Enhancing Data Governance in Data Trustees Through ODRL-Based End-of-Life Policies

Michael Steinert[1,2][a] and Daniel Tebernum[1][b]
[1]*Fraunhofer Institute for Software and Systems Engineering ISST, Dortmund, Germany*
[2]*TU Dortmund University, Dortmund, Germany*

Keywords: ODRL, End-of-Life Data Management, Data Trustees, Data Spaces, Data Governance.

Abstract: While data sharing drives innovation, ensuring compliance with legal, regulatory, and trust requirements presents significant challenges. Research identifies data trustees as intermediaries between providers and consumers, facilitating compliant and trusted data sharing. However, an underserved aspect is managing the end-of-life (EoL) of shared data, where standardized, machine-interpretable mechanisms for detailed EoL policies are lacking. To address this gap, we propose an extension of the Open Digital Rights Language (ODRL) to incorporate semantically rich EoL policies. This enables the specification of data deletion requirements, supporting legal and regulatory obligations. Data trustees can use these enhanced policies to coordinate EoL actions among all parties. The explicit semantics within these policies facilitate clearer accountability and support the creation of auditable logs by making EoL obligations machine-interpretable and unambiguous. Our ODRL extension has been evaluated by ODRL and data governance experts, ensuring its robustness and relevance for practical implementation. This work contributes to the standardization of EoL data management by analyzing and articulating the detailed requirements for EoL policies in the context of data trustees, and by proposing a specific ODRL extension to meet these requirements. For practitioners using ODRL, our extension provides enhanced, machine-interpretable EoL capabilities, improving compliance and trust.

1 INTRODUCTION

In times of constantly growing data volumes and ever more complex legal requirements, efficient end-of-life (EoL) data management is becoming increasingly important. This is particularly relevant with the widespread adoption of generative AI, which can produce vast quantities of data, further emphasizing the need for systematic data deletion strategies. A systematic approach to data deletion is essential (Tebernum and Howar, 2023), especially for data trustees who facilitate trusted and neutral data access between data providers and data consumers while complying with legal requirements (Specht-Riemenschneider and Kerber, 2022). In this process, data trustees must maintain the trust and interests of all stakeholders. Acting neutrally does not mean having no interests of their own; rather, it means balancing the various stakeholders' interests (Schinke et al., 2023), a core aspect emphasized in defining their trusted role (Lindner and Straub, 2023). To address these interests, data trustees can operate in a specific sector (e.g., the building sector) to meet stakeholder needs and understand data and regulatory requirements (Steinert et al., 2025). Incorporating EoL considerations offers data trustees technical advantages (such as improved resource and cost optimization, reduced attack surfaces, higher data quality (Tebernum and Howar, 2025)) and strengthens stakeholder trust (Steinert and Tebernum, 2025).

To ensure the protection and targeted use of data, data providers and consumers conclude data processing agreements with data trustees to govern data flows, enforce agreed-upon rules, and manage consent or usage rights (Stachon et al., 2023). These agreements (also called data contracts) contain comprehensive usage guidelines in which prohibitions, permissions, and duties are defined (Iannella and Villata, 2018). One way of making such policies machine-readable and automatically enforceable is to map them in Resource Description Framework (RDF) using Open Digital Rights Language (ODRL). Similar access and usage policy mechanisms are already

[a] https://orcid.org/0009-0008-3888-2092
[b] https://orcid.org/0000-0002-4772-9099

Steinert, M. and Tebernum, D.
Enhancing Data Governance in Data Trustees Through ODRL-Based End-of-Life Policies.
DOI: 10.5220/0013669800003985

In *Proceedings of the 21st International Conference on Web Information Systems and Technologies (WEBIST 2025)*, pages 27-38
ISBN: 978-989-758-772-6; ISSN: 2184-3252

being used in data spaces (Dam et al., 2023; Eitel et al., 2021), i.e., digital infrastructures that enable the secure, controlled, and sovereign exchange of data between different stakeholders (Otto et al., 2022). However, while ODRL provides a robust foundation for general usage policies, including a basic *odrl:delete* action, specifying the complex nuances required for verifiable and compliant EoL data management presents a challenge within the standard ODRL vocabulary. Specifically, important aspects such as the explicit rationale behind a deletion requirement (e.g., legal mandate vs. user consent withdrawal vs. data quality remediation), the mandated deletion method (e.g., logical deletion vs. cryptographic shredding), requirements for proof of deletion or the handling of data dependencies are not natively supported in a standardized, machine-interpretable way. Representing these EoL specifics often requires extensions or remains implicit within policy definitions. This semantic gap hinders interoperability and the automated traceability of deletion duties and auditable compliance trails, particularly within multi-stakeholder environments, such as data spaces, facilitated by data trustees.

The importance of EoL data management is underscored by external research, such as a survey of 35 data trustee experts, which found that respondents consider reliable and traceable data deletion to be a part of data trustees' activities[1]. Acknowledging these findings and the identified semantic gap in current policy languages, we aim to enhance EoL data governance for data trustees by addressing the following two research questions:

RQ1: What specific requirements must an EoL policy vocabulary fulfill to enable EoL data management by data trustees, particularly within data spaces?

RQ2: How can the ODRL standard be extended with a semantically rich vocabulary to meet these EoL requirements and facilitate machine-interpretable specification of deletion policies?

To answer these questions, we first analyze and articulate the requirements for a uniform, semantically rich vocabulary for data deletion policies suitable for data trustees. Subsequently, we propose an ODRL extension incorporating these requirements, drawing inspiration from existing EoL frameworks like the DestroyClaims concept (Tebernum and Howar, 2025)[2]. Our contribution focuses on defining the essential components of this vocabulary to facilitate machine-interpretable EoL policies. The aim is to establish a more comprehensive data governance model, enabling automated compliance with EoL duties and thereby reinforcing trust in data trustees.

The remainder of this paper is structured as follows: Section 2 provides the theoretical background on data trustees, data spaces, and the intricacies of EoL data management. Section 3 delves into the requirements for an EoL policy vocabulary tailored for data trustees, directly addressing RQ1, and also presents our proposed ODRL extension, detailing the vocabulary and associated mechanisms needed to fulfill these requirements, thereby answering RQ2. Section 4 describes the methodology and findings of our qualitative evaluation, focusing on an expert workshop to assess the feasibility and conceptual soundness of our ODRL extension. Section 5 discusses the implications, potential benefits, and limitations of our ODRL extension. Finally, Section 6 summarizes our findings and outlines directions for future work.

