Manufacturing Data Security, Trustiness and Traceability

Javier Pérez-Soler[1], Pau Garrigues[1], Imanol Fuidio[2], Stefan Wellsandt[3], Gennady Laventman[4], Mohammad Amin Khodamoradi[5], Santiago Gálvez-Settier[1] and Juan-Carlos Pérez-Cortés[1]

[1] *Instituto Tecnológico de Informática, Camino de Vera s/n, Valencia, 46022, Spain*
[2] *Ikerlan Technology Research Center, Pº J.M. Arizmendiarrieta 2, Arrasate, 20500, Spain*
[3] *Bremer Institut für Produktion und Logistik GmbH at the University of Bremen, Hochschulring 20, Bremen, 28359, Germany*
[4] *IBM Israel-Science and Technology Ltd. 94, Derech Em-Hamoshavot, Petach Tikva, 49527, Israel*
[5] *Instituto de Desenvolvimento de Novas Tecnologias, Campus da FCT NOVA, Caparica, 2829-516, Portugal*

Abstract

European manufacturing companies collect a large amount of data during their manufacturing processes thanks to the increasing use of sensors, actuators and instruments. This amount of data is very valuable for improving manufacturing quality, with a view to the Zero Defects objective. In order to take full advantage of the vast amount of accumulated data, quality must be targeted, and its Security, Trustiness and Traceability must be preserved. Data quality is influenced by several factors: including human errors, communication problems or inaccuracies, so any system used must be sufficiently reliable.

This paper describes the technical approach to develop solutions to control manufacturing Data Security, Trustiness and Traceability with a zero-defect approach and provided by the i4Q project. The approach is exemplified with the industrial production line of injected plastic spare parts for the automotive sector.

Keywords

Product quality, machine learning, artificial intelligence, zero-defect manufacturing, non-destructive inspection, data security, trusted networks, data traceability, data repository

1. Introduction

The i4Q Project [1] presents a complete solution consisting of sustainable IoT-based Reliable Industrial Data Services able to manage the huge amount of industrial data coming from cost-effective, smart, and small size interconnected factory devices for supporting manufacturing online monitoring and control [2].

This work describes the necessary strategies, methods, and key technologies to ensure data quality, which is influenced by many factors, including human error, communication issues or inaccuracies. Monitoring systems are expected to be reliable enough for decision-making, but sensors are susceptible to provide unreliable information.

Trusted networks allow ensuring the reliability, integrity and privacy of the data exchanged. The heterogeneity in technologies leads to a complex ecosystem of data models that is difficult to contextualize, leading to misinterpretations and incomplete data analysis. Setting the guidelines for

Proceedings of the Workshop of I-ESA'22, March 23–24, 2022, Valencia, Spain
EMAIL: javierperez@iti.es (J. Pérez-Soler); pgarrigues@iti.es (P. Garrigues); ifuidio@ikerlan.es (I. Fuidio), wel@biba.uni-bremen.de (S. Wellsandt), gennady@il.ibm.com (G. Laventman), a.khodamoradi@uninova.pt (M.A. Khodamoradi), sgalvez@iti.es (S. Gálvez-Settier), jcperez@iti.es (J.C. Pérez-Cortés)
ORCID: 0000-0002-4195-9491 (J. Pérez-Soler); 0000-0003-3408-3249 (P. Garrigues); 0000-0003-4761-9532 (I. Fuidio), 0000-0002-0797-0718 (S. Wellsandt), 0000-0002-2700-5384 (M.A. Khodamoradi), 0000-0003-3479-8367 (S. Gálvez-Settier), 0000-0001-6506-090X (J.C. Pérez-Cortés)

information models, needed metadata for traceability and interoperability is crucial for ensuring data quality.

Finally, the increasing amount of collected data requires flexible data storage and data analysis infrastructure able to process information without affecting data quality. Distributed systems can help improve the performance of data storage and data processing, opening the door to the use of Distributed Ledger Technologies.

The proposed approach is exemplified with the use case named "Automatic Advanced Inspection of Automotive Plastic Parts", and industrial production line of injected plastic spare parts for the automotive sector.

2. Technical approach

The i4Q Project aims at providing methodologies, tools and infrastructure to ensure the necessary data quality to enable operational intelligence and improve data analysis results effectiveness.

