Speed, the double-edged sword of the Industry 4.0

Marion Toussaint [1,2], Sylvere Krima [3], Allison Barnard Feeney [4] and Herve Panetto [2]
[1]*Associate, NIST, 100 Bureau Drive, Gaithersburg, MD, 20899, USA*
[2] *Université de Lorraine, CNRS, CRAN, 54000 Nancy, France*
[3] *Georgetown University, Washington, DC, 20057, USA*
[4] *NIST, 100 Bureau Drive, Gaithersburg, MD, 20899, USA*

Keywords: industry 4.0, data exchange, data interoperability, data traceability

Abstract: The recent and ongoing digital transformation of the manufacturing world has led to numerous benefits, from higher quality products to increased productivity and reduced time to market. In this digital world, data has become a critical element in many essential decisions and processes within and across organizations. Data exchange is now a key process for the organizations' communication, collaboration, and efficiency. Industry 4.0/Industry of the Future adoption of modern communication technologies has made data available and shareable at a speed faster than we can consume or track it. This speed is a double edge sword and comes with key challenges, such as data interoperability and data traceability, which manufacturers need to understand in order to adopt the best mitigation strategies. This paper is a summarized introduction to these challenges, their origins, and what they mean to manufacturers.

1 INTRODUCTION

Over the centuries, technological advancement has changed the production methods that humans use. New techniques and production processes have radically changed people's working conditions and lifestyles.

The First Industrial Revolution marked the birth of mechanization through the use of water and steam power. The Second Industrial Revolution reflected the emergence of mass production possible through the discovery of electricity. The Third marked the emergence of automation in production processes through the introduction of electronics and information technology. Finally, the Fourth Industrial Revolution, also known as "Industry 4.0", was formed by the digital revolution that started during the Third Industrial Revolution based on cyber-physical systems (CPS). It is also characterized by the interconnectivity of the systems and access to real-time data.

The digital revolution in the world of manufacturing is fueled by advances in information and communication technologies. Paper-based 2D drawings and unstructured data sources (e.g., spreadsheets, text documents, email, …) have been replaced by structured digital data models containing various types of information (e.g., product design, manufacturing equipment, process data …). On the same principle, automated processes to collect and analyze data in real-time have succeeded the manual methods formerly used.

The digitalization of manufacturing and the adoption of IoT/CPS technology (e.g., smart sensors, smart actuators, machine learning) and cyber-physical systems have facilitated and resulted in the generation and acquisition of large volumes of heterogeneous data (Reinsel et al, 2018) (e.g., product models or telemetry data). Organizations produce, consume, and exchange massive volumes of data as part of their daily operations. Data now has the power to instantly turn into information, knowledge, and educated decisions, in an effort to boost performances (e.g., reducing cost or optimizing resources) (Rüßmann et al., 2016). For instance, tooling data can now be processed and analyzed by AI agents to optimize machine performance and energy efficiency in real-time. Digital data has become an essential player in many decision-making processes and a critical enabler to improving manufacturing competitiveness (Tao et al., 2018).

Industry 4.0 necessitates and enables fast access and exchange of that product data among a variety of applications and information systems - within and across organizations. A product creates and relies on a large amount of data during its lifecycle in response

to different processes (e.g., design, manufacturing, inspection) and business needs (e.g., technical, commercial, regulatory). Every organization involved in the product lifecycle relies on this data to perform its function. It represents the "fuel" behind the organization's contribution, efficiency, and value: organizations can create more value and drive faster innovation by exchanging data across them, facilitating collaboration.

Unfortunately, fast and reliable data exchange is also a complex operation that comes with multiple challenges (Panetto et al., 2019), each of which can have drastic consequences on organizations, their operations, their products, and their collaborators. In this paper, we define and discuss the risks associated with two major challenges, data interoperability and data traceability. In the next section we introduce the data interoperability issue and discuss why the traditional information standard development process is inadequate to support the Industry 4.0 fast-paced environment. We follow by discussing cyber threats, why manufacturing is a viable target, and how appropriate data traceability can help mitigate these risks in this complex environment. Finally, we conclude and discuss future directions.

