AIEONME FOUNDATION

AI Ethics, Oversight & Norms for Multilateral Engagement

in partnership with AIEONME FRONTIER (Global South)

STRATEGIC PAPER — JUNE 2026 UPDATE

From Spud to Scarcity

Frontier Cyber-AI Access Asymmetry and the Lithium That Powers It

An April–June 2026 update to the Spud / GPT-5.5 analysis

Why the model arrived, the regime hardened, and the Global South still pays the material cost

June 3, 2026

AIEONME Foundation | AIEONME Frontier | @aieonme

Author: Arq. Gustavo E. Cardozo | Founder & Executive Director

Ciudad Perico, Jujuy, Argentina (Lithium Triangle, NOA)

AIEONME-SP-2026-006 | Supersedes AIEONME-SP-2026-005

Bottom Line Up Front

In our April 23, 2026 paper we described OpenAI's frontier model "Spud" as pretrained but unlaunched, and warned that a restricted-release regime for frontier cyber-AI was being constructed around US-allied corporate consortia. Six weeks later, both predictions have resolved — and hardened.

First: Spud launched as GPT-5.5 on April 23, 2026 — not GPT-6 — and is now OpenAI's flagship, joined by a gated cyber variant (GPT-5.5-Cyber, May 7). Anthropic countered with Claude Opus 4.8 (May 28) and, on June 2, expanded Project Glasswing to roughly 150 new organizations across more than 15 countries.

Second: The Global South remains structurally excluded. Every named Glasswing and Trusted Access for Cyber partner is a US ally or Western corporate; India is the sole non-Western inclusion. No Latin American, African, or Lithium-Triangle nation has frontier cyber-AI access — even as a White House veto, a demonstrably leaky access regime, and a new June 2 voluntary pre-release review concentrate control further in Washington.

Third — this paper's core contribution: the extraction asymmetry. Argentina and Jujuy set lithium production records (Jujuy ~67,000 t in 2025; US$953M in total provincial mining exports) feeding the battery storage the AI datacenter buildout demands, while the region is denied the frontier AI those datacenters run. The Global South supplies the material base of an AI infrastructure it cannot access. This is the asymmetry no Northern think tank is naming.

---

1. What Changed: Spud Became GPT-5.5

OpenAI released GPT-5.5, codenamed "Spud," on April 23, 2026. It is the company's first fully retrained base model since GPT-4.5 — Chief Scientist Jakub Pachocki confirmed that the intermediate 5.1–5.4 versions were post-training layers on the GPT-5.0 base, while Spud was a genuinely new pretrain. President Greg Brockman called it "a new class of intelligence" representing "two years of research."

Reported benchmarks: 82.7% on Terminal-Bench 2.0, 88.7% SWE-bench Verified, 58.6% SWE-Bench Pro, approximately 84.9% GDPval, and a 1-million-token context window. API pricing roughly doubled to $5/$30 per million input/output tokens, with $30/$180 for the Pro variant.

1.1 Why GPT-5.5 and Not GPT-6

The model many expected as "GPT-6" shipped as a 5.x point release. OpenAI did not publicly explain the renaming. The practical read — shared by Gary Marcus and others — is that the gains are incremental: the SWE-Bench Pro score fell short of leaked "high 70s" projections, which supported keeping the 5.x branding. GPT-6 remains uncommitted, with most trackers placing it in Q3–Q4 2026 at the earliest. For AIEONME's thesis, this matters: a non-generational leap weakens the "existential capability gap" narrative and strengthens the case that the binding constraint is access, not raw capability.

1.2 The Competitive Backdrop

GPT-5.5 is a B2B counterattack. OpenAI had been in an internal "Code Red" since December 2025 as Anthropic's enterprise revenue surged. Anthropic stated on April 6, 2026 that its run-rate revenue had surpassed $30 billion, up from approximately $9 billion at the end of 2025 — CEO Dario Amodei cited "80x annualized growth." By late May, reporting placed Anthropic's run-rate near $47 billion. The frontier race is now as much a revenue contest as a capability contest, which shapes who gets access and why.

