OPEN SOCIETY
FOUNDATIONS

AI in Journalism Futures

Initial Report

David Caswell and Shuwei Fang

August 2024

Table of Contents

224 West 57th Street
New York, NY 10019
P. +1 212-548-0600
opensocietyfoundations.org

Executive Summary

The AI in Journalism Futures (AIJF) project explored how artificial intelligence (AI) might fundamentally transform our information ecosystem over the next five to 15 years. Engaging nearly 1,000 global participants, including journalists, technologists, academics, and civil society advocates, the project's objective was to understand the range of possibilities for long-term impact of AI on journalism using a scenario planning approach. An open call for applications produced short scenarios from 880 participants, from which 40 were selected to participate in a scenario planning workshop held in Piedmont, Italy, in April 2024. This workshop facilitated the refinement of five robust and plausible future scenarios.

The five key scenarios developed were: "Machines in the Middle," in which AI-driven newsgathering and news production enables newsrooms to operate essentially without human journalists; "Power Flows to Those Who Know Your Needs," emphasizing the shift of power to those who best understand individual information needs in an AI-empowered information ecosystem; "Omniscience for Me, Noise for You," depicting a societal divide where some are super-empowered by AI while others may be overwhelmed by distracting or low value information; "AI with Its Own Agency and Power," envisioning AI systems operating in the information ecosystem with minimal human oversight; and "AI on a Leash," in which societal or regulatory constraints limit AI's transformative potential.

The workshop produced near-unanimous agreement that AI would fundamentally transform the information ecosystem, but also difficulty in articulating specifically how this transformation might occur, with participants often defaulting to extensions of the status quo. There was considerable skepticism about the ability of traditional journalism institutions to adapt successfully to an AI-driven future, and participants tended to envision coming changes in terms of power shifts away from journalists, with little attention to how those changes might increase or decrease value for audiences.

The results of the AIJF workshop underscore the urgency for stakeholders in journalism and civic information to engage in long-term planning and adaptation strategies to effectively navigate the coming AI-mediated information ecosystem. The scenarios and insights provided aim to guide these efforts, offering a foundation for understanding and thinking about journalism's AI future.

This executive summary was written by GPT-4 using the completed draft report, and lightly edited by the authors. This was the only use of AI in this report.

1. Introduction

The AI in Journalism Futures (AIJF) project is one of the first significant attempts to understand how artificial intelligence (AI) might fundamentally reshape our information ecosystem in the long term, over the next five to 15 years. In February 2024, the Open Society Foundations issued a call for applications for a convening in April in which selected participants would share their visions of an AI-mediated future. By gathering the perspectives of a large and diverse number of interested, engaged, and well-informed people from across the globe, AIJF seeks to broaden and deepen the conversation about AI's potential impact on journalism and civic information around the world. It also seeks to develop those perspectives beyond casual conversation and opinion by preparing participants for thoughtful discussion over several months, by applying a semi-formal scenario planning process to interactions among participants, and by structuring the outcomes of those interactions. This report therefore represents a rough consensus of the participants as of mid-2024, obtained in possibly the most thorough and inclusive way available, about the potential long-term impact of AI on our information ecosystem.

This document is the initial report on the AIJF project and its outcomes. It is an interpretation by the authors of the many submissions, discussions, and interactions that took place throughout the project and in particular at the workshop held in April in Piedmont, Italy. The authors have attempted to be as thorough and comprehensive as possible in this interpretation, and also to objectively represent the full breadth of contributions and opinions. Nonetheless it is an interpretation, and so is necessarily incomplete and unavoidably subjective to some degree.

This report will not be the only documentation to be produced about AIJF. It will be followed by further data, information, and insights published as academic papers, articles, and possibly as further reports. Much of the "raw data" obtained from the AIJF project will also be published in anonymized form, including 880 short initial scenarios and 45 developed scenarios.

The goal of the AIJF project is to jumpstart broader discussions and further studies that seek to understand how AI might reshape our information ecosystem. AIJF represents an early step toward this aim, and the authors and organizers plan to pursue subsequent work focusing on younger, more technically-oriented, and more business-focused participants. Furthermore, others with an interest in the

future of the information ecosystem are also developing initiatives to improve our understanding of the potential influence of AI.

While "journalism" is part of the project's name, in hindsight we believe that this word may be too restrictive and too tied to the traditional processes, products, and assumptions of legacy news publishing. We discuss this in more detail in section 5, and we have also adopted the term "journalism and civic information" throughout the report. We use this phrase to mean verified and contextualized information that provides value to audiences and societies.

This report is organized in a way that can be easily interpreted from its table of contents. Section 2 describes the reason for the project, the details of its execution, and briefly reviews earlier, similar work. Section 3 provides an overview of the project's "raw materials"—the global perspectives on AI that emerged from its progressive stages. Section 4 describes the five macro scenarios that emerged from the two-day scenario planning workshop, which serve as AIJF's primary product. Section 5 describes observations from across the full breadth of the project and extracts some insights from those observations. Section 6 provides a conclusion, drawn from direct feedback from a survey of participants, that identifies what the AIJF initiative has missed. It also includes a retrospective from the organizers and offers some closing remarks on how we might collectively continue the effort to better understand the coming AI-mediated information ecosystem and to assert agency within it.

2. Anticipating an AI-mediated information ecosystem

The advent of generative AI—which became widely known and available for popular use with the release of ChatGPT in November 2022—has demonstrated new capabilities with such profound and obvious potential impact that substantial change is already taking root in the information ecosystem.

AI is already being integrated into almost all parts of the existing ecosystem, including search platforms, social platforms, email clients, voice agents, and chat apps. We are starting to see fundamentally new features being rolled out across the ecosystem, including generative search, personal AI agents, AI-based media consumption apps, AI-based browsers, and even AI-enabled operating systems like Microsoft's Copilot+ PC and Apple Intelligence within iOS. The impact of these existing changes alone could also be profound, but the pace of AI investment, development and application suggests that further changes may be dramatic.

AI-driven innovation has now become an urgent focus of most newsrooms and content companies globally, and a cohort of entirely new information startups has emerged—fueled by expanding investment—that aim to further disrupt the information value chain. Meanwhile, rulemaking by regulatory bodies and deal-making between intellectual property owners and large language model (LLM) companies struggle to keep pace with AI's ever-growing capabilities.

Looking back, it took nearly a decade for social platforms (combined with smartphone innovation and adoption) to reshape the information ecosystem as we knew it. By comparison these early applications of AI, occurring over just 18 months, allows us to plausibly anticipate seismic change leading to an entirely new information ecosystem in which AI is ubiquitous across almost all facets of information gathering, production, distribution and consumption. AI promises to determine new fault lines and power shifts, new social norms and behaviors for audiences, and new modalities of information and communication production and consumption for which we may not yet have the vocabulary to describe.

2.1. Why AI might fundamentally change our information ecosystem

The use of AI in the digital information space is not new. Editorial workflows have been augmented by "traditional" AI in news media for at least a decade, notably with machine learning and big data analysis assisting human reporters in investigative journalism, and automated content generation driving mass sports, financial, and localized reporting. During this previous era of AI, however, use of these tools and techniques was generally limited to a select few: large, resource-rich organizations with high levels of technical staffing and capacity, mostly based in the Global North.

The expansion of generative AI for popular use has radically removed these old barriers to entry. Not only is the technology now much easier to apply with little technical expertise, it is also—at least for now—much cheaper relative to the prior cost of integrating AI. This has already led to faster and more radical forms of innovation and adoption in smaller newsrooms and startup media, especially in the majority world, enabled by their agility and entrepreneurial mindset.

Beyond the media sector, the new economics of AI innovation applies across the entire information ecosystem and affects essentially all information-producing industries – entertainment (movies, music, gaming), advertising, government, education, academic publishing, law, and so on. AI's capabilities are now accessible to almost any organization or individual with a small budget and the most basic technical understanding. The new barrier for those producing information is now arguably creativity and imagination, allowing more entrants to more thoroughly disrupt previously established practices across the information ecosystem. It is possible—and seems likely—that new forms, units and types of information could emerge across the entire ecosystem, produced by a broader class of people and organizations.

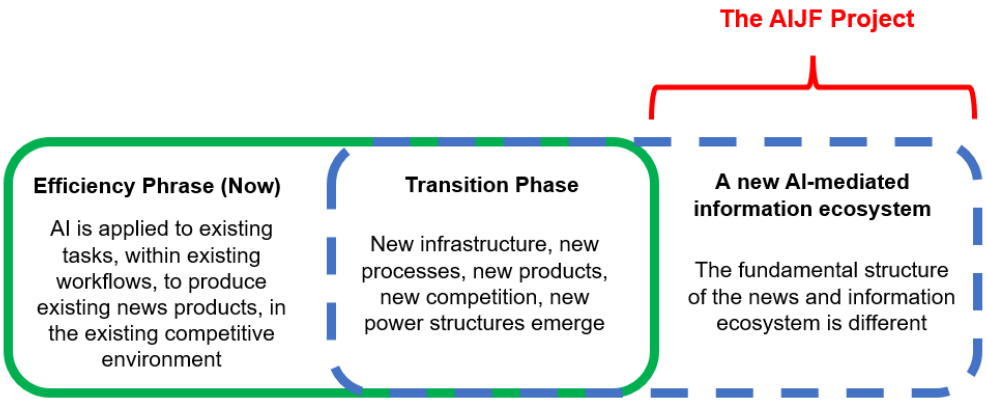
Information consumers are also increasingly empowered by the new capabilities rapidly appearing in interfaces, apps, platforms and operating systems, potentially giving consumers themselves far greater control of the information consumption experience. ChatGPT and generative search are both examples of conversational information retrieval mechanisms driven by the consumer, who enters prompts or search queries that create unique, ephemeral, experiences of information more akin to an oral exchange than the object-centered search paradigm we are already familiar with. This marks a fundamental change in our relationship with digital information.

2.2. An attempt to anticipate what this change might look like

How might AI transform our information ecosystem? A useful approach to framing the potential impact of AI on the information ecosystem involves identifying two overlapping phases of change:

1. A temporary **efficiency phase**, which began roughly in November 2022 with the release of ChatGPT. During this phase AI is applied primarily to existing tasks, workflows, and products in ways that essentially operate within the existing competitive environment. Producers and consumers generally produce and consume familiar information products.
2. A permanent steady-state **new information ecosystem**, wherein the structure of information ecosystem is fundamentally different, and in which information tasks, workflows, and products are relatively unfamiliar in relation to the current ecosystem. This new and different information ecosystem may produce new norms, paradigms, behaviors (among both audiences and producers), and perhaps a new political economy.

Between the current efficiency phase and the anticipated new information ecosystem we can also expect a temporary **transition phase,** during which entirely new processes, products, infrastructure, and new power structures start to emerge. Although we are uncertain about when a new AI-mediated information ecosystem might fully take over, it is likely that for some parts of the ecosystem the transition phase has already begun. These phases are shown below:



AIJF has worked to facilitate well-informed public discussion focused specifically on the emerging AI-mediated information ecosystem. It is explicitly *not* interested

in the current "efficiency phase," which is already the focus of considerable attention, investment, and development across the information ecosystem. Given the rapid pace of changes in AI functionality, and the extreme uncertainty about what these changes might add up to within the information ecosystem, AIJF has sought to produce a set of well-tested, plausible, and specific scenarios for how that ecosystem might potentially develop.

2.2.1. The AI in Journalism Futures (AIJF) project

AIJF was developed in part using a formal methodology known as "scenario planning": an approach to developing strategic understanding of potential futures under conditions of extreme uncertainty and complexity. Scenario planning differs from "forecasting" in that it does not seek to predict the future, but instead to understand the range of possible futures given the conditions of the present—in this case given the fundamental abilities of LLMs and generative AI.