In order to achieve this goal, the work is divided into five work areas with different purposes:

- To provide a common information model that ensures the quality and traceability of collected data, including metadata regarding accuracy and reliability of the data measured. This helps to consolidate the information and implements best practices.
- To provide security tools and mechanisms, such as Public Key Infrastructure and project level certificates, as well as best practices and recommendations, to ensure security and privacy on every step of the data cycle.
- To provide a secure and safe oriented communication infrastructure that guarantees integrity, reliability and privacy of manufacturing data gathered.
- To analyze and assess the suitability of Distributed Ledger Technologies for ensuring security, trustiness and immutability of data in the i4Q framework.
- To provide the distributed storage architecture and infrastructure capable of handling an exponential increase in data without affecting its quality.

2.1. Manufacturing data quality strategy

This area of work focuses on a strategy to manage data quality. Its first part focuses on a guideline outlining key aspects of data and information quality management in manufacturing. It describes production systems in a way that data and information activities become more visible. Besides, it clarifies key stakeholders and how they can manage information quality issues, for instance, by training and technical measures integrated in information systems. The second part of this work area develops an Open-Source software tool called QualiExplore to raise awareness for information quality in manufacturing.

QualiExplore grounds on a knowledge base containing information quality factors. These factors describe phenomena that influence data characteristics and organizations should manage them to minimize information quality problems. QualiExplore organizes data quality factors using a data life cycle model and acknowledged quality characteristics, such as accuracy, precision, timeliness, and completeness. An optional conversational user interface will allow users to interact with QualiExplore intuitively via natural language. For instance, users in the exemplified use case will be able to describe their work and interest in information quality in a plastic spare parts production line and receive the matching factors they should manage. i4Q's complementary quality management guidelines will contain approaches and techniques the users could use to manage the identified factors.

The manufacturing data quality strategy focuses on two main data quality standards for Industry [4]: ISO 8000 [5] for data quality in the production line and automation systems and ISO/IEC 5259 [6] for data quality for analytics and machine learning (ML). i4Q uses these standards because they are widely used in organizations and also because ISO is an acknowledged standardization body.

The ISO 8000 standard series contains several documents introducing vocabulary, concepts, and processes concerning data quality. One important document outlines a related management

framework that introduces roles and responsibilities and elaborates some principles to exchange data characteristics like defining specific syntax, encoding data semantically, and conformance to data specification. Besides, the ISO 8000 series gives some guidelines to measure data completeness and accuracy.

The ISO/IEC 5259 standard package focuses on providing qualitative data for analytics and ML. This standard overviews data quality measures and provides terminology, examples and details for frameworks to create data quality processes while providing management guidelines and measurement indicators to monitor the state of such data quality processes.

2.2. Manufacturing data trustiness and traceability

This area of work provides easier, trustable and traceable access to data coming from many different sources. For this, it is necessary to analyze and model incoming data, in order to standardize the information models and metadata required.

The solution provides tools to ensure data trustiness and full traceability. The goal is to enhance the level of trust by employing a blockchain based data service, to support data traceability. This enhances trust and acceptability by providing security and trust in the data that flows directly to the blockchain, serving as a single point of truth, preserving provenance and supporting non-repudiation. Information stored on the blockchain cannot be changed or erased and is proved to be authentic. Blockchain-based services ensure that the information is not to be tampered with. Moreover, differential visibility scope of information shall be supported for cases in which a subset of the participating entities needs to share some information. The blockchain capabilities shall be exposed to other components and microservice applications via REST interfaces, with possibility to use several programmable SDKs.

The underlying blockchain solution shall use Orion Blockchain DB [7], an open-source project that provides a Key-Value (KV) datastore functionality with the addition of rich provenance and blockchain capabilities. Orion provides a multiparty approval mechanism between required participants using its multi-sign transactions and use its provenance and blockchain capabilities to govern the actions that take place upon the arrival of new data. Using its Read-Write Key Level Access Control mechanism, different visibility scope over existing data and its history is defined. The REST interface [8] provides API similar to regular KV databases, thus providing an easy way to invoke transactions to manipulate data in KV store, and each one of these manipulations automatically stored to the blockchain ledger. Same REST interface provides support for various rich queries to KV data, provenance data and blockchain ledger data, including rich range queries, history queries and non-repudiation and data integrity proofs.

