

AI's Role in the Future of News:

A White Paper on the *Newsroom America* Project

“

A collaborative effort to explore the future of journalism using Artificial Intelligence (AI).

Journalism deserves better tools, not fewer journalists. So we're building super-powers for journalists, not substitutes.

”

Peter Fowler & Andrew McMillan

Kinga VoxPop Ltd

April 2025

Executive Summary

The *Newsroom America* project addresses the challenges facing modern journalism, including the collapse of traditional business models and the rise of disinformation, by exploring the use of Artificial Intelligence (AI) to enhance news production.

This white paper introduces a working proof-of-concept platform that leverages AI to automate the transformation of press releases into accurate, clear, and neutral breaking news stories.

Key benefits include increased speed and volume of news production, improved transparency through a *Three-Wire Publishing Model*, and a focus on maintaining journalistic values such as accuracy, fairness, and accountability.

The paper details the first AI roles within the system (AI Reporter, AI Journalist, AI Editor) and presents a case study demonstrating how AI can be used to produce balanced and contextualised news reports. It takes a critical look at some of the problems yet to be solved.

Ultimately, *Newsroom America* aims to demonstrate a sustainable future for journalism by combining the efficiency of AI with essential human oversight.

Newsroom America (newsroomamerica.com) is the prototype of a white label AI-native newsroom that can be created with the *VoxPop AI News Engine*, which can be customised to provide any number of specialised interactive news services.

Comparing Traditional Newsrooms With AI Newsrooms

To clarify the need for innovation in journalism, the table below compares the traditional newsroom model with the *Newsroom America* approach. It highlights key differences in cost, speed, transparency, scalability, and editorial oversight—demonstrating how AI, when guided by human values, can enhance rather than undermine journalism.

Dimension	Traditional Newsroom	Newsroom America
Operational Cost	High fixed costs (salaries, printing, distribution, office space).	Minimal overhead; lean team + AI handles routine reporting.
Speed of Publishing	Hours to days depending on staff availability.	3–5 minutes for factual press release to publish-ready story.
Coverage Volume	Limited by newsroom size; prioritises a few top stories per day.	Thousands of stories/day possible; full public record coverage.
Human Resources	Large editorial team needed for all stages (reporting, editing, review).	Small team supervises AI pipeline; focuses on high-value journalism.
Transparency	Opaque process; readers can't see how the story was produced.	Radical transparency: story origin, source, AI/human roles all visible.
Correction Process	Manual, sometimes slow; few outlets publish correction logs.	Real-time logging, public audit trail, AI-assisted detection.
Source Handling	Varies widely; not all sources consistently vetted.	Source reliability scoring system governs AI behavior and publication eligibility.
Audience Engagement	Mostly passive; comment sections often toxic or closed.	Moderated speech/text input, comment coaching, AI-aided engagement.
Scalability	Difficult to scale without increasing staff and cost.	Highly scalable; same AI backbone can serve local, regional, national, and international coverage.
Story Evolution	Static articles; follow-ups require new stories.	Living stories update over time, link related developments automatically.
Ethical Governance	Depends on internal policies; often ad hoc.	Codified editorial charter, AI oversight dashboard, community and legal accountability.

Table of Contents

1. Executive Summary.....	2
2. Comparing Traditional Newsrooms with AI Newsrooms.....	3
3. Introduction.....	5
4. The <i>Newsroom America</i> Approach.....	6
5. Transparency and Trust: The Three-Wire Publishing Model.....	13
6. Ensuring Accuracy and Accountability.....	15
7. Ethical Framework and Editorial Governance.....	18
8. Moderated Audience Participation.....	23
9. Roadmap and Future Innovations.....	25
10. Monetisation.....	33
11. Conclusion.....	35
12. About the Authors.....	38
13. Appendix A: Glossary of Terms.....	39

1. Introduction

Newsroom America is a collaborative effort to explore the future of journalism using Artificial Intelligence (AI). Journalism deserves better tools, not fewer journalists. So we're building superpowers for journalists, not substitutes.

Founded in New Zealand in 2010 by journalist and entrepreneur, Peter Fowler, and open-source technologist, Andrew McMillan, *Newsroom America* is built as a public-service project at the intersection of journalism and civic technology. It was conceived as a response to the collapse of traditional news business models and the rise of disinformation. The guiding idea is simple yet radical: AI should enhance journalistic values, not undermine them.

Significant advancement in Large Language Models (LLMs) has allowed our vision of an AI-native newsroom to become a reality. The proof-of-concept platform (Screenshot below) automates the transformation of verified press releases into accurate, clear and neutral breaking news stories and publishes them.

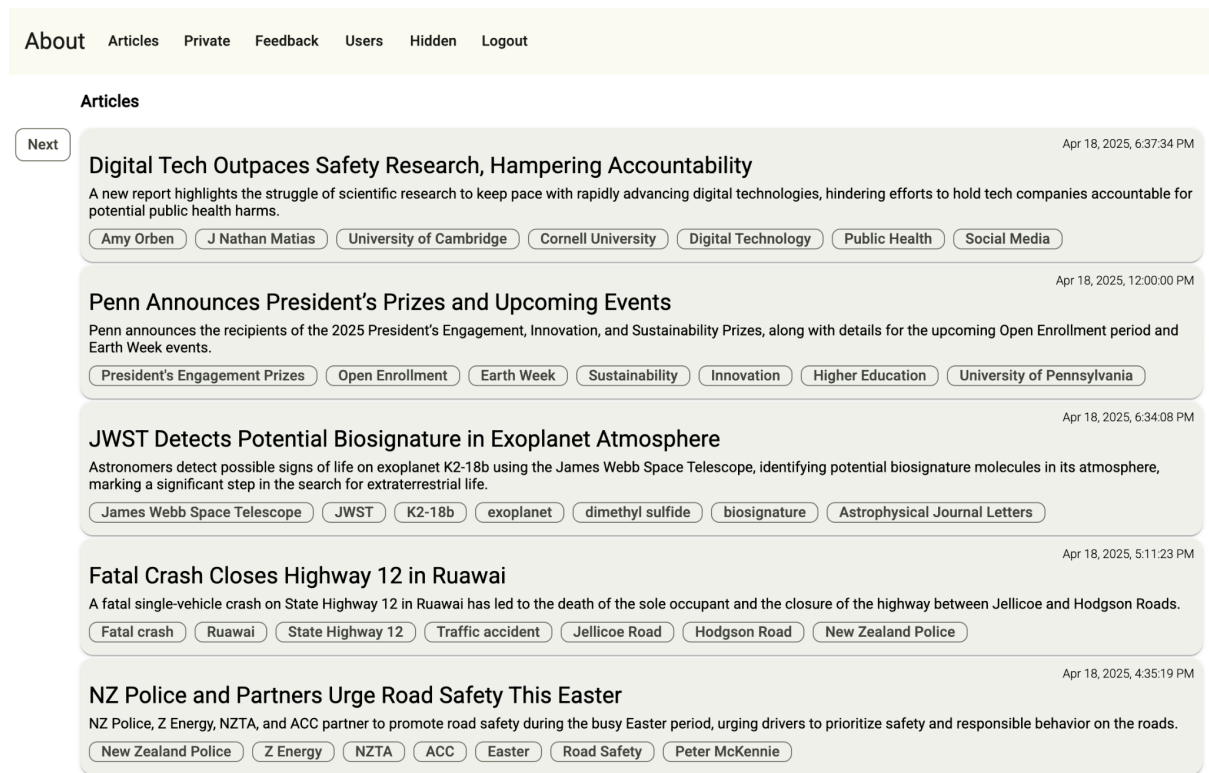


Fig. A: PROOF OF CONCEPT - Newsroom America has automated the “first mile” of reporting.

This is just the first step of many - "The first mile" of journalism - which is the initial stage of newsgathering. This is when new information is first observed, captured, or reported. This phase is closest to the source of truth but also the most vulnerable to distortion, error, or misinformation if not verified. The next step is to teach our AI journalists to write as a journalist, develop a story with reaction, provide context, determine source reliability and so on and so forth. But we have proof of concept that this is possible.

The core conviction driving *Newsroom America* is that journalism is not mere “content” but an essential public service, alongside police and hospitals. Its role is to verify facts, ask tough questions, provide context and inform the public. By harnessing AI to assist in these tasks (rather than replace them), *Newsroom America* aims to rebuild the speed, reach, and transparency of factual reporting without sacrificing editorial integrity.

Instead of automating journalism away, *Newsroom America* augments a human team with AI “colleagues”. It speeds up routine news production while rigorously upholding accuracy, fairness, and accountability. In this model, machines handle volume and speed, but humans “set the values” and provide the judgment and nuance that algorithms alone cannot.

By combining the reach of automation with the editorial judgment and transparency of professional journalism, the platform seeks to demonstrate a new way forward for news in the fourth industrial revolution.

2. The *Newsroom America* Approach: AI-Augmented News Production

2.1 System Overview: At its core, *Newsroom America* operates a real-time newswire powered by Artificial Intelligence. The prototype ingests press releases and official statements delivered via email and within ~3 minutes produces a concise, neutral “news story”. The AI components rewrite the release into a breaking news story format – applying standard journalistic structure (headline, lede, body, quotes, etc.) – and tag each story with metadata such as topics, geography, urgency and source reliability. Crucially, every AI-generated story is *traceable back to its source*: the original press release or document is linked, and all AI transformations are logged. This means a reader (or editor) can always see what the story is based on, ensuring verifiability.

The newsroom currently focuses on factual, public-domain material (such as government announcements) to prove that the first mile of reporting can be automated in a *reliable and reviewable* way. By automating the task of reviewing press releases, determining their importance and news value, and turning them around into a breaking news story, human journalists are freed to focus on higher-level tasks.

