



Evaluation of Industry 4.0 Data formats for Digital Twin of Optical Components

Arno Schmetz¹ · Tae Hun Lee¹ · Maximilian Hoeren¹ · Marvin Berger¹ · Susanne Ehret¹ · Daniel Zontar¹ · Soo-Hong Min² · Sung-Hoon Ahn^{2,3} · Christian Brecher^{1,4}

Received: 17 September 2019 / Revised: 5 December 2019 / Accepted: 20 January 2020 / Published online: 2 March 2020
© The Author(s) 2020

Abstract

A wide range of software and hardware components are present in today's production systems and plants using a variety of interfaces and data formats for information exchange on different levels of the system. To increase the traceability, the lifecycle management and providing a single point of source of component-specific data, the Digital Twin technology is proposed, linking different data sets tailored to the requirements of different kind of users (e.g., machines, technicians, logistics, manufacturing execution systems). The data exchange between entities in the manufacturing network relies on machine-readable, flexible and self-describing data formats. When implementing or integrating different components into complex systems, the interoperability challenge is a major concern to address by the system designers and becomes a central task for the creation and integration of Digital Twin technology. In this paper, we evaluate different formats that are used in real environments and create a requirements framework for an ideal format for exchanging flexible and self-describing data in context of optical components manufacturing process and their special requirements.

Keywords Digital twin · Industry 4.0 · Lifecycle support · Smart manufacturing · Sustainable development

1 Introduction

In today's production systems and manufacturing lines, the data emerging from production increases, allowing data-centric extensions. Such extensions to the production can focus the offline and online monitoring of a process and the machine, prediction techniques for quality and wear or the optimization of the production itself to name only some. All these techniques rely on data and need to handle the data accordingly. Understanding provided datasets is a crucial task throughout a digital production but there is no overall accepted and implemented standard for data exchange

formats. While some standards and fairly spread formats exist, features that are not supported by these need to be integrated manually, resulting in an additional set of converters and conversion layers to be implemented [1].

Tackling the challenge of rising global carbon dioxide emissions, research in green manufacturing focused on developing new energy sources, process-oriented reduction of waste and increase of energy efficiency throughout the production. The digitalization of the production enables additional potentials for data-centric optimizations to achieve eco-friendly manufacturing [2, 3]. For the smart manufacturing throughout all sectors, efficient data exchange and energy-efficiency are priority actions and linked together [4].

One key technology for smart manufacturing is the Digital Twin, which merges data from different sources to create a digital representation of processes, products, machines or components. Such a Digital Twin is directly linked to the lifecycle management, enabling new concepts of waste reduction resulting from analyses of the product lifecycle [2, 5].

✉ Arno Schmetz
arno.schmetz@ipt.fraunhofer.de

¹ Fraunhofer Institute for Production Technology IPT, 52074 Aachen, Germany

² Department of Mechanical and Aerospace Engineering, Seoul National University, Seoul 08826, Republic of Korea

³ Institute of Advanced Machines and Design, Seoul National University, Seoul 08826, Republic of Korea

⁴ Laboratory of Machine Tools and Production Engineering (WZL), RWTH Aachen, 52074 Aachen, Germany

2 Data Exchange and Representation

The task of describing data for exchange is not specific to the manufacturing domain. Other domains like biology and agriculture have active research ongoing towards data exchange and data representation, including Synthetic Biology Open Language SBOL [6] and agro Extensible Markup Language agroXML [7] as examples. In this paper, we focus on the domain of manufacturing, especially the special area of optical precision assembly.

With the current shift towards smart manufacturing, efficient data exchange becomes a significant part in the production ramp-up, optimization and resource-efficiency including waste-reduction. Current approaches for data exchange and data representation in such smart manufacturing aim for well-known technical exchange formats in combination with defined rulesets and ontologies.

Ontologies are used to organize knowledge by structuring the information. The aim is to classify knowledge into data, relations and other components to provide a controlled vocabulary for knowledge representation. Originally arising from metaphysics, ontologies became a major research topic in the computer science domain and are used in multiple domains in different shapes [8]. A prominent example of such an ontology is the Dublin Core ontology, describing metadata of digital and physical resources like creator, language and title and became the ISO standard 15836-1:2017 [9].

3 Technology Evaluation

To select suitable and efficient data exchange format for the digital twin for precision assembly of optical component, we define a use-case from real production and extract its requirements to evaluate selected data exchange formats. The data exchange formats are selected based on the recent research approaches of the industry. Doing this, we focus on the data formats, the entangled ontologies schemas and data representations but do not scope data exchange protocols like Simple Object Access Protocol SOAP, REpresentational State Transfer REST, and message-oriented approaches.

3.1 Use-Case Laser Diode Fast-Axis Collimation Optic Assembly

The laser-based applications require precise beam shaping and collimation to achieve required quality. The diode laser systems emit light from their emitters which is not collimated in their vertical and horizontal axis (Fig. 1). In comparison, the divergence in the y-axis is significantly higher

