SKELLIG

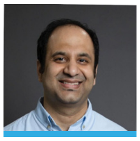
4.0 WHITE PAPERS: AN EDUCATIONAL SERIES

CONNECTING EQUIPMENT:

The First Step Towards Implementing an Industry 4.0 Structure in a Lab Environment

In today's competitive business landscape, efficient workflows, agility, and a competitive edge are crucial. Digitizing laboratory equipment and processes is the first step toward achieving these goals. Automating data collection offers benefits like error reduction, data integrity, traceability, and operational efficiency. However, labs often face challenges accessing and leveraging data trapped in silos due to outdated egress options, vendor-specific software packages, or limited data extraction capabilities. This white paper discusses an approach to integrate such equipment, creating a cohesive infrastructure that aims to connect, collect, store, and analyze data seamlessly, eliminating point-to-point integrations, saving time, and reducing costs.

ABOUT THE AUTHOR



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.

Challenges Faced in the Current Scenario

INTEROPERABILITY

One of the significant challenges in achieving seamless data integration arises from the diverse array of equipment, software, and data formats used by different Original Equipment Manufacturers (OEMs). Each OEM may have its own proprietary formats, communication protocols, and data management practices, making it difficult to establish effective interoperability between different systems. Standardizing data formats, communication protocols, and data management practices is essential for enhancing digitization efforts.

DATA QUALITY AND CONSISTENCY

Inaccurate, incomplete, or inconsistent data hinders the reliability and usefulness of the digitized information, undermining the integrity of conclusions. Ensuring data quality is a critical aspect of any digitization initiative.

DATA SECURITY AND PRIVACY

Digitizing lab data raises concerns about data security and privacy. Labs and OEMs need to implement robust data security measures to protect sensitive information.

COST AND RESOURCES

The process of digitizing lab data, integrating different systems, and maintaining these systems can be resource-intensive. Smaller labs, or those with limited budgets, may struggle to invest in the necessary infrastructure and personnel.

CHANGE MANAGEMENT

Transitioning from traditional lab practices to digitized workflows may require significant organizational and cultural changes. This can be a challenge for labs that are resistant to change or have entrenched workflows.

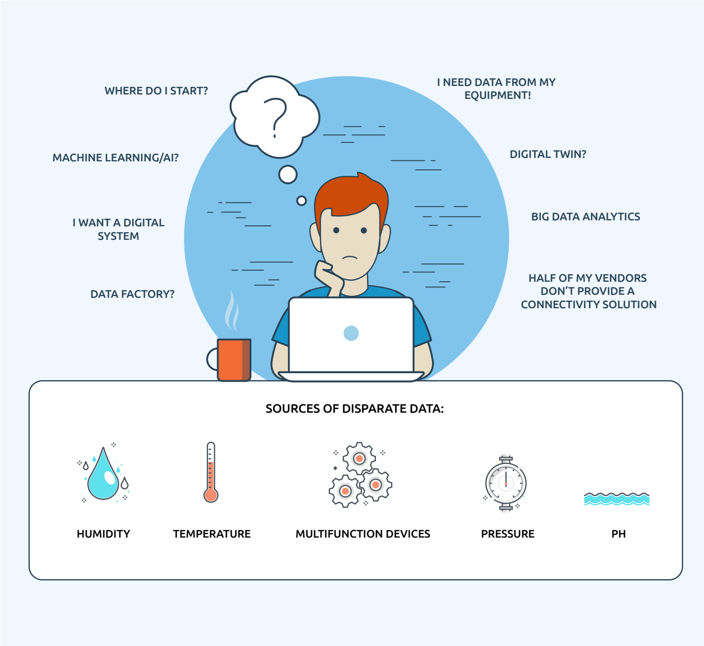


Figure 1: Various challenges in lab data integration often lead to confusion about where to begin.

Solution to the Challenges

The ensuing discourse presents the recommended steps for effectively planning and executing a lab equipment integration project. By highlighting the intricate considerations and challenges, it aims to provide a thorough understanding of the problems that may arise during the project. The outlined steps offer a well-structured and successful integration process, taking into account the multifaceted aspects involved in seamlessly integrating lab equipment.

