
AI AND DEMOCRACY

Resources on misinformation, public trust, civic life, media literacy, elections, public institutions, and democratic participation in the age of AI.

AI & Election Integrity

Resources on how AI is already affecting election administration, voter information, and the security of the vote.

Resource / Activity	Description & Application	Source
AI Democracy Projects Chatbot Study	A collaborative test in which election officials and AI researchers asked five leading chatbots basic voter questions (where to vote, registration deadlines) and found false or misleading answers at least half the time — including telling Nevada voters same-day registration wasn't available when it has been law since 2019. A concrete case study for why AI chatbots shouldn't be treated as authoritative sources on civic logistics.	<i>AI Democracy Projects: Proof News and the Science, Technology, and Social Values Lab, Institute for Advanced Study (2024)</i>
AI Elections Accord	A voluntary agreement among more than 20 major technology companies — including Google, Meta, Microsoft, OpenAI, and TikTok — to build detection tools for deceptive AI election content, support provenance labeling, and fund public media-literacy efforts ahead of 2024's global elections. Useful as a primary document for discussing the gap between voluntary industry commitments and enforceable regulation.	<i>Tech Accord to Combat Deceptive Use of AI in 2024 Elections, Munich Security Conference (February 2024)</i>
Election Official AI Tabletop Exercise	A crisis-simulation format in which participants (election officials, in the original version) work through realistic AI-enabled attack scenarios — a deepfake of an official spreading false voting instructions, an AI-generated phishing wave — and practice a coordinated response. Adaptable as a classroom simulation for a public administration or elections course.	<i>Brennan Center for Justice, the Elections Group, and Institute for the Future, AI scenario planner (2023–2024)</i>
Safeguards for AI in Election Administration	A framework for election offices considering AI tools for their own operations (not just defending against outside misuse), covering vendor transparency, testing before deployment, and ongoing monitoring. Directly useful for a case study on responsible AI procurement in a public-sector setting.	<i>Brennan Center for Justice, Safeguards for Using Artificial Intelligence in Election Administration (2023)</i>
State Deepfake Election Law Tracker	Students review the rapid, uneven landscape of state laws — over 150 bills introduced in 2024 alone — regulating AI-generated deceptive content in political communications, and compare a few states' approaches side by side. Shows students that "AI and elections" is already an active, fast-moving area of state policymaking, not just a hypothetical future concern.	<i>Brennan Center for Justice, States Take the Lead in Regulating AI in Elections (2024)</i>

Synthetic Media & Provenance

Tools and standards for tracing where a piece of media actually came from in an environment where realistic fakes are cheap to produce.

Resource / Activity	Description & Application	Source
Content Credentials / C2PA Standard	An open technical standard, backed by Adobe, Microsoft, Google, OpenAI, and others, that attaches cryptographically signed provenance data to a file recording what created or edited it. Students can examine a Content Credentials label on a real image and discuss what it does and doesn't prove.	<i>Coalition for Content Provenance and Authenticity (C2PA), formed 2021</i>
Watermarking Generative Output	Students compare two technical approaches to labeling AI-generated content — embedded watermarking (invisible signals baked into the pixels or audio) versus attached metadata (a removable record traveling with the file) — and discuss why bad actors can strip one but not always the other.	<i>NIST AI 100-4, Reducing Risks Posed by Synthetic Content (2024); Google DeepMind's SynthID</i>
Reading a Synthetic Media Regulation	Students read a single state's actual deepfake-in-elections statute and identify what specifically counts as regulated content, who can be held liable, and what the enforcement mechanism is — surfacing how much these details vary even among laws aimed at the same underlying problem.	<i>Adaptable using any current state statute tracked by the Brennan Center's state legislation database</i>

AI-Mediated Deliberation & Civic Participation

Emerging tools that use AI to help large, disagreeing groups find common ground rather than to generate persuasive content for one side.

