



Brussels, 12 July 2023

Data and Artificial Intelligence: Researchers and innovators need legal clarity, not another copyright reform

The European Alliance for Research Excellence (EARE) was built by companies and researchers united by the idea that **access and use of data should be as open as possible, and as closed as necessary**. This is what guided our advocacy work around the last reform of copyright rules in Europe, and more specifically on the need for a **broad exception to copyright for Text and Data Mining (TDM)**.

We commended the latest Copyright Directive, as it included what we saw as a **future-proof exception for TDM** (articles 3 and 4 of the [Copyright in the Digital Single Market Directive](#)). As a reminder, TDM is a building block for both machine and deep learning: without the ability of computers to analyze very large amounts of data and employ cognitive technologies to allow the learning of patterns, Artificial Intelligence (AI) is not possible. **It is now widely understood that the quality of AI algorithms depends on access to large and diverse datasets**. This is something that is recognized for instance in the USA and [Japan](#), where TDM activities are broadly allowed and even encouraged.

Recent developments in the sphere of AI have revived the debate in Europe about access rights to data used to train AI algorithms. There have been calls to limit access and use of data to train algorithms to data that is in the public domain only. Some websites are [considering](#) charging researchers to access data that is in the public domain. Most recently, some EU policymakers have pushed for the AI Act to include an obligation to provide a detailed disclosure of training data specifically protected under Copyright laws.

Transparency and reproducibility of research is core to the values of EARE. However, **this last-minute amendment by the European Parliament, just before negotiations are to take place between the EU institutions, would open a Pandora's box** of questions related to how this obligation would be implemented on a practical level, and how it would interact with existing Copyright laws. By way of comparison, Japan took 15 years to agree on a policy which would balance copyright rules with the development and commercialization of AI.

Obligations which touch upon transparency and openness of AI tools need to be proportionate, and any potential ramifications need to be properly investigated, understood, and balanced with international norms via a dedicated impact assessment. Switching from allowing free public domain data access for researchers to costly subscription models and **shoehorning ad-hoc requirements on the disclosure of training data specifically protected**



by copyright for AI tools in EU legislation, would not only hamper researchers' and academics' ability to innovate, but also go against the standards set by the European Commission through the Digital Services Act.

EARE believes that governments need to maintain the right balance between protecting copyright and rightsholders' productions to create an enabling environment for Text and Data Mining and AI-based innovations to thrive. And as highlighted by Commissioner Breton in this recent [response](#) to a Parliamentary Question, we wish to recall that **the provisions included in the Copyright Directive, and which are still being transposed across the EU, have already achieved this balance.** According to the existing legal framework, anyone can conduct text and data mining activities on data they have legal access to, without needing an additional license or authorization, or being placed under additional obligations. Rightsholders also benefit from an opt-out solution, whereby they can request that their content not be mined.

As the global race to responsibly develop powerful AI systems is accelerating, there is again a risk that the European Union may fall behind other geographies that host AI-powerhouses.

What European innovators and researchers need is legal clarity, not another copyright reform and new rules that may lack certainty and predictability. We call on EU policymakers to look first at existing legal frameworks and to ensure their effective implementation before suggesting new ones.

Signatories:

- BSA | The Software Alliance;
- Recherche - Association of European Research Libraries (LIBER);
- Society of College, National and University Libraries (SCONUL);
- UCL Library Services.