



Unifying data model of manufacturing processes for cross-domain collaboration in the Internet of Production

Jiyoung Moon¹ · Lucia Ortjohann¹ · Philipp Niemietz¹ · Thomas Bergs^{1,2} · Alexander Peitz³ · Michael Emonts³ · Christian Brecher^{2,3,4} · Jana Sasse⁵ · Yannik Lockner^{5,8} · Malte Seefeldt⁵ · Mauritius Schmitz⁵ · Christian Hopmann^{3,5} · Moritz Kröger⁶ · Tamoghna Majumder⁷ · Nilesh Thakare⁷ · David Bailly⁷ · Emad Scharif⁷

Received: 9 October 2025 / Accepted: 28 April 2026
© The Author(s) 2026

Abstract

Modern production heavily relies on data generation, automated processing, and decision-making to enhance quality and cost efficiency. Key elements for achieving fully autonomous, connected, and digital production, as envisioned by the Internet of Production (IoP), include data-driven inspection, machine-to-machine communication, and machine learning for process control. Current research focuses on concepts like cyber-physical systems, ontologies, digital shadows, and digital twins, often considering limited manufacturing processes, which restricts wider application due to high adaptation needs. This paper introduces a process-independent data model for describing industrial processes through interoperable data formats and a method for defining generic atomic production operations. This format captures data from general production runs to individual measurements, enabling comprehensive data representation and exchange across diverse manufacturing sectors, thus facilitating data sharing and digital value creation. The data model's applicability is showcased within the IoP *Cluster of Excellence* using six distinct manufacturing processes.

Keywords Digital shadow · Industrial production · Data model · Meta-model · Internet of production · Cross-domain

1 Introduction

Modern manufacturing processes generate and collect vast volumes of process data, facilitated by advancements in technologies such as enhanced sensors and the Industrial Internet of Things (IIoT). These developments provide a foundation for integrating machine learning and computer science methodologies into production systems. A key enabler of smart manufacturing is the availability of extensive datasets, as highlighted in prior research [1].

The primary objective of smart manufacturing is to transform data captured throughout the product life cycle into actionable production knowledge. This transformation is most effective when manufacturing data is computationally analyzed, enabling data-driven decision-making to optimize production processes [2].

The Internet of Production (IoP) [3] denotes a globally interconnected production infrastructure that enables real-time exchange, storage, and processing of heterogeneous

manufacturing data across organizational and technological boundaries. In this context, IoP represents a connected and integrated approach to manufacturing, where the physical and digital worlds converge to provide an efficient, agile, and flexible production environment. By leveraging IIoT, cloud computing, and artificial intelligence technologies, the IoP enables manufacturers to optimize their production processes, improve quality control, reduce downtime, and enhance supply chain management [4]. It also facilitates cross-domain collaboration, providing stakeholders with real-time access to production data, enhancing decision-making and shortening time-to-market [3].

Essential to the IoP are data models, which are used for organizing, defining and managing the complex relationships between data elements [3]. These models should not only facilitate efficient data processing, storage, and analysis, but also serve as the basis for implementing advanced data-driven strategies aimed at a broad range of applications such as increasing the efficiency and agility of

Extended author information available on the last page of the article

manufacturing organizations. As large volumes of heterogeneous data are generated in production environments, data models provide the necessary structure for processing and storing data streams that originate from different sources and formats.

Another opportunity given by using data in the IoP is the creation of comprehensive digital representations of processes and products, called digital shadows, which lead to precise monitoring and control of manufacturing [5]. These digital shadows are crucial for combining different data sources and enable a comprehensive, digital representation of production. By using digital shadows, manufacturing companies gain a unified view on process parameters, material behavior and resulting product quality [6]. While data models provide the structural and semantic foundation that ensures consistency, interoperability, and traceability of production data, digital shadows use these data models to create dynamic, digital representations of production processes and products in real time. In this way, digital shadows rely on data models to enable real-time synchronization, cross-domain data integration, and analytical reuse across the production lifecycle. Therefore, the connection between digital shadows and data models is of central importance in IoP.

This work presents a conceptual reference architecture and data model whose primary objective is structural generalization across manufacturing domains rather than empirical performance evaluation.

2 Literature review

To leverage the aforementioned potentials of the IoP, acquiring, managing and analyzing process and machine data in order to generate process knowledge is crucial. A multitude of different approaches and concepts for data and knowledge management are a focus of current research and often enough are used with intersecting meaning. For this reason, the following subsection aims to give an overview of the researched fields and defines the frequently used terminology.

2.1 Terminology and concepts

2.1.1 Cyber-physical production systems

The concept of a *Cyber-Physical Production System* (CPPS), in the broadest sense, describes a system in a manufacturing environment that is composed of human resources, production equipment and aggregated products, to which there exists one or several interaction interfaces for monitoring and control [7]. In such CPPS', the physical and virtual (cyber) component interact through data acquisition

methods, usually by automated reading of sensor data or operator inputs, in combination with feedback and control loops either directly employing physical actors or through decision-making support to the operator [8].

2.1.2 Asset administration shells

In the scope of the Reference Architectural Model for Industry 4.0 (RAMI 4.0), a related concept, the *Asset Administration Shell* (AAS) has been introduced. An AAS is described as the "virtual, digital and active representation of an asset in an I4.0-system" [9]. The shell therefore provides generic access to underlying, usually proprietary device drivers and information channels. This can include meta-data to describe the device. The device status can be archived, creating historical data for life cycle management [10]. However, there is no common data model for specific systems, and therefore a wide variety of implementation exist [11].

2.1.3 Digital twins

The term *Digital twin* describes a dual system consisting of a physical and digital object. These objects are interconnected at the digital level through a fully automatic and bidirectional data flow, where physical changes are reflected in the digital model and digital changes can influence the physical system [12]. Therefore, physical changes are reflected onto the digital model, and digital changes to the model automatically reflect onto the physical system [13]. Digital twins show a broad horizon of applications, ranging from digital modeling of individually manufactured parts over whole manufacturing plants [14]. However, due to the high degree of interconnection between virtual and real asset, the creation and maintenance of digital twins usually requires high implementation and integration effort.

2.1.4 Digital shadows

Digital shadows show a strong uptake in current research. Kritzinger et. al. defined the digital shadow as a virtual asset representation, including only a unidirectional automated data flow from real to virtual representation [14]. Further, digital shadows usually only partially model given processes or assets because their creation is coupled to a defined purpose and use-case [15]. Therefore, digital shadows show a wide variety of applications in different fields. These include, but are not limited to, manufacturing. For example, digital shadows have been used for monitoring and prediction of tool wear in machining systems [16]. Digital shadows have also been applied for the representation and analysis of machine tool process data [17]. Further applications include plastics processing [18] and composites

manufacturing [19]. Applications have also been reported in buildings and infrastructure [20] as well as in health [21] and comfort [22] systems.

2.1.5 Ontologies

Another prerequisite for smart manufacturing and cross-domain solutions is the transferability of data and knowledge. This can be made possible through the use of so-called ontologies, which define a common dictionary for individual domains. The concept was first introduced by Berners-Lee et al. [23] and defines not only the vocabulary and the standardized terms for talking about or storing data within a domain but also the interplay between components through semantic relationships. As a result, data can be categorized and universally described. While some ontologies span the universal concept of a production domain, enabling shared semantic representations across manufacturing environments [24], others focus on specific applications or automation use cases where ontologies model domain entities, their properties, and relationships for a particular task [25], some of which are discussed in section 2.2.

2.2 Processes and process-related digital shadows

While the concepts introduced in Section 2.1 provide general frameworks for digital representation and interoperability in Industry 4.0 environments, previous research has often focused on individual production processes, resulting in different degrees of progress towards the IoP for different applications. This section highlights the production processes considered in this paper and their individual state of the art regarding digital shadows and data models.

2.2.1 Automated fiber placement

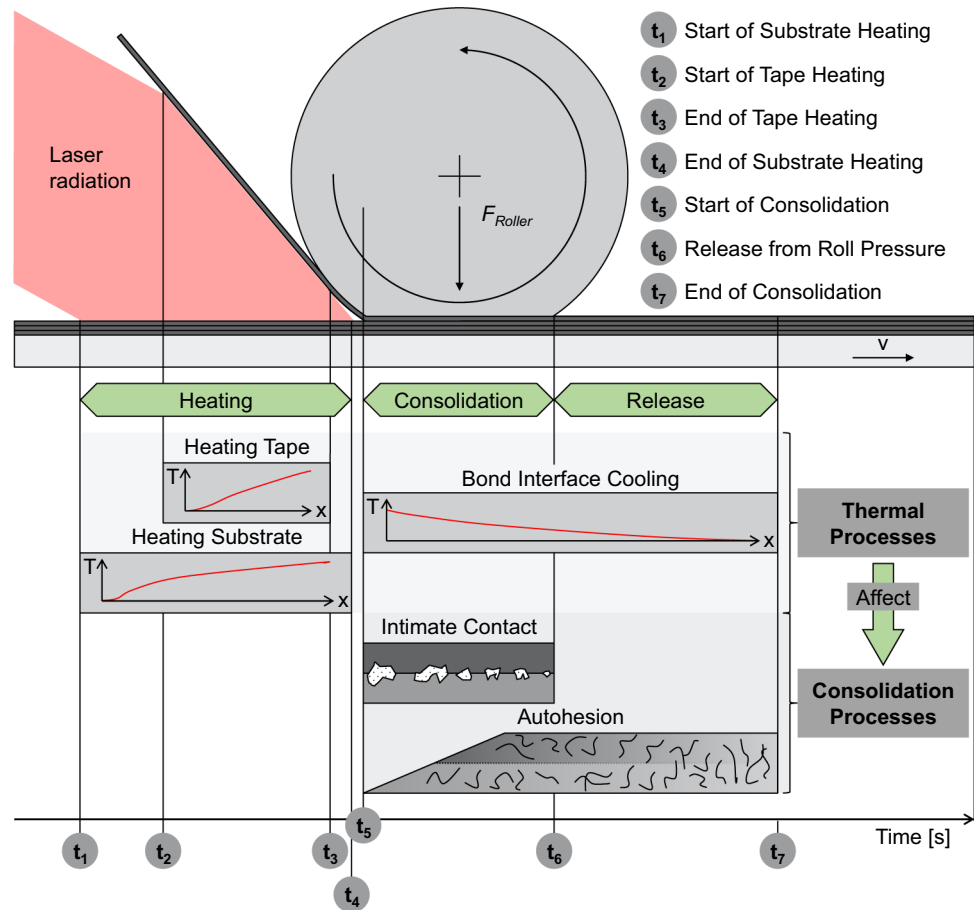
Automated fiber placement (AFP) is a discontinuous process in which strips of unidirectionally reinforced, thermoplastic strip material (thickness approx 50-250 μ m), so-called UD-Tapes, are additively joined layer-by-layer into so-called laminates. By using a laser system for material heating, the material can be partially or fully welded together, creating a strong material bond. This is known as in-situ consolidation. When heated over the melting temperature by the laser, a consolidation roller, usually coated with an elastomer, applies the required pressure simultaneously onto the tape and substrate to enable the bonding process. The tapes can be flexibly oriented to realize varying fiber angle orientations. Figure 1 schematically shows the physical process.