Orion uses x.509 certificate [9] based PKI mechanism for both user and server authentication.

This area of work enhances the level of trust in the exemplified use case by employing a blockchain based data service, to support data traceability in the data that flows directly to the blockchain, thus serving as a single point of truth, preserving provenance and supporting non-repudiation.

2.3. Manufacturing data trusted communication and distribution

The trusted networks area of work implements reliable data collection, providing connectivity to industrial data sources through Trusted Networks able to assess and ensure precision, accuracy, and reliability. Technologies such as Time Sensitive Networks (TSN) for wired communications, and wireless access networks (Industrial Wireless Sensor Network (IWSN), Low Power Wide Area Networks (LPWAN), ad-hoc connections, etc.), are integrated or merged with other solutions such as Software Defined Network (SDN), Network Function Virtualisation, or Network Slicing in order to improve reliability of the communication infrastructure and therefore the integrity and reliability of data collected.

This trusted network infrastructure, in which the TSN and the low power IWSN resources are orchestrated with a centralized SDN controller, will be validated ensuring that the expected Quality of Service (QoS) is achieved.

This area of work includes the development of software-defined wireless industrial interfaces for data communication, paying special attention to requirements such as predictability and determinism, high reliability and trustability, low consumption and efficient and fast network deployment [10], while reducing the installation cost of new-wired infrastructure. Time-sensitive transmission of data over deterministic Ethernet networks is also required for applications that require very low transmission latency and high availability and can use the floor plant wired network infrastructure.

The wireless solutions based on WSN shall be used in the exemplified use case, and LPWAN for long range requirements, are envisaged to merge and implement state of the art techniques to enhance the QoS and robustness in these types of networks, with the introduction of SDN and adaptative mechanisms.

2.4. Manufacturing data security

Collected data can be delivered by diverse types of devices and be transmitted through and processed by a significant number of layers and technologies. The type of devices in an Industrial Automation and Control System (IACS) are mainly DCS (Distributed Control Systems), RTU (Remote Terminal Units), PLC (Programable Logic Controler) and SCADA (Supervisory Control and Data Adquisition). The operation of these devices translates into the necessity of recommendations and guidelines to enable multilayer cyber security features in Industrial Internet of Things, as well as tools to implement these recommendations, enabling Industrial Internet of Things devices to interact with the whole system securely in all stages of a manufacturing scenario.

This area of work comprises the analysis of the architecture and methodology to provision signed certificates with Hardware Security Module and trusted material to devices with a Trusted Platform Module or Secure Element using asymmetric encryption architecture. Furthermore, the solution targets OPC-UA certificate management mechanism [11] in order to provide a real-world example where is envisioned the union between the state-of-the-art security technology and a protocol specification.

The security mechanisms may differ between communication solutions for Distributed Ledger Technology tools proposed, which means applying security by design during development, adjusting security, and safety policies at different levels to ensure the trustability and privacy of data.

In the exemplified use case, this is a piece of software that distributes trust using x.509 certificates[9] and asymmetric cryptography. Once the trust is distributed, every module has a digital identity that uses to provide security between different endpoints using asymmetric cryptography considering different trust policies, adjusting security and safety policies at different levels. In the case of OPC-UA, the Global Discovery Server (GDS) provides role management policies which can be integrated with existing user and role management systems. Roles have access permission for nodes within the OPC UA Information Model.

The solution enriches the current state of the art with tools to provide seamless management of security by design components in an IACS environment with Hardware Security Modules (HSM).

2.5. Manufacturing data storage and use

This data repository solution is devoted to design and implement a distributed storage system considering those aspects unique to the Industry 4.0 paradigm [12]. One of the main aspects to consider is the high degree of digitization expected in companies, resulting in most manufacturing devices acting as sensors or actuators and generating vast amounts of data.

The infrastructure is, then, able to absorb large volumes of data coming into the system at high speeds. Similarly, it is as elastic as possible to adapt the computing resources, it requires to the existing demand and be ready to use additional resources, either local to the factory or from remote systems like public or private clouds if needed, although bearing in mind possible data privacy restrictions. In addition to the storage capabilities, the system must provide easy ways to access this data so other components and microservice applications can easily consume and use it to improve the efficiency of the system.

