N° 16 // December 2025

MEDIATION

Understanding through information and dialogue

Biannual publication of Fondation Hirondelle

AI news anchor introducing himself at the 5th World Internet Conference in Wuzhen (China). © STR / AFP

Championing an AI designed for the public good

As a technological revolution, the advent of artificial intelligence (AI) is often compared to the discovery of electricity. However, generative AI is moving faster than an electrical current, revealing numerous risks, especially with regard to trust in the factuality and relevance of created content.

Another challenge lies in helping journalists—especially those with languages and cultures neglected by these new tools—to get to grips with them, while applying them in a responsible manner and in the public interest.

Fondation Hirondelle is taking an active part in this ongoing process: we have implemented a directive on the use of generative AI for our media, and we are using language models to improve content analysis in the local languages of the areas where we work. In addition, we are considering how to optimise certain editorial tasks in order to increase the time available for fact checking and field reporting.

This issue of Mediation provides a glimpse of a precise moment in time: AI is developing so rapidly that today's tools may be outdated in the near future. However, what must remain current is the goal of defending digital models developed for the public good; models that compensate their sources (particularly journalists) and represent all sectors of society, especially the most vulnerable, who are already socially, economically and technologically excluded from the benefits of digital progress.

Caroline Vuillemin, General Director

INFORMATION IN THE AGE OF AI

AI was already being used to produce and consult information, but in 2022 it made a significant leap forward with the advent of open-source generative AI. As this technology makes it easy to produce and widely spread plausible-looking content (whether real or fabricated), journalists would like to see its individual and collective use regulated.

Open AI made ChatGPT available online for free in November 2022, and in the three years since, the use of open-source generative AI has become widespread. It can be used to produce all types of content from information found online, combining text, image, audio and video formats. It is an incredible, low-cost production tool, but what is its relevance for an audience seeking information, and for journalists whose role consists of delivering an accurate reflection of reality?

Whatever the model (American Open AI, Chinese DeepSeek, Swiss Apertus, etc.), the responses provided by generative AI depend mainly on the question asked, the data available, and the algorithm that guides its selection from this data. In an online context favouring dominant cultures, it is equally possible to receive either accurate information or a plausible but false idea of reality. As early as 2023, multiple disinformation campaigns were using generative AI to deceive, in order to discredit individuals and organisations and to influence election results.

Journalists quickly realised the extent of this phenomenon, its potential and the threat that it posed to the credibility of their work. NGOs, notably Reporters Without Borders, took the lead in creating charters and practical models for an AI use that ensures reliable and transparent information. This was an important first step, but can AI be used to make further improvements? While it has the potential to generate an infinite number of deepfakes, might it also be capable of fostering the production and widespread broadcast of quality information accessible to the largest possible audience? What technical conditions would be necessary for this to happen? What private and public investments would be required, and what energy parameters would pertain?

These ambitious questions broadly reflect those addressed by regulation that is still in its infancy. Should they remain unanswered, the outlook is grim: AI will be fed on various forms of disinformation and will worsen the spiral of information chaos, eventually creating social rifts which will have an even greater impact in fragile societies. To avoid this outcome, we decided to bring together, in this issue of Mediation, the reflections of Fondation Hirondelle and of experts on the use of AI for public interest information, in the global South as well as the North. ■

Interview

AI ENABLES A MORE DIVERSIFIED NEWS ECOSYSTEM

© DR

Urvashi Aneja is the founder and head of **Digital Futures Lab**, a Goa-based think tank dedicated to building "futures where technology serves everyone, not just a few". In this interview, she outlines how artificial intelligence is reshaping the dissemination of information in India, highlighting both its potential and its pitfalls.

Idol of the Hindu Goddess Durga decorated with the theme of Artificial Intelligence, Kolkata, September 2025. © SARKAR / AFP

Figures show that there is a strong attraction for AI news in India[1]. Which opportunities does AI present for the dissemination of information in this country?

