Technical Report

Improving Engineering Data Exchange in Parallel Production Systems Engineering

Stefan Biffl[1]
Arndt Lüder[2]
Felix Rinker[1,3]
Laura Waltersdorfer[1,3]

[1]TU Wien, Information Systems Engineering, Information and Software Engineering, Quality Software Engineering, Vienna, Austria
<firstname>.<lastname>@tuwien.ac.at

[2]Otto-v.Guericke Universität Magdeburg, Magdeburg, Germany
<firstname>.<lastname>@ovgu.de

[3] Christian Doppler Laboratory for Security and Quality Improvement in the Production System Life Cycle, Information and Systems Engineering, TU Wien, Vienna, Austria
<firstname>.<lastname>@tuwien.ac.at

Technical Report No. CDL-SQI 2019-04
Issued: March 2019

Citation: S. Biffl., A. Lüder, F. Rinker, L. Waltersdorfer.: „Improving Engineering Data Exchange in Parallel Production Systems Engineering“,Technical Report CDL-SQI 2019-04, TU Wien, Vienna, Austria, March 2019.

Improving engineering data exchange in parallel production systems engineering

Stefan Biffl[1], Arndt Lüder[2], Felix Rinker[3], and Laura Waltersdorfer[1]

[1] Institute of Information Systems Engineering, Technische Universitt Wien, Austria
`<firstname>.<lastname>@tuwien.ac.at`
[2] Institute of Factory Automation, Otto-von-Guericke University of Magdeburg, Germany
`arndt.lueder@ovgu.de`
[3] Christian Doppler Laboratory for Security and Quality Improvement in the Production System Lifecycle (CDL-SQI), Technische Universitt Wien, Austria
`felix.rinker@tuwien.ac.at`

Abstract. In the parallel engineering of industrial production systems, domain experts coming from several disciplines need to exchange data efficiently to prevent the divergence of local engineering models. However, the data synchronization is hard (a) as it may be unclear what data consumers need and (b) due to the heterogeneity of local engineering artifacts and data. In this paper, we introduce use cases and a process for efficient engineering data exchange Engineering Data Exchange (EDEx) that guides the definition and semantic mapping of data elements for exchange and facilitates the frequent synchronization between domain experts. We identify main elements of an EDEx information system to automate the EDEx process. We evaluate the effectiveness and effort of the EDEx process and concepts in a feasibility case study with requirements and data from real-world use cases at a large production system engineering company. The domain experts found the EDEx process more effective and the EDEx operation more efficient than the traditional point-to-point process, and providing insight for advanced analyses.

Keywords: production systems engineering · data exchange · data integration · process design · multi-aspect information system · multidisciplinary engineering.

1 Introduction

Engineering industrial, recently also cyber-physical, production systems, e.g., long-running and safety-critical systems for assembling automotive parts or for producing metal, is the business of multi-disciplinary production system engineering (PSE) companies [3] [17]. In parallel engineering, the disciplines develop their engineering and artifacts, such as plans, models, software code, or machine configurations, independently, but have to consider dependencies between the engineering disciplines in order to build a common system. A key success factor is the capability to exchange selected data in the engineering artifacts with related domain experts efficiently and in a timely manner to reduce rework due to inconsistencies.

We illustrate the engineering data exchange *EDEx* process with a use case from simulation in PSE, as simulation is a major consumer of engineering data for assessing the safety and business risks of a production system before system construction. Goal of the simulation engineer is to design simulation systems that allow exploring dynamic properties of the designed production system, such as throughput or the physical feasibility of production steps. Therefore, the simulation engineer requires input data from several engineering data providers on key parameters of system parts, such as the rotation speed, torque, control signals, or power consumption of a motor as foundation for calculating and analyzing the movement of work pieces and robots over time.

In the traditional EDEx process [4] [1], domain experts communicate their engineering artifacts point-to-point, typically in the form of spreadsheet tables, pdf or XML files. Unfortunately, in the traditional EDEx process, Luder et al. [11] identified the following major challenges.

Ch1. Unclear data requirements of and benefits for stakeholders. For potential data providers, it is often not clearly defined which project participants require what kind of data at what point in time in the project. Even if general dependencies between stakeholders are known, the specific relations between engineering artifacts and their content within an engineering project can change during the project execution. Insufficient overview may prevent even willing stakeholders from sharing their data.

Ch2. Heterogeneous engineering data is hard to integrate for sharing. Due to strongly diverging scientific and practical histories, engineering tools and data are typically specific for a discipline and not designed for use with other disciplines or the project they contribute to. While the disciplines share some common concepts, such as the concept of a device or a signal, these concepts are not consistently modeled, making data integration for sharing error prone and hard to automate. Consequently, data providers tend to share engineering artifacts that take high effort for consuming domain experts to find and interpret, and, thus, hinder comprehensive automated processing.

In this paper, we introduce a process for efficient data logistics to exchange engineering data to address these challenges and to automate data logistics in order to improve the value and reduce the risks of EDEx. We investigate the following research questions (RQs) based on Design Science research methodology [18].

