Conceptional Framework to Reduce Misinformation in Production System Engineering

S. Karch[1(✉)], A. Strahilov[1], and A. Lüder[2]

[1] let's dev GmbH & Co. KG, Alter Schlachthof 33, 76131 Karlsruhe, Germany
sabrina.karch@letsdev.de
[2] Chair of Production Systems and Automation, Otto-von-Guericke-University, Universitätspl. 2, 39106 Magdeburg, Germany

Abstract. In addition to the efficient use of materials, the significance of information as a resource increase in digital age. Digital tools are seen as a catalyst in generating and processing of digital data and information. One of the main challenges is to find the right information in the increasing oversupply of information as well as to manage this resource sustainably throughout the entire product life cycle. Misinformation is identified to be one of the main drivers of inefficiency in an engineering activity and is influenced by quantity and quality of data and information.

To further reduce this inefficiency, the study presents a conceptual framework that pursues the approach of demand-orientation as well as the smallest possible and appropriate quantum of data and information. Senior experts confirm the conceptual framework in semi-structured interviews. The outlook refers to use cases for investigation in a multi case study.

Keywords: Production Systems Engineering · Data and Information flow · Lean Digitalization

1 Introduction

The exchange of data and information on demand is the main focus, especially when experts from different disciplines work together collaboratively, such as in production systems engineering. The effectiveness and efficiency of collaboration as well as in each individual engineering activity depends on the logistical principle of information: right information, at the right time, in the right quantity, at the right place, in the right quality [1].

In today's processes, native data of interim or final results are converted to PDF or exchanged as MS Office file using email [2, 30].

In a currently discussed ideal, a process participant receives required data in a usable format by a tool-independent Common Data Model, a multilateral data platform. In this way, many actors have the possibility to collaborate on engineering data simultaneously and across companies based on current and consistent data [3].

T. Bauernhansl et al. (Eds.): WGP 2023, LNPE, pp. 3–13, 2024.
https://doi.org/10.1007/978-3-031-47394-4_1

The requirements for a data unit, named as data quantum, have to be defined and standardized. For this purpose, the study aims to develop and validate a fundamental framework for a specific data quantum and follows the research question:

How to define the smallest possible and appropriate quantum of data from the user's point of view?

First, an overview of the theoretical background and related work is given (Sect. 2). Based on this literature-centered study a conceptional framework is developed (Sect. 3). Section 4 raises the research methodology and Sect. 5 presents the research results validated by the semi-structured interviews. Section 6 discusses the results and limitations. Finally, a summary and an outlook for further work is given (Sect. 7).

2 Theoretical Background and Related Work

2.1 Data Management

Data management is an integral part of working with data quanta. Hanksche [4] considers data management to be "the planning, design, monitoring and control of the use, distribution and communication of information in organizations in order to achieve strategic goals". Data management includes all concepts that enable the organizational and technical implementation to provide data reliably, enable effective data use, and ensure data quality, data protection, and information security [4]. In addition, data management includes, e. g., data security, data distribution and access control [5].

Data are syntactic entities, input to interpretation processes and become information as interpreted data within contextual, semantic relations. Knowledge results from individual interpretation of information, experience and learning processes. Consequently, information is the input and output of a decision, which is finally made based on knowledge [6]. This is also the understanding behind Dippold's definition of data management. Value is created by translating information into decisions and actions. Data management has the function of making data available just in time to every process participant [7].

Relevant sub-aspects are implemented in different use cases, e. g. in the project SemAnz 4.0 with the target to describe data points semantically unambiguously [8]. ECLASS is a data standard for the classification of products and services using standardized ISO-compliant characteristics [9]. The data receiver using ECLASS can easily interpret the received data, but cannot evaluate the quality.

The project management can provide guidance on the question of the right amount of data by analogy with release planning. However, operational solutions are not discussed in this context [10].

The presented use cases cover relevant aspects of data management and working with data quanta, but do not represent a holistic view. The target is to gather significant aspects and transfer them into a generally valid framework.

2.2 Lean Paradigm in Data Flow

Lean is known for the paradigm of eliminating waste and work without added value [11] and addresses the effectiveness and efficiency of activities and processes.

The customer is the starting point for lean systems approach, which is summarized by Womack and Jones in five principles [12].

