Semantic Data Management Maturity Survey – Initial Results Analysis

Gerhard Goldbeck* and Vikki Cantrill

Goldbeck Consulting Ltd, UK

* Corresponding Author: gerhard@goldbeck-consulting.com

SemanticMaterials

Funded by the European Union

Contents

Executive summary

This report provides a first analysis and evaluation of the initial responses received for the Semantic Data Management Maturity Assessment Service[1] survey launched in September of 2025 as part of a service provided by Semantic Materials[2].

The current work builds upon a larger benefits analysis and maturity study undertaken in 2019 that looked into strategies for industry to engage in materials modelling[3]. This earlier paper explains how industry can make materials modelling and simulation a routine part of research, product development, scaling, and product life-cycle management. It summarises future developments and the barriers to adoption, and suggests strategies to overcome them to boost innovation. Consequently, this current work compliments and looks at parallels within semantic data management.

The Semantic Data Management Maturity Assessment service[1] is a self-assessment tool that enables individuals to evaluate their company's current position and aspirations towards adopting and implementing effective semantic data and knowledge management systems within today's artificial intelligence (AI)-driven innovation environment.

The survey is structured to probe the key issues of knowledge management[4] of data, processes, people, and tools and technologies. From the survey responses received, the following report provides a European perspective in this area.

This report is a brief appraisal of the results to date[5] and will be used to inform future work by us in this area. Briefly, current maturity levels are relatively low, with fragmented data management preventing efficient collaboration. The "people" dimension shows highest maturity through growing awareness, however "processes" lag behind, likely owing to a lack of standardised workflows.

Pleasingly, significant improvements and progress is expected within the next 3–5 years as strategies begin to be implemented across business functions.

Key opportunities in this quickly developing area include transforming data into integrated, semantically rich resources; faster, more efficient innovation, improve product quality, and strengthen market competitiveness. It should be noted however that there are currently substantial barriers to overcome. Barriers include limited resources and budget, insufficient corporate buy-in, poor data quality, and a lack of shared terminologies. Despite these challenges, substantial advances are anticipated as companies begin to recognise the strategic benefit of semantic data management.

The maturity model

Data and/or knowledge management is crucial across many organisations and becoming ever more so within the area of materials science[6]. Scientific research and developments relies on excellence in data quality, reproducibility and collaboration, and, as such, demands effective data management to benefit research investments. During the next 10 years data management is expected to become more automated and AI-driven with stringent requirements, like real-time data sharing and governance.

Consequently, understanding the current situation experienced within materials research and development is essential.

Maturity models offer a means to evaluate a number of strategic areas within an organisation and monitor changes over time. Maturity models typically review defined business areas and rank them, so that performance and development can be mapped.

There are numerous maturity models available for businesses to work with, each with their own strengths and shortcomings[4,7]. So, choosing a model that suits business need by being sufficiently comprehensive and flexible to meet evolving needs is important.

Through evaluating an organisation's knowledge management capabilities, by looking specifically at the areas (also called maturity model dimensions) of data, people, processes, and tools[4] and technologies, we can assess how developed and how well structured and organisation's practices are. Indeed, a benefits analysis and maturity study undertaken in 2019 that looked into strategies for industry to engage in materials modelling used very similar maturity model dimensions and descriptors to evaluate that area also[3].

Here, we developed a survey[1] as a tool to undertake a self-assessment of semantic data management maturity to enable organisations to evaluate their current position and aspirations towards adopting and implementing effective semantic data and knowledge management systems.

Semantic data management is based on the concepts of the Semantic Web and its related standards and tools. The latter are overseen by the World Wide Web Consortium (W3C)[8], an international public-interest non-profit organisation. In particular, according to Simperl et al "the Semantic Web refers to a set of standards and technologies that are used to enable the machine-driven design and operation of a wide range of aspects of data management, from the description and organisation of collections of digital artifacts, to the access to distributed digital repositories, as well as search, ranking, and recommendation."[9]

Furthermore, according to W3C, semantics and related knowledge engineering technologies:

- provide the tools to achieve a higher degree of automation of the underlying computational tasks, thus enabling scalability,
- alters the ways the corresponding applications are developed,
- changes how information is created and evolves as part of these applications,
- facilitates information reusability,
- facilitates mediation across distributed information spaces, which are discovered and explored at run time—independently of the specific data organisation (as in traditional databases).

The core features of semantically enabled applications are that:

- data is self-described, encoded as RDF (Resource Description Framework) using a triple-based data model and accessed via Web standards and protocols, such as URIs (Uniform Resource Identifier) and HTTP (Hypertext Transfer Protocol),

- the meaning of data is specified by using ontologies. Ontologies capture the types of things of interest in a given vertical domain, as well as their main attributes and relationships,
- reasoning can be used to infer new facts from a Semantic Web knowledge base, to classify new information according to an ontological schema, and to identify potential inconsistencies and other mismatches in data quality,
- data management services are enabled to operate in an open environment, where new data sources can be discovered, integrated, and used on the fly.

