Computer Law & Security Review 62 (2026) 106372

Contents lists available at ScienceDirect

Computer Law & Security Review: The International Journal of Technology Law and Practice

journal homepage: www.elsevier.com/locate/clsr

# Traceability for privacy: A foundational pillar for an effective and competitive EU digital rulebook

Maitrayee Pathak

*Research Group Public Law and Computer Science, Zentrum für Angewandte Rechtswissenschaft (ZAR), Karlsruhe Institute of Technology (KIT), Vincenz-Prießnitz-Straße 3, 76131, Karlsruhe, Germany*

ARTICLE INFO

*Keywords:*
GDPR
AI Act
Data governance
EuroStack
Digital identity
OTrace
W3C verifiable credentials

ABSTRACT

The European Union has established a comprehensive digital rulebook (GDPR, DSA, DMA, DGA, Data Act, AI Act), to develop a digital single market that respects fundamental rights. An investigation into this legal framework, however, reveals significant barriers to its effectiveness, including inconsistent enforcement, "gold-plating," and a burdensome compliance landscape for SMEs. This paper's analysis establishes that a standardized, interoperable, and privacy-preserving traceability layer is a critical infrastructural prerequisite to bridge the gap between legal principles and technical enforceability. Traceability, defined as the ability to document the data lifecycle and link processing to a verifiable legal basis, provides the necessary foundation for accountability, a gap in the institutional enforcement system. This paper conducts a legal-technical analysis, first establishing through a systematic review that traceability is a foundational requirement across the EU digital rulebook. Second, it compares technical architectures based on blockchains, attestations (OTrace) and cryptographic claims (W3C Verifiable Credentials). The analysis concludes that while public blockchains present GDPR compliance paradoxes and require workarounds, a careful merger of OTrace and VCs, and an operational framework synthesizing the EU Digital Identity Wallet and Data Governance Act data intermediaries offers a pragmatic, legally-aligned solution to the current institutional deficit. Finally, it connects this "EU-Trace" model to the EuroStack's strategic vision, concluding that a risk-based mandate for traceability can mitigate the negative effects of regulatory fragmentation by harmonizing the mechanism of compliance, thereby enhancing competitiveness for a transparent, accountable digital Europe.

## 1. Introduction: the EU's regulatory effectiveness gap and the case for traceability

### 1.1. The EU's ambitious digital rulebook and its implementation deficit

A comprehensive "digital rulebook" encompassing the General Data Protection Regulation (GDPR), Digital Services Act (DSA), Digital Markets Act (DMA), Data Governance Act (DGA), Data Act, Artificial Intelligence Act (AI Act) was adopted by the EU to develop a single market that is both competitive and protective of fundamental rights [1]. This suite of legislation forms a distinct European vision for digital sovereignty [2].

However, this legal framework faces a significant implementation deficit. High-level analyses, such as the 2024 Draghi Report, have identified critical barriers to the framework's effectiveness, which this investigation concludes are symptoms of an underlying institutional and operational deficit. These barriers include inconsistent enforcement across Member States, the practice of "gold-plating" regulations [3] with additional national requirements, and an asymmetric and burdensome compliance landscape that disproportionately affects Small and Medium-sized Enterprises (SMEs) [4]. These challenges reveal a deeper, structural disconnect between abstract legal principles and the technical reality of today's data ecosystems. The coexistence of multiple legal regimes with overlapping supervisory mandates risks fragmenting enforcement, resulting in inconsistent application of the rules, and due to “privatization of enforcement” even risking the rule of law [5].

The core principles of accountability, purpose limitation, transparency, and user control, which are the foundations of the EU digital rulebook [6], are difficult to audit in part because they were fundamentally designed to regulate bilateral relationships between a single data controller and a data subject, rather than the complex realities that often necessitate non-standard, reflexive regulatory approaches [7]. Modern data processing and artificial intelligence systems, however, are not usually deployed by a single, monolithic actor. Rather, they are

*E-mail address:* maitrayee.pathak@kit.edu.

https://doi.org/10.1016/j.clsr.2026.106372

Available online 1 July 2026

produced, deployed, and operated across complex, multi-layered algorithmic and data supply chains by "many hands" [8]. These supply chains facilitate data flowing iteratively between multiple parties, fundamentally muddying traditional accountability analyses which often assume a single regulatory target with the power to control a system's inputs, structure, or outputs. This fragmentation creates an "accountability horizon," where individual actors within the supply chain have severely limited visibility into the operations of their upstream providers or downstream deployers. Furthermore, recent scholarship has acknowledged the complicated legal agreements typical of the digital economy and turned to contract specification languages and smart contracts as potential solutions [9].

Recent European enforcement actions also demonstrate that complex supply chains are inherently difficult to audit. In the ad-tech ecosystem, the Belgian Data Protection Authority's action against IAB Europe revealed that once a user's privacy preference is broadcast into the Real-Time Bidding network, it disappears into closed-source downstream databases, making it practically unfeasible to empirically verify if brokers actually delete data upon a withdrawn consent signal [10]. Similarly, in its fines against Google [11] and Shein [12], the CNIL bypassed this opacity; rather than untangling granular technical faults across layered architectures, the regulator applied the single economic unit doctrine and cross-border joint controllership to hold parent companies directly. Finally, this accountability horizon is especially severe in generative AI ecosystems. As demonstrated by the Italian Garante's action against OpenAI for training ChatGPT on unlawfully processed personal data, upstream model providers often operate with severe opacity [13]. This upstream non-compliance creates a compliance paradox for downstream deployers: under EDPB guidance [14], deployers are legally accountable for ensuring lawful data provenance, yet they remain technically powerless to audit the developer's proprietary, black-box training logs. While the courts and national Data Protection Authorities (DPAs) have increasingly labelled all actors participating in these supply chains joint controllers regardless of their original contractual identification as processors, they are merely clarifying the responsibility, but not solving the auditability issues.

The gap between legal intent and technical verifiability and operative enforceability contributes to inconsistent enforcement that undermines regulatory effectiveness, stifles innovation by creating legal uncertainty, and ultimately harms the EU's competitiveness [4]. The subsequent systematic analysis of this implementation deficit establishes that bridging the gap requires a foundational infrastructural layer largely underexplored in policy discussions: data traceability. This investigation posits that while such a technical layer cannot eliminate the political sources of fragmentation [15], it can substantially mitigate their negative economic effects by harmonizing *the mechanism of compliance*, thereby creating a predictable and unified operational environment for businesses, even where the underlying legal environment remains fragmented.

### 1.2. Defining traceability as a legal-technical construct

In the context of data privacy and regulation, this paper uses the following definition:

Traceability is a legal-technical construct defined as the ability to create a verifiable record of the complete lifecycle of personal data (monitoring all access, modifications, processing, and storage) and linking these activities back to a verifiable legal basis, such as consent.

It is a concept that extends beyond the purely technical notions of *data provenance* and *data lineage* which provide a detailed historical record of a data item's origins and transformations [16]. In contrast, traceability in the legal context is fundamentally about demonstrating accountability [17]. It provides the verifiable evidence (in other words: the *digital paper trail*) necessary to prove that data processing aligns with legal obligations to "demonstrate compliance" according to 5(2) of the GDPR. While traceability provides verifiable evidence, it is important to note that the core technical mechanism of traceability is purpose-agnostic. Its effectiveness in protecting privacy or enhancing competitiveness is determined entirely by the legal and governance regime in which it is implemented.

### 1.3. Thesis and paper roadmap

This paper's investigation posits that the integration of a standardized, interoperable, and privacy-preserving traceability layer into Europe's digital infrastructure is an essential prerequisite for enhancing regulatory effectiveness, developing a truly competitive single market for data, and successfully realizing the strategic vision of the EuroStack initiative, a proposal for a common European digital infrastructure designed to enhance digital sovereignty [18]. To substantiate this thesis, the paper offers three key contributions. First, it conducts a systematic, cross-regulatory legal analysis establishing the imperative for traceability as a necessary remedy to overcome current institutional, procedural failures. Second, it performs a detailed comparative study of the main candidates for technical architectures for traceability, cross-examining them against the legal imperatives identified in the first step, and the EU governance challenges identified by the Draghi Report. Third, it proposes a novel and viable operational framework for traceability ("EU-Trace"), which synthesizes two proven technical architectures, and the EU's unique legal and public infrastructure. Finally, it presents a strategic policy vision for integrating this architecture as a foundational governance technology within the EuroStack, advocating for a risk-based mandate to ensure proportional and effective accountability.

To develop this argument, the paper is structured as follows. Section 2 will establish the legal imperative for traceability across the EU digital rulebook. Section 3 will provide a critical comparative analysis of technical architectures for traceability, and make a case for the proposed hybrid model as the basis for a new operational framework. Section 4 will connect this analysis to the EU's strategic goals, merging the visions of the EuroStack and Data Commons and analyzing the governance of the proposed traceability layer as a piece of Digital Public Infrastructure, including how a risk-based mandate would function. Finally, Section 5 will conclude by summarizing the paper's contributions and offering policy recommendations.

## 2. The legal imperative for traceability across the EU digital rulebook

### 2.1. Introduction: from accountability to demonstrable traceability in the EU digital single market

The European Union's ambitious legislative project to construct a digital single market has produced an overlapping web of regulations [1]. While each instrument (from the foundational GDPR to the technologically-specific AI Act) addresses different policy objectives, the following systematic analysis reveals the emergence of a common, unifying imperative: the requirement for end-to-end traceability of personal data.

For the purposes of this analysis, traceability is defined as the capacity to create and maintain a verifiable and auditable record of the entire lifecycle of personal data. This encompasses its provenance (origin), the specific legal basis for each processing activity, all transformations and transfers, its use in automated decision-making and algorithmic systems, and its eventual deletion.

While it is unlikely that legislators in 2016 envisioned the precise technical requirements of the 2024 AI Act, this paper will argue that the cumulative effect of this legislative journey has been the creation of a de facto mandate for end-to-end data traceability.The analysis now proceeds to analyze each stage of this legislative evolution, demonstrating how the abstract principle of accountability has been translated into a concrete, technical-operational imperative for personal data

traceability.

### 2.2. The foundational mandate for traceability in GDPR

An examination of the GDPR establishes it as the cornerstone of the EU's digital rulebook and the legal bedrock upon which all subsequent traceability requirements have been built. Although the term traceability is not explicitly mentioned within its provisions, an investigation of the GDPR reveals the establishment of a comprehensive framework that makes traceability a functional necessity for the demonstration of compliance. This is achieved through the operationalization of four interconnected requirements: the overarching principle of accountability, the specific documentation mandate for Records of Processing Activities (RoPA), the architectural requirement of data protection by design and by default, and the operational demands of the fulfillment of data subject rights. The imperative for traceability within the GDPR is therefore mainly driven by the evidentiary burden on the data controller to demonstrate compliance to supervisory authorities.

#### 2.2.1. The accountability principle as the legal bedrock

The formalization of the accountability principle was a main requirement introduced by the GDPR. Article 5(2) states that "The controller shall be responsible for, and be able to demonstrate compliance with, paragraph 1 ('accountability')" [6]. This operative phrase requires the controller to create and maintain sufficient evidence to prove its adherence to the data processing principles outlined in Article 5(1), such as lawfulness, fairness, transparency, purpose limitation, and data minimisation.