In the context of data and information, three aspects are of key importance: customer orientation, elimination of waste and the ideal of one-piece flow.

Customer Orientation. Ohno underlines the role of the customer by affirming that all activities, including the processing of information, are only relevant if they result in products, services or an activity, customers want to buy (T. Ohno according to [13]). In reference to this customer centricity, a downstream process step is also to be considered as an internal customer whose needs and requirements must be taken int account [14]. In the context of transformation of data and information, these internal customers have information needs. Consequently, needs orientation is to be seen as a prerequisite for an effective and efficient processing of data and information in accordance with the information logistics principles.

Elimination of Waste. Waste, or muda in Japanese, means "to toil" or "pointless effort", has an immediate impact on effectiveness and efficiency and includes all activities that do not benefit the customer [14]. Misinformation is identified to be one of the main drivers of inefficiency in an engineering activity and is influenced by quantity and quality of data and information [15]. In the increasing availability of data and information, misinformation must be eliminated to ensure an efficient process.

One Piece Flow. An ideal from a lean point of view is the so-called one-piece flow or the supply with the smallest possible units so that the customer's needs are met on demand [14]. This idea is also to pursue in the context of data in order to meet demand orientation and the logistical principles of data and information [1].

The resulting framework should meet these requirements of lean paradigm.

2.3 Data-Driven Engineering

Production System Engineering (PSE) is a model-based and multi-disciplinary process involving various domain experts [16] and considering a Production System of Systems (PSoS). PSE is highly complex, often concerning coevolving products and system components [17]. To handle the complexity and ensure effective and efficient engineering, different engineering process methods have been developed and standardized considering the designed PSoS as a digital value creation chain [18]. Major representatives of engineering guidelines are ISO 18828 [19], VDI/VDE Guidelines 2221 [20], 2206 [21], and 4499 [22].

PSE engineers in different domains take design decisions based on suitable engineering knowledge [23]. The engineering activities form a network, linked by data exchange of externalized engineering knowledge [24]. The interaction of engineers is challenged [25] by the different engineering habits, especially limiting efficient and effective information exchange and utilization. Especially the engineering quality evaluation is impeded by a missing integrated information representation covering the individual knowledge of the experts and their cross-discipline.

3 Conceptional Framework

The study proposes a conceptional framework based on the literature reviewed, to respond the research question. The framework is intended to meet the requirements of the lean paradigm (see Fig. 1) described in Sect. 2.2.

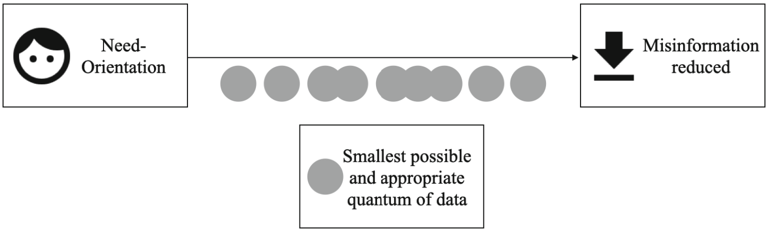

Fig. 1. Correlation of lean paradigm and misinformation.

The framework intends a description of a data quantum through five main aspects. Each data quantum contains information about security, transaction, data type, rights and roles, and data quality (see Fig. 2).

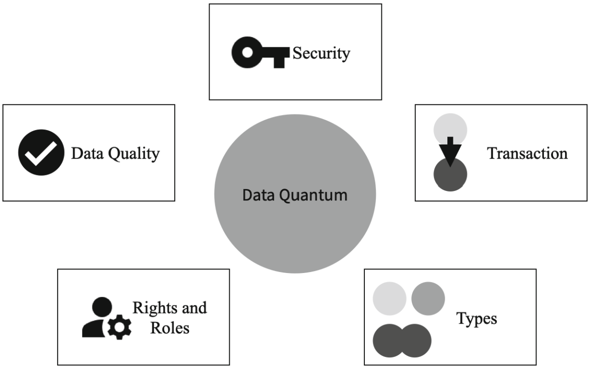

Fig. 2. Main aspects of a data quantum.

