GEN AI AND JOURNALISM

UTS Centre for Media Transition

PROFESSOR MONICA ATTARD is the Co–Director of the Centre for Media Transition (CMT), Professor of Journalism Practice and recently head of journalism at UTS. She spent 28 years at the ABC, working across radio and television and was the ABC's Russia correspondent from 1989 to 1994 and 2003 to 2005. She is the author of *Russia, Which Way Paradise?*

DR MICHAEL DAVIS is a research fellow at the UTS Centre for Media Transition. His research focuses on understanding and developing effective responses to misinformation. He previously worked at the Australian Communications and Media Authority, where he provided regulatory advice on online misinformation and disinformation, the news media bargaining code, and impartiality and commercial influence in news.

LISA MAIN is a former journalist and award-winning producer in Australia, the US and the Middle East. Lisa was the Director of Grants at the Judith Neilson Institute for Journalism and Ideas and prior to this, ABC Digital Editor. She has worked on international investigations including the Panama Papers and founded the ABC's Fact Check Unit. Lisa completed a fellowship at Oxford University examining the impact of metadata legislation and holds a Masters in International Relations.

CONTENTS

01

OVERVIEW

The decision to hold a global AI summit at Bletchley Park in Buckinghamshire, England in October 2023 must surely have been deliberate. The historic, hallowed estate, which became the headquarters of Allied code-breaking efforts in World War II, housed the UK Government's Code and Cypher School, which broke into the secret communications of the Axis powers – and shortened the war. That this was the location chosen by the UK Prime Minister, Rishi Sunak, to bring world leaders and tech giants together to discuss AI, speaks volumes about how seriously the technology is challenging the business, legal and social norms we take for granted. Few areas of human enterprise will be untouched. Journalism is amongst them.

Indeed, it is safe to assume that understanding the opportunities and risks of generative AI for journalism will soon be, if it is not already, a core skill for editorial policy managers across the world. Editorial managers will need to deal with questions concerning copyright of material produced by journalists, producers and editors in newsrooms, and for media managers and owners, whether the scraping of content from websites owned by journalism outlets should be compensated by the makers of large language models, and what upside generative AI poses for media organisations in both the creation and distribution of their journalism.

Questions also arise concerning redundancy of human personnel in the production cycle of journalism, as well as the ability or inability of LLMs to generate unbiased and accurate information without human intervention, and the positive – or more likely negative – impact this may have on the declining levels of public trust in journalism. The ever-changing ethical parameters within which journalists are required to operate are also likely to be impacted by the extent and manner of the adoption of generative AI in the workflow of public interest journalism.

The publication of algorithmically generated news articles on sport, weather and even finance has been a feature of the news industry for more than a decade (Nishal and Diakopolous, 2023). These have used basic template-based programs which input structured data to generate text. However, this technology might soon be surpassed by large language models which produce substantive articles, using language that is convincingly human and generally grammatically correct, synthesised from information scraped from the multiple sources upon which the LLMs are trained, and posing the risk of unintended error, bias and copyright breach along with opportunities to rethink how journalism is produced. It is, however, the downside that is grabbing attention: generative AI has created the possibility that any question or task request appropriately posed, can produce an answer or output that may be wrong because the model producing it is a statistical one built by algorithms at scale. Probabilistic in nature, it is capable of writing coherent sentences or drawing coherent images, rather than always and necessarily delivering accurate responses.

There is at least one Australian news organisation, News Corp, which had said it was using generative AI to produce 'service content', including some 3,000 weather, traffic and finance reports each week (Meade, 2023), though it has since walked this statement back (Jaspan, 2023). Using generative AI might not be problematic in the case of weather or traffic reports because whilst the number of such generated articles is large, they are repeating information verbatim from official sources such as the weather bureau or the roads and traffic authorities. Finance content is unlikely to be changed distribution point to distribution point and can be checked by a human being. The original statement by News Corp led many editors, amongst them many we spoke to for this report, to question whether generative AI, when used with appropriate human oversight, poses more of a 'time tax' on cash-strapped news-media organisations.

Whilst non-generative AI has long been used to improve workflows across a number of different newsroom activities – audio transcription, summarisation of complex documents, social-media scraping and augmentation of subscriber information including personalised content recommendations, to name a few – its use is considered by editorial managers to be time saving with clear benefit for the news production cycle. The arrival of generative AI may change the game, even if the move from non-generative AI to generative AI may in some cases be so subtle that newsrooms and their journalists might not realise that the tools they are using are in fact based on large language models or other generative AI technology.

The risks associated with the use of generative AI for the production of journalism are crystalising, from fear that 'robots' will soon be producing journalism, to reality – if impending. The guardrails required to protect information integrity and the benefits of human-produced public interest journalism are becoming clearer as organisations begin to experiment with the technology and create teams to investigate the opportunity–risk quotient. For example, newsrooms better understand the behaviour of algorithms and their capacity to tamper with information integrity as a result of the experience with digital platforms that have adversely impacted business models. However, it is clear from the research for this report and those

produced by other centres such as Polis (Beckett & Yaseen, 2023) that newsroom leaders are still pondering what needs to be understood and interrogated about generative AI, particularly whether it will improve commercial outcomes and industry sustainability. As Polis also found, there is broad recognition that a strategic approach is required and newsrooms will need to soon undergo internal analysis of their capabilities. In Australia, this process has begun. Strategic planning will inform if, how and when newsrooms will be able to harness the benefits of generative AI.

However, this 'internal analysis' that newsrooms have begun is not new to them. Indeed, since the early 1990's, when the world wide web arrived on humanity's doorstep, newsrooms have been forced to adapt, over and again. The arrival of the internet triggered profound disruptions to the news business, which required constant innovation to adapt to online distribution, and to cope with the collapse of siloed news production and distribution.

This first wave of disruption, which left newsrooms struggling to fund the purchase of new content management systems as their workflows changed, in some cases profoundly, to accommodate online distribution, and to train staff in these new systems, was not a one-off. Ten years later, social media platforms arrived on the scene, changing the way audiences consumed news and striking another blow to an already battered commercial media business model by bleeding advertising revenues almost dry. As social media innovated, it changed the way news is produced; anyone with a smart phone could produce content for a mass audience. The capacity social media provided for people to publish what they witnessed, particularly in zones of conflict, changed the way traditional media reported, forcing news media into a more participatory landscape which now operates around the clock. News media is no longer just in competition with other news media. It is also in competition with social media, which has ushered in a new, digitally inspired form of information disorder and the need for news media to up its game in its ability to verify information.

These battles, on multiple fronts, have left news media organisations in some cases fighting for survival. In recognition of this existential crisis, Australia introduced a world-first News Media Bargaining Code (NMBC) which allowed money to flow back to news media from digital platforms. To date, the obligations under the legislation have not been triggered because the dominant digital platforms – Google and Meta – have arrived at commercial arrangements to compensate news media for their content. However, as these 3-year deals come to an end, news media will be looking for new agreements. Some will be looking to other digital platforms to bring under the NMBC's umbrella, either to make up the shortfall in the event that Meta decides carrying news is too expensive or because, with AI companies commercialising a product that is trained on freely available – but costly to produce – news content, news media decides to seek compensation.

The arrival of the era of generative AI in November 2022 is viewed by most in the news industry as yet another wave of technological change capable of triggering profound disruption and requiring adaptation through the allocation of yet more resources and know-how. Over the decades of disruption to the news industry, each successive technological change has required journalists to adopt new skill sets, often to support the menial tasks of meta-tagging, SEO optimisation, image searching and resizing, and newsletter formatting, and sometimes more complex tasks such as verification. The non-generative AI systems adopted by newsrooms over the past decade have enabled them to automate some tasks associated with advanced content

THE RISKS ASSOCIATED WITH THE USE OF GENERATIVE AI FOR THE PRODUCTION OF JOURNALISM ARE CRYSTALISING, FROM FEAR THAT 'ROBOTS' WILL SOON BE PRODUCING JOURNALISM, TO REALITY, IF IMPENDING

management systems and workflows, even if verification tools have fared less well, outrun by the wiles of those determined to pollute the information ecosystem. However, optimistically, most editors CMT interviewed for this report expressed the hope that generative AI will distinguish itself from previous technological disruption by offering more relief to journalists and in turn, allowing more time for the production of actual journalism.

The purpose of the interviews conducted for this report was to glean from editors and those working in product development for news how they are currently using AI, how they see their newsrooms adopting generative AI in the future and what work yet needs to be done to ensure the editorial robustness of their news offering, if and when the new technologies are adopted.

As found in a recent global survey of newsroom usage of AI (Beckett & Yaseen, 2023) most Australian newsrooms we talked with use AI at some point across the news production process, in newsgathering, production or distribution. Most editors are clear-eyed about the benefits they see in generative AI use, particularly in relation to the time it would free up for journalists to perform more editorially analytical work. Whilst the larger legacy organisations are already assigning task forces to investigate the editorial, legal and ethical frameworks needed to protect both the news product and consumer, initial investigations are still very much based on a 'wait and see' approach. Smaller news outlets are also choosing to adopt a 'wait and see'

approach. It was unsurprising, given the relative 'newness' of the technology, to see that most editors do not believe their organisations are 'generative-AI ready' and that the solution will be found only in greater proprietorial research and development, and staff training. Despite the fact that no newsroom editors CMT spoke with are as yet contemplating using generative AI to produce journalism, all expressed ethical concerns, which ranged from informing consumers that the news they are consuming is AI-generated to editorial interventions to curtail algorithmic bias. Whilst they are under consideration, they remain without applicable solution. Some editors expressed concern at possible job losses resulting from broad generative AI adoption, whilst others countered that this impact can be mitigated by a more streamlined adoption of new technology into existing uses of AI. Across the board, the common factor in responses was that the human factor would remain of critical importance; most editors are cognisant of the decline in audience trust (Newman et al., 2023) and believed that without human oversight, trust might be further compromised. Indeed, an important finding of this research is that newsroom leaders are viewing the advent of generative AI in an already polluted information ecosystem as an opportunity to position themselves as a trusted news source. This indicates that the desire to be trusted might act as a brake on overzealous adoption of generative AI.

Against this background of as-yet-unknown risk and opportunity of AI for all human endeavour, the Bletchley Park AI summit, attended by US Vice President Kamala Harris, European Commission president Ursula von der Leyen, computer scientists and the executives of leading AI companies, has produced an international declaration, signed by 28 nations, recognising the need to be alert to the possible harms. Australia is one of the signatories to the so-called Bletchley Declaration, which commits to

collective AI safety research. The summit will convene again next year in France to review what the next steps will be. In the meantime, the nations represented at the summit have canvassed their differences and their points of agreement and all claimed that AI regulation was on their national agendas, if with differing degrees of commitment. The United States is indicating regulatory hesitation as it waits for research from its newly founded AI Safety Institute, which aims to develop standards for safety, security and testing of AI models and for authenticating AI-generated content. While the UK made clear its belief that regulation cannot precede understanding the nature of the harms which might flow from AI, the European Union is in the process of passing an AI bill. The Albanese government in Australia launched an enquiry in early 2023 to examine whether regulations are needed to ensure that AI development proceeds safely. More than 500 submissions were received by the enquiry. To date there is no indication if or when legislation might be presented to the Parliament. Despite the different pace of thinking about regulatory frameworks, each nation which attended the Bletchley Park Summit, including China, agreed that defining the problems that may be encountered by the use of AI is paramount.

On this, there would be agreement from Australian newsroom leaders. Defining the potential problems for information integrity and copyright, and acting on evidence, appears as important to the processes for the production of public interest journalism as to the health or banking sectors.

This report contains an overview of global generative AI developments as they pertain to the production of news and information, a survey of Australian newsroom attitudes and current and projected usage, and an overview of regulatory interventions and responses.

AN IMPORTANT FINDING OF THIS RESEARCH IS THAT NEWSROOM LEADERS ARE VIEWING THE ADVENT OF GENERATIVE AI IN AN ALREADY POLLUTED INFORMATION ECOSYSTEM AS AN OPPORTUNITY TO POSITION THEMSELVES AS A TRUSTED NEWS SOURCE

EXECUTIVE SUMMARY

1 Editors are taking a test and learn approach at this early phase of the adoption cycle. The test and learn environment was structured to varying degrees. Some organisations are taking a top down approach whilst others are encouraging their teams to experiment and report up.

2 Of the news organisations we spoke with none have officially deployed generative AI tools for research or the wholescale writing of articles.

3 Editors see an opportunity for fact-based verified news sources to increase in value as audiences are exposed to low-quality AI generated content. However, this opportunity is tempered by concern over the proliferation of sophisticated deep fakes and disinformation campaigns, now that AI-generated content is easily created.

4 Most editors feel we have already arrived at the 'beigeification' of news and that genAI will only accelerate this trend.

5 In some cases, editors are more concerned about the increase of misinformation and disinformation in the broader ecosystem than they are excited about the opportunities genAI might offer.

6 The opportunity most editors spoke about was automating arduous back-end processes like SEO and meta-tagging.

7 Some news organisations have identified an opportunity to serve new and different audiences via improved accessibility, automating text to voice and vice versa, for audiences who may have hearing difficulties or suffer from impaired vision.

8 There is alarm over the increased sophistication of deep fakes and the lack of tools available to newsrooms to spot deepfakes or determine if audio is authentic or fake.

9 Editors are looking at how generative AI might 'augment' the practice of journalism rather than replace journalists.

10 Every organisation has clear oversight guidelines for any piloting and experimentation of genAI.

11 Every news organisation is in the process of developing guidelines for genAI. Most organisations are taking a holistic approach including expertise from legal, technical, product development and editorial.

12 Most news organisations we spoke with stressed the importance of transparency, including making their guidelines publicly available.

13 Product teams were interested in seeing how genAI improved their existing tool kit and were also looking at developing some priority genAI tools to assist with personalisation and distribution.

02

GENERATIVE AI AND JOURNALISM
THE GLOBAL VIEW

This chapter provides a detailed overview of the impact of generative AI on public-interest journalism around the world, sourced from academic and other literature as well as media coverage. We begin with an outline of the technology itself, including its capabilities and limitations, and consider its potential impact on journalism and on the broader information environment. In this chapter, we examine how it is already being implemented in newsrooms globally, and explore other potential uses that are being considered. We then look at potential risks: for editorial practices, for the news industry, and for the information environment. Finally, we examine how newsrooms outside of Australia are mitigating these risks.

What is generative AI?

Generative AI is a term used to describe AI that generates content. It thus describes an AI system in terms of its function, rather than, for example, the type of algorithms it uses or the data it is trained on. Other functional examples include analytic or discriminative AI. Typically, however, generative AI refers to a type of machine learning system that mathematically models patterns in training data and uses those patterns to generate new content (Feuerriegel et al., 2023). Generative AI systems have been deployed to generate almost all forms of traditionally human-only outputs, including natural language text or audio, realistic images and video, and even music, in response to a user prompt.

Generative AI systems are probabilistic, meaning the system generates an output on the basis of statistical relationships modelled from training data. Text generators are typically built on what is known as large language models (LLMs) – a model built from vast amounts of text-based training data that allow it to predict the most-likely next chunk of text in a sequence. Image generators are trained on images, and again built on a statistical model of relationships between data points in the training images. Multimodal generative AI refers to systems which take a prompt in one form, such as text, and deliver a response in another form, such as an image. Foundation models, such as GPT-3 and GPT-4, are very large, general-purpose models that can be used as is, augmented via prompt engineering or additional data, or fine-tuned to provide outputs for specific applications (Feuerriegel et al., 2023). Other models have more-limited training sets and more-specific applications.

Released on 30 November 2022, ChatGPT is a generative AI system incorporating a chatbot user interface that makes it very easy to use. The free public release is built on version 3.5 of the GPT model, while a paid subscription enables access

to GPT-4. ChatGPT was the fastest-growing consumer application ever, achieving 100 million users in just two months, until the Meta-owned social media platform Threads hit 100 million only two weeks after its release in July 2023.

Non-generative AI is largely used for analytic functions such as decision making. Generative AI is able to generate novel outputs, giving it an agent-like ability to interact with users (Feuerriegel et al., 2023). The extraordinary ability of recent LLM-based AI systems to generate appropriate and convincing responses to user prompts led some to believe that the systems were or would soon be conscious, even super-intelligent, generating a flurry of discourse over an existential risk to humanity. But, as noted by Charlie Beckett and Mira Yaseen, 'Some of the more extreme dystopian critiques and over-heated marketing hype have distracted from a proper debate about immediate concerns' (Beckett & Yaseen, 2023, p. 10). Some of these concerns stem from the capabilities of generative AI; others from its limitations.

The capabilities and limitations of generative AI

The chief capabilities of generative AI are the ability to very quickly generate plausible and appropriate output in response to user prompts, and to adjust this output on the basis of further prompts. Input prompts can be quite complex, and include large amounts of data or other information.

What implications might these capabilities have for journalism? Francesco Marconi, co-founder of journalism AI startup AppliedXL and former head of R&D at the Wall Street Journal, suggests that generative AI is likely to lead to a more interactive delivery and consumption of news (AppliedXL, 2023). Bertrand Pecquerie, founder of the Global Editors Network, writes of voice-powered AI (such as through smart speakers like Amazon Alexa or Google Home) as 'the birth of conversational journalism', with journalists themselves interacting with smart speakers for research and story generation (Pecquerie, 2018).

The limitations of generative AI stem from how the models work and the data on which they have been trained. Many generative models are trained on data sourced from the internet and are not updated afterwards, effectively placing a cutoff date or recency constraint on the information they are able to synthesise in their output. Madhumita Murgia, AI editor at the Financial Times, says, 'Based on where it is today, it's not original. It's not breaking anything new. It's based on existing information' (Adami, 2023). Because of this, generative AI can't meet audience demand for more analysis or a more developed take on a subject. According to Francesco Marconi, this recency constraint makes many generative AI systems 'useless for journalistic research' (AppliedXL, 2023), though it must be noted that some generative models are now connected to real-time information sources and can provide more up-to-date information.

Another significant limitation is that generative AI is prone to produce errors of fact or nonsensical output dressed up in plausible-sounding text or realistic imagery. This limitation was observed very quickly after ChatGPT's release, with errors ranging from an inability to perform arithmetic or to comprehend causal relationships, to outright confabulation of 'facts', people and materials, leading to defamation lawsuits (Kaye, 2023) and a lot of egg on face (Belot, 2023; Bohannon,

2023). This problem, commonly called 'hallucination', essentially renders it unsuitable for use as a totally automated content-generation tool in any application where accuracy or facticity are required, such as news journalism (AppliedXL, 2023). The limitation arises from the nature of the technology itself – AI text generation is designed to generate natural-sounding text by synthesising lexical structures from an array of sources, not to provide accurate or factual responses by assessing the certainty of a particular claim. The risk of error is exacerbated if the output relates to current events or real-time data. This suggests that current generative-AI tools are unsuited to breaking news reporting, a complex and expensive operation that requires careful fact-checking and cross-referencing of information (Adami, 2023).

