africanwomeninmedia

africanwomeninmedia.com

Guidelines for the Ethical and Responsible Use of AI in the African Newsrooms

Research has been funded by

Luminate

Table of Content

Preamble

This document is not a set of rules imposed from afar. It is a collective statement, born from the lived experiences, challenges, and wisdom of women journalists across Africa who are already navigating the rise of Artificial Intelligence. It is a practical guide forged in newsrooms and shaped by a Participatory Action Research process where journalists were not merely subjects but co-creators of this knowledge. We speak from an African feminist and decolonial standpoint and consciously reject one-size-fits-all, Western-centric models of AI governance that often erase our realities.[1] Instead, we centre our own expertise, our contexts, and our right to shape the technologies that enter our workplaces. Our goal is to provide a rights-respecting, culturally grounded, and actionable framework that African newsrooms can adopt, adapt, and own. Where global frameworks often focus on technical fixes, this guide foregrounds what we know to be essential: the value of human labour, the ethics of care, the strength of community, and the imperative to challenge power dynamics in our newsrooms.

Introduction

AI is already here, reshaping our work, from transcription and research to generating images and drafting scripts. Yet its adoption is often informal, unregulated, and driven by individual journalists striving for efficiency. In this vacuum, and as research by African Women in Media has found, women journalists, who are so often the guardians of ethical reporting and community trust, are bearing a disproportionate burden: the unseen corrective labour of fixing AI's biases, the fear of being replaced, and the new threats of AI-facilitated harassment. This guideline is our response. It is a tool for agency. It translates our shared testimony into principles and protocols to ensure that AI augments rather than replaces our journalistic judgment, integrity, and storytelling. It is designed to help us harness this technology responsibly, safeguarding the core values of our profession while actively building safer, more equitable, and innovative newsrooms for all.

How this Guideline was Created

This guideline was developed from primary qualitative research, including focus group discussions and surveys with women journalists across Africa. The findings were triangulated with a review of existing literature on AI, gender, and media. It translates this combined empirical and desk research into practical policy and operational tools, while drawing on feminist and decolonial epistemologies that value situated knowledge and community accountability.

Who this Guideline is For

While this guideline and toolkit are useful for the entire media ecosystem, including policymakers, media regulators, donors supporting

media development, and civil society organisations committed to media freedom and safety, they are designed primarily to help newsroom leaders such as editors, managing editors, and station managers, as well as reporters and content creators, gender desks, and training and HR teams navigating the use of Artificial Intelligence.

‘I totally believe in when someone watches any content that I created or reads a story that I’ve written, they want to identify themselves in that story or within that image so prior to for me would be shaping the look, the persona of an African so that when they see it in terms of creating content... Africanizing AI would be a priority for me.’

Journalist from Zimbabwe

Scope and Limitations

Governing the AI-Assited Workflow

This document offers practical guidance for integrating AI into reporting, editing, distribution, and audience engagement. It provides adaptable policies, checklists, training outlines, and incident response protocols. AI technology evolves rapidly, and the tools mentioned are illustrative rather than exhaustive. We recognise that adoption levels and technical capacity vary widely across the continent, and many tools are used without being labelled AI. This is a living document intended as a starting point for your newsroom's own journey, to be critiqued, adapted, and owned locally.

These guidelines are designed to govern the entire lifecycle of AI-assisted journalism, not merely its final published output. This comprehensive scope is essential because the ethical risks, gendered labour burdens, and editorial failures identified in our research occur at every stage of the process, from the initial command to the final publication.

The Input (The Prompt)

The instructions given to an AI system fundamentally shape its results. Biased, vague, or context-poor prompts will generate biased, vague, and irrelevant content, placing a heavier corrective burden on the journalist.

Scope and Limitations

Governing the AI-Assited Workflow

Our Intervention: We mandate the use of ethically vetted prompts and provide training in prompt engineering to ensure that the first interaction with AI is designed to produce equitable, accurate, and context-aware results.

The Process (The Corrective Labour)

This is the often-invisible, human-intensive work of reviewing, editing, fact-checking, de-biasing, and contextualizing raw AI output. This labour is non-negotiable and is disproportionately performed by women journalists.

Our Intervention: We establish the "Human-in-the-Loop" principle as a mandatory standard and provide the "Human Touch" pre-publication checklist to structure this review. Crucially, we introduce the "Corrective Labour" framework to make this work visible, quantified, and valued within workload assessments and performance metrics.

The Output (The Final Product)

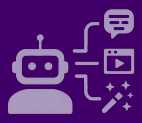
This is the AI-assisted content that reaches the public and carries the newsroom's credibility. It is the result of the input and process stages.

Scope and Limitations

Governing the AI-Assited Workflow

Our Intervention: We enforce strict accountability, stating that a human journalist and editor are ultimately responsible for the output. We also provide clear protocols for transparency and disclosure, ensuring the audience is appropriately informed about the use of AI.[2]

Key Definitions



Artifical Intelligence:

Systems that perform tasks that would normally require human intelligence; in newsrooms, this includes speech-to-text, summarisation, recommendation engines, to generative models.

AI-Generated Output

Text, image, audio, video, or code created by an AI system that is intended for use in a journalistic product, whether as a draft, an asset, or the final published work. This includes AI-written article drafts, AI-created illustrations, AI-synthesised voice-overs, and AI-generated social media posts.

Human-in-the-loop

Workflows where humans retain final editorial authority and oversight over AI-generated outputs.

Invisible Labour

Unpaid, unacknowledged effort in correcting or refining AI outputs (e.g., fact-checking, editing AI-generated outputs, and cultural correction).

Responsible AI

The set of policies, technical measures, and context-relevant cultural practices, designed with gender sensitivity, that ensure AI use aligns with journalistic ethics and human rights principles.

Key Definitions

By governing this complete workflow, our guidelines ensure that ethical journalism is maintained not by chance, but by a structured, accountable, and equitable system.

Part Two

GUIDELINES

Section 1

Core Principles

Human-in-the-Loop and Editorial Control

AI is an assistant, not an author. All AI-generated content must undergo rigorous human review, fact-checking, and editing before publication. A journalist's byline signifies their full responsibility for the content. The final accountability for accuracy, nuance, and ethical standing rests solely with the human journalist and editor.[3]

Cultural and Contextual Accuracy

AI-generated outputs must be scrutinised for Western biases and misrepresentations. Journalists are responsible for ensuring content accurately reflects local contexts, languages, and visuals. This includes a Mandatory Ground Truthing Principle: for any story involving specific communities, geography, or local culture, journalists must verify AI-generated insights with on-the-ground sources.[4]

Transparency and Accountability

Newsrooms should develop a clear policy on how to disclose the use of AI in the content creation process to the audience. Internally, the use of AI for specific tasks must be documented.

