Challenges of Implementing a Unified Namespace in the Life Sciences

This paper describes the requirements for an ideal implementation of a Unified Namespace (UNS). It further discusses the MQTT communication standard as a preferred core communication infrastructure, while shedding light on limitations and benefits of the OPC standard. As most 3.0 factories will not meet the minimum requirements for an ideal UNS, a hybrid-UNS approach is proposed along with a hypothetical example.

ABOUT THE AUTHORS

Amy Williams is an automation engineer experienced in on-site automation support and technical leadership for distributed control systems integration projects in life sciences. She is also a certified Industry 4.0 professional from 4.0 Solutions. Amy has been working with Skellig since 2020.

Purnendu Saptarshi is a senior automation engineer with over a decade of experience in designing and implementing process control systems in life sciences. Additionally, he has experience integrating equipment with SCADA systems. He is a certified Industry 4.0 professional from 4.0 Solutions. Purnendu has been working with Skellig since 2014.

The ideal UNS implementation is within reach

Implementing the Unified Namespace requires network and communication architecture that can seamlessly transmit data across the automation stack. All available data needed for operation should be accessible to all nodes regardless of location in the functional area. Nodes are any points in the communication infrastructure that produce or consume data. These nodes can range from field devices, such as temperature probes, to quality personnel consuming data for analysis.

In the effort to bring the benefits of an Industry 4.0 architecture to bear, it is important to only use software and hardware that support the four key requirements of a successful UNS: (1) being edge-driven, (2) utilizing report by exception, (3) being lightweight, and (4) utilizing open architecture.[1]

To evaluate if equipment or a software program meets these four key requirements, compatibility with data standards must be determined. A data standard is the “rulebook” for how data should be formatted, transported, or manipulated.

Two standards which inherently support UNS architecture are MQTT and MQTT with Sparkplug B

MQTT is an open-source, lightweight data standard that requires data to be transferred by publish and subscribe (PUB/SUB) messaging via report by exception over TCP/IP.

The MQTT standard was first developed in 1999 as a message protocol for oil and gas SCADA systems. At the time, satellite links were used to transmit data. This type of network was expensive and had low bandwidth (or, rates of data transfer). MQTT developers, Andy Stanford-Clark and Arlen Nipper, recognized the power of data for the oil and gas industry. They addressed the need for consistent, reliable, and affordable data transmission by inventing MQTT.

The MQTT standard was designed to be as lean and flexible as possible while maintaining both reliability and security. To achieve this, the MQTT standard does not require a specific format for topic naming, or the data being transmitted (the payload). MQTT also does not enforce security requirements or status monitoring. It can be configured to require login information and utilize heartbeat monitoring for status. These are optional as they increase data package size, consuming more bandwidth.

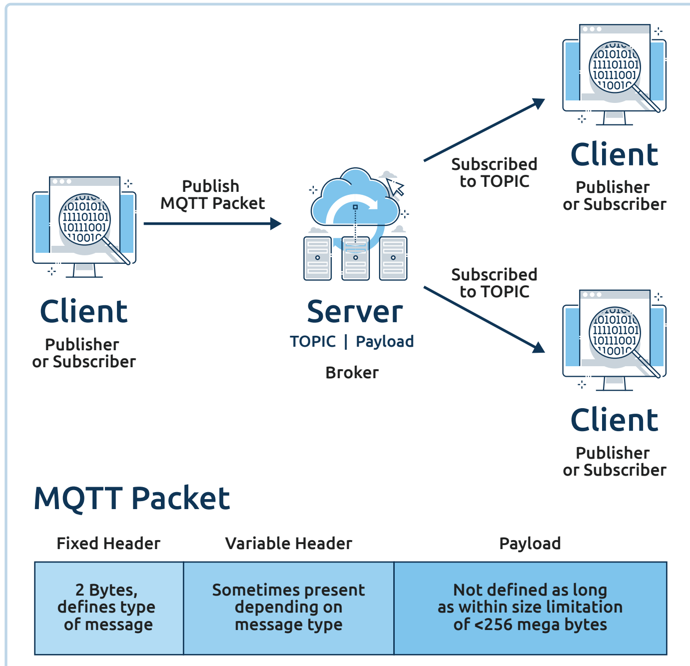

Figure 1: Overview of MQTT. **Note:** A client can both publish and subscribe to different topics simultaneously.