Variations in key AFP process parameters such as heating temperature, placement rate, and consolidation pressure influence the bonding mechanisms and resulting laminate

quality [26]. In particular, the development of intimate contact and subsequent polymer diffusion during consolidation are strongly governed by the process temperature and pressure conditions [27]. Experimental studies have further shown that processing temperature and placement rate significantly affect interlaminar strength and laminate properties in laser-based tape placement processes [28]. Further, as a process manufacturing primarily semi-finished products, the produced laminates are supplied to subsequent forming processes such as thermoforming or injection molding. Interchanging processing history to and from those processes can significantly improve part quality, process reliability and defect traceability.

Employing digital models and data-driven methods in automated fiber placement has been gaining relevance in recent research. Current research on data-driven models, digital twins and digital shadows is strongly focused on the topics parameter history and optimization, part defect detection and manufacturing cell/machine setup through 3D component modeling and simulation.

Thermal processes and the thermal process history is one of the main factors influencing bond quality and part warpage. Data-driven models to predict thermal phenomena and bonding quality are shown by Schäkel et. al. [29]. Schulz et al. demonstrate how temperature-dependent process data can be integrated into a digital representation of the AFP process to estimate bonding quality during manufacturing [30]. Further work by Schulz et al. links spatially resolved temperature measurements with laminate geometry to analyze process-induced effects in tape laying processes [31]. Geometric defects during manufacturing such as gaps, overlaps, folds, wrinkles or foreign objects are a strong research focus. Oromiehie et al. analyze processing-induced defects in automated fiber placement, including gaps, overlaps and waviness caused by process parameter variations [32]. Fu and Yao review common defect types in fiber reinforced composites, including fiber misalignment, voids, and interfacial defects that affect structural integrity [33]. However, little work is done concerning digital modeling of detected defects, a crucial step for information exchange to subsequent forming processes. Schuster et. al. show a system to detect defects and hint at modeling these in a database structure [34]. Brasington et al. present AI-supported process planning systems for automated fiber placement that integrate process knowledge and simulation-based decision support for manufacturing configuration [35]. Francis et al. develop a digital simulation environment for AFP that predicts the influence of process parameters on defect formation and product quality [36]. Brasington et al. further propose methods for mapping multimodal manufacturing data within the AFP process to support defect analysis and process optimization [37].

Fig. 1 Schematic description of automated fiber placement

However, existing approaches typically focus on isolated aspects and lack a comprehensive data-structural framework that enables cross-process information exchange. As such, we propose a generic data model for automated fiber placement comprising the manufacturing conditions and the parameter history.

2.2.2 Injection molding

The injection molding (IM) process is one of the most significant plastic manufacturing methods, known for its complex and non-linear process dynamics. IM products are commonly used in automotive, medical, electrical, and household applications. This is a batch manufacturing process where polymer material is fed into the plasticizing unit's barrel by a rotating screw. The material melts from friction and heat supplied by heaters around the barrel, collecting in front of the screw tip in an area called the antechamber. As this melting and dosing occurs, the screw retracts, allowing material to accumulate under applied back pressure to homogenize the melt's characteristics. Once enough molten plastic has accumulated, the screw stops rotating and then injects the melt at high pressure through a nozzle into a temperature-controlled mold. The mold contains one or more

cavities which are filled during injection that shape, cool, and release the solidified part when the mold opens. Heat exchange occurs through a temperature-controlled coolant circulating in cooling channels within the mold. During the holding pressure phase, the screw continues to inject melt to make up for volumetric shrinkage caused by the melt's pressure-volume-temperature (pVT) characteristics. The plasticizing unit then replenishes the injected melt during the cooling and ejection phase, preparing for a new cycle as the mold closes again (see Fig. 2) [38].

Previous models for process control in IM have been described in the context of digital shadow data structures and production data modeling [25]. Process control strategies and model-based monitoring approaches for injection molding are further discussed in [39] and [40]. Ontology-based approaches for representing manufacturing knowledge and enabling semantic interoperability have been proposed in [41]. Sapel et al. built a holistic description of manufacturing with an ontology [42]. This ontology defines all assets, the properties and the relations between the assets. Further this holistic ontology was built up to connect existing ontologies from different manufacturing processes via cross-domains. For example, the two manufacturing processes of IM and drilling are combined using

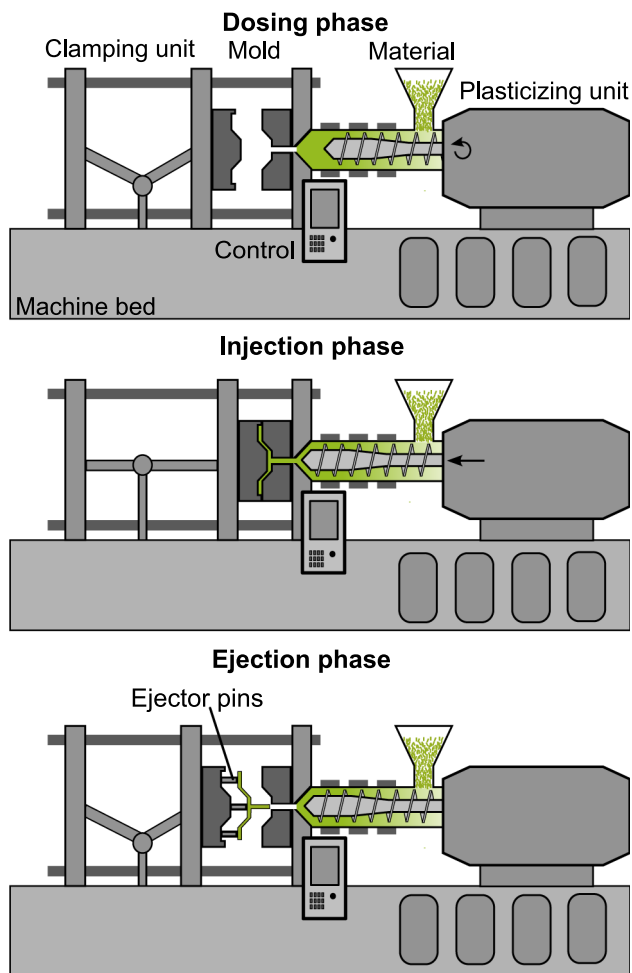


Fig. 2 Injection molding process

the built ontology in [42]. The IM machine and the drilling machine are connected via the class machine. Both subclasses include the machines available in the company and their describing data.

In addition to the machine, this ontology focuses on the physical object, which is described by the classes raw material, product, machine and mold. These classes are subclasses of the physical object.

The physical object can be fully described by the connections and the data included in the individual classes [42]. In addition to the ontology described in [42], there are IM-specific interfaces for process description. For this

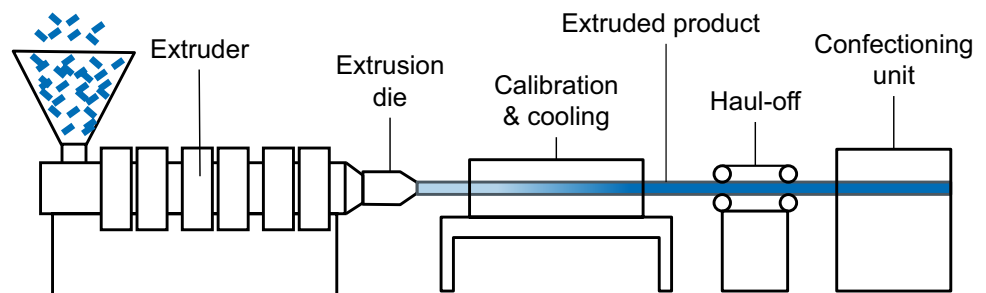
interfaces, most data models and standards are focused on machine data within one production line. A standard protocol for machine data is specified by the EUROMAP standards. EUROMAP specifies general concepts and interfaces for Industry 4.0 integration in plastics and rubber machinery [43]. EUROMAP 77 defines the OPC UA-based data interface for injection molding machines and manufacturing execution systems [44]. While these protocols are extensive, there is no similarly accepted standard for data regarding the produced part after it leaves the machine, e.g., for quality control purposes.

2.2.3 Plastics profile extrusion

Extrusion is the most common method for producing continuous plastic products such as films, sheets, pipes, and profiles. A simple profile extrusion line is depicted in Fig. 3. In thermoplastics extrusion, plastic granules are melted in an extruder to provide a continuous melt stream. To form a specific product, the melt stream is pushed through a die and cooled in a calibration section before being cut and stacked for further processing. The complexity in this process from a data models perspective stems from the modularity of extrusion lines. Many extrusion lines consist of multiple extruders, melt pumps, mixing elements, or have additional pre- or post-processing steps associated with them.

In plastics processing, the state of the art regarding industry 4.0 fitness varies by processing method. While there are numerous publications in IM, there is only little research on extrusion applications. This also represents the current situation in the industry, where innovation is often first adopted in IM. Ramírez-Durán et al. describe an ontology for extrusion [45]. This ontology contains modules for the components of an extrusion line, the spatial relationships between the components, and the features and 3D models of those components. The modules are completed by a module for data captured by sensors in the extrusion line. Domain experts from industry were consulted during development. Most data models and standards in plastics processing are focused on machine data within one production line. A standard protocol for machine data is specified by the EUROMAP standards [43], which specify the OPC UA interfaces. EUROMAP 84 is specific to extrusion systems, while

Fig. 3 Profile extrusion line



EUROMAP 86 (material supply systems) holds capacities for specifying material data. Although these protocols provide comprehensive specifications for machine-level communication, no comparably established standard exists for representing product-related data beyond the machine boundary, for example in the context of quality inspection.

2.2.4 Fine blanking

Fine blanking is a forming technology for manufacturing high-quality sheet-metal components, meeting the stringent demands of industries such as automotive, aerospace and electronics. In fine blanking, the metal strip is first straightened from the coil before performing the shearing operation (see Fig. 4). Precision tools in fine blanking allow the production of large quantities of parts with optimal dimensional and shape tolerances, as well as an up to 100% smooth sheared surface. Unlike conventional shearing, fine blanking employs a counter punch, a V-ring, a small die clearance and three independently acting process forces: punch force, counter punch force, and V-ring force. These features improve the quality of sheared edges while eliminating the need for secondary operations, saving time and reducing costs [46].

In the field of sheet-metal forming, there is a significant emphasis on data-driven methods for process monitoring to enable real-time control and integration of closely linked production processes [47]. In addition, Liewald et al. highlight the potential of these methods to complement and enhance traditional modelling techniques such as simulations [48]. Addressing fluctuations in material properties requires non-destructive testing techniques, combined with advanced techniques for data analysis [49]. By leveraging process data from forming technology, data-driven approaches can reduce process uncertainties, increase production efficiency, and ensure product quality.

