STRENGTHENING INFORMATION INTEGRITY

Issue Brief 2: Advertising, Artificial Intelligence and the Global Information Crisis

April 2026

United Nations

United Nations Department of Global Communications

Introduction to the series

This series of issue briefs builds on the foundation of the *UN Global Principles for Information Integrity* (*UN Global Principles*) and related normative frameworks, including the Pact for the Future and the Global Digital Compact.

Each issue brief focuses on a specific, timely topic, drawing from research findings and expert insights. The series aims to open pathways for action, moving from principles to practice. This is the second issue brief in the series. Drawing from a growing body of policy and practice across diverse contexts, Issue Brief 1 outlines key concepts and challenges and offers solution frameworks and practical implementation through the 3R (Research, Risk, Response) model.

About the authors

The Information Integrity Team in United Nations Global Communications works at the intersection of policy, research and communications to uphold and strengthen information integrity, a key foundation for United Nations thematic priorities and operational mandates. The team's work contributes to emerging normative frameworks for information integrity, actionable initiatives, coalition-building, strategic communications and tools to help policymakers and practitioners find effective solutions to address evolving risks across information landscapes.

The Conscious Advertising Network (CAN) is a UK-based network of over 200 members taking bold steps to ensure effective advertising works for everyone. CAN works for advertising that inspires, innovates, and drives real results while respecting human rights. By tackling waste and fraud, promoting societal cohesion and environmental sustainability, and championing supply chain transparency, CAN unlocks new opportunities for growth and innovation.

TABLE OF CONTENTS

Terminology

For the purposes of this issue brief, the key terms below are used as follows:

Artificial intelligence (AI)
This paper discusses a range of AI technologies, including generative AI, AI recommender systems, agentic AI, large language models (LLMs), chatbots, etc. The paper does not delve into the technical differences of these uses of AI, but looks at the dynamics driving their adoption and impact on information integrity.

Agentic AI
Agentic AI is an artificial intelligence system that can accomplish a specific goal with limited supervision.[1]

Generative AI
AI systems that can create new content, such as text, images, or music, without direct human input.[2]

AI surface
The portion of a screen where advertisements appear within or alongside AI-generated content (e.g. chatbot responses, AI search summaries or AI-generated media).

Ad industry
A sector of organizations and companies creating and placing advertising or providing analysis of advertising. This includes advertising agencies, marketing departments within organizations, AdTech companies and trade bodies that seek to persuade the public to respond to products, services or causes.

Advertisers
Companies or individuals who purchase advertising, such as the marketing department of organizations or advertising firms. In this paper, this term excludes technology service providers, including AI companies and AdTech.

AdTech
The ecosystem of technology, tools and platforms that monetize online content.

Ad verification
Technology, including practices, software, tools and algorithms, which ensures ads appear on intended sites, reach targeted audiences and are viewed by real people.

Attention economy

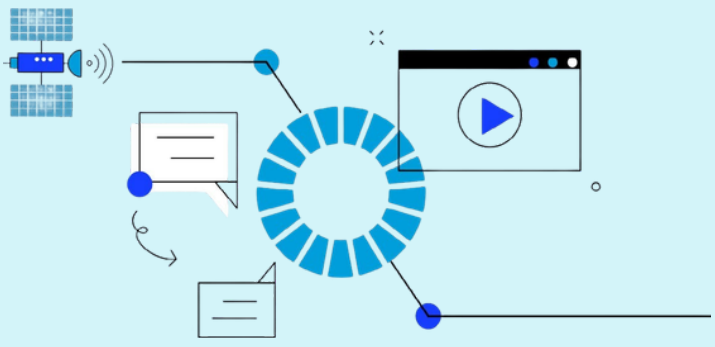
An approach to the management of information that treats human attention as a scarce commodity. To address the scarcity of people's attention, technologies have been increasingly aimed at strategic capture of private attention aided by systematic collection and analysis of personal data, which has become a profitable business model.[3]

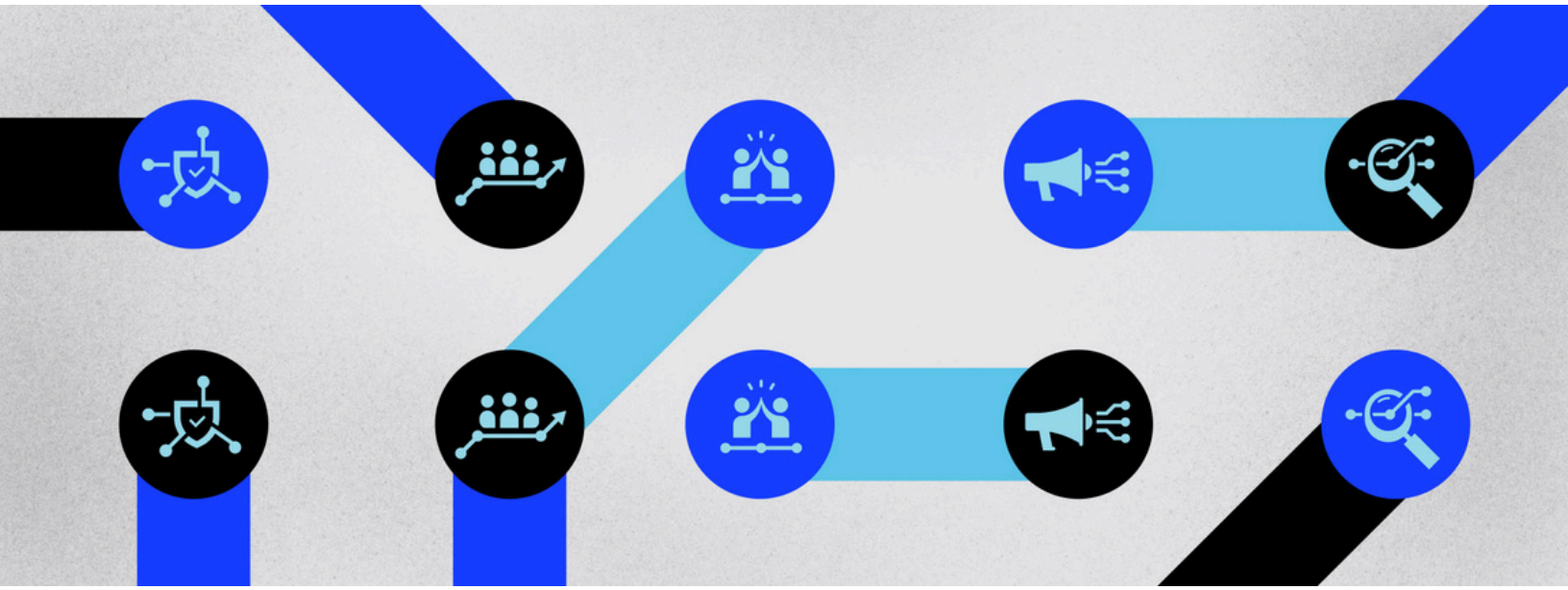
Brand safety

Practices and technology, including AI, that ensure advertising is placed in the appropriate context, with the intention to protect advertiser reputation.