2 THEORETICAL FOUNDATION

This section first outlines the core concepts that underpin our work: data trustees and data spaces. Understanding these concepts is crucial for contextualizing our proposed ODRL extension for EoL data management. We also discuss the importance of EoL data management.

2.1 Data Trustees

Data trustees represent a specific form of data intermediary designed to act in the best interests of external stakeholders, namely data providers and data consumers, without pursuing commercial goals concerning the data itself (Specht-Riemenschneider and Kerber, 2022). Literature distinguishes between a narrower focus on fiduciary data management and a broader view emphasizing the trustee's role in facilitating trusted data sharing (Reiberg et al., 2023; Lindner and Straub, 2023). Their fundamental role is to serve as a trusted entity facilitating interactions between these parties (Schinke et al., 2023), ensuring neutrality and transparency in their operations (Lindner and Straub, 2023). By fulfilling this role, data trustees ensure that data access and utilization occur responsibly and compliantly.

Typical responsibilities of a data trustee include mediating between data supply and demand, supporting the technicalities of data exchange, and fostering trust among participants (Specht-Riemenschneider et al., 2021). They offer an alternative to platform models, thereby helping to counteract the formation of platform monopolies (European Commission, 2022). This contributes to market pluralization and

[1] https://doi.org/10.5281/zenodo.14992879

[2] https://github.com/DaTebe/destroyclaims

enables smaller or specialized actors to participate effectively in the data economy, which can strengthen market innovation (Shaharudin et al., 2024). The operations of data trustees are grounded in technical, legal, ethical, and organizational standards that guarantee security and reliability. Functionally, they can either hold the data directly or facilitate the direct sharing of data between the parties involved (Lauf et al., 2023). In either scenario, the objective is to safeguard the interests and rights of all participants, cultivate trust within the data ecosystem, and ensure fair and transparent data sharing (Stachon et al., 2023). Finally, data trustees should have domain knowledge, such as in the building sector, which can further enhance their ability to meet stakeholder needs and navigate specific regulatory requirements (Steinert et al., 2025).

2.2 Data Spaces

The concept emerged partly from the need to manage increasingly diverse and distributed data sources beyond traditional centralized databases (Franklin et al., 2005). Data space technology offers a technical foundation for implementing the functionalities required by data trustees, enabling secure and sovereign data sharing (Otto et al., 2022). CEN and CENELEC, the European bodies responsible for developing technical standards, define a data space as an "interoperable framework, based on common governance principles, standards, practices, and enabling services, that enables trusted data transactions between participants" (European Committee for Standardization, 2024). The goal of data spaces is to facilitate the controlled sharing and cooperative use of data between different organizations, irrespective of the data types involved, while ensuring data providers retain sovereignty over their assets (Otto, 2022). Technical implementations often follow architectures like the International Data Spaces Reference Architecture Model (IDSA RAM) (Otto et al., 2019), which outlines components and layers for secure data sharing. Consequently, data spaces empower organizations and individuals to gain deeper insights from their data, thereby improving decision-making processes and fostering innovation (Bacco et al., 2024).

2.3 The Interplay of Data Trustees and Data Spaces

Data trustees and data spaces are complementary concepts (Steinert and Altendeitering, 2024). In addition, the broader category of data intermediaries, which includes data trustees, is also seen as part of the role model of data spaces (Gemein et al., 2023). Data trustees provide the organizational, legal, and trust-building framework necessary to govern sovereign data sharing according to agreed rules and policies. Conversely, data spaces provide the technical infrastructure required for secure, efficient, and sovereign data management and sharing. The data trustee establishes the governance and trust layer by defining the rules of engagement (often codified in ODRL policies), while the data space provides the technical means to enforce these rules, manage data flows, and facilitate secure interaction. This interplay ensures that data can be protected effectively while being utilized purposefully, benefiting all participants and contributing to an innovative and competitive data economy. Our work enhances this interplay by introducing EoL data management within data trustees that operate within or interact with data spaces.

2.4 End-of-Life Data Management

Data deletion, the deliberate and often irreversible action of removing or obliterating information, constitutes the final stage of the data lifecycle (Tebernum and Howar, 2023) (also referred to as EoL data management). It transcends the simplicity implied by a 'delete' command, representing a necessary process driven by multifaceted requirements and conveying significant implicit meaning. Despite its increasing importance, research indicates this area remains under-addressed compared to other lifecycle stages (Tebernum et al., 2021, 2023). Understanding the depth of data deletion is paramount, particularly within trust-based ecosystems facilitated by data trustees.

Controlled data deletion ensures compliance with legal and regulatory requirements, such as the GDPR's "right to be forgotten" (European Commission, 2016), mitigating significant legal and financial risks. It also improves security by reducing the attack surface; removing redundant or sensitive data minimizes the potential damage from breaches. Moreover, it maintains data quality (Wang and Strong, 1996) and utility, ensuring that decisions are based on current, relevant, and accurate information, which is important, for example, when needing to remove specific data points from trained models (Ginart et al., 2019). Furthermore, it optimizes resource utilization, reducing storage costs and potentially improving system performance, aligning with principles of efficient resource management and Green IT (Van Bussel and Smit, 2014). Finally, ethical considerations, organizational policies, or evolving risk assessments may require data removal.

Given its destructive and irreversible nature, data deletion demands a systematic and precise approach. A simple instruction to delete is often insufficient. Effective deletion requires detailed specifications, including: precise identification of the target data (e.g., specific records, files identified by hashes, data within certain systems); explicit documentation of the rationale for deletion (e.g., compliance, user request, quality issue); definition of preconditions (e.g., temporal constraints, geographical location, completion of related processes); specification of the deletion method (e.g., logical deletion, cryptographic wiping, physical destruction); assignment of responsibility; and implementation of safeguards (e.g., simulations, explicit approvals) to prevent errors.

