



18. FACHTAGUNG • EKA 2024

ENTWURF KOMPLEXER AUTOMATISIERUNGSSYSTEME

BESCHREIBUNGSMITTEL, METHODEN, WERKZEUGE UND ANWENDUNGEN

ULRICH JUMAR, CHRISTIAN DIEDRICH (HRSG.)

■ EINE KOOPERATION VON:

ifak



Impressum

18. Fachtagung

«ENTWURF KOMPLEXER AUTOMATISIERUNGSSYSTEME»
(EKA 2024)

14.-15. Mai 2024 • Magdeburg

DOI: <http://dx.doi.org/10.25673/116027>



HERAUSGEBER

Ulrich Jumar, Magdeburg
Christian Diedrich, Magdeburg

Institut für Automation und
Kommunikation e.V. Magdeburg
An-Institut der Otto-von-Guericke-Universität
Magdeburg
Werner-Heisenberg-Straße 1
39106 Magdeburg

Telefon: +49 (0)391 990140

Fax: +49 (0)391 9901590

Internet: www.ifak.eu

Vorwort

Tagungsband der 18. Fachtagung

Entwurf komplexer Automatisierungssysteme – EKA 2024

Die Tagung EKA – Entwurf komplexer Automatisierungssysteme – wurde 1992, also vor nunmehr bereits 32 Jahren, von Prof. Dr.-Ing. Dr. h.c. Eckehard Schnieder, dem damaligen Leiter des Instituts für Regelungs- und Automatisierungstechnik an der TU Braunschweig gegründet und dort in regelmäßiger Folge bis zur 9. EKA 2006 veranstaltet. Seit der Neuausrichtung des Instituts in Braunschweig wird die Tagungsreihe gemeinsam vom Institut ifak e.V. und der Otto-von-Guericke-Universität in Magdeburg fortgesetzt. So fand die 10. EKA im Jahr 2008 erstmals in der Wirkungsstätte des ifak, in der Denkfabrik im Wissenschaftshafen Magdeburg, statt.

Neben thematisch breit angelegten Kongressen, die naturgemäß einen möglichst großen Kreis von Teilnehmenden ansprechen, haben kleine und spezielle Fachtagungen wie die EKA einen festen Platz in der Tagungslandschaft. Die Tradition der EKA ist ein Indiz dafür, dass der Entwurf komplexer Automatisierungssysteme nicht erst seit der umfassenden Digitalisierung wissenschaftliche und praktische Herausforderungen bereithält. Wie der Untertitel verdeutlicht, hat die Fachtagung EKA das Ziel, Beschreibungsmittel, Methoden und Werkzeuge für den Entwurf komplexer Automatisierungssysteme im Dialog von theoretischen Erkenntnissen und praktischer Nutzung zu thematisieren. Die Schwerpunktwahl unterstreicht, dass ein von der spezifischen Anwendung getriebenes Denken durch ein fundiertes methodisches Vorgehen und einen substantiellen theoretischen Hintergrund ergänzt werden muss. Folglich ist ein gelingender Brückenschlag zwischen theoretischen Erkenntnissen und deren Nutzung in der Praxis ein wichtiges Anliegen der Tagung.

Im Ergebnis der Begutachtung der erweiterten Beitragskurzfassungen hat das Programmkomitee das EKA-Tagungsprogramm zusammengestellt. Integriert in die Vortragsitzungen der 18. EKA ist eine Vorlesung von Prof. Christian Diedrich, mit der er sich an der Otto-von-Guericke-Universität in den wohlverdienten Ruhestand verabschiedet. Die Anzahl der im elektronischen Tagungsband zusammengestellten Beiträge ist deshalb etwas geringer als bei den vorausgegangenen Tagungen. Beibehalten wird die bewährte Praxis, dass ausgewählte EKA-Beiträge in einer überarbeiteten und erweiterten Form später in einer „Special Issue EKA 2024“ der wissenschaftlichen Zeitschrift *at-Automatisierungstechnik* im Verlag De Gruyter Oldenbourg erscheinen werden.

Die Tagungsleiter Prof. Diedrich und Prof. Jumar wünschen Ihnen eine aufschlussreiche Lektüre des elektronischen Tagungsbandes der 18. EKA. Die Magdeburger Fachkollegen freuen sich, Sie im Jahr 2026 zur nächsten EKA erneut begrüßen zu dürfen.

Prof. Dr. Ulrich Jumar

*ifak - Institut für Automation und
Kommunikation e.V.*

Prof. Dr. Christian Diedrich

Otto-von-Guericke-Universität Magdeburg

EKA 2024 – Entwurf komplexer Automatisierungssysteme

1 Methods to Describe, Assign and Derive Capabilities from the Capability, Skill and Service (CSS) Model

Michael Winter¹, Jürgen Bock², Johannes Hoos³, Tobias Klausmann⁴, Boris Schnebel⁵, Siwara Schmitt⁶, Marco Simon⁷, Melanie Stolze⁸, Luis Miguel Vieira da Silva⁹, Tobias Kleinert¹ (¹RWTH Aachen; ²TH Ingolstadt; ³Festo SE & Co. KG; ⁴Lenze SE; ⁵Fraunhofer IOSB; ⁶Fraunhofer IESE; ⁷Technologie-Initiative SmartFactory KL e.V.; ⁸ifak e.V.; ⁹Helmut-Schmidt-Universität/ Universität der Bundeswehr Hamburg)

2 Analyse und Konzeption von Informationsmodellen für Soft Robotics

Linus Witucki, Mike Barth (Karlsruher Institut für Technologie)

3 SysML-Profil zur automatisierten Generierung von PDDL-Beschreibungen aus Systemmodellen

Hamied Nabizada¹, Lasse Beers¹, Maximilian Weigand¹, Felix Gehlhoff¹, Alexander Fay² (¹Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg; ²Ruhr-Universität Bochum)

4 SIM-CIP: Konzept eines räumlichen Informationsmodells für komplexe industrielle Großanlagen

Alexander Auer^{1,2}, Birgit Vogel-Heuser¹, Tobias Knödler², Dorothea Pantförder¹ (¹TU München; ²MAN Energy Solutions SE)

5 Integration verschiedener Informationstechnologien in der Verwaltungsschale am Beispiel von AAS und AutomationML

Arndt Lüder¹, Jan Blume², Miriam Schleipen³, Matthias Freund⁴, David Hoffmann¹, Pooja Gupta¹ (¹Otto von Guericke Universität; ²Fraunhofer IOSB; ³EKS Intec; ⁴Festo SE & Co. KG)

6 Ansätze für die Modellierung von Streckeninformationen zur Verifikation im Steuerungsentwurf

Aron Schnakenbeck¹, Robin Mroß², Marcus Völker², Stefan Kowalewski², Alexander Fay^{1,3} (¹Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg; ²RWTH Aachen University; ³Ruhr-Universität Bochum)

7 Konzepte zur Realisierung proaktiver Verwaltungsschalen und deren Kommunikation

Melanie Stolze¹, Gustavo Cainelli¹, Christian Kosel², Alexander Belyaev³, Christian Diedrich³ (¹ifak e.V.; ²Arena 2036 e.V.; ³Otto-von-Guericke-Universität Magdeburg)

8 Analytics and Visualization of Industrial Asset Condition using Asset Administration Shell submodels.

Sandeep Rudra¹, Kazeem Oladipupo¹, Christian Diedrich¹, Sebastian Kaebisch² (¹Otto-von-Guericke-Universität Magdeburg; ²Siemens AG)

9 Praktikable Ermittlung des produktbezogenen CO₂-Fußabdrucks mithilfe grüner Digitaler Energiezwillinge

Florian Balduf, Zai Müller-Zhang, Thomas Kuhn (Fraunhofer-Institut IESE)

10 Modelling of NOA 178 Verification of Request concept using Pro-Active Asset Administration Shells

Harish Kumar Pakala, Muhammad Rizwan Kamleen, Christian Diedrich (Otto-von-Guericke-Universität Magdeburg)

11 Integration von Zeitreihendaten in die Verwaltungsschale - Pragmatische Umsetzung des Standards für die Anwendung in der Automatisierungsindustrie

Aaron Zielstorff, Dirk Schöttke, Fiona Helena Büttner, Thomas Kämpfe, Stephan Schäfer (HTW Berlin)

-
- 12 Webbasiertes Werkzeug für das modellbasierte Engineering von KI-Anwendungen für Automatisierungssysteme**
Marvin Schieseck¹, Philip Topalis¹, Alexander Fay² (¹Helmut-Schmidt-Universität/Universität der Bundeswehr in Hamburg; ²Ruhr-Universität Bochum)
- 13 Industrielle Sicherheit durch Zertifikatsmanagement-Konzepte im Lebenszyklus einer Produktions-Gray-Box**
Marwin Madsen, Mike Barth (Karlsruher Institut für Technologie)
- 14 Analyse dezentraler Fahrzeugsteuerungsarchitekturen mittels variierender Autonomiegrade an kritischen Punkten von Intralogistiksystemen**
Felix Gehlhoff¹, Bjarne Huth¹, Alexander Wenzel² (¹Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg; ²Leibniz Universität Hannover)
- 15 Simplified Model for Data Collection of Liquid Products in Process Manufacturing**
Wei Guo¹, Wan Li¹, Swantje Thiele², Dominik Orzol², Thomas Bewer³, Tobias Kleinert¹ (¹RWTH Aachen; ²IPGR e.V.; ³Bucher Emhart Glass, Switzerland)
- 16 Effiziente Integration von Elektrolyseuren in modularen Elektrolyseanlagen**
Julius Lorenz¹, Tobias Kock¹, Lukas Grahl², Vincent Henkel³, Lena Scholz³, Katharina Stark⁴, Rene Lorenz⁵, Anselm Klose⁶, Leon Urbas¹ (¹TU Dresden; ²Semodia GmbH; ³Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg; ⁴ABB AG Forschungszentrum; ⁵DLR e.V., Institut für Technische Thermodynamik; ⁶Process-to-Order Lab, TU Dresden)
- 17 Koordinierter Betrieb eines Pumpspeicherkraftwerkes mit reversibler Pumpturbine und einem Batteriespeicher an einem geteilten Standort**
Leon Tadayon^{1,2}, Josef Meiers¹, Lukas Ibing², Kevin Erdelkamp², Georg Frey¹ (¹Universität des Saarlandes; ²RWE Generation SE)
- 18 Data-Driven Soft Sensors for Electrical Machines**
Nada Sahlab¹, Arzam Kotriwala¹, Andrew Habib¹, Victor Mukherjee² (¹ABB Corporate Research Germany; ²ABB Motors and Generators Finland)
- 19 Modellbasierter Entwurf und Anwendung einer Eigenschaftsregelung zur Automatisierung eines innovativen Drückwalzprozesses**
Lukas Kersting¹, Bahman Arian², Julian Rozo Vasquez³, Ansgar Trächtler^{1,4}, Werner Homberg², Frank Walther³ (¹Fraunhofer IEM; ²Universität Paderborn; ³TU Dortmund; ⁴Heinz Nixdorf Institut, Universität Paderborn)

„Ulrich Jumar, Christian Diedrich (Hrsg.):
EKA 2024 - Entwurf komplexer Automatisierungssysteme, 18. Fachtagung“

Methods to Describe, Assign and Derive Capabilities from the Capability, Skill and Service (CSS) Model

Themengebiet: Modellierung mit formalen Beschreibungen

Michael Winter¹, Jürgen Bock², Johannes Hoos³, Tobias Klausmann⁴, Boris Schnebel⁵, Siwara Schmitt⁶, Marco Simon⁷, Melanie Stolze⁸, Luis Miguel Vieira da Silva⁹ und Tobias Kleinert¹⁰

Abstract: For the use of the CSS model, workflows are needed to describe, assign and derive capabilities. This contribution provides a Capability Determining method that encompasses the three underlying technology-independent methods to guide stakeholders in describing, assigning and deriving capabilities for manufacturing. The first underlying method, Describing Capabilities, describes a workflow for the formal semantic description of capabilities. The workflow for assigning capabilities to a production resource is explained using the second method Assigning Capability. In the last workflow, capabilities are derived from a production specification or a process description using the method Deriving Capability.

Keywords: Industry 4.0, Capability description, Resource function, Product description, Automation engineering

¹ RWTH Aachen University, Chair of Information and Automation Systems for Process and Material Technology (IAT), Turmstraße 46, 52064 Aachen, m.winter@iat.rwth-aachen.de

² Technische Hochschule Ingolstadt, AI-Applications in innovative Production and Logistics, Esplanade 10, 85049 Ingolstadt, juergen.bock@thi.de

³ Festo SE & Co. KG, Advanced Develop. Digital Engineering, Rüter Straße 82, 73734 Esslingen, johannes.hoos@fest.com

⁴ Lenze SE, R&D Systems Engineering, Hans-Lenze-Straße 1, 31855 Aerzen, tobias.klausmann@lenze.com

⁵ Fraunhofer IOSB, Optronik, Systemtechnik und Bildauswertung, Fraunhoferstr. 1, 76131 Karlsruhe, boris.schnebel@iosb-extern.fraunhofer.de

⁶ Fraunhofer IESE, Embedded Systems, Fraunhofer-Platz 1, 67663 Kaiserslautern, siwara.schmitt@iese.fraunhofer.de

⁷ Technologie-Initiative SmartFactory KL e.V., Trippstadter Straße 122, 67663 Kaiserslautern, marco.simon@smartfactory.de

⁸ ifak - Institut für Automation und Kommunikation e.V., Werner-Heisenberg-Str. 1, 39106 Magdeburg, melanie.stolze@ifak.eu

⁹ Helmut Schmidt University, Institute of Automation Technology, Holstenhofweg 85, 22043 Hamburg, miguel.vieira@hsu-hh.de

¹⁰ RWTH Aachen University, Chair of Information and Automation Systems for Process and Material Technology (IAT), Turmstraße 46, 52064 Aachen, kleinert@iat.rwth-aachen.de

1 Introduction

In the rapidly evolving and volatile environment of manufacturing, flexibility plays a central role to be successful. The term flexibility encompasses the incorporation of more product variants through one production facility, the ability to accommodate smaller batch sizes of different product variants and the reorganization of production orders with a reduction in associated efforts. To fulfill these requirements for flexibility, the authors are applying the Capability, Skill, and Service (CSS) model for the implementation of capability-based engineering. Extended from the product, process and resource (PPR) paradigm, the CSS model decouples product design from production engineering [Kö23]. A capability is considered the implementation-independent specification of a function, while a skill is defined as the implemented function specified by a capability, cf. [Di22]. In order to produce products by applying a respective production procedure, capabilities are required to realize the individual steps of the procedure. These required capabilities, in turn, have to be matched to the capabilities offered by available production resources [Di22]. To determine capabilities for product fabrication and ultimately to find a suitable sequence of feasible production steps, a matching between required capabilities and those offered by the resources is performed. In the current state of research, there are no technology-agnostic and multi-purpose methodologies for the formal description of capabilities to achieve this matching. Therefore, this contribution by the Plattform Industrie 4.0 working group "Semantic and Interaction of I4.0 Components" addresses the following research questions:

- How to methodically approach the formal description of capabilities?
- How to methodically approach the assignment of capabilities to resources?
- How to methodically derive capabilities from product or process descriptions?

The objective is to provide three technology-independent methods to guide stakeholders through describing, assigning and deriving capabilities for manufacturing.

The remainder of this paper begins with an overview of related publications in Section 2. In Section 3 the methods for the formal description of capabilities are presented. Finally, Section 4 concludes the presented methods and outlines future research.

2 Related Work

In [Fr22], an overview of capability and skill-based manufacturing systems is given by conducting a literature review. Research in the area of capabilities and skills is justified by the need for adaptability and flexibility in manufacturing and the associated process planning. In addition to a formal description of capabilities and skills, the property of being suitably designed for matching of requirements and functions of production resources – ultimately for process planning – was identified as one of the most common requirements for capabilities and skills. It is shown that different modeling methods, e.g., ontologies or

UML, and different technologies, e.g., AutomationML or OWL, are used to model capabilities. However, modeling issues such as inconsistent modeled capabilities or different levels of abstraction are named as the most common challenges for the use of capabilities and skills in manufacturing. Therefore, there is a need for a method to formally describe capabilities that is ideally technology-independent and thus can be used for different capability-based approaches. [Fr22]

A number of contributions to methodical modeling of capabilities use ontologies: The authors of [Jäl19] present a systematic development process of an ontology to describe capabilities of manufacturing resources. The presented method is based on the ontology engineering methodology of [Su09] and consists of five phases: feasibility study, kickoff, refinement, evaluation, and application and evolution. The main focus of this method is to model general concepts of a capability model to represent the domain under consideration, and it only considers modeling with ontologies. Actual capability instances are then modeled in the evaluation phase, but no methodological approach is presented. Furthermore, only the view of a resource and its provided capabilities is described, so that e.g. the derivation of capabilities from product specifications or process descriptions is not considered. In [Kö20] a method is presented that automatically generates the sub-aspects machine structure, capabilities and skills of a capability model from different artifacts in order to generate the entire capability model. In particular, the machine structure and skills aspects can be created automatically. The capability aspect is created semi-automatically by using a modelling tool for capabilities based on the VDI 3682 standard for formalized process description and then automatically transferring the created JSON file to the ontology. Accordingly, this approach allows the creation of a specific capability model in an ontology with a specific tool, without considering other standards for processes or other approaches to capability modeling, such as resource or product-based ones.

Other approaches use the Asset Administration Shell (AAS) to describe capabilities of assets: In [HDM21], a graphical modeling framework is presented for modeling systems according to the AAS standard. The framework enables the modeling of AAS without requiring an in-depth understanding of AAS. It also allows the modeling of capabilities or processes using BPMN. However, no method is presented on how to specifically model capabilities, and no standards are used to describe or derive capabilities. The authors of [Gr23] emphasize the importance of engineering processes and the embedding of engineering information from numerous standards in such capability models or specifically in digital twins (like the AAS) for continuous engineering. Important general suggestions for information modeling are given, such as the integration of standards into the AAS, reusable and extensible information modeling, but also the updating of information across all phases of engineering. However, there is no specific methodological proposal regarding capabilities and which standards should be used from PPR perspective to model them.

3 Methods for Determining Capabilities

The Capability Determining method proposed in this contribution entails the multiple use of underlying methods for describing, assigning and deriving capabilities. The user can decide whether a capability shall be described formally, assigned to a production resource, derived from a product specification, or be derived from a process description, see Fig. 1.

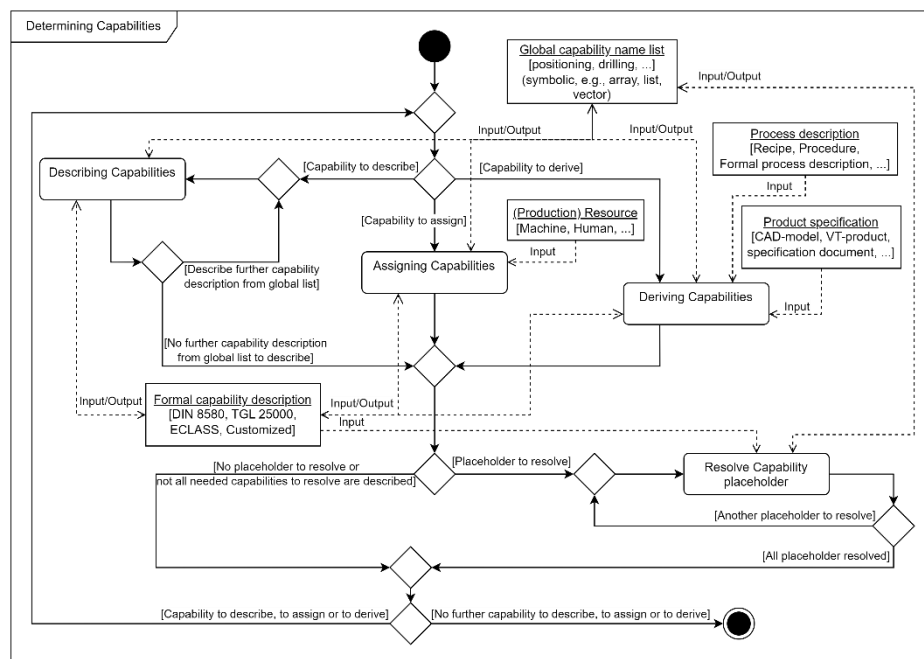


Fig. 1: Workflow for the method of Determining Capabilities

In the Capability Determining method various artifacts are used as inputs or outputs for the underlying methods. The main artifact that is used as an input and output for the underlying three methods Describing, Assigning and Deriving Capabilities is the formal capability description. This artifact can be, e.g., a capability submodel of the Asset Administration Shell or an ontology, which can consider several standards like DIN 8580, TGL 25000, ECLASS or customized function descriptions. Another artifact that serves as input and output for the underlying methods is the global capability name list. In this name list all formally described capabilities are available as key value pair, where the key is the capabilities unique id and the value is the capabilities informal name, e.g., “Drilling”. Depending on the underlying methods the name list might contain capability placeholders, where the placeholders are a temporary addendum. The capabilities for these placeholders

must be described in the method Describing Capabilities and the placeholders must be resolved.

For a comprehensive understanding of the workflow regarding the description of a capability, the underlying method Describing Capabilities is used to formally describe capabilities, see Section 3.1.

The method Assigning Capabilities is used for assigning formally described capabilities to a resource to be used in capability-based production approaches, see Section 3.2. For this method a production resource, e.g., a machine or human, is used as input.

The method Deriving Capabilities allows the derivation of capabilities from product specifications or process descriptions for a possible automated match with assigned capabilities of resources, see Section 3.3.

In the case that no suitable capability description was found in the methods Assigning Capabilities or Deriving Capabilities, a capability placeholder (with a unique id) will be created. This placeholder is used as an entry in the global capability name list. The placeholder, which represents the capability to be described, is then formalized in the method Describing Capabilities. After the capability description is formalized with another unique id, the activity Resolve Capability placeholder is to apply. As the formal capability description might differ from the preliminary capability placeholder, the placeholder must be exchanged with the final unique id of the formal capability previously described. This activity can be repeated until all capability placeholders are resolved.

3.1 Describing Capabilities

The method Describing Capabilities is one of the underlying methods of the Capability Determining method. As shown in Fig. 2, the workflow of the method Describing Capabilities includes the general activities for a formal and semantic description of a capability.

The description of a capability specifies the context in which the capability will be applied. Depending on the context, capabilities can differ in their descriptions and properties, e.g., “grasping” a tomato (food industry) or “grasping” a component (manufacturing industry). The context can be refined by analyzing a production process or by comparing similar production processes, as well as by interviewing stakeholders to identify experiences and challenges.

Existing capability descriptions should be considered to avoid the creation of duplicates. Therefore, the introduced global capability name is to check whether the desired capability name already exists. Subsequently, if the capability name is identified within the list, the referenced formal capability description should be checked to see whether it matches the capability to be created. If the capability name is not found in the global capability name list, a thorough examination of other capability descriptions within the list is warranted to ascertain whether they encapsulate the desired capability but are listed under a different (synonymous) name. In this case, the capability name used or the capability name in the global list should be reconsidered. If no correspondence is discerned within the global

capability name list for the queried capability, alternative databases such as triple stores, Asset Administration Shell repositories, or publicly available sources including ontologies, ECLASS, or the IEC Common Data Dictionary (CDD) should be explored for further investigation.

If one of the identified existing capability descriptions suits well to the capability needed to be described, the workflow of describing capabilities is finished. For this, an existing capability description is evaluated to be sufficient by checking corresponding activities of the method Describing Capabilities and with a successful validation, the capability is resolved in the superior method Determining Capabilities to the global capability name list.

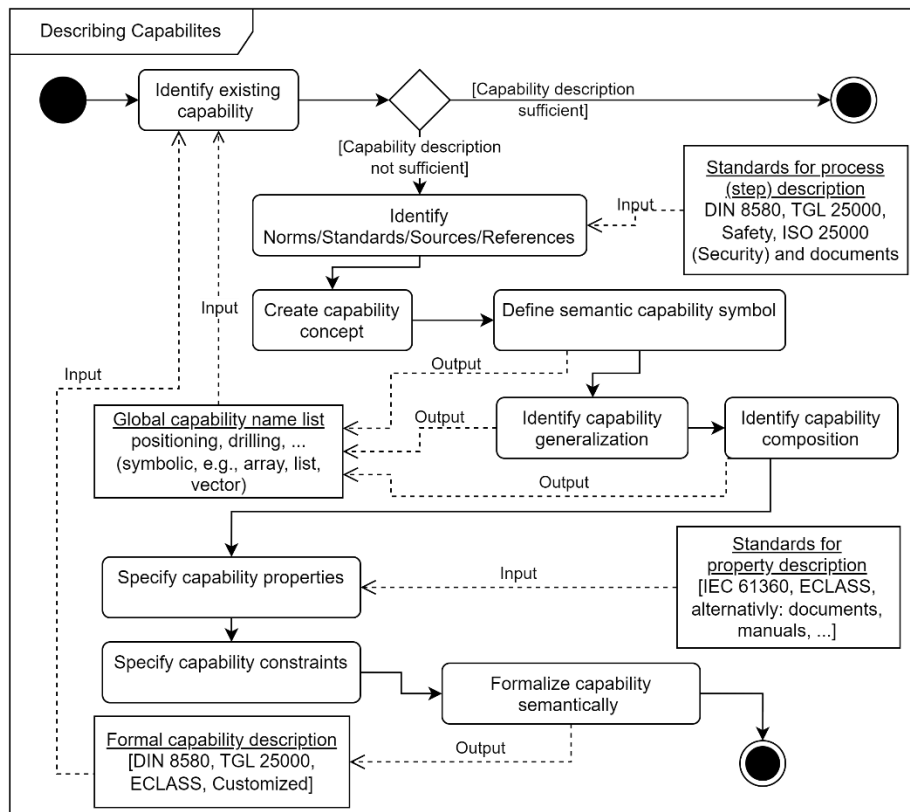


Fig. 2: Workflow for the method of Describing Capabilities

If no suitable existing capability is found, the next activity is to identify norms, standards, sources, or references that may contain information for the description of the capability. In the field of engineering and manufacturing, compliance with established norms,

standards and references is crucial to ensure consistent quality and reliability of capabilities and processes. One such recognized standard is the DIN 8580, which defines the terminology for manufacturing processes. In process industry, the TGL 25000 is a standard that provides description of unit operations by two distinct classification schemas. Furthermore, safety and security information can be essential. This can include ISO 25000 for defining system and software quality requirements and their evaluation, or other internal company security documents that define specific policies and standards.

For generating the capability concept, the model of the semiotic triangle will be used. The semiotic triangle is a model in semiotics and linguistics that distinguishes between the sign (or symbol), the thought (or reference), and the referent (or object) while also defining the relationships between them [Ja13]. The creation of the capability concept is an activity where the thought is generated with their attributes to be filled out in further steps of the method, i.e., the thought embodies the specification or description of the symbol "drilling" with their attributes that's needs to be filled out, e.g., name, properties, constraints, parent relationship and composition.

After the capability concept is created and the capability symbol is defined as well as added to the global capability name list, a capability description at different hierarchical levels can be required. One reason for this may be that required capabilities are described in a more abstract form, while offered capabilities are described more specifically. Consequently, different levels of modeling depths are necessary for matching within capability-based engineering. In order to describe the capabilities at different hierarchical levels, it is necessary to identify the capability generalization. Therefore, the upper class (superclass) must be linked to the capability to be described by using the aforementioned activity of identifying existing capabilities. In this activity of identifying generalized capabilities, it is possible that multiple generalizations may be identified either symbolically, which need to be described formally, or else identified with a complete capability description. With a successful identification of generalized capabilities, they have to be referenced. In the case that neither a capability is identified symbolically nor a complete capability description, the aforementioned activity to identify norms, standards, sources, or references is used to introduce a new entry into the global capability name list and to connect an upper-class capability from these other sources to the capability to be described.

A capability can also be composed of other capabilities. The composition of capabilities is a concept that involves combining two or more capabilities to create a new capability. Analog to the activity of identifying the capability generalization, the aforementioned activity of identifying existing capabilities and the activity to identify norms, standards, sources, or references can be used to identify capability compositions.

Properties are used to describe a capability in more detail. These Properties cover a broad spectrum from physical dimensions to sustainability data. To derive a suitable set of properties for a given capability, documents and manuals of production resources can be referred to as well as norms and standards such as IEC 61360 or ECLASS.

Capabilities can be constrained by either property constraints or transition constraints. Both constraint types can be specified by a condition. These conditions are either pre-, post- or invariant conditions. The property constraint is defined by a property value statement, which specifies the context in which a capability may be applied. For example, a capability “Brick burning” might only be applicable in a temperature range of 950°C – 1000°C. Additionally, the oven temperature is to reach before and held while the bricks are burned. Therefore, the property constraint applies pre (before) and invariant (while) conditions on the property temperature. The transition constraint allows to describe a dependency of a capability on another capability. For example, to cut a thread, it is first necessary to drill a hole. Therefore, when utilizing the capability “Threading”, a precondition for the capability “Drilling” might be apply.

Finally, semantic formalization which enables matching of required capabilities (derived from, e.g., product specification or process descriptions) and offered capabilities (assigned to production resources) has to be provided. Therefore, the formalized capability description is used for the following methods to either assign a capability to a production resource or else to derive a capability form a product specification or a process description.

3.2 Assigning Capabilities

Assigning capabilities to operational resources is a systematic approach for manufacturers or operators of a machine, to derive resource capability model instances, in order to allow for advanced automated production planning. In Fig. 3, the workflow of the method Assigning Capabilities with their activities and artifacts is shown.

The first activity involves identifying production resources. In the scope of this activity, the description of the production resource and their environment (if necessary) is used for selection. For acquiring the description, machine-readable descriptions, e.g., formal information models such as an AAS, or resource non-formal documentation, e.g., operation manuals, are used. The information can be gathered by using internal and external sources, e.g., manufacturers product catalogue.

Typically, resources already possess skills, which allows to operate them for the processing of products. Hence, the identification of resource skills becomes paramount, as skills serve as the executable implementations of encapsulated functions specified by capabilities [Di22]. As a prerequisite to the following activities, resource skill descriptions are made available by using the skill descriptions, e.g., by reading information models, such as Module Type Package, PackML or Control Components. The information gathered from the skill description is used to extract further information to determine the effect on the product by processing it in the production resource.

It is essential to ascertain the effect of each resource so that the alternative solutions of the production machines can be abstracted to a capability via a description (of the effect). Therefore, one goal of the production resource is to impose an effect onto the product

through skill execution. The effect is described as a change between two states¹¹, e.g., the change of a product geometry or a temperature change in a process.

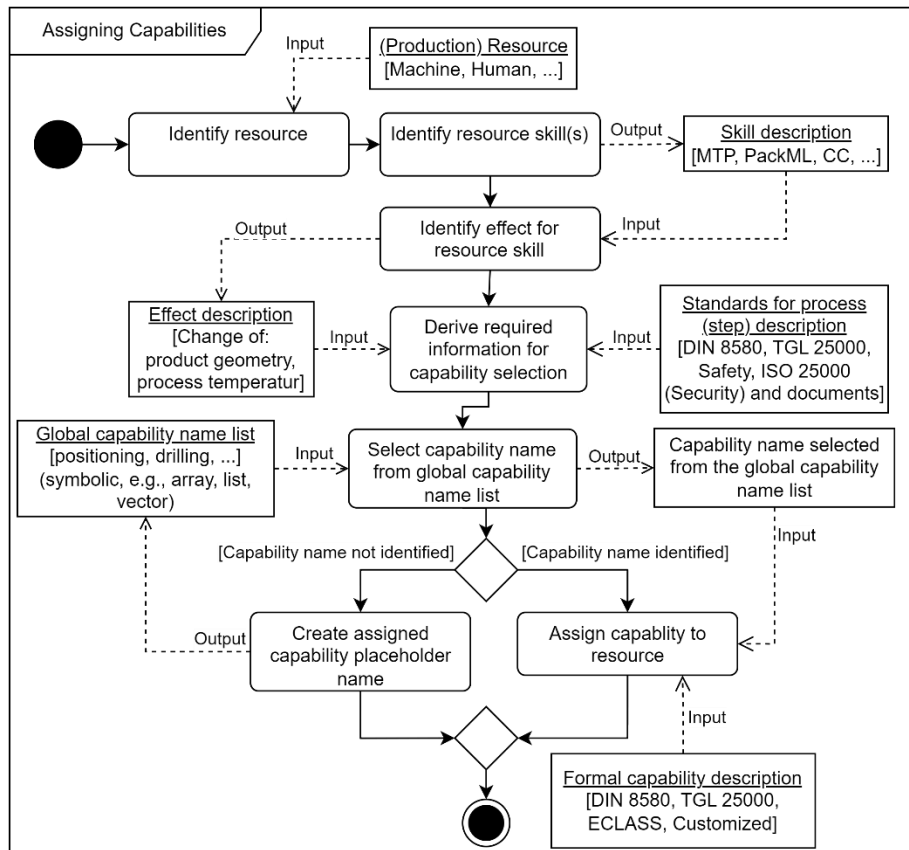


Fig. 3: Workflow for the method of Assigning Capabilities

For the selection of the capability to be assigned, further information to derive it is required in addition to the effect description. This information is gathered from standards or other documents describing processes or operations such as DIN 8580 or TGL 25000. Prepared with the derived information, the capability is to be selected from the global capability name list, as explained before. In addition, the capability name to select, is to ascertain whether the capability aligns with the resource's behavior and effect, as well as whether the capability is described sufficiently for the resource. Either the capability name is identified for an effect, or the capability name is not sufficient.

¹¹ The effect, as a formal element, is currently under research.

If the capability name for an effect is not identified, a new capability placeholder name has to be created in the global capability name list and it has to be assigned to the production resource. The placeholder is used as a provisional entity until a formal capability description for the given resource is provided by the method Describing Capabilities, see Section 3.1. In case that the capability name for an effect is identified and is also formally described, the capability can be directly assigned to the production resource using the capability name from the global capability name list. For the use of the formal capability description, the selected capabilities for the production resource are made accessible in a capability description model, e.g., the AAS capability description submodel or a resource ontology.

This method of assigning capabilities ensures a structured alignment with resources, fostering an effective utilization of own or external contractor assets, enabled by optimized allocation of these resources.

3.3 Deriving Capabilities

In addition to the two underlying methods described above, the third method focuses on extracting required capabilities from existing descriptions of production procedures or product properties, i.e. product specifications or process descriptions consisting of a logical sequence of process steps for manufacturing the product. The workflow of the method Deriving Capabilities is shown in Fig. 4.

A product specification or process description is necessary to derive required capabilities, which may be available in various forms, such as a computer-aided design (CAD) model, a digital twin, a product data model, a recipe, a procedure, a formal description of a process or other technical specifications. However, these specifications or descriptions alone are not always sufficient to fully identify the required capabilities and their sequence of execution for manufacturing a product.

After identifying the specifications and descriptions, a decision is made whether capabilities have to be derived from a product specification or from a process description. The product can have a modular structure. Consequently, the presence of product parts must be identified. A product part itself is a component of a product to be manufactured. Therefore, a product part is considered equivalent to a product in this contribution.

The objective of the identified products is to ascertain the effects on the product, thereby enabling the derivation of information about possible capabilities for manufacturing the product via a description of the effect. An effect imposed on a product can be, to give a specific example, the change imposed by drilling a desired hole into a piece of metal. The effect description is an artefact that will be used for deriving required information with regard to the capability selection. The next activity in the workflow identifies the process(es) either directly from process description or by the process(es) identified from product specification to be manufactured. Analogously to the product parts, a process can

be structured into process steps, which also have to be identified. Existing recipes, procedures or formal descriptions of the process allow the identification of individual process steps. Afterwards, the effects of the process(es) (process steps) have to be identified. For example, an effect that has to be imposed onto the process and the product can be their temperature which has to be within a certain range in order to drill the desired hole in the piece of metal.

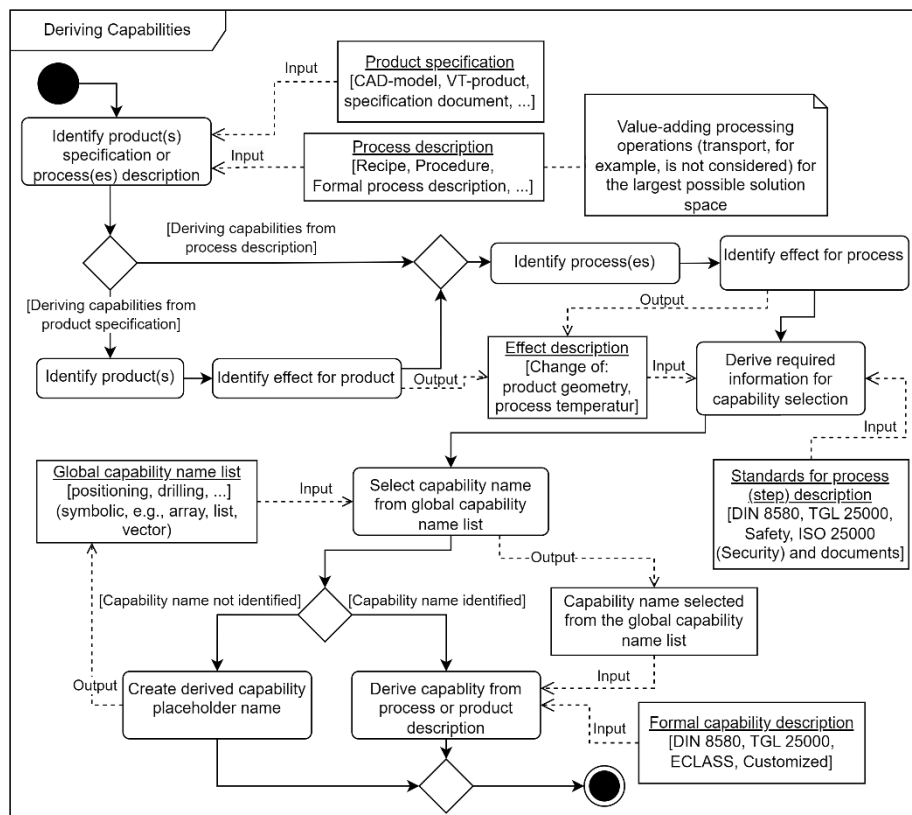


Fig. 4: Workflow for the method of Deriving Capabilities

The gathered information, including that pertaining to the identified effects as described in the effect description and to process-relevant information from standards such as DIN 8580 or TGL 25000, must be considered in order to derive the required information for capability selection. Analogously to the method Assigning Capabilities, the capability has to be selected from the global capability name list. Additionally, the capability name must be selected to determine whether the capability aligns with the effect of the product specification or the process description. The selection of a capability from the global capability name list is dependent on whether the capability id is identified as sufficient or

insufficient for an effect. In correspondence to the Assigning Capabilities method, if the capability name for an effect is not identified, a new capability placeholder name is to create in the global capability name list and be the derived capability. Accordingly, the placeholder is to be used as a provisional entity until a formal capability description for the derived capability is provided using the Describing Capabilities method, see Section 3.1. In cases where the name of a capability that imposes the desired effect is identified and also formally described, the capability can be directly derived from the product specification or process description. The method of deriving capabilities ensures a structured derivation of capabilities from product specifications or process descriptions.

4 Conclusion and Outlook

Facing rapidly evolving environments in manufacturing requires flexibility in production. The use of the CSS model as a means to implement Capability-based engineering enables decoupling of product design and production planning and execution. This contribution provides methods for the formal description of capabilities, their assignment to resources, and their derivation from product specifications or process descriptions. The methodical approach through the Capability Determining Method proposes a structure to encompass the description, assignment und derivation of capabilities. This technology-independent method guidelines stakeholders for navigating the complexities of capability-based engineering.

Our research identifies topics, which has to be validated by the working group of Plattform Industrie 4.0. A pilot implementation for the interaction with the CSS model as well as the validation of the methods is currently under development. Therefore, tool support is needed, also to involve various stakeholders. In addition to the current capability concept, the concept of effects was identified and included into the methods to some extent. The concept of effects will be elaborated in further publications.

References

- [Di22] Diedrich, C. et.al.: Information Model for Capabilities, Skills & Services, Plattform Industrie 4.0, Berlin, 2022, <https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/CapabilitiesSkillsServices.html>, accessed: 20.01.2024.
- [Fr22] Froschauer, R.; Köcher, A.; Meixner, K.; Schmitt, S.; Spitzer, F.: Capabilities and Skills in Manufacturing: A Survey Over the Last Decade of ETFA, 2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA), Stuttgart, Germany, pp. 1-8, 2022.
- [Gr23] Grüner, S. et. al.: Towards asset administration shell-based continuous engineering in process industries, at-Automatisierungstechnik; 71(8), S. 689-708, De Gruyter (O), Berlin, 2023.

-
- [HDM21] Huang, Y.; Dhouib, S. and Malenfant, J.: An AAS Modeling Tool for Capability-Based Engineering of Flexible Production Lines. In IECON 2021 – 47th Annual Conference of the IEEE Industrial Electronics Society, Toronto, Canada, pp. 1-6, 2021.
- [Jä19] Järvenpää, E.; Sitala, N.; Hylli, O.; Lanz, M.: The development of an ontology for describing the capabilities of manufacturing resources. *Journal of Intelligent Manufacturing*, vol. 30, no. 2, pp. 959–978, 2019.
- [Kö20] Köcher, A. et.al.: Automating the Development of Machine Skills and their Semantic Descriptions. In 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 1013–11018, 2020.
- [Kö23] Köcher, A. et.al.: A reference model for common understanding of capabilities and skills in manufacturing, *at–Automatisierungstechnik*; 71(2), S. 94–104, De Gruyter (O), Berlin, 2023.
- [Su09] Sure-Vetter, Y. et.al.: Ontology engineering methodology. In: Staab, S, Studer, R (Eds.), *Handbook on Ontologies*, pp. 135–152, 2009.
- [Ja13] Jakus, G. et.al.: *Concepts, Ontologies, and Knowledge Representation*. SpringerBriefs in Computer Science, DOI: 10.1007/978-1-4614-7822-5_1, 2013.

„Ulrich Jumar, Christian Diedrich (Hrsg.):
EKA 2024 - Entwurf komplexer Automatisierungssysteme, 18. Fachtagung“

Analyse und Konzeption von Informationsmodellen für Soft Robotics

Unified Soft Robotics Description Format

Linus Witucki¹, Mike Barth²

Abstract: Moderne flexible Fertigungs- und Montagelinien werden zunehmend durch Anwendungen aus dem Bereich der Mensch-Maschine-Kooperation und -Kollaboration ergänzt. Dabei spielen Cobots unterschiedlichster Hersteller eine wesentliche Rolle und bilden den Stand der Technik. In der Forschung werden neue Roboterklassen untersucht, welche das Zusammenspiel von Mensch und Maschine inhärent sicher ermöglichen. Hierzu zählt das Konzept von Soft Robots. In diesem Beitrag stellen die Autoren ein Konzept für deren Integration in modell-basierte Engineering-Methoden vor. Hierzu werden Informationsmodelle für Rigid Body Robotik, welche aktuell die Basis für die Robotik-Modellierung bilden, auf ihre Adaptierbarkeit auf „Piece-Wise Constant Curvature“ (PCC) Modelle der Soft Robotik analysiert. Diese mathematische Modellierung soll in bekannte Informationsmodelle integriert werden. Als Grundlage dient dabei das als neutraler Standard klassifizierte Unified Robotics Description Format (URDF). Zum Abschluss werden die kritischen Elemente der Erweiterung des Informationsmodells URDF identifiziert.


Keywords: Robotik, Soft Robotik, Informationsmodelle, URDF

1 Einleitung und Motivation

1.1 Notwendigkeit von Soft Robotics für moderne Fertigung

Im Kontext moderner flexibler und wandlungsfähiger Produktions- bzw. Montagekonzepte ist ein Fokusbereich der Forschung die Mensch-Roboter Kollaboration [Vz18]. Hierbei führen Mensch und Roboter parallel und physisch gekoppelt gemeinsam eine Aufgabe aus. Dabei werden die Vorteile des Roboters, Bewegungen schnell, präzise und häufig auszuführen, mit der Flexibilität, Kognition und Kreativität des Menschen verbunden [Vz18]. Den Stand der Technik für die Kollaboration bilden sogenannte Cobots, welche die Sicherheit durch eine umfangreiche Sensorik (z.B. Kraft-Moment-Sensoren in den Gelenken) sicherstellen [Vz18]. Als neuartige Roboterklasse, welche die

¹ Karlsruher Institut für Technologie, Institut für Regelungs- und Steuerungssysteme, Fritz-Haber-Weg 1,
76131 Karlsruhe, linus.witucki@kit.edu,  <https://orcid.org/0009-0003-4973-9471>

² Karlsruher Institut für Technologie, Institut für Regelungs- und Steuerungssysteme, Fritz-Haber-Weg 1,
76131 Karlsruhe, Mike.Barth@kit.edu,  <https://orcid.org/0000-0003-2337-063X>

Mensch-Maschine-Kollaboration inhärent konstruktiv sicherstellt, können dabei zukünftig auch sogenannte Soft Robots zur Anwendung kommen [Lz17]. Diese sind aus weichen flexiblen Materialien, wie (Kunst-)Stoff oder Silikon, gefertigt und werden zum Beispiel über Pneumatik angesteuert [Lz17]. Ein solcher Roboter ist in Abb. 1 dargestellt.

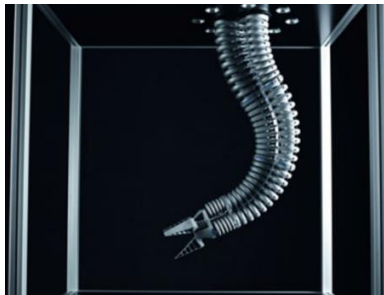


Abb. 1: Festo's Bionic Handling Assistant Quelle: Festo AG.

Durch den Aufbau des Roboters führen äußere Kräfte zu einer Verformung des Roboters, welcher ein Feder-Dämpfer-Verhalten annähert. Somit sind Soft-Robots inhärent sicher für die physische Kopplung von Mensch und Roboter [Lz17].

1.2 Notwendigkeit von Informationsmodelle für Engineering von Robotik Systemen

Im Kontext der Fertigungs- und Prozessindustrie kommen zahlreiche Roboter zum Einsatz. Dabei kann der Hersteller und Typ der Roboter innerhalb einer Fertigung stark variieren. Um die Heterogenität und Komplexität der Robotersysteme im Engineering und Betrieb beherrschbar und handhabbar zu halten, müssen verwendete Software- und Administrationstools mit den unterschiedlichen Systemen umgehen können [Cz16]. Für die Handhabung solcher komplexen Systeme werden bereits Strukturen wie die Verwaltungsschale entwickelt [Yz21]. Bei einer hierdurch entstehenden domänen- und systemübergreifenden Struktur steigt jedoch auch die Bedeutung des Informationsaustausches [Sz21, Yz21]. Hierzu können Informationsmodelle eingesetzt werden, um den Austausch von Engineering-Artefakten oder Betriebsdaten zu strukturieren und mittels einer definierten Semantik Fehlinterpretationen vorzubeugen [Sz21, Yz21]. Somit kann zum Beispiel der Einsatz von Heterogenen Roboterteams vereinfacht werden [Lz17b]. Anhand der eingesetzten Informationsmodelle können die Eigenschaften eines Roboters strukturiert und maschinenlesbar modelliert werden. Für herkömmliche Knickarmroboter, existiert unter anderem das Unified Robotics Description Format (URDF) zur Modellierung relevanter Merkmale, wie beispielsweise Armlängen, Gelenkarten und Belastungsgrenzen. Für den Einsatz von Soft Robots im industriellen Kontext existieren bislang keine eigenständigen Informationsmodelle. Gleichzeitig sind, wie im Folgenden erläutert wird, die bestehenden Modelle nicht geeignet um die Eigenschaften von Soft Robots abzubilden.

2 Bestehende Informationsmodelle der Robotik

Im Umgang mit herkömmlichen Knickarmrobotern wird maßgeblich das Unified Robotics Description Format (URDF) als Informationsmodell eingesetzt [TD24]. Mit diesem auf XML basierenden Informationsmodell werden Roboter als Kette von starren Verbindungen (sogenannte Links) und punktförmigen Gelenken (sogenannte Joints) beschreiben [TD24]. An dieser Stelle sollen weitere Vertreter der Informationsmodelle für Rigid Body Roboterarme genannt werden. Hierzu zählen das Simulation Description Format (SDF) und das Universal Scene Description (USD) Format [TD24]. Diese sind in der Lage, neben dem Knickarmroboter, weitere Aspekte, wie die direkte Umgebung des Roboters zu beschreiben. Für die Definition eines Roboters werden jedoch grundsätzlich die gleichen Elemente und Attribute verwendet. Da in dieser Ausarbeitung nur auf die Modellierung der Roboterarme an sich eingegangen wird, werden die erweiterten Aspekte der Informationsmodelle USD und SDF nicht betrachtet. Im Folgenden wird in dieser Ausarbeitung eine Analyse des URDF durchgeführt. Zunächst folgt eine konzeptionelle Analyse der Modellierungsmethode des URDF-Informationsmodells. Auf Basis dieser und der mathematischen Auseinandersetzung kann anschließend der Handlungsbedarf in Bezug auf die Modellierung von Soft Robots abgeleitet werden.

2.1 Analyse und Bewertung des Stands der Technik

Ein im URDF modellierter Roboter setzt sich aus zwei Komponenten zusammen. Dem steuerbaren Joint und der starren Verbindung zwischen zwei Joints: dem Link. Ein Joint beschreibt dabei ein Gelenk als punktförmig, d.h. ohne räumliche Ausdehnung. Die Definition von diskreten Gelenken unterscheidet einen Knickarmroboter wesentlich von einem kontinuierlichen Soft Robot. Dem Joint werden im URDF die Elemente „origin“, „parent“, „child“ und weitere optionale Elemente zugeordnet. Auf Basis dieser drei erforderlichen Elemente, sowie der Rotationsachse definiert im Element „axis“, lässt sich die kinematische Kette des Roboters ableiten. Über ein weiteres Attribut „type“ kann definiert werden, um welche Gelenkart es sich handelt. Für die Definition von Knickarmrobotern kommt maßgeblich der Typ „revolute“, über den ein Drehgelenk mit Achsbegrenzungen definiert ist, zum Einsatz. Zusätzlich sind beispielsweise prismatische Gelenke definiert. Neben dem Joint ist ein Link als starre Verbindung zwischen Joints im URDF enthalten. Dabei weist der Link im Gegensatz zum Joint eine räumliche Ausdehnung auf, welche durch die Elemente „visual“ und „collision“ repräsentiert wird. Des Weiteren können über das Element „inertial“ die dynamischen Eigenschaften des Links abgebildet werden. Dieses beinhaltet die Definition einer punktförmigen Masse und des dadurch entstehenden Trägheitsmoments. Mittels einer Abfolge von Links und Joints kann somit ein beliebiger Knickarmroboter beschrieben werden. Die erfolgte Aufzählung betrachtet nur Elemente und Attribute mit direkter Relevanz für die Steuerung und Regelung von Knickarmrobotern sind. Dementsprechend ist die Aufzählung nicht vollständig.

2.2 Zusammenhang mit mathematischen Grundlagen von Knickarmrobotern

Für die Regelung eines Roboterarms in kartesischen Koordinaten werden die Vorwärts- und Inverse-Kinematik verwendet. Hierbei handelt es sich um die mathematischen Transformationen vom Gelenkwinkel- in den Arbeitsraum, der in kartesischen Koordinaten die Pose eines definierten Punkts entlang des Roboters beschreibt. Als Definitionspunkt der Roboterbewegung im Arbeitsraum wird der Endeffektor bzw. der „Tool Center Point“ (TCP) als charakteristischer Punkt des Endeffektors gewählt. Die Berechnungen basieren auf der kinematischen Kette des Roboterarms, welche die mechanischen Zusammenhänge der einzelnen Abschnitte eines Roboters beschreibt [DP04]. Somit kann der Zusammenhang zwischen dem kartesischen Arbeitsraum und dem Gelenkwinkelraum des Roboters hergestellt werden. Die kinematische Kette kann für einen Knickarmroboter anhand der in Abschnitt 1.2 benannten „parent“ und „child“ Elemente der Joints kann ein hierarchischer Zusammenhang der Gelenke bestimmt werden. Darauf aufbauend können mittels der Elemente „origin“ und „axis“ die Rotation definiert werden.

Mittels der Bewegungsgleichung (1) können die für eine Bewegung des Roboters benötigten Drehmomente τ der Gelenke bestimmt werden.

$$J(q)\ddot{q} + C(q, \dot{q})\dot{q} + g(q) = \tau \quad (1)$$

Die Matrix des Trägheitsmoments $J(q)$ ist im Gelenkwinkelraum q definiert. Die Matrix $C(q, \dot{q})$ vereint Zentrifugal- und Coriolis-Effekte und ist abhängig von der partiellen Ableitung der Einträge der Masse $M(q)$. Der Vektor $g(q)$ beschreibt den Einfluss der Gravitation auf die starren Abschnitte und ist dementsprechend abhängig von der Masse eines jeden Abschnitts [H80]. Für die Drehmomentregelung eines Knickarmroboters müssen demnach diese Parameter für alle Gelenke im Informationsmodell abgebildet sein. Aus der URDF Datei eines Roboters können diese Parameter anhand des „inertial“ Element der Links entnommen werden. In den zusätzlich genannten Informationsmodellen liegen die beschriebenen Parameter in gleicher Weise vor. Somit ist das erläuterte Vorgehen äquivalent.

Auf Basis dieser Grundlagen kann in den folgenden Abschnitten analysiert werden, in wie weit bestehende Informationsmodelle geeignet sind um Soft Robots darstellen zu können. Dafür wird als nächstes der Handlungsbedarf erörtert.

3 Modellierung von Soft Robot Kontinuummanipulatoren

Ein Soft Robot wie in Abb. 1 dargestellt weist eine kontinuierliche Krümmung auf. Somit bildet der Soft Robot ein unendlich dimensionales kontinuierliches System [Sz23]. Ein solcher Roboter wird in dieser Arbeit daher als Soft Robot Kontinuummanipulator (SRKM) bezeichnet. Des Weiteren wird ein SRKM in dieser Arbeit in einige

grundlegende Komponenten unterteilt. Die Pose des SRKM kann über die in Abb. 2 eingezeichnete theoretische infinitesimal dünne Krümmungslinie, die im inneren des Roboters verläuft, definiert werden. Der reale SRKM weist jedoch eine räumliche Ausdehnung auf, die durch ein zusätzliches Volumen um die Krümmungslinie herum definiert werden kann. Da ein SRKM in der Anwendung voraussichtlich aus mehreren aktuierten Segmenten bestehen wird, existieren diskrete Segmentlinien, die den Übergang zweier Segmente definiert.

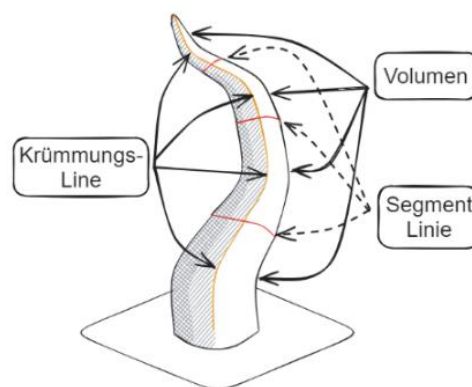


Abb. 2: Abstrahierte Darstellung eines Soft Robot Kontinuummanipulators

Für die mathematische Beschreibung der Krümmungslinie gibt es in der Literatur eine Vielzahl an Möglichkeiten. Neben Ansätzen zur exakten Beschreibung, wie beispielsweise die „Cosserat Rod“ Theorie [Rz10] werden hierbei auch Methoden mit reduzierter Modellordnung [Sz23b] verwendet. Unter den Ansätzen reduzierter Modellordnung ist auch die sogenannte „Pseudo Rigid-Body“ Methode [S09]. Hierbei handelt es sich um die Modellierung eines kontinuierlichen Roboterarmes durch eine finite Anzahl an starren Verbindungen und Gelenken. Durch diese Annahme werden systematische Fehler in die Berechnung der Vorwärts- und Rückwärts-Kinematik eingebracht. Diese äußern neben einem Fehler in der Position des Endeffektors insbesondere durch eine abweichende Form des Roboters [S09]. Durch die Annäherung des SRKM als Knickarmroboter kann ein so modellierter Roboterarm theoretisch im URDF abgebildet werden. Diese Methode wird nicht weiter erläutert, da dieser Beitrag Modellierungsmethoden mit kontinuierlicher Krümmung betrachtet.

In dieser ersten Betrachtung von Informationsmodellen für SRKM soll die sogenannten „Piece-Wise Constant Curvature“ (PCC) Approximation verwendet werden. Hierbei handelt es sich um die Approximation der Krümmung entlang des SRKM durch Segmente mit konstanter Krümmung. Diese Approximation vereinfacht die Berechnung der Pose des Roboters erheblich [Cz20] und die Beschreibung der Krümmung im Informationsmodell. Auf diese Eigenschaft wird in Abschnitt 4 noch einmal aufgegriffen und erläutert. Die beschriebene PCC Methode wird bereits erfolgreich zur Steuerung von bestehenden SRKM eingesetzt [WJ10].

3.1 Berechnung der Pose eines PCC modellierten Kontinuumsmanipulators

Wie bei der Beschreibung von Knickarm Robotern ist auch für SRKM die Definition einer Vorwärts-Kinematik möglich. Diese basiert auf einem Tupel aus charakteristischen Variablen $\mathbf{k}_i = [\phi_i \ \kappa_i \ l_i]^T$, mit dem der geometrische Zusammenhang von Referenzkoordinatensystemen entlang des Roboterarms beschrieben wird. Dabei handelt es sich um die Krümmung κ_i , die Segmentlänge l_i und die Verdrehung der Krümmungsebene ϕ_i eines Segmentes des SRKM. In Abb. 3a ist der Konfigurationsraum eines schematischen SRKM abgebildet. Die gesamte Pose des Roboters lässt sich dann durch die Tupel \mathbf{k}_1 bis \mathbf{k}_4 , also den Parametern jedes Segmentes beschreiben. [Lz23, Sz23]

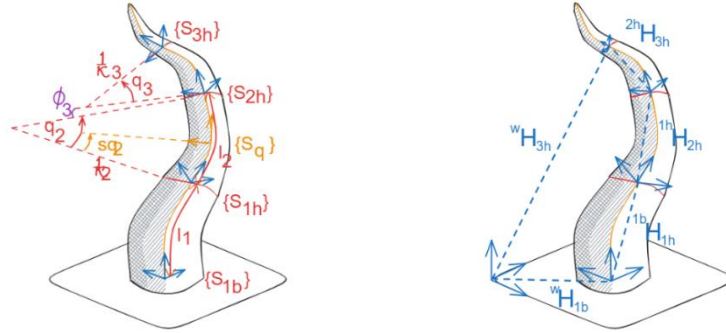


Abb. 3: a) Konfigurationszustände und b) Koordinatentransformation eines PCC SRKM [Fz15a]

Anhand der der Konfigurationsvariablen kann, wie in Abb. 3b dargestellt, für jedes Segment des SRKM eine Transformationsmatrix jH_i gebildet werden, die den geometrischen Zusammenhang von Segment S_j zum folgenden Segment S_{i+1} bildet. [Lz23]

$${}^jH_i = \begin{bmatrix} {}^jR_i & {}^j\sigma_i \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4} \quad (3)$$

Wobei mit $q_i = \frac{1}{\kappa_i}$ gilt:

$${}^jR_i = \begin{bmatrix} \cos^2 \phi_i (\cos q_i - 1) + 1 & \sin \phi_i \cos \phi_i (\cos q_i - 1) & \cos \phi_i \sin q_i \\ \sin \phi_i \cos \phi_i (1 - \cos q_i) & \sin^2 \phi_i (\cos q_i - 1) + 1 & \sin \phi_i \sin q_i \\ -\cos \phi_i \sin q_i & -\sin \phi_i \sin q_i & \cos q_i \end{bmatrix}, \quad (4)$$

$${}^j\sigma_i = \frac{l_i}{q_i} \begin{bmatrix} \cos \phi_i (1 - \cos q_i) \\ \sin \phi_i (1 - \cos q_i) \\ \sin q_i \end{bmatrix} \quad (5)$$

Mit dieser Definition der Vorwärts-Kinematik kann die Koordinatentransformation der unterschiedlichen Abschnitte eines PCC SRKM zueinander und zur globalen Referenz, der Basis, wie in Abb. 3 b) dargestellt, stattfinden. Die Berechnungen basieren dabei auf dem oben definierten Tupel \mathbf{k}_i . Um die Bestimmung einer Pose des SRKM mittels eines Informationsmodells zu ermöglichen, müssen diese Variablen entweder als zeitunabhängig oder zeitabhängige Komponente definiert werden. [Lz23]

3.2 Berechnung der Dynamik eines PCC modellierten Kontinuummanipulators

Neben der statischen Pose kann auch die Dynamik des SRKM berechnet werden. Hierzu wird über die Ableitung der Vorwärts-Kinematik die Jacobi-Matrix $J_i = \frac{\delta H}{\delta q}$ gebildet, welche benötigt wird, um das Trägheitsmoment $M(q)$ des Systems zu bestimmen. Die Berechnung kann in [Wz17], [Fz15a] und [Fz15b] nachvollzogen werden. Ebenso wird für die Berechnung der Coriolis-Effekte $C(q, \dot{q})$ auf [Wz17], [Fz15a] und [Fz15b] verwiesen. Ein Soft Robot kann durch seine Bauweise als Masse-Feder-Dämpfer System approximiert werden. Daher existieren, zusätzlich zu den in Abschnitt 3.1 dargestellten Matrizen, Feder $K(q)$ und Dämpfer $D(q)$ Einflüsse auf die Bewegungsgleichung.

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) + D(q)\dot{q} + K(q) = A(q)\tau$$

Die Matrix $A(q)$ beschreibt die Aktuiertheit des Systems. Ist jedes Segment separat ansteuerbar, dann ist die Matrix $A(q)$ quadratisch und weist einen vollen Rang auf [Wz17]. Insbesondere die Feder- und Dämpfer-Beschreibung stellt bei dieser Berechnung, aufgrund der Nichtlinearitäten, eine bedeutende Herausforderung dar. Für die Modellierung der Eigenschaft existieren einige Ansätze. Darunter Ansätze der Linearisierung [Kz19] und Identifikation nichtlinearer Parameter [Fz15a].

Damit die mathematischen Berechnungen von Pose und Dynamik möglich sind, müssen die notwendigen Parameter in einem Informationsmodell explizit enthalten oder daraus implizit ableitbar sein.

3.3 Abgeleiteter Handlungsbedarf für Informationsmodelle

In Abschnitt 2.2 wurde gezeigt, dass es mit dem URDF Informationsmodell möglich ist einen Knickarmroboter abzubilden. Die Analyse aus Abschnitt 3.1 und 3.2 zeigt jedoch grundlegende Unterschiede zu SRKM. Diese führen dazu, dass ein SRKM nicht ausreichend mit dem URDF abgebildet werden kann.

Eine konzeptionelle Basis des bestehenden Informationsmodells ist die grundlegende Unterteilung in Joint und Link mit jeweils exklusiven Elementen und Attributen. Stattdessen weist ein SRKM eine kontinuierliche Krümmung auf. Im Hinblick auf das URDF liegt dementsprechend in jedem Segment eine Kombination aus Joint und Link vor. Dies kann so im bestehenden URDF nicht abgebildet werden. Des Weiteren ist die

Definition des Joints auf ein diskretes Set an Gelenken beschränkt. Darunter ist kein Gelenktyp mit dem eine mathematische Beschreibung der ausgedehnten Krümmung eines Soft Robot möglich ist. Darüber hinaus wird bislang die Masseverteilung stets punktförmig approximiert. Für die Berechnung exakte Berechnung der Krümmung eines Soft Robots im Einfluss von Gravitation ist jedoch die Masseverteilung entlang eines Segmentes zu berücksichtigen [Sz23]. Somit muss für SRKM auch diese Beschränkung konzeptionell überwunden werden. Als weiter wichtiger Punkt kann die Feder-Dämpfer Eigenschaft des SRKM aufgeführt werden. Es findet sich im URDF kein korrespondierendes Element zu den in Abschnitt 3.2 beschriebenen Parameter Parametern wieder. Aus diesen Umständen heraus entsteht der Bedarf nach einen Informationsmodell, welches in der Lage ist SRKM abzubilden.

4 Konzeptionelle Erweiterung zur PCC Informationsmodellierung

Aus Abschnitt 3.3 geht hervor, dass eine Modellierung von SRKM mit dem URDF nicht ausreichend möglich ist. Hier folgt nun eine Analyse der relevanten Parameter für die Anpassung der bestehenden Informationsmodell. Hierbei werden im ersten Entwicklungsschritt nur PCC modellierte SRKM betrachtet.

4.1 Notwendige Parameter der „Piece-Wise Constant Curvature“ Modellierung

Die in Abschnitt 3 dargelegten mathematischen Zusammenhänge ermöglichen die Definition von charakteristischen Parametern. Diese müssen in einem Informationsmodell vorhanden sein, um den SRKM vollständig beschreiben zu können. Darunter die Masseverteilung $M(q)$ und das Trägheitsmoment $J(q)$ eines Segmentes, sowie die beiden Parameter zur Beschreibung der Steifigkeit $K(q)$ und Dämpfung $D(q)$. Des Weiteren muss eindeutig definiert werden, welche zeitlich variablen Stellgrößen vorliegen. Bei einem herkömmlichen Knickarmroboter mit „revolute“ Joints sind das z.B. die Gelenkwinkel. Dies wird durch das „type“ Attribut des Joints im URDF gekennzeichnet. Um zu definieren, dass bei einem SRKM die Krümmung eine Stellgröße darstellt muss eine entsprechende Definition möglich sein. Ebenso kann bei einem SRKM die Länge eines Segments und die Verschiebung der Krümmungsebenen gestellt werden. Eine Definition der jeweiligen physikalischen Grenzen muss wie beim „revolute“ Joint ebenfalls möglich sein. Da ein SRKM aus mehreren Segmenten bestehen kann muss die Angabe der Reihenfolge möglich sein. Neben den Parametern zur mathematischen Berechnung existieren noch weitere, die für ein Informationsmodell wichtig sind. Dazu gehört die räumliche Ausdehnung des Soft Robots, z.B. definiert durch ein Mesh. Für die Anwendung in physikalischen Simulationen kann es wichtig sein, eine separate Beschreibung der Kollisionsform des Roboterarmes zu definieren. Da die Berechnung von Kollisionen mit komplexen Geometrien z.B. über die Approximation von konvexen Hüllen eine Komplexität von $\mathcal{O}(n \log(n))$ [Bz96].

Diese beschriebenen Aspekte und Parameter müssen bei der Erweiterung eines Informationsmodells für SRKM berücksichtigt werden. Da das URDF einige Elemente, wie Volumen-Meshes und Elemente zur Beschreibung der Reihenfolge von Roboterkomponenten, bereits aufweist, kann das URDF als Basis zur Erweiterung verwendet werden.

5 Zusammenfassung und Ausblick






In dieser Arbeit wurde auf Basis der Analyse von herkömmlichen Informationsmodellen der Robotik und den mathematischen Grundlagen von Knickarmrobotern und SRKM der Handlungsbedarf für die Erweiterung der bestehenden Informationsmodelle hergeleitet. Anschließend wurden die für die Erweiterung des URDF relevanten Aspekte der SRKM herausgearbeitet. Für die Umsetzung müssen die Eigenschaften in Elementen und anschließend Attributen formalisiert werden. Damit möglichst viele unterschiedliche Modellierungsmethoden von SRKM mit dem Informationsmodell abgedeckt werden können, bedarf es einer entsprechend strukturierten Übersicht dieser Modellierungsmethoden. Zur Validierung können Tools entwickelt werden, die das neue Informationsmodell anwenden. Darunter z.B. ein Tool für die automatisierte Parametrierung einer Simulationsumgebung für Soft Robots.

6 Literaturverzeichnis

- [Bz96] Barber, C. B., Dobkin, D. P., Huhdanpaa, H.: The quickhull algorithm for convex hulls. In *ACM Transactions on Mathematical Software (TOMS)* 22.4: 469–483, 1996.
- [Cz16] Cai, Y.; Tang, Z.; Ding, Y.; Qian, B.: Theory and application of multi-robot service-oriented architecture. In *IEEE/CAA Journal of Automatica Sinica*, 3(1), 15–25, 2016.
- [DP04] Géradin, M.; Duysinx, P.: *An introduction to robotics: mechanical aspects*, 2004.
- [Fz15a] Falkenhahn, V.; Mahl, T.; Hildebrandt, A.; Neumann, R.; Sawodny, O.: Dynamic Modeling of Bellows-Actuated Continuum Robots Using the Euler–Lagrange Formalism. In *IEEE Transactions on Robotics*, Bd. 31, Nr. 6, S. 1483–96, 2015.
- [Fz15b] Falkenhahn, V.; Hildebrandt, A.; Neumann, R.; Sawodny, O.: Model-based feedforward position control of constant curvature continuum robots using feedback linearization. In *IEEE International Conference on Robotics and Automation (ICRA)*, S. 762–67, 2015.
- [Kz19] Katzschmann, R. K.; Santina, C. D.; Toshimitsu, Y.; Bicchi, A.; Rus, D.: Dynamic Motion Control of Multi-Segment Soft Robots Using Piecewise Constant Curvature Matched with an Augmented Rigid Body Model. In *2nd IEEE International Conference on Soft Robotics*, S. 454–461, 2019.
- [Lz23] Li, Y.; Myszka, D. H.; Murray, A.: The Kinematics of Constant Curvature Continuum Robots Through Three Segments. In *IEEE Robotics and Automation Letters*, Bd. 8, Nr. 11, S. 7631–38. 2023

-
- [Sz21] Schmied, S.; Mathias, S. G.; Großmann, D.; Müller, R. K.; Jumar, U.: Information modelling with focus on existing manufacturing systems. In *Annual Reviews in Control*, 51, 392-400, 2021.
- [Sz23] Della Santina, C.; Duriez, C.; Rus, D.: Model-Based Control of Soft Robots: A Survey of the State of the Art and Open Challenges. In *IEEE Control Systems Magazine*, Bd. 43, Nr. 3, S. 30–65, 2023.
- [Wz17] Wang, H.; Wang, C.; Chen, W., Liang, X., Liu, Y.: Three-Dimensional Dynamics for Cable-Driven Soft Manipulator. In *IEEE/ASME Transactions on Mechatronics*, Bd. 22, Nr. 1, S. 18–28, 2017.
- [Lz17] Lee, C.; Kim, M.; Kim, Y. J.; Hong, N.; Ryu, S.; Kim, H. J.; Kim, S.: Soft robot review. In *International Journal of Control, Automation and Systems*, 15(1), 3–15, 2017.
- [TD24] Tola, D.; Corke, P.: Understanding URDF: A Dataset and Analysis. *IEEE Robotics and Automation Letters*, 9(5), 4479–4486, 2024.
- [Vz18] Villani, V.; Pini, F.; Leali, F.; Secchi, C.: Survey on human–robot collaboration, In industrial settings: Safety, intuitive interfaces and applications. *Mechatronics*, 55, 248–266, 2018.
- [S09] Su, H.-J.: A Pseudorigid-Body 3R Model for Determining Large Deflection of Cantilever Beams Subject to Tip Loads. In *Journal of Mechanisms and Robotics*, 1(021008), 2009.
- [Cz20] Cheng, H.; Liu, H.; Wang, X.; Liang, B.: Approximate Piecewise Constant Curvature Equivalent Model and Their Application to Continuum Robot Configuration Estimation. In *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1929–1936, 2020.
- [WJ10] Webster, R. J.; Jones, B. A.: Design and Kinematic Modeling of Constant Curvature Continuum Robots: A Review. *The International Journal of Robotics Research*, 29(13), 1661–1683, 2010.
- [Lz17b] Li, X.; Bilbao, S.; Martín-Wanton, T.; Bastos, J.; Rodriguez, J.: SWARMS Ontology: A Common Information Model for the Cooperation of Underwater Robots. In *Sensors*, 17(3), Article 3, 2017.
- [Rz10] Rucker, D. C.; Jones, B. A.; Webster III, R. J.: A Geometrically Exact Model for Externally Loaded Concentric-Tube Continuum Robots. In *IEEE Transactions on Robotics*, 26(5), 769–780, 2010.
- [Sz23b] Sadati, S. M. H.; Naghibi, S. E.; da Cruz, L.; Bergeles, C.: Reduced order modeling and model order reduction for continuum manipulators: An overview. In *Frontiers in Robotics and AI*, 10, 1094114, 2023.
- [Yz21] Ye, X.; Hong, S. H.; Song, W. S.; Kim, Y. C.; Zhang, X.: An Industry 4.0 Asset Administration Shell-Enabled Digital Solution for Robot-Based Manufacturing Systems. *IEEE Access*, 9, 154448–154459, 2021.
- [H80] Hollerbach, J. M.: A Recursive Lagrangian Formulation of Manipulator Dynamics and a Comparative Study of Dynamics Formulation Complexity. In *IEEE Transactions on Systems, Man, and Cybernetics*, 10(11), 730–736, 1980.

SysML-Profil zur automatisierten Generierung von PDDL-Beschreibungen aus Systemmodellen

Hamied Nabizada ¹, Lasse Beers ¹, Maximilian Weigand ¹, Felix Gehlhoff ¹ und Alexander Fay ²





Abstract: In diesem Beitrag wird die Entwicklung eines dedizierten SysML-Profiles vorgestellt, das die Einbettung von Planning Domain Definition Language (PDDL) Konstrukten direkt in SysML-Modelle ermöglicht. Dieser Ansatz unterstützt die automatisierte Prozessablaufplanung, indem es eine Schnittstelle zwischen Systemmodellierung und Planungsalgorithmen schafft. Durch die Anwendung des Profils in einem Anwendungsfall aus dem Flugzeugbau wird gezeigt, wie Planungsbeschreibungen effizient erstellt und für Optimierungsprozesse in PDDL-Solvern genutzt werden können.


Keywords: AI Planning, Modellbasierte Systementwicklung, PDDL, SysML

1 Einleitung

Die modellbasierte Systementwicklung (MBSE) hat sich als Alternative zur traditionellen, dokumentenbasierten Systementwicklung etabliert, indem sie den Einsatz von Modellen zur detaillierten und durchgängigen Beschreibung von Produktionssystemen fördert [HS21]. Durch diese einheitliche Datenbasis wird eine effiziente Zusammenarbeit der beteiligten Gewerke im Entwicklungsprozess geschaffen [SS20].

Methoden der MBSE, wie z.B. die Systems Modeling Language (SysML)-gestützten Modellierungs-Vorgehensweisen von [We16] oder [Ei15], bieten strukturierte Ansätze zur Erstellung von Systemmodellen. Diese strukturierten Vorgehensweisen reduzieren die Heterogenität der Inhalte und der Darstellung einzelner Teilmodelle des Systems und fördern somit die Wiederverwendbarkeit und Vergleichbarkeit der erstellten Modelle [Es07]. Allerdings erfordern diese Methoden oft viele manuelle Schritte, wie z.B. die Zuordnung von vielen einzelnen Prozessschritten zu möglichen technischen Ressourcen, die diese Prozesse umsetzen können, welches ein hohes Maß an Expertenwissen benötigt. Dies trifft insbesondere zu, wenn verschiedene Systemvarianten verglichen werden sollen, um die ideale Konfiguration für spezifische Optimierungsziele zu bestimmen. Bei der Analyse verschiedener Systemvarianten wird sowohl die Gesamtstruktur als auch das Verhalten

1 Helmut-Schmidt-Universität / Universität der Bundeswehr, Institut für Automatisierungstechnik, Holstenhofweg 85, 22043 Hamburg, Deutschland, hamied.nabizada@hsu-hh.de,  <https://orcid.org/0000-0001-8251-837X>; lasse.beers@hsu-hh.de,  <https://orcid.org/0000-0003-0946-7742>; maximilian.weigand@hsu-hh.de,  <https://orcid.org/0009-0000-1602-4873>; felix.gehlhoff@hsu-hh.de,  <https://orcid.org/0000-0002-8383-5323>

2 Ruhr-Universität Bochum, Lehrstuhl für Automatisierungstechnik, Universitätsstrasse 150, 44801 Bochum, Deutschland, alexander.fay@rub.de,  <https://orcid.org/0000-0002-1922-654X>

der Subsysteme betrachtet. Die beiden Aspekte, die Gesamtstruktur des Systems und das Verhalten der einzelnen Subsysteme, müssen in Beziehung zueinander gesetzt werden, um untersuchen zu können, wie Änderungen in einem Teil des Systems das Verhalten des Gesamtsystems beeinflussen. Zum Beispiel kann eine Variante eines Robotik-Systems, die über mehrere parallel arbeitende Endeffektoren verfügt, einen Produktionsprozess schneller durchführen als eine Variante mit nur einem Endeffektor. Allerdings sind bei der erstgenannten Variante Leerlaufzeiten der zusätzlichen Endeffektoren zu erwarten. Sollte in einem Optimierungsszenario die Maximierung der Ressourcenauslastung angestrebt werden, so würde der zusätzliche Endeffektor diesem Ziel widersprechen.

Im Rahmen der Systementwicklung muss ermittelt werden, welche Systemvariante am besten geeignet ist, um eine festgelegte Menge von Prozessen unter Berücksichtigung von Prozessbedingungen effizient durchzuführen [Tö17]. Dafür kann beispielsweise die Durchlaufzeit herangezogen werden. Um diese berechnen zu können, wird ein Prozessablaufplan benötigt, der die genaue Abfolge der auszuführenden Aktionen definiert und diese den entsprechenden Subsystemen zuordnet. Eine Automatisierung dieser Prozessablaufplanung ist erstrebenswert, um den zeitlichen Aufwand für die Bewertung verschiedener Systemvarianten zu reduzieren.

Um diese Automatisierung zu ermöglichen, ist es notwendig spezifische Planungsaspekte in die Systemmodellierung zu integrieren. Planungsaspekte beziehen sich auf die Definition von Aktionen, die zur Erreichung spezifischer Ziele innerhalb des modellierten Systems notwendig sind. Diese Aspekte sind zentral im Forschungsbereich des *AI Plannings*, der sich mit der Entwicklung von Algorithmen und Sprachen zur Lösung von Planungsproblemen beschäftigt [GNT16]. Aus diesem Forschungsbereich stammt die *Planning Domain Definition Language (PDDL)*, die sich als de-facto Standard für die Beschreibung von Planungsproblemen etabliert hat [Ma22b]. Daher wird in diesem Beitrag ein speziell entwickeltes *SysML*-Profil für PDDL vorgestellt. Dieses Profil ermöglicht es, PDDL-spezifischen Konstrukte direkt in *SysML*-Modelle zu integrieren und somit die Automatisierung der Prozessplanung in der Systementwicklung unterstützt.

In Abschnitt 2 werden zunächst die technischen Grundlagen dargelegt und eine Übersicht über verwandte Arbeiten gegeben. Das entwickelte *SysML*-Profil für PDDL wird in Abschnitt 3 vorgestellt. Anschließend erfolgt in Abschnitt 4 die Anwendung dieses Profils auf die Strukturmontage im Flugzeugbau. Der Beitrag schließt mit einer Zusammenfassung und einem Ausblick in Abschnitt 5 ab.

2 Grundlagen und verwandte Arbeiten

2.1 SysML und Profilmechanismus

Die Unified Modeling Language (UML) ist eine von der Object Management Group (OMG) standardisierte grafische Modellierungssprache, die ursprünglich für die Softwareentwick-

lung entwickelt wurde. Sie bietet eine breite Vielfalt an Elementen und Diagrammen, um verschiedene Aspekte eines Systems zu modellieren [OM19]. Diese Elemente und Diagramme ermöglichen die grafische Darstellung der Struktur, des Verhaltens und der Interaktion eines Softwaresystems.

Der UML-Profilmechanismus, ein zentraler Bestandteil des UML-Metamodells, ermöglicht die Erweiterung der UML mit spracheigenen Mitteln. Auf diese Weise kann die Sprache durch die Integration domänenspezifischer Konstrukte und Notationen an die spezifischen Anforderungen bestimmter Systeme angepasst werden [We08]. Das bekannteste Beispiel für ein UML-Profil ist die Sprache SysML. Sie wurde speziell für die Modellierung komplexer Systeme entwickelt und hat sich mittlerweile als führende Modellierungssprache im Bereich der MBSE etabliert [Ma22a]. Der Profilmeechanismus bietet jedoch auch die Möglichkeit zur individuellen Erweiterung der Sprache. Ein Profil definiert dabei Stereotypen, Tagged Values, Derived Properties, Constraints und weitere Konstrukte, die zur Erfüllung der spezifischen Anforderungen notwendig sind [Se15].

2.2 Planning Domain Definition Language (PDDL)

Die Planning Domain Definition Language (PDDL) [Ha19] ist eine standardisierte Sprache, die speziell für die Beschreibung von Planungsproblemen im Bereich der symbolischen KI entwickelt wurde. Sie strukturiert die Problemstellung, indem sie zwischen *Domänen (Domain)* und *Problemen (Problem)* unterscheidet: Die Domäne stellt das strukturelle Gerüst für Planungsaufgaben bereit, indem sie allgemeine Definitionen enthält, die über einzelne Planungsprobleme hinweg gültig sind. Ein *Problem* hingegen konkretisiert die Planungsaufgabe innerhalb einer bestimmten Domäne. Es legt konkrete Objekte fest, beschreibt den Anfangszustand und definiert den gewünschten Zielzustand. Dadurch wird die Domäne auf eine spezifische Situation angewendet, die gelöst werden soll.

Die Kernelemente einer Domänenbeschreibung umfassen u.a. Objekttypen, Prädikate und Aktionen. *Typen (Type)* dienen zur strukturierten Organisation von Objekten in der Planungsdomäne und helfen bei der klaren Zuordnung von Eigenschaften und Aktionen zu spezifischen Objektklassen. *Aktionen (Action)* sind zentrale Bestandteile in PDDL und beschreiben mögliche Handlungen innerhalb der Planungsdomäne. Jede Aktion definiert spezifische *Parameter*, die variabel sind und die beteiligten Objekte spezifizieren, sowie *Vorbedingungen*, die erfüllt sein müssen, bevor die Aktion ausgeführt werden kann. Die *Effekte* beschreiben die Veränderungen im Zustand der Planungsdomäne nach der Ausführung der Aktion. Dies ermöglicht es, aus einer Reihe von Aktionen Pläne zu erstellen, die auf das Erreichen eines gewünschten Zielzustands ausgerichtet sind. *Prädikate (Predicate)* in PDDL dienen dazu, spezifische Bedingungen oder Zustände in der Planungsdomäne zu beschreiben. Sie sind wesentliche Elemente zur Formulierung von Vorbedingungen und Effekten von Aktionen. Ein Prädikat kann als eine Art Aussage verstanden werden, die entweder wahr oder falsch ist und sich auf die Objekte oder Zustände bezieht, die durch die Aktionen beeinflusst werden. Durch die Verwendung von Prädikaten können

Planungssysteme logische Zusammenhänge innerhalb der Domäne überprüfen und entscheiden, ob bestimmte Aktionen unter den gegebenen Umständen zulässig und wirksam sind. *Funktionen (Function)* speichern numerische Informationen, die auch quantitative Zustände wie Kosten oder Distanzen darstellen können. Diese werden genutzt, um quantitative Vorbedingungen für Aktionen zu definieren und deren Effekte quantitativ zu bewerten, was die Grundlage für die Optimierung von Plänen nach spezifischen Kriterien schafft.³

2.3 Verwandte Arbeiten

In [HVC13] wird eine SysML-Taxonomie für Montageaufgaben vorgestellt, die zur Beschreibung von Systemfähigkeiten herangezogen wird. Diese Systemfähigkeiten bilden die Basis für die Ableitung von PDDL-Aktionen. Der Beitrag stellt zwar eine PDDL-Beschreibung zur Steuerung eines Robotiksystems in Montageumgebungen vor, jedoch erfolgt die Erstellung der PDDL-Beschreibung manuell und ist auf die Taxonomie beschränkt.

Viera et al. präsentieren in [Vi23] einen fähigkeitsbasierten Ansatz zur Generierung von PDDL-Beschreibungen. Auf Basis einer Ontologie werden benötigte Fähigkeiten zur Herstellung von Produkten und angebotene Fähigkeiten von Produktionsressourcen zugeordnet, woraus dann PDDL-Beschreibungen automatisch generiert und verarbeitet werden. Dieser Ansatz erfordert jedoch ein spezifisches Fähigkeitsmodell und beruht nicht auf einem Systemmodell aus der MBSE.

In [Ri21] wird ein Ansatz vorgestellt, um Planungsprobleme in Hierarchical Domain Definition Language (HDDL), einem Dialekt von PDDL, zu vereinfachen. Dafür vergleichen die Autoren zunächst HDDL und SysML Elemente und leiten anschließend einen Workflow für die Modellierung von HDDL Modellen ab. Die Beschreibung erfolgt jedoch konzeptionell und eine Transformation der Elemente muss manuell durchgeführt werden.