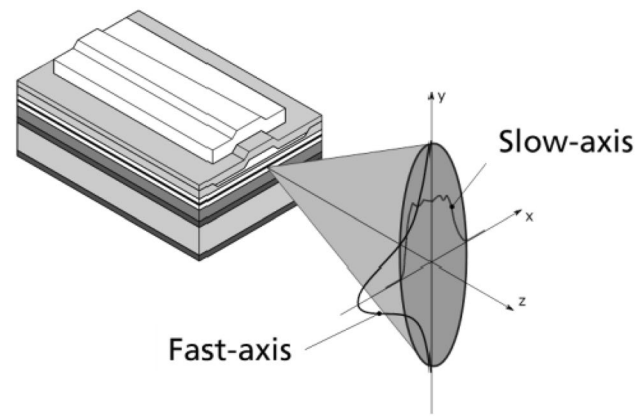
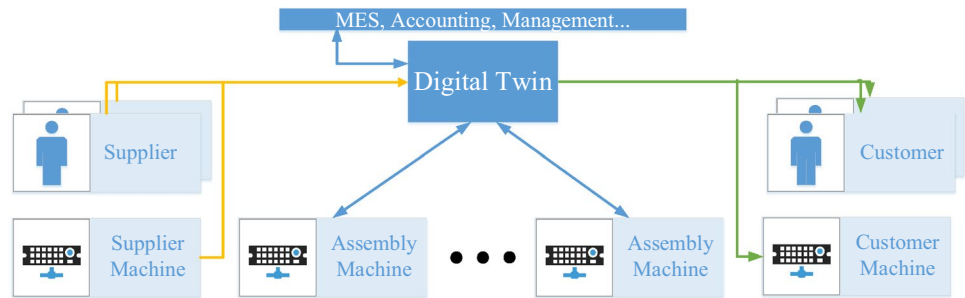


Fig. 1 A high-power diode laser and its emitter characteristics

than the divergence in the x-axis. Due to this observation, the y-axis is called fast-axis while the x-axis is called slow-axis. High power diode laser manufacturers provide specifications for the divergence distribution of the lasers. For example, typical high-power diode laser systems series like Jenoptik JOLD [10] specify the fast-axis divergence for 95% of the emitted light between 40° and 70° while the slow-axis divergence is between 5° and 10° . This shows also the dominance of the fast-axis divergence quantitatively. For applications, such diode lasers need to be assembled and bonded with a fast-axis collimation optic (FAC). For small high precision systems, this task can be done by function-oriented assembly strategies dealing with the tolerances, manufacturing inaccuracies and handling imprecisions [11]. For the assembly and bonding procedure, measurements of optical functions and properties are required inside the assembly line which can consist of multiple machines. Manufacturers who produce or apply the assembled laser with FAC into their products often measure the optical properties after some critical processes to adapt the next production steps accordingly. This procedure is very important for precision optical devices, since some negative effects are not preventable (e.g. adhesive shrinking) [12, 13].

Thus, a Digital Twin for precision optics assembly should aggregate the data from measurements on different steps of the production chain with the properties of the applied components such as laser, FAC lens and be available to provide the users relevant information to adapt their next production steps according to the actual status of the optics. This also enables customers to adapt their production to the actual properties of the assembled systems without the requirement of measuring the components again and delaying the production. This concept is shown in Fig. 2. Suppliers provide data regarding the components (yellow arrows). This can be done directly via an API (Application Programming Interface) to the Digital Twin or by providing data sheets requiring a manual input into the Digital Twin at the assembly company.

Fig. 2 Concept and data flow of the Digital Twin for an assembly use-case with suppliers and customers



The machines along the process chain of the assembly company can access the data, create additional data and merge datasets inside the Digital Twin. In addition to the machines, this kind of access is also available to services like a MES system (blue arrows). The relevant data are made available via the Digital Twin to the customers, their machines or services directly by an API of the Digital Twin (green arrows). By this, the overall production line from first supply to final product of the customers is sped up by providing aggregated datasets for an optimal data pipeline.

3.2 Feature Extraction

Together with an industry company, we identified a set of criteria shown in Fig. 3 for the investigation of different data formats based on the defined use-case:

Human-readable (R1): The data exchange format can be accessed, read and interpreted by humans. This focuses the usability for manual monitoring or fault tracing, not a visualization or graphical interaction with the data. In best case, a format is always human-readable and allows interpretation by a user. In worst case, the format is completely closed, unreadable and can only be accessed with proprietary software.

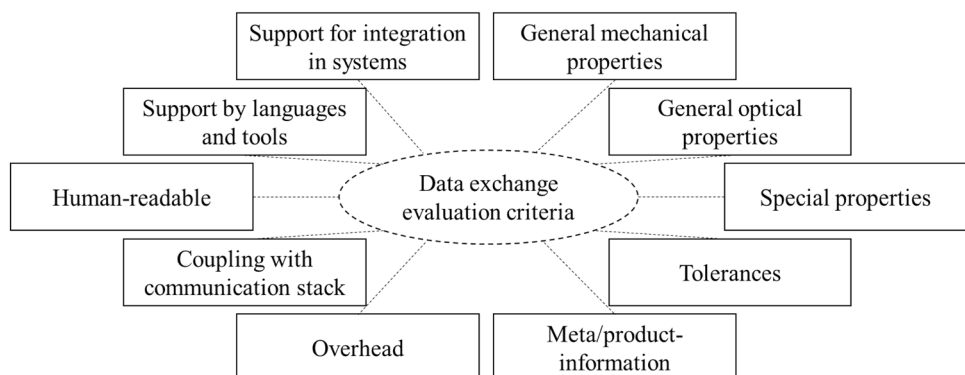
Support by Languages and Tools (R2): The format is supported by programming languages/frameworks and tools to ensure fast and easy interaction. This criterion is important especially for the applications which require low latency process interaction, or applies sensors with

high frequency such as acoustic-emission sensor. Costs like licensing also have impact on this Criterion. In best case, the format comes with a large set of free tools, software libraries for different programming languages and support documentation. In worst case, there are no tools and libraries available besides the actual specification or general information.

Support for integration in (manufacturing) systems (R3): The information can be integrated into existing tools and systems. This is especially important for PLC (Programmable Logic Controller) or Cloud providers targeting the manufacturing industry and being capable of handling incoming data accordingly. In best case, there are existing extensions and plug-ins for existing software systems like MES, PLCs and IIoT (Industrial Internet of Things) platforms or native support for the information provided by the format. In worst case, there is no existing integration available, such that each integration and access has to be provided by the users themselves.

