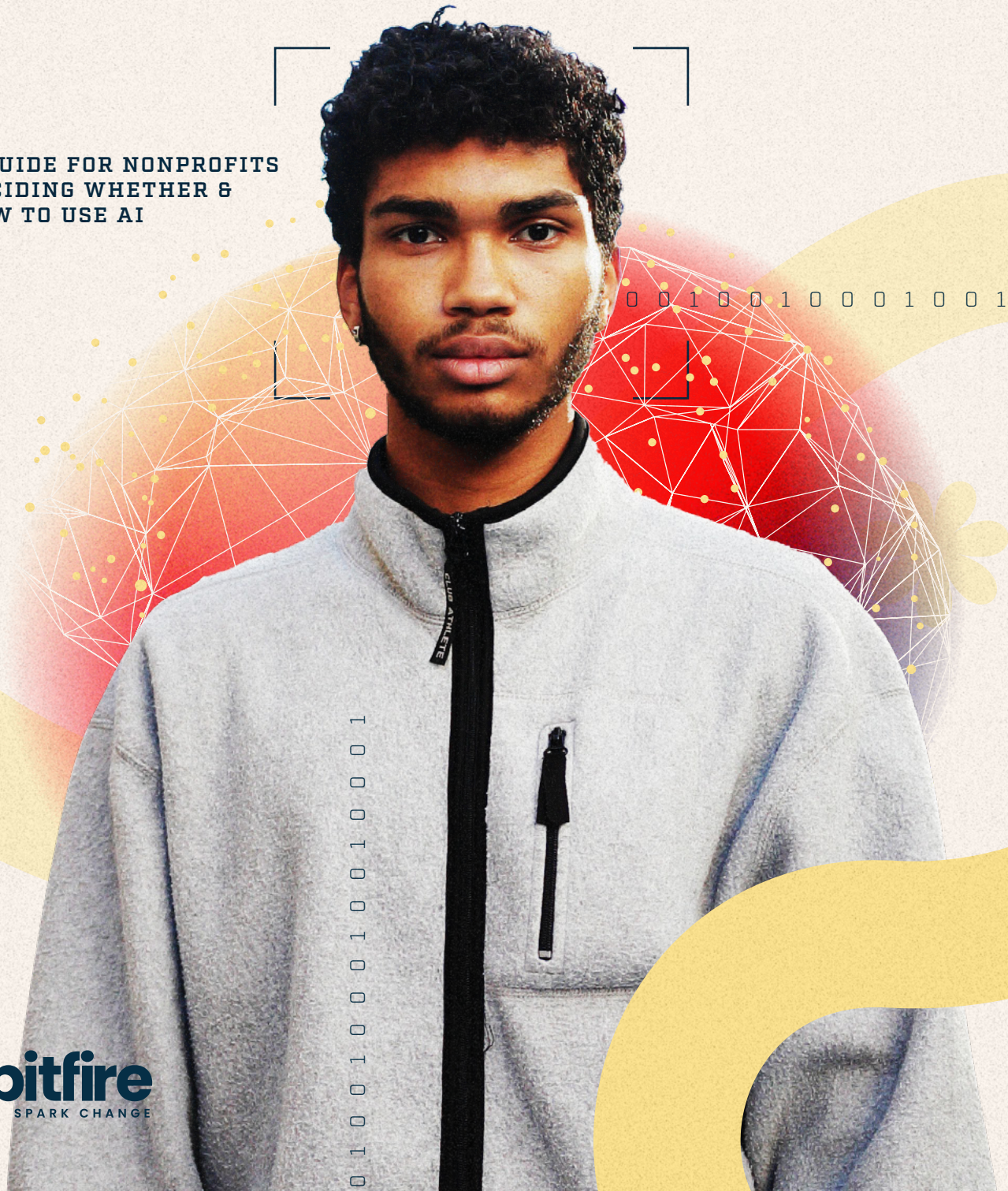
CTRL+ALT+ETHICS

2025

A GUIDE FOR NONPROFITS DECIDING WHETHER & HOW TO USE AI

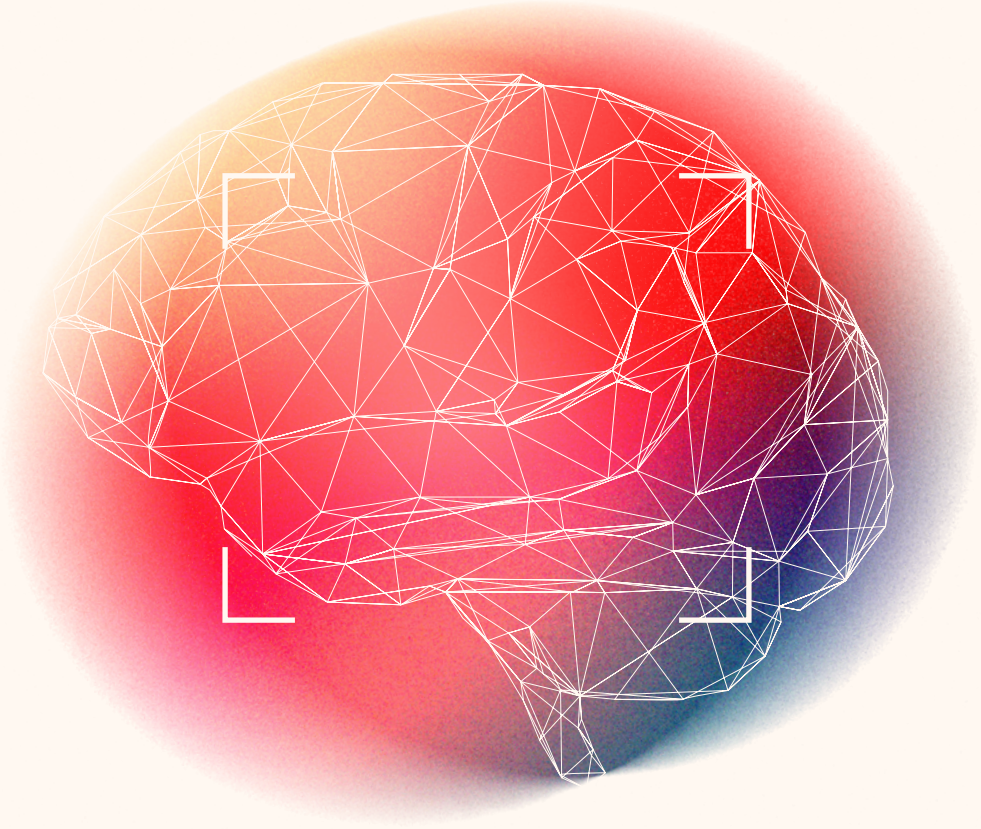


spitfire
SPARK CHANGE

spitfire

Contents

A GUIDE FOR NONPROFITS DECIDING WHETHER & HOW TO USE AI



Purpose 03

Primer on AI 04

Harms and Possibilities of Generative AI 06