However, data acquisition in this context remains challenging due to high production rates, dynamic effects, and significant process forces, making it an ongoing topic of research. Molitor et al. developed a procedural method specifically for data acquisition, preparation, and transformation tailored for machine learning applications [50]. Moreover, the performance of classification models is heavily dependent on the quality of the dataset and its acquisition procedure. Key factors include sensor type selection, measurement methods used, and positioning within the tool or press [51].

Research on digital shadows or twins in sheet-metal forming processes is limited, and there is a lack of standardization compared to other manufacturing processes [52]. Efforts to combine simulations and process data in industrial stamping plants are underway [53], but standardized data storage and sharing remain unestablished.

2.2.5 Hot rolling

Hot rolling is a continuous forming process that involves heating the metal above its recrystallization temperature and then rolling it into desired shapes and sizes. The hot rolling process includes multiple passes of height reduction where the material height is decreased in every step until the final desired height is achieved. The used raw material for producing those strips are casted slabs. These casted slabs experience geometric changes like reduction in thickness and increase in length during the rolling process, especially in the head and tail parts [54]. Research work has shown that process parameters such as rolling temperature, rolling speed as well as material composition influence product quality such as profile and flatness of rolled sheets and space in between chatter marks [55]. A schematic description of the hot rolling process is illustrated in Fig. 5.

Fig. 4 Fine blanking line

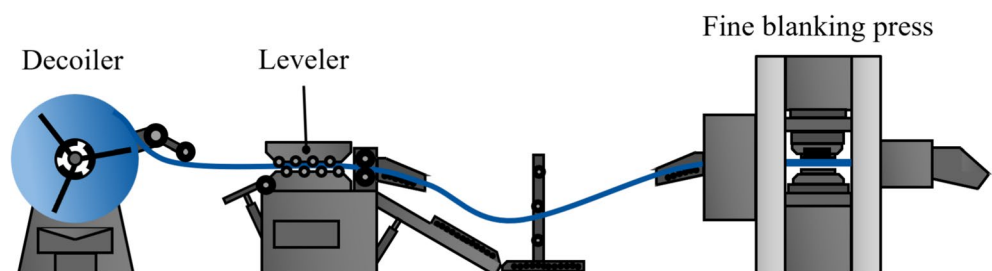
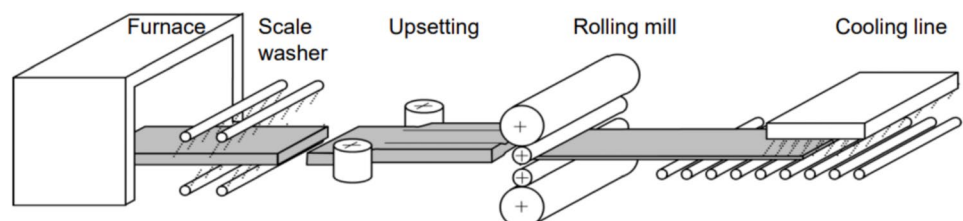


Fig. 5 Schematic illustration of hot rolling



Recent study has shown that using data-driven techniques and digital models for hot rolling processes is becoming more and more relevant. In [56], a reinforcement learning algorithm deep deterministic policy gradient is used to design the pass schedule automatically for hot rolling process to achieve final mechanical properties. In [57], using backpropagation neural network and genetic algorithms, quantitative relationships between parameters are established that are present in the models for mean flow stress, meta-dynamic recrystallization and static recrystallization with chemical and deformation conditions of Ti micro alloyed steel. In [58], a deep learning model based on process parameters and chemical composition is developed that is able to predict mechanical properties of hot rolled steel plate.

Research on digital-twin in the hot-rolling process can be deeply integrated and interconnected through the physical and virtual worlds. It has the capacity to forecast and regulate product states while monitoring the production process [59]. The hot-rolling digital twin system is significant to encourage the development of smart factories [54].

2.2.6 Ultra-short pulse laser ablation

In ultra-short pulse (USP) laser ablation, femto- to picosecond laser pulses are used to remove material from a workpiece through highly precise ablation processes [60]. USP laser radiation with pulse durations below 10 ps enables material removal with minimal thermal influence due to the extremely short interaction time between laser energy and the material [61]. The overall working principal is schematically shown in Fig. 6. Typical ablation crater sizes range from sub micrometer to several tenth of micrometers and ablation depth of a few hundred nanometers [61]. By overlapping consecutive pulses, a 2.5D shape can be ablated into the material. Typically this process works in layers, where the 2.5D structure is sliced. These slices are afterwards

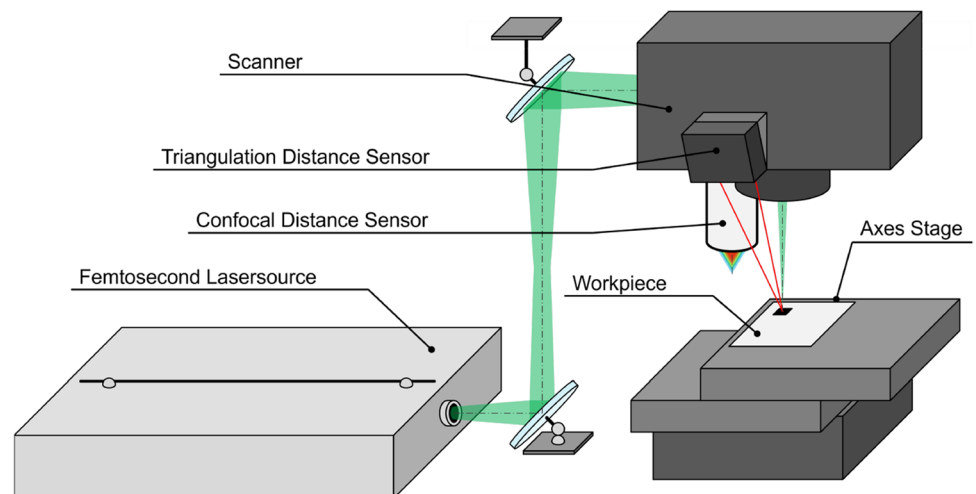
hatched with vectors to form a single layer of the process. The vector list is afterwards loaded onto the machine and the focal spot is moved along the vector path ablating material at this specific area. To this day the path geometry is typically not altered during the process but is precalculated and loaded onto the machine. However most systems also include high frequency emission observation systems, which map process emission to a specific point of the workpiece [62] or other quality control systems like confocal or white light interferometers, which allow for the evaluation of depth and roughness at certain parts of the workpiece [63]. These sensor systems could be used in order to alter the beam path depending on certain readings, but to this day a unified data model which allows for such advanced optimization techniques is not in place.

Most systems use G-Code or proprietary data formats for storing vector paths. Some laser machining software environments implement such approaches for integrating machine control and process data [64]. Other industrial platforms provide proprietary frameworks for linking sensor information with machining data streams [65]. Comparable solutions for scan system control and process monitoring are also implemented in commercial laser control software [66]. However, some open source formats do exist which try to open up the underlying vector model [67].

2.3 Data models for generic digital shadows

Research on frameworks for the digital shadow within the IoP was initiated in 2021 by Brecher et al., who proposed a conceptual model to define the relationships among assets, data, and the objectives of a digital shadow [68]. Becker et al. [39] applied this concept to real-world IM scenarios. Building on this model further, the Digital Shadow Reference Model (DSRM) by Michael et al. [69] provides a structured framework for modelling data elements, relationships, and resources across production domains. This framework

Fig. 6 Ultra-short laser processing



integrates digital shadows throughout product lifecycle stages and employs ontologies for semantic alignment, thereby ensuring data interoperability across IoP applications.

In addition, Riesener et al. [70] focus on digital shadows within product lifecycle management, proposing a methodology for consolidating and structuring data from multiple sources. Complementing this approach, Van der Aalst et al. [71] present process mining as a practical DS model within the IoP, highlighting its flexible design. Their object-centric metamodel captures multiple objects per event, thereby overcoming limitations of traditional event logs.

Furthermore, Brecher et al. [72] further employed ontology-based data access (OBDA) to create use-case-specific digital shadows, which organizes raw sensor data into semantically consistent structures. While Sapel and Hopmann [73] highlighted the need for a standardized vocabulary in IM, proposing an ontology-based dictionary to improve communication and data integration across systems.

Nevertheless, these papers provide conceptual frameworks but lack concrete data structures to support practical implementation. Therefore Van der Aalst [71] introduced the concept of *Data-Models-in-the-Middle* (DMM) as a practical solution to address complexity and limited reusability in data handling within the IoP. DMM have the goal to standardize and harmonize diverse data sources, enabling immediate analytics capabilities without the need for further data transformations. Tacke et al. [74] used this approach and proposed a minimal meta-model specifically designed for low-level machine data, aimed at enhancing standardization and interoperability across Industry 4.0 solutions. Their DMM uses a CSV-based format to ensure accessibility and ease of use, facilitating data sharing, processing, and visualization directly from raw machine outputs.

While these papers provide a foundation for DMM, they do not fully address the incorporation of metadata for production processes. To bridge this gap, the model proposed in this paper explicitly integrates metadata, advancing these concepts and ensuring a more comprehensive representation of production dynamics.

3 Data architecture for digital shadows in production

As can be seen in the current state of the art, strong emphasis is put on developing digital shadows to model and combine quality, simulation and machine data usually from one individual process for means of quality control and process optimization. However, reusability of machine and sensor data from different machines or different processes is usually limited due to the data being tailored to the specific use-case, the process and the machine. Data-driven optimization

methods require large amounts of data and the relative cost of producing this data through dedicated experimental runs is usually high. Therefore, utilizing process and machine data from one production environment to another shows potentials to be highly beneficial [75]. In this section, a general, process-independent data model to universally describe production and machine data will be presented (Fig. 7).

This model introduces the core classes and attributes required to structurally represent production environments. The main entities comprise ProductionRuns, denominators, Assets, Measurement instances, Educts, and Products, together with their relationships. Attributes define identifiers, temporal references, and contextual metadata, while associations describe how machine data, process states, and produced entities are linked within a unified structure. The following subsections explain these elements step by step, beginning with the denominator as the fundamental reference entity.

