MarketImpact
SYSTEMS. STRATEGY. SAFETY.

WORKING PAPER WP-2026-03

Infrastructure Sovereignty and Humanitarian Neutrality

Commercial, Legal, and Political Risks of Jurisdictional Dependency in Humanitarian Digital Infrastructure

Thomas Byrnes, MarketImpact Digital Solutions Ltd, Nicosia
Version 1.2, July 2026. Working paper for discussion; not peer reviewed. An essay version of this analysis was circulated as "Your AI Has a Passport Problem" (13 June 2026, revised 3 July 2026).
Contact: thomas@marketimpact.org | marketimpact.org
Licence: CC BY-NC 4.0
Suggested citation: Byrnes, T. (2026), "Infrastructure Sovereignty and Humanitarian Neutrality: Commercial, Legal, and Political Risks of Jurisdictional Dependency in Humanitarian Digital Infrastructure", MarketImpact Working Paper WP-2026-03, Nicosia: MarketImpact Digital Solutions Ltd.

Abstract

In June 2026, a US export-control directive required Anthropic to restrict access to two frontier AI deployments by foreign nationals. Because the company said it could not verify nationality at scale in real time, it suspended access for all users. The controls were lifted on 30 June 2026, after government review, new safeguards, and new government-facing testing commitments. This paper argues that the episode exposed a wider condition: many common humanitarian workflows, including email, file storage, coordination platforms, analytics, and AI-assisted tools, depend on providers subject to foreign legal compulsion, commercial discretion, sanctions, export controls, and national-security authorities. The paper develops a three-part framework of sovereignty risk: commercial risk from provider discretion over continuity and terms; legal risk from extraterritorial data-production and separate foreign-intelligence authorities; and political risk from state restrictions on access to frontier capabilities. It argues that these risks converge on perceived humanitarian neutrality and access where host authorities perceive that beneficiary or operational data may sit within reach of a state with strategic, military, intelligence, sanctions, donor, or alliance interests in the crisis. The 2019 WFP-Houthi biometric-registration dispute in Yemen illustrates the access mechanism: disputes over control of beneficiary-data systems can interrupt assistance. The analysis is conceptual and documentary; it uses no primary data. It closes with practical steps for jurisdictional review, sensitive-data separation, infrastructure diversification, portable AI-methodology training, and contingency planning.

Keywords: humanitarian technology, jurisdictional dependency, digital sovereignty, data governance, perceived neutrality, CLOUD Act, FISA Section 702, export controls, artificial intelligence

1. Introduction: the June 2026 episode

On 12 June 2026, Anthropic said it had received a US government directive requiring it to suspend access to Fable 5 and Mythos 5, two frontier model deployments sharing an underlying model, one general-use and one a restricted defensive-cyber deployment, for foreign nationals, including foreign-national employees inside and outside the United States [1]. Anthropic said the directive arrived at 5:21pm on a Friday and did not explain the specific national-security concern [1]. Because the company said it had no reliable way to verify nationality across users and deployments in real time, it suspended both deployments for all customers [1][2]. The government did not order a global shutdown; a nationality-based access condition produced one through the provider's compliance response.

On 30 June 2026, the controls were lifted. Fable 5 returned to global availability from 1 July. Mythos 5 returned more narrowly, initially to a set of US organisations following government approval, and remained subject to staged restoration [3]. Anthropic also announced new safeguards, deeper information-sharing arrangements with government, and pre-release government testing commitments for future models [3]. The Commerce Secretary described the process as working with Anthropic to "analyze and approve" the model and "ensure alignment across the US Government" [4].

The outage was temporary. The governance lesson was not. Access to frontier AI can be altered quickly by state authority, and restoration can be conditioned on government-facing review and safeguards. Access did not simply resume automatically. It was restored after government review, new safeguards, and new testing commitments.

Most sector commentary treated this as an AI story that resolved itself: a capability restricted, a nervous government, a negotiation, access restored. This paper argues that reading misses the structural condition the episode exposed. The Anthropic case is a symptom. The condition is that many common components of humanitarian operational infrastructure, the email, the file storage, the coordination platforms, the databases, the analytics, are supplied by a small number of providers with material exposure to the legal process and political authority of the same foreign jurisdiction.

Two clarifications frame everything that follows.

First, terminology. This paper uses **jurisdictional dependency** to mean dependence on providers that can be compelled, restricted, or sanctioned by a foreign state. It does not mean that the state operates the infrastructure or has direct access to customer data. The relevant state capacities differ by layer: stored-data production through legal process, foreign-intelligence collection under separate authorities, export controls on capabilities, and sanctions on transactions. The June episode is a visible example of state override capacity at one layer; other legal mechanisms reach other layers in different ways.

Second, the destination of the argument. The practical risk examined here is not only disclosure. It is access. If a host authority believes that beneficiary or operational data sits within legal reach of a foreign state with interests in the conflict, the agency's infrastructure can become part of the negotiation over humanitarian access. In that setting, platform jurisdiction becomes a neutrality question whether or not any data has actually been requested. Sections 2 to 6 build the risk framework; Section 7 develops this neutrality argument, which is the paper's centre of gravity.