In 2019, a new standard called Sparkplug B was introduced. Sparkplug B is an extension of MQTT version 3.3 based on end-user feedback to improve industrial application. This standard is tailored for plant floor devices and SCADA communications.

A few features of Sparkplug B include specific definitions of topic naming and payloads. By standardizing hierarchy and data format, interoperability (the ability to easily transmit data across different systems) increases. Sparkplug B also requires state monitoring with the introduction of birth and death certificates.

In addition, this standard allows compression of data to maximize bandwidth efficiency and enables discoverability—the ability to publish new tags automatically without requiring manual configuration. This is how "self-aware" SCADA systems can be implemented.

Since Sparkplug B is an extension of MQTT, any broker that supports Sparkplug B will also support MQTT 3.3. This compatibility is useful as enterprise implementation often uses both Sparkplug B and plain MQTT.

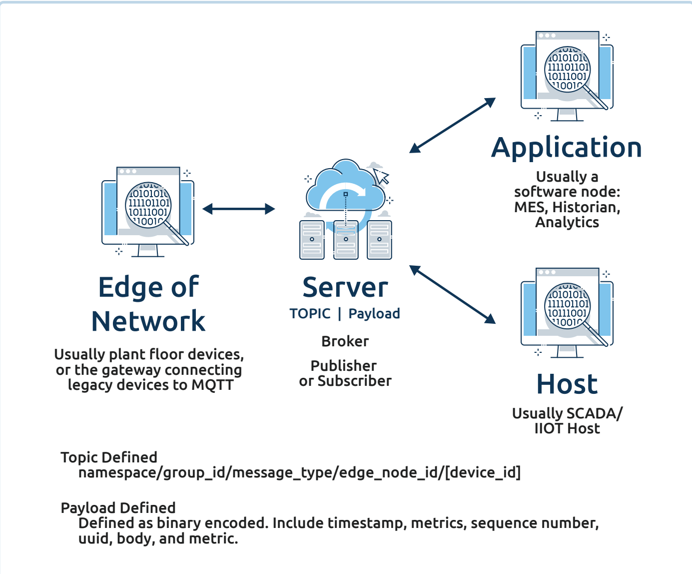

Figure 2: Overview of MQTT with Sparkplug B. **Note:** Sparkplug B is MQTT with additional definitions implemented to increase standardization and interoperability. The standard also recognizes that clients will have different needs and defines specific roles for clients as outlined in the figure from left to right: Edge of Network client, Application client, and Host client.

Sparkplug B is best used at the edge, with plain MQTT for higher levels of the enterprise. Plain MQTT is less restrictive on topic naming conventions, allowing it to better handle more complex transactional data communications than Sparkplug B.

Another common standard, OPC-UA, can also be configured to meet UNS requirements.[2] This standard is maintained by the OPC Foundation and has been widely implemented across 3.0 factories. OPC-UA can be beneficial on the plant floor. However, when choosing the protocol for UNS enterprise data transfer, MQTT is preferred since OPC-UA does not guarantee compliance with the four key requirements. OPC-UA based solutions can comply, but it's not a guarantee they will.