Urvashi Aneja: One big opportuny is around linguistic diversity. These tools offer the ability for people who don't speak the dominant languages to access news in their local language and that's a big thing because India has more than 200 languages. Nonetheless, we're still far from these tools actually being able to perform well in all those languages. Another aspect is from a literacy perspective: you have low levels of literacy for a large part of the population in India. And the fact that you can now access these tools with multimodal formats means that there are new opportunities for people to participate in the information ecosystem, even while they may not have classic literacy. Both as listeners or viewers, and also as people who can produce news content in different formats. Third, AI also brings down the cost of news production. You don't necessarily need a big studio or a lot of people any more. Just a single person can shoot the story, record the interviews, edit it and publish it. Of course producing a video alone and releasing it on a platform was possible before AI. But AI has made these actions much easier for individuals or very small media outlets. And the lowering of cost of production means that a lot more voices, geographies, perspectives can potentially be represented, which is not when you have a more monolithic and centralized news industry.

What are the main risks of this use of AI in the news?

The really well-established problem is around misinformation and disinformation, with the growing ease of creating content that is misleading, intended to persuade people. The trouble is that many of the technical tools that are being put on the table to authentificate content, such as watermarks, are soon proving to be ineffective. And so we really don't know how to help people figure out what is real and what is fake. Besides, the quality of information produced by classic news organizations has declined, for a variety of reasons, including because their ad revenues are being decimated by social media and now being further decimated by generative AI. Most young people today in India are no more watching the news or reading newspapers or even news websites. Many just look at the generative AI results. And the trustworthiness of that has really declined. If we don't fix the problem, there's also the risk that states come down with very high-handed regulation that has an impact on freedom of speech, or otherwise let social media platforms become the arbiters to verify what is accurate.

The really well-established problem with AI in the news is around disinformation

How AI should be regulated to face those risks?

The answer is not going to come from regulation alone. We also need to push harder on technical solutions. We've got to deepen the ways to allow the viewer to see the provenance of the data, to check how a media piece has been designed at every layer of this creation, to identify and track the different elements of its composition…

But there's certain kinds of regulation that we definitely need. On social media platforms at least, one of the big issues is that content that is disinformational, sensational or defamatory goes viral. It's not that the content exists, but

(1) According to Reuters Digital News Report 2025, 44% of people in India feel comfortable with AI-generated information supervised by humans.

The larger fallout is that people stop engaging in public debate, because they constantly wonder if all that is real

that it gets amplified and everyone sees it, and it gathers more and more momentum. And that virality has a lot to do with the recommendation algorithms on these platforms, and those algorithms are optimized to target individuals and to share content based on people's very personalized kind of preferences. I think data regulation has a huge role to play there so that we're not then targeting people to show them that kind of content. In the EU, there is already a legislation that has this capacity. But the enforcement of it is a whole other challenge.

What about AI regulation in India?

There have been requirements for social platforms to ensure that they do not become spaces where disinformation and misinformation spread. Those guidelines have mostly been motivated by the need to reduce political misinformation and its impact on elections. A new set of rules issued by the Ministry of Information Technology now requires creators of AI-generated content to clearly label it as such. However, implementing these rules is far from straightforward. It is often difficult to determine what should count as "AI-generated." Many common tools for graphic design or text editing already integrate AI features. If someone uses such software to create a genuine news post for social media, does that qualify as AI-generated content? And if the same tools are used to create a fake image, does the same label still make sense? In that case, the label itself may lose much of its meaning.

Finally, AI is a massive change in the ecosystem of information. It will have impact that we don't measure right now. But I also have some faith in people's resilience and capacity for falling completely prey to what AI generated misinformation. Given the potential that you could now fabricate anything like Donald Trump riding a unicorn going into Iraq or whatever, people are not buying it. But the larger fallout is that people then stop engaging as much in public debate, because they begin to just constantly wonder if all that is real. That, I think, is the really worrying aspect of it. This has very serious long-term impacts for the health of democracy. ■

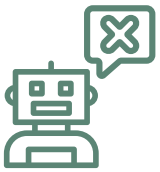
Countering disinformation and hate speech with AI

In Eastern Democratic Republic of Congo, Fondation Hirondelle led a pilot project using AI to detect and counter hate speech being broadcast on 10 community radio stations. With the help of natural language processing (NLP), radio programme content was recorded, transcribed in Swahili and analysed to detect violent speech. In response, Studio Hirondelle DRC produced and broadcast special programmes aimed at reinforcing social cohesion. With support from the European Commission's Service for Foreign Policy Instruments (FPI), this initiative was presented at the AI Action Summit in Paris in February 2025, and received an award at the AI for Good Global Summit in Geneva in July 2025.