General Mechanical Properties (R4): Sizing, relative positions of parts/subparts and further mechanical-based properties can be modeled in the format including suitable units of measurement (down to nanometer scale). In addition, this includes modeling of a system consisting of subsystems and their assembly and bonding method. In best case, the format allows the modeling of all mechanical properties and the units of measurement and provides a set of shared base models and vocabulary to be used and interpreted by every user of the format. In worst case, no support

Fig. 3 The evaluation criteria for the data exchange formats



for adding such information is available and no method of adding such function is available.

General Optical Properties (R5): Typical properties of optical components like laser power, wavelengths, beam orientations and transmission can be modeled in the format. In best case, the format allows the modeling of all optical properties and the units of measurement. In worst case, no support for adding such information is available and no method of adding such function is available.

Special Properties (R6): Key performance indicators resulting from the function-oriented assembly process, where the assembly steps are not determined by mechanical or relative references but the actual laser collimation output [8], such as the smile (indicating a specific performance pattern in the use-case) of the resulting beam shape can be modeled in this format. In best case, the format supports custom properties that identify key indicators for an object including a description of the indicator. In worst case, no support for adding custom properties is available.

Tolerances (R7): Tolerances of the measurements and mechanical properties can be modeled and adopted. In best case, the format supports the integration of tolerances, its specification and referencing international standards natively. In worst case, no support for tolerances is available at all, including adding self-defined tolerances.

To fulfill the criteria R4 to R7, especially a common implementation of the properties according to existing standards is required to enable the data exchange without further definitions.

Meta/Product-Information (R8): Targeting the exchange of data with the users of the assembled system, the format can be used for defining product-oriented information like product id, product name or serial numbers. In best case, the format defines a separate area for meta-information including native support of meta-information ontologies like Dublin Core. In worst case, the format does not support meta-information natively, requiring the users to specify those information inside the main data manually.

Overhead (R9): For sending a small set of actual information, how much overhead (e.g., additional format-specific data) has to be sent as well. In best case, a data set is represented with no or minimal overhead, resulting in efficient storage usage. In worst case, a data set is represented with big overhead and useless information resulting in large storage requirements and high communication latency due to inefficiency.

Coupling with communication stack (R10): Some data formats are coupled with a certain communication stack or service resulting in tools and resources may focus the communication or service omitting the data modeling aspect. In best case, the format is completely independent from the communication and all tools and resources can be used without focus on communication. In worst case, the format

is tightly coupled and deeply integrated into a complex communication stack, so that the tools, support and usage are communication-centric.

The defined best and worst case scenarios define the range regarding the later evaluation of data formats for these criteria. Based on the implementation of a format, some of the criteria might be linked to each other (like the criteria R4 to R7) or have potential to contradict each other, for example a better human readability could cause additional overhead in the format. These potentials of contrasting criteria requirements and linking of criteria can be addressed in different techniques. The first one is to have a prioritization of the criteria done by the different stakeholders, which is already part of the evaluation methodology described in the next subsections. The other one is the deduction of linking and contrasts of the criteria to model them and their priority inheritance accordingly. As this deduction is depending on stakeholder and use-case, no general quantified deduction can be presented and is omitted in this paper.

3.3 Format Selection

For this evaluation, we selected exchange formats which are present in the manufacturing industry at the moment omitting the exchange formats of other domains. Coming from the domain of computer science, different ontologies and formats did arise and are integrated in the manufacturing-oriented formats. More specialized formats targeting the computer science domain primarily like the Semantic Web [14] are therefore omitted in this paper. Formats being present in the manufacturing industry come with an existing ontology and nomenclature already providing a base for manufacturing-oriented data to be modeled. Computer science oriented formats do not target such behaviors, standards or nomenclatures and would require the users to create the whole ontology for manufacturing on their own. Thus, the support in the manufacturing industry for these computer science oriented formats would be strictly limited. On the other hand, formats of computer science domain may have advantages, if data of different domains e.g. a multi-domain production have to be merged. By that, computer science oriented formats could manage the multi-domain data exchange while manufacturing-oriented formats provide better acceptance in the manufacturing industry, legacy devices integrating parts of the computer science domain concepts. Therefore, we focus on the formats already existing in the manufacturing industry in this paper.

The selected formats for this paper include:

AutomationML (F1): Automation Markup Language is an open XML-based format aiming for the interconnectivity and data exchange between different tools in manufacturing. To achieve this, AutomationML describes a top-level format and a set of sub-formats. Thus, AutomationML integrates

Fig. 4 The structure of AutomationML

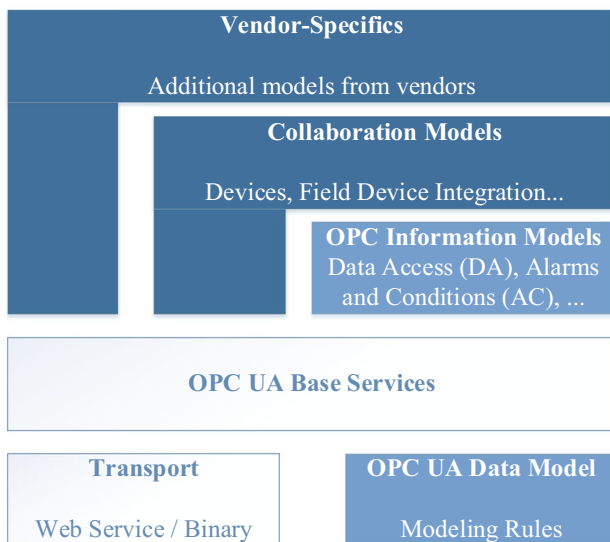
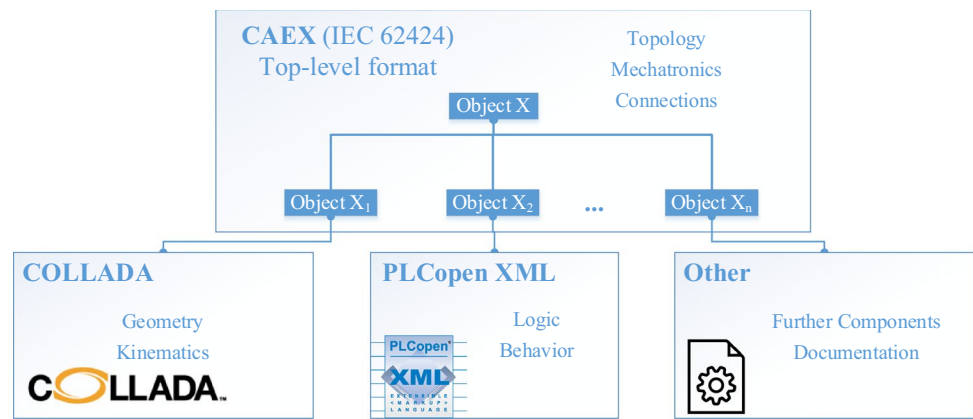