Much of the sector's visible debate over the past two years has centred on whether it is safe to paste beneficiary data into a chatbot. The larger question is whether the governance of entire operational environments matches the jurisdictional exposure they carry.

Method and scope

This paper reviews public provider statements and policies, primary legal authorities, policy documents, humanitarian doctrine, and selected contemporary reporting. It does not include interviews, confidential contracts, provider audits, or a systematic survey of agency infrastructure. Claims about sector practice therefore describe the visible public conversation and documented cases rather than measured sector behaviour, and are flagged as such where they occur. Section 10 sets out the resulting limitations in full.

2. Position in the literature

This paper sits at the intersection of three bodies of work that rarely meet.

Humanitarian data-governance scholarship has documented harm, experimentation, protection, and responsible processing in crisis settings. Sandvik, Jacobsen and McDonald's taxonomy of humanitarian experimentation established the field's central caution about testing technologies on crisis-affected populations [5]. Jacobsen and Fast's work on humanitarian technology governance examined how digital systems blur access, control, and care [6]. Sandvik's 2026 Global Policy commentary extended the critique to AI-assisted aid through the concepts of embedded experimentality, maldata, and data extrapolation [7]. The ICRC's Handbook on Data Protection in Humanitarian Action remains a principal doctrinal reference, and its third edition addresses government access to humanitarian data through public cloud services, the mechanism at the heart of this paper's legal analysis [8].

Digital sovereignty scholarship, anchored by Pohle and Thiel, frames sovereignty as a political concept concerned with state authority and governance over digital infrastructures [9]. That literature has focused mainly on states and European policy debates. Humanitarian actors present a distinct case: they do not seek sovereignty in the state sense, but sufficient operational autonomy, confidentiality, and independence to maintain principled access. The application of the sovereignty lens to principled humanitarian actors remains thin.

Legal scholarship on extraterritorial data access, principally analyses of the CLOUD Act and FISA Section 702 [10][11][12], explains the mechanisms but rarely connects them to humanitarian access negotiations.

The contribution of this paper is to join these three conversations through the concept of the **perceived neutrality of operational infrastructure**: to show that jurisdictional dependency is not only a data-protection or procurement question but an access question, and that the mechanism through which it operates is host-authority perception rather than confirmed data access.

3. A three-risk framework

Jurisdictional dependency generates three distinct but compounding sovereignty risks. Table 1 summarises the framework.

Table 1. Three sovereignty risks of humanitarian jurisdictional dependency

Risk	Mechanism	Demonstrated precedent	Humanitarian consequence	Time horizon
Commercial	Provider discretion over product continuity, licensing terms, and pricing, bounded only by contract and commercial strategy	Microsoft's retirement of Skype (May 2025) after documented historical humanitarian field use [13][14]	Tools suited to low-bandwidth, high-risk contexts can be withdrawn by product decision; dependency migrates rather than shrinks	Continuous
Legal	Stored-data production through valid legal process where a provider has possession, custody, or control of data; separate foreign-intelligence collection through authorities such as FISA Section 702	CLOUD Act cross-border provision, 18 U.S.C. § 2713 [10]; Section 702 certifications and compelled provider assistance [11][12]	Some content, records, metadata, or communications may be reachable, depending on legal authority, data type, provider control, encryption and key custody, privileges and immunities, notice rules, and ability to challenge	CLOUD Act cross-border provision since 2018; Section 702 subject to reauthorisation, lapse, and continuation rules
Political	Export-control and sanctions authority used to restrict access to capabilities, with restoration conditional on state review	June 2026 suspension and conditional restoration of Anthropic frontier deployments [1][3][4]	Operational AI capability can be altered quickly by state authority; restoration can be conditioned on government-facing review	Demonstrated June 2026

The risks compound as the framework descends: commercial risk is often more visible and contract-manageable, legal risk is structural, and political risk has now been demonstrated at the frontier-AI layer. Sections 4 to 6 develop each; Section 7 shows their convergence on neutrality; Section 8 addresses objections.

4. Risk one: commercial sovereignty

A provider you depend on can decide your use case no longer serves its interests.

On 5 May 2025, Microsoft retired Skype, having announced the retirement in February with a general migration path toward Teams [13]. Skype is best read as an illustration of the pattern rather than proof of sector-wide 2025 reliance: a commercially rational end-of-life decision can remove a tool whose technical characteristics once made it useful in low-bandwidth humanitarian settings. UNHCR documented low-bandwidth Skype deployments for staff in hardship locations as early as 2010 [14]. The conditions that made such a tool useful, conflict zones, remote sites, and underfunded country offices working on constrained connections, have not gone away. There was public notice and a general migration path, but no humanitarian-specific transition plan, and the author has found no humanitarian-specific assessment of the retirement's operational effect. No systematic data on current sector usage of specific collaboration tools was identified for this paper.

