



Digital Democracy
Initiative

FUTURE-PROOFING ELECTIONS AGAINST DEEPFAKE DISINFORMATION

ADVOCACY BRIEF

ADVOCACY RECOMMENDATIONS

The following recommendations build directly on the evidence presented in this report and on comparative insights from case studies in Namibia, Ecuador, Singapore and Germany. These country-specific contexts illustrate how deepfake and synthetic media disinformation affect the information space differently and depending on the strength of democratic institutions, media ecosystems, and the capacity of civil society. In Namibia, gendered cheapfakes exposed weaknesses in local-language moderation; in Ecuador, AI-generated narratives intersected with violence and criminal networks; in Singapore, tight regulation protected electoral processes but risked restricting free expression; and in Germany, strong regulation and independent fact-checking partnerships proved effective in maintaining public trust.

Together, these case studies demonstrate that the threat of deepfake disinformation is not defined solely by technology but also by the social and institutional systems surrounding it. The recommendations that follow take these lessons to form a set of global advocacy priorities grounded in transparency, equity, collaboration and human rights—principles essential for building democratic resilience in the age of generative AI.

The recommendations are structured around five interconnected action pillars designed to foster proactive, rather than reactive, approaches to information integrity.

- 1. Platform Accountability:** Transparency and shared responsibility. Platforms must be transparent about how content circulates and empower independent oversight to ensure fairness across all regions and languages.
- 2. AI Ethics and Human Impact**
Assessments: Accountability through ethics and inclusion. AI companies must be held to measurable ethical standards that

uphold human rights, safety, and inclusivity, especially the inclusion of voices from the Global South.

- 3. Regulatory Reform:** Protect rights while preventing harm. Regulatory approaches must safeguard free expression, creativity, and press freedom as well as mitigate the malicious use of synthetic media.
- 4. Public Resilience and Literacy:**
Empowerment through knowledge. Public resilience is democracy's first line of defence. Strengthening citizens' ability to identify and critically assess synthetic content strengthens immunity against manipulation
- 5. Leverage AI to Strengthen Civil Society:** Use technology responsibly for transparency and inclusion. AI should be a tool for civic empowerment, not only a source of risk.

5.1 Platform Accountability:

Recommendation 1: Ensure Global Transparency and Accountability of Social Media Platforms

Rationale:

Social media platforms are the primary vectors for deepfake disinformation. In more than **80%** of countries that held elections in 2024–2025, GenAI content increased and was amplified by opaque algorithms. This uneven moderation, particularly in non-English contexts such as Ecuador and Namibia, undermines electoral integrity and fuels polarisation.

Action:

Establish a CSO-led global campaign, in partnership with multilateral bodies, to advocate for open data access, decentralised moderation, and consistent election safeguards across all countries.

Key Priorities:

- Open data access: Require platforms to provide affordable, research-grade datasets for independent monitoring of algorithmic

performance and content moderation.

- Decentralised moderation: Adopt geofenced systems with local-language reviewers to capture cultural nuance and improve accuracy.
- Election parity: Guarantee that every national election benefits from a dedicated election-integrity team and real-time coordination with civil society.
- After each election, platforms should publish datasets on detections of synthetic media, related enforcement actions, and content moderation timelines to inform evidence-based reform.

Recommendation 2: Guarantee Equitable Data Access and Localised Moderation Standards

Rationale:

Without consistent data access and multilingual coverage, civil society cannot hold platforms accountable. Equity in data and content moderation is essential for fair and transparent digital governance.

Action:

Embed equitable access and localisation requirements in global and national policy frameworks to ensure all regions, particularly the Global South, benefit from the same transparency and protection standards.

Key Priorities:

- Enforce mandatory, affordable data access for CSOs and researchers.
- Mandate localised, language-specific moderation to close linguistic gaps.
- Require post-election transparency reports from platforms be made public in every country.

Recommendation 3: Build a Global Coalition for Platform Reform and Capacity Strengthening

Rationale:

Sustained reform depends on coordinated pressure and shared expertise. Cross-sector coalitions can combine advocacy, monitoring,

and capacity-building to drive systemic change.

Action:

Launch a multisectoral “Open Data for Democracy” campaign that brings together civil society, governments, researchers and international organisations to advocate for transparency and accountability in platform governance.