2 DATA INTEROPERABILITY

Following this digital transformation of the industry and the modernization of the adopted communication technologies, data is now available from all, to all, and in a multitude of formats. Organizations can easily connect different software and physical systems, internally and within their network of collaborators, as long as these systems speak a common language.

Unfortunately, today's manufacturing organizations are characterized by complex environments consisting of domain-specific components such as systems, networks, or machines, clustered in heterogeneous groups. While the interaction of these components is crucial for manufacturing as it supports production processes, effective interoperability across all elements of the product lifecycle is a growing challenge (Panetto, 2007). The amount of data produced and consumed continues to increase due to this growing ecosystem (of machines, systems, and networks), but so does the number of data formats. These data are collected from distributed data sources and therefore do not necessarily share the same format. Data heterogeneity is an important factor in data exchange. The different components of an organization's environment must be able to unambiguously interpret, use, integrate, and compare the information exchanged.

These different systems need a common language to exchange and understand information. The use of neutral model-based data standards helps provide a common data format, and thus facilitates interoperability between all parties involved in an exchange. Standards are essential for properly integrating, exchanging, and interpreting data manufacturers rely on (Sapp et al., 2021). Standards define an agreed-upon language (data format, definitions, etc.) for data exchange between the different systems that consume, process, and produce data. The lack of standardization results in a multiplication of information formats that are not necessarily compatible with each other, making it difficult for stakeholders to communicate and exchange data.

Information standards are an important asset for organizations because they help facilitate business interaction and support interoperability between systems, people, and organizations. Information standardization also saves time and reduces costs by eliminating the need to have separate translators for each pair of systems that need to exchange data. The adoption and implementation of standards by organizations improves performance, competitiveness, and transparency given that standards promote the accessibility of information by all stakeholders. Information standards are powerful tools for innovation and productivity and are therefore key enablers to the evolution and digitalization of the manufacturing sector. Nowadays, standards support the full product lifecycle. Product definition data is represented in ISO 10303 (informally known as STEP) (ISO, 2020). Manufacturing planning systems can read in STEP data and generate manufacturing instructions in G-code (ISO, 2009) or ISO 10303-238 (STEP-NC) (ISO, 2007). MTConnect Agents (MTConnect Institute, n.d.) stream machine execution data that represents an as-manufactured product. Coordinate measurement system software can read in STEP product definition data and generate inspection plans and inspection results represented in the Quality Information Framework (QIF) (DMSC, 2016).

Despite this, information standards present a major challenge, which can impact their adoption and implementation by organizations: the complexity and current development process length of prominent standards are incompatible not only with the needs and pace of the industry but also with the lifespan of data.

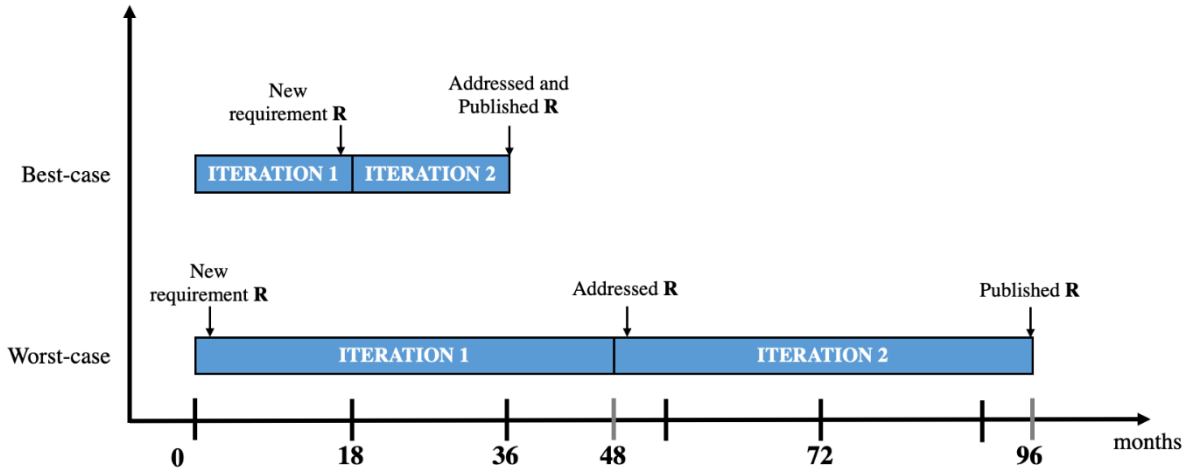

Figure 1: Requirement management in standards development process.