Thanks to funding obtained by a new research project developed in partnership with the Federal Polytechnic School (ETH) in Zurich, this system of monitoring community radio with AI will now be extended to more than 100 radio stations in Eastern DRC.

Generative AI: an industry with limited factuality

20 %



Percentage of significant inaccuracies in the responses of major American AI chatbots to questions about the news.

Source: News integrity in AI assistants. European Broadcasting Union, October 2025. Responses to 2,709 test questions asked of 4 American AI chatbots.

A steep increase in electricity consumption

Is the planet able to produce enough electricity to power generative AI models for the daily use of the general public? The latest report from the International Energy Agency* contains some information that responds to this question. The report reveals that, on average, LLMs (Large Language Models) use between 0.1 and 0.5 Wh to generate an answer of a few lines in response to a simple question. Image generation consumes around 2 Wh, while video generation is much more energy intensive, using 115 Wh to generate a short, relatively low-quality video (6 seconds in length, at 8 frames per second): enough electricity to recharge two laptops.

This electricity is mainly used by data centres, which receive requests and respond by processing billions of pieces of data stored in the probability calculation database. These centres' consumption represented 1.5% of global electricity demand in 2024 and is set to rise to 3% of global demand by 2030. Generative AI is responsible for more than half of this consumption.

Will these alarming trends come up against other physical limitations linked to the carbon intensity of the electricity produced to supply data centres, the water used to cool them, and the metals used to manufacture the infinite number of chips they contain? The wider development of this technology also depends on the medium-term cost-effectiveness of generative AI models, which is still uncertain.

