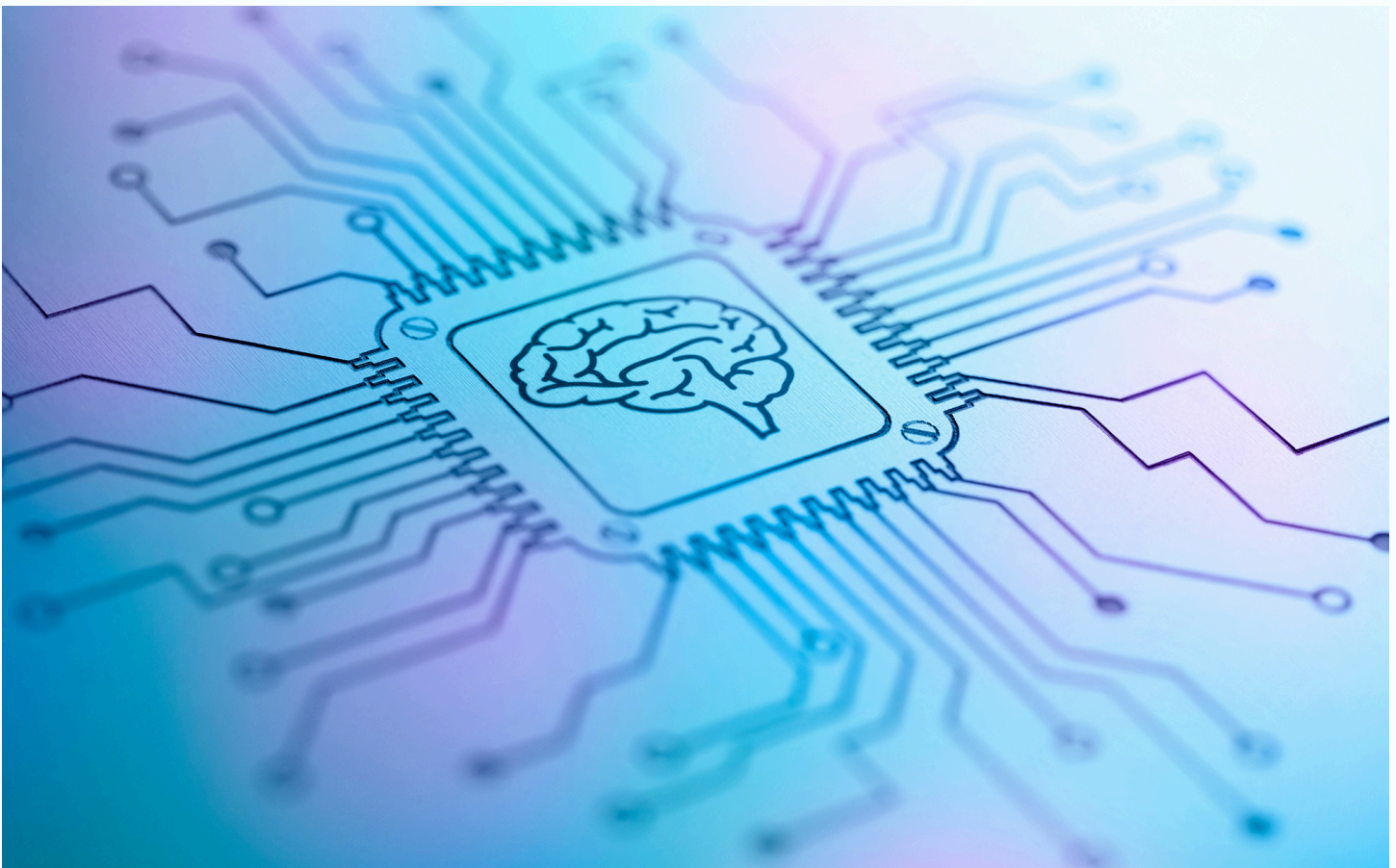# Artificial Intelligence: Its potential and ethics in the practice of public participation

A Research Paper by the International Association for Public Participation Canada

A computer chip with a brain in the center to represent artificial intelligence.