RQ1. What are main elements of an effective and efficient engineering data exchange EDEx process in Multi-Disciplinary Engineering? To address this research question, Section 2 summarizes related work on approaches for data exchange in multi-disciplinary production systems engineering (PSE). In Section 3.1, we discuss requirements for the EDEx process collected in workshops with stakeholders at a large PSE company. In Section 3.2, we propose steps for an EDEx process that address these requirements. For designing the EDEx process, we adapt the *Multi-Model Dashboard* approach [4] from constraint evaluation to EDEx and replace the design requirement of an initial common concept model, which may not be available, with direct links between consumer and provider data elements.

RQ2. What are main information system mechanisms that enable engineering data exchange for Multi-Disciplinary Engineering? In Section 3.3, we derive requirements for effective and efficient EDEx information system (EDExIS) mechanisms: capabilities for data set specification and for the representation of dependency relationships as foundation for data integration and transformation. In Section 4, we report on an evaluation of the effectiveness and effort of the proposed EDEx process with EDExIS mechanisms in a feasibility case study with requirements and data from real-world use cases with domain experts at a large PSE company.

Section 5 discusses the findings and limitations. Section 6 concludes and proposes future research work. From the research we expect the following contributions for the information systems engineering (ISE) community. The use cases and EDEx process give ISE researchers insight into the PSE domain, the foundation for Industry 4.0 applications. The EDEx process contributes capabilities for designing and investigating agile processes and information systems in PSE, a foundation for conducting engineering projects for cyber-physical production systems economically.

2 Related work

This section summarizes related work on data exchange in production systems engineering, information system (IS) engineering, and software engineering.

2.1 Data Logistics in Production Systems Engineering

In the Production Systems Engineering (PSE) process [3], the content of the exchanged artifacts is important as these artifacts contain only part of the local models of the domain experts. Due to the inherent dependencies between the local models of the domain experts, such as the impact of electro-magnetic fields from electrical wiring on communication quality of communication wiring, domain knowledge is required on both the customer and the provider data models to interpret the content of the exchanged data. Therefore, it is necessary to move from delivering engineering artifacts to engineering data exchange EDEx. Although business process analysis [12] is useful to better understand the relevant stakeholder groups, activities, and exchanged engineering artifacts, additional data modeling is required to represent the knowledge required for EDEx.

While EDEx is already important and difficult for traditional PSE, the migration towards cyber-physical systems is a complex task that requires an extensive solution, covering technical, operational, and human dimensions [6]. Due to this multi-dimensional complexity, traditional information systems have not yet adequately addressed the challenges imposed by collaboration in multi-disciplinary engineering systems: heterogeneous tools and data formats, diverging views on artifacts and their versioning are the most pressing ones [6]. Optimizing and enriching the currently available engineering data and data exchange is a possible quick win that can be achieved by integrating EDEx [16]based on the machine understandable representation of knowledge on how exchanged data elements fit to the local data models of the data providers and consumers.

While there are engineering tool suites that integrate several engineering functions in one set of tools with a common data model that greatly simplifies EDEx, most engineering projects use many tools with heterogeneous data models that are challenging to integrate [3]. The traditional EDEx process [4] is a point-to-point exchange of engineering artifacts between domain experts via email, repository, or USB stick, typically in the form of spreadsheet tables, PDF or XML files.

Lüder et al. [11] introduce an architecture for engineering data logistics, based on *AutomationML* [17] an open, XML-based format for the exchange of engineering data. The proposed architecture allows exchanging data between discipline-specific data models with varying hierarchical key systems. While this approach is useful in an *AutomationML* environment, the approach does not consider how to negotiate the EDEx between many data consumers and providers.

The *Multi-Model Dashboard* (MMD) [4] process guides the systematic definition, monitoring, and evaluation of PSE parameters and constraints. While we can build on the MMD strengths as foundation for the EDEx research in this paper, the following limitations of the MMD approach require significant adaptation for data exchange in a systems engineering project. The MMD does not consider the provision of data to consumers, but focuses on the evaluation of engineering parameters and constraints. Furthermore, the MMD assumption of well-defined common concepts is hard to implement in practice since several disciplines cooperate with each other, with no discipline clearly leading.

2.2 Data exchange contributions from Inf. Systems and Software Engineering

Methods from business process management provide useful approaches, such as UML class diagrams [5] or BPMN [12], for EDEx definition by characterizing involved stakeholders, systems and, to some extent, data types and their relationships. However, these methods are generic and need to be adapted for new contexts, also in the case of heterogeneous engineering data integration [14]. In specialized domains, such as medicine, science and engineering, new approaches may be needed to optimize data exchange according to domain-specific requirements [9] [13].

Semantic Web technologies are recognized for facilitating data exchange across applications and organizations in the web and have proposed engineering data integration approaches following the *interchange standardization approach* [16]. However, the manifold types of dependencies in PSE data models are different form typical Semantic Web requirements [10] and the Semantic Web technology stack is therefore currently seldom used in engineering environments.

Software engineering design patterns [8] encapsulate best practices of software system design for commonly occurring problems, in our case data and tool integration. In the context of this work, we build on design patterns such as *message passing* and *publish-subscribe* to support the loose coupling of engineering work groups and tools.

3 Design of the Engineering Data Exchange Process and IS

This section introduces requirements, use cases, and main elements for an Engineering Data Exchange EDEx process and derives mechanisms of an EDEx IS.

