GLOBAL AI GOVERNANCE

The Need for an AI Stability Board

Three summits, three years, and each declaration weaker than the last. What a real global stability body would need to do differently.

What this deck covers

01 Three summits, each with less binding language than the last

02 Why the UN's 2024 advisory body did not close the gap

03 Three working precedents: the FSB, the IAEA, the Montreal Protocol

04 Four functions no current body performs

05 A workable path that does not require universal agreement first

BLETCHLEY PARK, NOVEMBER 2023

28 + EU

countries signed the first AI Safety Declaration

A statement of shared concern about frontier AI risk. No enforcement mechanism, no standing body, deliberately.

SOURCE: THE BLETCHLEY DECLARATION, UK GOVERNMENT, NOV 2023

PARIS AI ACTION SUMMIT, FEBRUARY 2025

62 → 2

nations signed; the US and UK refused

The two governments that hosted the prior summits and set the safety agenda would not sign the third declaration.

SOURCE: BBC NEWS / CNBC, 11 FEB 2025

UK AI SAFETY INSTITUTE

15

months from opening to dropping "safety" from its name

Renamed the AI Security Institute in February 2025. The international network it coordinated was renamed the same week.

SOURCE: MLEX, FEB 2025

WORKING PRECEDENTS

Three institutions already show the pieces

Financial Stability Board

IAEA

EST. 1957

Montreal Protocol

1987

G20, 2009

Coordinates national regulators and monitors systemic risk across borders. Does not regulate a single bank directly.

Treaty-backed right to inspect, regardless of whether the country being inspected wants it performed.

A binding phase-down schedule tied to an independent scientific finding, with trade measures for non-parties.

What no current body performs

1. **Independent, compelled evaluation**
   Not a lab grading its own thresholds. Standing access before release, not a voluntary demo.

2. **Cross-border incident reporting**
   A finding in one country currently has no obligatory channel to reach the others.

3. **A common threshold registry**
   Twenty companies each define "severe risk" themselves. No shared, auditable standard exists.

4. **Escalation with real consequence**
   A public finding, a required disclosure, a coordinated response among recognising jurisdictions.

"

A scientific panel with better branding is not a stability body. The connective piece is a disclosure obligation and an escalation path, not another report.

On what the UN's 2024 advisory body left unbuilt

Start from what already exists

1 **Formalise the existing institute network**

The UK, US, Japan, France, Germany, Italy, Singapore, South Korea, Australia, Canada and the EU already coordinate. Give it mutual recognition and joint disclosure.

2 Audit the voluntary thresholds

Twenty companies already publish their own. Convert self-report into an externally audited registry.

3 Let a coalition of the willing move first

The Montreal Protocol opened with a smaller group and used market access to bring holdouts in later.

The institution can still be built before the crisis that would otherwise force it.

Every prior global risk regime connected a piece already in place to an enforcement mechanism that had been missing all along. AI governance has the report, the network, and the voluntary thresholds. It is missing the connective piece.

Terence Kok
AI & Innovation Strategist

SAVE · COMMENT · FOLLOW