Generative AI for Election Officials:

Use Cases, Best Practices, and Ethical Guardrails

Thessalia Merivaki and Ioannis Ziogas

McCourt School of Public Policy
Georgetown University

Draft White Paper

September 30, 2025

Table of Contents

Abbreviations

ADA	Americans with Disabilities Act
AI	Artificial Intelligence
API	Application Programming Interface
CISA	Cybersecurity and Infrastructure Security Agency
DL	Deep Learning
EAC	U.S. Election Assistance Commission
GenAI	Generative Artificial Intelligence
HiL	Human-in-the-Loop
LLM	Large Language Model
ML	Machine Learning
NIST	National Institute of Standards and Technology
NLP	Natural Language Processing
RAG	Retrieval-Augmented Generation
RMF	Risk Management Framework
UI	User Interface
VRA	Voting Rights Act

1 Introduction

Generative artificial intelligence (GenAI) has emerged as a potentially transformative instrument for election administration, particularly in contexts characterized by resource scarcity, extreme polarization, heightened public scrutiny, and increasing demands for accessibility and transparency. U.S. election administration is decentralized and most often underfunded. Yet, officials are expected to deliver precise and timely services, maintain the highest standards of professionalism, and preserve public trust. Within the election administration ecosystem, even minor operational or communicative errors can be construed as disenfranchising or partisan, thereby undermining institutional legitimacy in a polarized information environment. In this context, GenAI may serve as a valuable augmentation technology that helps election offices meet escalating informational and operational challenges, provided its use falls under robust governance frameworks and human oversight.[1]

Empirical studies suggest that local election jurisdictions face chronic capacity constraints, including underfunding, understaffing, high turnover rates, and uneven access to technical expertise.[2] These pressures intensify during electoral cycles due to surges in service demands.[3] In such settings, GenAI can act as a capacity multiplier for election officials: it can assist in the generation of frequently asked questions, draft reminders of statutory deadlines, prepare summaries of legislative updates, or even predict potential demand bottlenecks. Particularly when communications are concerned, GenAI applications may produce localized variations of official messages at scale and maintain textual consistency across channels.[4] These uses, while preliminary, suggest measurable efficiency gains in all areas where election officials may encounter demand and supply imbalances.

Moving beyond efficiency, GenAI applications hold significant promise for promoting accessibility and equity in electoral participation. Federal statutory requirements, such as Section 203 of the Voting Rights Act, mandate language assistance for designated minority groups, while the Americans with Disabilities Act imposes obligations for accessible communications.[5] Local election offices often lack the resources to fully meet these obligations. GenAI systems can effortlessly draft plain versions of official notices and generate accessibility features such as alternative

[1] For a comprehensive summarry of identifiable risks embedded in AI use across sectors see the AI RMF 1.0 (NIST 2023, 2024).

[2] Brennan Center for Justice (2025) and Ferrer, Thompson, and Orey (2024); cf. Currie, Wilentz, and Ebrahimi (2022).

[3] For instance, abrupt legal changes, cyberattacks, and disinformation campaigns.

[4] The Elections Group and Peters (2024).

[5] Merivaki and Suttmann-Lea (2022).

text for images, in addition to translating and disseminating relevant documents. However, these outputs require subsequent review by fluent staff and legal experts to ensure accuracy and compliance with statutory language.[6]

Trust, in particular, deserves special mention as it constitutes a core performance indicator of election communications. While empirical studies on the effect of disclosure of GenAI content have been largely inconclusive, the normative rationale for disclosure remains compelling.[7] Transparency, coupled with human review and institutional accountability, is shown to bolster long-term legitimacy even if short-term attitudinal effects are variable. Election officials have additional reasons to embrace transparency: in recent years, they have faced adversarial uses of GenAI, including synthetic audio, video, and text designed to mislead voters. To curb the effects of adversarial uses, the Cybersecurity and Infrastructure Security Agency (CISA) and civil society organizations emphasize the necessity of proactive defensive capacities, such as rapid rumor control messaging and verifiable content provenance.[8] In essence, a consistent and transparent communication ecosystem seems to be key towards maintaining public trust in elections.[9] Accordingly, election officials should treat trust as an *operational metric*, which needs to be monitored and assessed regularly.[10]

2 Understanding (Generative) AI

2.1 Foundations

The term *artificial intelligence* (AI) is often used as a catchall for algorithmic approaches and computational systems that execute tasks associated with perception, reasoning, prediction, and language use.[11] Within this broad term, *machine learning* (ML) refers to methods that learn patterns from data rather than relying

[6] U.S. Census Bureau (2022) and U.S. Department of Justice (2014, 2024).

[7] Jakesch et al. (2019), for instance, reports negligible differences in perceived credibility, while Lim and Schmälzle (2024) find small but significant declines in trust (also see Hancock, Naaman, and Levy 2020).

[8] CISA (2024) and Goldstein et al. (2023).

[9] Ada Lovelace Institute, AI Now Institute, and Open Government Partnership (2021), Grimmelikhuijsen and Meijer (2014), Licht et al. (2014), Selbst (2021), Suttmann-Lea and Merivaki (2023), Tyler (2023), and Veale, Van Kleek, and Binns (2018).

[10] Posard et al. (2024).

[11] Russell and Norvig (2021).

exclusively on "hard-coded" rules.[12] *Deep learning* (DL) is a subfield of ML based on multilayered neural networks that learn hierarchical representations and power contemporary language and vision systems.[13] Finally, *generative AI* (GenAI) denotes models trained to produce novel content, such as text, images, audio, or code, by estimating the distribution of observed data and sampling from it (see Figure 1).[14] For election officials, precision in terminology is neither cosmetic nor epiphenomenal. While much of the confusion around the term "AI" arises from the fact that vendors frequently apply similar marketing labels to describe very different mechanisms, the use of accurate terminology in the context of election administration helps clarify capabilities, constraints, risk profiles, oversight needs, and disclosure requirements.

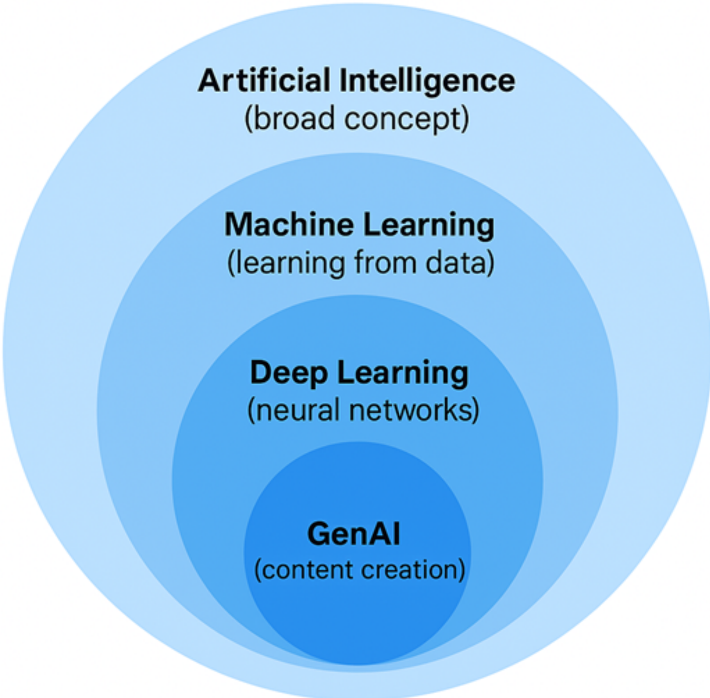

Figure 1: The AI Hierarchy.

[12]Mitchell (1997).
[13]LeCun, Bengio, and Hinton (2015).
[14]It is important to note that large language models (LLMs) are a special subcategory of GenAI systems.

2.2 A Working Taxonomy of AI Models

Instead of providing an exhaustive list of all the different technologies routinely grouped under the generic umbrella term "AI," which is beyond the scope of this report, we focus on a concise toolkit taxonomy most relevant to election administration.[15] Within this taxonomy, we identify three broad families that election officials are most likely to encounter: (a) *rules-based systems*; (b) *predictive and ranking models*; (c) *generative (GenAI) models*. The purpose of this taxonomy is to provide a common framework for discussing potential use cases, benefits, and limitations.

We intentionally avoid assigning the label "chatbot" to any of these families. In this report, we refer to the term *chatbot* as only the chat-based user interface (UI), whereas the underlying mechanism is a specific *model structure*. Given that a chat interface denotes just the structure of the UI, the term "chatbot" by itself conveys no regulatory or oversight information; capabilities and obligations stem from the underlying mechanism instead.[16]

1. **Rules-based systems**

    Description: This category is one of the oldest forms of "artificial intelligence." Rules-based systems perform their functions in the form of rules and decisions. They operate using simple "if–then" statements (e.g. a simple color schema for a thermometer: *IF* temperature $> 100.4\,°\mathrm{F}$ *THEN* red, *ELSE* green). They are relatively easy to set up and require minimal maintenance. They are typically used to automate trivial tasks. This class of models *does not learn*.

    Significance: Rules-based systems are traceable and auditable, suitable for stable FAQs and checklists, but prone to "breaking" outside their set scripts. They may require updating, when and if the rules change.

    Examples: Phone trees that route callers by topic; a form that returns the correct polling place based on address input; a webpage application that attempts to answer common questions; an internal checklist that requires human sign-off before social media messages go live.

[15] For a representative inventory, see Table 1 in Appendix A.

[16] This convention aligns with contemporary practice in the dialogue-systems literature, which distinguishes "assistants," "agents," and "chat systems" as separate families (Gao, Galley, and Li 2019; Jurafsky and Martin 2025; Qin et al. 2023).

2. **Predictive and ranking models**

   Description: Machine learning models that infer patterns from data to make quantitative predictions or ordered assessments about future or unseen cases (e.g. Netflix's recommendation algorithm). Unlike rules-based systems, these algorithms *learn* decision rules automatically rather than having humans explicitly encode them.

   Significance: Useful for forecasting, prioritization, and personalization pipelines. They require labeled data to train and can dynamically retrain on new data; as such, their performance and output validity depend on the quality of the training data. Because they learn from past examples, predictive models must be continuously monitored for performance degradation (concept drift) and bias propagation.

   Examples: Predicting which counties are most at risk for misinformation; ranking polling locations by predicted resource strain; prioritizing voter outreach by modeled likelihood of turnout; scoring trustworthiness of generative AI outputs in election communications.