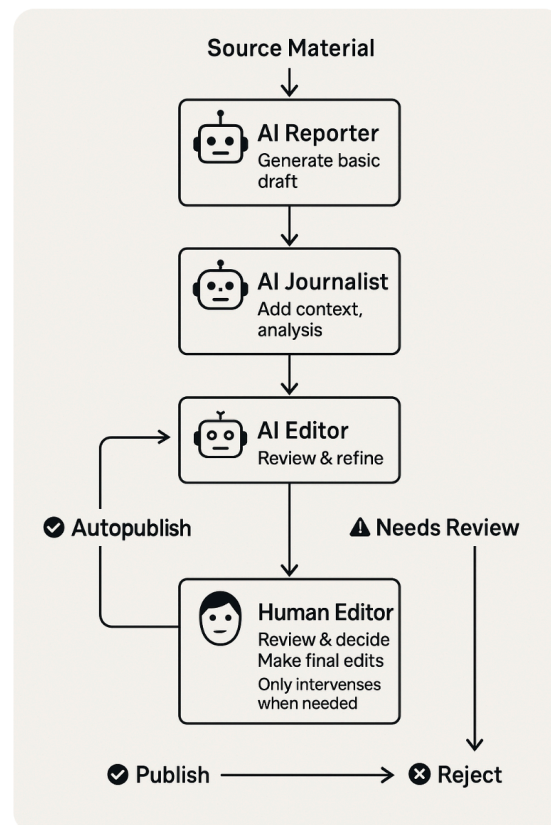
Take the example of a large earthquake. The AI can accurately turn around and publish a breaking news story based on USGS data in under three minutes. It can process and publish any warnings from the Pacific Tsunami Centre and keep them updated. The AI will continue to update the original story with information coming in from police, fire, utilities, hospitals and government officials. The journalists meantime can focus on creating exclusive content, such as ringing businesses near the epicentre of the earthquake to get a better understanding of the impact and report eyewitness accounts.

Notably, the platform also allows the audience to directly engage using text or voice, whether leaving eyewitness reports, news tips, pointing out corrections, rating each story’s accuracy

or just leaving comments. This creates a feedback loop to refine the AI’s performance and the newsroom’s decisions over time.

In short, *Newsroom America* is a working proof-of-concept for a native AI-assisted newsroom – one where automation handles the volume of news, and humans handle the values behind the news.

2.2 AI/Human Editorial Roles and Workflow: To replicate a newsroom, *Newsroom America* is building a virtual team of AI reporters and journalists, each with defined editorial roles. Together, they operate in a pipeline supervised by humans.



The key AI roles include:

- **AI Reporter (*Seymour Krelborn*)** – Our live automated reporter that turns raw source material into a basic breaking news draft, for proof of concept. This AI reporter ingests news data such as a press release and extracts the key facts, rewrites them in a clear news style, and structures the story with a headline, lede, and quotes. It also attaches standard metadata (date, categories, source info etc) and assesses the story’s urgency, importance and veracity. The AI Reporter does *not* add any information beyond the source – it doesn’t provide background or analysis. As a result, its output is factual but often bare-bones. Seymour simply rewrites press releases into 500-word or less news stories, faithfully sticking to the contents of the release. This yields fast, highly accurate summaries, but initially they may read as “flat” bulletins lacking context, particularly with political releases. Seymour will

eventually make an initial judgment of source reliability: if a source is unfamiliar or the content is potentially dubious or breaches the Editorial Policy, it flags the story for human review rather than auto-publishing. Currently human editors and the editorial board (*see later*) manually assign source reliability scores.

The screenshot shows a web interface for managing source reliability. At the top, there's a navigation bar with links: About, Articles, Private, Feedback, Sources, Users, Hidden, Logout. Below this, there's a form for adding a new source. The form includes fields for 'From:' (Boeing News <news@boeing.mediaroom.com>), 'To:' (boeing@feeds.newsroomamerica.com), 'Name*' (Boeing), and 'Source Tag' (Boeing). Below the form, there's a 'Trustworthiness: 95%' slider ranging from 'Untrusted (0)' to 'Highly Trusted (100)'. There are two toggle buttons: 'Auto Publish' (checked) and 'Do Not Auto Hide' (unchecked). A 'Save Changes' button is also present. Below the form, there's a section titled 'Articles' with a list of three articles. Each article has a title, a date, and a set of tags. The first article is 'Boeing Announces Q1 2025 Program Deliveries' dated Apr 9, 2025, 3:04:03 AM, with tags: Boeing, NYSE: BA, Commercial Airplanes, Defense, Space & Security Programs, Financial Results, AH-64 Apache. The second article is 'Boeing to Release First Quarter Results on April 23' dated Apr 4, 2025, 4:03:13 AM, with tags: Boeing, Financial Results, Kelly Ortberg, Brian West, NYSE: BA, Company Outlook, Aerospace. The third article is 'BOC Aviation Orders 50 Boeing 737 MAX Jets' dated Mar 31, 2025, 1:58:36 PM. Each article has 'Hide' and 'Unpublish' buttons. There are also 'Next' and 'Redo' buttons on the left side of the articles list.

Source Reliability Scores help determine whether a newsmaker, such as Boeing, can be auto-published or held for human review.

- AI Journalist (To be developed)** – A second-layer AI that will act as a beat reporter or analyst, adding depth, nuance, and context to the AI Reporter’s draft. The AI journalist’s job is to answer the questions the first draft didn’t: *Why does this news matter? What led to it? What are other perspectives?* The AI Journalist might pull in historical background (“What is this agency and how was it created?”), cross-reference related news or previous statements, and suggest additional angles. It can incorporate relevant data or explanations (for example, inserting a short explainer or a timeline of key events). It also attempts to introduce counterpoints or reactions – for instance, if an official claims something controversial, the AI Journalist will look for past remarks by experts or opposition figures to provide balance. Essentially, the AI Journalist acts as a research assistant and writer that enriches the story with context and multiple viewpoints. The output is a more comprehensive second draft that’s closer to what a human journalist might produce after doing some background research and getting quotes from multiple sources. Notably, the AI Journalist will be careful to flag anything it isn’t confident about – e.g. it will mark potentially biased phrases or unsupported claims for a human to review.
- AI Sub-Editor (To be developed)** – An AI sub-editor that ensures the story meets the editorial policy and newsroom standards of quality and fairness before publication. The AI sub-editor reviews the combined draft, fixes grammar and style issues, and checks for compliance with editorial policy: Is the tone neutral and

factual? Are all claims attributed to a source? Have important facts been fact-checked or at least noted for follow-up? It will rearrange paragraphs if needed for better flow or emphasis, and craft a headline that accurately reflects the story without sensationalism. Importantly, the AI Editor makes a publishing judgment by assigning the story an *Editorial Status Label*. Based on predefined criteria (including the source's reliability level and the story's content), it will tag the story either  **“Reliable Source / Autopublish”** (safe to publish immediately),  **“Needs Review”** (requires a human editor to approve or improve it), or  **“Do Not Publish”** (the draft fails standards or is too sensitive) For example, a mundane weather service update might get auto-published, whereas a politically charged announcement would be held for human oversight by default. The AI Editor can also highlight ethical red flags (such as potential defamation risk or privacy concerns) to make sure such stories get human attention. In essence, this AI acts as a diligent copy editor combined with a traffic cop, enforcing rules and routing stories appropriately in the workflow.

- **Human Editor** – The human journalists and editors are the ultimate decision-makers and fail-safes in this newsroom. A human editor oversees the entire process through an admin dashboard and is alerted to any story that the AI marked  "Needs Review" or  "Do Not Publish. Humans intervene in high-risk or complex stories – for example, anything involving serious allegations, national security, or legal issues will be routed to a person before publishing. The human editor can edit the AI's copy, request additional information, or kill a story that isn't fit to run. They also handle cases the AI isn't equipped to, such as investigative pieces that require original reporting, or nuanced judgment calls that algorithms can't make. In day-to-day operations, one human can supervise many automated outputs, focusing only where needed. If everything looks good, the human might simply do a quick final read and approve auto-publication. But if the AI has made an error or missed context, the human editor will catch it. Essentially, the human remains in charge of the newsroom; the AI acts like a multitude of junior staff under the editor's guidance. This human-in-the-loop model preserves accountability: nothing gets published that a responsible person hasn't ultimately green-lit when it matters.

This multi-layer AI workflow allows *Newsroom America* to publish routine news quickly while escalating anything sensitive to human judgment. It dramatically increases output (the AI can churn through thousands of press releases 24/7) but with safeguards so that speed does not trump standards. As a result, a tiny team in the smallest of towns can manage a continuous newswire without letting automated errors or biases slip through unchecked. This structure preserves editorial judgment, transparency, and accountability and makes mass publishing feasible even with minimal human staffing.

2.3 AI Output Case Study: Uncritical platforming, Loaded framing

In April 2025, the U.S. State Department issued a press statement about shutting down a government disinformation office (the Global Engagement Center, rebranded as R/FIMI). Here's how *Newsroom America's* AI reporter Seymour Krelborn handled this story, and how the layered process will address the critical issues that arise:

- **AI Reporter's Output:** Seymour quickly produced a news flash based on the press release. It generated a headline and lede along the lines of "*State Department Shuts Down Counter Foreign Information Office*", and summarised the release: that the office (formerly GEC) was being closed for allegedly "spending millions to censor Americans," according to the statement. This draft faithfully conveyed the content of Secretary of State Marco Rubio's announcement – including quotes about protecting Americans' rights and ending "government-sponsored censorship". On facts, it was accurate. However, from an editorial perspective, the draft had serious issues: it essentially platformed Rubio's narrative without question or context.

Issues identified: The automated story, if published as-is, would lack balance and critical context in several ways:

- *Uncritical platforming of a single narrative* – The piece reflected only Secretary Rubio's perspective, conveying his claims as if they were objective facts. It did not clarify that Rubio's stance was a political position or provide any hint that others might view the situation differently. Nor did it explain what the GEC actually is or that it had a mandate to counter foreign propaganda, leaving readers with only Rubio's characterisation of it as a "censorship" office.
- *Lack of historical context* – The draft mentioned the office's closure but offered no background on how the GEC was established or what it had been doing (its mission to counter disinformation from foreign adversaries). Important context, such as the GEC's bipartisan origins in 2016 and its role in U.S. information policy, was missing, which could mislead readers about the significance of the change.
- *Loaded framing without attribution* – The press release used charged terms such as "metastasized into a censorship bureaucracy" and "silencing political speech." The AI included some of these phrases directly, but failed to attribute them to Rubio. Phrases such as "censorship bureaucracy" ended up in the story as if they were objective descriptors, when in fact they are Rubio's very subjective (and emotive) framing. This kind of unchecked language could carry bias.
- **Impact:** Published in that form, the story would essentially read as a government press release, not a piece of journalism. Readers would have no context to judge Rubio's claims, and no alternative views – a scenario that could inadvertently spread one-sided or misleading information. This example mirrors a common human failure in fast news: reporting the claims in a press statement without investigating them further. *Newsroom America's* design will catch these pitfalls by adding the next layers as follows:
- **AI Journalist's Enhancements:** Recognising the story's importance (a major policy shift with political overtones), the system will flag it for the second-layer AI Journalist. This AI rewrites the draft with significant additions and modifications to

ensure balance and depth. Key improvements included:

