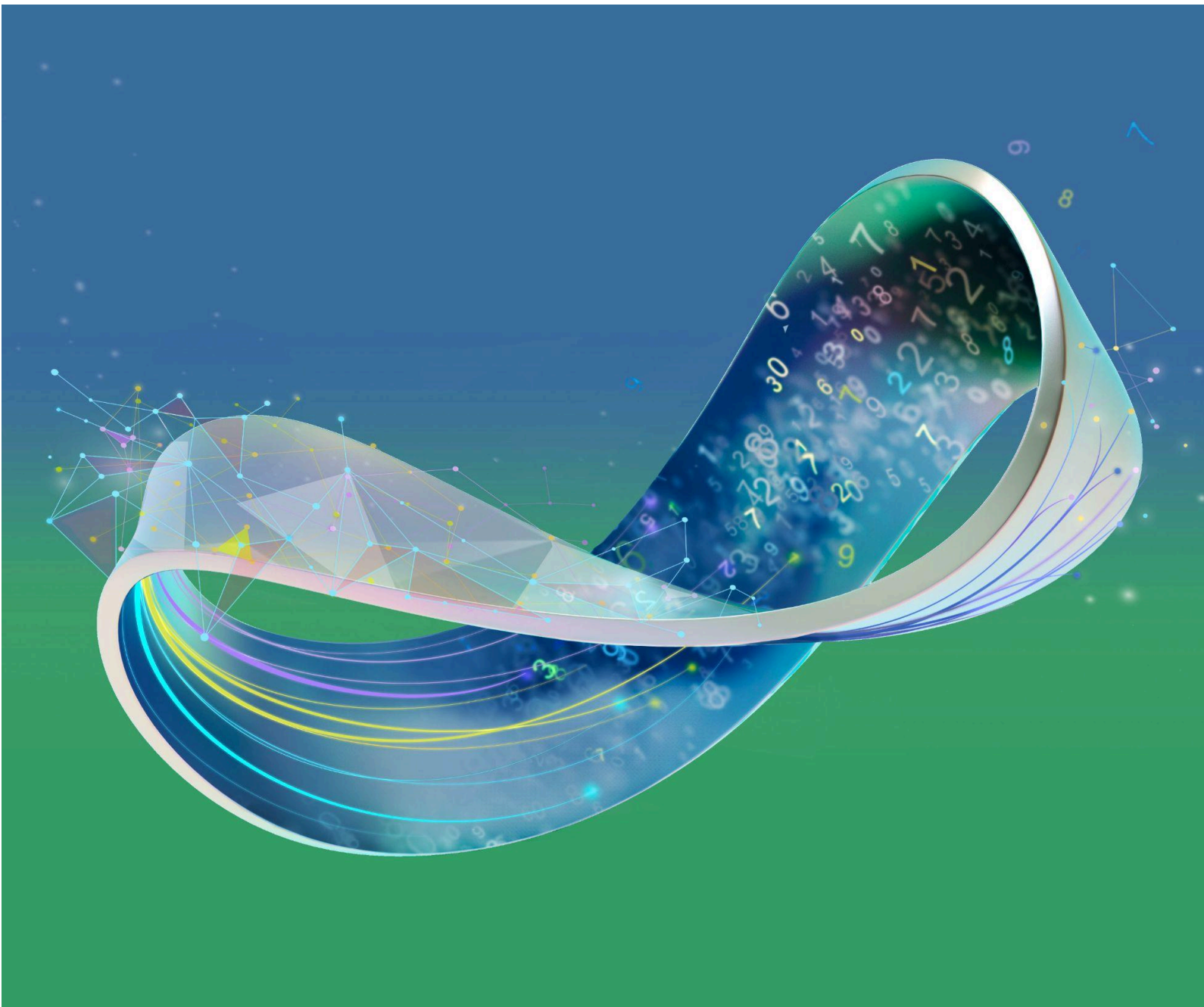
EMBL

Apply AI Strategy

Call For Evidence | European Commission

June 2025



Artificial Intelligence (AI) is transforming how science and industry intersect, offering vast potential to accelerate discovery, improve productivity, and drive innovation across sectors. ***Nowhere is this more evident than in the life sciences****, where data-intensive research increasingly intersects with industrial applications, from precision medicine and biomanufacturing to diagnostics and agriculture. These domains are foundational for public health and planetary sustainability as well as for Europe's economic resilience and industrial competitiveness. The* ***life sciences*** *with their wealth of high-impact, societally relevant challenges, and an abundance of open, high-quality data, represent a* ***high-potential domain for industrial AI adoption****.*

About EMBL

The European Molecular Biology Laboratory (**EMBL**), Europe's intergovernmental organisation for molecular biology, and the world's largest repository for biomolecular data (**EMBL-EBI**), has long championed the development and integration of emerging technologies to advance innovation and serve public good.

