebnissen von UAP2.1 abgeleitet, mit existierenden Open Source Komponenten z. B. aus dem Eclipse IIoT-Ökosystem in Verbindung gesetzt, so dass die Plattform auf möglichst vielen bewährten Komponenten aufsetzen und diese integrieren kann.

Arbeitsschritt 2.2.2: Konzepte Datendienste

Im Zentrum der IIP-Ecosphere Plattform steht das Aufsammeln von Daten von Maschinen und Geräten über Konnektoren, das Transportieren und Weiterverarbeiten von Daten, z. B., durch verteilbare Dienste, die Künstliche Intelligenz realisieren, sowie das Ablegen, weiterreichen bzw. Visualisieren von Daten. Hierbei treffen verschiedenste Ziele aufeinander: a) Standard-basierte und offene Kommunikation mit Maschinen und Geräten, z. B. über MQTT, OPC UA oder Verwaltungsschalen; b) flexible und offene Kommunikation innerhalb der Plattform, bei der der Anwender durch Konfigurationsmöglichkeiten Protokoll und Wire-Format festlegt, c) flexible und validierte Komposition von anwendungs-spezifischen oder von der Plattform bereitgestellten Diensten zu Applikationen d) dynamische Verteilbarkeit einzelner Dienste oder von Dienst-Ensembles auf Verarbeitungs-Ressourcen, die von der Plattform über Plattformdienste verwaltet wird, insbesondere Edge-Devices. Diese Grundkonzepte wurden in den Anforderungsdokumenten [STI21, EIC21a], deren Design und Implementierung im Plattformhandbuch [EIC20, EIC21b, EIC22d, EIC23a, EIC23c] dokumentiert.

Weiterhin wurden in diesem Arbeitsschritt die Kernkonzepte für Data Sharing und sicheren Datenaustausch basierend auf Konzepten des IDS in enger Zusammenarbeit mit AP3 und AP6 (semantische Datenintegration) abgeleitet und detailliert. Diese Konzepte wurden in Form von funktionalen Anforderungen und Qualitätsanforderungen in [EIC21a] beschrieben und in [EIC22d] verfeinert. Als Teil der Forschung auf dem Gebiet des sicheren Datenaustauschs wurde die Bachelorarbeit mit dem Titel "Enhancing the Security Design of Industrial IoT Platforms" (2022, Universität Koblenz-Landau, Institut für Softwaretechnik) betreut, diese Arbeit beschäftigte sich mit der Untersuchung von 12 Cloud Plattformen des Industriellen Internet der Dinge (IIoT), davon 3 Plattformen aus Deutschland und eine Plattform aus

dem Europäischen Wirtschaftsraum (EWR), auf Sicherheits- und Datenschutzmechanismen. Die Ergebnisse sind im Plattform-Handbuch dokumentiert.

Arbeitsschritt 2.2.3: Konzepte Sicherheit

In diesem Arbeitsschritt wurden die Sicherheitskonzepte (Zugriff, Usage Control) für die virtuelle Plattform in enger Zusammenarbeit mit AP3 und AP6 erarbeitet, detailliert und mit den anderen Arbeitsschritten und existierenden Open Source Komponenten abgestimmt. Die Ergebnisse wurden in Form von funktionalen Anforderungen und Qualitätsanforderungen in [EIC21a] beschrieben und in [EIC22d] verfeinert.

Um die Sicherheit und den Datenschutz der Plattform zu verbessern, wurde ein spezielles UML-Profil für Sicherheit und Datenschutz erstellt. Der Zweck dieses UML-Profiles ist es, einen Katalog von Sicherheits- und Datenschutzmechanismen bereitzustellen, um die entsprechenden UML-Modelle und die Architektur der Plattform zu annotieren. Mit einem solchen Annotationsprozess können geeignete Mechanismen auf einer hohen Abstraktionsebene ausgedrückt werden. Auf diese Weise geben wir den Entwicklern und Designern des Systems einen Überblick über einen vollständigen Rahmen für Sicherheit und Datenschutz. Darüber hinaus ermöglicht dieses spezielle Profil Datenschutz- und Sicherheitsprüfungen, wie z. B. automatische Analysen, sowie manuelle Audits. Das führt zu einer erhöhten Softwarequalität. Der durch das Profil dargestellte Katalog bietet die Möglichkeit, Datenschutz- und Sicherheitsmechanismen auf mehreren Abstraktionsebenen zu strukturieren. Es handelt sich dabei um Entwurfsstrategien (Strategies), Teilstrategien (Substrategies), Muster (Pattern) und Technologien zur Verbesserung des Datenschutzes (privacy enhancing technologies, PET), die beim Entwurf der Plattform angewendet werden können.

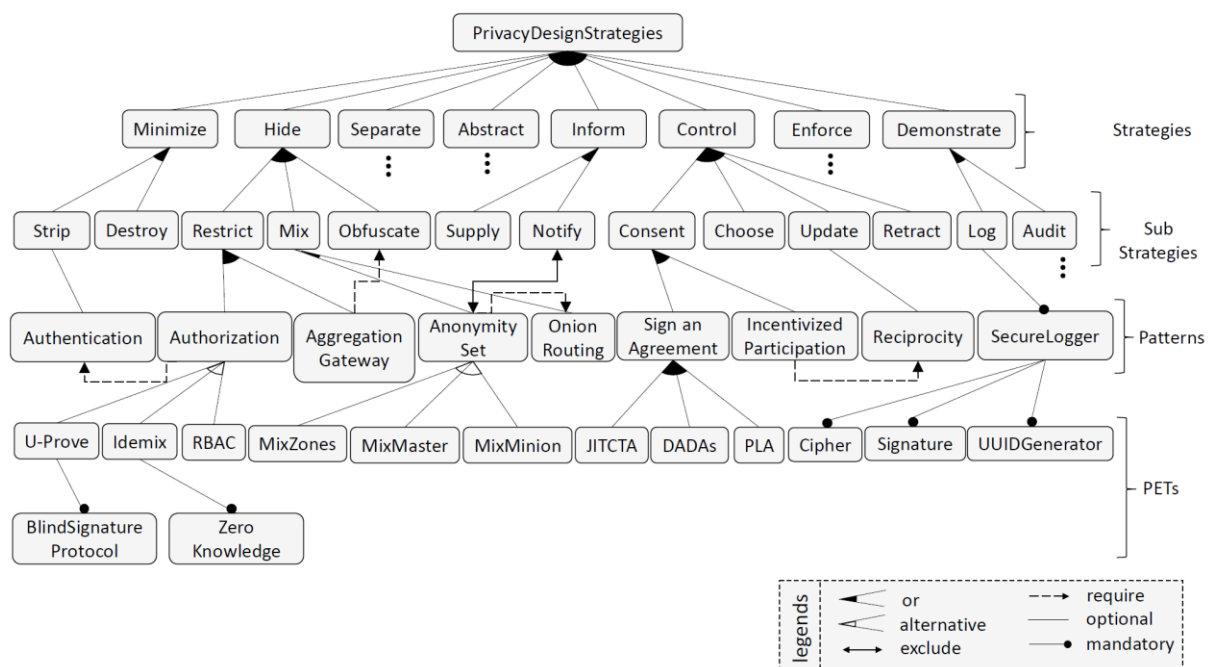


Abbildung 3: Ein Auszug aus dem Feature-Model mit Strategien, Teilstrategien und Mustern zum Schutz der Privatsphäre, und PETs (vgl. [EIC21a]).

Das Feature-Modell in Abbildung 3 die vier Abstraktionsebenen des Datenschutz- und Sicherheitskonzepts. Ein Feature-Modell beschreibt eine Reihe von Merkmalen und deren Beziehungen, in diesem Fall Merkmale zum Schutz der Privatsphäre und der Sicherheit und deren Beziehungen. Die Hauptstruktur des Feature-Modells ist hierarchisch. Es gibt jedoch Beziehungen, die nicht der hierarchischen Struktur folgen. Einige Merkmale können andere Merkmale erfordern. Zum Beispiel erfordert die *Auto-risierung (Authorization)* die Authentifizierung (*Authentication*). Andere Merkmale können sich gegenseitig ausschließen, z. B. Anonymität (*Anonymity Set*) und Benachrichtigung (*Notify*).

Das Ergebnis ist eine integrierte und umfassende Konzeptsdokumentation die einem zusammenfassenden Artefakt, dem Plattform-Handbuch präsentiert wird.

Arbeitsschritt 2.2.4: Konzepte Konfigurierbarkeit

Anpassbare und flexible Plattformen sind zurzeit nicht im Fokus der industriellen Entwicklung von IIoT/Industrie 4.0 Plattformen wie wir in [SAU21] in Form einer Analyse von 21 industriellen Plattformen diskutiert haben. Im Gegensatz dazu setzt die IIP-Ecosphere Plattform auf späte Entscheidungen, die nicht vom Entwickler bzw. Plattformanbieter, sondern vom Anwender getroffen werden, automatisierter Anpassung der Plattform an die Anwenderbedürfnisse, z. B., dem Entfernen nicht-benötigter Funktionalität auch um das Vertrauen in die Plattform zu erhöhen, sowie automatisierte Anwendungserstellung und -integration, um Applikationsentwickler von wiederholten, oft fehlerträchtigen Entwicklungstätigkeiten zu entlasten.

Das Konzept für die Konfigurierbarkeit sieht hierbei insbesondere ein durchgängiges Konfigurationsmodell „von den Devices bis zur Applikation“ vor (vereinfachte Darstellung in Abbildung 4 zeigt), das mit Validierungsbedingungen (Constraints) ausgestattet ist, und die Gültigkeit einer Plattform- bzw. Applikationskonfiguration vor Instantiierung überprüfen bzw. Werte zwischen Konfigurationsentscheidungen propagieren kann. Die Instantiierung erzeugt automatisiert die gewünschte, installierbare Variante der Plattform, Applikations-Templates zur vereinfachten Implementierung von Diensten (hier werden erforderliche technologie-bedingte Zusammenhänge etabliert um den Entwickler zu entlasten) und schließlich die automatisierte Integration von Diensten zu ausführbaren Applikationen. Zur Konzeptualisierung gehört neben der Konfigurationsmodellierung auch die zu instantiierenden Artefakte, die von Quellcode-Dateien für Produktions- und Testcode (Java, Python) über Service- und Applikationsdeskriptoren, ausführbare Build-Prozessen bis hin zu Plattform- und Applikationscontainern mit den jeweils benötigten Python-Abhängigkeiten.

Die Konzepte wurden in Form von funktionalen Anforderungen und Qualitätsanforderungen in [EIC21a] beschrieben und schrittweise in [EIC20, EIC21b, EIC22d, EIC23a, EIC23c] verfeinert.

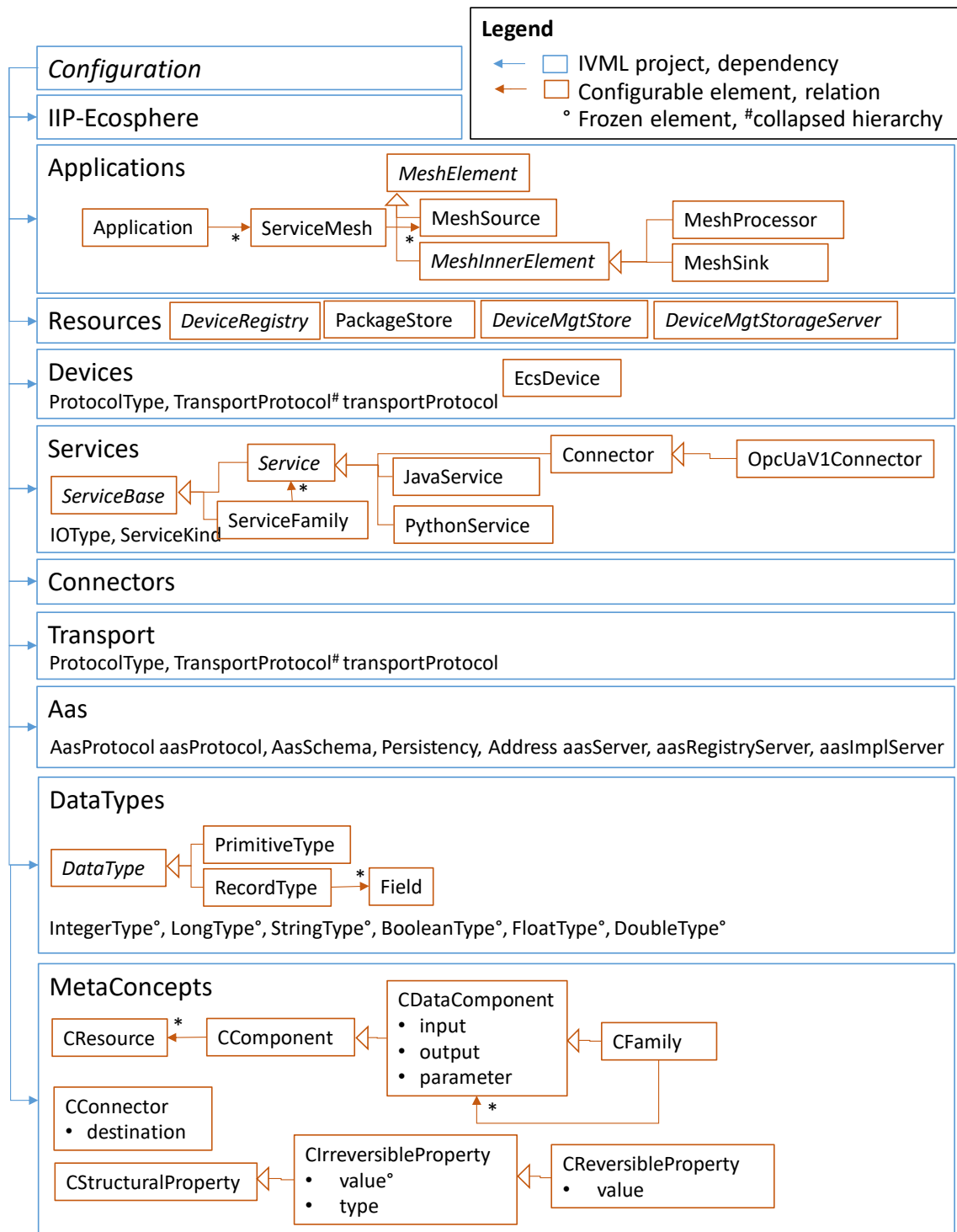


Abbildung 4: Stark vereinfachte Übersicht der Konfigurationskonzepte, d.h., des IIP-Ecosphere Konfigurations-Metamodells.

Arbeitsschritt 2.2.5: Konzepte Selbstanpassung

Installierte Plattform-Instanzen und Applikationen müssen der Evolution, z. B., durch neue Geräte und Standards, als auch der benötigten Flexibilität in Industrie 4.0 Umgebungen Rechnung tragen. Während die Evolution der Plattform über die Zeit, z. B. durch Integration neuer Funktionalitäten sowie Anstöße aus Arbeitsschritt 2.1.3 durch Nachführung des Konfigurationsmodells und der Instantiierung (Konzepte,

siehe Arbeitsschritt 2.2.4) sowie Neu-Generierung der Plattform-Instanzen bzw. Applikationen abgebildet werden kann, sind dennoch oft Anpassungen beim Starten einer Anlage bzw. der Applikationen oder sogar im laufenden Betrieb erforderlich. Hierbei zeigte sich insbesondere in den funktionalen Anforderungen [EIC21a] die Notwendigkeit für explizite Freigaben der Bediener, sowohl für die Anpassung von KI-Modellen als auch für die Re-Parametrierung von Diensten oder dem Re-Deployment. Für die Selbstanpassung wurden daher zwei Grundkonzepte vorgesehen:

- 1) Das dynamische, heterogene Deployment von Applikationen und Diensten auf verteilte Rechen-Ressourcen, insbesondere Edge-Geräte unterschiedlicher Hersteller. Dabei legt die Applikations-Konfiguration nicht fest, auf welchen Ressourcen die Ausführung letztlich stattfinden soll. Dies erfolgt im Rahmen des sogenannten Deployment-Plans, der erst nach der Anwendungsentwicklung definiert werden kann. Insbesondere können mehrere Pläne pro Anwendung festgelegt werden, die beim Start einer Applikation ausgeführt werden. Mit geeigneten Mechanismen, z. B. Verwaltungsschalenoperationen zum Passivieren/Aktivieren von Services bzw. einer geeigneten Wahl der technischen Grundlage der Stromprozessierung in der Plattform, können diese Festlegungen auch zur Laufzeit angepasst werden.
- 2) Service-Parameter, die im Konfigurationsmodell festgelegt und über Verwaltungsschalen-Operationen des jeweiligen Service gesetzt werden können. Die Grund-Werte der Parameter sind im Konfigurationsmodell festgelegt, können aber über den jeweiligen Deployment-Plan zur Applikationsstartzeit bzw. später zur Laufzeit geändert werden.

Die unterliegenden technischen Grundlagen erlauben zudem weitere Anpassungen zur Laufzeit, z. B., um dynamisch das Ziel einzelner Datenströme festzulegen und damit dynamische Service-Migration zu ermöglichen. Diese Möglichkeiten wurde für den Fall vorgesehen, um flexibel auf entsprechenden Bedarf reagieren zu können. Die Konzepte zur Selbstanpassung wurden in Form von funktionalen Anforderungen und Qualitätsanforderungen in [EIC21a] beschrieben und schrittweise in [EIC20, EIC21b, EIC22d, EIC23a, EIC23c] verfeinert.

Architektur der virtuellen Plattform (UAP 2.3)

In diesem UAP wurden die abgestimmten und integrierten Konzepte für die virtuelle Plattform in eine statische und dynamische Grob-/Feinarchitektur übersetzt. Ziel war es, die Konzepte so zu detaillieren, dass eine reibungslose technische Umsetzung basierend auf ausgewählten Open Source Komponenten ermöglicht wird. Die Architektur selbst wurde agil und inkrementell entwickelt, d.h., ausgehend von einer Grobarchitektur in Form aufbauender Schichten (Schichten-Architektur), wurde eine Feinplanung der tiefsten noch nicht realisierten Schicht vorgenommen, umgesetzt, getestet, integriert, die Ergebnisse in die Architektur zurückdokumentiert und die nächste Schicht angegangen.

Die einzelnen Plattform-Bereiche aus UAP2.2 wurden dabei, basierend auf einer Grobarchitektur, zunächst in Form von Teilarchitekturen, im Regelfall als UML-Klassendiagramme und, wo angebracht, UML-Sequenzdiagramme oder UML-Zustandsdiagramme technisch detailliert und in eine vorläufige Gesamt-Schichtenarchitektur basierend auf RAMI 4.0 Basiskonzepten integriert. Hierzu wurde ein mit den Partnern abgesprochenes OpenSource UML-Werkzeug (Papyrus) verwendet, d.h., das integrierte Architekturmodell ist in maschinenlesbarer Form verfügbar und war Sicherheitsprüfungen durch CARiSMA (siehe Arbeitsschritt 2.3.3) zugänglich. Einzelne Zwischenversionen der Plattformarchitektur

wurden insbesondere mit dem KI-Accelerator (AP 6) diskutiert. Abbildung 5 stellt eine stark vereinfachte Version der der Gesamt-Architektur dar.

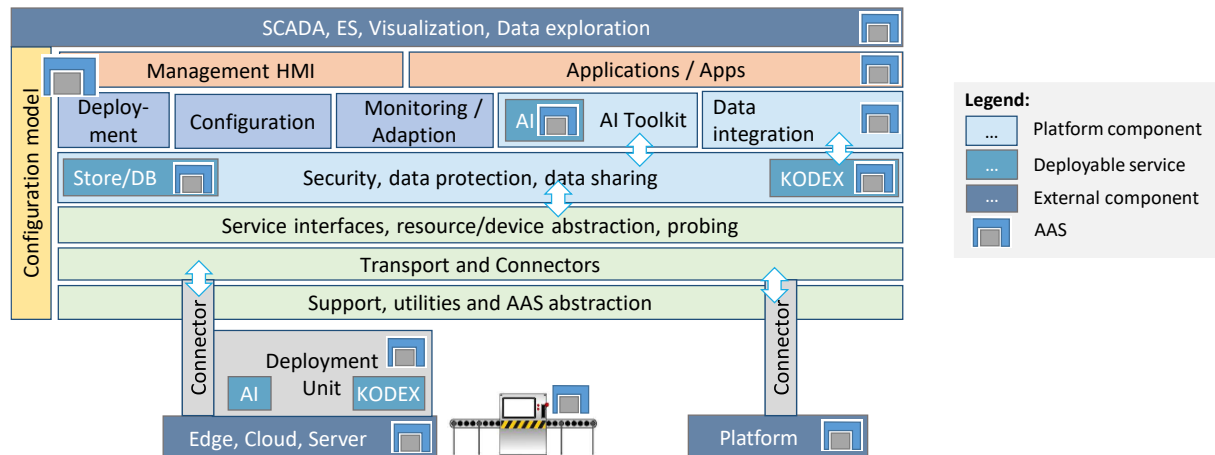


Abbildung 5: Finale Übersichtsarchitektur der IIP-Ecosphere Plattform

Die tiefste Schicht der Plattform stellt Verwaltungsschalen-Funktionalität in für die Plattform einfach nutzbarer Form bereit, d.h., durch eine Abstraktere Schnittstelle, und erlaubt damit lokale Anpassungen bei Updates von Eclipse BaSyx sowie, im Extremfall, einen Austausch der verwendeten Implementierung. Zudem werden Basismechanismen z. B. zur Identitätsverwaltung, zum Komponenten-Lebenszyklus, zum Laden optionaler bzw. alternativer Komponenten eingeführt. Darüber liegt der Plattform-interne Datentransport (wählbar sind AMQP und MQTT in verschiedenen Versionen bzw. Realisierungen) sowie die Konnektoren zu externen Datenquellen bzw. -senken (insbesondere OPC UA, Verwaltungsschalen, MQTT, Beckhoff ADS). Als nächste Schicht folgen die Verwaltungsmechanismen für verteilbare Services (ServiceManager), die Service-Ausführungsumgebungen (Java, Python), die Ressourcenabstraktion (ECS-Runtime) und Techniken zum Monitoring auf Service-Ebene. Danach folgen zwei Schichten für wiederverwendbare Dienste, die die Datensicherheit erhöhen (z. B. eine Integration von KIPROTECT KODEX) und das Data Sharing ermöglichen. In der folgenden Schicht sind der KI-Werkzeugkasten (insbesondere Föderiertes Lernen, Integration von RapidMiner RTSA), die semantische Datenintegration (siehe AP6), das zentrale Monitoring, die Verwaltung und Instantiierung des Konfigurationsmodells sowie das dynamische, heterogene Deployment vorgesehen. Applikationen basieren auf diesen Mechanismen und können die in den darunterliegenden Schichten definierten Services verwenden, aber auch eigene Services einführen. Auf der auf dieser Ebene angesiedelten Plattform-Verwaltungsschale, die insbesondere aus den verfügbaren Services, den Ressourcen (ECS-Runtime und ServiceManager), einer Abbildung des Konfigurationsmodells und weiteren technischen Informationen besteht, setzt letztlich die Benutzerschnittstelle zur Plattform-Verwaltung und zum Applikationsdesign auf. Das Konfigurationsmodell überspannt, wie bereits in Arbeitsschritt 2.2.4 angesprochen, alle Schichten der Plattformarchitektur.

Wie oben beschrieben, wurde einzelne Komponenten agil entlang dieser der Schichtenarchitektur in Unterarbeitspakt 2.5 inkrementell realisiert, die Realisierung mit dem ursprünglichen Entwurf abgeglichen und im Plattformhandbuch [EIC20, EIC21b, EIC22d, EIC23a, EIC23c] mit den jeweiligen Beweggründen und Entscheidungen dokumentiert. Dies führte zu insgesamt 16 Detail-Diagrammen, die einzelne Plattform-Komponenten bzw. Teilarchitekturen detaillieren.

Arbeitsschritt 2.3.1: Teilarchitektur Plattformdienste

Die Plattformdienste überdecken verschiedene Schichten der Gesamt-Plattformarchitektur, insbesondere höhere Schichten (in Abbildung 5 in abstrakter Form mit „Deployment“, „Configuration“ oder „Monitoring“ bezeichnet). In diesem Arbeitsschritt wurden insbesondere die zentralen Plattformdienste wie das zentrale Monitoring, die ECS-Runtime und der Service-Manager geplant. Im Entwurf jedes Plattformdienstes sind die zugrundeliegenden bzw. zu integrierenden OpenSource-Komponenten verzeichnet. Entsprechend der Konfigurations-Konzepte aus Arbeitsschritt 2.2.4 wurden zudem diverse Variationspunkte vorgesehen, insbesondere für die Verwaltungsschalen-Implementierung, gerätespezifische Netzwerk- und Monitoring-Mechanismen oder um das Framework zur verteilten Service-Ausführung austauschen zu können. Hierfür waren einfache, einheitliche Variabilitätsmechanismen zu identifizieren und einzuplanen. Weiterhin wurden die Verwaltungsschalen-Teilmodelle für diese Plattformdienste entworfen und, wo bereits möglich, gegen (standardisierte oder in Standardisierung befindliche) Konzepte der IDTA abgeglichen. Die initialen Entwürfe wurden, je nach Entwicklungsstand, im Plattformhandbuch [EIC20, EIC21b, EIC22d, EIC23a, EIC23c] mit den jeweiligen Entscheidungen und Besonderheiten dokumentiert. Abbildung 6 stellt beispielhaft das Architekturdiagramm der ECS-Runtime als Plattformservice dar.

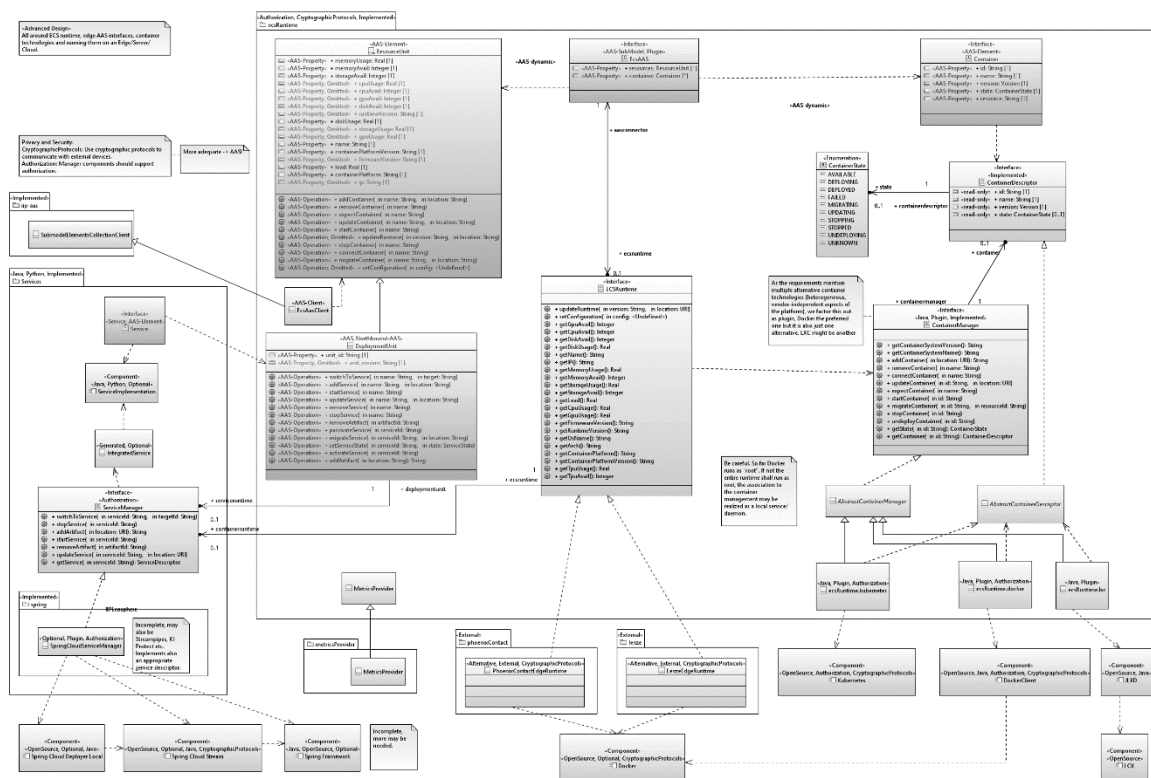


Abbildung 6: Ausschnitt aus der UML-Teilarchitektur der Plattformdienste, hier die ECS-Runtime

Arbeitsschritt 2.3.2: Teilarchitektur Datendienste

Die Teilarchitektur der Datendienste überdeckt sowohl die unteren Schichten der Gesamtarchitektur (Datentransport, externe Konnektoren) als auch die Grundlagen der höheren Schichten für wiederverwendbare Datendienste. Hierzu gehören insbesondere der Datentransport (mit Transport-Konnektoren

für MQTT und AMQP sowie die protokoll-unabhängigen format-spezifischen Daten-Serialisierer), die externen Konnektoren (für OPC UA, MQTT, Verwaltungsschalen und Beckhoff ADS), die Service-Schnittstellen und das lokale Service-Monitoring. Weiterhin zählen hierzu wiederverwendbare Service-Integrationen für Python- und Betriebssystemdienste (das sogenannte Service-Environment für Java und Python, in Abbildung 7 dargestellt), sowie generische Dienste insbesondere zur Repräsentation und Steuerung von Applikationen über Verwaltungsschalen.

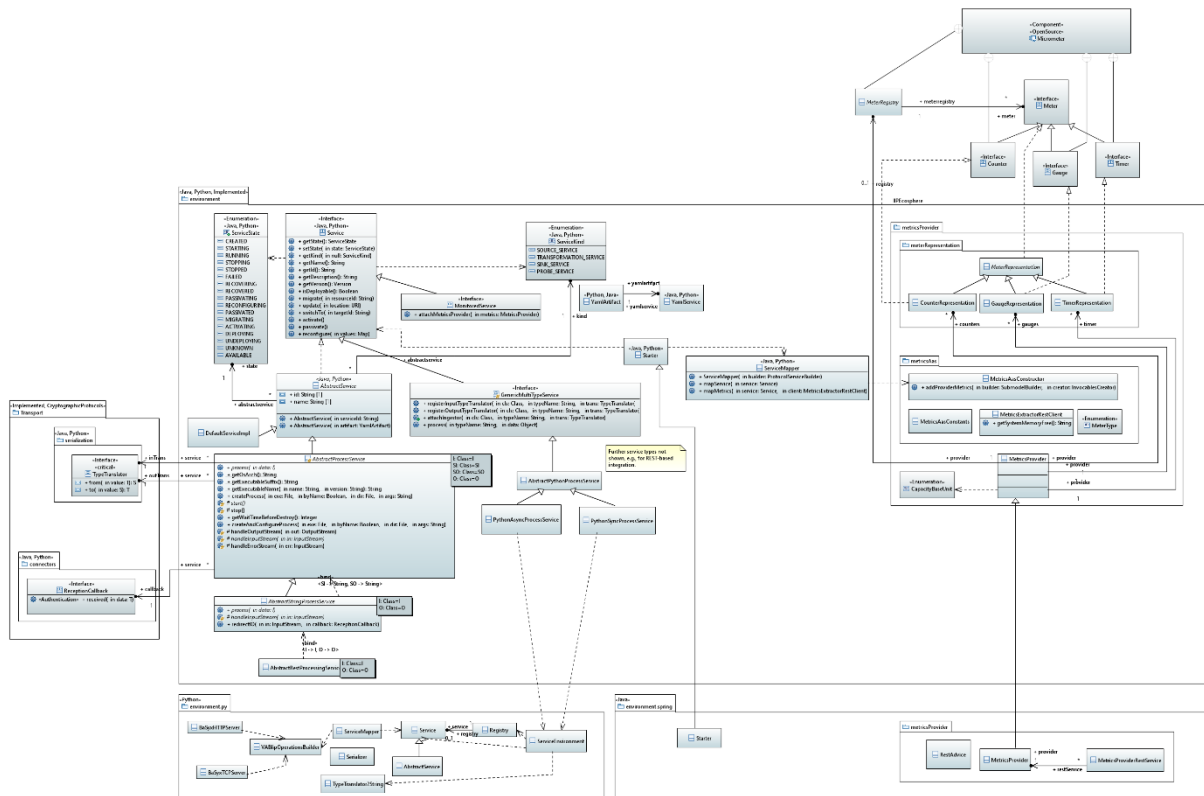


Abbildung 7: Teilarchitektur der gemeinsamen Service-Ausführungsumgebung für Java, Python und Betriebssystemdienste

Eng verwoben mit den Datendiensten ist die modell-getriebene Konfiguration und Instantiierung der Applikationen. Hierfür wurden insbesondere Mechanismen zur automatisierten Generierung von Konnektoren, Applikations-Containern sowie zum (generierten) Testen von Applikationen auf verschiedenen Ebenen (Service-Implementierung, Service, Applikation, verteilte Applikation) entworfen.

Zu diesen Themen fanden auch diverse Austausche zwischen dem BMWi Projekt DaPro sowie den ThinkTanks „Daten“ (AP3) und „KI und Produktion“ (AP4) statt.

Arbeitsschritt 2.3.3: Teilarchitektur Sicherheit

In diesem Arbeitsschritt wurden die Sicherheits-Konzepte aus Arbeitsschritt 2.2.3 in Form einer Teilarchitektur detailliert und mit Ergebnissen aus AP6 zu Anonymisierung und Pseudonymisierung integriert. Beispielsweise wurde hier die Integrations-Architektur für das OpenSource Anonymisierungs- und Pseudonymisierungsprogramm KODEX des Partners KIPROTECT über das Service-Environment aus Arbeitsschritt 2.3.2 als Betriebssystemprogramm entworfen. Dabei ist zu berücksichtigen, dass KODEX

als generischer Service bereitgestellt wird, d.h., bei Verwendung des Kodex-Dienstes im Konfigurationsmodell einer Applikation, erzeugt die Instanziierung notwendige applikations-spezifische Artefakte und passt damit den generischen Kodex-Dienst automatisiert in die jeweilige Anwendung ein.

Weiter wurden basierend auf Arbeitsschritt 2.2.5 für das Projekt relevante Sicherheits-Konzepte ausgewählt und in ein UML-Profil überführt (siehe Abbildung 8). Dies ermöglicht es die im Projekt entstandenen UML-Modelle mit den Sicherheits-Konzepten, zu annotieren, also so zu ergänzen, dass zusätzliche Sicherheitsinformationen direkt im Modell verfügbar sind. Das ermöglicht es die Sicherheitskonzepte während des gesamten Lebenszyklus der Softwarekomponenten einzusehen.

In diesem Arbeitsabschnitt wurde speziell der Security and Data Protektion Layer gestaltet und mit den Sicherheitskonzepten erweitert. Die besondere Aufmerksamkeit liegt hier bei Anonymisierung, Pseudonymisierung und Verschlüsselung von schützenswerten Daten.

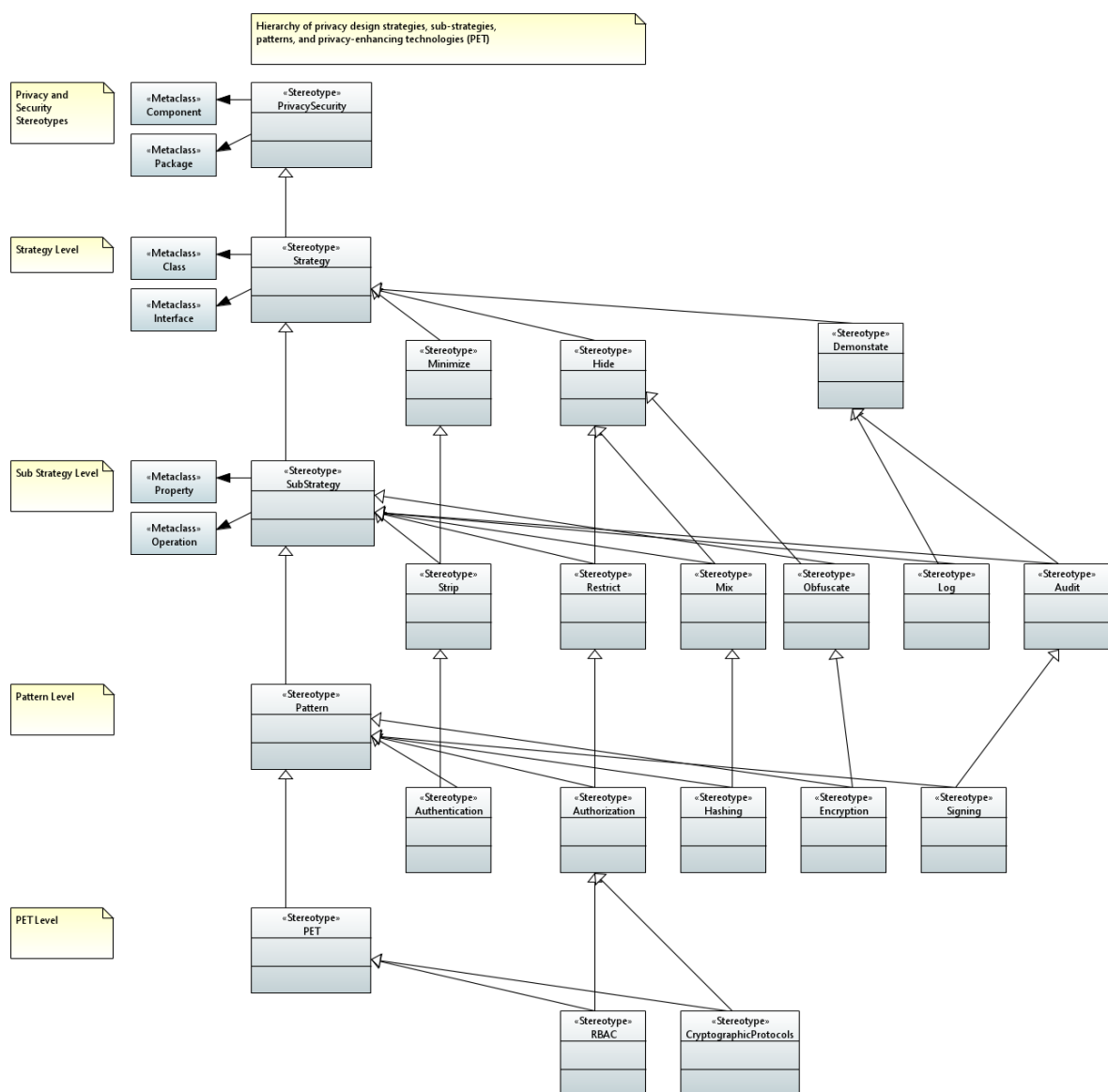


Abbildung 8: Das in ein UML-Profil überführte Feature-Model mit Strategien, Teilstrategien und Mustern zum Schutz der Privatsphäre, und PETs

Im Rahmen der Architektur, konnte hierbei die UMLsec Methodologie verwendet werden. UMLsec ist eine Erweiterung der UML zur Integration von sicherheitsrelevanten Informationen in UML-Diagrammen. Diese Informationen können für die modellbasierte Sicherheitsanalyse verwendet werden. Die meisten Sicherheitsinformationen werden mithilfe von Stereotypen hinzugefügt und decken viele Sicherheitseigenschaften ab, darunter sicherer Informationsfluss, Vertraulichkeit und Zugriffskontrolle. Mit Hilfe eines Angreifermodells können diese Eigenschaften auf Modellebene überprüft werden. Die UMLsec Methodologie und deren Analysen (Checks) sind in dem Open-Source Tool CARISMA implementiert. Hierzu wurden im Papyrus-Architekturmodell der Plattform sicherheitsrelevante Komponenten, wie Konnektoren (Connectors) zu anderen Systemen auf die Sicherheitsaspekte Informationsfluss, Vertraulichkeit und Zugriffskontrolle untersucht. Ein Resultat war, z. B. das der Nachrichtenaustausch mittels HMAC (hash-based message authentication code) gesichert werden soll.

Da verschiedene Industrie 4.0 Protokolle und deren jeweilige Implementierung Sicherheitsmechanismen auf verschiedene Weise realisieren und Details von Sicherheitsmechanismen wie Passworte, Access-Tokens oder Zertifikate nicht im Konfigurationsmodell enthalten sein sollen, wurde im Rahmen dieses Arbeitsabschnitts der sogenannte IdentityStore entworfen, ein abstrakter Zugriff auf sicherheitsrelevante Daten. Konkrete Realisierungen können übliche Mechanismen wie LDAP, Schlüsselringe (keyring) oder Zertifikatsserver sein. Insbesondere zur Fehlersuche aber auch für Regressionstests im Rahmen der kontinuierlichen Integration ist es hier hilfreich, auch unsichere, nicht-verschlüsselte, einfach zu verwendende Realisierungen bzw. Kommunikationsformen aktivieren zu können.

Arbeitsschritt 2.3.4: Teilarchitektur Selbstadaption

Entsprechend der zwei Kern-Konzepte aus Arbeitsschritt 2.2.5 wurden die Teilarchitekturen der Plattform so ausgelegt, so dass die vorgesehene Form der Selbstadaption ermöglicht wird. Zum einen wurde die Struktur der Deployment-Pläne entworfen und in die zentralen Plattformdienste, insbesondere die zentrale Verwaltungsschale der Plattform so integriert, dass Deployment-Pläne (via Verwaltungs-Benutzeroberfläche oder Kommandozeile) ausgeführt werden können. Damit werden die unterliegenden Plattformdienste wie ECS-Runtime und Service-Manager instruiert und letztlich Applikationen und Services verteilt orchestriert. Zum anderen wurden die Service-Schnittstellen und deren Verwaltungsschalen so verfeinert, dass Aktivierung, Passivierung und Parametrierung von Services zur Start- und zur Laufzeit ermöglicht wird.

Arbeitsschritt 2.3.5: API für KI-Dienste

Die ursprüngliche Idee, eine vereinheitlichte Programmierschnittstelle (API) als Grundlage für Dienste, die Künstliche Intelligenz umsetzen, hat sich im Laufe des Projekts grundlegend gewandelt. Dazu trugen verschiedene Diskussionen mit dem KI-Accelerator (AP6) bei, die das Konzept einer „easy-to-use“ KI in Form einzelner Services in den Vordergrund gestellt haben. Leider wurde dabei offensichtlich, dass breite Feld der künstlichen Intelligenz selbst im Anwendungsfeld der industriellen Produktion nicht so einfach zu erfassen ist, so dass sich die Partner in Abstimmung mit AP2 auf eine zunächst pragmatische und experimentelle Vorgehensweise geeinigt haben: Basierend auf den vier IIP-Ecosphere-Demonstratoren und den öffentlichen Plattform-Demonstratoren sollen Beispiel-Dienste realisiert und in einen wiederverwendbaren KI-Baukasten (KI-Toolkit) integriert werden. Damit wurde eine eigene Teilarchitektur für den KI-Baukasten erforderlich. Diese besteht aus:

- Einer Funktions-Bibliothek zur vereinfachten Realisierung von Datenverarbeitungsaufgaben, z. B. zur Identifikation von QR-Codes mit Java oder Python, zur Vorverarbeitung und Recodierung von Bildern oder zur generischen Aggregation von Zeitreihen.
- Einer generischen Integration der KI-Laufzeitumgebung des Partners RapidMiner, dem Real-Time Scoring Agent (RTSA) mit einem ähnlichen Ansatz wie für KODEX in Arbeitsschritt 2.3.3 beschrieben. Der RTSA führt KI-Deployments aus, die mit dem RapidMiner Studio entwickelt und getestet werden. Die Service-Integration übersetzt daher Eingangs- und Ausgangsdaten in geeigneter Weise, passt damit den RTSA in eine gegebene Applikation ein und verwendet RapidMiner bzw. RTSA als fortgeschrittenen, eingebetteten KI-Baukasten.
- Einer generischen und teil-generierten Integration des Python Frameworks Flower zur Realisierung von Föderiertem Lernen. Im Gegensatz zu generischen Diensten wie KODEX oder RTSA handelt es sich bei der Flower-Integration nicht um abgeschlossene Funktionalität, sondern selbst um ein Framework, das selbst wiederum geeignet programmiert werden muss. Um die Programmierung und die anschließende Integration in eine Plattform-Applikation zu erleichtern, sieht diese Teilarchitektur vor, dass zur jeweiligen Anwendung des Dienstes passender Python-Rahmencode generiert und über das Python-Service-Environment aus Arbeitsschritt 2.3.2 ausgeführt wird. Der Rahmencode umfasst sowohl den Federated Learning Server als auch den Federated Learning Client und erleichtert die Realisierung von Federated Learning Applikationen, wie z. B. wir dies im Rahmen des Hannover Messe 2022 Demonstrators oder des dynamischen Demonstrators mit NovoAI demonstriert und evaluiert haben.

Arbeitsschritt 2.3.6: API für Datenmarktplatz

In diesem Schritt wurde mit dem Think Tank Daten (AP 3) eine mögliche Schnittstelle zwischen IIP-Ecosphere Plattform und Datenmarktplatz diskutiert. Besonderer Fokus lag auf einer übergreifenden Integration mit Industrie 4.0 Verwaltungsschalen, z. B., zur Repräsentation typisierter Daten und Meta-Daten.

Arbeitsschritt 2.3.7: Gesamtarchitektur virtuelle Plattform

Neben den bereits beschriebenen Teilarchitekturen erforderte die Gesamtarchitektur für die IIP-Ecosphere Plattform auch den Entwurf diverser, weiterer Komponenten und Teilfunktionen, die die Teilarchitekturen verbinden. Dazu gehören

- Ein Service-Lebenszyklusmechanismus, der die flexible Plattformkonfiguration und -instantiierung unterstützt und das Starten bzw. das Stoppen der Plattformdienste ermöglicht.
- Ein erweiterter Ressourcen-Mechanismus der Programm-Ressourcen in unterschiedlichen, alternativen Quellen finden kann. Dies vereinheitlicht auch das Testen von Komponenten.
- Ein Mechanismus der die Installation von Betriebssystemprogrammen abstrahiert und damit z. B., Diensten erforderliche, alternative Java- oder Python-Versionen bereitstellt.
- Ein Mechanismus zum Auflösen semantischer Identifikationen in Verwaltungsschalen (siehe auch AP 6). Hierfür wurden neben der Abfrage von Online-Katalogen, z. B. EClass via REST auch lokale, Internet-Unabhängige Mechanismen vorgesehen.

- Ein Mechanismus zum Verfolgen des Ausführungs-Status von Aufgaben in einem verteilten System, z. B., zur Darstellung des aktuellen Status der Ausführung eines Deployment Plans beim Starten einer Applikation in der Verwaltungs-Benutzerschnittstelle. Dieser Mechanismus wurde mit den üblicherweise synchronen Verwaltungsschalen-Operationen integriert.
- Erweiterungen der Build-Umgebung Maven, so dass mit einem Java-Build-System auch Python-Programme syntaktisch überprüft, validiert und in vereinheitlichter Weise Konfigurations-Instantiierungen ausgeführt werden können. Diese Erweiterungen wurden aufgrund des Feedbacks der Industriepartner im KI-Accelerator (AP 6) entworfen.
- Eine Teilarchitektur für die Verwaltungs-Benutzeroberfläche (siehe Arbeitsschritt 2.5.9) der IIP-Ecosphere-Plattform basierend auf der Gesamtverwaltungsschale der IIP-Ecosphere Plattform.

Integration und Anpassung existierender Komponenten (UAP 2.4)

Die Realisierung eines komplexen Software-Systems wie der IIP-Ecosphere-Plattform ohne Verwendung bzw. Integration bereits bestehender Software-Komponenten ist in der vorgegebenen Projektlaufzeit nicht effizient, da viele Teile wie Industrie 4.0 Protokolle oder Datenstromverarbeitung bereits in Form stabiler Open Source Komponenten existieren. Ziel dieses UAP ist es, basierend auf der Architektur aus UAP 2.3 und den zur Verfügung stehenden Open Source Komponenten in möglichst kurzer Zeit bei höchstmöglicher Qualität eine funktionsfähige Plattform zu entwickeln (UAP 2.5).

Zunächst wurde hierzu der Kern der Plattform bestehend aus Komponenten wie BaSyx-Integration, Transport, Konnektoren und Service-Schnittstellen in Open Source bereitgestellt und inkrementell um weitere Schichten und Komponenten erweitert. Dabei wurden ebenso die jeweils erforderlichen funktionalen Regressionstests entwickelt und integriert. Regressionstests sind für eine stabile Softwareentwicklung erforderlich, da dann notwendige Änderungen und Ergänzungen von Komponentenrealisierungen im Rahmen einer kontinuierlichen Integration die Ausführung der Tests der (potentiell) betroffenen Plattform-Komponenten anstößt, um möglichst frühzeitig, unbeabsichtigtes Fehlverhalten zu erkennen und zu korrigieren.

Arbeitsschritt 2.4.1: Integration/Anpassung Eclipse.IoT

Das Eclipse.IoT Ökosystem beinhaltet verschiedene, für IIoT-Systeme interessante Projekte, die insbesondere als Grundlage für die IIP-Ecosphere Plattform verwendet werden können. Allerdings sind nicht alle Komponenten gleichermaßen relevant oder integrierbar. Zunächst wurden, basierend auf der Plattformarchitektur aus UAP 2.3 pro vorgesehener Komponente ein kleines Beispiel realisiert und damit die Komponente auf Anwendbarkeit und spätere Integrierbarkeit getestet. Danach wurde die jeweilige Komponente je nach Teilarchitektur integriert und entsprechende Regressionstests entwickelt. Integriert wurden unter anderem:

- Eclipse BaSyx als Verwaltungsschalen Client bzw. Server mit optionaler Verschlüsselung. Hierbei ist zu berücksichtigen, dass wir uns zur Risikominimierung aufgrund der sich noch entwickelnden Verwaltungsschalen-Standardisierung für eine Abstraktion von BaSyx entschieden haben, d.h., BaSyx wird in der Plattform nicht direkt sondern durch Schnittstellen angesprochen, die für die Zwecke der Plattform auf einem günstigeren Abstraktionsniveau liegen. Beispiels-

weise verwenden wir Entwurfsmustern, die für uns übliche Verwaltungsschalenkonstrukte sowohl knapper (rund 30% weniger Programmcode erforderlich) als auch bei Erstellung auf typische Fehlverwendung prüfen. Zudem wurden größere, wiederverwendbare Konzepte realisiert und abstrahiert, z. B. zum Remote-Deployment einzelner Verwaltungsschalen. Diese Abstraktion hat sich als günstig erwiesen, da Anpassungen in der Verwendung von BaSyx, die durch erforderliche Aktualisierungen von BaSyx notwendig waren, direkt auf dieser Ebene und nicht etwas bei jeder Verwendung umgesetzt werden konnten [EIC23b].

- Eclipse Paho als MQTT Transport-Konnektor und als externer MQTT-Konnektor, jeweils optionaler Verschlüsselung. Die generische Konnektor-Integration ist die Basis für eine automatisierte Generierung konkreter Konnektoren in individuellen Anwendungen.
- Eclipse Milo als externer OPC UA Konnektor, ebenso mit optionaler Verschlüsselung. Auch hier ist die generische Konnektor-Integration ist die Basis für eine automatisierte Generierung konkreter Konnektoren in individuellen Anwendungen.

Neben den genannten Komponenten wurden weitere Open Source Komponenten außerhalb des Eclipse.IoT Ökosystems identifiziert und integriert.

- RabbitMQ als AMQP Transport-Konnektor (mit optionaler Verschlüsselung). Analog zu der Integration von Eclipse Paho legt die generische Konnektor-Integration die Basis für eine automatisierte Generierung konkreter Konnektoren in individuellen Anwendungen.
- Apache Qpid J für als Vermittler für das AMQP-Protokoll, HiveMq und Moquette als Vermittler für das MQTT-Protokoll. Insbesondere Apache Qpid J und HiveMq werden im Rahmen einer Plattform-Instantiierung als mitgelieferte Kommunikations-Broker angepasst und bereitgestellt. Zudem werden die integrierten Komponenten insbesondere auch in Experimenten und Regressionstests eingesetzt.
- Micrometer als Framework um Laufzeit-Messungen auf (Edge-)Geräten interoperabel zugänglich zu machen. Dabei wurden in [CAS21] Micrometer, Spring Cloud Streams und Industrie 4.0 Verwaltungsschalen zum Überwachen von Ressourcen integriert und Integrationsmuster mit unterschiedlichen Performanzeigenschaften diskutiert.
- Integration des Konfigurations- und Produktlinienwerkzeugs EASy-Producer zur systematischen Konfiguration von Software-Systemen. EASy-Producer realisiert die formalen Sprachen und die Laufzeitumgebung, die in der IIP-Ecosphere-Plattform zur Konfigurationsmodellierung und Instantiierung verwendet werden. Um mit EASy-Producer Plattform- und Applikations-Container instantiieren zu können (siehe Arbeitsschritte 2.2.4 und 2.3.2), wurden entsprechende Client-Bibliotheken für Docker und LXC in EASy-Producer integriert.
- Optionale Integration von ThingsBoard zum IoT-Gerätemanagement (z. B., On/Offboarding) sowie MinIO und S3Mock zur zentralen Bereitstellung großer Dateien, z. B., für Docker-Images, Service-Images oder Firmware-Updates.
- Integration des Open Source Frameworks Prometheus für das zentrale Plattform-Monitoring.
- Integration des Web-Frameworks Angular für die Realisierung einer Verwaltungsschalen-basierten Benutzeroberfläche für die Plattform.

Arbeitsschritt 2.4.2: Integration/Anpassung Eclipse-Daten

Ursprünglich war die Integration von speziellen Eclipse.IoT Komponenten zur strombasierten, verteilten Datenverarbeitung vorgesehen. Nach einem Vergleich von Kandidaten auch außerhalb des Eclipse.IoT Ökosystem hat sich Spring Cloud Stream als der beste Kandidat herausgestellt wie in [EIC20] diskutiert. Daher erfolgte eine modellbasierte Integration von Spring Cloud Stream zur strombasierten, verteilten Datenverarbeitung. Weiterhin erfolgte eine Integration mit den Transport-Konnektoren um Spring Cloud Stream mit dem gleichen Transport-Protokoll zu betreiben wie die Transport-Schicht der Plattform.

Als ein Ergebnis der Umsetzung des Demonstrators für die Hannover Messe 2022 [EIC22a] hat sich herausgestellt, dass rückwärtsgerichtete Datenströme, z. B., für Kommandoflüsse, in der verwendeten Version von Spring Cloud Stream nur unzuverlässig funktionieren, was auch durch ein Update der Komponente nicht gelöst werden konnten. Allerdings war es möglich, die Datenströme direkt über den Datentransport der Plattform zu führen und die notwendigen Anpassungen an den Applikationen durch eine Anpassung der Instantiierung/Generierung automatisiert durchzuführen,

Arbeitsschritt 2.4.3: Integration/Anpassung Eclipse-Deployment

Ursprünglich war ebenso die Integration von speziellen Eclipse.IoT Komponenten zur Unterstützung und Optimierung des Deployments vorgesehen, z. B., durch besondere Update-Techniken für Applikations-Container. Im Verlauf des Projekts hat sich herausgestellt, dass zunächst eine direkte Integration von Docker als Container-Manager in die IIP-Ecosphere-Plattform der einfachste und universellste Ansatz ist. Diese Integration in die ECS-Runtime ermöglicht es ebenso, die lizenzfreundlichere Variante podman zu verwenden. Als weitere experimentelle Alternative wurden Linux Container (LXC) integriert. Die generativen Techniken der Applikations-Instantiierung wurden entsprechend erweitert, so dass sowohl Plattform- als auch Applikationscontainer mit ihren individuellen Programm- und Python-Abhängigkeiten automatisiert insbesondere für Docker, als Proof-of-Concept auch für LXC erzeugt werden können. Weiterhin konnte durch geeignete Mechanismen während der Generierung der Applikations-Container eine Optimierung des Zeitbedarfs für die Container-Erstellung und für das Deployment auf Zielgeräte erreicht werden [ALA23].

Zudem wurde eine experimentelle Integration des Container-Orchestrators Kubernetes entwickelt. Im Gegensatz zu Docker und LXC, die in der IIP-Ecosphere-Plattform über die Geräte-Verwaltungsschalen angesprochen werden und damit einer Industrie 4.0 kompatiblen Kommunikation folgen, realisiert Docker ein eigenes proprietäres Protokoll, was unter Umständen aufgrund zukünftiger strikterer Standardisierung nicht mehr zulässig sein könnte. Als Ansatz wurde eine Kubernetes-Erweiterung entwickelt und evaluiert, die eine Ausführung von Kubernetes über Industrie 4.0 Protokolle wie MQTT, AMQP oder Verwaltungsschalen ermöglicht [ALA23]. Da die Docker-Integration der IIP-Ecosphere Plattform für die Projektziele ausreichend ist, wurde zunächst von der Fortführung der Integration von Kubernetes abgesehen, auch da die Komplexität der Installation und des Betriebes als auch der Ressourcenbedarf auf Edge-Geräten einer aktuellen praktischen Verwendung entgegenstehen.

Arbeitsschritt 2.4.5: Integration IDS

In diesem Schritt wurden Ergänzungen für das Data Sharing in der IIP-Ecosphere Plattform vorbereitet, insbesondere Komponenten, die für die Realisierung von IDS-Konzepten und -Konnektoren erforderlich

sind. In diesem Kontext wurden von der Universität Koblenz-Landau eine Vielzahl an studentischen Abschlussarbeiten betreut. Die diente zum Wissenstransfer der Projekteinhalte.

Arbeitsschritt 2.4.6: Aufsetzen und Wartung der kontinuierlichen Integration

Ein komplexes Software-System wie die IIP-Ecosphere-Plattform erfordert kontinuierliche Validierung durch Regressionstests anlässlich jeder Änderung des Quelltexts bzw. von Komponenten. Hierfür wurde eine kontinuierliche Integration auf den Servern der Arbeitsgruppe SSE der Universität Hildesheim eingerichtet. Der Quelltext der IIP-Ecosphere-Plattform wird öffentlich auf github verwaltet und Änderungen von dort bezogen. Bei erfolgreichen Tests werden die neuen Komponentenversionen automatisiert (in vereinheitlichter Weise inklusive Nativ- und Pythonanteilen) als Snapshots in ein öffentlich verfügbares Maven-Artefakt-Repository deployt. Im Gegensatz dazu werden Release-Komponenten ins öffentliche Maven Central Repository deployt und Test- bzw. Entwicklungs-Container auf dockerhub bereitgestellt.

Über die Laufzeit des Projekts ist zudem eine Erweiterung der kontinuierlichen Integration auf neue Plattform-Komponenten sowie Wartung der kontinuierlichen Integration und des Snapshot-Deployments erforderlich. Beispielsweise musste aufgrund einer anstehenden Technologieänderung bei github das Zugriffs- und Abfrageprotokoll für die mehr als 40 IIP-Ecosphere-Komponenten und damit deren Build-Prozess angepasst werden. Zur Wartung gehört ebenso die Erstellung von Containern auf github mit automatisiertem Deployment nach dockerhub. Diese Erweiterungs- und Wartungstätigkeiten wurden bei Bedarf durchgeführt.

Umsetzung der virtuellen Plattform (UAP 2.5)

In diesem UAP wurden die weiteren Komponenten und insbesondere die Integration der Plattform-Komponenten realisiert und getestet. Dabei entstanden bis zum Projektende 7 Release-Versionen der Plattform, dazu (für die späteren Versionen) je eine aktualisierte Version des Plattform-Handbuchs [EIC20, EIC21b, EIC22d, EIC23a, EIC23c]. Das früheste Release wurde im August 2021, das letzte innerhalb von IIP-Ecosphere wird im Oktober 2023 bereitgestellt.

In diesem Umfeld wurden diverse studentische Arbeiten an der Universität Hildesheim durchgeführt, z. B., zur Verwaltungsschalen-Compliance, zur Verwaltungsschalen-Skalierbarkeit, zum Service-Monitoring auf Edge-Geräten mit Verwaltungsschalen, zur automatisierten Container-Erstellung, zur Transformation von OPC UA Companion Specifications in die Konfigurationssprache der Plattform, zu alternativen Service-Integrationen oder zu alternativen Datentransportformaten und deren Performanz.

Arbeitsschritt 2.5.1: Umsetzung/Integration der Plattformkomponenten/dienste

In diesem Arbeitsschritt wurden alle geplanten Kern-Plattformkomponenten und dienste realisiert und Regressionstests bereitgestellt. Dazu gehören u.a. die Transport-Konnektoren, die externen Konnektoren, die Service-Umgebungen, das Service-Monitoring sowie der Service-Manager, die ECS-Runtime und die zentralen Plattform-Services, die die Plattform-Verwaltungsschale bereitstellen. Für die Plattformkomponenten wurden entsprechende Verwaltungsschalen als Daten- und Operationsschnittstellen realisiert [EIC23b], so dass eine verteilte Verwaltung der Geräte, der Container und der Services ermöglicht wurde. Hierzu kamen Mechanismen um Verwaltungsschalen für Komponenten externer Hersteller wie Geräte (die oft noch ohne Verwaltungsschale ausgeliefert werden), Services und Konnektoren, oder Applikationen einfach erstellen zu können.

Arbeitsschritt 2.5.2: Umsetzung/Integration Konfigurierbarkeit

Grundlage für die Realisierung der Konfigurierbarkeit und der Instantiierung ist die Konzeptionierung und Modellierung des Konfigurationsmetamodells in der Integrated Variability Modeling Language (IVML). Hierbei wurde auch die unterliegende Implementierung, d.h., das EASy-Producer Framework angepasst, verbessert und erweitert, insbesondere um dynamisch Modellerweiterungen hinzuladen zu können, und so Kernmodell und Erweiterungen (z. B., für KODEX, RapidMiner, MIP technologies oder NovoAI) zu trennen. Inzwischen enthält das Konfigurations-Metamodell mehr als 218 strukturierte Typen, mehr als 620 Variablen und um die 100 Validierungsbedingungen. Insbesondere handelt es sich bei diesem Modell um ein topologisches Konfigurationsmodell, das Datenflussgraphen zwischen Services in Applikationen modellieren und validieren kann.

Zur automatisierten Instantiierung einer IVML-Plattformkonfiguration ist weiterhin ein Instantiierungsprozess erforderlich. Für die IIP-Ecosphere-Plattform wurde ein entsprechend komplexer Instantiierungsprozess entworfen und realisiert, der mehr als 15 Artefakttypen u.a. Schnittstellen, Datenserialisierer, Service-Templates, Projekt-Templates Service-Integrationen, Test-Rahmen für Java und Python-Services, Build-Prozesse, Applikations-Artefakte und Applikations-Container erstellen kann. Insgesamt handelt es sich dabei um mehr als 140 Templates, die verschiedene Aspekte einer Plattform- bzw. einer Applikationsinstantiierung in Java und Python generieren können. Dabei werden insbesondere Datenservices, die mit der Plattform ausgeliefert werden, geeignet in Applikationen eingepasst.

Neben 9 Instantiierungs-Testfällen mit eigenem Modell enthält die Plattform instantiierbare 7 Anwendungsbeispiele, darunter den mit den Partnern entwickelten Demonstrator für die Hannover Messe 2022 und die Tage der Digitalen Technologien 2022, die wiederum als Regressionstests in der kontinuierlichen Integration verwendet werden. Die Quellen für den Demonstrator für die Hannover Messe 2023 und seine Erweiterung auf zwei Applikationen für die EMO'23 soll nach Projektabschluss veröffentlicht werden.

Arbeitsschritt 2.5.3: Umsetzung/Integration besonderer Konnektoren

Im Industrie 4.0 Bereich sind verschiedenste Protokolle und Informationsmodelle gängig, insbesondere auf der sogenannten Feldebene. Fokus der IIP-Ecosphere-Plattform, festgelegt durch die Plattform-Anforderungen [EIC21a], sind allerdings die höheren Ebenen, die insbesondere von IT-Komponenten verwendet werden, während die Feldebene zunächst nicht zu betrachten ist. Daher war primär die Integration der laut [SAU21] zwei „großen“ Industrie 4.0 Protokolle OPC UA und MQTT vorgesehen. Diese sind insbesondere als Konnektoren verfügbar, die automatisiert in der Applikationsinstantiierung in Applikationen eingepasst werden, d.h., für Standardfälle ist keine Programmierung (low code, sogar no code) für diese Konnektoren erforderlich. Weiterhin waren diese Konnektoren zunächst als Startpunkt der Applikationsdatenflüsse gedacht. Im Laufe des Projekts hat es sich gezeigt, dass insbesondere die MQTT Konnektoren für weitere Aufgaben eingesetzt werden können, z. B. zur Übergabe von Ergebnissen an ein Dashboard, d.h., am Ende eines Datenflusses oder sogar in der Mitte, wenn Daten z. B. in eine Cloud übergeben und wieder empfangen werden sollen.

Neben der Nutzung der Verwaltungsschalen innerhalb der Plattform, insbesondere zum verteilten Steuern der Plattform, gibt es auch Anwendungsfälle, in denen Verwaltungsschalen-Konnektoren benötigt werden, z. B., zum zielgerichteten Abfragen von Informationen (in den IIP-Ecosphere Plattform-De-

monstratoren). Analog zu den OPC UA Konnektoren können Verwaltungsschalen-Konnektoren im Regelfall automatisch generiert werden. Zudem hat es sich gezeigt, dass es erforderlich ist, zusätzliche Funktionen zur Parametrierung von Konnektoren zu ermöglichen, z. B. um relevante Informationen für eine Abfrage aus einem anderen Datenformat zu extrahieren und an den Konnektor weiterzugeben.

Da die Konnektoren und deren Eingangs/Ausgangsdatentypen in IVML spezifiziert werden, ist für eine Integration der OPC UA Companion Specifications eigentlich eine manuelle Übersetzung von XML nach IVML erforderlich. Dies ist insbesondere für die großen Companion Specifications, die mehrere tausend Zeilen umfassen, unrealistisch. Diese Lücke hat eine studentische Arbeit an der Universität Hildesheim geschlossen, die eine automatisierte Übersetzung der OPC UA Companion Specifications nach IVML und damit die automatisierte Generierung der zugehörigen besonderen Konnektoren ermöglicht. Damit existiert ein Proof-of-Concept, der in Zukunft auf IDTA-standardisierte Verwaltungsschalen ausgedehnt werden könnte.

Arbeitsschritt 2.5.4: Umsetzung/Integration dynamisches Deployment

Das dynamische, heterogene Deployment der Applikationen in der IIP-Ecosphere Plattform basiert im Wesentlichen auf der ECS-Runtime und dem Service-Manger sowie deren Verwaltungsschalen. Die Orchestrierung wird durch Deployment-Pläne definiert erfolgt entweder per Knopfdruck in der Verwaltungs-Benutzerschnittstelle oder über die Kommandozeilenschnittstelle (hier können auch einzelne Services deployt werden).

Die Ausführung von Services, insbesondere von Python-Services setzt die Installation entsprechender Software-Pakete auf dem Zielgerät voraus. Auch wenn die manuelle Installation eine Möglichkeit ist, wird im Regelfall eine virtualisierte Ausführung bevorzugt, auch da dann gefahr- und restlos installierte Bibliotheken entfernt werden können. Die Containerisierung von Applikationen unterstützt die Plattform in Form automatisch generierter Container (siehe auch Arbeitsschritte 2.2.3 und 2.3.2). Die erforderlichen Software-Komponenten werden dabei als Teil der Service-Spezifikation im Konfigurationsmodell definiert und während der Container-Generierung berücksichtigt. Dabei sind neuerdings sowohl konfliktierende Abhängigkeiten (in Form verschiedener conda-Environments) möglich und die Herangehensweise versucht die Erstellungszeit als auch die Zeit für das Deployment der Container auf die Geräte zu optimieren [ALA23]. Die Container werden in einer lokalen Docker-Registry für alle Geräte bereitgestellt, so dass Container gefahrlos auch lizenzierte oder geschützte Programme enthalten können.

Arbeitsschritt 2.5.5: Umsetzung/Integration Datenkomponenten

Für die stromorientierte Datenverarbeitung wurde Spring Cloud Streams ausgewählt und in Arbeitsschritt 2.4.2 integriert. In diesem Arbeitsschritt wurde die Integration von Spring Cloud Streams mit der Umsetzung der Konfigurierbarkeit in Arbeitsschritt 2.5.2 verbunden, so dass (prinzipiell) selbst der Strom-Prozessor bei einer Plattform-Instantiierung austauschbar ist. Dabei muss der Ansatz zur Generierung von Service-Integrationen entsprechend generisch sein. Service-Integrationen stellen die Brücke zwischen dem Framework zur Datenprozessierung (und dessen technischen Anforderungen) und der aktuellen Service-Implementierung, z. B., in Java, Python oder als Betriebssystemprogramm, dar.

Zudem wurden in diesem Arbeitsschritt das KI-Toolkit realisiert, d.h., die Datenverarbeitungsbibliothek, die generische RTSA-Integration und die Generierung der Flower-Templates für eine vereinfachte Verwendung von Förderiertem Lernen.

Arbeitsschritt 2.5.6: Umsetzung/Integration Sicherheitskomponenten

In diesem Abschnitt erfolgte im Wesentlichen die Realisierung des vereinheitlichten Mechanismus zur Bereitstellung von Authentifizierungs-Token und Sicherheitszertifikaten. Dieser Mechanismus wurde mit dem Datentransport, den Konnektoren und den Verwaltungsschalen integriert und getestet. Zudem wurden in diesem Arbeitsschritt gemeinsame Tätigkeiten mit AP6, z. B., die generische KODEX-Integration oder eine Integration des Hyper-Frameworks vorbereitet.

Arbeitsschritt 2.5.7: Umsetzung/Integration Selbstanpassung

Die Selbstanpassung wurde wie in Arbeitsschritt 2.3.4 geplant umgesetzt und getestet. Diese Fähigkeiten wurden insbesondere in den IIP-Ecosphere-Demonstratoren für die Hannover Messe 2023, den Tag der Information 2023 in Hildesheim wie auch für Nürnberg Digital 2023 eingesetzt. Hier war es erforderlich, mehrere Instanzen derselben Applikation für einen Anwendungsfall im Federated Learning (gemeinsam mit AP4) auszuführen, d.h., ein Cobot lernt von einem zweiten Cobot, wobei jeder Cobot von einer eigenen Applikationsinstanz gesteuert wird. Die entsprechende Applikation soll im Konfigurationsmodell nur einmal beschrieben und anwendungsspezifische Services natürlich auch nur einmal implementiert werden. Die Applikationsinstanzen werden dabei im jeweiligen Deployment-Plan so parametrisiert, dass Services und insbesondere Konnektoren auf die zum jeweiligen Cobot gehörenden OPC UA Datenpfade zugreifen können, ohne dass dies pro Applikationsinstanz im Konfigurationsmodell oder in der Implementierung erfolgen muss. Theoretisch wäre es mit dem Ansatz auch denkbar, dass die Cobots im benannten Demonstrator während der Ausführung der Applikationen ihre Rollen tauschen.

Arbeitsschritt 2.5.8: Umsetzung/Integration adaptives Deployment

Die Anforderungsermittlung für die IIP-Ecosphere Plattform hat gezeigt, dass innovative produzierende Firmen ein überraschendes Interesse an Systemen haben, die ihr Verhalten zur Laufzeit ändern, dies aber unter expliziten menschlichen Freigaben erfolgen soll. Diese Einsicht hatte zur Folge, unsere initialen Pläne für ein adaptives, sich zur Laufzeit selbsttätig änderndes Deployment zu reduzieren und den Fokus auf eher pragmatische Adaptionsunterstützung zu legen. Deployment-Pläne mit Startzeit-Anpassungen sowie Parametrierung von Services und Container zur Laufzeit in den Arbeitsschritten 2.5.4 und 2.5.7 legen hierzu die Basis. Weiterhin zeigte der Federated Learning Anwendungsfall für die Hannover Messe 2023, den Tag der Information 2023 in Hildesheim wie auch für Nürnberg Digital 2023, wie notwendig ein (unter expliziter Freigabe liegendes) Laufzeit-Deployment von KI-Modellen in der Praxis sein kann. Daher in diesem Arbeitsschritt entsprechende Mechanismen in die Plattform integriert, die den gezielten Austausch von KI-Modellen im laufenden Betrieb ergeben. In Kollaboration mit AP 4 und AP 6 entstanden daher neben einer initialen pragmatischen Umsetzung auch die Konzeption eines Verwaltungsschalen-basierten KI-Modell-Repositories, das sowohl Industrie-Anforderungen wie Versionierung als auch eine einfache Plattform-Integration ermöglicht.

Arbeitsschritt 2.5.9: Umsetzung einer Benutzerschnittstelle für die IIP-Ecosphere Plattform

Dieser Arbeitsschritt ist im ursprünglichen Arbeitsplan für IIP-Ecosphere nicht enthalten und wurde im Rahmen der Budgetübertragung zwischen KIPROTECT und UHi dem Gesamtarbeitsplan hinzugefügt. Diskussionen mit interessierten Firmen zeigten früh, dass eine geeignete Verwaltungs-Benutzerschnittstelle für die Plattform für eine erfolgreiche Verwertung erforderlich ist. Da die Gesamt-Verwaltungs-

schale der Plattform hinreichend Information für eine Einsicht in den laufenden Betrieb bereitstellt (bekannte Ressourcen, Services, Applikationen und deren Laufzeitzustand), bildet diese eine gute Informationsgrundlage (siehe Abbildung 9).

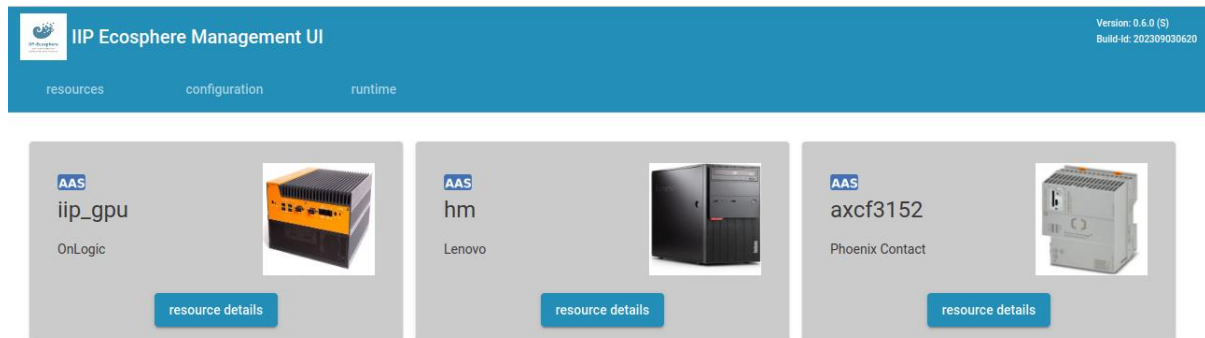


Abbildung 9: Benutzerschnittstelle der IIP-Ecosphere-Plattform, hier Ressourcen-Übersicht

Zudem sollte die Benutzerschnittstelle das Entwerfen und Erstellen von Applikationen für die IIP-Ecosphere Plattform erleichtern. Hierfür wurde, auch auf Initiative des Partners Sennheiser, ein eigenes Verwaltungsschalen-Teilmodell für die Konfiguration entworfen, dass jede Konfigurationseinstellung von den Ressourcen bis zu den Datenflüssen der Applikationen repräsentieren, konsistente Modelländerungen und selbst Instantiierungsprozesse durchführen kann. Darauf aufbauend wurde eine Benutzerschnittstelle basierend auf dem bekannten Web-Framework Angular entwickelt. Diese Benutzerschnittstelle ist für die meisten Elemente und Konfigurations-Konzepte Tabellen-basiert, erlaubt aber für die Datenflusseditoren typisches Drag & Drop (siehe Abbildung 10) ähnlich zu bekannten Systemen wie Node-RED. Die Benutzerschnittstelle selbst wird durch den Plattform-Instantiierungsprozess auf die jeweilige Installation vorbereitet, ist also selbst wiederum konfigurierbar, und sorgt dafür, dass, zusammen mit den zentralen Plattformdiensten, ein automatischer Start ermöglicht wird.

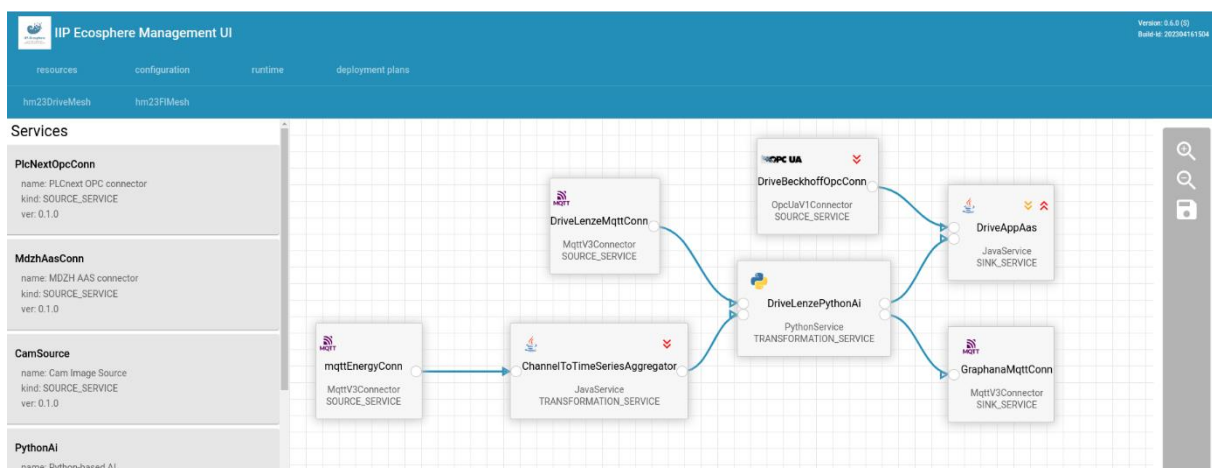


Abbildung 10: Benutzerschnittstelle der IIP-Ecosphere-Plattform, hier Applikations- bzw. Datenfluss-Designer

Evaluation der virtuellen Plattform (UAP 2.6)

In diesem UAP wurden die Komponenten der virtuellen Plattform auf ihre Funktion, ihr Zusammenspiel und ihre Qualitätseigenschaften evaluiert. Ziel ist es, die besonderen Eigenschaften der virtuellen Plattform zu analysieren und zu belegen.

Grundlage der Evaluation ist die Validierung der Funktionalität, die bereits in UAP 2.5 durch Regressions-tests für alle integrierten Open Source Komponenten, in Arbeitsschritt 2.4.6. durch Vorbereiten und Warten einer kontinuierlichen Integration und in UAP 2.5 durch die Entwicklung von regressiven Integrationstests für die Plattformkomponenten auf unterschiedlichsten Granularitätsebenen - von Einzel-funktionen bis zu den Plattform-Beispielen - abgedeckt wird.

Diverse Eigenschaften der Plattform wurden insbesondere qualitativ und quantitativ bei der Vorbereitung für und im Betrieb auf drei Messen (Hannover Messe 2022, Hannover Messe 2023, EMO 2023), auf den Tagen der Digitalen Technologien 2022 in Berlin, dem Tag der Informatik 2023 in Hildesheim und Nürnberg Digital 2023 evaluiert und belegt.

Arbeitsschritt 2.6.1: Evaluation der Plattformfunktionalität

Neben der grundlegenden Funktionalität wurden in diesem Arbeitsschritt insbesondere die Evaluation der Performanz verschiedener Integrationsmuster für Verwaltungsschalen im Bereich des Service/-Geräte-Monitoring [CAS21], der Skalierbarkeit von Verwaltungsschalen [SAU23], der Anwendbarkeit für die KI-Applikationsentwicklung [EIC22a] und der generellen Anwendbarkeit vertikal-Integrierter Verwaltungsschalen für Industrie 4.0 Plattformen [EIC23b] gezeigt. Auf den oben angesprochenen Messen und öffentlichen Veranstaltungen konnten wir zudem die Stabilität der Plattform und der Demonstrations-Applikationen über Tage hinweg beobachten. Im Vorfeld identifizierte Funktions- und Performanzprobleme (insbesondere das Bereitstellen von Statusupdates als Datenstrom in einem Verwaltungsschalen-Teilmodell) wurden analysiert und behoben.

2.6.2: Evaluation der Konfigurierbarkeit

Das Konfigurationsmodell und der Instantiierungsprozess ist Kernbestandteil jeder Applikationsentwicklung mit der IIP-Ecosphere-Plattform. Insbesondere während der Entwicklung der Plattform-Beispiele, der mit der Konfigurierbarkeit in Verbindung stehenden Plattform-Komponenten aber auch der Plattform-Demonstrations-Anwendungen identifizierte fehlende Konfigurationskonzepte oder Probleme im Instantiierungsprozess konnten behoben und somit der gesamte Ansatz stabilisiert werden.

Im Detail konnten wir das automatisierte Deployment von Plattformkomponenten, Services und generierten Container auf unterschiedlichen Edge-Geräten, d.h., in einer heterogenen, teilweise ressourcenbeschränkten Umgebung, erproben und evaluieren. Für die Generierung von Containern im Build-Prozess der Plattform zeigen wir in [ALA23], dass eine Kombination von Techniken wie Container-Stacking und Datei-Fingerprinting sowohl Container-Erstellungszeit als auch Container-Deploy/Updatezeit auf den Ressourcen signifikant verbessern kann. Insbesondere evaluierten wir in [EIC22b] die Anwendbarkeit, Speichernutzung und Performanz eines KI-Services des Hannover Messe 2022 Demonstrators auf einer Phoenix Contact AXC F 3152 PLC-Edge. Zudem hat das Demonstrator-Team in [EIC22a] aus den Erfahrungen verschiedene Erkenntnisse für die allgemeine Realisierung KI-basierter Industrie 4.0 Anwendungen abgeleitet und diskutiert. Weiterhin wurde der Fall Industrie 4.0-konformer Kommunikationsprotokolle (MQTT, AMQP, Verwaltungsschalen) für Container-Orchestratoren in [ALA23] diskutiert

und gezeigt, dass ein Proxy-basierter Ansatz ohne Eingriffe in den Orchestrator selbst keine wesentlichen Performanznachteile nach sich zieht. Eine studentische Arbeit beschäftigte sich mit einer Erweiterung der Evaluation für OPC UA.

Für die neu hinzugekommene Benutzeroberfläche (Arbeitsschritt 2.5.9) konnte die Konfigurierbarkeit für die meisten Elemente gezeigt werden, insbesondere für die Konfiguration der Applikations-Datenflüsse.

Um die Integration externer Komponenten zu erleichtern, wurde zudem ein mehrstufiges Testkonzept entwickelt, umgesetzt und mit der Applikationsinstantiierung integriert. Hierzu gehören insbesondere generierte Tests für Python-Services, da KI-Implementierungen oft ohne geeignete Integrationstests geliefert werden. Zudem wurde eine verteilte Test- und Evaluationsumgebung für die Plattform geschaffen, die automatisiert die Plattform sowie gegebene Applikationen instantiiert, ausführt, beobachtet und auch wieder rückstandsfrei von den Ausführungsressourcen entfernt.

Die Eigenschaften der Konfigurierbarkeit wurden ebenso in Kollaboration mit Partnern evaluiert. Einerseits wurde im November 2022 ein Service-Integrationsworkshop (siehe auch AP6) an der Universität Hildesheim zur Evaluation des Plattformansatzes und zur Sammlung von Feedback durchgeführt. Diskutierte Verbesserungen wie eine Integration von Python in die Applikations-Build-Prozesse, die Erstellung eines Entwicklungs-Containers mit Plattform und Entwicklungswerkzeugen sowie die Bereitstellung eines Video-Tutorials wurden zeitnah umgesetzt.

Weiterhin wurde die Konfigurierbarkeit (und die Funktionalität) in den zwei an die Plattform angebundenen dynamischen Demonstratoren (MIP technology, NovoAI) evaluiert, insbesondere auch im eigenen Umgang der jeweiligen Partner. Hierbei zeigte sich, dass die Plattform bis dahin unbekannten Anwendungsfällen, insbesondere dem Wechselspiel zwischen MIP Sensorsoftware und KI-Service ohne Probleme gewachsen ist. Allerdings zeigte sich ebenso, dass neben dem Video-Tutorial noch mehr Einstiegsdokumentation erforderlich ist. Sobald allerdings diese Hürde (momentan interaktiv mit Unterstützung von Mitarbeitern der UHi) gemeistert ist und die Struktur und die möglichen Einstellungen im Konfigurationsmodell verstanden wurden, eigene Applikationen oder Dienste in effizienter Weise entwickelt bzw. integriert werden können. Erkenntnisse aus den dynamischen Demonstratoren werden in die Plattform-Dokumentation auch nach Projektabschluss eingearbeitet. Dabei sollen über die Zeit ähnliche Themen vom Plattformhandbuch in die Online-Dokumentation auf github verschoben werden, so dass der Zugang erleichtert und das Plattformhandbuch mehr für eine tiefere Auseinandersetzung mit der Architektur genutzt werden kann.

