Algorithms and Journalism

Dr. Dani Madrid-Morales, d.madrid-morales@sheffield.ac.uk

Shandong Normal University, 9 April 2026

University of Sheffield

A World TOP 100 University

The Plan for this Week

8 April

Yesterday we talked about key approaches and principles of digital storytelling, and I introduced you to some examples from around the World.

9 April

Today we'll talk about algorithms and data, and how both are impacting the production and distribution of news content.

10 April

During the seminar, I'll share some best practices in data journalism, and we'll work together on an example. Bring your own laptop if you can.

Today's Learning Objectives

By the end of today's session, you should …

1. Understand the various roles algorithms play in journalistic content **production**, **newsgathering**, and **distribution**.
2. Analyse how algorithmic tools and systems are transforming **journalistic practices** and **decision-making**.
3. Examine the implications of **algorithm-driven journalism** for journalistic labour, ethics, and legal responsibility.

University of Sheffield

Part 1: Automation

Define what algorithms are, why they matter for journalism, and discuss how they are used to produce news content (**automated reporting**) and consider whether machines could replace human journalists.

What is an Algorithm?

- In simple terms, it's like a **recipe** or a step-by-step instruction list to accomplish a task (as Gillespie (2014) puts it, an algorithm is an "encoded procedure for transforming input data into a desired output").
  - Algorithms sort our social media newsfeeds, recommend videos, or even help us navigate (think Facebook's feed algorithm or Google Maps directions).
  - They are **central to the digital environment** journalists operate in, shaping not only the content production but also the flow of information to users globally (Brake, 2017).
- Algorithms take input data and produce an output based on calculations, and they now permeate all aspects of life, including **how we get news** (Napoli, 2014)

Algorithms in Journalism

Digital Transformation: As news went digital, journalistic processes have become increasingly mediated by algorithms (Hermida et al., 2012)

From CAR to Computational: Journalism has evolved from Computer-Assisted Reporting to data journalism and now computational journalism; a turn towards quantitative, algorithm-driven methods (Coddington, 2015)

Algorithms as Institutions: Algorithms in media act with institutional power as they shape behaviors, decision-making and even what counts as news, much like traditional media institutions do (Napoli, 2014).

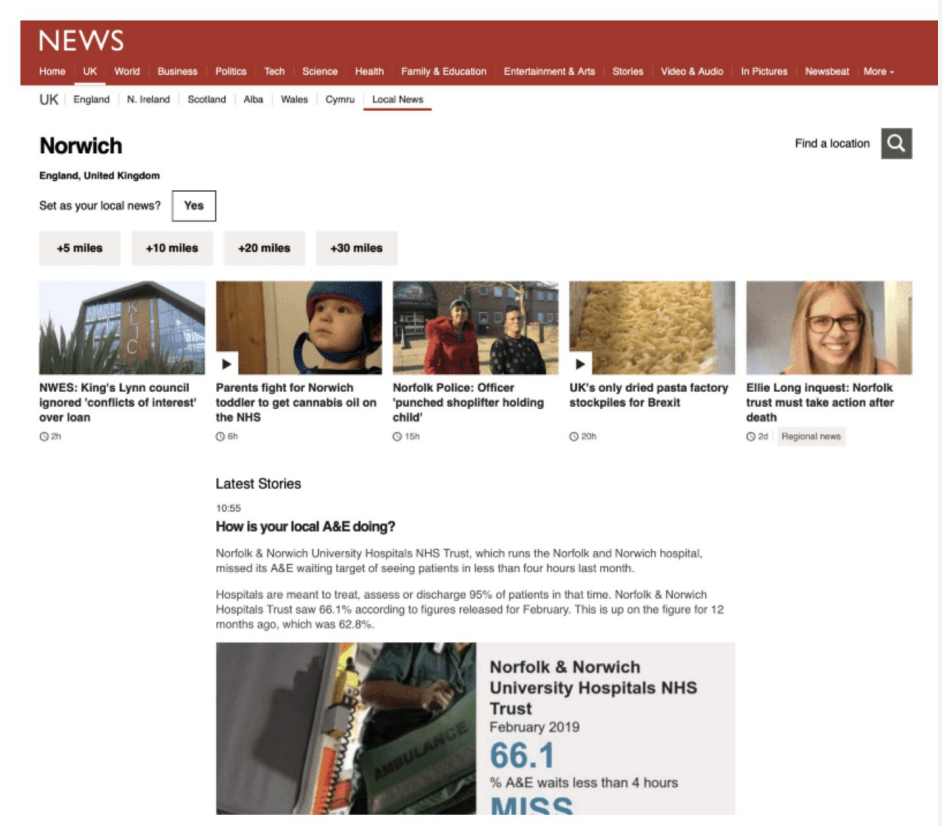
Not Neutral: Algorithms are designed with specific goals and data inputs, and they carry biases from their creators or training data (Feenberg, 1999; Brake, 2017)

