



# From Hype to Reality

Critical perspectives on AI

AREJA<sup>®</sup>  
Instituto de  
pesquisas aplicadas

  
The AI  
Accountability  
Network

 Pulitzer Center

# From Hype to Reality

Critical perspectives on AI

## Credits

### **President and CEO**

Lisa Gibbs

### **Research concept and executive coordination**

Flora Pereira

### **Authors and research leads**

Bianca Fermiano and Paula Goerg

### **Project coordination**

Bruna Wagner and Maria Kariénova

### **Digital researchers**

Alonso Balbuena, Gustavo Faleiros, Jonatan Rodríguez,  
Maria Kariénova, and Sarah Swan

### **Graphic design**

Natan Aquino

### **Editors**

Bruna Wagner, Donnalie Jamnah, Flora Pereira,  
Maria Kariénova, Marina Walker Guevara, and Sarah Swan

### **Proofreaders**

Alexandra Waddell and Dana Thompson

# Summary

## 1 Introduction and methodology pg 4

- Qualitative pg 6
- Quantitative pg 8
- Digital assessment pg 11
- Limitations pg 13

## 2 Executive summary pg 14

## 3 The landscape of AI reporting pg 16

- How did we get here? pg 19
- The journalists' input pg 22
- The good examples pg 31
- Journalists worth following pg 34

## 4 Pressing issues in AI today pg 36

- Labor and human rights pg 37
- Bias and discrimination pg 39
- Environmental impact pg 40
- Data governance and sovereignty pg 40
- Privacy and security pg 43
- Communication and (mis)information pg 44
- Complexity and fallibility pg 45

## 5 AI governance and regulation pg 47

- Brainstorming solutions pg 51
- Grievances and redressing mechanisms pg 52

## 6 Engaging the public and building awareness pg 55

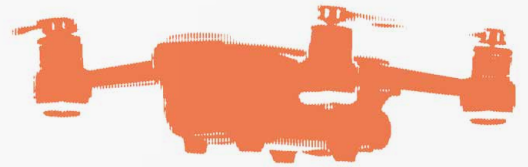
- Improving AI coverage pg 58
- What journalists would like to see pg 60
- The AI Accountability Network pg 62
- Rethinking education pg 64
- Engaging audiences pg 66
- Who needs to be engaged (and how)? pg 69
- The big, powerful, and influencing elephant in the room pg 73
- Measuring impact pg 75

## 7 Conclusions and recommendations pg 76

- Recommendations pg 78

## Appendices pg 80

- Literature Review pg 80
- Endnotes pg 82



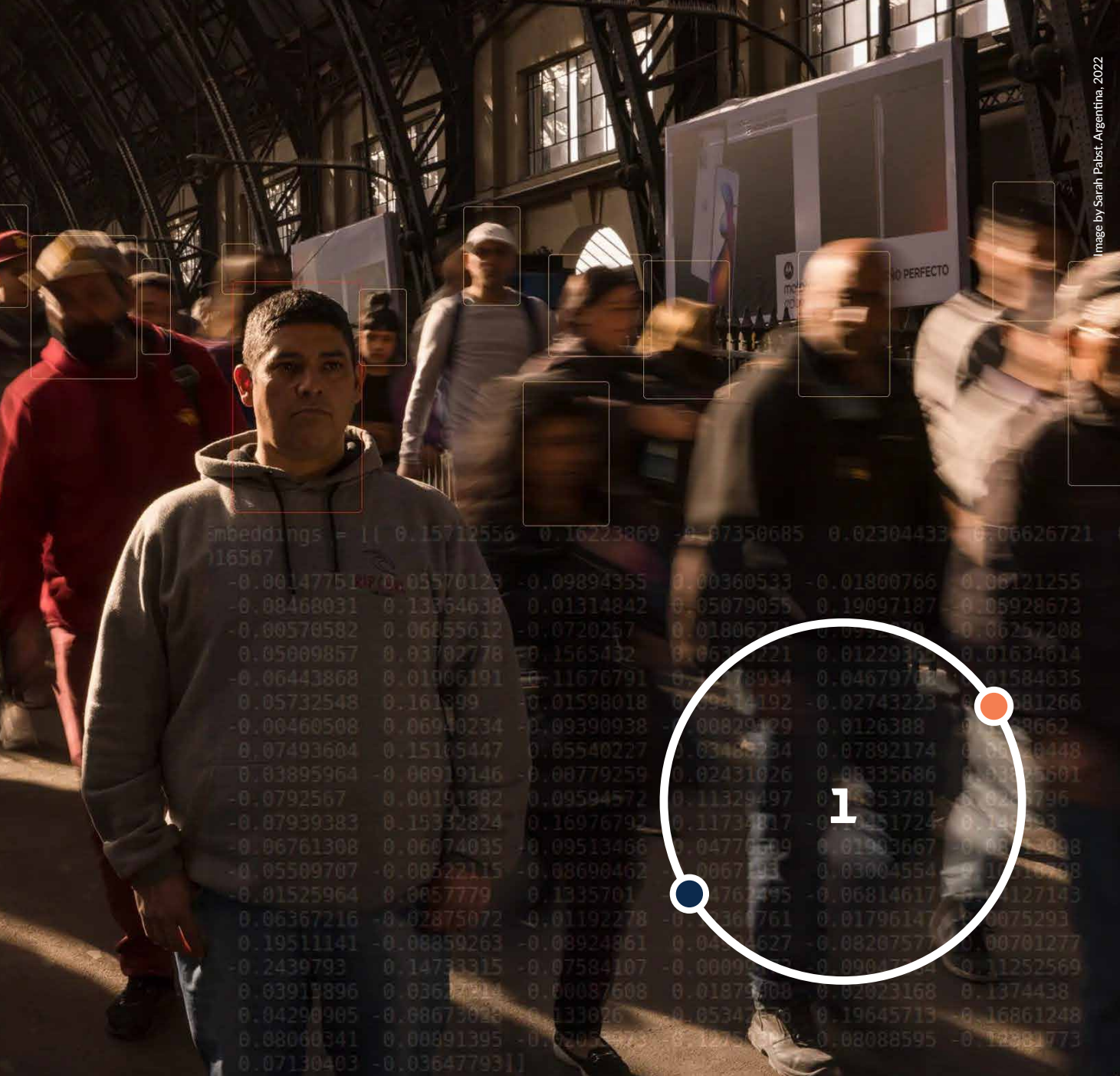


Image by Sarah Pabst, Argentina, 2022

# Introduction and Methodology

The ubiquity of artificial intelligence (AI) in modern society has significantly changed the way we live, work, and interact. From setting a thermostat to managing finances, from observing one's sleep patterns to monitoring crops, from translating a menu for a tourist to enhancing precision and control in a surgical procedure, AI has been woven into the fabric of our lives, creating the perception that “intelligent” machines could be a solution for nearly all of humanity’s challenges. Although these machines are capable of learning, adapting, and making decisions that used to be the sole domain of human intellect, there are still a significant number of issues surrounding AI that are not thoroughly discussed and understood. Nevertheless, the heightened enthusiasm for the technology’s potential in contemporary discussions is seldom accompanied by a practical view of limitations and ethical concerns, a phenomenon known as AI hype.

AI hype is particularly disseminated by industry and media narratives that emphasize the revolutionary powers of AI technology without critical scrutiny of the consequences of its development and deployment, nor the actual effectiveness of these tools. To an extent, the interest from tech companies in highlighting the benefits and potential of their products is fathomable. The lack of a more judicious engagement of the media stories, however, calls for an evaluation of the journalistic landscape surrounding AI issues and how that coverage could be improved. Journalism is indispensable for raising questions about transparency, ethical implications, regulation, and accountability in the deployment of AI systems. It also has the potential to demystify the complexities of AI technologies and educate audiences. When the media is caught in the hype, the already imbalanced relationship between a powerful industry navigating in a fast-paced, uncharted territory and the society that is supposed to benefit from it loses an important ally when it comes to safeguarding fundamental principles and rights.

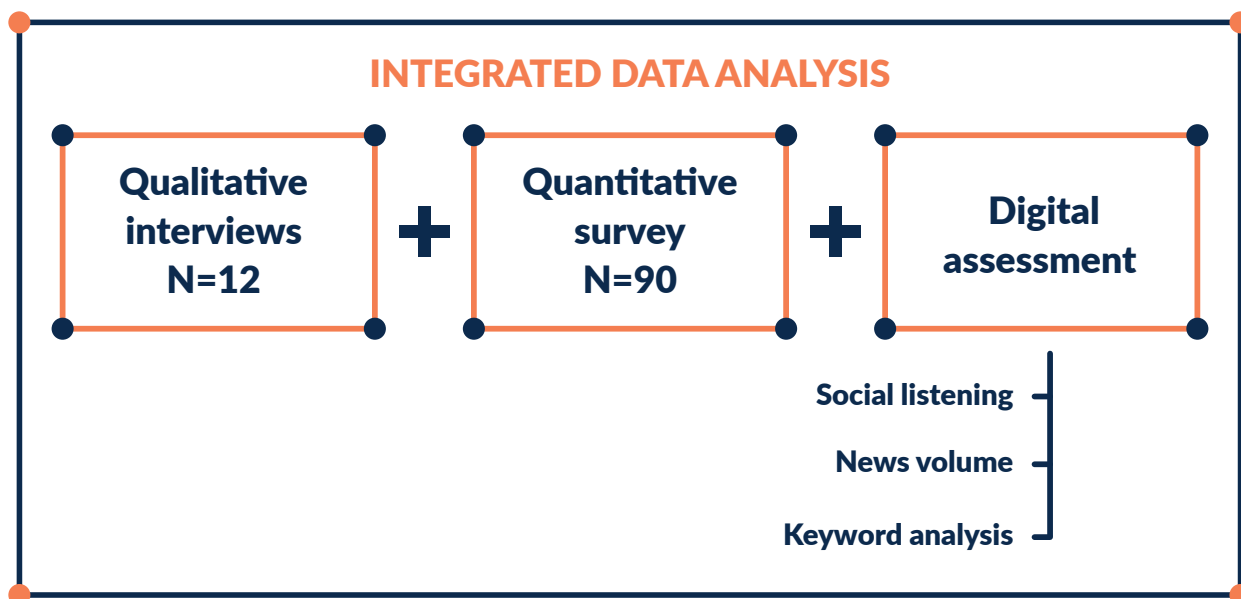
Aiming to address this shortfall, the Pulitzer Center has developed the AI Accountability Network, which seeks to foster robust and nuanced journalism about artificial intelligence and to develop a strong audience engagement program around the reporting to raise awareness and inspire action. This research study was planned and conducted to inform and strengthen the program’s strategies, considering the following goals:

- Establishing a baseline of knowledge across the organization to inform staff, partners, and stakeholders, and guide priorities on AI reporting and civic engagement
- Gathering the perspectives of a diverse group of experts about AI reporting, understanding key topics and gaps in coverage and knowledge
- Identifying opportunities to improve and increase coverage, build capacity, collaborate and investigate, both in terms of reporting and raising public awareness
- Understanding the central messages that need to be communicated around accountability and socioeconomic impacts while taking underrepresented voices into consideration
- Identifying key stakeholders and audiences and gauging public levels of knowledge and awareness while exploring ways to measure the impact
- Finding opportunities for engagement and networking

In order to obtain a more holistic understanding of these research goals, a mixed-methods study was designed and conducted between June and September 2024.



**Graphic 1** Research methodologies applied.



## Qualitative

In the qualitative step, a total of 12 in-depth interviews were conducted via videoconference with experts whose work is centered in varying AI-related fields, such as AI governance and policy advocacy, AI implementation and societal impact, AI in journalism and media, or human rights, and ethical AI research and development. This approach allowed for a comprehensive understanding of the pressing issues related to the topic and for the contextualization of the nuances and complexities of AI-related challenges and how they are portrayed in the media. Specialists from Africa, Asia, Europe, and North and South America contributed their perspectives, facilitating a wide range of insights into the current state of AI coverage worldwide.

The specialists invited to participate in the study all have worked in AI specifically or with emerging technology issues within areas such as journalism, policy and governance, research, human rights, and ethics. A more detailed description of their professional background is provided as follows:

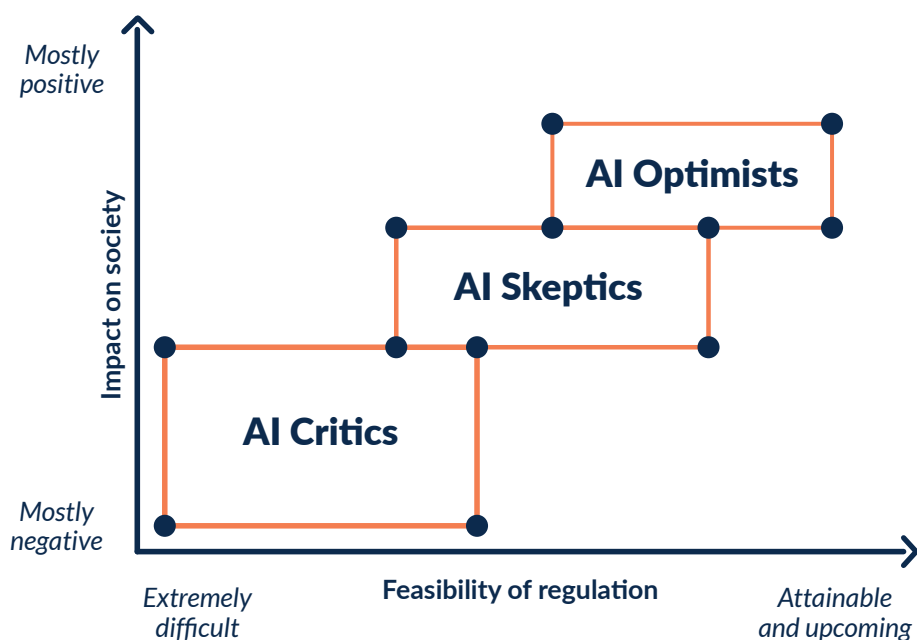
- An executive director of an NGO working with the promotion of digital inclusion and the safe use of technological tools. With a master's in economics and a doctorate in Social and Political Thought, this person has been a reference in the field of the internet and human rights. The NGO leader has been working in human development for over five decades and has assisted in the formulation of public policies for the ethical and responsible use of AI in the leader's region.
- A digital policy expert with a legal background, who contributes to discussions on AI governance and regulation, focusing on privacy, ethics, and inclusion. The expert's work impacts the formulation of guidelines that balance technological innovation and fundamental rights, influencing strategic decisions globally.
- A CEO of a data journalism laboratory in a local community within a large urban center, this person has been involved in collective actions and community organizations for nearly two decades. The CEO's work focuses on the intersection of technology and social justice, researching topics such as data colonialism, digital rights, internet governance, and data protection.

- A researcher with a doctorate in communication, who works in the fields of digital rights and AI governance. The researcher's contributions guide public policies for the ethical use of technology, promoting inclusion and digital rights, considering the sociopolitical challenges posed by artificial intelligence.
- A journalism professor and a researcher focusing on automation in the media sector, immersive technologies, civic data science, and governance of technologies. The professor's expertise is supported by a robust academic background, extensive research, and practical experience in journalism, contributing significantly to understanding AI's societal and technological impacts.
- A researcher specializing in AI governance, focusing on the responsibility and transparency of automated systems. The researcher contributes to initiatives that integrate engineering standards, ethics, and digital rights, guiding the development of public policies for emerging technologies.
- A lawyer specialized in digital rights, who works with strategic litigation to ensure fair and accessible internet in the lawyer's region. The person has been involved in lawsuits challenging legislation that restricts freedom of expression and contesting harmful content moderation practices by major tech platforms. Additionally, the lawyer advocates for the establishment of internet access as a fundamental right.
- A co-executive director of a digital rights NGO, who specializes in data governance and is a journalist with a master's in social sciences. The director collaborates with social movements and activist networks to build effective and fair regional technological alternatives and advocate for internet access as a fundamental right.
- A cognitive scientist who leads research in the field of ethical AI and alternative epistemologies, who works to promote a more critical understanding of the impact of AI systems on society. Their publications include studies on AI ethics, algorithmic injustice, participatory AI, data justice, and the decolonization of computational sciences.
- A technology consultant and executive supporting the development of prosocial technology, who assists governments and international organizations in adopting innovative technologies. The consultant's research seeks to balance technological innovation with social responsibility and sustainability.
- A policy and research expert focusing on technology, human rights, and public policies, primarily working in the civic sector, which includes NGOs and public interest organizations. The expert's experience encompasses security oversight, including security institutions and surveillance technologies.
- A journalist and analyst specialized in AI and ethics, who explores the intersection of algorithms and society. The journalist's stories highlight how emerging technologies shape global politics and economics, offering a critical perspective on the use of AI by large corporations and governments.

In terms of their assessment of the impact of AI on society, three main positions can be identified. A small parcel of the 12 interviewed respondents can be considered AI Optimists. Although acknowledging there is a problematic side to the AI hype and deeper issues to be brought to public attention, they believe the technology is necessary for progress and that it currently offers more benefits than drawbacks. Moreover, they are optimistic about the potential for effective regulation. A second and larger group may be regarded as AI Skeptics. They are more cautious when discussing AI impacts, presenting positive and negative aspects. There is a tendency to emphasize the challenges in regulation and the need for more critical coverage and transparency, but also to recognize that AI is necessary and an irrevocable part of the future. The last group is composed of AI Critics and is slightly larger than the prior. These experts view AI's impact as mostly negative, questioning its overwhelming and ubiquitous necessity, and expressing significant concerns about the feasibility of effective regulation. They focus on the ethical, social, and labor-related issues associated with AI.

## Graphic 2

Qualitative respondents' profile, according to their assessment of AI for its impact on society and feasibility of regulation.



The interview questions covered a range of topics aimed at understanding AI reporting and its societal impacts, including how it affects privacy, ethical concerns, and current media coverage, seeking insights on improving AI reporting, outreach and communication, and on demystifying common misconceptions. It also addressed gaps in coverage, public policies, industry accountability, and AI governance, and discussed strategies for engaging various audiences, and the role of an informed public in holding tech companies accountable.

## Quantitative

The quantitative survey was designed to systematically obtain data on journalists' experiences, challenges, and needs in covering AI-related stories. Providing an essential view of the current media landscape, this step allowed for an analysis grounded in the perspectives of the segment of professionals the AI Accountability Network aims to support and enabled the recognition of prevalent trends that could help explain existing gaps and opportunities in AI accountability reporting. The final sample totaled 90 responses from journalists from 32 different countries.

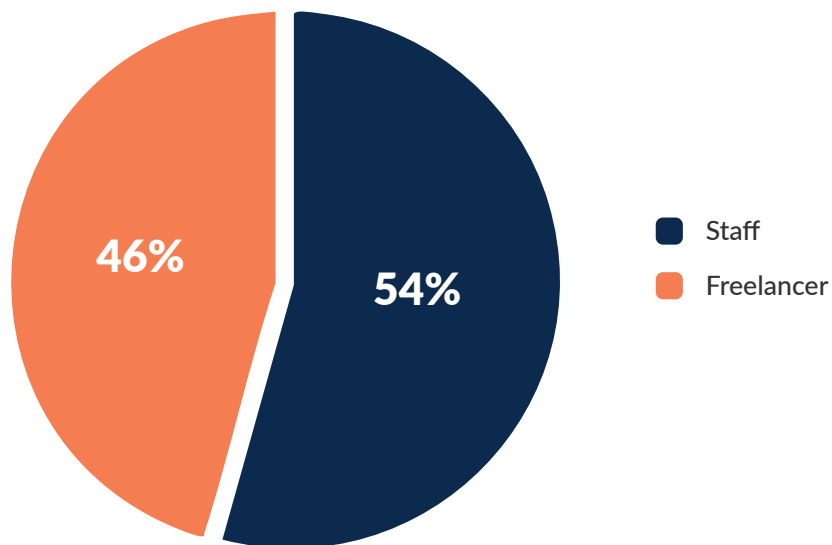
Journalists were contacted by email and social media by the Pulitzer Center and invited to participate in a voluntary survey using a link with the questionnaire, which was available in five languages: English, Spanish, Portuguese, French, and bahasa Indonesia.

The questions aimed to understand journalists' perspectives on the impact of AI technologies, the frequency and depth of AI coverage, and the main sources of information on AI issues. It also explored the most-reported AI topics, barriers to in-depth AI reporting, and the types of training and support needed by journalists. Additionally, the survey aimed to identify priority audiences for AI stories and the media outlets that consistently cover AI issues.

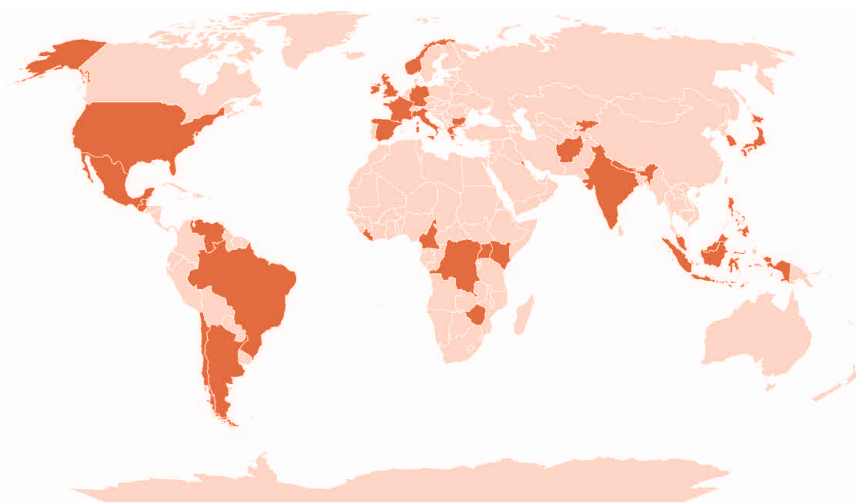
Just over half of the participants are part of an organization's staff (54%), and 46% are freelancers. Three in 10 journalists are from Asia (31%), with another segment of the same size being from North America (29%). Two in 10 are from Europe (20%), 12% from Africa, and 8% from South America. The United States is highlighted as the country with the largest number of respondents, with nearly a quarter of the sample (23%).