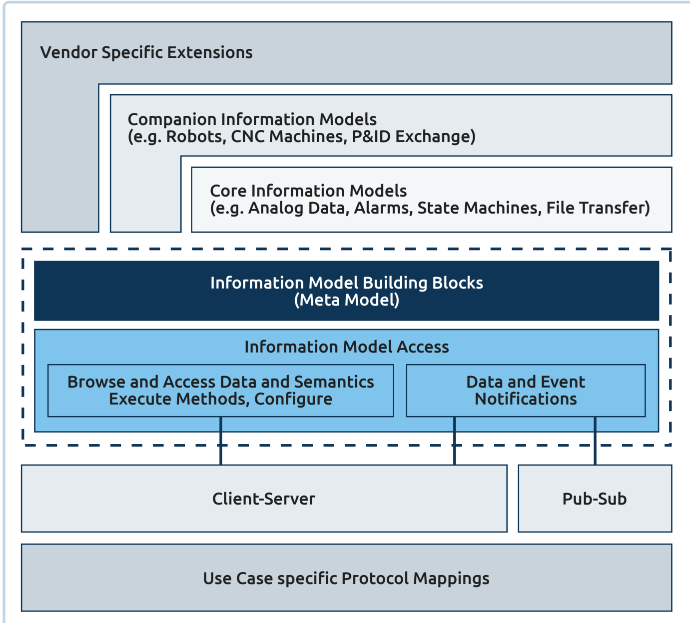

Figure 3: Overview of OPC-UA (2019, OPC Foundation). **Note:** The OPC-UA standard is complex and can be compatible with a multitude of other data standards and protocols, including Client-Server or Pub-Sub.

Compliance allows previously nonviable data to be collected, dispersed, and analyzed across an enterprise in a financially feasible way. Business architecture that uses heavy protocols, discrete connections, or polling generates significant engineering and hardware costs that prohibit sufficient data collection. Unlike MQTT, OPC-UA can be configured to not be industry 4.0 compatible. OPC-UA has a variety of elements and numerous companion specifications or custom vendor information models which instruct how to support a range of functionality. In addition, various software applications and hardware components that are OPC-certified do not have to comply with all OPC standard requirements. Because OPC-UA is so versatile, it is not always lightweight or report by exception. This decreases interoperability, since one OPC-UA solution can be polling, while another is PUB/SUB.

Therefore, the ideal implementation of a UNS is one that uses MQTT-compatible solutions across the enterprise. If needed, existing installations with different communication standards/protocols (e.g. Foundation Fieldbus, Profibus, and/or OPC DA/UA) should be translated to these open source protocols for enterprise communication.

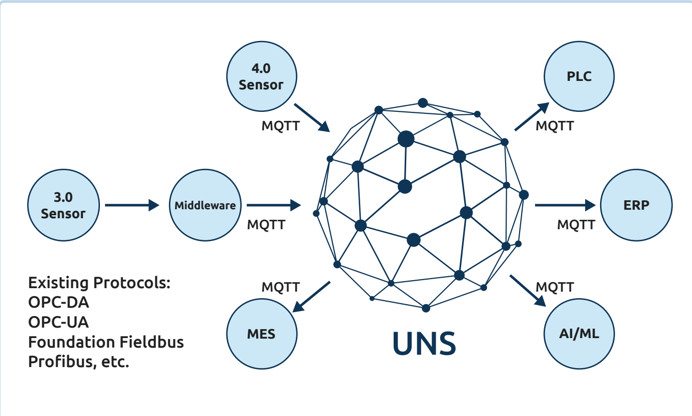

Figure 4: Overview of Ideal UNS use of data protocols.

Once the ideal implementation of the UNS is achieved, Industry 4.0 tools are immediately unlocked

Current real-world applications of UNS architecture built per the outlined requirements, and which support Industry 4.0 benefits include:

- MES systems that can automatically load recipes with the correct recipe, batch ID, formula, and equipment
- Electronic SOPs/MBRs/Logbooks that automatically transfer data between shifts
- Predictive maintenance that can generate work orders in the ERP or CMMS before equipment breaks
- AI that can suggest real-time parameter adjustments to improve yield before the batch is completed
- ERP schedules which can update automatically based on real-time floor data
- Self-aware UNS that can auto-configure SCADA automation upon detection of new data points, drastically reducing automation engineering design time and cost
- Machine Learning algorithms utilizing Big Data analytics that assess health of the overall business to improve efficiency, increase productivity, and reduce waste – in short, "do more with less"

These solutions immediately leverage the advantages of Industry 4.0 enabled by the underpinning UNS architecture, bringing improvements to quality and productivity.