The data repository area of work includes a number of off-the-shelf tools. Currently all of them are open-source. Some of them are database managers (like MongoDB, MySQL, etc.). Others offer other types of storage (such as MinIO, which offers file and blob storage). The database managers cover a variety of structures and formats, including standard relational SQL-based data, JSON documents and other structured formats (like key-value pairs, graphs, etc.). Each tool offers its own indexing mechanisms.

In the scope of the exemplified use case, the objective of this area of work is to create a suitable storage system for the collected data of plastic spare parts for the automotive sector. This system or repository oversees receiving, storing, and serving the data in a proper way to the other components in the architecture. It provides the proper tools for administrators to consult and transform the information contained inside it, as well as the ways for data scientist to use this data in their experimentation. These tools are provided through a suitable user interface. The solution also oversees data protection, serving as a secure system for the information by means of encryption, both in flight and at rest.

The result is an efficient repository, ready to provide its service to the rest of the components present in the system.

3. Conclusions

This article has introduced the i4Q project's areas of work for data security, trustiness and traceability for an industrial inspection approach as conceived in the exemplified use case "Automatic Advanced Inspection of Automotive Plastic Parts".

Inspecting and predicting techniques for control of product quality are based on ML models analyzing secured, trustable and traceable data collected from production in-field sensors and other production related information.

These outcomes are available in the scope of i4Q to allow the design and construction of zero-defect production lines to maximize product quality.

The other i4Q use cases cover five more industrial scenarios, namely white goods, wood equipment, metal machining, ceramics pressing, and metal equipment. Each case makes use of the here described functionalities for Data Security, Trustiness and Traceability, in order to ensure prediction, inspection, and supervision of a given production line.

4. Acknowledgements

The research leading to these results received funding from the European Union H2020 Program under grant agreement No. 958205 "Industrial Data Services for Quality Control in Smart Manufacturing (i4Q)"

5. References

[1] i4Q Project Website, 2021. URL: https://www.i4q-project.eu/

[2] R. Poler, A. Karakostas, S. Vrochidis, A. Marguglio, S. Galvez-Settier, P. Figueiras, A. Gomez-Gonzalez, B. Mandler, S. Wellsandt, R. Trevino, S. Bassoumi, C. Agostinho, An IoT-based Reliable Industrial Data Services for Manufacturing Quality Control, in: 2021 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), IEEE, Cardiff, 2021, pp. 1–8. doi: 10.1109/ICE/ITMC52061.2021.9570203.

[3] A. Karakostas, R. Poler, F. Fraile, S. Vrochidis, Industrial data services for quality control in smart manufacturing–the i4q framework, in: 2021 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4. 0&IoT), IEEE, Rome, 2021, pp. 454-457. doi: 10.1109/MetroInd4.0IoT51437.2021.9488490.

[4] R. Perez-Castillo, A. G. Carretero, M. Rodriguez, I. Caballero, M. Piattini, A. Mate, S. Kim, D. Lee, Data Quality Best Practices in IoT Environments, in: 11th International Conference on the

Quality of Information and Communications Technology (QUATIC), IEEE, Coimbra, 2018, pp. 272-275. doi: 10.1109/QUATIC.2018.00048.
[5] ISO 8000. URL: https://www.iso.org/standard/50798.html
[6] ISO/IEC 5259. URL: https://www.iso.org/standard/81093.html
[7] Orion. URL: https://github.com/hyperledger-labs/orion-server
[8] Orion REST example. URL: http://labs.hyperledger.org/orion-server/docs/getting-started/transactions/curl/datatx
[9] x.509 certificate. URL: https://en.wikipedia.org/wiki/X.509
[10] J. Vera-Pérez, J. Silvestre-Blanes, V. Sempere-Payá, D. Cuesta-Frau, Multihop Latency Model for Industrial Wireless Sensor Networks Based on Interfering Nodes, Applied Sciences 11 (2021) 8790. doi: 10.3390/app11198790
[11] OPC Foundation, OPC 10000-2 - Part 2: Security Model. URL: https://reference.opcfoundation.org/v104/Core/docs/Part2/.
[12] Industries 4.0 paradigm. URL: https://en.wikipedia.org/wiki/Fourth_Industrial_Revolution