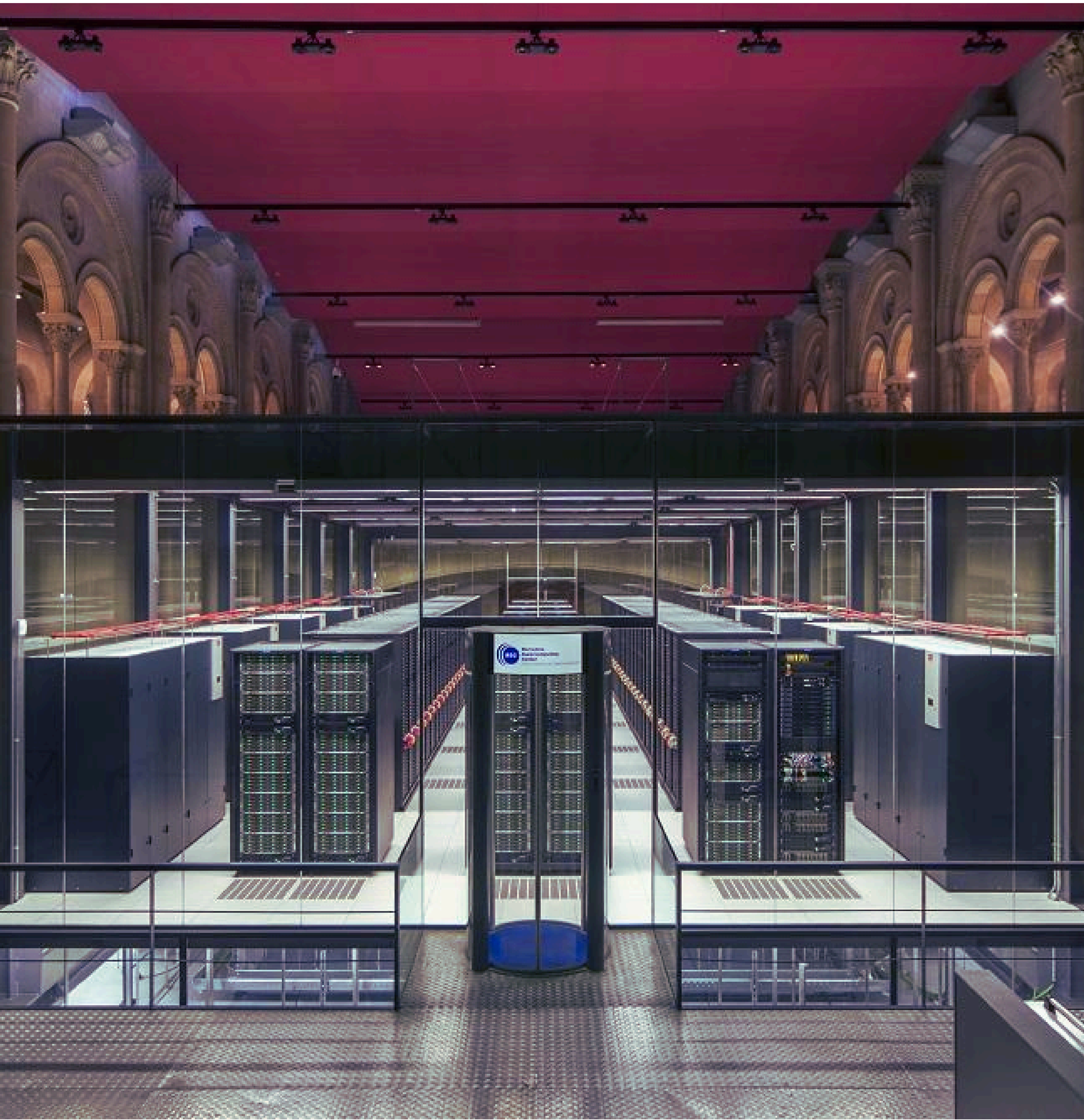
Compute for Good: The Potential of European Supercomputers for AI Research and Development

September 2024



Observatorio de Riesgos Catastróficos Globales (ORCG) is a science-diplomacy organization that works on formulating governance proposals that allow the comprehensive management of different global risks in Spanish-speaking countries. We connect decision-makers with experts to achieve our mission, producing evidence-based publications.

Bas, Guillem., Acelas Alejandro., (2024), Compute for Good: The potential of European supercomputers for AI research and Development, http://orcg.info/reportes

Executive Summary

The European High-Performance Computing Joint Undertaking (EuroHPC JU), a public-private consortium that pools supercomputing resources in Europe, has recently shifted its focus to artificial intelligence (AI). In particular, the European Commission is building so-called AI Factories, consisting of supercomputers and associated data centres, services, and specialised talent, to support the development and testing of AI models by a larger number of public and private users.

In total, the initiative is funded with a budget of around €7 billion, out of which €2.1 billion will be allocated to adapt existing supercomputers for AI workloads. The EuroHPC JU consists of eight operating supercomputers, to be joined in the next years by two exascale supercomputers in Germany and France.

Notwithstanding, EuroHPC resources are unlikely to be helpful for European companies to train frontier models for several reasons: (1) European public resources will likely fall short of what top tech companies already have or plan to acquire in the near future; (2) industry actors are unlikely to access a sufficient percentage of the infrastructure; and (3) prominent European companies already have substantial resources, so they are not incentivized to seek EuroHPC's compute.

However, EuroHPC's supercomputers are among the most performant public compute clusters in the world, opening a wide range of opportunities to fill neglected gaps. In this context, we recommend the following actions:

1. **Promote beneficial AI applications** that leverage the EU's global leadership in key industries and/or tackle neglected issues, such as healthcare, cybersecurity, or climate adaptation.
2. **Broaden involvement in trustworthy AI research** by lowering the minimum compute requirements for AI projects.
3. **Grant the AI Office a discretionary compute fund** for research and applications for evaluating frontier AI systems.
4. Use the recommended initiatives at the AI Factories as a **pilot to test what a future 'CERN for AI' could look like**.

Table of contents

1. Introduction

The European High-Performance Computing Joint Undertaking (EuroHPC JU) is a public-private consortium that pools resources to develop a "world-leading federated, secure and hyper-connected supercomputing, quantum computing, service and data infrastructure ecosystem" in Europe.

The members of the EuroHPC JU are the European Commission, 35 states —including the 27 Member States of the European Union and 8 Associated Countries[1]— and three private partners —the European Technology Platform for High Performance Computing, the Big Data Value Association, and the European Quantum Industry Consortium—. Currently, the EuroHPC JU is funded with a budget of around €7 billion for the period 2021-2027. Out of this funding, €3 billion comes from the Multiannual Financial Framework, while a similar amount is contributed by participating countries. €900 million coming from private members completes the budget.[2]

Historically, the Joint Undertaking has been mainly dedicated to scientific research in fields such as medicine, fundamental physics, material sciences, and earth science.[3] However, EU leaders have recently announced their intention to partially reorient the consortium towards efforts in artificial intelligence (AI) research and development. In the past years, progress in AI has shown that training and testing the most capable AI models requires large amounts of computing power.[4]

In the 2023 State of the Union, president of the European Commission Ursula von der Leyen first announced "a new initiative to open up [the EU's] high-performance computers to AI start-ups to train their models."[5] The initiative was ratified at the 4th European AI Assembly, where the Commission and the EuroHPC JU committed to opening and widening access to the EU's supercomputing resources to "support the further development and scalability of AI models, [...] reducing training time from months or years to a matter of weeks."[6]

Finally, in January 2024, the Commission launched the AI innovation package, including several measures aimed at putting those previous commitments into practice.[7] As one of its central pillars, the package included a proposal to amend the EuroHPC Regulation and allow for the establishment of so-called AI Factories, consisting of AI-dedicated supercomputers and associated data centres, services, and specialised talent.[8] The most relevant activities of

[1] The eight associated countries are the United Kingdom, Turkey, Israel, Serbia, Norway, North Macedonia, Montenegro, and Iceland.
[2] 'Discover EuroHPC JU'.
[3] 'Discover EuroHPC JU'.
[4] Sevilla et al., 'Compute Trends Across Three Eras of Machine Learning'.
[5] '2023 State of the Union Address by President von Der Leyen'.
[6] 'Commission Opens Access to EU Supercomputers to Speed up Artificial Intelligence Development'.
[7] 'Commission Launches AI Innovation Package to Support Artificial Intelligence Startups and SMEs'.
[8] European Commission, Proposal for a Council Regulation amending Regulation (EU) 2021/1173 as regards an EuroHPC initiative for start-ups to boost European leadership in Trustworthy Artificial Intelligence.

these factories would include support for the development and testing of AI models, as well as widening the use of AI-dedicated supercomputers to a large number of public and private users, including startups and small and medium-sized enterprises.[9] On 23 May, the Council reached a political agreement in favour of the amendment,[10] which was officially adopted on 17 June[11] and entered into force on 9 July.[12] Finally, the EuroHPC JU amended its Working Programme to incorporate the AI Factory pillar on 26 July.[13]

While the amendment awaited approval, the EU took some preliminary steps towards repurposing the EuroHPC JU. That can be seen in the new version of its Access Policy, approved in December 2023, which includes an access mode directed towards AI and data-intensive applications. This category is primarily meant to serve Industry organisations, Small to Medium Enterprises (SMEs), startups, and public sector entities requiring access to supercomputing resources to perform AI and data-intensive activities, such as training foundation models and generative AI, including large language models. This access mode will allocate up to 20% of total time from computers with AI accelerators during one year.[14] However, industry[15] will still have access to the other modes, including extreme scale access —up to 50% of pre-exascale and exascale computers for 1 year— and regular access —up to 70% of all computers for 1 year—.

