A REVIEW ON THE DEVELOPMENT OF DATASPACE CONNECTORS USING MICROSERVICES CROSS-COMPANY SECURED DATA EXCHANGE

Sze-Kai Gan[1*], Thein-Lai Wong[1], Ching-Pang Goh[1], Wah-Pheng Lee[2] and Yee-Mei Lim[3]

[1] Faculty of Computing and Information Technology,
[2] Centre for Postgraduate Studies and Research, Tunku Abdul Rahman University College,
Kampus Utama, Jalan Genting Kelang, 53300, Wilayah Persekutuan Kuala Lumpur, Malaysia
[3] Tech Innovation and Production, GMCM Sdn. Bhd., Kuala Lumpur, Malaysia
*Corresponding author: gansk-wr21@student.tarc.edu.my

ABSTRACT

Industry Revolutions 4.0 (IR4.0) involves digital transformation of businesses to deliver quality products and services supported by real time decision making. It does not only impact the manufacturing industry but involves everyone in any value chains, from producers to consumers in all industries. With the integration of Internet of Things (IoT) devices and business applications, data generated provide valuable information for business insights, hence referred as assets. Industrial Data Space Reference Architecture Model (IDS-RAM) aims to achieve a reliable exchange of data between organizations and platforms developed by different vendors, hence enabling seamless value chain integration. Data flow between organizations are usually crucial, high load and real time. Therefore, the data exchange architecture that involves the dataspace connectors that reside in every organization, must enable uninterrupted quality service, real time streaming and secured usage control of data with high volume, high velocity and high availability. To handle the data communication, whether they are stored in premises, on the cloud, or on edge, with different computing capacity and capabilities, all systems need a flexible, light and reliable connector architecture that can scale and optimize resources effectively based on demand. This paper reviews the existing researches that are related to the development of dataspace connectors in the past 10 years.

Keywords: High Availability, Industry 4.0, Reactive, Microservice, Connector

1.0 INTRODUCTION

The Industrial Revolution (IR) began in the 18th century has been changing the working conditions and lifestyles of humans tremendously. From steam-powered factories to mass production and eventually computer technology, the revolution increased human productivity while reducing the cost significantly. The emergence of IR 4.0 has given rise to the new wave of fundamental change in industry revolution(Ahuett-Garza and Kurfess, 2018). IR 4.0 represents the result of automation combined with information technology and operation technology (Lu, 2017), such as Industry Internet of Things (IIoT) and cloud computing to improve the accessibility to real-time data. IIoT uses the power of smart sensors, actuators, and real-time data analytics to improve the manufacturing process. The smart sensors installed at different stations are interconnected and controlled by computer software to gather huge amounts of data for real-time analysis. This allows the company to monitor, identify and amend the inefficiencies and problems in manufacturing processes without any delay. IIoT has been widely used in the field of manufacturing, warehouse, smart utilities, smart city, etc., which required full automation and monitoring.

Reference Architecture Model Industry 4.0 (RAMI 4.0)[1](Federal Ministry for Economic Affairs and Energy, 2019) was published in 2015 by the German Electrical and Electronic Manufacturers' Association (ZVEI), to support the Industry 4.0 efforts. RAMI 4.0 emphasizes industrial production as the primary area of application as shown in Figure 1. According to DIN SPEC 91345 (2016), the Information Layer describes the data that is used, generated or modified by the technical functionality. This includes a runtime environment for pre-processing of events, consistent integration of different data, formal description of models and rules, etc. The Communication Layer describes Industry 4.0-compliant access to the information and functions of the connected assets. In other words, it manages the exchange of data between company i.e., which data is used, where it is used and when it is distributed.

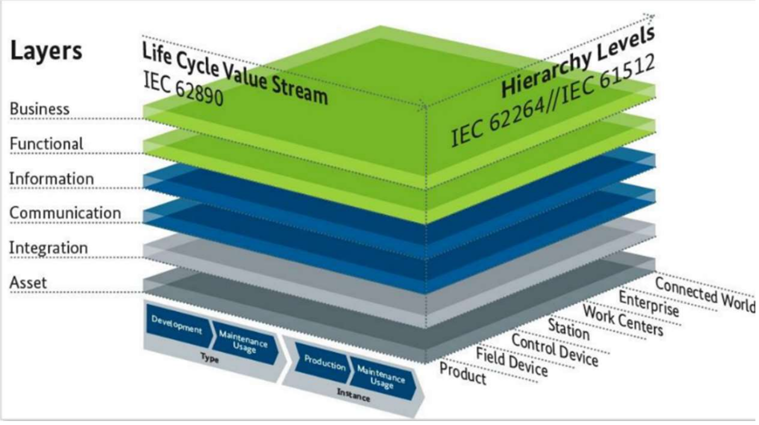

Figure 1. Reference Architecture Model for Industry 4.0 (Federal Ministry for Economic Affairs and Energy, 2019).

