

NPL REPORT MS 59

**NPL'S DATA QUALITY FRAMEWORK AND ITS INTEGRATION WITH
HVMC: A PROPOSAL FOR ENHANCING NATIONAL PROGRAMMES**

**JOÃO GREGÓRIO
DANIEL POVEY**

FEBRUARY 2025

NPL's Data Quality Framework and its Integration with the HVMC:
A proposal for Enhancing National Programmes

João Gregório¹ and Daniel Povey²

¹Informatics, Data Science, National Physical Laboratory

²Manufacturing Metrology, Materials and Mechanical Metrology, National
Physical Laboratory

© NPL Management Limited, 2025

ISSN 1754-2960

<https://doi.org/10.47120/npl.MS59>

National Physical Laboratory
Hampton Road, Teddington, Middlesex, TW11 0LW

This work was funded by the UK Government's Department for Science, Innovation & Technology through the UK's National Measurement System programmes.

Extracts from this report may be reproduced provided the source is acknowledged and the extract is not taken out of context.

Approved on behalf of NPLML by
Ed Quinn, Group Leader.

CONTENTS

GLOSSARY/ABBREVIATIONS
EXECUTIVE SUMMARY

1 INTRODUCTION.....1

2 DATA QUALITY1

3 CERTIFICATION BY ANALYSIS3

4 MODEL-BASED ENTERPRISE.....4

5 DISCUSSION.....5

6 CONCLUSION7

7 AKNOWLEDGEMENTS7

8 REFERENCES.....7

GLOSSARY/ABBREVIATIONS

Term	Definition
ASME	American Society of Mechanical Engineers
BSI	British Standards Institution
CbA	Certification by Analysis
C-MBEE	Connected Model-Based Enterprise Environment
DT	Digital Twin
FAIR	Findable, Accessible, Interoperable, and Reusable
HVMC	High Value Manufacturing Catapult
IoT	Internet of Things
ISO	International Organization for Standardization
IUK	Innovate UK
KPI	Key Performance Indicator
MBD	Model-Based Definition
MBE	Model-Based Enterprise
ML	Machine Learning
NAFEMS	National Agency for Finite Element Methods and Standards
NPL	National Physical Laboratory
PMI	Product Manufacturing Information

EXECUTIVE SUMMARY

The manufacturing sector is undergoing a significant transformation driven by the adoption of advanced digital tools and technologies, collectively known as Industry 4.0. This shift emphasises the integration of cyber-physical systems, the Internet of Things (IoT), and cloud computing to create smart factories. Central to this transformation is the concept of data quality, which is essential for making informed decisions, ensuring operational efficiency, and maintaining product quality.

This report explores the integration of the National Physical Laboratory's (NPL) Data Quality Framework with the High Value Manufacturing Catapult (HVMC) and the wider catapult network. It highlights NPL's offerings and outlines pathways for future collaboration on national programmes. The report focuses on two key work groups: Certification by Analysis (CbA) and Model-Based Enterprise (MBE).

The Certification by Analysis (CbA) committee emphasises the use of digital tools to support and enhance traditional certification processes. It highlights the importance of data quality in virtual testing and certification, ensuring that data-driven decisions are reliable and traceable. The report discusses the development of a data-driven credibility assessment framework, which evaluates the dependability of sources, data, and methodologies, focusing on robustness and uncertainty for data quality and simulation confidence.

The Model-Based Enterprise (MBE) user group focuses on the integration of digital models and data management practices to create a seamless and traceable digital thread throughout the supply chain. The report examines the Connected Model-Based Enterprise Environment (C-MBEE) testbed programme, which aims to enhance digital supply chain integration through software interoperability analysis and data traceability. It discusses the development of vocabulary mapping software and the practical application of hashing algorithms to improve data integrity and traceability.

By adopting Model-Based Engineering (MBE) practices and leveraging frameworks like the C-MBEE testbed programme, manufacturers can significantly improve digital supply chain integration, enhance data traceability, and ensure the integrity of their data. These efforts align with the broader goals of digital transformation and competitive advantage in the manufacturing sector, highlighting the critical importance of robust data management and traceability in modern manufacturing environments [1].