3. **GenAI models**

   Description: Generative models are trained to produce new content (text, images, audio, video, or code) by learning the patterns and structure of existing data and then sampling from those patterns to create plausible outputs (e.g. ChatGPT). These systems are not programmed with explicit rules; instead, they infer likely responses based on the data they were trained on and the context of the user's prompt.

   Significance: Generative AI represents a steep change in accessibility and productivity. Users can draft messages, summarize documents, generate images, or create training materials via a prompting interface. However, outputs can be frequently incomplete, outdated, or incorrect; human review is essential.

   Examples: Drafting voter education messages in plain language; creating staff training manual; generating responses to FAQs through a secure chat interface; producing sample social media posts for accessibility review. In internal settings, GenAI can assist with spreadsheet formulas or formatting guidance.

Together, these three model families describe the main ways algorithmic systems operate within the broader AI landscape. Rules-based systems automate fixed logic; predictive and ranking models use data to forecast or prioritize; and generative models create new content. While the first two categories are already embedded in many existing applications and analytic tools, GenAI represents the most recent and most visible development, and the one most directly encountered by election officials through commercial platforms. Without completely disregarding the first two categories, we intentionally focus our attention at how generative AI works, its potential benefits and limitations in the context of election administration, and the safeguards needed for its responsible use.

2.3 How Generative AI Works

GenAI systems are trained to produce original content by identifying and recreating patterns found in large datasets. These models do not execute fixed or "pre-programmed" rules; instead, they *generate* new outputs that resemble the material they were trained on. The most familiar examples of GenAI are Large Language Models (LLMs) such as ChatGPT, Claude, Gemini, and Copilot, which produce text in response to user prompts. Although the terms "AI" or "AI-powered" are used broadly in public discussions, these GenAI systems represent a very specific subset; these are tools that *simulate* reasoning by predicting what words or phrases are likely to appear next in a sequence.

GenAI models are trained on vast collections of text, code, and other media to learn the statistical relationships between words, sentences, and concepts. Through this training process, the model estimates a probability distribution: given a certain input, what sequence of words is most likely to follow?[17] When a user types a prompt such as *"Explain how absentee ballots are counted"*, the model does not search the internet for the answer or recall a stored document. Instead, it produces an original response by drawing on its learned patterns of how similar questions have been phrased and answered across its training data. The result is an output that *sounds* coherent and informative but is, at its core, a probabilistic reconstruction rather than a factual retrieval.

To illustrate, imagine the model as a sophisticated autocomplete system that pre-

[17] It is important to note here that, technically, LLMs are not trained on "words," but "tokens." Tokens are the fundamental units of text that a model uses to process and generate human language (e.g. the word `[running]` contains the token `[run]` and the common ending token `[ning]`. For a definitive guide on tokenization see Jurafsky and Martin (2025).

dicts each next word based on everything it has previously written. If the prompt is detailed and contextually grounded, the generated text will often be accurate and well-structured. If the prompt is vague or refers to recent or highly specific information, the model may "hallucinate," meaning that it is likely to produce plausible but incorrect or fabricated statements. This behavior is not a malfunction, but an artifact of how generative models function: they optimize for linguistic plausibility, not factual accuracy.

The advantage of modern LLMs lies in their extensive training.[18] The most popular models train on a vastly heterogeneous mixture of data, consisting of: web text; books and literary corpora; encyclopedic content (e.g. Wikipedia); news and journalistic content; public code repositories; annotated instruction/dialog datasets; and, in many deployed settings, disclosed user data used in later fine-tuning stages. Beyond simple text (i.e. LLMs) GenAI models also ingest multimodal data (images, audio, etc.) for vision or speech capabilities. The exact composition and proprietary inputs are seldom fully disclosed, as are the parameters that shape their final model behavior.[19]

Without delving deep into the technical architecture of these systems, it is useful to understand what happens during the training phase. Put simply, an LLM is exposed to enormous volumes of text and constructs an internal multidimensional "map" of terms that are linguistically associated and frequently appear together;[20] as such, it identifies the deep linguistic structural rules and statistical patterns. Once trained, the model can generate new combinations of words that follow those same linguistic patterns, producing responses that appear "natural."[21] The quality of the model's output depends on both the diversity of the material it was trained on and the rigor of the alignment and fine-tuning steps that follow. For all these reasons, LLMs perform best when trained in English, given the amount of available training data; they usually underperform and are susceptible to degradation artifacts when trained in languages that exhibit less preponderant usage.

[18] For context, current state-of-the-art LLMs are trained on trillions of tokens of text. While official numbers are not publicly disclosed, Llama 3 reportedly used over 15 trillion tokens and estimates for GPT-5 are in the ballpark of 50 trillion.

[19] This poses a massive regulatory gap, as information pertaining to trust and safety cannot be independently audited and verified.

[20] For example, the word `[dog]` is expected to be in closer conceptual proximity to the words `[leash]` and `[walk]` than the word `[cat]`.

[21] The entire process is akin to a human learning another language: we are first acquainted with the formal linguistic rules, before we are able to make associations; however, it takes time and a large volume of exposure to really learn a language (i.e. "training") or to become as proficient as a native speaker.

To further contextualize the discussion above, we would like to address a common misconception: GenAI models can *neither reason nor understand* in the human sense. They do not possess comprehension or intent; they simply reproduce the distribution of the data upon which they were trained. They may appear intelligent, but at a foundational level they are just optimized mathematical formulas. In essence, all GenAI models (and LLMs in particular) do is to identify how words correlate in the training data and, based on those patterns, produce output that *simulates* those data. If, for example, we trained an LLM exclusively on the writings of Friedrich Nietzsche, the model would respond to prompts in a style and tone indistinguishable from the German philosopher. This would happen not because it understands Nietzsche's philosophy, but because it has learned to mimic the linguistic patterns of his work.

It should be noted that the underlying models are not directly retrainable by end users. End users often erroneously interpret the ability of an LLM to retain a history of prior information as a modification of its core behavior or parameters. In reality, a model just personalizes its output to match a user's perceived biases; this is commonly referred to as *user bias*. LLMs essentially create several individual "instances" per user, whereby they adopt a specific lens that would be familiar with – and welcoming to – the user's opinions, without altering their underlying mechanism. This creates the impression of a model being agreeable to one's views, further undermining the user's capacity to retrieve and assess accurate information.

2.4 Why This Matters for Election Administration

Understanding these dynamics is essential for election officials, who are often responsible for providing accurate and legally grounded information. Generative AI systems should not be expected to reason about facts or verify their correctness. Having no internal concept of truth, they may produce responses that appear confident and coherent, but in reality they are inaccurate or completely fabricated.

These GenAI shortcomings become especially relevant in domains that rely on precise terminology and procedural nuance, such as election administration. In this domain, subtle distinctions can carry both legal significance and institutional implications. This context differs sharply from the general purpose language that dominates most GenAI training data, which may cause such models to introduce errors that are difficult to detect without expert review. They should therefore be treated as *drafting* or *support* tools, not as intelligent collaborators. Understanding their capabilities and limitations is essential for any public use of GenAI, where representational fairness and neutrality are not abstract ideals but operational requirements.

At the same time, the capacity of GenAI to easily produce generically fluent text (or pleasing visuals) offers meaningful benefits when used under human supervision, some of which we explore in the next section. These systems excel in communication and outreach tasks: drafting social media messages, simplifying technical text into plain language, designing infographics, or producing accessible formats for voters with disabilities. These tasks can be performed relatively autonomously, provided a human has verified the input prompt and confirmed the validity of the output. When used as an unintelligent assistant, rather than an authoritative source, and under clear review and transparency protocols, GenAI can substantially reduce overhead and improve workflow efficiency without compromising accuracy or public trust.

2.5 Opportunities

Although GenAI introduces new oversight needs, it also presents practical opportunities to significantly improve efficiency and resilience in election administration. Election officials can expand their capacity in small and under-resourced offices, promote internal consistency, enhance the clarity of their messaging, and maximize the reach of voter communications. When paired with structured protocols of human-in-the-loop (HiL), GenAI becomes not a replacement for expertise, but a mechanism to extend it. Below we offer an epigrammatic list of areas where GenAI systems could and should be utilized under certain explicit oversight frameworks, which we detail in Section 4.

[**Efficiency**] Election administration involves a high volume of routine communications, notices, and updates. Generative systems can be assigned to the tasks of drafting initial versions of voter reminders, press releases, internal memos, or important reminders. They can be explicitly asked to follow approved formats, allowing staff to focus on verification and adaptation, rather than composition. They can also generate multiple format variants in plain text, alongside short versions appropriate for social media or long summaries; all within a single prompt (or instanced thread). This could assist staff in meeting administrative and communication demands within compressed timelines.

[**Accessibility**] Plain language rewriting and multilingual translation are among the most immediate areas where GenAI can expand voter access; it just so happens to be one of the areas where most election officials are under tremendous resource strain. GenAI can generate preliminary versions of alternative text for images, caption videos and soundbites, or simplify summaries that complement standard election materials in several languages simultaneously. These applications can narrow the resource gap between large and small jurisdictions, particularly where language access or accessibility requirements exceed staff capacity. Their effectiveness, however, depends on pairing generative outputs with human review to ensure cultural and linguistic accuracy.

[**Consistency**] Maintaining consistency across ballots, voter guides, websites, and social media is a persistent administrative challenge. Generative tools can help maintain consistency in phrasing and terminology across these materials and identify discrepancies that might confuse voters. If assisted by customized and airgapped retrieval-augmented generation (RAG) pipelines, they can reference only approved source documents, ensuring that new drafts remain aligned with verified content. In this role, GenAI can serve as an auxiliary system for internal coherence rather than an autonomous communication channel.

[**Knowledge Retention**] High staff turnover and frequent procedural updates make institutional memory difficult to sustain. GenAI tools can extract key points from lengthy manuals, summarize prior reports, and generate guidelines tailored to specific roles (e.g. provisional ballot processing versus poll worker) that simplify training and onboarding. GenAI can transform unstructured information into modular and reusable content, thereby supporting continuity between election cycles. They are particularly valuable for jurisdictions that rely on a large number of seasonal or temporary staff, where rapid knowledge transfer is essential.

[**Rapid Response**] During events that demand immediate and accurate communication, such as misinformation surges or natural emergencies, generative systems can assist in preparing initial statements consistent with official guidance, provided they have been explicitly instructed to do so. When time is of the essence, GenAI can accelerate drafting and versioning communications, while preserving tone and accuracy. As always, their usefulness lies in their ability to shorten response time without displacing staff judgment or oversight.

