GOVTECH DECODED

EPISODE 11
REINVENTING WORK WITH AGENTIC AI

Host: Adriana Chan
Guests: Lois JI, Wan Ding Yao, Jessica Foo
Date aired: 10 November 2025

[Ding Yao] I think it is clear that AI safety and security takes a whole of government, maybe even a whole of nation effort.

(Intro music)

[Adriana] Hi everyone, welcome to GovTech Decoded. In this series we'll discuss hot tech topics and how the Singapore government leverages technologies to build tech for public good. I'm GovTechie Adriana and your host for this episode.

Today we're going to be diving into the buzzy world of AI. It's a technology that really is becoming very integral to our day-to-day lives. I believe quite a lot of us use it.

The government is no stranger to its adoption either, so we are going to be talking all about that. I'm excited to be joined by my guests. We've got Lois, Ding Yao and Jessica.

[Lois] Hi everyone, I'm Lois. I'm a data scientist from GovTech. At the same time, I'm forward deployed to SkillsFuture Singapore, SSG, where I lead a team of data scientists and engineers.

Our team focuses on capturing the latest AI technology and applying it to enhance our day-to-day work.

[Ding Yao] Hi, I'm Ding Yao. I'm a cyber security engineer at GovTech. I've been appointed to CSA to work on emerging tech strategy, covering AI and quantum.

So my day-to-day job surrounds identifying strategic trends in emerging tech, as well as working with government partners and industry partners to conceptualise, implement and manage emerging tech strategy and policy.

[Jessica] Hi, I'm Jessica. I'm from GovTech's AI practice. We focus on building deep AI capabilities within GovTech.

We experiment with the latest technologies in order to ensure that we bring about meaningful change across various domains in government. I work specifically on responsible AI, in which we try and ensure that the AI that we use and we put out for public use is safe, secure and trustworthy.

[Adriana] So we've got a little game that we're going to play to warm us up. How it's going to work is that a few true or false questions, as well as a few decode ones. First one is, ooh, true or false? AI systems are infallible. Who would like to take that?

[Ding Yao] So true or false? I would say false. And the reason is that if there are controls that are not in place to ensure that AI systems are safe, secure and trustworthy, AI systems can fail us.

[Adriana] True or false again? AI is out to steal all our jobs. Scary, scary.

[Jessica] Who wants to take this?

[Lois] I'll take this.

[Jessica] Okay.

[Lois] So wearing my SSG hat, my SkillsFuture hat, AI is not out there to steal your jobs. So AI can't possibly replace everything that we do. In fact, it's already here to change the way we work and sometimes the way we live.

So do not be fearful of AI. Instead, learn how to use it and learn how to go beyond just AI.

[Adriana] I think we have our first decode question. So it's a fill in the blank one. And yeah, so I'm just gonna pass this over to you. To me, responsible use of AI means... I think Jessica.

[Jessica] Yeah, okay. I will take that since I'm from the responsible AI team. So in the responsible AI team, we have defined six principles of responsible AI.

So the first is safety. The second is robustness. The third is fairness. The fourth is explainable. The fifth is privacy. And the last is transparency.

So these six principles, as long as we uphold them, when we actually deploy AI or use AI, we believe that that ensures the responsible use of AI.

[Ding Yao] To me, responsible use of AI means using AI in a way that allows society to benefit while minimising the negative consequences. In short, AI for public good. This means building and using AI in a way that allows our citizens to trust the AI systems that we build.

And this means that we need to ensure that the six principles that Jessica just mentioned are actually incorporated within our AI systems.

[Adriana] I love that. We went from principle to application to a human-centred answer. Wonderful.

Okay, so the last decode question that I have is one that I am not familiar of a term and I'll need you to help me with. How is Agentic AI different from Gen AI?

[Jessica] There are a lot of different definitions around agentic AI. But I think for us, we feel that the main difference between Agentic AI and Gen AI is that there is a sense in which agents in an agentic system have decision-making abilities. They also have planning abilities.

And because they have decision-making abilities, they are also equipped with the capabilities and the tools to actually perform tasks autonomously. So I think that is the biggest differentiator between Gen AI and Agentic AI.

[Adriana] So in earlier episodes, we of course had brought up AI. You can't go anywhere without talking about AI, right? So you talked about Launchpad.

We've talked about Pair before. What other applications does AI have within the government and what kind of benefits do you see for the wider Singapore?

[Jessica] Yeah, so I would broadly categorise the type of AI applications we have in government into two categories. The first is internal facing applications. So these applications focus mostly on allowing public officers to do their work more efficiently.