Arbeitsschritt 2.6.3: Evaluation der Datenflüsse

Eine grundlegende Evaluation der Performanz des Datentransports mit Industrie 4.0 Protokollen im Bereich der Stromverarbeitung zeigt [EIC22d], dass die entsprechenden Anforderungen aus [EIC21a] mit hinreichend Leistungsreserve eingehalten werden können. Analog zur Konfigurierbarkeit waren die Datenflüsse durch KI-Komponenten und das Service-Environment der Plattform Kernbestandteil der Entwicklung der Plattform-Demonstratoren sowie Gegenstand verschiedener Evaluationen [EIC22a, EICH22b, EICH22c]. Insbesondere anhand des Demonstrators für die Hannover Messe 2022 konnte

Verbesserungen für die Datenflüsse, die Generierung, den Container-Einsatz und das Testen bzw. Integrierten von Services abgeleitet werden [EIC22a], die durch entsprechende Anpassungen an der Plattform gelöst werden konnten bzw. zu einer verbesserten und stabileren Umsetzung führten. Die verschiedenen Plattform-Demonstratoren zeigten in praxisrelevanter Weise unter Verwendung von Industriekomponenten insbesondere beidseitige Konnektivität zur Feld-Ebene (Brownfield) sowie zur IT-Ebene in Form eines üblichen Dashboard-Setups (Node-RED, Influx, Grafana).

Arbeitsschritt 2.6.4: Evaluation der Sicherheit

Im Bereich der Sicherheit konnte die Funktionalität und die Effektivität der Plattform-Mechanismen zur Verwaltung von Benutzeridentitäten, Krypto-Zertifikaten sowie zur Authentifizierung und Verschlüsselung der Daten im Bereich des Datentransports und der Konnektoren gezeigt werden. Weiterhin wurde die KODEX-Integration (aus AP 6) im Kontext der Plattform validiert und evaluiert, z. B., im Rahmen einer studentischen Arbeit in Form alternativer Integrationen (Konsolenströme, REST) und deren Anwendbarkeit und Performanz.

Arbeitsschritt 2.6.5: Evaluation der Selbstanpassung

Die Konzeption und Umsetzung der Selbstanpassung (siehe Arbeitsschritt 2.5.7 und 2.5.8) wurden insbesondere in den IIP-Ecosphere Platfordemonstratoren des Jahres 2023 praktisch und qualitativ evaluiert. Die erwartete Funktionalität bzw. die erwarteten Effekte konnten bestätigt werden.

Arbeitsschritt 2.6.6: Sicherheits-Codereview

Durch den geänderten Fokus, insbesondere auf ein verteiltes, heterogenes System und KI-Ausführung auf der Edge wie von den KI-Accelerator-Partnern angeregt aber auch durch die Entwicklung dreier öffentlicher Demonstratoren mit all ihren praktischen Details und Problemen (z. B. lange unklaren Funktionsproblemen im Hannover Messe 2023 Demonstrator ausgelöst durch eine auch von Experten zunächst für unproblematisch eingestufte Netzwerk-Topologie) mussten wir die Ressourcen im Think Tank „Plattformen“ geeignet umverteilen. Ein Sicherheits-Codereview wurde daher letztlich nicht durchgeführt.

Arbeitsschritt 2.6.7: Evaluation: Sicherheit der Nutzungsszenarien

In diesem Arbeitsschritt wurde die Sicherheit der intendierten Nutzungsszenarien evaluiert und analysiert. Dazu wurden die Erkenntnisse in weiteren, anschlussfähigen Projekten wie DataPorts, einer Datenplattform für digitale, vernetzte Häfen und XAPS, ein Projekt mit dem Ziel durch Methoden des maschinellen Lernens komplexe Abhängigkeiten zwischen Fehlern und Bearbeitung in Produktionssystemen frühzeitig zu entdecken und diese mit KI-Methoden zu erklären, validiert. Das Ergebnis ist ein dokumentiertes Vorgehen zur Evaluation sowie deren Evaluationsergebnisse.

2.1.3 Think Tank „Daten“ (AP 3)

Ziel des Think Tanks „Daten“ (AP 3) war die Umsetzung der Sammlung, Verwaltung, Bewertung, Nutzung und das Sharing von Daten in der intelligenten Produktion innerhalb der dafür bestehenden rechtlichen Rahmenbedingungen unter Berücksichtigung damit verbundener Potenziale und Herausforderungen. Dabei sollten Konzepte, Methoden und Rahmenbedingungen für das Data Sharing und für die Datenbewertung sowie ein Überblick über relevante Datenformate und Optionen für die Datenverwaltung erarbeitet werden.

State-of-the Art Analyse (UAP 3.1)

Ziel des UAP 3.1 ist es, durch eine gründliche State-of-the-Art Analyse zu den unterschiedlichen im Think Tank „Daten“ bearbeiteten Themen eine Grundlage für die Arbeiten zu legen, Innovationspotentiale und Forschungs- und Entwicklungslücken zu identifizieren sowie die Innovation der entwickelten Methoden sicherzustellen.

Die Analyse des State-of-the Art im Bereich Data Sharing unterteilt sich in zwei Stränge. Einerseits wurde der aktuelle Stand der wissenschaftlichen Literatur analysiert. Andererseits wurden, getrieben durch Praxisrelevanz des Themas, existierende technische Lösungen analysiert und gegenübergestellt. Im Bereich Data Sharing erfolgte dabei unter anderem die Analyse und Gegenüberstellung dreier im Rahmen der International Data Spaces entwickelten Frameworks. Diese beschäftigen sich mit der technischen Durchsetzung von Nutzungsrichtlinien während des Datenaustauschs. Im Kontext des Datenmarktplatzes wurden verfügbare Datenmarktplatzlösungen und deren Charakteristika identifiziert.

Da sich die Bereitschaft zum Data Sharing über Unternehmensgrenzen hinweg in der Interaktion mit den beteiligten Unternehmen und auch in der von IIP-Ecosphere mit 73 Unternehmen durchgeführten Umfrage als kritischer Punkt herausgestellt hat, wurden auch Alternativen zum Data-Sharing betrachtet. Dabei geht es darum kontrolliert Wissen zu teilen, ohne die Daten selbst zu teilen. Hierzu wurden ebenfalls die Literatur gesichtet, insbesondere im Bereich der Methode des föderierten Lernens (Federated Learning), bei dem Modellparameter statt Daten geteilt werden.

Im Bereich der rechtlichen Begleitforschung wurden zu Beginn der Projektlaufzeit von IIP-Ecosphere eine Literatur-Recherche durchgeführt, die als Grundlage für eine juristische Ersteinschätzung und Schwerpunktsetzung diene. Als Forschungsschwerpunkte wurden in diesem Zusammenhang Fragen der Rechtsgebiete Datenschutz, Wettbewerbs- und Kartellrecht und Geistiges Eigentum identifiziert.

Rechtliche Begleitforschung (UAP 3.2)

Für die rechtliche Begleitforschung wurden die Voraussetzungen und möglichen Risiken beim Sharing von Daten im Kontext der intelligenten Produktion analysiert. Bei der Analyse der in UAP 3.1 herausgearbeiteten Forschungsfragen ergab sich, dass neben der Frage, ob die Daten im Einzelfall Personenbezug aufweisen und dementsprechend die Regelungen der Datenschutz-Grundverordnung (DSGVO) sowie anderer Datenschutzgesetze Anwendung finden, auch die Verhältnisse bezüglich eines möglichen Dateneigentums und der Übertragung von Nutzungsrechten zu beleuchten waren, die je nach Argumentation im Ergebnis eine Anwendung des Material- bzw. Immaterialgüterrechts zur Folge haben. Außerdem wurde auf die Möglichkeit von wettbewerbs- und kartellrechtlich relevanten Rückschlüssen durch Produktions- und Betriebsdaten hingewiesen. Daneben wurden der Verlust und Diebstahl der zur Verfügung gestellten Daten als hohe unternehmerische Risiken festgestellt, die die Erforderlichkeit technischer Maßnahmen zur Anonymisierung der Daten begründen.

Einen bedeutenden Teil der Analyse stellte die durch den Think Tank Daten durchgeführte, umfassende Befragung von Partnern und Assoziierten dar, deren Ergebnisse Aufschluss über den juristischen Wissens- und Erfahrungsstand, relevante Problemfelder und die Vorstellungen über die von AP 3 erzielten Ergebnisse gaben. Darüber hinaus trug der Austausch im Arbeitskreis/CoP „Rahmenbedingungen für Data Sharing“ zu weiteren Erkenntnissen bei der Erarbeitung von Lösungsansätzen für die identifizierten Herausforderungen bei. Der dadurch bestehende Überblick wurde im Rahmen (digitaler) Workshops

erweitert, indem die Vermeidung von Personenbezug verarbeiteter Daten (z. B. durch die Verwendung synthetisch hergestellter Daten zu Trainingszwecken), (juristische) Verantwortlichkeiten und mögliche Use-Cases im Bereich des Geistigen Eigentums thematisiert wurden.

Neben der Nutzung anonymisierter oder synthetischer Daten wurde insbesondere die Methode des „Federated Learning“ als Privacy-Enhancing-Technology (PET) untersucht, mit der ein datenschutzrechtskonformes Training von KI-Modellen erfolgen kann [PUS22].

Die gesammelten Erkenntnisse konnten bei der Festlegung der rechtlichen Rahmenbedingungen der Verarbeitung von Daten im Projekt, vor allem auf der IIP-Ecosphere-Plattform, eingesetzt werden. Im Laufe von IIP-Ecosphere wurde die rechtliche Begleitforschung außerdem einer Regulierung durch EU-Gesetzgebungsinitiativen, insbesondere durch den EU-Data-Act, den EU-AI-Act und den Data Governance Act angepasst. Auch die EU-Gesetzentwürfe des „Digital Services Act“ und des „Digital Markets Act“ wurden dabei berücksichtigt.

Konzepte für Datenschutz und Datensicherheit (UAP 3.3)

Ziel des UAP 3.3 ist es, zu untersuchen, wie in einem Ökosystem der Schutz und die Sicherheit der dort zur Verfügung gestellten Daten gewährleistet werden kann. Dabei sind sowohl technische als auch organisatorische Maßnahmen zu berücksichtigen und im Hinblick auf ihre Machbarkeit im IIP-Ecosphere Kontext zu bewerten.

Als Startpunkt der Arbeiten erfolgte die Identifikation möglicher Herausforderungen im Hinblick auf den Datenschutz und die Datensicherheit in Datenökosystemen unter Berücksichtigung der konkret in IIP-Ecosphere geplanten Instrumente und Anwendungsfälle. Auf Basis der Herausforderungen wurden anschließend konkrete organisatorische und technische Maßnahmen respektive Instrumente zu Datenschutz und Datensicherheit herausgestellt und analysiert. Insbesondere wurden dabei folgende Maßnahmen und Instrumente betrachtet: Mechanismen aus den International Data Spaces (insbesondere Konnektoren und Usage Control), Architekturen von Plattformökosystemen, Identitätsmanagement, (Data) Governance in Ökosystemen, Umgang mit personenbezogenen Daten (Anonymisierung und Pseudonymisierung) sowie Datenrisikomanagement. Für einzelne Aspekte, wie z. B. für Plattformarchitekturen, wurden mögliche Design-Optionen gegenübergestellt und deren Vor- und Nachteile analysiert.

Zur Evaluation und Bewertung einzelner Maßnahmen und Instrumente im Kontext des IIP-Ecosphere Projekts wurde das Konsortium im Rahmen einer gemeinsam mit der rechtlichen Begleitforschung durchgeführten Umfrage eingebunden. Auf Basis der Umfrage konnten im Anschluss Präferenzen hinsichtlich der Auslegung einzelner Instrumente und Werkzeuge in Bezug auf Datenschutz und Datensicherheit abgeleitet und Handlungsempfehlungen hinsichtlich der IIP-Ecosphere Plattform, des Datenmarktplatzes und der Anwendungsfälle aufgestellt werden.

Die getätigten Recherche- und Umfrageergebnisse sowie Gestaltungs- und Handlungsempfehlungen wurden mit dem Whitepaper „Datenschutz und Datensicherheit in Datenökosystemen“ [BRU21] publiziert und dem Konsortium sowie der interessierten Öffentlichkeit umfassend zur Verfügung gestellt. Die Ergebnisse und Handlungsempfehlungen liefern dabei grundlegende Anforderungen für die Gestaltung und Auslegung des prototypischen IIP-Ecosphere Datenmarktplatzes (UAP 3.5). Nicht zuletzt sind die Handlungsempfehlungen aber auch für das Design weiterer Datenökosysteme in der Domäne der industriellen Produktion nutzbar.

Bewertung von Daten (UAP 3.4)

Ziel des UAP 3.4 ist es, Methoden zur Bewertung von Daten zu entwickeln, die als wichtiger Faktor zu einer sozialen Datenmarktwirtschaft beitragen können, indem sie eine faire Preisgestaltung von Daten als wirtschaftlicher Ressource ermöglichen.

Im Bereich der Datenbewertung wurde zunächst eine State-of-the-Art Analyse durchgeführt. Die Nutzung moderner datenbasierter KI-Methoden schafft einen Bedarf an hochwertigen Daten, da die Effektivität von ML-Methoden stark von der Qualität der verfügbaren Daten abhängt, mit denen sie für eine bestimmte Zielaufgabe trainiert werden können. Bei der Nutzung der Daten aus externen Quellen stellt sich die Frage nach dem Wert der Daten. Die Bewertung von Daten externer Partner ist nicht trivial. In der Literatur werden zur Bewertung von Daten eine Reihe von Ansätzen vorgeschlagen. Dabei soll der Wert von Daten für das Training von ML-Modellen quantifiziert werden. Zu diesen Methoden gehören z. B. Methoden zur effizienten Berechnung von Shapley-Werten und Reinforcement-Learning-Ansätze zur Schätzung der Beiträge einzelner Datenpunkte. Es gibt auch eine große Menge an Literatur, die sich mit der Bewertung von Beiträgen innerhalb von föderierten Lernkonfigurationen beschäftigt, bei denen mehrere Teilnehmer, die ihre Daten privat halten wollen, ein zentralisiertes Modell trainieren, indem sie lokal berechnete Parameter-Updates bereitstellen. Auf der Grundlage der durchgeführten State-of-the-Art Analyse wurde sich ein Überblick zum Thema Datenbewertung verschafft. Der Überblick über die aktuellen Methoden zur Bewertung von Daten wird dabei durch eine erweiterte Analyse dieser Methoden ergänzt, welche deren praktischen Nutzen erhöht: i.) Herausforderungen in Bezug auf Implementierung und Rechenaufwand; ii.) praktischer Leitfaden für die Anwendung der diskutierten Methoden in der Praxis; iii.) Fahrplan in Bezug auf die Anwendung der diskutierten Ansätze auf die anderen Paradigmen des maschinellen Lernens z. B. des verstärkten und unüberwachten Lernens.

Die gemeinsame Nutzung von Daten durch föderiertes Lernen ist ein attraktives Modell, da hierbei die Daten gegenüber von anderen Formen des Data Sharing nicht preisgegeben werden müssen. Aus diesem Grund wurde das Thema Datenbewertung für diese Form des Lernens weiter untersucht. Beim föderierten Lernen ist es besonders im Produktionsumfeld von großem Interesse einen fairen Wert der bereitgestellten Daten zu ermitteln, um Maschinenbetreiber zur Teilnahme zu motivieren. Grundsätzlich gibt es zur Datenbewertung beim föderierten Lernen zwei verschiedene Ansätze. Beim ersten Ansatz erhalten alle Teilnehmer nach erfolgtem Training, dasselbe Modell als Ergebnis und mittels einer Metrik wird ermittelt, wie groß der Beitrag der einzelnen Datensätze dazu war. Zu diesem Zweck wurden bereits verschiedene Metriken vorgeschlagen (z. B. Bewertungsregeln, Peer Prediction, Bayesian Truth Serum oder Correlated Agreement). Anschließend schaffen die Teilnehmer entsprechend ihres Beitrages einen monetären Ausgleich. Ein anderer Ansatz ist es eine intrinsische Motivation zu schaffen, nützliche Daten bereitzustellen, indem die Teilnehmer am Ende nicht alle dasselbe Modell erhalten, sondern eines das ihrem individuellen Beitrag entspricht. Die mit föderiertem Lernen erstellten Modelle sind in der Regel besser als wenn jeder Teilnehmer unabhängig ein Modell nur auf seinen Daten trainiert hätte. Auf diese Weise bietet das föderierte Lernen allen Teilnehmern einen Vorteil ohne Gefahr zu laufen, sensible Daten preiszugeben. Wegen des bereits genannten Mangels an öffentlichen Datensätzen existiert bisher jedoch keine Untersuchung dazu, welches Verfahren zur Datenbewertung im Produktionsumfeld am besten geeignet ist.

Neben der Bewertung von Daten haben wir uns auch mit der Generierung von Daten beschäftigt, einem Thema, das ebenfalls eng mit der Verfügbarkeit von Daten und dem Data Sharing verbunden ist. Wie bereits beschrieben, sind besonders aus dem Bereich der Produktion bisher nur wenige für föderierte Verfahren geeignete Datensätze öffentlich verfügbar. Denn selbst wenn Unternehmen solche Daten erheben, haben sie nur eine geringe Neigung diese zu veröffentlichen, aus Sorge Wettbewerbsvorteile zu verlieren. Öffentlich verfügbare Datensätze sind für den wissenschaftlichen Diskurs aber unerlässlich, um ML-Methoden vergleichen zu können. Einen Beitrag, um dieses Problem zu lösen, können synthetische Datensätze sein z. B. auf der Basis von CAD-Modelle typischer Bauteile und Werkstücke. Hierzu wurden im Projekt erste Arbeiten durchgeführt.

Weiterhin wurde auch das Thema Datenbewertung für die wichtige KI-Methode Reinforcement Learning betrachtet. In vielen realen Szenarien wie der intelligenten industriellen Produktion sind sowohl Kosten als auch Risiken mit dem Sammeln von Proben für das Training von Reinforcement Learning Agenten verbunden. Das Gebiet des Batch Reinforcement Learning geht diese Probleme an, indem es die Sammlung von Daten an einen Fachmann oder ein sorgfältig überwachtes Programm auslagert und anschließend nach einer optimalen Strategie sucht. Obwohl Ansätze wie Batch Constrained Deep Q-learning (BCQ) vielversprechend sind, sind sie auf sorgfältig konstruierte Datensätze angewiesen, die gut auf die beabsichtigten Zieldomänen abgestimmt sind. Die Erstellung eines geeigneten Datensatzes kann kostspielig sein und erfordert Fachwissen.

Mit dem Aufkommen von Datenmärkten besteht eine Alternative zur Erstellung eines eigenen Datensatzes darin, vorhandene Daten zu kaufen. Dies wirft jedoch Fragen hinsichtlich der Übertragbarkeit und Robustheit eines auf diesen Daten trainierten Agenten auf. In den letzten Jahren gab es eine Reihe von Bemühungen innerhalb des Paradigmas des überwachten Lernens zur Bewertung von Daten, einschließlich Data Shapley und Data Valuation mit Reinforcement Learning (DVRL). Wir haben evaluiert, inwieweit DVRL verwendet werden kann, um hochwertige Proben für das Training eines Batch Constrained Deep Reinforcement Learning Agenten für eine bestimmte Zieldomäne zu identifizieren. Wir führen eine empirische Evaluierung in drei MuJoCo-Domänen durch, wobei wir Szenarien untersuchen, in denen Quell- und Zieldomäne identisch sind. Zudem haben wir auch Szenarien des Transferlernens (erforderlich in AP 4.5) untersucht, bei denen sich die Parametrisierung Quell- und Zieldomäne unterscheiden. Die Evaluierung ergab, dass unser entwickelter Ansatz, DVRL-BCQ, andere Ansätze in diesem Bereich wie Vanilla BCQ, BAER, CQL und TD3-BC übertrifft. Die Ergebnisse flossen in den Artikel [ABO22] ein.

Zusätzlich zur Untersuchung von technischen Methoden zur Datenbewertung wurden ebenso organisatorische bzw. bilanzierungsrechtliche Rahmenwerke zur Bemessung des Datenwertes untersucht. Aus organisatorischer Perspektive wurde zunächst Datenwert als multidimensionales Konstrukt definiert, dass es ermöglicht die Potentiale eines bestimmten Datenobjekts für die Organisation zu bemessen. Anschließend erfolgte die Ermittlung von Dimensionen und möglichen Metriken von Datenwert im organisatorischen Kontext. Neben rein ökonomischen Metriken können bspw. Inhalt, Seltenheitswert, Nutzung, Nützlichkeit oder Umfang des Datensatzes eine Rolle bei der Bewertung von Daten darstellen. Parallel wurde bilanzierungsrechtlicher Perspektive der Datenbewertung betrachtet. Hier erfolgte eine Analyse des aktuellen Status und der zu erwartenden Entwicklungen. Abschließend erfolgte die Bewertung auf die Anwendbarkeit der etablierten Verfahren auf den Kontext von Datenmarktplätzen.

Insgesamt zeigte sich, dass weder organisatorische noch bilanzierungsrechtliche Mechanismen geeignet sind, um eine faire Preisfindung von Daten auf einem Datenmarktplatz zu etablieren. Dies ist insbesondere durch die speziellen Charakteristika von Daten als Kredenzgut, dass sich mit marginalen Grenzkosten vervielfältigen lässt, bedingt. Zudem hängt der Wert von Daten für ein Unternehmen stark vom anvisierten Anwendungsfall ab.

Entsprechend stellen marktbasierende Methoden zur Datenwertermittlung eine vielversprechende Methode dar, um den „realen“ Datenwert zu ermitteln. Dies ist beispielsweise bereits für existierende Datenmärkte und Datenbroker im Bereich von Personen- oder Finanzdaten gängige Praxis. Einzelne Marktplatzmechanismen, wie beispielsweise Verhandlung und „Competitive Pricing“ sollen im Rahmen der Implementierung des IIP-Ecosphere Datenmarktplatz umgesetzt werden. Neben den Vorarbeiten zum Datenmarktplatz lieferten die im Rahmen von IIP-Ecosphere erzielten Ergebnisse zudem die initiale Grundlage für einen durch das Fraunhofer ISST publizierten Bericht im Rahmen des Projekts IEDS zum Status quo der Datenbewertung [HUP22], der insbesondere die bilanzierungsrechtliche Perspektive fokussiert.

Kontrolliertes Data Sharing und Daten-Marktplatz (UAP 3.5)

Ziel des UAP 3.5 ist die Entwicklung geeigneter Lösungen für das Data Sharing für den Produktionskontext. Dies beinhaltet insbesondere die iterative Entwicklung einer prototypischen Datenmarktplatzlösung für das IIP-Ecosphere Ökosystem. Neben dem Datenmarktplatz wurde in IIP-Ecosphere auch an dem Thema förderiertes Lernen als gangbare Alternative zum tatsächlichen Teilen von Daten gearbeitet.

Datenmarktplatz: Basierend auf der Erstellung von Anwendungsfällen und User Stories sind verschiedene Module des Datenmarktplatzes (darunter Angebotsportal, Transaktionslog, Abrechnungsmodul, Benutzeroberfläche, Konnektoren und Mechanismen zu Usage Control) zu implementieren.

Zur Konzeptionierung des Datenmarktplatzes wurden User Stories über verschiedenste Kanäle aufgenommen, priorisiert und kontinuierlich auf deren Umsetzbarkeit geprüft sowie ggf. verfeinert. Zu diesen Kanälen zählen u.a. die Konsortialbefragung im Rahmen des UAP 3.2, die Spiegelung mit anderen Projektteilnehmern im Rahmen bilateraler Absprachen und der Konsortialtreffen, die Einbindung externer Interessenten aus dem Industrial Advisory Board und Beiträge im Rahmen der Open Core Sessions. Initial vorgesehene Praxisformate (bspw. Lego Serious Play) wurden dabei zum Teil durch die Beschränkungen in der Corona-Pandemie verhindert und durch entsprechende virtuelle Formate ersetzt. Die Anforderungen aus Sicht der wirtschaftlichen Verwertung des Angebots wurden durch interne Value Proposition Workshops identifiziert.

Basierend auf den Recherchearbeiten und dem Dialog mit potentiellen Anwendern und Stakeholdern wurden folgende Design-Prinzipien definiert: 1) Vertrauenswürdigkeit, 2) Datensouveränität 3) Dezentralisierung; 4) Förderiert, 5) Open-source, 6) Ease-of-Use und z) Standardkonformität.

Da der Aufbau eines Datenmarktplatzes von Grund auf mit einem enormen Aufwand verbunden und im Rahmen der zur Verfügung stehenden Mittel nicht leistbar ist, wurden auf Basis des ersten stabilen Satzes an User Stories Recherchen zur Identifizierung von nutzbaren open-source Basistechnologien als Grundlage für den Datenmarktplatz begonnen. Insgesamt wurden 17 Softwarelösungen genauer

betrachtet. Basierend auf den genutzten Technologien, Lizenzen, Funktionsumfang und weiteren Qualitätskriterien von open-source Softwareprojekten wie Wartbarkeit, Portabilität und Nachhaltigkeit wurde schlussendlich die E-commerce Plattform Mailchimp Open Commerce (ehem. Reaction Commerce) ausgewählt und als Grundlage für die Umsetzung des „IIP-Ecosphere Data Marketplace“ verwendet. Dabei ist festzuhalten, dass Mailchimp Open Commerce hier nur als rudimentäre Basiskomponente verwendet werden konnte und zur Umsetzung eines föderierten Datenmarktplatzes umfangreiche Anpassungen an der Mailchimp Open Commerce selbst und die Entwicklung weiterer Anwendungskomponenten (s. Systemsicht unten) notwendig wurden. Die Entwicklungsarbeiten wurden basierend auf den User Stories in iterativer Form durchgeführt. Die Erweiterung der Mailchimp Open Commerce Plattform umfasste insbesondere die Entwicklung von Plugins für digitale Güter, für digitales Fulfillment via Distributionen, für dynamische Metadaten Templates, für Angebotsanfragen, für die Zugriffsregelung auf die Angebote sowie für Anbieter- und Produktrezensionen. Außerdem wurde die Preismodellierung und die Suchfunktionalität erweitert (um Faceted Search) sowie ein Modul für die Generierung der Rechnungen und ein Zahlungsplugin für Banküberweisung entwickelt.

Das Webinterface des Shops wurde auf dem Framework Vue-Storefront 2 aufgebaut und erweitert, um die oben aufgeführten Funktionalitäten zu unterstützen. Die Weboberfläche der Administrationsanwendung wurde dagegen von Grund auf entwickelt.

Die resultierende Softwarekomponente „IIP-Ecosphere Data Marketplace“ wird folgend aus zwei Perspektiven beschrieben. Die Systemperspektive erläutert die Zusammensetzung des IIP-Ecosphere Data Marketplace aus technischer Sicht, geht auf die einzelnen Komponenten ein und beschreibt die genutzten Technologien. Die funktionale Sicht beschreibt den konkreten Funktionsumfang des Datenmarktplatzes und die implementierten Features.

Systemsicht

Das Gesamtsystem IIP-Ecosphere Data Marketplace besteht aus insgesamt fünf Teilkomponenten. Eine Übersicht der Komponenten ist in Tabelle 1 dargestellt. Deren Zusammenhang wird folgend beschrieben.

Tabelle 1: Komponenten des IIP-Ecosphere Data Marketplace (Systemsicht)

Komponente	Beschreibung	Technologien
Admin application	Die Admin application ist eine Browseranwendung, die vom Marktplatz-Administrator verwendet wird, um Daten-Assets, Benutzer und den Marktplatz selbst zu managen.	Nuxt 3, Vue 3, Vuetify 3, TypeScript
Storefront	Die Storefront (SF) application ist eine Browseranwendung, welche die Shop-Ansicht für externe des Marktplatzes darstellt. Potentielle Kunden können dort die Angebote durchsuchen, finden und bestellen.	Nuxt 2, Vue-Storefront 2, Tailwindcss 3, TypeScript

Broker	Der Broker ist eine Browseranwendung, die den Zusammenschluss von Datenmarktplätzen ermöglicht, indem Datenangebote von registrierten Datenmarktplätzen durchsucht und unter Einhaltung von Zugangsberechtigungen einheitlich dargestellt werden.	Nuxt 3, Vue 3, ElementUI, TypeScript
Backend Server	Der Backend-Server enthält die Geschäftslogik eines Datenmarktplatzes und stellt eine GraphQL-API zur Verfügung, die von Admin- und SF-Anwendungen genutzt wird, um relevante Informationen darstellen zu können.	Nodejs 14, Mailchimp open commerce (ehem. Reaction Commerce)
Metadata Storage	Der Metadata-Storage speichert die Metadaten über Datenangebote, kontrollierte Listen und Templates zur Bereitstellung für den Datenmarktplatz.	MongoDB

Abbildung 11 zeigt schematisch die Zusammensetzung der Systemkomponenten. Die Admin App und die Storefront (SF) App sind SSR (Server Side Rendering) Webapplikationen. Sie bestehen jeweils aus einem Server, der die Kommunikation mit dem Backend Server abstrahiert und das entsprechende Webinterface zur Ausführung in einem Browser zur Verfügung stellt. Der Admin- und der SF-Server greifen über eine GraphQL-Schnittstelle auf die Geschäftslogik im Backend-Server zu. Der Broker ist ebenfalls eine SSR Web Applikation und stellt eine optionale Systemkomponente dar. Mehrere Broker-Anwendungen können über die GraphQL-Schnittstelle die Angebote von mehreren Backend-Servern aggregieren und den Zugriff auf die Assets autorisieren.

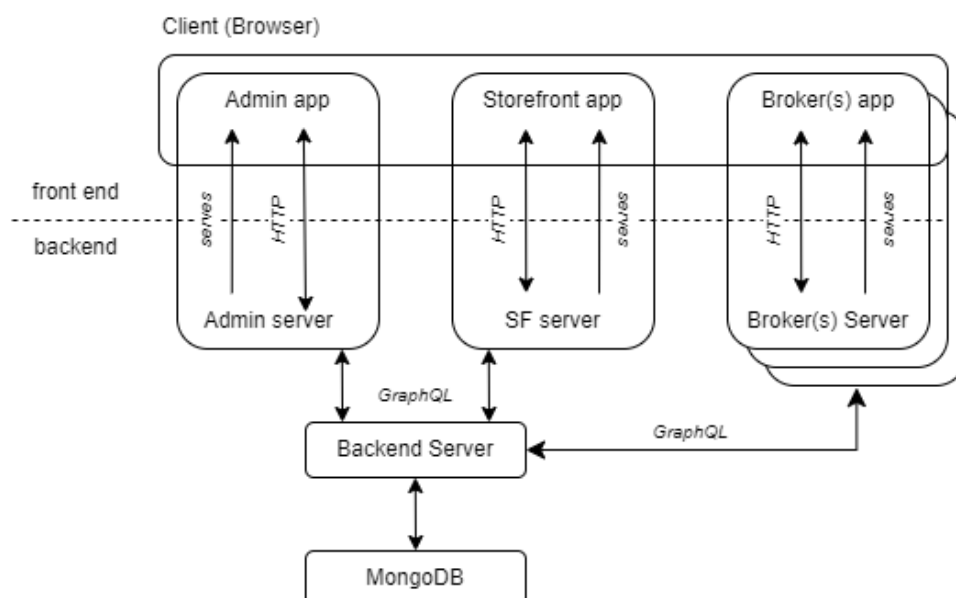


Abbildung 11: Interaktion der Systemkomponenten (Datenmarktplatz)

Damit auch Unternehmen mit geringen Ressourcen den Datenmarktplatz schnell und einfach produktiv einsetzen können, wird die Gesamtanwendung als Docker-Compose Konfiguration, bestehend aus den einzelnen Images, bereitgestellt.

Funktionale Sicht

Wie bereits zuvor beschrieben, handelt es sich beim IIP-Ecosphere Data Marketplace nicht um eine zentralisierte, sondern um eine föderierte Anwendung, die es Unternehmen der industriellen Produktion ermöglicht selbstbestimmt und transparent ihre Daten zu teilen. Zur Listung der von Datenangeboten stellt ein Industrieunternehmen (Datenanbieter) einen eigenen Marktplatz bestehend aus Storefront und Admin App bereit und richtet diese ein. Datenangebote werden hier nur mittels Metadaten beschrieben, d.h. die Daten verbleiben bis zum eigentlichen Austausch in der Umgebung des Anbieters. Zudem kann der Datenanbieter bestimmen, wer am Datenmarktplatz teilnehmen und das Angebot sehen kann. Für den initialen Einsatz des Datenmarktplatzes wird davon ausgegangen, dass zunächst Partner aus bestehenden Geschäftsbeziehungen des Industrieunternehmens (Tier 1 Supplier, Tier 1 Customer) ein hohes Interesse an den Daten besitzen und der Datenmarktplatz zunächst an diesen Partnern ausgerichtet werden sollte. Insbesondere erleichtert dies den Aufbau von Vertrauen zwischen den Parteien. Eine Skalierung des Datenmarktplatzes, beispielsweise über das gesamte Wertschöpfungsnetzwerk hinweg, kann im zweiten Schritt über die Broker-Komponente erfolgen. Diese aggregiert die Angebote registrierter Datenmarktplätze und kann entsprechend über ein höheres Angebot verfügen. Eine Zusammenfassung der beschriebenen Rollen ist in Abbildung 12 dargestellt.

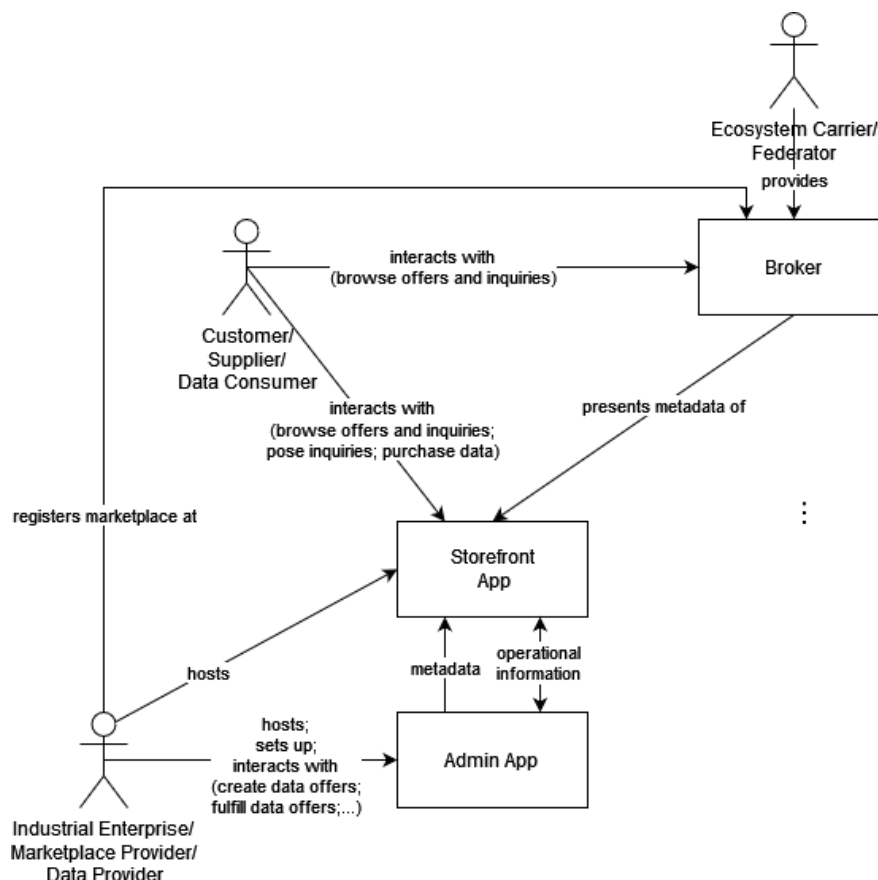


Abbildung 12: Rollenmodell des Datenmarktplatzes aus funktionaler Sicht

Der IIP-Ecosphere Data Marketplace ermöglicht in seiner aktuellen Form als Minimum Viable Product die Ende-zu-Ende Abwicklung eines Datentransfers auf Metadatenebene. Die einzelnen implementierten Funktionalitäten der Prozessschritte werden folgend komprimiert erläutert.

Die Erstellung und Konfiguration des Stores stellt einen notwendigen Schritt zur Bereitstellung des Datenmarktplatzes dar. Dieser erfolgt in der Admin App durch den Datenanbieter und entsprechend auch Store-Owner. Dieser Prozessschritt beinhaltet die Festlegung grundlegender Eigenschaften wie Name und Währung des Stores, die Beschreibung von Kontaktinformationen und rechnungsrelevanter Informationen, und die Festlegung möglicher Fulfillment-Wege und deren Preise. Letzteres ermöglicht die Internalisierung entstehender Kosten, die ggf. durch die Aufbereitung und Bereitstellung der Daten anfallen. Falls gewünscht, kann der Datenmarktplatz zur Aggregation durch einen Broker freigegeben werden.

Datenangebote werden ebenfalls durch den Datenanbieter in der Admin-App (*Angebotsportal*) erzeugt (siehe Abbildung 13). Dazu müssen drei Konzepte durch den Datenanbieter bedient werden. Ein Data Asset ist eine abstrakte Beschreibung der angebotenen Daten und kann mehrere Varianten besitzen. Eine Variante beschreibt eine konkrete Sammlung von Daten in Bezug auf das übergeordnete Data Asset. Beispielsweise können Varianten mit personenbezogenen Daten existieren, die in der Forschung verwendet werden können, und solche bei denen Personendaten anonymisiert wurden. Schlussendlich muss zu jeder Variante mindestens eine Distribution existieren, welche den Zugriff auf eine Variante technisch realisiert. Beispielsweise kann eine Datenbank über die Weitergabe von Zugangsinformationen oder den Transfer eines Backups zur Verfügung gestellt werden.

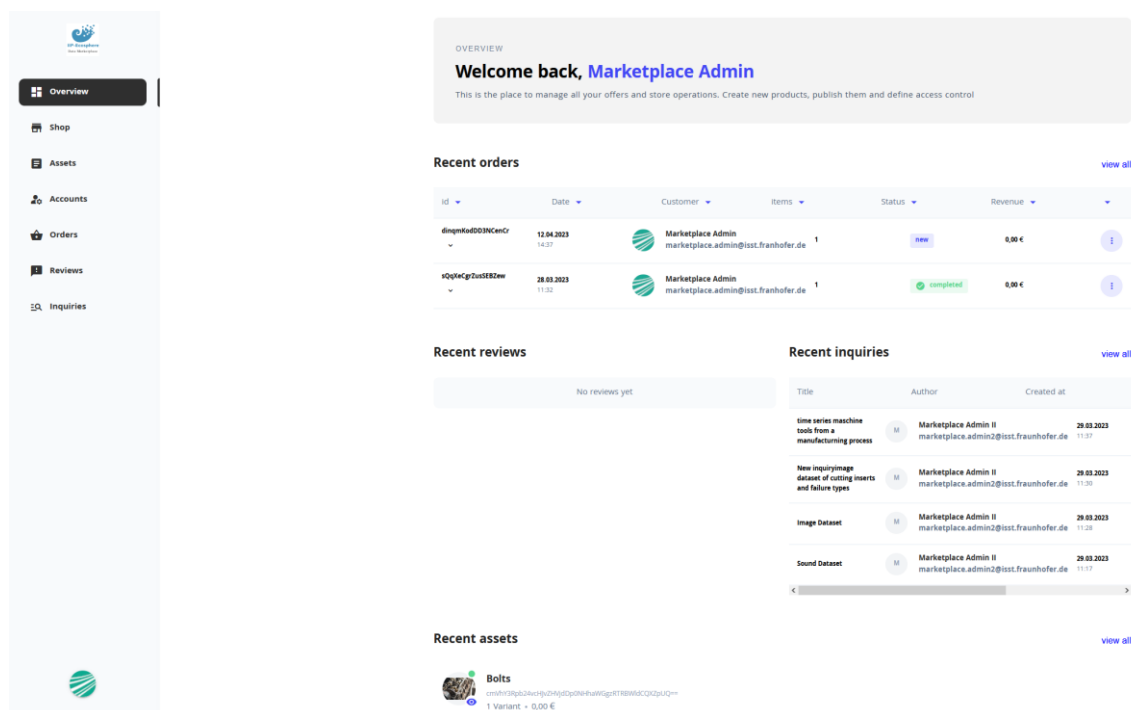


Abbildung 13: Oberfläche der Admin App (Datenmarktplatz)

Auf jeder Konzeptebene erhalten die Datenanbieter die Möglichkeit verschiedene verpflichtende und optionale Attribute zu pflegen. Dazu gehören insbesondere:

- **Templates:** Um die Erstellung von Datenangeboten zu erleichtern und deren Vergleich und Filterung während des Datenfindungsprozesses zu ermöglichen wurden für den Datenmarktplatz Attributsammlungen bereitgestellt, die für die ganzheitliche Bestimmter Data Asset-Typen verwendet werden können. Beispiele für implementierte Templates sind „Labeled Image Data“ und „Asset Administration Shell“. Hierbei wurde entsprechend den Anforderungen der industriellen Produktion Rechenschaft getragen.
- **Datennutzungsbedingungen:** Definition von Datennutzungsrichtlinien (*Usage Control Policies*) für fachliche Nutzer und technische Nutzer durch Beschreibung der relevantesten Policies aus fachlicher Sicht oder in der auch in den International Data Spaces-Standardisierungsaktivitäten geförderten ODRL-Policy-Sprache. Hierbei wurden für die in der Konsortialbefragung als wichtig identifizierten Nutzungsbedingungen Templates definiert.
- **Preise:** Beschreibung verschiedener denkbarer Preismodelle wie bspw. „frei Verfügbar“, „Einmalzahlung“, „Nutzungsbasiert“ und „Abo“ und Festlegung der Preise.
- **Sichtbarkeit:** Einstellung der Sichtbarkeit der Daten Assets für spezifische User (zwei Optionen: sichtbar ausschließlich für X, verbergen spezifisch für Y). Dies ist insbesondere für die industrielle Produktion relevant, da teilweise von den Metadaten auf die Struktur von Produkten (beispielsweise eingesetzte Sensorik) geschlossen werden kann. Denkbar ist es diese Funktionalität zukünftig in Attributbasierte Zugangskontrollen auszuweiten.

Die Erzeugten und freigegebenen Datenobjekte können durch Datenkonsumenten in dem spezifischen Store eines Datennutzers oder mithilfe des Brokers in einer aggregierten Sicht gefunden werden. Angebotserstellung und Angebotsevaluation erhalten somit separate *Benutzeroberflächen*. Dazu müssen die Nutzer mit einem vorher durch den Datenanbieter vergebenen Benutzer angemeldet sein und die notwendigen Stammdaten wie bspw. Adresse vergeben haben. Zur einfachen Findung existierender Datenangebote können zudem bekannte Mechanismen wie Suche und Filterung genutzt werden. Einzelne Datenangebote können in der Detailansicht noch einmal genauer untersucht werden.

Der Kaufprozess der Daten wird in der Storefront abgewickelt. Entsprechend müssen Datennutzer zum Kauf der angebotenen Daten aus dem Broker in die Storefront abstimmen. Zum Kauf der Daten werden aus herkömmlichen E-commerce Aktivitäten bekannte Mechanismen, wie beispielsweise ein virtueller Einkaufswagen genutzt. In der derzeitigen Version unterstützt der mit dem Rechnungskauf die in Deutschland am häufigsten genutzte Zahlungsweise im B2B-Geschäft. Zur Durchführung der Transaktion muss diese durch den Datenanbieter bestätigt werden. Der Datenanbieter kann die Daten entweder direkt an den Käufer außerhalb des Datenmarktplatzes bereitstellen oder den Fulfillment-Prozess des Datenmarktplatzes nutzen und dort die Credentials bereitstellen. Dies ermöglicht zudem die langfristige Einsehbarkeit im *Transaktionslog*. Für die im Rahmen des Kaufprozesses vereinbarten finanziellen Rahmenbedingungen wird im Datenmarktplatz automatisch eine gesetzeskonforme Rechnung im *Abrechnungsmodule* erzeugt.

Neben der herkömmlichen angebotsbasierten Funktionsweise bietet der Datenmarktplatz auch die Möglichkeit Datenanfragen an den ursprünglichen Datenanbieter bzw. an ursprüngliche Datennachfrager zu stellen, welches eine Anforderung des Konsortiums darstellt. Datennachfragen werden in der Storefront formuliert und können durch einen Broker aggregiert werden. Da Datennachfragen meist weniger kon-

kret als Datenangebote formuliert werden, wurde an dieser Stelle ein leichtgewichtiges Verfahren implementiert. Datennachfragen können durch alle Teilnehmer gesichtet werden. Bei Interesse können diese im Rahmen eines implementierten Nachrichtenmoduls auf die Datennachfrage antworten, um bspw. Kontaktdaten oder weitere Informationen anzufragen. Ebenso kann der Datenanbieter direkt mit einem auf dem Datenmarktplatz existierenden Datenangebot auf die Nachfrage antworten, welche der potentielle Konsument annehmen oder ablehnen kann. Bei einer Annahme erfolgt das Fulfillment analog zum herkömmlichen Angebotsprozess.

Sollen *Konnektoren* nach dem Verständnis der IDSA als zusätzlicher Sicherheits- und Vertrauensmechanismus eingesetzt werden, so ist eine Anbindung an den Datenmarktplatz problemlos möglich. Ein Verpflichtender Einsatz von Konnektoren erhöht jedoch die Einstiegsbarrieren, insbesondere für kleine und mittelständische Unternehmen. Die Entwicklung eines eigenen IIP-Ecosphere Konnektors erschien insbesondere aus Interoperabilitätssicht nicht sinnvoll. Bereits jetzt existieren mehr als 20 IDS-Konnektoren, die nur teilweise in der Lage sind miteinander zu kommunizieren. Entsprechend integrierten sich die Arbeiten zu Konnektoren in IIP-Ecosphere nach Analyse der Konnektorenlandschaft in die Weiterentwicklung des Eclipse Dataspace Connector (nun Eclipse Dataspace Components [EDC]) als open-source Projekt bei der Eclipse Foundation. Dabei wurden insbesondere die Anforderungen der Verarbeitenden Industrie nach Selbstbestimmung und Einfachheit berücksichtigt, sodass die EDC als wichtige Grundlage für ein zukünftiges Manufacturing-X Ökosystem gesehen werden¹. Das Gesamtsystem IIP-Ecosphere Data Marketplace steht interessierten Parteien unter der GNU GPL Lizenz als open-source Software im github des Fraunhofer ISST zur Verfügung.

Kontrolliertes Data Sharing via Föderiertes Lernen: Eine Möglichkeit, ein verallgemeinertes Modell zu erhalten, besteht in der Zusammenführung von Produkt- und Prozessdatensätzen aus verschiedenen Fertigungsunternehmen (z. B. bei einem Maschinenbauer). Doch trotz der gut dokumentierten Vorteile der Zusammenführung von Daten zögern Unternehmen oft, Daten gemeinsam zu nutzen. Hier bietet das föderierte Lernen (FL) eine Alternative, die es Dateninhabern ermöglicht, gemeinsam ein globales Modell zu trainieren, ohne ihre Daten physisch zu teilen. Stattdessen trainieren die Teilnehmer/Unternehmen ein lokales Modell mit ihren lokalen Daten und senden dann die Parameter des trainierten Modells an einen zentralen Server. Der Server führt dann die Aggregation dieser Parameter durch, bevor er die aggregierten Parameter an die Teilnehmer zurücksendet, damit diese das globale Modell mit den aggregierten Parametern unter Verwendung ihrer eigenen Daten erneut trainieren. Nach mehreren Runden erreicht das globale Modell seine Konvergenz, ohne dass Daten an den Server gesendet werden müssen.

Im Projekt haben wir untersucht, inwieweit sich föderiertes Lernen für die intelligente industrielle Produktion eignet, indem wir Experimente mit CNNs auf Schalldaten durchgeführt haben, die in Log-Mel-Spektrogrammbilder umgewandelt wurden, und zwar mit einer großen Anzahl von Clients (68). Wir stellten fest, dass föderiertes Lernen bei diesem Datensatz ähnlich gut abschneidet wie zentralisiertes Lernen. Außerdem wurden Forschungsarbeiten im Bereich der Erhöhung der Sicherheit für das föderierte Lernen durchgeführt. Dabei wurde die Erkenntnis erlangt, dass in diesem Zusammenhang adversariales

¹ Bauplanstudie Manufacturing-X: https://www.iosb.fraunhofer.de/content/dam/iosb/iosbtest/documents/presse/Bauplanstudie-Manufacturing-X_dt.pdf

Training und Randomisierungstechniken eingesetzt werden können, um White-Box-Angriffe erfolgreich abzuwehren und gleichzeitig die hohe Leistung der trainierten Modelle zu erhalten.

Zudem wurden Aktivitäten durchgeführt, um das föderierte Lernen in die virtuelle IIoT Plattform zu integrieren, die in IIP-Ecosphere entstanden ist. Hier ist es Ziel, die Methoden des föderierten Lernens im Produktionsumfeld einfacher zugänglich zu machen.

Kontrolliertes Data Sharing via Datengenerierung: Kontrolliertes Data Sharing muss sicherstellen, dass keine ungewollte Information preisgegeben wird. Traditionelle Ansätze zum Schutz der Privatsphäre vor Weitergabe der Daten, bestehen in der Aggregation von Daten, dem Entfernen von Identifikatoren, dem Ändern von Quasi-Identifikatoren und dem Stören von Werten. Diese Methoden sind jedoch anfällig für Angreifer mit Hintergrundwissen, z. B. durch die Verknüpfung von Datensätzen mit anderen nicht de-identifiziert Datensätzen. Eine Alternative besteht darin, datenschutzfreundliche synthetische Daten zu generieren, die den realen Quelldaten ausreichend ähnlich sind und zum Trainieren von ML-Modellen mit einem minimalen Leistungsverlust verwendet werden können. In letzter Zeit gab es vielversprechende Bemühungen in diesem Bereich, die Generative Adversarial Networks (GANs) und deren Weiterentwicklungen zur Erfassung der Verteilung eines Datensatzes verwenden. Trotz der oben genannten Fortschritte bleiben Fragen bezüglich des Ausmaßes, in dem GANs anfällig für Angriffe auf Mitgliedschaftsinferenzen sind, unbeantwortet. Die Literatur unterscheidet vier Angriffsmodelle: i.) Vollständiger Black-Box-Generator, ii.) Partieller Black-Box-Generator, iii.) White-Box-Generator und iv.) Zugriff auf die gesamte Architektur des GANs. Von den vier Angriffsmodellen wurde der partielle Blackbox-Generator-Angriff in der Literatur selten berücksichtigt. Wir haben untersucht, inwieweit man einen kontinuierlichen Kontrollagenten darauf trainieren kann, geeignete Samples zu finden. Dazu wurde der latente Eingaberaum exploriert, um Samples zu finden, die sich gegebenen Zielsamples annähern. Dabei haben wir vielversprechende Ergebnisse erhalten, bei denen die Netze gelernt haben, Seed-Werte zu finden, die Zielbilder für einen Generator approximieren können, der auf MNIST und den Adult-Datensätzen trainiert wurde. Zusätzlich wurden im Bereich der Datengenerierung für kontrolliertes Data Sharing mit GANs auch Arbeiten zur Adressierung von Modus-Kollaps in synthetischen Datensätzen mit Generative Adversarial Network Games und Parallel Nash Memory durchgeführt.

Datenformate und Datenverwaltung (UAP 3.6)

Der effiziente Austausch von Informationen zwischen Maschinen, Edge-Geräten und maschinellen Lernmodellen ist für moderne intelligente industrielle Produktionsumgebungen unerlässlich. Um die Kommunikation zwischen Maschinen und Geräten zu erleichtern, wurde in den letzten Jahren eine Reihe offener Standards entwickelt. In einem Survey-Paper geben wir einen Überblick über die Datenformate, die heute in der Produktion am häufigsten verwendet werden. Insbesondere betrachten wir die praktischen Herausforderungen bei der Definition von Datenformaten, die für intelligente industrielle Produktionsumgebungen geeignet sind, mit einem kritischen Blick. Zu diesem Zweck geben wir einen ganzheitlichen Überblick über den Datenfluss in der intelligenten Produktion, von der Kommunikation zwischen Maschinen und Geräten bis hin zu den Schritten, die unternommen werden, um sicherzustellen, dass die Daten in einem brauchbaren Format für die Verarbeitung durch moderne maschinelle Lernalgorithmen zur intelligenten Entscheidungsfindung vorliegen.

Ein weiteres wichtiges Thema im Bereich der Datenformate sind der aktuell an Interesse gewinnende (entstehende) Standard der Verwaltungsschale (Asset Administration Shells). Das IIP-Ecosphere Team

hat sich in Zusammenarbeit mit den Partnern an der Universität Hildesheim und mit Partnern aus der Industrie in den Standard eingearbeitet und das Potential des Standards für IIP-Ecosphere und insbesondere für die IIP-Ecosphere Plattform untersucht. In diesem Kontext wurde im April 2022 vom IIP-Ecosphere Team auch ein Event zum Thema Verwaltungsschalen organisiert (siehe auch Abschnitt 1.9). Verwaltungsschalen sind inzwischen ein wichtiger Baustein der IIP-Ecosphere Plattform [EIC23b].

2.1.4 Think Tank „KI & Produktion“ (AP 4)

Das Ziel des Think Tank „KI& Produktion“ ist die anwendungsorientierte Erforschung und Entwicklung von KI-Methoden, ihrer Potenziale, und ihrer Anwendungsvoraussetzungen für die Produktion. Dabei bilden KI-basierte, selbstoptimierende Lösungen für die intelligente Produktion einen besonderen Schwerpunkt. Die Arbeiten sind dabei in verschiedene Schwerpunkte strukturiert. Das umfasst Methoden zur Selbstoptimierung (UAP 4.1), Ebenen-übergreifende Optimierung (UAP 4.2), Selbststeuernde Prozessketten und Qualitätsregelkreise (UAP 4.3), Interpretierbarkeit und Erklärbarkeit (UAP4.4), Transfer von Wissen und Modellen (UAP4.5) und Easy-to-Use KI (UAP4.6).

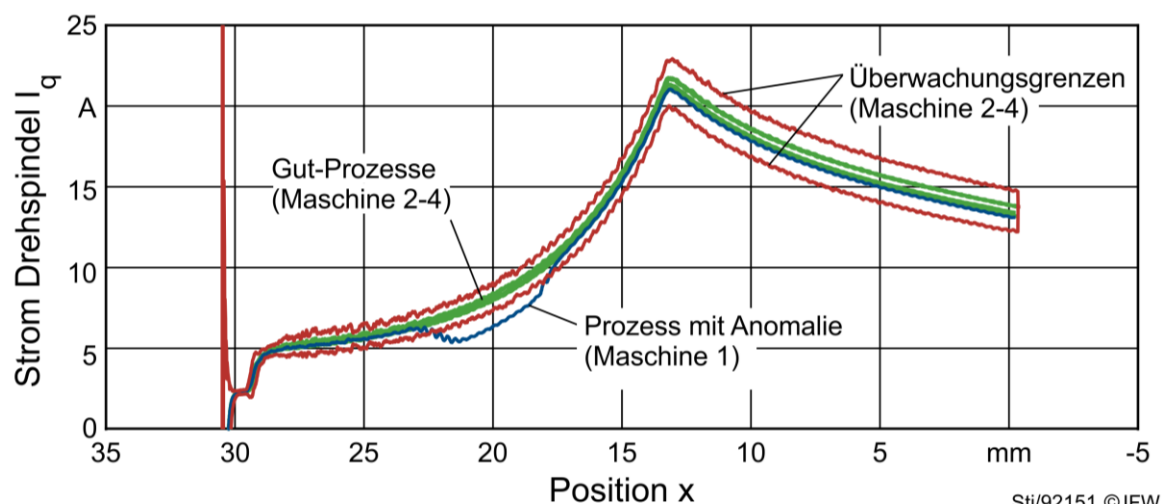
Methoden zur Selbstoptimierung (UAP 4.1)

Die Überwachung von Werkzeug-, Prozess- und Maschinenzuständen reduziert Kosten, welche durch Prozessfehler, unzureichende Qualität und plötzliche Maschinenausfälle entstehen können. Führen die Maschinen die Überwachung selbstständig mit hinreichender Leistung aus, optimieren diese Maschinen Qualitätsverluste und die Prozesssicherheit. Überwachung ist damit eine Voraussetzung für den (teil)autonomen Betrieb von Werkzeugmaschinen. Ziel dieses Unterarbeitspaketes ist es Methoden zu entwickeln, die Prozesse und Maschinen mithilfe des maschinellen Lernens autonom und erfahrungsbasiert überwachen und dieses Wissen auf weitere Prozesse und Maschinen übertragen. Dafür wurden zwei wesentliche Anforderungsszenarien identifiziert und betrachtet: Erstens, eine prozessparallele Überwachung reaktionszeitkritischer Prozessfehler in Zerspanprozessen. Diese stoppt den Prozess im Fall einer erkannten Anomalie, sodass der Bediener die Möglichkeit hat einzugreifen, bevor größerer Schaden entsteht. Somit werden schwere Maschinenschäden verhindern, die zu kostspieligen Reparaturen und langen, ungeplanten Stillstandszeiten führen. Das zweite Anwendungsszenario ist die Bestimmung von Werkzeugverschleiß ohne kritische Reaktionszeit. Die Verschleißbestimmung ist industriell relevant, da sie es ermöglicht Werkzeuge zu tauschen, bevor diese verschleißbedingt die Oberflächen von Werkstücken beschädigen oder sogar im Prozess brechen. Damit reduziert die Verschleißbestimmung Qualitätsverluste und trägt zum teilautonomen Betrieb von Werkzeugmaschinen bei.

Prozessparallele Überwachung reaktionszeitkritischer Prozessfehler

Für die prozessparallele Überwachung reaktionszeitkritischer Fehler in Zerspanprozessen wurde als Beispielprozess das Längs- und Querplandrehen ausgewählt. Die Prozesse werden in der Industrie sehr häufig eingesetzt, haben aber eine noch handhabbare Komplexität. Wesentliche Herausforderung beim Transfer von Wissen sind zeitliche Unterschiede in den Prozessabläufen der verschiedenen Maschinen, wie sie etwa durch unterschiedliche Bearbeitungsgeschwindigkeiten oder aufgrund von Spezifika der jeweiligen Maschinensteuerung entstehen. Unterschiede in den zeitlichen Abläufen und der Werkzeugwege wurden systematisch jeweils für drei verschiedene Fräs- sowie Drehzentren untersucht und veröffentlicht [DEN23b]. Auf Basis dieser Anforderungsanalyse wurde ein neuartiger Ansatz entwi-

ckelt, der statistische Grenzwerte für die prozessparallele und reaktionszeitkritische Überwachung zwischen Maschinen überträgt. Durch die Methode werden die Überwachungsgrenzen mit den Daten einer Maschine parametrisiert, auf eine andere Maschine transferiert und dort für eine prozessparallele Überwachung genutzt. Die Methode wurde softwaretechnisch umgesetzt und an vier baugleichen Drehmaschinen erprobt, wo sie erfolgreich zur Detektion von Materialanomalien eingesetzt wurde (Abbildung 14). Hierfür stellte ein assoziierter Partner von IIP-Ecosphere seine Maschinen in einer Industrieumgebung zur Verfügung. Anomalien in Form von Einstichen wurden erfolgreich erkannt. Dabei war die Überwachung nur auf Basis von Antriebssignalen und ohne zusätzliche Sensorik möglich. Das unterstreicht die Eignung des Ansatzes für industriell verbreitete Werkzeugmaschinen. Als wesentliche Herausforderungen bei der Umsetzung ist die maschinenübergreifend einheitliche Erfassung von Daten zu nennen. Die Leistungsfähigkeit der Überwachung wurde zudem durch die unterschiedlichen Wartungs- und Verschleißzustände der Maschinen beeinträchtigt. Die Ergebnisse wurden veröffentlicht [DEN21a, DEN23b]. Weiter wurde die Robustheit gegenüber verschiedenen Störgrößen und variierenden Fehlerarten beim Transfer und damit das Transferpotential untersucht und veröffentlicht [DEN23a].



Überwachung

Zielmaschine: CTX beta 1250 TC
Datensender: 3 x CTX beta 1250 TC
Grenzen: statistisch nach BRI09

Anomalie

Materiallücke


Prozess

$v_c = 240 \text{ m/min}$
 $a_p = 1 \text{ mm}$
 $f = 0,3 \text{ mm}$

Material: S355JR

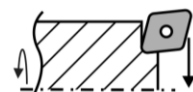
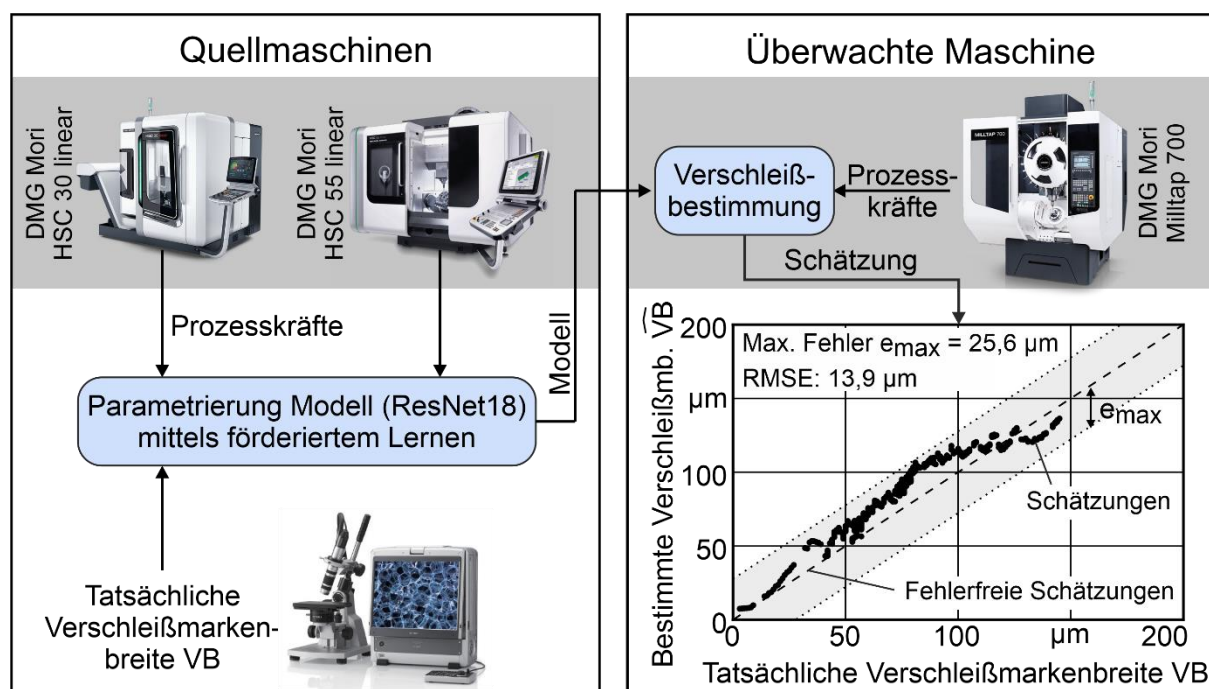


Abbildung 14: Erfolgreiche Überwachung reaktionszeitkritischer Prozessanomalien beim Drehen mit Überwachungs-grenzen, die zwischen Maschinen transferiert wurden.

Bestimmung von Werkzeugverschleiß

Das zweite Anwendungsszenario ist die zeitunkritische Gewinnung planungsrelevanter Prozessinformationen durch die Auswertung von Prozessdaten. Dafür wurde als Beispielprozess die Verschleißbestimmung beim Fräsen unter maschinenübergreifender Nutzung von Prozesswissen betrachtet. Für die Entwicklung der Verschleißbestimmung wurde ein Demonstrator aus drei Fräszentren aufgebaut (Abbildung 15). Für alle Maschinen wurden Werkstücke aus dem gleichen Werkstoff, gleiche Werkzeuge (VHM-Schaftfräser) und gleiche NC-Anweisungen genutzt. An zwei Maschinen, den Quellmaschinen, sind aufgezeichnete Prozesssignale und der in regelmäßigen Abständen gemessene Verschleiß bekannt (Verschleißmarkenbreite VB). An den Quellmaschinen wurden insgesamt sechs neue Fräser

durch Schulterfräsen ohne Kühlschmierstoff vollständig verschlissen (Einsatzgrenze bei $VB = 150 \mu\text{m}$ angenommen). Mit den dabei aufgezeichneten 4.652 multivariaten Prozesssignalverläufen wurde ein tiefes neuronales Netz (ResNet18) parametrierung. Tiefe neuronale Netze leiten wichtige Zusammenhänge eigenständig aus den Prozessdaten ab und erfordern deshalb weniger Expertenwissen über die Anwendungsdomäne als klassische Ansätze. Dadurch sinkt der Qualifikationsbedarf von Fachpersonal hinsichtlich domänenspezifischen Wissens. Eingangsgröße des Modells sind die Verläufe der Prozesskräfte als Zeit-Frequenz-Darstellung (STFT). Ausgangsgröße des Modells ist die Verschleißmarkenbreite VB in Mikrometern. Die Parametrierung des Modells erfolgte mithilfe des föderierten Lernens. Das föderierte Lernen ermöglicht das Teilen von Wissen zwischen Firmen, ohne dass die Daten selbst geteilt werden müssen, was für viele Firmen eine Voraussetzung für den Einsatz unternehmensübergreifender Lösungen ist. An der dritten Maschine wird das Modell anschließend genutzt, um die Verschleißmarkenbreite VB auf Basis der Prozesskräfte zu bestimmen. Eine Anlernphase an der überwachten Maschine gibt es nicht. In Abbildung 15 (rechts unten) ist die erfolgreiche Bestimmung der Verschleißmarkenbreite über die Nutzungsdauer eines Werkzeugs dargestellt. Die Diagonale markiert fehlerfreie Ergebnisse. Der maximale Fehler e_{max} lag bei $25,6 \mu\text{m}$ (31 %, bezogen auf den tatsächlichen Verschleiß bei $VB = 80 \mu\text{m}$). Die Ergebnisse wurden als industrienaher Veröffentlichung eingereicht [DEN23d]



Prozess

Material: Sphäroguss EN-GJS-600-3

Werkzeug: Schaftfräser D8 mm

[Bildquellen: DMG Mori AG, Keyence Corporation]

$v_c = 200 \text{ m/min}$

$a_p = 2 \text{ mm}$

$a_e = 4 \text{ mm}$

$f_z = 0,05 \text{ mm}$

Sti/118847@IFW

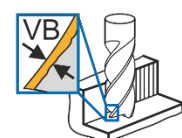


Abbildung 15: Überwachung eines Fräasers mit Prozessdaten anderer Maschinen

Die Methode zur Verschleißbestimmung wurde mit Docker containerisiert und ist damit auch für potentielle Anwender als Softwarebaustein verfügbar. Um die Nachvollziehbarkeit zu erhöhen und die Methode robuster und damit anwendungsorientierter zu gestalten wurde die Methode so konzipiert, dass sie Zeit-Frequenz-Diagramme nutzt (STFTs). Die Diagramme sind durch Ingenieurinnen und Ingenieure

interpretierbar. Damit kommt die Methode dem Feedback potentieller Anwender nach, hohe Nachvollziehbarkeit und Interpretierbarkeit zu gewährleisten. Zusätzlich wurden der Einsatz und Nutzen von Saliency-Maps für die STFSS untersucht. Saliency-Maps quantifizieren, welche Bereiche der Zeit-Frequenz-Darstellung (STFTs) von den tiefen neuronalen Netzen für die Verschleißbestimmung genutzt werden. Daraus werden Anwendungsvoraussetzung für die Methode extrahiert. Etwa werden vorzugsweise der Frequenzbereich bis etwa 1,5 kHz und die Kraft F_z für die Verschleißbestimmung genutzt. Entsprechend sollten an den Werkzeugmaschinen Sensoren verbaut sein, die diesen Frequenzbereich und diese Kraftrichtung erfassen. Die Veröffentlichung der Ergebnisse ist in Vorbereitung.

Die Durchführung der Versuche, Aufzeichnung der Daten sowie deren Aufbereitung und Strukturierung war mit sehr hohen Sachkosten und zeitlichen Aufwänden verbunden. Um Mehrwerte für die Community zu erzeugen, wurden die Daten, die für das Training und die Validierung der neuronalen Netze aus Abbildung 15 genutzt wurden, veröffentlicht [DEN23b]. Dafür wurden die erzeugten Daten aufbereitet und in die öffentlich zugängliche Datenbank „Mendeley Data“ überführt. Der Datensatz umfasst insgesamt 6.418 Prozesssegmente mit multivariaten Zeitreihen, für die als Label die Verschleißmarkenbreite gegeben ist. Einen Überblick zur Struktur des Datensatzes gibt Abbildung 16.

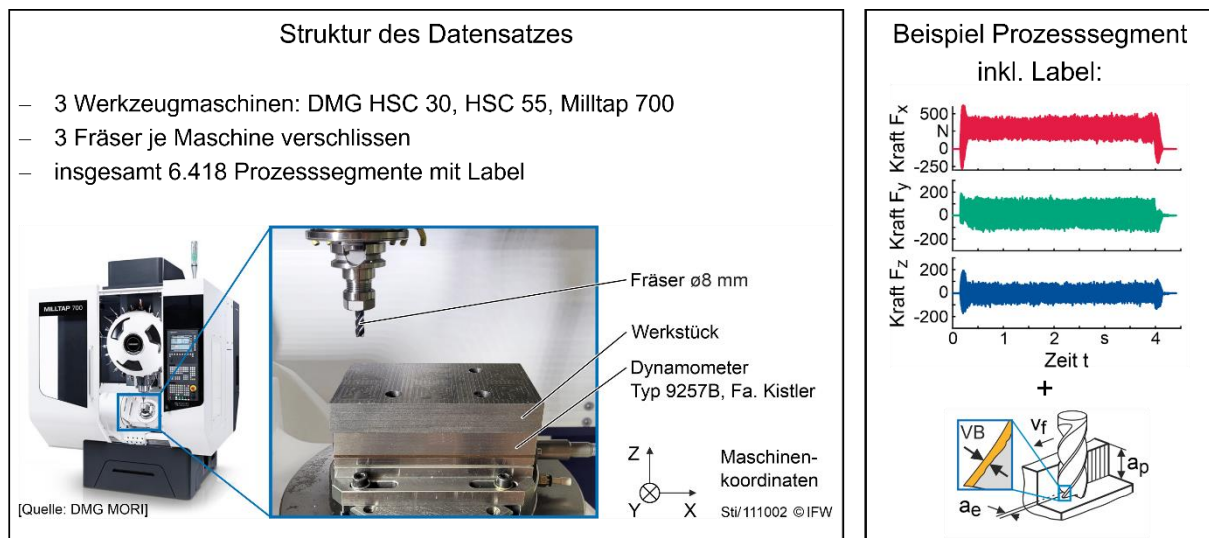


Abbildung 16: Zusammenfassung des veröffentlichten Datensatzes [DEN23b] mit Zeitreihen vom Fräsen unter variierendem Werkzeugverschleiß und auf variierenden Werkzeugmaschinen.

Mit dem Ziel die Robustheit datengetriebener Ansätze zu steigern, wurde auch die Übertragbarkeit von Prozesswissen zwischen zwei verschiedenen Prozessarterforschern. Dafür wurde ein Algorithmus entwickelt, der Prozessausschnitte nach Ähnlichkeit in einer Baumstruktur ordnet, wodurch eine Verknüpfung ähnlicher Prozessausschnitte entsteht. Dies ermöglicht es, Metainformationen, etwa über ein Fehlermuster oder Prozesszustand, zwischen den Prozessarten zu übertragen. Der Ansatz wurde erfolgreich für den Wissenstransfer zwischen Fräs- und Drehprozessen erprobt. So hat der entwickelte Algorithmus eine Anomalie beim Fräsen korrekt als Materialfehlstelle identifiziert, nur indem der Ansatz nach dem ähnlichsten Fehlermuster bei Drehprozessen sucht. Damit erhöht der entwickelte Ansatz die Robustheit beim Transfer von Prozesswissen zwischen variierenden Prozessen erheblich. Das ist insbesondere für die praktische Anwendung von hoher Bedeutung, da im industriellen Umfeld eine große Vielfalt an Prozessen eingesetzt wird. Die Ergebnisse wurden veröffentlicht [DEN21c].

Forschung zu unterstützenden KI-Methoden zur Selbstoptimierung

Im Rahmen der interdisziplinären Zusammenarbeit im Think Tank wurden Herausforderungen aus praktisch untersuchten Anwendungsfällen in die KI-Methodenforschung übertragen. Zum Lernen der Steuerung einer lokalen Suche für kombinatorische Optimierungsprobleme wurden drei unabhängige algorithmische Aspekte mit jeweils mehreren Alternativen identifiziert und ihre sequentielle Auswahl während eines iterativen Suchlaufs als Markov-Entscheidungsprozess (MDP) formalisiert. Es wurde ein tiefes neuronales Graphen-Netzwerk als Policy-Modell für dieses MDP entwickelt, das einen gelernten Controller für die lokale Suche namens NeuroLS hervorbringt [FAL22].

Außerdem wurde die Online-Klassifizierung von unregelmäßig abgetasteten multivariaten Zeitreihen mit fehlenden Werten erforscht. Für diese Aufgabe wurde ein Deep Set-basiertes Modell namens 'Deep Convolutional Set Functions' vorgeschlagen. Aufgrund der starken Unregelmäßigkeiten in den Zeitreihen ist die Verwendung von Indikatoren für fehlende Werte nicht sinnvoll, da sie den Daten zusätzliches Rauschen hinzufügen, die das Lernen des Modelles erschweren. Normale Deep Learning-Algorithmen können nicht angewendet werden, wenn die Daten fehlende Werte enthalten. Daher verwenden wir eine neue Architektur, die dieses Problem umgeht, indem wir die Kanäle in den Zeitreihen aufteilen und sie durch den gemeinsamen Encoder leiten. Schließlich kombinieren wir alle latenten Einbettungen und leiten sie durch den Decoder [YAL22]. Dieses Modell wurde an mehreren Datensätzen getestet und hat die besten Ergebnisse geliefert.

Ebenen-übergreifende Optimierung (UAP 4.2)

Ziel dieses Unterarbeitspaketes ist es, eine Gesamtmethode zu entwickeln und zu erforschen, durch die eine prozessübergreifende und damit ganzheitliche Reaktion auf Störgrößen sowie die Optimierung des Materialflusses vollständiger Fertigungssysteme realisiert wird. Dafür wurden bestehende, datengetriebene Methoden um intelligente KI-Lösungen erweitert, wodurch die Produktionssteuerung in die Lage versetzt wurde, einzelne Prozesse und Ebenen-übergreifende Prozesse selbstständig und autonom zu optimieren.

Um definierte Systemgrenzen zu erzeugen, wurde ein zusammenhängendes, ereignisdiskretes, flexibles Fertigungssystem spezifiziert. Die betroffenen Komponenten stellen Aufträge, Bearbeitungsstufen und Maschinen dar. Die Modellierung der flexiblen Fertigungssysteme (FFS) erfolgte in einer ereignisdiskreten Materialflusssimulation mit der Simulationssoftware Tecnomatix Plant Simulation. Dort wurde eine graphische Benutzerschnittstelle zur einfachen Modellierung und Variation von FFS und den Komponenten der Aufträge, Bearbeitungsstufen und Maschinen programmiert.

Als geeignete Versuchsumgebung zur Untersuchung über die Auswirkung von KI-Methoden in der Produktionssteuerung von flexiblen Fertigungssystemen wurde die ereignisdiskrete Materialflusssimulation identifiziert. Zur Quantifizierung der Systemleistung wurden in einer Recherche bestehende Leistungskennzahlen (KPI) systematisch analysiert und relevante KPI für das weitere Vorgehen extrahiert. Das daraus entwickelte, ganzheitliche Kennzahlensystem ist Tabelle 2 zu entnehmen.

Tabelle 2: Ganzheitliches Kennzahlensystem zur Bewertung von KI-Methoden in FFS

Kategorie	Kennzahl	Beschreibung
Zeitbasiert	Gesamtproduktionsdauer	Kumulierte Durchlaufzeit aller Aufträge in einem Fertigungssystem

Zeitbasiert	Durchlaufzeit	Summe der Bearbeitungs-, Transport- und Wartezeiten in allen Bearbeitungsstufen eines Auftrags in einem Fertigungssystem
Zeitbasiert	Liefertermintreue	Relative Einhaltung zugesagter Liefertermin
Zeitbasiert	Verspätung	Differenz zwischen SOLL- und IST-Liefertermin
Kostenbasiert	Produktionseinzelkosten	Entstehende Kosten, die einem Auftrag direkt zurechenbar sind
Kostenbasiert	Herstellkosten	Produkteinzel und Gemeinkosten, die zur Herstellung eines Auftrags entstehen
Qualitätsbasiert	Gesamtausschussmenge	Relativer Ausschuss

Basierend auf den Erkenntnissen aus den vorangegangenen Arbeitsschritten wurde ein ganzheitliches Optimierungskonzept entworfen, das auf Störung innerhalb diskreter FFS autonom reagiert. Das zugrundeliegende Optimierungsproblem ist als „Flexible Scheduling Problem“ (FSP) bzw. optimierte Belegungsplanung bekannt. Dieses Optimierungsproblem hat zum Ziel, die zielgrößenoptimale Allokation von Aufträgen auf einer Vielzahl von Maschinen zur Bearbeitung einer definierter Arbeitsvorgangsfolgen unter Berücksichtigung von Störgröße, wie z. B. Maschinenausfälle, sicherzustellen. Durch die Formulierung eines FSP kann die Produktionssteuerung eines FFS auf die Kernelemente Aufträge, Bearbeitungsschritte und Maschinen reduziert werden. Als wichtigste Zielgröße des FSP wurde die Kennzahl *Verspätung* übernommen.

Das spezifizierte Fertigungssystem wurde als dezentrales, Multiagentensystem aufgefasst (MAS). In MAS werden Entscheidungsprozesse auf eine Vielzahl von autonomen Agenten aufgeteilt. Das kooperative MAS zur Produktionssteuerung ist zwischen Produktionsleitebene und den physischen Werkzeugmaschinen eines flexiblen Produktionssystems verortet (Abbildung 17).

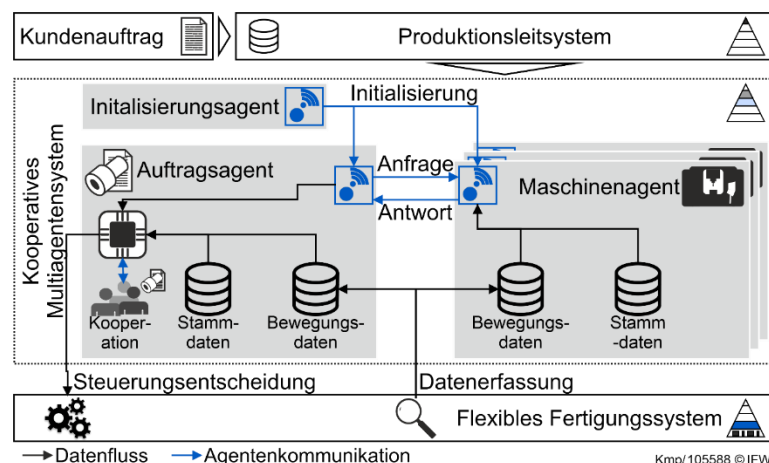


Abbildung 17: Kooperatives Multiagentensystem zur dezentralen und autonomen Produktionssteuerung

Kernelemente des Konzepts sind Aufträge und Maschinen, die im kooperativen MAS über entsprechende Softwareagenten abgebildet sind. Softwareagenten sind allgemein als digitale, flexible und gleichzeitig autonome Entitäten zu definieren. Für jeden Bearbeitungsschritt bzw. für jede Arbeitsvorgangsfolge erfolgt eine intelligente Verhandlung über die Bearbeitung von Aufträgen zwischen dem entsprechenden Auftrags- und in Frage kommenden Maschinenagenten. Unter der Berücksichtigung der Zielgröße Verspätung erfolgt auf dieser Verhandlungsbasis eine optimale Zuweisung von Aufträgen auf

Maschinen. Dabei erhöht der Informationsaustausch zwischen den Agenten, die sogenannte Kooperation, die Entscheidungsgüte. Die Modellierung des Fertigungssystems erfolgte über eine ereignisdiscrete Materialflusssimulation, die zukünftig durch ein reales FFS ersetzt werden könnte.

Durch die Kooperation der Agenten entsteht ein „NP-hartes“ Problem, dessen Lösung durch seine sehr hohe Komplexität überproportional viel Aufwand erfordert. In einer Literaturrecherche wurde das Deep Q Reinforcement Learning (DQN) als ein KI-Ansatz identifiziert, der bei dieser Problemkomplexität praktisch einsatzfähig ist. Zur Integration des DQN-Ansatzes in das MAS wurde der Auftragsagenten modifiziert, sodass ein künstliches Neuronales Netz (kNN) die Steuerungsentscheidungen trifft. Da das DQN-Verfahren grundsätzlich auf dem klassischen Reinforcement-Learning-Verfahren beruht, ist eine Explorationsphase notwendig. Der zugehörige Trainingsprozess erfolgt über eine Simulationsumgebung. Damit das Optimierungskonzept prinzipiell auch auf reale Fertigungssysteme angewandt werden kann, wurde ein separates Optimierungsmodell konzeptioniert, in dem das zentralisierte Training stattfindet. Dadurch kann die Simulationsumgebung zur Explorationsphase genutzt werden und bei guter Performanz durch eine Schnittstelle zum entsprechenden realen Fertigungssystem ersetzt werden. Zusätzlich ergibt sich dadurch die Möglichkeit ein zentralisiertes Training auf einer entsprechenden informationstechnischen Infrastruktur durchführen zu lassen und das kooperative MAS dezentral im Fertigungssystem zu betreiben (vgl. Abbildung 18).