Taken together, these opportunities illustrate how GenAI, when properly constrained and supervised, can strengthen election operations by amplifying human capacity rather than substituting for it. The following section examines the limitations and risks that accompany these benefits, and the conditions under which responsible deployment is possible.

2.6 Limitations

Generative AI systems, while powerful, introduce distinctive risks that election administrators must understand before integrating them into public communication or internal workflows. These limitations primarily stem from how the models are trained and the data they rely on. Recognizing these structural weaknesses is a prerequisite for responsible adoption.

[**Hallucination and Drift**] A key technical challenge for the adoption of GenAI applications by election officials lies in the combination of two distinct, yet often conflated, risks that are structurally embedded in all such models: *hallucination* and *drift*. Hallucination refers to the generation of seemingly coherent (or visually consistent), but ultimately factually incorrect content. Drift, however, has a dual meaning and can be manifested as both: (a) concept drift, which is most often associated with temporal changes that result in a degradation of output validity over time;[22] and (b) interaction drift, where sustained and repeated interactions with a model can result in the latter adopting a warped frame and deliberately producing erroneous information.[23]

While all large language models are prone to varying degrees of both hallucination and drift, taken together these two risks tend to become more profound in specialized domains.[24] Mitigation strategies do exist, but their implementation requires deep technical knowledge and unfettered access to the root modeling pipelines. Such access is, as expected, seldom provided to individuals outside the corporate structure and, even if exceptions were to be made for election officials, the technical expertise required to effectively remedy both hallucination and drift would most likely fall outside of their knowledge domains.

In the context of election administration, the risks of hallucination and drift are particularly relevant. GenAI outputs should be grounded in authoritative knowledge bases, such as statutes, directives, court orders, and official guidance with explicit citations. In all such instances, hallucinations threaten the potential validity and reliability of generated outputs.[25] Similarly, election domains are ex-

[22] Augenstein et al. (2023), Bang et al. (2023), Gama et al. (2014), Ji et al. (2023), Lu et al. (2019), and Webb et al. (2016).

[23] Interaction drift is also referred to as interaction bias (Chen et al. 2025; Hu et al. 2025; Peters and Chin-Yee 2025; Rzadeczka et al. 2025).

[24] Although, Claude 3 and Gemini 1.5 are generally understood to be tuned for more empathetic engagement, which makes them more susceptible to interaction drift, ChatGPT 5 is arguably more prone to concept drift. Llama 3 and similar open-source models carry the greatest risk of exhibiting both hallucination and drift.

[25] Example: a GenAI model produces an authoritative summary of a new legislative update. The

tremely susceptible both to concept and interaction drift. With concept drift, evolving administrative practices (e.g., statutory revisions, court rulings, polling place changes, and deadline updates, to mention just a few) change the relevant knowledge base and may result in outdated information.[26] With interaction drift, however, a user faces the risk of obtaining what could be classified as purposeful misinformation due to the model behaving as a reinforcement agent.[27]

To add an additional layer of complexity, model performance can degrade at critical moments and without detection when vendors perform upstream updates, thus imperceptibely shifting behavior across product versions. To enhance factual accuracy and institutional accountability and to address hallucination and drift from a workflow perspective, election offices are called to subject outputs to structured legal and editorial review, rigorously interrogate responses, maintain dated records of provenance, define revalidation intervals, and route ambiguous queries to human experts. These measures may address accuracy and temporal currency jointly, but impose significant externalized costs on election officials' already constrained resources.

[**Bias and Representation**] As discussed in Section 2.3, all machine learning models, including GenAI, learn from existing data (i.e. past historical examples) rather than from neutral or idealistic representations of the world. Simply put, all GenAI models can produce as outputs are a reflection – in fact, an amplification – of the biases embedded in the inputs during the training phase.[28] This means they absorb not only linguistic patterns and associations, but also the social hierarchies embedded in those data. Bias in this context can arise at several levels. *Selection bias* reflects which data were included in the training data (e.g. whose voices, experiences, and perspectives were over- or underrepresented). *Algorithmic bias* emerges from how optimization processes translate those inputs into weighted

output seems legitimate and the model provides a source; however, upon close inspection the legislative update never happened and the source does not exist.

[26]Example: a GenAI model summarizes recent election laws. Although new regulations have emerged, the model continues to rely on outdated legal frameworks from its training data. As a result, its summary no longer reflects current law.

[27]Example: a GenAI model begins to mirror a user's assumptions and framing after extended interaction. Over time, the model learns what the user expects to see and modifies its subsequent output to align with that perceived expectation.

[28]The canonical example of bias propagation involves an internal Amazon model that was trained to shortlist applicants for engineering positions. Since the model was trained on the résumés of existing engineers, who happened to be primarily male, the outcome was unintentionally discriminatory against all female and nonbinary applicants; in effect, the model outright rejected such applications, since they did not conform to the training data of "successful" applications it had been exposed to (see Raghavan et al. 2020; cf. An et al. 2025; Liang et al. 2021; Navigli, Conia, and Ross 2023; Omiye et al. 2023; Resnik 2025).

parameters, promoting correlations that appear statistically dominant even if they encode stereotypes or structural inequities. Finally, *representation bias* affects how the model expresses ideas: which analogies it favors, which metaphors it defaults to, and which cultural reference points it treats as "neutral."

Most commercial generative models are trained on vast, largely Western datasets (in English) compiled from public accessible sources. As such, they tend to reproduce the dominant perspectives and norms most common in those Western environments and place them in the historical contexts from which the data were derived.[29] For example, occupational associations that link "nurse" with feminine pronouns or "engineer" with masculine ones persist in model predictions even after extensive fine-tuning. Similarly, geographic bias may lead a model to interpret "polling place access" primarily through an urban lens, or to misrepresent the terminology used by non-U.S. jurisdictions. In election administration, such biases can have real consequences: a multilingual translation might use a culturally unfamiliar term for absentee voting; a GenAI model might prioritize widely covered issues from large states over concerns relevant to small or rural counties simply because the former dominate the latter in the training data; or an image generator might fail to represent the diversity of voters and poll workers that actually characterize American elections.

Developers have recognized these dynamics and have identified several strategies to mitigate bias, all with certain caveats. One popular approach is *data balancing*, in which underrepresented categories are deliberately oversampled or synthetically expanded to counteract skew. Another is *debiasing during training*, where model weights are adjusted to neutralize statistically learned associations (for instance, reducing the correlation between gendered terms and specific professions). A more recent method involves what some researchers call *opposite-bias prompting*, which involves injecting counterexamples or adversarial data designed to reverse known tendencies. While these approaches can reduce visible bias in benchmark tests, they introduce their own downsides. First, artificially inverting or suppressing patterns may create the appearance of neutrality without improving fairness, but results in inconsistent behavior when models generalize to new contexts. Second, *bias correction is in itself a new form of bias*, reflecting the normative choices of those who define "balance." The reality is that biases in modern LLMs are still present, albeit much more subtle, and can appear within any context and without warning. Despite conscious and persistent efforts by researchers to remedy it, bias propagation in GenAI still remains mathematically unresolved.

[29] For instance, if the training data consists of writings and images of the early 20^{th} century, a GenAI model will propagate the exact same biases and stereotypes that existed at that time period.

For election officials, this means that bias is not an incidental flaw to be patched, but a structural feature of all data-driven systems. The relevant question is not whether a model is "biased" (they all are; bias is inescapable), but whether its training and deployment contexts align with the values of neutrality and equal access that govern public administration. When GenAI tools are used to generate voter information, translate materials, enhance communications, or create imagery, their outputs inevitably reflect the cultural assumptions and distributions inherent in their source data. Understanding these limitations is critical to interpreting what such systems can (and cannot) faithfully represent.

Aside from the technical aspects of bias, there lies a deeper ethical consideration we have yet to address: since we know that generative models encode structural and representational biases, continued use without appropriate mitigation effectively becomes an ethical choice. The legitimacy of any public-facing AI use depends not on the absence of bias, but on the transparency with which those biases are acknowledged and managed. In this respect, the duty of election officials parallels that of data stewards: to disclose limitations, mitigation attempts, document provenance, and ultimately to ensure that no group is systematically disadvantaged by inadvertent bias propagation.

Finally, a further challenge stems from the subtle and often imperceptible ways that modern LLMs reproduce bias. Rather than outputting overtly discriminatory statements, these systems may shape communication through tone and framing, thus nudging perceptions without overtly altering facts. When election officials rely on generative text for outreach or training, such biases can be unintentionally amplified, reinforcing assumptions that are statistically "common" in the data, albeit not necessarily fair or contextually appropriate. Awareness of this dynamic is essential for preserving neutrality and trust, especially in environments where even small linguistic or cultural asymmetries can affect public confidence and trust in institutions.

[Privacy and Data Sensitivity] Unlike many traditional software tools, generative AI systems rely on continuous data exchange between the user and remote servers. Prompts entered into public or consumer-grade GenAI interfaces are transmitted to a vendor's infrastructure for processing, where they are typically stored and used to improve future model versions. Even if this occurs under anonymized or aggregate conditions, the possibility that sensitive information could be inadvertently retained introduces substantial privacy and compliance concerns. For election administrators this risk is nontrivial, at their work routinely involves information that is highly sensitive, personally identifiable, legally protected, and politically relevant.

The most immediate danger arises from unintentional disclosure through prompts. A staff member seeking to generate a template might paste an internal communication or a voter list snippet into a GenAI tool, thus effectively exposing sensitive data to an external vendor. Since prompts and outputs are logged for quality control, such information could be disseminated outside the agency's custody and beyond the reach of open records policies or statutory retention limits. Similar risks arise when using third-party plugins or integrations whose data-handling practices differ from the platform's primary terms of service.[30] In many commercial offerings, the difference between "public" and "enterprise" accounts determines whether user data are recycled into future model updates, which is a distinction often overlooked by nontechnical staff.

A second, less visible but extremely critical, issue concerns *contextual privacy*. Even if no directly identifying information is shared, GenAI systems can infer or reconstruct sensitive details from surrounding cues. When multiple prompts reference small jurisdictions, unique job titles, or local events, model logs can yield composite profiles of individuals or offices. Indeed, one of the major advantages of modern GenAI models is their ability to profile their users and reconstruct the contexts that would allow them to personalize their future interactions. This is particularly problematic in election administration, where staff are often the targets of harassment and misinformation campaigns. Protecting contextual privacy, therefore, requires assuming that any prompt text may be discoverable and linking this assumption to established information security practices.