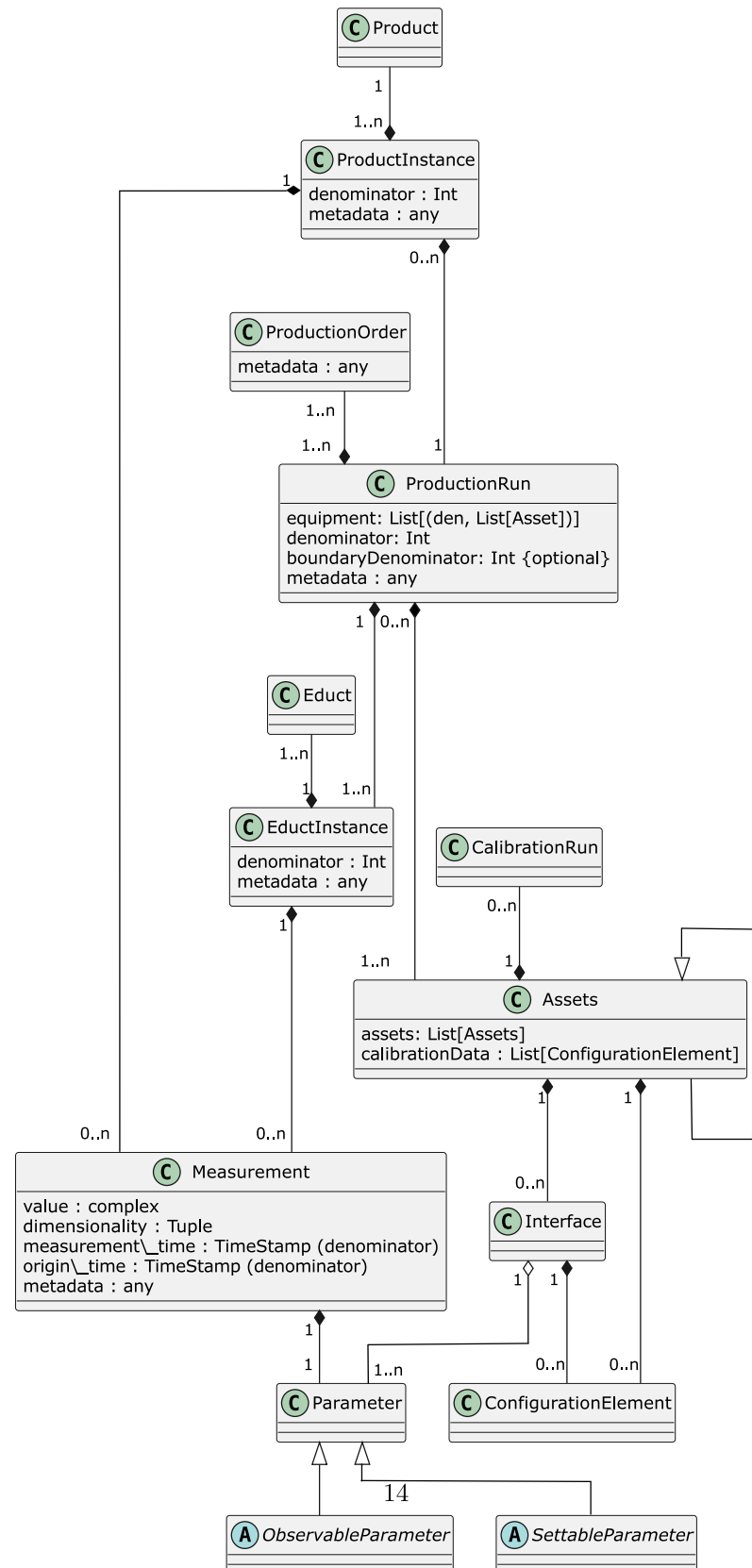
3.1 Common denominator

Referencing machine data to production steps and produced goods is a process-specific task which varies strongly with the type of process. In general, two types of processes are differentiated:

1. Continuous processes
2. Discontinuous processes

In continuous processes, material is usually fed using a continuous motion through the plant, e.g., in roll-to-roll processes. Discontinuous processes are composed of a distinct set of sequential, and possibly parallel, production operations.

Therefore, the present study proposes to define a common denominator as a process- and application-specific ground reference. This denominator defines the smallest relevant unit of production to be individually tracked. The denominator can be bound to a specific unit or value, or be unitless, defined purely in a nominal way. Possible definitions include e.g. a defined length of produced material, the execution of a specific process step or the completion of a cyclic operation in a connected assembly line. Thereby, the choice of the denominator is subjective, decided by the data scientist responsible for the process. The denominator should be fixed at design time and remain static over the production phase. A special case occurs when the denominator of the considered process defines either the start or the end of a process step, while the process step in itself remains variable in time. In this case, either the start of the first process step or the end of the last process step is not well defined. Therefore,

Fig. 7 Universal process data model

we propose an optional `boundaryDenominator`, enabling the encapsulation of `ProductionRuns` into a well defined denominator interval. In Fig. 8, exemplary denominators are schematically depicted.

Choosing the denominator yields an inherent trade-off between data complexity and accuracy. When the denominator is too large, adding new data points that require a lower resolution is not seamless, requiring adaptation to the data structure. However, a low denominator increases complexity and data synchronization, especially with large production systems. The required overhead takes up a non-negligibly large amount of storage. Lastly, the optimal denominator size is application- and domain-dependent, i.e., in research or prototyping applications it could be lower than in later production environments.

The denominator can be reset through dedicated events which interrupt regular production and ensue a change in the production environment, such as material changes, emergency stops, planned and unplanned maintenance.

3.2 Machine-specific data structures

An essential consideration in production system design is the resolution at which `Assets` are defined. The definition of `Assets` directly impacts the observable parameters and the characteristics of the production run. When the asset boundary is set at a coarse level—for instance, encompassing an entire unit such as a fine blanking machine or a USP structuring machine—this decision inherently limits the variability that can be captured within the production process. Under such a configuration, the `ProductionRun` data structure can only describe processes occurring between distinct machines.

Conversely, defining `Assets` at a finer resolution, such as at the sensor and actuator level, enables the production run to be described in much greater detail. This granularity facilitates the design of more flexible production processes

within individual machines. However, this increased flexibility introduces additional operational complexity, as the management of assets, production runs, and measurements becomes more intricate.

This trade-off underscores the importance of machine-specific data structures, which must be tailored to the operational environment of the production machine. The requirements for data and asset structure vary significantly across different production environments. For example, a cleanroom environment for wafer production using USP may emphasize parameters and configurations that prioritize precision and contamination control, with flexible production steps within a machine. In contrast, a standard manufacturing facility focused on mass production may necessitate a simpler, static configuration, where the production steps are fixed and less variable.

3.3 Measurement of machine data

Measurement is the key component of process analysis and optimization, and lie at the core of the digital shadow. However, measurements can differ greatly in their acquisition, format and purpose. Extensive work exists, focusing on process-specific measurement data. As such, the herein developed model does not aim to holistically define measurement data, but rather defines a process-independent, common interface for identifying and referencing it.

An interesting approach to enable data processing of raw measurement data using higher-order process models is addressed by the model-in-the-middle from Tacke et al. [74]. Therein, the authors aim to attribute the characteristics when, where and what to machine data. This is realized by integrating timestamps (when), Object Identifiers (where) and Observations (what) to the data structure. This is seamlessly compatible with the herein proposed Measurement structure, it being linked to the what and where

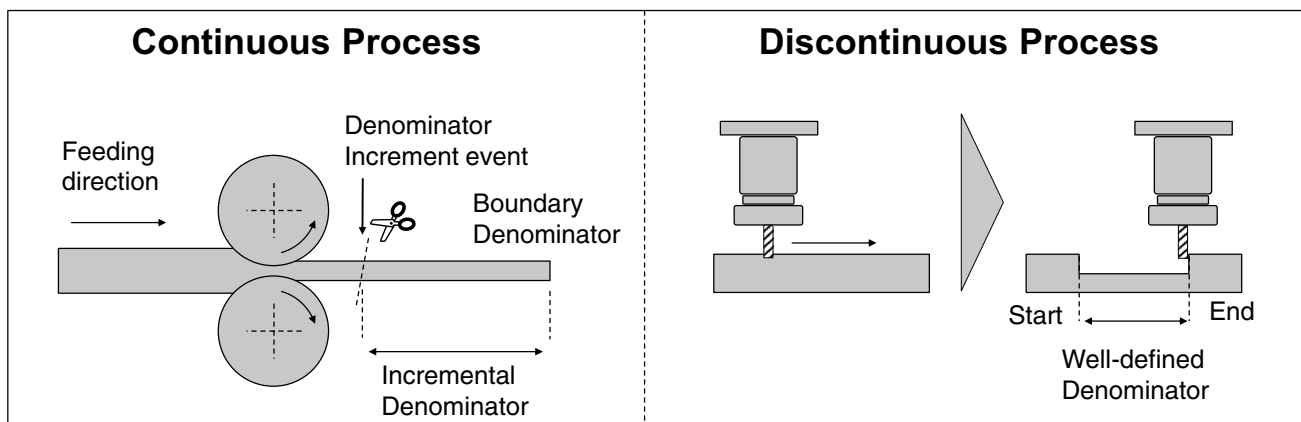


Fig. 8 Exemplary denominator selection

through the relationship to the Parameter and the actual value of the measurement.

In the presented model, the notion of when is extended to include two different concepts of time. The `measurement_time` refers to the timestamp relating to the time of acquisition of the data point, which, of course, differs for all data points even if they relate to the same denominator. Therefore, each data point is additionally associated with a `origin_time`, enabling all relevant data points to be mapped onto the same denominator, independent of their acquisition time. This `origin_time` is arbitrary but fixed, e.g., the time of the initial creation of a produced part.

The herein described Measurement is not meant to solely represent machine data. The data structure should encompass machine data as well as i.e. multi-dimensional quality data (such as microscopy images, 3D point clouds, thermal videos, CT- and X-Ray Scans) and mechanical, chemical and physiological material characteristics (i.e. datasheet data). To adequately describe these data structures with the goal to simplify data exchange and later evaluation, the data's dimensionality and thereby its interpretability needs to be known. As such, we propose to include a `dimensionality` tuple in the data model as a descriptor of the measurement's shape, in conjunction with a freely definable `metadata` field, which can include further important information for data interpretation (e.g. maximum physical ranges represented, units...).

3.4 Educt-specific data structures

Educts describe materials or components that serve as an input to the process and are usually consumed during production. This includes not only raw materials or semi-finished products that are transformed into a product, but also parts or materials that are required and frequently replaced for the operation of the machines.

Educt-specific data have a complex relation to the denominator since, in many cases, the educts are destroyed or modified before the denominator assigned to the according product is formally created. This applies e.g. to processes where educts are pre-processed and reshaped (melted, mixed) before the actual product is created.

In practice, educt-specific data often combine data sheets from the manufacturer with data collected offline, i.e., data that exists independent of the `ProductionRun`. It might also contain references to previous manufacturing processes, e.g., when the Educts of this process are products of a prior process. As a result, the Educt data sets are stored in a database separate from the `ProductionRun`, and the association of a specific `EductInstance` to a `ProductionRun` occurs retroactively through the denominator.

3.5 Product-specific data structures

For this purpose, a Product is defined as anything produced or created during processing. While each resulting Product exists beyond the process's scope, the related data for modeling purposes are associated via a specific `ProductInstance`.

Product-specific data can come from multiple sources, such as quality control measurements occurring during production or offline, e.g., in a lab setting. As a result, these data do not necessarily have the same dimensionality or formatting and need to be documented carefully (see Section 3.3). Offline data might not exist for all product instances since quality control measures in lab settings are only performed for a small sample size of the production batch. In contrast, inline data usually exists for each product instance that refers to the common denominator. Product-specific data must default to the common denominator (see 3.1).