However, successfully implementing a UNS is not the end of an overall digital transformation strategy

The purpose of the UNS is to provide the required infrastructure to utilize any future 4.0 tools. As these technologies are applied, the business will gain more insight, spurring new goals and needs. Therefore, the company will have to remain vigilant and continue to evaluate and implement new technologies.

In Life Sciences, this will also require proper vetting of any new tools for compliance. The FDA's Advancement of Emerging Technology Applications for Pharmaceutical Innovation and Modernization Guidance acknowledges that FDA reviewers will need to adapt to the use of novel technologies.[3] To promote this advancement, the FDA created the Emerging Technology Program with the goal that, "based on experience gained during the program, FDA intends to develop guidance and standards, as necessary, on emerging technologies and approaches to encourage and facilitate the innovation and modernization in pharmaceutical industry."[3]

As industry and regulations evolve, it is imperative that Life Sciences manufacturers prioritize involving QA early for digitization strategies to ensure compliance. Companies should also be flexible about modifying SOPs and procedures as needed to accommodate new technologies. Moreover, companies must prioritize providing employees proper training to gain the technical skills to adapt to new tools.

If a 3.0 factory has legacy technology which does not meet the UNS four key requirements, it can still succeed in Industry 4.0 with a Hybrid UNS

In most cases, Industry 3.0 infrastructure will not meet the four key requirements of the ideal UNS. Many existing factories have highly integrated preferred provider stacks from the plant floor to supervisory levels that utilize discrete connections via OPC-UA.

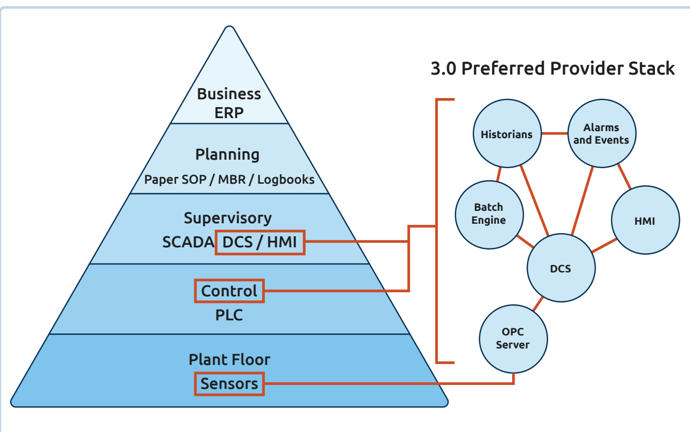

Figure 5: Hypothetical 3.0 automation stack with preferred provider solution.

Large amounts of capital were spent purchasing licenses required to reach their current levels of connectivity. It is not feasible to re-route all existing connections through the UNS.

It would be a mistake to disregard UNS as an option because existing connectivity is "sufficient." As discussed in Skellig's second white paper, this is not the case.

The traditional 3.0 preferred provider stack will not provide the data connectivity needed to benefit from Industry 4.0 productivity. First, it relies on discrete connections that silo data and do not easily scale, which also makes mapping for context more difficult. Second, maintaining these connections often requires costly continual purchasing of licenses. Moreover, these systems are usually built on closed software platforms that only expose data to other preferred vendors' systems. This prevents selecting best-in-class solutions and inhibits implementation of a self-aware ecosystem.