In practice, dependency often migrates to other commercial collaboration tools - Teams, WhatsApp, Zoom, Signal, Google Workspace, or local equivalents - but this paper identified no systematic dataset on current sector usage.

The risk applies across the sector's SaaS estate. Microsoft can change its humanitarian licensing terms. Google can alter what Workspace includes at the nonprofit tier. Palantir can reprice. Any continuity these providers owe is contractual, limited, and defined by product terms, service-level commitments, and

commercial strategy, not by humanitarian necessity. The sector treats these platforms as infrastructure. The providers treat them as products. That mismatch is the risk.

5. Risk two: legal sovereignty

This risk is quieter than a product retirement or an export control. Its principal instrument has been in force since 2018, and much of the sector's public discussion appears not to have engaged with it.

5.1 Stored-data production: the CLOUD Act

The US CLOUD Act requires covered US-jurisdiction electronic communication and remote-computing service providers, through valid legal process, to preserve, back up, or disclose customer communications and records within their possession, custody, or control, regardless of whether the data is stored inside or outside the United States [10]. SharePoint files on a Microsoft server in Frankfurt remain within reach of US legal process. So do Outlook mailboxes, OneDrive documents, Teams records, Google Drive files, and Gmail, to the extent they sit within a covered provider's possession, custody, or control.

The CLOUD Act does not make exposure uniform. It creates a legal pathway to data within a provider's possession, custody, or control; what can actually be produced depends on service architecture, customer-held keys, encryption model, tenant configuration, logs, metadata, contract terms, and the form of legal process. Customer-held keys and client-side encryption change what can be produced in content form, though metadata, account records, and access logs usually remain reachable even where content is not. This paper identified no public evidence that such mapping is routinely performed across the sector.

Nor do the providers grant direct, unfettered access. Requests must go through legal process, and Microsoft and Google both state that they review, narrow, reject, or challenge requests where they can, and notify customers where legally permitted [15]. In some cases, however, notice to the customer can be delayed or prohibited by court order [16].

EU data residency complicates rather than resolves the position. GDPR Article 48 provides that a third-country judgment or decision requiring disclosure may only be recognised or enforceable if based on an international agreement such as a mutual legal-assistance treaty, and European Data Protection Board guidance emphasises that a foreign request does not by itself create a lawful basis or transfer ground [17]. The result is a genuine conflict-of-laws problem: a provider may face US disclosure obligations while also needing a lawful basis and transfer mechanism under EU law. The practical risk for humanitarian agencies is therefore not simple US access, but uncertain, contested, and sometimes secret legal process across overlapping regimes. For a host authority assessing whether to trust an agency's systems, that uncertainty is itself part of the problem.

5.2 Foreign-intelligence collection: FISA Section 702

Section 702 should be treated separately from the CLOUD Act. It is not a general stored-data production mechanism for a SharePoint folder or a mailbox. It authorises targeted foreign-intelligence collection against non-US persons reasonably believed to be outside the United States, with compelled assistance from electronic communication service providers, under certifications and targeting and minimisation procedures approved annually by the Foreign Intelligence Surveillance Court rather than individual warrants [11]. ODNI reported an estimated 349,823 Section 702 targets in calendar year 2025 [12].

Its relevance to humanitarian operations is narrower than the CLOUD Act's but still material: humanitarian communications may be acquired incidentally where humanitarian actors communicate

with, or through selectors tasked against, lawful foreign-intelligence targets, and providers may be obliged to assist under a distinct national-security regime. As of July 2026, the authority's status is unsettled. The statute lapsed on 12 June 2026 after Congress declined a further extension, but collection reportedly continues under existing FISC-approved certifications into approximately March 2027, while the legal position of providers assisting under lapsed authority in new cases is disputed [11][18]. This paper describes that position as reported and provisional, not as a stable baseline.

Chronology note. The final Section 702 extension and the Anthropic export-control directive both fell on 12 June 2026. This paper does not suggest a causal link. The juxtaposition matters only because two distinct forms of US state capacity, foreign-intelligence authority and export-control authority over AI capability, were visible on the same day.

5.3 Who the exposure applies to

Exposure also differs by organisational legal status. UN entities, the ICRC, international organisations, INGOs, national NGOs, and local partners do not face the same legal position. Privileges and immunities, headquarters agreements, inviolability of archives, contractual architecture, and whether a provider is acting for an international organisation can materially alter the analysis; the literature on the ICRC's privileges-and-immunities agreement with Luxembourg discusses precisely how inviolability concepts can extend to data, cyber infrastructure, and data centres used by a protected organisation [19]. This paper therefore treats jurisdictional exposure as a risk factor requiring organisation-specific review, not a single conclusion applicable equally to all humanitarian actors.