Der Ansatz von Batarseh und McGinnis [BM12] nutzt ein SysML-Profil zur automatisierten Generierung von Simulationsmodellen für Arena, basierend auf SysML-Systemmodellen. Dieser Ansatz zielt darauf ab, die Modellierungskomplexität zu reduzieren und die Effizienz zu steigern. Planungsprobleme werden jedoch nicht adressiert.

In [Wa19] wird ein Ansatz beschrieben, der das Modell eines Fertigungssystems automatisiert in einen Produktionsplan umwandelt. Dieser Ansatz nutzt dafür die ISA-95 und generiert daraus automatisiert PDDL-Beschreibungen. Daher ist dieser Ansatz speziell auf Fertigungssysteme ausgelegt, die nach der ISA-95 modelliert sind. Daher ist sowohl eine Anwendbarkeit auf Systeme, die nicht nach dieser Norm modelliert sind, sowie eine Integration in SysML-basierte Systemmodelle nicht möglich.

Die Modellierung von PDDL-Domänenbeschreibungen wird häufig als besonders herausfordernd, zeitaufwändig und fehleranfällig betrachtet [Li23]. Dennoch bietet keiner der

³ Eine vollständige Erläuterung der Planungssprache kann [Ha19] entnommen werden.

bestehenden Ansätze eine direkte Integration von Planungslogiken in SysML-basierte Systemmodelle, die in der Regel detaillierte Informationen über das System bereitstellen und somit eine wertvolle Basis für PDDL-Domänenbeschreibungen liefern können. Daher zeigt dieser Beitrag, wie entsprechende Elemente aus dem Systemmodell mittels eines SysML-Profiles für PDDL mit planungsrelevanten Informationen annotiert werden können, um eine automatisierte Generierung von PDDL-Domänen zu ermöglichen.

3 SysML-Profil für PDDL

Die Integration von PDDL in das SysML-Umfeld erfordert die Entwicklung eines dedizierten SysML-Profiles, das die spezifischen Konstrukte und Zusammenhänge von PDDL abbildet. Als Grundlage für dieses Profil dient die Backus-Naur-Form (BNF)-Definition von PDDL 3.1 [Ko11], durch die die Kompatibilität mit dem PDDL-Standard sichergestellt werden kann. Wie in [BW13] dargestellt wird, ist die automatisierte Überführung der BNF-Definition in ein Metamodell, wie beispielsweise das Profil, jedoch nicht ohne Weiteres möglich, da die BNF-Definition nicht ausreichend ist, um die syntaktischen und semantischen Unterschiede zwischen den beiden Modellierungssprachen zu überbrücken. Daher erfolgte die Entwicklung des Profils manuell, stets unter Einhaltung der BNF-Definition von PDDL 3.1, um die Kernkonstrukte von PDDL nach SysML zu übertragen.

Besonders für die Beschreibung der PDDL-Domäne liefert ein Systemmodell die erforderlichen Informationen. Die Daten für die Problembeschreibung stammen üblicherweise aus einem Produktmodell, das nicht direkt Teil des Systemmodells ist. Das Profil kann sowohl die Domänen- als auch die Problembeschreibung darstellen, vorausgesetzt, das Produktmodell wird in die Systemmodellierungsumgebung integriert. Dieser Beitrag fokussiert auf die Darstellung der Domänenbeschreibung.

Die zentralen Konstrukte der Domänenbeschreibung umfassen Aktionen, Prädikate, Funktionen und Typen, die als Grundlage für die Planungsaufgaben in PDDL dienen. Jedes dieser Elemente wurde im Rahmen des SysML-Profiles als Stereotyp modelliert. Stereotypen erlauben die Erweiterung bestehender UML-/SysML-Metamodellelemente bzw. UML-/SysML-Metaklassen um zusätzliche Eigenschaften, die für PDDL-spezifische Zusammenhänge notwendig sind. Ein Auszug dieser Stereotypen für die Beschreibung von PDDL-Domänen wird in Tabelle 1 aufgelistet und den entsprechenden UML-/SysML-Metaklassen zugeordnet.⁴

In PDDL definiert die **Domain** die Umgebung und die Regeln, innerhalb derer die Planung stattfindet. Dies umfasst spezifische Aktionen, Typen und Bedingungen, die in Planungsprozessen benötigt werden. Für die Modellierung in SysML wird das entsprechende Element durch den Stereotyp <<PDDL_Domain>> repräsentiert, welches die UML-Metaklassen Package oder Model erweitert. Indem der übergeordnete Ordner des Systemmodells als PDDL-Domain

⁴ Das vollständige Profil ist zur weiteren Einsicht und Nutzung verfügbar unter <https://github.com/hsu-aut/SysML-Profile-PDDL>.

Tab. 1: Zuordnung von PDDL-Elementen zu UML-/SysML-Metaklassen

PDDL-Element	PDDL-Stereotyp	UML-/SysML-Metaklassen
Domain	<<PDDL_Domain>>	Model, Package
Type	<<PDDL_Type>>	Class
Predicate	<<PDDL_Predicate>>	ObjectFlow, ControlFlow
Function	<<PDDL_Function>>	ObjectFlow, ControlFlow
Action	<<PDDL_Action>>	CallBehaviorAction

stereotypisiert wird, ist es möglich die PDDL-Domain aus dem Systemmodell abzuleiten, da die planungsspezifische Struktur und deren Logik sowie benötigte Attribute Teil des Systemmodells sind. Dafür werden sogenannte *derivedProperties* verwendet, welche auf die weiteren PDDL-Domain Elemente wie **Type** oder **Action** verweisen.

Ein **Type** in PDDL definiert eine Klasse von Objekten, ähnlich wie Datentypen in Programmiersprachen verwendet werden. Mit dem entsprechenden Stereotyp <<PDDL_Type>> wird die UML-Metaklasse `Class` erweitert, da diese genutzt wird, um strukturelle Aspekte des Systemmodells abzubilden, beispielsweise die Komponenten des Systems. Durch die Erweiterung der UML-Metaklasse `Class` können Objekteigenschaften modelliert sowie Typenhierarchien definiert werden. Solche Hierarchien erlauben es, Objekte in Über- und Untertypen zu strukturieren, wodurch Subtyp-Strukturen innerhalb des Systems abgebildet und verwaltet werden können.

Eine **Action** beschreibt in PDDL eine spezifische Aktion, die ausgeführt werden kann, um den Zustand des Planungssystems zu verändern. In Systemmodellen wird in der Regel ein Aktivitätsdiagramm verwendet, um das Verhalten des Systems zu beschreiben [Be24]. Daher erweitert der Stereotyp <<PDDL_Action>>, welcher in Abbildung 1 dargestellt ist, die UML-Metaklasse `CallBehaviorAction`, die in bestehenden Systemmodellen zur Abbildung von Prozessen bzw. Systemfunktionen genutzt wird. Jede Aktion in PDDL definiert außerdem die notwendigen Vorbedingungen, die erfüllt sein müssen, bevor die Aktion ausgeführt werden kann, den Effekt, welcher aus der Aktion resultiert, sowie Parameter in Form von typisierten Variablen, welche in den Vorbedingungen und Effekten verwendet werden.

Im Profil werden diese Zusammenhänge durch die Erweiterung des Stereotyps mit Hilfe eines `Customization` Elements umgesetzt. Innerhalb des `Customization` Elements ist es möglich `derivedProperty` hinzuzufügen. Ein `derivedProperty` ist eine Eigenschaft, dessen Wert durch seine Beziehung zu anderen Elementen definiert ist. Beispielsweise verweist das `derivedProperty` `Precondition` auf alle `Object`- und `ControlFlows` mit den Stereotypen <<PDDL_Predicate>> oder <<PDDL_Function>>, welche in ein Element mit dem Stereotypen <<PDDL_Action>> eingehen, und das `derivedProperty` <<PDDL_Effect>> auf alle `Object`- und `Controlflows`, die aus dem <<PDDL_Action>> Element herausgehen. Auf diese Weise können die einzelnen Elemente des PDDL-Profiles miteinander verknüpft und in Beziehung gesetzt werden.

Es ist ebenso möglich, die Nutzung von Stereotypen einzuschränken. So wird bspw. die `possibleOwner` Eigenschaft des `Customization` Elements verwendet, um die Nutzung des `<<PDDL_Action>>` Stereotyps einzuschränken. In diesem Fall darf der Stereotyp nur innerhalb der `<<PDDL_Domain>>`, `Packages` sowie `Activities` genutzt werden. Dies dient als Richtlinie für den Anwender und verhindert die Nutzung des Elements an einer falschen Stelle.

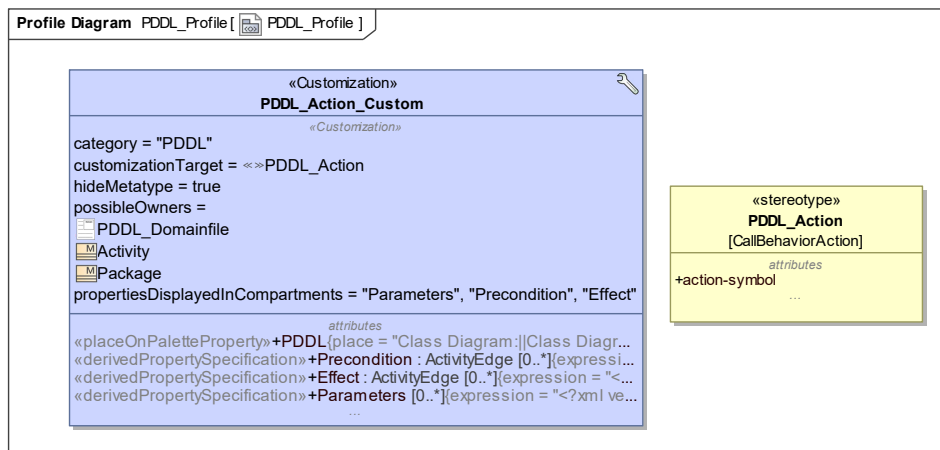


Abb. 1: `<<PDDL_Action>>` inklusive dem Customization-Element

In PDDL wird ein **Predicate** verwendet, um Bedingungen oder Zustände innerhalb einer Planungsdomäne zu beschreiben, die entweder zutreffen oder nicht zutreffen. Diese Prädikate werden für die Definition der Vorbedingungen und Effekte von Aktionen genutzt. Der entsprechende Stereotyp `<<PDDL_Predicate>>` erweitert die UML-Metaklassen `ObjectFlow` und `ControlFlow`. Sie werden genutzt, um Prädikate als Vorbedingungen oder Effekte für PDDL-Aktionen zu modellieren.

Im Gegensatz dazu repräsentiert eine **Function** in PDDL einen numerischen Wert. Dieses Element wird verwendet, um eine Vorbedingung oder einen Effekt einer Aktion abbilden zu können, entsprechend erweitert der `<<PDDL_Function>>` dieselben Metaklassen wie der Stereotyp `<<PDDL_Predicate>>`.

Das vorgestellte Profil ermöglicht eine PDDL-spezifische Erweiterung bestehender SysML-Systemmodelle. Durch die Verwendung von `derivedProperties`, um die Zusammenhänge der einzelnen PDDL Elemente dynamisch zu erfassen, führen Änderungen an das Systemmodell, sofern sie mit dem PDDL-Profil annotiert sind, automatisch auch zu Änderungen der entsprechenden PDDL-Elemente. Dies sichert die Konsistenz zwischen der Systementwicklung und der Planungslogik.

Das so angereicherte Systemmodell bildet die Grundlage für die automatisierte Generierung von PDDL-Beschreibungen. Mit diesem angereicherten Systemmodell können die relevan-

ten Planungsinformationen extrahiert und einem Algorithmus zur Generierung übergeben werden. Die Beschreibung der Methode zur systematischen Anreicherung von Systemmodellen sowie ein Algorithmus zur Generierung von PDDL-Beschreibungen basierend auf den angereicherten Systemmodellen werden in separaten Veröffentlichungen beschrieben.

4 Anwendungsfall im Flugzeugbau

Zur Validierung des entwickelten Profils wurde ein Anwendungsfall aus der Strukturmontage im Flugzeugbau herangezogen, bei dem ein verfahrbarer UR10-Roboterarm innerhalb des Flugzeugrumpfs positioniert ist. Dieses Robotik-System wird eingesetzt, um verschiedene Kollare auf Nieten zu drehen, wobei jeder Kollartyp einen spezifischen Endeffektor benötigt.⁵

Im Rahmen des Anwendungsfalles sollte ein Prozessablaufplan zur Optimierung der Durchlaufzeit identifiziert werden. Dabei war zu berücksichtigen, dass eine Umkonfiguration der Hardware erforderlich ist, sobald ein anderer Endeffektor für das Aufdrehen der Kollare benötigt wird.

In Abbildung 2 wird ein Ausschnitt des Aktivitätsdiagramms aus dem angereicherten Systemmodell gezeigt, der das Zusammenspiel zwischen dem modellierten Systemverhalten und dem PDDL-Profil veranschaulicht.

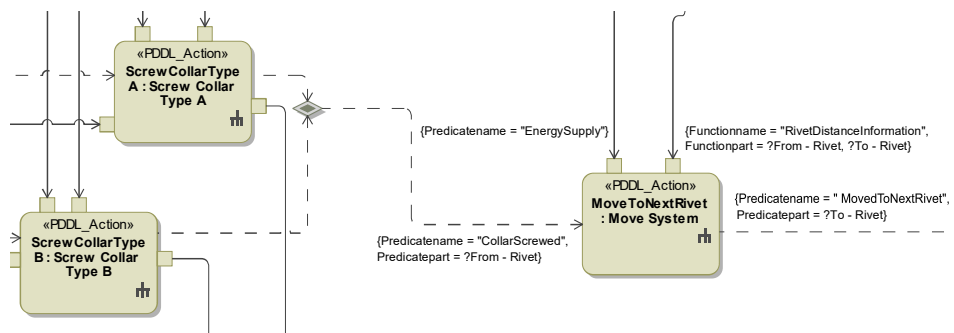


Abb. 2: Ausschnitt aus dem angereicherten Systemmodell

Der Ausschnitt zeigt drei «PDDL_Action»-Elemente sowie die dazugehörigen Vorbedingungen und Effekte. Die «PDDL_Action»-Elemente „ScrewCollarType A“ bzw. „ScrewCollarType B“ beschreiben den Prozessschritt für das Aufdrehen der Kollare. Sobald dieser Prozessschritt abgeschlossen ist, bewegt sich das System in der «PDDL_Action» „MoveToNextRivet“ zum nächsten Niet. Diese Schritte werden zyklisch vielfach durchlaufen.

Um die Aktion „MoveToNextRivet“ auszuführen, die den Endeffektor von einem Niet zum nächsten führt, fließen folgende Vorbedingungen in die Aktion ein:

⁵ Eine Vorgehensweise zur Erstellung des zugehörigen Systemmodells ist in [Be23] dargestellt.

1. Das System muss mit Energie versorgt sein. Dafür fließt das Prädikat „EnergySupply“ in die Aktion hinein.
2. Der Kollar, an dem sich der Endeffektor aktuell befindet, muss bereits verschraubt sein. Dafür sind die Aktionen „ScrewCollarTypeA“ bzw. „ScrewCollarTypeB“ zuständig, die das Prädikat „CollarScrewed“ auf „True“ setzen, sobald die Aktion abgeschlossen ist.
3. Die Koordinaten der Niete und die damit verbundene Distanz zwischen den Nieten ist in der Function „RivetDistanceInformation“ definiert und wird von dort aus an die MoveToNextRivet <<PDDL_Action>> übergeben. Ein PDDL-Solver kann auf der Basis der Distanzen und Niettypen den Niet identifizieren, welcher als nächstes bearbeitet werden sollte.

Sobald die Aktion durchgeführt wurde, hat sie als Effekt, dass der Endeffektor sich auf einer neuen Arbeitsposition befindet. Das entsprechende Prädikat „MovedToNextRivet“ wird mit der neuen Nietposition auf „True“ gesetzt.

Als Modellierungssoftware wurde im Anwendungsfall Magic Systems of Systems Architect (MSoSA) für die Erstellung des Systemmodells sowie für die Erstellung des Profils verwendet. Das Systemmodell wurde anschließend innerhalb dieser Software mit dem Profil angereichert. Durch die Anreicherung des Systemmodells kann es systematisch nach Planungsinformationen abgefragt werden. Zur automatisierten Generierung der Planungsbeschreibungen kann beispielsweise der *Report Wizard* in MSoSA verwendet werden, der auf die Velocity Engine aufbaut. Dafür werden Templates in der Velocity Template Language (VTL) benötigt. Im Rahmen des Anwendungsfalls wurden entsprechende VTL-Templates für die PDDL-Domain und PDDL-Problem geschrieben. Diese Templates beinhalten die syntaktische Struktur von PDDL mit entsprechenden Platzhaltern, die durch die Engine gefüllt werden. In List. 1 wird ein Ausschnitt der damit erzeugten Domänenbeschreibung gezeigt, der die zuvor beschriebene Aktion aus Abbildung 2 in PDDL-Syntax zeigt.

```
(:action MoveToNextRivet
  :parameters (?From - Rivet ?To - Rivet)
  :precondition (and
    (CollarScrewed ?From)
    (EnergySupply)
  )
  :effect (and
    (MovedToNextRivet ?To )
    (not (CollarScrewed ?From ))
    (increase (total-cost) (RivetDistanceInformation ?From ?To))
  )
)
```

List. 1: PDDL-Syntax für die Aktion "MoveToNextRivet"

Anschließend kann diese standardisierte Beschreibung an einen Solver zur Lösung des Planungsproblems übergeben werden. Für den Anwendungsfall kam beispielsweise der Solver *Delfi* [Ka18] zum Einsatz, der von [Mu22] bereitgestellt wird. Mit dem resultierenden Plan des Solvers kann das System abschließend hinsichtlich Optimierungszielen wie der Minimierung der Durchlaufzeit oder dem erwarteten Energieverbrauch untersucht werden. Die daraus gewonnenen Daten dienen dem Systementwickler als Basis für Entscheidungen in der Systemauslegung. Sollte beispielsweise statt eines UR10 ein UR16 in Betracht gezogen werden, lässt sich ein direkter Vergleich dieser beiden Varianten durchführen. Dafür muss das Systemmodell mit den Informationen für den UR16 erweitert werden, jedoch kann die gleiche Planungslogik ohne Mehraufwand wiederverwendet werden.

5 Zusammenfassung und Ausblick

Im Rahmen dieses Beitrags wurde ein SysML-Profil vorgestellt, das speziell darauf ausgerichtet ist, die Modellierung und Planung komplexer Produktionssysteme durch Integration von Planungslogiken mittels PDDL zu verbessern. Das entwickelte Profil ermöglicht eine PDDL-spezifische Erweiterung bestehender SysML-basierter Systemmodelle, was zu einer konsistenten und synchronisierten Aktualisierung von Planungs- und Systeminformationen führt. Das heißt, dass Änderungen im Systemmodell automatisch zu Änderungen in den PDDL-Komponenten führen. Die Anwendbarkeit des Profils wurde in einem realen Anwendungsfall im Flugzeugbau erfolgreich validiert.

Die Nutzung dieses Profils setzt jedoch eine Vorgehensmethode voraus, die beschreibt, welche Schritte notwendig sind um Systemmodelle systematisch mit diesem Profil zu erweitern. Zudem ist ein Algorithmus erforderlich, der die angereicherten Systemmodelldaten liest und in PDDL-Syntax umwandelt. Diese automatisierte Generierung von PDDL-Beschreibungen wurde bereits experimentell mit VTL-Templates implementiert. Sie beschränkt sich jedoch auf die Verwendung in der Modellierungssoftware MSoSA. Da Informationen für die PDDL-Problembeschreibung in der Regel nicht direkt im Systemmodell enthalten sind, sondern sich oft in einem Produktmodell befinden, empfiehlt sich ein Ansatz, der diese Informationen aus verschiedenen Engineering-Artefakten extrahieren kann. Insbesondere weil die Informationen aus der Domänen- und Problembeschreibung in Beziehung gesetzt werden müssen, werden zukünftige Arbeiten darauf abzielen, diese algorithmische Generierung zu standardisieren und flexible Transformationsprogramme zu entwickeln, die beispielsweise auf den Ansatz von [WF22] basieren, um weitere Datenquellen berücksichtigen zu können. Sowohl die Vorgehensmethode als auch der Algorithmus sind Gegenstand zukünftiger Veröffentlichungen.

Förderhinweis

Die Autoren [Projekt iMOD] bedanken sich für die Förderung bei dtec.bw – Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr. dtec.bw wird von der Europäischen Union – NextGenerationEU finanziert.

Literatur

- [Be23] Beers, L.; Weigand, M.; Nabizada, H.; Fay, A.: MBSE Modeling Workflow for the Development of Automated Aircraft Production Systems. In: 28th International Conference on Emerging Technologies and Factory Automation (ETFA). S. 1–8, 2023.
- [Be24] Beers, L.; Nabizada, H.; Weigand, M.; Gehlhoff, F.; Fay, A.: A SysML Profile for the Standardized Description of Processes during System Development. In: International Systems Conference (SysCon). Im Erscheinen, 2024.
- [BM12] Batarseh, O.; McGinnis, L. F.: System modeling in SysML and system analysis in Arena. In: Proceedings of the Winter Simulation Conference (WSC). S. 1–12, 2012.
- [BW13] Bergmayr, A.; Wimmer, M.: Generating Metamodels from Grammars by Chaining Translational and By-Example Techniques. In: MDEBE@ MoDELS. S. 22–31, 2013.
- [Ei15] Eigner, M.; Dickopf, T.; Schulte, T.; Schneider, M.: mecPro2-Entwurf einer Beschreibungssystematik zur Entwicklung cybertronischer Systeme mit SysML. Tag des Systems Engineering (Hanser)/, S. 163–172, 2015.
- [Es07] Estefan, J. A.: Survey of Model-based Systems Engineering (MBSE) Methodologies. IncoSE MBSE Focus Group 25/8, S. 1–12, 2007.
- [GNT16] Ghallab, M.; Nau, D.; Traverso, P.: Automated Planning and Acting. Cambridge University Press, 2016.
- [Ha19] Haslum, P.; Lipovetzky, N.; Magazzeni, D.; Muise, C.: An Introduction to the Planning Domain Definition Language. Springer International Publishing, Cham, 2019.
- [HS21] Henderson, K.; Salado, A.: Value and benefits of model-based systems engineering (MBSE): Evidence from the literature. Systems Engineering 24/1, S. 51–66, 2021.
- [HVC13] Huckaby, J.; Vassos, S.; Christensen, H. I.: Planning with a task modeling framework in manufacturing robotics. In: International Conference on Intelligent Robots and Systems. S. 5787–5794, 2013.
- [Ka18] Katz, M.; Sohrabi, S.; Samulowitz, H.; Sievers, S.: Delfi: Online Planner Selection for Cost-Optimal Planning. In: Ninth International Planning Competition. S. 57–64, 2018.
- [Ko11] Kovacs, D. L.: Complete BNF description of PDDL 3.1. Language Specification, Department of Measurement and Information Systems, Budapest University of Technology and Economics/, 2011.
- [Li23] Lindsay, A.: On Using Action Inheritance and Modularity in PDDL Domain Modelling. Proceedings of the International Conference on Automated Planning and Scheduling 33/1, S. 259–267, 2023.

- [Ma22a] Ma, J.; Wang, G.; Lu, J.; Vangheluwe, H.; Kiritsis, D.; Yan, Y.: Systematic Literature Review of MBSE Tool-Chains. *Applied Sciences* 12/7, S. 3431, 2022.
- [Ma22b] Mayr-Dorn, C.; Egyed, A.; Winterer, M.; Salomon, C.; Fürschuß, H.: Evaluating PDDL for programming production cells: a case study. In: 4th International Workshop on Robotics Software Engineering (RoSE). S. 17–24, 2022.
- [Mu22] Muise, C.; Pommerening, F.; Seipp, J.; Katz, M.: PLANUTILS: Bringing Planning to the Masses. In: 32nd International Conference on Automated Planning and Scheduling (ICAPS 2022). System Demonstrations. 2022.
- [OM19] OMG: Systems Modeling Language (SysML™ 1.6), 2019.
- [Ri21] Rimani, J.; Lesire, C.; Lizy-Destrez, S.; Viola, N.: Application of MBSE to model Hierarchical AI Planning problems in HDDL. In: International Conference on Automated Planning and Scheduling (ICAPS). 2021.
- [Se15] Seidl, M.; Scholz, M.; Huemer, C.; Kappel, G.: *UML @ Classroom: An Introduction to Object-Oriented Modeling*. Springer, 2015.
- [SS20] Schmidt, M. M.; Stark, R.: Model-Based Systems Engineering (MBSE) as computer-supported approach for cooperative systems development. In: Proceedings of 18th European Conference on Computer-Supported Cooperative Work. 2020.
- [Tö17] Törmänen, M.; Hägglund, A.; Rocha, T.; Drenth, E.: Integrating Multi-Disciplinary Optimization into the Product Development Process using Model-Based Systems Engineering (MBSE). In: NAFEMS World Congress. 2017.
- [Vi23] Vieira da Silva, L. M.; Heesch, R.; Köcher, A.; Fay, A.: Transformation eines Fähigkeitsmodells in einen PDDL-Planungsansatz. *at-Automatisierungstechnik* 71/2, S. 105–115, 2023.
- [Wa19] Wally, B.; Vyskočil, J.; Novák, P.; Huemer, C.; Šindelár, R.; Kadera, P.; Mazak, A.; Wimmer, M.: Flexible production systems: Automated generation of operations plans based on ISA-95 and PDDL. *IEEE Robotics and Automation Letters* 4/4, S. 4062–4069, 2019.
- [We08] Weikiens, T.: *Systems engineering with SysML/UML: Modeling, analysis, design*. Morgan Kaufmann/Elsevier, Amsterdam und Boston, 2008.
- [We16] Weikiens, T.: *SYSMOD - The Systems Modeling Toolbox: Pragmatic MBSE with SysML*. Lulu.com, 2016.
- [WF22] Weigand, M.; Fay, A.: Creating Virtual Knowledge Graphs from Software-Internal Data. In: 48th Annual Conference of the IEEE Industrial Electronics Society. S. 1–6, 2022.

„Ulrich Jumar, Christian Diedrich (Hrsg.):
EKA 2024 - Entwurf komplexer Automatisierungssysteme, 18. Fachtagung“

SIM-CIP: Konzept eines räumlichen Informationsmodells für komplexe industrielle Großanlagen

Alexander Auer^{1,2}, Birgit Vogel-Heuser³, Tobias Knödler⁴ und Dorothea Pantförder⁵

Abstract: Die richtige Information zur richtigen Zeit am richtigen Ort – dieser Leitgedanke erhält im Rahmen der Industrie 5.0 eine zunehmende Bedeutung, insbesondere für den grundlegenden Wert der Menschzentrierung. Komplexe industrielle Anlagen, wie die in diesem Beitrag behandelte Wärmepumpenanlage, enthalten zahlreiche und vielfältige technische Informationen. Ein Instandhaltungstechniker, der auf Alarmsignale reagiert und ortsbezogene Maßnahmen innerhalb einer Anlage ergreifen muss, steht vor der Herausforderung in diesem komplexen Umfeld zielgerichtet an nötige Informationen zu gelangen, um anfallende Kosten zu minimieren. Standortbezogene Dienste (LBS) bieten hierfür potenzielle Lösungen, allerdings ist die Anwendung dieser im industriellen Bereich defizitär, da bestehende Systeme und Forschungsarbeiten sich vorwiegend auf Verbraucheranwendungen fokussieren. Dieser Beitrag legt den Schwerpunkt auf die industrielle Anwendungslösung mittels LBSs. Hierzu wird am Beispiel einer komplexen Wärmepumpenanlage das Konzept eines räumlichen Informationsmodells (SIM-CIP) für industrielle LBSs verdeutlicht. SIM-CIP integriert räumliche Informationen mit attributiven Informationen und geht zusätzlich über die rein ortsbezogene Erweiterung hinaus. So werden mögliche kontextabhängige Faktoren, die einen industriellen Arbeiter betreffen, berücksichtigt. Das konzeptionelle Modell umfasst fünf grundlegende Systeme – ein Datenbanksystem, ein Kartierungssystem, ein Positionierungssystem, ein Kontextmanagementsystem und eine Benutzerschnittstelle – die im Rahmen dieses Beitrages näher erläutert werden.

Keywords: Spatial Information System, Context-Aware Location Based Service, Human Machine Interface, Human-Centered Engineering

1 Einleitung und Motivation

Die Bereitstellung maßgeschneiderter Informationen zum optimalen Zeitpunkt am geeigneten Ort gewinnt angesichts der Industrie 5.0 [BDP21] und der darin fokussierten Menschzentrierung, zunehmend an Bedeutung. In diesem Kontext bedeutet das eine adaptive Anpassung der eingesetzten technologischen Lösungen an die individuellen Bedürfnisse eines Arbeiters [Xu21]. Die für den Menschen erleichternde Bereitstellung gezielter ortsbezogener Informationen ist Teil eines solchen Bedürfnisses. Komplexe

¹ TUM, Lehrstuhl AIS, Boltzmannstr. 15, 85748 Garching, alexander.auer@tum.de

² MAN Energy Solutions SE, Multimedia Sol., Stadtbachstr. 1, 86153 Augsburg, alexander.auer@man-es.com

³ TUM, Lehrstuhl AIS, Boltzmannstr. 15, 85748 Garching, vogel-heuser@tum.de

⁴ MAN Energy Solutions SE, Multimedia Sol., Stadtbachstr. 1, 86153 Augsburg, tobias.knoedler@man-es.com

⁵ TUM, Lehrstuhl AIS, Boltzmannstr. 15, 85748 Garching, pantfoerder@tum.de

industrielle Großanlagen erstrecken sich über eine Fläche von mehreren tausend Quadratmetern, häufig verteilt auf unterschiedliche Ebenen. Die in diesen Anlagen auftretenden Informationen sind vielfältig und zahlreich. Hierzu gehören umfassende Montagehandbücher und Betriebsanleitungen, detaillierte Pläne und Schemata, Echtzeit-Prozessdaten und Warnsignale, Sicherheitsbereiche, sowie dokumentierende Kontrollberichte. Auch die steigende Komplexität der Prozesse in Form von verstärkter Automatisierung oder fortschreitender Systemintegration leisten dazu ihren Beitrag. Diese Informationsmenge und -vielfalt ist somit für einen Menschen ohne fortschrittliche technologische Unterstützung nicht effizient zugänglich. So steht beispielsweise ein Instandhaltungstechniker, der auf Alarmsignale oder Fehlermeldungen reagiert, für eine Aufgabenlösung vor der Herausforderung in diesem komplexen Umfeld effizient und zielgerichtet relevante Informationen aus umfangreichen Dokumentationen und Datenbanksystemen zu erhalten. Dies betrifft sowohl vielschichtige aufwendige Maßnahmen als auch Routinetätigkeiten. Ohne Weiterentwicklung der verwendeten Schnittstellentechnologie wird es für ihn zunehmend aufwendiger an notwendige Informationen zu gelangen oder relevante zu beachten. Für die Minimierung anfallender Kosten ist ein direkter Zugriff auf nutzbringende Informationen ebenfalls beachtenswert.

Neben einer tätigkeitsspezifischen Informationsfilterung ist dafür eine Filterung bezogen auf die örtliche Position ein erstrebenswerter Lösungsansatz. So können beispielsweise einem Techniker im Feldeinsatz notwendige Pläne oder Daten für Zulieferteile direkt bei örtlicher Nähe an einem Bauteil oder anderen relevanten Instandhaltungs- und Wartungsorten gefiltert bereitgestellt werden. Dies erlaubt eine unmittelbare positionsbezogene Zugänglichkeit zu Informationen, ohne dass seitens eines Technikers eine aufwendige manuelle Selektion durchgeführt werden muss. Für mobile Anlagenteile, wie mobile Roboter, kann eine solche Art der Informationsbereitstellung ebenfalls nutzbringend sein. Zudem lassen sich informative Sicherheitsmechanismen für sicherheitsrelevante Bereiche zum Schutz eines Technikers oder der Integrität einer Anlage einrichten. Weitere Benutzergruppen können ebenfalls von einer ortsbezogenen Filterung von Informationen profitieren. Beispielsweise kann ein solcher Filter bereits bei Plänen für die Erbauung einer Anlage eingesetzt werden, um Bauarbeiter, respektive einen Bauleiter bei Informationsverteilungen zu unterstützen. Auch hinsichtlich modernster Assistenzsysteme mittels erweiterter Realität (Extended Reality, XR) ist ein ortsbezogener Filter attraktiv, um eine Koordinierung dieser Anwendungen zu realisieren.

Für eine entsprechende Umsetzung muss ein zugrundeliegendes Modell verschiedene Anforderungen erfüllen. Hierfür wurden im Rahmen dieses Beitrages vier spezifische Kategorien ausgewählt: Standort, Datenschema, Kontext und Vermittlung. Die wesentlichen Anforderungen an ein Modell innerhalb dieser Kategorien sind wie folgt:

- **Standort (Req-A):** Ortsbezogene Informationsbereitstellung basierend auf der Position eines Nutzers mit Anpassbarkeit an variable Anlagengröße
- **Datenschema (Req-B):** Struktur, die eine Repräsentation, Ergänzung, Integration und Aktualisierung von Informationen ermöglicht

-
- **Kontext (Req-C):** Informationen stehen einem Nutzer basierend auf der Relevanz für seine Rolle oder Aufgabe zur Verfügung und sind nach dessen individuellen Bedürfnissen adaptiv und personalisierbar
 - **Vermittlung (Req-D):** Informationen werden für einen Nutzer assistierend zur Aufgabenlösung bereitgestellt

Unter der Berücksichtigung dieser Anforderungen wird innerhalb dieses Beitrages ein Konzept für ein räumliches Informationsmodell im Kontext komplexer industrieller Großanlagen (Spatial Information Model for Complex Industrial Plants, SIM-CIP) vorgestellt, um standortbezogene Dienste zu ermöglichen.

2 Standortbezogene Dienste

Standortbezogene Dienste (Location Based Service, LBS) sind Anwendungen, die den Standort eines Nutzers verwenden, um darauf basierende digitale Dienste bereitzustellen [We14]. Typischerweise werden diese Anwendungen auf portablen Endgeräten wie Tablets, Handys und XR-Brillen genutzt. Die Anwendungsfelder reichen von Marketing oder Unterhaltung bis hin zu Navigation oder Bestandsverwaltung [Fa18]. Mit steigender Digitalisierung haben sich LBSs zu einem bedeutenden Forschungsfeld entwickelt. Vor allem Aspekte der Positionierung, Kartierung und Kontextabhängigkeit werden darin thematisiert.

Für die Positionierung von LBSs können verschiedene Technologien eingesetzt werden [YR15, Fa22]. Im freien Gelände ist das Global Positioning System (GPS) als Standard etabliert. Für eine Innenraumpositionierung (IPS) ist es größtenteils notwendig gesonderte Funktechnologie mit eigener Infrastruktur oder rechenaufwändige computergestützte Bildverarbeitung (Computer Vision, CV) einzusetzen. Funksysteme für eine IPS sind beispielsweise WiFi, Bluetooth Low Energy (BLE), Ultra-Wideband (UWB) und Radio-Frequency Identification (RFID). Überwiegend werden dabei RSSI-Signale (Received Signal Strength Indicator, RSSI) ausgewertet. Beispiel für ein LBS mittels Funktechnologie im Innenraum ist das System von Estimote [Es12]. Bei einer erforderlichen, möglichst exakten Bestimmung der Position eines Nutzergerätes (<1cm Abweichung), ist es nötig Lösungen aus der computergestützten Bildverarbeitung zu verwenden. Dabei werden visuelle Merkmale der Umgebung erfasst und interpretiert. Bei diesem Genauigkeitsanspruch ist diese Technologie sowohl für die Innenraumpositionierung als auch für die Positionierung im freien Gelände erforderlich. Zusätzlich können, falls vorhanden, interne Messeinheiten (Internal Measurement Units, IMUs) eines Gerätes für einen zusätzlichen Genauigkeitsgewinn respektive die simultane Lokalisierung und Kartierung (Simultaneous Localization and Mapping, SLAM) verwendet werden. Je nach Anforderung an einen LBS sind unterschiedliche Arten der Positionsbestimmung [We14] erforderlich. Für einige Anwendungen ist eine Positionserkennung durch Annäherung an einen Funksignalgeber (z. B. Bluetooth-Beacon) ausreichend. Andere Anwendungen benötigen eine Positionsbestimmung in der

Ebene oder im Raum ohne Orientierung, die beispielsweise durch Methoden der Lateration oder Angulation auf Basis mehrerer Funksignale erfolgen kann [Zh22]. LBSs, die XR-Anwendungen zur Verfügung stellen, benötigen eine Positionsbestimmung mit räumlicher Orientierung. Dafür sind zusätzliche Sensoren wie Gyroskope oder Magnetometer nötig, oder es werden Verfahren aus der CV, wie etwa Stereovision, verwendet. Microsoft entwickelte ein auf Bildverarbeitung basierendes System mit räumlichen Anker (Azure Spatial Anchors, [MS22]), um LBSs für Extended Reality Anwendungen zu ermöglichen. Für die Kartierung von LBSs können unterschiedliche Systeme verwendet werden. LBS-Anwendungen, die sich auf Nutzung von 2D-Karten mit niedriger Auflösung beschränken, verwenden geographische Informationssysteme (GIS) wie beispielsweise Google Maps [CL20]. Bei LBS-Anwendungen auf Basis von 3D-Karten werden häufig hochgradig individualisierte Systeme genutzt [Li20]. Anhand dieser Karten erfolgt Geofencing, also eine Ausweisung von LBS-Zonen für Interessenspunkte, nach unterschiedlichen Prinzipien [We14]. Zum einen können Zonen nach symbolischen Koordinaten ausgewiesen werden, bei denen eine Ortsausweisung nach Benennungen erfolgt (z. B. Zentralplattform am Verdichter). Zum anderen können Zonen nach geometrischen Koordinaten definiert werden. Typischerweise erfolgt dies über ein Rasterkoordinatensystem, das am GPS-System orientiert ist. Bei beiden Prinzipien wird angenommen, dass sich ein Interessenspunkt mit hoher Wahrscheinlichkeit innerhalb einer definierten Zone befindet. Kontextabhängige LBSs berücksichtigen neben den Standort des Nutzers, relevante Faktoren wie Aktivität, Umgebung, Zeit und persönliche Präferenzen. Die Forschungsschwerpunkte in diesem Bereich liegen insbesondere auf der kontextuellen Erfassung, Darstellung, Schlussfolgerung und Anpassung [GDF18].

Bei Betrachtung der Forschungsliteratur wird deutlich, dass sich LBSs vorwiegend auf Verbraucheranwendungen wie Navigation, Tourismus, standortbasierte Werbung oder Geotagging konzentrieren, während Potenziale in industriellen Prozessen bisher untergeordnet betrachtet und ein Defizit in der Anwendung in industriellen Sektoren wie dem Maschinen- und Anlagenbau besteht [Fa18]. Begründen lässt sich das mit der herausfordernden Integration von LBSs im industriellen Umfeld. Neben einer fehlenden Nachfrage nach LBSs in industriellen Prozessen durch ein mangelndes Bewusstsein für die Technologie, ist eine direkte Übertragung von LBS-Verbraucheranwendungen auf Industrieanlagen kaum möglich. Der Anspruch an ein LBS im industriellen Kontext unterscheidet sich dafür zu deutlich. Während ein LBS basierend auf einer Karte von Google Maps für Verbraucheranwendungen ausreicht, sind in industriellen Anlagen präzisere und detailliertere Karten erforderlich. Zudem müssen teilweise Schnittstellen hinsichtlich Soft- und Hardware, für industrielle LBSs speziell entwickelt oder angepasst werden. Darüber hinaus gelten in industriellen Anlagen oft strengere Regeln und Vorschriften hinsichtlich Datenschutz und -sicherheit als im Verbrauchermfeld. Das in diesem Beitrag vorgestellte Konzept eines räumlichen Informationsmodells zielt darauf ab, verschiedene Komponenten eines kontextabhängigen LBSs für eine Umsetzung im industriellen Bereich zu identifizieren und vorzustellen. Dies soll dazu beitragen, die Anwendung eines solchen Systems im industriellen Umfeld hinsichtlich Industrie 5.0 und der damit verbundenen Menschzentrierung, zu verbessern.

3 Konzept eines räumlichen Informationsmodells am Beispiel einer komplexen Wärmepumpenanlage

Im Rahmen dieses Beitrages wird für das Einbringen von kontextabhängigen LBSs in komplexen industriellen Anlagen ein räumliches Informationsmodell vorgeschlagen. Dies soll eine Struktur für die Verknüpfung von räumlichen Informationen mit weiteren attributiven Informationen beschreiben, um somit eine Information um den realen Faktor Ort, an dem diese auftritt, respektive benötigt wird, zu erweitern. Im Folgenden wird zuerst ein Anwendungsfall an einer komplexen Wärmepumpenanlage beschrieben für den kontextabhängige LBSs eingesetzt werden sollen. Anschließend folgt die Beschreibung der Bestandteile des Modells in Bezug auf den Anwendungsfall.

3.1 Instandhaltungs- und Wartungsvorgänge an einer Wärmepumpenanlage

In dem in diesem Beitrag betrachteten Anwendungsfall einer komplexen Wärmepumpenanlage umfasst die Technische Dokumentation mehrere Tausend Seiten. Die Gesamtdokumentation ist in verschiedene Teildokumentationen untergliedert, wobei ein Großteil auf Dokumentationen für Zulieferteile entfällt, die wiederum Sektionen beinhalten. In Hinblick auf Instandhaltungs- oder Wartungstätigkeiten sind daraus jedoch, abhängig von der auszuübenden Aufgabe, lediglich spezifische Auszüge relevant. So muss eine Anlagenaufsicht im Falle eines Alarmsignals aus der Anlage, beispielsweise in Form eines Fehlercodes, die umfangreiche Technische Dokumentation durchsuchen und sich gegebenenfalls mit anderen Anlagenexperten abstimmen, um potenziell mögliche Ursachen des Fehlers zu ermitteln. Diese werden anschließend an einen Techniker weitergeleitet, der im Anlagenbereich nach der tatsächlichen Ursache des Fehlers sucht. Daraufhin muss der Techniker die notwendigen Wartungsdokumente in der Technischen Dokumentation finden, um eine Problembehebung nach Anleitung durchzuführen.

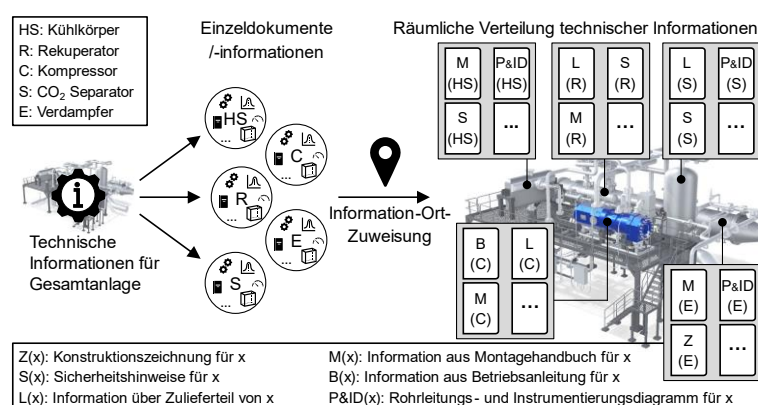


Abb. 3.1: Gezielte räumliche Bereitstellung technischer Informationen für die Instandhaltung am Beispiel einer komplexen Wärmepumpenanlage (Bildquelle: MAN ES SE)

Um die beschriebene Prozesskette effizienter zu gestalten, muss ein Techniker im Feld dabei unterstützt werden auf relevante Informationen schnell und zielgerichtet zuzugreifen. Kontextabhängige LBSs, also standortbezogene Informationsbereitstellungen (Abb. 3.1) mit nutzenbringenden Filtermöglichkeiten, können hier eine Möglichkeit der Unterstützung sein. Bezogen auf den beschriebenen Anwendungsfall können beispielsweise Wartungsdokumente direkt bei örtlicher Nähe und auf Basis der Rolle des Technikers an einem betreffenden Wartungsort gefiltert bereitgestellt werden. Ein Techniker soll damit ohne aufwendige manuelle Selektion die für ihn relevanten Informationen erhalten, zusätzliche sicherheitsrelevante Warnhinweise bekommen und gegebenenfalls durch Technologie, wie XR, assistiert werden.

3.2 Beschreibung des räumlichen Informationsmodells

Das Konzept des räumlichen Informationsmodells enthält die in Abb. 3.2 dargestellten Komponenten, um bei der hier behandelten Wärmepumpenanlage kontextabhängige LBSs zu realisieren. Darin enthalten ist ein Datenbanksystem, ein Positionierungssystem, ein Kartierungssystem, ein Kontextmanagementsystem und eine Benutzerschnittstelle. Es beschreibt somit den Weg vom Informationsbaustein bis zum Benutzer. Die spezifischen Merkmale der Einzelkomponenten werden im Folgenden erläutert.

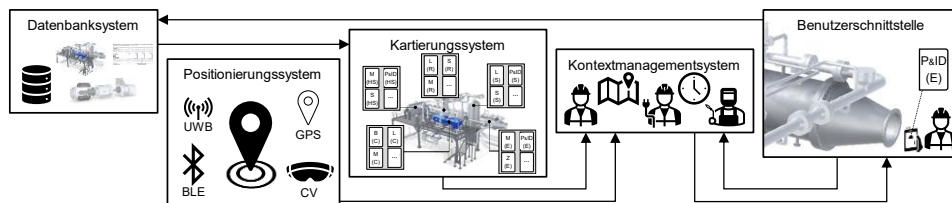


Abb. 3.2: Konzept des räumlichen Informationsmodells unterteilt in die grundlegenden Systeme

Datenbanksystem. Das Datenbanksystem enthält sämtliche Informationen, die für verschiedene Aufgaben innerhalb der Anlage nötig sind (Req-B). Im Falle der hier behandelten Wärmepumpenanlage zählen dazu beispielsweise Montagehandbücher, Betriebsanleitungen, Pläne, Echtzeit-Prozessdaten, Warnsignale, Sicherheitsbereiche und Kontrollberichte. Die enthaltenen Informationen innerhalb des Systems sind untergliedert in Informationsbausteine (ein Informationsbaustein ist im Rahmen dieses Beitrages eine einzelne, abgeschlossene Information, die einen grundlegenden modularen Bestandteil übergeordneter Informationsstrukturen, wie beispielsweise Wissensgraphen, darstellt). Informationsbausteine können in ihrem Umfang unterschiedlich ausgeprägt sein. Diese reichen von umfassenden PDF-Dokumenten über einzelne Kapitel, Textausschnitte oder Grafiken bis hin zu einzelnen Werten. Organisiert werden diese Bausteine beispielsweise anhand von Metadaten, Tags, semantischen Beschreibungen oder räumlichen Koordinaten. In dem hier beschriebenen Konzept bilden Informationsbausteine die Grundlage für das Kartierungs- und Kontextmanagementsystem.

Positionierungssystem. Das Positionierungssystem bestimmt die Position eines Nutzergerätes innerhalb der Anlage durch die Auswertung von im Folgenden beschriebenen Sensordaten (Req-A). Bei der hier betrachteten Wärmepumpenanlage sind deshalb für eine Positionsbestimmung verschiedene Bedingungen zu berücksichtigen. So spielt die Größe der Anlage eine entscheidende Rolle, ebenso wie dessen Belichtung und die darin auftretenden Reflexionen von metallischen Oberflächen. Interferenzen mit anderen Geräten sind ebenfalls möglich. Es handelt sich hierbei um eine statische Anlage, wodurch kaum Änderungen in der Anlagenumgebung zu erwarten sind. Außerdem existieren Bereiche innerhalb der Anlage von unterschiedlicher Relevanz. So werden Informationen beispielsweise am Kompressor vermehrt benötigt und eine exakte Informationseinbettung kann sinnvoll sein, wohingegen in Zwischengängen größtenteils eine rein zonenbezogene Informationsbereitstellung ausreicht. Im Anwendungsfall sind somit folgende Stufen der Ortung angedacht:

- **Stufe I – Annäherung an Funksignalegeber (Beacon):** Überwiegend für Warn- und Sicherheitshinweise oder Statusmeldungen bei Signalempfang, sowie für die mobile Informationsbereitstellung (mobile Beacons)
- **Stufe II – Lateration von Funksignalen (mehrere Beacons):** Bildet die Basis der ortsbezogenen Informationsbereitstellung, um Zonen für beispielsweise Dokumentenbereitstellungen auszuweisen und liefert die Koordinierungsgrundlage für die in Stufe III verwendete CV
- **Stufe III – CV mit SLAM (Kamera und IMU-Nutzung):** Für präzise örtliche Einbettung von Informationen mittels XR-Visualisierungen

Innerhalb dieses Anwendungsfalles ist primär der Einsatz von Stufe II vorgesehen. Es kann jedoch wahlweise das Prinzip aus Stufe I genutzt werden oder eine ergänzende Feinpositionierung durch eine Stufe III Ortung erfolgen. Die resultierenden Standortdaten werden anschließend für eine koordinierte Informationsbereitstellung dem Kontextmanagementsystem übermittelt.

Kartierungssystem. Innerhalb des Kartierungssystems wird eine 3D-Karte der Wärmepumpenanlage anhand von CAD-Modelldaten mit Informationsbausteinen aus dem Datenbanksystem erstellt (Req-A & Req-B). Hierbei findet eine attributive Ortskennzeichnung und Ortszuweisung von Informationen zu den Bereichen der Wärmepumpenanlage statt. Idealerweise besteht die 3D-Karte aus einer vollständigen CAD-Modellabbildung der Anlage, die jedoch nicht voraussetzend ist, da die visuelle Darstellung der Anlage überwiegend als Unterstützung für das Hinterlegen von Informationsbausteinen innerhalb der Karte dient. Modelldaten sind allerdings essenziell, falls Modellerkennungsverfahren genutzt werden, um Informationsbausteine visuell an beispielsweise Bauteilen einzubetten. In diesem Anwendungsfall erfolgt eine Informationshinterlegung primär über Rasterkoordinaten. Es wird davon ausgegangen, dass ein Informationsbaustein mit hoher Wahrscheinlichkeit in einem definierten Sektor benötigt, respektive sich ein dazugehöriges Bauteil darin befindet. Somit können Informationen unabhängig von einem Bauteil referenziert werden, was beispielsweise eine

zonenbezogene Ausweisung von Sicherheitsinformationen ermöglicht. Auch die Koordinierung von Bereichen, an denen ein Modellerkennungsverfahren für die Bereitstellung eingebetteter Informationen über CV eingesetzt werden soll, erfolgt durch eine Ausweisung innerhalb des Rasterkoordinatensystems. Durch diese kombinierte Nutzung aus Rasterkoordinaten und Modellerkennungsverfahren soll ein skalierbares ressourcenschonendes System ermöglicht werden. Eine Bereichsdefinition, um Informationsbausteine innerhalb der Karte zu hinterlegen und somit Informationscluster zu bilden, erfolgt nach verschiedenen Aspekten. Einmal werden Bereiche anhand physischer Raummerkmale definiert, die sich aus baulichen Gegebenheiten wie Raumbegrenzungen durch Wände oder unterschiedliche Ebenen ergeben. Ebenso werden Bereiche um einzelne Bauteile respektive Baugruppen ausgewiesen. Dadurch ist es beispielsweise denkbar, dass im Anlagenbau etablierte Kraftwerk-Kennzeichensystem (KKS) zu integrieren, da darin alle Bauteile eindeutig bezeichnet und identifiziert sind und somit korrespondierende Informationen entsprechend zugeordnet werden können. Weitere Bereichseinteilungen sind nach Funktionen, Prozessen oder Sicherheitsaspekten möglich. Hierfür ist aber jeweils eine eindeutige örtliche Identifikation Voraussetzung. Die definierten Bereiche werden zum Abruf der Informationsbausteine mit beispielsweise Metadaten oder Tags versehen, um eine Referenz zu ermöglichen. Diese Kartenstruktur mit den Informationsclustern wird dem Kontextmanagementsystem letztendlich bereitgestellt.

Kontextmanagementsystem. Innerhalb des Kontextmanagementsystems werden die verschiedenen Rollen und weitere potenziell nützliche Filter (z. B. Aufgabe, Zeit, etc.) definiert, um den Zugriff auf Informationsbausteine der Wärmepumpenanlage zu regeln (Req-C). Damit können diese zielgerichtet gefiltert beispielsweise einem Techniker über die Benutzerschnittstelle bereitgestellt werden. Hierfür werden sämtliche Attribute, mit denen Informationsbausteine im Modellrahmen verknüpft sind, innerhalb dieses Systems koordiniert, sodass diese in den definierten Bereichen der 3D-Karte bereitgestellt werden können. Das Kontextmanagementsystem übernimmt hier ebenfalls die Koordinierung der Position auf Basis der erstellten 3D-Karte und der erhaltenen Standortwerte des Nutzers. Aufgrund der Vielfalt industrieller Anlagen existieren keine allgemeingültigen Rollen respektive Filter, weshalb eine jeweilige systematische Definition dieser erforderlich ist. Ein geeignetes Vorgehen besteht darin, zunächst notwendige Benutzergruppen zu identifizieren. Bei der hier behandelten Wärmepumpenanlage wären das unter anderem die Anlagenaufsicht, Mechaniker und Elektriker. Anschließend werden Aufgaben und Berechtigungen für jede dieser Benutzergruppen analysiert. Basierend darauf lassen sich einzelne Rollen mit dazugehörigen spezifischen Berechtigungen erstellen. In der hier behandelten Anlage zählen dazu beispielsweise „Anlagenaufsicht“, „Mechanik - Inspektion“, „Mechanik - Wartung“.

Benutzerschnittstelle. Vervollständigt wird das Konzept durch die Benutzerschnittstelle, die auf eine Verwendung von Tablets, Handys oder XR-Brillen ausgerichtet ist. Darin werden sämtliche relevanten Informationen einem Nutzer zugänglich gemacht und spezifische Konfigurationen zur Verfügung gestellt (Req-C & Req-D). Dafür wird auf die

verschiedenen Benutzerrollen und weitere Filter aus dem Kontextmanagementsystem zurückgegriffen. Unabhängig vom Endgerät werden diese, sowie weitere Informationen seitens des Kartierungs- und Kontextmanagementsystems, für den LBS bereitgestellt. Lediglich die Client-Anwendung, die unter anderem die grafische Benutzeroberfläche enthält, muss entsprechend des Endgerätes angepasst sein. Ein Hinzufügen von Informationsbausteinen zum Datenbanksystem über die grafische Benutzeroberfläche respektive Client-Anwendung, wie beispielsweise Prüfberichte, ist ebenfalls angedacht.

4 Evaluation

Das Konzept des räumlichen Informationsmodells wurde mit Experten aus Fachbereichen der Technischen Dokumentation, der multimedialen Anwendung und des Anlagenbaus diskutiert. Es wurden überwiegend Einzelgespräche mit neun Experten geführt, die in ihrem jeweiligen Fachbereich größtenteils mehr als zehn Jahre tätig waren. Alle Fachbereichsexperten geben dem Modell eine allgemeine Zustimmung und betonen die Bedeutung eines ganzheitlichen Ansatzes, um Informationen strukturiert und zielgerichtet bereitzustellen. Die Experten der Technischen Dokumentation (vier Personen) wiesen insbesondere auf die Schwierigkeit hin, Informationsbausteine zu definieren und zu extrahieren. Vor allem hinsichtlich der Rollendefinition im Kontextmanagement sehen sie das als herausfordernd an. Zudem raten diese bei einer Systementwicklung das im Kraftwerksbau überwiegend eingesetzte KKS zu beachten. Die Experten für multimediale Anwendung (drei Personen) unterstrichen die Bedeutung des Hinzufügens von Informationsbausteinen in das Datenbanksystem seitens eines Benutzers. Die Experten des Anlagenbaus (zwei Personen) halten einen Einsatz der Modellumsetzung bereits beim Erbauen einer Anlage für denkbar, um es anschließend für Wartungsvorgänge weiter zu nutzen respektive zu erweitern. Alle Experten wiesen auf die Schnittstellenproblematik, bedingt durch die Vielzahl an Systemen mit unterschiedlichen Datenformaten innerhalb einer Anlage, hin. Ein Großteil sieht aber das Modell als eine Grundlagenstruktur für eine ganzheitliche informationsbezogene Anlagenübersicht.

5 Fazit und Diskussion

In diesem Beitrag wurde ein Konzept für ein räumliches Informationsmodell für komplexe industrielle Anlagen vorgestellt. Dafür wurden Instandhaltungs- und Wartungsvorgänge bei einer komplexen Wärmepumpenanlage betrachtet. Daran wurden mögliche Anforderungen an ein Modell identifiziert und Bestandteile erläutert. Der daraus resultierende Modellansatz besteht aus einem Datenbanksystem, einem Positionierungssystem, einem Kartierungssystem, einem Kontextmanagementsystem und einer Benutzerschnittstelle. Die Systeme stehen in Wechselwirkung und sollen eine Grundlage für kontextabhängige LBSs innerhalb industrieller Anlagen bilden. Durch den Fokus des Beitrages auf die Modellkonzeptionierung wurde davon ausgegangen, dass benötigte Informationsbausteine bereits für bestimmte Anlagenbereiche vorgegeben sind.

Dies ist allerdings ein aufwendiger Prozess, der zusätzlich stark abhängig von einer Anlage ist und deshalb gesondert betrachtet werden sollte. Außerdem wird im Modell von einer statischen Karte ausgegangen, die durch eine dynamische Positionserkennung ergänzt werden kann. Zukünftig sollte, auch hinsichtlich der Aktualität einer Karte, eine dynamische Kartierung behandelt werden. Durch die enormen Fortschritte im Bereich der künstlichen Intelligenz sollte auch dessen Einsatz bei der Wechselwirkung zwischen den Systemen oder innerhalb eines jeweiligen Systems in Betracht gezogen werden.

Literaturverzeichnis

- [BDP21] Breque, M.; De Nul, L.; Petridis, A.: Industry 5.0: towards a sustainable human-centric and resilient European industry. LU: European Commission, Directorate-General for Research and Innovation, Luxembourg, 2021.
- [CL20] Chan, C. M.; Lau, S. L.: Using augmented reality and location-awareness to enhance visitor experience: A case study of a theme park app. In 2020 IEEE Graphics and Multimedia (GAME), IEEE, S. 49-54, 2020.
- [Es12] Estimote Inc., www.estimote.com, Stand: 15.04.2024.
- [Fa18] Falkowski, T. et al.: Toward A Desing Approach For Industrial Indoor Location-Based Services (I²LBS). In (Marjanović, D., Hrsg.): DS 92: Proceedings of the DESIGN 2018 15th International Design Conference. S. 1043-1054, 2018.
- [Fa22] Farahsari, P. S. et al.: A survey on indoor positioning systems for IoT-based applications. IEEE Internet of Things Journal, 9(10), S. 7680-7699, 2022.
- [GDF18] Grifoni, P.; D'Ulizia, A.; Ferri, F.: Context-awareness in location based services in the big data era. In (Xhafa, F., Hrsg.): Mobile Big Data: A Roadmap from Models to Technologies. Lecture Notes on Data Engineering and Communications Technologies, vol 10., Springer, S. 85-127, 2018.
- [Li20] Liu, J. et al.: A bim based hybrid 3d indoor map model for indoor positioning and navigation. ISPRS international journal of geo-information, 9(12), S. 747. 2020.
- [MS22] MS, Microsoft Corporation, www.learn.microsoft.com/en-us/azure/spatial-anchors/overview, Stand: 15.04.2024.
- [We14] Werner, M.: Indoor location-based services. Prerequisites and foundations, Springer, New York, 2014.
- [Xu21] Xu, X. et al.: Industry 4.0 and Industry 5.0—Inception, conception and perception. Journal of Manufacturing Systems, 61, S. 530-535, 2021.
- [YR15] Yassin, M.; Rachid, E.: A survey of positioning techniques and location based services in wireless networks. In 2015 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES), S. 1-5, 2015.
- [Zh22] Zhuang, Y. et al.: Bluetooth localization technology: Principles, applications, and future trends. IEEE Internet of Things Journal, 9(23), S. 23506-23524. 2022.

Integration verschiedener Informationstechnologien in der Verwaltungsschale am Beispiel von AAS und AutomationML

Arndt Lüder* (Universität Magdeburg, Austrian Center for Digital Production Wien), Jan Blume (Fraunhofer IOSB), Miriam Schleipen (EKS Intec), Matthias Freund (Festo SE & Co. KG), David Hoffmann, Pooja Gupta (Universität Magdeburg)

*Otto-v.-Guericke Universität /FMB/IAF/PSA
Universitätsplatz 2, 39106 Magdeburg, Deutschland
arndt.lueder@ovgu.de

1 Einleitung

Die zunehmende Digitalisierung von Produktionssystemen nimmt unter dem Schlagwort Industrie 4.0 eine zunehmende Dynamik an. Es entstehen dabei eine wachsende Anzahl von Demonstratoren, Technologiestudien, Anwendungsempfehlungen und betrachteten Anwendungsfällen [KW2022], die in der Mehrheit hin zu standardisierten Lösungen zur Informationsbereitstellung, -übertragung, und -nutzung weiterentwickelt werden sollen [VEP2019].

Die dabei entstehende Vielfalt von Entwicklungen wird jedoch zunehmend zum Problem. In teilweise parallel-laufenden Anstrengungen werden dieselben oder stark überlappende Anwendungsfälle betrachtet, konkurrierende Informationsmodelle geschaffen, verschiedene Technologien für denselben Anwendungsfall propagiert und sich teilweise widersprechende Standards angestrebt [Hoppe2023]. Dies ist vor dem Hintergrund der verschiedenen Standards, die zu berücksichtigen und ggf. geeignet zu kombinieren sind [DIN2023], zwar richtig, jedoch kann und wird dieses unkoordinierte Vorgehen für Endanwendende, die in die Standardisierungsaktivitäten nicht involviert sind und denen eventuell der tiefere Einblick fehlt, zunehmend zum Problem. Diese stellen sich also vermehrt die Frage, welche Investitionen die richtigen sind und welche Entwicklungsrichtungen für sie den nachhaltigsten Nutzeffekt erbringen können. Dies führt derzeit leider sehr häufig zum Aufschieben von Investitionsentscheidungen, was einen negativen Einfluss auf die Dynamik der Entwicklung der Industrie 4.0 hat [JEK2019].

Erste Arbeiten zur Harmonisierung der verschiedenen Technologien und Standards werden bereits seit einigen Jahren verfolgt und haben bereits erste Ergebnisse erbracht [HS2014]. Um diesem Problem jedoch nachhaltig entgegenzutreten, wird in einem gemeinsamen Diskussionspapier von VDMA, IDTA, OPC Foundation und AutomationML e.V sowie einigen größeren Unternehmen eine Integrationsstruktur für verschiedene Datenaustauschtechnologien vorgeschlagen [DMH2023]. Das Diskussionspapier beschreibt dabei eine Vorgehensweise, mit welcher die 3 Technologien AutomationML, OPC-UA, und AAS interoperabel verwendet werden können, um sich gegenseitig zu ergänzen und damit die jeweiligen Stärken bestmöglich auszunutzen. Um diese zum Leben zu erwecken sind jedoch einige Designentscheidungen zu treffen, die einen erheblichen Einfluss auf die Nutzbarkeit der Integration in der praktischen Anwendung besitzen.

Der AutomationML e.V. hat nun mit Blick auf die Datenaustauschtechnologien AAS und AutomationML einen ersten Vorschlag vorgelegt. Dieser betrachtet vorrangig die Integration von Engineering-Daten in eine AAS-basierte Industrie 4.0 Verwaltungsschale ohne jedoch eine Anwendung der entwickelten Ideen auch auf Verwaltungsschalen allgemein zu verhindern. Im Rahmen

dieses Beitrages soll diese Integrationsmethode vorgestellt und deren Auswirkungen an einem Implementierungsbeispiel diskutiert werden.

2 Grundlagen

Im Rahmen der Industrie 4.0 wird mit dem Referenzarchitekturmodell Industrie 4.0 [DIN2016] ein Framework spezifiziert, indem mit der Industrie 4.0 Komponente intelligente und mit self* Eigenschaften ausgestattete Funktionseinheiten kooperativ flexible Produktionssysteme realisiert werden. Diese Industrie 4.0 Komponenten bestehen aus zwei Grundelementen, dem eigentlichen cyber-physischen System, das eine Funktionalität im Produktionssystem bereitstellt und als Asset bezeichnet wird, sowie der Industrie 4.0 Verwaltungsschale, die als digitales Abbild und digitaler Zugang zur Komponente anzusehen ist [WEG2017]. Diese Struktur wird sowohl international diskutiert [MAP2018] als auch in ersten Prototypen realisiert [BCB2021, JBM2023].

Aufgabe der Industrie 4.0 Verwaltungsschale ist es, das von ihr beschriebene Asset und alle über dieses zum aktuellen Zeitpunkt bekannten Informationen für Nutzer gemäß den Ideen der IEC 62890 verfügbar zu machen [IEC2017]. Dazu stellt das Asset eine Informationsstruktur bereit [PI2022], die in Grundzügen in Bild 1 dargestellt ist. Neben einer Informationsmenge zur Identifikation des Assets werden mit den sogenannten Teilmodellen (Engl.: Submodels) Informationen bereitgestellt, die jeweils für verschiedene Anwendungsfälle entlang des Lebenszyklus des Assets dienen. Dabei werden in [PI2020] noch keine spezifischen Teilmodelle festgelegt. Dies erfolgt derzeit im Rahmen der IDTA (Industrial Digital Twin Association), die sich eine koordinierte Erstellung der Teilmodelle auf die Fahnen geschrieben hat (siehe <https://industrialdigitaltwin.org/content-hub/teilmodelle>).

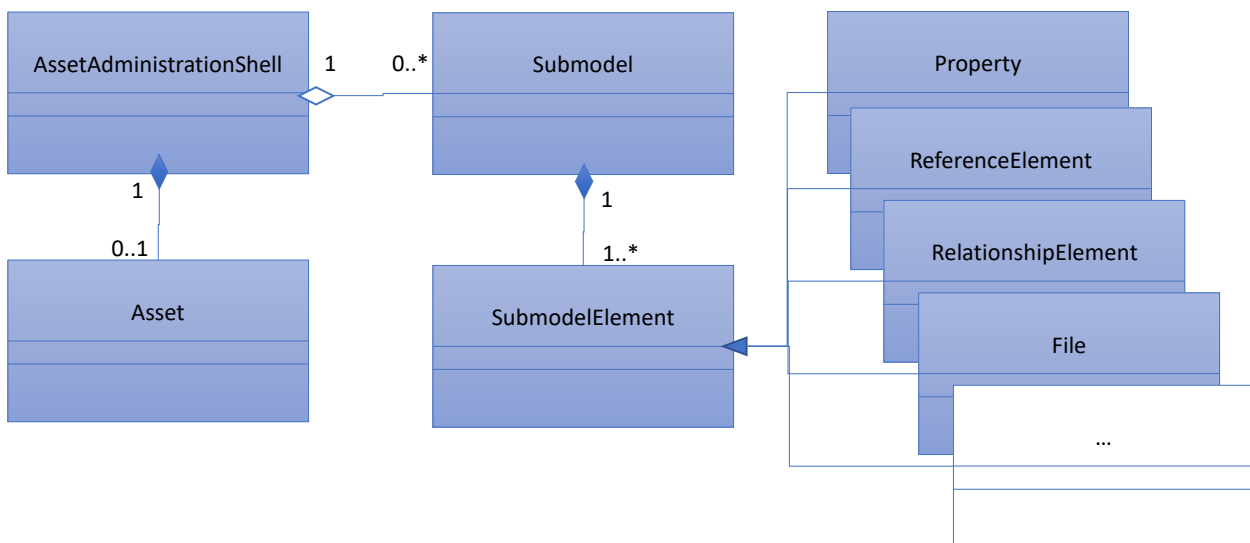


Bild 1: Informationsstruktur der Verwaltungsschale (in Anlehnung an [PI2022])

Zur technischen Realisierung einer Verwaltungsschale werden in [PI2022] neben dem AASX Format auch Formate wie AutomationML und OPC UA Knotenmodelle vorgeschlagen. Dies löst jedoch nicht die oben genannte Problematik der fehlenden Harmonisierung.

In diesem Dokument soll der Fokus auf AASX als Basis zur Umsetzung einer lebenszyklusübergreifenden Verwaltungsschale liegen. Eine Abbildung aller im Lebenszyklus relevanter Informationen, würde jedoch zu einer Neumodellierung von bekannten und bereits betrachteten Informationsmengen in einem weiteren Datenformat führen. Dies muss vermieden werden. In [DMH2023] wird vorgeschlagen, dass für Informationsmengen zu einem Asset eher eine quellen- und senkenorientierte Auswahl der Implementierungstechnologien zu verwenden ist. Dabei wird

vorgeschlagen, Informationsmengen, mit starkem Bezug zur Laufzeitphase eines Assets über OPC UA Knotenmodelle, Informationsmengen mit starkem Bezug zur Entwurfsphase eines Assets über AutomationML, und Informationsmengen, die einen organisation- bzw. lebenszyklusübergreifenden Bezug besitzen, über AAS abzubilden. Zur Integration soll dann, wie in Bild 2 gezeigt, zwischen Teilmodelle und Detailmodellen unterschieden werden. Für Entwurfs- und Laufzeitdaten sollen die Detailmodelle alle relevanten Daten enthalten und in den entsprechenden Technologien (AutomationML und OPC UA) verfügbar sein. Diese werden dann über Teilmodelle, in denen nur der lebenszyklusübergreifende Anteil dieser Informationen direkt publiziert wird, referenziert werden. Im AutomationML e.V. wird derzeit (in Kooperation mit dem DIAMOND Projekt¹) an der Realisierung der entsprechenden Teilmodell-Vorlagen (Submodel Templates) als Blaupause für die Integration gearbeitet, die dann in der IDTA standardisiert werden sollen.

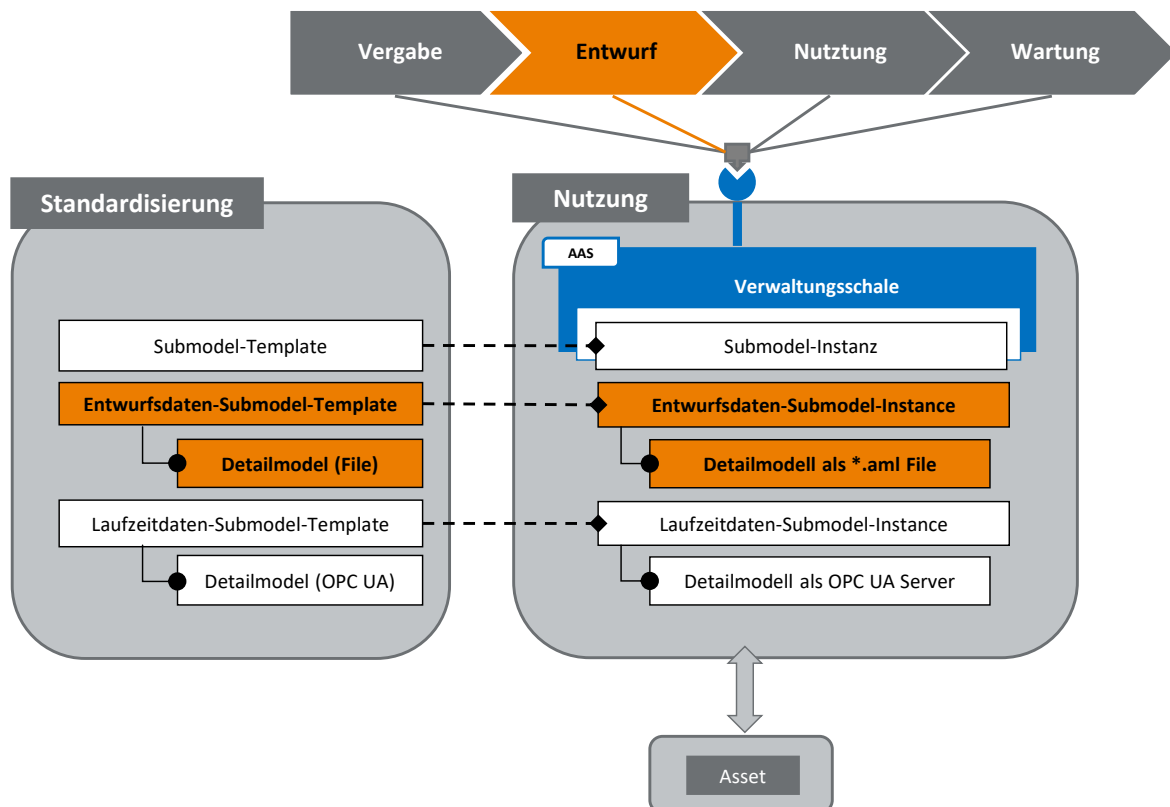


Bild 2: Integrationsstruktur der Verwaltungsschale (in Anlehnung an [DMH2023])

Das AutomationML Datenaustauschformat wird vom AutomationML e.V. als offene und kostenfreie, herstellernerneutrale und XML-basierte Datenaustauschtechnologie für den Informationsaustausch entlang der Entwurfsnetzwerke für Produktionssysteme entwickelt. Es ermöglicht die Abbildung von Objekthierarchien als sogenannte *InstanceHierarchy*, in deren Rahmen für Objekte als (ggf. von *SystemUnitClasses* abgeleiteten) *InternalElements* abgebildet werden. Diese *InternalElements* können Semantiken als *RoleClass*, Eigenschaften als *Attribute* und Beziehungen zu anderen Objekten als Kombination aus (von *InterfaceClasses* abgeleiteten) *ExternalInterfaces* und *InternalLinks* besitzen [Drath2022].

Für die Abbildung von Entwurfsdaten für spezifische Anwendungsfälle im Entwurf von Produktionssystemen stellt AutomationML anwendungsfallspezifische *Role-* und

¹ diamond-project.de

InterfaceClassLibraries bereit, die in den verschiedenen Teilen des AutomationML Standards beschrieben werden [LSD2017].

Wie in [LBB2020, BRL2021, LMB2022] beschrieben wird, kann AutomationML als Datenaustauschformat für Entwurfsnetzwerke in einfacher Form verwendet werden. Dabei wird eine, als Common Concepts benannte und an die Teilmodell-Struktur der Verwaltungsschale angelehnte Modellierungsmethode verwendet. Bild 3 zeigt dabei beispielhaft die sich ergebende Datenstruktur für zwei durch ein Kabel verbundene Automatisierungsgeräte. Für jedes Objekt des zu modellierenden Produktionssystems entsteht ein *InternalElement*, das jeweils anwendungsfallsspezifische *InternalElements* mit den anwendungsfallsspezifischen Semantiken, Attributen und Relationen enthält.

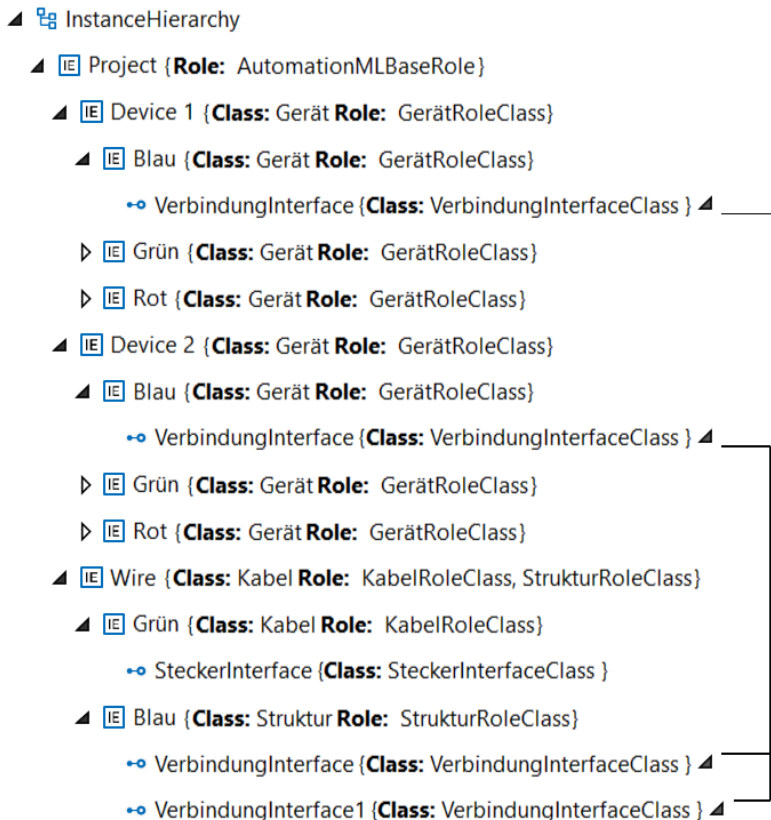


Bild 3: Common Concept basierte AutomationML Modellstruktur für Entwurfsnetzwerke (in Anlehnung an [LMB2022])

3 Integrationsmethode

Ziel der hier vorgestellten Integrationsmethode für AutomationML Projekte in AAS ist die transparente Integration von Entwurfsdaten, welche in einem oder mehreren AutomationML-Projekten enthaltenen sind, in einem Teilmodell einer Verwaltungsschale. Dazu werden derzeit (wie in Bild 4 gezeigt) drei Integrationsbausteine definiert, die gemeinsam verwendet werden können.

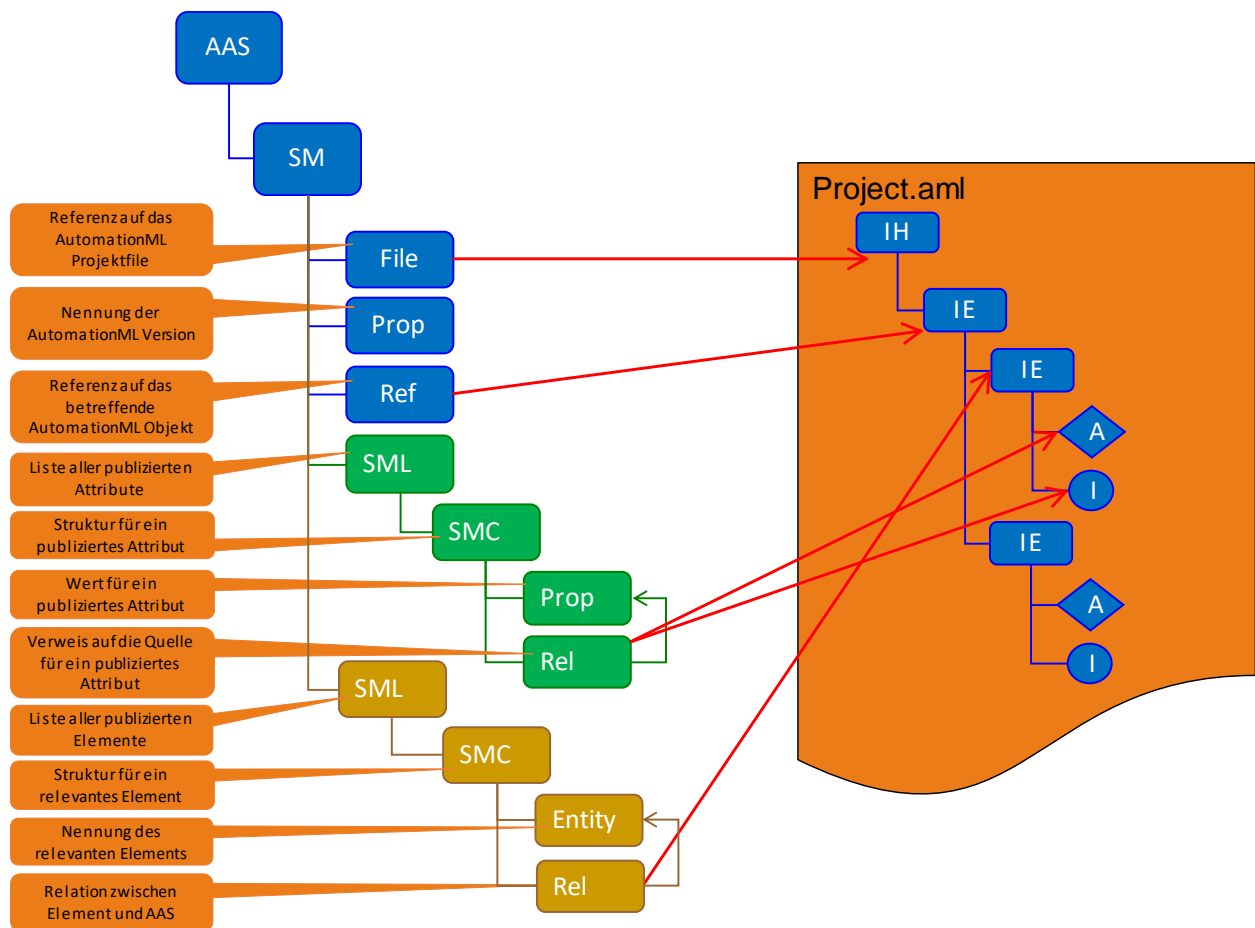


Bild 4: Generelle Integrationsstruktur

Der erste Integrationsbaustein (blauer Anteil in Bild 4) ermöglicht die Integration von AutomationML-Dateien sowie das einfache Referenzieren von Inhalten dieser Dateien aus einem Teilmodell heraus. Damit wird die einfache Weitergabe der Entwurfsdaten über eine Verwaltungsschale adressiert. Dazu soll dieses Teilmodell zwei verpflichtende und ein optionales Teilmodell-Element (Engl.: Submodel Element) enthalten. Das erste verpflichtende Submodelelement ist ein Datei-Element (Engl.: File Element). Es soll als *AutomationMLData* benannt sein, eine Referenz auf das AutomationML Projektfile sowie als *SemanticID* einen String der Form https://www.automationml.org/AutomationML_AAS_SemanticIDs/General/FileReference enthalten. Das zweite verpflichtende Teilmodell-Element ist ein als *AutomationMLVersion* benanntes Property Element, das die im AutomationML Projekt verwendete AutomationML Version angibt und als *SemanticID* einen String der Form https://www.automationml.org/AutomationML_AAS_SemanticIDs/General/VersionNumber enthält. Auf Basis dieser Property können Parser entscheiden, ob sie die AutomationML Daten verarbeiten können. Das optionale Element wird durch ein Referenzelement gebildet. Dieses soll als *AutomationMLElementReference* benannt sein, als *SemanticID* einen String der Form https://www.automationml.org/AutomationML_AAS_SemanticIDs/General/ElementReference enthalten und der Benennung des relevanten Einstiegspunktes zur Nutzung des AutomationML Files dienen, falls dies nicht die Wurzel der *InstanceHierarchy* ist.

Der zweite Integrationsbaustein (grüner Anteil in Bild 4) ermöglicht das Publizieren von einzelnen Attributwerten aus einem AutomationML Projekt in einem Teilmodell zwecks lebenszyklusübergreifender Weitergabe. Dazu muss ein Teilmodell ein verpflichtendes Teilmodell-Element vom Typ *SubmodelElementList* enthalten. Dieses soll *AutomationMLAttributesList* heißen und als *SemanticID* einen String der Form https://www.automationml.org/AutomationML_AAS_SemanticIDs/General/AttributeList enthalten. Inhalt dieser Liste sollen Teilmodell-

Elemente vom Typ `SubmodelElementCollection` sein. Jedes dieser Elemente dient zum Publizieren eines Attributes und soll als `SemanticID` einen String der Form `https://www.automationml.org/AutomationML_AAS_SemanticIDs/General/AttributeContainer` enthalten. Jedes dieser attributbezogenen Elemente besteht dann aus zwei `SubmodelElements`. Das erste Teilmodell-Element ist vom Typ `Property`, trägt den Namen des zu repräsentierenden Attributes und enthält dessen Wert und hat als `SemanticID` einen String der Form `https://www.automationml.org/AutomationML_AAS_SemanticIDs/General/AutomationMLAttribute`. Das zweite Teilmodell-Element ist vom Typ `RelationshipElement`, trägt den Namen `RelationBetweenPropertyAndAttribute`, hat als `SemanticID` einen String der Form `https://www.automationml.org/AutomationML_AAS_SemanticIDs/General/ElementReference` und enthält eine `ExternalReference` auf das AutomationML Attribute über eine `FragmentReference` sowie eine `ModelReference` auf das `PropertyElement`.

Der dritte Integrationsbaustein (gelber Anteil in Bild 4) ermöglicht das Identifizieren relevanter Teilstrukturen bzw. Elemente in einem AutomationML Projekt aus einem Teilmodell heraus. Dazu muss dieses Teilmodell ein verpflichtendes Teilmodell-Element vom Typ `SubmodelElementList` enthalten. Dieses soll `AutomationMLElementList` heißen und als `SemanticID` einen String der Form `https://www.automationml.org/AutomationML_AAS_SemanticIDs/General/ElementList` enthalten. Inhalt dieser Liste sollen Teilmodell-Elemente vom Typ `SubmodelElementCollection` sein. Diese enthalten als `SemanticID` den String `https://www.automationml.org/AutomationML_AAS_SemanticIDs/General/ElementContainer`. Jede `SubmodelElementCollection` besteht dann aus einem `Entity` Element und einem `RelationshipElement`. Das Entity Element soll das entsprechende AutomationML Element abbilden, den Namen des betreffenden AutomationML Elementes tragen und als `SemanticID` einen String der Form `https://www.automationml.org/AutomationML_AAS_SemanticIDs/General/ElementReference` enthalten. Das `Relation` Element dient der Verlinkung des Entity Elementes und des AutomationML Elementes, trägt den Namen `RelationBetweenElementAndAAS`, enthält eine `SemanticID` der Form `https://www.automationml.org/AutomationML_AAS_SemanticIDs/General/AASRelation` und verweist über eine `ExternalReference` vom Typ `FragmentReference` auf das betreffende AutomationML Element sowie über eine `ModelReference` auf das Entity Element.

