Artificial Intelligence and Election Integrity in 2024

Assessing ethical questions, concerns, and trade-offs as 2 billion voters cast their ballots

CARNEGIE
ETHICS
FELLOWS

CARNEGIE
COUNCIL for Ethics in International Affairs

Contents

The following project was produced by a working group from the inaugural Carnegie Ethics Fellows cohort reflecting nearly two years of convenings, collaboration, and research. Each report in this special series examines a critical issue at the intersection of ethics and international affairs. The ***Carnegie Ethics Fellowship*** *aims to develop the next generation of ethical leaders across business, government, academia, and non-governmental organizations.*

Authors

TRAVIS GIDADO

Carnegie Ethics Fellow and Associate at Kirkland & Ellis LLP

Travis Gidado is an associate with Kirkland & Ellis LLP. His practice focuses on diverse corporate matters, including mergers and acquisitions, corporate governance, and minority investments. Gidado graduated from Yale University with a BA in ethics, politics, and economics, and his first job was with Goldman, Sachs & Co. as a legal analyst. He also holds an MPhil in public policy from King's College, Cambridge, and a Master of Law in China studies from the Yenching Academy of Peking University, where he was a Yenching Scholar. Gidado received a JD-MBA from the University of Chicago Law School and the Booth School of Business, respectively. He has worked for a wide range of public and private sector institutions, including the Center for Strategic and International Studies, the Obama administration, and Atlantic Media.

CHRISTINE JAKOBSON

Carnegie Ethics Fellow and Principal at Principia

Christine Jakobson is an ethics researcher and advisor, helping to make the world a better place through philosophy. She received a Ph.D. focusing on moral philosophy from the University of Cambridge and an MSt from the University of Oxford. Jakobson is a principal at Principia where she works across technology, finance, energy, law, audit, and insurance, advising executive leaders on the ethics of technology, ethical decision-making and leadership, ethical culture, and strategy. Jakobson was selected for the 2022 research sprint at the Berkman Klein Center for Internet Studies at Harvard University and was also awarded a coaching fellowship from the Women's Impact Alliance.

HINH TRAN

Carnegie Ethics Fellow and Lead Counsel, Litigation & Employment at Ramp

Hinh Tran is lead counsel, litigation & employment at Ramp, a finance automation startup based in New York City. Previously, he was a litigation associate at Keker, Van Nest & Peters and a judicial law clerk for U.S. District Judge Dale A. Drozd. Tran also served as a Bates Legal Fellow for the World Intellectual Property Organization in Geneva, Switzerland, and as a Google Public Policy Fellow. He holds a JD from the University of Michigan Law School, a BA from the University of California, Berkeley, and he also studied at Stanford University's Hoover Institution and the National University of Singapore's Faculty of Law.

A Historic Electoral Milestone

Marked by an unprecedented number of elections, 2024 represents a historic milestone in democratic governance for scholars and citizens alike. Approximately 64 countries—representing over 2 billion people—will have elections this year. Among those polities, India, the United States, Indonesia, Brazil, and Mexico stand out as the largest and most economically prosperous democracies with significant electoral processes—not to mention the European Parliament. And political mobilization of this scale should not be taken for granted: This level of collective electoral participation will not be seen again until 2048.

Although those who support democracy as the best political framework for facilitating economic prosperity and safeguarding basic human rights should be heartened by widespread suffrage, political and economic volatility remain major countervailing concerns. Perhaps the biggest threat to democratic governance is the least understood: Societies still do not know how to sufficiently combat the use of artificial intelligence (AI) by malevolent actors who are deploying this technology to fuel manipulation, misinformation, and disinformation campaigns. From deepfake images to AI bots masquerading as real people on social media platforms, there are plenty of tools for malicious actors to use if they wish to undermine democratic regimes—and worse, most of these activities have become extremely difficult to circumvent. For example, during the 2020 U.S. presidential election, AI-generated profiles were used to spread misinformation and create confusion among voters. Notable cases involved AI-generated fake news about political candidates, amplified by bots across media platforms, making it harder for voters to discern legitimate information from fabricated content. These activities have become quite sophisticated and remarkably challenging to detect and mitigate, undermining democratic integrity in the process.

