The Human Side of Service Engineering, Vol. 62, 2022, 319–327
https://doi.org/10.54941/ahfe1002573

Towards a Maturity Model to Measure Data Consistency in the Manufacturing Industry

Maximilian Feike and Philipp Christel

Fraunhofer IAO Stuttgart, 70569, Germany

ABSTRACT

The deployment of Artificial Intelligence (AI) in the manufacturing context is said to provide significant benefits to organizations. However, many manufacturers struggle to meet the requirements necessary for the use of AI technologies within their company. A major challenge is linking and processing the existing data into a way in which it can be reliably processed by AI algorithms. This is especially relevant in established companies characterized by a historically grown bulky and decentralized IT infrastructure. Moreover, in many of these companies there is no common understanding of data consistency. Therefore, we investigate the diverse dimensions of data consistency and set the foundation for a maturity model to assess a company division's status quo. Based on our literature review and four interdisciplinary and iterative workshops conducted with experts from an automotive OEM, we developed the concept of a maturity model for data consistency that provides situation-specific recommendations for further improvement.

Keywords: Data consistency, Data continuity, Enterprise legacy systems, Future factories, AI services, AI requirements, Data set, Maturity model

INTRODUCTION

With the rise of AI in recent years, the importance of data as basis for operational and strategic excellence of manufacturers has increased. However, many companies find themselves in a world of (outdated) legacy systems with an amount of data that is continuously increasing (Steven, 2018). Especially in industries with a high level of customization, resulting in complex products and production processes, manufacturers still struggle to reach system interoperability and continuity all over their software infrastructure (Huang et al., 2020).

A company-wide interconnected data set is the base for advanced information exploitation and new technological possibilities, such as for advanced analytics, smart applications, or highly efficient AI solutions (Kuhn and Franke, 2021). Hence, it is necessary to connect existing systems and overcome data and information silos (Breivold, 2019). To drive improvement, it is necessary to measure progress from a data perspective. This is especially relevant for companies implementing a uniform data governance.

Therefore, our work aims to develop a multi-level maturity model, which enables divisions or business units to classify themselves in terms of their current level of data consistency and provides situational recommendations towards seamlessly interconnected enterprise IT systems. The application of our instrument should allow a structured and holistic evaluation of a business unit's current position on its journey to full data consistency. A catalogue of criteria for an objective measurement is to be developed and is supposed to be applicable by industry experts without further guidance. The recommendations for action will allow its users to derive specific measures tailored to the current level of maturity.

NEED FOR A MATURITY MODEL TO MEASURE DATA CONSISTENCY

In the manufacturing context, the modernization of legacy enterprise systems is a challenge faced by many organizations looking to leverage new technologies in order to meet the demands of high scalability and high availability (Furda et al., 2018). Often, companies still operate legacy systems characterized by old processes, systems, or application programs, even though newer technologies are now available (Jha et al., 2014). One of the overarching goals is to connect all manufacturing and automation processes and seamlessly integrate systems to overcome the still predominant information silos (Breivold, 2019).

Achieving a high level of data consistency as an established manufacturer is a success critical highly complex undertaking that requires a step-by-step approach (Witsch and Vogel-Heuser, 2012). To be effective, the measures to be taken must be tailored to the current state of development. This requires an assessment of the organization's current maturity level and the identification of appropriate measures that contribute to accomplish the next level.

The maturity model developed is intended to contribute to this assessment. As the establishment of a company-wide interconnected data set is a fundamental requirement for advance technologies like AI services and machine learning solutions (Kuhn and Franke, 2021) our maturity model aims to enable manufacturing companies to benefit from the potential of these promising solutions. Furthermore, our work is intended to contribute to a common understanding of data consistency.

CONCEPTUAL BACKGROUND

In literature, there is uncertainty about a common definition of data consistency, its properties, and dimensions (Wada et al., 2011). Furthermore, the terms "data consistency" and "data continuity" are used synonymously (Furda, 2018; Kizipinar, 2021; Wada et al., 2011) whereas others use the paraphrase of a seamless and fully integrated system (IEC, 2015). The variety of terms used in literature points out that there is a lack of a uniform terminology. In most cases, the understanding of data consistency only covers the technical level. However, this isolated view leaves out important aspects that are also relevant for understanding and achieving data consistency in established companies. For this reason, we want to contribute to a holistic and interdisciplinary understanding of the concept of data consistency. We use the HTO Framework (Karltun et al., 2017; Karwowski et al., 1994) as a starting point to investigate the diverse dimensions of the term data consistency. This concept allows to gain a holistic understanding of what data consistency means in manufacturing related organizations considering human, organizational and technological aspects.