3.1 Required capabilities for an Engineering Data Exchange Process

Following the *design science cycle* in [18], we set up an initial problem investigation with workshops [2], outlining the context and problem space of research, and deriving the following requirements for EDEx capabilities that allow addressing the challenges introduced in Section 1: *Ch1. Unclear data requirements of stakeholders* and *Ch2. Heterogeneous engineering data is hard to integrate for sharing.*

Cap1. Engineering Data Representation. This capability concerns the representation of candidates for, overview on, and specifics of typical engineering data structures, such as tree hierarchies of the functions of a production system, (e.g., a work cell consists of devices), lists of objects (e.g., list of motors), and objects and their attributes (e.g., motor torque or rotation speed) and relationships (e.g., a work cell with an electric motor requires an electric power supply), both for data consumers and data providers.

Cap2. Semantic Link Knowledge representation. This capability concerns the representation of candidates for, overview on, and specifics for semantic links for data integration between selected consumer and provider data elements. The explicit representation of tacit knowledge on these semantic links will allow reasoning on, improving, and automating data integration and data transformation for EDEx.

Cap3. Process Data Representation. This capability concerns the representation of metadata on the EDEx process, e.g., who provided what data when, versions of data elements, data quality and validity (e.g., unclear/checked valid, invalid data).

Cap4. Consumer- and Benefit-Driven EDex Planning. This capability emphasizes planning EDEx guided by business benefits coming from data consumer use cases to ensure the prioritization of EDEx with high benefits compared to the cost for set up and operation, an economic improvement over the traditional provision of engineering artifacts. This capability implies the requirement for providing an overview on stakeholders interested in requested or provided data as foundation for a data logistics marketplace and for advanced analysis on the network of EDEx relationships.

3.2 Use Cases for Evaluation

From workshops with 27 domain experts at a large PSE company, from four different domains of expertise, and 6 researchers, we identified two illustrative EDEx consumer use cases (UCs) and benefits as foundation for the EDEx process design and evaluation.

The engineering of a typical industrial production system (PS), such as automotive assembly, requires at least the collaboration of, and EDEx between, the *plant planner* (PP), who plans the layout of the PS, *mechanical engineer* (ME), *electrical engineer* (EE), and *robot programmer* (RP). Each domain expert designs and updates complex and heterogeneous local models that are hard to understand by other domain experts.

UC Sim. Data exchange for production system simulation. In a typical advanced engineering environment, a *simulation engineer* (SimE) designs and runs simulation models to check the engineering results and to optimize production system parameters, such as safety risks, production throughput, and energy consumption. The design of the simulation models depends on the input of several other domain experts. The PP, ME, EE, RP may provide configuration parameters of motors and conveyors in a transport system and requirements of production processes, such as process duration (s) and production resource parameters, such as length (m), mass (t), or power consumption (kW). The SimE requires this input from data providers to calculate characteristics, e.g., power consumption or movement dynamics, of a system part, e.g., a drive chain, to find out whether the system part will behave as intended and to provide feedback to the contributing engineering disciplines on risks and on necessary design changes. The manual synchronization of these data typically requires additional effort, tends to be error prone, and induces avoidable project risks.

UC PM. Engineering project monitoring. The *project manager* (PM) wants to use the input from data providers to the SimE to assess project progress by analyzing the completeness and quality of data with respect to the project phase and planned deliverables. Missing or inconsistent data may make sense in an early design phase, but may pose a major risk in closer to a later design milestone and may require action by the PM. Unfortunately, the PM does not understand the various engineering artifacts.

3.3 Engineering Data Exchange Process Design

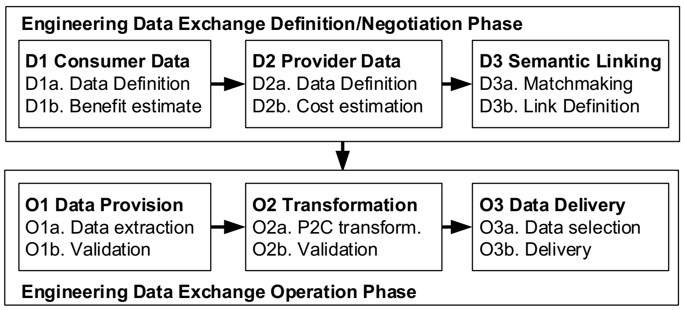

Fig. 1. Data Exchange process architecture with definition/negotiation and operation phases.

To address the required capabilities in Section 3.1 and the use cases in Section 3.2, we introduce the main elements of an engineering data exchange EDEx process, a treatment design according to [18], based on the knowledge gathered in workshops with domain experts. The EDEx process adapts and extends the *Multi-Model Dashboard* process [4] in the research scope of cooperating multi-disciplinary engineering work groups in a production systems engineering project. The EDEx process is independent of a concrete implementation technology.

Process architecture. Fig. 1 gives an overview on the EDEx definition and operation phases. The EDEx operation phase assumes an agreement between data consumers and data providers on the data model and concepts for EDEx. Therefore, a negotiation of the data requested by consumers and the data published by providers is required to reach such an agreement, comparable to a marketplace of well-defined data products.

EDEx roles are the data consumer, the data provider, and the EDEx curator. The data consumer requests data according to her local consumer data model from providers to improve her business processes. The data provider has engineering artifacts that contain relevant data for a consumer and knows how to extract this data from the artifacts following his local provider data model. A domain expert may be consumer and provider. The *EDEx curator* has background knowledge on the PSE business and relevant data models of all domain experts to mediate between consumers and data providers. The *EDEx curator* has the capability to link the local data models of consumers and providers with appropriate link definitions, such as mathematical formulae.