This report provides a comprehensive understanding of how data quality frameworks can support the digital transformation of the manufacturing sector. It underscores the need for high-quality data to drive reliable and traceable decisions, ultimately contributing to the sector's long-term success and sustainability.

1 INTRODUCTION

Currently, the manufacturing sector is undergoing a significant transformation driven by the adoption of digital tools and technologies. This shift, often referred to as Industry 4.0, emphasises the integration of cyber-physical systems, the Internet of Things (IoT), and cloud computing to create smart factories. These advancements enable manufacturers to optimise processes, enhance productivity, and maintain a competitive edge in the global market.

To fully leverage these emergent digital technologies, high-quality data is essential for making informed decisions, ensuring operational efficiency, and maintaining product quality. As manufacturers increasingly rely on data-driven methods, the integrity, accuracy, and traceability of data become critical factors in their success. This report explores the integration of the National Physical Laboratory's (NPL) Data Quality Framework [2] with the High Value Manufacturing Catapult (HVMC) and the wider catapult network, highlighting NPL's offerings and outlining pathways for future collaboration on national programmes.

Two key work groups within this context are the Certification by Analysis (CbA) committee and the Model-Based Enterprise (MBE) user group. Both work groups are formed under the Smart Design Innovation Network (HVMC, BSI, NAFEMS, NPL, and the Turing Institute). The CbA work group focuses on the use of digital tools to support and enhance traditional certification processes, ensuring that data quality underpins the infrastructure used for virtual testing and related tasks. The MBE user group, on the other hand, emphasises the integration of digital models and data management practices to create a seamless and traceable digital thread throughout digital supply chain environments. Both work groups aim to improve data provenance and trustworthiness, which are essential components of a robust data quality framework.

The report is structured as follows: the second section looks into the importance of data quality in manufacturing, discussing key dimensions and standards. The third section focuses on CbA, emphasising the role of data quality in virtual testing and certification processes. The fourth section examines the MBE concept, exploring how digital supply chain integration and data traceability can be enhanced through initiatives such as the Connected Model-Based Enterprise Environment (C-MBEE) flagship. The final sections provide a discussion on the current state of data quality in manufacturing and conclude with recommendations for future collaboration and implementation.

By addressing these topics, this report aims to provide a comprehensive understanding of how data quality frameworks can support the digital transformation of the manufacturing sector, ensuring that data-driven decisions are reliable, traceable, and aligned with industry standards.

2 DATA QUALITY

Data quality is a cornerstone of modern manufacturing, playing a crucial role in ensuring efficient and accurate operations. High-quality data is essential for monitoring processes, making informed decisions, and maintaining a competitive advantage in the industry.

In manufacturing, data quality impacts various aspects, from production to supply chain management. For instance, sensors and IoT devices generate vast amounts of data that are vital for real-time monitoring and optimising manufacturing processes. Ensuring the accuracy, completeness, and consistency of this data is critical for maintaining operational efficiency and product quality. NPL's approach to data quality in manufacturing builds on the ISO 25012