The information standards development process is complex. This process is generally long, irregular, and difficult to plan. Firstly, the waterfall methodology for project management is prominent, which implies that 1) the entire deliverable is only available (for review) at the end of the development iteration and 2) the requirements must be defined at the beginning of the project and do not change throughout the entire iteration. According to ISO itself, standard development iterations can last between 18 and 48 months (ISO, n.d.). This means that in a worst-case scenario, if a new requirement is identified after a new iteration just started, it will not be addressed for another 48 months, only once the iteration is complete, and will therefore be published up to 96 months later. Similarly, in the best-case scenario, a new requirement could be addressed and published in 18 months (see Figure 1). But in both cases, additional time must be given to software vendors to implement, test, and deliver updated software solutions.

Secondly, standards are developed by experts that are working for different organizations. The contribution and participation of these experts to the standards development process are entirely voluntary. The resources available depend on the experts' schedules and their organizations' needs, making the development process irregular and difficult to plan.

The duration and management of the standards development process are not aligned with the needs of the industry. Strong market competition results in shortened product life cycles and requirements that change often and faster than the pace of standards development. The standards development process is incompatible with the data lifespan. Industry 4.0 values speed and rapid innovation. Consequently, manufacturers need standards development organizations to accelerate and simplify the standards development process, so the resulting standards represent current industry needs and are eagerly adopted.

3 DATA TRACEABILITY

Manufacturing has become more automated, connected, and data-centric. Industry 4.0 is characterized by the networking of machines, systems, and products and the convergence of physical, digital, and virtual environments. This continuous networking and emergence of cyber-physical environments allow data to be more quickly accessible and facilitates the fast and timely exchange of data between the systems that require the information and the systems that have the information (inray Industriesoftware GmbH, 2018). These data exchanges are both intra- and inter-organizational, and can be characterized as high-speed, high-volume, high-frequency, and low latency exchanges. For instance, on the manufacturing floors, complex instructions and monitoring data are exchanged in real-time between the different manufacturing systems. Similarly, the integration with other technologies, such as artificial intelligence (AI), means that data is used to generate and share decisions at a pace and volume significantly greater than anything humans can manually validate or track.

This pace and volume of data exchange in the manufacturing world comes with significant challenges. The heavy reliance on data-driven decisions and the integration of new technologies have made organizations more vulnerable to cyber

threats, a major concern for companies regardless of their size and sector. The manufacturing sector generates large amounts of data and relies heavily on it, which makes this sector an ideal target for cyber-attacks. To no surprise, the manufacturing sector was particularly impacted by cybercrime in 2020 and 2021. According to the IBM Security's X-ForceThreat Intelligence Index 2022 report, manufacturing was the most attacked sector in 2021 (with 23,2% of all attacks), while it was ranked second in 2020 (with 17,7% of all attacks) and eighth in 2019 (with 8,1% of all attacks) (IBM Security 2022).

Generally, security threats are classified according to the governing principles of the CIA triad security model: confidentiality, integrity, and availability (Ham, 2021; Nweke, 2017). Data confidentiality requires that data remains secret or private, data integrity requires that the data is trustworthy and free from tampering, and finally, data availability requires that data is always accessible to authorized access when it is needed. Threats and vulnerability are assessed based on the type of risks1 associated with and the potential damage they can cause to an organization's assets, such as data, applications, and systems.

These risks cannot always be averted and are a significant challenge to identify and contain. The IBM's Cost of a Data Breach Report 2022 shows that in 2021, the mean time to identify (MTTI) a data breach was 212 days and 75 days to be contained (MTTC), for a total lifecycle of 287 days. This represents a slight increase over 2020, when the average time to identify and contain a data breach was 280 days (an average of 207 days to identify and an average of 73 days to contain) (IBM Security, 2021a).