Object-level transparency

Detailed data availability on specific ad placement and effectiveness, treating each placement as a discrete, auditable object. This level of transparency enables advertisers to reconcile ad spend directly to exact posts, feeds, creators and AI outputs that carried their ads.[4]





I. INTRODUCTION

The integrity of the global information ecosystem is at a crisis point, and the rapid adoption of Artificial Intelligence technologies (AI) is intensifying this deterioration at a pace that demands urgent attention. The UN Secretary General's High-Level Advisory Body on Artificial Intelligence, in its final report *Governing AI for Humanity* (2024), identified "damage to information integrity" as a top area of concern and risk.[5] The report warned that AI-generated disinformation threatens democratic institutions and processes, among other impacts, and called for preventative and mitigation measures to increase societal resilience. The first *UN Global Risk Report* (2025), reinforced this finding, ranking misinformation and disinformation as a top global vulnerability – a serious, already unfolding threat that the international community remains insufficiently prepared to address.[6]

Information Integrity
Information integrity refers to an information ecosystem in which reliable and accurate information is available to all, enabling people to engage meaningfully in public life, make informed decisions and exercise their rights. This ecosystem is shaped by the actions of a diverse range of actors, including governments, technology companies, media, civil society and individuals.

Strengthening information integrity means protecting the right to freedom of expression and access to information, ensuring inclusive access to a range of information sources and enabling people to navigate information spaces safely, with privacy and freedom. It involves building resilient societies that foster trust, knowledge and public empowerment. Challenges to information integrity encompass a spectrum of risks, such as disinformation, hate speech, restrictions on press freedom and the malicious use of technologies.

Yet, the transformation of the information ecosystem is occurring without adequate guardrails and without transparency about how AI tools work, what data is being collected, or what sources they draw on or exclude. AI is integrated into aspects of everyday life, often without users' knowledge or meaningful consent, consolidating control of the information ecosystem among the few large technology companies. People are increasingly relying on AI to make decisions and shape their understanding of the world without the literacy or information needed to assess its safety and reliability. These dynamics are compounded by the use of AI to create and distribute false and hateful content at scale, contributing to a broader erosion of trust in information sources.[7]

Advertising may seem a surprising entry point for addressing these challenges. Yet advertising is the dominant business model of the digital information ecosystem, funding all kinds of content, from pluralistic media and high-quality entertainment to hate speech and disinformation. It is the core revenue model of platforms where billions of people access news, connect with others and form their understanding of the world. These platforms are run by algorithms optimised to maximize user attention and advertising exposure. Content that keeps people engaged generates revenue, whether or not it is accurate, reliable or safe.

The *UN Global Principles for Information Integrity*, launched by the UN Secretary-General in June 2024, explicitly recognized advertisers as key stakeholders with the power to strengthen the information ecosystem. This issue brief builds on the *UN Global Principles* to create a resource for advertising industry leaders and policymakers, examining how AI in advertising can be approached in ways that uphold information integrity.

This paper argues that advertisers should view AI not only as a new technology to adopt, but as a force that is already reshaping the information environments on which their businesses depend. Strengthening information integrity is essential for commercial success.

II. ADVERTISING'S ROLE IN THE INFORMATION ECOSYSTEM

Advertising is the primary funding model of the digital information ecosystem and therefore shapes the financial incentives that determine what is produced, distributed and amplified online. This gives advertisers significant, though often unrecognized, influence over the health and integrity of information environments.

Digital advertising is powered by a complex network of intermediaries that facilitates the delivery of targeted online advertisements. When a brand purchases digital advertising, its money passes through multiple intermediaries – demand-side platforms, ad exchanges, supply-side platforms, and verification services – before the advertisement reaches a single person. At each step, fees are extracted and revenue is distributed to platforms, advertising intermediaries,[8] and to a lesser extent, content creators and publishers.

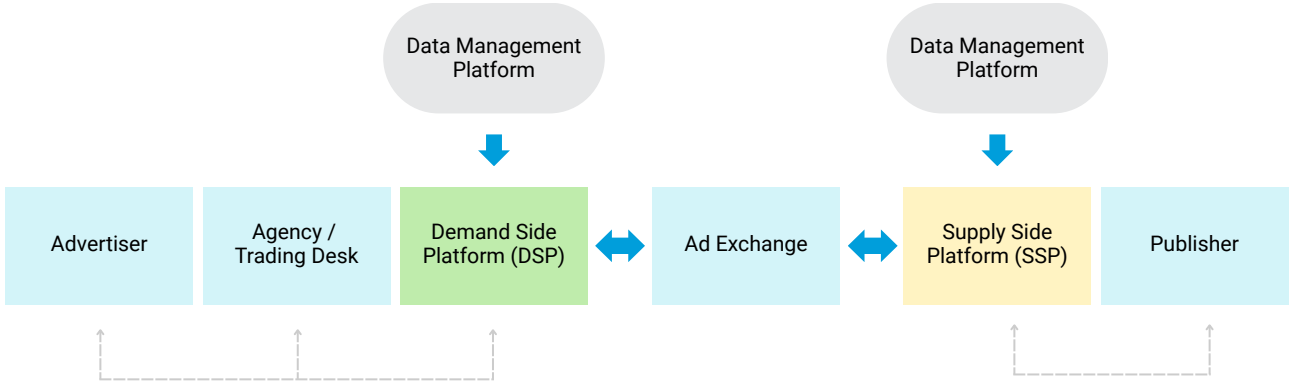


Figure 1. AdTech supply chain

Three characteristics define how the digital advertising supply chain functions: attention, scale and opacity.

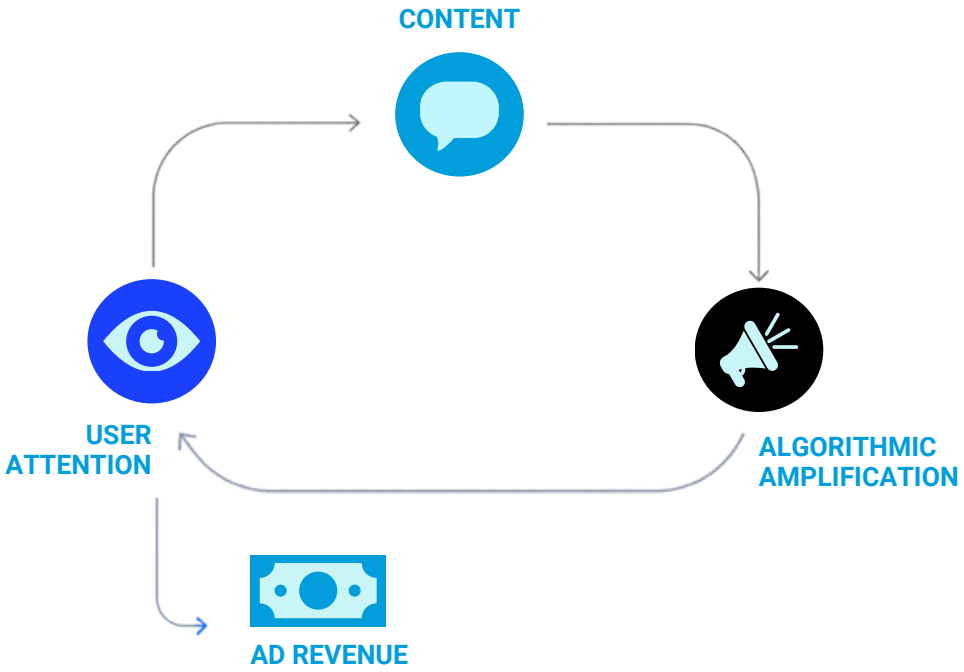


Figure 2. Attention Economy

Attention

The longer people look at their smartphone, television or other digital device, the more advertising can be shown to them, and the more revenue a digital platform generates. This dynamic has been described as the "attention economy," in which human attention is the commodity being bought and sold.[9] To maximize attention, digital platforms collect vast amounts of data from users, from their interests and relationships to behaviours and spending habits, to keep users engaged for as long as possible. This model is now being recreated and scaled out through AI assistants and chatbots.