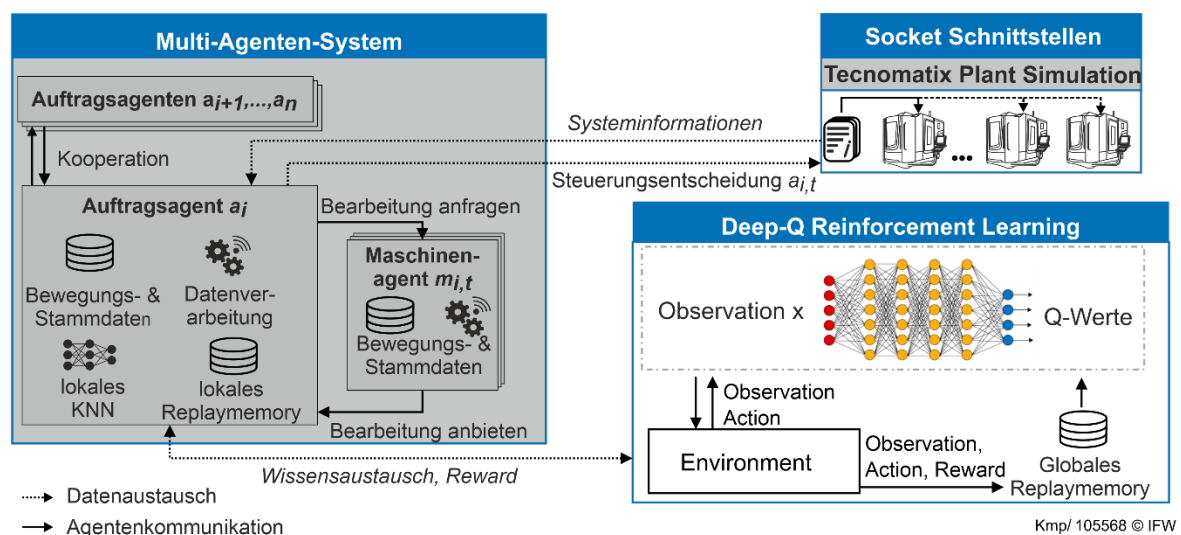
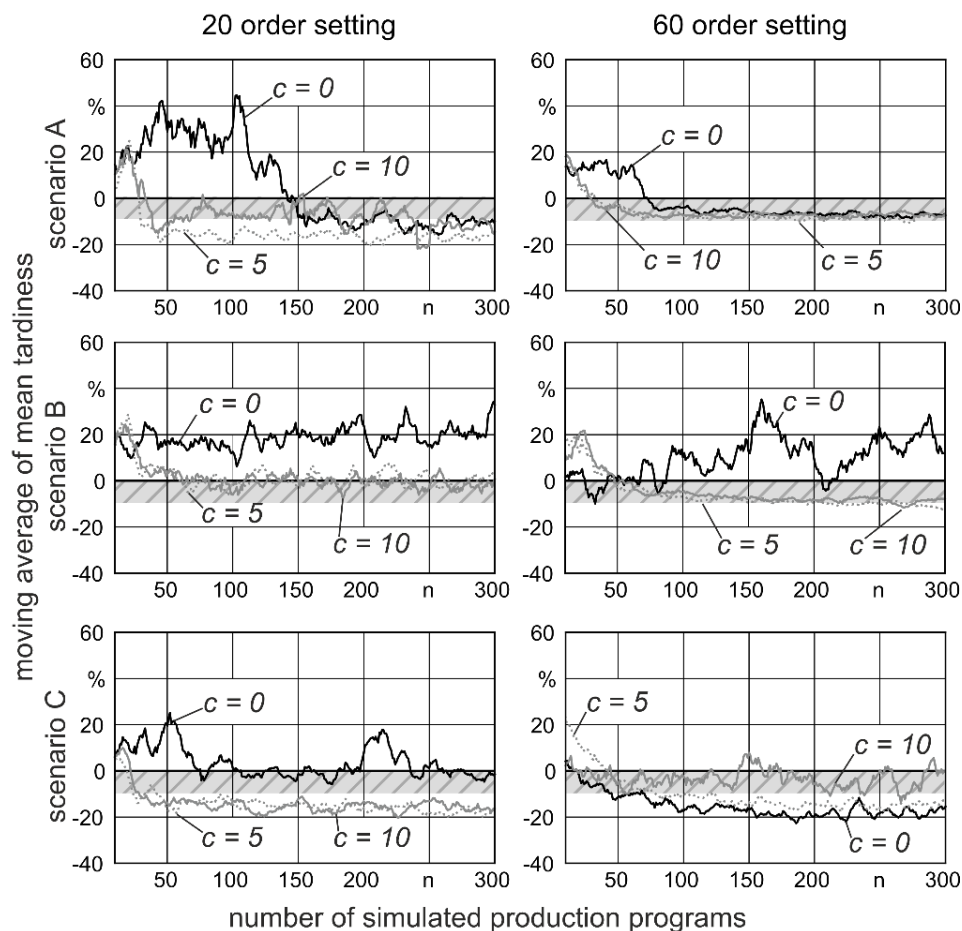


Abbildung 18: Integration der DQN-Ansätze in das kooperative Multiagentensystem

Zur Exploration des entwickelten Optimierungskonzepts wurde eine umfangreichere Simulationsstudie durchgeführt, deren Ergebnisse in [DEN21e] veröffentlicht sind. Darin wurden die Auswirkungen bzw. der Effekt eines erweiterten Kooperationsgrads zwischen Auftragsagenten auf das DQN-Verfahren analysiert. Es wurden insgesamt sechs verschiedene Konfigurationen eines FFS mit unterschiedlicher Anzahl von Maschinen, Bearbeitungsschritten und zu bearbeitenden Aufträgen eines Produktionsprogramms (20 und 60) zusammengestellt. In allen Konfigurationen wurde der Kooperationsgrad c zwischen keiner Kooperation ($c = 0$), einer Kooperation zwischen fünf aufeinander folgenden Aufträgen ($c = 5$) und einer Kooperation zwischen zehn aufeinander folgenden Aufträgen ($c = 10$) variiert. Alle Experimente wurden als Mehrwert für die Community auf Github zur weiteren Verwendung lizenzfrei zur Verfügung gestellt [PAL21].

Im niedrig-komplexen Szenario A, in dem ein FFS mit insgesamt neun Maschinen in drei Bearbeitungsstufen sowie drei unterschiedliche Auftragsstypen simuliert wurde, ist zu erkennen, dass eine angestrebte Verspätung (schraffierter Bereich in Abbildung 19) in allen betrachteten Kooperationsgraden c erreicht wurde. Das DQN-Verfahren benötigt ohne Kooperation signifikant länger als mit Kooperation. Im Falle des 60 Aufträge umfassenden Produktionsprogramms ist zudem zu erkennen, dass der Lernerfolg der Auftragsagenten sich mit zunehmender Informationsgrundlage stabilisiert.

In Szenario B wurde die Komplexität des FFS deutlich erhöht. Insgesamt wurden hier 18 Maschinen in drei Bearbeitungsstufen im FFS simuliert. Die Komplexität der Produktionsprogramme entsprach Szenario A. Durch die Erhöhung der Komplexität des FFS wurde festgestellt, dass der Verzicht auf eine Kooperation ($c = 0$) den Lernerfolg gänzlich verhindert, sodass die angestrebte Verspätung nicht stabil erreicht wurde. Im Gegensatz dazu konnte über einen Kooperationsgrad von $c = 5$ bereits ein vergleichsweise stabiler Lernerfolg erzielt werden, der über eine Erhöhung des Kooperationsgrads noch leicht verbessert werden konnte.



Details of the flexible manufacturing system

	Scenario A	Scenario B	Scenario C
Number of processes p :	3	3	5
Machine tools per process m :	3	6	6
Order types in production program t :	3	3	6
Size of neural network layers:	8x8x8	20x10x6	70x35x12

Kmp/ 105576 © IFW

Abbildung 19: Ergebnisse der Simulationsstudie zum Einfluss des Kooperationsgrads

In Szenario C wurde der Komplexitätsgrad des FFS auf insgesamt 30 Maschinen in fünf Bearbeitungsstufen erhöht. Zusätzlich wurde ebenfalls die Komplexität des Produktionsprogramms auf insgesamt 6 Auftragsstypen erhöht. Zu Verifizierung der Signifikanz der bisherigen Ergebnisse wurden alle Experimente dieses Szenarios zehnfach wiederholt und in Abbildung 19 die Mittelwerte dieser Experimente aufgeführt. Es ist zu erkennen, dass ohne Kooperation nicht der gewünschte Lernerfolg hinsichtlich der Zielgröße Verspätung realisiert werden konnte. Für die Kooperationsgrade $c = 5$ und $c = 10$ konnte hingegen ähnliche Beobachtungen wie für die vorherigen Szenarien A und B verifiziert werden.

Insgesamt ist als Ergebnis festzuhalten, dass eine explizite Berücksichtigung des Kooperationsgrads eine signifikante Performance-Steigerung für komplexe Optimierungsprobleme, die durch das DQN-Verfahren und kooperativen MAS gelöst werden, hervorbringt.

Unterstützende Methodenforschung im Bereich ebenen-übergreifende Optimierung

Prognosen sind ein wichtiger Aspekt, der in den gezeigten Anwendungsfällen einen Mehrwert bietet. Dies kann auf viele verschiedene Anwendungsfälle in der Produktion angewendet werden: von der vorausschauenden Wartung über die optimierte Ressourcenplanung bis hin zur Qualitätsprognose. Zunächst haben wir ein Modell für die deterministische Vorhersage von Zeitreihen entwickelt. Inspiriert von der U-Net-Architektur haben wir das Y-former Modell implementiert, eine auf Aufmerksamkeit basierende Architektur. Das aktuelle Modell 'Informer' leidet unter dem Problem der Überanpassung. Y-Former umgeht dieses Problem durch eine bessere Architektur und beste Generalisierung. Des Weiteren übertraf das Y-former Modell die Informer Architektur für mehrere Datensätzen signifikant [MAD22].

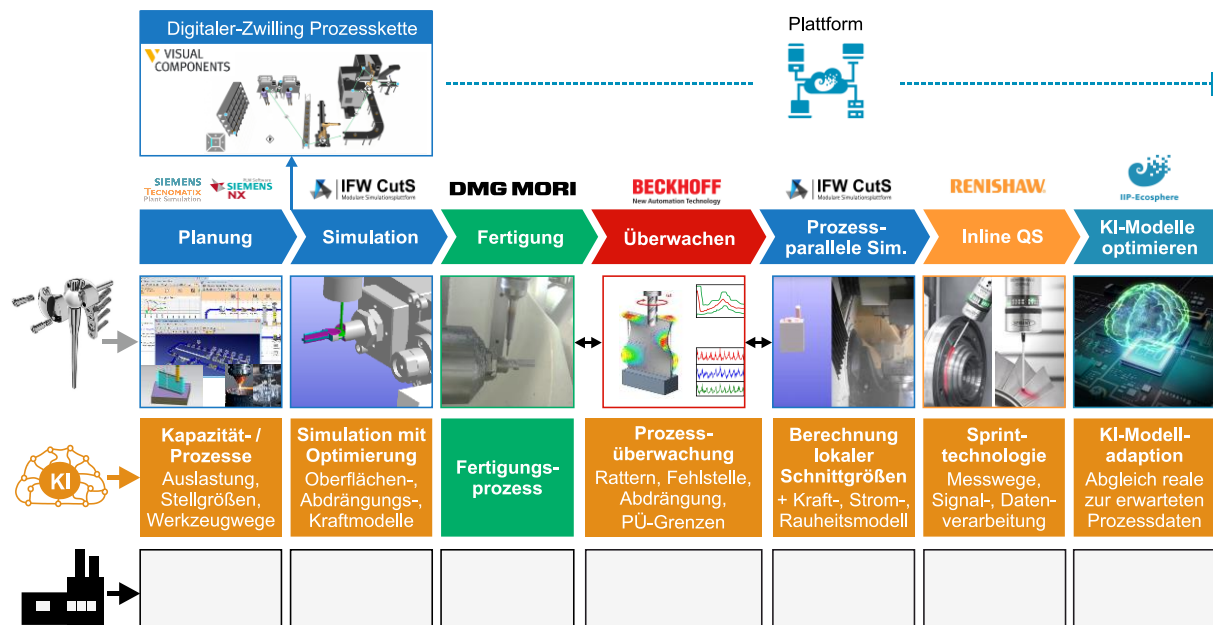
Zudem haben wir uns mit dem Aspekt der deterministischen Vorhersage unregelmäßiger multivariater Zeitreihen mit fehlenden Werten beschäftigt. Wie oben dargelegt, ist die Analyse dieser Art von Zeitreihen aufgrund der Unregelmäßigkeit der Beobachtungen und der Spärlichkeit schwierig. Alle bestehenden Modelle sind aufgrund der rekursiven Natur der Lösung sehr langsam. Um dies zu umgehen, haben wir GraFITi vorgeschlagen, ein schnelles graphenbasiertes Modell. Wir haben hierbei Aufmerksamkeitsnetzwerke als Basis für unsere Architektur gewählt. GraFITi ist nicht nur genau, sondern auch effizient im Vergleich zu den bestehenden Modellen [YAL24].

Selbststeuernde Prozessketten und Qualitätsregelkreise (AP 4.3)

Ziel des APs 4.3 ist die methodische Entwicklung und Erforschung ganzheitlicher Qualitätsregelkreise, die durch Einsatz von KI-Lösungen Produktionsprozesse qualitätsorientiert anpassen und steuern.

Referenzanwendung, Qualitätszielgrößen und Wirkzusammenhänge

Zur Entwicklung und Erforschung selbststeuernder Prozessketten und Qualitätsregelkreise wurde im Rahmen der Projektlaufzeit eine Rahmenprozesskette ausgearbeitet (Abbildung 35). Die Prozesskette stellt mögliche KI-Ansätze und die im Institut einsetzbare Soft- und Hardware für die jeweiligen Prozessschritte im Anwendungsgebiet der Medizintechnik – spanende Implantatsfertigung vor. Implantate weisen besonders hohe Qualitätsanforderungen bei sehr niedrigen Toleranzen auf. Hierzu zählen eine hohe Maß-, Form- und Lagegenauigkeit sowie eine geringe Oberflächenrauheit. Für die experimentellen Untersuchungen wurde der Fokus auf die Qualitätszielgrößen Maß- und Formgenauigkeit gelegt und die Einflussgrößen auf den Zielwert mittels einer Ursache- und Wirkungsanalyse identifiziert.

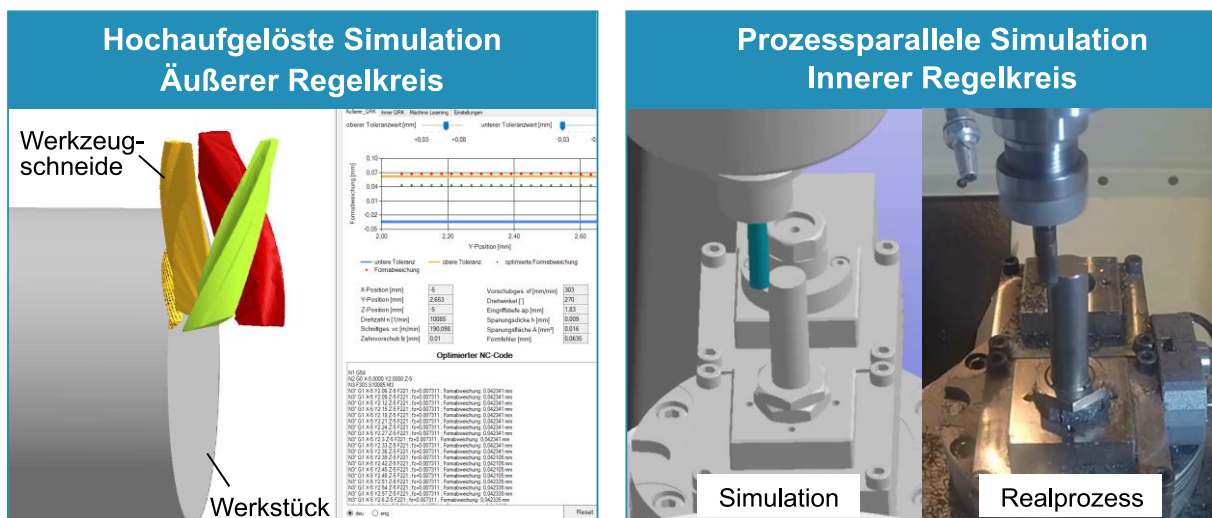


Bildquellen: Jkoo, Siemens, Techpilot, DMG Mori, Visual Components, Fraunhofer, Springer, Komet, Beckhoff, Renishaw, Wallpaperaccess

He/109553 © IFW

Abbildung 20: Rahmenprozesskette – Spanende Implantatsfertigung

Das Erreichen der genannten Zielgrößen ist durch die komplexen Geometrien patientenspezifischer Implantate besonders anspruchsvoll. Die spanende Herstellung bedarf mehrachsige Fertigungsprozesse und speziell abgestimmte Prozessparameter. Zur Erreichung der Qualitätsziele wurde daher eine ganzheitliche, qualitätsorientierte Anpassung und Steuerung innerhalb der Prozessplanung – äußerer Regelkreis – sowie in der Prozessüberwachung – innerer Regelkreis – mittels des Einsatzes von KI-Lösungen untersucht. Innerhalb der Prozessauslegung erfolgte eine KI-basierte Prognose der Maß- und Formgenauigkeit und die Adaption von Prozessstellgrößen in Form eines optimierten NC-Codes mittels einer realitätsnahen, hochauflösenden Materialabtragsimulation. Zur Umsetzung der Prozessüberwachung wurde eine prozessparallele Simulation mithilfe eines digitalen Zwillings erarbeitet (Abbildung 21). Der entwickelte digitale Zwilling ist in der Lage Steuerungs- und Maschinenparameter in Echtzeit auszulesen und zugleich zu modifizieren, sodass direkt in den laufenden Fertigungsprozess eingegriffen werden konnte.



He/109554 © IFW

Abbildung 21: Simulationsansätze – Äußerer und innerer Regelkreis

Zur Erreichung einer hohen Modellprognosegüte wurden verschiedene KI-Methoden (maschinelles Lernen), wie Entscheidungsbäume, lineare und generalisierte Modelle sowie verschiedene Hyperparameter-Konfigurationen untersucht (Softwarebibliothek Microsoft ML.Net). Die Generierung des Trainings-, Test und Validierungsdatensatz erfolgte mittels experimenteller Zerspanungsversuche. In den Versuchen wurden Prozessparameter, wie der Zahnvorschub und Schnittstrategie, variiert sowie prozessparallel interne Maschinensteuerungssignale (ML-Eingangsgrößen), wie Achsposition und -Ströme, ausgelesen und lokale Schnittgrößen innerhalb der Simulation berechnet. Nach der spannenden Bearbeitung erfolgte eine hochauflösende Messung von Formabweichung und Maßhaltigkeit (ML-Ausgangsgröße) mit einem maschinenintegrierten Hochgeschwindigkeits-Messtaster (Renishaw Sprint). Nach einer anschließenden Datenaufbereitung erfolgte das Modell-Training, die Modell-Evaluation und Implementierung in die Simulationssysteme (Abbildung 22). Das Anlernen und Auswahl des einzusetzenden KI-Modells erfolgte mit der Methode des automatisierten Lernens – AutoML. Als Versuchsumgebung für die Datengenerierung und der Erforschung der ganzheitlichen Qualitätsregelkreise diente das in AP 11.2 aufgestellte Experimentierfeld (vgl. Abschnitt 2.1.11).

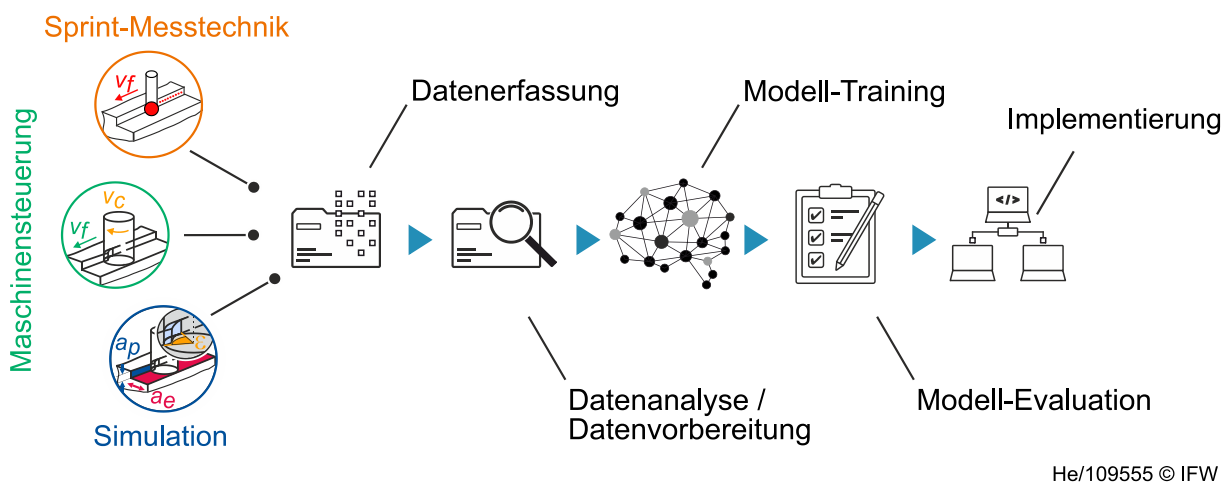


Abbildung 22: Prozess zur Erstellung der ML-Modelle

Erprobung, Exploration und Transferpotential

Zur Erprobung und Erforschung der Qualitätsregelkreise wurde exemplarisch ein Teilbereich des DRUJ (eng: distal radioulnar joint) - Implantat mit und ohne KI-modellunterstützender, ganzheitlicher Qualitätsregelkreise gefertigt. Dabei konnte festgestellt werden, dass die engen Fertigungstoleranzen ohne den Einsatz eines Qualitätsregelkreises nicht eingehalten werden konnten (Abbildung 23).

Experimentelle Fräsversuche mit aktiven Qualitätsregelkreis – innerer, äußerer und kombiniertem Regelkreis – zeigen auf, dass die zu erzielenden Qualitätszielgrößen mit Hilfe der in der Simulation integrierten KI-Modellierungen und der Prozessstellgrößenadaption erzielbar sind (Abbildung 23). Die Adaption der Stellgrößen im NC-Code (äußerer Regelkreis) oder über einer prozentualen Änderung von Maschinensteuerungsvariablen (innerer Regelkreis) erfolgt über einen aufgestellten und in der Simulation integrierten Algorithmus. Innerhalb des Algorithmus werden im jedem definierten Simulationsintervall Prozessstellgrößen iterative angepasst, bis ein zulässiger Qualitätszielwert vom KI-Modell prognostiziert wird oder eine maximale Iterationsanzahl erreicht wurde.

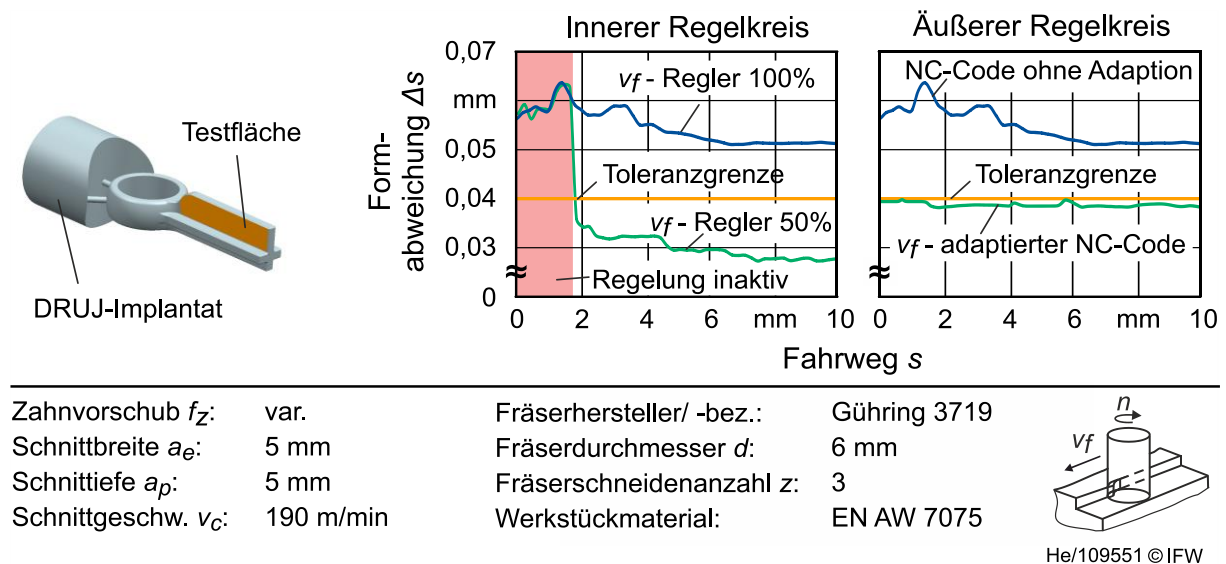


Abbildung 23: DRUJ-Implantat – Fertigung mit und ohne KI-modellunterstützenden Qualitätsregelkreis

Der vorgestellte Ansatz, bestehend aus hochauflösender und prozessparalleler Simulation mit integrierten KI-Prognosemodellen bietet großes Potential in der Einzelteil- und Serienfertigung. Dies gilt insbesondere für den Einsatz von künstlicher Intelligenz innerhalb eines digitalen Zwillings, wodurch Maschinensteuerungsdaten in Echtzeit gelesen und geschrieben, lokale Schnittgrößen berechnet und Prognosen zu Qualitätszielgrößen vorgenommen und entsprechende Prozessstellgrößen adaptiert werden können. Aktuelle Herausforderung bilden die umfangreiche Datenbasis zum Anlernen der KI-Modelle und dessen Gültigkeit bei veränderten oder neuen Eingangsgrößen. Des Weiteren ist die erforderliche Simulationsschwindigkeit zur Ermittlung lokaler Schnittgrößen und optimaler Prozessstellgrößen zu nennen.

Zur Bewertung des industriellen Einsatzes und Transferpotentials der aufgezeigten, ganzheitlichen Qualitätsregelkreise erfolgten Live-Demonstrationen und Diskussion auf der Weltleitmesse der Produktionstechnologie – EMO Hannover 2023. In Form von begleitenden Expertengesprächen wurde das Potenzial bei den Industriepartnern erkannt und eine gemeinsame Übertragung des Ansatzes in das industrielle Produktionsumfeld angestrebt.

Interpretierbarkeit und Erklärbarkeit (UAP 4.4)

Interpretierbarkeit und Erklärbarkeit bei Zeitreihen

Erklärbarkeit und Interpretierbarkeit von KI-Methoden in der intelligenten Produktion sind Gegenstand aktueller Forschung. In diesem Arbeitspaket wurde an einen Überblick über bestehende Methoden der erklärbaren künstlichen Intelligenz (XAI) gearbeitet, die auf Zeitreihendaten angewendet werden. Dabei wurden verschiedene Erklärungsmethoden, gängige Datensätze sowie unterschiedliche Evaluationsmethoden der Erklärungen identifiziert. Auch wurde bei der Sichtung Wert auf eine Unterscheidung von Ansätzen zur Erklärbarkeit im Gegensatz zu Ansätzen für die Interpretierbarkeit durchgeführt. So wurde z. B. analysiert wie eine Konzept-basierte Interpretierbarkeit in existierenden KI-Ansätzen und Architekturen eingebettet wird und wie eine Post-Hoc Interpretierbarkeit nach Nutzen eines Blackbox KI-Modells hergestellt werden kann. Zudem wurde eine Übersichtsarbeit über die Ansätze zur Bewertung der durch

XAI-Methoden erzeugten Erklärungen zusammengestellt. Außerdem wurden bestehende Probleme von XAI-Methoden diskutiert und reflektiert.

So ist es z. B. wichtig sich den Erklärungsbedarf im jeweiligen Produktionskontext zu erfassen und genau zu überprüfen, welchen Erklärungsbedarf die jeweilige Nutzergruppen hat, wie Erklärungen ohne zu stören oder abzulenken in die jeweiligen Prozesse eingebracht werden können und in welcher Form Erklärungen am besten bereitgestellt werden (z. B. textuell, visuell, explorierbar). Es wird erwartet, dass die aktuell sehr populären großen Sprachmodelle eine wichtige Rolle für die weitere Entwicklung der Erklärbarkeit von KI spielen werden, insbesondere im Bereich der Erklärungen für Endnutzer (nicht KI-Entwickler). Diese Entwicklungen kamen jedoch im Projekt zu spät, um noch produktiv aufgegriffen werden zu können.

Methoden zur Quantifizierung von Unsicherheiten von Vorhersagen

Die Quantifizierung der Unsicherheit bei Prognosen hilft dabei, die Zuverlässigkeit der Vorhersagen zu verstehen. Mit den zusätzlichen Informationen kann der Benutzer entscheiden, ob er die von einem maschinellen Lernmodell gelieferte Vorhersage für seinen Entscheidungsprozess nutzt. Oft werden unsichere Vorhersagen ignoriert und sehr sichere Vorhersagen berücksichtigt. Zu diesem Zweck haben wir ein generatives Modell namens GQFormer entwickelt, das diese Unsicherheit quantifizieren kann. Der GQFormer ist ein Multitasking-Modell, das Parameter in der Architektur teilt, die implizit mehrere Quantile effizient modelliert [JAW22, JAW22a]. Wir untersuchten auch die Unsicherheitsquantifizierung bei der Vorhersage unregelmäßiger multivariater Zeitreihen mit fehlenden Werten. Es wurde hierfür ein Modell entwickelt, das auf einem bedingten Normalizing Flow basiert und eine neuartige invertierbare Aufmerksamkeitsschicht nutzt [YAL24a].

Erklärung von Anomalien

Die Erkennung von Anomalien ist in der Produktion eine wichtige Aufgabe.. Zu diesem Zweck haben wir eine unüberwachte Methode zur Erkennung von Anomalien entwickelt, um anomale Zeitreihen zu erkennen, die von einem intralogistischen System erzeugt werden. Für diese Aufgabe haben wir mit dem Unternehmen Lenze zusammengearbeitet. Es wurde eine Zeitreihensegmentierung verwendet, so dass man nicht nur die Anomalie erkennen, sondern auch erklären kann, warum es sich um eine Anomalie handelt.

Wir haben uns außerdem mit der Erkennung von unbekannten Klassen bei Klassifizierungs-Problemenstellungen von Zeitreihen beschäftigt. Dies kann ebenfalls im Kontext der Anomalierkennung genutzt werden. Hier besteht das Ziel darin, Zeitreihen zu finden, die zu keiner der bekannten Klassen gehören. Wir verwenden die DTW-Distanz und die Kreuzkorrelation zwischen den klassenspezifischen Baryzentren und der Eingabe verwendet, um die unbekannten Klassen während des Tests zu erkennen [AKA22].

Transfer von Wissen und Modellen (UAP 4.5)

In Anbetracht der Kosten und des Ressourcenbedarfs für das Trainieren von KI-Modellen und auch in Anbetracht des oft herrschenden Datenmangels für spezielle Anwendungen, spielt der Transfer von (Teil-)Modellen und der Transfer von Wissen, durch den ausführliches Training vermieden werden kann und zudem robustere Modelle erhalten werden können eine wichtige Rolle in der KI. In diesem Unterarbeitspaket wurden verschiedene experimentelle Arbeiten im Bereich des Transfers von Wissen und

Modellen durchgeführt. Zudem wurde das Thema Transferlernen auch in Zusammenarbeit AP 3.4 bearbeitet. Siehe dazu auch die Beschreibung der Arbeiten zur Datenbewertung (Data Valuation for Batch Constrained Reinforcement Learning).

Transfer-Lernen für die Klassifizierung von Zeitreihen: Wegen der großen Bedeutung von Zeitreihen-Daten für die Produktion wurde das Transferlernen für die Zeitreihenklassifikation untersucht. Dabei spielt insbesondere die Übertragbarkeit von Deep-Learning-Modellen für die Zeitreihenklassifizierung eine wichtige Rolle, einschließlich Convolutional Neural Networks, Residual Neural Networks, Encoder und Inception. Um zuverlässige Aussagen zu erhalten, wurde das entwickelte Framework gründlich auf dem größten öffentlich verfügbaren Benchmark für die Zeitreihenanalyse evaluiert: dem UCR-Archiv, das aus 85 Datensätzen aus verschiedenen realen Domänen besteht. Unsere Experimente zeigten, dass Resnet das am besten übertragbare Modell ist.

Kombination von robustem Adversarial Reinforcement Learning mit Curriculum Learning: Modellfreie Deep Reinforcement Learning-Ansätze sind für das Training von Simulationen sehr erfolgreich, wenn in kurzer Zeit eine große Menge an Zustandsübergängen für das Training der Agenten gesammelt werden kann. Das Training von Agenten in der realen Welt ist jedoch aufgrund der Kosten für Ressourcen und Zeit eine größere Herausforderung. Ein hybrider Ansatz besteht darin, den Agenten zunächst innerhalb eines digitalen Zwillings der Zieldomäne zu trainieren. Wenn jedoch ein großer Modellierungsfehler zwischen Trainings- und Zieldomäne besteht, verhalten sich die übertragenen Agenten oft suboptimal. Um diese Schwäche zu reduzieren, wird im Think Tank an Ansätzen gearbeitet, Robust Adversarial Reinforcement Learning (RARL) mit Curriculum Learning zu kombinieren. Evaluierungen wurden derzeit innerhalb einer für RARL modifizierten Version der kontinuierlichen Kontrolldomänen von MuJoCo durchgeführt (<https://github.com/lerrel/rllab-adv>).

Off-Policy-DRL innerhalb nicht-stationärer Domänen: Derzeit basieren Steuerungssysteme hauptsächlich auf statischen und modellbasierten Heuristiken. Solche Systeme berücksichtigen auch instationäre Zustände der Umgebung. Reinforcement Learning bietet eine alternative Methode zur Steuerung von Produktionssystemen. Eine große Anzahl von Deep Reinforcement Learning- Algorithmen basiert jedoch auf Erfahrungswiedergabespeichern. Nicht-stationäre Domänen bringen das Risiko mit sich, dass aktuelle Speicherstrukturen veraltete Muster speichern. Wir arbeiten zurzeit an einem Ansatz, der aus einem kleinen Speicher für die letzten Übergänge und einem großen Speicher für die alten Übergänge besteht. Dies hilft uns, einen Kompromiss zwischen der Verwendung der neuen Übergänge und dem Abrufen der alten Übergänge besser zu finden. Dies verringert auch das Problem des katastrophalen Vergessens, da ein solches Dual-Memory-Modell die Auswahl alter Übergänge ermöglicht, die den neuen ähnlich sind. Wir verwenden einen Fingerabdruck basierenden Ansatz, der es dem Lerner ermöglicht, zu verschiedenen Zeitpunkten gespeicherte Samples zu disambiguieren. Der Fingerabdruck wird auch eine effizientere Suche nach relevanten Zustandsübergängen innerhalb unseres Duell-Speicher-Modells ermöglichen.

Zeitreihenprognose: Das Erlernen komplexer Zeitreihenprognosemodelle erfordert in der Regel eine große Menge an Daten, da jedes Modell für jede Aufgabe/jeden Datensatz von Grund auf neu trainiert wird. In der Industrie ist es oft nicht möglich, große Datenmengen zu sammeln, so dass Modelle erforderlich sind, die auf kleinen verfügbaren Datensätzen lernen. Die Nutzung von Lernerfahrungen mit ähnlichen Datensätzen ist eine gut etablierte Technik für Klassifizierungsprobleme, die so genannte

"few-shot classification". Die bestehenden Ansätze können jedoch nicht auf Zeitreihenprognosen angewandt werden, da i) multivariate Zeitreihendaten unterschiedliche Kanäle aufweisen und ii) sich Prognosen grundsätzlich von der Klassifizierung unterscheiden. In Erweiterung jüngster Arbeiten zu heterogenen Attributen in Vektordaten entwickeln wir ein Modell, das aus permutationsinvarianten Deep-Set Blöcken besteht, die eine zeitliche Einbettung enthalten. Unser Transfer-Learning-Framework übertraf die aus einfacheren Szenarien übernommenen Baselines, die entweder nicht aufgabenübergreifend lernen können oder denen zeitliche Informationen fehlen [BRI22].

Wir haben uns auch mit der Aufgabe der Zeitreihenprognose mit nur wenigen beobachteten Ereignissen beschäftigt. In Fällen, in denen jedoch ein bestehender Satz von Datensätzen verfügbar ist, können Meta-Learning Methoden angewendet werden. In diesem Paper entwickeln wir eine Meta-Learning Methode, die Stichproben aus zusätzlichen Datensätzen nutzt und lernt, Zeitreihen durch adversariales Lernen als Zusatzaufgabe für den Zieldatensatz zu erweitern. Unser Modell (FEML) ist mit einem geteilten Convolutional-Block ausgestattet, das latente Merkmale für unterschiedlich lange Eingaben aus verschiedenen Datensätzen lernt. Diese werden im Anschluss von einem datensatzspezifischen Decoder ausgewertet und zur Vorhersage für unterschiedliche Ausgabelängen verwendet. Wir zeigen, dass FEML datensatzübergreifend lernen kann und durch zusätzliches Lernen auf generierten Hilfsstichproben für den Zieldatensatz die Vorhersageleistung im Vergleich zu Single-Task-Lernen und verschiedenen Lösungen aus Joint Learning, Multi-Task-Lernen und klassischen Vorhersage-Baselines verbessern kann [JAW23].

Easy-to-use KI (UAP 4.6)

Die Entwicklung und Anwendung von Verfahren der KI in der Produktion erfordert hohes Domänenwissen in Produktionstechnik und KI. Zum Beispiel sind für die Entwicklung von KI-basierten Lösungen Anomalieerkennung das zeitintensive Aufsetzen, Durchführen und Bewerten von Experimenten notwendig. Hinzu kommt, dass jeder Use-Case individuell mit einem Datensatz und mit einer (oder ggf. mehreren) KI-Methoden erprobt wird. Um das Einstiegsniveau in die KI-Domäne zu reduzieren und das zeitintensive Experimentieren zur Detektion von Anomalien zu erleichtern, wurde im Rahmen des ein Anomaly Detection Framework entwickelt. Das Framework soll dabei das zeitintensive Aufsetzen, Durchführen und Bewerten von Experimenten für verschiedene KI-basierte Ansätze in der Anomalieerkennung in der Produktion erleichtern, indem das Einstiegsniveau und die Zeitintensität herabgesetzt werden. Ziel hierbei war, dass ein automatisches Benchmarking über bestehende KI-Verfahren mit bisher verfügbaren produktionstechnische Datensätze einfach durchgeführt werden kann. Außerdem können neue Datensätze in die prototypische Implementierung durch einfaches Hochladen eingefügt werden. Das Framework Werkzeug ein Framework entwickelt und mit ersten Testnutzern erprobt, deren Feedback zur weiteren Verbesserung des Frameworks genutzt werden wird.

Eine zweite Arbeit im Bereich easy-to-use KI fand für das Thema föderiertes Lernen statt. Wir haben ein einfach zu bedienendes KI-Framework für föderiertes Lernen entwickelt. Über eine grafische Benutzeroberfläche kann ein Benutzer ein föderiertes Lernszenario ausführen. Ziel war es hier föderiertes Lernen einfach erfahrbar zu machen, das Erproben der Methode mit verschiedenen Parametern zu erleichtern und somit den ersten Schritt zur Erprobung von föderierten Lernen im eigenen Kontext zu leisten. Zu den Konfigurationsoptionen gehören die Auswahl des Datensatzes, Anzahl der Clients, die zu verwendende Architektur des neuronalen Netzes und verschiedene andere für föderiertes Lernen

relevante Hyperparameter (z. B. Anzahl der Trainingsepochen pro Runde). Abbildung 24 zeigt das einfach zu bedienende Front-End des Frameworks. Zusätzlich unterstützt die Benutzerschnittstelle auch die

Echtzeitüberwachung des Fortschritts und Erfolg des föderierten Lernprozesses.

The image shows a user interface for configuring a federated learning process. It includes radio button selections for Dataset (Lenze, Bosch, Cifar10) and Model (LSTM AutoEncoder, Support Vector Machine, MobileNetV2). There are also sliders for FL Clients (set to 10), FL Rounds (set to 9), and Epochs (set to 15). A blue 'Start FL' button is at the bottom right.

Dataset:	☒ Lenze	FL Clients:	<input type="range" value="10"/>	10
	☐ Bosch	FL Rounds:	<input type="range" value="9"/>	9
	☐ Cifar10	Epochs:	<input type="range" value="15"/>	15
Model:	☒ LSTM AutoEncoder	Start FL		
	☐ Support Vector Machine			
	☐ MobileNetV2			

Abbildung 24: Front End des Easy-to-Use Frameworks zum föderierten Lernen

Als Grundlage wurde hier das Flower Framework verwendet, das für seine Effizienz und Flexibilität bekannt ist. Das auf dieser Basis erstellte IIP-Ecosphere Framework unterstützt einen FL-Server und mehrere FL-Clients (Docker) und ist höchst anpassbar. Sowohl der Server als auch die Clients können sowohl in Containern (Docker) als auch stand-alone ausgeführt werden. Unterschiedliche KI-Modelle wie z. B. lassen sich einfach integrieren. Zudem wird ein automatisiertes Aufteilen der Datensätze zwischen den FL-Clients ermöglicht, wodurch sich unterschiedliche FL-Szenarien simulieren lassen. Derzeit kann das Framework mit den folgenden Datensätzen ausgeführt werden: cifar10, demos, iemocap und einem von der Lenze-Gruppe bereitgestellten Datensatz.

Ebenfalls als Teil von UAP 4.6 wurde die Entwicklung von KI-Modellen für die Hannover-Messe 2022, die Hannover Messe 2023 und für die EMO 2023 durchgeführt. Es handelt sich dabei zum einen um eine Reihe von KI-Modellen zur Qualitätsprüfung für die in den Demonstratoren verwendeten, in Aluminium produzierten Modellautos und um KI-Modelle zur Prüfung der Konformität mit der Spezifikation der Modellautos (Zahl der Fenster, Format, Farbe der Reifen). Hier kamen neben einer geeigneten Vorverarbeitung Neuronale Netze zur Bild Analyse (Deep Learning) zum Einsatz. Trainingsdatensätze wurden produziert (Erfassung, Annotation), KI-Modelle Trainiert, getestet und in mehreren Iterationen verfeinert.

Außerdem wurden für die EMO 2023 KI-Modelle zur verbesserten Erkennung von Magnetidentifikationen entwickelt. Für die Magnetidentifikation liefert der Sensor zwei Arten von Signalen, das Clock Signal und das gelesene Signal zum Magnetidentifikationsstreifen. Die Signale werden mit einem Savitzky-Golay vorverarbeitet und Noise zu entfernen und die weitere Verarbeitung zu erleichtern. Danach werden die Signale mit Hilfe des Clock-Signals, weiterer Filter und Vorverarbeitungsschritte sequenziert. Um das KI-Modelle zum Erkennen der Magnetidentifikationen zu trainieren wir eine fenster-basierter Ansatz verwendet (*sliding windows*). Auf diese Weise wurde aus annotierten Sequenzen ein Trainingsdatensatz bestehend Drei-bit Folgen erzeugt und mit diesem Datensatz ein KI-Modell zum Erkennen/Übersetzen des Signals in die kodierte Identifikation trainiert. Für das Training wurde ein Neuronales Netzwerk mit 3 Layern (Fully-connected) verwendet und in 50 Epochen trainiert. Durch das KI-Modell konnte die Erkennungsrate erheblich gesteigert werden. Das KI-Modell wurde in die Prozesse integriert und bei der Demonstration bei der EMO 2023 eingesetzt.

Anwendung von Meta-Lernen für Easy-to-use KI

Wir haben auch untersucht, wie wir das Meta-Learning vereinfachen können. Als datengesteuerter Ansatz erfordert Meta-Learning Meta-Merkmale, die die primären Lernaufgaben oder Datensätze repräsentieren, die traditionell als technische Datensatzstatistiken geschätzt werden, die maßgeschneidertes Fachwissen für jede Meta-Aufgabe erfordern. Für diese Problemstellung haben wir zunächst einen Extraktor für Meta-Merkmale namens Dataset2Vec vorgeschlagen, der die Vielseitigkeit von Meta-Merkmalen für Datensätze mit der Ausdruckskraft von Meta-Merkmalen kombiniert, die von tiefen neuronalen Netzen gelernt werden. In einem Experiment mit einer groß angelegten Hyperparameter-Optimierungsaufgabe für 120 UCI-Datensätze mit unterschiedlichen Schemata als Meta-Lernaufgabe zeigen wir, dass die Meta-Merkmale von Dataset2Vec die von Experten entwickelten Meta-Merkmale übertreffen, und demonstrieren so zum ersten Mal die Nützlichkeit von gelernten Meta-Merkmalen für Datensätze mit unterschiedlichen Schemata [JOM21].

Wir haben auch den ersten Metadatenatz von 40 multivariaten Zeitseriendatenätzen zusammengestellt, darunter auch Industriedatenätze. Wir bewerteten die Leistung verschiedener Algorithmen auf dieser Grundlage. Außerdem stellten wir einen installierbaren Benchmark-Datenatz bereit und generierten die Leistung für Datensätze mit einer Vielzahl von Algorithmen. Dies hilft den Benutzern, verschiedene Modelle für ihre Aufgabe nahtlos zu testen und das beste Modell für ihre Aufgabe auszuwählen [BRI22].

2.1.5 Think Tank „Geschäftsmodelle“ (AP 5)

Im Rahmen des Think Tank „Geschäftsmodelle“ wurden Innovationen im Kontext KI-basierter Lösungen für die intelligente Produktion sowie Geschäftsmodelle für deren Verwertung erforscht. Die entsprechenden Erkenntnisse wurden in Form eines KI-Innovationsbaukastens zusammengeführt und so für die praktische Unterstützung von KI-basierten Innovationen im produzierenden Gewerbe nutzbar gemacht. Der Baukasten unterstützt Firmen bei der Entwicklung von KI-basierten Lösungen von der Ideenkonkretisierung bis hin zur geschäftlichen Betrachtung der Lösung im Ökosystem, in dem diese angeboten und operiert werden soll. Nachfolgend erfolgt eine Zusammenfassung der Ergebnisse und Erkenntnisse aus AP 5 des Think Tanks „Geschäftsmodelle“ und wie diese von Unternehmen genutzt werden können.

Typisierung von KI-basierten Geschäftsmodellen für die intelligente Produktion (UAP 5.1)

Als Ergebnis des AP 5.1 wurde ein Framework zur Typisierung und detaillierten Charakterisierung von KI-Geschäftsmodellen entwickelt (vgl. Abbildung 25). Dieses Framework führt acht grundlegenden Elemente auf, welche KI-Geschäftsmodelle beschreiben, und gibt einen Überblick über 33 Ausprägungen, welche diese Elemente annehmen können. Es strukturiert die Analyse von eigenen KI-Geschäftsmodellen oder KI-Geschäftsmodellen von Konkurrenzunternehmen und bietet damit die Grundlage für faktenbasierte strategische Entscheidungen sowie Geschäftsmodellinnovationen. Die Ergebnisse des AP 5.1 wurden in eine wissenschaftliche Publikation überführt (siehe [WIL21b]) und bei der European Academy of Management Conference 2021 eingereicht, angenommen und präsentiert (AP 13.3).² Außerdem ist auf Basis der Ergebnisse ein Werkzeug zur Charakterisierung und Typisierung von KI-Ge-

² Das Framework ist zudem über folgenden Link aufrufbar: <https://ecosystem-tools.de/AI-Toolbox/AI-business-models.html>

schäftsmodellen entwickelt worden, welches die gewonnenen Erkenntnisse auf vielfältige Weise multimedial für die Praxis operationalisiert. Die Konzeption und der Mehrwert des Werkszeugs für Unternehmen werden ausführlich in AP 5.4. erläutert.

Open framework for the characterization of artificial intelligence business models											
Dimensions of AI-Debei and Avison (2010)	AIBM Element	Characteristics									
<i>Value proposition</i>	Value proposition	Optimization		Individualization		Matching		Novelty		...	
<i>Value architecture</i>	Key resources	Proprietary		Stakeholder-provided		Public		Acquired/Rented		...	
	Core AI capability	Clustering/ Categorization	Classification		Detection		Generation		Prediction	...	
	AI integration type	Assistance			Augmentation			Automation			...
<i>Value network</i>	Customer	B2B			B2C			B2G			...
	Key partner role	Data provider		Technology provider		Integrator			Investor		...
<i>Value finance</i>	Sales model	Ownership			Use/Availability			Result			...
	Revenue model	Asset sale	Lending/ Leasing/ Licensing	Subscription	Usage fee	Third-party	Freemium	Open source	...		

Abbildung 25: Framework zur Typisierung und Charakterisierung von KI-Geschäftsmodellen [WIL21]

Erfassung von Rahmenbedingungen und Status quo KI-basierter Geschäftsmodelle (UAP 5.2)

Für die Umsetzung KI-basierter Geschäftsmodelle sind eine Reihe von Rahmenbedingungen notwendig, die Unternehmen eine effektive Innovation von KI-Lösungen ermöglichen. Um diese genauer benennen zu können, wurden im Zuge von AP 5.2.1 Herausforderungen und Potenzialen der Implementierung von (KI-)Geschäftsmodellen erhoben. Dies geschah mit Hilfe eines literaturbasierten Leitfadens, Interviews mit Unternehmensvertretern, sowie im Rahmen eines Workshops. Die Ergebnisse wurden anschließend in ein Buchkapitel überführt (siehe [JAL21]). Hierin wurden zusätzlich ein bisher wenig erschlossenes Potenzial des Einsatzes von KI in industriellen und industrie-nahen Dienstleistungen identifiziert sowie die Notwendigkeit der Bewertung von Lösungen mit einem ganzheitlichen Wertverständnis festgestellt, welches über monetäre Faktoren hinausgeht. Die Potentiale können durch den Einsatz von Grenzobjekten in der Entwicklung von KI-Geschäftsmodellen und die Einbindung interdisziplinärer Stakeholder durch Gestaltung von (Sub-)Ökosystemen erschlossen werden. Hierbei lag der Fokus im Beitrag von [JAL21] insbesondere auf zwei Kategorien von Grenzobjekten: Demonstratoren und persuasive Technologien. Demonstratoren, wie der Industrie 4.0-Demonstrator *Portable Industrial Demonstrator for Cyber-physical Systems – PID4CPS* oder die in IIP-Ecosphere involvierten Demonstratoren, dienen im Innovationsprozess der Schaffung eines gemeinsamen Verständnisses zwischen heterogenen Stakeholdergruppen sowie der Erprobung innovativer, technologie-basierter Lösungen vor ihrer finalen Einführung [OKS19]. Persuasive Technologien unterstützen die Strukturierung von Innovationsprozessen sowie den Austausch innerhalb dieser Innovationsprozesse, z. B. durch deren Einsatz in Workshops [JAL19; JAL20]. Diese und zusätzlich identifizierte Faktoren informierten die Gestaltung der Werkzeuge und Workshops in AP 5.4.

In Interviews und Workshops mit Konsortial- und assoziierten Partnern wurden außerdem KI-basierte Geschäftsmodelle erhoben (AP 5.2.2). Dazu wurden im Kontext der IIP-Ecosphere Plattform (AP 2) sieben Rollenverständnisse identifiziert, welche die Basis für repräsentative KI-Geschäftsmodelle in der intelligenten industriellen Produktion bieten. Die sieben Geschäftsmodelltypen sowie die Wertschöpfungsmechanismen der jeweiligen Rollen wurden in Geschäftsmodellkarten (siehe Abbildung 26 links) überführt und entsprechend öffentlich dokumentiert³. Die Geschäftsmodellrollen bieten Unternehmen die Möglichkeit KI-Innovationsökosysteme anhand von typischen Rollen von Firmen im Bereich der intelligenten Produktion zu modellieren. Zudem können diese für die eigenen Lösungen angepasst werden und als Überprüfung genutzt werden um herauszufinden, ob alle relevanten Kompetenzen im Ökosystem vorhanden sind.



Abbildung 26: KI-Geschäftsmodellkarten und Mini-Cases der identifizierten Best Practices

Obwohl es einige etablierte Vorgehensweisen für die Entwicklung von Geschäftsmodellen gibt, sind es doch häufig die speziellen Rahmenbedingungen eines Unternehmens die diese Vorgehensweisen erschweren. Speziell wenn es um für das Unternehmen neuartige datenbasierte Lösungen geht, stellen etablierte Vorgehensweisen nur einen Teil der Möglichkeiten dar. Während diese meist mit wissenschaftlichen Grundlagen fundiert sind, existiert auch ein weiterer Erfahrungsschatz im Umgang mit derartigen Lösungen, der von der Forschung noch nicht erfasst wurde. Dieser Erfahrungsschatz wird mit Best Practices bezeichnet und bietet wichtige Einblicke in den Wertschöpfung von KI innerhalb von Unter-

³ Die Karten können über folgenden Link eingesehen werden: <https://ecosystem-tools.de/AI-Toolbox/AI-business-model-roles.html>

nehmen. Best Practices für KI-Wertschöpfungsmechanismen wurden daher im Rahmen einer Interviewstudie erhoben und in einen Katalog überführt (AP 5.2.3). Es wurden 30 Best Practices identifiziert. Da Teilnehmer teilweise ähnliche Aussagen tätigten, wurden die Best Practices die sich thematisch und inhaltlich überschneiden, zusammengefasst. Diese konsolidierten Best Practices wurden in Form von Mini-Cases (AP 5.2.4) konsolidiert (Abbildung 26 rechts) und in das Whitepaper „Künstliche Intelligenz in der intelligenten, industriellen Produktion - Best Practice beim Rollout von KI-basierten Lösungen“ integriert, das im Rahmen von AP 11 erstellt wurde.

Erhebung spezifischer Innovationsbedarfe in den Unternehmen (UAP 5.3)

Auch wenn die Rahmenbedingungen für den erfolgreichen Aufbau eines KI-basierten Geschäftsmodells je nach Unternehmen unterschiedlich sein können, ist es dennoch von großer Wichtigkeit abzuleiten, welche Innovationsbedarfe häufig auftreten und an welchen Stellschrauben Unternehmen drehen sollten, um derartige Vorhaben umzusetzen. Dazu wurden zunächst Chancen und Herausforderungen bei der Umsetzung von KI-basierten Geschäftsmodellen auf Basis einer strukturierten Interviewstudie identifiziert (AP 5.3.1 & AP 5.3.2). Um die Innovationsbedarfe von Unternehmen aus dem Ökosystem der industriellen Fertigung abzuleiten wurden die Chancen und Herausforderungen mit bereits bekannten Innovationstools, Frameworks und Vorgehensweisen abgeglichen. Damit wurde eine Wissensbasis geschaffen aus der abgeleitet werden konnte, für welche Bereiche methodische Unterstützung am dringendsten benötigt wird. Bspw. konnte erkannt werden, dass die Beschreibung des Mehrwertes einer KI-basierten Lösung, die Definition von Anforderungen an diese oder auch die Kommunikation und Abstimmung zwischen verschiedenen Akteuren in einem KI-Projekt einer der wichtigsten Bedarfe darstellen. Zusätzlich wurde eine quantitative Studie zum Thema Innovation in der Produktentwicklung sowie zum Thema Scheitern von Innovationsvorhaben durchgeführt [RII23]. Darin wurde unter anderem festgestellt, dass es zum einen häufig offener Innovationsprozesse bedarf, um alle relevanten Kompetenzen für ein Vorhaben zusammenzubringen. Sollten derartige offenen Innovationsprozesse jedoch scheitern, kann sich dies negativ auf zukünftige Produktneuentwicklung auswirken. D.h. ein gescheitertes KI-Innovationsprojekt kann dazu führen, dass Organisationen dies nicht noch einmal wiederholen, obwohl evtl. weiterhin offene Innovationsansätze benötigt würden. Der KI-Baukasten beinhaltet genau dazu Innovationswerkzeuge, die verhindern sollen, dass derartige Projekte scheitern. Ein weiterer wesentlicher Aspekt ist hierbei die Innovationskultur, wie aus der Studie hervorgeht, aber auch der Fokus einer Organisation auf fortwährendes Lernen. Entsprechend sind die entwickelten Werkzeuge darauf ausgelegt, möglichst leicht adoptiert werden zu können. Dies wurde bspw. insofern umgesetzt, indem nicht komplett neue Werkzeuge entwickelt wurden, sondern existierende Werkzeuge und Konzepte berücksichtigt bzw. weiterentwickelt wurden.

Auf dieser Basis wurde die Entwicklung der KI-Innovationswerkzeuge und -methoden vorangetrieben (AP 5.3.3). Bspw. wurde für das Thema der Mehrwertbestimmung und -beschreibung der KI-V^{di}P entwickelt. Für die Definition von Anforderungen wurde die KI-Use-Case-Ontologie gestaltet. Auch für die Entwicklung von Geschäftsmodellen zeigte sich ein Bedarf für weitere methodische Unterstützung, wozu ein Charakterisierungsmodell für KI-Geschäftsmodelle entwickelt wurde. Als letztes großes Thema bestand Bedarf für die Integration verschiedener Akteure in die Entwicklung von KI-basierten Lösungen und deren Geschäftsmodelle. Dazu wurde eine Methodik für die KI-Ökosystemanalyse und -entwicklung erforscht. Die entwickelten Werkzeuge und Methoden werden nachstehend näher vorge-

stellt und sind gesammelt als KI-Innovationsbaukasten öffentlich zugänglich⁴. Der Baukasten stellt neben den wissenschaftlichen Veröffentlichungen das zentrale Artefakt dar, in dem die Arbeiten des Think Tanks Geschäftsmodelle konsolidiert zur Verfügung gestellt werden.

Methoden, Konzepte und Werkzeuge zur Entwicklung von KI-basierten Geschäftsmodellen (UAP 5.4)

Basierend auf den Erkenntnissen aus AP 5.2 & AP 5.3 wurden Methoden, Konzepte und Werkzeuge zur Entwicklung von KI-basierten Geschäftsmodellen entwickelt, um Unternehmen bei der erfolgreichen Entwicklung und Nutzung von KI-basierten Lösungen zu unterstützen. Diese werden über den genannten KI-Baukasten bereitgestellt. Als Leitbild für die Entwicklung von KI-basierten Geschäftsmodellen wurde ein trianguläres Referenzmodell erstellt (vgl. Abbildung 27). Je nach Situation kann ein Unternehmen anders in das Modell einsteigen und dann iterativ die Erkenntnisse in die jeweiligen anderen Werkzeuge transferieren. So ergibt sich am Ende ein Geschäftsmodell für ein KI-basiertes Wertversprechen, das in einem Ökosystem eingebettet erfolgreich umgesetzt werden kann.

Für die Entwicklung der Werkzeuge und Methoden wurden jeweils bereits bestehende Konzepte auf deren Eignung überprüft (AP 5.4.1), ggf. konzeptionell angepasst oder es wurden neue Konzepte entwickelt (AP 5.4.2). Dadurch entstanden sechs Werkzeuge von denen drei mit einer Methodik ergänzt und in anwendungsorientierte Workshopformate überführt wurden (AP 5.4.3 & AP 5.4.4).

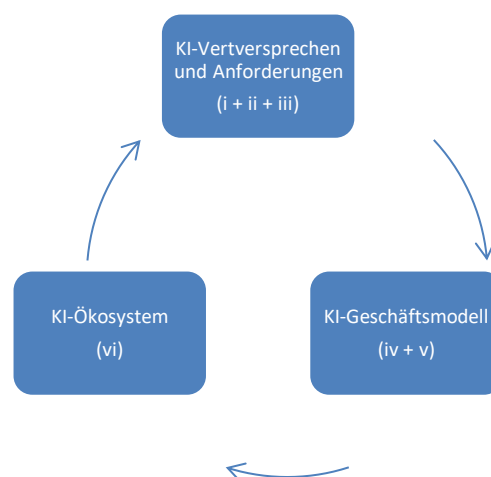


Abbildung 27: Rahmenmodell für die Konzeption der Werkzeuge und Methoden

Der KI-Innovationsbaukasten wurde prototypisch bei zwei Demonstratoren, drei Teilprojekten (IIP-Plattform AP 2, Daten-Marktplatz UAP 3.5 und Experimentierfeld UAP 11.2) sowie bei Projektpartnern erprobt sowie evaluiert (AP 5.4.9). Im Zuge eines Unterauftrags wurde der KI-Innovationsbaukasten in Form einer Innovationsinsel im offenen Innovationslabor JOSEPHS in der Nürnberger Innenstadt vom 28.08.2023 – 30.09.2023 evaluiert. Die Platzierung des KI-Innovationsbaukasten im JOSEPHS wurde durch die dortigen Innovation Guides begleitet, welche interessierte Akteure zum Ausprobieren motiviert haben, die einzelnen Bestandteile näher erläutern und schließlich auch Feedback eingeholt sowie dokumentiert haben. Der Öffentlichkeitstransfer im JOSEPHS wurde außerdem durch ein Event des RIH-Nürnberg am 27.09.2023 begleitet, welches die regionalen Akteure des IIP-Ecosphere Ökosystems

⁴ <https://ecosystem-tools.de/AI-Toolbox/AI-Toolbox.html>

zu einem Austausch zum KI-Innovationsbaukasten sowie aktuellen Trends in der intelligenten industriellen Produktion eingeladen hat. Der bereitgestellte digitale KI-Innovationsbaukasten (AP 5.4.5, AP 5.4.6, AP 5.4.7 & AP 5.4.8) umfasst sowohl die recherchierten und bereits zur Verfügung stehenden Werkzeuge, Methoden und Workshopformate aus externen Quellen als auch die im Projekt entwickelten Lösungen, die nachstehend beschrieben werden: i. The KI-VdiP Template for value proposition design, ii. AI-Use-Case Requirements Taxonomy, iii. Inspiration cards on roles of AI-based intelligent agents, iv. AI Business Model Characterisation Tool, v. Card game for designing AI business models und vi. Ecosystem Design Workshop und Ecosystem-Designer

AI-Toolbox

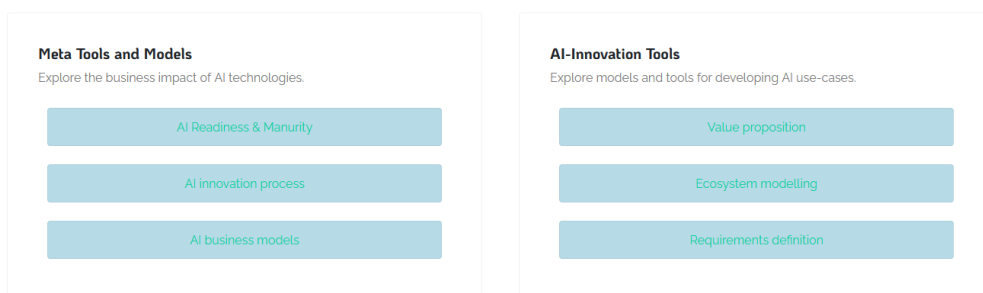


Abbildung 28: Darstellung des KI-Innovationsbaukastens auf der Website

Die entwickelten Werkzeuge, Methoden und Workshop-Konzepte des Baukastens werden nachfolgend im Detail vorgestellt. Eine detaillierte Beschreibung der extern gesammelten Werkzeuge und Methoden ist nicht vorgesehen. Diese können jedoch im vollständigen Baukasten (vgl. Abbildung 28) eingesehen werden.⁵

i) Gestaltung und Innovation von KI-basierten Dienstleistungen (KI-V^{di}P)

Das im Zuge von AP 5.2 und AP 5.4 weiterentwickelte und in [JAL21] wissenschaftliche beschriebene Werkzeug KI-V^{di}P (vgl. Abbildung 29) ermöglicht Unternehmen eine strukturierte Formulierung von prägnanten Wertversprechen, welche gleichzeitig den Mehrwert sowie die einzigartigen Facetten der entwickelten KI-Lösung abbilden.

⁵ <https://ecosystem-tools.de/AI-Toolbox/AI-Toolbox.html>


Abbildung 29: Das VdiP-Tool zur Entwicklung KI-basierter Wertversprechen⁶

Es hilft bei der Fokussierung auf die wesentlichen Faktoren einer Lösung. Die Operationalisierung als Canvas, welches eine allgemein bekannte und vielgenutzte Darstellungsart für Wissenswerkzeuge ist, ermöglicht eine schnelle und einfache Nutzung ohne großen Schulungsaufwand. Gleichzeitig wurde ein begleitendes Workshopkonzept erstellt. Das Format wird durch Experten begleitet und möchte, neben der Formulierung eines KI-basierten Wertversprechens, den Austausch zwischen den beteiligten Akteuren fördern, um dadurch in einem Entwicklungsteam ein besseres Verständnis für die Sichtweisen und Fokuspunkte des Gegenübers zu schaffen. Der Workshop wurde im Laufe des Förderprojekts mit Konsortialpartnern und Unternehmen erfolgreich durchgeführt sowie in der Lehre im Zuge von Executive Education Programmen eingesetzt. Ergänzend wurden außerdem hybride Inspirationskarten gestaltet, mit deren Hilfe Nutzer oder Workshopmoderatoren die Kreativität im Formulierungsprozess erhöhen oder festgefahrene Diskussionen auflösen können.

ii) Taxonomie zur systematischen Beschreibung, Dokumentation und Kommunikation von KI-Anwendungsfällen

Um den Bedarf nach Definitionsmöglichkeiten von Anforderungen an KI-Anwendungsfälle zu adressieren, wurde ein Taxonomie-Artefakt (siehe Abbildung 30) iterativ entwickelt. Zunächst wurden 87 Veröffentlichungen aus der Praxis (Fach-Magazine, Artikel, Berichte, etc.) recherchiert und analysiert als auch bekannte Plattformen wie *huggingface.com* und *kaggle.com* nach relevanten Inhalten durchsucht. Als zweite Quelle wurden 22 Interviews mit Experten aus der Praxis durchgeführt. Auf Basis dieser Ergebnisse wurde eine Ontologie entwickelt, mit der KI-Anforderungen multi-dimensional beschrieben

⁶ Das Tool steht unter anderem hier zur Verfügung: <https://ecosystem-tools.de/AI-Toolbox/AI-value-proposition.html>

werden können. Diese wurde in zwei Fokusgruppen evaluiert und weiterentwickelt. Die Taxonomie wurde in [KIR22] Veröffentlicht und steht zudem über den KI-Innovationsbaukasten zur Verfügung⁷.

AI-Use-Case Requirements Taxonomy

Implementing an AI use-case means considering more than just technological requirements and thus involves interdisciplinary teams that define these requirements. A requirements taxonomy for AI-use-cases was developed by Kirschbaum et al. (2022), that can be explored below:



Abbildung 30: Taxonomie-Auszug aus der Ontologie zu Anforderungen an KI-Anwendungsfälle

Das Tool dient Unternehmen als Template und Sammlung von bekannten (Best Practice) Anforderungen an KI-Anwendungsfälle in der intelligenten Produktion. Nicht immer sind für einen Anwendungsfall alle Anforderungen relevant, bieten aber einen Überblick über häufig genutzt und berücksichtigte Faktoren. Diese reichen von organisationalen Faktoren über technische und soziale bis hin zu ethischen und rechtlichen Gesichtspunkten.

iii) Rollenbilder für die Charakterisierung verschiedener intelligenter Agenten zur Unterstützung der Integration solcher in Wertschöpfungssysteme der intelligenten Produktion

Eine weitere Studie beleuchtet Integrationsoptionen für KI-basierte intelligente Agenten, welche in produktionsnahen Prozessen und Dienstleistungen wertschöpfend eingesetzt werden können. Hierbei wurde die Plattform *Crunchbase* als Datenquelle genutzt, um Rollen solcher intelligenter Agenten herauszuarbeiten, indem bisherige Einsatzkontexte untersucht werden. Dabei wurden über 100 Unternehmen mit über 130 Lösungen analysiert, um die verschiedenen Rollen zu typisieren, welche intelligente Agenten in Wertschöpfungssystemen einnehmen können. Die Inspirationskarten zu Rollen von KI-basierten intelligenten Agenten (vgl. Abbildung 32) dienen als Erweiterung des Baukastens um eine Übersicht von Gestaltungsoptionen für die Diversifizierung von Wertversprechen und die Innovation von Geschäftsmodellen. Sie stellen aus wissenschaftlicher Sicht ein Grenzobjekt dar, welches den Informations- und Ideenaustausch zwischen homogenen sowie heterogenen Akteuren fördern soll. Konkret unterstützt das Forschungsergebnis als praktischen Beitrag, Manager sowohl in der Gestaltung als auch im Management datenbasierter Geschäftsmodelle, indem es Einsatzszenarien für intelligente Agenten aufzeigt. Gleichzeitig ermöglichen die Rollenbilder eine bessere Abschätzung von potenziellen Auswirkungen auf sozio-technische Systeme und helfen Managern bei der Planung des Change-Managements gegenüber Mitarbeitern und weiteren Stakeholdern während der Einführung solcher auf intelli-

⁷ Link zum Tool: <https://ecosystem-tools.de/AI-Toolbox/AI-requirements-taxonomy.html>

genten Agenten basierender KI-Geschäftsmodelle und -prozesse. Die Nutzbarkeit der sechs Inspirationskarten als Grenzobjekt wurde im Zuge von 7 Fokusgruppen mit über 50 Teilnehmern aus Praxis und Wissenschaft evaluiert. Die Erkenntnisse der Forschung zu Integrationsoptionen für KI-basierte intelligente Agenten in Form der leicht zugänglichen Inspirationskarten wurden im Rahmen eines Workshops auf der 18. Internationalen Tagung Wirtschaftsinformatik in Paderborn präsentiert. Es ist geplant die Ergebnisse als wissenschaftlichen Beitrag bei der 19. Internationalen Tagung Wirtschaftsinformatik in Würzburg einzureichen.



Abbildung 31: Exemplarische Darstellung der Vorder- und Rückseite einer der sechs entwickelten Inspirationskarten zu Rollen von KI-basierten intelligenten Agenten

iv.) Charakterisierung und Typisierung von KI-Geschäftsmodellen und v. Kartenspiel zur Gestaltung von KI-Geschäftsmodellen

A framework for the characterization of artificial intelligence business models

Dimension	AIBM Element	Characteristics							
Value proposition	Value proposition	Optimization	Individualization	Matching	Novelty	add new			
Value architecture	Key resources	Proprietary	Stakeholder-provided	Public	Acquired/Rented	add new			
Value architecture	Core AI capability	Clustering/Categorization	Classification	Detection	Generation	Prediction	add new		
Value architecture	AI integration type	Assistance	Augmentation	Automation	add new				
Value network	Customer	B2B	B2C	B2G	add new				
Value network	Key partner role	Data provider	Technology provider	Integrator	Investor	add new			
Value finance	Sales model	Ownership	Use/Availability	Result	add new				
Value finance	Revenue model	Asset sale	Lending/Leasing/Licensing	Subscription	Usage fee	Third-party	Freemium	Open source	add new

Export Data

Datei auswählen

Keine ausgewählt

Upload

Abbildung 32: Tool zur Charakterisierung und Typisierung von KI-Geschäftsmodellen⁸

⁸ Das Tool kann über nachstehenden Link genutzt werden: <https://ecosystem-tools.de/AI-Toolbox/AIBM-characteristics>

Das in AP5.1 entwickelte und in [WIL21] veröffentlichte Framework zur Charakterisierung und Typisierung von KI-Geschäftsmodellen wurde auf vielfältige Weise multimedial für die Praxis operationalisiert:

- (A) Überführung in ein digitales, interaktives Werkzeug (vgl. Abbildung 32). Dieses Werkzeug kann von Unternehmen eigenständig oder begleitet durch ein von der FAU entwickeltes Workshopkonzept zum Design von KI-Geschäftsmodellen angewandt werden. Der Gestaltungsprozess dieses Werkzeugs wurde in [JAL21] wissenschaftlich verwertet.



Abbildung 33: Druckvorlage für das Kartenspiel zur Gestaltung von KI-Geschäftsmodellen⁹

- (B) Gestaltung als Kartenspiel. In Zusammenarbeit mit einem Grafikdesigner wurden die acht zentralen Elemente von KI-Geschäftsmodellen und die 33 Ausprägungen, welche die Geschäftsmodellelemente annehmen können, als Kartenspiel gestaltet. Druckvorlagen für den Selbstausdruck und Selbstbeschnitt auf Papier oder dünner Magnetfolie für Verwendung an einem Whiteboard werden Unternehmen über die Website zur Verfügung gestellt (vgl. Abbildung 33). Die Vorlagen sind sowohl

⁹ Das Tool kann über nachstehenden Link genutzt werden: <https://ecosystem-tools.de/AI-Toolbox/AIBM-Card-game>

auf Deutsch als auch auf Englisch verfügbar. Das Kartenspiel als Papiausdruck und als Magnete sind ebenfalls Teil des von der FAU entwickelten Workshopkonzepts zum Design von KI-Geschäftsmodellen.

- (C) Verfügbarkeit als Präsentationsfolie. Das Framework zur Charakterisierung von KI-Geschäftsmodellen ist auch als Präsentationsfolie über die Website downloadbar. Diese kann mit einem Videoprojektor entweder auf ein mit lösbaren Stiften beschreibbares Whiteboard oder eine Wand, welche mit Post-Its beklebt werden kann, projiziert werden.

Das im Projekt entwickelte Werkzeug kann auf mindestens zwei verschiedene Arten verwendet werden. Zum einen unterstützt das Werkzeug die Gestaltung und Innovation von eigenen KI-Geschäftsmodellen, indem es, durch Vorgabe der wichtigsten Elemente, den Gestaltungsprozess strukturiert und durch Aufzeigen von Gestaltungsoptionen die Kreativität fördert. Zum anderen kann das Werkzeug dazu dienen, die KI-Geschäftsmodelle von Konkurrenzunternehmen strukturiert zu analysieren und sein eigenes Geschäftsmodell, die Positionierung der eigenen Produkte und Dienstleistungen im Markt sowie die korrespondierende Unternehmensstrategie entsprechend anzupassen. Außerdem dient das Werkzeug innerhalb von Unternehmen dazu, ein gemeinsames Denkmodell in Bezug auf Geschäftsmodelle sowie gemeinsame Bezeichnungen für die jeweiligen Elemente zu etablieren, um den Austausch zwischen interdisziplinären Unternehmensakteuren mit unterschiedlicher fachlicher Sozialisation zu vereinfachen.

vi) Entwicklung von KI-Innovationsökosystemen

Für die Erhebung von Geschäftsmodellen und Innovationsbedarfen wurde auf Basis des Ökosystemmodells von [TAL20] eine Workshop-Methodik entwickelt (siehe Abbildung 34), um in Multi-Akteur-Systemen KI-basierte Wertversprechen weiterzuentwickeln. Die Methodik wurde unter anderem auf die in IIP-Ecosphere entwickelte IIP-Ecosphere Plattform (AP 2), den Daten-Marktplatz (UAP 3.5) und das Experimentierfeld (UAP 11.2) angewendet. Dabei wurden Geschäftsmodellrollenbilder entwickelt¹⁰, die im Rahmen der Publikation [KIR23] veröffentlicht wurden.

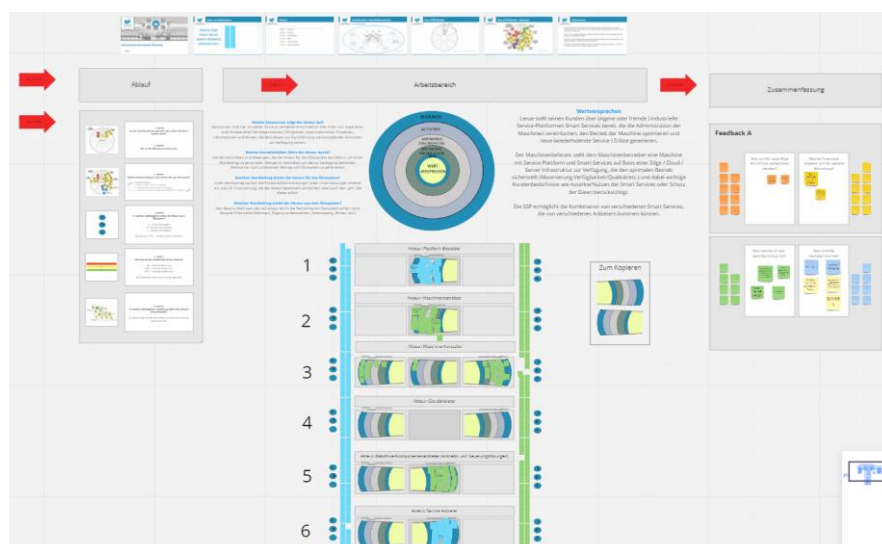


Abbildung 34: Auszug aus dem Workshop zu KI-Innovationsökosystemen

¹⁰ <https://roletool.ecosystem-tools.de/Ecosystem-Designer.html>.

Der Workshop kann zum einen genutzt werden, um existierende Ökosysteme zu analysieren oder um neue zu entwickeln. Zum einen werden Unternehmen dadurch befähigt über eine isolierte Betrachtung ihrer KI-basierten Lösung hinauszudenken. Dies wird durch die Berücksichtigung unterschiedlichster für das Ökosystem relevanter Akteure erreicht, die das Wertversprechen einer Lösung beeinflussen. Zum anderen werden die Kompetenzen der jeweiligen Akteure systematisch analysiert und modelliert. Dadurch können zum einen Kompetenzlücken aufgedeckt, aber auch Zuständigkeiten mit Partnern klar definiert und zugewiesen werden. Zuletzt ermöglicht der Workshop eine Analyse von Abhängigkeiten zwischen den Akteuren und unterstützt somit die strategische und operative Entscheidungsfindung. Sowohl die Ergebnisse der Workshops als auch die entwickelte Workshop-Methodik sind im Baukasten verfügbar.¹¹

Chatbot für die KI-basierte Unterstützung bei der Einführung der IIP-Plattform

Im Rahmen der Zusammenarbeit mit dem Think Tank Plattformen wurde ein Chatbot-Prototyp entwickelt, der beim Onboarding von neuen Nutzern unterstützt.¹² Der Chatbot hat Zugriff auf die Dokumentation der IIP-Plattform, als auch auf das Plattform-Handbuch. Damit können Nutzer Fragen zur IIP-Plattform in natürlicher Sprache stellen und bekommen eine entsprechende Antwort auf diese.

Es können Anwendungsfälle wie Verständnisfragen, Anleitungen oder auch konkrete Informationen zu Fehlern beim Aufsetzen der Plattform und FAQs gestellt werden. Der Chatbot wird in Onboarding-Sessions eingesetzt und kann Nutzern auch nach den Workshops noch Fragen beantworten. Er erleichtert damit den Support-Aufwand des Entwicklerteams und unterstützt die Integration der IIP-Plattform in Unternehmen.

2.1.6 KI-Accelerator (AP 6)

Der KI-Accelerator zielte auf eine Beschleunigung der Anwendung von Methoden der Künstlichen Intelligenz in der Produktion, insbesondere durch praktische Unterstützung wie einen KI-Lösungskatalog, Machbarkeitsstudien aber auch Arbeiten zu Referenz- und Standardkomponenten. Da sich hier deutliche Synergien zu den Arbeiten des Think Tanks Plattform und der IIP-Ecosphere-Plattform (AP2) abzeichneten, wurden von Beginn des Projekts Meetings zum KI-Accelerator durchgehend mit Beteiligung von AP2 durchgeführt und insbesondere durch eine gemeinsame AP-Leitung von AP2 und AP6 verstärkt.

Der Kickoff des KI-Accelerators fand am 12.03.2020 in Hildesheim statt. Weitere Besprechungen zu gemeinsamen Arbeiten fanden im drei-Wochen-Rhythmus als Online-Meetings statt. Neben der Berichterstattung zu aktuellen Arbeiten beider Bereiche wurden an verschiedenen Stellen gemeinsame Arbeiten (KI-Lösungskatalog, Datenformate, I4.0 Verwaltungsschalen, Plattform-Vision, Plattform-Architektur, Messe-Demonstratoren) identifiziert und gemeinsam bearbeitet. Zudem fanden auf Einladung Vorträge zu Arbeiten anderer Projekte im gleichen Förderprogramm (z. B. zu SPAICER von SAP) aber auch zu relevanten dynamischen Demonstratoren (NovoAI, MIP) statt. Das breite Themenspektrum aber auch die interessanten Diskussionen motivierten andere Projektpartner, sich dauerhaft an den Meetings des KI-Accelerator teilzunehmen, z. B., Sennheiser und Phoenix Contact.

¹¹ Link zum Tool: <https://ecosystem-tools.de/Ecosystem-Modelling/Ecosystem-Modelling.html>

¹² <https://iip.ecosystem-tools.de/> (Login: iip, PW: iip)

KI-Lösungskatalog (& Industrieumfrage) (UAP6.1)

Wie bereits in der Einleitung erwähnt ergaben sich aus der regelmäßigen Diskussion im Kreis von interessierten Partnern aus Industrie, Forschung und Verbänden auch zahlreichen Anregungen für weitere Aktivitäten in IIP-Ecosphere. Dazu gehört insbesondere die Durchführung einer größeren, ursprünglich nicht geplante Studie zu KI-Nutzung in Form von einer Industrieumfrage, die hier ebenfalls beschrieben wird.

Bedarfsermittlung und Industrieumfrage

Die Idee des Lösungskatalogs und der damit verbundenen Bedarf in Bezug wurden im KI-Accelerator intensiv diskutiert. Die Ergebnisse dieser Diskussion sind in den Entwurf des Beschreibungsschemas für die KI-Lösungen und in den Entwurf für den Prototyps des KI-Lösungskatalogs eingeflossen (s.u.)

Aus der Diskussion im KI-Accelerator und insbesondere auf Anregung des Projektpartners VDW ergab sich die Idee eine Studie zur KI-Nutzung in Form einer Industrie-Umfrage durchzuführen. Ziel der Studie war, die Anforderungen für die Arbeiten im Accelerator und im IIP-Ecosphere Projekt insgesamt noch weiter zu schärfen und auf dieser Grundlage die Arbeiten in IIP-Ecosphere noch stärker zu priorisieren. Die Industrie-Umfrage umfasst neben Fragen zum Thema KI auch Fragen zu Themen, die eng mit der Nutzung von KI-Methoden verknüpft sind wie etwa zur Datennutzung, zu Industrie 4.0 Plattformen und zum Cloud-Computing. Zudem enthält sie auch Fragen mit direktem Bezug zum K-Lösungskatalog. Die in 6 Unterkategorien organisierten Fragen gemeinsam im KI-Accelerator erarbeitet und in Diskussionen schrittweise verfeinert.

Schließlich wurde die Umfrage in ihrer endgültigen Form mit 28 Fragen als online-Survey aufgesetzt. In enger Zusammenarbeit mit dem Projektpartner VDW wurde die Umfrage aktiv beworben: Die Industrie-Umfrage wurde durch den VDW in 2020/2021 auf verschiedenen Kanälen (direkt und indirekt) im Umfeld des VDW und VDMA verbreitet u.a. über E-Mail-Direktkontakte aus der Werkzeugmaschinenbranche; über die VDW eigene Zeitschrift Branchenreport, über die VDW-Homepage und Social Media Kanäle sowie über den VDMA Arbeitskreis Industrie 4.0 und den VDMA Industrie 4.0 Newsletter. Eine weitere Verbreitung fand auch über andere Projektpartner statt (u.a. IHK, Deutsche Messe).

Durch die aktive Bewerbung konnte eine hohe Beteiligung von Insgesamt 75 Unternehmen erreicht werden (vollständige Beantwortung). Die Ergebnisse der Umfrage wurden analysiert und in einem Whitepaper [NIE21a] zusammengefasst, das auf der Web-Seite des Projekts publiziert wurde. Es ergaben sich wichtige Erkenntnisse insbesondere in den Bereichen Erwartungen zu KI, Diversifizierung von KI-Anwendungen, Datenflut und Datenmangel, Cloud-Nutzung, IIoT Plattformen sowie "Mensch und KI".

Zudem gab es einen Medien-Roundtable zur Umfrage (mit nachfolgenden Artikeln in Fachmedien), der unter der Leitung des VDW organisiert wurde und die Sichtbarkeit der Umfrageergebnisse weiter erhöht hat. Aus diesem Event ergaben sich eine Anzahl von Artikeln in Fachzeitschriften und zwar im Industrieanzeiger, in „Produktion - Technik und Wirtschaft für die deutsche Industrie“, im DIMA-Magazin und in NC-Fertigung

Die Ergebnisse der Umfrage wurden zudem als Anforderungen und zur Priorisierung von Aktivitäten in die weitere Gestaltung des Projektes eingebracht (siehe Diskussion im Whitepaper).

Beschreibung, Kategorisierung und Strukturierung

Der KI-Lösungskatalog soll den Zugang zu einer Vielzahl von bereits existierenden und im Projekt entstehenden KI-Lösungen erleichtern. Ein wichtiger Aspekt des Aufbaus Lösungskatalogs ist die Frage, wie die einzustellenden Lösungen in geeigneter Form beschrieben und strukturiert werden sollen. Ziel war es dabei sowohl eine einfache Auffindbarkeit von Lösungen zu unterstützen als auch eine initiale Entscheidung über die Einsetzbarkeit einer aufgefundenen Lösung im eigenen Anwendungskontext zu ermöglichen. Als Grundlage für die Erarbeitung des--Beschreibungsschemas wurden existierende KI-Lösungen gesammelt und gesichtet. Dazu wurde eine Literaturrecherche in Zusammenarbeit mit Projektpartnern des KI-Accelerators durchgeführt, die auch relevante Beschreibungsformate umfasst hat (z. B. aktuelle Publikationen zu KI, I4.0, KPIs, Digital Twins aus VDMA, VDI, CECIMO u.a.). Dies umfasste auch die Sichtung von Aktivitäten in Normung und Regulierung von KI-Anwendungen (z. B. DIN-KI Normungs-Roadmap; EU-Informationsmaterial zum Artificial Intelligence Act, VDMA Analyse und Positionspapier).

Auf dieser Grundlage wurde inkrementell ein Beschreibungsschema für den KI-Lösungskatalog abgeleitet (Kategorisierung und Strukturierung), unter Verwendung existierender Taxonomien und Standards. Dies umfasst auch die Auswahl von Vokabularen für die Wertedefinition unter Berücksichtigung existierender Taxonomien. Abbildung 35 fasst das erarbeitete Beschreibungsschema, das in fünf Teilbereiche strukturiert ist, zusammen.