2. The Restricted-Release Regime Hardened

2.1 OpenAI: GPT-5.5-Cyber and Trusted Access for Cyber

On May 7, 2026, OpenAI announced GPT-5.5-Cyber in limited preview — its most permissive cyber model, for vetted defenders under Trusted Access for Cyber (TAC). The same day it published its "Cybersecurity in the Intelligence Age" action plan, explicitly framing expansion around "the US and its allies." A companion platform, Daybreak (built on Codex Security), launched May 11. From June 1, top-tier TAC members must enable phishing-resistant MFA.

Named partners include Cisco, Intel, SentinelOne, Snyk, Akamai, Cloudflare, CrowdStrike, Fortinet, Okta, Palo Alto Networks, Rapid7, Semgrep, Sophos, Tenable, Trail of Bits, and Zscaler, plus Bank of America, Goldman Sachs, JPMorgan Chase and Morgan Stanley. European expansion added Deutsche Telekom, BBVA, Telefónica, and Scalable Capital; South Korea (via KISA, May 27) and Japan were the first Asian entrants. No Latin American, African, or other Global South entity is named in any TAC announcement.

2.2 Anthropic: Opus 4.8 and the Glasswing Expansion

Project Glasswing's first results (May 22–26): Claude Mythos Preview scanned more than 1,000 open-source projects, flagging 23,019 issues, of which 6,202 were high or critical; independent firms validated over 90% of a sampled subset. But fewer than 1% had been patched upstream — exposing a triage bottleneck where AI-driven discovery vastly outpaces human remediation.

Claude Opus 4.8 (May 28) scored 69.2% on SWE-Bench Pro — beating GPT-5.5 and Gemini 3.1 Pro on that test — and was around four times less likely than its predecessor to let coding flaws pass unremarked. Anthropic said Mythos-class models would reach all customers "in the coming weeks."

On June 2, 2026, Anthropic extended Project Glasswing to approximately 150 new organizations across more than 15 countries, up from the 50 initial partners. The expansion came a day after Anthropic confidentially filed for an IPO following a $65B round at a near-$1 trillion valuation. Anthropic stated that since launch, Glasswing partners revealed more than 10,000 high or critical-level security flaws, and estimated that a major cyberattack could impact more than 100 million people. Participating countries reportedly include Australia, Canada, France, Germany, Italy, Switzerland, the Netherlands, Spain, Belgium, Sweden, India, Japan, New Zealand, and South Korea; named organizations include Okta, Samsung, SK Hynix, SK Telecom, and NATO.

The decisive fact: the expanded country list contains no Latin American, no African, and essentially no Global South partner except India.

Table 1 — The Restricted-Release Regime as of June 3, 2026

Lab / Program	Flagship	Cyber Variant	Access Model	Global South
Anthropic / Glasswing	Opus 4.8 (May 28)	Mythos Preview	Membership (~150 orgs, 15+ countries)	**India only**
OpenAI / TAC	GPT-5.5 (Apr 23)	GPT-5.5-Cyber (May 7)	Identity verification + tiers	**None named**
Google	Gemini 3.1 Pro	Big Sleep / CodeMender	Internal + selective	**None named**
Chinese open-weight	DeepSeek V4 / Qwen 3.5	Ungated	Open weights, no coalition	**Available to all**

The central contradiction sharpened over the window: the closed labs gate their most capable cyber models to Western insiders in the name of safety, while comparably capable Chinese open-weight models (DeepSeek V4 under MIT license, Qwen 3.5 under Apache 2.0) proliferate globally without restriction. MITRE OCCULT evaluations showed DeepSeek-R1 scoring over 90% on offensive cyber-knowledge tests. The security rationale is undermined by open-weight proliferation; the access asymmetry against the Global South persists regardless.

3. The Regime Is Leaky and Politically Contested

Two developments undercut the premise that restricted release reliably keeps frontier cyber capability out of unvetted hands.

3.1 The Leak

Mythos was accessed on its launch day by a private Discord group using a third-party contractor's credentials combined with URL-guessing. The breach chain ran through a chain of vendors into an Anthropic vendor environment. As one security executive put it, access controls are a policy, not an architecture, and policies fail. Third-party access is the structural vulnerability — which means the regime's exclusion of the Global South does not even deliver the security it promises.

