

Global Principles on Artificial Intelligence (AI)

August 2023

Introduction

AI developers and regulators have a unique opportunity to establish an ethical AI framework to boost innovation and create new business opportunities, while ensuring that AI develops in a way that is responsible and sustainable. To achieve this, it is essential that AI systems are trained on content and data which is accessed lawfully, including by appropriate prior authorisations obtained for the use of copyright protected works and other subject matter, and that the content and sources used to train the systems are clearly identified. This document sets out principles that the undersigned publisher organisations believe should govern the development, deployment, and regulation of Artificial Intelligence systems and applications. These principles cover issues related to intellectual property, transparency, accountability, quality and integrity, fairness, safety, design, and sustainable development.

The proliferation of **AI Systems, especially Generative Artificial Intelligence (GAI)**, present a sea change in how we interact with and deploy technology and creative content. While AI technologies will provide substantial benefits to the public, content creators, businesses, and society at large, they also pose risks for the sustainability of the creative industries, the public's trust in knowledge, journalism, and science, and the health of our democracies.

We, the undersigned organisations, fully embrace the opportunities AI will bring to our sector and call for the responsible development and deployment of AI systems and applications. We strongly believe that these new tools will facilitate innovative breakthroughs when developed in accordance with established principles and laws that protect publishers' intellectual property (IP), valuable brands, trusted consumer relationships, and investments. The indiscriminate appropriation of our intellectual property by AI systems is unethical, harmful, and an infringement of our protected rights.

Our organisations represent thousands of creative professionals around the world, including news, magazine, and book publishers and the academic publishing industry such as learned societies and university presses. Our members invest considerable time and resources creating high-quality content that keeps our communities informed, entertained, and engaged. These principles – applying to the use of our content to train and deploy AI systems, as they are understood and used today – are aimed at ensuring our continued ability to innovate, create and disseminate such content, while facilitating the responsible development of trustworthy AI systems.

INTELLECTUAL PROPERTY

- 1) Developers, operators, and deployers of AI systems must respect intellectual property rights,** which protect the rights holders' investments in original content. These rights include all applicable copyright, ancillary rights, and other legal protections, as well as contractual restrictions or limitations imposed by rightsholders on the access to and use of their content. Therefore, developers, operators, and deployers of AI systems—as well as legislators, regulators, and other parties involved in drafting laws and policies regulating AI—must respect the value of creators' and owners' proprietary content in order to protect the livelihoods of creators and rightsholders.

- 2) **Publishers are entitled to negotiate for and receive adequate remuneration for use of their IP.** AI system developers, operators, and deployers should not be crawling, ingesting, or using our proprietary creative content without express authorisation. Use of intellectual property by AI systems for training, surfacing, or synthesising is usually expressly prohibited in online terms and conditions of the rightsholders, and not covered by pre-existing licensing agreements. Where developers have been permitted to crawl content for one purpose (for example, indexing for search), they must seek express authorisation for use of the IP for other purposes, such as inclusion within LLMs. These agreements should also account for harms that AI systems may cause, or have already caused, to creators, owners, and the public.
- 3) **Copyright and ancillary rights protect content creators and owners from the unlicensed use of their content.** Like all other uses of protected works, use of protected works in AI systems is subject to compliance with the relevant laws concerning copyrights, ancillary rights, and permissions within protocols. To ensure that access to content for use in AI systems is lawful, including through appropriate licenses and permissions obtained from relevant rightsholders, it is essential that rightsholders are able effectively to enforce their rights, and where applicable, require attribution and remuneration.
- 4) **Existing markets for licensing creators' and rightsholders' content should be recognised.** Valuing publishers' legitimate IP interests need not impede AI innovation because frameworks already exist to permit use in return for payment, including through licensing. We encourage efficient licensing models that can facilitate training of trustworthy and high-quality AI systems.

