

Advocating for a Better Balance Between Research Data Protection and Research Data Openness: Toward a Gradual Opening of Research Data

Good data management: an asset for sovereignty

Current geopolitical tensions and attacks on researchers working on sensitive topics now require the implementation of reasonable precautions before opening up research data. The goal is to prevent interference, protect researchers and scientific values, ensure research sovereignty, and ultimately guarantee academic freedom.

The latest recommendations—from the Council of the European Union¹ and the French National Strategic Review²—are moving toward strengthening research security, a development that our institutions must take into account. In this context, calling for caution in making research results available, it appears both urgent and necessary to consolidate policies for good data management and to strengthen national and European infrastructures for research data.

Balancing Open Science and Research Protection

The European Open Data Directive and the 2016 French Law for a Digital Republic have enabled universities to develop ambitious policies in support of open science, many of which include a component dedicated to opening up research data. Making these validated datasets openly available helps ensure more transparent and cumulative science, as well as compliance with research integrity requirements³: data accessibility enables traceability, reproducibility, and reuse.

At the same time, the protection of certain data is imperative, whether it involves personal data, safeguarding the nation's scientific and technical potential, commercializing research results (particularly through patents), or maintaining medical confidentiality. Even if data remains closed or shared on a restricted basis, it must be managed according to FAIR principles for verification and cost-efficiency purposes⁴ because the responsibility for preserving research data now falls on institutions⁵. Data openness and data protection are closely intertwined issues. However, today they are often treated separately within universities and managed by professional groups with different work cultures. As a result, they too often appear contradictory to research communities, which feel that their governing bodies and funders are issuing conflicting directives.

¹ Council Recommendation of 23 May 2024 on enhancing research security. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024H03510>

² French National Strategic Review 2025.

https://www.sgdsn.gouv.fr/files/files/Publications/20250713_NP_SGDSN_RNS2025_EN_1_0.pdf

³ OFIS, Intégrité scientifique et sécurité de la recherche : pour une approche coordonnée face à l'ingérence étrangère, 2026.

<https://www.ofis-france.fr/wp-content/uploads/2026/02/Fevrier2026-Ofis-Note-Ingerences-Etrangeres.pdf> [in French]

⁴ European Commission: Directorate-General for Research and Innovation and PwC EU Services, Cost-Benefit Analysis for FAIR Research Data – Cost of Not Having FAIR Research Data, Publications Office, 2018, <https://data.europa.eu/doi/10.2777/02999>

⁵ Décret n° 2021-1572 du 3 décembre 2021 relatif au respect des exigences de l'intégrité scientifique par les établissements publics contribuant au service public de la recherche et les fondations reconnues d'utilité publique ayant pour activité principale la recherche publique, <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000044411360> [in French]

A poor balance between protection and openness is not without consequences. When datasets opened too soon after their scientific validation, or are not sufficiently vetted before being opened, sensitive data may be disclosed, or the confidentiality required prior to a patent filing may no longer be maintained. However, restricting access to datasets does not guarantee their protection, and restricting access without proper data management often results in the data quickly becoming unusable.

Taking a Strategic Approach to the Gradual Sharing of Data

Science is cumulative and can only benefit from the open access to research data. This must be viewed as a gradual, long-term process requiring careful consideration in advance regarding proper data management and protection before making the data available. Each stage, from the strictest confidentiality to restricted sharing among peers or teams, and finally to public and unrestricted dissemination via a data repository, takes varying amounts of time depending on the nature and sensitivity of the data. Even in the case of sensitive data, confidentiality is not perpetual, and data eventually always becomes public. In France, the archives of the Defence Historical Service are the best-known example of this, with confidentiality periods that are certainly long but rarely exceed half a century. This aspect of data's future is often overlooked in discussions about data protection and openness, even though it is essential to address it within research-intensive universities. The future of data must be considered over the long term. The decision to keep data closed comes at a cost to institutions: financial, when it comes to acquiring ever-greater secure storage capacity; reputational or legal, when data is lost, altered, or leaked. Regardless of the solution chosen by the researcher or the institution, the future value of the data lies in best practices established from the moment of its creation.

Raising awareness among all stakeholders about responsible and ethical data management

Raising awareness about proper data management is an excellent way to foster collaboration among the services responsible for research data protection and those responsible for open science⁶: chief data officers, data protection officers, defence security officials, chief information security officers, scientific integrity officers, librarians, archivists, and technology transfer officers. Concerted action by all these stakeholders is necessary to carry out collaborative work on data management that provides institutions with the tools to prepare for the future of their data.

⁶ Data management clusters are an ideal setting for raising awareness among communities about these topics; however, support for projects involving data that must be protected can only be handled internally within institutions to ensure confidentiality.
<https://recherche.data.gouv.fr/en/data-management-cluster>