In 2017[2], Industrial Data Space Reference Architecture Model (IDS-RAM) enables a reliable exchange of data between company with common rules to be implemented in the Communication Layer of RAMI 4.0. Data sovereignty is, thus, a central aspect for IDS. Besides specifying the rules to ensure the data sovereignty within data ecosystems, the IDS-RAM, published by International Data Spaces Association (IDSA) (Otto *et al.*, 2019), consists of five layers: business layer, functional layer, process layer, information layer and system layer. The standards materialize in the IDS-RAM and DIN SPEC 27070:2020-03 (2020) define methods for secure data exchange between the various Industrial Data Space connectors. Figure 2 shows each connector is able to communicate with every other connector or component in the ecosystem of the Industrial Data Space, to allow any organization to communicate to the outside world real-time.

---

[1] https://www.i-scoop.eu/industry-4-0/#The_Reference_Architectural_Model_Industrie_40_RAMI_40

[2] https://www.i-scoop.eu/industry-4-0/industrial-data-space/

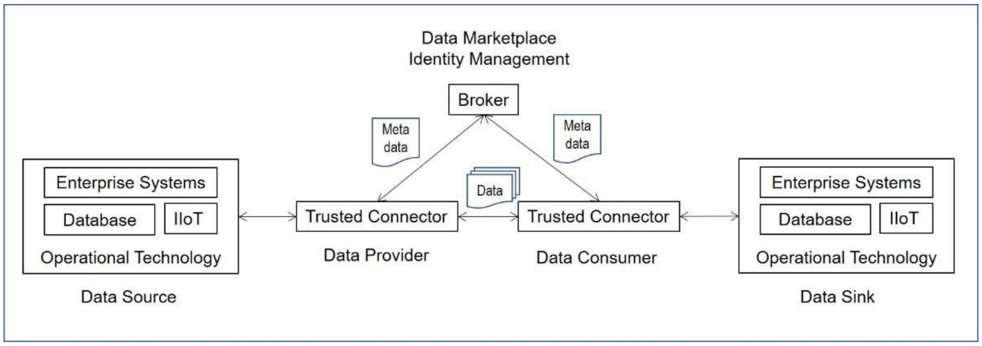

Figure 2. Data Exchange Architecture in IDS environment

This paper examines the current research related to dataspace connector development and the issues of data security, high availability, low latency and scalable resources of cross-company data exchange between two trusted connected are not addressed in IDS-RAM. It is particularly crucial for the development of digital twin for timely distributed business operations in the Industry 4.0. Hence, this paper will review and propose the reactive microservices as an architecture model to overcome the challenges of using Data Exchange Architecture in the cross-company secured data exchange in the manufacturing environment.

2.0 LITERATURE REVIEW

Detailed research related to trusted connector development over the recent decade are studied in this section. The research findings are categorized into three general areas which are service-oriented architecture (SOA) and microservices, orchestration and security.

2.1 Service-Oriented Architecture (SOA) and Microservices

Microservices architecture (Fowler and Lewis, 2014) is a variant of the service-oriented architecture (SOA) (Xiao, Wijegunaratne and Qiang, 2017), which is built from a collection of services. Each individual service contains very minimal functions and runs on its own process autonomously. In contrast to the monolithic architecture which comprises all modules into a single application, the collection of the microservices is loosely coupled; hence it provides the flexibility to modify and upgrade the services without interrupting other modules. In addition, microservices can achieve autonomy, isolation, and resilience with the help of external services, for instance load balancer. A recent study done by Santana, Alencar and Prazeres (2018) showed microservices architecture started being used in IoT applications to improve the performances of large-scale systems. Petrasch (2017) studied the microservices architecture with Enterprise Integration Patterns (EIP)(G. Hohpe, 2003) for inter-service communication. They demonstrated how domain driven design, model-driven and pattern-based approach for composing microservices to reduce the complexity of microservices architecture.