In this hypothetical case, a Hybrid UNS solution should be implemented as a first step towards transforming to true Industry 4.0 architecture. This solution would leverage existing connections for direct control and selectively publish data to the UNS. It is important to understand that a UNS will not eliminate all discrete connections. There will be cases where using discrete connections makes sense. The exact details of how to apply a UNS will vary from site to site.

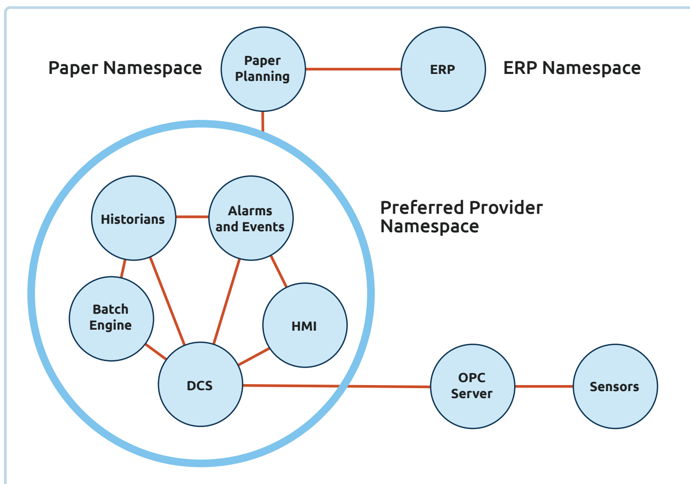

Figure 6: Hypothetical 3.0 stack with a closely integrated preferred provider stack as a node view. **Note:** This hypothetical assumes only OPC was used on the plant floor. However, this scenario still applies if a combination of OPC, Foundation Fieldbus, Profibus, and HART were used in conjunction from sensors to the preferred provider namespace.

A Hybrid UNS approach can leverage existing solutions and revolutionize 3.0 facilities

In many cases, plant floor data should be published as read only to the UNS so it can be consumed elsewhere for analysis. Artificial Intelligence (AI) and Machine Learning (ML) applications require plant floor data for models.

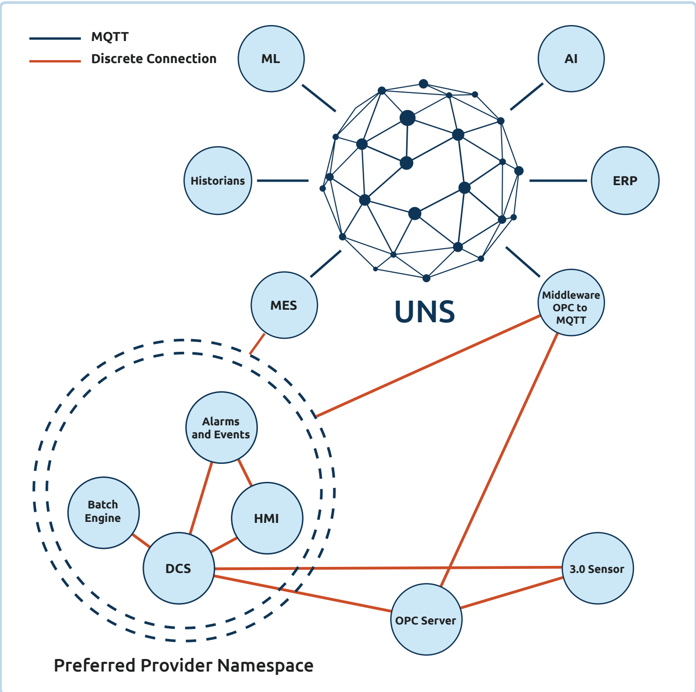

Figure 7: Hypothetical Hybrid UNS. **Note:** For this hypothetical, it is assumed 3.0 Sensors can be OPC-UA, Foundation Fieldbus, or Profibus.