Questions to Consider 11

Spitfire Resources 13

Purpose

With the sudden and rapid application of artificial intelligence (AI) across industries, many activists, social movement leaders, foundations and others find themselves with a critical question:

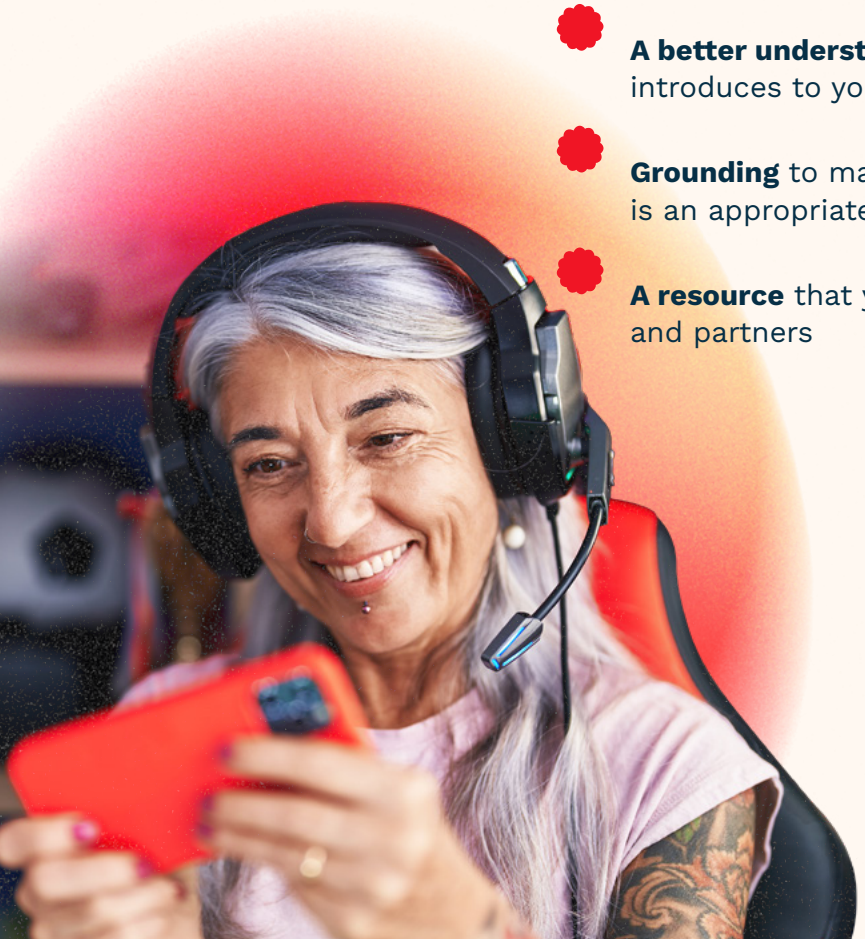
Where does AI fit into our communications strategy?