Scale

The scale at which attention is monetized is substantial. Global advertising spend reached an estimated US$1.14 trillion in 2025. Advertising accounted for 75 per cent of Google's US$350 billion revenue in 2024, and 98.6 per cent of Meta's US$162.6 billion revenue in 2025. Advertising is, in effect, the core business model of the largest technology companies. Without it, their business models would not be viable.

Opacity

The trillion-dollar advertising ecosystem is largely opaque. A lack of transparency means that advertisers rarely have a complete view of where their advertisements appear or what they actually

cost, enabling waste and fraud.[11] Further integration of AI into the supply chain and decision-making processes risks deepening this opacity.

> Industry analysis indicates that approximately **16 to 17 per cent of programmatic ad transactions are flagged as fraudulent,** with invalid traffic accounting for roughly 8.5 per cent of ad impressions globally.[12] Independent analysis suggests that implementing transparency practices can deliver double-digit percentage improvements in return on ad spend.[13]

The consequences for the information ecosystem are significant. Without effective advertiser oversight, advertising revenue flows indiscriminately, funding content that is best able to attract and retain user attention regardless of quality or accuracy. Independent journalism and other reliable information sources struggle to compete with polarizing disinformation or misleading AI-generated content for readership and funding. Opacity reduces both advertisers' and audiences' capacity to make informed choices about the information they engage with and finance.

III. INCENTIVES SHAPING THE DIGITAL INFORMATION ECOSYSTEM

These three characteristics (attention, scale and opacity) serve as incentives to prioritize maximizing screen time and revenue over implementing guardrails and safe practices. While large platforms describe their approach to monetizing attention as "topic agnostic,"[14] this neutral-sounding framing obscures how AI recommendation systems function in practice. Content that triggers strong emotional reactions or is polarizing tends to generate higher engagement. AI recommendation algorithms are therefore designed to amplify such content, regardless of its veracity, reliability or safety. Platforms, in turn, are incentivized to maintain this practice, which underpins their profitability.

Evidence indicates that platforms frequently fail to moderate illegal content or content that violates their own standards and policies, in some cases prioritizing revenue retention over user safety.[15] The result is a reinforcing cycle: high engagement of extreme content drives amplification, which generates revenue, which in turn, creates demand for more of the same.

The implications for children and young people, who spend significant time in these digital spaces, are particularly serious. They face heightened exposure to risks, from the proliferation of AI-generated sexualized imagery depicting minors to algorithms directing users towards suicide-related content.[17]

The “topic agnostic” approach is a misnomer. In reality, this attention monetizing approach falls short of mitigating risks and, in a growing number of cases, has shown to heighten the risk of online harms, especially for vulnerable and marginalized communities.

THE GOVERNANCE GAP

The governance and regulatory landscape of AI technologies in advertising has not kept pace with the speed of adoption. Innovation without safety parameters exports risk to users and to the wider

public, while AI actors profit. Safety-by-design, embedded from the outset, is a precondition for durable and equitable adoption of any new technology with wide scale implications. The current "global governance deficit"[19] makes the *attention-scale-opacity* challenge harder to address effectively.

Regulatory fragmentation

Advertising governance and regulation remain fragmented across jurisdictions, consisting largely of a patchwork of measures addressing discrete elements of the ecosystem. Few States have regulated AI in advertising specifically, although several have introduced legislation mandating statutory labelling of AI-generated content, including advertising. While such efforts are notable, none to date confronts the central *attention-scale-opacity* challenge posed by the advertising business model, or its consequences for information integrity.[20]

Industry self-regulation of digital advertising and AI is similarly limited in scope, application and uptake. Voluntary standards have been developed to address ad fraud, brand safety, transparency, supply-chain traceability and sustainability.[21] These efforts show progress but are limited by their inconsistent uptake and limited accountability mechanisms. While the most forward-looking approaches to AI integration take into account a range of risks, including online harms, bias, data handling and workforce impacts, none yet extends to the holistic risk monitoring or systematic data sharing needed for comprehensive accountability.[22]

New industry standards are also emerging on the use of agentic AI in media buying but risk factors have not yet been fully addressed in these early efforts, including the poor quality of data available to AI agents, short-term corporate decision-making, opacity in agent decision-making, and loopholes that already affect media buying practices.[23]

Inequitable application

Many geographies and settings face particular challenges and communities in under-resourced contexts are disproportionately exposed to online risks. Large technology companies consistently under-invest in content moderation capacity in languages other than English. In many under-resourced contexts, ad verification infrastructure and related legal frameworks are also lacking, leaving policymakers and civil society with limited tools to intervene or to advocate effectively.

Inclusive and adaptive industry standards are therefore essential so that safeguards are applied equitably across jurisdictions and settings, and so that lower-income markets do not become testing grounds for AI systems considered too risky to deploy elsewhere.

IV. ARTIFICIAL INTELLIGENCE IN ADVERTISING: THE INDUSTRY TRAJECTORY

The advertising industry is adopting AI at a rapid pace, with significant implications for skills, efficiency, employment and broader society. A 2025 census by the UK's Institute of Practitioners in Advertising recorded a 14 per cent year-on-year drop in the number of employees at creative agencies, part of which was attributed directly to AI.[24]

Recent advancements have accelerated AI integration across marketing operations. Creative studios now use large-language and image models to generate videos, banner variants and copy. A 2025 survey of advertising executives found that 83 per cent said their company had deployed AI in the creative process, up from 60 per cent in 2024.[25] Widespread adoption of generative and agentic tools is also creating new digital spaces to be monetized through advertising, new advertising technology products and new market participants.

Although adoption has been high, many newer AI applications remain in initial testing phases. Questions persist over effectiveness, with a growing body of evidence challenging the assumption that AI-enabled personalized and targeted advertising delivers better results.[26] One recent study of advertising effectiveness found that linear television delivered the strongest returns, while significant investments in paid social media produced declining returns.[27] This research suggests that investments in news, community content and traditional linear television may better guarantee effectiveness.

Studies of advertising on AI chats and other AI surfaces are currently limited. Available research suggests that the effectiveness of AI-based advertisements and of AI advertising environments should be monitored carefully. User trust in AI surfaces is often found to be low, and, on that basis, likely to rank low in advertising effectiveness, owing to errors, bias in responses and broader societal concerns.[28]

RISKS TO INFORMATION INTEGRITY FROM AI IN ADVERTISING

The advertising industry is embracing AI rapidly, positioning it as the solution to staying competitive, reducing costs and increasing effectiveness. Yet AI adoption with no guardrails accelerates risks to information integrity, undermining the effectiveness of advertising and the functioning of the broader information ecosystem. Resolving this tension is the central challenge this issue brief seeks to address.

The United Nations Framework for Action

The *UN Global Principles for Information Integrity* provides the analytical and normative framework for this paper. The Principles are organized around five pillars, each of which has direct relevance to the AI-advertising nexus and carries recommendations for advertisers.

