



AI and the future of journalism

An issue brief for stakeholders

Dr. Anya Schiffrin, Columbia University

KEY MESSAGES:

- Journalism organizations are embracing Generative AI as part of newsroom practice but wary of the economic impact of large companies profiting from their content if used without proper permission and compensation.
- Media outlets around the world have issued codes of conduct that stress respect for audience data, authenticity of content, disclosure when AI is used, transparency, diversity and integrity of information and the right to remuneration by the AI firms.
- Without agreements on copyright and intellectual property, the current business model of journalistic creation will be undermined and with it a serious threat to cultural diversity. The models themselves will be unreliable if not trained on quality information.
- As well as being affected by the growth of Generative AI, news organizations and journalists are covering the topic and educating audiences about the risks and potential benefits.
- Concentration in the AI sector will have profound implications for the entire world and competition authorities are weighing whether new regulations are needed.
- Creators, publishers and journalists have made the case for updating copyright regulations or at least stringently enforcing the ones that exist and making sure they take into account new technologies.
- Some journalism organizations are signing licensing deals with large Generative AI firms. In doing so, these firms are holding the line in requiring compensation and setting a market for licensing content.
- Key areas to consider include: preserving authenticity, diversity of languages and cultures and transparency of information. Maintaining competitive markets may help avoid the further deterioration of the cultural and news ecosystem.

I. Journalists cautiously embracing Generative AI

Developments in Artificial Intelligence (AI) and Generative AI are changing constantly. Governments, educators and the public struggle to keep up. Designing policies that will not get out of date seems almost impossible. Generative AI could transform (or even destroy) journalism as we know it, so the journalism community has been fully focused on many aspects of this phenomenon.



Generative AI refers to artificial intelligence systems capable of generating content, such as text, sounds, images, and code, based on input data. These systems use algorithms to produce new data samples that mimic the patterns and structures of the input data they were trained on. Current examples of generative AI tools include ChatGPT and Midjourney.

Source: UNESCO

- Many newsrooms have embraced Generative AI as a reporting tool. Generative AI is used for analyzing data and working with large volumes of material and techniques are being taught in many journalism schools and in workshops around the world. UNESCO has published a handbook on [Reporting on Artificial Intelligence](#).
- Journalism associations and newsrooms have published numerous guiding principles, codes of conduct, and charters relating to the use of Generative AI. These stress the importance of disclosure when Generative AI is used to create text or audio/visual materials shown to audiences.
- Accuracy and credibility are paramount for quality news outlets and worries about the “hallucinations” of Generative AI outputs are top of mind. Human involvement and oversight are essential whenever Generative AI is used in order to prevent the dissemination of inaccurate or misleading information. UNESCO has also pointed out the risks of “[oversimplification](#)” or highlighting of one relatively inconsequential point rather than a main idea.
- Journalists play a key role in educating the public about Generative AI and raising awareness about some of the risks. The dangers of deepfakes have been discussed widely in the journalism community, particularly in 2024 because of the danger posed to elections and to Democracy. Journalists around the world have become deeply involved in the media and information literacy movement.
- The dangers of deepfakes are personally experienced by many in the journalism community. False content has been created to discredit and attack well-known journalists such as Nobel laureate Maria Ressa and Patricia Campos Mello from Brazil. Indeed, UNESCO’s 2023 [report](#) on gender-based violence and generative AI warns that deepfake technology can exacerbate the problem of online violence against women.
- Governments and researchers struggle to predict the tremendous economic effects and potential effects on employment and the [labor market](#), education and even social stability if large numbers of jobs are lost with large consequences for income inequality. So too in the world of journalism, education and culture where there is uncertainty

as to which organizations will use Generative AI and how. It's already widely used in marketing and communications and jobs for writers, reporters, editors and translators have been cut or transformed as a result.

- The knock-on effects are also tremendous as the companies that own Large Language Models (LLM), and make the most comprehensive generative AI tools, may well end up being a few monopolies mostly concentrated in one or two countries. Even supply chain questions, such as access to chips needed for training and running the models, have taken on a new urgency. These topics will become key for regulators and governments as well as the private sector.

“The days when it mattered whether a company was third or fourth on the Google search page are over because AI agents will scrape all of the web to get results.”

Toshit Panigrahi, co-founder of licensing firm TollBit, which works with publishers to monetize their data.

- Another important subject is payment for use of intellectual property including the question of copyright for inputs into the LLMs. The failure to compensate adequately the originators of information and knowledge already used to train the LLMs may itself have large consequences for the information and knowledge ecosystem.
- Freeriding on cultural and journalistic creation is a key threat as it will undermine the incentives to produce information. Fair payments to publishers and artists will help safeguard the information and cultural ecosystem. Without it, we face model deterioration with potentially significant adverse consequences.