A third significant limitation of generative AI is bias (Feuerriegel et al., 2023). This is also found in analytic AI systems, and has a range of causes, including biases in the datasets on which AI models are trained, biases in the design or selection of algorithms, or biases in prompt design or interpretation of outputs (Feuerriegel et al., 2023; for a detailed typology see van Giffen et al., 2022). In part this is a function of availability – the internet is dominated by English-language content, for example, and AI systems trained on the internet without any debiasing processes are likely to reflect that dominance in their output. In part it is a function of social biases that are reflected in the training data – such as the association of occupations with a particular gender or race. And where limited datasets are used for training, biased selection of data may occur. Bias in generative AI means that it may provide inaccurate or misleading responses, or responses that unfairly characterise or exclude a group of people. OpenAI's report issued on the release of GPT-2 in 2019 stated, 'We expect that internet-scale generative models will require increasingly complex and large-scale bias evaluations' (Solaiman et al., 2019).

ANOTHER SIGNIFICANT LIMITATION IS THAT GENERATIVE AI IS PRONE TO PRODUCE ERRORS OF FACT OR NONSENSICAL OUTPUT DRESSED UP IN PLAUSIBLE-SOUNDING TEXT OR REALISTIC IMAGERY

The use of publicly available but copyright material to train generative AI models has raised legal questions that have important implications for news media. These largely remain unresolved, though cases are underway that may provide some clarity, at least in the US (Golding, 2023).

Perhaps the most worrisome aspect of generative AI is the potential for malicious use that undermines the integrity of the information ecosystem. A study co-authored by researchers from Cornell and Stanford universities in collaboration with OpenAI found that people are 'largely incapable of distinguishing between AI- and human-generated' news stories, which points to potential malicious use through the propagation of misinformation, propaganda and other influence campaigns (Kreps et al., 2022). While such malicious activity is not new, the danger of generative AI tools lies in the potential scale and velocity of influence campaigns: 'the ability to produce large volumes of credible-sounding misinformation quickly, then to leverage networks to distribute it expeditiously online' (Kreps et al., 2022, p. 114).

A recent report from OpenAI sets out a number of ways in which generative AI might facilitate disinformation campaigns and other influence operations (Goldstein et al., 2023).

Generative models will decrease the cost of disinformation and propaganda campaigns, potentially leading to an increase in the number and diversity of malicious actors, while commercial operations may gain new competitive advantages. Campaigns will become easier to scale and transfer between platforms. New tactics might emerge, including personalised, real-time production of misinformation via chatbots, leveraging the fact that people find generative-AI chatbots relatively trustworthy – in some cases, even more trustworthy than other people (Deiana et al., 2023). Indeed, people trust the 'average' synthetic faces that AI generates more than the irregular ones that nature does (Nightingale & Farid, 2022). Further, generative models may improve the effectiveness of messaging by allowing it to be more easily tailored to particular cultural and political contexts, particularly by those lacking local knowledge (Goldstein et al., 2023). And disinformation may be less discoverable as the ability to easily alter messaging means malicious actors will be less likely to rely on boilerplate text. The report notes also that the possibility to fine tune general language models for disinformation purposes may yield still higher risk.

Generative AI may also lead to the pollution of the information ecosystem due to a vast increase in the amount of information produced online. Francesco Marconi notes, 'The explosion of data from sources such as the web, sensors, mobile devices, and satellites has created a world where there is simply too much information. Some experts predict that by 2026, 90 per cent of online content could be machine-generated'. (Sharman, 2023b).

How can these limitations be addressed?

To address biases, potentially offensive output, and legal and ethical issues, many AI developers undertake 'alignment' processes in an effort to ensure AI output conforms with ethical principles (Gent, 2023). This can involve adding filters, fine-tuning algorithms, surgically adjusting training data or retraining models on modified datasets.

Retrieval-augmented generation is an AI framework that seeks to improve the quality of LLM-generated responses by using external sources of knowledge to supplement the model's internal representation of information. It can thus provide the model with the most current and reliable information. It can also provide users with access to the model's sources, enabling its claims to be checked for accuracy (Martineau, 2023).

Some, including Meta CEO Mark Zuckerberg, argue that open-source development of LLMs will accelerate improvement and promote the incorporation of better safeguards, an argument that some suggest may be self-serving, helping Meta to catch up with OpenAI and Google without committing the resources it would otherwise need to (Guaglione, 2023b). Meta has publicly shared the code for its LLM, Llama 2, allowing users to build their own customised chatbots from the source code. AppliedXL used open-source models, including Llama 2, to build their LLM, for reasons of cost, transparency and the ability it gives them to connect the models to real-time data. Using open-source models also allows organisations to avoid making their data available to AI companies (Guaglione, 2023b).

An OpenAI report suggests that, as they improve, language models are likely to become more usable, reliable and efficient. This may reduce the risk of unintended errors, but it may also make the models more useful for malicious actors (Goldstein et al., 2023). The report notes that there is no silver bullet to address the risks of malicious use, and that significant research and scrutiny will be required to develop effective interventions. Interventions are likely to be needed at a range of different points in the development and implementation process, including:

- model design and training – for example building more reliable models, making generative output detectable, or restricting data collection

- access – by imposing usage restrictions and developing principles for the release of new models

- content dissemination – through coordination between digital platforms and AI providers to identify AI content, requiring proof of personhood to post content, or adopting digital provenance standards

- audience susceptibility – through digital literacy training and the development of consumer-focused AI tools.

Detecting AI-generated content is difficult (Kirchner et al., 2023) and likely to get harder as models improve. AI-generated text is currently more difficult to detect than images or video (Goldstein et al., 2023). AI expert Chenhao Tan is sceptical that AI-detection tools will be of much use in the long run, and others warn that detection tools can only ever play catch up in what will become a technological arms race (Thompson & Hsu, 2023). An alternative may be building in automatic watermarking into generated content (Belanger, 2023) – though other tools may be built to remove them or models without them used instead (Goldstein et al., 2023).

A more promising route would be to watermark authentic content and use filters on digital platforms to restrict uploads to watermarked images or promote authenticated news stories (Zhao et al., 2023). The industry Coalition for Content Provenance and Authenticity (C2PA), which unifies the Adobe-led Content Authenticity Initiative and the Microsoft–BBC Project Origin under the auspices of the Linux Foundation, is collaborating on industry standards for authentication and digital provenance. While this is promising for newly created content, it would not help authenticate existing 'legacy' content archived on the internet (Goldstein et al., 2023).

Thus, a wide range of interventions are likely to be necessary to mitigate potential risks, addressing both supply and

demand sides. Addressing the supply side requires education and increased media literacy: 'While people rarely demand to be misinformed directly, information consumers often demand information that is cheap and useful for their goals – something influence operations can tailor to with greater freedom from the constraints of reality' (Goldstein et al., 2023, p. 65).

Many argue that there is a need for ethical principles for AI development, whether voluntary or enforced through regulation. We will discuss this further in Chapter 4.

How are newsrooms using AI?

AI can be seen as a third wave of digitalisation in journalism (Beckett & Yaseen, 2023). As articulated in Chapter 1, the first was the move online and the shift to mobile. The second arrived with social media, which has had an enormous impact on how content is consumed and created, as well as on the revenue of news publishers. Those revenue pressures are a significant driver of AI take up.

According to Francesco Marconi, three distinct waves of AI uptake in journalism are also apparent: automation, augmentation and generation (Adami, 2023). During the automation phase, the focus was on automating data-driven news such as financial and sports results and weather, using natural language generation and article templates. The second wave saw reporting augmented through machine learning and natural language processing to analyse large datasets and identify trends. The third phase is the rise of generative AI.

Polis, a journalism research centre at the London School of Economics, conducted a global survey of newsroom engagement with AI in the first half of 2023 that covered both

generative and other forms of AI (Beckett & Yaseen, 2023). More than 75 per cent of respondents reported using AI – of all types – in at least one area of news gathering, production and distribution, with 90 per cent using it in production, 80 per cent in distribution and 75 per cent in newsgathering. However, the take up of AI – and its social and economic benefits – are much greater in the global north.

Non-generative AI

An earlier Polis survey (Beckett, 2019) had found extensive and accelerating non-generative AI use in newsrooms, particularly for large-scale but relatively basic functions such as social-media scraping, data analysis and simple automated content creation. The most successful users of non-generative AI were organisations that 'took a strategic, holistic approach and who recognised that these technologies required fundamental self-analysis of the organisation's capabilities and future planning' (Beckett, 2019, p. 9). Indeed, the 2023 survey finds that most organisations are taking a more strategic approach to generative AI, given the lessons they have learned from dealing with non-generative AI and other technologies (Beckett & Yaseen, 2023).

In newsgathering, the most common uses of non-generative AI are in text analysis such as optical character recognition and audio transcription, and in data mining for news discovery and investigations. Web and social media scrapers help journalists identify trending topics and potential stories. Many saw this as the area of greatest success for existing analytical AI tools. On a finer-grained level they can help isolate trends amongst particular audiences and demographics. Most newsrooms are using third-party tools, though some have developed in-house systems trained on specific data (Beckett & Yaseen, 2023).

In production, common uses include data retrieval for verification and fact-checking, proofreading and information analysis using natural language processing (NLP), and content creation based on structured data such as finance, sports and weather information (Beckett & Yaseen, 2023).

The widest range of uses is in distribution, including: AI-powered social-media distribution and SEO, response to user queries through chatbots, content customisation, personalisation and recommendation systems, and speech-to-text media conversion (Beckett & Yaseen, 2023).

The benefits of non-generative AI centre on efficiency and the automation of tedious and repetitive tasks, freeing up journalists to focus on reporting. Large-scale data analysis is an area where AI tools can yield insights that may be undetectable to humans. Few see many quality advantages in using AI (Beckett & Yaseen, 2023).

Reuters is using machine learning to generate transcripts, translate audio and identify public figures on videos uploaded to its distribution service (Reuters Communications, 2023).

Generative AI

There has been rapid uptake of generative AI in the media industry outside of news. Ingenio, a US lifestyle publisher, has been using generative AI since the end of 2021, when it integrated OpenAI's ChatGPT3 into its content management system (Guaglione, 2023a). It has since used it to publish thousands of articles, with 'some editorial oversight' to check accuracy. Ingenio's head of media, Josh Jaffe, said in March this year that they could now publish 1000 articles for the same cost as a single article formerly. Despite this, he thought that AI would not take people's jobs, and if it does, they 'will be tedious ones'. Ingenio has now also integrated ChatGPT into

TAKE UP OF GENERATIVE AI IN THE NEWS INDUSTRY HAS NOT BEEN QUITE SO COMPREHENSIVE, THOUGH THE SPEED OF TAKE UP IS PERHAPS UNPRECEDENTED

a 'spiritual guide chatbot', which can personalise its responses based on a user's birth chart. Jaffe sees the future of media being more interactive in just this way, with less consumption of content in article form. He also sees a commercial opportunity, seeing the chatbot as a potential subscription product and a vehicle for native advertising. Perhaps tellingly, the explosion of content produced by the company has not yet resulted in significant gains in audience, with indexing and SEO not yet driving traffic to their sites (Guaglione, 2023a).

Take up in the news industry has not been quite so comprehensive, though the speed of the take up is perhaps unprecedented. Although 85 per cent of respondents to the global Polis survey had experimented with generative AI in the newsroom (Beckett & Yaseen, 2023), a survey of 101 journalists, editors and other newsroom staff from around the world, released in May 2023 by the World Association of Newsrooms (WAN-IFRA), found that slightly less than half (49%) of newsrooms are actively working with generative AI tools (Roper et al., 2023). These may seem like high figures; however, few respondents to either survey were frequently using generative AI or had integrated it into their production processes. According to WAN-IFRA, these figures show both the importance of the novel technology for publishers, as well as a strong degree of caution.

Some news publishers have experimented with using generative AI to write entire articles. Tech website CNET famously published a raft of articles that were later found to contain significant errors, forcing a retraction as well as the quick development of AI guidelines (Menegus, 2023). Other publishers, such as UK local and regional news publisher Reach, have been careful to maintain human editorial oversight (Sharman, 2023a). Italian newspaper Il Foglio published a series of short articles over a month, and challenged its readers to correctly identify each AI-generated article in return for a free subscription and a bottle of champagne (Adami, 2023).

RADAR AI is a media agency which combines a small team of journalists with AI-generated content to deliver customised local news stories across multiple outlets. Since 2018 it has filed over 500,000 articles for over 300 news outlets.

Several news publishers are using generative AI to generate non-news content. Buzzfeed, for example, is using AI to power its personality quizzes and the New York Times has used ChatGPT to create a Valentine's Day message generator (Adami, 2023).

Both the WAN-IFRA and Polis surveys found that generative AI was mainly used for research and workflow and efficiency improvements. Thirty-two per cent of respondents to the WAN-IFRA survey said that their newsrooms were using it for topic ideation, article creation and translation, while 19 per cent said they were using it for content personalisation and audience interaction. Other uses include minor text correction or copy editing, concept generation in design, and data analysis. The Polis survey, which was finer-grained, saw newsrooms using generative AI for headline suggestions and SEO, 'brainstorming' ideas, producing summaries and infoboxes for newsletters, generating copy and gathering background information.

While most newsrooms are using publicly available models, the BBC has been using a model trained internally on BBC content to summarise information. No content thus produced is released directly to the public without human oversight (Arguedas & Simon, 2023)

Some newsrooms, such as Indonesia's TVOne, China's Xinhua news agency and India's Aaj Tak, have used generative AI to create and animate artificial weather presenters and even news anchors, complete with their own social media profiles and personalities, and also to write their scripts and convert them to synthetic speech (Beckett & Yaseen, 2023; Samosir, 2023). Others were using it to generate video but dubbed with a human, rather than synthesised, voice. This is not just for the sake of novelty: it can make for higher quality audio than recording in the field, and it can facilitate distribution in different languages (Samosir, 2023).

Other perceived opportunities for genAI

Although generative AI represents a new and currently somewhat unpredictable phase in AI adoption, but most see its use as fairly continuous with other types of AI (Beckett & Yaseen, 2023). In content creation, for example, generative AI may supplant AI tools that use structured data to populate article templates.

Aside from the uses which have already found an application in news production, many newsrooms see great potential as generative AI develops further, particularly in personalisation and customisation of content for different audiences, experimenting with novel formats and increasing impact (Beckett & Yaseen, 2023). Many also see opportunities arising from the accessibility and ease of use of generative AI, which is a clear difference from other AI technologies that often require programming expertise. In this way, some think that generative AI might help level the playing field for less well-resourced newsrooms. In addition, some think that the ability of large language models to be sensitive to context may help automate fact-checking processes, while others see this same feature as an advantage in customisation and personalisation (Beckett & Yaseen, 2023).
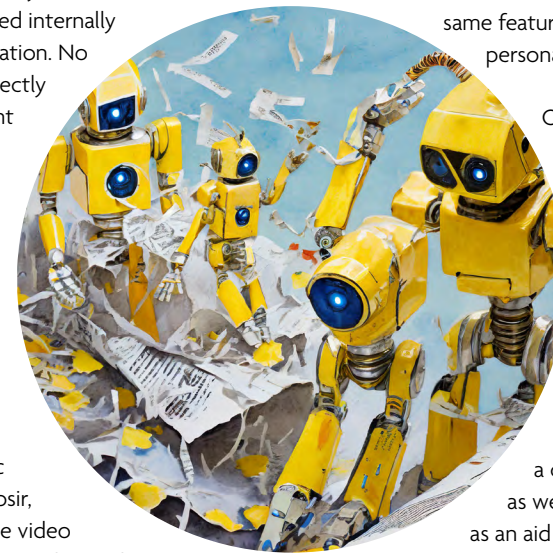
Over 70 percent of respondents to the WAN-IFRA global survey see generative AI as a helpful tool in the short term, while 18 percent said it needs more development (Roper et al., 2023). Half of respondents see generative AI mainly in a supportive role for improving workflow, while 39 per cent thought it would both improve workflow and help improve the quality of content (Roper et al., 2023) – a contrast with non-generative AI, which, as we saw above, was almost wholly viewed as an aid for workflow (Beckett & Yaseen, 2023).

The areas of greatest perceived opportunity are text creation and correction, research and improving workflow efficiency, closely followed by translation, topic ideation and personalisation (Roper et al., 2023).

Francesco Marconi argues that it is important to integrate different AI tools to account for their inherent strengths and weaknesses. For example, analytical AI tools can be used to detect, capture and analyse real-time information, for example on breaking events. This can be combined with generative technologies powered by LLMs: 'Combining these event detection systems with large language models will pave the way for an entirely new approach to journalism' (Adami, 2023).

Personalisation is the area where it is currently being used the least (19%), which perhaps reflects a need to develop more-tailored systems and workflows to take full advantage of the technology's potential to increase personalisation. Nonetheless, many think AI will eventually be capable of personalising content delivery in ways that were never previously feasible (Roper et al., 2023). A similar opportunity lies in the facility to localise content to specific audiences – both through text generation and synthetic voice. This would markedly reduce the time needed to tailor content to different audiences, particularly audio (Roper et al., 2023).

In terms of text generation, many see useful opportunities in automating and personalising text for push notifications, social-media posts and messaging platforms (Beckett & Yaseen, 2023), while others are looking at enhancing automated article creation in areas such as financial market reporting. More finely tuned models than those currently available publicly may allow generative AI to be tailored to the needs of specific newsrooms or even smaller teams within them (Roper et al., 2023).

Potentially novel uses for generative AI include using chatbots to conduct preliminary interviews with the public to gauge sentiment (Beckett & Yaseen, 2023). Generative AI can also be used to automate layout and digital production processes. This increases the potential for personalisation of news content, for example to deliver a unique issue to each user (Roper et al., 2023). Finally, news aggregation services and republishers also see opportunities in generative AI. The Newsroom, for example, is an app that delivers AI-generated summaries of the day's big news stories, drawing on multiple stories to find common facts and highlight different perspectives (Adami, 2023).

Implementation strategy

Only around one third of respondents to the Polis survey said their organisations had or were developing an AI strategy. For these newsrooms, AI strategy and implementation are mostly being led by dedicated cross-functional or digital-innovation teams (Beckett & Yaseen, 2023). Many are reviewing their strategies with the rise of generative AI. Some have established partnerships with AI vendors, sometimes temporarily while they build in-house capabilities.

Most newsrooms surveyed for the Polis report noted the importance of staff training to increase AI literacy. This was particularly the case for less well-resourced newsrooms that are only beginning to adopt AI. But the arrival of generative AI has seen a need for further training even in well-resourced newsrooms, with a focus on prompt design and experimentation with large language models, as well as understanding legal, ethical and business implications (Beckett & Yaseen, 2023).

Eighty-five per cent of respondents to the Polis survey thought more collaboration on AI amongst newsrooms and between newsrooms, academia and tech companies would be useful, particularly to help bridge knowledge and resource gaps in small newsrooms and in the Global South. For example, universities can help with research on effective and ethical uses of AI in journalism (Beckett & Yaseen, 2023).

Challenges to implementation

Over half of respondents to the Polis survey said they were not ready or only partially ready to deal with the challenges of AI integration, particularly with the arrival of generative AI. Financial constraints and lack of technical expertise were key

challenges. Technical expertise is a matter not only of hiring AI specialists but also training staff so that AI can be successfully integrated across the organisation. AI specialists tend not to understand journalistic applications, while journalists don't understand how best to make use of AI (Beckett & Yaseen, 2023).

Similarly, technical challenges arise from the need to integrate AI into existing systems to achieve interoperability. This requires a great deal of development and testing. These technical challenges are a reason why most publishers are choosing proprietary AI tools rather than adopting open-source models, which don't come with technical support. Instead, publishers are increasingly signing deals with AI companies, possibly exchanging data for access to quality models (Guaglione, 2023b).

