

# Proposition for a Scientific Institution for the Development of Frontier AI

November 23

**Caitlin Amy Russell**

**Sofia Surla**

Following the recent UK Artificial Intelligence (AI) Safety Summit, there has been widespread criticism of the inability of government to effectively govern AI.<sup>1</sup> This is attributed to an inability of academic or government bodies to truly understand Frontier AI development.<sup>2</sup> Development of Frontier AI relies on three factors: compute, talent, and data.<sup>3</sup> Governments have data, and academic bodies have (some) talent, but without expensive and difficult-to-obtain computing resources, it's impossible to understand the true frontier and therefore regulate effectively.

To remedy this shortfall, thinktanks have proposed a scientific organisation, akin to CERN or CEPI whereby centralising data, compute, and talent, they could produce Frontier AI for citizens' common good.<sup>4</sup> Despite the fact that an international scientific organisation could democratise the benefits of AI and improve regulation, no concrete steps have been made towards this. Current proposals are rendered impossible by the desire to include China, the US, and current leading AI companies, yet in history only European-founded institutes have been successful in developing international research in scientific domains with military potential.<sup>5</sup>

---

<sup>1</sup> Ciaran Martin et al., "Expert Comment: Oxford AI Experts Comment on the Outcomes of the UK AI Safety Summit," University of Oxford: News and Events, November 3, 2023, <https://www.ox.ac.uk/news/2023-11-03-expert-comment-oxford-ai-experts-comment-outcomes-uk-ai-safety-summit>.

<sup>2</sup> Brianna Rosen et al., "AI Safety Summit: Trends, Challenges, and Opportunities" (Oxford, 2023).

<sup>3</sup> Ben Buchanan, "The AI Triad and What It Means for National Security Strategy," 2020.

<sup>4</sup> Daniel Zhang et al., "Enhancing International Cooperation in AI Research: The Case for a Multilateral AI Research Institute," 2022; Lewis Ho et al., "International Institutions for Advanced AI," July 10, 2023, <http://arxiv.org/abs/2307.04699>.

<sup>5</sup> Importantly, CERN is the only international scientific institute which had a focus on nuclear technology, which had civilian and military purposes.

Proposal: a European-led institute, without the involvement of current AI superpowers, could avoid current cooperation roadblocks and prove more effective.

Initial proposals were consistently expected to include all relevant State stakeholders, in particular the USA and China.

We offer an alternative option whereby the EU and adjacent countries would collaborate on their own AI organisation, the European Centre of Frontier AI Development (CEDIF).

The goals of such an organisation would be:

- Encouraging intra-European cooperation and knowledge-sharing
- Creating an example of publicly-funded safe Frontier AI development
- Making Europe a global player in the AI sphere and minimise talent brain drain

Instead of relying on the UN or other international organisations containing potentially at-odds players, the CEDIF would allow European countries to build their own capabilities in AI outside of the sphere of influence of the USA and China.

Taking this into consideration, this paper will look at the potential political pros and cons of such an endeavour. Part of the reasoning for excluding the two largest players in the AI landscape is the ongoing mistrust in Sino-American relations. <sup>6</sup>

---

<sup>6</sup> Choking off China's Access to the Future of AI, CSIRO 2022, <https://www.csis.org/analysis/choking-chinas-access-future-ai>

In the past, agencies such as the IAEA were created in order to “oppose” what were considered rival countries, the CEDIF on the other hand would exist with the purpose of building EU-internal capabilities.

---

**In trying to implement an institute as advocated for above there are several issues at play, first in feasibility (the how) and secondly if this would actually be an improvement on the current status quo.**

**Feasibility of a European-led institute which could act as a counter-hegemonic power:**

**1. Financing**

There is currently a huge discrepancy in the amount that the EU proposes to spend on developing AI, in comparison to China or the US:

In 2018, the EU launched a Coordinated Plan on AI in an effort to create a greater degree of coordination, attract more private investments and complement national investments with 1.5 billion euro by 2020, which is 70 per cent more than in the previous period (2014–17).<sup>29</sup>

The initial review of the Coordinated Plan on AI is expected in the first quarter of 2021. In the new budget for 2021–7, the EU will invest 2.2 billion euro in AI – to be complemented by funds coming from the 95 billion-euro Horizon Europe programme, of which about 20 per cent will go into R&D of the digital agenda.<sup>30</sup>

China has set itself the objective of becoming an AI superpower by 2030. In 2017, it adopted a strategy on AI that foresaw the investment of tens of billions in development as well as the application of AI. The objective set for 2020 was for the country’s core AI industries to

exceed 23 billion US dollars and related industries 150 billion. By 2025, China's core AI industries are to exceed 60 billion US dollars in value; its related industries, 760 billion – by 2030, they are expected to exceed, respectively, 150 billion and 1.5 trillion US dollars.<sup>31</sup>

(Ghiretti, Technology, 2021, pg. 10)

## 2. Talent

“In AI, the US and China hold the primacy – although the two excel in different areas. For example, the US has an advantage in the production of patents, R&D and in language-processing AI.<sup>26</sup> China, for its part, is more advanced in the application of AI as well as in the amount of investments that go into its R&D. Nonetheless, there are other areas in which competition between the two is still open – including the number of highly skilled experts and the application of AI to the military. Interestingly, despite China being the top provider of AI researchers, the US remains the top destination of researchers – including those from China – in the field (two thirds of whom select the US as their country of choice for work). (Ghiretti, Technological Competition, 2021, pg. 9.