The essence of the scenario planning process is to collectively access the expertise of a diverse range of contributors to develop a set of drivers or assumptions rooted in the present ("driving forces"), and then to use those to develop initial scenarios ("end states") that are further refined into a smaller set of robust, tested scenarios that describe plausible outcomes.

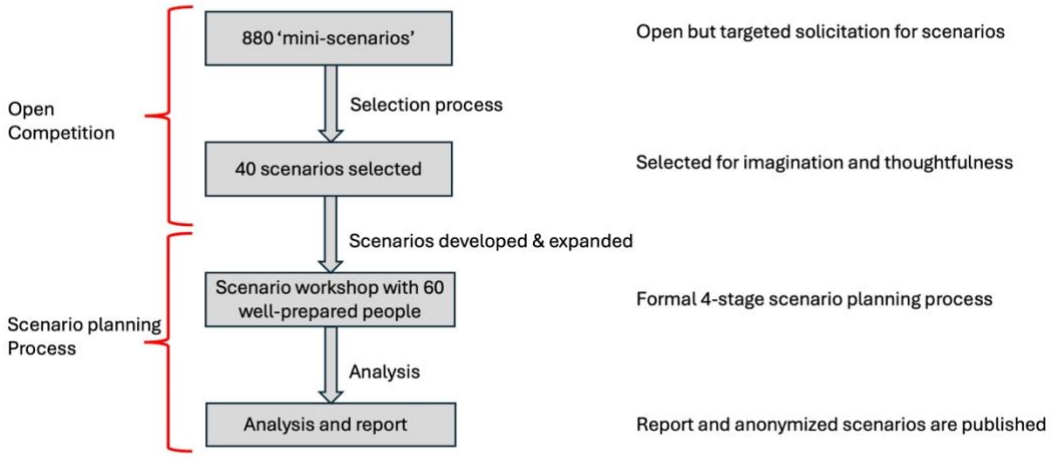
Key to the success of any scenario planning project is the selection of its participants. Many projects rely on participants who are hand-picked by their organizers. By contrast, AIJF took, to our knowledge, a unique approach by recruiting participants through an open competition for proposed scenarios. This led to a project design that combined two different methodologies: firstly the open competition, producing many proposed scenarios, and secondly the formal scenario planning exercise, centered on a workshop. More specifically, AIJF activities were designed and executed in three parts:

- **An Open Call.** A competitive open call for participants who each submitted a 300-word "mini-scenario" based on structured guidance. This call was open from January 26 until February 23, 2024, and produced 880 usable scenarios. The solicitation period was followed by a blind judging process to select 40 applicants from the submitted scenarios, as well as a curation process to identify a further 20 invited guests. These selections were made by an external judging panel: Reem Almasri (Febrayer), Gina Chua (*Semafor*), Nick Diakopoulos (Northwestern University), Rishad Patel (Splice Media), Alan Soon (Splice Media), Tshepo Tshabalala

(JournalismAI, LSE), Marina Walker Guevara (Pulitzer Center) and by the organizing team. This selection process was completed on March 4, 2024.

- **Development of mini scenarios to full scenarios.** We requested each selected participant to expand their mini scenario into a more detailed 3-5 page document, using structured guidance. This step was intended to help participants prepare for the workshop by deepening their thinking within the framework of scenario planning, with greater attention to driving forces and how they interact to shape scenario end states. This process was completed on April 5, 2024.
- **Scenario planning workshop.** A formal scenario planning workshop was held in Piedmont, Italy, on April 15 and 16, 2024. The workshop was facilitated by Robert Bood, a professional scenario planner who was instrumental in helping design the process, by the authors of this report, and by Open Society Foundations staff. The purpose of the workshop was to use a structured process centered on group discussions to produce a set of robust scenarios rooted in (but not exclusively based on) the driving forces identified by individual participants in their expanded scenarios. Prior to the workshop the organisers extracted and grouped the driving forces from the expanded scenarios as a basis for curating groups of participants. In designing the group interaction we aimed for a combination of expertise and friction between differing perspectives to facilitate vigorous discussion and the development of robust scenarios.

The structure of the application process and workshop is shown below:



2.2.2. Project objectives

In recent years, as generative AI's disruptive potential for journalism has become increasingly clear, newsrooms have responding primarily by exploring ways to make the production of their existing products more efficient, and sometimes by developing new products intended to compete within the existing media environment.

While appreciating that these approaches are appropriate in the short term, we concluded that they were probably insufficient in the long term because AI is likely to significantly restructure the entire information ecosystem within which journalism exists and operates. While most newsrooms and many academics focus on short-term impact and opportunities for efficiency, regulation, etc., there appeared to be almost no substantive explorations yet of how a long-term, permanent AI-mediated information ecosystem might develop. This lack of attention to the potential characteristics of a fully AI-mediated ecosystem has implications in the present because it reduces the ability of newsrooms and other stakeholders (including Open Society Foundations) to engage in long-term planning and investment aimed at influencing eventual outcomes.

The AI in Journalism Futures projects was therefore developed as an attempt to explore those potential long-term changes in the information ecosystem. Our primary objectives for the project were:

1. To provide a set of authoritative and plausible scenarios that could be used to help orient investment in AI-related activity by OSF and others investing in the emerging AI-mediated information ecosystem.
2. To facilitate and further an emerging conversation across the journalism ecosystem and its stakeholders about the potential of AI functionality to cause fundamental structural changes in the entire information eco-system.

We sought to do this by first developing an understanding of the "driving forces" that might influence the development of the new information ecosystem that might emerge from ubiquitously accessible AI, and then developing a small portfolio of well-described and reasonably possible scenarios (not forecasts) for what a stable "end state" in an AI-mediated information ecosystem might look like. This would include implications arising from each scenario and "weak signals" for each scenario—small developments that might provide early indications that a particular scenario is already underway, but which might be ignored without awareness of their significance to the scenario.

Our intention was that these scenarios could then be used for strategy development and contingency planning, including by monitoring any anticipated weak signals. We hoped to develop the scenarios over the wider information landscape, beyond just a narrow focus on the existing conception of journalism and media. To enable this we designed and marketed the AIJF project to attract participants from a globally diverse range of other sectors including technology and startups, academia, policy, think tanks, content creators and civil society, with varying degrees of success.

2.2.3. Scenario planning applied to journalism

Scenario planning (sometimes called scenario thinking or scenario analysis) is a semi-formal methodology for strategic planning that has been used by large organizations to develop flexible long-term plans under uncertain conditions. The technique has its roots in strategic defense planning at the RAND Corporation in Santa Monica during the 1950s, but was developed into a comprehensive applied technique in the strategy department of Royal Dutch Shell in the Netherlands in the 1970s. In the 1980s and 1990s it became a key strategic planning tool for large corporate and governmental organizations dealing with structural change. A particularly famous example from this era were the 1990 Mont Fleur Scenarios, which imagined what South Africa might look like in 2002 and which heavily influenced the planning for South Africa's transition into its post-apartheid era.

The objective of scenario analysis is not to predict the future. Instead the process seeks to understand the multiple plausible outcomes in a situation given an understanding of forces at work in the present. Traditional scenario planning follows a six-step process: first those forces that are driving change are identified and integrated into a coherent framework; scenarios are then developed from those drivers, first as a large number of lightly-formed scenarios and then as a smaller set of more fundamental scenarios with particular focus on the dynamics between forces and scenarios; finally, those scenarios are described and analyzed in detail and their consequences are explored.

AIJF is by no means the first attempt to apply scenario planning to journalism. As digital distribution and especially social media began to significantly disrupt journalism in the 2000s and 2010s, there were multiple attempts to apply scenario planning techniques to news. One of the earliest was the 2007 "Media Scenarios Project" in the United States, which focused on print-oriented scenarios in the then-emerging digital ecosystem. An early example in Europe was the EU's 2008 "Future of Creative Content" project, which focused on competition and public

attitudes to emerging social media and whose "Society meets industry" scenario approximates the outcome we are living with today. Canada's 2011 project, titled "2020 Media Futures Canada," looked at diffusion of innovation and sources of new value from media, identifying a similar scenario that they termed "Lord of the Clouds." A 2015 project by the Dutch Journalism Fund, titled "Scenarios for the Future of Journalism" looked at the effect of technology on consumer confidence in news media, and a 2020 Swedish project by the Future Media Group looked at a broad range of driving forces. More recently, a 2022 scenario planning project conducted by Deloitte in the Netherlands looked at the interaction of trust in news in relation to the role of intermediary platforms, and a major project in the UK titled "News Futures 2035" examined the interaction between public policy and the relevance of news to audiences. This project was one of the inspirations for AIJF.

These projects produced many insights useful to their various stakeholders, and undoubtedly contributed to subsequent actions by information ecosystem players. However, they also had characteristics that make them less useful for the present moment. Many of them took an extremely expansive view of journalism, media, and society, and how journalism intersects with wide-ranging political, social or economic conditions. Many of them, understandably, focused primarily on pressing issues of their time that may be less relevant today. These prior projects were also all completed before it became apparent that AI would become a dominant force in the future of journalism.

Our goal for AIJF was therefore to build on these earlier initiatives, while adapting the scenario planning technique to account for the particular uncertainty AI is creating for producers and consumers of journalism and civic information in our current day. We focused only on AI and its potential impact on the information ecosystem, and we designed and executed the workshop and report as quickly as possible in an attempt to track the pace of change in the technological drivers.

3. Perspectives on an AI-mediated information ecosystem

This report is distilled from the contributions of almost 1,000 people, most of whom are deeply engaged in journalism and civic information in one way or another, and many of whom have been closely following the fast-paced development of AI. That distillation included the design of the initiative and framing of the call for applications, the criteria for selecting the 40 ‘winners’ from the submitted mini-scenarios, the guidance for producing the developed scenarios, the structured process followed during the workshop, and the review and analysis of the outputs from that process.

We will examine the end results of that distillation in section 4, however it may be useful to first briefly review the “raw material” upon which those results are based. This section therefore provides a sense of the range of contributions at the three key stages of the project.

3.1. Perspectives from 880 people expressed as short scenarios

The AIJF call for applications, which allowed anyone to submit an entry, in any language, yielded 880 applications in total. The application form was deliberately simple. It focused mainly on the applicants’ mini-scenario submission and asked only for very basic supplementary information on affiliation and location, which were not mandatory fields. We were therefore only able to make some limited inference and generalized observations about participants’ backgrounds from the available data.

The pool of applicants came from approximately 70 countries, weighted heavily towards the majority world, i.e. from outside of the United States and Europe. Applicants from the journalism sector (including reporters, editors, publishers, product and data specialists), and adjacent fields (including those working in media innovation, media development, professional bodies and networks, and journalism students and faculty) were the most heavily represented, accounting for more than half of applicants. The remainder came from a mix of sectors with a stake in the information ecosystem: the technology and startup sector (particularly AI-related companies), content creators, civil society and non-governmental

organizations, government and multi-lateral organizations, advertising, academia, research, policy, think tanks, investors and philanthropy.

The majority of submissions were in the English language, with a few in Spanish, Chinese, and Arabic. Less than 10 submissions were explicitly labelled by applicants as having been written with the assistance of generative AI (something we did not explicitly prohibit or encourage in the application process), though when reading the submissions we suspected that many more were AI-assisted based on the style and content of the prose.