LABORATORY EQUIPMENT SURVEY

It is important to perform a survey of the equipment prior to embarking on the journey of connecting them. This analysis involves gathering data for all instrument installations, including the device manufacturer, software/ hardware versions, and data ingress/egress capabilities. This information can typically be obtained through an export of the existing asset management system, if available.

CATEGORIZATION OF EQUIPMENT

After collecting the equipment information, they should be categorized depending on their integrability using a complexity-based approach with the following suggested guidance:

CLASS 0

LOW COMPLEXITY

Easy integration

- Equipment has vendor provided integration methods
- Data egress is available over common industrial protocols readily consumable by a wide selection of middleware softwares such as OPC DA[1], OPC UA[2], MQTT[3]

CLASS 1

MEDIUM COMPLEXITY

Integration requiring minor custom code

- Integration needs minor custom code, but data egress is possible with widely available software options for data processing and contextualization
- Examples include API integration, SQL/ODBC or another type of database integration
- This category also considers the availability of third party middleware for easy integration, with the added burden of applicable cost to the project

CLASS 2

MEDIUM COMPLEXITY

Custom scripting requirements

- Integration needs custom code, or application development for extracting data
- Examples include serial data interface development, custom drivers configuration that may be provided by third party middleware
- This category also includes equipment for which vendor provided integration options may be too costly and have closed architecture. In this case, the integrator may choose to develop custom interface solutions.

CLASS 3

EXTREME COMPLEXITY

Low probability of integration

- Equipment for which data egress is not provided by the vendor or is extremely impractical/costly
- The cost of a custom solution outweighs commercial off-the-shelf options

IMPLEMENTING A POC TO ESTIMATE EFFORT FOR LARGE SCALE INTEGRATION

Before undertaking a large-scale lab equipment integration project, it is often a good idea to conduct a "Proof of Concept" (POC) portion of the project to ensure that the proposed solution will meet the user's requirements. This is particularly important when integrating "first of kind" equipment with a central system, as there may be unforeseen challenges that need to be addressed. By conducting a POC, users can identify any potential issues early on, and work with their system integrator to address them before they become larger, more costly problems.

The insights gained from a POC can also be used to estimate the level of effort required for a scaled project. By conducting a smaller scale test first, users will gain valuable learnings that can be leveraged to accurately estimate the time and resources needed for a larger project. This can help avoid any unexpected delays or costs that could impact project success.

Often, the POC also shows cross functional stakeholders the value of the data obtained from equipment. This provides an early win for the integration efforts, which helps fuel future projects.

CHOOSING A VENDOR AGNOSTIC INTEGRATOR

The benefit of choosing a vendor agnostic system integrator is their ability to work in the best interests of the customer. Because these system integrators are not tied to any specific manufacturer or supplier, they are able to offer a more impartial evaluation of different solutions available in the market. This means that they can choose the most appropriate solution for the project's specific needs, rather than being limited to a specific brand or manufacturer.

This will ensure that the solution selected is tailored to the user's unique requirements, and that the lab equipment integration project is completed in a manner that is transparent and unbiased. Additionally, vendor agnostic system integrators may be better positioned to negotiate pricing and terms with multiple vendors, ultimately providing users with more competitive pricing and ensuring that they receive the best possible value for their investment.