This requirement for "demonstrable compliance" necessitates a complete audit trail of data processing activities. Legal-technical scholarship has affirmed that accountability is an ongoing, proactive process, not a one-time, "check-the-box" exercise [19]. According to Article 24 of the GDPR, controllers are expected to implement comprehensive privacy management programs, including policies, procedures, and technical controls, and to document these measures thoroughly. This collection of documented evidence is what allows a controller to demonstrate compliance.

Technically, this means that for any given piece of personal data, the controller must be able to trace its processing history and link each stage to a corresponding compliance justification. The ability to show that data was collected for a specified, explicit, and legitimate purpose (purpose limitation), that only the necessary data was collected (data minimisation), that it is accurate, and that it is not stored for longer than necessary (storage limitation) has to be evidenced. The practical necessity of this approach is evidenced by Article 88 on Logging requirements from Regulation (EU) 2018/1725 that prescribes that EU institutions should have accurate logging (traceability) for “the collection, alteration, access, consultation, disclosure, including transfers, combination and erasure of operational personal data” [20].

This evidentiary burden has been strictly enforced, with DPAs treating specific technical failures as breaches of the accountability principle itself. For instance, the Polish DPA fined the Szczecin-Centrum District Court for violations of Articles 25 (Data Protection by Design) and 32 (Security of processing), linking these failures to an infringement of Article 5(2) [21]. Similarly, the Spanish DPA imposed a fine on EDP ENERGIA, S.A.U., citing a violation of Article 25 because the company had not designed its contracting process with appropriate operative measures to trace and verify authorization when acting on behalf of another person [22]. The regulatory bodies who can enforce this demonstrable compliance has been further broadened by the Court of Justice of the European Union in Meta Platforms (C-252/21), which concluded that even national competition authorities can assess GDPR compliance [23].

#### 2.2.2. Records of processing activities as the primary instrument of traceability

If Article 5(2) provides the legal basis for traceability, Article 30 provides the primary instrument for its operationalization. This article mandates a specific, detailed instrument of traceability: the Record of Processing Activities (RoPA) [6]. The RoPA can be understood as the GDPR's most explicit tool for the creation of a comprehensive map of an organisation's data processing activities. Under Article 30(1), the maintenance of a record is required from each controller, containing, at a minimum: the name and contact details of the controller, joint controller, representative, and Data Protection Officer, the specific purposes of the processing (e.g., payroll management, customer marketing), a description of the categories of data subjects (e.g., employees, clients) and the categories of personal data (e.g., contact details, health records), the categories of recipients to whom the personal data have been or will be disclosed, including those in third countries, details of transfers to third countries and the documentation of suitable safeguards, envisaged time limits for the erasure of different data categories, and a general description of the technical and organisational security measures implemented [6].

This detailed list effectively requires the controller to trace and document the entire planned lifecycle of each category of personal data it processes. The RoPA serves as the central repository of this traceability information, acting, in the words of the Irish Data Protection Commission, as a "a living, dynamic, document, which is continuously updated to reflect the current position of the organisation at all times with regard to their processing of personal data" [24]. The explicit obligation in Article 30(4) for the controller or processor to "make the record available to the supervisory authority on request" makes it a key piece of evidence for the demonstration of compliance [6]. Failure to maintain an adequate RoPA constitutes a direct breach, as was demonstrated by the CNIL’s fine against the data broker TAGADAMEDIA [25]. The sanction was imposed partly because its RoPA was deficient, failing to specify which of two companies was acting as the data controller and thus failing to clearly trace responsibility.

#### 2.2.3. Data protection by design and by default and the architectural imperative

Article 25 of the GDPR raises the concept of traceability to a proactive, architectural principle. It mandates that data controllers, "both at the time of the determination of the means for processing and at the time of the processing itself, shall implement appropriate technical and organisational measures… which are designed to implement data-protection principles… in an effective manner" [6]. This is the principle of Data Protection by Design. Furthermore, the implementation of measures is required to ensure that "by default, only personal data which are necessary for each specific purpose of the processing are processed" [6]. This is Data Protection by Default.

Together, these principles require organisations to build traceability into their systems and processes. The EDPB Guidelines 4/2019 state that effective implementation requires integrating safeguards "throughout the processing lifecycle," connecting the legal principle to the technical necessity of data lifecycle tracking [26]. Examples of the implementation of data protection by design that directly support traceability include: legal basis and data minimisation through systems that collect only the essential data fields required for a specific purpose, thereby creating a traceable link between each data point and its lawful purpose; lifecycle security from collection to secure destruction, with logs and audit trails to trace access and modifications; autonomy for giving users granular controls within the legal basis, requiring the system to trace and apply these preferences to all relevant processing activities [26].

Article 25 thus makes traceability an engineering imperative as well.

#### 2.2.4. The functional necessity of traceability for data subject rights

The final pillar establishing the traceability mandate in the GDPR is the practical impossibility of the fulfillment of Data Subject Rights (Art.

15–22) without it. These rights place specific operational demands on data controllers.

Regarding the Right of Access under Article 15 a recent CJEU jurisprudence has interpreted its scope so expansively as to make comprehensive, contextual traceability an operational necessity. In Case C-487/21 (CRIF), the Court ruled that the transparency requirement of Article 12(1) of the GDPR includes receiving even "extracts from documents or even entire documents or extracts from databases" if essential for the data subject to understand the context of the processing [27]. This requires traceability back to source documents. Further, in Case C-579/21 (J.M. v. Pankki S), the CJEU held that system logs (detailing who accessed personal data and when) are themselves the personal data of the individual whose information was accessed [28]. Finally, in Case C-634/21 (SCHUFA Holding), the Court affirmed the right to "meaningful information about the logic involved" in automated decisions, necessitating the ability to trace a specific algorithmic output back to its data inputs [29]. Fulfilling an access request requires organisations to take 'reasonable steps' to locate information across their systems; without robust processes using traceability, handling complex or large-scale requests can cause a 'disproportionate or unjustified level of disruption' and become a significant operational burden [30].

The Right to Erasure under Article 17 requires the controller to trace and delete all copies of the data, including notifying downstream recipients to whom the data has been disclosed. The Right to Data Portability under Article 20 requires the ability to trace, combine, and export a specific user's data in a machine-readable format [6]. These rights presume in practice that the controller has a complete and accurate understanding of where all personal data resides and how it moves through its systems—in other words, traceability.

### 2.3. Extending traceability to platform ecosystems: the DMA and DSA

While the GDPR established the principles of accountability (and traceability), the DMA and DSA adapted these principles to large online platforms and digital ecosystems. This analysis reveals a shift in the focus of traceability from an internal, accountability-driven exercise for the demonstration of compliance to regulators, towards a more external, transparency-oriented function to empower users and ensure fairness in the market.

#### 2.3.1. DMA: traceability of granular consent as a precondition for business models

The DMA targets the market power of large online platforms designated as gatekeepers. A key provision for data traceability is Article 5(2) which prohibits a gatekeeper from combining personal data from one of its core platform services with personal data from other services, cross-using personal data from one core platform service in another separate service, or signing in end users to other services in order to combine personal data [31]. The crucial exception to this prohibition is when "the end user has been presented with the specific choice and has given consent within the meaning of Article 4, point (11), and Article 7 of Regulation (EU) 2016/679″ [31]. This reference to the GDPR's definition of consent means that the consent must be freely given, specific, informed, and unambiguous [6]. This necessitates a granular consent management system capable of tracing the specific choice presented, the user's action (cross-using, combining or signing in the user), the scope of consent, and the consent lifecycle, including any instances of withdrawal and ensuring that requests for consent are not repeated for the same purpose within a one-year period if refused.

Legal scholarship has highlighted the significant challenges of obtaining truly "free" consent from users in the context of a power imbalance with a gatekeeper, making the quality and traceability of the consent record even more critical for legal defence [32]. The DMA, therefore, makes granular traceability of consent a precondition for engaging in data processing activities fundamental to many large platform business models. The practical enforcement was demonstrated in the European Commission's non-compliance investigation into Meta's "pay or consent" model. The Commission's found that the binary choice presented to users failed to provide a genuine alternative for those who do not consent, thereby not meeting the high standard for freely given consent required by the DMA, resulting in the first-ever fine under the DMA in April 2025 [33].

#### 2.3.2. The DSA: traceability for transparency and safety

The DSA aims to create a safer and more transparent online environment by regulating the responsibilities of intermediary services [34]. Its traceability mandate builds upon GDPR principles, focused on the output of algorithmic systems used for advertising [35].

Article 26 of the DSA mandates that providers of online platforms presenting advertising must ensure that for each specific ad shown to each individual user, the user can identify, in a clear and real-time manner: that the information is an advertisement; the natural or legal person on whose behalf the advertisement is presented; the natural or legal person who paid for the advertisement; and "Meaningful information directly and easily accessible from the advertisement about the main parameters used to determine the recipient to whom the advertisement is presented" [34].

This last requirement creates a mandate for *output traceability*. This necessitates an auditable trail linking the ad (the individual output) to the specific personal data points (e.g., age, location, interests) and the profiling logic used for targeting. The EDPB's Guidelines 3/2025 on the interplay between the DSA and the GDPR confirm this, clarifying that DSA transparency is provided in real-time and is ad-specific, a more dynamic and user-facing form of traceability than that required by the GDPR [35]. Industry bodies such as IAB Europe have released detailed implementation guidelines to help the ad-tech supply chain comply with the DSA's ad transparency obligations [36].

Furthermore, for Very Large Online Platforms, Article 39 requires the creation of a publicly accessible repository of advertisements, containing details such as the content of the ad, the advertiser, the period of dissemination, and the main parameters used for targeting [34]. This creates a long-term, searchable record of advertising activities, embedding traceability as a tool for both public and regulatory scrutiny. The DSA also bans targeted advertising based on profiling using special categories of personal data or data from minors, requiring proactive, dynamic traceability as opposed to consent traceability previously required under the GDPR [34].

### 2.4. Traceability for the data economy: the DGA and data act

Building on the foundations of the GDPR and the platform-specific rules of the DMA and DSA, the DGA and the Data Act aim to develop a single market for data. In this new, multi-party ecosystem, traceability becomes the legal and technical mechanism for establishing neutrality, transparency and trust.

#### 2.4.1. DGA: traceability in governed data sharing

The DGA aims to facilitate data sharing by creating new, trustworthy structures for data exchange through Common European Data Spaces [37]. These multi-party ecosystems require significant trust, which is technically and legally enforced through neutrality, transparency and traceability.

The DGA introduces *data intermediation services* as neutral third parties connecting data holders and users. While the European Commission's practical guidance does not explicitly use the term 'audit trail,' it clarifies that intermediaries must be able to prove compliance with Article 12′s strict neutrality obligations [38]. This functionally necessitates a complete and verifiable record of all data flows to demonstrate that data has not been used for any other purpose.

This principle is now being implemented in sectoral data spaces. The European Health Data Space Regulation grants patients granular transparency: “Electronic health data access services should provide detailed

information on access to data, such as when and which entity or natural person accessed data and which data were accessed." [39] This is a user-facing traceability mandate that is impossible to fulfill without a complete audit log of access events. The technical foundation for the data spaces is being developed by the Data Spaces Support Centre, whose Blueprint defines among its technical "building blocks" traceability and provenance tracking for evidence "on the origin of the data and the treatment of the data during its processing in the value chain" and "what was done with the data." [40]

Where personal data is involved, the DGA defers to the GDPR in Article 1(3) [37]. This means intermediaries must trace the data flow as well as demonstrate a valid legal basis under the GDPR.