So with these applications, they can turn out their outputs faster and it allows them to be able to do their work faster. So one example of that is AI bots. So AI bots is actually a no-code platform that allows any public officer to create a chatbot out of their necessary documents.

So it means like any public officer would be able to actually upload their documents. I've seen a lot of public officers do great things with them. One use case is actually the IM8 chatbot.

So IM8 is actually a series of cybersecurity regulations we have around how we develop applications in government. And for a lot of developers, it's quite difficult to navigate. And so having an IM8 chatbot is great because we just ask it questions and then we are able to get the answers and it allows us to build our applications faster.

So another example of an internal facing application is Smart Compose, which basically allows public officers to write emails with AI. At the end of the day, the public officers still have to evaluate the content that's produced by the AI and then send it out. But it really reduces the time it takes for them to actually generate the initial draft.

The other category of AI applications that I would say are mostly used and deployed in government today are those that are outward and public facing. So one of the projects that I work on is Project Pensive. Project Pensive is essentially an AI tool that enables us to actually detect dementia from a series of drawings.

So the elderly just have to do like a series of drawings and from these drawings, we are able to produce a detection or rather a classification of whether the elderly has dementia or not. In the past, this sort of detection and diagnosis process used to take like one or two hours at the polyclinic. And we've actually shortened that time to just five minutes on a digital tool.

So that is great because it allows us to actually detect dementia faster for these people and that allows us to intervene faster and it reduces the probability of deterioration of dementia. But that project currently is still in beta phase. We're still like (progressive), thinking of how we can progressively roll it out across the Singapore population.

[Adriana] That's really cool. Really cool applications. Do you have something else that you want to add?

[Ding Yao] Yes. So we've seen this through the use of AI in cybersecurity and we are making big bets. Some of the key use cases include red teaming, threat modelling and threat hunting.

And much of the work goes into codifying the expertise of our cyber officers so that AI agents can actually automate some of the more routine tasks. And we find that this frees up the capacity of our limited manpower resources, very expensive, so that they can focus on doing higher value work. One example is in vulnerability research, which typically requires more human ingenuity and creativity.

[Adriana] Wow, that's super cool. You're just using AI as sort of like your backup. In my head, it sounds like an army that you have at your disposal.

[Ding Yao] But we see more as a human-AI collaboration and we don't think that AI will take over our jobs.

[Adriana] Of course, we've already decoded that.

[Jessica] Yeah, and I think on top of partnering with other commercial providers or AI leaders, international organisations, we also at AI practice, work a lot in sort of coming up with different playbooks to enable us to all have a sort of common understanding and grounding in all these agentic AI terminology. And so that enables us to level up everyone's capabilities at the same time and to scale up that sort of capability building across government.

So we have published the agentic AI playbook, which defines what are agentic AI architectures, how do we go about building a good agentic AI system. At the same time, we have also developed an open source, the agentic risk capability framework, in which we have defined different sort of risks as well as technical controls that exist in agentic AI systems. The purpose of that framework is to actually enable developers as well as organisations to actually identify risks and to mitigate them so that they can actually use agentic AI for greater good.

[Adriana] Beautiful. This agentic playbook is something that I can read?

[Jessica] Yeah.

[Adriana] Yeah? Okay, fantastic.

[Adriana] I think you found your target audience right here. Okay. So AI, lots of big words, a lot of buzzwords, lots of things moving super, super fast.

Could you introduce some of the new tech terms or technologies that we should know? And how are we staying at the forefront with everything moving so fast?

[Jessica] I think we've touched on it briefly, but definitely I think agentic AI is inevitably one of the biggest trends of the year. And I think for us, how we feel that we should actually be at the forefront of change is to first develop a common understanding across government about what we even mean about new AI technologies. And that's why the AI practice actually put out the agentic AI primer.

It's actually freely available for any developer across the government to look at and reference. And in the primer, we actually set out certain definitions of what we mean by what is agentic AI. So this could, for us, for example, we actually break it down into a component level.

We see agentic AI comprises of the LLM itself, which is kind of like the brains behind the system. But at the same time, the reason why agentic AI can actually autonomously execute tasks is because they're equipped with tools. So tools enable capabilities.

So these are things like being able to send an email, being able to book a restaurant or send a calendar invite. And lastly, another component of agentic AI system is actually memory. So the ability to actually remember information from past interactions and to store it in memory and to then retrieve the information for future plans as well as future outputs.

