Factory+

Literature Review

for

Catapult

AIN4918

22nd June 2020

<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author:</td>
<td>Alex Godbehere</td>
</tr>
<tr>
<td></td>
<td>Approved through Windchill</td>
</tr>
<tr>
<td>Technical review:</td>
<td>Harry Burroughes</td>
</tr>
<tr>
<td></td>
<td>Approved through Windchill</td>
</tr>
<tr>
<td>Authorised:</td>
<td>Chris Greaves</td>
</tr>
<tr>
<td></td>
<td>Approved through Windchill</td>
</tr>
</tbody>
</table>

Authorised distribution: The University of Sheffield proprietary information

AMRC proprietary information.
AMRC
Advanced Manufacturing Park
Wallis Way
Rotherham
S60 5TZ
United Kingdom

The AMRC has prepared this report on behalf of (Partner/Customer).

This report may only be reproduced in full, except with the express permission of the AMRC.

The AMRC has exercised due care in conducting this report but has not, apart from as specifically stated, independently verified information provided by others. No other warranty, express or implied, is made in relation to the contents of this report. Therefore, the AMRC assumes no liability for any loss resulting from errors, omissions, or misrepresentations made by others. Any recommendations, opinions or findings stated in this report are based on circumstances and facts as they existed at the time the AMRC performed the work. Any changes in such circumstances and facts upon which this report is based may adversely affect any recommendations, opinions or findings contained in this report. Where experiments have been carried out, these have been restricted to a level of detail required to achieve the stated objectives of the statement of work. This work has been undertaken in accordance with the AMRC’s quality system.

Document revision history

<table>
<thead>
<tr>
<th>Revision</th>
<th>Date of issue</th>
<th>Reason for revision or original issue</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td></td>
<td>Original issue</td>
</tr>
</tbody>
</table>

Confidential. Copyright © The University of Sheffield / Jul-20 AMRC.
Template detail: AMRC.RP Revision 16 (December 2018). For information about this template contact: qande@amrc.co.uk
Executive summary

Due to the rising popularity of the Industrial Internet of Things (IIoT) and Industry 4.0 (I4.0), a plethora of frameworks, protocols and methods have been developed to aid the adoption of these new promising digital technologies. Although each of the methods is undoubtedly effective when deployed correctly, the sheer number of choices available is making the landscape ‘jungle-like’, reducing agility and stifling the adoption of IIoT in UK manufacturing (1).

This document aims to condense the current landscape into a concise format, using the latest literature to answer a series of questions surrounding the current state of IIoT and the Smart Factory. This document will not offer recommendations; its sole purpose is to provide a succinct summary of the state of current technologies.
Contents

1 Introduction ............................................................................................................................. 6

2 Frameworks ............................................................................................................................ 7
   2.1 Why is a framework required? .......................................................................................... 7
   2.2 What frameworks should be considered? ......................................................................... 7
   2.3 What are the differences between IIAR and RAMI4.0 and is there interoperability? .... 8
   2.4 Is ISA-95 still relevant in IIoT? .................................................................................... 8

3 Protocols ................................................................................................................................ 8
   3.1 Should a factory have a single core connectivity standard? ............................................. 9
   3.2 What protocols should be adopted for secure, reliable, future-proof factory communications? .......... 9
   3.3 What is OPC UA PubSub? ............................................................................................ 10

4 Infrastructure ........................................................................................................................ 10
   4.1 Primary Application ....................................................................................................... 10
   4.2 Edge Architecture ......................................................................................................... 10
   4.3 Should a gateway be used as a single point of entry to a cell? ....................................... 11
   4.4 Are there any emerging technologies poised to disrupt future factory architectures? .... 11
      4.4.1 5G .......................................................................................................................... 11
      4.4.2 Time Sensitive Networks ....................................................................................... 11
      4.4.3 Software-Defined Networking .............................................................................. 12

5 Deployment & Containerisation ............................................................................................. 12
   5.1 What functions should a smart factory perform? ............................................................ 12
   5.2 How can containerisation be used to help deploy and run IIoT software? .................... 13

6 Legacy .................................................................................................................................. 13
   6.1 Why is supporting legacy devices so important? ............................................................ 13
   6.2 How should legacy devices be connected to a modern smart factory? ......................... 14
   6.3 Can legacy machine integration be achieved on a budget? ............................................ 14
<table>
<thead>
<tr>
<th>Document</th>
<th>Revision</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMG005497</td>
<td>Release</td>
<td>5 of 23</td>
</tr>
</tbody>
</table>

7 Security ........................................................................................................................................... 14

7.1 How can data be kept safe in transit? ...................................................................................... 14

7.2 How can data at the edge be protected? ................................................................................... 15

Appendix A – Culture ............................................................................................................................ 18

Terms and Definitions ............................................................................................................................ 19

References ........................................................................................................................................... 20
1 Introduction

Industry 4.0 is not a new concept. The term was first coined almost ten years ago in 2011 at Hannover Fair by the German government (2), defined as “Utilising the power of communications technology and innovative inventions to boost the development of the manufacturing industry” (3). This newfound power not only promised to drive business return on investment and drive top-line growth (4) but also to increase efficiency, reduce waste and lower energy consumption (5).