The DGA also provides a framework for data altruism, where individuals or companies voluntarily make their data available for purposes of general interest, such as scientific research using a *European data altruism consent form* to ensure that consent can be given and withdrawn easily and in a standardized manner across Member States [37]. Recognised data altruism organisations must maintain transparent records of how the data is used, requiring a traceable link from the data subject's consent to the specific research or public interest project that uses their data. In the DGA's framework, traceability is therefore a mechanism of transparency for data holders to participate in these new data-sharing ecosystems.

#### 2.4.2. The data act: technical traceability for the internet of things

The Data Act creates a new right for users (both consumers and businesses) to access the data generated by their use of a connected product (often referred to as the Internet of Things (IoT)) or related service and to share that data with third parties of their choice [41].

Article 3 of the Data Act mandates that products must be designed and manufactured, and related services provided, in such a manner that data generated by their use is, by default, easily, securely, and directly accessible to the user. The Data Act requires manufacturers to build systems capable of technically tracing, isolating, and extracting the specific data stream generated by a single user's device and to provide it in a structured, commonly used, and machine-readable format [41].

The interplay with the GDPR is again explicitly addressed. Article 1 (5) of the Data Act clarifies that it is without prejudice to EU data protection law, and where the user is not the data subject whose personal data is requested, that data may only be made available if there is a valid legal basis under the GDPR, such as the data subject's consent. The data holder must be able to, technically, trace the data to a specific product and user, and trace the valid legal basis, such as the consent of the data subject(s) involved.

The Data Act thus embeds traceability at the hardware and software design level, making it a prerequisite for placing connected products on the EU market. Together, the DGA and Data Act establish that in a multi-actor data ecosystem, traceability enables transparency, facilitates data sharing, and ensures that the rights of individuals remain protected.

### 2.5. Ultimate traceability: data provenance and logging in the AI act

The EU AI Act, the world's first comprehensive law on artificial intelligence introduces one of the most granular and technically prescriptive traceability requirements in the entire EU digital rulebook. The AI Act completes the evolutionary trajectory of traceability by moving beyond tracing the data's lifecycle (GDPR), the user's granular consent (DMA), or its commercial flow (DGA/Data Act). In line with the output traceability concept of the DSA advertisement transparency requirements, it mandates the traceability of the automated system's behaviour, targeting two critical and often opaque aspects of AI: the data used to train it and the operational decisions it makes. This is achieved through stringent requirements for data provenance and automatic event logging for high-risk AI systems [42].

#### 2.5.1. Data governance and provenance for AI training data

Article 10 of the AI Act requires that the training, validation, and testing datasets used to develop high-risk AI systems are subject to "appropriate data governance and management practices" [42]. The governance must track "data collection processes and the origin of data, and in the case of personal data, the original purpose of the data collection" and "data-preparation processing operations, such as annotation, labelling, cleaning, updating, enrichment and aggregation" [42]. Providers of high-risk AI systems must be able to trace their training data back to its source, demonstrate its lawfulness and suitability for the intended purpose. This includes design choices regarding the datasets, data preparation operations, examination for possible biases, and identification of data gaps [42].

This mandate directly addresses the "black box" problem at the training stage. Given that AI models are often trained on vast datasets scraped from the internet or aggregated from multiple sources, establishing provenance is a significant technical and legal challenge [43]. However, it is essential for mitigating algorithmic bias, ensuring fairness, and verifying compliance with other laws, such as copyright and the GDPR. For general-purpose AI (GPAI) models, providers are also required to draw up and make publicly available a summary of the content used for training, further reinforcing the transparency of data provenance [42]. The AI Act thus makes the traceability of data inputs a prerequisite for developing and deploying high-risk AI in the EU. This requirement has been operationalized through the mandatory GPAI Template, published by the European Commission. This template legally obligates all providers to publicly disclose a "sufficiently detailed summary" of the content used for training, including their data sources and collection processes [44].

#### 2.5.2. Automatic event logging for high-risk systems

Beyond the inputs, the AI Act also mandates *output traceability*. Article 12 requires that high-risk AI systems "shall technically allow for the automatic recording of events (logs) over the lifetime of the system" [42]. In practice these logs must be automatic, immutable, and sufficiently detailed to ensure a level of traceability appropriate to the system's intended purpose [45]. The goal is the creation of a verifiable record of the system's functioning to facilitate post-market monitoring, investigate incidents, and monitor operations [42]. The precise technical specifications for these logging capabilities will be defined in harmonized standards developed by the designated European standards bodies, specifically CEN—CENELEC Joint Technical Committee 21 (JTC 21) [46].

For certain high-risk systems, such as those used for remote biometric identification, Article 12(3) specifies the minimum information that must be logged: the period of each use of the system, the reference database against which input data was checked, the specific input data that resulted in a match, and the identification of the natural persons involved in the verification of the results, linking the automated decision to human oversight [42].

Article 19 reinforces this obligation by requiring providers to keep these automatically generated logs for an appropriate period, which must be at least six months, unless other laws require a longer retention period [42].

This combination of mandatory, automatic logging creates the most robust and prescriptive form of traceability found in EU digital law. It establishes an immutable audit trail of the AI system's decisions, linking specific inputs to specific outputs and human interventions. This operational traceability is the ultimate mechanism for accountability. When an AI system causes harm or makes a questionable decision, these logs provide the necessary evidence to investigate what happened, why it happened, and who is responsible. The AI Act, through its dual requirements of data provenance and operational logging, ensures that the entire lifecycle of a high-risk AI system, from the data it learned from to the decisions it makes, is fully traceable and accountable. This represents the culmination of the legal principle that began with the GDPR: to

ensure trust in complex digital systems, their inner workings must be demonstrable, auditable and traceable by design.

### 2.6. The evolving legal imperative of traceability

The following table synthesizes the evolution of the core traceability requirements and the nature of traceability through the EU digital regulations.

As summarized in Table 1, this legislative evolution demonstrates a trajectory from static, documentary accountability under the GDPR to dynamic, systemic and increasingly real-time operational traceability. Demonstrating compliance across multi-party algorithmic supply chains is a further technical architecture challenge, necessitating a wider compliance lens than the traditional controller-based accountability approach of the GDPR. In the European digital single market, accountability is no longer an abstract principle, but it is a technical reality, and its name is traceability.

## 3. A comparative analysis of technical architectures for traceability in the EU

Privacy compliance scholarship has revealed several technical approaches to translating the GDPR into machine-readable formats. For example, semantic web technologies and legal reasoning tools like PrOnto (Privacy Ontology for Legal Reasoning) [47] support privacy-by-default and legal compliance. PrOnto provides a legal knowledge modeling of privacy agents, data types, processing operations, rights, and obligations and incorporates references to GDPR concepts such as sensitive data, anonymization techniques, and processing purposes. Similarly, Natural Language Processing (NLP) has successfully automated certain compliance checks of legal documents. For instance, approaches like DIKAIO [48] utilized several machine learning techniques to evaluate Data Processing Agreements against the GDPR, effectively predicting compliance breaches based on the presence or absence of required legal information types.

**Table 1**
The evolving legal imperative for personal data traceability in EU digital regulations.

| Regulation | Key Articles | Core Traceability Requirement | Nature of Traceability | Primary Context/ Addressee |
|---|---|---|---|---|
| GDPR | Art. 5 (2), 24, 30, 25, 15–22 | Demonstrable accountability. Records of Processing Activities (ROPA). Data Protection by Design and by Default. Exercise of Data Subject Rights. | Documentary, Accountability-Based | All data controllers and processors. |
| DMA | Art. 5 (2) | Traceable, GDPR-standard consent for combining or cross-using personal data. | Consent-Based | Designated 'gatekeeper' platforms. |
| DSA | Art. 26 | Transparency for advertising parameters | Output-Based | Online platforms and marketplaces. |
| DGA | Art. 12 | Neutral and transparent data flows managed by data intermediaries. | Flow-Based | Data sharing and intermediation services. |
| Data Act | Art. 4, 5 | User access to generated data; Sharing data with third parties upon user request. | Technical, Access-Based | Manufacturers of connected products (IoT). |
| AI Act | Art. 10, 12, 19 | Data provenance for training data; Automatic event logging for system operations. | Provenance & Operational | Providers/ deployers of high-risk AI systems. |

However, while semantic ontologies and NLP compliance checking are essential for ex-ante compliance (mathematically defining the taxonomy of the law and automating the legal audit of static documents like Data Processing Agreements) they do not inherently address the problem of ex-post audit through evidence. Ontologies define what the system should do, and NLP verifies the intent of the data controller as expressed in their policy agreements, but neither provides the operational ledger necessary to prove what the system actually did in real-time. Traceability serves as the necessary operational counterpart to semantic modeling and NLP compliance checking to provide the audit trail for the physical execution of the policy agreements across the algorithmic supply chain.

Moving from conceptual necessity to practical implementation, the legal imperative for traceability requires an investigation into a technical architecture capable of operating at scale within the European regulatory and infrastructural landscape. Apart from providing the means of traceability for the traceability requirements analyzed in Chapter 2, any proposed architecture must also provide a practical response to the two key challenges of inconsistent enforcement and legislative gold-plating highlighted by the Draghi Report [4].

### 3.1. How traceability architectures address inconsistent enforcement

Inconsistent enforcement of the EU digital rulebook stems from fragmented supervision, divergent legal interpretations, and a deficit of resources and technical expertise within national supervisory authorities (SAs) [49]. A proposed technical architecture needs to be able to at a minimum mitigate this by establishing a standardized, evidence-based foundation for oversight.

A mechanism of efficiency of such an architecture would be to create a "single source of truth" for regulators. Currently, an SA's investigation relies on a company's internal documentation, such as RoPAs, logs, and policies, which vary dramatically in quality and format [50]. The reflexive approach followed by some DPAs with the aim of cooperatively addressing compliance deficiencies further necessitates non-standard enforcement if built on these varied internal documents [7]. This variance can make enforcement subjective and inconsistent. A proposed traceability architecture should address this variance by creating a standardized, machine-readable, and cryptographically verifiable audit trail of data processing events. A regulator in Ireland and a regulator in Spain investigating the same company would therefore examine the exact same type of evidence, generated according to the same open standards. The harmonization of the evidence base can be a mechanism to lead to more consistent enforcement processes.

This proposed architecture should also lower the barrier for technical audits. SAs are often hindered by a lack of technical expertise [50]. The technical architecture should act as a force multiplier for these under-resourced bodies. Instead of requiring deep forensic expertise to analyze complex, proprietary system logs, a regulator could query the user's designated traceability system for a clear record of data lifecycle events. Opaque technical processes should thus be translated into a legible, standardized format, making it easier for SAs to audit compliance and verify a company's claims.

Finally, a standardized evidentiary framework should enable effective cross-border enforcement, where inconsistency is most acute [50]. Ideally the architecture should be linked to pan-European public infrastructure(such as the emerging European Digital Identity (EUDI) Wallet [51], provided it achieves successful deployment) to create an interoperable system by design, rather than relying on member-state-specific

solutions. This could significantly improve cooperation between SAs under the GDPR's one-stop-shop mechanism. The Lead Supervisory Authority could share and interpret verifiable evidence with other concerned SAs, as all would be operating with the same open standards, in other words the same “language of proof".