Automated Journalism (“robot reporting”)

Using software or algorithms to convert structured data into news stories with minimal human intervention (Dörr, 2016; Graefe, 2016)

- Clean, structured data in ⇒ algorithm applies a narrative template ⇒ news story out. Often used for **repetitive topics** like financial earnings, sports scores, weather, earthquakes
- Automation enables news to be produced **faster and at a far greater volume** than human journalists could manage (Lokot & Diakopoulos, 2016)
- The advantages include consistency, efficiency, and the ability to cover many localised or data-heavy stories (e.g., stock reports for hundreds of companies) that might otherwise go unreported (Diakopoulos & Koliska, 2017)

Automated Journalism | Example



https://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP398.pdf

Automated Journalism | Other Examples

***Los Angeles Times* (USA)**: Early adopter with "Quakebot" – an algorithm that instantly reports earthquake alerts (first used in 2014), and homicide report bots generating short crime news since 2010 (Young & Hermida, 2015)

Reuters & Others (UK/Korea): Major news agencies like Reuters and outlets like Forbes use automated systems for market reports and other routine news. In South Korea, the *Financial News* developed robot writers for stock news (Jia, 2020)

Tencent's 'Dreamwriter' (China): In 2015, Chinese tech giant Tencent debuted Dreamwriter, an algorithm that published a complete news article (one of the first examples of automated journalism in Chinese media) (Jia, 2020)

Activity #1: Will Robots Replace Journalists?

1. In pairs, read **this news story**, and **this new story**. One was generated by a human journalist and the other by an automated system. Discuss and guess which is which.
2. After a few minutes, you'll be asked to share the key points of your discussion around the following questions:
   a. What clues helped you decide?
   b. Would most readers notice?
   c. What implications does this have for trust and transparency?

Limits of automated reporting (I)

Routine Only: Automated systems handle repetitive, structured stories well but struggle with complex, unstructured news. They excel at straight facts (Lokot & Diakopoulos, 2016).

Lack of "Human" Judgment: Most algorithms can't yet replicate a journalist's news judgment, critical thinking or curiosity. They struggle to ask questions, chase down leads.

Formulaic Output: Robot-written stories tend to be template-based and uniform in style. They may lack storytelling flair, often resulting in "dry" narratives angles (Jia, 2020)

Data Dependency: If the input data is poor, incomplete or biased, the output story will be flawed. Most algorithms cannot independently verify data accuracy or cross-check facts outside their dataset (Jamil, 2023)

Limits of automated reporting (II)

Human Touch: Creativity, narrative framing, and investigative instincts remain largely human skills, especially in accountability journalism.

Hybrid approaches are emerging, where AI drafts and humans refine or contextualise, such as seen in Germany's Tagesschau or Brazil's Agência Pública experiments.

- ARD's Tagesschau uses an AI-driven metadata and search system (called Medas) to streamline archival research.
- Agência Pública developed an AI-assisted workflow to measure the real-world impact of its journalism.

University of Sheffield

Questions?

University of Sheffield

Part 2: Algorithms

Explore how algorithms assist journalists in **finding** and **analysing** news (beyond writing) and how algorithms act as gatekeepers in disseminating news to audiences.

