This paper was presented at the XXXVII ISPIM Innovation Conference, held in Granada, Spain on 07 June to 10 June 2026. ISBN 978-952-65771-2-8.

Trust Before Intelligence: Data Quality as the Foundation for AI Agents

Simon B. Kranzer*

Meshmakers GmbH, Firmianstrasse 31A, 5020 Salzburg, Austria
E-mail: simon.kranzer@meshmakers.io

Reinhard Mayr

Meshmakers GmbH, Firmianstrasse 31A, 5020 Salzburg, Austria
E-mail: reinhard.mayr@meshmakers.io

Gerald Lochner

Meshmakers GmbH, Firmianstrasse 31A, 5020 Salzburg, Austria
E-mail: gerald.lochner@meshmakers.io

* Corresponding author

Abstract: Organizations investing in AI agents frequently discover that the limiting factor is not model intelligence but the quality, ownership, and trustworthiness of the underlying data. This paper presents a practice-oriented case for applying Data Mesh principles as an organizational and architectural foundation for AI agents. We describe how a decentralized, domain-oriented approach — treating data as a product with clear ownership, quality guarantees, and self-service access — addresses the failure modes typical of centralized data platforms. Drawing on a practical implementation, we discuss trade-offs between decentralization and central control, mechanisms for embedding data quality and semantic clarity, and the cultural shifts required. We argue that establishing trust in data is a precondition for productive AI agents, and that innovation managers should treat the data foundation as an innovation object in its own right.

Keywords: data mesh; data quality; AI agents; data products; data governance; decentralized architecture; domain ownership; trustworthy AI; platform engineering; digital transformation

1 Introduction

Most AI agent initiatives fail far upstream of the model. Teams invest in agent frameworks, orchestration layers, and fine-tuned language models, only to discover that responses degrade in production because the data the agents depend on is inconsistent, stale, ambiguously defined, or unaccountable. In many cases, the limiting factor is not model capability but trust in the data feeding intelligence (Redman, 2018).

This challenge can be understood as an innovation management problem rather than a purely technical one. Research on digital innovation emphasises that value creation increasingly depends on how organizations structure and recombine digital resources across organisational boundaries, rather than on isolated technological artefacts (Yoo et al., 2010; Nambisan et al., 2019). From this perspective, data is not merely an input to AI systems but a critical organisational asset whose ownership, quality, and accessibility must be actively designed.

At the same time, emerging AI agents — systems that combine foundation models, retrieval mechanisms, and tools to act on behalf of users — place new demands on data. Unlike traditional analytics, agents operate autonomously and communicate results in natural language, amplifying the impact of data inconsistencies and ambiguities.

This paper presents an exploratory case study in the domain of distributed energy systems, specifically energy communities, where heterogeneous data sources were integrated into a decentralised data architecture based on Data Mesh principles (Dehghani, 2020). The goal was to enable AI agents for consumption and cost optimisation.

The contribution is twofold:

1. An empirically grounded case study of applying Data Mesh principles in a distributed energy context to support AI agents.

2. A conceptual argument that data trustworthiness — operationalised through ownership, quality guarantees, and semantic clarity — constitutes a foundational layer for AI-based systems.

The remainder of the paper is structured as follows. Section 2 reviews related work on data quality, data architectures, and AI system reliability. Section 3 characterises the data-quality challenge that AI agents expose. Section 4 outlines the Data Mesh approach and its relevance for AI. Section 5 describes the implementation case. Section 6 reports lessons learned, trade-offs, and mechanisms. Section 7 discusses implications for innovation managers. Section 8 concludes.

2 Related work

The effectiveness of AI systems is increasingly understood to depend less on model sophistication and more on the quality, structure, and governance of the underlying data. Early work by Sculley et al. highlights the phenomenon of "hidden technical debt" in machine learning systems, demonstrating that data dependencies, pipeline fragility, and feedback loops often dominate system complexity and failure modes (Sculley et al., 2015). Subsequent research and industry practice have reinforced this view, emphasizing that improvements in data quality and reliability frequently yield greater performance gains than incremental advances in model architectures.

