EARE
European Alliance for Research Excellence

EARE policy priorities for the next mandate (2024-2029)

Pioneering Excellence in Research Across Europe

The European Alliance for Research Excellence (EARE) brings together public and private actors from the European research ecosystem.

At EARE, our primary objective is to advocate for a policy and regulatory environment which facilitates data sharing and use, fosters innovation, and promotes a competitive European research and innovation ecosystem. Based on our members' feedback, we have developed our policy recommendations for the next five years.The upcoming mandate will be pivotal in shaping the future of research and development, especially with the development of Artificial Intelligence (AI). It will play a crucial role in enhancing the EU's competitiveness and strategic autonomy, ensuring that Europe remains at the forefront of innovation while safeguarding its values and interests (see here). We call on the European Commission to help ensure that AI contributes positively to society and the economy and reinforces Europe's position as a leader in innovation.

THE CHALLENGES ENCOUNTERED BY RESEARCHERS IN EUROPE

Text and Data Mining (TDM) represents the next wave of innovation enabling European governments, researchers, and small businesses to unlock the power of data. The Integration of TDM practices, in particular of AI, in the research process is crucial as most cutting-edge research uses AI tools and techniques. While AI is increasingly ubiquitous, researchers and innovators face difficulties relating to the access of copyrighted materials, which severely restricts their work and hampers their ability to innovate in Europe.

Limited access to data hampers innovation and research

One of the most common challenges facing researchers and innovators in Europe is the inability to access the data they need to develop trustworthy AI systems. Access to vast amounts of data permits large-scale analyses, which in turn helps accelerate robust scientific discoveries.

Moreover, the risk of bias in AI systems will only increase if developers are unable to access copyright-protected data. Access to broad and varied data is needed to support the development of high-quality AI systems (see here). Large publicly available data sets, including web scraped data by third party aggregators are an important resource for research. However, access to such data sets are jeopardised by efforts to prevent web crawlers or to restrict access to publicly available data. Additionally, researchers find paywalls and electronic fences (Technological Protection Measures or TPMs) to be a major obstacle in accessing copyrighted online resources for which they have paid and to which they have a right of access, which can undermine the exception included in Article 4 of the DSM Directive. A survey conducted by EARE's member LIBER Europe found that TPMs hinder research by causing prolonged lock-outs to works that been paid for and which are legally accessible. This leaves university access uncertain, even after resolution. These disruptions prevent essential activities like TDM, crucial for modern research.

Legal fragmentation and uncertainties that impede the use of AI & machine learning

Researchers all too often refrain from using research tools such as AI and machine learning "because they are afraid of copyright infringement" (see here). The implementation of the TDM exceptions and of the Directive on copyright and related rights in the Digital Single Market (DSMD) remains fragmented in the EU. Moreover, the delineation between commercial and non-commercial uses create legal uncertainty for researchers and AI providers.

A lack of education about the benefits of data sharing and translational research

EARE's members also identified a lack of awareness and education about open data, the legality of data sharing, and related topics as one of the difficulties researchers and innovators face. Despite recent progress, persistent challenges and gaps demonstrate that further effort is needed to increase data collaboration.

Key Findings from European Commission's Survey on Copyright and Data Use by Researchers and Research organizations

81.4%
supported the idea of an open-ended clause in copyright law permitting the use of copyright-protected knowledge resources for all kinds of research purposes.

80%
of researchers face significant barriers due to a lack of subscriptions to access copyrighted knowledge resources. This underlines the need to introduce a Secondary Publishing Right in European law.

90%
of research organizations voiced support for the provision of further guidance to researchers regarding how existing copyright exceptions can facilitate TDM.

EARE POLICY PRIORITIES

Addressing these issues would require the European Commission to create a more conducive environment for research and innovation, to enable better access to data to ensure an efficient use of public funds and technology in science. It would require bringing down barriers, developing incentives, and enabling a competitive research and innovation ecosystem to spread knowledge and ensure data access. EARE members have identified seven principles to provide policymakers with options to foster the responsible uptake of AI in science and research in the EU, and to position Europe at the forefront of research and innovation.

