O5 PARTNERS

WHITE PAPER · 2026

A WHITE PAPER ON IT AND AI SOVEREIGNTY

The Sovereignty Imperative

Owning IT and AI Infrastructure in 2026 and Beyond

O5 Partners s.r.o.
o5partners.com
Published 2026

Powerful technology.
Private by design.
Yours to control.

CONTENTS

What is inside

EXECUTIVE SUMMARY

A clear-eyed look at where the technology businesses depend on actually sits — and what to do about it.

The IT and AI tools most European businesses depend on sit largely outside their control — in jurisdictions and on infrastructure they do not own. That arrangement worked while it was invisible. In 2026, it is no longer invisible.

The decisions European businesses made about IT and AI over the past decade were reasonable ones. The tools delivered real productivity gains, and the alternatives at the time were limited. What has shifted, in a way that few decision-makers anticipated, is the surrounding landscape. Regulation has tightened substantially. Geopolitics has rearranged itself. The commercial models behind cloud AI have begun to reveal their long-term shape. And the technical alternatives — sovereign infrastructure that businesses can genuinely own and operate — have matured to a point where they are practically achievable rather than aspirational.

This paper lays out the changed landscape, the regulatory pressure converging on European organisations, and the practical answer to a question that more business leaders are quietly asking: *what does it mean to own the technology we depend on, rather than rent it?*

SIX FINDINGS AT A GLANCE

1. The dependency is structural, not incidental. Approximately **70% of the European cloud market** is controlled by three US hyperscalers, and around **80% of EU corporate software and cloud spending flows to US vendors**. Collectively, European organisations spend an estimated **€265 billion per year**

on US software and services — close to 1.5% of EU GDP *(European Parliament, 2025; Cigref & Asterès, 2025).*

2. "Sovereign cloud" offerings from hyperscalers do not fully resolve the question. All three major US providers launched sovereign cloud initiatives in 2025, but the parent companies remain subject to US legal jurisdiction — including the CLOUD Act — regardless of where data physically sits *(Exoscale, 2026; UpCloud, 2025).*

3. Regulation is no longer optional. NIS2, the EU Data Act, the EU AI Act and intensifying GDPR enforcement now converge on the same operational picture: organisations must know where their data lives, who can access it, and under whose terms. GDPR fines have reached **€7.1 billion cumulatively**, NIS2 carries penalties up to **€10 million or 2% of global turnover**, and the EU AI Act introduces fines up to **€35 million or 7% of turnover** — higher than GDPR *(ASEE, 2026; European Commission, 2025; Legiscope, 2026).*

4. The cost curve of cloud AI penalises adoption. Per-token and per-seat pricing scales costs in direct proportion to value delivered. Lenovo's 2026 TCO analysis finds that self-hosted inference can be **up to 18× cheaper than cloud APIs** over a 3-year period for high-utilisation workloads, with a breakeven point reached in under four months *(Lenovo, 2026).*

5. Sovereign infrastructure has become practically achievable. Modern small and medium AI models, deployed on modest local hardware, now match or exceed cloud-hosted models for the specific tasks most organisations actually need. The trade-off between capability and control — real until 2022 or 2023 — has materially narrowed.

6. The question is universal. European businesses face a version of the same question that emerging-market organisations are asking with greater urgency. **61% of Western European CIOs** now plan to shift toward local or regional providers, with geopolitical concerns cited as a primary driver *(Gartner via The Register, 2025).*

WHAT TO DO NEXT

For decision-makers, three steps follow from this picture. First, conduct a **structured dependency assessment**: which systems run on foreign-controlled

infrastructure, what data passes through them, and what jurisdictional reach applies. Second, map **regulatory exposure**: which obligations apply now, which apply by 2026 and 2027, and which require infrastructure changes rather than process changes. Third, evaluate **sovereign alternatives** that did not exist three years ago: locally-deployed AI, owned infrastructure, hybrid architectures designed for sovereignty by default. The transition does not need to be sudden. But it does need to start somewhere.

I.

INTRODUCTION

The landscape has changed.

The pace of IT and AI adoption across Europe over the past decade has been substantial. Businesses moved quickly to deploy tools that genuinely made them more productive, more capable and more competitive. None of those decisions were mistakes.

What has shifted — quietly, and in parallel with that adoption — is the surrounding landscape. Regulation has tightened in ways that few business leaders fully anticipated when they first signed cloud contracts in 2018 or 2020. Geopolitics has rearranged itself. The commercial models behind cloud and AI have begun to reveal their long-term shape. And what was once a technical choice — where data is processed, on whose servers, under whose laws — has become a strategic and legal question that boards are starting to ask in earnest.

This paper is about that changed landscape. It is not an argument against the cloud, against AI, or against the technology stack most European businesses currently use. It is an argument for clarity: understanding what an organisation actually depends on, what risks that dependency creates today (rather than in some hypothetical future), and what the practical alternatives now look like.

What sovereignty means in this context

The word is used loosely, so it is worth being precise.

Digital sovereignty is the ability of an organisation to control its own data, technology and infrastructure without depending on external entities to grant that control. Data residency — where servers physically sit — is only the surface of it. True sovereignty includes **jurisdiction** (whose laws can compel access), **operational control** (whose decisions affect availability and pricing), **auditability**

(whether the systems can be inspected) and **continuity** (whether access can be removed by a third party).

Two or three years ago, this distinction was mostly of interest to specialists. In 2026, it has become a board-level question — not because anyone made the wrong call, but because the rules and the risks have evolved.