These challenges are more keenly felt in smaller newsrooms and those in the Global South (Arguedas & Simon, 2023; Beckett & Yaseen, 2023).

Other challenges include ethical concerns over responsible use of AI, cultural factors including scepticism and resistance to change, and managerial challenges relating to conflicting priorities. These have arisen particularly with the arrival of generative AI.

Some industry organisations are helping smaller newsrooms to improve their AI capabilities. In the US, for example, the Associated Press Local News AI initiative, which began in 2021, is helping small newsrooms to implement AI tools to automate time-consuming tasks (Associated Press, n.d.). The non-profit, industry-funded Partnership on AI has developed a guide to help local newsrooms with implementation (Partnership on AI, 2023).

OTHER CHALLENGES INCLUDE ETHICAL CONCERNS OVER RESPONSIBLE USE OF AI, CULTURAL FACTORS INCLUDING SCEPTICISM AND RESISTANCE TO CHANGE, AND MANAGERIAL CHALLENGES RELATING TO CONFLICTING PRIORITIES

Impact on newsroom workflows and roles

Around a quarter of respondents to the Polis survey said AI adoption has already had a significant impact on workflows, saving a great deal of time in processes like fact-checking, social-media monitoring and content distribution, and freeing up journalists to spend more time on newsgathering and editorial tasks. Automation has changed the nature of production work rather than replacing it altogether. Automated processes still require staff attention, but at a broader, systems level, such as adjusting parameters or rules rather than needing, for example, to schedule distribution of posts individually (Beckett & Yaseen, 2023).

Many newsrooms noted that they are hiring AI-related roles including prompt designers, AI engineers and data analysts. Other staff are being trained in AI use, particularly in prompt design, and see their roles changing as a result of AI adoption.

Forty-five per cent of respondents to the WAN-IFRA survey thought that newsroom jobs would change significantly as a

result of the technology, while only 14 per cent thought they wouldn't change at all (Roper et al., 2023).

Many argue that generative AI will require a pivot in focus for journalists. With much mundane reporting being automatable, Axel Springer CEO Mathias Doepfner believes that news media must focus on investigative journalism and original commentary (Yerushalmy, 2023).

Perceived risks of generative AI

Many journalists and newsrooms fear that generative AI will exacerbate the risks of other forms of AI, and potentially lead to new risks.

Editorial risks

Eighty-five per cent of respondents to the WAN-IFRA survey were most concerned about inaccuracy and the quality of content produced by generative AI, while 67 per cent were concerned about plagiarism and copyright infringement. Only

38 per cent were concerned about threats to job security while 46 per cent were concerned about privacy and data protection.

Eighty-two per cent of respondents to the Polis survey raised concerns about editorial quality. Many feel that generative AI is likely to exacerbate the risks posed by other types of AI, such as bias, inaccuracy, and privacy. Being conscious of and acting to mitigate these risks is critical to avoid negative impact on public trust (Beckett & Yaseen, 2023).

Algorithmic bias is a major concern, with many worried that reliance on AI could increase bias in news coverage. Biases in training data could lead to biased content recommendations, skewed perspectives, or misrepresentation in news coverage, particularly of marginalised groups and in languages other than English. This risk is more severe in many countries in the Global South (Arguedas & Simon, 2023; Beckett & Yaseen, 2023).

Many are concerned about accuracy and verification, particularly of third-party content obtained from social media during breaking news events (Arguedas & Simon, 2023). Most note that human editorial oversight of generated content is critical, and that AI cannot be relied upon for fact-checking. The New York Times recently tested five AI-detection tools, which often produced both false positives and false negatives, likely because they can't process contextual clues like humans do (Thompson & Hsu, 2023).

Some argue that verification is a standard journalistic process and that newsrooms are thus well equipped to address these problems. Chris Looft, coordinating producer of visual verification at ABC News (US), says that ABC News 'already has a team working around the clock, checking the veracity of online video', and that the arrival of generative AI models doesn't change this. 'The process remains the same, to combine reporting with visual techniques to confirm veracity of video. This means picking up the phone and talking to eye witnesses or analyzing meta data' (Sherman & Rizzo, 2023).

Dealing with manipulated or completely fabricated images is also a serious concern, as there is no original material to compare a fabricated image with and there is a need for tools to reliably detect AI-generated content (Beckett & Yaseen, 2023).

Another concern that carries legal risk in many jurisdictions is privacy. The training sets of many generative AI models include publicly available personal information obtained from social media and other websites. OpenAI has been sued for breaches of the EU's General Data Protection Regulation (GDPR) as well as class actions under US law (Brittain, 2023; Woollacott, 2023). Although the information it was trained on was publicly available, it may still breach privacy law, which often requires that private information is not used or revealed outside the context in which it was originally produced or collected (Gal, 2023). In addition, many AI systems collect the data from user prompts to further train or fine-tune their models. Concerns for newsrooms arise from publication of private information in AI-generated content and from information used in prompts, whether it is information about a journalist, a source or a subject in the news, as well as information selected for in-house AI-training sets.

Industry and business risks

Risks to the journalism industry are a substantial concern. Some fear that AI will incentivise the mass production of poor-quality journalism such as clickbait or even politically biased commentary and other polarising content, feeling that the quest for personalisation will bear the same rotten fruit as the quest for engagement that drives social media (Beckett & Yaseen, 2023). Likewise, AI trained on commercial media also 'risks replicating all the biases of old media' (Warren, 2023).

Similarly, some are concerned about the commercial motives of AI companies, suggesting these may outweigh ethical concerns such as privacy, fairness, accountability and transparency. Others worry that newsrooms will become dependent on proprietary technologies developed by a small number of companies, effectively making AI companies another gatekeeper for news and introducing AI risk into core newsroom infrastructure, particularly when their functioning is not well understood (Arguedas & Simon, 2023). This is exacerbated by concerns about the values of AI companies being imposed on system users through their setting of system filters and other rules (Beckett & Yaseen, 2023).

Sustainability is a risk that many see for the industry. Small organisations may find it even more difficult to compete with large players who can leverage AI to dominate news markets.

On the other hand, as noted earlier, some thought that the accessibility of generative AI may help small newsrooms to experiment and extend their reach, while failure to adapt may threaten larger companies, particularly those that have been slow to digitise their businesses (Beckett & Yaseen, 2023).

Some newsrooms expect the new generations of AI to result in job losses as more tasks become automated; others say that while AI will require journalists to learn new skills, it will not replace jobs (Beckett & Yaseen, 2023). German publisher Axel Springer has forecast likely job losses, with CEO Mathias Doepfner stating in a letter to employees that AI has the 'potential to make independent journalism better than it ever was – or simply replace it', noting that 'only those who create the best original content will survive' (Yerushalmy, 2023). News Corp global CEO Robert Thomson has predicted a 'tsunami' of job losses (Jaspan, 2023).

There is concern that global disparities in technological capability will be reflected in the rate and level of AI adoption, leading to a greater impact on media sustainability and reach in less wealthy countries. A report from consultancy firm Oxford Insights observes that many countries in the Global South lack an appropriate supporting environment for AI development and integration, which includes a robust technology sector, adequate data and communications infrastructure, strategic vision and attention to governance and ethics (Rogerson et al., 2022). Publicly released AI tools based on foundation models like ChatGPT break down some of these disparities. However, they tend to be proprietary and controlled by large companies, perhaps increasingly so. They are also mostly trained on English-language data and therefore of limited use in other languages or cultural contexts. For AI solutions to be useful and to mitigate potential shortcomings, large amounts of local data are required to adapt systems to the local context. Yet such data is also less readily available in resource-constrained regions, and training large AI models is expensive and requires technical expertise (Yu et al., 2023). In such environments most media companies lack the resources to develop in-house systems or to hire in-demand experts. International competition between economic powers may again worsen the situation, with publicly accessible systems such as ChatGPT not available in many areas of the world due to sanctions or governance concerns (Beckett & Yaseen, 2023).

There is also concern about the potential impact of generative AI on news industry revenues. As generative AI is incorporated into open digital platforms like search engines, it may exacerbate the 'walled garden' problem where users increasingly stay within the search engine environment instead of navigating to a news website, leading to a reduction in digital advertising and subscription revenue (Bruell, 2023).

Gary Rogers, founder of RADAR AI, suggests that 'part of what makes generative AI so disruptive is that the tools of content creation are moving toward the audience, which could fundamentally change the role of news organisations and their ability to monetise on their content. If we all have generative AI tools at home … the role of news organisations may be to create the information people can build on, shifting from news production to news gathering' (Arguedas & Simon, 2023, pp. 14–15).

Relatedly, AI companies are reaping commercial benefits from news because it is openly available on the web, while it is news companies and other information sources that bear the cost of producing it. This has led many news companies to block AI webscrapers from accessing their websites, and to a series of copyright challenges around the world, the outcome of which will likely have a significant impact on the operations of both the AI and news industries. Francesco Marconi argues that to have an AI industry that is sustainable in the long term, 'we also need an equally sustainable news sector, because it is the source of high-quality information' that is used to train LLMs (AppliedXL, 2023).

Some of the world's largest news companies, including Associated Press and News Corp, have made deals or are reportedly in talks with OpenAI, and several, including News Corp, Axel Springer, Vox Media and Advance (Condé Nast) are reportedly forming a coalition for the purpose (Buckley & Wilson, 2023). In Australia, some media companies, including News Corp, have suggested that AI firms should be brought within the News Media Bargaining Code or a similar framework that will incentivise them to compensate news producers for their content. Industry bodies have also issued statements of principles for AI development and governance that call for compensation for use of news and other media content, such as the News/Media Alliance (2023) and Digital Content Next (2023).

'On the one hand [journalism] is around the world a sector under great commercial, political and competitive pressure. It is weak in resources compared to the giant corporations developing this technology. The potential for deep structural threats to journalism in the future must be part of our thinking now' (Beckett & Yaseen, 2023, p. 10).

THERE IS CONCERN THAT GLOBAL DISPARITIES IN TECHNOLOGICAL CAPABILITY WILL BE REFLECTED IN THE RATE AND LEVEL OF AI ADOPTION, LEADING TO A GREATER IMPACT ON MEDIA SUSTAINABILITY AND REACH IN LESS WEALTHY COUNTRIES

Risks to the information environment

Many newsrooms are wary of risks to the information environment. The concern is with both irresponsible use of generative AI by news media (i.e. editorial risks writ large), and increased propagation of misinformation in the broader information ecosystem (Beckett & Yaseen, 2023). As News Corp CEO Robert Thomson put it, 'The danger is, it's rubbish in, rubbish out, and, in this case, rubbish all about' (Jaspan, 2023). In the newsroom, the primary concerns are algorithmic bias, inaccuracies in AI-generated output, and the problems of verification discussed above. For the broader ecosystem, the common view is that generative AI will allow the dissemination of misinformation at much larger scale. 'Unchecked algorithmic creation presents major risks as it relates to a healthy information ecosystem,' Marconi says. 'This doesn't mean that generative AI has no role in journalism, but that we can't solely rely on it' (Adami, 2023).

Gordon Crovitz, of NewsGuard, says that generative AI 'is going to be the most powerful tool for spreading misinformation that has ever been on the internet. Crafting a new false narrative can now be done at dramatic scale, and much more frequently' (Hsu & Thompson, 2023a). It will be easier for malicious actors to generate convincing articles that counterfeit the branding and format of genuine news products. Chatbots could share conspiracy theories in increasingly credible and persuasive ways, whether responding to innocent user queries or used to craft targeted disinformation campaigns, while deepfaked images, video and voice will become increasingly sophisticated. But the problem does not solely reside in the production of convincingly deceptive content; pollution of the information ecosystem by, as Steve Bannon put it (Stelter, 2021), 'flooding the zone with shit' makes it more difficult for users to discern quality information, thus disrupting the 'marketplace of ideas'. This parallels problems being seen already in the creative industries, with AI-produced material being produced quickly and at high volume, but with low quality (United States Federal Trade Commission, 2023).

A 2019 OpenAI report accompanying the release of ChatGPT precursor GPT-2 cautioned that 'easy-to-use interfaces can enable malicious use from otherwise unskilled actors'. Such an interface is, of course, one of ChatGPT's notable features.

A January 2023 OpenAI report on the risks of malicious use of generative AI concluded that language models could 'significantly affect how influence operations are waged in the future'.

Aside from the potential for harm arising from the use of generative AI in disinformation campaigns, the information environment is already being polluted by the creation of AI-powered content farms which exist principally to attract digital advertising revenue (Brewster et al., 2023).

Erosion of trust in news, science and official sources of information is a well-recognised consequence of a polluted information environment, though this is difficult to quantify. Simply being aware that the information environment is polluted by manipulated content and malicious actors has been shown to lower trust and lead people to believe that authentic content is actually fake (Ternovski et al., 2022). There is evidence of this effect occurring during the 2023 Israel–Hamas conflict (Hsu & Thompson, 2023b).

However, some argue that the misinformation threat from generative AI is overblown. One study contends that despite the potential for generative AI to create misinformation more easily and at greater volume and scale, what matters for harm is how much misinformation is consumed and what effect it has on those who consume it (Simon et al., 2023). Consumption, the authors argue, is mostly limited by demand and not supply. Thus, partisanship and identity, rather than access to reliable information, are key factors in whether people believe and share misinformation (Altay, Berriche, Heuer, et al., 2023). The problem with those who share misinformation is not that they 'do not have access to high-quality information but instead that they reject high-quality information and favor misinformation' (Simon et al., 2023, p. 3). At the same time, other people are more likely to be uninformed than misinformed (Altay, Berriche, & Acerbi, 2023).

This perhaps provides an additional reason in support of the view expressed by several newsrooms that the advent of generative AI presents an opportunity for trusted news sources (Beckett & Yaseen, 2023). Indeed, some researchers argue that interventions aimed at increasing trust in reliable news sources are likely to be more effective than trying to address misinformation or manipulated media directly (Acerbi et al., 2022).

A report from the Balliol Interdisciplinary Institute sees the increased personalisation potential of generative AI as a two-edged sword, with the ability to tailor responses to individual preferences possibly limiting output to that which reinforces a user's worldview, in a reanimation of the 'filter bubble' problem that is a (somewhat disputed) consequence of recommender algorithms driven by user engagement (Arguedas & Simon, 2023) or to contribute to polarisation in the public sphere (Ventura-Pocino, 2021).

Global disparity in access to new AI systems may exacerbate problems in the information environment. Countries in the Global South tend to experience worse problems of information pollution due to less-robust democratic frameworks, including fewer regulatory controls on digital platforms (Von Nordheim et al., 2023). Independent news media are often small and poorly resourced organisations with low internet penetration, making it difficult for them to counteract local and foreign state propaganda and disinformation, particularly where these benefit from AI technologies. This leads to a lack of trust in AI technologies within independent news organisations (Beckett & Yaseen, 2023).

How are newsrooms mitigating the risks of AI?

As a result of the editorial risks set out above, many newsrooms have decided against using AI in editorial functions or decision-making (Beckett & Yaseen, 2023).

However, there doesn't appear to be a uniform approach for how newsrooms are enforcing the use of generative AI tools. Nearly half of the participants in the WAN-IFRA survey said that their journalists have the freedom to use generative AI as they see fit, while only 20 per cent have internal guidelines in place. Three per cent of respondents said that they don't allow generative AI to be used at all (Roper et al., 2023).

The number of news organisations that have developed internal guidelines to govern how they use AI is, of course, difficult to ascertain as news media are generally under no obligation to publish internal guidelines. However, by April 2023, one study identified seven organisations with publicly available guidelines (Becker, 2023), including Associated Press,
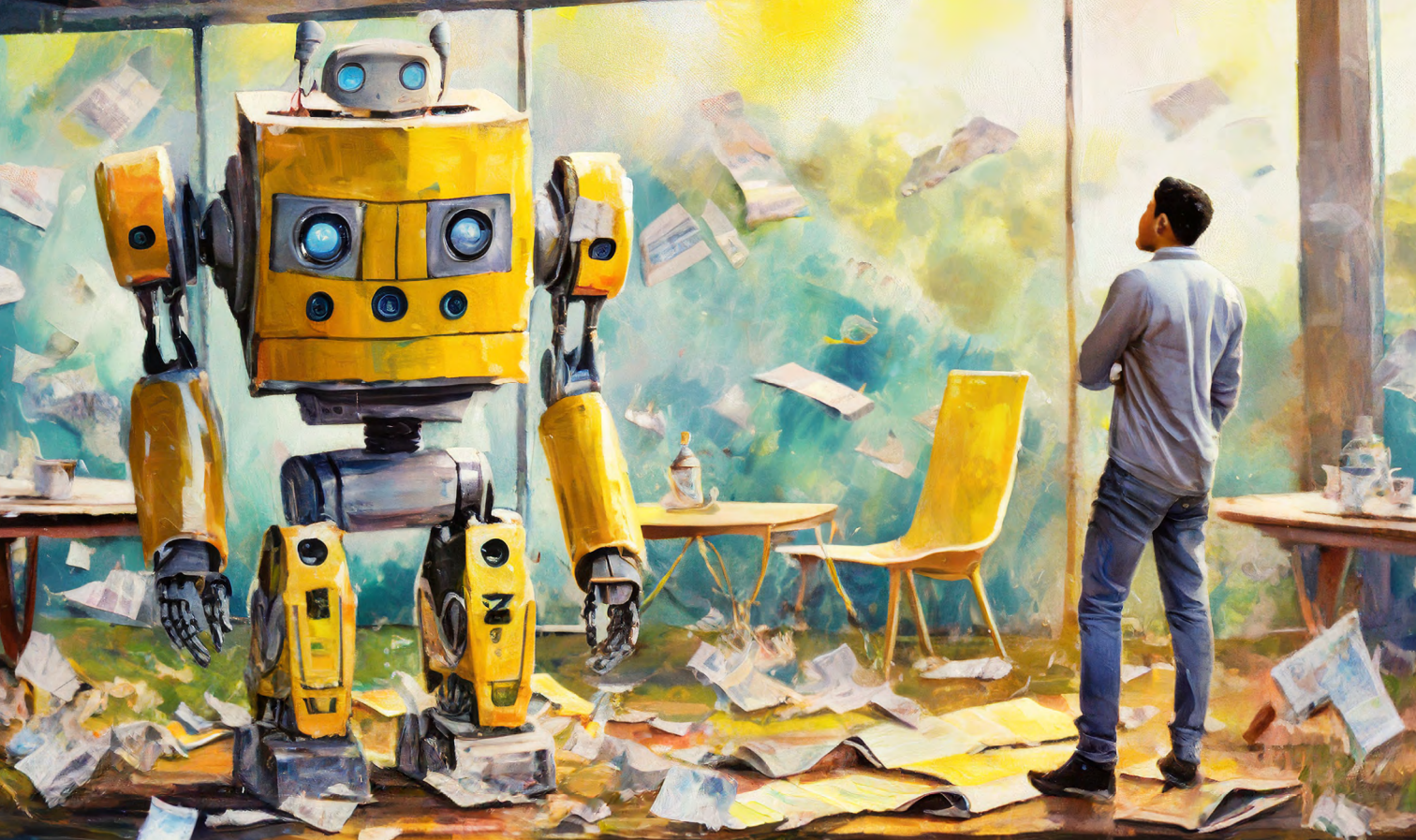
Thomson Reuters, Bayerischer Rundfunk, the BBC, Deutsche Presse Agentur and Wired. Most of these were published before the release of ChatGPT, and thus do not consider generative AI specifically.