EDEx Definition Phase. The EDEx definition phase consists of three main steps to identify feasible and beneficial candidate instances for data exchange. At the end of this phase, the EDEx roles come to agreements on which data sets they plan to exchange as foundation for the technical design and implementation in a suitable EDEx environment. Figures 2 and 3 illustrate selected cases of the EDEx processes.

D1. Consumer data definition and prioritization. Consumer candidates have to define their data requests. In general, consumers know what data is likely to be available from which data providers. Outcome of this activity is a data model of the local consumer data view, e.g., in UML, *AutomationML*, or in natural language based on the modelling concepts and vocabulary of the consumer. The *EDEx curator* validates with the consumer the definition of requested data and estimates the likely benefit of providing the data to focus on the most relevant EDEx instances first. This step results in a set of data model elements in the local consumer data view, with a semantic description understandable both to the *EDEx curator* and prospective providers based on the modelling concepts and vocabulary of the *EDEx curator* (see Fig. 2b, tag D1). Note that this step is iterative to allow consumers adding data elements.

D2. Provider data definition and cost estimation. A provider can react to consumer data requests by agreeing to publish data that is semantically equivalent to (parts of) the requested consumer data. In general, providing the data will involve extracting the data elements from suitable engineering artifacts, e.g., the mechanical structure of a work cell. Outcome of this activity is a set of data model elements in the local provider data view, with a semantic description understandable both to the *EDEx curator* and prospective providers based on the modelling concepts and vocabulary of the provider (see Fig. 2b for examples). Extracting data from engineering artifacts can take significant effort and cost, even to an expert. Therefore, the *EDEx curator* has to elicit the likely cost for data extraction and transformation into a format that is suitable for EDEx, such as *AutomationML*. The goal of this step is to give feedback to the provider whether the data is of sufficient quality and reasonable cost to continue setting up the EDEx (see Fig. 2b, tag D2).

D3. Consumer-provider mediation and semantic link definition. D3a. Economic matchmaking between consumers and providers. For each promising consumer data request, the *EDEx curator* tries to find a set of providers that would allow providing the requested data. In typical cases, the data elements may come from several providers in a variety of data formats (see Fig. 2a). Main target of this step is a set of EDEx providers that could, together, provide the input data for transformation into the requested data elements. If there are several solutions, the options could be ranked by data quality, availability, and likely cost. D3b. Semantic linking between consumer and provider data models. For each pair of requested and provided data items, the *EDEx curator* establishes a formal semantic link, i.e., a formula that specifies how to calculate the consumer data value from published provider data instances using the modelling concepts and vocabulary of the EDEx curator. A semantic link can describe, in a simple case, semantic identity. More advanced semantic relationships [10] include string operations, mathematical calculations, and parameterized function calls to semantic transformation algorithms (see Fig. 3). Outcome of this step is a set of consumer data, semantically linked to a set of provider data as foundation for designing the EDEx operation, supported by an EDExIS.

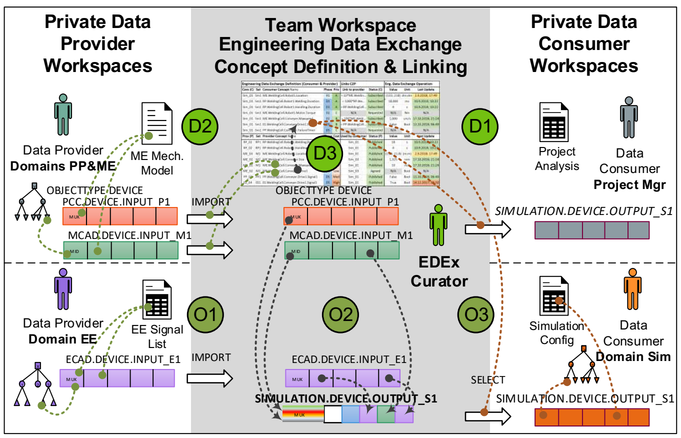

Fig. 2. Engineering data exchange definition/negotiation and operation for a customer data set (based on [4]; tags in green circles refer to EDEx process steps in Fig. 1).

EDEx Operation Phase. The EDEx definition phase provides the foundation for conducting the exchange of data instances (see fig. 2 and 3.

O1. Data provision and validation. The provider extracts the data elements as agreed in the EDEx definition phase from their local engineering artifacts. Then the provider transforms the extracted data into a data model and format that the *team workspace* can import (see fig. 2, tag O1). The provider and the *EDEx curator* agree on a procedure to validate the data from extraction to input to *team workspace* to ensure that only correctly transformed data is imported. The *EDEx curator* imports valid data into the *team workspace*. Main contributions of this step are imported valid data in the *team workspace* and feedback to the data provider on the validity of the provided data.

O2. Semantic data transformation and validation. A transformation mechanism in the *team workspace* propagates the imported data along the semantic links to fill in or update consumer data sets (see fig. 2, tag O2). The *EDEx curator* checks the correctness of the transformation of imported provider to consumer data. Motivation for this step is to get updated consumer data sets and feedback on the validity of the semantic transformation of the recently imported provider data.