Crucially, the action of deletion, particularly when formally specified and recorded, carries inherent meaning beyond the mere absence of data. The documented specification itself becomes valuable metadata. The stated *rationale* reveals the intent behind the deletion (compliance, risk reduction, etc.). The decision to delete implies a *value assessment* - the data is no longer deemed necessary, useful, or its retention poses unacceptable risk according to defined policies. Specified *conditions* provide context about operational workflows or regulatory constraints. The chosen *deletion method* reflects the perceived sensitivity of the data and the required level of assurance. Furthermore, the existence of a structured, documented deletion process signifies the organization's *governance maturity* (Tallon et al., 2013) and commitment to the responsibilities of data trustees, providing an auditable trail that is important for accountability.

However, standard policy languages like ODRL, with basic actions such as *odrl:delete*, lack the semantic depth required for comprehensive EoL data management. This gap in expressiveness, particularly in defining the rich context of deletion for traceable and auditable processes, motivates the ODRL extension detailed in this work.

3 PROPOSED ODRL EXTENSION FOR EOL POLICIES

Building upon the analysis of data trustee needs (Section 2) and the identified semantic limitations of standard ODRL for EoL management (Section 2.4), this section details our ODRL extension. Our goal is to create a semantically rich and machine-interpretable vocabulary that enables the precise specification, automated processing, and auditing of data deletion policies by data trustees, particularly within data spaces.

This ODRL extension addresses **RQ1: What specific requirements must an EoL policy vocabulary fulfill to enable EoL data management by data trustees, particularly within data spaces?** by fulfilling the specific requirements derived from our analysis. It then answers **RQ2: How can the ODRL standard be extended with a semantically rich vocabulary to meet these EoL requirements and facilitate machine-interpretable specification of deletion policies?** through the proposed *eol:* vocabulary and its mechanisms. Specifically, the ODRL extension is designed to fulfill the following requirements:

- **REQ1: Semantic Richness for Rationale:** Explicitly capture the reason for the EoL action (e.g., GDPR request, contract end, data quality).
- **REQ2: Specification of Deletion Method:** Allow specification of the required deletion method (e.g., logical delete, wipe, physical destruction).
- **REQ3: Granular Data Identification:** Enable precise identification of data assets, potentially beyond URIs (e.g., via hashes, record IDs).
- **REQ4: Clear Temporal and Spatial Specification:** Support unambiguous temporal and spatial constraints for deletion actions.
- **REQ5: Machine Interpretability and Automation:** Be defined formally (e.g., RDF) for unambiguous machine interpretation and automation.
- **REQ6: Support for Traceability and Auditing:** Ensure traceability and auditability of EoL actions (e.g., via logs or confirmations) to enable confirmation processes.
- **REQ7: Compatibility and Extensibility within ODRL:** Integrate seamlessly with ODRL and follow its design principles as a modular extension.
- **REQ8: Contextual Applicability for Data Trustees:** Be suitable for multi-party scenarios managed by data trustees within data spaces.
- **REQ9: Operational Policy Management Features:** Incorporate attributes for practical policy management at the rule level, allowing for control over execution behavior.

To meet these requirements, we propose defining this ODRL extension using RDF, introducing a new vocabulary under the namespace *eol:*.

3.1 The *eol:* Vocabulary Extension

The *eol:* vocabulary complements standard ODRL by adding specific classes and properties tailored for EoL data management, addressing the requirements outlined above.

3.1.1 Rule-Level Operational Attributes (Addressing REQ9)

To provide operational control over individual rules within a policy, we introduce several properties applicable directly to instances of *odrl:Rule* (such as an *odrl:obligation* or *odrl:permission*):

- *eol:simulated* (*xsd:boolean*): Indicates if the rule execution should only be simulated (e.g., for testing or impact analysis) rather than enacted.
- *eol:notification* (*xsd:boolean*): Specifies if the assignee should receive a notification before the data associated with the rule is deleted.
- *eol:optIn* (*xsd:boolean*): Indicates whether the deletion of data as per the rule must be explicitly accepted (opted-in) by the assignee, rather than being processed automatically.
- *eol:strict* (*xsd:boolean*): Mandates that rule execution should only proceed automatically if the processing system understands all classes, properties, and values within the rule. This prevents unintended actions due to partial interpretation.

3.1.2 Specifying WHAT to Delete: The *eol:Locator* (Addressing REQ3)

Standard ODRL uses *odrl:target* to identify the asset. However, EoL actions often require more granular identification, especially for distributed data or specific data fragments. We introduce:

- *eol:Locator* (Class): Represents a detailed description of an asset's location or identity beyond a simple URI. Instances of *eol:Locator* can have various specific properties to accommodate different types of location or identification information (e.g., content hashes, file paths, persistent identifiers, database UIDs, URLs).
- *eol:hasLocator* (Property): Links an *odrl:Asset* (and therefore also an *odrl:AssetCollection*, as it is a subclass) to one or more *eol:Locator* instances. Its domain is explicitly set to *odrl:Asset*, allowing locators for both individual assets and collections.
- *eol:locatorSHA256Hash* (Property): This property serves as one example of how an asset can be identified within an *eol:Locator* instance, specifically using its SHA256 content hash (*xsd:string*). Other properties could be defined for different identification schemes.

This allows policies to precisely target specific data instances or collections for deletion, even if individual items lack persistent URIs, using precise identifiers like content hashes or other suitable locator properties.

3.1.3 Specifying WHY: The *eol:Reason* (Addressing REQ1)

Understanding the rationale behind a deletion requirement is important for compliance, auditing, and appropriate handling (e.g., notifying stakeholders). We introduce:

- *eol:Reason* (Class): Represents the justification for the policy, rule, or action. Instances of this class would typically be URIs representing concepts like *eol:reason:LegalObligation*, *eol:reason:UserConsentWithdrawal*, *eol:reason:DataQualityIssue*, etc.
- *eol:hasReason* (Property): Links an *odrl:Rule* (or an *odrl:Action* within it) to an instance of *eol:Reason*, providing machine-interpretable context.