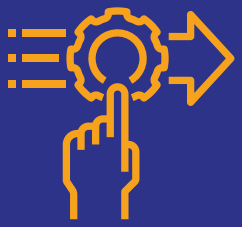
Align policies with the realities of research today, and move beyond the obsolete differentiation of commercial & non-commercial research

In its approach to copyright and open data policies, the European Commission, must recognise the reality of research today. Research projects often have mixed funding sources involving staff working across public-private boundaries. Projects which have goals blurring the lines between commercial and non-commercial purposes find their innovative potential to be limited by this artificial dichotomy. This artificial differentiation raises doubts about the applicability of TDM exceptions to public-private partnerships (PPPs) and down-stream commercialisation, which ultimately hinder innovation in Europe.

Ensure an open access to data is essential for research and development

The access rights to data used to train algorithms, governed by the Directive on copyright and related rights in the DSMD, have been questioned during the recent negotiations on the AI Act. There were heated debates on how to limit access to and use of data that is used to train AI, while other recent data legislation (Data Act, Data Governance Act) sought to promote data openness to allow AI development and create value in Europe. It is imperative that AI tools are accurate, reliable, reproducible, truthful, and free from bias – with data as the backbone that underpins this. Increased access to broad and varied data should therefore be prioritised to unlock AI development and benefit researchers and innovators. As recognised by the Council of the EU and the European Commission (see here), adopting an open approach to data will be crucial to ensure Europe's competitiveness and future prosperity. Instead of restricting data access, the European Commission should promote open data policies and adopt data access policies and infrastructures that facilitate data sharing while maintaining security. In this context, Europe should adopt a EU-wide Secondary Publication Rights which would enable immediate access to publicly funded research covering all types of uses (see here).

Prioritize the proper implementation of TDM exceptions introduced by the Copyright Directive

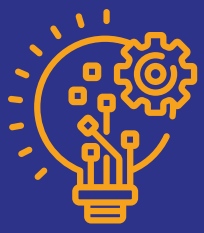


TDM exceptions in Articles 3 and 4 of the DSM Directive were introduced to contribute to the development of artificial intelligence in the EU (see European Parliament's summary, the European Commission's Q&A and press release, EARE's position papers). TDM exceptions were clearly intended to permit the training of systems such as large language models. The opt-outs within Article 4 of the DSM Directive were expressly intended to allow rights holders to reserve their rights with respect to TDM activities for commercial purposes.

While a more open approach based on the principle that "the right to read is the right to mine" would have been an ideal scenario, we believe that the current legislation gives rights holders the ability to participate in AI in a manner that reflects their preferences. The Commission should pay particular attention to the implementation of TDM exceptions. In particular, the Commission should ensure that exceptions are balanced and do not place a disproportionate burden on TDM users, and that they are not stifling innovation or leading to entrenched domination by stakeholders with the financial means to access restricted data.

Develop opt-out standards in a way that benefits innovation and research

It is recognised that opt-outs exercised under Article 4 of the DSM Directive can reduce the amount of data available for training which can in turn lead to less varied datasets and poorer AI models. Therefore, any opt out must be clearly defined, and not lead to ambiguity that unduly further limits the data that is used for training. Policy makers should therefore support industry efforts between rightsholders and AI developers to develop standard technological solutions for declaring and reading machine readable opt outs under Article 4 of the DSM Directive. There is currently no consistent technology in place, leading to confusion, uncertainty, risk, and increased costs. European institutions and standardisation organisations should pay special attention to SMEs and startups and develop feasible and scalable standards based on what already exists. Stakeholders' collaboration in developing standard approaches for opt outs would provide clarity for rights holders and organisations, regarding the data used to train models and help reassure content providers, developers, researchers, and educators alike.



Enable a competitive Research & Innovation Ecosystem

AI is a driver of innovation in all sectors of the economy. Intellectual Property should foster innovation instead of limiting it. Copyright should not be used to solve issues raised by the development of AI. Countries such as Japan and Singapore have passed pro-innovation copyright laws that strongly support machine learning and minimize bias while US case law confirms that the analysis of data constitutes fair use. In contrast, in Europe, the reservations of rights under Article 4 for commercial TDM reduce the data a model can train on and leads to market concentrations, hampers competition by creating barriers to entry for smaller companies and startups and stifle privately funded research. Open data policies can alleviate this issue to some extent by making data available for all players – leading to a more competitive AI landscape that promotes faster advances and the responsible use and deployment of AI technology.

