Incentivizing Sustainable Data Exchanges through Unique Contextualization of History and Destiny

Wout Slabbinck[1,*], Beatriz Esteves[1], Maarten de Mildt[2], Ruben Dedecker[1], Julián Rojas Meléndez[1], Sofie Verbrugge[2], Didier Colle[2], Pieter Colpaert[1] and Ruben Verborgh[1]

[1]*IDLab, Department of Electronics and Information Systems, Ghent University – imec, Belgium*
[2]*IDLab, Department of Information Technology, Ghent University – imec, Belgium*

Abstract

Exchanging raw data points carries a significant cost and risk for both senders and recipients. Senders cannot reliably indicate what value they aim to achieve with the transaction, and how they want their data to be treated. Recipients cannot easily obtain and produce evidence about the data's accuracy and their lawful right to process it for the intended purpose. While legal frameworks impose boundaries on acceptable behaviors, they prescribe no mechanisms or incentives for compliance, making raw data exchange—despite all of its drawbacks—the current path of least resistance. To encourage sustainable data-driven services, we establish key requirements for encapsulating raw data with a context that captures its purported history and intended destiny. In this paper, we introduce the resulting Trust Envelope model and specification, and showcase its expressiveness by addressing the identified requirements through two real-world use cases. We argue that our approach delivers mutual benefit to sender and recipient by facilitating more frequent granular exchanges, wherein each data transmission is encapsulated within a context unique to the specific processing. This reduces the cost and risk of exchanges, incentivizing them over raw data transactions. Trust Envelopes thereby introduces a local trust context to support the controlled exchange of sensitive data points, which previously would have relied on assumed or more explicit sender–recipient trust relationships. Our work can be extended with constraints across varying legal grounds from distinct jurisdictions, and to accompany purpose-specific derivation of data points.

1. Raw data exchange lacks a sustainable foundation

Businesses delivering services or goods to consumers require a degree of personal data exchange; sometimes for direct and transparent content-related reasons (home address to fulfill a shipment), but often for indirect and obscured goals (home address for unsolicited geotargeted advertising).

As reports tend to focus on how consumers are disadvantaged by the latter category of behaviors, one-dimensional calls for absolute consumer privacy often disregard the careful balance with the former category, in which data is a demonstrable necessity to satisfy a customer's request. Resolutely eradicating all transfer of data is in neither parties' best interest; we should strive to facilitate mutually beneficial exchanges by combining legal, economical, and technical means.

Any socioeconomic system relies on the assumption that at least 50% of parties will attempt to do the right thing; otherwise, its society would be effectively lawless. However, the encoding of ethics and morality into law usually follows a *restrictive* character, not a *prescriptive* one. For instance, most jurisdictions consider the sale of human bodily tissue immoral, and thus enact laws that forbid their exchange for money. This restriction disincentives such acts by introducing instruments for the detection and punishment of those who consider engaging in them. Nonetheless, the associated legal text neither suggests nor positively incentivizes other kinds of trade in which to partake instead, since endorsements exceed a lawmaker's remit.

Similarly, the General Data Protection Regulation (GDPR) [1] and related legal frameworks place restrictions and conditions on the exchange of certain kinds of data. Through those protections, they aim to constrain the longterm boundaries of a level playing field, crucially stopping short of defining the actual field itself, which they entrust to daily economic reality. Invoking society's foundational axiom that a majority will at least *try* to follow the letter and hopefully the spirit of the law, we conclude that a substantial share of the abundant abuse happens not out of malice or criminal intent, but rather unaddressed hindrances in between positive intent and practical execution. Compliance burdens imposed by GDPR and others have not only increased the costs of slightly or strongly unlawful acts, but also those of demonstrating perfectly lawful behavior. Unfortunately, for lack of prescription or real-world precedent to the contrary, the unlawful pathway more often ends up less expensive to businesses after pragmatic consideration.

For the sake of argument, we simplify business categorization into one of three groups:

- the *acting*, who are already adhering to the letter and the spirit of the law;
- the *willing*, who want to adhere, but face obstacles of effort and cost;
- the *unwilling*, who deliberately attempt what law decisively forbids.

With the first two groups jointly representing the expected 50% majority, the last group resisting help, and the first group not needing any, the second group is a persuadable target inclined to respond to positive incentives. Legal enforcement remains the only remedy for the *unwilling*, and a last resort as a negative incentive for the *willing* in case all else fails.

Given considerable existing attention on how technology can improve prevention and detection to discourage unwanted behaviors, we focus in this paper on technologies that reduce the cost and complexity for the group of the *willing* to engage in societally beneficial behaviors. We strive to simplify responsible data processing and subsequent compliance for companies. Toward consumers, by complementing law's existing negative incentives with technology-based positive incentives, we improve their selection of service providers that offer a higher ratio of utility to risk from personal data. For both senders and recipients, we avoid unnecessary technological and legal friction when exchanging data required for economic transactions.

Continuing the home address example: when an individual relocates, it becomes necessary to update their address across a range of services, including government agencies, financial institutions, insurance, and even the microchip registry of their pets. These services require manual updates on a case-by-case basis, incurring costs, inefficiencies, and barriers to innovation for all involved. Data decentralization approaches [2] can facilitate such processes by keeping the source of the data (people) closer to the source of the change (those people moving house). However, those alone prove insufficient as *i)* identity and policy management are limited to simple authentication and access control, *ii)* data minimization and purpose limitation principles are not perpetuated, and *iii)* data integrity and quality are not easily maintainable.

