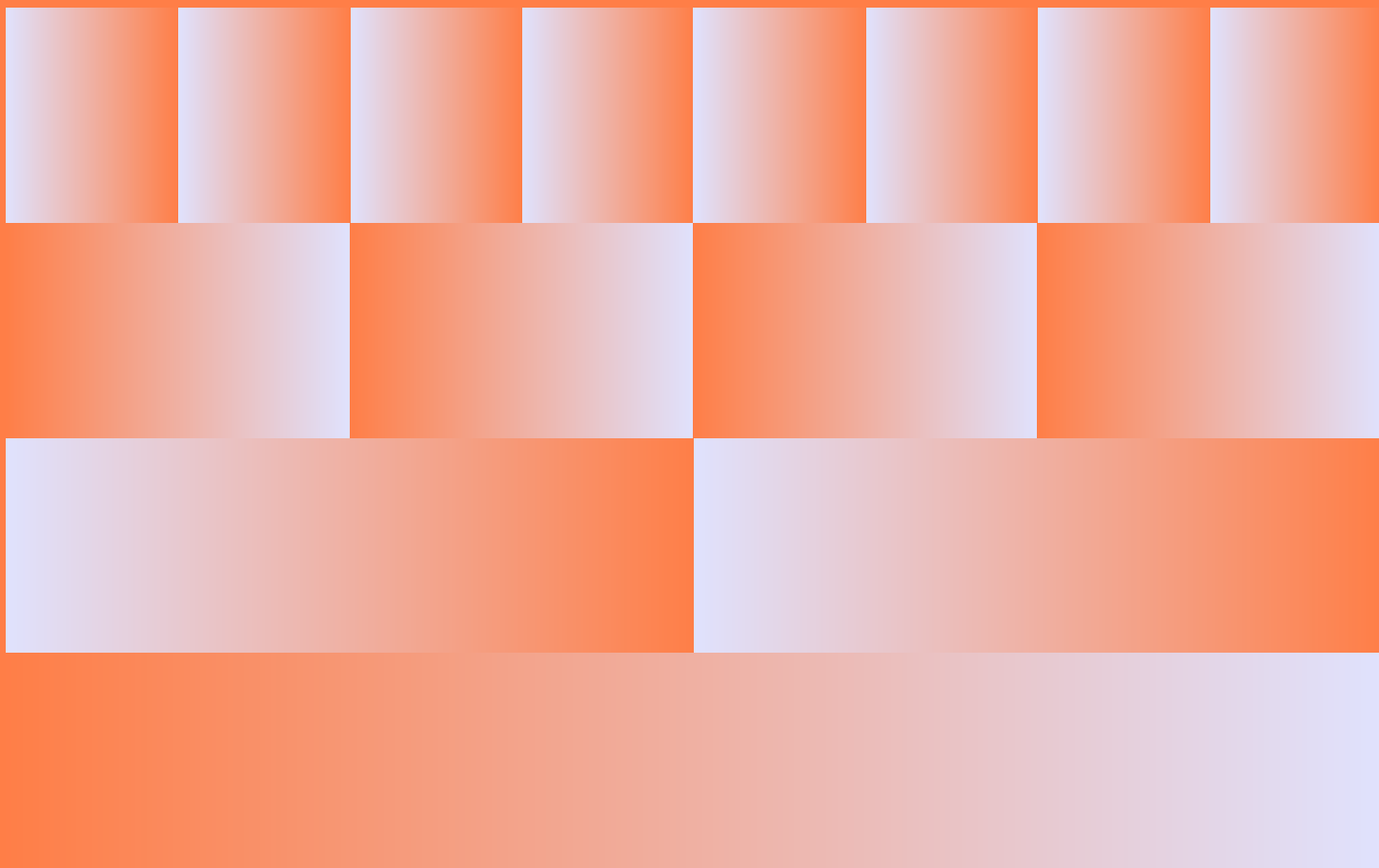




POLICY BRIEF

The European Union's AI Factories

Lessons for Public Investment in AI Infrastructure
in Europe

Nicole Lemke, Catherine Schneider

October 30, 2025

Table of Contents

1. Executive Summary	3
2. Introduction	4
3. AI Factories in Focus	6
3.1. Scale and Access	6
3.2. Consortia and Services	7
3.3. Ecosystem Integration	10
3.4. Learning from European AI factories	12
4. Conclusion	15

Executive Summary

This policy brief examines the EU's AI factory network, analysing the thirteen AI factories selected before October 2025 with regard to the scale of and access to their infrastructure, the composition of their consortia, and the innovation ecosystems they are located in. In our analysis, we found an inherent tension in this compute infrastructure, noting that although AI factories are suited to supporting research in training medium-sized AI models, the factories are not sufficient to boost commercial AI innovation across the EU at scale. We reveal several key features of the AI factories, deriving learnings for AI factories and future AI gigafactories, such as:

Partnership consortia are primarily composed of research institutions, rather than commercial actors.