In July, a Nieman Lab article covered 21 organisations with published guidelines or advice on AI (Cools & Diakopoulos, 2023). Some of these, such as CBC/Radio Canada (Fenlon, 2023), are relatively developed sets of guidelines, while others, such as the Guardian (Viner & Bateson, 2023) and the Financial Times (Khalaf, 2023), are broad advice or principles. Some are published in a letter from the editor or a blog post rather than in complete form. Internal documents may be more comprehensive.

Industry bodies have also moved to develop guidelines on AI. These include Finland's Council for Mass Media, Belgium's Raad voor de Journalistiek, Germany's Journalisten-Verband, and the Radio Television Digital News Association in the US. In Australia, the Media Entertainment and Arts Alliance has issued a position statement on AI which contains several statements on journalistic process (Media, Entertainment & Arts Alliance, 2023).

Becker et al. (2023) analysed a sample of 52 internal and industry AI guidelines and policies from Europe, the USA and Canada, the majority of which were published after April 2023. Eight-seven per cent of the sample documents explicitly state where AI is permitted in the journalistic process, while 67 per cent specify where it is not permitted. Eighty-five per cent require human supervision in some form, though only 65 per cent always require it. Seventy per cent specify risks of AI use, including error and hallucination, bias and copyright violation. Ninety per cent refer to transparency, but only 17 per cent specify exactly how this will be achieved, such as through a byline or an endnote. Fifty-four per cent refer to data privacy and the same to source protection. Thirty-seven per cent of guidelines note internal coordination between different departments on AI, while 17 percent mention external cooperation with technology companies, consultants, researchers or government. The majority do not refer to dependency or independence. Seventy-three percent mention responsible professional roles such as editor-in-chief or legal staff, but only eight per cent of the documents refer to specific compliance mechanisms.

Comparing different organisations, the authors observe that commercial media are somewhat more permissive with the use of AI than public-service media, but the former tend to have finer-grained guidelines and more frequently call for transparency and human oversight of content than the latter, while public-service media guidelines refer more to human control over algorithms. There is also some variation across countries.

The authors note several blind spots in the guidelines. These include a lack of enforcement and accountability mechanisms, a lack of oversight over algorithms and AI systems as opposed to content output, and a lack of explicit directives on external collaborations. Given the reliance on external expertise, the authors argue that guidelines could 'include provisions for transparent and ethical engagement with such actors' (Becker et al., 2023, p. 21). The authors also observe that, despite references to serving audiences and the general industry emphasis on greater audience engagement, no guidelines refer to the need to engage audiences or seek their feedback on AI use or guidelines. There was also little awareness of broader social equity issues with AI.

Developing ethical guidelines and responsible practices is seen by many as one of the most challenging areas in terms of its complexity and demand on resources. For example, while bias is a well-known shortcoming of AI, knowing how to address it is another matter. Some worry the complexity of the task means it is likely to be ignored (Beckett & Yaseen, 2023).

One respondent said that with the need to have dedicated data teams there was a concomitant need for a value-alignment unit to ensure data teams operated in accordance with editorial values. As well as being intrinsically wrong, misalignment will 'result in the brakes being pulled (rightfully) before AI-prototypes are employed at any real scale in most legacy newsrooms' (Beckett & Yaseen, 2023).

A small number of newsrooms are employing rudimentary de-biasing techniques for recommender systems and NLP systems, but generative AI is still too new and dynamic for biases to be systematically assessed or addressed. Engaging in red-teaming work is very time consuming and difficult to implement in newsrooms due to a lack of dedicated staff (Beckett & Yaseen, 2023).

Most newsrooms believe human oversight is essential to responsible use of AI, particularly to avoid bias and inaccuracy. The importance of context and interpretation also limit reliance on automation, with 'untuned' foundation models in particular unlikely to be sufficiently sensitive to local context (Beckett & Yaseen, 2023). Charlie Beckett says that 'AI is not about the total automation of content production from start to finish: it is about augmentation to give professionals and creatives the tools to work faster, freeing them up to spend more time on what humans do best … Human journalism is also full of flaws and we mitigate the risks through editing. The same applies to AI' (Adami, 2023).

How much human oversight, and what kind, may depend on the content and the purpose. The Newsroom is planning to 'streamline' its human-review process, having different tiers of oversight 'depending on the topic'. Sports, for example,

DEVELOPING ETHICAL GUIDELINES AND RESPONSIBLE PRACTICES IS SEEN BY MANY AS ONE OF THE MOST CHALLENGING AREAS IN COMPLEXITY AND DEMAND ON RESOURCES. FOR EXAMPLE, WHILES BIAS IS A WELL-KNOWN SHORTCOMING OF AI, KNOWING HOW TO ADDRESS IT IS ANOTHER MATTER

will likely not need the oversight of geopolitics and climate coverage, while breaking news, in particular, is an area where AI performs poorly and where its application should be severely limited. One of the Newsroom's founders, Pedro Henriques, states, 'When news is breaking, there's still not enough information for us to be able to really validate it properly' (Adami, 2023).

Some are sceptical of the possibility of 'aligning' AI models with ethical values, which requires the operationalisation of values into algorithm parameters. Some noted that values are not universal and can be politically contested – some of the filters that OpenAI has implemented in ChatGPT reflect a US-centric worldview and are aligned with the company's commercial values (Beckett & Yaseen, 2023). Academic research has suggested an in-built 'left-libertarian' bias to ChatGPT (Hartmann et al., 2023; Motoki et al., 2023) while a study from the Brookings Institution noted the markedly different responses from Google Bard and ChatGPT on political questions (West, 2023).

However, some AI companies are developing AI tools specifically for journalism, with journalists involved in development. For example, AppliedXL purports to ground its models with foundational journalism principles, including transparency and accuracy, as well as principles for identifying newsworthiness, and keeps humans in the loop to monitor and retrain the system. The Boston Globe is one news organisation making use of AppliedXL tools (Guaglione, 2023b). It is notable though that its current tools have a very specialised application to news about ongoing medical research. AppliedXL co-founder Francesco Marconi told the Reuters Institute that 'the news industry must be actively engaged in the AI revolution. In fact, media companies have an opportunity to become a major player in the space – they possess some of the most valuable assets for AI development: text data for training models and ethical principles for creating reliable and trustworthy systems' (Adami, 2023). Marconi also urged the industry to use AI

to help filter out 'noise' and disinformation from an ecosystem polluted by too much unwanted information. 'This marks an inflection point, where we now must focus on building machines that filter out noise, distinguish fact from fiction, and highlight what is significant'. Using customised third-party tools may in many cases allow publishers to keep control of their training data (Guaglione, 2023b).

Most suggest that transparency is a key element of ethical practice and critical for maintaining trust. Noting that it is impossible to know exactly how a system generates its output or makes decisions, many say that newsrooms must be fully transparent when they use AI in content creation or other editorial tasks. At the same time, some worry that a public that is sceptical of AI will reject newsrooms that openly use it.

Similarly, many believe AI companies must be more transparent about how their systems work, and should focus on explainability. Some said they should also proactively engage with journalists and other stakeholders to improve efforts at value alignment, including in less attractive markets to bridge digital divides, particularly for vulnerable communities (Beckett & Yaseen, 2023). OpenAI recommended in its release report for GPT2 that new AI releases with an impact outside the AI community 'should undergo interdisciplinary analyses among researchers and broader society'. Principled decision-making frameworks and 'infrastructure for distributed risk analysis', such as legal agreements that balance interests may facilitate such pre-release analyses (Solaiman et al., 2019).

In the next chapter we will look at how media organisations in Australia are thinking about and using generative AI, as well as the precautionary moves they are making to protect the integrity of their content.

03

THE VIEW FROM
AUSTRALIAN NEWSROOMS

This chapter presents the findings of qualitative research based on interviews with Australian newsroom editors, producers and product development teams.

The objective was to gain insight into how Australian newsrooms are thinking about generative AI and if they are beginning to adapt their intra-newsroom workflows to accommodate the technology. This chapter explores this and any related changes Australian newsrooms are making to journalistic methodology and practice to mitigate risks. We also discuss the use of non-generative AI tools that have been deployed in newsrooms for some time and how generative AI might augment or replace these capabilities.

With newsrooms in the global north actively turning their attention to the challenges and opportunities of generative AI for their work processes and outputs, our intention was to ascertain where Australian media organisations are situated with respect to generative AI.

Following the insights gained in our review of international developments in chapter 2, our interviews explored the following questions:

- How newsrooms are using or are planning to use generative AI tools in newsgathering, editorial and production processes and how this builds on current uses of non-generative AI, as well as broader opportunities they are considering but have not yet begun to implement.
- How news organisations are thinking about generative AI in the context of editorial processes, the business side of news production and the broader information environment.
- Any related changes newsrooms are making, or considering, to journalistic methodology, guidelines or business practices to mitigate any risks arising from generative AI.

Report methodology

This report is based on a series of semi-structured interviews with news editors and product teams across news media organisations, ranging from those which can be broadly defined as legacy media organisations through to smaller publishers, both traditional and online-only. The organisations span the publicly funded national broadcasters and commercial organisations, large and small. Editors and product personnel from the following organisations were interviewed:

- Australian Broadcasting Corporation (ABC)
- Sydney Morning Herald (SMH)
- Guardian Australia
- The Daily Mail
- The SBS
- Southern Cross Austereo (SCA)

which is interesting. Maybe that's because they understand it better than others. Maybe because they don't have the benefit of seeing various changes and disruption in the newsroom over the past 20 or 30 years. And God knows there's been lots of it." (edited for clarity)

Shock and awe – the arrival of generative AI

Editors are acutely aware we are in the early adoption cycle of generative AI, but they also understand that this won't last long because the pace of development is faster than anything that has come before it.

"I think any organisation that told you that they were ready for this challenge would be lying. It's happening too fast. It's just extraordinary how quickly things are changing." Lenore Taylor, Editor, Guardian Australia

The interviews reflected an awareness that the external environment in which journalism operates, whether online, on social media or behind subscriber paywalls, has significant influence over how news organisations structure themselves internally. This has implications for investment decisions around technology stacks and how newsroom workflows are organised.

All news editors we spoke with recognised generative AI technologies would have a significant impact on the news industry. For the most part, editors felt it was too early in the adoption cycle to articulate precisely how the news industry might change, but no one was prepared to underestimate how big the shift would be.

"I think it's going to change many industries. And it's going to change the opportunities available in many industries, the work, the jobs available, in ways that I'm not educated on it enough to tell you. I just know that it's just getting cleverer and more refined all the time. So I'm anticipating massive upheaval, which our industry has seen ever since I joined it, really. But I'm anticipating more." Lisa Allan, Editor, Newcastle Herald

Everyone we spoke to said generative AI would significantly reshape the broader news ecosystem. All recognised the threat of what the founder of Google's DeepMind, a leading AI laboratory, refers to as the "misinformation apocalypse" where disinformation is created at a pace, scale and sophistication that is unprecedented (Thorbecke, 2023).

"I think it's going to have massive implications, bigger implications than we could imagine now, and I think it will have implications on everything from the media's reliance on Google Search referrals, through to our sense of what information is, and if it can be trusted or not." Justin Stevens, Director, ABC News.

"I think there's going to be challenges. Like, the challenges around maintaining the integrity of our work. So I think it's going to be hugely challenging." Lisa Allan, Editor, Newcastle Herald (edited for clarity)

Most news editors we spoke to were focused on the immediate threat of an increased disinformation challenge and a general decay of the digital information ecosystem. Most were clear that they could not predict precisely how generative AI was going to reshape the broader news ecosystem other than opining it would deliver more disinformation and deep fakes in the near term. This was

- The Newcastle Herald
- Australian Community Media (ACM)

The editors and product personnel CMT interviewed were:

- **Lenore Taylor**, Editor, Guardian Australia
- **Bevan Shields**, Editor, SMH
- **Chris Paine**, Head of Premium Content, Metro Publishing Nine
- **Sophia Phan**, Growth Content Editor, SMH, The Age, Brisbane Times, WAToday
- **Justin Stevens**, Director of News, ABC
- **Angela Stengel**, Head, Digital Content & Innovation, ABC
- **Gina McKeon**, Editor, ABC Innovation Lab
- **Erin Reimer,** Senior Editorial Lead for digital news and current affairs, SBS
- **Melanie Withnall**, Head of News and Information, SCA
- **Barclay Crawford**, Editor, Daily Mail Australia
- **Lisa Allan**, Editor, Newcastle Herald
- **Saffron Howden**, National Editorial Training Manager, ACM

CMT also requested participation in this study from News Corp Australia; however, the organisation declined. News Corp Australia Executive Chair Michael Miller is on the public record detailing what he initially said was the company's adoption of generative AI in limited newsroom settings (Meade, 2023), before stepping back from those comments (Jaspan, 2023). Where relevant, this report refers to those public comments.

Editors and newsroom leaders we spoke to are cautiously optimistic. Many are energised by the opportunity for innovation and at the same time thinking deeply about how generative AI will challenge many of journalism's fundamentals, including the requirement to verify information. This sentiment was encapsulated by Bevan Shields, editor of the SMH:

"I've seen all sorts of things come and go, you know, the end of X platform or Y platform. And I think those people are relatively relaxed about this and what it might mean. Younger reporters are more anxious than excited about it,

EVERYONE WE SPOKE TO SAID GENERATIVE AI WOULD SIGNIFICANTLY RESHAPE THE BROADER NEWS ECOSYSTEM. ALL RECOGNISED THE THREAT OF WHAT THE FOUNDER OF GOOGLE'S DEEPMIND REFERS TO AS THE "MISINFORMATION APOCALYPSE"

what concerned them the most – the increased flood of low-quality, unverified news content and the impact that would have on public discourse and democracy more broadly.

"My concern is that as information sources splinter, the morass people will have to wade through to try to find reliable information is just getting deeper and deeper and deeper." Lenore Taylor, Editor, Guardian Australia

"So I don't know where that's going to end up. But I do worry, because I think you need an agreed set of facts to be able to make decisions around politics, around what really happened. There's always going to be an opinion and opinion is great and fabulous and we have really strong opinion-based media in Australia. But I just think, how are we going to navigate all of this? And that's the thing that I probably worry about more than the tools themselves and what the tools can do." Melanie Withnall, Head of News and Information, SCA

One newsroom editor recalled several instances where journalists had been sent images which were not flagged by their standard image verification processes such as reverse image search, only to find later that the images were, in fact, fake.

Other news editors raised questions about our reliance on Google searches to find information and discover news stories. One editor spoke about what happens to discoverability of news articles once audiences evolve their habits to using AI chatbots such as ChatGPT instead of search engines. ChatGPT will answer questions from

the information its crawlers have scraped from the internet, including news articles, but it might not offer links to the original news story in the same way as traditional internet search engines.

"I think we're not blissfully ignorant, but I think the whole industry here and abroad will be caught flat footed. In particular, the reliance that organisations have on referrals from Google search." Justin Stevens, Director, ABC News

Similarly, some editors asked what happens when audience habits evolve to trusting generative AI tools to discover their news sources for them.

"Well, I think there's a lot of potential harms, in that they give the appearance of being authoritative information. And sometimes they're accurate, but sometimes they're not. The unreliability of the information at the moment is, you know, the major impediment of using it. The fact that they (AI crawlers) can ingest content from reliable sources, and then serve it back in unreliable ways. So I think there's a lot in terms of the information ecosystem while it's still so unregulated and uncertain. I think there's a lot of dangers." Lenore Taylor, Editor, Guardian Australia

BITING THE BULLET? NOT QUITE YET …

Editors and product teams described how they are thinking about generative AI as it applies to their editorial processes, including research, writing, newsgathering and production, as well as audience development and distribution. The first clear observation to make is that deploying generative AI to write news articles without human oversight is not something newsrooms are considering in this early phase.

"Integrity is so important to the journalism that we do. So, I struggle to see, as far as our storytelling goes, that we will be doing much with it (generative AI) for a little while, just because we're not ready. Integrity is very, very important. However, I think it would be very unwise to ignore it, as well." Lisa Allan, Editor, Newcastle Herald

"The general policy is we don't want journalists using ChatGPT for their journalism." Justin Stevens, Director, ABC News

"Our focus is enterprise journalism, original journalism. We're not doing company results. And I just cannot really see how in the short to medium term that will be replaced (by generative AI)." Bevan Shields, Editor, SMH

In that context, the comments made by News Corp chief Michael Miller in August 2023 about the use of generative AI to weekly produce some 3000 weather, traffic and finance reports took many by surprise, including News Corp staff, who promptly asked the journalists' union, the Media, Arts and Entertainment Alliance (MEAA), to seek clarification from the company. Mr Miller later said that generative AI was not being used, and that it would be more accurate to describe the articles as providing 'service information', an area where using AI is 'nothing new' (Buckley & Wilson, 2023).

AI is not new to Australian newsrooms. Most are operating with some level of non-generative AI in their systems already. This is usually found in audience-personalisation systems, content-recommendation algorithms, automated-distribution systems, and image and speech recognition. And most news organisations see potential opportunities to apply generative AI in similar backend functionality: workflow optimisation, news distribution and audience recommendation rather than editorial output.

This is not because organisations are not thinking about the editorial implications of generative AI. It is clear they are. In our discussions, editors discussed how the technology, even in its nascent form, is throwing up many critically important questions that need to be considered and answered before news organisations will consider formally deploying it into their workflows.