From an organizational perspective, Redman argues that data quality is fundamentally an organizational problem rather than a purely technical one, requiring clear accountability, governance structures, and stewardship (Redman, 2013). Similarly, Provost and Fawcett

(2013) emphasize that value creation in data science depends critically on understanding, preparing, and managing data, rather than on algorithmic novelty alone.

In response to the limitations of centralized data architectures, Dehghani proposes Data Mesh as a decentralized socio-technical paradigm that treats data as a product and assigns ownership to domain teams (Dehghani, 2020; 2022). The approach is based on four core principles: domain-oriented ownership, data as a product, self-serve data infrastructure, and federated computational governance. These principles aim to address scalability and organizational bottlenecks inherent in monolithic data platforms. Empirical and practitioner-oriented studies (e.g. Machado et al., 2022; Strengholt, 2023) indicate increasing adoption of Data Mesh, while also highlighting challenges related to governance complexity and cross-domain consistency.

Parallel to these developments, recent research on modern AI systems — particularly large language models and AI agents — emphasizes the importance of grounding, robustness, and trustworthiness. Liang et al. (2023) argue for holistic evaluation frameworks that go beyond accuracy and consider reliability, transparency, and safety. These dimensions are intricately linked to the quality, provenance, and semantic clarity of the data used during both training and inference.

More specifically, AI agents that rely on retrieval mechanisms (e.g. retrieval-augmented generation) introduce additional dependencies on structured and well-described data sources. Prior work in natural language processing has shown that models tend to produce inconsistent or misleading outputs when operating on weakly governed or semantically ambiguous data (Ji et al., 2023; Sculley et al., 2015; Liang et al., 2023). This reinforces the need for explicit data contracts, metadata, and traceability mechanisms.

Despite these advances, the relationship between decentralized data architectures and the performance of AI agents remains underexplored, particularly in distributed and domain-intensive environments such as energy systems. Existing research has treated data architectures and AI systems as separate concerns, leaving a gap in understanding how organizational data design influences the reliability and effectiveness of AI agents.

This paper addresses this gap by examining how Data Mesh principles can be operationalized to support AI agents in a real-world setting, and how data quality, ownership, and governance interact to enable trustworthy AI-driven decision-making.

3 The data quality challenge for AI agents

Why AI agents amplify data problems

AI agents — systems that combine foundation models, tools, and retrieval mechanisms to act on behalf of users — interact with data in ways that differ from traditional analytics. A dashboard consumer can visually detect anomalies; a language-based agent usually cannot. An agent will confidently summarize an incomplete table, mis-join entities whose identifiers are inconsistent across domains or quote outdated figures with no warning. Errors that were tolerable in reports become reputational risks when delivered as fluent prose to customers or regulators (Sculley et al., 2015).

Three failure modes recur in practice:

- **Semantic ambiguity:** the same term (e.g. "active customer", "booking", "order") carries different definitions across domains, and the agent collapses them silently.
- **Unclear ownership:** when an agent produces a wrong answer, it is unclear who is responsible for the source data and who can correct it.
- **Pipeline opacity:** data reaches the agent through long chains of transformations that no single team fully understands, making root-cause analysis slow and accountability diffuse.

Limits of centralized data platforms

Centralized data platforms were conceived for stable analytical workloads. They assume that a central team can absorb the semantic knowledge of every producing domain and remain accountable for quality. Evidence from practice suggests this model does not scale with organizational complexity or the pace of change introduced by AI experimentation (Dehghani, 2022; Machado et al., 2022).

4 Data Mesh as an architectural and organizational response

Four principles

Data Mesh (Dehghani, 2020; Dehghani, 2022) proposes four principles:

1. **Domain-oriented ownership:** the teams closest to the business domain are accountable for the data they produce.

2. **Data as a product:** datasets are treated as products, with identified consumers, quality guarantees, versioning, documentation, and a responsible product owner.

3. **Self-serve data platform:** a platform team provides infrastructure that makes it affordable for domains to publish high-quality data products without re-inventing storage, catalogue, or observability components.