Despite global AI-specific compute infrastructure being largely held by private sector actors, the AI factory ecosystem sees academia and research institutions dominating the composition of the partnership consortia. Only two of the thirteen factories examined have consortium without a research or academic partner, reflecting how well the factories are placed to support research and public innovation. However, a limited involvement of industry may hamper the adoption of AI by the startups and SMEs that the EU hopes to attract, failing to include the perspectives of crucial actors and industries.

Factories claim specialised focuses, but reality shows a generalist approach.

Although many regions hosting the factories expect to see regional industries benefit from increased compute capacity, nine of the thirteen factories examined in the paper focus on at least five sectors, a more generalist approach that rarely takes into account the strength of the surrounding ecosystem. Only one factory, HammerHAI, exclusively targets industries that reflect the strengths of Stuttgart's automotive industry. This reveals the AI factories aim primarily to provide generalist AI infrastructure, rather than deeply engage with regional strengths.

Location matters, and so does how the EU approaches strengthening them.

AI factories located near large AI talent pools or in strong innovative regions may be better suited to leverage their existing talent to increase the number of commercial users and researchers, while smaller, less-technologically adept regions might struggle to develop industrial proximity to compute. For this reason, it would benefit the EU to more deeply consider the ecosystem in which they seek to place AI factories, and co-locate gigafactories to strengthen promising ecosystems.

The structure of the AI factories has implications for the future gigafactories.

If the AI factories may see their impact on AI innovation limited by their engagement with the private sector, it is all the more important that the gigafactories, which will offer four times the compute capacity, consider the needs of commercial actors. Gigafactories must be more adapted to offer enterprise users the flexibility they need in compute access, if they are to meaningfully offer alternatives to the growing private compute infrastructure landscape. They also need to be placed more wisely, either in promising AI ecosystems or strategically leveraged to bridge gaps in compute for private enterprises.

In the global race to build AI capacity and incentivize data-driven commercial enterprises, the EU's investment in publicly subsidised compute infrastructure remains a big bet to drive AI innovation across Europe. This is one step in the right direction, but commercial innovation cannot emerge from this infrastructure alone; only robust understanding of commercial needs and strengthening tech hubs will truly lead to the AI landscape that the EU hopes to achieve.

For a complementary perspective on the EU's AI gigafactories initiative, see the policy brief by [Julia Hess](#) (interface) and [Felix Sieker](#) (Bertelsmann Stiftung): [Built for Purpose? Demand-Led Scenarios for Europe's AI Gigafactories](#).

Introduction

Earlier this year, the European Union (EU) announced its [plan](#) to position the EU as a global leader in AI by boosting technological development and strengthening global competitiveness. The establishment of AI factories and gigafactories is a central component of this plan. AI factories and gigafactories are increasing the EU's AI-specific compute infrastructure, by building new or upgrading existing supercomputers across the bloc. They are part of an effort to “foster an innovative European AI ecosystem.” This policy brief focuses on AI factories, which offer medium to large-scale AI-specific compute infrastructure as a structured and subsidized entry point to this infrastructure for researchers, SMEs, and startups, in addition to skill development, and sector-specific expertise. AI factories are therefore aimed at supporting both AI research and commercial innovation in Europe, balancing the needs of research entities with those of commercial actors.

While AI factory is a term often used to denote specialized compute infrastructure in general, in this policy brief, it refers to the specific term used by the EU to describe AI-optimized supercomputers — a type of computer that is higher performing than a standard computer, often handling computationally complex tasks — and their surrounding ecosystems made up of up to 25,000 AI-optimised chips. AI gigafactories refers to the term used by the EU for large-scale

supercomputer facilities with up to 100,000 AI-optimised chips specifically used to the development and training of next-generation AI models containing trillions of parameters.