**Graphic 3** Employment status of the quantitative sample.

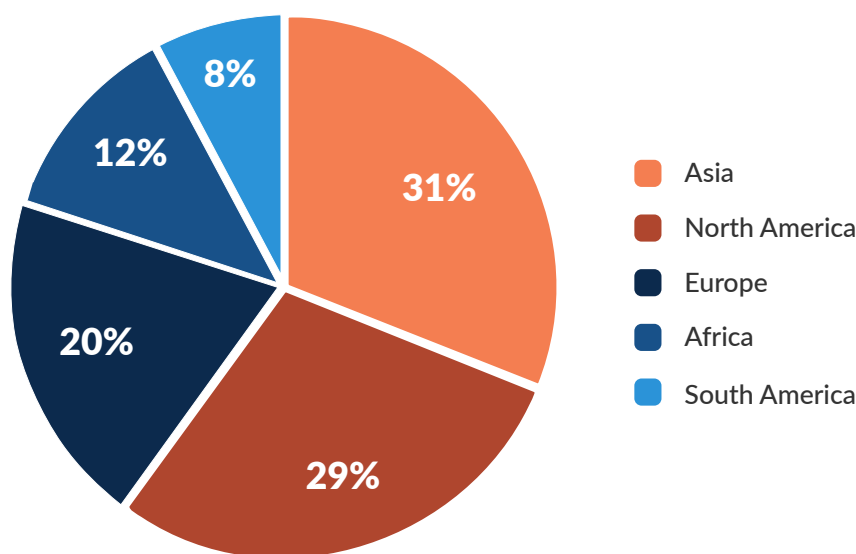


**Graphic 4** Country distribution of the quantitative sample (Frequency).

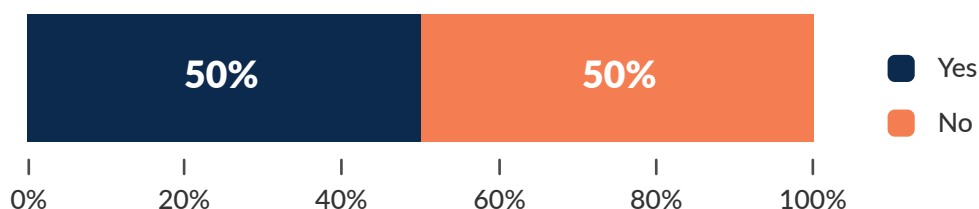


United States of America 21	Spain 3	Afghanistan 1	Japan 1
India 10	Uganda 3	Bulgaria 1	Kuwait 1
Philippines 6	Brazil 2	Cameroon 1	Kyrgyzstan 1
United Kingdom 5	France 2	Chile 1	Liberia 1
Kenya 4	Germany 2	Democratic Republic of the Congo 1	Norway 1
Mexico 4	Greece 2	Guatemala 1	South Korea 1
Argentina 3	Malaysia 2	Ireland 1	Venezuela 1
Indonesia 3	Nepal 2	Italy 1	Zimbabwe 1

**Graphic 5** Continental distribution of the quantitative sample.



**Graphic 6** Journalists in the survey who have written one or more stories about AI (N=90 journalists).



Half of the respondents have written at least one story about AI in the past. An analysis of the keywords contained in these stories reveals that “ethics” is the term most frequently mentioned as the topic covered. Other terms like regulation, prejudice, surveillance, transparency, technology's impact, data security, digital cloning, and fake news point to concerns about potential harm from the use of technology. Such a tendency, allied with the opt-in nature of the survey, indicates that the obtained sample might be better informed about the AI issues being discussed in the study, which limits generalization to a wider group of journalists. Nonetheless, part of the survey participants have also shared stories written about technological applications of AI, with stronger focus on benefits and solutions, and a reduced or nonexistent critical perspective.

## Graphic 7

Main keywords of the stories provided by journalists, keywords mentioned in two or more stories (N=45 stories).



## Digital assessment

The digital assessment consisted of three elements: a social listening exercise, an assessment of news volume, and a keyword analysis related to AI tools and resources using Google Keyword Planner. This approach enabled a comprehensive understanding of how the topic is currently being discussed and perceived across various digital platforms. The news volume analysis helped identify trends and the level of public interest, which are key for determining the relevance and timeliness of coverage. Social listening allowed for gauging public sentiment and engagement, uncovering the most pressing concerns and popular discussions related to the topic.

The keyword analysis pointed to the most searched and discussed terms, which can potentially ensure that coverage aligns with what audiences are actively seeking information about. Integrating these digital assessment tools into the analysis in this research study is crucial for developing strategies to create informed, relevant, and impactful stories that will resonate with the audience. A more detailed description of each element is provided as follows.



## SOCIAL LISTENING

Using BuzzSumo and Brand24, the online presence, engagement, and sentiment surrounding the keywords “AI bias”, “AI harm”, “AI technology”, and “artificial intelligence” were evaluated. BuzzSumo’s content analyzer was applied to search for the keywords across various social media platforms, based on shares, links, and comments, providing insights into content engagement, volume, and evergreen scores. Brand24’s analytic media monitoring tool, although limited to some platforms, complemented this by examining online discussions and identifying influential profiles, emerging trends, and impactful content. The combined data from these tools helped uncover public perception, trending topics, and opportunities for digital outreach, offering a comprehensive view of AI’s impact and development in engaging new audiences.



## KEYWORD ANALYSIS

The keyword analysis involved categorizing and observing relevant keyword sets related to AI tools and resources, generative AI services, and their applications in various fields such as health care, education, and business. Data was retrieved from the Google Keyword Planner, a search engine tool. The analysis focused on identifying the most prevalent search terms, such as “AI image generator,” and understanding user intentions behind these searches. It also included examining the search volume for different categories, such as AI accountability, ethics, regulations, and the impact of AI on society. The data was then used to map user pathways to discovery, highlighting how people search for AI-related information and the implications for journalistic coverage.



## NEWS VOLUME

The Media Cloud tool allows for searching words and combinations of words, collecting data points such as attention over time and total attention, providing a snapshot of media attention and public interest in a topic. The analysis focused on AI accountability keywords, such as “AI + bias”, “AI + discrimination”, “AI + black box”, “AI + labelling”, “AI + facial recognition”, “AI + biometric”, “AI + harm”, “AI + disproportionate”, “AI + algorithm harm”, and “AI + colonialism”. The study used the Wayback Machine News Archive for the period from August 8, 2022, to July 25, 2024, and focused on English-language sources. Experimental data points like top words, top languages, and top sources were used as references to cross-check findings, but the main data points were attention over time, total attention, and sample content. The keywords “election” and “Gaza” were used as a baseline for comparison since these are topics pervasive in news worldwide. “AI technology” was also added to the analysis to provide an idea of the news volume surrounding the topic overall.

## Limitations

Given the myriad of ways in which AI is embedded into modern society's life and the broad range of issues and people impacted by it, this study is not an all-encompassing review of journalistic coverage of the technology and its consequences. Limited sample sizes and geographical representation, language barriers, self-selection bias, temporal constraints, and the still-developing tools and methodologies of the digital assessment are limitations that need to be acknowledged as potentially skewing the findings and restricting generalizations to a wider context. Moreover, an extensive literature review and secondary data gathering from case studies were conducted to inform the research design and to provide a substantial background to guide the results' discussions. Due to time and scope restrictions, this stage of desk research has not been fully integrated into this analysis, but the articles and studies reviewed are available in the appendices, under Literature Review.

Future research is recommended to address the aforementioned gaps and build on the findings of this study to provide a more comprehensive understanding of AI reporting and its societal impacts. Nevertheless, the mixed research methods integrated to the analysis and the cross-examination of the data obtained with the available literature allows this study to provide a well-grounded pathway to enhance the discussion about AI accountability and how to tell better stories and engage audiences.





# Executive summary

The study findings reveal that growing discussions around AI have created a distorted or limited perception of its societal role, scope, and risks. The excitement and optimism surrounding the topic reveal a tendency to lean toward positive evaluations, often overlooking the potential harms of this technology. It also represents a risk of a superficial or even unrealistic understanding of what these tools can or cannot do. This phenomenon can obscure criticisms and hinder the conscious access of the general population to these innovations. Journalists face a challenge, often unconsciously, when trying to write about the critical points of AI, especially as they are frequently "swimming against the tide" of an industry that does not always reveal all the necessary aspects for a more accurate analysis. The lack of industry transparency, economic pressure, and corporate influence, combined with technical complexity and limited access to independent experts (due to knowledge or financial constraints) make the balanced and ethical discussion of AI an arduous and complex journey.

The simplified and sometimes sensationalist narrative about AI, usually based on what the media provides, diverts journalism's focus from immediate and tangible impacts, such as algorithmic bias and labor implications, favoring a narrow view of the technology. Journalists and experts agree that the complexity and lack of training are barriers to more accurate AI coverage. The limited top-of-mind associations with journalists or outlets specialized in AI coverage reflects a gap in tech journalism, but the work of local and community reporting, which highlights social impacts and brings marginalized community perspectives to light, is crucial to promoting a more balanced and conscious public discourse.

AI intersects with many dimensions in society, and some of the most pressing issues identified include labor and human rights, bias and discrimination, environmental impact, data governance and sovereignty, privacy and security, misinformation, and complexity and fallibility of AI systems. Creating effective solutions for AI governance requires the inclusive participation of all stakeholders, as behind-closed-doors discussions limit transparency and the emergence of innovative ideas. Initiatives like central agencies, regulatory sandboxes, and inclusive governance are essential to balancing power and incorporating diverse perspectives in the decision-making process.

Journalistic AI training programs, like the one from the Pulitzer Center, are viewed positively for providing training and support for journalists and amplifying voices. There is concern about the concentration of expertise within tech companies and their financial influence on AI research, which generates apprehension among experts as to the size of the challenge in demystifying AI. Nevertheless, enabling journalists to make the connection of AI with its ethical and social implications in their stories is seen as a necessary step toward a more balanced and realistic view of the technology within society.

The work of demystifying AI to empower future professionals and conscious citizens can begin from early school years. Integrating AI accountability journalism and resources into education allows for both exposure to critical emerging technologies and a promotion of ethical and holistic understandings of their use. Engaging stakeholders through information and initiatives where they can explore diverse views and share their own expertise might be a way to ensure a more straightforward path to responsible AI.



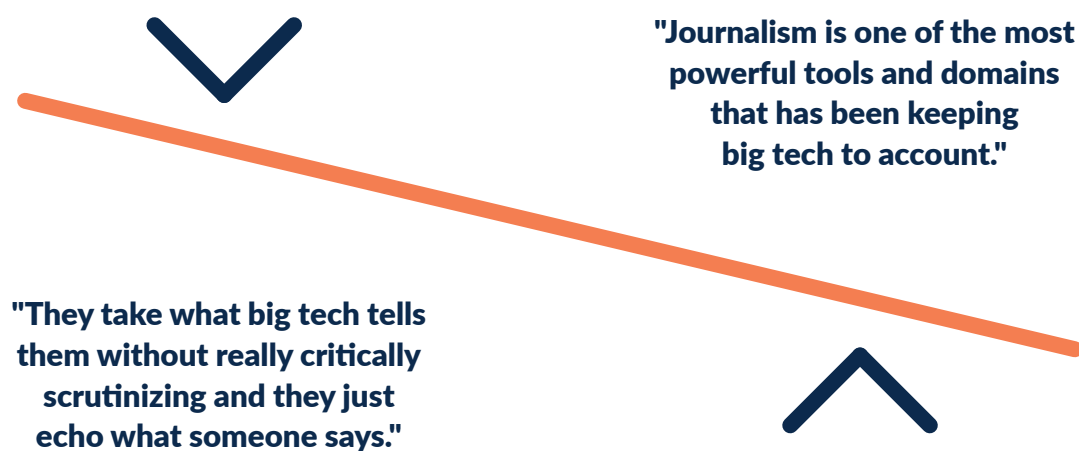


Image by Madeline Cronje, South Africa, 2022

# The landscape of AI reporting

According to the interviewed experts, while the role of the media in shaping public discourse about AI cannot be overstated, another aspect of journalism has been leaving room for questions when it comes to commitment to the public interest: Are all the aspects of AI's impact in society being scrutinized to inform audiences and ensure accountability?

**Graphic 8** Verbatim samples of the qualitative interviews about AI media coverage.



The rapid advance of AI tools in various economic sectors and the appeal of celebrating technological breakthroughs, aligned with limited technical knowledge and the rush to be at the cusp of innovations, have made AI hype a constant factor in mainstream communication. In tandem, speculations about the potentially tragic fate of humanity under a machine-led civilization frequently cloud much of the reasonable scrutiny needed when evaluating AI's impact on society, particularly with regards to marginalized communities. According to the interviewed experts, critical pieces that provide a well-rounded view of the issues are the exception in most of the media today, a phenomenon that is also identified in the news volume assessment.

---

*"Journalism generally covers products with an extremely optimistic or an alarmist perspective, but it is a kind of alarmism focused on the exercise of futurology."*

---

During the qualitative interviews, a few experts pointed to some misconceptions observed in journalists that they believe strongly impact their reporting:

- **AI as a monolithic entity**

The idea of AI as a single, unified technology, rather than a diverse set of tools and applications with varying capabilities and limitations, which leads to oversimplified reporting.

- **Overemphasis on ChatGPT and LLMs**

There is a tendency to focus on large language models (LLM) like ChatGPT, assuming they represent the entirety of AI. This overlooks other significant AI technologies and their impacts and overstates the effectiveness and accuracy of those tools.

- **Existential risk focus**

Prioritizing existential risks posed by AI, influenced by prominent voices in the tech industry, which overshadows more immediate and tangible issues such as algorithmic bias, privacy concerns, and labor impacts.

- **Scientific purity**

There is a misconception that AI is a purely scientific discipline, uncorrupted by politics and ideology when, in reality, AI development is heavily influenced by corporate interests, funding, and political agendas. It is also significantly impacted by human biases.

- **Underestimating ethical and social implications**

Not fully grasping the ethical and social implications of AI technologies leads to reporting that lacks depth in these critical areas.

The Media Cloud analysis helps to illustrate part of that phenomenon. While keyword searches for “AI + technology” indicate an overall high attention rate in terms of news volume, with a story count only 17% smaller than the baseline “Gaza” for the same period, the accountability-related search with the highest number of stories observed (“AI + bias”) is found with a frequency 95% lower than the control expression. Other combinations associating the technology with accountability are found in a volume even more reduced. Although the tool cannot provide a deeper analysis of the frequency or quality of the stories reaching the public, such results corroborate the experts’ perception of an imbalance in the type of stories being released.

The digital assessment also provides insight on where the public interest lies. The keyword analysis shows that the wish to learn more about AI tools and resources and how they can be applied to one’s daily life is the most prevalent, followed by the need to understand what the technology is, from basic definitions to specific types of AI. More importantly, when a value is added to the search, “benefits of AI” largely outweighs the search for potential negative impacts or harm. The eagerness to understand AI and how it can affect their lives exists among the public. Yet, when this search is met by superficial or partial results, there can be a tendency to detach AI, for example, from its complicated and often cumbersome supply chain.

Another evidence of this trend is the overall positive sentiment surrounding the discussions about AI in social media, which indicates high public engagement and interest in its advancements without significant ethical concerns. Traditional media outlets, alongside YouTube, are the platforms that generate higher engagement rates in the online dialogue about AI, which is a substantial opportunity to generate educational content to increase the critical assessment of the topic and demystify the hype. There are indications that such reflections are beginning to gain traction, especially for topics such as AI regulation and AI bias for specific audiences, but the prevailing impression among specialists is that the industry’s marketing, lobbying power, and unlimited resources have been quicker and more effective in dominating the narrative, and that there is hardly a comparable force to counterbalance it.



# How did we get here?

The interviewed experts identify a number of factors that contribute to the unbalanced tone of AI reporting. These are mostly related to the nature of the technology itself, the industry influence, and issues coming from the newsrooms.

## Industry PR and power dynamics

The undeniable power of the industry and the vast resources applied to market their products and services have a strong influence in the content that reaches the media. Journalists often rely on press releases and industry events to learn about the latest developments, leading to a repetition of corporate claims. Governments adopting AI technologies also often shape their narratives in a positive and fairly superficial position to explain public spending. Additionally, there is a predominant perception that most media outlets and many of the tech specialists sourced for information have sponsoring ties with big tech companies.

---

*"The overwhelming amount of industry PR that shapes the information ecosystem."*

---

---

*"If I'm a social scientist who is a legal scholar, who studies AI, and I would like to say something, I probably don't have any journalists in the room. It is going to be my peers or students who are going to know about what I'd like to say, because the journalist is already in the company's conference."*

---

## Assumed complexity and fast-evolving nature of AI technology

The technical and ethical intricacies of AI represent a challenge in the sense that journalists are often unaware of what questions to ask. Intelligence is a concept that is still not completely explained by science and attributing it to the technology seems to have given it an aura of mystery that should be privy only to its creators. The AI supply chain is vastly unknown by non-specialists, a shortfall that leads to most of society overlooking the consequences of its dynamics. Many of the AI-related fields, such as machine learning, neural networks, computer vision, natural language processing, robotics, etc., frequently sound as too much of a challenge for a reader to delve into with confidence. Finally, the rapid development of the technologies requires a lot of resources from newsrooms to keep up with the latest advancements and implications, resulting in outdated or oversimplified narratives.

---

*"I'm not saying that there is some dishonesty in the reporting, right? I believe that the journalists make the best effort to understand the issue and report it as he or she sees it, but because the technology is moving so fast, a lot of assumptions or understandings that they use in the analysis or in the reporting may be inaccurate, unfortunately."*

---

## **Lack of training or resources and problematic media practices**

There is a consensus among specialists that journalists need better training and that having access to specialized knowledge would aid them in covering AI accurately and critically. However, there is also a widespread understanding that there are limitations of time and funds in most newsrooms to enable that. Still, a few experts note that the investigative nature of journalism does not fail other relevant topics to society, such as politics or economics. And since AI has become part of most areas of life, they do expect reporters to be better prepared to ask the relevant questions.

---

*“In general, I don't have the impression that when I see a curriculum for a journalism program, I see a lot of tech in there as a subject of study in all of its complexity.”*

---

Another challenge in asking the necessary questions is related to industry transparency and access to information. Independent verification of the claims made by AI developers or vendors is nearly impossible, an issue that is still very much tied to regulatory frameworks that are yet to be established or improved.

---

*“We have this kind of awful concept that is the black box of technology. (...) We don't know how a lot of the platforms that we use work. (...) And so, the tech industry will say: ‘Well, this is proprietary information. If I tell you how my code works, I'm gonna lose a lot of money’.”*

---

The “cool” factor is also strongly associated with any tech coverage by the experts. The excitement about innovative aspects of AI is considered a sellable perspective, and one that feeds into the techno-solutionism that is prevalent in society today. The idea that technology may solve all problems and make humans perform better is appealing to most of the audience, which leads to a one-sided presentation of the facts.

---

*“Particularly, investigative journalism is doing very well in that, but most of them also have journalists that cover the flashy last iPhone. Because those articles have huge responses online, that's what people want to read.”*

---

## Lack of training or resources and problematic media practices

Most experts agree that the coverage coming from the ground is only presented when that ground belongs to the Silicon Valley or any other major business hubs. A geographical and socioeconomic bias when it comes to looking for AI-related leads results in neglected perspectives and impacts on the global majority and marginalized communities. Alternative media and smaller outlets are mentioned as often providing more nuanced coverage, but one that lacks the same reach of mainstream media.

---

*"I think a lot of AI coverage really focuses just on the Bay Area and just on the companies that exist there or, you know, the broader Silicon Valley like in London and places like that. But there's not really enough people that are on the ground in any communities, (...) particularly communities in the Global South that are also very much part of the supply chain of AI development, and impacted, by the products of AI development."*

---

Due to the technical aspect of AI issues, sourcing specialists to help explain issues can also be a challenge for reporters. A few experts point out that often, when a specialist that has no ties with the industry or political interests is interviewed, they present a critical view that confronts the overall optimism and raises key questions. Yet, these articles tend to create a dichotomy that is a sellable but not necessarily helpful stance: a "yes or no to AI" issue. Presenting polarized views without providing a balanced analysis of trade-offs and monitoring mechanisms ends up alienating audiences and reducing the engagement with real, tangible discussions.

---

*"There are examples, unfortunately, of situations where a journalist or a specific outlet has that kind of inherent skepticism I spoke about, and they just don't have the capacity to get to the bottom of it. And so, they will call up someone like me to be like: 'Hey, could you just kind of tell me why this is bad? I know that I'm supposed to be concerned about this, but I don't really know why. Can you just quickly tell me why this thing is bad?' And obviously it's actually not in depth. It's a kind of different iteration of the same problem."*

---

Finally, in a context where newsrooms are under pressure to enhance engagement, there is a certain understanding from the experts to the need for producing attention-grabbing content. Nevertheless, reporting fueled either by theoretical alarmism or industry self-interest and focus on high-profile figures making bold declarations often overshadow more nuanced and critical discussions, not necessarily because they receive attention, but because such attention is not capitalized to then bring important issues to the surface.

---

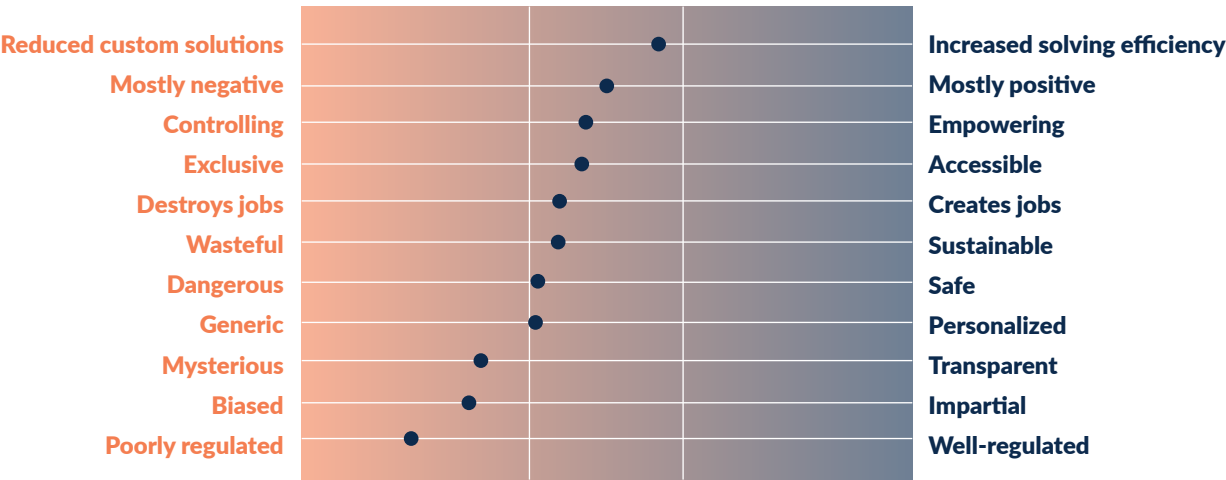
*"In terms of the framing of AI, I think that the media can be sensational, because that's how you get people's attention. But I think that they should really have more of a focus on existing harms, more than existential harms. Because, essentially, I feel it would make them go down to the field, instead of focusing on those talking heads. (...) Because these guys, when they say something, it's like throwing a big rock into the lake and it creates a lot of ripples. It's easy to report."*

---

# The journalists' input

When evaluating a series of attributes about AI technologies and considering whether their effects would be positive or negative, journalists present mostly neutral opinions. The positive trisection remains empty, with the idea that AI increases efficiency in problem solving being the only one closer to a favorable placement. Three issues are positioned in the section of negative evaluations: the lack of regulation of AI technologies, the lack of impartiality of the technology (bias), and the lack of understanding about how it works.

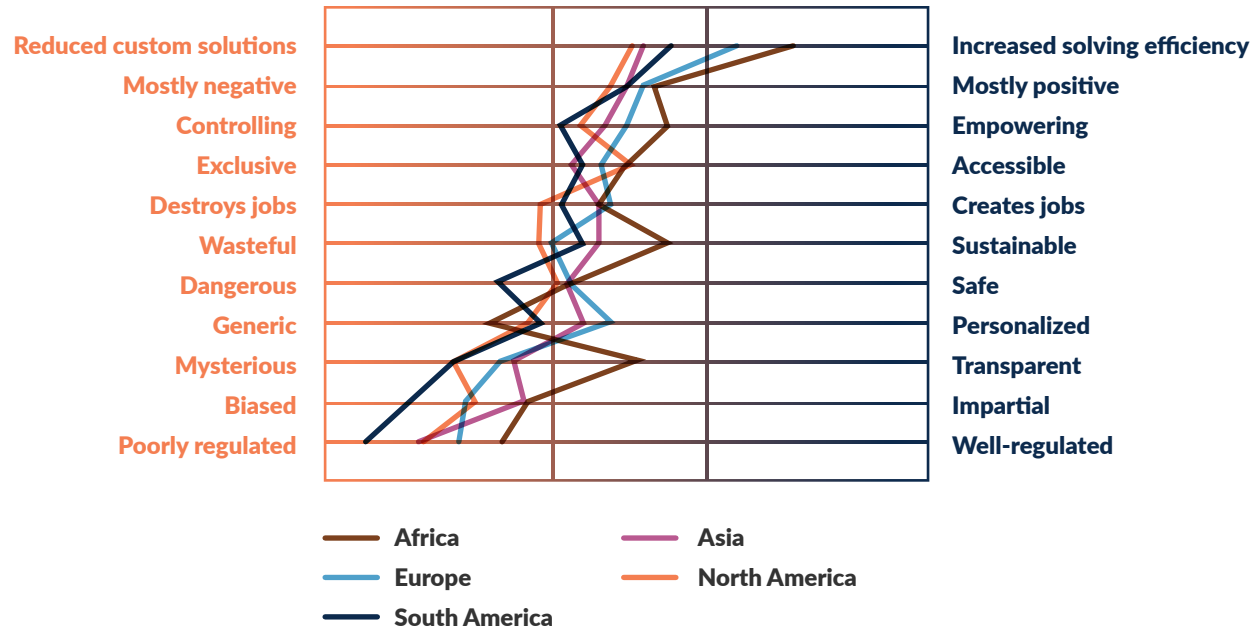
**Graphic 9** Journalists' perceptions about AI, mean (N=90).



