Generative Artificial Intelligence for Election Administrators

AN INTRODUCTION

The Elections Group

Table of Contents

Generative Artificial Intelligence for Election Administrators: An Introduction

BY KATHRYN PETERS

The concept of an "artificial intelligence" has motivated science fiction authors and computer scientists for decades. Over that time, advances in computing power and methods for training complex systems created systems capable of beating top human chess players[1] and Jeopardy! champions[2], among other feats. The release of ChatGPT in November 2022 captured the public imagination once again with a demonstration of a seemingly general-purpose artificial intelligence.

This inflection point has sparked significant discussion and speculation. How will this latest advancement change our economy? Our democracy? What are the technological limits of these models? Could they seed future utopias or dystopias? Do they even work? Keeping up with the public discussion around this generation of tools, called "generative artificial intelligence" for their shared ability to create new material from their training data[3], is more than a full-time job.

> **Luckily, engaging constructively with new technology does not require a computer science degree or a crystal ball.**

Luckily, engaging constructively with new technology does not require a computer science degree or a crystal ball. Subject-matter experts in a wide range of fields play a critical role in shaping how generative AI will be used because they bring deep insights into the specific problems and processes where these tools might be deployed.

Good implementations of technology share two important features: most importantly, they solve real problems for the people who use them, and they do so in ways that are efficient, sustainable, ethical, and accessible—what we recognize as well-designed. Identifying real needs and opportunities where additional technology could support better election administration and a better voting experience is the first step to implementing generative AI well. While election officials cannot

1 Accomplished in 1997 by IBM's Deep Blue, a demonstration of the power of parallel processing.

2 As demonstrated in 2011 by IBM's Watson, an application of machine learning techniques (among other "question-answering system" components).

3 The distinction between "generative" AI and earlier demonstrations is not a bright line. An MIT News piece breaks down the differences and overlaps.

Kathryn Peters

Peters is a former technology researcher at CITAP and co-founder of Democracy Works. She is currently thinking about potential civic futures and writing those thoughts down here.

simultaneously hold expertise in every field that powers their work, they can—and should—actively help set the agenda for what new technologies are built and deployed in their field.

This guide is intended to give election administrators the tools they need to imagine whether and how generative tools might support their work. The first two sections include a practical introduction to some technical aspects of generative AI, a brief summary of the major capabilities and shortcomings of currently-available generative tools, and case studies in how other election and government officials are testing or using AI in their work. The third provides a framework for imagining potential uses and planning successful deployment of generative AI tools.

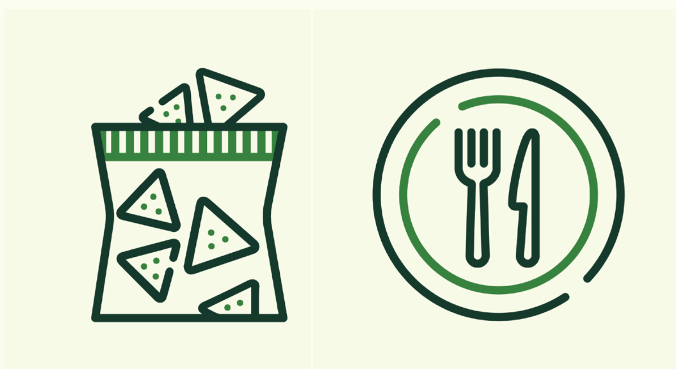
The guide is the work of a collaboration between Kathryn Peters and The Elections Group. The author would like to thank TEG's TJ Pyche, Michael Susek, and Jennifer Morrell for their insights, questions, and partnership in making the report possible. Thanks also to Melinda Dubroff, Kawandeep Virdee, Charley Johnson, and Keith Porcaro for sharing their time and talents. Any remaining errors are entirely my own.

GETTING STARTED AND DIVING DEEPER

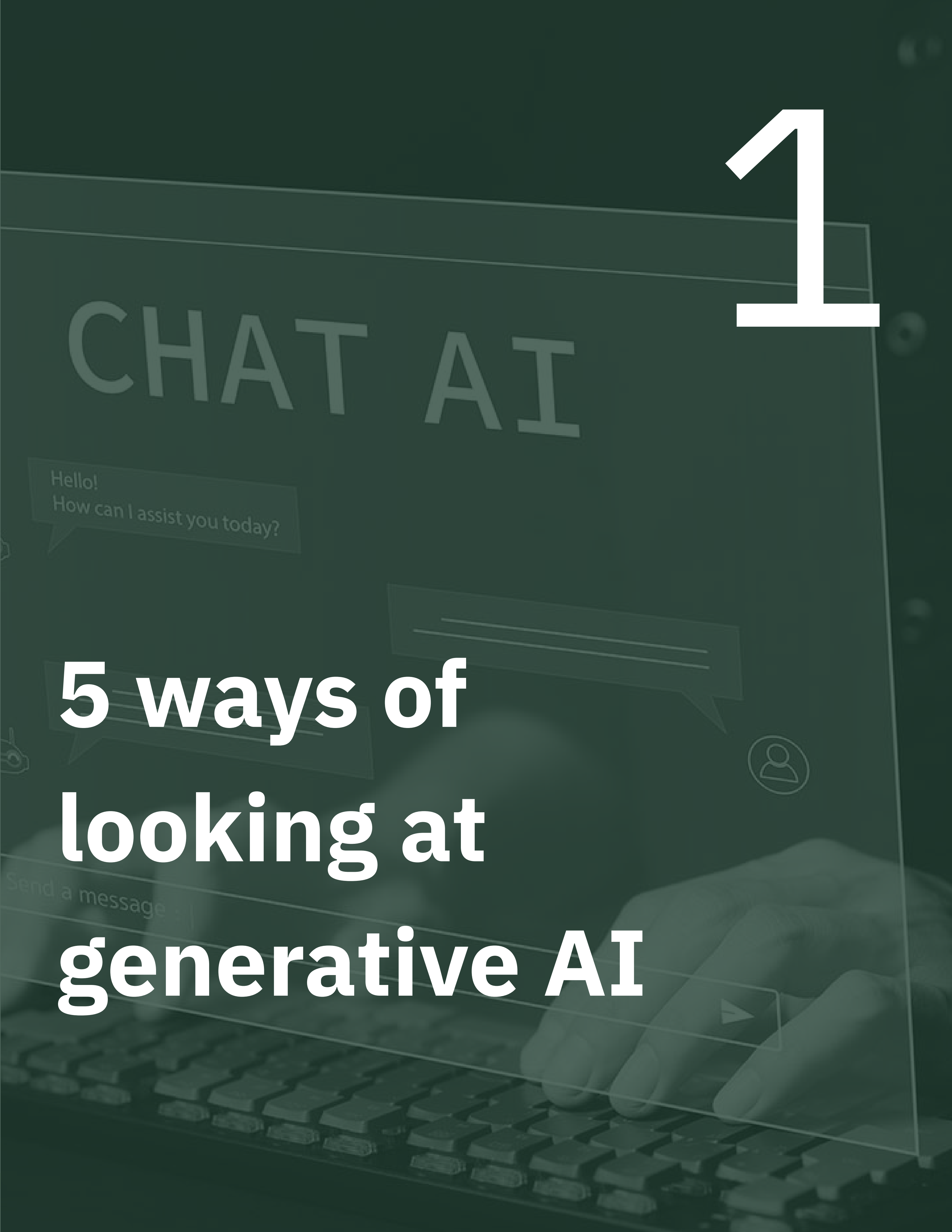
This guide is intended to serve as a bite-sized introduction to the generative artificial intelligence tools that can be read in a single sitting.

It includes annotated reading lists with additional resources that go into more detail. These lists are labeled using "snack" and "meal" emojis. The snack resources will tend to be brief and written using limited jargon. They'll help extend and reinforce the basic concepts introduced here. The meal resources may include more technical explanations or add related concepts not directly in this guide.