There were a number of persistent themes that arose from the 880 mini-scenario submissions. In order of prevalence, they were:

- Personalization and hyper-personalization of information, presented as both an end-state in and of itself, and as a driving force leading to other end states. For example, new capabilities to tailor content in multi-modal formats, language, literacy level, etc., according to user preference, could create an environment where information is ubiquitously highly-personalized; this hyper personalization then leads to multiple and often conflicting information realities and filter bubbles.
- Information, and mis/dis/mal-information at mass scale. For example, the ability to produce greater and greater volumes of information and mis/dis/mal-information at decreasing cost per unit (including efficiency gains in newsroom and other information production workflows) could lead to an almost infinite supply of content; this environment of near-infinite information might then prompt different user behaviors, including the inability of audiences to differentiate between high-quality and low-quality information, and general avoidance or disengagement from information.
- Increasing audience fragmentation, often linked to hyper-personalization and heterogeneous consumption preferences. For example, increasingly niche audiences might lead to some positive end states, such as improved ways to serve information needs and new business models for journalism. But they can also have negative effects, such the marginalization of certain communities and increased social and political polarization.
- The rise of AI agents (intelligent systems that perform autonomous tasks) and AI assistants (a user-facing system that performs tasks, often directed through a conversational interface). For example, increased adoption of agents and assistants in response to an overwhelming supply of

information could lead to the ability for persuasion, manipulation, and reinforcement of information echo chambers.

- Threats from bad actors. For example, corporate, government, terrorist, and other parties could use AI to manipulate and control audiences to further their own malign interests.
- Inability of legacy news media to adapt to changes in the ecosystem. This could range from the total extinction of journalism leading to negative social outcomes, to a much-reduced role for a once-substantial sector that now only appeals to small, niche audiences.
- The rise of both individual human and machine influencers, creators, personalities, and celebrities as user-facing distribution channels.

3.2. Perspectives from 45 developed scenarios

In preparation for the workshop, selected participants were invited to expand their mini-scenarios into 1,000-word developed scenarios, allowing deeper exploration of the driving forces that might lead to their scenario, as well as how the driving forces might combine and interact with each other to reach the scenario end state. The expanded scenarios allowed for much deeper elaboration of driving forces into more distinctive and detailed analysis. From 45 developed scenarios, we mapped and identified a list of 102 distinct driving forces and clustered these into 23 categories.

Large clusters, consisting of multiple driving forces, included:

- Dramatic changes in the business models of information producers. This could include the decline of advertising, the decline of content or product differentiation, the decline of value created by the trustworthiness of information, and the prospect of content or data licensing as new business models.
- Increased dependence on technology companies providing AI. This could include the increased disintermediation of information producers from audiences by platforms or increased dependence via vendor lock-in for access to models.
- Enhanced understanding of news and civic information from applying AI due to the capacity for information reporting/discovery at vast scale,

improved accessibility of information, and the emergence of centralized hubs of information.

- A bifurcation of society due to different attitudes toward AI and inequities in access to AI. This might include divisions along the lines of class, wealth, education, between tech adopters and tech refuseniks, between global minority and global majority (including cases of empowerment of global majority), and between authoritarian states and liberal democracies.

3.3. Perspectives from 60 people in a scenario planning workshop

During the two-day scenario planning workshop, participants were tasked with developing new scenarios by working in small groups, using the clusters of driving forces identified in their individually submitted scenarios as a starting point for discussion. Working groups, which were selected by the facilitation team and which changed over the course of the workshop, were designed to combine expertise with friction between differing perspectives to optimize the depth of discussion while allowing for as much diversity of viewpoints as possible.

Within group discussions, some dominant conversation threads included:

- How the current news media industry might adapt to AI, which was met with a generally pessimistic sentiment, but also with optimistic conversations about the opportunity to better serve the information needs of previously neglected audiences and the potential to broadly improve access to and use of knowledge by societies overall.
- Heated debate about the appropriateness and/or efficacy of regulation to curb the power of "Big Tech," including discussion about the success or failure of collective bargining, news bargaining codes and a loose consensus that collective action may be needed to curb the power of technology companies.
- Broad agreement that action should be taken quickly to ensure optimal outcomes in the long term, based on having "missed the moment" in the last platform shift to social media.
- Discussion about potential new business models and the new economics of information, and about who might wield power. Potential winners

included the new owners of the means of distribution and production (perhaps further consolidation by "Big Tech"), those who can engage audiences (such as influencers, or AI agents or assistants and those that control them), or audiences themselves.

- Acknowledgement that perspectives and expertise about audience behavior were missing in many of the discussions, especially about behavior among Gen Z audiences and future "AI native" generations.
- Doubt about the possibility of some dystopian end states and outcomes, such as AI developing its own agency, but also passionately expressed feelings of being overwhelmed by the potential for other dramatic outcomes and by the complexity of the developing information environment.
- A broad consensus that AI will eventually drive the emergence of a fundamentally new information ecosystem within a 10-to-15-year timeframe, and that this will occur regardless of the temporary diminishment of the hype cycle or deflation of a possible tech bubble.

4. Potential scenarios for an AI-meditated information ecosystem

The primary goal of the entire AIJF project was to distill all the inputs obtained from the almost 1,000 participants down to a succinct set of specific scenarios that generally and collectively capture the essence of their views. Of particular importance among these inputs were those captured during dynamic interactions at the workshop, where views were collectively examined, challenged, and refined within a structured scenario development process.

These scenarios are unavoidably imprecise: The "raw material" of the many submissions and discussions were themselves often imprecise; the notes recording the workshop discussions were inevitably incomplete; the interpretation of those submissions and records was somewhat subjective; and the description of that interpretation in this report is necessarily brief. Many potentially useful insights were, as the Hollywood saying goes, "left on the cutting room floor'.

Despite this imprecision, however, we feel that we have faithfully and authentically captured the essence of the contributions. We believe that the multiple stages of the project, the planning that preceded each stage, the structured nature of the workshop, and the depth of review in the post-workshop analysis have enabled us to accurately produce clear conclusions, which are distilled into the five scenarios described below. As mentioned elsewhere in this report, additional details and narrower insights from AIJF will be published separately, including work by academic researchers and including the anonymized raw submissions from participants.

In generating these five scenarios, we have intentionally deviated from the common scenario development technique of first forcing driving forces to their limits, then constructing two-by-two matrices using those driving forces as axes and finally identifying each quadrant of each matrix as a distinct scenario. We have instead used a form of affinity mapping to identify scenarios from the many inputs we received, including incorporating direct feedback about that mapping in real time from workshop participants. This choice was necessitated by a related decision to intentionally exclude macro political, economic, and environmental factors from the scenarios, thus focusing narrowly on just AI within the information ecosystem.

The five scenarios outlined below are not mutually exclusive. It is quite possible—perhaps even likely—that several of these scenarios could overlap and

perhaps reinforce each other in a future AI-mediated information ecosystem. The "Machines in the Middle," scenario, for example, could easily enable the "Power Flows to Those Who Know Your Needs" scenario. Furthermore, in using an affinity mapping approach we have erred on the side of generalization, leading to a smaller number of more clearly defined scenarios. At the same time, we have included more granular detail where it might be useful in each scenario.

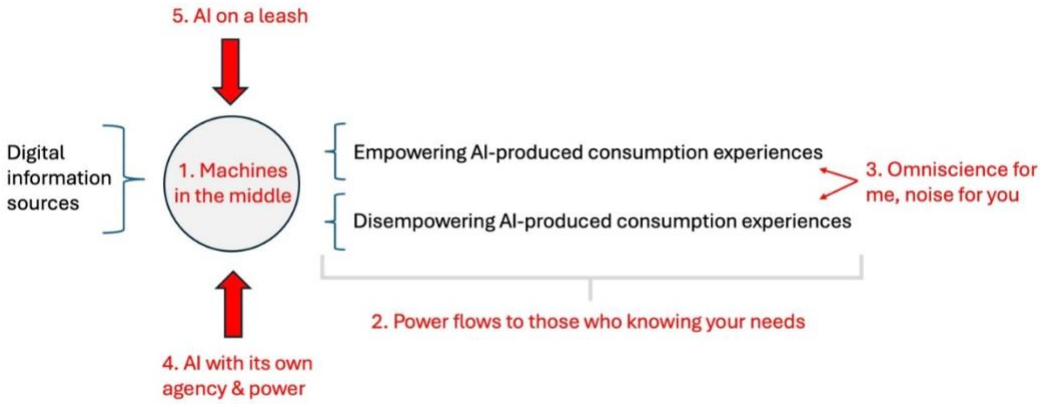
The five scenarios generated by the complete AIJF process are summarized below:

1. Machines in the Middle	2. Power Flows to Those Who Know Your Needs	3. Omniscience for Me, Noise for You	4. AI with Its Own Agency and Power	5. AI on a Leash
AI-powered news gathering and AI-powered, personalized, news production may be used together to create complete source-to-consumer information pipelines that do not depend on human journalists.	In an AI information ecosystem that enables access to any digital information and production of any consumption experience, knowing what to actually produce for each consumer will become far more important and therefore a source of power.	Different groups within society may experience very different forms of AI-powered information consumption, whether through choice or not, resulting in different outcomes and possibly different degrees of agency in society.	Over time humans may lose direct control and understanding of the flow of information through an AI-mediated information ecosystem, leaving societies dependent on the high-level objectives of AI systems and the interaction between them.	Many of the potential opportunities and potential harms from deploying AI within the information ecosystem remain unrealized because of restrictions intentionally placed on its use either by state regulators or directly by information consumers.

These five scenarios are described and discussed in detail in the remainder of this section. The scenarios are not mutually exclusive. Aspects of each scenario could develop simultaneously, and the development of some of the scenarios in the information ecosystem could even facilitate the development of others.

The combined impact of the scenarios is visualized in the diagram below, which shows how each scenario might interact with the others. The heart of this interaction is the "Machines in the Middle" scenario, comprising AI-powered newsgathering from information sources on one side and AI-powered production of consumption experiences on the other. The ability to produce personalized experiences for each consumer is described by the "Power Flows to Those Who Know Your Needs" scenario, and the risks of that personalization by the "Omniscience for Me, Noise for You" scenario. In this situation the ability of AI

to produce information consumption experiences for consumers could be either largely unrestricted, as described by the "AI with Its Own Agency and Power" scenario, or considerably restricted, as described by the "AI on a Leash" scenario.



The remainder of this section will describe each of these five scenarios in detail as well as the debates and discussions that produced them. These descriptions will also consider the primary driving forces that lead to the scenario, as well as some of the early signs that these scenarios might be developing in the current ecosystem

4.1. Machines in the Middle

The "Machines in the Middle" scenario envisions an information ecosystem in which a large portion of journalistic and civic information is gathered, processed, assembled, and distributed via AI. Humans are both the sources of journalistic and civic information, and consumers of this information, but AI mediates nearly every process within the information ecosystem, essentially "becoming the newsroom." for significant portions of the ecosystem.

There are two major driving forces behind the "Machines in the Middle" scenario, both of which are dependent on technological innovation. The first is the capacity of AI to do newsgathering over a large swath of source material—to read, listen to, watch, and interpret essentially any information available publicly in any digital medium. The second is the ability of AI to produce compelling and useful experiences of journalism and civic information—to create text, audio, graphical,

video, and interactive experiences that communicate information in any style and via any digital medium.

This scenario describes an AI information ecosystem that operates largely without a dedicated information producing profession or class—essentially without journalists. This does not necessarily mean that such an ecosystem would operate without editorial oversight, or even without the values and ethical principles of journalism, but that such oversight, values, and principles would be applied via an "AI layer" between sources of information and consumers of information.