Establish transparent mechanisms to resolve TDM-related blocks and timely restoration

To prevent Technical Protection Measures (TPMs) from becoming a major obstacle in accessing paid and legally accessible copyrighted online resources and to ensure the exception included in Article 4 of the DSM Directive applies, EARE calls on Member States to ensure lock outs are resolved within 72 hours. We also call for transparent mechanisms to quickly resolve TDM-related blocks, advocating for improved circumvention measures and timely access restoration.

Promote education and AI uptake

To understand and address societal challenges, such as climate change of pandemics, we need to support cooperation, encourage interdisciplinary knowledge-sharing, invest in knowledge valorisation (creating social and economic value by transforming data and research results into beneficial solutions for society), and to foster AI uptake in science (see here). A multi-faceted strategy is required to make this a reality, including policy initiatives, funding, incentives not to opt out from AI development, and education. As part of this strategy, the EU should encourage voluntary data sharing by providing incentives for individuals and organisations to share data and educating them about the societal benefits of data sharing.

The Path Forward

As the global race to responsibly develop powerful AI systems accelerates, the European Union risks of falling behind other regions. In this regard, EARE welcomes the commitment by the President of the European Commission to elaborate a 'strategy for a European Data Union' within the next five years, to support Europe's rise to power in the AI sector.

However, to maximise Europe's innovative potential, the European Commission should consider all of these recommendations during its next mandate. We call on the EU to address the challenges detailed above, ensuring access to large data sets for AI training, but also to ending the outdated distinction between commercial and non-commercial users and purposes, and to promote knowledge sharing to foster European strategic autonomy and innovation.

ABOUT US

The European Alliance for Research Excellence (EARE) is a coalition of companies and research organisations committed to fostering excellence in research and innovation in Europe.

Established in 2017 with the support of Microsoft, Allied for Startups, and BSA | The Software Alliance, EARE brings together esteemed members from the European research and innovation ecosystem, including LIBER Europe, Research Libraries UK, UCL Library Services, SCONUL, and the UK Libraries and Archives Copyright Alliance (LACA).

The Secretariat of the Alliance is provided by APCO. EARE's priorities and positioning are always based on consensus from its members – whose input is collected on a regular basis. Microsoft funds the Secretariat's activities, and EARE is governed by by-laws which can be consulted upon request. The Alliance is listed in the EU Transparency Register (n. 193384748034-24).

OUR MEMBERS

Allied FOR STARTUPS

The Software Alliance BSA

LIBER

EBLIDA

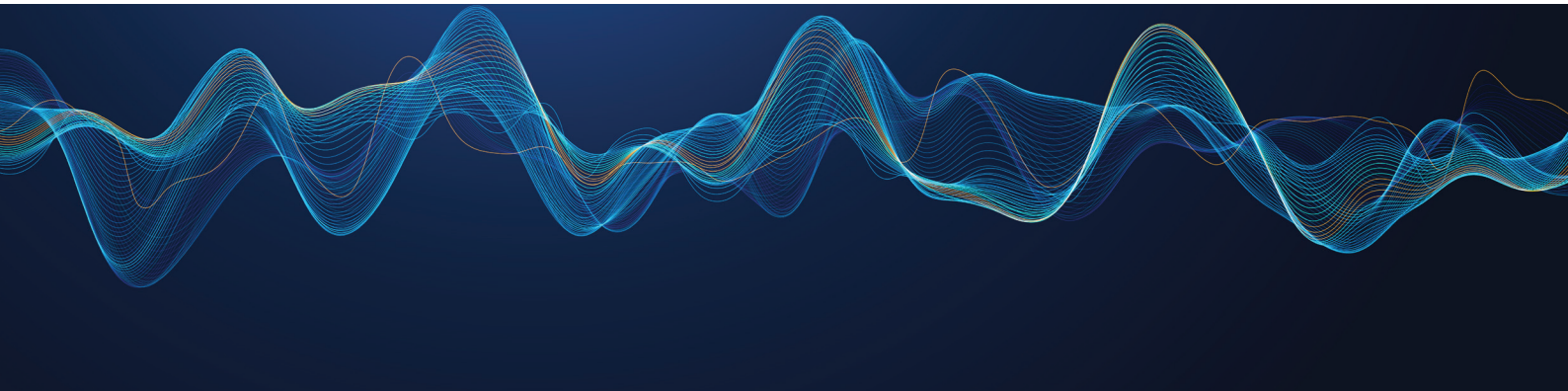
L Libraries & A Archives C Copyright A Alliance

