

Responsible AI Principles for the Media Sector

The RMF's **Responsible Artificial Intelligence (AI) Principles** outline the key areas that the media sector should consider when developing and implementing AI. They consider AI technology which is commercially implemented at the time of writing (2026), including, for example, generative and agentic AI. They are intended as a practical framework to support media companies to apply AI responsibly, rather than formal commitments. They cover both operational uses of AI and AI-assisted content creation.

The principles were developed in collaboration with RMF participants through the Responsible AI working group, which convened throughout 2026. These principles reflect the perspectives and practices of a diverse set of media sub-sectors represented within the RMF. As a result, some principles may be more relevant at an individual company level, rather than across the sector.

Definitions¹:

- Artificial Intelligence (AI): The field of computing that enables machines to mimic human intelligence, including learning, reasoning and problem solving.
- Generative AI (genAI): AI approaches that create new content, such as text, images or programming code, based on patterns learning from training data.
- Agentic AI: An agent is an AI system that can independently reason, plan, adapt and pursue objectives or goals without direct human intervention over an extended time horizon and interact with tools.

AI is used as a catch-all phrase in the below principles unless further specified.

1. We are committed to transparency and explaining where AI is used

What it means

- Be transparent about how and where AI-assisted outputs and processes are used, depending on the use case. For example, labelling content accurately to audiences.
- Explain this through clear communication to relevant stakeholders, including employees, audiences and the wider public. Ensure stakeholders understand why AI is used and how it may affect them.

Why it matters

- Builds trust amongst audiences, the sector's stakeholders and in the AI solutions that are used.
- A lack of transparency and explainability can lead to lower trust, conceal accountability and lead to uncertainty about how AI influences content that users engage with.

Caveats and considerations

- Transparency is an important aspect of the EU AI Act.

¹ Ray Eitel-Porter, Paul Dongha and Miriam Vogel, *Governing the Machine: How to Navigate the Risks of AI and Unlock Its True Potential*, Bloomsbury Publishing, 2025

- Transparency expectations depend on the use-case. For example, transparency over audience-facing content should have higher priority than operational uses or content prioritisation and targeting, as there are higher risks to trust.
- There may be intellectual property or commercial sensitivity implications.
- Many advanced AI models are complex, making it difficult to provide clear yet technically accurate explanations to non-experts. Still, companies should explain their use of AI models to their audiences and customers.

2. We are committed to understanding and responsibly managing the impact of AI on our colleagues, society, and the environment

What it means

- Our colleagues: Provide organisation-wide and role-specific training to support employees to productively, safely and beneficially engage with AI tools, and create awareness of the wider impacts of AI.
- Society: Promote AI literacy and transparency amongst our audiences and customers, to build and maintain trust in the tools that we use and their outputs.
- Environment: Develop a better understanding of the usage of AI in our organisations as a first step. As a second step, work towards measuring the environmental impacts of AI, with energy as the priority, by engaging with AI system providers to improve disclosure of environmental impacts, assumptions and methodologies. At the same time, identify opportunities for efficiencies in AI use to reduce impacts.

Why it matters

- Our colleagues: Create awareness amongst colleagues, preserve human judgement and accountability whilst enhancing productivity and innovation.
- Society: AI literacy creates widespread understanding and builds trust amongst stakeholders, to ensure AI is used successfully, with proper due diligence and provides trusted outputs.
- Environment: Identify opportunities to mitigate and manage negative environmental impacts resulting from AI use.

Caveats and considerations

- Our colleagues: Role-specific training will vary across organisations for content creation, depending on if and what type of content is produced.
- Society: The type of AI-generated or assisted content will vary between sub-sectors. It is important to be able to explain in what types of content AI is used, and if applicable, which parts of the production process.
- Environment: There is a lack of consistent data and methodologies to measure and compare the environmental impact of AI within the industry. Limited disclosures from AI providers means it can be difficult to collect accurate data related to the environmental impacts of AI at an individual company level.