Individuals and organizations across the board are using this technology — from government agencies for surveillance purposes to marketing professionals to optimize their work. Currently valued at $100 billion in the U.S., the market for AI is expected to grow twentyfold by 2030, and most business leaders believe AI is critical for success.

Beyond strategic goals, everyone who is a communicating leader should have access to the critical, relevant information needed to decide on its use in their organizations. Spitfire has seen technology revolutionize the way we connect a time or two, and we know that there is a lot to learn, consider and apply when thinking about how AI will impact your team dynamic, the work product and the overall connection with key audiences.

This guide is made for the communicating leader who doesn't have time to drink from the firehose of daily AI updates, innovations or breakthroughs. The guidance in this checklist is to give your team:

- **A better understanding** of the pros and cons that AI introduces to your work and communication
- **Grounding** to make an informed decision about whether AI is an appropriate tool for your organization
- **A resource** that you can use when onboarding new staff and partners



If you have any questions or would like to discuss specific scenarios your team is facing with regard to AI, get in touch with us at **info@spitfirestrategies.com**

Primer on AI

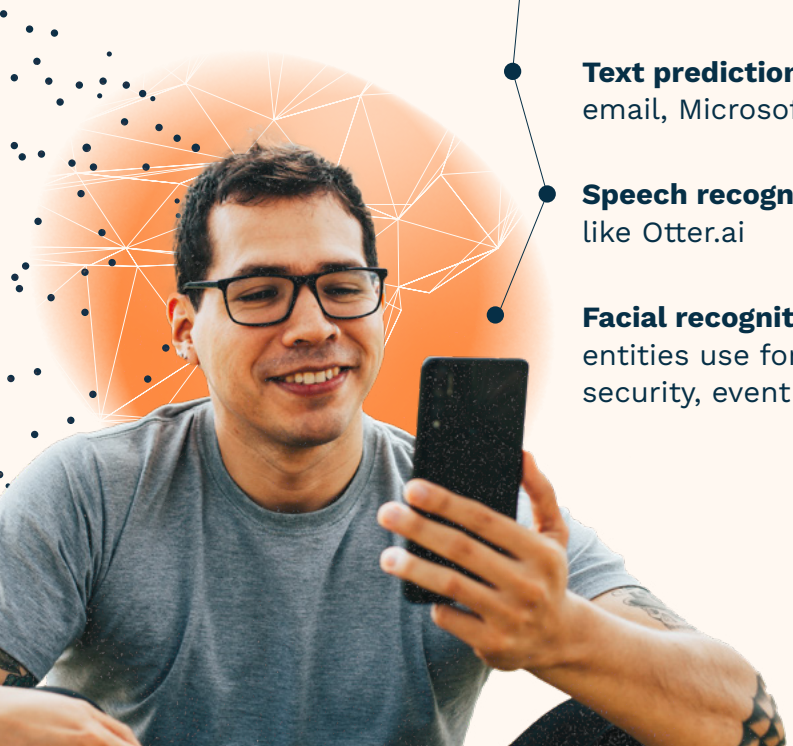
Types of AI

AI is the use of computer systems to simulate human mental processes, including interpreting and generating language. There are three broad categories of AI:

Predictive

Predictive AI uses machine learning — a type of AI that imitates how humans learn — and statistical analysis to analyze available data and predict future outcomes or trends. This model manifests through tools including:

- **Social media algorithms** to anticipate what you'd want to see based on your online and demographic data
- **Algorithmic decision-making tools** to determine whether you're denied or approved for a mortgage loan or rental property, a job that you applied for, public benefits (Medicaid, SNAP, etc.) or as another example how law enforcement uses these tools to determine the likelihood that someone will commit a crime
- **Text prediction** that completes your sentences, such as on email, Microsoft Word and Google Docs
- **Speech recognition** such as Siri and voice transcription tools, like Otter.ai
- **Facial recognition** that government agencies and private entities use for public surveillance, including at airport security, event venues, workplaces and apartment buildings



Today, many people associate AI with generative AI tools like ChatGPT, but predictive AI technologies — such as facial recognition and audio detection tools — are often used for public surveillance and crime prediction. This has furthered the cycle of biased overpolicing, particularly in Black and brown communities, and can have catastrophic consequences for individuals when these tools fail. In other use cases, such as the design of social media algorithms built to maximize attention, it has contributed to increased polarization, has created digital echo chambers, and has facilitated the spread of mis- and disinformation.

Analytical

Analytical AI is based on statistical machine learning that uncovers patterns and trends to help users make informed decisions. Businesses as well as the finance and healthcare industries use this model to:

- **Identify** what products and customer segments are generating more sales.
- **Segment** customers and personalize marketing campaigns.
- **Predict** stock market trends and assess credit risks.
- **Diagnose** diseases based on medical imaging and patient data.

Generative

Generative AI responds to a user's requests (made through "prompts") to create "new" content, including images, videos, software code and text. These models are trained on massive amounts of data and, from this, draw patterns to understand what the user is asking for and use this information to create content. Generative AI models are normally used by chatbots and image bots, which are powered by large language models (LLMs) — systems that communicate through language in a human-like way. **Examples of tools that use this model include:**

- Owned by OpenAI, **ChatGPT** is a chatbot and language model designed to engage in conversation through prompts. Anthropic provides a similar service called Claude.
- Also part of Microsoft's suite of products, **CoPilot** is a chatbot that helps support work tasks, from brainstorming to coding.
- **Google Gemini** is a suite of AI models that power several services and products like Gemini, a chatbot similar to ChatGPT.
- **Stable Diffusion** is a tool that uses deep learning — a type of machine learning that uses artificial neural networks (a model that imitates the structure of the brain) to learn from data — to generate photorealistic images from text prompts.

Harms and Possibilities of Generative AI

Harms of AI

With public discourse around AI lauding this technology as both powerful beyond human capabilities and able to **solve** some of our biggest challenges and also calling it out for causing great **harm** and **disrupting** our way of life, it's no wonder that **most of us** are cautious about it.

It is important to demystify the hype cycle and apply AI on our own terms. Below are the known, common challenges with this technology as well as what may be useful ways to apply it that are in-line with our values.

Furthers the climate crisis and injustice

The rapid development and use of generative AI tools is **directly contributing to the climate crisis, because these tools rely on tons of energy and water to operate.** These tools are powered by computer systems and software housed in spaces across the country, known as data centers, and these systems and data centers are constantly running. This creates a cycle of conscious heating and cooling, because massive amounts of air conditioning and water are required to offset the energy demand and prevent data centers from overheating.

While data centers are not new, the inflated investment in generative AI tools — a direct result of the marketing hype cycle that Big Tech companies led — is leading to increased development

of data centers to keep up with the perceived demand for this technology. This not only disrupts our climate, it **negatively impacts people who live near a data center** — from dealing with increased noise from cooling fans to receiving a **higher home electric bill** to offset the energy costs. Recent **research and analysis** suggests that the growing demand for data centers to expand generative AI will have a **disproportionately harmful impact on rural and historically Black communities.**

Trades creativity, culture and critical thinking for speed

Much of the language used to market generative AI suggests that these tools can improve efficiency and productivity. For resource-constrained organizations and advocates, this can make the technology appear to add value to their work, because less time drafting can potentially create more headroom for creativity and strategy. While these tools do generate content quickly, the reality is that what they produce is a regurgitation of text scraped from the internet.