"I would not want to be an editor who's necessarily run out in front on using AI to produce copy. I think that's quite a leap. This thing (generative AI) only really appeared this year; probably to be producing news stories out of it is slightly scary to me. I think there are other applications and uses for it before we went straight to copy." Bevan Shields, Editor, SMH

Editors are also curious about how generative AI will apply not only to the practice of journalism but also the broader information ecosystem in which journalism exists. Most organisations are already testing the technology but remain
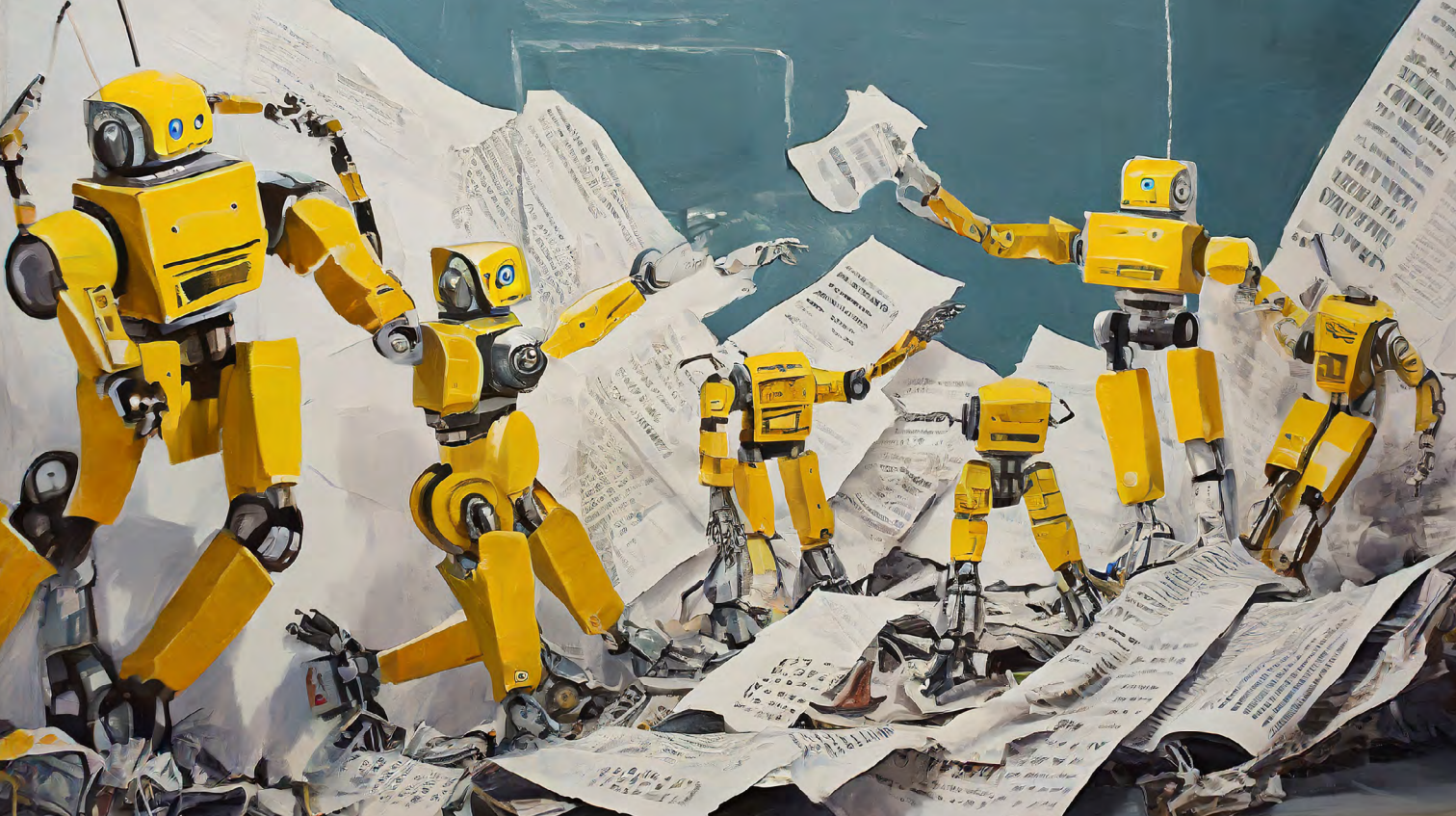
wary of deploying it until they have fully considered the consequences of doing so in this early phase of the adoption cycle.

For example, every editor was aware of the limitations of generative AI, including inaccuracy and 'hallucinations', and its inability to take account of context, to identify facts or to weigh evidence.

 "We have sent at least one email to staff making clear our current position on it, which is to proceed with caution and that we don't want to use it in our news or journalism without us knowing about it. We are extremely interested in the potential opportunity but we are not at a point where we're using it or deploying it." Justin Stevens, Director, ABC News

However, editors also recognise the appeal for generative AI to optimise newsroom workflows and reduce costs.

"It's something that can't be ignored by the industry, and particularly in regional media, where we have very limited resources compared to other news organisations. We want to be able to serve our communities with all of the information that they need and want. And we really struggle to stretch those resources every single day to achieve that. So anything that can help us with that is always going to be something that we would look at." Saffron Howden, National Editorial Training Manager, ACM

Test-and-learn approach

Most organisations we spoke to have adopted a 'test-and-learn' approach where journalists can experiment with generative AI tools like ChatGPT or OpenAI's DALL-E, but only when the resulting product has human oversight and, in some cases, has been upwardly referred for approval. All these newsrooms indicated they have restricted experimentation to production uses, such as creating images, rather than in editorial.

"We are doing some mild experiments within teams in the newsroom but we're not doing this in an official way, in terms of using ChatGPT in our workflows – and experiments have been for production, rather than editorial uses." Erin Reimer, Senior Editorial Lead for digital news and current affairs, SBS

"If we were using it for idea generation, you know, give us five titles for a podcast episode title speaking to women under 35. The attitude at the moment is give it a try, use it and let us know how you're using it, and what you're using it for." Melanie Withnall, Head of News and Information, SCA

On the whole, this experimentation is aimed at improving workflows and efficiency, saving time for journalists by automating repetitive and menial tasks. Editors and product teams are focused on finding solutions to their existing workflow pain points – and there are many, from SEO, meta tagging and reformatting for multi-platform publishing.

"We're experimenting and looking at it and we've not made any decisions yet about whether we will or won't use it. It's very clear that it won't ever replace journalists when it's not reliable. We are thinking about different ways where it might take away, or help with menial or repetitive or time-consuming tasks, like summarising huge amounts of information." Lenore Taylor, Editor, Guardian Australia

"We've got a tech team and we're looking at it. We're looking at it all the time, mostly for optimising workflows." Barclay Crawford, Editor, Daily Mail Australia

One newsroom recently held a hackathon and came up with a solution that saved an enormous amount of time for image editors. Using machine learning, the team developed a tool to search and identify emotions in photos. They trained it on their proprietary image library then integrated it into their content management system. It delivers generous time savings when journalists are looking for particular emotions in photos.

"So you can search 'Anthony Albanese', 'happy', 'upbeat', or 'frustrated', or whatever. The AI got quite good and quite quick at selecting these photos that match the description. I think that kind of use case is really helpful. Because it speeds up the process of me looking through thousands of Anthony Albanese photos, so it's helping me do that job, but it's not generating an image." Chris Paine, Head of Premium Content, Metro Publishing Nine

Newsroom leaders and editors pointed out that building and maintaining trust with the audience guided their decision-making about if and how they would deploy ChatGPT.

"For us, the key thing that will be top of mind every step of the way, is safeguarding and retaining trust. And then that being at the centre of every decision we take in relation to generative AI, but that not being at the cost of potential efficiencies or things that could actually help the audience." Justin Stevens, Director, ABC News

"Quality is a fundamental part of the value proposition for a subscriber, you know, so anything that has the potential to mess with quality, introduce errors or doubt around quality, authenticity, accuracy or the origins of a story is a real threat to the brand proposition that we have." Bevan Shields, Editor SMH

Some news organisations are more structured than others in this early test-and-learn phase .

"We're a bit of an informal company, so we have a group that's working on learning these tools, implementing these tools, and it is a collaboration between content and product." Melanie Withnall, Head of News and Information, SCA

EVERY EDITOR WAS AWARE OF THE LIMITATIONS OF GENERATIVE AI, INCLUDING ACCURACY AND HALLUCINATIONS, AND ITS INABILITY TO TAKE ACCOUNT OF CONTEXT, TO IDENTIFY FACTS OR TO WEIGH EVIDENCE.

Product innovation

Each news organisation expressed enthusiasm about what generative AI could unleash in terms of product innovation, but some are more focused on exploiting opportunities in this area.

"We've got a fantastic product and technology team at the moment. And so I think there is a good culture of road testing where there's capability for deploying generative AI. So that might be everything from how we can deploy the metadata more powerfully or front-facing audience discovery." Justin Stevens, Director, ABC News (edited for clarity)

"We have a huge technology, product and digital team here, obviously, because we make our own app. We are really trying to understand how the tech works, what we might build in house, what we might use, what we might license, etc., etc. And so, we're really in that testing phase. And I think we're almost getting to the point where we'll be putting those policies together." Melanie Withnall, Head of News and Information, SCA

Synthetic voice

Improvements to accessibility were also identified as an early opportunity. The ABC is exploring how to make their predominantly text-based news service more accessible to people with vision and hearing impairment. Some virtual voice products are being piloted that enable the user to listen to the news article via a synthetic voice – the artificial production of human speech – rather than reading it. At the same time, the ABC has developed its own product and, in the process, discovered that this product is significantly more accurate because the AI is trained on the ABC's archive rather than the internet. As a result, the synthetic voice has become familiar

with local names of cities and towns, as well as Indigenous words. The off-the-shelf products trained on larger global databases were not always able to accurately capture Indigenous words and at times delivered offensive results.

"The examples which turned up during the pilot highlighted the need for a human to be at the beginning and end of any process using AI." Angela Stengel, Head of Innovation and Digital Content, ABC

Another opportunity identified by the ABC was about how it might use generative AI's translation capabilities to better service Australians who identify with more than 270 Aboriginal and other ancestries. The ABC is piloting synthetic voice translation that would enable it to publish news articles in multiple languages. The concept is very much in its 'test and learn' phase.

"The ABC is piloting synthetic voice for a range of applications, including to support accessibility considerations. While the potential is there to service audiences in their preferred language other than English with ABC News, connecting with new communities to build that audience requires more than AI technology." Angela Stengel, Head of Innovation and Digital Content, ABC

Southern Cross Austereo is looking at how synthetic voice could be deployed for seemingly simple information services like short weather reports that take a human being significant time to produce.

"If you think about a weather report, which literally goes for less than 10 seconds on air – and we're mandated to do it – we're just saying it's 26 degrees and sunny, and it will be 26 degrees and sunny for the entire week. We are not talking about a developing situation like a cyclone coming into Cairns. It's 26 degrees and sunny, so a very short sentence. But it takes a lot of someone's time. They have to go to the Bureau of Meteorology, they have to write it, put it in the system, they have to then record it, then they have to upload it into our playout system. So there are actually quite a lot of touchpoints. Whereas if you could automate that process, and you've got 99 radio stations, you could be saving a good couple of hours of someone's time. Then they can spend that time doing the thing that the tech can't do, which is go and speak to someone in the community and

develop an original story." Melanie Withnall, Head of News and Information, SCA

Only a few news organisations we spoke to were in the early test-and-learn phase of synthetic voice. Building an accurate synthetic voice product to serve non-English speaking users or those with vision or hearing impairments is one thing. Building an audience around this offering is another. And newsrooms are acutely aware of any impact it might have on audience trust or reputation.

"So all of this piloting has been happening in earnest … Where there's not much upside, we just park it, or if the pilot demonstrates we could not do this potentially because it would compromise our editorial policies or our general framework around making sure we don't do anything that brings into question trust." Justin Stevens, Director of News, ABC

PERCEIVED OPPORTUNITIES

Aside from the early test-and-learn experiments currently underway in newsrooms, editors are considering other opportunities with generative AI.

"This is the first time … that I've seen something that really has some good applications for journalism … but we're not going to truly know until we really spend some time experimenting with it." Melanie Withnall, Head of News and Information, SCA

There is also a lot of early optimism about what generative AI might deliver in terms of new formats.

"I'm excited about the prospect of how generated imagery could help enhance the storytelling experience from a production point of view, subject to there being clear guardrails." Justin Stevens, Director, ABC News

SBS has undertaken a business-wide sprint to consider a raft of AI solutions in such areas as translations, video production and publishing workflows. Data visualisation, in particular, is an area where it is looking at possible opportunities for generative AI, particularly to improve existing third-party tools.

"We are looking at opportunities for visualisation of data. Those that would be much faster than the tools we currently have and those which would provide a much more contextual experience for the audience." (edited for clarity) Erin Reimer, Senior Editorial Lead for digital news and current affairs, SBS

Optimising workflows and content management systems

For the newsroom editors we spoke with, gaining backend efficiencies is the most immediate opportunity presented by generative AI. Editors repeatedly said they hoped to automate more of the monotonous, low-value, repetitive tasks to enable journalists to engage in the higher-value, original and creative journalism work.

"When you look at the last decade, and we've become a multi-platform news provider, we've added a lot to the expectations around people's skills and their output. And where we need to reset that and try and free up time –

however possible – to make sure we safeguard the quality of what they do, but also make it sustainable so staff aren't run ragged. So, you know, where there are tools to free them up to focus more on journalism, which goes to sort of functional tedious workflows, which doesn't compromise the information." Justin Stevens, Director, ABC News

Common examples given were automating SEO, meta-tagging, translations, and podcast summaries.

"I'm looking at how (generative) AI can help us, or prompt a journalist with an SEO headline, based on best practice, so that they don't have to spend 15 minutes thinking about it, you know, stuff like that." Saffron Howden, National Editorial Training Manager, ACM

"A lot of the tasks that our teams do are quite low value but take a lot of time. So we want to free people up to be more creative, to do more things that are more exciting to create greater value for the business." Melanie Withnall, Head of News and Information, SCA

Audience personalisation, distribution and analytics

Most newsrooms have already embedded third-party, non-generative AI tools such as Social Flow and Echobox into their newsroom workflows. These third-party tools offer a range of uses from social media analytics to distribution. Newsrooms are keeping a close eye on opportunities that might arise as generative AI is integrated into these existing tools.

"When we have a fantastic story, I'm interested in how we can make sure that we make that story relevant and accessible to our Gen Z audience on social media, as well as the more rusted-on audience on our homepage, and then our newsletter audience and all that kind of thing." Saffron Howden, National Editorial Training Manager, ACM

The different business models of the various media organisations influence how they are thinking about the opportunities of personalisation. News organisations that have a subscription model are less interested in using generative AI to enhance their personalisation capabilities, whereas those that do not are interested in how generative AI might help them better understand audience behaviours and habits to help them reach the target audience at the right time.

"Anything that you can do that will help surface stories that are of value to listeners that are interesting for them, that helped them in their lives, yes we would absolutely look at doing it." Melanie Withnall, Head of News and Information, SCA

"It would potentially be good for audience development if we can use it to pick up and analyse search and story trends

in real time." Barclay Crawford, Editor Daily Mail Australia

The ABC viewed opportunities in personalisation through an accessibility lens.

"There are interesting trials (mostly with non-generative AI) around the potential accessibility upsides. Things like captioning and translation, for instance, automatic audio of the written form and stuff like that. So all of that piloting has sort of been happening in earnest and we'll just figure out where is there an audience gain or where it could potentially compromise our editorial policies or our general framework around making sure we don't do anything that brings into question trust." Justin Stevens, Director, ABC News

Using generative AI in Research

Non-generative AI has been used for research in journalism for several years now; however, some editors are already considering the possible opportunities for using generative AI for deep research in investigative journalism. For example, some identified an opportunity for generative AI to improve existing tools that enable journalists to interrogate large data sets and scour open intelligence sources, or data visualisation tools that help journalists visualise and interrogate large-scale document leaks.

Other editors spoke about pattern recognition and the abilities of generative AI to surface a line of inquiry that would otherwise be near impossible to spot. Again, these capabilities of generative AI are considered a potential opportunity because of the capacity to save enormous amounts of time for journalists who would otherwise have to trawl through and organise large datasets and documents.

EDITORS REPEATEDLY SAID THEY HOPED TO AUTOMATE MORE OF THE MONOTONOUS, LOW-VALUE, REPETITIVE TASKS TO ENABLE JOURNALIST TO ENGAGE IN THE HIGHER VALUE, ORIGINAL AND CREATIVE JOURNALISM WORK

"The ability to be able to scrape publicly available information more effectively, and identify where there are gaps, recurring themes or a pattern that would otherwise take, you know, months for an investigative journalist to do. I don't think that AI necessarily gives us an answer straight away, not one that we could publish, but it might help identify or connect things that warrant proper investigation. And, you know, human-led investigation." Justin Stevens, Director, ABC News

PERCEIVED RISKS

Most news organisations have chosen not to deploy generative AI to write news articles because they say they are focused on the possible risks to the news industry generally and the broader information environment rather than internal editorial risk that might arise.

The one area where news organisations are alert to editorial risk is verification. As generative AI supercharges the proliferation of sophisticated deep fakes, including in imagery, the problem of verifying source material presents a clear and present editorial danger.

Verification and deep fakes

User-generated content is a growing source of material used for news stories; in particular, the smartphone era has enabled people to capture critical moments on camera before journalists are deployed to the scene. These images have important news value, but need to be verified before newsrooms can publish them. This can be problematic in the cut and thrust of the 24-hour news cycle where there is pressure to break a story.

Now, with the arrival of generative AI, editors are concerned that existing verification tools, such as Google and Bing reverse image search, are no longer effective or efficient.

Our interviews revealed that most newsrooms are grappling with the fear of heightened verification challenges, particularly given that deep fakes are more sophisticated, harder to identify and being published at a pace and scale that the news industry has not previously experienced.

"I think this is where it's anxiety-inducing, because it does potentially bring into question the authenticity of a lot of things. So particularly in an investigative capacity. We do deploy a great deal of training around identifying misinformation and authentication of materials, but I think it will have to go up a notch as soon as possible with generated imagery. When will be the first big, deepfake piece of news where the media is tricked into thinking something was real when it wasn't?" Justin Stevens, Director of News, ABC

"Maybe you can use AI to identify AI? I don't know, but we need tools to figure out if it's real or not because convincing fakes are really worrying." Lenore Taylor, Editor, Guardian Australia

In some newsrooms the arrival of deep fakes has dramatically slowed the processes of verification.

"I reckon it (generative AI) has slowed everything down." Barclay Crawford, Editor, Daily Mail Australia

Already, one news organisation has experienced several close calls, pulling stories just minutes before publication after discovering the supplied images were deep fakes. This occurred despite the newsroom following existing editorial and verification checks. The images were almost published because the verification tools – including reverse image searches on both Bing and Google – failed to pick up the fact that the image had been tampered with.

As a result, news organisations are feeling more vulnerable. Their primary concern is that they, along with the audience, could be caught out. The lack of confidence in a newsroom's ability to identify deepfakes is new and adds legal as well as editorial risk.

"There have always been attempts to hoodwink the media with fake images and video. But now deep fakes are better and our current tools like reverse image searches aren't picking them up. The possibilities of AI hoodwinking the media are now limitless and the fakes are coming at us all the time. It's something we all need to be aware of and to remind our staff about. It's an old saying, but it's still as relevant in 2023 as ever: If an image or video seems too good to be true, it probably is." Barclay Crawford, Editor, Daily Mail Australia

Deep fakes and media personalities

There is also concern about high-profile journalists and media personalities increasingly becoming the subject of deep fakes. Aside from potential damage to personal reputations, news organisations fear that the public's trust in their high-profile media personalities, warranted or not, could also decline as a result, and that they might suffer organisational brand damage themselves.

"Trust is the key to building audience relationships with shows and talent. There is a huge risk with deep fakes for journalism, but also for entertainment. Think of creators like Christian Hull who put their life online in their podcast, there's hours of content that could be utilised. That's a risk to the business but more importantly to the creator, and the audience. Especially when you think about cloned voices, and how much we invest in creativity. I think younger audiences are good at spotting a fake, but the fakes are getting better." Melanie Withnall, Head of News and Information, SCA

"A large part of our media culture is driven by personality and high-profile individuals. And given how easy it for people's images and voice to be faked, in that sense, that's a massive concern." Justin Stevens, Director, ABC News

The proliferation of deep fakes will continue to be challenging not only for individual newsrooms and the news industry, but also for society as a whole. Deep fakes create confusion about important issues, can target high-profile journalists and

make it more difficult for audiences to decipher what's real and what's not.

Lack of editorial control in AI-automated distribution tools

Most organisations are currently using distribution tools that have non-generative AI components. These tools will likely increase in sophistication as generative AI capabilities are integrated into them.

A problem with AI-powered distribution tools is that the AI algorithms are not transparent about how they make editorial distribution decisions. In many cases, editors do not know how or why AI is choosing to send out a particular story at a particular moment in time.