One threat particularly relevant is data manipulation, an attack that focuses on subtly altering

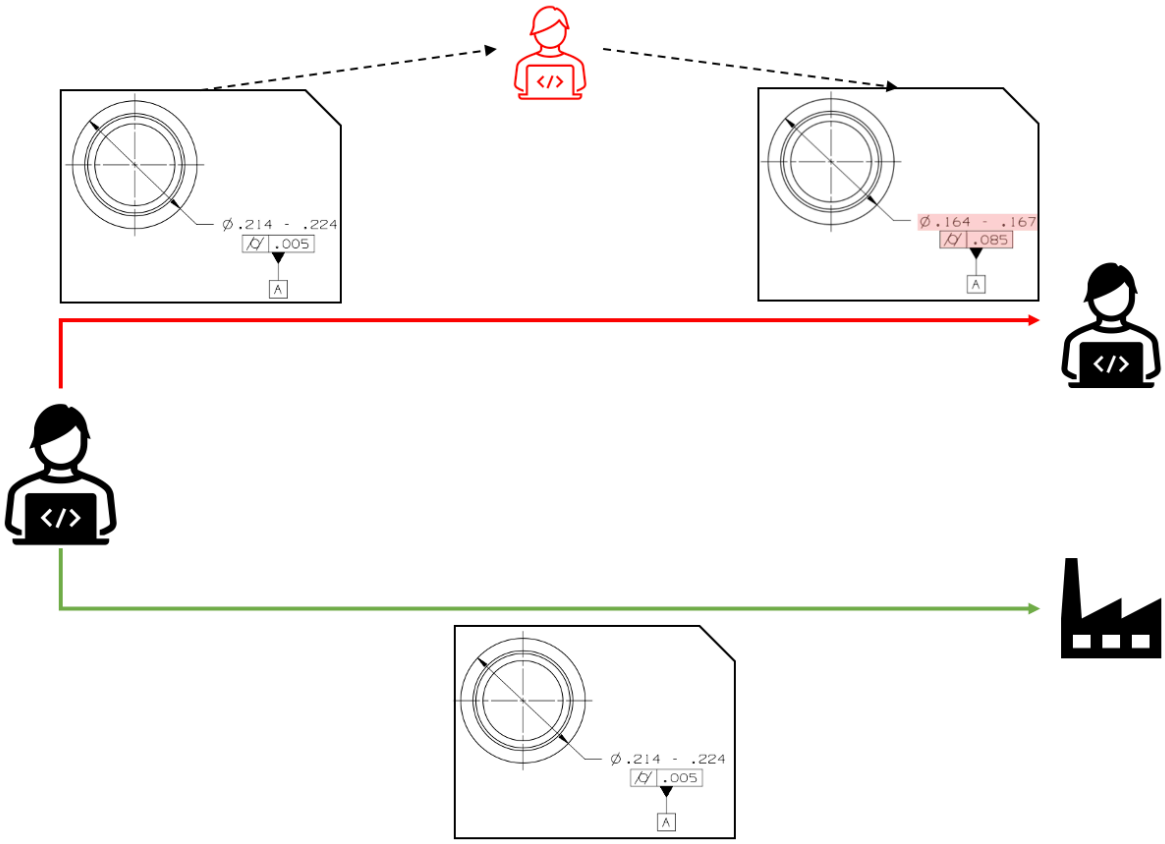

Figure 2: Example of data manipulation during a data exchange – the red flow indicates a malicious actor tampering PMI on a 2D drawing using a Man In The Middle (MITM)[1] attack.

[1]NIST Computer Security Resource Center glossary https://csrc.nist.gov/glossary/term/man_in_the_middle_attack

data (Wu et al., 2018) with the objective of manipulating data-driven decisions and relies on data exchange to propagate tampered data and decisions across an organization and its network. This tampering can result in corruption, modification, and/or destruction of the data, ultimately causing a loss of trust in the data (IBM Security, 2021b) and the decisions derived from it. Data manipulation can also potentially lead to different manufactured products.