TRANSPARENCY

- 5) **AI systems should provide granular transparency to creators, rightsholders, and users.** It is essential that strong regulations are put in place to require developers of AI systems to keep detailed records of publisher works and associated metadata, alongside the legal basis on which they were accessed, and to make this information available to the extent necessary for publishers to enforce their rights where their content is included in training datasets. The obligation to keep accurate records should go back to the start of the AI development to provide a full chain of use regardless of the jurisdiction in which the training or testing may have taken place. Failure to keep detailed records should give rise to a presumption of use of the data in question. When datasets or applications developed by non-profit, research, or educational third parties are used to power commercial AI systems, this must be clearly disclosed so that publishers can enforce their rights. Where developers use AI tools as a component into the process of generating knowledge from knowledge, there should be transparency on the application of these tools, including appropriate and clear accountability and provenance mechanisms, as well as clear attribution where appropriate in accordance with the terms and conditions of the publishers of the original content. Without limiting and subject to paragraphs 6 and 9, AI developers should work with publishers to develop mutually acceptable attribution and navigation standards and formats. Users should also be provided with comprehensible information about how such systems operate to make judgments about system and output quality and trustworthiness.

ACCOUNTABILITY

- 6) **Providers and deployers of AI systems should cooperate to ensure accountability for system outputs.** AI systems pose risks for competition and public trust in the quality and accuracy of informational and scientific content. This can be compounded by AI systems generating content that improperly attributes false information to publishers. Deployers of AI systems providing informational or scientific content should provide all essential and relevant information to ensure accountability and should not be shielded from liability for their outputs, including through limited liability regimes and safe harbours.

QUALITY AND INTEGRITY

- 7) **Ensuring quality and integrity is fundamental to establishing trust in the application of AI tools and services.** These values should be at the heart of the AI lifecycle, from the design and building of algorithms, to inputs used to train AI tools and services, to those used in the practical application of AI. A fundamental principle of computing is that a process can only be as good or unbiased as the input used to teach the system (rubbish-in-rubbish-out). AI developers and deployers should recognise that publishers are an invaluable part of their supply chain, generating high-quality content for training, and also for surfacing and synthesising. Use of high-quality content upstream will contribute to high-quality outputs for downstream users.

FAIRNESS

- 8) **AI systems should not create, or risk creating, unfair market or competition outcomes.** AI systems should be designed, trained, deployed, and used in a way that is compliant with the law, including competition laws and principles. Developers and deployers should also be required to ensure that AI models are not used for anti-competitive purposes. The deployment of AI systems by very large online platforms must not be used to entrench their market power, facilitate abuses of dominance, or exclude rivals from the marketplace. Platforms must adhere to the concept of non-discrimination when it comes to publishers exercising their right to choose how their content is used.

SAFETY

- 9) **AI systems should be trustworthy.** AI systems and models should be designed to promote trusted and reliable sources of information produced according to the same professional standards that apply to publishers and media companies. AI developers and deployers must use best efforts to ensure that AI generated content is accurate, correct and complete. Importantly, AI systems must ensure that original works are not misrepresented. This is necessary to preserve the value and integrity of original works, and to maintain public trust.
- 10) **AI systems should be safe and address privacy risks.** AI systems and models in particular should be designed to respect the privacy of users who interact with them. Collection and use of personal data in AI system design, training, and use should be lawful with full disclosure to users in an easily understandable manner. Systems should not reinforce biases or facilitate discrimination.

BY DESIGN

- 11) **These principles should be incorporated by design into all AI systems, including general purpose AI systems, foundation models, and GAI systems.** They should be significant elements of the design, and not considered as an afterthought or a minor concern to be addressed when convenient or when a third party brings a claim.

SUSTAINABLE DEVELOPMENT

- 12) The multi-disciplinary nature of AI systems ideally positions them to address areas of global concern. AI systems bear the promise to benefit all humans, including future generations, but only to the extent they are aligned to human values and operate in accordance with global laws. Long-term funding and other incentives for suppliers of high-quality input data can help to align systems with societal aims and extract the most important, up-to-date, and actionable knowledge.

The following organisations endorse these Principles:



AMI – Colombian News Media Association



Asociación de Entidades Periodísticas Argentinas (Adepa)



Association of Learned & Professional Society Publishers



Brazilian Newspaper Association (ANJ)



Digital Content Next



European Magazine Media Association



European Newspaper Publishers' Assn



European Publishers Council



FIPP



Grupo de Diarios América



Hungarian Publishers' Association



Inter American Press Association



Korean Association of Newspapers



News/Media Alliance



News Media Association



News Media Canada



News Media Europe



News Publishers' Association



The Japan Newspaper Publishers & Editors Association



Professional Publishers Association



WAN-IFRA