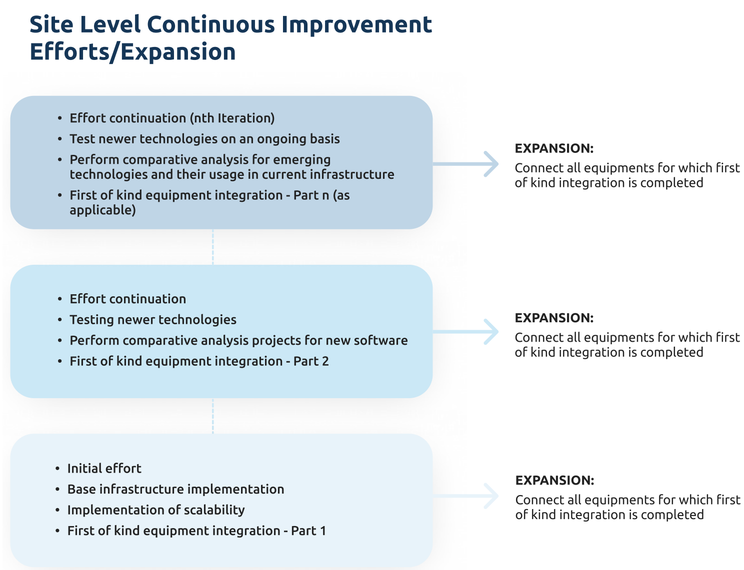


Figure 2: Implementing a Proof of Concept at a small scale before transitioning to a large scale effort.

IMPLEMENTING OPEN ARCHITECTURE

Open architecture is a software design approach that allows for the integration of various add-on products and components to work seamlessly with original implementation of software. The greatest advantage of open architecture is that it provides flexibility to choose from a variety of compatible products and components to build a customized solution that meets specific business needs.

"Open architecture framework gives you the power to choose from a wide variety of software and hardware solutions so you're never locked in and can gradually invest in your security profile. An open, unified solution can be used across multiple systems and locations - compatibility and interoperability are guaranteed." [4]

For data connectivity solutions, such an architecture allows users to benefit from a flexible and modular infrastructure that is not tied to a specific vendor or platform. This leads to increased efficiency, cost savings, and improved data management capabilities. Additionally, it provides opportunities for continuous improvement over time, as technology evolves, to better align with market needs.

Open architecture also facilitates interoperability and compatibility with multiple systems, resulting in better management and securing of data across different environments, allowing for more comprehensive and informed risk-based decisions during the design process.

SCALABILITY

Point-to-point integration is not scalable due to the complexities and inefficiencies it introduces. Custom connections between the network and each piece of equipment result in a growing web of complexities, making it challenging to manage and maintain as the number of devices increases.

Moreover, the lack of standardized approaches and the fragmentation of integration methods hinder expandability and upgrades. Data silos are formed, limiting data accessibility and collaboration. Additionally, the time-consuming nature of point-to-point integration increases deployment time and resource requirements.

To avoid these challenges, the project must consider the scalability of the implemented solution. This is highly important, as a POC is executed to prove equipment integration capability. Emphasis should be given to develop modular solutions that are easily instantiable to integrate and maintain a similar type of equipment. A simple class-based approach should be taken wherever possible, to effectively implement and manage the integration environment, as more devices are added to the ecosystem. Prioritizing scalability in the design from the start leads to lower maintenance costs, better user experience, and increased agility.[5]

MANAGING THE SOFTWARE DEVELOPMENT LIFE CYCLE (SDLC)

"The software development life cycle (SDLC) is the process of planning, writing, and modifying software. It encompasses a set of procedures, methods, and techniques used in software development. When considering a vendor's capability for managing the software development life cycle, it is important to assess their ability to provide end-to-end services that cover all stages of the SDLC. This includes planning, design, development, testing, deployment, and maintenance. This can be achieved by evaluating the vendor's track record of successful project deliveries, as well as their expertise in the latest software development methodologies and technologies." [6]

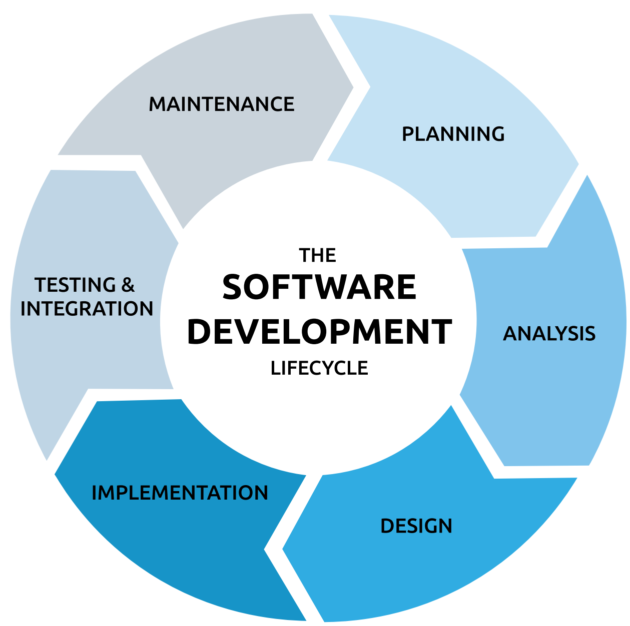


Figure 3: Software Development Lifecycle (SDLC) Process.[7]