As mentioned, MQTT is the preferred protocol for transporting data to and from the UNS. However, OPC can still be useful or even beneficial for plant floor controls. For a hybrid approach, OPC should be left as-is for plant floor connections. Middleware can be used to convert OPC into MQTT for easy distribution by the UNS. The main issue with OPC-UA is that legacy plant floor configurations are usually heavier than MQTT. To distribute plant floor data over OPC-UA across the business would most likely require either throttling or additional configuration to make it more lightweight. Therefore, it should be converted to MQTT prior to distribution.

The next step is to evaluate what data from the Supervisory and Control systems should be published as read only to the UNS. This varies greatly between clients based on their needs and current connectivity. A general rule of thumb is to connect data that will be consumed elsewhere. Data used for direct control, such as execution time of a step, most likely would not need to be added to the UNS, while alarms, events, and setpoints would be.

MES should be implemented to reduce paper usage and increase data accessibility. MES will most likely require the use of discrete connections to properly integrate with legacy control software. An evaluation determining what to connect from MES to UNS would be required. Connections that provide new data, such as calculations or events, should be connected to the UNS.

Once the selected data is published into the UNS, historians would be configured to subscribe to those topics in the UNS. Again, it is worth configuring historian communication to the UNS as its own connection to avoid needing discrete connections moving forward.

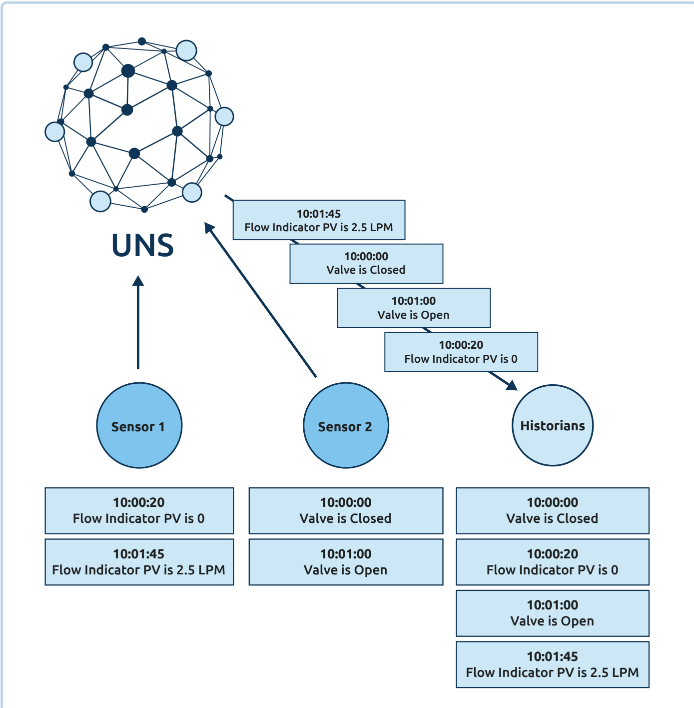

Figure 8: Visualization of data packets historized based on timestamp, not order of arrival to the historian.

A common misconception is that utilizing a historian within a UNS could cause inaccurate event logging. The MQTT data protocol requires the message to contain a timestamp for each data point generation. Regardless of the order of messages reaching the historian, the historian will log events in the correct chronological order based off the message's timestamp.

Once the hybrid UNS is configured, the once-3.0 factory will have the minimum connectivity needed to implement 4.0 solutions.

It is important to recognize that each factory installation is unique, therefore each factory will require its own evaluation. The path forward may look different from the sample case described above. The purpose of this hypothetical scenario is to illustrate the types of challenges and thinking that will be required for implementing a UNS.

The goal of this approach is to include as many companies as possible in the fourth industrial revolution, or Industry 4.0. Existing 3.0 architecture can provide local interconnectivity, but this is not sufficient for 4.0 technology.

A hybrid UNS approach provides a practical entry point to Industry 4.0 while respecting cost and time boundaries

This does mean that once a UNS is implemented the factory must only use solutions which meet the UNS four key requirements: open architecture, lightweight, report by exception, and edge-driven data.