Consequently, the product-specific data structures require careful collection of metadata. This includes documenting data formats, dimensionality, and the data's relationship with the common denominator. The metadata structures must conform with existing standards within the individual application, provided such standards exist. Necessary components within the metadata structure to ensure efficient access include, but are not limited to, a reference to existing standards (if applicable), the data source(s), e.g., lab equipment or inline sensors (see 3.2), the data's dimension and format, the data's relation to the common denominator and its completeness concerning the denominator, i.e., whether these product-specific data exist for each instance of the product or just for a subset.

4 Data model demonstration

The unifying data architecture presented in Section 3 is fairly generic. The following sections demonstrate how the data model fits each of the investigated production processes and how special cases for individual processes are handled.

4.1 Automated fiber placement

4.1.1 Denominator description

Since AFP is a discontinuous process, a distinct process cycle can be defined as the denominator. A process cycle can be one or multiple (when executed simultaneously) of the following operations:

- Orientation of substrate
- Pre-Heating of substrate
- Placement and consolidation of tape

This denominator choice enables a detailed resolution of process parameter measurements for individual process steps as well as correlating those steps to one another, creating insights about the process history. Due to the discontinuous, well-defined structure of these process operations, a `boundaryDenominator` is not required for the AFP process.

4.1.2 Assets and machine-specific data structures

A multitude of different machine technologies for AFP are currently available on the market. However, for manufacturing thermoplastic laminates with in-situ consolidation, a set of core components are required within each machine type [76]. These include a heating system, a tape guiding and feeding system, an NC-system for positioning and a pressure application system. All these components are usually integrated into a defined machine system and as such can be represented by the `Assets`, `AssetInterface` and `Parameters` data structures. Additionally, the use of inline-monitoring sensors such as temperature- and part surface sensors is gaining in popularity and can be modeled using the described approach.

4.1.3 Educt and product-specific data structures

Multiple interpretations of a `ProductionOrder` and subsequently a `ProductionRun` are possible in practice. For consistency with industrial applications, a `ProductionOrder` in AFP is defined as manufacturing a set number of identical laminates for a specific customer. Accordingly, during a `ProductionRun`, one or multiple laminates of the same or different types can be partially or fully manufactured. According to the data model introduced in Section 3, a `ProductionRun` manages multiple `ProductInstances` over its duration. However, according to the model specification, a `ProductInstance` can only be associated to one specific `ProductionRun`. This leads to challenges when e.g. a material (educt) change triggers a new `ProductionRun` while unfinished `ProductInstances` remain in production, a common scenario in series production. Exchanging the UD-tape spool is done on-the-fly, either through splicing or using advanced spool exchange systems. To handle this, the current `ProductInstances` are transformed into `Educts` of the next `ProductionRun`. This effectively splits the production history of a laminate into a nested structure. Due to the general definition of the denominator representing a process cycle, a `ProductInstance` does not necessarily contain a regular sequence of denominators but rather a rising, partial and irregularly spaced collection of denominators. Further, multiple `ProductInstances` can contain the same denominator, due to different operations being executed in

parallel (i.e. orientation of one laminate while a placement on another laminate is ongoing).

The main process Educt is the employed UD-Tape, which differs in e.g. matrix material, fiber material, width, thickness and additives. The `EductInstance` is an important data structure, tracking i.e. the batch number or storage time over defined environmental conditions, factors influencing later processing and part quality. Additional `Educts` can include adhesives to fix the first layer or cleaning chemicals of the tooling. The use of these `Educts` can influence later surface properties.

4.2 Plastics profile extrusion

4.2.1 Denominator description

Extrusion processes are continuous and, thus, do not have a 'natural' denominator; therefore, an artificial `denominator` must be defined. We propose one discrete unit as a single piece after confection, i.e., relating data to a roll for flat film, blown film, or a piece of extruded profile after the saw. Consequently, machine-specific data are collected over a set period and mapped onto the same common denominator. While this reduces accuracy, it represents the process since a piece of extruded profile is the smallest possible unit passing quality control. This denominator definition requires the use of an additional `boundaryDenominator`, which transcribes the denominator interval for the continuous process step, i.e., the start and end of the extrudate belonging to the denominator instance. The origin time is arbitrary but fixed, e.g., the timestamp where the extrudate leaves the die. This choice is associated with few consequences for educt-specific and product-specific data structures since it is the natural unit for product-specific data structures, and any change in the educt-specific data happens at a larger scale than this common denominator. In addition, any parts produced during an educt change are usually automatically removed from the production batch.

4.2.2 Assets and machine-specific data structures

Extrusion lines have a modular setup consisting of different components, and while some standard configurations exist, the model must be able to accommodate different extrusion line setups. However, existing ontologies [45] can be re-used to organize assets. Machine data are collected in conformation with EUROMAP 84 [43].

4.2.3 Educt and product-specific data structures

The `Educts` can be one or multiple base polymers or other additives. They are typically either supplied by raw material

manufacturers or produced in an in-house compounding process. The latter results in a special case since the compounding process itself occurs on a type of extrusion line, making this a prime example of a case where Products from previous processes are turned into Educts for the following process in a different ProductionRun. In addition, many extrusion lines are equipped with gravimetric systems that supply multiple educts to the same machine. Typical educt-specific data include safety data sheets, material data from offline lab characterizations, compound recipes, or batch IDs. Information can be collected in conformation with EUROMAP 86 [43].

In each ProductionRun, a specified number of products is created in one continuous process. The created products usually belong to one ProductionOrder, and stopping continuous production stops the ProductionRun in question. The Products of the extrusion process can vary greatly depending on the individual process, both in production time per part and individual part size. This is, in part, determined by the throughput of the extruder or the haul-off speed, but also by the frequency with which the extrudate is cut into individual units. For example, one piece of an extruded profile is extruded over a short period of time compared to a roll of blown film but over a long period of time compared to individual granulate pellets that were produced for subsequent processes (e.g. IM, compression moulding). As a result, the amount and resolution of product-specific data with respect to the denominator depends on the process in question.

Extrusion processes have no universal standards for product-specific data structures. Therefore, all data collected for products must be documented carefully and extensively with appropriate metadata that is understandable across different environments. Examples of typical product-specific data in extrusion are inline characterizations of the extrudate geometry, color, product IDs, etc., but also offline analyses of smaller product batches in a lab environment, e.g., for the characterization of crystallinity, orientations, or mechanical properties.

4.3 Injection molding

4.3.1 Denominator description

The IM process is a discontinuous process, meaning that a natural denominator choice is the duration of the repetition of variable state changes, called a machine cycle. Based on the provided model, the denomination of these cycles can be integrated within a ProductionRun. In the progress of one cycle, the discontinuous IM process covers all parameter state changes that occur during the iterative conduction of process cycles while assuming that the process is fully thermally balanced and not subject to erroneous state changes

due to ramp up process or external influences. Each *educt*- and *ProductInstance*, such as the used material batch during production or the produced parts during one cycle can therefore be unambiguously mapped to the correct cycle. The *boundaryDenominator* may be used to store the integer representation of a timestamp which indicates the start time of production. The sequence of denominated machine cycles then unambiguously defines the progress of a ProductionRun.

4.3.2 Assets and machine-specific data structures

The detail of asset and machine-specific data structures must satisfy the need to unambiguously map emitted data for the ProductionRun to the emitting asset or asset-component. On the other hand, the model should retain a high level of intuitiveness for the subsequent users, engulfing engineers, applicants and only peripheral data modelers. Therefore, the modeling of the IM machine, peripheral assets and the mold are analyzed. The typical separation of the IM machine core components into machine bed, plasticizing unit, clamping unit and control [77] is not suitable for modeling here. Machine bed and clamping unit do not usually get assembled in varying constellation depending on the process. The plasticizing unit, on the other hand, can be exchanged before production if required, for example if a lower shot volume is needed. The change may directly impact the available data due to a differing number of temperature sensors in the heating bands. Accordingly, we propose to divide the IM machine into the plasticizing unit, machine control and base machine. Apart from the machine, several types of peripheral assets can be modeled, such as tempering units, dryers, dosing units, scales, cameras and many more. These peripherals can be assumed to either offer data interfaces themselves, be readable by hardwiring or standardized interface within the IM machine. Another intricate part of the IM process is the IM mold. The mold itself can be further dissected into a base mold and exchangeable inserts for different part geometries. Each base mold or insert can be equipped with additional sensor technology to measure process parameters such as temperature or pressure. The base mold - insert relationship is of character $n:m$, intending that n base molds may be equipped with m inserts. In many cases, however, inserts are fit to be installed in exactly one base mold. Furthermore, the sensors are commonly not exchanged between molds and inserts. For simplicity, base mold, insert and sensor are not modeled separately. Instead, a *Mold* is modeled as the assortment of specific (base) mold, insert and sensor equipment used within a ProductionRun and able to generate a certain set of data. This data is not subject to polymorphism, making data storage and semantic mapping

efficient. The disadvantage of multiple *Mold* instances with duplicate inherent components outweighs the added complexity to the data model if modeled separately.

4.3.3 Educt- and product-specific data structures

Common ingredients of the used material mixture entail the polymer, color pigments, flow stabilizers, anti-statics or fibers for reinforcement. The material is, by nature, a passive asset, i.e. does not actively emit data. A separate modeling of all ingredients would only be rectified to avoid data repetition when using an abundance of different mixtures. For simplicity reasons the plain material mixtures could be modeled as an entity. The products per machine cycle are typically between one for a single cavity mold and up to 96 or even more in very specialized cases such as the production of PET bottle caps. Each product instance resulting from the production run is described by its mechanical properties, dimensions, and weight. A wide range of measurement methods can be applied in a regressive manner, ranging from scalar outputs (1-dimensional) to images or even videos (3- or 4-dimensional). Commonly used measurements are scalar in nature and describe discrete distances between components to evaluate shrinkage, warpage, or part weight. During a cycle, the machine logs time-series data, such as the progression of the injection curve or the clamping force. These measurements are 2-dimensional consisting of the relative time of measurement during the cycle and the measured value.

4.4 Fine blanking

4.4.1 Denominator description

Fine blanking is a cyclic process that produces multiple identical or varied parts per stroke. The denominator is defined as one stroke, providing a straightforward reference point. While a stroke consists of distinct process phases that are valuable for detailed analysis, using a more complex denominator is unnecessary for establishing a standard reference. Segmentation details can instead be stored as meta-data within the measurement class.

4.4.2 Assets and machine-specific data structures

The fine blanking periphery comprises two primary machines: the roller leveler and the fine blanking machine. Both are designated as Assets, monitored, and controlled via Interface entities. Various sensors, such as force sensors and acoustic emission sensors, are also included as Assets. Force, for example, is both a settable parameter via the machine interface and an observable parameter through

sensors. During each stroke, time series data is recorded with machine and sensor measurements in order to characterize the fine blanking process. The different parts of the fine blanking tool can be also included as Assets, giving information of the wear state of the tool.