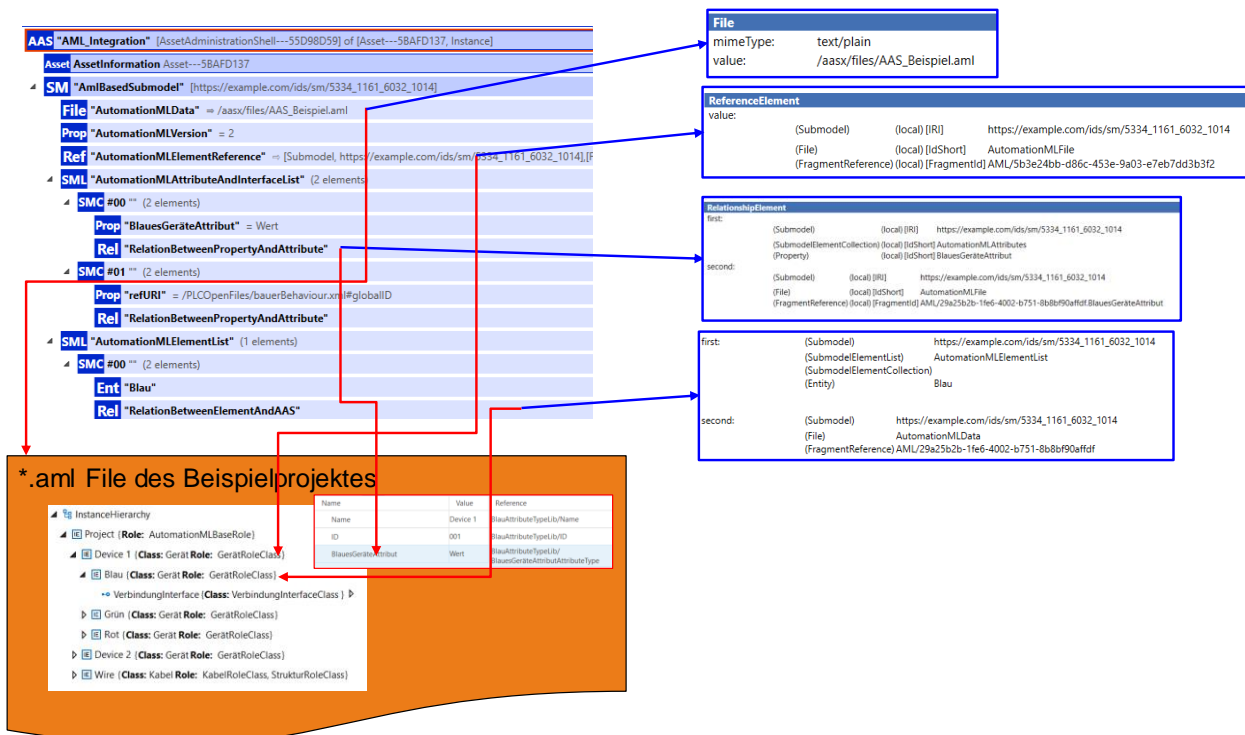


Bild 5: generisches Anwendungsbeispiel

Bild 5 zeigt ein Beispiel, in dem alle drei Integrationsbausteine auf das AutomationML Projektbeispiel aus Bild 3 angewendet werden.

In bestimmten Fällen können bestehende Teilmodelle bereits die im AutomationML Projekt enthaltenen Attribute und Elemente abbilden. Um hier Redundanzen zu vermeiden können diese in die oben beschriebene Modellierung eingebunden werden. Im Falle von Attributen entfällt dann das *Property* Element und es wird im *Relationship* Element auf das entsprechende *Property* Element des bestehenden Teilmodelles verwiesen. Im Falle eines AutomationML Elementes (das dann als Entity bereits in einem anderen Teilmodell benannt ist) wird analog verfahren und das *Entity* Element entfällt.

Ebenso ist es wichtig zu wissen, ob eine Änderung von Inhalten im AutomationML Projekt die Inhalte im Teilmodell ändern muss oder ob dies umgekehrt der Fall ist, wer also als Data-Owner zu sehen ist. Dies wird über die *Relationship* Elemente in den *SubmodelElementCollection* abgebildet. Die im *first* Teil des *Relation* Elementes genannte Seite ist für diesen Informationsinhalt zuständig. Die andere hat „nur“ Leseberechtigung.

4 Beispiel einer technischen Realisierung

Als Validierung der beschriebenen Methode wurde durch die Autoren der Otto-von-Guericke Universität die in Bild 6 dargestellt prototypische Realisierung für die Integration implementiert und getestet.

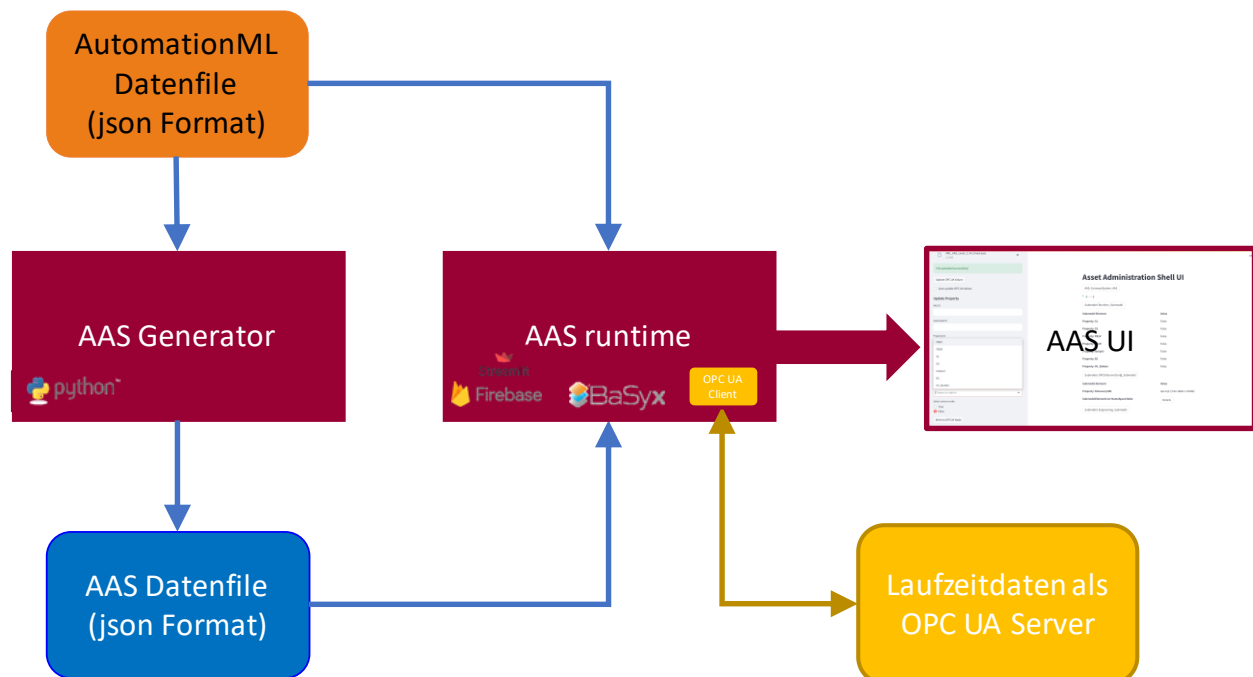


Bild 6: Prototypische Realisierung

Grundideen dieser Implementierung sind (a) eine Zweiteilung hinsichtlich Informationsmodellierung und Informationsnutzung, (b) eine klare Aufgabenverteilung entlang der 3 Technologien AutomationML, OPC-UA, und AAS entlang des Lebenszyklus von Produktionssystemen und (c) die Nutzung frei verfügbarer Implementierungstechnologien. Entsprechend entstanden zwei Softwaresysteme.

Das erste Softwaresystem implementiert in Python die automatische Generierung einer Verwaltungsschale entsprechend der obigen Beschreibung auf Basis eines AutomationML Datenfiles, das

als json File verfügbar ist. Dabei wird davon ausgegangen, dass im AutomationML Datenfile die Assets über eine AutomationML Struktur in Analogie zu [LBB2020] und [BRL2021] gestaltet sind und in deren Substrukturen die Teilmodelle über entsprechende Rollen aus [AML2019] und die in AAS abzubildenden Attribute über eine AAS Facette gekennzeichnet sind. Im Ergebnis erzeugt das erste Softwaresystem ein entsprechendes AAS Datenfile, das ebenfalls als json File gestaltet ist.

Das zweite Softwaresystem implementiert unter Nutzung des BaSyx Python SDK, Streamlit und Firebase eine AAS Laufzeitumgebung. Diese Laufzeitumgebung nutzt (a) explizit die Key Value Pair Struktur von json und die damit mögliche Integration beider Informationssätze in einer Firebase Datenbank, die zu einer Transparenz der Herkunft von Informationen führt, und (b) integriert sie OPC UA Clients zur Gewinnung von Laufzeitdaten auf Basis der Darstellung des Zugangsweges zu OPC UA Knoten in einem OPC UA Server gemäß [AML2017].

Mit den beiden Softwaresystemen ist es möglich aus Entwurfsdaten automatisch Verwaltungsschalen zu generieren, diese mit OPC UA basierten Laufzeitdaten zu verknüpfen und die dabei entstandene Verwaltungsschale als Server zur Nutzung zur Verfügung zu stellen.

5 Erste Anwendungsfälle

Im Rahmen des Diamond Projekts werden derzeit gemeinsam mit der Tool Chain Arbeitsgruppe das AutomationML e.V. erste Teilmodelle standardisiert. Diese gliedern sich in den Entwurfsprozess von Produktionssystemen ein und stellen dessen Ergebnisse für spätere Lebenszyklusphasen zur Verfügung. Als ein Anwendungsbeispiel kann hier die Nutzung von Entwurfsdaten zum Assetmanagement genannt werden.

Im Rahmen der Dissertationsschrift [Langosch2024] wurde ein Informationsmodell vorgeschlagen, das die für das Assetmanagement im Automobilbau relevanten Entwurfswerte abbildet. Diese entstammen den Entwurfsschritten des Basic Engineering, der mechanischen und elektrischen Konstruktion sowie der Steuerungsprogrammierung oder kommen über Gerätebeschreibungen vom Lieferanten entsprechender Komponenten.

Im Rahmen des Diamond Projektes werden derzeit gemeinsam mit der Tool Chain Arbeitsgruppe das AutomationML e.V. die dazu passenden AutomationML Informationsmodelle sowie die dazu gehörenden AAS Teilmodelle

- Planning data for operators and maintenance,
- Detailed structure of production systems,
- MCAD,
- ECAD,
- PLC programming, and
- Robot online programming and simulation

entwickelt um diese gemeinsam mit der IDTA zu standardisieren. Damit wird ein verlustfreier Übergang von Entwurfsdaten zum Assetmanagement ermöglicht.

6 Zusammenfassung

Im vorliegenden Beitrag wird der aktuelle Stand der vom AutomationML e.V. entwickelten Integrationsmethode von AutomationML und AAS in einer Verwaltungsschale vorgestellt, die gemäß dem Diskussionspapier der Verbände [DMH2023] mit dem Ziel entsteht, den größten Nutzen aus den jeweils integrierten Technologien für die Industrie 4.0 zu ziehen. Er zeigt, dass eine solche Integration problemlos möglich und technisch umsetzbar ist. In den aktuellen Arbeiten das

AutomationML e.V. sowie des DIAMOND Projektes wird diese Integrationsmethode weiter verfeinert und auf die ersten AutomationML basierten Informationsmodelle angewendet. Dies bildet die Basis für die Umsetzung entsprechender Teilmodell-Vorlagen für die Integration bestimmter Informationsmodelle, die in der IDTA standardisiert werden (siehe industrialdigitaltwin.org/content-hub/teilmodelle).

Acknowledgement

Diese Arbeit entstand mit Unterstützung der Christian Doppler Forschungsgesellschaft, dem Österreichischen Bundesministerium für Digitalisierung und Finanzen und der National Foundation for Research, Technology and Development Österreichs. Das Kompetenzzentrum Center for Digital Production wird im Rahmen des COMET – Competence Centers for Excellent Technologies Programmes durch das BMVIT, BMDW, und das Österreichische Bundesland unter Management der FFG gefördert.






7 Literatur

- [KW2022] H. Kagermann, W. Wahlster: Ten Years of Industrie 4.0, in: Sci, Volume 4, Issue 3, 2022, <https://doi.org/10.3390/sci4030026>.
- [VEP2019] N. Velasquez Villigran, E. Estevez, P. Pesado, J. de Juanes Marquez: Standardization: A Key Factor of Industry 4.0, in: 2019 Sixth International Conference on eDemocracy & eGovernment (ICEDEG), 2019, P. 350–354. <https://doi.org/10.1109/ICEDEG.2019.8734339>
- [DIN2023] DIN: Deutsche Normungsroadmap Industrie 4.0 – Version 5, Januar 2023, <https://www.din.de/resource/blob/907746/be4b76ec37bc6b18b46dd15082865e92/nrm-industrie-4-0-version-5-2023-final-web-data.pdf>
- [JEK2019] R. Joppen, S. von Enzberg, A. Kühn, R. Dumitrescu: Investitionsentscheidungen vor dem Hintergrund der Digitalisierung am Beispiel Schaltschrankbau, Zeitschrift für wirtschaftlichen Fabrikbetrieb, vol. 114, no. 7-8, 2019, pp. 483-487. <https://doi.org/10.3139/104.112119>
- [Hoppe2023] S. Hoppe: Das Bindeglied schlechthin zwischen OT und IT, Automatisierung, Jahrgang 2023, Heft 5, S. 35-37, <http://www.austromatisierung.at/Ausgabe523/webpaper.html>.
- [HS2014] R. Henssen, M. Schleipen: Interoperability between OPC-UA and AutomationML. Disruptive Innovation in Manufacturing Engineering towards the 4th Industrial Revolution, 8th International CIRP Conference on Digital Enterprise Technology, Procedia CIRP, Volume 25, 2014, Pages 297–304, <https://doi.org/10.1016/j.procir.2014.10.042>.
- [DMH2023] R. Drath, C. Mosch, S. Hoppe, A. Faath, E. Barnstedt, B. Fiebiger, W. Schlögl: Diskussionspapier – Interoperabilität mit der Verwaltungsschale, OPC UA und AutomationML: Zielbild und Handlungsempfehlungen für industrielle Interoperabilität, <https://www.automationml.org/wpcontent/uploads/2023/04/Diskussionspapier-Zielbild-und-Handlungsempfehlungen-fuer-industrielle-Interoperabilitaet-5.3.pdf>.
- [PI2022] Plattform Industrie 4.0: Details of the Asset Administration Shell - Part 1 - The exchange of information between partners in the value chain of Industrie 4.0, https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/Details_of_the_Asset_Administration_Shell_Part1_V3.pdf, Mai 2022.
- [DIN2016] DIN Spec 91345: Reference Architecture Model Industry 4.0 (RAMI 4.0), Beuth publisher, April 2016,

https://www.iec.ch/dyn/www/f?p=103:38:508042177081439:::FSP_ORG_ID,FSP_APEX_PAGE,FSP_PROJECT_ID:1250,20,23819.

- [WGE 2017] Wagner C., Grothoff J., Epple U., Drath R., Malakuti S., Grüner S., Hoffmeister M., Zimmermann P.: The role of the Industry 4.0 Asset Administration Shell and the Digital Twin during the life cycle of a plant. In: Proceedings of the 2017 IEEE Conference on Emerging Technologies Factory Automation (ETFA 2017), Limassol, Cyprus, 2017.
- [MAP2018] Ministry of Economy and Finances (France), Alliance Industrie du Futur (Italy), Federal Ministry for Economic Affairs and Energy (BMWi) and Plattform Industrie 4.0 (Germany): The Structure of the Administration Shell: TRILATERAL PERSPECTIVES from France, Italy and Germany, 2018, <https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/hm-2018-trilaterale-coop.html>
- [BCB2021] S. Beden, Q. Cao and A. Beckmann, "Semantic Asset Administration Shells in Industry 4.0: A Survey," *2021 4th IEEE International Conference on Industrial Cyber-Physical Systems (ICPS)*, Victoria, BC, Canada, 2021, pp. 31-38, doi: 10.1109/ICPS49255.2021.9468266.
- [JBM2023] Jacoby M, Baumann M, Bischoff T, Mees H, Müller J, Stojanovic L, Volz F. Open-Source Implementations of the Reactive Asset Administration Shell: A Survey. *Sensors*. 2023; 23(11):5229. <https://doi.org/10.3390/s23115229>
- [IEC2017] International Electrotechnical Commission: IEC 62890 - Life-cycle management for systems and products used in industrial-process measurement, control and automation, Beuth publisher, April 2017.
- [Drath2022] Drath, R.: AutomationML: Das Lehrbuch für Studium und Praxis, Berlin, Boston: De Gruyter Oldenbourg, 2022. <https://doi.org/10.1515/9783110782998>.
- [LSD2017] A. Lüder, N. Schmidt, R. Drath: Standardized information exchange within production system engineering, In: Multi-disciplinary engineering for cyber-physical production systems: data models and software solutions for handling complex engineering projects - Cham: Springer-Verlag, S. 235-257, 2017
- [LBB2020] A. Lüder, L. Baumann, A. -K. Behnert, F. Rinker and S. Biffel, "Paving Pathways for Digitalization in Engineering: Common Concepts in Engineering Chains," *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, Vienna, Austria, 2020, pp. 1401-1404, doi: 10.1109/ETFA46521.2020.9212009.
- [BRL2021] A. -K. Behnert, F. Rinker, A. Lüder and S. Biffel, "Migrating Engineering Tools Towards an AutomationML-Based Engineering Pipeline," *2021 IEEE 19th International Conference on Industrial Informatics (INDIN)*, Palma de Mallorca, Spain, 2021, pp. 1-7, doi: 10.1109/INDIN45523.2021.9557517.
- [LMB2022] A. Lüder, K. Meixner, S. Biffel: Engineering Data Treasures, Their Collection and Use, IFAC-PapersOnLine, Volume 55, Issue 10, 2022, doi: 10.1016/j.ifacol.2022.10.105.
- [AML2019] AutomationML e.V.: Application Recommendation - Asset Administration Shell (AAS) Representation Version 1.0.0, November 2019, https://www.automationml.org/wp-content/uploads/2022/04/Asset-Administration-Shell-Representation-V1_0_0.zip.
- [AML2017] AutomationML e.V.: Best Practice Recommendation – Data Variable, Version 1.0.0, May 2017, https://www.automationml.org/wp-content/uploads/2022/05/BPR_007E_BPR_DataVariable_V1.0.0.zip
- [Langosch2024] M. Langosch: Ideales Informationsmodell für die Übergabe digitaler Anlagenmodelle aus dem Engineering in den Fabrikbetrieb, Promotion, Otto-v.-Guericke Universität, Fakultät Maschinenbau, 2024, eingereicht.

Ansätze für die Modellierung von Streckeninformationen zur Verifikation im Steuerungsentwurf


Aron Schnakenbeck ¹, Robin Mroß ², Marcus Völker ², Stefan Kowalewski ² und Alexander Fay ³




Abstract: Während der Entwicklung von automatisierten Anlagen ist die Spezifikation der Steuerungslogik ein wichtiger Schritt. Um diesen Prozess mit formalen Methoden zu unterstützen, kann die Steuerungslogik zunächst mithilfe eines formalen Modells spezifiziert werden, was eine formale Verifikation ermöglicht. In einem Vorgängerbeitrag der Autoren wurde eine solche Verifikation vorgestellt, die es ermöglicht, GRAFCET-Spezifikationen zu analysieren. In diesem Beitrag wird dieses Konzept um die Modellierung von Streckeninformationen erweitert. Die Modellierung von Streckeninformationen ermöglicht eine genauere Verifikation und in Teilen eine Reduktion des Zustandsraums. Anhand eines durchgängigen Beispiels werden zwei Ansätze für die Modellierung von Streckeninformationen vorgestellt und verglichen: die Modellierung mittels Invarianten und die Modellierung mittels GRAFCET. Bei der Modellierung mittels GRAFCET wird gezeigt, wie neben dem Verhalten der Steuerstrecke auch zusätzliche Eigenschaften des kombinierten Verhaltens modelliert werden können, wie beispielsweise die korrekte Abarbeitung des SPS-Zyklus. Um diese zusätzlichen Eigenschaften anwenderfreundlich modellieren zu können, wird eine Erweiterung von GRAFCET vorgestellt und gezeigt, wie dieses Konzept zur Modellierung von Streckeninformationen in die bereits vorgestellte Werkzeugkette zur Verifikation von GRAFCET-Spezifikationen integriert werden kann.


Keywords: GRAFCET, Verifikation, formale Methoden, Streckenmodellierung

1 Einleitung

Während der Entwicklung von automatisierten Anlagen ist die Spezifikation der Steuerungslogik ein wichtiger Schritt. Allerdings ist die in der Praxis verbreitete, manuelle Implementierung dieser Steuerungslogik aus informellen Anforderungen heraus fehleranfällig [SF14]. Formale Methoden können diesen Prozess unterstützen, indem zunächst die Steuerungslogik mithilfe eines formalen Modells spezifiziert wird. Dieses Vorgehen hat die Vorteile, dass die formale Spezifikation als Dokumentation für den Code genutzt werden kann und eine automatisierte Codegenerierung möglich wird [Ju19]. Zudem kann eine formale Verifikation der Steuerungslogik früh in den Entwicklungsprozess eingebunden

1 Helmut-Schmidt-Universität, Institut für Automatisierungstechnik, Holstenhofweg 85, 22043 Hamburg, aron.schnakenbeck@hsu-hh.de,  <https://orcid.org/0000-0002-3507-2851>

2 RWTH Aachen University, Lehrstuhl Informatik 11, Ahornstraße 55, 52074 Aachen, mross@embedded.rwth-aachen.de,  <https://orcid.org/0000-0002-1007-5597>; voelker@embedded.rwth-aachen.de,  <https://orcid.org/0000-0001-7348-0146>; kowalewski@embedded.rwth-aachen.de,  <https://orcid.org/0000-0002-7184-4804>

3 Ruhr-Universität Bochum, Universitätsstraße 150, 44801 Bochum, alexander.fay@rub.de,  <https://orcid.org/0000-0002-1922-654X>

werden. In der Vergangenheit haben die Autoren dieses Beitrags in [Sc23] eine Werkzeugkette vorgestellt, die einen solchen formalen Entwicklungsprozess von Steuerungscode abbildet und eine formale Verifikation enthält. Allerdings berücksichtigt der vorgestellte Ansatz bei der Verifikation nur das Steuerungsverhalten. Das Verhalten der Steuerstrecke wurde nicht modelliert, sondern es wurde angenommen, dass sich die Eingangssignale der Steuerung nichtdeterministisch verhalten. Während der Evaluation dieses Modellierungsansatzes stellte sich heraus, dass gewisse Eigenschaften ohne weitere Informationen über das Streckenverhalten nicht definitiv nachgewiesen werden können. Dazu gehört beispielsweise, dass das einschränkende Verhalten der Steuerstrecke zu Verklemmungen führen kann, also dass es keine Transition mehr gibt, die noch schalten kann. Ein weiterer Vorteil das Streckenverhalten zu modellieren kann sein, dass es den Zustandsraum während der Verifikation einschränkt und so einer Zustandsraumexplosion entgegenwirken kann. Daher wird in diesem Beitrag die Modellierung von Streckenverhalten für die formale Verifikation diskutiert. Zudem wird gezeigt, wie die in [Sc23] vorgestellte Werkzeugkette erweitert werden kann, um die Modellierung des Streckenverhaltens während der Verifikation zu berücksichtigen.

Da die Akzeptanz von formaler Verifikation in der Praxis immer noch gering ist [Vo14], wird in diesem Beitrag, sowie in dem von den Autoren vorgestellten Verifikationsansatz [Sc23], GRAFCET als Beschreibungsmittel verwendet, das in der automatisierungstechnischen Praxis anerkannt ist [SF14]. GRAFCET ist eine in [DIN EN 60848] genormte, graphische Spezifikationsprache, die neben sequenziellem, alternativem und nebenläufigem Verhalten auch die Modellierung von hierarchischen Steuerungsstrukturen (mittels sogenannter *Teilgrafcets*) ermöglicht.

In ihrer Studie zu existierenden Modellierungsansätzen von Streckenverhalten in der Literatur zeigen Galvão et al. [Ga19], dass der Grad der Granularität entscheidend ist, mit dem das Streckenverhalten modelliert wird. Zum einen schreiben die Autoren, dass eine detailliertere Modellierung einen erhöhten Ressourcenbedarf bei der Verifikation erfordert und, dass der Modellierungsgrad beeinflusst, welche Eigenschaften verifiziert werden können. Daher wird in dieser Arbeit anhand eines durchgängigen Beispiels gezeigt, wie die Streckenmodellierung über zwei Ansätze verschiedener Granularität erfolgen kann: die Modellierung über boolesche Invarianten und die Modellierung mittels GRAFCET. Der Aufwand zur Erstellung eines vollständigen formalen Modells der Steuerstrecke ist allerdings hoch, und es ist nicht davon auszugehen, dass ein solches in der Praxis vorliegt. Daher ist das Ziel dieses Ansatzes die Modellierung von Teilm Informationen, die für die Verifikation von besonderer Bedeutung sind, und in einer Form, die für den Anwender möglichst eingängig ist, was die Verwendung von beispielsweise temporalen Logiken ausschließt.

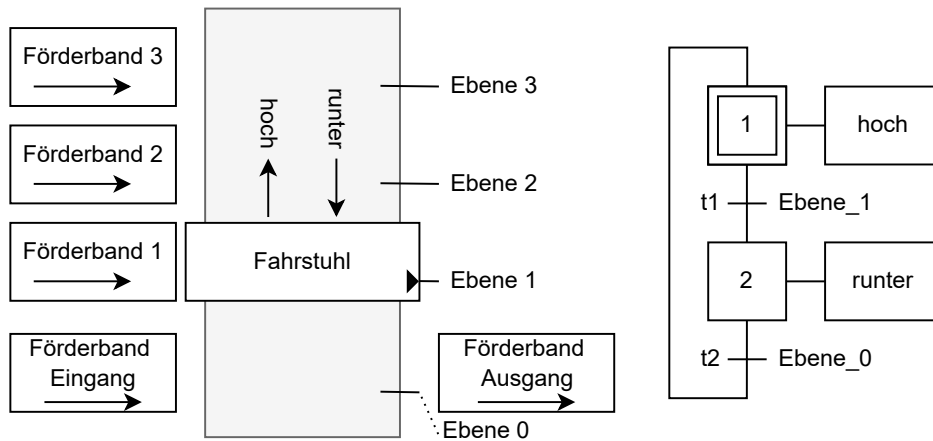


Abb. 1: Lastenfahrstuhl als Steuerstrecke und die zugehörige Steuerung als GRAFCET-Instanz.

2 Streckenmodellierung

Als Beispiel wurde ein Lastenfahrstuhl gewählt (siehe Abb. 1). Dieser hat auf vier verschiedenen Ebenen Förderbänder als Eingänge und auf der untersten Ebene ein Förderband als Ausgang. Der Fahrstuhl verfügt über zwei Aktoren, die den Fahrstuhl herauf- und herunterbewegen, und mehrere Sensoren, die die Anwesenheit des Fahrstuhls in einer bestimmten Höhe registrieren. Die vereinfacht dargestellte Steuerung sorgt dafür, dass der Fahrstuhl nach oben fährt, bis die erste Ebene erreicht ist, und im Anschluss nach unten fährt, bis die unterste Ebene erreicht ist.

2.1 Modellierung von Streckeneigenschaften mit Invarianten

Der Zusammenhang zwischen den Sensorvariablen, der sich aus der Geometrie der Strecke ergibt, kann über Invarianten in Form von einfachen booleschen Bedingungen formuliert werden. Im Beispiel des Fahrstuhls (siehe Abb. 1) kann zu jedem Zeitpunkt nur maximal einer der Sensoren den Fahrstuhl registrieren. So ist die entsprechende Variable genau dann 1, wenn der Fahrstuhl registriert wird, anderenfalls 0:

$$(Ebene_0 + Ebene_1 + Ebene_2 + Ebene_3 \leq 1) = wahr \quad (1)$$

Eine wichtige Anforderung für den Steuerungsentwurf ist, dass die Steuerung immer eine nächste stabile Situation erreichen kann. Für die Steuerung in Abb. 1 ist das gegeben, wenn die beiden Transitionsbedingungen nicht gleichzeitig *wahr* sein können, also *Ebene_0* und *Ebene_1* nicht gleichzeitig *wahr* sein können. Die Verifizierung dieser Anforderung wird unter der Berücksichtigung der Invariante möglich. Vergleichbares gilt bei der Verifikation,

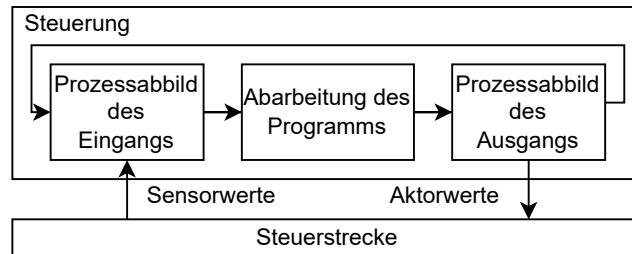


Abb. 2: Kombiniertes Verhalten aus Steuerung und Steuerstrecke. Darstellung nach [Lu20].

dass Alternativverzweigungen nur exklusiv schalten können, wie von der Norm [DIN EN 60848] vorgegeben.

In vorherigen Arbeiten der Autoren [Sc23] wurden Transformationsregeln vorgestellt, die Grafsets in Modelle formuliert in Guarded Action Language (GAL) [Th16] überführen, mit dem Ziel der Modellverifikation, mittels des Model Checkers ITS-Tools [Th15]. Die Invarianten, die die Streckeninformationen abbilden, müssen zum Zeitpunkt der Erzeugung des GAL-Modells berücksichtigt werden. In der bisherigen Transformation nach GAL wird eine Komponente generiert, die jeder booleschen Eingangsvariable einen zufälligen Wert zuweist. Diese Zuweisung ist - ohne Wissen über die Strecke - nichtdeterministisch und erzeugt im resultierenden Transitionssystem alle möglichen Belegungen dieser Variablen als Folgezustände. Invarianten wie in Gleichung 1 schränken die möglichen Folgezustände ein: Jeder Folgezustand, der diese Invarianten nicht erfüllt, darf nicht erzeugt werden. Um sicherzustellen, dass diese nicht erzeugt werden, wird nach der Zuweisung der neuen Werte überprüft, ob die Invariante verletzt wird. Wenn dies der Fall ist, so werden die involvierten Eingangsvariablen so gesetzt, dass die Invariante erfüllt wird. Wenn zum Beispiel die Eingänge *Ebene_0*, *Ebene_1*, *Ebene_2* und *Ebene_3* den Wert *wahr* erhalten haben, dann genügt es all diese Variablen auf *falsch* zu setzen. Für den Fall, dass eine Invariante unerfüllbar ist, oder dass mehrere Invarianten im Widerspruch zueinander stehen, kann in GAL Gebrauch von dem Prädikat *abort* gemacht werden, welches keinen Folgezustand generiert. Das resultierende Modell ist dann allerdings der Art, dass eine in der Regel nicht wünschenswerte Verklemmung vorliegt. Dies kann anschließend mittels Model Checking überprüft werden.

2.2 Modellierung von Streckeneigenschaften mit GRAFCET

Eine weitere Anforderung an das System in Abbildung 1 kann sein, dass es in der Steuerung zu keiner Verklemmung kommen darf, also immer eine nächste Situation erreichbar ist. Dies ist nicht gegeben, wenn keine Transition mehr schalten kann. Angenommen, während der Spezifikation in Abbildung 1 entsteht ein Fehler und die Aktion an Schritt 2 setzt ebenfalls die Variable *hoch* auf *wahr*, dann ist zu beobachten, dass *Ebene_0* niemals

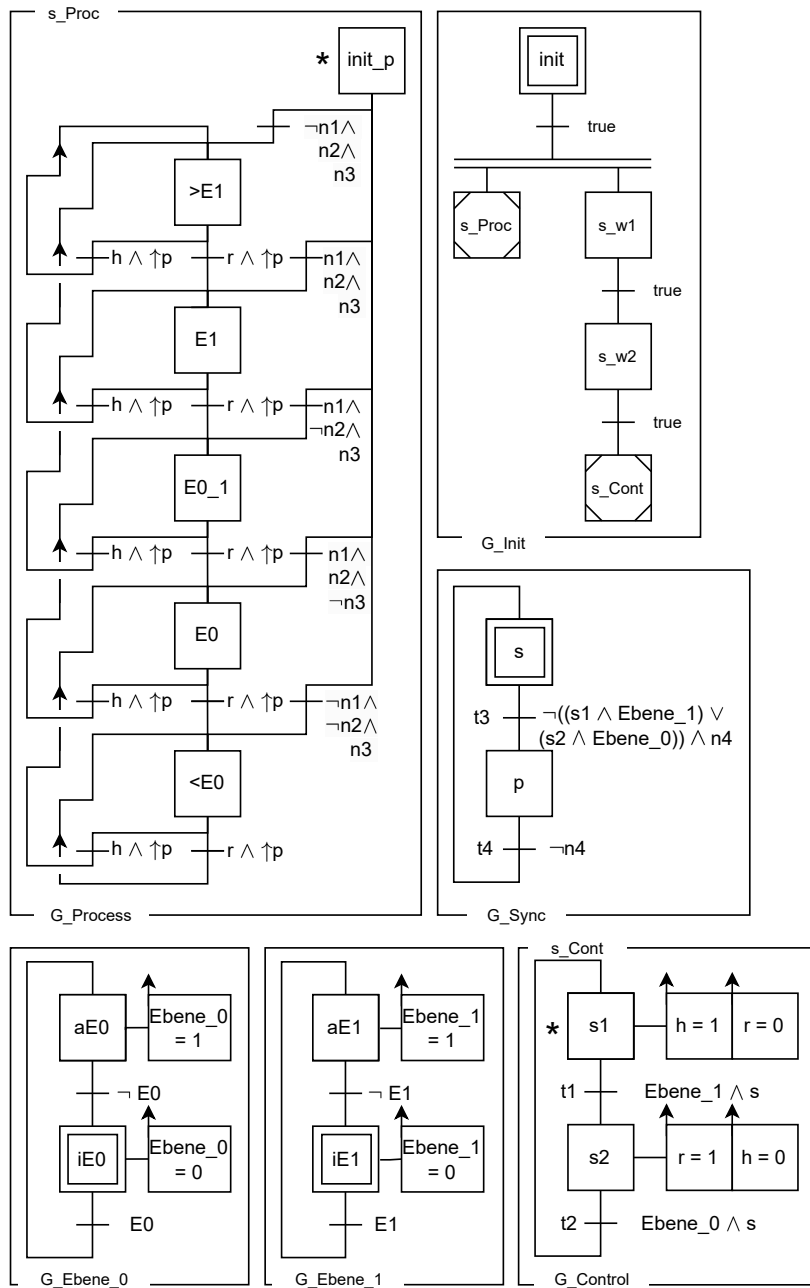


Abb. 3: In GRAFCET modelliertes kombiniertes Verhalten der in Abbildung 1 dargestellten Steuerung und Strecke mit r für runter und h für hoch.

wahr wird, da der Fahrstuhl niemals so angesteuert wird, dass er nach unten fährt. Die Invariante in Gleichung 1 ist nicht geeignet, diese Anforderung zu verifizieren, da sie keine Auskunft über den Zusammenhang zwischen Sensorik und Aktorik gibt. Um diese benötigten Informationen zu modellieren, soll in dieser Arbeit GRAFCET sowohl für die Modellierung der Steuerung als auch für die Modellierung der Steuerstrecke genutzt werden. Dies hat den Vorteil, dass der Anwender durchgängig ein Beschreibungsmittel verwenden kann. Um das kombinierte Verhalten aus Steuerung und Strecke innerhalb eines Grafsets⁴ korrekt modellieren zu können, muss zusätzlich zum (I) Steuerungscode auch das Verhalten der (II) Strecke und das (III) Zusammenspiel aus Steuerung und Strecke mit GRAFCET modelliert werden. In Abbildung 2 ist das Zusammenspiel der Komponenten dargestellt.

Für das Beispiel des Lastenfahrstuhls ist die Modellierung des (I) Steuerungscode bereits in Abbildung 1 erfolgt und wird im Folgenden noch geringfügig angepasst. Die Modellierung der Strecke (II) ist in Abbildung 3 in dem Teilgrafcet $G_Process$ dargestellt. Dabei bildet das Modell nicht das vollständige Verhalten des Lastenfahrstuhls ab, sondern nur die für die Verifikation nützlichen Informationen. Der Fahrstuhl wird mit fünf Schritten modelliert, die anzeigen, ob der Fahrstuhl von einem Sensor detektiert wird ($E1, E0$), oder ob er sich unter, zwischen, oder über den Sensoren befindet ($<E0, E0_1, >E1$). Je nachdem, ob die Steuerung die Befehle *hoch* (h) oder *runter* (r) sendet, wird ein entsprechend anderer Schritt des Strecken-Grafsets aktiviert. GRAFCET ist dazu entworfen, deterministisches Verhalten zu modellieren, was für den Entwurf von Steuerungen auch angemessen ist, da nichtdeterministisches Verhalten in Steuerungen üblicherweise unerwünscht ist. Steuerstrecken können sich aber durchaus nichtdeterministisch verhalten, beispielsweise wenn sich eine Strecke in einer nicht definierten Ausgangssituation befindet. Um dies abzubilden, wird ein einzelner Anfangsschritt $init_p$ definiert, der einen der Schritte der Strecke über eine Transition aktiviert. Dies wird zum einen erreicht, da sich die Transitionsbedingungen dieser Transitionen gegenseitig ausschließen und so immer nur einer der Schritte des Strecken-Grafsets aktiviert wird. Dass während der Verifikation im Zustandsraum alle möglichen initialen Situationen der Strecke abgedeckt sind, wird dadurch sichergestellt, dass die Transitionsbedingungen, die auf den Initialschritt folgen, mit als nichtdeterministisch modellierten Eingangsvariablen belegt sind.

Für die Modellierung des Zusammenspiels (III) zwischen Strecke und Steuerung wird Teilgrafcet G_Sync eingefügt. Dieser hat zwei Schritte s und p , die dafür sorgen, dass sich entweder die Strecke oder die Steuerung entwickeln kann. Auf Seiten der Steuerung garantiert dies, dass keine Zwischenwerte während der Abarbeitung des Steuerungscode an die Strecke gelangen (wie in Abbildung 2 dargestellt). Die Transitionsbedingungen in $G_Control$ werden dann um das Konjunkt der Schrittvariable s erweitert. Auf Seiten der Strecke kann umgesetzt werden, dass immer nur ein Zustandsübergang zurzeit stattfinden kann, auf den die Steuerung dann wieder reagiert, indem die Transitionsbedingungen in $G_Process$ um das Konjunkt $\uparrow p$ ⁵ erweitert werden. Dieses Verhalten modelliert dann

4 $GRAFCET$ bezeichnet das Beschreibungsmittel und $Grafcet$ eine Instanz des Beschreibungsmittels.

5 Der Operator $\uparrow x$ wird genau dann *wahr*, wenn eine steigende Flanke im Signal x auftritt.

eine realitätsnahe Synchronisation zwischen Steuerung und Strecke. Ein Wechsel von der Abarbeitung der Steuerung zur Abarbeitung der Strecke und umgekehrt erfolgt immer dann, wenn *G_Control* oder *G_Process* eine stabile Situation erreicht haben. Dies wird über die Transitionsbedingungen in *G_Sync* sichergestellt.

Zudem sind noch syntaktische Regeln von GRAFCET zu berücksichtigen. Werden Steuerung und Strecke in einer GRAFCET-Instanz abgebildet, so können Eingangs- und Ausgangsvariablen in GRAFCET, die üblicherweise zu Sensorik und Aktorik korrespondieren, nicht wie gewohnt verwendet werden, da die Eingangsvariablen der Steuerung die Ausgangsvariablen der Strecke sind (siehe Abbildung 2). Die Variablen, mit denen Steuerung und Strecke kommunizieren, werden daher in dem hier vorgestellten Ansatz in interne Variablen umgewandelt. Dies hat die Konsequenz, dass kontinuierlich wirkende Aktionen nicht mehr verwendet werden können, da diese syntaktisch nur Ausgangsvariablen schreiben können. Jede Variable, die mit kontinuierlich wirkenden Aktionen geschrieben wird, wird daher durch ein Konstrukt substituiert, das in Abbildung 3 für die Variablen *Ebene_0* und *Ebene_1* mit *G_Ebene_0* und *G_Ebene_1* dargestellt ist. Dies hat nun die Folge, dass es in der internen Abarbeitung von GRAFCET zwei sogenannte *transiente* Abläufe dauert, bis die Variablen (hier *Ebene_0* und *Ebene_1*) korrekt beschrieben werden. Daher wird zur korrekten Initialisierung das Konstrukt *G_Init* eingeführt, welches *G_Prozess* und *G_Control* aktiviert, wenn die Schritte *s_Proc* und *s_Cont* aktiv sind.

Das dargestellte Beispiel bildet eine normkonforme GRAFCET-Instanz, die das kombinierte Verhalten von Steuerung und Strecke korrekt wiedergibt, wobei die Modellierung von physikalischer Zeit allerdings nicht betrachtet wird.

Es ist festzustellen, dass die Modellierung der in Abbildung 3 dargestellten Konstrukte einen erheblichen Modellierungsaufwand bedeuten würde, welcher aber mittels einer Modelltransformation automatisiert werden kann. Dazu schlagen wir in diesem Beitrag eine Erweiterung von GRAFCET vor, die wir in Anlehnung an Frey et al. [FL98] als *Prozesstechnisch Interpretierter GRAFCET* bezeichnen und die eine auf die Strecke ausgerichtete Syntax und Semantik enthält, wobei die Semantik im Wesentlichen über die beispielhaft vorgestellte Modelltransformation definiert ist. Dieser *Prozesstechnisch Interpretierte GRAFCET* kann dann über eine Modelltransformation in einen regulären GRAFCET nach [DIN EN 60848] transformiert werden, wie in Abbildung 3 für das Beispiel gezeigt, sodass sich der Nutzer nicht mit der Modellierung des SPS-Steuerzyklus beschäftigen muss. Formal handelt es sich um eine Erweiterung des GRAFCET-Metamodells wie in [Sc23] vorgestellt, dem eine Klasse für den *Prozesstechnisch Interpretierten GRAFCET* hinzugefügt wird, die einzig anzeigt, dass in ihr eine Strecke modelliert wird, sodass eine Transformation möglich wird. Die entstehenden Grafcets können im Anschluss mit der bestehenden, in [Sc23] vorgestellten, Werkzeugkette verifiziert werden.

Zudem ist in Bezug auf den Modellierungsaufwand zu beachten, dass es sich bei den für die Verifikation interessanten Streckeninformationen um recht einfache Zusammenhänge zwischen Sensorvariablen handeln kann, die sich aus der Streckengeometrie ergeben. Dies

ist beispielsweise bei Endlagensensoren von Kolben, bei der Detektion mehrerer Füllstände eines Tanks, oder bei dem gezeigten Fahrstuhl der Fall. All diese Streckeninformationen sind in GRAFCET ähnlich modellierbar, sodass eine Art Bibliothek definiert werden kann, aus der der Anwender das entsprechende GRAFCET-Konstrukt auswählt und nur noch die Variablenbenennung vornehmen muss. Die entsprechende Transformation des *Prozesstechnisch Interpretierten GRAFCET* erfolgt dann automatisiert in normkonformen GRAFCET.

3 Diskussion

Eine fehlende Modellierung von einschränkenden Streckeninformationen führt zur größtmöglichen Überapproximation des Streckenverhaltens, was für falsche Alarmer in der Verifikation sorgen kann. Mit beiden vorgestellten Ansätzen wird diese Überapproximation verringert. Für Invarianten ist die Überapproximation aber höher, als bei der Modellierung mittels GRAFCET, da die Invarianten keine Zustände abbilden können. Es konnte beispielhaft gezeigt werden, dass die Zuhilfenahme von Invarianten dazu führt, genauere Verifikationsergebnisse zu erzielen, wenn transiente Abläufe oder die Exklusivität von Alternativverzweigungen analysiert werden. Mit der Streckenmodellierung als Grafcet konnten diese beiden Anforderungen ebenfalls genauer verifiziert werden, und darüber hinaus, ob es im kombinierten Verhalten aus Steuerung und Strecke zu Verklemmungen kommt. Weiter kann die Modellierung mittels Grafcet dazu genutzt werden, die Strecke betreffende Anforderungen zu verifizieren, wie beispielsweise die Erreichbarkeit von Zuständen in der Strecke. Für den Lastenfahrstuhl wäre eine solche Anforderung, dass der Fahrstuhl die obere/untere Endlage nicht über-/unterschreiten darf. In Bezug auf den Modellierungsaufwand sind die vorgestellten Invarianten durch den Anwender vergleichsweise einfach in boolescher Logik zu formulieren. Die Beschreibung der Strecke mit GRAFCET ist für einfache Zusammenhänge unter Nutzung der Modelltransformation kein großer zusätzlicher Modellierungsaufwand, da für ähnliche Steuerungsgeometrien die Grafcets auch wiederverwendet werden können. Sollen allerdings komplexere Streckenmodelle umgesetzt werden, ist auch ein höherer Modellierungsaufwand nötig.

Die Modellierung der Strecke als normkonformen GRAFCET hat hingegen den Vorteil, dass sie unabhängig von der Implementierung der Verifikationsmethode ist, sodass jedes bestehende GRAFCET-Verifikationswerkzeug auch auf einen Grafcet angewendet werden kann, welcher Streckeninformationen enthält. Für die Invarianten hingegen muss beim Model Checking während der Generierung des Zustandsraums sichergestellt werden, dass sie auch gelten, was sich nur werkzeugspezifisch umsetzen lässt.

Für das Ziel der Zustandsraumverkleinerung ist die Formulierung von Invarianten vorteilhaft, da sie die Kombinatorik der Eingangsvariablen tendenziell einschränken. Für die Modellierung von Streckeninformation in GRAFCET ist dies allerdings nicht der Fall. Für das gezeigte Beispiel vergrößert sich der Zustandsraum um ein Vielfaches.

4 Zusammenfassung und Ausblick

Es wurden zwei Ansätze zur Modellierung von Streckeninformation vorgestellt: Invarianten und die Modellierung mittels GRAFCET. Für Invarianten wurde dargelegt, wie diese beispielhaft formuliert und dann bei der Erstellung des Zustandsraums in GAL für das Model Checking berücksichtigt werden können. Für die Modellierung mittels GRAFCET wurde dargelegt, wie neben den Streckeninformationen noch weitere Konstrukte modelliert werden müssen, sodass das kombinierte Verhalten aus Steuerung und Strecke sichergestellt werden kann. Dies ermöglicht die Modellierung des SPS-Zyklus und die Einhaltung oder auch Verletzung von Echtzeitkriterien. Um dies mittels einer Modelltransformation zu unterstützen, wurde der *Prozesstechnisch Interpretierte GRAFCET* eingeführt, in dem das Streckenverhalten modelliert wird. In der Diskussion wurde gezeigt, dass die Modellierung mittels GRAFCET zu bevorzugen ist, wenn die höchste Genauigkeit der Verifikationsergebnisse erzielt werden soll. Invarianten eignen sich, wenn der Modellierungsaufwand gering gehalten werden soll, oder wenn eine Reduzierung des Zustandsraums erzielt werden soll.

Für zukünftige Arbeiten ist es sinnvoll, die Verifikation des Zeitverhaltens von GRAFCET zu untersuchen, was in [Sc23] nicht berücksichtigt wurde. Dies ermöglicht es während der Verifikation auch das Zeitverhalten der Strecke zu berücksichtigen. Auch kann es sinnvoll sein, die Streckenmodellierung basierend auf Zuständen in dem Formalismus durchzuführen, der für das Model Checking genutzt wird (in diesem Fall GAL), um eine Begrenzung der Zustandsraumgröße zu erzielen. Auch kann es sinnvoll sein, die Streckenmodelle aus dem vorgestellten *Prozesstechnisch Interpretierten GRAFCET* direkt in GAL zu transformieren, um eine bessere Kontrolle über die Zustandsgröße zu erhalten und diese möglicherweise stärker zu begrenzen als in dem vorgestellten Ansatz.

Literatur

- [DIN EN 60848] DIN Deutsches Institut für Normung e. V.: GRAFCET, Spezifikations-sprache für Funktionspläne der Ablaufsteuerung, DIN EN 60848, 2014-12.
- [FL98] Frey, G.; Litz, L.: Verification and validation of control algorithms by coupling of interpreted Petri nets. In: SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics. Bd. 1, 7–12 vol.1, 1998.
- [Ga19] Galvão, J.; Oliveira, C.; Lopes, H.; Tiainen, L.: Formal Verification: Focused on the Verification Using a Plant Model. In (Machado, J.; Soares, F.; Veiga, G., Hrsg.): Innovation, Engineering and Entrepreneurship. Springer International Publishing, Cham, S. 124–131, 2019.
- [Ju19] Julius, R.; Fink, V.; Uelzen, S.; Fay, A.: Konzept zur bidirektionalen Transformation zwischen GRAFCET-Spezifikationen und IEC 61131-3 Steuerungscode. at - Automatisierungstechnik 67/3, S. 208–217, 2019.

- [Lu20] Lunze, J.: Automatisierungstechnik - Methoden für die Überwachung und Steuerung kontinuierlicher und ereignisdiskreter Systeme. De Gruyter, Berlin, Boston, 2020, ISBN: 9783110465624.
- [Sc23] Schnakenbeck, A.; Mroß, R.; Völker, M.; Kowalewski, S.; Fay, A.: Transformation von GRAFCET in GAL auf Basis eines ausführlichen Metamodells zur Verifikation von Entwurfsfehlern. at - Automatisierungstechnik 71/1, S. 56–68, 2023.
- [SF14] Schumacher, F.; Fay, A.: Petrinetzmodell für die Formalisierung von GRAFCET-Spezifikationen. at - Automatisierungstechnik 62/6, S. 385–393, 2014.
- [Th15] Thierry-Mieg, Y.: Symbolic Model-Checking Using ITS-Tools. In (Baier, C.; Tinelli, C., Hrsg.): Tools and Algorithms for the Construction and Analysis of Systems. Springer Berlin Heidelberg, Berlin, Heidelberg, S. 231–237, 2015.
- [Th16] Thierry Mieg, Y.: From Symbolic Verification To Domain Specific Languages, Diss., Sorbonne Université , UPMC ; Laboratoire d’informatique de Paris 6 [LIP6], 2016.
- [Vo14] Vogel-Heuser, B.; Diedrich, C.; Fay, A.; Jeschke, S.; Kowalewski, S.; Wollschlaeger, M.; Göhner, P.: Challenges for Software Engineering in Automation. Journal of Software Engineering and Applications 7/, S. 440–451, 2014.

„Ulrich Jumar, Christian Diedrich (Hrsg.):
EKA 2024 - Entwurf komplexer Automatisierungssysteme, 18. Fachtagung“

Konzepte zur Realisierung proaktiver Verwaltungsschalen und deren Kommunikation

Untersuchung der Einsatzmöglichkeiten von Events und Operationen für den Nachrichtenaustausch

Melanie Stolze¹, Gustavo Cainelli², Christian Kosel³, Alexander Belyaev⁴, Christian
Diedrich⁵

Abstract: Sowohl die passive als auch die reaktive Verwaltungsschale (VWS) halten immer mehr Einzug in die Praxis. Ebenso wird der Einsatz und die Entwicklung proaktiver VWS, deren Umsetzung zum aktuellen Zeitpunkt noch nicht spezifiziert ist, in aktuellen Forschungsprojekten evaluiert. Das Ziel ist, einen Überblick über die derzeit bestehenden Möglichkeiten für die Implementierung einer proaktiven VWS zu geben und somit einen Teil zu deren Spezifizierung beizutragen. Die drei Schwerpunkte liegen dabei auf der Realisierung der Architektur und Schnittstelle, sowie auf der flexiblen Integration verschiedener Interaktionsmuster für proaktive VWS.

Keywords: Verwaltungsschale, Interaktion, I4.0-Sprache, Digitaler Zwilling

1 Motivation und Problemstellung

Mittlerweile hält die Verwaltungsschale (VWS) immer mehr Einzug in die Praxis. Die VWS kann dabei in 3 Arten (passiv, reaktiv und proaktiv) realisiert werden, wie in Abbildung 1 dargestellt ist. Komponentenhersteller bieten ihren Kunden bereits sowohl passive als auch reaktive VWS über ihre Weboberflächen an [WI24], [WA24]. Ebenso wird der Einsatz und die Weiterentwicklung der VWS in Forschungsprojekten der Konjunktur- und Zukunftspakete KoPa35c [BA24], Gaia-X [BM24] und Manufacturing-X [PI24] des Bundesministeriums für Wirtschaft und Klimaschutz vorangetrieben. Forschungsprojekte aus KoPa35c, wie „Verwaltungsschale für den Leitungssatz“ VWS4LS [AR24], fokussieren die Nutzung passiver, reaktiver als auch proaktiver VWS, während sich die anderen genannten Zukunftspakete mit Datenökosystemen beschäftigen. Arbeiten wie [Ju23] zeigen bereits sehr gut, wie eine Brücke zwischen den in den geförderten Zukunftspaketen entwickelten Ergebnissen geschlagen werden kann. Dafür

¹ Institut für Automation und Kommunikation e.V., IKT & Automation, Werner-Heisenberg-Str. 1, 39106 Magdeburg, melanie.stolze@ifak.eu

² Institut für Automation und Kommunikation e.V., IKT & Automation, Werner-Heisenberg-Str. 1, 39106 Magdeburg, gustavo.cainelli@ifak.eu

³ Arena 2036 e.V., Pfaffenwaldring 19, 70569 Stuttgart, christian.kosel@arena2036.de

⁴ Otto von Guericke Universität Magdeburg, 39106 Magdeburg, alexander.belyaev@ovgu.de

⁵ Otto von Guericke Universität Magdeburg, 39106 Magdeburg, christian.diedrich@ovgu.de

verbindet [Ju23] die Eigenschaften proaktiver VWS und ihrer I4.0-Sprache mit denen in einem Datenökosystem vorhandenen Konnektoren zum unternehmensübergreifenden Datenaustausch.

Diese Arbeit legt den Fokus jedoch auf die Implementierungsmöglichkeiten der proaktiven VWS und der I4.0-Sprache und präsentiert unter anderem Forschungsergebnisse aus dem Projekt VWS4LS [AR24], in dem die Leitungssatzentwicklung und -fertigung mit Hilfe von VWS effektiver gestaltet werden soll. In der Leitungssatzfertigung steigt durch heterogene Maschinenparks und Produkte der Losgröße 1 der Zeitaufwand für die Aufgaben in der Fertigungsplanung wesentlich an. Unregelmäßigkeiten im Produktionsablauf, wie z. B. ungeplante Instandsetzungen, Materialflussschwierigkeiten oder kurzfristige Bedarfsschwankungen führen zu neuen Ablaufplänen und erschweren die Planung. Zur Lösung des Problems wird in [AR24] der Einsatz proaktiver VWS evaluiert, die die Fertigungsplaner in ihren Tätigkeiten unterstützen sollen, teilautomatisiert geeignete Ressourcen für ein zu fertigendes Produkt zu finden. Bei dieser Art von VWS erfolgt eine eigenständige Kommunikation zwischen VWS mittels der in der VDI/VDE 2193-1 spezifizierten I4.0-Sprache [VD20a]. Allein mit [VD20a] ist die proaktive VWS jedoch noch nicht spezifiziert und somit existieren verschiedene Lösungen für deren reale Umsetzung, die in diesem Bericht zusammengefasst werden. Beginnend führt Kapitel 0 in Architekturvarianten ein, die sowohl aus der Literatur sowie eigenen im Projekt VWS4LS entwickelten Konzepten bekannt sind. Kapitel 3 beleuchtet die Einsatzmöglichkeiten der im VWS-Metamodell [ID23a] spezifizierten Elemente „Operation“ und „Event“ und deren Vor- und Nachteile, um die proaktive VWS mit Kommunikationsfähigkeiten auszustatten. Dabei wird auch bezüglich der spezifizierten Schnittstelle der VWS [ID23b] untersucht, inwieweit sie die Übergabe von I4.0-Nachrichten gemäß der VDI/VDE 2193-1 ermöglicht. Zum Ende wird in Kapitel 0 die Integration verschiedener Interaktionsmuster in den vorgeschlagenen Architekturen diskutiert.

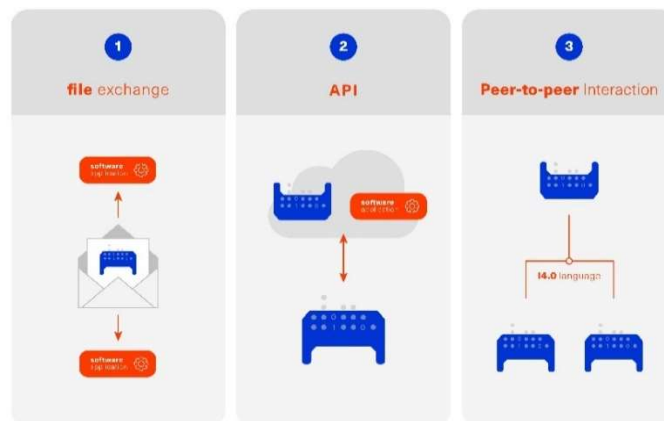


Abbildung 1: Die drei Arten von Verwaltungsschalen: passiv (1), reaktiv (2) und proaktiv (3) [ID23a]

2 Architekturen für die proaktive Verwaltungsschale

In [BD19] wird ein abstraktes Konzept für die Architektur einer proaktiven VWS vorgestellt. Dort setzt sich die proaktive VWS aus einem aktiven und passiven Part zusammen. Während der passive Part die Daten des Assets über die VWS-Teilmodelle in Form eines Datenmodells bereitstellt, übernimmt der aktive Part die Ausführung von Algorithmen, deren Orchestrierung, sowie die Aussendung und das Empfangen von I4.0-Nachrichten. [BD19]

Zu sehen ist das abstrakte Konzept in Abbildung 2. Die Komponenten des aktiven Parts werden in [BD19] wie folgt beschrieben werden:

- Ein Interaktionsmanager, der die Zustandsmaschinen für ausgewählte Interaktionsmuster implementiert,
- mehrere Algorithmen, die Berechnungen in einzelnen Zuständen der Zustandsmaschine mit Hilfe der Daten aus dem passiven Part übernehmen,
- ein Komponentenmanager, der die Kommunikation zwischen dem Interaktionsmanager und den Algorithmen orchestriert, sowie
- ein Messenger als Kommunikationsschnittstelle, der zum einen die Nachrichten von dem Interaktionsmanager entgegennimmt, diese gemäß der I4.0-Sprache aufbereitet und nach außen versendet und ebenso I4.0-Nachrichten von außen entgegennimmt und wiederum an den Interaktionsmanager weitergibt.

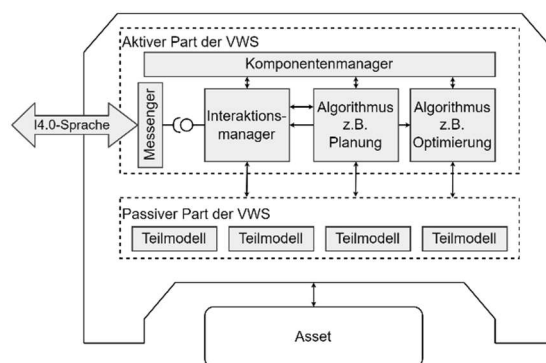


Abbildung 2: Abstrakte Architektur für eine Typ 3 VWS (in Anlehnung an [BD19])

Für die folgenden Abschnitte wird folgender Grundsatz für die Implementierung einer proaktiven VWS angenommen. Der passive Part wird durch einen VWS-Server realisiert, der zum einen eine Schnittstelle für die Assetanbindung und zum anderen eine Schnittstelle zu anderen Applikationen bereitstellt, der VWS-API (Application Programming Interface) [ID23b]. Die Realisierung des Servers kann mit Hilfe des Eclipse BaSyx™ Frameworks [EB24] geschehen. Die bis zu diesem Punkt implementierten Funktionalitäten repräsentieren eine reaktive VWS. Dabei kann ein Server nur eine

(genannt VWS-Server) oder aber mehrere VWS (genannt VWS-Repository) verwalten. Damit aus einer reaktiven VWS eine proaktive VWS wird, muss diese in die Lage versetzt werden, I4.0-Nachrichten zu verschicken, wie es in Abbildung 1 unter Nummer 3 und in [PI20] definiert ist. Dazu werden VWS-Elemente in der proaktiven VWS modelliert, die den späteren Austausch von I4.0-Nachrichten ermöglichen. Diese Elemente werden für die folgenden Abschnitte in jeder proaktiven VWS als gegeben betrachtet und in Kapitel 3 näher erläutert. Auf dem beschriebenen Grundsatz basierend werden in den nächsten Abschnitten drei Arten von konzeptionellen Architekturen gezeigt, mit denen eine proaktive VWS implementiert werden kann.

2.1 Monolithischer Ansatz

Bei dem monolithischen Ansatz ist laut [Ge15] im Allgemeinen eine einzige zentrale Applikation zur Koordinierung und Steuerung aller Funktionen und Prozesse zuständig, wobei die Applikation konzeptionell als eine zusammenhängende und unteilbare Einheit behandelt wird. Implementierungsmäßig können die zentral gesteuerten Funktionen jedoch auf mehrere Prozesse und Hardwarekomponenten verteilt sein. Für die Implementierung einer proaktiven VWS bedeutet dies, dass der passive Part mit Hilfe des VWS-Servers/-Repositories bereitgestellt wird. Zudem übernimmt in diesem Fall der VWS-Server/-Repository auch die Realisierung der Algorithmen. Die Algorithmen werden in Form von VWS-Services oder Teilmodell-Services programmiert, die beim Start des Servers eingelesen und mit einem eingehenden Trigger durch den Server ausgeführt [Gr22] werden. Aufgrund der laufenden Entwicklungsarbeiten am Eclipse BaSyx™ Framework [EB24] wurden in [Gr22] die Services noch als „Lambda Funktionen“ bezeichnet. Abbildung 3 zeigt anhand eines VWS-Servers/-Repositories die beiden Möglichkeiten zur Umsetzung einer proaktiven VWS. Die Anteile, aus denen sich eine proaktive VWS zusammensetzt sind in allen folgenden Architekturabbildungen rot markiert. Während der VWS-Server als eine Einheit eine proaktive VWS bildet, setzt sich diese in einem VWS-Repository aus Einem von mehreren aktiven und passiven Parts zusammen sowie dem aktiven Part mit Algorithmen, der auf alle VWS angewendet werden kann. Es ist ersichtlich, dass der Programmcode bei Nutzung mehrerer VWS-Server dupliziert wird. Dies erschwert mit zunehmender Komplexität den Wartungsaufwand der Applikationen. Um dies zu umgehen, können mit Hilfe von VWS-Repositories und durch die spezifizierten VWS-Teilmodelle, Algorithmen generisch programmiert und zur Laufzeit für jede VWS individuell ausgeführt werden. Jedoch duplizieren sich auch hier die Algorithmen, wenn man mehr als ein VWS-Repository instanziiert. In [Gr22] wurde beschrieben, dass ein solcher Ansatz bei kleinen Anwendungen noch einfach zu handhaben ist. Für komplexere proaktive VWS, deren Funktionalitäten sich über die Lebensdauer der VWS erweitern, ist dieser Ansatz laut [Gr22] jedoch nicht zu empfehlen. Ein Grund dafür kann sein, dass neu entwickelte Algorithmen unter der Beachtung von Abhängigkeiten in bereits bestehende Funktionen

und Komponenten integriert werden müssen. Mit einem wachsenden Gesamtsystem, steigt auch die Anzahl der Abhängigkeiten.

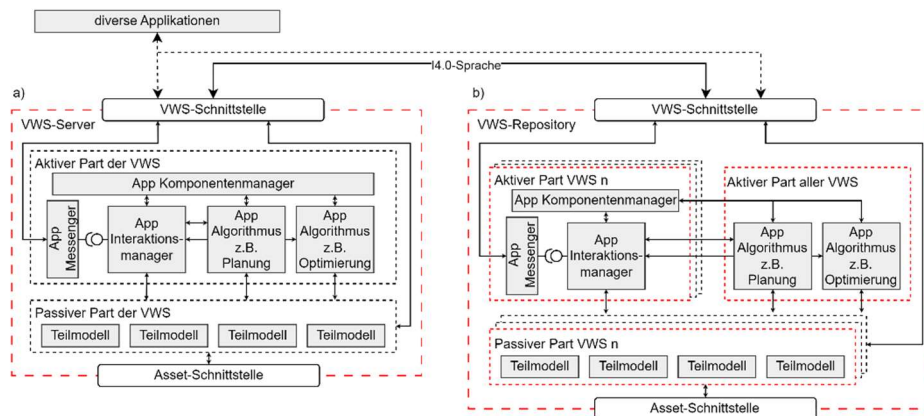


Abbildung 3: Proaktive VWS, realisiert mit dem monolithischen Ansatz. a) Verwendung von VWS-Server, b) Verwendung von VWS-Repository

Die in Abbildung 3 dargestellte VWS-Schnittstelle des VWS-Servers/-Repositories, realisiert in dieser Arbeit mehrere Aufgaben. Zum einen stellt sie die von der IDTA spezifizierte VWS-API [ID23b] bereit, um Daten aus der VWS über http/REST (Hypertext Transfer Protocol / Representational State Transfer) zu lesen und zu schreiben. Diese Kommunikation ist durch gestrichelte Pfeile gekennzeichnet. Zum anderen ermöglicht die Schnittstelle mit Hilfe des App Messengers das aktive Versenden und Empfangen von I4.0-Nachrichten (durchgezogene Pfeile).

2.2 Agentenbasierter Ansatz

Der agentenbasierte Ansatz ist durch die Präsenz von autonom agierenden Agenten charakterisiert. Jeder Agent beinhaltet die Fähigkeiten, Schlussfolgerungen zu ziehen, Entscheidungen zu treffen und Informationen auszusenden, um z. B. in Interaktion mit anderen Agenten gemeinsame Ziele zu erreichen. Entgegen dem monolithischen Ansatz bietet dieser mehr Flexibilität, Dezentralisierung und Ausfallsicherheit für ein System. Jedoch ist das Management der miteinander interagierenden Agenten komplex, vor allem in verteilten Systemen. Zudem steigen mit erhöhtem Kommunikationsbedarf auch die Latenzzeiten im System an. [SLP22], [CF21] Angewandt auf die proaktive VWS kann eine agentenbasierte Architektur wie in Abbildung 4 realisiert werden. Dabei sind die Komponenten, die von Agenten repräsentiert werden durch ein entsprechendes Agentensymbol markiert. Zudem gibt es keinen Komponentenmanager mehr, der die Agenten orchestrieren würde.

In der Literatur gibt es durchgehend agentenbasierte Ansätze zur Realisierung einer proaktiven VWS. Dabei gibt es Arbeiten wie [Is23], [SLP22], [SPL23], [Ca23] und [Gr22], die deutlich von der Implementierung einer proaktiven VWS sprechen. Andere Arbeiten wie [YH20] und [Ju23] sprechen von reaktiven VWS. Die zuletzt genannten Arbeiten realisieren jedoch durch die Zusammenschaltung der reaktiven VWS mit anderen Applikationen indirekt Teile einer proaktiven VWS. Nur wenige der genannten Arbeiten beziehen jedoch die I4.0-Sprache in ihr Konzept mit ein, die als ein Hauptmerkmal der proaktiven VWS in [PI20] definiert ist. Durch die Verlagerung des Interaktionsmanagers in einen Agenten, können VWS flexibel an Interaktionsmuster angebunden bzw. wieder entfernt werden, wenn sie die für die Interaktion notwendigen Daten in den VWS-Teilmodellen abbilden.

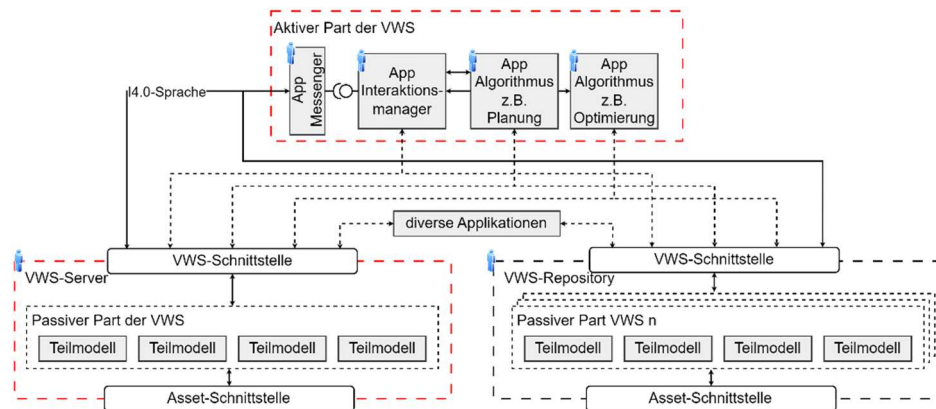


Abbildung 4: Proaktive VWS, realisiert mit dem agentenbasierten Ansatz

In [SLP22] und [SPL23] wird eine Methodologie vorgeschlagen, VWS mit Multi-Agenten-System (MAS) zu verbinden. Darin wird eine Richtlinie beschrieben, die die Entwicklung von proaktiven VWS mit Hilfe der Vorteile von MAS forciert. In [Gr22] ist der aktive Part mittels externer Applikationen realisiert, die die für eine Interaktion notwendigen Informationen aus dem VWS-Server/-Repository weiterverarbeiten. Wiederum eine andere Variante wird in [Is23] beschrieben, wo ganze VWS oder einzelne Teilmodelle mit Agenten ausgestattet werden. Zusätzlich wird in [Ca23] und [YH20] der VWS-Server nicht mit Hilfe des Eclipse BaSyx™ Frameworks [EB24] implementiert, sondern über einen OPC UA Server, der das VWS Modell verwaltet. Der aktive Part wird in diesem Fall durch einen OPC UA Client übernommen, der die Businesslogik implementiert.

Der Vorteil der agentenbasierten Architektur ist, dass einzelne Agenten für mehrere VWS verwendet werden können und somit nicht mehrfach entwickelt werden müssen. Jedoch entsteht durch die Anwendung eines Agenten auf viele VWS ein zentralistischer Ansatz, der bei Ausfall eines Agenten eine Beeinträchtigung vieler VWS nach sich zieht. Zudem benötigen die separierten Agenten ein integriertes Management.

2.3 Orchestrierender Ansatz

In einem orchestrierenden Ansatz wird eine zentrale Entität verwendet, auch Orchestrator genannt, der die Kombination von Komponenten in einem Interaktionsmuster übernimmt. Da bisher noch keine Publikationen bzgl. eines solchen Ansatzes zur Implementierung einer proaktiven VWS existieren, hat das Projekt VWS4LS [AR24] ein Konzept für diesen Ansatz entwickelt. Dabei wird der Orchestrator durch ein Workflow Management System (WMS) repräsentiert, das sowohl den Messenger als auch den Komponenten- und Interaktionsmanager in einem abbildet. Das WMS implementiert hierbei die Interaktionsmuster und ermöglicht wie der agentenbasierte Ansatz die flexible Einbindung

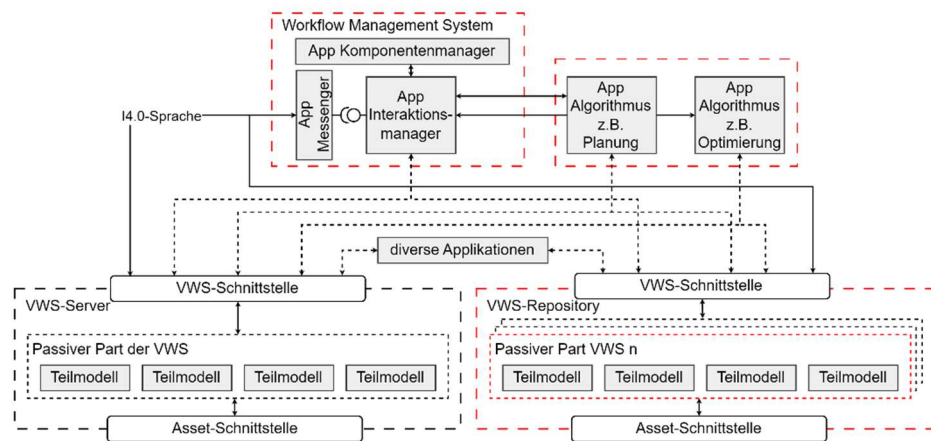


Abbildung 5: Proaktive VWS, realisiert mit dem orchestrierenden Ansatz

von neuen VWS in die Interaktionsmuster. Die Algorithmen werden auf andere Applikationen verteilt, wie z. B. einem Produktionsleitsystem. Der passive Part ist wie bei allen anderen Ansätzen durch einen VWS-Server/-Repository implementiert. Dieser Ansatz liefert einige Vorteile, z. B. die vom WMS unterstützte horizontale Skalierbarkeit, die es ermöglicht, VWS nach Bedarf hinzuzufügen oder zu entfernen. Zudem gibt es etablierte Softwaretools für die Erstellung von Workflows, die hier zum Einsatz kommen können. Unter anderem implementieren bereits erste Softwareunternehmen die VWS-API in ihre WMS-Lösungen. Andererseits besteht hier ähnlich dem agentenbasierten Ansatz die Gefahr, dass bei Ausfall des WMS alle an das WMS gekoppelten VWS ihre Funktion als proaktive VWS verlieren.

2.4 Fazit

Jede der vorgestellten Architekturen besitzt ihre Vor- und Nachteile und es kann nicht eine Architektur als die beste Lösung forciert werden. Vielmehr muss abgestimmt auf den Anwendungsfall die bestmögliche Architektur ausgewählt werden. Beginnend mit dem

monolithischen Ansatz, bietet er für leichtgewichtige Funktionen einen guten Lösungsansatz. Durch die Verwendung eines VWS-Repositories ist es möglich, mehrfach anwendbare Algorithmen nur einmal zu integrieren. Trotzdem können bei der Verwendung mehrerer VWS-Repositories, in denen die gleichen Algorithmen benötigt werden, eine Redundanz von Programmcode entstehen, die einen erhöhten Wartungsaufwand fordert. Bei Nutzung mehrerer dezentral gehaltener VWS-Repositories behalten hingegen bei Ausfall eines VWS-Repositories die proaktiven VWS in allen anderen VWS-Repositories ihre Funktionalitäten. Dieser Vorteil ist bei dem agentenbasierten und orchestrierenden Ansatz nicht per se gegeben. Wird für einen Algorithmus bzw. eine Funktionalität, z. B. nur ein Agent integriert, verlieren alle proaktiven VWS diese Funktionalität, sobald der Agent ausfällt. Dies ähnelt eher einem zentralisierten Ansatz. Vielmehr müsste die Integration eines Agenten-Typen forciert werden, von denen Agenten-Instanzen abgeleitet werden, die jeweils nur für eine Gruppe oder eine einzelne proaktive VWS verantwortlich sind. Der Vorteil wiederum ist, dass der Programmcode für die Agenten nur einmalig entwickelt werden muss und somit der Wartungsaufwand geringer als bei dem monolithischen Ansatz ist. Jedoch steigt mit mehreren Agenten auch die Kommunikationslast in einem Netz.

Das Management aller Komponenten wird bei dem monolithischen Ansatz von einer zentralen Applikation übernommen. Das hat zum Nachteil, dass bei Ausfall dieser Applikation alle Prozesse zum Erliegen kommen. Der agentenbasierte Ansatz ist demgegenüber vorteilhafter, da die Agenten sich jeweils selbst managen. Jedoch ist die übergeordnete Organisation aller Agenten nicht leicht. Die Vor- und Nachteile beider Ansätze werden durch den orchestrierenden Ansatz verbunden. Zum einen gibt es das WMS als zentrale Instanz, dass die verteilten Komponenten organisiert. Fällt das WMS jedoch aus, verlieren die proaktiven VWS ihre Funktionalität. Um diesen Nachteil zu mindern, müssten mehrere verteilte und voneinander unabhängige WMS genutzt werden.

3 Schnittstellen zum Austausch von I4.0-Nachrichten

Im vorhergehenden Kapitel wurden Architekturen zur Implementierung proaktiver VWS beschrieben. Nun folgt ein Einblick in die Implementierungsdetails des aktiven und passiven Parts, um die autonome Interaktion zwischen proaktiven VWS mit Hilfe der I4.0-Sprache mit den zurzeit spezifizierten Mitteln zu ermöglichen.

3.1 Wahl des Kommunikationsprotokolls

Für die Realisierung des Nachrichtenaustauschs werden die Kommunikationsprotokolle MQTT (Message Queuing Telemetry Transport) und http/REST betrachtet.

MQTT ermöglicht den Austausch von Topics mit flexiblem Datenrumpf. Für die Übertragung der I4.0-Nachricht sind in [BD19] bereits Beispiele für die Implementierung

des Nachrichtenrumpfs in JSON-Formaten beschrieben, die in [Gr22] und [Is23] wiederverwendet wurden. Problematisch ist bisher jedoch, dass es keine VWS-API Spezifikation für MQTT gibt. Somit besteht die Gefahr, proprietäre Schnittstellen zu entwickeln, die dann nicht miteinander kompatibel sind. Die REST-Schnittstelle der VWS hingegen ist in [ID23b] spezifiziert und die Struktur der Datenrumpfe für die Schnittstellenaufrufe sind fest vorgeschrieben. Jedoch stellt die REST-API keine Operationen bereit, die speziell für die Übergabe von I4.0-Nachrichten gedacht sind. Deshalb werden im Folgenden andere Lösungen betrachtet, die unter anderem die Interoperabilität proaktiver VWS gewährleisten sollen.

3.2 Nutzung von Events und Operationen für die Übergabe der I4.0-Nachricht

In einer proaktiven VWS muss es Funktionen geben, die die Übertragung von I4.0-Nachrichten in einer Interaktion zwischen mehr als zwei proaktiven VWS ermöglichen. Diese Funktionen werden von dem Messenger als aktive Komponente übernommen. Zur Realisierung der Funktionen gibt es die VWS-Elemente „Event“ und „Operation“. Das Event wird in der Spezifikation des VWS-Metamodells unter dem Reifegrad „experimentelle Phase“ [ID23a] geführt. Zudem ist weder in [ID23a] ein Eventtyp beschrieben, mit dem die Übertragung einer I4.0-Nachricht möglich wäre, noch eine API-Operation zum Aufruf von Events. Das führt dazu, dass der Nachrichtenaustausch über Events für die Praxis bisher noch ungeeignet ist. Die Operation hingegen besitzt eine eigene API-Operation, über die die Operation aufgerufen werden kann. Damit die Operation möglichst interoperabel für den Nachrichtenaustausch genutzt werden kann wird die Einführung eines in dem Projekt VWS4LS entwickelten Teilmodells namens „Message Participant“ vorgeschlagen (siehe Abbildung 6). Dieses Teilmodell bildet die Beschreibung für den Messenger aus dem Konzept von [BD19] und implementiert zusätzlich Operationen für das Anstoßen zusätzlicher, in der Interaktion benötigter, Algorithmen.

Unter „userOperation“ werden benutzerdefinierte Operationen modelliert, die die für ein Interaktionsmuster zusätzlich benötigten Algorithmen anstoßen können. Diese Operationen können mit den implementierten Interaktionsmustern variieren. Die Operation „newMessage“ ist für die Übertragung der I4.0-Nachricht zuständig und bleibt in ihrer Struktur konstant. Die einzigen variablen Größen bezüglich des übermittelten Inhalts bilden die „interactionElements“. Zusätzlich erhält das Teilmodell eine Auflistung der Protokolle und Rollen, an denen die proaktive VWS teilnehmen kann.

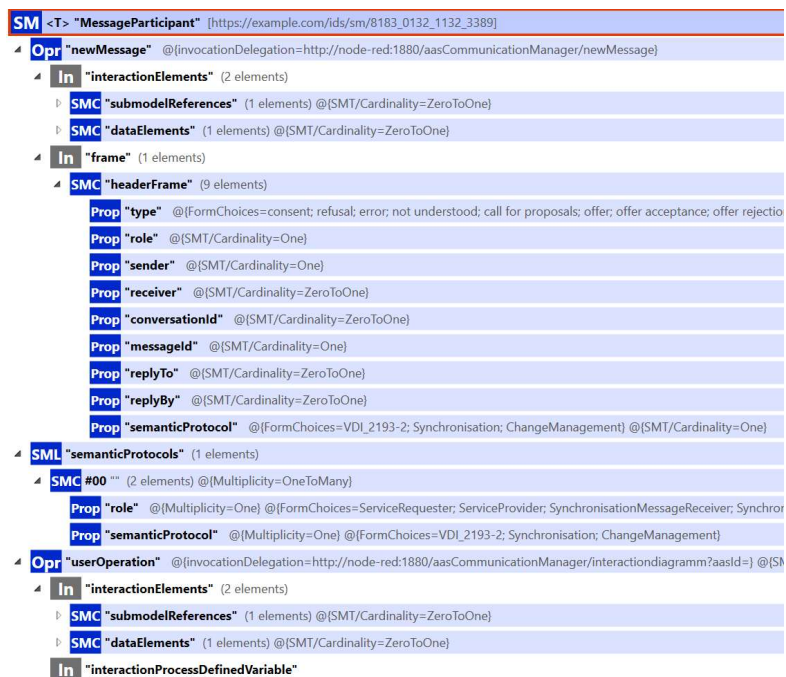


Abbildung 6: Teilmodell "Message Participant" für die Umsetzung proaktiver VWS

3.3 Fazit

Für die Interaktion mit VWS und ihren Inhalten ist bisher nur die VWS-API [ID23b] mit dem http/REST-Protokoll spezifiziert. Die Nutzung anderer Kommunikationsprotokolle ist somit vorerst proprietär. Betrachtet man die VWS-Elemente, zur Ausführung von Algorithmen, steht derzeit nur die „Operation“ zur Verfügung, da sich das „Event“ noch in der experimentellen Phase befindet. Damit jedoch I4.0-Nachrichten über die bestehenden VWS-Technologien gesendet und proprietäre Formate umgangen werden können, wurde das im Projekt VWS4LS entwickelte Teilmodell „Message Participant“ als erster Vorschlag für die interoperable Umsetzung proaktiver VWS vorgestellt. Mit Hilfe

dieses Teilmodells sind sowohl die Protokolle, mit den von der VWS eingenommenen Rollen, als auch die Struktur zur Übergabe von I4.0-Nachrichten spezifiziert.

4 Interoperabilität trotz vielfältiger Interaktionsmuster ermöglichen

Bisher wurde in der VDI/VDE 2193-2 [VD20b] das Ausschreibungsverfahren zwischen einem Auftraggeber und einem Auftragnehmer als Interaktionsmuster spezifiziert. Diese Interaktion ermöglicht sowohl die unternehmensinterne als auch unternehmensübergreifende Verhandlung für eine automatisierte Auftragsvergabe zwischen einem zu fertigenden Produkt (mit geforderten Fähigkeiten) und der Ressource (mit angebotenen Fähigkeiten). Neben dem genannten Interaktionsmuster gibt es weitere Interaktionsmuster, die aus anderen Anwendungsfällen für proaktive VWS entstehen. Darunter fallen z. B. die automatisierte Anmeldung von Wartungs- und Instandhaltungsmaßnahmen [Is23] oder das automatisierte Änderungsmanagement in der Entwicklung [AR24]. Es kann erwartet werden, dass die Anzahl der Interaktionsmuster in der Zukunft stetig steigen wird. Aufgrund dessen müssen die in den vorherigen Abschnitten vorgestellten Konzepte die Integration der individuellen Interaktionsmuster über den Interaktionsmanager als aktive Komponente ermöglichen. Ein leichtgewichtiger Lösungsansatz hierfür ist die Nutzung von Workflow Management Systemen, wie es in dem orchestrierenden Ansatz bereits erwähnt wurde. Dabei werden zunächst die zu einem Interaktionsmuster zugehörigen Zustandsmaschinen der einzelnen Interaktionsteilnehmer abgeleitet und darauf basierend Workflows modelliert. Die einzelnen Workflowobjekte werden mit Businesslogik ausgestattet, die das Versenden von I4.0-Nachrichten über die in Abschnitt 3.2 genannte Operation „newMessage“, deren Empfang, sowie das Anstoßen der benutzerdefinierten Operationen aus dem Teilmodell „Message Participant“ übernehmen. Die einzelnen Zustandsmaschinen können jedoch genauso in dem monolithischen Ansatz in einem VWS-Server/-Repository oder in dem agentenbasierten Ansatz in einem Agenten-Typen programmiert werden. Vorteil der WMS ist jedoch, dass sie für den Entwickler der Zustandsmaschinen einfacher verständlich sind, da sie eine grafische Repräsentation der Zustandsmaschinen ermöglichen. Zudem stellen viele WMS-Lösungen bereits Schnittstellen zu Applikationen der Produktions-, Betriebs- und Unternehmensleitebene her, sodass übergreifende Interaktionsmuster einfacher zu implementieren sind.