5.4 The scope of the sector's current conversation

The visible sector conversation on data governance has focused heavily on AI chatbots: do not paste sensitive data into a chatbot, check the processing agreement, be careful with prompts. That concern is real, but it covers a fraction of the exposure. The question is not whether someone pastes data into a chatbot. It is which data, in which services, under which provider's possession, custody, or control, can be reached through which legal authority, and under which safeguards. For agencies running Microsoft 365 or Google Workspace, significant portions of communications, collaboration records, account metadata, and operational data may sit within legal or intelligence-access pathways whose applicability depends on provider, service, configuration, and data type. The AI layer being added on top does not create the jurisdictional exposure. It inherits it.

6. Risk three: political sovereignty

What happens when a state uses the override capacity it already holds. June 2026 demonstrated it twice.

First, the restriction. Reporting linked the directive to a claimed jailbreak [20]. Anthropic disputed the characterisation, describing the demonstrated technique as narrow and non-universal, surfacing previously known minor vulnerabilities that other publicly available models could also find [1][3]. The directive did not order a global shutdown. It imposed a nationality condition, and Anthropic said it had no reliable way to enforce that condition at deployment scale, so it suspended access for everyone. That detail sharpens the structural point: a restriction aimed at one category of user produced a global outage, because the infrastructure did not come with a nationality filter.

Second, the restoration. Access resumed only after government review of the company's safeguards, new information-sharing arrangements, and pre-release government testing commitments for future models, with Mythos 5 restored narrowly to government-approved organisations [3][4]. The episode proves that the mechanism exists: state authority can alter access to frontier capabilities quickly, and

restoration can be conditioned on government-facing review. Whether and how the mechanism would be repeated across other providers, platforms, or product types depends on the legal authority invoked, provider architecture, product type, and political context. One demonstration is not a general regime. It is, however, a precedent that did not exist before June 2026.

No public evidence was identified that the less-than-three-week Fable 5/Mythos 5 suspension materially disrupted humanitarian operations. That absence does not dissolve the risk; it locates it. The outage was survivable because frontier AI is not yet deeply embedded in most humanitarian workflows. The warning is prospective: as AI assistants become part of search, triage, drafting, analysis, translation, case-management support, and coordination workflows, the operational cost of a similar restriction will rise.

Sanctions create a related but distinct access-restriction risk. They may affect whether a provider can serve particular geographies, entities, individuals, payment channels, or counterparties. This paper does not analyse sanctions doctrine in detail, but treats sanctions as part of the wider class of state powers that can affect platform availability and continuity.

The dependency question extends beyond standalone models, and here the risk is prospective rather than demonstrated. AI assistants are being embedded into the productivity suites the sector already runs on: Microsoft 365 Copilot across Word, Outlook, Teams, and SharePoint, and Gemini across Google Workspace applications [21]. Separately, humanitarian analytics platforms illustrate a different dependency type: WFP has described its DOTS platform as an enterprise data platform powered by Palantir Foundry, while stating that Palantir does not have access to beneficiary information, which is held in WFP's own systems [22]. Today, the AI layer is still often administratively separable from the underlying platform; Copilot is an assignable add-on licence, and restricting it would not shut Teams down [21]. The risk is prospective: as workflows, automations, search, summarisation, triage, and decision support are rebuilt around AI features, removal or restriction of the AI layer may degrade operational capacity even where the underlying email or file system remains available. The sector is deepening dependency on AI-integrated tools without asking what happens if the AI layer becomes a point of state restriction.

Further research: connectivity sovereignty. Connectivity raises a related but separate sovereignty problem that this paper does not treat in detail. Starlink's growing role in disaster and conflict connectivity, combined with its US jurisdiction and its ownership structure following SpaceX's February 2026 acquisition of xAI [23], creates a further dependency layer worth examining. The Ukraine episodes reported since 2023 indicate that Starlink availability and permitted use in conflict settings can turn on corporate policy and owner-level decisions, including geofencing and refusal to enable service in particular areas; the exact facts of individual episodes remain contested in public reporting [24]. The connectivity layer deserves its own properly sourced analysis.

7. Where the risks converge: perceived neutrality

Any one of these risks is manageable alone. A tool is retired, another is found. A law creates exposure, contracts and configurations are renegotiated. An export control interrupts a model, an earlier version bridges the suspension. Agencies have absorbed each individually for years. The problem is their combination, and the place they combine is not the server room. It is the negotiation over access.

Consider an agency operating where the United States has strategic, military, intelligence, sanctions, donor, or alliance interests. Its operational environment runs substantially on US-jurisdiction platforms. Under the CLOUD Act, valid legal process can compel production of data within those providers' possession, custody, or control. Under Section 702, humanitarian communications may be acquired

incidentally where humanitarian actors communicate with, or through selectors tasked against, lawful foreign-intelligence targets. Under the export-control regime demonstrated in June, capabilities can be restricted and their restoration conditioned. None of this establishes that any humanitarian data has been sought. It establishes that lawful pathways exist.