With organisations keen to benefit from these promises, a plethora of frameworks, protocols and methods have been developed to aid the adoption of these new digital technologies. Although each of the methods is undoubtedly effective when deployed correctly, the sheer number of choices available is making the landscape ‘jungle-like’, reducing agility and stifling the adoption of IIoT in UK manufacturing (1).

In addition to the influx of new protocols and frameworks are a host of new buzzwords and phrases pertaining to the ‘Fourth Industrial Revolution’, many of which are used interchangeably. For example, IIoT and I4.0 are frequently used to refer to Industry 4.0 and the ‘Smart Factory’ but this isn’t necessarily the case. IIoT represents the concept of all devices being connected, enabling them to ‘talk’ to each other to make intelligent decisions and optimise their performance, whereas I4.0 encompasses the advances and innovations that are enabled by said connected technologies. IIoT is a component of I4.0, not an interchangeable concept. I4.0 would not exist without IIoT and IIoT would be limited in use without the bigger picture of I4.0 (6).

Like IIoT, a Smart Factory is a key component of I4.0. A Smart Factory embodies the very IIoT devices that enable all of the benefits of the Industry 4.0 concept. One of the key characteristics of a ‘smart’ industrial ecosystem such as the one offered by Industry 4.0 and embodied within the Smart Factory is seamless end-to-end communication from the enterprise level down to edge devices and sensors on the shop floor, enabling the manufacture of highly-customisable products that are realised through flexible mass production and minimising downtime (4).

Interoperability of systems in this fragmented landscape is one of the most commonly identified barriers to adoption of I4.0 technology - technology which can improve productivity, reduce waste, and improve life for the workforce. To demonstrate how the adoption of a common framework can significantly reduce complexity when implementing a Smart Factory, the AMRC are building on the cutting-edge research from previous projects and are launching ‘Factory+’, a site-wide implementation of a digitally connected factory across the AMRC’s Factory 2050 facility. The implementation aims to standardise the connection between devices on the shop floor and act as an operational tool, a research sandbox and a development testbed for future IIoT research within the AMRC. The scope of work outlined within the Factory+ project has been strategically defined to reduce the barriers to adoption (data standards, interoperability and knowledge capture) as outlined in the AMRC’s ‘Development of a 4IR (Industry 4.0) Ready Machining Facility’ (7).

This report is a review of the latest literature to enable informed decisions to be made for the architecture definition phase of the AMRC’s Factory+ project using the latest information and best practices.
2 Frameworks

Multiple international consortiums exist that produce Industry 4.0 frameworks and reference architectures. This section will investigate the requirement for a framework, review two of the most popular frameworks and summarise the differences between them.

2.1 Why is a framework required?

Diversity of standards is largely expected in the manufacturing sector due to the large variety of machines and equipment used in factories all over the world. To support an Industry 4.0 vision, however, this diversity requires that all devices are integrated into a common system requiring different types of data acquisition methods, file transfer protocols, data storage techniques and data transmission protocols to be individually connected, resulting in increased development time and reducing the very flexibility to change and adapt that the Industry 4.0 paradigm seeks to provide. By utilising a standard framework for data collection, transport and storage, factory communications can remain efficient (due to no longer relying on cross-protocol translation layers), data models can be standardised, data can remain consistent (by belonging to a single source of truth) and information can be standardised for consumption in the Enterprise Information System domain (1).

2.2 What frameworks should be considered?

In 2017 the AMRC produced a report that provided a best-practice framework for the development of a use-case specific Enterprise Architecture (7). This report identified The Industrial Internet Consortium (IIC) and the Working Group for Industry 4.0 as the two main organisations providing recommendations and guidelines for Industry 4.0 architectures. Both IIC and The Working Group for Industry 4.0 have produced frameworks for adopting Industry 4.0 (IIRA and RAMI4.0 respectively).

The Industrial Internet Reference Architecture (IIRA) is a generalised blueprint architecture template to enable IIoT architects to design systems according to a common set of standards, concepts and best practices. It addresses the need for a common framework to develop interoperable IIoT systems across a broad range of cross-industry verticals. (8). The concepts that it presents are high-level and abstract, providing the architect with maximum flexibility. IIRA outlines the requirement for a standardised communication architecture and protocol, citing the N² problem (9) when faced with connecting applications that communicate over different connectivity technologies. In addition, it outlines the requirement of a Publish/Subscribe architecture (as opposed to a more traditional Client/Server) to achieve the reliability, scale and performance demanded by IIoT systems in addition to enabling incremental updating and evolution, a crucial requirement for a research facility.