5 Ausblick

Die vorgestellten Konzepte haben verschiedene Implementierungsmöglichkeiten für proaktive VWS in Hinsicht auf deren Architektur betrachtet. Die meisten Forschungsarbeiten bedienen sich dabei dem agentenbasierten Ansatz. Der monolithische und orchestrierende Ansatz sind bisher wenig untersucht. Zudem wurde die Integration

der I4.0-Sprache aus der VDI/VDE 2193-1 mit Hilfe der bestehenden VWS Spezifikationen betrachtet. Da nur wenige der genannten Arbeiten die Integration der I4.0-Sprache verfolgten wurde ein in [AR24] entwickeltes Konzept zur Übertragung der I4.0-Nachricht in dieser Arbeit vorgestellt. Die Lösung bietet mit dem VWS-Element „Operation“ einen leichtgewichtigen Ansatz für die Übermittlung von I4.0-Nachrichten über die bereits spezifizierte VWS-API und sollte in weiteren Arbeiten auf deren Allgemeingültigkeit hin untersucht werden. Auf den Einsatz des VWS-Elements „Event“ zur Nachrichtenübertragung sollte solange verzichtet werden, bis dieses ausreichend spezifiziert ist. Ansonsten werden die erarbeiteten Lösungen in der Praxis nur schwer Einzug halten.

Zukünftig sollten praktische Evaluierungen der hier vorgestellten Implementierungskonzepte einer proaktiven VWS erfolgen. Dabei sollten unter anderem die Anforderungen an Flexibilität, Performance, Integration von Zugriffskontrollen und Stabilität jedes vorgestellten Ansatzes untersucht werden. Zudem wird eine Spezifikation der proaktiven VWS benötigt, die einen Konsens über die Anforderungen an eine proaktive VWS schafft, um auch die in der Industrie 4.0 angestrebte Interoperabilität zu fördern und nicht zu behindern. Dies muss zudem zeitnah geschehen, damit vor allem die Implementierung der proaktiven VWS in der industriellen Umgebung ein Erfolg wird.

Literaturverzeichnis

- [AR24] Verwaltungsschale für den Leitungssatz. Zugegriffen: 15. April 2024. [Online]. Verfügbar unter: <https://arena2036.de/de/verwaltungsschale-fuer-den-leitungssatz>
- [BA24] BAFA - Modernisierung Fahrzeughersteller / Zulieferindustrie. Zugegriffen: 22. Januar 2024. [Online]. Verfügbar unter: https://www.bafa.de/DE/Wirtschaft/Handwerk_Industrie/Modernisierung_Fahrzeughersteller_Zulieferindustrie/modernisierung_fahrzeughersteller_zulieferindustrie_node.html
- [BD19] A. Belyaev und C. Diedrich: Specification „Demonstrator I4.0-Language“ v3.0. 2019.
- [BM24] B.-B. für W. und Klimaschutz: Das Gaia-X Ökosystem - Souveräne Dateninfrastruktur für Europa. Zugegriffen: 22. Januar 2024. [Online]. Verfügbar unter: <https://www.bmwk.de/Redaktion/DE/Dossier/gaia-x.html>
- [CF21] Cardoso, R.C.; Ferrando, A.: A Review of Agent-Based Programming for Multi-Agent Systems. Computers 2021, 10, 16. <https://doi.org/10.3390/computers10020016>
- [EB24] Eclipse BaSyx™. Zugegriffen: 15. April 2024. [Online]. Verfügbar unter: <https://github.com/eclipse-basyx>

-
- [Ge15] M. Geirhos: Entwurfsmuster - Das umfassende Handbuch, 1. Aufl. Bonn: Rheinwerk Verlag, 2015.
- [Ca23] G. P. Cainelli, P. Yazdani, L. Underberg, U. Jumar, and C. E. Pereira: Managing the 5G system based on production requirements using an Asset Administration Shell approach. IFAC-PapersOnLine, vol. 56, no. 2, pp. 2108–2114, 2023, doi: 10.1016/j.ifacol.2023.10.1113.
- [Gr22] S. Grunau, M. Redeker, D. Göllner, and L. Wisniewski: The Implementation of Proactive Asset Administration Shells: Evaluation of Possibilities and Realization in an Order Driven Production. pp. 131–144, 2022, doi: 10.1007/978-3-662-64283-2_10.
- [ID23a] Industrial Digital Twin Association, Hrsg.: Asset Administration Shell Specification - Part 1: Metamodel Schema. März 2023.
- [ID23b] Industrial Digital Twin Association, Hrsg.: Asset Administration Shell Specification - Part 2: Application Programming Interfaces. Juni 2023.
- [Is23] R. Islam u. a.: Erfahrungsbericht bei der Umsetzung der VWS Type 3 : Interaktionen in einer Maintenance-Anwendung. 2023, doi: 10.25673/111634.
- [Ju23] S. Jungbluth, A. Witton, J. Hermann and M. Ruskowski: Architecture for Shared Production Leveraging Asset Administration Shell and Gaia-X. 2023 IEEE 21st International Conference on Industrial Informatics (INDIN), Lemgo, Germany, 2023, pp. 1-8, doi: 10.1109/INDIN51400.2023.10218150.
- [PI24] Manufacturing-X Initiative. Zugegriffen: 23. Januar 2024. [Online]. Verfügbar unter: <https://www.plattform-i40.de/IP/Navigation/DE/Manufacturing-X/Initiative/initiative-manufacturing-x.html>
- [PI20] Plattform Industrie 4.0: Verwaltungsschale in der Praxis – Wie definiere ich Teilmodelle, beispielhafte Teilmodelle und Interaktion zwischen Verwaltungsschalen (Version 1.0). Berlin: BMWi, 2020.
- [SLP22] L. Sakurada, P. Leitao and F. De La Prieta: Engineering a Multi-agent Systems Approach for Realizing Collaborative Asset Administration Shells. 2022 IEEE International Conference on Industrial Technology (ICIT), Shanghai, China, 2022, pp. 1-6, doi: 10.1109/ICIT48603.2022.10002770.
- [SPL23] L. Sakurada, F. D. L. Prieta and P. Leitao: A Methodology for Integrating Asset Administration Shells and Multi-agent Systems. 2023 IEEE 32nd International Symposium on Industrial Electronics (ISIE), Helsinki, Finland, 2023, pp. 1-6, doi: 10.1109/ISIE51358.2023.10227964.
- [VD20a] VDI, Hrsg.: VDI/VDE 2193 Blatt 1- Sprache für I4.0-Komponenten - Struktur von Nachrichten. 2020.
- [VD20b] VDI, Hrsg.: VDI/VDE 2193 Blatt 2 - Sprache für I4.0-Komponenten - Interaktionsprotokoll für Ausschreibungsverfahren. 2020.

- [WA24] Digitales Engineering, WAGO Deutschland. Zugegriffen: 22. Januar 2024. [Online]. Verfügbar unter: <https://www.wago.com/de/digitalisierung/digitales-engineering>
- [WI24] Digitaler Zwilling - WITTENSTEIN SE. Zugegriffen: 22. Januar 2024. [Online]. Verfügbar unter: <https://www.wittenstein.de/de-de/unternehmen/digitalization-center/digitaler-zwilling/>
- [YH20] X. Ye and S. H. Hong: Toward the Plug-and-Produce Capability for Industry 4.0: An Asset Administration Shell Approach. no. December, 2020.

„Kazeem Oladipupo, Sandeep Rudra, Christian Diedrich, Sebastian Kaebisch (Hrsg.):
EKA 2024 - Entwurf komplexer Automatisierungssysteme, 18. Fachtagung“

Analytics and Visualization of Industrial Asset Condition using Asset Administration Shell submodels.

Kazeem Oladipupo¹, Sandeep Rudra¹, Christian Diedrich¹, Sebastian Kaebisch²

Abstract: The advent of industry 4.0 necessitates a paradigm shift toward autonomous industrial operations, prompting the integration of digital twin to orchestrate machines and processes within automation systems. The Asset Administration Shell (AAS) emerges as a pivotal digital twin intended to encapsulate an asset's lifecycle. It comprises of several submodels defining different aspects of machines or processes that in turn empowers continuous autonomy.

This potential of AAS opens the scope to extend the standardization of machine processes that are common within a factory. One of these processes is a self-diagnostic and continuous analysis of the state of a machine. This imperative not only mitigates machine downtime but expeditiously detect faults or anomalies in production processes.

This study explores the integration of Asset Administration Shell (AAS) into industrial machine analytics, aligning with Industry 4.0 paradigms. It emphasizes the application of AAS submodels for enabling machines to self-diagnose and continuously analyze their state, focusing on a universal analytics approach. This is achieved through a standardized submodel based on ISO 22400-2:2014 Key Performance Indicators (KPIs), facilitating a vendor-agnostic solution for machine analytics.

The paper highlights the effectiveness of these standardized submodels in improving machine efficiency, predictive maintenance, and operational effectiveness in industrial processes. It also discusses the challenges and practical applications of these submodels, offering insights into their real-world implementation.

Keywords: Industry 4.0, Asset Administration Shell, Industrial analytics, ISO 22400-2, data integration.

1 Introduction

In the words of Clive Humby, a British mathematician, “Data is the new oil. It’s valuable, yet if unprocessed, it remains of little use. Just like oil must be converted into fuel or plastic, data too needs to be refined and analyzed to unlock its true potential.” [5]. This concept, introduced in 2006, serves as a cornerstone in the realm of data-driven approaches. In the pursuit of Industry 4.0 objectives, manufacturers globally are focusing on enhancing operational efficiency and product quality.

The adoption of Industry 4.0 is driven by the inclusion of cutting-edge technologies such as the Internet of Things (IoT), Artificial Intelligence (AI), and big data analytics. Within this framework, digital twins, facilitated through the Asset Administration Shell (AAS),

¹ Otto Von Guericke Universität, Integrated Automation, Universitätspatz 2, 39106. Magdeburg,
{christian.diedrich & [kazeem.oladipupo](mailto:kazeem.oladipupo@ovgu.de)}@ovgu.de
sandeep.rudra@st.ovgu.de

² Siemens AG, Technology, Otto-Hahn-Ring 6, 81739, München, sebastian.kaebisch@siemens.com

stand out as pivotal for defining assets in Industry 4.0. The AAS, with its submodels, provides a standardized framework for capturing essential characteristics of industrial assets [6].

The impetus for this paper springs from the synergy between extensive manufacturing knowledge and the innovative potential of Industry 4.0 technologies. The research aims to delve into the capabilities of the AAS and leverage the power of its submodels to bolster analytics capabilities, offering standardized descriptions for analytical KPIs. This endeavor seeks to bridge the gap between the abundance of available data and the extraction of practical insights. It advocates methodologies for depicting analytical KPIs and storing these within the AAS.

The structure of this paper is organized by beginning with the introduction which explains addresses potentials of data driven application and AAS, Section 2 presents a brief background on key topics central to the paper. Section 3 provides related work which review existing literature in area of analytics. Section 4 investigates proposed methodology for implementing the findings. It describes an analytical framework for defining KPIs within AAS and industrial analytics domain. Section 5 provides details about the implementation of the proposed methodology. Section 6 discusses some findings and provides recommendation. Section 7 concludes the paper and summarizes the research.

2 Background

2.1 Asset Administration Shell

The AAS [1] is a core element in the concept of the I4.0 Component, as evident from the AAS's definition. According to Plattform I4.0, an I4.0 Component is "*an entity with a global unique identifier and communication abilities, comprising an administration shell and an asset in an I4.0 system, offering services with specified QoS (quality of service).*" [7].

Grasping the connection between the I4.0 Component and the AAS in the context of manufacturing and Industry 4.0 requires certain critical insights. Hoffmeister underlines the crucial role of the I4.0 Component in the development of Smart Factories, marking its importance [8]. Cyber-Physical Systems are recognized for their amalgamation of the physical and digital realms. The Asset Administration Shell (AAS) symbolizes the digital facet of a CPS. Within the framework of Industry 4.0, it functions as the interface that connects the physical and virtual worlds.

The metamodel of an AAS is defined to contain collection of *submodels* which defines different aspect of the asset. in the submodel, is collections of *submodelElements* (SME). Some type of SMEs are *submodelElementCollection*(SMC) *property*, *range*, *submodelElementList*(SML), *referenceElement*, etc. in this paper, SMC and SME will be used to denote *submodelElementCollection* and other *submodelElements* types respectively.

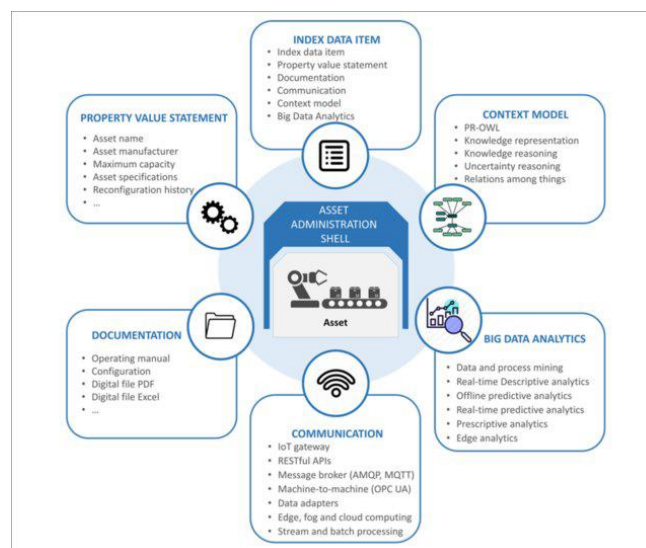


Figure 1: The role of AAS in I4.0 [9].

The AAS exist in three distinct types each offering different levels of complexity and interactivity.

The Type 1 AAS are the most basic form. They are essentially serialized files, often in formats like XML or JSON. These shells contain static information about the asset, which doesn't change over time [10]. The Type 2 AAS represent a more dynamic and interactive category. Existing as runtime instances, they are hosted on servers and can contain both static and dynamic information [10]. The Type 3 AAS are more intelligent type of AAS built on type 2 AAS to be autonomous and have decision making capabilities. They can interact with other AAS using industry 4.0 language.

2.2 Overview of Industrial Analytics

The manufacturing sector worldwide is experiencing a profound shift moving from traditional physical production to focusing on data-driven processes and products. This digital shift is creating vast amount of diverse data throughout the industrial value chain, ranging from simulation data in the design phase to sensor data during manufacturing, and sensor data during the usage of the product therefore huge amount of data is generated during the complete life cycle of a product, encompassing all stages from initial design to eventual recycling. Gaining valuable insights and knowledge from this data is becoming a crucial factor for success in the industry, such as for process optimization and product enhancement [11]. This process is known as industrial analytics, which involves applying data analytics specifically for generating industrial value. Sitting at the intersection of data science and industrial engineering, industrial analytics is fundamental to Industry 4.0 [12].

The phrase "industrial analytics" typically denotes the use of data analytics for generating value in the industrial sector, often in the context of Industry 4.0. It's also interchangeably

referred to as "Industry 4.0 analytics" and "industrial intelligence". Common forms of analytics are descriptive analytics, diagnostic analytics, prescriptive analytics, and predictive analytics [13]. Table 1 shows the focus and basic analytic question attached to these analytics types.

Table 1 : Types of Analytics and their focus areas

	Descriptive Analytics	Diagnostic Analytics	Predictive Analytics	Prescriptive Analytics
Focus	Transparency	Root Cause	Forecast	Action
Analytical Question	What has happened? What is happening?	Why has it happened?	What will happen?	How can we make it happen?
Example	What is the current first pass yield (FPY)?	What has decreased in certain regions?	What will be FPY in the next quarter?	How can we increase FPY?

3 Related Work

Predictive Maintenance (PdM) plays a crucial role in the framework of Industry 4.0, particularly within the context of smart manufacturing. The AAS is suggested as an essential element in attaining interoperability and reducing the complexity of operational technology. This is crucial for harmonizing various PdM strategies in the ever-changing industrial sector [14] [15]. An important aspect of authors research is a case study that examines the application of machine learning in predictive maintenance (PdM) of milling machines. This case study demonstrates the practicality of the methodology by using data from diverse sources to perform predictive analytics. The model's potential in various industrial applications and its effectiveness in solving real-world difficulties are highlighted by these practical demonstrations [14][15].

The development of ISO 22400 was driven by the need for a standardized approach to performance measurement in the manufacturing industry. Prior to its introduction, manufacturers often relied on a variety of inconsistent and non-comparable metrics to assess their operations. This lack of standardization led to significant challenges in benchmarking and optimizing performance. In response, ISO 22400 emerged, providing a structured and universally applicable set of KPIs. These KPIs align with the evolving needs and technological advancements in the manufacturing sector, offering a solution to the previously fragmented approach to performance measurement. The standard's development also reflects a growing trend towards globalization in manufacturing, necessitating a common framework that can be applied across diverse manufacturing environments [3]

The discussion of ISO 22400[3] is pertinent as this standard provides the framework for defining analytical KPIs in the context of the analytic submodel. This standard is

instrumental in establishing a structured approach for the description of key performance indicators, thereby facilitating a standardized and coherent method for developing the analytic submodel.

4 KPI description with AAS submodel

In the context of AAS, the overall architecture used in this paper is provided in Figure 2. In this architecture is an AAS server hosting a type 2 AAS. This AAS contains some submodels that collaborate to provide an analytics value for the machine user. The analytics submodel is the focus of this paper. Other submodels shown in Figure 2 are in one way or the other helping the analytics service (Analytics App) to carry out its function real-time and on-demand.

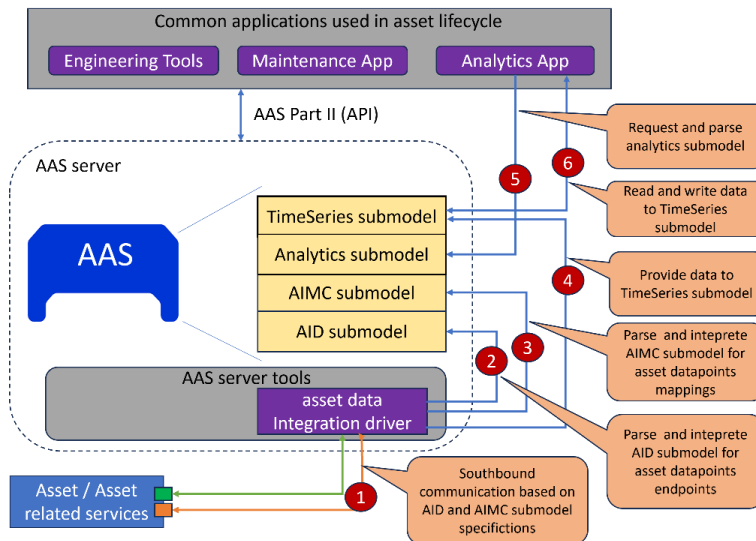


Figure 2: Proposed methodology.

4.1 Analytics Submodel

Providing a top-down mapping of the KPI structure provided in ISO-22400[3] into analytics submodel. As seen in Figure 3, the Top SMC defines the name of the KPI and inside the SMC, other information related to the KPI is provided. The *Datapoints* (an SMC containing Reference SMEs) and *Value* (a Reference SME) are extension of the original KPI structure. They are used to provided references to parameters used in calculating KPIs and storing the calculated outcome respectively.

Other information in the *KPI SMC* are used by the analytics service as a guidance on how the calculation will be performed. For example, the *Timing* SME can define if a KPI should

be calculated in real-time, periodically or on-demand. The Rating SME tells the upper and lower limits trend of the KPI.

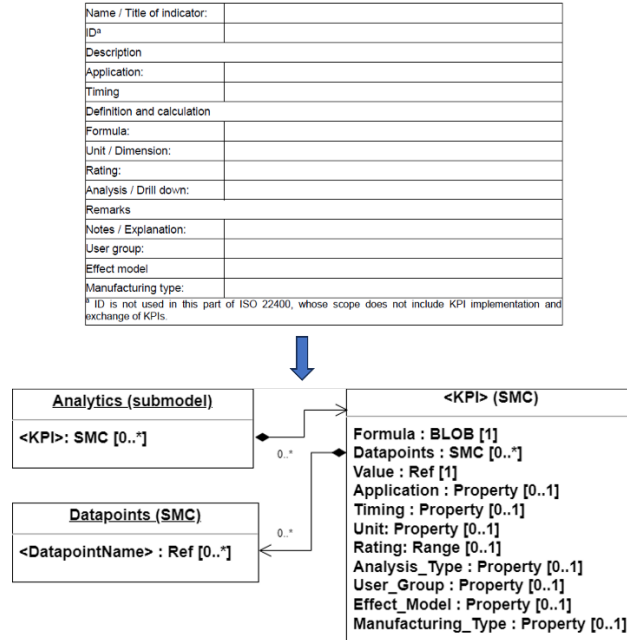


Figure 3: Mapping of ISO 22400 KPI structure into Analytics Submodel UML

4.2 Asset Interfaces Description (AID) and Asset Interfaces Mapping Configuration Submodels (AIMC).

The AID [4] and AIMC submodels shown in Figure 2 helps the AAS integrate its corresponding asset's operational data. The AID submodel contains description of interfaces related to asset datapoints and/or related services and AIMC provides mapping of asset payload to a designated location within the AAS.

Since all the metrics needed to calculate a machine KPI are all time dependent information, bringing them into the AAS needs to be defined as time dependent parameters. The *TimeSeries* [2] submodel offers such possibilities. The *TimeSeries* submodel allows AAS elements to be recorded with respect to the time.

This paper combines the potentials of already defined submodels like *TimeSeries* submodel, AID and AIMC submodel to achieve its analytics goal.

5 KPI implementation with AAS analytics service.

This chapter focuses on the implementation of the submodel modelled in section 4. This implementation consists of two interconnected parts: The development of an AAS with necessary submodels as seen in Figure 2 and the integration of analytical service as a proof of concept to the potential of analytics submodel. The metric that would be focused on this development is the *Overall Equipment Efficiency* (OEE).

5.1 AAS and Submodels Development:

Leveraging the AAS package Explorer, the AAS and its submodels are developed to encapsulate all the core aspects of an asset. Amongst these aspects is the asset's operational metrics. The *OEE* is one of the important metrics used to evaluate how well a machine is operating with respect to its expected production output. To evaluate a machine's OEE, metrics like *availability*, *quality* and *performance* of the asset is captured or calculated. Figure 4 shows the overview of an asset's AAS containing the analytics submodel where holds the descriptions of some KPIs used to evaluate the machines overall efficiency.

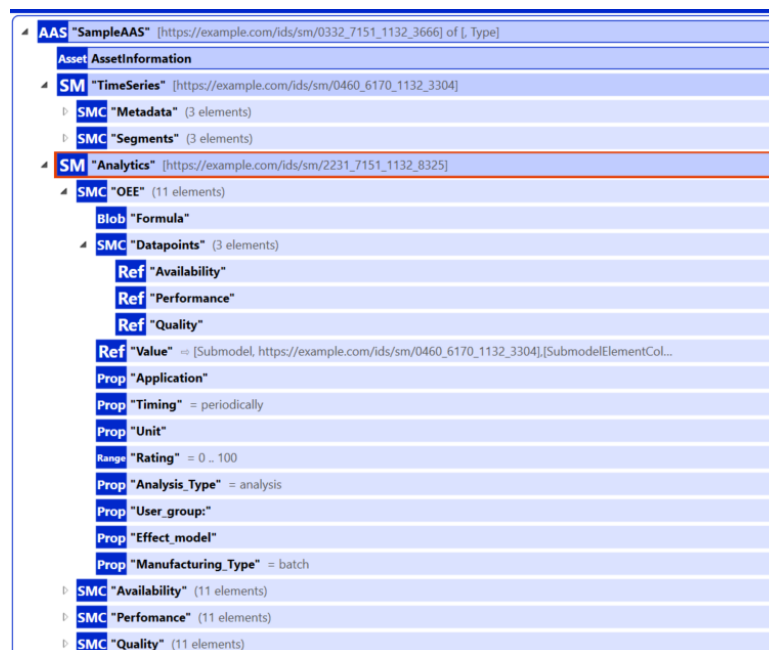


Figure 4: Sample AAS with Analytics Submodel as described in section 4 and Figure 3

As seen in Figure 4, the Formula (BLOB SME) contains the formula for the OEE in a blob format and the datapoints SMC provides references to the parameter to be used in the formula. For this implementation, the datapoints are time-based metrics so they are provided in the *TimeSeries* submodel and referenced in each metrics (Availability SMC,

Performance SMC and Quality SMC) datapoint. The value of the calculated OEE is also provided in the *TimeSeries* submodel because it is a time-based metric.

The OEE which is a well-known KPI in manufacturing, combining availability, effectiveness, and quality ratio. It shows the percentage of productive manufacturing time. A perfect OEE score of 100 percent means manufacturing high-quality parts efficiently and without stoppages. OEE is ideal for improving equipment productivity, identifying losses, or benchmarking performance. It is calculated with the following formula:

$$OEE\ Index = Availability * Quality\ ratio * Performance\ [3]$$

5.2 Analytic Service

The analytic service uses the AAS analytics submodel to run analytics on an AAS specific to an asset. Figure 2 shows the overall picture of how the analytics service is implemented.

The flowchart of the analytics service is provided in Figure 5 considering that AAS server is running, and asset data integration driver shown in Figure 2 is integrating asset datapoints into the AAS. The flowchart serves as a guide to the systematic process the analytic service implementation, detailing each step from data input to processing and output.

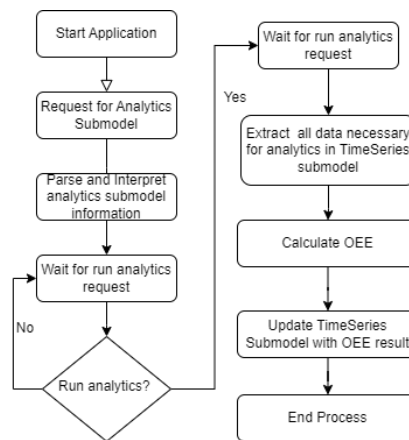


Figure 5: Flowchart for analytics service.

6 Discussion and Recommendation

During the development of the analytics submodel, it was seen that the *TimeSeries* submodel is an important submodel for its implementation because that is the only submodel that stores time-related data within AAS ecosystem.

One of the drawbacks of the *TimeSeries* submodel is that it is structured to store collection of data within one time stamp. The metadata SMC does not allow the creation of multiple records for different data collection. For instance, asset datapoints or asset related services data can be created with one record and an opportunity to create another record that is not specific to asset datapoints.

It is recommended that this feature is investigated because if AAS will capture different aspects of asset information across its life cycle, it is certain that some of these aspects might need TimeSeries submodel at different time by different applications.

The *analytical* submodel is an ideation of how asset analytics can be achieved with AAS. There is still some work to be done on the modelling most especially on the Formula SME. Some rules on how the expression will be presented must be defined for all users to know how to properly provide their formulas.

Also, because the formula SME is using BLOB, the information provided in the BLOB needs to be properly and intelligently interpreted by analytics service according to the expression rules.

7 Conclusion and Outlook

This paper has provided a review of fundamental elements of Industry 4.0, specifically emphasizing on the Asset Administration Shell (AAS) and its use in the domain of industrial analytics. As industrial automation technologies continue to advance, the importance of on-the-fly analytics that can be used to evaluate assets is paramount. Providing a standardized description for well-known KPIs to enable *plug-and-analyze* capabilities for applications on the shop floor will help facilitate early detection of problems in machines.

The proposed methodology aligns with current Industrial approach for implementing the AAS. The methodology architecture has been designed to create a connection between *Analytics*, and *TimeSeries* submodels, with each submodel serving a distinct function in the asset. The method included the process of defining asset datapoints in AID and AIMC submodels that facilitates data integration from asset.

The implementation section provides an overview of how the proposed methodology was implemented. This involved the development of the AAS and its submodels using the AASX Package Explorer, defining important KPIs and their parameters, which are then used by analytical services.

Though the analytics submodel in this paper focuses on assets KPIs. In future work, an extension of how this submodel can be used to analyze a part (component) of asset from the information provided in either the asset's PDF manual or *TechnicalData* submodel with the use of technology like Large Language Models (LLMs) is expected to be explored.

Bibliography

- [1] IDTA. Specification of the asset administration shell part 1: Metamodel - idta number 01001 3.0. Industrial Digital Twin Association, 2023.
- [2] Industrial Digital Twin Association. IDTA-02008-1-1 Submodel TimeSeries-Data. https://industrialdigitaltwin.org/en/wp-content/uploads/sites/2/2023/03/IDTA-02008-1-1_Submodel_TimeSeriesData.pdf, 2023. Accessed: [your access date].
- [3] Automation systems and integration, Key Performance Indicators (KPIs) for manufacturing operations management Part 2: Definitions and descriptions. International Organization for Standardization (ISO). Geneva. 01-2014
- [4] Industrial Digital Twin Association, IDTA 02017-1-0 Asset Interfaces Description. Industrial Digital Twin Association, 2024.
- [5] M. Palmer. Data is the new oil, May 2006. https://ana.blogs.com/maestros/2006/11/data_is_the_new.html.
- [6] M. Rüßmann, M. Lorenz, P. Gerbert, M. Waldner, J. Justus, P. Engel, and M. Harnisch. Industry 4.0: The future of productivity and growth in manufacturing industries. Boston Consulting Group, (9), 2015.
- [7] E.K.K. Pepels. Adding smartness to smart factories. Eindhoven University of Technology Institutional Repository.
- [8] Michael Hoffmeister. The Industrie 4.0 Component. ZVEI, 2015.
- [9] Alexandros Bousdekis and Gregoris Mentzas. Enterprise integration and interoperability for big data-driven processes in the frame of industry 4.0. *Frontiers in Big Data*, 4, 2021. https://www.researchgate.net/publication/352112470_Enterprise_Integration_and_Interoperability_for_Big_Data-Driven_Processes_in_the_Frame_of_Industry_40
- [10] Eclipse BaSyx. BaSyx / Documentation / Asset Administration Shell, 2020. https://wiki.eclipse.org/BaSyx/_/Documentation/_/AssetAdministrationShell.
- [11] A. Dogan and D. Birant. Machine learning and data mining in manufacturing. *Expert Systems with Applications*, 166, 2021.
- [12] Y. Cui, S. Kara, and K. C. Chan. Manufacturing big data ecosystem: a systematic literature review. *Robotics and Computer Integrated Manufacturing*, 62, 2020.
- [13] Tarig Khalid. Data analytics capsules, 2022. https://www.researchgate.net/publication/359109194_Data_Analytics_Capsules.
- [14] Salvatore Cavalieri and Marco Giuseppe Salafia. A model for predictive maintenance based on asset administration shell. *Sensors*, 20(21):6028, 2020. <https://doi.org/10.3390/s20216028>.
- [15] J.R. Rahal, A. Schwarz, B. Sahelices, et al. The asset administration shell as enabler for predictive maintenance: A review. *Journal of Intelligent Manufacturing*, 2023. Received: 09 March 2023; Accepted: 06 October 2023; Published: 08 November 2023.

„Ulrich Jumar, Christian Diedrich (Hrsg.):
EKA 2024 - Entwurf komplexer Automatisierungssysteme, 18. Fachtagung“

Praktikable Ermittlung des produktbezogenen CO₂-Fußabdrucks mithilfe grüner Digitaler Energiezwillinge

Florian Balduf¹, Zai Müller-Zhang², Thomas Kuhn³

Keywords: Digitale Zwillinge, CO₂-Fußabdruck, Nachhaltigkeit, Verwaltungsschale.

Abstract

In dieser Veröffentlichung führen wir grüne Digitale Energiezwillinge (gDEZ) als einen vielversprechenden Ansatz zur Ermittlung des produktbezogenen CO₂-Fußabdrucks in der Fertigungsindustrie gemäß etablierten Standards ein. Angesichts der zunehmenden Dringlichkeit, Treibhausgasemissionen zu reduzieren, ist es entscheidend, effektive Methoden zur Bewertung und Optimierung von Produktionsprozessen zu entwickeln. Der Einsatz von digitalen Zwillingen zur effizienten Ressourcennutzung steht im Zentrum unseres Ansatzes, mit dem Ziel, die Treibhausgasemissionen der hergestellten Produkte zu ermitteln und Potenziale für deren Reduzierung zu erschließen. Der gDEZ-Ansatz integriert Datenmodelle mit den Aspekten Ressource, Produkt und Prozess und Berechnungsmodelle mit unterschiedlichen Detailgraden, um eine praktikable Ermittlung des CO₂-Fußabdrucks auf Produktions- und Produktebene zu ermöglichen. Der Artikel skizziert das Datenmodell der gDEZ, zeigt Anwendungsszenarien sowie das Konzept der Berechnungsmodelle auf und diskutiert zukünftige Entwicklungen in diesem Bereich.

1 Einleitung

Die globalen Klimaveränderungen haben ein Umdenken in der Nutzung von Produkt- und Bestandteilen innerhalb der Kreislaufwirtschaft sowie in der Erfassung umweltbezogener Kosten für Produkte, wie beispielsweise Carbon Credits, angestoßen [Fi23]. Dies stellt eine Herausforderung für die Wettbewerbsfähigkeit dar, da Unternehmen verpflichtet sind, Produktinformationen über den gesamten Lebenszyklus hinweg zu dokumentieren. Die Potenziale der Industrie 4.0, insbesondere das Konzept der servicebasierten Architekturen, können hierbei einen erheblichen Beitrag leisten [TA17, KSA20].

¹ Fraunhofer IESE, Embedded Systems Engineering, Fraunhofer-Platz 1, 67663 Kaiserslautern, florian.balduf@iese.fraunhofer.de, <https://orcid.org/0000-0002-3367-1690>

² Fraunhofer IESE, Virtual Engineering, Fraunhofer-Platz 1, 67663 Kaiserslautern, zai.mueller-zhang@iese.fraunhofer.de, <https://orcid.org/0000-0002-0228-943X>

³ Fraunhofer IESE, Embedded Systems, Fraunhofer-Platz 1, 67663 Kaiserslautern, thomas.kuhn@iese.fraunhofer.de, <https://orcid.org/0000-0001-9677-9992>

Ein zentrales Anliegen, insbesondere für kleine und mittlere Unternehmen (KMU), besteht in der praktikablen Ermittlung des CO₂-Fußabdrucks auf Basis etablierter Berechnungsstandards. Aktuell greifen KMU oft auf individuelle Lösungen z.B. unter Verwendung von MS Excel zurück, was zu einer Fragmentierung und Inkonsistenz der Daten führen kann. Die Einführung einer standardisierten Datenablage könnte nicht nur die Berechnung der THG-Emissionen vereinfachen, sondern auch in anderen Bereichen bedeutende Vorteile bieten. Dies würde zu einem einheitlichen Vorgehen führen und es ermöglichen, Informationen vergleichbar zu machen und über die Wertschöpfungskette hinweg zu verwenden. In diesem Artikel präsentieren wir einen praxisnahen Ansatz zur Ermittlung des CO₂-Fußabdrucks. Der Ansatz basiert auf grünen Digitalen Energiezwillingen, welche Daten und Berechnungsmodelle beinhalten.

Der Aufbau des Artikels gestaltet sich wie folgt. Nach der Einleitung in Abschnitt I führt Abschnitt II in den Hintergrund des digitalen Zwillings und der Verwaltungsschale ein. Ein besonderer Schwerpunkt liegt auf bisherigen Arbeiten im Bereich Product Carbon Footprint (PCF), und verwandte Arbeiten im Bereich Industrie 4.0 und Nachhaltigkeit werden untersucht. In Abschnitt III wird die Problemstellung beleuchtet und aufgezeigt, warum die praktikable Ermittlung des CO₂-Fußabdrucks viele Unternehmen vor große Herausforderungen stellt. Abschnitt IV stellt unser Lösungskonzept vor und erläutert, warum die Datenmodelle der gDEZ ein Schritt hin zur praxistauglichen Ermittlung produktbezogener Treibhausgasemissionen sind. In Abschnitt VI wird der Artikel zusammengefasst und zukünftige Arbeiten werden diskutiert.

2 Stand der Technik

In der aktuellen Forschung und industriellen Praxis haben digitale Zwillinge eine zentrale Bedeutung erlangt, insbesondere in Bezug auf Nachhaltigkeitsmodelle und deren Anwendung in der Fertigungsindustrie. Die Arbeiten von Sigh et al. und Jabbour et al. haben gezeigt, dass digitale Zwillinge die Resilienz von Lieferketten signifikant verbessern und zur Nachhaltigkeit der Produktionsprozesse beitragen können. [Si23] [Ja20]

Die Nutzung von Big Data und diskreten Ereignissimulationsmodellen, wie von Chávez et al. vorgestellt, ermöglicht eine detaillierte Bewertung und Optimierung von Nachhaltigkeitsleistungen in Fertigungssystemen [Ch22]. Edrisi et al. betonen die Fähigkeit digitaler Zwillinge, umfangreiche Datenmengen zu analysieren und Prognosen über Nachhaltigkeitsindikatoren zu erstellen, was besonders bei zunehmender Anzahl beteiligter Unternehmen und Datenmengen von Vorteil ist [EA23]. Die Integration dieser Technologien ermöglicht eine präzisere Überwachung und Steuerung von Umweltauswirkungen auf Produktebene, was für die Einhaltung von Nachhaltigkeitszielen und die Implementierung von Kreislaufwirtschaftskonzepten unerlässlich ist.

Die Verwaltungsschale (englisch: Asset Administration Shell, AAS) bietet eine technologische Grundlage für den Austausch Asset-zentrierter Informationen über den gesamten Lebenszyklus hinweg. Sie wird durch die Industrie 4.0 Initiative gefördert und zielt darauf ab, Interoperabilität durch standardisierte Datenbeschreibungen zu gewährleisten. Die Möglichkeit, mithilfe der AAS PCF-Informationen auszutauschen und zu optimieren, wie durch die PCF-Submodellvorlagen unterstützt, trägt wesentlich zur Transparenz und Nachvollziehbarkeit bei der Berechnung von und der Berichterstattung zu CO₂-Fußabdrücken bei.

Verschiedene Standards wie ISO 14044, ISO 14067 und das Greenhouse Gas Protocol haben das Ziel, eine Anleitung zur Berechnung der produktbezogenen THG-Emissionen bereitzustellen und dadurch einen internationalen Standard zu etablieren. Diese Standards bieten eine wichtige Orientierungshilfe für Unternehmen, erfordern jedoch eine spezifische Anpassung an die betrieblichen Gegebenheiten. Das GHG Protocol und die Catena-X PCF-Regeln bieten Frameworks für eine ganzheitliche Betrachtung des Lebenszyklus von Produkten. Die derzeitigen Standards liefern jedoch keine detaillierten Informationen, wie der Beitrag der THG-Emissionen der Produktion zum PCF einzelner Produkte ermittelt und zugeordnet werden soll. Unser Ansatz zielt darauf ab, diese Lücke zu schließen, indem wir Aufwendungen, die nicht direkt einem einzelnen Produkt zugeordnet werden können, identifizieren und Berechnungsmodelle entwickeln, wie diese sinnvoll auf das Produkt umgelegt werden können. Dies ermöglicht eine einfache, praxisnahe und dennoch standardkonforme Zuordnung von THG-Emissionen und unterstützt Unternehmen bei der Ermittlung des produktbezogenen CO₂-Fußabdrucks, und dadurch wiederum die Erreichung der Nachhaltigkeitsziele

3 Lösungskonzept

In diesem Abschnitt führen wir die Datenmodelle und Berechnungsmodelle der grünen Digitalen Energiezwillinge ein. Die **Datenmodelle** beinhalten die Aktivitätsdaten für die Ermittlung des CO₂-Fußabdrucks. Die Genauigkeit der Berechnungsmodelle basiert auf dem Digitalisierungsgrad des Produktionssystems. Das bedeutet, dass für eine gut digitalisierte Produktion, in der eine Vielzahl an Aktivitätsdaten digitalisiert ist, ein Berechnungsmodell mit höherer Genauigkeit gewählt werden kann. Andernfalls wird ein Modell mit niedrigerer Genauigkeit ausgewählt.

Mit dem gDEZ-Ansatz wollen wir einen Schritt weiter gehen als bisherige Vorhaben und eine praxistaugliche Berechnung des CO₂-Fußabdrucks auf Produktions- und Produktebene ermöglichen. Hierfür werden die verschiedenen Quellen von Treibhausgasen wie z.B. energiebezogene Emissionen oder Emissionen durch den Transport identifiziert und mithilfe des später beschriebenen Whitebox-Ansatz-Hotspots identifiziert (Datenpunkte, die eine hohe Relevanz für die Berechnung der THG-Emissionen haben). Zur Umsetzung setzen wir auf die standardisierten Datenmodelle der Verwaltungsschale (Asset Administration Shells, AAS) [In22]. Hierdurch ist ein

Datenaustausch entlang der digitalen Lieferkette gegeben und die Datenmodelle können in Form von Teilmodellen wiederverwendet werden. Die Datenmodelle ermöglichen zudem eine Zuordnung der Informationen zu den jeweiligen Verantwortungsbereichen. Es findet also eine Unterscheidung der TGH-Emissionen statt, die direkt oder indirekt der betrachteten Produktion zuzuordnen sind (vergl. Scope 1 und Scope 2) oder die der vor- bzw. nachgelagerten Lieferkette zugeordnet werden (Scope 3).

Der AAS-Standard gewährleistet die Interoperabilität der Datenmodelle und ermöglicht den Austausch von Datenmodellen zwischen verschiedenen Organisationen. Auf diese Weise kann ein nahtloser Datenaustausch innerhalb digitaler Lieferketten durch die Einbettung der PCF-Datenmodelle in die Digitalen Produktpässe (DPP) erleichtert werden. Um öffentliche und nichtöffentliche Informationen zu trennen, gibt es verschiedene Sichten auf die Datenmodelle der gDEZ. Externe Sichten enthalten öffentlich zugänglich Informationen wie aggregierte Daten zum CO₂-Fußabdruck und Recyclingdaten, welche zu einem späteren Zeitpunkt in Digitale Produktpässe (DPP) einfließen sollen. Interne Sichten enthalten Informationen über spezifische Parameter und Prozessinformationen auf Produktebene, die aus Wettbewerbsgründen nicht geteilt werden sollen. Die Europäische Kommission befürwortet Digitale Produktpässe im "Aktionsplan für die Kreislaufwirtschaft", wo sie als innovative Lösungen positioniert sind, um kreislaufwirtschaftliche Praktiken zu verbessern [EC20]. In dieser Arbeit legen wir den Fokus auf die gDEZ-Datenmodelle und führen ferner Anwendungsfälle sowie Berechnungsmodelle ein. Um die Logik hinter dem Datenmodell zu beschreiben und zu zeigen, dass die vorgeschlagenen Eigenschaften bestimmte Zwecke erfüllen, präsentieren wir nachfolgend zwei Anwendungsfälle, die wir zusammen mit unseren Industriepartnern identifiziert haben.

Anwendungsfall 1: Der Hersteller muss PCF-Daten an den Originalausrüstungshersteller (OEM) offenlegen. In diesem Anwendungsfall liegt der Fokus auf der Verbesserung von Transparenz und Nachhaltigkeit innerhalb der Lieferkette durch die vorgeschriebene Weitergabe des PCF von Herstellern auf der Zuliefererebene an Originalausrüstungshersteller (OEM). Dieser Anwendungsfall hebt die entscheidende Rolle produktspezifischer Daten guter Qualität hervor. Dies ist ein Schritt zur Erreichung von Nachhaltigkeitszielen und beantwortet die steigende Nachfrage nach Umweltverantwortung in der Industrie. Das Ziel dieses Anwendungsfalles ist es, auf die Notwendigkeit einer besseren Methode zur Erfassung und gemeinsamen Nutzung dieser Daten in einem gemeinsamen Format hinzuweisen, insbesondere im Hinblick auf den CO₂-Fußabdruck von Materialien und Komponenten. Dies ermöglicht es OEMs, fundierte Entscheidungen zu treffen, die die Umweltauswirkungen ihrer Produkte verringern. Der Ansatz unterstützt breitere Nachhaltigkeitsinitiativen, wie die Reduzierung von Treibhausgasemissionen, die Optimierung der Ressourcennutzung und die Förderung der Kreislaufwirtschaft. Darüber hinaus reagiert dieser Anwendungsfall geschickt auf die Anforderungen neu entstehender Nachhaltigkeitsvorschriften, wie sie beispielsweise

durch den Green Deal der Europäischen Union und die Richtlinie zur Nachhaltigkeitsberichterstattung von Unternehmen (CSRD) veranschaulicht werden.

Anwendungsfall 2: Optimierung des Produktionsprozesses für Energieeffizienz. Dies zielt auf die Verbesserung der Energieeffizienz in Produktionsprozessen durch Analyse und Optimierung des Energieverbrauchs, unterstützt durch gDEZ. Dieser Anwendungsfall unterstreicht die Bedeutung eines praktischen Ansatzes zur Datenaggregation als grundlegenden Schritt zur Erzielung messbarer Energieeinsparungen und Reduzierung des CO₂-Fußabdrucks in Produktionsprozessen. Wenn dies erreicht wird, können Ineffizienzen identifiziert und angegangen werden, was den Energieverbrauch und den CO₂-Fußabdruck von Produktionsbetrieben reduziert. Dieser Anwendungsfall ist besonders relevant als Antwort auf steigende Energiekosten, Nachhaltigkeitsvorschriften und unternehmerische Umweltverpflichtungen.

Um den PCF auf Produktebene angemessen zu bestimmen, reicht es nicht aus, nur die Informationen des Produkts selbst zu bewerten, wie beispielsweise Gewicht und Abmessungen. Auch die Aktivitätsdaten von Ressourcen und Prozessen sind relevant. In einer Fabrik werden Fertigungsprozesse von Ressourcen durchgeführt, die einen Mehrwert für die Produkte schaffen. Wir berücksichtigen auch die Energie, die von den Ressourcen verbraucht wird, um die Produkte in der PCF-Berechnung auf Produktebene zu verarbeiten. Darüber hinaus sind zur Bestimmung des Energiebedarfs eines Produkts auch relevante Prozessinformationen erforderlich. Daher ist es notwendig, relevante Informationen aus Produkt, Prozess und Ressource (PPR) [PSB13] in die Datenmodelle zu integrieren. Die folgenden Eigenschaften werden weiter iteriert und dienen als initialer Vorschlag, den wir mit unseren Industriepartnern validieren konnten. Um die Datenstruktur der Informationen, die durch die gDEZ dargestellt werden, zu identifizieren, muss das Szenario klar definiert sein. Die Datenmodelle beschreiben Fabrikinformationen, wie die verfügbaren Ressourcen, und wie sie mit dem Produkt, den Prozessen und seiner Umgebung interagieren.

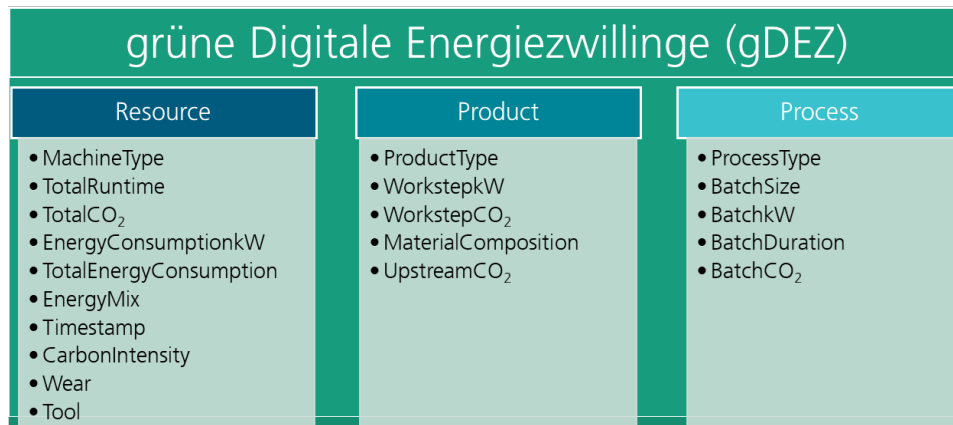


Abb. 1: Vereinfachte Darstellung des gDEZ-Datenmodells

Die Informationen des gDEZ-Ressourcenaspekts bieten einen umfassenden Einblick in die operativen Aspekte von Produktionsanlagen. Sie umfassen die folgenden Eigenschaften:

- **MachineType:** Beschreibt den Typ der Maschine oder Ressource.
- **TotalRuntime:** Gibt die Gesamtbetriebszeit der Ressource an.
- **TotalCO₂:** Zeigt die Gesamtemissionen von Kohlendioxid (CO₂) der Ressource an.
- **EnergyConsumptionkW:** Stellt den Energieverbrauch der Ressource in Kilowatt (kW) dar.
- **TotalEnergyConsumption:** Gibt den Gesamtenergieverbrauch der Ressource an.
- **EnergyMix:** Beschreibt die Zusammensetzung des Energiemixes, der von der Ressource verwendet wird.
- **Timestamp:** Gibt den Zeitpunkt an, zu dem die Daten erfasst wurden.
- **CarbonIntensity:** Stellt die spezifische CO₂-Emissionsintensität der Ressource dar.
- **Wear:** Beschreibt den Verschleißzustand der Ressource.
- **Tool:** Gibt an, ob die Ressource ein Werkzeug oder eine bestimmte Ausrüstung verwendet.

Diese Daten ermöglichen eine detaillierte Analyse des Energieverbrauchs, der damit verbundenen THG-Emissionen und der Betriebszeiten einzelner Ressourcen. Durch die Bereitstellung und Übergabe dieser Informationen können Hersteller Optimierungspotenziale erarbeiten, um die Energieeffizienz zu steigern und den CO₂-Fußabdruck der Produktion ihrer Produkte zu verringern. Darüber hinaus ermöglicht die Verfügbarkeit von Zeitstempeln eine zeitliche Analyse des Energieverbrauchs, was wiederum eine bessere Planung und Steuerung der Produktionsabläufe unterstützt.

Der Produktaspekt des gDEZ bietet eine detaillierte Analyse der Produkte in Bezug auf deren Herstellung und Zusammensetzung. Hier sind die zugehörigen Eigenschaften:

- **ProductType:** Beschreibt den Typ oder die Kategorie des Produkts.
- **WorkstepkW:** Gibt den Energieverbrauch des Arbeitsschritts für die Herstellung des Produkts in Kilowatt (kW) an.
- **WorkstepCO2:** Zeigt die damit verbundenen CO₂-Emissionen des Arbeitsschritts für die Herstellung des Produkts an.
- **MaterialComposition:** Beschreibt die Zusammensetzung der Materialien, aus denen das Produkt hergestellt ist.
- **UpstreamCO2:** Gibt die CO₂-Emissionen an, die bei der Herstellung der Vorprodukte oder bei vorangehenden Prozessen entstehen.

Diese Informationen ermöglichen eine detaillierte Analyse des Energieverbrauchs, der CO₂-Emissionen und der Materialzusammensetzung jedes Produkts.

Der Prozessaspekt des gDEZ liefert umfassende Einblicke in die verschiedenen Fertigungsprozesse, die zur Herstellung von Produkten verwendet werden. Hier sind die entsprechenden Eigenschaften:

- **ProcessType:** Beschreibt den Typ oder die Art des Fertigungsprozesses.
- **BatchSize:** Gibt die Anzahl der Einheiten an, die in einem Fertigungszyklus verarbeitet werden.
- **BatchkW:** Zeigt den Energieverbrauch der Charge während des Fertigungsprozesses in Kilowatt (kW) an.
- **BatchDuration:** Gibt die Dauer des Fertigungszyklus oder der Charge an.
- **BatchCO2:** Zeigt die damit verbundenen CO₂-Emissionen während des Fertigungsprozesses oder der Charge an.

Für eine praktikable **Berechnung** des CO₂-Fußabdrucks von Produktionsprozessen schlagen wir einen mehrstufigen Ansatz vor. Hiermit wollen die unterschiedlichen Fertigungsumgebungen, einschließlich der unterschiedlichen Grade der Digitalisierung und Integration. Um diesen unterschiedlichen Voraussetzungen zu begegnen und eine praktikable Anwendbarkeit sicherzustellen, grenzen wir mindestens drei unterschiedliche detaillierte Berechnungsmodelle voneinander ab, deren Konzepte wir folgend einführen.

Das **Direktemissionsmodell** konzentriert sich darauf, direkte Emissionen aus Fertigungsprozessen mithilfe von Aktivitätsdaten zu quantifizieren. Durch die Konzentration auf unmittelbare Emissionsquellen innerhalb der Produktion ermöglicht dieses Modell die Berechnung des CO₂-Fußabdrucks pro Batch (*BatchCO2*), durch das Nutzen des Energieverbrauchs der Ressource (*EnergyConsumptionkW*) und den Emissionsfaktoren im Strommix (*CarbonIntensity*).

Das **erweiterte Modell** integriert zu den direkten Emissionen, die Dimensionen von Existenz und Verschleiß. Dieser Ansatz berücksichtigt die Lebenszyklusauswirkungen von Geräten und berücksichtigt die Emissionen im Zusammenhang mit Herstellung,

Wartung und Stilllegung oder Ersatz. Durch die Berücksichtigung dieser Faktoren bietet das Modell eine genauere Sicht auf den CO₂-Fußabdruck der erzeugten Produkte.

Das **umfassende Modell** integriert zu dem erweiterten Modell, zusätzlich die Informationen aus Vorprozessen. Dies erstreckt sich über die Grenzen der Fabrik hinaus, um die Emissionen von Materialien und Komponenten zu erfassen, die in der Fertigung verwendet werden. Dieses Modell nutzt Lieferketten-Daten, um Emissionen aus Rohstoffgewinnung, Transport und Verarbeitung einzubeziehen.

4 Ausblick

In dieser Arbeit haben wir die gDEZ-Datenmodelle vorgestellt, die zur Modellierung der erforderlichen Aktivitätsdaten für die PCF-Berechnung benötigt werden. Für die Datenmodelle haben wir illustriert, welche Aktivitätsdaten notwendig sind und wie sie in den Datenmodellen dargestellt werden. Um die Datenqualität zu gewährleisten, sollten entsprechende Datenquellen integriert werden, um vernetzte Daten für die Eigenschaften in den Datenmodellen bereitzustellen. In einem weiteren Schritt wollen wir, basierend auf den gDEZ-Datenmodellen, unterschiedlich genaue Berechnungsmodelle zur Verfügung stellen. In der Praxis können Hersteller ein geeignetes Berechnungsmodell basierend auf dem Digitalisierungsgrad ihres Produktionssystems auswählen. Dies vermeidet unnötige Investitionen in die Digitalisierung und Aufrüstung des Produktionssystems. Die Anwendung der gDEZ-Modelle auf mehrere Produktionsschritte entlang einer Wertschöpfungskette gewährleistet konsistente PCF-Berechnungsergebnisse.

Die zukünftigen Arbeiten umfassen mehrere Aspekte. Zunächst werden wir die zweiten und dritten Berechnungsmodelle um Gleichungen erweitern, um die Berechnung zu konkretisieren. Gleichzeitig werden wir die PCF-Berechnung unter Berücksichtigung der Gemeinkosten wie Personalkosten erweitern, die nicht direkt einem einzelnen Produkt zugeordnet werden können. Zweitens werden wir das Berechnungsmodell um ein White-Box-Modell erweitern. Das White-Box-Modell berücksichtigt einen umfassenden Satz von Aktivitätsdaten für die Berechnung. Das White-Box-Modell hat daher eine hohe Genauigkeit. Allerdings ist die Implementierung eines White-Box-Modells in der Praxis meist aufgrund des begrenzten Digitalisierungsgrades eines Produktionssystems nicht möglich. Der Zweck des White-Box-Modells ist es, den Fehlerbereich der gDEZ-Berechnungsmodelle zu evaluieren. Wir werden eine Reihe von Experimenten einrichten, um die Fehler zu bewerten, indem wir die Berechnungsergebnisse des White-Box-Modells mit den bisherigen Berechnungsmodellen vergleichen. Bei den Berechnungsmodellen werden wir bestimmte Parameter weglassen oder sie durch Daten geringerer Qualität ersetzen und die Berechnungsergebnisse mit dem White-Box-Modell vergleichen. Basierend auf den Ergebnissen können wir feststellen, welche Parameter für die PCF-Berechnung relevant sind und mit hoher Genauigkeit bereitgestellt werden müssen.

Drittens werden wir die aktuelle Bewertung mit dem White-Box-Modell erweitern, um die relevanten Parameter zu identifizieren, die eine hohe Genauigkeit erfordern, sowie die Parameter, die weggelassen oder mit niedriger Genauigkeit bereitgestellt werden können.

Das White-Box-Modell kombiniert alle verfügbaren Informationen und dient als Baseline, um die Genauigkeit der Berechnungsmodelle festzustellen. Deshalb enthält das White-Box-Modell auch Informationen, die nur theoretisch verfügbar sind. Da es sehr aufwendig ist ein vollständiges White-Box-Modell aufzubauen, führen wir effiziente Berechnungsmodelle ein. Wir führen Metriken ein, um die Abweichungen zwischen den realen Berechnungsmodellen (White-Box-Modellen) und den idealisierten Berechnungsmodellen zu messen. Diese Metriken ermöglichen es, CO₂-Hotspots im Produktionsprozess zu identifizieren und zu bestimmen, welche Datenpunkte präzise gemessen und integriert werden müssen. Dies ist besonders relevant für die Identifikation von Bereichen, in denen geschätzte Werte ausreichend sind, also das Modell vereinfacht werden kann. Hierdurch wollen wir Berechnungsmodelle erstellen, die genau genug und gleichzeitig praktikabel sind. Bei unserem Praxispartner wollen wir die praktische Anwendbarkeit und Effektivität unserer Modelle im Kontext einer Pumpenproduktion evaluieren.

5 Literaturverzeichnis

- [Ch22] Chávez, C. A. G.; Barring, M.; Frantzén, M.; Annepavar, A.; Gopalakrishnan, D.; Johansson, B.: Achieving Sustainable Manufacturing by Embedding Sustainability KPIs in Digital Twins. In: 2022 Winter Simulation Conference (WSC), 2022.
- [EA23] Edrisi, F.; Azari, M. S.: Digital Twin for Sustainability Assessment and Policy Evaluation: A Systematic Literature Review. In: 2023 IEEE/ACM 7th International Workshop on Green And Sustainable Software (GREENS), 2023.
- [EC20] European Commission: Digital Product Passport: Facilitating the Circular Economy within the European Union, 2020. Verfügbar unter: [https://ec.europa.eu/environment/circular-economy/pdf/digital_product_passport.pdf] Zugriff am 13. April 2024.
- [Fi23] Fischer, K.: Kapitel 15. Globalisierung: Globale Warenketten und Arbeitsteilung. In (Görg, C., Hrsg.): APCC Special Report: Strukturen für ein klimafreundliches Leben. Springer Spektrum, Berlin, Heidelberg, 2023.
- [In22] Plattform Industrie 4.0: Details of the Asset Administration Shell: Part 1 - The Exchange of Information between Partners in the Value Chain of Industrie 4.0 (Version 3.0rc02). Plattform Industrie 4.0, 2022.
- [PSB13] J. Pfrommer, M. Schleipen und J. Beyerer, "PPRS: Produktionskompetenzen und ihre Beziehung zu Produkt, Prozess und Ressource," in International Conference on Emerging Technologies and Factory Automation (ETFA), Cagliari, 2013.

- [Ja20] Jabbour, C. J. C.; Fiorini, P. D. C.; Ndubisi, N. O.; Queiroz, M. M.; Piato, É. L.: Digitally-enabled sustainable supply chains in the 21st century: A review and a research agenda. In: *Science of the Total Environment*, Bd. 725, Art. Nr. 138177, 2020.
- [KSA20] Kuhn, T.; Schnicke, F.; Antonino, P. O.: Service-based Architectures in Production Systems: Challenges, Solutions & Experiences. In: *2020 ITU Kaleidoscope: Industry-Driven Digital Transformation (ITU K)*. IEEE, S. 1-7, Dezember 2020.
- [Si23] Singh, G.; Singh, S.; Daultani, Y.; Chouhan, M.: Measuring the influence of digital twins on the sustainability of manufacturing supply chain: A mediating role of supply chain resilience and performance. In: *Computers and Industrial Engineering*, Bd. 186, Art. Nr. 109711, 2023.
- [TA17] Tantik, E.; Anderl, R.: Potentials of the Asset Administration Shell of Industrie 4.0 for Service-Oriented Business Models. In: *Procedia CIRP*, Bd. 64, S. 363-368, 2017.

Modelling of NOA 178 Verification of Request concept using Pro-Active Asset Administration Shells

H.K. Pakala*, Ch. Diedrich*

**Otto von Guericke University Magdeburg, Magdeburg, LSA 39106
Germany (e-mail: harish.pakala@ovgu.de)
Germany (e-mail: christian.diedrich@ovgu.de).*

Abstract:

The NAMUR automation pyramid is the current state of the art in the process control industry. The pyramid is a monolithic structure with information flowing either in top-down or bottom-up directions. Introducing new technologies into this structure would imply modifying or replacing existing equipment and the control software. NE 175 opens up this closed structure with a new open architecture (NOA) to provide access to external interfaces without changing the existing structure safely, securely and reliably. In continuation of the NOA, the NE 178 presents an architecture for the verification of the request concept (VOR) to allow plant-specific M+O applications to make requests for a change in functioning of the plant by-passing the strong hierarchy of the pyramid. This paper attempts to model the actors involved in this concept and their interactions using the Pro-Active AAS concept, it introduces a new interaction protocol and demonstrates the concept using the single water tank case study.

Keywords: automation systems, cyber physical systems, Asset Administration Shell, NOA, verification of request, automation pyramid

1. Introduction

Normen-Arbeitsgemeinschaft für Mess-und-Regeltechnik in der Chemischen Industrie (NAMUR) automation pyramid is the traditional, well-accepted and proven structure within the process control industry. This structure emphasizes and ensures a hierarchical information flow from the field devices at the bottom to the MES at the top, as shown in Figure 1. The field devices, programmable logic controllers, SCADA and the MES applications within the four levels of the pyramid, as pointed out in Figure 1, are tightly coupled and interrelated. Introducing new technologies into the existing plant would require modifications at all levels by the appropriate manufacturers, involving huge man-hours and financial investment by all the players involved.

Within the Information technology (IT) industry during the last two decades, there were significant advancements like better user experience interfaces, virtual reality, blockchain, enterprise integration, ontologies and the semantic web, automation verification of the systems, Chat GPT and many more. Combined with the Artificial Intelligence sectors like energy, health, transportation and supply chain are integrating and benefitting from the IT advancements. The article (Klettner et al. 2017) points out that within the automation sector in reference to the NAMUR ecosystem, there is no significant adoption of IT technologies. In this context, one of the NAMUR working groups has published the (NE 175), which introduces the concept of NAMUR open architecture (NOA). NOA is safe, secure, and reliable and intends to open up the automation pyramid to introduce the latest features and products from the IT industry. The (NE 175) professes that NOA is open, but the pyramid structure stays intact and the scope for the new technologies is rather parallel to the existing structure.

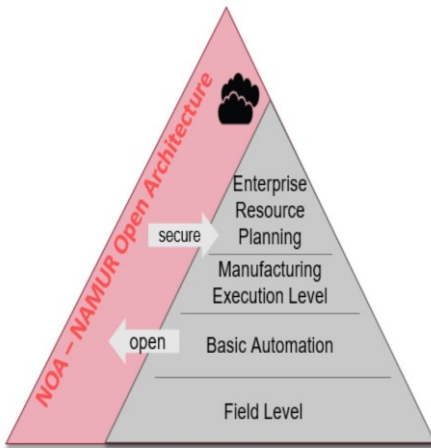


Figure 1 NOA Pyramid (Klettner et al. 2017)

NE 178 further strengthens the idea of the autonomy of the plant and the plant owner and introduces an architecture for the NOA Verification of Request (VOR) concept. The actual document for NE 178 is in draft stage and the details of the VOR concept are published in (Iatrou et al. 2022). It is the systems of systems architecture for the transmission of information between the plant and the interfaces of new IT infrastructure. The aim of the VOR concept is to systematically allow external applications to make requests for a change in the functioning of the plant and accordingly, the NE 178 identifies the actors involved and defines their roles and responsibilities. The need for structured requests and the relevant feedback messages, information about security and authentication mechanisms, step-by-step processing of the incoming requests, and the logging of the requests such that everything gets tracked within the entire system and other aspects are outlined by NE 178 as described in (Iatrou et al. 2022)

(Specification IEC 16593-1) describes an Industry 4.0 (I4.0) component as a combination of multiple sub-components (I4.0). Here, each sub-component or the component as a whole has a specific behaviour. The I4.0 component interacts and reacts to the events generated within its environment. The behaviour of these I4.0 components may be modelled, using explicit procedures or finite-state machines, as shown in Figure 2.

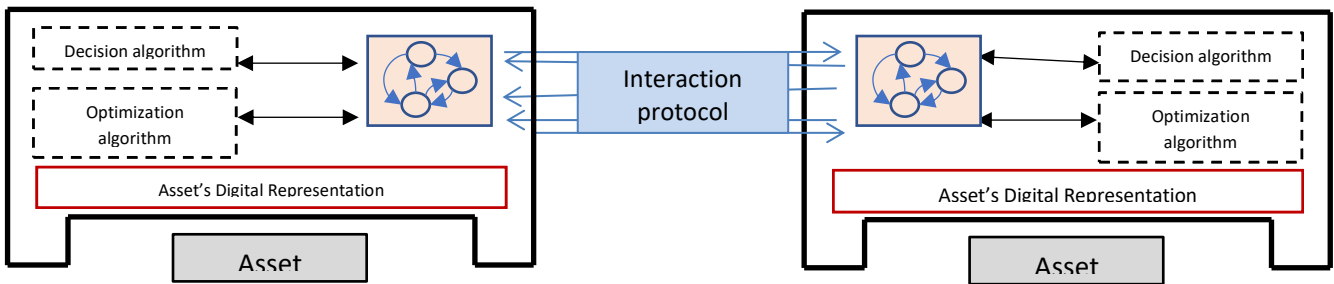


Figure 2 I4.0 environment consisting two Pro-Active AAS each consisting state machines modelling respective behaviours

(INSTITUT FÜR AUTOMATISIERUNGSTECHNIK 2019) address the aspect of pro-active asset administration shells (AAS) furthering the idea of type 3 AAS as presented in (Specification 01002-3-0). Pro-Active AAS are the I4.0 components that communicate with each other using the I4.0 language semantics (VDI-Richtlinien) and have the capability to take independent decisions. In this article, we aim to demonstrate the VOR Systems of Systems architecture using a set of Pro-Active AAS each communicating with each other over a defined interaction protocol. Accordingly, this paper introduces a new interaction protocol with relevant sequence diagrams. A prototype implementation of VOR architecture is presented with CPC domain represented by a single water tank case study (Larsen et al. 2016).

The rest of the paper is organized into seven sections. Section 2 presents a brief summary of NOA concept, AAS and the interaction protocols. Section 3 presents a detailed reported on the VOR concept as presented in (Iatrou et al. 2022). Section 4 presents a description of the single water tank use-case introduced in (Larsen et al. 2016). Section 5 demonstrates the implementation of VOR concept using Pro-Active AAS. Section 6 presents a brief overview of literature of past works. Section 7 concludes the paper.

2. Background

NAMR Open Architecture (NOA)

The NAMUR pyramid shown in Figure 1 is a monolithic and closed structure, with four different layers. As a system integrator, a plant owner has to own the responsibility of the interconnection between different levels, making his job more important and, the same applies to the device and the control software manufacturers. Such a hierarchical structure enables tight coupling between manufacturers of different products required for automation and has worked well over many decades. With the pyramid structure

in place, for new technology to be integrated the existing devices should be upgraded, modified or replaced by the new ones, involving high investment from all the players involved.

In such a background, (NE 175) just like its previous documents, is introduced by one of the NAMUR working groups. The NE 175 document introduces the NAMUR open architecture (NOA) concept. The primary goal of NOA is the introduction of modern IT technologies into the automation architectures without compromising the safety, security and reliability of the existing and future plants. The NE 175 clearly states that the ability to add new technologies is only complementary to the existing structure and the intent is not to replace the existing structure. In Figure 1, the tilted bar on one of the sides of the automation pyramid proposed by NE 175 acts like a plugin to the existing one. For the

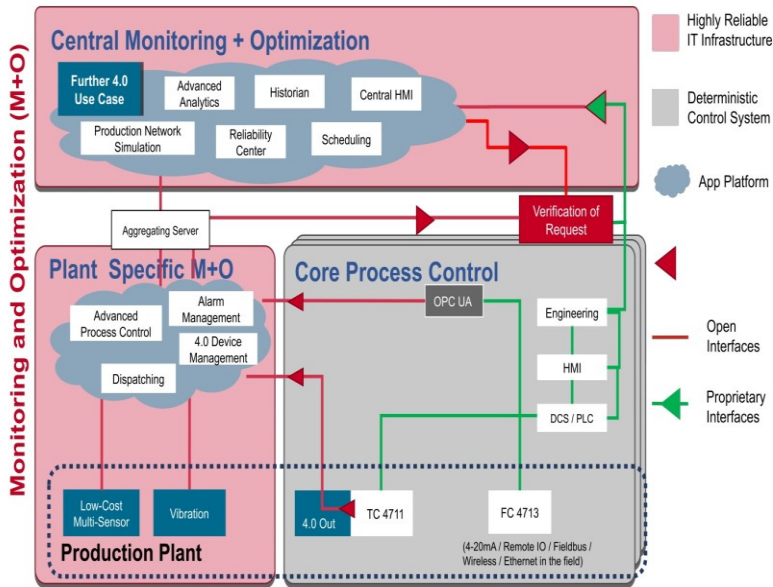


Figure 3 NOA Architecture

realization of such an idea, the NE 175 introduces a new architecture (as shown in Figure 3) with the concept of three communicating domains, namely the Core Process Control (CPC), Plant Specific Monitoring and Optimization (psM+O) and, Central optimization and monitoring (M+O). The CPC domain is the automation pyramid itself within the plant. The psM+O is the communication domain owned by the plant owner that is responsible for transmitting of information from the CPC domain to the external interfaces and acts like the demilitarized zone for the plant, where the plant owner has clear control over what kind of information enters and leaves the CPC domain. The M+O is the actual domain purported to be the additional bar, presented in Figure 1, that aims to introduce the technological advances from the IT industry into the automation pyramid.

Asset Administration Shell (AAS)

The concept of Asset Administration Shell, referred to as the standardized digital representation of an asset, is introduced by Platform Industry 4.0 (Plattform Industrie 4.0 2016). It aims to standardize the exchange of information between different partners, primarily within the automation industry. In order to achieve this, a standardized information model is introduced in (Specification 01001-3-0) with a core element called the Submodel that could contain elements of different types (Property, Range, MultiLanguageProperty, SubmodelElementCollection (SMC)). The term Submodel in the context of encapsulation of information about an asset could represent either a Nameplate or a Technical Data or a Documentation or asset access-related information. The information about an asset modelled using the AAS information model could be exchanged as files (Type 1 AAS), or a software component can consume the data and provide standardized interfaces for access (Type 2 AAS) (Specification 01002-3-0).

Interaction Protocols

(VDI-Richtlinien) characterizes the interactions of an active AAS with its environment as interaction protocols (IP) with I4.0 language format. IP is a structured sequence of interactions (or the exchange of messages) between I4.0 components, attaching defined protocols to the interactions makes the communication restricted to a set of message types. Where an I4.0 message as defined in the (VDI-Richtlinien) is organized into two parts frame (representing sender and receiver information, message type, conversation and messages IDs and a reference to the protocol definition) and the interaction elements (constituting the actual data).

3. NE 178 Verification of Request Architecture

In this section, we present different steps of the NE 178 as presented in (Larsen et al. 2016).

The architecture identifies three main actors within the scope of a plant in a process control industry, namely the requester issuer (this could be either from the psM+O or M+O domain, as pointed out in section 2), the VOR server and VAS server (VAS). NE 178 defines a sequence of steps, starting with, authentication until the request execution and provides guidelines on who (the specific actor) should perform them.

Authentication and Authorization Step: Every request issuer (RI) shall authenticate itself with the VOR server and then directly issue the requests. The VOR server first authorizes the RI against its credentials for every incoming request. Successful authentication would result in VOR sending feedback to the RI. Feedback is not issued for the failed authentication step. The credentials may be just the user ID and password or RI-specific certificates and, these may be managed using existing state-of-the-art solutions.

Verification Step: The VOR server first verifies the authenticated request, probably using a set of Boolean expressions. The plant owner is responsible for specifying these expressions.

Mapping Step: A request from the M+O user is more on a higher-level construct, indicating that the issuer does not know the internal dynamics or the operation of the plant. For example, a request from psM+O or M+O domain could be to switch off the valve that inlets the water into the tank. A set of endpoints represents the CPC domain, through which the specific aspects can be accessed and modified. Either the VOR server or the VAS maps an authorized request to the suitable CPC endpoints, in this paper we model the VAS server to perform this task.

Acceptance Step: This step checks whether the current state of the plant is valid and whether the plant is prepared enough to execute the request. One of the situations where a request is not accepted could be, for example, the absence of the operation personnel on the plant who is required during the execution process for monitoring and logging environment situations. NE 178 specifies that efforts must be made to automate the acceptance and not bring in human interference.

Mapping Verification: The mapped request is then checked for relevance if it requires execution, for example, does the specific valve within the plant need to be switched off? In case a single request consists of changes to multiple CPC endpoints then the request needs to be executed only if all the changes are possible or else the request is completely rejected. This step finalizes the entire step and forwards the changes to the CPC endpoints.

Feedback Processing: Any request processing step can generate a feedback message. Every feedback submitted by an actor in a specific step is first verified and then dispatched. A feedback created by the VOR server will be dispatched to the RI and the feedback created by VAS will be dispatched to the VOR server only. There is no direct interaction between the VAS and the RI. At every step, the request and the feedback messages are timestamped and stored for traceability and verification in future.

4. CaseStudy

Figure 4 shows the P&ID diagram of a process automation plant with a single water tank, a use case as presented in (Larsen et al. 2016). A pump supplies water to the tank at a specific rate, and the valve (0041) attached between the tank and the pump controls the water flow into the tank. The tank has an outlet to drain the water, and another valve (0043) controls the water flow leaving the tank. The water inlet into the tank and the draining may not happen simultaneously. The two valves in Figure 4 are digitally accessible and have an OPC UA interface. The valves may be closed or opened by updating a boolean variable over the OPC UA interface. A level sensor is attached to the water tank, records the water level (in quantity) in the tank and sends the values as messages to the controller component at regular intervals, not shown in Figure 4. The controller component maintains the water level within maximum and minimum thresholds. Every time the controller receives a message from the level sensor, it checks the level against the threshold. If the value equals (maximum - 1), valve 0041 is closed and valve 0043 is opened. If the value equals (minimum + 1), valve 0041 is opened and valve 0043 is closed. The level sensor attached to the tank has an OPC UA interface, a single variable representing the water level in the tank.

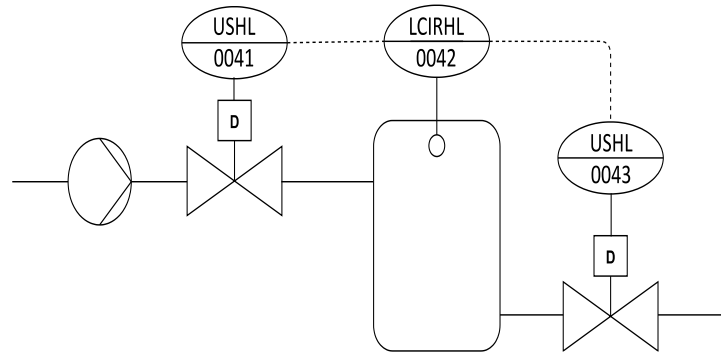


Figure 4 P&ID for the Single Water Tank (SWT) use-case

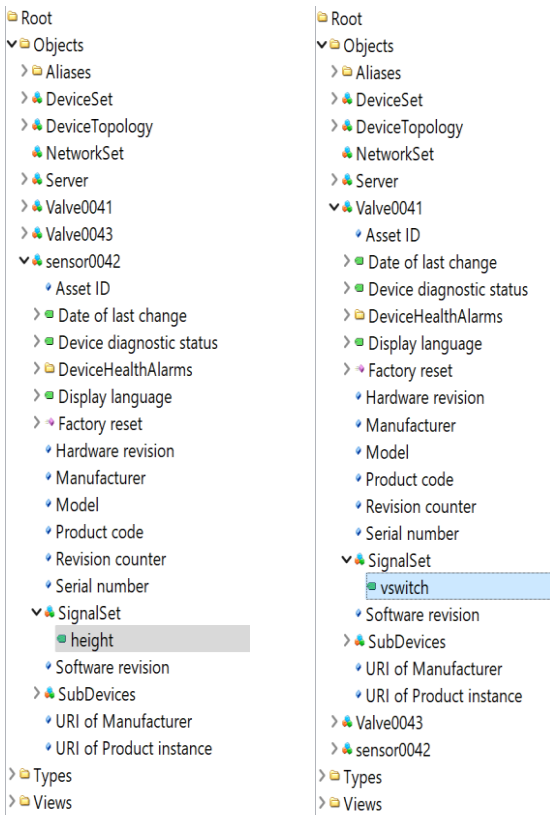


Figure 5 OPC Server representation for the single water tank use case

Figure 5 shows the screenshot from the UAExpert (OPCUA client) interface. The two valves and the level sensor for each, an object of type 'PADIMType' is instantiated. The sensor0042 has the variable 'height' of type LevelMeasurementVariableType under the SignalSet object, it represents the water level in the tank (with real-time updates). The valves 0041 and 0043 have the variable 'vswitch' of type boolean under the 'SignalSet' object and are used to open/close the respective valves. The PADIM-OPCUA companion specification can be accessed from (OPC Foundation 2024). Figure 5 indicates for the three devices, only one namespace is used, representing the entire plant. Based on the NE 175 guidelines, only read-only access to the OPCUA server is provided to the external M&O applications.

In this paper, based on guidelines from NE 178, we present a framework based on the concept of proactive AAS to access the functionality of the controller unit safely and securely. We suppose a scenario in which, due to unforeseen circumstances, external M&O applications may initiate a request to switch off the supply to the water tank. To this scenario we add an additional check, the water level in the tank shall always be above the average of maximum and minimum thresholds, when the source of the water is unexpectedly closed for-ever.

5. Realization of the NE 178 Concept for single water tank use case using Pro-Active AAS

In this section, we demonstrate the NE 178 concept using the single water tank use case, In this context, a new interaction protocol is introduced.

In this paper, we utilize a set of Proactive AAS to demonstrate the NE 178 VOR concept. We model the NE 178 actors RI, VOR server, and VAS server as individual Pro-Active AAS. These representations do not represent a specific asset but encapsulate defined processes. Each AAS has a unique identifier and documents the information about the underlying process using the submodels. Figure 6 presents the screenshot of the package explorer (PE), it shows three AASs, and each AAS has a set of submodels. The RI AAS has three submodels, namely RequesterInformation, Requests and FeedBacks@Vor. The RequesterInformation submodel contains an identifier for the RI and its security-related credentials. The Requests submodel contains a list of SMCs, each representing a request issued by the RI. The FeedBacks@Vor submodel also contains a sequence of SMCs, each representing the feedback received by the RI from the VOR server. The VOR server AAS has four submodels, namely, a) 'Requests@RI', which contains a list of SMCs, each representing the requests received from RI. b) Requests@Vor, which contains a list SMCs, each representing the requests sent to the VAS server. c) FeedBacks@RI, which contains a list SMCs, each representing the feedback sent to RI. d) FeedBacks@Vas contains a list of SMCs, each representing the feedback received from the VAS server. The VAS server AAS has three submodels namely, a) Requests@Vor, which contains a list SMCs, each representing the requests received from the VOR server, b) FeedBacks@Vor, which contains a list of SMCs, each representing the feedback sent to the VOR server.

Env "AdministrationShells"	
▲	AAS "RequestIssuer" [https://example.com/ids/sm/1590_6070_5042_0183] of [https://e
	Asset AssetInformation https://example.com/ids/asset/8533_9080_6022_7862
▶	SM "RequesterInformation" [https://example.com/ids/sm/9454_1121_2122_6308]
▶	SM "Requests" [https://example.com/ids/sm/1545_1121_2122_4807]
▶	SM "FeedBacks@Vor" [https://example.com/ids/sm/1590_6070_5042_2104]
▲	AAS "VORServer" [https://example.com/ids/sm/3441_6070_5042_1861] of [https://exan
	Asset AssetInformation https://example.com/ids/asset/8533_9080_6022_7862
▶	SM "Requests@RI" [https://example.com/ids/sm/8401_6070_5042_4406]
▶	SM "Requests@Vor" [https://example.com/ids/sm/5324_0121_2122_4530]
▶	SM "FeedBacks@RI" [https://example.com/ids/sm/5041_1121_2122_9789]
▶	SM "FeedBacks@VAS" [https://example.com/ids/sm/8421_6070_5042_8446]
▶	SM "RequestTypes" [https://example.com/ids/sm/3441_6070_5042_2861]
▲	AAS "VirtualAutomationServer" [https://example.com/ids/sm/2431_6070_5042_1789]
	Asset AssetInformation https://example.com/ids/asset/8533_9080_6022_7862
▶	SM "Requests@VOR" [https://example.com/ids/sm/5282_8121_2122_2515]
▶	SM "NodeIds@VoR" [https://example.com/ids/sm/6164_0282_1032_5955]
▶	SM "FeedBacks@VAS" [https://example.com/ids/sm/2431_6070_5042_3789]

Figure 6 Screen shot from the Package Explorer representing three AAS and their respective submodels

Similar to the interaction protocol presented in (INSTITUT FÜR AUTOMATISIERUNGSTECHNIK 2019), in this section, we present a protocol representing a series of interactions between the three NE 178 actors, as shown in the sequence diagram from Figure 7. The behaviour of each AAS within the context of the interaction protocol is modelled using the finite state machines (FSM). A specific interaction between any two actors involves the transfer of a message from one actor to the other and is identified by a unique message type. The labels identifying the directed arrows in Figure 7 are the message types. The sequence of the message types between all the actors makes up the interaction protocol. Each message exchanged between the actors over a specific interaction is an I4.0 message as described in (VDI-Richtlinien). The actual content of each I4.0 message is either a submodel or any of the AAS submodel elements.

Authentication and Authorization: As part of the SWT demonstrator, we have maintained a single authentication server. Each Pro-Active AAS representing the NE 178 actor is provided with a certificate. The Pro-Active AAS uses this certificate to get an access token from the authentication server, it then uses it for interaction with the other AAS. Every un-authenticated interaction would be rejected.

We have used PythonAASxServer framework for the implementation of the Pro-Active AAS (Harish Kumar Pakala 2024). A detailed description of the framework is provided in (Weiss et al. 2023). The Framework provides an order management interface, this is used for creating a request.

The interaction protocol in the Figure 7 begins when the RI, receives an order to create a new request. The RI then creates a request message, that contains an SMC (request1) constituting the elements, identifier, description, impact, the priority of the request and the timestamp of the request creation, as shown in Figure 8 and transmits the message to the VOR server.