In implementing an open architecture, it is critical that the chosen integrator is able to fully support the developed solution. Inevitably, changes happen throughout the lifecycle of the end equipment, often due to external factors such as OS upgrades, patching, vendor firmware upgrades, middleware upgrades, and other potential factors. The solution needs to be easily adaptable to the modified integration requirements, in order to maintain data integration requirements and performance of the implemented system.

Overall, selecting a vendor with a comprehensive understanding of the software development life cycle and a proven track record of delivering high-quality software solutions will ensure a successful partnership. Figure 3 provides visual representation of the SDLC process. The stages shown are always part of an integration project, with heavy emphasis on the maintenance phase.

DECIDING PROJECT SUCCESS CRITERIA

The following suggested success criteria for the integration efforts are essential to ensure the project produces specific, measurable results which are achievable, relevant, and time-bound.

These key points are baselines to be considered while executing an integration project:

- **Listing Equipment to be Integrated to Define the Project Scope:** List the equipment to be integrated (and their alternatives) with specific scope for the project. This helps in determining the end results when equipment is integrated with the data acquisition system, producing meaningful data.
- **Defining Data Contextualization as Close to the Equipment Layer as Possible:** It is always preferable to define data contextualization as close to the equipment layer as possible. One of the project's success criteria should be how well the data has been contextualized and how easily it is available for use.
- **Timebound Completion of the Project:** As with all projects, equipment should be integrated within specific timelines, and appropriate changes to the schedule should be made to reflect actual project progress. It is important to be aware that, in some rare cases, a piece of equipment may become non integrable due to various limitations. The effect of such changes over schedule should be minimized.
- **End User Involvement and Satisfaction:** End users are the most critical stakeholders of an integration project, and should be involved from the beginning, both for their equipment expertise and process input. The solution must significantly improve the users' day to day work, and provide ready access to all necessary data in a way that is convenient to them. Therefore, end user satisfaction is a crucial part of project success criteria.
- **Comprehensive Documentation:** Due to the nature of open architecture, there are a multitude of ways in which a solution can be implemented to obtain data from equipment. It is imperative to develop a comprehensive documentation package that, at minimum, details the following key points:
    1. *Network architecture – depicting overarching networking architecture, along with end device connections*
    2. *List of software packages used, including versions*
    3. *List of hardware components used (such as middleware/protocol converters, IIOT devices)*
    4. *Documentation around software configuration – data contextualization, custom driver configuration, third party software configuration details*
    5. *User interface details (if developed) – to provide a guideline for the end users on how to utilize acquired data, review, represent etc.*

THE VISION: HEADING TOWARDS UNIFIED NAMESPACE (UNS) IMPLEMETATION

Given the design considerations above, a Unified Namespace implementation is an advantageous solution.

Unified Namespace implementation allows for the creation of a single source of all data and information relating to a uniquely identified organizational element. Each element within the organization is uniquely identified within the UNS. Thus, the current state of the organization/business is captured within the UNS, which contains every element within the organization and its related data.