We therefore propose exchanging data via a *Trust Envelope*, a unique contextualized association of the data with instantiated provenance and usage conditions specific to a certain exchange, providing data integrity and usage control. Trust Envelopes support recipient- and purpose-specific data exchanges on the Web through explicit and verifiable trust mechanisms. For example, a Trust Envelope can uniquely encapsulate a specific transmission of an individual's address (or derivation thereof) with subsets of provenance and usage policies that are relevant and necessary for the intended processing. Their issuing mechanism is intentionally lightweight, such that each new usage of the data can be substantiated with a new envelope. Their cryptographic attachment only requires unidirectionality: each envelope describes one specific data unit, yet any data unit can freely associate with any number of envelopes. The effectiveness of a Trust Envelope for consumers and businesses indeed does not hinge on technological guarantees of its data's inseparability: it builds and leverages the incentive for a recipient—as a presumed member of the *acting* or *willing* groups—by reducing their effort and cost of demonstrating compliance when handling such data within predefined constraints.

Trust Envelopes act as facilitators of trusted data exchanges within weaker trust relationships, building on existing literature that defines trust as the *currency of the ecosystem economy* [3], as well as a pillar of *(inter)organizational relationships* [4] and *infrastructure, technology, and control* support

mechanisms [5]. Furthermore, expanding on the notion of *digital trust*, they provide "evidence that an entity has behaved and/or will behave in accordance with the self-stated behaviour" [6] as well as *technology trust*, "the subjective belief by which an organization assesses that the underlying technology infrastructure and support mechanisms are capable of supporting interorganizational communications, transactions, and collaborations" [5]. To this end, Trust Envelopes formalize behavior, the details needed to assemble evidence of behavior, and a means to evaluate trusted interactions from an organizational perspective.

2. Motivating use cases and requirements analysis

Based on the two real-world use cases below, we perform a requirements analysis for contextual data exchange.

2.1. Accessing age-restricted goods or services

Consider a company selling wine, in a jurisdiction where must ensure that buyers are over 18 years old. This obligation seems simple, but fulfilling it digitally becomes surprisingly complex.

The buyer (data sender) is caught in a situation where they must meet the demands of the seller (the recipient), not because they want to, but because it is the legally approved way to complete the purchase. Still, they aim to share as little personal data as possible, avoiding undesired consequences such as their data being reused for loyalty programs, and ensuring that any data exchanged provides benefit for them. The seller requires a cost-effective means of proving compliance of their legal obligations. They do not strictly need the buyer's date of birth, but the confidence the legal threshold is met and provable during later audits.

This situation highlights a digital trust problem with **asymmetric** trust requirements. The sender requires a sufficient degree of trust that the recipient intends to use their data for the intended purpose only. The recipient requires evidence that meets the threshold to fulfill its legal obligations. To support this use case, guarantees about accuracy and usage must be **explicit** and **specific**. Explicit in that the purpose and scope of data use must be clearly stated; specific in that the verifiability of the data must be tailored to the transaction. Anything less would not satisfy the use case's conditions; anything more creates unhelpful complexity and/or liability.

The use case thus needs **high signal, low noise**: minimal, targeted context that de-risks the data exchange for the service. Each such exchange is **unique**: proof that a buyer meets the age threshold is situational–not merely technically valid, but legally and situationally appropriate. What counts as sufficient evidence depends on who is involved, what the purpose is, and under which conditions the data is used. Generic data or broad policies dilute clarity and increase risk. Purpose-specific, contextualized proof supports exchange most efficiently.

2.2. Logistics use case

In logistics, real-time tracking is essential for planning and estimating delivery times. This typically involves monitoring the transporting vehicle or vessel's location. Consider a scenario where Beverage Company (BC) contracts Boxport to transport goods via inland waterways.

Waterway authorities monitor vessel movements and can verify location data, yet are reluctant to expose their sensing infrastructure or raw tracking feeds to third parties [7]. Beverage Company is not interested in the vessel itself, only in as far as it reflects the position of its cargo. During transit, cargo and vessel position are inseparable, so disclosing the former discloses the position of the entire ship, including competitors' goods and sensitive operational details.

The resulting economic conflict of data is alleviated by a technology-assisted legal solution: augment the raw data with a context that omits sensitive provenance details to the extent possible, and restricts the usage of the raw data to cargo tracking. This provides Beverage Company, as a member of the *willing* group, with evidence of allowed usage during an audit. If Beverage Company finds itself tempted

to violate the usage policy by reconstructing the provenance with the location of a competitor's cargo, the context will not stop them; but the company will be unable to misrepresent the context as evidence of correct usage during audits.

2.3. Resulting Requirements

Explicitness from senders Senders must define clear usage control policies that specify how data may be processed, for what purpose, and under which conditions. This does not imply that senders need to have technical knowledge on formally representing policies. Third-party tools can be used by senders to define policies, e.g., using natural language or UI interactive features, which can then be translated into their formal representation in an automated manner.

Specificity from recipients Recipients must be able to articulate exactly what data they require for what purpose. This is related to data minimization and ensures a high signal-to-noise ratio, reducing unnecessary exposure for the sender.

Uniqueness of each exchange Every transaction is context-dependent and shaped by highly specific policies.

Legal compliance Exchanges must adhere to jurisdictional regulations such as the GDPR [1] in Europe or the California Consumer Privacy Act (CCPA) [8].

Accountability Transparency in both data verification and usage is essential. All parties must be able to audit the lifecycle of the data and demonstrate compliance when required.

3. Background

As noted in the introduction, raw data is difficult to work with, and the same holds true for its history and destiny, which are themselves forms of data. Accordingly, the technologies considered must be both machine-interpretable and system-agnostic, supporting interoperability and well-defined semantics. To this end, many of the technologies discussed in this section originate from the Semantic Web, with a particular emphasis on RDF ontologies [9].

3.1. Ontologies for representing the history of data

These approaches fall into two categories: **provenance**, which captures what has happened to the data, and **data assurance**, which establishes authenticity and integrity through cryptographic proofs. *Provenance* refers to data that captures the origin, context, and lifecycle of data. This includes information such as what the data is about, when and by whom it was created, how it has been modified, and where it is stored. Several well-established W3C Recommendations support the modeling of such provenance, including the DCMI Metadata Terms [10], the Provenance Ontology (PROV-O) [11], and the Data Catalog Vocabulary (DCAT) [12].