> *The decisions European businesses made about IT and AI were reasonable. The information available now is different.*

How this paper is organised

What follows is structured to be read in any order, but it tells a connected story. Section II maps the dependency stack — what European businesses actually run on, and where the foreign concentration sits. Section III lays out the regulatory environment, with the four major instruments that now converge on operational IT decisions. Section IV addresses the "sovereign cloud" offerings that hyperscalers launched in 2025, and why they do not fully resolve the question. Section V extends the analysis to emerging markets, where the same question lands with sharper edges. Sections VI through VIII move from problem to answer: what sovereignty actually looks like in practice, what the economics are, and how a realistic transition path works. Section IX offers four anonymous case scenarios. Section X concludes.

Every quantitative claim is sourced. References use Harvard style and are listed in full at the end of the paper.

II.

THE DEPENDENCY STACK

What European businesses actually run on.

The conversation about technology sovereignty often gets abstract. The numbers are not. Layer by layer, the IT stack that most European businesses depend on is dominated by a small number of foreign companies — and the economics of that arrangement are quantifiable.

The infrastructure layer

At the foundation sits cloud and data centre infrastructure. Approximately **70% of the European cloud market is held by three US hyperscalers** — Amazon Web Services, Microsoft Azure and Google Cloud. European cloud providers collectively hold around **15%**, with SAP and Deutsche Telekom — Europe's largest players — accounting for just 2% each *(European Parliament, 2025; Synergy Research Group via CNBC, 2026).*

The picture extends to physical infrastructure. The vast majority of Europe's data centre capacity is operated by, or commercially aligned with, US-headquartered companies. Where the data physically sits is — as we will see in Section IV — only part of the question. But the operational reality is that the foundation of European digital infrastructure is, by ownership and governance, predominantly American.

The platform and productivity layer

Above the infrastructure sits the software that everyday work actually runs on. Here the foreign concentration is even more pronounced. Windows holds **73% of desktop operating systems** across Europe. Google Search commands **over**

89% of the web search market. Mobile operating systems — iOS and Android — are effectively entirely US-controlled *(European Parliament, 2025).*

In productivity software, Microsoft 365 and Google Workspace dominate. In customer relationship management, Salesforce. In databases and enterprise applications, Oracle, Microsoft and IBM. In analytics, the same names. SAP is the only prominent European enterprise software vendor at meaningful scale.

The cybersecurity layer

The tools most organisations use to defend themselves are also concentrated. Firewalls, endpoint protection, identity and access management, security information and event management — these markets are dominated by US vendors and a small number of non-EU specialists. European cybersecurity firms specialise mainly in services and niche product areas, rather than competing in the core platform tiers *(European Parliament, 2025).*

The newest layer: AI

The most recent addition to the stack — large language models and the AI tools built on them — is, predictably, even more concentrated. OpenAI, Anthropic and Google are the dominant providers of frontier AI capabilities, with Microsoft positioned as the largest commercial reseller through Copilot integration. Per-token economics mean that costs scale directly with adoption; data sent to these services travels to foreign training and inference infrastructure; and the underlying models themselves are proprietary, with limited auditability.

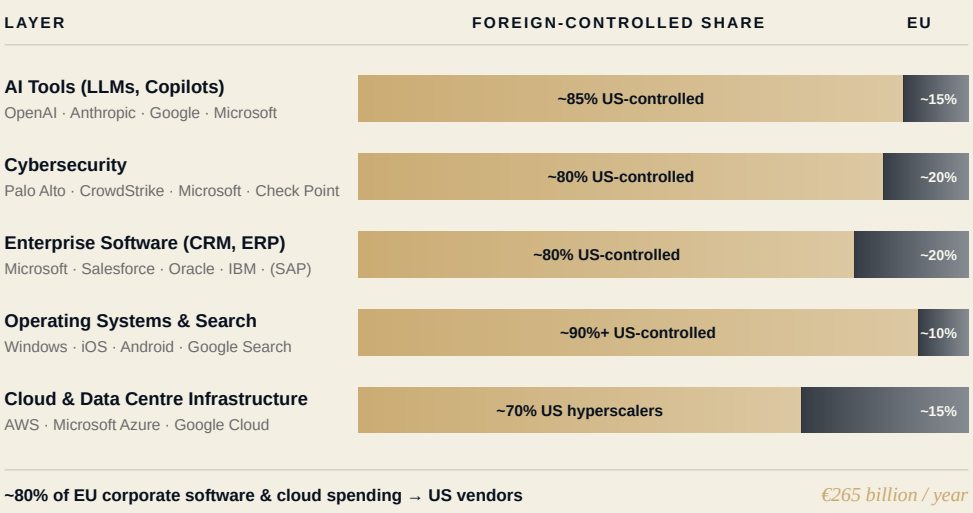
What this looks like financially

A 2025 study by the French business association Cigref and Asterès calculated that **European organisations collectively spend an estimated €265 billion per year on US software and cloud services**. That is approximately 1.5% of EU GDP, flowing each year from European economies to a small number of foreign vendors *(Cigref & Asterès, 2025; Computerworld, 2026).*

At the individual organisational level, the numbers are also striking. Germany's federal government alone spends **€481 million per year on Microsoft licenses** — and that figure does not include the German federal states *(Computerworld, 2026).*

DIAGRAM 1

The European Technology Dependency Stack



Sources: European Parliament (2025); Synergy Research Group via CNBC (2026); Cigref & Asterès (2025).

The point of this picture is not that the technology is bad. Most of these tools work well — that is why they have become dominant. The point is that **the architecture of European business depends, layer by layer, on a small number of companies headquartered outside the EU's legal and political jurisdiction**. As the next sections will show, this matters in 2026 in ways it did not in 2018.

III.

THE REGULATORY ENVIRONMENT

What was once preference is now law.

For most of the past decade, where European businesses stored and processed data was an organisational preference. In 2026, four converging EU instruments — NIS2, the EU Data Act, the EU AI Act and intensifying GDPR enforcement — have turned it into a legal obligation with substantial penalties for non-compliance.