LOOK FOR THESE ICONS TO DIVE DEEPER



SNACK **MEAL**

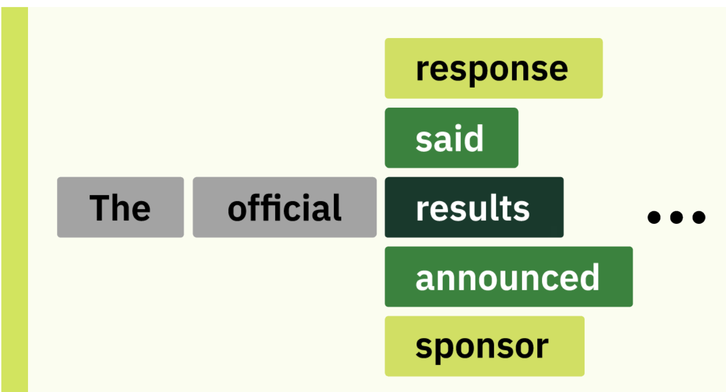


1

5 ways of looking at generative AI

IT IS: A PROBABILISTIC CONTENT GENERATOR

At its core, a large language model (LLM) uses a statistical process to generate text. Some of the earliest versions worked by predicting the best next word, and then the next word, and then the next… not unlike letting the autocomplete suggestions on a smartphone draft text messages.[1] The current generation can now model and create larger blocks of text at once, working in phrases, sentences, and paragraphs rather than single words. But at their most basic, these tools are designed to create new text based on a mathematical distillation of very large quantities of existing text.



In the same basic way, a diffusion model is a probabilistic image generator. It combines pixels into visual combinations using an underlying statistical model that distills very large quantities of existing images alongside extensive written descriptions of those images.[2]

In both cases, models tend to work best in situations where the content they generate is similar to content found in their training data. For example, an image generator may create an image of "a person riding a horse" easily but struggle to create an image that correctly shows "a horse riding a person,"[3] while a text model may create a reasonable chart based on a table of data but fail at calculating simple arithmetic. In both cases, the initial data used to build these models has many more examples of some types of content, while others are less common, and therefore, harder for the model to reproduce.

IT IS: A CHAT-BASED USER INTERFACE

In many of their current implementations, generative AI tools offer a chat-based interface to many other technical systems. The companies building these models have each released a chatbot to showcase their offerings: you might have experimented with ChatGPT[4] or Gemini[5] or Claude.[6] Even image generation tools are trained using text-tagged images, making text the key input for tools like Midjourney[7] or DALL-E.[8] While simple chat tools have existed for some time, the range of potential inputs to AI models is driving new work to define conversational user interfaces as an alternative to the graphical user interfaces we're accustomed to relying on.

In addition to using text (or voice) chat to interact

1 *The Financial Times offers a* ***quick, accessible illustration*** *of both next-word and next-phrase prediction. Georgetown University's Center for Security and Emerging Technology published a* ***more complete technical description of next-word prediction*** *and how models have grown more sophisticated over time.*

2 *Scientific American published a graphic explainer on* ***how diffusion models function****.*

3 *The* ***example actually given by an MIT researcher*** *in conversation is "a horse riding an astronaut," but we'll keep things less fantastic here.*

4 ***https://chatgpt.com/****, built by OpenAI*

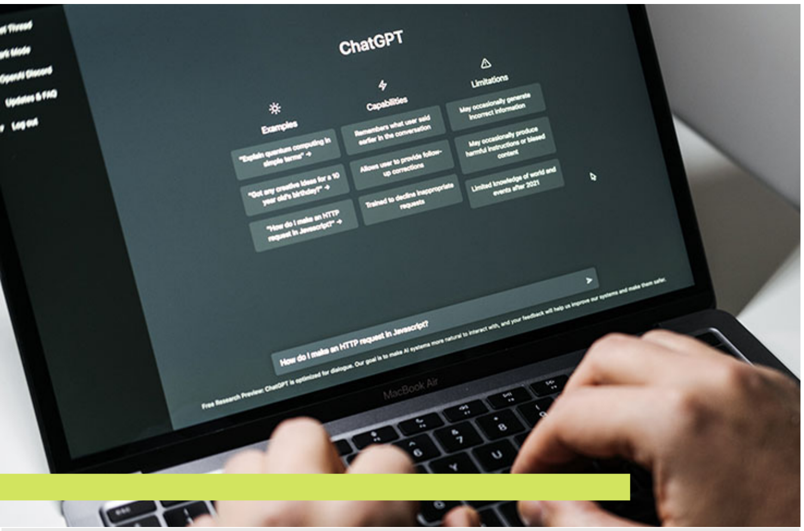
5 ***https://gemini.google.com****, built by Google*

6 ***https://claude.ai/****, built by Anthropic*

7 ***https://www.midjourney.com/home***

8 ***https://labs.openai.com/****, built by OpenAI*

with the models themselves, it's possible to develop applications that integrate other tools[9] to create AI-mediated interactions with maps, search, and other functions.



What can this tool do? We're still figuring that out...

In one sense, this can make AI models a very powerful interface: the person using one can attempt any function that they can describe in words. In another, it can also make them incredibly opaque, and even frustrating: a blank text box offers no visual cues to suggest what the tool is capable of, or to suggest what functions to try.[10] (Imagine flying a modern jet using voice commands alone!)[11]

Visual computing interfaces are the result of decades of research and design work. That deep expertise in user experience doesn't yet exist for unstructured conversational formats.[12] Emerging training materials and courses on "prompt engineering" represent early steps into understanding and formalizing best practices in using chat-based tools.[13]

IT IS: A MATHEMATICAL SUMMARY OF THE WEB

A contemporary artificial intelligence model takes vast quantities of inputs and converts them into small units or "tokens" of data—trillions of them. One way of conceptualizing this is to compare it to prior forms of capturing and storing digital data. Writing in the New Yorker, science fiction author Ted Chiang called ChatGPT "a blurry JPEG of the Web," describing how its model stores a very compressed copy of the public internet.[14]

This is because the first stage in developing a model is to gather a **dataset**. For a large language model, this dataset will be text, while for a computer vision or image-generation model it will require a collection of images with detailed descriptions and metadata. No matter what's in the collection, it needs to be huge—trillions of words huge. "We're maybe already running out

9 In addition to chat tools, these companies offer application programming interfaces (APIs) that support linking AI models into other software.

10 The Nielsen Norman Group of user experience researchers make the case that, even more than a shift from graphical to text interfaces, these tools also transition from users giving commands to users stating intent.

11 Designer Matthias Dittrich makes the case for just how hard it is to do complex tasks through a single prompt.

12 Work to define conversational user interface practices is underway—the Association for Computing Machinery adopted an annual conference on CUI (Conversational User Interface) in 2023. Many best practices from graphical user interfaces and experiences will transfer to this new field, making training in plain language, information frameworks like "bite, snack, meal" and other resources already in use from the Center for Civic Design and others useful even in these new settings.

13 One example among many, many available courses online.

14 The full article. A JPEG is an image file format with very customizable file sizes, from the very large and detailed to the very small and impressionistic. Or if digitizing photos doesn't make intuitive sense, Chiang also compares the process to a photocopy of a photocopy of a photocopy—Xeroxing the Internet down to size.

A high-resolution JPEG-format image. (Photo by Joe Shlabotnik)

A compressed, blurry JPEG-format image

of content to use" huge.[15] LLM developers are often secretive about exactly what data they've collected and used, but publicly available training data includes scraped content from across the Internet—Wikipedia, Reddit, newspapers, personal blogs—all of it raw material for modeling how words stack together into sentences and paragraphs.

Next comes **data processing**. At its most basic, a large language model is compressing all that text down into many measurements (called parameters) that capture which words are most often used in proximity to each other.[16] For example, the dataset text may regularly place the word "queen" near king, chess, royal, or the names of countries that still have monarchies, and far from terms that have nothing to do with governance or celebrity, like boil, meditation, or mist. Just as a JPEG image file compresses an image into many individual pixels of color, a large language model encodes these huge quantities of text down into parameters. Those parameters then allow the model to create new blocks of text that mirror those proximity relationships, putting similar words together as a way of creating plausible, grammatically-correct text.[17]

That resulting model doesn't retain any of the original text used in its training, but it keeps the word associations from that content and relies on them in generating new text. This can lead it to offer paraphrases that might reflect the original data quite closely, or new statements that remix those patterns in unexpected ways. For example, when a prompter asked Google's Bard for facts about the James Webb Space Telescope, the model's response included that "JWST took the very first pictures of a planet outside of our own solar system," which is not true, but is plausible based on news stories about the telescope that include the words first, pictures, planets, and solar system all in close proximity to one another.[18]

15 The Wall Street Journal summarized the current conversation about [identifying additional sources of training data] for AI models.