Reference Architectural Model Industrie 4.0 (RAMI4.0) is much more manufacturing-focused than IIRA. It ensures that all participants involved share a common perspective and develop a common understanding. As a whole, RAMI4.0 is represented as a three-dimensional map showing the most important aspects of Industry 4.0, comprising of Business, Functional, Information, Communication, Integration and Asset layers. For this evaluation, the focus will primarily be on the Communication, Integration and Asset layers. RAMI4.0 supports the idea of a mesh network, utilising ‘Administration Shells’ to access legacy devices and assets, as opposed to the more traditional hierarchal approach. Traditionally this hierarchal ecosystem is represented by part 1 the ISA-95 standard that breaks down the entire system into a set of levels, with data flowing from the process level (Level 0), through the sensing (Level 1), monitoring (Level 2), Manufacturing Operations Management (Level 3) and Business Planning levels (Level 4). The RAMI4.0 architecture declares this as the ‘old’ way of doing things and is a strong advocate of a more distributed approach (10).

To adopt a distributed approach, the concepts of object types and administration shells must be understood. Dr Madhusudan Pai summarises these in the white paper ‘Interoperability Between IIC Architecture & Industry 4.0’.
Reference Architecture for Industrial Assets’ (4) as ‘An object can be a product, asset, software, machine, or even a factory. Objects that have the ability to communicate independently using I4.0 compliant communication are called I4.0 compliant components. Non-I4.0 compliant components can be made I4.0 compliant by deploying an administration shell, which essentially provides a virtual representation and a description of the entire life-cycle of the object or asset, in addition to I4.0 compliant communication with the rest of the value chain.’

2.3 What are the differences between IIRA and RAMI4.0 and is there interoperability?

Since both RAMI4.0 and IIRA both serve to define reference architectures for an IIoT ecosystem, they naturally have some crossover resulting in several discussions regarding their interoperability gaining prominence (7) (4).

- IIRA is much more generalised than RAMI4.0, focusing on only a critical few characteristics – namely safety, security, and resilience. It adopts general concepts from the ISO/IEC/IEEE 42010:2011 standard which includes concerns, framework and viewpoints.

- Although both IIRA and RAMI4.0 both focus more on the bigger picture, they both stipulate the need for a service-oriented-architecture (SOA) to encapsulate functionalities into services, ensuring that any SOA-based protocol that is adopted (e.g. OPC UA) provides semantic interoperability by default.

- RAMI4.0 is primarily focussed on creating smart manufacturing value chains, whereas IIRA has much broader focus areas such as energy, healthcare and transportation. The scope of the work that this literature review is designed to inform surrounds factory connectivity in the IIoT space, therefore RAMI4.0 provides more concrete concepts for architectures and connectivity patterns.

It is widely acknowledged that IIoT solutions and frameworks will develop that are a hybrid of RAMI4.0 and IIRA, leading to comprehensive end-to-end IIoT ecosystems. It is important to note that the combination and expansion of techniques described in RAMI4.0 and IIRA is encouraged to suit specific needs and business cases (4).

2.4 Is ISA-95 still relevant in IIoT?

As technologies develop and Industry 4.0 concepts become more widely adopted, the standards that have been underpinning manufacturing for decades must also develop to remain relevant. Unlike technology, the information model that ISA-95 standard defines does not become dated over time and the concepts that it defines are as relevant as ever in the modern connected world.

There is a general tendency to take the reference model in part 1 of the ISA-95 standard too literally, with Cosman stating that the logical levels of the ISA-95 pyramid do not represent physical domains for implementation (11). It is important to decouple ISA-95 from the traditional layered network hardware model as in new emerging environments the middleware layer will effectively ‘evaporate’ to the cloud (12).

In summary, the ISA-95 standard is a critical building block to architectures such as RAMI4.0, not a replacement for it. Although the traditional physical network topologies that ISA-95 compliant systems were based on will change in the new world, the most important part of the ISA-95 standard, the information model, remains as relevant as ever.

3 Protocols

This section aims to uncover the current state of the landscape with regards to the industry-standard protocols and techniques recommended for use by the IIC, the Working Group for Industry 4.0 and other industrial leaders.
3.1 Should a factory have a single core connectivity standard?

Having to implement and manage multiple legacy systems and connectivity standards generates additional work, increases risk and stifles innovation (13). To keep the connectivity architecture manageable IIRA suggests a single core connectivity standard per functional domain but does make the case for multiple connectivity standards across domains if business rules dictate a requirement (9). In both cases, gateways should be used to connect non-compliant devices to the core connectivity standard in a single domain (analogous to administration shells in RAMI4.0) and to bridge connectivity standards between functional domains. By adopting a connectivity standard per domain, the $N^2$ problem of an exponentially increasing number of protocols and endpoints to manage as new devices and protocols are introduced becomes largely linear.

Similarly, RAMI4.0 recommends common communication structures (networks, protocols, rules, syntax and languages) in order to reduce complexity and enable standardised processes across the business and between administration shells (10). For example, the use of technologies such as OPC-UA, HTTPS and MQTT are recommended for use in the communication layer.