The SMH made an observation about its experience with an AI-powered third-party distribution tool which highlights this risk.

"One of our third-party platforms that we use for

distribution actually has an AI component, and a couple of years ago it was introduced, and it was pulling, like lead stuff for the sake of optimising content. But it doesn't really explain how the algorithm or the AI works. It's just what the AI has selected as the best sell for a story. And it kind of just shows how important it is to have a human editor or producer contextualize the story and actually say, oh, yeah, that line is good. But for the nature of the story, or its developments, in isolation it doesn't make sense, or it actually could carry legal risk." Sophia Phan, Growth Content Editor, SMH, The Age, Brisbane Times, WAToday

Copyright and intellectual property rights

News organisations are understandably keen to protect the intellectual property rights attached to their original reporting. Guardian Australia and ACM have already switched off the Open AI crawlers on their websites. SBS has taken the decision to block a small selection of AI crawlers.

Editors all flagged deep concern about copyright and

intellectual-property issues that the arrival of generative AI presents. Many news media organisations have legal teams investigating possible breaches and some are members of organisation-wide taskforces for generative AI that are interrogating legal implications with in-house legal specialists.

Loss of Jobs

Editors are primarily looking at how generative AI might augment the practice of journalism rather than replace the work of journalists.

"We're looking at this from a task and not a job perspective, you know, in terms of how we might use AI, from an operational point of view. It's more about what are the tasks that we can potentially look to improve with AI? And therefore, how can we enrich the roles that we have? I'm speaking very specifically about our journalists, and our newsroom units." Erin Reimer, Senior Editorial Lead for digital news and current affairs, SBS

Nevertheless, there is anxiety about job security, particularly while the impacts of the technology remain uncertain.

"As an editor, I've spent a lot of my time, rather than contemplating the good and the bad, sort of managing some people's anxiety about this in the newsroom. ... Younger reporters in particular are more anxious than excited about it." Bevan Shields, Editor, SMH

We've got people across our organisations, but I'm sure across all media, that worry about jobs, and probably across all industries, that are worried about their jobs. We keep saying that human oversight is still so important and there's things that it can't do, there's those relationships that it can't have and maintain. There's a lot that it can't do. But I am aware that there is an awful lot that it can do as well." Lisa

Allan, Editor, Newcastle Herald

Disparity between larger and smaller newsrooms

Product development with the current wave of generative AI will likely be more challenging for smaller newsrooms, which do not have the capacity to hire the deeply skilled teams required to develop, test and deploy AI products. Indeed, our conversations with mid-sized national newsrooms revealed they also feel this challenge. Local news organisations with a large, international head office felt well placed by comparison.

"I think I'm fortunate to be part of a big, global organisation, because the sorts of resources that you need to put into this and the sorts of expertise that you need to actually know what you're doing would be impossible for me to do here locally. Like, I just couldn't do it. And I think we're as active as we can be, I think it's just moving really quickly. It's moving so quickly that I don't think anyone could say, 'Yeah, I'm confident we've got this.'" Lenore Taylor, Editor, Guardian Australia

The loss of trust in news and information

Along with concerns about copyright and intellectual property, the risks generative AI poses to the broader information environment are significant for the editors we interviewed, who were quick to weigh the opportunities against a backdrop of an even more-polluted information ecosystem.

Most editors agree there is an opportunity for news organisations and audiences to rediscover the value of news, but they are far less confident that this opportunity will translate into journalism and the pursuit of truth winning the

EDITORS ARE PRIMARILY LOOKING AT HOW GENERATIVE AI MIGHT AUGMENT THE PRACTICE OF JOURNALISM RATHER THAN REPLACE THE WORK OF JOURNALISTS

day. The increased burden placed upon audiences to wade through a flood of junk content and misinformation produced by generative AI is a constant theme.

"I am more concerned about the dangers it poses for the news ecosystem as a whole. That's my major concern. But then, the more I look at it, the more I can see that, yes, maybe there are opportunities for how it might help us do journalism. But I think that threat is probably, you know, bigger, in my mind." Lenore Taylor, Editor, Guardian Australia

Editors understand that well-crafted and precisely targeted disinformation remains a powerful force, especially around big news events like the Israel–Gaza war, elections and referendums, including the recent Voice to Parliament referendum in Australia. Editors reaffirmed their commitment to delivering fact-checked and verified news whilst acknowledging this may not be enough to cut through a digital information ecosystem overrun by junk content, misinformation and sophisticated deepfakes.

One scenario related to political photojournalism: if, for example, an organisation had a photo of a political leader in an historic moment and the photo had been ingested by AI crawlers, the photo could re-appear in a convincing AI-generated mash-up that told a different, inaccurate story about the moment in history. The AI-generated image might be of such quality that audiences assume it is genuine.

"I think that's a problem. I think it's a real problem. I'm just not, I'm not across how we would go about fixing it. But obviously, it's a problem." Lenore Taylor, Editor, Guardian Australia

STRATEGIES TO MITIGATE EDITORIAL AND ETHICAL RISKS

Many organisations we spoke with were in the process of developing formal guidelines and policies governing how their organisations would engage with generative AI, including for legal issues such as copyright.

AI guidelines, principles and policies

Discussion about guidelines and policies strongly reflected disparities in newsroom size, experience with AI and the organisation's business objectives.

The Guardian parent group in the UK is the most advanced with the development of guidelines. In June 2023, The Guardian UK publicly released its 'Approach to Generative AI', the result of months of work by teams across the organisation including from editorial, creative, engineering, product, legal, commercial and partnerships (Viner & Bateson, 2023).

The Guardian UK established a small, experimental team — led by the Head of Innovation, Chris Moran, with developers, data scientists, and product managers — looking at how the organisation might responsibly use AI. This team made recommendations to a broader Guardian taskforce that includes editorial leaders, legal units and those with syndication expertise. This group is exploring how The Guardian might adopt generative AI in particular contexts. Australian editor Lenore Taylor is on it.

"The Taskforce group, they're working, they've got collaborations with various external people, and they're doing specific projects to look at specific types of potential uses, and then it will go back to ExCo (The Executive Committee of The Guardian) and then there will be decisions made about what we will or won't do." Lenore Taylor, Editor, Guardian Australia.

SBS and ABC are also in the process of bringing together multi-disciplinary teams to draft guidelines and principles on the use of generative AI.

"It's been a multidisciplinary group, with our codes team, legal team, news and current affairs representatives, technology and so on." Erin Reimer, Senior Editorial Lead for digital news and current affairs, SBS

The ABC has deployed cross-functional executive teams with expertise in legal, business, product development and editorial to work on a policy framework.

"A framework and AI policy is in development for where the ABC stands on it. So we're covering everything from ethics

Human oversight

All news editors we spoke with mentioned the requirement for human oversight as a non-negotiable, first principle for guidelines governing the use of generative AI. Almost unanimously, editors rejected the idea of replacing journalists with generative AI to write any copy beyond low-value outputs such as podcast summaries or basic weather reports. And even then, the need for meaningful human oversight in any application of generative AI in the editorial process, even for low-value content, is considered paramount. Editors also stressed the importance of critical thinking and reasoning as essential journalistic attributes that only a human journalist can perform.

"If I want to summarise a podcast, it still has human eyes go over it. You can't publish anything that doesn't have two sets of eyes on it." Melanie Withnall, Head of News and Information, SCA

Transparency

Throughout the discussions on guidelines and policy frameworks, we observed a consistent commitment by news organisations to ensure their organisation is clear and transparent with audiences about how it is using generative AI, now or in the future, particularly in the production of news.

"Our objective with guidelines is also to be transparent with our audience. And I think that's a really important thing as we, as an industry, approach AI. We have to be honest and upfront with our readers and our audience about what we're doing and how we're doing it. So that sort of underpinned the code of practice that we now have in place." (edited for clarity) Saffron Howden, National Editorial Training Manager, ACM

"Whenever it (generative AI) is deployed, for whatever reason, for any piloting, we will need to be and continue to be extremely transparent, internally and externally, when it's played a role and explain what role it's played and what the potential limitations are of it." Justin Stevens, Director of News, ABC

The commitment to transparency was often anchored within the broader issue of how newsrooms maintain trust with their audiences at a time when it will become more difficult to establish what is real and what is fake.

"It's so important that we bring our audiences with us on this journey. So they know every step of the way, this is where we're including it. This is how we're using it. You know, I think that transparency is going to be key not just for us personally, but the whole industry, so that we don't further erode trust." Saffron Howden, National Editorial Training Manager, ACM

"So at the risk of stating the obvious, it becomes incredibly important for any use of it internally, or for any outward-facing purpose, that whenever it is deployed, for whatever reason, even for any piloting for technological reasons, we will need to be and continue to be extremely transparent, internally and externally." Justin Stevens, Director of News, ABC

Whilst each newsroom was clear in a commitment to disclose if and when generative AI is used to create content, the discussions revealed differing approaches to the level of detail necessary in such disclosures.

Some editors adopted an absolutist approach and declared the need to go beyond communicating merely when generative AI is used in the production of a story

---

through to journalistic implications and the implications for editorial policies. Obviously, legal loomed large in that, as do product and technology. So there's a lot of work underway and we're trying to get alignment from all of those different perspectives, from a policy point of view, in terms of how it might be deployed in terms of product, or programming and journalism." Justin Stevens, Director of News, ABC

The SMH was unequivocal in stating that generative AI adoption in its editorial output is not a top priority because it did not align with its strategic business objectives. Instead, its interest in generative AI is around how it might optimise workflows and free journalists from menial, low-value tasks. Even so, the SMH is thinking about how to bring diverse teams together to think more strategically about how to adopt the technology and create a policy framework to govern the use of generative AI.

The Newcastle Herald's approach to AI falls under the ACM-wide AI Code of Practice which has been developed centrally and covers all ACM mastheads and employees. This code does not cover business objectives or strategy.

"Our company has recently done an AI, like, policy. So that sort of only came out in the last couple of months. That sort

of tells us what's acceptable and what's not an acceptable use of generative AI." Lisa Allan Newcastle Herald

The majority of newsrooms are planning to make their guidelines publicly available for reasons of transparency and to build trust with their audiences.

The SMH has already updated its freelance and contributor policy to make it very clear that the organisation would not accept any copy generated with AI. There is a clear business reason not to pay for material that is not written by a human.

"You can see very quickly where there could be disastrous consequences with that. Also, paying someone XXX dollars per word for something they are not actually be writing. So there's that too." Bevan Shields, Editor, SMH

Some themes surfaced in discussions about guidelines that offer insight into how news organisations are approaching their policy frameworks. Consistent themes being considered in policy and guidelines discussions are considered in what follows.

WHILST EVERY NEWSROOM WAS CLEAR IN A COMMITMENT TO DISCLOSE IF AND WHEN GENERATIVE AI WAS USED TO CREATE CONTENT, THE DISCUSSIONS REVEALED DIFFERING APPROACHES TO THE LEVEL OF DETAIL NECESSARY IN SUCH DISCLOSURES.

to communicating if generative AI is used to augment the practice of journalism. For example, if generative AI is used to assist in early research or brainstorming ideas for a particular story or headline.

Other editors questioned if it is necessary to disclose to the audience that generative AI was used for a podcast summary or to generate weather reports.

"Would I declare that a podcast title was created by the production team with the assistance of Chat-GPT in a brainstorm session, I don't know, probably not. But long-form investigative journalism or an article? Yes, I would declare if we used generative AI." Melanie Withnall, Head of News and Information, SCA

Training

Training is considered an important component of any guidelines as well as a stand-alone effort necessary for newsrooms to navigate and mitigate some of the risks generative AI presents.

Guardian Australia recently held online training sessions with its head of innovation in London and the Australian team. The organisation has regular emails updating staff about the most relevant developments in generative AI to help bolster general understanding of the technology and its application in the news environment.

"Chris Moran has done an online education session for all staff across the organisation. I sit in on it. There's also sort of a weekly email just going out to senior editors about here's all the developments in AI this week, because goodness knows, every week there's a lot. There's sort of this deep work going on in the UK, which I know about, have some oversight

in … *I don't have time to personally be sort of deeply involved in it. But I know the direction of travel."* Lenore Taylor, Editor, Guardian Australia

Most editors we spoke with identified the need for training on the legal issues raised by the use of generative AI. Some organisations are in the process of updating their existing legal and copyright training. Editorial policies are also being reviewed, particularly on verification and sourcing.

However, no Australian news organisation has yet developed bespoke training for how journalists might use generative AI to augment their work. This is most likely because newsrooms are still in a test-and-learn phase and not yet clear on how generative AI might be adopted and deployed.

Maintaining trust in news

Most editors we spoke with feel that generative AI will almost certainly place further strain on public trust in the news media. Gains in trust throughout the COVID-19 pandemic have since been reversed (Newman et al., 2023).

When considering what can be done to mitigate the harms that generative AI might inflict, which could further erode trust in news media, some editors feel that timeless editorial policies and practices could be adapted to this wave of disruption.

"Sometimes the tried and tested things are the best. So … if we are entering a period where people are going to have doubt or want to know where things originated from, then, you know, a good old sourcing policy never goes astray. We have a sourcing policy for copy. Maybe we need to look at the sourcing policy and how we disclose the source of the material that we're publishing." Bevan Shields, Editor, SMH

"When news started breaking on Twitter a few years ago, bombings, terrorism attacks, (we had to figure out) how to validate that information. There were processes like, when did this image originate? Have you reverse searched it? Is this account authentic? How long has this account been active? Have there been any other similar videos? I guess, Journalism 101, and [these processes] are still applicable to AI. So I think it's just a different sort of medium." Sophia Phan, Growth

Content Editor, SMH, The Age, Brisbane Times, WAToday

Some also suggested that consumers would eventually learn how to establish which news sources to trust in an era of generative AI.

"I think it's a good opportunity to educate people how to spot the difference between what's true and what's fake, and what sort of brands to associate trust with." Sophia Phan, Growth Content Editor, SMH, The Age, Brisbane Times, WAToday

Concern over possible further pollution of the information ecosystem is tempered by cautious optimism about the opportunity for news mastheads to reassert their core brand as a trusted news source. Indeed, we observed broad agreement amongst editors that in an era of AI and the accompanying deep fakes, quality journalism could experience a renaissance where fact-based, verified news becomes an increasingly rare and increasingly valuable commodity.

"In many ways, you know, it's a huge opportunity for journalism … that these kinds of trusted mastheads is where you can go for truth and for facts." Lenore Taylor, Editor, Guardian Australia

"Outlets like the ABC have the ability to stand out in coming years where there's going to be more information overload for the public, we also know there'll be an increase in deep fakes and misinformation as a result. And so, I think the more that we can define how we stand out in that climate, you can come to us to read, listen to or watch us and you know, you can trust every single thing we say because we've done the legwork to establish and confirmed what's occurred." Justin Stevens, Director of News, ABC

DISCUSSION

Australian newsrooms are taking early steps to prepare for the disruption that generative AI will create.

Editors are cautiously optimistic given they have all experienced successive waves of technological disruption over the past decade. Most editors acknowledge that it is too early to be prescriptive about how generative AI will reshape

the news industry, but they are thinking about the immediate challenges and opportunities that generative AI presents.

Each editor we spoke with understands the opportunity for news organisations to stand out in the broader information ecosystem, which they believe will be filled with more low-quality news-style content, making it even more difficult for the audience to wade through and find fact-based, verified news. In this increasingly polluted information environment, traditional news mastheads see a potential opportunity that their value as a trusted source of fact-based, verified information might increase. Some editors are confident that their 'brand of news' will thrive. Others are less confident.

Optimising workflows

To connect with readers, news organisations have long known they need much more than a great story. They need to understand the habits of audiences on each platform. They

need to shape their story in formats that work best on each platform and understand the ever-changing nuances of social media algorithms.

Most news organisations now have distribution teams. Publishing a story in a digital, multi-platformed environment has added a lot of menial tasks to a journalist's workload. Every organisation we spoke to is motivated by the goal of getting their journalists back to doing more journalism and less content management.

Optimising workflows is the clear upside for every news organisation. Some will investigate commercialising products developed in house, but attracting the necessary technical skills and AI expertise could prove difficult in a market where generative AI skills are scarce and demand is high. Partnerships with technology companies could emerge as a result. News organisations we spoke with seemed open to the idea of partnerships, but it was not something they had considered in any great detail at this stage.

Deep fakes and the verification challenge

The challenge of verification is the immediate, primary concern for news organisations. The proliferation of deep fakes will continue to be challenging not only for individual newsrooms and the news industry but for society as a whole, and is likely only to worsen. Deep fakes create confusion about important issues, can target high-profile journalists and make it more difficult to decipher what is real and what is not.

Previously, it was common and reasonable for news organisations to see their role as helping audiences navigate false and deep-fake content. However, the arrival of generative AI has increased the sophistication of deep fakes to the degree that news organisations no longer have confidence in their current tool kits and capabilities to identify them.

What is surprising is how vulnerable the larger metro and national news organisations feel in their ability to combat deep fakes. However, the Newcastle Herald felt less vulnerable to being caught out by deep fakes given the small communities in which they operate and their close, personal, human-to-human relationship with their readers.

In recent testimony to a US congressional hearing into AI, Sam Gregory, Executive Director of WITNESS, a nonprofit organisation that helps people use video and technology to protect and defend human rights said, "In the direct experience of my own organization in analysing high-profile suspected deepfakes encountered globally, it is challenging to support rapid, high-quality media forensics analysis; detection resources are not widely available to the media or the public; and the gap between analysis and timely public understanding

is wide and can be easily exploited by malicious actors (Gregory, 2023). This gap was front of mind for each editor we spoke with, and raises a direct challenge to their role of sorting fact from falsehood in the pursuit of public interest journalism.

The pathway forward for the news industry with regard to deep fakes is far from clear or settled. As discussed in Chapter 2, some experts say debunking deep fakes is futile, that digital watermarking will likely become easy to erase and that the news industry should focus on verifying what is real rather than what is fake.

International efforts to combat deep fakes include Project Origin, which was established in 2018 by the BBC, CBC/Radio Canada, The New York Times and Microsoft. In 2020, roughly two years before the release of ChatGPT, Project Origin partnered with the Adobe-led 'Content Authenticity initiative' to establish the Coalition for Content Provenance and Authenticity (C2PA). C2PA is building an end-to-end solution in the form of a technical standard that identifies the provenance and authenticity of different types of media.

Internationally, the BBC announced its new BBC Verify project in May 2023 as a response to audience requests for more transparency about how they verify their stories (Turness, 2023). BBC Verify is made up of a highly specialised team of around 60 journalists with a range of forensic and specialist open-source intelligence skills that enable them to go beyond conventional newsroom techniques.

The challenge of deep fakes will likely remain until media organisations are able to gain access to more sophisticated tools and develop internal capabilities to help identify them.

THE CHALLENGE OF VERIFICATION IS THE IMMEDIATE, PRIMARY CONCERN FOR NEWS ORGANISATIONS. THE PROLIFERATION OF DEEP FAKES WILL CONTINUE TO BE CHALLENGING NOT ONLY FOR INDIVIDUAL NEWSROOM BUT FOR SOCIETY AS A WHOLE...

With many elections scheduled in 2024, including one of the most consequential, the U.S Presidential election, we will likely learn if and how malicious actors will use generative AI to influence voters and how effective deep fakes are in the context of elections.