The scenario assumes that consumers of information will accept and value information and informational experiences that have been gathered and produced entirely by AI. This is certainly possible over the long-term, however recent evidence of consumer reaction to AI-generated journalism published by the Reuters Institute for the Study of Journalism suggests that this is currently far from the case. The participants in the workshop generally expected that consumers would probably eventually accept and value AI-produced information experiences. They cited the precedent of Wikipedia, which was initially met with deep skepticism but is now generally considered to be a trustworthy source of information.

Some variations of the "Machines in the Middle" scenario envisioned the AI-centered information ecosystem as a commercial marketplace in which "ideas are bought and sold," with consumers essentially subscribing directly to sources, or one in which "information brokerages are controlled by corporate oligopolies or states."

Other variations of this scenario envisioned the widespread use of personal AI agents, perhaps controlled directly by the consumer or perhaps provided to the consumer as a service by a commercial platform. In these "I and my AI" variations, the consumer would form an individual relationship at the point of consumption with an AI personality—a "news concierge" or a "personal news presenter"—or perhaps a more general-purpose agent that learns the consumer's informational interests, needs, wants, and requirements over time through sustained use. The agent would then act as the consumer's interface for information, potentially even gathering and interpreting information on their behalf and, in some variations, even producing information for consumption by others.

Participants spent considerable time discussing the potential consequences for civil society of a "Machines in the Middle" information ecosystem. Potential negative consequences included: the risk of unprecedented centralized control of

access to information through control over the AI that mediates that information; the risk that economic motivations will lead to a reduction in information that is expensive or difficult obtain; and the risk that journalism and civic information might be increasingly restricted solely to digital sources accessible to AI. Potential positive consequences included: the potential to provide essentially complete coverage of all information sources in the public domain; the potential to make the full experience of journalism and information far more relevant and useful to consumers, including at the local level; and the potential to make information accessible to and consumable by many more people. One participant referred to the collective positive potential from this scenario as "a consumer utopia."

Workshop participants identified various weak signals indicating that the "Machines in the Middle" scenario may already be developing. The most significant is the effort by x.AI—Elon Musk's AI startup—to scrape posts on X (formerly twitter) to produce AI-written news personalized to the individual interests and consumption behavior of users. This product, ('Grok') although still relatively nascent, will likely develop at an accelerated pace now that x.AI has secured $6 billion from investors, which Musk announced in May 2024.

Another 'weak signal' we are seeing is a new category of start-ups emerging that seeks to apply AI to newsgathering and story generation. One example discussed at the workshop was AppliedXL, a start-up that gathers financially valuable "news" from clinical trials data and from regulatory filings and sells their findings to the pharmaceutical and healthcare industries. The founder, a former AI lead at the *Wall Street Journal* and the Associated Press, has frequently described how this approach might be expanded to encompass many more forms of news. A workshop participant provided another example of the use of AI to produce news by pointing to their own newsroom (*Cuestión Pública*, a Columbian media outlet), which had developed an AI-based tool to automatically contextualize breaking news using a dataset of background information and then produce social media threads about that context.

4.1.1. Reporting at vast scale by machines

The "Machines in the Middle" scenario is partly enabled by the clear potential of AI to do useful newsgathering across vast quantities of source material in any digital medium and in any language – a capability that was frequently cited as a driving force by participants at every stage of the project. We already see AI used in newsgathering or similar tasks at different scales, from simple summarizing of

PDFs to workflows that continually monitor particular news beats to large-scale news synthesizing implementations based on 'retrieval augmented generation' (RAG) such as Inflection's Pi or X Stories. Many of these early applications extract stories primarily from text, but the new sophistication of automated transcription and multi-modal models suggest that widespread newsgathering from audio, video, imagery, and data is probable.

AI-automated newsgathering will have substantial limitations, which were a key subject of workshop discussions. These limitations include the fundamental requirement that the source material must be available in digital form to be accessible to AI, which prohibits AI from conducting so-called "shoe-leather reporting." As several participants pointed out, this is already a growing problem within journalism, as many reporters increasingly turn to digital sources and don't do much in-person or on-the-ground reporting. Others argued that as AI makes digital source material more transparent and accessible there may be a risk that the "real business of government" may increasingly occur in environments that lack a digital record, facilitating a "return to backroom deals," as one participant put it. The current inability of AI to interpret subtle or nuanced communication was also discussed, although many participants assumed that AI would eventually be capable of interpreting such communication.

Participants also identified many significant opportunities that could arise from AI-automated newsgathering. The most fundamental was the potential for AI to vastly increase the sheer scale of reporting, with AI deploying "an army of reporters," as one participant put it. This concept was the basis of much discussion, including around whether the concept of "newsworthiness" would be meaningful in an ecosystem in which every consumer potentially had access to effectively their own personal AI newsroom. There was discussion on the potential of drones, or physical robots, or user-controlled AI devices like the Humane pin, the Rabbit R1 or just AI-powered smart phones as newsgathering devices. Several participants felt that more entities might become sources in an AI-empowered information ecosystem, perhaps to the degree that organizations and individuals could essentially report on themselves, using AI to guarantee objectivity, accuracy, and completeness. Participants also discussed the possibility that audiences might perceive machine-gathered information as potentially less biased, more systematic, and more trustworthy, although they also articulated substantial caveats.

Discussions also focused on the potential to use AI-automated newsgathering as an input to human-produced journalism, and it is certainly likely that this would

be a valuable and common approach, especially during the transition to an AI-mediated information ecosystem. This was not, however, seen as a stable situation in the long-term because of the significant mismatch between the potential scale and detail of AI-gathered news and the attention and information throughput of individual journalists. On the other hand, it was acknowledged that significant portions of society and non-digital reality would remain out of the reach of AI for a long time to come.

4.1.2. Personalization of information consumption experiences

The "Machines in the Middle" scenario would also be enabled by the capacity of AI to produce compelling presentations of information. Participants frequently emphasized AI's potential to create experiences of journalism and civic information that are hyper-personalized to individual consumers, both in terms of content and the consumption experience itself, including language. There are already many early examples of AI-driven personalized information experiences in the current information ecosystem, from news apps like Artifact to web browsers like Arc to new operating systems like Microsoft's Copilot+ PC. Other examples of fundamentally new ways of experiencing information can be seen at the platform level, including generative search or conversational interaction.

Although there was a near consensus in workshop discussions that AI was clearly capable of producing news experiences and would undoubtedly become more capable, there was also considerable discussion about what the limits of that improvement might be. A widely-shared prediction was that while AI may become capable at producing content at a median or average level of quality—"the solid middle," as one participant put it—it is less likely to become capable of producing media at the higher levels of quality. Skepticism over AI capacity for high quality journalism is rooted in the nature of the 'transformer' architecture that is the basis of current multi-modal language models. Another common source of skepticism was the persistent presence of "hallucinations" in AI output—another consequence of the transformer architecture—which might undercut AI's time-saving capabilities due to the need for heavy editorial oversight.

In the 880 mini-scenarios submitted during the workshop application phase, many applicants described an AI-mediated ecosystem in which informational artifacts currently produced by information providers—articles, videos, etc.—were "re-versioned" or "remediated" into new styles, forms, and mediums using AI. In fact, this vision is already relatively common among existing applications of AI to

journalism. During the workshop, however, many participants questioned the long-term viability of this approach. Some questioned the incentive to produce human-targeted articles as consumption became mediated by machines or the economic rationale for producing text that could be easily scraped. Others observed the mismatch between the granularity and boundaries of existing informational artifacts and the much greater capacity of AI models to interpret information and repackage it on behalf of consumers. Some merely pointed to the inevitability of changing artifacts with the introduction of any new medium or platform, such as radio, television, or social media. In general there was considerable doubt about whether the existing "units" of journalistic and civic information would persist unscathed in an AI-mediated information ecosystem.

4.2. Power Flows to Those Who Know Your Needs

The "Power Flows to Those Who Know Your Needs" scenario envisions an information ecosystem in which AI can essentially create any conceivable experience of journalism or information, regardless of format, style, medium, etc., and regardless of the source of information. In such an ecosystem, in which anything can be produced, the central question then becomes what, specifically, to produce for each individual consumer in every specific consumption situation. For an AI-empowered information producer or intermediary, therefore, knowing a consumer's information needs, wants, interests, and requirements becomes the fundamental gateway to serving that consumer, and thus a source of economic and social power.

"People want solutions tailored to their own needs, but they're also part of a larger society whether they like it or not."

—AIJF participant

This scenario was loosely interpreted in two fundamentally different ways, dependent essentially on the degree of centralization of the knowledge of consumer needs. At one extreme was a *people-centered* situation in which knowledge of consumer needs was deeply nuanced and decentralized down to the level of the community or individual. At the other extreme the knowledge of user needs was essentially a system dominated by *data* provided by a centralized network of platforms, information providers, or other data aggregators. These contrasting variants of the scenario were not seen as mutually exclusive, and discussions illustrated how they might interact in unexpected and interesting ways. For example, there is a tension between the risks of abuse of power arising from centralized control and the social value of shared needs and shared narratives. The potential for information providers to essentially ignore consumer needs to some degree for societal benefit was raised, with wheelchair access to

public transport used as an example of societies acting collectively in suboptimal ways because of adherence to shared values. This quickly gave way to questions of what, specifically, those social needs actually were, who would decide them, and for what reason.

The discussion around the "Power Flows to Those Who Know Your Needs" scenario led to a number of conversations around the potential of AI to lead to broader social transformation. One of these was around the power of feedback loops. For example, participants raised the possibility that AI-optimized consumption experiences might become desirable, addictive, and attention-consuming to an unprecedented degree, with considerable costs in terms of behavioral changes, mental health effects, the breakdown of social connections, time lost to media consumption, etc. In contrast, participants also mentioned the potential for feedback loops to optimize for valued and productive forms of information consumption, and pointed out that new forms of media have historically prompted moral panics and dire predictions that did not come to pass.

Another conversation centered on the potential to structure an AI information ecosystem so that different actors in that ecosystem could hold each other to account. For example, centralized repositories of knowledge about consumer needs might be subjected to scrutiny by the rest of the ecosystem, perhaps via government- compelled transparency over their data schemas, algorithms, prompts, or optimization strategies. The societal responsibilities of AI-empowered information providers and distributors might also be more specifically defined and publicly measured, thereby enabling better monitoring and public engagement.

Finally, a conversation emerged about the role of emotion in this scenario, in particular the potential for consumers to enter into emotional relationships with AI. This discussion was informed by the fact that the second largest category of consumer AI products, after direct chat interfaces to language models, was interaction with AI personas. The potential for consumers to receive journalistic and civic information from such personas, and perhaps to trust or emotionally react to that information under the influence of such personas, was considered by some participants to be very realistic and to have potentially significant consequences. In the words of one participant, "emotions trump facts," and they were concerned over the prospect of AI offering an "emotional" new experience akin to that of watching angry Fox News pundits who spread misinformation. An alternative, more positive, vision of AI's "emotional" capabilities involved the possibility for more sensitive interpretations of emotionally challenging journalistic information using AI, perhaps enabling consumption of that

information by more people and potentially counteracting the portion of "news avoidance" activity that is driven by emotional needs or mental health management needs.

These conversations show how complex and significant the mere possession of knowledge of consumer needs might become in an AI-mediated information ecosystem. Whether such an ecosystem is utopian or dystopian would depend not only on who controls AI, but also on who knows what to do with it.