This paper explores the current state of AI in advertising in the framework of these five principles, identifying five top risks the industry and policy makers can address to strengthen information integrity:

UN Global Principles for Information Integrity	Risk to Advertisers
 Societal Trust and Resilience	The Investment Gap
 Healthy Incentives	Brand Safety
 Public Empowerment	Information Infrastructure Instability
 Independent, Free and Pluralistic Media	The Collapsing Open Web and Media Choice
 Transparency and Research	Systemic Opacity

Societal Trust and Resilience — The Investment Gap

The role of advertising in shaping AI tool design is an important and underexamined dimension of efforts to strengthen information integrity. Despite large capital inflows, technology companies have yet to establish profitable business models from generative AI,[29] creating pressure for monetization and accelerating product development. Some AI tools have incorporated design features considered addictive,[30] manipulative or harmful to users.[31] Advertisers who place their advertising within these environments are, whether intentionally or not, providing the revenue underwriting these design choices. Meaningful action in this area is an important step toward addressing the current trust deficit and for restoring advertising effectiveness.[32]

A reduction in advertising effectiveness is not the only consideration: the trust deficit contributes to broader conditions for social instability. The *World Social Report* (2025) warns that economic insecurity, inequality and declining trust are destabilizing societies worldwide, with the rapid spread of misinformation and disinformation identified as exacerbating these trends.[33]

These dynamics are not abstract for the advertising industry. Evidence from the International Monetary Fund shows that GDP remains depressed by around 1.5 per cent following major protests, and that strong institutions are needed to absorb and recover from such shocks.[34] An information ecosystem weakened by amplified disinformation and declining trust accelerates the conditions for unrest, increases its frequency and slows recovery. This reinforces a cycle of fragmentation and distrust that compounds the commercial risks advertisers face. Investing in information integrity is therefore not only a moral imperative but a prerequisite for the stable, prosperous societies in which advertising-dependent businesses can thrive.

Generative AI is not yet a sustainable business, with actors facing significant revenue shortfalls and investor pressure to recoup capital. While monetization approaches differ by company, some are choosing to host advertising.[35]

This dynamic creates interdependence. AI companies rely on advertiser revenue and corporate partnerships with advertising agencies or brands; while advertisers are seeking access to emerging audiences. This interdependence gives the advertising market significant leverage to influence AI product design. **In line with the *UN Global Principles*, advertisers can direct investment towards AI tools that demonstrably uphold information integrity, contributing to incentives for responsible innovation and toward a more resilient information ecosystem.**

Healthy Incentives — Brand Safety

The uses of AI in advertising include:

- **Brand safety:** AI tools, such as natural language processing (NLP)[36] or image recognition, to scan content before or during ad placement to ensure appropriate placement.
- **Video and image creation:** Ads created partially or fully by generative AI.
- **AI surfaces:** Advertising placed next to AI-generated content, such as images, text queries or music.
- **Ad tailoring, targeting and placement:** Deploying AI to generate hyper-personalized creative work (tailoring), predicting user behavior with precision (targeting) and automating media buying through autonomous agentic agents (placement).

Brand safety is one area where AI performance has fallen short of developers' initial claims. AI tools can miss nuance, sarcasm and high-risk AI-generated content, leading to advertisements being placed in polarizing or brand unsafe environments. Platform-provided brand-safety tools may also prioritize platform revenue, leading to continued unsuitable placements to sustain engagement.

High-risk AI-generated content is flooding digital platforms and the open web. On platforms, AI systems optimized for engagement have demonstrably favoured controversial or sensationalist AI-generated content, which drives higher engagement and thus higher ad viewing figures.[37] Meanwhile, brand-safety AI operates in direct tension with platforms' core revenue-optimizing algorithms. In effect, advertising revenue helps underwrite business models that amplify high-risk content, while high-quality media, independent journalism and community content — which evidence suggests deliver more durable returns — remain comparatively underfunded.

Advertisers face several additional reputational considerations related to the use of AI. Consumer acceptance of AI-generated advertising is currently low, and campaigns from major brands have faced consumer backlash and been withdrawn. The prevalence of errors and bias in generative AI outputs also raises questions about the reliability of AI surfaces from a brand-safety perspective. Gender, racial and socioeconomic bias in AI text, image and video generators can expose brands to additional risk, while also risking harm to groups in vulnerable or marginalized situations.

The *UN Global Principles* have clear recommendations for advertisers in this context. **Advertisers should demand transparency on digital advertising processes and call for published criteria to support human-rights responsible advertising placements and monetization. Given the scale of investment, advertisers should also vet the content and creators that they support and fund.**

Public Empowerment — Information Infrastructure Instability

Public empowerment initiatives offer an opportunity for advertisers, policymakers and civil society groups to find common ground in support of information spaces where people have genuine control over and understanding of their online experiences. Civil society has raised broad and serious concerns about AI's impact across a range of issues, including privacy, algorithmic bias, echo chambers, disinformation, mental health and employment. Public concern over the collection and use of personal data is particularly acute, heightened by high-profile cases of AI-enabled manipulation in which personal data has been exploited to target, deceive or influence individuals, with devastating consequences.

Advertisers can contribute meaningfully to advancing public empowerment. By supporting AI and digital literacy initiatives, adopting transparent data practices and avoiding discriminatory targeting, advertisers can help equip people with the capacity and tools to better navigate information spaces. Literacy initiatives that are accessible across age groups, demographics and languages enable people to understand how AI works, how their personal data is collected and used, and what rights they have online. This extends to literacy on how platform tools, advertising policies and data practices are communicated. Plain, accessible language supports public empowerment and efforts to restore trust in information environments.

With increasing investment in and reliance on AI systems, and no real option to opt out due to lack of consumer choice or control, the public also faces broader concerns over financial stability and the resilience of critical public information systems. Significant capital is flowing into AI development at a pace that many analysts have characterized as a speculative bubble, with valuations increasingly disconnected from demonstrated profitability. The information ecosystem and the advertising industry are being rewired around AI infrastructure that is physically fragile, environmentally resource-intensive and consolidates critical dependencies with a small number of dominant technology companies. The widespread integration of AI into critical information infrastructure before viable alternatives or adequate safeguards are in place creates structural dependencies that are difficult to reverse and potentially destabilizing for individual markets and for the global economy.[38]

> The United States alone accounts for approximately 45 per cent of the world's data centres, with Europe hosting a further 15 per cent, leaving most of the world significantly dependent on infrastructure controlled by a small number of dominant actors.[39]

The *UN Global Principles* call on the advertising industry to support technologies, platforms and publications that strengthen public empowerment, including support for literacy initiatives, transparent data practices and avoidance of discriminatory targeting. Advertisers can also support practices for the development of sustainable and more resilient information infrastructure for better long-term commercial and societal outcomes.

Independent, Free and Pluralistic Media — The Collapsing Open Web and Media Choice

The shift in advertising revenue from traditional media to digital platforms has weakened the financial viability of independent and public interest journalism, which are vital for information integrity. Public interest media outlets are competing for attention and algorithmic amplification with disinformation, influencer content, AI slop and entertainment.

The growing use of AI search summaries and generative AI tools in place of traditional search engines is placing further strain on online publishers and news outlets. Publishers derive advertising revenue in part from user views and clicks on advertising on web pages. Traditional search journeys saw users search for a topic, click through to a website and scroll through content, including advertisements, to find what they needed.

Most generative AI tools and AI-driven search summaries bypass the pages they summarize, compensating neither the publishers nor the creators of the content they process. Their use is substantially affecting visitor numbers across news and other websites. Candr Media Group has warned that AI search features could force a third of Independent Publishers Alliance members to close within the coming year.[40] The UK's Daily Mail Group has also warned of an 89 per cent decrease in visits to its sites.[41] The impact of these phenomena is particularly pronounced for outlets serving marginalized communities.

Many news outlets, particularly at the local level, have been unable to sustain operations, giving rise to so-called news deserts. The resulting information vacuum is often filled by AI-generated articles, reduced-quality journalistic content or content imitating news. The resulting reduction in content scope, quality and depth has implications for democratic governance, informed civic discourse, accountability and the protection of human rights. A recent survey found that 54 per cent of advertisers believe that generative AI has contributed to a decline in overall media quality.[42]

Some advertising practices contribute to these trends. Approaches to brand safety that focus on avoiding adjacency to unsuitable content often lead to restrictions on advertising placement in news — for example, on climate change coverage and on content serving marginalized groups.[43] AI-driven brand-safety technologies can also be calibrated in ways that further reduce advertising revenue to news outlets. Local news is under particular threat. The reduction of quality online content is, in turn, contributing to the growth of high-risk content and AI slop.