NIS2: cybersecurity becomes board-level

The NIS2 Directive, which had to be transposed into national law by EU Member States by **17 October 2024**, significantly expanded the scope of EU cybersecurity regulation. Where its predecessor covered a relatively narrow set of critical operators, NIS2 covers tens of thousands of additional mid-sized businesses across sectors including healthcare, finance, manufacturing, transport, energy, digital infrastructure, public administration and digital service providers *(European Commission, 2024).*

The obligations are substantive: structured risk management, encryption, multi-factor authentication, supply chain security, incident reporting within tight deadlines, and crucially, **personal accountability for senior management**. Penalties reach **up to €10 million or 2% of global annual turnover** for essential entities, whichever is higher *(Greenberg Traurig, 2025).*

Implementation has been demanding. By mid-2025, **13 of the 27 EU Member States had still not completed NIS2 transposition** into national law, prompting the European Commission to issue formal "reasoned opinions" — legal warnings — to those falling behind *(Skadden, 2025).* In June 2025, ENISA, the EU's cybersecurity agency, issued detailed technical guidance on the specific security

measures regulated entities are expected to have in place — and for many newly regulated organisations, meeting these requirements will require infrastructure investment that hadn't been budgeted for.

The EU Data Act: ending lock-in by law

The EU Data Act, which became applicable on **12 September 2025**, addresses a related but distinct problem: vendor lock-in in cloud and software services. Among other provisions, it now legally requires cloud and software providers to **actively support customer switching** between services, and mandates the **phasing out of switching fees entirely by 2026–2027** *(European Commission, 2025; Gislen Software, 2026).*

That EU regulators felt it necessary to legislate against lock-in directly is itself a significant signal. The Data Act also includes provisions designed to block unlawful third-country government access to data stored in the EU — implicitly acknowledging that the CLOUD Act and similar foreign legal instruments are not theoretical concerns.

The EU AI Act: a phased rollout with penalties higher than GDPR

The EU AI Act, in force since **1 August 2024**, applies through a phased implementation timeline. Prohibited practices (manipulative AI, social scoring, predictive policing) became unlawful from **2 February 2025**. Rules for general-purpose AI models entered into application on **2 August 2025**. The majority of obligations — covering high-risk AI systems in sectors such as healthcare, education, critical infrastructure, law enforcement and employment — apply from **2 August 2026**. Full application to AI embedded in regulated products follows on **2 August 2027** *(European Commission, 2025).*

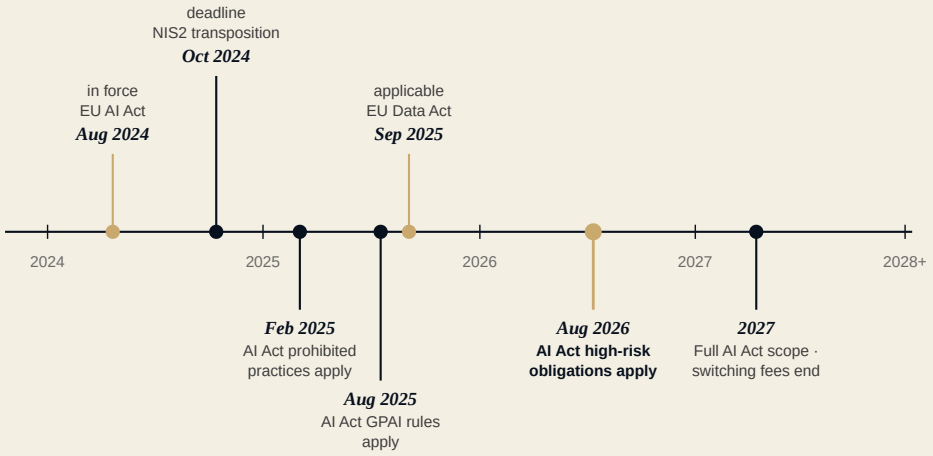
Penalties under the AI Act reach **up to €35 million or 7% of global annual turnover** — materially higher than GDPR. Both frameworks apply concurrently to AI systems that process personal data *(Legiscope, 2026).*

GDPR: enforcement is no longer theoretical

GDPR has been in force since 2018, but the past three years have transformed enforcement. Cumulative GDPR fines reached **€7.1 billion as of January 2026**. Data protection authorities in Austria, France and Italy have issued rulings specifically against US-based tools for violating GDPR through transatlantic data transfers — creating real and immediate compliance exposure for organisations using those tools *(ASEE, 2026).*

DIAGRAM 2

EU Regulatory Timeline: IT and AI Sovereignty



Gold markers indicate landmark dates with significant compliance impact. Sources: European Commission (2024, 2025); Legiscope (2026).

COMPLIANCE EXPOSURE

Maximum penalty thresholds, 2026

GDPR: **€20M or 4% of global turnover**. NIS2: **€10M or 2% of global turnover** for essential entities. EU AI Act: **€35M or 7% of global turnover**. All three can apply concurrently to the same organisation.

For business leaders, the practical implication is unambiguous. The four instruments do not require the same actions, but they converge on the same

operational question: **does your organisation know where its data lives, who can access it, and under whose terms?** If the answer is unclear, the compliance risk is no longer theoretical.

IV.

THE 'SOVEREIGN CLOUD' MIRAGE

Why hyperscaler sovereignty offerings don't fully solve the problem.

In 2025, all three major US hyperscalers launched "sovereign cloud" offerings specifically marketed to European customers concerned about data sovereignty. These are commercially reasonable responses to a real customer demand. They are also, by their structural nature, unable to fully resolve the question they propose to answer.