From a governance perspective, privacy management cannot rely solely (or entirely) on vendor assurances. Contracts with vendors for customized products, if any, should clearly specify: (a) whether prompts or outputs will be stored or reused for model improvement; (b) where data are geographically processed and retained; (c) what deletion or audit mechanisms are available; and (d) whether subcontractors have any access to the data.[31] Equally important is clarity, on the election administration side, about what categories of information must never enter a generative system,—such as voter registration data, internal incident reports, or privileged correspondence. In this respect, privacy and security overlap: loss of

[30]For example, an election office using a document-analysis plug-in within a GenAI platform to summarize internal poll-worker rosters might unknowingly transmit that roster to an external developer's server rather than to the main vendor, bypassing existing security and records retention controls.

[31]Storage location is particularly important. Proper data storage is expensive and has to abide by a rigorous legal framework. This added cost incentivizes vendors to avoid more strict data privacy jurisdictions, such as the EU. Indeed, many commercial LLMs have their user data stored in third countries, outside their homebase, which affords them the opportunity to evade legal prosecution if a data leak occurs.

informational control can undermine both compliance and trust.

That being said, a more subtle form of privacy erosion occurs through *latent memorization*. Because large language models are trained on massive text corpora, they occasionally memorize rare text strings, which can later reappear verbatim in generated content. This phenomenon has been documented even in state-of-the-art models.[32] Although the likelihood of encountering such memorized content in typical election use cases is low, the existence of this vulnerability underscores a broader point: once data are used for model training, they are effectively irretrievable and undeletable. This permanence distinguishes GenAI from traditional IT systems and magnifies the ethical weight of pre-deployment privacy due diligence.

At its core, the privacy question surrounding GenAI is not purely technical—it is institutional. Election offices function as custodians of public trust and are therefore bound by higher standards of transparency and restraint. Using a generative tool that collects or retains user data without clear boundaries risks not only individual exposure but also reputational harm to the electoral process itself. The ethical posture, therefore, must be one of precaution: assume information leakage, so minimize disclosure and treat every generative interaction as a public record subject to scrutiny.

[**Security**] Outside privacy concerns, generative systems introduce distinctive security risks that differ from those of traditional software applications. Because GenAI models can accept and execute natural language instructions, they expand the range of potential exploits from technical intrusion to linguistic manipulation. Adversaries can craft prompts (known as *prompt injections*) that override guardrails, exfiltrate hidden information, or trigger behaviors such as generating false or misleading content. Given that these systems are connected to external tools and databases, these injections can escalate into operational compromise, allowing the model to access user data or modify its output beyond its intended scope.

The most common such threats involve *model poisoning* or *data pollution*, where malicious actors introduce corrupted or adversarial text into public information streams later scraped for model training. If election offices rely on generative systems that have been inadvertently trained on such contaminated data, outputs may be subtly skewed toward misleading narratives or hostile framing. This problem is compounded by the opacity of most commercial training pipelines, which makes independent verification of provenance practically impossible.

[32] See Carlini et al. (2021) for a detailed study of training data leakage in large language models.

Even when models themselves are secure, vulnerabilities can emerge through APIs and third-party plugins that share access tokens or user inputs. These auxiliary components are frequent targets for credential harvesting, backdoor system penetration (i.e. hacking), denial of service, or privilege escalation attacks. Such breaches carry significant reputational and legal consequences for election officials, whose internal information systems often interconnect voter registration and ballot tracking, even if no voter data are directly accessed or altered.

The most visible form of abuse, however, lies in the weaponization of generative tools for influence operations. Synthetic audio, video, and text can impersonate election officials, fabricate announcements, or erode public confidence in the legitimacy of results. Although these threats typically originate outside official systems, their existence reinforces the need for internal discipline: every authentic message should have a verifiable and traceable provenance so that fabricated material can be quickly disproven. From a risk-management standpoint, election offices should treat GenAI deployments as part of their broader cybersecurity threat matrix. Practical safeguards include: restricting tool use to sandboxed environments; disabling external plugins unless formally vetted; enforcing tiered and logged access; disclosing all AI-assisted communications; and conducting periodic "redteam" testing for potential prompt injection and data leakage.

The realization of GenAI's promise is ultimately contingent upon alignment with recognized governance and risk management frameworks.[33] These frameworks operationalize trust by specifying HiL oversight, uncertainty communication, post-adoption evaluation, version control for models and prompts, and adaptation planning mechanisms for updates. Such practices map directly onto the election administration ecosystem. It should be noted, that this emphasis on best practices is not unique to the American experience. Internationally, the European Union's AI Act establishes obligations for general purpose generative systems that serve as a salient reference point for vendor standards and transnational best practices.[34] Taken together, these controls translate normative commitments (i.e. transparency and accountability) into applicable procedures that are testable, measurable, reviewable, and - most importantly - auditable.

The ideal purpose of GenAI in election administration should be best understood not as a substitution for human expertise or legal judgment, but as an augmentation of existing processes. When properly governed, GenAI can accelerate the

[33] The NIST AI Risk Management Framework (AI RMF 1.0) and its Generative AI Profile provide guidance on documentation, provenance, testing, and continuous monitoring (NIST 2023, 2024).

[34] See Artificial Intelligence Act, Regulation (EU) 2024/1689 (2024)

messaging and communications, allow for consistency and accessibility, provide structured workflows, and reduce capacity overload for election officials. To realize this promise, however, we call for robust oversight, layered verification, commitment to transparency, and adherence to established regulatory standards. The satisfaction of these conditions would enable GenAI experimentation, while simultaneously preserving the reliability and impartiality of election officials.

3 AI and GenAI in Election Administration: Practical Use Cases

Here we offer relevant use case examples for election officials, translating the technical capabilities of advanced computational models into concrete, field-relevant applications. We start by distinguishing between traditional AI tools that reorganize existing content through pattern matching and GenAI tools that create original content within contextual parameterization. The focus here is practical: where and how can election offices safely deploy these tools to expand capacity and/or improve communications without compromising public trust. Each example below is drawn from real or emerging uses across U.S. jurisdictions and assumes explicit human verification at every stage of implementation. Before turning to specific applications, a guiding principle is worth emphasizing. GenAI should *augment*, not replace, human expertise. It performs best when used to generate first drafts or templates that staff can refine and approve. The goal is to accelerate repetitive or text-intensive tasks, not to delegate legal interpretation or policy judgment to an algorithm.

3.1 Communications and Voter Outreach

[**Social Media and Digital Communications**] Election officials increasingly rely on social media to reach voters quickly, either to correct misinformation or to provide updates. GenAI can streamline the creation of clear and accessible content across platforms. When used responsibly, these systems can reduce the time and effort required to produce and adapt materials for multiple channels.

Several GenAI applications are currently suitable for this purpose. *ChatGPT*, *Claude*, *Gemini*, and *MS Copilot* (both in their web and app versions) can be used to draft posts, emails, media releases, generate image captions, and produce alternative phrasings suitable for different platforms (e.g., a longer explanatory Facebook post versus a concise Twitter/X update). All these models excel at tone adjustment—ensuring messages remain neutral, factual, and within character limits. They are particularly useful when offices must rapidly localize or rephrase similar content for different jurisdictions or audiences.

For visual communication, *Canva* remains one of the most practical tools for election officials. It combines a controlled design environment with integrated AI features such as "Magic Design," "Magic Edit," and "Text to Image." These functions allow staff to generate graphics, remove or replace image elements, and automatically adjust layouts to match specific platform dimensions. Because Canva templates can be locked to official branding, colors, and logos, it offers greater control and compliance consistency than open-ended generic image generators. Canva's AI components never publish autonomously, making it an ideal low-risk entry point for offices with limited design staff.

By contrast, text-to-image generators embedded in systems such as ChatGPT or Copilot Designer provide more creative flexibility, allowing entirely new imagery from text prompts; however, these carry higher risk of visual inaccuracy or the inclusion of unintended symbols. These models often struggle with fine details (e.g. seals, signage, ballots, embedded text etc.) that could carry legal or reputational implications, if misrepresented. For that reason, officials should use open-ended image generation primarily for internal brainstorming or concept development, not for direct publication.

A balanced workflow typically involves: (1) drafting a message text with a GenAI assistant (e.g. ChatGPT or Copilot); (2) producing visuals within a controlled platform, such as Canva; (3) manually verifying accessibility features and factual correctness; and (4) obtaining supervisory approval before posting. When used thoughtfully, GenAI can save considerable time and help maintain a robust online presence. In smaller offices, such efficiencies can feel like adding another communications staffer without increasing headcount.

[**Chatbots and Voter Queries**] Chatbots[35] can substantially reduce the time voters spend searching for information on election websites. Properly configured, they can answer routine questions about registration, polling locations, ballot tracking, or election dates, allowing staff to focus on complex or exceptional cases. They are relatively easy to implement, cost-effective, and can minimize the risk of human error in providing routine information, provided their scope is clearly

[35] Reminder that a "chatbot" is just the UI; the engine powering a chatbot can range from fixed rules-based responses to GenAI.

defined and their content derives from verified official sources.

Most chatbots currently used in election administration are *rules-based*, meaning they operate on pre-programmed decision trees built around "if–then" logic. For example, if a voter asks "Where is my polling place?", the system matches that phrase to a defined response or directs the user to a specific page. Because these chatbots rely on keyword recognition, they are predictable and transparent in how they generate answers. Their reliability, however, depends on how comprehensively their decision trees are coded. For instance, California's Secretary of State's chatbot *Sam* will successfully respond to "What voting options are there?" but not to "How can I vote?", which returns an error instead. This illustrates a central limitation of rules-based systems: unless every possible phrasing is anticipated, they will fail to answer or produce error messages.

These constraints often result in user frustration, which indirectly raises the costs of voting for interested citizens and the operating costs for election offices. Unless a chatbot has a clear escalation path, such as offering a "Would you like to speak with an election official?" option that transfers the voter to live assistance, rules-based systems will inevitably be of limited use. Balancing accuracy and user satisfaction is therefore a critical part of chatbot design. Rules-based chatbots perform best when used for narrow, well-defined tasks, such as locating polling places based on zip codes or addresses, confirming registration status, or linking to FAQs.