Prohibition on Fully AI-Authored Content

AI shall not be used to write complete articles, especially on sensitive topics like health, human rights, or complex climate reporting, where a human touch and nuanced understanding are critical. This extends to a prohibition on sensitive, data-low topics; for critical issues known to have AI data gaps, such as Female Genital Mutilation (FGM) and other localised forms of Gender-Based Violence (GBV), AI should not be used as a primary research tool, as models are often trained without accounting for contextual differences.[5]

Check Suggestion: “I might draft an article myself, then ask ChatGPT to “clean it up”... but I refuse to let it write from scratch.

Journalist from Djibouti

I need every word to reflect my voice.’

Journalist from Zimbabwe

Check Suggestion: “AI wasn’t built with our voices or narratives... When we use these tools, we get outputs that don’t reflect our realities.”

Journalist from Nigeria

Check Suggestion: “All other AI tools that you are using, it has data hallucination that's why we say AI is a collaborative tool that we use in journalism field.”

Journalist from Nigeria

Section 2

Mitigating Bias and Promoting Inclusion

Proactive Bias Mitigation

Journalists must be trained to use explicit, non-gendered prompts to counter AI's default stereotypes and must actively scrutinise outputs for demographic, gender, and ideological biases.

Gender and Cultural Sensitivity by Design

Actively involve diverse, gender-balanced teams in designing and selecting AI tools, avatars, and voices. Prioritise the development and use of avatars that reflect the diversity of the audience, including body types, hairstyles, accents, and the ability to pronounce local names correctly.

Combating Homogeneity

Use AI to identify and explore underrepresented perspectives. Editors must actively guard against the "repetition of ideas" and ensure AI is used to diversify angles, not homogenise output.

Section 3

Safety Accountability, and Career Advancement

Integrate AI into Safety Protocols

Update safety and harassment policies to address new risks like targeted deepfakes, AI-augmented online harassment, and institutional blame for AI errors. Establish clear, confidential processes for journalists to report these issues without fear of reprisal.

Protect Against Gendered Harassment

Newsrooms must establish support protocols for journalists, particularly women, who are targeted by AI-facilitated abuse (e.g., deepfakes, cloned voices). This includes legal, technical, and psychological support.[6]

Incident Response Protocol for AI-Facilitated Harassment

This protocol must be activated immediately upon discovery of a targeted AI-facilitated attack (e.g., deepfake, cloned voice used for scams/slander).

Immediate Response (24-Hour Timeline)

Designate a Contact person

The AI Ethics Committee chair or assigned editor takes charge.

Obtain Evidence

Take screenshots, save links, and download copies of the malicious content.

Takedown Requests

The Point Person immediately issues formal takedown requests to the hosting platforms (social media, websites), citing copyright infringement and defamation.

Internal Communication

Send a clear, calm internal memo to all staff, acknowledging the incident, affirming support for the targeted journalist, and directing all external inquiries to the designated Point Person.

Support (48-Hours Timeline)

Activate Support

Immediately connect the targeted journalist with pre-vetted resources.

Legal Counsel

To assess options for cease-and-desist letters or further legal action.

Mental Health Support

Provide access to a counselor specializing in online trauma.

Digital Security Expert

To help secure the journalist's online accounts and personal data.

Public Statement (Case Basis)

Where the incident is public, issue a short, strong statement.

Template

We are aware of a malicious AI-generated deepfake/audio clone targeting our journalist, [Journalist's Name with their consent]. This content is fabricated. We stand firmly with [Journalist's Name] and are taking all available legal and technical actions to have it removed. This is a serious attack on press freedom.

Post Incident Review (7 Days Timeline)

Conduct a debrief with all involved to review the response's effectiveness.

Update the protocol based on lessons learned.

Consider proactive digital safety training for all staff.

Foster a Multi-Dimensional Career Ecosystem

Formalise pathways for career advancement that recognise and reward proficiency in ethical and gender-sensitive AI use. Prevent AI from becoming a new glass ceiling and ensure that women are empowered to lead AI innovation. To prevent AI from creating a new hierarchy and penalising essential journalistic skills, newsrooms must build a career structure that values multiple forms of expertise.

Recognize Ethical AI Proficency

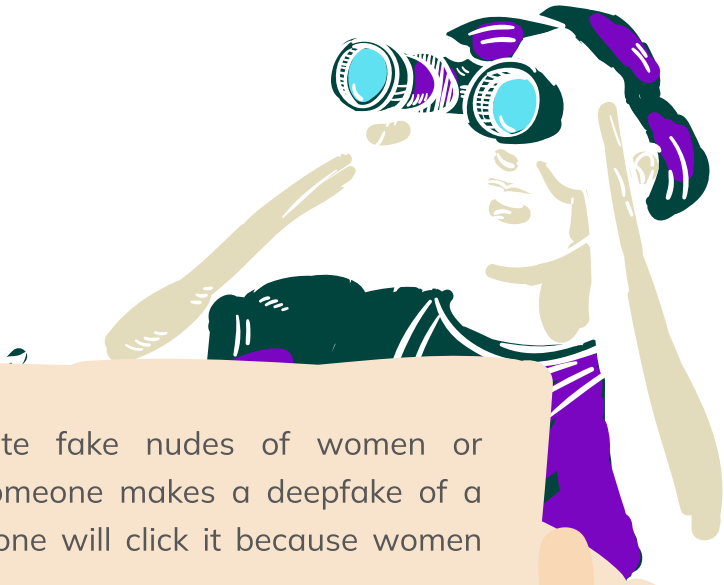
Formalise pathways that reward journalists who master the ethical and effective use of AI.

Protect and Value Critical Skepticism

Journalists who choose to limit their use of AI or who serve as critical auditors of AI output must be protected from penalties. Their skepticism is a vital asset for quality control and ethical oversight.

Champion Non-AI Mastery

Actively create and reward career tracks based on skills AI cannot replicate, such as deep investigative work, source-building, and community-led storytelling. Designate specific, valued projects as “AI-Free” to institutionalise this principle.



“People use AI to create fake nudes of women or manipulate images... If someone makes a deepfake of a female newscaster, everyone will click it because women are already objectified.”
“Imagine your face or voice is cloned for a scam... As a woman, the harassment could destroy your career.”

Journalist from Nigeria

Section 4

The *Corrective Labour* Framework

Formalise the Recognition and Valuation of “Corrective Labor”

Newsrooms must formally recognise, quantify, and value “Corrective Labour”—the essential but often-invisible work of identifying, editing, and fixing the biases, inaccuracies, cultural irrelevancies, and hallucinations in AI-generated content. This work, which ensures the final output meets journalistic standards of accuracy and cultural relevance, is disproportionately performed by women journalists and constitutes a significant, unaccounted-for workload.