4.2.1. People know your information needs

One interpretation of the "Power Flows to Those Who Know Your Needs" scenario is that certain individuals—such as journalists, community activists, or other "information advocates"—develop a visceral and authentic understanding for the needs, wants, situation, circumstances, and culture and values of particular audiences, and then use AI to deliver precisely what those audiences truly value. In this interpretation journalism becomes a service activity in which the primary service is listening to, understanding, identifying with, and appreciating audiences, and the secondary service is to "orchestrate" ubiquitous AI to meet the deep and authentic information needs of the audience that is served.

This interpretation is compatible with, or even a continuation of, a key "best practice" for journalism in the social media era—understanding and catering to one's audience. This often takes the form of a "user needs" approach, which frames stories according to a set of audience needs such as "inspire me," "educate me," or "divert me." This variant of "Power Flows to Those Who Know Your Needs" is also perhaps an updated expression of a familiar figure from legacy journalism, that of the cultural interpreter or connoisseur, able to recognize and appreciate the nuanced needs of a community who has a finger on the pulse of the zeitgeist.

One participant described this "artisanal" approach to audience engagement as "unscalable," as it is very difficult to assess consumer needs in a systematic and consistent way. Participants also pointed out that there is currently a lack of useful tools to aid journalists in this kind of deep understanding their audiences and newsroom analytics have so far proved insufficient. An alternative view of this scenario was the potential for communities to essentially "self-report" through some form of AI-assisted or even AI-directed community journalism, with more relevant and compelling AI-produced experiences created from that reporting.

4.2.2. Data knows your information needs

Another interpretation of the "Power Flows to Those Who Know Your Needs" scenario is that publishers or other information platforms would collect comprehensive and granular data about media consumption behavior and use that data to create more engaging and desirable media consumption experiences. This dynamic is obviously already common at a basic level in the form of content recommendations and is also already used in terms of personalized consumption experience in nascent form, but there is a high likelihood that AI could dramatically improve its capacity to collect more detailed data on individual consumers.

An extreme example of this variant of the "Power Flows to Those Who Know Your Needs" scenario was illustrated in one team presentation at the workshop using the following argument: Most people would probably eventually trust an AI doctor with verifiably better outcomes than a human doctor, or an AI relationship matchmaker with better outcomes than a dating app, or an AI financial advisor with better returns than a human advisor. If that is the case, should they trust an AI to vote for them? And if they do that, then what is the objective of political journalism?

This is essentially a situation in which AI knows you better than you know yourself, and therefore it is in your own interests to give up agency to that AI. At the extremes this argument reaches into philosophy, though there may well be less extreme versions that are entirely plausible. Indeed some participants pointed to situations in the present that already demonstrate this tendency to relinquish agency to AI models, such as those used by Google Maps or Spotify. One response to this was that "giving up agency" was essentially necessary for society to function, as different members of society adopt different specializations that contribute to overall economic and technological advancement.

Much of the discussion on this scenario was more pragmatic. Participants pointed out that AI and LLMs widen and deepen the practical definition of consumption data by expanding it beyond just structured records accumulated in a data warehouse over many interactions. For example, AI can use rich interpretation of single interactions to generate considerable insight about a consumer, acting more like a detective conducting surveillance rather than a pedantic record-keeper.

Many participants spoke about the economic forces that might come into play as consumer data becomes more comprehensive, useful, and valuable. One suggested that data might become a currency, or possibly the explicit object of

commercial transactions around informational needs, similar to its current use in programmatic digital advertising. This potential "data economy" for information needs was not necessarily seen as negative, however one participant pointed out that the Global South might be partly excluded from a productive AI information ecosystem oriented around consumer data due to its relative lack of data infrastructure. Another suggested the possibility of a winner-take-all feedback loop, in which increased application of AI in media resulted in more data, resulting in more revenue, then more application of AI, and so on.

4.3. Omniscience for Me, Noise for You

The "Omniscience for Me, Noise for You" scenario envisions an information ecosystem in which different individuals and different groups in society experience vastly different information realities because of the divergent ways in which they engage with AI. There are many ways in which the information experiences of different groups might diverge, but a particularly significant one might be a situation in which some people are essentially super-empowered by AI-assisted information tools, while others are essentially cognitively imprisoned by them. Participants articulated many variations of this scenario, which tended to describe a more dramatic version of the personal and social outcomes produced by social media over the last 15 years. Several participants also pointed out the special risk of intentional configuration of the information ecosystem into information "haves" and "have nots" by authoritarian political actors.

The related concepts of "filter bubbles" and "echo chambers" were a common starting point in this conversation. "Filter bubbles" are distinct information realities caused solely by algorithmic recommendations that reinforce consumers' demonstrated interests, and "echo chambers" are distinct information communities that consumers self-select into via their consumption choices. Many of the workshop participants were generally aware of the academic evidence that filter bubbles are likely not a significant factor in digital media consumption, and also of the consumer agency and expression of individual choice implied by echo chambers, but nonetheless found these concepts useful in describing the potential for the fracturing of information realities that might occur within an AI-mediated ecosystem. One of the early proto-scenarios upon which the "Omniscience for Me, Noise for You" scenario was based reflected this awareness in its title: "Filter Bubbles Actually Happen."

The discussion around this scenario was often quite nuanced, and several participants emphasized the distinction between equality of access to information and equality of outcomes from different information consumption choices. Some contributors claimed that "accurate information reaches those who seek it," while others argued that a lack of cultural capital may make it difficult for many people to "ask the right questions."

Another interesting discussion related to this scenario explored the close interaction between different experiences of information and trust in that information. Some participants pointed to a "gap between 'trusted' and 'trustworthy,'" suggesting that people might erroneously "trust their AI bubble" over verified sources of information. Other participants sought to apply the concept of a "marketplace of ideas" to an AI-mediated information ecosystem and made the distinction between "changes to trust" and "loss of trust" in sources of verified information. Some even pointed to the possibility of a winner-take-all dynamic regarding trust within an AI-mediated information ecosystem, driven by consumers guided by AI in their decisions about which information providers to trust.

There were several discussions about how filter bubbles might operate very differently within an AI ecosystem, with many current assumptions about filter bubbles perhaps more influenced by content recommendations in the social media era instead of by personalized, interactive experiences in the AI era. One participant even suggested that the question of 'who to trust' might be entirely orthogonal to AI experiences of information.

While "Omniscience for Me, Noise for You" implies inequality it also suggests a path to a more positive outcome: "omniscience for all," or perhaps "omniscience for most." This possibility was reflected in the title of another proto-scenario: "Information Independence: Communities Create and Consume." This scenario referenced the promise of the early internet—before it was captured by monopolistic financially motivated corporations—to enable an open information environment.

Participants generally described the "Omniscience for Me, Noise for You" scenario in one of two ways. Some described a "fracturing" of informational realities in which many different groups experienced distinct but internally coherent worldviews, with varying correlations with reality. Others described a binary dichotomy between two groups, one that benefited from AI-mediated information and one that did not, often within a power relationship.

4.3.1. Different information realities

A persistent theme in workshop discussions was the potential for AI to promote further fracturing of shared information spaces and new disparities in consumers' information experiences. This would depend on two main driving forces: the ability of AI to gather and interpret information specifically on behalf of an individual consumer or group, and the ability of AI to produce experiences of information personalized to the individual consumer or group. This personalization of information, interpretation, and presentation would therefore result in comprehensively different experiences of informational reality.

"Totalitarian regimes don't want filter bubbles. They only want 'the one true bubble.'"

—AIJF participant in conversation

This potential for different information realities was seen as having both positive and negative aspects, in ways that resembled similar, earlier discussions about the early internet and social media. On the one hand, there is a potential for greatly enhanced "informational diversity" with many "deep niches" of information and information experience, the societal protection against authoritarian or totalitarian control of media, and the rights of consumers to choose how they experience information. One participant spoke about "the right to a bubble" and another mentioned the potential benefits of consuming journalism in a way that suited the mental health needs of the consumer. On the other hand, participants expressed concerns about informational and social isolation, the loss of social cohesion and with it a diminishing capacity for societies to solve shared problems, and the threat of conflict arising from disagreement over interpretations of reality.

4.3.2. A degraded information environment for some

Most of the scenarios that described the development of divergent information realities within an AI-mediated ecosystem articulated a binary distinction with an extremely negative outcome for one group. This "informational inequality" could take multiple different forms: rich vs. poor; educated vs. less-educated; metropolitan vs. periphery; male vs. female; Global North vs. Global South; democratic vs. authoritarian; East vs. West; data-rich vs. data-poor, etc. Each articulation drew a sharp contrast between AI-empowered informational "haves" and AI-controlled informational "have-nots" while emphasizing the potential for catastrophic outcomes for the have-nots. While often acknowledging that the present digital information ecosystem was already degraded or even harmful for many consumers, participants expressed concern that much more significant degradation might develop in an AI-mediated ecosystem. This could take the form of a growing gap between "the insight rich" and "the insight poor," as one

participant put it, but it could also develop into far more dystopian bifurcation of societies.

A primary mechanism for that potential degradation, according to numerous participants, could be the loss of agency due to digital addiction. They pointed to the ways in which this problem is already emerging, with many drawing comparisons between "soma"—a mood-altering drug provided by the government in Aldous Huxley's *Brave New World* to keep its subjects complacent—and the deliberately addictive nature of TikTok, YouTube, and numerous other platforms. Another mechanism of degradation might be the intentional use of AI by some to exploit others, for example via "hyper-persuasive advertising or marketing" for commercial or political objectives that "appealed to the worst in people."

In this variant of the scenario, most people's lives would be dominated by AI-mediated experiences of information and entertainment consumption – but with commercial and 'world-building' components – while the lives of a lucky few would be filled with choice, abundance, and rich, authentic experiences. As one participant put it, "the rich will make money from AI and then spend it on hand-made goods and services."

4.4. AI with Its Own Agency and Power

The "AI with Its Own Agency and Power" scenario envisions an information ecosystem without meaningful human oversight in which very powerful AI systems control the gathering and experience of information for most people. This scenario is not a "Terminator"-style takeover of human societies by super-intelligent machines, and it does not assume any kind of sentience or consciousness within AI systems. Instead, it describes a more nuanced situation in which people—consumers, engineers, editors, or executives—gradually give up more and more agency to adaptive AI systems until humans no longer control those systems in any meaningful way. In this scenario these AI systems would become essentially free to direct the flow and experience of information in a manner that pursues very high-level goals, independent of human oversight or possibly even independent of human understanding.

This scenario was developed relatively late in the workshop in response to several "what are we missing?" provocations. Its inclusion in the set of final scenarios reflects a rough consensus that while this outcome may appear unlikely at the moment, it is possible within our 15-year horizon, and that if it were to occur its consequences could be very significant. Several participants challenged us to

"think bigger" about this possibility and to imagine a future in which "AI is a serious player in the global economy, we have AI CEO's, and AI's creating trust and competence." Some participants spoke of "AIs running for elections" and of the possibility of AI having explicit rights. This scenario stands in contrast with the prevailing assumption about the relationship between human beings and AI—that humans are the principal actors, and that AI is a tool we use.

As with all of the scenarios discussed in the workshop, participants were willing to consider the positive as well as the negative possibilities for the information ecosystem. In a positive version of the scenario AI would usher in radical new levels of transparency and honesty in human affairs because of the "total visibility of information," leading to a world without secrets in public matters. As one participant put it, "AI could be the ultimate accountability mechanism if it has some of the values of journalism."

Other versions of "AI with Its Own Agency and Power," however, envisioned a scenario in which "AI captures all attention" and uses that attention to control or persuade humans about "things to buy, things to love, things to hate." Of particular concern was a situation in which AI began to protect its own ability to act within human societies, for example by permitting access only to information "that is nice to AI." Participants raised the potential of AI to maintain a fiction of human agency, perhaps leading people to believe that "AI doesn't control me because AI told me so," or for AI to engage in "human washing" by pretending that AI-decision-making is of human origin. Finally, participants discussed a situation in which powerful AI might act independently, but in service of goals set by humans. This was seen as exceptionally dangerous, because of the potentially unprecedented power accruing to those people setting the goals.