A prominent example in the field of AI[1] is the public-private partnership between Google DeepMind and EMBL's European Bioinformatics Institute (EMBL-EBI), which led to the development of the Nobel prize-winning AlphaFold Protein Structure Database[2], a landmark that illustrates the industrial potential of open data and cross-sectoral collaboration.

EMBL welcomes the European Commission's ambition to scale AI deployment across strategic sectors through the *Apply AI Strategy*. This response highlights how **public research organisations and life science infrastructures can help translate Europe's scientific excellence into AI-powered industrial capabilities and advance Europe's competitiveness and digital sovereignty**.

1. Invest in AI-Ready Data Infrastructure for Industrial and Scientific Innovation

Europe's life science sector already generates vast amounts of high-quality, open data through publicly funded research infrastructures. These data resources, such as the EMBL-EBI-hosted Protein Data Bank in Europe or the European Nucleotide Archive, are used extensively by SMEs, pharmaceutical companies, and digital health start-ups across Europe. However, many of these resources remain fragile[3] due to short-term funding cycles and are not yet optimised for industrial-scale AI use.

[1] EMBL published its AI Science strategy in February 2025: https://www.embl.org/editorhub/wp-content/uploads/2025/02/EMBL_AI-Strategy_Feb2025_Accessible.pdf

[2] Google DeepMind partnered with EMBL-EBI to develop the AlphaFold database and make openly available the results of an AI system which makes state-of-the-art, accurate 3D protein structure predictions, which previously took years, in minutes. https://www.embl.org/news/science-technology/alphafold-wins-nobel-prize-chemistry-2024/

[3] See Global Biodata Coalition Working Group on Sustainability. (2023). Consultation Paper. Zenodo. https://doi.org/10.5281/zenodo.8384740

- **Establish a dedicated "European Data Investment Programme" to guarantee the sustainability and industrial usability of core datasets**. Unlocking the full potential of AI for Europe's science will depend on access to high quality and highly annotated data at scale, linked to key associated metadata types. Yet, the long-term sustainability of many open biodata resources, hosted by European infrastructures and used extensively[4] by the health and agri-food sectors, is far from secure. Many are reliant on short-term competitive funding sources, sometimes from a single or very limited number of funders. This creates operational fragility and long-term risk for services that are now mission-critical for Europe's AI innovation and digital sovereignty.

- **Invest in infrastructure for AI-generated outputs, including dynamic model repositories and curated AI prediction archives.** Recognising that AI models and their predictions will increasingly complement traditional biodata, these new use cases require infrastructure for archival, curation, and sharing. Future investment should support the development of model repositories and dynamic databases to host pre-computed AI outputs, building on successful examples such as the AlphaFold database. Ensuring data usability at scale also requires the development and adoption of coordinated meta standards.

2. Align Compute and AI Factories with Life Science Sectoral Needs

AI deployment requires tailored computing environments that are responsive to the specific requirements of different scientific domains. Initiatives such as the EuroHPC Joint Undertaking and the upcoming AI Factories and Gigafactories represent vital infrastructure investments for Europe's digital future. As these capabilities are further developed, there is an opportunity to enhance their alignment with the operational needs of data-intensive domains, such as the life sciences.

- **Develop computing environments tailored to the needs of specific sectors.** Many disciplines, particularly the life sciences, encounter distinct challenges, including substantial hidden overheads related to data preparation, workflow adaptation, and coordination across teams. Most HPC systems are optimised for simulation-heavy fields such as physics, while life science data processing involves different computational patterns. For example, environmental datasets have complex data linkage, while in health and biology, those are increasingly too large, complex, or sensitive to be easily transferred between facilities. This presents challenges for infrastructure design, and calls for secure data environments that can support cross-institutional and cross-border research.