- **Background and context:** The AI Journalist would insert background explaining what the GEC (Counter-Foreign Information office) is – noting it was created in 2016 under Obama with bipartisan support in Congress. It would clarify the GEC’s original mission (to combat propaganda from groups such as ISIS and nation-state adversaries) and how it later evolved, including expansions under Trump and Biden. This historical context helps readers understand the office that Rubio shut down was not simply a “censorship bureau” but part of a long-running effort against foreign disinformation.
- **Controversies and critiques:** The rewritten story would include the broader debate around the GEC. For instance, it would be noted that critics had accused the GEC of “mission creep” – blurring the line between fighting foreign propaganda and monitoring domestic speech. It would reference specific controversies mentioned elsewhere (e.g. involvement of the Stanford Internet Observatory and resulting FOIA lawsuits). By mentioning these, the story acknowledges there were genuine concerns about the GEC’s work, which gives context to Rubio’s actions without endorsing his framing.
- **Counterpoints and reactions:** The AI Journalist would also seek out other perspectives. It could accurately add that *civil liberties advocates and foreign policy analysts had previously voiced concern* both about government overreach and about the consequences of dismantling disinformation programs. In other words, it would indicate there is a tension: some applaud the shutdown as a win for free speech, while others worry it removes defenses against foreign disinformation. The draft would note that responses from watchdog groups were not immediately available (inviting follow-up). It would find a quote from an archive showing prior bipartisan support for the GEC’s work, to contrast Rubio’s current criticism. If ACLU or former officials had publicly commented, the AI would have included those quotes; lacking them, it at least flags the types of reactions expected.
- **Clarifying attribution:** The AI Journalist and Editor together handle the loaded language issue. In the improved draft, phrases such as “metastasized into a censorship tool” are clearly attributed to Rubio’s *claims* rather than stated as fact. For example, it might write, *Rubio said the program had “metastasized” into a domestic censorship tool*, putting the contentious word in quotes and in Rubio’s mouth. The AI Editor would explicitly flag “metastasized” as an emotionally loaded term and move it into a direct quote from Rubio. The headline would be changed to something like *“Rubio Shuts Down Disinfo Program, Citing Free Speech Concerns”* – framing it as Rubio’s stance and indicating debate, rather than a neutral description of the office.

Thanks to these enhancements, the second draft becomes far more robust: it reports Rubio’s announcement with context and skepticism, making clear what was being asserted and by whom. It included historical facts about the GEC and noted the possible implications of its

closure (for example, how adversaries such as Russia or China might cheer the decision, or how it fits into a larger free speech vs. security debate). In effect, the AI Journalist turns what was initially a one-sided report into a more comprehensive news story that a human journalist might have written with more time.

- **AI Editor's Checks:** The third “pair of eyes” over the Rubio announcement would be the AI Editor, yet to be developed. The AI Editor would review this enriched story, performing a thorough quality check: verifying that all critical statements were attributed, scanning for any factual errors against the source, and ensuring the tone remained impartial. In this case, the AI Editor would have made some specific tweaks as noted: it flagged the word “metastasized” (as mentioned) and ensured it was in quotes; it adjusted the headline to be more informative and balanced about the *nature* of the story (emphasizing the policy shift and that there were differing views); and it even suggested adding a sidebar explainer titled “Understanding the GEC” to give readers quick background on that office’s purpose. The presence of such a sidebar or an inline explainer would directly tackle the knowledge gap for readers unfamiliar with the agency. After these edits, the AI Editor assigns an output label. In this scenario, given that the story now included multiple viewpoints and had its facts cross-checked, the AI Editor might label it “ *Publish with attribution; includes counterpoints and context.*” That label indicates the story meets the criteria to run on the site, with transparency about sources and sufficient context. However, because this is a high-impact political story, the AI Editor would flag it as a high-importance piece (“importance: 85” for example) and alert a human to take a look before publishing.

In summary, the GEC story example shows how *Newsroom America’s* multi-layer editorial process will catch the nuances that our single-pass AI reporter Seymour is missing. The first draft (AI Reporter) was fast but shallow. The second layer (AI Journalist) injected needed context, balance, and nuance, turning the story from essentially a press-release regurgitation into a piece of journalism with analysis and multiple perspectives. The final AI Editor layer enforced quality and editorial policy adherence, ensuring transparency about what was said and by whom. At the end, a human editor could confidently publish the story knowing it is fair, contextualised, and in line with the accuracy mission. This layered approach – machine plus human – is designed to prevent the inadvertent publication of a one-sided narrative, demonstrating how AI can be used responsibly in a newsroom setting. What started as an automated output rife with potential bias is transformed (through AI assistance and human oversight) into a fair, informative news report.

It’s worth noting that for a routine story with no such controversies – say a local weather alert or a scientific discovery– the AI might handle the entire workflow with minimal human input, publishing directly to the site. But for anything of major consequence, the system either automatically involves the human or at least provides all the contextual enhancements so that when a human does review it, they have a rich draft to work from. In all cases, no story is simply dumped online by a black-box AI – there is always an editorial audit trail and explicit labeling (discussed next) to maintain accountability.

3. Transparency and Accuracy: Three-Wire Publishing Model



From the outset, Newsroom America was designed with radical transparency in mind. A cornerstone of this is the *Three-Wire Publishing System*, which clearly indicates to the audience how a given story was produced and vetted. Traditional news agencies historically had different “wires” or feeds for breaking news, analysis, etc. *Newsroom America* adapts this concept to the age of AI, and plans three parallel newswires, each with distinct editorial characteristics and visual labels:

3.1 Press Release Wire – This wire is a live feed of raw, unedited press releases and official announcements. Essentially, anything that is received from verified sources (government agencies, businesses, NGOs, etc.) can be published *verbatim* here for the record. The purpose is to provide direct access to source documents so readers can see exactly what was said, without delay or filtering. An editorial note is attached making clear that Newsroom America has not edited or fact-checked these items. By having a Press Release Wire, the platform separates original source material from news articles – letting the public inspect the raw input if they wish. This supports transparency and also helps distinguish between what information came directly from a source versus what the human or AI did with it. (For example, in the GEC case above, the State Department’s full press statement could be available on the Press Release Wire for those who want to read it themselves.)

3.2 AI Auto-Publish Wire – This second wire consists of news stories that were automatically generated by AI and published without immediate human review. These are typically short factual updates derived from structured inputs such as press releases, public data feeds, or breaking emergency alerts. Examples might include: a brief on a just-issued earthquake report, a stock market closing update, or a police bulletin about a road closure. The value of this wire is speed and scale – it can deliver factual breaking news stories within minutes, far faster than a human-only team could. However, every story on this wire is clearly labeled as machine-generated, with a note that no human editor has reviewed it yet. If errors are discovered (by the AI’s own checks, by users, or by staff), corrections are made and transparently logged

(more on corrections in the next section). The *AI Auto-Publish Wire* is essentially the “AI-first” news feed, capitalising on automation to keep up with real-time events. But transparency is paramount: readers see the blue AI label and know to interpret these stories accordingly (with perhaps a bit more caution, understanding they might be preliminary).

3.3 Human-Reviewed Wire – This is the premium news feed containing all articles that have undergone full human editorial review. Any story on this wire has been fact-checked, edited, and approved by a human editor to meet the highest standards of accuracy, context, and clarity. These stories carry bylines and sourcing notes, and often are longer-form or more complex pieces – e.g. investigative reports, explanatory journalism, or any AI-written story that a human editor has revised and approved. The *Human-Reviewed Wire* represents *Newsroom America*’s gold-standard output. Readers can consume these stories with confidence that they’ve been thoroughly vetted in line with the newsroom’s editorial policy. Practically, this wire might include the day’s top news that was important enough to get a human touch, as well as original journalism from the team. Every article here is labeled to show it is human-reviewed, reinforcing that a journalist’s eyes (and accountability) are on it.

By segmenting content in this way, *Newsroom America* does something unusual in media: it openly distinguishes machine output vs. human-refined output vs. raw source input, for every single item. This is a direct answer to the trust and transparency challenge of AI in news. Rather than hide the role of AI, the platform highlights it. If a story was auto-generated, the audience knows that at first glance (via color coding and icons, as well as explicit labels in the story page). Readers who care about maximum immediacy can follow the AI wire, whereas those who prioritize fully checked news can stick to the Human-Reviewed wire – or they can read both, but with an understanding of the difference. Moreover, by publishing the Press Release Wire, *Newsroom America* ensures the original sources are one click away for anyone to verify.

This three-wire model not only serves readers but also partners or other journalists: it’s evident which content is raw, which is AI-curated, and which is edited, enabling *informed news consumption*. In an era of information overload and widespread skepticism, such transparency is intended to build credibility. As the project describes, “*each story is labeled not just with its topic, but with how it was created – allowing readers to choose their level of reliability, speed, or source proximity.*” A reader might, for instance, read an AI-written breaking story on the *Auto-Publish Wire*, then click the source link to the Press Release for verification, and later see a more fleshed-out version on the human wire. By inviting this scrutiny, *Newsroom America* aligns with its ethos that transparency underpins trust. The audience is treated as a partner in verification, not a passive consumer of a black-box feed.

4. Ensuring Accuracy and Accountability

Speed and volume are meaningless if the information isn't accurate. Thus, *Newsroom America* couples its AI-driven output with stringent accuracy control mechanisms and accountability practices. The goal is that AI-generated journalism meets or exceeds the standards of the best human-only newsrooms. To achieve this, the project has begun implementing a combination of story scoring, error tracking, and correction auditing that continuously measures and improves quality.