3.2 What protocols should be adopted for secure, reliable, future-proof factory communications?

Although data structuring formats such as JSON and XML have been demonstrated to function effectively in a factory environment over IoT protocols such as HTTP, MQTT (14) and OPC Unified Architecture (OPC UA) are becoming increasingly established as the industrial standards for vendor-neutral industrial communication (15), with OPC UA being defined as ‘one of the pillars of homogenous industrial communication’ by modern literature (16).

In addition to the best practices documented by the literature, both the IIC and the Working Group for Industry 4.0 indicate OPC UA as the international standard to be used for machine to machine communication in IIRA and RAMI4.0 frameworks (17). The IIC’s IIRA defines OPC UA as playing a strategic role in the standardisation of IIoT devices by identifying it as one of the core connectivity standards and the Working Group for Industry 4.0’s RAMI4.0 indicates for OPC UA or MQTT to be used to standardise communication (10) (9). OPC UA is recommended by these organisations in part due to the flexibility of the technology. It can provide robust, secure communications between machines on the shop floor in a ‘vertically-integrated factory set up’ in addition to also providing heterogeneous communications at the enterprise level (4).

Due to OPC UA’s deep, native integration with PLCs and existing control system infrastructure, in addition to being highly compatible and compliant with industrial fieldbus protocols like EtherCAT, it has gained the support of organisations such as Microsoft, who define OPC UA as ‘the established, worldwide data modelling standard for Industrial IoT’ (18), the Fraunhofer Society via OpenIoTFog (19) and the ZVEI manufacturers’ association, who define OPC UA as part of the ‘Criteria for Industry 4.0 Products’ (20).

Although OPC UA is currently recommended by many to connect IIoT assets, other technologies are becoming increasingly mentioned in the literature. MQTT, for example, is widely regarded as the de-facto protocol for IoT due to widespread support, lightweight requirements, quality of service options, and a many-to-many, publish-subscribe model. Standard MQTT message payloads often utilise plaintext JSON-encoded ASCII strings which can be read by a very broad range of applications; however, this is less efficient than binary encoding and does not provide end-to-end encryption of the message contents by default to keep the specification flexible and lightweight. The MQTT protocol, therefore, offers unique advantages at the transport layer, but additional work is required to secure the message contents.

One of the reasons that OPC UA is currently the most widely used standard is due to the interoperability between OPC UA and other IoT communication systems such as MQTT and HTTPS. Although MQTT was invented back in 1999, it is only since the adoption of the technology by Facebook for its Messenger platform that it has started to see widespread use in the industrial sector, despite originally being invented in to reliably connect thousands of industrial devices over huge networks. Due to the plethora of benefits that MQTT provides, it is starting to be considered by many to be the foundation of the ‘next-generation’ of smart factory networks.
Many of the shortcomings of MQTT when compared to OPC UA (inefficient encoding, lack of security and lack of information models) can be solved by implementing a standard such as Sparkplug. Sparkplug (21) is a topic and payload definition specification for industrial devices communicating over MQTT. The latest revision of Sparkplug (B) provides an open and freely available specification for how Edge of Network (EoN) gateways or native MQTT enabled end devices and MQTT Applications communicate bi-directionally within an MQTT Infrastructure. It solves many of the problems that using MQTT for IIoT introduces such as state management, efficient encoding (using Google Protocol Buffers) and payload structure. When implementing an MQTT/Sparkplug stack the network bandwidth savings over a protocol like Modbus can be as large as 99.5% (22), which is critical as the expected number of devices fighting for the same amount of bandwidth is set to skyrocket as IIoT is adopted across the manufacturing industry.

3.3 What is OPC UA PubSub?

Similarly, many of the features that are making architects seriously consider technologies like MQTT (e.g. report by exception) are being added into the OPC UA specification with additions such as Part 14 introducing Publish and Subscribe functionality. As of this recently published addition to the OPC UA Specification (23), OPC UA supports the same many-to-many, publish-subscribe model that MQTT adopts, making it even more suitable for the distributed approach of the 14.0 vision defined by RAMI4.0. An aspect of this recent PubSub addition to the OPC UA specification is to use MQTT to deliver OPC UA binary messages, enabling the benefits of topic-based many-to-many communication without sacrificing the feature set that makes OPC UA so popular and fit for purpose in an industrial environment.

Although OPC UA PubSub has been earmarked to solve many challenges in the IIoT space such as reducing downtime in time-critical applications (24), enabling seamless connectivity at the enterprise level (25) and accelerating the development and adoption of Time-Sensitive Networking (TSN) within manufacturing (26) it is worth noting that OPC UA over PubSub is noticeably slower than MQTT/Sparkplug for data transmission in time-critical applications due to the complexity and overhead of the OPC UA stack (27).

4 Infrastructure

This section aims to investigate whether an industry-standard network topology and architecture exists that allows for optimised communications, enables efficient maintenance and adheres to strict cyber-security policies whilst still being flexible enough to adhere to diverse business rules.