What, then, is left for the EU? The Union generates a good share of skilled people in the sector; however, they often move abroad – again, mostly to the US – in order to continue studying and to work. Eighteen per cent of the world's top researchers in the field of AI come from Europe, a proportion that increases to 22 per cent if the UK is included, but only 10 per cent of them – 14 per cent in the case of the UK – work in Europe.<sup>28</sup> The US offers a more fertile environment for R&D as well as for applications, easing the possibility of a rewarding career both in terms of fame and money. The issue of the impossibility of scaling

in the EU is more or less consistent throughout the technology sector. One of the problems, albeit not the only one, is that existing competition rules prevent the formation of so-called champions within the EU while allowing extra-EU large firms not only to operate in the Union but also to acquire its enterprises. The result is that highly innovative EU SMEs or startups are acquired by foreign, often American and Chinese, giants. (Ghiretti, Technological Competition, 2021, pg. 9-10)

### **3. European Countries as Stakeholders:**

This is an area that needs more work:

1. Who are national stakeholders in EU? Might any country oppose the introduction of CERN for AI?
2. Business Interests?
3. Scope for increasing budget – would this be possible, or not?

### **4. Location of such an institute:**

Another issue that is highly contentious – benefits of an institute are likely to be linked to its location – could this actually move forward

## **What could be the implications of such an institute?**

### **1. Increasing likelihood of great-power conflict:**

- a. Increase likelihood of Chinese dangerous developments in AI because they feel threatened by growing European power.
- b. Realist Theory of IR: Multipolar Systems are by their nature destabilising

- c. Robert Trager – information sharing, if shared information shows close levels of development it increases likelihood of conflict, otherwise it decreases likelihood of conflict *Uncertainty, Information and Risk in International Technology Races*, 2023, <https://journals.sagepub.com/doi/epub/10.1177/00220027231214996>)
- d. Provides the US and China with increasing information for which to develop, an imbalance in European Transparency, with high barriers to information sharing from the other side.

## **2. Decreasing likelihood of great-power conflict:**

- a. EU acts as a multilateral power, rather than a “great power” (EU Multilateralism in a multipolar world) – third argument. This is as a multilateral institution that the EU acts fundamentally multilaterally in everything it does, which is more of a constructivist view of IR (the idea that international institutions and norms can determine behaviours of individual nation states). This means that even if the EU did develop a military edge, they would not use this increase international tensions, and likelihood of Great Power Conflict.
- b. Improved information sharing undermines the military “lead” of the US or China, as this only comes from a military capability that the other power does not have – relative power is more important than power.
- c. An EU lead might be the only way to get the other powers to “stand down” from increased brinkmanship, and by providing a counterweight to current Sino-US ambitions diffuse tensions

- d. Eventually, like CERN, the clear advances of this institution in comparison to others could promote transparency and push information sharing between China and the US (arXiv:2001.00463v2 [cs.CY] 9 Jan 2020)

**What is Europe's current position as a "balancing" power?** – much of the thinking comes from *EU Multilateralism in a multipolar world*

- 1. Non-superpower – benefited from protection of current superpowers.**
- 2. Third, or fourth, "great power"**
- 3. As a multilateral institution its contributions are fundamentally multilateral in the world.**

#### **Comparisons to the foundation of CERN:**

- 1. Historical Context:**

The war had profoundly affected developments in nuclear physics, sharply accelerating the pace of research. Countries such as Italy, France, Denmark and Germany, which had been at the cutting edge in the 1930s, were excluded from those developments by the war. The stakes surrounding the atomic bomb were such that the United States of America and the United Kingdom monopolized and maintained a veil of secrecy around nuclear research, especially after the bombing of Hiroshima and Nagasaki. (Pg. 1, Patrick Petitjean. Pierre Auger and the Founding of CERN. Petitjean, P., Zharov, V., Glaser, G., Richardson, J., de

Padirac, B. and Archibald, G. (eds). Sixty Years of Sciences at Unesco, 1945-2005, Unesco, pp.57-60, 2006. halshs-00166533 )

From 1949, European physicists, in particular French and Italian (notably Edoardo Amaldi, the mainstay of the project), joined forces to bridge the gap and to draw up a plan for a European laboratory. Their aim was to attain a level of research equivalent to that of the USA. The project was examined in Geneva in December 1949 at the European Cultural Conference, organized by the European Movement. It was endorsed by scientists, diplomats and science officials, but had not yet won governmental support. (Pg. 2, Patrick Petitjean. Pierre Auger and the Founding of CERN. Petitjean, P., Zharov, V., Glaser, G., Richardson, J., de Padirac, B. and Archibald, G. (eds). Sixty Years of Sciences at Unesco, 1945-2005, Unesco, pp.57-60, 2006. halshs-00166533 )

This implies that there was competitive intentions, not just cooperative ones.

## **2. Success of CERN?**

## **3. Cooperation vs. non-cooperation**

**What are the central questions which need more research to decide whether this would be an effective policy?**

1. Historical examples of counterweight military hegemonies.
2. Informational Transparency on increasing or decreasing tensions.
3. European ability and interest in acting as a feasible counterweight.
4. Informational Asymmetries in historical context.

5. Third-power and race-dynamics in IR.