3.1.4 Specifying HOW: *eol:destructionMethod* and Refinements (Addressing REQ2)

The core *odrl:delete* action lacks specificity regarding the deletion method. We leverage ODRL's *odrl:refinement* mechanism within a *odrl:Constraint* attached to the rule governing the *odrl:delete* action. We introduce:

- *eol:destructionMethod* (LeftOperand): Defined as an *odrl:LeftOperand*, this represents the concept of the required method or level of destruction. It is used as the *odrl:leftOperand* in a refinement constraint associated with an action (like *odrl:delete*).
- Specific Destruction Method Identifiers (RightOperands): The required method is specified using an IRI as the *odrl:rightOperand* in the refinement constraint (typically with *odrl:operator odrl:eq*). These IRIs can represent:
  - Abstract destruction levels, such as *eol:method:Recycled* (implying data might be restorable), *eol:method:Deleted* (standard logical deletion), *eol:method:Wiped* (data is overwritten and not practically restorable), or *eol:method:PhysicallyDestroyed* (the storage medium itself is destroyed). These abstract levels of recoverability are inspired by Cantrell and Runs Through (2019).
  - Concrete algorithms or techniques, for example, an IRI representing a specific wiping algorithm (e.g., a Gutmann method variant) or a destruction level from standards such as NIST SP 800-88, ISO/IEC 27040 or DIN 66399.

While the specific *eol:* vocabulary elements detailed above directly address REQ1 (Rationale), REQ2 (Method), REQ3 (Identification), and REQ9 (Operational Attributes), the overall design, based on RDF and ODRL's extensibility, inherently supports other requirements. Standard ODRL temporal (*odrl:dateTime*) and spatial (*odrl:spatial*) constraints are leveraged to meet REQ4 (Temporal and Spatial Specification). The use of RDF ensures REQ5 (Machine Interpretability), which, combined with the explicit semantics, provides a foundation for REQ6 (Traceability and Auditing). Compatibility (REQ7) is maintained by adhering to ODRL patterns. How these elements integrate to support data trustee operations in multi-party contexts (REQ8) is further detailed in Section 3.3.

3.2 Illustrative Example EoL Policy

Listing 1 demonstrates how these extensions integrate within an ODRL policy, reflecting the structure from our formal definition (provided in the Appendix of this paper) and incorporating terms shown in the following JSON-LD structure. The policy references an ODRL profile (via the *odrl:profile*), which defines or imports the *eol:* vocabulary terms and their usage. This ensures that all parties understand the semantics of the EoL extension used. It shows an *odrl:Agreement* where a data trustee assigns an obligation to a data consumer. The obligation rule within the policy specifies that its execution will not be simulated (*eol:simulated* is false), requires strict interpretation (*eol:strict* is true), mandates assignee opt-in for deletion (*eol:optIn* is true, meaning it is not fully automated without acceptance), and triggers a notification to the assignee (*eol:notification* is true). The obligation is to delete (*odrl:delete*) an asset (*odrl:Asset*, which uses *eol:hasLocator* to refer to a *eol:locatorSHA256Hash*, triggered by user consent withdrawal (*eol:hasReason* pointing to *eol:reason:UserConsentWithdrawal*). This obligation is constrained temporally, active from January 1, 2025, and expiring by December 31, 2026, and spatially limited to the EU (via an *odrl:spatial* isA constraint). Furthermore, the deletion action is refined to require a specific destruction method (*eol:destructionMethod* equal to a URI representing the Gutmann method).

3.3 Integration Within Data Trustees

This extended ODRL vocabulary integrates into the operational workflows of data trustees, potentially within data spaces. While this section outlines how this integration would function, its concrete implementation and empirical validation are planned for future work. Data trustees use these policies in data contracts to specify EoL duties. Technical components (e.g., a data space connector enhanced with policy enforcement capabilities) can parse and act on these machine-interpretable policies (REQ5). The explicit semantics, such as *eol:hasReason* and the specific *eol:destructionMethod*, along with rule-level operational attributes like *eol:simulated* or *eol:notification*, guide implementation (e.g., triggering specific deletion scripts or workflows) and facilitate targeted notifications or compliance checks (REQ9). Crucially, this structure enables automated logging and auditing based on policy terms (e.g., recording when a policy with a specific reason and method was executed), enhancing traceability and auditing (REQ6). This approach aligns with ODRL principles (REQ7) and directly supports the data trustee's governance role in managing multi-party data lifecycles (REQ8).

In summary, this proposed ODRL extension, characterized by the *eol:* vocabulary and driven by the identified requirements, provides a concrete mechanism to facilitate expressive, machine-interpretable EoL policies, thereby strengthening data governance for data trustees.

4 EVALUATION

To assess the feasibility and appropriateness of the proposed ODRL extension for EoL policies, and to validate the underlying requirements, a qualitative evaluation study was conducted. Nine experts participated in semi-structured interviews. These experts were selected to represent a diverse range of relevant domains: one data scientist, two mobility and smart city experts, one manufacturing expert, one cloud infrastructure expert, and four data space experts. Crucially, all participants had knowledge of and practical experience with data governance and policy management, as well as expertise with ODRL. The interviews were guided by a framework covering six evaluation criteria: Relevance & Completeness of Requirements (addressing RQ1), Suitability of the Solution Approach, Semantic Clarity & Expressiveness, Technical Feasibility & Integration, and Potential Impact & Usefulness (all addressing RQ2). Interview transcripts were analyzed using qualitative content analysis to identify recurring themes, patterns, and critical insights regarding the proposed ODRL extension, following established procedures for deductive content analysis and systematic qualitative analysis of text data (McKibben et al., 2022; Puppis, 2019).

```
{
    // Context definitions omitted to save space
    "@type": "Agreement",
    "uid": "urn:policy:example01-delete-obligation",
    // Profile would define or import the eol: vocabulary terms and their usage
    "odrl:profile": { "@id": "http://example.com/odrl/profile/data-trustee" },
    "dct:description": { "@value": "Delete user data (consent withdrawal)" },
    "dct:issued": "2025-01-01",
    // REQ7: Compatibility / REQ5: Interpretability / REQ6: Auditing
    "obligation": [{
        "@id": "urn:policy:example01-delete-obligation#obligation",
        // REQ1: WHY - Semantic Richness for Rationale
        "eol:hasReason": { "@id": "eol:UserConsentWithdrawal" },
        // REQ9: Operational Policy Management Features
        "eol:simulated": false,
        "eol:notification": true,
        "eol:optIn": true,
        "eol:strict": true,
        // REQ8: Contextual Applicability (Data Trustee Context)
        "assigner": "urn:party:trustee",
        "assignee": "urn:party:consumer",
        "target": {
            "@type": "Asset",
            "@id": "urn:asset:userXYZ",
            "dct:title": "User XYZ Data Set",
            "eol:hasLocator": {
                "@type": "eol:Locator",
                // Locator can use various properties; this is one example
                "eol:locatorSHA256Hash": "db2dc5..."
            }
        },
        "action": {
            "@id": "odrl:delete",
            "refinement": [{
                // REQ2: HOW - Refinement specifying the Deletion Method
                "leftOperand": "eol:destructionMethod",
                "operator": "odrl:eq",
                "rightOperand": "http://example.com/odrl-eol/method/gutmann-method"
            }]
        }
        // REQ4: WHEN & WHERE - Constraints on the Obligation's applicability
        "constraint": [
            {
                "leftOperand": "odrl:dateTime",
                "operator": "odrl:gteq",
                "rightOperand": "2025-01-01"
            },
            {
                "leftOperand": "odrl:spatial",
                "operator": "odrl:isA",
                "rightOperand": "urn:location:EU"
            },
            {
                "leftOperand": "odrl:dateTime",
                "operator": "odrl:lteq",
                "rightOperand": "2026-12-31",
                "skos:note": "Rule Expiration Date"
            }
        ]
    }]
}
```