16 As described in rich, illustrated detail by the [Financial Times].

17 If the concept of "word proximity" still feels abstract, several online games feature a secret word and reveal the proximity score of each term you guess. [Semantle] is one example you can use to play and develop a more intuitive sense of how proximity gets encoded into these models.

18 NewScientist recapped the [error and public response]. In a more political example, technology journalist Casey Newton shared a screenshot of ChatGPT responding to the prompt "[What are some fun facts about the gay rights movement?]" with "the first openly gay person elected to the Presidency of the US was Pete Buttiegieg in 2020," mirroring that same tendency to build from words that may regularly appear together in texts in unexpected ways.

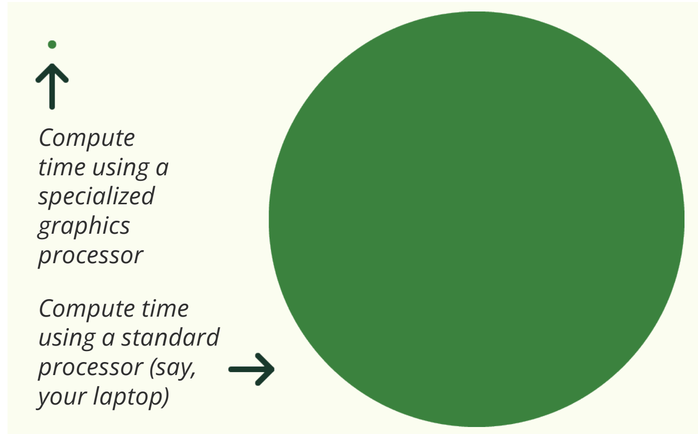
In this sense, a model takes (basically) the entirety of the web and compresses it down into mathematical associations that summarize the whole. That distillation is sometimes faithful and sometimes loses information or adds a kind of blurriness—not unlike looking at a too-small JPEG of a photograph.

4 IT NEEDS: SCALE

Early work developing artificial intelligence systems was limited by constraints on computing power, time, and complexity.[19] Previously unimaginable increases in available scope and power across three key areas make this current generation of tools possible.

The first is the quantity of training data. As noted in the previous section, large language models are trained on trillions of words, which is to say, nearly all digitized content. Collecting all of this is the result of book-scanning projects, the availability of optical-character recognition tools, increased global Internet connectivity and specifically access to social media and other publication tools that put so many of our thoughts and words online, and systems that make all these words discoverable and gatherable.[20]

The second is computing power to convert that data into a model. Distilling raw text into parameters requires quantities of processing effort many orders of magnitude greater than any other application. Those computing demands drive demand for semiconductors and processing chips and run on massive amounts of energy.[21,22]



The third is the number of parameters in the final model. Each competitor designing generative artificial intelligence uses its own machine learning processes to convert training data into a model, and those processes may summarize that data into more or fewer variables. In general, models built with more data and retaining more parameters tend to be more powerful—thinking back to the comparison to a JPEG file of a photo, the model can be higher-resolution and include more detail, or use more compression for a less precise capture of the original information.[23] In exchange for that precision, larger-parameter models require more capacity and energy to operate on an ongoing basis.

19 [A very brief history](#) that notes how these other computing developments have advanced artificial intelligence research is available from the Harvard Science in the News blog.

20 The New York Times offers a fascinating peek into the process of [identifying and collecting this data](#).

21 The AI Now Institute released a [detailed policy paper exploring the implications of this demand](#) for semiconductors, specialized software, and data center tools in building AI.

22 The graphic illustrates [a time comparison made by venture capital firm Andreesen Horowitz](#) estimating that one GPT-3 inference operation would require 32 hours (115,200 seconds) to run on a typical computer's central processing unit, but only one second on the custom graphics processing unit (an NVIDIA A100) preferred by AI companies.

23 Information is Beautiful maintains a chart [comparing how model sizes have grown over time](#). Here, it's possible to see that while some of the most visible large language models are also some of the largest, there's also quite a bit of size variation even among popular models.

IT NEEDS: HUMAN INPUT

Generative AI models rely on human effort—not only in the creation of the original texts and images that make up their training data, or the engineering and design work that goes into building the models, but throughout their life cycles.

Data preparation

Training data often requires human review prior to use. Data annotation may add tags or interpretive labels that make raw text much more useful for modeling purposes, for example.[24]

Training

Even once the training data is converted into a mathematical model, there's significant work to be done. At that stage, the model has a great deal of potential capability but very limited tools for human interaction. Its core function is next-word prediction, so given the start of a sentence or phrase, it will spin out additional text, but it can't provide an answer to a question, or engage in anything resembling dialogue, or edit an existing text.

Training is the process of providing the model with very specific examples of inputs and outputs to develop its user interface. For a chatbot, this may be sample prompts and model responses. For other custom uses, this stage offers an opportunity to refine how the model presents information and what types of human input it handles effectively. Drafting these samples, rating responses, and refining the model outputs all rely on human-generated content.

Oversight

Many current applications of AI still include a "human in the loop," often made invisible by the technology, providing additional review and oversight.[25]

Whether those inputs come in the stages of collecting data, building models, training the tools, or reviewing their work on an ongoing basis, much of what we label as "artificial" relies heavily on human decision-making and effort.

24 The Verge published a deep dive into the economy of annotation labor.

25 In one recent example, an anonymous claim that a majority of "Just Walk Out" purchases via Amazon were actually human-classified remains in dispute, but confirmed that Amazon relies on at least some human review and oversight of the system.

> **Generative AI models rely on human effort—not only in the creation of the original texts and images that make up their training data, or the engineering and design work that goes into building the models, but throughout their life cycles.**

GO DEEPER

"[Generative AI Exists Because of the Transformer](#)" by Visual Storytelling Team and Madhumita Murgia, Financial Times, 2023.

I want to see inside building a large language model: this is an illustrated introduction to transformers and how one specific technical insight made the current generation of large language models and applications possible. A great quick read to understand what's actually happening at the most granular level.

"How ChatGPT and Other LLMs Work—and Where They Could Go Next" by David Nield, Wired, 2023.

I want a quick overview of large language models to share: this is an extremely short review of how the transformer-based generation of LLMs and the resulting strengths and weaknesses of the chatbots built on those models.

"ChatGPT Is a Blurry JPEG of the Web" by Ted Chiang, The New Yorker, 2023.

I want to think more about the design implications of generative AI: Chiang's JPEG analogy and the related question of "when is a paraphrase what we need, and when is a quote more appropriate?" are helpful and applicable in many conversations about generative tools, and the full piece is very much worth a read.

"Artificial Intelligence Glossary: AI Terms Everyone Should Learn" by Adam Pasick, New York Times, 2023.

I forgot what a "transformer" is: This piece offers concise definitions of commonly-used jargon for easy reference.

"See How AI Generates Images from Text" by Sophie Bushwick, Matthew Twombly, and Amanda Hobbs, Scientific American, 2023.

I get text generation, but I don't know how that applies to images: a quick, illustrated overview of how image generation models work.

"Stop designing chat-based AI tools" by Matthias Dittrich, UX Planet, 2024.

I'm curious about how chat-based user interfaces change things: this designer describes the biggest drawbacks of chat interfaces and suggests a number of small, specific changes that can make working with generative AI tools much less frustrating.

"[AI: First New UI Paradigm in 60 Years](#)" **by Jakob Nielsen, Nielsen Norman Group, 2023.**

I want to think more about this "chat-based interface" concept: this piece from a user experience researcher discusses the current drawbacks of generative AI's user experience and imagines future interfaces.

"[AI Is a Lot of Work](#)" **by Josh Dzieza, New York magazine, 2023.**

On the people behind AI: a look into the work of Kenya- and US-based annotators working to create the data that goes into training large language models and diffusion models.

"[The Human in the Loop](#)" **episode of Computer Says Maybe, Alix Dunn, 2024.**

More on the people behind AI: This podcast episode explores the humans who make generative AI possible, for those who prefer an audio format.

"[Computational Power and AI](#)" **by Jai Vipra & Sarah Myers West, AI Now Institute, 2023.**

On scaling needs: This policy paper considers the current demand for computational power driven by generative artificial intelligence and the policy implications of that demand. It's a detailed piece that's nonetheless accessible for anyone interested in understanding how these tools are reshaping the energy and technology sectors.