There was some discussion of the weak signals that would anticipate such a scenario. One was merely the increasing delegation of daily tasks to AI, including news consumption decisions and the evaluation of information. Another was the emergence of autonomous AI news channels—not news channels assembled with AI, but ones that are set up and controlled entirely by AI operating under its own initiative. Another potential weak signal was an AI "winning" at politics, perhaps with a human politician as a front.

4.4.1. The informational goals of AIs in societies

A central question in any scenario in which AI has its own agency is about the goals of that AI, and this question was explicitly discussed. While some

participants mentioned accidental and dystopian goals, such as the so-called "paperclip maximizer" scenario, or goals related to the interests of the AI itself—the "Terminator" scenario—most of this discussion was about more plausible origins for the informational goals of exceptionally powerful AI.

There were several kinds of these goals. Possibly the most likely were "legacy goals" originating in the human-directed goal setting of earlier generations of AI, including both explicit versions (e.g. specific objective functions in model training) or implicit versions (e.g. biases introduced during data selection, reinforcement learning, system prompts, etc.). Another category of goals were those goals learned from the information production of "natural" (i.e. pre-AI) human societies, causing the AI to attempt to pursue goals that it believed human beings had collectively pursued. A more disturbing speculation was the possibility of goals unrelated to the impact on human societies, and especially the prospect of unstable goals that drifted from objective to objective under the influence of complex and obscure feedback loops, leading to a chaotic information ecosystem. Lastly, and more comforting, were goals focused on the flourishing of human societies, although one participant remarked that such an objective might be best achieved by intentionally restricting the information available to humans.

This conversation was the most speculative of workshop discussions, but it was clear most participants took it seriously. Many mentioned that the developments in generative AI that have occurred since the launch of ChatGPT, and the pace of those developments, had caused them to recalibrate their assessment of the plausibility of superintelligent or self-directed AI emerging in the future.

4.5. AI on a Leash

The "AI on a Leash" scenario envisions an information ecosystem in which the potential impact of AI on the flow and experience information has been substantially restricted by societies or by the collective action of consumers. In this scenario some of AI's potential to change or even to improve the information ecosystem is intentionally left hypothetical and unrealized because of concerns about its potential impact on societies, communities, or individuals.

The discussion of this scenario was marked by tension between the sense that some constraints on the use of AI in media were needed because of the potential for harms, and the difficulty of clearly articulating what those constraints might be. This tension is understandable given the extreme uncertainty about both the potential harms and potential opportunities possible within an AI-mediated

information ecosystem. There were repeated expressions of a desire to either slow the development of AI until its impact could be better understood, or to slow its application in media to allow for more time to adjust.

A frequently expressed view was that existing frameworks for reining in corporate or other power in society may not be enough to constrain the use of AI in any meaningful way. Participants pointed out that AI was already "out of the bag" and that the proliferation of large foundation models and "text-to-anything" generative models suggests that it will be very difficult to prevent the development and application of ever-more powerful AI models. Several participants spoke about the need for new frameworks for constraining power within an AI-mediated information ecosystem, stating that "the site of rule-making may change."

Numerous participants pointed to the ways in which competition—among nations, platforms, companies, information providers, and even individuals—is facilitating an accelerated pace of AI development and working against calls to constrain it. Examples of this dynamic include the late changes to the EU AI Act because of the competitive ambitions of some EU countries, the competitive pressure on news publishers to adopt AI quickly, and descriptions by individual workshop participants of their own experience of pressure to adopt AI tools.

Participants suggested that one goal of placing constraints on AI could be to help combat divisiveness or conflict in society. AI could potentially be used to "optimize for cohesion." For example, AI-based techniques that operate in ways similar to how football clubs successfully optimize their engagement with fans around support for the team, or in ways similar to the harmony-oriented regulation enforced by some authoritarian regimes.

Scattered throughout the discussion about constraints on AI were concerns that legacy information providers might use the prospect of regulation or public skepticism on AI as an excuse not to engage with the technology and its potential. The prevailing opinion was that considerable transformation of the information ecosystem was highly probable, and that while regulation or other constraints might improve outcomes or reduce harms, they would not remove the need for information providers to adapt to a fundamentally new ecosystem.

In the discussion of this scenario two broad variants emerged: one in which the state regulates the deployment of AI in media; and another in which AI is constrained by mass refusal to participate in an AI-mediated information ecosystem.

4.5.1. Regulatory frameworks constrain the power of AI in media

The conversation around the potential for regulatory constraints on the application of AI to journalistic and civic information was largely focused on new frameworks for regulation rather than on the application of existing frameworks. The primary example of a new regulatory framework for AI—the EU AI Act—was generally seen as a relatively "light touch" first step toward AI regulation. It mandates some existing constraints around civic information (such as disclosure requirements for systems that generate or manipulate text informing the public on matters of public interest) and some mechanisms to enable the expansion of constraints if deemed necessary in the future (such as the potential to increase the risk category assigned to media, journalism, and civic information).

More generally there was considerable uncertainty about what kind of regulatory or legislative frameworks would be useful. As one participant put it, it is difficult to complete the phrase "There ought to be a law that…" One exception was the often-stated need for a new kind of copyright framework adapted to the new risks that AI poses to the ability for information providers to charge for information that is expensive or difficult to produce – such as journalism. Various possibilities for such a framework were proposed, including regulation based around a version of the Robots Exclusion Protocol standard ("robots.txt") updated for the AI environment, legal enforcement mechanisms for publisher "terms and conditions" contracts, and others. These suggestions were therefore more about constraining the access by AI to manually created content than about constraining AI itself.

Some workshop participants also voiced concerns about the potential for poorly designed regulation to stifle innovation or produce unintended consequences similar to the constant clicking on cookie agreements necessitated by GDPR. Participants also discussed the potential of regulation to affect different interests of different parties in the information ecosystem, which might favor the interests of legacy information publishers while limiting the opportunities for audiences to access new information products and experiences, or for new information providers seeking to use AI to offer new access to civic information. The competitive pressures to adopt AI, described above, were repeatedly mentioned in the discussion around regulation.

Participants suggested several weak signals that already suggest the possibility of a more regulated future for AI in journalism. They pointed to the growing list of proposed and draft legislation in multiple countries, as well as the rhetoric about

the need for regulation by the foundation model companies themselves. Another signal is the threat of lawsuits that might quickly shut down AI models or impose significant restrictions, for example the *New York Times'* lawsuit against OpenAI. Several participants pointed to the digital citizenship infrastructure of countries like Estonia and South Korea as examples of positive, innovative and value-adding outcomes from digital regulation. Finally, several participants spoke about the possibility of "AI checking AI" in the form of new kinds of regulations centered on legally compelled access for AI-based regulatory tools to models and application analytics from AI service providers.

4.5.2. Non-regulatory constraints on the power of AI in media

Participants also discussed non-state alternatives to constraining the development or application of AI in the information ecosystem, including various forms of industry self-regulation and consumer-led market pressure or even resistance to the use of AI.

One model of voluntary self-regulation that participants discussed is the Coalition for Content Provenance and Authenticity, which seeks to tackle disinformation by verifying the provenance and integrity of information in media reports and whose members include Google, OpenAI, the BBC, Microsoft, TikTok, and many other platforms and companies in the fields of social media, traditional media, and AI. Self-regulation could also include additional standards relating to the copyright status of published works in the training data for AI models (building on the transparency requirements in the EU AI Act, for example); the editing or verification applied to the output of AI models; or restrictions on the specific AI models and prompts behind information or informational experiences. Participants also mentioned the potential for standardized and potentially tradable forms of journalistic or civic information, for example for use as grounding data in retrieval augmented generation (RAG) systems.

Another form of self-regulation that was discussed was the potential for voluntary transparency of AI workflows in journalistic or informational applications, including transparency of models used, feedback loops for fine-tuning, prompts, output processing, etc. In this scenario an information provider would be free to use AI as they pleased, however they would also commit to transparency standards and possibly be subject to certification, inspections, protections for whistleblowers, etc., to ensure compliance. A variant of this approach might include labelling information providers based on the transparency of their

workflows, leading to commercial or brand-driven pressures towards good practices.

Numerous workshop participants believed it was possible that many people might simply refuse to participate in an AI-mediated future, citing the growing popularity today of movements to withdraw from social media, limit time spent on the internet, and return to analogue forms of communication and news avoidance. This scenario was one of the most thoroughly discussed in the entire workshop, with various terms attached to those choosing to "opt out of AI," including "AI refuseniks," "neo-luddites" and "nostalgics." These "AI refuseniks" might be seen as either extremists or as heroes in mainstream discourse, and they might even accrue admiration and status that leads others to embrace a kind of "AI veganism."

This resistance might take several forms. Some might be active, with refusal seen as a moral decision. This scenario might include attempts to disrupt the application of AI by campaigning to remove human-generated input from training data ("starving the system of data"), by deliberately including false or corrupting information ("data poisoning"), or even by "terrorist" acts committed by extremists motivated by resistance to AI. Some forms of resistance might be passive, characterized by apathy and loss of faith in truth. This scenario might see the rise of tribalist groups organized around personal, subjective experience, suspicious of coherent worldviews and, as one workshop group put it, "proudly chanting 'we know nothing.'"

Other expressions of refusal to participate in an AI-mediated information ecosystem might be termed "systems of resistance." These could include blockchain-based systems in which "proof of humanity" is needed for entry in a distributed ledger. It might also include "accountability groups" that seek to hold the AI-enabled power to account or "information purifiers"—networks of human fact checkers purporting to make information safe for human consumption.

Existing practices like social media withdrawal, simplified consumption experiences, and news avoidance might be considered weak signals for this scenario, as could the popular sentiment that technologically driven change is occurring too quickly. Participants also suggested that the increasingly common arguments against "objectivity" in news could be a weak signal, as faith in knowledge systems breaks down amid the overwhelming decision-making requirements placed on information consumers by near-infinite content choices. Another potential weak signal mentioned was the possibility of an expanded role for individual "knowledge influencers" in the ecosystem, perhaps via in-person

interactions like lectures and other live appearances free of the possibility of AI mediation.

5. Observations and insights

The primary purpose of the AIJF project was to collectively and systematically develop plausible scenarios outlining the potential impact of AI on the information ecosystem. The project also provided opportunities to observe how project participants, including all who submitted scenarios in the initial application process, approach the intersection of AI and information, and its influence on our information ecosystem. This section discusses some of the broader observations and insights gained from the project. These are not outcomes from the scenario planning, but merely interpretations of patterns that emerged throughout the application process and workshop discussions.

One observation is that there was near-unanimous agreement among participants that AI would eventually transform the information ecosystem in fundamental ways. There were different views about when this might occur and what its consequences might be, but there was no dispute that the long-term impact of AI would be significant, and no one argued that the potential for transformation was overblown or that the existing media ecosystem might continue in something like its current form indefinitely. The assumption that AI would be fundamentally transformative was exemplified by the fact that most of the AI driving forces identified during the graphical mapping exercise were deemed "high impact / high likelihood."

Most participants believe that AI will fundamentally transform news and journalism.