Fig. 5 The structure of OPC UA

different task-specific formats and ontologies as shown in Fig. 4. General topology is described with the Computer Aided Engineering Exchange (CAEX) format, while using COLLADA (COLLABORATIVE Design Activity) and PLCopen for geometric, mechanical and logical information [15].

OPC UA Data Model (F2): OPC UA (Open Platform Communications Unified Architecture) aims to be the standard for data exchange in service-oriented systems. OPC UA is tightly coupled to the services and communication. In contrast to its predecessor OPC, the data model allows self-description by integrating a base ontology and the use of information models. These information models create new ontologies, tailored for special domains inside the manufacturing [16]. This extendibility and the coupling to the communication and services is shown in Fig. 5.

BatchML/B2MML (F3): The Business to Manufacturing Markup Language (B2MML) and the Batch Markup Language (BatchML) implement a set of international standards aiming for the connection of enterprise and control systems.

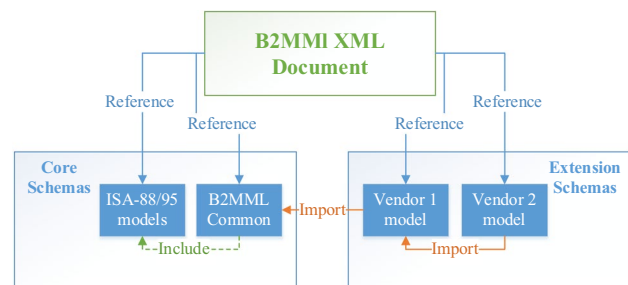


Fig. 6 The structure of B2MML/BatchML

For BatchML/B2MML documents, common and ISA-88/95 models are used and can be extended by further definitions and references. Both languages were merged to enable the usage of general recipes while enabling the data model completely compatible with each other with the option of extending as shown in Fig. 6 [17].

MTConnect (F4): MTConnect is an open XML-based standard especially for the monitoring and analysis of data and connectivity of machines. The information model of MTConnect is focused on machine tools and provides also some extensions for partial interoperability with OPC UA and B2MML. The specification defines general conventions and some specific modeling conventions of machine tools in detail as seen in Fig. 7 [19].

3.4 Evaluation Methodology

The method of evaluation contain qualitative and quantitative aspects while the priority of them may differs depending on the actual role in production line or other optics-use-cases. Therefore, we build up following methodology which can be also applied for other user-cases, exchange formats and ontologies.

Step 1: Stakeholder Requirements

For the evaluation, we take into account different stakeholders in the whole process. The stakeholders in the production are represented by $S = \{S_i | i \in \{1 \dots N_S\}\}$, where N_S is the



Fig. 7 The structure of MTConnect (Fig. from [18])

amount of stakeholders. For each stakeholder, the criteria are evaluated regarding their relevance. The criteria are written as $R = \{R_j | j \in \{1 \dots N_R\}\}$, where N_R is the amount of criteria (for the selection above holds $N_R = 10$). For a criterion R_j and a stakeholder S_i , we can define $A_{i,j} \in \{0,1\}$ indicating the relevance of the criterion (not relevant: 0 and relevant: 1). Analogously, we define the weight $W_{i,j} \in \{0.1, 0.2, \dots, 1.0\}$ indicating the actual priority of criteria to each stakeholder. The sets A and W describe the sets of all $A_{i,j}$, respectively $W_{i,j}$ as described above.

Step 2: Evaluation of criteria

The formats F_k are to be evaluated after the defined criteria defined. This evaluation is quantified by a rating $C_{k,j} \in \mathbb{Q}_{\leq 1}^+$, where j is given by the criteria R_j .

Step 3: Scoring calculation

Based on the rating for all relevant criteria from step 2 and the relevance from step 1, we define a scoring function

$$score(i, k) = \frac{\sum_{R_j \in R} A_{i,j} * W_{i,j} * C_{k,j}}{\sum_{R_j \in R} A_{i,j}}$$

which describes the score of a format F_k for a stakeholder S_i . Finally, the overall score for a format regarding the full use-case is calculated by

$$score(k) = \frac{\sum_{S_i \in S} score(i, k)}{N_S}$$

These scoring functions allow the assessment of data formats regarding a defined use-case with a numerical value $score(i, k) \in \mathbb{Q}_{\leq 1}^+ \wedge score(k) \in \mathbb{Q}_{\leq 1}^+$.

Table 1 The weights of the criteria R_j for the different stakeholders S_i of the use-case

	$S_i \in S$				
	S_1	S_2	S_3	S_4	S_5
$R_j \in R$					
R_1	0.1	0.3	0.3	0.5	0.1
R_2	0.2	0.6	0.2	0.2	0.8
R_3	0.3	0.4	0.7		0.4
R_4	0.8	0.5		0.6	0.5
R_5	0.8	0.8	0.2	0.6	1.0
R_6	0.8	1.0	0.4	1.0	1.0
R_7	0.5	0.5	0.1	0.4	0.3
R_8	0.3	0.3	0.9	1.0	1.0
R_9		0.4	0.8		0.2
R_{10}	0.2	0.7	0.2	0.2	0.6

4 Use-Case Appliance

For the appliance, we use the use-case defined before, with a designed data flow as shown in Fig. 2. The use-case focuses on the optical assembly company and their customers, while the suppliers are omitted.