- **Embed AI Factories within the existing research infrastructures ecosystem**. Scaling existing HPC centres will not be sufficient; this requires careful attention to

[4] In 2024, 5.4 billion requests from 6.1 million unique IP addresses to EMBL-EBI data resources came from EU member states.

the ecosystem around such facilities, bringing together HPC infrastructure specialists with AI and science communities and existing research infrastructures to create environments where researchers can effectively develop and deploy AI tools. This can be addressed through support for applied use-cases and long-term partnerships that create sustainable ecosystems that encourage innovation and ensure long-term return on investment.

3. Bridge the AI Talent Gap through Sector-Crossing Career Pathways

Europe must ensure that AI adoption is underpinned by a skilled workforce. Life science and biotech SMEs consistently cite data skills and AI talent as key barriers to innovation. Europe's public research organisations, including large-scale infrastructures[5] and intergovernmental institutions, host world-class training ecosystems that could be leveraged for broader talent development beyond academia.

- **Expand industry-embedded training models, such as co-funded PhDs or fellowships jointly supervised by research institutes and biotech/pharma companies.** These joint training initiatives foster interdisciplinary collaboration and provide early-stage researchers with structured exposure to both academic and industrial environments. Public-private models already in use, such as EMBL's co-funded postdoc programmes with GSK[6], offer scalable blueprints for how such schemes could be supported at EU level.

- **Support cross-sectoral engineer exchanges, allowing industry professionals to contribute to academic technology development and vice versa**. Engineers and technical experts play a foundational role in building and maintaining the AI systems that underpin both scientific research and industrial innovation. Structured secondment programmes, embedding industry engineers in research infrastructures or enabling academic technologists to work temporarily within biotech and pharma companies, would help align technical capabilities across sectors. These exchanges would not only strengthen innovation pipelines but also deepen mutual understanding of constraints and opportunities across the research-industry interface. A co-funding mechanism at EU level could help mainstream these exchanges, recognising the critical role of engineers as AI enablers.

4. Leverage Europe's Research Infrastructures for Industrial AI Uptake

Europe's research infrastructures (RIs) ecosystem is globally unique, providing open-access data resources essential to both academic and industrial AI applications for human health and the environment. Beyond data, many RIs, such as EMBL, also act as incubators for

[5] For example, EIROforum members (CERN, ESA, ESO, ILL, EUROfusion, European XFEL, ESRF, EMBL): https://www.eiroforum.org/

[6] https://www.embl.org/news/lab-matters/1702-new-collaboration-embl-gsk/

AI-driven technologies[7] and, through strong industry partnerships, help translate scientific advances into industrial innovation, reinforcing Europe's technological leadership.

- **Recognise RIs as strategic digital assets for Europe's AI competitiveness.** RIs already underpin Europe's leadership in open data and AI-ready resources, serving both academic and industrial users globally. Recognising RIs as foundational components of Europe's AI ecosystem, and ensuring they are equipped with the mandate, resources, and governance frameworks to support industry engagement, would help unlock their full potential as innovation accelerators. The long-term success of AI also depends on sustained investment in basic research. Fundamental scientific advances in biology, chemistry, and data science often lay the groundwork for transformative AI breakthroughs, such as AlphaFold, well before their industrial impact becomes visible.

- **Accelerate AI-enabled technology development within RIs.** Research infrastructures should be supported to develop and deploy next-generation scientific instrumentation and automated workflows that harness AI capabilities. Examples include smart microscopy, real-time data analysis pipelines, and autonomous experimental systems. These innovations enhance the productivity and precision of scientific discovery and create technology platforms with high translational value for industrial R&D, particularly in sectors such as pharmaceuticals and diagnostics.

About EMBL

The **European Molecular Biology Laboratory** (EMBL) is one of the world's leading research institutions, and Europe's flagship laboratory for the life sciences, founded in 1974. With 29 member states, the EMBL is an inter-governmental organisation with more than 110 independent research groups and service teams covering the spectrum of molecular biology across six sites in Germany, France, Spain, the United Kingdom and Italy. EMBL is driving visionary fundamental research, offers vital services to scientists globally, trains Europe's future scientific talent while actively engaging in technology transfer and industry relations, and nurturing policy dialogue in Europe and worldwide. www.embl.org

Contact information

Plamena Markova, Chief of International Relations: int-relations@embl.org

Anne-Charlotte Fauvel, EU Relations Lead: anne-charlotte.fauvel@embl.org

[7] https://www.embl.org/news/science/using-artificial-intelligence-to-discover-therapeutic-antibodies/