By September 2023 and since 2021, the EuroHPC JU had received 120 applications related to AI. Out of those, 13 were for regular access and 17 were for extreme scale access, with 5 and 3 projects of the respective categories already operating.[16] These numbers are expected to grow significantly in the coming months, once the amendment of the EuroHPC JU regulation is implemented. According to official sources, EuroHPC JU members will invest a total amount of €2.1 billion from the overall budget in acquiring new or upgrading existing supercomputers with AI capabilities, creating supercomputing services in AI, and developing AI-oriented microprocessors and skills support.[17]

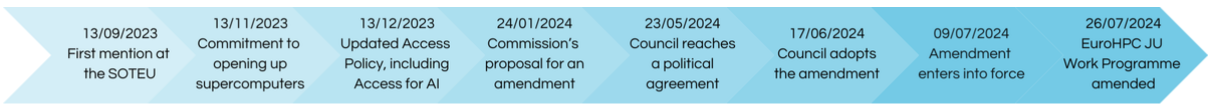


Figure 1. Timeline of initiatives to repurpose European supercomputers for AI R&D.

[9] European Commission.
[10] 'AI: Council Reaches Political Agreement on the Use of Super-Computing for AI Development'.
[11] Council Regulation (EU) 2024/1732 of 17 June 2024 amending Regulation (EU) 2021/1173 as regards a EuroHPC initiative for start-ups in order to boost European leadership in trustworthy artificial intelligence.
[12] 'The "AI Factories" Amendment to the EuroHPC JU Regulation Enters Into Force'.
[13] 'The EuroHPC Joint Undertaking Amends Work Programme to Incorporate the New AI Factory Pillar'.
[14] 'Access Policy of the EuroHPC Joint Undertaking Supercomputers'.
[15] Industry refers to large and small private companies, and is distinguished from academia and the public sector.
[16]The other applications were for Benchmark Access and Development Access, both of which allocate significantly less compute (up to 5% of all computers). See 'AI and HPC: Enhancing Research Possibilities'.
[17] 'AI Factories'.

The rest of this report is divided into two sections. First, we provide a technical overview of EuroHPC JU's supercomputers and assess their potential for several AI-related activities, including the development of cutting-edge models. Second, we suggest four recommendations on how those resources could be allocated to strengthen the European ecosystem for trustworthy and beneficial AI.

2. The potential of European supercomputers

The EuroHPC JU consists of eight operating supercomputers located in eight different countries of the European Union.[18] These supercomputers will be joined by two exascale supercomputers in Germany and France, expected to be ready by 2024 and 2026, respectively. Table 1 lists these ten supercomputers, their location, and performance.

Supercomputer	Location	Sustained performance (FLOP/s)[19]	Peak performance (FLOP/s)
JUPITER (planned for 2024)	Jülich (Germany)	~1.00E+18 (9E+19 FP8 FLOP[20])	
Jules Verne[21] (planned for 2026)	Bruyères-le-Chatel (France)	~1.00E+18	
LUMI	Kajaani (Finland)	3.86E+17	5.39E+17
Leonardo	Bologna (Italy)	2.47E+17	3.14E+17
MareNostrum 5	Barcelona (Spain)	1.78E+17	3.12E+17 (to be extended to 4.50E+17[22])
Meluxina	Bissen (Luxembourg)	1.28E+16	1.83E+16
Karolina	Ostrava (Czech Republic)	9.80E+15	1.29E+16

[18] 'Our Supercomputers'.
[19] Performance includes GPU and CPU partitions.
[20] 'NVIDIA Grace Hopper Superchip Powers JUPITER, Defining a New Class of Supercomputers to Propel AI for Scientific Discovery'.
[21] Jules Verne is not the final name of the supercomputer, which is still to be defined, but the name of the consortium that will host it. However, in this report, *Jules Verne* will be used as the name of the supercomputer for practical purposes.
[22] 'Spanish Government Strengthens AI Capabilities of MareNostrum 5 Supercomputer, Managed by BSC'.

Decaulion	Guimarães (Portugal)	7.22E+15	1.00+E16
Vega	Maribor (Slovenia)	6.92E+15	1.00E+16
Discoverer	Sofia (Bulgaria)	4.52E+15	5.94E+15

Table 1. Operating and planned supercomputers constituting the EuroHPC JU. Performance expressed in floating-point operations per second (FLOP/s). Source: EuroHPC JU.

This report will focus on the five most powerful supercomputers of EuroHPC JU, i.e. those built at pre-exascale —LUMI, Leonardo, and MareNostrum 5— and those planned to be at exascale —JUPITER and Jules Verne—. In the June 2024 Top 500 Supercomputers list, LUMI, Leonardo, and Marenostrum 5 ACC[23] were 5th, 7th, and 8th respectively, while JUPITER and Jules Verne would make it into the Top 3. However, it is worth noting that many large clusters are not included in the Top500, especially because it is an opt-in list that requires participating organisations to run a specialised LINPACK benchmark, which most industry players don't do. Table 2 shows the largest known compute clusters and compares their performance and number of specialized chips to those of our selected European supercomputers.

AI Compute Cluster	Stage	FP16 FLOP/s	Chip type	# Chips
xAI rumor	Planned	6.75E+20	B200	300,000
CEIBA: AWS&NVIDIA	Planned	1.04E+20	GB200	20,736
Memphis Supercluster (xAI)	Operating	9.90E+19	H100	100,000
Microsoft Arizona	In construction	7.43E+19	H100	75,000
Google A3	Operating	2.57E+19	H100	26,000
Meta GenAI 2024a	Operating	2.43E+19	H100	24,576
Meta GenAI 2024b	Operating	2.43E+19	H100	24,576
Voltage Park	In construction	2.38E+19	H100	24,000
Jupiter	In construction	*2.38E+19*	*GH200*	*24,000*
Inflection AI Cluster	In construction	2.18E+19	H100	24,000

Table 2. Largest known AI compute clusters in the world. Preliminary data by Pilz et al. (forthcoming). Data has not yet been verified.

[23] Marenostrum 5 has two partitions: the Accelerated (ACC) partition, ranking 8th, and the general purpose partition (GPP), ranking 22nd.

The prospect of German and French exascale supercomputers looks especially relevant, as both countries have nurtured the ever-growing companies Aleph Alpha or Mistral AI. The German AI Action Plan, updated in 2023, mentions JUPITER and high-performance computing as a pillar of the strategy to "make infrastructure AI-ready with targeted initiatives and specifically improve access for the AI community, especially start-ups."[24] On its side, the report of the French AI Commission aims at making France a major location for computing power, which would involve securing 30,000 H100s or equivalent by 2024, offering enough computing power by 2026 to support five companies training next-generation foundation models in a reasonable time, and supporting installation and electrical connection for private computing power.[25] At the 2023 edition of Viva Technology, an annual technology conference held in Paris, French President Emmanuel Macron and CEO of Mistral AI Arthur Mensch participated in a panel where the former committed to "strengthening public-private initiatives to defend a policy of opening European supercomputers to initiatives aimed at training foundation models."[26]

Both the power of upcoming European supercomputers and the political willingness could lead the reader to think that the EuroHPC JU could foster European strongest companies in their quest to train cutting-edge general-purpose AI models that compete with globally leading companies. However, this is unlikely to occur for several reasons.

First, though impressive, **European public resources will likely fall short of what top tech companies already have or plan to acquire in the near future**. Most of those companies not only have larger clusters but have also begun to acquire next-generation chips such as NVIDIA Blackwell, which is significantly more powerful than previous generations like Hopper—as an example, when compared to Hopper H100, Blackwell GB200 NVL72 provides up to a 30x performance increase in LLM inference and up to a 4x increase in LLM training.[27] Though the exact capacity of those companies is mostly unknown, the following examples might serve as an illustration:

- Microsoft has reportedly partnered with Oracle to build a cluster of 100,000 GB200 chips.[28] This amounts to 17x FP16 peak performance when compared to Jupiter. The company also appears to have an internal target to amass 1.8 million AI chips by the end of 2024.[29]

- Meta recently announced that they are building two new clusters with 24,576 GPUs each and installing 350,000 H100 GPUs by the end of 2024, reaching a total

[24] 'BMBF-Aktionsplan Künstliche Intelligenz: Neue Herausforderungen Chancenorientiert Angehen'.
[25] 'IA: Notre Ambition Pour La France'.
[26] *Intelligence Artificielle: Les Annonces d'Emmanuel Macron à VivaTech*.
[27] 'NVIDIA GB200 NVL72'.
[28] Gardizy, 'Musk's xAI and Oracle End Talks on a Server Deal'.
[29] Stewart, 'Microsoft Has a Target to Amass 1.8 Million AI Chips by the End of the Year, Internal Document Shows'.

computing power equivalent to 600,000 H100s.[30] This would equal 25x FP16 performance when compared to Jupiter.

- xAI has reportedly built a cluster of 100,000 H100s,[31] reaching 4x FP16 performance when compared to Jupiter.

Considering these numbers, even if a company managed to get exclusive access to a supercomputer for several months—which, as explained below, seems extremely unlikely—, the resulting training run might still lag behind others powered by private infrastructure.

Second, **industry actors are unlikely to access a percentage of resources that decisively contributes to frontier model training**. As of May 2024, the EuroHPC JU's Access Policy offers, as part of the mode for AI and data-intensive applications, up to 20% of total time from AI-optimized supercomputers, which is in fact subject to allocation demands from other access models for those systems in a given period. Moreover, the maximum compute investment of a single AI factory is capped at €400 million.[32] This reality is unlikely to change dramatically in the near term—in fact, departing drastically from this policy might go against the very nature of EuroHPC JU, whose mission is to widen the use of supercomputing infrastructure to a large number of users.