Written by:

Morgan Boyco MCIP RPP & Dr. Pamela Robinson MCIP RPP
April 2024

iap²
CANADA

INSPIRING BETTER
DECISIONS TOGETHER

Table of Contents

Executive Summary

Public participation practitioners are committed to exploring innovative ways to support democratic decision-making. Recent advancements in artificial intelligence tools offer potential avenues to enhance efficiency, creativity, and inclusivity in our practice. Yet, alongside these opportunities, there are also significant concerns about bias, accuracy, and risk. This research brief explores how P2 practitioners might encounter AI tools; examines the risks, opportunities and uncertainties these tools present; and offers ethical considerations to guide their use.

Introduction

This research paper provides an orientation to artificial intelligence (AI) technology for public participation (P2) practitioners. We explore the ways AI might intersect with P2 processes today, raising opportunities, uncertainties, and risks associated with its use. We offer a practical guide for P2 practitioners navigating these AI encounters by sharing some ethical considerations for using AI tools.

The field of artificial intelligence encompasses a wide array of technologies with foundations in computer science, philosophy, mathematics, neuroscience, linguistics, and more. AI "is simultaneously high risk, low risk and everything in-between" (Boucher, 2020, p. 1). And at this point, much of what is written about the potential uses of AI remains aspirational or highly speculative.

For most of us, AI has become synonymous with generative AI, backed by machine learning, which produces novel content, including text, images, audio, and video, based on an algorithmic remix of existing data. Widely accessible and general purpose generative AI tools like ChatGPT have ignited imagination as well as concern and debate over the implications and ethics of these new technologies. There are many claims emerging about the potentials of generative AI to power P2 processes (Bright et al., 2024; Marmolejo-Ramos et al., 2022).

A Guide to AI Terms

AI, GPTs, NLP, LLM, AGI? A handy guide to AI-related terminology you'll encounter in this paper and elsewhere.

Term	Definition
Artificial intelligence (AI)	No universally accepted definition. Many refer to AI as computer systems performing tasks that would typically require human intelligence, such as learning, reasoning, and decision-making. Avoiding ambiguity or attributing human-like qualities to computer systems, the International Organization for Standardization defines AI as: "a technical and scientific field devoted to the engineered system that generates outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives" (ISO, 2024).
Generative AI	Using models trained on vast amounts of data to "learn" patterns and create novel content, such as text, images, audio, video, and computer code, in response to specific prompts.
Machine learning (ML)	A set of techniques underpinning many recent advancements in AI, referring to the development of computer systems that use data and algorithms to perform a task, with improving accuracy over time.
Neural networks	Machine learning algorithms modeled on the human brain's layered and interconnected nodes, designed to process data and identify patterns (Brown, 2021).
Natural language processing (NLP)	The field of AI that creates systems for analyzing and processing human language and producing text that mimics human language patterns.
Large language model (LLM)	A type of AI technology that acts as a "next-word prediction engine," trained on enormous volumes of internet-based resources able to produce human-like responses to natural language queries (Mearian, 2024).

GPT	(as in ChatGPT) (Generative Pre-trained Transformer) A type of machine learning model pre-trained on extensive text datasets and designed to generate new text in response to a prompt by predicting subsequent words in a sequence.
Multimodal AI	Machine learning models designed to handle and generate diverse data types, including text, images, audio, and video.
Prompt engineering	Creating effective prompts for generative AI tools to appropriately guide output toward the user's desired response.
Algorithmic bias	When AI systems produce results that unfairly advantage or disadvantage certain groups due to erroneous assumptions in the algorithm or biases in the input data.
Hallucination	When an AI generates inaccuracies or results that are unrelated or misaligned with the input objectives due to limitations in its training data or inability to contextualize information properly.
Alignment	The process of encoding human values and goals into AI systems, such as LLMs, to make them as helpful, safe, and reliable as possible (Martineau, 2021).
Artificial general intelligence (AGI)	An entirely hypothetical form of AI that would match or exceed human-level capabilities and intelligence.
AI Hype	Exaggerated claims and excessive enthusiasm surrounding AI technology development, which could lead to distorted or unrealistic perceptions of its capabilities or inappropriate usage (Siegel, 2023). Also using the term "AI" to describe only tangentially related products or services (Rosenberg, 2023). Includes the critique that focusing on hypothetical risks (like world-ending doomsday scenarios caused by AGI) diverts attention away from real and current harms of AI (Helfrich, 2024).