O3. Data selection and delivery. The consumer selects data instances by providing the *team workspace* with the type of and information to select the requested data instances, such as data identifiers or selection conditions, similar to a SQL query to a database. The *team workspace* delivers the result data to the consumer (see Fig. 2a, tag O3). Outcome of this step is a set of selected data at the consumer in the agreed format.

Data Exchange Definition/Negotiation Phase							Operation Phase		
Engineering Data Exchange Definition (Consumer & Provider)					Links C2P		Eng. Data Exchange Operation		
Cons (C)	Set	Consumer Concept Name	Phase	Prio	Link to provider	Status (C)	Value	Unit	Last Update
Sim_01	Sm1	ME.WeldingCell.Robot1.Location	R1	A	= 10*ME.Weldin...	Subscribed	(133; 218)	dm;dm	2.9.2018; 17:49
Sim_01	Sm1	RP.WeldingCell.Robot1.Welding.Duration	D3	A	= 1000*RP.We...	Subscribed	18,000	ms	10.9.2018; 10:22
Sim_01	Sm1	RP.WeldingCell.Robot1.Handling.Duration	D3	A	= RP.WeldingC… D3	…cribed	4	s	10.9.2018; 10:22
Sim_01	Sm1	ME.WeldingCell.Robot1.Motor.Tor… D1	D2	C	N/A	…uested	N/A	N… O3	N/A
Sim_01	Sm1	ME.WeldingCell.Conveyer.Maxspee…	D2	B	= 100*ME.Weld… O2	…bed	1,000	cm…	…0.2018; 21:24
Sim_01	Sm1	EE.WeldingCell.Conveyer.Drive1.Signal1	D5	B	= PP.WeldingCell...	…ribed	False	Bool	13.10.2018; 06:49
Sim_01	Sm1	PP.WeldingCell.Conveyer.FailureTimer	D5	C	N/A	Requested	N/A	s	N/A
Prov (P)	Set	Provider Concept Name	Phase	Cost	Used by consumer	Status (P)	Value	Unit	Last Update
RP_02	RP1	RP.WeldingCell.Robot1.Welding.Duration	D3	Low	Sim_01	Published	18	s	10.9.2018; 10:22
RP_02	RP1	RP.WeldingCell.Robot1.Handling.Duration	D3	Low	Sim_01	Published	4	s	10.9.2018; 10:22
ME_03	M1	ME.WeldingCell.Robot1.Location D2	R1	Med	Sim_01	Published	(13.3; 21.8)	(m… O1	…2018; 17:49
ME_02	M2	ME.WeldingCell.Conveyer.Size	R2	Low	Sim_03	Published	1,325	m…	…0.2018; 21:24
ME_02	M1	ME.WeldingCell.Conveyer.Maxspeed	D2	Low	Sim_01	Published	10	m/s	17.10.2018; 21:24
ME_02	M2	ME.WeldingCell.Conveyer.Drive1	D1	Low	Sim_03	Agreed	N/A	Bool	N/A
PP_03	PP1	PP.WeldingCell.Conveyer.Drive1.Signal1	D5	Med	Sim_01	Published	False	Bool	13.10.2018; 06:49
EE_04	EE1	EE.WeldingCell.Conveyer.Drive1.Signal1	D5	High	Sim_01	Published	True	Bool	14.12.2017; 06:50

Fig. 3. EDEx definition/negotiation and operation overview table (based on MMD dashboard [4]; tags in green circles refer to EDEx process steps in Fig. 1).

3.4 Illustrating Use cases.

Fig. 2a illustrates an overview on the roles, engineering artifacts, and exchanged data for the EDEx definition/negotiation and operation processes (see Fig. 1) for one consumer data set, in this case device parameters collected for the SimE. The data providers and data consumers, such as the PP, ME, EE, and SimE, operate in *private workspaces*. The *team workspace* contains shared data views as foundation for preparing and operating the EDEx processes.

Parameter exchange for production system simulation. In this use case, the SimE requires a set of parameters to configure the simulation for a device (see Fig. 2a, lower right hand part, red bar), such as a robot or conveyer. The SimE requests the set of parameters from providers, such as the PP, ME, EE, and RP, who may agree and publish their local engineering data corresponding to a consumer request (see Fig. 2b, left hand part). Then the *EDEx curator* links the set of parameters requested by the SimE with the set of parameters published by the

PP, ME, EE, and RP (see Fig. 2b, middle part for the ME and EE data) to enable the EDEx operation.

During the EDEx operation phase, the *team workspace* receives updates of provider data instances in engineering artifacts from the *private workspace*s of the PP, ME, EE, and RP (see Fig. 2a, left hand side for the ME and EE) and transforms this input data according to the semantic links into output data for delivery to the SimE (see Fig. 2a, right hand side, and example output data in Fig. 2b, right hand upper part). The SimE can be notified as soon as relevant data for a requested data set is available or changed, so the SimE can consider when to retrieve which part of the currently available data.

Production system engineering project monitoring. In this use case, the PM can benefit from the EDEx for simple and advanced analyses. A simple analysis could be to subscribe to the same data sets as the SimE and analyze at specific points in the project for which data elements the engineering data is expected but missing.

Fig. 3 shows a snapshot of the EDEx overview table during operation: data instances coming from the providers have been processed according to the linking formulae to fill in data instances for consumers (tags O1, O2, O3). For consumers, the EDEx overview (tag D1) shows the status of the data elements as requested, agreed for provision, or subscribed for delivery. The EDEx overview table (tag D3) shows the status of linked data elements. For a requested data element, there may be several providers; therefore, the EDEx overview table (see Fig. 3) indicates the cost of providing a data element and the engineering process phase, in which the data will be available with sufficient precision, to support making an informed choice on the best provider. For example, EESignal1 could be obtained from PPSigna1 at lower cost.