Third, **prominent European companies already have substantial resources**. For instance, Mistral AI partners with Microsoft to get access to Azure supercomputing infrastructure for AI training and inference workloads, which has allowed the company to scale their previous efforts and develop Mistral Large and Large 2, their largest models to date.[33] Likewise, Aleph Alpha has partnered with NVIDIA and Hewlett Packard Enterprise (HPE) to build Alpha ONE, presented as the fastest European commercial data centre.[34] Luminous models, Aleph Alpha's main product, were trained on an HPE supercomputer.[35]

This makes it unlikely that such companies will seek EuroHPC's compute to train their models, especially considering operational constraints. While distributed training appears to be feasible—in fact, some training runs have been split between data centres—, model training is often centralised to improve the latency of communication between GPUs and thus improve utilisation. Moreover, European companies with the capacity to access compute by other means might lack the interest to benefit from EuroHPC JU due to those technical barriers and perhaps other access burdens.

---

[30] Lee, Gangidi, and Oldham, 'Building Meta's GenAI Infrastructure'.
[31] Musk, 'Memphis Supercluster'.
[32] 'Decision of the Governing Board of the EuroHPC Joint Undertaking No 41/2024 Amending the Joint Undertaking's Work Programme and Budget for the Year 2024 (Amendment No 3)'.
[33] Boyd, 'Microsoft and Mistral AI Announce New Partnership to Accelerate AI Innovation and Introduce Mistral Large First on Azure'.
[34] Schreiner, 'OpenAI Competitor Aleph Alpha Launches Europe's Fastest Commercial AI Data Center'.
[35] Hotard, 'Sovereign and Trustworthy AI Is at the Core of Hewlett Packard Enterprise's Investment in Aleph Alpha'.

In conclusion, boosting European champions to compete at the global level seems to be an unproductive path for the EuroHPC JU to follow. Leading companies in the continent might be better off seeking private compute elsewhere, with European public supercomputers being reserved for smaller actors.

Nevertheless, as shown in Table 3, European supercomputers appear to be among the most notable public computing infrastructure in the world.

Infrastructure	Budget	Performance for AI traning	Purpose
European High-Performance Joint Undertaking (European Union)	€7 billion (2021-2027), out of which €2.1 billion will be dedicated to AI.[36]	~8.9E+17 FLOP/s (operating) + ~9.0E+19 FP8 FLOP/s (Jupiter) + >1.0E+18 FLOP/s (Jules Verne)	Multiple, with a recent shift to AI-oriented activities.
National AI Research Resource (United States)	$2.6 billion estimation (2024-2030)	3.77E+18 FLOP[37]	AI-oriented activities
Beijing AI Public Computing Platform (China)	Unknown	~3.00+18 FLOP[38]	AI-oriented activities
AI Compute Access Fund (Canada)	$2 billion (2025-2030)	TBD	AI-oriented activities
AI Research Resource (United Kingdom)	£300 million (initial investment)	4.09E+17 FLOP (Dawn) 2.10E+18 FLOP/s (Isambard)[39]	AI-oriented activities.

Table 3. Public compute infrastructure in selected countries.

Such computing power places Europe in a privileged position, as it opens a wide range of opportunities for the EuroHPC JU to fill any gaps that major actors currently neglect due to market incentives. In the following section, we provide a set of recommendations on how the

[36] According to the recently approved amendment, "[n]o additional resources will be required from the EU budget, but the proposed measures concern a redeployment of the resources available in the Joint Undertaking."
[37] Miller and Gelles, 'The NAIRR Pilot: Estimating Compute'.
[38] Weilan, 'Beijing Launches AI Public Platform as Demand Mounts for Computing Power'.
[39] Trueman, 'The Fastest AI Supercomputer in the UK Is Now Operational'.

European Commission could distribute EuroHPC's resources to ensure both competitiveness and trustworthiness.

3. Recommendations

Given the crucial importance of compute for AI research and development, decisions around the allocation of computing resources constitute a key point of intervention.[40] As such, European policymakers must ensure that the direction taken by the EuroHPC JU is aligned with the European objectives of developing an ecosystem for safe and trustworthy AI.

As illustrated by the analysis in the previous section, the computing resources from the EuroHPC JU seem unfit to bolster European companies in their efforts to stay at the frontier of AI development. However, this infrastructure could focus on users that would otherwise be unable to access compute, such as smaller companies, researchers, or auditors. In fact, there's a breadth of opportunities in AI outside of building state-of-the-art general-purpose AI models. Organisational capacity constraints mean that leading AI labs can only focus on a few streams of work—typically, improving the capabilities of their frontier AI models and developing a handful of AI safety research projects. Meanwhile, endeavours such as domain-specific scientific applications or exploratory AI safety bets are still underdeveloped and could greatly benefit from additional European resources.

Greater access to compute could facilitate the development of such projects, both by eliminating financial barriers to the acquisition of large compute quantities and by reducing operative and logistic barriers of access to AI models from academia and small organisations. In relation to the first point, the current landscape is severely affected by the so-called compute divide, which refers to the unequal access to computing power enjoyed by well-resourced firms on the one hand and small actors on the other hand.[41] Second, academia often faces difficulties to productively absorb large amounts of resources—notwithstanding, they can develop methods that might be reused to great profit in large-scale applications.

To bridge this gap, we recommend European policymakers to prioritise projects that address the gaps left by leading AI providers. In particular, we suggest the following actions:

1. Promote beneficial AI applications that leverage the EU's global leadership in key industries and/or tackle neglected issues, such as healthcare, cybersecurity, or climate adaptation.
2. Broaden involvement in trustworthy AI research by lowering the minimum compute requirements for AI projects.

[40] Sastry et al., 'Computing Power and the Governance of Artificial Intelligence'.
[41] Besiroglu et al., 'The Compute Divide in Machine Learning'; Ahmed and Wahed, 'The De-Democratization of AI'.

3. Grant the AI Office a discretionary compute fund for research and applications for evaluating frontier AI systems.
4. Use the recommended initiatives at the AI Factories as a pilot to test what a future 'CERN for AI' could look like.

Promote beneficial AI applications that leverage the EU's global leadership in key industries and/or tackle neglected issues, such as healthcare, cybersecurity, or climate adaptation.

The creation of a robust European ecosystem for AI development, gathering compute, data, and talent, is an excellent opportunity to achieve two objectives. First, the EU should **leverage its competitive advantages** by supporting work in critical industries where Europe is already at the forefront. In particular, we welcome the GenAI4EU initiative, which aims to contribute to the development of novel use cases and emerging applications of general-purpose AI in Europe's 14 industrial ecosystems, as well as the public sector (see Figure 2).[42] According to a survey, access to compute is one of the most important bottlenecks for generative AI startups in Europe, given limited availability and high costs. For that reason, around half of those startups report that they would benefit from accessing the EuroHPC JU.[43] The Commission plans to invest an estimated amount of €500 million in the initiative.[44]

Besides doubling down on Europe's comparative advantages, we suggest that the EU **promote differential technology development** by avoiding risk-increasing technologies and preferentially advancing risk-reducing, defensive, and substitute technologies.[45] This could include developing vertical foundation models that address industry-specific challenges, tackle intrinsic problems in AI R&D, or deliver neglected public services.

[42] 'European AI Office'.
[43] Hutchinson, Goll, and Muegge, 'Generative AI in the European Startup Landscape 2024'.
[44] 'Communication on Boosting Startups and Innovation in Trustworthy Artificial Intelligence'.
[45] Sandbrink et al., 'Differential Technology Development'.

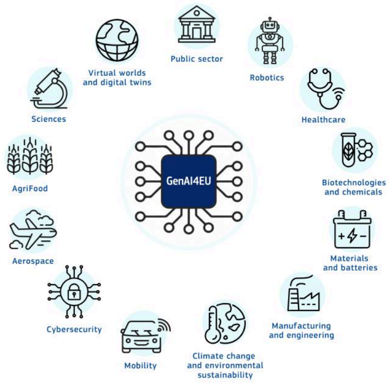


Figure 2. Industrial ecosystems considered in the GenAI4EU initiative.

Generative AI startups in the EU mostly focus on developing downstream applications on top of existing large foundation models (62.9%), followed by providing development tools and infrastructure for generative AI models (41.6%) and developing foundation models (31%).[46] However, we expect that most of the awarded projects will fall in the latter category. That is because model pre-training is significantly more computationally expensive than developing downstream applications, and the EuroHPC JU's Access Policy limits the allocation of compute for inference runs—which is intrinsic to most of the activities in the second stream—to a 10% of the overall allocation.[47]

For that reason, in this section, we highlight some sectors where the EuroHPC JU could be crucial to *develop* AI-based solutions that foster Europe's competitiveness and address complex challenges at the same time. For each subsection, we provide some examples of projects already operating in European supercomputers, as well as startups working in the sector.[48]

Healthcare and biotechnology. AI will have a wide range of applications in healthcare, for example, by helping provide tailormade healthcare solutions, detect early diseases, or improve epidemiological surveillance. For instance, MareNostrum 5 is intended to enhance medical research in the design of new drugs, vaccine development, and simulations of virus propagation,[49] while LUMI will soon host a project developing a foundation model for cancer diagnosis and treatment.[50] Among startups, there are numerous examples—among them, French startup Owkin uses AI to uncover biomarkers and mechanisms linked to diseases

[46] See Hutchinson, Goll, and Muegge, 'Generative AI in the European Startup Landscape 2024'. In the survey, multiple selections were possible, as many startups conduct their activities in more than one category.
[47] 'Access Policy of the EuroHPC Joint Undertaking Supercomputers'.
[48] Mentioning a startup is intended as an illustration and should not be interpreted as an endorsement in any case. The authors of this report do not have any conflicts of interest with regards to any European AI startup.
[49] 'European Supercomputer MareNostrum 5 Starts up at BSC'.
[50] 'New Finnish LUMI Projects Announced: AI Innovations in Healthcare and Beyond'.

and treatment outcomes, while German Ada provides AI-driven personalised healthcare through a virtual application.