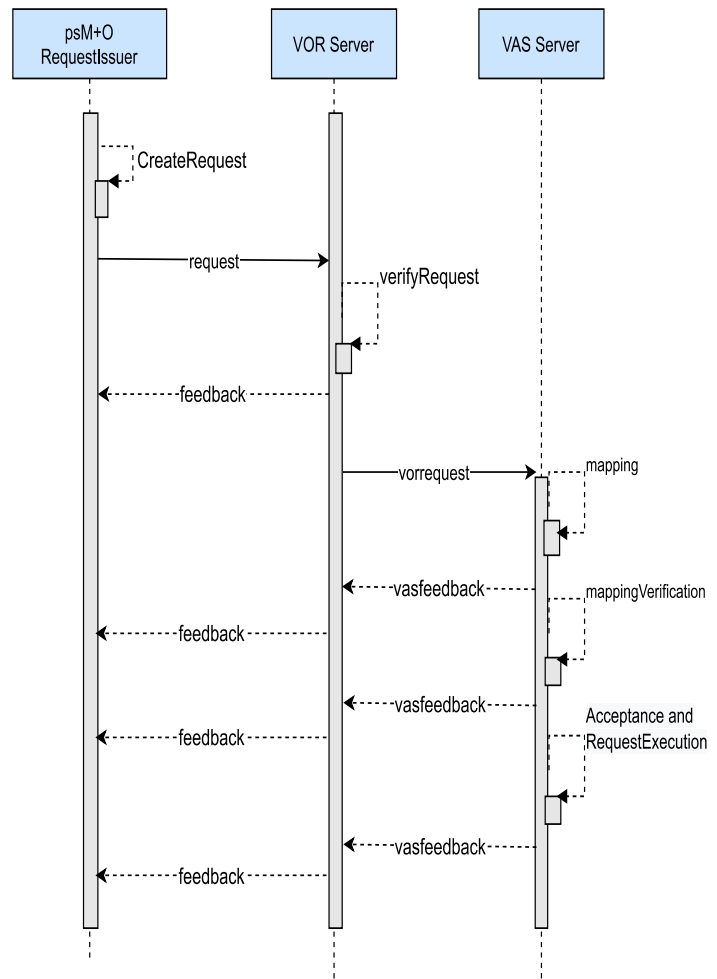


Figure 7 Sequence Diagram representing the VOR concept

SMC "Request1" (7 elements)
Prop "UniqueRequestIdentifier" = ABC-123-XXX-456
Prop "description" = Close the valve that regulates the water flow into the tank
Prop "impact" = Medium
Prop "priority" = High
Prop "GenerationTimestamp"
SMC "Request2" (7 elements)

Figure 8 AAS PE screen shot depicting a Request from RI

The RI may issue only two request types “Close the valve that regulates the water flow into the tank” and “Open the valve that regulates the water flow into the tank”. The VOR server first verifies the originator of the request message, it is the eligibility test for the RI to issue a specific request. For every verified request, the VOR server notifies the RI with a feedback message. The VOR server interprets the description of the message and turns it into an executable request, as shown in Figure 9. We suppose that the VOR server has complete knowledge of the structure of the plant. Figure 9 shows that valve0041 needs to be closed. The valve0041 is mentioned under the parameters SMC and, the change to the switch associated with the valve is mentioned under the modification SMC. The VOR server creates a new message of type ‘vorrequest’ and attaches to it the Request1 SMC shown in Figure 9.

In the case of the SWT demonstrator use case, suppose that the water supply into the tank needs to be closed due to some external reasons and to stop the flow of water into the tank. The RI specifies this information in the description property element as shown in the Figure 8. The priority of this request is set to high and, the impact of this request execution is set to medium. The NE 178 mandates that every request generated by the RI needs to be logged and accordingly, the Request1 SMC is added to the 'Requests' submodel.



Figure 9 AAS PE screen shot depicting a Request from RI

If yes, the request is rejected indicating the mapping verification is failed, and negative feedback is sent to the VOR server. In case the level is above the threshold, the request is accepted as mapping verified and pushed for execution.

In the execution step, both values 0041 and 0043 are closed one after the other immediately without any delay and positive feedback is sent back to the VOR server (assuming that modifications of the OPC UA variables and the changes to the underlying assets are precise without any errors). Upon receipt of the 'vasfeedback' message from the VAS server, the VOR server examines the message and send an appropriate message of type 'feedback' to the RI. Figure 10 shows the structure of a vasfeedback from the VAS server, screen shot from the package explorer. The intent of the message is specified as the 'implemented', indicating that the the change request is executed succesfully.

Benefits of using Pro-Active AAS

- Figure 11 shows the submodel RequestTypes representing the type of requests possible by the VOR concept (part of VOR server AAS). The plant owner may dynamically edit this submodel and delete or add new request types.
- The plant owner may also specify a mapping between Request Issuers and Request Types,

The VAS server examines the incoming vorrequest and attempts to map the request to an executable action, for example, to modify a specific variable from the OPCUA server. In case the mapping step fails, negative feedback of type 'vasfeedback' is sent to the VOR server. In the SWT use case, the request 'close switch related to the Valve0041' is mapped to nodeld of the variable switch, with a change to the variable as 'closed'. The list of the variables and related nodeld mappings are part of the VAS Server AAS shown in Figure 6. After the successful mapping step, as there is no requirement for personal in SWT use case the request is accepted and an acceptance notification is sent to the VOR sever.

In the SWT use case, it is assumed that the valve0041 may be closed only when it results in a scenario where the water level is above the average of maximum and minimum thresholds. The VAS server obtains the water level in the tank by reading the relevant OPCUA variable and checks if the level is already below the average.

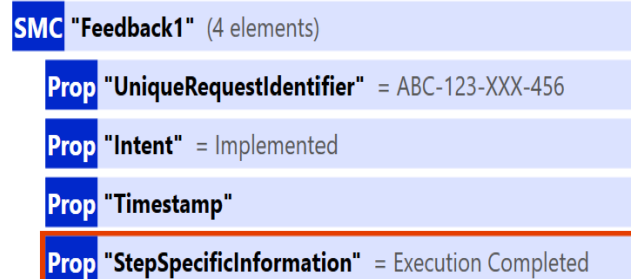


Figure 10 AAS PE screen shot depicting a Feedback from VAS Server

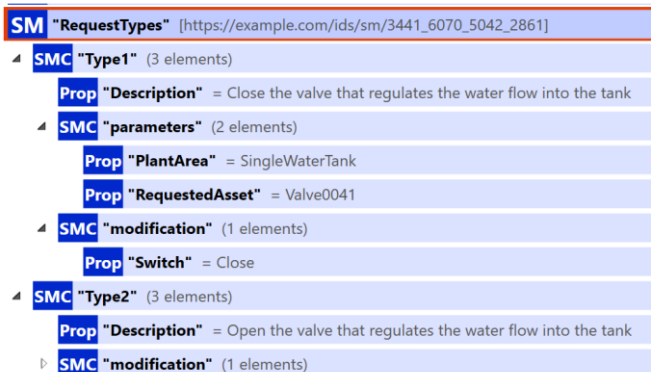


Figure 11 AAS PE screen shot depicting the Request Types from VOR Server AAS

allowing him to decide to deny a specific RI for a specific Request Type.

- All the messages exchanged within one execution of an interaction protocol are logged as submodel elements with respective AAS, as shown in Figure 6.

Single source of the truth for the Plant Owner.

- Apart from the central authentication server, role-based access to the submodels may be granted.

Access to information may be restricted tightly among the three NE 178 actors and also the external applications.

- Each asset shown in Figure 4 may be represented by AAS. Documenting technical properties of the asset using a submodel, may allow the plant owner to change, for example, the threshold values for the water level in the tank at any instant in time.

As part of the SWT demonstrator, we have modelled the assets shown in Figure 4 as AAS, the screenshots are not included in this paper.

FSM implementations for each of the NE 178 actors are implemented in the PythonAASxServer (Harish Kumar Pakala 2024) framework as pointed out in section 5. These FSMs fetch the information from the AAS repositories dynamically, every time they perform a specific operation. FSM implementations are cyclic, this allows the creation of new requests after completing another request without a change in the configurations.

6. Related Work

This section provides a short summary of the previous work within the domain of digital twins, industry 4.0 solutions for control within the automation pyramid, our intention is just to provide a flavor of approaches adapted by different works. (Azarmipour et al. 2020) aims to attach a digital twin (DT) at all levels of the automation pyramid. In the article, the authors have shown an example of DT at the process control level. The basic idea is that the DT requests information from the process level, performs the required simulations or optimizations and returns the result to the MES whenever it is requested.

To ensure the security of the pyramid structure, the authors assert the architecture enables a unidirectional flow of information from the control level to the DT, indicating that the DT will never have the opportunity to send any control signals to the process level. (Martinez et al. 2021) proposes a DT structure for the entire automation pyramid and introduces a virtual model to all the level and within in this context aims to leverage the benefit the AI. (Pessl and Rabel 2022) presents a use-case for dissolution of the automation pyramid within the industrial environment, where dissolution here means that instead of a top-bottom (and vice-versa) flow of information it would be interactions between all the level. To the best of the knowledge of the authors there is no paper which deals with a multi-phase write access to an automation component such a field device from an external application.

7. Conclusion

Industrial automation system migrates from a rigorous hierarchical architecture to an open system. In the process control domain this is accompanied by the NOA concept. It starts with a strict read only approach and is nowadays extended by a secure way of modifying parameters in automation devices. This concept is called "Verification of Request" (VOR). While defining the requirements in the according "NAMUR Empfehlung" NE 178 this paper propose a specification which fulfill the requirements. This is implemented as prototype using the Asset Administration Shell (AAS) which is a standardized specification of digital twins. The VOR is not restricted to the process control domain, it is suitable for a large variety of applications where a secure write access to the automation system is needed

8. Acknowledgement

This work of the project "Administration Shell Networked - Interoperability between I4.0 Components" was supported by the Federal Ministry for Economic Affairs and Energy in the Industry 4.0 initiative under the following numbers: FKZ: 13I40V001A.

Publication bibliography

Azarmipour, Mahyar; Elfaham, Haitham; Gries, Caspar; Kleinert, Tobias; Epple, Ulrich (2020): A Service-based Architecture for the Interaction of Control and MES Systems in Industry 4.0 Environment. In : 2020 IEEE 18th International Conference on Industrial Informatics (INDIN): IEEE.

Harish Kumar Pakala (2024): PythonAASxServer: Github. Available online at <https://github.com/harishpakala/PythonAASxServer>, checked on 4/26/2024.

Iatrou, Chris Paul; Hoppe, Henry; Erni, Klaus (2022): NOA Verification of Request: Automatisierte Anlagenoptimierung mit Feedback aus Edge und Cloud. In *atp* 64 (1-2), pp. 78–84.

INSTITUT FÜR AUTOMATISIERUNGSTECHNIK (2019): Specification Testbed "AAS networked". Proactive AAS - interaction according to the VDI/VDE 2193. INSTITUT FÜR AUTOMATISIERUNGSTECHNIK.

Klettner, Christian; Tauchnitz, Thomas; Epple, Ulrich; Nothdurft, Lars; Diedrich, Christian; Schröder, Tizian et al. (2017): Namur Open Architecture. In *atp* 59 (01-02), p. 17. DOI: 10.17560/atp.v59i01-02.620.

VDI-Richtlinien, 04 2020: Language for I4.0 Components - Structure of messages.

Larsen, Peter Gorm; Fitzgerald, John; Woodcock, Jim; Fritzson, Peter; Brauer, Jörg; Kleijn, Christian et al. (2016): Integrated tool chain for model-based design of Cyber-Physical Systems: The INTO-CPS project. In : 2016 2nd International Workshop on Modelling, Analysis, and Control of Complex CPS (CPS Data). IEEE, pp. 1–6.

Martinez, Edwin Mauricio; Ponce, Pedro; Macias, Israel; Molina, Arturo (2021): Automation pyramid as constructor for a complete digital twin, case study: A didactic manufacturing system. In *Sensors* 21 (14), p. 4656.

NE 175, 7/9/2020: NAMUR Open Architecture NOA Concept. Available online at <https://www.namur.net>, checked on 6/6/2023.

OPC Foundation (2024): OPC 30081: Process Automation Devices - PADIM (1.01). Available online at <https://reference.opcfoundation.org/nodesets/129>, updated on 2/1/2024, checked on 4/18/2024.

Pessl, Ernst; Rabel, Birgit (2022): Digitization in Production: a Use Case on a Cloud-based Manufacturing Execution System. In : 2022 8th International Conference on Computer Technology Applications. New York, NY, USA. New York, NY, USA: ACM.

Plattform Industrie 4.0 (2016): Fortschreibung der Anwendungsszenarien der Plattform Industrie 4.0. Ergebnispapier. Bundesministerium für Wirtschaft und Energie (BMWi).

Specification IEC 16593-1: RM-SA - Reference Model for Industrie 4.0 Service architectures.

Specification 01001-3-0, April 2023: Specification of the Asset Administration Shell. Available online at <https://industrialdigitaltwin.org/>, checked on June 2023.

Specification 01002-3-0, April 2023: Specification of the Asset Administration Shell, checked on June 2023.

Weiss, Marco; Wicke, Kai; Wende, Gerko; Pakala, Harish; Gill, Milapji (2023): MaSiMO-Development and Research of Industry 4.0 Components with a Focus on Experimental Applications of Proactive Asset Administration Shells in Data-Driven Maintenance Environments.

„Ulrich Jumar, Christian Diedrich (Hrsg.):
EKA 2024 - Entwurf komplexer Automatisierungssysteme, 18. Fachtagung“

Integration von Zeitreihendaten in die Verwaltungsschale

Pragmatische Umsetzung des Standards für die Anwendung in der Automatisierungsindustrie


Aaron Zielstorff¹, Dirk Schöttke², Fiona Helena Büttner³, Thomas Kämpfe⁴ und Stephan Schäfer⁵


Abstract: Die Digitalisierung von Prozessen führt zu einer Zunahme bei der Generierung umfangreicher Datenmengen. Diese ermöglichen weitere Analyseverfahren, wie zum Beispiel die Optimierung von Prozessparametern, die Erkennung von Anomalien, die Steigerung der Maschineneffizienz und die effizientere Planung von Wartungsarbeiten. Eine zentrale Herausforderung dabei ist die pragmatische Integration dieser Informationen in bestehende und zukünftige Systemlandschaften. Hierfür bietet sich die Nutzung der Verwaltungsschale (AAS) unter Verwendung der „Time Series Data“ Teilmodellvorlage an. Sie stellt einen einheitlichen Standard für die Integration und semantische Beschreibung von Zeitreihendaten innerhalb der Verwaltungsschale dar. Der vorliegende Beitrag illustriert die Einbindung von Zeitreihendaten in die AAS-Umgebung zur Wertschöpfung aus generierten Prozessinformationen.


Keywords: Digitaler Zwilling, Verwaltungsschale, Zeitreihen, Historische Daten, Digitale Wertschöpfung


1 Einleitung


Im Rahmen der Digitalisierung nimmt die Integration und effiziente Nutzung von Daten eine Schlüsselrolle in der Automatisierungstechnik ein. Die Verwaltungsschale (AAS) ist in diesem Kontext eine zentrale Komponente, die es als Standard ermöglicht, einen lebenszyklusübergreifenden Austausch von Informationen zu einem Asset zu gewährleisten [Dr23]. Die stetige Zunahme von Datenquellen, unter anderem ausgelöst durch die starke Zunahme von „smarten“ Sensoren und kostengünstigen

¹ HTW Berlin, FB1, Wilhelminenhofstraße 75A, 12459 Berlin, Aaron.Zielstorff@htw-berlin.de, 
<https://orcid.org/0009-0001-2476-8415>

² HTW Berlin, FB1, Wilhelminenhofstraße 75A, 12459 Berlin, Dirk.Schoettke@htw-berlin.de, 
<https://orcid.org/0009-0001-9401-9698>

³ HTW Berlin, FB1, Wilhelminenhofstraße 75A, 12459 Berlin, Fiona.Buettner@student.htw-berlin.de, 
<https://orcid.org/0009-0005-6993-8368>

⁴ HTW Berlin, FB1, Wilhelminenhofstraße 75A, 12459 Berlin, Thomas.Kaempfe@htw-berlin.de, 
<https://orcid.org/0009-0007-5703-0226>

⁵ HTW Berlin, FB1, Wilhelminenhofstraße 75A, 12459 Berlin, Stephan.Schaefer@htw-berlin.de, 
<https://orcid.org/0009-0008-6198-1582>

Speicheroptionen, führt zu einer wachsenden Generierung von Zeitreihendaten [SHS18]. Diese Daten bilden ein großes Potential für die Wertschöpfung, indem sie tiefere Einblicke in Betriebsprozesse ermöglichen und fortgeschrittene Analysemethoden wie prädiktive Wartung, Qualitätskontrolle und Prozessoptimierung ermöglichen. Die Herausforderung besteht nicht nur in der effektiven Erfassung von Zeitreihen- und historische Daten, sondern auch in deren konsistenter Bereitstellung. Hierfür bietet das „Time Series Data“ Submodel Template (Teilmodellvorlage) einen Standard für die Integration und semantische Beschreibung von Zeitreihendaten in der Verwaltungsschale [Id23]. Diese interoperable Beschreibung ermöglicht es, die generierten Datenpunkte im Lebenszyklus des Assets zu interpretieren und effektiv zu nutzen.

2 Stand der Forschung & Anwendungsfall

2.1 Time Series Data Submodel Template

Die Industrial Digital Twin Association (IDTA) definiert eine Reihe standardisierter Teilmodellvorlagen. Dabei handelt es sich um Datenmodelle, die eine semantisch definierte Datenstruktur zur Beschreibung eines spezifischen Zusammenhangs darstellen [Be20]. Zu diesen Teilmodellen zählen beispielsweise das „Digital Nameplate“ (Digitales Typenschild), „Handover Documentation“ (Übergabe Dokumentation) und auch „Time Series Data“ (Zeitreihendaten) [Pl21]. Das Teilmodell für Zeitreihendaten stellt einen Ansatz für die semantische Beschreibung dieser Daten über den Lebenszyklus eines Assets dar. Die Spezifikation definiert dabei die Integration externer Datenquellen (z.B. Datenbanksysteme) sowie die Speicherung von Zeitreihendaten in der Verwaltungsschale selbst [Id23]. Die Nutzung des Teilmodells beginnt bereits in der Engineering Phase eines Assets (Asset Typ).

In dieser Phase wird das Teilmodell basierend auf Metadaten des zugehörigen Produkttyps erstellt. Der Fokus des Time Series Data Teilmodells liegt jedoch auf der Phase nach der Instanziierung eines Produktes mit der Speicherung von erzeugten Datensätzen [Id23].

Das Teilmodell teilt Zeitreihendaten in sogenannte Zeitreihensegmente ein. Dabei ist ein Zeitreihensegment als eine Teilsequenz der Variablen innerhalb einer Zeitreihe definiert. Ein Segment besitzt generell einen Namen, eine Beschreibung, eine Eigenschaft zur Angabe der Anzahl der Datensätze, eine Startzeit und eine Endzeit [Id23]. Es können ebenfalls weitere benutzerdefinierte Eigenschaften eingebettet werden (z.B. Abstrakte oder Status, ob ein Segment bereits abgeschlossen ist).

Der Standard unterscheidet zwischen drei Typen von Zeitreihensegmenten (vgl. [Id23]).

1. **InternalSegment:** Dieser Typ erlaubt die Verwaltung der Struktur und des Inhalts von Zeitreihendaten in der Verwaltungsschale selbst. Er ist geeignet für die Integration weniger Datenpunkte. Dies ist z.B. bei der Übergabe von Zeitreihendaten unter Verwendung von Verwaltungsschalen der Fall.
2. **LinkedSegment:** Das LinkedSegment ermöglicht es einer Industrie 4.0-Anwendung, den Endpunkt eines externen Systems unter Angabe einer Abfrageklausel auszulesen. Hierbei wird die Zeitreihe nicht von der Verwaltungsschale verwaltet. Anwendungsfälle umfassen die Brownfield Integration, bei welcher Massendaten anfallen, die wiederum dynamischer Natur sein können (kontinuierliche Updates des externen Systems).
3. **ExternalSegment:** Der Typ ExternalSegment ermöglicht es eine Daten- oder BLOB-Datei zu finden, in welcher die Zeitreihendaten abgespeichert sind. Das Segment findet bei statischen Zeitreihendaten Anwendung (insbesondere in der Brownfield Integration). Anwendungsfall ist auch hier die Übergabe von Zeitreihendaten bei einer generell geringen Anzahl an Zugriffen auf den Datensatz.

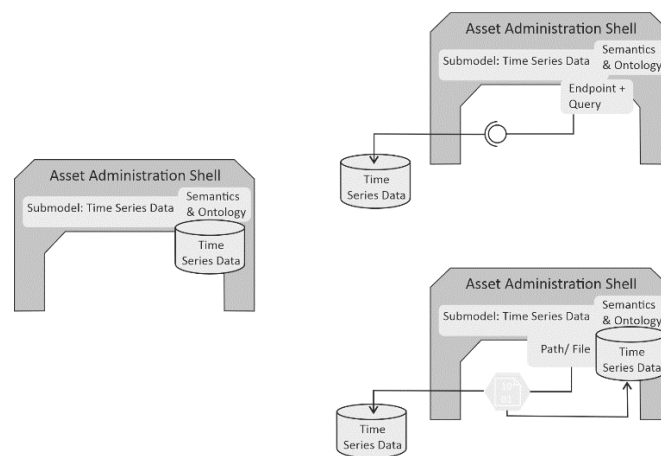


Abb. 1: Typen von Zeitreihensegmenten (Quelle: [ID02])

2.2 Industrieller Anwendungsfall

Im Rahmen eines Retrofit-Projektes wurde beim Projektpartner der EAW – Relaisstechnik GmbH ein Niet- und Schweißautomat (siehe Abbildung 2) vollständig modernisiert. Der Automat beinhaltet eine umlaufende Werkzeugkette, welche im Drei-Sekundentakt Federbleche entlang mehrerer Bearbeitungsstationen führt. An diesen Stationen werden

den Federblechen Nietschaltstücke bereitgestellt und vernietet, sowie ein Kontaktmesser angeschweißt. Abschließend werden die fertigen Kontaktfedern gebogen und sortiert.

Im Unternehmen erfolgt aktuell die schrittweise Aufbereitung des Digitalisierungsprozesses unter Nutzung von Verwaltungsschalen. Im Rahmen des Prozesses werden auch relevante Prozessinformationen zur Bewertung der Produktionsabläufe erfasst und visualisiert. Im Beispielautomat betrifft dies den Ressourcenverbrauch, die Überwachung der Versorgung, die Erfassung des Kühlkreislaufes, sowie die Gut- und Schlechteildetektion.

Die genannten Informationen sollen ohne Eingriff in die Prozesse bereitgestellt werden. Ziel ist die Implementierung eines digitalen Zwillings in Form einer Verwaltungsschale zur Prozessüberwachung wie auch zur Informationsbereitstellung für die Optimierung der Anlageneffizienz.



Abb. 2: Niet- und Schweißautomat der EAW Relaistechnik GmbH

Unter Nutzung der Verwaltungsschale werden hier diverse Teilmodelle angelegt. Im Kontext des Beitrages soll insbesondere das Time Series Data Teilmodell mittels der aufgenommen Zeitreihendaten Mehrwerte generieren. Folgenden Ziele werden angestrebt:

- Steigerung der Gesamtanlageneffektivität (Overall Equipment Effectiveness, OEE) durch Betriebszeiten- und Leistungsdatenüberwachung
- Prozessoptimierung über die Analyse von Zeitreihendaten bezogen auf die Produktqualität und den Materialverbrauch
- Zustandsüberwachung und Anlagenmonitoring von Schlüsselindikatoren (Key Performance Indicators, KPIs)
- Analyse der Energieverbrauchsdaten zur Steigerung der Energieeffizienz durch Verbrauchsoptimierungen
- Fehleridentifikation und Diagnose zur Minimierung von Ausfallzeiten und Steigerung der Zuverlässigkeit
- Optimierung der Wartungsintervalle basierend auf tatsächlichen Abnutzungswerten und Nutzmustern
- Überwachung der Produktqualität während der Produktion zur Sicherstellung der geforderten Spezifikationen
- Vorhersage des Werkzeugverschleißes zur prädiktiven Wartung und Verhinderung von Ausfallzeiten

Der Niet- und Schweißautomat ist zur Generierung der notwendigen Zeitreihendaten mit diversen Sensoren ausgestattet. Aufgenommen werden unter anderem die elektrischen Parameter, die Erfassung der Betriebsstunden, der Durchfluss des Kühlsystems, sowie die Qualität der Fertigung über die Gut-/Schlechteildetektion. Da als Steuerungskonzept eine Zustandsmaschine auf Basis der OMAC PackML implementiert wurden, können ebenfalls weitere Informationen wie die Anlagenzustände bereitgestellt werden.

Im konkreten Anwendungsfall handelt es sich um Geräte (u.a. Industriesteuerungen), die das MQTT- Protokoll unterstützen. Als Clients besitzen sie die Möglichkeit, ihre Messdaten an einen MQTT-Broker zu übermitteln (publish), so dass diese von einem anderen MQTT-Client entgegengenommen werden können (subscribe). Vorteile in der Nutzung des Protokolls ergeben sich aus der eventbasierten Kommunikationsinfrastruktur, da nur bei Zustandsänderungen Informationen über das Netzwerk transferiert werden und eine höhere Datenübertragungseffizienz zu erreichen [MQ00].

3 Architektur & Implementierung

Um einen Mehrwert aus den genannten Betriebsdaten eines Assets zu generieren, müssen diese zunächst persistiert werden, so dass auf aktuelle und historische Daten zugegriffen werden kann. Im beschriebenen Anwendungsfall werden kontinuierlich neue Daten generiert. Da die Zeitreihendaten somit nicht statischer Natur sind, empfiehlt es sich für die Integration in die Verwaltungsschale das „LinkedSegment“ aus der IDTA-Spezifikation anzuwenden.

3.1 Architektur für die dynamische Integration von Zeitreihendaten

Das Time Series Data Teilmodell benötigt in der Implementierung zwei wesentliche Architekturkomponenten, die für die Integration von Zeitreihendaten in der Verwaltungsschale von Bedeutung sind. Auf der einen Seite bedarf es der Implementierung der Verwaltungsschale inklusive der Teilmodelle und Concept Descriptions. Konkret handelt es sich dabei um das sog. AAS Environment, bestehend aus dem AAS Repository, dem Submodel Repository und dem Concept Description Repository [BA00]. Auf der anderen Seite wird ein Datenbanksystem zur Ablage der Zeitreihendaten benötigt.

BaSyx, mitsamt seiner Off-the-Shelf-Komponenten, stellt eine mögliche und frei verfügbare Implementierung der Verwaltungsschale dar. BaSyx setzt als Middleware die Spezifikationen der IDTA bezüglich des Verwaltungsschalenmetamodells, der Application Programming Interfaces (API) und der Embedded Data Specifications um. Es stellt die Möglichkeit bereit, Laufzeitinstanzen von Verwaltungsschalen zu starten, welche über ihre Schnittstellen (API) von Industrie 4.0 Applikationen ausgelesen werden können [BA01].

Zur Speicherung von Zeitreihendaten existieren mehrere dedizierte Datenbanklösungen. Diese zeichnen sich durch ihre Leistung bei der Handhabung großer Datenmengen und einer hochfrequenten Generierung von Datenpunkten aus. Eine beispielhafte Datenbank für Zeitreihendaten ist „InfluxDB“. Vorteile von InfluxDB sind die niedrige Latenzzeit für Abfragen, die optimierte Datenkompression und die Möglichkeit der Datenabfrage über http/REST [IN00].

Damit im Anwendungsfall aufgenommene Messwerte von den Sensoren in der Datenbank abgelegt werden können, müssen diese über Konnektoren angebunden werden. BaSyx stellt eine mögliche Komponente für die Anbindung an diverse industrielle Protokolle bereit. Die sog. „DataBridge“ unterstützt beispielsweise OPC UA, MQTT, Kafka und einige weitere Protokolle [DA00]. Jedoch beschränkt sich der Anwendungsfall der DataBridge bisher ausschließlich auf den Transfer von Livedaten zwischen dem Asset und der Verwaltungsschale selbst.

3.2 Telegraf

Im Bereich der Datenbankanbindung von Assets existieren ebenfalls Komponenten, die dieses Problem adressieren. So gibt es beispielsweise die Komponente „Telegraf“, die als Daten Konnektor für InfluxDB Datenbanken eingesetzt werden kann. Ähnlich wie auch bei der DataBridge werden diverse Kommunikationsprotokolle unterstützt. Diese finden unter anderem beim „Internet of Things“ (IoT) Verwendung. Ein Beispiel hierfür ist das MQTT-Protokoll, welches im beschriebenen Anwendungsfall eingesetzt wird. Dies kann ebenso von Telegraf eingesetzt werden.

Abb. 3 zeigt eine Komponentenübersicht für die Assetintegration beim Einsatz von Verwaltungsschalen während der operativen Phase eines Assets. Auf der rechten Seite ist die Übertragung von Livedaten über die DataBridge aufgeführt. Links zu erkennen in der Abbildung ist das Datenbanksystem (InfluxDB) und Telegraf für das Abfragen und Speichern von generierten Zeitreihendaten aus dem Asset bzw. den installierten Sensoren. Über das Time Series Data Teilmodell kann eine Industrie 4.0 Anwendung (hier das BaSyx AAS Web UI) auf den Endpunkt und den Abfrageterm für die Zeitreihendatenbank zugreifen und den Datensatz aus der Datenbank abfragen. Das bedeutet, dass auch bei einem eventuellem Verbindungsverlust zu dem Asset kein Error entstehen wird, sondern ein Null-Wert gespeichert werden kann und die Visualisierung dies auch zeigt. Außerdem ermöglicht es die längere Beobachtung von Prozessen, da die Daten bis zum Aufzeichnungsbeginn zurückverfolgt werden können.

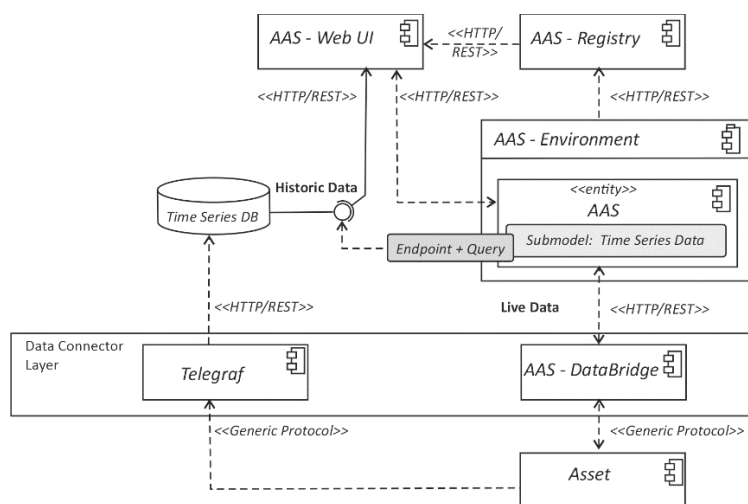


Abb. 3: Architektur für die Asset-Integration und Integration von Zeitreihendaten über das Time Series Data Teilmodell

3.3 PlugIn-Mechanismus und Visualisierung von Zeitreihen

Mit der Digitalisierung folgt zwangsläufig die Forderung, Informationen rollenspezifisch darzustellen. Die betrifft sowohl die Inhalte als auch die Form der Visualisierung. Mit einem PlugIn-Mechanismus ist es möglich, zugeschnittene Nutzerschnittstellen einzubetten.

PlugIns können Teilmodell-spezifisch erstellt und aufgrund der einfachen Einbindung firmenübergreifend genutzt werden. Die dynamische Einbindung erfolgt, indem das neu erstellte PlugIn in einem dafür vorgesehenen Ordner abgelegt wird. Um festzustellen, ob ein PlugIn für ein Teilmodell verfügbar ist, wird dessen SemanticId überprüft. Existieren PlugIns, werden sie beispielweise im AAS-WebUI dem Nutzer zur Verfügung gestellt.

Für das Teilmodell TimeSeriesData wurde ein entsprechendes Plugin umgesetzt, welches die Darstellung der Zeitreihendaten über Diagramme ermöglicht. Die Diagramme können vom Nutzer, mit Hilfe von Diagramm-spezifischen Optionen, angepasst werden.

4 Zusammenfassung und Ausblick

Die Integration von Zeitreihendaten in die Verwaltungsschale über das „Time Series Data“ Teilmodell bietet signifikante Vorteile gegenüber vergleichbaren proprietären Lösungen. Ein entscheidender Mehrwert entsteht durch die Skalierungseffekte, die sich aus der Anwendung dieses Standards ergeben. Der Einsatz einheitlicher und interoperabler Schnittstellen vereinfacht die Integration von Informationen aus heterogenen Datenquellen erheblich, was in komplexen Fertigungsumgebungen zur Reduktion von Integrationsaufwänden und -kosten führt.

Zudem fördert das Teilmodell die Sicherstellung der Datenqualität und -konsistenz durch eine harmonisierte und semantisch eindeutige Beschreibung von Zeitreihendaten. Im Gegensatz zu proprietären Lösungen, die oft isolierte Dateninseln schaffen, erleichtert der Standard die system- und unternehmensübergreifende Datenintegration. Die Verwaltungsschale fungiert dabei als zentrale Schnittstelle für sämtliche Asset-bezogenen Informationen, wodurch Anwendungen für fortschrittliche Analysemethoden, wie prädiktive Wartung und Qualitätskontrolle, optimierten Zugriff auf die benötigten Daten erhalten.

Darüber hinaus ermöglicht die standardisierte Einbindung von Zeitreihendaten in die AAS eine effizientere Nutzung von Machine Learning und KI-Technologien, was zu weiteren Optimierungen in der Prozessüberwachung und Anlagenwartung führt. Dadurch können Unternehmen ihre digitalen Transformationen effektiver vorantreiben und einen entscheidenden Wettbewerbsvorteil in einer zunehmend datengetriebenen Industrielandschaft erzielen.

Diese Forschungsarbeit wurde im Forschungsvorhaben „Basys4Transfer“ durch das Bundesministerium für Bildung und Forschung BMBF unter der Nummer (01|S22089G) gefördert. Die Autoren bedanken sich für die Unterstützung des BaSys4Transfer-Teams.

5 Literaturverzeichnis

- [Dr23] Drath, R. et.al.: Diskussionspapier – Interoperabilität mit der Verwaltungsschale, OPC UA und AutomationML, Zielbild und Handlungsempfehlungen für industrielle Interoperabilität, Techn. Ber., 2023.
- [SHS18] Schütze, A.; Helwig, N.; Schneider, T.: Sensors 4.0 – smart sensors and measurement technology enable Industry 4.0. *Journal of Sensors and Sensor Systems* 7/1, S. 359–371, 2018
- [Id23] Industrial Digital Twin Association e.V.: IDTA 02008-1-1 Time Series Data, Frankfurt am Main, 2023.
- [Be20] Bedenbender, H. et.al.: Verwaltungsschale in der Praxis, Techn. Ber., Plattform Industrie 4.0, 2020.
- [Pl21] Plattform Industrie 4.0: VWS-Referenzmodellierung, Exemplarische Modellierung einer fertigungstechnischen Anlage mit AASX Package Explorer auf Basis des VWS-Metamodells, Techn. Ber., 2021.
- [BP21] Bansal, M.; Priya: Performance Comparison of MQTT and CoAP Protocols in Different Simulation Environments. In (Ranganathan, G.; Chen, J.; Rocha, Á., Hrsg.): *Inventive Communication and Computational Technologies*. Springer Singapore, Singapore, S. 549–560, 2021.
- [Ec24] Eclipse BaSyx: AAS Environment, 2024, url: <https://github.com/eclipse-basyx/basyx-java-server-sdk/tree/main/basyx.aasenvironment>, Stand: 12.01.2024.
- [In24a] Influx Data: InfluxDB Time Series Data Platform, 2024, url: <https://www.influxdata.com/>, Stand: 12.01.2024.
- [SKA23] Schnicke, F.; Kuhn, T.; Antonino, P. O.: Eclipse BaSyx DataBridge – Datenintegration einfach machen! *atp magazin* 65/5, S. 50–58, 2023.
- [In24b] Influx Data: Telegraf - Time Series Data Collection, 2024, url: <https://www.influxdata.com/time-series-platform/telegraf/>, Stand: 12.01.2024.

Webbasiertes Werkzeug für das modellbasierte Engineering von KI-Anwendungen für Automatisierungssysteme

Marvin Schieseck¹, Philip Topalis¹ und Alexander Fay²

Abstract:

In diesem Beitrag wird ein Konzept vorgestellt, welches es der Industrie erlaubt, eigenständig KI-Anwendungen für Automatisierungssysteme samt Fertigungs- oder Verfahrensprozess zu modellieren und deren Entwicklung, Integration, Dokumentation und Betrieb auf Basis dieses Modells zu koordinieren. Das Konzept basiert auf der Kombination eines grafischen Beschreibungsmittels und einer Entwicklungsmethode, welche in ein webbasiertes Werkzeug zur grafischen Modellierung überführt werden. Anschließend wird zur Evaluation des Werkzeugs ein exemplarischer Anwendungsfall aus der Industrie modelliert. Die Evaluation zeigt den Nutzen des Werkzeugs bei der Koordination von Entwicklung, Integration und Betrieb industrieller KI-Anwendungen auf.

Keywords: KI-Engineering, Web-basierte Werkzeuge, grafische Modellierung

1 Einleitung

In den letzten Jahren wurde die Integration von Künstlicher Intelligenz (KI) in Automatisierungssysteme wie Produktions- und Prozessanlagen seitens der Unternehmen, der Forschung und der Politik vorangetrieben [Bu20; Bu23]. Dabei wurden eine Vielzahl von Anwendungsfällen identifiziert, bei denen der Einsatz von KI entweder die Effizienz bestehender Ansätze übertreffen konnte oder eine Lösung überhaupt erst ermöglichte. Beispiele dafür finden sich im Bereich der Bild- und Signalklassifikation sowie der Textverarbeitung [ISO/IEC TR 24030]. In der Folge konnten durch die Nutzung von KI industrielle Anwendungsfälle der prädiktiven Instandhaltung oder der optischen Qualitätsüberwachungen umgesetzt werden [Di20].

Trotz dieser Potenziale und den vielversprechenden Anwendungsfällen wird KI in der Industrie dennoch vergleichsweise wenig eingesetzt [MG23]. Häufig werden KI-Anwendungen im Rahmen von Forschungs- und Pilotprojekten entwickelt, diese können aber im Anschluss nur bedingt wirtschaftlich in die Praxis überführt werden. Dies kann insbesondere darauf zurückgeführt werden, dass die Integration von KI-Systemen in Automatisierungssysteme im Vergleich zur Integration von KI-Systemen in konventionelle Softwaresysteme mit zusätzlichen Herausforderungen konfrontiert ist. Selbiges gilt neben der Integration auch

¹ Helmut-Schmidt-Universität, Institut für Automatisierungstechnik, Holstenhofweg 85, 22043 Hamburg, Deutschland, {marvin.schieseck, philip.topalis}@hsu-hh.de,

² Ruhr-Universität Bochum, Lehrstuhl für Automatisierungstechnik, Universitätsstraße 150, 44801 Bochum, Deutschland, alexander.fay@rub.de,

für die Entwicklung von KI-Systemen für Automatisierungssysteme. Von Merkel-Kiss und Garrel [MG23] wurden diese Herausforderungen im Rahmen einer quantitativen Metastudie durch Auswertung von 24 nationalen Studien klar herausgestellt. Nach den Autoren sind die größten Herausforderungen, insbesondere für kleine und mittelständische Unternehmen: 1.) *Mangel an Kompetenzen in den Unternehmen und Mangel an Fachkräften*, 2.) *hohe Aufwände für die Konzeptionierung, Entwicklung, Implementierung und den Betrieb sowie damit einhergehend hohe Kosten und lange Amortisationszeiten*, 3.) *komplexe, technischen Infrastrukturen*, 4.) *komplexe Entscheidungsprozesse für die Unternehmen*, 5.) *niedrige Technologiereife und fehlende Standards*.

Um die Entwicklung und Integration von KI-Systemen zu unterstützen, existieren bereits verschiedene Ansätze. Beispielsweise existieren diverse Methoden, die eine zeitliche und logische Abfolge von Handlungen vorschlagen und so den Entwicklungsprozess durch gezielte Festlegung der Entwicklungsschritte unterstützen, wie beispielsweise der *Cross-Industry Standard Process for Data Mining (CRISP-DM)* [WH00], *Data mining methodology for engineering applications (DMME)* [Hu19] oder das *Process Model for AI Systems Engineering (PAISE)* [HS22]. Ein weiterer Ansatz sind grafische Beschreibungen des KI-Systems sowie des technischen Systems [KWS21] [STF23]. Neben Methoden und Beschreibungsmitteln existieren auch Werkzeuge, welche bei der Entwicklung von KI-Systemen unterstützen. Der Großteil dieser KI-Werkzeuge unterstützt den Anwender, indem sie es erleichtern, ein KI-Modell zu erstellen, den Code zu kapseln und zu versionieren und die Ressourcen für das Training bereitzustellen und zu verwalten [Ru21]. Beispiele dafür sind KI-Werkzeuge wie *MLFlow*³, *ClearML*⁴, *Knime*⁵ oder *RapidMiner*⁶, aber auch KI-Plattformen wie *Vertex AI*⁷ von Google, *Azure AI*⁸ von Microsoft oder *SageMaker*⁹ von Amazon. Diese Werkzeuge sind darauf spezialisiert, den Aufbau und das Training von KI-Modellen sowie die damit verbundene Datenvor- und -nachverarbeitung zu erleichtern und zu verwalten [KKH22]. Sie ermöglichen es jedoch nicht, die technische Anlage oder die technischen Prozesse eines Unternehmens konkret zu berücksichtigen und in den Entscheidungs- oder Modellierungsprozess mit einfließen zu lassen. Die bestehenden Werkzeuge nutzen keine für industrielle Anwendungen geeigneten Beschreibungsmittel und unterstützen nicht alle Schritte, die für die industrielle KI-Entwicklung erforderlich sind. Zusammenfassend folgt, dass weder die bestehenden KI-Werkzeuge noch die vorgestellten Methoden noch die Beschreibungsmittel den vorgestellten Herausforderungen in ausreichendem Maße begegnen können, insbesondere den hohen Aufwänden, den langen Amortisationszeiten sowie dem Umgang mit der komplexen technischen Infrastruktur. Keiner der bestehenden Ansätze ist hinreichend geeignet, ein verständliches Konzept einer KI-Anwendung für ein

3 <https://mlflow.org/>

4 <https://clear.ml/>

5 <https://www.knime.com/>

6 <https://docs.rapidminer.com>

7 <https://cloud.google.com/vertex-ai?hl=de>

8 <https://azure.microsoft.com/de-de/solutions/ai>

9 <https://aws.amazon.com/de/sagemaker/>

Automatisierungssystem zu entwickeln, die Entwicklung zu koordinieren, die Integration der KI in das Automatisierungssystem zu ermöglichen und abschließend den Betrieb zu gewährleisten.

Es fehlt ein Lösungsansatz, der es der Industrie erlaubt, eigenständig KI-Anwendungen für Automatisierungssysteme samt Fertigungs- oder Verfahrensprozesses zu modellieren und deren Entwicklung auf Basis des Modells zu koordinieren. Dieser Lösungsansatz muss als industrielle Schnittstelle zwischen den verschiedenen an der Entwicklung beteiligten Domänen geeignet sein und die Integration der bestehenden KI-Werkzeuge in die industrielle Entwicklung von KI ermöglichen. Dieser Handlungsbedarf wird auch ausführlich vom Forschungsbeirat der Plattform Industrie 4.0 [Ma21a] dargelegt.

In diesem Beitrag wird der Handlungsbedarf adressiert und ein Konzept zur Schließung dieser Lücke vorgestellt. Das Konzept basiert auf der Anwendung des Beschreibungsmittels von Schieseck et al. [STF23], welches in ein web-basiertes Werkzeug zur grafischen Modellierung überführt wird, um speziell die industrielle KI-Entwicklung nach DMME zu unterstützen. Es wird dabei nach dem von Schnieder [Sc99] vorgestellten Prinzip der Kombination von Beschreibungsmittel, Methode und Werkzeug vorgegangen. Das Werkzeug wurde mit React Flow¹⁰, einer Bibliothek für interaktive und browserbasierte grafische Modellierung, realisiert¹¹.

2 Stand der Technik und Forschung

In diesem Abschnitt des Beitrags werden verschiedene Ansätze, wie Entwicklungsmethoden, grafische Modellierungssprachen, KI-Werkzeuge, Community-Plattformen, Automated Machine Learning oder Design Patterns vorgestellt. Nachfolgend werden diese Ansätze erläutert und daraus ein gezielter Handlungsbedarf abgeleitet.

2.1 Entwicklungsmethoden

Unter einer Methode wird ein auf einem Regelsystem aufbauendes Verfahren zur Erlangung von theoretischen oder praktischen Ergebnissen verstanden. Im Kontext der Systementwicklung zeigen Methoden auf, wie eine Lösung planmäßig erarbeitet werden kann, welche Entwicklungsschritte in welcher Reihenfolge vollzogen werden sollten und welche Aktivitäten in den jeweiligen Schritten auszuführen sind. [Sc99]

Es existieren Methoden, welche darauf ausgelegt sind, die allgemeine KI-Entwicklung zu unterstützen, wie beispielsweise CRISP-DM [WH00], *Machine Learning Operations* [KKH23] (MLOps) oder *Engineering data driven applications* [He19] (EDDA). Darüber hinaus existieren Methoden, welche speziell für die industrielle KI-Entwicklung entwickelt

¹⁰ <https://reactflow.dev/>

¹¹ <https://github.com/schiesem/ai-modeling-tool>

wurden, wie beispielsweise *Process Model for AI Systems Engineering* PAISE [HS22], *Machine Learning for Production* [Ei22] (ML4P) oder DMME [Hu19].

Als de-facto Standard für allgemeine Data Mining- aber auch KI-Projekte im Allgemeinen hat sich in den letzten zwei Jahrzehnten der *Cross-Industry Standard Process for Data Mining* (CRISP-DM) [WH00] durchgesetzt [Ma21b; VDI/VDE 3714-1]. Das Vorgehen nach CRISP-DM ist in sechs Schritte unterteilt: 1. *Business Understanding*, indem die Geschäfts- und Businessziele definiert werden, 2. *Data Understanding*, indem Hypothesen auf der Grundlage von Erfahrungen und qualifizierten Annahmen mithilfe der Daten definiert werden, 3. *Data Preparation*, indem relevante Daten aufbereitet werden, 4. *Modeling*, indem eine geeignete KI-Methode ausgewählt und ein Modell trainiert wird, 5. *Evaluation*, indem das trainierte Modell mithilfe von Testdaten evaluiert wird und 6. *Deployment*, indem das Modell sowie die Datenvor- und Nachverarbeitung in Betrieb genommen wird.

CRISP-DM bezieht sich auf die Entwicklung von KI-Anwendungen für konventionelle Softwarebereiche und ist nur eingeschränkt für technische Systeme und damit die industrielle KI-Entwicklung geeignet, weil es die Gegebenheiten des technischen Systems und des zugrundeliegenden technischen Prozesses nicht ausreichend berücksichtigt [Hu19; Ma21b]. Deshalb wurde CRISP-DM von Huber et al. [Hu19] für technische Systeme angepasst, durch die Ergänzung von drei Schritten. Diese adaptierte Methode bezeichnen die Autoren als *Data mining methodology for engineering applications* (DMME) (siehe Abb. 1).

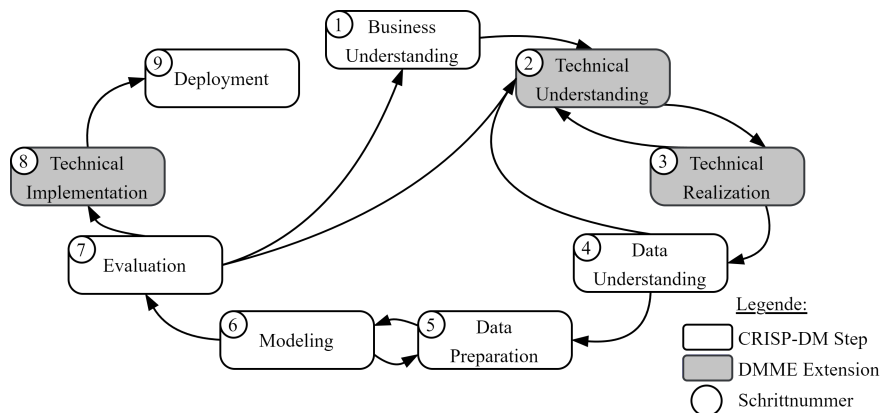


Abb. 1: DMME Methode nach Huber et al. [Hu19]

Zunächst werden nach dem Business Understanding zwei aufeinanderfolgende Schritte, das *Technical Understanding* und die *Technical Realization* ergänzt. Im *Technical Understanding* werden Informationen über die technische Anlage sowie die prozess- und verfahrenstechnischen Zusammenhänge zusammengetragen. Mithilfe dieser Informationen werden aus den Geschäftszielen dann technische Ziele formuliert. Danach werden im Schritt *Technical Realization* verschiedene Konzepte für mögliche KI-Anwendungen entworfen und bewertet. Parallel dazu werden die erforderlichen Daten innerhalb des technischen Systems

aufgenommen und gegebenenfalls Veränderungen der Anlage durchgeführt, beispielsweise das Installieren oder Kalibrieren von Sensorik. Der dritte Schritt, die *Technical Implementation*, wird nach der Evaluation ergänzt. Er dient dazu, das trainierte Modell sowie die dazugehörige Datenversorgung in das technische System zu integrieren. DMME kann als angepasster de-facto Standard für die industrielle KI-Entwicklung angesehen werden und stellt ein Mindestmaß an absolut notwendigen Schritten für die industrielle KI-Entwicklung bereit.

2.2 Grafische Modellierungssprachen und Beschreibungsmittel

Durch grafische Darstellungen kann der Kern eines Problems formal und verständlich für unterschiedliche Beteiligte dargestellt werden [Ha19; VDI/VDE 3681]. Grafische Beschreibungsmittel und Modellierungssprachen ermöglichen die Definition von Symbolen, Regeln und Semantik einer solchen grafischen Darstellung und damit eine verständliche Modellierung für alle beteiligten Experten. Das übergeordnete Ziel einer grafischen Modellierung ist, die Aufgabe und Lösung des Problems möglichst formal abzubilden und dadurch die Lösungsfindung sowie die Kommunikation aller an der Entwicklung beteiligten Akteure zu erleichtern [Sc99].

Grafische Modellierungssprachen werden in allgemeine Modellierungssprachen (General Purpose Modeling Language, GPMLs) und in domänenspezifische Modellierungssprachen (Domain Specific Modelling Languages, DSMLs) unterteilt. GPMLs wie die Unified Modeling Language (UML) oder die Systems Modeling Language (SysML) erlauben es Experten, verschiedene Probleme aus unterschiedlichen technischen Bereichen auf der Grundlage allgemeiner Konzepte wie Klassen, Attribute oder Vererbung zu modellieren. Diese Flexibilität verringert jedoch gleichzeitig auch die Produktivität des Modellierungsprozesses und die Verständlichkeit der grafischen Darstellung, weil die technischen Konzepte auf Domänenebene auf der Grundlage der allgemeinen Konzepte von Grund auf rekonstruiert werden müssen. DSMLs sind weniger flexibel, bieten dafür aber bereits die technischen Konzepte auf Domänenebene, die der Fachsprache des betreffenden Fachbereichs entsprechen. Sie steigern daher die Produktivität und Verständlichkeit der Modellierung. [Fr10; Ro19]

Ein Ansatz speziell für die grafische Beschreibung von KI-Systemen wurde von Kaymakci et al. [KWS21] vorgestellt, welcher auf der Darstellung der Datenquellen, der Datensinken sowie der datenverarbeitenden Elemente beruht. Dieser Ansatz wurde von Schieseck et al. [STF23] aufgegriffen und mit Elementen anderer Beschreibungsmittel wie der formalisierten Prozessbeschreibung [VDI/VDE 3682] kombiniert. Auf dieser Basis wurde ein grafisches Beschreibungsmittel für KI-Systeme innerhalb von Automatisierungssystemen vorgestellt, welches auf drei grundlegenden Elementgruppen basiert: Systemkomponenten, Systemfunktionen und Systemrelationen.

Die Systemkomponenten stellen die Bausteine dar, aus denen sich das technische System zusammensetzt. Sie werden weiter in technische Ressourcen und Produkte unterteilt.

Die technischen Ressourcen, wie beispielsweise Steuerungen, Edge-Geräte, Computer, Cloud, Aktoren und Sensoren führen die Systemfunktionen aus. Die Produkte hingegen sind die Gegenstände der realen Welt, die während des technischen Prozesses umgewandelt werden, beispielsweise Halbzeuge zu Fertigteilen. Systemfunktionen wandeln eine Eingabe in eine entsprechende Ausgabe um und ermöglichen damit die Zielerreichung des technischen Systems. Es werden sieben verschiedene Systemfunktionen definiert: Automatisieren, Transformieren von Produkten, Aufnehmen von Daten, Speichern von Daten, Verarbeiten von Daten, Trainieren und Inferieren. Systemrelationen beschreiben die Verbindungen und Beziehungen zwischen den Systemelementen und erlauben es, die Abhängigkeiten zu modellieren. Es werden drei Kategorien von Systemrelationen definiert: Kommunikation, Zuordnung und Material-/Energiefluss. Eine Kommunikation ist eine informationstechnische Verbindung zwischen zwei Systemkomponenten. Eine Zuordnung ist eine strukturelle Zuweisung von Systemfunktionen zu Systemkomponenten. Ein Material- oder Energiefluss stellt eine Verbindung zwischen Produkten und Transformationsprozessen dar. In Abb. 2 ist die definierte Symbolik der Elemente dargestellt. Die Systemfunktionen werden durch Rechtecke dargestellt. Die technischen Ressourcen werden durch Rechtecke mit abgerundeten Ecken repräsentiert und für die Produkte werden Kreise verwendet. Die Systemrelationen werden durch Linien repräsentiert. Die Zuweisung von Systemkomponenten und Funktionen erfolgt mithilfe einer Strichlinie. Der Materialfluss wird durch eine Punktlinie dargestellt. Die Symbolik und die Semantik für diese Elemente wird aus der VDI 3682 [VDI/VDE 3682] übernommen. Eine Kommunikation wird mithilfe einer Volllinie dargestellt. Zusätzlich haben Schieseck et al. [STF23] ein Meta-Modell vorgeschlagen, um das grafische Beschreibungsmittel in eine vollwertige DSML zu überführen. Durch dieses Meta-Modell werden die Klassenrelationen formal definiert.

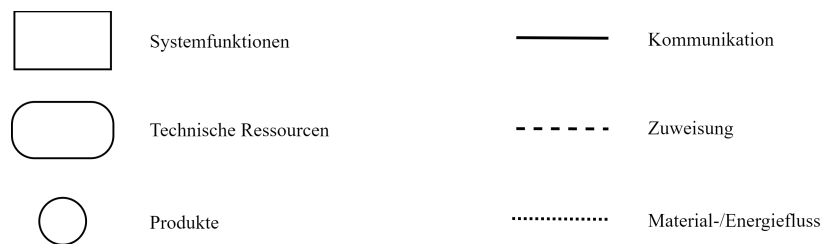


Abb. 2: Symbole des Beschreibungsmittels nach Schieseck et al. [STF23]

In Abb. 3 ist die Modellierung eines exemplarischen KI-Anwendungsfalls zur Überwachung eines Stanzprozesses abgebildet. Bei diesem Stanzprozess wird das Eingangsprodukt Blechplatte durch den Transformationsprozess Stanzen in ein fertiges Fertigteil und Restmaterial umgewandelt. Der Stanzprozess wird durch einen Antriebsmotor ausgeführt. Durch einen rechten und einen linken Lagesensor werden die Positionsdaten der jeweiligen Seite der Hubmechanik erfasst. Diese Daten werden von den Sensoren an die Steuerung kommuniziert. Dort erfolgt auf Basis dieser Werte die Lagesteuerung der Stanze, indem der Antriebsmotor

entsprechende Stellwerte erhält. Gleichzeitig kommuniziert die Steuerung die Positionsdaten in eine Cloud. Dort werden zunächst Merkmale aus den zeitabhängigen Positionsdaten erzeugt und als Merkmalreihe gespeichert. Diese Daten können für das Training des Modells verwendet werden oder für die Ausführung eines bereits trainierten Modells. Das Ausführen und Anwenden eines bereits trainierten Modells unter Verwendung von Produktionsdaten wird in diesem Kontext auch als Inferenz bezeichnet. Weil die Cloud von der Steuerung stetig neue Batches von Daten erhält, können diese kontinuierlich und parallel zum Stanzprozesses in der Cloud ausgewertet werden. Bei entsprechender Eignung der Daten und der Auswahl eines geeigneten Modells könnte über diesen Aufbau zum Beispiel eine KI-Anwendung zur prädiktiven Instandhaltung der Stanze realisiert werden.

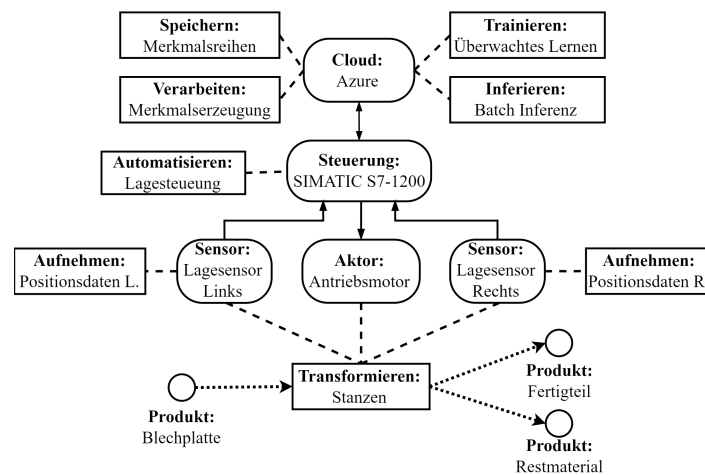


Abb. 3: Modellierung eines exemplarischen Anwendungsfalls mithilfe des Beschreibungsmittels nach Schieseck et al. [STF23]

2.3 KI-Werkzeuge

Ein Werkzeug ist nach Schnieder [Sc99] ein Hilfsmittel, das die Anwendung einer Methode sowie die Verwendung eines Beschreibungsmittels unterstützt, um die Entwicklungsaufwände und damit verbundenen Kosten zu reduzieren. Für die Entwicklung von KI-Anwendungen gibt es bereits eine Vielzahl derartiger Werkzeuge, sowohl kommerziell genutzte als auch open source Ansätze.

Ein bekanntes open source Werkzeug ist *KubeFlow*, welches auf die Orchestrierung von KI-Pipelines unter Verwendung von Kubernetes spezialisiert ist [Ru21]. *KubeFlow* ermöglicht, die Datenvor- und Nachverarbeitung, das Training und das Deployment des ausführbaren Modells zu kapseln und mithilfe einer webbasierten Plattform zu visualisieren. Ein weiteres open source Werkzeug, welches ebenfalls die Orchestrierung von KI-Pipelines

ermöglicht, ist *ClearML*. *ClearML* erlaubt wie *KubeFlow* die Kapselung der Datenvor- und Nachverarbeitung, des Trainings und des Deployments sowie eine webbasierte Visualisierung [CI23]. Neben *KubeFlow* und *ClearML* ist *MLFlow* ein Werkzeug, welches auf die teilautomatisierte Dokumentation und Versionierung des Trainings, der Evaluation und der Datenverarbeitung spezialisiert ist. Dieses Tool stellt ebenfalls eine webbasierte Visualisierung bereit.

Weitere open source Werkzeuge, die nicht webbasiert sind, sondern eine lokale Installation erfordern, sind *RapidMiner* und *Knime*. Beide Werkzeuge sind auf das Data-Mining und auf erweiterte Datenanalyse und spezialisiert und nicht auf die Entwicklung vollständiger KI-Anwendungen. Sie basieren auf einer grafischen Visualisierung und Modellierung von Data-Mining-Pipelines, bestehend aus Datenvor- und -nachverarbeitung und Training sowie Ausführung von Modellen.

Kommerzielle Werkzeuge sind beispielsweise *Azure AI* von Microsoft, *SageMaker* von Amazon oder *VertexAI* von Google. Diese Werkzeuge basieren auf einem nutzungsbasierten Preismodell (Pay-per-Use) und sind ebenfalls webbasiert. Sie erlauben ebenfalls die Orchestrierung von KI-Pipelines wie *KubeFlow* und *ClearML*. Darüber hinaus stellen sie auch teilautomatisierte Dokumentation und Versionierung wie *MLFlow* bereit. Im Gegensatz zu den open source Werkzeugen können die kommerziellen Werkzeuge nicht im gleichen Umfang angepasst und erweitert werden [HM22]. Dafür stellen sie einen umfangreichen Support und flexible Rechenleistungen zur Verfügung. Eine Übersicht über die KI-Werkzeug-Landschaft wird von der *AI Infrastructure Alliance*¹² gegeben und aktuell gehalten.

2.4 Community Plattformen

Community Plattformen ermöglichen einen Zusammenschluss von Nutzern und stellen spezielle Dienste für den Informationsaustausch zwischen den Mitgliedern bereit. Eine verbreitete Community-Plattform im Bereich der KI-Entwicklung stellt *OpenML*¹³ [Va13] dar. Dort können KI-Modelle, Daten sowie Evaluationsergebnisse hochgeladen und zwischen den Nutzern ausgetauscht werden. Zusätzlich stellt *OpenML* verschiedene Bibliotheken, beispielsweise für Python [Fe19], bereit, um die geteilten Inhalte möglichst einfach in individuelle Entwicklungs-Projekte integrieren zu können.

Eine weitere und insbesondere in der wissenschaftlichen Community verbreitete Plattform ist *paperswithcode*¹⁴. Dort werden Veröffentlichungen inklusive des dazugehörigen Codes hochgeladen, damit die entsprechenden Inhalte den Mitgliedern der Community zur Verfügung stehen. Darüber hinaus werden KI-Anwendungen nach verschiedenen

¹² <https://ai-infrastructure.org/>

¹³ <https://www.openml.org/>

¹⁴ <https://paperswithcode.com/>

Anwendungsfällen kategorisiert, beispielsweise Computer Vision, Natural Language Processing, Audio- oder Spracherkennung, um den Mitgliedern der Community die Suche zu erleichtern.

Eine der weltweit größten Community-Plattformen für Data-Science und KI-Anwendungen ist *Kaggle*¹⁵. Dort können ebenfalls Datensätze und Modelle hochgeladen und unter den Mitgliedern ausgetauscht werden. Eine Besonderheit bei Kaggle ist, dass auch Unternehmen eigene Datensätze hochladen können und anschließend einen Wettbewerb mit einer datenspezifischen Aufgabe ausschreiben können. Diese Wettbewerbe sind teilweise mit fünfstelligen Preisgeldern dotiert und jedes Mitglied von Kaggle kann an dem Wettbewerb teilnehmen.

2.5 Automated Machine Learning

Automated Machine Learning (AutoML) zielt darauf ab, Anwendungen für maschinelles Lernen automatisch zu erstellen, ohne dass umfassende Kenntnisse in Statistik und maschinellem Lernen erforderlich sind. Bei AutoML wird versucht, die derzeitige Art der Erstellung von ML-Anwendungen durch Automatisierung zu verbessern. Automatisiert werden vor allem das Auswählen, das Erstellen, das Trainieren und das Optimieren eines ML-Ansatzes sowie die Datenvor- und -nachverarbeitung. Aber auch ML-Experten können von AutoML profitieren, indem mühsame Aufgaben wie die Hyperparameter-Optimierung und Datenverarbeitung automatisiert werden. Fachfremde können von AutoML profitieren, weil sie dazu befähigt werden, selbständig diese Aufgaben zu erfüllen, ohne auf einen KI-Experten angewiesen zu sein. [ZH21]

2.6 Design-Patterns

Design-Patterns (dt. Entwurfsmuster) erfassen bewährte Lösungskonzepte für wiederkehrende Probleme in technischen Disziplinen. In der Softwareentwicklung sind Design-Patterns üblicherweise Entwürfe für bewährte Softwarearchitekturen, welche von Softwareentwicklern als Grundlage für eine eigene Lösung verwendet werden können. [Ga94]

Von Lakshmanan et al. [LRM20] werden derartige Design-Patterns speziell für Machine Learning Anwendungen vorgestellt, welche den Entwicklungsprozess von KI-Anwendungen unterstützen können. Es werden bewährte Entwürfe für die Aufbereitung und Verarbeitung von Daten, für das Training, das Verwalten und das Integrieren eines ML-Modells sowie für das Überwachen einer ML-Anwendung gezeigt. Daran anknüpfend werden von

¹⁵ <https://www.kaggle.com/>

Huyen [Hu22] verschiedene Softwarearchitekturen für KI-Anwendungen beschrieben und diskutiert, beispielsweise Edge-basierte, Cloud-basierte und hybride Architekturen. Diese Architekturen werden auch in der [ISO/IEC 22989] vorgestellt.

2.7 Bewertung und Handlungsbedarf

Schnieder [Sc99] argumentiert, dass der Einsatz von Methoden oder Beschreibungsmitteln ohne unterstützende Werkzeuge nur von begrenztem Nutzen für das industrielle Engineering ist, weil in diesem Fall die Mehraufwände und damit die Mehrkosten steigen. Im Idealfall sollten also für eine wirtschaftliche Entwicklung alle Schritte einer Methode durch Werkzeuge unterstützt werden. Diese Unterstützung kann entweder a.) durch ein Werkzeug für alle Schritte oder b.) durch mehrere untereinander kompatible Werkzeuge für einzelne Schritte erfolgen. Damit die Industrie beispielsweise auf der Basis von DMME eine wirtschaftliche KI-Entwicklung durchführen kann, ist also eine durchgängige Werkzeugunterstützung entweder nach a.) oder b.) wünschenswert. Aus diesem Grund wird nachfolgend bewertet, inwieweit die vorgestellten Ansätze die Industrie bereits in den erforderlichen Schritten nach DMME unterstützen und ob diese Unterstützung ausreichend ist.

Die bestehenden KI-Werkzeuge sind darauf ausgelegt, das Training der Modelle (S6), deren Evaluation (S7), die Datenvor- und Nachverarbeitung (S5) und das Produktivsetzen/Deployment (S9) der trainierten Modelle zu unterstützen sowie die Versionierung der Softwareelemente zu vereinfachen (S9). Beim Datenverständnis (S4) können die Werkzeuge durch statistische Methoden und grafische Darstellungen unterstützen. Die Werkzeuge setzen auf das Bereitstellen und Verwalten von Rechenressourcen und auf web-basierte grafische Oberflächen, um Softwareelemente zu kapseln und miteinander zu verbinden. Es besteht jedoch keine Möglichkeit, zusätzliche Informationen und Randbedingungen des zugrundeliegenden technischen Systems mit einfließen zu lassen. Dadurch können diese Werkzeuge nicht beim Identifizieren von Anwendungsfällen im eigenen Unternehmen (S1), dem Formulieren von technischen Zielen (S2), beim Entwurf von verschiedenen technischen Konzepten und Softwarearchitekturen (S3) sowie der Anpassung des technischen Systems (S3) helfen. Bei der Implementierung der KI-Anwendung in das technische System können die Werkzeuge eingeschränkt unterstützen, weil sie verschiedene Rechenressourcen bereitstellen, welche in die Anlage und das technische System integriert werden können (S8).

Das grafische Beschreibungsmittel nach Schieseck et al. [STF23] ermöglicht durch die definierte Symbolik zusammen mit dem Metamodell eine abstrakte Beschreibung von KI-Anwendungen auf Systemebene durch Darstellung der wechselseitigen Abhängigkeiten zwischen KI, technischen Systemkomponenten und dem unterlagerten Fertigungs- sowie Produktionsprozess. Prinzipiell ist es deshalb in der Lage, beim Identifizieren von Anwendungsfällen (S1) zu unterstützen. Darüber hinaus erlaubt die grafische Darstellung

der technischen Zusammenhänge eine Formulierung von technischen Zielen (S2). Darauf aufbauend können verschiedene technische Konzepte und Softwarearchitekturen entworfen (S3), grafisch dargestellt, verglichen, bewertet und getestet und die technische Anlage dahingehend angepasst werden (S3). Ebenso wird die Koordinierung und Durchführung der technischen Implementierung (S8) unterstützt. Das Beschreibungsmittel hilft nur eingeschränkt beim Datenverständnis, weil es zwar technische Zusammenhänge innerhalb der Anlage aufzeigt, jedoch keine konkrete Untersuchung der Daten durchführt (S4). Weil das Beschreibungsmittel keine algorithmischen Inhalte ausführt, hilft es nicht bei der Produktivsetzung (S9), nicht bei der Datenvorverarbeitung (S5), nicht bei der Modellierung (S6) und nicht bei der Evaluation (S7).

Die vorgestellten Community Plattformen erlauben es, die aufgezeichneten Daten des eigenen Anwendungsfalls innerhalb der Community bereitzustellen. Dadurch können die Mitglieder der Community an dem Anwendungsfall mitarbeiten. Darüber hinaus kann auf bestehende Ansätze und Modelle innerhalb der Community Plattformen zugegriffen werden, um diese für den eigenen Anwendungsfall zu nutzen. Diese Plattformen unterstützen deshalb durch Bereitstellung von Daten und Modellen die Datenvorverarbeitung (S5), das Modellieren (S6) und die Evaluation (S7) und in Teilen auch das Datenverständnis.

AutoML zielt darauf ab, die Datenvorverarbeitung (S5), die Modellierung (S6) und die Evaluation (S7) zu automatisieren, weshalb AutoML auch hauptsächlich in diesen Schritten unterstützt. Viele der bestehenden KI-Werkzeuge haben bereits AutoML-Ansätze integriert. Die bestehenden KI Design-Patterns unterstützen beim Entwurf von KI-Systemen auf abstrakter Systemebene (S3) sowie der späteren Integration (S8), zum Beispiel durch die Bereitstellung von Softwarearchitekturen für industrielle Anwendungsfälle unter Einbezug funktionaler und nichtfunktionaler Anforderungen.

Zusammenfassend zeigt die Bewertung auf, dass eine ausreichende Werkzeugunterstützung für die industrielle KI-Entwicklung bisher nicht für alle Entwicklungsschritte nach DM-ME ausreichend gewährleistet ist. Während die Schritte (S5), (S6) und (S7) bereits von bestehenden KI-Werkzeugen unterstützt werden, existiert für die Schritte (S4) und (S8) keine ausreichende und für die Schritte (S1) (S2) und (S3) gar keine Werkzeugunterstützung. Gleichzeitig sind insbesondere die Schritte (S1) und (S2) für die Initialisierung eines KI-Projekts und der Schritt (S3) für die Konzeptionierung der KI-Anwendung von entscheidender Bedeutung [Hu19]. Darüber hinaus ist der Schritt (S8) für die Integration und die Inbetriebnahme der KI-Anwendung und damit für die Wirtschaftlichkeit der Anwendung entscheidend [Hu19]. Die Bewertung wird in Tabelle 1 zusammengefasst. Zumindest in Teilen lassen sich die von Merkel-Kiss et al. [MG23] aufgezeigten Herausforderungen auf diesen Mangel einer entsprechender Werkzeugunterstützung zurückführen. Eine geeignete Werkzeugunterstützung könnte beispielsweise bei der Initiierung und Koordinierung der Entwicklung einer KI-Anwendung unterstützen und die damit die hohen Aufwände, Kosten und Amortisationszeiten reduzieren sowie die komplexen Entscheidungsprozesse innerhalb der Unternehmen vereinfachen.

Tab. 1: Bewertung des Stands der Technik und Wissenschaft. (BU=Business Understanding, TU=Technical Understanding, TR=Technical Realization, DU=Data Understanding, DP=Data Preparation, M=Modeling, E=Evaluation, TI=Technical Implementation, D=Deployment)

		BU	TU	TR	DU	DP	M	E	TI	D
Ansatz	Schritt	S1	S2	S3	S4	S5	S6	S7	S8	S9
	KI Werkzeuge		○	○	○	◐	●	●	●	◐
Beschreibungsmittel		●	●	●	◐	○	○	○	●	◐
Community Plattformen		○	○	○	◐	●	●	●	○	○
AutoML		○	○	○	○	●	●	●	○	○
KI Design-Patterns		○	○	●	○	○	○	○	●	○
		● Geeignet ◐ Teilweise geeignet ○ Nicht geeignet								

3 Anforderungsanalyse

Die Bewertung des Stands der Technik und Wissenschaft hat gezeigt, dass für die Industrie nicht in allen erforderlichen Entwicklungsschritten nach DMME eine ausreichende Werkzeugunterstützung vorhanden ist. Dies betrifft insbesondere die Schritte (S1), (S2), (S3) und (S8). Gleichzeitig hat die Bewertung auch gezeigt, dass diese Schritte durch Anwendung des grafischen Beschreibungsmittels nach [STF23] sowie durch den Einsatz von KI-Design-Patterns unterstützt werden können. Keiner von beiden Ansätzen wurde jedoch bisher in ein Werkzeug integriert und diese sind daher für die industrielle Praxis gegenwärtig nur eingeschränkt von Nutzen. Deshalb ist ein neues Werkzeug erforderlich. Nachfolgend werden die Anforderungen herausgearbeitet, die ein Werkzeug dafür erfüllen muss:

- A1 *Interaktive und grafische Modellierung*: Das Werkzeug muss eine interaktive und grafische Modellierung ermöglichen. Interaktiv bedeutet, dass der Anwender die grafische Modellierung direkt in der grafischen Darstellung, beispielsweise via Drag-and-Drop der grafischen Symbole, durchführen kann. Dadurch sollen die modellierungsbedingten Mehraufwände und die daraus resultierenden Kosten niedrig gehalten werden.
- A2 *Berücksichtigung der Formalisierung des Beschreibungsmittels*: Das Werkzeug muss auf die Formalisierung des Beschreibungsmittels begrenzt werden können, also eine definierte Menge von Symbolen sowie Regeln für deren zulässige Kombination. Dadurch soll das Werkzeug ausschließlich eine Modellierung mit den Symbolen und Regeln ermöglichen, die durch das Beschreibungsmittel vorgegeben sind. Dadurch soll die Reproduzierbarkeit und die Verständlichkeit der Modellierung gesteigert werden.

- A3 *Serialisierte Repräsentation der grafischen Modellierung*: Das Werkzeug muss eine serialisierte Repräsentation der Modellierung bereitstellen, um die Interoperabilität mit bestehenden domänenspezifischen Werkzeugen ermöglichen. Dadurch soll das Werkzeug eine Möglichkeit bereitstellen, bestehende KI-Werkzeuge und verwandte Ansätze an das Werkzeug anzubinden.
- A4 *Einfacher Zugang zum Werkzeug*: Es muss ein einfacher Zugang zu dem Werkzeug für alle an der Entwicklung beteiligten Akteure bestehen. Dadurch soll die interdisziplinäre Zusammenarbeit gefördert und gleichzeitig die mit der Verwendung des Werkzeugs verbundenen Aufwände und Kosten niedrig gehalten werden.

4 Auswahl von Beschreibungsmittel, Methode und Werkzeug

Im Stand der Technik wurde DMME mit den neun Schritten bereits als angepasster de-facto Standard für die industrielle KI-Entwicklung herausgestellt. DMME stellt ein Mindestmaß an absolut notwendigen Schritten für die industrielle KI-Entwicklung bereit. Ebenfalls wurde dargelegt, dass das Beschreibungsmittel nach Schieseck et al. [STF23] eine geeignete Möglichkeit zur grafischen Modellierung von KI-Anwendungen in Automatisierungssystemen ist. Darüber hinaus ist es in der Lage, eben die Entwicklungsschritte von DMME zu unterstützen, die gegenwärtig noch nicht ausreichend unterstützt werden.

Noch nicht betrachtet wurden, welche verschiedene Möglichkeiten zur praktischen Umsetzung und Ausgestaltung eines Werkzeugs existieren. Anhand der Anforderungen werden deshalb nachfolgend verschiedene Implementierungs- und Realisierungsmöglichkeiten bewertet.

Allgemein kann zwischen lokaler Installation und web-basierter **Ausführung des Werkzeugs** unterschieden werden. In beiden Fällen kann prinzipiell eine interaktive und grafische Modellierung bereitgestellt werden (A1). Auch die Berücksichtigung der Formalisierung (A2) sowie die serialisierte Repräsentation (A3) ist in beiden Fällen möglich. Bei einer web-basierten Ausführung kann der Nutzer mithilfe eines beliebigen Browsers auf das Werkzeug zugreifen. Der Aufwand einer lokalen Installation ist nicht vorhanden (A4). Zusätzlich kann eine web-basierte Ausführung einfach um eine Nutzerverwaltung erweitert werden, sodass mehrere Akteure der unterschiedlichen Domänen gleichzeitig Zugriff auf dieselbe Modellierung erhalten können. Darüber hinaus sind viele der bestehenden KI-Werkzeuge ebenfalls web-basiert, wodurch eine Integration in diese KI-Werkzeuge durch eine web-basierte Lösung im Vergleich zu einer lokalen Installation vereinfacht wird.

Bei der **technischen Umsetzung des Werkzeuges** stehen konfigurierbare Werkzeuge, wie beispielsweise Meta Edit, oder Frameworks, wie beispielsweise ReactFlow oder DiagramJS, zur Verfügung. Eine interaktive und grafische Modellierung kann in beiden Fällen realisiert werden (A1). Auch die Berücksichtigung der Formalisierung ist in beiden Fällen möglich (A2). Die konfigurierbaren Werkzeuge sind jedoch bei der Bereitstellung einer serialisierten Repräsentation eingeschränkt, weil diese Serialisierung dort nicht beliebig gestaltbar ist.

Bei Frameworks hingegen ist die serialisierte Repräsentation beliebig gestaltbar (A3). Ein einfacher Zugang wird von web-basierten Frameworks wie ReactFlow und DiagrammJs ermöglicht. Die konfigurierbaren Werkzeuge verfügen teilweise auch über Dienste, welche einen web-basierten Zugang zur Modellierung ermöglichen, beispielsweise EMF Cloud. Allerdings sind diese Dienste und damit die web-basierte Bereitstellung der Modellierung nicht in vergleichbarem Umfang gestaltbar, wie es bei den Frameworks der Fall ist (A4). Für das Werkzeug können verschiedene **grafische Benutzerschnittstellen** verwendet werden. Die einfachste Schnittstelle stellt hierbei ein Command-Line-Interface dar. Dort erfolgt die Interaktion mit dem Programm durch Befehle innerhalb eines Terminals. Eine nutzerfreundlichere Schnittstelle ist ein Editor-Interface, welches aus Textfeldern und Buttons aufgebaut ist, wie beispielsweise beim *AutomationML Editor*¹⁶. Dort erfolgt die Interaktion mit dem Programm durch das Betätigen von Buttons und das Ausfüllen von Textfeldern. Darüber hinaus besteht auch die Möglichkeit eines grafischen Diagramm-Interfaces, welche sich neben Buttons und Textfeldern vor allem durch eine interaktive Arbeitsfläche auszeichnet, auf der verschiedene grafische Elementen erstellt und durch von Drag-and-Drop-Funktionen verbunden werden können. Beispiele für derartige Diagramm-Interfaces sind MetaEdit, ReactFlow oder das Eclipse Modeling Framework. Eine interaktive, grafische Modellierung im Sinne von (A1) kann nur durch ein Diagramm-Interface gewährleistet werden. Zusammenfassend folgt, dass ein web-basiertes Werkzeug mit Diagramm-Interface unter Zuhilfenahme eines Frameworks eine anforderungsgerechte Lösung darstellt. Dieses Werkzeug soll die Methodenschritte (S1), (S2), (S3) und (S8) von DMME unter Anwendung des Beschreibungsmittels von Schieseck et al. [STF23] unterstützen.

5 Konzept des Werkzeugs und Implementierung

Im vorangegangenen Kapitel wurde herausgearbeitet, wie ein anforderungsgerechtes Werkzeug gestaltet werden kann. Nachfolgend wird ein Konzept für ein derartiges Werkzeug und eine Implementierung vorgestellt. Die Implementierung erfolgt auf Basis von React¹⁷ unter Verwendung des Frameworks React Flow und der JavaScript-Laufzeitumgebung Node.js¹⁸. Ein Einblick in die Oberfläche des Werkzeugs wird in Abb. 4 gegeben.

Die grundlegenden grafischen Elemente sind Funktion, Ressource, Produkt, Produkt-/ und Materialfluss, Zuweisung und Kommunikation. Die Elemente Funktion, Ressource und Produkt kann der Anwender mithilfe von Buttons auf der Oberfläche instanzieren. Den instanziierten Elementen können Schnittstellen hinzugefügt werden, über die mithilfe von Drag-and-Drop dann Systemrelationen, also Zuweisung, Kommunikation oder Produkt-/ und Materialfluss, instanziiert werden können. Beispielsweise wird durch das Verbinden von zwei Systemkomponenten automatisch eine Kommunikationsbeziehung zwischen beiden

¹⁶ <https://www.automationml.org/>

¹⁷ <https://react.dev/>

¹⁸ <https://nodejs.org/en>

Systemkomponenten instanziiert. Jedes grafische Element erhält eine eindeutige ID, welche ebenfalls automatisch erstellt wird. Notwendige Attribute für die Kennzeichnung der verwendeten Elemente werden automatisch in einem Datenmodell gespeichert. Dadurch wird ein konsistentes Datenmodell entsprechend des Meta-Modells von Schieseck et al. [STF23] gewährleistet. Bei der Modellierung werden die Modellierungsvorschriften entsprechend der Formalisierung des Beschreibungsmittels nach Schieseck et al. [STF23] berücksichtigt und Fehlmodellierungen durch die Software abgefangen. Mithilfe eines Drop-Down Menüs können für die Systemkomponenten und die Systemfunktionen die entsprechenden Unterklassen ausgewählt werden, beispielsweise die Unterklasse Sensor oder Aktor für die Klasse Systemkomponente. Zusätzlich können weitere Merkmale wie Namen vergeben werden.

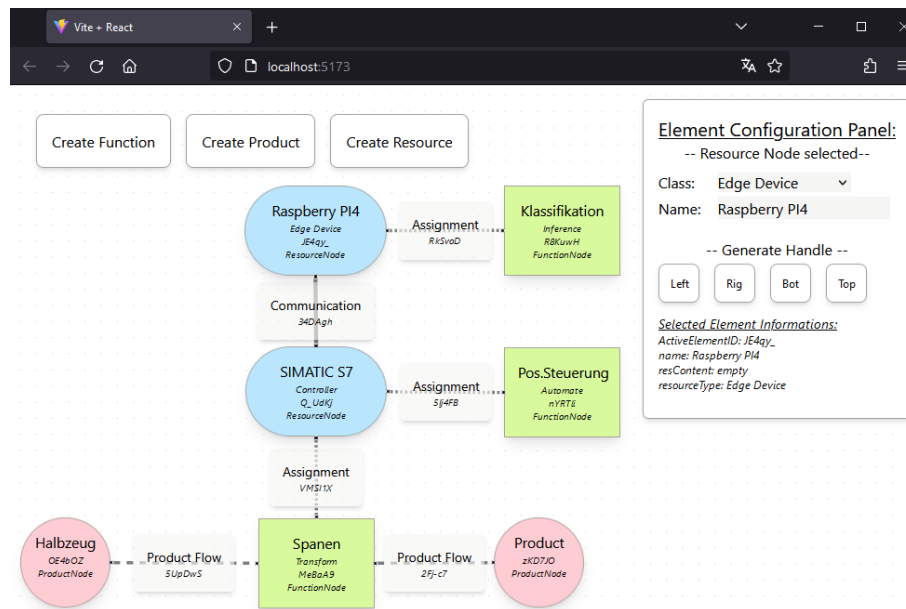


Abb. 4: Grafische Oberfläche des webbasierten Werkzeugs.

Der grafischen Modellierung wird ein JSON-Datenmodell zugrunde gelegt, welches sich schematisch an dem Meta-Modell von Schieseck et al. [STF23] orientiert und parallel während der Modellierung instanziiert wird. In Listing 1 ist ein exemplarischer Auszug des JSON-Datenmodells für die beispielhafte Modellierung aus Abb. 5 dargestellt. Die drei Elemente Funktionen, Produkte und Ressourcen werden in einem Array unter dem Key nodes abgelegt. Die Relationen werden in einem Array unter dem Key edges abgelegt und die Schnittstellen in einem Array unter dem Key handles. Welche Klasse von Element vorliegt, wird mit dem Key type spezifiziert. Die Unterklassen und weitere Attribute jedes Elements werden in einem Array unter dem Key data spezifiziert. Mithilfe weiterer

Keys wie `position`, `width` und `height` werden Informationen zur Positionierung der grafischen Elemente gespeichert. Diese werden jedoch aus Gründen der Übersichtlichkeit hier nicht genauer erläutert und sind in Listing 1 nicht mit aufgeführt. Das instanziierte JSON-Datenmodell kann exportiert werden, um die Modellierung zu speichern, zu teilen oder an andere Werkzeuge anzubinden.

```

{"state": {
  "nodes": [
    {
      "id": "JE4qy_",
      "type": ResourceNode,
      "data" {
        "name": Raspberry PI4,
        "resourceType": Edge Device
      },
      ...
    },
    { ... },
  ],
  "edges": [
    {
      "id": "34DAgh",
      "type": CommunicationEdge,
      "source": "JE4qy_",
      "sourceHandle": "oPrLTc",
      "target": "Q_Udkj",
      "targetHandle": "uSziUr",
      ...
    },
    { ... },
  ],
  handles: [ ... ],
}
}

```

List. 1: Auszug des JSON-Modells

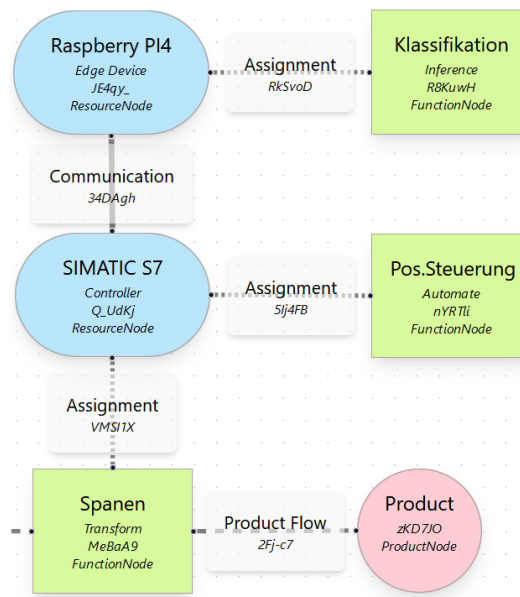


Abb. 5: Beispielhafte Modellierung mit dem Werkzeug

6 Evaluation am Fallbeispiel

Das Werkzeug wird zur Evaluation auf einen realen Anwendungsfall aus der Industrie angewendet. Bei dem Anwendungsfall handelt es sich um eine KI-Anwendung innerhalb einer technischen Anlage zur Überwachung der Produktqualität. Dieser Anwendungsfall wird im Rahmen des Forschungsprojekts *Engineering von künstlichen Intelligenzen*¹⁹ (EKI) umgesetzt und die Anwendung des Werkzeugs erfolgt in Zusammenarbeit mit einem Partner aus der Industrie.