Now take the host authority's perspective. A host authority in such a crisis asks the agency a straightforward question. You want to register our population for assistance. You will collect names, ID numbers, phone numbers, locations. That data enters your systems, which sync to SharePoint, back up to OneDrive, and coordinate through Teams, all supplied by a company subject to the legal process of a state with interests in this crisis, and in some cases without immediate notice to you. Why should we permit it?

Whether the foreign state is actually seeking the data is beside the point. The host authority has a reasonable basis for concern: the CLOUD Act is public law, and any government with a competent legal adviser can read it. The concern may be reasonable even where the authority's preferred alternative, local hosting, local database control, or direct state access, would create greater protection risks for affected people, a point Section 8 returns to. When the authority acts on its concern, whether by restricting access, demanding local systems, or blocking registration, the agency's infrastructure has become an obstacle to the assistance it was supposed to enable.

The mechanism is not hypothetical, though the closest precedent must be framed precisely. In 2019, WFP and the Houthi authorities in Yemen deadlocked over biometric registration, with the de facto authorities refusing arrangements that placed beneficiary data under WFP's control; WFP partially suspended assistance in Sanaa for around two months before agreement was reached [25]. The Yemen case is not evidence of US-jurisdiction platform risk. It is evidence of the broader access mechanism: disputes over who controls beneficiary-data systems can interrupt humanitarian assistance. This paper's argument is that jurisdictional dependency gives host authorities a new and legible reason to open exactly that kind of dispute. It is also the closest public analogy this paper identifies, not a direct precedent for the platform-jurisdiction claim.

The sector has debated one visible instance of this problem for years: whether WFP's partnership with Palantir compromises the neutrality of humanitarian data systems [22]. That debate matters, but its focus on one controversial vendor obscured the structural condition. Microsoft, Google, AWS, Oracle, and Palantir differ in services, data-access models, contracts, sub-processors, deployment options, key-management offerings, and transparency practices. At the jurisdictional layer, the same class of question arises for all of them, even though the answer may differ substantially by provider, service, contract, architecture, encryption model, and support-access arrangement. The difference between Palantir and Microsoft is that Palantir is controversial and visible while Microsoft is everywhere and unexamined. Nobody questions running a humanitarian response on Outlook. The question deserves asking in both cases, even where the answers differ.

The sector applies conflict-of-interest standards to staff, donors, and implementing partners. It appears rarely to have applied an equivalent standard to infrastructure. The ICRC is the closest available comparator, not because it offers a simple technical model, but because it has treated data, cyber infrastructure, and data-centre protection as questions of inviolability, privileges and immunities, confidentiality, and neutral humanitarian action. Its agreement with Luxembourg, signed 1 June 2022 and in force since 5 September 2023 alongside the establishment of the ICRC's Delegation for Cyberspace and later Global Cyber Hub, contains specific protections for ICRC data, cyber infrastructure, and data centres used by the organisation, including inviolability of data-centre premises and limits on interference "by cyber means" [19]. That approach does not eliminate security risk: the ICRC's 2022

breach exposed records of more than 515,000 people [26]. Sovereignty and security are different problems. Sovereignty architecture changes who may lawfully compel or interfere with data. Security determines who can steal it. Self-controlled or legally protected infrastructure carries the security burden with it.

The convergence point is this. Neutrality as a humanitarian principle and a host authority's perception of infrastructure alignment are not the same thing, and this paper's claim concerns the second. If infrastructure is the reason a host authority denies access, the agency's actual adherence to principle does not rescue the operation. The infrastructure has created a perceived-alignment problem in the forum that decides whether the agency operates.

8. Objections and trade-offs

Seven objections deserve direct answers, because each contains something true.

All jurisdictions create coercion risk. Correct. Moving away from US providers does not remove sovereignty risk; it changes which state, court system, intelligence framework, and corporate structure matter. Chinese, Russian, and EU jurisdictional reach over their respective platforms would each support a parallel analysis. The argument here is not that US jurisdiction is uniquely dangerous. It is that many of the sector's most common infrastructure dependencies run through providers exposed to the same jurisdiction, and that concentration itself is the governance failure.

Local hosting may be worse. Often true, and this paper says so. In conflict settings, host-state coercion or armed-group pressure on locally held beneficiary data may create greater risks for affected people than foreign legal process does. The host authority's concern can be reasonable while its preferred remedy is dangerous. That is an argument for deliberate, sensitivity-tiered architecture, not for localisation by default.

Provider safeguards matter. They do. Encryption, customer-held keys, contractual notice commitments, legal challenge practices, audit logs, transparency reporting, and data minimisation can materially reduce exposure, and the providers' own transparency reports document contested and narrowed requests [15]. The framework in this paper treats exposure as configurable, not binary. This paper identified no public evidence that such configuration analysis is routinely performed across the sector.

A non-US provider is not automatically safer. Correct, and some EU providers carry US parent entanglements or US sub-processors that reintroduce the exposure by another route. The right unit of analysis is the specific service's jurisdictional chain, not the brand's headquarters.