The technology exists to allow anyone, including 3.0 factories, to take advantage of the benefits offered by Industry 4.0. These tools can be used to achieve improvements in scalability, data collection, and efficiency that were previously impossible or very expensive within a 3.0 framework.

Life Sciences manufacturers can achieve a return on investment measured in months, not years

By embracing the principles of a 4.0 architecture, Life Sciences manufacturers have a lot to gain—specifically in the areas of implementation, agility, and efficiency. However, it can be difficult to predict a return on investment because it is tied to specific use cases, not the infrastructure itself.

Implementation is often the simplest use case to measure, though the value may not be visible at small scale. Proof of concept projects to develop necessary architecture and connect a few pieces of equipment can have a seemingly prohibitive cost, on the order of $100,000s. The benefit comes from the integration and validation costs for each additional piece of equipment being reduced by 80% (1 integration instead of 5). In this case, ROI can be measured in number of pieces of equipment instead of time.

One primary packaging manufacturer was only able to integrate equipment on time during the COVID-19 pandemic because of the unified namespace architecture in which they invested. In an effort to scale their process by more than 10,000%, the only way to meet deadlines was to perform the integration in this way. In this case, ROI was not measured in dollars but in time to market.

Personnel performed SAT and software IOQ on a new machine within three days of arriving on-site instead of requiring two weeks or more. Aligning the validation strategy around the UNS infrastructure was key to meeting these gains. Furthermore, the investment made in UNS infrastructure made adding functionality such as OEE and SPC a matter of hours and days instead of weeks and months.

For most Life Sciences manufacturers, OEE is not a critical KPI and SPC has no practical use. In biologics, batch titer and filtration efficiency are much more important—and much more difficult to measure. They often involve multivariate analysis to optimize. A Unified Namespace allows these processes to be connected to various AI and ML platforms with context. Unlocking the power of these platforms is critical to continuous improvement in manufacturing just as it has revolutionized financial services, tech, and retail.

Companies keep the details of their AI investments to themselves. As it is a competitive advantage, it can be difficult to measure the types of returns expected from these investments. The image below from a study by McKinsey & Co. shows cash flow by AI adoption cohort. The advantages of being an early adopter clearly accelerate over time.

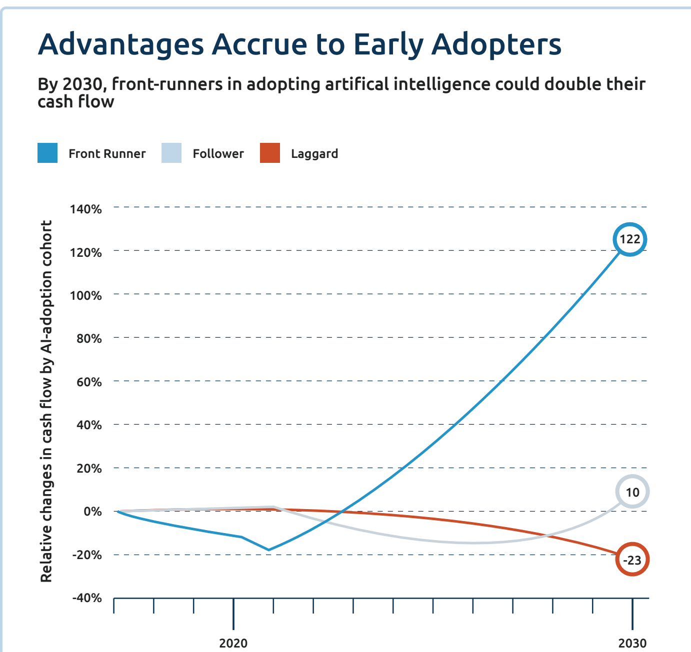

Figure 9: Results from a 2018 McKinsey & Co. study showing the financial advantages of early AI adoption.[4]

It is, of course, possible to deploy Artificial Intelligence and Machine Learning without a Unified Namespace. However, empowering people to do the most value-added work is one of the biggest advantages of Industry 4.0. This image from an Anaconda Study on Data Science[5] shows that data scientists spend 45% of their time cleansing and loading data. A UNS with appropriate software substantially reduces this time, allowing teams to spend more time training, scoring, and developing models, as well as working on other value-added tasks. Measuring ROI here could begin with looking at the reduction in non-value-added work, but the true value lies within the insights provided within time previously spent managing data.