4.1 Primary Application

The purpose of deploying a smart factory architecture is to ensure that data is unified in a single namespace with a single source of truth. Once the data has been collected and stored in the unified namespace, any number of Operational Technology (OT) or Information Technology (IT) entities can access it, providing they have the correct permissions, to process or present the data. A common approach that is advocated by the Sparkplug specification is to refer to any number of entities consuming data as ‘Applications’, with a single ‘Primary Application’ responsible for the monitoring and control of a given group of edge nodes (in addition to being able to consume data).

4.2 Edge Architecture

Edge devices can be categorised as either ‘near-edge’ or ‘far-edge’. Far-edge devices are used to connect the factory network to the cloud and act as a single access point to the factory network for the cloud. Near-edge devices are typically used at the other side of the factory network to handle the interaction between field devices (PLCs, controllers, sensors, etc.) and usually handle protocol translation logic, local buffering, logging, alert management and additional processing such as machine learning or data tagging.
4.3 Should a gateway be used as a single point of entry to a cell?

There are benefits and drawbacks to using gateways to connect devices in a cell to the wider factory network. RAMI4.0 is a strong advocate of a truly distributed approach whereby all devices are connected directly to the factory network (10) and using a gateway as the access point to the factory, however in their document ‘Secure implementation of OPC UA for operators, integrators and manufacturers’ the German Government suggests the use of gateways to connect machines and cells inside the factory when the factory is part of a wider cloud-based platform (20), citing security reasons.

The decision of whether or not to use a gateway-per-cell approach is also largely due to the type of organisation implementing it. For example, the requirement to implement legacy machines (see Legacy) in a brownfield site would traditionally result in a different architecture to that of one adopted by an organisation implementing a smart factory on a greenfield site. Is possible to use a low-cost (28) in-cell gateway when connecting legacy machines to a smart factory infrastructure to translate the legacy device’s native protocol and the core connectivity standard of the factory (administration shell) (15), however, if a gateway-per-cell architecture is adopted this functionality could be delegated to the cell gateway, standardising the architecture between greenfield and brownfield sites.

The gateway-per-cell model offers advantages in addition to solving the protocol incompatibility problem. By architecting the network topology in such a way that all data flowing from a cell flows through a gateway, the gateway can offer additional edge-based processing on the data, enabling functionality such as data tagging, subjective data transmission and edge analytics (20).

Choosing to adopt a gateway-per-cell architecture can also help support business rules that an organisation might have. For example, the AMRC frequently takes cells off-site to demonstrate them at tradeshows and events. By structuring cell assets to communicate only through the gateway, the responsibility for handling the connection back to the AMRC site via the cloud is delegated to the gateway resulting in a single configuration change. This single responsibility principle promotes agility within the organisation and is a key enabler for next-generation reconfigurable factories.

Ultimately the decision to use gateways should be one made based on the business rules and legacy requirements, however, the vast majority of the literature reviewed strongly recommended the adoption of a gateway-per-cell approach, including the official ‘Practical Security Recommendations for building OPC UA Applications’ guide by the OPC Foundation (29).

4.4 Are there any emerging technologies poised to disrupt future factory architectures?

IIoT and factory connectivity is a volatile, fast-moving field with advances continually threatening to disrupt the way that smart connected factories are architected. Although the technologies listed below are not mature enough for a production facility to benefit from at the time of writing, they are each likely to play a big part in the way we connect devices both inside and outside the factory in the future.

4.4.1 5G

Although the likely impact of 5G technologies is primarily to catalyse and accelerate the impact of IIoT in both greenfield and brownfield sites (30), it will be a key component in supporting the massive amounts of manufacturing data that will be generated as more facilities and machines are connected to smart factories around the world. Current 3rd generation mobile networks (3G and 4G) are unable to meet the predicted demands of future cyber-physical manufacturing systems (CPMS) such as high-data-rate, high reliability, high coverage and low-latency (31).

4.4.2 Time Sensitive Networks

One of the key components of the future smart factory is real-time communications. Advances in time-sensitive networks (TSN) are starting to influence the evolution of factory networks and will gain prominence in the coming
years due to their ability to achieve low transmission latency and high availability. Industrial TSN systems are in their infancy and current implementations rely on timing analysis using network calculus and trajectory approach methods that require a detailed blueprint of the network topology to function effectively. There has been some work investigating plug-and-play TSN however these don’t scale well and require a star-shaped network topology, reintroducing the $N^2$ problem (30).

4.4.3 Software-Defined Networking

Software-Defined Networking (SDN) is a concept that breaks the traditional, monolithic network architecture by enabling traffic routes to be software-defined, enabling rapid reconfigurability by a high-level SDN-controller, promoting agility and enabling the ability to cater to dynamically changing needs.