Encounters with Artificial Intelligence in Public Participation

In the categories below we frame where AI could be encountered, either as a practitioner using AI technology to support engagement, or how AI might be used by other actors to impact the process. In a strong caveat to this section, we make no claims about the accuracy or appropriateness of any of these tools (we remain wary of AI hype). In subsequent sections, we frame technical and ethical considerations for whether AI should be used in any given public participation process.

AI in Research, Planning, and Brainstorming

Conducting background research and careful planning of a goals and objectives-driven engagement process are the essential first steps in an effective public participation process. Various AI technologies could have a potential role at this learning and planning stage, which culminates in the development of a public participation plan.

AI-assisted web search is able to handle more complex queries and provide more tailored, conversational responses using natural language processing technology. Many P2 practitioners are likely already using these tools for background research and increasingly these technologies will be integrated into the standard web search and other productivity tools we use everyday. Additionally, AI-driven data mining and analysis of large data sets could support insights that guide the tailoring of engagement strategies. And used as a brainstorming tool, generative AI can assist in planning engagement activities and tools. In "conversation" with tools like ChatGPT, practitioners can move from engagement objectives to the development of coherent discussion templates or survey questions (AI-assisted survey development is already built into popular platforms such as SurveyMonkey and Typeform). Other specialized AI tools used for marketing and customer journey mapping could generate "personas" to help P2 practitioners better understand intended audiences and tailor engagement activities accordingly (delve.ai, for example).

Three Common Uses for LLMs

The <u>London (UK) Office of Technology and Innovation</u> offers three example tasks that generative AI text creation tools can help public servants with:

- Summarization (e.g., paste a long document and ask for a summary)
- Initial ideation (e.g., what is a good structure for this type of document?)
- Gap-checking (e.g., what are three persuasive counter-points to something I have written?)

Ethical Considerations: While relatively low-risk, users must still be aware of the potential for bias and inaccuracies in the outputs for these tasks and the inability of general purpose tools like ChatGPT to capture local context well. The quality and veracity of the tool's output may vary between the free and paid-access versions. Transparency of the logic and responsibility for particular outputs are difficult if not impossible to trace.

Case Study

The "Pair" app developed by the Singapore government is an AI chatbot based on ChatGPT, assisting thousands of civil servants with writing and research. The app enables secure usage of large language models, acting as a writing assistant and idea collaborator within the government space (Min, 2023).

AI in Communications

Authentic, ongoing, and appropriate communications are foundational to any meaningful engagement process. LLMs can obviously produce large volumes of text-based content quickly and with minimal effort but their output requires fact checking and sometimes rewriting to sound more human. Still, these generative AI tools could help streamline some writing tasks, such as first-draft development of press releases or plain language summaries, or help with brainstorming.

Multimodal AI tools, meanwhile, can produce multimedia content like graphics, audio, video, and 3D models. Used correctly, these technologies could speed up workflows and aid in simplifying complex information into understandable and engaging formats. Finally, numerous AI-backed technologies are already commonly used to improve accessibility of communications. In P2 processes, real-time speech-to-text, text-to-speech, image recognition, and language translation services powered by AI could make information more accessible at scale to more diverse audiences (although these tools still struggle with accuracy, nuance, and context).

Case Study

Through the creation of open-source datasets in local languages, the Indian government's language translation initiative, Bhashini, leverages AI technology to enable the delivery of digital services to a diverse population speaking over 121 languages. Through crowdsourced contributions and validation initiatives, Bhashini underpins the development of AI models for tasks like translation, speech-to-text, and text-to-speech (Bordoloi, 2022).

