



## **M20 Policy brief 2: AI's impact on the intellectual property rights of journalists**

**By Dr. Anya Schiffrin,**

**Columbia University's School of International and Public Affairs**

### **Overview- AI in a world of Solidarity, Equality and Sustainability**

Information integrity is key to economic development, governance and the stability of international relations. So too, is the intellectual property regime, which is why international agreements on Intellectual Property (IP) and copyright have existed for decades. The rapid growth of generative Artificial Intelligence (AI) has upended the global copyright system, necessitating a robust discussion of enforcement and protections for journalistic content used in Large Language Models (LLMs). Without a change of course, the current system, under which LLM/AI firms use journalism content without paying for it, will undermine the production of original news and erode the much-needed quality news ecosystem.

### **Introduction – Why innovation and the intellectual property of media creators are important to the G20**

AI has become global; therefore, global rules matter. Urgent global cooperation is needed to set out the norms and rules, especially as they apply to AI and copyright. The G20, representing the world's most powerful economies, is the natural place to begin these discussions, as complicated and contentious as they may be. This is especially relevant to the G20's [Digital Economy Working Group](#), which has already held a Workshop on generative AI and its ability to “produce high-quality deep fakes at a lower cost, and the impact on information integrity, and consideration of possible

recommendations”. It is also relevant to the G20’s Task Force on AI, data governance and innovation for sustainable development.

The economic development of societies is closely related to innovation and adoption of new technologies. Past agreements have tried to promote innovation and access to knowledge while also respecting the rights of creators. Without this balance, there will be no incentives for the creation of new ideas and tools. This is relevant for G20 countries that hope to promote innovation in a sustainable way, particularly with the use of AI systems.

The quality of the information ecosystem depends on the production of credible information, with AI promising better analysis and, possibly, dissemination of data. However, current arrangements attenuate incentives to produce high-quality information and thus are likely to seriously undermine news media, which desperately needs traffic and revenue. Without quality information, societies cannot function and democracy will be undermined.

The G20 is focused on economic development and political stability, and the unguided growth of AI is a threat to both. While AI promises economic progress, it could destroy jobs, increase inequality, and worsen the information ecosystem; all of these would have adverse economic and political effects.

There are also distributional consequences both within and between countries, especially with AI companies being concentrated in a few large countries, and the focus of their training being centered around the most widely employed languages and country datasets. This has implications for journalists too, as the news they produce ends up in LLMs without compensation.

Discussions about AI are similar to earlier discussions around social media platforms unfairly benefiting from news content. This time around, there is a new urgency because the pace of change will quickly bring about dramatic shifts to the information ecosystem.

## **Key issues for G20 discussions**

People in the news media and the public who recognize these new threats may want to agree on a shared position on copyright and journalism in the age of AI. Several “buckets” of issues need to be addressed to build consensus.

We list the issues, with suggestions for consensus positions and work programs for G20 task forces.

### *Intellectual Property*

Access to information has been good for economic development. However, the idea of “fair use” has been taken to an extreme and resulted in [theft of local intellectual](#)

[property](#) as LLMs run by massive AI corporations take intellectual property produced by others without fair compensation. Then they produce directly competing results. This was true not just in the training phase but also in the constant updating needed for the models as [Retrieval-Augmented Generation](#) (RAG) necessitates constant scraping of the web by “AI agents”. In other words, theft of intellectual property is ongoing and constant. In the same vein, the AI responses to search questions are replacing traditional search tools. Answers provided by AI searches have already eroded traffic to news sites and will likely cause even greater [destruction of traffic](#). Google’s new AI Overviews [tool](#) and AI Mode are just two products already shown to undermine traffic referrals to news sites.

AI-generated answers are often so complete (albeit sometimes incorrect because of the absence of fact-checking and the technical creation of hallucinated connections and content) that people don’t feel the need to visit a news site to get more information. Many chatbot interfaces to LLMs don’t provide links to click for more information. If the LLMs did reveal their sources, they might be asked for payment -- but to the extent that they avoid recognition, they will lack credibility. Ultimately, they will have a choice – to disclose and pay, or to avoid news as curated quality data.