In an announcement on 10 October the European Commission stated that six new AI factories will be joining the original cohort of thirteen selected in locations that either wish to develop an AI-optimised supercomputer, or those that wish to upgrade an existing EuroHPC supercomputer to increase AI capabilities. The factories are assembled under the umbrella of the European High Performance Computing Joint Undertaking (EuroHPC JU), a legal and funding entity for European supercomputing, creating a massive network of publicly funded supercomputers. Factories are located across all of Europe, from Barcelona, Spain to Sofia, Bulgaria and from Athens, Greece up to Kajaani, Finland.

Additionally, five AI gigafactories are planned to be selected by the end of this year, with the aim of providing alternatives to large scale, private compute infrastructure and allowing for the training of advanced models in Europe. With 76 expressions of interest submitted to host one of these gigafactories, it is clear that European cities hope to capitalize on this investment in compute infrastructure. In addition to AI factories and gigafactories, the EU is also working on the Cloud and AI Development Act, [aiming to triple its data centre capacity in the next 7 years](#).

Taken together, these investments represent the EU's response to a global data centre boom and massive investments in compute infrastructure for AI by other countries, such as the [\\$500 billion "Stargate" project](#) in the U.S., the UK's [£750 million investment](#) in a supercomputer in Edinburgh, or Singapore's commitment to its [national supercompute infrastructure](#). Just this year, U.S. tech companies have spent enough on data centres that ["the dollar value contributed to GDP growth by AI data center expenditure"](#) may have ["surpassed the total impact from all U.S. consumer spending."](#) With the considerable sum of [10 billion euros](#) already invested in AI factories, there are big hopes for their impact on European AI innovation ecosystems: new cutting-edge European AI models, increased collaboration on trustworthy AI, higher adoption by industry, and more equitable access to compute power for both research and commercial actors.

This policy brief analyses the thirteen AI factories selected before October 2025 — examining their specific characteristics including the scale of their infrastructure, structure of their consortia, and their relation to the regional innovation systems in which they are located. From this analysis, we derive lessons for AI factories and gigafactories examining the market fit of this infrastructure to drive private sector innovation. We argue that understanding the ecosystems in which the factories are located is a central facet of consideration for their monitoring and evaluation of how well they support the European Commission's goals. We conclude by making

recommendations for the selection of future AI gigafactories, exploring lessons to direct the [additional 20 billion](#) euro investment, and highlighting the potential limitations of the current approach to developing Europe's compute infrastructure landscape.

AI Factories in Focus

Scale and Access

Progress in the currently most popular state-of-the-art AI models lacks the large-scale compute infrastructure needed for training and deployment. To continue training ever larger models, AI increasingly relies on [powerful and specialized supercomputers](#). AI supercomputing is a form of high-performance computing (HPC), a field of computing that maximizes compute power and performance often measured in petaflops, i.e., to one thousand million million floating-point operations per second, or exaflops, i.e., one quintillion floating-point operations per second. Supercomputers run massive parallel workloads needed for training demanding AI models and running inference but also other compute-heavy tasks such as running simulations.

The AI factories mostly emerged from pre-existing supercomputers federated under the EuroHPC JU. The supercomputers at AI factory locations are generally highly performant: As of June 2025, three of them even make the top 10 of the [top 500 list of high-performance supercomputers](#) worldwide — Jupiter in Jülich (#4 with 930,00 peak petaflops), LUMI in Kajaani (#9 with 531,51 peak petaflops), and Leonardo in Bologna (#10 with 306.31 peak petaflops). However, AI supercomputers differ from other supercomputers in their hardware, which requires specific chips and networking infrastructure to support the most demanding AI workloads. Not all AI factories' hardware is AI-specific, so this list does not necessarily reflect their performance on AI workloads. Some supercomputers like those in Barcelona and Jülich have recently been upgraded with AI-optimized hardware.

Although the EU's AI factories host highly performing supercomputers, their AI-specific capacities are lower scale than many of leading supercomputers privately owned by leading AI labs and cloud providers. They do allow for training and deploying mid-sized AI models: [Jupiter in Jülich](#) now includes 24,000 NVIDIA GH200 Grace Hopper superchips, processors designed for HPC and large-scale AI, in its booster module. [Mare Nostrum 5 in Barcelona](#) currently includes 4,480 NVIDIA Hopper GPUs in its accelerated partition and is set to receive an additional general partition based on NVIDIA Grace CPUs. Furthermore, researchers in Switzerland recently trained a [70 billion parameter large language model](#) on the