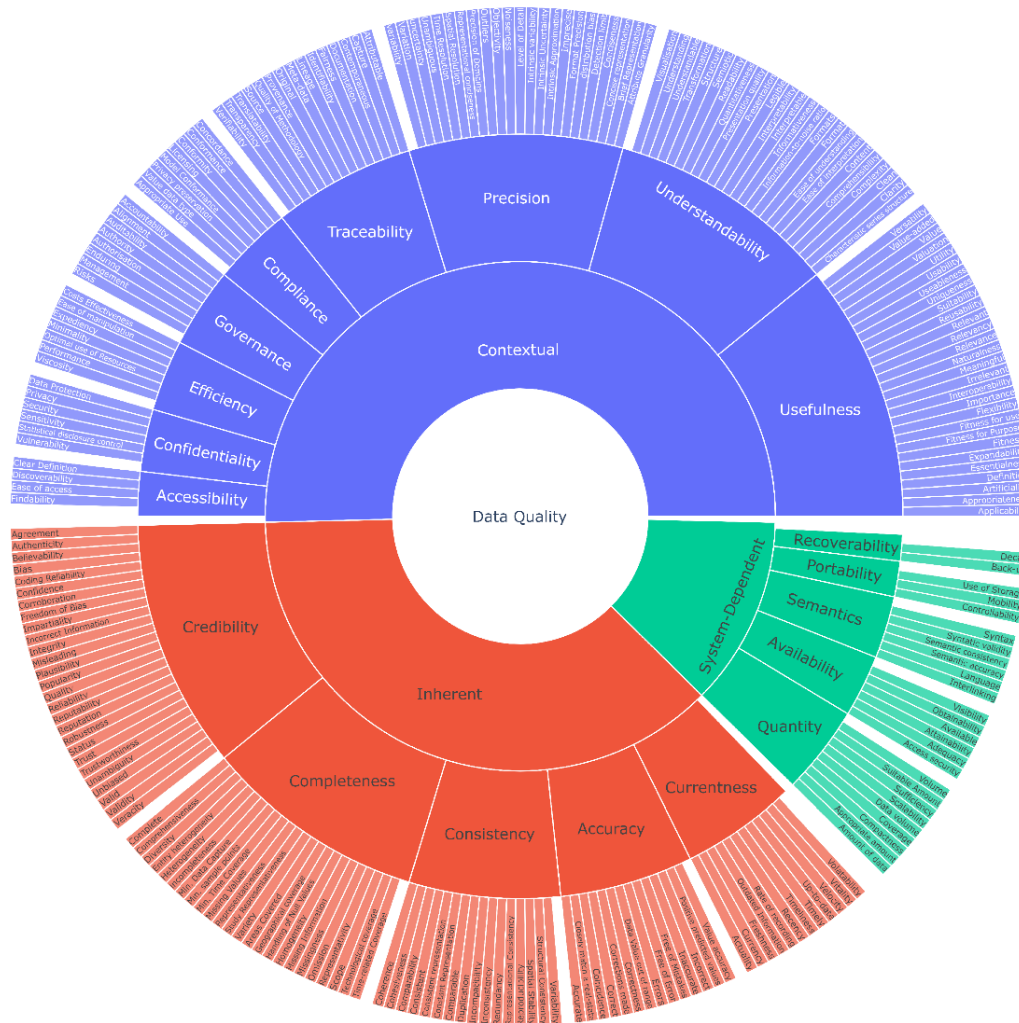


Figure 1: Data Quality Framework. The first concentric ring categorises data quality into three domains: inherent, contextual, and system dependent. Each domain is further split into core dimensions. Figure taken from reference [2] with permission.

standard, which provides a comprehensive framework for managing data quality in information systems [3]. This standard includes key data quality dimensions such as correctness, completeness, and consistency, which are essential for ensuring reliable data in manufacturing environments.

Building on the ISO 25012 framework, we suggested the addition of four more data quality dimensions: governance, usefulness, quantity, and semantics, expanding the framework as illustrated in Figure 1 [2]. These dimensions cover specific needs in manufacturing, such as ensuring data is governed for compliance and regulatory requirements, determining how useful data is for decision-making, use the amount of data that sensors produce to optimise performance, and knowing the semantics of data so that systems can communicate with each other across different functions of a supply chain environment.

High-quality data supports manufacturers to implement advanced techniques such as predictive maintenance, where data from sensors is used to predict equipment failures before they occur, reducing downtime and maintenance costs. It also supports the implementation of model-based enterprises, where digital models of products and processes are used to improve design, production, and lifecycle management.

Despite the specific data needs of the manufacturing sector, the general principles of data quality for information systems remain relevant. By adopting a structured approach to data

quality, manufacturers can ensure that their data is reliable and fit for purpose, supporting better decision-making and more efficient operations, while also leaving a demonstrable audit trail for streamlining approvals.