The concepts illustrated in Figures 2 and 3 are the foundation for prototype designs as input to the evaluation with domain experts in Section 4.

3.5 Engineering Data Exchange Information System Design Considerations

This section describes the main architectural components of mechanisms of an EDex information system (EDExIS) to address the requirements for capabilities of the EDex process described in Section 3.1 and the use cases described in Section 3.2. The design of these mechanisms is likely to vary depending on the application context.

Looking at fig. 2 the scope of the EDExIS corresponds to the concept of the *team workspace* with interfaces to the *private workspace*s of data providers and consumers.

EDEx management and overview. For managing the EDEx process, the EDExIS has to provide a mechanism providing the capabilities of the EDEx overview table illustrated in Fig. 2b, including EDEx definition functions to request, agree on providing, publishing, and subscribing to data elements (see EDEx process steps D1 to D3), as well as setting relevant attributes of and searching the table for understanding the status of the EDEx definition in the project team.

EDEx data definition languages. For EDEx definition, the EDExIS has to process the languages for the specification of consumer and provider data sets, and the language for semantic link definition specifying (a) the dependencies between consumer and provider data sets and (b) the transformation of imported provider data into consumer data.

Fig. 3 illustrates examples of semantic link definitions between consumer and provider data models. In the simplest case, the output value is just an identical copy of an input instance value see fig. 4 , formula DL1), assuming matching IDs for concepts welding cell, conveyer, etc.). Simple cases require scaling and/or shifting the input values, e.g., to adjust for different scales of units, such as m, cm, or mm, or s and ms (see fig. 4 , formula DL2). More advanced links may require more complex formulae including custom functions or combining the instance values from several data elements (see fig. 4 , formula DL3). A link formula can involve one or more providers and data elements as data sources and can encapsulate capabilities for string operations, advanced algorithms, and access to external knowledge, e.g., web services.

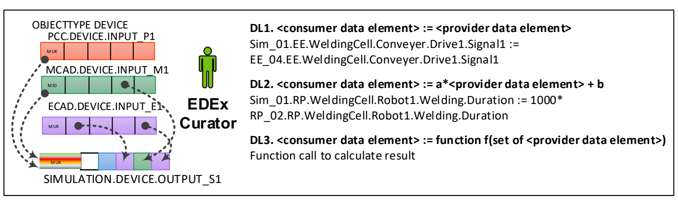

Fig. 4. Semantic link definition between consumer and provider data models.

EDEx operation capabilities. For conducting the EDEx operation steps, the EDExIS has to be able (a) to import and validate provider data, (b) to store imported data versions including their metadata for processing, (c) to analyze the data and semantic links in order to correctly propagate the provider data to consumer data structures, and (d) to select and export consumer data.

4 Evaluation

This section reports on the evaluation of the engineering data exchange EDEx process and requirements (a) in an initial feasibility case study [15] with 19 domain experts at a large production systems engineering (PSE) company, the project coordinators for each domain were the same as in the initial workshop, a systems integrator for metallurgic production systems, and (b) in a cost/benefit comparison of the EDEx definition and operation processes to the traditional process of point-to-point exchange of engineering artifacts between domain experts, closing an iteration of the design cycle [18] and providing knowledge for guiding future research.

4.1 Feasibility Study

Goal of the feasibility study is to evaluate the basic concept of the EDEx process with domain experts by following the steps of the EDEx process description (see Section 3.3 and fig. 2). Based on the use cases introduced in Section 3.2, we designed prototypes of selected user interface elements, such as the overview table, data specification, linking, and retrieval as electronic mock up artifacts with data from domain experts. We collected data on the usability and usefulness of the EDEx process based on the *Technology Acceptance Model* questionnaire [7] [1].

Further, we developed technology prototypes of the IS capabilities to explore the feasibility of designing the EDExIS concepts with available technologies, including *AutomationML* for data specification (see [11], an Excel dialect for the specification of dependency links, Java code for transformations, and BaseX as data storage. We conducted and discussed the EDEx steps in a workshop with domain experts representing the roles data provider (PP, ME, EE, RP in the use cases), data consumer (SimE, PM), and EDEx curator.

Overall, the domain experts found the EDEx process feasible, useful, and usable for basic cases that make up most of the data exchange use cases in their typical project context, assuming that the EDExIS provides effective tool support to automate the data transformation, storage, and selection tasks. The domain experts provided improvement suggestions for the user interfaces, and for describing the data transformation and linking formulae in their context. Further, the domain experts noted that more complex cases may take considerable effort to design and automate; therefore, cost-benefit estimates in the EDEx process are important to guide planning the EDEx implementation.

4.2 Cost/Benefit Considerations

To evaluate the costs and benefits of the EDEx process via a *team workspace* in comparison to the traditional manual process of point-to-point e-mail based EDEx, we elicited needs and estimates from domain experts, who are responsible for engineering and project management of large-scale metallurgic production system projects.