The rise of generative AI has also supported the emergence of AI-generated "made for advertising" sites — low-cost platforms designed to attract clicks and advertising revenue.[43] One notable example was the AI-generated news site Channel3Now, which falsely identified the perpetrator of the UK’s 2024 Southport attack as an asylum seeker. The false claim spread

rapidly online and preceded widespread riots. A UK parliamentary inquiry subsequently found that the unrest was triggered in part by disinformation amplified by algorithmic recommendation systems, and criticised social media companies for uneven responses.[46]

Governments have a responsibility to protect independent, free and pluralistic media.[47] Policymakers should address how AI companies can adequately compensate for the journalistic content on which their systems are trained and depend. **The *UN Global Principles* calls on advertisers to ensure transparency across their media supply chains and to make deliberate choices to fund independent, free and pluralistic media. Partnerships with journalistic institutions to develop sustainable funding models can help the media industry adapt to rapid technological change.**

Transparency & Research — Systemic Opacity

Despite advertiser calls for greater oversight, AI-driven advertising infrastructure remains opaque across many dimensions, from carbon reporting to media-buying decisions to ad libraries. The complexity and opacity of the digital advertising system already enables fraud and facilitates consumer protection problems such as scams.[48]

Agentic media buying, where AI systems autonomously purchase and optimize advertising placements, removes human judgement from content-distribution decisions. When algorithms bid on inventory in milliseconds without transparency, it becomes difficult for organizations to trace who is funding which content and why they are seeing it.

This represents a structural vulnerability. Opacity compounds when multiple autonomous systems interact. It becomes nearly impossible to trace how disinformation or scams propagate through the ecosystem. Where algorithms and AI systems curate content without disclosing commercial incentives or decision-making logic, the distinction between organic content and paid influence becomes difficult — or in some cases impossible — to establish.

Platform ad libraries, which are intended to provide transparency on who is advertising on platforms, what advertising looks like and who is targeted, are insufficiently transparent. They are inadequate for rigorous research, creating blind spots that undermine information integrity and slow efforts to identify harmful advertising such as scams.

Addressing transparency is essential for accountability. The technical infrastructure to deliver such transparency already exists. Digital systems track every micro-transaction with precision in the financial sector. What is missing is the regulatory and commercial incentive to apply equivalent standards in digital advertising.

Meta has publicly acknowledged that scam advertisements on its platforms have historically been responsible for around 3 to 4 per cent of revenue. Independent research indicates that this figure could be as high as 10 per cent, and Meta's own internal documents have suggested that its platforms expose users to an estimated 15 billion "higher risk" scam advertisements each day.[50] Verification of advertisers is a key step in reducing scams. The company reported having verified 55 per cent of advertisers in 2024 and estimated that verification rates would reach 70 per cent in 2025.

Implications for advertisers

The analysis above reflects an evolving commercial and regulatory environment in which the case for strengthening information integrity is reinforced by practical realities facing advertisers.

Regulatory developments across major jurisdictions are actively shaping commercial obligations. For advertisers and platforms operating in multiple markets, responsible deployment of AI and attention to information integrity are shifting from voluntary considerations to compliance requirements. Early alignment of internal practices with information integrity standards — through platform selection, vendor due diligence and supply-chain transparency — can reduce exposure to regulatory change and to the costs of reactive compliance, including remediation, legal exposure and operational disruption.

Capital markets are also increasingly treating information integrity as a material factor in corporate disclosure and risk assessment, rather than a reputational one. Shareholder resolutions addressing platform accountability and information integrity are gaining traction, reflecting investor attention to the financial implications of information risks.[51]

V. RECOMMENDATIONS FOR ADVERTISERS

> *'Advertisers can benefit the information ecosystem in a way that both strengthens information integrity and makes good business sense. While technology companies are unlikely to readily abandon current business models, healthier incentives can be achieved through greater transparency for advertisers into advertising processes and the adherence to human-rights responsible advertising policies by advert deliverers. By gaining more control of a transparent supply chain, advertisers can also see a better return on their investment.'*
>
> *UN Global Principles for Information Integrity* (2024)

Embed Transparent, Human Rights-Responsible Advertising

Advertisers should use the current window of opportunity to make information integrity a condition of AI uptake, prioritizing, where possible, tools developed with clear guardrails and safety measures at the design and development stage rather than retrofitted after the fact.

Measures include:

- Require AI and AdTech companies to adopt transparency standards that enable end-to-end validation of the AdTech supply chain, including object-level transparency, so that advertisers can make informed decisions about cost and effectiveness.
- Require Agentic AI companies to make the decision-making behind media buys transparent and auditable.
- Require AdTech companies to carry out independent third-party audits and vetting of ad exchange supply partners.

- Require full disclosure of any AI-facilitated or AI-generated monetized output, including:
  - the output itself, whether text, image, sound or video, and whether it was monetized and by which advertisers;
  - whether paid content influenced ranking, summarization or generation, as well as the provenance of the content: who or what created it, whether it was altered, and whether it met the platform's own safety thresholds;
  - contextual snapshots taken at the moment an advertisement was served, showing the exact AI-generated content it appeared alongside, while protecting user privacy.
- Require platforms to maintain clear audit trails that enable brands to reconstruct incidents quickly, including the full decision trace for each transaction.

The Conscious Advertising Network's Guiding Principles are an example of an industry and civil society partnership that seeks to align advertising with human rights.

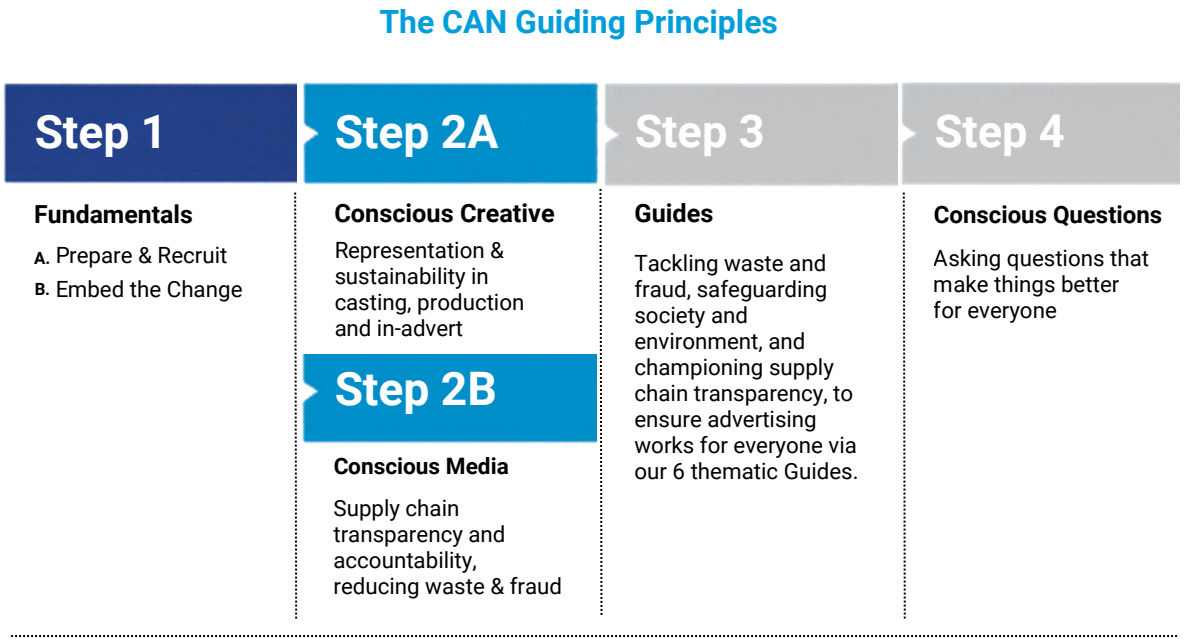


Figure 3. The Conscious Advertising Network's Guiding Principles

Advertise with Media Outlets, Creators, AI Tools and Platforms that Bolster Information Integrity

Advertisers should make information integrity a core component of media placement strategies and support measures that ensure quality content creators and publishers are adequately remunerated by AI companies for the content on which their systems depend.