3 CERTIFICATION BY ANALYSIS

As manufacturers increasingly adopt digital tools and technologies, ensuring confidence in the information and outcomes they facilitate becomes crucial. These tools and technologies generate data that informs or automates decisions. Understanding the traceability of such decisions requires quantifying the quality of the data they use. This is particularly important in applications where these tools replace or supplement traditional certification processes, such as in CbA. Because of this, the infrastructure used for virtual testing and related tasks needs to be based on tools and methods for evaluating data quality.

The importance of CbA in the future of manufacturing is well recognised. The HVMC has identified CbA as a priority technology under the Smart Design umbrella and is actively delivering projects to accelerate this capability development in the UK [4]. This focus has led to collaboration between NPL and the HVMC, resulting in the publication of a report titled “Uncertainty Quantification: An Essential Aspect of Certification by Analysis” which is available on request to the authors. This publication highlights the importance of uncertainty quantification in CbA and discusses several methods for evaluating the uncertainty associated with modelled values. Quantifying uncertainty in these applications is critical, and understanding acceptable levels of uncertainty for any given application is complex.

Following the Innovate UK (IUK) Materials and Manufacturing Vision 2050 publication [5], the HVMC has partnered with the British Standards Institution (BSI), NAFEMS, and NPL to establish a national Smart Design Innovation Network. The goal of this network, which includes a dedicated CbA subcommittee, is to drive the transformation of product design and through-life certification. The network has engaged key industrial stakeholders through several events and has received interest across multiple sectors. It has identified modelling, simulation, and digitalisation capabilities as key enabling technologies for achieving competitive engineering transformations in the UK. These data-driven technologies must demonstrate sufficient trustworthiness to replace or contribute to certification processes. Quantifying confidence in these solutions is a challenge that data quality frameworks are designed to address.

In line with these efforts, NPL and the National Manufacturing Institute Scotland (NMIS) – an HVMC centre – developed a framework for assessing the credibility of data based on its provenance. Models and simulations using this data inherit this associated credibility to support decision-making in remanufacturing [6]. A data-driven credibility assessment framework was developed to quantify the quality of models and simulations in industrial part remanufacturing. This framework, which was built using the ReMake testbed, evaluates the dependability of sources, data, and methodologies, focusing on robustness and uncertainty for data quality and simulation confidence. The testbed aimed to replace physical experiments with digital testing to reduce manufacturing costs and promote a circular economy closing in on net-zero goals. A conceptual data model designed using competency questions was developed that maps data requirements, digitalises credibility evaluations, and promotes data traceability and accessibility. This model allows for the automatic generation of a credibility assessment, which increases the reliability and trust of simulations used to check the feasibility for reconditioning worn-out industrial components.

As the HVMC continues to focus on engineering transformations, particularly the virtualisation of certification processes, the development and implementation of data quality frameworks presents a valuable opportunity for collaboration.

4 MODEL-BASED ENTERPRISE

The C-MBEE programme, part of the Made Smarter Innovation Hubs, was designed to improve digital supply chain integration by enabling the testing and demonstration of digital solutions in a collaborative environment [7]. This program aimed to solve issues that arise when suppliers need to share data, especially Product Manufacturing Information (PMI). It does so by proposing the development of new manufacturing data-sharing standards, making it easier to evaluate the tool-based capabilities of each supplier, improving collaboration in the supply chain and mitigating interoperability problems. It focuses on the exchange of PMI features between big manufacturers and their suppliers.

At the centre of an MBE framework are Model-Based Definition (MBD) models, which are increasingly recognised and used by manufacturers. MBD models consist of 3D CAD designs that contain all the necessary information to manufacture a part, including details typically found on 2D drawings. These designs act as master files or the “single source of truth” in the supply chain. Both the American Society of Mechanical Engineers (ASME) [8] and the International Standards Organization (ISO) [9] have created standards to define and capture data that is difficult to outline on engineering drawings, ensuring universal interpretation.