Africa is the continent that attributes the most positive evaluations to AI. Its journalists perceive the highest degree of transparency in AI technologies, are not as focused on questions of regulation, perceive these technologies as more sustainable, more generic (not personalized), more empowering, and give a positive evaluation of efficiency in problem solving, an item that also receives a positive evaluation in Europe. Journalists from South America and North America have the most critical opinions about AI, being the audience that presents the most negative evaluation about the lack of transparency. In South America, journalists tend to be more critical regarding the lack of regulation, the lack of impartiality, and the possible harms involved in the use of these technologies.

# Graphic 10

Journalists' perceptions about AI per continent, mean  
(Africa N=11, Asia N=28, Europe N=18, North America N=26, South America N=7).

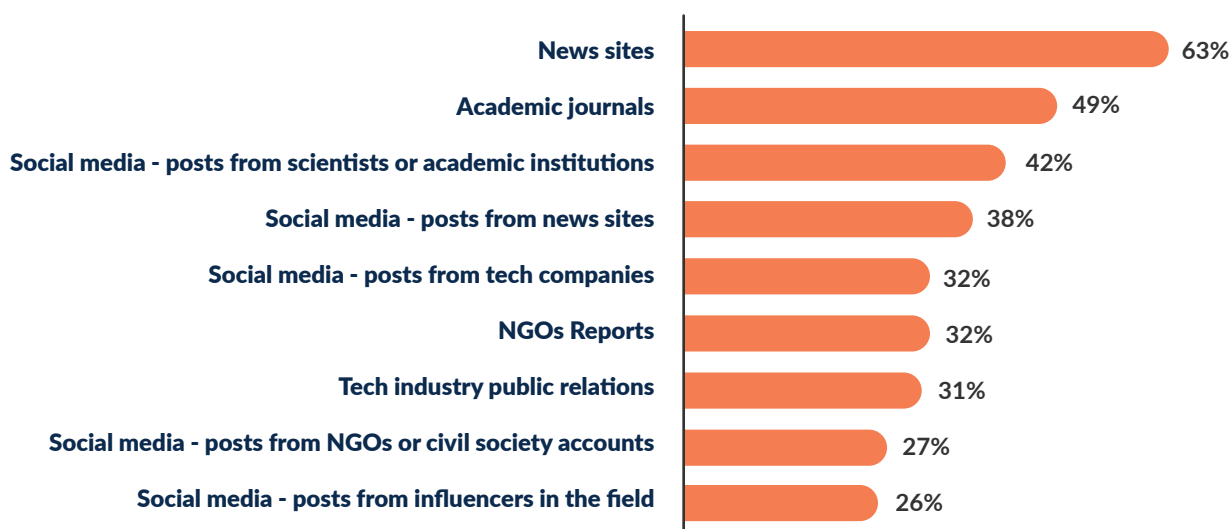


Understanding journalists' perceptions about the technology itself is valuable to identify the roots of some of the problems in the current coverage of AI issues. There is an overall neutrality with regards to the technology having a positive or negative impact, providing power or controlling users' lives, being accessible or not, impacting jobs, being environmentally sustainable, safe or dangerous, or being able to tailor experiences to one's preferences versus providing generic solutions. Such results could stem from the search for impartiality that is intrinsic to the profession. Yet, it can also represent a lack of sufficient information to present a more consolidated position.

As opposed to the specialists, who distrust mainstream media coverage of AI, journalists highlight news websites as their main source of information about the topic. More than half of the respondents mention the websites themselves (63%), and another 38% cite posts from these sites on social media. The second source of information in this area comes from the academic field, with nearly half of the journalists (49%) obtaining information from academic articles and 42% from posts by scientists and academic institutions on social media. Academic journals are more prominent in South America (57%). Information from tech companies and NGOs are consumed by about a third of journalists (32%), with the first resonating more in Europe (50%) and Africa (46%), while NGOs are more relevant in South America (43%). Social media is an important communication channel, consolidating AI information from multiple sources in one place.

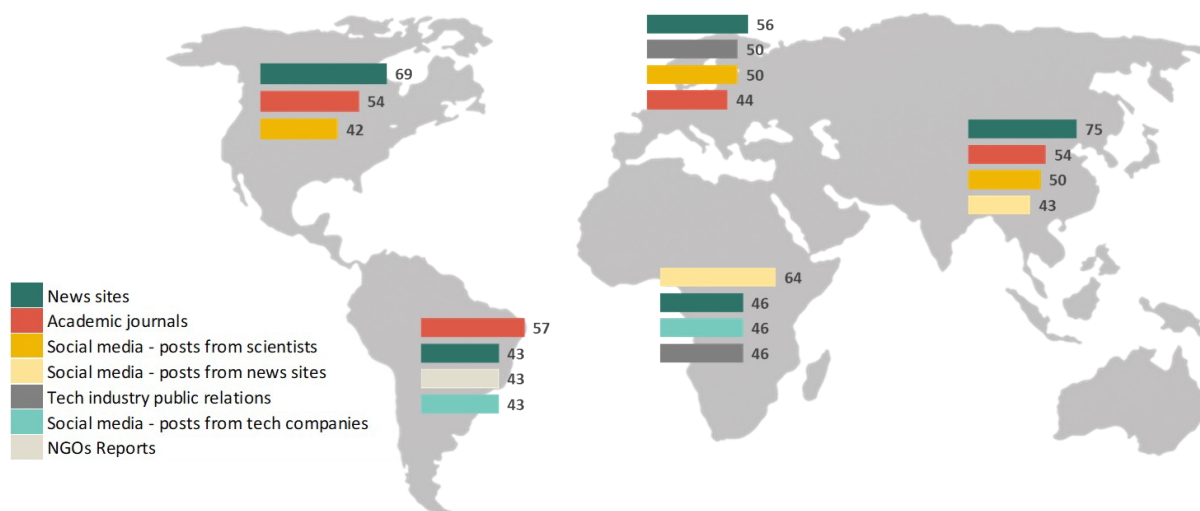


**Graphic 11** Main sources of information on AI issues, % of cases (N=90).



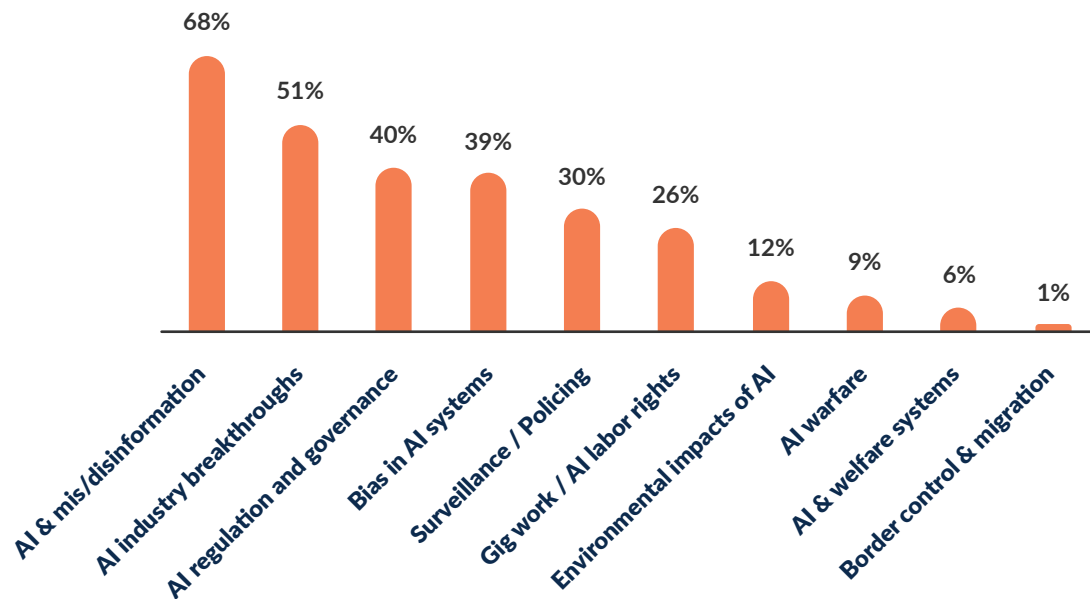
Other answers mentioned spontaneously: Conversation with AI sources (6%), newsletters (6%), conferences/events (3%), podcasts (3%), AI journalism (1%), expert online forums (1%), expert reports (1%), recent books (1%), scientists (1%), social media - posts from social media platforms (1%), source-based information (1%).

**Graphic 12** Main sources of information on AI issues per continent, % of cases, options with a frequency of 40% or more (Africa N=11, Asia N=28, Europe N=18, North America N=26, South America N=7).



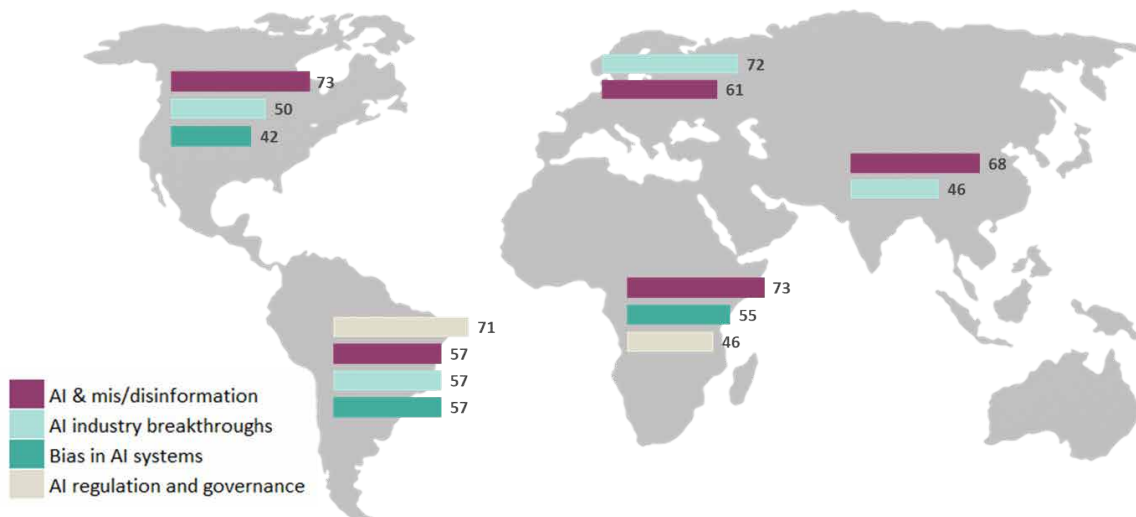
The journalists' observation of the issues more frequently seen in the media indicates their attentiveness to some of the problems resulting from AI use. Nearly seven out of 10 consider AI disinformation as the most covered topic. In second place, half of the journalists cite AI advancements (51%). Issues related to regulation and bias in AI systems are highlighted by four in 10 journalists (40% and 39%, respectively), followed by topics of surveillance (30%) and labor rights (26%). Professionals in South America highlight AI regulation and governance as the most present issue in the media (71%). Problems related to bias in AI systems are more reported in South America and Africa (57% and 55%, respectively).

**Graphic 13** Most-reported AI topics in the media, % of cases (N=90).



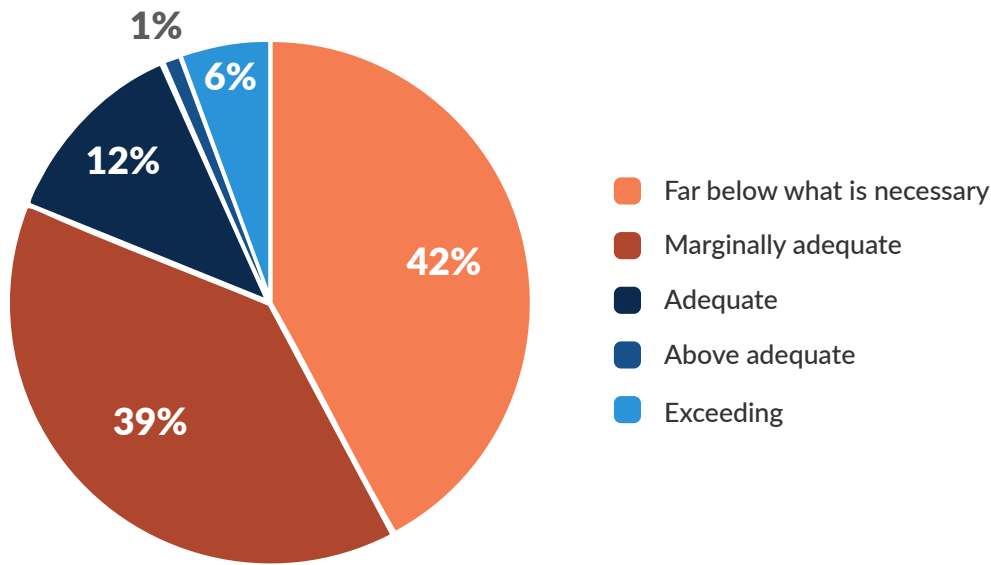
Other answers mentioned spontaneously: Consciousness and philosophical aspects (1%), general tools (1%), technology developers (1%), the companies creating AI products 1%.

**Graphic 14** Most-reported AI topics in the media per continent, according to journalists, % of cases, options with a frequency of 40% or more (Africa N=11, Asia N=28, Europe N=18, North America N=26, South America N=7).

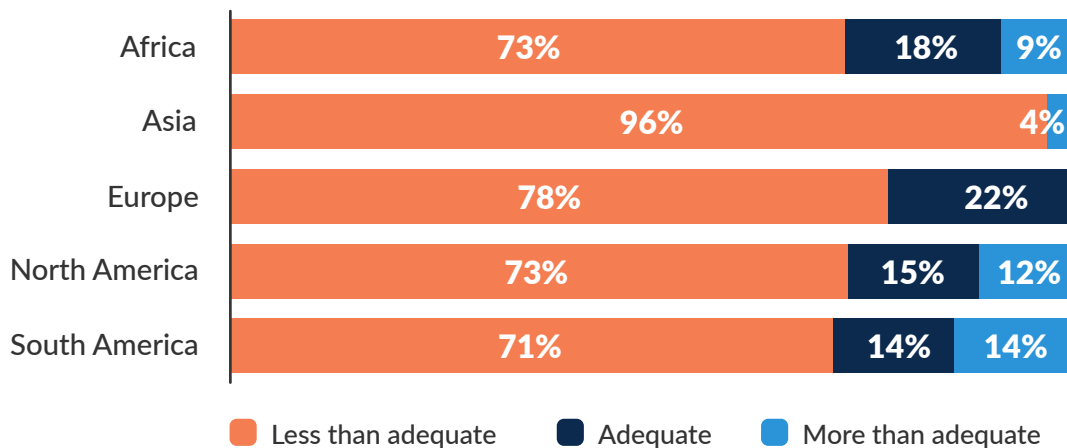


The vast majority of journalists around the world believe that the current frequency of AI technologies coverage is below adequate (81%), with this position being strongest in Asia (96%). Europe is the continent where the largest proportion of journalists consider AI coverage to be adequate (22% in Europe, compared to 12% worldwide). In South America and North America, a higher proportion of journalists consider coverage to be above adequate (14% and 12% respectively, compared to 7% worldwide).

**Graphic 15** Current frequency of AI technologies coverage, % of cases (N=90).






**Graphic 16** Current frequency of AI technologies coverage per continent, % of cases (Africa N=11, Asia N=28, Europe N=18, North America N=26, South America N=7).



Concurrently, when asked to name outlets that cover AI stories with consistency in their region, 24% of journalists were unable to name a media organization, and 13% declare there is none meeting this criteria. In the social listening analysis from the digital assessment, the need to extend the shelf life of accountability-related stories to keep relevant discussions going is highlighted. The significant number of journalists indicating a lack of consistent representation of AI stories overall in their regions could suggest an opportunity to occupy these spaces with stories that could make a difference in audience engagement in the long run.

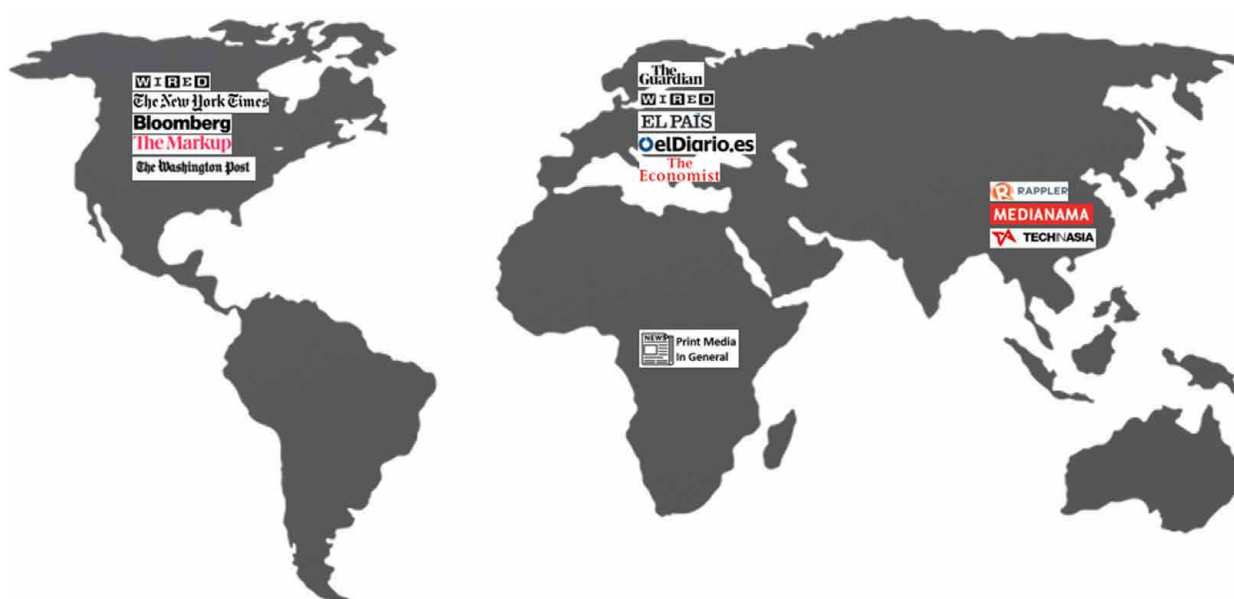
Among those who recalled an outlet, *Wired* magazine stands out as the main reference, spontaneously cited by 12% of journalists. Its presence is stronger among the audience in North America and Europe. *The Guardian* and *The New York Times* follow, mentioned by 7%, with *The Guardian* being the main reference in Europe and *The New York Times* the second-most-cited outlet in North America.

**Graphic 17** Media outlets covering AI stories per region or country on a fairly consistent basis, % of cases that mentioned each outlet (N=90).

01 <b>WIRED</b> 12%	02 <b>The Guardian</b> 7%	03 <b>The New York Times</b> 7%	04  National Media In General 4%	05 <b>Bloomberg</b> 3%
06 <b>EL PAÍS</b> 3%	07  RAPPLER 3%	08 <b>The Markup</b> 3%	09 <b>The Washington Post</b> 3%	10  Print Media In General 3%

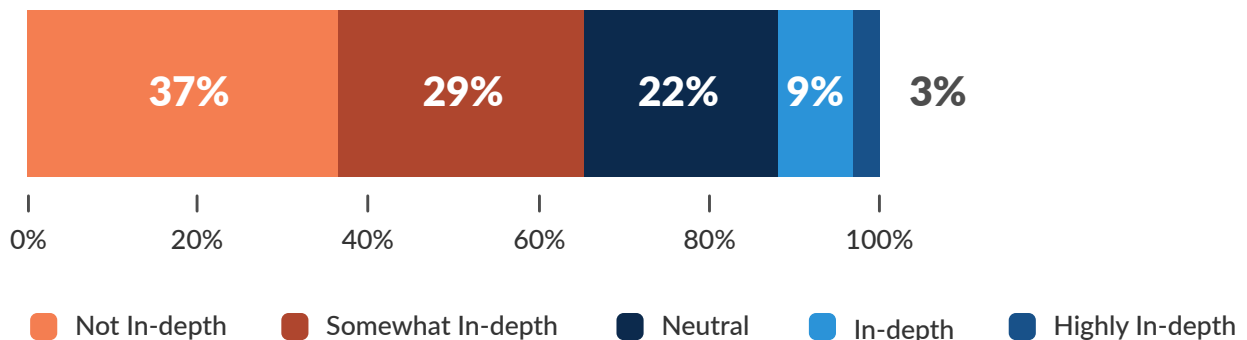
There is a regionalization concerning outlets around the globe. In North America, the standout outlets cited are WIRED, The New York Times, Bloomberg, The Markup, and The Washington Post. In Europe, The Guardian, WIRED, El País, elDiario, and The Economist are mentioned. In Asia, Rappler, MediaNama, and Tech in Asia. Respondents did not cite any major outlets covering these topics in Africa and South America.

**Graphic 18** Main media outlets covering AI stories per region or country on a fairly consistent basis per continent (Africa N=11, Asia N=28, Europe N=18, North America N=26, South America N=7).



Regarding the depth of AI reporting, the proportion of neutral evaluations stands out. While two-thirds of the consulted journalists say there is insufficient depth in the coverage (37% say current pieces are usually not in-depth and 29% say they are somewhat in-depth), at least one out of five (22%) declares neutrality as to the complexity of the reporting, indicating again a possible lack of awareness of the issues to pick a more definite stance

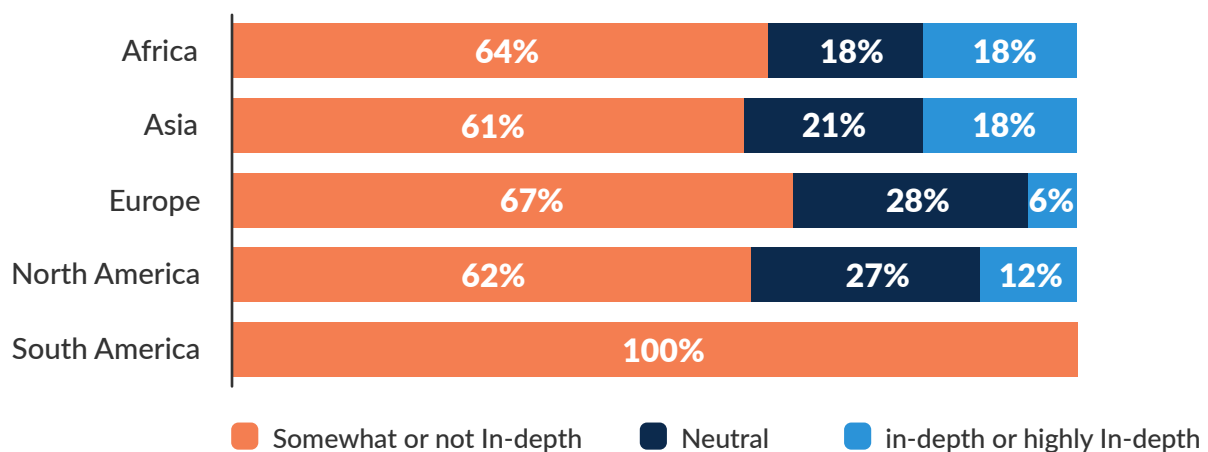
**Graphic 19** Depth of current reporting on AI\*, % of cases (N=90).