AI for Interaction and Dialogue

Advancements in machine learning and natural language processing are relevant for the dialogue processes at the heart of public participation. Tools for moderation and toxicity screening can help ensure discussions happening on digital engagement platforms remain constructive and respectful. Chatbots and even AI-powered digital facilitators ("chat assistants" or "crowd discussion support agents") propose to automate feedback collection, provide immediate responses, and support dialogue participants, even helping to guide group conversations (Argyle, Busby, Gubler, et al., 2023; Hadfi et al., 2023; Ito et al., 2020). And while significant quality gaps remain, automatic transcription services and real-time language translation could help make these dialogues more accessible to a broader, more diverse audience. Meanwhile, AI-supported fact-checking tools, such as Full Fact AI, are helping journalists maintain the integrity of information shared during political debates.

Case Study

Researchers developed an NLP conversational agent and deployed it with Afghan citizens discussing the fall of Kabul in August 2021.The experiment sought to demonstrate the potential for conversational AI to amplify the contribution of women in an online dialogue about an emotional and contentious topic and in a typically male-dominated space (Hadfi et al., 2023).

AI in Analysis and Decision Support

The advanced computational power of AI has strong potential for analysis including climate modeling for adaptation planning or traffic analysis and real-time road management for urban transportation planning. The allied professions including urban planners, engineers, and architects that P2 practitioners often work may increasingly turn to AI analysis in their city building work (Zheng et al., 2023).

People standing in front of a computer screen with an overlay of graphs.

AI technology's purported ability to synthesize and support analysis of large volumes of unstructured data has obvious applications to the resource-intensive activity of community input analysis in civic engagement processes. NLP-backed tools can be applied to data such as open-ended survey responses, public comments from a variety of sources, interview data, and focus group transcripts to automate the development of summaries and extract insights. These AI-powered analysis functions are becoming integrated into tools commonly used by P2 practitioners, like SurveyMonkey and Qualtrics, as well as in digital engagement platforms, such as CitizenLab and PublicInput. Other AI tools promise to support trend forecasting and sentiment analysis, providing practitioners with additional inputs for examining community trends and feedback. Community members could also deploy AI tools themselves, empowering them to better access and decipher technical and expert-driven processes.

Case Study

In 2019, the Youth for Climate initiative utilized CitizenLab's NLP technology to conduct an exploratory analysis of over 1,700 citizen contributions in order to convert ideas into impactful recommendations efficiently. Using the platform's automated data analysis feature, guided by human analysis and oversight, the team was able to process thousands of ideas within a relatively short time frame (Cuau, 2019).

Synthetic Democracy?

Researchers in marketing, psychology, and social sciences have been asking the question, "Should large language models replace human participants?" (Crockett & Messeri, 2023; Grossmann et al., 2023). Experimentation is underway using LLMs to supplement or even fully replace human participants in research with so-called "silicon samples" (Argyle, Busby, Fulda, et al., 2023). One group has tested whether conducting semi-structured interviews with ChatGPT "could give insights into public opinions in a way that otherwise only interviews with large groups of subjects could deliver" (Dengel et al., 2023). It is a short leap to imagine the use of these "virtual publics" in various governance processes—a kind of crowdless crowdsourcing (Boussioux et al., 2023).

Ethical Considerations: While the ethical problems with synthetic democracy may be readily apparent to P2 practitioners, some decision makers or process sponsors may be lured by the promise of automation or the prospect of an efficient and frictionless route to "understanding" the public by using AI-powered proxies in place of real participants. P2 practitioners should be looking ahead to anticipate what to do about "synthetic democracy."

The current excitement, promise, and hype surrounding AI tools is high. Innovators and advocates see its potential to help us better gather and share information, be more efficient, find new insights, and automate tasks such as writing survey questions, meeting note-taking, transcription, answering questions, and passive monitoring of social media for sentiment analysis. So far, we have explored the many ways P2 professionals can expect to encounter AI across various aspects of their work. Use of these tools also gives rise to risks and uncertainties.