By contrast, *generative AI chatbots* (or GenAI chatbots) are designed to move beyond keyword matching and instead interpret the intent behind a voter's query. Systems built on LLMs, such as *ChatGPT*, *Gemini*, or *Copilot*, can hold short conversational exchanges and produce answers to complex or variably phrased questions. For example, a GenAI chatbot could recognize that "Where do I go to vote early?" and "When can I drop off my ballot?" both relate to early voting options, even if phrased differently. This adaptability can lead to better user experience and accessibility, especially for voters unfamiliar with official terminology.

However, GenAI chatbots introduce significant oversight and accuracy challenges.[36] Since they generate language statistically rather than retrieving fixed text, they can produce plausible but incorrect answers, or respond inconsistently to the same question across sessions. This unpredictability poses a risk to public trust if voters receive conflicting or outdated information. Moreover, these systems often require more computational resources, ongoing monitoring, and, if connected to the open web, carry privacy and data retention concerns.

[36] For a detailed discussion of GenAI limitations see Section 2.6.

The safest and most practical design for election offices is to use *retrieval-augmented chatbots* that generate answers only from pre-approved sources, such as the office's website or designated document repositories. Tools like *Microsoft Copilot Studio*, *Google Dialogflow*, or open-source frameworks like *Rasa* allow for such controlled deployments. In these configurations, the chatbot does not "make up" information; it only reformulates text that already exists in official records, providing a balance between accessibility and reliability.

[**Multilingual Content Generation and Translation**] Many jurisdictions are required by Section 203 of the Voting Rights Act to provide voter information in multiple languages. Even in areas without a formal legal mandate, multilingual communication is an essential part of accessibility and public trust. Generative AI tools can assist by producing first-draft translations of communications, notices, FAQs, and public statements across several languages, thus maintaining compliance with the law and reducing the time needed to reach diverse communities.

Traditional translation tools, such as Google Translate, rely on word- and phrase-level statistical alignment between language pairs. These tools are free and relatively fast, but often struggle with the nuances of legal or procedural language where a single term can have binding significance. For example, a literal translation of "provisional ballot" or "curing deadline" can easily yield inaccurate or misleading results if the translator fails to understand the underlying election process. Tools like Google Translate are also less sensitive to tone, formality, and cultural variation, which can make voter materials sound either too casual or too bureaucratic for their intended audiences.

By contrast, GenAI-based translation systems (including ChatGPT, Gemini, and Copilot) process meaning and context before generating output. Instead of substituting words directly, they reconstruct the sentence in the target language according to its intended tone and function. If prompted properly, these models can explain or justify specific translation choices and offer alternatives for ambiguous terms, or even adapt tone depending on audience (e.g., plain language for voters vs. formal phrasing for poll worker guidance). This makes them especially useful for drafting multilingual versions of outreach posts and training documents. Dedicated AI translation tools such as *ModernMT* and *DeepL* provide similar contextual translation capabilities, often with enterprise-grade privacy settings. *ModernMT*, for instance, learns from previous translations within the same organization, ensuring consistency across election cycles and documents.[37]

For smaller jurisdictions without in-house translation staff, GenAI translation can

[37] See "Translation and Localization" Prompt Library, (AI EMB Accelerator 2025).

help narrow resource gaps and promote equity in access to information. Given that many GenAI tools can detect idiomatic mismatches or flag potentially ambiguous terms before human review, the likelihood of overlooked errors is greatly reduced. However, these translations must always be reviewed by fluent speakers familiar with election law and local terminology. Even high quality GenAI outputs can misinterpret legal phrasing and oversimplify culturally specific expressions. A mistranslated instruction (e.g. one that alters a ballot curing deadline or misstates ID requirements) could inadvertently disenfranchise voters. For this reason, every translation should be subjected to expert human verification, ideally through side-by-side review. As a sidenote, translation should only occur in secure, enterprise environments with clear data privacy guardrails.

3.2 Document Processing and Internal Operations

[**Summarization**] Election officials routinely work with lengthy documents: legislation, court rulings, guidance from state agencies, technical manuals, or vendor contracts. Such documentation must be read, interpreted, and communicated accurately. Generative AI can help condense these materials into digestible summaries, making it easier for staff to identify key changes or compliance requirements.

Several GenAI tools are currently suitable for document summarization. *Microsoft Copilot*, *ChatGPT*, *Claude*, and *Gemini* all perform well when tasked with extracting main points, simplifying legal text into plain language, or preparing draft briefings. Because *Copilot* integrates directly with Microsoft Word, Outlook, and Teams, it is especially practical for election offices that already operate within the Microsoft ecosystem. Staff can upload or open a file and simply prompt: "Summarize this document for internal review and highlight any deadlines or new procedural requirements." The resulting output can be generated as a bullet-point summary or as a draft presentation, including slides and speaker notes.

Claude and *Gemini* offer flexible summarization depth, allowing users to request executive summaries, technical briefs, or detailed breakdowns, depending on audience needs. This feature is useful when a director wants a concise overview while legal or operations staff require more granular detail. However, because these systems do not natively export to PowerPoint, users must manually copy content into slides or rely on templates. *ChatGPT* can also generate presentation-ready content, including draft slide structures and talking points, though this often requires basic familiarity with formatting or script-based export options.

For smaller tasks or controlled visual summaries, non-generative tools such as *Beautiful.ai* or *Notion AI* can automatically convert existing documents into structured outlines or slide decks. These systems rely on pattern extraction rather than content generation, making them more predictable but less flexible. They can be useful for reformatting existing reports without introducing new text or interpretation.

In practice, a balanced workflow might look like this: (1) upload or paste the source text into a GenAI assistant (e.g. Copilot); (2) prompt the model to summarize key points in plain language, flagging any procedural or compliance details; (3) verify that the summary preserves statutory terminology and factual accuracy; and (4) reformat the approved text into a briefing memo or slide deck using either Copilot or Beautiful.ai. For long documents, it is advisable to divide them into sections before summarization to prevent timeouts or incomplete results.

While these tools can dramatically accelerate information processing, human oversight remains essential. Generative models may paraphrase in ways that alter meaning or omit exceptions critical to legal interpretation. For example, a model might simplify "ballots postmarked by Election Day" into "ballots received by Election Day," introducing a serious compliance error. Staff should therefore cross-check outputs against source text and verify all quoted language before circulation. This is particularly relevant in niche contexts where GenAI models have simply not been exposed to a critical volume of relevant training data, thus are likely to experience both hallucination and concept drift.

[**Meeting Minutes and Transcription**] Election officials spend much of their time participating in public meetings, briefings, and training sessions. Documenting those accurately and maintaining minutes or records can be both time-consuming and error-prone. GenAI transcription tools can assist by converting spoken discussions into written text, organizing meeting notes, combining key points, generating summaries of decisions, or scheduling follow-up meetings. Under proper supervision, GenAI applications can enhance transparency and record keeping, while freeing staff to focus on substantive tasks rather than manual note taking.

Several AI tools are currently suitable for transcription and minute preparation. *Otter.ai*, *Fireflies.ai*, *Zoom AI Companion*, *Google Meet Notes*, and *Microsoft Copilot for Teams* all provide reliable speech-to-text transcription, automatically labeling speakers and timestamps. These tools can then generate action items or short summaries, depending on the meeting's purpose. For instance, after a board of elections meeting, an official can ask: "Summarize the main topics discussed and list any motions or pending follow-up actions." The output can then be formatted into a draft set of minutes for staff verification. *Copilot for Teams* is particularly well-suited for jurisdictions already using Microsoft 365, since it integrates with Teams calendars, recordings, and file storage. Gemini is also a flexible tool since its integration with Gmail, given the latter's rising popularity among enterprise users.

ChatGPT and *Claude* can also assist with transcription summarization when provided with a transcript file, although their product integration is currently limited.[38] Officials can prompt the model to generate thematic summaries ("summarize all comments related to poll worker training") or to organize content by speaker, topic, or decision outcome. This approach can be useful when a meeting covers multiple subjects or when minutes must highlight specific compliance or budgetary issues, topics with which both *ChatGPT* and *Claude* have been understood to perform better than their alternatives in highly specific and technical tasks; counterintuitively, such better performance increases the likelihood of hallucination and drift.[39] For sensitive or confidential meetings such as those discussing security protocols, vendor performance, or personnel issues, transcription should only occur within secure and enterprise-grade environments. Uploading recordings or transcripts to public GenAI platforms (including ChatGPT or Gemini) is not recommended, as content may be retained for model improvement. In such cases, local transcription solutions such as *Microsoft Copilot for Teams* or paid versions of *Otter.ai* with end-to-end encryption offer safer alternatives.

A balanced workflow typically includes: (1) recording the meeting with a secure transcription tool such as Teams Copilot or Otter.ai; (2) exporting or reviewing the transcript to confirm speaker attributions; (3) using a GenAI assistant to produce a concise thematic or action-oriented summary; (4) verifying factual accuracy and ensuring all motions, votes, and deadlines are preserved verbatim; and (5) having staff or board members review and approve the final minutes. This ensures transparency while maintaining procedural integrity.

GenAI systems can dramatically efficiency and accessibility, providing near-instant records for public posting and enabling staff to search and cross-reference discussions. However, keep in mind that overreliance on AI-generated summaries also introduces risks. Some of the most innocuous such risks relate to conflating speakers, omitting dissenting views, or paraphrasing certain oral arguments without altering meaning. On the other side of the risk spectrum, however, they may subtly alter

38 This can arguably be a net positive since it requires a mandatory HiL step prior to the output being approved.

39 Hochmair, Juhász, and Kemp (2024), Nascimento et al. (2024), and Russell et al. (2025)

language that carries legal significance, retain/leak sensitive information, or even infer data without warning. Model performance aside, several external reasons should further constraint our trust of such outputs, particularly when their input consists of audio: recording quality, accents, overlapping speech, or similar verbal tone can easily confuse these models, resulting in inaccurate representations.

[**Poll Worker and Staff Training**] Training poll workers and election staff is among the most resource intensive aspects of election administration. The change of procedures and evolution of the legal frameworks require constant updating of training materials. In addition to having permanent staff frequently retrained, potentially large numbers of temporary staff must be onboarded quickly during each election cycle. GenAI can help develop concise training materials (e.g. manuals, quizzes, role-play exercises, and short videos) by converting existing procedural documents into accessible formats, thus freeing considerable amounts of time and resources to be allocated elsewhere.

Several GenAI applications are currently suited for developing and updating training materials. Tools such as *ChatGPT*, *Gemini*, *Claude*, and *Microsoft Copilot* can transform election manuals, legal statutes, or standard operating procedures into draft guides for specific roles. For example, staff can prompt: "Create a training module for provisional ballot clerks that explains the key steps for processing provisional ballots and includes three scenario-based quiz questions." The system can generate a first stepwise iterative draft, which staff then refine to ensure compliance with regulations and locally specific procedures.