\*Note: The question regarding the depth of current reporting was not associated with any specific outlet, but as a general view.

All respondents from South America agreed that local coverage is somewhat or not in-depth. Africa and Asia present a higher share of journalists rating their coverage as in-depth or highly in-depth (18% each continent).

**Graphic 20** Depth of current reporting on AI per continent, % of cases (Africa N=11, Asia N=28, Europe N=18, North America N=26, South America N=7).





A series of factors are perceived by journalists as barriers to pursuing in-depth reporting on AI technologies and their impact on society, ranging from more basic issues like the lack of understanding of AI concepts (mentioned by 51% of journalists) and the complexity of the topic (40%), to issues revealing a lack of technical structure, such as the lack of training opportunities on AI accountability reporting (53%), lack of newsroom awareness/support (39%), difficulty accessing data or sources (34%), and not having enough local examples of AI reporting as inspiration (30%). Issues related to the lack of understanding and the complexity of AI are more common barriers in South America and North America. The lack of knowledge and technical support is more evident in Africa, Asia, and Europe.

**Graphic 21**

Main barriers journalists face in your region to pursue in-depth reporting on AI technologies and their impact on society, % of cases (N=90).

**Lack of training opportunities on AI accountability reporting**



**Lack of understanding of basic AI concepts**



**Perceived complexity of the topic**



**Lack of newsroom awareness/support**



**Difficulty accessing data or sources**



**Not enough local examples of AI reporting to use as inspiration**



**Lack of audience awareness/interest in the topic**

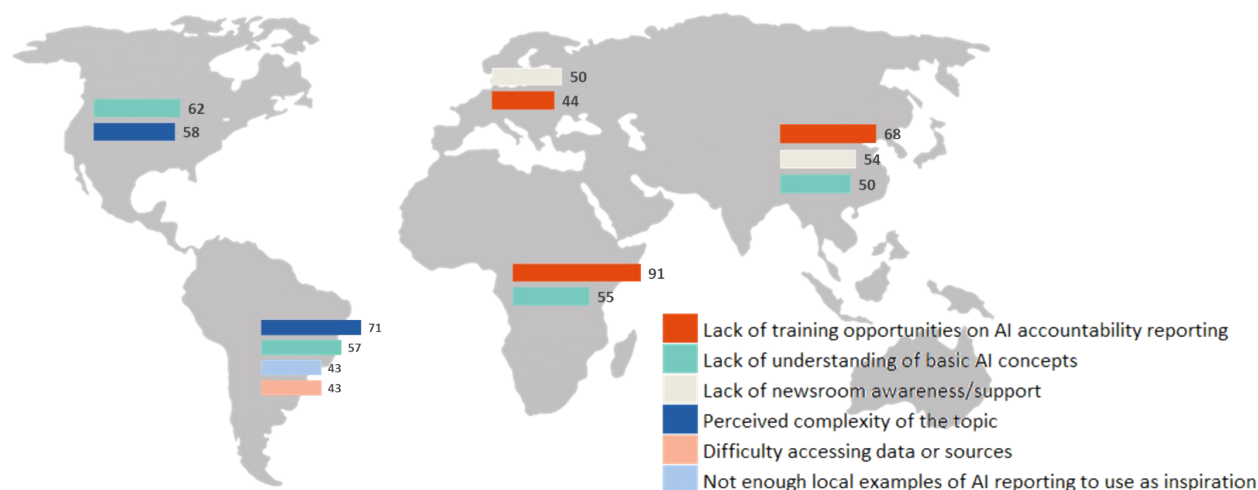


**Too much AI hype**



Other answers mentioned spontaneously: Funding for reporting (1%), fusion of technologies (1%), industry-funded misinformation (1%), lack of critical journalism (1%), lack of transparency (1%).

**Graphic 22** Main barriers journalists face in their regions to pursue in-depth reporting on AI technologies and their impact on society per continent, % of cases, options with a frequency of 40% or more (Africa N=11, Asia N=28, Europe N=18, North America N=26, South America N=7).



In terms of training, a few of the qualitative interview participants highlight that many journalism programs do not include comprehensive technology courses, leaving students unprepared to cover complex technological topics like AI. Additionally, much of the available and updated material on AI and technology is in English, and language barriers can often hamper the discussions within many classrooms worldwide. As a result, educators fall back on traditional ethics debates without necessarily addressing specific ethical challenges posed by new technologies. Furthermore, the concentration of AI expertise within companies can lead to narrow perspectives among new AI researchers, who are often educated by professors with ties to these companies. These gaps in training contribute to journalists' struggles with understanding and critically covering technological advancements and their societal impacts.

# The good examples

Many of the interviewed experts found it difficult to name outlets or journalists that could be associated with good AI coverage at the top of their minds. A similar phenomenon is observed among the journalists consulted in the survey: 18% could not name an outlet and 61% could not name a leading journalist. Regardless, when given the time to search for pieces that caught their attention, there is work praised by in-depth reporting on AI and its societal impacts, particularly coming from local and community-based outlets. Experts believe that having the perspectives of marginalized communities challenges the existing power structure, helping to bring some balance to the public discourse. Investigative work that brings the real-world impacts of AI deployment is also key to raise critical awareness of the issues, enabling a more grounded understanding that can be extremely valuable to the general public.

**Graphic 23** Examples of good AI coverage, according to experts.<sup>1</sup>

<b>MIT Technology Review</b> Highlighted for its in-depth analysis and coverage of AI-related topics, particularly the Data Bricks group, which provides insightful publications on AI's impact on businesses and the potential for companies to develop their own AI systems.	<b>Jacobin</b> Praised for its detailed articles and ability to foster debates among readers. Critical issues such as information integrity and the challenges of fake news were recalled.	<b>Rest of World</b> Commended for its deeply reported, on-the-ground stories about AI's impact outside of Silicon Valley, including data workers in China and data center resistance in Latin America.
<b>La Barra Espaciadora</b> Works on popularizing concerns about technology and human rights.	<b>Ojo Público</b> Conduct investigation work using AI, such as identifying government contracts for geolocation monitoring and testing data anonymization claims.	<b>Agência Mural, Amazônia Real, Marco Zero Comunicação, and Datalab</b> Local and community-based outlets providing critical reporting on AI from the perspectives of marginalized communities.
<b>The Intercept Brasil</b> Noted for its investigative journalism and critical approach to AI and technology issues.	<b>ProPublica</b> Mentioned for its excellent investigative work over the past few years.	<b>The Markup</b> Praised by its significant investigative journalism on AI.
<b>O Joio e o Trigo</b> Addresses technology-related issues, including AI, while primarily focusing on food and politics.	<b>Agência Pública, Repórter Brasil, O Núcleo, and Politics.org.br</b> Local outlets highlighted for reporting on various social issues impacted by AI and for their comprehensive and critical coverage.	<b>AzMina and Gênero e Número</b> Noted for offering perspectives on AI from a gender lens, making their work particularly valuable.
<b>404 Media</b> Recently emerged as a noteworthy outlet for its critical coverage of AI.	<b>The New York Times</b> A facial recognition experiment of The Privacy Project was mentioned as worthy illustrative way to discuss privacy and how lives have been affected by information shared with and without consent.	<b>The Guardian</b> Noted for its interesting and diverse coverage of AI.
<b>Time Magazine</b> First to run a story on big tech work in Kenya and consistently covers AI issues with detailed and contextual reporting, including communities affected by AI outside of Silicon Valley.	<b>Financial Times</b> Consistent in covering AI-related stories, particularly those highlighting the human impact of AI technologies.	
<b>Wired</b> Mentioned for incentivizing readers to debate the issues covered.	<b>The Wall Street Journal</b> Noted for exemplary pieces providing detailed and contextual reporting on AI issues.	

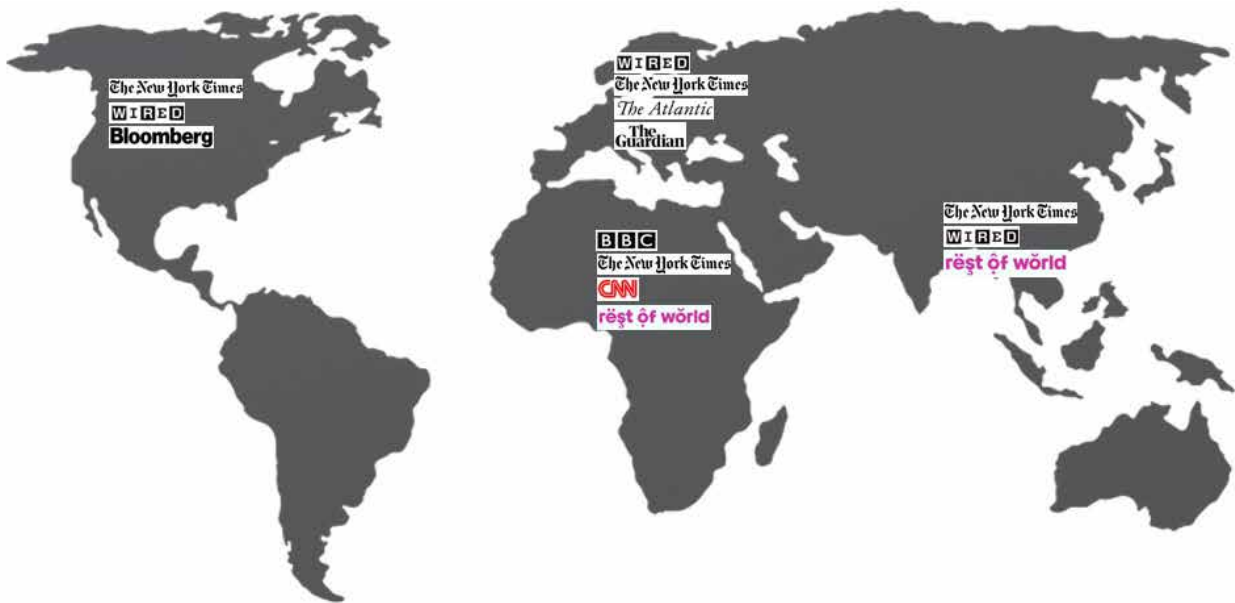
These outlets were mentioned by the experts as examples of continuous good coverage or specific projects that have gained their attention. For the journalists in the survey, a ranking of prominent global outlets presenting a strong coverage of AI stories can be calculated. *The New York Times* is mentioned by 27%. With a strong global presence, it is highlighted in North America, Europe, Asia, and Africa. *Wired* magazine occupies second place (18%) and receives the spotlight in Europe, North America, and Asia. Rest of World and *The Guardian* follow with 11% of mentions each. The former is more prevalent among Asian respondents, while the latter is stronger in Europe. The BBC (9%), *MIT Technology Review* (8%), and *The Washington Post* (7%) also reach notable positions in the global ranking. BBC and CNN are recalled more often in Africa, which is also the case for *The Atlantic* in Europe and Bloomberg in North America.

**Graphic 24** Most prominent global outlets you see doing strong AI reporting, % of cases that mentioned each outlet (N=90).

01 <b>The New York Times</b> 27%	02 <b>WIRED</b> 18%	03 <b>röst of world</b> 11%	04 <b>The Guardian</b> 11%	05 <b>BBC</b> 9%
06 <b>MIT Technology Review</b> 8%	07 <b>The Washington Post</b> 7%	08 <b>A</b> <i>The Atlantic</i> 6%	09 <b>The Markup</b> 6%	10 <small>THE WALL STREET JOURNAL.</small> <b>WSJ</b> 6%
11 <b>Bloomberg</b> 4%	12 <b>FT</b> FINANCIAL TIMES 4%	13 <b>CNN</b> 3%	14 <b>L</b> LIGHTHOUSE REPORTS 3%	15 <b>REUTERS®</b> 3%
16 <b>TC</b> TechCrunch 3%				

## Graphic 25

Most prominent global outlets you see doing strong AI reporting per continent (Africa N=11, Asia N=28, Europe N=18, North America N=26, South America N=7).



A total of eight outlets coincide as a top-of-mind of good examples for both experts and journalists. *The New York Times*, *Wired*, *Rest of World*, *The Guardian*, *MIT Technology Review*, *The Markup*, *The Wall Street Journal*, and *Financial Times* seem to have consolidated their position as media to be observed and approached for partnership in engagement initiatives.

The reasons for pointing out these media vehicles as doing work that is worth following include elements that help create a more nuanced and comprehensive understanding of AI's impact on society. According to experts, these can be the coverage of overlooked communities, especially those in the Global South; an inversion of power, focusing on perspectives from the bottom of the food chain rather than just the executives and big-name companies; a critical perspective that juxtaposes different views or questions companies' claims; and a grounded understanding, relating AI to people's day-to-day lives, making the technology more understandable and relevant to the general public.

It is important to note that identifying outlets as references for AI journalism does not exempt these same channels from criticism for occasionally skimming the surface of AI issues. Echoing big tech claims without scrutiny or focusing on product features rather than the underlying politics or economics is the most frequently observed faux pas. However, even coverage of problem resolutions sometimes falls into a superficial look, such as focusing on aspects like court dates and judge statements when reporting on a litigated case, instead of delving into how society is impacted by the object of the lawsuit.

# Journalists worth following

It is a challenge for experts and journalists to name a professional who has been standing out for doing a solid job reporting AI stories. In the survey, 61% of respondents could not recall any leading journalists covering the issues in their region. Among the experts, only a few had a person in mind immediately, indicating a strong association with appreciated work. Such assessments were justified by community-focused stories, linking questionable practices with everyday life consequences, critical assessments of large language models (LLMs) and their societal impacts, contextual reporting on AI workers, and the use of accessible language to explain AI concepts. Those mentioned by name in the qualitative sample and who were cited at least twice in the quantitative study are presented below.

**Graphic 26** Leading journalists covering AI issues (journalists cited at least twice), % of cases that mentioned each journalist (N=90). Experts' mentions are not quantified but are indicated beneath each journalist's photo.



Mentioned by experts

Karen Hao is an award-winning journalist who covers artificial intelligence. She was the first journalist to profile OpenAI. She is currently working on a book about the company and the AI industry for Penguin Press, set to be published in 2025. She is also a contributing writer for *The Atlantic* and leader of the AI Spotlight Series, a program she created with the Pulitzer Center to train journalists on how to cover AI.



Cade Metz is an experienced journalist specializing in technology who currently works as a reporter for *The New York Times* in San Francisco. Having worked for *Wired* as a senior staff writer, he covers the emergence of new ideas, the startup scene, and the major players in the internet industry, such as Google and Facebook.



Karol Ilagan is a Filipino journalist and journalism educator. Before joining the faculty of the University of the Philippines Diliman, she led investigations and collaborations at the Philippine Center for Investigative Journalism. She is a member of the International Consortium of Investigative Journalists and part of the Pulitzer Center's Rainforest Investigations Network and AI Accountability Network.



Jaemark Tordecilla spent the last year at the Nieman Foundation at Harvard University as a fellow, studying AI. He is a winner of the TOYM Award in the Philippines for his work in digital journalism. He has worked for nine years as the head of digital media at GMA News in the Philippines, where he oversaw all online publishing and audience development activities and founded the network's Digital Video Lab.



Kashmir Hill is a reporter for *The New York Times*, where she covers privacy and technology. She is currently working on a book about facial recognition technology to be published by Random House. She writes about the impending technological dystopia and how we can avoid it. She has written for *Gizmodo Media Group*, *Forbes* magazine, and *Above The Law*, and was the editor of *Real Future*, Fusion's technology vertical.





2%

Leon Yin is an award-winning journalist for *Bloomberg News*, where he creates datasets and develops research methods to investigate the impact of technology on society. As a data journalist, he writes for *Inspect Element*, a practical guide to auditing algorithms. His work quantifying Google and Amazon's self-preferencing has been cited numerous times by lawmakers, academia, and popular media. Yin began his career working with news at *The Markup* and conducting research at NASA, writing Fortran scripts.



Mentioned by experts

Billy Perrigo is a technology journalist for *TIME* magazine, covering the technology sector. He is well-known for his investigations into the practices of large technology companies and the social impact of AI. He is the winner of the Orwell Prize 2022 for his report "Inside Facebook's African Sweatshop."



Mentioned by experts

Cory Doctorow is a Canadian sci-fi writer and activist, as well as a contributing writer to *Wired*, *Financial Times*, *The New York Times*, among others. He is known for advocating for freedom of information in the digital age. A fierce critic of restrictive copyright laws and an advocate for the free sharing of knowledge, he is an important voice in the defense of free culture and freedom of expression in the digital age.



Mentioned by experts

Garrison Lovely is an American freelance writer and host of "The Most Interesting People I Know" podcast. Known for pieces like "Can Humanity Survive AI?", he has published stories in *The Nation*, *Jacobin*, *BBC Future*, *TIME*, and numerous other newspapers and magazines.

## Graphic 27

Leading journalists covering AI issues (journalists cited at least twice) by origin of mention (Africa N=11, Asia N=28, Europe N=18, North America N=26, South America N=7).

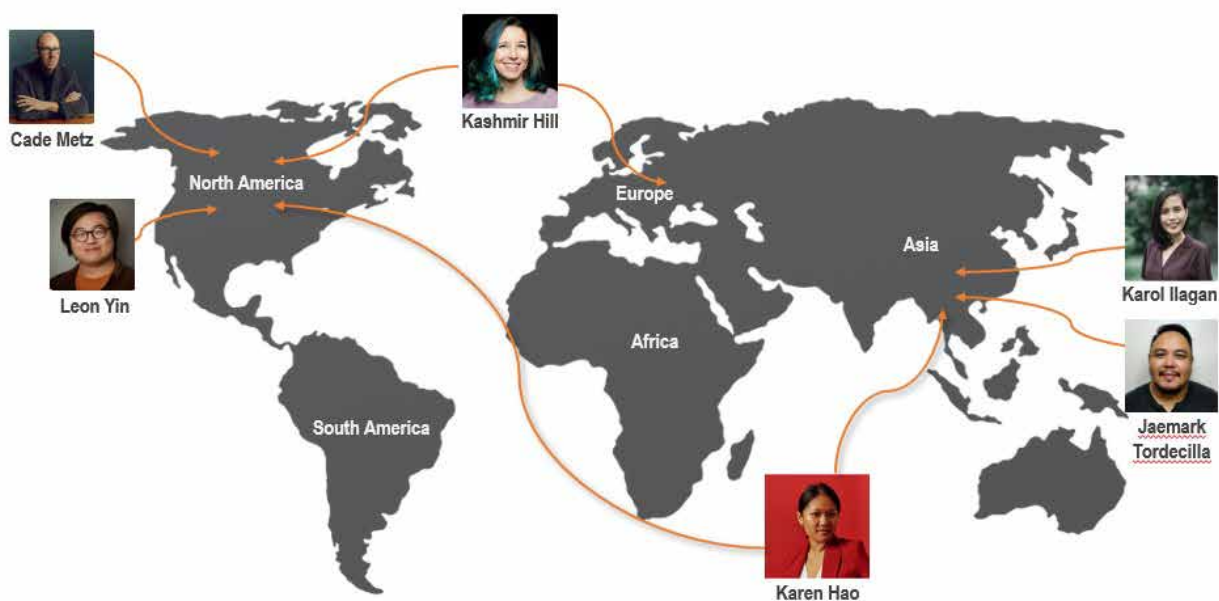






Image by Sarah Pabst, Brazil, 2024

# Pressing issues in AI today

In light of the challenges and shortcomings detected in current AI stories, an overview of the most pressing AI-related issues is imperative to provide guidance for reporters willing to delve into more critical coverage of the topic. AI is part of an industry dominated by a few giant tech corporations that concentrate mammoth power and resources, setting the rules and practices in the supply chain with minimal oversight. As a result, profits are often prioritized over transparency and ethical considerations, and reporting on these issues represents raising awareness, informing the public and policymakers, and catalyzing discussions on accountability.

The experts interviewed identified a number of intersections of the technology with important aspects of society's life, including labor and human rights, bias and discrimination, environmental impact, policy and governance, communication and information, and privacy and security. It is a shared perspective that AI's rapid advancement and pervasive influence across various sectors require deeper analysis, and that the media can play a pivotal role in fostering accountability and ensuring the public is well-informed about ethical, social, and economic implications, moving beyond superficial narratives and sensationalism. This chapter will explore these critical issues and suggest more nuanced approaches to matters that already impact many communities and that are often off the radar of public interest.

---

*"You don't have to use a computer or a smartphone to be affected by AI."*

---

## Labor and human rights

At first glance, the intersection of AI and labor tends to lead to reflections on job displacement or, according to the most sensational headlines, how robots are going to steal your job. Although there are estimates that the technology will soon displace millions of jobs worldwide, the same researchers claim AI will have created new jobs in nearly double the lost amount ([The Future of Jobs Report](#), by the World Economic Forum, 2023). The context in which some of these new jobs come to fruition, however, can be highly problematic.

For the "intelligence" feature of the technology to be possible, data needs to be labeled to train AI models,<sup>2</sup> content has to be moderated to flag problems or ensure compliance, errors require correction, and algorithms and models must be developed and defined for each specific application. Additionally, the physical infrastructure that supports an AI system needs to be managed and maintained, voice samples are necessary for AI-generated speech, texts require sourcing and proofreading, and visual concepts need to be designed. Human labor is necessary to make all of these tasks possible, despite product descriptions often failing to explain how the completely automated decisions of their machines rely on it.

As is the case with other tasks that require a large workforce performing repetitive work as quickly as possible, the AI global market relies heavily on a large network of low-paid and overworked people, usually from countries where lower wages and fewer labor protections optimize development costs. Besides often performing monotonous and undervalued roles, with little room for professional growth, these workers face challenging conditions and are frequently unaware of the final impact of their work due to the fragmented nature of the assignments. Despite a few cases gaining global notoriety in the recent past due to litigation actions, the interviewed experts provided examples of issues that require much more attention than what is currently seen in public discourse:

● **Low wages and poor working conditions**

Workers are paid very low wages for long hours of repetitive, monotonous or highly stressing tasks with no opportunities for career progression.

● **Mental health impact**

Workers are often exposed to disturbing and traumatic content, leading to mental health issues such as anxiety, depression, and PTSD. When companies provide mental health support, it is often inadequate or insufficient.

● **Lack of labor protections**

Workers are employed as contractors without benefits, job security, or protections under labor laws. People from impoverished or marginalized communities are often more susceptible to accepting these conditions, which perpetuates their vulnerability.<sup>3</sup>

● **Surveillance and control**

AI systems are used to monitor and control workers, reducing their tasks to repetitive, low-skill activities and increasing their workloads.

● **Global inequity**

While workers in vulnerable communities, especially from the Global South, carry out the undervalued labor behind sophisticated AI products, tech companies in the Global North, especially in the Silicon Valley, reap the benefits.

# Bias and discrimination

One of the AI shortfalls that has been more frequently discussed in recent years is how the training data for algorithms largely reflects already existing biases and leads to discriminatory outcomes. Within this scope, the experts suggest nuances in which these biases are impacting results that merit media and public attention.