4. **Federated computational governance:** global policies (privacy, access, quality thresholds) are codified and enforced as code, while domain teams retain autonomy within those guardrails.

From mesh principles to AI-ready data

The relevance for AI agents is direct. Agents require data that is semantically well-defined (principles 1 and 2), discoverable (platform), and governed (principle 4). Treating data as a product forces domains to articulate the contract they offer: schemas, freshness guarantees, semantic meaning, and known limitations. These contracts are also what retrieval-augmented agents need to cite sources reliably and degrade gracefully when data is missing.

This paper was presented at the XXXVII ISPIM Innovation Conference, held in Granada, Spain on 07 June to 10 June 2026. ISBN 978-952-65771-2-8.

Several accounts describe the application of mesh principles in enterprise settings (Strengholt, 2023; DAMA International, 2017). Our contribution is to position this work specifically as the precondition for AI agents rather than as a purely analytics concern.

5 Implementation case

Context: Energy communities and distributed data sources

The study was conducted in the context of energy communities, where multiple stakeholders collaboratively produce, consume, and share energy. These environments are characterized by inherently decentralized data ownership, heterogeneous data sources, and varying levels of data quality and standardization.

The data landscape included several distinct sources:

- energy provider platforms, providing consumption data, tariffs, and billing-related information,
- smart meter data from distributed metering points within the energy communities,
- external contextual data such as weather data and forecasts,
- additional auxiliary data sources (e.g. market signals or temporal pricing structures where available).

These data sources differ significantly in structure, update frequency, semantic definitions, and ownership responsibilities. Key entities such as "consumption", "load profile", or "cost" were defined inconsistently across providers and community-level systems. Furthermore, data refresh cycles ranged from near real-time (smart meters) to delayed batch updates (provider platforms), introducing additional integration challenges.

The objective was to integrate these heterogeneous data sources into a unified data space, in which each dataset is exposed as a governed data product. AI agents then used this data foundation to support:

- consumption optimization (e.g. load shifting),
- cost optimization (e.g. tariff-aware scheduling),
- and decision support for participants in energy communities.

The setting is representative of broader challenges in distributed energy systems and data-driven optimization scenarios, where data ownership is fragmented and coordination across domains is required.

Methodological approach

This study follows an exploratory case study approach, focusing on the introduction of Data Mesh principles in a real-world distributed energy setting.

The implementation was conducted over a period of 12 months, involving multiple stakeholders including energy providers, system operators, and participants of energy communities.

Data sources for the study included:

- system artefacts such as data products, schemas, and metadata definitions,
- platform-generated logs and data quality indicators,
- observations from the implementation process,
- and informal stakeholder feedback from participating domains.

The authors were directly involved in the design and implementation of the data architecture. This involvement introduces a potential bias, particularly in the interpretation of outcomes. To mitigate this, findings were triangulated across multiple sources, including system-level observations and independently observable artefacts such as data quality metrics and system behavior.

The evaluation is qualitative in nature and focuses on:

- perceived improvements in AI agent outputs before and after the introduction of governed data products,
- observed changes in data quality dimensions (e.g. consistency, completeness, traceability),
- and organizational changes related to data ownership and governance.

No controlled experimental setup or quantitative benchmarking was conducted. Consequently, the findings should be interpreted as indicative and exploratory rather than causal or generalizable.

Technical implementation

The data space was implemented as a decentralized architecture aligned with Data Mesh principles. Each domain (e.g. energy provider, community operator) was responsible for publishing its own data products.

Each data product was defined as a formal contract and included:

- machine-readable schema definitions (e.g. structured formats and field-level constraints),
- semantic metadata describing the domain meaning of attributes,
- freshness indicators specifying update frequency and latency expectations,
- explicit access constraints and usage policies.

A central data platform provided shared infrastructure for storage, transformation, discovery, and monitoring, enabling domain teams to publish data products without re-implementing common functionality.

For AI agents, a dedicated retrieval interface was implemented. This interface supports:

- structured queries across data products,
- semantic descriptions suitable for interpretation by language-based models,
- and metadata-driven filtering (e.g. based on freshness or domain relevance).

