The importance of open data infrastructures for European competitiveness in life sciences

ELIXIR Position Paper

Europe's ability to compete globally in life sciences depends not only on excellent research, but on the infrastructure that enables it. Open data resources—such as deposition databases and knowledge bases of curated data—are essential to research and innovation and Europe's competitiveness. Yet these resources are often invisible, undervalued and underfunded. At a time when global uncertainty threatens these scientific assets, Europe has an opportunity to act decisively and invest in its own data infrastructures.

This position paper outlines why open data is a strategic asset, how it underpins European competitiveness and AI capability, and what actions are needed to secure its future.

I. Open data resources are the foundation for life science innovation: Europe should invest in open data as a critical enabler of innovation

Open data infrastructures, such as deposition databases and curated knowledge bases, are critical enablers of life science research and innovation. These resources allow scientists in both academia and industry to analyse larger, richer, more reliable and more consistent datasets, make comparisons across different data types and derive new knowledge faster and more reliably. The impact of these resources is systemic: breakthroughs like AlphaFold, mRNA vaccine development and antimicrobial resistance tracking would not have been possible without decades of open, structured data from resources like the Protein Data Bank (PDB), UniProt/Swiss-Prot, the European Nucleotide Archive (ENA) and other ELIXIR Core Data Resources[1]. This generates long-term benefits and immediate socio-economic impact[2,3], representing a significant return on investment for governments and funders.

[1] https://elixir-europe.org/platforms/data/core-data-resources
[2] See Impact Assessment of EMBL-EBI: https://www.embl.org/documents/wp-content/uploads/2021/10/EMBL-EBI-impact-report-2021.pdf
[3] See Open Data: A Driving Force for Innovation in the Life Sciences: https://f1000research.com/documents/10-828

And yet, these infrastructures are at risk. Historically, the US, particularly through the National Institutes of Health (NIH), the National Science Foundation (NSF) and the United States Department of Agriculture (USDA), has played a central role in sustaining open data infrastructures. However, recent policy changes—like blanket bans on specific institutions or on foreign subawards—have led to disruptions to support of internationally shared infrastructures. Many of Europe's core bioinformatics resources are tightly integrated with international collaborations. The retreat of US funding agencies from shared governance or cost-sharing could destabilise these collaborations, putting at risk innovation in Europe, particularly future scientific and economic developments that are based on accessing high-quality data such as developments in AI.

Recommendations:

- Explore transitional funding mechanisms to support critical European resources following reduction in US support.
- Recognise data resources as infrastructure, not research projects, with suitable funding instruments and lifecycle management.

II. To reach their full potential, AI investments must be anchored in sustainable data infrastructure

The Draghi report, *The Future of European Competitiveness* (2024)[4], stresses the strategic role of research infrastructures and data. It frames competitiveness not as low wages or more hours, but as the capacity to innovate. Enrico Letta's high-level report, *Much More than a Market* (2024)[5], reinforces this vision by calling for a "fifth freedom" in the EU Single Market: the free circulation of research, innovation, data, knowledge and education. This reframes data not just as an enabler, but as a strategic factor of production—on a par with goods, services, capital and labour.

Open data infrastructures thus play a critical role—and nowhere more than in AI, where Europe is investing billions into AI factories, data spaces, and digital sovereignty. But these efforts will stall without high-quality, interoperable and well-annotated data. AI does not

[4] https://commission.europa.eu/topics/eu-competitiveness/draghi-report_en
[5] https://european-research-area.ec.europa.eu/documents/letta-report-much-more-market-april-2024

thrive on data volume alone: it requires standards, structure, provenance and curation—features provided by trustworthy life science resources like the ELIXIR Core Data Resources.

Europe's investments in AI and compute capacity must be matched by sustained support for the upstream data infrastructures that are essential for their operation. Without high-quality, curated and interoperable data, flagship AI initiatives risk under-delivering on their potential.

Recommendations:

- Anchor investments in AI and data spaces in sustainable, FAIR[6]-compliant and openly accessible datasets—particularly in life sciences, where data complexity is highest and potential impact greatest.

III. Secure Europe's data sovereignty through investing in world-leading data resources, a skilled workforce and national capacity aligned with European infrastructure

Databases and digital resources are more than tools — they are critical components of an interconnected research infrastructure that enables discoveries and innovation across the life sciences.

As data becomes central to scientific progress and technological competitiveness, sovereignty concerns have rightly gained prominence. Ensuring that Europe is not dependent on data controlled elsewhere is essential for long-term research, innovation and economic resilience. But data sovereignty is not about isolation. On the contrary, real and lasting protection is driven by leadership, trust and influence.

When international partners align their data formats, identifiers and standards to European infrastructure, such as those coordinated by ELIXIR, Europe's strategic position is strengthened.

Sovereignty also depends on having the expertise to interpret, curate and govern data effectively. EU member states must therefore invest not only in transnational and national

[6] FAIR stands for Findable, Accessible, Interoperable, Reusable

infrastructure, but also in people and skills to sustain it. A skilled workforce of bioinformaticians, data scientists and infrastructure specialists is essential to turn digital assets into knowledge and innovation.

Recommendations:

- Reinforce Europe's sovereignty by strengthening Europe's role as a net data contributor, ensuring others continue to align with European infrastructures and open data values.
- Invest in human capital by sustaining a digital workforce with expertise in managing, curating and exploiting complex datasets.
- Amplify national efforts by incentivising alignment of national capacities with EU-level strategies (e.g. the EU Life Science Strategy, the European Strategy Forum for Research Infrastructures (ESFRI) and the Common European Data Spaces including the European Open Science Cloud (EOSC)).

IV. Europe should play the role of global leader in open science, supporting multilateral collaboration and championing the benefits of open data

Europe has a unique opportunity to assert global leadership in open science. As other actors scale back or disengage from shared scientific infrastructures, Europe can step forward to provide continuity, trust and long-term support.

Investing in open data infrastructures is not only a scientific priority—it also builds a diplomatic and geopolitical asset. By visibly supporting open access EU research infrastructures, Europe can reinforce multilateral international collaboration and the scientific commons, and reinforce its status as a reliable, values-driven partner and a hub for academic freedom and innovation.

At the same time, the design of the EU's next research framework, FP10, offers a rare opportunity to translate rhetoric on open science into concrete commitments. Sustained investment in globally connected, open infrastructures would send a clear signal: that Europe is prepared to lead, not just advocate.

Recommendations*:*

- Support multilateral science by providing stable, long-term funding for international open data infrastructures with strong European participation.
- Leverage diplomacy through openness by positioning Europe as a principled and dependable steward of global scientific collaboration.
- Lead through action by making open science and shared data infrastructures central, well-funded pillars of the next Framework Programme—signalling to the world that Europe values open data and global collaboration.

ELIXIR Hub, Hinxton, UK - 20 June 2025
https://www.elixir-europe.org