A comparison with past electoral cycles reveals the extent to which technological concerns have taken precedence over other issues. Historically, election monitors were primarily focused on ensuring voters' physical security, combating classic forms of election fraud (e.g., stuffing ballot boxes with illegal votes on behalf of a favored candidate or party), and confirming the results of manual vote counts. Now, election monitors must pay closer attention to social media posts spreading lies about voting locations or doctored images of candidates accompanied with untrue statements about their political views.

With generative AI platforms such as ChatGPT growing more adept by the day, widespread adoption of this technology alongside substantial efforts to get out the vote in 2024 is set to meaningfully reshape the election landscape and potentially influence democratic processes in unforeseen ways. As their impact intensifies, it will be important to understand the challenges that adoption of this technology presents while also understanding how societies can become more resilient in facing such challenges. This article provides a summary of the trade-offs associated with managing AI's potential impact on elections, examining broad global and national implications for safeguarding electoral integrity.

Balancing Competing Values and Key Trade-offs

In spite of the legitimate concerns AI presents—many of which have been mentioned at the outset of this article—these technologies also offer promising opportunities for safeguarding electoral security and democratic integrity. AI can be employed to detect and mitigate cyber attacks, such as hacking attempts on voting systems and databases. Machine-learning algorithms can analyze vast amounts of data to identify patterns indicative of fraud or interference, enabling faster responses to possible threats. Automated systems are also capable of **processing votes with much higher accuracy** than manual programs, **minimizing latent human error risks**. Furthermore, AI-driven tools can improve verification processes for voter identification and registration, thereby reducing fraud risks and ensuring compliance with legal and ethical standards. Proactive implementation of AI in these areas (among many others) can foster greater public trust in electoral outcomes by helping make democratic processes more transparent and reliable.

Of course, we cannot ignore the risk that self-interested actors will try to leverage AI for their own gain from an electoral perspective. For example, leaders of fragile democracies may seek to exploit AI technologies to entrench incumbent power by monitoring and suppressing opposition activities, manipulating public opinion, and skewing electoral outcomes in favor of the ruling regime. By undermining democratic processes in this way, indefinite, prolonged incumbency may result in the erosion of basic human rights over time. Therefore, we believe there is an urgent need for cooperation between national governments, multinational technology companies, NGOs, academic institutions, and media outfits to counter AI-related threats to democracy in a collaborative fashion. Beyond basic knowledge-sharing, these cross-cutting partnerships can help bridge technological divides wherever they exist and ensure greater equanimity in access to the tools and infrastructure needed to protect sensitive electoral processes.

To help encourage the development of multilateral, multi-stakeholder partnerships that can manage the threats AI presents to democracy while harnessing its positive potential, we have identified six key trade-offs for interested parties to consider: privacy versus transparency, security versus accessibility, innovation versus stability, safeguards versus chilling effects, efficiency versus accountability, and centralization versus decentralization. These trade-offs may serve as useful lenses through which to understand how AI may be leveraged to mount a positive (i.e., proactive) defense of democracy as malevolent actors aim to goad decision-makers into negative (i.e., reactive) responses, recognizing the inescapable value judgments prompted by this multifaceted analysis.

Privacy versus Transparency

When applying AI technologies to the electoral process, there is a critical trade-off between privacy and transparency. While AI can improve transparency by making electoral processes more open and data-driven, it can also infringe on individual privacy. For example, AI-driven voter verification systems can ensure that only eligible voters participate, therefore enhancing transparency at a time when political parties are challenging the resiliency of their own systems with greater aplomb (e.g., U.S. President Donald Trump's assertion that votes were rigged against him during the 2020 election). However, these systems will likely collect and store other sensitive personal data, raising privacy and data security concerns. Striking the right balance requires careful regulation to protect privacy without compromising transparency, and such a nuanced approach will be difficult to achieve at the start of any effort to leverage novel technologies.