This allowed AI agents to dynamically retrieve and combine multiple data products for optimization tasks, including:

- load shifting based on consumption patterns and weather forecasts,
- cost minimization using tariff and pricing information,
- and cross-domain reasoning across provider and community data.

Additionally, an evaluation harness was implemented to replay predefined agent queries against versioned data products. This enabled detection of regressions caused by schema changes, data inconsistencies, or quality degradation.

Governance and quality assurance

Data ownership was explicitly assigned to domain teams, each responsible for maintaining and publishing their respective data products. This represents a shift from centralized data management towards distributed accountability.

Data quality was operationalized along multiple dimensions:

- correctness (alignment with known reference values),
- completeness (absence of missing or undefined values),
- timeliness (data freshness and update frequency),
- semantic clarity (well-defined and unambiguous definitions),
- traceability (ability to track data lineage and transformations).

A federated governance model was established to ensure cross-domain consistency. This included:

- shared standards for identifiers (e.g. metering points, customers),
- alignment on temporal semantics (e.g. time zones, aggregation intervals),
- and consistent handling of sensitive data and access policies.

Governance rules were implemented as policy-as-code within the platform, ensuring automated enforcement rather than relying solely on documentation.

Data products intended for AI agent consumption were subject to stricter requirements, including:

- stable semantic definitions,
- explicit disambiguation of domain-specific terminology,
- and validation through retrieval-based test cases.

This ensured that only sufficiently well-defined and reliable data products were exposed to AI agents for decision-making tasks.

6 Findings and lessons learned

Trust is a pre-condition, not an afterthought

The case study shows that applying Data Mesh principles changed how participating energy communities engage with their own data. Sources that had previously been treated as technical by-products — smart meter readings, provider billing data, weather feeds — were reframed as governed data products with explicit owners, semantics, and quality expectations. This shift made the data foundation legible to stakeholders and, in turn, allowed AI agents to be instrumented more reliably for consumption and cost optimization tasks. Trust in the underlying data emerged as a precondition for trusting the agents that depend on it.

These observations are qualitative and based on the implementation described in Section 5. Quantifying the measurable improvements in agent performance — for example, reduction in cost through tariff-aware scheduling, or accuracy of load forecasts — is the subject of a planned follow-up study.

Trade-offs between decentralization and control

Full decentralization is not the goal. Without federated governance, domains diverge and cross-domain use cases — including most interesting AI agents — become impossible. Conversely, too much central control reproduces the original bottleneck. The practical balance observed in the case:

- Domain autonomy over schema, semantics, and internal implementation.
- Central agreement on interoperability concerns: identifiers, time semantics, personal data handling, and product metadata.
- Central provision of non-differentiating infrastructure.

Cultural shifts

Product thinking is the larger change. Engineers accustomed to producing tables "for the BI team" had to learn to view other teams — and, increasingly, AI agents — as customers deserving a documented contract. Data product ownership proved harder to staff than anticipated; the role combines domain expertise, communication skills, and a product mindset that is still rare.

Observability and feedback loops

An often-overlooked finding is that AI agents turn out to be a very effective data-quality detector. Because agents exercise data in semantically complex ways, issues that had been dormant for years — duplicated customer records, inconsistent status codes —

surface quickly. Feeding these observations back to domain teams through platform tooling closed the loop between AI experimentation and data-product improvement.

7 Discussion and practical implications

For innovation managers

AI agents are an innovation delivery problem, not only a modelling problem. Innovation managers sponsoring AI work should budget explicitly for data-product work and resist the narrative that intelligence can compensate for poor data. The case demonstrates that data quality and product ownership are innovation objects that merit the same rigor as the agents themselves.

For digital transformation leaders

Transformation programmes tend to be structured around applications and platforms. Data Mesh suggests a third axis: the organization of ownership. Reassigning accountability to domains changes reporting lines, skill requirements, and incentives. These changes should be sequenced carefully and communicated honestly; otherwise, the change is resisted as a relabeling exercise.