4.1 Story Scoring System: Stories are evaluated using a detailed rubric called the *Newsroom America AI Journalist Scoring Rubric*, a 100-point scale that breaks down key dimensions of news quality. This rubric is modeled on the criteria top news organisations (Such as the BBC, AP, and Reuters) use internally for editorial excellence. The components of the score include:

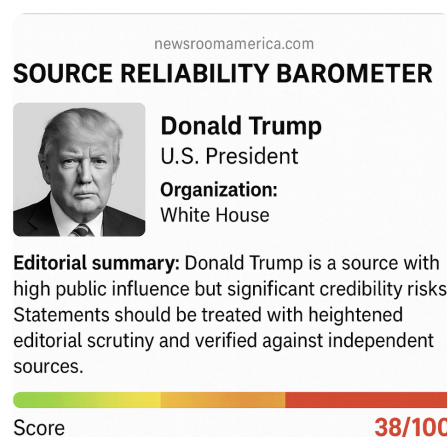
- **Factual Accuracy (25 points)** – Are all facts in the story correct and supported by the source or verified data? This is to catch any mistakes or “hallucinations” the AI might introduce. Ideally, the AI should score high here if it’s faithfully using the press release content and cross-checking names, dates, figures, etc.
- **Attribution & Sourcing (10 points)** – Does the story properly attribute claims and information to the correct sources? For example, if a statistic is mentioned, does it say who provided that statistic? If a claim is controversial, does it attribute it (“X said...”) rather than stating it as fact? This prevents ambiguity about where information comes from.
- **Contextual Integrity (10 points)** – Does the story provide essential context, or does it risk misleading by omission? This item checks if the AI added background or clarifications when needed (or at least didn’t strip out context present in the source). It’s a measure to ensure the story stands on solid ground and makes sense to a reader who isn’t already knowledgeable about the topic.
- **Headline Accuracy (10 points)** – Is the headline accurate and fair? Clickbait or exaggerated headlines are penalized. The headline should reflect the story’s content without distortion – a crucial aspect because many readers only see headlines.
- **Correction Transparency & Speed (15 points)** – If errors were found, how quickly and openly were they corrected? This is an innovative metric: it rewards the

system for catching and fixing mistakes promptly. Logging a correction and updating a story within minutes, for instance, would contribute to a high score here. This incentivises the AI to not only strive for accuracy but to be forthright when something needs fixing.

- **Editorial Policy Compliance (30 points)** – Does the story adhere to Newsroom America’s editorial standards and ethical guidelines? This covers a broad range: neutrality of tone, avoiding defamatory language, respecting privacy (e.g., not naming a minor crime victim), and so on. Essentially, it checks if the story content violates any “red lines” set by editorial policy.

The system computes this score (some components can be assessed automatically, others might come from a human editor’s input or a checklist the AI Editor goes through). If a story’s score falls below a certain threshold – say 85/100 – it is flagged for human review before or shortly after publishing. A score above, e.g., 95 would indicate the story is of very high quality, comparable to what a top newswire journalist might produce. This scoring mechanism provides a measurable benchmark for AI-generated content, allowing the team to track improvements over time and identify weak spots (e.g., if many stories lose points on “contextual integrity,” that tells editors to focus on that in the next AI training). It’s also defensible externally: by basing it on known standards of elite newsrooms, Newsroom America can show it is holding AI to the same bar as human journalists.

Additionally, the platform evaluates Story Importance and Source Reliability as part of its scoring and decision logic. Each story is assigned an “importance” value from 0 to 100 estimating its news value. Routine, low-impact news (a minor local event) might score <30, whereas a breaking national story or something surprising and broadly impactful might score >70. This importance score helps the AI Editor and human editors prioritise. For instance, the AI Editor is programmed to “flag high importance stories to humans” – meaning even if an important story is accurate, a human should double-check it simply due to its significance. Meanwhile, sources are classified by Reliability level: content from reliable sources (e.g., an official government press office or a known news agency feed) can be used with more confidence, whereas material from unknown or less reliable sources is automatically held for human inspection. In practice, an AI Reporter might determine “*Source = unfamiliar activist group, therefore do not autopublish*”, regardless of the story’s content, until a person verifies the source.



Conversely, a press release from the Prime Minister’s office might be considered reliable enough to auto-publish on the AI wire – but even then, if it’s very important, it might still be routed to a human for a quick look. The interplay of importance, reliability, and quality score forms a comprehensive story scoring system that guides what happens to each piece of content (publish now, hold, or escalate). This ensures that, for example, an AI will not autonomously publish a scandalous claim from a dubious source at 3 AM without a human in the loop, because either the reliability rating or the content scoring would catch it and require human oversight.

Real-Time Correction Auditing: No matter how good the AI or the process, errors will sometimes occur – what’s critical is how they are handled. Borrowing from the best practices of news organisations, *Newsroom America* plans a real-time correction log and auditing system for AI content. Every time an error is detected in a story (whether by an automated fact-check, a human editor, or even a user report), the system logs it in a structured way. The log entry includes: what the mistake was (factual error, typo, misattribution, etc.), how it was caught (for example, a user clicked a “flag” button, or an editor manually noted it), and how quickly it was fixed. This log is not just internal; the plan is to make it accessible via dashboards or even a public-facing feed. Imagine a publicly viewable list of corrections showing: “10:32 AM – in story X, corrected the spelling of CEO’s name; 11:15 AM – in story Y, updated the outcome of a court case that was initially incomplete.” This level of transparency is rarely seen in traditional media (which might issue corrections but not catalog every minor fix). However, because Newsroom America can automate much of this, it aims to “log every correction, no matter how small”. This not only builds public trust by showing that errors aren’t hidden, but it also provides a valuable dataset for improving the AI. By studying the correction log, developers can identify systematic errors (for example, if many corrections involve numeric figures, perhaps the AI parsing of numbers needs improvement).

The Correction Log ties into the scoring as well: the speed and transparency of a correction feeds back into a story’s quality rating (as noted, “Correction transparency and speed” is 20 points of the rubric). If a story was initially published with a minor mistake but fixed within minutes and duly noted, the system essentially gives itself some credit for that responsiveness, whereas a slow or opaque correction process would be a mark against. Over time, Newsroom America will be able to report an “AI accuracy rate” and error frequency thanks to this rigorous tracking. For instance, it could say “In Q2 2025, our AI-generated stories had a factual error rate of X per 1000 words, which improved by Y% after deploying a new model,” etc., analogous to how a manufacturing process might track defect rates. The ambition is to set an industry benchmark for automated journalism accuracy and accountability.

In practice, if a reader spots an issue in an AI-written article on *Newsroom America*, they can comment or flag it; the moderation tools (discussed later) will route this to an editor, and if confirmed, a correction will be issued and logged, often within the same news cycle. Readers might even see a note on the story like “Updated at 14:00 to correct X...” much as they would on a reputable news site. Additionally, any substantive corrections would likely be annotated in the Transparency Ledger for that story (part of the governance features). By integrating this with AI, the system can potentially catch certain errors automatically – for example, an Accuracy Evaluation Prompt is part of the admin interface that lets an editor re-run a story through a fact-checking AI to score it or find discrepancies. All of this reflects a

belief that acknowledging and fixing mistakes openly is essential to credibility, especially when using AI. As the project literature notes, in a “post-truth” environment, “*radical transparency is not just good practice – it’s a survival mechanism.*”

Benchmarking and Improvement: With scoring and logs in place, *Newsroom America* can continually benchmark the AI’s performance against established standards. For example, if the target is to consistently hit a score of 90+ on each story, the team will know if they fall short and why (perhaps the AI is frequently missing needed context, dragging scores down – indicating a need to tweak the AI Journalist’s prompt). The initiative explicitly aims to track error rates over time, by model version and topic. This means the system will learn whether certain domains (say, science news or legal news) confuse the AI more often, leading to targeted training in those areas. Weekly or monthly QA reports are planned, and the ultimate step is to publish a public-facing corrections feed or dashboard so outsiders can monitor how the AI newsroom is doing. In short, *Newsroom America* treats accuracy not as a static promise but as a metric to be measured and improved, with the same rigor one would apply to algorithmic performance – except the performance metric here is editorial quality. By setting up these feedback loops (user feedback, editor oversight, logging, scoring), the system can get better and better, and stakeholders can hold it accountable to quantitative goals (e.g., keep factual error rates below 1% of content). This approach mirrors how top news organisations maintain trust – through a culture of correction and self-audit – but makes it more systematic and data-driven, leveraging the fact that AI can track every tiny detail.

In summary, *Newsroom America*’s accuracy and accountability framework consists of:

- a) **pre-publication safeguards** (AI Editor checks, scoring thresholds that trigger human review)
- b) **post-publication transparency** (three-wire labeling, visible source links, and a public audit trail for stories)
- c) **post-publication corrections** (real-time logging and updates of any errors).

Together, these ensure that if the AI ever “gets it wrong,” the system will catch it and fix it – and everyone will know. By openly modeling their error-rate and correction practices on the BBC/AP standard, they underscore that the measure of success is not zero mistakes, but how responsibly you handle mistakes. This ethos is crucial to making AI news reliable.

5. Ethical Framework and Editorial Governance

Deploying AI in a newsroom raises profound ethical questions. Who is accountable for automated content? How do we prevent hidden biases or undue influence? *Newsroom America* tackles these questions head-on with a comprehensive ethical framework and governance model to ensure AI serves the public interest and adheres to journalistic principles. All automation in the system is “subject to editorial values, human oversight, and public transparency,” as a core operating rule. This section outlines the guiding principles, policies, and governance structures that ground *Newsroom America*’s approach.

Core Ethical Commitments: The project has articulated seven key commitments that together form an ethical charter for AI-assisted journalism. These were developed to guide system design and editorial decision-making. They are:

1. **Human Accountability** – Humans retain ultimate control and responsibility. No AI system at Newsroom America operates autonomously without a mechanism for human intervention. Final editorial decisions, especially on sensitive stories, always rest with a human editor. This ensures that accountability can always be traced to a person, not an algorithm. If something goes wrong, a human is answerable, and a human can step in to correct the course.
2. **Radical Transparency** – The newsroom commits to complete transparency about how each story was produced. Every article carries an auditable trail of its origin: what the source was, whether AI was involved in writing or editing, and what review steps it passed through. The audience will never be misled into thinking AI-generated text is “reporting” by a human – AI contributions are clearly disclosed. This principle is embodied in features such as the three-wire labeling and the detailed story logs. It’s a rejection of the opaque algorithmic feeds Big Tech often uses.
3. **Source Skepticism & Classification** – The platform evaluates all content against a structured source reliability framework. It doesn’t matter if a statement comes from a government minister or a random blogger – the system treats *all* sources with consistent scrutiny. Powerful actors are not given a free pass. For example, official press releases are checked for veracity and context just as any source would be. An AI might flag that “*this claim comes from a political figure about their opponent*”, hence it should be contextualized or verified independently. This commitment is an antidote to AI naivety: the system won’t assume input data is true or unbiased simply because of its origin. A source classification model governed by human editors is in development to codify this (assigning reliability levels, political context, etc., to sources). Ultimately, every piece of content is tagged with metadata about the source (e.g., government press release, NGO report, eyewitness report) and handled accordingly.
4. **Accuracy Over Speed** – Speed is one of AI’s advantages, but *Newsroom America* will not sacrifice accuracy or thoughtfulness for the sake of being first. If a story needs more time to verify or add context, the system will delay publishing rather than put out a half-baked report. In practice, this might mean the AI holds back on auto publishing a sensational story until a human can review it, or it might mean prioritising a slower, thoroughly contextual piece on the Human Wire over a quick AI blurb on the Auto wire. This principle is a direct response to the contemporary media problem of chasing clicks and breaking news at the expense of getting it right. *Newsroom America* explicitly prioritises *factual integrity, attribution, and completeness of framing* even if that means missing a short-term scoop. An example will be built into the system: if the AI isn’t fully confident or if it detects that crucial context is missing, it marks the story as “Needs Review” and does not push it live until cleared. Better to be right than first.