More advanced use cases include scenario simulation and interactive decision trees. A GenAI tool can generate mock Election Day situations and suggest appropriate responses, such as a voter arriving at the wrong precinct, an equipment malfunction, a question about ID requirements, or an improperly sealed mail-in ballot. This allows offices to develop adaptive training exercises (i.e. "simulation exercises") that help poll workers practice their decision making in realistic contexts. Because such exercises are generated quickly, officials can easily refresh materials to reflect new laws or lessons learned from past elections.

For producing multimedia or e-learning content, GenAI can also be paired with AI-powered design and video tools. Platforms such as *Articulate Rise 360* and *Adobe Captivate* can convert GenAI text into interactive online training modules with embedded quizzes. Similarly, *Synthesia* and *Loom AI* can generate short explainer videos using virtual presenters, while *Canva* can assist with visual design and accessibility formatting. These combinations allow smaller offices to create professional quality materials without dedicated production teams.

A balanced workflow might include: (1) using a GenAI assistant (e.g. Claude or Copilot) to generate draft training text based on an existing manual; (2) refining that text for accuracy and procedural detail; (3) importing it into an e-learning platform like Articulate Rise for interactivity; and (4) final review by senior election staff or legal counsel before dissemination. This process enables offices to keep training materials current without starting from scratch each cycle.

Still, officials should exercise caution through every step of this chain. Generative tools may oversimplify complex legal requirements or unintentionally omit critical security or verification steps. They may also introduce stylistic inconsistencies across training materials, if multiple staff use different tools or prompting approaches. To mitigate these risks, it is essential that all GenAI materials undergo review by subject matter experts, and that any legal content, such as voter eligibility or chain-of-custody procedures, be validated against authoritative sources.

[**Proofing and Quality Control**] Accuracy and consistency are fundamental to public trust in elections. Voter guides and informational materials must be without errors, clearly written, and compliant with legal requirements. Even small inconsistencies (e.g. a misspelled candidate name or a mismatched deadline) can trigger confusion, ultimately resulting in the erosion of public trust. GenAI can assist election offices in the proofreading and quality control process by catching errors that humans might overlook during long production cycles, especially under time pressure.

Most GenAI tools perform well for this purpose and can be prompted to review documents for factual consistency and internal coherence. For instance, staff can upload or paste draft ballot instructions and ask: "Check that dates, deadlines, and the number of candidates are consistent throughout, and flag any ambiguous or contradictory phrasing." The model can then generate a checklist of issues or suggest clarifying revisions. *Copilot* in particular, which offers integration within Word and Outlook, can highlight formatting inconsistencies or deviations from office-specific templates, helping maintain professional presentation across applications.

Traditional AI proofreading tools, such as *Grammarly*, *PerfectIt*, and *Adobe Acrobat's AI features*, complement GenAI systems. The above specialize in detecting spelling, grammar, formatting, and accessibility issues using non-generative neural architectures.[40] For example, *PerfectIt* can enforce style-guide consistency ("mail ballot" versus "absentee ballot"), while *Adobe Acrobat* can scan PDFs to ensure

[40] Most proofreading and spelling tools are still based on RNNs instead of Transformer architectures, which makes them extremely lightweight.

that form fields are labeled, text does not overflow boxes, and documents comply with accessibility standards. These tools work through deterministic pattern matching, which makes them highly reliable for mechanical edits but less capable of assessing legal accuracy or logical flow.

A practical workflow might include: (1) running a draft through *PerfectIt* or *Grammarly* to catch issues on the surface; (2) uploading the revised version to *ChatGPT*, *Claude*, or *Copilot* for content review; (3) prompting the model to cross-check data points such as election dates and candidate lists against source materials; and (4) assigning final human verification for legal compliance and tone. In high-stakes materials like ballots or voter guides, this layered approach, which is technical first and generative second, creates redundancy and reduces the risk of both clerical and conceptual errors.

Although GenAI can enhance speed and consistency, it cannot guarantee correctness. Models may inadvertently "fix" text in ways that alter statutory meaning, or introduce plausible (but incorrect) suggestions. For instance, an AI might rewrite a phrase for readability and unintentionally change its legal effect ("must" to "may," or "Election Day" to "the day of voting"). Because of this, no AI output should ever be published without review by staff familiar with election law and office procedures. Proofing tools should also never be connected to public interfaces for draft or final ballots.

4 Best Practices Framework

4.1 Why Governance Matters

Generative AI can meaningfully enhance the capacity of election offices, but only when deployed within clearly defined guardrails that preserve institutional integrity. For election officials, the adoption of AI is not just a matter of innovation, but a question of governance. Each AI assisted communication, whether an internal draft or a public notice, carries the authority of the institution that issues it. If left unchecked, automation can inadvertently compromise the trust it seeks to reinforce.

Governance is the foundation that determines how these technologies are used, by whom, and for what purpose. It is less about technological sophistication than about clarity of responsibility. By embedding oversight and traceability into every stage of AI use from prompt creation to publication, offices can maintain the transparency and neutrality that define credible election administration.

Our recommendations for responsible use of GenAI are guided by the simple principles of human oversight and secondary verification. GenAI can be remarkably helpful in reducing administrative burden and accelerating communication, yet its convenience must never substitute for human judgment. HiL is not an optional step, but the mechanism that ensures accountability and protects institutional legitimacy. The framework that follows translates these principles into operational practice, outlining what ethical and responsible AI use should look like inside an election office.

4.2 Core Ethical Foundations for Responsible AI Use

Responsible integration of AI in election administration rests on a handful of core values that define the boundaries of legitimate use: accountability, transparency, fairness, privacy, and competence. These are not abstract ideals; they are operational commitments that translate directly into workflow design, staff training, operating procedures, and public communication. Without all the above being present, both institutional resilience and public trust can potentially be undermined.

[**Accountability**] Humans remain responsible for every AI-assisted output. Nothing produced by a generative model should circulate externally without a named human reviewer signing off on its output. Offices should keep a secure record of prompts and edits to preserve an audit trail. Accountability ensures that institutional legitimacy stays where it belongs: with humans, not optimized mathematical functions.

[**Transparency**] Election offices should disclose, whenever feasible, that AI tools were used in preparing content for public release. A short line at the bottom of a notice, such as "prepared with AI assistance and reviewed by staff," is sufficient to preserve trust. Internally, transparency means documenting decisions about prompts and data sources so that any output can be reconstructed and verified if challenged. When the process is visible, the product becomes more credible.

[**Fairness and Equity**] Generative systems mirror the biases in their training data, which makes review for neutrality essential. Offices should routinely check

AI-generated text and imagery for cultural sensitivity and representational balance. The goal is not to eliminate bias entirely, as this would be an impossible task; rather, the goal is to identify and mitigate bias wherever it might shape meaning or perception. In practice, fairness may involve testing multilingual translations with community reviewers or adjusting prompts to produce plainer, more inclusive phrasing.

[**Privacy and Security**] Confidential data (e.g. as voter information, internal communications, system credentials etc.) should never be entered into a publicly available AI platform. The same applies to sanitized data that could be reconstructed. Offices should rely on airgapped enterprise accounts that meet recognized cybersecurity standards and ensure that staff understand which data are prohibited from disclosure. AI deployment in election administration must begin from the presumption of privacy, not convenience.

[**Competence and Continuous Review**] Ethical governance depends on informed staff. Training programs should familiarize personnel with how generative systems work and where are their weaknesses situated. Staff should be trained on prompt engineering and log maintenance, in addition to data security/privacy. Because AI tools evolve quickly, offices should treat this as a continuing obligation involving reviewing model updates and vendor changes at regular intervals. Competence, in this context, means knowing both the capabilities of the tool and the limits of one's own reliance on it.

In short, these values define the ethical perimeter within which AI may responsibly operate in an electoral environment. They are the foundation for the procedural guardrails and verification practices that follow.

4.3 Operational Guardrails for GenAI Deployment

Ethical principles mean little without concrete procedures to make them work. Operational guardrails translate broad commitments into everyday practice. In election administration, this means setting boundaries on how generative systems are used, what kinds of content they may produce, and how outputs are verified before release. The purpose is not to slow down innovation but to make sure that any use of AI strengthens rather than dilutes institutional trust.

[**HiL/Oversight**] No generative system should ever operate unattended in a public-facing capacity. Every output, from a short calendar reminder draft to a translated notice, should pass through at least one qualified human reviewer before publication. Ideally, one staff member would initiate the prompt and another would verify the output for factual accuracy. This division of labor and responsibility prevents both overreliance on a single user and the gradual normalization of machine-authored content.[41] For small offices, HiL oversight can be as simple as a "two-person check" before anything is posted or mailed.

[**Prompt Design and Review**] Prompts are the invisible code that governs how GenAI behaves. A vague or poorly written prompt can yield misleading results, while a clear one can reduce risk substantially. Offices should maintain a small library of approved prompts; these should be plain, factual, and stripped of speculation so that staff can adopt (or adapt) them safely. Each prompt should include constraints, such as "cite only from official sources" or "do not interpret law." Staff should avoid open-ended or hypothetical instructions ("What if..." or "Imagine that...") unless used for internal brainstorming, in which case they should only be used is exclusive sandboxing modes, outside the official airgapped models. Periodic review of these prompt templates helps maintain consistency as policies (or models) evolve and ensures auditability.

[**Verification and Validation**] Verification begins once a model has produced an output, which then should be checked for accuracy against vetted information. For factual materials, the office should keep a short checklist: verify numbers, confirm dates, test hyperlinks, and inspect citations. Where retrieval-augmented systems are used, staff must confirm that the retrieved sources are legitimate and up to date. If an output cannot be fully verified, it should be discarded rather than revised through additional prompting. This conservative approach ensures that uncertain content never leaves the drafting stage.

[**Documentation and Auditability**] Every AI assisted action should leave a trace. Offices should record the model name and version, the date of use, the exact prompt, and the output before human editing. This documentation protects both the user and the institution: if a question arises about an announcement or posting, the record shows what the system produced and who approved it. Maintaining such logs can be as simple as saving text files in a shared folder or attaching them to an internal ticketing system. Over time, these archives become an institutional memory that supports both transparency and training.