### 3.2. How traceability architectures mitigate the effects of gold-plating

"Gold-plating" is the practice of Member States adding country-specific requirements to EU regulations, fragmenting the Single Market and creating burdensome compliance landscapes for businesses [3]. While a technical architecture cannot prevent a Member State from enacting such laws, it should provide a mechanism to mitigate their detrimental consequences.

The architecture should harmonize the *mechanism of compliance*, creating a predictable and unified operational environment even where underlying legal rules remain fragmented. A company operating in Germany and France might face different national data protection nuances [49]. With a traceability architecture, the technical process for demonstrating compliance would be nearly identical in both countries. This has the potential to reduce the operational complexity, especially for cross-border operations by creating a single compliance architecture that sits underneath the fragmented legal layer.

Traceability architectures may also create a market for specialized service providers, such as smart contract providers or traceability agents. These providers would have to develop expertise in navigating the specific gold-plated requirements of multiple Member States and offer a unified "compliance-as-a-service" solution. A company could then rely on its service provider to handle the nuances of national laws, effectively outsourcing and centralizing the management of this fragmentation.

This approach can shift the focus of an audit from non-standard legal interpretation to concrete, verifiable events. Gold-plating often creates legal uncertainty and debate over the interpretation of national rules [3]. A traceability system changes the crucial question from "What is the precise interpretation of this national law?" to "Is there a verifiable, cryptographically signed record of consent for this specific data processing event?". This focus on a factual, auditable data trail has the potential to cut through the ambiguity created by gold-plating and can provide a clearer path to demonstrating compliance.

### 3.3. Comparative study of data privacy traceability architectures

#### 3.3.1. Concepts

In the context of this analysis, a traceability architecture is defined as the structural and operational framework of technologies, protocols, and data models designed to systematically capture, store, and verify key personal data events (ranging from consent events like creation, modification, revocation to downstream processing operations). It functions as the socio-technical mechanism that translates abstract legal compliance requirements (such as demonstrable accountability) into an auditable, evidentiary record even across distributed algorithmic supply chains.

To overcome the 'accountability horizon' inherent in multi-party algorithmic supply chains, traditional centralized logging (where a single entity controls a siloed database) is structurally insufficient [8]. Therefore, this analysis focuses on architectures capable of cross-boundary verification, specifically evaluating the three primary paradigms currently proposed in the literature:

- Blockchain-based Systems: These are decentralized, distributed ledgers utilizing cryptographic consensus mechanisms. In the context of regulatory compliance, it is necessary to explicitly distinguish between public, permissionless chains (which face insurmountable controller-identification issues under the GDPR) and private, permissioned consortium chains (which are favored in enterprise tracking but sacrifice true decentralization) [52].
- Attestation-based Protocols (e.g., OTrace): These are decoupled architectures where users delegate the continuous, high-volume auditing of backend data events to an independent, third-party "Traceability Agent." This approach effectively separates the high-integrity act of granting consent from the high-volume task of logging operations [53].
- W3C Verifiable Credentials (VCs): This is a standardized data model for cryptographic claims made by an issuer about a subject. VCs are fundamentally user-centric, typically stored in an edge device (such as a digital wallet), and specifically designed for the selective, privacy-preserving disclosure of attributes, rather than the continuous logging of backend system events [54].

There is significant scholarship on the privacy compliance of blockchain architectures. Systematic literature reviews on GDPR-compliant blockchains [52] categorize the fundamental conflicts between distributed ledgers and the GDPR into six major thematic groups: Data Deletion and Modification (Articles 16, 17, 18), Protection by Design and Default (Article 25), Responsibilities of Controllers and Processors (Articles 24, 26, 28), Consent Management (Article 7), Data Processing Principles (Articles 5, 6, 12), and Territorial Scope (Article 3). To mitigate these paradoxes, the literature details how architects must propose extensive, complex workarounds. For instance, achieving compliance often requires dividing data into personal and non-personal categories, storing all personal data off-chain in local databases, and only storing cryptographic proofs of existence (hashes) on the immutable ledger. Furthermore, to comply with the right to be forgotten, systems must rely on logical deletion—destroying the encryption keys linked to the off-chain data rather than altering the ledger itself [52].

#### 3.3.2. Comparative analysis

This comparative analysis evaluates the three primary technical architectures for data privacy traceability against a dual-lens framework. Each architecture is assessed on its capacity to fulfill the legal compliance imperatives of the EU digital rulebook identified in Table 1 of Chapter 2, and its potential to address inconsistent enforcement and legislative gold-plating. To illustrate the practical implications of the compliance characteristics of each technology, this analysis introduces a hypothetical "Smart Diagnostic AI App" supply chain as a representative testing model. Consider a European hospital (the deployer) utilizing a diagnostic application developed by a German health-tech SME (the provider). This SME relies on a foundation model accessed via an API from a US-based technology conglomerate (the model provider), whose training data was originally aggregated by a third-party global data broker. The following analysis specifically examines how each architecture traces data flows across this complex, multi-party ecosystem, while also considering their alignment with the EU's public digital infrastructure defined in the EuroStack [18].

*GDPR: Demonstrable Accountability (Art. 5(2)):* Public blockchains demonstrate a poor capacity for native accountability due to obscure controller identities [55], requiring complex workarounds via permissioned chains to ensure a contractually designated controller is identifiable [56]. In practice, if the European hospital attempts to log a patient's consent on a public network, it cannot prove to a regulator which nodes control the data, forcing the hospital and the German SME to abandon it for a private chain. Attestation-based protocols offer a highly aligned candidate for creating a demonstrable record of processing, though they require strong institutional enforcement to ensure their "covert-accountability" mechanisms are respected [53]. Applied to our hypothetical, the hospital, the German SME, and the US model provider all route cryptographically signed attestations of their respective processing steps to a registered Traceability Agent, creating a unified, end-to-end operational record of accountability. Verifiable Credentials (VCs) mitigate this challenge of establishing legal identity by

providing cryptographically verifiable proof of a legal basis where the issuer is an accountable entity natively recognized by the system [54]. When the hospital issues a VC to the patient (a cryptographic claim documenting the exact terms of the patient's consent) its verified eIDAS identity [57] and the legal basis are mathematically baked directly into the credential.

*GDPR: Records of Processing Activities (ROPA) (Art. 30):* Public blockchains are unsuitable for dynamic, sensitive ROPA data due to fundamental transparency and immutability constraints, making them highly inefficient [55]. Compliance requires off-chain storage for the ROPA data, recording only immutable hashes on the ledger [56]. Consequently, the German SME cannot store its constantly updating processing records on-chain without exposing trade secrets or violating privacy. Conversely, an attestation agent's log serves as a strong operational counterpart, functioning as a dynamic, real-time, user-centric ROPA for the entire algorithmic supply chain [53]. This allows the Traceability Agent to seamlessly map the data's journey from the hospital to the SME and finally to the US provider. VCs moderately support this requirement; while they represent discrete authorizations, they do not constitute a complete, continuous record of all processing activities [54]. The patient can hand the hospital a VC authorizing treatment, but that credential will not track the subsequent backend processing by the US model provider.

*GDPR: Data Protection by Design and by Default (Art. 25):* The data replication inherent in public chains fundamentally conflicts with data minimization [55,43], necessitating off-chain storage and Zero-Knowledge Proofs to verify claims without revealing sensitive data [58]. If the hospital's app logs sensitive diagnostic data on a blockchain, developers must build complex cryptographic workarounds to hide the actual health metrics from the public network. Attestation protocols provide a strong structural defense, using "covert-accountability" as a design principle to deter data misuse across the supply chain [53]. This discourages the SME or US provider from exploiting the data, as their actions are securely logged by the Traceability Agent. VCs provide a robust mechanism in this domain, natively supporting data minimization via "selective disclosure" mechanisms [59]. The patient can use selective disclosure to share only the specific health attribute required by the hospital without revealing their entire medical history.

*GDPR: Exercise of Data Subject Rights (Art. 15–22):* The mathematical immutability of blockchains directly conflicts with the Right to Erasure [55], requiring logical deletion workarounds applied to off-chain data [55]. If a patient requests the hospital to delete their diagnostic data, the on-chain hash remains in the blockchain. This is insufficient for compliance, as the EDPB classifies cryptographic hashes as pseudonymized, rather than anonymized, personal data [56]. Attestation protocols structurally support data subject rights, designed to log the exercise of rights requests via "Rights Attestations" [53]. The agent seamlessly logs the transmission of the patient's erasure request from the hospital down to the SME and to the US model provider. VCs offer a robust user-centric model, empowering data subjects with control, access, and deletion of their credentials immediately within their digital wallet [57], but they do not record the cascading controller and processor notifications required. While the patient can revoke or delete the consent credential from their digital wallet, the wallet does not automatically notify the hospital, nor does it track whether the SME and the model provider actually deleted the patient's personal data.

*DMA: Traceable, GDPR-Standard Consent for Combining or Cross-use (Art. 5(2)):* Blockchains can provide a strong immutable log of granular consent, provided they employ off-chain and permissioned workarounds to establish controller accountability [55]. If the US-based model provider acts as a designated gatekeeper combining the hospital's diagnostic data stream with their other services (e.g. a Google Health service), it must rely on these private ledger setups to prove compliance. Attestation protocols offer an aligned approach to create a granular, high-volume log of consent and its subsequent utilization [53]. The Traceability Agent easily logs both the patient's initial consent and the gatekeeper's specific use of the data across its services. VCs are equally aligned, providing a cryptographically verifiable and highly portable record of granular consent required for compliance [59]. The patient generates a verifiable consent credential that the US provider can ingest and audit to prove DMA adherence for combining personal data from the hospital with their other services.

*DSA: Transparency for Advertising Parameters (Art. 26):* While blockchains establish a transparent audit trail, their inherent scalability and cost limitations create severe barriers for high-volume, real-time advertising ecosystems, theoretically requiring modular or Layer-2 solutions [60]. If the app developer uses targeted advertising in the diagnostic app, logging real-time ad parameters on-chain is architecturally prohibitive. Attestation protocols are strongly suited for this mandate, capable of efficiently logging the targeting parameters for each individual ad through processing attestations [53]. The Traceability Agent can act as a high-speed receiver, recording the specific demographic targets for every single ad impression served. VCs offer only a moderate fit; they can securely support user-managed ad preferences but do not provide the output traceability necessary to track specific ad impressions [59]. The patient's wallet securely stores their general ad preferences, but cannot log the real-time outputs generated by the platform.

*DGA: Neutral and Transparent Data Flows (Art. 12):* Public blockchains are poorly suited for governed data sharing due to obscure identity mapping [55], but permissioned chains offer a strong framework for secure, neutral, and auditable ledgers among trusted participants [56]. A DGA-regulated data intermediary facilitating flows between the hospital and the German SME must build a private consortium chain to maintain secure, identifiable governance. Attestation protocols are highly compatible here, as the Traceability Agent directly mirrors the technical and legal role of a multi-party data intermediary [38]. The agent centrally logs the data exchanges between the hospital and the SME, functioning exactly as the DGA envisions. VCs effectively manage permissions and access rights within a data space but lack the capacity to log the continuous data flows themselves [59]. VCs authorize the SME to initially enter the data space, but cannot track the subsequent machine-to-machine data exchanges.