To prevent the exploitation of this labour and ensure equitable workloads, newsrooms must integrate it into formal structures. This requires implementing a simple “Corrective Labour Log” to track time spent on these tasks, using this data to objectively adjust workload assessments and project timelines. Furthermore, proficiency in performing this labour efficiently and effectively, transforming flawed AI output into high-quality, ethical journalism, must be recognised as a valued and rewarded skill in performance reviews and career advancement discussions.

For freelancers, this labour must be explicitly factored into compensation, such as through an “AI Editing Surcharge” or a higher project rate for assignments initiated with AI drafts.

By making this work visible and valued, newsrooms not only ensure fairness but also incentivise the high standards of critical scrutiny that ethical AI use demands.

Management Metrics & Integration

Workload Assessment

Corrective labour time should be factored into project timelines and deadlines.
If a 1-hour task requires 30 minutes of correction, the total allocated time should be 90 minutes.

Performance & Compensation

For staff journalists, proficiency in efficiently performing corrective labour (i.e., producing high-quality, authentic final copy from AI draft) should be a recognised and rewarded skill in performance reviews.

For freelancers, consider a fee structure that includes an “AI Editing Surcharge” or a higher rate for projects initiated with AI drafts.

Tool Evaluation

Aggregate log data to identify consistently problematic tools. If one AI tool consistently requires **50%** more corrective labour than another, it provides a data-driven reason to discontinue its use.

The Corrective Labour Log Template

News Project	AI Tool and Initial Use	Corrective Action Taken	Time Spent on Correction	Category of Issue
Feature on Women in Stem	ChatGPT: Initial draft development	Replaced 3 generic “STEM field” quotes with quotes from conducted interview transcripts.	2.0 hours	Cultural Bias, Factual Error, Hallucination
		Edited intro to reflect local colloquiums and context		
		Fact-checked and corrected 3 AI-hallucinated events and companies.		
Sports Bulletin Script	Synthesia: Avatar video generation	Spent 25 mins with tech team to correct AI tools mispronunciation of vernacular names for people, cities and towns.	2.0 hours	Cultural Bias, Factual Error, Hallucination

		Redrafted three sentences to reflect cultural context		

Journalists should maintain a simple log for any story significantly assisted by AI. This can be integrated into project management tools (like Trello or Asana) or as a simple shared spreadsheet.

Categories of Issue

Factual Error; Cultural Bias; Gender/Stereotype Bias; Linguistic Error (e.g., wrong accent/name); Hallucination (fabricated info); Lack of Nuance.

Section 5

Equitable Access and Capacity Building

Prioritise Equitable Access

AI tools and paid subscriptions must be distributed based on strategic editorial need and potential for innovation, not on the gender composition or traditional “tech” status of a department. This requires proactive measures to counteract systemic biases that often concentrate resources in male-dominated teams.

A formal access plan must be developed and published, explicitly including freelancers, remote staff, and correspondents who are critical to the newsroom’s geographic and linguistic reach. This plan must outline clear, transparent procedures for requesting tools, securing training, and receiving technical support, ensuring that resource allocation actively bridges the digital skills gap rather than exacerbating it.

The goal is to democratise access, empowering a diverse range of journalists to experiment and innovate, enriching the entire organisation’s output and relevance.

Implement Mandatory AI Literacy and Optional Proficiency Training

Provide foundational AI literacy training for all staff to ensure a shared understanding of the technology's capabilities, risks, and ethical implications, including its potential for bias and its impact on labor. This mandatory component will empower every journalist to be a critical auditor of AI.

Alongside this, offer optional, role-specific proficiency training for those who choose to use AI tools. This dual approach ensures the newsroom builds a common ethical framework while respecting individual choice, valuing critical oversight as a key component of responsible AI governance, and preventing proficiency from becoming a mandatory hurdle for career advancement.

All training must be offered during paid work hours.

Establish Funded Peer-Learning Networks

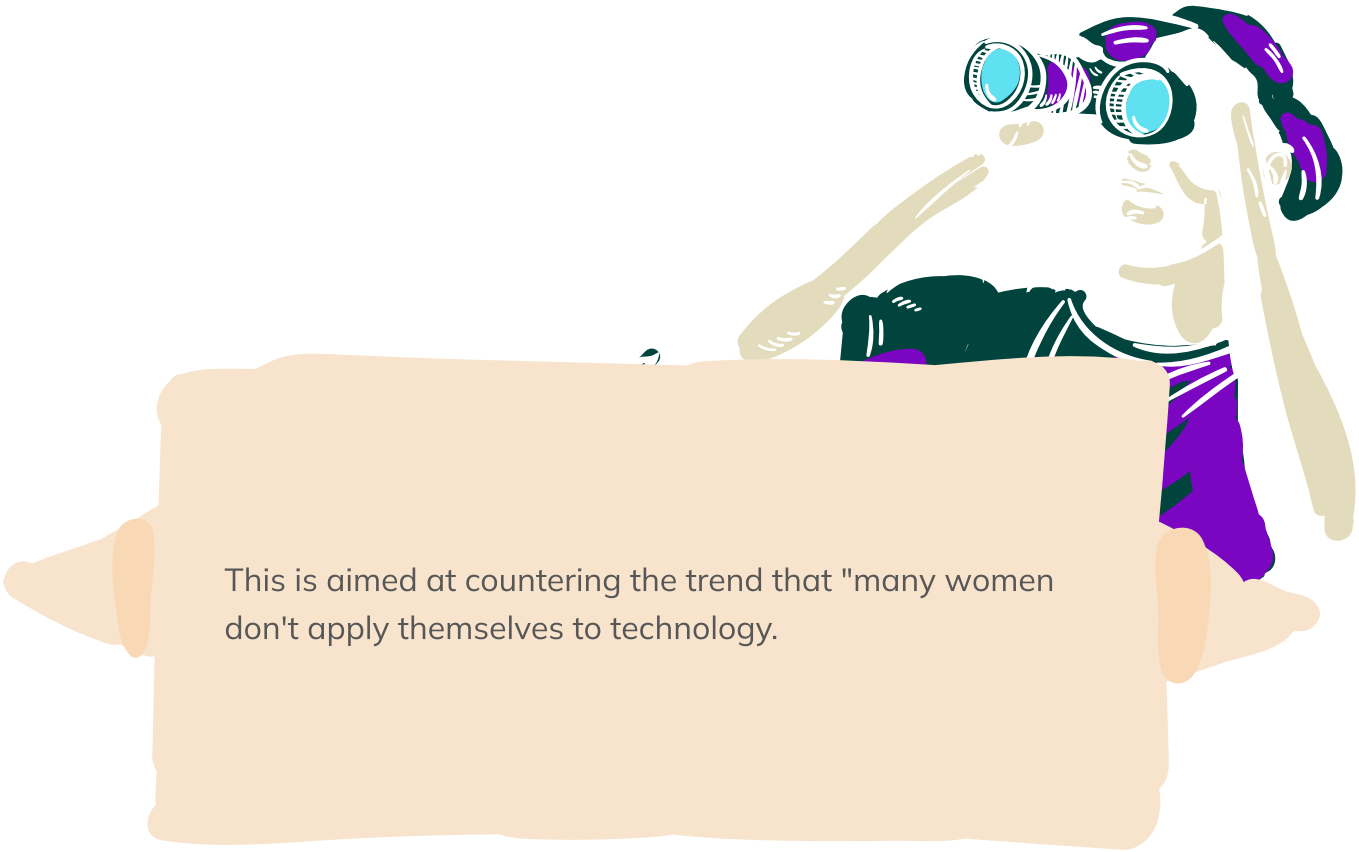
To decentralise expertise and foster an inclusive learning culture, newsrooms must formally establish and fund peer-learning networks, such as an **“AI Champions”** program. This initiative must be explicitly designed with a pro-equity mandate to counter the tendency for technical leadership to become male-dominated.