So just even having this dictionary, this glossary of common terms for us to understand what is agentic AI really sets the foundation for us as a whole of government to actually move forward and actually develop real use cases around agentic AI.

[Lois] Well, speaking of real use cases, in SkillsFuture, we actually explored the application of agentic systems to the space of customer relationship management, where we developed a team of AI agents to inform policymakers on potential improvements they can make through the lens of public and user feedbacks. So for this AI agents we have developed, each of them is set up in a way to model after the function within the customer relationship management team, effectively creating what we call a digital twin of the team. So in this case, we are also not neglecting the fact that we need to do so responsibly.

So we do have officers, actual officers in the loop to fact check some of the information that's presented by these agents in the system. So as this project has real implications to the policies, we actually want to have more robust testing done before releasing into the production system. So currently the project is still in its beta testing stage, and we have to really thank our business users where there is a very strong partnership with the AI team as well as the business team to bring about all this robustness checks.

So just imagine tomorrow when a director asks them, hey, can you share with me what's the emerging topics in the CRM space and what are some of the divisions that needs to be brought on board to resolve a particular policy review? The business users could simply query this agentic system and get a quick answer before going back to their bosses.

[Adriana] That's super cool. It's like you need to have memory, you need a brain, you have the hands, which are the tools, but you also then have a job scope, which is a specialisation, and you are partnering with actual people to deliver those answers. That's right.

[Jessica] Fantastic. So another way in which we feel it's important for us to actually keep at the forefront of changes is to actually publish in the academic community because this allows us to get feedback on whether our methodologies are sound. They really like to hear actually on the ground real practical use cases, but they were also able to give us really good feedback on how we can improve our methodologies and actually level up our capabilities.

So we feel like actually publishing widely with the international academic community is a great way for us to also build and improve ourselves.

[Ding Yao]
We can just add that the government is not doing this alone. So it was recently announced that the MDDI family, our ministry, will be partnering with Google to build agentic AI on top of Google's infrastructure. And this will allow us to not reinvent the wheel, but really apply agentic AI to realise the benefits that it can bring to our citizens and society.

[Adriana]
So from privacy concerns to ethical concerns, I think there has been a lot of controversy around AI. What are the main risks that you can see with AI? How do we mitigate them? And sort of like, what is the role that government has in this sort of risk management?

[Ding Yao]
So we have discussed quite extensively about the benefits that AI can bring to society and our citizens. But in our push for AI adoption, and it is a very strong push, we need to also ensure that the safety and security of our citizens are not compromised. So we see that there are two main groups of risks when it comes to AI. AI safety and AI security. So AI safety is about developing and deploying and using AI in a way that reduces the negative consequences. And this entails broader considerations, such as ensuring alignment of AI systems to the public good. Whereas AI security is about ensuring the confidentiality, integrity and availability, or CIA, of our AI systems. So the CIA of our AI systems can be threatened by attacks that seek to exfiltrate sensitive data or manipulate AI systems to perform in unintended ways and resulting in disruption to its operations.

So for AI security, we have been receiving actually quite a number of reports that suggest that nation state actors are trying to perform very sophisticated attacks on AI companies to steal their model weights and plant malicious backdoors that they can exploit later. So it's quite a matter that we take seriously. And for AI security, our immediate priority is really to elevate the baseline of our systems to ensure that they are protected against these threats. So one example of this effort is on how GovTech has worked very closely with CSA to develop security guidance for agentic AI systems for both government and industry. So this covers the processes of identifying and assessing risks across the agentic AI workflows as well as capabilities. And it also proposes practical controls that developers can consider to ensure that their systems are protected against AI security threats.

[Jessica] And to delve into AI safety, I think for us, we also feel that the first line of defence is definitely policy officers themselves. And so we have published a responsible AI playbook in which we outline certain best practices when it comes to evaluating the safety of the AI applications. So this could include dimensions like whether there is unsafe content that is generated, but it could also include dimensions such as whether the content that is produced is fair or biased. Our team also builds capabilities in specific sort of models as well as guardrails. And one of them, for example, is LionGuard. And LionGuard actually detects Singlish toxic or unsafe content online.

We first actually released it last year. And this year, we have actually released a V2 in which this is actually multilingual. We're able to detect unsafe content in Tamil, Malay, as well as Mandarin.

And so this really allows us to actually share this capability with the rest of government and they can actually incorporate these guardrails into their AI applications. And in order to do that more efficiently, we actually have central products in government. And that is, for example, through products such as AI Guardian.