Measures include:

- Prioritize media outlets, AI tools and platforms that demonstrate robust information integrity safeguards, including public interest journalism and AI systems that use independent verification methods, such as ad verification tools and manual vetting.
- Assess and exclude placements that present risks to information integrity, including made-for-advertising sites, unverified generative-AI interfaces and programmatic placements with documented fraud patterns.
- Establish dedicated funds to support journalism and creators, including in developing countries, separate from standard media buying, including multi-year commitments to enable planning and direct partnerships.

Support Policy that Strengthens Information Integrity

Advertisers should support accountability mechanisms that make information integrity measurable, auditable and enforceable across advertising services, and work with civil society and other AI stakeholders to support policy that drives transparency and accountability through the advertising supply chain.

Measures include:

- Independently verify algorithmic systems for bias and discrimination while protecting privacy.
- Ensure adequate compensation of creators and publishers.
- Require disclosure of AI-generated media and machine-readable labeling standards that identify the underlying model and responsible party, with mandatory human-in-the loop for high risk categories.
- Require AI vendors to disclose environmental impact data, including information on infrastructure stability, as part of procurement due diligence.
- Invest in training programmes that equip workers with the skills to use AI tools in ways that uphold information integrity, embedding human judgement at key moments.

Case Study:

The Belém Declaration on Information Integrity on Climate Change is a result of the **Global Initiative for Information Integrity on Climate Change**, a multi-stakeholder coalition co-led by the Government of Brazil, the UN and UNESCO. The Initiative helped achieve historic integration of information integrity into international climate governance at COP30 — including by formally embedding it in the COP outcome, the Mutirão Decision. Launched at COP30 and initially signed by more than 20 countries, the Declaration established shared international commitments to address climate disinformation and promote accurate, evidence-based information on climate change.

The Declaration makes explicit calls on the private sector to commit to the integrity of information on climate change in their business practices, in line with the *UN Guiding Principles on Business and Human Rights* (2011), and to ensure transparent, human rights-responsible advertising practices that bolster climate information integrity and support reliable information and journalism. The Charter for Accountable Climate Advertising, being developed under the Global Initiative, provides a delivery mechanism for these commitments.

For advertisers, the Declaration represents both a normative framework and a practical commitment mechanism — a signal from 20+ governments that transparent, human rights-respecting advertising practices are now a matter of international policy.

VI. RECOMMENDATIONS FOR POLICYMAKERS

Embed Transparent, Human Rights-Responsible Advertising

Align governance and policy on AI and advertising with the *UN Global Principles for Information Integrity*.

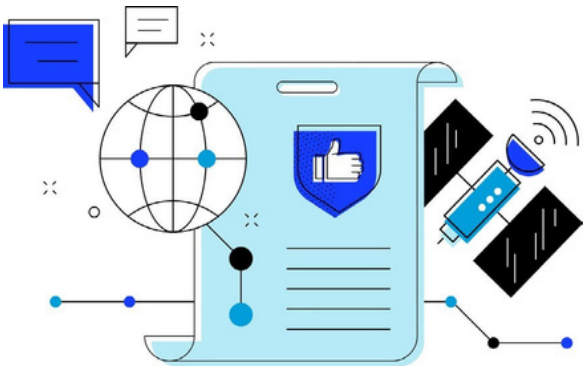
- Undertake multi-stakeholder and cross-border collaborations to raise awareness of the role of advertising in shaping the incentives that drive the information ecosystem, including in the context of safe, transparent and trustworthy AI.
- Support the UN framework for action by publicly endorsing and implementing the *UN Global Principles for Information Integrity*.
- Support the Belém Declaration on Information Integrity on Climate Change and the Global Initiative for Information Integrity on Climate Change as vehicles for translating commitments into practice for climate action.
- Harness existing industry standards to develop clear policies that minimize risks to information integrity and support conditions for brand safety.

Drive Transparency and Strengthen Accountability Across the Digital Advertising Ecosystem

Policymakers can collaborate with civil society and the advertising industry to establish measures that promote transparency and accountability across the digital advertising supply chain, including:

- Enforceable object-level transparency standards for advertising placement, including data minimization practices that uphold the right to privacy and full disclosure of advertising campaign data, ensuring that every monetized impression is auditable.[54]

- Independent, third-party algorithmic impact assessments[55] for large-scale AdTech systems and generative-AI models used in advertising contexts, evaluating risks to information integrity, with publicly disclosed results.
- AI-generated media disclosure and machine-readable labelling standards, with human-in-the loop for advertising creative and media placement decision-making.
- Accountability standards comparable to those applied in other large markets, with compliance mechanisms that are robust, independently verifiable and consistently applied, including searchable public registries for AI advertising systems to track model versions, training-data provenance, known biases and compliance status.
- Publicly accessible and searchable ad libraries on AI surfaces.



VII. A Practical Approach: Adopt the United Nations 3R Approach to Information Integrity Risk Identification & Management

The UN 3R framework — Research, Risk, Response — provides a sophisticated yet accessible structure to the identification, management and evaluation of information integrity risk through which the advertising industry can operationalize the analysis set out in this brief.[56]

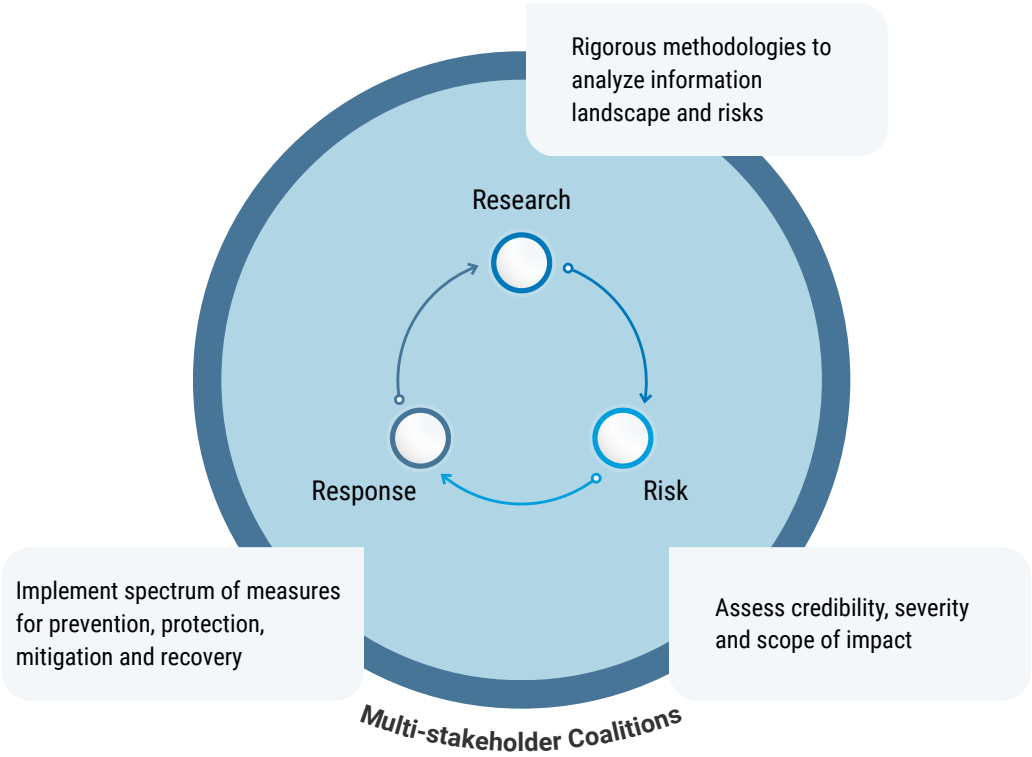


Figure 4. Research, Risk assessment, Response Approach. See Issue Brief 1.