Key Priorities:

- Coalition building: Partner with relevant organisations to ensure Global South leadership.
- Public advocacy: Combine public mobilisation with policy engagement at UN and regional forums, supported by expert webinars and joint reports.
- Decentralised moderation advocacy: Promote in-country moderation teams fluent in local languages; coordinate open letters setting measurable benchmarks; link executive incentives to election-integrity goals; and publish annual public scorecards grading platforms by coverage, speed, and transparency.
- Capacity building: Train CSOs in OSINT and social-media forensics and develop a global dashboard to track platform compliance.

Recommendation 4: Institutionalise International Standards and Monitoring Mechanisms

Rationale:

Long-term accountability requires global norms and regular, transparent performance evaluation. Standard-setting must ensure fairness, transparency, and human rights compliance across digital ecosystems.

Action:

Codify international standards on algorithmic transparency, equitable moderation, and election safeguards through UN and regional bodies, supported by continuous CSO monitoring and independent audits.

Key Priorities:

- Short term: Build coalitions, launch the campaign, and secure pilot data-sharing

agreements with at least two major platforms, including one Global South election.

- Medium term: Embed decentralised moderation and transparency requirements in platform policies and regulatory frameworks.
- Long term: Adopt international norms on data access, algorithmic transparency, and election protection through UN and regional instruments.
- Metrics: Track the number of platforms offering free research-grade application programming interfaces (APIs), post-election datasets, and measurable reductions in language-disparity complaints verified through annual CSO and user surveys.

5.2 AI Ethics and Human Impact Assessments

Demanding ethical standards from tech companies moves governance from reactive to proactive ethics, embedding accountability throughout the AI lifecycle. By coupling civil society advocacy with international cooperation, this approach balances innovation with harm prevention and ensures that the evolution of GenAI aligns with democratic principles and human rights values.

Recommendation 1: Mandate Ethical Standards and Accountability Across the AI Lifecycle

Rationale:

Companies such as OpenAI, Google, and xAI have introduced powerful tools without consistent ethical safeguards. Ethical principles are often reactive and impact assessments typically occur after deployment, when harms such as gendered disinformation or non-consensual imagery have already spread.

Action:

Establish binding ethical requirements across the entire AI lifecycle. This includes mandatory human rights and psychosocial impact assessments (HIAs) at the design, development, and deployment stages, as well as employ

continuous post-release monitoring.

Key Priorities:

- Ethics by design: Shift from reactive governance to prevention through early-stage ethical review.
- Lifecycle assessments: Ensure HIAs evaluate not only technical risks but also social and cognitive impacts, including susceptibility to manipulation.
- International alignment: Coordinate with global frameworks such as UNESCO's Recommendation on the Ethics of AI to harmonise standards.

Recommendation 2: Establish Independent Ethics Boards within Technology Companies

Rationale:

Corporate self-regulation has proven inadequate. Ethical oversight must be independent, transparent, and inclusive. Lessons from Meta's Oversight Board show that legitimacy depends on autonomy, diversity, and binding authority. A Global AI Ethics Coalition, including academic and civil society partners, could develop common standards for these boards. The Coalition would define prohibited uses (e.g., electoral manipulation or targeted suppression), establish risk registers and release notes, and oversee regular, regionally disaggregated audits.

Action:

Require major AI firms to establish independent ethics boards with representation from civil society and affected communities. These boards should review high-risk AI systems—especially election-related tools such as voice cloning, targeting systems, and image generation—and publish their findings in the public domain.

Key Priorities:

- Boards must include journalists, human rights lawyers, technical experts and affected communities.
- Clear mandates should require “comply or explain” responses to recommendations.

- Public case summaries and rationales must be published regularly to ensure transparency.

Recommendation 3: Embed Global Ethical Standards in AI Governance Frameworks

Rationale:

Without shared norms, AI governance risks reinforcing Global North dominance and deepening inequities. Global coordination is essential to ensure that ethical standards are inclusive, enforceable, and locally relevant.

Action:

Promote the adoption of international norms through a UN-led AI Ethics Pact that embeds ethical principles into global governance frameworks and national regulations. This will institutionalise standards and make ethics a requirement for access to public or multilateral funding.

Key Priorities:

- Establish clear definitions of harm and accountability mechanisms that prioritise prevention over profit.
- Ensure representation from Global South experts and gender-diverse voices in all norm-setting processes.
- Link compliance with access to public procurement or partnership eligibility.

Recommendation 4: Strengthen Monitoring and Accountability Mechanisms

Rationale:

Ethical standards are only practical when compliance is independently verified. Regular reporting, third-party audits, and public transparency are crucial to maintaining credibility and trust.