What was launched

Microsoft launched its Sovereign Cloud in **June 2025**. AWS announced its European Sovereign Cloud for end-2025, with the first region in Germany. Google Cloud introduced its Data Boundary for Europe in **May 2025**. All three offerings emphasise that customer data, encryption keys and operational control remain within the EU — and all three were promoted, with varying language, as solutions to European sovereignty concerns *(Exoscale, 2026).*

By summer 2025, however, it had become broadly understood across European policy and legal analysis that these initiatives **cannot fully guarantee European data sovereignty** *(UpCloud, 2025; Digital Samba, 2025).* The reason is structural rather than commercial.

The CLOUD Act and FISA reach

The Clarifying Lawful Overseas Use of Data Act (the CLOUD Act), enacted by the United States in 2018, requires US-headquartered companies to provide US authorities with access to data under their control — regardless of where that

data is physically stored. FISA Section 702 and Executive Order 12333 provide additional intelligence-collection authorities. These instruments apply to the parent companies of the major hyperscalers, and through them to any subsidiary or operational structure those parent companies control *(Digital Samba, 2025).*

A European subsidiary running on European hardware in European data centres remains, in this legal sense, an extension of its US parent. The sovereignty offering does not change the corporate structure. It changes the marketing and the operational arrangement. The underlying legal reach remains.

> *A European subsidiary running on European hardware in European data centres remains, in this legal sense, an extension of its US parent.*

The GitHub Copilot incident: a case in point

In early 2025, Microsoft restricted access to GitHub Copilot for developers located in regions subject to US export controls and sanctions. The restrictions were not announced in advance and were applied automatically, leaving affected users — including developers at organisations with no connection to the sanctioned activity — unable to access tools they depended on for daily work. The incident drew significant attention across European developer and policy communities, prompting renewed questions about how much operational confidence could be placed in cloud-based professional tools governed by non-EU law *(Digital Samba, 2025).*

The lesson for European business decision-makers is not that Microsoft, AWS or Google are untrustworthy partners. They are large, professional companies operating in good faith. The lesson is that **good faith does not override jurisdiction**. A US-headquartered company, when ordered by a US court or agency, must comply. The sovereign cloud structures do not — and cannot — change this.

What genuine sovereignty requires

A genuine sovereign solution requires ownership of the entire stack — software, infrastructure, operational governance — by an entity not subject to the controlling jurisdiction. Specifically:

The **infrastructure** must be owned and operated by an entity headquartered in the relevant jurisdiction, with no parent-company legal exposure to foreign authorities. The **software** must be auditable — either fully open-source or with source-code access — rather than a black box whose behaviour and data flows must be trusted. The **operational governance** must be under direct customer or jurisdictionally-aligned control, with no remote administrative access by a foreign entity that could be compelled. And the **data** must never leave the controlled environment — not for analytics, not for training, not for "service improvement," not for any reason.

This is a high bar. It is achievable. But it is not what the hyperscaler sovereign offerings provide, and treating those offerings as equivalent — as some marketing materials suggest — produces a false sense of resolution that may not survive a serious legal or operational review.

V.

THE GLOBAL SOUTH CASE

The same question, sharper edges.

The sovereignty conversation that has gained traction in Europe is also being held — often with greater urgency — by organisations and governments far beyond Brussels. Where European businesses are responding to regulatory pressure and emerging risk, organisations in emerging economies are responding to costs that are already biting and an asymmetry that is already visible.

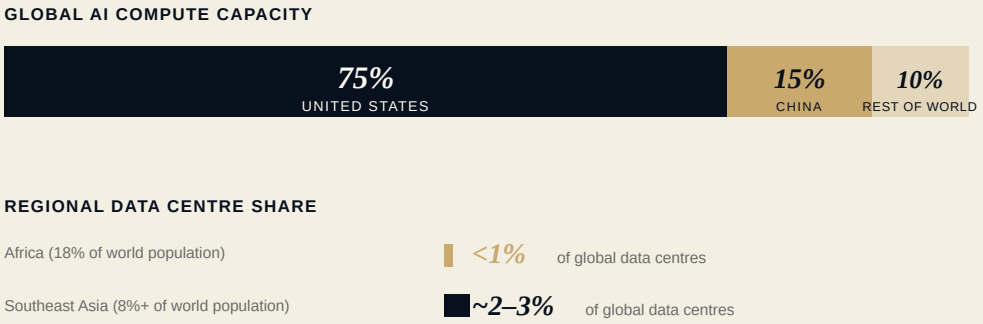
The shape of the global asymmetry

The geography of AI today is striking. According to World Economic Forum reporting, approximately **75% of global AI compute capacity is concentrated in the United States and a further 15% in China**. The remaining 10% is spread across the entire rest of the world *(World Economic Forum, 2026).*

The picture extends to physical infrastructure. The Center for Strategic and International Studies notes that **Africa, home to 18% of the world's population, accounts for less than 1% of global data centre capacity**. Southeast Asia, despite hosting over 8% of the global population across rapidly growing digital economies, accounts for approximately 2–3% of global data centre capacity — a gap that is only beginning to close *(CSIS, 2025).*

DIAGRAM 3

The Compute Asymmetry: Where AI Capacity Sits



Sources: World Economic Forum (2026); Center for Strategic and International Studies (2025).

Why the costs of dependency are higher

The financial burden of cloud and platform dependency falls harder in emerging economies. Cloud and AI services are typically priced in US dollars; for organisations operating in currencies that float against the dollar, every period of depreciation translates into rising IT costs without any improvement in service. Foreign-currency exchange exposure becomes a structural cost of using foreign cloud — a cost that does not exist when infrastructure is owned domestically.

The regulatory environment in many emerging economies is also moving fast. Data localisation laws have been adopted or strengthened across Southeast Asia, Latin America and Africa in recent years, often with explicit sovereignty language and sometimes with shorter implementation timelines than equivalent European measures.

Building AI vs. owning AI

The conversation about AI in emerging economies is often framed as a race for compute and national capability — large data centres, sovereign foundation models, training facilities. This framing matters at the national policy level. But it

can obscure something important for individual organisations: **you do not need to build AI to own it**.