This results in content that might read well-enough but that is removed from your human imagination, creativity, and unique voice and perspective. It **homogenizes and flattens lived experiences, identities and the communities** that make up our world. There is also some emerging evidence that routine use of generative AI tools in work settings **reduces critical thinking and makes people more prone to mistakes.** This is an emerging field of study and one that we'll keep an eye on.

It's true that for some generative AI models, it is possible to build a custom tool that is tailored to generate content that better reflects your voice, but this process is **labor intensive and comes at an additional cost.**

Built on stolen data and human work

The most readily available and integrated generative AI tools (e.g., ChatGPT, Google Gemini, Meta Llama, etc.) are built using LLMs — the underlying technology that powers these tools and their ability to recognize, translate and generate text at a large scale. These tools are designed for broad use, which means they require massive amounts of data to function even moderately well. To meet the data demand, technology companies building these tools **scrape available content from the internet — without consent or regard for copyright laws.**

Technology companies are also **not required to be transparent about the data** they use to train these tools. This allows them to take advantage of any content that is publicly accessible on the internet and feed it into their technology **without compensation for the creators they're stealing from.** This creates ethical challenges around the use of generative AI tools that are designed using people's work without consent and compensation.

Speaks with authority, but results are not trustworthy

Like any technology, generative AI tools are **likely to make mistakes.** These tools generate responses that can appear to be thoughtful, comprehensive and accurate, but **they are not reliable sources of truth and information.** This is because LLMs are a technology designed to generate text and images based on patterns identified in massive amounts of data, not human consciousness and when returning results about complicated issues, **the AI can misinterpret or simplify nuance.**

Ahead of the 2024 election, research found that when asking questions about the election and voting using two of the largest generative AI tools on the market — ChatGPT and Gemini — the tools generated responses that were **wrong 27% of the time.** Furthermore, because we don't know what data is used to train the models, the results may offer up propaganda or censored information, **with many outlets reporting about DeepSeek.**

While generative AI tools can support with quickly developing materials, it's important to not rely on the content these tools produce. You should always check that the information they produce is accurate.

Possibilities of AI

Most of us are using and interacting with AI every day. But how can we effectively and strategically use it in our work? Deciding whether to use AI will depend on your personal consideration of tradeoffs listed above. Below we explore a few hopeful use cases and considerations, including some experiments we've run. In general, we focus this section on uses of generative AI because it is in high demand. Generative AI has gotten the most attention recently because of ongoing efforts to add it to a wide variety of products and because of the release of several popular AI tools, including ChatGPT, DeepSeek and Claude.

Capturing the "flood of data"

When the Puerto Rico Supreme Court made publicly available every use-of-force report, the organization Kilómetro Cero had a problem. It was flooded with paper reports and didn't know how to process them all. Enter the Human Rights Data Analysis Group (HRDAG), which **leveraged machine learning and big data to make this huge collection of 10,000 reports available and searchable.** With carefully constructed machine learning techniques, HRDAG was able to not only make it possible to look for specific words and phrases in these records, but their investigators could also identify details like which officers were involved in a complaint or investigation and click through to see other cases where that officer was involved. This tool helped investigators enhance accountability by understanding patterns in the department's use of force to a degree never before possible.

The language machine: Small model translation that outperforms big LLMs

One of the areas where AI shows some promise is in the area of translation. LLMs, like ChatGPT, are billed as a one-stop shop for translation between any number of languages. And yet, these tools often don't capture the nuance of language that human translators understand intuitively, and they struggle to accurately translate underserved languages where tech companies don't invest enough in expertise or have adequate data to capture. Africa alone has hundreds of underserved languages, which are at risk of being shut out of the global information ecosystem as machine translation becomes more common. **Huniki** is a **DAIR Institute** project to build super **high-performing small language models that outperform the tech giants.** Working in languages like Tigrinya and Amharic, Huniki is making powerful language tools available by partnering with local tech companies.