For more details on the benefits of a UNS implementation, please refer to Paper 2 of our white paper series, ***Unlocking Industry 4.0 Across Life Sciences Starts with the Uni ied Namespace (UNS).***[8]

In the current context, the uniquely-identified organizational elements are the individual lab equipment units. By connecting them within a UNS framework, it is possible to access data associated with each unit across the entire organization. A complete vision of such an implementation at a given laboratory is provided by the following infographic:

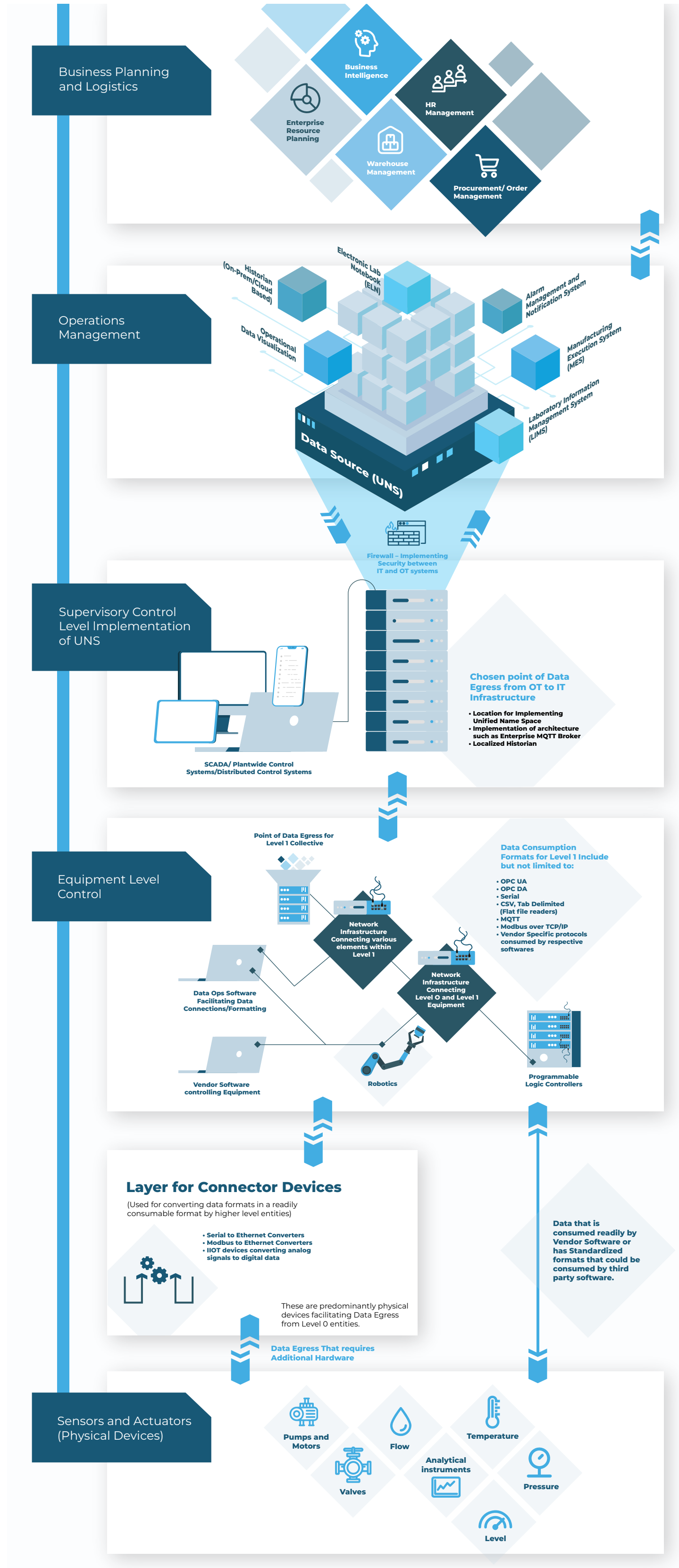


Figure 4: Visualization of comprehensive integration of lab equipment and associated data in a central infrastructure.

This concept can be further expanded by considering an individual lab as an element, thereby connecting multiple labs into a central infrastructure on a global scale. This emphasizes the importance of considering all aspects of the solution and project success criteria in order to demonstrate scalability at an ever expanding level.