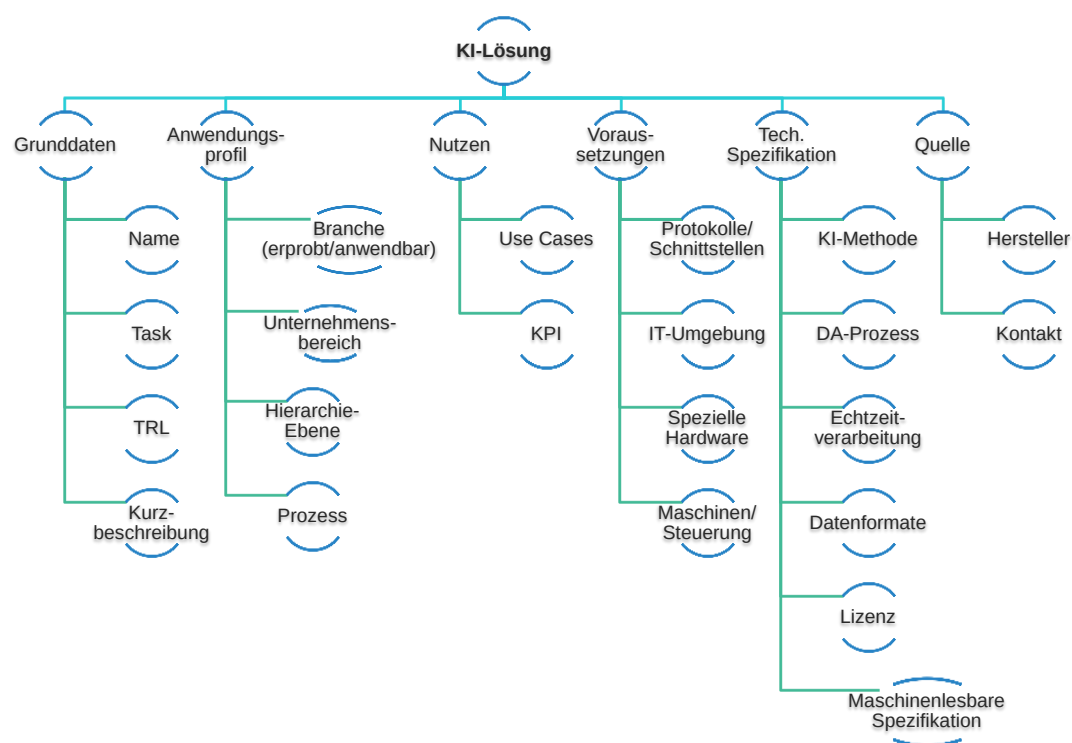


Abbildung 35: Beschreibungsformat für KI-Lösungen im Lösungskatalog

Die Eignung des Beschreibungsformat wurde gemeinsam mit den Projektpartnern aus der Industrie im KI-Accelerator durch Nutzung des „Strukturierungs-Prototyps“ für die Beschreibung eigener KI-Lösungen prototypisch erprobt (Erste Evaluation/Pre-Experiment).

Entwurf und Entwicklung des Lösungskatalogs

Der KI-Lösungskatalog soll den Zugang zu einer Vielzahl von bereits existierenden und im Projekt entstehenden KI-Lösungen erleichtern. Zur Vorbereitung der Umsetzung wurde das entwickelte Beschreibungsschema (s.o.) inkrementell in einen Entwurf für ein Schema für die Speicherung der Daten überführt. Zudem wurde gemeinsam eine erste Konzeption für die im Lösungskatalog benötigten Funktionen entwickelt und ein Entwurf einer webbasierten Benutzerschnittstelle erarbeitet.

Auf der Grundlage des Entwurfs wurde ein funktionaler Online-Prototyp unter Nutzung existierender Frameworks implementiert. Dabei wurde insbesondere die Nutzungsprozesse und die Benutzerschnittstelle intensiv diskutiert, erprobt und in mehreren Iterationen verfeinert. Um den Lösungskatalog attraktiver zu machen, wurde auch das Design mehrfach in Kooperation mit Partnern aus der Industrie mehrfach verbessert. Die Umsetzung des KI-Lösungskatalogs wurde im Projekt erfolgreich abgeschlossen und der Lösungskatalog steht zur Verfügung.

Evaluation

Dem Thema Evaluation wurde bei der Entwicklung des KI-Lösungskatalogs besondere Beachtung geschenkt. Bereits in der frühen Phase wurden die Beschreibungsformate über einen Strukturierungsprototypen evaluiert (s.o.). Nach Fertigstellung einer ersten Version des Prototyps Zur Evaluation wurde der Lösungskatalog probeweise mit ersten Lösungseinträgen der Partner befüllt ist. Hier wurden Beschränkungen erfasst und zur weiteren Verfeinerung des Katalogs genutzt. Auch für spätere Versionen wurde kontinuierlich Feedback von den Projektpartnern zu Funktionalität und Nutzbarkeit gesammelt und zur weiteren Verbesserung des Lösungskatalogs verwendet.

Darüber hinaus wurden auch andere Dinge im Kontext des Lösungskatalogs evaluiert. Hierzu zählen die Möglichkeiten mittelfristig Verwaltungsschalen im Lösungskatalog zu benutzen. Hier wird insbesondere Potential im Bereich der technischen Spezifikation (siehe Abbildung oben zum Beschreibungsschema) gesehen. Außerdem wurde in enger Kooperation mit AP2 diskutiert, wie mögliche Interaktionen (bzw. eine Integration des Lösungskatalogs mit der IIP-Ecosphere Plattform aussehen könnten. Hier wurde z. B. die Idee eines Service Store diskutiert.

Machbarkeitsstudien (UAP 6.2)

In diesem UAP wurden verschiedene Studien zu KI-Anwendungen in verschiedenen Kontexten durchgeführt, insbesondere um Aspekte von Datenerfassung, Datentransport, Formaten und KI-Services auch in den IIP-Ecosphere-Demonstratoren zu erproben und zu diskutieren, bevor diese in weiteren Schritten z. B. als Referenzkomponenten in UAP 6.3 weiterentwickelt werden. Diese Schritte sind ebenso Grundlage der Erprobung der IIP-Ecosphere Plattform in industriellen Anwendungen bzw. im industriellen Kontext und dienen daher als geeigneter Rahmen, Plattform-Aspekte außerhalb von AP2 vorzubereiten, zu diskutieren und gemeinsam zu entwickeln. Daher leistete insbesondere SSE@UHI Beiträge zu diesem UAP, auch wenn im Projektplan keine formale Beteiligung vorgesehen war.

Im Rahmen der Identifikation des Bedarfs (Arbeitsschritt 6.2.1) erfolgte eine Bedarfsermittlung. Als Grundlage dienten insbesondere anwendungsgetriebene Bedarfe aus den IIP-Ecosphere Demonstratoren, z. B., für Pseudofehlererkennung in der automatisierten Endkontrolle (Slashwhy, Rapidminer), Platinenpositionserkennung (Bitmotec) oder Peakerkennung in Datensätzen (Slashwhy, L3S, IFW).

Hinzu kamen Anwendungsfälle aus der eigenen Produktion bzw. Entwicklung, wie beispielsweise die Verarbeitung von Akustikdaten im End-of-Line-Test einer realen Produktionslinie basierend auf einer dauerhaften Datenaufzeichnung mit entsprechender automatisierter Fehlererkennung (Lenze), einer Getriebequalitätserkennung mit bzw. ohne föderiertem Lernen (Lenze, L3S) oder der Diskussion einer Anwendung der IIP-Ecosphere-Arbeiten auf den bereits existierenden Lenze-Demonstrator der mobilen Förderstrecke (Lenze, UHi mit vorbereitender studentischer Arbeit). Weiterhin wurden von allen Partnern Anwendungsfälle für die IIP-Ecosphere Messedemonstratoren diskutiert, hier beispielsweise erforderliche KI-Services für visuelle Qualitätserkennung oder für Condition-Monitoring für einen Linearantrieb (insbesondere Lenze, IFW). Weiterhin wurden Anwendungsfälle der IIP-Ecosphere Plattform und sich daraus ergebende initiale Machbarkeitsstudien mit den Partnern, mit assoziierten Partnern (insbesondere Phoenix Contact, Trevisto) aber auch mit weiteren Unternehmen (z. B. embedded ocean GmbH, Hoerbiger Antriebstechnik GmbH) diskutiert. Besondere synergistische Impulse aus den gemeinsamen Diskussionen waren der Fokus auf Industrie 4.0 Verwaltungsschalen (getrieben durch Lenze) wie auch die gemeinsame Vision der geplanten Plattform-Funktionalitäten.

In Arbeitsschritt 6.2.2 (Planung) wurden die identifizierten Bedarfe priorisiert und konkrete Arbeitsschritte der Partner beschlossen. Insbesondere wurden Zusammenarbeiten und Abgleiche mit den IIP-Ecosphere-Demonstratoren sowie mögliche Integration von Diensten in die IIP-Ecosphere-Plattform besprochen. Insbesondere wurden Synergien und Beiträge der Machbarkeitsstudien zu den öffentlichen Demonstrationen für die Hannover Messe 22, die Tage der digitalen Technologien 2022 in Berlin, die Hannover Messe 2023, den Tag der Informatik in Hildesheim, die Vorstellung bei Nürnberg Digital 2022 und die EMO 2023 diskutiert und geplant. Beispiele hierfür sind die verschiedene KI-Services für die visuelle Qualitätskontrolle (L3S), eine Echtzeit-Cobotsteuerung (Phoenix Contact), eine Antriebssteuerung mit integrierter Messwerterfassung (Lenze, IFW), ein Condition-Monitoring für die Antriebssteuerung (Lenze, Rapidminer, ISMLL@SUH), eine Energiemessung (Phoenix Contact, IFW), eine visuelle Erkennung von Werkstücken (NovoAI) sowie eine magnetische Produktidentifikation (MIP), die insbesondere mit AP2 (SSE@UHi) abgestimmt wurden. Weiterhin wurden individuelle Machbarkeitsstudien (Lenze, Phoenix Contact, embedded Ocean), die Umsetzung und Integration von Plattformdiensten neben UAP 6.5 und UAP 6.6 (Rapidminer, Trevisto) sowie gemeinsame Arbeiten zur Performanzevaluierung der Plattform (Hoerbiger) abgesprochen und geplant.

In Arbeitsschritt 6.2.3 wurden geplante Machbarkeitsstudien aus Arbeitsschritt 6.2.2 durchgeführt. In den IIP-Ecosphere-Demonstratoren wurde beispielsweise die Machbarkeit von KI unterstützte Analyse der Leiterplattenendkontrolle mit verschiedenen KI-Verfahren (Decision Trees, Random Forests, bagging, SVNs, deep neural networks) erprobt und bestätigt (Slashwhy, Rapidminer, Bitmotec) und letztlich in die IT-Struktur bei Sennheiser integriert und öffentlich demonstriert. Weiterhin wurde die Peakerkennung für die Takzeioptimierung im VNW Demonstrator erprobt und demonstriert (Slashwhy). Insbesondere wurden diverse Arbeiten durchgeführt, die letztlich im Rahmen der Messe-Demonstratoren öffentlich vorgestellt wurden. Eine Grundlage hierzu war insbesondere die Durchführung einer studentischen Arbeit zur Realisierung/Evaluation von Plattform-Komponenten für die modulare Förderstrecke (Lenze, SSE@UHi), die im Rahmen des Condition Monitoring für die EMO 2023 Verwendung fand. Bitmotec installierte im Rahmen der Demonstratoren für die Hannover Messe 2023, den Tag der Informatik, Nürnberg Digital und EMO 2023 ihre Betriebssystem Bitmotec auf einem eigens beschafften industriellen Edge-Gerät, das von IFW in den Demonstratoraufbau integriert wurde. Hierzu gehörten

insbesondere die Virtualisierung von Bitmotec, das Aufsetzen von NodeRED, Influx und Grafana als Basis für eine Integration mit der IIP-Ecosphere Plattform sowie die gemeinsame Entwicklung anpassbarer Visualisierungen für die genannten Demonstratoren. Weiterhin realisierten die beteiligten Partner (L3S, IFW, Lenze, Rapidminer, SUH, MIP) die in Arbeitsschritt 6.2.2 benannten Services. Die Services wurden mit von der IIP-Ecosphere-Plattform generierten Tests für die Demonstrator-Anwendungen validiert und erfolgreich validierte Services wurden integriert und öffentlich demonstriert. Zur

Die gewonnenen Erfahrungen, Erkenntnisse und Ergebnisse wurden schließlich in Arbeitsschritt 6.2.4 (Analyse der Ergebnisse) reflektiert. Hierzu gehören insbesondere die Diskussion von Cobot-Integrationsproblemen mit Universal Robots aufgrund der Erfahrungen zur Hannover Messe 2022 (Phoenix Contact, SSE@UHi). Die Probleme wurden im Rahmen der Echtzeit-Cobotsteuerung für die Hannover Messe 2023 und die folgenden Demonstratoren erfolgreich gelöst. Die Erkenntnisse aus den Vorbereitungsarbeiten für die modulare Förderstrecke wurden in die Diskussion der Ausrichtung des IIP-Ecosphere Experimentierfelds eingebracht (Lenze, Phoenix Contact, SSE@UHi). Neben technischen Machbarkeitsstudien wurde insbesondere mit den IIP-Ecosphere Partnern Siemens, Lenze, IFW ein Service-Integrationsworkshop für die IIP-Ecosphere Plattform durchgeführt und damit die Plattform sowie pragmatische Schritte zur Realisierung und Integration von Referenzimplementierungen gemeinsam exploriert. Das Feedback aus diesem Workshop war für die weiteren Arbeiten an Machbarkeitsstudien und der Plattform höchst relevant und wertvoll. Ergebnisse aus den Machbarkeitsstudien wurden unter anderem in verschiedenen Versionen des Plattformhandbuchs [EIC20, EIC21a, EIC22a, EIC22d, EIC22e, EIC23a, EIC23c], bezüglich KI auf einem industriellen Edge-Device in [EIC22b] und bezüglich Integration verschiedenster Komponenten in [EIC22a, EIC22c] dokumentiert.

Ergänzend hierzu wurden vom VDW in Arbeitsschritt 6.2.1 zur Sammlung von Informationen drei Literaturrecherchen durchgeführt:

- In 2020 zu KI und Industrie 4.0 (z. B. aktuelle Publikationen und Studien zu KI, I4.0, KPIs, Digital Twins aus VDMA, VDI, CECIMO u.a.) sowie auch Standards und Normen (z. B. VDI Richtlinien, VDMA Einheitsblätter, ISO-Normen), insbesondere auch um die Entwicklung des Lösungskatalogs und das Sammeln von KI-Lösungen (UAP 6.1.2) zu unterstützen.
- In 2021 lag der Fokus auf Hinweisen zu Aktivitäten in Normung und Regulierung von KI-Anwendungen (z. B. DIN-KI Normungs-Roadmap; EU-Informationsmaterial zum Artificial Intelligence Act, VDMA Analyse und Positionspapier).
- In 2022 ging es hauptsächlich um die Sammlung von Beispielen für KI in der Produktion (d.h. konkrete Anwendungsbeispiele in Forschung und Industrie).

Die Literaturrecherchen waren ebenso Grundlage für die gemeinsame Gestaltung und Durchführung einer Online Industrie-Befragung zum Thema KI-Einsatz in der intelligenten Produktion (siehe auch UAP 6.1.1) in 2021 mit einer Neuauflage zur EMO im Jahr 2023. Die Industrie-Online Umfrage enthielt weiterhin Informationen, die für die Planung und Ausrichtung der Machbarkeitsstudien und der Demonstratoren (insbesondere Arbeitsschritt 6.2.1 und 6.2.2) sowie für die Anforderungen an die IIP-Ecosphere Plattform und die Ausrichtung der Umsetzung in AP2 relevant war.

Zudem wurde vom VDW in der Evaluation der Informationstypen (Arbeitsschritt 6.2.2) eng mit UAP 6.4 und AP2 auf dem Gebiet der OPC UA Companion Specifications zusammengearbeitet (VDW,

SSE@UHi). Insbesondere wurden die maschinenlesbaren XML-Spezifikationen der OPC UA Companion Specifications automatisiert in das Plattform-Konfigurationsmodell (IVML) übersetzt und zur generativen Ableitung von OPC UA Konnektoren (auch aus Teilmodellen der OPC UA Companion Specifications) genutzt. Die gemeinsame Validierung des Ansatzes mit SSE@UHi zeigte, dass alle 55 aktuell bekannten OPC UA Companion Specifications korrekt in Java-Code übersetzt werden können und bereits im UMATI OPC UA Testserver implementierte Companions Specs auch korrekt gelesen werden können.

Referenzimplementierungen (UAP 6.3)

Ziel der Referenzimplementierungen war es, hilfreiche dokumentierte (Referenz-)Beispiele zur Implementierung ausgewählter Funktionen aus den Machbarkeitsstudien zu schaffen, die an sich bereits einen Mehrwert darstellen.

Für die Referenzimplementierungen wurde, basierend auf den Erfahrungen in UAP 6.2, in Arbeitsschritt 6.3.1 der Bedarf identifiziert. Als aussichtsreiche Kandidaten für Referenzimplementierung wurden die Pseudofehler-Erkennung aus dem Sennheiser-Demonstrator (Slashwhy, Rapidminer), die Platinenpositionserkennung (Bitmotec), das Föderierte Lernen als generische, wiederverwendbare Implementierung (L3S, die Getriebefehlererkennung (L3S, Lenze), die Peakerkennung (Slashwhy, L3S, IFW), das Condition Monitoring für Getriebemotoren (Lenze), eine Anomalieerkennung für Portalroboter (Lenze, ISMLL@UHi), ein robustes End-of-Line-Testing gegenüber Hintergrundgeräuschen (Lenze), das Hochregallager im MCC (Lenze), ein Deep-Representation-Learning für Condition Monitoring (Lenze), die Datenerfassung mit dem Drive-Data-Hub (Lenze), sowie weitere Use Cases in der Qualitätsinspektion für Getriebemotoren (Lenze) identifiziert. Für KI-Komponenten wurde eine Anforderungsanalyse durchgeführt. Analog zur Bedarfsermittlung der Machbarkeitsstudien in Arbeitsschritt 6.2.1 wurde für die benötigten Services für die öffentlichen (Plattform-)Demonstratoren mitgeplant, d.h., insbesondere für visuelle Qualitätsinspektionen oder für das Condition Monitoring (Lenze).

In Arbeitsschritt 6.3.2 (Planung) wurden initialen Pläne für die Umsetzung von Kandidaten aus Arbeitsschritt 6.3.1 erstellt und Vorarbeiten durchgeführt. Dies betraf insbesondere im Rahmen des Sennheiser-Demonstrators die Identifikation und Konzeptionierung besonders relevanter Submodule, hier Test- bzw. Prüfmodul, Materialbereitstellung, Materialabfuhr und -transport, Robotersystem und die Leiterplattenidentifikation (Bitmotec) sowie die Planung der Umsetzungsschritte für eine Referenzimplementierung für die intelligente Leiterplattenendkontrolle (Slashwhy). Weiterhin wurde für den Use Case für die Qualitätsinspektion für Getriebemotoren geplant (Lenze). Für die (Plattform-)Demonstratoren wurde ein Linear-Verfahrentrieb für einen Cobot im Demonstratortisch (Lenze, IFW), die Evolution des Services für die visuelle Qualitätsinspektion (von neuronalen Netzwerken über föderiertes Lernen zur Erhöhung der Robustheit), eine Familie alternative Condition Monitoring Services für EMO'23, sowie technisch erforderliche Daten- und Applikationsservices (Kamera-Service, MQTT und OPC UA Konnektoren, Applikationssteuerung mit Verwaltungsschalen-Repräsentation; Lenze, SSE@UHi, L3S, IFW) geplant. Insbesondere für die alternativen Condition Monitoring Services, die von Lenze, Rapidminer und ISMLL@UHi im Rahmen der gleichen Demonstrator-Applikation unabhängig realisiert werden sollten, als auch für die Visualisierung der Resultate war eine Absprache des Datenformats des Antriebs, der Eingangs- und Ausgangsdatenformate sowie der Interaktion mit den restlichen Demonstrator-Services mit den restlichen Demonstratorpartnern (SSE@UHi, L3S, IFW) erforderlich.

Im Rahmen des Arbeitsschritts 6.3.3 (Implementierung) wurde für den Sennheiser-Demonstrator zusätzlich zum späteren Einsatzort wurde bei Bitmotec vorab ein Versuchsstand konzipiert und konstruiert, der den Teil des Prüfplatzes mit der Tray Ablage und dem Kamerasystem maßstabsgetreu abbildet. Dadurch konnte bei Bitmotec bereits parallel zur Entwicklung der Prüfselle an der dem PCB-Erkennungssystem (Leiterplattenidentifikation) gearbeitet werden. Zudem wurden Referenzimplementierung der KI-Services für mehrere Leiterplatten auf zwei Prüfstationen entwickelt (Slashwhy). Lenze arbeitete am Use Case der Qualitätsinspektion für Getriebemotoren. Für die öffentlichen (Plattform-)Demonstratoren wurden in enger Zusammenarbeit mit AP2 die Applikationsmodelle realisiert, das Konfigurationsmetamodell und die Code-Generierung erweitert (insbesondere für eine modell-basierte generische Integration von förderiertem Lernen in die IIP-Ecosphere-Plattform durch L3S und SSE@UHi) und gemeinsam KI- und Anwendungsservices realisiert: Die visuelle Qualitätsinspektion wurde als Python-Service für Tensorflow und Tensorflow-lite (auf PLCNext) umgesetzt. Das Condition Monitoring wurde als Python-basierter Plattform-Service unter Integration des Lenze-Antriebs und des Drive-Data Hub realisiert. Zudem wurde insbesondere mit dem dynamischen Demonstrator zur magnetischen Produktidentifikation ein unterstützender KI-Service und ein entsprechender Datenfluss für die Demonstrationsapplikation entwickelt. Für alle Services wurden die Datenbereitstellung und der Datendurchfluss getestet, die Funktionalität mit generierten Tests und in Testläufen der jeweiligen Applikation Validierung und entsprechende Deployment-Container generiert und auch diese getestet.

Die Evaluation (Arbeitsschritt 6.3.4) bestätigte die erfolgreiche Umsetzung relevanter Referenzimplementierungen. Das PCB-Erkennungssystem (Bitmotec) wurde im Laufe des vierten Quartals 2022 final erfolgreich in den Sennheiser-Demonstrator integriert. Durch diese Einbindung konnte die Funktionalität des PCB-Erkennungssystems in einer realen Umgebung demonstriert werden. Ebenso wurde die Referenzimplementierung für mehrere Leiterplatten auf zwei Prüfstationen validiert (Slashwhy). Die Validierungen in der Sennheiser Demonstratorzelle in Kombination mit dem Roboter-system werden über das Projektende hinaus fortgeführt, um eine spätere Verwertung der Ergebnisse zu ermöglichen. Analog wurde der Use Case für Qualitätsinspektion für Getriebemotoren erfolgreich validiert (Lenze) Die KI- und Applikationsservices wurden im (Plattform-)Demonstrator-Aufbau erfolgreich öffentlich demonstriert, in verschiedenen Versionen auf der Hannover Messe 2022, den Tagen der digitalen Technologien 2022 in Berlin, der Hannover Messe 2023, dem Tag der Informatik 2023 in Hildesheim, Nürnberg Digital 2023 sowie der EMO 2023 (SSE@UHi, L3S, IFW, Phoenix Contact, Lenze, MIP, Bitmotec). Die automatisch integrierten Applikationen mit den jeweiligen Services liefen je nach Veranstaltung tagelang, zur EMO beispielsweise insgesamt fast 1000 Komplettdurchläufe, wobei Ausführungsinformationen für weitere Verbesserungen aufgezeichnet wurden, z. B., aufgenommene Bilder und Erkennungsergebnisse. Detailevaluationen z. B. zur Ressourcennutzung und zur Performance wurden im Anschluss mit diesen Service-Versionen durchgeführt.

In Arbeitsschritt 6.3.5 erfolgte die Dokumentation der Referenzimplementierung. Für den Sennheiser-Demonstrator wurde über den Projektverlauf wurden die Softwarebestandteile nach aktuellen Standards der Programmierung dokumentiert (Bitmotec, Slashwhy). Die (Plattform-)Demonstrator-Komponenten liegen als Open Source Plattform-Beispiele vor und wurden insbesondere im Plattform-handbuch [EIC20, EIC21a, EIC22a, EIC22d, EIC22e, EIC23a, EIC23c] dokumentiert. Detaillergebnisse wurden u.a. in [EIC22b, EIC22a, EIC22c] veröffentlicht. Neben den öffentlichen Präsentationen auf den

Messen und Veranstaltungen wurden ausgewählte Ergebnisse zudem in öffentlichen Vorträgen präsentiert, z. B., der Demonstrator und die IIP-Ecosphere Plattform (SSE@UHi) in der Begleitveranstaltung zu Nürnberg Digital, die IIP-Ecosphere Plattform (SSE@UHi), das Condition Monitoring (Lenze) und die KI-gestützte magnetische Produktidentifikation (MIP) auf einem Vortrag auf der EMO-Messe (SSE@UHi, IFT) und dem IIP-Ecosphere Abschlussposium.

Zusätzlich wurden in den Arbeitsschritten 6.3.1, 6.3.2 und 6.3.4 auch Arbeiten im Kontext von OPC UA durchgeführt. Dabei stand in Arbeitsschritt 6.3.1 die Erstellung einer maschinenlesbaren Parameterspezifikation im Fokus. Hierzu wurde der Bedarf für maschinenlesbare Parameter ermittelt, Erkenntnisse und Erfahrungen aus den IIP-Ecosphere Demonstratoren und den Plattform-Demonstratoren berücksichtigt und Anforderungen aus der Industrie-Umfrage (siehe Arbeitsschritt 6.1.1) entnommen. Im Rahmen der IIP-Ecosphere Plattform wurden maschinenlesbare Parameterspezifikationen insbesondere im Rahmen der Service-Parametrierung, der Konnektor-Parametrierung (hier insbesondere über OPC UA companion specifications) und der (Konnektor-)Datenflüsse als relevant identifiziert.

Die maschinenlesbaren Parameterspezifikationen wurden in Arbeitsschritt 6.3.2 im Rahmen einer Referenzimplementierung umgesetzt. Hierzu gehört sowohl die automatisierte Integration von OPC UA companion specifications in das Plattform-Konfigurationsmodell, die Anwendung der integrierten Strukturen im Rahmen konfigurierter und generierter OPC UA Konnektoren als auch die Parametrisierung von Services und Konnektoren. Die jeweilige Spezifikation findet im Rahmen des Plattform-Konfigurationsmodells für die jeweilige Anwendung statt. Parameteränderungen können insbesondere über Verwaltungsschalen-Operationen zur Laufzeit eingesteuert werden (gemeinsam mit SSE@UHi). Laufzeitparameteränderungen wurden insbesondere im Hannover Messe 2023 Plattform-Demonstrator relevant, wo mehrere Instanzen derselben Applikation unterschiedliche Fabrikstandorte repräsentieren und dort föderiertes Lernen zeigen. Die Parametrisierung erlaubt die Zuordnung der verwendeten Kameras und Cobots und damit die Auswahl der zu verwendenden Teile der OPC UA Modelle in den verwendeten Geräten.

In Arbeitsschritt 6.3.4 wurden die Ansätze zur maschinenlesbaren Parameterspezifikation evaluiert und abschließend dokumentiert. Die automatisierte Integration von OPC UA companion specifications konnte für alle 55 verfügbaren Spezifikationen erfolgreich ausgeführt werden und führte zu validen Konfigurationsmodellen. Ausgewählte Modelle konnten gegen den UMATI Testserver im Betrieb validiert werden. Die Laufzeit-Parametrierung von Konnektoren und Services im Rahmen des Hannover Messe 2023 Demonstrators war ebenso erfolgreich. Die Ergebnisse wurden in einer studentischen Projektarbeit (SSE@UHi) wie auch im Plattformhandbuch [EIC20, EIC21b, EIC22d, EIC23a, EIC23c] dokumentiert.

Überführung in anpassbare Standardkomponenten (UAP 6.4)

In diesem UAP wurden Ansätze identifiziert und Techniken näher betrachtet, die über Referenzimplementierungen hinaus Standards aufnehmen und verwenden, um KI- und Datenanalysekomponenten im Produktionsbereich noch breiter anwendbar, wiederverwendbar, parametrierbar und interoperabler gemacht. Ein besonderer Fokus wurde hierbei auf OPC UA und UMATI gelegt.

In Arbeitsschritt 6.4.1 wurden frühzeitig Anforderungen für Standardkomponenten identifiziert. Die Aktivitäten hier überlappten synergistisch mit der Anforderungssammlung für die IIP-Ecosphere-Plattform in AP2, den Machbarkeitsstudien, den ersten Schritten zu Referenzkomponenten sowie der Entwicklung

des KI-Lösungskatalogs (UAP 6.1). Bestehende Standards wie OPC UA, OPC UA companion specifications, die Arbeiten rund um UMATI und Ansätze wie ECLASS waren daher ersten und offensichtliche Kandidaten. Hinzu kam der aktuelle Trend zu Industrie 4.0 Verwaltungsschalen (im Projekt vorangetrieben durch Lenze), initial mit Blick auf das Software-Typenschild, die Dokumentation von Komponenten, das Geräte-Typenschild. Hinzu kamen diverse projektübergreifende Diskussionen zu KI-Schnittstellen (Rapidminer) oder mit ECLASS, befeuert durch eine gemeinsame öffentliche Veranstaltung rund um die Industrie 4.0 Verwaltungsschale mit dem BMBF HAISEM-Lab Projekt (mit Beiträgen von LNI4.0, ZVEI, IDTA, Lenze und IIP-Ecosphere). Weiterhin fanden intensive Diskussionen zum Thema KI-Software-Services, zur Identifikation von Konzepten zur Strukturierung und Integration von KI-Diensten (mit Beiträgen der IIP-Ecosphere Demonstratoren und der Best Practices) statt, die beispielsweise in Veranstaltungen der Begleitforschung vorgestellt wurden. Beispielsweise wurden im Sennheiser-Demonstrator die Anforderungen und UseCases auf Beiträge zu KI-Standardkomponenten hinsichtlich ihrer Übertragbarkeit als anpassbare Standardkomponenten analysiert (Slashwhy). Kernanforderungen wurden insbesondere in [EIC21a] dokumentiert und in [EIC20, EIC21b, EIC22d, EIC23a, EIC23c] diskutiert bzw. im Rahmen der Plattform-Realisierung fortgeschrieben.

In Arbeitsschritt 6.4.2 (Domänenexploration und Anpassungsmechanismen) wurden existierende und benötigte Techniken und Mechanismen für Standardkomponenten identifiziert diskutiert und mit dem Think Tank Plattformen (AP2) abgestimmt. Neben gewünschten Anpassungsmechanismen für eine Industrie 4.0 Plattform wurden insbesondere Domänenexplorationen für (Maschinen-)Konnektoren, für anpassbare Anwendungsdienste und für wiederverwendbare KI-Services durchgeführt. Für Plattformkomponenten fanden hierzu diverse Gespräche und Diskussionen mit Partnern (VDW, Lenze, Rapidminer), assoziierten Partnern (Phoenix Contact, MDZH) und externen Interessenten (incari, Schunk, Alromatic, Codewerk) statt. Insbesondere wurden weitere Erläuterungen und auch weiteres Informationsmaterial ausgetauscht und präsentiert (VDW), u.a. zum „OPC UA Online Reference- Online Versions of OPC UA Specifications and Information Models“ sowie zur „OPC UA for Machine Tools Spezifikation (UA4MT)“. Die eingebrachten Gedanken, Ideen, Erfahrungen und Anforderungen führten in weiteren Schritten zum Ansatz automatisch generierbarer Konnektoren. Der Bedarf an wiederverwendbaren und anpassbaren Diensten führte beispielsweise zur generischen Integration von Sicherheitskomponenten in UAP 6.6. Zudem erlaubte eine Domänenexploration rund um den Begriff „KI-Service“ (initial begonnen auf einer IIP-Ecosphere Klausurtagung) zu einer schärferen Charakterisierung und zur Validierung an konkreten Beispielen. Zudem wurde das Konzept eines „KI-Meta-Services“ identifiziert, eines Services der selbst KI-Services nutzt, um sie z. B. miteinander zu kombinieren. Beispiele für solche KI-Meta-Services sind föderiertes Lernen, Ensemble-Lernen und das Deployment von Services. Aus diesen Diskussionen ergaben sich initiale Planungen zur generischen, wiederverwendbaren Integration des föderierten Lernens in die IIP-Plattform.

Die Umsetzung beispielhafter Standardkomponenten fand in Arbeitsschritt 6.4.3 statt. Abweichend vom ursprünglichen Gedanken ist die Pseudofehlererkennung wie sie bei Sennheiser gebraucht wird nicht übertragbar auf andere Produktionsszenarien. Gründe hierfür sind geringe Stückzahl, hohe Variantenvielfalt, die so bei anderen Leiterplattenproduzenten nicht auftauchen. Der Plan, hieraus eine Standardkomponente abzuleiten, wurde letztlich aufgegeben (Slashwhy). Dennoch wurden Ideen aus IIP-Ecosphere (insbesondere Sennheiser-Demonstrator, KI-Accelerator und ThinkTank Plattform) bei Slashwhy übernommen und als modulares Baukastensystem in deren „Styx“-Ansatz integriert. Es handelt sich

dabei um eine Unterstützung der Arbeitsschritte, die im Rahmen des Entwicklungsprozesses einer KI-Lösung unabhängig von den genutzten KI-Methoden und dem Anwendungsfall benötigt werden. Styx basiert auf einem generischen Code-Grundgerüst in Form eines Projekttemplates in Python (unter Verwendung von Cookiecutter, Streamlit, scikit-learn, tensorflow, aced, pymongo, omniboard) und unterstützt insbesondere containerisiertes Training und Inferenz.

Für die weiteren identifizierten Möglichkeiten wurden diverse Komponenten entworfen und umgesetzt:

- Die automatisierte Generierung applikationsspezifischer Maschinen-Konnektoren, die im Gegensatz zu üblichen, händisch programmierten Konnektoren leichtere und schnellere Anpassbarkeit sowie Datensparsamkeit und bessere Performanz bei erhöhter Konsistenz erlauben.
- Als Grundlage für automatisiert generierte OPC UA Konnektoren (SSE@UHI, VDW), die automatisierte Übersetzung und Integration von OPC UA companion specifications in das Plattform-Konfigurationsmodell. Dies ermöglicht einfache (Wieder-)Verwendung komplexer OPC UA companion specifications bzw. dafür definierter Strukturen und damit die automatisierte Generierung zugehöriger OPC UA Konnektoren.
- Erste Komponenten für das Plattform-KI-Toolkit in Form wiederverwendbarer und anpassbarer Datenvorverarbeitungskomponenten, generierter Federated Learning Integration, RapidMiner RTSA Integration.
- Bereitstellung von Plattform-Schnittstellen zur Auflösung semantischer Identifizierer (semantic Identifier) in Verwaltungsschalen für beliebige Kataloge. Beispielhafte Umsetzung der Auflösung für ECLASS und „admin-shell.io“ als lokalen Mechanismen zur Auflösung ohne Internet-Verbindung sowie des ECLASS-Webservice zur Auflösung semantischer Identifier in Verwaltungsschalen (gemeinsam mit ECLASS).
- Initiale Entwicklung einer Geräte-Verwaltungsschale (diskutiert mit FORTISS in einem IDTA-Meeting) sowie Bereitstellung von Schnittstellen für gerätespezifische Laufzeit-Messwerte. Exemplarische Umsetzung gerätespezifischer Laufzeit-Messwerte für Phoenix Contact PLCnext (mit Phoenix Contact) und Bitmoteco (mit Bitmotec).

In Arbeitsschritt 6.4.4 erfolgte die Validierung und Evaluation der Umsetzungen aus Arbeitsschritt 6.4.3. Die beschriebenen Komponenten wurden erfolgreich gegen entsprechende Vorgaben getestet und die Tests in wiederholbare Regressionstests überführt. Die automatisierte Übersetzung und Anwendung der OPC UA companion specifications wurde erfolgreich gegen den OPC-UA Testserver "Sample-Server" der UMATI Community getestet (SSE@UHI, VDW). In diesem sind verschiedene Companion Specifications mit einer zugehörigen simulierten Testinstanz implementiert. Der Partner VDW leistete insbesondere Hilfe und Support bei der Umsetzung. Verschiedene Instanzen der generierten Konnektoren für MQTT, OPC UA, Verwaltungsschalen wurden in den IIP-Ecosphere Plattform-Demonstratoren verwendet, validiert und öffentlich demonstriert. Selbst größere Datenformate wie die des Drive Data Hubs (Lenze) konnten problemlos generiert und verwendet werden. Analog wurden automatisch integrierte applikationsspezifische Instanzen von Diensten des KI-Toolkits, insbesondere des föderierten Lernens validiert und erprobt. Die Ergebnisse wurden im Plattformhandbuch [EIC20, EIC21b, EIC22d, EIC23a, EIC23c] sowie in [EIC22a][EIC22b][EIC22c][EIC23b] diskutiert und veröffentlicht. Angedacht war auch eine technische Live-Integration des EMO'23 Demonstrators mit der UMATI-Community, was trotz aller

Bemühungen der beteiligten Partner letztlich an inkompatiblen OpenSource OPC UA Komponenten scheiterte.

Die Ansätze zu Standardkomponenten wurden in Arbeitsschritt 6.4.5 (Rückführung und Evaluation in Demonstratoren) angewendet. Insbesondere die automatisiert generierten Konnektoren waren hilfreich in den dynamischen Demonstratoren zur magnetischen Produktidentifikation (MIP) und dem dynamischen Demonstrator zu föderiertem Lernen von NovoAI. Beide Anwendungen basieren auf größeren Datenstrukturen, für die Konnektoren generiert und erfolgreich angewendet werden konnten. In diesen Anwendungen ergaben sich neue Aspekte, die erprobt und wiederum in die Plattform zurückgeführt werden können. Sowohl MIP als auch NovoAI (hier das AVA-Produkt) bieten spezielle Hardware an, die durch vorkonfigurierte Konnektoren einfacher integriert werden können. Das Plattform-Konfigurationsmodell wurde daher mit einer optionalen Erweiterung für beide Partner ergänzt. Insbesondere bei MIP ist es erforderlich, die Hardware durch geeignete Nachrichten explizit anzusteuern. Hieraus ergab sich die Notwendigkeit einer Standardbibliothek, die wiederum auf dem Konfigurationsmodell bzw. den darin definierten Typen für MIP basiert. Bisher wurden auf dieser Ebene nur Applikationen gebaut, jetzt kann die Realisierung einer Applikation durch Verwendung der Standardbibliothek vereinfacht werden. Zudem erforderte der dynamische Demonstrator von MIP eine Verhandlung des Ergebnisses mit der eigentlichen Hardware, d.h., eine KI in der Plattform schlägt eine Verbesserung der Identifikation vor, die von der Hardware geprüft und ggf. freigegeben wird. Ein derartiges Ping-Pong zwischen Hardware, Konnektor und KI-Service war bisher in keiner Applikation aufgeschienen und konnte hier erfolgreich validiert und auf der EMO 23 öffentlich demonstriert werden.

Die Ideen und Erkenntnisse aus den beschriebenen Arbeiten wurden in Arbeitsschritt 6.4.6 Standardisierungsaktivitäten zugeführt. Hierzu nahmen Vertreter des KI-Accelerator an folgenden Veranstaltungen aktiv teil:

- IDTA WG Submodels Meeting am 16. Februar 2022
- LNI 4.0 Klausurtagung am 30. Juni 2022
- AK STD 1941.0.8 Expertenrat für künstliche Intelligenz in industriellen Anwendungen

Weitere Diskussionen, z. B., mit ECLASS regte beispielsweise zu bislang fehlenden Katalogeinträgen an. Weiterhin wurden Ergebnisse des Plattform Usage View [STI21], der auf dem LNI 4.0 Edge Configuration Usage View basiert, in die Arbeiten des LNI 4.0 zurückgeführt (Lenze) und in eine überarbeitete Version des LNI 4.0 Dokuments integriert.

Zusätzlicher Arbeitsschritt 6.4.7: Proprietäre Maschinenprotokolle

Dieser Arbeitsschritt ist im ursprünglichen Arbeitsplan für IIP-Ecosphere nicht enthalten und wurde im Rahmen der Budgetübertragung zwischen KIPROTECT und UHi dem Gesamtarbeitsplan hinzugefügt. Während der Fokus des Think Tank „Plattformen“ (AP 2) im Wesentlichen auf Integration von Standard-Protokollen wie OPC UA, MQTT, AMQP oder Verwaltungsschalen liegt, gibt es natürlich diverse proprietäre Protokolle verschiedenster Hersteller, die damit nicht abgedeckt werden, deren Existenz aber dennoch berechtigt ist und deren Integration für die IIP-Ecosphere-Plattform relevant und interessant ist. Daher war das Ziel dieses Arbeitsschritts, die Auswahl für und die Integration proprietärer Maschinenprotokolle in die IIP-Ecosphere Plattform durchzuführen.

In Zusammenarbeit mit dem Think Tank „KI & Produktion“ (AP 4) wurde hier als aussichtsreiches Protokoll das TwinCAT ADS Protokoll von Beckhoff identifiziert, das Daten in Hochgeschwindigkeit (Sampling-Raten von 1kHz bis über 10kHz) bereitstellen kann. Eine beispielhafte Programmiersprachen-Integration in Python liegt in der Form des Open Source Projekts pyAds vor. Da die IIP-Ecosphere-Plattform im Konnektor-Bereich im Wesentlichen auf das (üblicherweise performantere) Java setzt, wurde hier zunächst eine beispielhafte Übertragung von pyAds nach Java unter Zuhilfenahme der proprietären Windows-ADS-Bibliothek vorgenommen. Eine Evaluierung dieser Java-Bibliothek gegen die TwinCat Simulations-Umgebung sowie gegen das Beckhoff-Edge-Device im Plattform-Demonstrator für die EMO 2023 ergab insbesondere bei lokalem Zugriff einen Geschwindigkeitsvorteil von Faktor 2.5-3 für die IIP-Ecosphere-Lösung [WEB23]. In einem zweiten Schritt wurde die ADS-Bibliothek für Java in die IIP-Ecosphere-Plattform und dort in einen externen Konnektor integriert. Hier ist insbesondere interessant, dass die Daten-Zugriffe bei Hochgeschwindigkeitsprotokollen weniger generisch realisiert werden als z. B. bei üblichen OPC UA Client-Implementierungen. Daher wurde die IIP-Ecosphere-Konnektor-Schnittstelle geeignet erweitert.

Semantische Datenintegration (UAP 6.5)

IIP-Ecosphere soll den einfachen Einsatz von KI-basierten-Methoden, wie z. B. für Predictive Maintenance, in einer „intelligenten“ Produktion bei mittelständischen Produktionsunternehmen ermöglichen. Da im mittelständischen Kontext oft die Ressourcen für die SW-Entwicklung und Daten-Integration fehlen, wurden in diesem Unterarbeitspakt folgende Resultate erarbeitet, um eine Integration von KI-Methoden mit geringstem Aufwand zu ermöglichen:

- Semantische Datenintegration (SDI),
- Integration der SDI in unterschiedliche Streaming-Plattformen,
- AI-Applikation, mit der über ChatGPT eine interaktiv und explorativ Datenanalyse ermöglicht wird.

Semantische Datenintegration

Das Ziel der Semantische Datenintegration (SDI) ist die Erweiterung existierender Streaming-Plattformen wie auch der IIP-Ecosphere Plattform mit der Fähigkeit heterogene Daten (z. B. Sensordaten, Logs und Engineering-Daten einer Maschine) „semantisch“ zu annotieren und diese über eine standardisierte Abfragesprache (z. B. SPARQL) dynamisch zugänglich zu machen. Die SDI sowie die Abfragesprache sollen in einer föderierten Umgebung einsetzbar sein, falls mehrere Komponenten (z. B. Maschinen) in einem Produktionsprozess zusammengeschlossen sind. Folgende detaillierte Ergebnisse im Kontext der SDI wurden erarbeitet:

- Analyse und Konzeption der technischen Anforderungen, welche für die SDI in IIP-Ecosphere erforderlich sind. Hierfür wurde eine Literaturanalyse zu den Themen Industrie 4.0 Referenzarchitekturen, Asset Administration Shell, Semantic Web Technologien und Knowledge Graphs erstellt. In der Konzeption ist eine Architektur auf Basis von einem virtuellen Knowledge Graph (VKG) skizziert worden, wobei der VKG in unterschiedlicher Form, z. B., über MQTT-Konnektoren in die IIP-Plattform eingebunden werden kann. Der technische Report „Semantic Data Integration in Industry 4.0“ beschreibt diese Ergebnisse und ist online verfügbar [FIS21].

- Als Integrations-Schema der SDI und des VKG wurde eine Ontologie auf Basis der standardisierten SOSA-Ontologie modelliert und mit Versuchsaufbau-spezifischen Klassen erweitert. Die Sensor- und Log-Daten des Aufbaus wurden dynamisch über Mappings mit der Ontologie verbunden und bilden als Gesamtheit den VKG. Komplexe Abfragen über den VKG können dann über vordefinierte oder benutzerspezifizierte SPARQL-Queries erstellt werden, wo z. B. Eventmuster formuliert werden können. Der VKG wurde mit der Plattform Ontop realisiert und wurde mit Versuchsaufbau-spezifischen Daten getestet.
- Um die Verarbeitung in einer föderierten Umgebung einzuordnen, wurden zwei Ansätze zur Datenföderation analysiert und umgesetzt. Der erste Ansatz, genannt „Query-level Federation“, basiert auf der Idee, dass einzelne VKGs über unterschiedliche Query-Endpoints zur Verfügung gestellt werden, welche dynamisch während der Query-Evaluierung abgefragt werden. Der zweite Ansatz, genannt „Infrastructure-level Federation“, integriert die Datenquelle über eine „Distributed Database Engine“, wie z. B. TEIID (<https://teiid.io/>), wobei die Mappings ausschließlich auf der Distributed Database Engine abgebildet werden. Für beide Ansätze wurden die Stärken und Schwächen analysiert und Guidelines für die Wahl des Ansatzes erarbeitet. Die Resultate wurden im Whitepaper [FIS22] erläutert und unseren Partnern bei Siemens Digital Industries vorgestellt.
- Da je nach Unternehmen unterschiedliche IOT-Plattformen und Protokolle vorhanden sind, wurde eine Integration nebst der IIP-Plattform selbst, auch eine Integration über StreamPipes evaluiert. Zusätzlich wurde eine ressourcenschonende Integration direkt über einen MQTT-Broker erarbeitet. Für die Integration mit der IIP-Plattform wurde ein exemplarischer Java-basierter Producer im Framework entwickelt, aber die MQTT-Broker-basierte Integration für den Versuchsaufbau vorgezogen. Über vordefinierte SPARQL-Queries konnte gezeigt werden, wie Sensor-Daten aggregiert und spezifische Eventmuster erkannt werden. Die Resultate sind im Erfolgsbericht detailliert beschrieben und online verfügbar.

AI-Applikation: Integration von ChatGPT & Streaming

Um den einfachen Einsatz von KI-basierten-Methoden mit einer „low-code“ Integration zu erreichen, verbinden wir die Semantischen Datenintegration mit einem Frontend das auf einem Chatbot aufbaut, welcher in unserem Fall ChatGPT ist. Dieser dient dem End-User als Frontend und erlaubt, die Datenanalyse interaktiv und explorativ zu gestalten, ohne selbst Programmierkenntnisse zu besitzen.

Der Einsatz von ChatGPT im industriellen Kontext wirft neue Herausforderungen auf, da die Technologie kein implizites Wissen über industrielle Daten besitzt. Zum Beispiel kann ChatGPT keine dynamischen, maschinenspezifischen Logs oder Sensordaten lesen und verarbeiten. Um dennoch einen umfangreichen Zugriff auf diese Daten zu ermöglichen, können diese über ein externes Interface (auch Plugin genannt), auf Abruf zur Verfügung gestellt werden.

Im Rahmen dieses Unterarbeitspakets betrachteten wir, wie weit ChatGPT für die dynamische Analyse von industriellen Daten verwendet werden kann, die in heterogener Form und mit einer „undefinierten“ Semantik auftreten. Die erste Erkenntnis war, dass eine direkte Einbindung der unterschiedlichen Datenströme schwierig ist, da Plugins zu unterschiedlichen Protokollen entwickelt werden müssten. Hier kommt die Semantische Datenintegration (SDI) ins Spiel, welche eine klare Beschreibung der Daten über eine Ontologie, und die standardisierte Abfragesprache SPARQL zur Verfügung stellt, wobei GPT-4

in der Lage ist gültige SPARQL-Abfragen zur generieren. Um die SDI mit ChatGPT zu verbinden, entwickelten wir für den Demonstrator ein KG-Plugin, welches ChatGPT folgende Informationen zur Verfügung stellen kann:

- Schema-API, welche das Vokabular, zum Beispiel Klassen, Properties, Instanzen, im Kontext der industriellen Daten erzeugt. Hierdurch wird eine systematische Beschreibung der verwendeten Assets (die Ontologie) in Form von Relationen zur Verfügung gestellt.
- SPARQL-API, welches die von ChatGPT generierten SPARQL-Abfrage auf Basis von User-Interaktion an einen vordefinierten KGs weiterreicht, und die Resultate als Relationen retourniert.

Nach erfolgreichen Tests mit dem KG-Plugin, integrierten wir das ChatGPT-Frontend mit unseren Versuchsaufbau, indem wir über die Schema-API die SOSA-Ontologie ChatGPT bekannt gaben und die SPARQL-API den Zugriff auf den KG unseres Demonstrators erlaubten. Hierdurch konnte gezeigt werden, dass Benutzer explorativ mit der Maschine kommunizieren können, und einfache statistisch Analysen mit Hilfe eines Wolfram-Alpha-Plugin durchführbar sind. Wir arbeiten derzeit an einer Publikation, um unseren Ansatz und den Demonstrator zu beschreiben. Zusätzlich wurde eine Kurzfassung der Publikation im Erfolgsbericht beigefügt.

Versuchsaufbau Siemens

Ein Versuchsaufbau wurde entwickelt, um die Fähigkeiten der semantischen Datenintegration und der Integration mit ChatGPT zu demonstrieren. Die Basis hierfür bildet ein 3D-Drucker (siehe Abbildung 36), welcher über diverse Temperatur- und Positions-Sensoren verfügt. Ergänzt wurde dieser um weitere Komponenten, wie einer industriellen Wage für die Messung des Verbrauchs an Material, sowie eines Moduls zur Messung des Stromverbrauchs. Ziel des Versuchsaufbaus ist es, diesen möglichst realistisch an Szenarien innerhalb der Industrie zu gestalten. Durch die Verwendung von heterogenen Datenquellen, Protokollen und Standards wurden reale Herausforderungen der Produktion reproduziert.

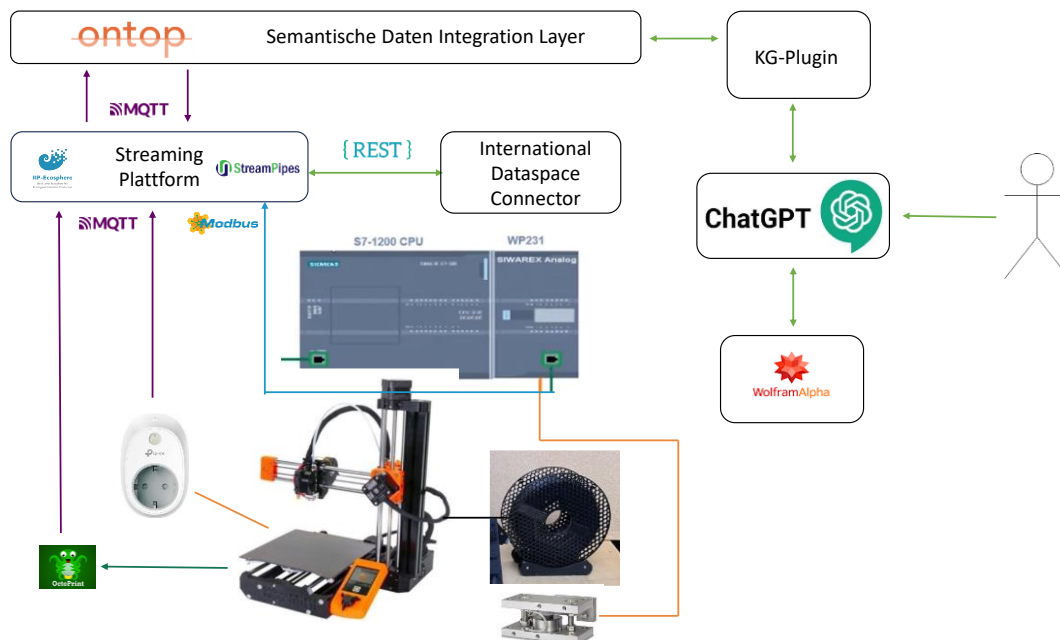


Abbildung 36: Detaillierter Versuchsaufbau (Siemens)

Der Schwerpunkt liegt auf der Demonstration einer flexiblen Lösungsstrategie, welche statische und dynamische Daten aus einem Fertigungssystem in einen KG integriert, und Datenanalyse-Methoden dabei helfen, diese Daten verständlich und nachvollziehbar zu analysieren. Eine solche Form der Datenintegration und Datenanalyse reduziert die Einstiegshürde und den Umgang mit KI-Lösungen im Bereich industrielle Automatisierung und in industriellen Produktionsanlagen. Abschließend wurde die SDI exemplarisch im Siemens Versuchsaufbau (3D Drucker) umgesetzt.

Im Rahmen von Arbeitsschritt 6.5.2 wurde in mehreren Iterationen auch die Ergebnisse eine gemeinsame Konzeption sowie eine Absprache von Schnittstellen, Datenformaten und Transportprotokollen, insbesondere von Verwaltungsschalen, auch für den KI-Lösungskatalog. Hierzu gehörten gemeinsam mit dem Partner Lenze Ansätze zur semantischen Annotation von Informationen in Verwaltungsschalen, z. B., zur Datenherkunft. Die daraus ableitbaren Anforderungen für die Plattform wurden im Rahmen von UAP 2.2 im Rahmen der Plattform-Anforderungssammlung erfasst und gemeinsam als Whitepaper veröffentlicht. Anhand der in UAP 2.2 entwickelten Konzepte und der in UAP 2.3 entwickelten Plattformarchitektur wurden in Arbeitsschritt 6.5.3 Beiträge zur IIP-Ecosphere Plattform diskutiert. Besprochen wurde eine generische SPARQL-Integration über mehrere Datenbanken unterschiedlichen Typs, z. B. Zeitreihen oder NoSQL sowie die Integration eines bereits evaluierten, experimentellen GAIA-X-Konnektors. Als Grundlage wurde hierfür ein gemeinsames, virtuelles Hands-on-Meeting zur Verwendung der IIP-Ecosphere Plattform mit anschließenden Detaildiskussionen durchgeführt. Zudem nahm ein Vertreter von Siemens am Plattform-Integrationsworkshop in Hildesheim teil.

Plattform & Sicherheit (UAP 6.6)

Die Kern-Partner dieses UAPs, KIProtect und SSE@UHi führten in mehreren Iterationen zunächst Diskussionen durch, um gemeinsam Kandidaten für Komponenten zur Anonymisierung, Pseudonymisierung und transparenten Verschlüsselung von Daten in der intelligenten Produktion zu identifizieren (Arbeitsschritte 6.6.1 und 6.6.2). Initiale Ergebnisse hieraus, insbesondere zur Anonymisierung und Pseudonymisierung wurden im Rahmen von (Arbeitsschritt 6.6.1) gemeinsam erfasst und in die Sammlung

von Plattformanforderungen (siehe UAP 2.2) eingebracht, integriert und in Form der funktionalen Plattformanforderungen als Whitepaper dokumentiert.

Als relevante Technologiekomponenten wurden Kodex [1], die Open-Core Plattform für Privacy- & Security-Engineering sowie Hyper [2] als sichere, Ende-zu-Ende verschlüsselte Kommunikation identifiziert (Arbeitsschritt 6.6.1). Für eine Integration in die IIP-Ecosphere-Plattform wurden im Rahmen von Arbeitsschritt 6.6.3 zusammen mit UAP 2.3 die Integration der Komponenten in die Plattformarchitektur diskutiert, beschlossen und im Rahmen des IIP-Ecosphere Plattformhandbuchs dokumentiert. Kodex [1] soll dabei als generischer, anpassbarer Plattformdienst integriert werden, Hyper hingegen als optionale, transparente Ergänzung der Transportschicht.

KIProtect hat im Rahmen des Projekts Kodex [1] gemäß den gemeinsamen Beschlüssen angepasst und weiterentwickelt. Kodex erlaubt es, Daten effektiv durch verschiedene Methoden wie Pseudonymisierung, De-Identifikation und Verschlüsselung bei der Verarbeitung zu schützen. Der Fokus wurde hierbei auf Pseudonymisierung gelegt, da eine vollständige Anonymisierung der komplexen und oft hochdimensionalen Daten wie sie bei Industrieanwendungen anfallen oft nicht möglich ist. Weiterhin wurde Kodex im Rahmen des Projektes (Arbeitsschritt 6.6.5) technologisch weiterentwickelt und überarbeitet, mehrere neue Ansätze zur Pseudonymisierung (Arbeitsschritt 6.6.6) von Daten wurden im Rahmen des Projekts erprobt und umgesetzt, so z. B. ein Verfahren basierend auf der Stromverschlüsselung von Daten, welches die teilweise Erhaltung struktureller Eigenschaften der Ursprungsdaten bei der Pseudonymisierung erlaubt und so insbesondere für komplexe numerische Daten relevant ist (Arbeitsschritte 6.6.8 und 6.6.10) [Dewes 2022].

Letztlich wurde Kodex gemeinsam mit SSE@UHi an die IIP-Ecosphere Plattform integriert (Arbeitsschritt 6.6.4). Technisch geschah dies als modell-basierte Integration von Kodex in die IIP-Ecosphere Plattform als Blueprint für Dienste, die als lokale Betriebssystemprozesse vorliegen und Daten mittels textueller Ströme (JSON) austauschen. Die initiale Integration erfolgte über Kommandozeilenströme. Im Rahmen einer studentischen Projektarbeit wurde hier eine alternative Integration von KODEX via REST realisiert sowie die Performanz beider Integrationen analysiert und verglichen (Arbeitsschritt 6.6.8). Auch wenn die REST-Variante zusätzliche Prozessressourcen erfordert, ist die Performanz sowohl auf Windows als auch auf Linux-Systemen besser.

Weiterhin hat KIProtect im Rahmen des Projekts eine Plattform für die sichere, Ende-zu-Ende verschlüsselte Kommunikation entwickelt, Hyper [2], die genutzt werden kann um mehrere Instanzen der IIP-Ecosphere Software über Unternehmensgrenzen hinweg zu verbinden und für den Datenaustausch einen hohen Sicherheitsstandard zu gewährleisten. Hyper nutzt hierfür moderne Technologien wie das QUIC-Internetprotokoll und kombiniert diese mit einem innovativen, dezentralen Vertrauensansatz der auf modernen kryptographischen Methoden und einem Privacy-/Security-By-Design Ansatz basiert. Insgesamt wurde Hyper im Rahmen der gemeinsamen Planung eines Sicherheitsschilds sowie eines sicheren Device-Onboarding-Verfahrens für die IIP-Ecosphere Plattform geplant und erste praktische Schritte durchgeführt (Arbeitsschritt 6.6.7). Arbeitsschritt 6.6.7 war im Rahmen einer Budgetübertragung zwischen KIPROTECT und UHi in Richtung Plattform neu geplant. Ziel war es, mit Hyper als Plattformkomponente ein transparent-verschlüsselndes Overlay-Netzwerk aufzubauen, dass Geräten sicheren und authentifizierten Zugang zur Plattformkommunikation und zur Plattform-Verwaltungsschale bietet.

Initiale Versionen von Hyper zeigten bereits dessen Potential und die Möglichkeiten aber auch die Stärkung der IT-Sicherheit, die die IIP-Ecosphere-Plattform dadurch erhalten würde. Hierzu wurden verschiedene Online-Besprechungen zur konzeptionellen und praktischen Integration durchgeführt. Zudem wurden grundlegende Mechanismen zur Integration eines externen Identitätsmanagements (über den IdentityStore der Plattform) vorbereitet und mit Hyper sowie anderen Technologien erprobt. Die entsprechende alternative Komponente wurde Teil der Plattform auf github. Gegen Projektende war eine technische Integration mit der IIP-Ecosphere Plattform greifbar nahe und erste Integrationsexperimente wurden durchgeführt. Leider lag Hyper erst gegen Projektende und nach gemeinsamen Anpassungen und Fehlerbereinigungen in einem stabilen, integrationsfähigen Zustand vor. Durch den klaren Fokus auf die gemeinsame Präsentation der Partnerergebnisse auf der EMO'23 Messe und die hierfür notwendigen, aufwändigen Integrationsarbeiten sowie durch überraschende Krankenstände, konnte letztlich eine Integration von Hyper in die IIP-Ecosphere-Plattform nicht voll-ständig umgesetzt werden. Auch wenn der KODEX-Service für alle Plattform-Demonstratoren bereits früh verfügbar war (auch für Arbeitsschritt 6.6.10), war eine Integration in den EMO'23 Demonstrator wie ursprünglich im Rahmen von Arbeitsschritt 6.6.9 geplant nicht mehr möglich. Da ein hohes Potenzial bei der Nutzung von Hyper in der IIP-Ecosphere-Plattform besteht, wurde der aktuelle Stand detail-reich dokumentiert und es besteht Bereitschaft der involvierten Partner, das Thema in der Zukunft weiterzuverfolgen. Die Sicherheitsauditierung aus Arbeitsschritt 6.6.11 wurde im Rahmen der beantragten Budgetübertragung ausgesetzt und letztlich nicht durchgeführt.

KIProtect arbeitet weiter aktiv an der Vermarktung beider Technologien und hat für Kodex bereits einen ersten großen Pilotkunden gewinnen können. Die im Rahmen des IIP-Ecosphere Projekts entwickelten Technologien und Ansätze konnten auch in anderen Bereichen in die Praxis eingebracht werden. SSE@UHi hat im Umfeld der Verwertung der IIP-Ecosphere Plattform weiterhin starkes Interesse an einer Integration von Hyper und plant die Arbeiten nach Projektende und nach weiterer Konsolidierung der Plattform fortzuführen.

2.1.7 Demonstrator „Werkübergreifende Optimierung der Taktzeiten“ (AP 7)

Dieser IIP-Demonstrator ist bei Volkswagen Nutzfahrzeuge angesiedelt. In der industriellen Großserienfertigung sind Störungen mit hohen Kosten verbunden. Um Stillständen entgegenzuwirken, konzentrierte sich die Forschung und Entwicklung bisher verstärkt auf die Vermeidung von Ausfallzeiten, zum Beispiel durch vorausschauende Wartung der Anlagen und den Einsatz alternativer Routen durch den Produktionsprozess. Diese Ansätze vernachlässigen jedoch Verzögerungen (Mikrostörungen) in einzelnen Prozessschritten, die über die gesamte Prozesskette hinweg propagiert und dabei oft weiter verstärkt werden. Auf diese Weise können bereits minimale, vom Mitarbeiter kaum wahrnehmbare, Verzögerungen in hohen Zusatzkosten resultieren.

Taktzeitenüberschreitungen in Produktionslinien sind häufig auf technische Unzulänglichkeiten oder Fehlfunktionen zurückzuführen, wie beispielsweise Fehler an einem Roboter innerhalb einer bestimmten Station. Ein solcher Fehler kann nicht nur direkte Verzögerungen verursachen, sondern auch eine Kaskade an Folgefehlern auslösen. Diese wiederum können sich kumulativ auf die gesamte Produktionslinie auswirken, wodurch die ursprünglich geplante Taktzeit erheblich negativ beeinflusst wird (siehe Abbildung 37). Die Identifikation solcher Fehlerquellen bietet ein hohes Potential die Produktivität zu

steigern. Ziel des Demonstrators mit Volkswagen Nutzfahrzeuge war es daher, ein intelligentes, Fertigungsanalysesystem zu entwickeln, mit dessen Hilfe Verzögerungen und Störungen frühzeitig und automatisiert erkannt und visualisiert werden. Dabei wird im Demonstrator konkret die Karosseriefertigung bei Volkswagen Nutzfahrzeuge im Werk Hannover betrachtet.

Anforderungen und Problemanalyse

In der Anlaufphase wurden zunächst die Karosserieproduktion des Volkswagen T6 und T7 als Anwendungsfälle analysiert. Beide Modelle werden dynamisch mit in mehreren Varianten produziert. Die Sequenz der produzierenden Varianten wird dabei durch den Bestellzeitpunkt festgelegt.

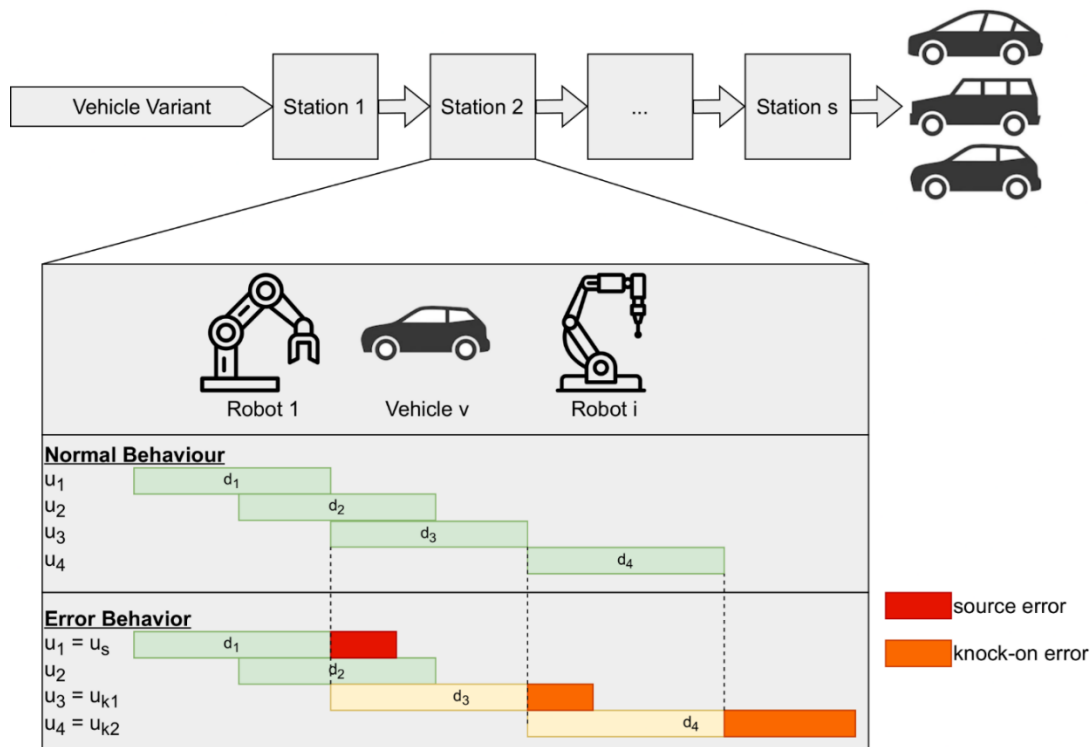


Abbildung 37: Folgefehler und Taktzeitverschleppung in der Linienfertigung

Als geeigneter KI-Anwendungsfall für die werkübergreifende Optimierung von Taktzeiten wurde in UAP 7.1 die vollautomatisierte Montage der Türen und Heckklappen der Fahrzeuge identifiziert. Hier werden Ursache-Fehler-Wirkungsketten und Auswirkungen auf die Taktzeit bereits in manuellen Analysen untersucht, sodass eine gute Wissensgrundlage für die spätere Validierung von KI-Ansätzen vorhanden ist.

Ein erster Datensatz aus der Produktion wurde von VWN zur Exploration der Daten bereitgestellt. Die Datenbasis wurde durch dekodiert, die hinterlegten Daten durch eigens entwickelte Skripte mit Hilfe von Python inkl. Pandas Bibliothek analysiert, ihre Eignung zur Umsetzung des Anwendungsfalls bestätigt und die Erkenntnisse regelmäßig mit VWN diskutiert. In diesem Rahmen wurde die Arbeitshypothese aufgestellt, dass zwischen Ursache- und Wirkungsfehler unterschieden werden muss. Letztere stellen Taktzeitverschleppungen durch erstere dar. Als Grobkonzept für diesen Demonstrator wurde daher festgelegt, dass eine KI-Anwendung implementiert werden soll, die statistisch signifikante Häufungen von

Fehlern erkennt, im nächsten Schritt Ursache- und Wirkungsfehler ohne umfassende manuelle Analysen unterscheidet und in der Folge automatisiert zeitnahe Warnungen bei Auftreten von Ursachefehlern über eine geeignete Visualisierung ausgibt.

Die Anforderungsanalyse zur Bewältigung des Problems der Taktzeitenüberschreitung hat zudem ergeben, dass eine zusätzliche systematische Datenerhebung und -analyse erforderlich ist. Insbesondere sollen Daten aus verschiedenen Quellen wie VASS, Event und letztendlich auch Prozessdaten gesammelt werden. Die Zusammenführung dieser Daten bietet die Möglichkeit, sie zum Training von KI-Modellen zu verwenden. Ziel ist es, durch die Kombination dieser Daten eine Prädiktion von Folgefehlern nach einem initialen Ursprungsfehler zu ermöglichen. Das KI-Modell soll in der Lage sein, Muster im Zeitverhalten zu erkennen und Vorhersagen darüber zu treffen, wie sich bestimmte Fehler auf die gesamte Produktionslinie auswirken könnten. Dadurch wird nicht nur eine reaktive Fehlerbehebung ermöglicht, sondern auch proaktive Maßnahmen sollen ergriffen werden, um Ausfallzeiten und Ineffizienzen zu minimieren.

Zusätzliche Datenerfassung und -visualisierung

Gemäß den identifizierten Anforderungen wurde eine zusätzliche systematische Datenerfassung und zugehörige Datenvisualisierung umgesetzt. Abbildung 38 zeigt einen Überblick über die erfassten Datenquellen und typische Datenvolumina. Die Erweiterung der Datenbasis durch zusätzliche Events- bzw. Maschinendatenerfassung wurde im System robmon (Erweiterte Prozessgerätediagnose) implementiert. (UAP 7.1. bis 7.4.)

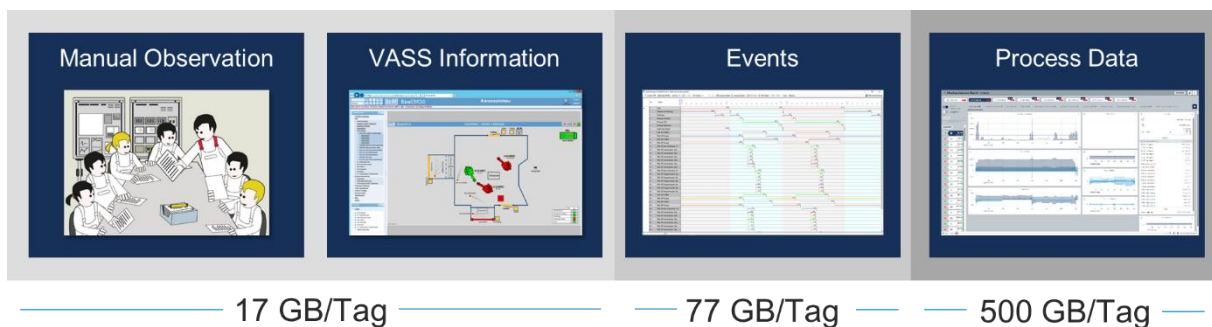


Abbildung 38: Erfasste Datenquellen und Volumina im Werk Hannover

Abbildung 39 zeigt eine Übersicht über das Applikationskonzept für Datenerfassung und Visualisierung. Die Hauptaufgabe des robmon Ecosystems besteht darin, die Prozessdaten der Roboter und Prozessgeräte einzusammeln und persistent zu sichern.

Applikationskonzept

Robot & PLC - Die SPS mit den zugehörigen Robotern und Prozessgeräten. Die SPS hat Bausteine eingebunden die die Daten über TCP/2007 bereitstellen.

stream-grabber - Der Prozess der den Bytestrom der SPS entgegen nimmt verarbeitet und in den stream-storage schreibt.

stream-manager-rpc - Das Bindeglied aller Prozesse im EcoSystem von robmon ist dafür zuständig Steuerbefehle an die richtigen Teilnehmer zu leiten.

stream-manager-ui - Die Benutzeroberfläche für die System Betreiber um das EcoSystem zu managen.

stream-manager-db - Die Datenhaltung der Metriken, Systemkonfigurationen und Prozessbestandteilen.

stream-storage - Die Datenhaltung der Prozessdaten

data-api - HTTP Datenendpunkte für Konsumierende Applikationen wie ZAU, CTSS oder das robmon-dashboard selbst.

dashboard - Benutzeroberfläche zur Auswertung der Prozessdaten

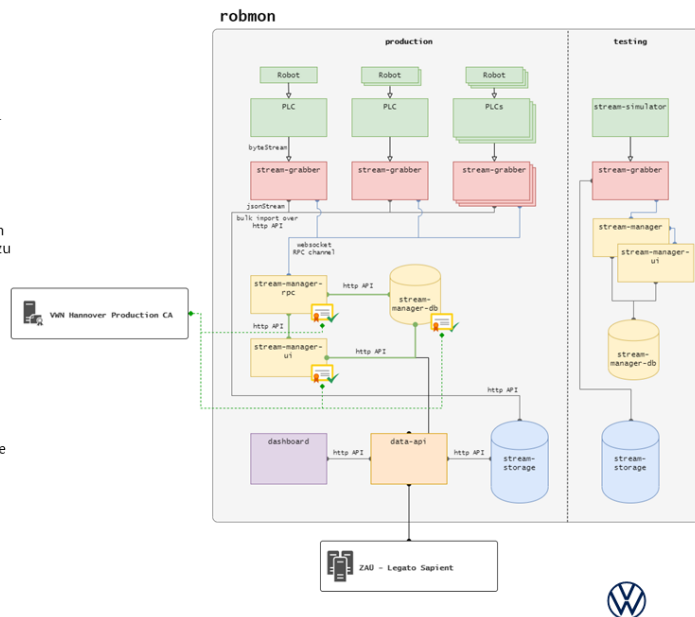


Abbildung 39: Übersicht über das Applikationskonzept zur Datensammlung

Die SPS sendet nur die reinen Rohdaten im 50 ms Takt an den `stream-grabber` Prozess. Die Aufgabe des `stream-grabbers` ist, diese Daten schnellstmöglich entgegenzunehmen und in ein menschenlesbares Format zu parsen. Die Daten entsprechen einem proprietären Protokoll welches auf Seiten der SPS und auf Seiten der `stream-grabber` eingehalten werden muss. Das Produkt ist ein mit Kopfdaten angereichertes JSON Format welches in einem kleinen Puffer im `stream-grabber` gehalten wird bis die maximale Füllgrenze des Puffers erreicht ist. Ist dies der Fall, wird der Block persistent in den `stream-storage` geschrieben. Um die ganzen `stream-grabber` zu kontrollieren und zu überwachen wurde ein `stream-managereingeführt`.

Die produktive robmon Infrastruktur wurde flächendeckend im Karosseriebau T7 und ID BUZZ installiert. Hierfür wurde eine virtuelle Server Infrastruktur auf der Grundlage des robmon Applikationskonzept aufgebaut. Insgesamt wurden in diesem Zusammenhang 8 virtuelle Maschinen installiert. Zudem wurden zentrale Komponenten aufgesetzt. Hierzu zählen zwei Stream-Grabber Gateways (Karosseriebau-T7, Karosseriebau-ID) zur Datenerfassung, eine erweiterbare Time-Series-Database mit 25 TB Speicher, welche für 42 Tage Datenhaltung der anfallenden Zeitreihendaten ausgelegt ist. Zur Umsetzung der robmon Anwendung wurden zentrale robmon Applikationsbestandteile installiert: Manager-RPC, Manager-UI, Manager-DB, Data-API und das Dashboard.

Seitens der Anlagen wurden die SPS Bausteine eingespielt, parametriert und im System durch Anpassung der Gateway IP-Adresse eingebunden. Zudem wurden alle Robote für das Weiterleiten der Prozessdaten angepasst (EA-Bereich) und das System wurde für die KI-Integration vorbereitet z. B. durch das Anlegen eines Modell-Repositories.

Um die gespeicherten Daten für Analyse- oder Visualisierungszwecke zu erfassen wurde eine `data-api` umgesetzt, welche die Daten möglichst einfach über eine REST-API zur Verfügung stellt. Diese API bildet auch die Grundlage um die Daten für den Endanwender möglichst übersichtlich darzustellen, um die Erkenntnisse über die gewonnenen Daten zu erlangen.

Dashboards zu Datenanalyse

Zur Visualisierung und Analyse der Daten wurde ein robmon-Dashboard Konzept umgesetzt.



Abbildung 40: Screenshot des robmon-dashboards

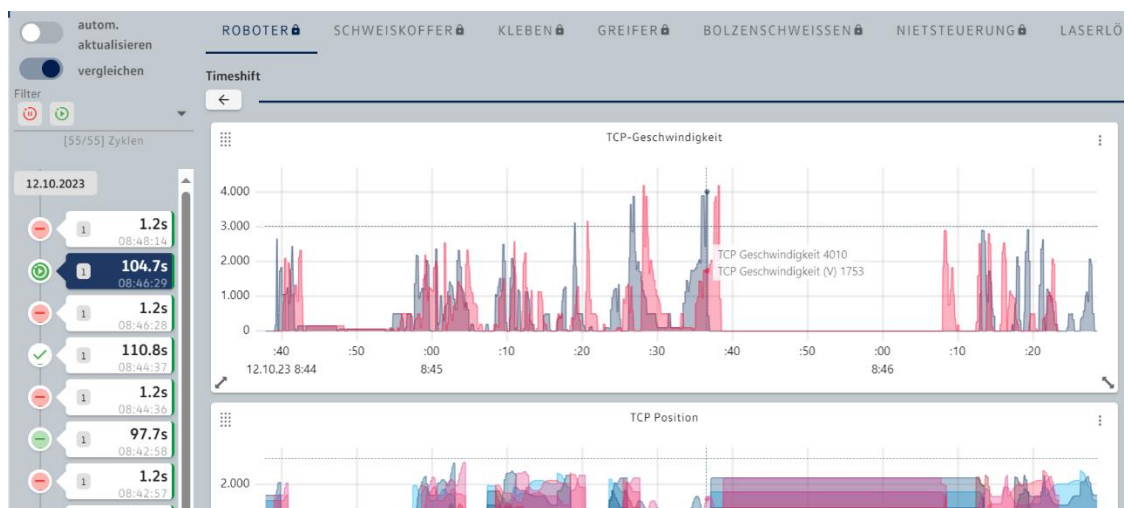


Abbildung 41: Analyse von Prozessen im robmon

In dem robmon-dashbaord (Abbildung 40) ist es dem Endanwender möglich sich sein eigenes Dashboard zusammenzustellen. Hierbei hat dieser die Möglichkeit erfasste Datenpunkte insbesondere Zeitreihen in ein Diagramm-Widget zu laden. Es beinhaltet auch Funktionen um zwei gleichartige Prozesse miteinander zu vergleichen um ggf. die Unterschiede einfach zu erkennen (vgl. Abbildung 41). Dies kann helfen den Prozess zu bewerten, Verzögerungen zu erkennen und mögliche Gründe für Verzögerungen schneller zu identifizieren.

Zudem kann der Endanwender mithilfe einer Time Shift Funktion (Abbildung 42) die zu vergleichenden Prozesse zu einem gewählten Ankerpunkt zeitlich synchronisieren. Diese Funktion unterstützt den Nutzer dabei den genauen Zeitpunkt im Prozesse zu finden, an der die Problematik potentiell auftritt.



Abbildung 42: Analyse in Prozessen in robmon

Das System ist auch bereits für die Einbindung von KI-Modellen vorbereitet. So können z. B. die Ausgaben, die durch ein KI-Model generiert werden, über die Datenschnittstelle ins System zurückgespielt werden. Damit lassen sich die KI-Ergebnisse wiederum einheitlich ebenfalls im robmon-dashboard visualisieren.

Klassifikation und -prädiktion von Taktzeitfehlern

Für den verbesserten Umgang mit Taktzeitfehlern wurden KI-Modelle erprobt und entwickelt. Der hierfür gewählte Ansatz zur KI-basierten Prädiktion von Taktzeiten ist in Abbildung 43 zusammengefasst. Eine zentrale Herausforderung für die Entwicklung der KI-Ansätze besteht darin, dass die Ursache für Störungen á priori nicht bekannt ist und damit das Labeling der Daten erschwert ist. Es ist zum einen erforderlich zwischen ursprünglichen Fehlern und Folgefehlern zu unterscheiden. Zum anderen ist es auch wichtig sich auf relevante Fehler zu konzentrieren. Daher wurden neben Neuronalen Netz Architektur ein regelbasiertes Verfahren für die automatische Dateninterpretation entwickelt. Dies ist notwendig um Prädiktionen von Fehlern durchzuführen, da diese ausreichend große Trainingsdatensätze benötigen, um Ursache-Wirkungszusammenhänge zu erkennen.

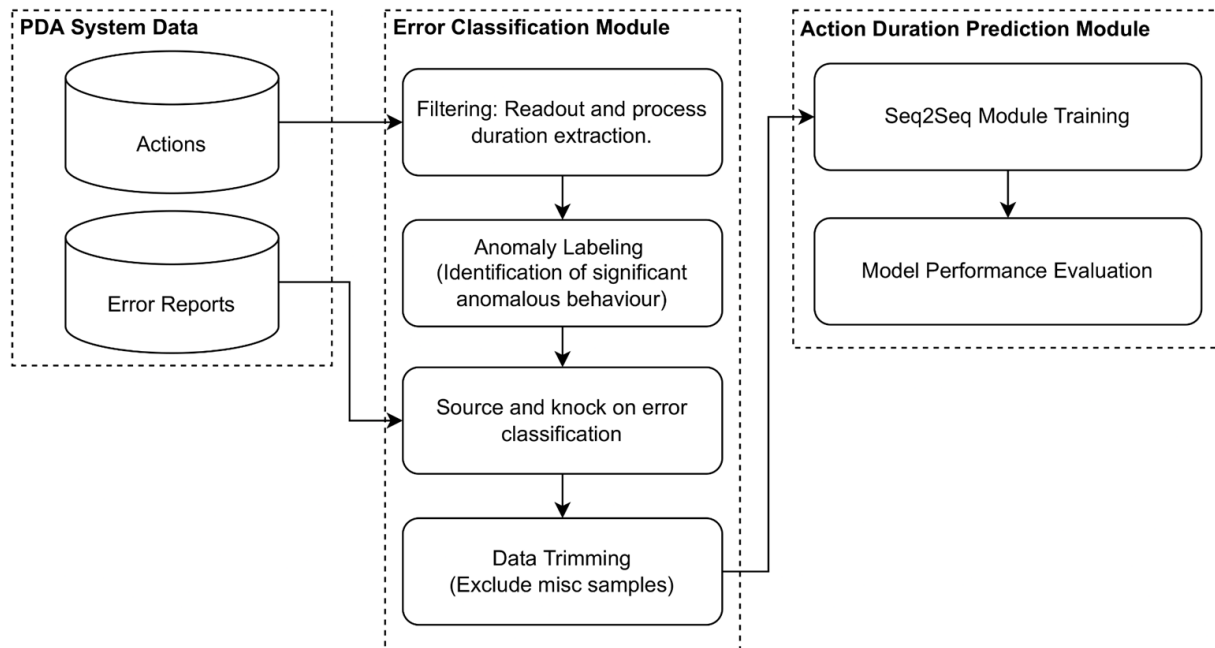


Abbildung 43: Ansatz zur Prädiktion von Taktzeiten

Auf der Grundlage dieser Trainingsdaten wurden KI-Modelle trainiert um 1) Fehler in Ursprungs- und Folgefehler zu klassifizieren (Source and knock on classification) und 2) Prädiktionen für die Dauer von Aktion (Prädiktion von Taktzeitfehlern, Seq2seq module training) durchzuführen. Für diese Aufgabe wurde mehrere KI-Ansätze ausgewählt, trainiert und in der Leistung miteinander verglichen, um das optimale Modell zu identifizieren. Zu diesem Zweck wurden neben Long-Short-Term-Memory (LSTM)- und TadGAN-basierte auch Gated-Recurrent-Unit- sowie Transformer-basierte Architekturen herangezogen.

