

KR21 PRINCIPLES ON ARTIFICIAL INTELLIGENCE, SCIENCE & RESEARCH



KR21 Principles on Artificial Intelligence, Science and Research

The European Union's Artificial Intelligence Act (EU AI Act) 2024 is not the first, and certainly unlikely to be the last effort by legislators to shape the development of AI as a technology, or the business models that build up around it.

Given that much cutting-edge research in AI takes place in universities and publicly-funded research organisations, and that AI represents an important technological tool in and of itself, it is vitally important that the needs of research and development (R&D) and the research performing organisation (RPO) sector are taken into account. This is essential if we are to avoid inadvertent consequences and collateral damage that ultimately throttles innovation, knowledge valorisation and slows long-term growth and development.

This document proposes a number of principles which we believe should be applied in the making of AI regulation and policy. They represent a call for a nuanced approach that takes account of legitimate concerns and addresses these in the most appropriate and practical way. They support the EU's stated goal of making it a world-class hub for AI – boosting our research and industrial capacity in an environment of excellence and trust.¹

Background – AI, Research and Science, Public-Private Partnerships (PPPs)

In the popular imagination, AI is primarily associated with major tech companies. In reality, however, it is much more than this – it is both a crucial tool for research and the product of it.

It is worth underlining that the definition of AI itself is sometimes hard to pin down. From the enthusiasm of companies talking about AI in order to inflate share prices (“AI washing”), to clumsy efforts that suggest that specific types of AI – such as generative AI – are the entirety of the machine learning landscape.

In particular, of late the entertainment industry has been effective in capturing the debate. We already see the distorting effect this has had at the policy level, where current discussions are almost entirely focussed on generative AI models that produce artistic outputs, with seemingly no recognition of the fact AI is deployed across all sectors of the economy and society and the many positive contributions it brings.² This “copyright capture” of the debate is consistent with the entertainment industry's tradition of using intellectual property to overstretch its influence on how knowledge is accessed and used. Such digital policy capture is often to the detriment of the wider public interest in scientific and technological advancement.

¹ European Commission. (25/4/2018). Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. Artificial Intelligence for Europe (Communication COM(2018) 237 final). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:237:FIN>

² See the open statement from the IP Federation regarding the AI Code of Conduct discussions in the UK. This reflects similar levels of entertainment lobbying activities in Washington as well as in Brussels during the EU AI Act debate. <https://www.ipfederation.com/download/text-data-mining-tdm-uk/>

A further impact of this distortion is the narrative that copyright law represents a suitable tool for the detailed regulation of AI. While exceptions to allow TDM are vitally important, above and beyond that, we argue that using copyright as a tool to solve all the issues AI poses risks exacerbating market inequalities, reinforcing market concentration, and encouraging the production of poor-quality AI.

Crucially, in line with wider trends, policy must develop conscious of the fact that R&D in this space takes place within an ecosystem including both public and private actors. Much new and cutting-edge AI research is taking place in public research institutions in collaboration with private actors, and as part of knowledge transfer and valorisation initiatives.

This occurs in the context of the tens of billions of euros of taxpayer investment in research made each year across Europe, with a view to creating new businesses and jobs further down the line. The fact is that while the EU and national funding policies stress such collaboration and support commercial uptake of research, it is undermined by a European information access regime which artificially divides commercial research from non-commercial research.³

This makes R&D, PPPs and knowledge valorisation activities around AI in Europe far more complex and onerous in comparison to our competitors in the US, Japan, South Korea, Taiwan, Switzerland and Singapore, who make no distinction in copyright law between types of research. This is already having long-term negative implications for science, the economy and public welfare, and should be revisited if Europe wants to be a world-leader in AI.

Given the importance of both public and private sector research, it is vitally important that AI policy solutions avoid a one-size-fits-all approach. AI isn't just big tech versus the entertainment industry. To make Europe a world-class place to undertake AI, it is vitally important that the needs and specificities of the research sector and knowledge transfer are reflected in all regulatory developments.

This paper focuses on R&D and the scientific application of AI in particular, and sets out solutions for these sectors.

KR21 Principles on Research and AI Policy Formation

We believe that the principles set out below should apply when discussing and taking legislative and non-legislative policy decisions that have an impact on AI in research:

1. **The needs of science and research should underpin AI policy formulation:** it is perhaps an obvious point, but it should not be forgotten that research rights are not just fundamental human rights, but also that research is essential to the achievement of any number of wider social and economic policy goals. The insights produced by researchers – including those related to AI – save lives, boost understanding, and offer solutions to many of the existential challenges we face today. The needs of medicine, bioscience, environment, agriculture, technology, engineering, R&D and education, to name just some relevant

³ See, for example, Arts. 3 and 4 in the Copyright and the Digital Single Market Directive (2019/790), making this distinction in the context of AI. See also Art. 5.3(a) of the Information Society Directive 2001/29/EC: "Member States may provide for exceptions or limitations ...[for] scientific research ... to the extent justified by the non-commercial purpose to be achieved."

sectors, need to be central to AI policy formation. A one-size-fits-all approach driven by one industry must be avoided at all cost.