[Alps supercomputer in Lugano](#) equipped with 10,752 NVIDIA Grace Hopper superchips. Similarly, the LUMI and Jupiter supercomputers have been [used to train TildeOpen LLM](#) with [more than 30 billion parameters](#). Yet, these supercomputers cannot match the sheer AI-specific capacity of privately operated AI supercomputers that are used to train large-scale commercial models with [hundreds of billions or more of parameters](#). xAI's AI supercomputer Colossus in Memphis, for example, is built with 200,000 AI-specialized chips, while other planned projects from AI behemoths often include more than that. **Thus, AI factories provide substantial compute capacity and some AI-optimized hardware, but their infrastructure at present remains geared toward mixed HPC workloads which includes AI but also other types of computational problems** such as simulations.

Access to AI factories can be local with users within the hosting lab or remote via dedicated networks or national access programs. [Access is granted at different scales and with different durations](#): Playground access which grants access to one partition of the hardware within 2 working days, fast lane which grants up to 50,000 GPU hours for a maximum of 3 months within 4 working days, and large-scale access which involves more than 50,000 GPU hours for 3, 6, or 12 months within 10 working days. AI startups, SMEs, and researchers benefit from free access. While these different access modes clearly aim to provide fast and flexible access, they cannot match the on-demand, elastic scaling of private cloud compute providers, where money can pay to bypass bottlenecks and capacity limits. The distinction remains: **AI factories are valuable in making subsidized medium-scale AI-specific high-performance compute publicly available for a larger user base, yet they cannot replicate the scale and elasticity of private AI-specific supercomputers or cloud providers.**

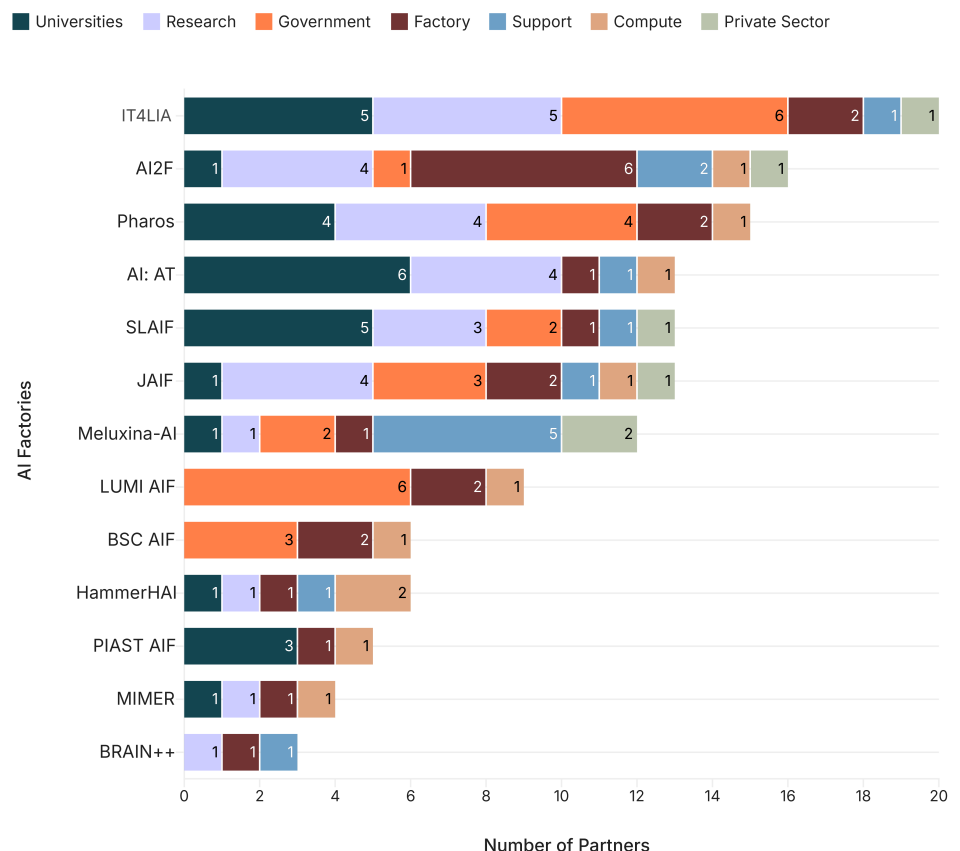
AI factories' mixed use and their history as large-scale research infrastructure suggest that **their main strength lies in building additional AI-specific compute capacity for public research** — an important step, not least after the success Swiss researchers had with training an open 70 billion parameter model on comparable public infrastructure. It is therefore important for policymakers to understand that AI factories' potential to support business use cases is likely limited, at least beyond certain use cases such as prototyping or fine-tuning certain models. While subsidized compute infrastructure is a good first step, additional measures are needed to support commercial AI innovation in Europe.