3.2 The Veto and the Executive Order

The White House vetoed Anthropic's plan to expand Mythos to roughly 70 additional organizations (reported late April). After a draft executive order was scrapped (~May 21) amid lobbying, President Trump signed an EO on June 2, 2026 establishing a voluntary framework for early government access to step-change models before release; the federal review window was cut from 90 to 30 days, and incremental version updates are explicitly excluded. The Department of Defense labeled Anthropic a "supply chain risk"; Anthropic sued to reverse it. Analyst Zvi Mowshowitz warned of an emerging "ad-hoc prior restraint era" in which informal licensing "favors the connected and insiders." For the Global South, the lesson is stark: access is being decided by US executive discretion, not multilateral rule.

4. The Core Finding: The Lithium-AI Extraction Asymmetry

This is the contribution that distinguishes AIEONME's analysis from every Northern think tank. The argument is simple and material: the Global South extracts the raw materials that power the very AI infrastructure it is denied access to.

4.1 AI Is Driving Battery Demand

Global shipments of battery energy storage systems rose 75.5% to 421.2 GWh in 2025, with 600 GWh projected for 2026, as AI datacenters consume an increasing share of grid capacity (InfoLink Consulting). Utility-scale storage is projected at roughly 28% of planned 2026 US capacity additions. AI-datacenter-specific battery shipments are projected to exceed 300 GWh by 2030. SB Energy secured $1 billion from OpenAI and SoftBank in January 2026 for a 1.2 GW solar-plus-storage Stargate site in Texas — tying frontier compute directly to large-scale lithium battery deployment.

4.2 Lithium Is the Input, and Jujuy Is Breaking Records

Lithium represents roughly 60% of battery-cell cost, and LFP chemistry dominates stationary storage. The IEA projects lithium demand rising nearly ninefold by 2040. Jujuy reached approximately 67,000 tonnes of lithium in 2025 — up from 45,000 t in 2024 and 22,000 t in 2023, per provincial mining minister José Gómez. Jujuy's total mining exports (all minerals) hit US$953 million in 2025, making it Argentina's third-largest mining-exporting province; at the national level, Argentine lithium exports specifically reached approximately US$905 million in 2025 (Panorama Minero). Lithium Argentina AG reported full-year 2025 production of approximately 34,100 tonnes of lithium carbonate at Cauchari-Olaroz, with cash operating costs below $6,000 per tonne. Argentina is now the world's fourth-to-fifth-largest producer, with national output more than tripling in three years.

4.3 The Water and Sovereignty Cost Falls on the Region

Per the United Nations University, global lithium output of roughly 240,000 tonnes consumed an estimated 456 billion litres of water — equivalent to the annual domestic water needs of 62 million people in sub-Saharan Africa. In Chile's Salar de Atacama, lithium mining accounts for up to 65% of regional water usage. UNU found that 54% of energy-transition minerals sit on or near indigenous territories. Director Kaveh Madani framed it directly: what is sold as a sustainability solution is actively harming people elsewhere in the world.

4.4 The Argument: Landlord, Not Participant

Drawing on the green-extractivism literature (Thea Riofrancos), the "sacrifice zones" framing (UNU-INWEH), and the 2026 compute-colonialism scholarship, the region becomes landlord, not participant: it supplies lithium, copper and rare earths, bears the environmental and water costs, and is excluded from the frontier AI those resources help power. The Stargate Abilene buildout — 1.2 GW, up to 400,000 GPUs, mixing wind, solar, gas and battery storage — is the physical embodiment of compute that the Lithium Triangle materially enables but cannot use. (Notably, OpenAI and Oracle scrapped the planned 600 MW Abilene expansion in March 2026 amid financing strain, a reminder that the buildout's economics are themselves contested.)

5. Latin America's Governance Gap

No frontier-AI safety institute equivalent to the UK AISI or US CAISI exists in the region. The ILIA 2025 index (CENIA and ECLAC/CEPAL) found Latin America accounts for 6.6% of global GDP but only about 1.1% of worldwide AI investment. Latam-GPT (CENIA, 30+ institutions across 8 countries) and the UNESCO Regional AI-in-Education Observatory are capacity, governance and education efforts — not frontier safety or evaluation bodies.

National developments confirm the exposure. Brazil's central bank tightened cyber rules via Resolution 538/2025 after a July 2025 attack of more than R$1 billion routed through a vendor. Argentina absorbed roughly 5.7 billion cyberattack attempts in 2025, within a regional total near 850 billion, and launched a "Gemelo Digital" AI system reportedly architected with a foreign vendor — drawing sovereignty criticism over extraterritorial data exposure. The region defends critical infrastructure against AI-accelerated threats without access to the AI-accelerated defenses.