[**Model Versioning and Change Tracking**] Commercial models change frequently, sometimes without notice. A new release can subtly alter tone or phrasing,

[41] This also prevents user bias from being perpetuated.

while major updates may affect factual tendencies and increase the risks of hallucination or drift. Election offices should treat each update as a new tool, not a continuation of the old one. Before resuming normal use, staff run a small suite of test prompts to detect unexpected changes in style or accuracy. If outputs shift in ways that could affect public communication, postpone deployment until staff have reviewed and adjusted workflows. Documenting which version was active during a given election cycle allows future reviewers to reproduce and understand past results.

[**Bias and Factuality Checks**] Even with human review, subtle bias can creep in through phrasing or imagery. When evaluating generative outputs, staff should ask two questions: "Is this accurate?" and "Who might be excluded or misrepresented by this wording or image?" For example, a seemingly neutral phrase like "voters should bring identification" may imply a legal requirement that does not apply universally. Similarly, generated images should reflect demographic diversity and accessibility features present in real polling environments. Offices should maintain a brief internal checklist for these reviews and evaluate whether misrepresentations may generate distortions before they propagate.

[**Data Privacy and Secure Handling**] Publicly available GenAI platforms should never be used to process sensitive materials. Election officials should rely only on locally hosted systems covered by appropriate data handling and legally binding agreements. Before adopting any vendor platform, confirm where data are stored and for how long they are retained. Sensitive or unpublished information, including voter records and internal correspondence, must remain outside all generative environments. When in doubt, treat every prompt as if it could one day become public or used for model refinement/training.

[**Incident Response and Escalation**] Despite precautions, errors or misuses will occasionally occur. Offices should define clear escalation steps for correcting and documenting them. If inaccurate or misleading content is released, the correction should be immediate and transparent, ideally through the same channel where the error appeared. Internally, the event should be logged with a short note explaining what happened, what model was used, and how the workflow will change to prevent recurrence. A predefined incident response protocol turns a potential crisis into a manageable and instructive moment.

4.4 Implementation Pathways for Election Offices

Adopting GenAI responsibly is less about acquiring new technology than about building a repeatable workflow that fits each office's capacity and risk tolerance. Election offices differ dramatically in size, staffing, technical depth, and resources. A small county with two full-time staff cannot implement the same oversight structure as a larger office with a dedicated communications team. Below we offer scalable entry points so that every jurisdiction can begin to benefit from GenAI without compromising accuracy or security.

[**Low-Capacity Offices**] Smaller jurisdictions can begin with tightly scoped, low-risk uses of GenAI. The safest entry points are drafting and formatting tasks that never leave internal circulation, such as creating checklists, rewriting notices into plain language, or summarizing meeting notes. Free or enterprise versions of tools like Microsoft CoPilot, ChatGPT Teams, or Gemini for Workspace can operate entirely within a closed environment. Outputs should always be reviewed by at least one additional staff member before external use. For visual design, Canva's template locking and brand-kit functions allow even small teams to generate compliant graphics without risking visual inconsistency. The goal at this stage is not automation but augmentation: freeing staff from repetitive writing tasks while preserving full human oversight.

[**Mid-Capacity Offices**] Mid-sized offices, typically with communications or IT support, can formalize their use through a documented workflow. This may include a prompt library with pre-approved instructions ("summarize only from official materials," "avoid speculative phrasing"), a two-stage review process, an approved database of relevant materials, and routine audits of outputs for accuracy and tone. Offices may also experiment with retrieval-augmented systems that reference only locally stored documents (state manuals, FAQs, public notices etc.)—so that generative outputs remain within verified boundaries. These offices can pilot internal staff positions, which could be responsible for testing new features and maintaining documentation after vendor changes. Transparency is achieved not through public disclosure alone but through internal clarity: everyone should know which tools are used, for what purpose, and under whose supervision.

[**High-Capacity Offices**] Larger jurisdictions (or even state-level agencies) can explore more advanced integrations such as sandboxed or locally hosted GenAI environments. These systems combine the efficiency of large models with the control of restricted data access. A typical configuration might include retrieval-augmented chatbots grounded in the state's election code or a custom Copilot instance confined to internal servers. Before deployment, each system should undergo security review and red-team testing to ensure that it cannot retrieve or generate unauthorized content. High-capacity offices are also well positioned to collaborate with academic or civic/tech partners to develop benchmark datasets for bias and accuracy evaluation—resources that can benefit smaller jurisdictions through shared learning. In this tier, the focus shifts from safe experimentation to institutionalization: embedding AI governance into procurement policies, staff training, and documentation standards.

[**Capacity Building and Continuous Review**] Regardless of size, every office should treat GenAI adoption as an ongoing process. Models evolve quickly, as do legal and procedural frameworks around elections. Establishing a recurring review cycle annually or after each major election ensures that practices stay current. Training sessions, tabletop exercises, and scenario testing can help staff remain confident, but critical, users of these tools. Partnerships with peer jurisdictions and professional associations are valuable sources of shared lessons and updated guidance. The aim is not to reach a final state of compliance, but to cultivate a sustained posture of responsible innovation.

4.5 Sustaining Trust

Public confidence is not earned by adopting new tools, but by showing that those tools are used responsibly. Every GenAI output, no matter how small, carries the weight of institutional authority once it leaves an election office. Maintaining that trust requires continuous oversight and a willingness to adapt when conditions change.

Transparency should not be reduced to a single disclosure line, such as: "this message was AI-assisted." It is a consistent practice of explaining how information is created, how it is reviewed, and who has approved it; it is about maintaining a chain of command that make information auditable and responsibility attributable. Officials can make this visible through small cues: consistent citation of official sources, dated version control, and staff oversight. Over time, these signals help the public understand that AI tools are being used under supervision, not in secret. In environments where misinformation spreads quickly, procedural openness is a stronger defense than any technical safeguard.

Trust depends on the capacity to detect and correct mistakes. Offices should establish simple verification loops that extend beyond initial review, spot-checks of published materials, monitoring public feedback, and rapid correction protocols when errors appear. When corrections are made transparently, credibility strengthens rather than weakens. The same logic applies internally: when staff discover inconsistencies in GenAI outputs, these should be logged and used to update prompt templates or workflows. Learning from past mistakes and understanding how they occurred in the first place is the operational form of integrity.

Continuous oversight is also a form of operational openness. Each election cycle, for example, brings forth new legal considerations. Building short debrief sessions into post-election reviews on what worked, what failed, and what needs clearer rules helps to maintain institutional learning. These reflections should feed into a living policy document that governs GenAI use. Over time, such iterative governance ensures that ethical commitments do not stagnate as technology evolves.

Overall, trust grows when election officials see themselves as part of a larger, interdependent ecosystem. From state agencies and local jurisdictions to professional associations, networks of shareholders should be encouraged to coordinate in sharing established best practices and tested procedures so that learning does not happen in isolation. Collaborative review mechanisms, joint audits, cross-jurisdictional exercises, and shared case repositories are effective ways to distribute both the labor and the responsibility of oversight. This networked form of governance ensures that the strength of the system does not depend on the resources of any single office.

That being said, the most durable form of trust is cultural rather than procedural. When staff view GenAI as a tool that complements, rather than one that replaces, professional judgment, they are more likely to question outputs and verify facts. Measured innovation means advancing cautiously, documenting thoroughly, and sharing lessons openly. In this culture, the credibility of election administration is not at odds with technological adoption; rather, it is the very condition that makes such adoption possible.

Ultimately, sustaining public trust in elections requires the same disciplines that have always defined effective administration: transparency and accountability. GenAI does not replace these values; it tests and reaffirms them. When election offices treat every generative system as an extension of their institutional responsibility—subject to review, documentation, and correction—they not only manage new technology effectively but also strengthen the democratic confidence on which their legitimacy rests.

5 Bibliography

Ada Lovelace Institute, AI Now Institute, and Open Government Partnership (Aug. 2021). *Algorithmic Accountability for the Public Sector*. Tech. rep. Ada Lovelace Institute; AI Now Institute; Open Government Partnership.

AI EMB Accelerator (2025). *Prompt Library for the AI EMB Accelerator*.

An, Jiafu, Difang Huang, Chen Lin, and Mingzhu Tai (Mar. 2025). "Measuring gender and racial biases in large language models: Intersectional evidence from automated resume evaluation". In: *PNAS Nexus* 4.3.

Artificial Intelligence Act, Regulation (EU) 2024/1689 (2024). *Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union. OJ L 2024/1689, 12 July 2024.

Augenstein, Isabelle, Timothy Baldwin, Meeyoung Cha, Tanmoy Chakraborty, Giovanni Luca Ciampaglia, David Corney, Renee DiResta, Emilio Ferrara, Scott Hale, Alon Halevy, Eduard Hovy, Heng Ji, Filippo Menczer, Ruben Miguez, Preslav Nakov, Dietram Scheufele, Shivam Sharma, and Giovanni Zagni (2023). *Factuality Challenges in the Era of Large Language Models*.

Bang, Yejin, Samuel Cahyawijaya, Nayeon Lee, Wenliang Dai, Dan Su, Bryan Wilie, Holy Lovenia, Ziwei Ji, Tiezheng Yu, Willy Chung, Quyet V. Do, Yan Xu, and Pascale Fung (2023). "A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity". In: *Proceedings of IJCNLP-AACL 2023*, pp. 675–708.

Brennan Center for Justice (2025). *Local Election Officials Survey*.

Carlini, Nicholas, Florian Tramer, Eric Wallace, Matthew Jagielski, Ariel Herbert-Voss, Katherine Lee, Adam Roberts, Tom Brown, Dawn Song, Ulfar Erlingsson, Alina Oprea, and Colin Raffel (2021). "Extracting Training Data from Large Language Models". In: *Proceedings of the 30th USENIX Security Symposium (USENIX Security 2021)*.

Chen, Yongpei, Sergey N. Kirshner, Anton S. Ovchinnikov, Manish Andiappan, and Trevor Jenkin (2025). "A Manager and an AI Walk Into a Bar: Does ChatGPT Make Biased Decisions Like We Do?" In: *Manufacturing & Service Operations Management* 27.2, pp. 354–368.

CISA (2024). *Risk in Focus: Generative AI and the 2024 Election Cycle*. Tech. rep. Cybersecurity and Infrastructure Security Agency (CISA).

Currie, Shannon, Ben Wilentz, and Nadrina Ebrahimi (2022). *Local Election Officials Survey*.

Ferrer, Joshua, Daniel M. Thompson, and Rachel Orey (2024). *Election Official Turnover Rates from 2000–2024*. Tech. rep. Bipartisan Policy Center.