### **Implications for AI-generated search**

However, while traffic to news sites is likely to continue to fall, the LLMs rely on quality news and the ongoing theft of IP will undermine incentives for the production of quality news. If all the value goes to the LLM firms, news publishers will battle to survive. Without quality information going into AI systems, there cannot be good information coming out, no matter how plausible, and even “brilliant”, the AI products. In this respect, AI will be worse than search engines and social media, because they typically do not refer readers to original sources.

#### *Mis- and disinformation/hallucination*

Both intentionally and unintentionally, AI may well worsen the information ecosystem by engaging in massive [pollution](#), injecting and [spreading mis and disinformation](#). AI has enabled deepfakes, making the [detection](#) of mis- and disinformation harder. It has enabled better [targeting](#) of mis- and disinformation, making the pollution of the information ecosystem more dangerous. The situation will worsen when incorrect AI outputs re-enter the system as new training data, at the same time as news media inputs become scarcer.

#### *Market power*

With data being the central input into AI, those firms with more data may have more powerful models, and be able to exert more market power. This means there may be attenuation of competition not just in the AI market, but other markets too. The problems could turn out to be even worse than evidenced in social media and search

platforms, which amassed enormous data for their content and advertising operations, and use this data to feed into their generative AI facilities -- like Grok, Meta AI and Google's Gemini.

AI firms (including previous big players in social media and search) have been ingeniously developing new ways of [amplifying and entrenching](#) market power, and [worked hard against](#) attempts to circumscribe this power. It is hard for individual media houses to stand up against these giant corporations that have taken and continue to take their content, even in defiance of technical signals like Robots.txt. This is surely a competition issue, particularly since the extraction also produces a rival service to the victims.

### *Regulation and Pace of Innovation*

Careful regulation, such as watermarking and labeling of content as well as spending to ensure accuracy of AI-generated results, can help create user trust in content provided by LLMs. Fair payments also provide incentives for publishers to continue providing accurate information for society. Without these incentives, the deluge of AI-generated [slop](#), which is already damaging the information ecosystem, can be disastrous for democracy. To prevent this from happening, the government has a role to play by enforcing copyright and payments for high-quality data, such as news. Collective agreements and collective negotiations with the AI firms will also be needed.

## Paths to fair compensation

### Three Key Considerations

**Fair Use policy:** The US has an expansive definition of [fair use](#), which is far broader than other countries. There is always a balance between promoting access to knowledge without undermining revenues earned from innovation. Fair Use was intended to share access to knowledge, but has evolved from allowing, say, college professors in Africa to use a couple of pages from an expensive US textbook, to wholesale theft of original content. The [TRIPS agreement](#) on intellectual property is part of the World Trade Organization. The G20 can consider how their efforts can reinforce the protection of intellectual property.

**Payment for IP:** Without incentives, there will be no quality information for LLM. They will continue to extract without payment or proper payment. This is terribly important not just for news media but for governments. *Solution:* Make AI firms pay for content. *Prepare:* for platform pushback/attempts to buy off opposition and derisive comments about how you don't understand how smart they are.

**Credit news sources and include links:** AI search and other tools need to show the source of information and provide accurate links as well.

Looking to the future and the creation of a fair system, we see four policy paths to compensation. The current free-for-all, where AI firms scrape whatever they find online and creators and publishers have no protection. This is essentially what happened during the training period of the AI models, but continues today as companies are continuously in need of new outputs to add to their datasets. A second path would be a strict policy of no AI company being allowed to draw from the intellectual property of news media. This seems unrealistic.

This leaves two most likely options, both of which involve paying for the use of IP.

1. **A fixed scale of pre-determined fees:** Payments to pharmaceutical companies during periods of compulsory licensing of medications are one example. (Compulsory licensing occurs, for instance, when there is an epidemic, and there is a need for the production of a drug beyond the level that the owner of the IP can produce. Other firms are given the right to produce the drug and sell it competitively, paying the IP owner a royalty.) Payment to musicians for each download is another example of a royalty or fixed fee for usage. The idea is that some property rights don't allow the owner to exclude others from using their property, so the property can be used, but must be paid for.