** Energy and AI, World Energy Outlook Special Report. International Energy Agency, Paris, April 2025.*

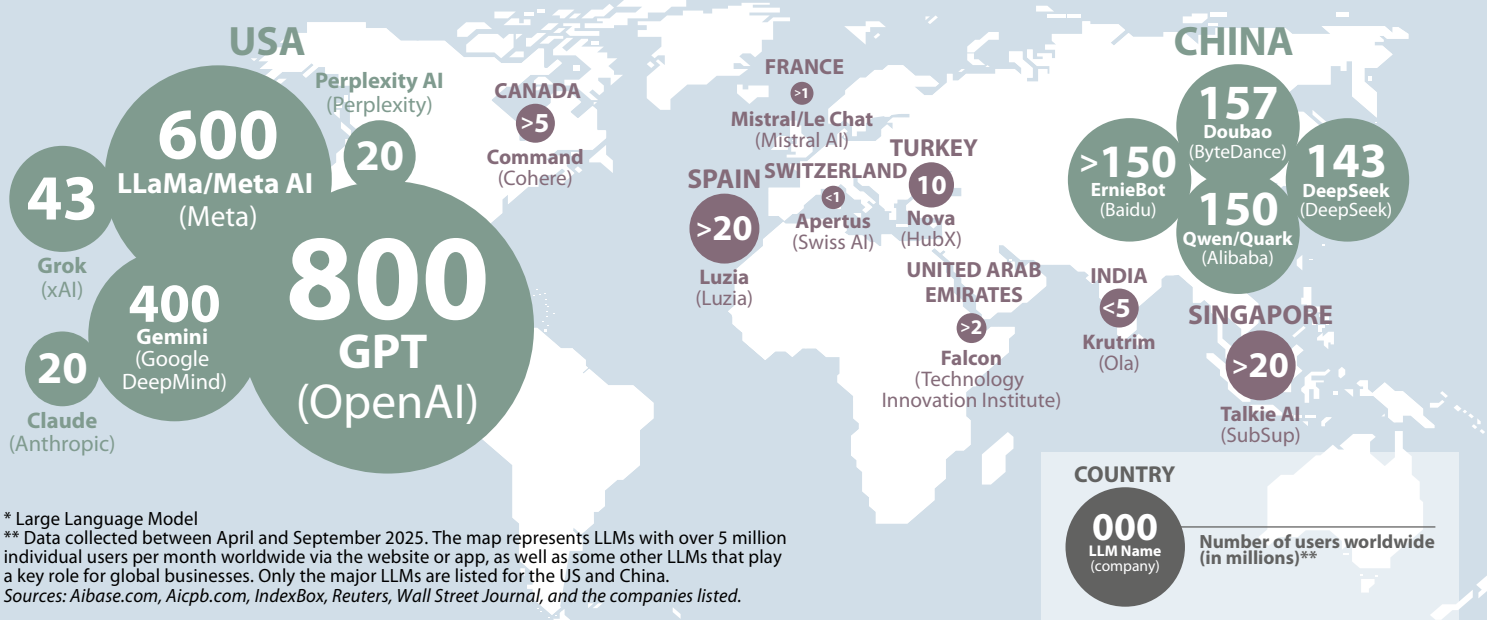
Recent and piecemeal regulation

According to Kent Walker, Google's President of Global Affairs, "AI is too important not to regulate and too important not to regulate well." From the EU's pioneering data protection regulation (GDPR) in 2018, to Joe Biden's Executive Order in October 2023 (revoked by President Trump in January 2025) and the EU's Artificial Intelligence Act of June 2024, lawmakers around the world have tried to produce legislation to regulate three stages of the AI process: the collection and use of data for AI, the use of algorithms, and published content.

However, these efforts rarely share a common spirit. "Regulatory approaches in different countries reflect their own socio-economic priorities, legal traditions, and governance models," says Amlan Mohanty, associate fellow at the Centre for Responsible AI based at IIT Madras, India. "The EU's rights-based approach, in the form of a comprehensive statutory law, seeks to 'protect health, safety, and fundamental rights.' China, with its strong desire for state control, prioritizes social order and the protection of 'Socialist Core Values' in its rules on AI-generated content. Japan has professed a 'human-centric' approach that aligns with its broader societal goals. Singapore and the United Kingdom have adopted a principle-based approach that reflects a pragmatic style, tailoring rules to specific industries." Adopted in September 2024, the Council of Europe's Framework Convention on AI represents an attempt at global consensus. Though it has been signed by most of the countries mentioned above, none of them have yet ratified it.

Big Data

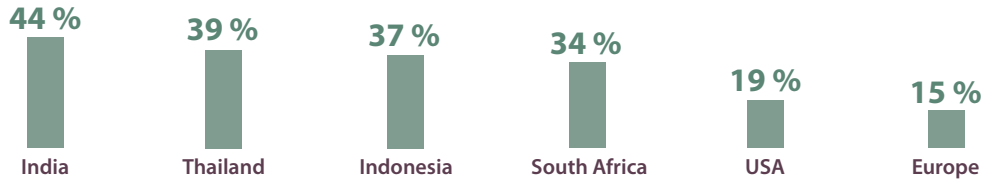
US and China dominate the field of chatbots and open-source LLMs*



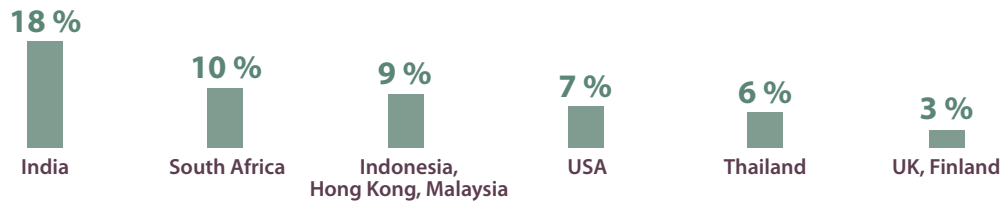
* Large Language Model
** Data collected between April and September 2025. The map represents LLMs with over 5 million individual users per month worldwide via the website or app, as well as some other LLMs that play a key role for global businesses. Only the major LLMs are listed for the US and China.
Sources: Aibase.com, Aicpb.com, IndexBox, Reuters, Wall Street Journal, and the companies listed.