Research goes beyond brand safety. The impacts of AI on information integrity require monitoring and managing across a wider set of exposures, drawing on the multi-stakeholder approaches and the expertise of civil society partners alongside commercial mechanisms already available to the industry.

Spheres of information integrity risks to be monitored by the advertising community:

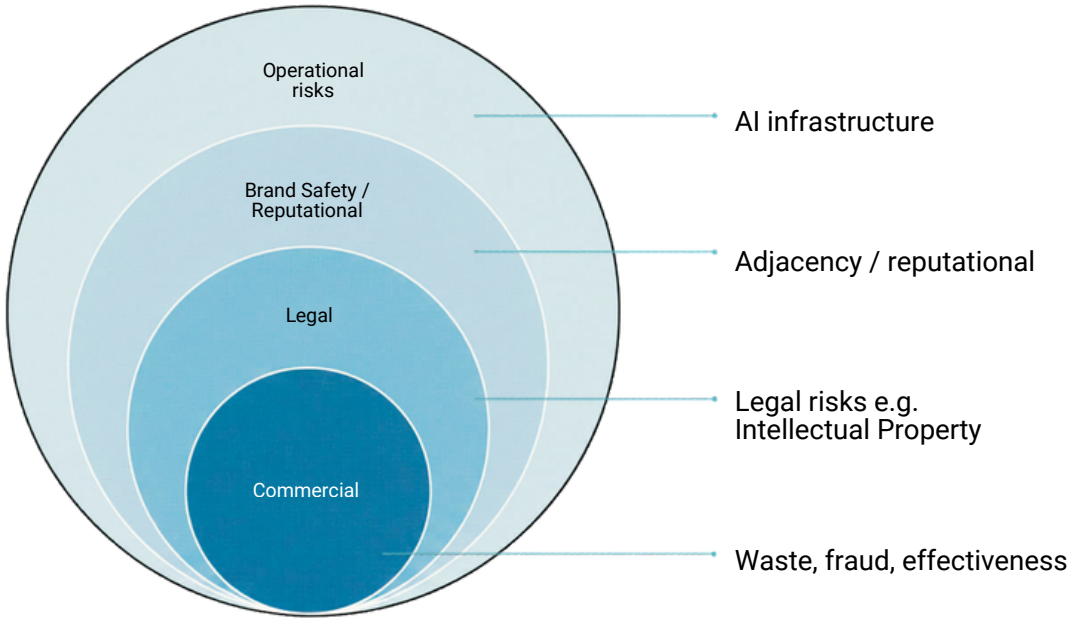


Figure 5. Spheres of information integrity risks

Risk management for information integrity warrants board or C-suite oversight, alongside applicable corporate responsibility and due diligence frameworks. Given the emergent nature of these technologies, advertisers should take a proactive approach to identifying practices likely to generate risk before they materialize. Publishing risk assessments and findings publicly and directly engaging with civil society strengthens the collective evidence base and grounds risk mitigation in independent verification rather than industry consensus alone.

Response covers a life cycle approach to risk management — prevention, protection, mitigation and recovery. Advertisers' decisions about where and how to spend can be informed by multi-stakeholder analysis and directed towards tools and platforms that demonstrably uphold information integrity.

For additional information on the 3R framework, see Issue Brief 1.

ENDNOTES

1. IBM. (n.d.). What is agentic AI? IBM Think. https://www.ibm.com/think/topics/agentic-ai
2. UN Innovation Network. (n.d.). Generative AI. United Nations. https://www.uninnovation.network/un-group-pages/generative-ai
3. United Nations Economist Network. (2021). The attention economy. United Nations. https://www.un.org/sites/un2.un.org/files/attention_economy_feb.pdf
4. Conscious Advertising Network. (2026). CAN Policy Platform 2026. https://www.consciousadnetwork.com/wp-content/uploads/2026/04/CAN-Policy-Platform-2026.pdf
5. United Nations Artificial Intelligence Advisory Body. (2024). Governing AI for Humanity: Final Report. United Nations. https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf
6. United Nations. (2025). UN Global Risk Report 2025. https://unglobalriskreport.org
7. United Nations. (2025). Information integrity in the AI age. We the Peoples. https://medium.com/we-the-peoples/information-integrity-in-the-ai-age-51494645e765
8. Advertising intermediaries include data management platforms (DMPs), which gather, sort and house information for use by marketers, publishers and other businesses; demand-side platforms (DSPs), software used by advertisers and agencies to automate the purchase of digital ad inventory; and supply-side platforms (SSPs), software used by publishers to automate the management, selling, and optimization of their digital ad inventory.
9. Centre for Humane Technology. (2024). The attention economy: Why do tech companies fight for our attention? https://www.humanetech.com/youth/the-attention-economy
10. WPP Media. (2025). Global advertising forecast: Global ad revenue set to reach $1.14 trillion. https://www.wppmedia.com/news/report-this-year-next-year-december-2025
11. Check My Ads. (2026). Restoring fairness for Europe's news publishers. https://checkmyads.org/report/check-my-ads-submits-paper-to-forum-on-information-and-democracy-outlining-solutions-for-a-fair-digital-ad-market/
12. AdWeek. (2025). Why ad fraud rates haven't budged in 15 years. https://www.adweek.com/programmatic/why-ad-fraud-rates-havent-budged-in-15-years/
13. Ebiquity. (2025). Media Transparency Report 2025. https://ebiquity.com/guides/media-agency-msa-transparency-report-2025/
14. BBC. (2022). Molly Russell: Instagram posts seen by teen were safe, Meta says. https://www.bbc.com/news/uk-england-london-63034300
15. Reuters. (2025, November 6). Meta is earning a fortune on a deluge of fraudulent ads, documents show. https://www.reuters.com/investigations/meta-is-earning-fortune-deluge-fraudulent-ads-documents-show-2025-11-06
16. UNICEF. (2026). Deepfake abuse is abuse: Statement by UNICEF on AI-generated sexualised images of children. https://www.unicef.org/press-releases/deepfake-abuse-is-abuse
17. United States Senate Committee on the Judiciary. (2025, September 16). Examining the harm of AI chatbots. https://www.judiciary.senate.gov/committee-activity/hearings/examining-the-harm-of-ai-chatbots
18. Knight-Georgetown Institute. (2026, March). What US lawsuits reveal about platform design that DSA reports don't. https://kgi.georgetown.edu/research-and-commentary/what-us-lawsuits-reveal-about-platform-design-that-dsa-reports-dont