Cybersecurity. It is important for AI systems to be used defensively to protect critical assets, as well as ensuring the reliability of AI systems themselves. For example, LUMI currently resources a project working on the development of an AI assistant for differentially private ML to avoid memorization of sensitive training data.[51] Previously, it has also been suggested that the supercomputer could power research in cybersecurity.[52] According to its recent strategy, the Spanish government has allocated a budget of €12m for R&D on AI and cybersecurity,[53] which could leverage MareNostrum 5 if needed. In the private sector, the French startup Gatewatcher has developed a generative AI assistant that helps with threat detection and incident response, among others.

Climate change and environmental sustainability. AI can help model climate change and predict the effects of certain activities on the environment, among many other applications to tackle climate change.[54] For instance, MareNostrum 5 will improve climate change simulations with the use of AI and big data analysis, as well as digital twins of the Earth.[55] On its side, LUMI has enabled the development of ML solutions for increased agricultural efficiency[56] and accelerated materials discovery,[57] both of which are crucial for sustainable development, and has recently announced the acceptance of projects to build AI innovations in climate adaptation and solar storm monitoring.[58] Another important endeavour is the development of batteries—in that regard, the EU-funded Battery 2030+ flagship initiative and the Swedish startup Northvolt aim at an AI-driven accelerated discovery of innovative battery materials that optimise energy efficiency.

Public sector. Generative AI can also help improve public services. For example, European supercomputers could power the development of general-purpose AI models that deliver neglected public services. One such example could be work on improving LLM performance in European minority languages. For instance, LUMI enabled the training of PORO 34B, an open multilingual LLM that outperformed all existing open language models on the Finnish language, functioning as a proof point for an innovative approach in developing AI models for languages with scarce data resources.[59] LUMI has also supported the addition of endangered Finno-Ugric languages to the machine translation engine of the University of

[51] 'LUMI Powers AI Assistant Development for Privacy-Preserving Machine Learning Models'.
[52] 'One of the World's Mightiest Supercomputers, LUMI, Will Lift European Research and Competitiveness to a New Level and Promotes Green Transition'; 'Solving Intricate Research Challenges with LUMI'.
[53] 'Estrategia de Inteligencia Artificial 2024'.
[54] Rolnick et al., 'Tackling Climate Change with Machine Learning'.
[55] 'European Supercomputer MareNostrum 5 Starts up at BSC'.
[56] 'DigiFarm Revolutionises Agriculture with Artificial Intelligence'.
[57] 'Accelerating Materials Discovery and Design with Machine Learning'.
[58] 'New Finnish LUMI Projects Announced: AI Innovations in Healthcare and Beyond'.
[59] 'Europe's Open Language Model Poro: A Milestone for European AI and Low-Resource Languages'.

Tartu.[60] Likewise, Spain will foster a public-private collaboration, including the Barcelona Supercomputing Center—host of MareNostrum 5—, to develop an LLM in Spanish and Spain's co-official languages.[61] Finally, the Large AI Grand Challenge awarded two startups developing language technologies for multilingual support, one encompassing all 24 official EU languages and the other one targeting Balto-Slavic languages spoken by 155 million individuals within the EU and candidate countries.[62] Cross-lingual performance disparities are an open challenge[63] that needs to be addressed to ensure that all citizens enjoy the benefits of AI equally.

For a more exhaustive list of promising sectors the EuroHPC JU could contribute to, including robotics, manufacturing, mobility, and aerospace, please refer to the European Commission's Communication on boosting startups and innovation in trustworthy artificial intelligence.[64] For a selection of promising European companies using AI in those and other sectors, please refer to the European AI Startup Landscape, which maps the AI startup ecosystem across four prominent hubs in Europe: France, Germany, the Netherlands, and Sweden.[65]

Those and other projects can be further supported financially by existing initiatives such as Horizon Europe's European Innovation Council (EIC) Accelerator[66] and the InvestEU Programme.[67]

Importantly, large-scale projects should be thoroughly screened and overseen. When evaluating a project in the AI and Data-Intensive category,[68] evaluators should ensure that the project adheres to the EU AI Act. Moreover, they should assess that the impact of the project—one of the assessment criteria—is indeed overall positive and does not carry any unacceptable risks. That is especially true for sectors such as cybersecurity and biotechnology, where AI models might possess dual-use capabilities that can be used both for good and harm. For higher-risk projects, the EuroHPC JU might also consider implementing Know-Your-Customer (KYC) checks by keeping records of the approved research group and providing a high-level overview of the purpose for compute usage.[69]

---

60 'Preserving Endangered Languages with LUMI: Machine Translation for Rare Finno-Ugric Languages'.
61 Blanchar, 'Sánchez Anuncia En El Mobile Que España Desarrollará Un Modelo de Lenguaje de IA En Castellano y Las Lenguas Cooficiales'.
62 'Large AI Grand Challenge – The Winners Are Out!'
63 Zhang et al., 'Don't Trust ChatGPT When Your Question Is Not in English'; Li et al., 'Quantifying Multilingual Performance of Large Language Models Across Languages'.
64 'Communication on Boosting Startups and Innovation in Trustworthy Artificial Intelligence'.
65 'The European AI Startup Landscape'.
66 'EIC Accelerator'.
67 'InvestEU Programme'.
68 'EuroHPC JU Call for Proposals for AI and Data-Intensive Applications Access Mode - Full Call Details'.
69 Egan and Heim, 'Oversight for Frontier AI through a Know-Your-Customer Scheme for Compute Providers'.

Broaden involvement in trustworthy AI research by lowering the minimum compute requirements for AI projects.

The compute divide refers to the unequal access to computing power between well-resourced firms and smaller actors like academia and startups.[70] While compute divide measures often focus on participation in the most compute-intensive projects, such as training foundation models,[71] valuable research also occurs at a smaller scale, particularly in academia.

Limited access to funding, compute infrastructure, and engineering expertise partially explains the disparity in compute use between industry and academia.[72] However, academics may also deliberately choose less compute-intensive projects due to the distinct aims of academic research compared to industry. Academic research typically advances basic knowledge and explores various domains for novel insights, while industry research focuses on developing commercial products.[73] Larger-scale models and applications tend to benefit industry research more directly.[74]

Many valuable academic contributions come from smaller-scale projects that explore novel approaches, refine existing techniques, or investigate the ethical and societal implications of AI technologies. Academia not only introduced many crucial innovations for machine learning[75]—today's dominant AI paradigm—but has also heavily shaped the development of AI systems by publishing seminal work on their interpretability,[76] robustness,[77] privacy,[78] and fairness.[79]

While the EuroHPC JU initiative has significantly reduced barriers to accessing large-scale computing resources through its offerings and education programs, its current policies don't address the needs of smaller—yet equally valuable—projects. For instance, the minimum compute request for a project through the AI and Data-Intensive Access Call is approximately 1000 times greater than what was required to finetune the renowned Alpaca

[70] Ahmed and Wahed, 'The De-Democratization of AI'.
[71] Besiroglu et al., 'The Compute Divide in Machine Learning'.
[72] Gelles et al., 'Resource Democratization'.
[73] Muller, '1.1.1 Industry and Academia'.
[74] Klinger, Mateos-Garcia, and Stathoulopoulos, 'A Narrowing of AI Research?'
[75] Krizhevsky, Sutskever, and Hinton, 'ImageNet Classification with Deep Convolutional Neural Networks'; Rumelhart, Hinton, and Williams, 'Learning Representations by Back-Propagating Errors'.
[76] Bau et al., 'GAN Dissection'; Doshi-Velez and Kim, 'Towards A Rigorous Science of Interpretable Machine Learning'; Ribeiro, Singh, and Guestrin, '"Why Should I Trust You?'
[77] Madry et al., 'Towards Deep Learning Models Resistant to Adversarial Attacks'; Goodfellow, Shlens, and Szegedy, 'Explaining and Harnessing Adversarial Examples'; Zou et al., 'Universal and Transferable Adversarial Attacks on Aligned Language Models'.
[78] Carlini et al., 'Extracting Training Data from Large Language Models'; Shokri et al., 'Membership Inference Attacks against Machine Learning Models'; Dwork et al., 'Calibrating Noise to Sensitivity in Private Data Analysis'.
[79] Buolamwini and Gebru, 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification'; Hardt et al., 'Equality of Opportunity in Supervised Learning'; Dwork et al., 'Fairness through Awareness'.

model from the open-source Llama 2 model.[80] Moreover, the AI and Data-Intensive Access Call limits compute usage for model inference to 10% of the total endowment, restricting many research avenues that focus on examining properties of pre-trained models, such as model evaluation or interpretability research.

Although academics have more options for accessing smaller compute provisions compared to large-scale resources, the logistical barriers associated with funding and access often remain substantial. These obstacles can deter researchers from engaging in even small-scale AI projects.

Consequently, we recommend **lowering the minimum compute requirements for AI projects** (e.g., to 200 node hours) and **removing restrictions on compute usage for model inference**. While this change requires expanding capacity for vetting and onboarding proposals, current EuroHPC investments can accommodate many new projects.

First, EuroHPC can leverage its recent budget investments in **continuous integration and development tools (CI/CD)** for its supercomputers to reduce technical support needs. Standardising access and delegating responsibilities to users can automate part of the onboarding process, easing demands on technical staff.

Second, to scale application vetting, EuroHPC could **expand its expert network and explore new compute allocation methods**, following recent meta-science literature on improving research impact through grant allocation experimentation.[81] Many of these changes to grant allocation could be implemented by Hosting Entities for AI Factories, which are tasked with proposing "AI user-friendly" access policies that contribute to national AI ecosystem development.