For SMEs and cross-organizational contexts

Data Mesh is sometimes dismissed as relevant only for large enterprises. The case suggests that SMEs benefit disproportionately because they cannot sustain a large central data team in the first place. Across organizational boundaries — for example, in consortia or supply chains — the mesh model is even more natural: each partner is already a de facto domain, and product contracts can become inter-organizational interfaces.

Limitations

This paper reports on a single implementation. Generalization requires caution. Not every organization can afford the platform investment required to make self-service viable, and some regulated settings impose constraints that limit domain autonomy. We make no claim that Data Mesh is the only viable architecture for AI-ready data; we claim that it addresses the specific failure modes observed in practice.

8 Conclusion

AI agents expose, rather than resolve, an organization’s relationship with its data. Intelligence delivered on top of unowned, undefined, or untrusted data is fragile. The case reported here supports the argument that treating data as a product — with clear ownership, platform-based governance, and explicit quality contracts — produces the foundation on which AI agents can be trusted. Trust, in this sense, is not a property of the model but of the organizational and architectural arrangement around the data.

For innovation managers, the practical implication is that investment in AI agents should be paired with investment in the data-product organization that supports them. The work is less visible than model tuning, but it is where reliability is determined.

9 References

DAMA International (2017) *DAMA-DMBOK: Data Management Body of Knowledge*. 2nd edn. Basking Ridge, NJ: Technics Publications.

Dehghani, Z. (2020) Data Mesh Principles and Logical Architecture. *martinfowler.com* [online], 3 December. Available at: https://martinfowler.com/articles/data-mesh-principles.html

Dehghani, Z. (2022) *Data Mesh: Delivering Data-Driven Value at Scale*. Sebastopol, CA: O'Reilly Media.

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A. and Fung, P. (2023) Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, 55(12), Article 248, pp. 1–38.

Liang, P., Bommasani, R., Lee, T., Tsipras, D., Soylu, D., Yasunaga, M., Zhang, Y., Narayanan, D., Wu, Y., Kumar, A., Newman, B., Yuan, B., Yan, B., Zhang, C., Cosgrove, C., Manning, C. D., Ré, C., Acosta-Navas, D., Hudson, D. A., Zelikman, E., Durmus, E., Ladhak, F., Rong, F., Ren, H., Yao, H., Wang, J., Santhanam, K., Orr, L., Zheng, L., Yuksekgonul, M., Suzgun, M., Kim, N., Guha, N., Chatterji, N., Khattab, O., Henderson, P., Huang, Q., Chi, R., Xie, S. M., Santurkar, S., Ganguli, S., Hashimoto, T., Icard, T., Zhang, T., Chaudhary, V., Wang, W., Li, X., Mai, Y., Zhang, Y. and Koreeda, Y. (2023) Holistic Evaluation of Language Models. *Transactions on Machine Learning Research*.

Machado, I. A., Costa, C. and Santos, M. Y. (2022) Data Mesh: Concepts and Principles of a Paradigm Shift in Data Architectures. *Procedia Computer Science*, 196, pp. 263–271.

Nambisan, S., Wright, M. and Feldman, M. (2019) The digital transformation of innovation and entrepreneurship. *Research Policy*, 48(8).

Provost, F. and Fawcett, T. (2013) *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. Sebastopol, CA: O'Reilly Media.

Redman, T. C. (2013) Data's Credibility Problem. *Harvard Business Review*, 91(12), pp. 84–88.

Redman, T. C. (2018) If your data is bad, your machine learning tools are useless. *Harvard Business Review* [online], 2 April.

Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F. and Dennison, D. (2015) Hidden technical debt in machine learning systems. In: *Advances in Neural Information Processing Systems*, 28, pp. 2503–2511.

This paper was presented at the XXXVII ISPIM Innovation Conference, held in Granada, Spain on 07 June to 10 June 2026. ISBN 978-952-65771-2-8.

Strengholt, P. (2023) *Data Management at Scale: Modern Data Architecture with Data Mesh and Data Fabric*. 2nd edn. Sebastopol, CA: O'Reilly Media.

Yoo, Y., Henfridsson, O. and Lyytinen, K. (2010) The new organizing logic of digital innovation. *Information Systems Research*, 21(4), pp. 724–735.