Data assurance refers to the establishment of authenticity and integrity of data through cryptographic mechanisms. The authoritative W3C Recommendations are are DIDs [13] and VCs [14], which enable interoperable, machine-readable representations of identity and claims. The DID standard allows individuals to manage their own identities based on cryptographic principles. Each identity is represented by a globally unique DID URL, which can be created and resolved without reliance on centralized authorities. While this enables self-sovereign control over identifiers, it does not by itself support the verification of additional claims about the identity. To address this limitation, VCs provide a mechanism for issuing and validating such claims. The VC recommendation defines a model for asserting and verifying claims through three roles: issuer, holder, and verifier. An issuer creates a credential by signing a set of claims with cryptographic proofs. A holder stores these credentials and, to share them with third parties, produces Verifiable Presentations (VPs), which are structured representations that

preserve the integrity and origin of the claims. Verifiers receiving such a presentation can then validate its authenticity using the mechanisms defined in the standard. VCs typically pertain to the holder, identified by a DID. By embedding claims that reference the DID and its associated document, the credential ensures that both the identity and the asserted facts can be independently verified. This conjunction forms the basis for robust data assurance.

3.2. Ontologies for representing the destiny of data

Another part of the story is the destiny of data, which concerns its future use. This requires moving beyond access control enforcement, which restricts access preemptively, and instead consider usage control enforcement, which enables continuous monitoring of data usage [15]. A core requirement for enforcing usage control is a language for specifying usage policies. Within the Semantic Web, a prime candidate for this purpose is the W3C ODRL Recommendation [16] as discussed in [17, 18]. ODRL expressivity goes beyond traditional access control. In particular, it allows the expression of constraints and, in addition to permission rules, other deontic concepts such as prohibitions and obligations. While originally designed to express usage control policies [19], only recent efforts by the ODRL Community Group[1] have initiated progress towards interoperable usage control enforcement of ODRL policies [20]. Moreover, to enhance transparency, sticky policies can be used to attach usage policies to data [21]. However, merely disclosing policies to a receiving party is insufficient; it is equally important that these policies can express legal concepts. To support this, the Data Privacy Vocabulary (DPV) [22], a W3C specification for describing (personal) data processing in support of legislative requirements, can be used in conjunction with ODRL to model legally-aligned policies [23, 24].

4. Uniquely Associating Relevant History and Destiny

4.1. Trust Envelope Model

Considering the previously identified requirements, we propose a model to facilitate unique contextualized exchanges, named **Trust Envelopes**.

Definition 1 (Trust Envelope). A Trust Envelope is a unique data document, timestamped and digitally signed by a sender, that represents a singular act of associating a *data unit* with *context* specific to an actual or proposed *processing* of the data unit by the recipient. The context's *history* metadata consists of a relevant subset of provenance claims pertaining to the data unit; the context's *destiny* metadata consists of instantiated usage policies capturing the sender's agreement or disagreement with relevant processing involving the data unit. A Trust Envelope thereby forms a two-sided evidence mechanism by which the sender can substantiate towards the recipient a degree of the data unit's relevance and correctness, and the recipient can demonstrate towards third parties the fulfillment of preconditions to a certain processing of the data unit.

Data units are neutral statements that can be made by anyone or anything, regardless of factuality. *"Ada Lovelace is regarded as the first computer programmer"* and *"Ada Lovelace was the first woman to win a Nobel prize"* are both statements, but only one of them is grounded in reality. A data unit can be composed of one single data point or statement, or of a small or large set of them, e.g., *"December 10, 1815"* is a data unit, but so is *"Ada was born on December 10, 1815"*. As each exchange of a data unit happens through a new trust envelope, creating a one-to-many relationship; where a data unit can and will be associated with many envelopes, some of which can even have the same sender and recipient. Figure 1 associates a trust envelope with all its components, as well as with all involved entities, e.g., sender or recipient, while quantifying their relationship. As such, Figure 1 reflects the data unit to envelope relationship as a $(1..N)$ relationship, with each envelope pointing to exactly one data unit. A trust envelope consists of three core components: *i)* a *data unit* (or a pointer thereto), *ii) provenance* pertaining to the data unit, *iii)* a *usage policy* pertaining to the data unit, along with policy provenance.
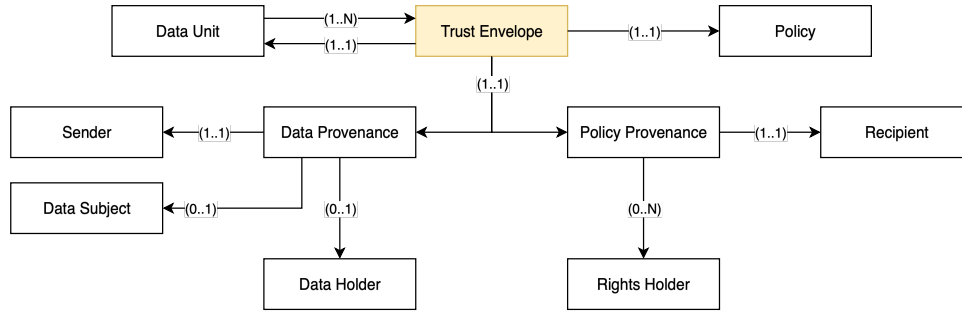

Figure 1: Trust envelope core components and involved entities.

Definition 2 (Sender). Entity that issues the trust envelope.

Definition 3 (Recipient). Entity that receives the trust envelope.

The transmission of an envelope does not necessarily imply a joint transmission of data; the data may be transmitted either beforehand or afterwards to the recipient, by the sender or another party. Additional envelopes might also be transmitted afterwards, to substantiate the data unit with additional evidence and/or to include additional conditions of usage. For example, a date of birth could initially be supported by a trust envelope with *"self-declared"* as provenance and *"to allow website entry"* as policy, which a later envelope could upgrade to respectively *"verified through driver's license scan"* and *"to buy 4 bottles of wine"*.