Beschreibung des Anwendungsfalls: In dem Anwendungsfall führt eine technische Anlage

¹⁹ <https://dtecbw.de/home/forschung/hsu/projekt-eki>

einen Fertigungsprozess aus, bei dem ein Schaum aus Polyurethan (PUR) auf den Rand einer Glasscheibe aufgetragen wird. Bevor das PUR aufgetragen werden kann, muss ein Primer auf die Glasscheibe aufgetragen werden. Dieser dient als Haftvermittler zwischen dem PUR und der Scheibe. Der Primerauftrag erfolgt in einer Fertigungszelle bestehend aus einem Roboter, einer Auftragseinheit und einer Kamera. Die Auftragseinheit besteht aus einem Schwamm für den Auftrag des Primers und einer Dosiereinheit zur Befeuchtung des Schwammes. Um die Haftung des PUR auf der Glasscheibe zu gewährleisten, muss die Auftragsfläche des Primers vollständig und ohne Fehlstellen benetzt sein. Ansonsten besteht die Gefahr, dass sich das PUR von der Glasscheibe löst. Das Ziel der KI-Anwendung ist es deshalb, den Primerauftrag zu überwachen und dadurch die Produktqualität der mit Polyurethan (PUR) umschäumten Glasscheibe zu gewährleisten.

Beschreibung des Fertigungsprozesses und der Modellierung: Zu Beginn des Auftragsprozesses nimmt ein Roboter die Scheibe auf und führt diese an dem Schwamm der Auftragseinheit entlang. Jede Scheibe besitzt einen eigenen RFID-Tag und kann so eindeutig zugeordnet werden. Hinter der Auftragseinheit ist eine Kamera montiert, mit deren Hilfe ein Bildstrom des Auftragsprozesses aufgenommen wird. Die Robotersteuerung übernimmt die Steuerung des Roboters und der Dosiereinheit. Die Auswertung der Bilddaten erfolgt auf einem Edge-Gerät mithilfe eines KI-Modells. Wenn die Auswertung der Bilddaten eine Fehlstelle identifiziert, dann erfolgt ein erneuter, punktueller Primerauftrag an der entsprechenden Fehlstelle. Da es sich hierbei um eine zeitkritische Auswertung der Daten bezogen auf den Produktionszyklus handelt, ist es sinnvoll, die Inferenz des KI-Modells nahe am Prozess auf dem Edge-Steuergerät auszuführen. Bevor das KI-Modell ausgeführt werden kann, ist allerdings ein rechenintensives Training des Modells erforderlich. Dieses Training kann nicht auf dem Edge-Gerät durchgeführt werden, weil es nicht über ausreichend Rechenressourcen verfügt. Deshalb werden die aufgenommenen Bilder über das Edge-Steuergerät ebenfalls in die Cloud übertragen und dort vorverarbeitet und gespeichert. Das Training des Modells erfolgt dann in der Cloud, weil dort flexible Rechenressourcen verfügbar sind. Anschließend werden das trainierte Modell und die Datenvorverarbeitung auf das Edge-Device für die Inferenz verschoben und können dort ausgeführt werden. Der beschriebene Anwendungsfall wurde mithilfe des Werkzeugs gemeinsam mit dem Industriepartner modelliert, das Ergebnis ist in Abb. 6 dargestellt.

Durchführung der Modellierung und des Entwicklungsprojekts: Der Industriepartner konnte das Werkzeug nutzen, um zunächst seinen Fertigungsprozess sowie die dazugehörigen Produkte darzustellen. Anschließend konnte er die am Fertigungsprozess beteiligten technischen Systemkomponenten sowie deren Relationen einzeichnen. Dadurch konnte sich der Industriepartner einen Überblick über das technische System verschaffen und zur Koordination und Initiierung des Entwicklungsprojektes nutzen. Der Industriepartner hat darauf aufbauend weitere Akteure wie Softwareentwickler und KI-Experten gezielt in die Konzeptionierung und Entwicklung eingebunden. So konnte in domänenübergreifender Zusammenarbeit mithilfe des Modells geklärt werden, an welchen Stellen und aus welchen Gründen welche Funktionen und Softwarekomponenten sinnvoll sind. Beispielsweise, dass

die Inferenz in diesem Anwendungsfall auf einem Edge-Gerät ausgeführt werden sollte, gleichzeitig das Training allerdings in der Cloud erfolgen muss. Anschließend konnten die jeweiligen Experten der Domänen ihre Entwicklungsaufgaben mit domänenspezifischen Werkzeugen fortführen. Beispielsweise haben die KI-Experten ihre Modelle unter Verwendung von MLFlow entwickelt und evaluiert. Die Automatisierungstechniker haben die Pfadsteuerung des Roboters innerhalb eines domänenspezifischen Steuerungstools angepasst. In regelmäßigen Terminen koordiniert durch das Industrieunternehmen war es so möglich, dass auf Basis der Modellierung verschiedene Experten parallel an der Umsetzung der jeweiligen Funktionen arbeiteten und die KI-Anwendung erfolgreich entwickelt wurde.

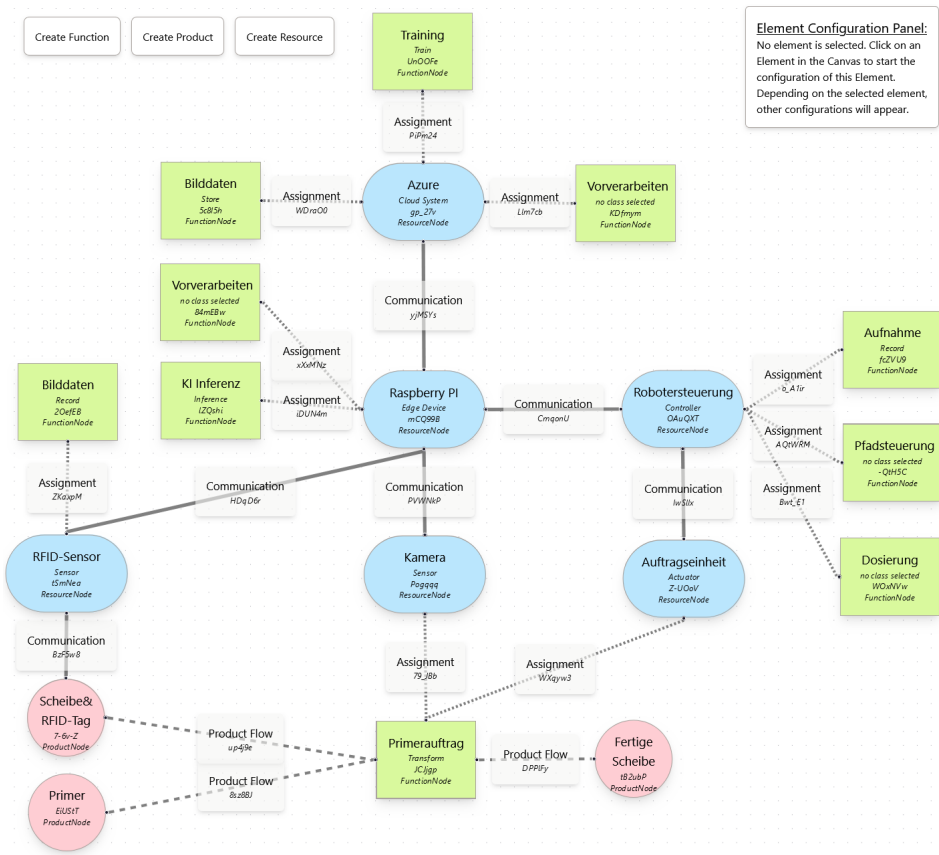


Abb. 6: Modellierung des Anwendungsfalls aus dem Evaluationsbeispiel

7 Diskussion

Auf Basis der Evaluation am Fallbeispiel wird diskutiert, ob das Werkzeug die gestellten Anforderungen A1 bis A4 erfüllt und welchen Beitrag das Werkzeug zum Lösen des Handlungsbedarfs leistet.

Das Werkzeug ermöglicht dem Anwender eine interaktive grafische Modellierung durch Buttons, Drop-Down-Menüs, Eingabefelder und Drag-and-Drop-Funktionen (A1).

Die Modellelemente und Verbindungen sind auf die von Schieseck et al. [STF23] vorgegebenen Elemente beschränkt. Zusätzlich können nur Verbindungen oder Elemente erstellt werden, welche dem Meta-Modell und damit Formalisierung entsprechen. Davon abweichende und fehlerhafte Verbindungen und Elemente werden automatisch vom Werkzeug abgefangen und verhindert (A2).

Während der Modellierung wird automatisch ein Datenmodell im JSON-Format instanziiert und dadurch eine serialisierte Repräsentation der Modellierung erstellt (A3). Das Datenmodell ermöglicht es, die Modellierungsinformationen zu speichern und zwischen verschiedenen Entwicklern und Werkzeugen auszutauschen. Innerhalb bestehenden KI-Werkzeuge wie *KubeFlow* könnten Schnittstellen geschaffen oder angepasst werden, um die Informationen im JSON-Format einzulesen. Der KI-Entwickler würde dann innerhalb von *KubeFlow* über Informationen der geplanten Hardware- und Kommunikationsbedingungen verfügen und könnte diese bei der Entwicklung und Auswahl des KI-Algorithmus berücksichtigen. Beispielsweise könnte es sein, dass die Inferenz in einem Anwendungsfall auf einem Edge Gerät ausgeführt werden muss. Durch den Import könnten die Kennwerte der Hardware und damit etwaige Einschränkungen hinsichtlich der Rechenleistung innerhalb von *KubeFlow* berücksichtigt und für Simulationen genutzt werden.

Über einen Webbrowser haben verschiedene Entwickler die Möglichkeit, auf das Werkzeug zuzugreifen und die grafische Modellierung einzusehen (A4).

Die Evaluation hat gezeigt, dass das Modell und das Werkzeug als Grundlage für die Entwicklung der einzelnen Funktionen und Softwarekomponenten sowie für die Koordination des Projektes verwendet werden kann. Das Werkzeug und die Modellierung können vor allem dabei unterstützen, den Anwendungsfall und das Geschäfts- sowie das technische Ziel zu beschreiben (S1), verständlich darzustellen (S2), ein domänenübergreifendes ein Lösungskonzept zu erarbeiten (S3) und dieses zu integrieren (S8). Durch die webbasierte Bereitstellung des Werkzeugs sowie den interaktiven grafischen Modellierungsansatz waren die Mehraufwände, die durch die Verwendung des Werkzeugs entstanden sind, gering. Insofern wird das Werkzeug als geeigneter Ansatz bewertet, um die Industrie in den bisher nicht ausreichend werkzeuggestützten Entwicklungsschritten der industriellen KI-Entwicklung zu unterstützen. Die Evaluation erfolgte nur an einem Anwendungsfall mit nur einem Industriepartner. Für eine präzisere Evaluierung des Werkzeugs und dessen Nutzens müssen weitere KI-Entwicklungsprojekte von weiteren Industrieunternehmen durchgeführt werden.

8 Zusammenfassung und Ausblick

In diesem Beitrag wurde ein webbasiertes Werkzeug zur Modellierung von KI-Anwendungen für Automatisierungssysteme vorgestellt. Es erlaubt der Industrie eine interaktive, grafische Modellierung von KI-Anwendungen auf Basis eines Beschreibungsmittels für KI-Anwendungen. Die grafische Modellierung besteht aus Systemkomponenten, Systemfunktionen und Systemrelationen, welche durch definierte Symbole dargestellt werden, sowie Regeln zu deren Kombination. Durch die webbasierte Implementierung des Werkzeugs und die interaktive Modellierung wird der zusätzliche Aufwand durch die Modellierung selbst gering gehalten. Das Modell kann in einer JSON-Datei gespeichert werden und dadurch zwischen Entwicklern und Werkzeugen ausgetauscht werden. Durch das Werkzeug werden insbesondere die Schritte der industriellen KI-Entwicklung unterstützt, die durch bestehende KI-Werkzeuge gegenwärtig nicht ausreichend oder gar nicht unterstützt werden. Dadurch wird die Initiierung und Koordination der Entwicklung von KI-Anwendungen aus der Industrie heraus gefördert.

Das Werkzeug stellt gegenwärtig noch ein Funktionsmuster dar und ist ein erster Schritt, um das modellbasierte Engineering von KI-Anwendungen zu ermöglichen. Für die Zukunft ist geplant, das Werkzeug zu erweitern und dem Anwender die Möglichkeit zu bieten, weitere Elemente und Attribute flexibel hinzuzufügen, beispielsweise durch Angabe des Kommunikationstyps oder der Kommunikationsgeschwindigkeit. Zusätzlich ist geplant, das Werkzeug um ein Expertensystem zu erweitern, welches dem Anwender parallel zur Modellierung Hinweise gibt, indem es beispielsweise geeignete Design-Patterns für den aktuellen Anwendungsfall vorschlägt. Außerdem ist für die Zukunft geplant, die Evaluation auszuweiten. Einerseits durch die Modellierung von weiteren realen Anwendungsfällen durch verschiedene Industrieunternehmen und andererseits durch die Durchführung von Umfragen unter Fachleuten.

Acknowledgements

Diese Forschung wird von dtec.bw - Forschungszentrum Digitalisierung und Technik der Bundeswehr gefördert. dtec.bw wird von der Europäischen Union - NextGenerationEU gefördert.

Literatur

- [Bu20] Bundesministerium für Wirtschaft und Klimaschutz (BMWK): Strategie Künstliche Intelligenz der Bundesregierung, 2020.
- [Bu23] Bundesministerium für Bildung und Forschung (BMBF): Aktionsplan Künstliche Intelligenz, 2023.

- [CI23] ClearML: ClearML - Your entire MLOps stack in one open-source tool, Software available from <http://github.com/allegroai/clearml>, 2023, URL: <https://clear.ml/>, Stand: 24.01.2024.
- [Di20] Diemer, J.; Elmer, S.; Gaertler, M.; Gamer, T.; Görg, C.; Grotepass, J.; Kalhoff, J.; Kramer, S.; Legat, C.; Meyer-Kahlen, J.-P.; Nettsträter, A.; Niehörster, O.; Schmidt, B.; Schweichhart, K.; Ulrich, M.; Weitschat, R.; Winter, J.: KI in der Industrie 4.0: Orientierung, Anwendungsbeispiele, Handlungsempfehlungen, Techn. Ber., Plattform Industrie 4.0, 2020.
- [Ei22] Eisenträger, M.; Frey, C.; Herzog, A.; Moghiseh, A.; Morand, L.; Pfrommer, J.; Stephani, H.; Stoll, A.; Wessels, L.: ML4P-Vorgehensmodell Machine Learning for Production, 2022.
- [Fe19] Feurer, M.; van Rijn, J. N.; Kadra, A.; Gijsbers, P.; Mallik, N.; Ravi, S.; Mueller, A.; Vanschoren, J.; Hutter, F.: OpenML-Python: an extensible Python API for OpenML. arXiv 1911.02490/, 2019, URL: <https://arxiv.org/pdf/1911.02490.pdf>.
- [Fr10] Frank, U.: Outline of a method for designing domain-specific modelling languages, Techn. Ber., ICB-research report, 2010.
- [Ga94] Gama, E.; Helm, R.; Johnson, R.; Vlissides, J.: Design patterns, elements of reusable object-oriented software. Addison-Wesley, 1994.
- [Ha19] Habermann, R.; Nagel, P.; Becker, M.; Büchel, A.; von Massow, H.: Systems engineering. Springer, 2019.
- [He19] Hesenius, M.; Schwenzfeier, N.; Meyer, O.; Koop, W.; Gruhn, V.: Towards a software engineering process for developing data-driven applications. In: 2019 IEEE/ACM 7th International Workshop on Realizing Artificial Intelligence Synergies in Software Engineering (RAISE). IEEE, S. 35–41, 2019.
- [HM22] Hewage, N.; Meedeniya, D.: Machine Learning Operations: A Survey on MLOps Tool Support. arXiv 2022. arXiv preprint arXiv:2202.10169/, 2022.
- [HS22] Hasterok, C.; Stompe, J.: PAISE®–process model for AI systems engineering. at-Automatisierungstechnik 70/9, S. 777–786, 2022.
- [Hu19] Huber, S.; Wiemer, H.; Schneider, D.; Ihlenfeldt, S.: DMME: Data mining methodology for engineering applications—a holistic extension to the CRISP-DM model. Procedia CIRP 79/, S. 403–408, 2019.

- [Hu22] Huyen, C.: Designing machine learning systems. O’Reilly Media, 2022.
- [ISO/IEC 22989] International Organization for Standardization (ISO): Artificial Intelligence concepts and terminology, ISO/IEC TR 22989, 2022.
- [ISO/IEC TR 24030] International Organization for Standardization (ISO): Information technology — Artificial intelligence (AI) — Use cases, ISO/IEC TR 24030, 2021.
- [KKH22] Kreuzberger, D.; Kühl, N.; Hirschl, S.: Machine Learning Operations (MLOps): Overview, Definition, and Architecture. arXiv 20222205/, 2022.
- [KKH23] Kreuzberger, D.; Kühl, N.; Hirschl, S.: Machine learning operations (mlops): Overview, definition, and architecture. IEEE Access/, 2023.
- [KWS21] Kaymakci, C.; Wenninger, S.; Sauer, A.: A Holistic Framework for AI Systems in Industrial Applications. In: Innovation Through Information Systems. Springer International Publishing, 2021.
- [LRM20] Lakshmanan, V.; Robinson, S.; Munn, M.: Machine learning design patterns. O’Reilly Media, 2020.
- [Ma21a] Manoury, M.; Brünnhäüßer, J.; Geiger, A.; Lindow, K.; Stark, R.: Modellierungs- und Simulationsbedarfe der intelligenten Fabrik, Techn. Ber., Forschungsbeirat der Plattform Industrie 4.0 / acatech – Deutsche Akademie der Technikwissenschaften, 2021.
- [Ma21b] Martínez-Plumed, F.; Contreras-Ochando, L.; Ferri, C.; Orallo, J. H.; Kull, M.; Lachiche, N.; Quintana, M. J. R.; Flach, P. A.: CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories. IEEE Transactions on Knowledge and Data Engineering 33/, S. 3048–3061, 2021, URL: <https://api.semanticscholar.org/CorpusID:212932421>.
- [MG23] Merkel-Kiss, M.; Garrel, J.: Systematische Literaturanalyse zum KI-Einsatz und KI-basierten Geschäftsmodellen in produzierenden kleinen und mittleren Unternehmen. Zeitschrift für Arbeitswissenschaft 77/3, S. 453–468, 2023.
- [Ro19] Roy Chaudhuri, S.; Natarajan, S.; Banerjee, A.; Choppella, V.: Methodology to develop domain specific modeling languages. In: Proceedings of the 17th ACM SIGPLAN International Workshop on Domain-Specific Modeling. S. 1–10, 2019.
- [Ru21] Ruf, P.; Madan, M.; Reich, C.; Ould-Abdeslam, D.: Demystifying mlops and presenting a recipe for the selection of open-source tools. Applied Sciences 11/19, S. 8861, 2021.

- [Sc99] Schnieder, E.: Methoden der Automatisierung. Vieweg+Teubner Verlag, Wiesbaden, 1999.
- [STF23] Schieseck, M.; Topalis, P.; Fay, A.: A Graphical Modeling Language for Artificial Intelligence Applications in Automation Systems. In: 2023 IEEE 21st International Conference on Industrial Informatics (INDIN). Lemgo, DE, 2023.
- [Va13] Vanschoren, J.; van Rijn, J.N.; Bischl, B.; Torgo, L.: OpenML: Networked Science in Machine Learning. SIGKDD Explorations 15/2, S. 49–60, 2013, URL: <http://doi.acm.org/10.1145/2641190.2641198>.
- [VDI/VDE 3681] Verein Deutscher Ingenieure e.V. (VDI): Einordnung und Bewertung der Beschreibungsmittel aus der Automatisierungstechnik, VDI/VDE 3681, 2022.
- [VDI/VDE 3682] Verein Deutscher Ingenieure e.V. (VDI): Formalisierte Prozessbeschreibungen, VDI/VDE 3682, 2015.
- [VDI/VDE 3714-1] Verein Deutscher Ingenieure e.V. (VDI): Implementation and operation of Big Data application in the manufacturing industry - Implementation of Big Data projects, VDI/VDE 3714-1, 2022.
- [WH00] Wirth, R.; Hipp, J.: CRISP-DM: Towards a standard process model for data mining. In: Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining. Bd. 1, S. 29–39, 2000.
- [ZH21] Zöllner, M.-A.; Huber, M. F.: Benchmark and survey of automated machine learning frameworks. Journal of artificial intelligence research 70/, S. 409–472, 2021.

„Ulrich Jumar, Christian Diedrich (Hrsg.):
EKA 2024 - Entwurf komplexer Automatisierungssysteme, 18. Fachtagung“

Industrielle Sicherheit durch Zertifikatsmanagement- Konzepte im Lebenszyklus einer Produktions-Gray-Box

Marwin Madsen¹, Mike Barth²

Abstract: Die zunehmende IT/OT-Security-Bedrohung für Industrieanlagen führt zu neuen Regelungen wie dem EU Cyber Resilience Act. Gleichzeitig werden Innovationen wie modulare Produktionssysteme vorangetrieben, was zu Veränderungen der Automatisierungsarchitektur und damit auch zu bestehenden Angriffsvektoren führt. Bestehende Sicherheitskonzepte müssen für diese Entwicklungen überprüft werden, wobei die Verwendung von Zertifikaten und damit die Nutzung einer Public-Key-Infrastruktur für viele Sicherheitsmechanismen grundlegend ist. In diesem Artikel werden insbesondere neue Systemarchitekturen wie modulare Produktionszellen oder vernetzte Maschinenteile unter dem Aspekt der Sicherheit analysiert. Dazu werden verschiedene Architekturprinzipien wie Gray- und Black-Box als Grundlage herangezogen und in den Kontext des Zertifikatsmanagements gestellt.

Keywords: IT/OT-Security, PKI, Modularisierung

1 Einleitung

Während die Modularisierung in der diskreten Fertigungsindustrie bereits Stand der Technik ist, werden auch in der Prozessindustrie erste Ansätze verfolgt. Eines der Hauptziele von Modularisierungsstrategien ist es, ein Modul mit einem geringen Engineering-Aufwand hinsichtlich der notwendigen Integration oder Rekonfiguration bereitzustellen. In diesem Zusammenhang werden häufig Begriffe wie Gray- und Black-Box verwendet. Diese Konzepte dienen zwei Zwecken: erstens der Verringerung des Aufwands für den Betreiber und zweitens der Ermöglichung einer generischen Produktion für Hersteller ohne Offenlegung des geistigen Eigentums [Pr19]. Es ist jedoch wichtig, ein Gleichgewicht zwischen Flexibilität und Sicherheit zu finden und da die Sicherheit eines Systems nicht zu Gunsten seiner Flexibilität vernachlässigt werden darf, sind umfangreiche Zusatzkonzepte erforderlich. In diesem Beitrag wird Sicherheit in der Ausprägung IT/OT-Security thematisiert. Solche Zusatzkonzepte für die Sicherheit sind abgestimmt auf die Modularisierung lediglich vereinzelt veröffentlicht vorhanden [Ma23]. Dies steht jedoch im Widerspruch zu Konzepten wie „Security by Design“ und aufkommenden Regulierungen, beispielsweise in Form des European Cyber Resilience Act (CRA) [Ch23]. Darüber hinaus weisen jüngste Berichte auf zahlreiche Schwachstellen in klassischen Gerätefamilien und die kontinuierliche Entdeckung neuer Malware hin

¹ Karlsruher Institut für Technologie, Institut für Reglun- und Steuersysteme, Fritz-Haber-Weg 1, 76131

Karlsruhe, marwin.madsen@kit.edu,  <https://orcid.org/0009-0006-9953-2382>

² Karlsruher Institut für Technologie, Institut für Reglun- und Steuersysteme, Fritz-Haber-Weg 1, 76131

Karlsruhe, mike.barth@kit.edu,  <https://orcid.org/0000-0003-2337-063X>

[Gt23]. Die richtige Balance zwischen Flexibilität und Sicherheit zu finden, bleibt ein kritischer Aspekt moderner industrieller Prozesse.

Eine zentrale Komponente der Sicherheit in der OT ist die Verwendung von X.509-Zertifikaten für Aspekte wie Verschlüsselung, Authentizität und Integrität [Se23]. Diese Zertifikate erfordern sowohl eine Public-Key-Infrastruktur (PKI) als auch Mechanismen zur Verteilung der Zertifikate. Dies kann von manueller Interaktion bis hin zu vollautomatischen Zertifikatsmanagement (ZM) wie dem Certificate Management Protocol (CMP) [Ad05] reichen. Viele der in der OT verwendeten Technologien oder Protokolle erfordern für die Sicherheitsmechanismen entsprechend ausgerollte Zertifikate. Insbesondere die Standardisierungsorganisation PI (PROFIBUS\&PROFINET International), ODVA, Inc., OPC Foundation und FCG (FieldComm Group), die Protokolle wie PROFINET, Ethernet/IP, OPC UA und HART-IP spezifizieren, setzen auf diese als Schlüsselement [Wa22, Op23, Ni19]. Einige dieser Organisationen definieren auch ihre eigenen Spezifikationen für ZM. So beschreibt beispielsweise die OPC Foundation sowohl grundlegendes ZM in OPC UA Part 2 [Op23] als auch einen so genannten CertificateManager in OPC UA Part 12 [Op22a] sowie die Authentifizierung eines neuen Gerätes und das Ausrollen von Zertifikaten durch den Betreiber (Onboarding) in OPC UA Part 21 [Op22b]. Obwohl diese Ansätze für ihre jeweilige Industriedomäne bzw. die vorgesehene Anlagenarchitektur funktionieren, bleiben konzeptionelle Fragen im Hinblick auf eine generische und architekturunabhängige Forschungsperspektive bestehen. Die einzelnen Ansätze führen zu impliziten Architekturentscheidungen wie dem direkten Kontakt mit einem Registrar in OPC UA Part 21 oder einer PKI Management Entity in CMP.

Im Folgenden soll der Begriff "*Modul*" nicht auf ein unternehmens- oder domänenspezifisches Konzept oder ein entsprechendes Informationsmodell, wie z.B. das Module Type Package [Au20] aus der Prozessindustrie, bezogen werden. Um diese Unterscheidung zu betonen, wird im Folgenden anstelle des Begriffs Modul der Begriff Produktions-Gray-Box (PGB) verwendet. Die PGB ist eine Gerätekomposition mit mindestens einer dedizierten Modulschnittstelle (MS). Zusätzlich enthält diese Komposition weitere sogenannte modulinterne intelligente Einheiten (MIE). Intelligent bezieht sich in diesem Zusammenhang auf die zugehörige Kommunikation auf einer Ebene, die die Verwendung von Zertifikaten einschließt. Wenn diese untergeordneten MIE nach außen hin nicht offengelegt werden, ergeben sich aus Sicht des Lebenszyklus von Zertifikaten erhebliche Sicherheitsprobleme. Abb. 1 zeigt die Ausgangssituation beim Onboarding einer PGB. Der Hersteller ist für die funktionierende Kommunikation in der PGB verantwortlich und hat daher bereits entsprechende Zertifikate ausgerollt. Die einzige Kommunikationsmöglichkeit für die PKI des Betreibers ist über die MS, weshalb Zertifikate nicht direkt an die MIE verteilt werden können.

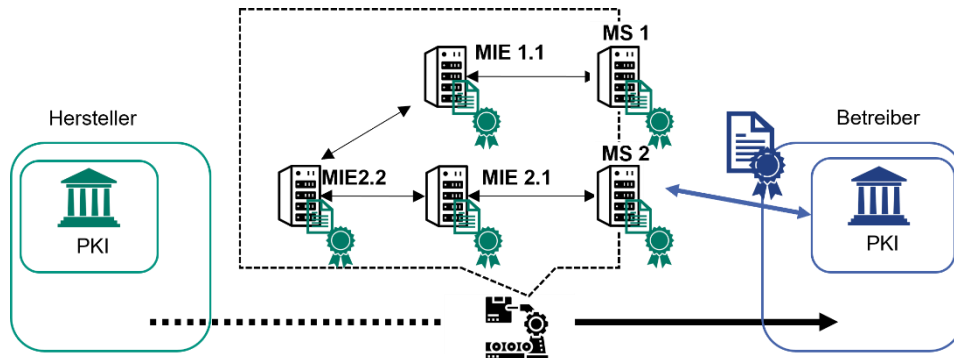


Abb. 1: Betrachtete modulare Architektur und Ausgangssituation während des Onboardings

Es ist zu betonen, dass die hier betrachtete Architektur einer PGB eine entscheidende Voraussetzung ist und nicht in jedem Fall zutrifft. So kann beispielsweise eine PGB, die neben der MS intern nur mit Motoren oder Sensoren mittels 4-20 mA Stromstärke kommuniziert, keine Zertifikate nutzen. Moderne IoT-Konzepte tendieren jedoch zu immer komplexeren PGB mit zahlreichen MIE. Sobald ein Plug&Produce-Konzept angestrebt wird, ist die hierarchische Struktur aus Abb. 1 eine sinnvolle Annahme. Einfache Beispiele finden sich bereits in der Praxis [Si24], in Veröffentlichungen von Standardisierungsorganisationen wie NIST [St22] und wurden auch in Gesprächen mit Industriepartnern des Instituts für Regelungs- und Steuerungssysteme verifiziert. Im Maschinenbau ist es üblich, einen Motion Controller einzusetzen, der die Antriebe, Aktoren und Sensoren steuert. Der Maschinenbauer installiert den Motion Controller dann als Ethernet-basiertes Feldbusgerät unterhalb einer speicherprogrammierbaren Steuerung (SPS). Ein Beispiel für den Feldbus wäre PROFINET. Die Antriebe können in der Regel über eine geroutete Engineering-Kommunikation erreicht werden, um deren Zustand zu überwachen oder Änderungen an deren Konfiguration vorzunehmen. Diese Maschinenzelle wird dann in die Anlage des Betreibers integriert, wobei die SPS als Schnittstelle zur Maschine fungiert. Darüber hinaus sind die von der SPS vom Anlagennetzwerk abgekapselten Komponenten oft noch direkt mit einem mobilen Engineering-Gerät z.B. über eine USB-Schnittstelle erreichbar. Dies entspricht dem einfachsten Fall der in Abb. 1 betrachteten Architektur, in der es eine MS und eine MIE sowie eine große Anzahl weiterer, nicht im Fokus stehender Komponenten gibt.

2 Threat Modelling

Die Notwendigkeit von Zertifikaten im Allgemeinen lässt sich aus bestehenden Sicherheitskonzepten und deren Abhängigkeit von der Verwendung von Zertifikaten ableiten [Wa22, Ni19, Op23]. Allerdings muss die Bedrohungslage berücksichtigt werden, wenn Entitäten innerhalb einer PGB keinen Zugang zu Zertifikaten haben.

Eine sorgfältige Netzwerksegmentierung mit geeigneten Intrusion-Prevention-Systemen verhindert, dass ein Angriffsvektor in die PGB geschaffen wird, da die Kommunikation nur über die dedizierte MS möglich ist. Bestehende und bewährte ZM-Konzepte sind vollständig auf die MS anwendbar, da ein direkter Kontakt mit der PKI möglich ist. Allerdings ist zu bedenken, dass es zwar zunächst keinen Kommunikationskanal auf Softwareebene zu den MIE neben der MS gibt, aber ein physischer Zugriff möglich sein kann. Zu diesem Zweck wurde früher das Konzept der Sicherheit durch Abschottung verwendet, das heute jedoch als durchlässig gilt [Hu19]. Obwohl das Ziel darin besteht, so weit wie möglich eine Black-Box zu schaffen, ist dies nicht ganz zweckmäßig. So muss es beispielsweise möglich sein, defekte MIE auszutauschen, und in manchen Umgebungen kann eine bewusste Entscheidung gegen Remote-Updates getroffen werden, was bedeutet, dass ein lokaler Zugriff zur Installation dieser Updates möglich sein muss. Die MS kann auch durch sogenannte Zero-Day-Exploits (Angriffe am ersten Tag nach der Entwicklung eines neuen Angriffsvektors) kompromittiert werden, insbesondere wenn sie für die Kommunikation in der Cloud bestimmt ist. Eine Kompromittierung dieser würde sich bis in die PGB fortsetzen, da eine zentrale Instanz nicht wie bei Zertifikaten nach der Entdeckung eine Revokation vornehmen kann, sondern ein manueller Eingriff notwendig ist. Nach dem Defense-in-Depth-Konzept [KF06] sollte es mehrere Sicherheitsschichten geben. Bei automatisiertem ZM wäre es sinnvoll, die Vorteile der Zertifikate bis in die PGB hinein zu nutzen. Insbesondere wenn PGB durch die neuen Entwicklungen einen dynamischeren Kontext aufweisen als klassische Varianten, müssen Fallback-Mechanismen für mögliche Fehlerpotentiale berücksichtigt werden.

Zahlreiche Bedrohungsanalysen der letzten Jahre sind auf die hier betrachtete Architektur adaptierbar und behalten Gültigkeit in Ihrer Argumentation [NM22]. Darüber hinaus betonen sowohl Kommunikationstechnologien als auch Sicherheitsstandards die Bedeutung einer PKI [Wa22, Se23]. Insbesondere die Arbeiten im Bereich Ethernet-APL zeigen den Bedarf an Zertifikaten. Diese basieren auf einem konvergenten Netzwerk, also dass alle Komponenten einer Produktionszelle an das gleiche Netz angeschlossen sind und ein einheitliches Protokoll verwendet wird. Das heißt, die hierarchische Struktur klassischer Automatisierungssysteme wird aufgebrochen und alle Komponenten können alle anderen zumindest kommunikativ erreichen. So hätten beispielsweise Engineering- und Operator-Stationen direkten Zugriff auf Feldgeräte. Dies berücksichtigt jedoch nicht die Entwicklung flexibler Automatisierungsarchitekturen, in denen zwar ein konvergentes Netzwerk möglich ist, aber dennoch hierarchische Strukturen zwischen MS und einer Orchestrierungsebene bestehen. Wird dieses Merkmal in die Architektur der Bedrohungsanalyse integriert, bleibt die Notwendigkeit für den Einsatz von Zertifikaten bestehen: Die bisher verwendeten Protokolle zwischen Feld und Steuerung - nämlich PROFIBUS DP, PROFIBUS PA und HART - ermöglichten keine sichere Kommunikation. Angesichts der zunehmenden Verbreitung von IP-basierten Protokollen, die durch höhere Datenraten und eine vereinfachte Installation, Konfiguration und Wartung von Geräten und Automatisierungstechnik getrieben werden, wäre das einzige Gegenargument gegen den Einsatz ihrer Sicherheitsmechanismen, wenn ein unverhältnismäßig hoher Aufwand für deren Implementierung notwendig wäre [Se23].

3 Analyse bestehender Spezifikationen

Bestehende Spezifikationen gehen bereits umfassend auf das ZM während des gesamten Lebenszyklus eines Geräts ein. Insbesondere das Onboarding und die nachfolgenden Interaktionen mit dem Betreiber werden detailliert betrachtet. Das Ausrollen von Zertifikaten in die MS ist bereits ein üblicher Betrachtungspunkt, da sie direkt mit dem Netz des Betreibers kommuniziert, welches eine Komponente zur Anbindung an die PKI enthalten sollte. Im Folgenden soll geprüft werden, ob sich die beschriebene Architektur einer PGB nahtlos in die bestehende Landschaft einfügt, insbesondere wenn auch die MIE mit Zertifikaten versorgt werden sollen.

3.1 Certificate Management Protocol

Es gibt eine Vielzahl von ZM-Protokollen, die hier nicht behandelt werden können. Eines der umfassendsten Protokolle ist das CMP, das den größten Teil des Lebenszyklus eines Zertifikats abdeckt. Es ist zu beachten, dass der Schwerpunkt des Lebenszyklus auf dem Zertifikat und nicht auf dem Gerät liegt. Es basiert auf authentifizierten, in sich geschlossene Nachrichten und ist daher nicht auf spezielle Transportprotokolle wie TLS beschränkt. Stattdessen kann es auch in Umgebungen mit eingeschränkten Ressourcen verwendet werden. Für CMP werden in diesem Abschnitt sowohl der ursprüngliche Standard in RFC 4210 [Ad05] als auch der industriegetriebene Entwurf eines Lightweight CMP Profile [BOF23] analysiert. RFC 4210 klassifiziert PKI-Entitäten in End Entities (EE), Registration Authorities (RA) und Certification Authorities (CA). Die beiden letztgenannten Entitäten sind Teil des PKI-Managements, und wenn keine Notwendigkeit besteht, zwischen ihnen zu unterscheiden, werden sie als PKI-Management-Entität bezeichnet. Im RFC 4210 wird nicht explizit auf die Möglichkeit von Proxys zwischen der EE und einer PKI Management Entity eingegangen. RFC 6712 hingegen enthält eine Spezifikation von HTTP als Transportprotokoll für CMP und weist darauf hin, dass HTTP die Überwindung von Netzwerkgrenzen durch die Verwendung von allgegenwärtigen Proxys erleichtert [KP12]. CMP-Nachrichten enthalten zwar keine sensiblen Informationen, die die PKI-Sicherheit gefährden könnten, aber es besteht die Möglichkeit, dass vertrauliche technische oder geschäftskritische Informationen beobachtet werden. Daher wird die Verwendung von TLS oder VPNs empfohlen. Die MS genießt zwar aufgrund ihres Onboardings ein gewisses Vertrauen, doch sollte dieses nicht allumfassend sein. Stattdessen sollte es für die vorgesehenen Aufgaben nach dem Prinzip der geringsten Privilegien eingeschränkt werden. Dieser Ansatz stellt sicher, dass nur die notwendigen Berechtigungen erteilt werden und mindert das Risiko eines möglichen Missbrauchs der Proxy-Stellung. Ein zentraler Baustein einer PKI ist die Sperrung von Zertifikaten und der damit verbundene Zugriff auf Sperrlisten. In diesem Zusammenhang wäre ein solcher Zugriff jedoch nur über die MS möglich. Folglich würde sich jede Kompromittierung der MS direkt auf die gesamte PGB auswirken. Ein anderer Ansatz wäre die MS als RA für die PGB zu nutzen. Es wird sogar ausdrücklich betont, dass die Einrichtung von RAs ein Implementierungsproblem ist, und es lediglich vorausgesetzt wird, dass eine RA

zertifiziert ist und über einen privaten Schlüssel zum Signieren verfügt. Ein MS könnte daher nach dem Onboarding als RA betrachtet werden. Die CMP-Spezifikation RFC 4210 definiert jedoch mehrere obligatorische PKI-Funktionen (sowie weitere empfohlene und optionale Funktionen), die RAs bereitstellen müssen und die keinen sinnvollen Overhead für MS darstellen würden.

Zusätzlich zu den bestehenden Entitäten spezifiziert das Lightweight CMP Profile auch eine Local Registration Authority (LRA), die eine RA mit Nähe zur Endentität darstellt. Dies hebt zum einen hervor, dass Produktionssysteme hierarchisch aufgebaut sind und die meisten Entitäten keinen direkten Kontakt zu zentralen Instanzen haben und zum anderen, dass PKI-Verwaltungsinstanzen wie CAs in besonders abgeschotteten Umgebungen gesichert werden müssen. Anstatt Zertifikatsanfragen durch mehrere Netzwerke zu einer zentralen CA oder RA zu schleifen, kann eine LRA bereits innerhalb desselben Netzwerks wie die Entität Prüfungen durchführen. Allerdings wird auch in diesem Zusammenhang nicht diskutiert, ob sich eine LRA zwingend im gleichen Netz befinden muss oder ob CMP-Nachrichten ohne Sicherheitsbedenken über einen Proxy an die nächste PKI-Verwaltungsinstanz gesendet werden können. In der hier betrachteten Architektur ist jedoch genau dies der Fall. Die MS ist nämlich keine klassische LRA, sondern ein notwendiger Proxy. Wenn die MS an die Stelle einer LRA treten sollte, gibt es den Anwendungsfall "FwdKeep", bei dem die CMP-Nachricht ohne Änderungen weitergeleitet wird und somit de-facto als Proxy fungiert. Obwohl die MS diese Proxy-Funktionalität bereitstellen könnte, müssen LRAs auch mindestens "FwdAddS" ermöglichen, d. h. das Hinzufügen einer Schutzschicht zur bestehenden CMP-Nachricht vor deren Weiterleitung. Die ursprüngliche Authentifizierung wird also nicht verändert, während die LRA gemäß der Spezifikation die Anfrage erfolgreich validiert und genehmigt. Darüber hinaus sollte eine LRA auch weitere PKI-Verwaltungsfunktionen implementieren, einschließlich der Fähigkeit, Zertifikate im Namen von Endteilnehmern zu widerrufen.

Es bleibt unklar, ob bei der Kommunikation zwischen EE und LRA eine Berücksichtigung von Proxys vorgesehen ist oder wie eine LRA eingerichtet wird. Insbesondere da der Proxy kein vorheriger Bestandteil des Anlagennetzwerk des Betreibers wäre. Beide Ansätze sind jedoch gültig und sollten bewertet werden.

3.2 OPC UA

Als Kommunikationstechnologie verfügt OPC UA über eine umfassende Spezifikation für Zertifikate, einschließlich Part 2 [Op23], 12 [Op22a] und 21 [Op22b], die in Bezug auf die PGB-Architektur analysiert wurden. Teil 2 unterstreicht, dass Zertifikate die grundlegende Komponente zur Gewährleistung der Sicherheit in OPC UA sind. Der Lebenszyklus eines Geräts nach Part 21 verdeutlicht, dass Sicherheitskonzepte unter potenziell komplexen Lieferketten, mehreren beteiligten Akteuren und komplexen Architekturen berücksichtigt werden müssen. Wo es früher nur ein Gerät mit einer Herstellerzertifikate als Ausgangspunkt gab, gibt es jetzt ein zusammengesetztes Gerät

oder, in der Bezeichnung dieser Arbeit, eine PGB, die ein oder mehrere MS und MIE hat. OPC UA Part 21 erkennt an, dass in der Industrie nicht nur isolierte Einheiten, sondern miteinander verbundene Systeme involviert sind. Diese Systeme bestehen aus Gerätenetzwerken, in denen PGB durch andere Maschinenbauer zu neuen PGB erweitert werden können. Es wird davon ausgegangen, dass PGB eine Abstraktion innerhalb eines Netzes sind und nur von einem oder mehreren nach außen sichtbaren Geräten, d. h. von ihren MS, physisch erreicht werden können. Es wird davon ausgegangen, dass die internen Interaktionen zwischen MIE vom Maschinenbauer konfiguriert werden.

Die entscheidende Divergenz zwischen Part 21 und der bestehenden Bedrohungsanalyse in Abschnitt 2 besteht darin, dass OPC UA allenfalls die Bereitstellung von Zertifikaten an die MS anerkennt. Dies spiegelt sich in der Definition der Composite Identity wider, in welchem bereits die Provisionierung der MS nur in die Kategorie „may“ eingeordnet wird und MIE nicht erwähnt werden. Außerdem wird das Composite-Konzept in erster Linie nur für den Authentifizierungsprozess verwendet, nicht aber für das anschließende Ausrollen der Zertifikate. Die Device Configuration Application (DCA) wird als zuständige Instanz für die Verwaltung der Zertifikate für die Anwendungen definiert und läuft auf Geräteebene. Aufgrund der Annahmen in der Composite Identity, dass höchstens MS, provisioniert werden, gibt es keine Spezifikationen des Zusammenspiels der DCA, die auf dem MS läuft, und den MIE. Dies setzt sich in Part 12 fort, wo, ähnlich wie bei den ZM-Protokollen, eine direkte Kommunikation zu einem CertificateManager gefordert wird, der die Form einer RA darstellt. Somit bietet OPC UA derzeit keine expliziten Mechanismen für die Provisionierung der MIE außer manuellen Methoden.

4 Ableitung von Forschungsfragen und Lösungsvektoren

Ein naheliegendes Konzept besteht darin, die MIE in das bestehende ZM zu integrieren. Hierbei können viele Betrachtungen aus klassischen Systemen adaptiert werden. Es bleibt zu klären, ob die MS das ZM für alle MIE gegenüber der PKI des Betreibers übernehmen kann, oder ob eine vollständige Implementierung bis hinunter zu jeder MIE notwendig ist. Typischerweise ist die Beantragung im Namen einer anderen Entität nur für besonders vertrauenswürdige Geräte zulässig, da sonst zusätzliche Angriffsvektoren geschaffen würden. Es ist daher notwendig, die Bedingungen für die sichere Einrichtung einer RA zu spezifizieren und dann zu prüfen, ob bestehende Onboarding-Prozesse der MS ein ausreichendes Sicherheitsniveau bieten können. Es ergibt sich die folgende Forschungsfrage:

R1. Was sind die Anforderungen an eine MS und einen Onboarding-Prozess um das ZM für MIE zu übernehmen?

Die Alternative wäre, dass die MIE das gleiche Onboarding durchlaufen wie die MS. MIE sind jedoch in der Regel nicht Teil desselben Netzes wie die PKI oder die RA. Folglich müssten sie die MS als Stellvertreter nutzen. Dies zwar theoretisch machbar, würde aber verfahrenstechnische Anpassungen des Onboardings erfordern. Beispielsweise spielt bei

OPC UA Part 21 ein Discovery-Service eine zentrale Rolle. Eine Anpassung könnte auch vorsehen, dass die PGB über die MS hinreichend authentifiziert werden, so dass die MIE nur noch in das Ausrollen der Zertifikate integriert werden müssten. So könnte die MS nach erfolgreichem Onboarding die MIE mit Daten über die PKI-Management-Entity versorgen. Anschließend könnten die MIE den Prozess der Zertifikatsbeantragung mit der MS als Proxy einleiten. Wie jedoch in Abschnitt 3 hervorgehoben wurde, fehlt in der derzeitigen Landschaft eine explizite Sicherheitsbetrachtung hinsichtlich der Verwendung von Proxys im ZM. Insbesondere die Revokation ist noch nicht hinreichend für diesen Fall betrachtet. Auch das Zusammenspiel zwischen MS und MIE müsste spezifiziert werden, da automatisierte Prozesse nur über standardisierte Schnittstellen möglich sind.

R2. Wie können neue Angriffsvektoren abgeschwächt werden, wenn die MS als Proxy für das ZM der MIE verwendet wird?

Eine weitere Konzeptvariante im Sinne der Modularisierung sieht die Entwicklung einer modulminimalistischen PKI vor. Diese PKI soll auf möglichst wenige Funktionalitäten reduziert werden, so dass sie ZM intern innerhalb eines PGB übernehmen kann. Ähnliche Ansätze werden bereits in anderen Domänen verwendet, z.B. in mobilen Umgebungen [TB08]. Der Grund dafür liegt darin, dass die MIE selbst nicht mit dem System des Betreibers interagiert. Sie benötigt lediglich Zertifikate für die interne Kommunikation. Es besteht daher keine zwingende Notwendigkeit, diese in die PKI des Betreibers zu integrieren. Allerdings bedeutet dieser Ansatz zunächst einen technologischen und prozessualen Mehraufwand für die PGB, der in späteren Arbeiten detailliert untersucht werden soll. Automatisierte Engineering-Methoden können diesen Prozess vereinfachen. Ermöglicht wird diese Konzeptvariante insbesondere durch die von dem CRA vorgeschriebenen Updates [Ch22]. Diese Updates ermöglichen eine kontinuierliche Anpassung der internen PKI an die laufende Evaluierung von kryptographischen Algorithmen. Von besonderem Interesse für eine solche modulminimalistische PKI sind Weiterentwicklungen von PKI-Ansätzen aus der Informatik. Während die Herausforderungen einer PKI in der OT - vor allem die Komplexität - und in der IT divergieren, bietet eine PGB ein relativ geschlossenes System und Anpassungspotential.

R3. Wie können bestehende ZM-Ansätze so angepasst oder erweitert werden, dass sie die autonome Natur der betrachteten Architektur nutzen können?

Die Analyse der bestehenden Spezifikationen zeigt einen weiteren Bedarf auf. So zielen beispielsweise Modularisierungsstrategien auf einen Plug&Produce-Prozess ab, bei dem Schnittstellen definiert werden und Konfigurationsparameter möglichst automatisch übernommen werden können. Die Verwendung von Zertifikaten allein bis hin zur MS, geschweige denn innerhalb der PGB, unterscheidet sich jedoch je nach PGB-Typ und der PKI-Unterstützung des Betreibers für das Onboarding und das nachfolgende ZM. Selbst wenn die PKI des Betreibers mehrere der zahlreichen Technologien unterstützt, gibt es keine Möglichkeit, dies in den bestehenden Informationsmodellen automatisch auszulesen. Selbst die viel diskutierte Asset Administration Shell bietet noch kein Submodell, das den Modularisierungsansatz auf die Sicherheit überträgt. Zwar ist ein

Teilmodell namens Security Engineering in der Entwicklung, doch liegt dessen Schwerpunkt eher auf der Ermöglichung strukturierter Sicherheitsentscheidungen durch die Hersteller [Re24]. Dies schließt natürlich die Anwendbarkeit auf die dargestellten Herausforderungen nicht aus, aber es ist offensichtlich, dass das Erreichen von Plug&Produce derzeit mit ZM nicht möglich ist, ganz zu schweigen von den verschiedenen anderen Sicherheitsmechanismen.

R4. Inwieweit ist das Ziel der Modularisierung, d.h. eine herstellerunabhängige, automatische Integration mit geringem Aufwand, für Sicherheitsmechanismen realisierbar und welche Modifikationen sind erforderlich?

5 Zusammenfassung & Ausblick

In dieser Arbeit wurde die Architektur einer PGB definiert und untersucht. Bei dieser Architektur könnte es sich zum Beispiel um eine Maschinezelle mit mehreren internen Geräten und internen Kommunikationsprotokollen handeln, die auf X.509-Zertifikaten für die Sicherheit beruhen. Die PGB-Architektur wird in der Industrie bereits eingesetzt und dürfte sich aufgrund der fortschreitenden technologischen Entwicklung weiterverbreiten. Es wurde eine prägnante Bedrohungsanalyse speziell für diese Architektur durchgeführt, die auf das Rational bestehender Bedrohungsanalysen für Zertifikate in OT-Systemen zurückgreift. Die Ergebnisse zeigen, dass Zertifikate nicht nur für die PGB-Schnittstelle im Anlagennetzwerk, sondern auch für die internen Geräte erforderlich sind. Anschließend wurde die Anwendbarkeit bestehender und etablierter ZM-Protokolle und ZM aus Kommunikationstechnologien im Zusammenhang mit dieser Architektur bewertet. Es ergibt sich, dass diese bestehenden Ansätze nicht direkt auf die PGB-Architektur angewendet werden können, ohne dass zusätzliche Sicherheitsbewertungen oder eine explizite Berücksichtigung möglicher Probleme vorgenommen werden. Abschließend wurden mögliche Anpassungsansätze identifiziert und Forschungsfragen skizziert, die in zukünftigen Arbeiten beantwortet werden.

6 Literaturverzeichnis

- [Pr19] Process Industry 4.0: The Age of Modular Production - On the Doorstep to Market Launch. ZVEI – Zentralverband Elektrotechnik und Elektronikindustrie e. V., 2019.
- [Ma23] Madsen, M.; Palmin, A.; Stutz, A.; Barth, M.: Security Analysis of the Module Type Package Concept. In: 21st IEEE International Conference on Industrial Informatics. S. 1-8, 2023.
- [Ch23] Chiara, P.G.: The Cyber Resilience Act: the EU Commission’s proposal for a horizontal regulation on cybersecurity for products with digital elements. In: Int. Cybersecur. Law Rev. 3, S. 255–272, 2022.

-
- [Gt23] Ginter, A.; Hale, G.; Machtemes, R.; Molina, J.; Wallhof, M.; Schneider, C.: OT Cyber Threats Report, Waterfall Security Solutions, 2023.
- [Se23] Security for Industrial Automation and Control Systems. International Electrotechnical Commission (IEC), IEC 62443, 2009–2020.
- [Ad05] Adams, C.; Farrell, S.; Kause, T.; Mononen, T.: Certificate Management Protocol (CMP). Network Working Group, RFC 4210, 2005.
- [Wa22] Walz, A.; Berndt, D.; Visoky, J.; Koppers, J.; Wiberg, J.; Armstrong, R.; Vincent, S.; Merklin, S.: FAQ on Industrial Ethernet Security Concepts. Industrial Ethernet Security Harmonization Group, 2022.
- [Ni19] Niemann, K.-H.: IT security extensions for PROFINET.: In IEEE 17th International Conference on Industrial Informatics (INDIN). S. 407-412, 2019.
- [Op23] OPC 10000-2: UA Part 2: Security. OPC Foundation, 2023.
- [Op22a] OPC 10000-12: UA Part 12: Discovery and Global Services. OPC Foundation, 2022.
- [Op22b] OPC 10000-21: UA Part 21: Device Onboarding. OPC Foundation, 2022.
- [Au20] Automatisierungstechnisches Engineering modularer Anlagen in der Prozessindustrie - Allgemeines Konzept und Schnittstellen. Gesellschaft Mess- und Automatisierungstechnik, VDI/VDE/NAMUR 2658, 2013/2020.
- [Si24] SIMOTION D - Drive-based, Siemens AG, mall.industry.siemens.com/mall/de/WW/Catalog/Products/10014179, Stand: 30.04.2024.
- [St22] Stouffer, K.; Pease, M.; Tang, C.; Zimmerman, T.; Pillitteri, V.; Lightman, S.: Guide to Operational Technology (OT) Security. National Institute of Standards and Technology, NIST SP 800-82 Rev. 3, 2022.
- [Hu19] Hunt, T.; Jia, Z.; Miller, V.; Rossbach, C. J.; Witchel, E.: Isolation and Beyond: Challenges for System Security. In: Proceedings of the Workshop on Hot Topics in Operating Systems. S. 96-104, 2019.
- [KF06] Kuipers, S.; Fabro, M.: Control Systems Cyber Security: Defense-in-Depth Strategies. Idaho National Lab. (INL), 2006.
- [NM22] Niemann, K.-H.; Merklin, S.: OT security requirements for Ethernet-APL field devices. atp Magazin, 5/2022, 2022.
- [BOF23] Brockhaus, H.; von Oheimb, S.; Fries, S.: Lightweight Certificate Management Protocol (CMP) Profile. Internet Engineering Task Force (IETF), RFC 9483, 2023.
- [KP12] Kause, T.; Peylo, M.: HTTP Transfer for the Certificate Management Protocol (CMP). Internet Engineering Task Force (IETF), RFC 6712, 2012.
- [TB08] Toorani, M.; Beheshti, A.: LPKI - A Lightweight Public Key Infrastructure for the Mobile Environments. In: 11th IEEE Singapore International Conference on Communication Systems. S. 162-166, 2008.
- [Re24] Registered AAS Submodel Templates, Industrial Digital Twin Association e.V., <https://industrialdigitaltwin.org/content-hub/teilm Modelle>, Stand: 30.04.2024.

Analyse dezentraler Fahrzeugsteuerungsarchitekturen mittels variierender Autonomiegrade an kritischen Punkten von Intralogistiksystemen

Felix Gehlhoff¹, Bjarne Huth¹, Alexander Wenzel²

Abstract: Produktions- und Logistikunternehmen sind gezwungen, ihre Strategien aufgrund zunehmender Produktvielfalt, kürzerer Produktlebenszyklen und eines Mangels an Fachkräften zu überdenken. Herkömmliche Systemstrukturen mit zentralen Steuerungsinstanzen sind zu unflexibel, um schnell und effizient auf dynamische Einflüsse wie kurzfristige Anpassungen von Prozessabläufen oder individuelle Prozessbedingungen zu reagieren. Um diesen Einflüssen entgegenzuwirken, werden zunehmend intelligente und koordinierte Systeme mit spezifischen Kompetenzbereichen und Funktionalitäten in Produktions- und Logistikprozesse integriert. Auf der Grundlage der Analyse von Herausforderungen in der Intralogistik präsentiert diese Arbeit ein Konzept zur Steuerung sogenannter Koordinationsbedarfspunkte. An diesen Punkten werden die Interaktionen zwischen Systemobjekten im internen Materialtransport koordiniert, was beispielhaft in einer Simulationsumgebung implementiert wird. Durch die Variation von Komplexitätseigenschaften wird die Erreichung logistischer Ziele bei verschiedenen Graden der Steuerungsautonomie analysiert. Dabei wurde ein positiver Einfluss der zunehmenden Autonomie und der Einbeziehung auftragsrelevanter Informationen festgestellt.

Keywords: Dezentrale Steuerung, Autonomiegrade, Intralogistik, Simulation

1 Einleitung

Der Logistiksektor und Logistiksysteme haben in den letzten Jahrzehnten erhebliche Veränderungen erfahren. Märkte und Kundenanforderungen sind zunehmend volatil geworden, was alle Formen des Bestandsmanagements kompliziert. Dies, zusammen mit den hohen Kapitalanforderungen für Bestände und Fortschritten in schlanken Methoden (Reduzierung von Beständen und Vermeidung von Verschwendung), hat dazu geführt, dass die Branche auf ein sehr präzises und schlankes Logistiksystem umsteigen musste, das Materialien Just in Sequence liefert. Eine solche komplexe Lieferkette ist anfällig für Ausfälle, wenn Störungen in einzelnen Untersystemen auftreten. Dies hat zu einem Bedarf an dezentraler, agiler und autonomer Materialhandhabung in diesen Untersystemen

¹ Helmut-Schmidt-Universität Hamburg, Institut für Automatisierungstechnik, Holstenhofweg 85, 22043 Hamburg, felix.gehlhoff@hsu-hh.de, <https://orcid.org/0000-0002-8383-5323>

² Leibniz Universität Hannover, Institut für Fabrikanlagen und Logistik, An der Universität 2, 30823 Garbsen, wenzel@ifa.uni-hannover.de, <https://orcid.org/0009-0008-7658-6956>

geführt, um eine insgesamt zuverlässige Intralogistik zu ermöglichen [Sc20]. Die Intralogistik, die Funktionen wie Steuerung, Ausführung und Optimierung interner Material- und Informationsflüsse umfasst [FO21], verzeichnet einen zunehmenden Trend zu höheren Automatisierungsgraden und zur Anwendung moderner Technologien. Mit dem Ziel, die Effizienz interner Prozesse zu verbessern und Kosten zu reduzieren, investieren Unternehmen in automatisierte Lager, intelligente Roboter und andere fortschrittliche Technologien wie AGVs und AMRs, um die wachsenden Dynamiken und Turbulenzen der Marktanforderungen zu bewältigen und eine digitale Transformation der Prozesse zu erreichen [FO21]. Diese Fahrzeuge sind mit den notwendigen Sensoren, Kommunikationsgeräten und Technologien zur Wissensverarbeitung ausgestattet, um komplexe Transportanforderungen sowohl automatisiert als auch möglicherweise autonom zu erfüllen, einschließlich der Koordination mit anderen Fahrzeugen, [RVD20]. Solche intelligenten Subsysteme können lokal operieren und Entscheidungen treffen, wodurch die Gesamtkomplexität des Intralogistiksystems durch die Verteilung von Steuerungskompetenzen auf spezifische Funktionsbereiche reduziert wird [BE16]. Durch die Kombination von Material- und Informationsflüssen ermöglichen sogenannte "aware objects" eine dezentrale Steuerung durch Informationsetiketten, Softwareagenten, verteiltes Wissen und die Fähigkeit, mit anderen Objekten in einem Netzwerk zu kommunizieren. Die intelligenten Systemobjekte sind nicht nur auf Transportobjekte beschränkt, sondern beziehen sich auch auf Umschlagpunkte, Güter und Lastenträger sowie auf Transportwege und Interaktionsbereiche [HO06]. Solche bewussten oder intelligenten Objekte und Entitäten in solchen Systemen können allgemein durch Softwareagenten dargestellt werden, die autonomes und intelligentes Entscheiden sowie flexible Interaktionen mit anderen Agenten und Entitäten im System ermöglichen [RE23].

Die notwendige Koordination umfasst das Aushandeln von Bestellungen und Vorfahrtsrechten, das Bestimmen optimaler Routen durch das System und das Identifizieren von Abhol- und Lieferorten. Die AGVs operieren auf diesen Routen im Transportnetzwerk. Allerdings führt der Einfluss der AGVs aufeinander am Shopfloor zu Situationen, die Interaktion erfordern. Diese Situationen können zu suboptimalem Verhalten und potenziell sogar zu Blockaden führen [HE22].

Traditionell wird die Kollisionsvermeidung durch zentrale Direktiven eines Steuerungssystems verwaltet [HE22]. Eine zentrale Kollisionsvermeidungssteuerung erfasst Informationen über die aktuellen Positionen, Ziele und geplanten Routen aller Fahrzeuge innerhalb eines Systems, um Kollisionspotenziale kreuzender Verkehrswege zu identifizieren [LK12]. Wenn Kollisionspotenziale auftreten, wird entweder die gesamte Route der AGVs neu definiert, der Fahrweg eines Fahrzeugs geändert oder die Fahreigenschaften der Fahrzeuge, wie ihre Geschwindigkeit, angepasst. Dieser Optimierungsprozess wird so lange wiederholt, bis keine Kollisionspotenziale mehr auftreten. Aufgrund der dynamischen Positionsdaten der Fahrzeuge ist ein regelmäßiger Steuerungsprozess erforderlich und entsprechend zeitaufwändig. Herausforderungen bei der zentralen Kollisionsvermeidung treten auf, wenn eine große Anzahl von Fahrzeugen beteiligt ist, was die zu verarbeitende Informationsmenge und damit die Berechnungszeit

erhöht [HE22]. Zusätzlich werden nur statische und dynamische Objekte sowie Hindernisse berücksichtigt, die bereits in der virtuellen Umgebung erfasst wurden [RVD20].

Eine autonome Kollisionsvermeidung durch die Fahrzeuge, gängig für AMRs, ist ebenfalls möglich. Dies beinhaltet eine situationspezifische Reaktion der Fahrzeuge auf die lokale Situation [SC20]. Mögliche Reaktionen der Fahrzeuge umfassen entweder eine Geschwindigkeitsreduktion mit einem endgültigen Stopp in definierter Entfernung zum nächsten Fahrzeug oder Ausweichmanöver und Routenanpassungen. Obwohl diese Methode im Vergleich zur zentralen Kollisionsvermeidung deutlich schnellere Reaktionszeiten aufweist [RVD20], ist sie hinsichtlich der gewünschten Priorisierung bestimmter Auftragsarten oder Fahrzeuge unflexibel und garantiert nicht frei von Deadlocks.

Daher besteht der Bedarf für einen Ansatz zur situativen und dezentralen Koordination von Stellen im Transportnetzwerk mit erhöhtem Koordinationsbedarf, im Folgenden als Koordinationsbedarfsstellen (KBS) bezeichnet. Die Steuerung von KBS, ähnlich den Zonensteuerungsansätzen [RVD20], kann in der Struktur variieren und hängt vom Umfang der Entscheidungskompetenzen der Steuerungsinstanzen ab, wobei unterschiedliche Autonomiegrade gezeigt werden. In diesem Papier bezieht sich die Struktur der Steuerung auf den Unterschied in der Entscheidungsfindungslotation und dem Autonomiegrad, während eine Steuerungsstrategie das angewendete Regelwerk zur Ausübung der Kontrolle in einer gegebenen Struktur beschreibt. Somit ist eine autonome, fahrzeuginterne Kollisionsvermeidung per Definition autonomer als ein dezentraler, zonenbasierter Ansatz, da sie nicht erfordert, dass sich die Fahrzeuge einer Hierarchie unterordnen. Allerdings kann selbst eine gegebene Struktur (z.B. dezentral) je nach den Entscheidungen, die noch von den Fahrzeugen getroffen werden, einen unterschiedlichen Grad an Fahrzeugautonomie aufweisen. Die Gestaltung der KBS und die Steuerung von KBS beeinflussen die Komplexität innerhalb des Systems und den notwendigen Informationsaustausch.

Basierend auf einer Analyse der Herausforderungen der modernen Intralogistik mit Fokus auf interne Transportsysteme in den Abschnitten 1 und 2, präsentiert diese Arbeit Eigenschaften der dezentralen KBS-Koordinationssteuerung (Abschnitt 3) und ein Simulationskonzept, das KBS unter verschiedenen Komplexitätseigenschaften berücksichtigt sowie die Ergebnisse der durchgeführten Experimente (Abschnitt 4). Der Beitrag schließt mit einem Fazit und Ausblick in Abschnitt 5.

2 Stand der Wissenschaft

Leitao et al. setzen variable Autonomie im Falle von Störungen ein, indem sie den beteiligten Entitäten eine höhere Autonomie gewähren, wenn Störungen auftreten [LR05]. Dies ermöglicht es den verschiedenen Entitäten, lokal zu operieren, ohne strikt zentralen Direktiven zu folgen, wodurch sie die Möglichkeit erhalten, Ressourcenzuweisungen neu

zu verhandeln oder dynamische Layoutänderungen vorzunehmen. Sobald die Störungen behoben sind, wird das Autonomielevel wieder reduziert [BA15]. Die Autoren zeigen, dass variable Autonomie im Vergleich zu strikt zentralisierten oder dezentralisierten Ansätzen die Durchlaufzeit, die Zykluszeit und die Stabilität verbessert, wenn Systemstörungen auftreten [LR05].

Gronau et al. führen einen Index ein, um den Grad der Autonomie in einem Wertstromproduktionssystem zu messen, indem sie den Anteil autonomer Interaktionen an den Gesamtinteraktionen vergleichen [GT16]. Dieses Konzept basiert auf dem Lean-Index aus der Wertstromanalyse, der die Wertschöpfung eines Prozesses durch den Anteil der nützlichen Zeit an der Gesamtprozesszeit misst. Je näher dieser Wert bei 1 liegt, desto effizienter ist die Zeitnutzung, wodurch unproduktive Zeit durch Verzögerungen minimiert wird.

Windt et al. postulieren, dass die Problemkomplexität mit dem angemessenen Grad der Selbstregulierung oder Dezentralisierung korreliert [W10]. Die Autoren legen dar, dass geringe Problemkomplexität bei hoher Autonomie oder Dezentralisierung zu schlechten Ergebnissen führt. Steigen jedoch Autonomie und Problemkomplexität, verbessert sich die Leistung zunächst, erreicht einen Höhepunkt und fällt dann wieder ab. Diese Annahmen werden durch verschiedene Simulationsstudien getestet, die auch die Rechenzeiten untersuchen und zeigen, dass die Anzahl der erforderlichen Berechnungen in zentralisierten Systemen exponentiell ansteigt, während sie in dezentralisierten Systemen nur linear zunimmt. Das Kommunikationsverhalten und die Rechenkapazitäten zentraler und dezentraler Einheiten werden in diesen Befunden nicht berücksichtigt.

Obwohl diese Ansätze erste Einblicke in das optimale Niveau der Autonomie bieten, ist der Vergleich verschiedener Autonomiegrade aufgrund des unterschiedlichen Designs und der Ergebnisse der Ansätze etwas eingeschränkt. Keine dieser Studien hat die vergleichende Systemleistung für dezentrale, aber hierarchische Systeme im Vergleich zum autonomen Verhalten der Fahrzeuge analysiert. Um eine solche Analyse zu ermöglichen, bietet der folgende Abschnitt eine detaillierte Analyse der Eigenschaften von KBS. Dieser Rahmen kann in Zukunft von beliebigen Algorithmen genutzt werden, um die Implementierung und den Vergleich verschiedener Algorithmen mit unterschiedlichen Graden der Dezentralisierung und Autonomie zu ermöglichen.

3 Eigenschaften von Koordinationsbedarfsstellen

KBS in modernen Intralogistiksystemen, die AGVs und AMRs verwenden, erfordern aufgrund geteilter Transportwege einen erhöhten Koordinationsaufwand. Diese KBS treten auf, wo Fahrzeuginteraktionen und Interaktionen mit anderen Systemobjekten stattfinden. Jede KBS erfordert spezifische Entscheidungen, um Sicherheit zu gewährleisten und den Verteilungsprozess effizient zu gestalten. Besonders kritisch sind Kreuzungen, wo mehrere Transportwege zusammentreffen, und eine erhöhte

Kommunikations- und Steuerungsanforderung besteht. Der Erfolg der lokalen Koordination zeigt sich in kollisionsfreien Durchgängen, termintreuer Lieferung, also des Abschluss des Auftrags vor Ende des avisierten Liefertermins, und minimalen Wartezeiten. Wartezeiten beeinflussen nicht nur die Termintreue der Transportaufträge, sondern auch die Energieeffizienz der Fahrzeuge negativ.

Die detaillierte Analyse der oben genannten KBS, fokussiert auf Kreuzungen in Intralogistiksystemen, erfordert eine umfassende Untersuchung der Eigenschaften solcher Punkte. Der Bereich, in dem Materialflüsse zusammentreffen, kann als eine Art Ressource mit begrenzter physischer Kapazität interpretiert werden. Um den Interaktionsbereich und den Zugriff auf einer KBS effektiv zu gestalten, wird ein Rahmen zur Erfassung von Konfliktstellen innerhalb der Interaktionsfläche eines Kreuzungsbereichs bereitgestellt. Anschließend wird eine Auswahl an Steuerungsmöglichkeiten zur Bestimmung der Reihenfolge der Durchgänge vorgestellt. Zusätzlich wird die Auswirkung der Koordinationsentscheidungen auf das individuelle Fahrverhalten dargestellt und die Unterschiede in der Steuerungsstruktur werden präsentiert.

Die Koordination einer KBS kann durch einen iterativen Prozess implementiert werden:

- Identifizierung von Konflikten zwischen den AGVs
- Entscheidung der Reihenfolge basierend auf einer Steuerungsstrategie
- Anpassung des Fahrverhaltens des untergeordneten AGVs
- Neubewertung der Routenkonflikte

Dieser Koordinationsprozess wird wiederholt, bis keine zeitlichen Konflikte in der Nutzung der Kreuzung mehr bestehen. Abhängig vom Design der KBS und der Anzahl der aktiven Eingänge entstehen Konflikte der Transportwege. Fahrtwegkonflikte oder Zielkonflikte bezeichnen die lokalen kritischen Interaktionen der Fahrtwege innerhalb von KBS. Sie werden in drei Arten eingeteilt und bringen das Verhalten von zwei befahrenen Wegen in Verhältnis:

- Divergierende Wege: Die Ausgänge an einer KBS stehen in keinem Konflikt; die Fahrtwege kreuzen sich nicht.
- Kreuzende Wege: Die Ausgänge an einer KBS stehen in keinem Konflikt; die Fahrtwege kreuzen sich.
- Identische Wege: Die Ausgänge an einer KBS stehen in einem Konflikt (identisch); die Fahrtwege kreuzen sich nicht.

Die Häufigkeit und Konfiguration von Zielwegkonflikten variieren je nach Ziel des Fahrzeugs innerhalb des Interaktionsbereichs, der Anzahl an Interaktionspartnern und der Auslegung der KBS. Wie in Abbildung 1 dargestellt, gibt es im Fall einer

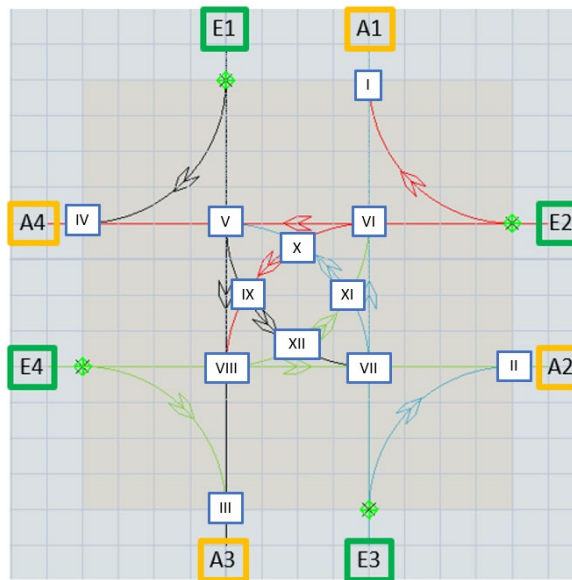


Abb. 1: Interaktionsbereich mit Entscheidungspunkten I bis XIII an einer omnidirektionalen Kreuzung.

omnidirektionalen Kreuzung jeweils vier Eingänge (E) und Ausgänge (A). Dies führt zu vielen potenziellen Teilnehmern, Entscheidungspunkten (EP) I bis XII und Zielwegkonflikten.

Aufgrund der zwölf EP innerhalb des Interaktionsbereichs ergeben sich folgende allgemeine Zeitbeziehungen zum Erreichen der jeweiligen EP je Fahrtweg. Diese Zusammenhänge erfolgen unter der Annahme eines quadratischen Interaktionsbereichs mit der Seitenlänge d und der Fahrtgeschwindigkeit v_i pro Fahrzeug i . Ein Abbiegevorgang nach links (LT) beansprucht fünf EP mit folgenden Zeiten. Diese sind in den Gleichungen (EQ) in Tab. 1 in der Reihenfolge der Durchfahrt des Fahrzeugs dargestellt und am Beispiel des Abbiegemanövers von E1 nach A2 und unter der Annahme eines Viertelkreisbogens für den Fahrtwegwechsel exemplarisch dargestellt. Ein Beibehalten des Fahrtweges auf einer geraden Durchfahrt von E1 nach A3 erfordert das Beanspruchen von drei EP auf dem Fahrtweg (siehe auch Tab. 1).

Gleichung	Entscheidungspunkt	Manöver	Zeitbeziehung [s]
1	V	LT	$t_{i,V} = \frac{1}{v_i} \cdot \left(\frac{d}{3}\right)$
2	IX	LT	$t_{i,IX} = \frac{1}{v_i} \cdot \left(\frac{d}{3} + \frac{1}{18} \cdot d \cdot \pi\right)$
3	XII	LT	$t_{i,XII} = \frac{1}{v_i} \cdot \left(\frac{d}{3} + \frac{1}{9} \cdot d \cdot \pi\right)$

4	VII	LT	$t_{i,VII} = \frac{1}{v_i} \cdot \left(\frac{d}{3} + \frac{1}{6} \cdot d \cdot \pi \right)$
5	II	LT	$t_{i,II} = \frac{1}{v_i} \cdot \left(\frac{2 \cdot d}{3} + \frac{1}{6} \cdot d \cdot \pi \right)$
6	V	SP	$t_{i,V} = \frac{1}{v_i} \cdot \left(\frac{d}{3} \right)$
7	VIII	SP	$t_{i,VIII} = \frac{1}{v_i} \cdot \left(\frac{2 \cdot d}{3} \right)$
8	III	SP	$t_{i,III} = \frac{d}{v_i}$

Tab. 1: Beziehungen der Zeit innerhalb des Kreuzungsbereichs

Für alle Fahrzeuge, die den Interaktionsbereich befahren möchten, müssen mehrere Vergleiche durchgeführt werden. Dies erfolgt anhand der Beziehungen aus der Distanz und der Anzahl an potenziellen Zielkonflikten auf dem Weg.

Um Koordinationsbedürfnisse anzusprechen und eine Sequenz in der Nutzung von EP zu erreichen, ist die Kontrolle der Reihenfolge, in der die EP passiert werden, notwendig. Dieser Unterabschnitt präsentiert eine Auswahl an Koordinationssteuerungsstrategien und diskutiert die damit verbundenen Anforderungen an Informationsdaten.

- Die "Fahrzeugintegrierte Kollisionsvermeidung" nutzt Sensoren zur Umgebungserkennung und ermöglicht dem Fahrzeug, autonom auf mögliche Konflikte zu reagieren, wobei auch eine Routenneuberechnung zur Reduzierung von Stop-Zeiten in Betracht gezogen wird.
- Die "First-in-First-out (FiFo)"-Strategie setzt auf zeitbasierte Kriterien, um die Vorfahrtsreihenfolge zu bestimmen, und erfordert Echtzeitinformationen zu den Fahrzeugen, wobei Schwankungen in der Geschwindigkeit berücksichtigt werden müssen.
- Die "Priorisierung auf Grundlage von Auftragsinformationen" integriert individuelle Transportauftragsdetails, um Dringlichkeiten zu berücksichtigen und Lieferzeiten zu optimieren, wobei eine kontinuierliche Bewertung des Fortschritts erforderlich ist.

Diese Steuerungsmethoden bieten eine Reihe von potenziellen Ansätzen, um eine effektive Kollisionsvermeidung zu gewährleisten und sichere sowie termintreue Transportprozesse zu gewährleisten. Zudem kann das Fahrverhalten der Fahrzeuge angepasst werden. Um eine kollisionsfreie Anpassung der Fahrwege zu gewährleisten, bedarf diese Anpassung der Berücksichtigung der Fahrzeuglänge, der Geschwindigkeit sowie eines einzuhaltenen Sicherheitsabstandes.

4 Simulation von Koordinationsbedarfsstellen

Die in den vorherigen Abschnitten beschriebene KBS der omnidirektionalen Kreuzung wurde in eine Simulationsumgebung implementiert. Das Simulationsmodell konzentriert sich auf die Koordinationsbedürfnisse an den Schnittstellen des Materialflusses, die Teil eines intralogistischen Transportnetzwerks sind. Ziel ist es, dezentrale und autonome Steuerungsstrukturen in das Simulationsmodell zu integrieren. Dadurch ermöglicht die Simulation die Analyse verschiedener Grade der Fahrzeugautonomie auf die Gesamtleistung der Logistik. Zu beachten ist, dass eine umfassende Analyse möglicher Steuerungsstrategien (d.h. weitere Regelwerke der dezentralen Steuerung) nicht im Rahmen dieses Beitrags liegt.

Simulationskomplexitätsparameter: Die Simulation sollte unterschiedliche Komplexitätsgrade ermöglichen, basierend auf der Anzahl der Fahrzeuge sowie der Anzahl möglicher Startpunkte und Ziele. Das Modell liefert Daten zu logistischen Kennzahlen innerhalb eines festgelegten Zeitrahmens mit festgelegten Start- und Endzeitpunkten, die einen quantitativen Vergleich ermöglichen. Kennzahlen für den quantitativen Vergleich der Leistung der Steuerungsstrategie unter verschiedenen Komplexitätsszenarien umfassen die Wartezeit vor dem Interaktionsbereich, die Abweichung von der Termintreue der Transportaufträge, die Durchlaufzeit des Interaktionsbereichs und die Auslastung des Transportsystems, d.h. die Anzahl der Fahrzeuge im System. Der Prioritätsindex gibt den Anteil hoch priorisierter Aufträge an. Die folgende Tabelle gibt einen Überblick über die simulierten Szenarien..

Szenario	Anzahl der Transportziele	Anzahl der Lager (Start)	Flotten-größe	Anzahl der Aufträge	Prioritäts-index [%]
1	Niedrig	Niedrig	Niedrig	80	58,8
2	Hoch	Niedrig	Niedrig	80	58,8
3	Niedrig	Hoch	Niedrig	160	56,9
4	Niedrig	Niedrig	Hoch	80	58,8
5	Niedrig	Hoch	Hoch	160	56,9
6	Hoch	Hoch	Niedrig	160	56,9
7	Hoch	Niedrig	Hoch	80	56,9
8	Hoch	Hoch	Hoch	160	56,9

Tab. 2: Szenarienbeschreibung

Die Simulationsergebnisse sind in Abb. 2 zusammengefasst. Die Ergebnisse zeigen, dass im implementierten Simulationsmodell für die Szenarien eins bis vier und sechs bis sieben mit den Steuerungsmethoden der Auftragspriorität und der autonomen Kollisionsvermeidung eine hohe Termintreue erreicht wird. In den Szenarien fünf und acht führt jedoch der Einfluss der erhöhten Flottengröße und der aktiven Transportaufträge, die an allen Eingangspunkten eintreffen, zu einer deutlichen Reduzierung der Termintreue bei

den dezentralen FiFo- und prioritätsbasierten Steuerungsmethoden. Die hohen Wartezeiten aufgrund der Kapazitätsbeschränkung des Interaktionsbereichs führen zu einer Verschlechterung der Termintreue, da Aufträge mit hoher Priorität hinter Aufträgen mit niedriger Priorität warten müssen, bis diese eine Freigabe erhalten. Fahrzeuge mit aktiven Aufträgen können den Interaktionsbereich nicht befahren, was zu Lieferverzögerungen führt und die Einhaltung von Zeitplänen verhindert. Darüber hinaus wirken sich zusätzliche auftragsspezifische Daten positiv auf den Koordinationsprozess aus. Unter identischen Kapazitätsbeschränkungen kann die prioritätsbasierte Steuerung in den Szenarien vier und sieben für beide Auftragsarten eine Termintreue von über 90 % erreichen. Ohne zusätzliche auftragsspezifische Daten fällt die Termintreue bei einer rein auf der Ankunftszeit basierenden Steuerung auf unter 70 % bzw. sogar unter 50 %.

Auf der einfachsten Komplexitätsebene (Szenario 1) gibt es nur geringe Unterschiede in den Ergebnissen zwischen den Steuerungsmethoden aufgrund weniger Kollisionspotenziale und Konflikten. In Szenario 3 werden viele hochpriorisierte Aufträge bereitgestellt, was dazu führt, dass die prioritätsbasierte Methode häufig Vorfahrtsentscheidungen nach dem FiFo-Prinzip in Konfliktsituationen trifft, was die reduzierte Termintreue von unter 90 % für hochpriorisierte Transportaufträge erklärt. Mit der autonomen Kollisionsvermeidung passen die Fahrzeuge ihr Verhalten an, verhindern Kollisionen bei nachfolgenden Transporten, optimieren den Fluss zwischen den Stationen und eliminieren den Koordinationsbedarf, was zu einer hohen Termintreue führt. Eine erhöhte Komplexität verschlechtert die Termintreue über alle Steuerungsmethoden hinweg. Ein erhöhtes Verkehrsaufkommen im Interaktionsbereich führt zu häufigen Entscheidungsfindungen und anschließenden Verhaltensanpassungen. Bei autonomer Kollisionsvermeidung führt dies dazu, dass Fahrzeuge anhalten, um Kollisionen zu verhindern, was zu individuellen Brems- und Beschleunigungsvorgängen innerhalb des Interaktionsbereichs führt und somit die Lieferzeiten verlängert. Hohe Auftragsvolumen führen oft zu FiFo-basierten Entscheidungen in den Szenarien 3, 5, 6 und 8, wodurch hochpriorisierte Aufträge mehrere Durchfahrten auf eine Freigabe warten müssen und sich die Termintreue der Aufträge verschlechtert. Mit zunehmender Anzahl potenzieller Zielorte in den Szenarien 2 sowie 6 bis 8 verschlechtert sich die Transportleistung bei einer autonomen Kollisionsvermeidungsstrategie aufgrund des erhöhten Verkehrsaufkommens. Je nach Abbiegemanöver werden vermehrt gemeinsame Entscheidungsstellen unterschiedlicher Fahrwege beansprucht, wodurch der Koordinationsbedarf innerhalb der Interaktionszone während der Durchquerung steigt.

Der Vergleich von Aufträgen mit hoher Priorität mit Aufträgen niedriger Priorität (dargestellt in Abb. 2 links gegenüber rechts) zeigt, dass die Ergebnisse für FiFo und autonome Steuerung sehr ähnlich sind. Nur die prioritätsbasierte Strategie in einer dezentralen Struktur erreicht eine höhere Termintreue bei Aufträgen mit hoher Priorität (z.B. sichtbar in den Szenarien 5 und 8).