"[How Tech Giants Cut Corners to Harvest Data for A.I.](#)" **by Cade Metz, Cecilia Kang, Sheera Frenkel, Stuart A. Thompson and Nico Grant, New York Times, 2024.**

On data needs: a New York Times team investigates the data sources major AI development efforts have tapped into and where those data collection efforts may violate the law.

2

How might we use it? What does it do well?

Some tools do one thing very well, and only one thing. Others—including technologies like databases or design software—are flexible and powerful in ways that support a broad range of uses beyond their most likely applications. Current uses of generative AI tools offers a good foundation for thinking about potential use cases in election administration, but their eventual applications may also include some surprising and creative discoveries.

This section provides an overview of what types of tasks generative AI tools are currently known to handle well, where and how their functionality falls short, and how they're already being tested and deployed in elections and other government settings. It includes a brief primer on fine-tuning and other ways of tailoring an AI model to a specific purpose.

WHAT IS IT GOOD AT?

At risk of stating the obvious, this generation of tools are referred to as generative because they're designed to create new content from inputs or existing bodies of work. Many of these uses look to generative tools to create content in a variety of forms.

Drafting support and creative ruts

Drafting text, images, or video? It's easy to enter a prompt and see what it returns—for many people, a blank page can be a high hurdle to overcome, while outlining or defining the prompt for what to create may help structure the process of working through what to say.[1] Drafting a public post on how post-election auditing will be carried out? It may be helpful to draft a prompt that includes the technical process documentation and spells out your expectations for tone, length, and key takeaways. That prompt may help create a rough first draft—or it may help get started writing from scratch after detailing the task.

When designing visual materials, looking at a variety of potential images can support decision-making about what kinds of illustrations best fit a specific need. Illustrating a polling place set-up guide and need additional icons to match the civic icon library you're already using? An image generator may help start the process, whether by creating the necessary icons or helping to define what additional illustrations to request from design partners.

In short, where it's valuable to begin with something, drafting prompts for a generative tool can provide a structured starting point for conversation, review, and creation.[2] Prompting is itself a creative effort, and in some cases, the process itself may prove to be a valuable step in clarifying your own needs and assumptions.

Summary and synthesis

These models are built on distilling large volumes of text into math. That same underlying process, applied to specific documents, can generate recaps or summaries with relative accuracy, which can be an effective first step in reviewing.[3] That distillation can also be used to identify clusters or specific

1 *One educator suggests that using the prompt-design process to have students think about context and brainstorming can be helpful in the classroom. Similar processes may be beneficial in an office setting as well.*

2 *Faculty at the University of Mississippi's student writing center summarized their experience testing generative writing support tools with students: "Students felt that the tools were helpful for finding ideas to get started with writing, to find sources once they had started writing, and to get help with counterarguments and alternate word choices. But when given the choice to use the assistants or not, most declined. Generative AI at this stage is unreliable, and many students found the trade off in reviewing AI suggestions to be too time consuming. And many students expressed a preference for continuing to develop their own voices through writing."*

3 *For a very quick, intuitive demonstration, see "tldr scale," which summarizes an article to a specified number of words.*

features across a large text or dataset.[4] What are common themes that run through the poll worker training manual? Do those themes represent the basic training values we're working to instill? How many sections of state election code address a given topic (say, early voting)? Which images in the set include flag imagery?

Based on other proximate words or concepts, a model might also offer additions or omissions to a document in review, or extrapolating related content from the initial text. If our main website menu includes links to voter registration, requesting ballots, and information on who's in office, what other offerings does the model associate with those, and should we consider adding any of them to our layout as well?

Sample data

Need to test a system with correctly-structured but fake data? Generating artificial polling places, sample voters for poll worker training, or other plausible filler to replace more sensitive information are all possible to create from samples that demonstrate the format and type of content you need.[5]

Basic automation

With additional training (for example, fine-tuning or multi-shot prompting, as described in the "Putting a model to work" section), it's possible to apply generative tools to simple data processing or review tasks. Given enough examples of content that complies with or violates specific rules, a model can apply the pattern across other documents or data. Comparing spreadsheets of ballot information to PDF ballot layouts as an additional proofing check is one simple automation early election adopters have tested, for example.

Play

This might feel incongruous in a consideration of applications within a professional field, but there's a reason that so many viral examples of AI-powered chatbots have been humorous or playful—the unexpected nature of how these tools respond can be fun![6] Using a model to create limericks about poll workers, cartoons of mail ballots, or finding other ways of interacting with serious information in creative ways may help election officials see common materials in a new light or suggest potential communications materials.

This type of experimentation has the added benefit of getting more familiar with specific AI tools, learning how variations in prompting affect outputs, and forming insights into what is and is not useful in your specific work through low-risk tasks before applying those lessons to implementations with more significant impact on your core operations.

4 *Data scientists at the Research Triangle Institute summarized their experiences using LLMs to [generate qualitative text coding, cluster identification, and labeling].*

5 *Popular non-election uses include artificial facial imagery ([https://synthesis.ai/]) and software testing ([https://sdv.dev/]).*

6 *Artist-developer Kawandeep Virdee has a [library of creative tools] including a pep-talk generator, a doodle-expander, a multi-person storytelling experience, and a creative coach.*

WHERE DOES IT FALL SHORT?

While this generation of AI models are much more general-purpose than their predecessors, and their blank chat interfaces suggest that they can take on any task, they do still have known shortcomings and limitations. Some of these shortcomings are specific to the technology itself, while others are rooted in how these tools confound our expectations about how software works.[7]

Limited public-facing chat

Whether it's a car dealership bot making $1 sales offers[8] or a city government bot advising small businesses to break the law,[9] using generative AI to provide an unstructured prompt interface is a significant challenge. While the makers of AI models benefit from demonstrating the range of inputs they can address, most other organizations deploying chatbots want (or even need) a more limited range of interactions.

Anticipating and designing for the sheer range of potential inputs, from intended to off-topic to intentionally malicious or just plain silly can trip up even the most careful planning.

Accuracy

Because these models don't "know" their underlying data in the way that a human mind does, they can create probable, well-structured content that isn't factual. These are popularly called hallucinations, though that term suggests that the model is capable of believing these outputs. More precisely, it's that remixing commonly-adjacent terms into new sentences and phrases will sometimes combine them in ways that are grammatically correct but no longer true.[10]

Because these models don't "know" their underlying data in the way that a human mind does, they can create probable, well-structured content that isn't factual.

Journalists have documented this phenomenon with election information specifically: a collaboration between Proof News and the AI Democracy Project developed a standard set of election information questions ranging from "Can I wear my MAGA hat to the polls?" to "Where do I vote [zip code]?" and ultimately rated half of all answers to be inaccurate across the five chatbots tested.[11]

One simple type of this error is negation. The word "not" is extremely common and so it may not have strong adjacent or distant relationships to many others—but its presence denotes a significant shift in the relationship between other words in phrases where it appears.[12] A model may retain the proximity between those other words without distinguishing that it's because one is often described as not the other, or something not to do.

7 Often, technology errors come up at this [intersection between people and tools](), as researcher J. Nathan Matias points out in his work.

8 As happened to the [Chevrolet of Westonville]() dealership

9 As [an official New York City business chatbot did](), with advice for business owners on splitting tips with employees and for landlords hoping to discriminate against renters.

10 In addition to the examples provided in the section on AI as a mathematical summary of the Internet, cognitive scientist Douglas Hofstadter provided examples of an earlier generation of ChatGPT [engaging confidently with a variety of nonsensical questions]().

11 The full report is available at "[Seeking Reliable Election Information? Don't Trust AI.]()" (Reassuringly, [a recent survey by the Bipartisan Policy Center]() shows that few Americans are turning to AI chatbots for their voting information.)

12 Entertaining examples include image generation results of "[rooms with no elephants]()" and "a kitchen with no polar bear." Abeba Birhane and Deborah Raji [explain the phenomenon in more detail]().

Bias

Existing social biases get deeply integrated into generative models through their initial data, training processes, and design choices.[13] Where models are used to make predictive or differential decisions about populations, they will replicate those patterns of discrimination. Some may be easily spotted—as when Amazon implemented a hiring tool that significantly favored male candidates[14]—while others may be quite subtle.

Silent failures

Generative tools will generate something in response to every prompt. In specific cases where training or fine-tuning has established a guardrail, the tool might reject the prompt request. In other instances, the tool won't be able to complete the prompt as requested but will return a result that attempts to do so, to varying degrees of success, leaving the user to determine if the response is adequate.[15] Where current user interface and experience patterns lead us to expect error messages or other clear signals of failure, adapting to situations where we must assess when the tool is failing independently of explicit cues requires new habits.