*Data Act: User Access and Third-Party Data Sharing (Art. 4, 5):* Blockchains can generate tamper-proof logs of IoT data access requests, but processing high-volume IoT telemetry on-chain is architecturally impractical [60], demanding an off-chain data and on-chain metadata configuration [58]. If a patient uses a connected biometric IoT device feeding continuous data into the hospital's app, putting that telemetry directly on-chain causes unsustainable data volume. Attestation protocols are a strong structural fit, capable of logging attestations for high-volume IoT data access requests and transfers [53]. The agent effortlessly logs the high-volume stream of data transfers from the biometric device as they occur. VCs perform poorly in this operational context; they can provision device authorization but cannot log the continuous data streams [59]. The patient uses a VC to initially pair their device with the hospital's app, but the VC cannot record the ongoing flow of heart rate or diagnostic metrics.

*AI Act: Data Provenance for Training Data (Art. 10):* The immutable chain of custody provided by blockchains is a moderate technical fit for tracking data provenance [43], but relies entirely on privacy workarounds (storing personal training sets off-chain and cryptographic hashes on-chain) to function legally [56]. In our practical example, when regulators audit the global data broker to verify the AI model's training data, the broker must rely on these hash-based workarounds to hide the massive personal datasets. Attestation protocols are moderately useful, optimizing for discrete processing events rather than the detailed, multi-stage transformations characteristic of AI training data [43]. The agent easily logs the patient's hospital visit, but struggles to map the complex data aggregation steps performed by the broker. VCs similarly offer promising utility; they can certify a dataset's high-level origin and potentially granular, record-level data provenance in a

processing chain [59]. The data broker could use a VC to certify the dataset's legal acquisition, and potentially the granular provenance required for a deep algorithmic audit as well.

*AI Act: Automatic Event Logging for System Operations (Art. 12 & 19):* Blockchains conceptually fit the requirement for automatic, immutable event logging, but structural scalability and cost severely limit their application for high-volume operational logging, requiring complex high-throughput blockchain setups [60]. As the US-based foundation model makes continuous, automated diagnostic decisions, the sheer volume of backend operations makes public blockchain logging too slow and expensive. Attestation protocols are highly compatible; their core architectural function of receiving and logging event attestations strongly supports the demands of automated backend AI logging [53]. For the diagnostic app the Traceability Agent is purpose-built to act as a high-speed receiver for these continuous diagnostic logs. VCs are unsuited for this mandate, lacking the design necessary for high-volume, automated back-end tracking [59]. A user's digital wallet is fundamentally not designed to track the high-volume, automated outputs generated by the US provider's server-side AI model.

*Addressing Inconsistent Enforcement:* Public blockchains fail to address enforcement inconsistencies due to inherent legal non-compliance, forcing reliance on private chains [56]. When different national regulators attempt to audit the cross-border interactions between the German SME and the US provider, navigating these bespoke private chains is likely to create additional auditing friction. Attestation protocols could moderately standardize evidence if universally adopted, but relying on market-based adoption risks further fragmenting the compliance landscape. Without a critical mass of EU patients, hospitals, diagnostic apps and underlying AI model providers, cross-border supply chain audits remain fragmented by the persistence of traditional, non-standardized enforcement methods. VCs offer the highest capacity to structurally mitigate this deficit by standardizing the evidence base via open standards and the mandatory EUDI Wallet, harmonizing the mechanism of compliance [57]. Because the VC infrastructure is legally standardized across the EU, regulators could more seamlessly audit (at least the consent part of) the supply chain regardless of whether the actors are in Germany, France, or beyond.

*Mitigating Gold-Plating:* Blockchains perform poorly here, potentially generating further legal ambiguity regarding the interpretation of immutability [55]. This creates friction if the jurisdictions governing the hospital and the SME interpret absolute on-chain immutability differently. Attestation agents moderately mitigate gold-plating by abstracting compliance complexity, though they may face pressure to create localized variants of their open standards. The agent shields the hospital, SME and model provider from conflicting national rules, provided member states don't force the agent into divergent local compliance structures. VCs offer a high structural mitigation, harmonizing the mechanism of compliance and reducing cross-border operational burdens regardless of underlying legal fragmentation. A VC-based system works identically for the German SME and the European hospital, bypassing national discrepancies.

*Alignment with EU Digital Infrastructure:* Blockchains demonstrate only moderate alignment; while compliance guidelines exist [56] and they face foundational legal misalignments, the EuroStack nonetheless highlights them as a technology requiring EU involvement [18]. Thus, while the hospital could theoretically use EU-backed blockchain initiatives, the fundamental GDPR conflicts remain unresolved or require costly workarounds. Attestation protocols are conceptually aligned with EU goals but lack enforcement power if relegated to an entirely voluntary market framework [53]. Without a regulatory mandate, the Traceability Agent architecture remains an optional best practice rather than hard infrastructure. VCs possess strong alignment, integrating natively with the mandatory EUDI Wallet and the eIDAS 2.0 framework [57]. In accordance with the EUDI Wallet initiative, European health providers like our hypothetical hospital are intended to link to the wallet via critical use cases such as e-Prescriptions and the EU Health Insurance Card. Consequently, this integration positions the architecture as a foundational, interoperable-by-design layer for European digital sovereignty.

#### 3.3.3. Architectural tradeoffs

This systematic comparative analysis reveals core architectural tradeoffs. Blockchains offer robust mathematical trust and permissioned private or consortium blockchains with several other workarounds can be engineered for GDPR compliance, they still require significant latency trade-offs that make them unsuitable for high-volume AI operational (AI Act) and algorithmic decision (DSA) logging. Verifiable Credentials offer robust user-centric privacy, selective disclosure, and consent management, but were not designed to track high-volume, backend algorithmic data flows because they reside exclusively on the user's edge device. Attestation protocols like OTrace handle high-volume backend logging effectively but lack a native, standardized, and legally recognized mechanism for the user to assert their identity and consent seamlessly across borders.

### 3.4. A proposed hybrid architecture: EU-Trace

The preceding analysis shows that while blockchains can be significantly enhanced with workarounds to address the compliance issues documented in legal-technical scholarship, the private, permissioned, and layer-2 workarounds are essentially necessitating the use of blockchain as a private high-throughput database which it was not originally designed for. The OTrace and Verifiable Credentials have high levels of fit with some of the legal requirements, but neither can fulfill all legal and strategic requirements.

This paper proposes a novel hybrid architecture, "EU-Trace," that synthesizes the strongest components of the technical architectures analyzed. The core innovation is a decoupled pattern combining the secure, user-controlled authorization function of Verifiable Credentials with the continuous, delegated logging of attestation-based protocols (Fig. 1).

The EU-Trace model is a legal-technical pattern. While its technological design may be viable in a purely private, market-based context, it has the potential to be more trustworthy, interoperable, and effective when integrated with the EU's legal frameworks and emerging public digital infrastructure as we will discuss in the next chapters. However, aligning novel technical architectures to fit emerging, untested legal frameworks and standards is an inherently imperfect science. Therefore, this architecture is not presented as a technological panacea, but rather as a resilient compromise designed to operate within the fragmented realities of the European digital single market.

#### 3.4.1. The core pattern: a decoupled architecture

The proposed architecture is based on two decoupled functions:

*Authorization via Verifiable Credentials*: The management of permissions is conducted by the user through a secure, user-controlled mechanism using Verifiable Credentials. This provides a cryptographically verifiable record of consent.
*Logging via a Delegated Agent*: The high-volume task of logging subsequent data processing events is delegated by the user to a Traceability Agent, which collects attestations from service providers.

This pattern is versatile and could be implemented in a purely private, market-based ecosystem where companies compete to offer wallet and agent services under contractual agreements. The viability of this technical architecture does not strictly depend on a public-law embedding. Its strength lies in separating authorization from continuous logging. A purely private implementation, however, would likely replicate the problems the EU's digital rulebook seeks to solve: a fragmented market of non-interoperable solutions, a significant trust deficit for users, and possibly even the formation of new data monopolies. The EU-

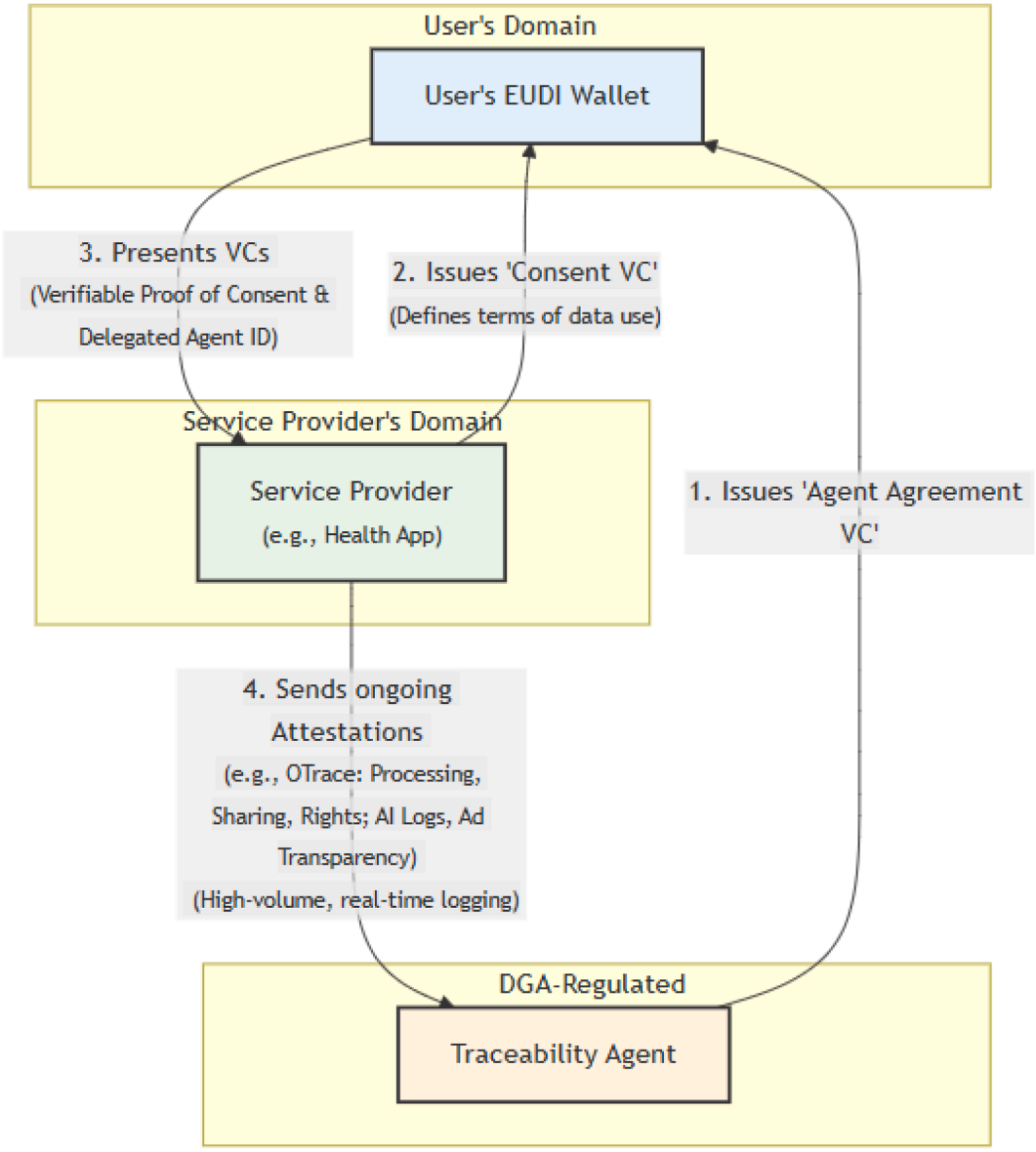


**Fig. 1.** The EU-trace hybrid architecture.

Trace model integrates this technical pattern with the EU's public digital infrastructure to solve these inherent risks.