Consortia and Services

Globally, the vast majority of AI-specific compute is in the hands of only a few private actors. This includes commercial cloud infrastructure that is available to a broader user base but also fully privatized infrastructure in the hands of single

companies. By contrast, **AI factories are run by multi-organization consortia and provide publicly subsidized access to compute infrastructure which is free of charge for certain users such as research institutions, SMEs and startups.** AI factories' consortia are highly diverse, involving private actors, universities, research organizations, governments, compute providers, and other supporting organizations. Figure 1 shows all AI factories' consortia with the types of partner organizations involved. While the actual governance structures or individual organisational responsibilities are not yet fully clear for the AI factories, their consortia do reflect the types of actors convening to support compute ecosystems in various national contexts. It must be noted that the below graphic only captures the thirteen original AI factories, as the analysis was conducted prior to the announcement of the six new factories.

AI Factories: Organizational Composition and Strategic Orientation



Academic institutions, with universities and research institutions form the backbone of most AI factories' consortia. Ten AI factories have at least one university or one research partner. Only two factories, BSC AI in Spain and LUMI in Finland, do not have one of these partners. This strong representation is likely

explained by AI factories' history as large-scale research infrastructures. Moreover, academic and research institutions are key to running the technical infrastructure at the core of AI factories — and they are often the ones hosting it. Some AI factories have formal partnerships with other AI factories, creating stronger linkages between various nodes of European compute. The six new factories seem to reflect the same trend, with consortia members represented primarily by academic or research institutions. Taken together with the free access they offer for AI for science and collaborative projects; most AI factories are likely excellently placed to support AI research and public AI innovation in Europe.

But without further support from the private sector or intermediary organizations, AI factories may struggle to induce commercial usage: The long-term and open goals of traditional research do not necessarily align with the more short-term, applied, and commercial interests of private actors. MIMER in Sweden, Pharos in Greece, and PIAST in Poland are entirely led by research institutions, with universities and research organizations making up the majority of their partners. Similar concerns exist for consortia led by one or multiple government organizations, like BSC AI and LUMI. While these organizations may provide financial aid and credibility, they alone are not necessarily sufficient to connect the AI factory to the needs of startups and SMEs. This is also indicated by AI factories governed by research or government-led consortia, which may focus on capability building like training and education as well as knowledge brokering activities like connecting different types of actors and facilitating knowledge exchange, rather than market-oriented activities. While this does not mean that the factories cannot still serve commercial interests while promoting research and knowledge generation, it does reflect a tension between what factories are expected to generate against what they might be best positioned to produce.

Slightly more than half of the original thirteen AI factories have at least one consortium partner from the private sector or a partner whose role is to support private actors. This often includes organizations who consult and support startups and SMEs — likely crucial in helping them make the most of access to an AI factory or organizations representing the private sector, not private companies themselves who can fulfil a coordinating function between the AI factories and potential users from the private sector. The lack of involvement of private partners is notable across all AI factories: only three offer access to funder networks or connections to a broader grant ecosystem, while one offers funding for proof of concepts developed at the factory. This picture is further confirmed by the six newly nominated AI factories.

The diversity of consortium partners among AI factories points to a fundamental challenge they face in balancing the needs of researchers with those of SMEs and startups. Depending on their consortium partners, AI factories may have different

strengths in reaching these groups of users and engaging in partner ecosystems. Against this background, the diversity among AI factories could thus play out as a strength rather than a problem to be solved: Instead of trying to serve all potential users across all thirteen AI factories, they could capitalize on their individual strengths and form a strong network to ensure all the different users receive the service that suits them most. The addition of the six additional factories in Czechia, Lithuania, the Netherlands, Romania, Spain, and Poland may influence the diversity of actors involved in this HPC initiative, but also may further entrench the AI factories as tools best suited for researchers rather than the dominant industries in those countries. Nevertheless, the extensive variation in AI factories' consortia and the limited involvement of private actors also makes it clear that **they are not a sufficient solution toward commercially viable AI innovation in Europe at scale.**

Ecosystem Integration

Where AI factories are located matters as much as who operates them. A region's pre-existing capabilities shape its ability to develop and adopt complex technologies, and regions with robust ICT knowledge bases are better positioned to leverage AI. This can be seen in cities like Paris or Stuttgart, which boast robust pools of STEM talent that can aid the adoption rates of AI in industry. Adoption tends to cluster around industrial hot spots and within broader knowledge networks, rather than being evenly distributed. Consequently, an AI factory's location and geographic scope are likely to influence both its immediate ecosystem and, potentially, the broader European innovation landscape.