2. **Volume and veracity of data matter:** what determines the quality of AI are the three 'Vs' – velocity, volume and veracity. It is a matter of fact that, where appropriately processed, greater access to information leads to better results, while restrictions on the availability of content, as a rule, will increase the chance of bias and poor model predictions.⁴ While this may not be a major concern for the entertainment industry, it is in scientific areas such as medicine, health, environment, agriculture, etc. A reduction in access to “high-quality” information also makes it harder to identify poor-quality information. It is obvious that if a model cannot train on accurate news reporting, it cannot identify “fake news.”

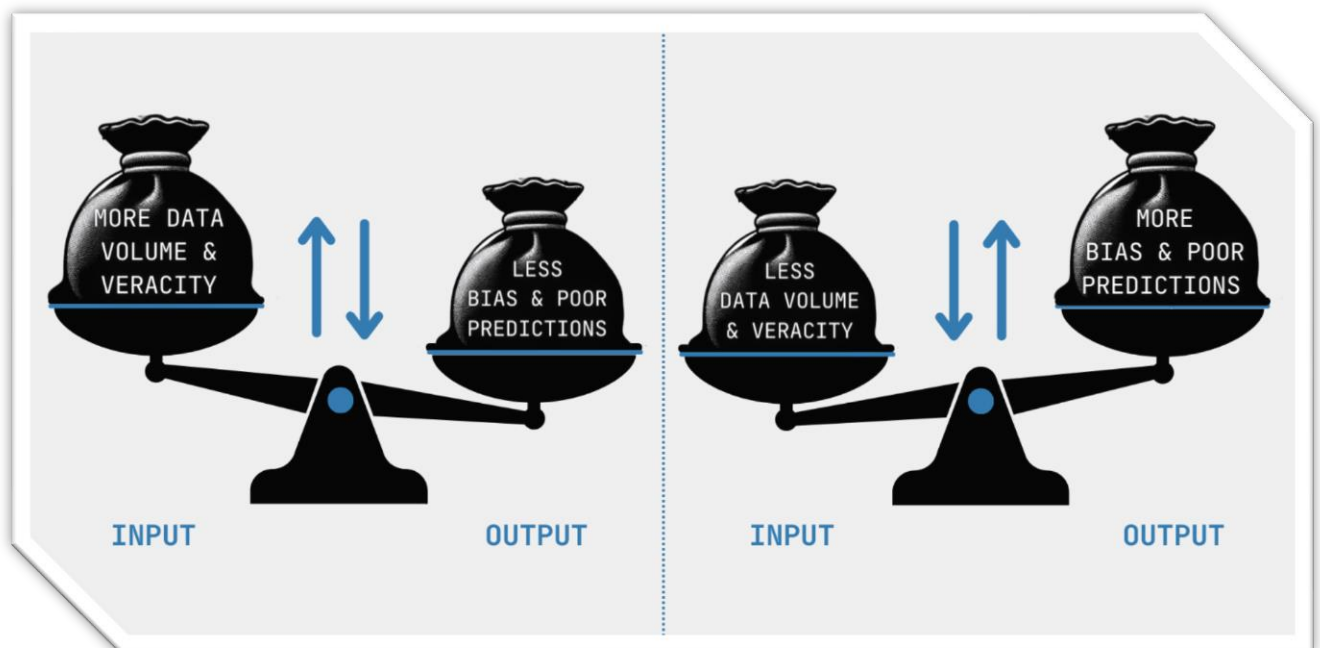


Figure 1: Illustration of the impact data volume and veracity on the input side have on the accuracy of AI's output

3. **Opt-outs increase bias and poor AI in certain sectors:** it follows that the possibility to use copyright law to opt out from enabling the use of legitimately and lawfully accessed content for training AI risks leading to more bias and less accurate outcomes. In the fields represented by the entertainment industries this will perhaps have few societal impacts. However, where the real-world application of AI is dependent on research and access to information required to build high-quality, safe and valuable products and services, opt-outs will have a negative impact on the quality of AI models and markets.

⁴ See Levendowski, Amanda. “How copyright law can fix artificial intelligence's implicit bias problem.” *Wash. L. Rev.* 93 (2018): 579. (arguing how flexible copyright exceptions in the US that maximise the use of information for AI training reduces bias.)