4.4.3 Educt- and product-specific data structures

Educts for fine blanking include coils and lubricants. While lubricant data is typically limited to product specifications, coil properties can be measured non-destructively at specific locations. These measurements can then be mapped to specific stroke locations, allowing for correlation with stroke data through the defined denominator. Product quality, measured inline or offline, can similarly be linked to strokes when the parts produced are associated with their respective strokes.

A *ProductionOrder* refers to a set of identical or distinct parts produced using a fine blanking machine. Since different parts can be manufactured with the same tool, an order may include a variety of parts. A *ProductionRun* is defined as a continuous series of strokes without interruption, which is crucial for subsequent data analysis since any downtime may affect sensor calibration and data integrity.

4.5 Hot rolling

4.5.1 Denominator description

The denominator range in the data model for the hot rolling process spans one complete process cycle. In one cycle, material is processed through multiple so-called passes. As a result, the material experiences multiple height reductions until the final desired thickness is achieved. Every pass increments the denominator. On completion of one process cycle, the denominator gets reset.

4.5.2 Educt and product-specific data structures

The Educts are the raw materials required for the hot rolling process, typically cast metal blocks. The cast slabs are then heated in a reheating furnace. These furnaces can process one or multiple cast slabs simultaneously. The heated cast slabs then represent the *EductInstance* for the hot rolling process *ProductionRun*.

One *ProductionOrder* consist of a set number of hot rolled strip batches as ordered by the customer. Further, we limit a single *ProductionOrder* to one set material. To fulfill a *ProductionOrder*, processing of multiple cast slabs can be required. Therefore, we define a *ProductionRun* as the process of achieving the full height reduction for a cast metal slab, subsequent to heat treatment of the cast slab.

Each pass in the hot-rolling process leads to a set height reduction of the cast slab, in turn causing substantial changes in the material's structure. Accordingly, after every pass, a new `ProductInstance` is created, incrementally capturing the material transformation upon achieving the final height.

The final height of the hot rolled strip is the `Product` that is finally achieved at the end of one `ProductionRun`.

4.5.3 Assets and machine-specific data structures

Assets refer to the rolling machines utilized in the hot rolling process, encompassing both roughing mills and finishing mills. Following the heating process, the material first passes through the roughing mill, where it undergoes significant thickness reduction, transforming into a thin slab. To achieve the final geometry and ensure high surface quality, the material is subsequently processed in the finishing mills. These finishing mills consist of multiple successive stands, which further reduce the thin slab into a narrow strip.

The configuration of the roughing and finishing mills may include two to several rollers, depending on the desired final thickness of the material.

4.6 Ultra-short pulse laser ablation

4.6.1 Denominator description

USP laser ablation is a 2.5D ablation process that works layer by layer. In a production step a layer of vectors is loaded onto the machine and these vectors are afterwards ablated onto the material with a scanning device. By repeating this layer-wise ablation over and over a 2.5D structure is formed. The size of the `denominator` in this case was therefore chosen to be a single layer of the process of a single product. During production runs, a layer typically is observed by multiple different sensor systems which map corresponding spatial resolved measurement data on the vectors of a single production layer. Typical sensors for process observation are spectroscopes for backscattered light and thermal emissions, optical coherence tomography and confocal sensors for depth measurement or white light interferometry for metrology analysis. All of these sensor measurements can be mapped to a certain point in our layer. The chosen denominator therefore allows the three-dimensional reconstruction of the ablation process and its corresponding measurements at a certain point in the workpiece.

4.6.2 Assets and machine-specific data structures

The Assets in this use-case are chosen to be modeled at the subsystem level. The machine itself is composed of a workpiece movement system, a scanning device for laser

focus placement, a laser source and several sensors. Adding each of these sub systems as a single Asset allows the data structure to be flexible and seamlessly adapted to include new components. New sensors can be added on demand and other components like beam shapers, which control the laser focal distribution could be upgraded or modified without breaking the data structure. This allows for a flexible system that can be tuned onto specific materials and products, since these require different quality inspection sensors.

4.6.3 Educt- and product-specific data structures

The `Product` of a USP ablation machine is a 2.5D generated structure on a workpiece. In theory, a workpiece holder could incorporate multiple workpieces which an ablation system would machine in successive order. However a single workpiece could be cut into multiple pieces and hold multiple products. In the use-case, the `ProductionRun` was defined as a single ablated layer. Therefore a `ProductionRun` always points to a single `EductInstance` and a single `ProductInstance`. The `Product`, `Educt`, `ProductInstance` and `EductInstance` are defined as single regions on a single workpiece. The `ProductInstance` of a `ProductionRun` layer becomes the `EductInstance` of the next run.

5 Discussion

5.1 Model positioning and architectural implications

The proposed data architecture aims to provide a process-independent reference structure for representing manufacturing data across heterogeneous production environments and thereby complements Industry 4.0 concepts such as CPPS, AASs, digital twins, digital shadows, and ontologies, which primarily address system architectures, asset representations, or semantic interoperability. In contrast to process-specific digital shadow implementations, the model abstracts production systems through a generic denominator, asset hierarchy, and unified measurement structure, enabling consistent representation of production, educt, and product relations independent of the underlying process physics.

Compared to the DSRM [69], which focuses on lifecycle-oriented structuring and semantic integration, the present approach emphasizes structural generalization across fundamentally different manufacturing paradigms. The denominator concept serves as a unifying abstraction layer that links production runs, machine states, and product instances without relying on process-specific temporal assumptions. The modeled `ProductInstance` and `EductInstance` represent

individual digital shadows of parts from a specific process. Through the presented modeling approach, the collected *dataTraces* of these digital shadows are extended by including a reference not only to the production system but also the processing phase. This simplifies the task of transferring knowledge e.g. from one machine system to another.

Special emphasis is put on ensuring compatibility with lower and higher order Models from the IoP. Through the fairly generic Measurement data structure, the DMM [71], containing time-series as well as event data can be seamlessly connected with the herein developed, superimposed process model through both the measurement and origin timestamps.

A possible limitation of the model appears when a process exhibits multiple phases with greatly varying time frames. This could be evident in multi-stage manufacturing processes such as hot forging with intermediate reheating cycles, where short deformation phases are interrupted by significantly longer heating periods. In this case, the denominator's resolution is either too low for one process phase or too high for the other process phase.

Further, it has to be noted that multiple, domain-specific approaches to enable generic process data management are currently in development or are already released, e.g., OPC UA Companion Specifications (IM, Machine Tools, Additive in development). A detailed screening of all such approaches is out of this work's scope. Therefore, coordinating the structure and implementing the presented model could ensue challenges and require model adaptations.

Lastly, there exists only very limited research on cross-domain data collaboration. Preliminary, basic research in the IoP *Cluster of Excellence* [3] shows the high potential of such a collaboration due to the thereof exponentially increasing available production data. However, semantic synchronization still remains a challenge. The herein developed data model aims to set the basis for such collaboration, but is expected to require extensions and/or modifications with increasing research work.

5.2 Validation scope and limitations

The presented use cases constitute a structural validation of the proposed data model by instantiating it across fundamentally different manufacturing paradigms, rather than a performance or implementation validation. Validation is performed through systematic instantiation across six fundamentally different manufacturing paradigms, including continuous and discontinuous, cyclic and layer-based processes. In each case, denominator definition, asset structure, educt-product relations, and measurement representations were mapped without modifying the core meta-model. This consistent applicability across heterogeneous production environments demonstrates the architectural generality of

the approach. In particular, the validation focuses on the conceptual alignment between process information and process structure, as well as on assessing the internal coherence and representational expressiveness of the proposed data model.

This work does not provide deployment-level or performance-based validation. In particular, no cross-site implementation, runtime analysis, or interoperability stress testing was conducted. Such validation would require the integration of heterogeneous IT infrastructures across multiple institutes and industrial environments and is therefore beyond the scope of this architectural contribution.

The claimed generalization capability thus refers to consistent data-structural representation across domains. Future work will address implementation-driven validation, including interoperability studies and complexity analysis under real production data conditions.

6 Conclusion and outlook

In this paper, a unified data model was proposed to address the challenges of data heterogeneity and limited cross-domain reusability in manufacturing processes. By defining interoperable data formats and a generic process-independent denominator, the model facilitates a standardized approach to describe production, machine, and quality data across a variety of manufacturing sectors. This model was structurally instantiated by applying it onto six diverse manufacturing processes, showing its versatility and potential for applicability with respect to data representation and semantic alignment. The ability to integrate data spanning from production runs to individual measurement points provides a robust framework for advancing digital production environments.

The flexibility to define a custom denominator effectively manages the trade-offs between data resolution, computational complexity and storage requirements, granting the possibility to optimize data management not only process-but also domain-specifically, such as series production, industrial research and development or academic research.

By integrating generic educt- and product-specific data structures, the transformative nature of production processes can be adequately represented, enabling cause-effect analyses of resulting product quality. Compatibility with the DMM of IoP is ensured by the inclusion of a fairly generic structure for measurement data.

Moving forward, the adoption of the proposed data model presents opportunities for accelerating the development of digital shadows in manufacturing. Future work should focus on implementation-driven validation in real-world manufacturing environments, exploring its impact on production efficiency, data-driven process optimization, and cross-domain collaboration.

Acknowledgements Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC-2023 Internet of Production - 390621612.