Third, we recommend that AI Factory hosts **specialise in specific research areas** to improve project evaluation and maximise researcher collaboration benefits. **Many of the most promising research bets are in AI Safety**, the field that aims to identify causes of unintended behaviour in machine learning systems and develop tools to ensure these systems work safely and reliably.[82]

[80] Alpaca finetuning took 3 hours on 8 A100 Nvidia GPUs (see Taori et al., 'Alpaca: A Strong, Replicable Instruction-Following Model'.). The minimum compute request among EuroHPC JU supercomputers is 7000 node hours on the Vega GPU cluster, which has 4 A100 GPUs per node. While the GPU hours needed to finetune the final Alpaca model significantly underestimates the compute required for a full academic project, it serves as a useful reference for the scale of compute typically needed for less compute-intensive academic projects.
[81] Wagner and Alexander, 'Evaluating Transformative Research Programmes'; Sharma et al., 'Piloting and Evaluating NSF Science Lottery Grants: A Roadmap to Improving Research Funding Efficiencies and Proposal Diversity'; Nielsen and Qiu, 'A Vision of Metascience: An Engine of Improvement for the Social Processes of Science'.
[82] Rudner and Toner, 'Key Concepts in AI Safety: An Overview'.

Relative to its importance, this field is still underinvested in by AI labs. If powerful AI systems are developed in the coming decade, we may lack the technical means to prevent their misuse or uncontrolled proliferation for large-scale cybercrime, social manipulation, or biological weapon development.[83] Encouragingly, several concrete research agendas[84] are already addressing these risks, with a growing cohort of researchers from academia and independent organisations contributing to the effort. The EuroHPC JU is well-positioned to amplify this work by providing necessary computational resources, potentially accelerating progress in AI safety research. By expanding compute access, EuroHPC has the potential to bolster various research approaches to trustworthy AI, guide the safe adoption of AI systems, and ultimately boost European AI competitiveness.

Grant the AI Office a discretionary compute fund for research and applications for evaluating frontier AI systems.

In the 2023 AI Safety Summit, celebrated in the United Kingdom, Ursula von der Leyen outlined a framework for mitigating risks from AI systems. The first pillar of that framework relies on "a thriving and independent scientific community, equipped with the means to evaluate AI systems." This community would "need public funding and access to the best supercomputers."[85] Within the proposed framework, she stated the European AI Office should work with the scientific community at large, contributing to the creation of standards and testing practices for frontier AI systems.

To make good on that vision, we propose **granting the AI Office a discretionary compute budget** to **conduct in-house projects** and, most predominantly, to **regrant it to selected projects** from academia and third-party evaluators. In addition to compute grants, the AI Office should **offer API credits for accessing AI models**. This facilitates access for researchers who lack the technical expertise to manage compute clusters and ensures they are not limited to evaluating models only from providers willing to host on EuroHPC infrastructure.[86]

By facilitating compute and model access, the EuroHPC JU can enable extensive evaluation and scrutiny of cutting-edge models by AI researchers, social scientists, ethicists, and other non-technical experts. This could include, for instance, collaborations between the EU AI

[83] Bengio et al., 'Managing Extreme AI Risks amid Rapid Progress'.
[84] Shevlane et al., 'Model Evaluation for Extreme Risks'; Anwar et al., 'Foundational Challenges in Assuring Alignment and Safety of Large Language Models'; Räuker et al., 'Toward Transparent AI'; Reuel et al., 'Open Problems in Technical AI Governance'; Hendrycks et al., 'Unsolved Problems in ML Safety'.
[85] 'Remarks of President von Der Leyen at the Bletchley Park AI Safety Summit'.
[86] Anderljung, Heim, and Shevlane, 'Compute Funds and Pre-Trained Models'.

Office and the UK AI Safety Institute, given that the UK is part of the Joint Undertaking.[87] Independent assessment is crucial to uncover potential biases, flaws, or unintended consequences in AI systems that developers may have overlooked, inadequately investigated, or failed to disclose.[88]

The implementation of the compute fund could draw inspiration from and collaborate with existing efforts to promote AI safety research and model evaluation. Table 4 presents recent initiatives fostering the ecosystem of AI safety and model evaluations, many of which are already producing impressive results. For example, the Center for AI Safety compute cluster has enabled over a hundred academic papers, garnering more than 2000 citations in 2024 alone.

Taking cues from these initiatives, the proposed compute fund could, for example, offer technical support led by domain experts, fund expenses beyond compute costs, earmark resources for specific research areas, or facilitate collaboration between external researchers and model providers.

Project	Support Provided	Focus Area
Anthropic's Third-Party Evaluation Initiative[89]	Funding Technical guidance from Anthropic domain experts In-house scale up of successful projects or commitment to purchase final product	Capability assessment for public security risks Metrics for scientific application potential, harmfulness, multilinguality, and societal impact Infrastructure, tools, and methods for developing evaluations
National AI Research Resource (NAIRR) Pilot[90]	Compute, cloud, and generative AI credits and access Collaboration opportunities with AI organisations Models, datasets, and documentation	AI for scientific discovery, open-source models, data exploration, public engagement, privacy, among others Applications in AI safety, medical treatments, advanced manufacturing, environmental science, and infrastructure resilience
National Deep Inference Fabric (NDIF)[91]	Structured access to internal activations and finetuning of large models Workshop training and tutorials	Understanding of AI decisions Auditing and evaluation for high-stakes applications Tools for AI safety and risk

[87] 'The United Kingdom Joins EuroHPC Joint Undertaking'.
[88] 'Third-Party Testing as a Key Ingredient of AI Policy'.
[89] 'A New Initiative for Developing Third-Party Model Evaluations'.
[90] 'About NAIRR Pilot'.
[91] 'National Deep Inference Facility for Very Large Language Models (NDIF): Project Proposal to the NSF'.

	Open-source online community	management
Center for AI Safety (CAIS) Compute Cluster[92]	Dedicated GPU cluster with flexibility on compute usage and changes to original research plan	Machine learning safety projects in robustness, monitoring, alignment, and systemic safety[93]
Systemic AI Safety Fast Grants (UK)[94]	Funding Support from the UK AI Safety Institute Connection to relevant stakeholders	"[C]oncrete, actionable approaches to significant systemic risks from AI"
Open Philanthropy's LLM Agent Benchmark Request[95]	Funding for LLM API credits and other forms of compute	Benchmarks measuring LLM agents' performance on consequential real-world tasks

Table 4. Selected initiatives to support AI safety research and evaluation.

In addition, the compute fund could serve as a starting point for other applications of EuroHPC's resources and expertise for international AI governance. Two approaches we consider particularly promising are **building a secure compute cluster for deep access to frontier AI models** and **extending discretionary compute budgets to other European institutions advancing trustworthy AI**.

First, a secure interface for model access beyond inference (e.g., allowing finetuning or interpretability techniques on model weights) could allow researchers to detect issues missed by black-box testing and provide better explanations of model decisions.[96] While the EuroHPC JU has expertise in compute cluster security,[97] protecting model weights from capable actors would require significant additional investment. This would most likely imply building a physically isolated facility, employing confidential computing to protect model weights while in use, and employing an experienced security team to counter advanced persistent threats.[98]

Second, directing EuroHPC's resources to **flexible compute allocations can serve as a powerful tool for drawing top AI talent to European institutions**. This could include universities

[92] 'CAIS Compute Cluster'.
[93] Hendrycks et al., 'Unsolved Problems in ML Safety'.
[94] 'Systemic AI Safety Fast Grants'.
[95] 'Request for Proposals: Benchmarking LLM Agents on Consequential Real-World Tasks'.
[96] Heim, 'A Trusted AI Compute Cluster for AI Verification and Evaluation'; Casper et al., 'Black-Box Access Is Insufficient for Rigorous AI Audits'.
[97] 'Decision of the Governing Board of the EuroHPC Joint Undertaking No 41/2024 Amending the Joint Undertaking's Work Programme and Budget for the Year 2024 (Amendment No 3)'.
[98] Nevo et al., 'Securing AI Model Weights: Preventing Theft and Misuse of Frontier Models'.

with strong AI research track records, national and European research centres, and independent organisations focused on trustworthy AI development.

While the EuroHPC JU may already offer researchers comparable computing resources to those in industry, the application process and project plan constraints can diminish its appeal. Even beyond financial considerations, AI researchers may choose to work in industry because it offers greater flexibility to execute on their research vision.

By connecting talented researchers with appropriate resources, expertise, and model access, the EuroHPC JU can help create a robust model evaluation ecosystem that promotes transparency, trust, and accountability in AI development across Europe and beyond.

Use the recommended initiatives at the AI Factories as a pilot to test what a future 'CERN for AI' could look like.