Sam Gregory told the US congressional hearing, "Politicians will claim that real audio is faked and put the pressure on fact-checkers to debunk this claim, when they lack the tools or the speedy capacity to do this" (Gregory, 2023).

If news organisations are to play a role in identifying and debunking deep fakes they will need to not only gain access to more sophisticated detection tools but also develop deeper in-house expertise.

Intellectual property and copyright concerns

The issue of copyright and unauthorised use of intellectual property was of deep concern for each media organisation we spoke with. Guardian Australia and ACM have already blocked the Open AI crawlers from scraping data from their websites. We perceived some hesitancy when discussing this issue, likely because most organisations are considering their legal position. Some may also have commenced negotiations with AI companies to license access to their content, as has occurred internationally.

Reports suggesting a high-profile legal tussle between the New York Times and OpenAI will be instructive for

how Australian newsrooms respond to the copyright issue moving forward. If the case eventuates and a US federal judge finds that OpenAI illegally copied articles from the New York Times to train its AI model, the court could order OpenAI to destroy its dataset and recreate its models using only work that it is authorised to use or is in the public domain (Allyn, 2023).

The issue of copyright is likely to remain unresolved and contentious until a case is determined in court or generative AI companies and Australian media organisations negotiate an agreement. As discussed in Chapter 2, one possibility, canvassed by News Corp in particular, is the expansion of the News Media Bargaining Code (NMBC) to cover AI companies or licencees. If enacted, this may prompt commercial deals such as those arrived at between digital platforms and news media organisations in the shadow of the NMBC.

Business risk and jobs

The tension between reducing costs and maintaining human oversight of editorial processes was not of major concern mostly because the news organisations we spoke to are not currently considering using generative AI to research or write news articles.

At this early stage in the development cycle, the pendulum is firmly on the human-oversight side with editors assuring concerned staff that the arrival of generative AI would not replace their jobs. Again, editors were adopting a 'test and learn' approach so the prospect of generative AI augmenting journalism workflows is a near-term possibility. It will be interesting to watch how this develops as financial pressure on media organisations to reduce costs continues. Over the past decade every news organisation has contended with sustained financial pressure as audiences consumed news online and advertising revenues moved away from news organisations to large tech companies like Google and the social media giants. This was a trend that the News Media Bargaining Code was designed to address.

We are yet to fully understand how the arrival of generative AI will disrupt the business models of media organisations. What we have gleaned through our

conversations with editors is that they are anticipating more upheaval and are preparing to adapt and protect their business models.

Editorial challenges and development of AI guidelines

Drawing the line between journalism and other content will become increasingly important for newsrooms. Many editors discussed the possible use of generative AI to produce 'service information' – content such as weather or traffic reports. Many are, after all, already using other forms of AI for this purpose. However, when we delved into various scenarios, it became less clear where 'service information' or 'information-style content' ends and journalism begins. This line is neither immediately obvious nor clear cut.

Australian news organisations are at various stages in the development of guidelines, principles and frameworks. Some of the larger organisations are taking a multi-disciplinary approach to include legal, editorial, product and technical expertise. This will require more time to establish internal alignment and could see organisations fall behind the pace of generative AI development. In Australia there is no industry-wide effort yet to harmonise the basic operating principles of generative AI. Australian newsrooms stressed the importance of being open with audiences about when and how AI is used in the newsgathering or production process. However, there is variation in how much disclosure editors believe is required. This variation reflects the international experience.

In June 2023, The Guardian UK released three broad principles governing the use of generative AI within the organisation, stating, *"If we wish to include significant elements generated by AI in a piece of work, we will only do so with clear evidence of a specific benefit, human oversight, and the explicit permission of a senior editor. We will be open with our readers when we do this"* (Viner & Bateson, 2023).

In October 2023, the BBC released its guidelines for generative AI, stating, *"We will be transparent and clear with audiences when Generative AI output features in our content and services. Human oversight will be an important step in the publication of Generative AI content and we will never rely solely on AI-generated research in our output"* (Davies, 2023).
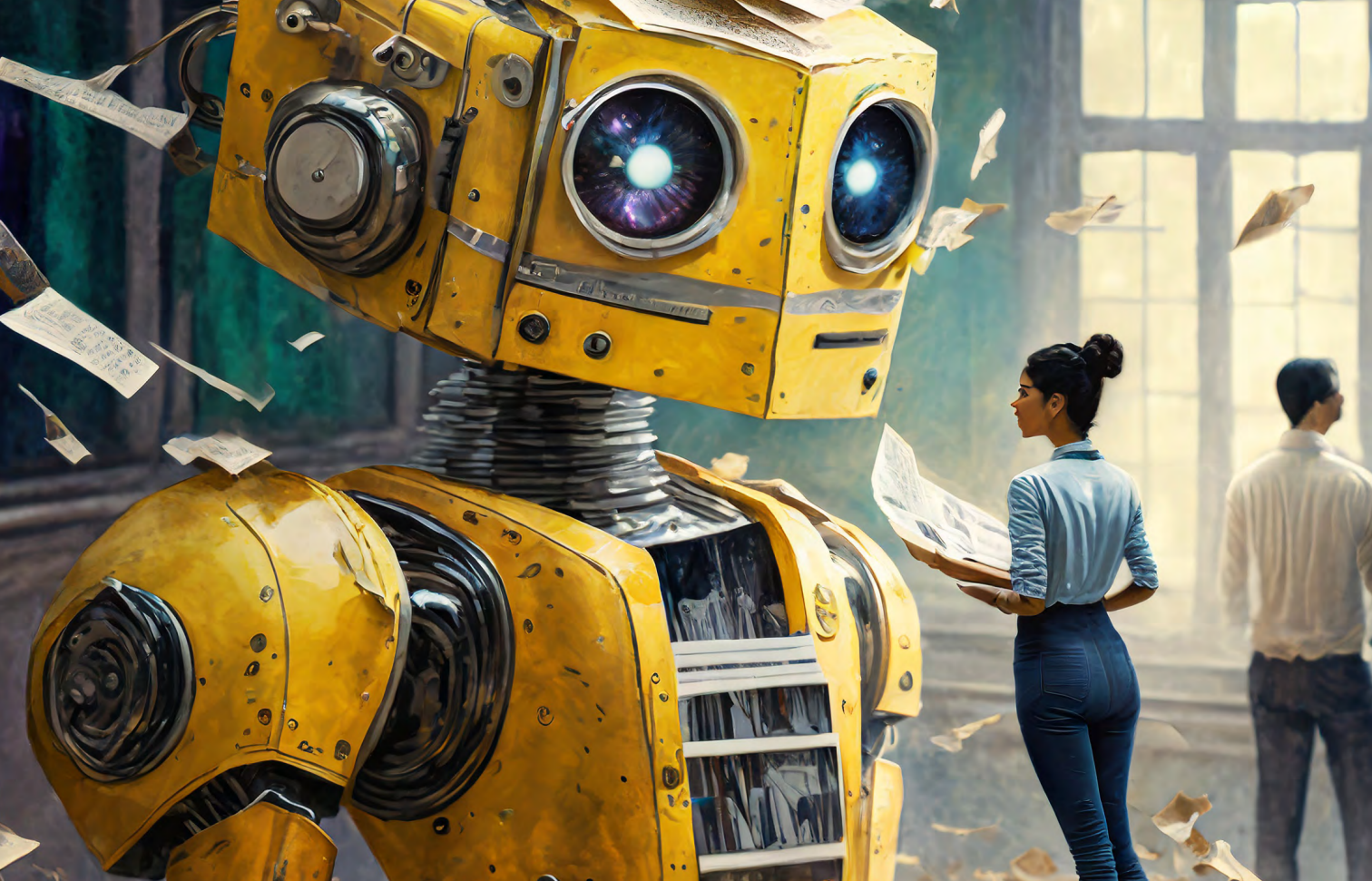
CBC (USA) takes a "no surprises" approach, promising "full disclosure" so that "audiences will be made aware of any AI-generated content before they listen, view or read it" (Fenlon, 2023).

Ringier, a media group in Switzerland, states that content generated by AI tools should be labelled. The group notes, however, that "labeling is not required in cases where an AI tool is used only as an aid," suggesting a different approach towards transparency when AI is used by journalists to help with brainstorming ideas, editing suggestions or any other task which might augment journalism workflows (Ringier AG, Corporate Communications, 2023).

Are Australian newsrooms ready for generative AI?

On more than one occasion throughout our discussions we heard the phrase, "Not all newsrooms are the New York Times." It is true no Australian newsroom can compete with the resources of the New York Times or BBC. However, local newsrooms clearly understand that developing AI fitness within their organisations is going to be necessary, and they need to quickly understand the opportunities and challenges of this fast-moving technology.

Guardian Australia has begun some early efforts at AI-literacy training. The ABC has significant strategic work underway to help them focus their limited technical resources and better understand how generative AI can build audience trust rather than diminish it.

Local newsrooms will need to dedicate resources to keep up with developments in AI, especially as it relates to the broader news industry. There may be situations where industry collaboration makes sense. Aligning and simplifying language and definitions in guidelines could be a starting point. Verification tools and training could be another.

Navigating concerns over verification and the quality of the broader information ecosystem will not be easy and may require some degree of industry-wide collaboration. Decisions about the future role of artificial intelligence in the news industry should consider all ethical, legal and societal dimensions, and there is an opportunity for the Australian news industry to begin having this conversation together as well as within their own organisations.

Yet again, the news industry finds itself at a turning point, and editorial leaders are clear-eyed about the risks. Finding the most efficient ways to mitigate those risks will help news organisations focus on innovation and unleash the potential opportunities of generative AI. It is encouraging to see a 'test-and-learn' culture and early product innovation already happening in many Australian news organisations.

The opportunities generative AI could present in optimising workflows and improving how news organisations research and investigate stories are enormous. Assuming generative AI adds more low-quality, news-style content to the broader information ecosystem, a focus on original journalism will be critical. Australian newsrooms are alert to this upside. If they can navigate and mitigate their well-founded concerns over verification and the quality of the broader information ecosystem, there could be a brighter and more innovative news industry on the other side.

04

REGULATORY DEVELOPMENTS

The release of ChatGPT in November 2022 accelerated moves to develop policy and regulations on AI. This chapter outlines the major developments around the world.

Canada

The Canadian government tabled a bill in December 2027 that included several proposed acts relating to the digital industry, including the Artificial Intelligence and Data Act. It aims to ensure the responsible design, development and deployment of AI systems. As of 30 November 2023, the bill had not yet passed. In September 2023 the government released the Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems (Government of Canada, 2023). The code sets out a range of measures to ensure accountability, safety, fairness and equity, transparency, human oversight, and validity and robustness.

China

China has issued a set of temporary measures, effective from August 15 2023, to regulate generative AI. It requires developers to submit security and risk assessments and receive clearance before releasing mass-market AI products. Following government approvals, four Chinese tech firms launched their AI chatbots to the public on August 31 (Reuters. 2023).

European Union

The EU proposed an AI Act in 2021 (European Commission, 2021). This grew out of what was initially a non-regulatory approach with the publication of non-binding Ethics Guidelines on Trustworthy AI in 2019 (European Commission, 2019), before a shift to regulation in a 2020 white paper (European Commission, 2020). The proposal was amended in June 2023, partly to accommodate concerns over its applicability to generative AI (Lennett, 2023). EU Parliament expects to finalise the law this year, but negotiations with member states were still underway at the time of publication (European Parliament, 2023). The general objective of the legislation is 'to ensure the proper functioning of the single market by creating the conditions for the development and use of trustworthy AI systems in the Union' (European Parliamentary Research Service, 2023, p. 3). The briefing note accompanying the draft notes the implications of AI systems for user safety and fundamental rights, including rights to non-discrimination, freedom of expression, human dignity, personal data protection and privacy (European Parliamentary Research Service, 2023).

The proposed legislative framework would enshrine a technology-neutral definition of AI systems and instead adopt a risk-based approach that classifies AI systems into four tiers from minimal to unacceptable risk. The act would ban harmful AI systems considered to be a clear threat to people's safety, livelihoods

and rights, such as those that exploit specific vulnerable groups. High-risk AI systems would be subject to conformity assessment and registration before being released. Most generative AI applications relevant to news media would be classed as limited risk and subject to a limited set of transparency obligations (European Parliament, 2023).

United Kingdom

In March 2023, the UK government released a white paper detailing its plans for a 'pro-innovation approach to AI regulation' (Department for Digital, Culture, Media & Sport, 2023). Despite the title, the proposed policy approach acknowledges the need to address risks arising from AI, including infringement of rights and the potential to undermine public trust. As such, the paper proposes a proportionate and mixed approach to regulation that includes the development of regulatory principles implemented by existing bodies, alongside technical standards, voluntary guidance and education. It aims for technological neutrality by regulating use, and notes the importance of a risk-based and context-specific approach. It proposes five 'values-focused, cross-sectoral principles': safety, security and robustness; appropriate transparency and explainability; fairness; accountability and governance; and contestability and redress. There is no specific reference to news, journalism or other elements of the information environment, though it does briefly consider a case of misinformation for the purposes of illustration. The need to address bias and protect individual privacy is considered.

United States

In October 2023, US president Joe Biden issued the Executive Order on Safe, Secure and Trustworthy Artificial Intelligence (The White House, 2023) a landmark Executive Order which sets out standards that seek to protect individual rights, promote innovation and competition, advance US interests and ensure responsible and effective government use of AI. Critically, it sets transparency requirements for developers of foundation models that pose serious risks, including the sharing of safety testing data. The National Institute of Standards and Technology is tasked with setting standards for safety testing, while government agencies will be charged with applying the standards to sectors for which they are responsible. Most relevantly for the news industry, the order seeks to establish standards and best practices for detecting AI-generated content and authenticating official content, for protecting privacy, to address algorithmic discrimination and to mitigate any harms to workers. There is no specific reference to protecting news and information integrity. The order followed the October 2022 release of the Blueprint for an AI Bill of Rights (Office of Science and Technology Policy, 2022).

US government agencies are also pursuing action against some AI developers under existing law. The Federal Trade Commission launched an investigation into OpenAI in July for potential violations of consumer protection laws. It asked OpenAI for details of consumer complaints regarding 'false, misleading, disparaging or harmful statements', and has asked for details on how it obtains data and trains its models, its processes for human feedback, risk assessment and mitigation, and its mechanisms for privacy protection (Susarla, 2023).

International agreements

As detailed in Chapter one, in November 2023, the UK government hosted an international AI summit focusing on frontier risks – those that emerge from the most advanced AI models, rather than from specific applications (Chatham House, 2022). The summit resulted in the Bletchley Declaration, a joint commitment from 29 governments to address frontier risks (UK Government, 2023a). The UK also announced an AI Safety Institute that would be tasked with testing the safety of emerging types of AI. Major AI companies agreed to submit frontier models for safety testing. Annual summits on AI safety are planned, as is a series of reports on the state of the technology (UK Government, 2023b).

In October 2023, the United Nations launched a high-level advisory body on artificial intelligence, noting that 'globally coordinated AI governance is the only way to harness AI for humanity, while addressing its risks and uncertainties' (United Nations, n.d.). The body will seek to build a global scientific consensus on risks and challenges, help harness AI for the UN Sustainable Development Goals, and strengthen international cooperation on AI governance. It plans to issue preliminary recommendations by the end of 2023, with final recommendations by mid-2024 (United Nations, 2023).

Australia

In June 2023, the Australian government released a discussion paper on supporting responsible AI. As of November 2023, it is considering feedback to inform any regulatory and policy responses (Department of Industry, Science and Resources, Australian Government, 2023).

CONCLUSION

The survey of Australian newsrooms in this report shows the way they are thinking about using generative AI and the the way they are considering which guardrails are needed to protect the ethical and editorial integrity of news content. The survey also points to the usefulness of pausing and taking stock of some of the unintended consequences of this new technology.

First amongst those consequences is the likely disparity between well-funded international newsrooms and small, underfunded newsrooms in the resources available for the development of products which could deliver a degree of autonomy from the makers and distributors of generative AI. This is likely to leave smaller operations highly dependent on the products produced by platforms or manufacturers. Smaller newsrooms may also struggle to develop ethical and editorial guidelines around the usage of generative AI, opening up another front of reliance on funding from those platforms distributing generative AI products.

There is a disparity too in Australia, as possibly elsewhere in the world, between metropolitan and regional newsrooms, where the latter's business models have been hard hit by successive waves of 2020–2022 Covid lockdowns. Many regional news outlets have been reduced to one- or two-person operations, with some already reliant on platform funding for innovation and business development. The connections between regional news media and metropolitan media have all but disappeared, except for the News Local network operated by News Corp, in which small, generally one-person bureaus are located regionally to report hyperlocal news content for the metro mastheads and which appear as drop-down tags on the online front page of each. The financial difficulties facing other regional newsrooms, which do not have the headquarters of a larger entity to rely upon, will leave them struggling to find the resources required to contemplate the editorial and ethical challenges posed by the use of generative AI to produce news content.

Finally, the challenge of verifying information flowing into newsrooms which are already significantly stressed is one that will impact all newsmakers – big and small. As information integrity becomes more stressed by new technologies, the ability of newsrooms to keep pace with more-sophisticated attacks will require specifically created tools. Again, the relative size and wealth of the news entity will play a role in how well they perform under such stress.