Security versus Accessibility

AI can enhance election security by identifying and mitigating cyber threats, but it may also limit accessibility to the political process on a local or national level. Advanced AI systems often require significant technological infrastructure and expertise, which may not be available to all prospective voters. Enhancing security through AI might inadvertently prove exclusionary, creating disparities in how different communities perceive electoral integrity. Returning to the United States as an illustrative example in this regard, imagine a world where prominent "blue" or Democratic-leaning states have robust election security measures in place, but key "red" or Republican-leaning states do not. If one of those red states has an issue with vote counts, it may be used by political parties to drive discourse around the results toward divisive ends, whether it is by claiming an illegitimate outcome or highlighting gaps between the "haves and have-nots" on a national scale. As a result, ensuring accessibility while maintaining security is a complex balancing act that necessitates creating inclusive technological solutions for all potential actors.

Innovation versus Stability

Rapid innovation around AI technologies presents a notable trade-off with electoral stability. While innovative AI applications can modernize and improve electoral processes, their introduction may also introduce unforeseen risks if co-opted by malign forces. New AI tools, such as chatbots meant to answer questions about how to cast one's ballot in a particular region, might malfunction or be exploited by malicious actors. If a chatbot intended to be a source of truth is manipulated to spread falsehoods at scale, there is no telling how far downstream the impacts may go. And if the impacts are significant enough, it may even necessitate a re-vote or extension of the promised electoral timeframe, sowing doubt in the process. Balancing the benefits of innovation with the need to ensure stable and predictable electoral processes is essential for maintaining public trust in democratic systems.

Safeguards versus Chilling Effects

Connected to the idea of innovation versus stability, maintaining safeguards against AI abuse will prove fundamental if democratic societies are to be protected from relevant threats. Yet, we must preserve some degree of nuance with respect to the safeguards implemented. Returning to the significant achievement ChatGPT represents, that technological leap was made possible by a business landscape and policy apparatus that strongly supported innovation. One could imagine a regulatory landscape that was so restrictive in its scope that it actually inhibited innovation and dissuaded entrepreneurs from taking the risks necessary to realize their visions. Therefore, even though AI must be kept within appropriate guardrails that can be constructed and policed, it must also be given the ability to develop in ways that can be beneficial for society—and its creators—long term. AI leadership will surely come to shape geopolitics, economic growth, and societal development over time, and regulators must thread the needle between protecting citizens and enabling technological creatives to test the limits of what is possible in the digital realm with all the latent capacity to improve society (including democratic governance) for the better.

Efficiency versus Accountability

AI systems have the potential to improve the efficiency of electoral processes, such as voter registration or vote-counting. However, these efficiencies can sometimes come at the cost of accountability. AI systems, particularly those based on complex algorithms, can operate as “black boxes” that make it difficult to understand their decision-making processes (for example, consider recommendation algorithms that determine what individual people see on Google or social media apps whose main feeds are based on complex inputs). If errors occur, determining responsibility and ensuring accountability becomes challenging. Finding a balance between leveraging AI for its promised efficiency gains while ensuring accountability when things go awry is crucial for reassuring citizens that election results can be trusted.

Centralization versus Decentralization

Finally, AI deployment in elections may be best managed by using centralized systems for data processing and decision-making. A centralized approach naturally can be understood to enhance coordination and effectiveness. At the same time, centralization would likely result in concentrated power, and such power increases the risk of potential abuse whether by actors with direct access to the underlying infrastructure or malign parties who can more easily target the resulting system. In comparison, decentralized systems may reduce the risk of concentrated power, making manipulation more difficult to achieve. This heightened resiliency against attack must be balanced against greater challenges in monitoring and potential inconsistencies across the system. Imagine an interconnected series of servers designed to manage a national election campaign, while local governments are given the freedom to determine which servers to select according to the resources available to them. The only stipulation is that these servers must be interoperable. Without a clear mandate for establishing baseline levels of quality and sophistication, one could envision an outcome where some regions have high-quality and largely secure servers while others end up with less robust infrastructure. If these lesser servers fail to live up to expectations while a live election is in process, such weak links may undermine the entire electoral framework. In summary, the sliding scale between centralized and decentralized technological systems for managing elections must involve establishing frameworks that are consistent in their quality, universally fair and not easily susceptible to mismanagement or power grabbing.