No direct public platform-jurisdiction precedent has been identified. A critic may fairly ask why, if US-jurisdiction platform exposure has been legible since the CLOUD Act's 2018 enactment, there is no documented public case of a host authority denying humanitarian access specifically because an agency used Microsoft, Google, AWS, Oracle, or another US-jurisdiction provider. That absence limits the claim. This paper does not argue that such disputes are already common or publicly documented. It argues that the legal basis for such concern is public, that disputes over beneficiary-data control have already interrupted assistance, and that platform jurisdiction supplies a plausible additional ground for future access disputes as host authorities become more legally and technically sophisticated. It is also possible that concerns already arise privately, or are expressed as demands for local control, data localisation, or direct database access rather than in the vocabulary of the CLOUD Act. The absence of a public precedent is therefore a limitation, not a refutation.

Architecture cannot solve bad faith. A jurisdictional review, encryption model, or host-authority explanation pack can answer good-faith concerns about legal exposure. It cannot by itself solve a pretextual access dispute, an armed actor's demand for direct database control, or a host authority's political decision to use data governance as leverage. In those cases, better architecture still matters, but differently: it reduces actual exposure, clarifies what the agency can honestly say, supports principled refusal of dangerous localisation or state-access demands, and gives donors and headquarters a tested basis for escalation. It is not a diplomatic shield.

The argument is about perception and access, not proven misuse. Yes, and explicitly so. This paper does not claim that any state has used these authorities against humanitarian data. Its claim is narrower and harder to dismiss: the authorities exist, they are public, host authorities can read them, and perception operates on what is legible. That is sufficient for the neutrality consequence, and it is the paper's strongest defence as well as its most careful claim.

9. Implications for practice and policy

None of this argues for abandoning US technology. The tools work, the alternatives are limited, and a rushed migration would create more risk than it removes. Six moves follow, ordered from diagnosis to resilience.

9.1 Run a jurisdictional sovereignty review. Reading the CLOUD Act is the start, not the review. A minimum viable sovereignty review, per platform, should record: provider and contracting legal entity; parent-company jurisdiction; governing law and dispute forum; data residency and backup locations; sub-processors and support locations; whether the provider has possession, custody, or control; whether the provider can access plaintext; customer-managed key, hold-your-own-key, or client-side encryption options; metadata, audit log, telemetry, and support-record exposure; the provider's government-request policy; the customer-notice policy; the provider's record of challenging demands; a GDPR Article 48 and transfer-conflict analysis; privileges and immunities where applicable; sanctions and export-control exposure; AI feature dependencies and disablement options; data export format and exit cost; offline or manual fallback; a tested outage scenario; and host-authority perception risk for the operating contexts concerned. The output is an evidence pack, including a host-authority explanation pack for sensitive contexts, not a policy statement. As Section 8 notes, this pack is not a diplomatic shield: it helps answer good-faith concerns and support principled escalation, but it cannot by itself solve pretextual or coercive demands for control. A companion Humanitarian Platform and AI Sovereignty Review Checklist, translating this review into a twenty-question operational screening tool, is in preparation and will be published separately.

9.2 Diversify deliberately. The sector diversifies funding so that no programme depends on a single donor; the same logic applies to infrastructure. EU-based platforms reduce US-jurisdiction exposure and operate under GDPR, while EU AI Act obligations phase in by category: prohibited-practice and AI-literacy rules began in 2025, general-purpose AI obligations began in 2025, general applicability remains centred on 2 August 2026 with exceptions, and high-risk-system obligations are now subject to later 2027-2028 application dates following the AI simplification package [27]. But sovereignty is a trade-off, not a destination, and a non-US provider is not automatically safer. The right question is not US or non-US, but which jurisdiction, provider, key arrangement, and fallback plan match the sensitivity of the operation.

9.3 Separate the sensitive. A headquarters budget spreadsheet is not a beneficiary registration database in an active conflict. The highest-risk data should sit on infrastructure the organisation controls or that enjoys legal protection appropriate to its status, in a jurisdiction assessed rather than assumed,

with access governance that does not depend on foreign law. Tier the architecture to the sensitivity, and accept that the highest tier carries its own security burden.

9.4 Build portable AI capability, not product dependency. Agencies adopting AI should build workflows around methods that transfer across models and providers: verification discipline, model comparison, adversarial review, prompt and workflow documentation, source checking, and tested fallback workflows across providers. The June suspension illustrates the point: workflows anchored to a single product inherited that product's political exposure, while method-anchored workflows could switch tools. Donors funding AI training should ask whether a programme teaches a product or a methodology, because product training builds a dependency while methodology training builds a capability that survives a restriction. Disclosure: the author runs a training programme built on this premise, and readers should weigh that interest when evaluating this recommendation.