SDN allows separation of control logic from the underlying physical devices, providing flexibility for communications and maintenance and allowing for easier, dynamic reconfiguration of plants as needed by smart manufacturing and highly customised products (30). SDN is an up-and-coming topic in the literature and is poised to accelerate the adoption of architectures such as RAMI4.0 due to the mesh-type, distributed network that it enables (32).

5 Deployment & Containerisation

Beneath the network topology of the smart factory lies software and data rules that drive the logic behind the transmission and storage of manufacturing data from the machines and devices on the shop floor. Although the concrete software requirements will vary between implementations and organisations the methods used to maintain, deploy and update these software packages should be standardised according to emerging best practices. This section outlines how modern containerisation and orchestration technologies can be used when deploying software within the smart factory.

5.1 What functions should a smart factory perform?

To assess whether containerisation is suitable for the smart factory, it is first important to understand what functions a smart factory needs to be able to perform.

Petit et al. (33) define that a smart factory must be able to:

- Store, retrieve and analyse data at the required granularity
- Standardise commonly used data across platforms and locations
- Enable the creation of digital mock-ups of various assets
- Enable access to data for visualisation and analytical tools
- Establish a data governance framework that can govern the principles of areas such as access, retention and deletion.

Except for points 2 and 5, the above functions can be extracted into ‘services’ offered to the end-user, similar to applications on a smartphone. By isolating each of these services into its own container the architecture can remain easy to manage and expand in addition to providing increased security as each application runs in a sandbox.

By architecting a common smart factory platform from points 2 and 5, it is possible to expose the gathered datasets to decoupled applications that wish to consume it (analysis, storage, archiving, reporting, digital twin, visualisation, etc.), a fitting candidate for the microservices architecture described below.
5.2 How can containerisation be used to help deploy and run IIoT software?

The migration to cloud platforms such as Microsoft Azure is becoming increasingly widespread in the literature as the benefits of distributing computational power to the cloud is realised and more organisations adopt the RAMI4.0 and IIIRA frameworks, both which propose distributed architectures (34).

Due to the modularity and scalability requirements of these new architectures, technology solutions such as Docker and Kubernetes are gaining prominence in the manufacturing environment (35) (36). This can be seen from the growing popularity of IIoT platforms that are built on these technologies such as Siemens MindSphere (37) and the trend of global IT and IIoT cloud providers such as Dell, HPE, Google Cloud Platform, Amazon AWS and Microsoft Azure all natively supporting containerisation technologies for IIoT applications.

In addition to the modularity and scalability that containerisation provides, the concept of microservices (many smaller containers working together) in part solves the requirement of devices on IIoT networks to operate cooperatively to gather and share data in a highly fault-tolerant, high-availability and self-healing fashion (36). Containers are extremely lightweight, enabling many of them to be run on low-cost, small-scale hardware, in large numbers to truly enable the IIoT and Industry 4.0.

Containers decouple the underlying application code from the infrastructure that it runs on, allowing for variances in hardware and computing ability, significantly reducing the effort required to upgrade hardware, scale both horizontally and vertically and recover from disasters as efficiently as possible. The use of containers in IIoT also has the potential to reduce the number of software bugs introduced into a production environment by enabling developers to test the code offline in a completely representative environment before deploying to prevent unexpected results when running the code on different hardware to that of the testing environment.

Building up a repository of containers for different key functions across the factory (data capture, storage, processing, analytics, etc.) enables software functions to easily be deployed on gateways or edge devices to repeatably process data at the machine/cell level in the same way that they would be deployed at a factory or cloud level. By adopting a containerised strategy on-site, an organisation is aligning its operations for effortless migration to the cloud (either on or off premises) if required in the future.

6 Legacy

Legacy, non-IIoT capable devices will be commonplace in manufacturing environments for years to come. With 80% of global companies feeling more positive about the potential of IIoT than they were 12 months ago (38), the integration of legacy devices into smart factories has been identified as one of the major barriers to adoption (13).

6.1 Why is supporting legacy devices so important?

Sectors such as aerospace are poised to benefit greatly from the digitisation of legacy systems to enhance the capability of supply chains, enable end-to-end integration via the capture and exploitation of manufacturing data and the implementation of artificial intelligence (AI) to support decision making, design, fabrication and operation (13). Aerospace has been identified as a prime industry to benefit from the digitisation of legacy systems due to the large number of ‘dumb’ devices on industrial shop floors (39), with Cisco estimating that 92% of all 64 million machines worldwide are not connected to a network of any kind (40).

The Made Smarter Review (13) highlights the importance of enabling standardisation and interoperability between legacy devices, stating that the first step to beginning a digital transformation journey is to adapt legacy systems into the IIoT paradigm.
6.2 How should legacy devices be connected to a modern smart factory?

As previously mentioned in 4.3, IIRA recommends that legacy devices are connected to a smart factory platform through connectivity gateways, bridging one or more connectivity standards or protocols (41). These gateways would traditionally be physical pieces of hardware running bespoke computer programs that speak both the language of the legacy device and the language of the factory’s core connectivity standard. RAMI4.0 recommends a very similar concept to IIRA in that ‘Administration Shells’ should be deployed (10) to act as an intermediary between the legacy asset and the smart factory network, performing much the same function as the gateway concept from IIRA.