Key Trade-offs to Consider

Privacy versus Transparency

While AI can improve transparency by making electoral processes more open and data-driven, it can also infringe on individual privacy.

Security versus Accessibility

AI can enhance election security by identifying and mitigating cyber threats, but it may also limit accessibility to the political process on a local or national level.

Innovation versus Stability

While AI innovations can modernize elections, they also introduce risks that could destabilize the process if misused, making it crucial to balance progress with electoral stability.

Safeguards versus Chilling Effects

While safeguards against AI abuse are essential for protecting democracy, overly restrictive regulations could stifle innovation and hinder beneficial developments.

Efficiency versus Accountability

While AI can enhance the efficiency of electoral processes, it risks accountability by operating as a “black box,” making it difficult to determine responsibility when errors occur.

Centralization versus Decentralization

While centralization enhances coordination, it risks power concentration and abuse, whereas decentralization increases security but introduces monitoring challenges and inconsistencies.

Deepfake Mitigation: Lessons from the 2024 Mexican Presidential Election

In the run up to a landmark general election that saw Claudia Sheinbaum become the first female president in Mexican history, concerns regarding the impact of artificial intelligence swirled around the race. Although it remains unclear how significant the impact of this nascent technology was, it certainly shaped discourse around the election as candidates were forced to debunk deepfakes and doctored posts that spread lies about their respective platforms. Malevolent actors leveraged AI-generated content to make false claims about Sheinbaum's campaign, with one famously claiming that her campaign was failing by **using audio** that was altered to sound like it was coming from the candidate herself. And it was not just campaign-oriented misinformation: Given the weight that any information purported as coming from Sheinbaum would carry in Mexican society during the campaign, fraudsters also saw value in leveraging her voice for financial gain. A well-circulated deepfake video of Sheinbaum was **used to spread investment-related scams**.

The most high-profile instances of deepfake use targeted the new president's campaign, but misinformation efforts affecting her opponent Xóchitl Gálvez added another level of complexity to the Mexican general election. President Sheinbaum's victory represents continuity for the ruling Morena party, previously led by the former president Andrés Manuel López Obrador, a charismatic left-wing leader who is no stranger to **leveraging misinformation opportunities for his benefit**. Given his popularity and the power of the "bully pulpit" that his presidency carried, AMLO (as he is commonly known) faced few restrictions in being able to parrot falsehoods about Gálvez's campaign—falsehoods that would eventually **make their way to supportive "troll" accounts** on X/Twitter and other social media platforms. Even though Gálvez did her best to debunk the lies, once they became social media fodder (thanks to the support of the president), they took on a new life, demonstrating the challenges of fighting misinformation once it is allowed to grow online.

Gálvez's experience also serves as a reminder that the greatest threats to electoral integrity can come from institutional actors: It is difficult enough to protect voters from third-party generated misinformation, but when the falsehoods come from leaders that should feel an obligation to protect their own citizens from such lies, safeguarding this process is almost impossible. Given President Sheinbaum's own experience with misinformation, one can only hope that she will be much more diligent than her predecessor in ensuring that the statements she issues are bereft of falsities that malevolent actors (or even her own party) can use to beguile citizens.