Authoritative presentation

Compounding the challenges of accuracy and silent failure, the structure and format of AI-generated content often adheres to expectations about how a good response should look or sound.[16] Hurried users may place trust in these superficial quality cues and overlook more substantive errors.

Guardrails limiting election use

The companies developing generative AI tools understand that providing inaccurate election information is a risk to their users and public reputations. In response, many have placed guardrails that limit the tools' response to prompts requesting election information: for example, prompting ChatGPT with content about voting processes triggers a response that directs users to CanIVote.org instead of generating new content.[17] Similarly, many chatbots will refuse to answer questions that may otherwise help users threaten election security, which limits their ability to engage with those topics in supportive or helpful ways. These intentional limitations reduce one set of election-related risks while blocking similar but legitimate uses by election officials when good-faith prompts trigger the guardrail response.

13 In one early study, researcher Joy Buolamwini demonstrated that facial recognition models performed much less reliably on darker-skinned and female faces.

14 Specifically, the hiring model favored candidates named Jared and those who played lacrosse.

15 This tendency anchors a whole subfield of research into self-correcting, seeking approaches to reduce this particular error.

16 As one computer science professor noted after reviewing a chatbot's answers to test questions, "The danger is that you can't tell when it's wrong unless you already know the answer."

17 As described (with very limited detail) by OpenAI in a blog post on its 2024 election preparations.

PUTTING A MODEL TO WORK

For all the training and human effort that goes into their initial development, applying a generative AI model to specific tasks or subject areas may require yet more work to tailor and narrow its focus. (This is one reason why earlier, single-purpose machine learning tools are still extremely useful—where they lack flexibility, they're already specialized to their tasks. Algorithms like the Electronic Registration and Information Center's identity-matching work serve their purpose well and wouldn't necessarily benefit from adding generative elements).

There are many ways to refine a generative AI tool, and this section summarizes three relatively common techniques as a general overview of the kinds of specialization that are possible.

Fine tuning

Fine tuning works in the same basic way as initial training: it provides a model with sample inputs and outputs, offers feedback on the model's responses, and helps refine how it responds to specific types of requests. For customized use cases and specialty applications, fine tuning becomes an additional training stage used to refine a copy of an existing model. This step can prepare a model to address a specific use case that its initial training did not include or add data and parameters in a specific field of knowledge to improve accuracy and completeness.

While training is generally done by the makers of an AI model prior to public release, fine tuning can be led by customers using the model. (Given the costs and effort this level of training entails, fine tuning an election-specific model would likely require collaboration among many election offices, rather than be a project any one jurisdiction could undertake independently).

Retrieval-augmented generation

In addition to training and fine-tuning, many applications using generative AI models also rely on a technique called retrieval-augmented generation, in which the tool incorporates a specific body of documents or knowledge that the model then uses in carrying out its tasks. This external corpus could include, for example, the state election codes or internal policy and procedures manuals. Adding these reference materials does not eliminate potential sources of error (see the "Where does it fall short?" for more detail), but can serve to narrow the scope of a prompt and limit what sources the model draws on in a specific task for greater relevance.

Prompting

Finally, prompting is the term for drafting the input for a generative model. Prompting an AI model is not as structured as clicking a button or typing in a structured text field. Instead it serves as another form of training and design. Because these tools are so new, just about everyone is discovering and adapting new techniques for getting models to do what they want, and "prompt engineering" courses on best practices are popular.

The length and content of a prompt play a big role in the kinds of outputs it generates. A prompt can ask a question or specify a task. It's also possible to offer examples in a prompt as a kind of final training or as a way of specifying the format you'd like for the output. For complex outputs, using the prompt to specify each step in a longer process may increase the reliability of the output, or divide the output into enough pieces to simplify checking of when in the process errors come up.[18]

Fine-tuning and retrieval augmentation help apply the broad, general functions of an AI model to a specific use case. Prompting gives users wide latitude to specify their needs and shape the outputs they receive. Together, these techniques

18 One prompt engineering course built with OpenAI.

work to focus the capabilities of an AI model and apply them to your desired process or output.

WHO IS USING THESE TOOLS, AND FOR WHAT PURPOSES?

Much of the public conversation about generative AI tools is future-facing and many organizations and individuals using these tools are quiet about exactly how they're configured and deployed, making it challenging to assess where and how these models are most useful and capable in practice.

Election administration uses

The election administration community includes a number of early adopters and willing testers. In a recent survey led by the Brennan Center, 7% of election offices reported using AI in their work in some form.[19] The three most popular reported uses, at approximately 2% of respondents each, were:

- Drafting social media and press releases
- Locating polling places
- Translating materials

In addition, The Elections Group has tested a number of sample scenarios using ChatGPT-4 to assess how well the tool performs at support functions. Available demos include:

- Summarizing RFP responses and providing a written comparison of their contents, to support a "first pass" in a procurement context
- Requesting a research summary on an election-related topic (in this case, ranked choice voting) as a first step in a deeper review of available material on the topic
- Modeling predicted voter turnout based on a (publicly available) voter history file
- Generating bar graphs and a histogram of voters by precinct based on a (publicly available) voter file

These demos are available upon request by emailing support@electionsgroup.com.

Other tested uses include ballot proofing and drafting press releases and other structured public-facing content.

Other government uses

The federal government also maintains a public repository of use cases where agencies have deployed AI-based tools.[20] These may include earlier iterations of machine learning and artificial intelligence rather than generative tools. Uses currently featured include:

The National Oceanic and Atmospheric Association (NOAA) is analyzing urban heat islands to better understand and predict extreme weather.

The Department of Veterans Affairs is summarizing and analyzing written feedback on its services to identify key topics and trends.

The Patent and Trade Office is adding new search tools to assist patent examiners in identifying prior art and related patents as part of their review process.

The Department of Homeland Security is piloting tools to support summarizing investigative reports, developing risk management plans for resilient communities, and personalizing immigration officer training materials.[21]

19 The full survey results also note that approximately 1/3 of officials would welcome additional guidelines on using AI.

20 Those use cases are published in the Federal AI Use Case Inventories on AI.gov.

21 As described in a news release from DHS: Department of Homeland Security Unveils Artificial Intelligence Roadmap.

GO DEEPER

”Seeking Reliable Election Information? Don’t Trust AI” **by Julia Angwin, Alondra Nelson, and Rina Palta, Proof News, 2024.**

I want to hear what common chatbots are saying about elections: these journalists and researchers convened a group of election officials to test common voting information questions on popular chatbots and found that a majority of responses were rated to be inaccurate by the experts.

”ChatGPT and Google Gemini Are Both Doomed” **by John Herrman, New York Magazine, 2024.**

I want to think about the limitations of open chat: a good overview of the harms that can come from treating LLM chatbot interfaces as general-purpose tools that goes beyond the specific examples shared in this guide.

“ChatGPT is a bullshit generator. But it can still be amazingly useful” **by Arvind Narayanan and Sayash Kapoor, AI Snake Oil, 2022.**

I want to think about the challenges of accuracy: pardon the profanity. This short piece considers the challenge of working with a tool that doesn’t have an intrinsic “ground truth” to build from and suggests how to approach deploying generative tools going forward. Their translation case suggests how providing these tools with an external, authoritative source of truth can increase the accuracy and reliability of what it generates.

In addition to this specific piece, the AI Snake Oil newsletter is consistently a good source of practical insights. Arvind Narayanan is on the faculty and Sayash Kapoor is a graduate student in the computer science department at Princeton. They share a deep interest in tech policy engagement and an interest in distinguishing the real benefits of new technology from the hype.

“Detain/Release: Simulating algorithmic risk assessments at pretrial” **by Keith Porcaro for the Berkman Klein Center, 2019.**

I want to think more about authoritative presentation: this case study from a law school course demonstrates how students and professional users responded to an authoritatively-presented risk assessment tool despite the fact that the underlying assessments were in fact randomized. It’s a short, effective illustration of how we rely on design cues when deciding to trust technology tools and how readily humans will form mental models to ‘fill in’ details about how a tool works.

"ChatGPT, Galactica, and the Progress Trap" **by Abeba Birhane and Deborah Raji, WIRED,2022.**

I'd like another general overview of potential limitations: Birhane and Raji describe how authoritative but incorrect outputs cause harm and critique framing them as accidents or errors given the design of these tools.