Listing 1: Example ODRL Policy with EoL Extensions in JSON-LD.

4.1 Relevance & Completeness of Requirements (RQ1)

There was a strong consensus among the experts regarding the relevance of addressing EoL data management with more granularity than standard ODRL provides. The identified set of nine requirements (REQ1-REQ9) was generally perceived as comprehensive and well-founded for specifying EoL policies in the context of data trustees.

Semantic Aspects (REQ1, REQ2, REQ3): The need to specify the rationale (REQ1) for deletion was acknowledged for context and auditability, although some experts questioned its direct technical necessity for policy execution engines. Specifying the deletion method (REQ2) was deemed crucial for compliance and operational clarity, prompting discussions on the technical capabilities required by the executing party. The requirement for granular data identification (REQ3) beyond simple URIs, potentially using hashes or internal identifiers via the proposed *eol:Locator*, was strongly supported as a necessary enhancement, although challenges regarding consistency of identifiers across systems were noted.

Contextual & Technical Aspects (REQ4, REQ5, REQ6, REQ7): Leveraging standard ODRL for temporal and spatial constraints (REQ4) was considered pragmatic. The fundamental need for machine interpretability (REQ5) through a formal RDF structure was undisputed. Supporting traceability and auditing (REQ6) was recognized as a critical goal, though experts highlighted the practical limitations and dependency on the logging mechanisms and trustworthiness of the executing environment. Ensuring compatibility with ODRL (REQ7) by using an extension vocabulary was seen as the correct approach.

Operational & Ecosystem Aspects (REQ8, REQ9): The suitability for multi-party scenarios typical for data trustees (REQ8) was confirmed by the inherent structure of ODRL for defining roles (assigner/assignee). The operational policy management features (REQ9), such as those for simulation, notification, opt-in, and strict interpretation at the rule level, were particularly well received and considered highly valuable for practical policy deployment and control. Discussions arose regarding the optimal placement of these attributes (policy vs. rule level) depending on the desired scope (e.g., EoL as part of a larger usage policy).

Potential Gaps Identified: A recurring theme was the challenge of addressing the EoL of derived data products (e.g., aggregated datasets, trained AI models), which was considered important but potentially beyond the scope of this initial vocabulary. The need for explicit notification or confirmation mechanisms upon successful deletion was also suggested by some participants, a point partially addressed by the *eol:notification* flag. Furthermore, a clearer definition of responsibilities for policy enforcement and verification within the ecosystem was deemed necessary.

4.2 Evaluation of the ODRL Extension (RQ2)

Suitability of Approach: Extending ODRL was broadly accepted as a suitable and pragmatic approach, given ODRL's prevalence in data spaces and related initiatives. Its flexibility was seen as an advantage, while its known complexity and lack of standardized tooling were acknowledged as general challenges. The proposed *eol:* vocabulary was found to be logically structured and consistent with the identified requirements.

Semantic Clarity & Expressiveness: The vocabulary was perceived as clear and understandable, particularly when presented with the policy examples. Experts confirmed its ability to effectively express the EoL nuances (reason, method, granular target, operational controls), representing a significant improvement over standard ODRL for this purpose. The limitation regarding derived data was reiterated here.

Technical Feasibility & Integration: While the vocabulary and its ODRL integration were considered sound, experts identified several practical implementation challenges. These primarily centered on the need to develop specific policy engine logic to interpret and act upon the *eol:* terms, the mapping of *eol:Locator* information to diverse target systems, ensuring target systems support the specified deletion methods, and potential limitations in existing APIs (e.g., for custom attributes). However, the overarching challenge of policy enforcement - ensuring that the specified actions are actually carried out reliably and verifiably in the target environment - was emphasized by almost all participants as a hurdle, independent of the policy language itself.

Potential Impact & Usefulness: The experts saw potential value in the proposed ODRL extension. Benefits highlighted included enhanced compliance support, improved data quality management, increased transparency in data handling, and fostering trust within data ecosystems, particularly for the governance role of data trustees. While the potential to improve current EoL practices was acknowledged, concerns about widespread, rapid adoption were also expressed, citing the general inertia of detailed policy implementation and the reliance on perceived need

and regulatory pressure.

General Acceptance & Suggestions: Overall feedback was positive, confirming the need and general direction of the proposed ODRL extension. Strengths were seen in the semantic richness and operational control features. The main perceived weakness is the reliance on effective enforcement mechanisms. Suggestions for improvement included considering more flexible placement of operational attributes and possibly including notification mechanisms. The critical role of clear governance frameworks defining EoL standards and responsibilities within data ecosystems was repeatedly emphasized.