## 1 How data is trained and by whom<sup>4</sup>

There is a perceived lack of transparency behind aspects of algorithm data training. First, a few experts note that many of the workers performing the labeling are not aware of what such data will be used for, working under guidelines that can be unclear as to already existing biases or obscure motivations behind the request. In conjunction, the workers themselves can impart their perceptions into the data set, inadvertently or not. A deeper issue has to do with the observed focus on specific functional goals and lack of human rights and ethical considerations when briefing data labeling teams, which can result in damaging outputs. The specialists reinforce that bias is intrinsic to human nature, but that transparency as to the parameters for data training would help auditors and users in identifying potential issues.<sup>5</sup>

## 2 Reproduction of historic bias<sup>6</sup>

Another issue with the quality of training data is crucial in understanding potential biases within AI models. Historical data, such as that from criminal justice systems, often reflects systemic racism, leading to biased outcomes when used in AI models. For instance, using such data to predict parole decisions can disproportionately harm Black communities due to their overrepresentation in the system. Similarly, social policies based on outdated educational or income data can perpetuate biases, affecting new students or neighborhoods by reinforcing stereotypes or ignoring current realities. These biases highlight the importance of critically assessing the data's origins and the variables emphasized during model training.

## 3 The pitfalls of facial recognition

One of the most frequently cited examples of the problematic side of AI due to its many documented cases of racial and gender bias leading to wrongful arrests<sup>7</sup> or other discriminatory practices. The way in which it disproportionately affects young, Black people in urban areas was mentioned by nearly all respondents. In addition, this type of technology is seen as potentially causing more harm than good when deployed in governmental areas such as social services or security.<sup>8</sup> Policymakers and companies often opt for facial recognition solutions as a show of diligence and innovation, despite the potential flaws in the technology that can exclude marginalized communities and infringe on individual freedoms. This results in expensive, sometimes unnecessary tools that can perpetuate discrimination and exclusion.

## 4 Internet access disparity

Considering that a lot of the data labeled and trained into AI algorithms comes from user-generated information, while there is no equitable and effective universal access to the internet, the chances are that there will always be groups excluded from decisions made by the applications. When a story circles around people's access to technology or the lack thereof, the digital inclusion angle usually conveys all the ways in which users can benefit from that access. Notwithstanding, a disparity in quantity and quality of data from marginalized groups available for training AI models can perpetuate inequalities, and the fact that not having internet access could imply excluding entire groups from data sets that decide public policy and security decisions that affect them, for example, is rarely discussed.

## 5 Non-English content

Another shortfall of the data training and content moderation in many AI systems available today is the limited non-English content available to provide quality outcomes. While most recently companies and communities have been directing efforts to create native non-English large language models, either for profit or for cultural preservation, content moderation in social media still represents a challenge both in capacity and in willingness from the industry to engage in actions to retain, reduce, or redress the damage caused by the spread of harmful content.

## Environmental impact

The environmental impact of AI seems to be a neglected issue. A few of the experts mention that local communities where data centers are installed are aware of the damages as they are directly suffering the consequences from it. Anywhere else, however, it seems that there is a tendency to adopt an “out of sight, out of mind” approach, especially considering the frequently overwhelming news already surrounding the environment’s health. Nevertheless, nearly all the interviewed experts believe AI’s carbon footprint needs addressing, particularly due to the significant amount of electricity required to train AI models and its resulting carbon emissions, and to the large volume of water necessary for cooling data centers, which often strains local freshwater resources. As the climate crisis advances to points of no return, there is a perception that the sustainability of the AI industry is seldom challenged by the mainstream media. Other matters such as the production of electronic waste, air and water pollution, and environmental inequality were mentioned less frequently, but indicate potential topics to be tackled, nonetheless.

---

*“So, data centers are very water consuming, right? They demand a lot of water. (...) There’s a lot of data centers in the north of Chile and there are a couple of journalists talking to the people that don’t have access to clean water anymore because these data centers are demanding or taking all the water that they need. So, people can use ChatGPT or Dall-E and create stupid images and that is in contrast to people not having access to clean water in northern Chile, which is a huge, huge problem already.”*

---

---

*“From the point of energy use, you can see that in countries like Ireland, energy required to run AI systems is outweighing the amount of energy that is required to run households, commercial spaces and so on. So, this becomes unsustainable from an environmental sustainability point of view, because we are burning down the planet to meet the energy demands of these systems.”*

---

## Data governance and sovereignty

AI’s performance and accuracy rely heavily on the vast datasets used to fuel algorithms and train models. The quality, quantity, and diversity of data fed into an AI system directly impacts how the machine will learn, make decisions, and be refined to improve or adapt to new scenarios. Therefore, understanding the way in which data is collected and what happens to each piece of information after it enters an AI model is intrinsically important to identify the issues surrounding this intersection.

**Graphic 28** Data life cycle and key issues related to AI, according to the interviewed experts.





There is a consensus among the experts that the public's "why-should-I-care" attitude when it comes to sharing data indicates the need for more educational information on the data life cycle and the ethical and practical implications each stage represents. They consider data a valuable economic resource, which can potentially be a matter of national security. Yet, they believe most of society remains unaware of how exactly this is exploited. Moreover, understanding how data is sourced, used, and by whom is a difficult task in terms of regulation and accountability, since the way in which AI supply chains operate means data flows constantly across borders, where domestic protection laws vary and may be harder to enforce. Maintaining data sovereignty (i.e., nations having the authority to govern the data within their borders) is a major concern when it comes to pushing for transparency from the AI industry, and a few experts believe this issue cannot be tackled without global regulation efforts.

---

*"I think that there's also another issue that is very important that I want to emphasize, which is that technology essentially crosses borders, right? So, meaning to say, that even if you have your house in order, if you have a great... How do you say it? A great organization between government, citizens, and all that, you're still going to face the problem of having issues outside of your control."*

---

The different standards of data protection among countries can also reinforce inequalities, since countries where civil liberties are better protected tend to impose restrictions on data flows to protect their interests, while users from places with a weaker or nonexistent regulatory framework are more vulnerable to privacy violations and data misuse. A few of the experts point out that alongside the data issues, there is an urgent need to delve into the mechanisms of the AI supply chain and the various stakeholders involved in it, so that tech companies, governments, and civil society can informedly engage in data governance.

Another problem raised in some of the interviews has to do with the lenient standards when it comes to consent and transparency on how the data is sourced and used. In a context where people need to accept terms of use to access products or services nearly once a week, and where governments are steadfast in adopting well-marketed technologies, a mammoth amount of data is generated and embedded into AI systems for the most varied sectors, such as retail, finance, health care, education, and social services. However, it is hardly possible to access this information and understand how it is processed once it enters a system, either because it evolves into a trade secret or due to the red tape surrounding it. Consequently, understanding how AI products operate and finding evidence of potential biases and harm becomes a significant challenge.



# Privacy and security

One of the most concerning issues intrinsic to data management in AI is privacy and security. Sensitive personal information is often aggregated into AI datasets with or without consent, and data protection efforts seem relevant to the media only when there is news of big data breaches. However, this peak of attention does not necessarily represent a scrupulous monitoring of companies' and governments' behavior over time. The increasing connectivity of homes and personal devices, the current race to attach every access one needs to function in society to biometric authentication, and the ample use of AI systems in mass surveillance and the military all come with a list of red flags that, according to the experts, is quite superficially discussed. Not respecting privacy is a prerequisite for the business model to work, citizens' personal data is continuously collected, monetized, and distributed without their awareness or approval,<sup>9</sup> which makes addressing these issues paramount for the experts.

**Graphic 29** Potential privacy and security issues related to AI use, according to the interviewed experts.



## Communication and (dis)information

Most of the experts note that there is no shortage of coverage on how AI has radically changed the way in which we communicate, and the same can be said about how this has introduced significant challenges related to disinformation. It seems to be a popular understanding that algorithms help to quickly disseminate information according to specific goals (often paid for by corporate interests), and also that they make it easier for false news to spread, regularly outpacing efforts to verify and correct it. Concurrently, AI and mis/disinformation is the topic most frequently recognized as receiving media attention, according to the surveyed journalists (68%, see Graphic 13). Several media outlets have been working relentlessly in the battle between fact-checking and their audiences' cognitive or confirmation biases. Nevertheless, a few experts believe there is room to better illustrate the link between AI algorithms for circulating content and their real-life impacts. That improvement may lie in palatable, step-by-step, accessible information on how these systems work, helping audiences to make connections with their own contexts and stepping outside the hype bubble.

---

*"It's kind of an algorithmization of everything, right? (...) Their activism, their social actions, their cultural engagements, they are more and more determined by the algorithmic models of the communication platforms and how they strategize their own goals. So, I think this is a serious intersection of AI. (...) These algorithms make messages reach the target people, or not, and that can result in misinformation issues, in adapting language and stories all the time to satisfy the interest of large corporations."*

---

One of the examples mentioned by the specialists is the use of the technology in political campaigns. A significant concern is the use of deep-fake technology, which can create convincing fake videos or audio recordings that might mislead voters or damage reputations. Additionally, AI can be used in micro-targeting voters with personalized political ads, potentially influencing voter behavior by exploiting personal data. The experts also highlight the risk of AI being used to manipulate public opinion through automated bots on social media platforms, which can amplify misinformation or biased narratives.<sup>10</sup> These technologies can undermine the integrity of elections by spreading false information and creating confusion among voters.<sup>11</sup>

# Complexity and fallibility

The increasing “complexity” of AI systems and their widespread use represent a challenge in ensuring their reliability to deliver continuous quality performance. The interviewed experts were unanimous in highlighting that AI is fallible and that blindly believing the technology is a problem-solving panacea often has direct consequences. The fact that questions pertaining to potential system errors, auditing practices, and safety measurements are rarely raised in public discussions when governmental contracts of AI technology are announced represents a dangerous indication that such scrutiny is not being undertaken. The specialists also tend to believe that this is an area in which the media can make a difference, bringing these issues to light, ensuring that the public and policymakers are aware of the potential dangers and the need for robust oversight. Some of the most prevalent issues are discussed below.

## Assumed complexity of AI systems

The perceived complexity of AI systems obstructs regulation and improvement efforts, leading to gaps and difficulties in ensuring AI systems are improved and maintained properly.

## Techno-solutionism

The belief that technology can solve all problems often overlooks potential negative impacts and broader societal implications. This can result in over-reliance on technology and neglect of non-technological, but adequate, solutions. The over-reliance on AI for policing, ignoring community-based approaches, and the use of AI in education without considering teacher input are examples.

## Dependence on AI

Increasing reliance on AI systems can lead to overdependence, reducing human oversight, which leads to a lack of critical thinking and human judgment in decision-making processes. The use of AI diagnostic tools in health care resulting in doctors trusting less in their expertise and judgment, or educators using AI to grade student work without assessing the nuanced aspects of performance such as creativity and critical thinking skills can illustrate this issue.

## Lack of measurement instruments for AI safety

AI lacks comprehensive safety measures, making it difficult to ensure safe operation in real-world settings, posing risks to users. Due to the secrecy surrounding algorithm design, verifying if an AI system will not produce results that could put a user in danger is hardly possible. Self-driving cars are the most obvious example, but the malfunctioning of automated tasks in manufacturing or agriculture, for example, can cause physical harm for workers.

## Algorithm auditing

Regularly auditing AI algorithms is necessary to ensure they are functioning as intended and do not perpetuate biases. Data can be dynamic, project needs may change, and ensuring the systems are up to date with information and the quality of results is essential.

## Reliability and hallucination in generative AI systems

Generative AI systems can produce unreliable outputs, sometimes fabricating information (hallucinations). This can lead to misinformation and potential misuse in critical areas, such as health care<sup>12</sup> and legal advice.

## Scalability of errors

Errors in AI systems can be rapidly scaled across multiple applications and users, which can amplify the negative effects of any single error, affecting a large number of people.

## Difficulty in accountability

When AI systems fail, it can be challenging to determine who is responsible, a problem that can hinder justice and proper redress for affected individuals. For instance, when a facial recognition system used by law enforcement misidentifies an individual, it can be difficult to pinpoint whether the fault lies with the software developers, the data providers, or the law enforcement officers who relied on the technology.

---

*“We know for a fact that they hallucinate, they produce incorrect and false information. Yet, there is a massive appetite to deploy and integrate generative systems into various domains, including law enforcement, including knowledge production systems, for example, companies are incorporating generative systems into search engines, which is resulting in catastrophic impacts. We are integrating generative systems like chatbots in health care, in medicine, without fully addressing that they hallucinate, they give incorrect information.”*

---

All of these topics represent the most pressing issues identified in the discussions with the interviewed experts. Although not representing an exhaustive list of points that need addressing, they provide a myriad of approaches for investigative journalism and engagement initiatives that could aid in demystifying the AI hype by provoking a more critical assessment of the technology and the consequences of its deployment. Nevertheless, there is a firm belief that AI governance and regulation are fundamental to enabling a fairer and more conscious use of the technology, and that there is still a long way to go in terms of developing and enforcing that. The following chapter presents an analysis of the discussions around this topic, which was also considered an urgent issue that requires media attention.





# AI governance and regulation

The current state of AI governance and regulation is characterized by a mix of emerging frameworks and ongoing challenges. Both experts and journalists interviewed recognize the industry is insufficiently regulated, and the first are nearly undivided in emphasizing the need for robust regulatory mechanisms to ensure the ethical deployment and accountability of AI technologies. The Media Cloud analysis shows that keywords such as “biometrics,” “surveillance,” and “facial recognition” obtain higher news volume, indicating attention to government-deployed AI technologies. Moreover, peaks in attention correlate with events like the European Union (EU) AI Act and United Nations (UN) resolutions on AI Governance. Similarly, keyword analysis shows that specific policies and laws such as the EU AI Act, executive orders in the U.S., and recommendations from the UN Secretary-General are frequently searched.

The social listening assessment also shows public interest in AI governance and regulation, particularly in terms of ethical practices, transparency, and accountability. Government responses to AI regulations have long-term relevance and high engagement, particularly when shared by government officials and business leaders. Content emphasizing the need for regulation or ethical AI practices tends to resonate with audiences, indicating a demand for more critical examinations of AI risks. Furthermore, litigation has been effective in raising the profile of AI accountability issues, particularly in the context of labor and AI supply chains, as in one of the case studies from the digital assessment, where a case about data labelers and moderators gained 39,000 engagements and a high evergreen score, obtaining worldwide prominence.<sup>13</sup>

Despite the identified interest the topic generates among audiences, and likely due to the fast-changing nature of the technology, AI governance and regulation frameworks are continuously in evolving stages, and uncertainty surrounds questions of compliance, enforcement, and the impact on innovation. The interviewed experts provided a few examples of regulatory frameworks available today and what they see as critical points for each:

#### ● **EU AI Act**

A comprehensive regulatory framework aimed at standardizing AI practices across the European Union. It categorizes AI systems into different risk levels and imposes requirements accordingly. High-risk AI systems must meet stringent requirements for data governance, transparency, and human oversight ([EU Artificial Intelligence Act](#)). The Act is still in the legislative process and has not yet been fully implemented. Critical points include its bureaucratic complexity and high compliance costs, which could stifle innovation.

#### ● **General Data Protection Regulation (GDPR)**

GDPR is a data protection regulation that impacts AI technologies by enforcing strict data handling and privacy standards. It applies to all organizations operating within the EU and those handling the data of EU citizens. Implemented in 2018, GDPR mandates transparency in data processing and grants individuals' rights over their personal data ([GDPR.EU](#)). Its stringent requirements have been considered a challenge for smaller organizations to meet.

#### ● **California Consumer Privacy Act (CCPA)**

CCPA is a state-level regulation in California that grants consumers rights over their personal data and imposes obligations on businesses handling such data ([State of California Department of Justice](#)). It impacts AI technologies by enforcing transparency and accountability in data processing. Implemented in 2020, CCPA is used within California but has influenced other U.S. states to adopt similar regulations. As with the GDPR, small businesses are likely put in disadvantage due to strict requirements.

#### ● **Singapore's AI Verify**

AI Verify is a framework developed by Singapore to encourage responsible AI deployment. It provides guidelines for AI governance and encourages vendors to self-report according to specified criteria ([Personal Data Protection Commission, Singapore](#)). This approach aims to foster transparency and accountability without imposing strict regulations. Critical points include its reliance on voluntary compliance, which may limit its effectiveness.

### ● **Brazil's Bill 2630**

Also known as the “Fake News Law,” Brazil’s Bill 2630 focuses on the regulation of digital platforms to combat misinformation. It includes provisions that impact AI technologies used for content moderation and dissemination ([Brazilian Legislative Chamber](#)). The bill is still under discussion and has not yet been fully implemented. The experts mention concerns about freedom of expression and the potential for government overreach have been raised by opposers of the bill. Although disagreeing with the issue, they highlight that this type of argument can easily gain a lot of traction in public discourse.

### ● **Regional Comprehensive Economic Partnership (RCEP)**

RCEP is a trade agreement among 15 Asia-Pacific countries that includes provisions for digital trade and data governance, indirectly impacting AI technologies. It aims to facilitate cross-border data flows while ensuring data protection. Implemented in 2022, RCEP is used across its member countries ([Asian Development Bank](#)). Its broad scope and the challenge of harmonizing regulations across diverse legal systems are critical points.

### ● **Comprehensive and Progressive Agreement for Trans-Pacific Partnership (CPTPP)**

CPTPP is a trade agreement among 12 countries that includes provisions for digital trade and data governance, affecting AI technologies. It promotes the free flow of data and prohibits data localization requirements. Implemented in 2018, the CPTPP is used across its member countries ([Investment Policy Hub](#), [UNCTAD](#)). Balancing data protection with the need for open data flows is the main challenge.

### ● **UN Resolution on AI Governance**

The UN Resolution on AI Governance aims to establish global guardrails for AI technologies, promoting ethical AI development and international cooperation. Introduced by the U.S. government and co-sponsored by 120 member states, the resolution calls for transparency, accountability, and human rights considerations in AI deployment ([UN News](#)). While not legally binding, it sets a framework for member states to develop their own regulations. Its broad and non-binding nature limits its enforceability.

The interviewed experts highlight that the EU has been at the forefront of the regulatory initiatives with comprehensive frameworks like the AI Act and the GDPR, bringing forward much-needed standards to ensure transparency and accountability. However, the bureaucratic complexity and high compliance costs are hurdles that many do not see being easily cleared. In contrast, the U.S. takes a piecemeal stance, with state-level regulations such as the CCPA and a reliance on self-regulation and industry standards. This method is presented as allowing for more flexibility and fostering innovation but also as providing fragmented and inconsistent protection across state and international borders, generating challenges in compliance, oversight, and accountability.

---

*“There’s a real debate to be had about a European style AI act, more regulation of models, versus a U.S. style lack thereof, and about ‘innovation’ versus ‘safety’. I would really love to see intelligent examination of those values trade-offs much more than I would just about anything else, because those give us some guidance about what we actually can do in the world.”*

---

A pragmatic realization that there is no optimal solution to all contexts is shared by the specialists. There is an apprehensive resignation that, regardless of the path chosen, the most vulnerable stakeholders, such as citizens and smaller businesses, will be at a disadvantage. Nevertheless, they reinforce the need to bring these regulatory issues into public discussion, and that the media serves a pivotal force in igniting debates that could help to shift the power balance between the tech industry interests and the protection of the public. The graphic below shows some of the challenges identified as facilitating the current poorly regulated AI landscape, and that the experts believe should be tackled to advance society into more productive discussions.



**Graphic 30** Challenges for AI governance and regulation, according to the interviewed experts.

### KNOWLEDGE AND ENGAGEMENT

**Lack of preparedness among policymakers:** The experts believe many decision-makers are not well-informed or adequately trained to discuss and decide on AI-related matters, which hampers effective regulation. This lack of preparedness can lead to uninformed decisions that fail to address the complexities of AI technologies.

**Stakeholder engagement:** According to the experts, effective AI governance requires meaningful engagement with a diverse range of stakeholders, including civil society, academic experts, and impacted communities. There is a perception that this is challenging due to the scale and assumed complexity of AI technologies.

### STRUCTURAL AND ORGANIZATIONAL

**Coordination issues:** The specialists point to historical difficulties in achieving collaboration between parliamentarians, public managers, and civil society that lead to fragmented efforts and ineffective policies. Moreover, they reinforce that sector-specific solutions are needed to adapt existing laws to AI-specific issues within different industries like health, finance, and education.

**Implementation, enforcement, and inequities:** Even when regulations are in place, the practicalities of implementation and enforcement can be problematic due to resource constraints and the need for specific guidance. A few interview participants argue that this leads to varying levels of protection and compliance across different regions, types and sizes of companies, and sectors.

### TECHNICAL AND GLOBAL STANDARDS

**Need for language and standardization:** Some of the interviewed experts mention that developing technical standards to help human auditors and regulators efficiently assess AI systems is crucial. This would include creating standardized forms for AI providers to communicate necessary information to regulators.

**Global governance:** A few specialists believe that the cross-border nature of AI technologies necessitates a global governance framework, but they also recognize that many countries lack comprehensive regulatory frameworks similar to the EU's. Differing priorities and power structures among countries add complexity to international cooperation.

### RESOURCE ALLOCATION

**Resource constraints:** The specialists note that many civil society organizations and affected communities lack the resources to engage effectively in AI governance processes. Ensuring that resources are available to monitor AI impacts and involve all stakeholders is considered a significant challenge.

**Data access for evidence generation:** A few of them also highlight that there is a significant challenge in obtaining access to meaningful data in order to understand the impact of AI systems on society. Without this data, it is difficult to generate the evidence necessary for informed policymaking.

### AGILITY AND ADAPTABILITY

**Lack of agility in government structures to keep up:** Many of the interviewed experts recognize that existing bureaucracies and institutions struggle to adapt to the fast-paced nature of AI technology, leading to coordination problems and outdated regulations.

**Incorporating ethics to innovation:** The specialists observe that a lot of the discussions surrounding AI governance reproduce the idea that regulation may slow down or prevent the development of certain technologies. While some of them believe this is a valid debate, many also point to the need to find ways to innovate ensuring ethical considerations.

### POWER DYNAMICS AND INTERESTS

**Concentration of power and interests:** Most of the experts agree that AI tends to concentrate benefits towards the powerful, leaving the powerless to deal with the repercussions without necessarily reaping the benefits. This concentration of power can lead to biased decision-making and lack of accountability.

# Brainstorming solutions

Generating ideas for solutions is also an exercise that requires the input of all impacted stakeholders in each context. The experts express concern about how much of these discussions still happen behind closed doors, which results in essential voices being excluded from understanding how decisions are made. This not only represents a transparency issue but also significantly limits creative and effective solutions. When asked to share ideas based on their experiences and fields of work, they do so with the reminder that while they stand at an advantageous point in terms of specific knowledge, their suggestions need to be balanced against localized and global expertise, including impacted communities, local administrations, business owners, and advocacy groups. This information can be helpful for journalism and engagement work in AI, providing an understanding beyond the gap and possible pathways to present to the audiences or for deeper investigations. Some of the shared ideas are presented below.

- **Central agency for AI issues**

Establishing a central agency to coordinate AI governance, facilitate information flow, and build partnerships among stakeholders. This agency would facilitate partnerships between technology developers, policymakers, and civil society.

- **Sector-specific regulations**

Instead of a horizontal regulatory framework, focusing on vertical sectors like health, finance, and education to adapt existing laws to AI-specific issues.

- **Dedicate efforts to build governance criteria**

Fast-tracking the research and development of technical standards to efficiently assess AI systems within a regulatory perspective and standardize the language necessary to establish fluent communication between AI providers and regulators.

- **Resource allocation from corporations**

Encouraging corporations benefiting from AI to reinvest profits into awareness campaigns, grievance mechanisms, and monitoring systems to ensure sustainable AI governance.

- **Regulatory sandboxes**

Creating environments where new technologies can be tested under regulatory supervision to understand their impacts before broader deployment.

- **Inclusive stakeholder governance**

Ensuring that all necessary stakeholders, including civil society, affected communities, and small players are involved in the AI governance process to address diverse perspectives and priorities.