The end of search? Website traffic may collapse and revenue for creators as well.

The new generative AI-powered searches with Google, Open AI and others are developing could mean the end of traffic to destination websites. This will have an enormous effect for news outlets dependent on search traffic. Not only would advertising be affected as fewer eyeballs means less advertising revenue but subscriptions too as readers often buy subscriptions after being directed to the website. In response, some publications are working with Open AI to ensure that future search engines link to their content. In July 2024, the US magazine, *the Atlantic*, announced it is also developing a search engine and browser extension.



Large Language Models

(LLM) are advanced AI models trained on vast amounts of data to understand and generate human-like text. These models, such as GPT by OpenAI and Llama by META, can perform a variety of language tasks, including translation, summarization, question answering, and more.

Source: UNESCO



The effect of Generative AI on competition

There is a debate about how competitive the LLM business will be, with some believing that it will be “commodified,” at least for most uses, while others arguing that there will be only a few firms with enormous monopoly profits. Even in the latter case, there are concerns about monopoly profits—in the upstream “bottlenecks,” e.g. the production of the chips used in the LLMs.

Portugal's competition authority 2023 [report](#) warned that nearly all aspects of the creation of generative AI have implications for competitive markets.

- Access to data is key, but it's disproportionately in the hands of a few powerful firms, which gives them a competitive advantage, enhancing their already strong market power.
- The use of massive amounts of data at their disposal also can collide with GDPR and intellectual property rights.
- Computing power requires access to affordable cloud services and hardware, but the cloud market is already [highly concentrated](#), with significant portions in the hands of the same firms that dominate in data, providing another avenue for amplification of market power. Whether this means that barriers to entry are high is debated.
- Having know-how and being able to experiment are essential for innovation but lack of access to data, contracts that handcuff employees and prevent them from switching companies, and no-poach agreements between firms are all possible choke points that can strangle competition.

II. Background on Generative AI and the effects on content creation

At the macroeconomic level, AI represents a continuation of the advances of knowledge associated with the digital revolution, raising questions about the extent to which it will *replace labor* or *augment its abilities* (is it intelligence assisting, like the microscope and telescope?). In the latter case, it could raise wages, in the former, increase inequality. AI is different from earlier innovations (like robots and machines) in that not only are these machines stronger and able to compute more rapidly, but they are also able to learn, viewed by some as the central human capability. They can learn faster—at least in certain particular domains, such as playing a well defined game like Go.

The extreme version of worry is associated with the attainment by LLMs of generalized artificial intelligence, where AI has learning capacities that go beyond well-defined arenas. There is a wide disparity of views on whether and when such capabilities might be attained. We will not discuss this issue further in this brief; but were this point reached (referred to in the literature as ‘the singularity’), it would be transformative of our entire society.

LLMs can synthesize and analyze data from large numbers of sources, which may discourage individuals from consulting the original sources. Compounding the existing [financial difficulties](#) faced by media outlets today (and covered in UNESCO’s [issue brief](#) on media viability, Finding the Funds for Journalism to Thrive), the fear is that if LLM companies don’t [compensate](#) the original sources, this will inevitably reduce their income sources, drying up their funding, and thus the production of original and valuable information. But matters are worse: there is an old expression GIGO—garbage in, garbage out. The quality of LLMs can only be as good as the quality of the data on which they are trained. The LLM companies, however, have not fully taken this on board: they seemingly do not want to compensate for the knowledge that they use—there is accordingly the risk that they will “kill the goose that lays the golden egg.” And this problem is even more acute if there are multiple LLMs, for then none of them will take responsibility for their collective effects on the information ecosystem. There is even the risk that they will undermine the business model of the search engines, which themselves have been undermining the quality information ecosystem. Thus some journalism organizations are now looking at investing in their own search engines and trying to plan for the day when search no longer brings traffic.

LLMs potentially reflect, and even amplify, prejudices and distortions existing in existing data, the data on which they are trained. If that data reflects longstanding racial or gender prejudices, those prejudices will be reflected in the output of the LLMs. The delegation of decision making to AI undermines the transparency of decision-making processes, because even those constructing these models don’t know fully how they work.

Aware of these dangers, organizations around the world have issued both analyses and recommendations about how to handle generative AI. Many of these focus on principles of cultural diversity, [human-rights](#) due diligence, equality- and safety built into design, transparency and accountability when Generative AI is used. Organizations such as the World Economic Forum, Reporters Without Borders and different UN agencies have provided important guiding principles and charters.