Action:

Require all major AI companies to publish annual ethics reports, independently audited and verified by civil society. CSOs should conduct shadow reporting to provide independent performance assessments and

expose gaps.

Key Priorities:

- An increased number of companies with functioning ethics boards.
- Evidence of HIA integration across AI product releases.
- Reduction in reported harms and bias incidents (tracked via IPIE and other observatories).

5.3 Building Public Resilience

Building public resilience transforms individuals into informed participants rather than passive recipients of information, creating a cumulative, society-wide shield against the evolving threat of deepfake disinformation.

Recommendation 1: Strengthen Digital Literacy and Critical Thinking

Rationale:

Deepfakes exploit universal cognitive biases and cultural specific norms of trust and sharing. The 2024 evidence shows that awareness and preparation helped limit their impact, yet vulnerabilities persist, especially among excluded groups and in information-saturated environments. Empowering citizens with critical thinking and media literacy skills is the most durable defence against manipulation.

Action:

Design pre-emptive education and awareness campaigns that teach citizens how to identify synthetic media, question sources, and verify information before sharing.

Key Priorities:

- Develop learning materials that explain typical manipulation techniques (voice cloning, spliced video, fabricated context) and simple verification steps: pause, source, date, trace.
- Produce resources in local languages and adapt to communication habits: radio and WhatsApp in rural areas, short videos in urban centres, posters in low-connectivity

zones.

- Partner with teachers' unions, community centres, public broadcasters and newsrooms to co-create workshops and public-service content for schools and youth programmes.

Recommendation 2: Build Collaborative Networks for Trusted Communication

Rationale:

People are more likely to accept corrections and prebunks from sources they already trust. Sustained collaboration between credible local actors strengthens collective immunity against disinformation.

Action:

Establish distributed networks of trusted messengers, such as community radio hosts, diaspora leaders, editors' forums, election authorities and digital rights groups, to issue rapid, credible corrections and prebunks.

Key Priorities:

- Formalise partnerships with clear escalation channels to platforms and election bodies.
- Provide partners with copy-ready prebunk scripts and adaptable broadcast formats.
- Encourage two-way feedback loops so that local observations feed into national and global early-warning systems.

Recommendation 3: Integrate Resilience into Education and Civic Infrastructure

Rationale:

One-off campaigns have a limited effect. Embedding resilience in education systems and civic institutions ensures continuity and long-term impact.

Action:

Mainstream digital-resilience curricula in national education policies and link them to broader civic-engagement initiatives.

Key Priorities:

- Incorporate media-literacy modules into school syllabi and teacher-training

programmes.

- Provide sustained funding for public-awareness units within election-related bodies and information-integrity CSOs.
- Align with UNESCO's Media and Information Literacy framework to harmonise standards globally.

Implementation Timeline and Metrics:

- Short term: Launch pilot campaigns in selected countries.
- Medium term: Scale programmes regionally and embed in national curricula.
- Long term: Codify media literacy and resilience standards into global education and governance frameworks.
- Metrics: Measured increases in public awareness and critical-thinking indicators; reduced misinformation-sharing rates in behavioural studies; and number and diversity of partnerships formed.

5.4 Leveraging AI for Civil Society

While much of the policy debate on AI focuses on its risks to democracy, the same technologies hold vast potential to strengthen civil society, expand participation, and enhance resilience.

As the report highlights, AI can be harnessed for social good: from detecting disinformation to improving accessibility, translating civic content across languages, and deepening public engagement. Harnessing these benefits requires proactive leadership from civil society itself, ensuring that AI development and deployment align with human rights, transparency, and inclusion. By investing in ethical adoption, capacity building, and governance frameworks, AI can become a tool that amplifies civic voices, accelerates accountability, and helps safeguard the integrity of democratic processes worldwide.

Recommendation 1: Harness AI for Social Good and Civil Society Capacity

Rationale:

While much of the discourse around AI focuses on its risks, the same technologies can significantly strengthen civil society’s reach, responsiveness, and resilience. AI can expand access to information, automate monitoring of disinformation, translate content across languages, and help identify emerging social or political risks. When guided by ethical principles and human rights, AI becomes a tool for empowerment rather than manipulation.

Action:

Encourage and resource civil society to adopt AI tools responsibly across key domains, while ensuring transparency, inclusion, and data protection.