Building AI — training foundation models, running national research programmes, developing competitive sovereign capability — is genuinely the domain of governments and major research institutions. According to UNESCO, **more than twenty Global South countries adopted national AI strategies or digital rights frameworks between 2018 and 2025**, up from just five at the start of that period *(UNESCO via Al Mokdad, 2025).*

Owning AI is something quite different — and more immediately achievable for any organisation, anywhere. It means using AI on infrastructure the organisation controls, with data that stays in the environment, on terms it fully controls. The models themselves can come from many sources: open-weight foundations, partnerships, internal fine-tuning, or commercial licensing. **What matters for sovereignty is not who built the model. It is where it runs, who can reach the data, and who controls the terms.**

The leapfrog opportunity

The historical parallel often cited is mobile telephony. Many African and South Asian countries leapfrogged landline infrastructure entirely, going straight to mobile because the older technology had become economically obsolete by the time those countries had the capital to invest. There is a similar window now — not to skip AI, but to skip the cloud-dependency model that most established economies built theirs on.

Organisations beginning serious AI deployment today have the option of doing so on infrastructure they own, with data that stays where it should, on terms they fully control. That option simply did not exist three years ago. The same shift in technical conditions that makes sovereignty achievable for European businesses is even more consequential for organisations operating in environments where the costs of cloud dependency are higher and the regulatory pressure is moving faster.

VI.

WHAT SOVEREIGNTY ACTUALLY LOOKS LIKE

Principles, not products.

Sovereign IT and AI infrastructure is not a single product. It is an architectural and operational pattern with several reinforcing principles — and like any architecture, it can be evaluated against clear criteria rather than marketing language.

What follows is a structured account of what genuine sovereignty requires in practice. None of these principles are exotic. Taken individually, each has existed in enterprise IT for decades. What is new is the maturity of the technical components that make them achievable together — at reasonable cost, at meaningful capability, across both IT infrastructure and AI workloads.

The principles

Data residency and control. Data never leaves the organisation's environment. Not for analytics, not for service improvement, not for model training, not for any third-party processing. This is the foundation; everything else depends on it.

Jurisdictional independence. The infrastructure and operational governance are not subject to foreign legal reach. This means hardware, hosting and operational control are held by entities headquartered in the relevant jurisdiction, with no parent-company exposure to foreign laws compelling data disclosure.

Auditability. Every layer — software, infrastructure, models, data flows — is open to inspection. This requires open-source or source-available software at the foundations, and transparent operational practices. Black-box vendor systems cannot, by definition, be sovereign.

Open foundations. Proprietary lock-in is avoided at every layer. Operating systems, databases, AI models, communication tools and infrastructure

orchestration are built on open-source foundations that can be inspected, modified and migrated.

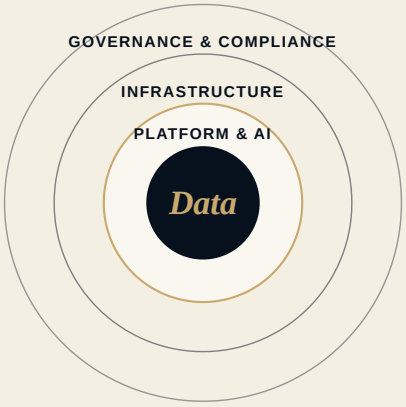
Predictable economics. Costs are fixed at the infrastructure level rather than scaling with usage. This makes multi-year budgeting accurate and aligns the cost of the technology with its capital nature rather than its operational use.

Operational continuity. The organisation can continue operating regardless of foreign provider business decisions, pricing changes, sanctions or geopolitical events. No critical function depends on the continuing willingness of an external party to provide service.

Human-led AI. Systems are semi-automated with meaningful human oversight maintained, rather than autonomous agents operating without inspection. This is both a practical safety measure and a hedge against the immaturity of fully autonomous AI in 2026.

DIAGRAM 4

The Layered Protection Model



EACH LAYER PROTECTS AGAINST:

- **Governance**
  Regulatory exposure, audit failure
- **Infrastructure**
  Foreign jurisdiction, provider lock-in
- **Platform & AI**
  Per-token costs, black-box behaviour
- **Data (core)**
  Breach, exfiltration, foreign access

Sovereign architecture treats protection as layered and inward-facing — every outer ring exists to keep the inner one inviolate.

A note on what sovereignty does not require

Sovereignty does not require building everything from scratch. It does not require rejecting useful technology, refusing all foreign vendors, or operating in isolation. Most sovereign architectures use foreign-developed open-source software — Linux, PostgreSQL, open-weight AI models — as foundational components. What sovereignty requires is that the **operational control, the data, and the legal accountability** remain with the organisation. Foreign technology becomes a tool rather than a dependency.

The distinction matters because it makes sovereignty practical. An organisation does not need to recreate the AI research frontier in order to deploy AI on its own infrastructure. It does not need to invent its own operating system to run sovereign servers. It needs to be deliberate about which components it owns, which it adopts, and where the boundary of control sits.

VII.

THE ECONOMICS

When the maths actually turns.

The case for sovereignty is often presented as a question of values — control, autonomy, security. It is also a question of finance. For most organisations doing substantive work with AI, the financial case for owned infrastructure has shifted decisively in 2025 and 2026.