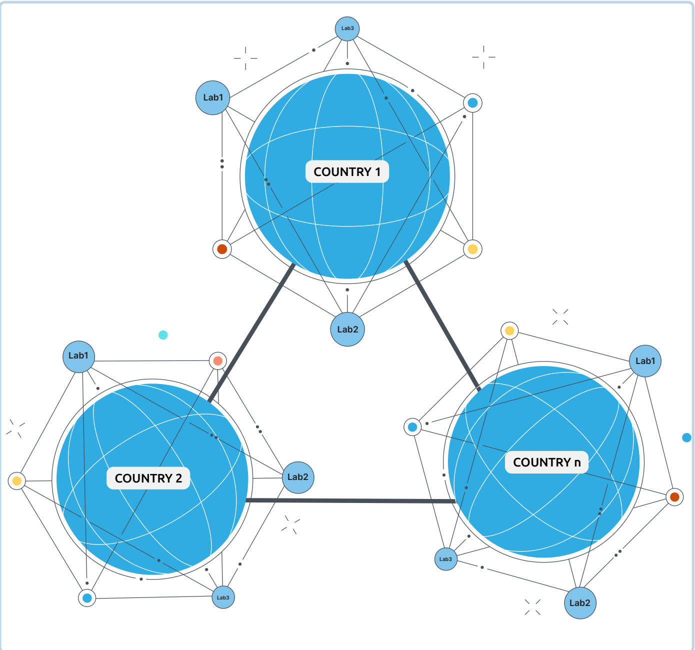


Figure 5: Visualization of scalability of lab equipment integration at a global level.

Conclusion

In conclusion, this paper provides a summary of the difficulties in gathering data from various OEM laboratory instruments. To get around these challenges, the article offers end users a step-by-step procedure for executing a project to consolidate data into a single platform for use in the automation stack hierarchy. Selecting a vendor-agnostic integrator that performs an equipment survey and offers alternative solutions is strongly advised to accomplish this goal.

To confirm the initial connectivity of the equipment with the centralized platform, it is also advised to carry out a proof of concept (POC) component as part of the larger project. This strategy offers an early success and a preview of the outcomes of the bigger project. During this smaller-scale effort, it is simpler to design and implement an open architecture solution, which removes compatibility restrictions for various solution platforms. Additionally, the solution's scalable design makes it easier and more convenient to integrate similar equipment types across the laboratory with significantly less time and effort, thereby lowering project costs.

Successful execution of an equipment integration project leads to digitization of the data, that can then be utilized in efforts toward streamlined workflows. It becomes easier to avail the advantages of a digital system, such as enhanced data quality and consistency, streamlined workflows, real-time monitoring of processes and analytics, and improved collaboration between cross functional teams. This invariably leads to better decision making based on factual and accurate data, reducing operational costs and improving efficiencies.

Citations (APA)

1. OPC Foundation. (2019, September 26). Classic. OPC Foundation. Retrieved May 15, 2023. https://opcfoundation.org/about/opc- technologies/opc-classic/

2. OPC Foundation. (2019, September 26). Unified architecture. OPC Foundation. Retrieved May 15, 2023. https://opcfoundation.org/ about/opc-technologies/opc-ua/

3. MQTT. MQTT: The Standard for IoT Messaging. (2022) Retrieved May 21, 2023. https://mqtt.org/

4. Genetec. (n.d.). Open architecture: the building blocks of a unified security platform. Retrieved May 21, 2023. https:// www.genetec.com/blog/products/open-architecture-the- building-blocks-of-a-unified-security-platform

5. Concepta. (n.d.). The importance of scalability in software design. Retrieved May 21, 2023. https://www.conceptatech.com/blog/ importance-of-scalability-in-software-design

6. What is the software development life cycle? SDLC explained. Coursera. (2023, June 16). Retrieved June 19, 2023. https:// www.coursera.org/articles/software-development-life-cycle

7. Tripathi, B. (2022, June 21). Software development client questionnaire - 10 questions to ask when developing software: Synotive. Synotive Technologies Blog. https://www.synotive.com/ blog/software-development-client-questionnaire

8. Saptarshi, P., & Williams, A. (2022). Unlocking Industry 4.0 across Life Sciences starts with the Unified Namespace (UNS) [White Paper]. Retrieved July 15, 2023 https://skellig.com/white-papers/