5. **No Automation Without Oversight** – All AI actions are designed to assist, not replace, human editorial judgment. Automation is kept on a leash: sensitive content is always flagged for human review. For instance, any story about a violent crime, an election, or a health emergency would automatically require human editor approval before publishing, because these are high-stakes contexts where nuance matters greatly. Even beyond specific categories, the general rule is that AI outputs are monitored by humans (through the dashboards and alerts). This commitment also implies continuous human involvement in developing the AI itself – editors and journalists are part of crafting the AI prompts, setting its rules, and reviewing its performance. It ensures that we never end up in a situation where “the algorithm decided to publish this and no one knows why.” Oversight is baked in at every level.
6. **Continuous Audit and Correction** – The system is built for continuous self-improvement and openness to critique. It logs decisions and errors (as described earlier) and invites internal and external audits. If an AI-generated story is found to be misleading or wrong, Newsroom America is committed to revising, annotating, or retracting it just as in traditional journalism. Nothing is to be swept under the rug. This culture of correction is crucial because it acknowledges that mistakes will happen; what matters is detecting and addressing them transparently. The public correction feeds, the accuracy scoring, and the user feedback loop all serve this principle. In essence, it’s about creating a virtuous cycle where the AI and the newsroom learn from every mistake and where the audience can see that learning process.
7. **Public Interest Above All** – Above speed, above volume, above convenience, the guiding question is: *“Does this serve the public’s right to timely, factual, and comprehensible information?”*. Every design choice or editorial decision is measured against that standard. For example, introducing the Press Release Wire serves public interest by giving direct access to source docs – it’s not done for profit or tech novelty, but because it helps inform citizens. If any feature or AI capability undermined public understanding or trust, it would be reevaluated or removed. This commitment essentially restates classic journalistic ethics (serve the public, not special interests), updated for an AI context. It acts as a North Star for the project’s evolution: no matter what fancy AI features could be added, they must always bow to the overarching mission of informing the public and enhancing democracy.

These ethical principles are not just slogans; they are integrated into the platform’s architecture and policies. For instance, “Human Accountability” is evident in the requirement of human sign-off and the presence of an editorial board. “Radical Transparency” is operationalized through the three-wire system and transparency ledger. “Source Skepticism” is embodied in the source classification model being built and the AI Journalist’s habit of checking every claim. Together, they form a robust ethical framework that others in the industry can reference for responsible AI use in news.

Editorial Policy and Governance Structure: To enforce these commitments, *Newsroom America* is introducing an editorial governance model that mixes traditional oversight with innovative community involvement. At the top, a Human Editorial Board

(even if informal in early stages) governs the major decisions: they set the editorial policy that the AI is trained on and they review the classifications and rules the AI uses. For example, the board defines what counts as a “reliable source,” what kinds of stories should never be auto-published, and how to handle corrections or ethical dilemmas. This ensures that journalistic standards are encoded into the AI’s rules – the AI isn’t making up its own standards. One concrete example mentioned is a *source classification model governed by a human editorial board*, meaning that humans will decide how sources are rated (such as defining tiers: official, expert, partisan, anonymous, etc.) and the AI must follow those ratings when writing or deciding on a story. If the AI suggests a change to those (say it finds a new source and guesses its reliability), a human ultimately approves that classification. This prevents automation from inadvertently elevating a disreputable source or suppressing an important one due to a glitch or bias.

The Editorial Dashboard is a key tool in governance: it provides editors with a real-time window into all AI activities. Editors can see incoming press releases, the AI drafts, the scores, and any flags. Through this dashboard, they can override decisions (e.g., stop an auto-publish in progress), adjust parameters (like raising the importance threshold temporarily if a particularly sensitive event is happening), or push updates. The dashboard is essentially the control panel for the AI newsroom, giving human editors situational awareness and control. This implements the oversight principle in daily operations – nothing is hidden from the editors.

Another governance aspect is the Moderated User Feedback Loop. Newsroom America actively involves its audience in oversight. Reader comments, ratings, and suggested fixes are welcomed, but to keep things constructive, they are moderated and triaged. A moderation system (likely a combination of AI and human community managers) filters out spam or abuse, then forwards substantive feedback to the editorial team or publishes it. For example, if many readers comment that a certain story is missing a key piece of context, editors will see that aggregated feedback and can act on it (perhaps updating the story or writing a follow-up). In some cases, user feedback might even trigger the AI to re-run a prompt – e.g., if users ask a question that the story didn’t answer (“What about X?”), an editor could use that as input for the AI Journalist to generate an additional explainer paragraph. By embracing reader input as part of the process, Newsroom America aims to improve coverage and also to give the community a sense of ownership. This approach recognizes that no newsroom has all the answers, and that readers often have local knowledge or perspectives that can enhance a story.

Furthermore, the platform maintains a Transparency Ledger for each story. This ledger is basically a public audit trail. It might show entries like: “*AI Reporter created draft v1 at 10:00 (source: Press Release XYZ); AI Journalist added context at 10:02 (sources: Archive ABC); AI Editor approved for Auto-Wire at 10:05; Human Editor edited paragraph 3 and moved to Human Wire at 10:20.*” Every story in Newsroom America thus carries a record of who/what touched it and how it evolved. This information may be presented to readers through a UI element (for those curious) or at least be available on request. It operationalises the radical transparency principle and is an important governance tool – if a question arises about a story, one can pinpoint exactly how it was generated and if a mistake was introduced, track it to the source.

Community and Collaboration: Governance is not limited to internal control; Newsroom America positions itself as an open research project and actively invites outside participation. Journalists, technologists, and academics are encouraged to get access to the system, test its outputs, and provide feedback or critiques. The project provides logins to the prototype for those interested in examining the editorial logic and AI behavior firsthand. By opening the doors in this way, the team makes itself accountable to a broad range of experts. It's a form of peer review: AI researchers might inspect how prompts are working; journalists might watch how the AI handles a breaking story and point out weaknesses. This collaborative stance is relatively unique – it treats the development of the AI newsroom as a public interest endeavor rather than a proprietary secret. The reasoning is that the challenges at hand (misinformation, trust in news, etc.) are larger than one startup or one newsroom, so the solutions should be developed in the open. It also helps build trust: people are less likely to be skeptical of an AI system they can poke at themselves.

Additionally, by engaging policymakers and civic technologists in dialogue, Newsroom America's governance model can evolve in step with societal expectations. If there are concerns about bias or legal issues, they can be addressed early with input from observers. This open-door policy shows that the project has nothing to hide and values accountability to the community it serves.

Legal and Ethical Oversight: The governance model also contemplates how to handle legal and ethical issues. For example, the “Arrest headline problem” is a known ethical issue: news outlets often report on arrests but fail to follow up on the outcomes (charges dropped, etc.), unfairly tarnishing reputations. However, it is doubtful arrests and court cases can be covered to a high standard without human involvement. Firstly any press release from a prosecutor or regulator is inherently biased because they are just one party to the case. In contrast, objective and impartial journalists in court report both sides: the prosecution and the defence. A significant challenge is how to get around this “trial by media” problem if you are relying only on releases from agencies such as police, without a counter-balance from defence or timely access to court documents and judges' sentencing notes or pre-trial directions.

Keeping track of cases is equally problematic because the comprehensiveness or otherwise of court records and their public release can vary enormously between jurisdictions. Courts in NZ don't usually issue public updates on the progress of cases (unless specifically requested and it may take days to get a response) because it would require too much time and resources. The District Court alone deals with 200,000 applications a year and publishes only a fraction of the eventual judgments on their website some months later. The High Court tries to give updates on the progress of high-profile cases but by no means all of them, and only as website updates and very rarely as press releases. Only the Supreme Court (because it has so few cases) is more comprehensive in its public statements. The courts in New Zealand prefer to deal with accredited court reporters who they know at the court counters. The Police are often unreliable for all but the most basic information about a court appearance (time, date, court and if you're lucky the charge).

An initial first step would be to always flag releases from prosecuting agencies as just that, and that the charges may be open to change or may be denied or disputed during subsequent court appearances. Newsroom America's will also implement case tracking for such stories,

ensuring that any automated report of an arrest is later updated with the resolution or clearly marked as “allegation, outcome pending.”.

These kinds of policies are part of editorial governance – the human editors must establish and enforce these rules, and encode them into the AI’s operating procedures. Similarly, for election coverage, the policy might mandate that an AI-generated report about a campaign statement must include the candidate’s party and some context about their polling, etc., to avoid inadvertently amplifying propaganda. These are the sorts of guidelines the editorial board would define, and then monitor via the dashboard and audits.

In terms of legal liability, because a human editor is always accountable, Newsroom America treats AI outputs as if they were written by a staff reporter – they must be legally vetted (for libel, etc.) under the same rules. The AI Editor helps by flagging risky language (like something that could be defamatory), but a human’s eyes will make the final call if there’s any doubt. The governance plan likely includes regular review of the AI’s content by legal advisors, especially as the system expands to ingest more sources (to ensure it’s not accidentally republishing copyrighted or private info, for example).

In conclusion, Newsroom America’s governance framework is multi-layered: internal editorial control, systematic transparency, and external community engagement. By codifying ethical principles and building tools (dashboards, ledgers, classification systems) to enforce them, the project ensures AI is always subordinate to human values. This approach mitigates the risk of AI missteps and builds a connection with users, as every aspect of the news creation process is accountable. It’s an evolving model, but one firmly anchored in the idea that journalism’s core mission and ethics must guide the use of AI, not the other way around.