*3.4.2. The "EU-Trace" synthesis: a resilient model for an imperfect world*

The EU-Trace synthesis leverages two pillars of the EU's digital strategy: the EUDI Wallet and the DGA as public trust enablers for the legal-technical architecture. This approach acknowledges that these pillars are imperfect but argues that their integration creates a system possibly more resilient to their known flaws (than the present form of the two pillars) and superior to a purely private alternative.

The EUDI Wallet acts as the mandatory, pan-European interoperability layer for all Member States by 2026 [51]. The wallet is a secure, user-controlled container for identity attributes and high-trust "electronic attestations of attributes," based on the W3C Verifiable Credentials data model. The EU-Trace model is designed to be resilient within the wallet's imperfect context, which includes security concerns, potential privacy leaks, scalability issues, and uncertain citizen uptake [61]. By deliberately offloading high-volume logging away from the wallet to the Traceability Agent, it mitigates the wallet's scalability and performance limitations. The EUDI Wallet's mandatory, standards-based nature is a possible path to overcoming the market fragmentation that would risk the effectiveness of a purely private system.

The Traceability Agent handles the high-volume task of collecting and storing granular traceability records. This model adopts the OTrace protocol and gives it a public law embedding within the DGA. The DGA creates a new type of regulated entity: the Data Intermediary [37]. Requiring the Traceability Agent to be a DGA-regulated entity has the possibility of transforming the system from a collection of private agreements into a trustworthy public ecosystem. The DGA's notification and registration process can provide a trust basis, while its legal mandate for neutrality and structural separation [37] has potential to solve the conflict-of-interest problem in having a data holder or user act as the record-keeper. Importantly, it also provides accountability through supervision by competent national authorities.

The EU-Trace model is technologically neutral with respect to the agent's internal architecture. While many agents might use traditional databases, the model does not prevent an agent from using a permissioned blockchain as its backend.

#### 3.4.3. A two-credential model for dynamic, user-controlled traceability

To make this architecture flexible and user-centric, another innovation is required: the user's EUDI Wallet to manage two distinct types of Verifiable Credentials. *The Consent VC* is the agreement between the data subject and the data controller. Issued by a service provider, it is a cryptographically verifiable record of the specific terms of a data processing agreement, detailing data categories, processing purposes, sharing permissions, and expiry date. *The Agent Agreement VC* is the auditor delegation. Issued by the user's chosen Traceability Agent, this VC acts as a pointer, containing the agent's identity and the secure endpoint where attestations should be sent.

This two-credential model is crucial for user control. A user can change their Traceability Agent at any time by acquiring a new Agent Agreement VC and revoking the old one, without affecting existing Consent VCs. This architecture must, however, address the "cold start" problem of VCs, including lack of effective, dependable and ubiquitous wallets [62]. Mandating the use of these VCs for access to the EuroStack and its Data Commons could create a catalyst for ecosystem development.

#### 3.4.4. Advancing the state of the art: from OTrace to "EU-Trace"

There are several aspects of novelty of the EU-Trace legal-technical architecture. It advances the state of the art through a novel technical synthesis and by grounding this synthesis in a unique legal and infrastructural context that makes it viable at a continental scale. It is an adoption of and an advancement beyond prior academic proposals like the OTrace protocol. The key differentiators are:

- *A Novel Hybrid Technical Pattern:* The core technical contribution is the combination of Verifiable Credentials for authorization with an OTrace-like agent for logging. This decouples the high-integrity act of granting consent from the high-volume task of recording data use.
- *Legal Embedding in the DGA:* While OTrace introduces the Traceability Agent as a core role, EU-Trace gives this agent a public law identity as a DGA-regulated Data Intermediary to enhance trust through strict neutrality and supervisory oversight.
- *Integration with Mandatory Public Infrastructure:* EU-Trace integrates its logic directly into mandatory, pan-European public infrastructure, the EUDI Wallet. This has the potential to transform a potentially fragmented, opt-in system to a foundational, interoperable-by-design feature of the digital single market.
- *Novel Two-Credential Architecture:* The explicit separation of the Consent VC and the Agent Agreement VC is a novel design contribution that allows a user to switch their Traceability Agent without re-negotiating every data-sharing consent, enhancing user control beyond the initial OTrace proposal.

Furthermore, EU-Trace acts as the necessary ex-post complement to ex-ante semantic ontologies and NLP compliance checking technologies (such as PrOnto and DIKAIO) discussed earlier. While those ex-ante tools are highly valuable for translating ambiguous legal text into structured system requirements during the design phase, EU-Trace provides the audit trail to prove that the deployed data supply chain actually adhered to those parameters in practice.

### 3.5. The EU-Trace operational framework

The EU-Trace legal-technical architecture functions as the operational framework necessary to fulfill the specific traceability imperatives identified across the EU digital rulebook. By decoupling secure authorization from continuous logging, the architecture satisfies three distinct regulatory tiers. First, for foundational GDPR data protection, the EUDI Wallet provides cryptographic proof of legal basis while the Agent dynamically maintains the Record of Processing Activities and data subject rights fulfillment logs. Second, for platform and ecosystem governance (DMA, DSA, DGA, Data Act), this model seamlessly handles granular consent for gatekeepers, receives output attestations for ad transparency, mediates third-party IoT data sharing, and technically implements the neutral data intermediation mandated by the DGA. Finally, under the AI Act, the Traceability Agent serves as the immutable backend ledger for automated event logging in high-risk systems, while also providing verifiable training data provenance for both high-risk and general-purpose AI models. Table 2 summarizes this operational mapping and cross-examines the EU-Trace framework against the legal imperatives of traceability identified in Chapter 2.

## 4. A traceability foundation for Europe's digital sovereignty

The legal imperatives for traceability and the EU-Trace architecture aligns with the EU's vision for digital sovereignty, the EuroStack initiative, to build a common European digital infrastructure to reduce technological dependencies and enhance competitiveness [18].

### 4.1. The eurostack and the data commons vision

The EuroStack initiative proposes a framework for a common digital stack built on sovereignty, security, and interoperability. Its vision of the Data Commons is being realized through Common European Data Spaces in areas like health, mobility, and finance [39].

The EuroStack vision, however, faces significant political and economic headwinds. The EU suffers from a major investment gap compared to its global competitors and persistent market fragmentation in key digital sectors, complicating the creation of unified infrastructure [18]. Overcoming these challenges requires technologies that can build trust and de-risk the ecosystem.

### 4.2. EU-Trace as a governance technology

The EU-Trace framework can act as a component for the EuroStack, potentially as a critical enabler. This integration reduces the friction of regulatory application. A standardized traceability layer aims to create a harmonized evidentiary baseline for auditing, providing regulators with a consistent, machine-readable data trail to enforce the digital rulebook more effectively across all Member States and tackling enforcement

**Table 2**
EU-trace operations for traceability imperatives.

| Regulation | Traceability Imperatives | EU-Trace Operations |
|---|---|---|
| GDPR | Demonstrable Accountability | Consent VC provides cryptographically verifiable proof of a valid legal basis. |
| GDPR | Records of Processing Activities | Traceability Agent log acts as a dynamic, real-time ROPA. |
| GDPR | Data Protection by Design and by Default | Decoupled architecture, selective disclosure via VCs enforce data minimization. |
| GDPR | Exercise of Data Subject Rights | Traceability Agent's log of Rights Attestations creates an audit trail for data subject requests. |
| DMA | Traceable Consent for Combining Data | Consent VC provides the specific, auditable proof of granular consent. |
| DSA | Transparency for Advertising Parameters | Traceability Agent receives "output attestations" from platforms detailing targeting parameters. |
| DGA | Neutral and Transparent Data Flows | Traceability Agent is the technical implementation of a DGA Data Intermediary. |
| Data Act | User Access and Third-Party Data Sharing | Consent VC from EUDI Wallet authorizes and logs third-party IoT data access. |
| AI Act | Automatic Event Logging for System Operations | High-risk AI systems send real-time event attestations to a designated Traceability Agent. |
| AI Act | Data Provenance for Training Data | AI developer generates a "Dataset Compliance Attestation" based on the verifiable consent history of all included data. |

fragmentation.

The architecture could also be combined with proposed political reforms to address the challenges of inconsistent enforcement and gold plating. It could act as governance technology to empower the proposed "granular coordination mechanisms." [15] The standardized, machine-readable audit trail from EU-Trace would provide a "Digital Regulation Coordinator" with a real-time, evidence-based view of the practical interactions between the regulations of the EU digital rulebook.

It offers a structural mechanism to enhance competitiveness, a key pillar of reform identified by CEPS/ECRI [15]. By offering "compliance-as-a-service" within the EuroStack, the model addresses a primary source of friction, the disproportionate regulatory burden on SMEs. Instead of investing in bespoke, fragmented compliance solutions, SMEs could connect to this infrastructure as a public utility.

Finally, integrating a traceability layer offers a uniform mechanism to verify compliance within the EuroStack, aligning data operations with European legal frameworks by design, rather than relying solely on post-hoc audits. This mechanism could create a proactive pathway to verify that Europe's digital infrastructure operates according to its own rules and values, not only reacting to sphere transgressions of big tech companies in the data ecosystems [63]. This approach offers a path to renewed global influence, not by exporting another regulation, but by pioneering an open, interoperable technical architecture for trustworthy data ecosystems. A verifiable, user-centric infrastructure built on open standards could be a privacy and AI-aligned candidate for global adoption. This would represent a new, more powerful form of the Brussels Effect [64], shifting from exporting rules to exporting functional, rights-preserving digital infrastructure.

### 4.3. A proportional, risk-based mandate for traceability

A one-size-fits-all mandate for traceability could impose a disproportionate burden on low-risk data processing activities and SMEs. A more nuanced approach, inspired by the risk-based model of the EU AI Act, is to implement a proportional, tiered model. This risk-based model would consist of two levels:

The first level is the *foundational traceability mandate***.** This baseline level applies to all data controllers. It mandates the implementation of robust, auditable internal logging systems sufficient to fulfill core GDPR obligations, particularly data subject rights. Chapter 2 has argued that this traceability requirement already exists in practice.
The second level is the *demonstrable traceability mandate*: This higher level of assurance could be made mandatory for activities classified as high-risk. Subject to policy debate, this could potentially include processing special categories of personal data (GDPR Article 9), sensitive financial data (PSD2, FIDA regulations), large-scale profiling, and any system designated "high-risk" under the AI Act. For these activities, controllers would be required to implement the full EU-Trace model, sending real-time, cryptographically signed attestations of all data processing events to a user-appointed, DGA-regulated Traceability Agent.

This tiered approach aligns with the EU's principle of proportionality, strengthens accountability where the risks of processing are high, and provides a scalable pathway for SMEs to participate in the data economy without facing undue compliance burdens.