For instance, in November 2021, UNESCO adopted the first-ever global standard on AI ethics, known as the "[Recommendation on the Ethics of Artificial Intelligence](#)". The Recommendation was followed by the publication of the [Guidelines for the Governance of Digital Platforms \(Nov 2023\)](#). Finally, this year 2024, the [United Nations published the Global Principles for Information Integrity](#). All of these instruments read in a comprehensive manner shed light on how to approach to governance of AI systems with a human rights and multistakeholder approach.



The most recent UNESCO approach to AI generated content platforms governance came with the **Guidelines for the Governance of Digital Platforms**, that clearly outline the importance of an independent governance systems whose operations and actions are aligned with international human rights standards, that is transparent and accountable, that has institutionalized checks and balances, that is open and accessible and that actively promotes cultural diversity.

Meanwhile platforms should perform due diligence, be transparent, accountable, provide tools to user to engage and act critically with their products and establish their systems and processes based on international human rights standards.

Stopping deepfakes: the role of Authenticity/Content provenance

LLMs are predictive models, i.e. given a sequence of words, they predict the next word that would follow, on the basis of the massive amount of data on which they have been trained. But in many areas, one needs accurate data, e.g. what time a particular show begins. Knowing that it is most likely at 7:30 pm doesn't suffice. LLMs can "hallucinate," simply making up references or attributing to, say, *New York Times* statements that were never there. Currently, there doesn't seem to be any system of accountability.

As a counterpoint to the false information, "hallucinations" and deepfakes produced by Generative AI, some have proposed that certification of accurate information is an essential tool. Some have suggested some sort of block chain or ledger that can track and demonstrate chain of custody of information so that people know where it came from and whether it was synthetically created. Watermarks placed on content have been considered though some freedom of expression groups object. The standards body is called the Coalition for Content Provenance and Authenticity and different firms and initiatives work to implement the standards.

[C2PA](#) was founded in February 2021 by Adobe, Arm, BBC, Intel, Microsoft and Truepic. The founders established a [Joint Development Foundation](#) project including more than 200 members to collectively build an end-to-end open technical standard to provide publishers, creators, and consumers with opt-in, flexible ways to understand the authenticity and provenance of different types of media. The Content Authenticity Initiative—led by Adobe, includes more than 3,000 members working to facilitate implementation of the C2PA standard by creating open source tools towards that end as well as sharing best practices. Project Origin was a similar community group to the CAI founded by Microsoft, New York Times and the BBC, focusing on newsroom implementation of C2PA. It is now working closely with the International Press Telecommunications Council, IPTC, to create and manage a C2PA compatible list of verified news publishers, and has published a list of trial participants including the BBC and CBC/Radio Canada. Their goal is to increase authenticity and transparency in digital content through an open provenance standard. Truepic is a US-based firm working with companies like Qualcomm, Microsoft, Hugging Face and more to implement the C2PA standards. (Sources: [Santiago Lyon and Mounir Ibrahim](#))

III. Draft legislation abounds

Draft legislation abounds and varies between region and country but in the USA (where many of the AI firms are based) there has been no legislation to date. States have passed their own AI laws and the US office of Science and Technology Policy has issued a Blueprint for an [AI Bill of Rights](#) while President Joe Biden issued an executive order. The European Union has issued a detailed AI Act and Mexico, China, Brazil, Japan have draft regulations. However, none of these regulations directly address the questions that the journalism community is focused on.

Type	European Union	United States	China	United Kingdom	Japan
Regulation Approach	Comprehensive law (AI Act) with a risk-based approach & strict requirements for high-risk AI systems.	Decentralized approach with sector specific laws and various agencies issuing guidelines and principles.	Centralized approach with strong government control over AI development & deployment.	A pro-innovation, flexible approach aims at fostering innovation while ensuring safety and ethics.	A risk-based & soft-law approach called " agile governance ".
Ethical Guidelines	A set of six ethical principles to foster trustworthy AI, such as human agency & oversight, transparency, diversity, non-discrimination & fairness.	Blueprint for an AI Bill of Rights focuses on fairness, transparency & accountability to ensure responsible AI use.	Ethical guidelines emphasize alignment with national goals and social stability.	The Centre for Data Ethics and Innovation (CDEI) issues roadmap to an effective AI assurance ecosystem.	AI Utilization Guidelines emphasizes ethical use of AI & fostering innovation.
Data Privacy	The General Data Protection Regulation (GDPR) applies to all industries & all personal data with robust requirements.	No federal law, but sector-specific & state-specific regulations like the California Consumer Privacy Act.	Stringent requirements with new laws like The Personal Information Protection Law and Data Security Law.	The UK GDPR , similar to the EU's GDPR.	The Act on the Protection of Personal Information (APPI) includes stringent measures regarding data breaches.
Funding	Significant development funding through Horizon Europe and Digital Europe programs.	Public investment through agencies like NSF & DARPA, heavy private sector investments, & academic research grants.	Substantial government investments, state-owned enterprises, and private investment.	National AI Strategy emphasizes both public & private funding. Gov't funding through UK Research and Innovation (UKRI).	Significant investment in AI through government initiatives like the Society 5.0 program.