Beyond the particular attacks Mexican presidential candidates faced using AI platforms, the Mexican general election also demonstrated the risks of AI as applied to the very institutions that have been entrusted with ensuring electoral integrity. Mexico's election authority, the Instituto Nacional Electoral (INE), has a broad mandate for organizing and overseeing elections at the federal level. During the recent election cycle, misinformation campaigns targeting the INE gained significant traction, with one claiming that it was possible **to erase the markers** the INE handed out to help voters cast their votes, therefore making it possible to vote multiple times (which would constitute fraud). If malevolent actors are able to undermine an independent organization tasked with ensuring the legitimacy of Mexican elections, then it is difficult to see how ordinary Mexicans will be encouraged to trust the outcome of democratic elections over time. Combine this with attempts to throw the INE into question levied by AMLO himself, and these actions pose perhaps the greatest threat to Mexican electoral integrity in the long-term. From the new president to the INE, key stakeholders must come together to strengthen Mexico's resiliency against future attacks on electoral integrity.

Amidst the outcome of the Mexican elections, there are some green shoots that should give observers hope for strong cross-cutting partnerships from an electoral perspective. One positive example is a recent multi-stakeholder approach bringing together policy advocates, journalists, and government officials. Representing a joint effort by Obturador Photo Agency, a collection of Mexican photojournalists; the Canadian Broadcasting Corporation (CBC); and the German Marshall Fund's (GMF) Technology Program, technological systems were introduced that would help **verify the authenticity of electoral images** using file metadata. By training editors to use these tools, it will better enable them to confirm whether images have been doctored or falsified, which will make it easier to separate fake photos from real images before they become widely circulated in articles. In other cases, it will enable photojournalists to confirm whether photos already in circulation are fake and decry their use accordingly. Although this technology is not guaranteed to capture every false image that enters (or could enter) the digital realm, it would certainly empower sophisticated actors on the frontlines of democratic speech to help safeguard democratic ideals one image at a time.

Conclusion

Integrating AI into democratic electoral processes presents a complex array of trade-offs with corresponding risks and benefits—all made more acute during this year of unprecedented political mobilization. While AI holds the promise of enhancing election security and integrity, it also introduces new challenges that must be carefully managed. Policymakers and election authorities must navigate such trade-offs with caution, ensuring that the deployment of AI does not undermine the very democratic principles it aims to protect. By fostering transparency, accountability, and equitable access to AI technologies, societies can harness the vast potential of these novel tools to strengthen democracy while mitigating their risks. As this year unfolds, the ethical and strategic deployment of AI to support free, fair elections may come to represent a crucial inflection point in democratic governance worldwide.

"While AI holds the promise of enhancing election security and integrity, it also introduces new challenges that must be carefully managed."

THE CARNEGIE ETHICS FELLOWSHIP

In today's world of geopolitical upheaval and global economic transformation, where can young leaders go to understand the power of ethical leadership, its impact on multilateral cooperation and collaboration, and how it applies to their professional and personal lives?

The Carnegie Ethics Fellowship is a space for talented young professionals to develop their capabilities and be examples of values-driven responsible leadership. Fellows collaborate on projects curated by Carnegie Council for Ethics in International Affairs, giving them the opportunity to contribute to work that has deep connections to both New York and the broader world.

The two-year Fellowship is structured to develop the next generation of ethical leaders from business, government, academia, and non-governmental organizations. The Fellowship is part of Carnegie Council's significant commitment to developing ethics in leadership and to the communities of experts that work toward this end, aligning the power of decision-making with reflective right action.

ABOUT CARNEGIE COUNCIL

Carnegie Council for Ethics in International Affairs is an independent nonprofit that works to empower ethics globally by identifying and addressing the most critical ethical issues of today and tomorrow. Founded by Andrew Carnegie over a century ago, we set the global ethical agenda and work for an ethical future by convening leading experts, building active communities, producing agenda-setting resources, and catalyzing the creation of ethical solutions to global problems. Join us in using the power of ethics to build a better world. Carnegie Council is a nonprofit 501(c) (3) institution. For more information, please visit CarnegieCouncil.org.

Carnegie Council for Ethics in International Affairs is an independent and nonpartisan nonprofit. The views expressed within this project are those of the authors and do not necessarily reflect the position of Carnegie Council.