Die Entscheidung für die KI-Architekturen LSTM, GRU und Transformer basierte auf ihrer fortgeschrittenen Fähigkeit, Sequenzdaten zu verarbeiten. Diese Modelle repräsentieren den aktuellen Stand der Technik in der Zeitreihenanalyse und haben sich in zahlreichen Anwendungen als besonders effektiv erwiesen. Das LSTM (Long Short-Term Memory) und das GRU (Gated Recurrent Unit) sind Varianten von rekurrenten neuronalen Netzen, die speziell entwickelt wurden, um langfristige Abhängigkeiten in Daten zu erkennen und zu modellieren. Die Transformer-Architekturen hingegen, ursprünglich für maschinelles Übersetzen konzipiert, hat durch seine self-attention Mechanismen die Fähigkeit, komplexe Sequenz-zu-Sequenz-Beziehungen effizient abzubilden.

Die Leistung der unterschiedlichen Ansätze wurde unter verschiedenen Versuchsbedingungen und Parametereinstellungen miteinander verglichen. Dabei zeigte sich zum einen – als sehr praxisrelevantes Ergebnis, dass mehr als 70 % der Sequenzen Taktzeitfehler (Ursprungs oder Folgefehler) enthalten. Zum anderen zeigte sich auch, dass die gewählten Ansätze gut geeignet sind Fehler vorherzusagen bei nicht zu weit in die Zukunft reichenden Vorhersagen. Hier sind insbesondere die Transformer-basierten Architekturen gut geeignet. Die Ergebnisse dieser Experimente wurden in einer wissenschaftlichen Publikation zusammengefasst und veröffentlicht [REI22].

Lessons learned und zukünftige Arbeiten

Im Rahmen des Demonstrators wurde über die bereits beschriebenen, sehr zeitaufwändigen Formen der Datenvorverarbeitung hinaus auch versucht, Prozessdaten, VASS und Eventdaten zu integrieren und in ein einheitliches Zeitreihenformat zu überführen, um eine noch reichere Datenbasis zu erhalten. Dabei stieß man jedoch trotz erfolgreicher Konvertierung auf erhebliche technische und logische Herausforderungen, insbesondere aufgrund der enormen Datenmenge und fehlender Informationen für eine geeignete Synchronisation.

Für den praktischen Einsatz der entwickelten KI-Modelle wurde zudem festgestellt, dass die Vorhersagen, welche mit den Daten trainierten KI-Modelle liefern, trotz guter Performanz nicht weit genug in die Zukunft reichen. Hier macht sich die bei prädiktiven Methoden oft zu beobachtende Tendenz, dass Prädiktionsstabilität und -güte mit der Länge des Vorhersagemodells rasch abnehmen, negativ bemerkbar. Um diese Modelle zum Einsatz zu bringen, müssen die KI-Modelle weiter verbessert werden, was beispielsweise durch die Kombination unterschiedlicher KI-Modelle und Prädiktionsmethoden gelingen dürfte. Weitere Arbeiten sind ebenfalls notwendig, um die Skalierbarkeit der exemplarischen Lösung zu erhöhen, damit sie effektiv auf die gesamten Produktionslinien oder sogar die gesamte Fabrik ausgerollt werden kann.

2.1.8 Demonstrator „KI-basierte Assistenzsysteme zur Zyklenoptimierung“ (AP 8)

Der Demonstrator 2 „Zyklenoptimierung“ wurde von den Partnern Gildemeister Drehmaschinen GmbH, Artis Marposs Monitoring Solutions und dem Institut für Fertigungstechnik und Werkzeugmaschinen (IFW) bearbeitet. Dabei wurde die Optimierung so genannter Zerspanzyklen (Programme) betrachtet, die Bewegungsabläufe von Werkstück und Werkzeug bei komplexen Zerspanprozessen steuern. Der Demonstrator betrachtet beispielhaft den Prozess des Wälzschälens, also die Fertigung von Verzahnungen. Dabei „schält“ ein rotierendes Werkzeug unter hoher Drehzahl eine Verzahnung in ein ebenfalls rotierendes Werkstück. Wie hoch die Produktivität des Prozesses und die Qualität der Bauteile sind, hängt von einer Vielzahl von Eingangsparametern ab. Die Parameter für das Wälzschälen optimal einzustellen ist aufwändig und erfordert tiefgehendes Prozesswissen. Aufgrund der hohen Drehzahl von Werkzeug und Werkstück ist eine Detektion von Fehlern (Überwachung) zudem nur für erfahrende Fachkräfte möglich. Das Ziel des Demonstrators „Zyklenoptimierung“ ist daher die Verfügbarkeit eines KI-basierten Assistenzsystem, das Wälzschälzyklen automatisiert parametriert und überwacht.

In Unterarbeitspaket (UAP) 8.1 wurden zunächst von allen Beteiligten Systemanforderungen definiert und in einem Anforderungsdokument zusammengefasst. Die Anforderungen umfassen u. a. Reaktionszeiten, aufgezeichnete Signale, Rechenressourcen und Schnittstellen zwischen den einzelnen Komponenten. Basierend auf den Anforderungen wurde ein Konzept für den KI-Assistenten entwickelt (s. Abbildung 44). Der KI-Assistent basiert auf einem Prozessmodell (1). Vergleichbar zum eigentlichen Prozess, gibt das Prozessmodell für gegebene Prozessparameter, Prozessdaten wie Motorströme oder Bauteilgeometrien zurück. Das Prozessmodell bildet die Grundlage für die Parametrierung (2) und die Prozessüberwachung (3). Die Parametrierung identifiziert zunächst mit verstärkendem Lernen basierend auf dem Prozessmodell die optimalen Prozessparameter. Im zweiten Schritt werden die Parameter dann basierend auf den real gemessenen Form- und Lagetoleranzen angepasst. Die Parametrierung wird offline, nach dem Prozess durchgeführt. Die Prozessüberwachung (3) führt prozessparallel einen

Soll-Ist-Abgleich zwischen modellierten und im Prozess gemessenen Daten durch. Liegt eine Anomalie, also ein unbestimmter Fehler vor, wird die Bauteile erneut vermessen und die Parametrierung erneut durchgeführt. Durch das zweistufige Vorgehen können 100%-Kontrollen des vermessenen Bauteils vermieden werden.

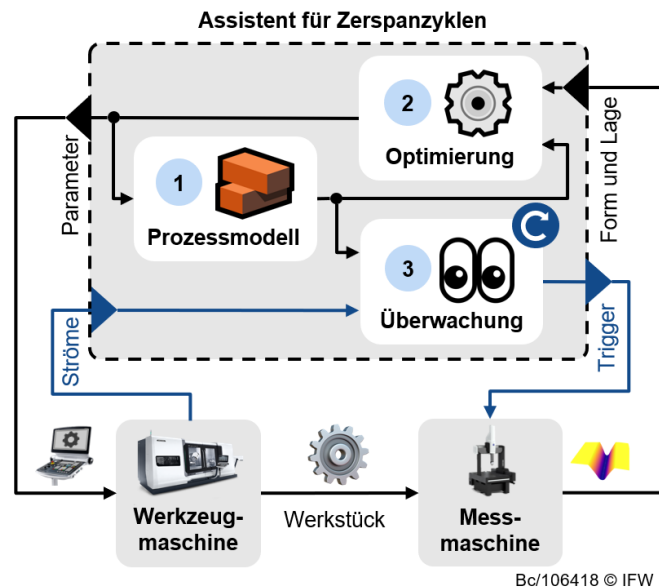


Abbildung 44: Aufbau des KI-Assistenzsystems

Der Wälzschälzyklus wurde im Projekt an den Werkzeugmaschinen DMG Mori NTX1000 und DMG Mori CTX 2000 Gamma am IFW integriert. Zusätzlich wurde an beiden Maschinen das Genior-Modular Modul des Partners Artis Marposs integriert. Das Modul wurde entsprechend angepasst, um das Logging von Daten zu ermöglichen. Zum Abschluss des Arbeitspakets wurden Trainings- und Arbeitsdatensätze aufgezeichnet. Dafür wurden Achskreuzwinkel, Kippwinkel, Werkzeugradius und die Zähnezahl des Werkstücks variiert. Aufgezeichnet wurden die Ströme, Regelfehler und Positionen der Antriebe. Anschließend wurden die Zahnräder vermessen, um die Zahndicke und Flankenlinienabweichung zu ermitteln.

In UAP 8.2 wurde die KI-basierte Prozessüberwachung (3) für den Assistenten entwickelt. Dafür wurde zunächst eine Segmentierung aufgebaut, die den Prozess basierend auf Sprüngen im Antriebsstrom der Vorschubachse in rekombinierbare Prozesssegmente aufteilt. Die Segmente werden nach ihrer Abfolge im Prozess unterschieden. Nachfolgend wurde eine Datenstruktur erarbeitet. Die Zeitreihendaten werden mit ihrem aktuellen Zeitstempel für jedes einzelne Signal abgelegt. In einer separaten Tabelle für Meta-Daten wurden die Parameter der einzelnen Segmente abgelegt. Die Segmente werden über den Start- und Stopzeitstempel den Zeitreihendaten zugeordnet. Im finalen Schritt wurde eine Überwachungsmethode entwickelt und umgesetzt.

Als Überwachungshardware kommt das Genior Modular System von Artis Marposs zum Einsatz. Dieses modulare Überwachungssystem besteht aus dem Hauptmodul GEMCPU und kann bei Bedarf um weitere Sensormodule erweitert werden. Im Idealfall kommt die Überwachung jedoch ohne zusätzliche Sensoren aus, indem die internen Daten der Maschinensteuerung als Signalquelle verwendet werden. Die Übertragung dieser Daten an Genior Modular erfolgt normalerweise über einen Feldbus wie Profinet oder Profibus, wie es im Fall der DMG Mori NTX1000 auch realisiert wurde. Da eine Installation von

Genior Modular über einen Feldbus jedoch eine tiefe Integration in die Maschinensteuerung erfordert, wurde für die DMG Mori CTX 2000 Gamma auf eine Retrofit-Lösung per Ethernet zurückgegriffen.

Die Idee bei dem neuen Überwachungsverfahren für die Wälzschälbearbeitung ist es, die von der KI erzeugte Referenzkurve direkt im NC-Programm auf der Steuerung zu hinterlegen und während der Bearbeitung an Genior Modular zu übertragen. Die Übertragung der Signalverläufe über das NC Programm an das Genior Modular Modul ist in Abbildung 45 dargestellt. Aus der Referenzkurve ergibt sich die obere Überwachungsgrenze, indem zu jedem Referenzwert eine einstellbare Sicherheit hinzugefügt wird. Da die komplette Bearbeitung in einem einzelnen Zyklusaufwurf stattfindet, erfolgt die Synchronisierung der Referenzwerte per Zeitstempel. Neben der Referenzkurve und dem Zeitstempel wird zusätzlich noch eine untere Grenze übertragen, sodass sich diese von der oberen Grenze unterscheiden kann.

Die Übertragung der Daten vom NC-Programm an die Überwachung erfolgt im Fall der Ethernet-Installation über R-Parameter und im Fall der Feldbusinstallation über den sogenannten Dualport-RAM der Siemens-Steuerung. Auf diesen speziellen Speicher ist wie bei den R-Parametern ein Zugriff vom NC-Programm aus möglich, sodass sich die Programme der beiden Maschinen nur geringfügig voneinander unterscheiden. Eine kurze Verweilzeit von 0,1 s pro Datensatz ist bei einer Taktrate von 10 ms ausreichend, um eine Übertragung der Daten zu gewährleisten. Ein Vorlesestopp STOPRE nach jedem Datensatz ist erforderlich, damit zum entsprechenden Zeitpunkt auch die Werte der richtigen NC-Zeile übertragen werden. Folgendes Beispiel zeigt drei Datensätze mit jeweils Referenzwert, unterer Überwachungsgrenze und Zeitstempel für die DMG CTX 2000 Gamma:

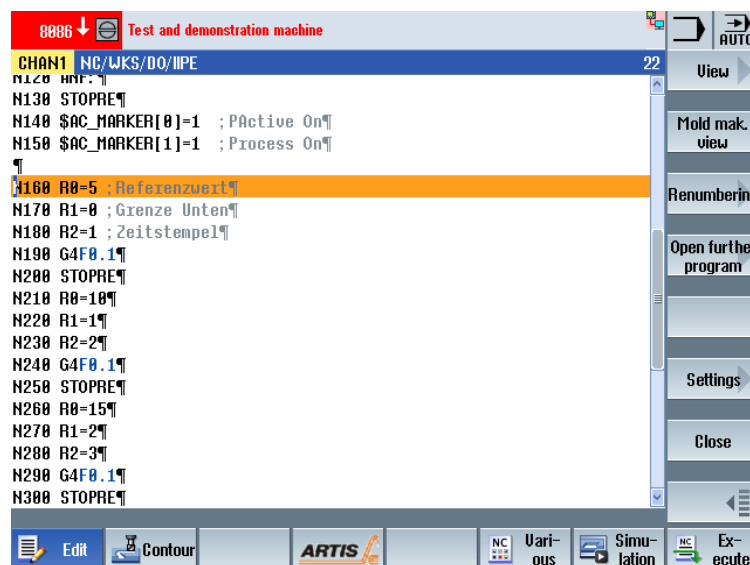


Abbildung 45: Übertragung der Referenzwerte über das NC-Programm

Der Ansatz der Übertragung der Referenzverläufe ist mit dem bisherigen Genior Modular nicht möglich. Aus diesem Grund musste das neue Verfahren als eigenes Überwachungsplugin implementiert werden. Dieses Plugin empfängt die Datensätze von der Maschine und führt darauf basierend eine eigene Überwachung mittels Soll-/Ist-Vergleich durch.

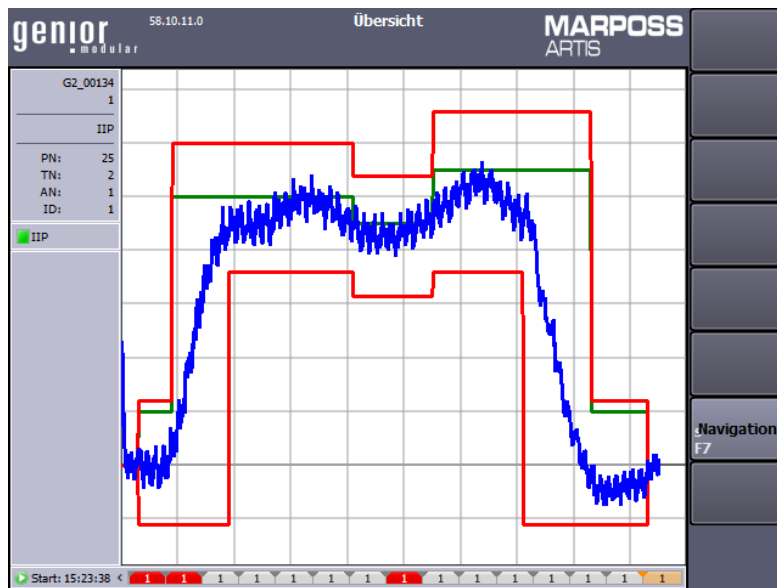


Abbildung 46: Signalverläufe während der Überwachung eines Wälzschälzyklus

In Abbildung 46 sind die Signalverläufe des Stroms der Vorschubachse während der Überwachung eines Wälzschälzyklus dargestellt. Innerhalb der ersten Sekunden erfolgt die Übertragung der Referenzdaten und mit Erreichen einer definierten Position wird die Überwachung scharfgeschaltet. Auf diese Weise hat die Dauer der Spindelsynchronisierung und damit der Startzeitpunkt der Bearbeitung keinen Einfluss mehr auf die Überwachung, der bei jedem Durchgang variiert und somit nicht vorhersehbar ist. Das Zeitverhalten der anschließenden Bearbeitung lässt sich hingegen gut abschätzen, so dass eine zeitgesteuerte realisiert werden konnte. Nach dem Prozess, wenn der gesamte Signalverlauf vorliegt, wird die Schätzung des Modells noch einmal aktualisiert (Precision Monitoring, rot). Dies ermöglicht die Überwachung mit einer erhöhten Genauigkeit.

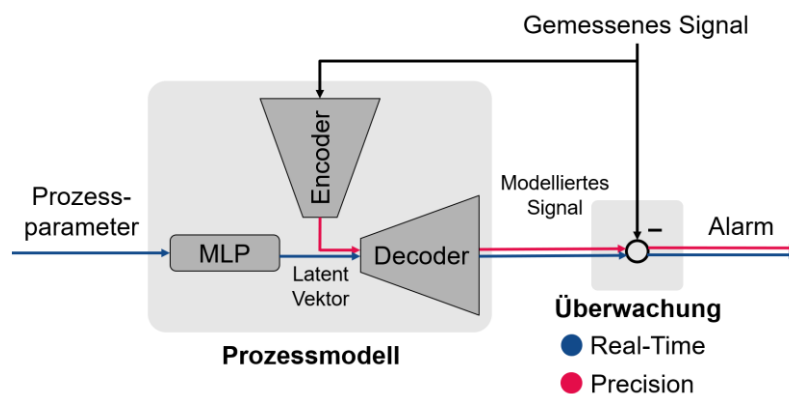


Abbildung 47: Aufbau der KI-Prozessüberwachung

In Abbildung 47 ist der Aufbau der Überwachung dargestellt. Die Hauptkomponente der Überwachung ist das Prozessmodell, das aus drei neuronalen Netzen besteht. Zunächst wird vor dem Prozess basierend auf den Prozessparametern ein Signalverlauf modelliert (Real-Time Monitoring, blau). Das Prozessmodell besteht in diesem Fall aus einem Multi-Layer-Perceptron (MLP) und dem Decoder eines Variational Auto Encoders. Während des Prozesses das modellierte mit dem gemessenen Signal in Echtzeit verglichen. Dieses Real-Time Monitoring wurde auf dem Artis Marposs Genior Modular umgesetzt. Nach dem Prozess, wenn der gesamte Signalverlauf vorliegt, wird die Schätzung des Modells

noch einmal aktualisiert (Precision Monitoring, rot). Dies ermöglicht die Überwachung mit einer erhöhten Genauigkeit. Dabei wurde gezeigt, dass die Überwachung in der Lage ist, kritischen Verschleiß des Werkzeuges zu detektieren.

Als überwachtes Signal wird der Strom der Vorschubachse verwendet. In einer Analyse mit dem beschriebenen Überwachungsansatz hat sich dieses Signal als besonders sensitiv erwiesen. Da die Werkzeugschneide mit zunehmendem Verschleiß kürzer wird taucht das Werkzeug einige Millisekunden später ins Werkstück ein, als bei einem arbeitsscharfen Werkstück. Da die Identifikation qualitätsrelevanter Prozesssignale insbesondere auf die Prozessüberwachung einzahlt, wurde das UAP 8.3.1 entsprechend in die Entwicklung der Prozessüberwachung (3) vorgezogen.

Die im Wälzschälzyklus hergestellten Verzahnungen müssen verschiedene Toleranzanforderungen erfüllen. Die Profillinienabweichung und Zahndicke gehören bei Lauf- und Passverzahnungen zu den wichtigsten Toleranzgrößen. Wie groß die Abweichung der gefertigten Verzahnungen ist, hängt insbesondere von Geometrieparametern ab, nämlich vom Achskreuzwinkel, vom Kippwinkel des Werkzeugs und vom Achsabstand. Die theoretisch idealen Parameter werden vom Werkzeughersteller vorgegeben. Aufgrund von Störungen, z. B. in der Maschinenkinematik, kann es jedoch vorkommen, dass in der Realität trotzdem Abweichungen auftreten. In diesem Fall ist bislang eine erfahrene Fachkraft notwendig, die die genannten Zyklus-Parameter adaptiert. Die Adaption ist zusätzlich von anderen Parametern abhängig, die aber nicht variiert werden können.

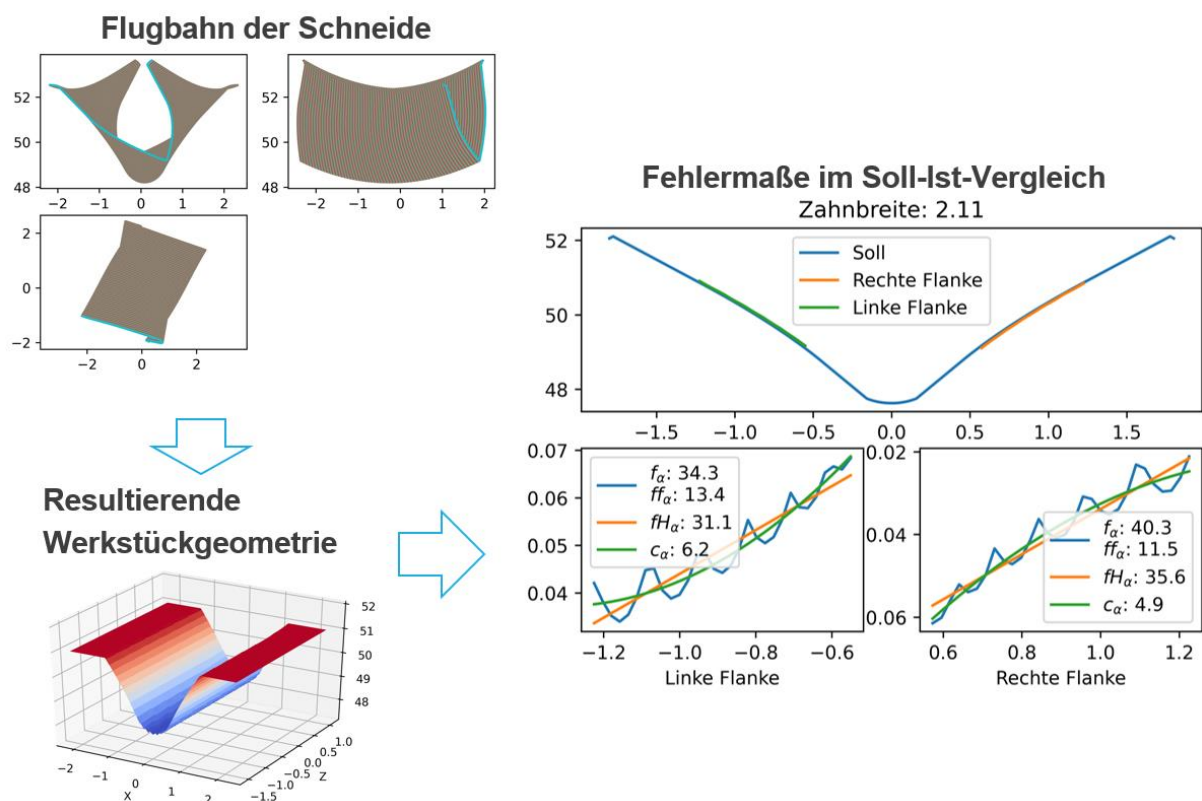


Abbildung 48: Geometrisches Prozessmodell für die KI-Parametrierung

Um den Bediener mit einem Vorschlag für die Parameter-Adaption zu unterstützen, wurde in UAP 8.3 eine KI-basierte Prozessparametrierung entwickelt. Für die Parametrierung relevant sind insbesondere Zyklus-Parameter. Die Basis der Parametrierung ist ein geometrisches Prozessmodell, das in Abbildung 48 dargestellt ist. Über eine Koordinatentransformation wird zunächst die Flugbahn der Schneide bestimmt. Aus der resultierenden Werkstückgeometrie werden dann unterschiedliche Fehlermaße berechnet.

Das Prozessmodell ist in der Lage für einen gegebenen Satz an Parametern, die Form und Lagetoleranzen vorherzusagen. Für die Parametrierung muss das Prozessmodell jedoch invertiert werden. Für eine gegebene Abweichung sollen die Form- und Lagetoleranzen berechnet werden. Für die Lösung des Problems wurden verschiedene KI-Methoden betrachtet und Reinforcement Learning ausgewählt. Verstärkendes Lernen wird benötigt, da es verschiedene Parameterkombinationen gibt, mit denen eine geringe Profillinienabweichung erreicht werden kann. Aufgrund der fehlenden Eindeutigkeit ist überwachtes Lernen daher nicht möglich. Somit wurde in UAP verstärkendes Lernen (Reinforcement Learning) fokussiert. Aus den Fehlermaßen wurde eine Verlustfunktion aufgestellt, die für das Training eines Policy Gradient Verfahrens (Deep Deterministic Policy Gradient) verwendet wurde. Somit ist es möglich für unterschiedliche Werkstücke, Werkzeuge und Prozessparameter die optimalen Prozessparameter zu berechnen.

Das ursprünglich eingeplante Bedienerfeedback zur Parametrierung wurde durch Qualitätsdaten der Werkstücke (Form- und Lagetoleranzen) ersetzt und das UAP 8.4 somit in das UAP 8.3 integriert. Die Größe eines Datensatzes, die mit Bedienerfeedback annotiert werden kann, ist sehr gering und mit hohen Kosten verbunden. Die Gildemeister Drehmaschinen GmbH hat in einer Berechnung mit Roh-teilkosten, Werkzeugkosten, Fertigungskosten, sowie Maschinen- und Personalstundensätzen Gesamtkosten von etwa 10 € pro Werkstück (entspricht einem Datenpunkt) errechnet. Reinforcement Learning benötigt hingegen Millionen von Datenpunkten. Entsprechende Datensätze könnten daher nur in einer Serienfertigung generiert werden. Im Projekt wurde daher ein Ansatz zur Datenakquise (siehe Abbildung 49) verfolgt, der im Versuch aufgezeichnete Daten um Simulationsdaten erweitert.

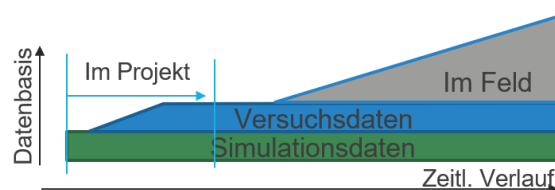


Abbildung 49: Ansatz zur Datenakquise mit Simulationsdaten

In UAP 8.5 wurden das Prozessüberwachungssystem und die Prozessparametrierung validiert. Die Überwachung wurde anhand der Prozessdaten der Fertigung von insgesamt 23 Werkstücken, davon elf mit verschlissenem Wälzschälwerkzeug und 12 Werkstücke mit arbeitsscharfem Werkzeug validiert. Die Prozessüberwachung könnte verschlissene und arbeitsscharfe Werkzeuge im zweiten Schnitt bei gleichen Verzahnungsparametern zu 100 % unterscheiden (s. Abbildung 50).

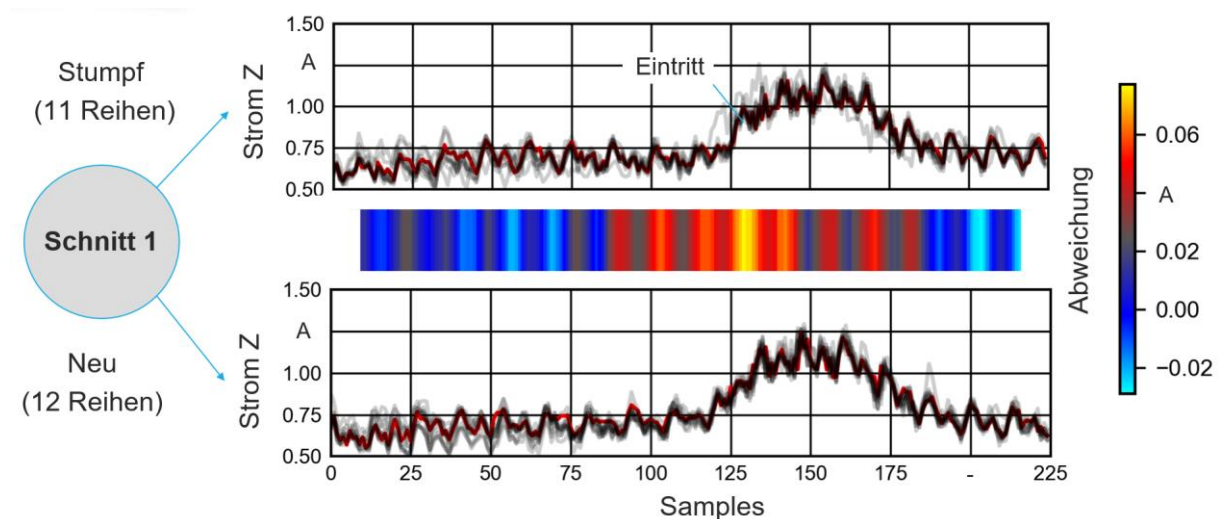


Abbildung 50: Validierung der Prozessüberwachung

Zur Validierung der KI-Parametrierung wurden Qualitätsdaten von 25 Verzahnungen mit variierendem Werkzeugdurchmesser und Achskreuzwinkel aufgenommen. In Abbildung 51 ist für einen beispielhaft gewählten Werkzeugdurchmesser die Gesamtprofilinienabweichung über dem Achskreuzwinkel dargestellt. Neben den gemessenen Profilinienabweichungen ist in grün ist zusätzlich die simulierte Profilinienabweichung dargestellt. Das Werkzeug ist für einen Achskreuzwinkel von 20° ausgelegt. In der Theorie ist die Profilinienabweichung bei 20° daher null. In der Messung beträgt die Profilinienabweichung jedoch $30\text{ }\mu\text{m}$. In der Simulation muss der Achskreuzwinkel um $1,2^\circ$ reduziert werden, um die minimale Profilinienabweichung zu erreichen. Entsprechend wird auch in der Realität der Achskreuzwinkel um $1,2^\circ$ reduziert, wodurch die Profilinienabweichung auf unter $20\text{ }\mu\text{m}$ sinkt. Dies entspricht einer Verbesserung von 33 %.

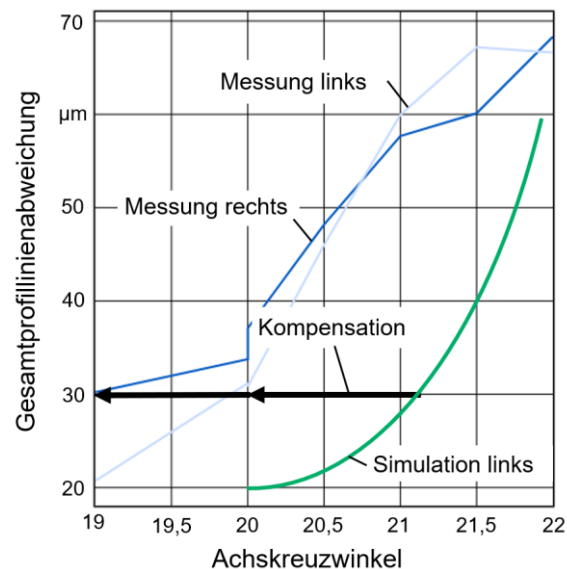


Abbildung 51: Gesamtprofilinienabweichung F_{h_alpha} über den Achskreuzwinkel

Das Ergebnis von 33 % Minimierung der Gesamtprofilinienabweichung ist nur für das gewählte Werkzeug und die verwendete Werkzeugmaschine NTX1000 repräsentativ. Abhängig von der Maschine

kann der Fehler deutlich höher sein. Die Versuche zeigen jedoch, dass es mit dem System möglich ist, die Flankenlinienabweichung deutlich zu minimieren.

Das Gesamtsystem wurde in einer Python-Umgebung implementiert und ist über ein Graphical User Interface (GUI) bedienbar, das mit der Python Bibliothek Streamlit programmiert wurde. Das GUI ist in Abbildung 52 dargestellt und umfasst alle beschriebenen Modelle des Demonstrators zusammen.

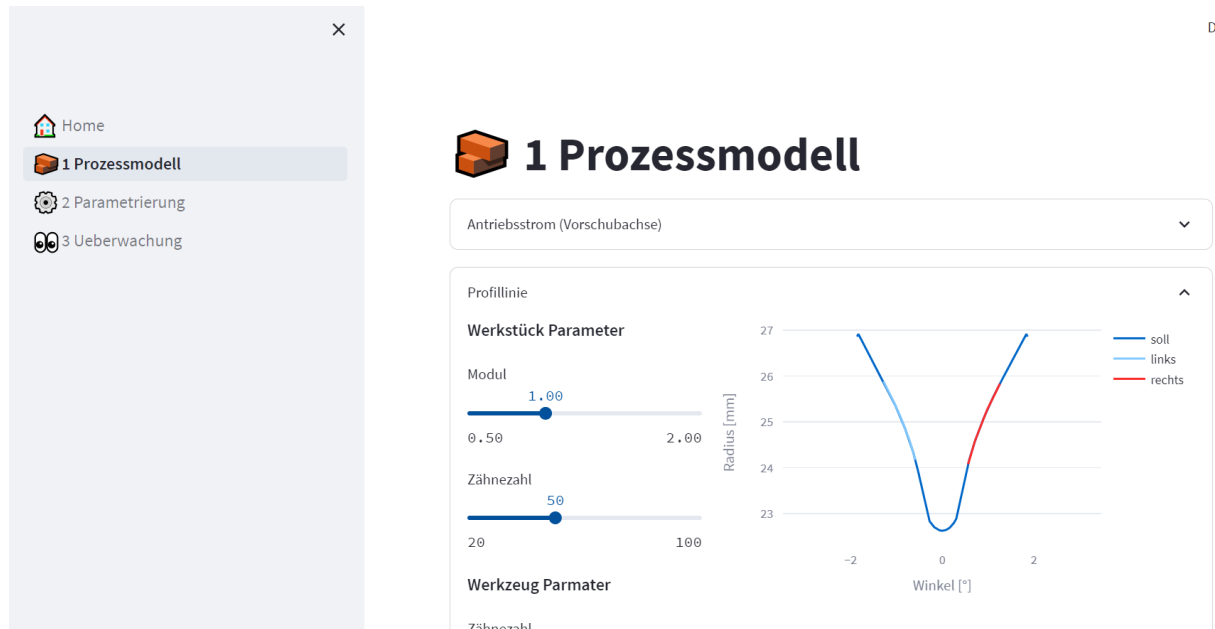


Abbildung 52: Graphische Benutzerschnittstelle des Demonstrators

In UAP 8.6 wurden die erzielten Ergebnisse der Öffentlichkeit zugänglich gemacht. Dafür wurden Veröffentlichungen im VDW Jahresbericht [DEN23c], auf der EMO Webseite [DEN22c], im PZH Magazin [BEC22] und im VDW Branchenreport erstellt [DEN22b]. Die erzielten Ergebnisse wurden zudem in einem Werkstattgespräch mit den Unternehmen Gildemeister Drehmaschinen GmbH und Marposs Monitoring Solutions GmbH präsentiert. Hierbei wurde der Demonstrator live vorgeführt und die erreichten Fortschritte in der KI-basierten Überwachung und Parametrierung erläutert. Ein öffentlich zugänglicher Live-Stream ermöglichte die Teilnahme einer breiteren Community. Beim der FDT Demo-Day Veranstaltung mit über 100 Teilnehmern in Berlin wurde der Demonstrator zudem einem breiten Fachpublikum vorgeführt.

2.1.9 Demonstrator „Autonome Roboterzellen zur Endprüfung in der Leiterplattenfertigung“ (AP 9)

Überblick über das Demonstrator Teilprojekt

Dieser Demonstrator ist bei der Firma Sennheiser electronic GmbH & Co. KG am Standort Wedemark (bei Hannover) angesiedelt und wurde von einem interdisziplinären Team der Unternehmen Bitmotec, Rapidminer (jetzt Altair), und slashwhy (KI-Accelerator) sowie des Instituts für Fertigungstechnik und Werkzeugmaschine (IFW) gemeinsam entwickelt. Erforscht werden sollen KI-Lösungen auf Arbeitsplatz- und Produktionssteuerungsebene, die im Demonstrator exemplarisch umgesetzt und innerhalb des Produktionssystems der Firma Sennheiser erprobt werden. Dort wurde eine vollständig autonome, agile und skalierbare Prü fzelle zur Leiterplattenendprüfung in der Sennheiser Future Factory aufgebaut

und zur Evaluation in das laufende Produktionssystem integriert. Zur Vermeidung monotoner, manueller Arbeiten werden solche Prü fzellen herkömmliche Arbeits- und Prüfplätze zukünftig sukzessive ergänzen und im Endeffekt effizient ersetzen.

Der Aufbau des Demonstrators erfolgte in vier Unterarbeitspaketen: Die Erstellung der Roboterzelle und seiner Elemente zur Materialbereitstellung, der automatischen Zu- und Abführung der zu prüfenden Leiterplatten und der dafür notwendigen Adapter (Austauschsätze), den Elementen zum Handling wie Werkzeuge und Werkzeugwechseleinheiten, sowie der intelligenten Identifikation von Systemschnittstellen wird im **Arbeitspaket 9.1 (Autonomes Rüsten)** adressiert. Die Erforschung und praktische Erprobung von KI-Lösungen auf Arbeitsplatz-Ebene wird in **Arbeitspaket 9.2 (Intelligente Leiterplattenprüfung)** gezeigt, in der KI-Modelle zur Identifikation von Pseudofehlern der Leiterplattenprüfung Schwerpunkt waren. Die notwendige Systemintegration der gesamten Zelle in das bestehende Produktionssystem inklusive der Modellierung des digitalen Zwillings erfolgt in **Arbeitspaket 9.3 (Systemintegration)**. Im Rahmen des **Arbeitspakets 9.4 (Produktionssystemmanagement)** galt es, ein Agentensystem zur dezentralen Steuerung und Optimierung einer zielgrößenoptimalen Entscheidungsfindung zu implementieren.

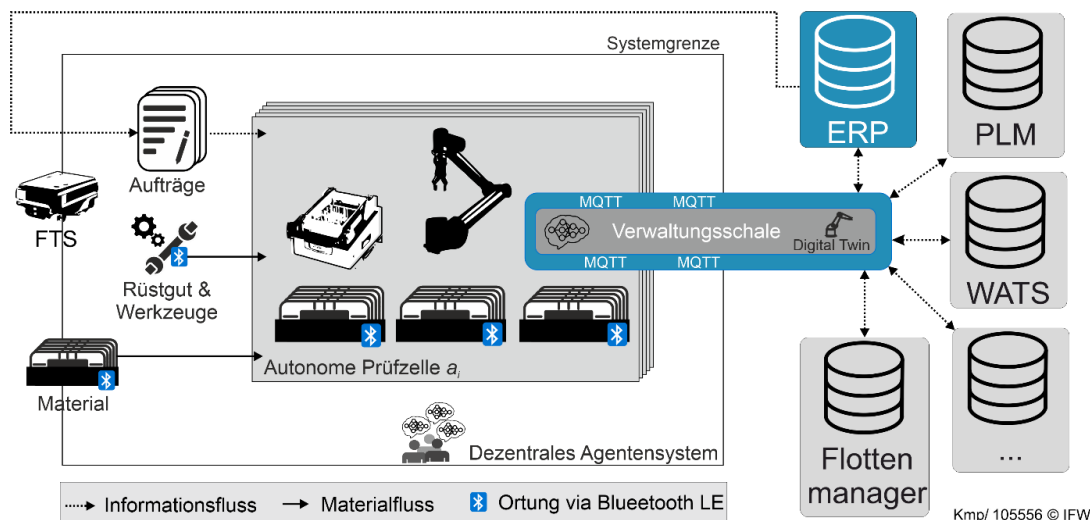


Abbildung 53: Schematische Darstellung einer autonomen Roboterzelle zur Leiterplattenprüfung

Arbeitspaket 9.1: Autonomes Rüsten

Zu Beginn des AP 9.1 wurden besonders relevante Submodule identifiziert. Die Identifikation startete mit der gemeinsamen Begehung des Shop-Floors und der gemeinsamen Analyse des gesamten Prüfprozesses und der darin zu durchlaufenden Arbeitsvorgangsfolge in der Leiterplattenendprüfung. Unter Zuhilfenahme zusätzlicher Expertengespräche, konnten die Module **Test- bzw. Prüfmodul**, **Materialbereitstellung**, **Materialabfuhr und -transport**, **Robotersystem** und die **Leiterplattenidentifikation** in gemeinsamen Workshops abstrahiert und somit als **relevante Submodule identifiziert** werden.

Für die verbleibenden o.g. relevanten Submodule wurden in gemeinsamen Workshops unter Sennheiser Verantwortung verschiedene Grobkonzepte zur Umsetzung erarbeitet. Dazu wurde für das Prüfmodul vorab eine Unterteilung des notwendigen Prüfequipments in einen universellen Grundadapter (der Firma INGUN) und in sog. Austauschsätze (Nadel- und Andruckplatten) beschlossen. Der Grundadapter wurde einem konkreten Prüfplatz zugeordnet, an dem weiteres Prüfequipment (Hard- und Software)

fest angeschlossen wird. Damit ist der Grundadapter universal einsetzbar und kann mittels der Austauschsätze für eine Vielzahl von Leiterplatten flexibel eingesetzt werden. Somit kann eine automatisierte Rüstung der Prü fzelle erfolgen. Auf dieser Annahme wurden anschließend Grobkonzepte erarbeitet, die aus Tabelle 3 und Tabelle 4 zu entnehmen sind. Eine Kombination dieser Grobkonzepte ergab insgesamt sechs mögliche Layoutvarianten, wie eine autonome Prü fzelle umgesetzt werden könnte (siehe Abbildung 54).

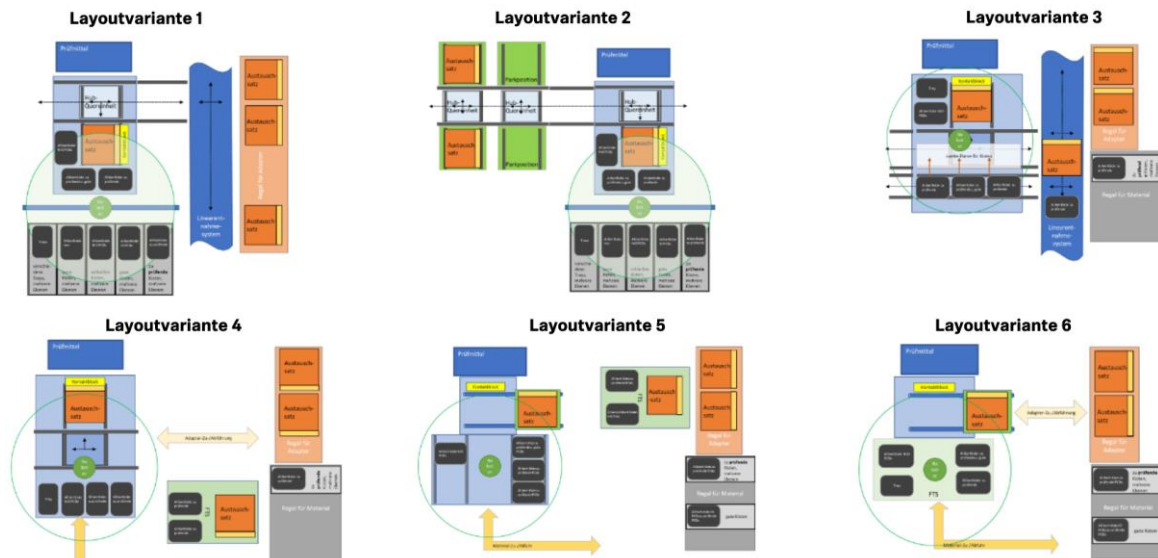
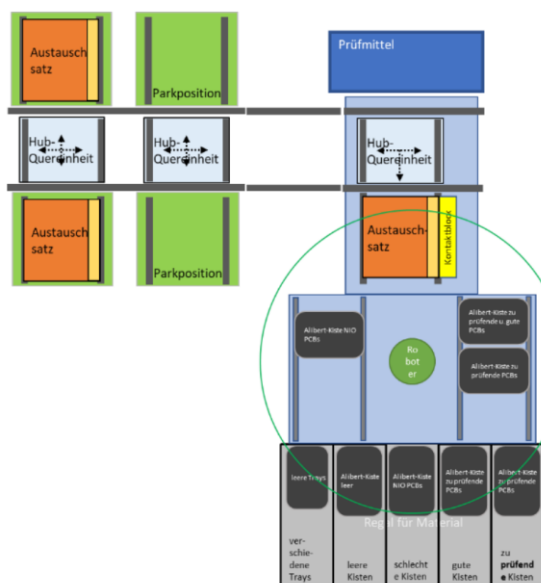


Abbildung 54: Layoutvarianten

Zur weiteren Konkretisierung wurden zusammen mit betroffenen Stakeholder-Gruppen Bewertungskriterien gesammelt und mithilfe eines Kreuzvergleichs hinsichtlich ihrer Priorität bewertet. Diese Bewertungskriterien wurden auf die drei Submodule angewendet, um geeignete Varianten zu identifizieren und ein ganzheitliches Grobkonzept der Prü fzelle zu entwerfen. Inklusiv einer ersten Kostenschätzung wurden zwei finale Varianten erarbeitet. Die erste finale Variante aus Abbildung 55 enthält aus Kostengründen kein fahrerloses Transportsystem (FTS).



Merkmale

- ▶ Prüfadapter
 - a) „Parkposition“ am Transfersystem
 - b) Zu-/Abführung über Transfersystem in Prü fposition
- ▶ Materialzu-/abfuhr
 - a) Direkt aus Regal mit Roboter
- ▶ Roboter
 - a) Auf flexiblem Wagen, gekoppelt
 - b) Ein Roboter für Kisten- und PCB-Handling

Vorteile

- ▶ Flexible Zu-/Abführung
- ▶ Flexibler, koppelbarer Robotertisch
- ▶ Geringe Komplexität Rüsten
- ▶ Einfache Erweiterung um weitere Plätze

Nachteile

- ▶ Material muss arbeitsplatzbezogen bereitgestellt werden
- ▶ Zugänglichkeit

Kosten Material (ohne Messtechnik, Materialkisten, Trays)

- ▶ 131.000€ vollautomatisiert
- ▶ 116.000€ teilautomatisiert (ohne Rüsten Prü fadapter)



Abbildung 55: Grobkonzept für autonome Prü fzelle (ohne FTS)

Die zweite finale Variante aus Abbildung 56 enthält ein FTS, wodurch die autonome Prüfwelle zusätzlich in eine flexible autonome Fertigungsstruktur eingebettet werden kann.

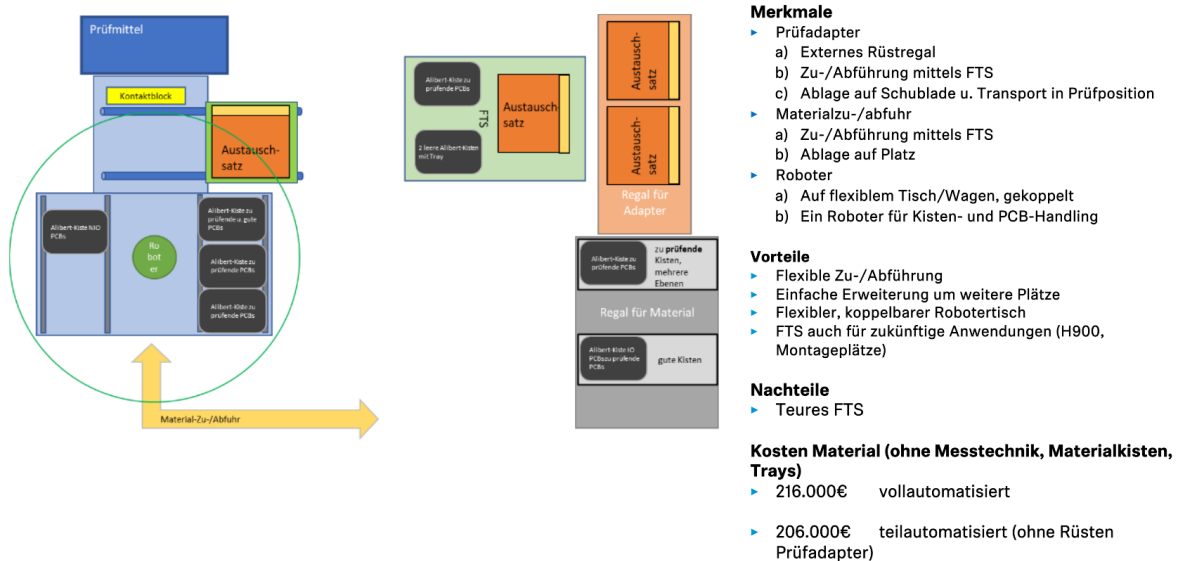


Abbildung 56: Grobkonzept für autonome Prüfwelle (mit FTS)

Tabelle 3: Erarbeitete Grobkonzepte für Prüfmodule

	Variante 1	Variante 2	Variante 3	Variante 4	Variante 5
Lagerung	Externes Regal	Auf Parkpositionen in einem Transfersystem	Externes Regal	Externe Regal	Externes Regal
Entnahme aus Lager	Linearentnahmesystem	Via Hub-Quer-Einheit	FTS mit Hub-Quer-Einheit	Roboter montiert auf FTS	FTS mit Hub-Quer-Einheit
Transport und Ablage zum/ am Prüfplatz	Ablage auf Transfersystem und Transport zur Prüfposition	Transfersystem inkl. Hub-Quer-Einheit am Prüfplatz	Transfersystem (FTS schleußt wieder ein)	Transport via FTS und Ablage durch Roboter des FTS auf Quer-Einheit	Transport via FTS und Ablage auf Hub-Quer-System des Prüfplatzes

Tabelle 4: Erarbeitete Grobkonzepte für automatisierte Materialbereitstellung und -transport

	Variante 1	Variante 2	Variante 3	Variante 4
Lagerung	Externes Lager am Platz	Externes Regal	Externes Regal	Externes Regal
Entnahme aus Lager	Durch Roboter	Linearentnahmesystem	FTS mit Hub-Quer-System	FTS mit Roboter
Transport und Ablage zum/ am Prüfplatz	Durch Roboter	Förderband / Transfersystem zum Prüfplatz (Handhabung anschl. Via Roboter)	FTS mit Hub-Quer-System	FTS mit Roboter (verbleibt am Prüfplatz)

Tabelle 5: Erarbeitete Grobkonzepte für Robotersystem

	Variante 1	Variante 2	Variante 3	Variante 4
Position des Roboters	Befestigung auf Linearachse	Überkopf an Traverse (fest am Prüfplatz)	Prüfmodule und Roboter auf extra Tischen (koppelbar)	Befestigung auf FTS (koppelbar mit Prüfplatz)
Handling von	Leiterplatten und Kisten	Leiterplatten und Kisten	Leiterplatten und Kisten	Leiterplatten und Kisten, Bereitstellung von Prüfadaptern

Kernergebnis dieses UAP ist vor allem die Erarbeitung eines Stufenkonzepts, das Einzug in den Anforderungskatalog gefunden hat. Das Stufenkonzept sieht eine stufenweise Erhöhung des Automatisierungs- bzw. Autonomiegrads vor. Da mit dem Start des Projektes noch kein Demonstrator bestand und die AP 9.1 bis 9.4 parallel begonnen wurden, sollten erste Produktionslenkungs- und KI-Lösungen schon an vorhandenen manuellen Prüfplätzen erprobt werden, um mit einem höheren Reifegrad im Demonstrator verwendet werden zu können.

Zusätzlich wurde bei Bitmotec ein Versuchsstand konzipiert und konstruiert, der den Teil des Prüfplatzes mit der Tray Ablage und dem Kamerasystem maßstabsgetreu abbildet. Dadurch konnte bei Bitmotec bereits parallel zu der Prüfzelle an dem PCB-Erkennungssystem (Leiterplattenidentifikation) gearbeitet werden. Diese Aktivitäten sind in der Stufe 0 beinhaltet. In Stufe 1 wurde das Grobkonzept auf eine vollständig autonome Prüfung von Leiterplatten mithilfe eines Roboters begrenzt. Die Rüstung von Austauschätzen und der Materialtransport laufen in dieser Stufe manuell ab. In Stufe 2 wurde der Materialtransport innerhalb der Zelle automatisiert und in Stufe 3 liegt eine vollautomatisiert und autonom belieferte Prüfzelle vor. Darüber hinaus erfolgte durch das IFW die Begleitung der Aufstellung der Detailkonzepte.

Auf Basis des erarbeiteten Stufenkonzeptes wurden durch Sennheiser in den UAP 9.1.3 bis 9.1.6 für jedes Submodul die Detailkonzepte innerhalb der Zelle erstellt. Hierbei wurden folgende Aspekte fokussiert:

1. Hardwaremodularität
2. Erweiterbarkeit der Hardware (manuelle Bedienung vs. autonome Steuerung und Bedienung) und mögliche Duplizierung/Spiegelung des Prüfmodules
3. Nutzung vorhandener Kommunikationsbausteine (MQTT, OPC-UA)
4. Softwaremodularität realisiert durch Steuerungsmodule innerhalb der SPS
5. Erweiterbarkeit der Software

Das favorisierte Kommunikationsdesign (ein Agent steuert die Zelle, die Submodelle kommunizieren nicht untereinander) erwartete die Definition einer Schnittstelle/API für jedes identifizierte Submodul. Dabei wurden unter Berücksichtigung der bereits bei Sennheiser vorhanden Technologien und Erfahrungen aus vorhergehenden Projekten die Schnittstellenprotokolle MQTT und OPC-UA genutzt. Abbildung 57 zeigt am Beispiel des Submoduls „Rüstmodul Austauschatz“ die prinzipielle Realisierung eines Submodules sowohl in der Hardware als auch in der Software.

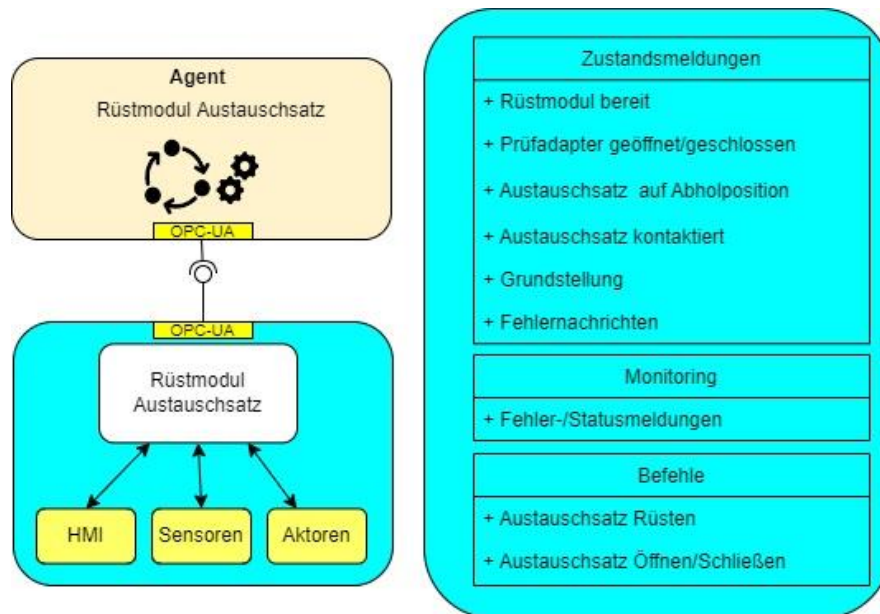


Abbildung 57: Prinzipielle Realisierung eines Submoduls am Beispiel „Rüstmodul Austauschsatz“

Die von Sennheiser erstellte Demonstratorzelle ist in Abbildung 58 dargestellt. Hier sind die Submodule

- Leiterplattenzufuhr,
- Leiterplattenabfuhr,
- Roboter,
- Rüstmodul
- und Prüfsystem

zu sehen.



Abbildung 58: Realisierung der Demonstratorzelle

Das Submodul Leiterplattenidentifikation wird mit dem von Bitmotec erstellten KI-basierten optischen Objekterkennungssystem (nachfolgend als PCB-Erkennungssystem bezeichnet) innerhalb der Demonstratorzelle in Abbildung 59 dargestellt.

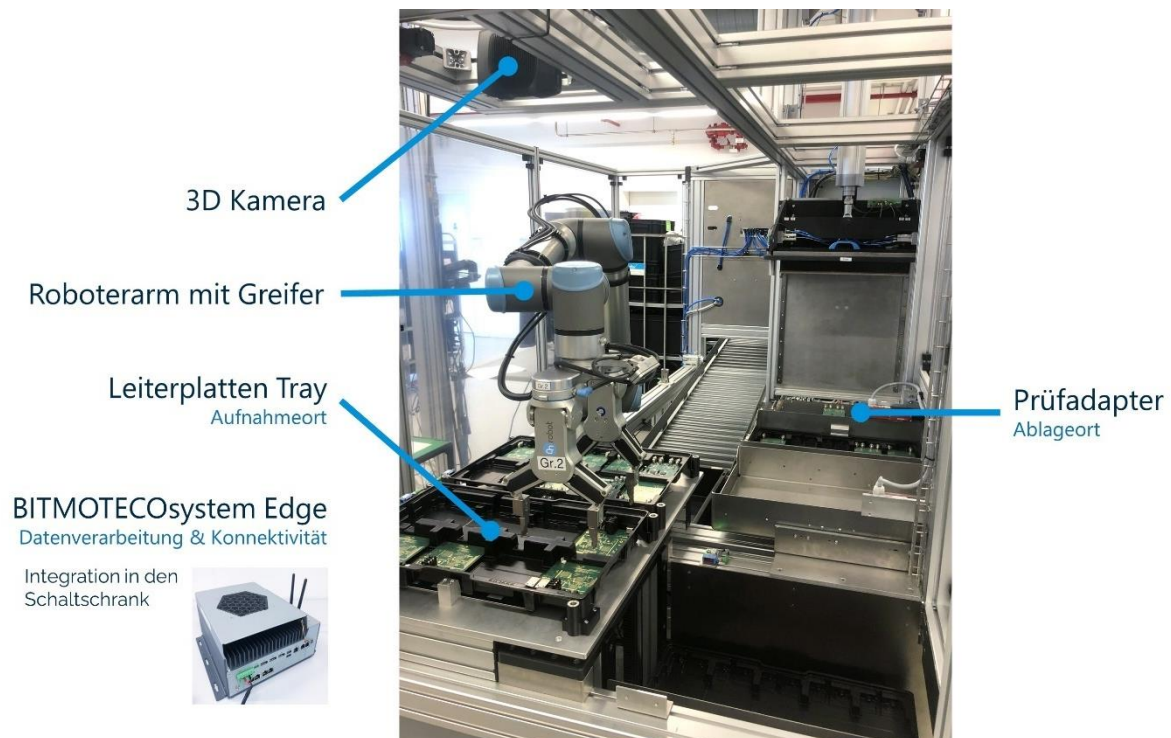


Abbildung 59: Demonstratorzelle mit integriertem PCB-Erkennungssystem bestehend aus 3D-Kamera und Edge-System zur lokalen Datenverarbeitung und Systemanbindung.

Die Umsetzung des Teilsystems eines autonomen, KI-basierten optischen Objekterkennungssystems zur robusten Klassifikation von Produkten (Leiterplatten – PCB) in der variantenreichen Fertigung erfolgte durch Bitmotec in mehreren Schritten.

Als Ausgangslage wurde ein separater Anforderungskatalog auf Basis der gewählten Layoutvariante der Prü fzelle für das PCB-Erkennungssystem erstellt. Hierbei wurden Systemanforderungen und funktionale Anforderungen unterschieden.

Systemanforderungen:

- Das System soll PCBs hinsichtlich Art, Lage und deren Greifpunkte analysieren können.
- Anbindung an das MORYX Framework zur Einbindung des Systems in die Produktionsumgebung sowie Asset-Kommunikation zum bidirektionalen Informationsaustausch über standardisierte I4.0 Schnittstellen.
- Übermitteln und Empfangen von Sicherheitszuständen.
- Das PCB-Objekterkennungssystem soll ein hohes Maß an Interoperabilität, Skalierbarkeit, Performanz, Transparenz und Sicherheit aufweisen.
- Auslegung als eigenständiges System mit Hard- und Software (Edge Computing)

Funktionale Anforderungen:

- **Tray Fächererkennung:** Es ist eine Funktion zur Fächererkennung von Trays bereitzustellen. Die Fächer der Trays können leer oder mit PCBs bestückt sein.
- **Tray-Fach Belegungserkennung:** Da die Trays teilweise unvollständig beladen sind, ist eine Prüfung der Fachbelegung erforderlich.
- **Leiterplattentyp Identifikation** Da mit dem System unterschiedliche Leiterplattentypen erkannt werden sollen, muss eine Leiterplattenidentifikationsfunktion bereitgestellt werden.
- **Leiterplatten Positions- & Lageerkennung + Greifpunkte:** Es ist eine Funktion zur Position und Lagererkennung einer Leiterplatte bereitzustellen. Darüber hinaus sollen die Greifpunktinformationen bereitgestellt werden.
- **Leiterplatten verrutscht/verkippt Erkennung:** Leiterplatten können in seltenen Fällen im Tray durch übermäßige Transporterschütterungen aus den dafür vorgesehenen "Tray-Schalen" rutschen. Diese liegen optisch auffällig und deutlich verrutscht im Tray und können mit großer Wahrscheinlichkeit nicht automatisiert gegriffen werden.
- **Erkennungszustände an Leitsystem übermitteln:** Ist es dem PCB-Erkennungssystem nicht möglich Art/Variante/Position der Leiterplatte zu erkennen, muss dies dem Leitsystem (MORYX) mitgeteilt werden.
- **Systemzustände an Leitsystem übermitteln:** Zustände des PCB-Erkennungssystems sind dem Leitsystem mitzuteilen, insbesondere Fehlerzustände.
- **System Shutdown:** Die Software soll eine Funktion erhalten, um das System geregelt herunterzufahren.
- **Auftragsannahme:** Das PCB-Erkennungssystem erhält vom Leitsystem (MORYX) die Angabe über das nächste zu inspizierende Fach des Trays. Die Erkennungs-Informationen bezüglich des ausgewählten Faches sind zurückzusenden (nicht alle Fächer gleichzeitig).
- **Sicherheitszustand:** Das Bilderkennungssystem erhält vom Leitsystem Informationen über den Sicherheitszustand des Gesamtsystems (z. B. Nothalt, etc.).
- **System mit CAD-Daten Anlernen:** Um den Einrichtungsaufwand möglichst gering zu halten, ist eine Funktion zum automatisierten Anlernen des Systems auf Basis von CAD-PCB-Daten wünschenswert. Das PCB-Erkennungssystem erhält vom Leitsystem oder zunächst testweise manuell die CAD-Daten der PCBs und erstellt sich auf dieser Datengrundlage synthetische Trainingsdaten. Auf dieser Basis wurde bei Bitmotec ein Versuchsstand konzipiert und konstruiert, der den Teil des Prüfplatzes mit der Tray Ablage und dem Kamerasystem maßstabsgetreu abbildet (siehe Abbildung 60).

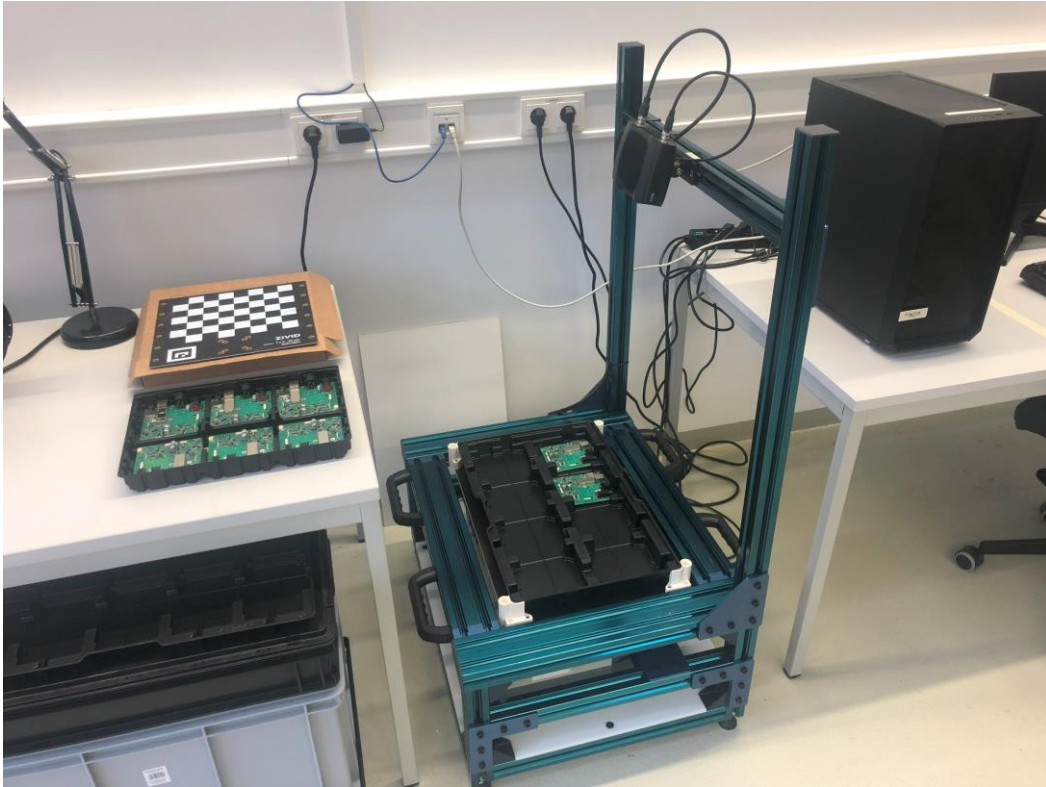


Abbildung 60: Versuchsstand von Bitmotec in der Laborumgebung

Es wurden Vorversuche zur Identifikation geeigneter Algorithmen für das KI-basierte „PCB-Erkennungssystem“ durchgeführt und geeignete Algorithmen identifiziert. In diesem Zuge wurden eigene Trainingsdatensätze durch Aufnahmen von den Leiterplatten und anschließendem Labeln erstellt. Die 2D-Trainingsdaten konnten durch Augmentierungsverfahren erweitert werden. Vorversuche haben gezeigt, dass in diesem Anwendungsfall mit einer ausschließlichen 2D-Objekterkennung nicht die erforderliche Genauigkeit und Zuverlässigkeit erzielt werden kann. Deshalb wurden neben den 2D Bilddaten auch 3D Bilddaten (Punktwolken) ausgewertet, um die Anforderungen an die Bildverarbeitung erfüllen zu können. Zum Aufnehmen der Rohdaten kam daher eine 3D Kamera mit Streifenprojektionsverfahren zum Einsatz, welche sowohl 2D als auch 3D Daten liefern kann.

Aus den Erkenntnissen der Vorversuche wurde das Softwarekonzept für die Umsetzung eines industrietauglichen KI-basierten „PCB-Erkennungssystems“ entworfen und umgesetzt. Die Integration der Softwarekomponenten erfolgte als spezielles Modul in das BITMOTECOSystem, welches als Edge Device die Datenverarbeitung direkt in der Demonstratorzelle ermöglicht. Nach der Abstimmung der Möglichkeiten der Asset-Kommunikation und deren Schnittstellen erfolgte die Planung, Spezifikation und Umsetzung der Schnittstelle an das übergeordnete Produktionsleitsystem (MORYX) mittels MQTT (vgl. Abbildung 61).

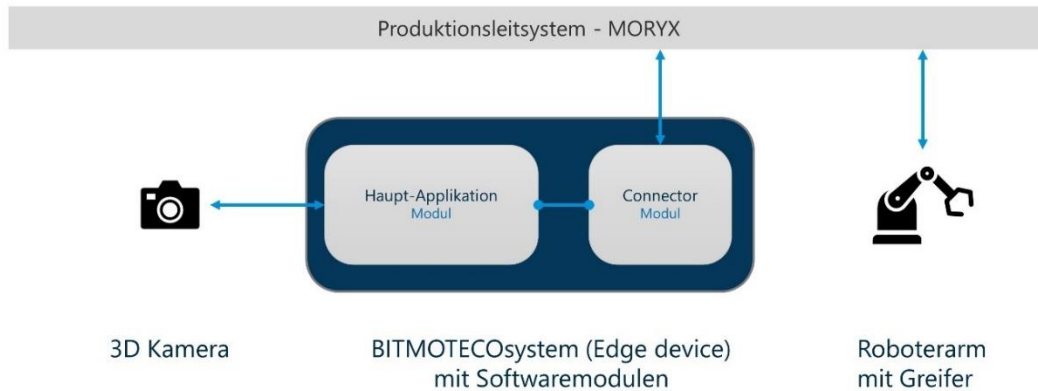


Abbildung 61: Systemkomponenten des PCB-Erkennungssystems: 3D Kamera angebunden an das Edge-System mit den Softwarekomponenten zur Objekterkennung und Klassifikation und der Konnektivität an das höhergelagerte Produktionsleitsystem.

Im Rahmen der Integration bei Sennheiser wurde das PCB-Erkennungssystem temporär abwechselnd sowohl am Sennheiser Demonstrator als auch am Versuchsstand bei Bitmotec eingesetzt. Diese Maßnahme diente dazu, umfangreiche Tests vorab bis zur finalen Integration im Demonstrator zu ermöglichen und sicherzustellen, dass alle Systemteile harmonisch zusammenarbeiten.

Das PCB-Erkennungssystem wurde im Laufe des vierten Quartals 2022 final erfolgreich in den Sennheiser-Demonstrator integriert. Diese Integration umfasste sowohl die Hardware- als auch die Software-seitige Anbindung an bereits bestehende Systeme bei Sennheiser (u.a. MORYX). Durch diese Einbindung konnte die Funktionalität des PCB-Erkennungssystems in einer realen Umgebung demonstriert werden.

Das PCB-Erkennungssystem wurde abschließend anhand der Anforderungen im Laborumfeld und unter Realbedingungen in der Sennheiser Demonstratorzelle validiert (vgl. Abbildung 62 und Abbildung 63). Die Validierungen in der Sennheiser Demonstratorzelle in Kombination mit dem Robotersystem werden über das Projektende hinaus fortgeführt, um eine spätere Verwertung der Ergebnisse zu ermöglichen.

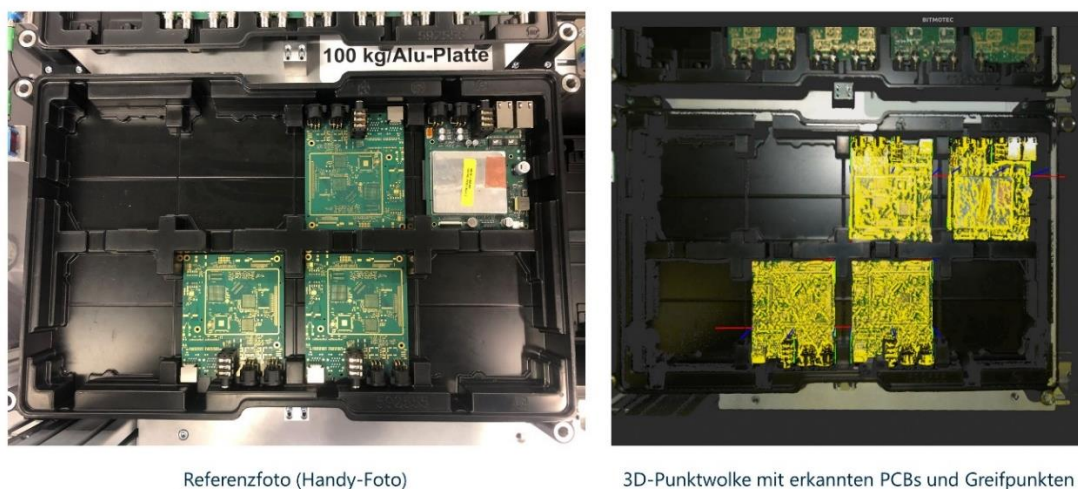


Abbildung 62: PCB-Erkennung: Referenzbild der Platinen-Anordnung im Tray (links), 3D-Visualisierung der Mess-punktwolke mit in Gelb eingefärbten PCB-Segmentationen (rechts).

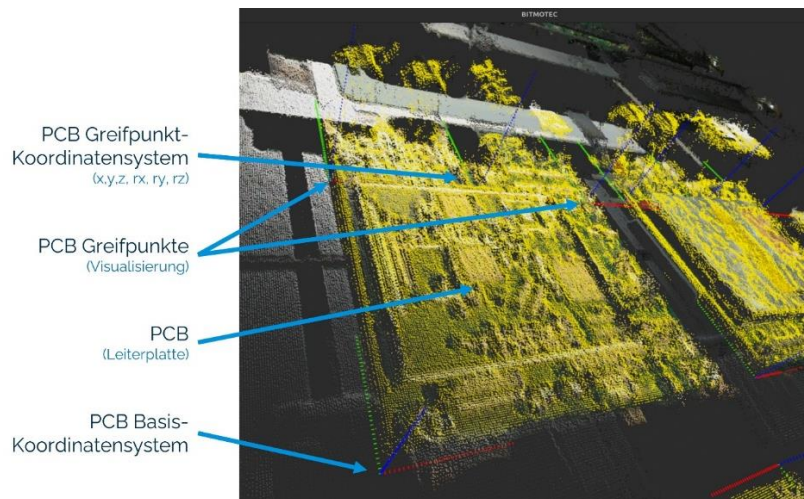


Abbildung 63: Detailansicht der automatischen 3D-PCB-Erkennung. Die erkannten PCBs sind in Gelb dargestellt. Die ermittelte Position des PCB-Greifpunkt-Koordinatensystems bildet sowohl die translatorischen- als auch die rotatorischen Transformationsinformationen der Leiterplatte ab

Fazit Arbeitspaket (AP) 9.1:

Während der Implementierungsphase des Demonstrators wurde - auf Basis einer Risikoanalyse und -bewertung sowie der Erstellung einer CE-Konformitätserklärung - die Sicherheit der Zelle nachgewiesen und sichergestellt (AS 9.1.7).

Abschließend wurde während des Probetriebs und der sich anschließenden Validierungsphase die korrekte Funktion der Applikation in Zusammenspiel mit dem Agenten geprüft. Die geforderte hohe Verfügbarkeit/Prozesssicherheit konnte nachgewiesen werden.

Arbeitspaket 9.2: Intelligente Leiterplattenprüfung

In AP 9.2, hat das Ziel die Identifikation von Pseudofehlern in der Leiterplattenprüfung mithilfe von KI-Modellen zu erkennen. Zu Beginn hat das Team aus Mitarbeitern und Mitarbeiterinnen der beteiligten Firmen die Zielsetzung für die intelligente Leiterplattenprüfung konkretisiert:

- Der Prüfplatz erkennt, ob Leiterplatten die im Test ausfallen, tatsächlich einen Fehler aufweisen oder ob es sich um ein Problem des Prüfplatzes handelt.
- Der Prüfplatz kann die mögliche Fehlerursache am Prüfplatz erkennen und darauf sinnvoll reagieren.

Zur Problemanalyse wurden in gemeinsamen Workshops in relevante Pseudofehler als Anwendungsfälle gesammelt und hinsichtlich der einsetzbaren KI-Methoden eingeordnet. Gemeinsam mit dem Verbundpartnern Rapidminer und slashwhy wurde zudem die Umsetzbarkeit aus Sicht der KI-Experten bewertet. Eine Zusammenfassung der Erkenntnisse ist Tabelle 6 zu entnehmen. Anwendungsfälle, die als machbar klassifiziert wurden und konkret einer KI-Methode zugeordnet werden konnten, wurden als relevante KI-Anwendungsfälle für das weitere Vorgehen im Projekt identifiziert. Aus Sicht von Sennheiser wurden die Anwendungsfälle daraufhin bewertet, wie oft sie üblicherweise in der Praxis vorkommen (siehe Tabelle 7).

Tabelle 6: Identifikation von relevanten Anwendungsfällen für KI

Nr.	Anwendungsfall	Einordnung in KI	Machbarkeit
1	Ausfälle bei Hochfrequenzmessung	Intelligente Zeitreihenanalyse	Ja
2	Zeitweise Ausfälle bei PLL Spannung	Intelligente Zeitreihenanalyse	Ja
3	Ausfall bei EM RF HF Messung	Intelligente Zeitreihenanalyse	Ja
4	Ausfall bei der RSSI Spannung	Intelligente Zeitreihenanalyse	Ja
5	Pausieren des Testprogramms nach Start	Kein KI-Problem	Nein
6	Fehlermeldung USB Comport	Kein KI-Problem	Ja
7	Runtime Error	Kein KI-Problem	Ja
8	Fehler bei RS485	Kein KI-Problem	Ja
9	Programmierung der Leiterplatte fällt aus	Kein KI-Problem	Ja

Tabelle 7: Häufigkeit von potentiellen KI Anwendungsfällen

Titel	Detailbeschreibung	Einschränkungen	Relevanz
Ausfälle bei der HF Messung	HF Nadel Massekontakte abgenutzt, HF Nadel ausgetauscht	Skywats Daten/ Weglaufende Messwerte	1x Monat
Alle Lp's fallen bei Step 10 Key Down aus	Nadel abgebrochen für TP 836	Aufeinander folgende Ausfälle	3-5x Monat
Ausfälle bei unterschiedlichen HF Messungen	PXI System HF Schaltkarten defekt (HF Relays schalten nicht richtig)	Skywats Daten/ Weglaufende Messwerte	1x Monat
Zeitweise Ausfälle bei PLL Spannung	VG Leiste Pin lose	Meist aufeinander folgende Ausfälle. (Es sind hier aber auch sporadisch	1x Monat
Unterschiedliche Ausfälle bei Spannungsmessungen und Pegel	Nadelbrett nicht richtig kontaktiert, Schrauben lose	Viele unterschiedliche Fehler, die meist aufeinander folgen.	10x Monat
Ausfall EM RF HF Messungen	Hf Stecker innen Pin abgenutzt.	Skywats Daten/ Weglaufende Messwerte	2x Monat
Ausfall bei der RSSI Spannung	HF Stecker vom Grundadapter zum Nadelbrett schief eingesteckt, Schraube vom HF Stecker oder Mutter vom Nadelbrett defekt .	Aufeinander folgende Ausfälle	3-3x Monat
Nach Teststart steht das Testprogramm	Es muss das aktuelle Testprogramm beendet werden und ein Testprogramm von einer anderen LP geladen werden und der erste Testschritt ausgeführt werden. Danach kann wieder das zu testende Programm geladen werden.	Bug im SEOI	5x Monat
Fehlermeldung USB Comport konnte nicht gefunden werden	Falsche Zuordnung der USB Ports im Gerätemanager	Fehlermeldung im SEOI	3-4x Monat
Runtime Error	Max Datenbank zerschossen /Gerät(Netzteil/Messgerät) nicht initialisiert	RTE log file zeigt Fehler	1x Monat
RS485 konnte nicht gefunden werden	Fehlende oder falsche Bezeichnung der Com Ports im Max	Fehlermeldung im SEOI	2-3x Monat
Runtime Error	PXI Systeme wurden nicht erkannt	NI MAX überprüfen, ob alle Geräte da sind	2x Monat
Alle Lp's fallen der Programierung aus	J-Link Adapter funktioniert nicht		1 2-3x Monat

Dabei zeigte sich eine große Vielfalt an Fehlerursachen, die oft nur wenige Male pro Monat auftreten und somit potentiell nur sehr wenige Trainingsdaten für KI-Modelle liefern können. Gemeinsam wurde beschlossen, sich zunächst auf die generelle Erkennung von Pseudofehlern im Leiterplattentest zu konzentrieren, ohne eine Berücksichtigung der konkreten Ursache.

Die KI am Prüfplatz hat die Aufgabe, auf Basis der Prüfdaten und ggf. weiterer Informationen eine passende Reaktion auszulösen. Bezogen auf den Anwendungsfall der grundsätzlichen Erkennung von Pseudofehlern bedeutet dies, dass die KI entscheiden muss, ob es sich bei einem in der Prüfung auf-

tretenden Fehler um einen Pseudofehler handelt oder einen tatsächlichen Fehler des Prüflings. Abhängig von dieser Entscheidung der KI erfolgen dann weitere Aktionen. Die Alternativlösung ohne KI wäre eine Entscheidung basierend auf festen Regeln. Um für das UAP 9.2.2 Kriterien für eine Bewertung der KI in einer Machbarkeitsuntersuchung zu finden, hat Sennheiser die im Folgenden beschriebene Vorgehensweise gewählt.

Die Wahrscheinlichkeit einen Fehler 1. oder 2. Art zu begehen, ist abhängig von der Auftretenswahrscheinlichkeit der beiden Hypothesen. Dies hat besonders gravierende Auswirkungen, wenn die Wahrscheinlichkeiten für das Eintreten einer der beiden Hypothesen sehr unterschiedlich sind. Wenn beispielsweise die Nullhypothese H_0 , dass es sich um einen Pseudofehler handelt, sehr selten zutrifft.

Um daraus ein Bewertungskriterium abzuleiten, ist auch die Alternative zu betrachten: In betrachteten Fall wäre die Alternative, dass es keine KI gibt, die eine Entscheidung fällt und somit dem Prüfplatz in seiner Entscheidung immer vertraut wird. Dies entspricht einer grundsätzlichen Entscheidung für die Alternativhypothese H_1 .

Ein Beispiel soll das illustrieren (vgl. Tabelle 8). In diesem Beispiel sei die Wahrscheinlichkeit für Hypothese H_0 (Pseudofehler) sehr gering mit 1%. Der Einfachheit halber sei die Wahrscheinlichkeit für eine richtige Entscheidung immer 95%. Dann würde die Entscheidung in 5,9% der Fälle für die Hypothese H_0 (Pseudofehler) getroffen werden und nur in 94,1% für H_1 (Fehler des Prüflings).

Tabelle 8: Beispiel Wahrscheinlichkeiten

$\alpha = 5\%$ $\beta = 5\%$		Wirklichkeit	
		Wahrscheinlichkeit von H_0 : 1%	Wahrscheinlichkeit von H_1 : 99%
Entscheidung ...	...mit einer Wahrscheinlichkeit von 5,9% für H_0	Wahrscheinlichkeit für eine richtige Entscheidung: $P(H_0) * (1 - \alpha)$ $= 1\% * 95\% = \mathbf{0,95\%}$	Wahrscheinlichkeit für Fehler 2. Art: $P(H_1) * \beta$ $= 99\% * 5\% = \mathbf{4,95\%}$
	...mit einer Wahrscheinlichkeit von 94,1% für H_1	Wahrscheinlichkeit für Fehler 1. Art: $P(H_0) * \alpha$ $= 1\% * 5\% = \mathbf{0,05\%}$	Wahrscheinlichkeit für eine richtige Entscheidung: $P(H_1) * (1 - \beta)$ $= 99\% * 95\% = \mathbf{94,05\%}$

Im alternativen Fall (keine KI) würde immer entschieden werden, dass es sich um einen Fehler des Prüflings handelt und die Entscheidung wäre dann in 99% der Fälle korrekt. Die Wahrscheinlichkeit für einen Fehler erster Art betrüge dann aber 1% und wäre damit erheblich größer als mit einer KI. Fehler 2. Art würden nicht auftreten. Es kommt bei der Bewertung ganz erheblich darauf an, wie wichtig die Verhinderung von Fehlern erster und zweiter Art ist. Übertragen auf den ersten Anwendungsfall sieht die Matrix (Tabelle 9) folgendermaßen aus:

Tabelle 9: Wahrscheinlichkeiten im ersten Anwendungsfall

		Wirklichkeit	
		Pseudofehler	Fehler des Prüflings
Entscheidung der KI...	...für Pseudofehler	Richtige Entscheidung mit folgenden möglichen Reaktionen:	Fehler 2. Art mit folgenden möglichen Konsequenzen: Fehlalarm beim zuständigen Mitarbeiter

		• Automatische Benachrichtigung des zuständigen Mitarbeiters • Automatische Ursachenanalyse • Ggf. Prüfplatz für die Dauer der Fehlerbehebung nicht verfügbar • Erneuter Test des Prüflings nachdem der Fehler behoben wurde	• Ggf. Prüfplatz nicht verfügbar • Erneuter Test des Prüflings mit vermutlich demselben Ergebnis
	...für Fehler des Prüflings	Fehler 1. Art mit folgenden möglichen Konsequenzen: • Ggf. unnötiger Reparaturversuch eines Technikers • Ggf. unnötiger Ausschuss	Richtige Entscheidung mit folgenden möglichen Reaktionen: • Weiterleitung zur Reparatur • Verschrottung der Leiterplatte

Die Wahrscheinlichkeit für einen Pseudofehler wurde für verschiedene Leiterplattentypen anhand der Reparaturdaten ermittelt. Die Kosten werden auch durch die ausgelöste Reaktion beeinflusst. Die KI könnte beispielsweise bei nur wenigen erkannten Pseudofehlern diese Fehler in Kauf nehmen und erst bei einer größeren Anzahl eine entsprechende Reaktion auslösen, aber auch hierfür werden die Kosten von Fehlern 1. und 2. Art berücksichtigt. Kosten für einen Fehler 1. Art:

- **Verschrottungskosten:** Sofern die Leiterplatte nicht repariert wird oder nicht repariert werden kann, fallen Verschrottungskosten an. Diese sind natürlich abhängig vom Wert der Leiterplatte. Hinzu kommen bei einem vergeblichen Reparaturversuch auch die Reparaturkosten.
- **Reparaturkosten:** Die Reparaturkosten sind im Wesentlichen abhängig vom benötigten Zeitaufwand. Der wiederum hängt von der Komplexität der Leiterplatte ab. Aus der Sennheiser Produktion gibt es eine Abschätzung für die Leiterplatten eines Testplatzes. Aus den ermittelten Werten wurde automatisch die Reparaturzeit aus den Stückkosten der Leiterplatte abgeschätzt (siehe Abbildung 64). Durch die Multiplikation mit dem Stundensatz entstehen daraus die Reparaturkosten. Der Stundensatz betrug zum Zeitpunkt der Ermittlung 46 €/h.