To determine a company's status quo in a specific area, the concept of maturity models has proven to be suitable (Wendler, 2012). In recent years, a variety of such models were developed for a wide range of applications. While some of these concepts focus on more very specific topics like digital business processes (Bitcom, 2020) others are designed to meet the needs of a single sector or industry (Machado et al., 2021). Some models follow a more generic approach to allow a holistic evaluation of a manufacturer's current state of digitalization (Schuh et al., 2020; Weber et al., 2017). Others use existing approaches from the field of data quality assessment, but with a narrower understanding (Bleiholder and Steinhäußer, 2011). However, none of these approaches provides a holistic and interdisciplinary understanding of data consistency.

RESEARCH METHOD

To achieve the goal of this work, a literature review is executed to identify existing maturity models from different research fields. In addition, our literature review covered existing stages of operational data management and architecture as well as methods to enable the quantifiability of data-related maturity. Based on our literature review, four interactive expert workshops were conducted to elaborate a common understanding of the term data consistency by following an iterative process. To gain a holistic view from different perspectives, project managers of an automotive OEM, IT business consultants, engineers of a large software provider and representatives of the scientific community were brought in for the first three workshops. The results were then reviewed and discussed with the OEM's senior management as well as representatives of a leading cloud provider. Adaptations, comments and change requests were brought together in a final workshop to create a common understanding of data consistency.

Based on this preliminary work, a first version of a five-level maturity model was developed. After an initial revision by experts with several years of experience within the field of IT development, the model was tested for comprehensibility and applicability. Therefore, we executed case studies in two production-related IT departments. In a last step, the results of our case studies were discussed.

A MATURITY MODEL FOR DATA CONSISTENCY

To develop a common understanding of the term data consistency among all participants, four interactive expert workshops were conducted to iteratively discuss the phenomenon with all its dimensions and characteristics in detail. As a result of this steps, the experts came up with a total of 25 elements that

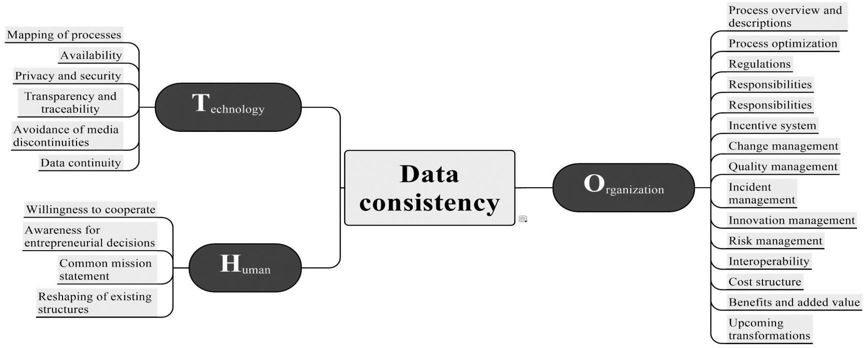

Figure 1: Elements of data consistency.

allow a comprehensive description of data consistency in a manufacturing context. According to the HTO approach these elements were allocated to the three dimensions Human, Technology and Organization (see fig. 1).

The path to a data-centric organization is individual for each company. The implementation of company-wide data consistency always requires a reorganization of IT systems and processes and is usually a transformation process lasting several years. At the beginning of every transformation, a current state analysis is necessary to clearly define the objectives. As mentioned before, maturity models are particularly suitable for recording the status quo of a transformation and making progress measurable. In Germany, the Acatech Industry 4.0 Maturity Index was developed to give guidance to companies on their way to the full implementation of Industry 4.0. This practically orientated maturity model examines companies from a cultural, technological, and organizational perspective (Schuh et al. 2017).

By combining our understanding of the term data consistency and its 25 elements with the underlying logic of the widely used Acatech Industry 4.0 Maturity Index, the result of our work is a maturity model for data consistency. It consists of the following levels: (1) Patchy and incomplete data basis, (2) Data silos: Lack of connectivity between areas, (3) Resolution of interface problems: Complexity reduction, (4) Data consistency: Transparency across all processes, (5) Data consistency as an enabler for (autonomous) self-optimization (Fig. 2).

Using a specifically developed questionnaire, our instrument enables a comprehensive assessment of the status quo based on the identified 25 elements of data consistency. Fig. 3 shows how the individual elements are interlinked.