### 4.4. Implementation challenges and limitations

The EU-Trace model's reliance on the EUDI Wallet and DGA-regulated intermediaries introduces dependencies on nascent ecosystems. While the core technical pattern of EU-Trace is viable on its own, its integration with the EU's public infrastructure (even with its flaws) offers a promising path to achieving a truly harmonized single market for data with the digital sovereignty of the Eurostack. However, both the EUDI Wallet and the DGA intermediary market brings inherent risks.

#### 4.4.1. The challenges of voluntary and mandatory adoption of the EUDI wallet

The EUDI Wallet is the designated interoperability layer for EU-Trace, but its successful, at-scale deployment is not guaranteed. Unresolved issues include the lack of clear business models, ongoing technical standardization challenges, and no guarantee of widespread citizen adoption [61]. This uncertainty is rooted in the central regulatory paradox of the eIDAS 2.0 regulation: while European Member States are strictly mandated to offer the wallet, and specific relying parties are mandated to accept it, its use remains strictly voluntary for citizens [51]. This dynamic introduces the risk of the "empty wallet" syndrome. Citizen indifference, user experience friction, and the lack of commercial incentives could cause adoption to stall [65]. While a niche group of highly privacy-conscious citizens could theoretically sustain the operational costs of Traceability Agents, a lack of widespread citizen adoption creates insurmountable service provider integration frictions. Without a critical mass of users, data controllers and gatekeepers lack the economic incentive (specifically, the reduction of compliance costs at scale) to integrate their backend APIs with the EU-Trace network. Should widescale EUDI Wallet adoption fail, the architecture would technically be forced to rely on commercial alternatives (such as Apple Wallet or Google Wallet) or more successful national digital ID solutions to issue Consent VCs. Relying on commercial alternatives would reinforce dependence on the very companies designated as "gatekeepers" under the DMA and contradict the digital sovereignty objectives of the EuroStack. The EUDI Wallet, despite its imperfections, remains the optimal path to a harmonized, pan-European evidentiary standard. In the absence of widespread adoption of the EUDI wallet, a viable sovereign fallback involves replacing it (at least in part) with more fragmented, but still digitally sovereign, nationally more successful alternatives to maintain interoperability and drive EU-Trace adoption. While theoretically feasible, technically addressing this "opening up" of the architecture to support multiple regional wallets while successfully integrating Verifiable Credentials requires additional research beyond the scope of this paper.

#### 4.4.2. The challenges of economic viability of DGA-regulated intermediaries

The market for DGA-regulated intermediaries is nascent and faces barriers to entry [66]. The DGA imposes strict rules of structural separation, legally forbidding these intermediaries from monetizing the data they hold or using it to improve their own ancillary service. This restriction raises a critical economic hurdle: what is the viable business model for a Traceability Agent? Without the ability to sell data, covering the server, cryptographic processing, and storage costs required to maintain real-time, high-volume operational logs for potentially millions of citizens requires alternative sustainability models. These agents might eventually operate as paid, premium subscription services for highly privacy-conscious citizens, or function as B2B "compliance-as-a-service" utilities funded by data controllers desperate to offload their regulatory reporting burdens, especially cross-border and especially considering the coming compliance obligations defined by the AI Act but not yet enforced. Alternatively, mitigating these profound commercial hurdles may ultimately require significant state subsidization to operate these intermediaries as foundational digital public infrastructure, potentially as part of the Eurostack. If the economic viability issues are addressed, the EU-Trace architecture could potentially serve as a catalyst for the DGA intermediary market. Thus, the proportional, risk-based mandate proposed in Section 4.3 could function as a mechanism to address both the regulatory deficit and the economic barriers of the DGA framework. Mandating traceability for high-risk AI and large-scale profiling forces data controllers to internalize these compliance costs, and could effectively bootstrap the B2B 'compliance-as-a-service' market required to sustain DGA-regulated

Traceability Agents.

#### 4.4.3. The risk of state surveillance and cyberattacks

For Traceability Agents to be trustworthy, legal compliance with the DGA is necessary but insufficient. Studies warn that digital identity systems utilizing a "core-complements" architecture inherently enable "platform-mediated surveillance" [67]. Applying this to the EU-Trace framework, the EUDI Wallet acts as the identity core and the Traceability Agent as the complement. Even if the underlying personal data remains decentralized at the service provider, the Traceability Agent's aggregation of metadata logs (detailing who interacted with what service, and when) creates a comprehensive map of a citizen's digital footprint. Consequently, data collected for a benign, legally mandated purpose becomes susceptible to weaponization by state power or malicious actors, as the architecture makes it technologically trivial to do so, a vulnerability seen when the Eurodac asylum database was made interoperable with national police databases [67].

There are well-established critiques of "surveillance capitalism" that relate to the aggregation of behavioral metadata [68]. While Zuboff warns against corporate digital platforms, the proposed architecture risks inadvertently replicating similar behavioral data "honeypots". Although the DGA forbids Traceability Agents from operating as commercial data brokers, the concentration of behavioral data may make these agents prime targets for malicious cyber actors, or foreign state intelligence agencies, similar to how data brokers became a "weak link in national security" [69].

A further concern is the ongoing debate regarding US extraterritorial laws and large cloud providers, which directly conflict with the GDPR and other EU laws. Recent disclosures in the French Senate highlighted that major US cloud providers cannot guarantee that European data will not be transferred to the US government if formally requested under US law [70]. While the EuroStack initiative explicitly encourages the use of sovereign, EU-based data centers to ensure digital autonomy, the current market dominance of United States-based providers may limit the practical technical and economic choices of Traceability Agents, including scalability and economic competitiveness, if they are forced to rely purely on smaller EU providers.

To mitigate state surveillance risks, relying solely on the DGA's legal mandate for neutrality is insufficient. Privacy-enhancing technologies, specifically end-to-end encryption, must be a mandated technical and operational measure for Traceability Agents. In practice, this means the Traceability Agent logs the existence of a data event (the timestamp and participants), but the actual details of the processing remain encrypted by an encryption key held at the Service Provider. If a data subject or a supervisory authority needs to read the log for an investigation, they must request the specific decryption key directly from the Service Provider. This mechanism can act as pragmatic tradeoff that is in line with current compliance reality: service providers are already responsible for supplying compliance information during audits or data subject requests. However, EU-Trace can make this process more efficient, because instead of a company spending weeks generating a bespoke, non-standard compliance report for a regulator, they simply hand over a decryption key to unlock the standardized, immutable traceability log held at the Traceability Agent. This maintains the centralized auditability of EU-Trace while mathematically ensuring the Agent itself possesses no readable behavioral data to surrender to mass surveillance requests or exploited by malicious actors. Alongside these technical safeguards, regulations must explicitly dictate that these encrypted logs (which inherently expose unencrypted metadata regarding the existence of user-provider relationships) exist exclusively for compliance verification. Any secondary access must require a specific judicial warrant.

## 5. Conclusion

### 5.1. Summary of contributions

This paper has argued that data traceability is a foundational pillar for the success of the EU's digital strategy. It established a legal imperative for traceability across the EU digital rulebook and proposed a concrete hybrid architecture ("EU-Trace") that leverages the EU's emerging digital identity infrastructure (EUDI Wallet) and legal framework for data sharing (DGA) to create a scalable and trustworthy system. This model was established as a significant evolution from prior academic proposals, tailored specifically to the EU's unique legal and strategic context. Finally, a strategic policy contribution was offered by articulating a vision for integrating this architecture as a foundational service layer within the EuroStack, using a proportional, risk-based governance model to transform compliance from a fragmented burden into a harmonized feature of Europe's public digital infrastructure.

### 5.2. Policy recommendations and future work

Based on this analysis, several concrete recommendations emerge. First, EU policymakers should formally recognize traceability as a critical digital infrastructure component. Adopt a risk-based approach to mandating traceability in the technical specifications for the EuroStack and the implementing acts for the EUDI Wallet, requiring demonstrable traceability for high-risk processing. Actively promote the role of DGA Data Intermediaries as the regulatory home for Traceability Agents, possibly even promoting publicly funded "utility" agents to support SMEs and bootstrap the market. Second, standards bodies (e.g., W3C, ETSI) should prioritize the development of standardized vocabularies and profiles for representing consent agreements, agent delegations, and data processing attestations using the W3C Verifiable Credentials data model to ensure pan-European interoperability. Finally, researchers should conduct further investigation into optimal hybrid governance models for a public traceability utility. The applicability of the EU-Trace model should be extended to other critical domains, most notably the governance of the AI lifecycle. The extension of traceability to AI training data is essential for solving the broader problem of AI explainability and accountability under the AI Act.

By embracing traceability as a core principle, the European Union can build a digital future that is not only innovative and competitive but also transparent and accountable.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

No data was used for the research described in the article.

## References

[1] Pathak M. Data governance redefined: the evolution of EU data regulations from the GDPR to the DMA, DSA, DGA, Data Act and AI Act. Eur Data Prot Law Rev 2024;10:43–56. https://doi.org/10.21552/edpl/2024/1/8.
[2] von DL. The European Union's pursuit of digital sovereignty through legislation. JIPITEC – J Intellect Prop Inf Technol E-Commer Law 2025;16.
[3] Magrani E, Alija N, Andrade F. Gold-plating' in the transposition of EU law. E-Pública 2021;8. https://doi.org/10.47345/v8n2art5.
[4] European Commission. The future of european competitiveness. part A, a competitiveness strategy for europe (Draghi report). Publications Office of the European Union; 2025. https://doi.org/10.2872/9356120.
[5] Demková GDG Simona. The enforcement dilemmas in europe's digital rulebook | techpolicy.press. Tech Policy Press; 2025. https://techpolicy.press/the-enforcement-dilemmas-in-europes-digital-rulebook. [Accessed 16 October 2025].