Author Contributions J.M., L.O., A.P., J.S., M.K., T.M.: Conceptualization, Model Development, Writing & Editing; P.N., T.B., M.E., C.B., Y.L., M.See., M.Sch., C.H., N.T., D.B., E.S.: Supervision, Validation.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data Availability The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Peng G, Cheng Y, Zhang Y, Shao J, Wang H, Shen W (2022) Industrial big data-driven mechanical performance prediction for hot-rolling steel using lower upper bound estimation method. *J Manuf Syst* 65:104–114. <https://doi.org/10.1016/j.jmsy.2022.08.014>
- Tao F, Qi Q, Liu A, Kusiak A (2018) Data-driven smart manufacturing. *J Manuf Syst* 48:157–169. <https://doi.org/10.1016/j.jmsy.2018.01.006>. Special Issue on Smart Manufacturing
- Pennekamp J, Glebke R, Henze M, Meisen T, Quix C, Hai R, Gleim L, Niemietz P, Rudack M, Knape S et al (2019) Towards an infrastructure enabling the internet of production. In: 2019 IEEE international conference on industrial cyber physical systems (ICPS). IEEE, pp 31–37
- Brecher C, Schuh G, van der Aalst W, Jarke M, Piller FT, Padberg M (eds) (2024) The internet of production: interdisciplinary visions and concepts for the production of tomorrow, 1st ed. 2024 edn. Interdisciplinary excellence accelerator series. Springer International Publishing and Imprint Springer
- Schuh G, Häfner C, Hopmann C, Rumpe B, Brockmann M, Wortmann A, Maibaum J, Dalibor M, Bibow P, Sapel P et al (2020) Effizientere produktion mit digitalen schatten. *Zeitschrift für wirtschaftlichen Fabrikbetrieb* 115(s1):105–107
- Idzik C, Hilger D, Hosters N, Kemmerling M, Niemietz P, Ortjohann L, Sasse J, Serafeim A, Wang J, Wolff D et al (2023) Decision support for the optimization of continuous processes using digital shadows. In: Internet of production: fundamentals, methods and applications. Springer, pp 281–301
- Ribeiro L (2017) Cyber-physical production systems' design challenges. In: 2017 IEEE 26th International symposium on industrial electronics (ISIE), pp 1189–1194. <https://doi.org/10.1109/ISIE.2017.8001414>
- Thiede S, Juraschek M, Herrmann C (2016) Implementing cyber-physical production systems in learning factories. *Procedia CIRP* 54:7–12. <https://doi.org/10.1016/j.procir.2016.04.098>. 6th CIRP Conference on Learning Factories
- DIN Deutsches Institut für Normung eV (2016) DIN SPEC 91345:2016-04: Referenzarchitekturmodell Industrie 4.0 (RAMI4.0)
- Pallasch C, Wein S, Hoffmann N, Obdenbusch M, Buchner T, Walzl J, Brecher C (2018) Edge powered industrial control: Concept for combining cloud and automation technologies. In: 2018 IEEE International conference on edge computing (EDGE), pp 130–134. <https://doi.org/10.1109/EDGE.2018.00026>
- Fimmers C, Blanke P, Wiecezorek M, Petrovic O, Herfs W (2023) Asset administration shells in tool lifecycle monitoring. *Procedia CIRP* 120:1209–1214. <https://doi.org/10.1016/j.procir.2023.09.150>. 56th CIRP International Conference on Manufacturing Systems 2023
- Fuller A, Fan Z, Day C, Barlow C (2020) Digital twin: enabling technologies, challenges and open research. *IEEE Access* 8:108952–108971. <https://doi.org/10.1109/ACCESS.2020.2998358>
- Cimino C, Negri E, Fumagalli L (2019) Review of digital twin applications in manufacturing. *Comput Ind* 113:103130. <https://doi.org/10.1016/j.compind.2019.103130>
- Kritzinger W, Karner M, Traar G, Henjes J, Sihn W (2018) Digital twin in manufacturing: a categorical literature review and classification. *IFAC-PapersOnLine* 51(11):1016–1022. <https://doi.org/10.1016/j.ifacol.2018.08.474>. 16th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2018
- AboElHassan A, Sakr AH, Yacout S (2023) General purpose digital twin framework using digital shadow and distributed system concepts. *Comput Ind Eng* 183:109534. <https://doi.org/10.1016/j.cie.2023.109534>
- Chehrehzad M, Kecibas G, Besirova C, Uresin U, Irican M, Lazoglu I (2024) Tool wear prediction through ai-assisted digital shadow using industrial edge device. *J Manuf Process* 113:117–130. <https://doi.org/10.1016/j.jmapro.2024.01.052>
- Bergs T, Gierlings S, Auerbach T, Klink A, Schraknepper D, Augspurger T (2021) The concept of digital twin and digital shadow in manufacturing. *Procedia CIRP* 101:81–84. <https://doi.org/10.1016/j.procir.2021.02.010>. 9th CIRP Conference on High Performance Cutting
- Sapel P, Gannouni A, Fulterer J, Hopmann C, Schmitz M, Lüticke D, Gützlaff A, Schuh G (2022) Towards digital shadows for production planning and control in injection molding. *CIRP J Manuf Sci Technol* 38:243–251. <https://doi.org/10.1016/j.cirpj.2022.05.003>
- Weckx S, Robyns S, Baake J, Kikken E, Geest RD, Birem M, Maes D (2022) A cloud-based digital twin for monitoring of an adaptive clamping mechanism used for high performance composite machining. *Proc Comput Sci* 200:227–236. <https://doi.org/10.1016/j.procs.2022.01.221>. 3rd International Conference on Industry 4.0 and Smart Manufacturing
- Adeagbo MO, Wang S-M, Ni Y-Q (2024) Revamping structural health monitoring of advanced rail transit systems: a paradigmatic shift from digital shadows to digital twins. *Adv Eng Inform* 61:102450. <https://doi.org/10.1016/j.aei.2024.102450>
- Rodríguez-Aguilar R, Marmolejo-Suacedo J-A, Rodríguez-Aguilar M, Marmolejo-Suacedo L (2024) Machine learning for digital shadow design in health insurance sector. *Mobile Networks and Applications* <https://doi.org/10.1007/s11036-023-02289-2>

22. Osman AH, Artus M, Alsaad H, Koch C, Voelker C (2023) Assessing a thermoelectric radiative cooling partition as a personalised comfort system using empirical experiments enhanced by digital shadow visualisation. *Build Environ* 245:110833. <https://doi.org/10.1016/j.buildenv.2023.110833>
23. Berners-Lee T, Hendler J, Lassila O (2001) The semantic web. *Sci Am* 284(5):34–43
24. Legat C, Seitz C, Lamparter S, Feldmann S (2014) Semantics to the shop floor: towards ontology modularization and reuse in the automation domain. *IFAC Proc Vol* 47(3):3444–3449. <https://doi.org/10.3182/20140824-6-ZA-1003.02512>
25. Michael J, Koren I, Dimitriadis I, Fulterer J, Gannouni A, Heithoff M, Hermann A, Hornberg K, Kröger M, Sapel P, Schäfer N, Theissen-Lipp J, Decker S, Hopmann C, Jarke M, Rumpe B, Schmitt RH, Schuh G (2023) A digital shadow reference model for worldwide production labs. In: Brecher C, Schuh G, van der Aalst W, Jarke M, Piller FT, Padberg M (eds) *Internet of production*. In: *Interdisciplinary excellence accelerator series*. Springer International Publishing, pp 1–29. https://doi.org/10.1007/978-3-030-98062-7_3-3
26. Stokes-Griffin CM, Compston P (2015) The effect of processing temperature and placement rate on the short beam strength of carbon fibre-peek manufactured using a laser tape placement process. *Compos A Appl Sci Manuf* 78:274–283. <https://doi.org/10.1016/j.compositesa.2015.08.008>
27. Schulz MB (2024) Digital twin methodology for the automated fiber placement with in-situ consolidation of thermoplastic tape; 1. Auflage. RWTH Aachen University. <https://doi.org/10.18154/RWTH-2024-00971>
28. Khodaei A, Shadmehri F (2022) Intimate contact development for automated fiber placement of thermoplastic composites. *Composites Part C Open Access* 8:100290. <https://doi.org/10.1016/j.jc.2022.100290>
29. Schäkel M, McNab J, Dodds N, Peters T, Janssen H, Brecher C (2018) Data collection and analysis for the creation of a digital shadow during the production of thermoplastic composite layers in unbonded flexible pipes. In: *International conference on off-shore mechanics and arctic engineering*, vol 5: Pipelines, Risers, and Subsea Systems, pp 005–04001. <https://doi.org/10.1115/OMAE2018-77011>
30. Schulz M, Janssen H, Brecher C (2019) A digital shadow for the infrared-based tape laying process of tailored blanks out of thermoplastic unidirectional tape. *Procedia CIRP* 85:224–229. <https://doi.org/10.1016/j.procir.2019.09.035>. 2nd CIRP Conference on Composite Material Parts Manufacturing, 10–11 October 2019, Advanced Manufacturing Research Centre, UK
31. Schulz M, Survila E, Janssen H, Brecher C (2021) A digital twin for estimating process quality during automated fiber placement of thermoplastic composites. In: *[SAMPE EUROPE Conference and Exhibition 2021 Baden/Zürich - Switzerland]*
32. Oromiehie E, Prusty BG, Compston P, Rajan G (2019) Automated fibre placement based composite structures: Review on the defects, impacts and inspections techniques. *Compos Struct* 224:110987. <https://doi.org/10.1016/j.compstruct.2019.110987>
33. Fu Y, Yao X (2022) A review on manufacturing defects and their detection of fiber reinforced resin matrix composites. *Composites Part C Open Access* 8:100276. <https://doi.org/10.1016/j.jc.2022.100276>
34. Schuster A, Mayer M, Willmeroth M, Brandt L, Kupke M (2020) Inline quality control for thermoplastic automated fibre placement. *Procedia Manuf* 51:505–511. <https://doi.org/10.1016/j.promfg.2020.10.071>. 30th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2021)
35. Brasington A, Godbold M, Harik R (2024) Artificially intelligent process planner for automated fiber placement. *Manuf Lett* 40:143–145. <https://doi.org/10.1016/j.mfglet.2024.03.020>
36. Francis B, Brasington A, Godbold M, Harik R (2024) A digital environment for simulating the automated fiber placement manufacturing process. *Manuf Lett* 40:146–149. <https://doi.org/10.1016/j.mfglet.2024.03.008>
37. Brasington A, Halbritter J, Godbold M, Kirkpatrick M, Sacco C, Harik R (2023) Mapping of multimodality data for manufacturing analyses in automated fiber placement. *Compos B Eng* 263:110881. <https://doi.org/10.1016/j.compositesb.2023.110881>
38. Rosato DV, Rosato MG (2012) *Injection molding handbook*, 2th edn. Springer
39. Becker F, Bibow P, Dalibor M, Gannouni A, Hahn V, Hopmann C, Jarke M, Koren I, Kröger M, Lipp J, Maibaum J, Michael J, Rumpe B, Sapel P, Schäfer N, Schmitz G.J, Schuh G, Wortmann A (2021) A conceptual model for digital shadows in industry and its application. In: Ghose A, Horkoff J, Silva Souza VE, Parsons J, Evermann J (eds) *Conceptual modeling. Lecture Notes in Computer Science*. Springer International Publishing, vol 13011, pp 271–281. https://doi.org/10.1007/978-3-030-89022-3_22
40. Zambrano V, Mueller-Roemer J, Sandberg M, Talasila P, Zanin D, Larsen PG, Loeschner E, Thronicke W, Pietraroia D, Landolfi G, Fontana A, Laspalas M, Antony J, Poser V, Kiss T, Bergweiler S, Pena Serna S, Izquierdo S, Viejo I, Juan A, Serrano F, Stork A (2022) Industrial digitalization in the industry 4.0 era: classification, reuse and authoring of digital models on digital twin platforms. *Array* 14:100176. <https://doi.org/10.1016/j.array.2022.100176>
41. Sapel P, Hopmann C (2023) Towards an ontology-based dictionary for production planning and control in the domain of injection molding as a basis for standardized asset administration shells. *J Ind Inf Integr* 35:100488. <https://doi.org/10.1016/j.jii.2023.100488>
42. Sapel P, Molinas Comet LT, Dimitriadis I, Hopmann C, Decker S (2024) A review and classification of manufacturing ontologies. *J Intell Manuf* 1–25. <https://doi.org/10.1007/s10845-024-02425-z>
43. European Plastics and Rubber Machinery: EUROMAP OPC UA Specifications. <https://www.euromap.org/i40/OPCUA>
44. Silva B, Marques R, Faustino D, Ilheu P, Santos T, Sousa J, Rocha AD (2023) Enhance the injection molding quality prediction with artificial intelligence to reach zero-defect manufacturing. *Processes* 11(1):62. <https://doi.org/10.3390/pr11010062>
45. Ramírez-Durán VJ, Berges I, Illarramendi A (2020) ExtruoInt: An ontology for describing a type of manufacturing machine for industry 4.0 systems. *Semantic Web* 11(6):887–909. <https://doi.org/10.3233/sw-200376>
46. Klocke F (2017) *Fertigungsverfahren* 4. Springer, Berlin Heidelberg
47. Unterberg M, Becker M, Niemietz P, Bergs T (2024) Data-driven indirect punch wear monitoring in sheet-metal stamping processes. *J Intell Manuf* 35(4):1721–1735
48. Liewald M, Bergs T, Groche P, Behrens B-A, Briesenick D, Müller M, Niemietz P, Kubik C, Müller F (2022) Perspectives on data-driven models and its potentials in metal forming and blanking technologies. *Prod Eng Res Devel* 16(5):607–625
49. Ortjohann L, Becker M, Niemietz P, Bergs T (2024) Material-data-driven prediction of sheared surface tears of fine blanked parts. *Mater Res Proc* 41
50. Molitor DA, Kubik C, Knoll M, Becker M, Groche P (2021) Ableitung eines vorgehensmodells zur systematischen wissensgenerierung aus sensordaten: Zustandsbewertung produktionstechnischer prozesse mittels maschineller lernverfahren. *Zeitschrift für wirtschaftlichen Fabrikbetrieb* 116(5):352–357
51. Kubik C, Knauer SM, Groche P (2022) Smart sheet metal forming: importance of data acquisition, preprocessing and transformation on the performance of a multiclass support vector machine for predicting wear states during blanking. *J Intell Manuf* 33(1):259–282