In 2019, Bigheti, Fernandes and Godoy proposed a Microservice-Oriented Architecture (MOA) for Control as a Service (CAAS). The authors simplified the MOA deployment, and provided a flexible, interoperable, and distributed architecture. It is a high-level design of the microservices as enhancement for ISA95 architecture, but they did not further address the drawbacks on this architecture and their solution. Dinh-Tuan, Beierle and Garzon (2019) implemented the microservices architecture in industrial manufacturing environments. They designed a decentralized architecture with the microservices based for industrial data analytic by developing the prototype, analyzed, and evaluated to provide further practical insights.

Reactive microservice further enhance the standard microservice to have more isolation and autonomy with its consensus protocol and asynchronous non-blocking messaging architecture[3]. Later in 2021, Santana *et al.* improved the limitation of high availability by utilizing the reactive and asynchronous programming paradigm in his research. The result showed that the reactive system improved only the availability of the services of distributed IIOT networks in agriculture 4.0. However, there is still lack of research using reactive microservices to improve the requirements of timely data exchange in the distributed manufacturing environment

To enhance and manage the complex architecture defined in IDS-RAM, each of the module shall work independently and autonomy with microservice architecture. Modules could be delegated to specialized team to develop and deliver.

2.2 Service Orchestration

Service orchestration is the arrangement, configuration and management of the system to ensure the system operates within its ecosystem with the expected behavior.[4] Girbea *et al.* (2014) designed a novel architecture in industrial automation based on the SOA. The new architecture was able to compute optimal production plans and executed the plans automatically. Besides, the development and the maintenance were also flexible and reusable. This enhanced the seamless transition from the current practice. Their paper focused on data monitoring in manufacturing processes without considering the data exchange between organizations. Theorin *et al.* (2017) developed an event-driven architecture for IR4.0, which was Line Information System Architecture (LISA). This architecture enabled the data exchange between organizations in more efficient and flexible ways. It also allowed data integration and utilization by using SOA and Enterprise Services Bus (ESB).

In the Manufacturing as a Service (Maas) domain, Landolfi *et al.* (2019) developed an ecosystem that acted as a virtual marketplace. This virtual marketplace promoted the virtual and physical assets so that the production demand was able to be achieved optimality. For the next evolution of the architecture, the authors suggested using FIWARE Generic Enablers (GE)(Alonso et al., 2018) and an extensive testing of their functionalities. FIWARE[5] offers an open standard platform and a set of standard APIs to support IoT and smart applications development in various domains. A FIREWARE GE is a set of general-purpose platform functions available through APIs.

Trunzer et al. (2019) discussed some system RAMI 4.0, American Industrial Internet Reference Architecture (IIRA) (Lin et al., 2017), Line Information System Architecture (Theorin *et al.*, 2017), SORRADES architecture (Karnouskos, Bangemann and Diedrich, 2009), etc. to address several issues in cross-enterprise data sharing, service orchestration and

---

[3] https://dzone.com/articles/diving-into-reactive-microservices

[4] https://cloudify.co/blog/why-service-orchestration-matters/

[5] https://www.fiware.org

real-time capabilities. These generic architectures bridged up the gap between reference and use-case specific architectures.

Microservices or SOA are highly flexible and dynamic and require more complex system architecture. Studies showed the needs of service orchestration to ensure the interconnectivity of the services and functions are performed as good as monolith application. Some modern orchestration tools such as Kubernetes, Chef, Ansible could help to manage those complex design.

2.3 Trusted Data Security

Security has always been prioritized in data exchange as the data is sensitive and it is the most important asset in an organization. A reliable security policy and algorithm is able to safeguard the data integrity, authority, and privacy especially in the manufacturing industry where operational and business data is proprietary and confidential. A study from Li, Xuan and Wen (2011) on the general architecture of trusted security system based on IoT materialized the trusted user module, trusted perception module, trusted terminal module, trusted network and trusted agent module to eliminate the various security threats in the application of IoT. Later in 2017, Teslya and Ryabchikov (2017) implemented the blockchain in IoT to enhance the security level of the IoT system. The authors designed an architecture that combined Smart-M3 information sharing platform, blockchain platform and smart contracts to process and store the information between smart space components. Nevertheless, this architecture has several limitations which are (1) lack of mechanisms to establish authorship, durability and unchangeability of information, (2) control over the exchange of resources in production, and (3) an integrated mechanism for reaching consensus among participants. (Brost *et al.*, 2018) developed a comprehensive security architecture which included Service Manager, Connection Manager, Routing Manager, Audit Logging, Dataflow control, Access control and Usage Control, Security Management as services bundled in OSGI framework (*OSGi Alliance*, 2018). The major drawback of the OSGI framework is it required all the services included in the single OSGI container for all the services to interact and run as a monolith application. Munoz-Arcentales *et al. (2019*, 2020) enabled the access and usage control in data-sharing ecosystems among multiple organizations in the food industry using the FIWARE European opensource platform. The Policy Enforcement Point (PEP) intercepted the incoming request from the Data Consumer to the Data Provider's infrastructure to enforce access control on resources.