Generative AI use is gaining ground among media users, especially in Asia and among young people

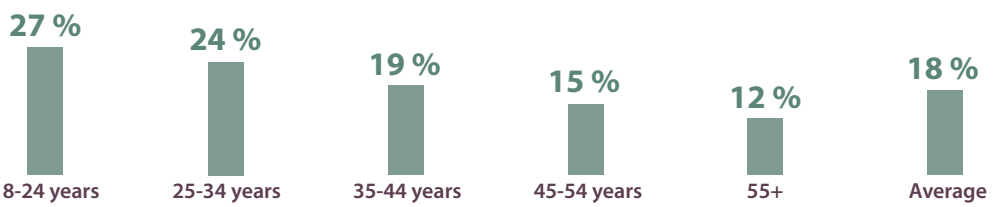
Proportion that say they are very or somewhat comfortable with news being produced mainly by AI with some human oversight



Proportion that uses an AI chatbot for news on a weekly basis



Proportion of the global population* interested in using an AI chatbot to get news



*Sample of 97,055 people in 48 countries on 4 continents.

Source: Reuters Digital News Report 2025.

Confidence in information generated by AI remains low

Net difference between the proportion of the global population* that thinks generative AI will make news more or less of each

*Sample of 54,638 people in 31 countries.
Source: Reuters Digital News Report 2025.

+ 29 % Cheaper to make

Transparent	Accurate	Trustworthy
- 8 %	- 8 %	- 18 %

Our experience

AI AND JOURNALISM: AN APPROACH CENTERED ON HUMAN RESPONSIBILITY

Video report by Studio Tamani at the "Namassa Danga" market in Bamako (Mali). © Harandane Dicko / Fondation Hirondelle

© Fondation Hirondelle

Fondation Hirondelle has adopted a directive defining how AI can, and cannot, be used in its newsrooms. Jacqueline Dalton, our Head of Editorial, reflects on the organization's stance and the principles guiding it.

Why did Fondation Hirondelle feel the need to reflect on AI?

Jacqueline Dalton: The discussion began when conversational agents like ChatGPT and generative image and video tools started proliferating. Our first step was to understand both the opportunities and the risks these technologies pose to journalism—especially in fragile contexts.

AI tools promise efficiency and cost savings, but they also carry significant risks

Because AI can amplify misinformation, we wanted to ensure that any use within our media would be transparent and responsible. Trust is the cornerstone of our relationship with audiences, so clarity about how and when AI is used is essential.

What were the main concerns?

An artificial voice would weaken weaken the connection with our listeners

AI tools promise efficiency and cost savings, but they also carry significant risks that clash with our commitment to independent, credible journalism rooted in local communities.

For instance, we considered using AI-generated voices to read the "community announcements" slot on our airwaves, which features messages written by individuals and small businesses. While this seemed practical and time-saving, we realized it would erode the authenticity that defines our media. Our audiences connect with real people—human voices they know and trust. An artificial voice would weaken that bond.

Beyond this, AI systems are shaped by biases embedded in their data and design. They aim to produce convincing answers, not necessarily accurate ones, and their sources often reflect cultural and informational imbalances—particularly in the countries where we work.

For these reasons, AI can only assist our teams, never replace them. Journalists remain fully responsible for editorial choices and for verifying every piece of content.

What are the key principles of the AI directive?

Fondation Hirondelle's approach is grounded in four principles:

• A Human-Centered Approach:

AI supports our journalists but never replaces them. Editorial responsibility and final validation always rest with humans.

• Transparency with the Public:

When content—text, image, audio, or video—is primarily generated by AI, this must be clearly disclosed.

• Respect for Quality and Integrity:

AI use must uphold our Code of Ethics and Professional Conduct. Factual accuracy and editorial quality are non-negotiable.