4.1 Stakeholder Identification

For the use-case, we identified a set of different stakeholders as follows. The quality control engineer S_1 is responsible for the specific data integration of metrology applied for the quality control of the base components (FAC, laser diode) and the assembled system in production (internal access) and product (external access). The assembly engineer S_2 needs to access the data from metrology as well as the production-specific data and the information for the modeling of the bonding/assembling process between subcomponents. The top-level management S_3 has to deal mainly with the meta and product level information as well as the interfacing between existing systems like the MES and Customer Relation Management (CRM). One of the customer of the assembled system S_4 focuses on the packaging of the laser system for further diode laser applications while the other S_5 uses the product as a main component in a complex system. Thus, S_4 and S_5 require the actual tolerances and metrology results of the assembled system to adapt this in their process chain. Using these stakeholders, we derive the relevance and weights for each criteria based on interviews and explicit requests on stakeholders. Table 1 shows the weights derived for the stakeholders, where empty cells denote a relevance value of 0.

For this use-case, the priorities of stakeholder differ from each other. Besides, some criteria are in minor interest of all stakeholders (like R_1) while some are in major interest (like R_6). The other criteria are showing a wide range throughout the stakeholders (like R_8).

4.2 Criteria Assessment

Figure 8 shows the target data model of the use-case represented as UML class diagram. This model contains all required components throughout the process and is extensible for additional information. The product is the central element and consists of sub-components that can also be bonded to each other. This data model is used as a reference for the criteria evaluation.

In the following, we will step briefly through the criteria that were evaluated mentioning the most significant differences or features.

Human-ReadableAll selected data formats support a data representation based on XML and the integration of ontologies using XML schema specification or RDF (Resource Description Framework) and OWL (Web Ontology Language). The OPC UA requires a binary representation, which is not readable for humans, but reduce memory consumption and computational power significantly, since the xml-parsing can be omitted. The XML-based data representation which enables the data human readable is common in OPC UA, but not required. Therefore, the rating of selected data formats is the same for this criterion except OPC UA which is rated than the others.

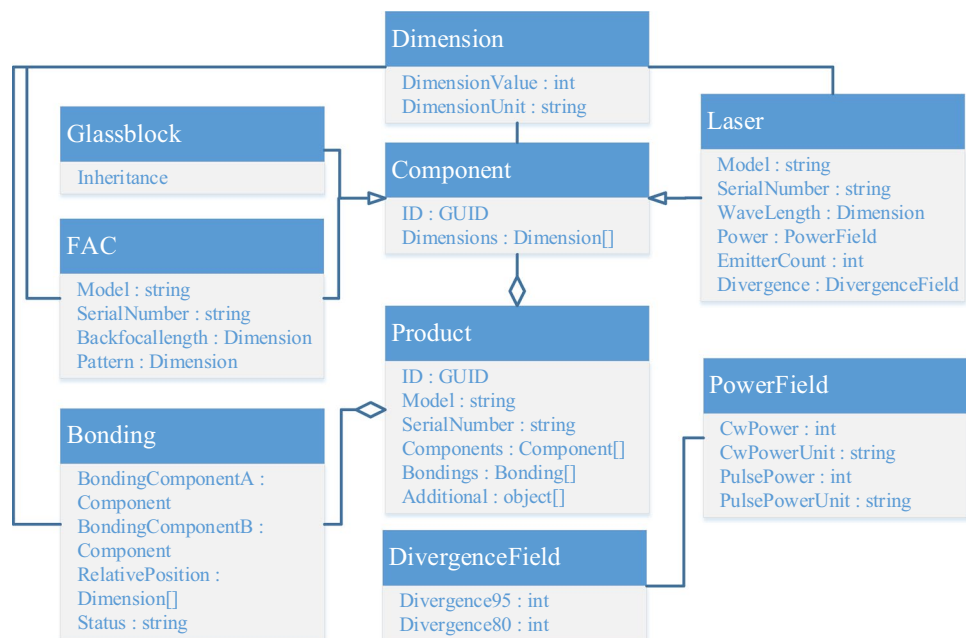
Support by Languages and ToolsAll formats are supported from different programming languages and tools. Typical languages like C#, C++ and Java are supported through official and user sources. Many of the tools and libraries are focusing on the Windows platform. For example, all official AutomationML tools are only targeting Windows. For OPC UA this changed when Microsoft contributed towards the OPC UA implementation adding support

for non-Windows systems. By this, the main tools and libraries of OPC UA are platform-independent allowing the usage for non-Windows environments [20].

Support for Integration in SystemsPLC-level interaction with the data and the integration with the manufacturing management systems of the stakeholders is a task of vertical interaction. An increasing number of PLC manufacturers like Beckhoff and Siemens supports OPC UA, while MTConnect and AutomationML is available limitedly. B2MML has broad support in MES systems and (as well as OPC UA) an increasing support in the field of energy monitoring. By integrating several standards, AutomationML is also indirectly supported by additional systems that use one of the sub-standards inside AutomationML.