2. **Determining fees in the context of a fair-playing field negotiations framework:** The larger competitive environment has a direct bearing on the ability of publishers to negotiate as a sector, so a framework is essential. A situation where powerful tech companies are on one side of the bargaining table and relatively powerless creators and publishers, who exclude those already in licensing deals with AI companies, are on the other, will not lead to fair outcomes. For this reason, the Australian Competition Commission created the News Media Bargaining Code, understanding that the power imbalance affected the price paid (or not) for journalism content. Government intervention and collective bargaining among publishers can also help ensure that small news outlets are not left out of deals struck with social media platforms and/or AI firms.

## **Proposed text for inclusion in G20 outputs**

### *For the Digital Ministers 2025 declaration*

Recognizing the importance of intellectual property and the production of quality information, the G20 calls for fair compensation for publishers and supports existing frameworks (where they are in place) as well as the development of new compensation frameworks where necessary.

Compensation and copyright protections should be given not just for content used for training, but also for grounding LLMs and RAG, and other data used for grounding these models. Frameworks need to be flexible enough to cover future uses of journalistic content in the LLM and AI products.

Recognizing that accuracy and provenance are key to building trust in journalism, AI and LLMs, the G20 supports efforts to show the sourcing and provenance of the information used in LLMs.

### *For the Heads of State (“Leaders’ declaration”)*

Information integrity is essential to the functioning of G20 member countries and providing incentives for the production of quality information is key. However, the use of journalistic content in LLMs without compensation will undermine the basis – and hence innovation incentives – for the media to produce quality information. For this reason, the G20 calls for the enforcement of existing frameworks for compensation for news publishers and for updating such frameworks where necessary.

## **Four recommendations for the media**

- Collective bargaining with AI firms will increase the chances of fair payouts
- Developing pricing criteria for licensing of news to AI firms.



- Lobbying national governments for robust enforcement and/or new legislation to protect copyright and IP
- Collaboration with other outlets to build small language models (or large ones) for news publishers to use.

### **Acknowledgements and call for comments**

*This policy brief was commissioned within the framework of the M20 ahead of the G20 Summit.*

*Thanks to Joseph E. Stiglitz for his comments and Lei Zhu for her research*

*The M20 initiative is a “shadow” parallel process set up to intersect with the G20 processes. The M20 seeks to persuade the G20 network of the most powerful global economies to recognise the news media’s relevance to their concerns.*

*As a collaborative M20 document, this paper is a working, live document. Share your suggestions or comments for consideration: [M20support@altadvisory.africa](mailto:M20support@altadvisory.africa).*

*For more information about the G20 process, which is hosted by South Africa in 2025, visit the website [here](#)*

### **Further Reading**

- Intellectual property issues in artificial intelligence trained on scraped data ([published](#) by OECD)
- AI and the future of journalism: an issue brief for stakeholders ([published](#) by UNESCO)
- AI-Enabled Influence Operations: [Safeguarding Future Elections](#)
- [Dismantling](#) AI Data Monopolies Before it's Too Late
- Generative AI [Is Challenging](#) a 234-Year-Old Law
- Are Tech's Generative AI "Fair Use" Dominoes [Starting to Fall?](#)
- AI Search Has A [Citation Problem](#)
- How AI Is Being Used to [Spread Misinformation](#) – and Counter It – During the L.A. Protests
- Big Tech [pushes for 10-year ban](#) on US states regulating AI
- [Who is suing and who is signing](#)

**This policy brief can be republished under [Creative Commons licence](#), - ie provided that you credit the source, indicate any changes to the text, and link back to the original article on the M20 site.**



**The M20 for the G20**  
 An initiative by SANEF, MMA and partners  
[M20support@altadvisory.africa](mailto:M20support@altadvisory.africa)  
[Media20.org](http://Media20.org)

