AGI PREPAREDNESS INSTITUTE

Building the Fourth International Geneva

Introducing the AGI Preparedness Institute



September 2026

1. Why AGI Preparedness, why Geneva, why now?

The AGI Preparedness Institute is a new Geneva-based non-profit that works to prepare national and international institutions for the consequences of advanced AI. This brief introduction document explains the reasoning behind the Institute, what it plans to do in its first year, and how you can get involved.

a) Why AGI Preparedness?

AGI Preparedness is about increasing the capacity of institutions to manage the increasing societal, economic, and geopolitical consequences of advanced AI. This includes reducing the risk of loss of control and catastrophic misuse of AI, enabling human flourishing in a post-labor economy, and retaining human freedom in an age of superpersuasion and supersurveillance.

The most operational way in which AGI Preparedness exists today is through corporate Responsible Scaling Policies. These are voluntary commitments by AI companies with an if-this-then-that logic with increasing requirements for risk mitigations, such as increased corporate cybersecurity, with increasing levels of AI capabilities. This is good. However, corporate social responsibility is not sufficient:

- Corporate responsible scaling commitments are voluntary, and some aspects of them would be game theoretically more credible if they can be verified and enforced.
- Many AI harms arise not from one AI model, but from the cumulative use of many systems across society.
- For threats requiring only a specific absolute AI capability level (vs. relative superiority to another AI), there are strong forces favoring open diffusion regardless of what any single actor does. This requires resilience and contingency plans.
- Certain risks are more effectively managed upstream or downstream of the model itself. AI model companies cannot regulate social media platforms, Internet service providers, or domain-specific bottlenecks, such as DNA synthesis providers, which turn digital files into biological material.

Consequently, we need a range of actors from national governments, to public-private partnerships, to international institutions to urgently step up their efforts for societal preparedness for advanced AI. The reason why the Institute explicitly uses the term Artificial General Intelligence (AGI) in its name is that AGI is a forward-looking term. AI cannot be adequately understood or governed by just looking at the state-of-the-art today. The total computing power of the stock of AI chips and the efficiency of that compute are increasing exponentially. Intelligence is getting cheaper by ca. 40x per year! AI models have gone from scoring 5% on high school math in 2021 to solving open research problems in mathematics, including the first Millenium Prize problem in

September 2026. From 17.5% on Cybench in 2024 to finding an unprecedented amount of real-world software vulnerabilities in 2026. Anything that wants to be future-proof needs to anticipate the possibility of continued fast AI progress.

It is equally important to be clear about what the Institute does not mean with AGI Preparedness. It does not mean preparing for a single "AGI day", a one-off event after which humanity has either solved all problems or gone extinct. My mental model is that the transition to an "age of AGI" in which machines will be in charge of the most consequential decisions, is closer to a "compressed Industrial Revolution", a period of sustained, accelerating change rather than a single threshold. The goal of preparedness is to shape this transition so that humanity flourishes and retains meaningful agency permanently.

b) Why Geneva?

Ecosystems of institutions and people can be surprisingly sticky. They create specialized talent pipelines and tacit know-how about how institutions actually work that only diffuses locally. So many have tried to build the next Silicon Valley, and yet these efforts largely failed. Despite all of its challenges and shortcomings, SF is still the clear AI capital of the world.

In a similar sense, **Geneva has been a sticky international governance ecosystem for more than a century and across (more than) three world orders.** The oldest International Geneva institutions go back to the 19th century. Most notably the International Committee of the Red Cross was set up in Geneva in 1863. The first formal and permanent intergovernmental organization, the International Telegraph Union, set up its offices in Bern in 1868 (and eventually morphed into the International Telecommunication Union based in Geneva).

After the end of World War I, in 1920, Geneva was selected as a neutral home of the first general-purpose multilateral organization, the League of Nations. This was **the First International Geneva**. At its greatest extent, the League of Nations had 58 members. However, it was weakened by the US refusal to join, and it ultimately proved incapable of preventing aggression by the Axis powers in the 1930s. The onset of the Second World War in 1939 showed that the League had failed its primary purpose.

