The File Format Interoperability Act

A legislative proposal for open file formats, consumer choice, and digital sovereignty

Ric Harvey

Draft v0.2 — April 2026

PROPOSED LEGISLATION · POLICY WHITE PAPER

The File Format Interoperability Act

A legislative proposal for open file formats, consumer choice, and digital sovereignty.

Draft v0.2 · April 2026

For circulation to legislators, civil society, and industry stakeholders

Executive Summary

Modern economies run on files. Contracts, medical records, architects' drawings, financial models, scientific data, court bundles, the everyday output of every desk worker in the country: all of it sits in formats designed and controlled by private vendors. When those formats are proprietary and undocumented, the data inside them belongs, in practice, to the vendor. The customer pays for the licence, creates the content, and still cannot leave.

This paper proposes the File Format Interoperability Act, a short and targeted statute. Any commercial vendor offering software in the jurisdiction must publish a complete, accurate technical specification of every file format that product reads or writes. The specification must be detailed enough that a competent third party could build a fully compatible reader and writer without reverse engineering.

The Act does not force anyone to adopt a particular open standard. Vendors remain free to invent, evolve, and differentiate. They just cannot hide the door.

The benefits show up immediately and compound. Customers gain a credible exit option, which restores the discipline of the market. Competitors gain a level field on which to win by building better products rather than steeper moats. Governments gain genuine procurement choice and a real path to digital sovereignty. Citizens, in the long run, gain the right to read their own files in fifty years' time without paying rent to whoever happens to own the format that decade.

Part I. The Problem

Vendor lock-in as market failure

The defining feature of file format lock-in is that the customer rarely encounters it at the point of sale. They encounter it years later, when migration costs have quietly compounded into the millions, when an entire department has been trained on a single tool, and when the only readable copy of a decade of institutional memory exists in a binary blob that nobody outside the vendor can fully parse.

This is not an incidental side-effect of software design. It is a deliberate commercial strategy, well documented in industry literature, and it produces predictable harms:

- Customers pay above-market prices because switching is expensive, not because the product is best in class.
- Competitors cannot enter markets even when they have a superior product, because no buyer can risk stranding their existing data.
- Public bodies become structurally dependent on vendors over which their elected representatives have no jurisdiction.
- Long-lived records (medical, legal, scientific, cultural) become unreadable when the originating product is discontinued or its licensing terms change.
- National infrastructure becomes hostage to the commercial decisions, sanctions regimes, and political preferences of foreign legislatures.

Why the market has not solved this

Markets correct lock-in only when buyers can see and price the cost of leaving at the moment of purchase. They cannot. The cost of leaving is unknown until you try, and by the time you try, you have already paid it. This is a textbook information asymmetry, and it is the kind of failure that legislation exists to correct.

Voluntary adoption of open formats has had partial success in narrow domains, notably document interchange and image encoding. It has failed almost entirely in the categories that matter most: line-of-business applications, computer-aided design, scientific instrumentation, healthcare records, financial reporting tools, and the bespoke formats produced by enterprise software at every layer of the stack.

> *"The customer pays for the licence, creates the content, and still cannot leave."*
>
> — The central asymmetry the Act seeks to correct

Digital sovereignty

Several governments, including those of the United Kingdom, France, Germany, and the European Union, have stated publicly that reducing dependence on a small number of foreign technology providers is a strategic priority. The stated goals include resilience against geopolitical disruption, jurisdictional control over citizen data, and the ability to enforce domestic law on critical national systems.

These goals cannot be achieved while the file formats underlying public-sector data are proprietary. A government may host its own servers, employ its own engineers, and buy its hardware from domestic suppliers, and still find itself unable to read its own records without licensing software from a single foreign company. Hardware sovereignty without format sovereignty is a half-built bridge.

The Act addresses this by ensuring that, regardless of which product a public body chooses, the underlying data remains technically accessible to any other competent vendor or in-house team. Sovereignty is restored not by favouring any single vendor, but by removing the advantage of market dominance by any single business.

Part II. Statement of Principles

The following principles inform every operative provision of the proposed Act. Where ambiguity arises in interpretation, courts and regulators should resolve it consistently with these principles.

1. Data created by a customer belongs to the customer. The format in which that data is stored should not be a mechanism for asserting a claim over it.
2. A specification, once published, is a public good. Publishing it costs the vendor little and yields the customer a great deal.
3. Competition between vendors should turn on the quality of features, performance, support, and price. Never on the inability of customers to leave.
4. Long-lived records of public, scientific, legal, and cultural value must remain readable beyond the commercial lifetime of any single product.
5. National sovereignty over data infrastructure requires format sovereignty as a precondition.
6. Innovation is not threatened by openness. It is sharpened by it. Vendors with genuinely superior products have nothing to fear from a customer who is free to leave and chooses not to.