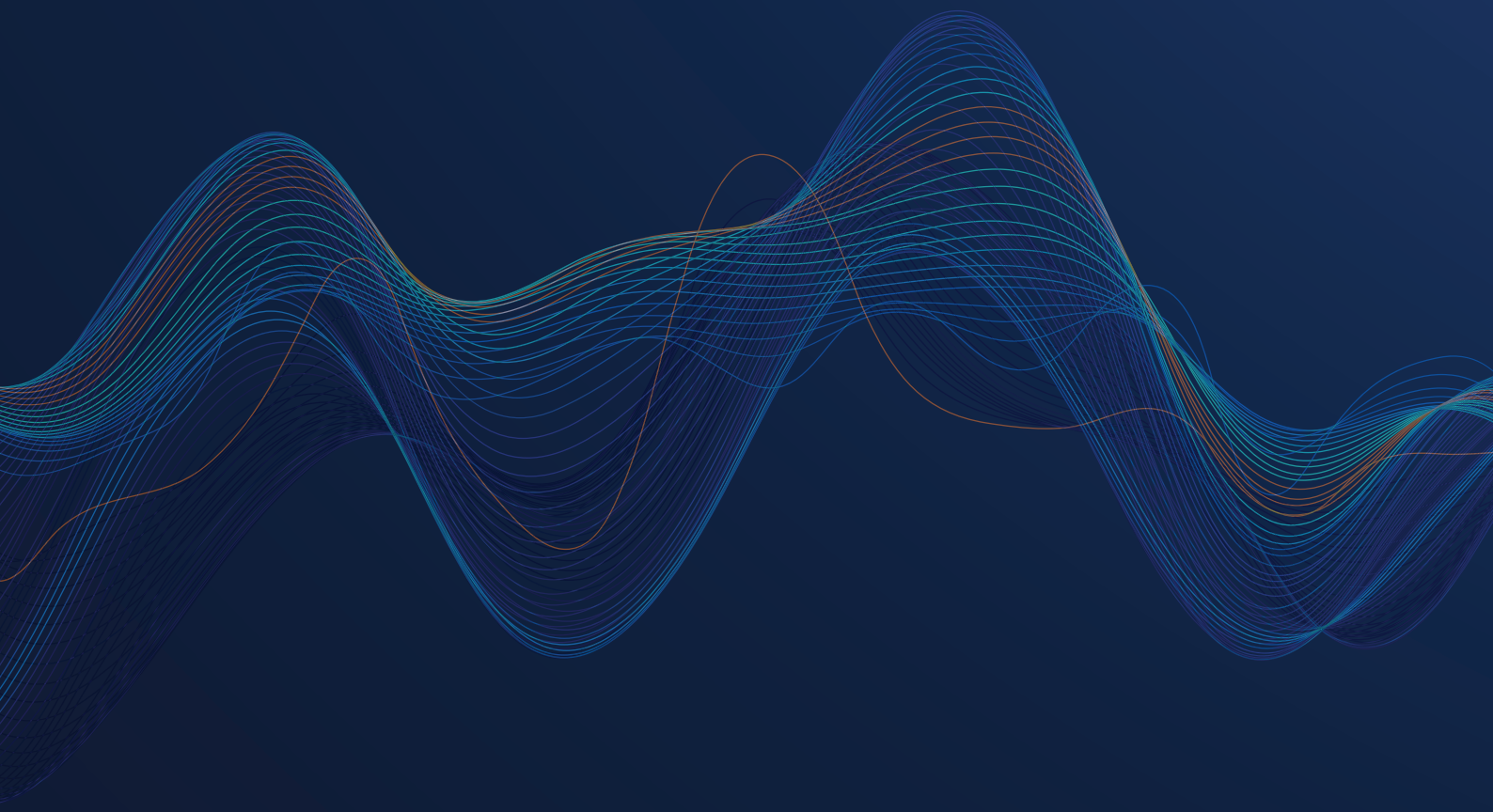
Microsoft

RLUK Research Libraries UK

SCONUL

UCL





European Alliance for Research Excellence (EARE)

Rue Montoyer 47, 1000 Bruxelles
Email: eare@apcoworldwide.com

EARE
European Alliance for Research Excellence