After the Second World War, the high-level political discussions moved to New York. The US, freshly irritated by Swiss banks' wartime dealings with Nazi Germany, wanted the newly launched United Nations headquartered on its own soil. And yet, **the Second International Geneva** still formed. The Palais des Nations stood ready, the International Labor Organization and the International Committee of the Red Cross had never left, and the technical machinery of the postwar multilateral order condensed around them: the UN's European office, GATT, the World Health Organization, the revised Geneva Conventions, CERN. The Cold War then increasingly turned Swiss neutrality from a liability into an asset: Geneva was one of the few tables where East and West would both sit, from the 1955 summit to Reagan–Gorbachev in 1985.

When the Cold War ended Geneva again renewed itself into **the Third International Geneva**. Economic globalization reached new heights on both ends of Lake Geneva with the rise and rise of the World Economic Forum and the newly established World Trade Organization. As a spillover of this, there has been a trend towards multistakeholderism and public-private partnerships rather than purely intergovernmental approaches. The creation of Gavi, the Global Fund, and the Internet Governance Forum fall into that bucket. Security discussions also became less state-centered and more human-centered. Theodor Winkler, a former Swiss intelligence officer, used the switch to an active neutrality policy in Switzerland to bring a new set of Genevan institutions into existence that would focus on a range of (non-traditional) security challenges from anti-personnel mines and the democratic control of armed forces to terrorism and crisis management.

The ways in which current international organizations work are not future-proof. However, the co-location of so many international organizations in Geneva still creates agglomeration effects. Similarly, the political stability and permanent neutrality of Switzerland remains a hard-to-replicate asset. As such, Geneva is a logical condensation point for adaptations and new institutions that can create another round of self-renewal, **the Fourth International Geneva,** to address international coordination issues in the age of AGI.

c) Why now?

AI is at an inflection point: 2026 is the year in which the misuse and loss of control potential of AI has become serious, as highlighted by AI models reaching superhuman levels in vulnerability discovery. The training wheels for AI governance are coming off. We have some level of recursive self-improvement without clarity on how strong the feedback loops will get. Increasingly autonomous agents are spreading through the economy faster than any prior technology. We are in the “foothills of the singularity”.

Digital International Geneva is at an inflection point: On June 20, 2025, the Federal Council backed a new strategy for International Geneva, which particularly emphasizes the goal of hosting events and organizations in Geneva with regard to digitalization, emerging technologies, and the environment. On June 25, 2025 the Swiss Federal Council decided that it wants to host the 2027 Geneva AI Summit, in line with the summits of Bletchley, Seoul, Paris, and Delhi. Starting from Fall 2026 the “Road to Geneva” starts and the Summit in June 2027 will be a key event for both international AI governance and for “Closing the Digital Geneva Gap”.

2. The AGI Preparedness Institute

The AGI Preparedness Institute is a Swiss non-profit organization founded in Geneva in 2026. Its purpose as defined in its statutes is to contribute to preparing international and national governance for developments in advanced AI and other transformative technologies, through:

a) research and analysis on the adaptation of international and national institutions;
b) publication and dissemination of work for the public, decision-makers, and International Geneva;
c) the design of and experimentation with new governance arrangements and mechanisms;
d) support for preparedness processes conducted by public institutions;
e) the organization of events, dialogues, and collaborations with academic, international, and civil-society institutions as well as private-sector actors.

2.1 Founder profile

The Institute is founded and directed by Kevin Kohler. In 2019, he was first author of a proposal for an "IPCC for AI". In 2024–25, as a researcher at the Simon Institute for Longterm Governance, he supported consultations during the UN General Assembly intergovernmental process on the Independent International Scientific Panel on AI and the Global Dialogue on AI Governance, as part of civil society. From April to July in 2026, as an individual consultant to UN University, he provided writing and editing support for the Panel's preliminary report. The AGI Preparedness Institute is independent of the Simon Institute and the UN AI Panel, and its views are its own.