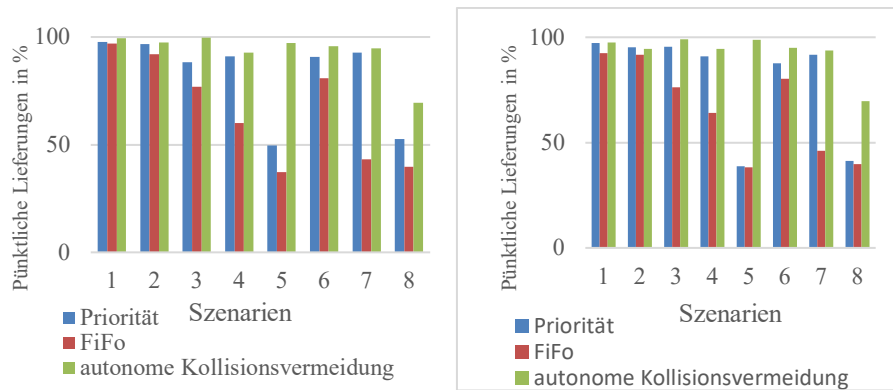


Abb. 2: Ergebnisse der Simulation, Eilaufträge (links) und Standardaufträge (rechts)

5 Zusammenfassung und Ausblick

Die Gestaltung von Steuerungsstrukturen an Materialflusskreuzungen oder von KBS als dezentrale hierarchische Systeme oder durch autonome Steuerung durch die Fahrzeuge selbst bietet vielversprechende Methoden zur Bewältigung von Konfliktsituationen in komplexen Transportumgebungen. Anpassungen des individuellen Fahrverhaltens beeinflussen nur die unmittelbar beteiligten Interaktionspartner, wodurch der Aufwand der KBS-Steuerung auf lokale Konfliktsituationen reduziert wird. Erste Simulationsergebnisse verdeutlichen, dass steigende Autonomie die Leistungsfähigkeit des Systems erhöht und auftragspezifische Daten sich positiv auf den Koordinationsprozess in Bezug auf die erreichte Termintreue auswirken. Ein umfassend vernetztes Transportsystem aus dezentralen KBS und frei navigierenden Fahrzeugen, integriert mit auftragspezifischen Informationen, kann einen zeiteffizienten, dezentral koordinierten Materialfluss ermöglichen. Basierend auf den Ergebnissen dieser Studie sollten weitere Untersuchungen zu dezentralen und autonomen AGVs durchgeführt werden, die sich nicht nur auf Anwendungen in der Intralogistik beschränken. Beispielsweise wurden in den Experimenten nur einfache regelbasierte Mechanismen für die dezentralen Steuerungsstrategien implementiert. Zukünftige Arbeiten werden elaboriertere Strategien implementieren, die die Leistungsfähigkeit intelligenter Softwareagenten nutzen.

Literatur

- [LK12] C. Liu, A. Kroll: *A Centralized Multi-Robot Task Allocation for Industrial Plant Inspection by Using A* and Genetic Algorithms*. In: Hutchison, Kanade et al. (Hrsg.): *Artificial Intelligence and Soft Computing*. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science), Vol. 7268, S. 466–474, 2012.

-
- [BE16] T. Beyer, P. Gohner, R. Yousefifar, K.-H. Wehking: *Agent-based dimensioning to support the planning of Intra-Logistics systems*. In: *2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*: IEEE, S. 1–4, 2016.
- [W110] K. Windt, T. Philipp, F. Böse, T. Becker: *Application of a three-component evaluation system for autonomous control in logistics*. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, Vol. 224 (8), 2010, S. 1267–1276.
- [RVD20] M. de Ryck, M. Versteyhe, F. Debrouwere: *Automated guided vehicle systems, state-of-the-art control algorithms and techniques*. Journal of Manufacturing Systems, Vol. 54, 2020, S. 152–173.
- [Fo21] J. Fottner, D. Clauer, F. Hormes, M. Freitag, T. Beinke, L. Overmeyer, S.N. Gottwald, R. Elbert, T. Sarnow, T. Schmidt, K.-B. Reith, H. Zadek, F. Thomas: *Autonomous Systems in Intralogistics: State of the Art and Future Research Challenges*. Logistics Research, Vol. 14 (2), 2021.
- [Ho06] M. ten Hompel: *Das Internet der Dinge – Potenziale autonomer Objekte und selbstorganisierender Systeme in der Intralogistik*. In: Arnold (Hrsg.): *Intralogistik: Potentiale, Perspektiven, Prognosen*. 1. Auflage. Dordrecht: Springer Netherland (VDI-Buch), 2006.
- [GT16] N. Gronau, H. Theuer: *Determination of the Optimal Degree of Autonomy in a Cyber-Physical Production System*. Procedia CIRP, Vol. 57, 2016, S. 110–115.
- [BA15] J. Barbosa, P. Leitão, E. Adam, D. Trentesaux: *Dynamic self-organization in holonic multi-agent manufacturing systems: The ADACOR evolution*. Computers in Industry, Vol. 66, 2015, S. 99–111.
- [LR05] P. Leitão, F. Restivo: *Experimental Validation of ADACOR Holonic Control System*. In: Mařík, Brennan (Hrsg.): *Proceedings of HoloMAS*. Berlin, Deutschland: Springer (Lecture notes in computer science Lecture notes in artificial intelligence, 3593), Vol. 3593, S. 121–132, 2005.
- [Sc20] T. Schmidt, K.-B. Reith, N. Klein, M. Däumler: *Research on Decentralized Control Strategies for Automated Vehicle-based In-house Transport Systems – a Survey*. Logistics Research, Vol. 13, 2020.
- [RE23] L.M. Reinpold, L.P. Wagner, F. Gehlhoff, M. Ramonat, M. Kiltbau, M.S. Gill, J.T. Reif, V. Henkel, L. Scholz, A. Fay: *Systematic Comparison of Software Agents and Digital Twins: Differences, Similarities, and Synergies in Industrial Production*, 2023.
- [HE22] B. Hein, M. Wesselhoft, A. Kirchheim, J. Hinckeldeyn: *Towards Industry-Inspired Use-Cases for Path Finding in Robotic Mobile Fulfillment Systems*. In: *2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA)*: IEEE, S. 1–4, 2022.

„Ulrich Jumar, Christian Diedrich (Hrsg.):
EKA 2024 - Entwurf komplexer Automatisierungssysteme, 18. Fachtagung“

Simplified Model for Data Collection of Liquid Products in Process Manufacturing

Wei Guo¹, Wan Li², Swantje Thiele³, Dominik Orzol⁴, Thomas Bewer⁵ and Tobias Kleinert⁶

Abstract: In the realm of process manufacturing, tracking product flow and data alignment remains important. Current technologies primarily cater to discrete and batch processing, leaving a gap in continuous process tracking. This paper introduces a simplified model designed to calculate the transit time slots as products traverse specific system passages, considering systems characterized by pipeline branching and mixing tanks. The model presented herein is based on certain idealized conditions and has not yet been empirically validated. Nevertheless, it provides a method for the calculation and allocation of comprehensive data histories of arbitrarily defined product units within continuous process manufacturing systems. It holds potential for future extension and refinement.

Keywords: data allocation, product tracking, process manufacturing, CSTR

1 Introduction

To optimize manufacturing processes across multiple process steps, the collection and integration of the data generated in each process step with respect to the product units is necessary. A product unit at the end of a production line is a result of all the process steps that it went through. Its quality characteristics are often related to multiple process steps whose parameters and behaviors are inter-related. Therefore, analyzing and optimizing product quality as well as production processes require product-oriented data integration across multiple process steps.

Existing methods enable product-oriented data integration through assigning product units with unique IDs at early process steps [Di19]. Such methods mainly focus on discrete manufacturing or batch manufacturing. In contrast, product-oriented data allocation in continuous manufacturing is less discussed, especially when handling e.g., liquid or

¹ RWTH Aachen University, Chair of Information and Automation Systems for Process and Material Technology, Turmstr. 46, 52064 Aachen, w.guo@iat.rwth-aachen.de


² RWTH Aachen University, Chair of Information and Automation Systems for Process and Material Technology, Turmstr. 46, 52064 Aachen, w.li@iat.rwth-aachen.de

³ IPGR e.V., Nerscheider Weg 170, 52076 Aachen, swantje.thiele@ipgr.com

⁴ IPGR e.V., Nerscheider Weg 170, 52076 Aachen, dominik.orzol@ipgr.com

⁵ Bucher Emhart Glass, Hinterbergstrasse 22, 6312 Steinhausen, Switzerland, thomas.bewer@emhartglass.com

⁶ RWTH Aachen University, Chair of Information and Automation Systems for Process and Material

Technology, Turmstr. 46, 52064 Aachen, kleinert@iat.rwth-aachen.de, 
<https://orcid.org/0000-0001-7441-4431>

gaseous materials. Hereby, one of the challenges is to track back the past positions of a product unit along the material flows in the production plant [TMK21].

Continuous Stirred Tank Reactors (CSTR) are widely used in continuous manufacturing as well as for representing the behavior of sections in continuous process which provide a product holdup. The residence time distribution (RTD) is often used to characterize the mixing dynamics inside the CSTR [RK08] and to trace disturbances introduced during operation [Kr18]. The past (i.e., historical) position of a product unit along the process route is not yet well studied. However, e.g., in the food and beverage industry or glass container manufacturing, data is collected and used for process monitoring and optimization [SKS23], supporting production planning and performance analysis [CV20] or inspecting defects [Li19], without considering the historical position of a given product unit in the production lines. This, however, becomes necessary, if historical time intervals of sensor and actuator signals along the production route of a plant shall be allocated to a specified product unit at the plant outlet, e.g., for reasons of product data collection or quality investigations.

In the tracking of discrete products, each product unit is typically conceptualized as a point. Its position can be calculated and tracked by measuring the distance between the product unit and some reference points (e.g. using tags and anchors [Sl19]). This is feasible, since a certain type of product usually has a regular size. Methods for tracking disturbances in continuous manufacturing [Kr18], however, cannot be used to calculate the transit time of a defined product unit passing by the passages (at certain places) in the whole process. Therefore, for comprehensive allocation of information pertaining to product fractions, a flexible method is required for localizing a variably definable quantity of product units.

To allocate the existing data more precisely to a given product unit, it is necessary to provide more accurate information about the relation of position and transit time of the considered product unit within the production process. This work introduces a simplified model that enables locating the start and the end of any liquid product units in continuous production processes, such that the relevant machine data time interval can be allocated to the corresponding product unit as the relevant product data generated during the product's production.

This article is structured as follows: Section 2 introduces on the important assumptions and required information for the calculation, providing the basics of the model used for data allocation in process manufacturing. Section 3 delves into the theoretical foundations of the model and the handling of CSTR within this paper.

Section 4 presents the methodologies for data allocation involving pipeline branching and the handling of CSTRs, showcasing the application and flexibility of the model. Section 5 discusses the limitations and challenges of the current model, proposing areas for future improvement and research.

2 Assumptions and required information

This section introduces and explains the assumptions underlying the model used in this work, which are applicable under ideal conditions. The scope of the paper focuses on developing a method to track and analyze the positions of liquid product units within pipes and tanks. Specifically, the paper aims to calculate transit time slots for a given product unit based on sensor or actuator position, sensor or actuator time-series data, and process events. This includes the calculation of the transit time slots for when a product unit passes a specific point in the system, such as a sensor, actuator, pipe branching etc. This paper also endeavors to address the complexities involved in product tracking and data allocation in continuous processes with branched pathways.

2.1 Assumptions

To facilitate the computation of historical product positions, the model designed in this work assumes that the behavior of the liquid product conforms to a plug flow regime. This assumption simplifies computational complexity by treating the flow of the liquid through the system as uniformly moving at the same velocity at a given point in the plant, without back mixing or variations in velocity across different sections of the conduit. The other complex fluid behavior is not considered in this work, such as the fluid behavior at connections or junctions.

In the production plant, the cross-sectional area at a given position is utilized to represent the smallest unit of the product. The midpoint of this cross-section serves as the computational location for determining the product's position within the system.

The direction of liquid flow is considered the positive direction for product positioning, aligning the movement of the product within the plant's flow dynamics. The trajectory of the product movement is represented by a line connecting the midpoints of the process equipment through which the product flows.

The model also assumes that the process equipment within the system is always fully saturated with the product, ensuring that there are no partial fills that might alter the flow dynamics. Additionally, it neglects the impact of height variations in the piping on the flow rate, assuming that gravitational effects and potential energy changes are minimal.

These assumptions streamline the calculation model by eliminating the need to account for complex fluid dynamics associated with varying fill levels and elevation changes, thereby focusing solely on the flow behavior under consistent and controlled conditions.

2.2 Required information

Besides the assumptions, some data are required to calculate the product position and allocate the data to the selected product unit. These parameters should be either readily

accessible or ascertainable through computation or measurement. They are categorized into three groups.

Static Data:

Geometry of Process Equipment: Knowing the exact dimensions and configurations is crucial. This includes diameters, lengths, curvature, and the layout of piping systems, as well as the shape and volume capacities of tanks.

Technical Specifications of Sensors and Actuators and Their Connections: Detailed information about the sensors and actuators involved in the system, such as their type, range, accuracy, and location, is necessary. This ensures that data collected is relevant and precise for the system's needs.

Dynamic Data:

Events During Processes: This includes changes in the flow paths, such as valve operations (open/close), shifts in direction of flow, and other process adjustments that impact the movement and position of the product.

Product Properties:

Flow Rates (Volume/Mass): Real-time data on how much product is moving through piping the system.

Density: Measurements or calculations of the product's density, which can affect flow dynamics and require adjustments in process handling.

3 Simplified model for calculating the position of liquid

To compute the transit time slot corresponding to a product unit passing through a predetermined location, it is necessary to determine the temporal markers denoting the beginning and the end of the observed product unit. Consequently, one must calculate the moment when a cross-section of the product traverses the location. This involves establishing a relationship between the known flow properties, the geometric configuration of the process system, and the temporal progression of the product within it.

3.1 Method to calculate the position of a product's cross-section

The relationship described above involves two cross-sections within a process system, specifically between an outlet at position x_a and another given position x_R (see Fig. 1 (a)).

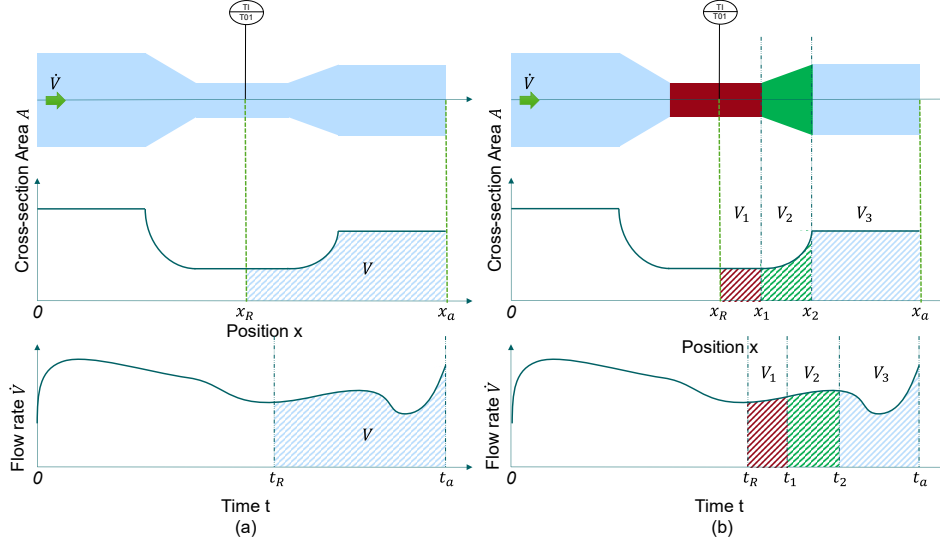


Fig. 1: (a) Theoretical foundation of calculating product position. (b) Calculation in segments.

The volume between the two cross-sections from x_R to x_a can be mathematically expressed as the integral of the cross-sectional area $A(x)$ with the positional difference between x_R and x_a . This is mathematically noted as: $V = \int_{x_R}^{x_a} A(x) \cdot dx$, where V represents the volume between the two cross-sections, $A(x)$ is the cross-sectional area, and dx is the infinitesimal segment of the path between x_R and x_a .

Furthermore, this volume can also be equated to the integral of the product flow rate $\dot{V}(t)$ over the time difference between t_R and t_a , which corresponds to the times when the product passes through x_R and x_a respectively: $V = \int_{t_R}^{t_a} \dot{V}(t) \cdot dt$, where $\dot{V}(t)$ is the time-variable flow rate of the product at the cross-section, t_R and t_a are the respective times at which the product is at positions x_R and x_a ⁷. We can deduce the following Equation (1):

$$\int_{x_R}^{x_a} A(x) \cdot dx = \int_{t_R}^{t_a} \dot{V}(t) \cdot dt \quad (1)$$

Based on the previously stated assumptions, the sole unknown variable t_R within the Equation (1) can be resolved mathematically.

In practical computation, it is not always feasible to calculate the volume of each process equipment by simple integration due to the complexity of certain geometries. Hence, the

⁷ Here is constant density and incompressible fluid considered.

target region may be subdivided into several contiguous segments, (see Fig. 1 (b)), with the volume of each segment calculated individually.

Based on the segmentation $V = V_1 + V_2 + V_3$ in Fig. 1 (b), the Equation (2) can be derived:

$$\int_{x_R}^{x_a} A(x) \cdot dx = \int_{x_R}^{x_1} A(x) \cdot dx + \int_{x_1}^{x_2} A(x) \cdot dx + \int_{x_2}^{x_a} A(x) \cdot dx \quad (2)$$

Equation (1), which provides the theoretical foundation for the model discussed in this work, does not mandate the use of integration for practical computations. This is practical for process equipment that is resistant to straightforward volumetric integration, wherein the entire part's volume and its length along the product path may be employed directly in the analysis. When the target location does not reside within a process equipment of complex geometry, it is advisable to employ the holistic approach for the computation of its volume. For process equipment with simple geometric shapes, such as straight pipelines, it is preferable to employ proportional calculations rather than integration methods when determining the required volume inside the equipment.

The following relationship can be derived by combining Equation (1) and Equation (2):

$$\int_{t_R}^{t_a} \dot{V}(t) \cdot dt = \int_{t_R}^{t_1} \dot{V}(t) \cdot dt + \int_{t_1}^{t_2} \dot{V}(t) \cdot dt + \int_{t_2}^{t_a} \dot{V}(t) \cdot dt \quad (3)$$

The variable t_R within the Equation (3) can be resolved step by step in a backward sequence:

$$t_R \leftarrow t_1 \leftarrow t_2 \leftarrow t_a \quad (4)$$

The Equation (4) also serves as the foundation for calculating when paths branch within the piping structure. It provides the necessary mathematical framework to assess how variables change at junctures where the flow diverges.

3.2 Handling of mixing time in CSTR

In an ideal CSTR, the assumption is that the mixing is instantaneous and complete. But in practice, the time it takes for the contents to be well mixed varies depending on the specific conditions and setup of the reactor. Generally, the mixing time in a practical CSTR can range from a few seconds to several minutes or more.

One simple empirical method for estimating mixing time θ_m is given by Equation (5):

$$\theta_m = \frac{K}{N} \quad (5)$$

where N is the impeller rotational speed and K is a constant that depends on the type of impeller and the reactor geometry [Ag15]. The constant K can vary depending on the specific setup and is often determined from experimental data or from literature if similar systems have been studied.

4 Data allocation with a combination of different process equipment

This section illustrates the methodology for calculating the flow of product streams through a complex piping system with a theoretical example. Using the method in Section 3, we model the transit times and volumes for each stream. By applying this approach to each segment, we demonstrate how to trace the path and position of multiple converging streams through a network of interconnected process equipment in a complex piping system.

4.1 Calculation with pipeline branching

The observed product at the outlet may comprise streams that have traversed distinct pathways, as indicated by the green and red routes in Fig. 2 (a). These streams do not consistently pass through the same passage and may spend varying durations within the production system. For instance, the green stream in Fig. 2 (a) follows a trajectory passing the positions of the sensors T01, T02, and T04, whereas the red stream passes along the positions of the sensors T01, T03, and T04. Depending on the flow rate and the geometry of the process equipment in both routes, those two product streams may have passed by T01 at different times. To accurately assign machine data to the respective product streams, it is essential to determine the precise moments when these streams passed specific passages, requiring the calculation of transit time and tracing of their historical positions throughout the production process.

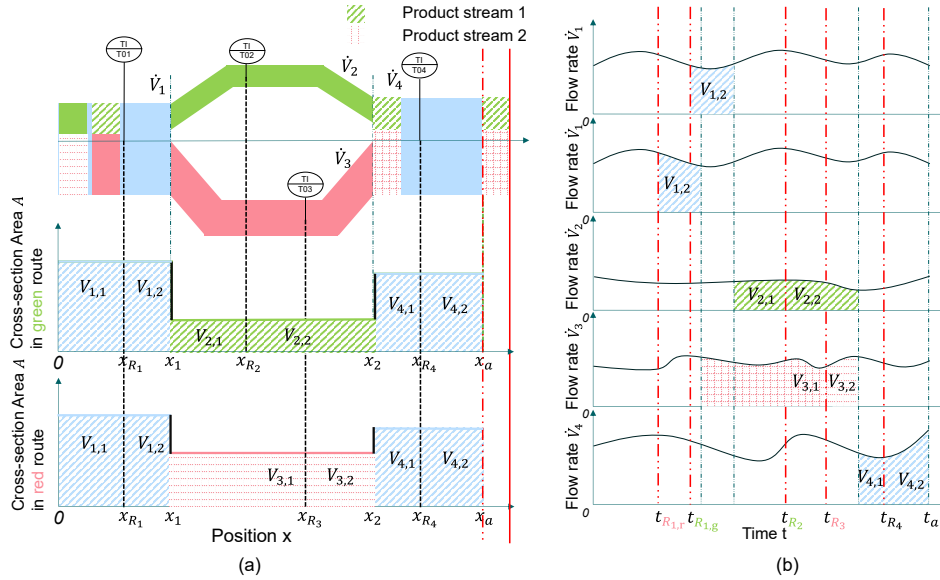


Fig. 2: (a) Product streams in a system with pipeline branching. (b) Results of the calculation

To calculate the transit time at which a product passes through a given passage within a particular route, it is necessary to address the separation of the product into different streams at the branching points. The paper employs Equation (3) at the junctures of process equipment to handle the separation of product flow, treating the product as two distinct portions pre- and post-branching. When the product is separated into multiple streams, the respective times for the different streams within their individual routes are calculated separately at the position where the streams diverge (e.g. at x_1 and x_2).

For instance, to calculate the transit time at which the product passes sensor T02, the time it takes for the product to reach position x_2 must be first determined using the flow rate \dot{V}_4 and the volume V_4 . Given the previously stated assumptions, x_a , t_a , \dot{V}_4 , \dot{V}_2 , and x_{R_2} are known. Since the geometry of all the process equipment in the system is also known, the volume $V_{2,2}$ (and $V_{2,1}$) can be also determined. Subsequently, the flow rate \dot{V}_2 and the volume $V_{2,2}$ are used to compute the time t_{R_2} , at which the product reaches the sensor at x_{R_2} (s. Fig. 2 (b)).

The t_{R_3} can be determined similarly. In the example provided within this paper, the times at which the two streams pass sensor T01 may differ. This is due to the difference in the time taken for them to reach position x_1 along the two routes, resulting in different passage times at T01 (s. Fig. 2 (b)), denoted as $t_{R_{1,r}}$ (in red route) and $t_{R_{1,g}}$ (in green route). The times $t_{R_{1,r}}$ and $t_{R_{1,g}}$ represent the moments when the end (the latter flowing part) of the product reaches sensor T01. The data measured by T01 is associated with these times (s. Fig. 3 (a)). To assign a time window that correlates with a considered unit (i.e. volume) of the product at the outlet of the process, the time when the first outflowing part of the product (indicated by the solid line in the diagram in Fig. 2 (a)) arrives at T01.

The part between the solid and dashed lines (s. Fig. 3 (b)) corresponds to the time window during which the observed product passes sensor T01.

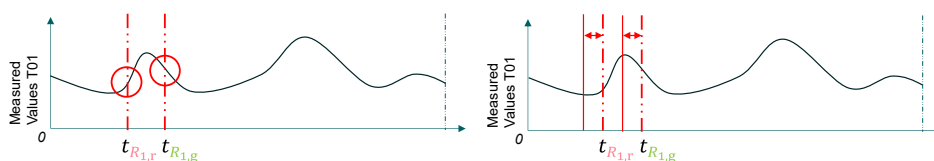


Fig. 3: (a) Data allocation for a product's cross-section. (b) Data allocation for the observed product.

4.2 Calculation with CSTR

In the context of a Continuous Stirred Tank Reactor (CSTR), the plug flow model is not applicable due to the inherent mixing dynamics within the reactor. Products in a CSTR are homogeneously mixed over a period, leading to a scenario where the composition of

the product exiting the reactor is essentially identical to that of the product entering. Given this situation, for the purposes of assigning product-related data with observed products, the paper treats the reactor in the following manner:

Regardless of the volume of the CSTR, once the designated mixing time has elapsed, it is assumed that the product at the inlet has effectively reached the outlet, where it, though, has the composition like in the CSTR, which, however, is not relevant for the transit time calculation that we address in this paper.

As shown in Fig. 4 (a), the system example includes several CSTRs (numbered with 1 to 3), which are assumed not to affect the overall length of the remaining parts of the pipeline in this paper. When calculating the transit time a product passes through a specified passage, it is essential to account for the mixing time associated with each CSTR based on its relative position to the passage.

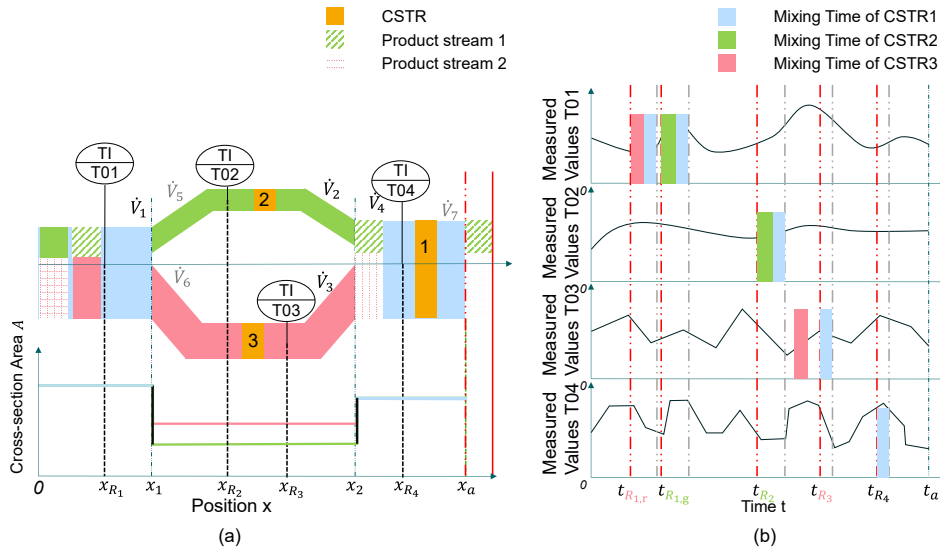


Fig. 4: Product streams in a system with pipeline branching and CSTR. (b) Results of the calculation.

As depicted in Fig. 4 (b), the presence of a CSTR introduces a delay in the timeline of calculations as compared to scenarios without a CSTR. Consequently, the transit time at which a product's cross-section passes through a passage is shifted from the position indicated by the grey dashed line to that marked by the red dashed line. For instance, sensors T01 and T02 are influenced by the delay induced by two CSTRs, while T03 and T04 are affected by only one CSTR. The method to calculate the transit time slot remains the same.

5 Discussion

This model provides a method to calculate and allocate the whole data history of product units with arbitrary defined quantities of product in continuous process manufacturing. Since this model is simplified and designed under several ideal conditions, it should be validated using either a practical system or simulation. The model can and should be extended to cover, e.g., parameter uncertainties and to allow for online applications.

Bibliography

- [Di19] DIN SPEC 91406:2019-12, Automatische Identifikation von physischen Objekten und Informationen zum physischen Objekt in IT-Systemen, insbesondere IoT-Systemen; Text Deutsch und Englisch, 2019. <https://doi.org/10.31030/3114151>
- [TMK21] Thies, M., Miny, T., Kleinert, T., 2021. Ein Konzept für Produkt-Verwaltungsschalen zur Produktbezogenen Integration von Produktionsdaten für die Qualitätsoptimierung, in: VDI Wissensforum GmbH (Ed.), Automation 2021. VDI Verlag, pp. 405–420. <https://doi.org/10.51202/9783181023921-405>
- [RK08] Roussinova, V., Kresta, S.M., 2008. Comparison of Continuous Blend Time and Residence Time Distribution Models for a Stirred Tank. *Ind. Eng. Chem. Res.* 47, 3532–3539. <https://doi.org/10.1021/ie070955r>
- [Kr18] Kruisz, J., Rehr, J., Faulhammer, E., Witschnigg, A., Khinast, J.G., 2018. Material tracking in a continuous direct capsule-filling process via residence time distribution measurements. *International Journal of Pharmaceutics* 550, 347–358. <https://doi.org/10.1016/j.ijpharm.2018.08.056>
- [SKS23] Sri Anbupalani, M., Kaliyannan, G.V., Sivaraj, S., 2023. IoT for Food and Beverage Manufacturing, in: Rajasekar, R., Moganapriya, C., Sathish Kumar, P., Harikrishna Kumar, M. (Eds.), *Integration of Mechanical and Manufacturing Engineering with IoT*. Wiley, pp. 141–157. <https://doi.org/10.1002/9781119865391.ch5>
- [CV20] Chen, X., Voigt, T., 2020. Implementation of the Manufacturing Execution System in the food and beverage industry. *Journal of Food Engineering* 278, 109932. <https://doi.org/10.1016/j.jfoodeng.2020.109932>
- [Li19] Liang, Q., Xiang, S., Long, J., Sun, W., Wang, Y., Zhang, D., 2019. Real-time comprehensive glass container inspection system based on deep learning framework. *Electronics Letters* 55, 131–132. <https://doi.org/10.1049/el.2018.6934>
- [Sl19] Slovák, J.; Vašek, P.; Šimovec, M.; Melicher, M.; Šišmišová, D. RTLS Tracking of Material Flow in Order to Reveal Weak Spots in Production Process. In *2019 22nd International Conference on Process Control (PC19)*; 2019; pp 234–238. <https://doi.org/10.1109/PC.2019.8815220>.
- [Ag15] Ascanio, G., 2015. Mixing time in stirred vessels: A review of experimental techniques. *Chinese Journal of Chemical Engineering* 23, 1065–1076. <https://doi.org/10.1016/j.cjche.2014.10.022>

„Ulrich Jumar, Christian Diedrich (Hrsg.):
EKA 2024 - Entwurf komplexer Automatisierungssysteme, 18. Fachtagung“


Effiziente Integration von Elektrolyseuren in modularen Elektrolyseanlagen

Einheitliche Schnittstellen zur Wasserstoffherzeugung


Julius Lorenz¹, Tobias Kock², Lukas Grahl³, Vincent Henkel⁴, Lena Scholz⁵, Katharina Stark⁶, Rene Lorenz⁷, Anselm Klose⁸, Leon Urbas⁹

Abstract: Die Erzeugung von Wasserstoff mittels Elektrolyse von Wasser bis in den Gigawatt-Maßstab, erfordert Anlagen, die aus einer Vielzahl von Elektrolyseuren modular aufgebaut sind. Für den Betrieb müssen alle prozessrelevanten Module der Anlage für die Prozessführung integriert werden. Eine aufwandsarme Integration mit allen für den Betrieb erforderlichen Informationen bildet die Grundlage für einen schnellen, effizienten Aufbau und spätere Anpassungen von Elektrolyse-Anlagen. Der Beitrag stellt die wesentlichen Integrationszenarien aus Betreiberperspektive vor. Dafür werden die Anforderungen der Integration bezüglich Information und Funktion abgeleitet. Resultat ist ein herstellerunabhängiges Standardintegrationsprofil. Das Standardintegrationsprofil bildet die Grundlage für die Umsetzung mittels Module Type Package aus der modularen Automation in der Prozessindustrie.

Keywords: Modulare Automation, Leitsystem-Integration, Prozessführung, Wasserstoff, Elektrolyse.

¹ Technische Universität Dresden, Chair of Process Control Systems & Group of Process Systems, Dresden, julius.lorenz@tu-dresden.de,  <https://orcid.org/0000-0001-7124-4844>


² Technische Universität Dresden, Chair of Process Control Systems & Group of Process Systems, Dresden, tobias.kock@tu-dresden.de

³ Semodia GmbH, Anschrift, Meißner Str.37, 01445 Radebeul, lukas.grahl@semodia.com,  <https://orcid.org/0000-0002-7335-0971>


⁴ Helmut-Schmidt-Universität Hamburg, Institut für Automatisierungstechnik, vincent.henkel@hsu-hh.de


⁵ Helmut-Schmidt-Universität Hamburg, Institut für Automatisierungstechnik, lena.scholz@hsu-hh.de

⁶ ABB AG Forschungszentrum Deutschland, Ladenburg, katharina.stark@de.abb.com

⁷ Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Institut für Technische Thermodynamik, Stuttgart, rene.lorenz@dlr.de,  <https://orcid.org/0000-0002-9539-4182>

⁸ Process-to-Order Lab, Technische Universität Dresden, Dresden, anselm.klose@tu-dresden.de,

 <https://orcid.org/0000-0003-3954-7786>

⁹ Technische Universität Dresden, Chair of Process Control Systems & Group of Process Systems, Dresden, leon.urbas@tu-dresden.de,  <https://orcid.org/0000-0001-5165-4459>

1 Einleitung

Die Erzeugung von grünem Wasserstoff mittels Elektrolyse von Wasser ist eine Schlüsseltechnologie für die Dekarbonisierung schwer elektrifizierbarer Sektoren wie der Chemie- und Schwerindustrie und dem Schwertransport [Sm18]. Um die dafür benötigten Wasserstoffmengen bereitzustellen, müssen die verfügbaren Elektrolysetechnologien hochskaliert und die Wasserstofferzeugungskapazität ausgebaut werden. Die Größe einer Elektrolyse-Anlage richtet sich vor allem nach der verfügbaren elektrischen Energie oder dem Bedarf eines Folgeprozesses [La23]. Anlagengrößen im notwendigen Gigawatt-Maßstab sind jedoch noch Stand von Forschung und Entwicklung [La24]. Da die Größe eines einzelnen Elektrolyseurs in der jeweiligen Technologie physikalisch begrenzt ist, muss eine solche Anlage aus mehreren Elektrolyseuren modular aufgebaut werden. Je nach anvisierter Leistung und Art der Anlage ergeben sich dafür unterschiedliche Modularisierungsansätze [Sc22]. Eine Gigawatt-Anlage aus derzeit verfügbaren Elektrolyse-Stacks aufzubauen, erfordert je nach Stackgröße die Integration vieler einzelner Stacks. Nutzt man beispielsweise Druck-Alkali-Elektrolyseure wie den „Sunfire-HyLink Alkaline“ der Sunfire GmbH mit je 10MW Leistung zum Aufbau ergeben sich rechnerisch 100 Stacks für 1GW Elektrolyseleistung [SU23].

Der modulare Aufbau einer Anlage erfordert die Integration der einzelnen Elektrolyseure; aktuell sind herstellerepezifische Beschreibungen und Integrationsszenarien der Normalfall. Die Integration einer Vielzahl von Elektrolyseuren mit einem herstellerepezifischen Ansatz resultiert damit in einer spezifischen Integration für jede erstellte Anlage, verbunden mit einer hohen Komplexität und Aufwänden. Diesem durch einheitliche Ansätze zu begegnen ist Zielstellung des in diesem Artikel vorgestellten Standardintegrationsprofils für Elektrolyseure. Gerade die Vielzahl redundanter Elemente mit ähnlichem Informations- und Funktionsumfang macht einen einheitlichen Ansatz im Bereich von Elektrolyseuren attraktiv.

2 Vorgehensmodell zur Entwicklung eines Standardintegrationsprofils

Für die Entwicklung eines Standardintegrationsprofils müssen die notwendige Information und die relevanten Funktionen ermittelt, formalisiert sowie verifiziert und validiert werden. In H2Giga-eModule wurde dafür ein sequentielles Vorgehensmodell mit sechs Schritten angewendet (Abbildung 1). Angesichts der klar umreißbaren Aufgabenstellung wurde eine Kombination von Wasserfallmodell [ISO24765] und inkrementellem Modell [ISO24765] für die Informationsmodellierung gewählt.

Im ersten Schritt werden die Integrationsszenarien für modulare Elektrolyse-Anlagen auf Basis von Literatur ermittelt. Dies geschieht einerseits mittels einer systematischen Literaturrecherche zu existierenden Beschreibungen von Elektrolyseuren und Elektrolyse-

Systemen und andererseits durch empirische Untersuchungen typischer Aufgaben und dem damit verbundenen Informations- und Funktionsbedarfs aus Betreibersicht. Im zweiten Schritt werden die notwendigen Anforderungen hinsichtlich Information und Funktionen innerhalb der Rahmenbedingungen und Systemgrenzen aus Schritt 1 bestimmt. Die daraus resultierende Informationsmenge und Funktionsumfang werden im dritten Schritt aufbereitet, damit im vierten Schritt das herstellerunabhängige Standardintegrationsprofil entwickelt werden kann. Das Standardintegrationsprofil bildet dann die Grundlage für die Erstellung spezifischer Lösungen. In Schritt 5 wird dies am Beispiel des Module Type Package (MTP) für den Teilaspekt der automatisierungstechnischen Integration gezeigt. Im letzten Schritt erfolgt die Verifikation und Validierung der konkreten Umsetzung auf Basis virtueller und physischer Demonstratoren.

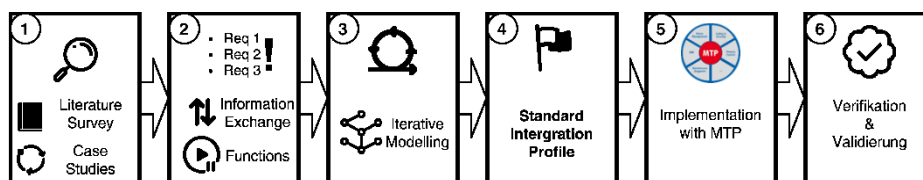


Abb. 1: Vorgehensmodell zur Entwicklung eines Standardintegrationsprofils

3 Integrationsszenarien

Eine Integration ist nach [IEEE] „[die] Zusammenführung oder Kombination von zwei oder mehr Elementen einer niedrigeren Ebene zu einem funktionierenden und einheitlichen Element einer höheren Ebene, bei dem die funktionalen und physischen Schnittstellen erfüllt sind.“. Im Anwendungsfall der modularen Elektrolyse-Anlagen ist dies das Einbinden einzelner Module in die Gesamtanlage Anlage zur Erzeugung von Wasserstoff. Hierbei müssen die Schnittstellen eines Moduls auf stofflicher, energetischer und informationstechnischer Ebene verbunden werden. Dies bedeutet, dass diese per Design so gestaltet sein müssen, dass die Kombination konfliktfrei möglich ist. Schwerpunkt der in diesem Artikel durchgeführten Analyse ist die informationstechnische Integration. Diese, im Folgenden kurz als Integration bezeichnet, behandelt den Austausch von Daten und das zielgerichtete Ansteuern der Funktionen der einzelnen Einheiten für den sicheren und wirtschaftlichen Betrieb der Gesamtanlage.

Je nach Integrationsszenario können sich die zu integrierenden Module unterscheiden. Ein wesentliches Kriterium neben anderen wie der Technologie ist die Granularität der Integrationsebene. Bei Elektrolyseanlagen lassen sich zwei wesentliche Arten unterscheiden, wie auch in Abbildung 2 dargestellt:

- A) Integration einzelner Prozesseinheiten v.a. Stacks und Balance-of-Plant (BoP) Komponenten analog zu [Bi22, La24]

- B) Integration mehrerer vorkonfigurierter Elektrolyse-Systeme (intern z.B. gestaltet wie A, aber als Ganzes inklusive anteiligem BoP integriert z.B. als Container)

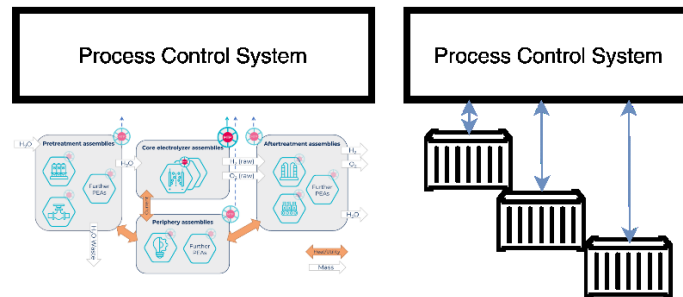


Abb. 2: Referenz-Integrationsszenarien von Elektrolysesystemen

Bei der Integration nach A müssen im Vergleich zu B mehr und auch im Hinblick auf die Energie- und Stoffwandlungsprozesse sehr unterschiedliche Module in das Leitsystem integriert werden. Bemisst man den Integrationsaufwand anhand der Schnittstellenanzahl und Anzahl unterschiedlicher Module ist dieser bei gleicher Produktionskapazität aufgrund der Anzahl der Schnittstellen bei A) höher als bei B). Gleichzeitig bestehen jedoch bei A) im Vergleich zu B) aus Betreibersicht mehr Freiheitsgrade, da durch einzelne Integration auch individuelle Anlagenkonfigurationen möglich sind. Das wesentliche Modul zur Erzeugung von Wasserstoff ist in Szenario A) der Stack. Hinzu kommen die BoP-Module, die ebenfalls über standardisierte informationstechnische Schnittstellen verfügen sollen.

In Szenario B) wird ein vorkonfiguriertes System aus z.B. Containern, die intern eigenständig für die Koordination der in A) genannten Module sorgen, integriert. Der Integrationsaufwand ist reduziert, da deutlich weniger Schnittstellen zwischen Modul und Leitsystem bestehen. Dieser Fall ergibt sich beispielsweise, wenn ein Betreiber seine Produktion auf Wasserstoff umstellen möchte und auf ein oder mehrere schlüsselfertige, vorautomatisierte Systeme zur Bereitstellung des Wasserstoffs setzt.

Mit einheitlichen Schnittstellen könnten zudem Blueprints für Prozessführungskonzepte entwickelt werden. Der Blueprint würde die grundsätzlichen Regeln für die Koordination der Module definieren und müsste für den Betrieb nur hinsichtlich der Anzahl der in einer konkreten Anlage enthaltenen Module mit den jeweiligen Eigenschaften gemäß Standardintegrationsprofil angepasst werden.

Das Standardintegrationsprofil sollte auch Möglichkeiten zur Anpassung und Optimierung bereitstellen. Ein Ansatz ist eine hohe Selbstbeschreibungsfähigkeit [IEEE]. Dies kann entweder durch zusätzliche Indikatoren, die eine bessere Vorgabe von Soll Arbeitspunkten ermöglichen oder von Parametern zur Anpassung des Betriebspunktes an die jeweiligen Einsatzbedingungen, erfolgen.

4 Anforderungen an die Modulbeschreibung

Das zu integrierende Modul muss hinreichend bezüglich seiner Schnittstellen, Eigenschaften und des Verhaltens beschrieben sein, um den Betrieb zu ermöglichen. Das Standardintegrationsprofil muss diese Informationen enthalten, um eine einheitliche Integration zu gewährleisten. Der Umfang wird dabei so gewählt, dass eine Eignung für möglichst viele Integrationsszenarien besteht. Aus diesem Grund bezieht sich die Spezifikation auf die Kernfunktionalität der Erzeugung und Bereitstellung von Wasserstoff am Modulausgang, welche die Möglichkeit zur Erweiterung bietet. Die inhaltliche Ausgestaltung und Struktur sollen zudem die Anforderungen der Prozessindustrie nach VDI 2770 „Mindestanforderungen an Herstellerinformationen für die Prozessindustrie“ [VDI20b] erfüllen. Die darin enthaltene funktionale/technische Spezifikation sollte als Teilmenge auch die Beschreibung der automatisierungstechnischen Schnittstelle zur Integration beinhalten. Die Informationen bei einer technischen Umsetzung sollten vorrangig in maschinenlesbaren Formaten vorliegen. Diese Forderung erweitert die Mindestanforderung nach PDF-A in VDI 2770 [VDI20b]. Beispiele für standardisierte Ansätze aus der Prozessindustrie und modularen Automation sind das MTP oder im Kontext der Industrie 4.0 und des Digitalen Zwillings die Verwaltungsschale.

5 Erhebung des Informations- und Funktionsbedarfs

Das Integrationsprofil selbst soll die notwendige Information für den Betrieb abbilden. Um die derzeit bereitgestellte Information über einen Elektrolyseur zu berücksichtigen, wurden 20 Spezifikationen verschiedener Elektrolyseure und Elektrolysesysteme hinsichtlich Eigenschaften und Funktionen ausgewertet und zur Identifikation der Kernfunktionalität die Schnittmenge gebildet, wie exemplarisch für zwei Hersteller in Abbildung 3 zu sehen ist.

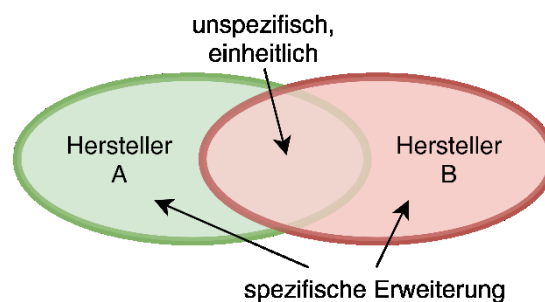


Abb. 2: Venn-Diagramm der Informationen zweier Hersteller

In der Schnittmenge enthalten waren u.a.

- Betriebsbereich,
- Teillastfähigkeit,
- Leistungsaufnahme,
- Wasserverbrauch und -qualität, sowie
- Wasserstoff-Produktionsmenge, -qualität und -druck.

Bei der konkreten Benennung und den Einheiten der Eigenschaften bestanden dabei allerdings Unterschiede. Bei einer nicht standardisierten Integration würde hier zusätzlicher Aufwand entstehen, weswegen an dieser Stelle eine Vereinheitlichung sinnvoll ist.

Betriebliche Funktionen, automatisierungstechnische Schnittstellen und Verhalten waren in den Spezifikationen nur in Einzelfällen vorhanden z.B. in Form von Anfahrzeiten. Aus diesem Grund wurde die Auswertung durch Untersuchung der notwendigen Daten und Funktionen, die im Betrieb bereitgestellt werden müssen, ergänzt. Dies beinhaltet Merkmale zur

- eindeutigen Identifikation der Module,
- des aktuell verfügbaren Funktionsumfangs,
- dem aktuellen Zustand,
- den verfügbaren Schnittstellen (Information, Stoff, Energie) sowie
- Beschreibungen des Sollverhalten (für Planung und Optimierung).

Für jeden dieser Bereiche leiten sich inhaltliche Anforderungen ab. Bei der Identifikation muss das Modul selbst als auch Information über die Elektrolyse-Technologie und Integrationsebene verfügbar sein. Die Bereiche der Funktion und Zustand besitzen eine enge Kopplung, da ein Betreiber mit den verfügbaren Funktionen den Zustand des Moduls steuern kann. Zu diesem Zweck muss bekannt sein, welche Funktion(en) in welchem Zustand zur Verfügung steht und welcher Zustand mit dem Funktionsaufruf erreicht wird. Zusätzlich sollte für Planung und Optimierung des Betriebs ein Modell des realisierten Verhaltens in Bezug auf die stofflichen, energetischen und informationstechnischen Schnittstellen bereitgestellt werden, damit diese bei Design und Ausführung der Prozessführung [Lo23] berücksichtigt werden kann. Diese Beschreibung ist nach heutigem Stand nur in wenigen Fällen in textueller Form, ergänzt durch Änderungsdaten verfügbar. Für einen reduzierten Abstimmungsbedarf, eine automatisierte Integration und die Reduktion manueller Schritte sollten diese und auch die anderen bereits genannten Inhalte in maschinenlesbaren Formaten/Modellen verfügbar sein.

Im Betrieb großskaliger Elektrolyse-Anlagen kann zudem angenommen werden, dass die Bedeutung der Interaktion mit einzelnen Aktoren und individuellen Bedienbildern gering ist, da Szenarien mit 100 oder 1000 Modulen der modularen Anlage in der gleichen Anzahl einzelner Ansichten resultieren würde. Unter der Annahme, dass die Steuerung der Anlage maßgeblich durch Rezepte und Algorithmen realisiert wird, wären Bedienbilder und Einzelsteuerelemente optionale Inhalte des Integrationsprofils, die bei der Inbetriebnahme, einer Fehleranalyse und bei der Instandhaltung genutzt werden.

6 Standardintegrationsprofil

Das Standardintegrationsprofil definiert die Mindestanforderungen bezüglich der Information, die für eine harmonisierte Integration der Elektrolyse-Module zur Verfügung gestellt werden sollen. Die technische Spezifikation und betrieblichen Schnittstelle, mit der Definition der Daten und Funktionen, stellt den größten Mehrwert für die effiziente Integration dar. Insgesamt werden im Standardintegrationsprofil 24 Merkmale und 8 Funktionen definiert. Je nach Granularität kann die konkrete inhaltliche Gestaltung des Standardintegrationsprofils unterschiedlich sein. Die damit verbundene Skalierbarkeit wird durch eine freie Angabe der konkreten Werte der Merkmale z.B. der Leistungsaufnahme für die jeweilige Granularität der Integrationsebene erreicht. Da der Großteil der Merkmale die in der Analyse ermittelten Informationen aus der Schnittmenge verschiedener Hersteller abbilden, sind diese für eine Umsetzung mit dem Standardintegrationsprofil bereits verfügbar. Zusatzaufwände entstehen vor allem bei der Vereinheitlichung von Merkmalen die nicht innerhalb der Schnittmenge liegen wie beispielsweise dem Betriebszustand. Die spezifizierten Betriebszustände *Idle*, *Starting (Cold-Start)*, *Production*, *Approaching Hot Standby*, *Hot-Standby*, *Starting (Hot-Start)*, *Cooldown* und *Error* müssen demnach realisiert werden. Bei der Bereitstellung der Schnittstellenbeschreibungen, den Funktionen und Verhaltensmodellen besteht voraussichtlich der größte Aufwand, da diese bisher nicht oder nur individuell realisiert sind. Das Standardintegrationsprofil sieht auch die Berücksichtigung der Funktionen, des Verhaltens und der Struktur nach dem Function-Behavior-Structure Modell [GE90] vor. Die je nach Granularität an der Modulgrenze verfügbaren Schnittstellen werden als Teil der Struktur beschrieben. Funktionen und Verhaltensmodelle referenzieren diese dann. Die Funktionen ändern den Betriebszustand des Moduls z.B. überführt die Funktion „*Start Production*“ das Modul aus dem Zustand *Idle* in *Production*. Gemäß der Beschreibung des Sollverhaltens durch Verhaltensmodelle ändern sich dabei die Prozessgrößen sowie die Stoff- und Energieströme mit Referenz auf die jeweilige Schnittstelle. Für die Planung des Betriebs und der Abstimmung der Schnittstellen mit angrenzenden Modulen ist die Angabe des Verhaltens in den Betriebszuständen zielführend z.B. typischer Temperatur- oder Druckverläufe während *Starting(Cold-Start)*.

Das vollständige Standardintegrationsprofil ist auf GitHub verfügbar (<https://github.com/p2o-lab/H2Giga-eModule-standard-integration-profile>).

7 Implementierung mittels Module Type Package

Die Problemstellung von Integration und Betrieb verschiedener Module besitzt bezüglich der Zielstellung eine große Analogie zur modularen Automation in der Prozessindustrie. Das MTP als Schnittstellenbeschreibung für Process Equipment Assemblies (PEAs), dient als einheitliches Metamodell für eine schnelle Integration von PEAs in der Prozessindustrie [VDI22, VDI20a]. Dieses bereits vorhandene und erprobte Konzept aus dem Bereich der Prozessindustrie soll zur Implementierung des Standardintegrationsprofils genutzt werden [Bi22]. Die Elemente eines Elektrolysesystems werden dabei als PEA definiert und realisiert die Automatisierungstechnischen Anforderungen. Dazu muss die Spezifikation des Integrationsprofils mit den MTP-Datenobjekten realisiert werden. Dazu werden die folgenden Regeln zur Repräsentation mittels MTP angewendet:

- Eindeutige Identifizierbarkeit durch das *PeaInformationLabel* mit der *DeviceClass* zur eindeutigen Beschreibung der Art des Elektrolysesystems
- Bereitgestellte Informationen durch *IndicatorElements* und *ProcessValueOuts*
- Eingaben realisiert durch *OperationElements* und *InputElements*
- Wertebeschränkungen realisiert am jeweiligen Objekt
- Betriebszustände realisiert als *Enumeration* im *Text-Set*
- Funktionen realisiert durch *Dienste*
- Optionale Bedienbilder realisiert als Teil des *HMI-Set*
- Optionale Einzelsteuerelemente als *OperationalElements* und *ActiveElements*

Die Transformation der spezifizierten Funktionen auf das Dienste-Modell bietet nochmals eine Erleichterung, da die einzelnen Funktionen auf einen Dienst übertragen werden, der die Funktion der Elektrolyse als „Electrolysis“ bereitstellt. Ein Dienst besitzt mindestens eine Prozedur, welche in diesem Fall kontinuierlich ausgeführt wird. Für die konkrete Ausgestaltung werden die Zustände des Integrationsprofils auf die Zustände des Dienste-Modells gemapped. Da das Dienste-Zustandsmodell bereits in der VDI/VDE/NAMUR 2658-4 Richtlinie standardisiert ist, wird eine herstellerunabhängige Ansteuerung des Dienstes gewährleistet. *Hot Standby* wird auf die Pause Loop des Dienste-Modells abgebildet und die Fehlerbehandlung ist über Konzepte im MTP möglich. Die Funktion der Arbeitspunktvorgabe wird durch ein *ParameterElement* realisiert, welches den Setpoint-Wert an den Dienst „Electrolysis“ übergibt.

Im MTP können nicht alle im Integrationsprofil spezifizierten Informationen enthalten sein, wie Verhaltensmodelle oder Strukturmodelle. Diese können als Attachment-Set ergänzt werden oder auch zukünftig als Teil einer PEA-Dokumentation bereitgestellt werden.

8 Zusammenfassung und Ausblick

Die effiziente Integration von Elektrolyseuren bildet die Grundlage für den schnellen Aufbau und Rekonfiguration von Elektrolyse-Anlagen. Der Beitrag zeigt, wie die dafür notwendigen Informationen und Funktionen ermittelt werden können. Ergebnis ist ein Vorgehensmodell sowie das resultierende umsetzungsneutrale Standardintegrationsprofil, welches als Metamodell für die konkrete Umsetzung dient, zum Beispiel mittels der Konzepte aus der modularen Automation in der Prozessindustrie. Die konkrete Umsetzung mittels MTP kann für die vereinfachte Integration und Ansteuerung genutzt werden [Lor23]. Im Rahmen des Forschungsprojektes H2Giga-eModule erfolgt als nächster Schritt die Verifikation und Validierung des Integrationsprofils auf einer Demonstratoranlage.

Zukünftig kann das vorgestellte Konzept durch die Verwendung einheitlicher, maschinenlesbarer Formate und Industriestandards für die Dokumentation wie die Verwaltungsschale im Kontext der Industrie 4.0, OPC UA Companion Specification und auch einheitliche Konzepte für Verhaltensmodelle z.B. durch standardisierte Simulationsmodelle verbessert werden. Zudem kann auch die Selbstbeschreibungsfähigkeit der Systemfunktionen z.B. durch die Ergänzung von System-Control Diagram (SCD) [Tk23] weiter verbessert werden. Für die konkrete Umsetzung mit MTP ist eine Erweiterung um noch in Entwicklung befindlicher Aspekte wie Safety, Alarming sowie Diagnose und Maintenance möglich.

9 Danksagung

Dem Bundesministerium für Bildung und Forschung sowie dem Projektträger Jülich sei für die finanzielle Unterstützung im Rahmen des Forschungsprojektes eModule (FKZ 03HY116A-E) der Leitplattform H2Giga gedankt.

Literaturverzeichnis

- [Bi22] Bittorf, L.; Beisswenger, L.; Erdmann, D.; Lorenz, J.; Klose, A.; Lange, H.; Urbas, L.; Markaj, A. and Fay, A.: "Upcoming domains for the MTP and an evaluation of its usability for electrolysis," in 2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA), Sep. 2022, pp. 1-4. doi: 10.1109/ETFA52439.2022.9921280
- [Ge90] Gero, J. S. (1990). Design Prototypes: A Knowledge Representation Schema for Design. AI Magazine, 11(4), 26. <https://doi.org/10.1609/aimag.v11i4.854>
- [IEEE] "The Authoritative Dictionary of IEEE Standards Terms, Seventh Edition," in IEEE Std 100-2000, vol., no., pp.1-1362, 11 Dec. 2000, doi: 10.1109/IEEESTD.2000.322230.
- [ISO24765] ISO/IEC/IEEE 24765, „System und Software-Engineering – Begriffe“, September 2017

-
- [La23] Lange, H.; Klose, A.; Lippmann, W. and L. Urbas: "Technical evaluation of the flexibility of water electrolysis systems to increase energy flexibility: A review," *International Journal of Hydrogen Energy*, vol. 48, no. 42, pp. 15 771-15 783, May 2023. [Online]. Available: [10.1016/j.ijhydene.2023.01.044](https://doi.org/10.1016/j.ijhydene.2023.01.044)
- [La24] Lange, H.; Klose, A.; Beisswenger, L.; Erdmann, D.; & Urbas, L. (2024). Modularization approach for large-scale electrolysis systems: a review. In *Sustainable Energy & Fuels* (Vol. 8, Issue 6, pp. 1208–1224). Royal Society of Chemistry (RSC). <https://doi.org/10.1039/d3se01588b>
- [Lo23] Lorenz, J.; Klose, A.; Lange, H.; Kock, T. and Urbas, L.: "Flexible Process Control for Scalable Electrolysis Systems," 2023 International Conference on Electrical, Computer and Energy Technologies (ICECET), Cape Town, South Africa, 2023, pp. 1-6, doi: [10.1109/ICECET58911.2023.10389401](https://doi.org/10.1109/ICECET58911.2023.10389401).
- [Tk23] Kock, T.; Klose, A.; Ingebringsten, I. P.; Bittorf, L.; Lorenz, J. and Urbas, L.: "Transformation of process functions to the Module Type Package utilizing System Control Diagrams," 2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA), Sinaia, Romania, 2023, pp. 1-4, doi: [10.1109/ETFA54631.2023.10275331](https://doi.org/10.1109/ETFA54631.2023.10275331).
- [Sc22] Scholz, L.; Markaj, A.; Fay, A.; Erdmann, D.; Beisswenger, L.; Lorenzen, L.; Urbas, L. and Bittorf, L.: "Anforderungen an modulare Elektrolyseanlagen: Erkenntnisse aus dem H2Giga - Projekt eModule", *atp magazin*, vol. 63, no. 6-7, pp. 62-70, Jun. 2022. [Online]. Available: [10.17560/atp.v63i6-7.2611](https://doi.org/10.17560/atp.v63i6-7.2611)
- [Sm18] Smolinka, T.; Wiebe, N.; Sterchele, P.; Palzer, A.; Lehner, F.; Jansen, M.; Kiemel, S.; Miede, R.; Wahren, S. and Zimmermann, F.: "Studie IndWEde. Industrialisierung der Wasserelektrolyse in Deutschland," 2018. [Online]. Available: <https://publica.fraunhofer.de/handle/publica/2993>
- [SU23] Sunfire GmbH. (Nov. 2023). *DatenBlatt Sunfire HyLink Alkaline*. Sunfire DatenBlatt HyLink Alkaline. Abgerufen am 31. Januar 2024, von https://www.sunfire.de/files/sunfire/images/content/Produkte_Technologie/factsheets/Sunfire-Factsheet-HyLink-Alkaline_2023Nov.pdf
- [VDI20a] VDI. 2776-1, "Process engineering plants - Modular plants - Fundamentals and planning modular plants," Nov. 2020.
- [VDI20b] VDI 2770-1, „Betrieb verfahrenstechnischer Anlagen - Mindestanforderungen an digitale Herstellerinformationen für die Prozessindustrie – Grundlagen“, Beuth, Apr. 2020
- [VDI22] VDI/VDE/NAMUR. 2658-1, "Automation engineering of modular systems in the process industry - General concept and interfaces," Tech. Rep., Jan. 2022.

Koordinierter Betrieb eines Pumpspeicherkraftwerkes mit reversibler Pumpturbine und einem Batteriespeicher an einem geteilten Standort

Leon Tadayon¹, Josef Meiers¹, Lukas Ibing², Kevin Erdelkamp² und Georg Frey¹

Abstract: Diese Veröffentlichung untersucht die Koordination von Pumpspeicherkraftwerken (PSW) und Batteriespeichern (BS) zur Steigerung der Wirtschaftlichkeit. PSW bieten eine hohe Speicherkapazität, jedoch begrenzte Reaktionszeiten, während BS niedrigere Kapazität, aber schnellere Reaktionszeiten aufweisen. Ein hybrides System aus beiden kann somit Synergien nutzen. Es wurde ein Mixed-Integer-Linear-Programming-Modell entwickelt, das den koordinierten Einsatz beider Systeme abbildet. Dabei agiert das PSW im Day-ahead-Markt und bietet Sekundärregelleistung, während der BS Primärregelleistung bereitstellt. Das Modell berücksichtigt Kosten durch Verschleiß und Alterungseffekte. Die Ergebnisse zeigen eine Ertragssteigerung von 12,35% durch die Koordination im Vergleich zum eigenständigen Betrieb beider Speichersysteme. Dieser Mehrwert kann durch eine effizientere Nutzung der Kapazitäten, insbesondere des BS, im koordinierten Einsatz realisiert werden.

Keywords: Pumpspeicherkraftwerk (PSW), Batteriespeicher (BS), Reversible Pumpturbine, Strommarkt, Simulation, Mixed-Integer-Linear-Programming (MILP)

1 Einleitung

Pumpspeicherkraftwerke (PSW) haben sich als eine etablierte und ausgereifte Technologie erwiesen und stellen den größten Teil der elektrischen Energiespeicherung weltweit dar [SK12]. Während sie eine große Speicherkapazität bieten, ist ihre Reaktionszeit jedoch begrenzt, und der Ausbau dieser Speicherform in Deutschland ist aufgrund erschöpfter Potenziale sowie ökologischer und finanzieller Risiken ins Stocken geraten. Im Gegensatz dazu sind die Kosten für Batteriespeicher (BS) in den letzten Jahren deutlich gesunken [In23], was sie zu einer attraktiven Technologie für stationäre Großspeicher macht. Batteriespeicher bieten zudem hohe Reaktionsgeschwindigkeiten durch ihre Leistungselektronik und eignen sich somit für Anwendungen, für die ein PSW zu träge wäre. Der deutsche Energiemarkt besteht aus verschiedenen Teilmärkten mit unterschiedlichen Produkten und Anforderungen, für die verschiedene Speichersysteme geeignet sind. Ein hybrides Speichersystem aus verschiedenen Technologien bietet daher die Möglichkeit, Synergien zu nutzen und die Vermarktungsmöglichkeiten zu verbessern [BSB19]. In dieser Studie wird der koordinierte Einsatz von PSW und BS als ein solches hybrides Energiespeichersystem betrachtet.

1 Universität des Saarlandes, Lehrstuhl für Automatisierungs- und Energiesysteme, Campus A5 1, 66123 Saarbrücken, Deutschland, leon.tadayon@aut.uni-saarland.de; josef.meiers@aut.uni-saarland.de; georg.frey@aut.uni-saarland.de

2 RWE Generation SE, Hydro & and Storage Asset Management, RWE Platz 3, 45141 Essen, Deutschland, lukas.ibing@rwe.com; kevin.erdelkamp@rwe.com

Dabei besteht der PSW aus einer reversiblen Pumpturbine und der BS wird als Lithium-Ionen-Batterie angenommen. Das Hauptziel in dieser Untersuchung ist die Modellierung und Optimierung des koordinierten Betriebs beider Systeme zur maximalen Ausnutzung ihrer jeweiligen Kapazitäten. Zu diesem Zweck wird ein Modell entwickelt, das die spezifischen Eigenschaften, Einschränkungen und Betriebszustände von PSW und BS berücksichtigt. Als Vorlage für die Modellierung dient der reale Standort eines PSW in Herdecke, an dem sich auch ein BS befindet [RW24a; RW24b]. Die Marktteilnahme am Großhandelsmarkt sowie die Bereitstellung von Systemdienstleistungen sind in das Modell integriert. Der PSW agiert im Day-ahead-Markt und stellt Sekundärregelleistung (SRL) bereit, während der BS Primärregelleistung (PRL) bereitstellt. Ein Koordinationsschema für den Einsatz beider Speichersysteme wird entwickelt, wobei die Kosten im Zusammenhang mit dem Verschleiß mechanischer Komponenten des PSW und der Alterung der Batterie berücksichtigt werden. Die Optimierung zielt darauf ab, den erwirtschafteten Ertrag des Systems zu maximieren, und wird mithilfe der Programmiersprache Python [VD14] und dem Modellierungspaket Pyomo [By21] erstellt. Der Solver Gurobi [Gu23] wird zur Lösung des Optimierungsproblems eingesetzt. Diese Simulation erstreckt sich über ein Jahr und nutzt historische Marktdaten im Zeitraum vom 01.01.2023 bis zum 31.12.2023, die in das Modell integriert sind. Die Simulation wird iterativ durchgeführt [BH16]. Durch die Aufteilung des Simulationszeitraums in kleinere Zeitabschnitte mit einer Dauer von einem Tag und die iterative Optimierung kann die Gesamtrechenzeit deutlich reduziert und die Teillösungen können zu einer optimalen Gesamtlösung kombiniert werden.

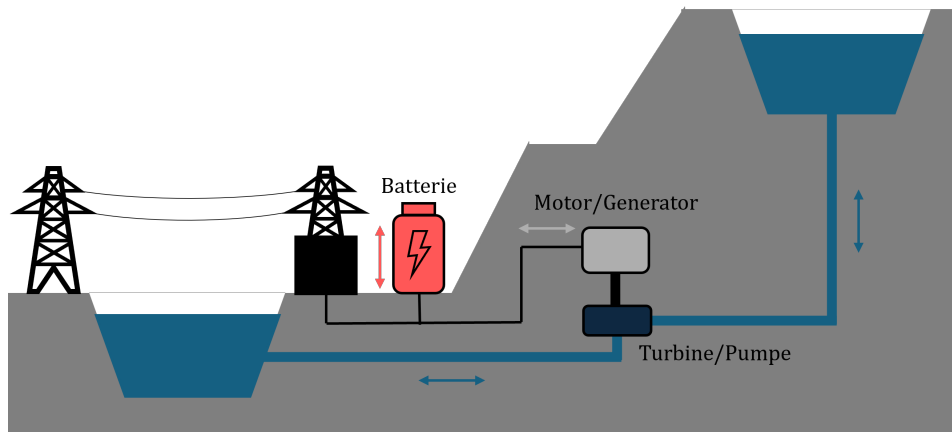


Abb. 1: Abbildung der Struktur des PSW mit reversibler Pumpturbine und elektrischer Maschine mit verbundenem BS.

2 Modellierung

Das kombinierte Speichersystem und die jeweiligen Anwendungen werden mittels Mixed-Integer-Linear-Programming (MILP) modelliert, wobei sowohl kontinuierliche als auch

ganzzahlige Variablen die Freiheitsgrade des Systems repräsentieren. Die Modellierung baut dabei auf einer einfacheren Modellierung [Ta] und Ansätzen aus der Literatur auf [BSB19; CPG16]. Das resultierende Optimierungsproblem wird durch eine definierte Zielfunktion gelöst, wodurch der optimale Betriebsplan des Systems für den analysierten Zeitraum ermittelt wird. Die Simulation erfolgt in Zeitschritten von jeweils fünf Minuten.

2.1 Pumpspeicherkraftwerk zur Bereitstellung von Sekundärregelleistung

Im Rahmen der PSW-Modellierung wird das Unterbecken der Anlage als ausreichend groß betrachtet, sodass es im Vergleich zum oberen Speicherbecken vernachlässigt werden kann. Somit ist für die Modellierung des PSW lediglich der Füllstand des Oberbeckens relevant. Der Füllstand des Beckens L_{PSW} [%] in einem bestimmten Zeitschritt wird durch Gl. 1 bestimmt. Hierbei steht V_{PSW}^N für das Speichervolumen des PSW [m³], Q_{Pump} für den Zufluss im Pumpbetrieb [m³/s], Q_{Turb} für den Abfluss im Turbinierbetrieb [m³/s] und t_s für die Zeitschrittlänge [s]. Start- und Endfüllstand des Beckens werden als halb voll angenommen.

$$L_{PSW}(t) = \frac{[Q_{Pump}(t) - Q_{Turb}(t)] \cdot t_s}{V_{PSW}^N} + L_{PSW}(t-1) \quad (1)$$

Im Betrieb als Pumpe kann die Leistung nicht reguliert werden, was bedeutet, dass die Pumpturbine im Pumpbetrieb lediglich in einem fixen Arbeitspunkt genutzt werden kann. Dieser Zufluss resultiert aus dem An/Aus-Zustand u_{Pump} [-] und dem Nenndurchfluss Q_{Pump}^N [m³/s] im Pumpbetrieb.

$$Q_{Pump}(t) = Q_{Pump}^N \cdot u_{Pump}(t) \quad (2)$$

Der Zusammenhang zwischen dem Durchfluss und der produzierten Leistung einer Turbine wird in dieser Studie als linear betrachtet. Der Durchfluss wird durch die Leistung P_{Turb} [MW] und einen konstanten Umrechnungsfaktor d_{Turb} [m³/(MW·s)] bestimmt.

$$Q_{Turb}(t) = P_{Turb}(t) \cdot d_{Turb} \quad (3)$$

Der Umrechnungsfaktor wird durch die Werte des maximalen und minimalen Durchflusses (Q_{Turb}^{max} und Q_{Turb}^{min} [m³/s]) sowie der maximalen und minimalen Leistung (P_{Turb}^{max} und P_{Turb}^{min} [MW]) bestimmt.

$$d_{Turb} = \frac{Q_{Turb}^{max} \cdot Q_{Turb}^{min}}{P_{Turb}^{max} \cdot P_{Turb}^{min}} \quad (4)$$

Die regulierbare Leistungsabgabe im Turbinenbetrieb wird in zwei Bereiche unterteilt. Die Leistung P_{Turb}^{plan} [MW] wird für vorab geplante Abrufe durch die Day-ahead-Vermarktung genutzt, während P_{Turb}^{flex} [MW] flexibel für die Bereitstellung von Regelleistung zur Verfügung

steht. Die Turbinenleistung im Betrieb muss innerhalb ihrer spezifischen Leistungsgrenzen gehalten werden, wobei u_{Turb} [-] den An/Aus-Zustand des Turbinenbetriebs darstellt.

$$P_{Turb}(t) = P_{Turb}^{plan}(t) + P_{Turb}^{flex}(t) \quad (5)$$

$$u_{Turb}(t) * P_{Turb}^{min} \leq P_{Turb}(t) \leq u_{Turb}(t) \cdot P_{Turb}^{max} \quad (6)$$

Durch den fixen Arbeitspunkt im Pumpbetrieb ergibt sich die Pumpleistung aus dem An/Aus-Zustand der Pumpe u_{Pump} [-]. Somit entspricht die generierte Leistung P_{Pump} [MW] der Nennleistung P_{Pump}^N [MW] oder der Pumpbetrieb ist aus.

$$P_{Pump}(t) = u_{Pump}(t) \cdot P_{Pump}^N \quad (7)$$

Der hydraulische Maschinensatz ist konstruktionsbedingt entweder ausgeschaltet, im Turbinier- oder im Pumpbetrieb aktiv. Diese Funktionsweise wird durch die folgende Gleichung in der Modellierung repräsentiert.

$$u_{Turb}(t) + u_{Pump}(t) \leq 1 \quad (8)$$

Der Start/Stop-Vorgang in einem Zeitschritt t wird durch die Einführung von Variablen für Start v_{Turb}, v_{Pump} [-] und Stop w_{Turb}, w_{Pump} [-] charakterisiert. Diese Variablen werden durch die An/Aus-Zustände ermittelt.

$$u_i(t) - u_i(t-1) = v_i(t) - w_i(t-1), \quad \forall i \in \{Turb, Pump\} \quad (9)$$

Die Leistungskapazität für die Bereitstellung von Regelleistung, sowohl positiv als auch negativ, wird durch den Abstand des geplanten Leistungspunktes zu den Leistungsgrenzen der Turbine bestimmt. Dieser Abstand bestimmt die verfügbare Leistung für die Regelleistungsbereitstellung. Es ist wichtig zu beachten, dass die für den Regelenenergiemarkt vermarktete Leistungskapazität über einen Zeitraum von vier Stunden konstant bleiben muss, gemäß der Produktstruktur von SRL.

$$C_{SRL}^{Pos}(t) \leq u_{Turb}(t) \cdot P_{Turb}^{max}(t) - P_{Turb}^{plan}(t) \quad (10)$$

$$C_{SRL}^{Neg}(t) \leq P_{Turb}^{plan}(t) - u_{Turb}(t) \cdot P_{Turb}^{min}(t) \quad (11)$$

Die effektive Bereitstellung von Regelleistung, angegeben als P_{SRL}^{Pos} und P_{SRL}^{Neg} [MW], entspricht den Abrufsignalen S_{SRL}^{Pos} [MW] und S_{SRL}^{Neg} [MW], solange sie die vorhandene Leistungskapazität nicht übersteigen. Falls der Abruf über diese Kapazität hinausgeht, wird weiterhin nur die Menge an Leistung bereitgestellt, die durch die vorgehaltene Kapazität limitiert ist.

$$P_{SRL}^j(t) \leq \min(S_{SRL}^j(t), C_{SRL}^j(t)), \quad \forall j \in \{Pos, Neg\} \quad (12)$$

Die Parameter des PSW wurden von der realen Anlage in Herdecke mit einem vergleichbaren Aufbau wie in der Modellierung entnommen [RW24b]. Fehlende Werte des realen PSW wurden basierend auf der Größenordnung von Werten von PSW Modellen aus der Literatur geschätzt [CPG16]. Die Angaben für die Kosten von Starts wurden ebenfalls aus der Literatur entnommen und mit der installierten Leistung skaliert [PS16]. Die entsprechenden Werte sind in der folgenden Tabelle aufgeführt.

Parameter	Wert	Parameter	Wert
P_{Turb}^{max}	162 MW	Q_{Pump}^N	101,7 m ³ /s
P_{Turb}^{min}	58,8 MW	V_{PSW}^N	1.530.000 m ³
Q_{Turb}^{max}	110 m ³ /s	c_{Turb}	541,1 €
Q_{Turb}^{min}	47,3 m ³ /s	c_{Pump}	1374,6 €
P_{Pump}^N	153,6 MW	-	-

Tab. 1: Parameter und ihre Werte im Modell des PSW.

2.2 Batteriespeicher zur Bereitstellung von Primärregelleistung

Der Ladezustand der Batterie, ausgedrückt als SoC_{BS} [%], für einen bestimmten Zeitschritt wird analog zum Füllstand eines Speicherbeckens berechnet. Dieser Wert hängt von den Leistungsabflüssen, dem Ladezustand des vorherigen Zeitschrittes, der Speicherkapazität E_{BS}^N [MWh] und t_h für die Zeitschrittlänge [h] ab. Die Leistung des Batteriespeichers wird in Leistungsbezug P_{BS}^{Pos} [MW] und Leistungsabgabe P_{BS}^{Neg} [MW] unterteilt, da der Wirkungsgrad η_{BS} [-] des Systems je nach Energiefluss unterschiedlich berücksichtigt werden muss.

$$SoC_{BS}(t) = \frac{[P_{BS}^{Neg}(t-1) \cdot \eta_{BS} - P_{BS}^{Pos}(t-1)/\eta_{BS}] \cdot t_h}{E_{BS}^N} + SoC_{BS}(t-1) \quad (13)$$

Die beiden Leistungsbereiche sind mit den An/Aus-Zuständen u_{Pos} [-] und u_{Neg} [-] ausgestattet. Sie können Werte zwischen null und der Nennleistung des Speichers P_{BS}^N [MW] annehmen, dürfen jedoch analog zu den Zuständen des PSW nicht gleichzeitig aktiv sein.

$$0 \leq P_k(t) \leq P_{BS}^N \cdot u_k(t), \quad \forall k \in \{Pos, Neg\} \quad (14)$$

$$u_{Pos}(t) + u_{Neg}(t) \leq 1 \quad (15)$$

Durch die zyklische Lebensdauer CL [-] ist es möglich, die zyklische Alterung zu berechnen. Mit der zusätzlichen Berücksichtigung einer konstanten kalendarischen Alterung CA [-]

lässt sich die Veränderung des Alterungszustands der Batterie ΔSoH [%] für jeden einzelnen Zeitabschnitt bestimmen [SS23].

$$\Delta SoH(t) = \frac{[P_{BS}^{Pos}(t) + P_{BS}^{Neg}(t)] \cdot t_h}{2 \cdot E_{BS}^N \cdot CL} + CA \quad (16)$$

Die initialen Kosten des Batteriespeichers (c_{BS} in €/MWh), zusammen mit der Entladetiefe DoD [%], ermöglichen die Berechnung der Alterungskosten c_{SoH} [€/‰] [SS23].

$$c_{SoH} = \frac{c_{BS} \cdot E_{BS}^N}{DoD} \quad (17)$$

In der Vermarktung der Leistungskapazität für PRL der Batterie C_{PRL} [MW] wird gewährleistet, dass die bereitgestellte Leistung zu jedem Zeitpunkt die Nennleistung des Batteriespeichers nicht überschreitet. Zusätzlich wird während des Betriebs des Speichers durch Gl. 18 sichergestellt, dass das 15-Minuten-Kriterium eingehalten wird [5022].

$$25\% \leq SoH(t) \leq 75\% \quad (18)$$

Die tatsächlich bereitgestellte PRL P_{PRL} [MW] ergibt sich gemäß den Vorschriften aus der Abweichung der Netzfrequenz Δf [Hz] von ihrem Sollwert von 50 Hz, wie in der folgenden Gleichung dargestellt.

$$P_{PRL}(t) = \begin{cases} C_{PRL}(t), & \forall \Delta f > 0,2 \text{ Hz} \\ C_{PRL}(t) \cdot \frac{\Delta f(t)}{0,2 \text{ Hz}}, & \forall 0,01 \text{ Hz} \leq |\Delta f| \leq 0,2 \text{ Hz} \\ 0, & \forall |\Delta f| \leq 0,01 \text{ Hz} \\ -C_{PRL}(t), & \forall \Delta f < -0,2 \text{ Hz} \end{cases} \quad (19)$$

Die Parameter des BS wurden entsprechend dem realen Speicher in Herdecke nachgebildet [RW24a], wobei fehlende Werte durch entsprechende Daten aus der Literatur ersetzt wurden [SS23]. Eine Aufstellung sämtlicher verwendeter Parameter ist in der folgenden Tabelle zu finden:

Parameter	Wert	Parameter	Wert
P_{BS}^N	7 MW	DoD	80 %
E_{BS}^N	7 MWh	CL	3500
η_{BS}	92,7 %	CA	1 %
c_{BS}	276.000 €/‰	-	-

Tab. 2: Parameter und ihre Werte im Modell des BS

2.3 Koordination beider Systeme

Um eine koordinierte Nutzung beider Speichersysteme zu gewährleisten, erfolgt deren Verbindung über eine Leistungsbilanz. Diese Bilanz ermöglicht es, die tatsächlichen Abrufe

von PRL und SRL durch die flexible Leistung des PSW sowie die Leistung des BS gemeinsam zu decken.

$$P_{PRL}(t) + P_{SRL}^{Pos}(t) - P_{SRL}^{Neg}(t) = P_{Turb}^{flex}(t) + P_{BS}^{Pos}(t) - P_{BS}^{Neg}(t) \quad (20)$$

2.4 Zielfunktion der Optimierung

Die verwendete Zielfunktion, welche den Gesamtertrag G_{total} [€] durch den Betrieb des Systems repräsentiert, wird in Gl.21 dargestellt. Die erste Zeile dieser Funktion veranschaulicht die realisierten Erträge am Day-ahead-Markt, welche sich aus der verkauften Energie und dem Marktpreis c_{DA} [€/MWh] zusammensetzen. Die Erlöse aus der Bereitstellung von PRL in der zweiten Zeile ergeben sich aus der vorgehaltenen Kapazität und dem entsprechenden Kapazitätspreis c_{PRL} [€/MW]. Die Einnahmen aus der Bereitstellung von SRL ergeben sich ähnlich zu denen der PRL, wobei jedoch zusätzliche Erträge durch die tatsächlich abgerufene Energie entstehen. Daher umfasst die Zielfunktion neben Kapazitätspreisen $c_{L,SRL}^{Pos}$, $c_{L,SRL}^{Neg}$ [€/MW] auch Arbeitspreise $c_{A,SRL}^{Pos}$, $c_{A,SRL}^{Neg}$ [€/MWh] für positive und negative SRL, wie in der dritten und vierten Zeile dargelegt wird. Die Kosten im Zusammenhang mit dem Verschleiß des PSW ergeben sich aus der Anzahl der Starts und den spezifischen Startkosten für Pump- c_{Pump} [€] und Turbinierbetrieb c_{Turb} [€]. Die Kosten für die Alterung des Batteriespeichers entspringen der Veränderung des Alterungszustands und den entsprechenden Alterungskosten.

$$\begin{aligned} Max \ G_{total} = & \sum_t [P_{Turb}^{Plan}(t) - P_{Pump}(t)] \cdot t_h \cdot c_{DA}(t) \\ & + \sum_t C_{PRL}(t) \cdot c_{PRL}(t) \\ & + \sum_t C_{SRL}^{Pos}(t) \cdot c_{L,SRL}^{Pos}(t) + C_{SRL}^{Neg}(t) \cdot c_{L,SRL}^{Neg}(t) \\ & + P_{SRL}^{Pos}(t) \cdot t_h \cdot c_{A,SRL}^{Pos}(t) + P_{SRL}^{Neg}(t) \cdot t_h \cdot c_{A,SRL}^{Neg}(t) \\ & - \sum_t v_{Turb}(t) \cdot c_{Turb} + v_{Pump}(t) \cdot c_{Pump} \\ & - \sum_t \Delta SoH(t) \cdot c_{SoH} \end{aligned} \quad (21)$$

3 Ergebnis

Im Zuge der Optimierung wurde ein effizienter Einsatzplan für beide Speichersysteme über den betrachteten Zeitraum erarbeitet. Es sei jedoch darauf hingewiesen, dass dieser Plan unter der Annahme einer fehlerfreien Preisprognose erstellt wurde. In der Realität könnten die tatsächlich erzielten Erträge von der Prognose abweichen, da eine perfekte Vorhersage nicht

möglich ist. Die resultierenden Einsatzpläne beider Speichersysteme im unkoordinierten Betrieb werden in Abbildung 2 für die ersten 5 Tage dargestellt. Während das eigenständige PSW Regelleistung beim Turbinieren bereitstellt, wobei sowohl positive als auch negative SRL abgerufen wird, liefert der Batteriespeicher kontinuierlich PRL und variiert dabei seine Leistung zwischen +/- 2 MW. Die Leistungsabrufe liegen somit deutlich unterhalb der maximalen Nennleistung des Batteriespeichers. Das Ergebnis der ersten fünf Tage für das koordinierte System wird in Abbildung 3 veranschaulicht. Der Einsatz des PSW zeigt kaum Veränderungen im Vergleich zum eigenständigen Einsatz. Dies ist größtenteils auf die relativ hohe Leistungskapazität des Batteriespeichers zurückzuführen, wodurch der Einfluss der Koordination auf die Betriebsweise des PSW gering ausfällt. Hingegen sind im Einsatz des Batteriespeichers deutlichere Unterschiede erkennbar. Es treten mehrere Zeitperioden auf, in denen die Leistung des Batteriespeichers null ist und somit die PRL-Abrufe durch das PSW abgedeckt werden. Zudem lassen sich markante Leistungsspitzen identifizieren, die auf eine intensivere Nutzung der Kapazitäten des Batteriespeichers hinweisen. Die Gesamterträge sowie die einzelnen Ertragsquellen sind in Tabelle 3 dargestellt. Durch die Koordination beider Speichersysteme konnte in dieser Simulationsstudie eine signifikante Steigerung der Gesamterträge um 12,35 % erzielt werden. Diese Ertragssteigerung spiegelt sich in jedem Teilsegment der Erlösstruktur wider, während gleichzeitig die Kosten durch Verschleiß des PSW und die Alterung des BS durch die Koordination gesenkt werden können. Die gesteigerte Wirtschaftlichkeit resultiert aus einer verbesserten Marktteilnahme und Effizienz, die bei der koordinierten Einsatzplanung unter Berücksichtigung des Verschleißes erzielt wird.