The levels of the maturity model reflect the increasing requirements on the IT systems of a company and its organization. This means that maturity level 2 can be achieved with less effort than level 4, for example. The maturity levels thus build on each other. It is also a question of corporate strategy which level is to be reached as transformation processes also take place against the background of a cost and benefit perspective.

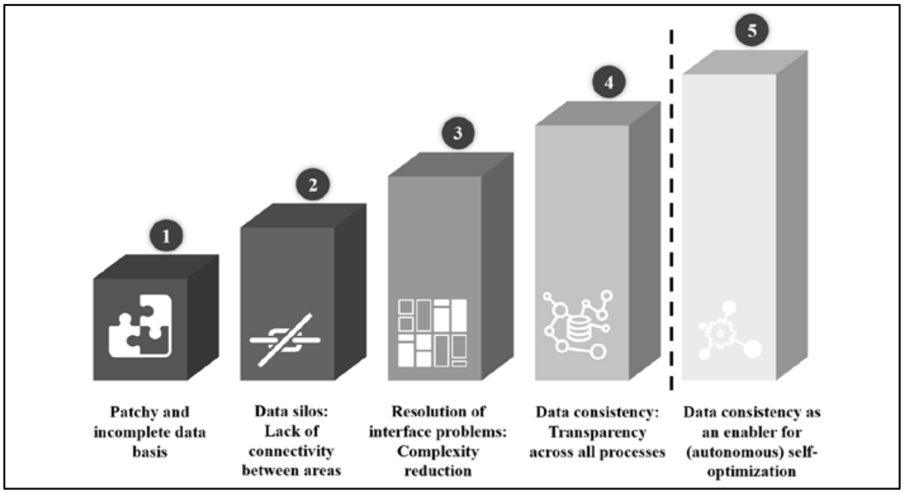

Figure 2: Multi-level maturity model for data consistency.

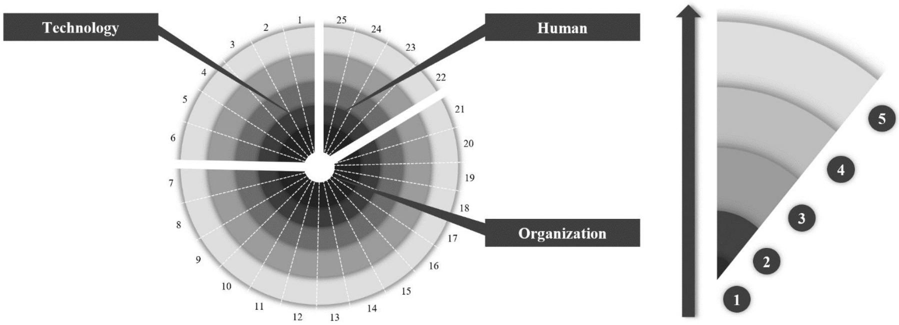

Figure 3: Structure and levels of the maturity models.

Levels of the Maturity Model

Patchy and incomplete data basis - Level 1: The result of structures that have grown over the years is a patchy and incomplete database. The pressure of digitization and the accumulation of ever new transitional solutions lead to an inhomogeneous IT system world that does not meet the existing requirements. Insufficient investments in IT systems result in low productivity and high manual effort. Organizational processes as well as system responsibilities are not clearly defined. Existing lack of understanding for business challenges and the relevance of information technology. Bureaucratic hurdles to innovation.

Data silos: Lack of connectivity between areas - Level 2: In this stage technical solutions towards data consistency are developed but only with focus on divisions or business units. Data silos are created by delimited system worlds. Improvements in the IT landscape already have a positive impact on productivity within the optimized areas. Regulations and processes are defined, but approaches differ across divisions. The willingness to cooperate and the awareness of business challenges only exist within the divisions (silo thinking).

Resolution of interface problems: Complexity reduction - Level 3: Level of customized interface solutions. Access to data is possible across divisions, but there are widespread interface problems between IT systems. Efficiency and productivity can be further improved, especially by optimizing interfaces. Roles and rights management is sometimes contradictory, and responsibilities are not defined across systems. Although working together on data consistency, business units act within the existing legacy systems.

Data consistency: Transparency across all processes - Level 4: Stage of a holistic solution for data consistency and the avoidance of media discontinuity. All processes are mapped precisely. Data is available anytime and anywhere. No inconsistencies between systems. Low maintenance effort and high saving potentials realized through optimized processes. IT systems ensure high productivity. Cross-divisional distribution of roles, rights, and responsibilities. An elaborated data governance concept is present. Understanding of the added value of data, the possibilities of IT and necessary organizational changes.