Part III. What the Act Does

What follows is a plain-English summary of the operative provisions intended for inclusion in the final statute. It is drafted to be jurisdiction-neutral; specific drafting will follow the conventions

of the enacting legislature.

Section 1. Definitions

For the purposes of this Act:

- "Covered software" means any software product offered for sale, licence, lease, subscription, or any other commercial arrangement to customers in the jurisdiction, where the product reads from or writes to a file, document, database, or other persistent data structure as part of its normal operation.
- "Covered format" means any file format, encoding scheme, schema, or data layout used by covered software to persist customer-generated data, configuration, or any other content the customer might reasonably wish to retain, transfer, or migrate.
- "Specification" means a written technical document, in machine-readable form, sufficient for a competent independent engineer to implement a fully compatible reader and writer of the format without recourse to reverse engineering, leaked materials, or undisclosed trade knowledge.
- "Commercial vendor" means any natural or legal person who offers covered software in exchange for consideration of any kind, including but not limited to direct payment, advertising revenue, data collection, or bundled subscription.

Section 2. Disclosure obligations

As a condition of offering covered software in the jurisdiction, a commercial vendor shall:

1. Publish a complete and accurate specification of every covered format used by the software, in a publicly accessible location, free of charge, and without registration or licensing requirements.
2. Update the specification at or before the moment any change to the format is released to customers in production builds.
3. Maintain historical versions of the specification for not less than fifteen years after the format ceases to be supported in shipping products.
4. Grant an irrevocable, royalty-free, worldwide patent licence to any party implementing a reader or writer of the format in accordance with the specification, for the limited purpose of such implementation.
5. Refrain from any contractual term, end-user licence provision, or technical measure that would prohibit, penalise, or materially impede a customer or third party from implementing, using, or distributing such an alternative reader or writer.

Section 3. Standards and quality

A specification published under this Act must:

- Be technically complete, including all binary structures, container formats, compression schemes, character encodings, and any cryptographic or integrity-checking mechanisms necessary to produce a valid file.
- Be unambiguous, such that two independent implementations following only the specification will produce interoperable output.
- Be available in at least one widely-used human-readable format such as Markdown, HTML, or PDF, and where applicable, in a machine-readable schema language.
- Disclose any dependency on third-party formats, libraries, or specifications, with reference to their respective public sources.
- Carry no terms that restrict commercial or non-commercial use of the resulting implementations.

Section 4. Enforcement

Enforcement should be light-touch but credible. The Act proposes a tiered remedy structure designed to encourage compliance rather than to punish good-faith effort:

1. A designated regulator, sitting within an existing competition or digital markets authority, accepts and investigates complaints from any party with standing, including customers, competitors, and civil society organisations.
2. On a finding of non-compliance, the regulator issues a notice requiring publication within a reasonable period, typically ninety to one hundred and eighty days depending on the complexity of the format.
3. Continued non-compliance after such notice attracts civil penalties scaled to the vendor's turnover from the affected product within the jurisdiction, with a statutory ceiling sufficient to deter wilful disregard.
4. Persistent and material non-compliance may result in an order prohibiting the offering of the affected product within the jurisdiction until the specification is published.

Section 5. Exemptions and limits

The Act recognises that disclosure obligations must respect competing interests. The following exemptions apply:

- Software produced by a single individual or small enterprise below a defined revenue threshold, on a sliding scale, with full obligations applying only above the threshold.
- Free and open-source software whose source code is already public under an OSI-approved licence is deemed to satisfy the disclosure requirement, provided the source itself documents the format with reasonable clarity.
- Formats whose disclosure would compromise the security of cryptographic systems, where the format is itself the cryptographic primitive. This exemption does not extend to file containers that merely happen to be encrypted.
- Formats used purely for internal vendor operations, that never persist customer data and are never exposed to customers as a portable artefact.

Part IV. Proof from History

Critics of this proposal often argue that opening up file formats will harm innovation. The historical record argues the opposite. Every major transition from a proprietary format to an open one has been followed by an explosion of adoption, a wave of new entrants, and a level of interoperability that the closed era could not produce. A few examples are worth dwelling on, because they are not edge cases. They are the foundations on which the modern internet runs.

