Despite some shortcomings, the current European copyright framework remains fit for purpose
Reflections from the European research & innovation ecosystem

In light of the recent questionnaire circulated by the Council Presidency on Copyright and Artificial Intelligence (AI) which aims to assess the state of play in the field of Artificial Intelligence (AI) and Copyright and understand the current trends, the European Alliance for Research Excellence (EARE) wishes to highlight that the Copyright Directive and the related Text & Data Mining exception, despite some clear shortcomings, remain the right instrument to encourage AI development in Europe.

1. **State of play in the field of AI and copyright**

- **EARE believes that the current rules on copyright & AI**, and in particular articles 3 & 4 which introduce Text and Data Mining (TDM) exceptions in the Directive on Copyright in the Digital Single Market (DSM Directive) provide **a good enough legislative framework** and there is no need to introduce further copyright-specific legislation. Nonetheless, the Directive needs to be properly implemented and opt-outs, since they exist, need to be standardized.

- **Text and Data Mining (TDM) is crucial for AI development**. Without the ability of computers to access and analyze very large amounts of data and employ cognitive technologies to allow the learning of patterns, technologies such as AI would simply not be possible.

- **TDM exceptions were specifically introduced** in the Directive on Copyright in the Digital Single Market (DSM Directive) **to contribute to the development of AI in the EU** and are limited to data and content on which users have "lawful access". Lawfully accessed content includes content made publicly available, or accessed in any other lawful manner, including but not limited to, access under agreement.

- **EARE believes these exceptions promote innovation and research by providing clarity that TDM practices** are clearly permitted under copyright law. By allowing researchers, universities and innovators to access and use copyrighted content for TDM without needing prior authorization, they facilitate scientific progress and innovation, while maintaining protection for copyright holders.

- Unduly **limiting TDM exceptions** by improper implementation of the opt out, or requiring an opt in approach, **would take Europe away from the evolving international legal landscape** when it comes to regulating the intersection between AI and Copyright and will significantly hamper AI development in Europe.

- The delineation between non-commercial and commercial TDM between articles 3 and 4 suggests that research, such as that undertaken by academia, should benefit from a broad TDM exception in article 3. However, **in reality it is not possible to delineate between commercial and non-commercial activities**. For example, university research is often privately funded. TDM activities may be undertaken in public private partnerships, and research projects which may start out as non-commercial efforts may later be commercialized. As such research cannot benefit readily from the broader exception under article 3, and the implications of complying with obligations under article 4 will apply to most areas of research, even when conducted by non-commercial entities.

- **Europe should not create its own hurdles with A**I. Machine learning applications and AI technology are pervasive in our everyday life, and adoption will only increase. This assessment and the need to ensure a broad access to high quality data are backed in two recent studies published and commissioned by the European Commission in 2024 ([here](#) and [here](#)), in Von der Leyen's recent policy priorities, in the Council strategic [agenda](#) and in recent [conclusions](#) adopted by the Council on

knowledge valorization. The Commission also considers the TDM exceptions included in the DSMD Directive as a [pro-competitive initiative](#) that should be considered important to reduce barriers to entry or limit their effects in the AI market.

2. **Training of AI models**

- While TDM exception can provides legal certainty, lack of consistency and standardization regarding implementation of the exception will nevertheless continue to present complications:

  o **Legal fragmentation regarding the implementation of the DSMD leading to uncertainties that impacts research and innovation:** Researchers all too often refrain from using research tools such as AI systems "because they are afraid of copyright infringement" and innovators face complexity, uncertainty and compliance costs which are also barriers to market entry for startups.

  o **The distinction between commercial and non-commercial use in EU copyright law hinders innovation and collaboration** - The commercial/non-commercial dichotomy contained in EU law and in the DSMD only adds further uncertainty, because it raises doubts about the applicability of TDM exceptions to public-private partnerships (PPPs), and down-stream commercialization.