Algorithms in News Discovery (I)

Data Mining for Stories: Journalists use algorithms to sift through large datasets, finding patterns or anomalies that could be newsworthy (Wu et al., 2019).

> Example: investigative teams running code to find suspicious financial transactions in leaked data

Computer-Assisted Reporting 2.0: Automation helps with tasks like scraping websites for data, merging databases, or analysing statistics, doing in seconds what might take days by hand (Diakopoulos & Koliska, 2017)

Algorithms in News Discovery (II)

Verification Tools: Newsrooms deploy algorithms to monitor and verify information in real-time.

> For example, they write computer programs that scan social media for breaking news clues or flag likely false information (Diakopoulos & Koliska, 2017)

Augmenting Human Effort: These tools assist journalists rather than write stories; they handle the heavy lifting of data crunching, allowing reporters to focus on interpretation, narrative and investigation

Case Study: Automated Document Analysis

Organisations like *The Guardian* and OCCRP use tools that scan large document leaks (e.g., Pandora Papers) to identify patterns and flag keywords.

These systems enable scalable investigations and global collaborations with hundreds of journalists analysing millions of files.

Algorithmic News Distribution (I)

Platform Algorithms: Tech platforms (Facebook, Google, Twitter, TikTok) use algorithms to decide which news stories people see in feeds and search results, they have become “invisible editors” of the news agenda (Brake, 2017)

Personalisation: These algorithms rank content based on user data (preferences, location, past clicks), creating a personalised news experience – one user’s top stories may differ from another’s, for better or worse (Napoli, 2014)

Algorithmic News Distribution (II)

Gatekeeping Shift: Traditionally editors decided what news was front page; now newsfeed and search algorithms act as gatekeepers, controlling information flow to the public (Brake, 2017). This power is often opaque and unaccountable.

> In a way, we can think of algorithms as “invisible gatekeepers”

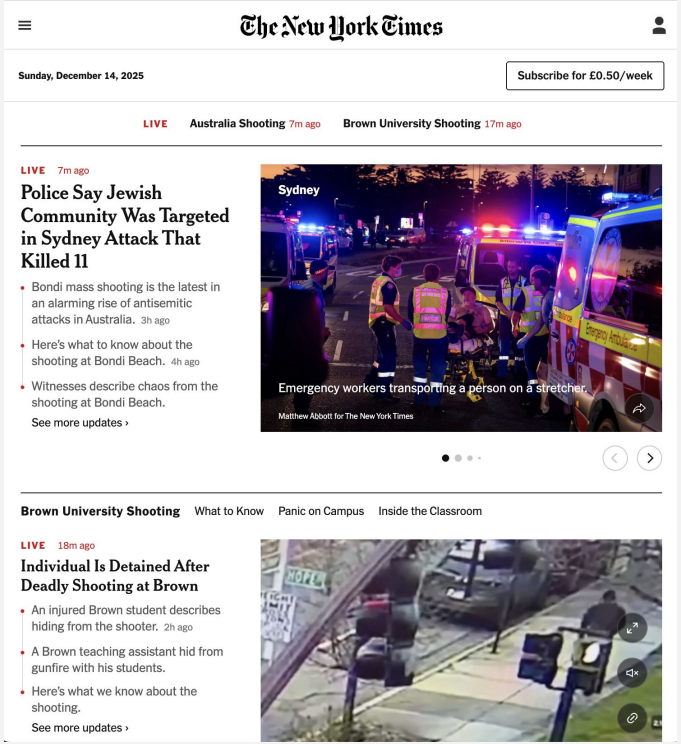
Newsroom Use of Algorithms: Media outlets themselves use algorithms to optimise distribution.

> For example, to schedule social media posts for peak audience times or using SEO (search engine optimisation) strategies so that their stories rank higher on Google (Diakopoulos & Koliska, 2017)

Case Study: The Tagging Algorithm

The New York Times' Editor tool automatically tags entities in articles to improve archiving and reader navigation.

Such backend algorithms enhance searchability and article recommendations, quietly shaping how users access knowledge.



University of Sheffield

Questions?