Table 1: Comparison of AI legislation in the EU, United States, China, United Kingdom, and Japan

Rather it is the **European Union that is leading the way**, introducing a regulatory regime that balances human rights and freedom of expression with responsibilities and guidelines. In May 2024 the Council of Europe agreed on a [treaty](#) on the use of AI systems. The convention, which is binding for 44 states, includes requirements for transparency and oversight and delineates responsibilities to respect international law and democratic institutions.

The EU's [Media Freedom Act](#) includes a provision in Article 18 requiring very large online platforms to include a “functionality” for media service producers to declare that they do not provide “content generated by artificial intelligence systems without subjecting it to human review or editorial control.”

The European Union's [AI Act](#) has **created tiers of risk** each of which require different actions. The four categories are unacceptable risk, high risk, limited risk and minimal risk. Riskiest areas need to provide risk assessments and explanations of what data was used. Regulators are aware that the ability of the LLMs to sweep up huge amounts of information will likely have important implications for privacy rights and state (or private sector) capacity for surveillance and Article Five includes strict controls about the use of biometric data.

The media is not classified as high risk but will still be subject to some transparency requirements such as [disclosing](#) when chatbots are used. Content moderation processes and recommendation algorithms are also subject to [transparency requirements](#).

Countries in the EU will have two years to implement the regulations and the Government of Spain, in collaboration with the European Commission, launched the first sandbox—controlled testing environment—of the requirements applicable to high-risk artificial intelligence systems in the proposed European regulation on Artificial Intelligence (AI)¹. This initiative is part of the Spanish digital transformation strategy, called Digital Spain Agenda 2026², which is part of the National Intelligence Strategy³.

[The Brussels Effect](#), Professor Anu Bradford's term describes how EU legislation sets standards outside the EU. This “Brussels Effect” is already visible in Brazil where some of the EU regulations are reflected in draft regulation of AI. Senate Bill No. 2338/2023, being reviewed by the Brazilian Senate, includes requirements for data protection, human oversight, transparency and non-discrimination on AI systems including mandatory algorithmic impact assessment. Penalties include fines and liability for the companies.

UNESCO guidelines on AI, governance and diversity

POLICY	DETAILS
The 2005 Convention On The Protection And Promotion Of The Diversity Of Cultural Expressions	<p>Calls for safeguarding cultural diversity and recognizes that culture is a key part of social cohesion.</p> <p>Examines some of the key questions that may affect scientists, artists, educators and journalists.</p>
2021 UNESCO Recommendation on the Ethics of Artificial Intelligence	<p>The protection of human rights and dignity is the cornerstone of the Recommendation, based on the advancement of fundamental principles such as transparency and fairness, always remembering the importance of human oversight of AI systems.</p>
2023 UNESCO Guidelines for the governance of digital platforms	<p>Focuses on safeguarding freedom of expression and access to information while promoting platform transparency, accountability, due diligence, user empowerment based on human rights standards.</p>

1 <https://portal.mineco.gob.es/es-es/comunicacion/Paginas/entorno-controlado-de-pruebas-sandbox-del-Reglamento-Europeo-de-Inteligencia-Artificial.aspx>

2 <https://espanadigital.gob.es/>

3 https://spain.representation.ec.europa.eu/noticias-eventos/noticias-0/las-claves-de-la-nueva-ley-de-inteligencia-artificial-2024-01-25_es

IV. Worries about copyright

LLMs are trained on massive amounts of data, but so far, the Generative AI firms have been loathe to properly compensate the originators of the data. In the US, they claim that they don't need to—their use is covered under the “fair use” exemption (see below). Other countries don't recognize “fair use,” so the LLM use of their data is a blatant violation. The particular cases are legal matters and some newspaper publishers are suing Open AI. But whether the law should be changed to clarify that what they are doing is not fair use—that the originators of the data should be compensated—is a policy matter.

Publishers and writers in many countries have objected to the blatant disregard for copyright evidenced by the indiscriminate crawling or “hoovering up” data from the internet. April 6, 2024 *New York Times* reported that Open AI transcribed more than one million hours of content from YouTube videos and put into ChatGPT-4. Google did too.

Google also changed its terms of service so it could use user-generated content from public Google docs and restaurant reviews. Meta discussed buying Simon & Schuster publishing house. According to *The New York Times*, MIT physicist Jared Kaplan's 2020 paper argued that scale is essential for model training. Because LLMs need trillions of “tokens”, AI firms constantly scrape content from all over the web, disregarding intellectual property rights and the enormous efforts and expense that goes into creating the information—a classic example of “free riding”.

