

Digital Agents: The Future of News Consumption

Expanded Scenario for Open Society Foundations

Tharin Pillay—April 2024

This piece was written as a contribution to the Open Society Foundations' AI in Journalism Futures Project. It speculates on how people will consume news in 5-10 years' time.

Written as if these predictions have already come true, it begins by describing an 'end state'—a vision of what information ecosystems will look like at a certain moment in time, and of how this will impact news consumption—before explaining how this end-state might arise. It closes by outlining some key assumptions that underlie the speculative scenario, and discussing to what extent we should expect these assumptions to hold.

The goal here is to paint a particular portrait of the future—to illustrate what may come to pass. I do not expect everything written below to happen (and certainly not to happen as written).

Advances in AI Systems

Foundation models—AI systems based on the 'transformer' architecture and trained using vast amounts of data and computing power—have improved steadily since they first came to public attention in 2022. These improvements were achieved primarily by scaling up the quality and quantity of data, and the amount of computing power, used to train these systems.

The most advanced of these models—frontier AI models—are fully multimodal: trained on and able to produce high-quality text, audio, and video content. They have a strong capacity to reason and—since they are connected to the internet, use various digital tools, and possess persistent memories—can take action in the world.

They are used in a wide range of tasks, including booking flights, designing buildings, diagnosing patients, and developing software. Even so, they're still not able to outperform the best humans in most professional domains—particularly those that require nuanced reasoning across long time horizons, such as executive leadership.

Digital Agents

These models have proved useful to enough people that they have become integrated into modern life. Most people think of them as 'digital agents', which serve as a combination of friend, adviser, personal assistant, employee, and companion.

The presence of such agents is ubiquitous across sectors and contexts, in the same way that the presence of the internet became ubiquitous decades earlier. In some cases, they have fully replaced human employees in the workplace; in others, while they have become integral to organisations' production processes, they are still closely supervised by humans. For example, they augment the work of many scientists and technical researchers, but are at present unable to replace them entirely.

Debates around the safety and sentience of these systems are ongoing and fierce, but this has done little to slow the rapid rate of adoption. American labs have led the way in AI advancements. Meanwhile, national and international governance efforts have largely failed to constrain the use and development of this technology (although the efficacy of regulatory regimes varies by region).

Personalised News Consumption

Most people have their own personalised agents, which they access through applications on their phone.

This is not the only way to interact with a digital agent: some hardware devices, such as wearable pins and glasses that enable agents to have real-time awareness of a person's physical environment, have been developed and adopted by those with the requisite capital and enthusiasm.

However they are accessed, these agents now intermediate almost all digital activity. For example, when searching for information, most people now speak to their agents rather than querying search engines. As the ability of these agents to fact-check themselves (and provide sources for their claims) has improved, 'AI hallucinations' have largely (but not entirely) become a problem of the past.

Markets for Models

Different organisations offer different kinds of agents: some are freely available—funded either by advertising (advertisers pay AI developers to have agents recommend their products) or by subsidies (as is common with companies backed by venture capital)—while others require a subscription. Open-source agents are also available for those with the technical savvy to set them up.

Beyond differences in how they are funded, digital agents vary greatly in their political and ideological leanings. There are liberal and conservative agents, socialist and libertarian agents, feminist and right-wing agents, and so on. Bias is a feature, not a bug. The baseline personalities of these agents can often be tweaked (depending on which company is offering them), and even the most politically 'neutral' agents sculpt themselves to fit individual preferences.

Most people consume news via their agents, who create tailored written, audio, and video content, drawing on news sources such as media companies and individual influencers. Most people favour short-form video content. The line between what does and does not constitute news has blurred almost beyond recognition. Even so, most measures of 'news' consumption reveal a steady decline, as people mostly prefer to spend their time engaging with personalised entertainment content or with the agents themselves, who are able to simulate human forms, and always available to chat via video.

The Media Landscape & Social Media's Decline

Meanwhile, the global media industry has been radically reshaped. Only a handful of outlets managed to weather the media sustainability crisis—a term which refers to media companies' struggles with financial sustainability, brought about because their primary funders, advertisers, realised that directing money to the technology companies that control content distribution platforms was much more efficient than paying media companies directly—and these outlets enjoy outsized influence.

Most small and mid-sized news media organisations have been forced to shut down. While it varies by country, broadcast news (radio and television) continues to be an important source of information, particularly for those who still do not have internet access.

Local news has seen a resurgence, as the dominant global outlets do not provide local coverage, and there is persistent demand for information that speaks to what is happening in people's local communities. Some of this is produced by traditional media organisations, who subsist on subscriptions and subsidies; but the majority comes from individual influencers, who are not bound by any media ethics codes, and who are often themselves advertiser-funded.

Most people no longer rely on social media for their news. As AI systems improved, people used them to flood social media platforms with AI-generated mis- and disinformation, propaganda, advertising, and other low-quality content; taking advantage of the fact that the platforms, designed to maximise metrics of attention to please advertisers, have little incentive to promote discourse that is factually accurate or grounded in reality.