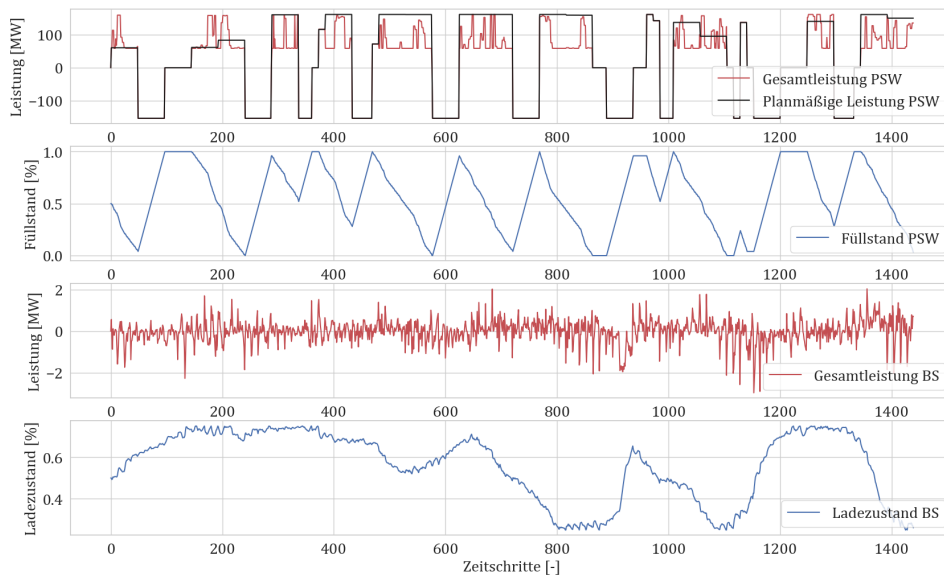


Abb. 2: Optimierter Einsatz beider Speichersysteme im eigenständigen Einsatz für die ersten fünf Tage der untersuchten Zeitperiode.

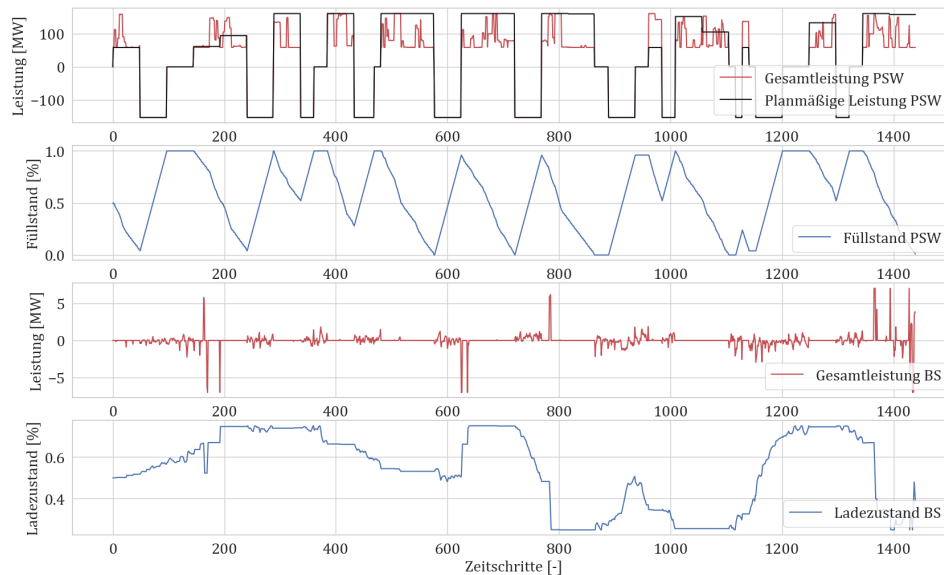


Abb. 3: Optimierter Einsatz beider Speichersysteme im koordinierten Einsatz für die ersten fünf Tage der untersuchten Zeitperiode.

Ertragsquelle	unkoordiniert	koordiniert	Unterschied
Gesamtertrag	38,699 Mio €	43,478 Mio €	+ 12,35 %
Day-ahead-Vermarktung	31,704 Mio €	34,051 Mio €	+ 7,40 %
PRL-Vermarktung	0,737 Mio €	0,759 Mio €	+ 2,95 %
SRL-Vermarktung	8,092 Mio €	10,249 Mio €	+ 27,89 %
Verschleiß/Alterung	- 1,835 Mio €	- 1,681 Mio €	- 8,39 %

Tab. 3: Erträge beider Einsatzstrategien der Speichersysteme.

4 Fazit

In der vorliegenden Simulationsstudie wurde die Effektivität des koordinierten Einsatzes eines PSW in Verbindung mit einem BS untersucht und mit eigenständigen Betriebsmodellen verglichen. Die Simulation basiert auf detaillierten Modellen für PSW und BS, die auf MILP-Problemen beruhen und in dieser Arbeit ausführlich erläutert wurden. Die Ergebnisse der Simulation zeigen deutlich, dass die koordinierte Betriebsführung zu einem signifikant höheren Ertrag von 12,35 % führt, was einen klaren Mehrwert darstellt. Diese Studie liefert wichtige Erkenntnisse über die technischen und wirtschaftlichen Potenziale der koordinierten Betriebsführung von PSW und BS und legt einen Grundstein für zukünftige Forschungen auf diesem Gebiet. Für weiterführende Untersuchungen wäre die Berücksichtigung von Vorhersagefehlern und die Entwicklung robuster Vermarktungsstrategien sinnvoll, wodurch

die Annahme einer perfekten Vorhersage fallen gelassen werden könnte. Zusätzlich könnte die Modellierung durch eine detailliertere Betrachtung komplexer technischer Aspekte wie der Turbinenkennlinie und dem Alterungsverhalten des BS erweitert werden. Dies könnte durch die Einbeziehung nicht linearer Mechanismen mittels linearer Approximation oder die Umstellung des Modells auf ein Mixed-Integer Nonlinear Programming (MINLP)-Problem erfolgen.

Literaturverzeichnis

- [5022] 50Hertz Transmission GmbH, Amprion GmbH, TenneT TSO GmbH, TransnetBW GmbH: Präqualifikationsverfahren für Regelreserveanbieter (FCR, aFRR, mFRR) in Deutschland ("PQ-Bedingungen"), 2022.
- [BH16] Braun, S.; Hoffmann, R.: Intraday optimization of pumped hydro power plants in the german electricity market. *Energy Procedia* 87, S. 45–52, 2016.
- [BSB19] Bahner, L.; Schreider, A.; Bucher, R.: Batteries and pumped-hydro: Pooling for synergies in the frequency response provisioning. In: *The 13th International Renewable Energy Storage Conference 2019 (IRES 2019)*. Atlantis Press, S. 109–118, 2019.
- [By21] Bynum, M. L.; Hackebeil, G. A.; Hart, W. E.; Laird, C. D.; Nicholson, B. L.; Siirola, J. D.; Watson, J. P.; Woodruff, D. L. et al.: *Pyomo-optimization modeling in python*. Springer, 2021.
- [CPG16] Chazarra, M.; Pérez-Díaz, J. I.; García-González, J.: Optimal energy and reserve scheduling of pumped-storage power plants considering hydraulic short-circuit operation. *IEEE Transactions on Power Systems* 32 (1), S. 344–353, 2016.
- [Gu23] Gurobi Optimization LLC: *Gurobi Optimizer Reference Manual*, 2023.
- [In23] International Energy Agency (IEA): *World Energy Investment 2023*, 2023.
- [PS16] Pellinger, C.; Schmid, T.: *Verbundforschungsvorhaben Merit Order der Energiespeicherung im Jahr 2030. Teil 2: Technoökonomische Analyse funktionaler Energiespeicher [Research Project on Merit Order of Energy Storage in the Year 2030. Part 2: Techno-Economic Analysis of Functional Energy Storage]*, 2016.
- [RW24a] RWE AG: *Batteriespeicher Herdecke*, <https://www.rwe.com/der-konzern/laender-und-standorte/batteriespeicher-herdecke/>, 2024.
- [RW24b] RWE AG: *Pumpspeicherkraftwerk Herdecke*, <https://www.rwe.com/der-konzern/laender-und-standorte/pumpspeicherkraftwerk-herdecke/>, 2024.
- [SK12] Simbolotti, G.; Kempener, R.: *Electricity Storage Technology Brief*. In: *International Renewable Energy Agency (IRENA) und International Energy Agency (IEA)*, 2012.
- [SS23] Schmidt, O.; Staffell, I.: *Monetizing Energy Storage: A Toolkit to Assess Future Cost and Value*. Oxford University Press, 2023.
- [Ta] Tadayon, L.; Meiers, J.; Ibing, L.; Erdelkamp, K.; Frey, G.: *Coordinated Dispatch of Battery and Pumped Hydro Energy Storage utilizing hydraulic short circuit*, unpublished.
- [VD14] Van Rossum, G.; Drake Jr, F. L.: *The python language reference*. Python Software Foundation, 2014.

„Ulrich Jumar, Christian Diedrich (Hrsg.):
EKA 2024 - Entwurf komplexer Automatisierungssysteme, 18. Fachtagung“

Data-Driven Soft Sensors for Electrical Machines

Nada Sahlab¹, Arzam Kotriwala¹, Andrew Habib¹ and Victor Mukherjee²

Abstract: As electrical machines are widespread in industrial automation, operating them efficiently has significant potential to improve sustainability. Due to the complexity of electrical machines, obtaining direct measurement of energy consumption is challenging and cost intensive. Soft sensors are useful in inferring variables using available measurements in industrial processes. The data-driven approach to developing soft sensors requires a sufficiently large and diverse training dataset. Given the high cost to obtain voluminous sensor data, turning to simulation data as an additional data source is less expensive, although possibly inaccurate. With this motivation, we explore the need and benefit of combining measurement data from intelligent sensors with electrical machine simulation data for building soft sensors. We present an approach to leverage both, sensor measurements and simulation data to develop a soft sensor for energy efficiency. The soft sensor implementation results for an induction motor support the feasibility of the approach.

Keywords: Electrical Machine, Soft Sensor, Simulation, Data Augmentation, Machine Learning.

1 Introduction

Electrical machines are widespread in industrial automation systems – they play a critical role in the overall performance, cost-effectiveness, and energy efficiency of industrial processes. Three phase induction machines are used in various industrial applications due to their simple design and reliability. They provide mechanical power from electrical power to drive loads. As electrical machines need to be supplied with high amounts of energy and with the goal of increasing sustainability, strict regulations for transparency regarding energy efficiency for electrical machines are required. Operating motors efficiently has significant potential to improve sustainability, especially since around 60% of installed motors operate below their rated power [Ch16]. According to one estimate, the adoption of high-efficiency motor systems could cut global electricity consumption by up to 7% [FE11]. With the increased installation and application of electrical machines in the future, energy efficiency becomes increasingly important. In fact, energy costs (as opposed to purchase costs) make up most of a machine’s lifecycle costs. So, providing more transparency on energy efficiency helps fulfill both sustainability and economic goals.

To ensure efficient operation of induction machines, dynamic modeling approaches are needed to assess and improve their performance. Electrical machines can be modeled

¹ ABB Corporate Research Center, 68309 Mannheim Germany, first.lastname@de.abb.com

² ABB Motors and Generators, 00380 Helsinki Finland, first.lastname@fi.abb.com

using numerical or analytical approaches, each offering a trade-off between computational effort and accuracy. Using such models, machine parameters can be derived to estimate the output power and overall efficiency from the supplied electrical input power.

Complementary to modeling approaches, monitoring approaches are needed to include sensed real values in the estimations. Within the vision of Industry 4.0, miniaturized sensors with communication and processing modules (intelligent sensors) have enabled the acquisition of data and analysis of industrial assets during operation. However, continuously measuring some variables using intelligent sensors is not always feasible due to interference with critical processes or complex machine construction. For instance, measuring the mechanical power of induction machines is not directly possible and usually relies on estimation approaches based on measured values such as flux, speed, or vibration. Additionally, sensors used to measure these values are placed in industrial environments and therefore may be subject to faults thereby introducing inaccuracy and decreasing their reliability.

In this regard, soft sensors or virtual sensors are useful in inferring variables based on available measurements in industrial processes. The data-driven approach to developing soft sensors requires a training dataset that is sufficiently large and diverse e.g., encompassing the range of operation of an electrical machine. Given that measured data can be costly to collect in large amounts, it is natural to turn to other sources of data, e.g., simulation data, which is much less expensive to collect but might be inaccurate depending on model used to represent the machine.

With this motivation, this paper explores the need and benefit of combining measurement data from intelligent sensors with data generated using electrical machine simulation for the purpose of building useful soft sensors for electrical machines. In our work, we use data assimilation to model differences between measurements and simulation followed by data augmentation to generate additional corrected simulation data. In the context of soft sensors, we can leverage data augmentation techniques in several scenarios, e.g.:

- To compensate for inaccurate or insufficient sensor measurements,
- Where direct measurement is infeasible due to high intrusiveness or costs e.g., measuring the output torque of an electrical motor, which can either require costly sensors, commonly not present in industrial settings, or high intrusiveness.

The remainder of the paper is organized as follows. Section 2, presents the related literature on electrical machine modeling approaches for energy efficiency estimations and soft sensors for condition monitoring and performance optimization of induction machines. Section 3 presents our approach and Section 4 discusses the results. Finally, Section 5 concludes the paper and discusses future work.

2 Related Work

2.1 Electrical Machines Modeling

To monitor and optimize the performance of electrical machines, a modeling approach is needed. Generally, electrical machines can be modeled in a numerical or analytical manner. Numerical methods, e.g., finite element method, are classified as accurate, although computationally expensive. Analytical methods encompass equations based on the machine equivalent circuit.

Simulations implement models of electrical machines as a function of time. For the scope of industrial automation, a simulation model is a virtual representation of one or multiple aspects of a physical asset, which mimics its expected behavior when giving possible inputs to a virtual model. Simulation models differ in type and fidelity level and can be complex and costly to run for industrial assets.

Electrical machines simulators are designed and used to simulate the operations of electrical machines and to collect operational data such as speed, power load, magnetic field, and other data.

A high-fidelity simulation of an electrical machine using finite element method is commonly applied and requires a geometrical model as well as knowledge about its material and properties. The simulation time, given through a number of steps and a given period is transformed into discrete time steps. Furthermore, loading points of the machine can be given as an input to compute output values such as the output electrical values, power or the flux density distribution within the magnetic field. [TKS20]

2.2 Energy Efficiency of Electrical Machines

Energy efficiency of electrical machines is the ratio of electrical input power to the mechanical output power delivered to the shaft to drive applications. There are different approaches to measure energy efficiency of induction motors, mostly consisting of indirect measurements and estimations due to the non-linearities exhibited within the electromagnetic field and machine construction. The nameplate method for example, relies on motor ratings to estimate the machine's efficiency profile [Ar22], whereas indirect measurements of the external flux or speed can be used to derive a linearized speed curve. Estimating different power losses by the machine is a further common approach to estimate its energy efficiency. The equivalent circuit approach uses parameters of the machine's equivalent circuit to calculate its output power given technical data or relying on lab tests. [Sa19]

It can be noted that all mentioned approaches with their variations differ in their accuracy and intrusiveness level. No single approach is collectively deemed best. In

recent years, approaches relying on artificial intelligence and data-driven methods have been increasingly considered.

In [AC23], a combination of a data-driven and physics-based model is used to rate the machine efficiency as part of its Digital Twin. In [Si23], search algorithms are compared and used to estimate parameters of the machine's equivalent circuit.

2.3 Soft Sensors

A soft sensor is an intelligent software module that uses existing physical sensors to generate similar data or derive new information by processing physically measured data using mathematical, physics-based, or data-driven approaches. It can be redundant to the physical sensor thus increasing its reliability, replace the physical sensor in case of high maintenance and calibration cost, or extend its functionality by deriving more information from raw measured data. One example is provided by [Li12], where an industrial soft sensor to measure nitrogen oxides based on process measurements was developed.

Data-driven soft sensors [KGS09], [SG21] apply data-processing and analysis techniques such as machine learning and artificial intelligence to measured quantities, obtained through physical sensors, to derive new data and information which are not directly measurable through the physical sensor.

The idea of data-driven soft sensors is already established. However, a key challenge of solely relying on historical data to train such sensors has been the difficulty in ensuring that a wide range of process states and/or conditions are sufficiently covered. This can be addressed by using adaptive soft sensing techniques [KGG10]. For instance, in our case, we leverage simulation data together with associated measurement data to account for any gaps in coverage of different operating conditions of electric machines.

Related to soft sensors is also the concept of surrogate models. They are simplified representations of complex systems or simulations. A common approach to obtain surrogate models is by fitting machine learning models on the input-output data resulting from simulation systems [TKS20], [TMK22]. Amongst the benefits of using surrogate models is to obtain instantaneous estimations, which would otherwise be more time-consuming. Moreover, they can be employed for various use cases, such as condition monitoring or as part of digital twin models. Surrogate models can be realized in a data-driven, hierarchical, or projection-based manner in which a high-fidelity simulation model is used as basis. Data-driven surrogate models using machine learning models can complement simulations for deriving output values of electrical machines. [TKS20]

3 Approach

We propose an approach that combines the use of sensor measurements and simulation data to develop soft sensors for electrical machines. Figure 1 shows an overview our approach to develop soft sensors for electrical machines. It can be used for different soft sensor applications e.g., speed estimation, motor health or fault detection, and energy efficiency estimation [Sa19].

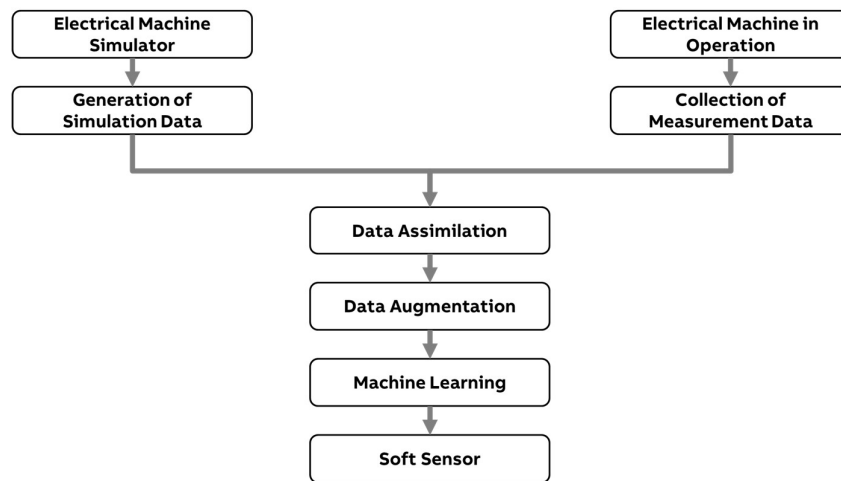


Figure 1: Workflow to develop data-driven soft sensors for electrical machines.

The approach consists of three steps for data collection and processing: gathering real data, generating simulation data, and assimilating and augmenting simulation data. Then, ML models are trained on the final dataset which consists of measurements and augmented simulation data. In the following sections, we provide more details on the proposed steps.

3.1 Collection of Sensor Measurements

We collect operational data, i.e. torque and speed in an experimental setup for an electrical machine. The data is collected using an external sensor attached to the operating motor. Section 4.1 gives more details about the data we collect in our experiments.

3.2 Generation of Simulation Data

We use a proprietary simulator for electrical machines to collect large amounts of simulation data. We generate data for the same operating conditions that were used while collecting the real data.

3.3 Data Assimilation

Based on the available measurement data, we first obtain simulation data for the same operating conditions. We seek to combine the measurement and simulation data such that the simulation data is adjusted to better match the associated real measurements. To this end, we employ “data assimilation” i.e., techniques to combine different sources of information such as integrating a numerical model with observations, with the goal of improving the forecasting capability of a given model [PVS22], [PVS22].

Different data assimilation techniques exist e.g., averaging values, applying priority rules to favor one source over another, or simply selecting the most recent data, fitting of error residuals and Kalman filter (see review of techniques: [PVS22]). We adopt a statistical data assimilation approach that has already been applied to the domain of machine learning for electrical machines (see [Bi24])

In particular, the error between recorded measurements at different loading conditions and associated simulation data for each model feature are computed. A normal distribution is then fitted to estimate the mean (μ) and standard deviation (σ) of the error for each loading point associated with each feature. Consequently, for any loading point, the simulator can be used to generate data, which can then be corrected using an estimate of the error from models fitted on the mean (μ) and standard deviation (σ) values.

3.4 Data Augmentation

Data augmentation refers to the process of creating variations of an existing data set for multiple purposes, e.g., fill gaps in data or increase the amount of data to enable the application of machine learning techniques. Examples of data augmentation approaches include oversampling, and synthetic minority oversampling technique (SMOTE).

In our work, data assimilation is a pre-requisite for data augmentation. Thus, once we have modelled the error between measurement and simulation data using data corresponding to the same operating conditions, we can then, also similar to [Bi24], use data augmentation to create new data that resembles measurement data, without needing to record additional measurements in a real setup.

This is done by generating simulation data for any number of additional operating conditions, and then correcting it to resemble real measurements. Not only is this beneficial because the effort to record additional measurements is eliminated but also

because machine learning modelling typically benefits from datasets that are larger in size and variety of conditions.

3.5 Machine Learning

The data assimilation and augmentation approach helps create an arbitrarily large dataset encompassing a wide range of operating conditions, which is then used to train a machine learning-based soft sensor.

Our work differs from [Bi24] in that we tackle a different problem: predicting load levels of electrical machines as a regression problem whereas they classify the health status of motors and fault conditions such as broken rotors.

4 Results

We developed a soft sensor for motor output power as output power directly correlates to motor efficiency and would otherwise need to be measured intrusively during operation, for instance, attaching a torque meter to the motor shaft to record torque and rotational speed.

4.1 Experimental Setup

To collect measurement data, we use a setup similar to the one presented in a reference work [Bi24] i.e., a “deriving” motor is controlled to act as a constant load torque to our “test” motor. We use an IE4 15 kW, 50 Hz, 400 V delta-connected induction motor and increase the load in steps. At each load level, torque meter measurements are recorded along with corresponding magnetic flux sensor measurements.

4.2 Data Generation, Assimilation, and Augmentation

We increased the load in steps of about 10% from no load to 110% to collect a few measurements samples for each distinct load – we consider different samples at about the same load as a ‘loading group’. For each of these measurements, we use the simulator to generate equivalent data. For both, measurements and simulations, we computed a number of statistical features over a number of windows of the flux data. For each feature, we assimilated the simulation and measurements by computing the error of at each loading group and then estimating the mean (μ) and standard deviation (σ) of the error.

In order to validate our approach, we leave out one loading group at a time and use the remaining available measurement data for data assimilation and augmentation. For the loading group we leave out, we generate simulation data, compute the same statistical

features and then correct them by adjusting the error estimated for the given loading point via data assimilation. This estimation and adjustment of error can be done an arbitrary number of times to augment the dataset. For instance, in our experiments, for every loading group, we have about 3 measurements, so we generated 5 augmented samples i.e., enough to increase the dataset but not so many that the machine learning model is biased by synthetic data in favor of real measurements.

4.3 Model Training and Validation

We trained machine learning models to learn loading from flux values, where loading denotes the ratio of output power to the rated power of the motor. The statistical features we used include mean, standard deviation, minimum, maximum, skewness, and kurtosis. We experimented with different modeling algorithms for regression such as linear models (linear and ridge) and ensemble models (random forest and gradient boosting). The best results we obtained were using gradient boosting regression with statistical features computed on both, X and Y components of flux.

Validation of our modeling was done by performing leave-one-group-out cross-validation [TAJ21], i.e., to determine how well our model would generalize to an unseen loading condition, we held out the measurements of that loading group. However, to realize the benefit of our data assimilation and augmentation, we included the simulation data we generated and corrected for the held-out loading group.

For each held-out loading group, the model predictions for each available measurement sample corresponding to the loading group are compared against ground truth torque meter measurements. The absolute errors are computed for these loading values and predictions and their mean values are plotted (see Figure 2).

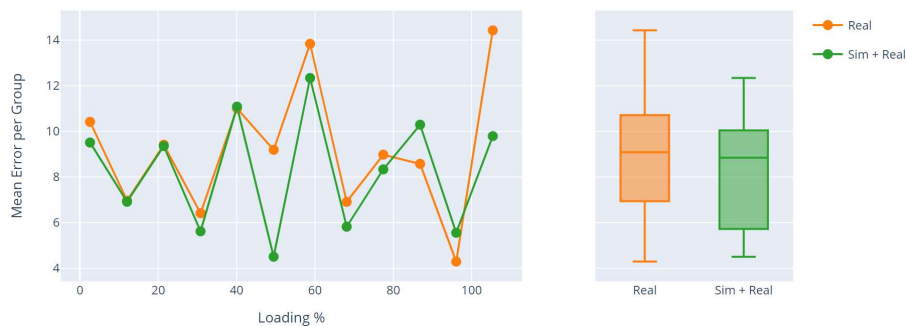


Figure 2 Comparison of modelling error distributions

The results show an overall error reduction when we use measurement data combined with simulation data ('Sim + Real') over only measurement data ('Real'). Most notable of these improvements include loading groups at the extreme ends. This is not surprising as machine learning models are naturally better suited for interpolation as opposed to extrapolation from the underlying training data distribution.

5 Conclusion & Future Work

Motivated by the need of transparency regarding energy efficiency of electrical machines and the challenges around respective data collection, an approach for developing data-driven soft sensors was pursued. The approach follows a 3-step process of collecting and processing data from both simulation and sensor measurements. We implemented the proposed data-driven soft sensor approach in a specific industrial application and our results support the feasibility and the promise of the approach.

Potential directions for further research may include investigating the relationship between data quantity and model performance as well as testing the use of such soft sensors for different applications. For instance, a soft sensor for machine losses may be trained using our approach, which could be used together with an optimizer to compute machine configuration parameters for a given set of operating conditions.

6 Acknowledgement

We would like to extend special thanks to our colleagues: Steven Jannasch, Panagiotis Kakosimos, and Ugur Aydin. Without their support and collaboration, this work would not have been possible.

7 References

- [AC23] Adamou, Adamou; Chakib, Alaoui „Energy efficiency model-based Digital shadow for Induction motors: Towards the implementation of a Digital Twin.“ *Engineering Science and Technology, an International Journal*, vol. 44 (2023).
- [Ar22] Aree, P. „Efficiency-Slip Curve Determination of Induction Motors Using Technical Data .“ *25th International Conference on Electrical Machines and Systems* . 2022.
- [Bi24] Billah, Md; Saberi, Alireza; Hemeida, Ahmed et al. „Generation of Unmeasured Loading Levels Data for Condition Monitoring of Induction Machine Using Machine Learning.“ *IEEE Transactions on Magnetics*, 2024.
- [PVS22] Pandya, Dishant; Vachharajani, Bhasha; Srivastava, Rohit. „A review of data assimilation techniques: Applications in engineering and agriculture.“ *Materials Today: Proceedings* 62 (2022): 7048-7052., 2022.

-
- [SG21] Sun, Qingqiang; Ge, Zhiqiang. „A Survey on Deep Learning for Data-Driven Soft Sensors.“ *IEEE Transactions on Industrial Informatics*, 2021: vol. 17, no. 9, pp. 5853-5866.
- [TAJ21] Tougui, Ilias; Jilbab, Abdelilah; El Mhamdi, Jamal. „Impact of the Choice of Cross-Validation Techniques on the Results of Machine Learning-Based Diagnostic Applications.“ *Healthcare informatics research*. 2021.
- [TKS20] Tahkola, Mikko; Keränen, Janne; Sedov, Dennis et al. „Surrogate Modeling of Electrical Machine Torque Using Artificial Neural Networks.“ *IEEE Access*, 2020.
- [TMK22] Tahkola, Mikko; Mukherjee, Victor; Keränen, Janne. „Transient Modeling of Induction Machine Using Artificial Neural Network Surrogate Models.“ *IEEE Transactions on Magnetics*, 2022.
- [Li12] Liukkonen, Mika; Hälikkää , Eero; Hiltunen, Teri; Hiltunen, Yrjö. „Dynamic soft sensors for NOx emissions in a circulating fluidized bed boiler.“ *Applied Energy*, 2012.
- [Ch16] Choudhary, P. K.; Dubey, S. P.; Tiwari, Bhavana; Dewangan, Babli. „Efficiency Optimization of Induction Motor Drive using Artificial Neural Network.“ 2016 International Conference on Energy Efficient Technologies for Sustainability (ICEETS). Nagercoil: IEEE, 2016.
- [KGS09] Kadlec, Petr; Gabrys, Bogdan; Strandt, Sibylle. „Data-driven Soft Sensors in the process industry.“ *Computers & Chemical Engineering*, 2009.
- [KGG10] Kadlec, Petr; Grbic', Ratko; Gabrys, Bogdan. „Review of adaptation mechanisms for data-driven soft sensors.“ *Computers & Chemical Engineering* 35.1. 1-24., 2010.
- [Si23] Singh, G. „Comparison Between Two Offline Artificial Intelligence Methods for an Efficiency Estimation of In-Service Induction Motors.“ *IETE Journal of Research*, 2023: 1-9.
- [FE11] Fletier, Tobias; W. Eichhammer, Wolfgang. *Energy efficiency in electric motor systems: Technical potentials and policy approaches for developing countries*. Vienna: United Nations Industrial Development, 2011.
- [Sa19] Santos, Vladimir; Eras, Juan; Gutierrez, Alexis; Ulloa, Mario. „Assessment of the energy efficiency estimation methods on induction motors considering real-time monitoring.“ *Measurement*, 2019.

„Ulrich Jumar, Christian Diedrich (Hrsg.):
EKA 2024 - Entwurf komplexer Automatisierungssysteme, 18. Fachtagung“

Modellbasierter Entwurf und Anwendung einer Eigenschaftsregelung zur Automatisierung eines innovativen Drückwalzprozesses

Lukas Kersting¹, Bahman Arian², Julian Rozo Vasquez³, Ansgar Trächtler⁴, Werner Homberg⁵ und Frank Walther⁶

Abstract: Das Konzept der Eigenschaftsregelung mit einer Online-Messung und Rückkopplung der Bauteileigenschaften ist ein derzeit in der Umformtechnik viel diskutiertes – aber bislang kaum umgesetztes Konzept, um den Automatisierungsgrad bei der Bauteilfertigung weiter zu erhöhen und um innovative, neue Produkte mit gradierten Eigenschaften herstellen zu können. Dieser Beitrag soll helfen, diese Lücke zu schließen und befasst sich daher mit den Schritten des Entwurfs bis hin zur Anwendung / Validierung einer neuen Eigenschaftsregelung für das Drückwalzen metastabiler austenitischer Edelstähle. Hierbei handelt es sich um einen Fertigungsprozess für rotationssymmetrische Edelstahlbauteile in kleinen bis mittleren Losgrößen. Der metastabile austenitische Edelstahl ändert während der Verformung seinen α' -Martensitgehalt und damit verbunden u.a. die magnetischen Eigenschaften. Deshalb soll über die Eigenschaftsregelung ein definiertes Einstellen des α' -Martensitgehaltes ermöglicht werden. Die Entwicklung der neuen Eigenschaftsregelung erfordert allerdings aufgrund der Prozesskomplexität ein modellbasiertes Vorgehen. Im Rahmen dieses Beitrags wird daher insbesondere gezeigt, wie mittels des modellbasierten Entwurfs und eines regelungstechnischen Systemmodells die Eigenschaftsregelung ausgelegt werden kann. Zudem beinhaltet der Beitrag erste experimentelle Validierungsergebnisse der zuvor entworfenen Eigenschaftsregelung unter Nutzung eines innovativen mikromagnetischen Softsensorsystems.

Keywords: Drückwalzprozess, Eigenschaftsregelung intelligenter Umformprozesse, modellbasierter Regelungsentwurf, experimentelle Validierung

1 Einleitung

Der wirtschaftliche Wandel zu Industrie 4.0 und cyberphysischen Systemen erfordert eine steigende Automatisierung industrieller Fertigungsprozesse. Insbesondere im Bereich der

¹ Fraunhofer Institut für Entwurfstechnik Mechatronik IEM, Scientific Automation, Zukunftsmeile 1, 33102 Paderborn, lukas.kersting@iem.fraunhofer.de, <https://orcid.org/0000-0002-9733-7528>

² Universität Paderborn, Umformende und Spanende Fertigungstechnik LUF, Warburger Straße 100, 33098 Paderborn, ba.luf@uni-paderborn.de

³ TU Dortmund, Lehrstuhl für Werkstoffprüftechnik (WPT), Baroper Straße 303, 44227 Dortmund, julian.rozo@tu-dortmund.de, <https://orcid.org/0000-0002-5377-1795>

⁴ Heinz Nixdorf Institut, Universität Paderborn, Fürstenallee 11, 33102 Paderborn, ansgar.traechtler@hni.upb.de, <https://orcid.org/0000-0001-9987-1655>

⁵ Universität Paderborn, Umformende und Spanende Fertigungstechnik LUF, Warburger Straße 100, 33098 Paderborn, wh.luf@uni-paderborn.de, <https://orcid.org/0000-0003-4081-1249>

⁶ TU Dortmund, Lehrstuhl für Werkstoffprüftechnik (WPT), Baroper Straße 303, 44227 Dortmund, frank.walther@tu-dortmund.de, <https://orcid.org/0000-0003-2287-2099>

Umformtechnik, einer Gruppe von Fertigungsverfahren, die aufgrund des fehlenden Materialabtrages besonders ressourcenschonend und somit zukunftsstrchtig erscheinen, wird aus diesem Grund in letzter Zeit das Konzept der Eigenschaftsregelung mit einer Online-Messung der Produkteigenschaften inkl. Rckkopplung in den Prozess diskutiert (siehe z.B. [Al16, St24]). Hierdurch ist es mglich, den Automatisierungsgrad nachhaltig zu erhhen, denn bislang erfordern einige Umformprozesse teils zeitintensive, manuelle Fertigungsschritte, um Produkte mit gewnschter Genauigkeit, Qualitt und definierten Eigenschaften herstellen zu knnen. Beispielhaft sei hier die Ermittlung der geeigneten Stellgroen / umformtechnischer Prozessparameter fr neue Produkte im Trial-and-Error-Verfahren genannt. Zudem ist teils manuelle Nacharbeit am gefertigten Produkt notwendig, um gewnschte Toleranzen zu erreichen.

1.1 Drckwalzen metastabiler austenitischer Edelsthle

Zu jenen aus Sicht der Prozessautomatisierung problematischen Umformprozessen gehrt das Drckwalzen metastabiler austenitischer Edelsthle. Hierbei handelt es sich um ein Verfahren zur Herstellung von rotationssymmetrischen Bauteilen aus zylindrischen Rohren. Das Drckwalzen wird generell bereits in der Luft- und Raumfahrtindustrie sowie in der Fahrzeugindustrie zur Herstellung von Strahltriebwerken, Hydraulikzylinder und Antriebswellen eingesetzt. Hier erfolgt die Bauteilfertigung auf CNC-gesteuerten Drckwalzmaschinen, die aus einer oder mehreren, in zwei Achsen verfahrbaren Drckwalzen sowie einem rotierenden Spindeltrieb mit Dorn bestehen [Ru93] (siehe Abb. 1).

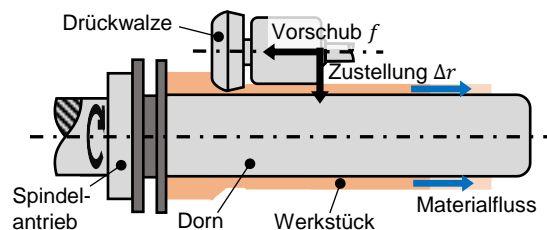


Abb. 1: Funktionsprinzip des Drckwalzens

An der Drckwalzmaschine wird im Verlauf der Bauteilfertigung ein rohrfrmiges Halbzeug auf den Dorn aufgesetzt und in Rotation versetzt. Dann wird die Drckwalze mit einem definierten axialen Vorschub und einer radialen Zustellung entlang des Bauteils bewegt und die Bauteiloberflche berwalzt. Hieraus resultieren eine charakteristische Wanddickenreduktion und eine Lngennderung. Bei den Halbzeugen aus metastabilen austenitischen Edelsthlen (z.B. 1.4307, AISI 304L, X2CrNi18-9) kommt es zudem zur verformungsinduzierten Phasenumwandlung des Gefges von Austenit zu α' -Martensit, verbunden mit einer nderung der Produkteigenschaften, wie z.B. der Hrte, Festigkeit und Magnetik. Die Menge an gebildetem α' -Martensit unterliegt neben den Prozessparametern Vorschub und Zustellung allerdings vielen weiteren Einflussfaktoren wie etwaigen Materialinhomogenitten, der exakten Legierungszusammensetzung der Werkstoffcharge

und der Umgebungstemperatur. Zudem ist auch die tatsächliche Geometrie des gefertigten Bauteils vielen Einflussfaktoren unterworfen, z.B. dem Rückfederungsverhalten der Maschine. Diese Prozesskomplexität macht den Prozess nur schwer reproduzierbar und führt zu Abweichungen, die in der industriellen Umsetzung zurzeit ansatzweise durch eine Kombination aus Erfahrungswissen, Trial-and-Error-Vorgehen und manueller Nachbearbeitung (z.B. Nachschleifen oder Abdrehen des Bauteils) kompensiert werden. Allerdings steht dieses Vorgehen der Automatisierung des Umformprozesses entgegen. Auch ist aufgrund der steigenden Entwicklungszeit die mögliche Produktkomplexität limitiert. Dies verhindert bspw. die Realisierung von innovativen funktionsintegrierten Hochleistungsbauteilen mit gradierten Eigenschaftsstrukturen (z.B. unsichtbare magnetische Bar- und QR-Codes) zur Vorbeugung von Bauteilfälschung. Somit bietet hier die Idee einer Eigenschaftsregelung viele neue Chancen und Möglichkeiten für die Prozessautomatisierung.

1.2 Stand der Forschung für eigenschaftsgerichtete Umformprozesse

Das Konzept einer Eigenschaftsregelung erweitert Umformprozesse um eine übergeordnete Reglerkaskade für die Werkstückeigenschaften der produzierten Bauteile [PAD14, A116]. Diese regelt die programmierte Werkzeugbahn der positionsgeregelten bzw. CNC-gesteuerten Werkzeugmaschine als Stellgröße abhängig einer Eigenschaftsmessung am Bauteil. Hierdurch sollen aufgrund der Prozessstörgrößen vorliegende Eigenschaftsabweichungen im Bauteil direkt online im Prozess erkannt und während der Bauteilfertigung kompensiert werden, um Ausschuss zu reduzieren bzw. manuelle Nachbearbeitung zu vermeiden und somit den Fertigungsprozess effizienter und nachhaltiger zu machen [St24].

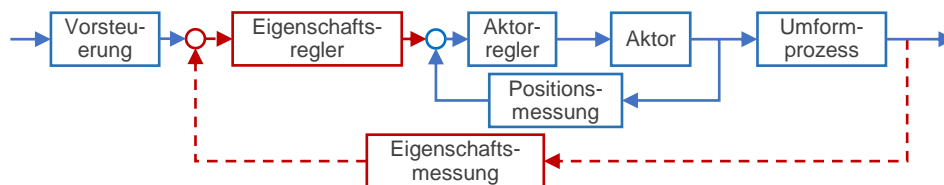


Abb. 2: Erweiterung des Umformprozesses (blau) um eine Eigenschaftsregelung (rot gestrichelt)

Die Regelgrößen der Eigenschaftsregelung umfassen generell sowohl geometrische Eigenschaften (z.B. Wanddicke, Länge) als auch Material- und Werkstückeigenschaften (z.B. Härte). Insbesondere im Bereich der geometrischen Eigenschaften sind bereits einige Anwendungen Stand der Forschung, insb. im Bereich des Tiefziehens. So regeln *Endelt et al.* und *Tricarico et al.* den Flanscheinzug beim Tiefziehen durch eine Online-Anpassung der Niederhaltkraft [ETD13, TP23]. Zudem sind weitere Anwendungen bekannt, bei denen allerdings die Stellgrößenänderung u.a. aufgrund der begrenzten Aktordynamik oder der Messdauer von Bauteil zu Bauteil erfolgt [En17, Fi20]. Eine weitere Anwendung der Geometrie-Eigenschaftsregelung, die zumindest in der Forschung weit verbreitet ist, findet sich im Incremental Sheet Forming. Hier erfolgt in mehreren Literaturquellen unterschiedlicher Autoren eine Regelung der Bauteilgeometrie zur Kompensation der Rückfederung im Prozess (siehe z.B. [HD11, Lu16, FBL23]). Hier muss allerdings

einschränkend ergänzt werden, dass es sich beim ISF um einen stark inkrementellen Prozess handelt mit mehreren sequenziell umgeformten Schichten. In der genannten Literatur erfolgt die Stellgrößenänderung immer von Schicht zu Schicht, unterbrochen durch die Messung mittels 3D-Scanner bzw. Stereokamera mit einer Messdauer- und Datenverarbeitungsdauer von 11 s [HD11] bis hin zu 15 min [Wa23]. Somit handelt es sich hier nur bedingt um Online-Regelung der Eigenschaften im eigentlichen Sinn.

Auch hinsichtlich der Regelung von Material- bzw. Werkstückeigenschaften existieren bereits einige Veröffentlichungen, wenngleich in einem deutlich geringeren Umfang als bei den geometrischen Eigenschaften. *Lafarge et al.* stellen eine Regelung der Härte beim thermomechanischen Ringwalzen vor und validieren das Konzept in Co-Simulation zwischen Python und LS-Dyna [La21]. *Ismael et al.* präsentieren einen Regelungsansatz für die kombinierte Regelung der Eigenspannungen und Krümmung beim Freiformbiegen und zeigen ersten Regelungsergebnisse anhand eines aus der FE-Simulation abgeleiteten Regressionsmodells in MATLAB/Simulink [Is22]. Zudem wird in [Spi23] ein Regelungskonzept für das Stanzlochwalzen vorgestellt. Hier ist – genauso wie bei den anderen genannten Veröffentlichungen – keine experimentelle Umsetzung und Validierung an einer Werkzeugmaschine als Stand der Forschung bekannt. Ausnahme ist der Prozess des Kaltwalzens eines Stahlbandes. In [Sc21] validieren *Schulte et al.* eine modellprädiktive Regelung der Oberflächenrauheit durch Anpassung des Walzspalts.

Für den Drückwalzprozess, der im Fokus dieses Beitrags steht, existieren Stand heute nur wenige Arbeiten hinsichtlich des Einsatzes von Regelungstechnik. Diese kommt beim Drückwalzen vor allem in Form des Lageregelkreises der CNC-gesteuerten Maschinenaktorik zum Einsatz [Ru93]. Forschungsarbeiten zur Regelung beim Drückwalzen erstrecken sich vor allem auf dem Gebiet der Regelung der Umformkräfte. In [Di92] wird eine adaptive Kraftregelung für das Drücken mit abschließendem Drückwalzprozess entworfen und umgesetzt. In [La24] wird eine Drückwalz-Kraftregelung mittels FE-Simulation entwickelt. Die hierdurch generierten Werkzeugbahnen werden in Form einer Steuerung auch experimentell an der Versuchsanlage validiert, allerdings erfolgt in [La24] keine Validierung im geschlossenen Regelkreis.

Zusammenfassend ist festzuhalten, dass das Konzept der Eigenschaftsregelung zwar bereits für eine Vielzahl von Umformprozessen existiert, aber insb. hinsichtlich der Regelung von Materialeigenschaften selten über das Konzeptstadium hinausgekommen ist und bislang kaum an Werkzeugmaschinen umgesetzt und validiert wurde. Für den Drückwalzprozess – insb. bei Verwendung von metastabilen austenitischen Edelstählen – ist aktuell keine Online-Eigenschaftsregelung bekannt. Um den Stand der Forschung zu ergänzen, wurde von den Autoren dieses Beitrags bereits in [Ke22a] ein Regelungskonzept für das Drückwalzen metastabiler austenitischer Edelstähle vorgestellt.

1.3 Konzept des eigenschaftsgerichteten Drückwalzprozesses

Das Konzept des eigenschaftsgerichteten Drückwalzprozesses metastabiler austenitischer Edelstähle sieht eine kombinierte Regelung der Wanddickenreduktion Δw als geometrische Regelgröße und des verformungsinduzierten α' -Martensitgehalts als zu regelnde Materialeigenschaft vor (Abb. 3).

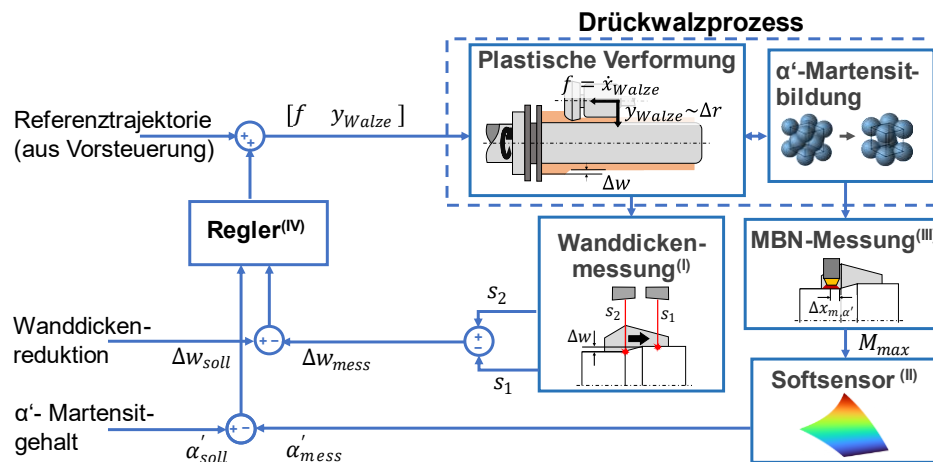


Abb. 3: Konzept des eigenschaftsgerichteten Drückwalzprozesses

Als Stellgröße der Regelung dient die Werkzeugbahn, d.h. der Vorschub $f = \dot{x}_{Walze}$ und die Zustellung Δr bzw. die radialen Walzenposition y_{Walze} . Die Werkzeugbahn wird für die Regelung online während der Bauteilfertigung, d.h. am selben Bauteil und im selben Überlauf, angepasst. Die Anpassung erfolgt hierbei auf Basis eines Vergleichs zwischen Soll- und Istwert der Regelgrößen α' -Martensitgehalt und Wanddickenreduktion Δw . Zur Online-Bestimmung dieser Regelgrößen wird ein bereits in [Ker22] und [Ro24] adaptiertes und validiertes Messkonzept genutzt. Die Messung der Wanddickenreduktion Δw erfolgt hier mittels zwei Laserdistanzsensoren S_1 und S_2 vom Typ OM70 der Fa. Baumer (siehe Abb. 3, I). Diese sind am Maschinensupport befestigt und verfahren somit mit der Drückwalze während des gesamten Umformprozesses. Hierbei ist Sensor S_1 mit ca. 90° Winkelversatz zur Umformzone auf den bereits ungeformten Bereich ausgerichtet, während Sensor S_2 den nicht umgeformten Bauteilbereich vermisst. Die Wanddickenreduktion lässt sich aus der Differenz beider Sensorsignale $\Delta w_{mess} = s_2 - s_1$ berechnen. Die Ermittlung des α' -Martensitgehaltes als zweite Regelgröße erfolgt indirekt mittels eines speziellen, von den Autoren dieses Beitrags in [Ro24] vorgestellten Softsensors (II). Der Softsensor besteht aus einem physikalischen (Hard-)Sensor und einem Softsensormodell. Als physikalischer Sensor wird ein 3MA-II-Sensor (Fraunhofer IZFP, III) genutzt. Dieser ist analog zu den Laserdistanzsensoren am Maschinensupport befestigt und misst kontinuierlich mit einem axialen Versatz $\Delta x_{m,\alpha'} \approx 8$ mm zur Umformzone. Als Messgröße wird für den Softsensor die maximale Amplitude M_{max} des mikromagnetischen

Barkhausen-Rauschens (MBN) genutzt. Ausgehend von M_{max} und dem Vorschub $f = \dot{x}_{Walze}$ erfolgt mittels der Softsensor-Modellgleichung aus [Ro24] die Berechnung des Martensitgehaltes α'_{mes} . Die so ermittelten Regelgrößen α'_{mes} und Δw_{mess} werden anschließend mit dem jeweiligen Sollwert α'_{soll} bzw. Δw_{soll} verglichen. Hierauf basierend erfolgt im Regler (IV) die Berechnung der Stellgrößenänderung.

2 Modellbasierter Entwurf der Eigenschaftsregelung

Für die Auslegung der Regelung wird im Rahmen dieses Beitrags aufgrund der Prozesskomplexität ein modellbasiertes Vorgehen gewählt. Das bedeutet, dass unter Nutzung des Modells die Realisierbarkeit der Regelung geprüft und die Parametrierung des Reglers vorgenommen wird. Dies hat den Vorteil, dass sich die Regelung effizienter auslegen lässt sowie Zeit und Kosten für experimentelle Versuche gespart werden können. Grundvoraussetzung für das modellbasierte Vorgehen ist ein geeignetes Modell der Regelstrecke, welches den vollständigen Regelkreis inkl. der Sensorik berücksichtigt und sich „erträglicher“ Rechenzeit simulieren lässt. Ein solches Modell wurde bereits in [Ke22a] von den Autoren dieses Beitrags vorgestellt und in [Ke23, Ke24] erweitert. Dieses Modell wird in diesem Beitrag als Entwurfsmodell für den Regelungsentwurf genutzt. Aus diesem Grund erfolgt im nachfolgenden Abschnitt eine kurze Zusammenfassung über wesentliche Charakteristika des Modells. Für tiefere Details sei auf die genannte Literatur verwiesen.

2.1 Systemmodell des eigenschaftsgeregelten Drückwalzprozesses

Das regelungstechnische Systemmodell bildet die Prozesskinematik und das nichtlineare Umformverhalten im eigenschaftsgeregelten Drückwalzprozess ab. Es ist als signalflussorientiertes Modell in MATLAB/Simulink implementiert und besteht grundsätzlich aus einer Kombination aus Geometriemodell und empirischen Kennlinien (siehe Abb. 4). Zusammenfassend besitzt das modellierte System insb. zwei Charakteristika:

- Es handelt es sich um ein ortsverteilt System. Dieser Aspekt wird im Modell mittels eines Geometriemodells (2) abgebildet. Die Mantelfläche des Bauteils ist hierbei ortsdiskret über $i = 1 \dots n$ Punkte mit einer Wanddicke w_i und axialen Position z_i approximiert. Die einzelnen Diskretisierungspunkte sind beim Drückwalzen aufgrund der spezifischen Prozesskinematik (helixförmige Walzenbahn) allerdings weder kontinuierlich aktuiert noch sensiert. Auch erfolgt die Messung mit einem örtlichen Versatz $\Delta x_{m,\alpha'}$ zur Umformung. Aus diesem Grund besitzt das Sensormodell (5) und (6) bzw. reale Messglied ein Totzeitverhalten ggü. der Zustandsänderung / Umformung. Die Totzeit T_{tot} aufgrund des axialen Versatzes errechnet sich nach [Ke23] aus dem axialen Sensorversatz $\Delta x_{m,\alpha'}$ und dem Vorschub des Aktors f über:

$$T_{tot} = \frac{\Delta x_{m,\alpha'}}{f} \quad (1)$$

- Die Regelstrecke wird im Modell insb. über die empirischen Teilmodelle (3) und (4) abgebildet. Wie aus Abb. 4 ersichtlich hängen die plastische Verformung und auch die α' -Martensitbildung sowohl von der Zustellung Δr als auch vom Vorschub f ab (im Fall der α' -Martensitbildung zumindest indirekt über Δw). Es handelt sich daher um ein verkoppeltes MIMO-System (Multiple Input Multiple Output) mit je zwei Eingangs- und Ausgangsgrößen.

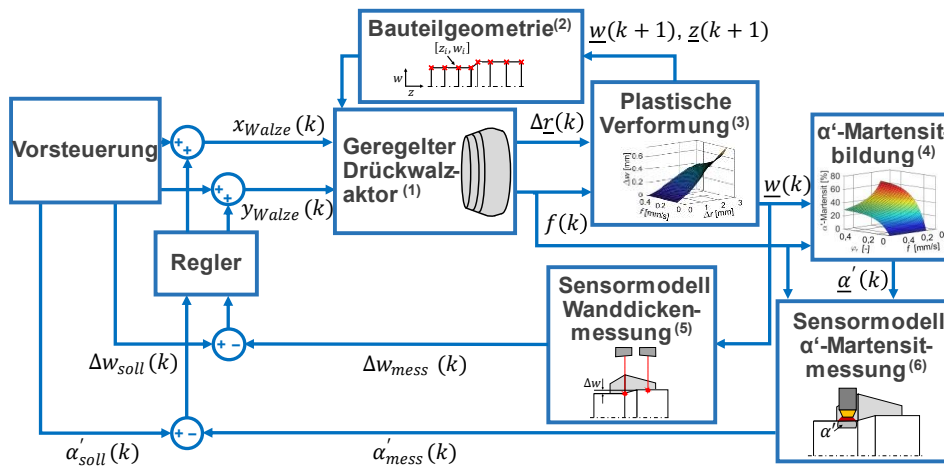


Abb. 4: Regelungstechnisches Modell inkl. Regeleinrichtung (Übersicht)

2.2 Festlegung der Reglerstruktur

Die Reglerstruktur eines MIMO-Systems hängt maßgeblich von dem Maß der Verkopplung zwischen den einzelnen Eingangs- und Ausgangsgrößen des Systems ab. Für das eigenschaftsgeregelte Drückwalzen haben Voruntersuchungen gezeigt, dass die stärksten Kopplungen zwischen Zustellung Δr und Wanddickenreduktion Δw (nachfolgend als nichtlineare Funktion G_{11} bezeichnet) sowie zwischen Vorschub f und Martensitgehalt α' (G_{22}) vorliegen, während die Querkopplung zwischen Δr und α' (G_{21}) bzw. f und Δw (G_{12}) eher gering ist. Ist ein System nur schwach gekoppelt, so kann es nach [Lu20] mutmaßlich über einschleifige Regler unter Vernachlässigung der Querkopplungen geregelt werden. Ein mögliches Indiz für die schwache Kopplung ist ein Kopplungsfaktor $\kappa < 1$ (siehe [Lu20]). Dies ist gemäß Voruntersuchungen für das hier vorliegende System erfüllt. Aus diesem Grund wird die Reglerstruktur für die Eigenschaftsregelung in Form einer sog. dezentralen Regelung (im Sinne der Definition nach [Lu20]) festgelegt (siehe Abb. 5).

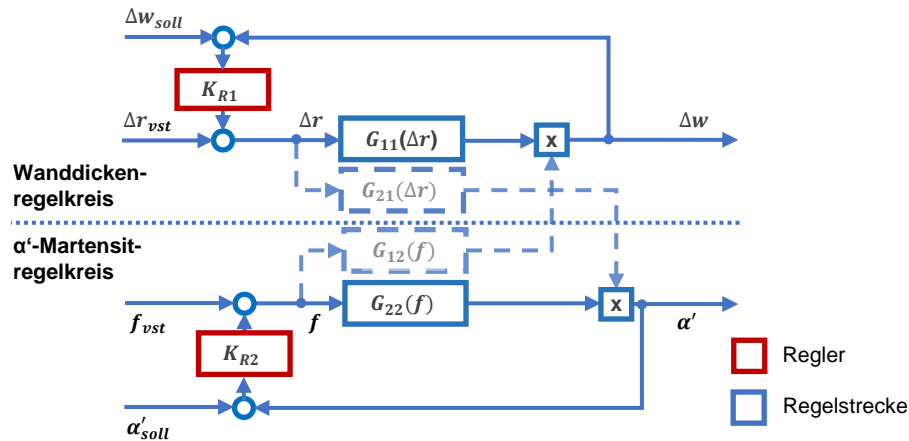


Abb. 5: Regelungsstruktur des Eigenschaftsreglers (Schema), adaptiert nach [Lu20]

Diese besteht aus zwei einschleifigen Regelungen, dem Wanddickenregelkreis und dem α' -Martensitregelkreis. Im Wanddickenregelkreis wird mithilfe der Stellgröße Zustellung Δr die Wanddickenreduktion Δw geregelt, im α' -Martensitregelkreis auf Basis der Stellgröße Vorschub f der Martensitgehalt α' . Jeder Regelkreis besitzt einen Eingrößenregler K_{R1} (Wanddickenregler) und K_{R2} (Martensitregler) zum Einregeln der jeweiligen Regelgröße Δw bzw. α' . Die Querkopplung zwischen den Regelkreisen (G_{12} und G_{21}) wird hierbei lediglich als Störgröße aufgefasst, die durch K_{R1} bzw. K_{R2} mit ausgeregelt wird.

2.3 Regelungsentwurf

Für den Entwurf der Regler K_{R1} und K_{R2} wird die Regelungsstruktur in das regelungstechnische Systemmodell in MATLAB/Simulink implementiert. Als Regler K_{R1} bzw. K_{R2} sind im ersten Ansatz aufgrund ihrer Robustheit ggü. Modellunsicherheiten und der stationären Genauigkeit PI-Regler ausgewählt. Insgesamt sind daher im Rahmen des Entwurfes zwei Parameter pro Regelkreis einzustellen: der Proportional- und Integralanteil für den Wanddickenregelkreis $K_{p,w}$ und $K_{i,w}$ sowie für den α' -Martensitregelkreis $K_{p,\alpha'}$ und $K_{i,\alpha'}$. Um diese Parameter zu ermitteln, sind aus der Literatur ein Vielzahl von Auslegungsverfahren bekannt, wie z.B. das Frequenzkennlinienverfahren, das Einstellen eines Betragsoptimums oder empirische Einstellregeln von *Ziegler* und *Nichols* [Fö22]. Allerdings lassen sich viele Einstellverfahren im Fall des eigenschaftsgeregelten Drückwalzens aufgrund der nichtlinearen Regelstrecke sowie dem komplexen dynamischen Verhalten inkl. der Totzeit zwischen Sensor und Walze nur eingeschränkt nutzen. Ein Verfahren, welches unter den vorhandenen Rahmenbedingungen auch aufgrund der Verfügbarkeit eines schnellen Simulationsmodells als prädestiniert erscheint, ist der Entwurf eines Reglers durch Parameteroptimierung mittels eines Gütemaßes, wie z.B. in [Fö22] dargestellt. Hierbei ist es generell Ziel, die Parameter eines PI-Reglers mittels numerischer Optimierung

so anzupassen, dass eine Gütefunktion J minimal wird. Für die Gütefunktion wird in [Fö22] das zeitliche Integral über den Regelfehler $e(t)^2 = (y_{soll} - y)^2$ mit dem (allgemeinen) Sollwert y_{soll} und dem Istwert y angesetzt, die sog. quadratische Regelfläche. Somit wird die Regelung zeitoptimal eingestellt. Für den im Rahmen dieses Beitrags vorliegenden Anwendungsfall erscheint es allerdings als sinnvoll, nicht den zeitlichen Verlauf des Regelfehlers, sondern den örtlichen Verlauf der quadratischen Regelfläche $e(z)^2$ als Gütemaß zu betrachten (Abb. 6). So ist es doch Ziel der Eigenschaftsregelung in der Umformtechnik, möglichst genaue Bauteile zu produzieren und im speziellen beim eigenschaftsgerichteten Drückwalzen möglichst orts aufgelöste Eigenschaftsstrukturen am Bauteil zu erhalten. Allerdings ist die Totzeit aufgrund des Sensorversatzes bei variierendem Vorschub f gemäß Gl. (1) zeitvariant. Aus diesem Grund ist jedoch nicht zweifelsfrei sichergestellt, dass ein zeitoptimaler Verlauf des Regelfehlers auch einem ortsoptimalen Verlauf entspricht. Daher wird in diesem Beitrag das Optimierungsproblem zur Bestimmung von $K_{p,w}$ und $K_{i,w}$ sowie $K_{p,\alpha'}$ und $K_{i,\alpha'}$ wie folgt definiert:

$$\min J = \int_0^{z_{end}} e(z)^2 dz, \quad (2)$$

wobei z die axiale Koordinate des finalen Bauteils bezeichnet und z_{end} das Ende des umgeformten Bereiches. Zudem wird beim Drückwalzen ein aperiodisches Reglerverhalten angestrebt. Der Regler sollte daher beim Einregeln eines Sollwertsprungs nicht den Sollwert überschreiten und überschwingen, da dies zu einer zu großen plastischen Verformung am Bauteil führt und diese irreversibel ist. Aus diesem Grund wird hier das Optimierungsproblem noch um folgende Nebenbedingung erweitert:

$$-e(z) \leq 0 \quad \forall \{z \in 0 \leq z \leq z_{end}\} \quad (3)$$

Das heißt, an jedem umgeformten Punkt auf dem Bauteil soll der Regelfehler positiv, also der Istwert der Regelgröße kleiner dem Sollwert, sein.

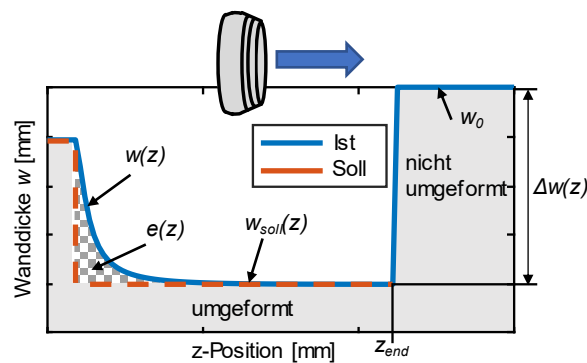


Abb. 6: Fehlerfläche beim Drückwalzen am Beispiel der Wanddicke ($y = w$), Schema

Gl. (2) und (3) stellen ein nichtlineares Optimierungsproblem mit Beschränkung dar, welches in MATLAB mit dem Algorithmus *fmincon* gelöst werden kann. Im Rahmen dieses

Beitrags wird unter Zuhilfenahme des Algorithmus das nichtlineare regelungstechnische Simulationsmodell in MATLAB/Simulink aufgerufen und unter Variation der Reglerparameter (im Fall des Wanddickenregelkreises $K_{p,w}$ und $K_{i,w}$, im Fall des α -Martensitregelkreises $K_{p,\alpha}$ und $K_{i,\alpha}$) so lange simuliert, bis die Funktion minimal ist. Der Algorithmus *fmincon* ist allerdings nur geeignet, um lokale Minima von J zu ermitteln [An09]. Aus diesem Grund wird die Optimierung in 20facher Wiederholung mit verschiedenen, zufällig gewählten Startpunkten wiederholt, um möglichst eine globale Lösung des Optimierungsproblems zu finden. Das Ergebnis ist im Folgenden für beide Regelkreise exemplarisch für jeweils einen Arbeitspunkt der Regelung dargestellt.

Wanddickenregelkreis

Abb. 7 zeigt das Simulationsergebnis des ausgelegten Wanddickenregelkreises am Beispiel einer Sollwanddickenreduktion $\Delta w_{soll} = 0,5$ mm, was bei einer Ausgangswanddicke $w_0 = 4$ mm einer Soll-Wanddicke w_{soll} von 3,5 mm entspricht. Hier wurde als anfängliche Zustellung durch die Vorsteuerung $\Delta r_{vst} = 1$ mm vorgegeben, wobei der Vorschub $f = \dot{x}_{Walze}$ gleichzeitig mit 0,1 mm/s konstant gehalten wurde. An dem genannten Arbeitspunkt führt die Lösung des Optimierungsproblems mit $e(z) = w_{soll}(z) - w(z)$ zu folgenden Reglerparametern: $K_{p,w} = 0,5$ und $K_{i,w} = 0,1062$.

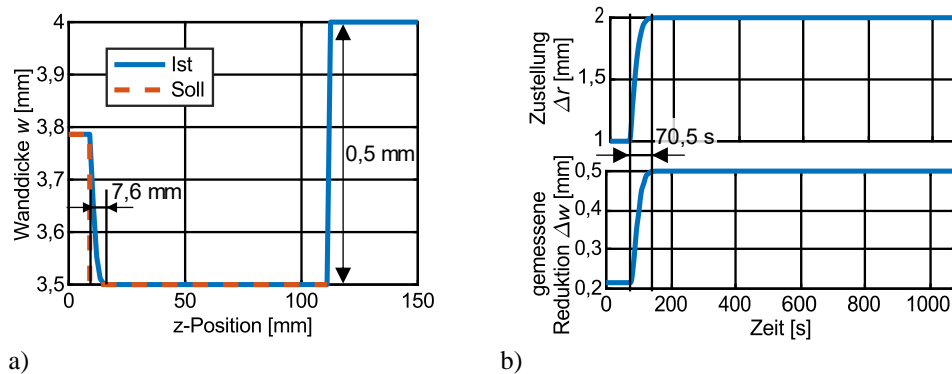


Abb. 7: Simulationsergebnis für die Auslegung des Wanddickenregelkreises

Der dazugehörige Wanddickenverlauf $w(z)$ ist in Abb. 7a genauer dargestellt. Vor Einschalten der Regelung beträgt die Wanddicke $w = 3,786$ mm. Nach dem Sollwertsprung bei $z = 9$ mm wird innerhalb von 7,6 mm der Sollwert w_{soll} erreicht. Die Regelung ist somit stationär genau, was auf den I-Anteil des PI-Reglers zurückzuführen ist. Die Istwanddicke w besitzt wie durch die Nebenbedingung Gl. (3) vorgegeben einen aperiodischen Verlauf ohne Überschwingen des Sollwertes. Somit entspricht das örtliche Verhalten der Regelung der vorgegebenen Solldynamik. In Abb. 7b ist zudem der Stellgrößenverlauf und das simulierte Sensorsignal für die Wanddickenreduktion Δw dargestellt. Die Einregelung der Sollwanddicke erfolgt demnach durch eine Erhöhung der Zustellung um 1 mm auf $\Delta r = 2$ mm innerhalb von 70,5 s, was relativ gut zu dem Einregelweg von

7,6 mm und dem konstanten Vorschub $f = 0,1$ mm/s korreliert. Eventuelle Abweichungen sind u.a. auf die Längenänderung des Bauteils aufgrund der Volumenkonstanz zurückzuführen. Somit lässt sich auch durch Vergleich von Abb. 7a und b festhalten, dass der örtliche Wanddickenverlauf näherungsweise einer örtlichen Projektion des zeitlichen Verlaufes entspricht.

α' -Martensitregelkreis

Abb. 8 zeigt das Simulationsergebnis des ausgelegten α' -Martensitregelkreises am Beispiel eines Soll- α' -Martensitgehaltes $\alpha'_{soll} = 50$ % (Vol.-%). Hier wurde als anfänglicher Vorschub $f_{vst} = 0,25$ mm/s bei gleichzeitig während des gesamten Versuches konstanter Zustellung $\Delta r_{vst} = 2$ mm vorgegeben. An dem genannten Arbeitspunkt führt die Lösung des Optimierungsproblems mit $e(z) = \alpha'_{soll}(z) - \alpha'(z)$ zu folgenden Reglerparametern: $K_{p,\alpha'} = 5,3 \cdot 10^{-3}$ und $K_{i,\alpha'} = 3,25 \cdot 10^{-4}$.

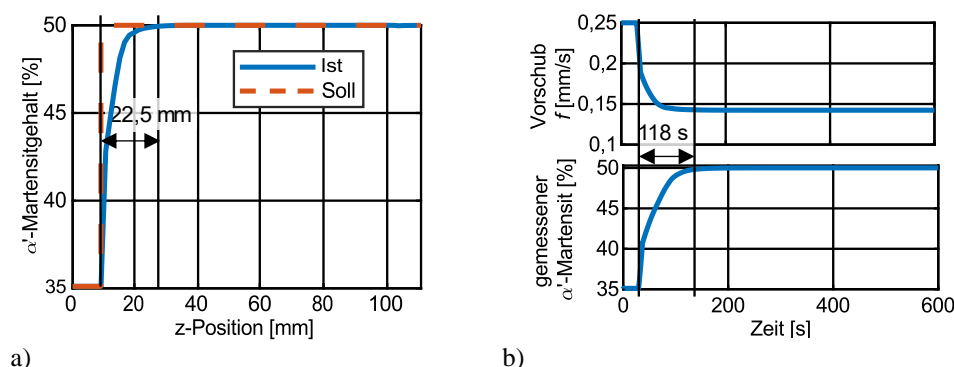


Abb. 8: Simulationsergebnis für die Auslegung des α' -Martensitregelkreises

Der dazugehörige Wanddickenverlauf $\alpha'(z)$ sowie die zeitlichen Verläufe von α' -Martensitmesssignal und Vorschub f sind in Abb. 8a und b genauer dargestellt. Am Bauteilanfang beträgt der α' -Martensitgehalt ca. 35%. Nach dem Sollwertsprung bei $z = 9$ mm wird durch eine Reduktion des Vorschubes von 0,25 mm/s auf 0,14 mm/s innerhalb von 22,5 mm α'_{soll} erreicht, was einer Dauer von ca. 118 s entspricht. Analog zum Wanddickenregelkreis ist das Regelungsverhalten aperiodisch und ohne Überschwingen. Auffällig ist allerdings, dass der optimal parametrisierte Regler beim α' -Martensitregelkreis mit dem o.g. Einregelweg von 22,5 mm deutlich langsamer ist als der nach gleichem Vorgehen ausgelegte Wanddickenregler. Somit wird der maximale α' -Martensitgehalt deutlich später erreicht. Dieses Phänomen lässt sich mit dem axialen Sensorversatz bei der α' -Martensitmessung mit $\Delta x_{m,\alpha'} \approx 8$ mm begründen. Somit liegt nach Gl. (1) beim α' -Martensitregelkreis aufgrund des Versatzes eine deutliche Totzeit vor. Dies führt wiederum, wie allgemein aus der Literatur bekannt, aufgrund der größeren Phasenverschiebung i.d.R. zu einer langsameren Regelung mit geringerer Bandbreite, zumindest wenn in beiden Fällen aperiodisches Verhalten und somit ein ähnlich gedämpftes Verhalten gefordert ist.

3 Experimentelle Validierung

Zur experimentellen Validierung der Eigenschaftsregelkreise wurde das Regelungskonzept auf die Drückwalzanlage PLB 400 (Leifeld) an der Universität Paderborn übertragen. Hierbei handelt es sich um eine Einwalzen-Drückwalzanlage mit hydraulischem Kreuzsupport. Der Kreuzsupport besitzt einen universellen Werkzeughalter, an dem die Drückwalze und über eine Tragekonstruktion die Zusatzsensorik für die Eigenschaftsregelung (3MA-II und Laserdistanzsensoren) befestigt sind. Der 3MA-II-Sensor ist hierbei in Kontakt zur Werkstückoberfläche positioniert, während die Laserdistanzsensoren mit leichtem Versatz in Umfangsrichtung zum 3MA-II Sensor berührungsfrei messen (Abb. 9).

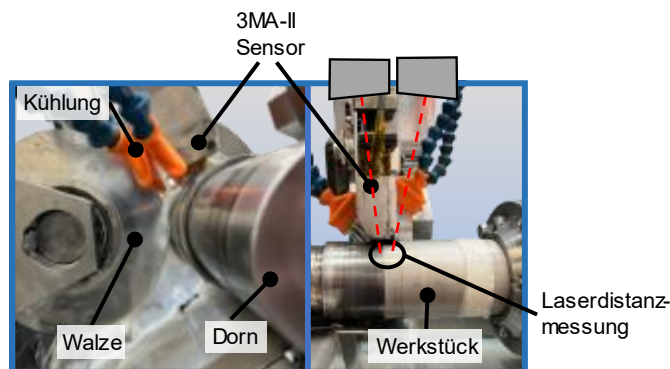


Abb. 9: Umsetzung der Eigenschaftsregelung an der Versuchsanlage

Die Ergebnisse der experimentellen Validierung an der Versuchsanlage sind nachfolgend analog zum modellbasierten Regelungsentwurf separat für den Wanddicken- und Martensitregelkreis und exemplarisch für jeweils einen Arbeitspunkt dargestellt. Beide Versuche wurden bei einer Drehzahl von $n = 30$ U/min durchgeführt.

Wanddickenregelkreis

Abb. 10 zeigt das Versuchsergebnis des Wanddickenregelkreises am Beispiel einer Sollreduktion $\Delta w_{soll} = 0,5$ mm. Hier wurde als anfängliche Zustellung $\Delta r_{vst} = 1$ mm vorgegeben, wobei der Vorschub $f = \dot{x}_{Walze}$ gleichzeitig mit 0,1 mm/s konstant gehalten wurde. Somit ist die Regelungsaufgabe prinzipiell identisch zur Simulation in 2.3. In Abb. 10 (unten) ist die online-gemessene Wanddickenreduktion, gemessen durch die zwei Laserdistanzsensoren, dargestellt. Zu Versuchsbeginn liegt der Messwert für die Wanddickenreduktion Δw bei ca. 0,25 mm. Nach dem Sollwertsprung bei $t = 17$ s kommt es innerhalb von 80 s zu einer Erhöhung der Wanddickenreduktion auf $\Delta w \approx 0,5$ mm durch eine Vergrößerung der Stellgröße Zustellung Δr um ca. 1 mm (Abb. 10 (oben)). Die Regelung ist bezogen auf das gefilterte Messsignal bzw. den Mittelwert stationär genau und stabil. Abweichungen im ungefilterten Signal können u.a. auf umfangsseitig unterschiedliche Ausgangswanddicken, welche bei der genutzten Drehzahl $n = 30$ U/min aufgrund

der begrenzten Aktordynamik nicht ausgeglet werden können, sowie auf Störeinflüsse z.B. durch Oberflächenbeschädigungen am Halbzeug und Vibrationen am Bauteil zurückgeführt werden. Insgesamt aber kann festgehalten werden, dass der Wanddickenregelkreis eine gute Übereinstimmung zur Simulation aufweist. Dies betrifft sowohl die Regelungs-dynamik, welche nur geringfügig langsamer als in der Simulation ist, als auch die notwendige Stellgrößenänderung von Δr zur Einregelung des Sollwertes.

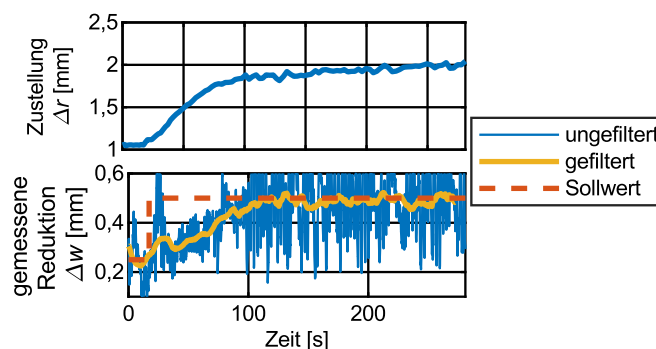


Abb. 10: Validierung des Wanddickenregelkreises

Martensitregelkreis

Abb. 11 zeigt das Versuchsergebnis des α' -Martensitregelkreises am Beispiel eines Soll- α' -Martensitgehaltes $\alpha'_{soll} = 50\%$. Hier wurde als anfänglicher Vorschub $f_{vst} = 0,25$ mm/s bei gleichzeitig durchgängig konstanter Zustellung $\Delta r = 2$ mm vorgegeben.

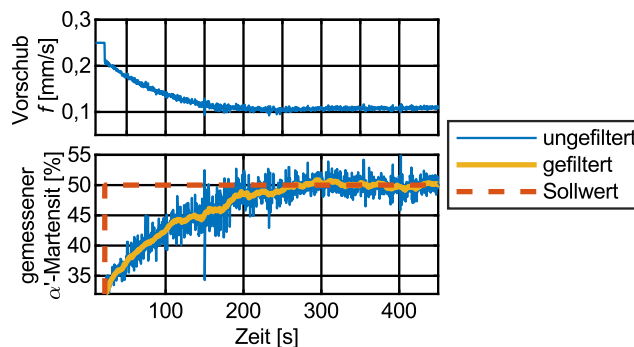


Abb. 11: Validierung des Martensitregelkreises

In Abb. 11 (unten) ist der online-gemessene α' -Martensitgehalt mittels 3MA-II und Softsensor dargestellt. Zu Versuchsbeginn liegt der α' -Martensitgehalt bei ca. 28%. Innerhalb von ca. 250 s ist der Sollwert von 50% durch Verringerung des Vorschubes f von 0,25 auf ca. 0,1 mm/s eingeregelt. Die Regelung ist somit im Rahmen der Messgenauigkeit von 3MA-II und Softsensor stabil und stationär genau. Allerdings weist das Messergebnis im Gegensatz zum Wanddickenregelkreis mehrere Unterschiede zur Simulation auf. Der

anfängliche α' -Martensitgehalt ist mit ca. 28 % fast 7 % geringer als in der Simulation und der Vorschub wird durch die Regelung deutlich stärker verringert (auf 0,1 statt 0,14 mm/s). Diese Abweichung ist allerdings nicht auf eine schlechte Modellgüte zurückzuführen, sondern durch die Tatsache bedingt, dass das α' -Martensitbildungsverhalten metastabiler austenitischer Edelstähle stark störgrößenbeeinflusst ist. So haben ergänzende Voruntersuchungen der Autoren gezeigt, dass die Affinität zur α' -Martensitbildung stark durch die genaue Legierungszusammensetzung des Halbzeugs beeinflusst ist und teils selbst innerhalb einer Halbzeugcharge variiert. Insofern ist positiv hervorzuheben, dass trotz dieser Störbeeinflussung der vorgegebene Sollwert von $\alpha'_{soll} = 50\%$ im Rahmen der Messgenauigkeit stationär genau durch die Regelung einregelt wird.

4 Zusammenfassung und Ausblick

In diesem Beitrag wurde der modellbasierte Entwurf einer neuen Eigenschaftsregelung für das Drückwalzen metastabiler austenitischer Edelstähle vorgestellt. Diese soll über zwei Regelkreise den α' -Martensitgehalt und die Wanddickenreduktion regeln. Hierzu wurden die Regler mittels numerischer Optimierung und Modellsimulation parametrisiert. Auch erfolgte eine erfolgreiche experimentelle Validierung beider Regelkreise am Realsystem. Somit führt dieser Beitrag zu einer Erweiterung des Standes der Forschung, denn bislang sind – wie eingangs gezeigt – Eigenschaftsregelung nur selten und im Fall des Drückwalzens gar nicht experimentell validiert worden. Die Ergebnisse offenbaren allerdings auch weitere Herausforderungen und Ansätze für die weitere Forschungsarbeit. So unterscheidet sich die Geschwindigkeit des Einregelvorgangs deutlich zwischen beiden Regelkreisen, wofür der bauraumbedingte Sensorversatz ursächlich ist. Hier könnte zukünftig die Nutzung prädiktiver Regelungsansätze (z.B. Smith-Prädiktor, MPC) helfen, die Auswirkung des Versatzes bzw. der resultierenden Totzeit auf die Regelungsdynamik zu minimieren und somit auch eine lokalere Abgrenzung von Bereichen unterschiedlichen α' -Martensitgehaltes zu ermöglichen. Weitere zukünftige Forschungsfelder könnten zudem sein, das Regelungskonzept auf umfangsseitige, angulare Gradierungen sowie Tiefengradierungen auszuweiten und somit komplexere Bauteilstrukturen zu ermöglichen. Auch ist es denkbar, den Chargeneinfluss auf die Regelung stärker im Regelungskonzept zu berücksichtigen, um z.B. über adaptive Regelungskonzepte eine optimale Einstelldauer der Regelung auch bei abweichender Legierungszusammensetzung zu garantieren.

Danksagung

Die Autoren danken der Deutschen Forschungsgemeinschaft (DFG) für die Förderung des Projekts Nr. 24335026 „Eigenschaftsorientierte Regelung von Verfestigungs- und Phasenumwandlungsprozessen beim Drücken und Drückwalzen metastabiler Austenite“ im Rahmen des Schwerpunktprogramms SPP 2183 „Eigenschaftsgerichtete Umformprozesse“.

Literaturverzeichnis

- [Al16] Allwood, J. M. et al.: Closed-loop control of product properties in metal forming. *CIRP Annals* 2/65, S. 573–596, 2016.
- [An09] Angermann, A. et al.: *MATLAB - Simulink - Stateflow. Grundlagen, Toolboxen, Beispiele*. Oldenbourg, München, 2009.
- [Di92] Dierig, H.: *CNC-Drücken mit adaptiver Regelung*. Zugl.: Dortmund, Univ., Diss. VDI-Verl., Düsseldorf, 1992.
- [En17] Endelt, B.: Design strategy for optimal iterative learning control applied on a deep drawing process. *The International Journal of Advanced Manufacturing Technology* 1-4/88, S. 3–18, 2017.
- [ETD13] Endelt, B.; Tommerup, S.; Danckert, J.: A novel feedback control system – Controlling the material flow in deep drawing using distributed blank-holder force. *Journal of Materials Processing Technology* 1/213, S. 36–50, 2013.
- [FBL23] D. Fischer, J.; A. Bristow, D.; G. Landers, R.: Layer-to-layer norm-optimal control of incremental sheet forming with a data-driven model. *Manufacturing Letters* 35, S. 1196–1202, 2023.
- [Fi20] Fischer, P. et al.: On part-to-part feedback optimal control in deep drawing. *Journal of Manufacturing Processes* 50, S. 403–411, 2020.
- [Fö22] Föllinger, O. et al.: *Regelungstechnik. Einführung in die Methoden und ihre Anwendung*. VDE Verlag GmbH, Berlin, Offenbach, 2022.
- [HD11] Hao, W.; Duncan, S. R.: Constrained model predictive control of an incremental sheet forming process. In (Sarangapani, J. Hrsg.): *2011 IEEE International Conference on Control Applications (CCA 2010)*. Denver, 28 - 30 September 2011. IEEE, Piscataway, S. 1288–1293, 2011.
- [Is22] Ismail, A. et al.: Control system design for a semi-finished product considering over- and underbending. *Engineering Proceedings* 1/26, 2022.
- [Ker22] Kersting, L. et al.: Innovative online measurement and modelling approach for property-controlled flow forming processes. *Key Engineering Materials* 926, S. 862–874, 2022.
- [Ke22a] Kersting, L. et al.: Echtzeitfähige Modellierung eines innovativen Drückwalzprozesses für die eigenschaftsgeregelte Herstellung gradierter Bauteile. In (Jumar, U.; Diedrich, C. Hrsg.): *Entwurf komplexer Automatisierungssysteme. Beschreibungsmittel, Methoden, Werkzeuge und Anwendungen*. Otto-von-Guericke-Universität Magdeburg, Magdeburg, 2022.
- [Ke23] Kersting, L. et al.: Echtzeitfähige Modellierung eines innovativen Drückwalzprozesses für die eigenschaftsgeregelte Bauteilfertigung. *at - Automatisierungstechnik* 1/71, S. 68–81, 2023.
- [Ke24] Kersting, L. et al.: Der eigenschaftsgeregelte Drückwalzprozess als neues mechatronisches System: Konzept und Modellierung. In (Bertram, T. et al. Hrsg.): *Mechatroniktagung 2024*. Technische Universität Dresden 14. und 15. März 2024, S. 47–52, 2024.

-
- [La21] Lafarge, R. et al.: Data based model predictive control for ring rolling. *Production Engineering* 6/15, S. 821–831, 2021.
- [La24] Laue, R.: Entwicklung des selbstregelnden Drückwalzens. Dissertation, Chemnitz, 2024.
- [Lu16] Lu, H. et al.: Model predictive control of incremental sheet forming for geometric accuracy improvement. *The International Journal of Advanced Manufacturing Technology* 9-12/82, S. 1781–1794, 2016.
- [Lu20] Lunze, J.: *Regelungstechnik 2. Mehrgrößensystem, Digitale Regelungen*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2020.
- [PAD14] Polyblank, J. A.; Allwood, J. M.; Duncan, S. R.: Closed-loop control of product properties in metal forming: A review and prospectus. *Journal of Materials Processing Technology* 11/214, S. 2333–2348, 2014.
- [Ro24] Rozo Vasquez, J. et al.: Softsensor model of phase transformation during flow forming of metastable austenitic steel AISI 304L. In (Kusiak, J.; Rauch, Ł.; Regulski, K. Hrsg.): *Numerical Methods in Industrial Forming Processes. Numiform 2023*. Springer, Cham, 2024.
- [Ru93] Runge, M.: *Drücken und Drückwalzen. Umformtechnik, Werkstückgestaltung, Maschinen, Steuerungskonzepte*. Verl. Moderne Industrie, Landsberg/Lech, 1993.
- [Sc21] Schulte, C. et al.: Model-based Control of the Strip Roughness in Cold Rolling. *IFAC-PapersOnLine* 11/54, S. 109–114, 2021.
- [Spi23] Spies, D. et al.: Conception of a multivariable product property control for punch-hole-rolling. In (Madej, L.; Sitko, M.; Perzyski, K. Hrsg.): *Material Forming. The 26th International ESAFORM Conference on Material Forming held in Krakow, Poland, April 19-21, 2023*. Materials Research Forum LLC, S. 2081–2090, 2023.
- [St24] Stebner, S. et al.: Monitoring the evolution of dimensional accuracy and product properties in property-controlled forming processes. *Advances in Industrial and Manufacturing Engineering* 8, S. 100133, 2024.
- [TP23] Tricarico, L.; Palmieri, M. E.: Robust design of deep drawing process through in-line feedback control of the draw-in. *Applied Sciences* 3/13, 2023.
- [Wa23] Wang, C. et al.: Learning-based model predictive control for two-point incremental sheet forming. *Journal of Manufacturing Processes* 105, S. 187–198, 2023.