In the Healthcare cyber physical system, Xu et al. (2020) proposed a privacy-preserving data integrity verification model by using lightweight streaming authenticated data structures. The authors discussed the design idea, architecture, formal definition, security definition and communication protocols in the model. In the Streaming Authenticated Data Structures (SADS), author provide six probability polynomial algorithms in to the model and leveraging the computing of the fully homomorphic encryption (FHMT) which shifts almost all the computation tasks to the server and little overhead for the client. Otto and Jarke (2019) designed a multi-sided data platform (MSP) based on IDS for secure and trusted data exchange but they only conducted a preliminary study and design on MSP's lifecycle

Data security and sovereignty are one of the key features for IDS-RAM, thus the data connector need to ensure the data exchange between manufactures are secured.

3.0 DISCUSSION AND MODELS

Although many researchers have developed different architectures to improve the manufacturing process and data exchanges for an organization, some other issues such as availability, fault tolerance and services discovery are still remained unsolved. Baboi, Iftene and Gîfu (2019) showed that microservices architecture can improve fault tolerance and scalability using various technologies from .Net to Java. Their studies showed how flexible and dynamic the microservices are, but the fault tolerance capability is built in a very basic way, which is to remove the server from the list if the server does not respond in the timely manner. Microservices architecture has been widely used in many fields to improve modularity, allow integration with different systems, platforms, and legacies. However, there are some other issues to be addressed. For instance, how to design an effective dynamic service allocation algorithm, to optimize the system's performance in runtime by balancing between computing power and latency. Besides, it is also important to have effective messaging techniques that could further improve the latency issue of microservices. Some other specific challenges when designing data centric multi-sided data platform (MSP), for example security, performance, data exchange optimisation, should be addressed too.

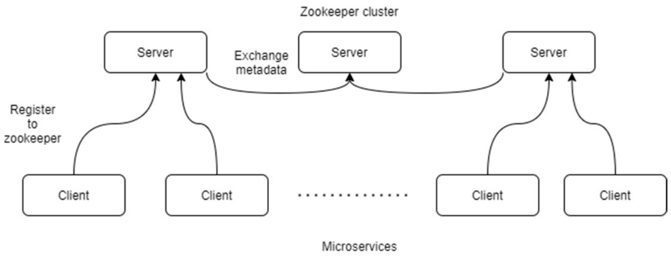

Figure 3. Zookeeper architecture

A centralized service discovery such as Zookeeper[6] can be developed to store the metadata of the services, service registration and discovery as shown in figure 3. Zookeeper ensures data consistency based on CAP theorem (Fox and Brewer, 1999). With this centralized approach, Zookeeper is able to provide actual status of the services and manage the services within the cluster. However, this architecture is highly dependent on the Zookeeper's performance. If the Zookeeper is down or delay in the metadata exchange when a new Zookeeper is added to the cluster, those services that are under management will not be able to communicate with each other. Besides, it has always been a challenge to design a secure microservices ecosystem that ensures data integrity and security as most of the services are communicated remotely. In addition, a trusted gateway needs to be built to verify and route the authorized request from the external world to the respective microservices. Therefore, a secure, responsive, high resilience and elasticity are very important criteria to build an ideal dataspace ecosystem.

To overcome the above-mentioned challenges, reactive systems, introduced in the Reactive Manifesto (Bonér *et al.*, 2014), enables consensus protocol, such as gossip protocol, to manage the distributed systems in a decentralized architecture to increase the isolation and autonomy of the system. Cluster modules manage its own cluster membership, load balancing and node partitioning, and achieve high availability based on CAP theorem. The combination

---

[6] https://www.ibm.com/analytics/hadoop/zookeeper

of the simplicity and flexibility of the microservices architecture with reactive system principles, also known as Reactive Microservices Architecture, may help to solve the challenges. The implementation of the reactive microservices should be done based on the Reactive Manifesto four principles: responsive, resilient, elastic and message-driven. The role of the service discovery like zookeeper, consul will act as service lookup to allow cluster modules to be access by external services.