Despite its benefits, the adoption of MBD has been slow in some industries due to a lack of understanding of relevant standards, availability of MBD-ready software, and challenges in data management to ensure interoperability and provenance of their data, including IP protections. Additionally, cultural resistance within companies poses a challenge to MBD adoption. To address these challenges, companies must develop data management plans that establish minimum requirements for data quality and handling, in accordance with current MBD standards. Training and educating less digitally literate suppliers are also crucial for facilitating supply chain integration. Implementing MBD practices and creating a robust data management plan can enhance collaboration with suppliers and improve data transactions and relationships across the supply chain. By defining data set quality requirements and maintaining quality standards, companies can ensure the interoperability of data across their supply chain [10].

As part of NPL's contributions towards the Connected MBEE programme, a framework, shown in Figure 2, for enhancing digital supply chain integration through software interoperability analysis was developed [11], addressing the challenges of data sharing between suppliers. It focuses on the quality aspects of manufacturing data, especially data traceability. The framework employs a systematic approach to create vocabulary mapping software that links interoperability elements with PMI features, cross-referencing with ISO and ASME standards. This enables efficient assessment of each supplier's tool-based capabilities, facilitating supply chain integration and pre-emptively addressing interoperability issues. The framework, presented as a data model using the C-MBEE testbed, focuses on the exchange of PMI features between large manufacturing companies and their suppliers.

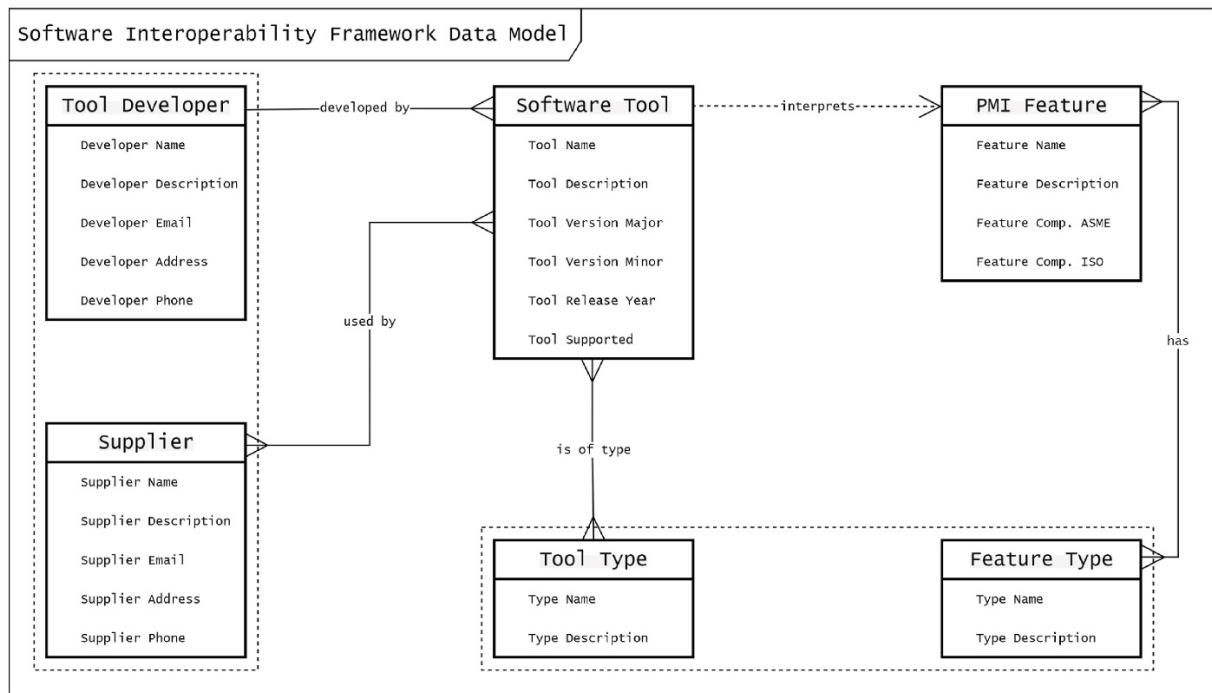


Figure 2 : Software interoperability framework for assessing PMI compatibility in digital supply chains. Adapted from [11] with permission.