4.3 Evaluation Summary

The expert interviews broadly confirmed the identified requirements for enhanced EoL data management (RQ1) and confirmed the suitability of the proposed ODRL extension (RQ2) to address these needs. The *eol:* vocabulary was found to be semantically expressive for EoL concepts and operationally valuable. However, the evaluation emphasized that the success of such a policy framework depends on addressing the persistent challenge of policy enforcement and establishing clear governance structures within the target data ecosystems. While the proposed *eol:* vocabulary provides the necessary means to specify detailed EoL policies, ensuring their implementation requires complementary technical and organizational measures. The results provide valuable input for refining the vocabulary and underscore the importance of integrating policy definition with robust enforcement and governance strategies in future work.

5 DISCUSSION

The introduction of the *eol:* vocabulary as an ODRL extension for EoL data management by data trustees presents several practical and theoretical implications. This section discusses these implications, the benefits, and limitations of the proposed approach, and its broader significance for data governance, building on the validation from our expert evaluation (Section 4).

Practical Implications: Leveraging ODRL for EoL Management. A design choice was to extend ODRL rather than proposing a new policy language. This decision carries practical advantages. Data trustees and participants in data spaces are familiar with ODRL for defining access and usage policies (Dam et al., 2023; Eitel et al., 2021). Integrating EoL specifications into ODRL avoids the overhead of developing, learning, implementing, and executing a separate language and its associated tooling. Systems already capable of parsing and interpreting ODRL (e.g., data space connectors) can potentially be adapted to handle the *eol:* extension, lowering the barrier to adoption. This pragmatic approach directly supports REQ7 (ODRL Compatibility) and enhances the likelihood of practical implementation within existing data governance frameworks. While conceptual integration into the ODRL model is straightforward, practical implementation of interpretation engines requires more work. Here, it is necessary that, over time, practical experience feeds back into research to identify the most important aspects, so that robust interpretation engines can be developed on this basis.

Towards Holistic Data Lifecycle Governance with ODRL. ODRL focuses on the phases of data access and usage. Our extension incorporates the final phase of the data lifecycle. By enabling the specification of *why* data should be deleted (*eol:Reason*), *how* it should be deleted (*eol:destructionMethod*), and precisely *what* data is targeted (*eol:Locator*), ODRL becomes a more comprehensive language for governing data throughout its entire lifecycle. This aligns with the need for holistic data lifecycle management (Tebernum and Howar, 2023) and moves ODRL towards being a language capable of expressing policies that span from data creation and sharing through to its eventual, traceable deletion. While inspired by concepts seen in frameworks like DestroyClaims[3], our integration within the ODRL standard facilitates broader interoperability and standardization potential.

Theoretical Implications: Formalizing EoL as Policy. Beyond the practical benefits, our work demonstrates the theoretical feasibility and utility of modeling EoL actions declaratively as machine-interpretable policies. Traditionally, data deletion might be handled through procedural scripts, manual processes, or implicit understandings. Formalizing EoL duties within ODRL elevates these actions to the level of explicit governance rules that are suitable for automated reasoning, traceability, and auditing. The ability to capture the meaning behind deletion (as discussed in Section 2.4) via properties such as *eol:Reason* adds semantic depth, transforming a simple delete command into a rich, context-aware governance instruction. This formalization is important for building trustworthy data trustees, where transparency and accountability are paramount.

Addressing the Need for a Uniform EoL Vocabulary and Standardization. The lack of a standardized vocabulary for EoL actions represents a significant barrier to reliable and verifiable data deletion across organizational boundaries, as highlighted in

[3]https://github.com/DaTebe/destroyclaims

our problem statement. Ambiguity regarding deletion requirements hinders automated enforcement, particularly when obligating data consumers to perform deletion, and complicates compliance auditing. Our work tackles this by systematically identifying the necessary semantic components (REQ1-REQ9) and proposing a concrete vocabulary (*eol:*) integrated within ODRL. While proposing a full W3C standard is beyond the scope of this initial work, the definition and validation of these requirements, coupled with the demonstration of their feasibility using ODRL extensions, represent a novel and necessary contribution. It lays the crucial groundwork for future standardization efforts, providing a candidate vocabulary that aims to foster a common understanding and improve trust and compliance regarding EoL duties in data spaces.

Facilitating Enforceability and Auditing. A persistent challenge in data sharing is ensuring that data consumers adhere to usage policies, including deletion duties. While our ODRL extension cannot guarantee technical enforcement on a consumer's system (which remains dependent on their implementation and willingness to comply), it facilitates enforceability and auditability. The machine-interpretable nature (REQ5) allows consumer systems to automatically recognize and process deletion duties. Explicit requirements for deletion methods (REQ2) and reasons (REQ1) clarify expectations. Crucially, the structured policy enables better auditing (REQ6). Data trustees can log the issuance of EoL policies, and mechanisms could be envisioned (potentially linked to data space components like clearing houses) to track acknowledgments or require confirmations of deletion from consumers, based on the policy rules. This provides a verifiable trail, enhancing accountability even if direct technical enforcement is limited.

Supporting Legal Compliance and Risk Management. The proposed ODRL extension ensures compliance with various legal and regulatory frameworks, including the GDPR and CCPA. The ability to specify the legal basis for deletion (e.g., *eol:UserConsentWithdrawal*, *eol:LegalObligation*) and the required deletion standard aligns with data protection principles. Furthermore, the rule-level operational attributes introduce risk management features. *eol:simulated* allows for impact analysis before actual deletion, mitigating the risk of accidental data loss. *eol:strict* prevents automated execution if the policy semantics are not fully understood by the processing system, reducing the chance of incorrect actions. *eol:optIn* ensures assignee consent for deletion when required, and *eol:notification* provides transparency. However, it is important to recognize that while these features help manage risk, they do not eliminate operational risk entirely; robust backup and recovery strategies, along with careful policy authoring and review processes, remain essential.

Limitations and Future Directions. Despite the positive feedback on our evaluation, we acknowledge the limitations of our approach. First, the challenge of technical enforcement on consumer systems remains. While our ODRL extension improves the clarity and auditability of the deletion duty, verifiable deletion still relies on consumer cooperation and additional technical mechanisms. Second, handling cases across multiple jurisdictions presents governance challenges that the *eol:* vocabulary cannot solve alone. Third, extending ODRL is pragmatic but adds complexity to the standard. This requires careful consideration of tooling support and potential interactions with other ODRL profiles. Fourth, the *eol:* vocabulary may require refinement to address more complex scenarios, such as intricate data dependencies, the need for specific proof-of-deletion artifacts, and partial data anonymization as an alternative to deletion. Finally, the success of our ODRL extension depends on community adoption and integration into data spaces. Further dissemination, the development of best practices, and pilot implementations are necessary next steps.