Security. The category security includes the technical and organizational protection of data itself as well as of data transmission. This includes, for example, the transmission technology, encryption or the API (Application Programming Interface) interface [26].

Transaction. In a transaction, a dataset is left in a consistent state after error-free execution of an instruction. Changed data can be transmitted via the pull or push principle as well as in a permanent transmission [26].

Data Type. This category gives an overview of how data can be organized to be exchanged. With the data types it is to be distinguished between values, attributes, properties, structure, class, list with values, files and formats. It can also be a set of data in the form of files in various formats.

Rights and Roles. With the rights and roles concept, it is important to distinguish which roles exist and which rights, such as read, write, export and convert, are permitted. The question of data owner also needs to be clarified. Rights and roles can also be assigned to data security, but are listed separately here due to their importance [5].

Data Quality. Data quality includes in particular, the topicality, completeness and consistency of data [5].

4 Method

To address the research question, the qualitative study is based on the Design Science Method and semi-structured interviews with senior experts. Interviews, as data collection method, are used to gather complex information, opinions and experiences. The semi-structured interview allows the respondents to answer in their own words and to contribute their ideas and experience. Expert opinion is a convenient way to reflect the conditions of practice and can be used to validate artifact designs at an early date [27, 28].

The interview guide is designed for an interview of about 30 to 45 min and follows a general structure of three main parts: welcome message with a note about recording, the main part with questions, a statement of gratitude for participation [28].

The interview questions are listed in the Table 1. Questions Q1 to Q5 refer to the categories presented in Sect. 3. In question Q6, the experts have the opportunity to add another category to the framework. Question Q7 and Q8 aims to ask the experts about barriers and what needs to be done to make the framework work.

Table 1. Overview of interview questions.

ID	Question
Q1	Can you confirm the category “security”? What is missing in your expertise?
Q2	Can you confirm the category “transaction”? What is missing in your expertise?
Q3	Can you confirm the category “data types”? What is missing in your expertise?
Q4	Can you confirm the category “roles and rights”? What is missing in your expertise?
Q5	Can you confirm the category “data quality”? What is missing in your expertise?
Q6	Would you like to add a category?
Q7	Why can’t the framework be successful from your point of view?
Q8	What to do to make the framework work?

Table 2 provides an overview of the interviews and give background information on the interview partners. These have experience in computer science and software development or consider the framework from a user perspective as an engineer.

The results of the interviews have been transcribed and analyzed using Mayring’s qualitative content analysis method. Inductive category development was applied, whereby the categories are formed spontaneously during the analysis of the transcribed interviews [29].

Table 2. Overview of interviews and interview partners.

Professional Activity	Professional Experience (years)	Position	ID
Software Development	22	Head of Development	I1
Software Development	16	CTO	I2
Engineering	28	Managing Director	I3
Engineering	28	Senior Researcher	I4

5 Results

The results of the expert interviews are presented and interpreted below, with a particular focus on the different opinions expressed by the experts, but also on the areas of agreement between them.

5.1 Results for the Content-Related Questions (Q1–Q6)

All experts confirm the importance of the framework, in particular the idea of standardization and the categories defined. As a result of their expertise, the following additions are of particular interest.

Question Q1 Security. In this category, the focus in the interviews is on three aspects, the protection of the data itself, the encryption of data during transmission, and the encrypted transmission path.

Manipulation of data is possible at the sender, on the transmission path as well as at the receiver, therefore an authenticity check should be conducted (I1), (I4). Signed data (I1) as well as copy protection is also a way to protect data (I2). To ensure that data does not get to unauthorized persons, authorized persons can verify themselves, e. g. via a digital ID card. A two-factor authentication can additionally protect data, in the simplest case via e-mail or using TOTP (Time-based On-time Passwords) (I2).

TOTP uses an algorithm to generate a temporary passcode using the current time of day as an authentication factor [34].

The experts agree that cost and benefit of *encrypting data or data packets* must be weighed up and should only be considered for particularly security-critical data in a collaborative use case (I1), (I2). Not separately worth protecting are, e. g. data of purchased parts. The closer a plant or PLC (programmable logic controller) program gets to the final state, the more protective the specific plant data are (I3). In standard cases, the *encryption of devices and transmission* channel should be secured, e. g. with HTTPS (Hypertext transfer protocol secure) encryption (I1), (I2).