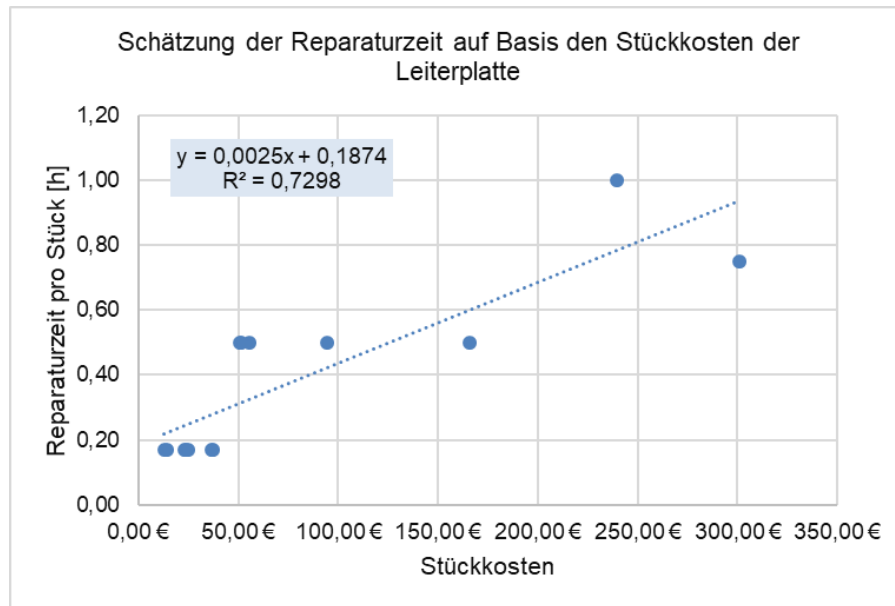


Abbildung 64: Zusammenhang zwischen den Kosten einer Leiterplatte und den abgeschätzten mittleren Reparaturzeiten

- **Pauschale Kosten für einen Fehler 1. Art:** Um aus beiden Kosten einen pauschalen Kostensatz für Fehler 1. Art zu machen, wäre zu klären gewesen, was tatsächlich bei einem Reparaturversuch an einer eigentlich funktionstüchtigen Leiterplatte passiert. Sofern nur festgestellt wird, dass die Leiterplatte in Ordnung ist, fallen nur die Reparaturkosten und die Kosten einer Wiederholprüfung an. Es könnte aber auch sein, dass die Reparatur gar nicht erst versucht wird und die Leiterplatte von vornherein verschrottet wird. Die Annahme für die weitere Berechnung war, dass beide Möglichkeiten gleich wahrscheinlich sind. Somit wurde ein Mittelwert gebildet.
- **Kosten für einen Fehler 2. Art:** Kosten durch Stillstand eines Prüfplatzes. Sofern Mitarbeiter anwesend sind, die nicht anderweitig beschäftigt werden können, fallen die Lohnkosten an. Dies sollte bei einem automatisierten Prüfplatz nicht maßgeblich sein. Sofern der Prüfplatz kein Engpass ist, fallen keine weiteren Kosten an.
- **Kosten einer Fehlersuche am Prüfplatz:** Ein pauschaler Zeitaufwand für Reparaturen an einem Prüfplatz ist von den Produktionstechnikern zusammen mit dem Prüfmittelservice auf 2 Stunden veranschlagt worden. Es fällt also der entsprechende Stundensatz für 2 Stunden an. Für den Prüfmittelservice galt ein Stundensatz von 55€/h.
- **Kosten einer Wiederholungsprüfung:** Während einer Wiederholungsprüfung ist der Prüfplatz nicht für andere Prüfungen verfügbar. Es fallen also die gleichen Kosten an, wie bei einem Stillstand des Prüfplatzes. Hinzu kommen Kosten für die Abnutzung, also der Maschinenstundensatz. Im konkreten Fall des Demonstrators wurde dieser auf 15 €/h (0,25 €/Minute) abgeschätzt. Um die Kosten einer Wiederholprüfung zu ermitteln, muss also auch die Prüfzeit bekannt sein.

Beispielrechnung

An einem Beispiel (Tabelle 10) sollen die obigen Überlegungen zusammengeführt werden unter der Annahme, dass die oben aufgeführten Konsequenzen sofort greifen. Für α und β werden 5% vorausgesetzt.

Tabelle 10: Beispielrechnung für Fehlerkosten und Wahrscheinlichkeiten

$\alpha = 5\%$ $\beta = 5\%$		Wirklichkeit	
		Pseudofehler Wahrscheinlichkeit: 97%	Fehler des Prüflings Wahrscheinlichkeit: 3%
Entscheidung der KI...	...für Pseudofehler Wahrscheinlichkeit: 92,3%	Richtige Entscheidung mit folgenden möglichen Reaktionen: - Automatische Benachrichtigung des zuständigen Mitarbeiters - Automatische Ursachenanalyse - Ggf. Prüfplatz für die Dauer der Fehlerbehebung nicht verfügbar - Erneuter Test des Prüflings nachdem der Fehler behoben wurde Wahrscheinlichkeit: $97\% * 95\% = 92,15\%$	Fehler 2. Art mit folgenden möglichen Konsequenzen: - Fehlalarm beim zuständigen Mitarbeiter - Ggf. Prüfplatz nicht verfügbar - Erneuter Test des Prüflings mit vermutlich demselben Ergebnis Wahrscheinlichkeit: $3\% * 5\% = 0,15\%$ Kosten: Fehlersuche Prüfplatz: 110 € Wiederholprüfung: 0,25 €
	...für Fehler des Prüflings Wahrscheinlichkeit: 7,7%	Fehler 1. Art mit folgenden möglichen Konsequenzen: - Ggf. unnötiger Reparaturversuch eines Technikers - Ggf. unnötiger Ausschuss Wahrscheinlichkeit: $97\% * 5\% = 4,85\%$ Kosten: Reparaturversuch + Wiederholprüfung: $7,82 € + 0,25 € = 8,07 €$ Verschrottung: 23,03 € Mittelwert: 15,55 €	Richtige Entscheidung mit folgenden möglichen Reaktionen: - Weiterleitung zur Reparatur - Verschrottung der Leiterplatte Wahrscheinlichkeit: $3\% * 95\% = 2,85\%$

Die Gesamtkosten der Fehlentscheidungen wären dann:

$$0,15\% * 110,25 € + 4,85\% * 15,55 € = 0,92 €$$

Im Vergleich dazu die Kosten ohne eine KI, also nur der Fehler 1. Art ($\alpha = 100\%$ und $\beta = 0\%$):

$$97\% * 100\% * 15,55 € = 15,08 €$$

In diesem Fall wäre ein System mit KI klar überlegen. Für die Machbarkeitsuntersuchung war die Frage zu beantworten, wie groß α und β maximal sein dürfen. Dafür wurden die Ergebnisse verallgemeinert.

Als Ziel hat Sennheiser festgesetzt, dass die Kosten mit Einsatz einer KI die Kosten ohne KI mindestens halbieren sollen:

$$\frac{K_{KI}}{K_{oKI}} \leq 0,5$$

Die Kosten der Fehlentscheidungen mit KI betragen:

$$K_{KI} = (1 - P_{PF}) \cdot \beta \cdot K_{2A} + P_{PF} \cdot \alpha \cdot K_{1A}$$

Die Kosten der Fehlentscheidungen ohne KI betragen:

$$K_{oKI} = P_{PF} \cdot 100\% \cdot K_{1A}$$

Das Ziel war also:

$$\frac{(1 - P_{PF}) \cdot \beta \cdot K_{2A} + P_{PF} \cdot \alpha \cdot K_{1A}}{P_{PF} \cdot K_{1A}} = \frac{(1 - P_{PF})}{P_{PF}} \cdot \frac{K_{2A}}{K_{1A}} \cdot \beta + \alpha \leq 1 - Z$$

Dabei gilt: P_{PF} steht für die Pseudofehlerrate, K_{1A} für Kosten eines Fehlers erster Art, K_{2A} für Kosten eines Fehlers zweiter Art, K_{KI} für Kosten der Fehlentscheidungen mit KI, K_{oKI} für Kosten der Fehlentscheidungen ohne KI. Im Fall Pseudofehlerrate von Null ist, bietet die KI keinerlei Vorteile. Diese Leiterplatten können aus der Betrachtung entfallen.

Die Kosten eines Fehlers 1. Art betragen:

$$K_{1A} = P_V \cdot K_V + (1 - P_V) \cdot (K_R + K_W)$$

Dabei sind K_R Reparaturkosten, K_V Verschrottungskosten, K_W Kosten einer Wiederholprüfung, P_V Wahrscheinlichkeit für eine direkte Verschrottung. Die Kosten eines Fehlers 2. Art betragen mit K_F für Fehlersuchkosten:

$$K_{2A} = K_F + K_W$$

Daraus folgt:

$$\frac{(1 - P_{PF})}{P_{PF}} \cdot \frac{K_{2A}}{K_{1A}} \cdot \beta + \alpha = \frac{(1 - P_{PF})}{P_{PF}} \cdot \frac{K_F + K_W}{P_V \cdot K_V + (1 - P_V) \cdot (K_R + K_W)} \cdot \beta + \alpha \leq 0,5$$

Der Faktor mit dem β multipliziert wird, ist abhängig vom Prüfling und wurde für eine Reihe von Leiterplatten berechnet. Das Ergebnis für das Verhältnis der Kosten mit Einsatz einer KI zu den Kosten ohne KI für die Leiterplatte mit dem kleinsten Faktor sieht folgendermaßen aus (Tabelle 11). Hier führen auch schon kleine Trefferquoten der KI zu einer deutlichen Kosteneinsparung.

Tabelle 11: Verhältnis der Kosten mit und ohne Einsatz von KI am Beispiel 576205

Leiterplatte:	576205											
Faktor:	0,22											
α / β	0%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	
0%	-	0,01	0,02	0,03	0,04	0,06	0,07	0,08	0,09	0,10	0,11	
5%	0,05	0,06	0,07	0,08	0,09	0,11	0,12	0,13	0,14	0,15	0,16	
10%	0,10	0,11	0,12	0,13	0,14	0,16	0,17	0,18	0,19	0,20	0,21	
15%	0,15	0,16	0,17	0,18	0,19	0,21	0,22	0,23	0,24	0,25	0,26	
20%	0,20	0,21	0,22	0,23	0,24	0,26	0,27	0,28	0,29	0,30	0,31	
25%	0,25	0,26	0,27	0,28	0,29	0,31	0,32	0,33	0,34	0,35	0,36	
30%	0,30	0,31	0,32	0,33	0,34	0,36	0,37	0,38	0,39	0,40	0,41	
35%	0,35	0,36	0,37	0,38	0,39	0,41	0,42	0,43	0,44	0,45	0,46	
40%	0,40	0,41	0,42	0,43	0,44	0,46	0,47	0,48	0,49	0,50	0,51	
45%	0,45	0,46	0,47	0,48	0,49	0,51	0,52	0,53	0,54	0,55	0,56	
50%	0,50	0,51	0,52	0,53	0,54	0,56	0,57	0,58	0,59	0,60	0,61	

Ganz anders sieht die Situation bei einer Leiterplatte mit einem sehr großen Faktor aus (man beachte die Änderung der Skalierung in Tabelle 12):

Tabelle 12: Verhältnis der Kosten mit und ohne Einsatz von KI am Beispiel 583888

Leiterplatte:	583888											
Faktor:	5,34											
α / β	0%	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%	
0%	-	0,05	0,11	0,16	0,21	0,27	0,32	0,37	0,43	0,48	0,53	
5%	0,05	0,10	0,16	0,21	0,26	0,32	0,37	0,42	0,48	0,53	0,58	
10%	0,10	0,15	0,21	0,26	0,31	0,37	0,42	0,47	0,53	0,58	0,63	
15%	0,15	0,20	0,26	0,31	0,36	0,42	0,47	0,52	0,58	0,63	0,68	
20%	0,20	0,25	0,31	0,36	0,41	0,47	0,52	0,57	0,63	0,68	0,73	
25%	0,25	0,30	0,36	0,41	0,46	0,52	0,57	0,62	0,68	0,73	0,78	
30%	0,30	0,35	0,41	0,46	0,51	0,57	0,62	0,67	0,73	0,78	0,83	
35%	0,35	0,40	0,46	0,51	0,56	0,62	0,67	0,72	0,78	0,83	0,88	
40%	0,40	0,45	0,51	0,56	0,61	0,67	0,72	0,77	0,83	0,88	0,93	
45%	0,45	0,50	0,56	0,61	0,66	0,72	0,77	0,82	0,88	0,93	0,98	
50%	0,50	0,55	0,61	0,66	0,71	0,77	0,82	0,87	0,93	0,98	1,03	

Hier sind Fehler 2. Art, also die Klassifizierung von Prüflingsfehlern als Pseudofehler mit einer darauf folgenden Fehlersuche am Prüfplatz, möglichst zu vermeiden. Dies kann auch durch entsprechende Reaktionen auf einen Pseudofehler erreicht werden. Die beiden Beispiele zeigen deutlich, dass die Anforderungen an die zu realisierende KI abhängig vom jeweiligen Prüfling sind.

Prüflinge, die mit ihren Messwerten nahe an den Toleranzgrenzen liegen, werden in der Praxis häufig durch eine Wiederholprüfung „gut“ getestet. Bei Betrachtung der Kosten für eine einzelne Wiederholprüfung ist dieses wirtschaftlich auch sinnvoll, so dass bei Sennheiser daraus in der Vergangenheit die Regel abgeleitet wurde, ausgefallene Prüfungen grundsätzlich einmal zu wiederholen. Auch beim Trainieren eines KI-Modells mit dem Ziel der Kostenminimierung ergibt sich schnell, dass die KI immer einen Pseudofehler vorhersagt, der sich durch Wiederholprüfung beheben lässt, weil dieses die geringsten Kosten verursacht. Ziel für das Training durfte deshalb nicht die Kostenminimierung sein, sondern die Genauigkeit der Vorhersage. Aufgrund der unsymmetrischen Verteilung der Daten, hat sich das Team entschlossen Cohen's Kappa und die Confusion Matrix (siehe Abbildung 65) für eine Bewertung der Modellergebnisse heranzuziehen.

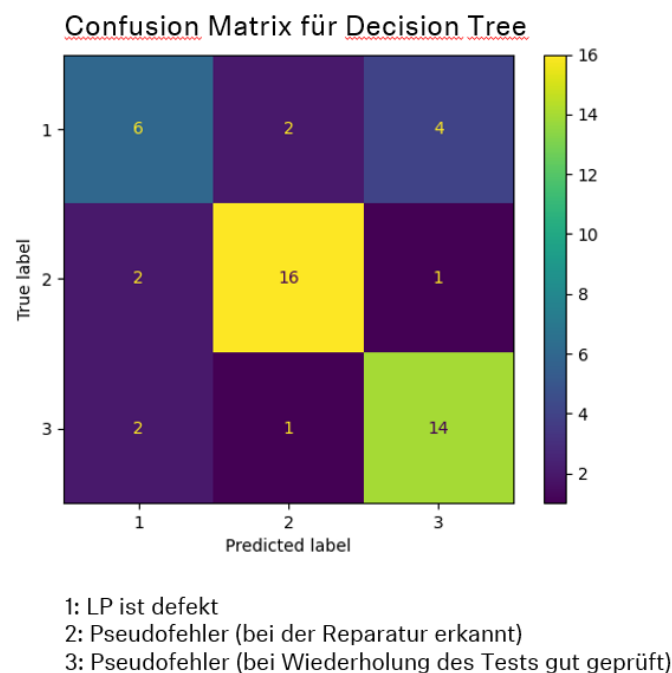


Abbildung 65: Confusion Matrix am Beispiel eines Entscheidungsbaums (Decision Tree Modell)

Darauf aufbauend wurde eine Machbarkeitsstudie für diesen Anwendungsfall mit verschiedenen KI-Methoden zur intelligenten Zeitreihenanalyse (allg. Klassifizierung von Pseudofehlern) durchgeführt. Da im Rahmen der Machbarkeitsanalyse keine vollumfängliche Studie über alle Leiterplattenvarianten hinweg durchgeführt werden konnte, wurde sich zunächst auf solche Varianten konzentriert, die eine besonders hohe Pseudofehlerrate aufgewiesen haben bei ausreichender produzierter Stückzahl. Durch dieses Vorgehen wurden die Varianten um das Modell „HORIZON-WIN LED“ und „HORIZON-WIN MID“ als besonders geeignet identifiziert.

Zum Abschluss der Machbarkeitsstudie wurden aufbauend auf dieser Analyse die KI-Methoden der Entscheidungsbäume, des Bagging-Verfahrens, des Random Forest, des Gradient Boosting, des AdaBoost, des Extra Trees, der Support-Vektor-Machine und der sog. Flood-Filling-Networks (FFN) auf historische, gelabelte Prüfdaten aus der Sennheiser Datenbank als Trainings- und Testdaten aufgebaut und trainiert. Die Ergebnisse dieser ersten Machbarkeitsstudie können aus Tabelle 13 entnommen werden.

Tabelle 13: Ergebnisse der Machbarkeitsstudie

KI-Methode	Genauigkeitsniveau
Entscheidungsbäume	75 Prozent
Bagging-Verfahren	83 Prozent
Random Forest	79 Prozent
Gradient Boosting	77 Prozent
AdaBoost	60 Prozent
Extra Trees	81 Prozent
Support-Vektor-Machine	62 Prozent
FFN	83 Prozent

Neben der Machbarkeitsstudie wurde durch das IFW im Rahmen der Erstellung eines Anforderungskatalogs die Entwicklung eines Integrationskonzepts zur autonomen Bewertung von Pseudofehlern auf Basis einer funktionierenden, KI-basierten Fehlererkennung und -klassifikation unterstützt. In diesem Konzept (Abbildung 66) wurden dazu umzusetzende Dienste erarbeitet und definiert, die bei einem Fehler innerhalb der Leiterplattenprüfung angesprochen werden. Neben der Erkennung und Klassifikation von Fehlern innerhalb der *Pseudofehlerbewertung*, wurde das Konzept eines Maßnahmenservice entwickelt, der eine intelligente Reaktion auf Pseudofehler durchführen kann. Dazu werden dem Maßnahmenservice Daten aus Simulationsstudien bereitgestellt, die näher in AP 9.4 erörtert werden.

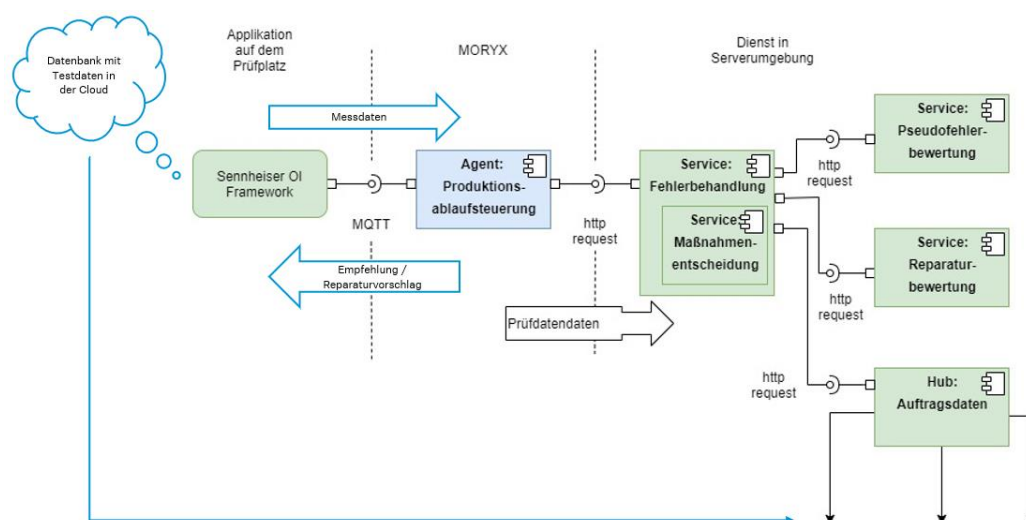


Abbildung 66: Integration der Pseudofehlererkennung in den Demonstrator

Gemeinsam mit den Projektpartnern Sennheiser, slashwhy, Rapidminer und IFW wurden im UAP 9.2.3 die Anforderungen an die umzusetzenden KI Anwendungsfälle formuliert. Neben den oben schon beschriebenen Genauigkeitsanforderungen waren hier auch die zu nutzende Entwicklungsumgebung, die Implementierung in Docker Containern und die Abstimmung der Schnittstellen zwischen KI Services und der bestehenden Software Infrastruktur relevant. Dies betraf nicht nur die zu schaffenden neuen Komponenten, sondern auch bestehende Systeme, die entsprechend erweitert werden mussten. Darüber hinaus war es wichtig, eine Möglichkeit zu schaffen, das Zusammenspiel der Komponenten und

die Kommunikation mit den KI Services zu überwachen und zu protokollieren. Zu einem späteren Zeitpunkt kam als ergänzende Anforderung hinzu, dass die KI-Modelle eine Konfidenz mitliefern. Diese Anforderung wurde notwendig, um dem später ergänzten Maßnahmenservice einen Anhaltspunkt zu liefern, wie vertrauenswürdig die Aussage des KI-Modells ist. Diese zusätzliche Information hat sich im Nachhinein auch bei der Überwachung der KI-Modelle bewährt (siehe UAP 9.2.14).

Sennheiser hatte für das Training der KI-Modelle im Rahmen der Machbarkeitsstudie Daten aus dem Cloud basierten Testdatenmanagement extrahiert und bereitgestellt. Für den Betrieb der KI Modelle in der realen Produktionsumgebung, wurde im UAP 9.2.4 definiert, die notwendigen Testdaten über die entsprechenden Schnittstellen, den KI Modellen zur Verfügung zu stellen.

Im UAP 9.2.6 war eine wesentliche Herausforderung im Umgang mit den Realdaten die Tatsache, dass Prüfabläufe und auch Toleranzen insbesondere beim Neuanlauf eines Produktes in der Produktion oft noch optimiert und angepasst werden. Dadurch ändern sich z. B. Bezeichnungen von Testschritten, Reihenfolgen von Testschritten und durch die Anpassung von Toleranzen möglicherweise auch die Entscheidung über bestanden oder nicht bestanden bei einem Testschritt. Die Lösung des Problems bestand in einer Eingangsprüfung für die Daten beim KI-Service, so dass solche Veränderungen bemerkt werden konnten und entsprechende Meldungen in den Logdateien protokolliert wurden.

Das Thema Datennormalisierung wurde den KI Services selbst zugeordnet, da nicht alle Modellarten tatsächlich eine Normalisierung benötigen. Für den Fall, dass sich die Datenschnittstelle für die KI Services nicht standardisieren ließe, wurde eine Verwaltungsschale für die KI-Modelle angedacht. Diese könnte die Spezifikation für die zu liefernden Daten und deren Format enthalten. Letztendlich wurden die Daten zwischen den verschiedenen Services im JSON Format übermittelt.

Die KI-Modelle wurden von slashwhy im UAP 9.2.7 erstellt und gemeinsam anhand der zuvor erstellten Kriterien bewertet. Da sich der Demonstrator selbst zu diesem Zeitpunkt noch im Bau befand, haben wir uns entschlossen, zwei in der Produktion im Einsatz befindliche manuelle Prüfplätze zur Erprobung der ersten KI-Modelle zu nutzen. Für sechs Leiterplatten lagen initial ausreichend Daten für ein Modelltraining vor. Sennheiser hat durch verstärkte Reparaturen dafür gesorgt, im Laufe des Projekts die Datenbasis entsprechend zu erweitern, so dass auch noch weitere Leiterplatten später ergänzt werden konnten.

Für die Offline-KI-Evaluation im UAP 9.2.8 wurde von slashwhy eine Matrix in Excel aufgebaut. Sennheiser hat auf Basis dieser Matrix eine Power BI Auswertung zur Visualisierung der Ergebnisse erstellt (siehe Abbildung 67 und Abbildung 68).

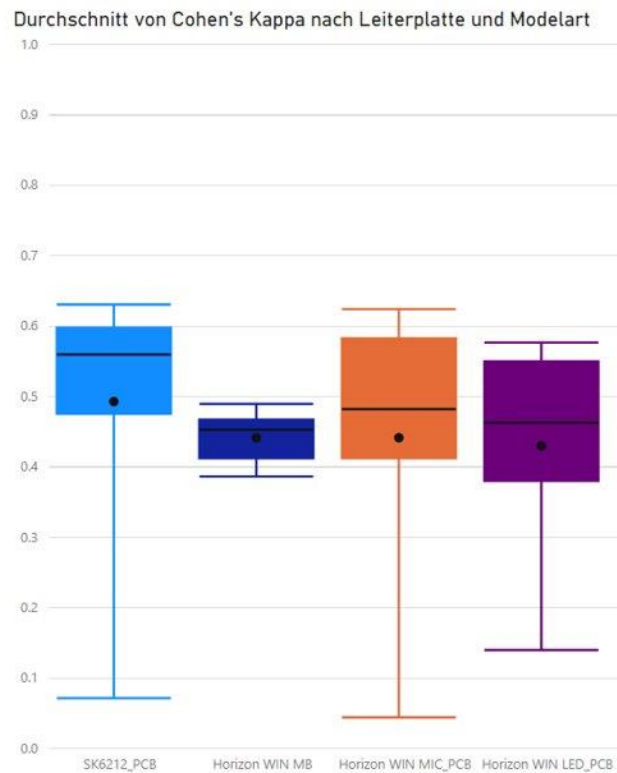


Abbildung 67: Cohen's Kappa Werte für unterschiedliche Leiterplatten (Auswertung in Power BI)

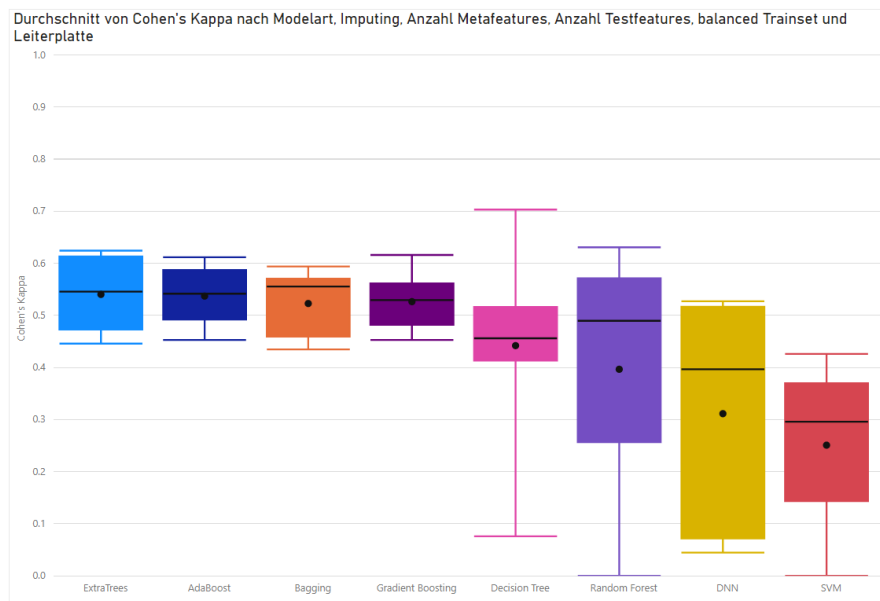


Abbildung 68: Cohen's Kappa Werte für unterschiedliche Modelltypen (Auswertung in Power BI)

Die Nutzung mehrerer Typen an Leiterplatten hat sich als sehr vorteilhaft erwiesen, da damit auch schnell deutlich wurde, dass eine Festlegung auf eine Modellart nicht hilfreich ist, sondern pro Leiterplattentyp eine Modellart ausgewählt werden muss. slashwhy konnte so auch an verschiedenen Leiterplattentypen Optimierungen der Modelle erproben.

Insgesamt zeigte sich oft ein recht niedriges Cohen's Kappa. Zum Vergleich wurde deshalb die Ersparnis ebenfalls berechnet. Dabei zeigte sich, dass der Zusammenhang zwischen Cohen's Kappa abhängig von der Art des Modelles und der Leiterplatte ist, so dass sich keine feste Umrechnung zwischen beiden Kennzahlen ergibt (siehe Abbildung 69). Zudem zeigte sich im Verlauf des Einsatzes der KI-Modelle, dass die Performance nach einer gewissen Zeit oft abnahm. Ursache dafür waren die schon bekannten Optimierungen der Testsequenzen, aber auch Änderungen bei Bauteillieferanten und damit verbundene Redesigns der Leiterplatten.

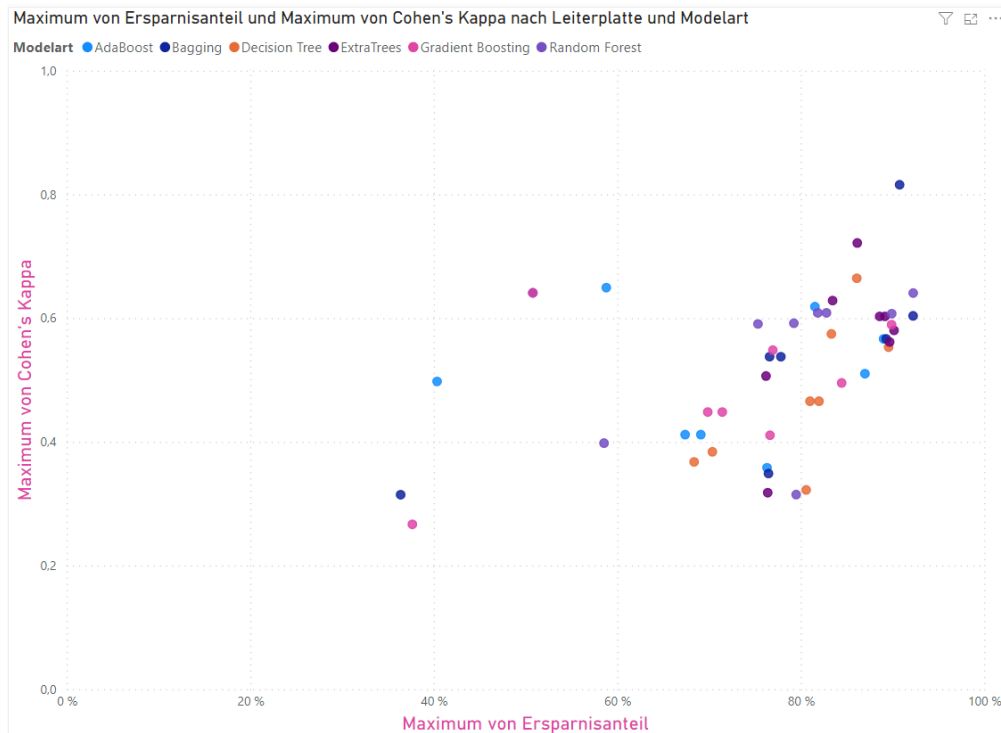


Abbildung 69: Zusammenhang zwischen Cohen's Kappa und Ersparnis (Auswertung in Power BI)

Für die Implementierung der geschaffenen KI Services und des Maßnahmenservices auf dem Demonstrator im UAP 9.2.9 musste im nächsten Schritt die Kommunikation zwischen der Teststation und den Services umgestellt werden von einer direkten Kommunikation auf eine Kommunikation über MORYX. An den verwendeten Live-Daten für die KI Services änderte sich jedoch nichts gegenüber der Implementierung auf einem manuellen Prüfplatz.

Im UAP 9.2.10 war die Realisierung der Randbedingungen für die Integration der KI Modelle vorgesehen. Auf dem Demonstrator war die Prüfung von Leiterplatten geplant, die sich zum Zeitpunkt der Erstellung in der Produktentwicklung befanden. Die bislang betrachteten Leiterplatten kamen dort nicht zum Einsatz, so dass komplett neue Modelle erstellt werden mussten. Für die neuen Leiterplatten waren jedoch zu dem Zeitpunkt nur sehr wenige Testdaten vorhanden und Reparaturdaten aus der Produktion waren bei Inbetriebnahme des Demonstrators noch nicht verfügbar. In dieser Situation hat sich der Maßnahmenservice bewährt, der in der Lage war, auch ohne KI-Modelle allein auf Basis der implementierten Regeln, sinnvolle Entscheidungen zu treffen.

Bei der Integration der neuen digitalen Services mit den neuen KI-Modellen in den Demonstrator im UAP 9.2.11 war auf Seiten von Sennheiser zu beachten, dass die Anzahl der Prüfplätze, die diese Services nutzen, sich nun verdoppelte, da zusätzlich noch ein weiterer manueller Prüfplatz für die neuen Produkte entstand. Jedoch war dieses ohne Schwierigkeiten möglich.

Im AS 9.2.12 war die Bereitstellung der Ergebnisse der KI-Modelle für die Nutzer dieser Information vorgesehen. Bei der ersten Implementierung der KI-Services auf den manuellen Prüfplätzen waren die Nutzer die Personen, die den Prüfplatz bedient haben, die also Leiterplatten in den Prüfadapter eingelegt haben, den Adapter geschlossen haben und die Prüfung gestartet haben. Diese Personen haben parallel zur Prüfung der Leiterplatten weitere Montageaufgaben ausgeführt. Die Ergebnisse des KI-Modells wurden nach Abschluss des Tests in Form eines Popups angezeigt. Das Popup konnte mit einer Bewertung des Ergebnisses wieder geschlossen werden.

Im weiteren Verlauf des Projektes wurde der angedachte Maßnahmenservice implementiert, so dass die Ergebnisse der KI-Modelle im JSON Format an den Maßnahmenservice übermittelt wurden. Ziel war es, der Produktionssteuerung MORYX mitzuteilen, was mit einer ausgefallenen Leiterplatte zu geschehen hat. Dieses erfolgte dann vom Maßnahmenservice auch wieder im JSON Format an die Produktionssteuerung.

Bereits bei der Offline Evaluation der KI-Modelle hatte sich gezeigt, dass eine Überwachung der KI Modelle notwendig ist. Die Data Scientists und Sennheiser als Betreiber des Demonstrators benötigten also die Möglichkeit, die Ergebnisse der KI-Modelle und des Maßnahmenservice zu dokumentieren. Dafür wurden die Logdateien verwendet, die von Sennheiser über eine weitere Power BI Auswertung analysiert wurden. Dies war für die Arbeit im Projektteam ausreichend. Für eine automatisierte Auswertung insbesondere auf Seiten der Data Scientists reichte dies nicht und so wurde die bislang schon existierende Erfassung der Testdaten in einer Cloud basierten Datenbank von Sennheiser um eine Erfassung der Ergebnisse der digitalen Services erweitert (siehe Abbildung 70).



Teststatus: Fehlerhaft

Seriennummer:
Teile-Nr.:
Revision:
Name des Produkts:
Auftragsnummer:
Start Datum/Uhrzeit:
UTC Start-datum/zeit:

KG00101995432059673383QA310530002
596733
0
LP.BEST EWD+_EM95_RF 606-694MHz (HW3)
01995432
2023-Feb-07 15:11:59
2023-Feb-07 14:11:59

Operator:
Ausführungszeit:
Testoperation:

049SYS40I_TC15
82.436 seconds (1 m, 22 s)
PCBA test (code:50)

► Stationsdaten

▼ Misc UUT Info (12)

Beschreibung
PCB_EW-DX_TB.seq
COMMON.CSV
596733_EW-DX_RF_606_694_MHZ.csv
CostCentre
Specific test time
Test mode
Type of Test
ATS TE Number
Service Decision
Defect Component
Pseudo error classification
Pseudo error certainty

Zeichenfolgen-Variablen
4.789.0
2022.11.01T07:05:22+01:00
2022.11.07T08:01:38+01:00
105400999
2023.02.07T15:13:23+01:00
Auto
Regular
TE000761
Reconnect
Unknown
Retest
0.451862752

Numerische Variable
0
0
0
0
0
0
0
0
0
0
0
0

Abbildung 70: Dokumentation der Ergebnisse der KI Modelle im Testdatenmanagementsystem

Diese Dokumentation war auch für die Verifikation der KI-Modelle im AS 9.2.13 erforderlich. Hierzu waren in einer frühen Phase des Projektes die Rückmeldungen über das Popup an den manuellen

Prüfplätzen (siehe UAP 9.2.12) ausgewertet worden. Dabei stellte sich die Frage, inwieweit diese Rückmeldung tatsächlich als Referenz dienen konnte, denn es zeigte sich eine starke Abhängigkeit von der Entscheidung des KI Service, des Prüfplatzes, der Schicht und des Wochentags (siehe Abbildung 71, Abbildung 72 und Abbildung 73)

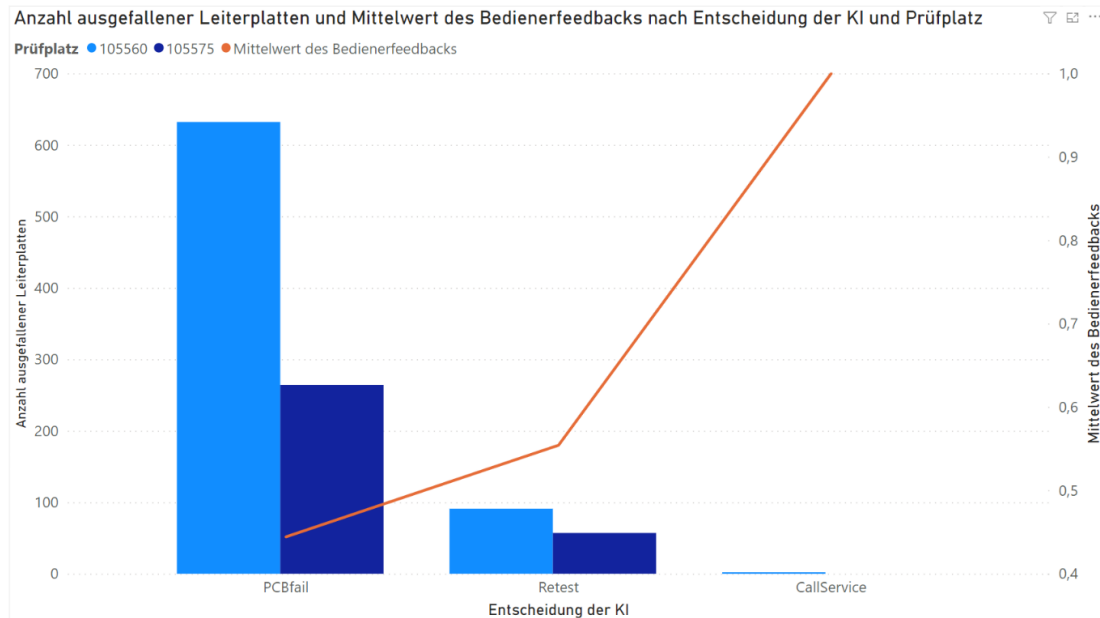


Abbildung 71: Mittelwert des Feedbacks in Abhängigkeit von der Serviceentscheidung (1 = zufrieden mit der Entscheidung der KI, 0 = nicht zufrieden mit der Entscheidung der KI)

Mittelwert des Bedienerfeedbacks nach Prüfplatz und Entscheidung der KI

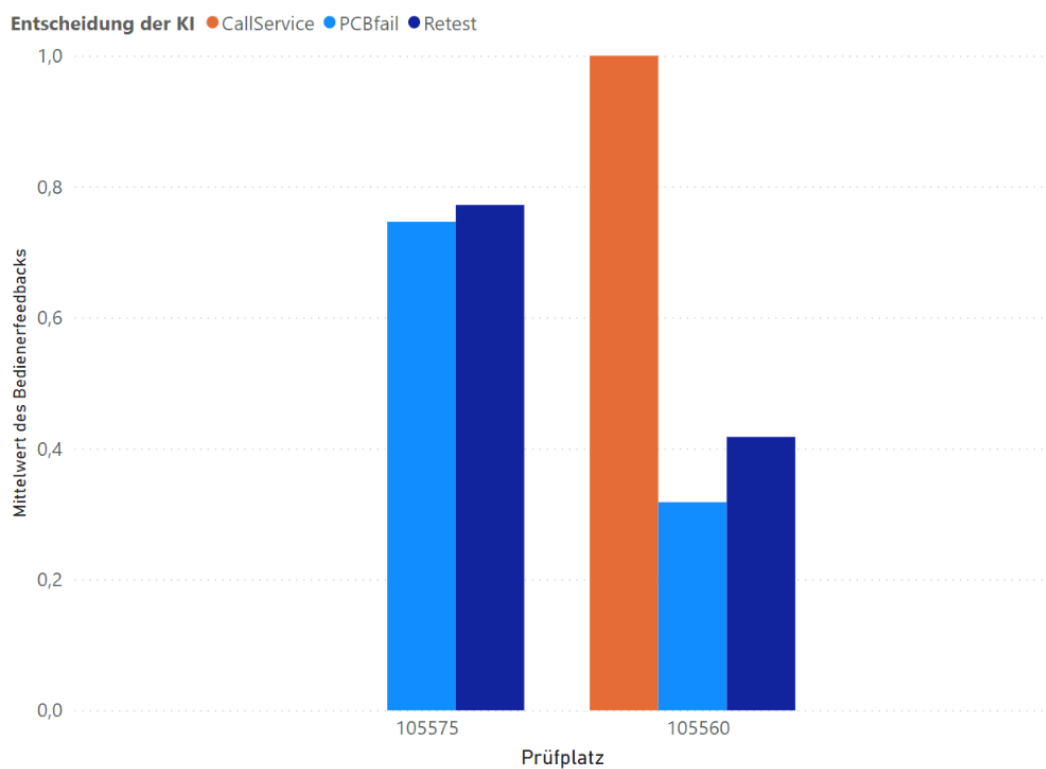


Abbildung 72: Mittelwert des Feedbacks in Abhängigkeit vom Prüfplatz

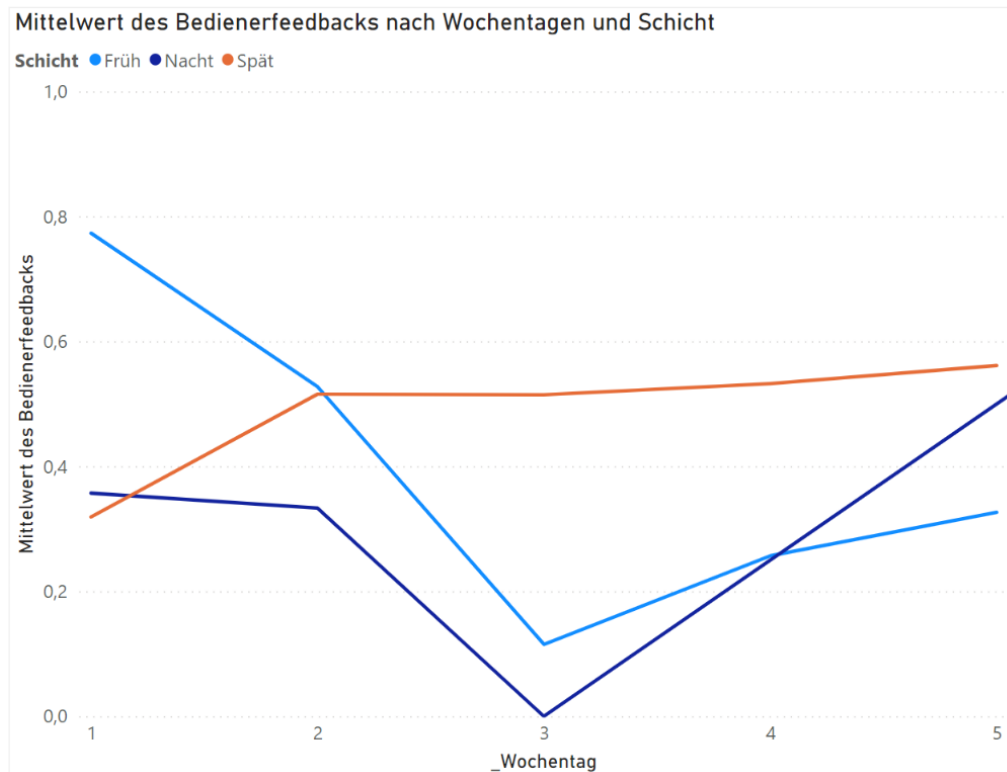


Abbildung 73: Mittelwert des Feedbacks in Abhängigkeit vom Wochentag (1 = Montag)

Idealerweise wäre für eine Verifikation der KI-Modelle die wahre Ursache für einen Ausfall einer Leiterplatte im Test bekannt. Dies ist jedoch aus mehreren Gründen nicht der Fall:

- Nicht alle Prüflinge werden repariert, einige werden auch verschrottet oder eingelagert und zu einem deutlich späteren Zeitpunkt repariert.
- Trotz Vorgabe kommt es vor, dass Reparaturen nicht dokumentiert sind.
- Nicht alle Serviceeinsätze an den Prüfplätzen sind dokumentiert.
- Auch wenn eine Wiederholprüfung zum Bestehen des Prüflings geführt hat, kann insbesondere an einem manuellen Prüfplatz nicht ausgeschlossen werden, dass Änderungen am Prüfling oder am Prüfplatz vorgenommen wurden.

Für die Validierung der KI-Modelle kam im UAP 9.2.14 wieder wie zu Beginn in der Machbarkeitsanalyse (siehe UAP 9.2.2) der Kostenaspekt hinzu. Diesmal jedoch nicht nur der Aspekt der Kosten von Fehlentscheidungen, sondern auch die notwendigen Kosten der Erstellung und Wartung der KI-Modelle. Aus den Erfahrungen, die das Projektteam gesammelt hatte, wurde deutlich, dass die KI-Modelle in der Performance im Laufe der Zeit nachlassen. Da eine belastbare Verifikation nicht möglich war, wurde stattdessen die Konfidenz der Modelle überwacht. Dabei zeigte sich, dass diese im Laufe der Zeit abnimmt und sich durch ein erneutes Training wieder erhöhen lässt (siehe Abbildung 74).

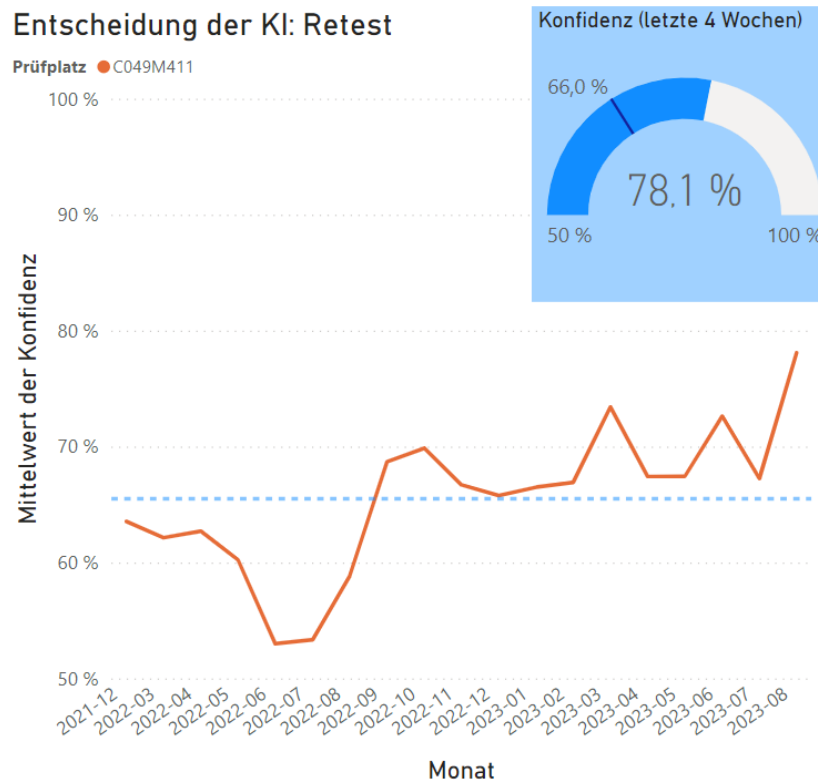


Abbildung 74: Beispiel für eine Überwachung eines KI Modells anhand der Konfidenz.

Daraus resultierte die Notwendigkeit, die KI-Modelle automatisiert zu trainieren und in die Anwendung zu bringen, um den zeitlichen Aufwand für die Wartung und die damit verbundenen Kosten zu minimieren. Zwei Vorgehensweisen sind möglich:

1. Ein regelmäßiges erneutes Training
2. Eine Überwachung der Modelle und bei Bedarf ein erneutes Training

Das Projektteam hat sich für die Überwachung der Modelle anhand der Konfidenz entschieden, um bei der potenziell sehr großen Anzahl von Modellen den Trainingsaufwand zu reduzieren und um andererseits bei Abweichungen auch schnell reagieren zu können. Idealerweise sollte auch das Monitoring und die Initiierung eines erneuten Trainings automatisiert erfolgen.

Die im UAP 9.2.15 eingeplante iterative Verbesserung der KI Modelle wurden von slashwhy und Rapidminer übernommen. Dabei wurden u.a. folgende Ansätze betrachtet:

- a. Übertragbarkeit der Modelle (Redesigns von Leiterplatten führten dazu, dass KI Modelle komplett neu erstellt werden mussten statt, die bislang trainierten Modelle weiterverwenden zu können)
- b. Normalisierung
- c. Umgang mit Toleranzänderungen in Prüfprogrammen

Arbeitspaket 9.3: Systemintegration

In UAP 9.3.1 wurde mit der Definition des Prototyps und dessen Systemgrenze gestartet. Dieser Prototyp ist Abbildung 53 zu entnehmen und wurde im gemeinsamen Kick-Off-Workshop unter Einbindung aller Verbundpartner erarbeitet. Als Systemgrenze wurde in diesem Zuge die autonome Roboterprüfzelle festgelegt, in der die Komponenten Test- bzw. Prüfmodul, Materialbereitstellung, -abfuhr und -

transport, Robotersystem und Leiterplattenidentifikation als Subkomponenten ermittelt werden konnten. Des Weiteren wurde festgelegt, dass die autonome Roboterprüfzelle inklusive der Subkomponenten über die Verwaltungsschale (UAP 9.3.3) als digitaler Zwilling abzubilden sind.

In UAP 9.3.2 wurde darauf aufbauend das Gesamtprozessdesign konzipiert. In gemeinsamen Workshops mit allen Verbundpartnern wurde dazu zunächst das gesamte Projektmanagement in *Azure DevOps* überführt und relevante User Stories für das AP 9 inklusive aller geplanten Unterarbeitspakete formuliert. Darüber hinaus wurde als zentraler Bestandteil des Gesamtprozesses ein Konzept zur Integration von autonomen Steuerungsagenten für die autonome Roboterzelle erarbeitet (siehe Abbildung 75). Der in diesem Projekt konzipierte Demonstrator bzw. alle zukünftig folgenden Roboterzellen verfügen in diesem Zusammenhang über einen eigenen übergeordneten Steuerungsagenten. Dieser Steuerungsagent agiert als zentrale Prozesssteuerung innerhalb der Roboterzelle und kennt vor diesem Hintergrund die gesamte Prozesslogik. Zusätzlich kommuniziert dieser Agent mit allen relevanten Komponenten (sog. Assets in Abbildung 75). Auch übernimmt dieser die zentrale Kommunikation mit dem Produktionsleitsystem. Von einer uneingeschränkt dezentralen Kommunikation wurde aus Gründen der Komplexität abgesehen.

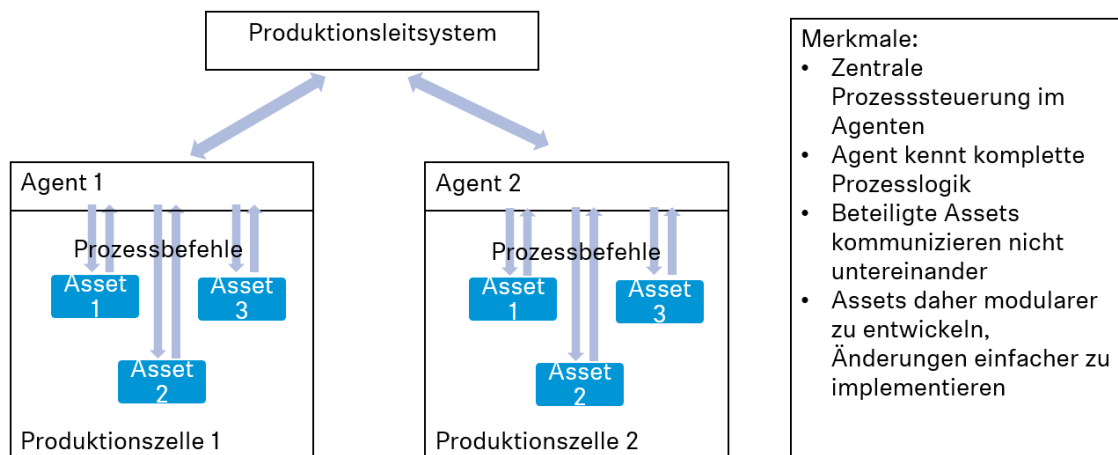


Abbildung 75: Nutzung von Steuerungsagenten im Gesamtprozessdesign

Aufbauend auf dem skizzierten Gesamtprozessdesign wurde in einem weiteren UAP der digitale Zwilling des Demonstrators konzipiert. Dazu wurde sich an der Technologie der Verwaltungsschale orientiert, die als standardisierter digitaler Zwilling der Industrie 4.0 zu betrachten ist. Über ein einheitliches Format ist sichergestellt, dass alle Komponenteneigenschaften homogen in der Verwaltungsschale abgelegt werden. Jede Verwaltungsschale ist in einen Header und einen Body unterteilt. Der Header beinhaltet die eindeutige Identifikation und beschreibt konsistent zur Verfügung stehende Eigenschaften und Fähigkeiten der I4.0-Komponente. Diese werden im Body über Submodelle spezifiziert, in welchen wesentliche Merkmale und Parameter einer I4.0-Komponente (hier: Roboterzelle und Subkomponenten) einheitlich beschrieben werden. Zusätzlich verfügt die Verwaltungsschale über eine standardisierte Programmierschnittstelle (API). Dieses Konzept wurde im Zuge des UAP 9.3.3 auf die Komponenten des Prüfmoduls, der Materialbereitstellung, des Robotersystems und der Leiterplattenidentifikation übertragen.

Eine beispielhafte Darstellung einer AAS (Asset Administration Shell = Verwaltungsschale) und deren Submodelle für das Prüfmodul kann Abbildung 76 entnommen werden. Der digitale Zwilling einer solchen Prü fzelle wird anhand hinterlegter Submodelle der Verwaltungsschale jeweils individuell konfiguriert und definiert. Neben den generischen Submodellen 'Typenschild', 'Identifikation', 'Dokumentation', 'Zustandsüberwachung' und 'Technische Daten' kommen auch komponentenspezifische Submodelle wie 'PCB' (Printed Circuit Board), 'Elektrischer Prüfadapter' und 'Elektrisches Prüfsystem' zum Einsatz. Die Verwaltungsschale eines Prüf adapters besteht bspw. aus einer Teilmenge der genannten Submodelle. So definiert das Submodell 'Elektrischer Prüfadapter' z. B. die Testfähigkeit eines Adapters über eine direkte Zuordnung der damit prüfbaren PCB-Teilenummern digital in einem Blockmerkmal. Im Submodell 'Elektrisches Testsystem' wird zudem die aktuelle Topologie des spezifischen Prüfsystems hinterlegt. Darüber ist der gerüstete Prüfadapter bestimmt. Bei einem Rüstvorgang wird dieses Merkmal im Submodell geändert und der digitale Zwilling dadurch aktualisiert.



Abbildung 76: Beispielhafter Auszug einer Verwaltungsschale für das „Rüstmodul Austauschsatz“

Die Verwaltung der Submodelle erfolgt per Product-Lifecycle-Management (PLM)-System und wird systemtechnisch getrennt vom operativen Produktionssystem realisiert. Dadurch werden die Submodelle einerseits in eine industrieübliche Versionsverwaltung überführt und andererseits die Konformität und Kompatibilität der Merkmalsdefinition und Merkmalsidentifikation nach dem eCl@ss-Standard sichergestellt. Die eigentliche Ausprägung bzw. Instanziierung der Verwaltungsschalen erfolgt erst im Produktionssystem (siehe Abbildung 77).

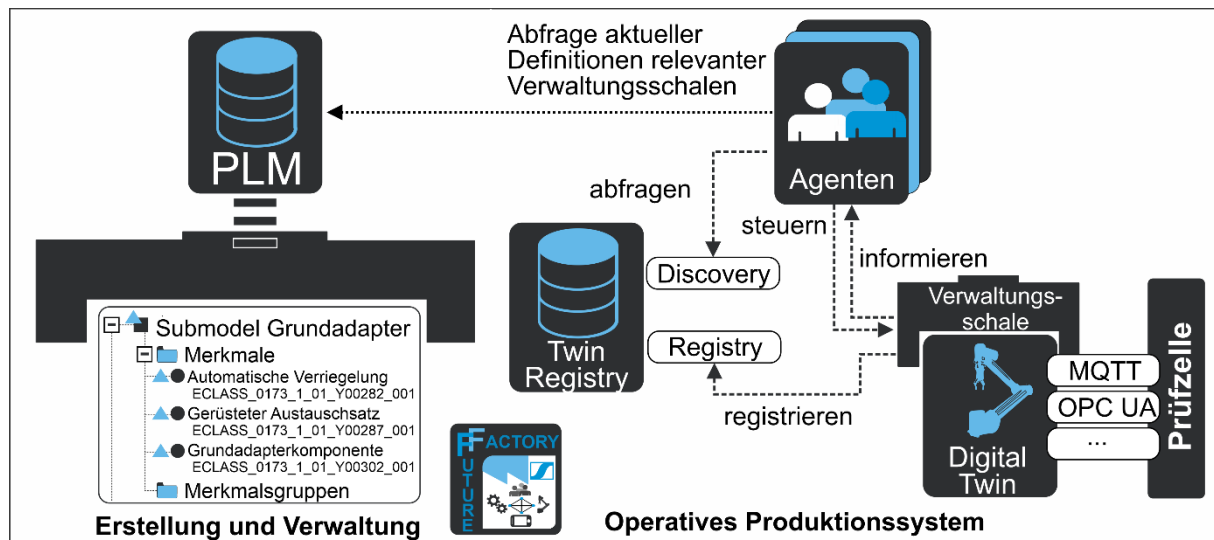
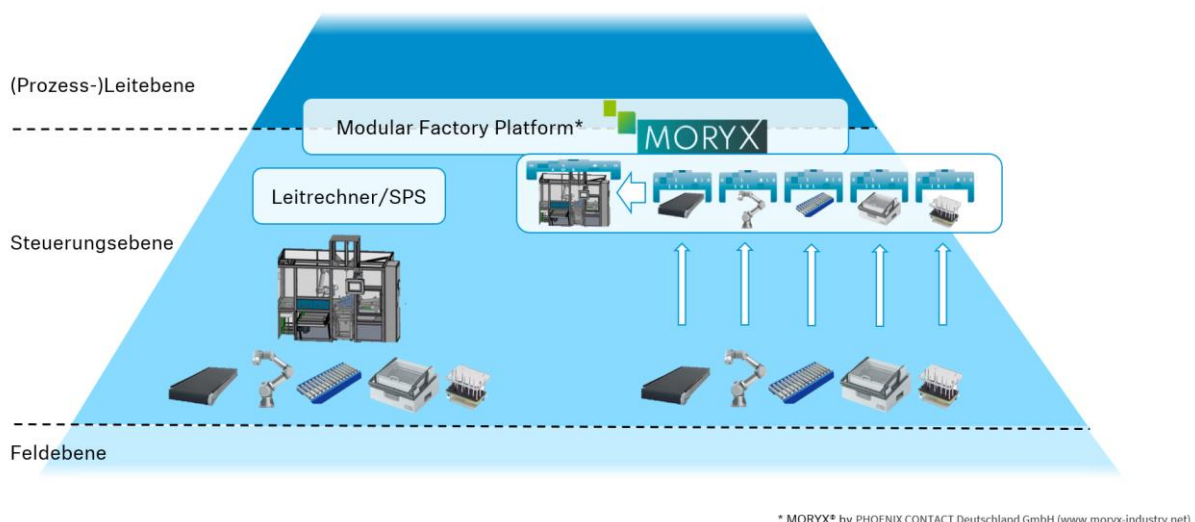


Abbildung 77: Idee des Gesamtprozessdesigns unter der Verwendung der Verwaltungsschale

Für die UAP 9.3.4 bis 9.3.8 hat sich im Projektverlauf eine Änderung vom geplanten Vorgehen ergeben. Da das skizzierte Steuerungskonzept vollständig mit der kommerziellen Software *MORYX* der Firma Phoenix Contact abzubilden ist, wurde sowohl das Kommunikationsdesign als auch das Informationsdesign vollständig übernommen. Über definierte API-Schnittstellen erfolgt zudem auch die Anbindung an relevante Systeme wie das PLM-System, die Sennheiser Testdatenbank und das Enterprise Resource Planning (ERP) System. Gleichzeitig wird das Konzept der Steuerungsagenten mittelfristig auf alle relevanten Arbeitsplätze (vgl. AP 9.4) ausgerollt. Für die Definition der Testszenarien (UAP 9.3.9) wurden vergangene Produktionsprogramme zweier Monate (15.01.2020 bis 15.02.2020) extrahiert und nach gefertigten Produktarten und Losgrößen klassifiziert. Dieser Referenzzeitraum erstreckt sich über eine Phase sehr hoher Kapazitätsauslastung.



* MORYX* by PHOENIX CONTACT Deutschland GmbH (www.moryx-industry.net)

Abbildung 78: Einbindung der Software MORYX der Firma Phoenix Contact GmbH in den Produktionsprozess

Die Software MORYX ist modular aufgebaut und erfüllt zum großen Teil die Anforderungen an den Produktionsprozess. So wurde in einem ersten Schritt zunächst auf Basis des digitalen Zwilling-Konzeptes die physikalische Welt in MORYX abgebildet (siehe Abbildung 78).

Im nächsten Schritt ermöglicht das PLM-System die Definition von Verwaltungsschalen, Metamodellen, Teilmodellen und Merkmalen für Klassentypen von Assets, um später Ausprägungen/Instanzen dieser zu generieren (die Twin-Registry). Die Kommunikation zu diesem PLM-System erfolgt über RESTful API-Endpunkte. Ebenso können in das PLM System eCl@ss-(Advanced)-Kataloge importiert werden.

Des Weiteren beinhaltet die Produktionssteuerungssoftware MORYX unter anderem die Auftragsabwicklung im Zusammenspiel mit dem ERP-System, die Prozesssteuerung, die Integration des PLM-Systems via API und die Verwaltung von digitalen Zwillingen. Abbildung 79 zeigt die prinzipielle Verbindung von MORYX mit der physikalischen Welt die Repräsentanz der physikalischen Welt in MORYX durch die digitalen Zwillinge und die Anbindung an das PLM-System. Die Nachbildung der Submodule in MORYX als digitaler Zwilling ist in diesem AP als der erste Schritt zur Nutzung von AAS und digitalen Zwillingen im Produktionsprozess zu verstehen. Die geplante Realisierung der AAS und der digitalen Zwillinge im PLM-System wird nach diesem Projekt weiterverfolgt. Das heißt, alle Submodule wurden zunächst in MORYX nachgebildet, um in einem späteren Schritt die Integration des digitalen Zwillings dieser Submodule in ein PLM-System zu realisieren.

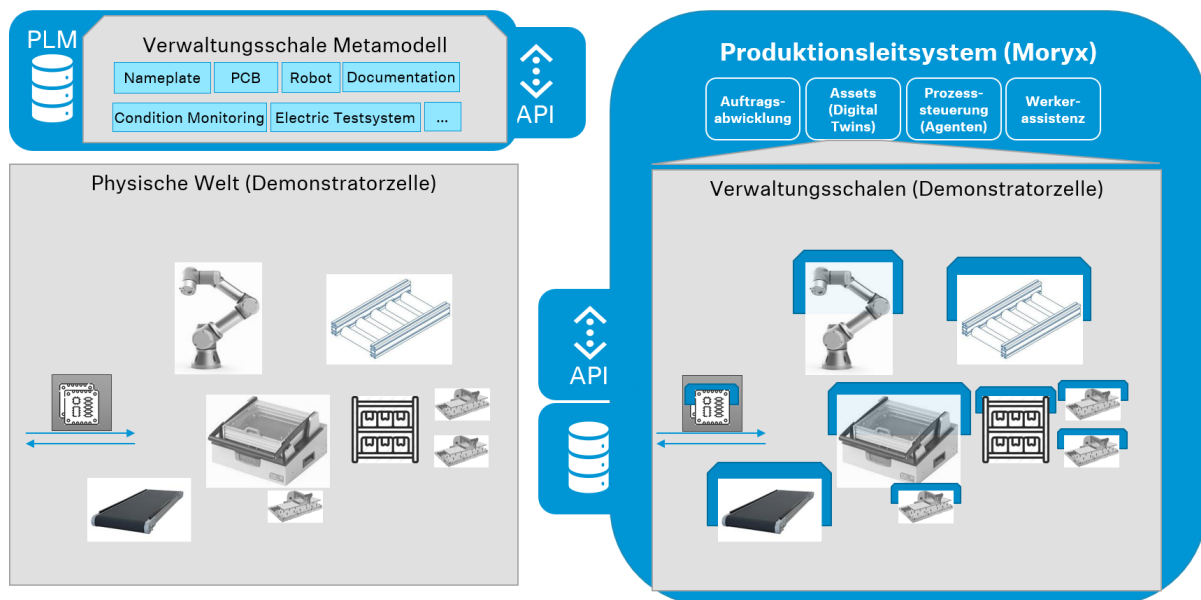


Abbildung 79: Einbindung der Software MORYX der Firma Phoenix Contact GmbH zur Abbildung des digitalen Zwillings

Die Teststrategie (UAP 9.3.14 und 9.3.15) für die Integration des Gesamtsystems in den Produktionsablauf gliederte sich in mehrere aufeinander aufbauende Teilaktivitäten:

- Test der verwendeten Kommunikationsprotokolle (MQTT, OPC-UA, REST-Endpunkte)
- Test der jeweiligen Schnittstelle/API zu den physikalischen Submodulen mit Hilfe von Simulationssoftware
- Test der jeweiligen Submodul-Schnittstelle/API zu MORYX mit Hilfe von Simulationssoftware
- Test der jeweiligen Asset-Implementierung und des digitalen Zwillings in MORYX (Asset-Modul) im Zusammenspiel mit den jeweiligen physikalischen Submodulen
- Test der Schnittstelle/API zur Auftragsabwicklung mit Simulationsdaten
- Sukzessive Einbindung aller Submodule und Komponenten in das Gesamtsystem

- Test der gesamten Prozesskette bzw. Prozesssteuerung

Im Laufe der UAP 9.3.14 und 9.3.15 zeigte sich wie erwartet, dass die Abbildung des Gesamtprozesses unter Berücksichtigung aller möglichen auftretenden Fehlerzustände im Gesamtprozess, die weitaus meiste Zeit bei der Verifikation in Anspruch nimmt. Dadurch nahm die iterative Fehlerbehebung (UAP 9.3.16) einen nicht unerheblichen Anteil in diesem AP ein. Schlussendlich konnte in enger Zusammenarbeit mit dem AP 9.1-Team eine hohe Verfügbarkeit der Demonstratorzelle sichergestellt werden.

Arbeitspaket 9.4: Produktionssystemmanagement

Das Ziel des AP 9.4 bestand in der Analyse des Produktionssystems und der Entwicklung eines Agentensystems, das die Vorteile dezentraler Optimierung und Entscheidungsfindung mit denen bisheriger zentraler Planungs- und Steuerungsansätzen verbindet. Dadurch sollte eine ganzheitliche Optimierung von Key Performance Indikatoren (KPI) erreicht werden.

In einem ersten UAP erfolgte dazu eine Analyse der Einflussgrößen im bisherigen Produktionssystem. Das Produktionssystem wurde diesbezüglich zunächst definiert und dessen Grenzen festgelegt. Das Sennheiser Produktionssystem am Standort Wedemark lässt sich primär in drei Bereiche gliedern. In einem ersten Bereich befinden sich mehrere automatisierte Produktionslinien zur vollautomatischen Bestückung von Leiterplatten. In einem zweiten Bereich liegt ein Bereich zur Leiterplattenprüfung. Dieser wird zum Teil mit Leiterplatten aus dem ersten Bereich und Leiterplatten aus externen Standorten beliefert. In einem dritten Bereich erfolgt die Montage der Leiterplatten sowie weiterer Peripherie, wie z. B. Gehäuse und Netzteile, zu fertigen Produkten. Zur Feststellung des Abhängigkeits- bzw. Verkettungsgrads der beiden Bereiche wurde vor Ort eine Materialfluss- bzw. Wertstromanalyse durchgeführt. Dadurch konnte erkannt werden, dass die drei Bereiche weitestgehend unabhängig voneinander agieren und organisatorisch getrennt funktionieren. Eine Verbindung wird über Eingangs- bzw. Ausgangslager hergestellt, in denen Bestände bzw. Aufträge bspw. durch Kanban Systeme gepuffert werden.

Da die autonome Roboterzelle im zweiten Bereich (Leiterplattenprüfung) zum Einsatz kommen wird und in diesem die Produktionsplanung und -steuerung weitestgehend eigenständig abläuft, wurde die Systemgrenze zur weiteren Betrachtung auf diesen Bereich festgelegt. Externe Einflüsse, d. h. Zu- und Abgänge von Material etc., werden als sog. Quelle bzw. Senke des Bereichs betrachtet. Diesbezüglich wurde für die weitere Analyse der Einflussfaktoren ausschließlich der Bereich der Leiterplattenprüfung betrachtet. Dieser Bereich besteht aus insgesamt 19 Arbeitsplätzen zur Leiterplattenprüfung, die über unterschiedliche Ausrüstung und Prüfequipment verfügen. Zum Zeitpunkt der Analyse waren vier der 19 Arbeitsstationen mit Robotern zur automatisierten Prüfung ausgestattet.

Nach der systematischen Fokussierung auf relevante Bereiche des Produktionssystems wurde der o.g. Bereich auf seine Einflussgrößen analysiert. Dazu wurden diese in mehreren Workshops gesammelt und bewertet. Prinzipiell lassen sich die Einflussgrößen in feste Stamm- und variable Bewegungsdaten separieren. Das Ergebnis der Sammlung ist Abbildung 80 zu entnehmen. Es wurde erkannt, dass das betrachtete Produktionssystem maßgeblich durch die Komponenten Fertigungsaufträge (Material) bzw. Betriebsmittel beeinflusst wird, die über die Komponente Arbeitsplan miteinander in Verbindung gebracht werden. Zusätzlich wurde erkannt, dass wesentliche Einschränkungen im Produktionssystem aus der Zuordnung zwischen den Artikelnummern der Leiterplatten und kompatiblen Betriebsmitteln zur Leiterplattenprüfung (hier: Prüfadapter und compatible Nadelbretter zur Leiterplatten Artikelnummer)

entstammen. Dadurch ergeben sich Abhängigkeiten zwischen Leiterplattenvarianten und Arbeitsplätzen, sodass die Prüfung bestimmter Leiterplatten an einen einzelnen Arbeitsplatz bzw. einen Pool von Arbeitsplätzen gekoppelt ist. Eine weitere Einflussgröße entstammt aus der Komponente Personal, das die Prüfung auf den 15 manuellen Arbeitsplätzen durchführt. Wesentliche Einschränkungen ergeben sich durch unterschiedliche Qualifikationsgrade des Personals zur Bedienung der verschiedenen Arbeitsplätze und zur Durchführung produktspezifischer Arbeitsvorgangsfolgen. Die sich daraus ergebene Produktionsteuerung ist in Abbildung 80 systematisch illustriert.

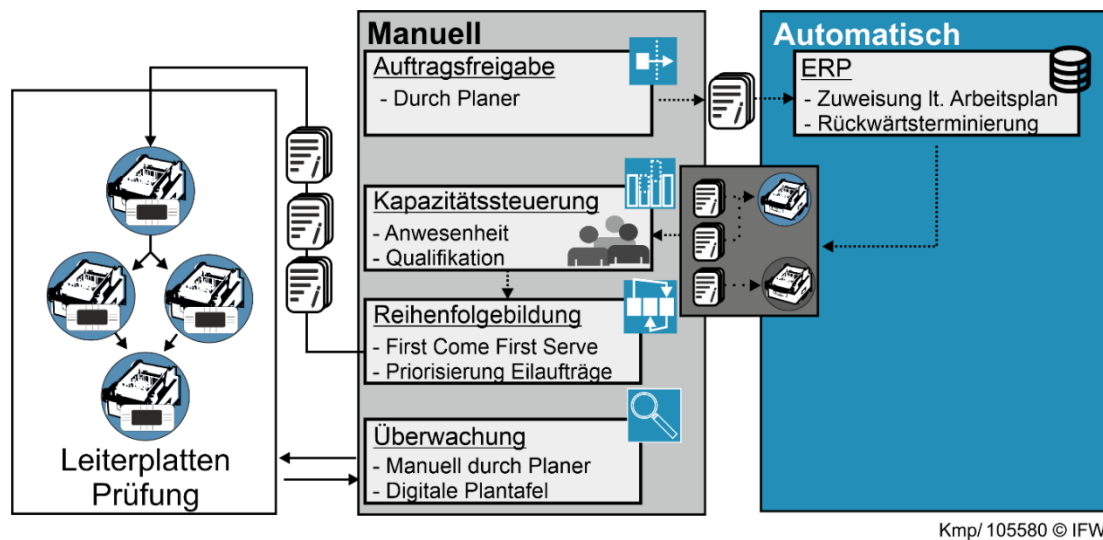


Abbildung 80: Manuelle Produktionssteuerung (Stand September 2021)

Tabelle 14: Einflussgrößen auf den Bereich Leiterplattenprüfung im Produktionssystem

Komponente	Stammdaten	Bewegungsdaten
Fertigungsaufträge	- Artikelnummer - Stückzahl - Bedarfsdatum - Startdatum - Stückzahl - Arbeitsgangsequenz	- Finale Stückzahl bzw. Ausschuss - Verbrauchten Menge - Bestände - Materialstatus - Lagerbestand
Arbeitsplan	- Meilensteinstartdatum - Meilensteindatum - Arbeitsplatz für Prüfung - Arbeitsgang-Sequenz - Arbeitsplanzeiten	- Arbeitsgangstatus - IST-Arbeitszeiten
Betriebsmittel	- Kennung - Bezeichnung - Stammlagerort - Zuordnung Adapter - Zuordnung Nadelbretter - Zuordnung Produkte bzw. Produktfamilie	- Betriebsmittelstatus - Betriebsmittel Lagerort
Personal	- Qualifikationsmatrix - Schichtplan	- Anwesenheitsstatus - Entwicklungsplan

Nach der Feststellung der wesentlichen Einflussgrößen (siehe Tabelle 14) erfolgte zudem eine Sammlung und Bewertung verfügbarer KPI zur anschließenden Festlegung einer Zielfunktion. Das Ergebnis der KPI-Sammlung ist

$$\frac{T_{IST}}{T_{Plan}} \cong 100 \% \quad (2)$$

$$\min_{x \in X} P_{HK}(x) \quad (3)$$

$$\min_{x \in X} B_{Prüfung}(x) \quad (4)$$

T_{IST} : Realisierter IST-Liefertermin

T_{Plan} : Zugesagter Plan-Liefertermin ggü. Bereich Montage und weiteren Kunden

$P_{HK}(x)$: Produktherstellkosten für Artikelnummer x

$B_{Prüfung}(x)$: Bestände für Artikelnummer x im Bereich der Leiterplattenprüfung

X : Grundgesamtheit aller zu produzierenden Artikelnummern

zu entnehmen. Über einen paarweisen Kreuzvergleich (vgl. AP 9.1) wurde anschließend eine Reduzierung und Priorisierung der gesammelten Kennzahlen durchgeführt. Durch dieses Vorgehen wurde die Kennzahl **Lieferperformance** als wichtigste KPI für das Produktionssystem definiert. Daher wurde diese KPI auch als Zielfunktion für das weitere Vorgehen in AP 9.4 übernommen. Zusätzliche Nebenbedingungen ergaben sich aus der Realisierung möglichst **niedriger Produktherstellkosten** und der Realisierung möglichst **geringen Beständen** innerhalb des Bereichs der Leiterplattenprüfung. Der Übersichtlichkeit können die definierte Zielfunktion (Formel 2) und die existierenden Nebenbedingung (Formel 3-4) unten entnommen werden.

$$\frac{T_{IST}}{T_{Plan}} \cong 100 \% \quad (2)$$

$$\min_{x \in X} P_{HK}(x) \quad (3)$$

$$\min_{x \in X} B_{Prüfung}(x) \quad (4)$$

T_{IST} : Realisierter IST-Liefertermin

T_{Plan} : Zugesagter Plan-Liefertermin ggü. Bereich Montage und weiteren Kunden

$P_{HK}(x)$: Produktherstellkosten für Artikelnummer x

$B_{Prüfung}(x)$: Bestände für Artikelnummer x im Bereich der Leiterplattenprüfung

X : Grundgesamtheit aller zu produzierenden Artikelnummern

In AS 9.4.2 wurde aufbauend auf der definierten Zielfunktion ein Controllingsystem bestimmt, das den Einfluss des beabsichtigten Agentensystems auf das Produktionssystem und dessen relevante KPI transparent darlegt. Da sowohl die Lieferperformance ermittelt als auch die Produktherstellkosten und die Bestände im Sennheiser ERP-System bereits vollständig erfasst wurden, kann das Controlling des Demonstrators zukünftig über die bestehenden Systeme sichergestellt werden. Zusätzlich wurden Abstimmungen mit den Vorgängen aus AP 9.3 vorgenommen, sodass diese Kennzahlen auch in den Services, die mit dem zu integrierenden MORYX-System kommunizieren, hinterlegt und abgefragt werden können. Das Gesamtkonzept ist in Abbildung 81 dargestellt. Die gewählten Kennzahlen sind priorisiert Tabelle 15 zu entnehmen.

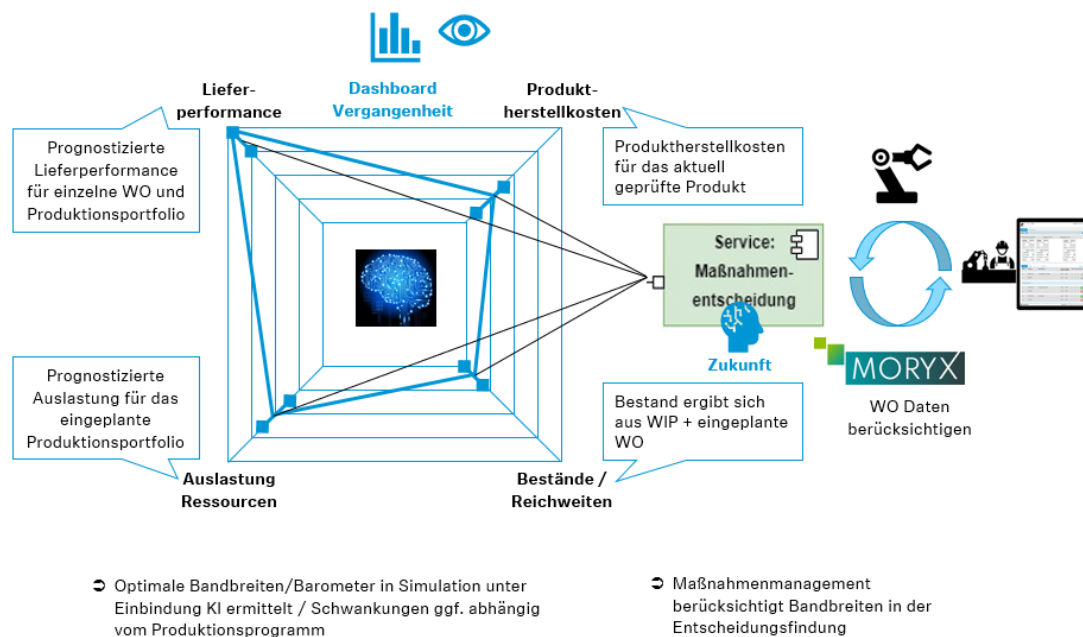


Abbildung 81: Berücksichtigung von KPI in den Services und der Produktionsplanung

In UAP 9.4.3 wurden die Zielfunktionen aus UAP 9.4.1 auf den Demonstrator angewendet. Aufgrund der thematisch engen Überschneidung beziehen sich die weiteren Ausführungen auf die UAP 9.4.4 und UAP 9.4.9. Um den autonomen Betrieb des Demonstrators und zukünftig weiteren Prüfzellen innerhalb des bestehenden Produktionssystems einen autonomen Betrieb zu ermöglichen, wurde das Konzept der „intelligenten Produktionssteuerung“ erarbeitet (vgl. Abbildung 82). Orientiert an dem Gesamtprozessdesign (vgl. Abbildung 75) und ausgerichtet am Konzept der auftragsgesteuerten Produktion verfügt jeder Arbeitsplatz über einen dezentral und autonome agierenden Steuerungsagenten. Im realen Produktionssystem werden diese nach erfolgreicher Systemintegration durch das Software-Framework MORYX (vgl. AP 9.3) realisiert. Zur wissenschaftlichen Untersuchung der „intelligenten Produktionssteuerung“ und zur Determinierung von notwendigen Steuerungs- und Regelparameter wurde in AP 9.4 zudem ein Agentensystem unter der Verwendung des Open-Source Software-Frameworks JADE programmiert und implementiert. Die Grundstruktur dieses Agentensystems ist Abbildung 83 zu entnehmen. In diesem Agentensystem existiert sowohl für jeden Auftrag ein Auftrags- bzw. für jeden Arbeitsplatz ein Arbeitsplatzagent. Auftragsagenten sind in diesem System in der Lage, in Frage kommende Arbeitsplatzagenten zu kontaktieren und eine Bearbeitung bzw. Prüfung entsprechend der Zielfunktion dezentral und autonom zu verhandeln. Zusätzlich erfolgt eine Kooperation mit anderen Auftragsagenten, um eine Priorisierung nachgelagerter Arbeitsvorgänge zur Entscheidungsfindung berücksichtigen zu können.

Tabelle 15: Kennzahlen im Sennheiser Produktionssystem

Nr.	Kennzahlen	Details
1	Lieferperformance	Relative Übereinstimmung zwischen zugesagtem Ablieferungstermin und IST-Abliefertermin
2	Terminverzug	Zeitlicher Verzug zwischen zugesagtem Ablieferungstermin und IST-Abliefertermin
3	Effizienz	Relative zeitliche Abweichung von IST-Zeiten ggü. Arbeitsplanzeiten

4	Bestände in der Fertigung und am Lager	-
5	Produktherstellkosten	Herstellkosten eines Auftrags
6	Durchlaufzeit der Einzel- und Gesamtaufträge	Zurückgemeldete Bearbeitungszeiten aus dem Shop Floor (Produktion)
7	Lean Index	Lean Index = Bearbeitungszeit / Durchlaufzeit
8	Flexibilitäts- und Qualifizierungsgrad der Mitarbeitenden (MA)	Der Flexibilitäts- und Qualifizierungsgrad der MA wird über eine sog. Qualifizierungsmatrix bestimmt, in der produkt- oder arbeitsplatzspezifische Qualifizierungen dem einzelnen MA zugeordnet werden. Kommt zum Stand 09/21 nicht steuernd zum Einsatz.
9	Auslastung der Ressourcen	Relative Auslastung der Arbeitsplätze
10	Verfügbarkeit der Ressourcen (Störungen)	Relative Verfügbarkeit bzw. Nicht-Verfügbarkeit der Arbeitsplätze aufgrund von Störungen
11	Lagerumschlag	Durchschnittliche Verweildauer von Aufträgen bzw. Artikelnummern im Eingangslager
12	Störungen nach Dauer und Ursache	Kategorisierung von Störungen
13	Rüstzeit und Häufigkeit	Anzahl von notwendigen Rüstvorgängen
14	Qualität	Relativer Anteil von Schlecht- vs. Gut-Prüfungen
15	Werkzeug-/ Maschinenstandzeiten	-
16	Materialverfügbarkeit	Relative Verfügbarkeit von Material

Zur Erfüllung der Zielfunktion wurde die Entscheidungsregel der kürzesten Bearbeitungszeit als Verhandlungsgrundlage zwischen Auftrags- und Arbeitsplatzagenten implementiert. Zur Kalkulation der kürzesten Bearbeitungszeit werden alle dem Arbeitsplatz bereits zugeordneten Aufträge (Warteschlange) berücksichtigt und der nächstmögliche Bearbeitungszeitpunkt bestimmt. Die Einhaltung dieser Regel stellt zudem sicher, dass die Zielfunktion vollständig erfüllt wird. Durch die Realisierung einer möglichst schnellen Abarbeitung von Aufträgen werden einerseits die Liefertermine bestmöglich bedient und Bestände durch die flexible Zuordnung der Aufträge auf Arbeitsstationen möglichst geringgehalten. Gleiches gilt in diesem Kontext auch für die Produktionskosten.