The main reason publishers around the world are worried about copyright protection is because the uncompensated use of publisher content threatens the viability of news outlets. Indeed the World Association of Newspapers August 2023 [Global Principles](#) argue that AI systems' *“indiscriminate misappropriation of our intellectual property is unethical, harmful, and an infringement of our protected rights.”*

It is expected that the new Generative AI search tools will damage revenue for publishers and creators even more than the older generation of search engines and social media did. There has been a major deterioration of revenues for legacy media because search and social media platforms that garnered the advertising revenue that had previously helped fund journalism. However, these often identified the source of the information and could drive some traffic to the legacy media. (More generally, the snippet provided on social media is all that individuals want) but with AI/LLMs there may be no attribution, and the likelihood of going to an original source may be even lower, with the consequences for the information ecosystem even greater.

Advocates are urging countries with compensation systems in place for publishers, such as Australia's News Media Bargaining Code and Canada's C-18 bill, to update them so as to include the generative AI firms that are using content without paying for it. They're also calling for “must-carry” provisions so that Generative AI firms which don't want to pay do not respond by dropping news as Meta has [has done](#) in Canada.

Valuing content

There is no global standard as to how to value content suggesting that perhaps UNESCO could usefully convene a meeting to discuss the economics of valuation. Mike Cragg, an economist who specializes in litigation and works for the tech consulting practice of [Keystone](#), notes that “Data used for training a status quo LLM base model, fine tuning such a model, training for a specific application using a retrieval augmented generation (RAG) model for real time or specialized application will all have different values based upon the application.” But while there may be disagreements about the precise value, what is clear is that the value is positive, and in some cases large. And so too for the potential harms to the information ecosystem.



Retrieval Augmented Generation (RAG) combines retrieval-based and generation-based approaches to improve AI’s ability to provide accurate and contextually appropriate responses. RAG systems retrieve relevant documents or pieces of information from a large dataset and use this information to generate more informed and precise outputs. This approach enhances the quality of the generated responses by grounding them in actual data.

A key issue in ascertaining how much should go to any publisher/creator is establishing where the data that is being used is coming from. Media outlets in Europe and the US are signing licensing deals for their content with Open AI, Perplexity and other companies. Some cash payments are being made but also the right to use the tools and tokens and design new products such as search tools.

News publishers had hoped for attribution and links to their content but it's not clear how the technology will develop. [Reports](#) have emerged that Perplexity and Open AI are unable to consistently cite and link to their sources.

One point of contention is whether the publishers or the creators should receive the payments from the large AI firms. In Brazil [musicians](#) have asked for a share of residuals. In Belgium, where European copyright directives call for creators to get paid while publisher negotiations with Google don't. Publishers feel strongly that they incur the costs and the risks and therefore payments should go to them. *Le Monde* signed an agreement with Open AI in 2024 and following the agreement, and in order to fulfill the neighbouring rights legislation, this leading French media organization decided to distribute 25% of the revenue related to the output to its journalists. Journalists can learn from other industries, such as music, where both the artists and their publishers have worked out arrangements.

Publishers are also trying to figure out what content is useful and how to tailor their content to the new prompts and a world in which [grounded data](#) (i.e. data that is grounded in information that has been verified) and up-to-date verified information are required. For example, the Axel Springer [agreement](#) with OpenAI isn't just about selling the rights to use the Axel Springer archive but will require Axel Springer to provide summaries, based on content in its publications, in response to ChatGPT prompts.

V. Journalists: Time to update Fair Use rules in key markets where AI companies are based

While AI companies claim their use of words or tokens lies well within the fair use doctrine, creators and publishers argue that copyright laws are needed to be updated or at least properly applied given the emergence of AI systems. The fair use doctrine is a US creation and was designed before Generative AI existed. It's not intended to deal with LLMs and web-based agents gobbling up masses of data — all of the data from, say, a newspaper — and then regurgitating something that embeds that data in ways that are not transparent about what the AI has hoovered up. There have been instances in which AI models come forth with whole passages, with or without attribution, and whole passages with incorrect attributions. Examples of such plagiarism has been well documented by writers and artists and in some of the lawsuits such as *The New York Times* lawsuit against Open AI.

Copyright regulations—journalism affected by the lack of global agreement on intellectual property and copyright protection for inputs	
US	"Fair Use" is a concept which doesn't exist in quite the same way anywhere else in the world.
Japan	Draft regulations there give no copyright protection for inputs into LLMs.
China	China's draft AI law provides limited copyright protection for inputs into LLMs. Article 21 permits the use of copyrighted materials for training AI models, provided that the use is justified and respects the rights of the original creators. However, the law primarily focuses on protecting AI-generated content and stipulates detailed guidelines for recognizing such content under copyright or patent law based on the user's contribution.
Europe	European Union countries have not implemented EU copyright directives in a consistent way. National differences remain.
UK	The current law does not provide explicit copyright protection for the data used to train LLMs.