4.0 CONCLUSIONS

Based on the findings of literature review in the four main areas, namely Service-Oriented Architecture (SOA) and microservices, service orchestration, systems interoperability and trusted security. The SOA and the microservices development have the following advantages: (1) more flexible, (2) could decouple the applications, (3) enhance the maintainability, (4) reduce complexity of the software architecture, (5) increase the scalability and performance. However, the studies showed that the architecture increases the complexity to manage and monitor the huge numbers of services running on the premises in the system. The main challenges which are not explored in the past researches is how to manage the multiple clusters of the services in more efficient way without losing any single data.

Reactive microservices architecture is believed to achieve a fault tolerant and fast data processing system, enabling value-chain integration and real-time data exchange as recommended in RAMI 4.0 and IDS-RAM. For interconnectivity and handling of unlimited stream of data from various sources such as IoT, on-premise servers, cloud, edge devices between data providers and data consumers, traditional HTTP or RESTFul protocol usually having difficulty to handle the huge load of the data, since each request needs a complete response result before it releases the connection and thread. The device or application that accepts huge data via the HTTP protocol required to keep the data in memory until the data was completely sent out. Therefore, it usually requires a large amount of memory with traditional monolith application architectural design. To address these issues, an asynchronized data streaming technology shall process the data in small chunk and to have the backpressure capability to control the amount of data transferred from the source to prevent data overflow.[7] The future work is to develop a dataspace connector with reactive microservice to offer a system with seamless data connectivity and high availability for data exchange between organizations in the industrial dataspace.

REFERENCES

Ahuett-Garza, H. and Kurfess, T. (2018) ‘A brief discussion on the trends of habilitating technologies for Industry 4.0 and Smart manufacturing’, Manufacturing Letters, 15, pp. 60–63. doi: 10.1016/j.mfglet.2018.02.011.

Alonso, Á. et al. (2018) ‘Industrial Data Space Architecture Implementation Using FIWARE’, Sensors, 18(7), p. 2226. doi: 10.3390/s18072226.

Baboi, M., Iftene, A. and Gîfu, D. (2019) ‘Dynamic microservices to create scalable and fault tolerance architecture’, Procedia Computer Science, 159, pp. 1035–1044. doi: 10.1016/j.procs.2019.09.271.

Bigheti, J. A., Fernandes, M. M. and Godoy, E. D. P. (2019) ‘Control as a Service: A Microservice Approach to Industry 4.0’, 2019 IEEE International Workshop on Metrology for Industry 4.0 and IoT, MetroInd 4.0 and IoT 2019 - Proceedings, pp. 438–443. doi: 10.1109/METROI4.2019.8792918.

---

[7] https://nordicapis.com/rest-vs-streaming-apis-how-they-differ/

Bonér, J. et al. (2014) ‘The Reactive Manifesto (Version 2.0)’, Reactivemanifesto.Org, 2(16 September 2014), p. pp.1-2. Available at: http://www.reactivemanifesto.org.

Brost, G. S. et al. (2018) ‘An ecosystem and IoT device architecture for building trust in the industrial data space’, CPSS 2018 - Proceedings of the 4th ACM Workshop on Cyber-Physical System Security, Co-located with ASIA CCS 2018, pp. 39–50. doi: 10.1145/3198458.3198459.

‘DIN SPEC 27070:2020-03’ (2020). doi: 10.31030/3139499.

‘DIN SPEC 91345’ (2016) in. Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.31030/2436156.

Dinh-Tuan, H., Beierle, F. and Garzon, S. R. (2019) ‘MAIA: A microservices-based architecture for industrial data analytics’, Proceedings - 2019 IEEE International Conference on Industrial Cyber Physical Systems, ICPS 2019, pp. 23–30. doi: 10.1109/ICPHYS.2019.8780345.

Federal Ministry for Economic Affairs and Energy (2019) ‘Plattform Industrie 4.0 - RAMI4.0 – a reference framework for digitalisation’, Plattform Industrie 4.0.

Fowler, M. and Lewis, J. (2014) Microservices. Available at: https://martinfowler.com/articles/microservices.html.

Fox, A. and Brewer, E. A. (1999) ‘Harvest, yield, and scalable tolerant systems’, Proceedings of the Workshop on Hot Topics in Operating Systems - HOTOS, pp. 174–178. doi: 10.1109/hotos.1999.798396.