Data consistency as an enabler for (autonomous) self-optimization-Level 5: In this stage IT Systems enable self-optimizing production and logistics processes. Together with a complete database, artificial intelligence programs can be brought in. Synchronized data lakes function as a single source of truth. There is no interference due to interface problems. Processes are automatically updated. New business models and massive saving potentials are created with the help of full data consistency. Group-wide data strategy serves as the basis for all IT-related decisions. Living data culture. The company's organizational structure follows the IT architecture.

A catalog of criteria was developed for the self-assessment of the maturity level (appendix, table 1). Industry experts can use this systematic approach to measure progress during the transformation. Looking at the results of an assessment, the areas examined are by no means always at the same maturity level. For example, the technological area may already be significantly more developed than the organizational level. However, the maturity model indicates which of the 25 elements require special attention and thus enables the prioritization of resources to the identified constraints. In terms of guidance, recommendations are presented to improve vulnerabilities. Especially for the current state analysis, the maturity model is a suitable tool to enable comparability across divisions or business units.

CONTRIBUTIONS, LIMITATIONS AND OUTLOOK

Our work contributes to an integrated understanding of the phenomenon of data consistency in an organizational context. Allocating the elements of data consistency to the three dimensions of the HTO-approach enables an extended consideration beyond a purely technological level. The maturity model enables organizations and departments to determine their status and derive situation-specific measures. In addition, our method allows to track progress over time and creates comparability between different units.

Of course, this research does not come without limitations. First, the generalizability of our results is limited, as we only used one case of a German

Table 1. Catalog of criteria for industry experts to self-assess the maturity level (example of the dimension "technology").

	Question	**Patchy and incomplete data basis**	**Data silos: Lack of connectivity between areas**	**Resolution of interface problems: Complexity reduction**	**Data consistency: Transparency across all processes**	**Data consistency as an enabler for (aut.) self-opt.**
Mapping of processes	Are processes up-to-date + precisely mapped?	Processes are not mapped.	Processes are only partially mapped.	Processes are mapped completely, but organized decentrally.	All processes are mapped realistically and precisely.	Processes update continuously.
Availability	Data access possible?	Access to the data is not possible.	Data is available in silos, but not in other areas.	Data is available, but with many interfaces between areas.	All data is available anytime and anywhere.	Data is provided proactively.
Privacy and security	Technical data protection concept defined?	Data protection concept does not exist.	Data protection is only elaborated in the own area.	The cross-domain concept exists, but with inconsistencies between systems.	Data protection concept exists across all areas. Authorizations are clarified.	Self-monitoring systems. Autom. monitoring mechanisms integrated into data protection concept.
Transparency and traceability	Data stored transparently in the system?	Clear allocations are not available. Documentation does not exist.	Data is stored in a traceable manner within the area. Documentation within the area exists.	Data is traceable across departments but interfaces cause problems. The documentation is not consistent across systems.	Uniform and transparent data flow in the system.Changes documented. Metadata described, data map available.	Traceability is given in all areas. A "single source of truth" has been established.
Avoidance of media discontinui-ties	Do interfaces exist between the legacy systems? (Connection of data systems)	Large number of media breaks in processes.	No media discontinuities within the own area. No interfaces to other areas (silo solution).	Interfaces are compatible. However, too many individual custom solutions still exist.	There are few media breaks, but interfaces are optimized for all areas. The data can be transferred.	Only one system available: Data lake. Interface problems are no longer a hurdle.
Data continuity	Data stored consistently in all systems?	Inconsistencies, and incomplete information / cause problems.	Duplicates/errors in the system exist, but rarely occur within a defined area.	Cross-domain data consistency. Manual maintenance effort due to discontinuities.	No inconsistencies between the system worlds. Machine data is described uniformly.	Achieved single source of truth enables real-time synchronization in all downstream systems.

OEM. Second, the recommendations, resulting from our maturity model do not consider the resources required for each measure and their constraints within an organization. Third, there are multiple interdependencies between the overall dimensions (human, technology, and organization) which makes an isolated consideration of some criteria difficult. The following implications for further research can be made. Further research and publications will be necessary to help companies understand the relevance of seamlessly interconnected IT systems, especially in regards of international competitiveness. As our research shows, the phenomenon of data consistency needs further explanation. In order to verify the validity of our results, the individual elements should be examined for their transferability to other industries.

APPENDIX

See Table 1.