Table 2: Copyright regulations in the United States, China, Japan, Europe, and United Kingdom

Data point

- Intellectual Property expert Giuseppe Mazziotti warns against the notion that generative AI can produce truly synthetic content with no implications for intellectual property. "Let's look at 'authorless works' with caution. At some stage in the production process there is often human selection, creation or involvement," he noted.



Different guidelines that have been issued

A number of media organizations, including EU Press Councils have issued ethical guidelines or added to existing ones. One point repeatedly stressed is the importance of transparency and human oversight. Indeed, the World Association of Newspapers released [Global Principles for Artificial Intelligence](#) in September 2023. Among other points, the principles call on AI developers, operators and deployers to:

- Respect intellectual property rights protecting the organizations' investment in original content.
- Leverage efficient licensing models that can facilitate innovation through training of trustworthy and high-quality AI systems.
- Provide granular transparency to allow publishers to enforce their rights when their content is included in training datasets
- Clearly attribute content to the original publishers of the content.
- Recognise publishers' invaluable role in generating high-quality content for training, and also for surfacing and synthesizing.
- Not misrepresent original works.
- Respect the privacy of users that interact with them and fully disclose the use of their personal data in AI system, design, training and use.

In July 2023, media scholars, journalists, publishers from all around the world came together in South Africa to draft the [Big Tech and Journalism: Principles for Fair Compensation](#). **These called upon platforms to be transparent and fair about their payments to news outlets, and support a plurality of outlets and diversity of views.**

Organization \ Aspect	Reporters Without Borders (RSF)	World Economic Forum (WEF)	World Association of Newspapers and News Publishers (WAN-IFRA)	Coalition of 17 News Organizations, including ICIJ
Policy	AI and the Right to Information	The Presidio Recommendations on Responsible Generative AI	Global Principles on AI	Paris Charter on AI and Journalism
Regulation	Advocates for strict regulations to prevent harms to the right to reliable information	Calls for a global AI governance initiative	Advocates for self-regulation with industry-wide standards and ethical guidelines	Calls for transparency, accountability, and ethical decision-making
Compensation	Right to opt-out from LLMs training; Fair compensation for content usage through independent & transparent distribution	Calls for updating of copyright laws to enable appropriate attribution, and ethical and legal reuse of existing content	Emphasizes intellectual property rights, adequate remuneration to publishers for use of their IP	Requires AI system owners to credit sources, respect intellectual property, ensure fair compensation for journalists
Disinformation	Calls for a clear accountability regime with criminal sanctions against harmful deepfakes	Promotes user feedback, disclosure of non-human interaction, content traceability	Promotes accountability for system outputs, including through limited liability regimes and safe harbours	Emphasizes human agency in editorial decisions, transparency in AI-generated content
Societal Implications	Highlights the need to safeguard journalism's role and prevent AI from undermining democratic values	Aims to balance AI innovation with ethical considerations to enhance societal benefit	Stresses publishers' role in generating high-quality content, the industry's sustainable development	Stresses the ethical use of AI to uphold human rights, peace, democracy, diversity

Table 3: Civil society and journalism recommendations

Spotlight on small languages

The [Windhoek + 30 Declaration of Information as a Public Good in 2021](#) reaffirmed the importance of information as a public good and also the importance of plurality and diversity. However, smaller news outlets will be unable to make money from generative AI and smaller languages are worried that they may be at risk. The LLMs are trained on large amounts of data—the larger the amount of data, the better they perform. But this may place societies that are small (Iceland) at a disadvantage. Will the LLMs accelerate the process by which English has become the dominant language? If so, what can be done to maintain cultural diversity?

The government of Iceland decided the best option is to give OpenAI lots of Icelandic content so it can be used in their models to make them more accurate and better quality. The government allocated 4 billion Icelandic kronur and the Miðeind language company has been creating spell checks and grammar checks in Icelandic.

“In many low-resource language communities, people realize and accept that our languages will not be adequately represented in cutting-edge language models if we are not willing to give up our data for their training. Many are of the opinion that language preservation must trump copyright considerations.”

Linda Heimisdóttir, CEO of Icelandic language company Miðeind

Other examples: creating small language models to archive and document a non-profit in New Zealand. Te Hiku Media, developed AI-trained language models to help prevent the Te Reo Māori language from further decline. Source: [How AI is helping revitalise indigenous languages](#) “ [AI Pirinka](#), created with a Japanese academic, is being used to preserve the unique language of the Ainu people, the indigenous inhabitants of Hokkaido in northeastern Japan. Google-funded [Woolaroo](#) (a UNESCO partnership) uses machine learning to teach and preserve languages like Yiddish and Louisiana Creole as well as showcasing speakers of endangered languages and how to pronounce them.

