

ChatGPT will be a tricky puzzle piece in public service

Where accountability and legitimacy are paramount, use of the AI chatbot requires an array of safeguards and good sense.

Ong Li Min and Jason Allen

Singapore's Civil Service is poised to embrace ChatGPT on an unprecedented scale.

Earlier this year, it was announced that an artificial intelligence (AI) app that taps the chatbot was being developed to assist in research and the crafting of speeches.

It carries the potential to revolutionise the work of approximately 90,000 civil servants, offering them an AI assistant for those very tasks.

The deployment of AI for public service delivery is not new.

Many are familiar with "Ask Jamie", the Government's natural language processing driven virtual assistant which assists the public with queries on policies such as MediShield Life.

It now also facilitates personalised transactions, such as directing the user to the personal tax portal to file taxes. In this way, the chatbot has enabled interactions between citizens and the state to become intuitive and seamless.

Meanwhile, generative AI (GenAI) has been percolating through public services across the globe.

But public sector use-cases of AI raise unique concerns affecting the citizen-state relationship, particularly accountability and legitimacy. And the solutions lie beyond general usage guidelines.

It is important to be clear about what AI tools can and cannot do well, given the current state of technology.

HALLUCINATIONS AND BEYOND

There is a reason why ChatGPT

hasn't been the most reliable of tools in the realm of legal proceedings, where truth and integrity are key pillars.

Large language models – or LLMs – are a type of AI algorithm that do not actually "comprehend" the text they produce, but simply predict the most likely next word in a sentence based on the training data.

ChatGPT uses this structure to string words together, in response to a prompt or query, based on a historical data set to produce plausible sentences. Each response is unique to each interaction.

The results can sometimes be eye-watering.

Recently, a United States judge imposed sanctions on two New York lawyers after they admitted to citing fictitious precedent cases to a US court, based on research using ChatGPT.

"We made a good faith mistake in failing to believe that a piece of technology could be making up cases out of whole cloth," the lawyers' firm said in a statement.

This case highlighted how the chatbot's responses can at times be seriously flawed. The danger is, they are stated in a convincing, authoritative tone. But they can do a lot of damage if given free rein.

In Australia, for example, the use of AI to automate debt assessment and recovery for welfare recipients has proved to be contentious. The country had to scrap a debt recovery scheme in 2020 and 470,000 wrongly issued debts had to be repaid.

For its part, OpenAI has prohibited the use of its GPT-4 in the context of what it calls, "high risk government decision-making, including law enforcement and criminal justice, migration and asylum".



OpenAI chief executive Sam Altman has said that that GenAI systems will require both global and local democratic input, to align them with people's values, history and culture. PHOTO: AFP

For now, the current state of technological development would preclude more "intelligent" uses like determining the applicable law or public decision-making.

Even when enterprise-grade systems are created, some deeper considerations remain in the context of public administration.

Using AI in public service communications might influence perceptions of the legitimacy of the government programmes too.

ACCOUNTABILITY IN HUMAN-AI WORKFLOWS

All this is not to undermine the potential of AI and how it can help deliver better service and justice, with the right guardrails in place.

Portugal is currently exploring the use of ChatGPT to help citizens on that front.

In the initial phase of rolling out

When AI tools are integrated into the workflow, questions over responsibility and accountability become harder. The number of facts to fact-check is going to rise exponentially, many of which cannot be attributable to a human author.

this conversational interface, it is envisaged that the language model will be able to answer citizens' queries intuitively

regarding the legal processes of marriage and divorce.

Eventually, the AI tool is expected to provide clarity about the services operated by Portugal's justice department as it guides users to the right ones for their needs.

Given the pace of development in these technologies, including techniques to detect machine-generated text, the risks and the challenges will hopefully be reduced, allowing us to harness the potential of the technology.

For example, ChatGPT's ability to summarise text opens up a range of possibilities.

Reports and other research material such as speeches or parliamentary debates can be summarised in a few paragraphs, thus substantially cutting down reading and research time. There is also scope for creative use like

generating multiple research questions and ideas if the user feels "stuck".

The tool can also be used to generate a first draft of a letter of decision, although a leaf might be taken from the European Commission's playbook not to replicate the response when it comes to public documents, given that language models run a real risk of plagiarising from their training sets.

In his recent visit to Singapore, OpenAI CEO Sam Altman said that GenAI systems will require both global and local democratic input, to align them with people's values, history and culture.

Nowhere is this more true than in the context of public administration. The Ministry of Communications and Information has clarified that usage guidelines have been introduced in the public service. It has also made it clear that officers are responsible for vetting AI-generated content and are accountable for their own work.

While this may resolve some issues of clarity, systematic incorporation of such tools into the public service will require deeper and more detailed reflection in due course.

When AI tools are integrated into the workflow, questions over responsibility and accountability become harder. The number of facts to fact-check is going to rise exponentially, many of which cannot be attributable to a human author.

Given that ChatGPT is a general purpose tool, it is the user who determines how they will use the programme. Human intervention to check the generated output is necessary to mitigate potential risks.

The problem is not the use of a GenAI tool per se but how it is used. Is the tool being used to generate ideas, frame problems, or search for facts?

Ideally, disruptive technology should be used to free up time for human intelligence and enhance quality of life.

As historian Louis Hyman suggested, automating tedious tasks can free us to do "more complicated, more rewarding, more human work".

Surely, this is something we want for our public servants.

• Ong Li Min is a research associate at the SMU Centre for AI & Data Governance.

• Jason Grant Allen is an associate professor of Law at SMU Yong Pung How School of Law and director of the SMU Centre for AI & Data Governance.