3. We ensure our use of AI is fair and accurate

What it means

- Actively recognise and understand the bias that exists in AI models and take measures to mitigate unintended bias by taking steps such as: examining quality of inputs and outputs, implementing controls, tools and metrics, tracking AI use.
- Ensure that outputs from AI are fair, accessible and accurate, minimising hallucinations where possible. Implement reviews of documentation and content to minimise risk of inaccurate outputs.
- Ensure impacts are not disparate across different user types or groups. Implement user safety measures and protect individuals from harmful content.

Why it matters

- As AI is trained on historical data, understanding the bias that exists helps companies to counter unintended bias. Bias can reinforce discrimination and inequality, leading to unjust outcomes that harm individuals and erode trust.

Caveats and considerations

- As Generative AI is trained on historical data from the internet, it can be difficult to pinpoint which biases it is learning and therefore accurately counter them.

4. We will prioritise safety and security of our AI systems, and protect Intellectual Property

What it means

- Safety & Security: Protect and secure individual and company data in AI systems that are received and used, in line with relevant data protection laws (e.g. EU and UK GDPR), policies and permissions.
- Intellectual property: Take steps to protect the Intellectual Property (IP) of content creators. Ensure confidential or copyrighted information is not uploaded into publicly accessible or unsecured AI tools without appropriate safeguards, contractual protections or organisational approval. This includes requiring permissions from AI developers before models are trained on copyrighted content.

Why it matters

- Safety & security: Prevents misuse, breaches, and loss of trust while protecting individuals' privacy and organisations from harm. Strong data protection and security practices reduce legal, financial and reputational risk.
- Intellectual property: Respects creators' work and ensure AI does not exploit or devalue human creativity and ownership. Respecting IP rights also reduces legal, financial and reputational risk.

Caveats and considerations

- Intellectual property: IP and confidentiality approaches will vary on an individual company basis. It can be difficult to determine whether training data was used with proper rights and consent.
- Safety and security: Safety and security approaches will vary on an individual company basis. AI systems require large volumes of data, increasing exposure to breaches, misuse, or unintended disclosure.

5. We implement robust AI governance processes, ensuring accountability and responsibility through human oversight

What it means

- Implement robust governance and risk management processes for AI deployment, including daily operating procedures and business functions, outlining the policies and standards users should follow. Ensure that risks are assessed in proportion to the AI use case and that appropriate controls are implemented. Use controls, tools and metrics to track where AI is being used and document the results.
- Maintain a human accountability structure which enables cooperation across an organisation yet assigns clear roles and responsibilities. Monitor and review AI systems continuously, recognising its fast-paced evolution.
- Extend responsible AI requirements and governance into the sourcing, selection, contracting and ongoing management of AI suppliers and AI-enabled services. Ensure appropriate safeguards, transparency and accountability across AI providers.

Why it matters

- Strong governance and clear accountability help to build stakeholder confidence and that the right controls are in place, ensuring decisions ultimately remain accountable to individuals.
- Allows errors or harms to be corrected and ethical standards or legal obligations to be upheld.
- Tracking and reporting help to record which AI systems are being used and prompts ongoing checks for continual improvement.
- Supplier governance helps reduce legal, ethical, operational and data-related risks, and supports improvement in supplier transparency on AI-related sustainability impacts.

Caveats and considerations

- AI solutions and tools may involve multiple systems and actors, making responsibility for outcomes harder to clearly assign.
- Governance, controls, and oversight should be proportionate to the AI use case and recognise that different applications of AI present different levels of risk and scrutiny.
- Governance frameworks must be flexible and adapt as AI technologies and use cases evolve.
- Media companies that operate under strict rules and standards for editorial control should integrate AI oversight into these existing accountability structures.
- As the adoption of agentic AI systems is growing, organisations may need to consider the unique governance challenges posed by them. This includes managing accountability separately, intervention protocols and establishing clear operational control boundaries for autonomous systems.
- Supplier transparency and control may vary across AI providers. Requirements should be proportionate to the risk and criticality of the AI use case.