That seven-year loop "spot the needed future institution early, engage the negotiations to design it, support its implementation" is the loop this Institute exists to run again, but faster, for the age of AGI.



2019 · Ideation — foraus: "IPCC for AI"



2024–25 · Institutional Design — Simon Institute: design options



2026 · Implementation — UN Panel: first report

Before founding the Institute, he worked at the Center for Security Studies at ETH Zurich on foresight, risk assessment, and crisis management for the Swiss federal administration; at the World Economic Forum, where he co-authored the Global Risks Report; and at the Simon Institute for Longterm Governance, where he led research

and multilateral engagement on the UN's AI governance track. He has published more than 50 essays on AGI futures on his blog "Machinocene".

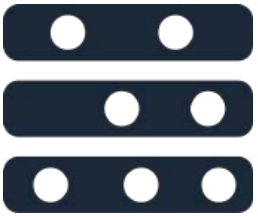
2.2 Initial activities

Between September 2026 and the Geneva AI Summit in June 2027, the Institute plans to work on the following areas:

- **Swiss AGI readiness.** Is Switzerland ready for advanced AI? A trilingual report (English, German, French) built on interviews across the federal administration, science, and industry, published in open access — and contributions on the road to the 2027 Geneva AI Summit.
- **Multilateral AGI readiness.** Publications and engagement on preparing the multilateral system for advanced AI — including contributions to the Global Dialogue on AI Governance and the Geneva AI Summit, work timed to the transition to a new UN Secretary-General, and experimentation with neutral, AI-native tools that make the public record of multilateral deliberations easier to navigate.
- **Forward-looking institutional and process ideas.** Developing governance designs whose time is coming — for example, a "Montreux Document" process restating how existing international law applies to advanced AI, or a Cosmic Endowment Fund for common-heritage space resources.
- **Workshops and briefings for policymakers.** Closed-door AGI scenario exercises for Swiss federal stakeholders, and a briefing series for diplomatic delegations in Geneva together with partner institutions.

2.3 Why our logo is a Swiss cheese

The logo of the Institute is a stack of three barriers, each pierced by holes that don't fully line up. It represents the Swiss cheese model of risk management, a model used in safety engineering from aviation to pandemic response. Its message is simple: No single protective layer is perfect. Every layer has holes. However, if the layers are independent and their holes do not align, hazards and threat actors still struggle to pass through. That is the Institute's view of AGI preparedness in one picture: safer AI models, prepared institutions, resilient societies are all necessary but not sufficient. An AI ecosystem that is aligned with human laws and values, corrigible, and resilient can only be achieved through a combination of technical AI safety and well-designed institutions. Also, for an institute based in Switzerland, a Swiss cheese was hard to resist.



3. Getting involved

The Institute is open for collaboration. If any of the following describes you, we would be happy to hear from you.

As a collaborator

If you work at a permanent mission, an international organization, a government, or a research institution and would like to co-host a briefing, request a scenario exercise, or explore joint research on the questions above, write to us.

As a builder

If you want to build AI-native tools for real governance processes or you sit inside an institution and have a concrete bottleneck that such a tool could relieve, the Institute intends to explore projects at this intersection. Expressions of interest are welcome.

As a contributor

The Institute is forming an advisory circle and also welcomes guest contributions. We care about rigor and relevance to live processes. Early-career authors with something real to say are explicitly welcome. Pitch us two paragraphs before you write the piece.

As a donor

The Institute is a Swiss non-profit association based in Geneva. Its work is funded by philanthropic grants and donations, and its resources are irrevocably dedicated to its statutory purpose. If you are considering supporting the mission, we will gladly share detailed plans and funding scenarios.

Contact: info@agipreparedness.org · agipreparedness.org · Geneva

Subscribe to the AGIPI newsletter



AGI PREPAREDNESS INSTITUTE

agipreparedness.org