University of Sheffield

Part 3: Challenges

Examining some structural pressures shaping digital journalism such as algorithmic platforms, AI transformation, and the tension between engagement metrics and public-interest reporting

(Platform) Systemic Challenges

The attention economy

Social media platforms prioritise content that generates clicks, likes, and shares.

In-depth reporting, which requires substantial resources, is disadvantaged under this algorithmic logic, which favors instant engagement over depth.

Smaller newsrooms

Small newsrooms, which operate on limited resources, might dismantle investigative desks and compromise public-interest journalism.

Editorial decisions increasingly reflect what the algorithm favors rather than what communities need to know.

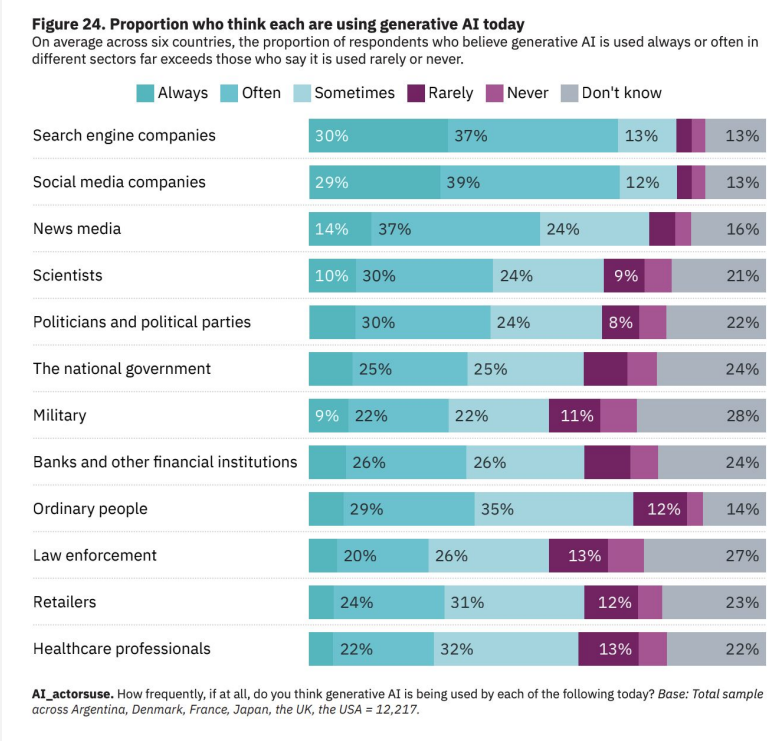
Algorithmic opacity

Audiences encounter journalism through opaque platform algorithms that determine what becomes visible.

This creates structural dependency: newsrooms must adapt storytelling to satisfy platform incentives while having no control over how stories are ranked, deprioritised or removed.

The AI “Revolution”

The role of Artificial Intelligence (AI) and Generative AI (GenAI) is rapidly becoming fundamental to digital narrative workflows. AI is seen by many media leaders as key to journalism's long-term survival due to its potential to improve production efficiency and facilitate complex data analysis at scale.