  o **Opt-outs result in the creation of data moats that hamper the development of high-quality AI by reducing the diversity and quantity of data.** This limitation can result in models that are less robust, less accurate, and more prone to biases, hindering research and innovation in Europe.

  o **Licensing practices which follow opts out hinder competition**. Licensing practices following opts out can create challenges for smaller AI players. Indeed, when individual users or organisations opt out, larger AI companies with resources and datasets might be less affected. This can create barriers to competition and dynamic where large companies maintain their dominance. Instead, promoting open data can foster healthy competition and innovation in the AI landscape by making available valuable datasets accessible to a wider range of developers, researchers and smaller AI companies.

- **Opt out standards should be developed in a way that benefits innovation and research**. Currently, no consistent technology for opt-outs is in place, leading to confusion, uncertainty, risk, and increased costs. Nonetheless, the European institutions and standardization organizations should pay special attention to SMEs and startups and develop feasible and scalable standards based on what already exists.

- **Transparency requirements can introduce additional compliance costs for research organisations.** As recently mentioned in European Commission's [study](#) on improving access to reuse of research results for scientific purposes, obligations such as those established in the AI Act, requiring detailed summaries of the data used for training, can "*add a layer of compliance costs for research organisations*". Indeed, it will also apply to Article 3 and this layer of compliance has not been tested yet.

3. **Remuneration**

- EARE members support the principle that **the right to read is the right to mine**. Therefore, there should not be additional licensing mechanisms to grant mining access to data one already has access to. This was the principle behind the adoption of articles 3 & 4 of the copyright directive. We strongly support open access policies, and advocate against the creation of licensing mechanisms to provide mining rights to data one already has access to – these would only favor larger players on the market, who can afford to pay for these licenses.

- **In addition to hindering competition, licensing agreements on opted out data can increase the cost of research.** Research in the EU struggles to clearly benefit from the TDM exception under Article

3 of the DSM Directive due to its limited scope (non-commercial v. commercial) and ambiguities (harmonized and acceptable machine readable opt-outs). Indeed, Article 3 only applies to non-profit research institutions and cultural heritage institutions. This means that commercial entities, or collaboration between research institutions and private companies do not benefit from the exceptions. And this is without mentioning the lack of harmonization across Member States, the ambiguity in defining what constitutes "scientific research" or "non-commercial" purposes and the issues raised by technical protection measures (TPMs). As a consequence, to avoid the risk of infringing copyright, researchers very often refrain from using Article 3 and instead resort to TDM under Article 4. They may need to negotiate licensing agreements to gain access to the data, leading to licensing costs and negotiations delays. The inability to fully benefit from the TDM exceptions due to opt-outs, licensing fees and administrative burdens increases the overall cost of conducting research and discourages innovation, particularly for smaller organisations.

- **Open data policies**, which promotes the availability, accessibility, and usability of data, 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. **Secondary Publishing Rights should also be introduced, and technical protection measures better regulated** to improve circumvention measures and timely access restoration.

4. **Policy and international context**

- **Global coherence to regulate AI and Copyright is needed.** The EU must ensure that the rules adopted align with international law, and other jurisdictions. AI systems are very often developed internationally involving multiple actors and countries and AI products are often operating in a global market.

- Internationally, some countries are considering introducing or have introduced TDM exceptions without a reservation of rights and lean on thoughtful safeguards to increase accessibility in copyright works by AI developers or users for developing and training their AI systems and boost AI industry (Hong Kong, Japan, US, China, Singapore). Given this situation, we would need to agree on common international standards that promote innovation to make sure Europe keeps up in the AI race globally. The last thing Europe needs is to go backwards on this issue.

About EARE: *The European Alliance for Research Excellence (EARE) was convened in 2017, and now brings together eight members from the research and innovation ecosystem in Europe, including BSA | The Software Alliance, Microsoft, Allied for Startups, LIBER, 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.*