HTTPS is based on HTTP and uses an additional protocol to encrypt data and transmission of data [32, 33].

Question Q2 Transaction. Each transaction must be provided with a *unique ID* and an authenticity check must be performed. Third party libraries make it possible to check the scope that has changed, to compare the entire document and to verify the update instead of checking data that has not changed with each modification (I2).

Versioning can be used, where access login can be tracked and meta-information about the changes can be provided (I1).

The advantage of *pull request* is that the data can be retrieved on demand. With a *push transaction*, the relevance must be configured in advance, similar to a PubSub (publishing subscribe) mechanism (I1). The challenge is to find out whether new data is relevant for a user. One approach is that this information must be read by all users and the access is logged, whereby data protection according to GDPR (General Data Protection Regulation) must be taken into account (I1). The GDPR regulates the processing of personal data of natural persons by natural persons, companies or organizations in the EU [31].

Information needs may change over the project duration. If the maintenance or operations department takes over responsibility from the project manager for a new plant, e. g. the project manager will only need information on changes (I3).

Question Q3 Types. In addition to classic data formats, a *stream* could also be considered for live data (I1), (I2). In interview (I4) the object reference was added, i. e. in relation to which artifact data is exchanged.

Question Q4 Rights and Roles. Data construct forms the fundament for the implementation of the *rights and role allocation.* An ACL (Access Control List) can be used to specify users and authorizations in regards to data (I2).

An example of the importance of roles and rights is the programming of the PLC software, especially for safety functions. It must be ensured that only trained employees can make changes. The transfer of responsibilities is also conceivable, e. g. if the PLC source code is transferred to the client (I3). The exchange of a subset of existing data, e. g., on an as-needed basis to external suppliers, can also be relevant and could be regulated via rights and roles (I4).

In the collaboration, it is important to establish *hierarchies or representative* regulations, so that, for example, all parties, engineers or lawyers, can accept modified data or specific parts of it depending on their area of expertise in a contract or operating manual (I2). In principle, the *review process* by the parties involved improves the quality of the changed data (I1).

A *reminder* can be a useful function in this context (I1), (I2) and an automatic reminder up to an automatic escalation is conceivable (I1). The challenge is to find balance between reminders and acceptance by users without forcing them too much into a process flow. Not to forget that an application built on the framework takes on a kind of social function (I1).

A *recommender* or an artificial intelligence in compliance with data protection could help to find data immediately or make suggestions about relevant data (I2).

The importance of *transparency* about given information, if necessary prioritized according to groups of people, is emphasized in the interview (I3), but it is important that changes can only be made by authorized persons.

Question Q5 Data Quality. A data quantum should always have an *identifier*, when and by whom it was created or changed, in the sense of versioning (I1). In general, data quality can be increased with the help of defined *responsibilities* (I2).

The *evaluation of quality* is challenging in the case of machine or automatically generated data, such as e. g. in a simulation (I2). Moreover, data quality is often a subjective perception, which makes it difficult to measure (I1).

In networked systems, a technical failure can also have a negative impact on data quality and *actuality* if, e. g. data has not yet been synchronized at a certain point. Nowadays, data actuality is supported with the help of the PubSub mechanism (I2).

It is essential for data quality and actuality, that all parties involved can access the *same data in one place* (I3). In terms of actuality, it is necessary to provide status information such as draft or final (I3).

Question Q6 Add a Category. In the interview (I2), the importance of *data protection* in accordance with GDPR is emphasized as a separate category for a data quantum.

GDPR compliance must be checked in two steps. First, is there data that is covered by data protection and second, do the technical solutions cover the requirements of the GDPR, such as regulated transfer, expiration dates, and deletion periods.

5.2 Results for the Implementation-Oriented Questions (Q7–Q8)

Question Q7 Why can't the Framework be Successful. From a *technical perspective*, the experts see no reason for the failing of the framework and a demand-driven supply of data quanta (I1), (I2).

Of particular importance, however, is user *acceptance* as a social aspect. If an application does not bring the user any tangible benefits, such as time savings, better information supply, or if user feel overregulated and limited in their competence or importance, the framework can fail (I1), (I2), (I4).