Conversational learning on your issues

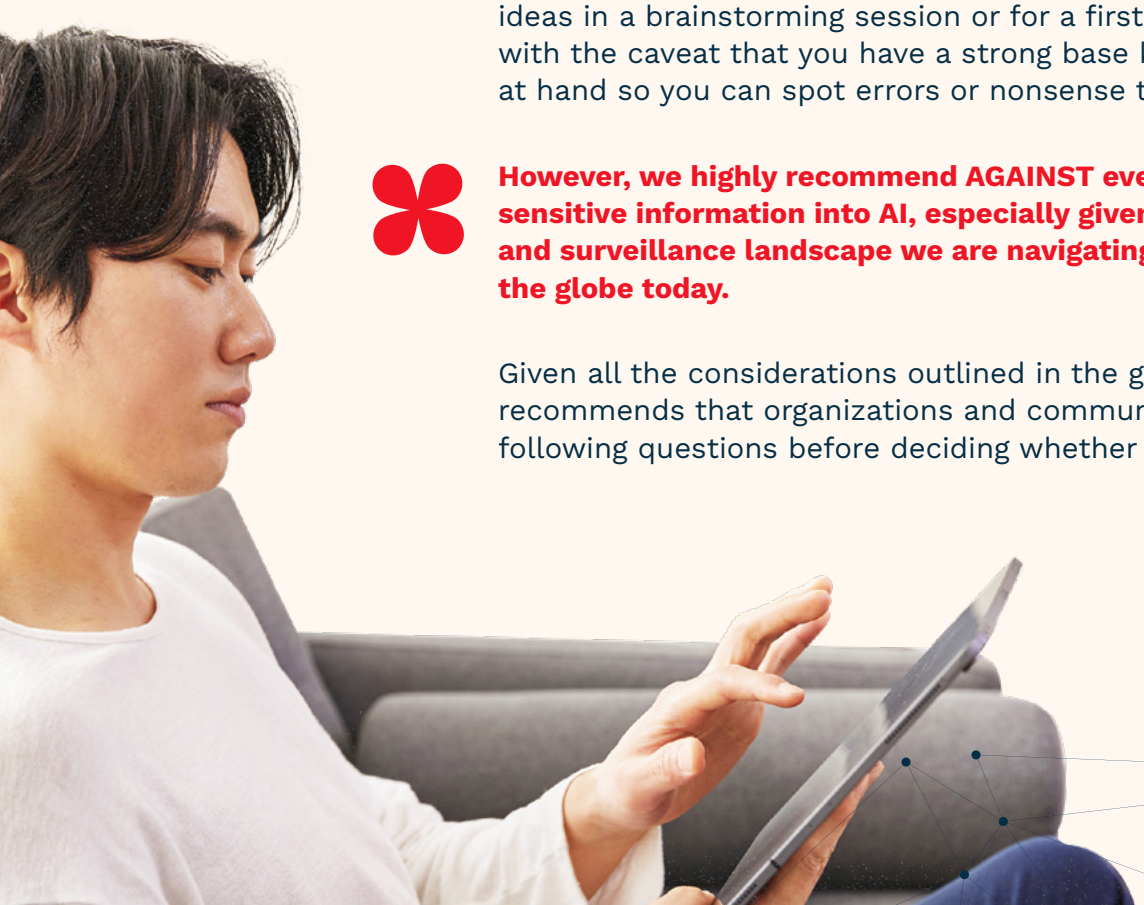
While the process of customization can be time-consuming and often requires partnership with a skilled technologist, it is possible to make your own version of a generative pre-trained transformer that helps people explore an issue in a friendly and conversational way, with a set of content that you curate so you know the results will be based in solid data. For example, Spitfire **created this tool** that **helps people learn skills related to bridge building, mediation and handling difficult civic conversations.** This tool was part of a project called **Best-Case Scenario**, which aimed to empower people to be more civically involved, and the tool synthesized dozens of resources and hundreds of pages of guidance we curated. We believe that more civic engagement, helping people find common ground and civility, and decreasing polarization are worth the costs across environmental and social concerns. The benefits of this type of model are that people can explore issues at their own pace and in whatever order they choose. This type of model has the potential to lower the barrier to participation on your issue. Organizations and individuals should not rely on this model to give legal advice on issues where a person's safety and livelihood are at risk.

Questions to consider before using an AI tool

Whether you are interested in making your work more efficient or you are conflicted about using AI in any form, with this guide, Spitfire aims to lay out various questions for you to consider when deciding how and when to use AI. Generative AI tools may facilitate ideas in a brainstorming session or for a first draft of a document, with the caveat that you have a strong base knowledge of the issue at hand so you can spot errors or nonsense the AI may include.