REFERENCES

Acerbi, A., Altay, S., & Mercier, H. (2022). Research note: Fighting misinformation or fighting for information? Harvard Kennedy School Misinformation Review. https://doi.org/10.37016/mr-2020-87

Adami, M. (2023, March 23). Is ChatGPT a threat or an opportunity for journalism? Five AI experts weigh in. Reuters Institute for the Study of Journalism. https://reutersinstitute.politics.ox.ac.uk/news/chatgpt-threat-or-opportunity-journalism-five-ai-experts-weigh

Allyn, B. (2023, August 16). 'New York Times' considers legal action against OpenAI as copyright tensions swirl. NPR. https://www.npr.org/2023/08/16/1194202562/new-york-times-considers-legal-action-against-openai-as-copyright-tensions-swirl

Altay, S., Berriche, M., & Acerbi, A. (2023). Misinformation on Misinformation: Conceptual and Methodological Challenges. Social Media + Society, 9(1), 20563051221150412. https://doi.org/10.1177/20563051221150412

Altay, S., Berriche, M., Heuer, H., Farkas, J., & Rathje, S. (2023). A survey of expert views on misinformation: Definitions, determinants, solutions, and future of the field. Harvard Kennedy School Misinformation Review. https://doi.org/10.37016/mr-2020-119

AppliedXL (Director). (2023, November 4). Francesco Marconi on AI in News: Testimony Before UK Parliament. https://www.youtube.com/watch?v=f9S32ugkgWs

Arguedas, A. R., & Simon, F. M. (2023). Automating democracy: Generative AI, journalism, and the future of democracy. https://ora.ox.ac.uk/objects/uuid:0965ad50-b55b-4591-8c3b-7be0c587d5e7

Associated Press. (n.d.). Local News AI. Associated Press. Retrieved 23 November 2023, from https://www.ap.org/discover/Local-News-AI

Beckett, C. (2019). New powers, new responsibilities. A global survey of journalism and artificial intelligence. London School of Economics and Political Science. https://blogs.lse.ac.uk/polis/2019/11/18/new-powers-new-responsibilities/

Beckett, C., & Yaseen, M. (2023). Generating Change: The Journalism AI report. Polis, London School of Economics and Political Science. https://static1.squarespace.com/static/64d60527c01ae7106f2646e9/t/6509b9a39a5ca70df9148eac/1695136164679/Generating+Change+_+The+Journalism+AI+report+_+English.pdf

Belanger, A. (2023, July 21). OpenAI, Google will watermark AI-generated content to hinder deepfakes, misinfo. Ars Technica. https://arstechnica.com/ai/2023/07/openai-google-will-watermark-ai-generated-content-to-hinder-deepfakes-misinfo/

Belot, H. (2023, November 3). Australian academics apologise for false AI-generated allegations against big four consultancy firms. The Guardian. https://www.theguardian.com/business/2023/nov/02/australian-academics-apologise-for-false-ai-generated-allegations-against-big-four-consultancy-firms

Bohannon, M. (2023, June 8). Lawyer Used ChatGPT In Court—And Cited Fake Cases. A Judge Is Considering Sanctions. Forbes. https://www.forbes.com/sites/mollybohannon/2023/06/08/lawyer-used-chatgpt-in-court-and-cited-fake-cases-a-judge-is-considering-sanctions/

Brewster, J., Fishman, Z., & Xu, E. (2023, June). Some of the World's Largest Blue Chip Brands Unintentionally Support the Spread of Unreliable AI-Generated News Websites. NewsGuard Misinformation Monitor. https://www.newsguardtech.com/misinformation-monitor/june-2023

Brittain, B. (2023, September 7). OpenAI, Microsoft hit with new US consumer privacy class action. Reuters. https://www.reuters.com/legal/litigation/openai-microsoft-hit-with-new-us-consumer-privacy-class-action-2023-09-06/

Bruell, A. (2023, June 28). Big News Publishers Look to Team Up to Address Impact of AI. Wall Street Journal. https://www.wsj.com/articles/big-news-publishers-look-to-team-up-to-address-impact-of-ai-80c3b88e

Buckley, J., & Wilson, C. (2023, August 9). News Corp staff blindsided by company's use of AI to produce articles for 'years'. Crikey. https://www.crikey.com.au/2023/08/09/news-corp-ai-articles-journalism/

Chatham House. (2022, May 25). UK AI Summit: What can it achieve? https://www.chathamhouse.org/events/all/members-event/uk-ai-summit-what-can-it-achieve

Davies, R. T. (2023, October 5). Generative AI at the BBC. https://www.bbc.co.uk/mediacentre/articles/2023/generative-ai-at-the-bbc/

Deiana, G., Dettori, M., Arghittu, A., Azara, A., Gabutti, G., & Castiglia, P. (2023). Artificial Intelligence and Public Health: Evaluating ChatGPT Responses to Vaccination Myths and Misconceptions. Vaccines, 11(7), 1217. https://doi.org/10.3390/vaccines11071217

Department for Digital, Culture, Media & Sport. (2023). A pro-innovation approach to AI regulation. https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper

Department of Industry, Science and Resources, Australian Government. (2023). Supporting responsible AI: discussion paper. https://consult.industry.gov.au/supporting-responsible-ai

Digital Content Next. (2023). Principles for Development and Governance of Generative AI. https://digitalcontentnext.org/blog/2023/06/05/dcns-principles-for-development-and-governance-of-generative-ai/

European Commission. (2019). Ethics guidelines for trustworthy AI. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

European Commission. (2020). White Paper On Artificial Intelligence—A European approach to excellence and trust (p. European Commission). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0065&WT_mc_id=Twitter

European Commission. (2021) Proposal for a Regulation laying down harmonised rules on artificial intelligence, COM/2021/206 (2021).

European Parliament. (2023, June 8). EU AI Act: First regulation on artificial intelligence. News - European Parliament. https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

European Parliamentary Research Service. (2023). Briefing: Artificial Intelligence Act (PE 698.792; EU Legislation in Progress).

Fenlon, B. (2023, June 12). How CBC News will manage the challenge of AI. CBS News Editor's Blog. https://www.cbc.ca/news/editorsblog/cbc-twitter-news-1.6873270

Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2023). Generative AI. Business & Information Systems Engineering. https://doi.org/10.1007/s12599-023-00834-7

Gal, U. (2023, February 8). ChatGPT is a data privacy nightmare. If you've ever posted online, you ought to be concerned. The Conversation. http://theconversation.com/chatgpt-is-a-data-privacy-nightmare-if-youve-ever-posted-online-you-ought-to-be-concerned-199283

Gent, E. (2023, May 10). What is the AI alignment problem and how can it be solved? New Scientist. https://www.newscientist.com/article/mg25834382-000-what-is-the-ai-alignment-problem-and-how-can-it-be-solved/

Golding, Y. T. (2023, October 18). The news media and AI: A new front in copyright law. Columbia Journalism Review. https://www.cjr.org/business_of_news/data-scraping-ai-litigation-lawsuit-artists-authors.php

Goldstein, J. A., Sastry, G., Musser, M., DiResta, R., Gentzel, M., & Sedova, K. (2023). Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations (arXiv:2301.04246). arXiv. http://arxiv.org/abs/2301.04246

Government of Canada. (2023). Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems. Innovation, Science and Economic Development Canada. https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems

Gregory, S. (2023). Testimony of Sam Gregory, Executive Director, WITNESS, before the U.S. Senate Committee on Commerce, Science and Transportation Subcommittee on Consumer Protection, Product Safety and Data Security. https://www.gen-ai.witness.org/wp-content/uploads/2023/09/Senate-Commerce-written-testimony_FINAL.pdf

Guaglione, S. (2023a, March 29). How one publisher is using generative AI to publish thousands of evergreen posts, create a chatbot. Digiday. https://digiday.com/media/how-one-publisher-is-using-generative-ai-to-publish-thousands-of-evergreen-posts-create-a-chatbot/

Guaglione, S. (2023b, August 14). The case for and against open-source large language models for use in newsrooms. Digiday. https://digiday.com/media/the-case-for-and-against-open-source-large-language-models-for-use-in-newsrooms/

Hartmann, J., Schwenzow, J., & Witte, M. (2023). The political ideology of conversational AI: Converging evidence on ChatGPT's pro-environmental, left-libertarian orientation (arXiv:2301.01768). arXiv. https://doi.org/10.48550/arXiv.2301.01768

Hsu, T., & Thompson, S. A. (2023a, February 8). Disinformation Researchers Raise Alarms About A.I. Chatbots. The New York Times. https://www.nytimes.com/2023/02/08/technology/ai-chatbots-disinformation.html

Hsu, T., & Thompson, S. A. (2023b, October 28). A.I. Muddies Israel-Hamas War in Unexpected Way. The New York Times. https://www.nytimes.com/2023/10/28/business/media/ai-muddies-israel-hamas-war-in-unexpected-way.html

Jaspan, C. (2023, September 14). Breaking news: How the Australian media is grappling with AI. The Sydney Morning Herald. https://www.smh.com.au/business/companies/breaking-news-how-the-australian-media-is-grappling-with-ai-20230914-p5e4nh.html

Kaye, B. (2023, April 6). Australian mayor readies world's first defamation lawsuit over ChatGPT content. Reuters. https://www.reuters.com/technology/australian-mayor-readies-worlds-first-defamation-lawsuit-over-chatgpt-content-2023-04-05/

Kirchner, J. H., Ahmad, L., Aaronson, S., & Leike, J. (2023, January 31). New AI classifier for indicating AI-written text. https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text

Kreps, S., McCain, R. M., & Brundage, M. (2022). All the News That's Fit to Fabricate: AI-Generated Text as a Tool of Media Misinformation. Journal of Experimental Political Science, 9(1), 104–117. https://doi.org/10.1017/XPS.2020.37

Lennett, B. (2023, June 15). Can Europe's Laws Keep Up with AI's New Arms Race? Tech Policy Press. https://techpolicy.press/can-europes-laws-keep-up-with-ais-new-arms-race/

Martineau, K. (2023, August 22). What is retrieval-augmented generation? IBM Research Blog. https://research.ibm.com/blog/retrieval-augmented-generation-RAG

Meade, A. (2023, July 31). News Corp using AI to produce 3,000 Australian local news stories a week. The Guardian. https://www.theguardian.com/media/2023/aug/01/news-corp-ai-chat-gpt-stories

Menegus. (2023, June 6). CNET's new guidelines for AI journalism met with union pushback. Engadget. https://www.engadget.com/cnets-new-guidelines-for-ai-journalism-met-with-union-pushback-152311269.html

Motoki, F., Pinho Neto, V., & Rodrigues, V. (2023). More human than human: Measuring ChatGPT political bias. Public Choice. https://doi.org/10.1007/s11127-023-01097-2

Newman, N., Fletcher, R., Eddy, K., Robertson, C. T., & Nielsen, R. K. (2023). Reuters Institute Digital News Report 2023. Reuters Institute for the Study of Journalism.

News/Media Alliance. (2023). News/Media Alliance AI Principles. https://www.newsmediaalliance.org/ai-principles/

Nightingale, S. J., & Farid, H. (2022). AI-synthesized faces are indistinguishable from real faces and more trustworthy. Proceedings of the National Academy of Sciences of the United States of America, 119(8), e2120481119. https://doi.org/10.1073/pnas.2120481119

Nishal, S., & Diakopoulos, N. (2023). Envisioning the Applications and Implications of Generative AI for News Media. Generative AI and HCI at CHI 2023, Hamburg.

Office of Science and Technology Policy. (2022). Blueprint for an AI Bill of Rights. https://www.whitehouse.gov/ostp/ai-bill-of-rights/

Partnership on AI. (2023). AI Adoption for Newsrooms: A 10-Step Guide. https://partnershiponai.org/ai-for-newsrooms

Pecquerie, B. (2018, February 4). AI is the new horizon for news. Global Editors Network. https://medium.com/global-editors-network/ai-is-the-new-horizon-for-news-22b5abb752e6

Reuters. (2023, September 11). What are governments doing to try to regulate AI? Euronews. https://www.euronews.com/next/2023/09/11/which-countries-are-trying-to-regulate-artificial-intelligence

Reuters Communications. (2023, July 10). Reuters launches AI-powered discoverability features for video library on Reuters Connect, accelerating discovery, editing and publishing. Reuters News Agency. https://www.reutersagency.com/en/media-center/reuters-launches-ai-powered-discoverability-features-for-video-library-on-reuters-connect-accelerating-discovery-editing-and-publishing/

Ringier AG, Corporate Communications. (2023, May 30). Ringier introduces clear guidelines for the use of artificial intelligence. Ringier. https://www.ringier.com/ringier-introduces-clear-guidelines-for-the-use-of-artificial-intelligence/

Rogerson, A., Hankins, E., Fuentes Nettel, P., & Rahim, S. (2022). Government AI Readiness Index 2022. Oxford Insights. https://oxfordinsights.com/ai-readiness/ai-readiness-index

Roper, D., Henriksson, T., Hälbich, K., & Martin, O. (2023). Gauging Generative AI's impact on newsrooms. World Association of News Publishers (WAN-IFRA). https://wan-ifra.org/insight/gauging-generative-ais-impact-in-newsrooms/

Samosir, H. (2023, July 13). AI news readers are becoming more common across Asia. Will Australia follow suit? ABC News. https://www.abc.net.au/news/2023-07-13/artificial-intelligence-news-readers-becoming-common-in-asia/102591790

Sharman, D. (2023a, March 10). Journalists say 'a lot of work to be done' as Reach starts AI test—Journalism News from HoldtheFrontPage. HoldtheFrontPage. https://www.holdthefrontpage.co.uk/2023/news/journalists-say-a-lot-of-work-to-be-done-as-publisher-starts-robot-reporting-test/

Sharman, D. (2023b, March 29). Publishers urged to become 'major players' in AI development—Journalism News from HoldtheFrontPage. HoldtheFrontPage. https://www.holdthefrontpage.co.uk/2023/news/publishers-urged-to-become-major-players-in-development-of-ai/

Sherman, A., & Rizzo, L. (2023, June 6). A.I. poses new threats to newsrooms, and they're taking action. CNBC. https://www.cnbc.com/2023/06/06/news-organizations-ai-disinformation.html

Simon, F. M., Altay, S., & Mercier, H. (2023). Misinformation reloaded? Fears about the impact of generative AI on misinformation are overblown. Harvard Kennedy School Misinformation Review. https://doi.org/10.37016/mr-2020-127

Solaiman, I., Brundage, M., Clark, J., Askell, A., Herbert-Voss, A., Wu, J., Radford, A., Krueger, G., Kim, J. W., Kreps, S., McCain, M., Newhouse, A., Blazakis, J., McGuffie, K., & Wang, J. (2019). Release Strategies and the Social Impacts of Language Models. (arXiv:1908.09203). arXiv. https://doi.org/10.48550/arXiv.1908.09203

Stelter, B. (2021, November 16). This infamous Steve Bannon quote is key to understanding America's crazy politics. CNN. https://www.cnn.com/2021/11/16/media/steve-bannon-reliable-sources/index.html

Susarla, A. (2023, July 18). FTC probe of OpenAI: Consumer protection is the opening salvo of US AI regulation. The Conversation. http://theconversation.com/ftc-probe-of-openai-consumer-protection-is-the-opening-salvo-of-us-ai-regulation-209821

Ternovski, J., Kalla, J., & Aronow, P. M. (2022). The Negative Consequences of Informing Voters about Deepfakes: Evidence from Two Survey Experiments. Journal of Online Trust and Safety, 1(2). https://doi.org/10.54501/jots.v1i2.28

The White House. (2023). Fact Sheet: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence. https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/

Thompson, S. A., & Hsu, T. (2023, June 28). How Easy Is It to Fool A.I.-Detection Tools? The New York Times. https://www.nytimes.com/interactive/2023/06/28/technology/ai-detection-midjourney-stable-diffusion-dalle.html

Thorbecke, C. (2023, June 16). Forget about the AI apocalypse. The real dangers are already here. CNN Business. https://edition.cnn.com/2023/06/16/tech/ai-apocalypse-warnings/index.html

Turness, D. (2023, May 21). Explaining the 'how'—The launch of BBC Verify. BBC News. https://www.bbc.com/news/uk-65650822

UK Government. (2023a, November 1). The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023. GOV.UK. https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023

UK Government. (2023b, November 2). World leaders, top AI companies set out plan for safety testing of frontier as first global AI Safety Summit concludes. https://www.gov.uk/government/news/world-leaders-top-ai-companies-set-out-plan-for-safety-testing-of-frontier-as-first-global-ai-safety-summit-concludes

United Nations. (n.d.). AI Advisory Body. United Nations; United Nations. Retrieved 27 November 2023, from https://www.un.org/ai-advisory-body

United Nations. (2023). UN Secretary-General launches AI Advisory Body on risks, opportunities, and international governance of artificial intelligence. https://www.un.org/sites/un2.un.org/files/231025_press-release-aiab.pdf

United States Federal Trade Commission. (2023). Creative Economy and Generative AI. Federal Trade Commission. https://www.ftc.gov/system/files/ftc_gov/pdf/creative-economy-and-generative-ai-transcript-october-4-2023.pdf

van Giffen, B., Herhausen, D., & Fahse, T. (2022). Overcoming the pitfalls and perils of algorithms: A classification of machine learning biases and mitigation methods. Journal of Business Research, 144, 93–106. https://doi.org/10.1016/j.jbusres.2022.01.076

Ventura-Pocino, P. (2021). Algorithms in the newsrooms: Challenges and recommendations for artificial intelligence with the ethical values of journalism. Catalan Press Council. https://fcic.periodistes.cat/wp-content/uploads/2022/03/venglishDIGITAL_ALGORITMES-A-LES-REDACCIONS_ENG-1.pdf

Viner, K., & Bateson, A. (2023, June 16). The Guardian's approach to generative AI. The Guardian. https://www.theguardian.com/help/insideguardian/2023/jun/16/the-guardians-approach-to-generative-ai

Von Nordheim, G., Bettels-Schwabbauer, T., Kleinen-von Königslöw, K., Barczyszyn-Madziarz, P., Budivska, H., Di Salvo, P., Dingerkus, F., Guazina, L. S., Krobea Asante, K., Ku , M., Lábová, S., Matei, A., Merkovity, N., Paulino, F. O., Petrovszki-Oláh, L., Serwornoo, M. Y. W., Valente, J., Wake, A., & Zakinszky Toma, V. (2023). The different worlds of Google – A comparison of search results on conspiracy theories in 12 countries. Convergence: The International Journal of Research into New Media Technologies, 13548565231203102. https://doi.org/10.1177/13548565231203102

Warren, C. (2023, July 18). Old media rub their hands in glee at an AI pay-off—But at what cost? Crikey. https://www.crikey.com.au/2023/07/19/artificial-intelligence-media-pay-off/

West, D. M. (2023, March 23). Comparing Google Bard with OpenAI's ChatGPT on political bias, facts, and morality. Brookings Institution. https://www.brookings.edu/articles/comparing-google-bard-with-openais-chatgpt-on-political-bias-facts-and-morality/

Woollacott, E. (2023, September 1). OpenAI Hit With New Lawsuit Over ChatGPT Training Data. Forbes. https://www.forbes.com/sites/emmawoollacott/2023/09/01/openai-hit-with-new-lawsuit-over-chatgpt-training-data/

Yerushalmy, J. (2023, March 1). German publisher Axel Springer says journalists could be replaced by AI. The Guardian. https://www.theguardian.com/technology/2023/mar/01/german-publisher-axel-springer-says-journalists-could-be-replaced-by-ai

Yu, D., Rosenfeld, H., & Gupta, A. (2023, January 16). The 'AI divide' between the Global North and Global South. World Economic Forum. https://www.weforum.org/agenda/2023/01/davos23-ai-divide-global-north-global-south/

Zhao, Y., Liu, B., Ding, M., Liu, B., Zhu, T., & Yu, X. (2023). Proactive Deepfake Defence via Identity Watermarking. 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 4591–4600. https://doi.org/10.1109/WACV56688.2023.00458

ABOUT THE CENTRE FOR MEDIA TRANSITION

The Centre for Media Transition (CMT) is an applied research unit based at the University of Technology Sydney (UTS).

Launched in 2017, the CMT is an interdisciplinary initiative of the Faculty of Arts and Social Sciences and the Faculty of Law. It sits at the intersection of media, journalism, technology, ethics, regulation and business. Working with industry, academia, government and others, the CMT aims to:

- Understand media transition and digital disruption, with a view to recommending legal reform and other measures that promote the public interest;

- Assist news media to adapt for a digital environment, including by identifying potentially sustainable business models;

- Develop suitable ethical and regulatory frameworks for a fast-changing digital ecosystem;

- Foster quality journalism, thereby enhancing democracy in Australia and the region;

- Develop a diverse media environment that embraces local/regional, international and transnational issues and debate;

- Combat misinformation and protect digital privacy; and

- Articulate contemporary formulations of the public interest informed by established and enduring principles such as accountability and the public's right to know.

The CMT's published works include reports on digital defamation, trust in news media, the state of regional news and news media innovation. Current projects include work on industry self-regulation, privacy, news verification, foreign reporting and press freedom.

The Centre regularly hosts public events, conferences and forums. You can sign up to our regular newsletter at **go.uts.edu.au/CMT-eNews-Signup**. Details of events and the CMT's work can be found on our website at **cmt.uts.edu.au**