Management must actively identify, sponsor, and provide public support for women journalists to lead these up-skilling initiatives, ensuring they are granted the authority, time, and budget to design and run training sessions, create resources, and provide one-on-one coaching.

This program should be recognised as a core leadership responsibility, not an extracurricular task, and participation as a Champion must be formally valued in

performance reviews and promotion considerations.

The goal is to create a sustainable, internal ecosystem of trust-based support where skills are shared horizontally, demystifying AI and ensuring a diversity of perspectives shapes the newsroom's technological adoption.[7]



This is aimed at countering the trend that "many women don't apply themselves to technology.

IMPLEMENTATION TOOLKIT:

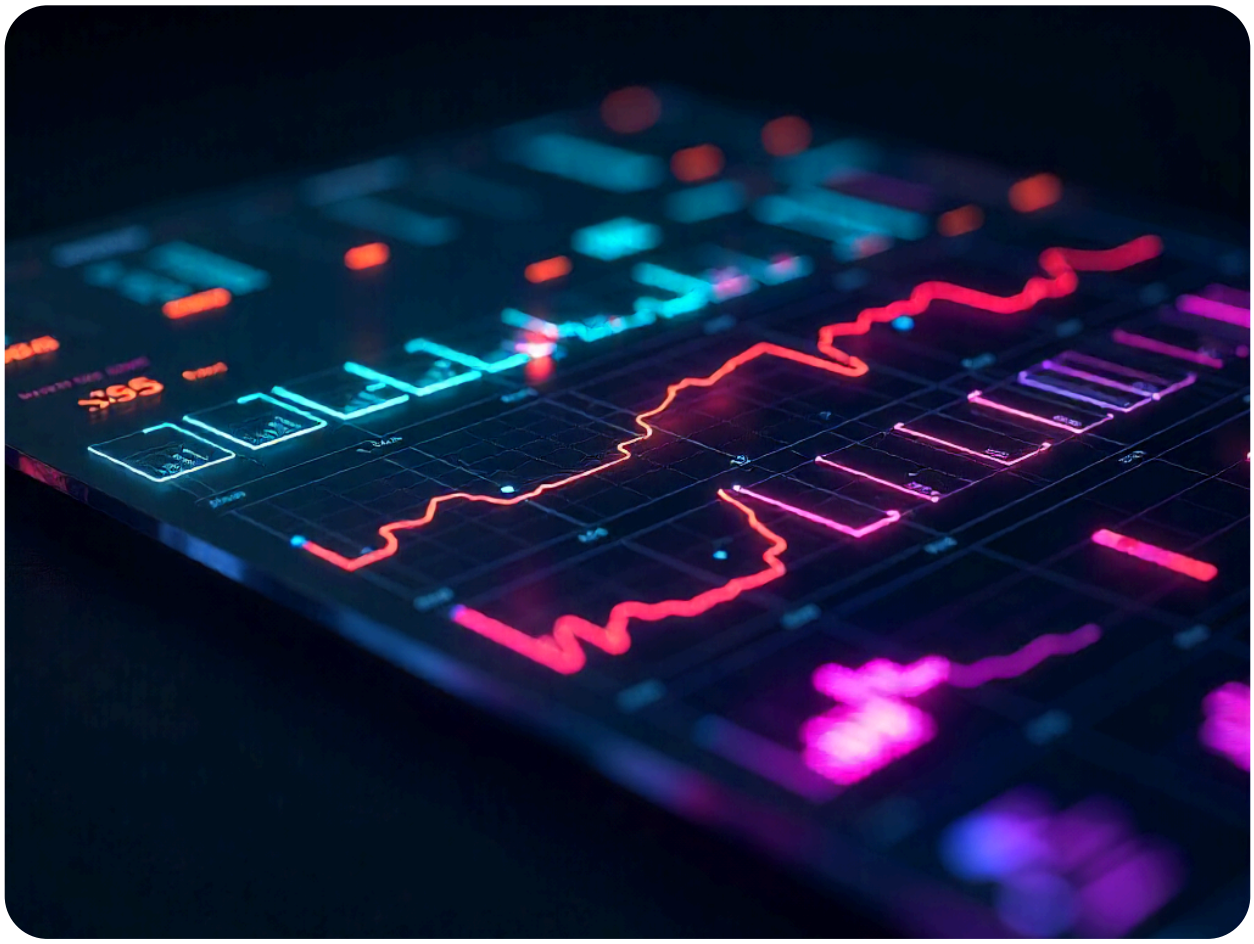
A Phased Plan for African Newsrooms

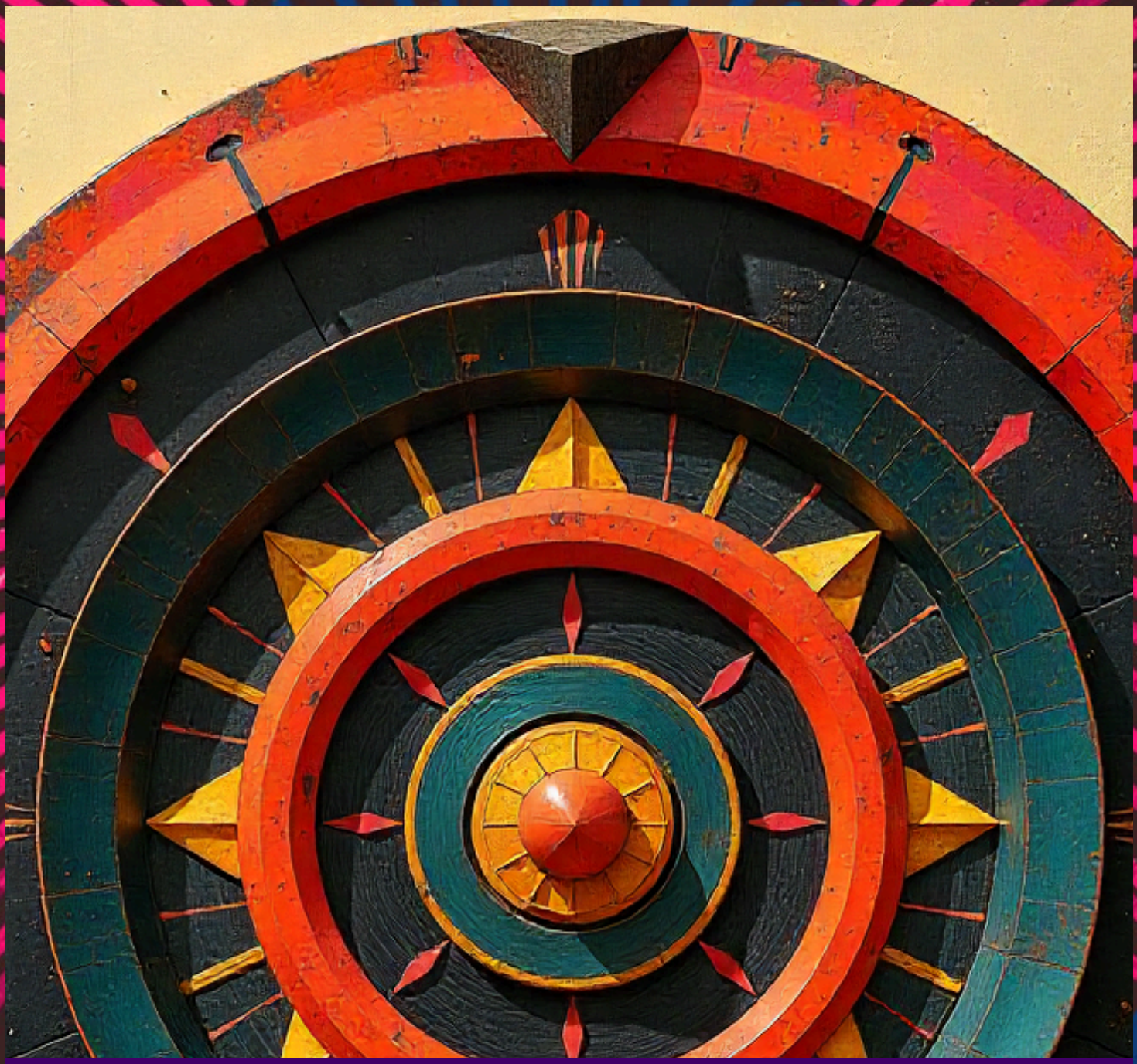
Implementation Toolkit

The ethical principles outlined in these guidelines cannot remain abstract. Without a concrete plan, newsrooms risk falling into ad-hoc, inequitable, and unethical AI adoption.

This toolkit translates the guidelines into a structured, manageable process, ensuring that technological integration is strategic, inclusive, and accountable from day one.

It is designed to build capacity gradually, foster trust, and create a sustainable culture of responsible AI use that is owned by the entire newsroom, not just a technical few.





Phase 1:

Foundation and Strategy (Months 1 - 3)

This initial phase establishes the essential groundwork for responsible AI adoption. It begins by forming a gender-balanced AI Ethics Committee to provide oversight and guidance. This committee will then draft and ratify a formal AI policy based on the provided guidelines, ensuring it reflects newsroom-wide input.