Additionally, the concept of the “digital thread” was explored to further enhance supply chain integration. This concept focuses on creating traceable data pathways associated with a product's lifecycle, emphasising the importance of data integrity and traceability. The C-MBEE testbed was used to demonstrate the practical application of hashing algorithms for enhancing data traceability [12]. Unique hash values, generated by analysing file contents, serve as unique digital identifiers that confirm file integrity. This method offers possibilities for future automation and application in various settings, contributing to improved data traceability in metrology and addressing issues related to data integrity and traceability.

By adopting MBD practices and leveraging learnings from programmes such as C-MBEE, manufacturers can improve digital supply chain integration, enhance data traceability, and ensure the integrity of their data. However, the adoption of MBE is underpinned by the need for extensive digitalisation in the manufacturing sector, which itself relies on robust data quality. Both the interoperability framework and the hashing algorithm for traceability aim to improve data provenance, thereby enhancing data trustworthiness. Data provenance is only one aspect of data quality, highlighting the need for comprehensive DQ frameworks.

5 DISCUSSION

This report has already established that data quality is a fundamental aspect of modern manufacturing. High-quality data supports aspects such as process monitoring and decision-making, allowing stakeholders to maintain a competitive edge in the industry. Utilisation of poor-quality data can result in production delays, increased manufacturing costs, and compromised product quality. Conversely, high-quality data enables manufacturers to implement advanced techniques such as predictive maintenance or informing process optimisations, reducing downtime and operational costs. Adoption of CbA and MBE workflows is used by this report as an example of the need for robust data quality assessment

methodologies for enabling manufacturers and suppliers to fully leverage the benefits of CbA and MBE.

Building on established quality frameworks, such as NPL's data quality framework [2], assessment tools for data quality can be developed and tailored to the specific needs of the manufacturing sector, thereby streamlining the confident adoption of workflows such as CbA or MBE. To achieve this, the following components are necessary to include in the assessment tool:

1. **Manual assessments:** Regular audits and reviews of data quality, involving subject matter experts to identify and rectify issues according to guidelines established by the data quality framework;
2. **Automated assessments:** Identification of data quality dimensions whose assessment could be automated and development of tools and algorithms to continuously monitor data quality;
3. **Metrics and KPIs:** Definition of Key Performance Indicators (KPIs) for each data quality dimension defined in NPL's data quality framework [2] to measure data quality over time.

By identifying the relevant data quality dimensions for manufacturing and pinpointing specific aspects within each dimension, manufacturers can employ targeted assessment methods. Considering the manual assessment aspect described above, a tool such as the MBE Maturity Index developed by Sandia National Laboratories [12] could be adapted and repurposed to assess the maturity of each individual data quality dimension. For instance, manufacturers can assess the accuracy of sensor data, ensure the completeness of datasets, and validate the consistency of data across different systems. This would also allow for the exclusion of irrelevant dimensions for specific applications within the manufacturing sector. The proximity of this "data quality maturity assessment" with that of Sandia's MBE Maturity Index would also enable the same review and audit process to be used, meaning that resources already held by manufacturers and suppliers could be used, minimising the need for staff training.

The repurposing of Sandia's Maturity Index into an analogous data quality assessment tool would also allow for the identification of specific data quality dimensions that could benefit from automated assessment methods. For example, machine learning models can be used to detect anomalies and inconsistencies in real time, but they require the data they ingest to meet specific requirements to produce quality outputs. At this stage, algorithms and tools could be used to ensure that the data used by predictive maintenance models consists of accurate measurements (accuracy) with appropriately minimised uncertainty (precision) and that the datasets do not have large amounts of missing values (completeness). Additionally, data should be consistent for these models to be able to detect trends and out-of-bounds values in the data (consistency).