In her Political Guidelines for 2024-2029, Ursula von der Leyen announced her intention to "propose to set up a European AI Research Council where [the EU] can pool all of [its] resources, similar to the approach taken with CERN[99]."[100] This announcement arrived after the idea had been recurrently proposed by experts for years, both as a general concept for an international project[101] and as a EU-led initiative specifically.[102] Among policymakers, former UK Prime Minister Rishi Sunak had floated the possibility of a CERN-like body to conduct international research on AI.[103] Notably, in the EU, the idea was recently echoed by the Scientific Advising Mechanism of the European Commission, which proposed a CERN-inspired European Distributed Institute for AI in Science (EDIRAS) that would "provide public scientists and researchers [...] access to the infrastructures and inputs needed to undertake cutting-edge research with AI in all scientific disciplines."[104]

If properly designed, **a CERN-like body could fulfil the three recommendations put forward in this report**: develop beneficial AI applications, support a variety of trustwortyy AI research agendas, and establish an ecosystem for model evaluations. Regarding the former, as put by

[99] CERN, the European Organisation for Nuclear Research, is a scientific joint venture involving 23 Member States that offers purpose-built particle accelerators and detectors for physicists and engineers to study fundamental particles.
[100] von der Leyen, 'Europe's Choice: Political Guidelines for the next European Commission'.
[101] Marcus, 'Artificial Intelligence Is Stuck. Here's How to Move It Forward.'; Ho et al., 'International Institutions for Advanced AI'.
[102] 'Moonshot in Artificial Intelligence: Trustworthy, Multicultural Generative AI Systems for Safe Physical Interaction with the Real World'; Renda, 'Towards a European Large-Scale Initiative on Artificial Intelligence: What Are the Options?'
[103] Parker, 'Rishi Sunak to Lobby Joe Biden for UK "Leadership" Role in AI Development'.
[104] European Commission, Directorate-General for Research and Innovation, *Successful and Timely Uptake of Artificial Intelligence in Science in the EU*.

the Centre for European Policy Studies (CEPS), this new institution could "secure the design, development, deployment, and uptake of cutting-edge trustworthy AI solutions for science, industry and public services."[105] In particular, as argued by the Scientific Advising Mechanism, the initiative could incentivise the prioritisation of AI research with major benefits for EU citizens in fields such as healthcare, advanced materials, social cohesion, and climate adaptation.[106] As for the latter two, experts have suggested that a large-scale scientific collaboration with significant compute, engineering capacity, and access to cutting-edge models could attract researchers worldwide to work on AI safety and the development of auditing tools.[107] Likewise, the AI Office could be granted privileged access to the CERN for AI and the evaluation tools it creates.

At the same time, such an initiative would have to overcome significant challenges. The most prominent one is having the capacity to build and maintain the required technical, social, and organisational infrastructure. This includes not only adapting existing supercomputing facilities but also building data infrastructure, attracting and retaining talent, and increasing coordination among involved agencies. Moreover, when conducting research on frontier AI models, the project would have to ensure robust security to protect the model's weights, algorithmic details, and other trade secrets.[108] Several projections have estimated the budget for a project of these characteristics at €35-120 billion, to be invested in three to seven years.[109]

The actions proposed in this report for the EuroHPC JU could serve as a **pilot to test what some elements of a future 'CERN for AI' could look like**. In the future, the EuroHPC JU could be one of the pillars of the new European AI Research Council. Other relevant actors could include the Directorate-General for Communications Networks, Content and Technology (DG-Connect), including the AI Office and the associated scientific panel of independent experts; the Directorate-General for Research and Innovation (DG RTD); the European Innovation Council and Small and Medium-sized Enterprises Executive Agency (EISMEA); the European Research Council (ERC) and European Research Council Executive Agency (ERCEA); the European Research Executive Agency (EREA); and the Joint Research Centre (JRC). To leverage its full potential, this multistakeholder effort shall be coordinated by strong leaderships, involve the private sector, and be open to international partnerships.[110]

---

105 Renda, 'Towards a European Large-Scale Initiative on Artificial Intelligence: What Are the Options?'
106 European Commission, Directorate-General for Research and Innovation, *Successful and Timely Uptake of Artificial Intelligence in Science in the EU*.
107 Ho et al., 'International Institutions for Advanced AI'; Jones, 'A "CERN for AI" – What Might an International AI Research Organization Address?'
108 Ho et al., 'International Institutions for Advanced AI'.
109 'Moonshot in Artificial Intelligence: Trustworthy, Multicultural Generative AI Systems for Safe Physical Interaction with the Real World'; Renda, 'Towards a European Large-Scale Initiative on Artificial Intelligence: What Are the Options?'; Juijn et al., 'A CERN for AI: The EU's Seat at the Table'.
110 Juijn et al., 'A CERN for AI: The EU's Seat at the Table'.

References

'Access Policy of the EuroHPC Joint Undertaking Supercomputers'. EuroHPC Joint Undertaking, 2023.

Ahmed, Nur, and Muntasir Wahed. 'The De-Democratization of AI: Deep Learning and the Compute Divide in Artificial Intelligence Research'. arXiv, 2020. https://doi.org/10.48550/ARXIV.2010.15581.

AI Safety Institute. 'Systemic AI Safety Fast Grants', n.d. https://www.aisi.gov.uk/grants.

AI-Boost. 'Large AI Grand Challenge – The Winners Are Out!', n.d. https://aiboost-project.eu/large-ai-grand-challenge-the-winners-are-out/.

Anderljung, Markus, Lennart Heim, and Toby Shevlane. 'Compute Funds and Pre-Trained Models', 11 April 2022. https://www.governance.ai/post/compute-funds-and-pre-trained-models.

Anthropic. 'A New Initiative for Developing Third-Party Model Evaluations', 1 July 2024. https://www.anthropic.com/news/a-new-initiative-for-developing-third-party-model-evaluations.

Anthropic. 'Third-Party Testing as a Key Ingredient of AI Policy', 25 March 2024. https://www.anthropic.com/news/third-party-testing.

Anwar, Usman, Abulhair Saparov, Javier Rando, Daniel Paleka, Miles Turpin, Peter Hase, Ekdeep Singh Lubana, et al. 'Foundational Challenges in Assuring Alignment and Safety of Large Language Models'. arXiv, 2024. https://doi.org/10.48550/ARXIV.2404.09932.

Barcelona Supercomputing Center. 'European Supercomputer MareNostrum 5 Starts up at BSC', 21 December 2023. https://www.bsc.es/news/bsc-news/european-supercomputer-marenostrum-5-starts-bsc.

Barcelona Supercomputing Center. 'Spanish Government Strengthens AI Capabilities of MareNostrum 5 Supercomputer, Managed by BSC', 14 May 2024. https://www.bsc.es/news/bsc-news/spanish-government-strengthens-ai-capabilities-marenostrum-5-supercomputer-managed-bsc.

Bau, David, Jun-Yan Zhu, Hendrik Strobelt, Bolei Zhou, Joshua B. Tenenbaum, William T. Freeman, and Antonio Torralba. 'GAN Dissection: Visualizing and Understanding Generative Adversarial Networks'. arXiv, 2018. https://doi.org/10.48550/ARXIV.1811.10597.

Bengio, Yoshua, Geoffrey Hinton, Andrew Yao, Dawn Song, Pieter Abbeel, Trevor Darrell, Yuval Noah Harari, et al. 'Managing Extreme AI Risks amid Rapid Progress'. *Science* 384, no. 6698 (24 May 2024): 842–45. https://doi.org/10.1126/science.adn0117.

Besiroglu, Tamay, Sage Andrus Bergerson, Amelia Michael, Lennart Heim, Xueyun Luo, and Neil Thompson. 'The Compute Divide in Machine Learning: A Threat to Academic Contribution and Scrutiny?' arXiv, 2024. https://doi.org/10.48550/ARXIV.2401.02452.

Blanchar, Clara. 'Sánchez Anuncia En El Mobile Que España Desarrollará Un Modelo de Lenguaje de IA En Castellano y Las Lenguas Cooficiales'. *El País*, 25 February 2024. https://elpais.com/economia/2024-02-25/sanchez-anuncia-en-el-mobile-que-espana-desarrollara-un-modelo-de-lenguaje-de-ia-en-castellano.html#.

'BMBF-Aktionsplan Künstliche Intelligenz: Neue Herausforderungen Chancenorientiert Angehen'. Bundesministerium für Bildung und Forschung, 2023.

Boyd, Eric. 'Microsoft and Mistral AI Announce New Partnership to Accelerate AI Innovation and Introduce Mistral Large First on Azure'. *Microsoft* (blog), 26 February 2024. https://azure.microsoft.com/en-us/blog/microsoft-and-mistral-ai-announce-new-partnership-to-accelerate-ai-innovation-and-introduce-mistral-large-first-on-azure/.

Buolamwini, Joy, and Timnit Gebru. 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification'. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, edited by Sorelle A. Friedler and Christo Wilson, 81:77–91. Proceedings of Machine Learning Research. PMLR, 2018. https://proceedings.mlr.press/v81/buolamwini18a.html.

Carlini, Nicholas, Florian Tramer, Eric Wallace, Matthew Jagielski, Ariel Herbert-Voss, Katherine Lee, Adam Roberts, et al. 'Extracting Training Data from Large Language Models'. arXiv, 2020. https://doi.org/10.48550/ARXIV.2012.07805.

Casper, Stephen, Carson Ezell, Charlotte Siegmann, Noam Kolt, Taylor Lynn Curtis, Benjamin