Key Priorities:

- Build awareness and technical literacy within CSOs to safely integrate AI tools.
- Promote open, affordable, and privacy-preserving AI applications tailored to civic needs.
- Establish ethical guidelines and accountability mechanisms to govern use

Below are some positive use cases of AI for civil society:

Application Area	AI Function / Tool	Example / Impact	Ethical & Governance Considerations
Disinformation Monitoring & Analysis	Machine learning for pattern recognition, content clustering, and synthetic media detection	Tools like Reality Defender or Deepware Scanner detect deepfakes early, enabling CSOs to flag harmful content before it spreads virally.	Ensure algorithmic transparency; avoid over-reliance on automated labelling.
Translation & Inclusion	Natural Language Processing (NLP) for real-time multilingual translation	AI-driven translation bridges linguistic gaps in countries like Namibia, improving outreach to rural and indigenous communities.	Prioritise cultural context and local dialect accuracy; address linguistic bias.
Media & Fact-Checking	AI-assisted verification; reverse image search automation	Tools like Truepic and InVID-WeVerify accelerate visual verification for journalists and CSOs.	Combine human review with AI checks; maintain privacy of image metadata.
Data Analysis for Advocacy	Predictive analytics to identify emerging social or political risks	AI-driven dashboards visualise trends in online hate speech or electoral discourse, helping with early intervention.	Require data anonymisation and informed consent for datasets.

Accessibility & Inclusion	Speech-to-text and text-to-speech tools; summarisation and adaptive interfaces	AI enhances accessibility for people with disabilities (e.g., voice interfaces for visually impaired voters).	Ensure compliance with data privacy and disability-access standards.
Crisis Response & Humanitarian Aid	AI forecasting for disaster response and spikes in misinformation	Predictive models assist humanitarian CSOs to anticipate floods or surges in misinformation.	Avoid surveillance misuse; ensure community consent in data collection.
Civic Engagement & Dialogue	Chatbots and conversational AI for public information campaigns	Civic bots provide verified election information or counter disinformation narratives in real time.	Ensure transparency so users know they are interacting with AI, not humans.

Recommendation 2: Build Ethical and Technical Capacity in the Civil Society Sector

Rationale:

For AI to be a force for good, civil society must not only have access to the tools but also understand and shape how they are built. The current asymmetry between AI developers and civic actors risks replicating existing digital inequities. By embedding ethical AI use into civil society practice, the sector can transform from a reactive actor to an innovative co-architect of the digital future, using the same technologies that spread disinformation to build trust, transparency, and resilience.

Action:

Develop regional AI Capacity Hubs for Civil Society, linking universities, technical experts, and grassroots organisations to co-design open-source civic AI tools.

Key Priorities:

- Provide training on AI ethics, data governance, and risk assessment for CSOs.
- Support the creation of shared repositories of open, local-language civic datasets.
- Partner with philanthropic donors and public institutions to fund civic-tech incubators.

Recommendation 3: Establish Governance Frameworks for Responsible AI Use

Rationale:

Civil society's adoption of AI must reflect the same accountability standards demanded of governments and corporations. Ethical guardrails ensure that civic applications of AI respect privacy, autonomy, and human dignity.

Action:

Develop a Civic AI Ethics Charter outlining clear principles for transparency, data protection, inclusivity, and public accountability.

Key Priorities:

- Mandate transparent disclosure of AI-assisted outputs and decision-making processes.
- Incorporate community feedback loops to assess impact and unintended consequences.
- Align governance standards with frameworks such as UNESCO's Recommendation on the Ethics of AI and the OECD's Principles on AI.

Implementation Timeline and Metrics:

Short term: Identify existing AI use cases and launch regional training pilots in 3–5 countries.

Medium term: Establish AI Capacity Hubs and release the Civic AI Ethics Charter.

Long term: Institutionalise AI literacy and ethical governance within civil society networks globally.

Metrics: Number of CSOs adopting AI tools ethically; documented improvements in digital inclusion; number of civic datasets or open tools developed; and independent audits confirming adherence to the ethics charter.



Digital Democracy Initiative

The Digital Democracy Initiative (DDI) is a programme to safeguard inclusive democracy and human rights in the digital age. It focuses on supporting local civil society in the Global South, particularly in countries undergoing democratic regression and where civic space is under pressure.

For more information visit:

digitaldemocracyinitiative.net



CIVICUS

CIVICUS is a global alliance of civil society organisations and activists working to strengthen citizen action and civil society throughout the world.

For more information visit:

civicus.org