Another reason for failure may be that the application based on the framework becomes too *complex*, so that the solution is not used. Also, unprofessionally maintained *responsibilities*, bring the framework to failure (I2).

Question Q8 What to do to Make the Framework Work. In the first place, it is necessary to create *acceptance* for the framework and related applications, by means of comprehensible documentation and a simple possibility of using the framework. A balance must be struck between technically feasible and sensible solutions, with a view to acceptance by users (I2).

An *MVP* (minimum viable product) must be defined that covers the minimum requirements of the parties and has at least the same acceptance as today's solutions, whereby a structural improvement in the background makes the decisive difference (I1), (I2). The framework can be successful if it is implemented in such a way that it not only meets the requirements of a research project but also takes *critical mass* into account (I1).

6 Discussion and Limitations

The framework was confirmed in technical feasibility. The proposed categories can be retained, supplemented by the category Data protection according to GDPR. The presented framework gives a basic, standardized ordering concept for working with data quanta and thus answers the research question presented at the beginning of this paper. The results of the interviews can be transferred into requirements for working with data quanta as well as applications based on the framework. In addition to technical aspects, social aspects that affect acceptance are also in the foreground, which emphasizes a further design from the user's point of view.

It cannot be excluded that other aspects might be relevant when further experts are interviewed. In addition, the focus of the study is on the design of the framework for a data quantum, the interconnection of multiple data quanta needs to be further investigated. Applying the framework to one or more use cases can also provide further insights.

7 Conclusion and Outlook

The demand-oriented supply of data and information out of the oversupply is one of the main challenges in digital age. In production systems engineering, data and information are the equivalent of material in production. An efficient, sustainable handling of these and the elimination of misinformation have a direct impact on the efficiency of the entire product life cycle. In order to be able to organize data in a needs-oriented manner, the research question of *how to define the smallest possible and appropriate quantum of data from the user's point of view* was investigated and the framework of the data quantum elaborated.

In semi-structured interviews with experts, this framework was confirmed and initial, detailed requirements were developed.

The framework is to be investigated in a multi-case study as part of further research activities. The aim is to analyze its operational applicability and to gain insights into its practicability. Attention will be paid to the provision of data quanta in line with demand-oriented requirements, technical implementation and acceptance by users.

Acknowledgement. Parts of this work were supported by the Federal Ministry of Education and Research (BMBF) within the research project H2Giga – FertiRob.

References

1. Augustin, S.: Information als Wettbewerbsfaktor – Informationslogistik. TÜV Rheinland, Köln (1990)
2. Strahilov, A., Hämmerle, H.: Engineering workflow and software tool chains of automated production systems. In: Biffl, S., Lüder, A., Gerhard, D. (eds.) Multi-Disciplinary Engineering for Cyber-Physical Production Systems, pp. 207–234. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56345-9_9