[6] European Parliament, Council. 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. GDPR; 2016.
[7] Pichlak M, Gaczoł K. Simple and advanced reflexivity in GDPR enforcement: empirical evidence from DPA activity. Int Data Priv Law 2023;13:267–83. https://doi.org/10.1093/idpl/ipad018.
[8] Cobbe J, Veale M, Singh J. Understanding accountability in algorithmic supply chains. In: 2023 ACM Conference on Fairness Accountability and Transparency. ACM; 2023. p. 1186–97. https://doi.org/10.1145/3593013.3594073.
[9] Baquero PM, Amariles DR, Amyot D, et al. The compliance gap in data supply chains: contract specification languages and smart contracts as compliance technologies. Artif Intell Law 2025. https://doi.org/10.1007/s10506-025-09464-8.
[10] C-604/22 IAB Europe v Gegevensbeschermingsautoriteit. 2024.
[11] CNIL. Cookies and advertisements inserted between emails: google fined 325 million euros. 2025. https://www.cnil.fr/en/cookies-and-advertisements-inserted-between-emails-google-fined-325-million-euros-cnil (Accessed 11 March 2026).
[12] CNIL. Cookies placed without consent: shein fined 150 million euros. 2025. https://www.cnil.fr/en/cookies-placed-without-consent-shein-fined-150-million-euros-cnil. [Accessed 11 March 2026].
[13] Garante. Comunicato stampa - ChatGPT, il Garante privacy chiude l'istruttoria. OpenAI dovrà realizzare una campagna informativa di sei mesi e pagare una sanzione di 15 milioni di euro. 2024. https://www.garanteprivacy.it:443/home/docweb/-/docweb-display/docweb/10085432 (Accessed 11 March 2026).
[14] European Data Protection Board. Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models. 2024. https://www.edpb.europa.eu/our-work-tools/our-documents/opinion-board-art-64/opinion-282024-certain-data-protection-aspects_en. [Accessed 11 March 2026].
[15] Arnal J. EU simplification will fail without better governance: three necessary reforms to make sure it doesn't fail. European Credit Research Institute (ECRI) and the Centre for European Policy Studies (CEPS); 2025. https://cdn.ceps.eu/2025/09/Simplification-and-governance-policy-brief-formatted.pdf (Accessed 16 October 2025).
[16] Pan B, Stakhanova N, Ray S. Data provenance in security and privacy. ACM Comput Surv 2023;55:1–35. https://doi.org/10.1145/3593294.
[17] Hellenic Data Protection Authority. The accountability principle. (n.d.) https://www.dpa.gr/en/Organisations/accountability (Accessed 15 October 2025).
[18] Bria F., Timmers P., Gernone F. EuroStack – a European alternative for digital sovereignty 2025:127 p https://doi.org/10.11586/2025006.
[19] Capodieci A, Mainetti L. A structured approach to GDPR compliance. In: Przegalinska A, Grippa F, Gloor PA, editors. Digit. transform. collab. Cham: Springer International Publishing; 2020. p. 233–43. https://doi.org/10.1007/978-3-030-48993-9_16.
[20] European Parliament, Council. Regulation (EU) 2018/1725 of the european parliament and of the council of 23 october 2018 on the protection of natural persons with regard to the processing of personal data by the union institutions, bodies, offices and agencies and on the free movement of such data, and repealing regulation (EC) no 45/2001 and decision no 1247/2002/EC. EUDPR; 2018.
[21] Personal Data Protection Office (UODO). Technical and organisational measures should be complementary. 2023. https://uodo.gov.pl/en/553/1486 (Accessed 16 October 2025).
[22] European Data Protection Board. Spanish DPA imposes fine of 1500,000 euros on EPD energía, S.A.U. for two infractions of the GDPR. 2021. https://www.edpb.europa.eu/news/national-news/2021/spanish-dpa-imposes-fine-1500000-euros-epd-energia-sau-two-infractions-gdpr_en. [Accessed 15 October 2025].
[23] C-252/21. Meta platforms Inc and others v bundeskartellamt. 2023.
[24] Data Protection Commission. Records of processing activities (RoPA) under article 30 GDPR. 2023.
[25] CNIL. Data brokers: tagadamedia fined €75,000. 2024. https://www.cnil.fr/en/data-brokers-tagadamedia-fined-eu75000. [Accessed 16 October 2025].
[26] European Data Protection Board. Guidelines 4/2019 on article 25 data protection by design and by default. 2020. https://www.edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-42019-article-25-data-protection-design-and_en. [Accessed 16 October 2025].
[27] C-487/21. F.F. v Österreichische datenschutzbehörde and crif gmbh. 2023.
[28] C-579/21. J.M. v. apulaistietosuojavaltuutettu, pankki s. 2023.
[29] C-634/21. OQ v land hessen, schufa holding ag. 2023.
[30] Information Commissioner's Office. Right of access. https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/individual-rights/right-of-access/(. [Accessed 16 October 2025].
[31] European Parliament, Council. Regulation (EU) 2022/1925 of the European parliament and of the council of 14 september 2022 on contestable and fair markets in the digital sector and amending directives (EU) 2019/1937 and (EU) 2020/1828. DMA; 2022.
[32] D'Amico AS. Market power and the GDPR: can consent given to dominant companies ever Be freely given? Eur Pap - J Law Integr 2023;8:611–29. https://doi.org/10.15166/2499-8249/678.
[33] European Commission. Commission finds apple and meta in breach of the digital markets act. 2025. https://ec.europa.eu/commission/presscorner/detail/en/ip_25_1085. [Accessed 16 October 2025].
[34] European Parliament, Council. Regulation (EU) 2022/2065 of the European parliament and of the council of 19 October 2022 on a single market for digital services and amending directive 2000/31/EC. DSA; 2022.
[35] European Data Protection Board. Guidelines 3/2025 on the interplay between the DSA and the GDPR. 2025. https://www.edpb.europa.eu/our-work-tools/documents/public-consultations/2025/guidelines-32025-interplay-between-dsa-and-gdpr_en (Accessed 15 October 2025).
[36] IAB Europe. IAB Europe releases DSA transparency implementation guidelines 1.1. 2025. https://iabeurope.eu/iab-europe-releases-dsa-transparency-implementation-guidelines-1-1/. [Accessed 16 October 2025].
[37] European Parliament, Council. 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. DGA; 2022.
[38] European Commission. Implementing the data governance act – guidance document. 2024.
[39] European Parliament, Council. Regulation (EU) 2025/327 of the European parliament and of the council of 11 February 2025 on the European health data space and amending directive 2011/24/EU and regulation (EU) 2024/2847. EHDS; 2025.
[40] Data Spaces Support Centre. Data spaces blueprint. 2024. | Version 0.5 | September 2023, https://dssc.eu/space/BPE/179175433. [Accessed 16 October 2025].
[41] European Parliament, Council. Regulation (EU) 2023/2854 of the European parliament and of the council of 13 December 2023 on harmonised rules on fair access to and use of data and amending regulation (EU) 2017/2394 and directive (EU) 2020/1828. Data Act; 2023.
[42] European Parliament, Council. Regulation (EU) 2024/1689 of the European parliament and of the council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending regulations (EC) no 300/2008, (EU) no 167/2013, (EU) no 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828. AI Act; 2024.
[43] Longpre S., Mahari R., Obeng-Marnu N., et al. Data authenticity, consent, & provenance for AI are all broken: what will it take to fix them? 2024. https://doi.org/10.48550/arXiv.2404.12691.
[44] European Commission. Explanatory notice and template for the public summary of training content for general-purpose AI models. 2025.
[45] Schwanke A. The EU AI Act: best practices for monitoring and logging. Medium 2025. https://medium.com/@axel.schwanke/compliance-under-the-eu-ai-act-best-practices-for-monitoring-and-logging-e098a3d6fe9d. Accessed 16 October 2025.
[46] CEN-CENELEC. CEN/CLC/JTC 21 work programme. (n.d.) https://www.cencenelec.eu/areas-of-work/cen-cenelec-topics/artificial-intelligence/ (Accessed 16 October 2025).
[47] Palmirani M, Martoni M, Rossi A, et al. PrOnto: privacy ontology for legal reasoning. In: Kő A, Francesconi E, editors. Electronic Government and the Information Systems Perspective.. Electronic Government and the Information Systems Perspective., 11032. Cham: Springer International Publishing; 2018. p. 139–52. https://doi.org/10.1007/978-3-319-98349-3_11.
[48] Cejas OA, Azeem MI, Abualhaija S, et al. NLP-based automated compliance checking of data processing agreements against GDPR. IEEE Trans Softw Eng 2023;49:4282–303. https://doi.org/10.1109/TSE.2023.3288901.
[49] Mildebrath H, Faucheux T. Second report on the application of the GDPR. 2024.
[50] European Union Agency for Fundamental Rights. GDPR in practice – experiences of data protection authorities. Publications Office of the European Union; 2024. https://doi.org/10.2811/768473.
[51] European Parliament, Council. 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. EUDI; 2024.
[52] Haque AB, Islam AKMN, Hyrynsalmi S, et al. GDPR compliant blockchains–a systematic literature review. IEEE Access 2021;9:50593–606. https://doi.org/10.1109/ACCESS.2021.3069877.
[53] Liao K., Thipireddy S., Weitzner D. Data traceability for privacy alignment 2025. https://doi.org/10.48550/arXiv.2503.09823.
[54] Sporny M, Longley D, Chadwick D, et al. Verifiable credentials data model v2.0. 2025.
[55] Finck M. Blockchain and the general data protection regulation: can distributed ledgers be squared with European data protection law? LU: European Parliamentary Research Service; 2019.
[56] European Data Protection Board. Guidelines 02/2025 on processing of personal data through blockchain technologies. 2025. https://www.edpb.europa.eu/our-work-tools/documents/public-consultations/2025/guidelines-022025-processing-personal-data_en. [Accessed 16 October 2025].
[57] European Digital Identity Cooperation Group. European digital identity wallet - architecture and reference framework 2025. https://eu-digital-identity-wallet.github.io/eudi-doc-architecture-and-reference-framework/2.6.0/architecture-and-reference-framework-main/ (Accessed 16 October 2025).
[58] Eberhardt J, Tai S. ZoKrates - scalable privacy-preserving off-chain computations. In: 2018 IEEE International Conference on Internet Things iThings IEEE Green Computing and Communication GreenCom IEEE Cyber Physical and Social Computing CPSCom IEEE Smart Data SmartData; 2018. p. 1084–91. https://doi.org/10.1109/Cybermatics_2018.2018.00199.
[59] Mazzocca C, Acar A, Uluagac S, et al. A survey on decentralized identifiers and verifiable credentials. IEEE Commun Surv Tutor 2025;27:3641–71. https://doi.org/10.1109/COMST.2025.3543197. –1.
[60] Song H, Qu Z, Wei Y. Advancing blockchain scalability: an introduction to layer 1 and layer 2 solutions. In: 2024 IEEE 2nd International Conference on Sensors Electronics and Computer Engineering ICSECE; 2024. p. 71–6. https://doi.org/10.1109/ICSECE61636.2024.10729396.
[61] Borak M. EUDI wallet sees progress but also criticism. https://www.biometricupdate.com/202502/eudi-wallet-sees-progress-but-also-criticism. [Accessed 16 October 2025].

[62] Bertocci V, Schenkelman D. Our take on verifiable credentials. Auth0 - Blog 2023. https://auth0.com/blog/our-take-on-verifiable-credentials/. [Accessed 17 October 2025].
[63] Stevens M, Kraaijeveld SR, Sharon T. Sphere transgressions: reflecting on the risks of big tech expansionism. Inf Commun Soc 2024;27:2587–99. https://doi.org/10.1080/1369118X.2024.2353782.
[64] Bradford A. The brussels effect: how the European union rules the world. Oxford University Press; 2020. https://doi.org/10.1093/oso/9780190088583.001.0001.
[65] Ølnes J. What could make the EU digital identity wallets fail? Biometric Update; 2026. https://www.biometricupdate.com/202602/what-could-make-the-eu-digital-identity-wallets-fail. [Accessed 11 March 2026].
[66] Johnston B, Milosavljevic N. The European commission's guidelines on the data governance act. Mason Hayes Curran 2024. https://www.mhc.ie/latest/insights/the-european-commissions-guidelines-on-the-data-governance-act. [Accessed 16 October 2025].
[67] Masiero S. Digital identity as platform-mediated surveillance. Big Data Soc 2023; 10:20539517221135176. https://doi.org/10.1177/20539517221135176.
[68] Zuboff S. The age of surveillance capitalism: the fight for a human future at the new frontier of power. London: Profile books; 2019.
[69] Reviglio U. The untamed and discreet role of data brokers in surveillance capitalism: a transnational and interdisciplinary overview. Internet Policy Rev 2022;11. https://doi.org/10.14763/2022.3.1670.
[70] Français S. CE commande publique : compte rendu de la semaine du 9 juin 2025. 2025. https://www.senat.fr/compte-rendu-commissions/20250609/ce_commande_publique.html. [Accessed 11 March 2026].