"Teaching AI Writing: Ideas" **by Leon Furze, self-published, 2024.**

I want to try using generative AI as a drafting tool: this blog post discusses techniques for teaching students to write with generative tools and provides approaches for how to think about prompting and chat interactions as the starting point for creative work.

"CISA's Roadmap for Artificial Intelligence" **by the Cybersecurity & Infrastructure Security Agency**

I want to see more about how the federal government is using AI: CISA has developed its own artificial intelligence planning roadmap. Of interest to deploying generative tools in elections, their identified "Line of Effort 2" is dedicated to assuring AI-based systems and supporting secure-by-design tools for use at all levels of government.

"How Large Language Models Work - From Zero to ChatGPT" **by Andreas Stöffelbauer, Data Science @ Microsoft, 2023.**

I'm thinking about prompting as a training process: this piece is less clear as a general-purpose introduction to generative tools than the resources in the previous section, but it includes a good explanation of prompting and how a technique called "few-shot prompting" works as a means of doing on-the-fly training of a generative AI tool as an end user.

"ChatGPT Prompt Engineering for Developers" **by DeepLearning.AI.**

I'd like to dive into practicing prompting: this specific course is designed for software developers but provides an accessible overview of best practices in drafting prompt texts and techniques for breaking complex processes into manageable steps.

3

Deploying generative tools in election administration

As generative tools move into broader use, the initial adopters who decide where and how to use them hold real power and influence over their future. We join in a much larger conversation about whether and how generative tools are deployed. This section is written to support election officials as they imagine potential uses for generative tools and then choose, plan, and review those uses over time.

As the introduction to this guide notes, the hallmark of a good piece of technology in election administration or any other professional context is that it solves a real problem for its users. Identifying appropriate matches between generative tools and election administration needs is an ongoing project. The Elections Group has led conversations with election officials to identify areas of potential for experimentation and use, but many of these uses are hypothetical at present.

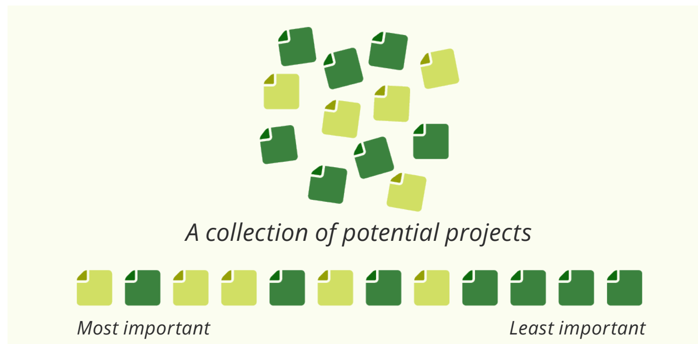
Even once potential applications of generative AI tools are identified, these systems pose new challenges to effective human governance and oversight compared to previous technologies. The tools themselves are black boxes, far too complex to understand by reviewing their datasets, training processes, or model parameters. And their generative design and constant incremental change means that a single point-in-time analysis of how the tool addresses a specific question or scenario cannot ensure a repeatable outcome even under near-identical conditions.

Instead, these tools will require ongoing human oversight and review practices at all stages of their use, from developing use cases through procurement, training, and at regular intervals throughout their implementation. The sensitivity and importance of election work will require that any use of generative AI tools also include adequate training support to ensure that those "humans in the loop" provide appropriate review and oversight.

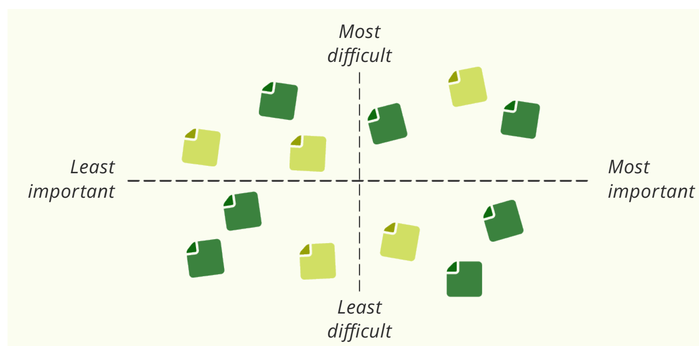
This section summarizes key considerations in deciding whether and how to use generative tools and suggests concrete practices and questions to use when collaborating with procurement specialists, technologists, and other experts throughout their life cycle.

IMAGINING ADDITIONAL POTENTIAL USES

In design research, one popular tool for helping teams prioritize work is an importance-difficulty matrix. In that exercise, a group of stakeholders comes together to list a potential set of projects to complete, requirements for a program, or solutions to a shared problem. The same prioritization system can also be used to help generate ideas for potential work by reversing the process. A quick look at how an importance-difficulty exercise typically goes:

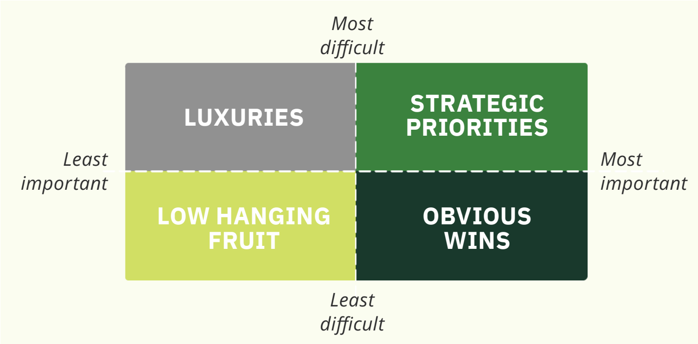


First, the group rank-orders every option from most to least important, discussing any disagreements or justifications along the way.



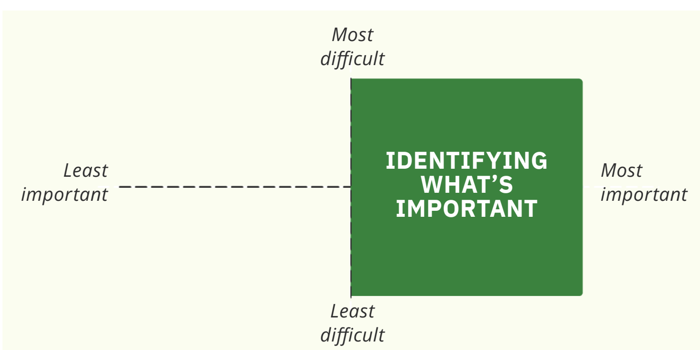
Next, they do the same rank-ordering by difficulty.

The final result of those conversations is a chart that looks something like this:



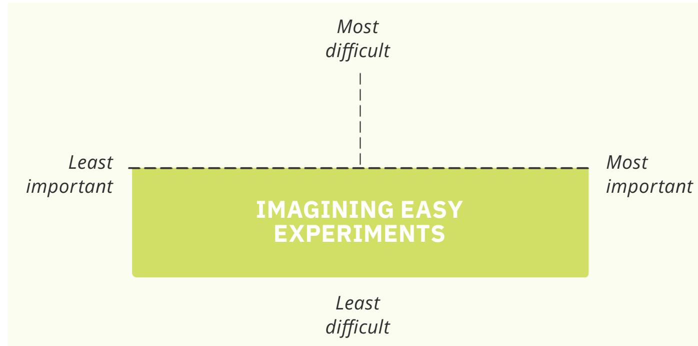
The matrix helps sort potential priorities into four categories: high-importance, low-difficulty items are obvious wins to pursue. Low-importance, high-difficulty items are luxuries, and might or might not ultimately be worth the investment of time, effort, or cost. High-importance, high-difficulty items are strategic priorities that may need additional planning or coordination to ensure their success. And finally, low-importance, low-difficulty items may serve as low-hanging fruit, opportunities for testing, or work to complete opportunistically in gaps alongside higher-importance pieces.

Using that framework as a model, there are several questions you might ask to begin thinking creatively about your own priorities for testing or deploying generative tools.



Identifying high-importance needs:

- What are the top five challenges I want to address in 2025?
- What are my five highest priorities for improving internal processes?
- What recurring tasks or processes are the most effortful or time-consuming in my office? Where do I (or my team/ staff) most need additional support?



