DATA, TECH & IP

MARCH 10, 2026

Op-Ed: “Portability Without Meaning: The Data Act’s Unfinished Architecture”

Ian Gauci

At critical moments in the EU’s digital regulatory history, ambition has outpaced architecture. The EU Data Act is the latest example. It did something quietly radical: it wrote semantic interoperability into the criteria for standards and specifications that can become mandatory for cloud and other data processing services. The question is whether the implementing architecture will honour that ambition, or quietly downgrade it into a checklist of file formats and transport protocols.

The European Interoperability Framework (EIF) defines semantic interoperability as ensuring that the precise format and meaning of exchanged data and information is preserved and understood, and that what is sent is what is ultimately intelligible. That is the real test: not that data passes from one system to another, but that the intelligence within it transits along and is understood. If the recipient cannot reconstruct the sender’s meaning, the exchange is not interoperable, even if it is technically flawless. The Interoperable Europe Act, which entered into force in April 2024, reinforces this framework. It establishes a governance structure – the Interoperable Europe Board, a central portal, mandatory interoperability assessments for trans-European digital public services – that takes semantic layers seriously, as the Commission’s first Annual Report on Interoperability confirms.

Article 35 calls for open interoperability specifications and harmonised standards that let services of the same type work together, make digital assets portable, support functional equivalence, and address semantic data interoperability. It also anchors the central Union standards repository. Once a reference is listed there, the Commission expects

providers to adapt within twelve months. Meaning, for the first time in EU data infrastructure law, it comes with a timetable.

Most cloud migrations fail in ways dashboards cannot detect. The files move, the containers start, the logs look clean, yet the business logic begins to drift because a field that used to mean one thing now means something else. Portability moves data. Interoperability ensures that data is information, thereby carrying meaning and value. That distinction is not semantic pedantry. It is the difference between a regulation that transforms markets and one that generates compliance paperwork.

The Repository will Decide whether the Promise is Real

The Commission has now published the results of its study on interoperability of data processing services, designed to support the first batch of repository content. The study builds an evaluation methodology drawing on the Common Assessment Method for Standards and Specifications (CAMSS) and the attributes in Annex II of the European Standardisation Regulation. That is sensible architecture. But sensible architecture is not the same as sufficient architecture. Recent academic analysis has already warned that standards embed policy choices. The repository is where those choices become obligations.

What the Semantic Web Taught us and What we Ignored

The irony is that Europe is arriving late to a debate the web has carried for decades. The original Semantic Web programme, conceived by Berners-Lee, Hendler and Lassila, promised a web where meaning could be represented in ways machines could reason about. It gave us durable foundations – RDF, OWL – and it inspired a long tradition of vocabulary design and ontology engineering. It also generated an equally durable scepticism, famously captured in Clay Shirky’s critique of the idea that the world’s categories can be cleanly encoded once and for all.

The data spaces ecosystem has been developing practical tools to address this problem. The Data Spaces Support Centre (DSSC) has been explicit that shared data models, vocabularies, and semantic mapping services are the scaffolding that makes exchange usable. The SEMIC

Support Centre has spent years doing the unglamorous work of keeping vocabularies and interoperability solutions findable and reusable, including the DCAT-AP family that many public data catalogues now rely on. At the W3C level, DCAT itself is defined as an RDF vocabulary designed to facilitate interoperability between data catalogues, most recently in DCAT Version 3.

The question is therefore not whether the tools exist but whether the Union repository will treat them as optional background reading, a footnote in a preparatory study, or as part of what it means to comply with Article 35. This is where Edwards and Veale’s insight becomes directly relevant: transparency mechanisms must move beyond individual rights-based approaches toward systemic regulatory frameworks capable of governing complex technical systems. The same logic applies to interoperability. An individual’s right to data portability means nothing if the receiving system cannot parse the data.

Why this Matters Now: Automated Agents, AI Act Exposure, and DORA’s Silent Assumption

Semantic drift used to be a human problem. Two teams defined the same label differently, someone eventually noticed, and the meeting that nobody wanted was scheduled. The emerging risk is machine speed. Automated agents can ingest, merge and act on data across services without ever asking what a term was supposed to mean. When an agent is wrong, it is confident.

This is precisely the regulatory exposure addressed by the AI Act. For high-risk systems, Article 10 requires training, validation and testing data sets to be relevant, sufficiently representative, and as free as possible from errors for the intended purpose. Semantic drift turns that obligation into a structural vulnerability. Data can be technically valid and fully portable under the EU Data Act, yet still be functionally wrong for the model that consumes it because the model was trained on one meaning and is now being fed another.

The same silent assumption underpins operational resilience requirements in the Digital Operational Resilience Act (DORA). DORA tests the availability, integrity and continuity of ICT systems, but it assumes that data flowing between those systems retains its semantic coherence. A system can pass resilience testing and remain operational

while its business logic quietly drifts. That is not resilience. It is an undetected degradation.

This is why Article 35 of the EU Data Act matters. By embedding semantic interoperability into standards that may become mandatory through the Union repository, it recognises that portability without preserved meaning is not interoperability. Without enforceable semantic artefacts and validation mechanisms, Europe risks building a switching regime that moves data efficiently while steadily eroding the intelligence embedded within it.

A Practical Way to Make Meaning Testable

The common mistake is to treat interface documentation as a semantic solution. An API description can tell you that a field is a number. It cannot tell you whether that number is a probability, a ranking, a threshold, or a vendor’s internal score. Standards that focus on transport and syntax are valuable, but they are not semantic standards. The Union repository cannot function without these enhancements. It needs a minimum that can be verified.

First, a declared semantics layer. Every data element listed in a standard should include a plain-language definition and, where available, a link to a controlled vocabulary that clarifies its meaning. Think of it as a data dictionary that travels with the specification, not buried in a developer wiki, but published as part of the standard itself. This is not a novel requirement. It is how the SEMIC Support Centre has operated for years, and it is how the Interoperable Europe Act envisages cross-border data exchange in the public sector. Without it, two providers can implement the same standard and still disagree on what a field means.

Second, reusable reference artefacts, meaning shared code lists, vocabularies, and ontology fragments that providers can adopt rather than reinvent in each contract. When a standard says ‘country’, there should be one agreed list of country codes, not a dozen proprietary variants. The DCAT-AP family already demonstrates that this approach is both feasible and scalable: common building blocks, maintained centrally, referenced by everyone.

Third, validation tools that can detect semantic divergence before it causes operational harm. The W3C’s SHACL, a language for describing and checking the structure and meaning of data, offers exactly this capability. It lets you write machine-readable rules that test whether data conforms not just to the right format but to the right meaning. The important point is not the acronym. It is the principle: if you can define what a data element should mean, you can build a test that catches the moment it stops meaning that.

Conclusion

Europe has made a constitutional move by writing semantic interoperability into binding criteria for standards that can become mandatory, but the implementation side is not there.

If the first repository wave sets only a syntactic floor, we will standardise misunderstandings and call the result compliance. If it insists on minimal semantic artefacts, validated by bodies with the relevant expertise, and testable against real-world data flows from day one, then switching can move data without losing the plot. The difference between those two outcomes is between a regulation that reshapes how Europe’s digital infrastructure communicates and one that merely reshuffles its filing systems.

Only one of those outcomes deserves the word interoperability. And only one is worthy of the constitutional ambition that Article 35 embodies.

Ian Gauci is Managing Partner at GTG, Malta, and Chair of Malta’s AI in Fintech Strategy Group. He co-authored Malta’s Virtual Financial Assets Act and advises on EU digital regulation across financial services, AI, and data governance.