The results from the assessment survey can enable enterprises to reflect on their current status and develop effective roadmaps to move from poor data integration and knowledge sharing to well-implemented and well-governed semantic data management systems.

Respondents to the survey were asked to reflect upon and rate how they perceived the adoption and implementation of semantic data or knowledge management systems where they worked. Information about four key areas was queried and ranked — data, people, processes, and tools and technologies. Level one ranking is for an area in its infancy with awareness, whereas level five reflects transformational change with full adoption and implementation[4].

Where respondents ranked responses, we offered the levels of semantic data management maturity based on information from Data Graphs (formerly Data Language, link not publicly available), with further inspiration based on Gartner[10] and the Technical University of Munich[11].

- Level 1. Awareness, but no implementation

The organisation knows about semantic data management but have not implemented it yet. The use of metadata is largely down to individual functions and IT systems. There is some awareness in the organisation of the benefits of semantic data management, e.g. knowledge graphs. There is also an awareness of the need to address a lack of interoperability, issues with data integration and the fact of knowledge being locked in people's heads. However, addressing these issues strategically with semantic data management is still kicked down the road.

- Level 2. Active, experimental projects ongoing

The organisation is playing with semantic data management informally and at the level of some functions. More enterprise-wide information governance is being considered. Activities are driven by the quest for AI implementation. However, legacy information systems are still dominant, semantic systems, such as knowledge graphs, are applied experimentally and in lighthouse projects.

- Level 3. Operational, clear objectives and implementation in some functions

The organisation has adopted clear objectives for implementation of semantic data management and it becomes more widely operational. It means that there are now common terms defined and understood across the entire organisation. There is some information governance in place. Tools, processes and guidelines to support the curation of high value information assets are implemented. Legacy information systems are still in place but operational semantic data management is more widely used.

- Level 4. Systemic, robust and scalable implementation across the business

Tools, such as knowledge graph platforms, are robust, scalable, and easy to integrate across business systems. This involves domain modelling, representing the concepts and relationships of a domain in an ontology that is understandable by both technical and non-technical stakeholders. All core assets are described and organised based on ontology (context, meaning). Tools also include semantics user interfaces that lets people manage and work on data models as well as search, explore and curate data. Reference data available via the user interface.

Legacy information systems have been replaced by semantic knowledge management systems; the latter enable quickly pivoting around a core business information model and support the use of application programming interfaces for end users.

Semantic data management platforms are strategic assets of the business, forming an authoritative source of all business information and reference data.

- Level 5. Transformational, demonstrated efficiency and effectiveness gains

Semantic data management is used pervasively, supports collaborations (internally and externally) and serves new insights. It supports more transformational use of data and the transforms the value of using machine learning/AI techniques. Research and development and new product development/launch are more efficient and effective owing to enhanced data availability and cooperation between functions.

Results

Respondents to the Semantic Data Management Maturity Assessment Service[1] provided qualitative judgements about their organisations as well as detailed feedback in response to specific questions.

The survey was circulated within Europe and replies requested. Respondents were primarily from materials and manufacturing industries, solutions providers and research institutions. The majority of respondents noted that, although they were not the sole owners of data management work, they did consider this a daily aspect of their work. And although overall maturity values are currently low, some data management strategies are now being implemented at the departmental or even enterprise scale, particularly within larger organisations.

From a high-level strategic perspective, we asked survey respondents a number of questions to ascertain how much of a priority a digitalisation strategy was within their organisation and what approaches were being followed. In general, strategies tend to be data-centric in approach by focusing on integration/interoperability and governance. However, much less of a priority was the need to become agnostic and independent from software vendors.

Conversely, projects citing AI or data integration benefit from widespread organisational awareness and are frequently discussed. Although here ontologies are given much less prominence, so here there is a tendency towards a lack of awareness.

To get further insights into a company’s current position semantic data and knowledge management implementation and to understand the importance that this topic plays strategically within in it for the near future, we asked respondents to evaluate the current state of play within their organisation and their best guess as to what levels may realistically be achieved in the next 3–5 years (Figure 1).