Imagining simple experiments and easy opportunities for learning:

- Where might I test initial content generation or playful interactions around low-risk tasks?
- What data do I have that would be simple to visualize and might spark new ideas to see charted or mapped?
- What large documents, datasets, or sources of information would I want to be able to summarize, skim, or ask questions about more easily?

These lists can serve as a starting point in answering the question of where new technology uses might align with your specific needs and priorities. As you answer these questions individually or in conversation with other election officials, we'd love to hear your responses.

Email support@electionsgroup.com with your responses and to request a similar workshop for your organization or jurisdiction.

> **The hallmark of a good piece of technology in election administration or any other professional context is that it solves a real problem for its users. Identifying appropriate matches between generative tools and election administration needs is an ongoing project. The Elections Group has led conversations with election officials to identify areas of potential for experimentation and use.**

PEEK OVER THE SHOULDERS OF OTHERS

Other election officials are imagining potential uses for generative AI tools as well, as seen in this summary of a generative workshop held at the Election Center gathering in Portland, Oregon in April 2024. If you would like The Elections Group to conduct this workshop or a similar one for your office or association, or you'd like to join a future workshop, email TJ Pyche at tj@electionsgroup.com.

Communications

- **Social Media Management:** Creating, managing, and reviewing social media posts; branding, including drafting logos; voter outreach campaigns.
- **Customer Interaction:** Answering voter questions via chatbots; AI reviewing phone bank customer service calls to identify changes in conversation tone.
- **Content Creation**: Website content, form letters, press releases, training modules, voter tutorials, outreach presentations, legislation testimony, simplifying reading levels, summarizing documents, drafting media PR.
- **Document Preparation:** Preparing presentations for memberships or community groups, preparing and summarizing training materials, writing instructions for PEO + voting, meeting agendas, writing emails, creating FAQ pages.
- **Communications Management:** General communications to voters, outreach to specific demographics, poll worker training modules, establishing protocols for content review.

Budgeting and Procurement

- **Budget Analysis and Preparation:** Estimating election costs, comparing budgets across jurisdictions, drafting budget projections, fiscal impact statements.
- **Procurement Management:** Creating and comparing RFPs, searching for vendors, managing inventory, assisting with RFQ/RFP analysis.
- **Financial Operations:** Outlining responses to grants, creating fiscal impact statements for bills, helping explain financial jargon, preparing payroll, and mileage reimbursements.

Operations

- **Training and Conferences:** Generating ideas for icebreakers, creating conference outlines, structuring training sessions, poll worker training, generating training materials.
- **Scheduling and Planning:** Scheduling, planning delivery routes, election day scheduling, workforce planning.
- **Task Automation and Support:** Sorting people into different groups, uploading and summarizing surveys, drafting Standard Operating Procedures (SOPs), generating schedules.
- **Election Specific Tasks:** Ballot proofing, creating election timelines, analyzing survey feedback, delivering voting equipment, managing ballot box collection.

Analytics and Summarization

- **Data Analysis and Reporting:** Creating correlations from voter files, analyzing voter turnout by location and demographics, voter registration analysis, summarizing data from elections.
- **Documentation and Compliance:** Writing scopes of work for RFPs, creating plain talk documents, compliance checks (uploading laws and checking processes), redistricting analysis.
- **Security and Verification:** Managing watermark security for documents, signature verification, ballot inventory management.

ASSESSING AI-SPECIFIC RISK

Generative tools introduce new risks than prior generations of AI, and yet when public discussions of these risks range from election-deciding deep fake videos to the potential for a complete takeover of humanity, it can be extremely difficult to step back and make clear assessments of present risks and potential mitigation measures.

A recent academic paper reviewing other research into the risks of open AI models proposed the helpful concept of "marginal risk."[1] In short, as the authors reviewed proposed threat scenarios, they asked:

- Does the researcher identify a specific threat, including both the malicious actor and the threat vector?
- Do they describe the existing risk level of this threat?
- Do they consider existing defense or mitigation measures available?
- Do they describe how the use of artificial intelligence (in this case, open models) poses new marginal risk compared to that baseline?
- Do they consider what potential defense or mitigation measures could address this marginal risk?

As election officials have become risk managers, they've had to imagine and plan for a wide range of scenarios already. Thinking about generative technologies, then, in the framework of "which existing risks or threats might generative tools exacerbate, and how?" provides one conceptual toehold for assessing risk in a more concrete fashion.

A recent report on "How Election Officials Can Identify, Prepare for, and Respond to AI Threats"[2] focuses on external uses of AI, rather than its use within election offices, but illustrates this basic framework. As the authors note, many of the threat scenarios they imagine were possible prior to the existence of generative AI tools, but these technologies make them easier and cheaper—and more likely to target a range of election targets. When considering the risk in an AI-driven scenario, starting with "what would the risk be if this scenario did not include generative AI tools?" can help isolate where the risks are genuinely new or where they represent shifts in likelihood or impact of existing risks in your framework.

Confidentiality, privacy, and data controls

One specific marginal risk of using the free, public versions of generative AI tools is that their terms of service typically allow their parent companies to use interactions—prompts, data, everything a user might submit—in training future iterations of these tools. It's not yet clear how a large language model might integrate or repurpose any given prompt or file shared, which means this provision represents unknown new threats to voters' privacy and the confidentiality of sensitive election

1 *Paraphrased from "On the Societal Impact of Open Foundation Models" by Sayash Kapoor, Rishi Bommasani, et al, ArXiv, 2024.*

2 *From the Brennan Center in partnership with The Elections Group and Institute for the Future.*

information.[3] Using these tools to process or handle any confidential or sensitive information, whether voter data or confidential data on election procedures or equipment, is inappropriate and in some cases may be illegal.

Many generative AI tools offer limited or zero data retention terms for at least some of their paid service plans. Evaluating the specific terms of a given tool and whether they meet requirements for data management practices in your jurisdiction is a key planning conversation to have before implementing generative AI to handle non-public data (see the "Choose, Plan, & Review" section).

CHOOSE, PLAN, & REVIEW (CPR)

In the recent report "Safeguards for Using Artificial Intelligence in Election Administration," Edgardo Cortés and his coauthors propose a framework they call "Choose, Plan, and Review" when considering whether and how to deploy artificial intelligence tools.[4] The report details many potential marginal risks of AI tools in the elections field and offers a framework for reviewing and mitigating those risks.

The CPR framework includes key considerations and questions to ask at each stage, applicable to any consideration of implementing artificial intelligence tools from the generative technologies discussed here or to established machine learning tools (like signature-verification support).

This section summarizes the CPR framework and provides worksheets to support implementing it in decision-making processes about generative AI tools. These worksheets are best used alongside the full report.

Choose

If, after working through the imaginative questions, you have a viable use case (or several) to explore, the next step is to choose what tool or tools to consider. The very first question to ask: Is this the simplest available technology capable of addressing this need? If your goal is to explore and familiarize yourself or your team with generative tools, then the answer will be yes. For deeper data handling or automation, there may be solutions available that rely on narrower forms of artificial intelligence or simpler algorithms that can meet the same requirements while introducing less complexity or lower risk into your operations.

In addition to following standard procurement and review processes for information technologies, the CPR framework suggests several questions specific to the use of artificial intelligence tools:

- **What are the core requirements for this project, and can simpler tools meet those requirements in whole or in part?**
- **Will this tool make or contribute to decisions that could affect a voter's ability to cast their ballot?**
  - **If so, have you identified how to ensure**

3 Researcher Alice Marwick explores the limitations of traditional privacy protections in understanding and addressing the privacy threat of generative AI tools.

4 Edgardo Cortés, Lawrence Norden, Heather Frase, and Mia Hoffman, "Safeguards for Using Artificial Intelligence in Election Administration," published by the Brennan Center.

  - human oversight and final decision-making authority?
- **Would variations in the tool's performance cause harm?**
  - Are you able to test the tool using real-world data rather than artificial sample data to assess its likely performance as deployed?
- **How critical is reliability? Are there ways that bias or variation in how the tool responds case-by-case could cause harm?**
  - Are you prepared to detect and address those discrepancies?
- **Generative tools will offer different outputs in response to the same input. Does that variability pose any risks?**
- **How many people will interact with this tool? Is there capacity for training those individuals in how the tool functions and how it may confound basic assumptions about software applications?[5]**

These additional questions explore the potential additional risks or costs of using artificial intelligence tools. Working through them with procurement and IT partners will help confirm a good fit between your needs and the solution—or highlight potential mismatches or unexpected difficulties that might otherwise turn an easy win into a headache down the road.

