Check Against Delivery

ARMA Conference
July 18, 2023
Chris Parsons, Manager, Technology Policy
Office of the Information and Privacy Commissioner of Ontario

Generative AI and the Future of Information Management

- Thank you for the opportunity to be here

- Information management is key to the operation of public and private sector governance, and generative AI technologies offer some novel uses and concerns to consider

- My remarks, today, will touch on some of the technologies that are in play, today, and then shift to discuss some common governance principles and how they might apply to the adoption of AI technologies. I'll conclude with a few emerging governance and policy considerations

- But first, I need to provide a caveat: all of the information I share, today, is for informational purposes only. Nothing said binds IPC's quasi-judicial tribunal, which may be called upon to independently investigate and decide on an individual complain or appeal based on specific facts of a case.

- How many of you are from outside of Ontario? OK, well for you I should probably explain what the IPC is in brief.

- We are a public sector regulator. Our mandate is to provide a right of access to information held by public sector organizations and, also, to protect individual privacy with respect to personal information held by public organizations.

- IPC is a modern and effective regulator. In part, this means that our staff undertake research to develop policy capacity on emerging technologies, such as AI systems.

- And this brings us to our present AI moment. Anyone who has read a newspaper, been to the movies, or simply not been hiding under a rock knows that we're in a bit of an "AI summer".

- Generative AI technologies are rapidly being injecting into the products that our organizations and colleagues are using; it is becoming an increasingly pressing technology for records and information management professionals

- Generative AI, generally, can be characterized by its ability to generate new content—such as images, text, audio, or even video—by learning patterns in underlying data and then being provided a prompt. Examples include ChatGPT or DALL-E

- We are seeing generative chat systems being integrated into professional environments. Examples include:

    - ChatGPT for employers
    - ChatGPT for providing information that is being used to guide organizational decisions

Check Against Delivery

ARMA Conference
July 18, 2023
Chris Parsons, Manager, Technology Policy
Office of the Information and Privacy Commissioner of Ontario

    - Online or private systems where an organization inputs files/records and then queries them
    - Use of generative systems to create new administrative records, including employment offers, recusals, or summaries of meetings

- At least some of what we're seeing is going to stick around in some form or another. What we see, today, is likely not the same as what we'll have in 5-10 years, but this kind of AI technology is unlikely to go away.

- So, let's now turn to some select principles of public administration that pertain to the public and private sector alike. I'm going to speak to five of them:

    - First, **transparency and accountability**. Per this principle, organizations assume responsibility for their actions and provide suitable justifications to support them. Transparency supports both scrutiny and, consequently, accountability.
    - Second, **efficiency**. It is important for organizations to make the most efficient use of resources to deliver substantive outcomes.
    - Third, **quality of output**. High quality outputs are required to build credibility in services or products.
    - Fourth, **predictability and reliability**. Organizations must behave with consistent and be predictable if they are to generate or maintain trust in their offerings or services.
    - Finally, **privacy and security**. Information must be collected, retained, disclosed, and protected in conformity with the law.

- So, these are some key — and I'm sure somewhat obvious — principles that organizations need to operate by. But what happens when we think about them in contrast to generative AI systems, such as ChatGPT?

- Well, let's first turn to **Transparency and accountability**.

    - LLMs, such as ChatGPT, are functionally black boxes. This means that we don't really understand how they work — we know the inputs and the outputs, but not how the inputs *get to* the outputs.

- What about **efficiency?**

    - It is very, very possible that ChatGPT-like systems could significantly improve the speed at which information management professionals operate. Think of a situation where you are in a M&A or processing a document under FOIA, and you need to know whether you can disclose certain information or whether there are additional legal considerations around its disclosure. A generative AI system some day in the future could assist in speeding up this process

Check Against Delivery

ARMA Conference
July 18, 2023
Chris Parsons, Manager, Technology Policy
Office of the Information and Privacy Commissioner of Ontario

- But, we're not there just yet. There is a challenge in the accuracy, reliability, and repeatability of outputs

- Moreover, while you may realize certain business efficiencies in processing information, your organization might also have environmental efficiencies you need to consider, and systems like ChatGPT are not presently environmentally efficient.

- When we turn to **quality of input,** we, also see benefits and concerns

  - It may be super handy to use a generative AI system to develop 1st drafts or take 1st runs at a given task. In the UK, the government released on June 29, 2023, guidance for public servant use of generative AI and openly acknowledges these kinds of uses.

  - But there are still concerns. Namely:

    - Outputs may sound convincing but actually be pretty superficial
    - The outputs may produce information that is seemingly true, but fake — and thus require high degrees of fact checking
    - There are not qualitative or quantitative ways of assessing the accuracy of these systems' outputs
    - In aggregate, this also has questions about whether the systems genuinely improve the efficiency of using these systems, today, beyond for a subset of tasks

- Similar kinds of potentials and concerns apply when turning to the principle of **predictability or reliability**.

  - We could see a world in the future where institutional memory could be quickly and easily recalled by just querying an organizations chatbot and associated knowledge repository. But we're not there yet, today.
  - Simple *and* complex inputs can lead to varying outputs, with the effect that it can be hard to trust the outputs.
  - Moreover: in the case of FOIA requests it can be harder to know what information was relied upon by decision makers unless all inputs (their queries) and outputs (the responses) are recorded somewhere

- Lastly, we turn to the principle of **privacy and security**

  - There is a risk that when you input information into these models that you may leak confidential information to the developer, who may use the information to train subsequent versions of the model

  - There is, also, the potential for these models like ChatGPT to memorize information even when you're deploying them just behind your own firewalls.

Check Against Delivery

ARMA Conference
July 18, 2023
Chris Parsons, Manager, Technology Policy
Office of the Information and Privacy Commissioner of Ontario

They can develop 'memories' from their training data with the effect that someone could pose a query and be presented with information they are not authorized to access. This could include confidential organizational planning or business records, personal information, or other materials.

- Finally, we don't really understand how to secure these kinds of systems, today. That's exciting from a computer science and engineering perspective, but likely far less good news for practitioners who actually need to manage or secure organizations' information

- So, in light of all this, what are a few emerging governance or policy considerations you could leave with today? I'm going to offer you four.

  - First, **explainability is key**. You need to consider whether you sufficiently understand how the systems work prior to adopting them.
  - Second, **what datasets were involved in the system's creation?** Many LLMs, like ChatGPT, are trained on public, English-dominant, data. Data is often scraped from the internet without consent or, in some cases, potentially without legitimate business purposes. This means that organizations need to question the lawfulness of the technology and, also, how it will deal with non-English texts. This point is particularly important for a bi-lingual country like Canada: can you guarantee that English and French inputs and outputs are being fairly processed?
  - Third, **what is today's model?** If you use a LLM today, and it's updated tomorrow, what does this mean for using it to create records? Can you retroactively re-create similar records as you could in the prior version? What does this do to predictability or reliability or quality? For managing the generation of records more broadly?
  - Fourth, **what's done with input data?** It bears repeating: how are your organizations inputs used to further train the model or create new datasets? This is of particular importance when talking about confidential or personal information.

- Just to conclude, I don't want anyone to think that generative AI is inherently some kind of bad thing. It's very exciting! But it's new and so should be approached with a degree of respect and caution, like all new technologies.

- The IPC is actively interested in this technology and policy area. If you're a member of a public sector organization, or the broader public sector, than we'd welcome an opportunity to speak with you to better understand your approach to AI technologies. This kind of engagement can help us should we develop guidance or other materials in the future about AI technologies or generative AI specifically.