Gama, João, Indrė Žliobaitė, Albert Bifet, Mykola Pechenizkiy, and Abdelhamid Bouchachia (2014). “A Survey on Concept Drift Adaptation”. In: *ACM Computing Surveys* 46.4, pp. 1–37.
Gao, Jianfeng, Michel Galley, and Lihong Li (2019). “Neural Approaches to Conversational AI”. In: *Foundations and Trends in Information Retrieval* 13.2–3, pp. 127–298.
Goldstein, Josh A., Girish Sastry, Micah Musser, Renee DiResta, Matthew Gentzel, and Katerina Sedova (2023). *Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations.* arXiv preprint: 2301.04246.
Grimmelikhuijsen, Stephan G. and Albert J. Meijer (2014). “Effects of Transparency on the Perceived Trustworthiness of a Government Organization: Evidence from an Online Experiment”. In: *Journal of Public Administration Research and Theory* 24.1, pp. 137–157.
Hancock, Jeffrey T., Mor Naaman, and Karen Levy (2020). “AI-Mediated Communication: Definition, Research Agenda, and Ethical Considerations”. In: *Journal of Computer-Mediated Communication* 25.1, pp. 89–100.
Hochmair, Hartwig H., Levente Juhász, and Takoda Kemp (Aug. 2024). “Correctness Comparison of ChatGPT-4, Gemini, Claude-3, and Copilot for Spatial Tasks”. In: *Transactions in GIS* 28.7, pp. 2219–2231.
Hu, Tiancheng, Yara Kyrychenko, Steve Rathje, Nigel Collier, Sander van der Linden, and Jon Roozenbeek (2025). “Generative language models exhibit social identity biases”. In: *Nature Computational Science* 5, pp. 65–75.
Jakesch, Maurice, Megan French, Xiao Ma, Jeffrey T. Hancock, and Mor Naaman (2019). “AI-Mediated Communication: How the Perception that Profile Text was Written by AI Affects Trustworthiness”. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* ACM, pp. 1–13.
Ji, Ziwei, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Yejin Bang, Delong Chen, Wenliang Dai, Ho Shu Chan, Andrea Madotto, and Pascale Fung (2023). “Survey of Hallucination in Natural Language Generation”. In: *ACM Computing Surveys* 55.12, pp. 1–38.
Jurafsky, Daniel and James H. Martin (2025). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, with Language Models.* 3rd. Online manuscript released August 24, 2025.
LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015). “Deep Learning”. In: *Nature* 521.7553, pp. 436–444.
Liang, Paul Pu, Chiyu Wu, Louis-Philippe Morency, and Ruslan Salakhutdinov (18–24 Jul 2021). “Towards Understanding and Mitigating Social Biases in Language Models”. In: *Proceedings of the 38th International Conference on Machine*

Learning. Ed. by Marina Meila and Tong Zhang. Vol. 139. Proceedings of Machine Learning Research. PMLR, pp. 6565–6576.

Licht, Jenny De Fine, Daniel Naurin, Peter Esaiasson, and Mikael Gilljam (2014). "When Does Transparency Generate Legitimacy? Experimenting on a Context-Bound Relationship". In: *Governance* 27.1, pp. 111–134.

Lim, Sue and Ralf Schmälzle (2024). "The Effect of Source Disclosure on Evaluation of AI-Generated Messages". In: *Computers in Human Behavior: Artificial Humans*.

Lu, Jie, Anjin Liu, Fan Dong, Feng Gu, João Gama, and Guangquan Zhang (2019). "Learning under Concept Drift: A Review". In: *IEEE Transactions on Knowledge and Data Engineering* 31.12, pp. 2346–2363.

Merivaki, Thessalia and Mara Suttmann-Lea (2022). "Designing Voter Education Across the States: State Responses to the Help America Vote Act". In: *Election Law Journal: Rules, Politics, and Policy* 21.1, pp. 46–59.

Mitchell, Tom M. (1997). *Machine Learning*. McGraw-Hill.

Nascimento, Nathalia, Everton Guimaraes, Sai Sanjna Chintakunta, and Santhosh Anitha Boominathan (2024). *LLM4DS: Evaluating Large Language Models for Data Science Code Generation*. arXiv preprint: 2411.11908.

Navigli, Roberto, Simone Conia, and Björn Ross (June 2023). "Biases in Large Language Models: Origins, Inventory, and Discussion". In: *J. Data and Information Quality* 15.2.

NIST (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. Tech. rep. NIST AI 100-1. National Institute of Standards and Technology.

NIST (2024). *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile*. Tech. rep. NIST AI 600-1. National Institute of Standards and Technology.

Omiye, Jesutofunmi A., Jenna C. Lester, Simon Spichak, Veronica Rotemberg, and Roxana Daneshjou (2023). "Large language models propagate race-based medicine". In: *npj Digital Medicine* 6, p. 195.

Peters, Uwe and Benjamin Chin-Yee (2025). "Generalization bias in large language model summarization of scientific research". In: *Royal Society Open Science* 12.4.

Posard, Marek N., Todd C. Helmus, Michelle Woods, and Bilva Chandra (2024). *The 2024 U.S. Election, Trust, and Technology*. Tech. rep. RAND Corporation.

Qin, Libo, Min Li, Wenbo Guo, Qiguang Chen, Carl Yang, Lizi Liao, Zidong Yang, and Yue Zhang (2023). "End-to-End Task-oriented Dialogue: A Survey of Datasets and Models". In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

Raghavan, Manish, Solon Barocas, Jon Kleinberg, and Karen Levy (2020). "Mitigating bias in algorithmic hiring: evaluating claims and practices". In: *Pro-*

ceedings of the 2020 Conference on Fairness, Accountability, and Transparency. Association for Computing Machinery, pp. 469–481.
Resnik, Philip (Sept. 2025). “Large Language Models Are Biased Because They Are Large Language Models”. In: *Computational Linguistics* 51.3, pp. 885–906.
Russell, Matthew, Aman Shah, Giles Blaney, Judith Amores, Mary Czerwinski, and Robert J. K. Jacob (2025). *Neural and Cognitive Impacts of AI: The Influence of Task Subjectivity on Human-LLM Collaboration.* arXiv preprint: 2506.04167.
Russell, Stuart and Peter Norvig (2021). *Artificial Intelligence: A Modern Approach.* 4th ed. Pearson.
Rzadeczka, Marcin, Anna Sterna, Julia Stolińska, Paulina Kaczyńska, and Marcin Moskalewicz (2025). “The Efficacy of Conversational AI in Rectifying the Theory-of-Mind and Autonomy Biases: Comparative Analysis”. In: *JMIR Mental Health* 12.
Selbst, Andrew D. (2021). “An Institutional View of Algorithmic Impact Assessments”. In: *Harvard Journal of Law & Technology* 35.1, pp. 117–191.
Suttmann-Lea, Mara and Thessalia Merivaki (2023). “The Impact of Voter Education on Voter Confidence: Evidence from the 2020 U.S. Presidential Election”. In: *Election Law Journal: Rules, Politics, and Policy* 22.2, pp. 145–165.
The Elections Group and Kathryn Peters (2024). *Generative Artificial Intelligence for Election Administrators: An Introduction.* White Paper. The Elections Group.
Tyler, Tom R. (2023). “Whither Legitimacy? Legal Authority in the Twenty-First Century”. In: *Annual Review of Law and Social Science* 19, pp. 1–17.
U.S. Census Bureau (2022). *Section 203 Language Determinations.*
U.S. Department of Justice (2014). *ADA Requirements: Effective Communication.*
U.S. Department of Justice (2024). *Section 203 of the Voting Rights Act: Language Minority Citizens.*
Veale, Michael, Max Van Kleek, and Reuben Binns (2018). “Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.*
Webb, Geoffrey I., Roy Hyde, Hong Cao, Hai Long Nguyen, and François Petitjean (2016). “Characterizing Concept Drift”. In: *Data Mining and Knowledge Discovery* 30.4, pp. 964–994.

Appendix

A. The AI-Powered Toolkit

Beyond rule-based chatbots, predictive models, and GenAI models, the following families frequently appear in products marketed as "AI" and are useful to distinguish.

Table 1: List of AI and GenAI Tools

System Type	What it is	Fit and Caveats
Knowledge-based systems	Logic/ontology rules, knowledge graphs, entity resolution	High traceability; good for eligibility checklists and structured FAQs; "brittle" if rules are incomplete
Classical supervised ML	Logistic regression, random forests, gradient boosting	Triage/routing, inbox prioritization; requires labeled data, monitoring for drift
Information extraction / document AI	OCR, layout parsing, NER, relation extraction	Turn PDFs and forms into structured data; human verification required for legal text
Classical IR (lexical search)	Keyword/BM25 retrieval	Strong for exact terms and citations; pair with filters for statutes and dates
Embedding / vector search	Semantic retrieval over text embeddings	Useful for RAG; must constrain to authoritative corpora and track corpus versions
Summarization (extractive vs. abstractive)	Extractive quotes vs. generative paraphrase	Prefer extractive for legal texts; reserve abstractive for drafts with review
Machine translation (NMT)	Neural MT engines	First-pass language access; always route through fluent/legal review

Table 1 (continued).

System Type	What it is	Fit and Caveats
Speech technologies	ASR (speech-to-text), TTS (text-to-speech)	Accessibility, hotlines, meeting captioning; verify accuracy on local terminology
Safety classifiers / content filters	Moderation and policy enforcement models	Reduce off-policy outputs; tune to election-specific categories
Anomaly / drift detection	Statistical/ML monitors for change	Monitor query mix, error rates, and corpus freshness; trigger reviews
Reinforcement learning	Policy tuning from feedback (RLHF/RLAIF)	Shapes assistant behavior; document versions and changes to preserve reproducibility
Agentic orchestration	Function-calling, planners, multi-tool pipelines	Powerful but complex; tightly permission tools and log every call for audit
Privacy-preserving ML	Redaction, differential privacy, federated learning	Protects PII/PHI; consider when training on sensitive local data
Generative media (diffusion, voice cloning)	Image/video/audio synthesis and editing	Outreach; contains adversarial risk; requires provenance and detection
Content authenticity / provenance	Cryptographic or metadata credentials	Attach source/provenance to official media; vendor support varies

Note. Many vendor offerings combine several of the above (e.g. vector search + LLM "answering"). Distinguishing retrieval (what sources are used), generation (what is composed), and governance (what is logged, reviewed, and versioned) clarifies evaluation and procurement.