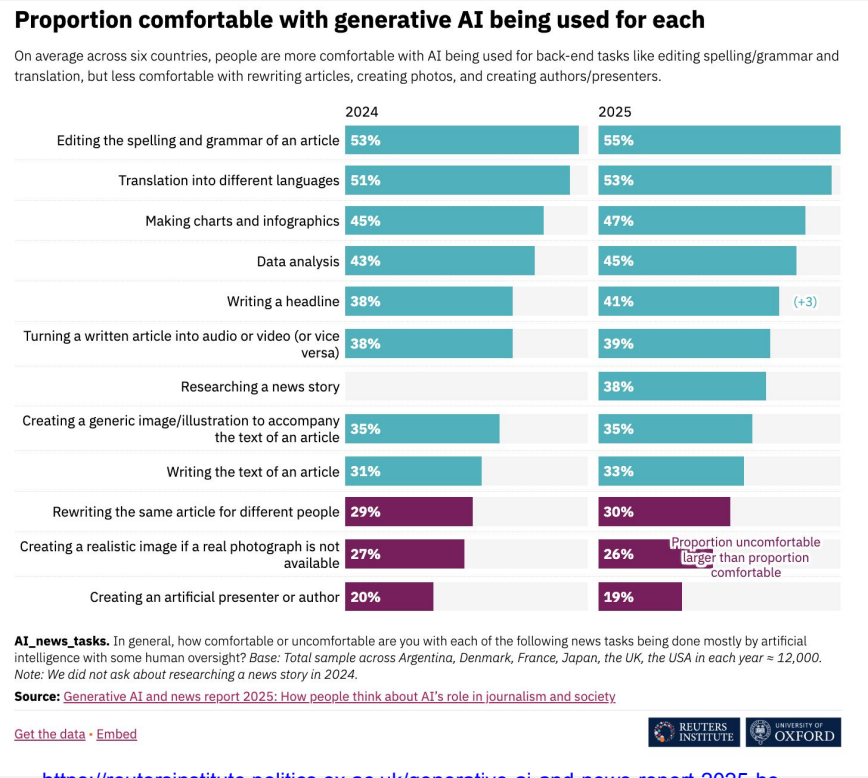


https://reutersinstitute.politics.ox.ac.uk/generative-ai-and-news-report-2025-how-people-think-about-ais-role-journalism-and-society

AI-related Challenges

Efficiency applications: AI now assists with…

- automated back-end work (transcribing interviews, tagging content, generating SEO-optimized headlines, copyediting, processing permissions)
- content generation (producing text, audio, video, and images for summarization, translation, and rewriting stories for different distribution channels)
- investigative assistance (analyzing millions of data points in seconds, identifying hidden patterns of corruption, using tools to detect synthetic deepfake audio)

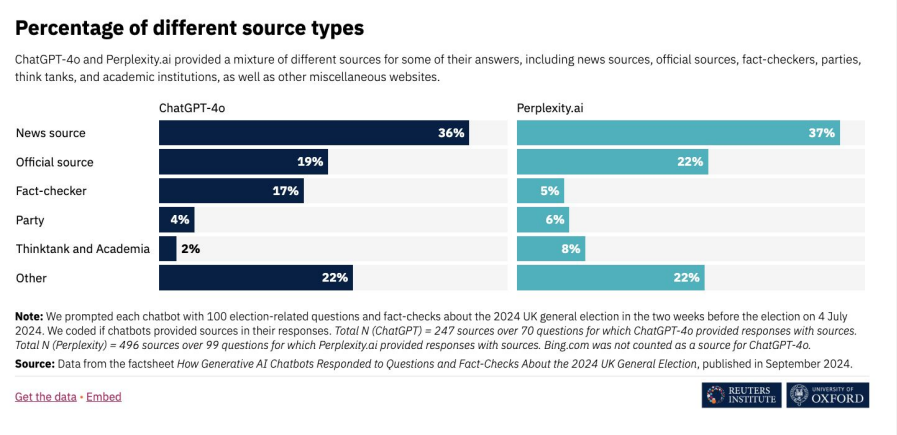


https://reutersinstitute.politics.ox.ac.uk/generative-ai-and-news-report-2025-how-people-think-about-ais-role-journalism-and-society

AI-related Challenges

The threat of homogeneity: When algorithms prioritise what is already popular or optimise content for engagement metrics, it poses risks of standardization and homogenisation, potentially relegating original stories or less-heard voices.

Journalists must evolve into “digital orchestrators” who combine human editorial judgment with AI analytical capabilities



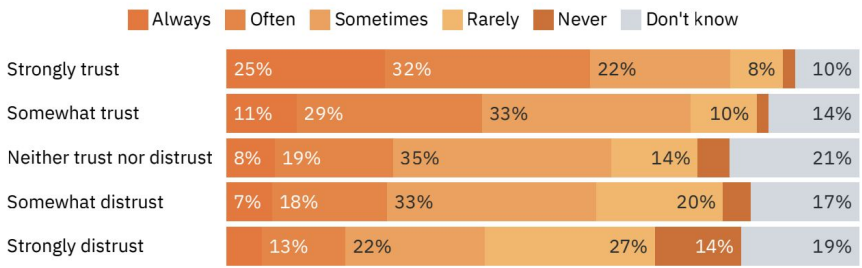
https://reutersinstitute.politics.ox.ac.uk/how-generative-ai-chatbots-responded-questions-and-fact-checks-about-2024-uk-general-election

AI-related Challenges

Trust and transparency paradox: AI cannot fully vet source quality or replicate nuanced judgment. Since AI models can hallucinate and spread disinformation, verification, interpretation, and critical questioning remain fundamentally human.