- **Continuous monitoring and adaptation**

Regularly reviewing and updating regulations to keep pace with technological advancements and emerging ethical concerns.

- **International cooperation**

Enhancing global governance mechanisms to ensure that AI regulations are consistent and effective across borders.

---

*“So, what we can do is to start trying to engineer some existing regulations to apply to AI technologies as they are deployed in real life. (...) For example, financial service providers. In Mauritius, they have an industry regulation that requires the financial service providers to disclose certain information about how they're using AI technology to reach financial advice. It doesn't really prevent them from doing it. It just kind of requires a certain amount of transparency and disclosure. That you can develop today.”*

---

---

*"I'm a regulator. Maybe I'm from the Food Safety Authority and I go to some restaurant, and I ask the manager: 'We want to make sure that your food is actually safe. Please produce evidence that your food is safe'. I think the manager is going to have a difficult situation. What kind of evidence is going to be sufficient or can please this officer? But if the officer says: 'Please show me the temperature of your freezer. Please show me things like that'. So, if you can come up with a list, right? We are probably in danger of being in a checklist regulation, but at least this is something."*

---

## Grievances and redressing mechanisms

The need for redress and grievance mechanisms in the AI industry is generally unquestioned by the interviewees; it is widely acknowledged as essential. These mechanisms are crucial for ensuring accountability, transparency, and fairness in the deployment of AI technologies. They provide a way for individuals and communities affected by AI to seek remedies and hold companies accountable for any harm caused. The experts emphasize the importance of these mechanisms in addressing issues such as human rights abuses, discrimination, privacy violations, environmental damage, among other negative impacts of AI.

Naming frameworks that are effective in offering and enforcing public protection, however, is more difficult than these specialists believe it should be. The fines calculated based on a company's global income under the EU's GDPR are mentioned as incentivizing compliance, but with a reminder that enforcement remains a challenge. Nevertheless, the experts note that the more advanced regulatory landscape and the higher public awareness and advocacy facilitate a more protected environment for these communities, mostly located in the Global North. In contrast, they believe that regions like Southeast Asia, Latin America, and Africa struggle with the lack of comprehensive regulation and often have to rely on long, costly, and very challenging litigation processes to obtain any form of redress, which is explained, in part, by the fact that those processes involve companies based in the Global North. Even so, without advocacy and media attention, the experts' perception is that human rights in marginalized communities will continue to be violated behind a veil of advancing towards a technologically enhanced future.

### AI and teenage pregnancy prediction<sup>14</sup>

In Salta, north of Argentina, a partnership with Microsoft signed in 2017 promised to deepen the province's digital transformation by preventing some of its 'most urgent social problems'. Its governor at the time proudly announced that "with this technology, it is possible to predict five or six years in advance, with name, surname, and home address, which girl—the future teenager—is 86% likely to have a teenage pregnancy." The criteria established for the algorithm to make these predictions, or what exactly happened to each girl identified as a potential 'risk', were less openly discussed. Girls aged 10-19 years old had sensitive personal information such as age, ethnicity, disability, and geographical coordinates of their households gathered and added to a system then called pioneering in the use of AI for social security policies. Social uproar against these privacy violations and the stigmatization of vulnerable groups, particularly from local feminist movements, seems to have tempered the enthusiasm of those eager to present results of the initiative. But without AI regulation and a formal grievance channel, there is no information about what happened with the collected data, whether the program was suspended, or any resolution in sight for the victims.

### AI and the safety of its content moderators

The exploitation of Kenyan workers in the AI industry has garnered significant attention, highlighting severe labor abuses. Workers employed by outsourced companies to moderate content for AI tools have faced harrowing conditions. These low-paid workers are exposed to graphic and distressing content, including violence, abuse, and serious criminal acts, leading to long-term psychological distress and mental health issues. Despite their crucial role in ensuring the safety of AI-generated content, these workers have received inadequate mental health support and faced abrupt dismissals when raising concerns. Complaints have been directed toward both their direct employers and the larger tech companies they serve. For example, content moderators for Facebook in Kenya faced union-busting tactics, with entire workforces being sacked when attempting to organize. Similarly, workers for ScaleAI were left unemployed and owed significant sums when the outsourcing firm Remotasks exited the African market overnight. The negligible worldwide awareness of AI supply chain processes, the nonexistent local regulatory framework, and the lack of an adequate grievance mechanism have exacerbated these issues, forcing workers to seek litigation to address their complaints. Media coverage and advocacy efforts have brought some attention to these problems, but resolutions remain limited. The absence of effective grievance mechanisms means that workers often endure prolonged legal battles to seek justice and fair treatment.

### AI and the defiance of data protection laws

The ban of X in Brazil serves as a significant case study in regulatory intervention and its cascading effects on both companies and users. The prohibition was primarily driven by concerns over data privacy and security, as well as the potential misuse of the platform, using AI algorithms to spread misinformation. Regulatory bodies in Brazil found that X was not in compliance with the country's stringent data protection laws, leading to an initial warning followed by a full-scale ban when the company failed to make the necessary adjustments. The situation escalated rapidly due to the company's perceived negligence in addressing the regulatory concerns, and its CEOs blatant attack of the justice behind the decisions. This not only led to the ban, fines, and threats to further legal actions, but also sparked widespread public debate and media coverage. The consequences for the company were severe, including loss of market share, revenue, and a tarnished reputation. For users, the ban meant losing access to a platform they relied on for various services, leading to frustration and a scramble for alternatives. Users coped by migrating to other platforms, and though this transition was not seamless, after a 38-days suspension the force of public outcry instigated and estimated by the company was not sufficient to make the Brazilian Supreme Court retreat. Ultimately, X complied with the legal demands to have the ban lifted and settled over 5 million US dollars in fines before it was allowed to resume its operations in the country.

The experts suggest that establishing and enforcing these measures does not necessarily mean coming up with an entirely new mechanism that would encompass all concerning issues, but rather a more concerted effort to uphold AI companies at least to the same standards companies in other sectors are expected to uphold, such as respecting workers' rights and environmental laws.

---

*"One thing is to comply with the law, environmental law, labor law, for instance, which are things we don't observe today. (...) Companies need to comply with the existing laws of the countries they operate in... In no way should we think that this sector would be exempt from complying with the existing legislation today."*

---

Additionally, they observe that localized governance efforts could be aided by examples from other industries, such as:

● **Financial services**

The financial industry often uses ombudsman services to handle customer complaints. AI companies could establish similar independent bodies to address grievances related to AI decisions.

● **Health care**

In health care, patient advocacy groups offer substantial support in addressing grievances. AI firms could create user advocacy groups to represent the interests of those affected by AI technologies.

● **Telecommunications**

Telecom companies are required to provide clear channels for customer complaints and have regulatory bodies overseeing them. AI companies could implement similar transparent complaint mechanisms and be subject to oversight by independent regulators.

● **Environmental regulation**

Environmental impact assessments are mandatory before large projects can proceed. AI companies could adopt similar practices by conducting ethical impact assessments before deploying new technologies.

● **Consumer protection**

Consumer protection agencies often have the authority to recall products and issue fines. AI regulators could be empowered to take similar actions against harmful AI applications.

An irrefutable point is that public awareness and knowledge about the impact of AI in different contexts are key to boosting discussions and decision-making about AI governance and regulation, and media coverage remains the most efficient tool to engage audiences with these issues. The social listening exercise indicates that government responses to AI regulation, content emphasizing ethical and transparency concerns, and regulation and litigation stories are already capable of obtaining relevance, sometimes for a longer period of time. Capitalizing on this tendency might be the key to deepening the discussions and generating a more long-standing impact in terms of public protection.





# Engaging the public and building awareness



Public awareness about AI has increased significantly over the past decade but is still considered low and misaligned with reality. The widespread media coverage of AI solutions and the speed at which these have advanced into many facets of daily life does not necessarily represent an enhanced understanding of how the technology works and all the people it impacts. The most common perceptions tend to be either as a miraculous solution to many problems or as an existential threat to humanity. Nuances such as current limitations, ethical considerations, and the need for governance structures are often overlooked.

A few respondents bring attention to what they call the *anthropomorphization* of AI, that is, a typical narrative from tech PR that often attributes human characteristics such as emotions or intentions to AI systems. Popular culture and the media are seen as contributing to it, and this can lead to misleading narratives, including:

#### ● **Overstating capabilities**

Marketing strategies frequently portray AI as having near-human intelligence and capabilities, which is often reproduced in popular media. For example, movies like *Ex Machina* and *Her* depict AI with advanced emotional intelligence and autonomy, which is far beyond what the technology can currently achieve.

#### ● **Existential threats**

Films such as *The Terminator* and *I, Robot* present AI as a sentient risk that could cause the collapse of human civilization. This narrative, which is also commonly found in media coverage, overshadows more immediate and realistic concerns, such as bias in algorithms and data privacy issues.

#### ● **Simplifying AI development**

Media often simplifies the complexity involved in developing AI, echoing corporate speech. Shows like *Westworld* present highly sophisticated AI without covering its creation process, overlooking the extensive research, ethical considerations, and technical challenges involved.

#### ● **Ignoring ethical issues**

Also following tech companies' script, many portrayals focus on the technological marvels of AI while neglecting the ethical dilemmas. For instance, the movie *Iron Man* showcases AI as a benevolent assistant without addressing the potential for misuse or the ethical implications of AI in warfare.

The resulting misconceptions and overexcitement about AI capabilities, the fear and distrust generated by dystopian narratives, the ethical and legal confusions to ensure accountability, and the neglect of human agency shape the public's apprehension and enthusiasm about AI, which heavily impacts policy decisions. Therefore, the compelling nature of demystifying AI hype lies in ensuring that journalistic coverage can contribute to directing these technological advancements and their corresponding governance structures towards a transparent and equitable distribution of responsibilities and benefits.

The interviewed experts emphasize the immense power AI technology has to both exacerbate and mitigate inequalities. On the one hand, the concentration of power and knowledge in the hands of a few tech giants leads to significant societal repercussions that often disproportionately affect marginalized communities, as observed in a previous chapter. On the other hand, AI's specific capabilities can be leveraged to address social issues such as healthcare disparities and educational inequalities, as long as there are needs assessments, model adaptation, human auditing and oversight, and compliance with regulations to protect user rights and security.

---

*"So, big tech is a darling to the African governments. In my field of work, I have found that when you vocalize that the big tech isn't doing something correctly in Africa, you're a liar and you are the enemy of progress. 'Big Tech is opening up opportunities.' 'It's employing young people.' Nobody is asking what kind of jobs they are. Nobody is asking what's really behind the scenes."*

---

The path to reducing the power imbalances between those who own the technology and the decision-making roles and those who should be benefiting from it, according to the interviews, starts by fostering a more critical view of AI, its realistic capabilities, and its shortcomings. This includes applying effective communication strategies to educate the public and policymakers, shifting the discussion from hypothetical risks and benefits to existing harms and practical applications. It requires engaging technologists, politicians, and civil society in a more pragmatic dialogue that goes beyond selling points and the urge to be at the digital forefront. The quality of AI reporting is considered directly correlated with the level of resistance from audiences to accept superficial or misleading information.

# Improving AI coverage

High-quality AI coverage has various elements and can be very context-specific. Experts state that good stories show a solid grasp of the potential and limitations of the technology but, more importantly, a holistic idea of its impact before and after deployment. Diverse perspectives that include more than companies' claims are expected, as well as those coming from workers, consumers, and any voices that can contribute to a critical understanding of how each issue affects people's daily lives.

---

*"You're already covering the economy like this. You already cover political or international news like this. There was a bombing somewhere, what are the consequences of that bombing for the relations of these two countries? Journalists are doing that, and you know that you need to do it because it's basic journalism. Why are you not doing it for tech?"*

---

Although critical of how journalism has been approaching the topic overall, the experts recognize the limitations in journalists' training, the lack of local examples, the assumed complexity of the topic, and the limited or misguided interest driven by the hype as barriers to reporters and editors, issues also highlighted by the surveyed journalists when asked about training needs (see Graphic 32). These demands indicate that there is still a path to be taken from initially understanding AI to then exploring the topic in depth, which represents an opportunity to improve AI coverage, demonstrating its impacts in society for better or worse. Support options that include autonomous learning (reading resources for self-teaching and AI reporting examples in one's language) do not resonate much with journalists in the sample.

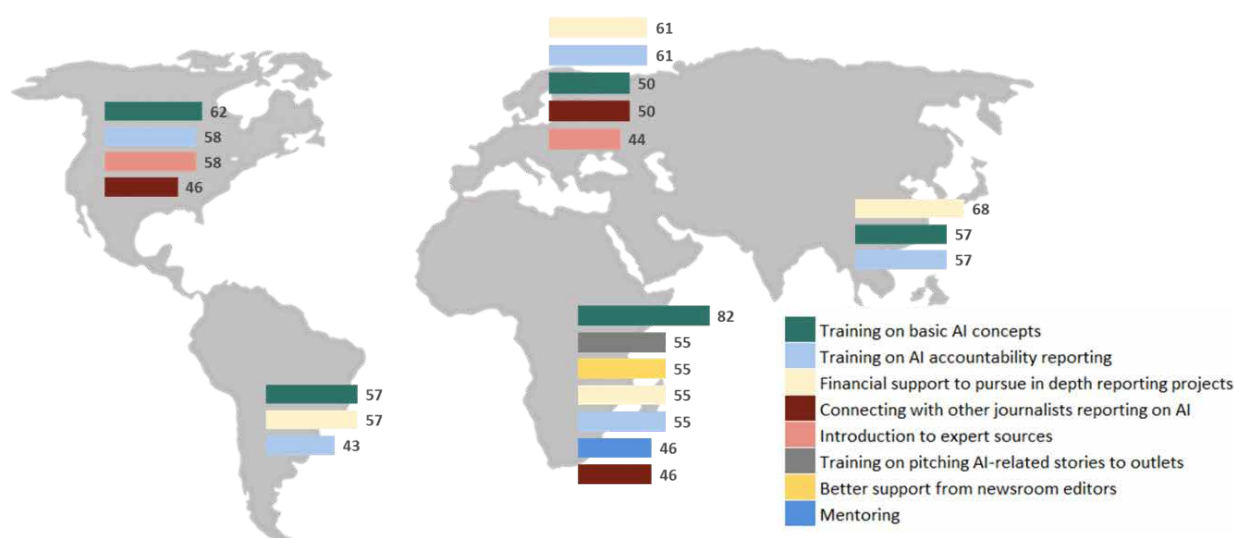
**Graphic 31** Training or other support journalists reporting on AI (or eager to do so) in your region need, % of cases (N=90).



Other answers mentioned spontaneously: Access to more free AI tools (1%), AI training linked to job opportunities for young people (1%), recognition and opportunities for new journalists (1%).

This greater demand for understanding basic AI concepts is concentrated in Africa, North America, and South America. The demand for financial support to pursue in-depth reporting projects is prioritized in Asia, Europe, and South America. Indeed, the interviewed experts from these regions mention a growing interest in AI governance and, for Asia and South America, significant concerns with government spending on AI solutions and its consequences in terms of racial bias and privacy violations. The African journalists who participated in the survey also concur with the specialists interviewed in the region regarding the need for a wider range of assistance, including ways to pitch AI stories to media outlets and receiving support from their editors.

**Graphic 32** Training or other support journalists reporting on AI (or eager to do so) in your region need per continent, % of cases, options with a frequency of 40% or more (Africa N=11, Asia N=28, Europe N=18, North America N=26, South America N=7).



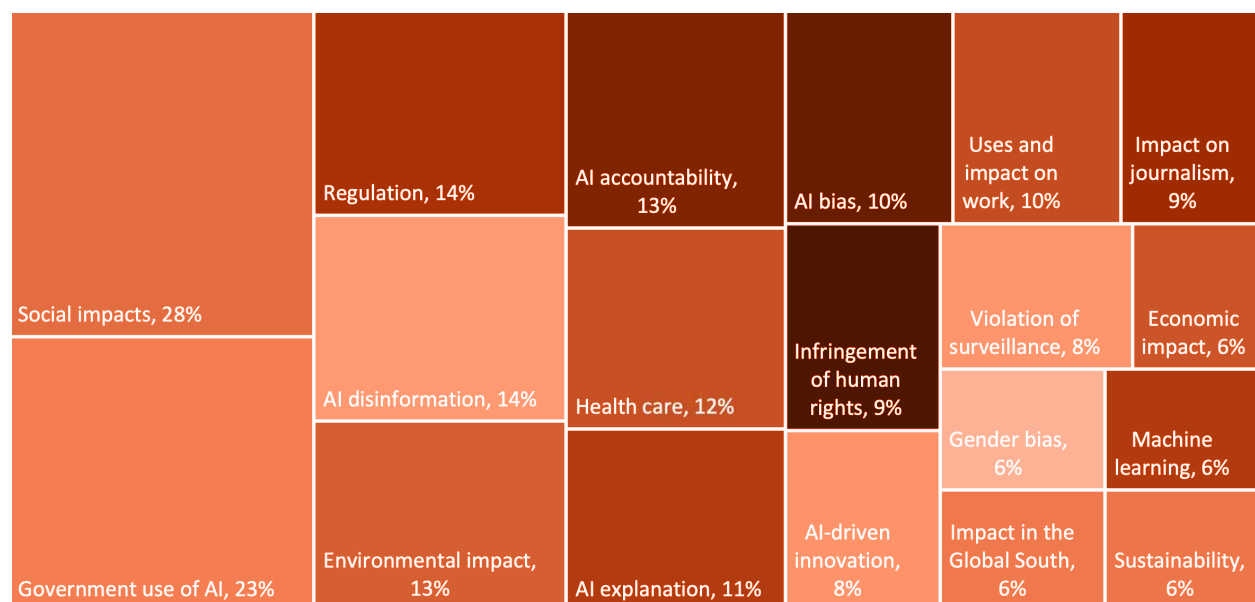
The experts point out that cultural narratives and social welfare represent an essential factor in developing training for journalists regarding how they will approach each story. Different attitudes toward technology and the societal issues it promises to solve need to be considered so that a more critical or educational approach does not backfire into a distrustful view of the media, instead of the topic itself. Regions that have been struggling with high crime rates and violence might be more vulnerable to believing, for example, that trading privacy rights for the security expected from an AI surveillance system is fair, even if it makes mistakes. On the other hand, in places with a stronger level of criticism of the technology and stricter regulations, a few specialists note that smaller businesses can be at a disadvantage due to compliance costs, which can create competitive disadvantages. Hence, the perspectives in journalists' training and the mining of sources need to go beyond the technological sphere and include local social scientists, community leaders, representatives of advocacy groups, small businesses, etc.

# What journalists would like to see

Journalists are interested both in the technological aspects of AI and in its broader societal implications. They seek to understand how AI can exacerbate or mitigate social inequalities, influence government policies, and affect everyday life. In the survey, they expressed a keen interest in a variety of AI-related topics. The most frequently mentioned area is the social impacts of AI, cited by 28% of respondents. Indeed, the Media Cloud analysis shows that issues such as AI bias and AI harm are underreported in comparison to the baseline key words, which could suggest that the lack of content that is crucial to understanding how AI affects different communities resonates among journalists.

The news volume study also indicates that keywords such as 'biometrics,' 'surveillance,' and 'facial recognition' are most linked to government-deployed AI technologies. Such a connection suggests an interest in how governments are using AI, a phenomenon that can also be observed in the second topic the surveyed journalists are most interested in pursuing or reading about in the media (23%). Other significant topics spontaneously mentioned in the survey include regulation, AI disinformation, environmental impact, AI accountability, and healthcare, each mentioned by 12% to 14% of journalists. None of these issues are identified as obtaining significant attention in the Media Cloud analysis, which could corroborate the perception of them being underrepresented in AI stories. Expressions related to healthcare, however, obtain a high search volume according to the keyword analysis, which could represent a curiosity among internet users about the impact of AI in the industry that is possibly not being met by content produced outside of AI providers.

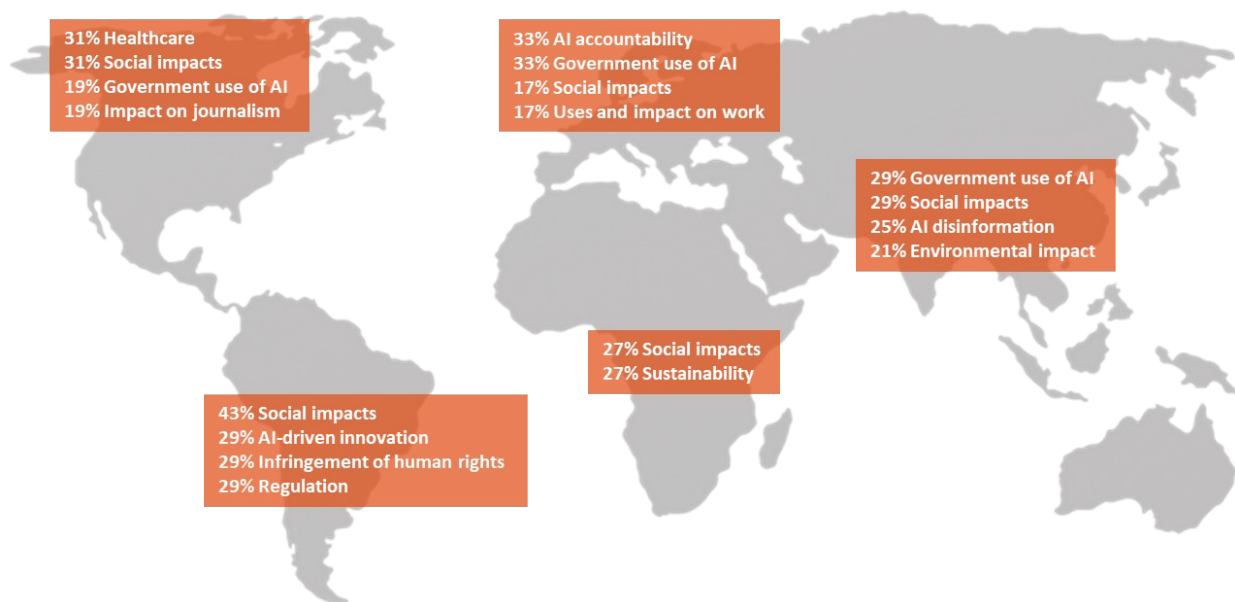
**Graphic 33** AI-related stories journalists wish to pursue or see reported on, % of cases (N=90).



Other answers mentioned include: Uses and impact on war (3%), recognizing AI generated content (2%), special needs communities (2%), control (1%), copyright abuses (1%), technical algorithmic audits done by journalists (1%), uses and impact on agriculture (1%).

The demand for information about the social impact of AI is particularly high in North and South America. In Asia and Europe, the primary focus is on the use of AI by governments. North American journalists also prioritize healthcare and the impact of AI on journalism. In South America, there is a strong interest in AI-driven innovation and regulation. European journalists are focused on AI accountability and its impact on work, while in Asia, environmental impact and AI disinformation are key concerns. In Africa, sustainability is an outstanding topic.

**Graphic 34** AI-related stories you wish to pursue or to see other journalists in your region report on per continent, % of cases (Africa N=11, Asia N=28, Europe N=18, North America N=26, South America N=7).





# The AI Accountability Network

Initiatives such as the Pulitzer Center's AI Accountability Network spark overall optimism among the consulted experts. A high standard is associated with the work of the Pulitzer Center, and they believe that training and support from the fellowship will help journalists create more accurate and in-depth stories, bringing attention to urgent concerns and underreported areas. They expect that the investigative stories to come could help increase public awareness and influence policies. Furthermore, establishing a network seems beneficial not only to journalists but also to civil society groups working to disseminate information more broadly and to create a sense of connection between issues across different regions. Engagement programs are also welcomed to target diverse stakeholders, which would be extremely valuable for shaping effective AI policies.