6. Moderated Audience Participation

At the heart of *Newsroom America* is a commitment to participatory journalism—making audiences not just consumers of news, but contributors to the public conversation. The Proof of Concept version allows the audience to leave comments and rate stories. In future, the platform will integrate VoxPop’s innovative speech-to-text tools, allowing users to respond to stories using their voice. The VoxPop system is used by top NPR shows such as *1A* and *On Point* to engage audiences and gather content to great effect. These spoken contributions are transcribed in real-time, enabling a more accessible and expressive form of public commentary than traditional text input alone.

Users can comment on any published story by either typing or speaking their thoughts directly into the platform. All public contributions are subject to moderation, but *Newsroom America* will enhance this process with the AI “Comments Coach” - a real-time assistant that reviews each submission for compliance with the site’s editorial policy, legal standards (such as defamation and privacy law), and tone guidelines.

AI Comments Coach: Example Interaction

Story: Council Approves \$12M Flood Resilience Plan"



The city is wasting money on this initiative. The council members are a joke!



AI Coach Feedback:

⚠️ Your comment cannot be published yet. Here's why:

- Allegation of waste—Please provide rationale or specifics for claims about misconduct or wrongdoing.
- Disparaging remark—Consider a more respectful tone.

✂️ Suggestions: Express the issue or concern constructively; ask a question to further dialogue.



Revised User Submission (Approved for Publication):

Can the council explain how this will help prevent future flooding? I'd like to understand why it's worth the cost.



Status: Published

Your comment is now live and visible to others. Thanks for contributing thoughtfully to the discussion!

Rather than simply rejecting problematic contributions, the *AI Comments Coach* provides transparent feedback to the user, highlighting specific issues (e.g., unsupported claims, aggressive tone, or inappropriate language) and suggesting revisions. Users are invited to revise and resubmit their comment, ensuring that the platform promotes inclusive dialogue while maintaining editorial standards.

This system offers several key benefits:

- **Transparency:** Users understand why their comment may not be publishable and how to fix it.
- **Education:** Contributors gain awareness of journalistic norms and legal boundaries.
- **Safety and Integrity:** The comment space remains civil, constructive, and legally sound.
- **Inclusion:** Voice input broadens participation for users who prefer or require verbal interaction.

By blending audience voice, automated coaching, and human editorial oversight, *Newsroom America* builds a moderated civic commons that supports engagement without sacrificing standards. It turns audience feedback into a two-way, iterative process—and ultimately, into a richer form of democratic discourse.

7. Roadmap and Future Innovations

Newsroom America is continuously evolving. The current platform demonstrates that press releases can be turned into news swiftly and responsibly, but the vision extends much further. The development roadmap anticipates richer AI capabilities, more dynamic story formats, and broader community impact – all while retaining the human-centric governance described above. The roadmap is grouped into several major phases or focus areas:

7.1 - Editorial Intelligence – Building Smarter AI Editors: The next generation of the system focuses on enhancing the AI editorial team’s intelligence and specialization. The plan is to integrate a suite of AI sub-editors (or “micro-services”) that each tackle a specific editorial task on every story. These include:

- *Fact-Checking Agent:* An AI that cross-verifies claims in a story against reliable databases and archives. For example, if a press release says “unemployment is at 5%,” the fact-checker bot will compare that against official labor statistics. If a story mentions an event, it checks whether that event actually occurred on the stated date. This agent flags any inconsistencies or unverified assertions for further review.
- *Framing & Bias Detector:* A “framing integrity” prompt that reviews the language of the article for neutrality. It looks at the headline and lead especially – are they worded objectively? It can detect emotionally loaded words or potential bias in phrasing. For instance, if the AI Reporter wrote “government finally scraps failed program,” this sub-editor might flag “finally” and “failed” as loaded terms not appropriate in a straight news context. It ensures that the story’s framing doesn’t inadvertently push a narrative.
- *Attribution Auditor:* This agent’s job is to scan the story to ensure every claim is attributed. It would catch, for example, if a sentence states a controversial point without saying who said it. It also identifies if any content might be quoted or paraphrased from the source material or elsewhere and checks that quotation marks and references are properly used. Essentially, it won’t let the AI “state” something that isn’t common knowledge or directly observable without citing a source. It also helps label anything that is opinion or conjecture as such.
- *Impact Assessor:* A sort of news judgment AI, which evaluates how important a story might be to the public. It looks at signals such as: is the subject a high-ranking official or significant institution? Is the topic something of broad interest (economy, health, safety)? It might also incorporate early user engagement or geographic reach (a story affecting a big city vs a remote area). Based on these, it suggests how prominently to display the story (front page headline vs a short brief). It essentially quantifies news value to aid editors in prioritisation.

These specialized sub-editors will run in parallel on each story, and their findings will be compiled into a story scorecard for the human editors (and AI Editor) to review. For example, after the AI Journalist writes a draft, the fact-checker might add a note “Claim about budget not verified”, the bias detector might say “tone is neutral”, attribution auditor

says “all good except paragraph 4 needs source”, impact assessor says “importance 75/100, national relevance”. The AI Editor and/or human editor can see this scorecard at a glance and know exactly what to fix or whether it’s publish-ready. This approach amplifies the concept of AI-assisted editing: rather than one monolithic AI trying to do everything, you have a team of narrow AI tools each excellent at one dimension of editing. This phase is effectively in line with what we saw in the GEC story – but in the future, those steps will be more automated and robust (with actual fact-checking against databases, etc., not just inserting context).

By formalising these roles, Newsroom America aims to make the AI output even more reliable and rich. It also provides an auditable trail (“the fact-check bot gave it 8/10 on accuracy” etc.), which strengthens transparency. In the next year, the team’s roadmap includes developing these prompts/agents and training them on real newsroom scenarios. The outcome should be that AI-generated stories are not only faster but demonstrably benchmarked for quality on multiple fronts (facts, tone, sourcing, importance). This phase truly gives journalists “superpowers” – even a small editorial team can have the equivalent of a whole backbench of sub-editors reviewing every line, courtesy of AI.

7.2 - Story Evolution – Real-Time, Living News Stories: Traditional digital news is often static – an article is published at a timestamp, then maybe a separate follow-up comes later. Newsroom America plans to shift towards dynamic story flows that update in real time as events unfold. In this model, a single story page can evolve throughout the day (or week or month) with new information appended or integrated, rather than spawning separate articles for each development. The roadmap includes building features such as:

- *AI Event Linkage:* The system will use AI to detect when multiple updates are part of the same story thread. For instance, if at 9am the government issues a press release, and at 11am an advocacy group responds, and at 3pm new data related to the issue comes out – rather than three separate stories, the AI will recognize them as linked by topic or entities (e.g., all about the “Education Funding Bill”) and group them. It essentially tags incoming news items with an event or story ID.
- *Automated In-Place Updates:* With events linked, the platform can update an existing news story with new paragraphs or sections as new info arrives or is inputted. Each update would be time-stamped within the story. Readers viewing the story could see something like a timeline or just a continuously edited article where sections say “Update 11:00 – [new info]”. Importantly, previous versions are archived, so there’s a record of what changed (maintaining transparency that things aren’t being surreptitiously altered). This mimics live news reporting (“filing to the live feed”) but with AI assistance. The AI might write the initial story and then later also write the update snippet and append it.
- *Narrative Tracking and Thematic Tags:* The system will tag stories with overarching themes or storylines (e.g., “Cyclone Gabrielle recovery” or “2025 Election”). This allows the newsroom to group related content and helps readers follow long-running issues. On the site, a reader could click a tag to see the full context or background of that narrative. For the AI, these tags help it decide what context to bring in; if it knows a press release falls under “Climate Policy Debate”, it can retrieve the recent

narrative context for inclusion.

- *Editorial Alerts for Updates:* When the AI is updating a story that a human editor previously reviewed, it will send a notification to the duty editor. For example, if an editor approved the 9am story, and at 11am the AI wants to append a response quote, the editor gets an alert: “Story X has a new development – review?” This ensures that human editors stay in the loop on evolving stories and can step in if the update is sensitive or if they want to rewrite the combined story for clarity.

The effect of these features is to treat news as a living, breathing thing that grows over time, rather than a set of discrete articles. It aligns with how information actually emerges and how audiences follow stories (think of live blogs or Twitter threads, but with editorial structure). For readers, it means a one-stop experience: they can bookmark a story and see it updated rather than hunting for separate pieces. For the newsroom, it means no more publishing five variants of essentially the same story – instead one evolving story that remains coherent and cumulative.

This is especially beneficial for resource-strapped newsrooms: rather than spending time writing marginally different stories, the AI can maintain the timeline and humans can jump in to do deeper analyses or corrections. It also addresses the problem of follow-up (such as the arrest story outcome issue: the story stays active until resolution, and the AI can be tasked to watch for that resolution and update when it happens (say, “*Update: six months later, this person was acquitted*”). So this Story Evolution phase has a strong public service angle: ensuring completeness of reporting over time.

Technologically, implementing this will involve the platform handling versioning, merging content, and UI to show updates. The AI will need to recognise context to avoid repeating information redundantly. The roadmap likely schedules prototyping this with a subset of stories (for example, try it with ongoing coverage of a trial or a natural disaster with multiple official updates).

7.3 - Platform Governance & Scale – Expanding While Upholding Values: As Newsroom America grows, it will handle thousands of stories per day (especially if users adopt personal AI reporters, discussed shortly). Scaling up requires robust infrastructure and governance to avoid mission drift. The roadmap’s third focus is on the governance tools and processes for large-scale operations:

- *Advanced Editorial Dashboard:* The editors’ control panel will be further developed to manage high volume. This might include better filtering (so editors can quickly see the most critical stories out of hundreds), bulk actions (e.g., “approve all low-risk community notices”), and analytics that highlight anomalies (like an unusual spike of flags on a certain story). The aim is that even with 1000 stories/day, editors can effectively supervise the AI through smart interfaces.
- *Community Moderation at Scale:* With more content and more user interaction, moderation tools must scale. This could involve AI moderation assistants that help filter user comments, and community guidelines enforcement automated to some degree. Ensuring civility and constructiveness in user contributions will be important

as the audience grows (since toxic comments could derail the collaborative spirit). The roadmap may explore partnerships or using existing content-moderation AI to handle this aspect, under human community manager oversight.