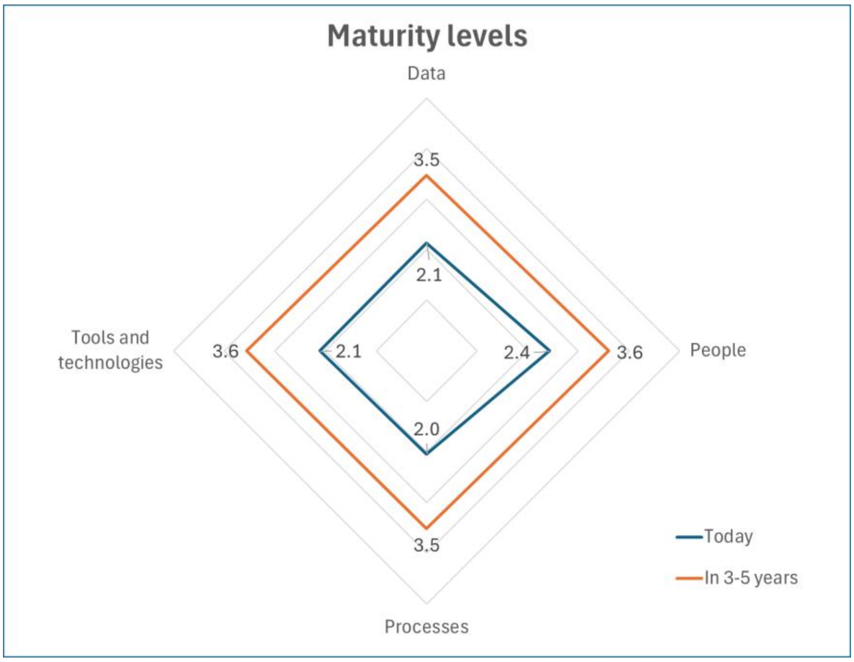

Figure 1 Semantic data and knowledge management maturity levels from survey respondents

Today, for the most part, it appears that the maturity levels within all areas — data, people, processes, and tools and technologies — is fairly low with a rating of a little over 2 or Active. Typically, the people maturity level is the highest rated area. It is worth noting that this trend is largely in line with that seen in the maturity study for materials modelling[3]. From the individuals' perspective, there is some knowledge about data and knowledge management and the changes and opportunities that may be on the horizon and how that may affect associated processes.

Subsequently, the lowest rated area tends to be within processes. This may well we because people have an awareness of what changes are needed and may even have undertaken preliminary training. However, the next steps are associated with considerable barriers whereby an organisation may well have access to appropriate tools, but aligning the tools and integrating them into the processes and all associated data systems can require considerable time and effort to achieve.

Indeed, within research and development areas it is the lack of standardised data formats and workflows that hinder the adoption of changes to established practices. So, transformation of processes tends to be the last area within which developments are noticeable.

Pleasingly, substantial steps towards implementing solutions are anticipated over the next 3–5 years, which is reflected in the overall increases in expected maturity levels. It is encouraging to note the improvements expected in the areas of people and tools and technologies, which will be at the cornerstone of ensuring good implementation and ensuring culture and process changes and developments into the future. Organisations aim to lift their maturity to between levels 3 and 4 across all areas, which means operations with clear objectives and implementation in some functions, working towards systemic, robust and scalable implementations across the business.

Alongside the qualitative questions posed, we also asked a number of more in depth questions for analysis. For context, it is important to understand who responded to our survey and the roles they currently take on within their organisations. The majority of the respondents identified quite strongly within the area of semantic technologies and data management or AI and work actively within these areas. A number of them being involved in the development or creation of semantic models, knowledge graphs, ontologies, controlled vocabularies, solutions for data analysis and digital archiving. We also received input from those working in strategic roles, such as training and consulting, as well as with technical delivery of the projects and company-wide change management. We therefore assume that the responses received accurately reflect the present circumstances within industry.

We asked, "what are your biggest challenges with respect to product data management?" The responses highlighted a fragmented and inconsistent approach for many that leads to multiple versions of the truth (which source data should be most trusted) and an inability to delve into business questions easily owing to a lack of business-critical definitions across an organisation's departments. Consequently, data that spans several functional areas typically needs to be collected and integrated manually. Here, a lack of shared terminology and concept definitions means jargon can inhibit both cross-team and external communication and collaboration. This is a particular concern with materials science because there is often a reliance on expert knowledge and the field is broad and complex, which can hinder and collaborative efforts. Furthermore, within the area, unlike chemistry that benefits from the IUPAC Gold Book[12] for definitions, there are few controlled vocabularies or agreed terminologies and there is a scarcity of curated data. One exception to this in the field of materials science is the Elementary Multiperspective Material Ontology[13,14], known as EMMO, which is a knowledge framework able to represent the materials world including its physics foundations.

Within our survey, we also asked what the biggest business drivers or opportunities posed by integration of data are that offer great value to your organisation and what are the biggest hurdles to ensuing this happens effectively?