For the data quality assessment tool to be effective, it is necessary that the metrics and KPIs used in the industry are directly linked to the existing data quality dimensions. This alignment ensures that the assessment tool accurately reflects the specific requirements and standards of the manufacturing sector. By mapping each metric and KPI to a corresponding data quality dimension, manufacturers can systematically evaluate and improve data quality across all relevant aspects. This approach not only enhances the precision and reliability of the assessments but also facilitates the automation of data quality monitoring, making it easier to maintain high standards consistently.

6 CONCLUSION

The integration of robust data quality frameworks is essential for promoting and enabling the digital transformation of the manufacturing sector. High-quality data enables manufacturers to make reliable and traceable decisions, enhancing operational efficiency and product quality. This report outlines CbA and MBD as core technology workflows that could directly benefit from the implementation of robust data quality frameworks. It also describes how NPL's existing data quality framework could be used as a base to create manufacturing-specific data quality assessment tools and frameworks by repurposing current MBE maturity assessment frameworks. By developing and implementing tailored assessment frameworks for evaluating data quality across the supply chain, both manufacturers and suppliers can achieve significant improvements in their digitalisation journeys, supporting the long-term success and sustainability of the move towards Industry 4.0 in the UK manufacturing landscape.

7 ACKNOWLEDGEMENTS

This work was funded by the UK Government Department for Science, Innovation and Technology through the UK's National Measurement System. Thanks is given to Russel Miller, Daniel Povey, and Paul Duncan for providing feedback on the original manuscript.

8 REFERENCES

- [1] UK Research and Innovation, "UK regulatory science and innovation networks: discovery phase," UKRI, 14 12 2023. [Online]. Available: <https://www.ukri.org/opportunity/uk-regulatory-science-and-innovation-networks-discovery-phase/>. [Accessed 18 02 2025].
- [2] R. Miller, H. Whelan, M. Chrubasik, D. Whittaker, P. Duncan and J. & Gregório, "A Framework for Current and New Data Quality Dimensions: An Overview," *Data*, vol. 9, no. 12, p. 151, 2024.
- [3] International Organization for Standardization, "ISO 25012 Data Quality Model," 2008.
- [4] HVMC, "Smart design: certification by analysis," High Value Manufacturing Catapult (HVMC), 2025. [Online]. Available: <https://hvm.catapult.org.uk/what-we-do/programmes/certification-by-analysis/>. [Accessed 06 02 2025].
- [5] UK Research and Innovation, "Materials and manufacturing vision 2050: Reimagining materials and manufacturing together," UKRI, 2023.
- [6] J. Gregório, M. Alsuleman, M. Chrubasik, P. Duncan and G. Bisland, "A competency question driven approach to conceptual data model design for digital verification and validation," in *Advanced Mathematical and Computational Tools in Metrology and Testing 2023*, Saravejo, 2024.
- [7] Made Smarter Innovation | Digital Supply Chain Hub, "Connected MBEE," 2024. [Online]. Available: <https://digitalsupplychainhub.uk/showcase/connected-mbee/>. [Accessed 28 February 2024].
- [8] American Society of Mechanical Engineers, "ASME Y14.41 Digital Product Definition Data Practices," ASME, New York, 2019.
- [9] ISO, "ISO 16792 Technical product documentation — Digital product definition data practices," ISO, 2021.
- [10] J. Gregório, J. Hughes and C. Lorch, "Model Based Definition – Recommendations for Supply Chain Integration," National Physical Laboratory, 2023.

- [11] R. Miller, J. Gregório and P. Duncan, "A framework for streamlining digital supply chain integration through software interoperability analysis," *Measurement: Sensors*, p. 101479, 2024.
- [12] J. Gregório, R. Miller and P. Duncan, "Hashing it out: Unravelling the threads of data in digital manufacturing," *Measurement: Sensors*, p. 101802, 2024.
- [13] Sandia National Laboratories, "NSC MBE Maturity Index Issue B," January 2022. [Online]. Available: <https://www.sandia.gov/mbit-mbe/2022/04/05/nsc-mbe-maturity-index-issue-b/>. [Accessed May 2024].