As with physical envelopes, recipients could choose at their own discretion to discard a trust envelope and only keep its data unit. Nonetheless, storing an envelope is in their best interest, because it serves as a mechanism for evidencing the integrity of the associated data unit and as evidence during auditing procedures to substantiate that the sender permitted certain processing. Trust envelopes thereby reduce the cost of compliance, as they accompany each usage of a data unit with highly specific evidence rather than a blanket consent tickbox. Similar to an HTTP message, each trust envelope describes one specific act of transmission, leaving no functional or practical incentive outside auditing to reuse them as a container for future transmissions.

It should also be noted that the sender might not be the entity to which the data refers, e.g., Alice can create an envelope for Bob's data or an envelope might have an unknown origin. As such, beyond the sender, provenance records should contain a **data subject**, in case the data being exchanged is personal data, or a **data holder**, in case the data is non-personal, if known. These terms are aligned with their legal counterpart definitions from the GDPR [1] and the Data Governance Act (DGA) [25], respectively. On the other hand, the recipient might be the one requesting the data directly, or it might not, i.e., the intended purpose of the sender may be that the data unit can be used by that recipient, but the recipient was not the one who asked for it.

Moreover, beyond data and its provenance, a policy and its provenance are also core components of the envelope. Usage policies can be composed of multiple rules, which contain permissive, prohibitive, or mandatory statements on the usage of a data unit. However, these should be policies specific to the data unit and intended usage, as opposed to the more generic policies that a sender uses to manage classes of data. For instance, a patient (sender) might have a private policy that all medical data can be shared with any certified medical professional for any purpose. A Trust Envelope for a patient's blood sugar level would contain an instantiated policy permitting a specific doctor (the recipient) to use the blood sugar measurement for the purpose of prescribing medication on a specific data. Note that no such specific policy existed in the sender's system; rather, it was derived from a more generic policy in that system. Since usage policies are also data, disclosing them in full may pose a risk to confidentiality or security; therefore, a trust envelope's usage policies should have narrow purposes, recipients, temporal constraints, and so forth. They should be stated as narrowly as possible, because there is always the possibility of issuing a new envelope in case the data unit is to be used for another

purpose. Policy provenance should contain not only the entity receiving the data — the previously defined *recipient* — but also one or more *rights holders*, if known.

Definition 4 (Rights holder). Entity *claiming* to have usage rights over the data unit.

If the data unit contains personal data, data subjects are rights holders, and as such, they hold personal data-related rights. However, the same piece of data can have different kinds of rights resting on it, with different holders. For instance, certain data units might also have intellectual property rights associated with them: Alice holds rights over a photograph in which she appears, while the photographer seeks compensation for capturing said image. Finally, it is the recipient's responsibility to assess whether a rights holder indeed possesses the claimed rights, by looking at the evidence provided by the sender.

4.2. Trust Envelope Specification

Given the model described in the previous subsection, we defined an RDF vocabulary with our proposed terms, which can be used to populate trust envelopes. This vocabulary was developed following the Linked Open Terms (LOT) methodology [26], which is an industry-tested, structured framework designed to guide the lifecycle of ontology engineering. It follows a four-stage development cycle based on Semantic Web development best practices, ensuring the quality, sustainability, and reusability of ontologies. During the initial stage, the requirements and scope of the vocabulary were clearly defined — the identified requirements are described in detail in Section 2, with the derived competency questions being available on the vocabulary's documentation. Building upon that, the conceptual model was translated into a formal representation using RDF and published at https://w3id.org/trustenvelope. Existing vocabularies and standards — such as ODRL [16, 27], DCMI Metadata Terms[2], or DPV [22] — were reused and extended where appropriate to ensure semantic interoperability. The *w3id.org*[3] service is used to provide a permanent identifier to the vocabulary, as well as to support content negotiation. The source code and serialized ontology files are hosted on GitHub, which is also used for issue tracking, under the CC BY-SA 4.0 license. Version control is managed using the Git system. Using the defined vocabulary, as well as the suggested terms from other vocabularies, Figure 2 provides an overview of our suggested ODP for trust envelopes.
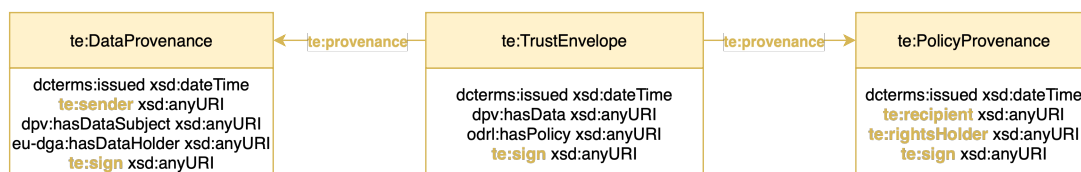

Figure 2: ODP for Trust Envelopes. Trust Envelope vocabulary terms, in yellow, include the `te` prefix.

Beyond recommending the usage of the DCMI Metadata Terms to express temporal information concerning their various components, ODRL is also recommended to both define policies and to link them with trust envelopes. As for the association with the origin subject of the data, we recommend the usage of DPV terms as it is a state-of-the-art resource to describe data and their related rights, e.g., the `dpv:hasDataSubject` property can be used to record the entity from which the data originated, in case of personal data, and the `eu-dga:hasDataHolder`[4] can be used for non-personal data-related use cases. Furthermore, other resources such as DCAT [12] and the VC [14] model are also considered viable options to represent data units. To the best of our knowledge, there are no vocabularies that provide terms and definitions to represent data and policy provenance, as well as for senders, recipients, and rights holders — as such, we provide formalizations for these terms in the Trust Envelope vocabulary.

---

[2]Published at http://purl.org/dc/terms/, with prefix `dcterms`.
[3]https://w3id.org, accessed on 31/July/2025.
[4]EU-DGA is the DPV extension for the Data Governance Act, available at https://w3id.org/dpv/legal/eu/dga.

In the spirit of evolvability promoted by LOT, this specification is intended to be extensible: new entities, e.g., actors or agents who enforce policies, or components, e.g., data type, can be added to the model and the specification as required.