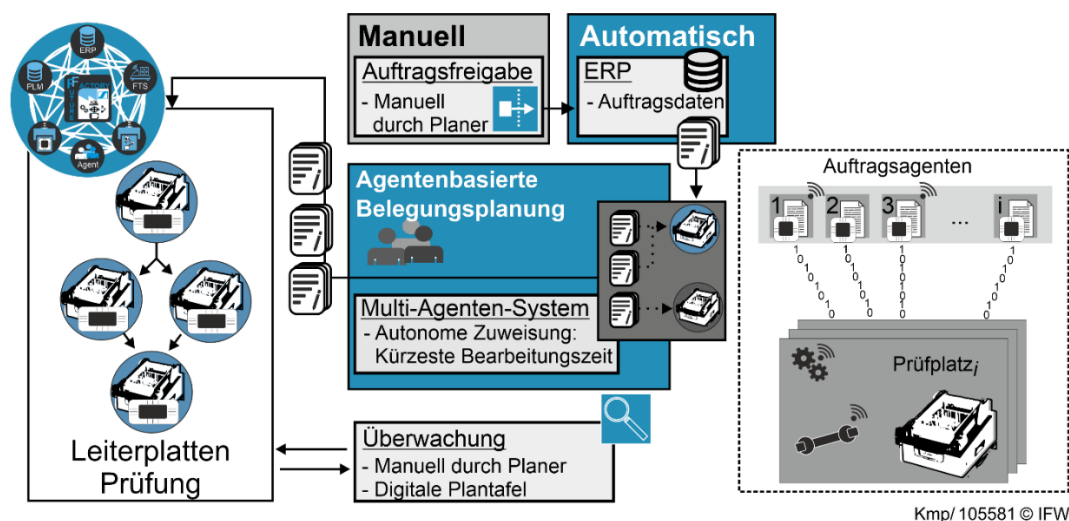
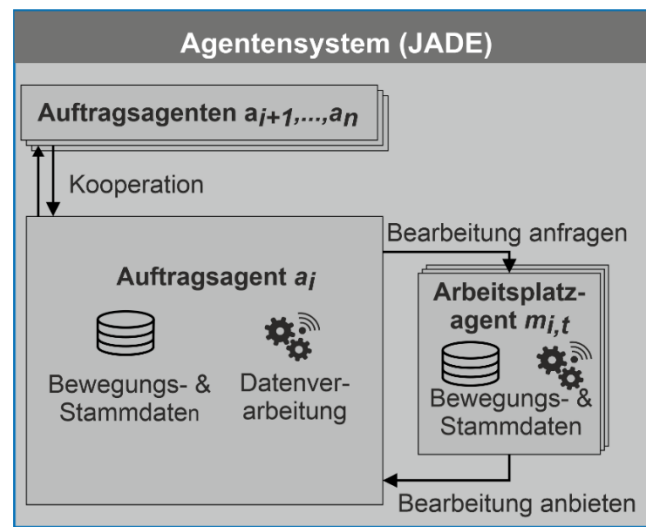


Abbildung 82: Intelligente Produktionssteuerung

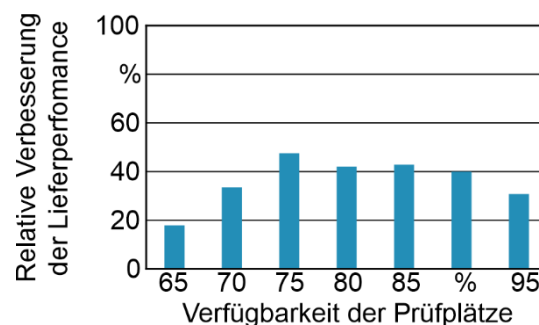
Zur weiteren Untersuchung wurde in den o. g. Arbeitsschritten zudem das zuvor skizzierte Produktionssystem in die Simulationssoftware *Tecnomatix Plant Simulation* überführt und mit dem Agentensystem über die Nutzung von sog. Socket-Schnittstellen kommunikationstechnisch verbunden. Anschließend wurde untersucht, welche Lieferperformance-Steigerungen und Potentiale auf Basis der „intelligenten Produktionssteuerung“ gegenüber der manuellen Produktionssteuerung zu erwarten sind. Es erfolgte eine Evaluierung anhand der definierten Testszenarien (vgl. AS 9.3.9). Zudem wurden die Ergebnisse mit den historischen Daten verglichen. Das historische Produktionsprogramm dieses Zeitraums wurde dazu in die Simulation importiert. Zusätzlich wurde der Ablauf der manuellen Produktionssteuerung in einer zweiten Simulation implementiert.



Kmp/ 105587 © IFW

Abbildung 83: Kommunikationsstruktur des programmierten Agentensystems

Die Ergebnisse dieser Simulationsstudie können Abbildung 84 entnommen werden.



Kmp/ 105582 © IFW

Abbildung 84: Ergebnisse der ersten Validierungsstudie

In diesem Bild ist die relative Verbesserung der Lieferperformance durch den Einsatz einer intelligenten Produktionssteuerung abgebildet. Um das Systemverhalten und dessen Auswirkungen auf den KPI bei stochastisch auftretenden Störungen von Arbeitsplätzen oder Personalengpässen differenziert zu untersuchen, wurden unterschiedliche Verfügbarkeitszenarien (X-Achse) definiert. In diesen Szenarien wurde die Verfügbarkeit von Arbeitsplätzen zwischen 65 und 95 Prozent variiert. Dadurch wurden in der

Simulation ordinäre Anomalien im Sennheiser-Produktionssystem in unterschiedlicher Häufigkeit berücksichtigt und eine Bandbreite möglicher realer Produktionsszenarien simuliert. Über alle Verfügbarkeitsszenarien hinweg ist zu erkennen, dass die Lieferperformance durch eine intelligente Produktionssteuerung im Referenzzeitraum zwischen ca. 19 % und 45 % verbessert werden konnte.

Im weiteren Verlauf des Projekts wurde das Vorgehen der skizzierten „intelligenten Produktionssteuerung“ mit dem MORYX-System (AP 9.3) realisiert und das gesamte System in Betrieb genommen. Eine Grundlage hierfür bildet die Abbildung der Zelle und einzelner Komponenten als digitale Zwillinge in diesem als Produktionslenkungsplattform zu bezeichnenden Systems (siehe AP 9.3).

Über die vorgesehenen Arbeitsinhalte hinaus wurde durch das IFW eine weitere Simulationsstudie vorbereitet, in der der skizzierte Maßnahmenservice aus AP 9.2 erforscht wurde. Über eine Erweiterung der zuvor skizzierten Simulationsumgebung in der *Software Tecnomatix Plant Simulation* wurde eine stochastische Prüfung durch eine Binomialverteilung basierend auf produktspezifische, historische Ausfallraten modelliert und integriert. Basierend auf den in Abbildung 85 dargestellten Service wurden die Zustände der Produktion erfasst und die entsprechende Maßnahme ausgelöst. Die Simulation wurde an den aus AS 9.3.9 erstellen Szenarien durchgeführt.

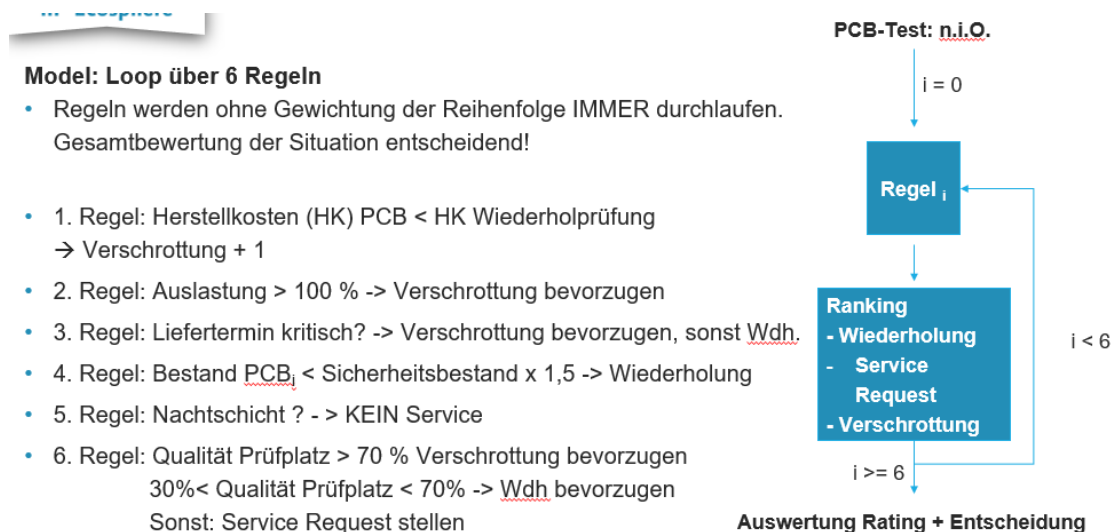


Abbildung 85: Maßnahmenservice zur Pseudofehlererkennung

Ausschlaggebende Kennzahlen waren die über den Service möglichen Kosteneinsparungen und die dazu in Abhängigkeit stehende Lieferperformance der Leiterplatten. Die Lieferperformance wurde in restriktive Kategorien, vor- und nachterminliche Fertigstellung, eingeteilt. Der Maßnahmenservice führte in einzelnen Produktionsprogrammen dazu, dass ca. 39 % der Aufträge verspätet waren. Insgesamt ist die Prüfung jedoch in der Produktion mit Lagerplätzen entkoppelt, sodass die Lieferperformance anhand der Lagerreichweite besser beurteilt werden kann. Als kritische Reichweite wurde ein Tag definiert, die in keinem Produktionsprogramm unterschritten wurde. Über die Simulation errechnete Produktionseinzelkosten und der Maschinenstundensätze für den Service wurde das monetäre Potenzial bestimmt. Das Einsparungspotenzial des Maßnahmenservices betrug in der Simulation zwischen ca. 200.000 € und 600.000 € je nach Produktionsprogramm (Szenario), welche ein Auftragsvolumen von einem Monat beinhalteten.

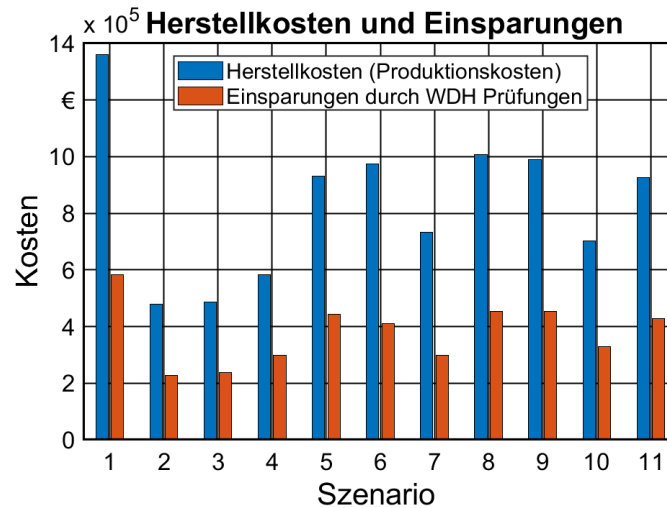


Abbildung 86: Herstellkosten und Einsparung der Leiterplatten

Arbeitspaket 9.5: Mehrwertbildung für die IIP-Ecosphere Community & Transfer

Mit dem Ziel der Erstellung einer Dokumentation mit geeigneter Darstellungsform zur Mehrwertbildung in der IIP-Ecosphere und geeigneter Präsentationsform für die Öffentlichkeit hat sich Sennheiser 1,2 PM beschäftigt. Dieser Aufwand inkludiert auch die Aufwände in AP 11.1 (Best Practices) und AP 14.4 (Individuelle Pläne für wirtschaftliche Verwertung). Hier sind insbesondere das sogenannte Werkstattgespräch zu nennen, welches in einem Live Stream direkt aus der Sennheiser Produktion übertragen wurde und im Internet über die Homepage <https://www.iip-eosphere.de/> weiterhin als YouTube Video abrufbar ist. Des Weiteren fanden diverse Publikationen mit Sennheiser Beteiligung in schriftlicher und mündlicher Form statt, die im Abschnitt 2.4 genannt sind.

2.1.10 Demonstrator „KI-basierte Prozess-Stabilität in der Glasformung“ (AP 10)

Anforderungsanalyse & Konzept (UAP 10.1)

Prozessdatenanalyse

In der ersten Umsetzungsphase des Projektes lag der Schwerpunkt auf der Erfassung von Prozessdaten aus realen Produktionsserien zur Spritzenherstellung bei dem Projektpartner Gerresheimer Bünde GmbH. Dazu wurde die beim Kooperationspartner vorhandene G3-Plattform so angepasst und erweitert, dass Prozessrohdaten aufgezeichnet und für weitere Verarbeitungsschritte genutzt werden können. Der Aufbau und die Datenerfassung der G3-Plattform sind in Abbildung 88 dargestellt. Mit Hilfe dieser Plattform konnte eine große Menge an Prozessdaten von mehreren Terabyte pro Tag und Maschine aus mehreren realen Produktionsserien erfasst werden. Die gesammelten Daten wurden anschließend umfassend untersucht und analysiert, um relevante Zusammenhänge zwischen den überwachten Größen zu identifizieren. Als Ergebnis der Datenanalyse wurden zwei Prozessschritte als vielversprechende Kandidaten für weitere Untersuchungen identifiziert, nämlich die Konusformung (Teil der Formmaschine) und der Siebdruckprozess (Teil der Bedruckungsstation), siehe Übersicht der Spritzenherstellung in Abbildung 87. Für beide Prozessschritte wurden digitale Zwillinge mit Hilfe neuronaler Netze erstellt.

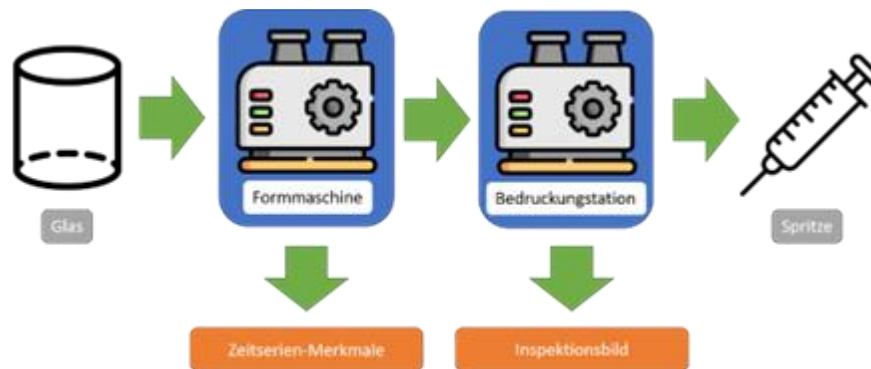


Abbildung 87: Vereinfachte Darstellung einer Produktionsstraße für Spritzen

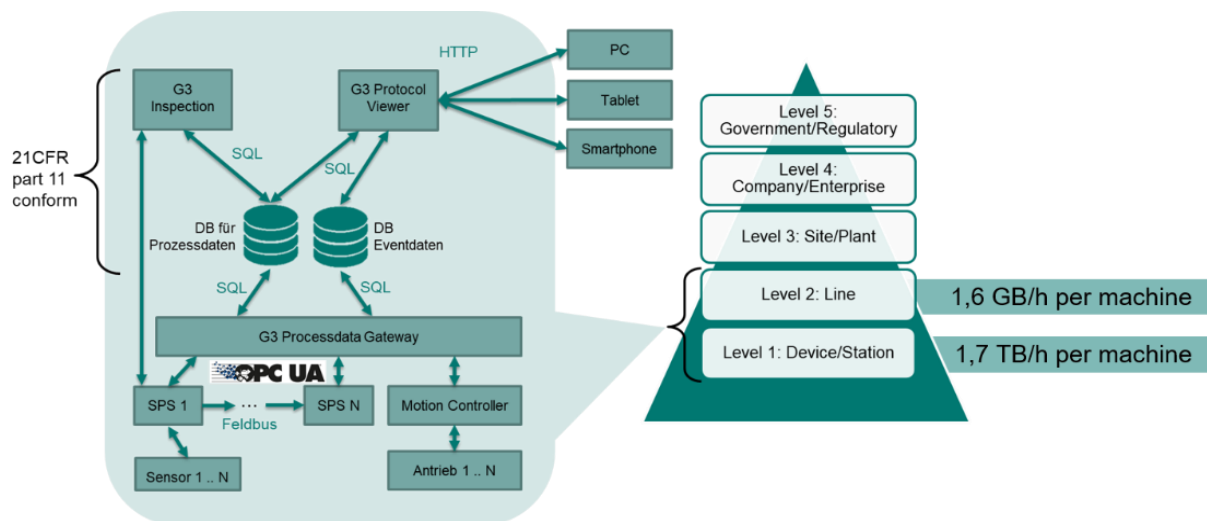


Abbildung 88: Plattform zur Datenerfassung (Quelle: Gerresheimer Bünde GmbH)

Modellierung des Glasformprozesses

Basierend auf der Datenanalyse wurden Prozessparameter und andere überwachte Größen, die einen signifikanten Einfluss auf die Ergebnisse des zugrunde liegenden Prozesses haben, für das Training der neuronalen Netze verwendet. Die Dimensionierung der Netzarchitekturen und anderer Parameter, die für das Modelltraining erforderlich sind, wurde mit Hilfe fortgeschrittener Bayes'scher Hyperparameter-Optimierungstechniken durchgeführt, die ein Gleichgewicht zwischen Exploration und Ausbeutung während der Hyperparametersuche gewährleisten. Weitere Einzelheiten zur Modellierung des Glasformungsprozesses sind in der Beschreibung von UAP 10.2 zu finden.

Simulation des Glasformprozesses zur Trainingsdatengewinnung

Zur Erzeugung weiterer synthetischer Daten für das Training wurde ein Signalgenerator entwickelt, der die in der Datenanalyse als relevant identifizierten Merkmale des realen Glasformungsprozesses nachbildet. Das Blockdiagramm des entwickelten Signalgenerators ist in Abbildung 89 dargestellt. Hierbei wird ein synthetisches Datensignal durch die Kombination einer linearen Funktion, die den Trend des Signals repräsentiert, mit zwei Sinusfunktionen, die die periodische Komponente des Signals repräsentieren, erzeugt. Alle für die Erzeugung der synthetischen Daten erforderlichen Parameter wie die Stei-

gung und die Achsenabschnitte der linearen Funktion sowie die Frequenzen und Amplituden der Sinusfunktionen basieren auf realen Prozessbeobachtungen und wurden entsprechend den statistischen Eigenschaften zufällig aus mehreren Gaußverteilungen entnommen. Diese synthetischen Daten wurden dann zur Anreicherung der Trainingsdaten verwendet, so dass die trainierten Modelle robustere Merkmale erlernen konnten, die zur Verbesserung der Modellgeneralisierung auf ungesehene Daten beitragen. Darüber hinaus hat sich in der Praxis gezeigt, dass die Anreicherung der synthetischen Daten mit dem entwickelten Signalgenerator das Konvergenzverhalten des Trainingsprozesses positiv beeinflusst.

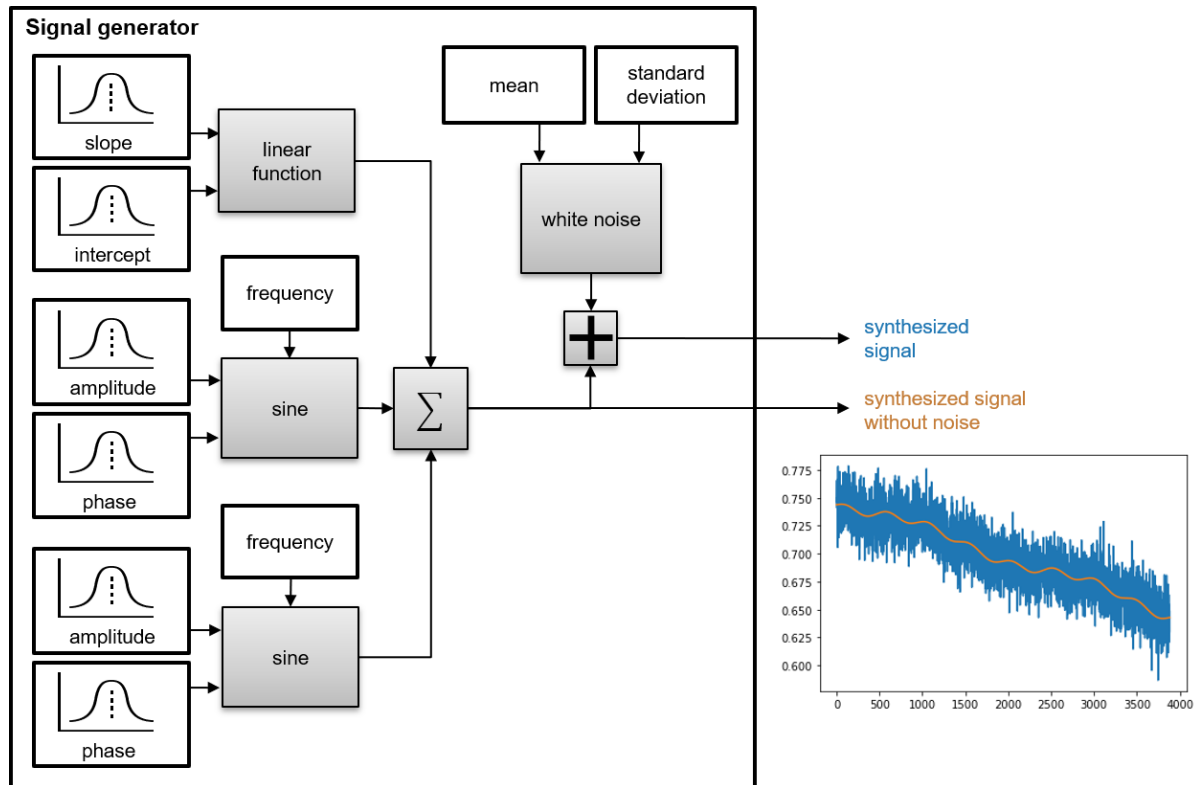


Abbildung 89: Blockdiagramm des entwickelten Signalgenerators

Entwicklung Teillösungen (UAP 10.2)

Training einfacher Entscheidungsbäume

Mehrere baumbasierte Methoden wurden für den Glasformungsprozess evaluiert und getestet. Dazu gehören Decision Tree, Random Forest und fortgeschrittene Boosting-basierte Baummethoden wie XGBoost und CatBoost. Das Ziel des Trainings ist die Vorhersage der Wartungszeit der Formmaschine basierend auf Prozessparametern und aktuellen Sensordaten, wie z. B. dem Innendurchmesser des Glases. Um die Leistungsfähigkeit und Verallgemeinerbarkeit der Modelle weiter zu verbessern, wurden Feature Engineering und Feature Selection durchgeführt. Darüber hinaus ermöglicht es das baumbasierte Modell, die Wichtigkeit jedes einzelnen Merkmals im Prozess und seinen Beitrag zum Endergebnis zu analysieren und weniger relevante Merkmale aus dem Modell zu entfernen (Feature Selection). Baumbasierte Modelle dienen als Grundlage für komplexere Modelle in den folgenden APs.

KNN zur Abbildung des Glasformungsprozesses

Es wurde eine neue Methode entwickelt, um das Black-Box-Modell des Produktionsprozesses aus den Daten zu lernen. Neuronale Netzwerkmodelle wurden entwickelt, um eine bessere Performanz zu erzielen, die durch das baumbasierte Modell begrenzt war. Die Methode basiert auf dem hierarchischen Variational Autocoder, einer neuronalen Netzwerkarchitektur, die hochdimensionale Daten in relevante latente Variablen kodiert. Im Gegensatz zum Vanilla Autocoder haben Methoden, die auf dem Variational Autocoder basieren, den Vorteil, dass die latenten Variablen voneinander unabhängig sind. Auf dieser Methode basierende digitale Zwillingsmodelle wurden erfolgreich für zwei Fertigungsprozesse entwickelt und integriert. Bei der Formdornverschleißvorhersage geht es darum, den Zeitpunkt des Formdornaustausches aufgrund des Endes der Formdornlebensdauer vorherzusagen (Abbildung 90). Das Ziel dieser Vorhersage ist die Maximierung der Formdornbetriebszeit unter variablen Verschleißbedingungen, die durch unterschiedliche Produkte (z. B. unterschiedliche Anforderungen) und Schwankungen der Prozessvariablen durch externe Faktoren verursacht werden. Dadurch werden die Kosten für unausgelastete Dorne minimiert und Ausfallzeiten reduziert. Darüber hinaus kann ein auf Autoencoder basierendes Modell leicht Anomalien erkennen, die durch externe Faktoren wie Bedienereingriffe verursacht werden.

Die zweite Methode bewertet den Drucksiebzustand automatisch (Abbildung 91). In diesem Fall werden Sensor- und visuelle Daten aus dem Inspektionssystem vorverarbeitet und als Input für das Modell verwendet. Das ungefähre Alter des Siebes wird für das Training verwendet, indem angenommen wird, dass das Drucksieb zu einem späteren Zeitpunkt in einem ähnlichen oder schlechteren Zustand sein sollte als das Drucksieb zu diesem Zeitpunkt (Abbildung 91 a)). Das KI-Modell lernt den relativen Zustand zwischen den Bildern x_1 und x_2 . Es prädiziert die Deep Degradation Metric (DDM) über die Zeit und kann einen geeigneten Austauschzeitpunkt vorhersagen, bevor das schlechte Produkt produziert wird (Abbildung 92). Gleichzeitig lernt es den relativen Prozess und kann das Produkt in Bezug auf den DDM-Wert vorhersagen (Abbildung 91 b)) und sich auf bestimmte Merkmale konzentrieren, die für die Vorhersage relevant sind (Abbildung 91 c)).

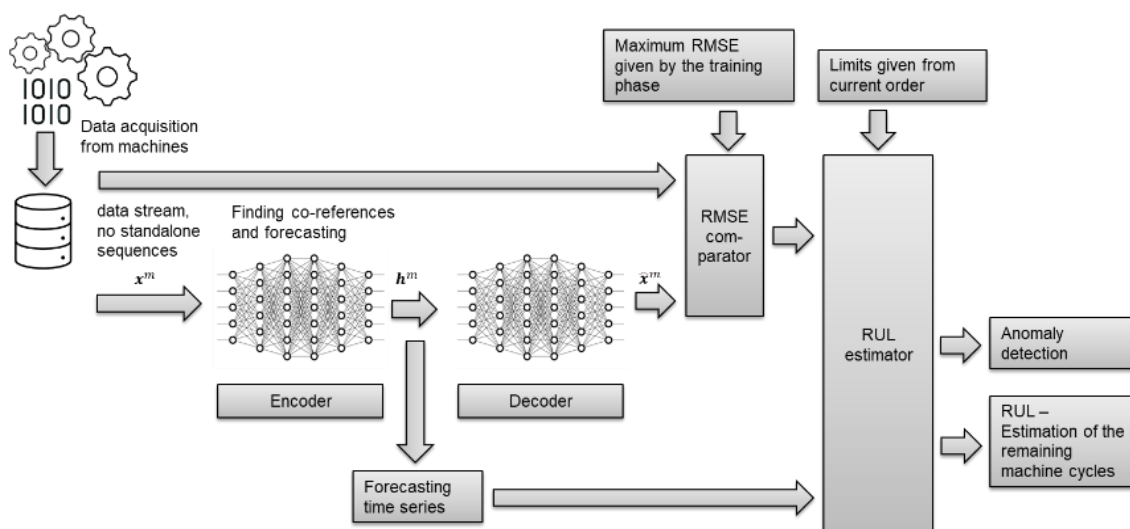


Abbildung 90: Schema des Autoencoders zur Prädiktion des Formdornverschleißes

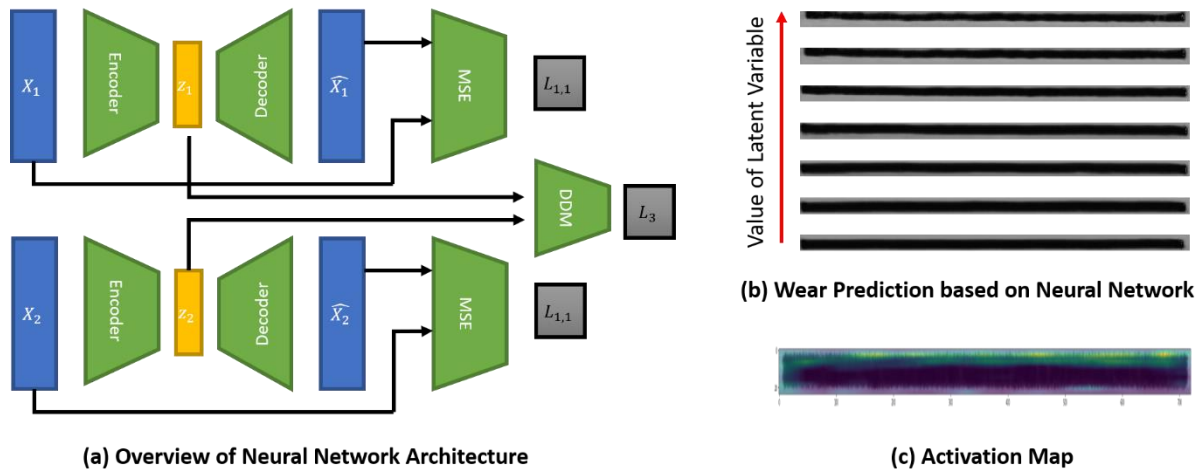


Abbildung 91: (a) Schema des Autoencoders für die Bewertung des Zustands des Drucksiebs. (b) Vorhersagen des Druckbildes bei Veränderung der latenten Variable, die den Verschleiß repräsentiert. (c) Aktivierungskarte der für die Vorhersage relevantesten Merkmale.

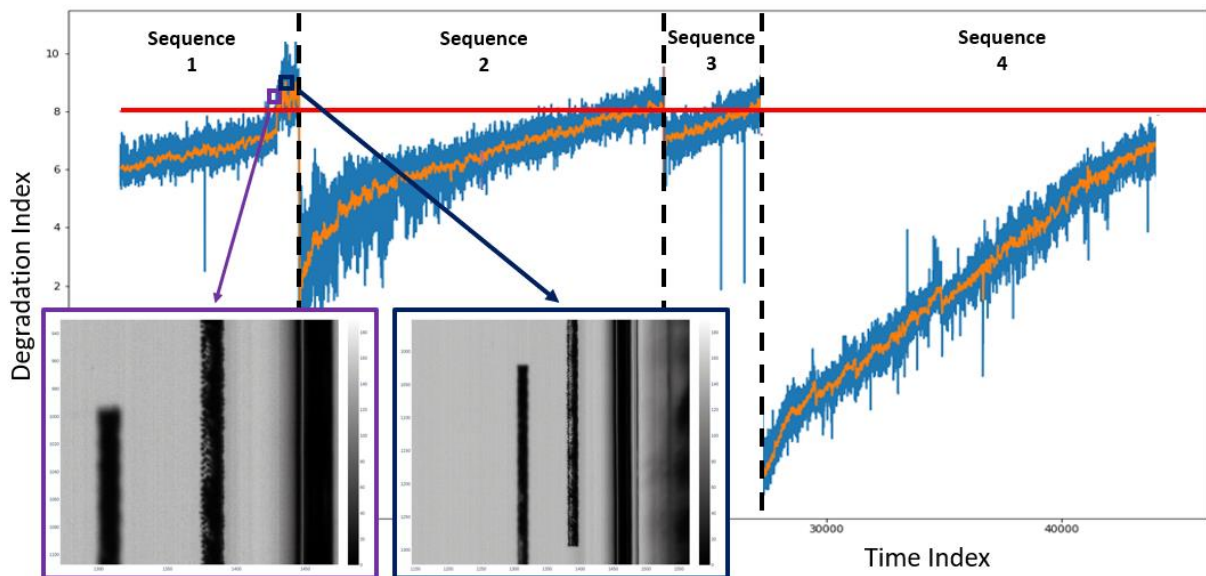


Abbildung 92: Die aus den Inspektionsbildern selbständig gelernte Deep Degredation Metric (DDM) (blau/orange) spiegelt sehr gut den Verschleißzustand des Drucksiebs wider. Die rote Linie stellt einen Schwellwert für eine notwendige Wartung dar. Die Sequenzen zeigen reale Wartungen (Drucksieb 1-2,3-4 oder Farbwechsel 2-3).

Modellerklärbarkeit

Die Modellerklärbarkeit ist ein entscheidender Aspekt, der bei der Anwendung der entwickelten Modelle in realen Szenarien berücksichtigt werden muss. Um in der Praxis die Gründe für die Modellentscheidungen sowie die Beiträge der einzelnen Eingabemerkmale zu den Modellergebnissen zu verstehen, haben wir Wichtigkeitsanalysen auf den Eingangsgrößen durchgeführt, die anhand Attribution Maps mit integralen Gradienten (IG) erfolgten. Aufbauend auf zwei Axiomen, nämlich der Sensitivität und der Implementierungsinvarianz, berechnet IG den durchschnittlichen Gradienten des Modelloutputs in Bezug auf seine Inputs entlang eines Pfades zwischen einer Baseline und dem betrachteten Input. Unsere empirischen Beobachtungen aus diesen Analysen haben gezeigt, dass die trainierten Modellen, z. B.

für den Konusformungsprozess, in der Lage waren, die dominanten Frequenzen in den Prozessdatensignalen automatisch zu erlernen. Darüber hinaus haben sie mehr Fokus auf die relevanten Merkmale gerichtet, deren Bedeutung anhand von Attributionskarten visualisiert werden kann (siehe Beispiel in Abbildung 93).

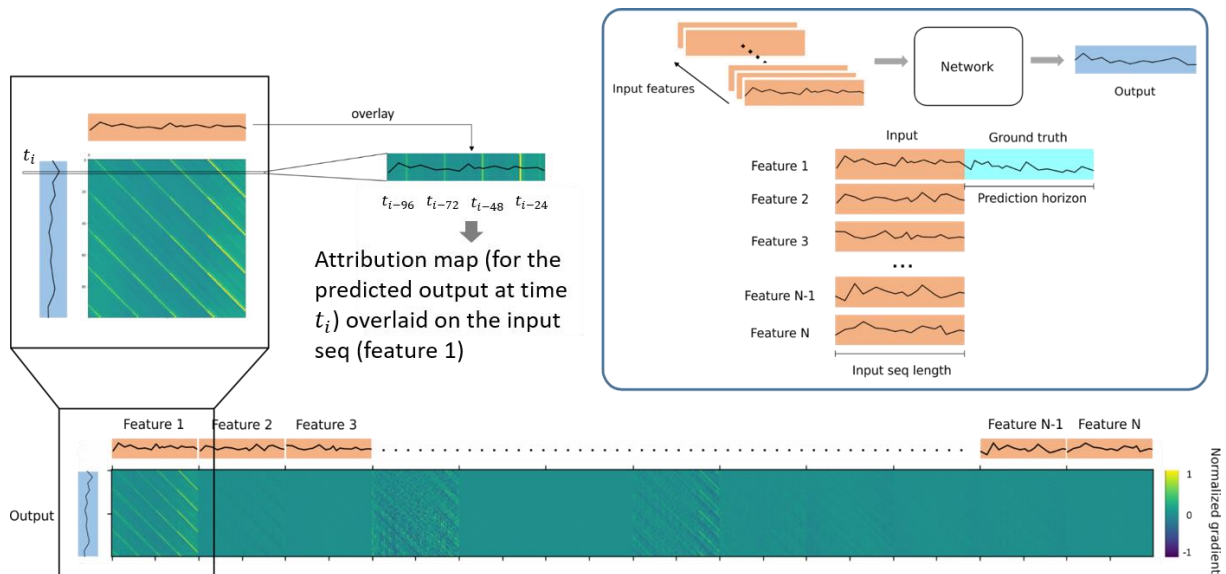


Abbildung 93: Wichtigkeitsanalyse der Eingangsgrößen zur Modelerklärbarkeit mittels Attributionskarten

Um für die Deep Degredation Metric eine Einschätzung zu erhalten, ob durch die entwickelte Methode tatsächlich selbständig aus den Inspektionsbildern eine plausible Verschleißmetrik angelernt werden kann, wurde die Eigenschaft des im Modell verwendeten Autoencoders genutzt, auf der Decoder-Seite aus dem latenten Variablen wieder ein Druckbild vorherzusagen, dass dem jeweiligen Zustand des latenten Raums entspricht. Um dies auszunutzen, wurde ein Tool entwickelt, welches dem Anwender die Möglichkeit gibt, über einen Schieberegler direkt die latente Variable zu manipulieren, die dem Verschleißwert entspricht. Der Decoder produziert anschließend ein mögliches Druckbild bei diesem Verschleißgrad, welches durch Experten auf Plausibilität überprüft werden kann (siehe Abbildung 94). Für die anderen latenten Variablen ist ein ähnliches Vorgehen möglich, um beispielsweise den Einfluss von Umgebungsbedingungen oder unterschiedlicher Produkttypen zu visualisieren.

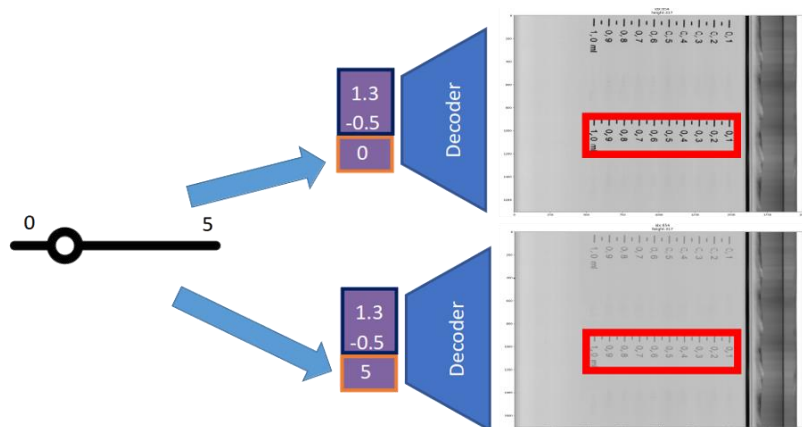


Abbildung 94: Tool zur Visualisierung des gelernten Wissens des Deep Degredation Metric Modells. Mittels des Schiebereglers kann ein Verschleiß gewählt werden, für das das Netz ein Druckbild prädiziert, welches auf Plausibilität geprüft werden kann.

Vorschlag neuer Prozessparameter mittels KNN Backpropagation

Aufbauend auf der Architektur neuronaler Netze, die im Rahmen des vorangegangenen UAP entwickelt wurde, wird das Modell so erweitert, dass der Prozess implizit aus den Daten gelernt wird, indem die Prozessparameter und das Ergebnis unter normalen Bedingungen variiert werden.

Das Ziel ist die Vorhersage besserer Prozessparameter für ein gewünschtes Ergebnis, z. B. vorgegeben durch eine Produktspezifikation, falls der Prozess aktuell nicht optimal läuft. Diese Art der Vorhersage unterstützt den Anwender bei der Auswahl der besten Parameter und macht Vorschläge, welche Prozessparameter getestet werden sollten, um den digitalen Zwilling bzw. den Prozess weiter zu verbessern. Diese Vorhersage kann durch Änderung des Wertes der latenten Variablen des trainierten Autoencoders erfolgen, der einer bestimmten Produktspezifikation eines Produkts entspricht.

Systemintegration und Implementierung (UAP 10.3)

Prädiktion einfacher Prozessparameter

Schnittstellen und APIs wurden implementiert und in den bestehenden Produktionsprozess integriert. Dies beinhaltet den Vorverarbeitungsschritt und das Speichersystem, um genügend Daten für die Entwicklung von baumbasierten und neuronalen netzwerkbasierenden Modellen zu sammeln. Nach Fertigstellung des Offline-Modellprototyps wurden die baumbasierten Modelle unter Verwendung der vorhandenen APIs in den Produktionsprozess in das System integriert.

Prädiktion von Prozesseinstellgrößen für das Gesamtsystem

Die Integration mit API für die Vorhersage von Prozessparametern wurde erweitert, um die Produktspezifikation für mehrere Teilprozesse des Glasformungsprozesses für die Vorhersage von Prozessparametern zu erfassen. Das entwickelte Modell wurde in das G3-System des Demonstrators integriert. Um den Benutzer bei der Entscheidungsfindung zu unterstützen, wurde ein einfaches Ampelsystem entwickelt, das anzeigt, wann ein Austausch erforderlich ist und wann eine Anomalie erkannt wurde.

Integration des KNN in den realen Glasumformungsprozess

Um den Produktionsprozess der Glasformung weiter zu verbessern, wurde ein Feedbacksystem in das Modell integriert. Dies ermöglicht eine direkte Parameteroptimierung des Gesamtsystems. Die Prozessparameter werden iterativ so lange optimiert, bis der Vorhersagefehler zwischen Zielvorgabe und simuliertem Ergebnis minimiert ist.

Verifikation & Validierung (UAP 10.4)

Validierung der Prozessvorhersage

Die entwickelten Modelle wurden in das G3-System integriert, was die Überprüfung der Leistungsfähigkeit der Modelle im laufenden Produktionsprozess ermöglichte. Die Abweichungen zwischen den vorhergesagten und den überwachten Ergebnissen konnten als Feedbacksignal für die weitere Optimierung der neuronalen Netzmodelle verwendet werden.

Validierung der Prozessparameteroptimierung

Durch die Validierung direkt am G3 System zeigte sich, dass die Rückmeldungen, die auf der Abweichung zwischen den Modellergebnissen und den gewünschten Ergebnissen gemäß den produktspezifischen Spezifikationen basieren, die Parameteroptimierung des Produktionsprozesses ermöglichen

Um die Generalisierung des trainierten Vorhersagemodells für andere Maschinen sowie für leicht veränderte Prozesse zu verbessern, wurde eine neue Methode zur Domänenanpassung (Transfer Learning) entwickelt und implementiert. Die Performance der entwickelten Methode wurde anhand eines öffentlich verfügbaren Benchmark-Datensatzes validiert. Mit dem gleichen Ziel, eine bessere Generalisierung für andere Maschinenkonfigurationen zu erreichen, wurde das für die Zustandsbewertung von Drucksieben entwickelte Modell so erweitert, dass es nicht nur für eine vorgegebene feste Maske, sondern für beliebige Drucklayouts eingesetzt werden kann.

Mehrwertbildung für die IIP-Ecosphere Community & Transfer (UAP 10.5)

Im Rahmen des IIP-Ecosphere Projekts wurden zwei Demonstratoren der in diesem Projekt entwickelten Technologie und ihrer Integration in den Spritzenherstellungsprozess entwickelt. Beide wurden auf der Messe EMO im September 2023 ausgestellt und vorgeführt.

Bei der Gerresheimer Bünde GmbH wurde das erste IIP Werkstattgespräch direkt im Werk veranstaltet, um dort dem ganzen Konsortium sowie dem interessierten Fachpublikum die in AP 10 entwickelten Technologien zu präsentiert sowie auftretende Fragen zu beantwortet. Für eine weitere Verbreitung und eine größere Reichweite wurde eine Aufzeichnung des Werkstattgesprächs über die youtube-Plattform der Allgemeinheit sowie dem interessierten Fachpublikum zur Verfügung gestellt.

Besonders wichtig war hierbei, die Notwendigkeit des strukturierten Vorgehens bei der Einführung von KI-Systemen in die Produktion zu vermitteln, also dem Weg von der richtigen Datenakquise über die sehr wichtige Datenaufbereitung, der Analyse der Daten hin zur Entwicklung spezieller für das Problem und die Daten geeigneter KI-Modelle. Zudem wurde dies bei diversen Projekttreffen innerhalb IIP und bei der Unterstützung anderer Gruppen weiter eingebracht.

2.1.11 Exploration und Best Practice (AP 11)

Best-Practices (UAP 11.1)

In diesem UAP ist es das Ziel, erfolgreiche Ansätze und Vorgehensweisen in Form von Best Practices zu sammeln, zu dokumentieren und zu systematisieren. Fokus ist dabei, Wissen zu etablierten und erfolgsversprechenden Vorgehensweisen bereitzustellen, welches die Umsetzung von KI-basierten Anwendungen in der Produktion erleichtert. Hierzu wurde zunächst ein gemeinsamer Prozesses zur Ermittlung, Beschreibung, Qualitätssicherung und Veröffentlichung von Best Practices definiert. Es wurden zwei Arbeitsgruppen zur Erfassung und Systematisierung von Best Practices zu den Themen „Umsetzung von KI-Projekten“ sowie „Akzeptanz von KI-Lösungen“ gegründet. Die Auswahl dieser zwei Themen basiert auf den meist-nachgefragten Problemstellungen aus der IIP-Ecosphere Industrieumfrage (siehe UAP 6.1). Eine Erkenntnis dabei war, dass in Unternehmen ein erhöhter Bedarf an Practices mit starkem Umsetzungsfokus besteht. Geringerer Bedarf besteht bei technischen Themen wie bspw. Datenbereinigung.

In den Arbeitsgruppen kollaborierten die Partner LUH, FAU und SSE mit Industriepartnern (bspw. Slashwhy, RapidMiner, Sennheiser), Vertretern aus Industriegremien (VDW) sowie dem BMBF Projekt HAISEM-Lab, das ähnliche Fragestellungen bearbeitet. Die Gruppen führten eine Literaturrecherche und Auswertung zu den initialen Best Practice Themen, Diskussion initialer Best Practices und zur ge-

meinsamen Sammlung von Best Practice Kandidaten und deren initialer Beschreibung durch. Insbesondere wurde ein Interviewleitfaden zur Befragung von Industrieunternehmen erstellt, um die Sichten auf und Anforderungen an Best Practices zu erfassen.

Thematisch erfolgte eine weitere Fokussierung auf Best Practices im Kontext des „Rollouts“ von KI-Lösungen, also dem Schritt von der prototypisch umgesetzten KI-Anwendung hin zum breiten Einsatz. Hierzu wurden unterschiedliche Praktiker aus dem Projekt und außerhalb des Projektes befragt. Im Anschluss an die Interviews erfolgte eine Dokumentation und Systematisierung. Hierzu wurde eine einheitlichen Beschreibungsvorlage zur Dokumentation von Best Practices entwickelt. Darüber hinaus erfolgte die Erstellung einer einheitlichen Rollenbeschreibung zur Anwendung in der Dokumentation von Best Practices sowie die Diskussion und Festlegung eines Vokabulars zur Dokumentation von Best Practices.

Eine Erkenntnis der Aggregation und der Interviews bestand darin, dass im Wesentlichen noch wenig etablierte und übertragbare Practices im o.g. Themenbereich existieren. Da die meisten befragten Unternehmen sich noch in einer frühen Phase der KI-Einführung befinden, fand eine Etablierung und Schärfung der Vorgehensweisen im Unternehmenskontext noch nicht statt. Dennoch konnten im Rahmen der Untersuchung Handlungsempfehlungen und Erfahrungen extrahiert werden. Es wird angestrebt die Erkenntnisse im Kontext des Oberthemas „Rollout von KI-Lösungen – Vom Prototypen zum produktiven Einsatz“ als Whitepaper zu veröffentlichen.

Aufbau des Experimentierfeldes (UAP 11.2)

Ziel des UAP 11.2 war der Aufbau eines Experimentierfeldes zur Erprobung von innovativen Technologien im Kontext KI-basierter intelligenter Produktion insbesondere durch KMU und Startups. Das Experimentierfeld besteht aus IT-Anteilen und produktionstechnischen Anteilen (IT-Experimentierfeld, PT-Experimentierfeld).

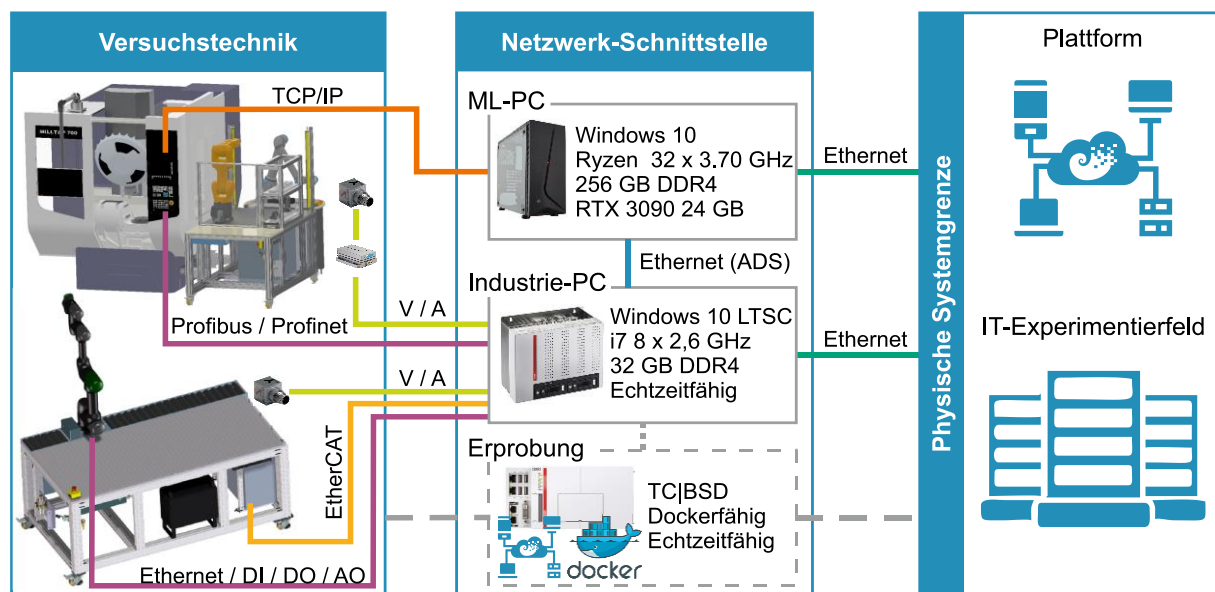
Im IIP-Experimentierfeld (IT-Experimentierfeld) wird die Möglichkeit geboten, KI-Methoden per Cloudzugang zu nutzen und ihr Potenzial für die intelligente industrielle Produktion direkt zu erproben. Als Use-Case wurde mit dem Föderierten Lernen (FL) ein vielversprechender neuer KI-Ansatz gewählt, der eine Alternative zum häufig datenschutzrechtlich schwierigen und auch von Unternehmen kritisch betrachteten Teilen von Daten für die Anwendung von KI-Methoden insbesondere von Methoden des Maschinellen Lernen (ML) darstellt. FL erlaubt unterschiedlichen Dateneinhabern, ein gemeinsames Modell zu trainieren, ohne ihre Daten zu teilen. Dies erlaubt es sogar Wettbewerbern, eine Datenallianz zu bilden und ein gemeinsames Modell zu trainieren, von dem alle Beteiligten profitieren können. Gleichzeitig sind auf der Grundlage von föderiertem Lernen auch neue Geschäftsmodelle möglich. In der Praxis kann dies auch Produktionsszenarien umfassen, in denen Unternehmen Daten mit identischen Maschinen erfassen und daher einen gemeinsamen Merkmalsraum haben. Hier kann föderiertes Lernen den Unternehmen helfen, ein robustes und effektives gemeinsames Modell zu erhalten, ohne Daten zu teilen.

Anstatt Daten zu teilen, trainieren beim föderierten Lernen die Teilnehmer eine lokale Kopie eines Modells mit unternehmenseigenen Daten und tauschen die daraus resultierenden Parameteraktualisierungen iterativ aus – in der Regel über eine vertrauenswürdige, zentralisierte Einrichtung. Aus den unterschiedlichen Parameter-Beiträgen der Teilnehmer wird dann ein aktualisiertes Modell berechnet und den Teilnehmern werden daraus wiederum verbesserte Parameter zur Verfügung gestellt.

Im Projekt wurde ein containerisiertes, föderiertes Lern-Framework implementiert, das es interessierten Parteien ermöglicht, die Vorteile dieses Lernansatzes zu explorieren. Dabei wurde zunächst ein container-basiertes, flexibles Back-End auf der Basis existierender FL-Ansätze umgesetzt, getestet und schrittweise verfeinert. Um das FL-Framework auch nutzbar zu machen, wurde zudem eine Benutzerschnittstelle umgesetzt, die es erlaubt eine FL-Prozess zu konfigurieren, zu starten und zu beobachten. Ziel war hier eine niedrigschwellige Möglichkeit zu Erprobung von FL-Prozessen. Deshalb wurden zusätzlich auch mehrere für die Produktion relevante Datensätze bereitgestellt, um ein direkte Erproben zu ermöglichen ohne eigene Daten preiszugeben. Im entwickelten System können Unternehmen somit ein solches Modell mit föderierten Lernen trainieren und die akteursübergreifende Kollaboration in einem neutralen Rahmen erproben.

Beispielszenario

Im PT-Experimentierfeld wurde aufbauend auf Expertengesprächen, Workshops und enger Zusammenarbeit im IIP-Konsortium ein Experimentierfeld geschaffen, welches aus einem Hochleistungsnetzwerk aus echtzeitfähigem Industrie-PC, Hochleistungsrechner, 5 Achs-Bearbeitungszentrum, diverser Sensorik und Schnittstellen besteht. Mit dem System können KI-Anwendung im Kontext das moderne, spannende Produktionsumfeld von KMUs, Startups und Forschende erprobt werden. Im Think Tank „KI & Produktion“ diente das Experimentierfelds zur Erforschung selbststeuernder Prozessketten und Qualitätsregelkreise (Kapitel 2.6.4).



Bildquellen: Beckhoff, Novatech, Docker, Kistler

He/109558 © IFW

Abbildung 95: Experimentierfeld – Vernetzung und Schnittstellen

Um breites Spektrum an Produktionsanwendungen abzudecken, erfolgte in enger Zusammenarbeit mit UAP 13.2 und anderen Arbeitspaketen die Erweiterung des Experimentierfelds (IT und PT) mittels einer mobilen, individuell erweiterbare Cobot-Handhabungszelle zu einer mobilen Messedemonstrators (Abbildung 95 und Abbildung 96). Die Cobot-Zelle verfügt über eine modifizierbare Linear-Positionierachse, wahlweise über ein oder zwei Cobot-Roboter und diverse technische Steuerungselemente. Integrierte Verstellmechanismen ermöglichen in das System gezielt Störgrößen einzubringen und die Einflüsse mit Hilfe von KI-Modellierungen im Anwendungsfeld der Zustandsüberwachung zu untersuchen.

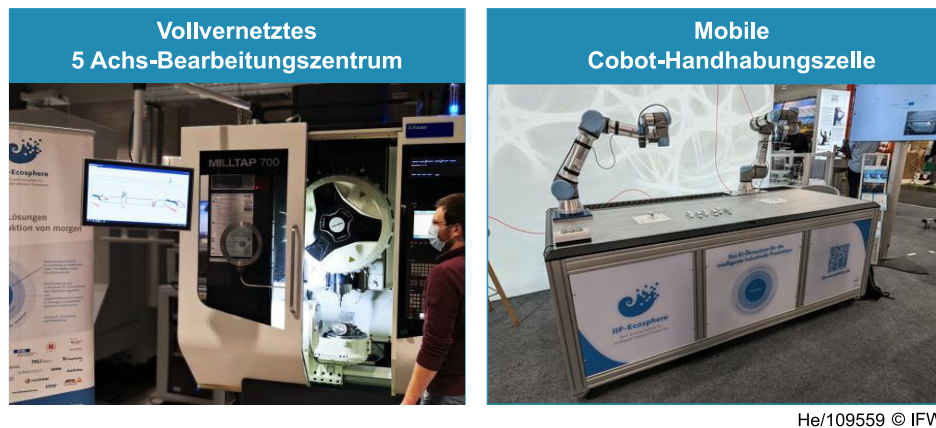


Abbildung 96: Experimentierfeld und Demonstrator (UAP13.2)

Die Cobot-Zelle diente beispielweise der Validierung der IIP-Ecosphere-Plattform sowie zur Präsentation unterschiedlicher Demonstrator-Anwendungsfällen auf Messen, Tagungen und Workshops (siehe Detail in UAP13.2). Insgesamt wurde das Experimentierfeld von diversen Firmen und Startups, wie u. a. INDAAQ Solutions, Maschibo, ATS Elektronik, Lenze, Tetralytix, Koehn AI Consulting, NovoAI und MIP genutzt, um neuartige KI-Methoden im Produktionsumfeld zu erproben.

Im Rahmen des Experimentierfeldes wurde auch der Datenmarktplatz mit Beispieldatensätzen (Open Data) aus dem Bereich der industriellen Produktion befüllt und dem Konsortium zu Testzwecken zur Verfügung gestellt (Abbildung 97). In diesem Zuge wurde zudem ein interaktives Benutzerhandbuch erstellt und für das Konsortium verfügbar gemacht. Aus Gründen der einfachen Verwaltung wurden dabei alle Komponenten auf den Servern des Fraunhofer ISST bereitgestellt.

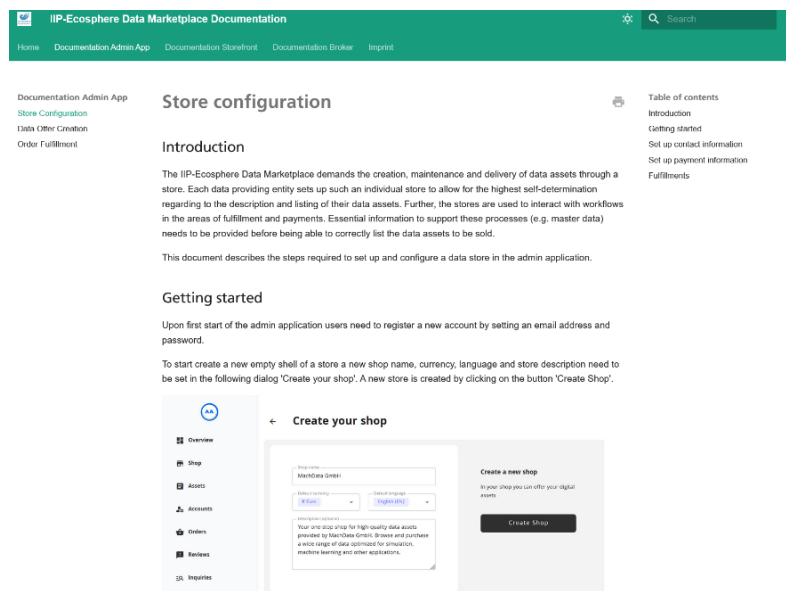


Abbildung 97: Interaktive Dokumentation Datenmarktplatz

Dynamisch definierte Demonstratoren (UAP 11.3)

Neben den bereits in der Antragsphase thematisch festgelegten vier Demonstratoren sind weitere kleinere Demonstratoren vorgesehen, die dynamisch während der Umsetzungsphase definiert und umge-

setzt werden. Basierend auf Ideen und Themen, die sich aus den Arbeiten in den TT und im KI-Accelerator ergeben, sollen hierfür Ideenwettbewerbe für Demonstratoren ausgeschrieben werden. Diese sollen sich insbesondere an KMU und Startups richten, wobei die Teilnahme nicht auf die assoziierten Partner beschränkt ist. Mit diesem offenen Ansatz möchte das Konsortium zusätzliche Anreize zur Nutzung von und Beteiligung an der IIP-Ecosphere schaffen und Ideen, die sich erst während der Umsetzungsphase in der Stakeholder-Community bilden, aufgreifen.

Die Ausschreibung für die dynamischen Demonstratoren (KI-Kleinprojekte) wurde konzipiert und erstellt. Die Ausschreibung wurde über Kommunikationskanäle innerhalb der IIP-Ecosphere beworben und auf der Homepage veröffentlicht. Beteiligte Partner bewarben die Ausschreibung zudem über ihre am Institut aufgebauten Vertriebskanäle (Newsletter, Homepage, Foren, Netzwerke, etc.). Die Ausschreibung war vom 18.10.2022 bis zum 30.11.2022 online. Dabei sind elf Bewerbungen aus der gesamten Bundesrepublik eingegangen. Für den reibungslosen Ablauf der Ausschreibung standen für kurzfristige organisatorische als auch technische Fragen die jeweiligen Expertinnen und Experten zur Verfügung. Zur Bewertung und Vorauswahl der potenzialstärksten Bewerbungen wurde eine Gutachter-sitzung mit Experten des Projektes IIP-Ecosphere durchgeführt. Mit den potenzialstärksten Bewerbungen wurden Termine im Januar 2023 vereinbart, um eine abschließende Beurteilung und Klärung von Auflagen vorzunehmen und offene Fragen zu klären. Daraufhin starteten die Projekte und wurden im Laufe des Jahres 2023 durchgeführt. Nachfolgend ist eine kurze Übersicht zu den ausgewählten Projekten zu finden:

KI-basierte Rattererkennung mittels smarterer Werkzeughalter, KOEHN AI Consulting

Die Erkennung von Ratterschwingungen in Zerspanprozessen ist eine wesentliche Herausforderung bei der Optimierung von Zerspanprozessen. Bekannte Systeme zur Rattererkennung, haben jedoch eine Reihe von Nachteilen. Die meisten Lösungen nicht sehr gut verallgemeinerbar in dem Sinne, dass sie auf bestimmte Maschinen- und Werkzeugkonfigurationen beschränkt sind. Andere Werkzeuge oder Maschinen können in der Regel nur durch ein erneutes Training des KI-Algorithmus einbezogen werden, das häufig auf leistungsstärkeren Maschinen durchgeführt werden muss. Im Demonstrator war daher das Ziel, eine Lösung zu entwickeln, die Ratterschwingungen mit einer größeren Robustheit gegenüber den genannten Randbedingungen bietet. Hierzu wurden im KI-Experimentierfeld eine Reihe von Datensätzen in Zerspanprozessen aufgezeichnet. Auf Basis der Daten wurde eine neuartige Ratter-KI angelernt. Eine Veröffentlichung der Ergebnisse ist in Vorbereitung.

Föderiertes Lernen für die Produktion (NovoAI)

Ziel des Kleinprojektes war es einen KI-Dienst aufzubauen und in die IIP-Ecosphere Plattform zu integrieren. Der KI-Dienst nutzt eine Methode zur Objekterkennung nutzen, die mit Federated Learning (FL) trainiert und eingesetzt wird. Für die FL-Architektur wurde ein zentraler Server für die Modellaggregation genutzt, der als KI-Dienst in der Plattform eingesetzt wird. Als Ergebnis steht eine FL-Integration und KI-Service für die IIP-Plattform bereit. Hierbei konnten weitere Erkenntnisse für die Nutzung der IIP-Plattform durch dritte gewonnen werden.

Synthetische Bilddaten (Synthetic Images)

Die Verfügbarkeit von Bilddaten, beziehungsweise der Aufwand zur Generierung einer ausreichenden Anzahl für Aufgabenstellungen in der optischen Qualitätskontrolle, stellt ein wesentliches Kostenhinder- nis dar. Im Rahmen des Kleinprojektes ist das Ziel, die Potentiale von synthetischen Bilddaten für den KI-Einsatz zu untersuchen und aufzuzeigen. Als Anwendungsfall wurde dazu in Zusammenarbeit mit dem TT KI&Produktion eine Segmentierung von Schweißzonen beim Wire Arc Additive Manufacturing (WAAM) ausgewählt. Die Information der Bereiche ist im Anwendungsfall für die Nachbearbeitung er- forderlich. Aufgrund des aufwendigen WAAM-Prozesses ist die Erzeugung von ausreichend vielen Trai- ningsbauteilen mit hohem Aufwand verbunden. Im Rahmen des Kleinprojektes erzeugte die Fa. Syn- thetic Images daher einen Bilddatensatz auf Basis von Rendertechniken in Verbindung mit einem KI- Style Transfer. Eine Veröffentlichung des Datensatzes und der Ergebnisse hinsichtlich Nutzen zum Training der KI ist in Vorbereitung.

Magnetische ID in der IIP-Ecosphere Plattform, (MIP Technology)

Im Rahmen des Kleinprojektes wurde ein Demonstrator entwickelt, der die magnetischen Informations- plattform (MIP Technology GmbH) in die IIP-Ecosphere Plattform anbindet. Dieser Demonstrator dient dazu die Robustheit der magnetischen ID im Produktionsumfeld zu zeigen und um die IIP-Ecosphere Plattform um Dienste zur Produktidentifikation zu erweitern. Der Demonstrator besteht hardwareseitig aus ID-Tags, einem Scanner und entsprechender Software zur Interpretation und Ausgabe der Infor- mation aus dem jeweiligen ID-Tag. Die Einbindung der erfassten Daten in die IIP-Ecosphere Plattform war wesentlicher Bestandteil der Arbeiten. Dies wurde erfolgreich umgesetzt und auch im Rahmen des Messedemonstrators auf der EMO 2023 demonstriert.

In einer Erweiterung wurde in Zusammenarbeit mit dem Think Tank Daten auch ein KI-Service für die Plattform entwickelt, der die Robustheit der Dekodierung der magnetischen Sensordaten erhöht. Auch dies wurde im Rahmen des EMO-Messedemonstrators gezeigt.

2.1.12 Innovation und Stakeholder-Einbindung (AP 12)

Regional Innovation Hubs (UAP 12.1)

Zur Einbindung und Integration der verschiedenen Akteure aus dem Ökosystem der intelligenten Pro- duktion wurden in den Regionen Nürnberg und Hannover Regional Innovation Hubs (RIHs) aufgebaut. Diese hatten das Ziel, regionale Akteure zu vernetzen, um Technologie-Transfer, Innovation und Grün- dungsaktivitäten in diesem Umfeld durch gemeinsame Aktivitäten voranzubringen. Dazu sollten die RIHs regionale Veranstaltungen mit weiteren Intermediären aus dem Bereich Produktion durchführen. Die Ausgestaltung der RIHs stand allerdings fast während der gesamten Projektlaufzeit im Spannungs- feld wenig absehbarer Beschränkungen aufgrund der COVID-19-Pandemie. Mithilfe einer zweigleisigen Strategie konnten trotz dieser Unsicherheiten die RIHs weiterentwickelt werden. Neben der synergeti- schen Zusammenarbeit der Konsortialpartner aus Hannover und Nürnberg bei der Konzeption für On- line-Veranstaltungen mit überregionaler Reichweite als Innovation Hub Deutschland wurde frühzeitig eine Regionalisierung der RIH auf hybridem Wege vorangetrieben. Eine Übersicht über die durch die RIHs konzipierten und durchgeführten Veranstaltungen findet sich in Abschnitt 1.6.13 unter Veranstal- tungen.

RIH Hannover

Die Partner des RIH Hannover trafen sich regelmäßig online zum Innovation Round-Table, um Maßnahmen und Veranstaltungskonzepte zu diskutieren und vorzubereiten. Außerdem war IIP-Ecosphere auf der Hannover Messe 2022 auf dem Stand des RIH-Partners hannoverimpuls vertreten. Zu den Aufgaben des RIH gehörte auch die Förderung von Startups. Zwei Ausgründungen aus der Leibniz Universität Hannover, NOVO AI und Tetralytics, die 2021 am L3S beziehungsweise am IFW mit dem EXIST-Gründerstipendium gefördert wurden, werden sowohl von RIH-Partnern als auch von IIP-Ecosphere betreut. Beide Ausgründungen nutzen auch das IIP-Experimentierfeld.

Um bei den Forschenden der akademischen Partner frühzeitig im Projekt das Thema „Gründung“ bewusst zu machen und die gegenseitige Vernetzung zu unterstützen, wurde im Juli 2020 gemeinsam mit dem Partner starting Business ein Gründungsworkshop mit 17 Teilnehmenden durchgeführt.

RIH Nürnberg

Durch Platzierung von Beiträgen aus dem Think Tank „Geschäftsmodelle“ ist eine Aktivierung der Community bei regionalen, digitalen Veranstaltungen von Multiplikatoren wie dem *Zukunftszentrum Süd* erfolgt (AP 12.1.2). Mit der absehbaren Lockerung der COVID-19-induzierten Beschränkungen wurden innovative Präsenzformate recherchiert und in studentischen Arbeiten entwickelt, um die Regionalisierung des RIHs zunehmend durch Präsenzveranstaltungen zu etablieren und sich von bereits bestehenden Formaten in der Metropolregion Nürnberg abzusetzen. Bevor die ursprünglich geplanten Veranstaltungs- und Partizipationskonzepte aufgrund der COVID-19-Pandemie in virtuelle Formate überführt wurden, wurden außerdem bereits drei Konzeptionsworkshops und zwei Kooperationsgespräche mit den assoziierten Partnern *JOSPEHS* und *ZOLLHOF* geführt.

Arbeitskreise (UAP 12.2)

Ziel der Arbeitskreise war es die Diskussion und Gestaltung relevanter Themen für KI in der Produktion mit relevanten Akteuren und Multiplikatoren zu etablieren. Durch die pandemische Situation wurden die Arbeitskreise überwiegend digital durchgeführt, was die Reichweite der Veranstaltungen erhöhte. Zudem ermöglichte dies, dass die Arbeitskreise von allen Think Tanks teilweise gemeinsam als übergreifende Veranstaltung unter dem Label „IIP-Ecosphere Open Core“ durchgeführt werden konnten. Dadurch konnten auch Stakeholder in den Austausch gebracht werden, die sonst nicht miteinander gesprochen hätten. Auch konnten Synergien zwischen den Arbeitskreisen besser genutzt werden.

Folgende Arbeitskreise wurden etabliert und besonders im Rahmen der „Open Core“-Veranstaltungen durchgeführt:

- Arbeitskreis Plattformen mit Schwerpunkt auf der Neuheiten zur IIoT-Plattform,
- Arbeitskreis Geschäftsmodelle mit Schwerpunkt auf der Erhebung von Innovationsbedarfen und Evaluation von Methoden und Werkzeugen des TT Geschäftsmodelle.
- Arbeitskreis des TT Daten mit Schwerpunkt auf Rahmenbedingungen des Data Sharing und zum Datenmarktplatz
- Arbeitskreis zum Experimentierfeld: Ergänzung und anwendungsgetriebene Weiterentwicklung des KI-Experimentierfeldes.

Community Plattform (UAP 12.3)

Um die Ergebnisse weithin von IIP-Ecosphere weithin sichtbar zu machen, gemeinsam die Idee von IIP voranzutreiben und den Austausch zwischen Stakeholdern zu verbessern, war es Ziel dieses UAP den Aufbau einer Community rund um IIP-Ecosphere zu fördern und insbesondere hierfür den Aufbau einer Community-Plattform zu koordinieren.

Um den Aufbau der Community und der geplanten Community-Plattform vorzubereiten wurden zunächst unter Leitung des Hafven Innovation Team wurden zwei Workshops zum Thema Aufbau von Communities, Strategieentwicklung, Stakeholder und Onboarding durchgeführt. Es wurde eine Roadmap für das Community-Building entwickelt. Auf Basis dessen wurden die Arbeiten zur Konzeptentwicklung der Community durchgeführt. Des Weiteren wurden kommerziell verfügbare Community Software-Lösungen verschiedener Anbieter hinsichtlich Erfüllung der zuvor erarbeiteten Anforderungen bewertet.

Nach einem Erfahrungsaustausch mit unterschiedlichen Akteuren wurde entschieden keine dedizierte Community-Plattform entwickeln zu lassen. Aufgrund der Eigenheiten der Zielgruppe im Produktionsfeld (eher konservativ) wurde es als eher unwahrscheinlich eingeschätzt eine kritische Masse an Akteuren auf einer solchen Plattform erreichen zu können. Vor diesem Hintergrund dieses Risikos wurde entschieden auf diese Investition zu verzichten und die Community über existierende Kanäle (Events, persönliche Gespräche, LinkedIn und andere Social Media Kanäle sowie die Web-Seite) zu adressieren und den Aufbau der Community auf andere Weise weiter voran zu treiben.

So wurde der Aufbau einer Community für das entstehende Ökosystem während der ganzen Projektlaufzeit aktiv vorangetrieben. Einen besonderen Schwerpunkt beim Community-Aufbau bildete dabei die virtuelle IIP-Ecosphere IIoT Plattform. Diese ist inzwischen so weit gediehen, dass sie vorgeführt und in Anwendungen und Anbindungen an existierende Komponenten (z. B. Anbindung Edge Device von Phoenix Contact) erprobt werden kann. Mit dem Ziel die Plattform bekannt zu machen und eine Community rund um die Plattform aufzubauen wurde diese in unterschiedlichen Kontexten vorgeführt und demonstriert. Dies geschah z. B. bei der Hannover Messe, den Tagen der digitalen Technologien und bei der Vorstellung der Plattform bei einem PIN-Treffen (<https://www.pin-hannover.de/>). Die Plattform stieß dabei auf viel positive Resonanz. Zudem wurde die Plattform als Open Source in GitHub bereitgestellt. Dadurch wird der Aufbau einer weiteren, stärker entwicklungs-orientierten Community gefördert, welche die IIP-Ecosphere Plattform nutzt und Beiträge zu ihrer Weiterentwicklung leistet. Die Aktivitäten rund um die Ausschreibung der dynamischen Demonstratoren dienen ebenfalls dem Community-Aufbau. Der Ausbau der Community rund um die Plattform, die ein wichtiges aus IIP-Ecosphere hervorgehendes Asset bildet, soll auch nach Projektende systematisch weiter betrieben und verstärkt werden.

2.1.13 Dissemination/Sichtbarkeit (AP 13)

Durchführung von Veranstaltungen (UAP 13.1)

Im Rahmen des Projektes wurde eine große Zahl an Veranstaltungen durchgeführt. Diese hatten die Information und Sensibilisierung für das Thema KI in der Produktion, die Einbindung von Stakeholdern und die Erhöhung der Sichtbarkeit von IIP-Ecosphere zum Ziel. Nachfolgend werden besonders erwähnenswerte Formate kurz vorgestellt:

Werkstattgespräche der Demonstratoren: 2022 startete die Online-Veranstaltungsreihe „Werkstattgespräche“, die Einblicke in die im Projekt entwickelten Demonstratoren gewährt und als Blueprint zeigt, wie KI-Lösungen in der Produktion implementiert werden können. Außerdem konnten die Zuschauer sich an der Diskussion beteiligen. Folgende Werkstattgespräche wurden angeboten und sind als Aufzeichnung im YouTube-Kanal von IIP-Ecosphere verfügbar:

- Das erste Werkstattgespräch fand am 18. Januar 2022 live vom Demonstrator zu KI-basierter Prozess-Stabilität in der Glasformung beim Konsortialpartner Gerresheimer am Standort Bünde statt.
- Das zweite Werkstattgespräch zum Demonstrator Autonome Roboterzelle zur Leiterplattenendprüfung wurde live am 29. Juni aus der Werkhalle der Sennheiser electronic GmbH & Co. KG am Standort Wedemark übertragen.
- Das dritte Werkstattgespräch zum Demonstrator KI-basierte Assistenzsysteme zur Zyklusoptimierung wurde am 18. Oktober mit GILDEMEISTER und Marposs Monitoring Solutions aus der Maschinenhalle des Instituts für Fertigungstechnik und Werkzeugmaschinen (IFW) in Garbsen gesendet.

Das Industrie 4.0 und KI-Event zu Demonstratoren und Anwendungsfällen wurde im Rahmen des *Nürnberg Digital Festivals* am 07. Juli 2023 in der *IHK Nürnberg* als Symposium mit Fachmesse veranstaltet, bei welchem wir über 85 Besucher begrüßen durften. Der Titel des Events war „Die Zukunft der intelligenten Fertigung? Anwendungsfälle zu Industrie 4.0 und KI in der Praxis“¹³. Sowohl etablierte Unternehmen wie *Sennheiser*, *Diehl Metering* oder auch *Schaeffler* waren mit Vorträgen zu KI-Anwendungsfällen als auch als Besucher vertreten. Aber auch Start-ups aus dem Bereich intelligente Fertigung wie *Alpamayo* und *PCB Arts* präsentierten sich neben weiteren 9 Ausstellern. Die Keynote von Prof. Dr.-Ing. Tim Hosenfeldt, Senior Vice President Central Technologies bei *Schaeffler*, wurde zum Highlight des Events, das von allen Beteiligten sehr intensiv auf Social-Media geteilt und bekannt gemacht wurde. Die *Universität Hannover*, sowie die *Deutsche Messe Technology Academy* waren neben den Multiplikatoren tatkräftig als Unterstützer beteiligt. Mit dem Event konnten nicht nur wichtige Stakeholder vernetzt werden, sondern auch wertvolle Ergebnisse und Erfahrungen aus den aktuellen Anwendungsszenarien in der Praxis geteilt werden. Durch die Einbindung von Startups und regionalen Unternehmen in die IIP-Ecosphere wurde auch das Netzwerk der beteiligten Partner erweitert.

Die jüngste Veranstaltung des RIH-Nürnberg in der Projektlaufzeit trug den Titel „Von ‚Do It Yourself‘ zu Intelligenten Lösungen“ und fand beim Kooperationspartner JOSEPHS am 27. September 2023 statt. Das Event war eine gesonderte Veranstaltung im Rahmen der einmonatigen Evaluationsphase im Innovationslabor des JOSEPHS in der Innenstadt von Nürnberg und diente dem Öffentlichkeitstransfer des KI-Innovationsbaukasten. Die Vorträge fokussierten insbesondere die neuesten Erkenntnisse des Think Tanks Geschäftsmodelle zu intelligenten Agenten und dem IIP-Plattform-Chatbot. Die Veranstaltung beinhaltete außerdem einen interaktiven Workshop, der die Teilnehmenden zu detaillierten Fragen zu den Werkzeugen des KI-Innovationsbaukasten einlud und gleichzeitig zur Evaluation der Ergebnisse genutzt wurde.

In diesem Zusammenhang wurden bereits während der COVID-19-Pandemie regionale Fachveranstaltungen im Bereich der industriellen Produktion identifiziert, die als Anker für eigene RIH-Veranstaltungen

¹³ Link zum Event: <https://www.iip-ecosphere.de/event-intelligente-fertigung/>

gen dienen können. Als Partner wurden unter anderem *Bayern-Innovativ*, das *Mittelstand-Digital Zentrum Augsburg* als auch die *IHK Nürnberg* kontaktiert, um die Inhalte von IIP-Ecosphere zu streuen und bei der Durchführung eines Industrie 4.0 und KI-Demonstratoren-Events mitzuwirken. Auf diese breite Grundlage konnte das Team für die Gestaltung der virtuellen Formate sowie für das große RIH-Event in der Metropolregion Nürnberg im Juli 2023 zurückgreifen.

Die folgende Tabelle gibt einen Überblick über Veranstaltungen, die von IIP-Ecosphere durchgeführt wurden. Arbeitskreistreffen sind hierbei nicht einzeln aufgeführt.

Datum	Veranstaltung
12. und 13.06.2020	Online-Event: Künstliche Intelligenz in der Produktion
27.11.2020	Gemeinsame Organisation eines Events zu KI und Spezialhardware mit dem BMBF-Projekt HAISEM-Lab
11. und 12.02.2021	Künstliche Intelligenz in der Produktion, IIP-Ecosphere in Kooperation mit der Deutsche Messe AG
24.06.2021	IIP-Ecosphere Open Core, Event in Kooperation mit der Deutschen Messe AG
05.10.2021	KI konkret: Predictive Quality in der Produktion
09.11.2021	Mensch und KI – gemeinsam erfolgreich?! Event in Kooperation mit der Deutsche Messe AG
24.11.2021	IIP-Ecosphere: Open Core (Vol. 2), Event in Kooperation mit der Deutsche Messe AG
15.12.2021	Auftaktveranstaltung IIP KI-Experimentierfeld
18.01.2022	Werkstattgespräch bei Gerresheimer: Prozess-Stabilisierung mit KI, Event in Kooperation mit der Deutsche Messe AG
24.03.2022	Lernende Systeme für KMU – Best Practices bei der Umsetzung von KI-Use-Cases beim Zukunftszentrum Süd
28.4.2022	Event zum Thema Verwaltungsschalen „Industrie-4.0-Verwaltungsschalen – was steckt dahinter? – gemeinsame Organisation mit dem Projekt HAISEM-Lab
29.06.2022	Werkstattgespräch bei Sennheiser, Event in Kooperation mit der Deutsche Messe AG
05.07.2022	Industrielle Innovationsprozesse der KI – Von Optimierung bis Geschäftsmodellen, Event der Zukunftsallianz Maschinenbau und Deutsche Messe AG in Kooperation mit IIP-Ecosphere
07.09.2022	3. Jahrestagung „KI in der industriellen Produktion“ der Zukunftsallianz Maschinenbau in Kooperation mit IIP-Ecosphere, Deutsche Messe Technology Academy Hannover
18.10.2022	Werkstattgespräch mit Gildemeister und MARPOSS Monitoring Solutions (ARTIS), Event in Kooperation mit der Deutsche Messe AG
13.12.2022	IIP Open Core – Forum KI in der Produktion (vol. 4), virtuell

19.04.2023	„Der heiße Stuhl“: Zwischen Begeisterung und Besorgnis – KI und die Zukunft der Industrie, Event auf der Hannover Messe in Kooperation mit der Deutsche Messe AG
07.07.2023	Die Zukunft der intelligenten Fertigung? Anwendungsfälle zu Industrie 4.0 und KI in der Praxis im Rahmen des Nürnberg Digital Festivals in der IHK Nürnberg
19.09.2023	Abschluss-Symposium IIP-Ecosphere auf der EMO Hannover 2023, Event in Kooperation mit dem VDW und der Deutschen Messe AG
20.09.2023	„Der heiße Stuhl“ vol.2: Zwischen Begeisterung und Besorgnis – KI und die Zukunft der Industrie, Event auf der EMO Hannover 2023 in Kooperation mit der Deutsche Messe AG
27.09.2023	Von „Do It Yourself“ zu Intelligenten Lösungen

Öffentlichkeitsarbeit und Medien (UAP 13.2)

Die folgende Liste fasst die wichtigsten Aktivitäten im Bereich der Öffentlichkeitsarbeit zusammen:

- Die Website wurde stetig erweitert und die meisten Unterseiten vollständig neugestaltet.
- Während der Projektlaufzeit wurden dreizehn Newsletter und neun gesonderte Veranstaltungshinweise an zuletzt 287 Empfänger verschickt.
- Der Social-Media-Auftritt auf Twitter hatte zuletzt 189 Follower. Im Jahr 2022 wurden 75 eigene Tweets veröffentlicht, im Jahr 2023 bis Oktober 39. Der LinkedIn-Auftritt hat zuletzt 187 Follower, es wurden im Jahr 2022 54 eigene Beiträge und im Jahr 2023 bis Oktober 28 eigene Beiträge veröffentlicht. Auf dem YouTube-Kanal wurden insgesamt 45 Videos veröffentlicht.
- In Zusammenarbeit mit dem VDW und der EMO Hannover wurden vier exklusive Artikel auf den Kanälen der Weltleitmesse für Produktion EMO Hannover veröffentlicht.
- Die Werkzeugmaschinenbranche wurde speziell mit regelmäßigen Beiträgen über IIP-Ecosphere im Branchenreport des VDW und der Zeitschrift „Industry Arena“ erreicht.
- Das L3S nutzt sein Forschungsmagazin Binaire (Auflage: bis 2000), um regelmäßig über IIP-Ecosphere zu berichten. Das Magazin wird hauptsächlich an Unternehmen, Multiplikatoren und Politiker verschickt. Ebenfalls werden auch die Magazine des PZH zur Ansprache des Netzwerkes genutzt, so beispielsweise das Magazin des Mittelstand Kompetenzzentrums Hannover „Zukunft Digital“ mit einem Beitrag zum KI-Lösungskatalog.
- Die RIH-Partner wurden mit Material zur zielgruppengerechten Ansprache von Unternehmen unterstützt. Insbesondere die IHK Hannover, MIN, DMTA und VDW stellten hierfür Kanäle wie Newsletter, Verbandszeitschriften und Twitter-Kanäle zur Verfügung.