RECOMMENDATIONS

Recommendations for AI companies

Human-rights based governance. In any kind of regulatory arrangements, AI companies should be able to demonstrate the systems or processes they have established to ensure ongoing human rights due diligence, as well as risk mitigation measures. These systems should be reviewed periodically and the review should be made public.

Media viability and diversity – Generative AI companies should create transparent frameworks and standards for collaboration with publishers and creators with a focus on diversity and inclusivity (including cooperation beyond dominant, English-speaking outlets). They should actively collaborate and seek journalists, publishers and media outlets inputs to improve existing features and develop new products to support high-quality and pluralistic journalistic content.

Transparency - Generative AI companies should regularly report to the public and the governance system on how they adhere to the principles of transparency. This includes transparency with regard to data collection and web scraping practices.

Web Scraping - Generative AI companies which use web crawlers and other scraping techniques to collect non-personal publicly available data should provide websites and content owners with effective tools to prevent unwanted automated data extraction (partially or entirely).

Attribution – Generative AI companies should urgently focus on improving attribution mechanisms and enable users to identify and connect with journalists, media sources and publishers by accurately and systematically citing and linking their content.

Provenance. Generative AI companies should develop effective mitigation mechanisms to address the risks of disseminating misleading information and hallucinations. They should also support efforts to certify accurate information and its provenance (source and history) and consider mechanisms to highlight and promote high-quality certified media content.

Build in regular impact assessments for existing and new products and services and particularly as these affect media outlets, the safety of journalists and overall freedom of expression as well as the privacy of audiences.

Respect intellectual property by seeking permission before using copyrighted content for training and and pay for the use of that content as mutually agreed to and negotiated between the parties.

At a minimum, respect opting-out decisions and robot.txt and other programs designed to block crawling. Consider designing the system so that it is based on “opting-in”.

Create transparent frameworks and standards for payments to publishers/creators and disclose AI licencing agreements publicly.

Be transparent about where content comes from and what data was used in training as well as on an ongoing basis.

Provide citations, disclose inputs and cite sources of information

Where there are different legislative frameworks enacted across the world, apply the laws consistently across jurisdictions, selecting the most supportive of human rights.

In terms of being free from bias and misinformation, respecting privacy and promoting freedom of expression.

Recommendations for Publishers/Newsrooms

Adopt clear policies, or update existing ones, on the use of generative AI and communicate these to audiences. Require human oversight and involvement at all levels of the process.

At a minimum, the use of Generative AI to report and analyze data should be clearly labelled and disclosed to audiences. In the same way artificially generated outputs including images and audio should be labelled as such.

As media has the power to significantly frame the public debate and shape the discussion, journalism needs to ask key questions, such as how AI can work for society and how it should work, and how it is applied not only to the immediate sphere of journalism but to other sectors of society.

Protect the privacy of audiences. All use of personalized data should follow privacy laws and, in regions where laws don't exist, then news outlets should follow the international human rights standards on privacy and data protection.

Participate in developing Media and Information Literacy policies.

Recommendations for States

Governments can strive, through regulatory processes, to prevent AI tools from being controlled by only a few entities and ensure that their development and deployment adhere to international human rights standards, including to protect privacy, intellectual property rights, labor rights and freedom of expression.

Regulatory authorities must also be careful that broad limitations on AI-generated content could end up creating greater restrictions freedom of expression.

AI Governance systems should promote dialogue with media, including for the investment in independent news media, and support the media ecosystem by making data available and supporting actions to bolster media sustainability, diversity, and plurality.

Governments should promote media and information literacy to enhance positive engagement with the Gen AI platforms with the aim of empowering users, in particular groups in situations of vulnerability and marginalization.

In keeping with UNESCO's guidelines, AI regulation should be designed to preserve human rights and freedom of expression as well as plurality of voices.

Regulate to ensure that markets remain competitive while human rights are respected, that creators are credited and that sources are clearly labelled.

Recommendations for intergovernmental organisations

UNESCO could convene multi stakeholder meetings to discuss standards of valuing content and the rights of journalism organizations and content creators.

UNESCO has [recommended](#) that governments promote media and information literacy.

UNESCO also supports a holistic approach to supporting media viability including financial support for public interest news with the proviso that editorial independence is respected.

CONCLUSION

At this writing it feels that the future of journalism—and many other areas related to culture, education and—will be upended. The questions are infinite and not yet answered. It seems that the pace of change is tremendous and the journalism community—always ready to innovate and adapt—is bracing itself for yet another transformation.