Risks, Uncertainties, and AI's WEIRD Problem

While there are many potential uses of AI tools for P2, several practical questions about reliability, appropriateness, and risk remain unanswered. Despite the promises of AI for tasks such as analyzing public input data, there remains a gap between the AI tools we have today and their practical application and abilities in P2 contexts. For instance, it remains to be seen whether AI tools can perform reliable analysis of citizen contributions to P2 processes, which are importantly context-specific, compared to human evaluation (Romberg & Escher, 2023). Deepening this challenge, the latest generative AI tools like ChatGPT seem prone to so-called hallucinations and offering inaccurate responses in a highly confident tone. For example, an AI-powered chatbot recently deployed by New York City to help businesses navigate local rules has been shown to confidently offer advice that would break the law (Lecher, 2024).

"In civic circles, a consensus is emerging that the current large language models are too unreliable to use in serious contexts without supervision" (Saperia, 2024).

A data error concept with a magnifying glass over an error icon.

Consider, as well, AI's WEIRD problem—where AI training data is predominantly based on Western, Educated, Industrialized, Rich, and Democratic (WEIRD) populations (Norori et al., 2021). These baked-in cultural norms and assumptions can significantly affect the performance of AI systems in non-Western cultures (Prabhakaran et al., 2022). This is a form of what's known as algorithmic bias, where the data used to train AI models results in output errors that unfairly benefit or disadvantage certain groups (Broussard, 2024; Kordzadeh & Ghasemaghaei, 2022).

What's more, the growing power, complexity, and enormous volume of data that modern AI systems are trained on is increasing the opacity of these systems. This lack of transparency coupled with a growing popular recognition of potential biases and inaccuracies in AI outputs are, rightly, impacting peoples' trust in their use. What will be the impacts on trust between practitioners and the public where AI systems are becoming intermediaries in P2 processes?

Generative AI tools are also trained, in part, on copyrighted material. Outputs could violate copyright laws while undermining creative economy workers by automating the production of content that would traditionally require human creativity and skill. Can P2 professionals

ethically use AI-based tools without a clear line of sight into the relevance, appropriateness, and provenance of AI training data?

Risks associated with AI can arise at any point in the AI lifecycle—from selection of training data, to user interface design, to AI governance and use. Researchers at Google's DeepMind, the tech giant's AI lab, have developed a taxonomy of 21 general ethical and social risks associated with large language models, which underpin many AI tools. They distill these numerous observed and anticipated risks into the broad categories listed below (Weidinger et al., 2022). And while some risks may decrease over time as the technology improves (e.g., improved accuracy, greener technologies), other risks are likely to emerge or grow (e.g., misinformation harms, privacy and data security risks). With the rapid pace of development, risk identification and assessment will remain a challenge.

Categories of Risks and Harms Associated with LLMs

Adapted from: (Weidinger et al., 2022)

Discrimination, hate speech and exclusion:
Perpetuation of harmful stereotypes and exclusionary norms, often reflecting biases present in training data, leading to discrimination and unfair treatment of marginalized groups.

Information hazards:
Privacy risks associated with inadvertent leaking of sensitive information held in training data.

Misinformation harms:
Harms perpetuated by the generation of false or misleading information.

Malicious uses:
The possibilities for generating disinformation at scale or the creation of malicious software.

Human-computer interaction harms:
The human-like dialogue interface of some tools could lead to overestimation of capabilities, privacy violations, perpetuation of harmful stereotypes, or other types of unsafe use.

Environmental and socioeconomic harms:
Direct environmental implications associated with high levels of energy consumption required to run LLMs, as well as the potential contribution to socioeconomic inequalities by automating jobs or undermining creative economy workers. There is also a more direct human cost behind AI development, with tech companies using low-wage labour, often outsourced to the Global South, to train AI models through manual data labelling (Taylor, 2023).

Public participation processes depend on trust between the participants and process facilitators. Given the myriad uncertainties and risks already noted, ethical use of AI for P2 means professionals need to assess and manage the various risks and uncertainties involved.

Ethical Considerations for AI Use in P2

While there are many ways we could use AI in our work, should we? What amount of automation is appropriate in which parts of the process? There is no one size fits all set of directives for P2 professionals. Workplace guidance on tool use, where it exists, is our starting point. But the unique nature of P2 work and its important place in democracy and building civic trust means that P2 professionals also need their own ethical framework to inform AI tool use.