Transparency is paramount: journalists must clearly disclose when content is AI-generated, though early studies show audiences often trust AI-labeled news content less.

Figure 38. Proportion who think human editors check AI outputs before publishing them
People who trust the news are considerably more likely to think that AI outputs are regularly checked.



News_Trust. Generally speaking, how much do you trust or distrust the news in your country? **AI_news_checking.** How often, if at all, do you think human editors check AI outputs to make sure they are correct or of a high standard before publishing them? *Base: Total sample across Argentina, Denmark, France, Japan, the UK, the USA who Strongly trust the news = 1020, Somewhat trust = 4817, Neither trust nor distrust = 3161, Somewhat distrust = 2304, Strongly distrust = 1025.*

https://reutersinstitute.politics.ox.ac.uk/generative-ai-and-news-report-2025-how-people-think-about-ais-role-journalism-and-society

The Case of Nick Lichtenberg - Overview

A *Fortune* magazine journalist was criticised recently for admitting that he uses AI in his work.

Many journalists have accused Lichtenberg of failing to do his work correctly.

However, Lichtenberg believes that AI is like any other technology – there'll be resistance first, but eventually it will be widely adopted.

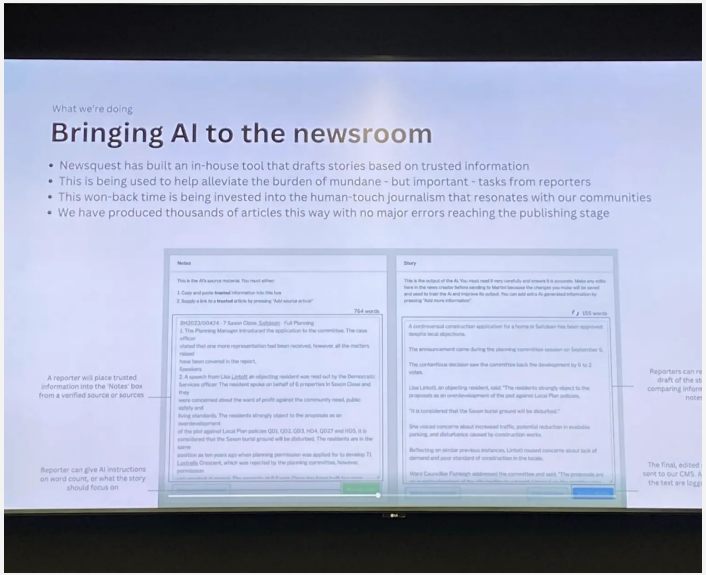
An illustration created with Google's AI tool Nano Banana, featuring a busy newsroom where every reporter, anchor, and radio host resembles Fortune's journalist Nick Lichtenberg. Source: https://reutersinstitute.politics.ox.ac.uk/news/viral-profile-AI-anxieties-Nick-Lichtenberg-fortune

The Case of Nick Lichtenberg - In his own words

"I've done long-form features for which AI helped me with **transcripts and with outlining** a structure of how I wanted to organise the information. I've had breaking news **stories based on a press release** about one of the biggest Fortune 500 companies. I've written stories on a general idea about something that I know to be happening, but I need **help gathering all the sourcing together** really quickly. There are other types of aggregation where I get a document, often a research note from an investment bank or a survey from Gallup or Deloitte or KPMG, and I can use AI to **synthesise that information**."