Because they were unwilling to change their core business model, the platforms failed to effectively moderate content at scale, even while using their own AI agents and systems. It became near-impossible to discern truth from falsehood on social media; many major platforms saw an exodus of users as a result.

Advertisers soon turned to paying AI companies directly, and the power of social media platforms accordingly declined. The dream of the 'fediverse', wherein different social media platforms are interoperable and able to communicate with one another, never came to fruition, as the fundamental problem for all open platforms (they are vulnerable to influence operations undertaken by bad-faith actors deploying digital agents at scale) remains unsolved.

Fragmentation, Polarisation, Walled Gardens

So while some still use social media platforms for entertainment and communication, most people now communicate in 'walled gardens': digital spaces hosted on private servers that require permission to access. These spaces allow humans to be certain they're communicating with other humans, although they are still occasionally infiltrated by digital agents. They are organised around various themes, including geographic location, hobbies, and beliefs.

Most organisations still maintain a digital presence through websites, which provide portals through which to interact with organisations' own agents. These sites are in turn mostly accessed by digital agents, who summarise content for their users as required.

Because most people's interactions are either mediated by inherently-biased digital agents or occur in walled gardens, information ecosystems are highly fragmented—and different fragments are strongly polarised against each other. Conspiratorial thinking is common.

This has had dire consequences for democratic processes and civic engagement. Since people inhabit distinct realities shaped by their personalised agents and their walled gardens, finding common ground and engaging in constructive public discourse is harder than ever.

Trust in democratic institutions and processes has steadily declined, as consensus reality has diluted. People disagree on basic facts, and interpret events through wildly divergent lenses. Social cohesion has frayed and continues to fray, as communities and interest groups become more insular and resistant to outside perspectives. There is widespread global distrust and confusion, for example, in how digital agents interact with electoral processes.

In sum: news consumption has been transformed. The ever-improving capabilities of foundation models, and their integration into daily life as personalised digital agents, has

fundamentally altered how people access and engage with information and with each other. The decline of both traditional media companies and social media platforms, unable to cope with the flood of AI-generated content and the exodus of users and advertisers, has further fractured the media landscape. And the rise of walled gardens as safe havens for human-to-human interaction has furthered the fragmentation and polarisation of information ecosystems, even while it has created new space for nuanced discussion. News consumption is highly personalised, decentralised, and divided.

Underlying assumptions

Several assumptions underlie the above scenario. This final section briefly considers the plausibility of each one.

1. ‘AI systems will continue to improve as more money is invested in them’—This is assumed because, since the advent of the transformer architecture, increasing the amount and quality of data, and the amount of computing power used to train systems, has resulted in improvements (a phenomenon often referred to as scaling laws’. Most experts on the matter expect this trend to continue (see for example [this paper](#), which discusses a recent comprehensive survey of the industry), but it is unclear whether further advances in algorithmic architecture will be necessary for further improvements, and if so, on what timescale such advances might occur. So there is no guarantee that AI systems continue to progress at the current rate (although the fact that enormous amounts of money are being invested in ever-larger training runs suggests progress will continue, at least in the near-term).
2. ‘Adoption will be rapid and widespread’—This is assumed because AI systems are already widely proliferating throughout the world. ChatGPT continues to be one of the [fastest growing services ever](#), so it seems reasonable to assume that adoption will only increase. This might be mistaken though, if for example regulatory hurdles prevent widespread diffusion, or if the costs of running these AI systems prove to be prohibitive when passed on to consumers.
3. ‘Regulation will be largely impotent’—This is probably the shakiest assumption made in the above piece. It is assumed to keep things simple and to focus on news consumption. But the AI governance space is rapidly evolving—see recent policy developments such as President Biden’s [Executive Order on AI](#) and international instruments such as the [Bletchley Declaration](#)—and its quite possible that regulation has more of an impact than the scenario gives credit for. In general, regulation tends to lag behind technological development—but robust international collaboration and coordination may be able to constrain this technology before it is further integrated into our digital lives, as the scenario describes.
4. ‘Hallucinations will be less of a problem with AI systems’— This assumption is based on the expectation that ongoing research and development efforts will enhance the robustness, transparency, and self-monitoring capabilities of AI systems. However, the inherent complexity and unpredictability of advanced AI means that some level of hallucination risk will likely persist; and new types of hallucination may become possible as systems continue to improve.
5. ‘Social media will decline in prominence’—This is assumed for the reasons discussed above (saturation with AI-generated content and a failure to moderate or adapt business models leading to an exodus of users and advertisers), but it may not hold if social media companies turn out to be better able to adapt to the rise of AI than predicted, for example by solving content moderation problems or developing new

business models.

6. 'Private digital spaces will become popular'—This follows from the idea that if publicly-accessible platforms become uninhabitable, people will turn to private spaces. This may be overstating the extent to which people are willing to shift their online behavioural habits, though, and is also predicated on the assumption that social media platforms will decline in prominence, which is of course not guaranteed.

ENDS