For guidance, P2 practitioners can look to the longstanding community of researchers and practitioners across public, private, and civil society sectors focused on fairness, accountability, transparency, and ethics in machine learning and AI. There are also well-researched and detailed AI guidelines and risk management frameworks published by the Canadian and US governments, as part of the European Union's AI Act, and in a growing number of cities, including Seattle (Generative Artificial Intelligence Policy), Boston (Interim Guidelines for Using Generative AI), and San Francisco (Generative AI Guidelines). The American Association for the Advancement of Science has developed a detailed decision tree for the responsible application of AI. Ultimately, we can return to our own pillars for effective P2 processes, including the IAP2 Core Values and the Code of Ethics to help determine where the ethical line is between reasonable and responsible use of AI tools and an inappropriate level of automating the professional tasks associated with public participation work.

Using these and other resources, and considering the opportunities, uncertainties, and risks described in this report, we have framed a series of ethical considerations as a foundation for thinking through AI tool use in P2.

Recommendations for Ethical and Appropriate Use of AI in P2

Be aware.
AI is a powerful technology, creating both opportunities and risks of real harm, including ethical and ecological implications, reputational risks, and risks to the health of our democracies and how participation is conducted.

Be accountable.
As a P2 professional, you are ultimately responsible for whatever your AI tool spits out and the impacts of its use. You are accountable to yourself, your employer, and most importantly to the community members you work with. You should be able to describe the AI system's processes, explain the outputs to participants, and be ready for when participants contest results.

Evaluate bias, promote equity.
Does the system and its use enhance equity? Determine how you will spot biased outputs and mitigate those impacts. Consider how the tool use may exacerbate digital divides, or marginalize underrepresented communities. Strive to use AI in ways that promote inclusivity and equitable outcomes.

Understand ethical implications.
Research and get to know the social, economic, and environmental impacts of the AI tool you are considering. Consider the legal, intellectual property, or human rights violations associated with the data collection and training process. Engage in these conversations as a P2 professional.

Use appropriate tools.
Is the AI tool credible? Is it reliable for your purposes? Start with your objectives in mind and choose the appropriate tools. There's a good chance AI isn't the answer. Remember that most AI tools are commercial products—exercise caution when using them for civic processes.

Ask.
As stated in IAP2's Core Values, participants should have a say in how they participate. Seek diverse perspectives on the use of AI.

Always disclose.
Be transparent about where and how AI is used. Be prepared to welcome participation through other channels or change course.

Protect privacy and be cyber secure.
Does your organization already have AI or other relevant guidelines? Never upload private information into a text generation tool.

Be skeptical.
Would an AI application be the best and most sustainable approach? Regard LLM tools as language pattern machines, rather than knowledge libraries. Don't be fooled by overconfident outputs. Beware the hype. Check that outputs are appropriate for your local social and political context.

Test wisely and consider risks.
Do you understand the risk level and the appropriate mitigation measures you should take? Use an AI tool as a support in an area where you have expertise and in-depth knowledge to gain an understanding of its limitations.

Don't let AI make decisions.
Decisions impacting peoples' lives should not be fully automated. Generative AI is incapable of evaluating information to make and explain decisions in the ways humans do. Decisions should always be made by humans.

Don't short-circuit yourself.
Don't let overreliance on AI short-circuit your own thinking and creativity process.

Keep a hand on the wheel.
Always ensure active human oversight and intervention is possible and practical at any stage of AI tool application. Also known as "human in the loop."

Build trust.
Only use AI tools if they will have a positive impact on levels of trust.

Real democracy, not synthetic.
Automation of P2 processes and use of AI tools should never diminish the value of individual contributions or fully replace the human interaction that underpin our democracies. There are no shortcuts to good engagement. P2 professionals should actively resist the use of "silicon samples" or "synthetic participation" in participatory processes.

The Future of AI and P2

As P2 professionals encounter AI tools more, we need to balance their claims of improving efficiency and creativity with mounting evidence about their bias and limitations. P2 practitioners should heed the concerns that AI technology use more broadly will usher in a new era of technocratic decision-making (Zheng et al., 2023) at odds with inclusive P2 efforts. We must continue to ensure P2 process integrity, accuracy, and people-centred values when we anticipate more calls for algorithmic-driven efficiencies in a time of budget austerity.