3. Lüder, A., Baumann, L., Behnert, A.-K., Rinker, F., Biffl, S.: Paving pathways for digitalization in engineering - common concepts in engineering chains. In: 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 1401–1404, ETFA, Vienna, Austria (2020)
4. Hanksche, I.: Digitaler Wandel- Lean & Systematisch. Springer Vieweg, Wiesbaden (2021)
5. Hildebrand, K., Gebauer, M., Mielke, M.: Daten- und Informationsqualität – Die Grundlage der Digitalisierung. Springer Vieweg, Wiesbaden (2021)
6. Aamodt, A., Nygård, M.: Different roles and mutual dependencies of data, information, and knowledge - an AI perspective on their integration. North-Holland Elsevier, Data Knowl. Eng. **16**, 191–222 (1995)
7. Dippold, R., Meier, A., Schnider, W., Schwinn, K.: Unternehmensweites Datenmanagement – Von der Datenbankadministration bis zum Informationsmanagement. Springer Vieweg, Wiesbaden (2005)
8. Semantische Allianz für Industrie 4.0.: https://www.hsu-hh.de/aut/forschung/forschungsthemen/semantische-allianz-fuer-industrie-4-0. Accessed 06 June 2023
9. Eclass. https://eclass.eu/. Accessed 06 June 2023
10. Felkai, R., Beiderwieden, A.: Projektmanagement für technische Projekte – Ein Leitfaden für Studium und Beruf. Springer, Heidelberg Dordrecht London New York (2015)
11. Ohno, T.: Toyota Production System – Beyond Large-Scale Production. Productivity Press, New York (1988)
12. Womack, J.P., Jones, and D. T.: Lean Thinking – Ballast abwerfen, Unternehmensgewinne steigern. Campus, Frankfurt am Main (2013)
13. Liker, J.K., Braun, A.: Der Toyota Weg - 14 Managementprinzipien des weltweit erfolgreichsten Automobilkonzerns. FBV, München (2014)
14. Bertagnolli, F.: Lean Management – Introduction and In-Depth Study of Japanese Management Philosophy. Springer, Wiesbaden (2022)
15. Karch, S., et al.: Lean Engineering – Identifying waste in engineering chains. In: 56th CIRP Conference on Manufacturing Systems, CIRP CMS, South Africa (2023). accepted
16. Biffl, S., Lüder, A., Gerhard, D.: Multi-disciplinary Engineering for Cyber-Physical Production Systems: Data Models and Software Solutions for Handling Complex Engineering Projects. Springer, Cham (2017)
17. Mehr, R., Lüder, A.: Managing Complexity within the Engineering of Product and Production Systems. In Biffl, S., Eckhart, M., Lüder, A., Weippl, E. (Eds.): Security and Quality in Cyber-Physical Systems Engineering, pp. 57–79. Springer, Cham (2019)https://doi.org/10.1007/978-3-030-25312-7_3
18. Paetzold, K.: Product and systems engineering/CA* tool chains. In: Biffl, S., Lüder, A., Gerhard, D. (eds.) Multi-Disciplinary Engineering for Cyber-Physical Production Systems, pp. 27–62. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56345-9_2
19. ISO/TR 18828: Industrial automation systems and integration - Standardized procedures for production systems engineering, Series of 5 parts (2018–2020)
20. VDI/VDE 2221–2019: Design of technical products and systems - Model of product design. Beuth Verlag (2019)
21. VDI/VDE 2206–2021: Development of mechatronic and cyber-physical systems. Beuth Verlag (2021)
22. VDI/VDE 4499 2008–2022: Digital factory. Beuth Verlag. 5 parts (2008–2022)
23. Drath, R., Barth, M., Fay, A.: Offenheitsmetrik für Engineering-Werkzeuge. Atp Edition, **54**(09), 46–55 (2012)
24. Behnert, A.-K., Rinker, F., Luder, A., Biffl, S.: Migrating engineering tools towards an automationml-based engineering pipeline. In: IEEE 19th International Conference on Industrial Informatics (INDIN), pp. 1–7 (2021)

25. Biffl, S., et al.: Experiences with Technical Debt in Parallel Multi-Disciplinary Systems Engineering Christian Doppler Laboratory for Security and Quality Improvement in the Production System Lifecycle. Technical Report (2019)
26. Fink, A., Schneidereit, G., Voß, S.: Grundlagen der Wirtschaftsinformatik. Physica-Verlag, Heidelberg (2005)
27. Wieringa, R. J.: Design Science Methodology for Information Systems and Software Engineering.Springer, Berlin Heidelberg (2014). https://doi.org/10.1007/978-3-662-43839-8
28. Johannesson, P., Perjons, E.: An Introduction to Design Science. Springer, Cham (2014)
29. Mayring, P.: Qualitative Inhaltsanalyse – Grundlagen und Techniken. Beltz, Weinheim (2010)
30. Schleipen, M., Mersch, T., Karch, S.: Engineering und Inbetriebnahme von Produktionsanlagen – Eine Analyse der beteiligten Prozesse und des entstehenden Datenflusses. Atp Magazin (2023) – accepted
31. GDPR. https://eur-lex.europa.eu/eli/reg/2016/679/oj. Accessed 25 Jul 2023
32. RFC Standard 2818. https://datatracker.ietf.org/doc/html/rfc2818. Accessed 25 Jul 2023
33. RFC Standard 2616. https://www.ietf.org/rfc/rfc2616.txt. Accessed 25 Jul 2023
34. RFC Standard 6238. https://datatracker.ietf.org/doc/html/rfc6238. Accessed 25 Jul 2023