The clear response here was the opportunity to transform data into an integrated, semantically rich resource for all. The benefits of connected intelligent data supported by metadata, workflows, and shared terminologies was highlighted. Such opportunities are expected to accelerate innovation, improve product quality and strengthen a company's position and competitive advantage in the relevant market sectors.

There are numerous hurdles to overcome that prevent effective data integration despite clear advantages. Such hurdles include, a lack of resources and budget, poor buy-in by leadership, and poor-quality data and workflows. Moreover, short-term strategic priorities and well-organised (though smaller) datasets reduces the perceived urgency.

In summary, the responses received to date for this survey align with our understanding of the field[15,16], how fast the area is evolving and the advances that are hoped for in the near future. Semantic data management maturity is low, but substantial advances are planned. Consequently, considerable change and improvement is expected overall within the materials sector over the coming decade. It is now our intention to expand the survey to glean further insights and monitor changes over the coming years to see how materials science adapts and manages data integration and innovation challenges going forward.

Acknowledgements

This study has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 862136 (OntoTrans) and the Horizon Europe Research and Innovation Programme under grant agreement no. 101137725 (BatCAT), and from UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee [grant number 10091190].

Disclaimer

All statements of fact, opinion, or analysis expressed in this report are those of the authors only and do not necessarily reflect those of the survey respondents or the European Union. Neither the European Union nor the granting authority can be held responsible for them. The information used and statements of fact made are not guarantees, warranties or representations as to their completeness or accuracy. The authors assume no liability for any short term or long terms decision made by any reader based on analysis included in this report. In many cases, the opinion expressed in the reports is our current opinion based on prevailing trends and is subject to change.

References

1. Cantrill, V. Semantic Data Management Maturity Assessment Service. *Semantic Materials* https://semanticmaterials.com/semantic-data-maturity-service/ (2025).
2. Semantic Materials website. *Semantic Materials* https://semanticmaterials.com/.
3. Goldbeck, G. & Simperler, A. Strategies for industry to engage in materials modelling. https://doi.org/10.5281/zenodo.3564455 (2019) doi:10.5281/zenodo.3564455.
4. Bougoulia, E. & Glykas, M. Knowledge management maturity assessment frameworks: A proposed holistic approach. *Knowledge and Process Management* **30**, 355–386 (2023).
5. *An Initial 15 Responses Have Been Included in This Analysis. Responses Were Received from Staff Working within Universities, Industry, and Software and Solutions Providers.*
6. Oliveira, O. N. & Oliveira, M. C. F. Materials Discovery With Machine Learning and Knowledge Discovery. *Front Chem* **10**, 930369 (2022).
7. Escrivão, G. & Silva, S. L. D. Knowledge management maturity models: identification of gaps and improvement proposal. *Gest. Prod.* **26**, e3890 (2019).
8. W3C. *W3C* https://www.w3.org/.

9. Simperl, E., Cuel, R. & Stein, M. Semantic Data Management: A Human-driven Process. in *Incentive-Centric Semantic Web Application Engineering* (eds Simperl, E., Cuel, R. & Stein, M.) 1–18 (Springer International Publishing, Cham, 2013). doi:10.1007/978-3-031-79441-4_1.
10. Understanding the Five Stages of Gartner's Maturity Model for Manufacturing Excellence. https://www.gartner.com/doc/3425317/understanding-stages-gartners-maturity-model.
11. Langer, B. Understanding Data & Analytics Maturity: A Systematic Review of Maturity Model Composition. *Schmalenbach J Bus Res* **77**, 205–227 (2025).
12. *The IUPAC Compendium of Chemical Terminology: The Gold Book*. (International Union of Pure and Applied Chemistry (IUPAC), Research Triangle Park, NC, 2025). doi:10.1351/goldbook.
13. Del Nostro, P. *et al.* Elementary Multiperspective Material Ontology: Leveraging Perspectives via a Showcase of EMMO-Based Domain and Application Ontologies: in *Proceedings of the 16th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management* 135–142 (SCITEPRESS - Science and Technology Publications, Porto, Portugal, 2024). doi:10.5220/0012910200003838.
14. Elementary Multiperspective Material Ontology. https://github.com/emmo-repo/EMMO (2024).
15. Cantrill, V., de Baas, A., Roscioni, O. & Goldbeck, G. *Developing Ontologies in Materials Science: Semantic Materials Workshop Report - Semantic Data and Knowledge Management for Chemicals and Materials Industries*. https://zenodo.org/doi/10.5281/zenodo.11127892 (2024) doi:10.5281/ZENODO.11127892.
16. Cantrill, V., Goldbeck, G., Kienberger, F. & Simperler, A. *Knowledge Organisation and Intelligent R&D for Chemicals and Materials Industries – A Workshop Report*. https://zenodo.org/doi/10.5281/zenodo.17749003 (2025) doi:10.5281/ZENODO.17749003.