As we think about how to navigate a possibly new world, it's important not to fall into the trap of thinking that government regulation will stifle innovation and therefore should not be pursued. Innovation is not an end in itself—there are innovations that enhance market power and exploitation. Innovation should be directed at increasing societal wellbeing.

This paper has identified a number of threats that AI poses to journalism, at least in some cases, promising regulatory frameworks that at least ameliorate these threats. As Guy Berger, the former policy and strategy director at UNESCO, puts it: “Cars could be cheaper without belts and bags, but customers would generally prefer those that have developed and installed such features—even when not required by regulation.”

The development of Generative AI and AI technology is still unfolding. Those in the industry often argue that it's a nascent industry and therefore its hands should not be tied too soon. Better to see how matters progress, and if there is a problem, address the problem when it has clearly emerged and been identified. We've argued that this perspective is wrong and dangerous. It's hard to undo monopolies and monopoly power, once established; it's better to prevent its emergence. Indeed, EU regulators stress the need for government to be able to regulate before firms become too large to regulate. We've argued, accordingly, that it's important not to tie government's hands too early on. Constant monitoring and assessment are required and the time to regulate is now.

References

IRCAI. (2024). Challenging systematic prejudices: an investigation into bias against women and girls in large language models. UNESCO.

UNESCO. (2023). Guidance for generative AI in education and research. UNESCO.

Further Reading

- In November 2021, UNESCO adopted the first global standard on AI ethics, known as the [“Recommendation on the Ethics of Artificial Intelligence.”](#)
- [United Nations Global Principles for Information Integrity](#)
- UNESCO organizes [the Global Forum on the Ethics of AI](#)
- [United Nations Global Principles for Information Integrity](#)
- [Forum on Information and Democracy working group on AI](#)
- [World Economic Forum has seven principles for human-centric AI](#)
- [The AI for Good Global Summit](#), organized by ITU in partnership with over 40 UN agencies
- [The WIPO Conversation on Intellectual Property and Artificial Intelligence](#)
- [Reviving Lost Tongues: How AI Battles Language Extinction](#)
- 1. [Newsrooms Are Already Using AI, But Ethical Considerations Are Uneven, AP Finds](#) by Forbes
- 2. [AI Journalism: Where Will the Rise of Automated News Writing and Fact-Checking Take the Industry?](#)
- 3. [Ethical and Legal Considerations in AI-Human Collaboration for Creative Writing](#)
- 4. [Artificial Intelligence in the News: How AI Retools, Rationalizes, and Reshapes Journalism and the Public Arena](#)
- 5. [Can journalism survive AI?](#)
- 6. [How the news ecosystem might look like in the age of generative AI](#)



With the support of the
UNESCO Multi-Donor Programme on Freedom of
Expression and Safety of Journalists (MDP)

About this brief

This brief comes as part of the UNESCO series World Trends in Freedom of Expression and Media Development. It discusses the challenges the development of Generative Artificial Intelligence poses to journalism, provides analysis of the key issues around its impact on copyright, working methods and business models and comes up with recommendations for different stakeholders on AI and the future of journalism. The brief is authored by Dr. Anya Schiffrin of Columbia University.

Funding for this work was provided by the Multi-Donor Programme on Freedom of Expression and the Safety of Journalists.

This document and the other issue briefs in the World Trends Report series on Freedom of Expression and Media Development can be downloaded at <https://www.unesco.org/en/world-media-trends/issue-briefs>.

Cover photo:

metamorworks/Shutterstock.com

Author:

Dr. Anya Schiffrin

Research lead:

Lei Zhu

Research help:

Isaiah Glick, Adalí Frias Deniz, Alexandre Azevedo.

Thanks to Guy Berger, Melissa Fleming, Courtney Radsch, Joseph E. Stiglitz.

Editorial coordination:

Guilherme Canela, Chief, Freedom of Expression and Safety of Journalists section, UNESCO

Kristjan Burgess, Senior Programme Specialist, Freedom of Expression and Safety of Journalists section, UNESCO

Ana Cristina Ruelas, Senior Programme Specialist, Freedom of Expression and Safety of Journalists section, UNESCO

Published in 2024 by the United Nations Educational, Scientific and Cultural Organization, 7, place de Fontenoy, 75352 Paris 07 SP, France ©UNESCO. This document is available in Open Access under the Attribution ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) License. By using the content of this publication, the users accept to be bound by the terms of use of the UNESCO Open Access Repository. The designations employed and the presentation of material throughout this publication do not imply the expression of any opinion whatsoever on the part of UNESCO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. The ideas and opinions expressed in this publication are those of the authors; they are not necessarily those of UNESCO and do not commit the Organization.