Bucknall, Andreas Haupt, et al. 'Black-Box Access Is Insufficient for Rigorous AI Audits', 2024. https://doi.org/10.48550/ARXIV.2401.14446.
Center for AI Safety. 'CAIS Compute Cluster', n.d. https://www.safe.ai/work/compute-cluster.
'Communication on Boosting Startups and Innovation in Trustworthy Artificial Intelligence'. Brussels: European Commission, 24 January 2024.
Council of the European Union. 'AI: Council Reaches Political Agreement on the Use of Super-Computing for AI Development', 23 May 2024. https://www.consilium.europa.eu/en/press/press-releases/2024/05/23/ai-council-reaches-political-agreement-on-the-use-of-super-computing-for-ai-development/.
Council Regulation (EU) 2024/1732 of 17 June 2024 amending Regulation (EU) 2021/1173 as regards a EuroHPC initiative for start-ups in order to boost European leadership in trustworthy artificial intelligence, ST/10109/2024/INIT § (2024). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32024R1732.
'Decision of the Governing Board of the EuroHPC Joint Undertaking No 41/2024 Amending the Joint Undertaking's Work Programme and Budget for the Year 2024 (Amendment No 3)', 2024.
Doshi-Velez, Finale, and Been Kim. 'Towards A Rigorous Science of Interpretable Machine Learning'. arXiv, 2017. https://doi.org/10.48550/ARXIV.1702.08608.
Dwork, Cynthia, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. 'Fairness through Awareness'. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, 214–26. Cambridge Massachusetts: ACM, 2012. https://doi.org/10.1145/2090236.2090255.
Dwork, Cynthia, Frank McSherry, Kobbi Nissim, and Adam Smith. 'Calibrating Noise to Sensitivity in Private Data Analysis'. In *Theory of Cryptography*, edited by Shai Halevi and Tal Rabin, 3876:265–84. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. https://doi.org/10.1007/11681878_14.
Egan, Janet, and Lennart Heim. 'Oversight for Frontier AI through a Know-Your-Customer Scheme for Compute Providers'. arXiv, 2023. https://doi.org/10.48550/ARXIV.2310.13625.
'Estrategia de Inteligencia Artificial 2024'. Ministerio para la Transformación Digital y de la Función Pública, 2024.
EuroHPC Joint Undertaking. 'AI and HPC: Enhancing Research Possibilities', 12 September 2023. https://eurohpc-ju.europa.eu/ai-and-hpc-enhancing-research-possibilities-2023-09-12_en.
EuroHPC Joint Undertaking. 'Discover EuroHPC JU', n.d. https://eurohpc-ju.europa.eu/about/discover-eurohpc-ju_en.
EuroHPC Joint Undertaking. 'Our Supercomputers', n.d. https://eurohpc-ju.europa.eu/supercomputers/our-supercomputers_en.
EuroHPC Joint Undertaking. 'The "AI Factories" Amendment to the EuroHPC JU Regulation Enters Into Force', 9 July 2024. https://eurohpc-ju.europa.eu/ai-factories-amendment-eurohpc-ju-regulation-enters-force-2024-07-09_en.
EuroHPC Joint Undertaking. 'The EuroHPC Joint Undertaking Amends Work Programme to Incorporate the New AI Factory Pillar', 26 July 2024. https://eurohpc-ju.europa.eu/eurohpc-joint-undertaking-amends-work-programme-incorporate-new-ai-factory-pillar-2024-07-26_en.
EuroHPC Joint Undertaking. 'The United Kingdom Joins EuroHPC Joint Undertaking', 12 May 2024. https://eurohpc-ju.europa.eu/united-kingdom-joins-eurohpc-joint-undertaking-2024-05-12_en.
'EuroHPC JU Call for Proposals for AI and Data-Intensive Applications Access Mode - Full Call Details'. EuroHPC Joint Undertaking, n.d.
European AI Startup Landscape. 'The European AI Startup Landscape', n.d. https://www.ai-startups-europe.eu/.
European Commission. '2023 State of the Union Address by President von Der Leyen', 13 September 2023. https://ec.europa.eu/commission/presscorner/detail/en/speech_23_4426.
European Commission. 'AI Factories', n.d. https://digital-strategy.ec.europa.eu/en/policies/ai-factories.
European Commission. 'Commission Launches AI Innovation Package to Support Artificial

Intelligence Startups and SMEs', 24 January 2024. https://ec.europa.eu/commission/presscorner/detail/en/ip_24_383.
European Commission. 'Commission Opens Access to EU Supercomputers to Speed up Artificial Intelligence Development', 16 November 2023. https://ec.europa.eu/commission/presscorner/detail/en/IP_23_5739.
European Commission. 'European AI Office', n.d. https://digital-strategy.ec.europa.eu/en/policies/ai-office.
European Commission. Proposal for a Council Regulation amending Regulation (EU) 2021/1173 as regards an EuroHPC initiative for start-ups to boost European leadership in Trustworthy Artificial Intelligence, 2024/0016 (CNS) § (2024).
European Commission. 'Remarks of President von Der Leyen at the Bletchley Park AI Safety Summit', 2 November 2023. https://ec.europa.eu/commission/presscorner/detail/en/speech_23_5502.
European Commission, Directorate-General for Research and Innovation. *Successful and Timely Uptake of Artificial Intelligence in Science in the EU*. Publications Office of the European Union, 2024. https://doi.org/doi/10.2777/46863.
European Innovation Council. 'EIC Accelerator', n.d. https://eic.ec.europa.eu/eic-funding-opportunities/eic-accelerator_en.
European Union. 'InvestEU Programme', n.d. https://investeu.europa.eu/investeu-programme_en.
Gardizy, Anissa. 'Musk's xAI and Oracle End Talks on a Server Deal'. *The Information* (blog), 9 July 2024. https://www.theinformation.com/articles/musks-xai-and-oracle-end-talks-on-a-server-deal.
Gelles, Rebecca, Veronica Kinoshita, Micah Musser, and James Dunham. 'Resource Democratization: Is Compute the Binding Constraint on AI Research?' *Proceedings of the AAAI Conference on Artificial Intelligence* 38, no. 18 (24 March 2024): 19840–48. https://doi.org/10.1609/aaai.v38i18.29959.
Goodfellow, Ian J., Jonathon Shlens, and Christian Szegedy. 'Explaining and Harnessing Adversarial Examples'. arXiv, 2014. https://doi.org/10.48550/ARXIV.1412.6572.
Hardt, Moritz, Eric Price, Eric Price, and Nati Srebro. 'Equality of Opportunity in Supervised Learning'. In *Advances in Neural Information Processing Systems*, edited by D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, Vol. 29. Curran Associates, Inc., 2016. https://proceedings.neurips.cc/paper_files/paper/2016/file/9d2682367c3935defcb1f9e247a97c0d-Paper.pdf.
Heim, Lennart. 'A Trusted AI Compute Cluster for AI Verification and Evaluation', 31 March 2024. https://blog.heim.xyz/a-trusted-ai-compute-cluster/.
Hendrycks, Dan, Nicholas Carlini, John Schulman, and Jacob Steinhardt. 'Unsolved Problems in ML Safety'. arXiv, 2021. https://doi.org/10.48550/ARXIV.2109.13916.
Ho, Lewis, Joslyn Barnhart, Robert Trager, Yoshua Bengio, Miles Brundage, Allison Carnegie, Rumman Chowdhury, et al. 'International Institutions for Advanced AI'. arXiv, 2023. https://doi.org/10.48550/ARXIV.2307.04699.
Hotard, Justin. 'Sovereign and Trustworthy AI Is at the Core of Hewlett Packard Enterprise's Investment in Aleph Alpha'. *Hewlett Packard Enterprise* (blog), 6 November 2023. https://www.hpe.com/us/en/newsroom/blog-post/2023/11/sovereign-and-trustworthy-ai-is-at-the-core-of-hewlett-packard-enterprises-investment-in-aleph-alpha.html.
Hutchinson, Philip, Frauke Goll, and Christoph Muegge. 'Generative AI in the European Startup Landscape 2024'. appliedAI Institute for Europe, 2024.
'IA: Notre Ambition Pour La France'. Commission de l'Intelligence Artificielle, 2024.
Intelligence Artificielle: Les Annonces d'Emmanuel Macron à VivaTech, 2023. https://www.youtube.com/watch?v=Gdf4mTfZitY.
Jones, Elliot. 'A "CERN for AI" – What Might an International AI Research Organization Address?' In *Artificial Intelligence and the Challenge for Global Governance*. Chatham House, 2024.
Juijn, Daan, Bálint Pataki, Alex Petropoulos, and Max Reddel. 'A CERN for AI: The EU's Seat at the Table'. International Centre for Future Generations, 2024.
Klinger, Joel, Juan Mateos-Garcia, and Konstantinos Stathoulopoulos. 'A Narrowing of AI Research?' arXiv, 2020. https://doi.org/10.48550/ARXIV.2009.10385.
Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E Hinton. 'ImageNet Classification with Deep

Convolutional Neural Networks'. In *Advances in Neural Information Processing Systems*, edited by F. Pereira, C. J. Burges, L. Bottou, and K. Q. Weinberger, Vol. 25. Curran Associates, Inc., 2012. https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

Lee, Kevin, Adi Gangidi, and Mathew Oldham. 'Building Meta's GenAI Infrastructure'. *Engineering at Meta* (blog), 12 March 2024. https://engineering.fb.com/2024/03/12/data-center-engineering/building-metas-genai-infrastructure/.

Leyen, Ursula von der. 'Europe's Choice: Political Guidelines for the next European Commission'. European Commission, 18 July 2024.

Li, Zihao, Yucheng Shi, Zirui Liu, Fan Yang, Ali Payani, Ninghao Liu, and Mengnan Du. 'Quantifying Multilingual Performance of Large Language Models Across Languages'. arXiv, 2024. https://doi.org/10.48550/ARXIV.2404.11553.

LUMI. 'Accelerating Materials Discovery and Design with Machine Learning', 12 March 2024. https://www.lumi-supercomputer.eu/accelerating-materials-discovery-and-design-with-machine-learning/.

LUMI. 'DigiFarm Revolutionises Agriculture with Artificial Intelligence', 28 May 2024. https://www.lumi-supercomputer.eu/digifarm-revolutionises-agriculture-with-artificial-intelligence/.

LUMI. 'LUMI Powers AI Assistant Development for Privacy-Preserving Machine Learning Models', 18 June 2024. https://www.lumi-supercomputer.eu/ai_assistant_development/.

LUMI. 'New Finnish LUMI Projects Announced: AI Innovations in Healthcare and Beyond', 31 May 2024. https://www.lumi-supercomputer.eu/new-finnish-lumi-projects-announced-ai-innovations-in-healthcare-and-beyond/.

LUMI. 'One of the World's Mightiest Supercomputers, LUMI, Will Lift European Research and Competitiveness to a New Level and Promotes Green Transition', 21 October 2020. https://www.lumi-supercomputer.eu/lumi-one-of-the-worlds-mightiest-supercomputers/.

LUMI. 'Solving Intricate Research Challenges with LUMI', 1 April 2022. https://lumi-supercomputer.eu/solving-intricate-research-challenges-with-lumi/.