• Ethics and Responsibility:

All AI-generated material must be verified independently. Sensitive or personal data must never be shared with AI systems.

How are teams being prepared?

We've identified potentially useful tools and created a Directive to guide their use. But guidelines alone aren't enough. We are also training our journalists in a way that reflects their daily realities. It is likely that we will continue to update the directive as technology and newsroom experiences evolve. ■

FONDATION HIRONDELLE
media for peace & human dignity

Eyewitness

THE CONCENTRATION OF OPINION-SHAPING POWER MUST BE ADDRESSED

Based in Berlin and Zurich, the NGO AlgorithmWatch is working for a world where algorithms are used to make societies more just, democratic, inclusive and sustainable. Policy Manager Estelle Pannatier warns against the power structures lying behind AI and their impact on access to information.

Estelle Pannatier © David Baechtold

Algorithm Watch advocates for algorithmic systems that promote human rights. What does this mean regarding AI and the human right to seek and receive information?

Estelle Pannatier: The use of algorithms and Artificial Intelligence poses several challenges to access reliable information. Non-transparent algorithms curate information on social media and search engines in which AI is more and more being integrated. AI-generated content (text, image and video) influences public debate. Furthermore, there is an enormous concentration of market and opinion-forming power among a few technology companies providing those services. This affects individuals, who struggle to access reliable information and may withdraw from online debate forums. It also impacts the media and journalism, which are subject to the interests of new gatekeepers, such as AI chatbots, between the media and the public. Furthermore, it impacts society and democracy, as our access to information increasingly depends on a few large companies.

Our access to information increasingly depends on a few large companies

Various models of AI regulations have already appeared worldwide. Do some of them favor more the dissemination of information and quality news for all audiences?

Some AI regulations, such as the Council of Europe's Framework Convention on AI, address the impact of algorithms and AI on human rights and democracy, what is encouraging. However, we observe that AI governance often lacks a comprehensive vision. While it is necessary to set rules for the development and use of AI, a more holistic approach is required. These regulations should indeed also include measures to strengthen the media and journalism, encourage research and the development of alternative platforms and algorithms that promote democracy, and foster democratic competencies in society. Finally, the concentration of market and opinion-shaping power must also be addressed. Only in this way can we guarantee access to reliable, high-quality, diverse information, and encourage constructive public debate.

Algorithm Watch also advocates for algorithmic systems that promote a sustainable way of living. How do you do so?

Today, increasingly large AI models are being brought on the market. At the same time, the energy consumption of these models has reached unprecedented levels. Rather, we should design AI systems that are sustainable throughout their value chain and which bring real added benefit to individuals and society. This means taking ecological, economic, and social aspects into account when developing and using AI systems. As an example of our advocacy work, we recently launched a survey in collaboration with civil society organizations in several European countries. Thousands of people told us how they perceive the risks posed by the energy hunger of data centers and what action they believe needs to be taken. Ultimately, it also is important to take a sober view of the many narratives surrounding AI and to help shape the social debate from a public interest perspective. ■

We should design sustainable AI systems which bring real added benefit to individuals and society

Fondation Hirondelle is a Swiss non-profit organisation working to ensure that people facing crises have access to reliable, local, independent information. It has been supporting local media and journalists in fragile contexts for over 30 years (since 1995). With a global remit and local roots, it creates spaces for dialogue, to strengthen social cohesion in fragmented societies.

Impressum

Mediation
Biannual publication
of Fondation Hirondelle

Director of publication:
Caroline Vuillemin

Editors:
Julie Crenn
Benjamin Bibas /
la fabrique documentaire

Graphic design:
Marek Zielinski

Translation:
Julie Windebank

Printing:
Groux & Graph'style

Avenue du Temple 19C
1012 Lausanne, Suisse

hirondelle.org
info@hirondelle.org
T. + 41 21 654 20 20

Support us:
UBS Switzerland AG
BIC - SWIFT: CRESCHZZ80A
IBAN:
CH32 0483 5041 8522 8100 4

ISSN 2624-8840 (Print)
ISSN 2624-8859 (Online)