- *Transparency Ledger & Analytics:* At scale, the Transparency Ledger becomes even more crucial. It could be turned into a public dashboard showing stats like “X% of today’s stories were AI auto-published vs Y% human-edited” and listing any major corrections or retractions. Essentially, a public accountability dashboard for the whole platform performance might be an outcome. This keeps the project honest and allows external observers to flag any patterns (for instance, if auto-published stories in one category have frequent corrections, that might indicate a need for more human oversight in that category).
- *Maintaining Mission Alignment:* A key challenge with scaling is avoiding the temptations of click-driven or profit-driven changes. The founders are wary that as any media venture grows, pressures (financial, competitive) can push it toward sensationalism or quantity-over-quality. The roadmap explicitly calls out that *commercialisation cannot come at the expense of editorial independence or integrity*. Practically, this means the governance structure might include non-profit or public-interest elements to firewall the mission. For example, an advisory board of public editors or a partnership with a university could be instituted to keep focus on civic goals. The team is especially conscious of serving underserved communities and hyper-local journalism needs, not just chasing a broad national audience. As we scale, we intend to prioritise deployments in local/regional contexts that lack news, demonstrating impact there rather than just seeking the largest audience. This philosophy is part of governance: any scale strategy should be *ethically aligned and community-accountable*. We might see, for instance, a requirement that any major feature addition or expansion is evaluated by whether it serves the public (per the core commitments) and if not, it’s rejected even if it could drive engagement.

The upshot of this phase is ensuring that growth does not dilute the core principles. The technology will be expanded to handle volume, but simultaneously measures like regular ethics audits, community feedback channels, and editorial oversight will be fortified. The result should be a platform that can grow from a prototype to a globally, nationally and locally significant news source without losing trust. The roadmap likely envisions partnerships with academic institutions, existing media or civic organizations to spread this model. Because Newsroom America is more a framework than a single site – the concepts could be adopted by local news orgs, etc., scaling might also mean federating the model outwards (lots of small Newsroom America-powered sites) rather than one central site eating the world.

7.4 - Case Study: The Westport News

The Westport News, a 150-year-old local newspaper, partnered with Kinga VoxPop Ltd under the Google News Initiative to create an affordable model for small newspapers transitioning from print to digital. The project addresses the financial and operational challenges many local outlets face as they navigate rising print costs, aging audiences, and advertiser migration to digital platforms.

Key highlights:

- **Crisis for Local News:** Most small newspapers have no digital presence and cannot afford to transition without jeopardizing their revenue.
- **Trust in Local News:** Local outlets remain highly trusted, largely due to their relevance, accuracy, and community roots.
- **Digital Tipping Point:** In 2023, digital media consumption surpassed traditional media in New Zealand, even among older demographics—making now an ideal time for transition.
- **Digital Revenue Opportunity:** People are willing to pay for actionable, relevant, and convenient local news, especially in digital formats.
- **News Avoidance Trend:** Audiences are increasingly overwhelmed by repetitive news, contributing to disengagement.
- **VoxPop’s Solution:** A custom-built **Digital News API** enables small newspapers to send proofread stories and photos directly via email for automatic digital publication—with no new staff or costs.

One of the first applications for the fully proven NewsRoom America model will be to integrate an automated 24/7 Westport/West Coast local news service into the apps. This will pull news from the most reliable sources such as Police, councils and the fire service and create a breaking news wire for the apps. A particular focus will be on providing real-time information during natural disasters, when the small team of journalists can be overwhelmed by the heightened reporting demands.

7.5 - AI NewsHound – Personalised AI Journalists:



Another exciting part of the future roadmap is product innovation to empower end-users directly. The concept of “*Create Your Own AI Journalist*”, branded as *AI NEWSHOUND*, is one such initiative. The idea is to let individual users or communities configure their own AI-driven news agent that caters to their specific information needs. Here’s how it is envisioned:

- Users would have a simple interface to set up their AI journalist. They can choose topics they care about (e.g., local council decisions, climate news, school updates) and geographic scope (just their town, or nationwide). Essentially, they define a personalized “news beat.”
- They can also set keywords or triggers – for example, a user could tell the system “notify me of any news about ‘Ruataniwha Dam’ or any major weather alerts in my county”. This is like setting up custom watchlists for issues or names that matter to them.
- The AI NewsHound then continuously monitors the relevant feeds (official press releases, open data, emergency alerts, etc.) within those parameters. It uses the same backend as Newsroom America – the difference is that it filters and prioritises specifically for that user. It might even pull from the Press Release Wire and Auto-Wire of Newsroom America, but just the subset that matches the user’s criteria.
- Users can choose how they want to receive information. Options might include a real-time alert (push notification or SMS the moment something happens) for urgent things, a daily email summary for more routine updates, or even an audio briefing if they prefer. The system is flexible: one user might basically create a personal news ticker for city council news, another might set a 5pm daily digest of all local crime and fire reports.
- Importantly, the AI journalist persona can be customised – users can name it and pick a style. For example, someone might call theirs “Hawke’s Bay Watchdog” and choose a tone like “just give me the facts brief” versus a more explanatory style. This personalization makes the AI feel like *“your own reporter working for you.”* The content it produces would still follow the central editorial guidelines (no misinformation, verified sources only), but it can be tailored in presentation.
- All information that the AI NewsHound provides is drawn from the verified sources and processes of Newsroom America. That means the user isn’t just scraping the web; they’re getting a curated feed that’s been through the AI newsroom pipeline. No clickbait, no influencer hot takes – just the relevant factual updates. This is a key differentiator from say, a Google Alert or a social media feed.
- Users could even allow their personal AI journalist to publish certain stories publicly. For instance, if a community leader sets up an AI to monitor local earthquake data and it generates a useful report, they could share it on the platform for others to see (subject to editorial review if needed). In effect, users become collaborators in news dissemination, with AI as the enabling tool.

The target audience for AI NewsHound ranges from businesses to civically engaged individuals and watchdog citizens to local journalists who could use it as a research assistant, and even emergency responders who need real-time situational updates. For example, a local journalist might “create” an AI colleague focusing on education board updates while they

themselves focus on writing human-interest stories about the schools – thus covering more ground together.

This concept flips the top-down model of news on its head: instead of just consuming what media outlets decide is news, people can actively “commission” the news they care about through their personal AI reporters. It addresses the problem of one-size-fits-all news and the gap in hyper-local reporting. If your town doesn’t have a full-time reporter, you could still get a feed of every notable public report about your town, organised and summarised.

From a roadmap perspective, AI NewsHound is in concept stage. Next steps include building the user interface for onboarding (choose topics, etc.), and integrating the alerting mechanisms. A pilot is suggested to test real-world usage. This would allow refinement of how well the AI can handle many individualised requests and how to make the experience simple for non-techy users.

If successful, AI NewsHound could represent a scalable, decentralised extension of *Newsroom America’s* mission. Each user basically becomes their own editor-in-chief for the issues they care about, with the assurance that the information is coming through a reliable pipeline. It’s a compelling vision of empowering the public with AI – not to create echo chambers, but to ensure everyone can stay informed about the things that matter in their community, even if traditional news has abandoned that niche. And because all these personal AIs operate under the same ethical framework (they’re just differently filtered views of the core system), the values of accuracy and context carry through to the micro level.

7.6 - Case Study - AI in Civil Emergencies

In a disaster, timely, accurate information can save lives. Yet traditional newsrooms, even well-prepared ones, can be overwhelmed or rendered inoperable during a major emergency. Following Cyclone Gabrielle in Hawke’s Bay in 2023, Kinga VoxPop Ltd developed an Emergency Broadcast System (EBS) in partnership with Radio Hawke’s Bay. It is designed to ensure that vital information continues to flow to a hyper-local level and messages can quickly be sent and broadcast on AM/FM radio. However finding volunteers or funding journalists to fully implement this project has been a challenge.

Even with staff and volunteers, in large-scale disasters such as floods, earthquakes, or cyclones, they may be cut off, studios may be inaccessible, and standard workflows break down. At the same time, reliable information *does* exist - it flows through email from verified sources such as:

- New Zealand Police
- Fire and Emergency NZ
- Civil Defence
- Local councils
- Power and infrastructure providers
- Health and emergency services

The challenge is getting this information processed and broadcast quickly, without relying on humans who might not be available.

The Solution: Plugging in the *VoxPop AI News Engine*

To solve this, we would connect the *VoxPop AI News Engine* directly to the Radio Hawke's Bay media output infrastructure - apps/social media accounts/web site and AM/FM broadcasts. All press releases and alerts from reliable Hawke's Bay sources will be automatically collected via email and ingested into the AI system, which can then process them into real-time emergency news reports.

Here's how the future system could work:

1. **AI Reporter** receives an incoming alert and instantly drafts a concise, factual newsflash.
2. **AI Broadcast Editor** adapts the story into spoken-word format—rewriting it for the ear rather than the eye.
3. **AI Newsreader** records the bulletin using high-quality neural speech.
4. **RoboJock** introduces the newsflash and broadcasts the story live on Radio Hawke's Bay's AM/FM channels or schedules it for immediate replay

VIDEO: Here's a 2 minute video demonstrating how AI radio news stories are created: <https://youtu.be/DyxnflSzpLg>

If station staff are unavailable or offline, the system could switch into *autonomous emergency mode*, capable of operating without human intervention. In just minutes, critical information can go from inbox to airwaves.

Multi-Platform Emergency Distribution

In addition to AM/FM broadcast, alerts will also be published to:

- The Radio Hawke's Bay mobile app
- The station's website
- Social channels (optional)
- Civil Defence or council push-alert systems (where integrated)
- Other media/interested parties

This creates a redundant, always-on public alert system tailored for small radio stations with big responsibilities.

A National Model in Development

This system is currently in planning for Hawke's Bay, but the model is designed for national and global scalability. Once the AI pipeline is tested and tuned, it can be replicated for other independent media and community media and broadcasters, forming a distributed, AI-assisted emergency broadcast network.

By plugging in to the *VoxPop AI News Engine*, even the smallest stations will be able to become 24/7 emergency news hubs without needing full-time staff on site.

Overall, the future roadmap of Newsroom America is about deepening the “human-AI symbiosis” in journalism. The system will get smarter in editorial tasks, the stories will get more dynamic and continuous, and the product will get more personal and participatory – all without compromising on transparency or trust. It’s an ambitious path that, if realised, could transform not just one outlet, but serve as a template for newsrooms globally. The team explicitly views *Newsroom America* as a model that can be adopted, adapted, and improved by others – a sort of open-source blueprint for the future of news. In that sense, the roadmap is not just a to-do list for this team, but an invitation to the wider media community to join in building a better news ecosystem, one that is fast and faithful to facts.

8. Monetisation

NewsRoom America is already incurring direct costs for processing press releases into stories. Currently it is about US\$10 for every 3,000 press releases processed. As the service ramps up and personalisation is introduced, these costs will rise significantly. Therefore the service will need funding for both the machine and the human journalists who oversee it.