PDF: a proprietary format that became universal once it opened

Adobe created PDF in 1993 and controlled it as a proprietary format for fifteen years. It did well in that period, but its real ascent began in 2008, when Adobe released the specification to the International Organization for Standardization. PDF 1.7 became ISO 32000-1, and the format passed under the control of an ISO committee on which Adobe holds one vote among many. PDF 2.0, published as ISO 32000-2 in 2017, removed the last of the proprietary references entirely[1]. Today the specification is freely available, and PDF is the closest thing the world has

[1] ISO 32000-2:2020, Document management, Portable document format, Part 2: PDF 2.0. Free copy hosted by the PDF Association at https://pdfa.org/resource/iso-32000-2/

to a universal document format. Adobe still sells excellent PDF tools, and competes with dozens of others on quality. The format is bigger, healthier, and more competitive than it ever was as a closed property.

This is the dynamic the Act is designed to engineer. PDF was opened voluntarily, after fifteen years of leverage. The Act asks vendors to do the same thing without waiting fifteen years.

PNG: open from the start, and unstoppable because of it

PNG was designed in the mid-1990s as a deliberate open replacement for GIF, which at the time was encumbered by a software patent on its compression algorithm. The PNG specification was published as a W3C Recommendation in 1996, became ISO/IEC 15948 in 2003, and was updated to a third edition in 2025[2]. Because the format was open and patent-free from day one, every browser, every operating system, every image library, and every camera manufacturer could implement it without paying a licence fee or asking permission. The result is that PNG is now ubiquitous, while the proprietary formats it competed against have either become open in turn or faded from use. Open standards do not lose. Closed ones either open up or die out.

TCP/IP, ODF, OCI: the same story, told repeatedly

The Transmission Control Protocol[3] and the Internet Protocol[4] were openly specified through the Internet Engineering Task Force, with no patent royalties and no licensing gatekeeper. They displaced a generation of proprietary networking stacks not because they were technically superior at the outset, but because anyone could implement them. The OpenDocument Format[5] did the same thing for office documents. The Open Container Initiative[6] did it for the container runtimes that power most of modern cloud computing. In every case, opening up the surface of integration produced more competition, faster innovation, and more durable infrastructure than closed alternatives.

The pattern is consistent enough to count as a rule. Open formats survive their creators. Closed ones do not.

Part V. Anticipated Benefits

Competition and innovation

When customers can leave, vendors must earn their renewal. Studies of markets that have moved toward greater interoperability, including telecommunications number portability, banking account switching, and the early-stage progress of the European Digital Markets Act, consistently show price reductions, faster product improvement, and a broader set of competitive entrants. The same dynamics will apply to software.

Critics argue that openness will reduce vendor incentive to invest. The historical record, set out in Part IV, suggests the opposite. The most innovative segments of the technology industry are

[2] Portable Network Graphics (PNG) Specification, Third Edition, W3C Recommendation, 24 June 2025. https://www.w3.org/TR/png-3/ Also published as ISO/IEC 15948.

[3] Transmission Control Protocol, RFC 9293, Internet Engineering Task Force, August 2022. https://www.rfc-editor.org/rfc/rfc9293

[4] Internet Protocol, RFC 791, Information Sciences Institute, September 1981. https://www.rfc-editor.org/rfc/rfc791

[5] OpenDocument Format for Office Applications, Version 1.3, OASIS Standard, 2021. Also ISO/IEC 26300. https://docs.oasis-open.org/office/OpenDocument/v1.3/

[6] Open Container Initiative Image Specification, maintained by the Linux Foundation. https://github.com/opencontainers/image-spec

precisely those built on open standards: the web, electronic mail, the Linux ecosystem, container runtimes, and the entire layer of internet protocols beneath them. Innovation flourishes when the surface of differentiation is real product capability rather than artificial barriers to exit.

Customer and business empowerment

For ordinary businesses, the Act removes a category of risk that today is invisible until it becomes catastrophic. A small architectural practice will be able to change CAD vendors. A medical clinic will be able to migrate its records to a competing platform without paying a ransom. A school will be able to retain twenty years of pupil work in a form that any future system can read. The cumulative effect on small and mid-sized organisations is, in aggregate, a substantial transfer of bargaining power back from vendors to customers.

Data sovereignty and the public sector

Public bodies can adopt the policy as a procurement requirement immediately, well before any statutory obligation comes into force. Any government tender for software that handles citizen data should require, as a precondition of participation, that the vendor publish a specification meeting the standard described above. This would place public-sector buying power behind the principle without waiting for the legislative cycle to complete.

Internationally, the Act aligns with stated policy goals across multiple democracies. It would be straightforward to coordinate enactment across allied jurisdictions, producing a transnational floor of openness that no single vendor could unilaterally route around.