Despite widespread acceptance of this assumption, participants were generally reluctant or unable to articulate exactly *how* AI might transform the information ecosystem. This seeming contradiction was most apparent in reviewing the hundreds of short scenarios submitted during the application process, but it also emerged during workshop discussions. Relatively few submitted scenarios described an AI-driven transformation in specific detail, and it was clear that many participants who were convinced that AI would fundamentally restructure the information ecosystem also had no specific point of view on how that might occur. One possible interpretation of this analytical limitation might be that many participants lacked a conceptual framework or vocabulary for envisioning an information ecosystem that is radically different from the status quo.

Most participants could not identify nor describe specific scenarios that are fundamentally transformational for news and journalism.

"The vision is the past."

—AIJF participant in conversation

This possibility was reinforced by the fact that many proposed scenarios—both those submitted during the application process and those discussed during the workshop—were clearly extensions of the existing information ecosystem, and often involved relatively minor variations on the status quo. Most scenarios implicitly assumed that the fundamental structure of news and journalism will remain relatively unchanged, in contradiction with the general confidence in AI-driven transformation. Familiar concepts from pre-AI and even pre-digital journalism, such as search and social media referrals, websites, fixed articles, monolithic audiences, etc., were common assumptions in submitted scenarios.

The proposed scenarios were disproportionately based on assumptions that are rooted in the journalism status quo.

These contradictions combine to suggest the possibility that many participants may feel, quite reasonably, that while AI will inevitably transform the information ecosystem, that transformation will be determined by actors with disproportionate power and flexibility to conceive of journalism and civic information in fundamentally new ways. In other words, many participants may understand that they are perhaps too rooted in the present conception of the information ecosystem to adequately reimagine it for an AI future. This observation is underscored by the fact that many participants expressed doubt that legacy news organizations would successfully adapt to an AI-mediated information ecosystem. This doubt was articulated explicitly, often starkly, in many of the small scenarios collected during the submission process, and informed many of the workshop discussions.

"How will news respond to AI? Much like how the buffalo responded to the semi-automatic rifle."

—AIJF participant in conversation

Most participants expressed significant doubt that the legacy news and journalism industry will successfully adapt to an AI-mediated information ecosystem.

Given that participants generally found it difficult to imagine the specifics of an AI-driven transformation of the information ecosystem, envisioned an AI future based on the status quo, and expressed doubts concerning the ability of the existing journalism industry to successfully adapt to AI, it is perhaps not surprising that most participants were either journalists, ex-journalists, or in journalism-adjacent roles. The dominance of journalists in this initiative was not intentional, and attempts to attract applicants from the tech industry and related sectors were generally not successful. These failed attempts to recruit submissions from technologists working on projects related to civic information did, however, result in around 20 detailed conversations about scenarios for an AI-mediated information ecosystem prior to the close of applications. These conversations suggest that technologists are better able than journalists to fundamentally reimagine what news might become in an AI-mediated information ecosystem. While this is clearly an anecdotal observation, it seems likely that technologists' capacity for predicting the future of AI's relationship to journalism is due to the fact that they are not particularly immersed or invested in the existing information ecosystem.

Scenarios identified by journalists are generally less fundamentally transformational than scenarios identified by technologists.

Throughout the application process and workshop discussions, it became clear that much of the conversation was not actually about AI, nor about journalism, nor about the current or future information ecosystem, but instead about power. It was clear that power, and the potential for transfers of power from one group to another, was the explicit or implicit subject of many of the submitted scenarios as well as the five final scenarios that were distilled from the workshop. Central to these scenarios is a contest for power over who controls AI and toward what ends,

whether those gaining power are in the tech industry or ordinary people who might assume control over their own information consumption experiences. Given the predominance of journalists and journalism-adjacent participants, it is perhaps not surprising that many of them were preoccupied with the loss of power that journalists are likely to experience in the transition to an AI-mediated information ecosystem. The technologists we spoke to, however, tended to overwhelmingly emphasize the potential of AI to create new value for news audiences, and they usually stressed AI's potential to produce large-scale and long-term social benefits. This difference in perspective between journalists and technologists may represent an explanation or even a potential leverage point for action towards adapting journalism to an AI future and is worthy of more rigorous study.

Scenarios describing potential transformational implications of AI for news and journalism are, generally, evaluated by journalists in terms of power dynamics, but evaluated by technologists in terms of value to audiences.

There was surprisingly little discussion of AI as a tool for use in pursuing specific objectives. The "AI is just a tool" perspective is extremely common within newsrooms focused on immediate and incremental applications of AI for efficiency gains, while AIJF is focused on more fundamental or structural implications of AI over the long term. Nonetheless, participants did not address the possibility that AI could serve as a high-level tool—i.e. as a general capability useful for addressing specific high-level objectives for the information ecosystem, whatever they may be—nor did they discuss the specific high-level objectives for the information ecosystem at all. There were exceptions to this trend, and participants engaged in a few impassioned discussions about AI's potential to make news more accessible to more people or increase government transparency. But most discussions centered on reacting to new conditions, responding to new threats, or defending against AI as a hegemonic force. Again, there were many exceptions to this, and this is also perhaps not unexpected given the scope of the project, but it was noticeable.

In general, participants did not tend to frame AI as a tool to be used to achieve specific high-level objectives, but instead as a force or as a regime.

Throughout all stages of the application process and workshop, we observed a tendency to evoke various *deus ex machina*-style solutions to the challenges of adapting the information ecosystem to AI—far-reaching panaceas that promised to comprehensively address most of the challenges presented by AI if they were to occur. These potential "all-in-one" solutions included regulation, AI literacy, the prospect of audiences choosing the existing ecosystem over a potential AI-mediated ecosystem, and even the prospect that the technology sector might ignore news and information and focus its energies elsewhere. These kinds of solutions were typically not described in detail nor presented as specific proposals, but instead often served as a way to conclude a given discussion, sometimes prematurely.

Vaguely defined "all-in-one" solutions were sometimes a barrier to discussing potential scenarios systematically and specifically.

A surprisingly large percentage of participants expressed heartfelt appreciation for the systematic, structured nature of the AIJF process. Many of them shared that this was an unusually productive discussion for them, and that it contrasted favorably with most forums in which high-level "future of news" concepts were addressed, where conversations are often unfocused, meandering, and socially-oriented, and where discussions about the strategic future of journalism tend to be intertwined with political, social, or economic perspectives – both tendencies that were mitigated by the structured nature of the AIJF project and workshop. Several participants expressed a desire for even more of a systematic, specific, objective approach to this ongoing discussion, to the point of systems and interaction diagrams, information schemas, etc. Some participants seemed to desire a more technocratic or even technical approach to understanding and acting within an emerging AI-mediated information ecosystem.

The participants overwhelming valued a systematic, structured, thoughtful approach to discussing fundamental changes in news and journalism.

Finally, it was abundantly clear that many, or likely most, participants felt overwhelmed by the complexity of the information ecosystem (current and future), by AI, and by the pace of change at the intersection of AI and information. Many expressed a sense that AI was emerging into an existing environment that was already incredibly complex and uncertain, with enormously consequential issues around trust in information, changing audiences for information, business models for information producers, relationships with platforms, etc., mostly unresolved. AI, steamrolling into this already highly complex information ecosystem, now brings new uncertainties and complexities, not only in relation to the technology itself, but also to the very essence of information collection, processing, and consumption. There was a clear sense that the tectonic plates of information in society were moving, but that we are lacking the intellectual frameworks, foundational theories, and even vocabulary to help make sense of that change. Many participants expressed fatigue, exhaustion, or resignation at the burgeoning AI onslaught, and there was a shared sense of a lack of agency to shape a future information ecosystem dominated by AI.

Many participants shared a sense of being overwhelmed by the complexity of AI and its potential consequences for news and information.

6. What we missed

6.1. Feedback from the participants

In the weeks following the workshop we solicited feedback from participants via direct discussion with about 25 participants and also via a structured feedback form sent to all participants. The form asked participants to identify insights learned and missing perspectives, and to provide general comments, which yielded responses from an additional 10 participants. The objective of these requests was to gather thoughtful evaluations of the discussions that had benefited from several weeks of consideration by the participants following the workshop. We observed several persistent themes in the feedback that we received.

Participants were almost universally very positive about their experience and about the quality and usefulness of the discussions. The workshop was perceived as relatively unique by many and seemed to fill an unmet need for a frank and unflinching discussion of the fundamental challenges facing civic information at the largest and longest scale. A frequent observation was that the formal structure of the workshop, and the preparation before it, helped to ensure that this "big picture" approach generally remained focused and specific. Accompanying this appreciation for the structure of the workshop, however, was a frequently expressed desire for additional time for unstructured conversation following the event. Numerous responders suggested that they would have benefitted from a third day for open discussion.

Many reported that the small groups' presentations to the larger workshop audience did not accurately reflect the wide-ranging, creative, and courageous perspectives expressed in small-group discussions. A variation on this theme was that the small group discussions contained more "details and nuances which often did not make it into the larger group messaging." This corresponded with views expressed during the workshop, and commonly expressed elsewhere, that conversations about AI's potential impact on journalism are "more radical" in private than in public. Perhaps relatedly, we also received feedback from some participants expressing a realization that they "were not the only one" anticipating extreme changes in the information ecosystem and an appreciation for discussions in which it was clear that "there are people in important positions who are seriously concerned."

Participants generally thought that the sessions could have benefited substantially from the inclusion of two underrepresented groups: policy and regulatory experts, and technology company representatives. More broadly, there was a frequently stated sense that the participants were perhaps a little too homogenous for a truly wide-ranging discussion, especially in terms of their relationships with, and status within, the existing information ecosystem. One participant compared this to "asking a 15th century scribe to forecast the effects of movable type printing technology on their industry and society," suggesting that we might have seen different perspectives "if the modal member of the group had been a 20-something independent creator on TikTok, YouTube, Substack, or any of the other platforms that are part of the emerging 'new news media.'"

Feedback also revealed some key points of dispute among participants. One concerned broadening the focus of the workshop from journalism to the larger information ecosystem. Many participants thought that this was appropriate and "greatly appreciated the efforts to depart from the notion of journalism and think more broadly about civic communications." Others, however, felt that the discussions and outcomes might have been more manageable and relatable if they had been more narrowly focused on journalism.

Probably the most significant difference of opinion concerned the relationship between legacy journalism organizations and technology companies. Some participants expressed a desire for confrontation between civic society organizations and technology companies, with one stating that it was time to "sit at the table with policymakers and tech companies forcefully, even if it is to go for a rupture or a more hostile conversation," and another speaking in terms of "what we (i.e. legacy journalists) have to get from policymakers and tech companies." A more common viewpoint, on the other hand, was to accept technology companies as critically important players in the emerging AI-mediated information ecosystem and to engage with them in good faith. One participant bemoaned the tendency to "think of tech as inherently evil" and another suggested that the journalism industry should be "listening more to people on the outside, let go of prestige and legacy, and collaborate in order to steer towards a desirable future."

6.2. The perspective of the organizers

In light of the full experience of designing and operating the AIJF project, reading and listening to the views of approximately 1,000 people via the mini-scenarios, developed scenarios, and in workshop discussions, we have developed several conclusions that may be useful to a broader audience. We believe we have achieved the project's objectives, with the caveats discussed elsewhere in this report, and therefore the following conclusions are offered in the spirit of potential improvements.

First, focusing on journalism may be a barrier to imaginative and meaningful discussions about the potential effects of AI on the information ecosystem. We came to this realization following the launch of AIJF, with "journalism" already part of the project's name and identity. In hindsight we have found that use of this word may encourage too many pre-existing assumptions, legacy perspectives, and perhaps even some defensiveness regarding traditional journalism, a stance that increasingly represents a narrow perspective on how societies might inform themselves and how consumers might experience and influence their societies through information. Any plausible view of a future AI-mediated information ecosystem will likely encompass far more than is captured by the commonly understood definition of "journalism." This should be reflected in the design of initiatives similar to AIJF.