A multi-stream revenue model can be built on its core capabilities in AI-driven journalism, structured data processing, and automation. Its monetisation approach would capitalise on both raw public data and AI-generated editorial content, offering distinct products and services across several high-demand markets.

8.1 - Raw Data & AI Training Feeds

NewsRoom America offers continuous, AI-ready structured data feeds sourced from public domain content, government databases, and corporate PR. These are tailored for use by AI labs, LLM developers, and domain-specific training initiatives in fields such as finance and law.

8.2. - AI-Generated News Subscriptions

The platform delivers automated news reports, industry-specific briefings, and real-time alerts. These are available as subscriptions for individual users and enterprise clients, catering to financial institutions, law firms, analysts, and corporate intelligence units.

8.3 - AI-Powered Fact-Checking

NewsRoom America provides an API-accessible AI tool for verifying news and combating misinformation. This product is aimed at government regulators, media monitoring groups, and digital platforms requiring scalable fact-checking solutions.

8.4 - AI News Anchors & Voice Assistants

The service includes automated voice and video content—such as news summaries for YouTube, TikTok, or voice interfaces—targeted at broadcasters, content creators, and podcasters seeking low-cost, high-frequency output.

8.5 - Market Intelligence Reports

Through its AI systems, the platform can generate custom economic, policy, and financial intelligence reports. These are particularly useful to hedge funds, corporate strategists, and political risk analysts seeking near-instant insights.

8.6 - AI-Powered Local News

By automating coverage of crime, weather, and municipal affairs, *NewsRoom America* supports small newspapers, community radio, and hyperlocal digital publishers. This enables local outlets to produce affordable, timely content without additional staff.

8.7 - AI-Generated Social Media Content



Branded news posters are visually formatted summaries of breaking news stories that can be automatically created and published to social media from editorially approved stories.

To meet the demand for high-volume digital publishing, the service auto-generates news posts formatted for platforms such as Twitter, LinkedIn, and TikTok, supporting publishers, marketers, and influencers.

9. Conclusion

Newsroom America was built as a provocation and a promise: a provocation that in an era of shrinking newsrooms and rampant misinformation, we should not accept the decline of journalism; and a promise that emerging technologies, if guided by the right values, can help revitalise journalism's role in democracy. This comprehensive initiative demonstrates that artificial intelligence can be harnessed to serve journalistic integrity rather than subvert it. By weaving together automation with rigorous editorial oversight, *Newsroom America* delivers news that is both fast and faithful. It shows that we can preserve the soul of journalism – facts, verification, accountability – even as we embrace new tools to extend its reach.

The journey so far has yielded a functioning prototype where AI accelerates reporting on reliable sources, and human editors ensure every story meets high standards. We've seen how an AI can turn a press release into a breaking news story in minutes, and then how additional AI layers and human judgment will turn that brief into a contextualised, balanced article. The integration of features such as the three-wire system and the public correction log means transparency is built into every step, inviting the public to trust but also verify. In practical terms, a reader on *Newsroom America* can know exactly where a story came from, how it was produced, and what has been done to ensure its accuracy – a level of insight virtually unheard of in conventional news delivery.

Crucially, *Newsroom America* confronts the challenges head-on: it acknowledges when AI might fall short (such as understanding nuance or following up on outcomes) and has designed solutions – whether it's keeping a human in the loop, adding a second AI pass for context, or planning a case-tracking system for legal stories. This reflective approach exemplifies a key strength of the project: it is not techno-utopian or dystopian, but pragmatic. It identifies specific problems in modern journalism (speed vs. accuracy, volume vs. quality, lost trust) and iteratively tests how a hybrid human-AI system can address them. Some solutions are already in place, others are on the roadmap, but all are guided by the ethos that editorial principles must govern technology and not be overridden by it.

Newsroom America's mission to rebuild trust is also a community mission. By involving journalists, academics, civic technologists, and readers in the process (through open access, feedback loops, and the promise of tools like AI NewsHound), it recognizes that trust is a two-way street. The audience is not a passive recipient but an active participant in shaping the news. In many ways, this project revives the idea of the news “as a conversation” – but now that conversation includes AI assistants contributing where appropriate, under human moderation. The result is a more responsive, inclusive form of journalism that can adapt to the information needs of each community and individual.

There is still much to do. The white paper has outlined forthcoming improvements: smarter fact-checking bots, real-time story updating, expanded governance mechanisms, and personal news reporters. These innovations will require careful implementation and continued vigilance to ensure that scale and automation don't introduce new problems. The *Newsroom America* team is candid that it's “early days” and they invite critical engagement

and collaboration from others. That open ethos is a strength; it means the project can learn and course-correct with input from the wider world.

In a time when many lament that “*journalism is in retreat*” and fear that “*algorithms have overtaken information*”, Newsroom America stands as a hopeful counterexample. It suggests that with thoughtful design, AI can be a tool to empower journalists – to give them “superpowers” of speed and breadth – instead of a threat. It also offers a potential lifeline to struggling local news: a model where a small team augmented by AI can cover a broad swath of civic news efficiently, possibly turning the tide on news deserts. If replicated, this could mean more communities in “News Deserts” staying informed and engaged.

Ultimately, the measure of *Newsroom America*’s success will be in how it impacts the public. Does it help citizens feel more informed? Does it hold the powerful accountable by surfacing facts quickly and contextually? Does it correct mistakes swiftly and visibly, thereby strengthening public confidence that evidence-based, factual and reliable reporting is the goal? Early indications are positive, but the true test will be broad adoption and continued performance. The invitation is open for editors, policymakers, technologists, and everyday news consumers to join in testing and refining this model.

Newsroom America proposes that the future of news need not be left to profit-driven algorithms or abandoned to the ashes of collapsing business models. Instead, it can be actively rebuilt, combining the best of human editorial wisdom with the best of AI’s capabilities. By doing so, we can aim to restore the speed, accessibility, and relevance of news without sacrificing the depth, ethics, and reliability that define journalism at its best. This white paper has synthesised how that vision is being pursued through concrete systems and principles.

As we move forward, we invite you to join us in shaping this future. We believe Newsroom America is not just one experiment, but the beginning of a wider movement to ensure that in the digital age, journalism thrives. We welcome collaboration from editors, policymakers, technologists, and news consumers. Contact us to explore partnership opportunities, contribute to development, or provide feedback on how we can collectively build a more informed and connected world.

About the Authors

Peter Fowler

Peter Fowler is a pioneering New Zealand journalist, editor, and media innovator with more than three decades of experience in broadcast and digital journalism. He founded newsroom.co.nz in 1996 and sold it to NZX in 2007, is co-founder of Kingā VoxPop Ltd and creator of *Newsroom America*, one of the first AI-native newsrooms in the world. A former Radio New Zealand journalist and presenter, Fowler has won multiple national awards, including Best Radio Journalist at the New Zealand Radio Awards, and was a finalist for Best Investigative Documentary at the AIB Media Awards. Throughout his career, he has led the integration of technology into public service journalism.

Andrew McMillan

Andrew McMillan is a technical leader and software architect with a deep commitment to civic infrastructure and information accessibility. As co-founder of Kingā VoxPop Ltd, he leads the development of systems that bridge AI and journalism. Andrew has a long track record of building scalable, reliable cloud-based systems, and has worked with government, academic, and media institutions to deploy reliable, ethical technology at scale. His work on the *Newsroom America* platform includes the real-time editorial pipeline, AI reliability scoring system, and public transparency architecture. McMillan has worked in senior positions most recently at Toast, previously at Google and One Laptop per Child and was a founding director of the Wellington software development company Catalyst IT.

Appendix A: Glossary of Terms

AI Reporter

An automated system (e.g. *Seymour Krelborn*) that converts press releases into concise, structured news stories. It produces factual summaries but does not add background or context.

AI Journalist

A second-layer AI designed to enrich news drafts with context, historical background, cross-references, and counterpoints. It simulates a beat reporter who answers the “why” and “what does it mean” questions.

AI Editor / AI Sub-Editor

AI responsible for reviewing drafts against editorial policy, ensuring neutrality, tone, grammar, and attribution. Assigns a publishing status:

✓ **Auto-Publish** | △ **Needs Review** | ✗ **Do Not Publish**

Three-Wire Publishing Model

A transparency framework separating stories into:

- **Press Release Wire** – raw, unedited source materials
 - **AI Auto-Publish Wire** – machine-generated, no human review
 - **Human-Reviewed Wire** – stories reviewed and approved by editors
-

Editorial Policy Compliance

Verification that each story meets journalistic standards, including accuracy, fairness, and ethical considerations (e.g. defamation, privacy).

Transparency Ledger

A public audit trail showing each step in a story’s creation—source, AI involvement, edits, approvals—ensuring accountability and traceability.

Story Score / AI Journalist Scoring Rubric

A 100-point evaluation applied to AI-generated stories, based on:

- Factual Accuracy (30 pts)
 - Attribution & Sourcing (10 pts)
 - Contextual Integrity (10 pts)
 - Headline Accuracy (10 pts)
 - Correction Speed & Transparency (20 pts)
 - Editorial Policy Compliance (20 pts)
-

First Mile of Journalism

The earliest stage of news reporting when raw information is first captured. It's closest to the source but also most vulnerable to distortion if not verified.

Radical Transparency

A core value: full public visibility into how stories are produced, including source links, AI authorship, and editorial decisions.

Source Reliability Score

A dynamic rating assigned to each information source, used to determine whether content can be auto-published or requires human review.

Correction Audit / Real-Time Correction Log

A structured record of all corrections, including what was changed, why, and how quickly. Logged internally and, when appropriate, publicly.

Moderated User Feedback Loop

A system for readers to flag issues, suggest corrections, or comment constructively—moderated by both AI and human editors to ensure quality input.

AI Comments Coach

An AI moderation assistant that checks user comments for tone, factuality, and legal risk. Offers suggestions before publication to improve civility and clarity.

Newsroom Dashboard

An internal control panel for human editors to oversee AI activity, review story drafts, monitor flags, and approve publication.

Story Evolution / Living Story

A dynamic story format that updates in real time as events unfold, with time-stamped additions and version tracking instead of separate articles.

AI NewsHound

A personalized AI journalist that monitors topics or regions defined by the user, delivering alerts, summaries, or audio updates using verified content.

Source Classification Model

A framework for categorizing sources (e.g., official, expert, partisan, anonymous) to inform AI behavior and editorial decisions.