EARE contribution to the upcoming Data Union Strategy

EARE welcomes the opportunity to contribute to the Data Union Strategy – an essential step in aligning Europe's data governance with technological progress. By fostering legal certainty, trust, and data sharing, this strategy can help position Europe as a global leader in AI. To fully realize this vision, key challenges facing researchers, startups, and innovators must be addressed.

1) Limited access to data hinders innovation

Access to large, diverse datasets is essential for effective AI development. Yet, researchers and innovators face barriers due to paywalls, copyright restrictions, and technological protection measures (TPMs), which can introduce bias and limit research quality. Moreover, a substantial portion of publicly funded research remains behind paywalls, despite being taxpayer-funded. Introducing secondary publishing rights (SPRs) at the EU level could expand data access, improving Europe's competitiveness and innovation capacity. SPRs would allow researchers to republish their work in open repositories after a short embargo, enhancing transparency, reproducibility, and dataset quality.

Our recommendation: Ensure open and fair access to public and publicly funded data, in line with the Open Data Directive. Introduce a Secondary Publishing Right to guarantee immediate open access to taxpayer-funded research. Making data accessible, interoperable, and reusable will empower EU researchers, innovators, and startups, fostering a transparent, collaborative, and innovation-driven ecosystem.

2) Strengthened TDM rights and open data policies are essential to unlock AI innovation in Europe

Text and data mining (TDM) enables researchers and innovators, including SMEs, to analyse large datasets and develop quality AI models. While Articles 3 and 4 of the EU Copyright Directive provide exceptions to balance innovation with rightsholders' interest, reservation rights for rightsholders under Article 4 and the growing push for licensing schemes risk turning freely usable content into restricted material. This reduces data availability and diversity, increasing biases in AI models, and hampering innovation. The lack of harmonisation and legal uncertainty surrounding Articles 3 and 4 add confusion, limiting their use and increasing data access costs, especially compared to countries like Japan, Singapore, and the U.S., where copyright laws better support innovation.

Our recommendation: Revisit Article 4 of the Copyright Directive, guided by the more flexible TDM exceptions seen in competitor jurisdictions. Clarifying the interpretation and implementation of TDM exceptions for both research and commercial use across Member States, including consistent guidance on lawful access, permissible reuse, and clear boundaries for opt-outs, would help ensure the EU remains competitive in AI development.

3) A modern data strategy must reflect the cross-sector nature of today's research and innovation landscape

Limiting TDM exceptions to non-commercial use creates legal uncertainty, especially in public-private partnerships (PPPs) or when non-commercial research leads to commercial outcomes (e.g. medical innovations). This restricts data sharing and disproportionately impacts startups and SMEs, hindering AI training and delaying real-world application of research.

Our recommendation: Support data use across sectors, especially in PPPs. Legal frameworks must ensure that TDM exceptions and open data policies apply to research with commercial potential. This would reduce ambiguity, encourage broader data sharing, and support innovation.

Conclusion

Europe's AI leadership relies on a data strategy that ensures open, secure access. To stay competitive, the EU must remove legal barriers, enable cross-border collaboration, and promote responsible data sharing. EARE urges the EU to expand access to large datasets and foster knowledge exchange to strengthen strategic autonomy and AI innovation.

About EARE: *The European Alliance for Research Excellence (EARE) was convened in 2017, and now brings together nine members from the research and innovation ecosystem in Europe, including Microsoft, the Association of European Research Libraries (LIBER Europe), the European Bureau of Library, Information and Documentation Associations (EBLIDA), Allied for Startups, BSA | The Software Alliance, LACA, Research Libraries UK, SCONUL (Society of College, National and University Libraries) and UCL (University College London) Library, advocating for the EU to live up to its innovation potential in the digital economy.*