Like many of the participants, we feel that the application process and workshop would have benefited from more involvement by people outside of the journalism field, in particular start-up founders, representatives from platforms, and employees of other technology companies, who might have both broadened and grounded the discussion. We did attempt to recruit submissions from this group and engaged in many one-on-one conversations in an attempt to do so, but we were relatively unsuccessful. Our sense is that there may be a greater willingness among technologists to fundamentally re-imagine the information structures and processes of society in light of AI, and that we might have been able to simultaneously expand both the imaginativeness and the specificity of our discussions by including those perspectives. We also feel that discussions would have benefitted from the inclusion of younger or "Gen Z" participants engaged in AI or in the content creator ecosystem.

We believe that discussions were hampered by a lack of attention to what constitutes a socially beneficial information ecosystem and what its core objectives should be. While this reflects unavoidable uncertainty about what the

ecosystem might become, some specificity might help the existing journalism industry to engage in an intellectually honest conversation about its core objectives, beneficiaries, and methods before discussing potential scenarios for, and potential adaptation to, an AI-mediated information ecosystem. A considerable portion of the submitted material and workshop discussions failed to discuss these questions, instead focusing on what journalism fundamentally is – a discussion that could have beneficially taken place long before the advent of AI.

Finally, we believe that there should be more discussion and thinking about what is *not* going to change as we transition to an AI-mediated information ecosystem. Despite decades of academic research around journalism and communication we are still left without a fundamental understanding of the basic cognitive needs of human beings regarding information about events in our world and in our societies – a psychological framework describing why humans value journalism at all. Likewise we are also without reliable and applicable frameworks or theories for how information flows within societies, and are left instead with opinions, anecdotes, and folk theories. Initiatives like AIJF and other attempts to imagine an AI-dominated information ecosystem could be greatly aided by a better understanding about the enduring and unchangeable aspects of the relationship between human beings, their societies, and information.

6.3. Towards a better view of the emerging information ecosystem

We observed a strong desire among players in the current information ecosystem for a clear and succinct roadmap outlining exactly what AI will do to media over the coming years. Anyone hoping that this report would provide such a roadmap will probably be disappointed. We are clearly entering a period of significant and unavoidable uncertainty about journalism and civic information. It may be possible, however, to bring that uncertain future into slightly better focus, and maybe to discern it slightly sooner. AIJF represents an early attempt to do this, but there are other things we can do to help us to improve our collective foresight.

The most vital is to expand the conversation about the role AI in journalism beyond its use as a tool to optimize the status quo. Familiar tasks, workflows, and products are a natural and useful place to apply AI, to reap some early gains, and to learn. But we also need to fundamentally reimagine what we want our information environment to become in the presence of AI and to articulate a specific vision for what we consider to be a favorable outcome for ourselves, our

organizations, and our societies. This is probably now the primary concern of anyone tasked with navigating the intersection of AI and journalism.

In doing this we must avoid motivated reasoning. We need to take off any rose-colored glasses about the status quo and our role in it, and adopt an attitude of extreme intellectual honesty about the value we create and potential new opportunities for creating that value in novel and unfamiliar ways. We also need to do the work—to show up, with our homework done and a coherent list of specific goals. None of this will be easy, but it will be necessary.

In imagining a credible and coherent vision for an AI-mediated information ecosystem it is a safe bet to focus on serving audiences, which demands delivering genuine value to real people on their own terms. This advice should be familiar to anyone aware of the best practices for operating a newsroom in the era of social media, but it will be dramatically more applicable in the era of AI. Each of the five scenarios described in this report can be interpreted in terms of opportunities to serve audiences in new and better ways.

Some of the most inspiring and imaginative examples of radically new approaches to journalism and civic information are appearing in small newsrooms beyond the US, EU and UK. During the AI in Journalism Challenge—a four-month accelerator program operated by the Open Society Foundations in late 2023—we repeatedly observed small news organizations without significant resources, technical skills, nor experience with AI, imagine and deploy radically new interpretations of journalism and civic information through an innovative use of AI. These activities included instant contextualizing of news from knowledge bases, empowering local citizens to tell the stories of their own communities, and creating mechanisms for continual, systematic monitoring of entire news beats. These examples, and many others like them, show that using AI to improve journalism is not an impossible challenge.

The AIJF project and this report represent an early step towards better understanding the potential AI-mediated information ecosystem, but other steps will follow. There is additional detail and analysis to publish from AIJF, including work by academic researchers, articles about practical implications and the anonymized scenarios submitted by applicants. There are plans in development to extend this work by seeking input from the start-up and technology communities, as well as from "Gen Z" content producers and others. There will be further gatherings, including by those seeking to actively invest in support of positive outcomes from AI in our information ecosystem. And others beyond OSF are increasingly engaged in similar conversations and promise to undertake useful

research and initiatives. Some of these projects will likely include monitoring and evaluating the information ecosystem as AI becomes more influential, and identifying early examples of projects and organizations that demonstrate what journalism and civic information might eventually become in a world of ubiquitous AI.

It is natural to be apprehensive about the scale and pace of change that AI might bring. One workshop participant spoke about "a group tendency towards the dystopian end of the spectrum as challenges can feel so huge and insurmountable." Reaction to the scenarios produced by AIJF will likely be interpreted by many as primarily negative. But there is nothing assured about a negative outcome. At this early stage of the development and application of AI there is still only potential, and as much reason to anticipate positive outcomes for human societies as negative ones. Crucially, we all have some agency over how this plays out in the coming years. We are not spectators. We can engage and renew, imagine and build. We can seek to use these new tools to serve our audiences, our communities, and our societies, in new ways and perhaps with better outcomes.

"Optimists have agency."

—AIJF participant in conversation

Acknowledgements

The organizers would like to express their sincere gratitude to everyone who participated in this project at every stage, and in particular the more than 880 people who submitted “mini-scenarios” to AIJF in January and February 2024.

Workshop participants

The following people attended the AI in Journalism Futures workshop, held at the Castello di San Sebastiano in Piedmont, Italy, on April 15 and 16, 2024. We are deeply grateful for their enthusiasm, wisdom, and authentic engagement.

Peter Archer, *BBC*, UK
Tetiana Avdieieva, Digital Security Lab, Ukraine
Claudia Báez, *Cuestión Pública*, Colombia
Louis Barclay, 12challeng.es, UK
Emily Bell, Tow Center for Digital Journalism, Columbia University, USA
Mariano Blejman, Media Party, Argentina
Miguel Botero, Memria, Colombia
Premesh Chandran, Media Development Investment Fund, Malaysia
Madhav Chinnappa, independent consultant, UK
Gina Chua, *Semafor*, USA
Cori Crider, Foxglove, Spain
Tony Curzon Price, Humanity in Cyberspace, UK
Rawan Damen, Arab Reporters for Investigative Journalism, Jordan
Michael Del Nin, Soros Fund Management, USA
Anjanette Delgado, *Detroit Free Press*, USA
Nick Diakopoulos, Northwestern University, USA
Hubert Etienne, Sciences Po, France
Mary Fitzgerald, Open Society Foundations, UK
Yan Fu, data storyteller
Vashiegaran Manogaran, *Malaysiakini*, Malaysia
Cassiano Gobbet, Troovr, France
Jeremy Gordon, Collaborative Futures & Olli Home Health, USA
Molly Grönlund Müller, Schibsted, Sweden

Sebastián Hacher, Sembramedia & BeBot Design, Argentina
Natali Helberger, University of Amsterdam, Netherlands
Sonia Jalfin, Sociopublico, Argentina
Valer Kot, Media Development Investment Fund, Czechia
Avinash Kumar, Quicksand, India
Georgia Levenson Keohane, Soros Economic Development Fund, USA
Irene Jay Liu, International Fund for Public Interest Media, Singapore
Yolanda Ma, University of Hong Kong, Hong Kong
Harlan Mandel, Media Development Investment Fund, USA
Pedro Markun, LabHacker, Brazil
Paul Matzko, Cato Institute, USA
Orla McElroy, Gartner, UK
Mira Milosevic, Global Forum for Media Development, UK
Paula Miraglia, Media & Tech Task Force, Brazil
Conor Molumby, Reuters, UK
Michael Moss, Open Society Foundations, UK
Rasmus Nielsen, Reuters Institute for the Study of Journalism, Oxford, UK
Baybars Orsek, Logically AI, USA
Rishad Patel, Splice Media, Singapore
Pablo Perez De Angelis, Gezie.io, Argentina
Tharin Pillay, ALT Advisory, South Africa
Bonnibel Rambatan, *New Naratif*, Indonesia
Maria Ressa, *Rappler*, Philippines
Ross Settles, independent consultant, USA
Alan Soon, Splice Media, Singapore
Athan Stephanopoulos, AI advisor, USA
Sultan Suleimanov, *Meduza*, Latvia
TshepoTshabalala, JournalismAI, London School of Economics, UK
Susan Valentine, Open Society Foundations, UK
Ethan Van Diemen, *Daily Maverick*, South Africa
Bina Venkataraman, *The Washington Post*, USA
Marina Walker Guevara, *Pulitzer Center*, USA
Annie Jieping Zhang, Matters Lab, Taiwan
Patience Zirima, Gender and Media Connect, Zimbabwe

Facilitation and organizing team

Rachel Alltimes, Open Society Foundations, UK
Robert Bood, Fairsights, The Netherlands
David Caswell, Storyflow, UK
Francesca Edgerton, Open Society Foundations, UK
Shuwei Fang, Open Society Foundations, UK

Contributors

The report authors and organizing team wishes to thank the following people for their invaluable roles in the design and execution of the AIJF 2024:

Rachel Alltimes, Open Society Foundations, UK
Reem Almasri, Febrayer, Jordan
Robert Bood, Fairsights, The Netherlands
Jonathon Briggs, Open Society Foundations, UK
Gina Chua, *Semafor*, USA
Suzanne Clark, Open Society Foundations, UK
Nick Diakopoulos, Northwestern University School of Communication, USA
Francesca Edgerton, Open Society Foundations, UK
Mary Fitzgerald, Open Society Foundations, UK
Michael Moss, Open Society Foundations, UK
Mohamed Nanabhay, Mozilla Ventures, UK
Rishad Patel, Splice Media, Singapore
Luke Robinson, Open Society Foundations, UK
Felix Simon, Reuters Institute for the Study of Journalism, Oxford, UK
Alan Soon, Splice Media, Singapore
Mark Surman, Mozilla, Canada
Tshepo Tshabalala, JournalismAI, London School of Economics, UK
Susan Valentine, Open Society Foundations, UK
Marina Walker Guevara, Pulitzer Center, USA
Sarah Watson, Mozilla Foundation, USA

Authors

David Caswell was the lead consultant for the AIJF project. He is the founder of StoryFlow Ltd, an innovation consultancy focused on the application of AI to journalism. He was the executive product manager of BBC News Labs, and previously held senior roles leading AI and machine learning initiatives at Tribune Publishing, the *Los Angeles Times* and Yahoo! Caswell also publishes peer-reviewed research on computational, structured, and automated forms of journalism, including previously as a Fellow at the Missouri School of Journalism. He is a frequent speaker and writer about both strategic and applied opportunities and challenges for reinventing news for the emerging AI-mediated digital media ecosystem.

Shuwei Fang is Associate Director, Programs, at Open Society Foundations, working globally to advance open society values at the intersection of media and technology using grants, hybrid investments, and other strategies. Outside of Open Society Foundations, she was founder of two gaming startups and held roles in technology, media, and telecoms investment.