19. United Nations AI Advisory Body. (2025). Governing AI for humanity: Final Report. United Nations. https://www.un.org/en/ai-advisory-body
20. UK Parliament. (2026). Social media, misinformation and harmful algorithms: Government response. https://publications.parliament.uk/pa/cm5901/cmselect/cmsctech/1397/report.html
21. Advertising Association. (2026). Best practice guide for the responsible use of generative AI in advertising. https://adassoc.org.uk/our-work/best-practice-guide-for-the-responsible-use-of-generative-ai-in-advertising
22. ISBA. (2025). Responsible Media Guide — Artificial Intelligence. https://www.isba.org.uk/knowledge/responsible-media-guide-artificial-intelligence
23. The advertising industry is experimenting with agentic decision-making, where agentic systems plan, execute, and optimise media across complex workflows. The intent is that humans set objectives, guardrails and risk tolerance, while agentic systems operate continuously within those boundaries. Critics have raised concerns including unpredictable actions, bias, and accountability gaps when systems operate independently.
24. IPA. (2025). Agency census 2025 shows workforce declines while diversity improves. https://ipa.co.uk/news/agency-census-2025
25. IAB. (2025). The AI gap widens. https://www.iab.com/insights/the-ai-gap-widens
26. Check My Ads. (2024). Google lets brands audit some GSP ads after Check My Ads pressure campaign. https://checkmyads.org/impact/google-gsp-pmax-transparency; Reuters. (2025, November 6). Meta is earning a fortune on a deluge of fraudulent ads, documents show. https://www.reuters.com/investigations/meta-is-earning-fortune-deluge-fraudulent-ads-documents-show-2025-11-06
27. Ebiquity, EssenceMediacom, Gain Theory, Mindshare, & Wavemaker UK. (2024). Profit ability 2: The new business case for advertising. Thinkbox. https://www.thinkbox.tv/research/thinkbox-research/profit-ability-2-the-new-business-case-for-advertising
28. Republique Francais. (2025). How do the French relate to artificial intelligence? https://labo.societenumerique.gouv.fr/en/articles/what-relationships-do-the-french-have-with-artificial-intelligence; Gallup. (2025). Reward, risk, and regulation: American attitudes toward artificial intelligence. https://www.gallup.com/analytics/695033/american-ai-attitudes.aspx
29. GIS Reports Online. (2026, February). The trillion-dollar question for AI business models. https://www.gisreportsonline.com/r/ai-business-models
30. MIT Technology Review. (2024, August). We need to prepare for 'addictive intelligence'. https://www.technologyreview.com/2024/08/05/1095600/we-need-to-prepare-for-addictive-intelligence
31. United States Senate Committee on the Judiciary. (2025, September 16). Examining the harm of AI chatbots. https://www.judiciary.senate.gov/committee-activity/hearings/examining-the-harm-of-ai-chatbots
32. World Federation of Advertisers. (2021). WFA launches the state of advertising report. https://wfanet.org/knowledge/item/2019/06/18/WFA-launches-The-State-of-Advertising-report
33. United Nations. (2025). World Social Report 2025. https://desapublications.un.org/publications/world-social-report-2025-new-policy-consensus-accelerate-social-progress
34. International Monetary Fund. (2021). The Macroeconomic Impact of Social Unrest. https://www.imf.org/en/publications/wp/issues/2021/05/07/the-macroeconomic-impact-of-social-unrest-50338
35. OpenAI. (2026). Our approach to advertising and expanding access to ChatGPT. https://openai.com/index/our-approach-to-advertising-and-expanding-access/
36. Natural Language Processing (NLP) combines computational linguistics, machine learning, and deep learning to analyze text and speech, allowing machines to perform tasks such as sentiment analysis, machine translation, and intent recognition.
37. Knight First Amendment Institute at Columbia University. (2023, September). The myth of the algorithm: A system-level view of algorithmic amplification. https://knightcolumbia.org/content/the-myth-of-the-algorithm-a-system-level-view-of-algorithmic-amplification

38. UN Environment Programme. (2025). AI has an environmental problem. Here's what the world can do about it. https://www.unep.org/news-and-stories/story/ai-has-environmental-problem-heres-what-world-can-do-about
39. Statista. (2026, April). Data centers worldwide by country. https://www.statista.com/statistics/1228433/data-centers-worldwide-by-country
40. The Media Leader. (2025, September). One-third of indie publishers could shut down by next year as AI search hits traffic. https://uk.themedialeader.com/one-third-of-indie-publishers-could-shut-down-by-next-year-as-ai-search-hits-traffic
41. Press Gazette. (2026, January 16). Five US publishers sue Google over deceptive and manipulative adtech practices. https://pressgazette.co.uk/marketing/five-us-publishers-sue-google-over-deceptive-and-manipulative-adtech-practices
42. DoubleVerify. (2024). Global Insights Report 2024. https://doubleverify.com/hubfs/46126064/content/DV_Report_GlobalInsights_2024_Global.pdf
43. See Stagwell for research on brand-safety tools and news de-funding: https://www.stagwellglobal.com/future-of-news
44. NewsGuard. (2025, February). Data shows advertising on news brings major benefits to ad performance — so why don't brands advertise on news? https://www.newsguardtech.com/insights/data-shows-advertising-on-news-brings-major-benefits-to-ad-performance-so-why-dont-brands-advertise-on-news/
45. MIT Technology Review. (2023, June). Junk websites filled with AI-generated text are pulling in money from programmatic ads. https://www.technologyreview.com/2023/06/26/1075504/junk-websites-filled-with-ai-generated-text-are-pulling-in-money-from-programmatic-ads
46. UK Parliament. (2024). Inquiry into social media, misinformation and harmful algorithms https://committees.parliament.uk/work/8641/social-media-misinformation-and-harmful-algorithms/news
47. United Nations Human Rights Committee. (2011, September 12). General comment No. 34: Article 19 — Freedoms of opinion and expression (CCPR/C/GC/34). United Nations. https://docs.un.org/en/CCPR/C/GC/34
48. JD Supra. (2026). Advertising law compliance in 2026: Five developments every advertiser should know. https://www.jdsupra.com/legalnews/advertising-law-compliance-in-2026-five-1601317; Federal Trade Commission. (2026). FTC announces crackdown on deceptive AI claims and schemes. https://www.ftc.gov/news-events/news/press-releases/2024/09/ftc-announces-crackdown-deceptive-ai-claims-schemes
49. The Media Leader. (2026). Meta admits revenue from fraud and scam ads might have accounted for 3–4 per cent of total revenue. https://uk.themedialeader.com/meta-admits-revenue-from-fraud-and-scam-ads-might-have-accounted-for-3-4-of-total-revenue
50. OpenMic. (2024). 2024 shareholder votes send strong message to Meta and Alphabet regarding artificial intelligence. https://www.openmic.org/news/2024/6/2024-shareholder-votes-send-strong-message-to-meta-and-alphabet-regarding-artificial-intelligence
51. Conscious Advertising Network. (2025). Guiding Principles. https://www.consciousadnetwork.com/can-guiding-principles/
52. See The Global initiative for Information Integrity on Climate Change. https://www.unesco.org/en/information-integrity-climate-change
53. United Nations Office of the High Commissioner for Human Rights. (2021). International Standards: Privacy in the Digital Age. https://www.ohchr.org/en/privacy-in-the-digital-age/international-standards
54. An algorithmic impact assessment is a process designed to help assess and mitigate the impacts associated with deploying an automated decision system.
55. United Nations. (2025, September). From Principles to Practice: Strengthening Information Integrity - Issue brief 1. https://www.un.org/sites/un2.un.org/files/un-information-integrity-issue-brief-01.pdf