When data is exchanged, it leaves the private and trusted system of the data owner to be sent to other systems. This process presents the critical risk that the data exchanged might have been tampered with by unauthorized parties (see Figure 2). It is therefore important to ensure the data remains accurate, authentic and trustworthy during the entire exchange process. Data integrity is the CIA triad aspect the most impacted by data exchanges, as integrity assists both the sender, who must ensure that data attributed to them is not tampered with, and the receiver, who needs "the guarantee that the message that is sent is the same as the message received and that the message is not altered in transit" (Agarwal and Agarwal, 2011).

Data integrity presents two main challenges: 1) validating the accuracy of the data and 2) tracking down inaccuracy. The former is commonly solved using digital signatures (Hedberg et al., 2016), while the latter is more complex and one that still needs to be addressed. The complexity (i.e., number of actors and steps involved), pace, and volume of data exchange that organizations are part of makes it impossible to manually account for and track down every single inaccuracy. Those same benefits and advantages that make manufacturers more competitive and innovative also make them more vulnerable to data integrity attacks.

4 CONCLUSION

The digital transformation of manufacturing has led to more connected, automated, and data-driven environments and processes. Data has become a key enabler to processes, exchanges, and decision-making. Manufacturing relies heavily on data and the exchange of this data between the different stakeholders, machines, and systems. By definition, data exchange refers to the process of sending and receiving data in a way in which the data content or meaning has not been altered during communication, in other words that the data received is an accurate representation of the data sent.

The digitalization of manufacturing has emphasized the importance of information management, data exchange, and the interoperability of the different actors in the manufacturing processes. The emergence of new technologies and networked data sources support new opportunities for organizational collaboration through high-speed and high-volume data exchange. In other words, this digital era helped improve the speed, volume, accuracy, and consistency of data exchange and innovations across and within organizations. But with great speed, came great challenges.

On one hand, faster innovation and collaboration are being hindered by the data interoperability challenges. Increased collaboration is associated with an increased number of heterogeneous systems that need to communicate with each other. While standards are a proven solution, their long and complex development process prevents them from keeping up with the fast-paced environment they need to support and provide interoperability for. Recent efforts (Sapp et al., 2021) promote a transition from predictive planning to adaptive project planning and the use of Agile methods to shorten the development iterations and increase the delivery velocity. These recommendations should drive manufacturers to favor standards that have adopted or are planning to adopt such methods.

On the other hand, data-driven decisions are exposed to the speed at which tampered data can propagate through organizations and corrupt these decisions. With the mean time to identify (MTTI) such a threat already close to 215 days (IBM Security, 2021a), the constant growth of data produced and exchanged is likely to push the MTTI upwards. While digital signatures have already proven their use in identifying such corruption, recent efforts (Krima et al., 2020; Ruland and Sassmannshausen, 2019; Cao et al., 2020) highlight the need for new formal data traceability methods and the use of data standards to automate the tracking of data exchange across large and complex networks of organizations and systems. Without such solutions, the mean time to contain (MTTC) tampered data and decisions will continue to increase with the quantity of data exchanges, perpetuating the current trend and continuing to put manufacturers at risk (IBM Security, 2021a). These efforts should drive manufacturers to favor standards over proprietary formats for their data exchange in order to enable maximum data traceability.

To conclude, the speed at which data exchanges can now be set up and performed has highlighted the need to reduce the time of 1) development and implementation of data interoperability solutions

(Sapp et al., 2021) and 2) data traceability operations (Krima et al., 2020; Ruland and Sassmannshausen, 2019; Cao et al., 2020) in response to cyber-attacks that manufacturers are victims of, which are two challenges we will focus on.

REFERENCES

Agarwal, A., & Agarwal, A. (2011). The Security Risks Associated with Cloud Computing. *International Journal of Computer Applications in Engineering Sciences*, *1*.