G. Hohpe, B. W. (2003) Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions. Addison-Wesley.

Girbea, A. et al. (2014) ‘Design and implementation of a service-oriented architecture for the optimization of industrial applications’, IEEE Transactions on Industrial Informatics, 10(1), pp. 185–196. doi: 10.1109/TII.2013.2253112.

Karnouskos, S., Bangemann, T. and Diedrich, C. (2009) ‘Integration of legacy devices in the future SOA-based factory’, IFAC Proceedings Volumes (IFAC-PapersOnline), 13(PART 1), pp. 2113–2118. doi: 10.3182/20090603-3-RU-2001.0487.

Landolfi, G. et al. (2019) ‘A MaaS platform architecture supporting data sovereignty in sustainability assessment of manufacturing systems’, Procedia Manufacturing, 38(Faim 2019), pp. 548–555. doi: 10.1016/j.promfg.2020.01.069.

Li, X., Xuan, Z. and Wen, L. (2011) ‘Research on the architecture of trusted security system based on the internet of things’, Proceedings - 4th International Conference on Intelligent Computation Technology and Automation, ICICTA 2011, 2, pp. 1172–1175. doi: 10.1109/ICICTA.2011.578.

Lin, S.-W. et al. (2017) ‘The Industrial Internet of Things Volume G1 : Reference Architecture’, Industrial Internet Consortium White Paper, Version 1., p. 58 Seiten.

Lu, Y. (2017) ‘Industry 4.0: A survey on technologies, applications and open research issues’, Journal of Industrial Information Integration, 6, pp. 1–10. doi: 10.1016/j.jii.2017.04.005.

Munoz-Arcentales, A. et al. (2020) ‘Data usage and access control in industrial data spaces: Implementation using FIWARE’, Sustainability (Switzerland), 12(9). doi: 10.3390/su12093885.

OSGi Alliance (2018). Available at: https://www.osgi.org/developer/architecture/.

Otto, B. et al. (2019) ‘International Data Space Reference Architecture Model Version 3.0’, (April), pp. 1–118. Available at: https://www.internationaldataspaces.org/wp-content/uploads/2019/03/IDS-Reference-Architecture-Model-3.0.pdf.

Otto, B. and Jarke, M. (2019) ‘Designing a multi-sided data platform: findings from the International Data Spaces case’, Electronic Markets, 29(4), pp. 561–580. doi: 10.1007/s12525-019-00362-x.

Petrasch, R. (2017) ‘Model-based engineering for microservice architectures using Enterprise

Integration Patterns for inter-service communication', Proceedings of the 2017 14th International Joint Conference on Computer Science and Software Engineering, JCSSE 2017, pp. 2–5. doi: 10.1109/JCSSE.2017.8025912.

Santana, C. et al. (2021) 'Increasing the availability of IoT applications with reactive microservices', Service Oriented Computing and Applications, 15(2), pp. 109–126. doi: 10.1007/s11761-020-00308-8.

Santana, C., Alencar, B. and Prazeres, C. (2018) 'Microservices: A mapping study for internet of things solutions', NCA 2018 - 2018 IEEE 17th International Symposium on Network Computing and Applications. doi: 10.1109/NCA.2018.8548331.

Teslya, N. and Ryabchikov, I. (2017) 'Blockchain-based platform architecture for industrial IoT', in 2017 21st Conference of Open Innovations Association (FRUCT). IEEE, pp. 321–329. doi: 10.23919/FRUCT.2017.8250199.

Theorin, A. et al. (2017) 'An event-driven manufacturing information system architecture for Industry 4.0', International Journal of Production Research, 55(5), pp. 1297–1311. doi: 10.1080/00207543.2016.1201604.

Trunzer, E. et al. (2019) 'System architectures for Industrie 4.0 applications: Derivation of a generic architecture proposal', Production Engineering, 13(3–4), pp. 247–257. doi: 10.1007/s11740-019-00902-6.

Xiao, Z., Wijegunaratne, I. and Qiang, X. (2017) 'Reflections on SOA and Microservices', Proceedings - 4th International Conference on Enterprise Systems: Advances in Enterprise Systems, ES 2016, pp. 60–67. doi: 10.1109/ES.2016.14.

Xu, J. et al. (2020) 'Privacy-preserving data integrity verification by using lightweight streaming authenticated data structures for healthcare cyber–physical system', Future Generation Computer Systems, 108, pp. 1287–1296. doi: 10.1016/j.future.2018.04.018.