---

*"I mean, AI is the next big thing, digital futures, right? Take such a big thing and demystify it. Demystify and make the connection. If this one country's political strategy is just to push for digital work, this is what it looks like. These are the ways in which that country is failing its people, by not having the proper checks and balances to make sure that the work that's been pumped in is not only ethical but also results in ethical outcomes. So, for me, I see a lot of value in getting journalists into a program where they quite understand all the moving parts of AI."*

---

Their enthusiasm is only tempered by an awareness of the substantial obstacles such initiatives may face. There is a perceived concentration of AI expertise within companies, which sometimes leads to the impression that it is difficult to find independent experts or research that are not tied to big tech corporations in any way. Additionally, there is concern about the time and funding required for journalists to develop a deep understanding of AI issues and the potential for AI reporting to be swayed by powerful tech companies. However, this is an association made with mainstream media in general, and not attributed to the Pulitzer Center's initiative. In that sense, it becomes key for the AI Accountability Network to communicate its independence both as a sign of credibility and as a demonstration of thriving projects that are not attached to big tech.

Nonetheless, the specialists praise the Pulitzer Center for its initiative to approach these issues, and tend to have faith that the organization is capable of producing important results in media coverage about AI. Within their scope of expertise, they provide some suggestions on how to overcome some of the hurdles cited.

**Graphic 35** Suggestions from the experts on how to widen the scope of journalists' training for reporting on AI issues.



### Solidarity journalism

Solidarity journalism has been mentioned by a few experts as a way of shifting the viewpoint and inverting the power distribution through these stories. It is an emerging approach that emphasizes reporting on issues from the perspective of those directly affected, particularly marginalized communities. This reporting seeks to amplify the voices of individuals who are often overlooked in mainstream media coverage. It is not necessarily about giving everyone a voice but rather approaching a bottom-up way of presenting important issues. In the context of AI, solidarity journalism can improve coverage by focusing on the real-world impacts of AI technologies on diverse populations. For example, rather than merely highlighting technological advancements, journalists practicing solidarity journalism would investigate how AI affects workers in industries like recycling or the availability of resources in communities near data centers. This method intentionally seeks to foster social justice by ensuring that the narratives around AI include the social, economic, and cultural consequences of these technologies, providing a more comprehensive and critical view.

# Rethinking education

For the specialists, the unquestionable need to bring deeper discussions about AI into all levels of education puts educational institutions in the spotlight. They are considered responsible for teaching future journalists, technologists, policymakers, and the general public foundational knowledge about AI and its ethical, cultural, and economic dimensions. However, this is yet another sphere that these experts see reduced to either too much faith in AI tools to guide learning goals or a flat rejection due to the use of technology as a shortcut for academic work.

The experts believe there is much to benefit from AI in learning environments, as long as conscientious use is employed. They also consider that there should be space to approach these issues at any school level. For instance, younger pupils should be introduced to basic concepts of AI and ethics in the curriculum. As with coding and programming, which are being incorporated earlier than ever into primary and secondary schools syllabuses, projects that require problem-solving and critical thinking combining AI education with social studies and ethics could provide young citizens with a more holistic understanding of the tools presented to them. Additionally, privacy issues and internet conduct need approaches that go beyond listing all the ways in which they should not behave, fostering their own reflection on how and why they engage with content.

---

*"And so what they need to do is ask ChatGPT in 500 words to argue whether (...) news recommendations are a positive development for society. So they will copy+paste the response from ChatGPT and then they need to contradict ChatGPT... If you ask ChatGPT to explain that audience metrics was a good development for journalism, then you need to argue that actually audience metrics was not necessarily a good development for journalism. And that worked really, really well because it does two things. The first thing is you can see students kind of know what they're talking about. Because the point of the class is to recognize all the challenges that these technologies come with. But second, it really shows the limitations of ChatGPT in the argument."*

---

When it comes to higher education and professional development, there is room for more directed efforts, such as specialized courses, research opportunities, continuous learning initiatives, and collaborative platforms to facilitate discussions and updates both for students and educators. According to the experts, educational institutions can enhance AI curricula and foster critical thinking through collaboration with civil society and industry experts in several ways:

● **Guest lectures and workshops**

Inviting independent researchers, industry professionals, journalists, and advocacy representatives to give guest lectures or conduct workshops can provide students with real-world insights and up-to-date knowledge about AI technologies or their implications.

● **Joint research projects**

Partnering with tech companies or civil society organizations on research projects can give students hands-on experience with AI tools and methodologies for measuring impact, while also contributing to the development of new technologies and ethical guidelines.

● **Curriculum development**

Developing and updating course contents with the collaboration of civil society and research experts ensures that the curriculum remains balanced, relevant, and aligned with current industry and regulatory standards and practices.

● **Ethical training**

Incorporating ethical training into AI courses, with input from advocacy groups and industry professionals, can help students understand the broader implications of AI technologies and the importance of responsible innovation.

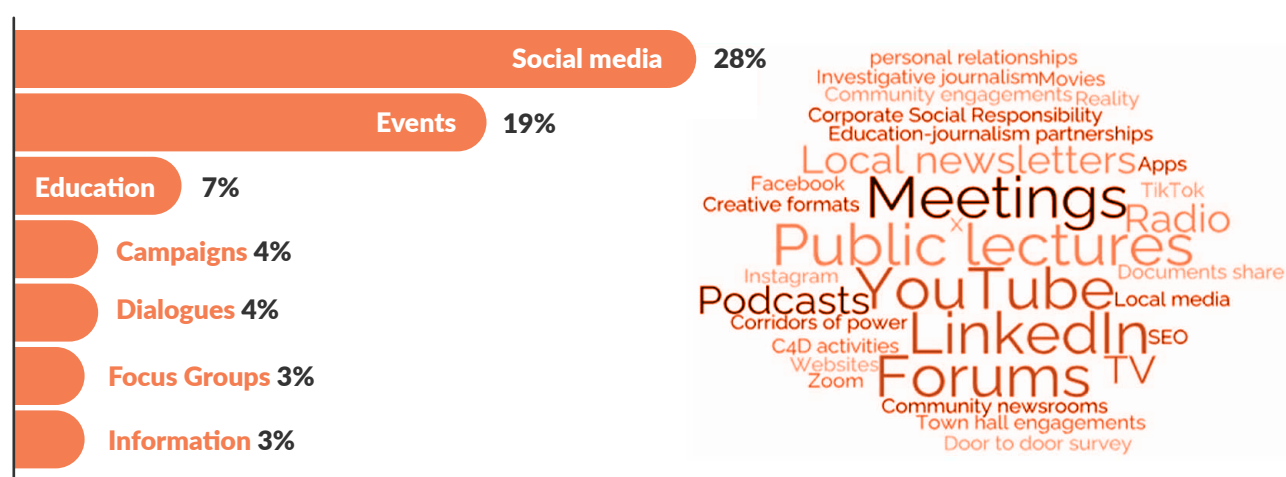
In the survey, two out of three journalists (66%, see Graphic 39 below) pointed to young people as one of the priority audiences for nuanced stories in their region, and that number reaches 82% in Africa and 72% in Europe. Additionally, 20% of respondents mentioned K12 students should also be prioritized.<sup>15</sup> These results reinforce the idea that, regardless of how advanced the technology availability and regulatory landscape are, there is a widespread concern in preparing the next generations to navigate and shape the future of AI in society.

# Engaging audiences

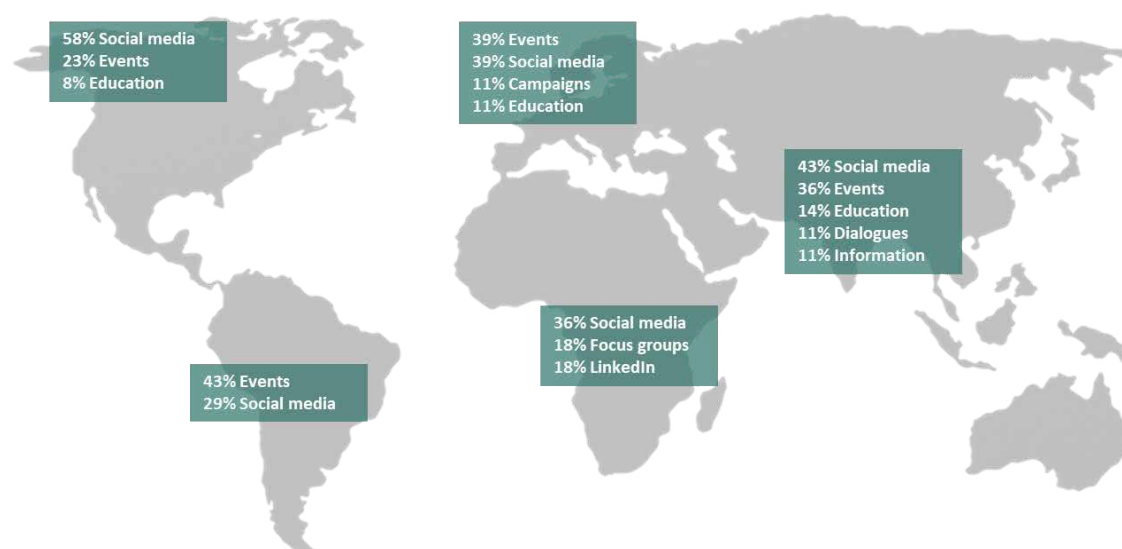
It is a shared view that audience engagement is critical to influence policymakers and the tech industry into adopting a more transparent and accountable conduct. The comprehensive and critical coverage of AI-related issues in media stories is indisputably important but cannot stand alone in the efforts for a better-informed relationship between society and technology. Experts suggest that highlighting immediate and tangible impacts, such as labor exploitation and energy consumption, can make AI issues more relatable and urgent. Public awareness campaigns, storytelling, and real-life examples can demystify AI and make it more accessible to the general public.

When considering the best way to reach audiences, two options stand out in the journalists' survey: social media (28%) and events (19%). Other less frequently mentioned forms of public contact include education, campaigns, dialogues, and focus groups. In Africa, communication options through discussion groups focusing on the issues and LinkedIn stand out.

**Graphic 36** How do you best reach the necessary audiences beyond media publication, % of cases (N=90). WordCloud presents answers mentioned by 1 to 3 journalists.



**Graphic 37** How do you best reach the necessary audiences beyond media publication per continent, % of cases (Africa N=11, Asia N=28, Europe N=18, North America N=26, South America N=7).



Given the relevance attributed to social media as a way of connecting with audiences, a few points from the digital assessment study stand out as potential guidelines for stories to generate engagement:

#### ● **Beginner-friendly reviews**

Public interest in cutting-edge AI developments, how they work, and their necessity is evidenced by the success of beginner-friendly content in the digital environment. Impact-focused stories could benefit from this format to introduce issues at a slower pace, as opposed to lengthy investigative pieces that retain attention from a smaller audience.

#### ● **Government response cases**

Government responses to AI regulations have a high evergreen score, particularly when shared by government officials and business leaders. The prolonged attention to these stories could be leveraged to monitor promises and consequences, ensuring consistency and following up on stories that require resolution.

#### ● **Ethical AI practices**

Considering the high search for regulations and ethics in the keyword analysis, there might be a demand for understanding how society can be protected from potential harms. Bringing stories that resonate with audiences' daily experiences could strengthen the connection between the technology and its impacts.

#### ● **Transparency concerns**

Growing concerns about AI transparency are leading to increased negative sentiment, suggesting a need for clearer information on AI systems' operations and decisions. Practical examples demonstrating how much the public is not aware of (see box below) could help retain interest.

#### ● **Regulation stories**

Stories about AI regulation maintain relevance longer than those about products and services, suggesting a focus on policy implications and comprehensive engagement strategies. Didactic explanations about the reasons and consequences of regulation in accessible language could help counterbalance industry claims of governance structures hindering technological progress.

#### ● **Litigation and accountability**

Litigation has been effective in raising the profile of AI accountability issues, particularly in the context of labor and AI supply chains. The worldwide examples of exploitation cases against big tech receive significant attention from mainstream media, but the connection of cases with local realities is still underused as a way of bringing attention to supply chain issues and engaging audiences in deeper discussions.



## The Privacy Project, by The New York Times

In 2019, the Times launched a series on privacy issues and the trade-offs being made to allow technology to enhance human potential. The series contained reporting discussing ideas, basic knowledge, public debate, and actions that could be taken for more conscious engagement with technology and fostering adequate regulation. One of these articles, "We Built an 'Unbelievable' (but Legal) Facial Recognition Machine" by Sahil Chinoy, discusses the implications of facial recognition technology. The author describes an experiment where public images of people near Bryant Park were collected and analyzed using Amazon's facial recognition service. The system identified 2,750 faces in a nine-hour period, including a match with Richard Madonna, a professor, with 89% similarity. This experiment highlights the ease with which individuals can be tracked without their knowledge, raising privacy concerns. Calling out how the technology's rapid advancement had outpaced legal regulations, the author debates stricter controls or even a ban on government use. The article explored the broader implications of facial recognition and noted how easy it was to build a photo database with public data and identify the professor, who was on his way to lunch when the system recognized his face from a partial profile image obtained by the cameras. This type of experiment can provide very tangible examples of the risks technology poses to civil liberties, prompting debates on the need for regulation or prohibition to protect privacy and free speech.

[See the full article here](#)



The frequent searches for examples of applications and how they can be used in one's work or lifestyle identified in the keyword analysis show an urge to discover the benefits of AI by the public. Exploring text that can lead to such search results, complemented with a more holistic view of all the ways in which the technology impacts the user and all those integrated into its ecosystem, could be a shortcut to bring attention to important issues. Additionally, the digital assessment shows that interviews, press releases, and why-posts (articles that begin with 'why' to drive curiosity or tease new information) about the topics are the ones more frequently shared, with a length of 2,000-3,000 words.

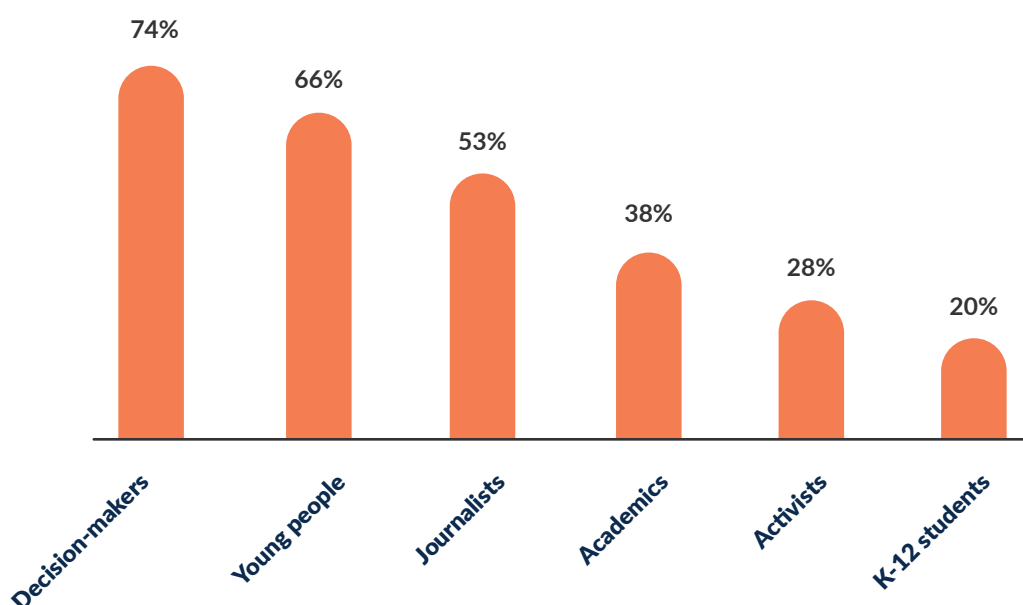
Journalists, academics, and activists are opinion leaders on accountability stories online, and the most shared pieces tend to come originally from traditional media outlets (e.g., AP News, The Guardian, NY Post), as well as from specialized outlets (Wired) and content produced on YouTube. Such information indicates there could be specific groups following these stories, but also that they may not be reaching all the relevant stakeholders to bring the topics onto the public agenda more effectively. Hence, finding ways to measure engagement of different target groups becomes essential.

Public opinion studies on knowledge and attitude towards AI are limited,<sup>16</sup> especially outside the Global North economies, which restricts the understanding of information gaps and the most effective ways to engage the general public. According to experts, this is essential to put governments and policymakers in check, both in terms of utilizing AI services and for regulating the industry. For instance, contracts of new AI technologies by governments seldom include a public consultation or needs assessment studies explaining the necessity of an AI tool, and even rarer are news following up on the application's performance and effectiveness. The same happens for private companies. Citizens and consumers, however, are provided with the good news of a new state-of-the-art tool that is supposed to make their lives easier or safer, and with scarce updates on how their extra expenses and forfeiting of privacy are being compensated. Therefore, learning more about public perceptions and expectations about AI will help develop communication strategies that will resonate with them.

## Who needs to be engaged (and how?)

For 74% of the surveyed journalists, decision-makers are the priority group of stakeholders to receive nuanced stories about the impacts of AI in their region. The experts tend to share that view, reinforcing that having the power to enact regulations and policies, combined with the perception of an insufficient understanding of AI's capacities and impacts, results in a sense of urgency for them to be approached. Among their suggested formats are detailed analyses, case studies, and evidence-based reports that highlight the regulatory implications and the consequences of inaction.

**Graphic 38** Priority audiences for nuanced stories about the impacts of AI technologies in your region, according to journalists, % of cases (N=90).



Other answers mentioned spontaneously: General public (3%), consumers (1%), groups between 40-60 (1%), investors (1%), people who are affected by AI decision-makers (1%), vulnerable community workers being exploited (1%), working class (1%).

For the specialists, the communities already harmed by AI technologies are often prioritized with the idea of empowering them for their own protection. Human-interest stories that highlight real-world implications on individuals and groups can make the abstract challenges of AI more relatable. Allowing these people to voice their own experiences and perspectives on AI is also paramount.

Another unanimous priority among the experts is the average user, consumer, or reader. They note that the public needs more tangible and complete information about how AI affects their daily lives. There is a belief that demands grounded in popular outcry find faster resolution among those able to make those decisions. Relatable stories that highlight personal impacts, such as effects on jobs or privacy, and explanatory journalism that demystifies AI can create familiarity with the issues. Simplifying complex concepts and providing practical examples are also useful tactics.

Engaging advocacy groups and civil society is more related to building partnerships than necessarily raising awareness. Specialists observe that these groups can help disseminate information and advocate for change. Stories that uncover ethical and social implications of AI, such as bias and discrimination, can mobilize advocacy efforts. Collaborating with civil society associations can expand the reach of the topic and the problems it brings.

For at least half of the journalists consulted in the survey (53%), it should be a priority to engage their own peers. Journalists shape public understanding and discourse around AI technologies, and initiatives such as the AI Accountability Network are considered commendable in capitalizing on that power to increase awareness, highlight ethical and social implications, and hold companies and policymakers accountable.

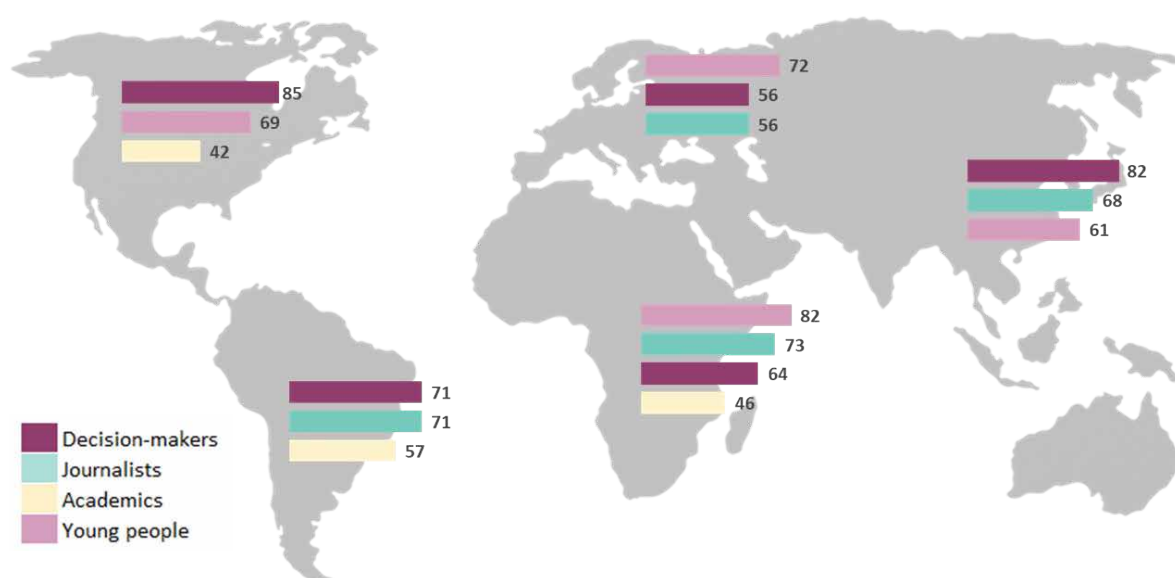
The success of litigation stories in shedding light on AI accountability in the online dialogue, according to the social listening exercise, indicates that engaging the legal community could raise opportunities to share stories and increase awareness about the most pressing issues. Potential benefits could be journalistic investigative pieces helping to generate evidence of AI harm when lawsuits, which are not considered an ideal redressing channel, are inevitable.

Within the same concept of establishing allyships with groups that can lend their expertise to advance ethical AI, according to the experts, engaging academics from varying fields is essential. Their contribution with valuable research, critical analyses, and evidence-based insights can inform policy, guide ethical considerations, and advance the agenda of explainable and accountable AI. Their involvement ensures that discussions are grounded in rigorous scholarship and can help bridge the gap between theoretical knowledge and practical applications. Concurrently, nearly 2 out of 5 journalists mentioned academics as a priority group for nuanced stories about AI technologies (38%), and these stakeholders are among the top three in South America (57%), Africa (46%), and North America (42%).

A few experts believe that small and medium-sized enterprises (SMEs) can often be at the forefront of innovation, yet mostly at a disadvantage when it comes to AI regulation. Besides their direct role in ensuring ethical development and deployment of AI, engaging SMEs could be an opportunity to foster a more inclusive and diverse AI ecosystem, driving innovation that benefits a broader range of communities.

Finally, young people and children are also considered priorities in terms of education and engagement with the quest for accountable AI by journalists in the survey (66% and 20%, respectively). Moreover, youth were highlighted as the top priority in Africa (82%) and Europe (72%). The interviewed experts share these views, and consider young audiences' engagement key for preparing the future workforce, fostering awareness and critical thinking, instilling ethical considerations, and promoting inclusivity. Early education on AI helps them understand its benefits and risks, enabling them to contribute responsibly and innovatively to the field while ensuring diverse perspectives are included in AI development and usage.

**Graphic 39** Priority audiences for nuanced stories about the impacts of AI technologies in your region per continent, % of cases, options with a frequency of 40% or more (Africa N=11, Asia N=28, Europe N=18, North America N=26, South America N=7).



Given the multitude of communities and topics that require attention and engagement initiatives, the experts recognize the work to be done is vast and needs collaboration from all spheres of society. Nonetheless, during the interviews a few suggestions of engagement initiatives were provided, and are summarized as follows, with an indication of the target public that could benefit the most from it.