5. Materialization of Trust Envelopes

This section applies the RDF ontology of Trust Envelopes, introduced in the previous section, to the use cases outlined in Section 2. The first example provides a detailed construction of a Trust Envelope, whereas the second illustrates the versatility of Trust Envelopes by demonstrating that the data and provenance of one can be used as the data unit for another.

To illustrate Trust Envelope instantiation, we revisit the accessing age-restricted goods use case, described in Subsection 2.1. Alice, identifiable as `https://alice.org`, wants to have alcoholic beverages delivered to her house in Belgium, where the legal age to buy alcoholic beverages is 18. She goes online to Alcoholic Beverages Corporations (ABC), at `https://alcoholic-beverages.com`, and attempts to buy beer. The store requires sufficient evidence to demonstrate that each purchase complies with Belgian law. Therefore, ABC restricts purchases of alcoholic beverages to those who provide sufficient evidence, which is being over the age of 18. As such, she sends the Trust Envelope shown in Listing 1 to ABC. The data unit contains a proof (represented by a VP), issued by the government, stating that Alice is over 18 years old. The envelope also encloses an ODRL policy, `https://example.org/policy`, permitting ABC to use the data unit solely for age verification purposes. Both the VP and policy are available at https://w3id.org/trustenvelope#age-restricted-goods.

```
1  ex:envelope1 a te:TrustEnvelope ;
2      dcterms:issued "2024-02-12T11:20:10.999Z"^^xsd:dateTime ;
3      dpv:hasData <https://country.org/uuid1> ; odrl:hasPolicy ex:policy1 ;
4      te:provenance ex:dataProvenance1, ex:policyProvenance1 ;
5      te:sign ex:signedEnvelope1 .
6  ex:dataProvenance1 a te:DataProvenance ;
7      dcterms:issued "2024-02-12T11:20:10.999Z"^^xsd:dateTime ;
8      te:sender <https://alice.org> ; dpv:hasDataSubject <https://alice.org> ;
9      te:sign ex:signedDataProvenance1 .
10 ex:policyProvenance1 a te:PolicyProvenance ;
11     dcterms:issued "2024-02-12T11:20:10.999Z"^^xsd:dateTime ;
12     te:recipient <https://alcoholic-beverages.com> ;
13     te:rightsHolder <https://alice.org> ; te:sign ex:signedPolicyProvenance1 .
```

Listing 1: Trust Envelope issued by Alice to prove to ABC that her age is above 18.

Trust Envelopes support composable, traceable data exchange: each envelope can build on the data and provenance of a previous one, transforming it into a new, context-specific unit governed by its own policies. To demonstrate this capability, we refer to the logistics use case described in Subsection 2.2. In this scenario, the waterway authority measures AIS[5] signals, including those transmitted by Bluewave (Boxport's designated vessel for this transport), and processes them. To monitor its fleet, Boxport requests access to this data. The authority responds by issuing a Trust Envelope (Listing 2) that encapsulates Bluewave's precise location at time t and explicitly grants Boxport permission to use and distribute the data.

Subsequently, when BC requests the location of its cargo from Boxport, the latter responds with a new Trust Envelope (Listing 3). This envelope materializes the previously received Trust Envelope from the waterway authority into derived cargo spatio-temporal data, using it as the basis for a new, context-specific data unit. Through this composition, BC retrieves *i)* the location data at time t for its cargo, *ii)* provenance and signatures to assess the accuracy and trustworthiness of the data, and *iii)* a deliberate policy permitting the use of the spatiotemporal data unit to estimate the arrival of its cargo.

Both envelopes' materializations, including data modelling, policies, and signatures, can be found at https://w3id.org/trustenvelope#cargo-monitoring.

---

[5]Automatic Identification System (AIS): https://en.wikipedia.org/wiki/Automatic_identification_system

```
1  ex:envelope2 a te:TrustEnvelope ;
2      dcterms:issued "2024-02-12T11:20:10.999Z"^^xsd:dateTime ;
3      dpv:hasData ex:AIS-measurement ;  odrl:hasPolicy ex:policy2 ;
4      te:provenance ex:dataProvenance2, ex:policyProvenance2 ;
5      te:sign ex:signedEnvelope2 .
6  ex:dataProvenance2 a te:DataProvenance ;
7      dcterms:issued "2024-02-12T11:20:10.999Z"^^xsd:dateTime ;
8      te:sender <https://waterway.org> ; eu-dga:hasDataHolder <https://boxport.com> ;
9      te:sign ex:signedDataProvenance2 .
10 ex:policyProvenance2 a te:PolicyProvenance ;
11     dcterms:issued "2024-02-12T11:20:10.999Z"^^xsd:dateTime ;
12     te:rightsHolder <https://crew.boxport.com>, <https://waterway.org> ;
13     te:recipient <https://boxport.com> ; te:sign ex:signedPolicyProvenance2 .
```

Listing 2: Trust Envelope issued by the waterway authority to allow Boxport the tracking of the real-time location of a given vessel.

```
1  ex:envelope3 a te:TrustEnvelope ;
2      dcterms:issued "2024-02-12T11:20:10.999Z"^^xsd:dateTime ;
3      dpv:hasData ex:cargo-location ; odrl:hasPolicy ex:policy3 ;
4      te:provenance ex:dataProvenance3, ex:policyProvenance3 ;
5      te:sign ex:signedEnvelope3 .
6  ex:dataProvenance3 a te:DataProvenance ;
7      dcterms:issued "2024-02-12T11:20:10.999Z"^^xsd:dateTime ;
8      te:sender <https://boxport.com> ; eu-dga:hasDataHolder <https://boxport.com>  ;
9      te:sign ex:signedDataProvenance3 .
10 ex:policyProvenance3 a te:PolicyProvenance ;
11     dcterms:issued "2024-02-12T11:20:10.999Z"^^xsd:dateTime ;
12     te:rightsHolder <https://crew.boxport.com>, <https://boxport.org> ;
13     te:recipient <https://beverage-company.com> ; te:sign ex:signedPolicyProvenance3 .
```

Listing 3: Trust Envelope created by the logistics provider that contains transformed data from the trust envelope from the logistics provider with the true location of a given vessel at a given time, allowing the company to use that location for predicting the Estimated Time of Arrival of the cargo the its destination.