If a gateway-per-cell approach is being adopted, then the responsibility for parsing non-I4.0 compliant communications could be delegated to a ‘collect’ layer in the cell gateway capable of running Docker containers. This layer would be responsible for interpreting and understanding communication from legacy devices and converting them to the common connectivity standard of the factory, albeit in a more centralised manner. By utilising a container repository for the legacy gateways, the architecture remains manageable and scalable. The vast majority of legacy equipment and the protocols that they use to communicate were often designed with little to no security in mind, therefore by adopting a containerised strategy for gateway deployment the communications and data can be sandboxed to only the adapter allocated to handle the traffic for that specific piece of equipment.

6.3 Can legacy machine integration be achieved on a budget?

The digitisation of legacy machines does not need to be expensive. Although there are benefits of adopting industry-grade digital systems such as robustness and support, the real value of I4.0 is the data, not how it is collected. In 2018, the AMRC published a report titled ‘Digitalisation of Legacy Machine Tools’ (42) to demonstrate how IIoT can be made available to small to medium-sized enterprises (SMEs) via the adoption of low-cost sensing and cloud capabilities. The report outlines how equipment from circa 1956 can be retrofitted with sensing equipment for no more than £500 to catapult it firmly into the I4.0 paradigm. The report also documents how the cost of digitising legacy machines can scale with the fidelity of information required by demonstrating a high-fidelity data capture retrofit for a Bridgeport Turret Mill (circa 1980) for less than £5,000.

Due to the open standard of OPC UA and some other industrial protocols, open-source libraries and software development kits (SDKs) exist that enable said industrial protocols to operate as effectively with low-cost, hobbyist devices as they do with industry grade solutions (43), ensuring that both I4.0 compliant and legacy equipment can communicate using the same standardised communication architecture as recommended by both RAMI4.0 and IIRA.

7 Security

With Allianz rating cybersecurity their number 1 corporate peril of 2020 (44) and the financial implications of breaches to cybersecurity becoming apparent via high-profile incidents (45), all modern smart factory or IIoT solutions must have robust cybersecurity measures embedded in their foundations as standard.

7.1 How can data be kept safe in transit?

Although in security management it can generally be assumed that 100% security is simply not possible (20), steps can be taken to significantly reduce the chance and impact of a cybersecurity attack. Traditionally, in a fully connected environment, cyberattacks can have more widespread impact and could be more difficult to protect against due to the number of connection points and their distributed nature. By adhering to and adopting a standardised framework such as RAMI4.0 or IIRA, which suggest using a standardised communication architecture, the attack vectors of CPMS are reduced to only a single set of moving parts, resulting in increased agility and higher efficiency for maintenance, updates and monitoring.
As discussed in 3.2, OPC UA is currently one of the recommended protocols for factory communications in the smart factory, in part due to the robust and proven security feature set it possesses. OPC UA is secure by design and comprises of both the CIA and AAA triad framework concepts as follows:

- **Confidentiality:** Messages are encrypted at the transport layer
- **Integrity:** Messages are signed at the transport layer
- **Availability:** Messages are restricted in size and no security-related codes are returned
- **Authentication:** Authentication via username and password or X.509 certificate on the application layer
- **Authorisation:** Address space is controlled by access rights, permissions and roles.
- **Accountability:** Auditable

It is for the security reasons listed above that OPC UA is recommended by the IIC (9) as one of the protocols to use in industrial communications over other eligible options such as HTTP/REST. IIRA (9) explains how OPC UA and MQTT are both equally desirable and fit for purpose in factory communications other than in the security domain where OPC UA is the clear winner due to in-built security features. This is primarily due to MQTT having no encryption by default making the sending of plaintext MQTT payloads over the internet unacceptable for most applications without the addition of payload encryption.

In addition to the OPC UA specification being secure-by-design, the OPC Foundation publishes materials outlining the best practices for the secure use and implementation of OPC UA (29) in addition to maintaining a comprehensive bulletin board of known security issues that impact the software that it maintains or distributes (46). It is for this reason that the German Federal Office for Information Security found it to have an extremely high level of security in contrast to many other industrial protocols during an independent audit (47).

### 7.2 How can data at the edge be protected?

As distributed architectures such as RAMI4.0 are adopted in factories all over the world more mission-critical functionality and business logic is pushed to the edge tier where emerging, sometimes unproven, technologies are deployed to leverage the benefit of IIoT (48). It is these layers holding an increasing amount of valuable information where new and advanced cybersecurity threats such as ransomware attacks are targeting. In addition to adopting perimeter best practices to reduce the opportunity of a malicious third party gaining physical access to the machine, countermeasures should be deployed such as implementing edge-based network monitoring solutions, deploying a next-generation firewall (48) and reducing the amount of data stored on the edge device (49).