Cao, Y., Jia, F., & Manogaran, G. (2020). Efficient traceability systems of steel products using blockchain-based industrial internet of things. *IEEE Transactions on Industrial Informatics*, *16*(9), 6004–6012. https://doi.org/10.1109/tii.2019.2942211

DMSC. (2016). *Digital metrology standards consortium QIF & DMIS standards*. DMSC. https://qifstandards.org/.

Ham, J. V. (2021). Toward a better understanding of "cybersecurity." *Digital Threats: Research and Practice*, *2(3*), [18], pp 1-3. https://doi.org/10.1145/3442445

Hedberg, T. D., Jr., Krima, S., & Camelio, J. A. (2016). Embedding X.509 Digital certificates in three-dimensional models for authentication, authorization, and traceability of product data. *Journal of Computing and Information Science in Engineering*, *17*(1). https://doi.org/10.1115/1.4034131

IBM Security (2021a). IBM: Cost of a data breach report. (2021). *Computer Fraud & Security*, *2021*(8), 4–4. https://doi.org/10.1016/s1361-3723(21)00082-8

IBM Security (2021b) IBM: 2021 X-Force Threat Intelligence index. (2021). *Network Security*, *2021*(3), 4–4. https://doi.org/10.1016/s1353-4858(21)00026-x

IBM Security (2022). IBM: 2022 X-Force Threat Intelligence index. *Computer Fraud & Security*, *2022*(3). https://doi.org/10.12968/s1361-3723(22)70561-1

inray Industriesoftware GmbH (2018). Industry 4.0 depends on data transfer. timely and target-oriented. *Medium*. Accessed: 2022-05-05, from https://medium.com/@inray.industriesoftware/industry-4-0-depends-on-data-transfer-timely-and-target-oriented-e725eeffa3e3

ISO. (n.d.). *Stages and resources for standards development*. ISO. Accessed: 2022-05-10, from https://www.iso.org/stages-and-resources-for-standards-development.html

ISO. (2007). *ISO 10303-238:2007*. ISO. Accessed: 2022-05-10, from https://www.iso.org/standard/38036.html.

ISO. (2009). *ISO 6983-1:2009*. ISO. Accessed: 2022-05-10, from https://www.iso.org/standard/34608.html

ISO. (2020). *ISO 10303-242:2020*. ISO. Accessed: 2022-05-10, from https://www.iso.org/standard/66654.html

Krima, S., Toussaint, M., & Feeney, A. B. (2020). Toward model-based integration specifications to secure the extended enterprise. *Smart and Sustainable Manufacturing Systems*, *4*(1). https://doi.org/10.1520/ssms20200022

MTConnect Institute. (n.d.). *MTConnect*. MTConnect. Accessed: 2022-05-10, from https://www.mtconnect.org/.

Nweke, Livinus. (2017). Using the CIA and AAA Models to explain Cybersecurity Activities. *PM World Journal, VI(XII).*

Panetto, H. (2007). Towards a Classification Framework for Interoperability of Enterprise Applications. *International Journal of Computer Integrated Manufacturing*, Taylor & Francis, 2007, 20 (8), pp.727-740. https://doi.org/10.1080/09511920600996419

Panetto, H., Iung, B., Ivanov, D., Weichhart, G., & Wang, X. (2019). Challenges for the cyber-physical manufacturing enterprises of the future. *Annual Reviews in Control*, *47*, 200–213. https://doi.org/10.1016/j.arcontrol.2019.02.002

Reinsel, D., Gantz, J., & Rydning, J. (2018). *The digitization of the world from edge to core*. IDC. https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf.

Ruland, C., & Sassmannshausen, J. (2019). System-wide traceability of commands and data exchange in smart grids. *2019 International Conference on Smart Energy Systems and Technologies (SEST)*. http://dx.doi.org/10.1109/sest.2019.8849108

Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2016). *Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries*.

Sapp, B., Harvey, M., Toussaint, M., Krima, S., Barnard Feeney, A., & Panetto, H. (2021). *Agile for model-based-standards development*. National Institute of Standards and Technology. http://dx.doi.org/10.6028/nist.ams.100-40

Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. *Journal of Manufacturing Systems*, *48*. https://doi.org/https://doi.org/10.1016/j.jmsy.2018.01.006.