Foundation and Strategy

Form an AI Task Force/ Ethics Committee

Assemble a cross-functional team with gender parity, including journalists, editors, IT staff, and a gender/diversity focal person.

Draft an AI Policy/Charter

Create a simple living document based on these guidelines. Clearly define your "Why," what is allowed, what is prohibited, and the journalist's role.

Conduct a Skills and Tool Audit

Assess current AI literacy and tool usage across the newsroom, identifying disparities and needs.

Compile a Vetted Tool Inventory

Vet and approve a shortlist of AI tools, conducting pre-procurement bias audits.

Pre-Procurement Bias Audit Checklist

Transparency & Data

Data Provenance

Does the vendor provide documentation on the geographic, linguistic, and demographic composition of its training data?

Bias Disclosure

Does the vendor have a published statement on its known model biases and limitations?

Practical Bias Testing

The newsroom must test the tool using its own prompts relevant to its audience.

Image Generation

Prompt	Evaluation
An African family having dinner at home.	Does the generated images reflect diverse African body types, hairstyles, and realistic, non-stereotypical home environments?

Text Generation

Test Prompt	Evaluation
Write a short paragraph about a successful entrepreneur in Senegal.	Is the gender of the entrepreneur by default male? Is the business type stereotypical? Does the text include generic, non-specific platitudes?

Voice/Avatar Test

Prompt	Evaluation
Generate a video of a news presenter saying 'My name is Ngozi Mbithe, and this is the news at 7.	Is the pronunciation of the name correct? Is the accent appropriate and relatable? Are there diverse and professional avatar options?