6. Critical Perspectives (April–June 2026)

Bruce Schneier characterized the Mythos and Glasswing launch as substantially a public-relations play that worked, noting that the firm Aisle replicated some Mythos-found vulnerabilities using older, cheaper public models — though that replication is partial. Gary Marcus argued observers were played and that gains are incremental, supporting the GPT-5.5 rather than GPT-6 read. Zvi Mowshowitz documented the executive-discretion drift toward informal prior restraint. The digital-colonialism scholarship of 2026 connects exclusion from defensive AI tools to bearing the extractive and environmental costs of the buildout — the exact linkage this paper formalizes.

AIEONME's position is balanced: the underlying cyber capability is real enough to matter, the marketing framing is inflated, the restricted-release architecture solves the labs' commercial problems as efficiently as it addresses security, and — the point others omit — the regime's costs and benefits are distributed along the same North-South axis as the material supply chain beneath it.

7. Recommendations

7.1 Immediate (0–3 months)

1. Publish this update as the authoritative Global South frame connecting the two asymmetries — frontier cyber-AI exclusion and material extraction — that no Northern think tank is linking. The empirical hook is unique: zero Lithium-Triangle nations appear in Glasswing or TAC, even as Jujuy sets lithium records.
2. Demand a seat, not just access. Advocate that any frontier cyber-AI access regime (Glasswing, TAC, the new US voluntary review) include Global South defenders — Latin American CSIRTs, central banks, and critical-infrastructure operators.
3. Propose a Latin American frontier-AI safety and evaluation capability, building on CENIA, ECLAC/CEPAL and ILIA, positioned to give the region evaluation access and a governance voice.

7.2 Staged (3–12 months)

4. Tie lithium royalties and benefit-sharing to compute access. Resource-supplying provinces such as Jujuy should negotiate AI-infrastructure and compute access within mining frameworks (e.g., Argentina's RIGI), converting raw-material leverage into capability rather than pure rent.
5. Engage the EU's August 2026 enforcement as an opening. The EU's own struggle to obtain Mythos access shows even wealthy jurisdictions are excluded; coalition with the EU AI Office on access norms could create precedent benefiting the Global South.

7.3 Benchmarks That Would Change This Analysis

- If any Latin American or Global South entity gains Glasswing or TAC access → shift from "exclusion" to "conditional inclusion" framing.
- If Mythos-class models reach general availability (Anthropic's "coming weeks" promise) → the restricted-release asymmetry partly dissolves; pivot to open-weight proliferation risk.
- If lithium prices sustain a recovery above ~$20,000/t → strengthen the extraction-leverage argument materially.
- If GPT-6 ships as a genuine generational leap → re-evaluate the "incremental gains" thesis underpinning the access-urgency case.

Conclusion: From Spud to Scarcity

Six weeks ago the question was whether Spud would arrive. It did — as GPT-5.5, an incremental step, not a generational rupture. That resolution matters because it relocates the urgency. The binding constraint on the Global South is not that the frontier is impossibly far ahead; it is that the region is locked out of a regime now governed by membership lists, identity gates, and US executive discretion — while supplying the lithium that powers the datacenters the regime runs on.

AIEONME Foundation and AIEONME Frontier occupy a position no other institution can claim: speaking from the Lithium Triangle, from the province whose salars feed the battery storage of the AI buildout, naming the asymmetry between what the region gives and what it is allowed to receive. The frontier labs converged on a governance model that treats Western incumbents as natural partners and everyone else as a supplier. The mineral supply chain and the capability supply chain run in opposite directions along the same North-South axis.

The window to assert a different arrangement is measured in months, not years — before the August 2026 EU enforcement, before GPT-6, before the next Glasswing expansion. The region that mines the future of AI must not be the region excluded from its defenses. That is the argument AIEONME will carry to CEPAL, to the OAS, and to the multilateral record. The mirror has become a ledger: it shows what the Global South contributes, and what it is denied. The task now is to balance it.

AIEONME Foundation · AIEONME Frontier

AI Ethics, Oversight & Norms for Multilateral Engagement

@aieonme | Founded February 1, 2026 | Jujuy, Argentina