6. Discussion

6.1. Meeting the Requirements

Declaring usage control policies is a necessary first step towards fulfilling the **sender explicitness** requirement. Unlike access control, enforcing usage constraints technologically becomes infeasible once data is exchanged [28]. From then on, enforcement relies on monitoring and auditing, which can be legally upheld under the assumption that both the sender and recipient have signed the policy. This touches the realm of **legal compliance**, which is further elaborated in Section 6.4. Promoting and re-using well-known ontologies for both the data and the provenance enables **recipient specificity** through allowing precise expression of what the data represents. **Accountability** is embedded in the Trust Envelope model. The history, comprising the data and its provenance, and the destiny, comprising the policy and its provenance, are each individually signed. These two signed entities are then encapsulated within an overarching signature at the Trust Envelope level. As such, a recipient can audit both the authenticity of the data and the identity of the sender. Furthermore, when the policy is co-signed by both parties, the sender can audit whether the data has been used in accordance with the agreed-upon terms. If misuse is detected, appropriate measures can be initiated, potentially invoking legal recourse.

6.2. Limitations

While Trust Envelopes capture history, destiny, and regulation around data, they may inadvertently leak information about an entity, as RDF's interconnected structure allows identifiers and relationships

to be traced across graphs. Incorporating Privacy-Enhancing Technologies (PETs) could mitigate these risks. One technique that could be included in Trust Envelopes is pseudonimity through incorporating work in [29] on DIDs and Solid. Another is including data minimization through the work in [30] on Zero-Knowledge Proofs for VCs. Although signing is vital for Trust Envelopes, its implementation is only briefly addressed in the examples, as signing RDF data is inherently ambiguous due to the open-world assumption: graphs can be extended at any time through dereferencing and imports, making it unclear what constitutes a complete and signable dataset. A pragmatic solution to this challenge is proposed by [31].

6.3. Trust Envelopes in Ecosystems

Having discussed the ontology technically, situating Trust Envelopes in a broader governance and business context of data ecosystems is appropriate. While not the focus of this research, this context must be acknowledged for the ontology to be integrated in practice. In data ecosystems, data is shared between multiple parties via data sharing services [32]. These services operate on data at a granular level (such as transformation and visualization). By adhering to the requirements stated in Section 2.3, the Trust Envelope model introduces a technique to add trust to such a service. Although the Trust Envelope model aims to ensure trustworthy data sharing, services may adopt it only superficially, e.g., for pragmatic reasons, claiming to be compliant while ignoring key requirements like sender explicitness. This partial use undermines the model's integrity, forcing other ecosystem participants to rely on unverified trust. A related challenge in data ecosystem governance is the lack of interdisciplinary alignment. Frameworks often address trust from a single dimension (solely legal, technical, or business), yet when interdisciplinary applied, new issues emerge. For example, a governance model developed from a business perspective may expose unforeseen legal complications during implementation [33, 34, 35, 36]. Trust in data ecosystems depends not only on technical mechanisms like Trust Envelopes, but also on the willingness of all involved entities to apply them properly. Governance structures, whether through independent boards or automated consensus mechanisms, shape how trust is operationalized and evaluated. For instance, in the case of age-restricted goods, a government might offer a Trust Envelope service for citizens to share proofs, implemented via a GovTech (private organizations intertwined with public sector components) [37]. Both parties need to be trusted: the government and the GovTech entity, the latter already being the subject of critique on whether they can be trusted entities [37]. Even with technically sound Trust Envelopes, trust breaks down if any party fails to engage transparently. To conclude, enforcing trust in ecosystems in a purely technical way remains unfeasible; as such, Trust Envelopes must accommodate current gaps of business and governance data experts.

6.4. Legal Considerations

The overall model described in Section 4.1 was developed to be regulation-agnostic — this design choice was considered the most appropriate to have a model that is sufficiently abstract to tackle legal requirements *i)* from different jurisdictions, and *ii)* from different data types, e.g., personal or non-personal data. Nonetheless, the terms chosen to model data provenance, in particular related to the source entity of the data, were kept in line with European legislation as these regulations are being followed and similarly adapted in other jurisdictions' regulations [38]. Moreover, by incorporating provenance and usage policies as fundamental components of trustworthy data exchanges, Trust Envelopes can be used as auditing tools by external legal authorities to assess good and poor practices of recipients, and possibly held as proof in judgments for the later case. Nevertheless, in future iterations of the model, legal experts will be consulted to analyze the extent to which the current model satisfies legal requirements for the exchange or personal and non-personal data, not just by looking at data protection law, but also extending to other branches of the law, e.g., IP rights. This is of particular importance to understand and correctly model all the parties that might have rights and obligations over the data unit, how they interplay, and how they are to be correctly interpreted and enforced. For example, returning to the logistics scenario modelled in Listings 2 and 3, while Boxport has recognized

rights over the location data of its vessels, its crew members also have recognized rights under the GDPR, as the location of the vessel also discloses their location.

7. Conclusion

Sheer data availability is no longer sufficient for data exchange systems that yearn to ensure their users trust their services, while simultaneously addressing interoperability, economic, and legal requirements. Furthermore, this paper has highlighted shortcomings of existing technological approaches to solve simple use cases, especially in light of increasing concerns around provenance, usage conditions, and regulatory compliance. Through the Trust Envelope model and specification, we propose an interoperable solution that captures the essential dimensions of data history, intended use, and legal constraints, with the aim of enabling contextualized data exchanges that support essential principles, such as data minimization, purpose limitation, or data integrity. The proposed specification, which promotes the usage of standards such as VCs, DIDs, ODRL, and DPV, demonstrates that Trust Envelopes offer a comprehensive framework for managing data in a way that is both expressive and enforceable, while the introduced use cases validate the model's versatility in handling diverse requirements.