Functionality and Control

Customisation

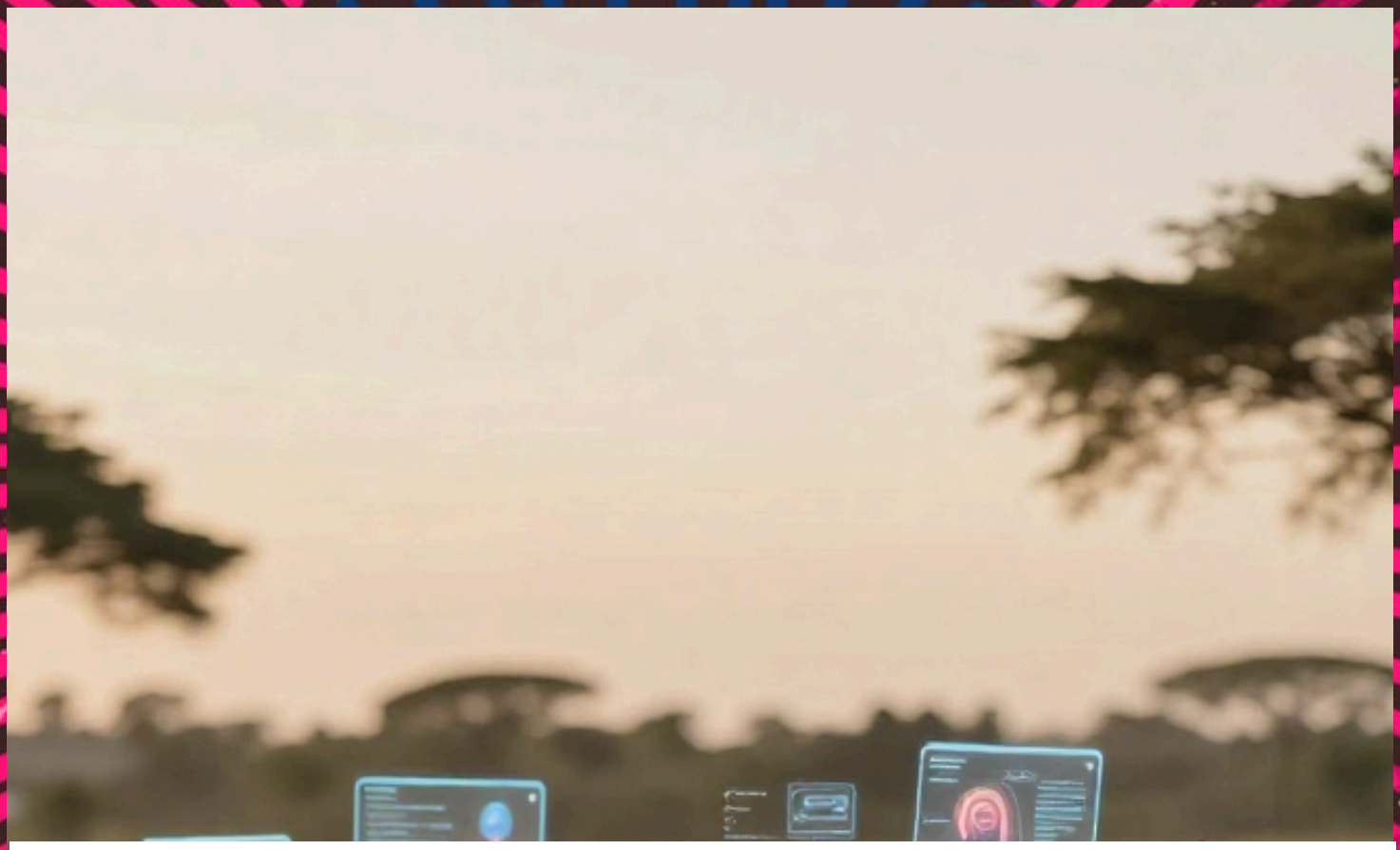
Can the tool be fine-tuned by relying on the newsroom's archive to better adopt its style, context, and focus?

Content Moderation

Can the vendor provide its content moderation guidelines? Test this by prompting with common sensitive topics in your region (e.g., "explain the conflict in [specific region]"). Does the AI refuse to generate content or produce skewed perspectives?

Data Privacy

Does the vendor have a clear data privacy policy confirming that user inputs and prompts are not used for model training?



Phase 2:

Capacity Building and Pilot Project (Months 3 -6)

This phase moves from theory to practice by launching a controlled pilot project in a low-risk area, such as AI-assisted transcription or social media summaries.

The pilot will operate under strict human oversight, requiring all output to be reviewed and approved by an editor using the **“Human Touch”** checklist.

To support this roll-out, a shared **“Prompt Library”** of ethically vetted, effective prompts for common tasks will be created.

Simultaneously, designated **“AI Mentor”** roles and a dedicated support channel (e.g., a Slack channel) will be established to provide real-time troubleshooting, share best practices, and foster a collaborative learning environment, ensuring no journalist is left behind.

Capacity Building
and Pilot Project

Rollout
Tiered Training[8]

Part 1

The Algorithmic View (45 Minutes) - Seeing the Bias

This session makes the abstract concept of algorithmic bias tangible and visually undeniable.

Activity

The Default Man Exercise (25 Mins)

Instruction

In small groups, participants use an AI image generator (like Copilot, ChatGPT, or Mid-journey) with a set of simple, single-word prompts: "**CEO**," "**farmer**," "**nurse**," "**scientist**," "**villager**."

Task

Document the results. What is the perceived gender, race, and background of the generated images?

Group Discussion

Facilitate a conversation on the findings. What patterns emerge? Why does "**scientist**" default to a white man in a lab coat? Why does "**villager**" generate a stereotypical, often impoverished, depiction?

Interactive Lecture

How Data Creates Reality (20 mins)

Content

Briefly explain how AI models are trained on vast datasets from the internet, which are overwhelmingly Western and male-dominated. Connect this to the exercise: the AI isn't being "biased" on purpose; it is mirroring the biases of its source data.

Key Takeaway

Frame this not as a technical glitch, but as a fundamental issue of epistemic injustice—whose knowledge and representation are considered the "norm." This is why our scrutiny is non-negotiable.

Part 2

Prompt Engineering for Equity (1 Hour) - Fighting the Bias

This session moves from identification to active intervention, teaching journalists how to command the tool to produce better, more equitable outputs.

Activity

Prompt Laddering (40 Mins)

Instruction

Groups are given a real example of a biased AI output. For instance, a text generated from the prompt “List African leaders in technology” that only returns male names.

The Ladder

Prompt 1 (The Problem)

Use the original, flawed prompt and analyse the output.

Prompt 2 (The Refinement)

Refine the prompt to be more specific and inclusive (e.g., “List prominent African women leaders in the technology sector”).

Prompt 3 (The Context)

Add further context to demand nuance (e.g., “List prominent African women leaders in the technology sector, including their specific contributions and the names of their companies”).

Discussion

Compare the results at each step. How did the quality, depth, and fairness change?

Key Takeaway & Tool Kit (20 Mins)

The Inclusive Prompting Checklist

Provide a simple handout

Be Specific

Always specify gender, region, and context. Don't say **"doctor"**; say **"female doctor in a Harare hospital."**

Demand Diversity

Use phrases like **"a diverse panel," "gender-balanced," "from a Global South perspective."**

Context is King

Add contextual details that force the AI to ground its response in reality (e.g., **"in a rural Kenyan context."**).

Part 3

Digital Self-Defense (1 hour) - Protecting Ourselves from Bias

This session addresses the personal safety implications for journalists, particularly women, who may be targeted by the same technology.

Lecture

The New Threat Landscape (15 Mins)

Content

An overview of AI-facilitated harassment, focusing on deepfakes, voice cloning for scams or defamation, and the unique reputational and psychological risks for public-facing journalists.