Modeling Properties (Mechanical, Optical, Special and Tolerances)All of the formats support mechanical, geometric, kinematic and more properties but with different approaches. While MTConnect focuses on machine tools directly with fixed ontologies, while other formats allow more flexible mappings and ontology definitions, even at runtime. AutomationML by its aggregation of sub-standards strategy could be used for domain specific properties and complete ontologies. For OPC UA, the OPC foundation defines the information models and maintains an official list of common models, which, however, do not include any information of optical components so far. Some of the optical properties may also be modeled inside some sub-standards like COLLADA with the AutomationML, but yet there is no exist of dedicated models for optical components in any of the selected formats. As a result, the extension of properties of optical components, which are not supported directly by the selected formats, can be done in the OPC

Fig. 8 A UML diagram for representation of a target data model in the use-case



UA and AutomationML more easily than the others. The MTConnect allows better management of traditional properties of machine tool applications.

*Meta- and Product Information*The integration of meta and product-specific information is more natively supported by the B2MML than the others, since it focuses in MES and top-level integration. The other formats also support such information for different requirements of different stakeholders, but with complex methods by modeling these information accordingly or integrating product-oriented schemas [21].

*Overhead*The overhead for data formats also depends on the actual implementation and usage schemas. As OPC UA supports a binary data representation, the overhead seems to

be smaller than the other formats. However, the binary representation limits the extendibility of the information models requiring the users to check on the limitations in detail against to custom adaptations. By this, OPC UA provides the possibility to reduce the overhead of the data while reducing the readability and extendibility. AutomationML has a slightly higher overhead on message level since the aggregation of different standards requires an additional layer of abstraction. In total, the formats all have a significant overhead due to their hierarchical modeling and variety of types and subtypes but the differences between the formats are limited.

*Coupling with the Communication Stack*While AutomationML and B2MML can be used mostly independent from communication technology, MTConnect and especially OPC UA have their data model tightly coupled to the communication stack. Therefore, major differences of the rating scores occurred here between the data formats.

The evaluated rating is shown in Table 2.

Table 2 The rating result for the formats against the criteria for this use-case

	$F_k \in F$			
	F_1	F_2	F_3	F_4
$R_j \in R$				
R_1	0.8	0.6	0.8	0.8
R_2	0.6	0.8	0.6	0.6
R_3	0.5	0.8	0.4	0.5
R_4	0.5	0.7	0.5	0.8
R_5	0.5	0.4	0.3	0.2
R_6	0.6	0.6	0.4	0.2
R_7	0.6	0.5	0.6	0.8
R_8	0.6	0.6	0.8	0.6
R_9	0.3	0.6	0.4	0.4
R_{10}	0.9	0.3	0.9	0.5

4.3 Assessment Results

Based on the calculated relevance of criteria for each stakeholder (Table 1), and the ratings (Table 2) above, the scorings are calculated in Fig. 9.

The final scoring function results per format for this use-case are presented in Table 3.

The Fig. 9 and Table 3 show the scorings by stakeholders and the final scoring results for each data format respectively. The highest score among the evaluated items was F2

Fig. 9 The scoring per data format for each of the stakeholders

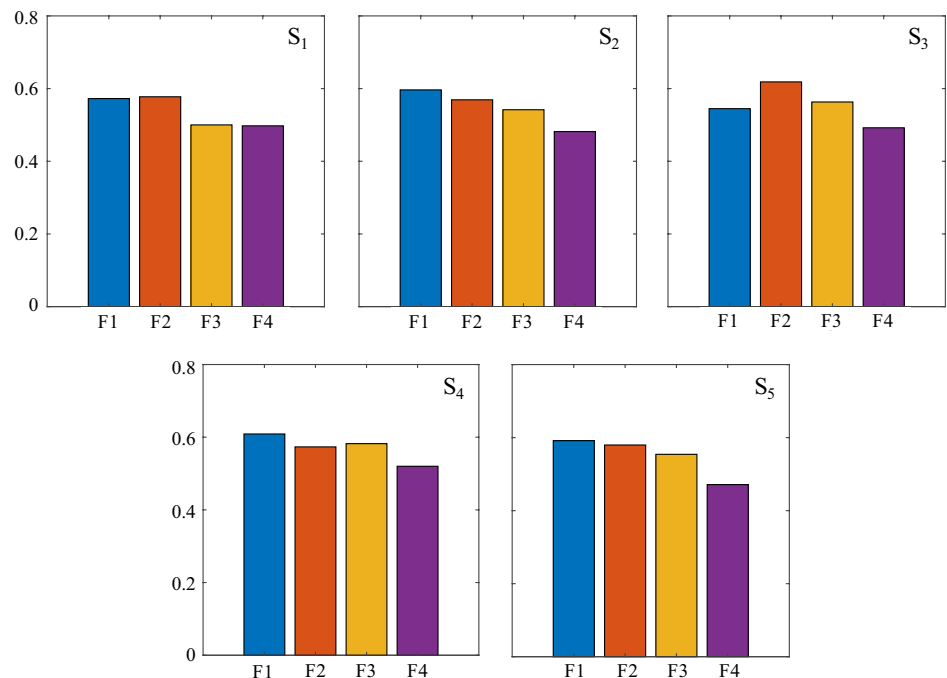


Table 3 The final score for each data format for the defined use-case

F1	F2	F3	F4
0.5828	0.5836	0.5483	0.4925

(OPC UA) which received 58.36% and the lowest score was F4 (MTConnect) which received 49.25%. However, considering that the difference between the best and the worst is only 9.11% and all scorings are around 50%, the data formats can be used for optical components only with additional efforts like explicit modeling of the optical properties (e.g., based upon ISO 23584 or ISO 10110). Equally, none of the proposed solutions provides a general and universal support for the use-case.

The selected use-case presents just one possible use-case in the field of precision assembly of optical microsystems and the stakeholders are derived from an ongoing project. Thus, adding more stakeholders or evaluation criteria may change the best fitting data exchange format.