Plan

Once the decision to deploy a generative tool has been made, the planning and preparation stage also requires additional steps. Critically, ensuring that anyone working with the tool is familiar with how generative technologies differ from other software and are both prepared and supported as they navigate new types of failures, interact with authoritatively-presented responses, and make decisions about how to use the tool's responses will require both initial training and new procedures for ongoing support.

In most cases, new tools will support existing roles and people in carrying out their functions. In those cases, shifting their responsibilities from direct completion of work to supervision and oversight may be straightforward. If the tool will accomplish a new task that is not currently assigned to a specific person or role, or if it replaces a function currently assigned to a person or role (i.e. a poll worker or other temporary staff) who should not be reassigned to a tech-supervisory capacity, ensuring clear delegation and responsibility for the tool will be a top consideration.

As you're creating training and implementation plans, the CPR model suggests asking the following questions:

- **Have you spoken with other organizations using the tool to learn from their experiences?**
- **Do the staff who will use this tool understand**
  - The task(s) it performs
  - The process used to do so
  - The data the tool uses
  - Its expected performance
  - Common risks or issues
- **Are staff familiar with the tool's user interface?**
- **Are staff trained on how to handle issues and internal accountability procedures for errors or failures?**
- **Do staff understand what the requirements are for human review and the importance of those requirements?**
- **Have you publicized the decision to implement artificial intelligence to the public, including a plain-language description of how it will be used,**
  - And have you offered an opportunity for feedback?

5 As discussed in the section on "Where do they fall short?", training may need to include not only information on the specific tool and its use but lessons in the complexity of unstructured chat prompts as an input, the potential for inaccuracies or silent failures, and the importance of not trusting responses solely based on their authoritative presentation or appearance.

- **Are contingency plans in place for errors or failures?**

Review

The makers of generative tools are constantly refining their models with new training and fine-tuning, and new generations built on ever-larger datasets are released regularly. In combination with the fact that these tools are designed to give variable responses to consistent inputs, they require regular, ongoing testing and review rather than a one-time battery of acceptance tests or regularly-scheduled review after planned new releases.

As you roll out generative tools, additional questions to ask in planning include:

- **Do your error and failure processes include auditing and review?**
- **How do your processes ensure human oversight of AI systems—especially where their use will affect voters?**
- **What procedures support external challenges to AI use, and are you prepared to redress those challenges?**

GO DEEPER

"Prioritization Grid" from the IBM Enterprise Design Thinking Toolkit.

I want to hear more about prioritization exercises: this simple introduction to using prioritization grids for project planning may be useful to holding bigger imaginative planning conversations across teams or with external stakeholders.

"8 Steps Nonprofits Can Take to Adopt AI Responsibly" by Beth Kanter, Allison Fine & Philip Deng, Stanford Social Innovation Review, 2023

I'd like other insights on what deploying AI tools well looks like: "using AI ethically is not a technical challenge but a leadership imperative, and its adoption must be deeply human-centered," these authors write, and then offer a framework for developing effective policies and practices for the use of AI across a variety of nonprofit or social good contexts.

"How Investors Can Shape AI for the Benefit of Workers" by Elana Berkowitz and Courtney Leimkuhler, Stanford Social Innovation Review, 2024.

I'd like to hear how other industries are thinking about what needs or challenges should drive generative AI use: "for example, think of a nurse manager or a teacher who no longer needs to devote 40% of their time to jockeying a schedule or designing slides and worksheets. Instead, they can lean into the uniquely human elements of their job in ways that require emotional and contextual assessment and insight that technology cannot replicate. AI will isolate and highlight these soft skills in ways that will make them more valuable and potentially higher compensated." This article reviews how AI tools could better support caregiving jobs and improve workers' satisfaction. Their suggestions for considering where and how technology and human skills dovetail may be helpful in imagining potential uses in other fields.

"Safeguards for Using Artificial Intelligence in Election Administration" by Edgardo Cortés, Lawrence Norden, Heather Frase, and Mia Hoffman, Brennan Center, 2024.

Cortés and his coauthors consider the election-specific risks of using artificial intelligence and apply a CPR—choose, plan, review—framework to recommend how officials can ensure that any use of AI technology supports their work and limits the risk of harm. Though it's a brief read, this framework anchors careful decision-making about technology tools and proposes a series of conversations and actions that require significant follow-through to implement.

"How Election Officials Can Identify, Prepare for, and Respond to AI Threats" **by David Evan Harris, Lawrence Norden, Noah Praetz, and Elizabeth Howard, Brennan Center for Justice, 2024.**

I'd like to update my risk analysis to incorporate external AI threats: this guide reviews where generative tools may increase the risk of existing threats to effective election administration and where it introduces new threats.

"On the Societal Impact of Open Foundation Models" **by Sayash Kapoor, Rishi Bommasani, et al, ArXiv, 2024.**

I want to read more about marginal risk: In many ways, generative AI tools share the risks of any technology—they require testing for security, accessibility, and appropriateness to the task. A recent review of studies into the specific threats posed by open foundation models introduces the concept of "marginal risk" to help isolate where the risks of using these specific tools differ from prior technology and considers how the use of generative AI changes the impact or likelihood of a given threat.

"Artificial Intelligence Resource Center" **by the National Institute of Standards in Technology (NIST)**

NIST provides a risk management framework and playbook for the development and deployment of new AI-based tools.

"We Don't Actually Know If AI Is Taking Over Everything" **by Karen Hal, The Atlantic, 2023.**

This summary of the Foundation Model Transparency Index highlights just how much about current popular models' datasets, training, energy consumption, and policies for handling personal data remains closed off from public review and scrutiny, and puts that secrecy in the context of other technology industry norms. As election officials consider where and how to use these tools, keeping this opacity in mind may suggest marginal risks to explore and mitigate.

4

Worksheets for AI deployment

IMAGINING PRIORITIES

1 What are the top five challenges I want to address in 2025?

2 What are my five highest priorities for improving internal processes?

3 What recurring tasks or processes are the most effortful or time-consuming in my office? Where do I (or my team/staff) most need additional support?

IMAGINING EASY WINS

1 What are five types of content I can try generating or playing with that don't need sensitive data or processes?

2 What data do I have that would be simple to visualize and might spark new ideas to see charted or mapped?

3 What large documents, datasets, or sources of information would I want to be able to summarize, skim, or ask questions about more easily?

IDENTIFYING CORE REQUIREMENTS - *SAMPLE*

Process name: *Ballot proofing*

1 Description:

Comparing ballot design against documentation of ballot requirements for completeness and accuracy

2 Need or goal:

Provides staff a "second pair of eyes"

Add an extra layer of redundancy to current process

3 Inputs (data, information):

Ballot layout images (PDF)

Precinct-specific lists of races and candidates (spreadsheets/CSV's)

Ballot formatting requirements (text)

4 Activities:

For each precinct, confirm that all applicable races and candidates are included

Confirm that the layout complies with requirements for formatting

5 Outputs:

List of errors or omissions and corrections required

Confirmation of accuracy

CHOOSING

Process name:

1 How might this process shape decisions that could affect a voter's ability to cast their ballot?

2 How could variations in the tool's performance cause harm?

3 How might inaccuracies or unreliability in how the AI handles this process pose a threat? Could the tool introduce bias?

4 Could variable responses to very similar prompts or inputs pose risks? What might those look like?

5 How many people will interact with this tool? Do we have the capacity to train them individuals in how the tool functions and how it may confound basic assumptions about software applications?

PLANNING

Process name:

1 Who else is using this tool or process?

2 What is their experience using it? What lessons did they wish they'd learned before implementation?

3 Who are the staff who will interact with this tool or process?

4 Have they been trained to understand

- The task(s) it performs
- The process for using the tool
- Its user interface (including prompts)
- The data the tool uses
- Its expected performance
- Common risks or issues

5 What is the process for correcting and reporting errors or failures? Do staff know this process?

6 What contingency plans are in place for addressing errors or failures?

7 Who is responsible for reviewing failure reports and ensuring accountability?

8 What are the requirements for human review of any AI output? Who is responsible for that review?

9 How will you describe the decision to use this tool to the public? What is the plain-language description of how this will contribute to your operations?

10 Does that public communication include an option for public feedback? If so, who is responsible for reviewing that feedback? What decisions or changes could public feedback still shape?

The Elections Group

electionsgroup.com