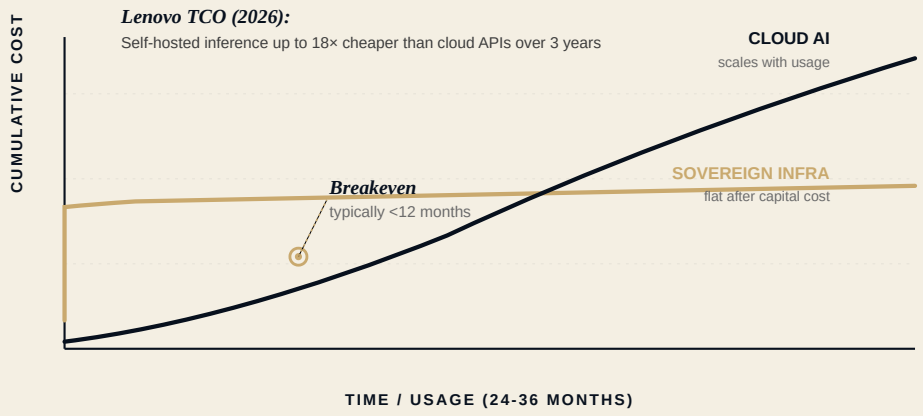
The shape of the curves

Cloud and on-premises infrastructure follow fundamentally different cost dynamics. Cloud costs scale linearly with usage: every additional query, every additional user, every additional gigabyte of data adds to the monthly bill. The model is elegant for low-utilisation, bursty workloads — and punishing for high-utilisation, sustained ones.

On-premises infrastructure follows the opposite pattern. The upfront investment is more significant: hardware, deployment, integration. But once that investment is made, the marginal cost of additional usage is close to zero. Power, cooling and maintenance are roughly constant. The cost per query, per token, per user falls as utilisation rises.

DIAGRAM 5

Cost Curves: Cloud AI vs. Owned Infrastructure



Illustrative cumulative cost over 24–36 months. Sources: Lenovo (2026); Anchoreo AI Research (2025); iFactory (2026).

The numbers

Lenovo's 2026 Total Cost of Ownership analysis, which compared on-premises infrastructure (using NVIDIA Hopper and Blackwell GPUs) against equivalent hyperscale cloud instances over a 5-year enterprise lifecycle, found that **on-premises infrastructure reaches breakeven in under four months for high-utilisation workloads**. Using a "token economics" framework, the analysis demonstrated that **self-hosted inference can be up to 18× cheaper than equivalent cloud API usage over a three-year period** for high-throughput AI workloads *(Lenovo, 2026).*

Independent analyses converge on similar conclusions. Anchoreo AI Research found **35% TCO savings for private AI data centres versus equivalent public cloud over a five-year horizon**, with operational expense reductions of 70% achievable for stable workloads *(Anchoreo via iFactory, 2026).* A 2025 industry analysis noted that renting a single NVIDIA H100 GPU in the cloud costs **between $5,000 and $75,000 per year** if used continuously — rivalling the $25,000–$30,000 outright purchase price of the same hardware *(Xenoss, 2025).*

Costs that don't appear on the cloud invoice

The visible cost of cloud is the monthly bill. The hidden costs are larger and harder to quantify. **Switching costs**: proprietary integrations, retrained teams, data migration, rebuilt workflows. **Compliance retrofitting**: investments required when regulatory requirements emerge for systems originally designed without them. **Breach exposure**: according to IBM's 2025 research, the average data breach now costs $10.22 million in the United States *(IBM via Tilkal, 2026).* Sovereign architectures that keep data in the controlled environment reduce this exposure structurally.

There is also the cost of **operational uncertainty**. The October 2025 AWS outage in the US-EAST-1 region disrupted UK government services including HMRC and the Government Gateway used by 50 million users, alongside major banks. Disruptions of this kind are rare but cumulative across time, and concentrated dependency makes them systemic rather than isolated.

THE CROSSOVER

For most mid-sized organisations doing real AI work

The breakeven point between cloud and owned infrastructure is reached within **4 to 18 months**, depending on utilisation intensity. Beyond that point, every additional month of cloud spending is a cost the organisation is choosing rather than absorbing.

The cost of doing nothing

The third financial picture worth considering is the *do-nothing* case. AI usage in most organisations is growing, not flattening. Cloud spending is therefore on an upward trajectory by default. Regulatory compliance retrofitting becomes more expensive the later it is undertaken. And the longer dependency persists, the higher the practical switching cost becomes. The economic case for sovereignty does not depend on cloud costs falling. It depends only on continuing to use the technology.

VIII.

THE TRANSITION PATH

From dependency to sovereignty, without disruption.

No sensible organisation moves from a fully cloud-dependent stack to fully sovereign infrastructure in a single change. Real transitions are phased, hybrid for an extended period, and honest about the complexity involved. The good news is that they are also well-understood — and the path looks broadly similar across organisation types.

A four-phase model

Phase 1 — Assessment. Before anything is built or migrated, the organisation needs a clear map of what it currently depends on. Which systems, which data flows, which regulatory obligations apply, and where the highest-risk dependencies sit. This phase produces a structured dependency inventory and a prioritised list of what should move first, what can wait, and what may reasonably remain on cloud infrastructure indefinitely.

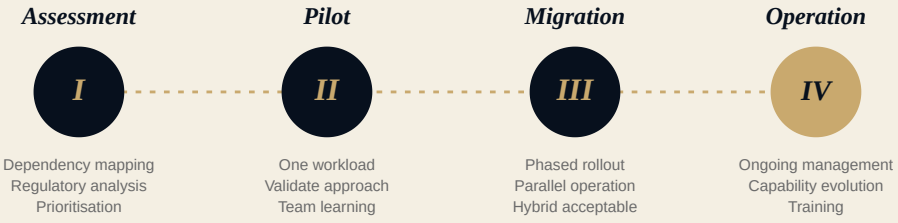
Phase 2 — Pilot. One workload — typically high in sovereignty value, modest in operational complexity — is moved to sovereign infrastructure to validate the approach. This is usually an AI workload (knowledge management, document analysis, research assistance) because the regulatory and cost pressures are concentrated there and the organisational learning value is high.