52. Zhou C, Zhang F, Wei B, Lin Y, He K, Du R (2021) Digital twin-based stamping system for incremental bending. *Int J Adv Manuf Technol* 116(1):389–401
53. Peinado-Asensi I, Montés N, García E (2024) Stamping process analysis in an industrial plant and its limitations to obtain an industrializable continuous twin. *Int J Mater Form* 17(1):12
54. Song C, Shen X, Xia J (2023) A digital twin model for automatic width control of hot rolling mill. *IEEE Access* 1–1. <https://doi.org/10.1109/ACCESS.2023.3306782>
55. Younes MA, Shahtout M, Damir MN (2006) A parameters design approach to improve product quality and equipment performance in hot rolling. *J Mater Process Technol* 171(1):83–92
56. Idzik C, Krämer A, Hirt G, Lohmar J (2024) Coupling of an analytical rolling model and reinforcement learning to design pass schedules: towards properties controlled hot rolling. *J Intell Manuf* 35(4):1469–1490
57. Cui C, Cao G, Li X, Gao Z, Zhou X, Liu Z (2022) The coupling machine learning for microstructural evolution and rolling force during hot strip rolling of steels. *J Mater Process Technol* 309:117736. <https://doi.org/10.1016/j.jmatprotec.2022.117736>
58. Xie Q, Suvama M, Li J, Zhu X, Cai J, Wang X (2021) Online prediction of mechanical properties of hot rolled steel plate using machine learning. *Mater Des* 197:109201. <https://doi.org/10.1016/j.matdes.2020.109201>
59. Desheng C, Jian S, Mingxin L, Sensen X (2023) Digital twin-based fault diagnosis platform for final rolling temperature in hot strip production. *Materials* 16. <https://doi.org/10.3390/ma16217021>
60. Finger J, Kalupka C, Reininghaus M (2015) High power ultra-short pulse laser ablation of in718 using high repetition rates. *J Mater Process Technol* 226:221–227. <https://doi.org/10.1016/j.jmatprotec.2015.07.014>
61. Jaeggi B, Neuenschwander B, Schmid M, Muralt M, Zuercher J, Hunziker U (2011) Influence of the pulse duration in the ps-regime on the ablation efficiency of metals. *Phys Procedia* 12:164–171. <https://doi.org/10.1016/j.phpro.2011.03.118>. *Lasers in Manufacturing 2011 - Proceedings of the Sixth International WLT Conference on Lasers in Manufacturing*
62. Zuric M, Nottrodt O, Abels P (2019) Multi-sensor system for real-time monitoring of laser micro-structuring. *J Laser Micro/Nanoeng* 14(3)
63. Kröger M, Lasogga C, Kratz M, Hinke C, Holly C (2024) Platform for adaptive integration of data-driven models and simulations into ultra short pulse manufacturing systems. *J Laser Micro/Nanoeng* 19(1)
64. Ultrashort Pulse Lasers & Micromachining | Pulsar Photonics — pulsar-photonics.de. <https://www.pulsar-photonics.de/en/>. Accessed 09 Dec 2024
65. GF Machining Solutions | Milling - EDM - Advanced Manufacturing — gfms.com. <https://www.gfms.com/com/en.html>. Accessed 09 Dec 2024
66. SCANmotionControl | Scanlab — scanlab.de. <https://www.scanlab.de/en/products/software/scanmotioncontrol>. Accessed 09 Dec 2024
67. Dirks S, Collet A, Schleifenbaum JH (2020) Open vector format for laser-based computer-aided manufacturing. *Progress Additive Manuf* 2020:117–130
68. Brecher C, Dalibor M, Rumpel B, Schilling K, Wortmann A (2021) An ecosystem for digital shadows in manufacturing. *Procedia CIRP* 104:833–838. <https://doi.org/10.1016/j.procir.2021.11.140>
69. Michael J, Koren I, Dimitriadis I, Fulterer J, Gannouni A, Heithoff M, Hermann A, Hornberg K, Kröger M, Sapel P, et al (2023) A digital shadow reference model for worldwide production labs. In: *Internet of production: fundamentals, methods and applications*. Springer, pp 61–89
70. Riesener M, Schuh G, Dölle C, Tönnies C (2019) The digital shadow as enabler for data analytics in product life cycle management. *Procedia CIRP* 80:729–734. <https://doi.org/10.1016/j.procir.2019.01.083>
71. van der Aalst WMP (2024) Experiences from the internet-of-production: Using “data-models-in-the-middle” to fight complexity and facilitate reuse. In: Weerd J, Pufahl L (eds) *Business Process Management Workshops. Lecture Notes in Business Information Processing*. Springer Nature Switzerland, vol 492, pp 87–91. https://doi.org/10.1007/978-3-031-50974-2_7
72. Brecher C, Buchsbaum M, Müller A, Schilling K, Obdenbusch M, Staudacher S, Albasatineh MC (2021) Gaining iiot insights by leveraging ontology-based modelling of raw data and digital shadows. In: *2021 4th IEEE international conference on industrial cyber-physical systems (ICPS)*. IEEE, pp 231–236. <https://doi.org/10.1109/ICPS49255.2021.9468116>
73. Sapel P, Hopmann C (2023) Towards an ontology-based dictionary for production planning and control in the domain of injection molding as a basis for standardized asset administration shells. *J Ind Inf Integr* 35:100488. <https://doi.org/10.1016/j.jii.2023.100488>
74. Tacke Genannt Unterberg L, Koren I, Wil MP, van der Aalst (2024) Maximizing Reuse and Interoperability in Industry 4.0 with a Minimal Data Exchange Format for Machine Data. *Gesellschaft für Informatik eV*
75. Zhang Y, Ding J, Li Y, Ren Z, Feng K (2024) Multi-modal data cross-domain fusion network for gearbox fault diagnosis under variable operating conditions. *Eng Appl Artif Intell* 133:108236. <https://doi.org/10.1016/j.engappai.2024.108236>
76. Peitz A, Striet P, Emonts M, Fischer K, Brecher C (2021) Design of a semantic information model for automated tape placement systems using opc-ua. In: *2021 26th IEEE international conference on emerging technologies and factory automation (ETFA)*, pp 1–8. <https://doi.org/10.1109/ETFA45728.2021.9613216>
77. Hopmann C, Michaeli W (2017) *Einführung in die Kunststoffverarbeitung*, vol 8. Carl Hanser Verlag GmbH & Co, KG

Authors and Affiliations

Jiyoung Moon¹ · Lucia Ortjohann¹ · Philipp Niemietz¹ · Thomas Bergs^{1,2} · Alexander Peitz³ · Michael Emonts³ · Christian Brecher^{2,3,4} · Jana Sasse⁵ · Yannik Lockner^{5,8} · Malte Seefeldt⁵ · Mauritius Schmitz⁵ · Christian Hopmann^{3,5} · Moritz Kröger⁶ · Tamoghna Majumder⁷ · Nilesch Thakare⁷ · David Bailly⁷ · Emad Scharifi⁷

✉ Jiyoung Moon
j.moon@mti.rwth-aachen.de

Lucia Ortjohann
l.ortjohann@mti.rwth-aachen.de

Philipp Niemietz
p.niemietz@mti.rwth-aachen.de

Thomas Bergs
t.bergs@mti.rwth-aachen.de

Alexander Peitz
alexander.peitz@azl.rwth-aachen.de

Michael Emonts
michael.emonts@azl.rwth-aachen.de

Christian Brecher
c.brecher@wzl.rwth-aachen.de

Jana Sasse
jana.sasse@rwth-aachen.de

Yannik Lockner
yannik.lockner@ikv.rwth-aachen.de

Malte Seefeldt
malte.seefeldt@ikv.rwth-aachen.de

Mauritius Schmitz
mauritus.schmitz@ikv.rwth-aachen.de

Christian Hopmann
christian.hopmann@ikv.rwth-aachen.de

Moritz Kröger
moritz.kroeger@ilt.rwth-aachen.de

Tamoghna Majumder
tamoghna.majumder@ibf.rwth-aachen.de

Nilesch Thakare
nilesch.thakare@ibf.rwth-aachen.de

David Bailly
david.bailly@ibf.rwth-aachen.de

Emad Scharifi
emad.scharifi@ibf.rwth-aachen.de

- ¹ Manufacturing Technology Institute (MTI), RWTH Aachen University, Campus-Boulevard 30, Aachen 52074, NRW, Germany
- ² Fraunhofer Institute for Production Technology IPT, Steinbachstraße 17, Aachen 52074, NRW, Germany
- ³ Aachen Center for Integrative Lightweight Production (AZL), RWTH Aachen University, Campus-Boulevard 30, Aachen 52074, NRW, Germany
- ⁴ Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Steinbachstraße 19, Aachen 52074, NRW, Germany
- ⁵ Institute for Plastics Processing (IKV), RWTH Aachen University, Seffenter Weg 201, Aachen 52074, NRW, Germany
- ⁶ Chair for Laser Technology (LLT), RWTH Aachen University, Alois-Riedler-Straße 4, Aachen 52074, NRW, Germany
- ⁷ Institute of Metal Forming (IBF), RWTH Aachen University, Intzestraße 10, Aachen 52072, NRW, Germany
- ⁸ OSPHIM GmbH, Seffenter Weg 201, Aachen 52074, NRW, Germany