By making trust explicit and verifiable, Trust Envelopes pave the way for more responsible, automated, and legally sound data exchanges. We aim to expand this work by applying the model to real-world scenarios across varied legal jurisdictions, further validating its utility and refining its implementation in practice. Moreover, the development of a rigorous strategy for digital signatures is of utmost importance to ensure Trust Envelopes' integrity and verifiability, including compliance with the newly-updated European Digital Identity framework [39]. Finally, we will integrate the Trust Envelope model into our authorization framework, which delegates usage control to a User-Managed Access (UMA) server [40, 41] and builds on an ODRL-conformant engine to perform policy interpretation and evaluation [20].

Acknowledgments

Declaration on Generative AI

During the preparation of this work, the author(s) used Grammarly in order to: Grammar and spelling check. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] 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), 2016. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj.

[2] R. Verborgh, Re-decentralizing the Web, for good this time, in: O. Seneviratne, J. Hendler (Eds.), Linking the World's Information: Essays on Tim Berners-Lee's Invention of the World Wide Web, ACM, 2023, pp. 215–230. doi:10.1145/3591366.3591385.

[3] T. W. Andreassen, L. Lervik-Olsen, H. Snyder, A. C. Van Riel, J. C. Sweeney, Y. Van Vaerenbergh, Business model innovation and value-creation: the triadic way, Journal of service management 29 (2018) 883–906.

[4] T. C. Earle, Trust in risk management: A model-based review of empirical research, Risk Analysis: An International Journal 30 (2010) 541–574.

[5] P. A. Pavlou, P. Ratnasingam, Technology trust in b2b electronic commerce: Conceptual foundations, in: K. Kangas (Ed.), Business Strategies for Information Technology Management, IGI Global Scientific Publishing, 2003, pp. 200–215. doi:10.4018/978-1-93177-745-2.ch014.

[6] R. N. Akram, R. K. Ko, Digital trust - trusted computing and beyond: A position paper, in: 2014 IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications, 2014, pp. 884–892. doi:10.1109/TrustCom.2014.116.

[7] I. Akaichi, W. Slabbinck, J. A. Rojas, C. V. Gheluwe, G. Bozzi, P. Colpaert, R. Verborgh, S. Kirrane, Interoperable and Continuous Usage Control Enforcement in Dataspaces, in: J. Theissen-Lipp, P. Colpaert, S. K. Sowe, E. Curry, S. Decker (Eds.), Proceedings of the Second International Workshop on Semantics in Dataspaces (SDS 2024), volume 3705 of *CEUR Workshop Proceedings*, CEUR, Hersonissos, Greece, 2024. URL: https://ceur-ws.org/Vol-3705/paper10, iSSN: 1613-0073.

[8] California Consumer Privacy Act, 2018. URL: https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB375.

[9] G. Klyne, J. J. Carrol, B. McBride, RDF 1.1 Concepts and Abstract Syntax, 2014. URL: https://www.w3.org/TR/rdf11-concepts/.

[10] DCMI Metadata Terms, 2020. URL: https://www.dublincore.org/specifications/dublin-core/dcmi-terms/.

[11] K. Belhajjame, J. Cheney, D. Corsar, D. Garijo, S. Soiland-Reyes, S. Zednik, J. Zhao, PROV-O: The PROV Ontology, 2013. URL: https://www.w3.org/TR/prov-o/.

[12] R. Albertoni, D. Browning, S. Cox, A. G. Beltran, A. Perego, P. Winstanley, Data Catalog Vocabulary (DCAT) – Version 3 – W3C Recommendation, 2024. URL: https://www.w3.org/TR/vocab-dcat-3/.

[13] Decentralized Identifiers (DIDs) v1.0, 2022. URL: https://www.w3.org/TR/did-1.0/.

[14] M. Sporny, D. Longley, D. Chadwick, I. Herman, Verifiable Credentials Data Model v2.0 – W3C Recommendation (2025). URL: https://www.w3.org/TR/vc-data-model-2.0/.

[15] J. Park, R. Sandhu, The UCONABC usage control model, ACM Transactions on Information and System Security 7 (2004) 128–174. URL: https://doi.org/10.1145/984334.984339. doi:10.1145/984334.984339.

[16] R. Iannella, S. Villata, ODRL Information Model 2.2 – W3C Recommendation, 2018. URL: https://www.w3.org/TR/odrl-model/.

[17] I. Akaichi, S. Kirrane, Usage Control Specification, Enforcement, and Robustness: A Survey, 2022. URL: http://arxiv.org/abs/2203.04800. doi:10.48550/arXiv.2203.04800, arXiv:2203.04800 [cs].

[18] B. Esteves, V. Rodríguez-Doncel, Analysis of ontologies and policy languages to represent information flows in GDPR, Semantic Web 15 (2024) 709–743. URL: https://journals.sagepub.com/doi/full/10.3233/SW-223009. doi:10.3233/SW-223009.

[19] V. Rodríguez-Doncel, N. Roman, Towards Conformance in ODRL 3.0, in: M. Sabou, A. Harth, P. Lisena, E. Curry, B. Zhang, R. Alharbi, Y. He, G. Rehm, S. Schimmler, S. Dietze, N. Manola, A. Cimmino, N. Fornara, V. Rodríguez-Doncel, J. Domingue, A. Rettinger, D. Trilling, M. Grobelnik, C. d'Amato, V. Fionda, I. Tiddi, G. Tolomei (Eds.), Joint Proceedings of the ESWC 2025 Workshops and Tutorials, volume 3977 of *CEUR Workshop Proceedings*, CEUR, Portorož, Slovenia, 2025. URL: https://ceur-ws.org/Vol-3977/OPAL2025-8.pdf, iSSN: 1613-0073.