6 CONCLUSION

This work addressed the gap in EoL data management within data trustee operations, particularly in the context of data spaces, by answering our research questions concerning the requirements for (RQ1) and the implementation of (RQ2) an expressive EoL policy vocabulary. Our proposed solution extends the ODRL standard with an *eol:* vocabulary, enabling the machine-interpretable specification of deletion policies. The qualitative evaluation via an expert workshop (Section 4) provided initial validation, indicating that the identified requirements are relevant and the proposed ODRL extension is perceived as a conceptually sound, technically feasible, and potentially impactful approach.

Building on this, we have shown that effective EoL data management is essential for data trustees, yet current policy languages such as ODRL lack the semantic depth to specify detailed, machine-interpretable EoL policies. This paper tackled this issue by first identifying the requirements for an EoL policy vocabulary (RQ1), covering the why (rationale), how (method), what (data identification), when/where (context), and operational aspects of EoL policies. We then proposed the *eol:* vocabulary, an ODRL extension designed to meet these require-

ments, thereby enabling the machine-interpretable specification of EoL policies (RQ2). Our qualitative evaluation with experts validated the relevance of these requirements and the conceptual soundness of the ODRL extension, underscoring its potential to enhance compliance, trust, and overall data governance.

The primary contribution, therefore, is the formalization of EoL policy requirements tailored for data trustees and a practical vocabulary to implement them. This approach facilitates more holistic data lifecycle governance by rendering EoL policies explicit, machine-interpretable, and auditable. While this work improves policy specification, ensuring technical enforcement in distributed systems and establishing robust ecosystem governance remain challenges. Future work should focus on refining the *eol:* vocabulary for data deletion, exploring mechanisms for verifiable EoL execution, and developing best practices for its integration within data spaces. As another part of EoL data management involves not only data deletion but also data retention, future work could also include the development of a complementary ODRL extension to specify retention policies. By promoting standardized, expressive policies for the full scope of EoL data management, encompassing both deletion and retention, we aim to further strengthen data governance and trust in data trustees.

REFERENCES

Bacco, M., Kocian, A., Chessa, S., Crivello, A., and Barsocchi, P. (2024). What are data spaces? systematic survey and future outlook. *Data in Brief*, 57:110969.

Cantrell, G. and Runs Through, J. (2019). The Five Levels of Data Destruction: A Paradigm for Introducing Data Recovery in a Computer Science Course . In *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*, pages 133–138. IEEE Computer Society.

Dam, T., Krimbacher, A., and Neumaier, S. (2023). Policy patterns for usage control in data spaces.

Eitel, A., Jung, C., Brandstädter, R., Hosseinzadeh, A., Bader, S., Kühnle, C., Birnstill, P., Brost, G., Gall, M., Bruckner, F., Weißenberg, N., and Korth, B. (2021). Usage control in the international data spaces.

European Commission (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

European Commission (2022). European data governance act.

European Committee for Standardization (2024). Trusted data transaction.

Franklin, M., Halevy, A., and Maier, D. (2005). From databases to dataspaces: a new abstraction for information management. *SIGMOD Rec.*, 34(4):27–33.

Gemein, O.-G., Hilberg, S. J., Lähteenoja, V., and Turpeinen, M. (2023). Reflections on the dga and data intermediaries.

Ginart, A., Guan, M. Y., Valiant, G., and Zou, J. (2019). Making ai forget you: Data deletion in machine learning.

Iannella, R. and Villata, S. (2018). ODRL information model 2.2.

Lauf, F., Scheider, S., Friese, J., Kilz, S., Radic, M., and Burmann, A. (2023). Exploring design characteristics of data trustees in healthcare - taxonomy and archetypes. In *Proceedings of the 31st European Conference on Information Systems (ECIS)*.

Lindner, M. and Straub, S. (2023). Datentreuhänderschaft status quo und entwicklungsperspektiven. Whitepaper, Begleitforschung Smarte Datenwirtschaft.

McKibben, W. B., Cade, R., Purgason, L. L., and and, E. W. (2022). How to conduct a deductive content analysis in counseling research. *Counseling Outcome Research and Evaluation*, 13(2):156–168.

Otto, B. (2022). The evolution of data spaces. In *Designing data spaces: The ecosystem approach to competitive advantage*, pages 3–15. Springer International Publishing Cham.

Otto, B., Steinbuss, S., Teuscher, A., and Lohmann, S. (2019). IDS Reference Architecture Model. Technical report, International Data Spaces Association.

Otto, B., ten Hompel, M., and Wrobel, S., editors (2022). *Designing Data Spaces: The Ecosystem Approach to Competitive Advantage*. Springer Nature.

Puppis, M. (2019). *Analyzing Talk and Text I: Qualitative Content Analysis*, pages 367–384. Springer International Publishing, Cham.

Reiberg, A., Appelt, D., Smoleń, A., and Kraemer, P. (2023). Datentreuhänder, datenvermittlungsdienste und gaia-x. Whitepaper, Gaia-X Hub Deutschland.

Schinke, L., Hoppen, M., Atanasyan, A., Gong, X., Heinze, F., Stollenwerk, K., and Roßmann, J. (2023). Trustful data sharing in the forest-based sector - opportunities and challenges for a data trustee. In *VLDB Workshops*.

Shaharudin, A., van Loenen, B., and Janssen, M. (2024). Exploring the contributions of open data intermediaries for a sustainable open data ecosystem. *Data & Policy*.

Specht-Riemenschneider, L., Blankertz, A., Sierek, P., Schneider, R., Knapp, J., and Henne, T. (2021). *Die Datentreuhand: ein Beitrag zur Modellbildung und rechtlichen Strukturierung zwecks Identifizierung der Regulierungserfordernisse für Datentreuhandmodelle*. Number 6 Beilage.

Specht-Riemenschneider, L. and Kerber, W. (2022). *Designing Data Trustees – A Purpose-Based Approach/ Datentreuhänder – Ein problemlösungsorientierter Ansatz*. Konrad-Adenauer-Stiftung e. V.