AI factories' geographic scopes are diverse, spanning regional, national, and international levels. Of the original thirteen factories, seven have highly centralized and regional consortia with three quarters of consortia partners being located in the same region the AI factory is. The majority of these factories stress the national investment and focus of building the AI factory, reflected in the consortia of partners that were assembled in the same region to ensure benefits were concentrated. Three AI factories have consortia with partners distributed nationally, rather than regionally concentrated. Three AI factories have consortia that include international partners, suggesting an international scope. BSC AIF and LUMI AIF have consortia made up entirely by international partners, reflecting their countries' supercomputing experience and capacity to form cross-border collaborations, while IT4LIA boasts multiple international partners, in addition to regional and national. While not extensively analysed, the new factories seem to mirror a concentration on national partners, with very few international partnerships in many of the emerging factory locations. **These variations indicate that AI factories' geographic scope shapes the networks they can build and the ecosystems they can influence, with many regional and national factories likely focusing on supporting local innovation systems even as pan-European integration is promoted.**



This is all the more important considering AI factories differ regarding the innovation capacities of the region they are located in. AI factories are mostly located in regions with strong innovation capacities. According to the EU Regional Innovation Scoreboard, five factories are situated in leading innovator regions, four in strong innovator regions, and four in moderate or emerging regions. About half of the factories are situated in urban centres, including four in capital regions and three in other metropolitan areas, while the remainder are in rural or hybrid regions. The more urban locations generally correspond to regions with higher

economic productivity. AI factories in regions with higher innovation and economic capacities stand a better chance at making an impact on the AI ecosystem (*see also infobox 1*). However, consortia partners may influence or support factories located in moderate or emerging innovation regions, helping to increase ecosystem capacity by infusing new expertise and networks into an AI factory.

Despite listing sectoral specializations, most AI factories are general-purpose compute infrastructures rather than focused on regional industries. Nine out of thirteen cover at least five sectors, with only HammerHAI in Stuttgart exclusively targeting engineering and manufacturing, reflecting local strengths like the automotive sector. A few others align with regional industries: LUMI with manufacturing in Kainuu, IT4LIA with agri-tech and agri-food in Emilia-Romagna, MIMER with gaming in Sweden, and SLAIF with upcycling in Maribor. This highlights the **generalist nature of most factories, with only select examples leveraging strong regional specialization**, underscoring that while factories may reflect regional industries their main purpose is providing generalist AI compute infrastructure.

AI2F and the Paris ecosystem

The French example shows how factories located in economically leading regions can capitalize on existing links with industry and are thus better positioned to support startups and SMEs with their infrastructure. This example is difficult to replicate by AI factories located in regions with less ideal starting conditions. AI2F is located in Île-de-France, a leading innovative and economically strong region in European comparison. Paris, located only about 40 km from the AI factories' location, is a European leader with a strong startup ecosystem with over 12.000 startups and 85% of France's R&D jobs. AI2F emphasizes this landscape, stating that it, "*capitalizes on a vivid ecosystem that created >1000 startups, 9 AI clusters, and significant research achievements*." Considering the greater potential of leading innovative and economic regions to take up AI innovation, AI factories are positioned to have a positive effect since innovation capacity and economic environment are already favourable. AI2F mentions the service of connecting their users to funding opportunities like access to VC and involves multiple supporting organizations that target start-ups in particular, taking advantage of its position in one of the strongest AI hubs in Europe.

Learning from European AI factories

AI factories represent a large investment in European compute infrastructure. The European Commission hopes that they will boost both AI research as well as commercial AI innovation. In this policy brief we explored all thirteen AI factories selected to date in depth to test this assumption. **Our main take-away is that their strength lies in providing European researchers with additional free and AI-specific compute infrastructure and services, building on already available HPC resources. In terms of their impact on commercial AI innovation, the new access pathways for**

startups and SMEs AI factories create are a positive first step, but more must be done to support commercial AI innovation in Europe.