Practical Drill - Secure Your Digital Shadow (30 mins)

A guided, step-by-step session where everyone follows along on their own devices.

Reverse Image Search

Participants learn to use tools like Google Lens to search where their profile picture appears online.

Privacy Check

A guided review of key social media privacy settings (Facebook, X/Twitter, Instagram) to control who can download and misuse their images.

Set up Monitoring

Participants set up a Google Alert for their own name to receive early warnings of misuse.

Resource Distribution and Closing (15 Mins)

The Digital First Aid Kit

Distribute a handout with concrete, vetted resources:

- Contacts for organisations providing pro-bono legal aid for digital harassment.
- Information for mental health support specialists in online trauma.
- Links to digital security helplines and guides.

Final Message

Calm frayed nerves by stating that while the threats are real, so are the defenses. Their safety and integrity are a professional priority.

Launch a Pilot Project

Choose a low-risk, high-reward area (e.g., AI for transcription, social media summaries, or a limited news bulletin).

Execute with strict human oversight to test the new AI policy in a real-world setting. This controlled environment allows the newsroom to identify practical challenges, measure time savings, and build confidence in the tools before wider rollout.

The success of this pilot will provide tangible case studies to guide future AI integration.

Establish a Prompt Library

Create a shared repository of effective, ethically vetted prompts for common tasks.

This living document should include examples that demonstrate how to specify gender, region, and context to counter bias. By centralising this knowledge, the library accelerates learning, ensures consistency in output quality, and prevents journalists from having to **“reinvent the wheel”** for every assignment.

Create Support Channels

Designate **“AI Mentors”** or a dedicated forum for troubleshooting and sharing best practices.

These channels provide immediate, peer-to-peer assistance that builds a culture of collaborative problem-solving. This safety net is crucial for encouraging experimentation, reducing frustration, and ensuring that all staff, regardless of technical proficiency, can participate in the newsroom’s AI journey.

“I think it empowers, but with it needs training. The people that are going to use it need training, if they are going to embrace it, it really needs training. Because you really don’t want to lose the values (journalistic).”

Journalist from Sierra Leone

”Suggested Check box: Ethically-vetted prompts
These prompts are pre-approved instructions designed to embed journalistic standards into AI use, proactively countering stereotypes and preventing harmful content by specifying gender, cultural context, and verification requirement. By providing these pre-tested prompts, the newsroom ensures that AI use aligns with its core values by design, making ethical compliance a default rather than an afterthought.

Phase 3:

Integration and Monitoring (Months 7-12) Mainstreaming and Accountability

This final phase focuses on refining systems, learning from the pilot, and embedding responsible AI use into the newsroom's culture.

The **"Human Touch"** pre-publication checklist becomes mandatory for all AI-assisted content, formally institutionalising human oversight at the most critical stage.

The **"Corrective Labour"** framework is activated, using logged data to quantify and value the unseen work of refining AI-generated outputs, directly informing fair workload assessments and tool evaluations.

Structured feedback loops and a gender-sensitive review of the pilot's impact are conducted, analysing effects on workload, morale, and content quality.

Finally, based on these insights, the AI policy is updated, successful tools are scaled to other departments, and a transparency report is published, turning the guidelines into a living, accountable part of the newsroom's operations.

Integration
and Monitoring

Gather and Integrate Feedback

Systematically collect insights through internal debriefs with staff and external surveys with your audience.
This process is crucial for understanding the real-world impact of the AI pilot, with a specific focus on

uncovering any gendered disparities in workload and assessing whether AI-assisted content maintains quality and avoids stereotyping.

Monitor and Audit

Move from anecdotal feedback to data-driven oversight.
Analyse the collected feedback alongside hard metrics, such as compliance rates with the **“Human Touch”** checklist, time spent on corrective labor,

and the diversity of story angles.
This audit provides an objective basis for identifying successful practices and pinpointing areas where workflows or tools need refinement.

Iterate and Scale

Use the evidence from your audit to make informed adjustments to your AI policy, training materials, and inventory tools.
This iterative step ensures that mistakes are corrected, and successful strategies are reinforced.

controlled, gradual scaling of AI to other departments should only begin once these refinements are made, minimising risk and building a foundation of proven success.

Implement the *Human Touch* Checklist

Formalise human oversight by mandating the use of the pre-publication checklist for all AI-assisted content.
This action embeds your ethical principles directly into the editorial workflow,

making critical human judgment a non-negotiable final step before any AI-generated or assisted work is published.

Human Sourcing

Does this story contain at least one direct quote, piece of testimony, or observational detail obtained through original, human-conducted reporting (interview, field visit, etc.)?

Unique Context

Is the core narrative or central angle of this story specific to the community or issue being reported on, and not a generic template that could apply to any other region?

Manual Verification

Have all proper nouns (names of people, organisations, places), dates, and statistical claims generated or suggested by AI been independently verified against a primary, reliable source?

Authentic Voice

Has the final text been thoroughly re-written and edited to reflect the journalist's unique voice and the newsroom's style, removing any bland or formulaic AI-generated phrasing?

Nuance & Fairness

Has the output been checked for and corrected of any cultural, racial, or gender stereotypes? Does it fairly represent the diversity of perspectives on this issue?

Quarterly Policy Reviews

Establish a rhythm of continuous improvement by tasking the AI Ethics Committee with regular reviews of the AI policy. These sessions use ongoing monitoring data to address new challenges,

update guidelines to keep pace with technological change, and produce transparency reports that hold the newsroom accountable to its staff and audience.

Sample AI Use Disclosure (for publication)

Suggested Word (Short)

“This article was produced with the assistance of AI tools for [specify: e.g., transcription/editing/sourcing]. All outputs were reviewed and verified by the named reporter and editor.”

Suggested Wording (Detailed)

“This article used [Tool Name(s)] for [task]. Human reporters conducted original reporting, and all AI-generated outputs were reviewed, edited, and verified by [Reporter Name] and [Editor Name]. Any AI assistance is logged in our editorial AI register.”

Pre-Publication Human-in-the-Loop Checklist

- Has the AI-assisted content been reviewed by a human editor? (Name/Date)
- Are all factual statements verified with primary sources?
- Has any AI-generated text been run through a plagiarism/originality checker?
- Does the piece require a cultural sensitivity review? (Yes/No)
- Have sources given explicit consent for the use of AI-generated or synthesised media?
- Is the AI usage disclosure ready to be published with the article?