Phase 3 — Migration. Remaining priority workloads are migrated in sequence, usually with a period of parallel operation. This phase typically lasts six to eighteen months depending on scale, and many organisations choose to leave certain low-sensitivity workloads on cloud infrastructure indefinitely rather than migrate them. Hybrid is a legitimate end state.

Phase 4 — Operation. Once the priority workloads are running on sovereign infrastructure, the work shifts to ongoing operation: maintenance, monitoring, capability evolution, team training. This phase is permanent. Sovereign infrastructure is owned in the same way other capital assets are owned — it requires ongoing care, but the foundation is in place.

DIAGRAM 6

The Four-Phase Transition



Hybrid operation between phases is the norm, not the exception. Realistic timelines: 12–36 months end-to-end depending on scale.

The two pieces of advice most often given by organisations that have completed this kind of transition are simple. Start with a real workload, not a proof-of-concept disconnected from operational reality. And budget the team time honestly — sovereign infrastructure is owned, and like any owned asset it requires people who know how to operate it. This is rarely the bottleneck many organisations expect, but it is always a real cost worth planning for.

IX.

CASE SCENARIOS

How this looks in practice.

The principles in this paper apply differently across organisation types. The four scenarios that follow are anonymous composites — plausible situations rather than specific client cases — designed to illustrate how the sovereignty question actually lands in different contexts.

Scenario A

Mid-sized European manufacturer

An 800-employee manufacturer in the automotive supply chain, headquartered in Central Europe with operations across three countries. Classified as an "important entity" under NIS2 due to its position in the manufacturing supply chain. Uses Microsoft 365 across the organisation, AWS for operational data, and has begun trialling Microsoft Copilot for engineering documentation.

The sovereignty question: Production specifications, supplier contracts and intellectual property flow through cloud-based tools whose CLOUD Act exposure makes them difficult to reconcile with new customer requirements from German OEMs demanding data residency guarantees. Engineering AI usage is growing, and per-seat costs have become noticeable.

What an answer looks like: A hybrid architecture. Engineering documentation, AI-assisted quality control and supplier IP move to a sovereign on-premises platform. General productivity (email, calendars, basic file sharing) remains on the existing cloud platform with stronger data governance. NIS2 compliance is built into the new architecture from the start. Estimated breakeven on the AI workload alone: 10 months.

Scenario B

Regional healthcare network

A network of four hospitals and twelve associated clinics serving a regional population of 1.4 million. Classified as an "essential entity" under NIS2, subject to GDPR with the highest sensitivity tier of personal health data, and beginning to deploy AI-assisted diagnostic tools.

The sovereignty question: Patient data flowing through any cloud-based AI tool creates direct GDPR exposure post-Schrems II. The clinical interest in AI for radiology and pathology is real, but every off-the-shelf option creates legal risk that hospital legal counsel finds difficult to sign off.

What an answer looks like: Sovereign on-premises deployment of medical AI models, fine-tuned on the network's anonymised data, with data never leaving hospital infrastructure. Operational continuity is independent of foreign provider availability. Compliance documentation is auditable end-to-end. The transition is phased over 18 months.

Scenario C

European research university consortium

A consortium of three European universities running collaborative research programmes in pharmaceuticals and advanced materials. Research data includes patentable findings, partner-company proprietary information and human subject data.

The sovereignty question: The convenience of cloud-hosted research and AI tools is in direct tension with the IP protection requirements of industrial partners. Several recent collaboration agreements have included explicit data-residency clauses that the consortium currently cannot honour with existing infrastructure.

What an answer looks like: A shared sovereign research platform, owned by the consortium, with strict access controls per project. AI capabilities (literature review, molecular structure analysis, patent search) are deployed on self-hosted infrastructure. Industrial partners are reassured by documented end-to-end control. Research throughput improves; partnership exposure decreases.

Scenario D

Public sector institution in an emerging economy

A national government agency in an East Asian country, responsible for citizen services and revenue administration. Recent national legislation has introduced data localisation requirements with a two-year compliance window. Budget is constrained and foreign-currency exposure on existing cloud contracts has become a material concern.

The sovereignty question: Current systems rely heavily on US hyperscaler infrastructure. The new legislation makes this position untenable within two years. Foreign-exchange volatility has caused IT costs to rise unpredictably. There is also genuine national interest in building local technical capability rather than continuing to export budget abroad.

What an answer looks like: A staged transition to sovereign infrastructure hosted in-country, beginning with citizen-facing services and revenue systems. AI capabilities (document processing, citizen query handling, fraud detection) are deployed on owned infrastructure. Cost predictability replaces foreign-currency exposure. Local technical operations create jobs and capability that compound over time.

X.

CONCLUSION

The sovereignty imperative.

The technology decisions European organisations made over the past decade were reasonable. The decisions they make in the next few years will determine whether they keep the capability they have built — on terms they actually control.

The case made across this paper is a single connected argument. The IT and AI tools most businesses depend on are concentrated in a small number of foreign companies, in foreign jurisdictions, on commercial terms that increasingly do not serve their customers. The regulatory environment has changed in ways that turn what was once a preference into a legal obligation. The "sovereign cloud" offerings launched in 2025 are commercial responses to real concerns but, by their structural nature, do not fully resolve them. The same dependency dynamics, scaled to emerging economies, are felt with greater urgency and with stakes that include national strategic autonomy. And — most importantly — the practical alternative has matured to the point where genuine sovereignty is achievable across IT and AI workloads, at reasonable cost, with capability that matches or exceeds what dependency offered.

The choice between powerful technology and meaningful control over it is, in 2026, no longer the trade-off it was even three years ago. Owned infrastructure, well-configured local AI models, transparent open-source foundations and predictable economics are all now within reach for organisations of essentially any size in essentially any geography. What remains is the question of timing.

