Guidelines for the Use of Artificial Intelligence in our Work

February 1, 2024

As people and organizations grapple with how to best use (and limit) rapidly evolving Artificial Intelligence (AI) technologies, here are a few principles to guide Evidence Action's adoption of such technologies as we *think big and act urgently* to reduce the burden of poverty for hundreds of millions.

Evidence Action permits staff to experiment with AI tools in their work. In doing so, you are encouraged to *iterate, again*, with the use of AI while understanding that formal approval is needed to institutionalize an AI tool into a system or process.

1. AI tools have the potential to augment our work by enhancing analysis, content generation, and other only-just-imagined uses. Try them out!
2. Prototype the integration of AI tools into workstreams in consultation with supervisors and IT and with careful attention to results and risks. Approval by the AI Steering Committee (see below) will be required to institutionalize an AI tool into a system or process.

Guidelines for use

1. **Humans make decisions and we will remain human-centered in our use of AI.**
   a. While AI tools might assist us in the analysis of data, creation of content, or performance of tasks, humans–rather than AI tools–will make decisions.
   b. We view AI tools as means to amplify and augment human abilities and as another avenue to *challenge convention* in pursuit of our mission.
2. **Everyone is accountable for their work products.**
   a. If you use AI tools in your work, you are still responsible for the final product and any errors therein.
   b. Use due care in evaluating products created with the assistance of AI tools.
3. **Be transparent in your use of AI for work.**
   a. As appropriate, acknowledge when and how you use AI tools in your work. Specific requirements for acknowledgment will be left to supervisors to establish based on the nature of the work, the significance of the contributions of specific AI tools, and our existing policies (more on these below).
   b. You are encouraged to share promising applications of (or concerns about) AI tools with your colleagues, for example via the #EvAcGPT Slack channel.
4. **As Evidence Action employees, we all have a responsibility to protect Evidence Action's confidential and proprietary information, including organizational intellectual property.**
   a. Different AI tools use your inputs differently and might expose information you input and/or use it to generate results for other users.

b. So, do not input confidential, internal, financial, personally identifiable, or proprietary (e.g., trademarks, unique programmatic practices, etc.) information into public models.
c. This includes Evidence Action's intellectual property - e.g., images, trademarks, brands, etc. - given how some AI tools use (and reuse) these sorts of IP assets.
d. Consult our IT Team for guidance on what tools might be appropriate for use with confidential or proprietary information or data.

Risks of use

Be knowledgeable and informed about the risks and limitations of generative AI tools. While these will continue to evolve, some risks to consider are:

1. **Generative AI tools regularly "hallucinate,"** providing false but plausible and correct-seeming information. This information can seem persuasive but could be entirely fabricated or inaccurate.
2. **AI tools may amplify biases** via their output because they are trained on data sets derived from the internet, and their data may include overrepresented, racist, misogynistic, and other biased views.
3. **Using generative AI tools might create reputational risks for the organization** when it produces work containing inaccurate information or for purposes that people might consider disingenuous or offensive (e.g. to write a "heartfelt message").
4. **There are risks and debates around ethical storytelling and AI**. For example, AI-generated but realistic images might be used to evoke sympathy, without representing real images or people.
5. **There are risks and debates around intellectual property** and how generative AI tools have incorporated the text and work of creators without compensation.

Preliminary Governance of Adoption

1. While the above guidelines and risks are not comprehensive, they are intended to provide practical guidance on our use of AI tools, which will evolve. Some tools, such as proofreading plug-ins (e.g. Grammarly) or fundraising tools are already in regular use and include AI-powered features; these guidelines are not meant to prevent their use. As noted above, consult your supervisor and our IT team for guidance around specific applications of AI tools. We will need to *iterate, again* together as we seek to responsibly adopt these technologies in furtherance of our mission.
2. These principles supplement and support rather than supplant the existing policies, regulations, and laws we already follow regarding the safeguarding of data, the use of digital tools, and how we communicate and attribute (most directly our Information Management, Confidentiality, IT Acceptable Use, and Communications policies and Privacy Statement).
3. If you have questions regarding these principles or seek approval to institutionalize an AI tool into a system or process, please send them to the Chief of Staff (zach.watson@evidenceaction.org), who will receive them on behalf of the Chief Operating Officer (COO) and refer them to an AI Steering Committee comprised of the COO, General Counsel, Director of IT, and relevant GLT Members to review any decisions to adopt such tools.
4. Finally, if you see something, say something: this is a rapidly evolving technology and our understanding of its use and capabilities is changing constantly. If you believe these principles are out of date or are missing something, please feel free to reach out to our Chief of Staff (zach.watson@evidenceaction.org) to suggest changes.