Given the scale of the compute resources provided by AI factories, it is likely that commercial clients will use their services for prototyping and testing ideas for medium-sized models and fine-tuning off-the-shelf models. Some users may take advantage of additional beneficial offerings like training, networking, and consulting services. Yet, the central challenge that emerges is that these services will likely not be enough to boost commercial AI innovation at scale. AI gigafactories are supposed to address this issue, but are located at the extreme end of AI-specific compute, leaving a gap in AI-specific compute capacities for commercial use. **This leaves a critical gap between early-stage exploration and hyperscaled applications that both AI factories and gigafactories fail to address.** Below, we outline further learnings for AI factories, AI gigafactories, and the broader landscape of AI compute in Europe.

AI Factories

Our learnings from AI factories are threefold:

1. In their conceptualization as one-stop-shops for AI development and adoption, **AI factories are tasked with balancing the needs of researchers with those of commercial actors.** In some cases, like AI2F in France or Meluxina in Luxembourg, the services offered even include support with prototyping, testing, and finding funding opportunities. **This indicates a useful, but partial, role in supporting innovation: AI factories are well-positioned as public research infrastructures that provide targeted support for startups and SMEs, but they cannot continually stretch their offerings to fully substitute for private expertise and investment in AI and compute infrastructure at different scales.**
2. Innovation flourishes in environments that facilitate the exchange of knowledge. Thus, **AI factories located near existing AI talent pools may, for example, be able to leverage their strong concentrations of talent to grow the number of AI users within various commercial enterprises.** With only two AI factories planned in or close to major AI European AI talent hubs like Paris and Barcelona, the other locations may pose a broader issue to the adoption of AI by SMEs and startups that may otherwise have little to no exposure through proximity to those developing AI systems. The approach may not lead to the promised technological innovation everywhere: [regions](#) without a strong base of expertise may face challenges in growing their AI capabilities and intra- and international innovation networks. Not having experts and business leaders who can tap into the AI factory offerings could lead to underutilisation, resulting in exacerbating AI knowledge concentration around areas with existing expertise rather than building out European capabilities and usage more broadly. Smaller existing talent pools will limit the network effects leading to more widespread adoption and accessing of the factories.
3. **AI factories' impact on commercial innovation will likely depend to a considerable extent on the innovation ecosystem they're embedded in.** AI innovation is more likely in innovation ecosystems with existing industry and knowledge bases in the digital sector. Moreover, sectoral AI innovation —for example in health or industrial engineering — is more likely to emerge in areas where these industries are already strong. Despite this, most AI factories provide a fairly uniform set of services without great specialization geared to the needs of their ecosystems, and many are located in

regions that are emerging or moderate innovators according to the EU. AI factories could benefit from instead refocusing and prioritising meaningful specialisation for regional and national industries, matching both in sectoral focus and regional needs. A generalist approach is likely unsatisfactory to actually drive forward a European innovation network, and more care must be given to how AI factories fit into existing industries, especially as future AI factories and gigafactories are selected.

AI Gigafactories and the European AI Compute Landscape

These learnings have important consequences for the selection of AI gigafactories scheduled for the end of this year as well as future policies regarding AI compute infrastructure:

1. **Policymakers would benefit from considering the ecosystem AI compute infrastructure is set to serve.** With four times the compute power of AI factories, the gigafactories will target a smaller and more specific set of users: companies capable of designing and training large-scale AI models. Yet, considering the innovative ecosystems and geographic location will influence the users requesting access to both AI factories and gigafactories, and thus more attention should be paid to the overall landscape of commercial interest, location and actors interested in developing and applying AI in their endeavours.
2. **This means a promising practice for the EU could be co-locating AI gigafactories with existing AI factories, further strengthening promising ecosystems, or at the very least ensuring strong communication between AI factories and gigafactories as well as commercial hyperscalers that can fill the gaps.** This is especially crucial as gigafactories seek to develop cross-sectoral frontier models; failure to ensure that different sectors are represented and can shape the gigafactories, especially the resources and data offered for training, may limit their capacity to develop the breakthrough models that the EU is betting on.
3. When it comes to future policies for AI compute infrastructure, **policymakers must further consider the applied needs of commercial AI innovation, including the preferences of private sector actors.** For some, it may be more desirable to pay for private compute, rather than apply and wait for public server space. Many of the leading AI labs are constructing private infrastructure to meet compute needs, while other enterprises may pay compute providers in order to get access quickly, without needing to make large up-front commitments to infrastructure. Moreover, as businesses grow, they may need access to alternative forms of infrastructure to continue that momentum, AI factories and gigafactories would be well advised to develop more flexible approaches for enterprise users, including off-ramps to hyperscalers for successful models. Moreover, application for access to AI factories is measured in months and capped at twelve months, depending on the duration needs of applicants. This may put users with less flexibility at a disadvantage, should their expectations of compute time or support services differ from projections. This may force enterprises to reapply and wait for access approval, which may be less appealing when compared to the continuous access of private infrastructure.