AI in Local Newsrooms

- Newsquest has expanded its workforce of "AI-assisted reporters" from seven to 36 in 3 years, embedding them into local newsrooms across the UK.
- These reporters use an AI-powered CMS to convert press releases into news stories, a move designed to free up traditional journalists to focus on “original content people are willing to pay for.”



https://pressgazette.co.uk/publishers/regional-newspapers/newsquest-36-ai-assisted-reporters-non-canon-news-disintermediation/

Activity #2: What would you do in this situation?

1. You are going to see three real world scenarios in which journalists may decide to use AI technologies.
2. For each scenario, you will have four options.
3. I need you to stand up and move around the room depending on your answer:
    a. If your answer is A, you will seat in the right area
    b. If your answer is B, you will seat in the left area
    c. If your answer is C, you will seat in the front
    d. If your answer is D, you will seat in the back
4. For each scenario, I will ask one volunteer to tell us why she/he chose each option.

Digital Journalism Dilemmas #1

Your newsroom is covering a drought crisis in East Africa. You need a compelling header image, but you don't have a photographer on the ground. A colleague suggests generating an "illustrative" image with Midjourney: cracked earth, a lone child, dramatic colour grading. The image would not depict a real place or person.

Option A — Use the AI image without disclosure. It communicates the reality emotionally; readers know feature images are sometimes symbolic.	**Option B** — Use the AI image but disclose clearly with the caption: "This image was generated using AI to illustrate drought conditions."
Option C — Do NOT use imagery at all. Better to have no visual than to use synthetic content in a real-world crisis story.	**Option D** — Commission or source authentic local imagery. Wait for real photos and delay publication; accuracy outweighs speed.

Digital Journalism Dilemmas #2

Your editor wants a quick explainer about a newly leaked UN report. Because of workload, they propose using an LLM to produce a first draft, which a journalist will then “lightly edit.” However, the LLM's draft includes fabricated links, generic phrasing, and misunderstandings of the report.

Option A — Use the AI draft and polish it. The human edits will fix errors; this speeds up production.	**Option B** — Use the AI draft only as a structural outline. Throw away the generated sentences but keep the suggested structure.
Option C — Do not use AI at all. The risk of hallucinations and subtle inaccuracies is unacceptable for public-interest journalism.	**Option D** — Use AI only after manually feeding it verified excerpts from the UN report. Force the system to “stay inside the box” of verified facts.

Digital Journalism Dilemmas #3

You are producing a short digital documentary about journalists detained in Myanmar. One of your interviewees fears retaliation if their real voice is broadcast. A colleague proposes using AI voice-cloning with pitch-shifting and style transfer to protect their identity.

Option A — Use the AI-generated voice and disclose it. This protects the source while maintaining transparency.	**Option B** — Use the AI-generated voice without disclosure. The emotional impact is preserved; disclosure may confuse audiences.
Option C — Use a human actor's voice instead. Avoid the ethical murkiness of AI, keep the story authentic.	**Option D** — Avoid audio altogether. Use text overlays or subtitles; safety outweighs audiovisual richness.

Ethical Issues in Algorithmic Journalism

Transparency & Disclosure: Should audiences be told when a news article is written by an algorithm or when their news is curated by one? Ethical journalism values transparency, yet many algorithmic processes are hidden (Diakopoulos & Koliska, 2017)

Bias and Fairness: Algorithms can inherit bias from their training data or design. They may skew coverage or miss stories affecting minorities if not carefully audited: the myth of algorithmic neutrality is false (Brake, 2017)

Ethical Issues in Algorithmic Journalism

Accountability: Who is responsible for errors or ethical lapses by “robot reporters”? The lack of a human byline complicates accountability when, say, an automated story publishes false information (Jamil, 2023)

Journalistic Values: Can traditional values like accuracy, truth, and minimising harm be upheld?