5 Conclusion

In this paper, we defined a methodology for the selection of a data exchange format and evaluated it against a use-case of precision assembly of optical systems. In contrast to conventional machine tool, this use-case requires significant utilization of properties from the optic and photonic domain and also custom defined performance indication properties which need to be integrated in the data format. Four existing major data formats, AutomationML, OPC UA, B2MML/ BatchML and MTConnect are evaluated based on several criteria and scoring methods which is defined together with the industrial partners of the use-case. The results shows that OPC UA is the most suitable data format with 58.36% and MTConnect the worst with 49.25% for the use-case. Direct application of a format allows the users to take the model as presented and create the data of the use-case using the provided features of fixed nomenclature and support by third party tools directly. However, all selected data formats are evaluated with similar scorings around 50% which means that all formats are not directly suitable for the use-case without any adoption.

The methodology has been shown to be applicable for the applied use-case in a specific manufacturing subdomain. To prove this methodology in advance, more use-cases from different manufacturing subdomains as well as additional data exchange formats are required for observation. Based on the results in this paper, the methodology can also be augmented with the explicit modeling of the criteria dependencies. In addition to the applicability, the reliability of the methodology has to be observed in detail. One stakeholder could

have a significant impact on the result. Thus, the evaluation could be intentionally manipulated by knowing the methodology (e.g. by prioritizing only the criteria that are supported widely by one format instead of the criteria relevant for the stakeholder). This characteristic can be solved by having a larger set of stakeholders, or even prioritize the stakeholders. The second option could be done by a responsible neutral person or organization project leading person or by deduction, like functions taking the available resources per stakeholder into account. The general reliability of this methodology has to be observed in additional use-cases.

To prove the reliability of the evaluation result, an implementation of the Digital Twin using all of these data exchange formats is also possible. The implementation per data format provides an overview of the actual support of the formats for the modeling of Digital Twin and allows the stakeholders to check how much their criteria are met by the resulting Digital Twins. This also allows a possible additional step of evaluation. In this step, the priorities, ratings and implementations can be evaluated by the stakeholders using the actual degree of fulfillment in Digital Twin.

More research using different optics-function-oriented use-cases is required to get a better insight on the eligibility of the data exchange formats. Furthermore, we focused on exchange formats in the domain of manufacturing. Considering the results of the evaluation, the selected data exchange formats could be extended to prominent exchange formats from other domains [e.g., JSON-LD (JavaScript Object Notation for Linked Data) with schema.org] regarding the possibilities of synergies of different domains. Also, the potentials of linking and contrasting criteria depends on the stakeholders and the use-cases, and thus, cannot be generalized. In the presented methodology, the assessment of these potentials is majorly a task for the stakeholders themselves. In further research on these potentials, their use-case and stakeholder-based deduction could support stakeholders in prioritizing the criteria.

In the field of photonic integrated circuits (PIC), a new data model was proposed, namely openEPDA based on Yet Another Markup Language (YAML) Version 1.2. This format comes from the photonics and electronics domain aiming for the data exchange format for PIC taking photonic properties of different production steps into account. While the aim perfectly matches the use-case defined in this paper, the current state of this format is still in draft focusing on chip-level first. Therefore, we omitted the format in the evaluation as it is not finished [22].

The methodology shown in this paper takes into account the amount of criteria, that are relevant to a stakeholder by increasing the influence of criteria for stakeholder according to the total number of criteria marked relevant in comparison to other stakeholders. A possible extension of the methodology could be a stakeholder weight function to avoid

additional relevance shifts based on groups of stakeholders from the same area. In this paper, we also omitted the scope of the data exchange communication, which has an impact on efficiency, security and performance, but recognize the need for research in the future as well.

Acknowledgements Open Access funding provided by Projekt DEAL. This work was supported by Basic Research Lab Program through the National Research Foundation (NRF) funded by the MSIT (NRF-2018R1A4A1059976).

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

- Park, K. T. et al. (2019). Cyber physical energy system for saving energy of the dyeing process with industrial internet of things and manufacturing big data. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 7, 1–20.
- Park, C.-W., et al. (2009). Energy consumption reduction technology in manufacturing—A selective review of policies, standards, and research. *International journal of precision engineering and manufacturing*, 10(5), 151–173.
- Beier, G. et al. (2017). Sustainability aspects of a digitalized industry—A comparative study from China and Germany. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 4(2), 227–234.
- Kang, H., Seok, et al. (2016). Smart manufacturing: Past research, present findings, and future directions. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 3(1), 111–128.
- Schroeder, G. N. et al. (2016). Digital twin data modeling with automationML and a communication methodology for data exchange. *IFAC-PapersOnLine*, 49(30), 12–17.
- Galdzicki, M. et al. (2014). The Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology. *Nature biotechnology*, 32(6), 545.
- Doluschitz, R., et al. (2005). agroXML-A Standardized data format for information flow in agriculture. *EFITA/WCCA 26*.
- Oberle, D. (2014). How ontologies benefit enterprise applications. *Semantic Web*, 5(6), 473–491.
- ISO 15836-1:2017 Information and documentation—The Dublin Core metadata element set—Part 1: Core elements.
- Jenoptik Optical Systems GmbH. Open heat sink diode lasers | JOLD-x-CPNN-1L | JOLD-QPNN-1L.
- Haag, S., et al. (2014). Engineering of automated assembly of beam-shaping optics. In: *Laser resonators, microresonators, and beam control XVI*. Vol. 8960. International Society for Optics and Photonics.
- Sauer, S., et al. (2019). Smart adhesive gap design to minimize volumetric shrinkage misalignment effects in the automated assembly of FAC to Bottom Tab subassemblies (Conference Presentation). In *High-power diode laser technology XVII*. Vol. 10900. International Society for Optics and Photonics, 2019.
- Sauer, S., et al. (2017). Individualized FAC on bottom tab subassemblies to minimize adhesive gap between emitter and optics. In *High-power diode laser technology XV*. Vol. 10086. International Society for Optics and Photonics.
- Berners-Lee, T., James, H., & Ora, L. (2001). The semantic web. *Scientific American* 284(5): 28–37.
- Drath, R., ed. (2009). *Datenaustausch in der Anlagenplanung mit AutomationML: Integration von CAEX, PLCopen XML und COL-LADA*. Springer, Berlin.
- Leitner, S.-H., & Mahnke, W. (2006). OPC UA—service-oriented architecture for industrial applications. *ABB Corporate Research Center*, 48, 61–66.
- Pipero, C., & Manjunath, K. (2006). ISA 95 Implementation best practices: Workflow descriptions using B2MML. In *WBF North American Conference*, Atlanta, GA.
- MTConnect Institute. MTConnect Standard Specification 4.1 Version 1.4.0. March 2018.
- Vijayaraghavan, A., et al. (2008). Improving machine tool interoperability using standardized interface protocols: MT connect.
- OPC Foundation. Microsoft contributing Open-Source OPC UA stack to OPC Foundation 04/2017. Retrieved from <https://opcfoundation.org/news/press-releases/microsoft-contributing-open-source-opc-ua-stack-opc-foundation/>. Accessed May 2019.
- Hepp, M. (2008). Goodrelations: An ontology for describing products and services offers on the web. *International Conference on Knowledge Engineering and Knowledge Management* (Springer, Berlin).
- Latkowski, S., et al. (2019). Open standards for automation of testing of photonic integrated circuits. *IEEE Journal of Selected Topics in Quantum Electronics*.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