Taken together, this policy brief provides insights to shape both the AI factories, and the planned AI gigafactories, which require that policymakers understand both the strengths and limitations of compute infrastructure in Europe. It remains to be seen how industry will approach and integrate new compute capacity into their operations. AI factories may try to induce commercial demand by offering services

from ideation, model development, and business case support, but industrial partners will need to develop their own unique value propositions of leveraging AI. Publicly subsidised compute infrastructure will only be able to provide support to make AI adoption acceptable from a financial and risk avoidance perspective if commercial actors actually decide and see value in leveraging AI, which is currently limited to a [13.48%](#) of all European businesses.

Despite their many offerings, AI factories cannot be the substitute for private sector investment and innovation. Various enterprises must consider and develop strategies that seek to leverage AI for the benefit of their business, rather than simply trying to capture some of the technological hype. Supporting SMEs and startups from early-stage ideation up through deployment requires extensive resources that companies themselves may be less willing to match. Committing to providing such extensive services may ultimately force the factories to provide more without establishing industry buy-in first, limiting their impact on the industries that historically have driven European economic growth. Ultimately, AI factories are best positioned to help private actors test the value of their ideas and if they are worth investing in their own AI infrastructure or using private on-demand services. **Public compute infrastructure cannot act as the panacea to inspire industry adoption and development, nor can it incentivise private sector actors to recognise legitimate value in AI.**

Conclusion

The European Commission has presented the AI factories as a critical tool in “strengthening the European AI innovation ecosystem.” However, **looking beyond official declarations and instead at the existing ecosystem, the picture is more mixed — showing both the potential and limitations of publicly subsidized compute infrastructure as a driver of large-scale commercial innovations.** The competition surrounding the selection process as well the public messaging around AI factories located in regions lagging in terms of innovative capacity suggest that regions have high hopes for the positive effects an AI factory might have for their regional innovation system directly, not only the European innovation ecosystem in general. Yet, the compute resources AI factories host are limited and will likely have to be distributed between regional, national, and international researchers, start-ups, SMEs, and industry interested in using them. This means that even as cities and regions jockey to be selected to host future AI factories or gigafactories, the return on investment may be limited both in terms of economic productivity and impact for local actors.

How different AI factories will negotiate between these various users’ needs remains to be seen. At the same time, the planned gigafactories offer an opportunity to learn

from the factories, but are equally unlikely to be a sufficient solution to the compute needs of commercial actors. The AI factories and gigafactories are yet to demonstrate their true value in supporting the development of commercial AI innovation in Europe. **The provision of publicly subsidized compute infrastructure is a good and necessary step toward supporting AI innovation in Europe. But alone it is not enough. Innovation doesn't arise from public research alone, for which AI factories are indeed well positioned — investment is also needed in commercial innovation, and this is where AI factories are less well equipped.** Only comprehensive examination of AI applications, critical commercial needs, and developing existing ecosystems will lead to the breakthrough innovations that the EU wants to see as a return on their investment.

Authors

Nicole Lemke

Senior Policy Reseacher AI

nlemke@interface-eu.org

Catherine Schneider

Policy Researcher AI

cschneider@interface-eu.org

Imprint

interface – Tech analysis and policy ideas for Europe
(formerly Stiftung Neue Verantwortung)

W www.interface-eu.org

E info@interface-eu.org

T +49 (0) 30 81 45 03 78 80

F +49 (0) 30 81 45 03 78 97

interface – Tech analysis and policy ideas for Europe e.V.
c/o Publix
Hermannstraße 90
D-12051 Berlin

This paper is published under Creative Commons License (CC BY-SA). This allows for copying, publishing, citing and translating the contents of the paper, as long as interface is named and all resulting publications are also published under the license “CC BY-SA”. Please refer to <http://creativecommons.org/licenses/by-sa/4.0/> for further information on the license and its terms and conditions.

Design by Make Studio

www.make.studio

Code by Convoy

www.convoyinteractive.com