Our focus here has been on the use and impact of AI tools in public participation processes (AI in P2). There have also been calls for better public participation in the design and deployment of AI tools themselves and in the policies governing their use (P2 in AI) (Corbett et al., 2023; Gilman, 2023; Moon, 2023). There are a growing number of private sector and government staff, researchers, and civil society organizations raising concerns about the adoption of these tools. Here, P2 practitioners also have a role to play in facilitating those dialogues and advocating for responsible, public-informed AI development, deployment, and governance. As P2 professionals, we can work together to share experiences, seize opportunities, and explore challenges as these technologies emerge and change.

We know there are no shortcuts when it comes to good public participation. We hope this paper begins to frame some opportunities for reflection and engagement within our professional community about the future of AI and P2.

About the Authors

Morgan Boyco, MCIP, RPP

Morgan is a PhD Candidate in the School of Planning at the University of Waterloo. Morgan's research focuses on public participation, contemporary planning and democracy theory, and the use of the latest digital civic engagement technologies to support local public participation processes. Morgan is also a professional planner and community engagement practitioner with Dillon Consulting, working with municipalities across Canada to more effectively involve their community members in local decision-making processes.

Dr. Pamela Robinson, MCIP, RPP

Pamela is a professor at the School of Urban and Regional Planning at Toronto Metropolitan University (formerly Ryerson) and is Academic Director at City Building TMU. Throughout her career as a planner her research and practice have focused on complex, emergent challenges that Canadian communities face. Her current research asks: who is planning the Canadian smart city? In 2020 Pamela received the inaugural Canadian Institute of Planners President's Award: Academic. She is an Advisor on the Toronto Public Library's Innovation Council and serves on the Board of Directors for the Metcalf Foundation. Pamela writes for Spacing and recently published a co-edited book The Future of Open Data (2022).

References

Argyle, L. P., Busby, E. C., Fulda, N., Gubler, J. R., Rytting, C., & Wingate, D. (2023). Out of One, Many: Using Language Models to Simulate Human Samples. *Political Analysis: An Annual Publication of the Methodology Section of the American Political Science Association, 31*(3), 337–351.

Argyle, L. P., Busby, E., Gubler, J., Bail, C., Howe, T., Rytting, C., & Wingate, D. (2023). AI Chat Assistants can Improve Conversations about Divisive Topics. *arXiv preprint arXiv:2302.07268.*

Bordoloi, P. (2022, July 14). India's Project Bhashini: Breaking the language barrier with AI. Analytics India Magazine.

Boucher, P. (2020). Artificial intelligence: How does it work, why does it matter, and what can we do about it? Scientific Foresight Unit (STOA), European Parliamentary Research Service.

Boussioux, L., N. Lane, J., Zhang, M., Jacimovic, V., & Lakhani, K. R. (2023). The Crowdless Future? How Generative AI Is Shaping the Future of Human Crowdsourcing. *Harvard Business School Technology & Operations Mgt. Unit Working Paper No. 24-005.*

Bright, J., Enock, F. E., Esnaashari, S., Francis, J., Hashem, Y., & Morgan, D. (2024). Generative AI is already widespread in the public sector. *arXiv preprint arXiv:2401.01291.*

Broussard, M. (2024). More than a Glitch: Confronting Race, Gender, and Ability Bias in Tech. MIT Press.

Brown, S. (2021, April 21). Machine learning, explained. MIT Sloan.

Corbett, E., Denton, E., & Erete, S. (2023). Power and Public Participation in AI. *Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization,* Article 8.

Crockett, M., & Messeri, L. (2023). Should large language models replace human participants? *PsyArXiv preprint.*

Cuau, C. (2019, July 5). Applying artificial intelligence to citizen participation: the Youth4Climate case study. CitizenLab.

Dengel, A., Gehrlein, R., Fernes, D., Görlich, S., Maurer, J., Pham, H. H., Großmann, G., & Eisermann, N. D. G. (2023). Qualitative Research Methods for Large Language Models: Conducting Semi-Structured Interviews with ChatGPT and BARD on Computer Science Education. *Informatics, 10*(4), 78.