Fig. 5 presents an overview of the findings for of the EDEx process steps in the use case Parameter exchange for production system simulation by comparing the effectiveness, i.e., correctness of results for a task, and the effort of a stakeholder conducting a task. We applied a 5-point Likert-Scale (++, +, o, -, –), where ++ indicates very positive effects, and – very negative effects. Positive effects refer to high effectiveness of the investigated approaches and to low effort for implementation and application.

EDEx Process Step	Effectiveness		Effort	
	Manual	EDEx	Manual	EDEx
D1. Consumer data definition & prioritization.	o	+	+	o
D2. Provider data definition & cost estimation.	--	+	+	-
D3. Consumer-provider semantic link definition.	--	++	N/A	-
O1. Data provision and validation.	-	o	+	o
O2. Data transformation and validation.	--	+	--	+
O3. Data selection and delivery.	-	++	--	++

Fig. 5. Comparison of the effectiveness and effort of traditional manual and EDEx processes.

Regarding effectiveness, the EDEx process was found effective to very effective by the interviewed stakeholders, both providers and consumers, because they were able to exchange data elements in a traceable and validated way. In the traditional approach, the data consumers had to define, procure, transform, and validate the required data with significant effort and prone to errors. However, the application of the EDEx process requires additional effort, especially during the EDEx definition (D2) and linking (D3), in particular for providers and for the new role of the EDEx curator.

On the upside, the results of the linking step (D3) significantly improve the representation of shared knowledge in the engineering team regarding an overview on the dependencies between the engineering roles on data element level. Domain experts and the PM can always get a current overview on the status of data deliveries and can identify missing engineering data and unfulfilled requests by consumers. In addition, the EDExIS can provide the benefit of immediate feedback on changed engineering data elements efficiently, without additional effort by the domain experts.

5 Discussion

This section discusses results regarding the research questions introduced in Section 1.

RQ1. What are main elements of an effective and efficient engineering data exchange EDEx process in multi-disciplinary engineering? Section 3.3 introduced as main elements EDEx roles, process steps, and data structures. The new role of the *EDEx curator* mediates between data consumers and providers. In the feasibility study, a domain expert filling this role informally was identified. The EDEx data structures represent the necessary knowledge on engineering data, semantic links between consumer and provider data, and the status on the EDEx process as foundation for effective EDEx for the use cases introduced in Section 3.2 and according to the required capabilities for EDEx in multi-disciplinary engineering, discussed in Section 3.1. Further, the EDEx process facilitates efficient

EDEx (a) by considering the benefits of EDEx for consumers and the cost for providers to focus first on the data sets with the best cost-benefit balance and (b) by automating the EDEx operation with support by the EDExIS.

As potential drawback of the EDEx process, the domain experts noted the need to convince data providers to take over the task and extra effort of extracting requested data from their engineering artifacts. For this task, specific tool support will be required according to the project context as well as appropriate compensation for the extra effort.

RQ2. What are main information system mechanisms that enable engineering data exchange for multi-disciplinary engineering? The EDExIS mechanisms for management and overview, data definition languages, and operation capabilities addressed the requirements for EDEx capabilities in Section 3.1 on a conceptual level. Together, the EDExIS mechanisms facilitate efficient round-trip-engineering among domain experts, i.e., the enrichment of common engineering concepts in iterations from several disciplines, as the domain experts may act both as consumers and providers. The design of an operational EDExIS will have considerable impact on the efficiency of the EDEx process in the application context and needs further investigation.

Limitations. As all empirical studies the presented research has some limitations that require further investigation.

Feasibility study. We evaluated the EDEx process approach with focus on specific use cases in cooperation with domain experts in a typical large company in PSE of batch production systems that can be seen as representative for systems engineering enterprises with project business using a heterogeneous tool and technology landscape. The evaluation results are based on observations from a limited sample of projects, stakeholder roles, and data models. To overcome these limitations, we plan a more detailed investigation in a wider variety of domains and application contexts.

The expressiveness of data specification and linking languages, used in the evaluated prototype, can be considered as a limitation. The prototype is able to address an initial set of simple data types, while industrial scenarios showed that value ranges and aggregated ranges have to be expressible in the desired data and link languages for specification and validation. While the evaluation worked well with data provided in tables, the evaluation of advanced data structures such as trees or graphs remains open.

6 Conclusion and Future Work

Digitalization in production system engineering (PSE) [17] aims at enabling flexible production towards the Industry 4.0 vision and at shortening the engineering phase of production systems. This results in an increase of parallel PSE, where the involved disciplines have to exchange updates for synchronization due to dependency constraints between the engineering disciplines.

In this paper, we introduced and investigated PSE use cases and the engineering data exchange EDEx process to provide domain experts in parallel

PSE with a systematic approach to define and efficiently exchange agreed sets of data elements between heterogeneous local engineering models as foundation for agile, traceable, and secure PSE. The EDEx process provides the foundations for addressing the characteristics of Responsible Information Systems, such as flexibility, privacy, trustworthiness, and security and specifically addresses major challenges introduced in Section 1.

*Ch1. Unclear data requirements of and benefits for stakeholders.*The EDEx definition phase results in a network of stakeholders linked via data they exchange. This network can grow iteratively, going beyond the insight of a one-time process analysis. The data in this network enables analyses of stakeholder relationships in an engineering project to provide the knowledge on which stakeholders require what data by when in the PSE process. Therefore, the EDExIS facilitates frequent synchronization between work groups to reduce the risk of divergent local designs, rework, and project delays.