9.5 Require dependency contingency planning, including degradation. Donors funding digital transformation should require dependency contingency plans as they require business continuity planning, and the scenarios should cover degradation as well as shutdown: an AI feature removed while the core product remains; one country office blocked by sanctions or export controls; a provider account locked pending compliance review; a legal hold or non-disclosure order constraining normal data-subject response; connectivity interruption; payment or licensing failure; a sub-processor or support location becoming politically unacceptable; and a host authority demanding a local copy or local administrative access. Three questions for every funded programme: which platforms and who has jurisdiction over them; what is the plan if access is restricted or degraded; has the fallback been tested.

9.6 Have the honest conversation. The sector needs to discuss openly what it means to run principled humanitarian operations on infrastructure subject to the legal authority of states with interests in the crises being responded to. That conversation has so far been had about one vendor. It needs to be had about the platforms every agency uses daily, with the nuance this paper has tried to model: exposure is configurable, organisational status matters, alternatives carry their own risks, and the neutrality consequence operates through perception.

10. Limitations

This is a conceptual and documentary analysis. It uses no primary data, no systematic survey of agency infrastructure, and no interviews; claims about sector tool usage are illustrative and flagged as such. Its jurisdictional focus is the United States, because that is where the dominant providers are incorporated and where the June 2026 episode occurred; equivalent analyses of other jurisdictions would extend the framework rather than contradict it. Organisational legal status materially changes the analysis, and this paper's framework is a starting point for organisation-specific review rather than a substitute for it. Events are ongoing: the Section 702 lapse-and-continuation position may change, the June 2026 precedent may be revised by subsequent policy, and claims about provider policies reflect the providers' public statements rather than independent audit.

11. Conclusion: the pattern

The sector treated beneficiary data as an asset until a breach proved it was a liability. It treated its communications tools as permanent until a commercial decision proved they were not. It treated AI access as a given until an export control proved it is a dependency, and the conditional restoration showed that the dependency can be managed by a state. And it has run its most sensitive operations on foreign-jurisdiction platforms for years while rarely examining the legal framework that governs what can be done with the data sitting on them.

Each time, the sector absorbed the shock, adjusted, and moved on without addressing the structural condition underneath. The condition is constant: common operational infrastructure supplied by providers the sector does not govern, within jurisdictions it cannot influence, under law it has rarely examined. For many common components of the operational stack, meaningful override capacity often sits with a short list of providers exposed to the same jurisdiction.

That is not a technology problem. It is a governance gap, and it is the sector's to close.

Disclosure and AI use statement

The author is the founder of MarketImpact Digital Solutions Ltd, a practitioner-led consultancy working on humanitarian crisis analysis, AI adoption, and digital systems, and runs the AidGPT responsible AI training programme, which trains humanitarian professionals in verification discipline and structured AI workflows across models and providers. This interest is directly relevant to the recommendation in Section 9.4 and is disclosed there in text.

Web research on the export-control episode and the US legal framework for cross-border data access, source cross-referencing, and initial structuring of the argument were AI-assisted using Claude Opus 4.6. The three-risk framework, the pivot from AI to operational infrastructure, the application of the CLOUD Act and FISA Section 702 analysis to humanitarian operations, the perceived-neutrality argument, and all editorial judgement are the author's own. The analysis underwent four internal adversarial review passes, including two independent structured reviews of this working paper version and a full revision after the export controls were lifted. This is a governance analysis, not legal advice. Any organisation's actual exposure depends on its legal status, providers, contracts, configuration, and key custody.

References

[1] Anthropic, "Statement on the US government directive to suspend access to Fable 5 and Mythos 5", 12 June 2026. https://www.anthropic.com/news/fable-mythos-access (accessed July 2026).

[2] Axios, "Trump admin blocks foreign access to Anthropic's most advanced AI", 12 June 2026; Ravie Lakshmanan, "U.S. Orders Anthropic to Suspend Fable 5 and Mythos 5 Access for Foreign Nationals", The Hacker News, 13 June 2026. Cited for contemporary reporting on the suspension and compliance decision.

[3] Anthropic, "Redeploying Claude Fable 5", 30 June 2026. https://www.anthropic.com/news/redeploying-fable-5 (accessed July 2026).

[4] Lutnick, H., Commerce Secretary, post on X, 30 June 2026; CIO, "US reverses export restrictions on Anthropic's Fable 5, Mythos 5 AI models", 1 July 2026; Reuters, "US removes curbs on Anthropic's latest Fable and Mythos AI models", by Chris Thomas, Karen Freifeld and Deepa Seetharaman, 30 June 2026, updated 1 July 2026.

[5] Sandvik, K.B., Jacobsen, K.L. and McDonald, S.M. (2017), "Do no harm: A taxonomy of the challenges of humanitarian experimentation", International Review of the Red Cross, 99(904), 319-344.

[6] Jacobsen, K.L. and Fast, L. (2019), "Rethinking access: how humanitarian technology governance blurs control and care", Disasters, 43(S2), S151-S168.

[7] Sandvik, K.B. (2026), "AI in Aid: Experimentality, Maldata, and Data Extrapolation", Global Policy Journal, commentary, 16 February 2026.