REFERENCES

Bitcom (2020), Reifegradmodell Digitale Geschäftsprozesse. Leitfaden. Available at: https://www.bitkom.org/sites/default/files/2020-04/200406_lf_reifegrad modell_digitale-geschaftsprozesse_final.pdf (Accessed: 28 January 2022)

Bleiholder, J., Steinhäußer, D. (2011). Reifegrad-Modell für Datenqualitäts-Assessment. Available at: https://www.doag.org/formes/pubfiles/2485831

/2011-01-BNews_Jens-Bleiholder-David-Steinhaeusser-Reifegrad-Modell-fur-Datenqualitaets-Assessment.pdf (Accessed 28 January 2022)

Breivold, H. P. (2019). Towards factories of the future: migration of industrial legacy automation systems in the cloud computing and Internet-of-things context. Enterprise Information Systems Vol 14(49, pp. 542–562

Furda, A., Fidge, C., Zimmermann, O., Kelly, W., Barros, A. (2018). Migrating Enterprise legacy Source Code to Microservices: On Multitenancy, Statefulness and Data Consistency. IEE software, 2018–05, Vol. 35 (3), pp. 63–72

Electrotechnical Commission

Huang, Z., Jowers, C., Dehghan-Manshadi, A. and Dargusch, M. (2020). Smart Manufacturing and DVSM Based on an Ontological Approach. Computers in Industry. Vol 117, p. 103189

IEC (2015). Factory of the future. Geneva, Switzerland: International

Jha, S., Jha MO'Brien, L. and Wells, M. (2014). Integrating Legacy Systems into Big Data Solutions: Time to make the Change. In Proceedings of the IEEE Conference on Asia-Pacific World Congress on Computer Science and Engineering. Washington DC: IEEE Computer Society; Nadi, Fiji, 4–5 November 2014.

Karltun, A., Karltun, J., Berglund, M., Eklund, J. (2017). HTO – A complementary ergonomics approach. Applied Ergonomics 59, pp. 182–190

Karwowski, W., Salvendy, G., Badham, R. Brodner, P., Clegg, C., Hwang, S.L., Iwasawa, J., Kidd, P.T., Kobayashi, N., Koubek, R., LaMarsh, J., Nagamachi, M., Naniwada, M., Salzman, H., Seppala, P., Schallock, B., Sheridan, T., Warschat, J. (1994). Integrating People, Organization, and Technology in Advanced Manufacturing: A Position Paper Based on the Joint View of Industrial Managers, Engineers, Consultants, and Researchers. The International Journal of Human Factors in Manufacturing, Vol. 4 (1), pp. 1–19

Kizilpinar, D. (2021). Data consistency in Microservices architecture. Medium. Available at: https://medium.com/garantibbva-teknoloji/data-consistency-in-microservices-architecture-5c67e0f65256 (Accessed 28 January 2022)

Kuhn, M., Franke, J. (2021). Data continuity and traceability in complex manufacturing systems: a graph-based modelling approach, International Journal of Computer Integrated Manufacturing. Vol. 34 (5), pp. 549–566

Machado, F., Duarte, N., Amaral, A. and Barros, T. (2021). Project Management Maturity Models for Construction Firms. Journal of Risk and Financial Management. Vol 14 (12), pp. 571–584

Schuh, G., Anderl, R., Gausemeier, J., ten Hompel, M. and Wahlster, W. (2017). Industrie 4.0 Maturity Index. Die digitale Transformation von Unternehmen gestalten. Munich: Acatech - Deutsche Akademie der Technikwissenschaften.

Steven, M. (2018). Industrie 4.0. Grundlagen – Teilbereiche – Perspektiven. Stuttgart: Kohlhammer

Wada, H., Fekete, A., Zhao, L., Lee, K., Liu, A. (2011). Data Consistency Properties and the trade-Offs in Commercial Cloud storages: the consumers' Perspective. Pacific Grove, California: CIDR Conference 2011, Asilomar, California

Weber, C., Königsberger, J., Kassner, L., Mitschang, B. (2017). M2DDM – A Maturity Model for Data-Driven Manufacturing. Procedia CIRP, 2017, Vol. 63, pp. 173–178

Wendler, R. (2012). The maturity of maturity model research: A systematic mapping study. Information software technology 54. Vol. 54 (12), pp. 1317–1339

Witsch, M. and Vogel-Heuser, B. (2012). Towards a Formal Specification Framework for Manufacturing Execution Systems. Transactions on Industrial Informatics 8. Vol. 8 (2), pp. 311–320