**Graphic 40** Initiatives suggested for stakeholder engagement.

Decision-makers & policy makers	Communities impacted by AI	Young people & children	General public	Advocacy groups & civil society	Journalists	Academics	Tech SMEs	Legal community
<p><b>Detailed reports and case studies:</b> Provide in-depth, evidence-based reports focusing on the trade-offs and impacts of AI technologies, including case studies and regulatory frameworks to illustrate the need for specific policies.</p>	<p><b>Workshops and training sessions:</b> Conduct workshops and training sessions to educate stakeholders on AI's technical aspects, risks, and benefits, bridging the knowledge gap for informed decision-making.</p>	<p><b>Collaborations with civil society and advocacy groups:</b> Partner with local and global civil society organizations and advocacy groups to amplify messages and provide broader perspectives on the societal impacts of AI.</p>						
<p><b>Direct presentations and briefings:</b> Organize interactive presentations and briefings for stakeholders to discuss the current state of AI, its challenges, and potential regulatory measures.</p>	<p><b>AI competitions and hackathons:</b> Host AI-themed competitions and hackathons for young people to inspire creativity and innovation while providing practical experience with AI technologies.</p>	<p><b>Regular updates and follow-ups:</b> Keep stakeholders informed with regular updates on AI developments and follow up on previous discussions to ensure continuity and sustained interest.</p>						
<p><b>Networking events and resource hubs:</b> Host networking events for small and medium tech business leaders and create online resource hubs with accessible information on AI regulations and ethical guidelines.</p>	<p><b>Multimedia content and interactive platforms:</b> Utilize various forms of media, including videos, infographics, podcasts, and social media, and create interactive platforms for public engagement with AI experts and journalists.</p>	<p><b>Community-led discussions:</b> Facilitate discussions led by community members who are directly impacted by AI technologies. Ensure that the conversation is grounded in lived experiences.</p>						
<p><b>Incentives and recognition:</b> Offer incentives such as grants, awards, or recognition programs for SMEs, journalists, school and community projects focused on responsible AI practices and education.</p>	<p><b>Investigative journalism and collaborative campaigns:</b> Focus on stories uncovering ethical and social implications of AI and partner with civil society groups to create awareness campaigns.</p>	<p><b>Educational campaigns and school partnerships:</b> Launch educational campaigns focusing on AI's real-world impacts and collaborate with schools to integrate AI education into the curriculum.</p>						
<p><b>Mentorship programs and AI clubs:</b> Establish mentorship programs and encourage the formation of AI clubs and extracurricular activities in schools and communities.</p>	<p><b>Feedback mechanisms:</b> Establish channels for community members to provide feedback and share their concerns about AI technologies.</p>	<p><b>Highlight landmark cases and legal forums:</b> Share detailed analyses of landmark AI-related litigation cases and participate in online legal forums to engage the legal community.</p>						
<p><b>Advisory roles:</b> Push for the involvement of diverse stakeholders in advisory roles for policy-making bodies, industry panels, and ethical review boards.</p>	<p><b>Interactive workshops and classes:</b> Organize hands-on workshops and classes that teach AI concepts through interactive activities, games, and projects.</p>	<p><b>Localized content:</b> Create educational content that is specific to the community's context, such as local languages, cultural references, and relevant examples.</p>						
<p><b>Research collaborations and conferences:</b> Foster partnerships between academic institutions and industry or government bodies for joint research on AI topics and organize conferences for interdisciplinary discussions.</p>	<p><b>Funding and publication opportunities:</b> Provide funding opportunities and grants specifically for research on AI impacts. Encourage and support publication in reputable journals and platforms dedicated to AI research.</p>	<p><b>Relatable stories and simplified explanations:</b> Use human-interest stories, accessible explanations, real-world implications, and advocacy success stories to make AI's abstract challenges more relatable and understandable.</p>						

# The big, powerful, and influencing elephant in the room

It is impossible to develop a discussion about AI impact without assessing the level of power and influence big tech companies have on the matter, and how difficult it is to ensure their responsible conduct when developing and deploying AI systems. There is palpable pessimism about the industry's lobby in government decisions, and strong skepticism about self-regulation attempts. These companies are, however, part of the fabric of society and are not completely immune to the consequences of their negligence or lack of awareness.

---

*"We don't have as much power, influence, or resources as big tech corporations. (...) We are not as organized as the other side."*

---

A few experts suggest that the antagonizing framework, although necessary and mostly fair, can only do so much to advance accountability discussions. The need for regulation is unquestionable, but there is a prevailing perception that society can benefit from the responsible advancement of AI. These specialists will consider progress in the public discourse when the tone of the discussion moves from "us versus them" to what trade-offs are being made for innovation and who needs to be involved in making these decisions. They believe that raising awareness and demystifying the AI hype is seen as a necessary balance, but effective change may only be possible when the large tech companies are successfully engaged in transparency and accountability.

---

*"We are constantly okay with sharing health data, where it actually results in benefits for the individual or benefits for society. We do large-scale healthcare studies with people's data and with electronic medical records because it results in a tangible benefit for people, even though that represents a gross violation of their privacy. We put in a bunch of safeguards to help limit any potential downsides from the violation of that privacy and from the fact that we're choosing one equity over another. But we go ahead and do it, because we think there's real social value in doing it."*

---

The interviews with the specialists yielded conversations about how to create a structured and organized approach to gather, prioritize, and present the needs and concerns of various groups to shape informed and balanced AI policies. It became clear that, besides providing tools for stakeholders to understand the issues, the movement to operationalize the required changes is still restricted to litigation or a few successful examples, such as the elaboration of the AI Act or the strong actions taken to force law compliance of X in Brazil. The following graphic contains more direct suggestions on how to reach big tech beyond those types of measures.



**Graphic 41** Suggestions for industry engagement, according to the interviewed experts.



---

*“— Is there any leverage that is yet to be explored for accountability?”*

*— Yes. Deciding what good actually is. Complaining is not actually generating a position of power, but deciding what good actually is helps to shift things.”*

---

# Measuring impact

The indicators of success for initiatives such as the AI Accountability Network are expected to be both of a subjective and objective nature. An increased number of nuanced AI-related stories, changes in policy, and public outcry are the most frequently mentioned in the interviews. Additional ways to monitor the progress of these debates are:

## • Audience reach and engagement

Using social listening and keyword analyses to track AI-related searches online, and analyze metrics such as the number of views, shares, comments, and likes on AI-related articles on a continuous basis.

## • Policy influence

Monitor any changes in AI-related policies or regulations that can be attributed to media coverage. This can include citations of articles in policy documents or statements from policymakers.

## • Public awareness and understanding

Conduct surveys or polls to assess changes in public familiarity and attitude towards AI issues before and after improved coverage. Look for increased knowledge about AI technologies and their societal impacts.

## • Journalistic practices

Evaluate changes in journalistic practices, such as volume analysis using tools like Media Cloud, the diversity of sources cited in AI stories, and the depth of investigative reporting.

## • Community impact

Assess the impact of AI stories on local communities, especially those in the Global South. This can include follow-up stories on community responses or changes in local policies and practices. Social listening assessments targeted on these areas or specific languages could complement the analysis.

## • Expert and stakeholder feedback

Gather feedback from AI experts, journalists, and other stakeholders on the quality and impact of AI coverage. Explore a variety of groups and fields of work and gather examples of changes observed in each area.

## • Training and capacity building

Measure the effectiveness of training programs for journalists on AI reporting. This can include pre- and post-training assessments to gauge improvements in knowledge and skills.

---

*“Well, we have research on AI journalistic coverage being conducted, and many point to a predominance of optimism. I think transforming this to bring a critical perspective, having indicators that there is indeed a cultural shift in how you approach technologies. Perhaps a success indicator is a step further: How these news pieces start to influence public discussions on the topic, both regulatory and judicial. (...) I think there is great potential once you can establish networks, strengthen networks, and ties with organizations that are monitoring this, with networks of human rights lawyers or activists.”*

---



Image by Selvaprasanth Lakshmanan, India, 2022

# Conclusions and recommendations

The integrated analysis from all the data points from this study underscores the pervasive influence of AI in modern society and the substantial gap in balanced and critical reporting surrounding AI technologies. The tendency towards either too optimistic or alarmist narratives is exacerbated by industry PR, the assumed complexity of the technology, and shortfalls in media professionals' specialized training. There is a clear need for AI governance and regulation, and for engaging diverse stakeholders in a more informed and productive public discussion about the impacts of technology and how to move forward with the most equitable distribution of benefits possible. The graphic below presents the main conclusions of this study, from a viewpoint of the needs and actions expected from each stakeholder on the way to responsible AI development and deployment.

**Graphic 42** Stakeholders' main needs and how they can contribute to responsible AI.

NEEDS	ACTIONS
<b>Connections:</b> Access to diverse specialists, community representatives, policymakers, tech companies to have a holistic view of the issues.	<b>Advance:</b> To lead the movement for responsible AI forward, either by research and education, news dissemination, regulation development, transparency, or accountability.
<b>Coordination:</b> Organizing and creating internal structures to make advancements in regulation and accountability possible.	<b>Be held accountable:</b> To be aware and take responsibility for the impact of actions or negligence, and to provide appropriate channels for grievances and redressing.
<b>Education:</b> Learning about capabilities, limitations, and impact of AI in useful and relatable ways.	<b>Be transparent:</b> To be open about the supply chain and processes behind a product. To clearly state limitations, and to allow for external auditing and quality control.
<b>Grievances channel:</b> a space for safely sharing complaints and problems and the certainty that will be addressed.	<b>Connect:</b> To link voices for improved media coverage, education, regulation, and resolution of issues.
<b>Information:</b> Knowing what is happening before, during, and after the deployment of AI systems with transparency.	<b>Demand:</b> To publicly demand for better information, clear governance structures, and public protection against current and potential harm.
<b>Integration:</b> A seat at the table when decisions about regulation and redressing mechanisms are being discussed and decided.	<b>Disseminate:</b> To help spread verified and uncomplicated information about the technology, potential problems and solutions, and impact cases.
<b>Language:</b> Establishing ways of communicating complex aspects of tech that will ease the process of understanding and regulating it.	<b>Legitimate:</b> To validate issues raised by lending a specialist or firsthand perspective and contributing to the credibility of the claims.
<b>Multidisciplinary approach:</b> Studies and reflections about AI from perspectives beyond technology engineering and economics.	<b>Measure:</b> To develop and apply metrics of impact, public awareness and attitude, available governance, and compliance.
<b>Pressure to act:</b> External motivation from other players demanding proper regulation and accountability.	<b>Share:</b> To share examples of real-life impact of AI in diverse communities. To exchange experiences seeking ideas for resolution.
<b>Research:</b> Evidence to understand the societal impacts, ethical considerations, and technical aspects of AI.	
<b>Specific demands:</b> Clearer regulatory attributes, operationalized requests and specific demands for more expedient resolutions.	
<b>Support:</b> Legal advice, media space, funding, recognition, and any type of support that will facilitate claiming for responsible AI.	



### Legal community

**They need:** Information, integration, research.  
**They can:** Advance, demand, and legitimate.



### Youth

**They need:** Education, information, and integration.  
**They can:** Demand, disseminate, and share.



### Tech SMEs

**They need:** Integration, language, specific demands.  
**They can:** Advance, be held accountable, be transparent.



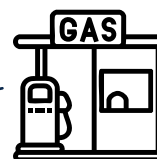
### Academia

**They need:** Information, multidisciplinary approach, research.  
**They can:** Advance, legitimate, measure.



### General population

**They need:** Education, grievances channel, and information.  
**They can:** Demand and disseminate.



### Big tech

**They need:** Coordination, language, multidisciplinary approach, pressure to act, research, specific demands.  
**They can:** Advance, be held accountable, be transparent, legitimate, and measure.

### Policymakers

**They need:** Coordination, education, language, pressure to act, research, and specific demands.  
**They can:** Advance, be held accountable, and be transparent.

### Impacted communities

**They need:** Grievances channel, information, integration, and support.  
**They can:** Demand, legitimate, and share.

### Advocacy & civil society groups

**They need:** Information, research, and support.  
**They can:** Advance, connect, disseminate.

### Media professionals

**They need:** Connections, education, information, multidisciplinary approach, research, and support.  
**They can:** Advance, connect, demand, disseminate, and share.

## Recommendations

Demystifying the AI hype and shaping conversations into a more present-centered assessment of products' capabilities, companies' conduct, and societal impacts will require helping audiences realize the pervasiveness of AI in their lives today, instead of the high tech, robotic, futuristic, or dystopian universe it is usually associated with. Articles focusing on how much of the technology is embedded in consumers' routine with examples of uses could be engaging. Helping users assess their interactions with digital platforms, for example, focusing on how much they understand what is going to happen with their data, how clear and accessible are these explanations, could not only motivate a more conscientious behavior, but also provide insights on how to improve explainability.

Contrasting pieces may also gain traction given the appeal such stories usually have with a wide range of audiences. They could help link, for instance, the easy and inconsequential use of an AI image generator for a cell phone sticker to each stage of the supply chain such a process incurs. A sensitive approach is recommended, however, since shaming users or overcomplicating stories may just alienate the audience. Helping the public slowly make the connections is key to provoking an active change in attitudes, and aggressive activism is likely to decrease receptivity.

Dichotomizing AI safety or usefulness for the human race is an approach that will probably maintain the *status quo* in terms of regulation and industry accountability. It is extremely necessary to shed light on the harm caused to many communities and on the dubious conduct within the supply chain. However, it is important to focus the narrative on the companies and decision-makers causing harm, instead of the technology. Ultimately, it is not AI that is making problematic decisions, but the corporations and people behind it. As much as society needs to rethink its tendency for techno-solutionism, techno-pessimism can stall the debate, collaboration, and innovation needed for advancement.

Approaching the necessary changes in governance and policy is challenging due to the difficulty in grasping all parts involved in the development and deployment of the technology, but also because both policymakers and the general public are AI users in one way or another. Initiatives that could help users define and identify “safe(r) AI” could be a way to establish examples of companies and products recommended to interact with,<sup>17</sup> and to develop a keener eye for identifying questionable practices. In addition, it could incentivize companies to review their conduct to avoid being tagged as a bad example.

In many countries, governments and gig economy workers have been in rooms discussing license to work, labor taxes, and social security. Yet, it is unlikely that most of these discussions included the way in which the algorithms work. Workers have been known to organize themselves to fight against the unjust working conditions in the platforms, but it is harder to find evidence of policymakers and gig workers learning from each other on the effects of the nearly completely unregulated realm of gig work algorithms. Making these types of connections is more of a job for advocacy or civil society groups, but media organizations such as the Pulitzer Center have the social capital to facilitate these encounters, either through journalism or engagement initiatives.

Finally, journalists seem to have been placed in a passive, misinformed position in terms of their responsibility for the current shortfalls of AI coverage. However, it is widely known that newsrooms face numerous challenges in terms of resources, and dedicating time to grasp the complexity of a phenomenon that is constantly changing is an acknowledged challenge. Moreover, AI brings significant changes into the profession itself, an issue that was not discussed in this study but that is likely to influence their attitudes towards the technology. Initiatives such as the AI Accountability Network and having peer references such as the ones listed on page 29 could go a long way in shifting media rooms’ practices for AI coverage. There is also a lot to be learned from smaller outlets and local media, which are perceived as measuring up to the task of providing valuable reporting.

Nevertheless, to keep providing journalists with training and confidence to go after intricate and obscure, but necessary, stories might require deconstructing issues into smaller, but integrated pieces and providing space for continuous updates. That implies, for example, not expecting a story on all the ways in which a community has been victimized by unfair police raids due to suspicious AI surveillance, but starting with parts of the problem. Perhaps investigating what is the internet access level in that community and how the datasets representing it look like. Next, a story about the purchase of such surveillance systems and how data was trained into it, followed by law enforcement training to use it, and so on. There is a sense that journalists are as overwhelmed by the hype as they are responsible for it, and they need the space and support to de-escalate, organize, and recreate the boundaries and priorities that are intrinsic to their profession.



# Appendices

## Literature Review

### AI AND NEWSROOM

- Becker, K., Simon, F., & Crum, C. (2023). Policies in parallel? A comparative study of journalistic AI policies in 52 global news organisations. Retrieved from <https://doi.org/10.31235/osf.io/c4af9>
- Dandurand, G. et al (2022). Training the news: Coverage of Canada's AI hype cycle (2012–2021). Shaping 21st-Century AI, Institut national de la recherche scientifique. Retrieved from [https://espace.inrs.ca/id/eprint/13149/1/report\\_ShapingAI\\_verJ.pdf](https://espace.inrs.ca/id/eprint/13149/1/report_ShapingAI_verJ.pdf)
- de-Lima-Santos, M.-F., Yeung, W., & Dodds, T. (2024). Guiding the way: A comprehensive examination of AI guidelines in global media. Retrieved from <https://doi.org/10.48550/arXiv.2405.04706>
- Dodds, T. (2021). Algorithmic-based technologies' impact on journalistic identities. Leiden Anthropology Blog. Retrieved from <https://www.leidenanthropologyblog.nl/articles/algorithmic-based-technologies-impact-on-journalistic-identities>
- Kahn, G. (2024). AI newsroom guidelines look very similar, says researcher who studied them. He thinks that's bad news. Reuters Institute. Retrieved from <https://reutersinstitute.politics.ox.ac.uk/news/ai-newsroom-guidelines-look-very-similar-says-researcher-who-studied-them-he-thinks-bad-news>
- Posetti, J., Simon, F., Shabbir, N. (2019). What if scale breaks community? Rebooting audience engagement when journalism is under pressure. Reuters Institute. Retrieved from <https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-10/Posetti%20What%20if%20FINAL.pdf>

### AI AND POLICY (including international governing documents)

- NPR. (2024, May 24). The impact of AI [Audio podcast]. Retrieved from <https://www.npr.org/2024/05/24/1196981155/the-impact-of-ai>
- Smuha, N. (2021). Beyond the individual: Governing AI's societal harm. Internet Policy Review. Retrieved from <https://policyreview.info/articles/analysis/beyond-individual-governing-ais-societal-harm>
- UNESCO. (2021). Recommendations on the ethics of artificial intelligence. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000381137.locale=en>
- United Nations. (2023). UN High-Level Advisory Body Interim Report. Retrieved from <https://www.un.org/ai-advisory-body>

### DIGITAL TECHNOLOGY & STRATEGIC LITIGATION

- Mutemi, M. (2018). Case asking the court to declare access to the internet a basic right. Retrieved from <https://www.mercymutemi.com/post/case-asking-the-court-to-declare-access-to-the-internet-a-basic-right>
- Mutemi, M. (2022). Case challenging the Facebook algorithm and content moderation. Retrieved from <https://www.mercymutemi.com/post/case-challenging-the-facebook-algorithm-and-content-moderation>
- Mutemi, M. (2022). Case challenging the unfair working conditions Facebook content moderators face. Retrieved from <https://www.mercymutemi.com/post/case-challenging-the-unfair-working-conditions-facebook-content-moderators-face>

## TOPIC-BASED PULITZER CENTER WORKS

- Bansal, V. (2022). India's gig workers and algorithms. Pulitzer Center. Retrieved from <https://pulitzercenter.org/projects/indias-gig-workers-and-algorithms>
- Burke, G., Dearen, J. (2022). Tech tool offers police 'mass surveillance on a budget'. Pulitzer Center. Retrieved from <https://pulitzercenter.org/stories/tech-tool-offers-police-mass-surveillance-budget>
- Burke, G. et al (2022). Police seize on COVID-19 tech to expand global surveillance. Pulitzer Center. Retrieved from <https://pulitzercenter.org/stories/police-seize-covid-19-tech-expand-global-surveillance>
- Emmanouilidou, L. (2022). Eye on the wall: Refugees and 'smart' borders. Pulitzer Center. Retrieved from <https://pulitzercenter.org/projects/eye-wall-refugees-and-smart-borders>
- Geiger, G. (2023). Unlocking Europe's welfare fraud algorithms. Pulitzer Center. Retrieved from <https://pulitzercenter.org/projects/unlocking-europes-welfare-fraud-algorithms>
- Hao, K. (2021). AI colonialism. Pulitzer Center. Retrieved from <https://pulitzercenter.org/projects/ai-colonialism>
- Ho, S., Burke, G. (2022). An algorithm that screens for child neglect raises concerns. Retrieved from <https://pulitzercenter.org/stories/algorithm-screens-child-neglect-raises-concerns>
- Jaswal, S. (2024). The journey of a propaganda message in Indian elections. Pulitzer Center. Retrieved from <https://pulitzercenter.org/projects/journey-propaganda-message-indian-elections>
- McIntyre, N. (2024). Spies in the loop: The workers building surveillance tech. Pulitzer Center. Retrieved from <https://pulitzercenter.org/projects/spies-loop-workers-building-surveillance-tech>
- Naundorf, K., Pabst, S. (2023). For the sake of security: The advance of facial recognition in South America. Pulitzer Center. Retrieved from <https://pulitzercenter.org/projects/sake-security-advance-facial-recognition-south-america>
- Pulitzer Center. (2021). Tracked and traced. Retrieved from <https://pulitzercenter.org/projects/tracked-and-traced>
- Sen, A. (2022). Peering into the black box. Pulitzer Center. Retrieved from <https://pulitzercenter.org/projects/peering-black-box>
- Shrivastava, K., Tapasya, Joshi, D. (2024). India's welfare machines. Pulitzer Center. Retrieved from <https://pulitzercenter.org/projects/indias-welfare-machines>
- Simpson, J. (2020). State of surveillance: Police use of AI and facial recognition in communities of color. Pulitzer Center. Retrieved from <https://pulitzercenter.org/projects/state-surveillance-police-use-ai-and-facial-recognition-communities-color>
- Thuy Vo, L. (2022). There goes the neighborhood. Pulitzer Center. Retrieved from <https://pulitzercenter.org/projects/there-goes-neighborhood>

# Endnotes

<sup>1</sup> The AI Accountability Network has supported projects with publications in many of the outlets mentioned by the study participants, and some of these stories can be found [here](#).

<sup>2</sup> Case studies on cheap labor behind data labeling and content moderation in [Brazil](#) and in [Russia](#).

<sup>3</sup> An example of [workers unionizing and attempting to change their conditions in Indonesia](#).

<sup>4</sup> An example of how bias from [who trains the data](#) can impact minorities.

<sup>5</sup> A case study on [data training](#).

<sup>6</sup> An example of [bias in the data trained](#) to alert the police about “suspicious” movements only in specific neighborhoods.

<sup>7</sup> A story on how innocent people identified based on a photo [ended up behind bars in Brazil](#).

<sup>8</sup> A story on how gig work is feeding [Russia's surveillance machine](#).

<sup>9</sup> Examples on [how data from users of popular websites and applications are used by companies to train AI models](#).

<sup>10</sup> Examples on the spread of [misinformation during India election](#).

<sup>11</sup> More examples on [how AI technology can affect elections through BJP's Whatsapp machine in India and AI-generated disinformation](#).

<sup>12</sup> See more about [hallucination in AI-powered transcription tools used in hospitals](#).

<sup>13</sup> An article about data workers pledging to [end modern day slavery](#).

<sup>14</sup> “The Case of the Creepy Algorithm That ‘Predicted’ Teen Pregnancy,” [Wired](#). “Inside Facebook’s African Sweatshop,” [Time](#).

“OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic,” [Time](#).

“Musk’s X Banned in Brazil,” [Lawfare](#). “Brazil Lifts Ban on Elon Musk’s X platform,” [Al Jazeera](#). “Volta do X: Entenda detalhes do bloqueio até a decisão de liberação da plataforma,” [CNN Brasil](#).

<sup>15</sup> Examples of how the Pulitzer Center’s is aiming at [prioritizing K-12 students and teachers](#).

<sup>16</sup> A few examples: [Pew Research Center](#), [Public First](#), [Gallup](#), [Artificial Intelligence Policy Institute](#), [UK’s Office for National Statistics](#), [European Center for Not-for-Profit Law](#), [Centre for the Governance of AI](#), [European Commission](#).

<sup>17</sup> An example of AI being used to [preserve Indigenous culture](#) and resist colonialism.