[8] Marelli, M. (ed.) (2024), Handbook on Data Protection in Humanitarian Action, 3rd edition, Geneva/ Cambridge: ICRC and Cambridge University Press.

[9] Pohle, J. and Thiel, T. (2020), "Digital sovereignty", Internet Policy Review, 9(4).

[10] Clarifying Lawful Overseas Use of Data (CLOUD) Act, PL 115-141, Division V (2018), codified at 18 U.S.C. § 2713; Congressional Research Service (2018), "Cross-Border Data Sharing Under the CLOUD Act", R45173.

[11] Brennan Center for Justice, "Section 702 of the Foreign Intelligence Surveillance Act, Explained" and "Section 702 of FISA: 2026 Resource Page"; Congressional Research Service (2025), "FISA Section 702 and the 2024 Reforming Intelligence and Securing America Act"; Reforming Intelligence and Securing America Act (RISAA), PL 118-49 (2024).

[12] Office of the Director of National Intelligence (2026), Annual Statistical Transparency Report Regarding the Intelligence Community's Use of National Security Surveillance Authorities, Calendar Year 2025, April 2026, Figure 4.

[13] Microsoft (2025), "The next chapter: Moving from Skype to Microsoft Teams", Microsoft 365 blog, 28 February 2025; Skype service retired 5 May 2025.

[14] Devex (2010), reporting on UNHCR's low-bandwidth Skype deployment for staff in hardship locations. Cited as historical evidence of field suitability, not as proof of 2025 reliance.

[15] Microsoft, Government Requests for Customer Data report and Law Enforcement Request policy (microsoft.com/corporate-responsibility/law-enforcement-requests-report); Google, Transparency Report, Requests for User Information (transparencyreport.google.com). Both accessed July 2026.

[16] Stored Communications Act, 18 U.S.C. § 2705, delayed notice and preclusion-of-notice orders.

[17] Regulation (EU) 2016/679 (GDPR), Article 48; European Data Protection Board, Guidelines 02/2024 on Article 48 GDPR, final version, 05 June 2025.

[18] Reuters, "A key surveillance law is expiring. What does that mean for U.S. intelligence?", 9 June 2026; Reuters, "US Congress rejects short-term FISA extension as expiration looms", 11 June 2026; Brennan Center for Justice, "Section 702 Surveillance Will Continue Until March 2027 Even if the Statute Lapses", 9 June 2026; CBS News, "A key spy authority, Section 702, expired due to inaction in Congress. What does that mean?", 13 June 2026.

[19] Raab-Gray, A. and Marelli, M. (2025), "Inviolability in the digital era: The ICRC's Agreement on Privileges and Immunities with Luxembourg", International Review of the Red Cross, 107(928), 401-428. doi:10.1017/S1816383125000190. https://international-review.icrc.org/articles/inviolability-in-the-digital-era-928 (accessed July 2026).

[20] Axios, "Trump admin blocks foreign access to Anthropic's most advanced AI", 12 June 2026; Axios, "How the world's top AI models were revived", 3 July 2026; Reuters, "US removes curbs on Anthropic's latest Fable and Mythos AI models", 30 June 2026, updated 1 July 2026; Anthropic, "Redeploying Claude Fable 5", 30 June 2026. Cited for reporting on the claimed trigger and Anthropic's rebuttal and restoration account.

[21] Microsoft, "Microsoft 365 Copilot service description", Microsoft Learn (learn.microsoft.com); Google, "Gemini for Google Workspace" documentation (workspace.google.com). Both accessed July 2026.

[22] WFP, public statements on the DOTS platform and the Palantir Foundry partnership, 2019 onward, including WFP Executive Board documentation. WFP states that Palantir does not have access to

beneficiary information, which is held in WFP's own systems; this reflects WFP's stated position rather than an independent audit.

[23] Reuters and CNBC, reporting on SpaceX's all-stock acquisition of xAI, February 2026; xAI acquired X in March 2025.

[24] Reuters, "Musk says he refused Kyiv request for Starlink use in attack on Russia", September 2023; Reuters, investigation into Starlink service decisions in Ukraine, July 2025; US Department of State, announcement of cooperation with Starlink on disaster response connectivity, June 2026. The facts of individual Ukraine episodes remain contested in public reporting.

[25] WFP (2019), "World Food Programme begins partial suspension of aid in Yemen", statement, 20 June 2019; Reuters and Associated Press, reporting on the biometric-registration dispute and its resolution, June-August 2019.

[26] ICRC (2022), statements on the January 2022 cyber attack against servers hosting Restoring Family Links data, affecting more than 515,000 people.

[27] European Commission, "AI Act", implementation timeline, noting entry into force on 1 August 2024 and general applicability from 2 August 2026 with exceptions; Council of the European Union, "Artificial Intelligence: Council gives final green light to simplify and streamline rules", 29 June 2026, noting 2 December 2027 and 2 August 2028 application dates for high-risk AI systems.