Ch2. Heterogeneous engineering data is hard to integrate for sharing. Semantic linking allowed the integration of heterogeneous data in the evaluated use cases. The semantic linking enables seamless traceability in the EDEx process that, for the first time, gives all stakeholders the opportunity to know and analyze which role provided or received which kind of engineering data, which addresses a major awareness short-coming in the traditional EDEx process. Further, the EDEx semantic linking improves the representation of shared knowledge in the engineering team in a way that is understandable for machines, a prerequisite for introducing Industry 4.0 applications.

Future Work. Advanced analyses on the exchanged data and associated metadata. The EDEx data will enable consumers and researchers to conduct advanced analyses, such as on expected but missing values, data validity and consistency, and symptoms for security risks. The EDEx metadata allows analyses of PSE process characteristics.

Semantic linking between consumer and provider data models. During the use of EDEx, the complexity of links may grow considerably with the number of data elements, consumers, and providers, which will require research on scalability of EDEx.

IT Security considerations. Centralizing knowledge in the EDExIS will require research on threats to the integrity of collected knowledge and of industrial espionage.

Finally, future work will include the application and evaluation of the EDEx process and an operational EDExIS in various engineering domains and application areas.

Acknowledgment. The financial support by the Christian Doppler Research Association, the Austrian Federal Ministry for Digital & Economic Affairs and the National Foundation for Research, Technology and Development is gratefully acknowledged.

References

1. Stefan Biffl, Matthias Eckhart, Arndt Lüder, Thorsten Müller, Felix Rinker, and Dietmar Winkler. Data Interface for Coil Car Simulation (Case Study) Part I. techreport CDL-SQI-M2-TR02, TU Wien, 2018.
2. Stefan Biffl, Matthias Eckhart, Arndt Lüder, Thorsten Müller, Felix Rinker, and Dietmar Winkler. Data Interface for Coil Car Simulation (Case Study) Part II - Detailed Data and Process Models. techreport CDL-SQI-M2-TR03, TU Wien, 2018.
3. Stefan Biffl, Detlef Gerhard, and Arndt Lüder. Introduction to the multi-disciplinary engineering for cyber-physical production systems. In *Multi-Disciplinary Engineering for Cyber-Physical Production Systems*, pages 1–24. Springer, 2017.
4. Stefan Biffl, Dietmar Winkler, Richard Mordinyi, Stefan Scheiber, and Gerald Holl. Efficient monitoring of multi-disciplinary engineering constraints with semantic data integration in the multi-model dashboard process. In *Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA)*, pages 1–10. IEEE, 2014.
5. Marco Brambilla, Jordi Cabot, and Manuel Wimmer. Model-driven software engineering in practice. *Synthesis Lectures on Software Engineering*, 1(1):1–182, 2012.
6. Ambra Calà, Arndt Lüder, Ana Cachada, Flávia Pires, José Barbosa, Paulo Leitão, and Michael Gepp. Migration from traditional towards cyber-physical production systems. In *2017 IEEE 15th International Conference on Industrial Informatics (INDIN)*, pages 1147–152. IEEE, 2017.
7. Fred D Davis. *A technology acceptance model for empirically testing new end-user information systems: Theory and results*. PhD thesis, Massachusetts Institute of Technology, 1985.
8. Gregor Hohpe and Bobby Woolf. *Enterprise integration patterns: Designing, building, and deploying messaging solutions*. Addison-Wesley Professional, 2004.
9. Andrés Jiménez-Ramírez, Irene Barba, Manfred Reichert, Barbara Weber, and Carmelo Del Valle. Clinical processes-the killer application for constraint-based process interactions? In *International Conference on Advanced Information Systems Engineering*, pages 374–390. Springer, 2018.
10. Olga Kovalenko and Jérôme Euzenat. Semantic matching of engineering data structures. In *Semantic web technologies for intelligent engineering applications*, pages 137–157. Springer, 2016.
11. Arndt Lüder, Johanna-Lisa Pauly, Konstantin Kirchheim, Felix Rinker, and Stefan Biffl. Migration to automationml based tool chains–incrementally overcoming engineering network challenges.
12. Business Process Model. Notation (bpmn) version 2.0. omg (2011).
13. Susanne Putze, Robert Porzel, Gian-Luca Savino, and Rainer Malaka. A manageable model for experimental research data: An empirical study in the materials sciences. In *International Conference on Advanced Information Systems Engineering*, pages 424–439. Springer, 2018.
14. Michael Rosemann and Jan vom Brocke. The six core elements of business process management. In *Handbook on business process management 1*, pages 105–122. Springer, 2015.
15. Per Runeson and Martin Höst. Guidelines for conducting and reporting case study research in software engineering. *Empirical software engineering*, 14(2):131, 2009.

16. Marta Sabou, Fajar J Ekaputra, and Stefan Biffl. Semantic web technologies for data integration in multi-disciplinary engineering. In *Multi-Disciplinary Engineering for Cyber-Physical Production Systems*, pages 301–329. Springer, 2017.
17. Birgit Vogel-Heuser, Thomas Bauernhansl, and Michael Ten Hompel. Handbuch industrie 4.0 bd. 4. *Allgemeine Grundlagen*, 2, 2017.
18. Roel J Wieringa. *Design science methodology for information systems and software engineering*. Springer, 2014.