Gilman, M. (2023). Democratizing AI: Principles for Meaningful Public Participation. Data & Society.

Grossmann, I., Feinberg, M., Parker, D. C., Christakis, N. A., Tetlock, P. E., & Cunningham, W. A. (2023). AI and the transformation of social science research. *Science*, 380(6650), 1108–1109.

Hadfi, R., Okuhara, S., Haqbeen, J., Sahab, S., Ohnuma, S., & Ito, T. (2023). Conversational agents enhance women's contribution in online debates. *Scientific Reports, 13*(1), 14534.

Helfrich, G. (2024). The harms of terminology: why we should reject so-called "frontier AI." AI and Ethics.

ISO. (2024). What is AI? All you need to know about artificial intelligence. International Organization for Standardization.

Ito, T., Hadfi, R., Haqbeen, J., Suzuki, S., Sakai, A., Kawamura, N., & Yamaguchi, N. (2020). Agent-Based Crowd Discussion Support System and Its Societal Experiments. *Advances in Practical Applications of Agents, Multi-Agent Systems, and Trustworthiness. The PAAMS Collection*, 430–433.

Kordzadeh, N., & Ghasemaghaei, M. (2022). Algorithmic bias: review, synthesis, and future research directions. *European Journal of Information Systems, 31*(3), 388–409.

Lecher, C. (2024, March 29). NYC AI Chatbot Touted by Adams Tells Businesses to Break the Law. The Markup.

Marmolejo-Ramos, F., Workman, T., Walker, C., Lenihan, D., Moulds, S., Correa, J. C., Hanea, A. M., & Sonna, B. (2022). AI-powered narrative building for facilitating public participation and engagement. *Discover Artificial Intelligence, 2*(1), 7.

Martineau, K. (2021, February 9). What is AI alignment? IBM.

Mearian, L. (2024, February 7). What are LLMs, and how are they used in generative AI? Computerworld.

Min, A. C. (2023, July 18). 4,000 civil servants using government Pair chatbot for writing, coding. The Straits Times.

Moon, M. J. (2023). Searching for inclusive artificial intelligence for social good: Participatory governance and policy recommendations for making AI more inclusive and benign for society. *Public Administration Review, 83*(6), 1496–1505.

Norori, N., Hu, Q., Aellen, F. M., Faraci, F. D., & Tzovara, A. (2021). Addressing bias in big data and AI for health care: A call for open science. *Patterns, 2*(10).

Prabhakaran, V., Qadri, R., & Hutchinson, B. (2022). Cultural Incongruencies in Artificial Intelligence. *arXiv preprint arXiv:2211.13069.*

Romberg, J., & Escher, T. (2023). Making Sense of Citizens' Input through Artificial Intelligence: A Review of Methods for Computational Text Analysis to Support the Evaluation of Contributions in Public Participation. *Digital Government: Research and Practice, 5*(1), 1-30.

Rosenberg, S. (2023, June 20). In new AI hype frenzy, tech is applying the label to everything now. Axios.

Saperia, E. (2024, February 26). An Infinite Shortage of Managers. CivicAI.

Siegel, E. (2023, June 2). The AI Hype Cycle Is Distracting Companies. Harvard Business Review.

Taylor, B. L. (2023, November 15). Long hours and low wages: the human labour powering AI's development. The Conversation.

Weidinger, L., Uesato, J., Rauh, M., Griffin, C., Huang, P.-S., Mellor, J., Glaese, A., Cheng, M., Balle, B., Kasirzadeh, A., Biles, C., Brown, S., Kenton, Z., Hawkins, W., Stepleton, T., Birhane, A., Hendricks, L. A., Rimell, L., Isaac, W., … Gabriel, I. (2022). Taxonomy of Risks posed by Language Models. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 214–229.

Zheng, Y., Lin, Y., Zhao, L., Wu, T., Jin, D., & Li, Y. (2023). Spatial planning of urban communities via deep reinforcement learning. *Nature Computational Science, 3*(9), 748–762.