4. **Opt-outs reinforce market concentration and a lack of competition:** when copyright-based opt-outs require the *re*-licensing of information an organisation *already* has legal access to not only does it increase bias and poor AI, but it disadvantages smaller players and reduces competition.⁵ The result is that often, only the richer AI players are likely to be able to achieve the velocity, volume and veracity necessary to offer competitive outputs, reinforcing existing AI market concentration and disequilibrium. Furthermore, opt-outs act as an incentive for technology companies to establish themselves outside of Europe in more AI-friendly legal regimes.⁶
5. **We need transparency, not monitoring:** transparency, in the shape of reproducibility and validation in the case of research, is a key feature of good scientific practice. Research ethics developed in an academic environment have already evolved and are continuing to evolve to address the challenges of AI. The complexity of AI systems and questions of transparency raises important practical issues which need to be overcome in ways that do not place impossible burdens on researchers and organisations. Led by the research sector and existing transparency frameworks, relevant domain-specific codes and sectoral guidelines provide a tried and tested means of avoiding poor practice.
6. **Uphold academic freedoms and wider human rights:** importantly, AI transparency needs to be balanced with academic freedom, and in particular the right of researchers to pursue their own goals without the chilling effect of being constantly monitored and questioned by content owners or their representatives. Transparency is vitally important but there must be safeguards against excessive interference, as it will undermine scientific research.

Similarly, we are deeply concerned by the case made by some rightsholders from the entertainment sector that they should be able to have control over all aspects of the AI development and deployment cycle.⁷ This appears to reflect a wider trend towards using copyright as a mechanism to exercise monopoly control over the technological possibilities the digital environment offers. In the interests of scientific and technological development this should be resisted. The cultural policy goals of copyright are important, but they should not undermine scientific advancement and wider public welfare.

7. **Promote a pro-research and modern information access regime:** as a more general point, it is well-established that efforts to promote more flexibility in information access regimes, such as copyright, are seen around the world as a key means of supporting R&D, scientific innovation and its industrial application.⁸ An approach based on principles rather than prescriptions offers a more sophisticated way of enabling scientific progress and innovation while protecting legitimate author interests. In doing so, it makes it possible to get beyond chilling distinctions which undermine Europe's significant public investment in

⁵ It is worth noting that when the EU decided to enable text and data mining (TDM) for limited forms of research purposes, it was responding to the fact that *re*-licensing content they *already* had legal access to would seriously undermine AI research.

⁶ For example, the US and technology-oriented countries in East Asia, such as Japan, Singapore, Taiwan, etc, where these companies can rely on flexible copyright exceptions that don't allow for opt-outs and don't distinguish between commercial and non-commercial research.

⁷ See n.2 above.

⁸ Mendis, D., White, B., & Hong, D. (2024). Copyright and Open Norms in Seven Jurisdictions: Benefits, Challenges & Policy Recommendations. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4728782

science, such as those in the European copyright regime which distinguish between commercial and non-commercial research.⁹

In the context of copyright law and the facilitation of data-driven innovation, Europe, with the exception of Switzerland, is an outlier in the global economy by not supporting commercial AI research or knowledge valorisation.¹⁰

8. **Take an evidence-based and domain-specific approach to regulation:** we recognise that the spread of AI does raise concerns in areas of data protection, privacy and the challenges to science posed by the creation of synthetic text and images.¹¹ However, we believe that approaches targeted at proven problems sector by sector offer the best way forward. The imperative of applying existing practices around research ethics and transparency led by specialists in academia has already been mentioned. We have seen in the context of the EU AI Act significant amendments made to the original text (which itself did not benefit from an adequate impact assessment as regards research¹²). These hastily adopted amendments, lacking informed debate and analysis of their impact, raise a variety of questions relating to AI, research, R&D and PPPs.¹³

Better regulation processes which recognise that a one-size-fits-all approach to AI is damaging, will reduce the risk of the voice of a single lobby with a specific ill-fitting solution capturing the discussion and inappropriate regulation being developed.

Policy makers must identify domain-specific issues based on the understanding that AI is deployed across all sectors of the economy and society – there is no one-size-fits-all.

⁹ See KR21. (4/3/2024). Open Norms Support Research and Innovation – Time for Europe to Modernise. <https://www.knowledgerights21.org/news-story/open-norms-support-research-and-innovation-time-for-europe-to-modernise/>

¹⁰ Art. 4 of the Copyright in the Digital Single Market Directive (2019/790) only allows commercial research where a rightsholder has not reserved their rights. This contrasts with fair use countries like Taiwan, South Korea, Singapore, USA, Israel, etc. and Japan and Switzerland which allow commercial research and AI. The construction of Arts. 3 and 4 hinder nearly all forms of knowledge valorisation and knowledge transfer activity. Moreover, even for non-commercial AI research, the fragmentation of implementation at the member state level means that cross-border collaboration is rendered extremely uncertain in terms of its legality.