Aufbau IIP-Modellwelt (UAP 13.3)

Ursprünglich war in IIP-Ecosphere der Aufbau einer Modellwelt als Demonstrator geplant, um die Ergebnisse des Projekts auf Messen und bei Events anschaulich darstellen zu können. Nach Diskussion unterschiedlicher Optionen mit beteiligten Unternehmen hat sich aber herausgestellt, dass es in Anbetracht der Zielgruppe sinnvoller ist. Einen etwas anderen Ansatz zu verfolgen: Um in der Zielgruppe

akzeptiert zu werden und Aufmerksamkeit zu erhalten, ist es wichtig, dass ein Demonstrator anders als die vorgesehene Modellwelt, reale Industriekomponenten enthält wie etwa Cobots, Edge Devices und Industrie-PCs. Es wurde deshalb in enger Zusammenarbeit mit UAP 11.2 (Experimentierfeld) eine flexible Demonstratorumgebung geplant, die es erlaubt IIP-Ecosphere Ergebnisse auf Messen und bei Events zu demonstrieren. Diese besteht aus einem größeren Tisch (siehe auch Abschnitt 2.1.11), zwei Cobots (UR5 Roboter) mit Greifer und Kameras sowie nach Industrie-Standard verbauten weiteren Industrie-Komponenten wie diverse Edge Devices, einem Industrie-PC, einem IT-Server und einer adäquaten Netzwerkanbindung. Über die beschriebenen Funktionalitäten hinaus, ermöglicht der Demonstrator auch als Hardwareplattform zur Validierung und Demonstration der IIP-Ecosphere-Plattform. Diese wurde dazu ebenfalls in den beschriebenen Demonstrator integriert.

Diese flexible Demonstrator-Basis wurde nacheinander für drei Messen in Kooperation mit weiteren Partnern im Projekt in Zusammenarbeit mit unterschiedlichen Arbeitspaketen, insbesondere mit AP 6 und AP 13 zu konkreten Demonstratoren für drei verwandte Demonstrator-Anwendungsfälle ausgebaut. Hierzu wurde ein Demonstrator-Anwendungsfall: Bei dem ersten Anwendungsfall, der auch den weiteren Anwendungsfällen zugrunde liegt, handelt es sich dabei um KI-gestützte visuelle Qualitätsprüfung. Für diesen Anwendungsfall wurde an der LUH der Aufbau vorbereitet, Skripte für die Steuerung des Roboters erstellt und für die KI zum Erkennen von Qualitätsproblemen und zum Erkennen von Abweichungen von der Spezifikation des Produkts Ansätze ausgewählt und effektive Modelle trainiert und evaluiert. Der Aufbau wurde in enger Kooperation mit UHi in mehreren Iterationen getestet und verfeinert, bevor er zur Präsentation auf der Hannover Messe 2022 verwendet werden konnte. Der entwickelte UseCase wurde zudem auch auf den Tagen der digitalen Technologien in Berlin im August 2022 demonstriert mit demselben Aufbau wie auf der Hannover Messe.

Dieser grundlegende Anwendungsfall wurde für die Hannover Messe 2023 erweitert, um ein Framework für das föderiertes Lernen für intelligente industrielle Produktion, das auch in die IIP-Ecosphere Plattform integriert wurde. Für die EMO 2023, wurde der Demonstrator nochmals angepasst und um weitere Funktionalitäten erweitert und zwar um eine lineare Translationsbewegung (7. Achse) des einen Cobot-Arms und um eine Anomalieerkennung für diese Linearchse des Cobot-Arms in enger Kooperation mit dem Projektpartner Lenze. Außerdem wurde das Ergebnis von einem der dynamischen Demonstratoren (siehe UAP 11.3) und zwar eine magnetische Produktidentifikation in den Demonstrator integriert. Die Ergebnisse und Erfahrungen aus dem Aufbau des Demonstrators auf der Grundlage der IIP-Ecosphere Plattform wurden zudem in zwei Publikationen zusammengefasst: [EIC22a, EIC22b].

Eine weitere Aktivität im UAP 13.3 beschäftigte sich mit der Kombination von Roboter und Augmented Reality. Wir haben eine Augmented-Reality-Schnittstelle für die Interaktion mit einem UR5 Cobot Arm implementiert. Über diese Schnittstelle kann ein Mensch dem Cobot über Handgesten Anweisungen geben. Dieses Ergebnis wurde ebenfalls bereits bei unterschiedlichen Ereignissen erfolgreich demonstriert.

Publikation und Verbreitung von Ergebnissen (UAP 13.4)

Die umfassenden Ergebnisse von IIP-Ecosphere wurden in einer großen Breite von Veröffentlichungen, Videos und Vorträgen verbreitet. Siehe dazu Abschnitt 2.4.

2.1.14 Verwertung (AP 14)

Verwertungsplanung (UAP 14.1)

Im Rahmen der Verwertungsplanung fanden Aktivitäten zur gezielten strategischen Planung der Verwertungs- und Verstetigungsaktivitäten statt. In der frühen Projektphase lag der Fokus auf der Schärfung der Wertversprechen und dem Aufbau eines breiten Spektrums von Angeboten, um die Attraktivität dieser Angebote noch während der Projektlaufzeit verproben zu können. Im Rahmen der Schärfung der Wertversprechen wurde beispielsweise in einem von der Begleitforschung organisierten Workshops die Vision von IIP-Ecosphere verfeinert. Auch wurden ein Workshops mit dem Hafven mit Schwerpunkt auf Community-Aufbau durchgeführt, die das weitere Vorgehen zum Aufbau einer Nutzercommunity informiert haben.

Ferner fanden im Rahmen von Klausurtagungen Workshops statt, um die Angebote von IIP-Ecosphere zu schärfen, sowie neue Angebote zu schaffen. In diesem Kontext wurde beispielsweise das Konzept für Werkstattgespräche definiert. Zu den Angeboten die im Rahmen der Verwertungsplanung innerhalb der Projektlaufzeit entwickelt und erprobt wurden, gehören der KI-Lösungskatalog, die IIoT-Plattform, der Datenmarktplatz, Workshops und Innovationstools, Communities of Practice, das KI-Experimentierfeld und weitere.

Für jedes der genannten IIP-Ecosphere Angebote wurden Wertversprechen erarbeitet, die eine klare Kommunikation des Mehrwerts des Angebots ermöglichen. Dies erfolgte in Zusammenarbeit der für die jeweiligen Angebote verantwortlichen Teilprojekte mit dem Think Tank „Geschäftsmodelle“, der hierfür die im Projekt entwickelten Tools „KI-V^{di}P“ und anderer Innovationswerkzeuge einsetzte. Im Zuge der Entwicklung dieser Wertversprechen wurden ebenfalls Rahmenbedingungen für die einzelnen Angebote sowie den Gesamtkontext der IIP-Ecosphere erfasst, was wiederum in den Planungsprozess einfluss.

Auf Basis der entwickelten Wertversprechen wurde für das KI-Experimentierfeld, für den IIP-Datenmarktplatz, als auch für die IIP-Plattform ein Ökosystem-Design-Workshop durchgeführt. Dadurch konnten die Angebote nicht nur weiterentwickelt werden, sondern wurden genutzt, um Rollenbilder für Geschäftsmodelle zu entwickeln, die in [KIR23] veröffentlicht wurden und im KI-Innovationsbaukasten zur Verfügung stehen.

Auf Basis der Erkenntnisse und Erfahrungen aus der o.g. Erprobung der Angebote wurde in Abstimmung mit den Konsortialpartnern schließlich das Szenario fokussiert, in welchem die IIP-Ecosphere IIoT-Plattform im Zentrum der Verwertung steht und die weiteren Angebote der IIP-Ecosphere darauf ausgerichtet sind, den Mehrwert des Angebots IIP-Plattform zu ergänzen und zu steigern. Dabei wurde ein instituts- und disziplinenübergreifendes Team gegründet, das das Ziel verfolgt auf die IIP-Plattform nach Projektende weiterzuentwickeln, im konkreten Anwendungsfällen zu validieren und perspektivisch auch auf der Basis der IIP-Plattform auszugründen. Der Think Tank Geschäftsmodelle unterstützte das erweiterte IIP IIoT-Plattform Team im Zuge einer Innovationsberatung durch Innovations- und Strategiewerkshops für die Ausgestaltung eines Geschäftsmodells unter Mitwirkung des Ökosystems. Außerdem wurde mit dem erweiterten IIP-Plattform Team eine Roadmap mit Meilensteinen erarbeitet, welche den Versuch eines Gründungsvorhabens auch über das Projekt hinaus unterstützt.

Mit dem Ziel der Ausgründung auf Basis der IIoT-Plattform wurde in den Workshops eine zweistufige Strategie entwickelt. In der ersten Stufe (1) wird die Validierung und Stabilisierung der Plattform angestrebt. Bei der IIoT-Plattform handelt es sich um ein Komplexes Softwareprodukt, welche vor einem Marktgang noch weiterentwickelt werden muss. Dies erfolgt anwendungsnah beispielsweise im Rahmen der Umsetzung von Use-Cases mit assoziierten Partner und im Validierungsprojekten (bspw. VIP+). In der zweiten Stufe (2) erfolgt eine Ausgründung und Erweiterung durch Dienste rund um die Plattform.

Vorbereitung nachhaltiger Betrieb und Weiterentwicklung des Geschäftsmodelle (UAP 14.2/14.3)

Im Rahmen der Vorbereitung des nachhaltigen Betriebs erfolgten Aktivitäten zur Stärkung und Bereitstellung der Mehrwertdienste. So wurde die Entwicklung des KI-Lösungskataloges abgeschlossen und der Katalog wurde über unterschiedliche Kanäle beworben. Auf Basis der erarbeiteten Wertversprechen wurden die Angebote aktiv beworben und die Anwenderakquise vorangetrieben. Auch die Aktivitäten im Rahmen des Community-Aufbaus (vgl. 1.9.3) dienen der Vorbereitung des nachhaltigen Betriebes.

In einem nächsten Schritt wurden im Rahmen eines ganztätigen Workshops in Hannover prototypische Geschäftsmodelle für die verschiedenen Angebote der IIP-Ecosphere skizziert und einzelne Angebote zueinander in Beziehung gesetzt. Basierend darauf wurden verschiedene Szenarien für die Verwertung des Gesamtgeschäftsmodells erarbeitet.

Mit der Priorisierung der Verwertung auf die IIP IIoT-Plattform erfolgte eine Fokussierung dieser Aktivitäten auf eine Geschäftsmodell für die IIoT-Plattform. Arbeitstreffen beschäftigten sich mit der Erarbeitung eines Wertversprechens, passenden Umsatzmodellen, Launch- und Wachstumsstrategien sowie einer Diskussion der in UAP 14.3.1 erhobenen Verwertungsmodelle. Zu Themen der Positionierung der Lösung und des Ökosystems im Markt sowie Möglichkeiten zur Verstetigung, wie etwa einer potenziellen Ausgründung, war der Think Tank zusammen mit dem erweiterten IIP-Plattform Team mit dem KI-Bundesverband sowie der Begleitforschung in Kontakt.

Zur Stärkung der Plattform fand auch eine Fokussierung des Ideenwettbewerbes für die Dynamischen Demonstratoren auf die Plattform statt. Im Rahmen der Dynamischen Demonstratoren entstanden so zwei weitere Use-Cases aus der Industrie, die mit Hilfe der Plattform umgesetzt wurden. Auch Aktivitäten zu den Messedemonstratoren und Messepräsentationen hatten einerseits die Weiterentwicklung der Plattform zum Ziel aber auch die Sicherstellung der Bedarfsgerechten Entwicklung, die durch engen Austausch mit Stakeholdern und potentiellen Nutzern auf dem Messen gewährleistet wird. Abschließend fand zur Vorbereitung eine Entwicklung eines neuen Markenauftrittes inklusive Website für die IIoT-Plattform nach Ende der Projektlaufzeit statt. Diese wird nach Projektende unter dem Namen „oktoflow“ weiterentwickelt¹⁴.

Individuelle Pläne für die wirtschaftliche Verwertung (UAP 14.4)

In Rahmen der individuellen Verwertung führten die Partner Aktivitäten durch, um die Verwertung der Projektergebnisse in Form neuer oder verbesserter Produkte und Services sowie Verbesserungen der in der eigenen Produktion sicherzustellen. Nähere Angaben hierzu sind den einzelnen Erfolgskontrollberichten der Partner zu entnehmen, da die konkreteten Planungen zur Verwertung der Ergebnisse in einzelnen Fällen als Vertrauliche Informatoin einzustufen sind.

¹⁴ <https://oktoflow.iip-ecosphere.de/>

Individuelle Pläne für die wissenschaftliche Verwertung (UAP 14.5)

Als wichtige Voraussetzung für die wissenschaftliche Anschlussfähigkeit wurde frühzeitig im Projekt ein Schwerpunkt auf OpenSource gelegt. Insbesondere die IIP-Ecosphere IIoT-Plattform steht vollständig offen zur Verfügung aber auch Artefakte wie die Software der Messedemonstratoren wurden unter einer Open Source Lizenz veröffentlicht. Im Bereich der Produktionstechnologie erfolgte beispielsweise die Veröffentlichung von Datensätzen. Diese Maßnahmen stellen sicher, dass eine wissenschaftliche Verwertung in weiterführenden Projekten und auch über das Konsortium hinaus erfolgen kann.

Im Rahmen der individuellen wissenschaftlichen Verwertung erfolgte darüber hinaus insbesondere durch die Forschungspartner die Identifikation von Fördermöglichkeiten und die Beteiligung an Projektausschreibungen mit dem Ziel die in IIP-Ecosphere erarbeiteten wissenschaftlichen Erkenntnisse sowohl im wissenschaftlichen Kontext weiterzuentwickeln als auch den Transfer aus der Forschung zu verstetigen. In diesem Kontext wurden durch die Projektpartner Mehrere Forschungsprojekten eingeworben, die direkt auf den Erfahrungen von IIP-Ecosphere aufbauen und teilweise die im Projekt etablierte Zusammenarbeit von Partnern fortführen. Im Einzelnen sind diese auf Grund der Vielzahl an Initiativen den Erfolgskontrollberichten der Partnern zu entnehmen. Nachfolgend werden einige herausragende Beispiele genannt:

Im Bereich der transferorientierten Vorhaben sind dabei insbesondere die erfolgreiche Einwerbung des ProKI-Zentrums Hannover und die des Innovation Hub DAISEC zu nennen. Die LUH baut in dem im Oktober 2022 gestarteten Vorhaben ProKI ein Demonstrations- und Transferzentrum für KI in der Zerspanung auf. In das ProKI-Zentrum fließen Erfahrungen aus IIP-Ecosphere ein und auch werden IIP-Ecosphere Demonstratoren mit dem Schwerpunkt Zerspanung im Rahmen des Zentrums weiterentwickelt und einer größeren Breite an Unternehmen zugänglich gemacht.

Ein weiteres wichtige Transfer-Projekt, welches auf den Erfahrungen und Ergebnissen des Projekts IIP-Ecosphere beruht, ist der European Digital Innovation Hub „DAISEC“ unter Leitung des Forschungszentrums L3S, welcher Unternehmen im Einsatz von KI und Zusätzlich auch im Bereich Sicherheit unterstützt. Die Produktion bildet dabei einen der Anwendungsbereiche, auf den sich DAISEC konzentriert. Aus IIP-Ecosphere werden hier insbesondere Erfahrungen im praktischen Einsatz von KI, die aktive Vernetzung mit den Multiplikatoren und Unternehmen der Region sowie Erkenntnisse über Unterstützungsbedarfe in das Projekt eingebracht.

Ebenfalls hervorzuheben ist die Beteiligung der Partner LUH, UHi, DMG, Lenze und ISS im Konsortium des Projektes Factory-X der Manufacturing-X Initiative des Bundes. Factory-X fokussiert den Aufbau eines föderierten Datenraumes für den deutschen Maschinenbau und die Wertschöpfung in Ökosystemen. In die Initiative bringen die Partner Erkenntnisse und Vorarbeiten zu beispielsweise Demonstratoren aus IIP-Ecosphere ein. Dadurch ist eine Weiterentwicklung der Methoden und Transfer in einen breiten Kreis von Unternehmen gewährleistet.

Basierend auf der IIP-Ecosphere Plattform wurde zudem im Herbst 2023 ein Antrag zum Aufbau einer Innovation Community rund um eine open source IIoT Plattform zur Förderung der Nachhaltigkeit der Produktion eingereicht. Idee ist es hier offene Innovationsprozesse in diesem Bereich zu fördern, um nachhaltige Lösungen einer breiten Palette von Unternehmen zugänglich zu machen. Dieser Antrag befindet sich noch in der Begutachtung.

Neben der Weiterführung der Projektthemen in Drittmittelprojekten stellen auch weitere Aktivitäten wie wissenschaftliche Veröffentlichungen, der Aufbau von Workshopangeboten (bspw. KI-Innovationsbaukasten der FAU) und die Verwertung im Rahmen der Industrienetzwerke (bspw. Production Innovation Network) wichtige Bausteine der wissenschaftlichen Verwertung dar. Die hohe Zahl von ca. neunzig wissenschaftlichen Veröffentlichungen sowie einer Vielzahl an Vorträgen aus IIP-Ecosphere unterstreicht den Erfolg der wissenschaftlichen Verwertung und den damit verbunden Kompetenzaufbau (vgl. Abschnitt 2.4).

2.1.15 Projektmanagement/Governance (AP 15)

Im Rahmen von AP 15 „Projektmanagement / Governance“ wurde die projektweite Koordination der Arbeiten durchgeführt. Im Kreis der Teilprojektleitungen fanden monatliche Statuskonferenzen statt sowie Statusgespräche zur Überwachung des Projektfortschrittes und zur Abstimmung im Projekt. Weiterhin wurden Tools zur Erfassung und Visualisierung der Projektkennzahlen, des Projektfortschrittes und Erfolgsindikatoren (KPI) weiterentwickelt und für die Arbeit im Steering Committee genutzt. Zur Zusammenarbeit im Projekt wurden Kommunikationskanäle wie unterschiedliche Mailverteiler, ein Cloud-Repository, Projektinformationen, Kontaktlisten geführt. Klausurtagungen zur Abstimmung innerhalb des Konsortiums wurden vorbereitet und durchgeführt. Bei den Klausurtagungen werden Ergebnisse aus den Teilprojekten vorgestellt und diskutiert als auch konsortiumsinterne Workshops zur Weiterentwicklung des Ökosystems durchgeführt.

Die strategische Steuerung des Projektes wurde durch ein fünfköpfiges Steering Committee (SC), bestehend aus den geschäftsführenden Leitungen der Institute L3S, IFW, SSE (UHI), FAU ermöglicht. Das SC befasste sich in etwa achtwöchigen Abständen mit aktuellen strategischen Themen im Projekt. Ein weiteres Gremium war das Industrial Advisory Board (IAB), bestehend aus: Herr Dr. Wladimir Bickel (MTU Maintenance), Dr. Jan Friedrichs (Digitalagentur Niedersachsen), Frau Anja Simon (LNI 4.0), Edgar Schüber (logicline GmbH), Christan Abt (HEITEC PTS GmbH), Dr. Karl Doreth (Doreth Engineering) und Dr. Michael Klimczak (Schaeffler Technologies). Das IAB deckt damit Kompetenzen aus den Bereichen Produktion und Automatisierung, Plattformen und Standardisierung ab und konnte so einen wertvollen Beitrag als Sparringspartner für das Projekt bieten. Das IAB führte jährliche Projektreviews durch. Auf Basis dessen wurde ein Review-Bericht verfasst und im Projekt verbreitet. Ebenfalls Teil der Arbeiten im AP 15 war die Koordination von jährlichen Meilensteintreffen sowie die Gewinnung und Abstimmung mit potentiellen neuen assoziierten Partnern.

2.2 Voraussichtlicher Nutzen und Verwertbarkeit

Besonders die intensive Beteiligung der Industriepartner hat gezeigt, dass weiterhin hohe Relevanz und Aktualität der Projekthinhalte aus Sicht der Wirtschaft gegeben sind. Von besonderer Relevanz ist die Weiterführung der Open Source IIoT-Plattform zu sehen, die im Projekt entstanden ist (IIP-Ecosphere Plattform). Für die weitere Verwertung dieser Plattform wurde ein tragfähiges Geschäftsmodell entwickelt, das die Verwertungsperspektiven bestätigt. Die Plattform verfügt über einige Alleinstellungsmerkmale (insbesondere die innovative Nutzung von Verwaltungsschalen, der low-code Ansatz, die Code-Generierung, die Erweiterbarkeit und die Nutzung aktueller und relevanter Standards), die einen großen Nutzen für die fertigende Industrie bieten. Für die zukünftige Verwertung der IIP-Ecosphere Plattform ist eine Ausgründung geplant. Hierfür wurden im Projekt bereits vorbereitende Schritte eingeleitet. Dazu

zählen die Erprobung der Plattform mit Industriepartnern, die bereits erwähnte Geschäftsmodellentwicklung sowie ein durchgeführtes Branding der Plattform als „oktoflow“ Plattform mit zugehörigem Logo, Namensregistrierung (Markenschutz) und Corporate Design (Logo siehe Abbildung).



Abbildung 98: Neues Logo der Plattform

Bei der Diskussion zur weiteren Verwertung der Plattform ist jedoch auch klar geworden, dass vor der Ausgründung eine Übergangsfinanzierung für den Übergang vom Prototyp zum Beta-Version eines Produktes notwendig ist. Hierzu wurde bereits ein Antrag eingereicht und ein weiterer Antrag im Rahmen geeigneter Förderlinien ist in Vorbereitung. Gleichzeitig wird bereits mit Unternehmen und Forschungseinrichtungen zusammengearbeitet, um weitere Anwendungen zu erarbeiten und zu implementieren. Weiterhin gibt es Aktivitäten, um die Sichtbarkeit der Plattform zu vergrößern und die Community zu erweitern (Web Site + Social Media) - dabei wird die Plattform kontinuierlich verbessert und erweitert (z. B. neue Konnektoren).

Auf Basis der großen Zahl erfolgreich umgesetzter Projekteinhalte, sind auch im Bereich der KI-Methoden Nutzen und Verwertungspotential als sehr gut anzusehen. In allen Demonstratoren und Think Tanks konnten bestehende technische Herausforderungen gelöst und erfolgreich KI-Ansätze implementiert werden. Eine Weiternutzung oder Überführung in neue Produkte und Dienstleistungen hat sich bei einer Vielzahl der im Projekt betrachteten Anwendungsfälle als machbar herausgestellt und ist geplant.

Die Ergebnisse der Think Tanks legen beispielsweise Grundlagen zur maschinenübergreifenden Nutzung von Prozesswissen und erleichtern damit die Entwicklung und Markteinführung datengetriebener Parametrierungs- und Überwachungslösungen bei Werkzeugmaschinen. Die Methoden wurden im industriellen Umfeld erprobt und können bei Herstellern von Maschinen und Überwachungssystemen Anwendung finden. Die im Rahmen des Projektes erzeugten Daten wurden veröffentlicht und stehen damit der Community für die Entwicklung von Produkten zur Verfügung. Die erforschten und dargestellten Methoden zu selbststeuernden Prozessketten und ganzheitlichen Qualitätsregelkreisen dienen insbesondere produzierenden Unternehmen in der spanenden Einzelteil- und Serienfertigung. Der Einsatz von Qualitätsregelkreisen ermöglicht die Fertigung von Werkstücken in hoher Präzision (feinen Toleranzen) und dient zur Minimierung von Nacharbeit und Ausschuss. Unternehmen können KI-modellunterstützte Qualitätsregelkreise in Form eines inneren, äußeren oder kombinierten Regelkreises selbst, mithilfe von externen Softwareherstellern oder in Zusammenarbeit mit Forschungsinstituten im eigenen Produktionsumfeld umsetzen.

Der im Rahmen des Projekts entwickelte Datenmarktplatz steht interessierten Parteien als open-source Software zur Verfügung. Der Datenmarktplatz kann einerseits von interessierten Unternehmen selbst

oder von externen Servicedienstleistern zur Bereitstellung von Datenangeboten implementiert werden. Im Datenmarktplatz können weitere Servicedienstleistungen entstehen, die initial in mehreren Workshops identifiziert wurden. Um die größtmögliche Verwertbarkeit der Arbeiten im Bereich von Datensouveränität und der International Data Spaces zu erreichen, wurden im Projekt allokierte Ressourcen auf den Aufbau und die Weiterentwicklung des zu den IDS konformen open-source Frameworks Eclipse Dataspace Components (EDC) gebündelt. Die Anforderungen der industriellen Produktion spielten im Rahmen dieser Tätigkeiten eine wichtige Rolle. Zukünftig sollen die EDC, gemäß der von VDMA und ZVEI in Auftrag gegebenen Bauplanstudie, die grundlegende Infrastrukturkomponente für entstehende Manufacturing-X Datenräume darstellen und das souveräne Teilen von Daten zwischen verschiedenen Akteuren im Bereich der industriellen Fertigung ermöglichen.

2.3 Relevante Fortschritte/Ergebnisse Dritter

Das IIP-Team führt eine regelmäßige Durchsicht der für dieses Vorhaben relevanten wissenschaftlichen und technischen Arbeiten in den unterschiedlichen Bereichen durch. Ergebnisse von Dritter Seite, die Ziele und Ergebnisse sowie Lösungswege und Vorgehensweisen grundsätzlich verändern, sind bis jetzt nicht bekannt. Für die in den Think Tanks entwickelten Methoden wurde stets der aktuelle Stand der Technik berücksichtigt. Dafür wurden wissenschaftlichen Arbeiten, national wie international, verfolgt. Ergebnisse von Dritter Seite haben sich dabei lediglich auf technische Details der Methoden ausgewirkt; wesentlichen Änderung der Ziele und Vorgehensweise wurden nicht vorgenommen.

2.4 Veröffentlichung der Ergebnisse

Erfolgte Veröffentlichungen

- [ABO22] Abolfazli, A., Palmer, G., & Kudenko, D. (2022). Data Valuation for Offline Reinforcement Learning. arXiv preprint arXiv:2205.09550. <https://doi.org/10.1038/s41598-022-26534-y>
- [ADH23] Adhisantoso, Y. G., Xuan, Q. L., Kellermann, C., Munderloh, M., & Ostermann, J. (2022). Introduction to deep degradation metric in smart production ecosystems. In 16th CIRP Conference on Intelligent Computation in Manufacturing Engineering (CIRP ICME).
- [AKA22] Akar, Tolga, et al. "Open set recognition for time series classification." *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Cham: Springer International Publishing, 2022.
- [ALA23] Alamoush, Ahmad; Eichelberger, Holger (2023): Adapting Kubernetes to IIoT and Industry 4.0 protocols - An initial performance analysis. In: Softwaretechnik-Trends, 43 (1), pp. 41-43, 2023, ISSN: 0702-8928.
- [ARA21] Arango, Sebastian Pineda, et al. "HPO-B: A Large-Scale Reproducible Benchmark for Black-Box HPO based on OpenML." *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*. 2021.
- [BEC22] Becker, J. M.: (2022): Produktionstechnische Expertensysteme: Wälzschalen mit Künstlicher Intelligenz, PZH 2022, Das Magazin des Produktionstechnischen Zentrums der Leibniz Universität Hannover / Jahresbericht 2021, S. 18.
- [BON21] Bonhage, Malte; Wilkens, Rainer; Denkena, Berend; Kemp, Daniel. „Der digitale Zwilling als Basis für ein intelligentes und skalierbares Produktionssystem“. SPS Magazin, vol. 6, 2021.
- [BRI22] Brinkmeyer, Lukas, et al. "Few-shot forecasting of time-series with heterogeneous channels." *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Cham: Springer Nature Switzerland, 2022.
- [BRU21] Fabian Bruckner,; Jahnke, Nils (2021): Datenschutz und Datensicherheit in Datenökosystemen. IIP-Ecosphere Whitepaper. doi:10.5281/zenodo.4588330.

- [CAS21] Casado, Miguel Gomez; Eichelberger, Holger (2021): Industry 4.0 Resource Monitoring - Experiences With Micrometer and Asset Administration Shells. In: CEUR-WS Proceedings of Symposium on Software Performance 2021 (SSP'21), CEUR-WS.org, 2021.
- [DEL23] Delventhal, Marlene „EU-Gesetzgebungsinitiative auf dem Prüfstand der Praxis: Folgen des Data Acts für IIoT-Plattformen in der intelligenten Produktion“. IT&Production, online: <https://www.it-production.com/allgemein/was-der-data-act-fuer-iiot-plattformen-bedeutet/>, 2023
- [DEN20a] Denkena B, Bergmann B, Reimer S, Schmidt A, Stiehl T, Witt W (2020): „KI-gestützte Prozessüberwachung in der Zerspanung“, ZWF 2020;115:295-298.
- [DEN20b] Denkena, B., Nejd, W., Stiehl, T. H.: „Ein digitales Ökosystem für die intelligente Produktion“. Technologie-Informationen 2020, 20-21.
- [DEN21] Denkena B, Bergmann B, Stiehl TH. (2021): „Artificial Wear for the Assessment of Monitoring Performance“, 54th CIRP conference on Manufacturing Systems 2021.
- [DEN21a] Denkena, B.; Bergmann, B.; Stiehl T.H. "Transfer of Process References between Machine Tools for Online Tool Condition Monitoring." Machines 2021, 9, 282. <https://doi.org/10.3390/machines9110282>
- [DEN21b] Denkena, B.; Bergmann, B.; Becker, J. M.; Reimer, S.; Stiehl, T. „Maschinen fehlerlos einfahren - Ausreißererkennung für Einzelteil- und Kleinserien“, IT&Production online, 22.04.2021, 4 Seiten.
- [DEN21c] Denkena, B.; Bergmann, B.; Becker, B.; Stiehl, T.H. "Time Series Search and Similarity Identification for Single Item Monitoring". In: Behrens B.A., Brosius A., Drossel W.G., Hintze W., Ihlenfeldt S., Nyhuis P. (eds) Production at the Leading Edge of Technology. WGP 2021. Lecture Notes in Production Engineering. Springer, Cham. DOI: 10.1007/978-3-030-78424-9_53
- [DEN21d] Denkena, B.; Bergmann, B.; Stiehl, T.H. "Artificial Wear for the Assessment of Monitoring Performance" Procedia CIRP, 2021, S. 1023-1028. DOI: 10.1016/j.procir.2021.11.172
- [DEN21e] Denkena, B.; Dittrich, M.-A.; Fohlmeister, S.; Kemp, D.; Palmer, G. "Scalable cooperative multi-agent- reinforcement-learning for order-controlled on schedule manufacturing in flexible manufacturing systems." 19. ASIM Fachtagung: Simulation in Produktion und Logistik, 2021, S. 305–314.
- [DEN21f] Denkena, B., Schinkel, F., Pirnay, J., Wilmsmeier, S. (2021): "Quantum algorithms for process parallel flexible job shop scheduling.". CIRP Journal of Manufacturing Science and Technology, Vol. 33, 2021, S. 100-114. DOI: 10.1016/j.cirpj.2021.03.006
- [DEN22a] Denkena, B., Klemme, H., Becker, J., Blech, H. "User-Specific Parameterization of Process Monitoring Systems" Production Engineering, 2022, 8
- [DEN22b] Denkena, B., Klemme, H., Becker, J.: (2022) „KI-basierte Überwachung und Parametrierung von Wälzschälzyklen im Projekt IIP Ecosphere“, VDW Branchenreport (2022), S. 8-11.
- [DEN22c] Denkena, B., Klemme, H., Becker, J. M., Litwinski, K.,(2022): Neuronale Netze: Von der Bildverarbeitung zur Prozessüberwachung, Emo Hannover (2022), online erschienen am 07. November 2022, 11 Seiten.
- [DEN23a] Denkena, B.; Klemme, H., Stiehl, T. H.: „Überwachung von Werkzeugverschleiß – Maschinenübergreifende Nutzung von Prozessdaten mithilfe von Maschinellem Lernen“. ZWF 118 (2023), 298-301.
- [DEN23b] Denkena, B.; Bergmann B.; Stiehl, T. H. (2023): „Identical NC-code on Different Machine Tools – Similarities and Differences in Timing and Positioning“. 16th CIRP Conference on Intelligent Computation in Manufacturing Engineering CIRP ICME, 13.-15. Juli 2022, Neapel, Italien.
- [DEN23f] Denkena, B., Klemme, H., Becker, J.: (2023): KI-basierte Überwachung und Parametrierung von Wälzschälzyklen im Projekt IPP-Ecosphere, VDW Jahresbericht 2022 (2023), S. 25-28
- [DIT20] Dittrich MA, Wilmsmeier S.: „Continuous and multi-criterial scheduling in job shop production“, CIRP Annals Manufacturing Technology 70(1)
- [DRU23] Drumond, Rafael Rego, Lukas Brinkmeyer, and Lars Schmidt-Thieme. "Few-shot human motion prediction for heterogeneous sensors." *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Cham: Springer Nature Switzerland, 2023.
- [EIC20] Eichelberger, Holger; Ahmadian, Amir Shayan; Dewes, Andreas; Ehl, Marco; Staciwa, Monika; Casado, Miguel Gómez (2022): IIP-Ecosphere Platform Handbook v0.2.0. IIP-Ecosphere Whitepaper. doi:<https://doi.org/10.5281/zenodo.5168946>.

- [EIC21a] Eichelberger, Holger; Sauer, Christian; Ahmadian, Amir Shayan; Schick Tanz, Michael; Dewes, Andreas; Palmer, Gregory; Niederée, Claudia (2021): IIP-Ecosphere Plattform – Anforderungen (Funktionale und Qualitäts-Sicht). IIP-Ecosphere Whitepaper. doi:10.5281/zenodo.4485774.
- [EIC21b] Eichelberger, Holger; Ahmadian, Amir Shayan; Dewes, Andreas; Ehl, Marco; Staciwa, Monika; Casado, Miguel Gómez (2022): IIP-Ecosphere Platform Handbook v0.3.0. IIP-Ecosphere Whitepaper. doi:https://doi.org/10.5281/zenodo.6620882.
- [EIC22a] Eichelberger, H.; Palmer, G.; Reimer, S.; Hieu Do, T. T. Vu; Laridi, S.; Weber, A.; Niederée, C.; Hildebrandt T., "Developing an AI-enabled IIoT platform - An early use case validation", In: Software Architecture. ECSA 2022 Tracks and Workshops Bd. 13928. S. 265-283. Springer International Publishing.
- [EIC22b] Eichelberger, H.; Palmer, G.; Niederée, C. (2022): Developing an AI-enabled Industry 4.0 platform - Performance experiences on deploying AI onto an industrial edge device, Symposium on Software Performance.
- [EIC22c] Eichelberger, H.; Reimer, S.; Niederée, C.; Palmer, G. (2022): Virtuelle IIoT-Plattform für die Digitalisierung der Fertigung, ZWF, Vol. 117 (2022), Nr. 12, S. 884-887.
- [EIC22d] Eichelberger, Holger; Ahmadian, Amir Shayan; Dewes, Andreas; Ehl, Marco; Alamoush, Ahmad; Staciwa, Monika; Casado, Miguel Gómez (2022): IIP-Ecosphere Platform Handbook v0.4.0. IIP-Ecosphere Whitepaper. doi:https://doi.org/10.5281/zenodo.7047640.
- [EIC22e] Holger Eichelberger, and Heiko Stichweh; Sauer, Christian (2022): Requirements for an AI-enabled Industry 4.0 Platform – Integrating Industrial and Scientific Views, ISBN: 978-1-61208-946.
- [EIC23a] Eichelberger, Holger; Ahmadian, Amir Shayan; Dewes, Andreas; Ehl, Marco; Alamoush, Ahmad; Staciwa, Monika; Casado, Miguel Gómez (2023): IIP-Ecosphere Platform Handbook v0.5.0. IIP-Ecosphere Whitepaper. doi:https://doi.org/10.5281/zenodo.7695124.
- [EIC23b] H. Eichelberger and C. Niederée, "Asset Administration Shells, Configuration, Code Generation: A power trio for Industry 4.0 Platforms," 2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA), Sinaia, Romania, 2023, S. 1-8, doi: 10.1109/ETFA54631.2023.10275339.
- [EIC23c] Eichelberger, Holger; Ahmadian, Amir Shayan; Dewes, Andreas; Ehl, Marco; Alamoush, Ahmad; Staciwa, Monika; Casado, Miguel Gómez (2023): IIP-Ecosphere Platform Handbook v0.6.0. IIP-Ecosphere Whitepaper. doi:https://doi.org/10.5281/zenodo.8429785.
- [FAL22] Falkner, Jonas K., et al. "Learning to Control Local Search for Combinatorial Optimization." *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Cham: Springer Nature Switzerland, 2022.
- [FAU22] Faubel, Leonhard; Schmid, Klaus; Eichelberger, Holger (2022): Is MLOps different in Industry 4.0? General and Specific Challenges. 3rd International Conference on Innovative Intelligent Industrial Production and Logistics (IN4PL).
- [FEI23] Feichtinger, Kevin; Meixner, Kristof; Rinker, Felix; Koren, István; Eichelberger, Holger; Heine mann, Tonja; Holtmann, Jörg; Konersmann, Marco; Michael, Judith; Neumann, Eva-Maria; Pfeiffer, Jérôme; Rabiser, Rick; Riebisch, Matthias; Schmid, Klaus (2023): Software in Cyberphysischen Produktionssystemen - Herausforderungen zur Umsetzung in der Industrie Journal Article. In: ATP-Magazin, 2023 (4), pp. 62-68, 2023, ISSN: 2190-4111.
- [GAR21a] Dold, Dominik; Garrido, Josep Soler (2021). An energy-based model for neuro-symbolic reasoning on knowledge graphs. In 20th IEEE International Conference on Machine Learning and Applications (ICMLA), 2021.
- [GAR21b] Garrido, Josep Soler; Dold, Dominik; Frank, Johannes (2021). Machine learning on knowledge graphs for context-aware security monitoring. In 2021 IEEE International Conference on Cyber Security and Resilience (CSR), 2021.
- [GAR21c] Garrido, Josep Soler; Dold, Dominik; Chian, Victor Caceres; Hildebrandt, Marcel; Runkler, Thomas (2022). Neuro-symbolic computing with spiking neural networks. In Neuro-symbolic computing with spiking neural networks, 2022.
- [GRA21] Graf, W., Wilmsmeier, S. „Quantum Technology in Flexible Job Shop Scheduling? A Field Report Using Digital Annealer“, Bitkom Events: Quantum Summit 2021, online, 26.-27.05.2021, 18 Seiten. https://www.youtube.com/watch?v=doylgwUXy-I

- [JAL20] Jalowski, M.; Schymanietz, M.; Möslin, K. (2020). Supporting Participants in Creative Processes: Opportunities for Persuasive Technology in Participatory Design. In: Proceedings of the Forty-First International Conference on Information Systems. Hyderabad, IN.
- [JAL21] Jalowski, Max; Roth, Angela; Oks, Sascha J.; Wilga, Matthäus (2021): Innovation KI-basierter Dienstleistungen für die industrielle Wertschöpfung – Ein artefaktzentrierter Ansatz.

In: Bruhn, Martin; Hadwich, Karsten (Ed.): Künstliche Intelligenz im Dienstleistungsmanagement. Forum Dienstleistungsmanagement., pp. 158-183, 2021, ISBN: 978-3-658-34324-8.
- [JAW21] Jawed, Shayan, et al. "Multi-task learning curve forecasting across hyperparameter configurations and datasets." *Machine Learning and Knowledge Discovery in Databases. Research Track: European Conference, ECML PKDD 2021, Bilbao, Spain, September 13–17, 2021, Proceedings, Part I* 21. Springer International Publishing, 2021.
- [JAW22] Jawed, Shayan, and Lars Schmidt-Thieme. "GQFormer: A Multi-Quantile Generative Transformer for Time Series Forecasting." *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 2022.
- [JAW22A] Jawed, Shayan, and Lars Schmidt-Thieme. "Fqformer: A fully quantile transformer for time series forecasting." *8th SIGKDD International Workshop on Mining and Learning from Time Series–Deep Forecasting: Models, Interpretability, and Applications*. 2022.
- [JAW23] Jawed, Shayan; Madhusudhanan, Kiran; Yalavarthi, Vijaya Krishna, and Lars Schmidt-Thieme (2023): Forecasting Early with Meta Learning. IJCNN 2023, June 18 - June 23, 2023.
- [JOM21] Jomaa, Hadi S.; Schmidt-Thieme, Lars; Grabocka, Josif (2021): Dataset2Vec: Learning Dataset Meta-Features. In: Data Mining and Knowledge Discovery, 10618 (0737), pp. 22, 2021.
- [KEL21a] Kellermann, Christoph, Eric Neumann, and Jörn Ostermann. "A New Preprocessing Approach to Reduce Computational Complexity for Time Series Forecasting with Neuronal Networks: Temporal Resolution Warping." *2021 International Symposium on Computer Science and Intelligent Controls (ISCSIC)*. IEEE, 2021.
- [KEL21b] Kellermann, Christoph und Jörn Ostermann. "Estimation of unknown system states based on an adaptive neural network and Kalman filter." *Procedia CIRP* 99 (2021): 656-661.
- [KEL21c] Kellermann, Christoph, et al. "Introduction to an Adaptive Remaining Useful Life Prediction for forming tools." *2021 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*. IEEE, 2021.
- [KEL22a] Kellermann, C., Selmi, A., Brown, D., & Ostermann, J. (2022). Fault Detection in Multi-stage Manufacturing to Improve Process Quality. *2022 International Conference on Control, Automation and Diagnosis (ICCAD)*, Lisbon, Portugal, S. 1-6, doi: 10.1109/ICCAD55197.2022.9853909
- [KEL22b] Kellermann, C., Neumann, E., & Ostermann, J. (2022). Prediction of variable forecast horizons with artificial neural networks by embedding the temporal resolution warping. *2022 International Conference on Control, Automation and Diagnosis (ICCAD)*, Lisbon, Portugal, S. 1-5, doi: 10.1109/ICCAD55197.2022.9853884
- [KEM21] Kemp, Daniel; Bonhage, Malte; Wilkens, Rainer (2021): Der digitale Zwilling als Basis für ein intelligentes und skalierbares Produktionssystem. In: SPS Magazin, 6 (June) 2021 , 2021.
- [KIR22] Kirschbaum, Julius; Posselt, Tim; and Roth, Angela (2022): USE-CASE-BASED INNOVATION FOR ARTIFICIAL INTELLIGENCE – AN ONTOLOGICAL APPROACH (2022). ECIS 2022 Research-in-Progress Papers. 64. https://aisel.aisnet.org/ecis2022_rip/64
- [KIR23] Kirschbaum, Julius; Posselt, Tim; and Moeslein, Kathrin (2023) "Designing innovation ecosystems with functional roles – The case of industrial, intelligent manufacturing. Handbook on Digital Platforms and Ecosystems in Manufacturing. Edward Elgar Publishing
- [KIR23] Kirschbaum, Julius; Posselt, Tim; and Moeslein, Kathrin (2023) "Designing innovation ecosystems with functional roles – The case of industrial, intelligent manufacturing. Handbook on Digital Platforms and Ecosystems in Manufacturing. Edward Elgar Publishing
- [MAD22] Madhusudhanan, Kiran, et al. "U-Net Inspired Transformer Architecture for Far Horizon Time Series Forecasting." *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. 2022.
- [NIE21a] Niederée, Claudia; Eichelberger, Holger; Schmees, Hans-Dieter; Broos, Alexander; Schreiber, Per. „KI in der Produktion – Quo vadis?“ IIP-Ecosphere Whitepaper, 2021.

- [NIE21b] Niederée, Claudia; Eichelberger, Holger; Schmees, Hans-Dieter; Broos, Alexander; Schreiber, Per. „Management Summary zu KI in der Produktion - Quo vadis?“ IIP-Ecosphere Whitepaper, 2021.
- [PAL21] Palmer G, Kemp D, Fohlmeister S (2021) Scalable cooperative Multi-Agent- Reinforcement-Learning for order-controlled on schedule manufacturing in flexible manufacturing systems. https://github.com/gjp1203/fms_marl (aufgerufen am 14.05.2021).
- [PUS22] Puschky, R. (2022): „Federated Learning – eine datenschutzfreundliche Methode zum Training von KI-Modellen?“ in ZD-Aktuell 2022, <https://beck-online.beck.de/?vpath=bib-data%2fzeits%2fZDAKTUELL%2f2022%2fcont%2fZDAKTUELL.2022.00019.htm>
- [REI22] Reimer, J., Wang, Y., Laridi, S., Urdich, J., Wilmsmeier, S., Palmer, G.: (2022): Identifying cause-and-effect relationships of manufacturing errors using sequence-to-sequence learning, Scientific Reports, Vol. 12 (2022), Nr. 22332, 11 Seiten. DOI: <https://doi.org/10.1038/s41598-022-26534-y>
- [RII23] Riihimäki, J. & Shahzad, K. (2023). The "Dark Side" of Open Innovation in Emerging Markets: An Integrated Framework. Proceedings of the 83rd Annual Meeting of the Academy of Management (AOM). Boston, Massachusetts, USA.
- [SAU21] Sauer, Christian; Eichelberger, Holger; Ahmadian, Amir Shayan; Dewes, Andreas; Jürjens, Jan (2021): Aktuelle Industrie 4.0 Plattformen – Eine Übersicht. IIP-Ecosphere Whitepaper . doi:10.5281/zenodo.4485756
- [SAU23] Sauer, Christian Severin; Eichelberger, Holger (2023): Performance Evaluation of BaSyx based Asset Administration Shells for Industry 4.0 Applications. In: Softwaretechnik-Trends, 43 (1), pp. 47-49, 2023, ISSN: 0702-8928.
- [STI21] Stichweh, Heiko; Sauer, Christian; Eichelberger, Holger (2021): IIP-Ecosphere Platform Requirements (Usage View). IIP-Ecosphere Whitepaper. doi:10.5281/zenodo.4485774
- [STI22] Stiehl, T. H. (2022): „Überwachung von Werkzeugverschleiß in der Kleinserie durch maschinenübergreifende Datennutzung“. IndustryArena
- [STI23] Stiehl, T. H. (2023): Multivariate time series data of milling processes with varying tool wear and machine tools. Mendeley Data. DOI: 10.17632/zpxs87bjt8.2
- [STO22] Stocker, Joachim, Herda, Nils, Jürjens, Jan. (2022) Life cycle and metrics to measure the resilience of business processes by considering resources. Bus. Process. Manag. J. 28(4): 10.1108/BPMJ-10-2021-0674
- [TUM22] Tuma, Katja & Peldszus, Sven & Strüber, Daniel & Scandariato, Riccardo & Jürjens, Jan. (2022). Checking security compliance between models and code. Software and Systems Modeling. 22. 1-24. 10.1007/s10270-022-00991-5.
- [WIL21a] Wilmsmeier, Sören (2021): Taktzeitoptimierung mithilfe von künstlicher Intelligenz Booklet. Phi-hannover. <https://www.phi-hannover.de/forschung/artikel/detail/taktzeitoptimierung-mithilfe-von-kuenstlicher-intelligenz>
- [WIL21b] Wilga, Matthäus; Jalowski, Max; Kirschbaum, Julius; Roth, Angela (2021): A Systematic Characterization of Artificial Intelligence Business Models as a Fundament for Business Model Innovation and Strategic Decision-Making. 21st European Academy of Management (EURAM) Conference 2021, 2021.
- [WIB22] Wißbrock, P., Richter, Y., Pelkmann, D., Ren, Z., & Palmer, G. (2022): "Cutting Through the Noise: An Empirical Comparison of Psychoacoustic and Envelope-based Features for Machinery Fault Detection". arXiv preprint arXiv:2211.01704.
- [WIB23] Wißbrock, Peter; Richter, Yvonne; Pelkmann, David; Ren, Zhao; Palmer, Gregory (2023): Cutting Through the Noise: An Empirical Comparison of Psycho-Acoustic and Envelope-based Features for Machinery Fault Detection. In: ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1-5, IEEE 2023.
- [XUA22] Xuan, Q. L., Adhisantoso, Y. G., Munderloh, M., & Ostermann, J. (2022). Uncertainty-Aware Remaining Useful Life Prediction for Predictive Maintenance Using Deep Learning. In 16th CIRP Conference on Intelligent Computation in Manufacturing Engineering (CIRP ICME).
- [YAL22] Yalavarthi, Vijaya Krishna, Johannes Burchert, and Lars Schmidt-Thieme. "DCSF: Deep Convolutional Set Functions for Classification of Asynchronous Time Series." *2022 IEEE 9th International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE, 2022.

Geplante und eingereichte Veröffentlichungen

- [DEN23a] Denkena, B., Klemme, H., Stiehl, T. H.: "Failure sensitivity and similarity of process signals among multiple machine tools". 56th CIRP Conference on Manufacturing Systems CIRP CMS '23, 24.-26. October 2023, Cape Town, South Africa. (accepted for publication)
- [DEN23b] Denkena, B., Klemme, H., Stiehl, T. H.: „Multidimensional position-based monitoring of machining using references of multiple machine tools". 17th CIRP Conference on Intelligent Computation in Manufacturing Engineering CIRP ICME, 12.-14. Juli 2023, Naples, Italy. (accepted for publication)
- [DEN23c] Denkena, B., Klemme, H., Stiehl, T. H.: „Multivariate time series data of milling processes with varying tool wear and machine tools". Data in Brief. (submitted for review)
- [DEN23d] Denkena, B., Klemme, H., Stiehl, T.H.: „Verschleißüberwachung beim Fräsen durch die maschinenübergreifende Nutzung von Prozessdaten". IndustryArena eMagazine EMO 2023. (eingereicht)
- [DEN23e] Denkena, B., Klemme, H., Becker, J.: „Process Monitoring with Variational Auto Encoders". 23rd Machining Innovations Conference for Aerospace Industry 2023 (MIC 2023), 29 -30. November 2023, Hannover, Germany
- [YAL24] Yalavarthi, Vijaya Krishna, et al. "GraFITi: Graphs for Forecasting Irregularly Sampled Time Series." Accepted for publishing in the 38th AAAI Conference on Artificial Intelligence (AAAI-24).
- [YAL24a] Yalavarthi, Vijaya Krishna, et al. "ProFITi: Probabilistic Forecasting of irregularly sampled time series via conditional flows." submitted for ICLR 2024
- [YAL23] Yalavarthi, Vijaya Krishna, et al. "Tripletformer for probabilistic interpolation of irregularly sampled time series." Accepted for publishing in 2023 IEEE International Conference on Big Data
- [Wiß24] Wißbrock, Peter and Ren, Zhao and Pelkmann, David, More than Spectrograms: Deep Representation Learning for Machinery Fault Detection. Available at SSRN: <https://ssrn.com/abstract=4679856> or <http://dx.doi.org/10.2139/ssrn.4679856> (under Review at Journal 'Applied Acoustics')

Vorträge

- Arnu, David; Klinkenberg, Ralf: "Industrial Data Science Platform and Applications in Electronics and Manufacturing, Online-Event AI in Manufacturing - Finnish-German Collaboration Initiatives von Business Finland und VDI/VDE Innovation + Technik GmbH, 14.12.2021.
- Becker, Jonas: „KI-Assistenzsystem für Werkzeugmaschinen", FDT-Demo Day, Berlin, 11.07.23.
- Brandt, Sebastian (Siemens Technology): "Ontology-based Access to Temporal Data" beim Stream Reasoning meets Cloud-Edge Intelligence Workshop 2022, 18.10.2022.
- Delventhal, Marlene: „Rechtliche Grundlagen bei der wirtschaftlichen Verwertung von KI", Panelvortrag im Webinar „Daten und KI-Modelle als Wirtschaftsgut" der VDI/VDE-IT, 23.06.2020.
- Eichelberger, Holger: „KI-Plattformen: KI-Module, Datenaustauschformate, Semantiken und Service" der Begleitforschung , UHi, 23.05.2022.
- Eichelberger, Holger: „PIN-Konferenz trifft Praxisforum Industrie 4.0", Vorstellung des IIP-Messedemonstrators bei der Gemeinschaftskonferenz des Produktion Innovation Network und der IHK Hannover, 08.06.22.
- Eichelberger, Holger; Schreiber, Per: Skalierbare und vertrauenswürdige KI: Zwei Beiträge bei der Vernetzungsveranstaltung des KI-Innovationswettbewerbes, 24.11.22.
- Hildebrandt, Petra (Sennheiser): "AI - From a Feasibility Study to a Sustainable Solution in Production" Rethink! Smart Manufacturing Europe -13./14.3.2023.
- Hildebrandt, Petra (Sennheiser): –"KI in der Produktion – ein Erfahrungsbericht", deutscher ingenieurinnenbund e.V. 29.3.2023.
- Hildebrandt, Petra (Sennheiser): „Aufwand und Nutzen – Einsatz künstlicher Intelligenz bei der Prüfung von Leiterplatten", Rethink! Smart Manufacturing D/A/C/H: 30.6./1.7.2022.
- Hildebrandt, Petra (Sennheiser): „KI - Von der Machbarkeitsstudie zu einer nachhaltigen Lösung in der Produktion". Veranstaltung „Die Zukunft der intelligenten Fertigung? Anwendungsfälle zu Industrie 4.0 und KI in der Praxis." Symposium mit Fachmesse Nürnberg – 7.7.2023.
- Hildebrandt, Petra: „KI allein genügt nicht!", Künstliche Intelligenz in der industriellen Produktion, Deutsche Messe AG, 24.09.2021.
- Hildebrandt, Petra: „Vollautomatischer Funktionstest bestückter Leiterplatten – Kann KI bei der Fehlerdiagnose helfen?", KI & Data Analytics in Manufacturing, 28.06.2021.

Hildebrandt, Petra; Puschmann, Jörg: "Auto MLOps...? - The path to productive ML model maintenance", Altair Data to Value Symposium, 19.4.2023.

Jalowski, Max: „Supporting Participants in Creative Processes: Opportunities for Persuasive Technology in Participatory Design“, Forty-First International Conference on Information Systems, Hyderabad/virtuell, 15.12.2020

Kirschbaum, Julius: "USE-CASE-BASED INNOVATION FOR ARTIFICIAL INTELLIGENCE – AN ONTOLOGICAL APPROACH" auf der European Conference of Information Systems (ECIS) 2022, Timisoara, Rumänien

Kirschbaum, Julius: „Effektiv KI in Unternehmen nutzen – Question Answering und Information Retrieval mit Large Language Models,,KI-Stammtisch des Mittelstand-Digital Zentrum Franken im Oktober in Nürnberg, 09.10.2023

Kirschbaum, Julius: „Lernende Systeme für KMU – Best Practices bei der Umsetzung von KI-Use-Cases“, Veranstaltung des Zukunftszentrum Süd für die Metropolregion Nürnberg/Erlangen, virtuell, 24.03.2022.

Kirschbaum, Julius: „QA-systems for Information Retrieval - Augmenting the IIP-Platform support service team with LLM-chatbot,, ChatGPT Night, 24.08.2023

Kirschbaum, Julius: „Wie wir in Zukunft mit Daten interagieren – Question-Answering und Information Retrieval mit Sprachmodellen (mit live Demo)“, KI-Stammtisch des KI-Hub-Bayern, IHK-Nürnberg, 09.10.2023

Kirschbaum, Julius; Wilga, Matthäus: "Der KI-Lösungskatalog von IIP-Ecosphere", 3. Jahrestagung „KI in der industriellen Produktion“ der Zukunftsallianz Maschinenbau in Kooperation mit IIP-Ecosphere, Deutsche Messe Technology Academy Hannover

Kirschbaum, Julius; Wilga, Matthäus: „Industrielle Innovationsprozesse der Künstlichen Intelligenz – Von Optimierung bis Geschäftsmodellen“, ZAM Meets AI, virtuell, 05.07.2022.

Niederée, Claudia; Eichelberger, Holger: „Technologie- und Wissenstransfer am Beispiel des Projekts IIP-Ecosphere zur Beschleunigung von KI-Lösungen in der Zusammenarbeit von Wissenschaft und Wirtschaft“, 15. Juli, IHK Ausschuss Industrie & Forschung, 15.07.2021.

Niederée, Claudia; Eichelberger, Holger: „Zutaten zur KI-Beschleunigung: Daten, KI-Methoden, Plattformen und Geschäftsmodelle“, METAV digital, 08.04.2021

Niederée, Claudia; Eichelberger, Holger: Sitzung des AK STD 1941.0.8 Expertenrat für künstliche Intelligenz in industriellen Anwendungen, 18.3.21

Palmer, Gregory: „Wie wertvoll sind meine Daten? Den Datenwert für Machine Learning messen“, Workshop der Begleitforschung: Daten in der Produktion – Datenqualität, Zertifizierung und Haftung: Vortrag von, 22.04.21.

Riihimäki, J. & Shahzad, K. (2023). The "Dark Side" of Open Innovation in Emerging Markets: An Integrated Framework. Proceedings of the 83rd Annual Meeting of the Academy of Management (AOM). Boston, Massachusetts, USA, 06.08.2023

Schreiber, Per: „Deep Dive“ des KI-Bundesverbandes: Vortrag zum IIP-Ecosphere Ökosystem, 08.11.22.

Schreiber, Per: „KI beschleunigt – Wie KI-Lösungen in der Produktion einfacher und schneller umgesetzt werden können“, Jahrestagung KI in der industriellen Produktion: 23.09.21.

Schreiber, Per: „KI beschleunigt – Wie KI-Lösungen in der Produktion einfacher und schneller umgesetzt werden können“, Mittelstandstag Schleswig-Holstein, 29.9.2021.

Schreiber, Per: Online-Event AI in Manufacturing - Finnish-German Collaboration Initiatives von Business Finland und VDI/VDE Innovation + Technik GmbH, 14.12.21.

Schreiber, Per: VDMA Webinar „Forschung für Mittelstand“, Vortrag zu IIP-Ecosphere, 05.10.2021.

Tage der digitalen Technologien: IIP-Ecosphere war mit einem Demonstrator zur visuellen KI-basierten Qualitätsinspektion auf Basis der IIoT-Plattform vertreten, 30.08.22.

Wilga, Matthäus: "Inspirationskarten auf Basis generischer Rollenbilder intelligenter Agenten", 18. Internationale Tagung Wirtschaftsinformatik, Paderborn, 18.09.2023

Wilga, Matthäus: „A Systematic Characterization of Artificial Intelligence Business Models“, 21st European Academy of Management (EURAM) Conference 2021. Montreal/virtuell, 16.06.2021

Wilga, Matthäus; Kirschbaum, Julius: "Der KI-Lösungskatalog von IIP-Ecosphere", 3. Jahrestagung „KI in der industriellen Produktion“ der Zukunftsallianz Maschinenbau in Kooperation mit IIP-Ecosphere, Deutsche Messe Technology Academy Hannover, 08.09.2022

Wilga, Matthäus; Kirschbaum, Julius: „Der KI-Lösungskatalog von IIP-Ecosphere“, 3. Jahrestagung „KI in der industriellen Produktion“ der Zukunftsallianz Maschinenbau in Kooperation mit IIP-Ecosphere, Deutsche Messe Technology Academy Hannover, 08.09.2022

Wilga, Matthäus; Kirschbaum, Julius: „Industrielle Innovationsprozesse der Künstlichen Intelligenz – Von Optimierung bis Geschäftsmodellen“, ZAM meets AI - Industrielle Innovationsprozesse der Künstlichen Intelligenz, 05.07.2022

Wilga, Matthäus; Kirschbaum, Julius: „Lernende Systeme für KMU – Best Practices bei der Umsetzung von KI-Use-Cases“, Zukunftszentrum Süd – Beratung zu Digitalisierung und KI für KMU, 24.03.2022

Wilkens, Rainer (Sennheiser): REFA Talk „Wandlungsfähige Produktion in einer turbulenten Zeit“: Vortrag „Wandlungsfähiges Produktionssystem – Gedanken und Ansätze“, 2.6.22.

Wilkens, Rainer: Industrial Communication Congress / Hannover –, <https://event.phoenixcontact.com/icc2021>, 10.6.2021.

Wilkens, Rainer; Puschmann, Jörg (Sennheiser): Sommercampus - Mehrwerte mit KI schaffen, Wertschöpfungsnetze nachhaltig transformieren,– Panel-Diskussion zu KI im Qualitätsmanagement. Berlin: 21.-22.6.2023

Lange, Dirk (Artis Marposs): Sommercampus - Mehrwerte mit KI schaffen, Wertschöpfungsnetze nachhaltig transformieren,– Panel-Diskussion. Berlin: 21.-22.6.2023

Kirschbaum, Julius: „Es kommt darauf an! – Anwendungsfälle für datenintensives maschinelles Lernen in der Produktion mit Hilfe von Plattformlösungen“, IHK-Webinar „KI & Dateneigentum in der Industrie: Auf der Suche nach klaren Antworten“ im Februar in Nürnberg, 06.02.2024

Weitere öffentliche Artefakte und Mehrwerte für die Community

Im Rahmen des Projektes wurde eine Reihe von öffentlich zugänglicher Software, Tools, Werkzeuge und weiteres erarbeitet. Eine Übersicht über diese „Artefakte“ aus den unterschiedlichen Projektteilen ist in nachfolgender Tabelle dargestellt.

Bezeichnung	Kurzbeschreibung mit Link	Art
IIP-Ecosphere Plattform	Open Source Umsetzung der IIP-Ecosphere Plattform https://github.com/iip-ecosphere/platform	Software / Open-Source Repository
Performance experiences	Developing an AI-enabled Industry 4.0 platform - Performance experiences on deploying AI onto an industrial edge device. Public experimental material https://zenodo.org/records/7157607	Datensatz / Code zu Veröffentlichung
Ressource Monitoring	Industry 4.0 Resource Monitoring - Experiences with Micrometer and Asset Administration Shells. Public experimental material. https://zenodo.org/records/5571817	Datensatz / Code zu Veröffentlichung
Plattform Handbuch	IIP-Ecosphere Platform Handbook. The platform handbook in various versions. https://zenodo.org/records/8429785	Handbuch
Performance Comparision	Data for "Performance comparison of TwinCat ADS for Python and Java". Public experimental material. https://zenodo.org/record/8421817	Datensatz / Code zu Veröffentlichung
Performance Analysis	Adapting Kubernetes to IIoT and Industry 4.0 protocols - An initial performance analysis. Public experimental material. https://doi.org/10.5281/zenodo.7158281	Datensatz / Code zu Veröffentlichung
EASy-Producer	Software configuration and instantiation toolkit, implementation of the Integrated Variability Modeling Language, Variability Instantation	Software / Open-Source Repository

	Language and Variability Template Language, maintained in IIP-Ecosphere as a central component of the IIP-Ecosphere platform. https://github.com/SSEHUB/EASyProducer	
Oktoflow Homepage	Open Source Nachfolger der IIP-Ecosphere Plattform. http://oktoflow.de	Software / Open-Source Repository
IIP-Ecosphere Data Marketplace	Interessierte Parteien erhalten mit dem MVP des föderierten IIP-Ecosphere Data Marketplace die Möglichkeit, selbst einen Datenmarktplatz für das eigene Unternehmen oder den unternehmensübergreifenden Datenaustausch bereitzustellen. Mithilfe der gelieferten Docker Compose-Konfiguration kann einfach in wenigen Minuten eine eigene Instanz implementiert werden. https://www.iip-ecosphere.de/angebote/datenmarktplatz/	Software / Open-Source Repository
Forschungsframework	Framework und Datensatz zur Untersuchung von Multi-Agenten Systemen im Umfeld der Produktionsplanung und Steuerung [PAL21] https://github.com/gjp1203/fms_marl .	Software / Datensatz / Open-Source Repository
Datensatz zu Fräsprozessen	Multivariate Zeitreihendaten von Fräsprozessen mit unterschiedlichem Werkzeugverschleiß und auf unterschiedlichen Werkzeugmaschinen. http://dx.doi.org/10.17632/zpxs87bjt8.2	Datensatz
AI-Use-Case Requirements Taxonomy	Implementing an AI use-case means considering more than just technological requirements and thus involves interdisciplinary teams that define these requirements. A requirements taxonomy for AI-use-cases was developed by Kirschbaum et al. [KIR22] that can be explored below. https://ecosystem-tools.de/AI-Toolbox/AI-requirements-taxonomy.html	Workshop/Tool
Charakterisierung von KI-Geschäftsmodellen	Wilga et al. [WIL21b] developed a framework for the characterisation of AI business models that can be used to explore and define the type of business model an organisation is developing. https://ecosystem-tools.de/AI-Toolbox/AI-business-models.html	Workshop/Tool
AI Business Model Roles in Manufacturing	In the project IIP-Ecosphere , seven archetypical business model roles were developed for the industrial intelligent manufacturing domain. https://ecosystem-tools.de/AI-Toolbox/AI-business-model-roles.html	Workshop/Tool
The AI-V ^{die} P Template for value proposition design	Jalowski et al. [JAL21] developed a tool for developing AI-Value-Propositions . The tool is also available as a Miro-Template . https://ecosystem-tools.de/AI-Toolbox/AI-value-proposition.html	Workshop/Tool
Ecosystem-Designer	Kirschbaum et al. [KIR23] developed a visualisation of innovation ecosystems that can be used for defining the roles of individual actors within the ecosystem. https://roletool.ecosystem-tools.de/Ecosystem-Designer.html	Workshop/Tool
Ecosystem Design Workshop	Based on the Ecosystem Pie Model by Talmar et al.[TAL20], this workshop format allows organisations to systematically develop the ecosystems around complex value propositions. https://miro.com/app/board/uX-jVMxwNqdM=?share_link_id=992684071933	Workshop/Tool

3 Literaturverzeichnis

- [ABD10] K. Abdennadher, P. Venet, G. Rojat, J. Retif, C., Rosset, A Real-Time Predictive-Maintenance System of Aluminum Electrolytic Capacitors Used in Uninterrupted Power Supplies. IEEE Trans. Ind. Appl. 46, 1644–1652, 2010
- [ABE10] J. V. Abellan-Nebot, F. Romero Subirón, A review of machining monitoring systems based on artificial intelligence process models, Int J Adv Manuf Technol 47 (1-4), pp. 237–257, 2010
- [ADA19] Adari S., Falk S., Sampson C., Germany's evolving platform landscape, Accenture / Working Group on Digital Business Models in Industrie 4.0, 2019
- [ALT17] Y. Altintas, D. Aslan, Integration of virtual and on-line machining process control and monitoring. CIRP Annals 66 (1), S. 349 - 352, 2017.
- [AMI19] M. Amini, S. Chang, Review of machine learning approaches for high dimensional process monitoring, Industrial and Systems Engineering Research Conference, 2019
- [ANA18] A. Anand, K. Bizer, A. Erlei, U. Gadiraju, C. Heinze, L. Meub, W. Nejd, B. Steinrötter, Effects of Algorithmic Decision-Making and Interpretability on Human Behavior: Experiments using Crowdsourcing, HCOMP'18, Zürich, 2018.
- [BAU16] T. Bauernhansl, J. Krüger, G. Reinhart, G. Schuh, WGP-Standpunkt Industrie 4.0, <https://wgp.de/de/wgpstandpunktpapier-fuehrt-durchs-schluesselloch-zu-industrie-4-0>, 2016
- [BIT19] Bitkom, Fraunhofer IAO, Potenziale von Industrie 4.0 im deutschen Maschinen- und Anlagenbau gemessen an der Bruttowertschöpfung bis zum Jahr 2025 (in Milliarden Euro), 2019, <https://de.statista.com/statistik/daten/studie/297992>
- [BIT19] Bitkom, Welche Hemmnisse sehen Sie beim Einsatz von Industrie-4.0-Anwendungen in Ihrem Unternehmen?, 2019, <https://de.statista.com/statistik/daten/studie/830813/umfrage/hemmnisse-beimeinsatz-von-industrie-40-anwendungen-in-deutschland/>
- [BMW16] BMWi, Forschungsagenda Industrie 4.0, <https://www.plattform-i40.de/I40/Redaktion/DE/Downloads/Publikation/forschungsagenda-i40.pdf>
- [BMW19] BMWi, Wachstumspfade bei der Digitalisierung von Geschäftsmodellen in Industrieunternehmen, 2019, <https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/Wachstumspfade-Digitalisierung-Gesch%C3%A4ftsmodelle.pdf>
- [BRI09] Brinkhaus J-W (2009) Statistische Verfahren zur selbstlernenden Überwachung spanender Bearbeitungen in Werkzeugmaschinen, Dr.-Ing. Diss., Garbsen, Leibniz Universität Hannover.
- [DEW22] Dewes, A. Verfahren zur Anonymisierung und Pseudonymisierung von Daten. In: Rohde, M., Bürger, M., Peneva, K., Mock, J. (eds) Datenwirtschaft und Datentechnologie. Springer Vieweg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-65232-9_14
- [DIN18] DIN e.V., Deutsche Normungsroadmap Industrie 4.0, Version 3, 2018
- [DIT19] M.-A. Dittrich, F. Uhlich, B. Denkena, Self-optimizing tool path generation for 5-axis machining processes. CIRP Journal of Manufacturing Science and Technology 24, S. 49 - 54, 2019.
- [EVE02] W. Eversheim, Organisation in der Produktionstechnik 3: Arbeitsvorbereitung. 4. Aufl., Springer, 2002.
- [GAB18] Gabriel P., Potenziale der künstlichen Intelligenz im Produzierenden Gewerbe in Deutschland, Institut für Innovation und Technik, 2018, <https://www.bmwi.de/Redaktion/DE/Publikationen/Studien/potenzialekuenstlichen-intelligenz-im-produzierenden-gewerbe-in-deutschland.pdf>
- [GEI18] R. Geirhos, C. R. M. Temme, J. Rauber, H. Schütt, M. Bethge, F. A. Wichmann, Generalisation in humans and deep neural networks, Advances in Neural Information Processing Systems, 2018
- [GOD09] I. Goodfellow, Y. Bengio, Q. Courville, Deep Learning, Adaptive Computation and Machine Learning series, 2016
- [HAS11] H. M. Hashemian, State-of-the-Art Predictive Maintenance Techniques. IEEE Trans. Instrum. Meas. 60, 226–236, 2011.
- [HUA15] S. H. Huang, Y. C. Pan, Automated visual inspection in the semiconductor industry: A survey. Computers in industry, 66, 1-10, 2015.
- [JAL 19] Jalowski, M./Fritzsche, A./Möslein, K.M. (2019): Applications for Persuasive Technologies in Participatory Design Processes, in: Oinas-Kukkonen, H./Win, K.T./Kara-panos, E./Karppinen, P./Kyza, E. (Hrsg.): Persuasive 2019 – Lecture Notes in Computer Science, Cham, S. 74-86.

- [JUN17] D. Jung, Z. Zhang, M. Winslett, Vibration Analysis for IoT Enabled Predictive Maintenance, International Conference on Data Engineering, pp. 1271–1282, 2017
- [KAG13] H. Kagermann, W. Wahlster, J. Helbig, Securing the future of German manufacturing industry. Recommendations for implementing the strategic initiative INDUSTRIE 4.0, Plattform Industrie 4.0, Frankfurt, 2013.
- [LIC19] K. Lichtblau, Plattformen – Infrastruktur der Digitalisierung, Vereinigung der Bayerischen Wirtschaft e.V., 2019
- [MIT15] 60. V. Mittelstädt, P. Brauner, M. Blum, M. Ziefle, On the visual design of ERP systems – the role of information complexity, presentation and human factors. Procedia Manufact. pp. 270–277, 2015.
- [MOB02] R. K. Mobley, An Introduction to Predictive Maintenance, Elsevier, 2002.
- [MON16] L. Monostori, B. Kádár, T. Bauernhansl, S. Kondoh, S. Kumara, G. Reinhart, O. Sauer, G. Schuh, W. Sihn, K. Ueda, Cyber-physical systems in manufacturing, CIRP Annals 65(2), pp. 621 - 641, 2016
- [MÜL16] F. G. Müller, M. Bressner, D. Görzig, T. Röber, Industrie 4.0: Entwicklungsfelder für den Mittelstand: aktuelle Hemmnisse und konkrete Bedarfe, Fraunhofer IPA, 2016.
- [OKS19] Oks, S.J./Jalowski, M./Fritzsche, A./Möslein, K.M. (2019): Cyber-Physical Modeling and Simulation – A Reference Architecture for Designing Demonstrators for Indus-trial Cyber-Physical Systems, in: Procedia CIRP, Vol. 84, No. 6, S. 257-264.
- [SAH14] B. Saha, D. Srivastava, Data quality: The other face of Big Data, Proceedings - International Conference on Data Engineering, pp. 1294-1297, 2014
- [SEV16] K. Severson, P. Chaiwatanodom, R. D. Braatz, Perspectives on process monitoring of industrial systems“, Annual Reviews in Control, 42, pp. 190-200, 2016
- [SUS15] G. A. Susto, A. Schirru, S. Pampuri, S. McLoone, A. Beghi, Machine Learning for Predictive Maintenance: A Multiple Classifier Approach. IEEE Trans. Ind. Inform. 11, 812–820, 2015.
- [TAL20] Talmar, M., Walrave, B., Podoynitsyna, K. S., Holmström, J., and Romme, A. G. L. 2020. “Mapping, Analyzing and Designing Innovation Ecosystems: The Ecosystem Pie Model,” Long Range Planning (53:4), p. 101850. (<https://doi.org/10.1016/j.lrp.2018.09.002>).
- [TRA17] 84. A.J.C. Trappey, C.V. Trappey, U. H. Govindarajan, A. C. Chuang, J. J. Sun, A review of essential standards and patent landscapes for the Internet of Things: A key enabler for Industry 4.0, Journal of Advanced Engineering Informatics, 33, 2017, pp. 208-229
- [UHL17] E. Uhlmann, E. Hohwieler, C. Geistert, Intelligent production systems in the Era of Industrie 4.0 – changing mindsets and business models. J. Mach. Eng. 17(2), 5–24, 2017
- [WES06] E. Westkämpfer, Einführung in die Organisation der Produktion. Springer, 2006.
- [WIC17] F. A. Wichmann, D. H. J. Janssen, R. Geirhos, G. Aguilar, H. H. Schütt, M. Maertens, M. Bethge, Methods and measurements to compare men against machines, Human Vision and Electronic Imaging, 2017
- [WIE09] H.-P. Wiendahl, Betriebsorganisation für Ingenieure. 7. Aufl., Hanser, 2009.
- [WU18] C. Wu, M. J. F. Gales, A. Ragni, P. Karanasou, K. C. Sim, Improving Interpretability and Regularization in Deep Learning, IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2018
- [XU18] J. Xu, X. He, H. Li, Deep Learning for Matching in Search and Recommendation, Intl. Conference on Research & Development in Information Retrieval, 2018
- [YU11] D. Yu, L. Deng, Deep Learning and Its Applications to Signal and Information Processing, IEEE Signal Processing Magazine, 2011

Berichtsblatt

1. ISBN oder ISSN	2. Berichtsart (Schlussbericht oder Veröffentlichung) Schlussbericht
3. Titel IIP-Ecosphere – Next Level Ecosphere for Intelligent Industrial Production: Gemeinsamer Schlussbericht	
4. Autor(en) [Name(n), Vorname(n)] Niederée, Claudia; Eichelberger, Holger; Schreiber, Per; Ahmadian, Amir Shayan; Broos, Alexander; Dewes, Andreas; Ehl, Marco; Heinke, André; Hildebrandt, Petra; Jahnke, Nils; Kirschbaum, Julius; Klinkenberg, Ralf; Lange, Dirk; Litwinski, Kai; Palmer, Gregory; Quardt, Daniela; Reimer, Jeff; Schicktan, Michael; Schmees, Hans-Dieter; Sieckman, Geeske; Suckro, Andrea; Stichweh, Heiko; Urdich, Jürgen; Wilga, Matthäus; Wilkens, Rainer; Wißbrock, Peter;	5. Abschlussdatum des Vorhabens 30.9.2023
	6. Veröffentlichungsdatum
	7. Form der Publikation
8. Durchführende Institution(en) (Name, Adresse) Leibniz Universität Hannover, Am Welfengarten 1, 30167 Hannover (Verbundkoordinator) Universität Hildesheim, Universitätsplatz 1, 31141 Hildesheim MARPOSS Monitoring Solutions GmbH, Buchenring 40, 21272 Egestorf Lenze Gruppe, Hans-Lenze-Straße 1, 31855 Aerzen Universität Koblenz-Landau, Rhabanusstraße 3, 5518 Mainz Friedrich-Alexander-Universität Erlangen-Nürnberg, Schloßplatz 4, 91054 Erlangen Siemens AG, Werner-von-Siemens-Straße 1 80333 München Volkswagen Nutzfahrzeuge AG, Mecklenheidestraße 74, 30419 Hannover Deutsche Messe Technology Academy GmbH, Messegelände / Pavillon 36, 30521 Hannover VDW Verein Deutscher Werkzeugmaschinenfabriken, Lyoner Str. 18, 60528 Frankfurt a. Main Bitmotec GmbH, Hollerithallee 17, 30419 Hannover Rapidminer GmbH, Westfalendamm 87, 44141 Dortmund KIProtect GmbH, Bismarckstraße 10-12, 10625 Berlin Sennheiser electronic GmbH & Co. KG, Am Labor 1, 30900 Wedemark Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., Hansastraße 27c, 80686 München DMG/Gildemeister Drehmaschinen GmbH, Gildemeisterstraße 60, D-33689 Bielefeld slashwhy GmbH & Co. KG, Netter Platz 2, 49090 Osnabrück	9. Ber. Nr. Durchführende Institution --
	10. Förderkennzeichen 01MK20006 A-Q
	11. Seitenzahl 177
12. Fördernde Institution (Name, Adresse) Bundesministerium für Wirtschaft und Klimaschutz (BMWK) 10115 Berlin	13. Literaturangaben 133
	14. Tabellen 17
	15. Abbildungen 98
16. Zusätzliche Angaben	
17. Vorgelegt bei (Titel, Ort, Datum)	
18. Kurzfassung Künstliche Intelligenz (KI) wird in der Produktion als wichtige Zukunftstechnologie gesehen. Ihre Anwendung im konkreten Produktionskontext stellt jedoch viele Unternehmen vor erhebliche Herausforderungen. Ziel des IIP-Ecosphere Projekts war es daher ein neuartiges Ökosystem – die Next Level Ecosphere for Intelligent Industrial Production (IIP-Ecosphere) – aufzubauen, welches die Nutzung von KI erleichtert und vorantreibt und damit zum Erreichen einer „nächste Ebene“ in der intelligenten Produktion beiträgt. Im Projekt wurden hierfür vielfältige Aktivitäten durchgeführt: In vier Think Tanks wurden innovative Methoden in den Bereichen KI-Methoden, Plattformen, Data Sharing und KI-Geschäftsmodelle erforscht und entwickelt. Beiträge des KI-Accelerators wie der KI-Lösungskatalog und KI-Services zielen auf die Beschleunigung der Umsetzung von KI-Lösungen ab. Zudem wurden vier größere Demonstratoren für die Anwendung innovativer KI-Methoden in realen Produktionskontexten mit namhaften Unternehmen als Blueprints erfolgreich umgesetzt ebenso wie vier kleinere, dynamisch im Projekt definierte Demonstratoren mit externen Partnern. Die Sichtbarkeit und Vernetzung wurden durch zwei Regional Innovation Hubs vorangetrieben, die unter anderem Arbeitskreise und vielfältige Events organisiert haben. Zudem wurden die Ergebnisse des Projekts mit größeren Demonstratoren auf drei wichtigen Messen (Hannover Messe 2022 und 2023, EMO 2023) eine Vielzahl von Unternehmen vorgeführt. Im Ergebnis ist ein Ökosystem – die IIP-Ecosphere - entstanden mit starker Vernetzung der Partner, reichem Erfahrungsschatz und wichtigen Diensten, Lösungen und Assets. Zudem resultieren aus den Aktivitäten im Projekt vielfältige weitere Ergebnisse, welche in Unternehmen genutzt, wissenschaftlich publiziert und in neuen Projekten weiterentwickelt werden. Besondere Erwähnung als Ergebnis des Projekts verdient die flexible und anpassbare open source IIoT-Plattform (Industrie 4.0 Plattform), die im Projekt auf der Grundlage der Anforderungen von Unternehmen konzipiert, entwickelt sowie erfolgreich erprobt wurde und unter dem Branding oktoflow-Plattform nach Projektende weiterentwickelt und verwertet wird.	
19. Schlagwörter Intelligente Produktion, KI, IIoT Plattform, Ökosystem, Industrie 4.0, Data Sharing	
20. Verlag	21. Preis

Document Control Sheet

1. ISBN or ISSN	2. type of document (e.g. report, publication) Final Report		
3. title IIP-Ecosphere – Next Level Ecosphere for Intelligent Industrial Production: Joint Final Report			
4. author(s) (family name, first name(s)) Niederée, Claudia; Eichelberger, Holger; Schreiber, Per; Ahmadian, Amir Shayan; Broos, Alexander; Dewes, Andreas; Ehl, Marco; Heinke, André; Hildebrandt, Petra; Jahnke, Nils; Kirschbaum, Julius; Klinkenberg, Ralf; Lange, Dirk; Litwinski, Kai; Palmer, Gregory; Quardt, Daniela; Reimer, Jeff; Schicktan, Michael; Schmees, Hans-Dieter; Sieckman, Geeske; Suckro, Andrea; Stichweh, Heiko; Urdich, Jürgen; Wilga, Matthäus; Wilkens, Rainer; Wißbrock, Peter;		5. end of project September 30, 2023	6. publication date --
8. performing organization(s) (name, address) Leibniz Universität Hannover, Am Welfengarten 1, 30167 Hannover (Verbundkoordinator) Universität Hildesheim, Universitätsplatz 1, 31141 Hildesheim MARPOSS Monitoring Solutions GmbH, Buchenring 40, 21272 Egestorf Lenze Gruppe, Hans-Lenze-Straße 1, 31855 Aerzen Universität Koblenz-Landau, Rhabanusstraße 3, 5518 Mainz Friedrich-Alexander-Universität Erlangen-Nürnberg, Schloßplatz 4, 91054 Erlangen Siemens AG, Werner-von-Siemens-Straße 1 80333 München Volkswagen Nutzfahrzeuge AG, Mecklenheidestraße 74, 30419 Hannover Deutsche Messe Technology Academy GmbH, Messegelände / Pavillon 36, 30521 Hannover VDW Verein Deutscher Werkzeugmaschinenfabriken, Lyoner Str. 18, 60528 Frankfurt a. Main Bitmotec GmbH, Hollerithallee 17, 30419 Hannover Rapidminer GmbH, Westfalendamm 87, 44141 Dortmund KIProtect GmbH, Bismarckstraße 10-12, 10625 Berlin Sennheiser electronic GmbH & Co. KG, Am Labor 1, 30900 Wedemark Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., Hansastraße 27c, 80686 München DMG/Gildemeister Drehmaschinen GmbH, Gildemeisterstraße 60, D-33689 Bielefeld slashwhy GmbH & Co. KG, Netter Platz 2, 49090 Osnabrück		7. form of publication --	9. originator's report no. --
		10. reference no. 01MK20006 A-Q	
		11. no. of pages 177	
12. sponsoring agency (name, address) Bundesministerium für Wirtschaft und Klimaschutz (BMWK) 10115 Berlin		13. no. of references 133	14. no. of tables 17
		15. no. of figures 98	
		16. supplementary notes --	
17. presented at (title, place, date) --			
18. abstract Artificial intelligence (AI) is considered an important future technology in production. However, its application in a specific production context is a challenging task for many companies. The aim of the IIP-Ecosphere project was, therefore, to build an innovative ecosystem - the Next Level Ecosphere for Intelligent Industrial Production (IIP-Ecosphere) - which facilitates and advances the use of AI and thus contributes to reaching a "next level" in intelligent production. A variety of activities were carried out in the project: innovative methods in the areas of AI methods, platforms, data sharing and AI business models were researched and developed in four think tanks. Contributions from the AI Accelerator aim to accelerate the implementation of AI solutions. In addition, four larger demonstrators for the application of innovative AI methods in real production contexts were successfully developed as blueprints with well-known companies, and four smaller, dynamically defined demonstrators were implemented with external partners. Visibility and networking were promoted by two Regional Innovation Hubs, which, among other things, organized working groups and a variety of events. In addition, the results of the project were demonstrated to various companies with larger demonstrators at three important trade fairs (Hannover Messe 2022 and 2023, EMO 2023). The project result is an ecosystem - the IIP-Ecosphere - with a strong network of partners, a wealth of experience and important services, solutions, and assets. In addition, the activities in the project produced in a variety of other results, which are used in companies, published scientifically and further developed in new projects. One of these results is worth a special mention: the flexible and customizable open source IIoT platform (Industry 4.0 platform), which was designed, developed, and successfully tested in the project based on the requirements of companies. It will be further developed and exploited under the branding "oktoflow platform" after the end of the project.			
19. keywords intelligent industrial production, AI, IIoT Plattform, ecosystem, Industry 4.0, Data Sharing			
20. publisher --		21. price --	