One of them is Sentinel. So Sentinel actually provides AI applications guardrails through a single endpoint. So instead of having to integrate different sorts of guardrails themselves, they only have to call one API call and select the sort of risk that they care about, whether it be unsafe content, it can be like off topic, or it can even be like privacy concerns. So one of the use cases we've worked with is actually safety testing for MOE's SLS (Student Learning Space) chatbot, which is actually a chatbot which teachers can upload documents and students can actually query and learn from the chatbot. So what we did was we actually created a data set in which we actually tested whether the chatbot was safe for minors. So you can imagine like this is actually like a data set that maybe other commercial providers may not be motivated to create.

But for us, because minor safety is so important, we created that data set and we performed that safety testing. And that actually assures MOE when they deploy the application that there's a minimum or rather baselines of safety and security that they think about when they deploy their application. And lastly, with agentic (AI) coming up, our team also actually worked with other partners, even across governments such as CSA, IMDA, and with the product teams as well.

And we actually developed the Agentic Risk Capability Framework, in which we outline different sort of agentic risks, as well as associated controls. And the idea is that developers and policy officers have sort of a place where they can immediately turn to in order to assess the risk, the safety, as well as the security risk of their agentic AI applications.

[Lois] I think over at SkillsFuture, we make sure that all the agentic exploration work are subjected to multiple layers of evaluation. So we are currently working towards a scalable way to do multi-layers, in total, three different layers of evaluation for our agentic systems.

The first layer is that individual agent layer. The second is the agent's interaction with tools. And the third layer being the agent-to-agent interaction, which within the community we refer to as orchestration layer. So there is also close scrutiny to the level of data and environment access that each of these agents are given. So this is because we do not want to grant additional privileges to any of these agents during the operation itself. And this also effectively contributes to the subsequent evaluation of the entire agentic system for reasons such as if we are able to detect responses from a particular agent that's clearly beyond its intended scope, then we know that this could be a sign of hallucination or a sign of poorly configured access rights. So in this case, as one of the earliest team to explore the agentic space within the public sector, how we actually started this effort is really to look at our previous use cases within the agency and to ask ourselves, are there certain limitations that we could overcome now that we have, you know, something like AI agents in place?

So ever since we developed the prototype for our beta tests, we have also been actively sharing our prototypes as well as our learning and takeaway during this journey at platforms such as Lorong AI, Stack Meetup. So we were really heartened by the community that grew organically within Singapore to really look at different agentic applications within both the public and the private sector.

[Adriana] It really sounds like agentic AI is the way forward, even from a security angle. Yeah.

[Ding Yao] If I can add, I think it is clear that AI safety and security takes a whole of government, maybe even a whole of nation effort.

[Adriana] This is the last question and I always like to end on a prospective note, right? So what is next for AI in your teams, in your domains?

[Jessica] So I think AI definitely will be extremely transformative in different domains as well as in the public service. And I think for us, we want to ensure that we remain at the forefront of innovation, that we continuously try to experiment, we evaluate, and we actually forecast trends ahead of time. And I feel like that will enable us to be able to actually use AI for good and to better advance the public interest.

[Lois] Well, I think this is not limited to just the forward deployed team within SSG, but on behalf of all the forward deployed teams in the various agencies, I'm confident to say that we have a very strong mandate as part of the GovTech AI practice group to help support and champion the public sector's efforts in moving towards cutting-edge technology such as the Agentic AI. And we will work very closely with our agency partners to realise the benefits of Agentic AI.

[Ding Yao] I think likewise, we will continue to stay at the forefront, monitor the threat landscape. And you know, AI is moving so fast. We're already seeing developments of embodied AI, where AI is integrated with robots to achieve effects on the physical environment. We haven't seen much of that yet, but it is something that we are looking out for. And you know, when the time comes and if adoption scales, we will definitely need to come up with guidelines as far as how to implement this and deploy this safely and securely.

[Adriana] Wonderful. The future really looks transformative, as you said.

We talked so much about the speed at which AI has moved through. We've come to the end of our episode as we move speedily to that conclusion. OK, so if you are keen to find out anything more about what we've discussed, head on over to go.gov.sg/govtechdecoded. If you really enjoyed this episode, please support us. Share this out with your friends and your family. Speaking of friends, these are new friends, so go ahead, add them on their LinkedIn pages, my esteemed colleagues. If you want to connect with GovTech, please head on over to go.gov.sg/connectwithgovtech to get all our social media platforms. Thank you again to the wonderful guests for your time, your energy, your expertise here. My name is Adriana, and I'll catch you at the next GovTech Decoded. Bye!

(Outro music)