M.Sc. Arno Schmetz studied computer science in Aachen (Germany) with a focus on communication and distributed systems. In April 2017, he started his career as research assistant at Fraunhofer IPT in the area of automation and connected adaptive productions. Arno Schmetz is responsible for a software framework for data acquisition, data aggregation, and preprocessing that is used in different projects and extended to their needs. In addition, he works in a team for developing platforms

and components for connected adaptive production, especially Smart Manufacturing Networks including Digital Twin technology.



Tae Hun Lee received his bachelor and master degree in electrical engineering from RWTH Aachen in 2016. Currently, he is working at the Fraunhofer Institute for Production Technology IPT as leader of the group special purpose and networked machines. His research is focusing on metrological analysis of machine tool behavior regarding the geometric, thermal, and dynamic influences.



Dr. Susanne Ehret received her Ph.D. from the University of Edinburgh in 2017 where she studied theoretical particle physics. She is researching in the field of precision assembly and automation at the Fraunhofer Institute of production technology since 2018.



Maximilian Hoeren started his career at the Fraunhofer IPT in Aachen in 2014, with the processing of his bachelor thesis regarding the automated active alignment of microoptical components. In 2017, he finished his masters in automation technology at the RWTH Aachen, and from then, he has been working as a research fellow at the institute. Since April 2019, Maximilian Hoeren is the Group Manager of precision assembly and automation. His expertise is in the field of process development

for automated active alignment.



Daniel Zontar has been an employee of the Fraunhofer Institute for Production Technology IPT since 2009. In the period from 2014 to 2017, he was responsible for “Active alignment for the assembly of optical systems” at the Fraunhofer IPT and took up the position as Manager of the Machine Development and Machine Networking Group in 2018. In his capacity, as the Coordinator of International Research and Industry Projects, he has developed a range of technologies

used to control assembly cells as well as solutions for special purpose and classical machine tools. Daniel Zontar has managed the Precision Engineering and Automation Department at the Fraunhofer IPT since June 2018.



Marvin Berger works since 2017 for Fraunhofer IPT in Aachen in the group of precision assembly and automation. He studied at the University Bayreuth from 2010 until 2016 for his Bachelor and Master of Science Physics in the areas of solid state bodies and biophysics. From October 2014, he worked for the Commonwealth Science and Industry Organization (CSIRO), Materials and Engineering in Clayton, Melbourne, and the Melbourne Center for Nanofabrication for 3 months.



Soo-Hong Min received his bachelor degree in mechanical engineering and business administration from Seoul National University in 2016. He is currently studying for a Ph.D. in the Department of Mechanical Engineering at Seoul National University. His research interests include micro-/nanoscale fabrication, flexible and printed sensor, and smart manufacturing system.



Sung-Hoon Ahn received a Ph.D. degree from Stanford University in 1997. He has been affiliated with UC Berkeley, Gyeongsang National University, University of Washington, and Hyundai WIA Corp. He is currently a full professor in the department of mechanical engineering at Seoul National University. His research interest include smart factory, green manufacturing, soft robotics, 3D printing, composite materials, micro/nano fabrication, renewable energy, and appropriate technology.



Prof. Dr.-Ing. Christian Brecher has been appointed as the Executive Director of the Fraunhofer Institute for Production Technology IPT, on January 1, 2018. Since January 1, 2004, he has already been the Ordinary Professor for Machine Tools at the Laboratory for Machine Tools and Production Engineering (WZL) of the RWTH Aachen as well as the Director of the Department for Production Machines at the Fraunhofer IPT. After finishing his academic studies in mechanical engineering,

he started his professional career first as a research assistant and later as a team leader in the department for machine investigation and

evaluation at the WZL. From 1999 to April 2001, he was responsible for the department of machine tools in his capacity as a Senior Engineer. After a short spell as a consultant in the aviation industry, Professor Brecher was appointed in August 2001 as the Director for Development at the DS Technologie Werkzeugmaschinenbau GmbH, Mönchengladbach, where he bore the responsibility for construction and development until December 2003. Prof. Brecher has received numerous honours and awards including the Springorum Commemorative Coin, the Borchers Medal of the RWTH Aachen, the Scholarship Award of the Association of German Tool Manufacturers (Verein Deutscher Werkzeugmaschinenfabriken VDW), and the Otto Kienzle Memorial Coin of the Scientific Society for Production Technology (Wissenschaftliche Gesellschaft für Produktionstechnik WGP).