National resilience

Critical national infrastructure today rests on a small number of vendors, several of which are headquartered in jurisdictions whose foreign policy may, in any given decade, diverge from the host country's. The Act does not require divestment from any vendor. It simply ensures that, in a contingency, a domestic alternative can be built and a transition can occur. This is the digital equivalent of maintaining strategic reserves: the existence of an exit option is itself a deterrent against the abuse of dependence.

Part VI. Implementation Timeline

The Act contemplates a phased rollout. Vendors get a reasonable period to comply without disrupting ongoing business operations:

- Months 0 to 6: Statute enacted. Designated regulator established. Consultation period for guidance documents.
- Months 6 to 18: Public-sector procurement clauses take effect for new contracts. Vendors of widely-deployed enterprise products receive priority engagement.
- Months 18 to 36: Disclosure obligations apply to all covered software above the small-enterprise threshold. Enforcement powers become operative.
- Months 36 onward: Sliding-scale obligations extend down through smaller vendors. Periodic statutory review every five years.

Part VII. International Alignment

Several existing legal instruments touch on related concerns and would benefit from explicit alignment with the proposed Act.

The European Union Digital Markets Act[7] establishes interoperability obligations for designated gatekeepers. The proposed Act extends a narrower obligation, format disclosure rather than full interoperability, to a wider class of vendors. Where the DMA targets the largest platforms, this Act addresses the long tail of vendor lock-in across the rest of the software industry.

The European Data Act[8], in force from 2025, provides rights of access and portability for data generated by connected products. Format disclosure is the logical and necessary complement: portability rights mean little if the format the data lands in is itself closed.

The General Data Protection Regulation[9] grants data subjects a right to portability of their personal data in a structured, commonly-used, machine-readable format. The proposed Act ensures that such formats actually exist and remain readable across vendor boundaries.

Public procurement frameworks in several Member States already preference open standards. The Act formalises and extends this principle, making it a baseline rather than a procurement preference.

Part VIII. Anticipated Objections

"This will harm innovation."

It will not. Innovation happens in the visible features customers buy, not in the hidden formats that hold them hostage. Part IV walks through the historical record in detail. Every major open-standard transition in computing, from networking to documents to container runtimes, has accelerated rather than slowed the pace of innovation.

"Trade secrets will be exposed."

A file format is not a trade secret in any defensible sense. It is the surface across which the product communicates with the data the customer has paid to create. The clever algorithms, optimisation techniques, and product designs that constitute genuine intellectual property remain entirely within the vendor's implementation. Publishing how a file is laid out exposes none of this.

"Compliance will be expensive."

Internal format documentation already exists at every serious software company. It has to, or the product could not be maintained. The cost of compliance is principally the cost of cleaning that documentation for publication. For new products, the cost is essentially zero if the discipline is built in from day one.

[7] Regulation (EU) 2022/1925 of 14 September 2022 on contestable and fair markets in the digital sector (Digital Markets Act). https://eur-lex.europa.eu/eli/reg/2022/1925/oj

[8] Regulation (EU) 2023/2854 of 13 December 2023 on harmonised rules on fair access to and use of data (Data Act). https://eur-lex.europa.eu/eli/reg/2023/2854/oj

[9] Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data (General Data Protection Regulation), Article 20 on the right to data portability. https://eur-lex.europa.eu/eli/reg/2016/679/oj

"This disadvantages domestic vendors against foreign ones."

On the contrary. The Act applies to any vendor offering covered software in the jurisdiction, regardless of where they are headquartered. It levels the field by removing a structural advantage that today accrues disproportionately to incumbents and to vendors with the resources to maintain proprietary moats.

"The big vendors will simply leave the market."

They will not. The Act applies wherever software is sold, and the markets it covers are too large and too lucrative to abandon. The same argument was made before every major consumer-protection law of the last fifty years. Vendors complied, kept selling, and in most cases prospered.

Part IX. Conclusion

The proposed Act is not a radical intervention. It does not nationalise software, mandate any particular technology, or favour any one vendor. It does one thing. It requires that the door out of every product be visible from inside it.

That single requirement, modest as it sounds, restores the basic conditions under which markets work and under which sovereign states can govern their own digital infrastructure. It returns to citizens and businesses something they assumed they already had: ownership of the data they create.

The technology industry will adapt to it, as it has adapted to every previous wave of consumer-protection legislation, and will, on the other side, be more competitive and more trustworthy than it is today. The cost of inaction is the steady, quiet, ongoing transfer of economic and democratic agency from the people who generate data to the small number of firms that happen to control the formats it sits in.

The case for action is strong. The case for delay grows weaker every year.

End of Proposal

Comments, drafting suggestions, and expressions of support
are welcomed during the consultation period.