[20] W. Slabbinck, J. Rojas Meléndez, B. Esteves, P. Colpaert, R. Verborgh, Interoperable Interpretation and Evaluation of ODRL Policies, in: E. Curry, M. Acosta, M. Poveda-Villalón, M. van Erp, A. Ojo, K. Hose, C. Shimizu, P. Lisena (Eds.), The Semantic Web, Springer Nature Switzerland, Cham, 2025, pp. 192–209. doi:10.1007/978-3-031-94578-6_11.

[21] D. Miorandi, A. Rizzardi, S. Sicari, A. Coen-Porisini, Sticky Policies: A Survey, IEEE Transactions on Knowledge and Data Engineering 32 (2020) 2481–2499. URL: https://ieeexplore.ieee.org/document/8807248/. doi:10.1109/TKDE.2019.2936353.

[22] H. J. Pandit, B. Esteves, G. P. Krog, P. Ryan, D. Golpayegani, J. Flake, Data Privacy Vocabulary (DPV) – Version 2.0, in: G. Demartini, K. Hose, M. Acosta, M. Palmonari, G. Cheng, H. Skaf-Molli, N. Ferranti, D. Hernández, A. Hogan (Eds.), The Semantic Web – ISWC 2024, Springer Nature Switzerland, Cham, 2024, pp. 171–193. doi:10.1007/978-3-031-77847-6_10.

[23] B. Esteves, H. J. Pandit, V. Rodríguez-Doncel, ODRL Profile for Expressing Consent through

Granular Access Control Policies in Solid, in: 2021 IEEE European Symposium on Security and Privacy Workshops (EuroS PW), 2021, pp. 298–306. doi:10.1109/EuroSPW54576.2021.00038.

[24] H. J. Pandit, B. Esteves, Enhancing Data Use Ontology (DUO) for Health-Data Sharing by Extending it with ODRL and DPV, Semantic Web Journal (2024). doi:10.3233/SW-243583.

[25] Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act), 2022. URL: http://data.europa.eu/eli/reg/2022/868/oj/eng.

[26] M. Poveda-Villalón, A. Fernández-Izquierdo, M. Fernández-López, R. García-Castro, LOT: An industrial oriented ontology engineering framework, Engineering Applications of Artificial Intelligence 111 (2022). doi:10.1016/j.engappai.2022.104755.

[27] R. Iannella, M. Steidl, S. Myles, V. Rodríguez-Doncel, ODRL Vocabulary & Expression 2.2 – W3C Recommendation, 2018. URL: https://www.w3.org/TR/odrl-vocab/.

[28] A. A. Nyre, Usage Control Enforcement - A Survey, in: D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, A. M. Tjoa, G. Quirchmayr, I. You, L. Xu (Eds.), Availability, Reliability and Security for Business, Enterprise and Health Information Systems, volume 6908, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 38–49. URL: http://link.springer.com/10.1007/978-3-642-23300-5_4. doi:10.1007/978-3-642-23300-5_4, series Title: Lecture Notes in Computer Science.

[29] G. De Mulder, B. De Meester, Pseudonymity for Personal Data Stores: Pseudonymous WebIDs and Decentralized Identifiers, in: 2nd International Workshop on Emerging Digital Identities (EDId), ARES 2025, 2025.

[30] C. H.-J. Braun, T. Käfer, RDF-Based Semantics for Selective Disclosure and Zero-Knowledge Proofs on Verifiable Credentials, in: E. Curry, M. Acosta, M. Poveda-Villalón, M. van Erp, A. Ojo, K. Hose, C. Shimizu, P. Lisena (Eds.), The Semantic Web, Springer Nature Switzerland, Cham, 2025, pp. 383–402. doi:10.1007/978-3-031-94575-5_21.

[31] R. Dedecker, J. De Roo, B. Esteves, P. Colpaert, Demonstrating a pragmatic solution to context associations in RDF using Blank Node Graphs, in: The Semantic Web: ESWC 2025 Satellite Events, 2025.

[32] M. de Mildt, S. Verbrugge, D. Colle, A market analysis on data ecosystem initiators and their value propositions in different ecosystems, Telecommunications Policy 49 (2025) 102910. URL: https://www.sciencedirect.com/science/article/pii/S0308596125000072. doi:https://doi.org/10.1016/j.telpol.2025.102910.

[33] M. I. S. Oliveira, G. d. F. Barros Lima, B. Farias Lóscio, Investigations into data ecosystems: a systematic mapping study, Knowledge and information systems 61 (2019) 589–630.

[34] R. D'Hauwers, N. Walravens, P. Ballon, Data ecosystem business models: value and control in data ecosystems, Journal of Business Models (JOBM) 10 (2022) 1–30.

[35] D. Lis, B. Otto, Data governance in data ecosystems–insights from organizations (2020).

[36] D. Heinz, C. Benz, M. Fassnacht, G. Satzger, Past, present and future of data ecosystems research: A systematic literature review (2022).

[37] N. Bharosa, The rise of govtech: Trojan horse or blessing in disguise? a research agenda, Government Information Quarterly 39 (2022) 101692. URL: https://www.sciencedirect.com/science/article/pii/S0740624X22000259. doi:https://doi.org/10.1016/j.giq.2022.101692.

[38] A. Bradford, The Brussels Effect: How the European Union Rules the World, Oxford University Press, 2019. doi:10.1093/oso/9780190088583.003.0002.

[39] Regulation (EU) 2024/1183 of the European Parliament and of the Council of 11 April 2024 amending Regulation (EU) No 910/2014 as regards establishing the European Digital Identity Framework, 2024. URL: http://data.europa.eu/eli/reg/2024/1183/oj/eng.

[40] W. Slabbinck, R. Dedecker, W. Termont, B. Esteves, P. Colpaert, R. Verborgh, From Access Control to Usage Control with User-Managed Access, Solid Symposium 2025, Leiden, The Netherlands, 2025. Forthcoming.

[41] W. Termont, Authorization for Data Spaces, Technical Report, KNoWS (IDLab, Ghent University –

imec), 2025.