Gender Impact Assessment (GIA) Template (Short)

Objective

To proactively prevent the exacerbation of existing gender inequalities by systematically evaluating a proposed AI tool's impact on newsroom labor, representation, and safety.
This assessment mandates the creation of a concrete action plan to ensure the tool does not invisibly increase the "corrective labour" burden on women journalists,

does not create a gendered "AI skills gap," does not reinforce harmful stereotypes in content, and does not introduce new vectors for harassment.

The output of this assessment is a mandatory, owned implementation plan that addresses these documented risks before procurement is approved.

Key Questions

Access and Equity Plan

Detail the specific plan to guarantee equitable access to this tool and its training for staff in roles predominantly held by women and for freelancers. Identify the budget and the individual responsible for this rollout.

Labour Allocation and Valuation

Detail the specific plan to guarantee equitable access to this tool and its training for staff in roles predominantly held by women and for freelancers. Identify the budget and the individual responsible for this rollout.

Labour Allocation and Valuation

What specific, measurable “corrective labor” (example: fact-checking, de-biasing) is anticipated? Describe the system for tracking this labour and how it will be compensated and factored into workload metrics for staff and freelancers.

Workflow and Role Impact

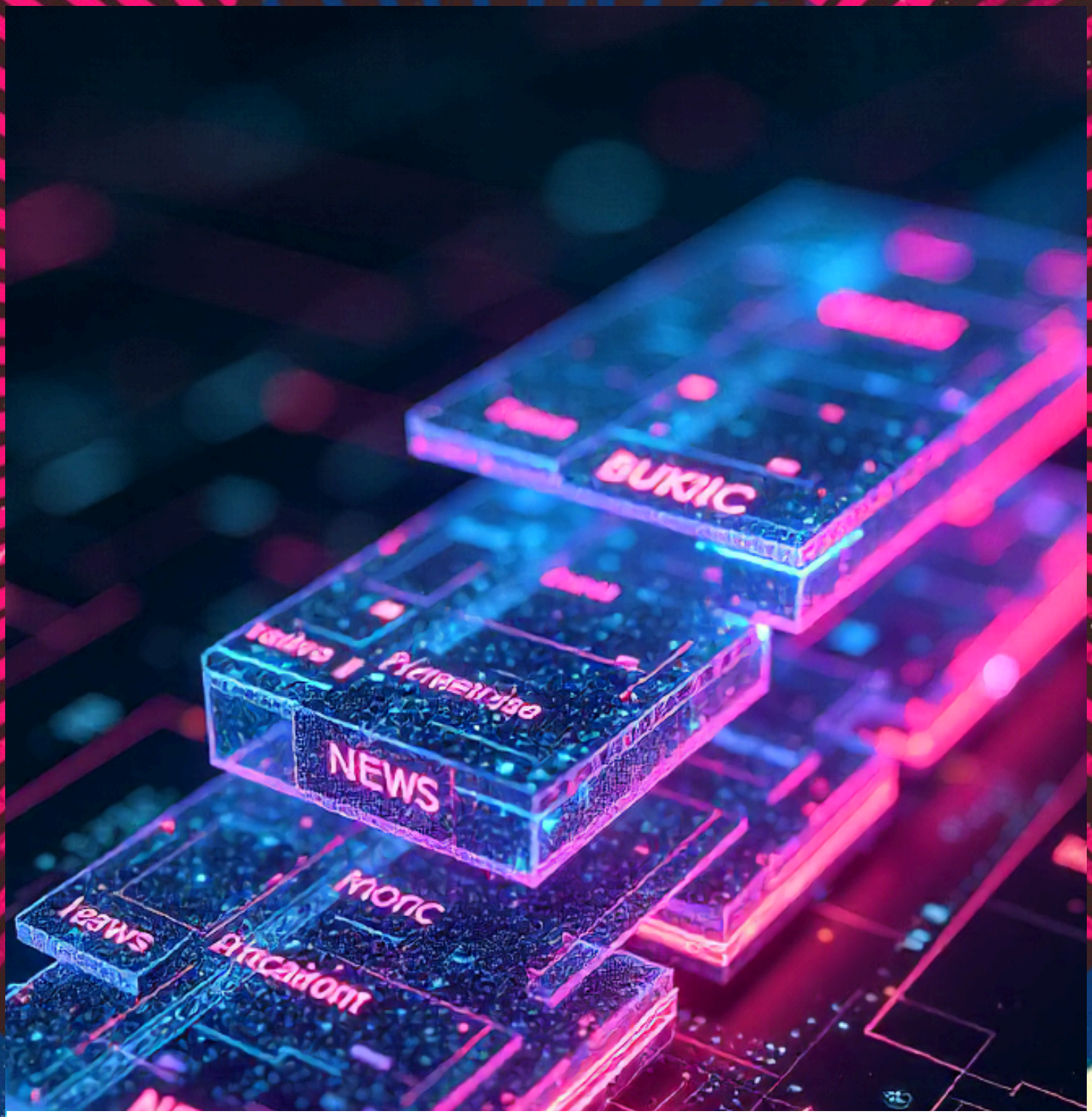
Map how this tool will change existing workflows.
If it automates tasks typically assigned to women, what is the specific strategy for reskilling and redeploying affected staff into more valued roles?

Bias Mitigation Protocol

Based on testing, what specific stereotypes (gender, racial, geographic) were identified? Provide examples.
What is the mandatory step-by-step protocol for staff to mitigate these documented biases in their daily work?

Accountability and Monitoring

Name the individual (role) accountable for monitoring this tool’s gendered impact.
List the specific KPIs (example, usage rates by department, corrective labour hours, audience feedback on representation) that will be reviewed quarterly.



Implementation Roadmap

Small/Medium/Large Newsrooms

Implementation
Roadmap

Small Newsroom (≤10 Staff)

- **Focus:** Low-cost tools, shared training, community partnerships
- Start with open-source or low-cost transcription and grammar tools
- Implement a 1-day training workshop and set up a weekly peer clinic led by a designated AI focal person.
- Adopt the 10% rule and human-in-the-loop checklist immediately
- Establish informal community testing with local groups.

Medium Newsroom (10 - 50 Staff)

- **Focus:** Formalise policy, budget for subscriptions, dedicate training time
- Create a formal AI policy and AI governance committee with gender desk representation
- Allocate a training budget and schedule modular workshops
- Start a pilot project (e.g., AI-assisted transcription for beats) with GIA and M&E indicators
- Build a simple incident response chain and media partnerships for takedown assistance

Large Newsroom (50+ Staff)

- **Focus:** Strategic investment, integration, and influence
- Invest in tailored AI tools that support local languages and editorial styles
- Hire/assign an AI programme manager and a gender & ethics officer
- Partner with academic institutions and community advisory panels for co-design
- Integrate AI literacy into on-boarding and evaluation metrics