However, we highly recommend AGAINST ever introducing any sensitive information into AI, especially given the current political and surveillance landscape we are navigating in the U.S. and around the globe today.

Given all the considerations outlined in the guide, Spitfire recommends that organizations and communicators consider the following questions before deciding whether to use AI.



Reputational risk

Are you prepared for backlash about the climate and ethical implications of an AI tool? Will it erode trust with your partners who are aware and against those impacts?

Do any of the harms conflict with your organization's stated values in a manner that is defensible, for example, conservation, labor, or economic and/or racial justice?

Clarity on use case

What AI platform are you considering using? How, exactly, will you use it?

Whom do you envision using the AI tool? Are they aware of all the risks and potential harms outlined in this guide?

Does your use of AI provide a service or meet a need that you otherwise could not meet? Is the potential gain in efficiency or speed worth the risk of harm (e.g., environmental costs, workforce disruption, etc.) and/or ethical tradeoffs (e.g., data theft and threat to creative rights, labor conditions for data workers, etc.)?

Reconcile capacity

If you decide to use AI, do you have a plan in place to ensure that the results are accurate and ethical? How will you determine whether there are important elements missing?

Are there tradeoffs in critical thinking that might weaken your team's long-term ability to work effectively? Can your team identify harmful narratives or bias in the AI output results?

Security and organizational risk

Is the data you plug into the AI tool protected and private? Are you prepared for the implications of others discovering that information and weaponizing it?

Do you have security measures in place to prevent bad actors from hacking into the tool and changing reference sources or using the data you input against you?

Do you know where the data informing the AI tool comes from, and can you guarantee it is accurate and free from bias? How often is the data updated and by whom?

REPORT CO-AUTHORS

Spitfire Team

A vice president, **Gabriel Rodriguez** leads major projects with a particular focus at the intersection of technology, higher education and democracy. This includes Spitfire's work on emerging issues such as AI. His specialization involves addressing complex communication challenges, including those within collaborative ecosystems, and spearheading narrative change efforts. He also supports public agencies in building their capacity to think and act through a lens of strategic communications

A director at Spitfire, **Lawrence Mason III** serves on the firm's Artificial Intelligence Ethics Committee and has supported clients dedicated to working for equity in and around innovative technologies, including AI for the People. He brings a decade of experience and specialties, including speechwriting and messaging.

A director at Spitfire, **Gaby Connor** serves on the firm's Artificial Intelligence Ethics Committee and supports leading technologists, organizations and advocates working to protect people from technological harms. She works closely with the public interest technology field, helping to shape messages that build broad public understanding about AI's far-reaching impact on our society. Her work has also included counseling field leaders through congressional meetings on AI.

A senior account executive at Spitfire, **Alex Williams** serves on the firm's Artificial Intelligence Ethics Committee and has worked at technology companies, with her role spanning executive and internal communications. Having supported leaders in sharing complex topics to broad audiences — including collaborating with machine learning engineers at Rakuten to share how the company is using artificial neural networks in its e-commerce business to better serve customers — she now applies those skills to organizations working in the public interest.

A managing senior vice president at Spitfire, **Ellie Klerlein** serves on the firm's Artificial Intelligence Ethics Committee and loves to conquer new challenges and enjoys working with clients who are willing to experiment and try new approaches to take on society's toughest problems. She worked with the Ford Foundation to support the expansion of the public interest technology field, which recognizes that technology is never neutral and includes technologists who know it is never enough to say tech is "for good." We must always ask: "good for whom?"

Stay in Touch

Let's Spark Change Together!

Spitfire is a strategic communication and advocacy firm rooted in the core principles that everyone belongs and has the power to spark change. We believe in human rights for everyone. Everywhere. We believe it takes all of us. We're driven by social justice, and that doesn't stop when we leave.

If this resource sparked ideas for how to shape your voice or reach new audiences, we'd love to stay in touch.

Follow us on LinkedIn

Subscribe to our Newsletter

Check out more of our tools

Email us to connect

THANK YOU

2025

A GUIDE FOR NONPROFITS DECIDING WHETHER & HOW TO USE AI

spitfire
SPARK CHANGE