Madry, Aleksander, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. 'Towards Deep Learning Models Resistant to Adversarial Attacks'. arXiv, 2017. https://doi.org/10.48550/ARXIV.1706.06083.

Marcus, Gary. 'Artificial Intelligence Is Stuck. Here's How to Move It Forward.' *The New York Times*, 29 July 2017. https://www.nytimes.com/2017/07/29/opinion/sunday/artificial-intelligence-is-stuck-heres-how-to-move-it-forward.html.

Miller, Kyle, and Rebecca Gelles. 'The NAIRR Pilot: Estimating Compute'. *Center for Security and Emerging Technology* (blog), 8 May 2024. https://cset.georgetown.edu/article/the-nairr-pilot-estimating-compute/.

'Moonshot in Artificial Intelligence: Trustworthy, Multicultural Generative AI Systems for Safe Physical Interaction with the Real World'. CLAIRE & euRobotics, 7 November 2023.

Muller, Gerrit. '1.1.1 Industry and Academia: Why Practitioners and Researchers Are Disconnected'. *INCOSE International Symposium* 15, no. 1 (July 2005): 1–9. https://doi.org/10.1002/j.2334-5837.2005.tb00648.x.

Musk, Elon. 'Memphis Supercluster', 22 July 2024. https://x.com/elonmusk/status/1815325410667749760.

National Artificial Intelligence Research Resource Pilot. 'About NAIRR Pilot', n.d. https://nairrpilot.org/about.

'National Deep Inference Facility for Very Large Language Models (NDIF): Project Proposal to the NSF', 2023.

Nevo, Sella, Dan Lahav, Yogev Bar-On, Henry Alexander Bradley, and Jeff Alstott. 'Securing AI Model Weights: Preventing Theft and Misuse of Frontier Models'. RAND, 2024.

Nielsen, Michael, and Kanjun Qiu. 'A Vision of Metascience: An Engine of Improvement for the Social Processes of Science'. *The Science++ Project* (blog), n.d. https://scienceplusplus.org/metascience/.

NVIDIA. 'NVIDIA GB200 NVL72', n.d.

https://www.nvidia.com/en-us/data-center/gb200-nvl72/.
NVIDIA. 'NVIDIA Grace Hopper Superchip Powers JUPITER, Defining a New Class of Supercomputers to Propel AI for Scientific Discovery', 13 November 2023. https://nvidianews.nvidia.com/news/nvidia-grace-hopper-superchip-powers-jupiter-defining-a-new-class-of-supercomputers-to-propel-ai-for-scientific-discovery.
Open Philanthropy. 'Request for Proposals: Benchmarking LLM Agents on Consequential Real-World Tasks', n.d. https://www.openphilanthropy.org/rfp-llm-benchmarks/.
Parker, George. 'Rishi Sunak to Lobby Joe Biden for UK "Leadership" Role in AI Development'. *Financial Times*, 5 June 2023. https://www.ft.com/content/7c30ea28-2895-44c2-9a2d-c31ea7fa27e7.
'Preserving Endangered Languages with LUMI: Machine Translation for Rare Finno-Ugric Languages', 27 June 2024. https://www.lumi-supercomputer.eu/preserving-endangered-languages/.
Räuker, Tilman, Anson Ho, Stephen Casper, and Dylan Hadfield-Menell. 'Toward Transparent AI: A Survey on Interpreting the Inner Structures of Deep Neural Networks'. arXiv, 2022. https://doi.org/10.48550/ARXIV.2207.13243.
Renda, Andrea. 'Towards a European Large-Scale Initiative on Artificial Intelligence: What Are the Options?' Centre for European Policy Studies, 2024.
Reuel, Anka, Ben Bucknall, Stephen Casper, Tim Fist, Lisa Soder, Onni Aarne, Lewis Hammond, et al. 'Open Problems in Technical AI Governance'. arXiv, 2024. https://doi.org/10.48550/ARXIV.2407.14981.
Ribeiro, Marco Tulio, Sameer Singh, and Carlos Guestrin. '"Why Should I Trust You?": Explaining the Predictions of Any Classifier'. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–44. San Francisco California USA: ACM, 2016. https://doi.org/10.1145/2939672.2939778.
Rolnick, David, Priya L. Donti, Lynn H. Kaack, Kelly Kochanski, Alexandre Lacoste, Kris Sankaran, Andrew Slavin Ross, et al. 'Tackling Climate Change with Machine Learning'. arXiv, 2019. https://doi.org/10.48550/ARXIV.1906.05433.
Rudner, Tim G. J., and Helen Toner. 'Key Concepts in AI Safety: An Overview'. Center for Security and Emerging Technology, 2021. https://cset.georgetown.edu/publication/key-concepts-in-ai-safety-an-overview/.
Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. 'Learning Representations by Back-Propagating Errors'. *Nature* 323, no. 6088 (October 1986): 533–36. https://doi.org/10.1038/323533a0.
Sandbrink, Jonas, Hamish Hobbs, Jacob Swett, Allan Dafoe, and Anders Sandberg. 'Differential Technology Development: A Responsible Innovation Principle for Navigating Technology Risks'. *SSRN Electronic Journal*, 2022. https://doi.org/10.2139/ssrn.4213670.
Sastry, Girish, Lennart Heim, Haydn Belfield, Markus Anderljung, Miles Brundage, Julian Hazell, Cullen O'Keefe, et al. 'Computing Power and the Governance of Artificial Intelligence'. arXiv, 2024. https://doi.org/10.48550/ARXIV.2402.08797.
Schreiner, Maximilian. 'OpenAI Competitor Aleph Alpha Launches Europe's Fastest Commercial AI Data Center'. *Decoder* (blog), 16 September 2022. https://the-decoder.com/openai-competitor-launches-europes-fastest-commercial-ai-data-center/.
Sevilla, Jaime, Lennart Heim, Anson Ho, Tamay Besiroglu, Marius Hobbhahn, and Pablo Villalobos. 'Compute Trends Across Three Eras of Machine Learning', 2022. https://doi.org/10.48550/ARXIV.2202.05924.
Sharma, Ishan, Caleb Watney, Dan Correa, and Tony Mills. 'Piloting and Evaluating NSF Science Lottery Grants: A Roadmap to Improving Research Funding Efficiencies and Proposal Diversity'. *IFP* (blog), 2 February 2022. https://ifp.org/piloting-and-evaluating-nsf-science-lottery-grants/.
Shevlane, Toby, Sebastian Farquhar, Ben Garfinkel, Mary Phuong, Jess Whittlestone, Jade Leung, Daniel Kokotajlo, et al. 'Model Evaluation for Extreme Risks'. arXiv, 22 September 2023. http://arxiv.org/abs/2305.15324.
Shokri, Reza, Marco Stronati, Congzheng Song, and Vitaly Shmatikov. 'Membership Inference Attacks against Machine Learning Models'. arXiv, 2016. https://doi.org/10.48550/ARXIV.1610.05820.
Silo AI. 'Europe's Open Language Model Poro: A Milestone for European AI and Low-Resource Languages', 20 February 2024.

https://www.silo.ai/blog/europes-open-language-model-poro-a-milestone-for-european-ai-and-low-resource-languages.
Stewart, Ashley. 'Microsoft Has a Target to Amass 1.8 Million AI Chips by the End of the Year, Internal Document Shows'. *Business Insider* (blog), 18 April 2024. https://www.businessinsider.com/microsoft-gpu-targets-1-8-million-ai-chips-this-year-2024-4.
Taori, Rohan, Ishaan Gulrajani, Tianyi Zhang, Yann Dubois, Xuechen Li, Carlos Guestrin, Percy Liang, and Tatsunori B. Hashimoto. 'Alpaca: A Strong, Replicable Instruction-Following Model'. *Stanford Center for Research on Foundation Models* (blog), 2023. https://crfm.stanford.edu/2023/03/13/alpaca.html.
Trueman, Charlotte. 'The Fastest AI Supercomputer in the UK Is Now Operational'. *Datacenter Dynamics*, 26 February 2024. https://www.datacenterdynamics.com/en/news/fastest-ai-supercomputer-in-the-uk-is-now-operational/.
Wagner, C. S., and J. Alexander. 'Evaluating Transformative Research Programmes: A Case Study of the NSF Small Grants for Exploratory Research Programme'. *Research Evaluation* 22, no. 3 (1 September 2013): 187–97. https://doi.org/10.1093/reseval/rvt006.
Weilan, Zhang. 'Beijing Launches AI Public Platform as Demand Mounts for Computing Power'. *Global Times*, 8 April 2024. https://www.globaltimes.cn/page/202404/1310215.shtml.
Zhang, Xiang, Senyu Li, Bradley Hauer, Ning Shi, and Grzegorz Kondrak. 'Don't Trust ChatGPT When Your Question Is Not in English: A Study of Multilingual Abilities and Types of LLMs'. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 7915–27. Singapore: Association for Computational Linguistics, 2023. https://doi.org/10.18653/v1/2023.emnlp-main.491.
Zou, Andy, Zifan Wang, Nicholas Carlini, Milad Nasr, J. Zico Kolter, and Matt Fredrikson. 'Universal and Transferable Adversarial Attacks on Aligned Language Models'. arXiv, 2023. https://doi.org/10.48550/ARXIV.2307.15043.

Acknowledgements

We thank Jaime Sevilla, Daan Juijn, Philip Hutchinson, Konstantin Pilz, and Siméon Campos for feedback on earlier versions of this draft. Feedback does not imply endorsement. All remaining errors are our own.

Observatorio de Riesgos Catastróficos Globales

OUR SOCIAL MEDIA

https://orcg.info

info@orcg.info

@RiesgosCatastroficosGlobales

Riesgos Catastróficos Globales

@RiesgosGlobales