> For example, an algorithm might publish names from a database (a crime report) without sensitivity a human editor would apply, raising questions of ethics and judgement (Ward, 2018)

Algorithmic Accountability in Journalism

Defamation & Liability: If an algorithmically generated news story defames someone (e.g., publishes a false claim that harms reputation), the news organisation could face libel liability. Determining fault is tricky – was it a programming error, the data source, or oversight? Unlike tech platforms, publishers can't easily escape liability by blaming "the algorithm" (Lewis et al., 2019)

Authorship & Copyright: Legally, algorithms cannot be authors. Automated content's copyright typically belongs to the organisation or programmer. However, the lack of a human author raises questions: Who owns or is accountable for the work? (Jamil, 2023)

Algorithmic Accountability in Journalism

No “Safe Harbor” for News Orgs: Internet platforms often have legal protections as neutral hosts (e.g., Section 230 in the US), but a news outlet that deploys a news-writing bot is considered the content creator. Therefore, it must meet legal standards for accuracy, attribution, libel, etc., just like any human journalist would (Lewis et al., 2019)

Regulatory Gaps: Most jurisdictions have outdated media laws that don’t explicitly address AI-generated content. This creates uncertainty – e.g., how to correct errors at scale, or whether mandatory disclosure of automated content is needed. There are growing calls to update regulations and guidelines to cover algorithmic journalism (Jamil, 2023)

Future Trends – Generative AI and the Next Wave

Generative AI Advances: New AI models (e.g. GPT-3/ChatGPT, GPT-4) can produce remarkably human-like text. This opens possibilities for more complex automated news stories (beyond templated data-to-text), but also new challenges with accuracy and editorial control

Enhanced Personalisation: Future news delivery might use AI to not just select stories but also tailor writing style or depth to individual readers. Creative variation (like automatically generating different versions of a story for novices vs experts) is on the horizon

Future Trends – Generative AI and the Next Wave

Deepfakes & Misinformation: As AI improves in generating text, audio, video, there's a flip side – the potential for fake news and deepfake content. Journalists will need advanced tools to verify authenticity and possibly deploy algorithms to fight AI-driven misinformation

Need for New Frameworks: The rapid evolution of AI in journalism is prompting industry and regulators to establish guidelines – from ethical AI use policies in newsrooms to possibly laws requiring AI content disclosure. The coming years will likely see clearer norms (Jamil, 2023)

Final Ideas - Critical Perspectives

1. **Journalistic Labour**: Automation can displace some routine reporting jobs, but it also creates new specialisations (data journalists, AI editors). Optimistically, it augments journalists (letting them focus on high-level work), but there's concern over job losses and deskilling in the newsroom
2. **News Content & Quality**: Automated content increases volume and coverage (e.g., hyper-local or data-heavy topics get reported), potentially enhancing the informational supply. However, risks include more homogenised, template-driven stories and less investigative depth or originality in automated outputs (Jia, 2020)

Final Ideas - Critical Perspectives

3. **Power and Gatekeeping**: Control over news shifts towards those who design and control algorithms (often tech firms). Platforms' algorithms influence which news gains visibility, raising issues of power concentration and democratic accountability (Brake, 2017). Journalistic authority is challenged by opaque systems determining news prominence
4. **Trust and Audience Perception**: Some studies show readers find automated news nearly as credible as human-written (Jia, 2020), yet transparency lapses or algorithmic errors could undermine trust in media. The growing role of unseen algorithms may distance audiences unless news outlets actively maintain credibility through accountability measures

Today's Learning Objectives

By now you should …

1. Understand the various roles algorithms play in journalistic content **production**, **newsgathering**, and **distribution**.
2. Analyse how algorithmic tools and systems are transforming **journalistic practices** and **decision-making**.
3. Examine the implications of **algorithm-driven journalism** for journalistic labour, ethics, and legal responsibility.

The Plan for this Week

8 April

Yesterday we talked about key approaches and principles of digital storytelling, and I introduced you to some examples from around the World.

9 April

Today we'll talk about algorithms and data, and how both are impacting the production and distribution of news content.

10 April

During the seminar, I'll share some best practices in data journalism, and we'll work together on an example. Bring your own laptop if you can.

University of Sheffield

Questions?

Algorithms and Journalism

Dr. Dani Madrid-Morales, d.madrid-morales@sheffield.ac.uk

Shandong Normal University, 9 April 2026

University of Sheffield

A World TOP 100 University