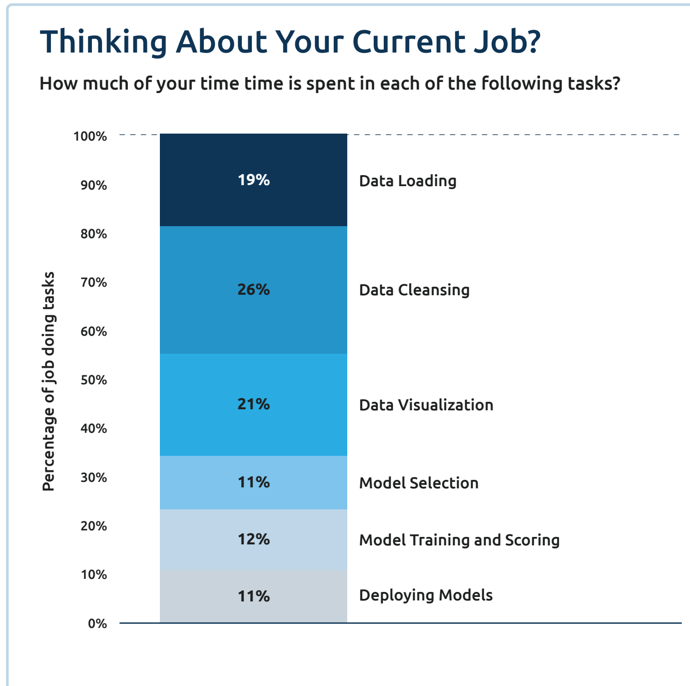

Figure 10: Results from a 2020 study showing how data scientists spend most of their time.[5] **Note:** For most respondents, data management tasks still consume a disproportionate amount of work time.

Technology alone cannot result in a successful 4.0 company. At the end of the day, companies are made of people. Success or failure are entirely dependent on the mindset of people within the organization. The number one challenge for leaders who wish to digitally transform will be addressing the "People Problem." Further discussion on what this means and methods to overcome it will be discussed in a future paper of Skellig's ongoing Industry 4.0 series.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Walker Reynolds. Walker Reynolds is the CEO and founder of 4.0 Solutions, and IIOT University. He has pioneered UNS implementation for manufacturing and played an influential role in growing the 4.0 movement. The technical foundation of the UNS guidelines discussed in this paper leverages the teachings of Walker Reynolds. However, any extrapolation to the life sciences is based off FDA guidelines, ISPE trends, and years of firsthand experience working as a systems integrator and site support within the industry. 4.0 Solutions did not fund or author this paper and are not responsible for the content or claims.

CITATIONS (APA)

1. 4.0 Solutions. (2020, November 6). Why your Industry 4.0 Applications are NOT scaling. YouTube. https://www.youtube.com/watch?v=8psBOqInulc

2. OPC Foundation. (2019, September 26). Unified architecture. OPC Foundation. Retrieved August 3, 2022, from https://opcfoundation.org/about/opc-technologies/opc-ua/

3. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER). (2017). Advancement of Emerging Technology Applications for Pharmaceutical Innovation and Modernization Guidance for Industry.

4. Advantages Accrue to Early Adopters. McKinsey & Co., September 2018.

5. Anaconda | State of Data Science 2020. (2020). Anaconda. Retrieved September 14, 2022, from https://www.anaconda.com/state-of-data-science-2020?utm_medium=press