Novel methods such as certificateless encryption and decryption are being developed (50) that may improve the security of data at the edge by encrypting it at source and advances in distributed ledger and blockchain technologies (51) (52) in manufacturing could be utilised to ensure data is integral and has not been intercepted or tampered with.

These developments are all the more important when considering MQTT as an industrial communication protocol. MQTT does not prescribe, nor provide, mechanisms for secure communications as standard but that does not mean they cannot be implemented. Transport Layer Security (TLS) is advised for MQTT connections between client and server, this prevents eavesdropping and man-in-the-middle style attacks, as the communication channel is secure, and the data encrypted in transit. However, the data is not secured at rest, either on the client or the server. For this reason, encryption of the MQTT payload should be implemented at the source to provide end-to-end encryption, reducing the risk of data being manipulated or forged by a malicious actor at any stage of the communication.

There are two forms of encryption, Asymmetric key and Symmetric key, with Asymmetric generally deemed the more secure due to the use of separate keys for encrypting and decrypting data. However, this added complexity adds
computational overheads each time data is encrypted/decrypted. For this reason, especially when considering limited-resource hardware, symmetric keys are most commonly implemented for high-frequency encryption. As the same key is used to encrypt and decrypt data, a secure method of sharing this key is required. A common method for this is to use a Digital Envelope, whereby Asymmetric encryption is utilised to enable the sharing of a symmetric key, which is subsequently used to encrypt further communications.

To summarise, it is widely acknowledged amongst the cybersecurity community that it is a case of ‘when’ not ‘if’ a cyber-attack happens and the key task is to ensure that a solution has the relevant tools in place to detect the breach and mitigate the impact of the attack. By implementing a cybersecurity strategy (53) and ensuring that key personnel can identify and classify a breach the impact of cyberattacks can be significantly mitigated (52).
Conclusion

Although the IIoT landscape is fast-paced and dynamic, there appears to be an emerging trend of best practices and techniques being adopted throughout the latest literature.

The two mainstream architecture frameworks (IIRA and RAMI4.0) are advocates of both MQTT and OPC UA technologies to standardise communications on the shop floor.

Upon reviewing the latest literature, it is evident that existing established technologies such as OPC UA are still widely adopted in industry and are continuing to be supported by the foundations that maintain and update them. OPC has been a standard in manufacturing for decades due to its security, standardisation and robust data modelling and appears to continue to dominate factory floors as I4.0 technologies are rolled out across factories globally.

Although OPC UA appears to be the current connectivity technology of choice, frameworks such as Sparkplug are making some more forward-thinking organisations reach for lighter-weight solutions such as MQTT as a replacement for traditional client-server, poll-response architectures to benefit from significantly reduced network traffic and better compatibility with IT systems. The recognition of MQTT as a communication standard is further solidified by the recommendation of the technology in both IIRA and RAMI4.0 specifications and the compatible offerings of market leaders such as Siemens.

Device and communication security and integrity should be considered as part of the foundation of a chosen communication protocol, in addition to legacy device connectivity, maintenance overhead and futureproofing by ensuring compatibility with cloud and distributed architectures.
Appendix A – Culture

The research surrounding the social, economic and environmental impacts of Industry 4.0 and its implementation is still a comparably young research field (52), however, it is widely understood that a smart factory does not translate to a ‘dark factory’, with people still expected to be key to operations. The development of new technologies that enable IIoT is causing the personnel roles in the factory to also require evolution and realignment to support new processes and capabilities (53).

For these new roles to be applicable in the IIoT world the personnel holding these roles must possess, amongst other key skills, a fundamental understanding of ICT and ICT security (5). To ensure that key personnel in these roles have the skills required, it is advised that an agile, long-term approach be adopted to bridge current cross-area skills gaps and that employee skills matrices be developed that outline current skillsets and aspiring future skills for employees (31).

Upskilling programs such as presentations, workshops, training courses and working groups are a proven way to bridge the skills gap (31) and digital information platforms such as blogs and message boards are a good way to keep employees informed (5).
Terms and Definitions

AI – Artificial Intelligence
AMRC – Advanced Manufacturing Research Centre
COA – Cost Of Ownership
CPMS – Cyber-Physical Manufacturing System
I4.0 – Industry 4.0
IIC - The Industrial Internet Consortium
IIIRA – Industrial Internet Reference Architecture
IloT – Industrial Internet of Things
IT – Information Technology
MQTT – MQ Telemetry Transport
OPC UA – OPC Unified Architecture
OT – Operational Technology
RAMI4.0 - Reference Architectural Model Industrie 4.0
SDK – Software Development Kit
SDN – Software Defined Networking
SME – Small to Medium-Sized Enterprise
TLS – Transport Layer Security
TSN – Time-Sensitive Networking
References


38. RACONTEUR. *IoT FOR BUSINESS.* s.l. : RACONTEUR, 2018. #0564.