Stachon, M., Möller, F., Guggenberger, T., Tomczyk, M., and Henning, J.-L. (2023). Understanding data trusts. In *Proceedings of the 31st European Conference on Information Systems (ECIS)*.

Steinert, M. and Altendeitering, M. (2024). Data trustees: A whitelisting approach for trusted data sharing. page 86–92. Association for Computing Machinery.

Steinert, M., Schleimer, A. M., Altendeitering, M., and Hick, D. (2025). Designing data trustees: A prototype in the building sector. In *Proceedings of the 11th International Conference on Information Systems Security and Privacy - Volume 1*, pages 157–166.

Steinert, M. and Tebernum, D. (2025). Questionnaire and survey results: Design and technical requirements for data trustees.

Tallon, P. P., Ramirez, R. V., and Short, J. E. (2013). The information artifact in it governance: Toward a theory of information governance. *Journal of Management Information Systems*, 30(3):141–178.

Tebernum, D., Altendeitering, M., and Howar, F. (2021). Derm: A reference model for data engineering. In *Proceedings of the 10th International Conference on Data Science, Technology and Applications - DATA*, pages 165–175. INSTICC, SciTePress.

Tebernum, D., Altendeitering, M., and Howar, F. (2023). A survey-based evaluation of the data engineering maturity in practice. In *Data Management Technologies and Applications*, pages 1–23, Cham. Springer Nature Switzerland.

Tebernum, D. and Howar, F. (2023). Structuring the end of the data life cycle. In *Proceedings of the 12th International Conference on Data Science, Technology and Applications - DATA*, pages 207–218. INSTICC, SciTePress.

Tebernum, D. and Howar, F. (2025). Treating the end of the data life cycle as a first-class citizen in data engineering. In *Solutions and Technologies for Responsible Digitalization*, pages 237–252, Cham. Springer Nature Switzerland.

Van Bussel, G. and Smit, N. (2014). Building a Green Archiving Model: Archival Retention Levels, Information Value Chain and Green Computing. In *Proceedings of the 8th European Conference on IS Management and Evaluation (ECIME 2014)*, volume I, pages 271–277. ACPI.

Wang, R. Y. and Strong, D. M. (1996). Beyond accuracy: what data quality means to data consumers. *J. Manage. Inf. Syst.*, 12(4):5–33.

APPENDIX

```
@prefix odrl: <www.w3.org/ns/odrl/2/> .
@prefix eol: <www.example.com/odrl-eol#> .
@prefix rdf: <www.w3.org/1999/02/22-rdf-synt
    ax-ns#> .
@prefix rdfs: <www.w3.org/2000/01/rdf-schema
    #> .
@prefix skos: <www.w3.org/2004/02/skos/core#
    > .
@prefix owl: <www.w3.org/2002/07/owl#> .
@prefix xsd: <www.w3.org/2001/XMLSchema#> .
@prefix dct: <purl.org/dc/terms/> .
# Policy-Level Attributes (REQ9)
eol:simulated a rdf:Property, owl:Data
    typeProperty, skos:Concept ;
    rdfs:label "Simulated"@en ;
    skos:definition "Rule execution is simul
        ated if true."@en ;
    rdfs:domain odrl:Rule ;
    rdfs:range xsd:boolean .
eol:notification a rdf:Property, owl:Data
    typeProperty, skos:Concept ;
    rdfs:label "Notification"@en ;
    skos:definition "Assignee is notified
        before data deletion if true."@en ;
    rdfs:domain odrl:Rule ;
    rdfs:range xsd:boolean .
eol:optIn a rdf:Property, owl:Data
    typeProperty, skos:Concept ;
    rdfs:label "Opt-In Required"@en ;
    skos:definition "Assignee must
        explicitly opt-in for deletion if
        true."@en ;
    rdfs:domain odrl:Rule ;
    rdfs:range xsd:boolean .
eol:strict a rdf:Property, owl:Data
    typeProperty, skos:Concept ;
    rdfs:label "Strict"@en ;
    skos:definition "Automated execution
        only if all terms are known."@en ;
    rdfs:domain odrl:Rule ;
    rdfs:range xsd:boolean .
# WHAT: Granular Data Identification (REQ3)
eol:Locator a rdfs:Class, owl:Class,
    skos:Concept ;
    rdfs:label "Locator"@en ;
    skos:definition "Represents a detailed d
        ata asset locator."@en ;
    skos:note "Describes asset locations."
        @en .
eol:hasLocator a rdf:Property,
    owl:ObjectProperty ;
    rdfs:label "hasLocator"@en ;
    skos:definition "Links an Asset to its
        Locator."@en ;
    rdfs:domain odrl:Asset ;
    rdfs:range eol:Locator .
# Example locator property
eol:locatorSHA256Hash a rdf:Property, owl:Da
    tatypeProperty, skos:Concept ;
    rdfs:label "SHA256 Hash Locator"@en ;
    skos:definition "SHA256 content hash of
        the Asset."@en ;
    rdfs:domain eol:Locator ;
    rdfs:range xsd:string .
# WHY: Rationale (REQ1)
eol:Reason a rdfs:Class, owl:Class,
    skos:Concept ;
    rdfs:label "Reason"@en ;
    skos:definition "Justification for a
        Policy, Rule, or Action."@en ;
    skos:note "Provides context, e.g., for
        informing a DPO."@en .
eol:hasReason a rdf:Property,
    owl:ObjectProperty, skos:Concept ;
    rdfs:label "hasReason"@en ;
    skos:definition "Links a Policy, Rule,
        or Action to its reason."@en ;
    rdfs:domain odrl:Rule ;
    rdfs:range eol:Reason .
# HOW: Deletion Method (REQ2)
eol:destructionMethod a rdf:Property,
    odrl:LeftOperand, skos:Concept ;
    rdfs:isDefinedBy eol: ;
    rdfs:label "Destruction Method"@en ;
    skos:definition "Specifies data deletion
         method or destruction level."@en ;
    skos:note "E.g., 'recycled', 'deleted',
        'wiped', 'physically destroyed', an
        algorithm like Gutmann method."@en .
```

Listing 2: Formal Definition of the EoL ODRL Extension Vocabulary in RDF/Turtle.