> *The organisations that engage with this shift on their own timing will be the ones who choose their position. Those who wait will have their position chosen for them by regulation, by costs, or by events.*

The path is structured: a clear assessment of current dependencies, an understanding of which regulatory obligations apply when, and a phased transition that begins with one workload of real operational value. The first step is rarely the hardest one. The first decision is.

That decision is, ultimately, simple. Sovereignty is not a rejection of progress. It is the foundation of strategic autonomy in a maturing technology landscape. The question is universal across geographies, scales and sectors. The answer is universal too. Own what you depend on — or accept that someone else does.

ABOUT O5 PARTNERS

Powerful technology, private by design, yours to control.

O5 Partners is a Czech company building private, data-sovereign IT and AI infrastructure for organisations that need to own the technology they depend on. We work at the intersection of cybersecurity, AI deployment, and IT consultancy — serving clients across Europe and, increasingly, in emerging markets.

Our flagship platform is deployed entirely on the client's own infrastructure. Data never leaves the client's environment, never reaches us, and is fully auditable at every level. There is no cloud dependency, no vendor lock-in, and no subscription to foreign AI platforms. EU regulatory compliance — including NIS2 — is built into the architecture from the ground up, not added as an afterthought.

We offer a complete proposition: software platform, infrastructure, expert consultancy, training and hardware — all under one roof. We do not compete with hyperscalers on price. We compete on sovereignty, security, auditability, and independence.

O5 PARTNERS

Powerful technology. Private by design. Yours to control.

O5 Partners s.r.o. · o5partners.com
General enquiries: info@o5.partners
Partnerships: partnership@o5.partners

REFERENCES

Sources and citations

References are presented in Harvard style. Online sources accessed between January and April 2026 unless otherwise noted.

1 Al Mokdad, A., 2025. *AI diplomacy in the Global South: Strategy or afterthought?* Medium. Available at: medium.com [citing UNESCO data].

2 Anchoreo AI Research, 2025. *On-Premises vs Cloud AI Models*. Cited in: iFactory, 2026. *Cloud vs On-Premise AI for Manufacturing: A Total Cost of Ownership (TCO) Comparison*. Available at: ifactoryapp.com.

3 ASEE, 2026. *EU Cloud Sovereignty: Why Businesses Are Moving Away from US Providers*. Available at: asee.io.

4 Cigref & Asterès, 2025. *Technological Dependence on American Software and Cloud Services: An Assessment of the Economic Consequences in Europe*. Paris: Cigref. Reported in: Computerworld, 2026.

5 Computerworld, 2026. *Gov't IT spending seen as key to building Europe's tech ecosystem*, March. Available at: computerworld.com.

6 Center for Strategic and International Studies (CSIS), 2025. *Divide and Deliver: How AI Can Serve the Global South*, October. Available at: csis.org.

7 Digital Samba, 2025. *Europe's Dependency on Microsoft: A Threat to Its Digital Sovereignty?* August. Available at: digitalsamba.com.

8 European Commission, 2024. *Directive (EU) 2022/2555 on Measures for a High Common Level of Cybersecurity Across the Union (NIS2 Directive)*. Available at: digital-strategy.ec.europa.eu.

9 European Commission, 2025. *Data Act (Regulation (EU) 2023/2854)*, applicable from 12 September 2025. Available at: digital-strategy.ec.europa.eu.

10 European Commission, 2025. *AI Act Service Desk: Timeline for the Implementation of the EU AI Act*. Available at: ai-act-service-desk.ec.europa.eu.

11 European Parliament Think Tank, 2025. *European Software and Cyber Dependencies* (PE 780.413), December. Available at: europarl.europa.eu.

12 Exoscale, 2026. *Sovereign Cloud and Data Sovereignty: An Overview*, January. Available at: exoscale.com.

13 Gartner, 2025. *Survey of Western European CIOs and IT Leaders*, May–July. Reported in: The Register, 2025. *Geopolitics push European CIOs to think local on cloud*, November. Available at: theregister.com.

14 Gislen Software, 2026. *Are US Cloud Services a Risk for European Companies?* March. Available at: gislen.com.

15 Greenberg Traurig, 2025. *EU NIS 2 Directive: Expanded Cybersecurity Obligations for Key Sectors*, August.

16 HarfangLab / Sapio Research, 2025. *State of European Cybersecurity Report Q2 2025*. Survey of 800 IT and cybersecurity decision-makers across Germany, France, Belgium and the Netherlands.

17 IBM, 2025. *Cost of a Data Breach Report 2025*. Reported in: Tilkal, 2026. *Cloud AI vs. On-Premise AI: The True Cost Comparison*. Available at: tilkal.co.

18 Legiscope, 2026. *EU AI Act Timeline: Key Dates and Deadlines*. Available at: legiscope.com.

19 Lenovo Press, 2026. *On-Premise vs Cloud: Generative AI Total Cost of Ownership (2026 Edition)*. Available at: lenovopress.lenovo.com.

20 Skadden, Arps, Slate, Meagher & Flom LLP, 2025. *NIS2 Update: EU Cyber Authority Sets Out Compliance Expectations*, August.

21 Synergy Research Group, 2025. *European Cloud Providers' Local Market Share*. Reported in: CNBC, 2026. *These four charts show how reliant Europe is on U.S. digital infrastructure*, February.

22 The Register, 2025. *UK politicians to draft outage blueprint after AWS calamity*, October. Coverage of the 20 October 2025 AWS DynamoDB DNS resolution failure in US-EAST-1.

23 UpCloud, 2025. *U.S. Hyperscalers and the European Cloud Data Sovereignty Gap*, October. Available at: upcloud.com.

24 World Economic Forum, 2026. *How the Global South is Reimagining the Future of AI*, February. Available at: weforum.org.

25 Xenoss, 2025. *Total Cost of Ownership for Enterprise AI: Hidden Costs and ROI Factors*, December. Available at: xenoss.io.