¹¹ See CESAER. (3/5/2023). Ensuring high-quality, transparent, open, trustworthy and equitable scholarly publishing. <https://www.cesaer.org/content/5-operations/2023/20230503-scholarly-publishing/20230503-cesaer-position-ensuring-high-quality-transparent-open-trustworthy-and-equitable-scholarly-publishing.pdf>

¹² See Science Europe. (31/1/2024). Open Letter: European Research and Higher Education Organisations Call On Commission Not to Neglect Their Needs in Lawmaking. <https://www.scienceeurope.org/our-resources/open-letter-european-research-and-higher-education-organisations-call-on-commission-not-to-neglect-their-needs-in-lawmaking/>

¹³ EU AI Act, Art. 52c.1(d)). The EU AI Office in developing frameworks for creating sufficiently detailed summaries should engage with institutions, libraries and staff who are involved in knowledge transfer and PPPs to ensure that transparency declarations are appropriate, sector specific and avoid chilling research.

9. **Support open source AI:** we welcome efforts to support the development of open source AI, both as a result of work in the research field, and as a tool to enable science. For example, we very much support that AI models produced through public funding be made openly accessible under legislation relating to secondary publication rights.¹⁴ Discussions here will benefit from greater clarity and consensus around what open AI actually is, and how it can be facilitated. We believe efforts around open AI should be augmented by research funders clearly mandating that publicly funded software of all types should be openly licensed and deposited in open access repositories – in the case of AI, this would mean the deposit of trained models and their parameters.

It is also of great importance that the software and algorithms behind AI used by public authorities, especially when it comes to informing decisions, are made publicly available as open source. This is essential for affected citizens and civil society, the media, etc to be able to monitor and examine the actions of public authorities.¹⁵

10. **Safeguard the open internet:** a discussion on AI and copyright is also a debate about how we use and shape the internet and therefore it is important to approach this debate conscious of the impact it will have on an open and free internet. Both public and private actors have used the internet from its advent to develop life-changing tools that serve society. It is also the largest research source of data and information in the world, crucially reflected in the CDSM Directive's reminder that AI models can be trained on content that is freely available online.¹⁶ Providers of all forms of search tools and interfaces use the internet everyday to generate better algorithms to improve access to knowledge. It is the world's most widely used source to identify new research, new and old phenomena affecting society as well as helping international collaboration, not least through the provision of translation software. Its international scope also means it creates a window through which bias in AI training can be reduced.

A free and open internet is important in itself, but the continued openness of the internet is also crucial for ensuring the positive evolution of AI models. AI models are dependent on a large and continuous flow of new data entering the digital sphere. Safeguarding the vitality of the internet is thus equally vital for the optimal development and performance of AI systems.

It is vital that the goal of an open equitable internet based on public interest continues to allow diverse voices to be heard and for research of all forms to continue.

¹⁴ See KR21. (February 2023). A Position Statement from Knowledge Rights 21 on Secondary Publishing Rights. <https://www.knowledgerights21.org/wp-content/uploads/2022/10/KR21-Secondary-Publishing-Rights-Position-paper-v1.1.pdf>; COMMUNIA association. (April 2024). Access to publicly funded research. <https://communia-association.org/wp-content/uploads/2024/04/Policy-Paper-17-on-access-to-publicly-funded-research.pdf>

¹⁵ See Kayser-Bril, N. (12/08/2019). Spain: Legal fight over an algorithm's code. *Algorithm Watch*. <https://algorithmwatch.org/en/spain-legal-fight-over-an-algorithms-code/>

¹⁶ Recital 14 of the Copyright in the Digital Single Market Directive (2019/790).

Further Reading

KR21

KR21 Primer on Artificial Intelligence ([source](#))

OECD

Artificial Intelligence in Science – Challenges, Opportunities and the Future of Research ([source](#))



European Commission

AI in science – Harnessing the power of AI to accelerate discovery and foster innovation: policy brief ([source](#))



About [KR21](#)

The KR21 programme is built on a conviction that **knowledge is essential for education, innovation and cultural participation**, and that everyone should have the possibility to access and use it, in particular through libraries, archives and digitally.

KR21 is a [Stichting IFLA Foundation](#) Programme, supported by the [Arcadia Foundation](#), in partnership with [IFLA](#), [LIBER](#) and [SPARC Europe](#).

Contact: info@knowledgerights21.org

FOLLOW US