Resource / Activity	Description & Application	Source
The Habermas Machine	An AI system, named for philosopher Jürgen Habermas, that mediates small-group discussions on contested political topics by drafting a group statement, incorporating participants' critiques, and iterating until it earns broad endorsement. In a 5,700-person UK study, participants rated its statements as fairer and clearer than ones from human mediators, and groups converged more than when discussing unmediated.	<i>Tessler et al., Google DeepMind with Oxford, Harvard, and Yale researchers, Science (2024)</i>
Pol.is and vTaiwan	An opinion-mapping platform where participants can only agree, disagree, or submit their own statement (no replies), which a clustering algorithm uses to surface where genuine common ground exists across a large, divided group rather than amplifying the loudest voices. Used in Taiwan's vTaiwan process to reach actionable consensus on more than 30 contested policy questions.	<i>Computational Democracy Project; Audrey Tang, Taiwan Ministry of Digital Affairs</i>
AI Mediator vs. Human Facilitator Debate	Students debate the tradeoffs the Habermas Machine research surfaced directly: an AI mediator can process a large group's input in seconds and may reduce personal bias, but a genuine consensus statement optimized for maximum endorsement raises its own question about whether "agreement" is the right goal of democratic deliberation in the first place.	<i>Adapted from critical responses to Tessler et al. (2024), including Habermas-theory critiques of consensus-optimization</i>

Public Trust & Institutions

Resources connecting AI's effects on the information ecosystem back to the institutions — local news, platforms, government offices — that public trust actually depends on.

Resource / Activity	Description & Application	Source
AI and the Local News Crisis	Students examine data on the collapse of local newsrooms — a third of U.S. newspapers gone since 2005 — alongside the emergence of AI tools that can both help remaining small newsrooms do more with less and threaten to further erode traffic and revenue by answering questions AI-summarized without a click-through. Connects an AI policy question directly to the civic information gap in students' own communities.	<i>Medill Local News Initiative, State of Local News Project, Northwestern University</i>
Santa Clara Principles	A civil-society-authored set of standards for what meaningful transparency and due process should look like when platforms moderate content at scale, including specific expectations for disclosing when moderation decisions were made by an automated system rather than a person.	<i>Coalition of civil society organizations and academics, Santa Clara Principles on Transparency and Accountability in Content Moderation (2018, revised 2021)</i>
Chatbots as a Civic Information Source Discussion	Students discuss the emerging pattern — documented in the AI Democracy Projects research — of citizens turning to general-purpose chatbots instead of official government sources for basic civic questions, and what that shift means for which institutions the public actually trusts to be accurate.	<i>Grounded in AI Democracy Projects (2024) and Pew Research Center trust-in-institutions tracking</i>
Designing a Responsible Civic AI Tool	As a capstone exercise, student teams design (on paper) an AI tool meant to help residents navigate a real local government process, specifying what data it would need, how it would handle a question it can't answer confidently, and what oversight would catch it if it started giving bad information at scale.	<i>Synthesis exercise drawing on Brennan Center election-administration safeguards and AI Democracy Projects findings</i>

Sources & Further Reading

AI Democracy Projects (Proof News and the Science, Technology, and Social Values Lab, Institute for Advanced Study). (2024). AI chatbots found inaccurate in answering voter queries.

Brennan Center for Justice. (2023). Safeguards for using artificial intelligence in election administration.

Brennan Center for Justice. (2024). States take the lead in regulating AI in elections — within limits.

Brennan Center for Justice, the Elections Group, and Institute for the Future. (2023–2024). AI scenario planner for election officials.

Coalition for Content Provenance and Authenticity (C2PA). (2021–present). Content Credentials specification. c2pa.org

Computational Democracy Project. Pol.is: An open-source platform for gathering and mapping public opinion at scale.

Medill Local News Initiative, Northwestern University. (2024–2025). State of Local News Project.

Munich Security Conference. (2024). AI Elections Accord: A tech accord to combat deceptive use of AI in 2024 elections.

National Institute of Standards and Technology (NIST). (2024). NIST AI 100-4: Reducing risks posed by synthetic content.

Santa Clara Principles Coalition. (2018, revised 2021). The Santa Clara Principles on transparency and accountability in content moderation.

Tang, A., & Weinberger, D. Plurality: The future of collaborative technology and democracy.

Tessler, M. H., et al. (2024). AI can help humans find common ground in democratic deliberation. *Science*, 386(6719).