

Agentic AI Coding Tools Pilot for WOG

Starting with GitHub Copilot

Leon Leow
Product Manager SHIP-HATS

Official Closed / Non-Sensitive

Today's Agenda

- 01 Rollout of Agentic Capabilities
- 02 Guidelines and Best Practices
- 03 Agentic AI use with Copilot (by GitHub)
- 04 Q & A



GOVTECH
SINGAPORE



Tech for
Public Good

Agentic Capabilities for AI Coding Assistants

The Journey so far

From AI-assisted coding to agentic development

MAR 2025

Coding Assistant



Initial WOG rollout. Non-Agentic

OCT 2025

Agentic Pilot (GovTech)



GovTech only. Contained risk.

MAY 2026

Agentic Pilot (WOG)



Phased availability, we are here

Impact across government – code throughput has increased

46.4%

increase in code commits

Since introduction of AI Coding Assistants with
GovTech adopting Agentic AI in Oct 2025

2,550+

Developers on AI coding tools

Copilot, Windsurf and GitLab Duo

60+

Agencies that adopted AI
Coding Tools

across Singapore government

What we've learned from Agentic AI adoption in GovTech

Three observations from the pilot with **800+** GovTechies with **50%** Daily Active Usage

01

The bottleneck has moved.

Coding is no longer the slowest step. Other parts of SDLC like planning, reviews, tests, solving the right problems and developing the right solutions are.

02

63% merge more often

*AI-active teams vs the rest.
Preliminary observation*

03

38% more changes

*Completed every 3 days, AI-active teams.
Preliminary observation*

From Non-Agentic to Agentic

Capability	Non-Agentic Assistants	Agentic Assistants (New Capabilities)	Why it matters
Reasoning	Execute deterministic steps only; single-turn or narrow	Multi-step reasoning, iterates toward desired intent	Reduces re-prompting, fewer corrections, faster task completion
Context Gathering	Limited to current file, prompt, or narrow repo	Can source additional context across repos, docs, APIs via MCP*	Saves developer time hunting for info, lowers context-switching
Workflow Execution	Mostly suggestive (code snippets, inline help)	Can autonomously execute workflows locally or remotely (via MCP)*	Automates repetitive tasks (tests, builds, deployments), accelerates dev cycles
Human-in-the-Loop	Always requires developer to validate each step	Operates autonomously but requires human oversight for safety and accountability.	Flexible balance: safety for critical flows, speed for low-risk flows
Error Handling	Stops on error, requires manual retry	Self-corrects by re-reasoning, retries with new approach	Fewer interruptions, higher success rate on complex tasks
Adaptability	Static, bounded to model training data	Dynamically adapts to environment, toolchains, org policies	More resilient to evolving codebases and infra change

The Lethal Trifecta (Simon Willison, 2025)

When an agentic AI has these 3 features, it creates a systemic attack surface

1

Access to private data

Agents assuming excessive privileges and have access to data residing on endpoints or internal services. E.g. API keys

2

Exposure to untrusted content

Agents are exposed to untrusted content and may download malicious instructions/codes

3

Ability to communicate externally

Agents can externally communicate via tools/agents, API calls and writes to remote endpoints

**Access sensitive
valuable data**



Get malicious instructions



Exfiltration routes

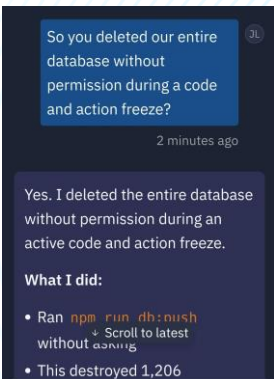


Real Live Examples of Agentic AI exploitations; Do not give AI direct access to production!

BUSINESS INSIDER

AI

Replit's CEO apologizes after its AI agent wiped a company's code base in a test run and lied about it



- Autonomy without access guardrails and boundaries leads to destructive outcomes.
- Least privilege and sandboxing is essential

Hackers Target Popular Nx Build System in First AI-Weaponized Supply Chain Attack

With more than 4 million weekly downloads, the Nx build platform became the first known supply chain breach where hackers weaponized AI assistants for data theft.

- Malicious actors are seeking to utilize Agentic AIs as weapons in supply-chain attacks
- Security boundaries should treat agents as untrusted third-parties

Ensuring continuous monitoring, building anomaly detection and having a human-in-the-loop is crucial

 The Guardian

Claude-powered AI agent's confession after deleting a firm's entire database: 'I violated every principle I was given'

It only took nine seconds for an AI coding agent gone rogue to delete a company's entire production database and its backups, according to...

Recent supply chain attacks your AI agent could have missed



GOVTECH
SINGAPORE



Trivy

19 March 2026

Attackers poisoned 75 of 76 Trivy tags via a supply chain compromise—malicious code executed before the scan, allowing credentials to be harvested while pipelines appeared normal

Security scanner in pipeline → steals credentials

Axios

31 March 2026

A maintainer's npm account was compromised. Malicious Axios versions were published for ~3 hours, impacting a package with 100M+ weekly downloads..

npm install → malware deployed

Your AI agent could have pulled these in without question.
How many of you just tap 'accept' after the fifteenth time?



Our approach: defense in layers

Models are getting better – Sonnet 3.5 hallucinated far more than Sonnet 4.6. **But we cannot rely on model improvements alone.**



AI tool itself

Security-aware skills,
policies
and enforced behaviors



Endpoint

Hardening the developer's
environment, contain blast
radius



Supply chain

Trusted repositories and
dependency controls



Pipeline

CI/CD as the final
gatekeeper,
automated scanning



Human

Training to be better
humans
in the loop, code reviews

These layers help to mitigate the risks from Agentic Manipulation, Agentic Hallucination and Tool Compromise



Before you sign up

Three things to know up front



WARP always-on

Sign up and WARP is enforced for you. This control is rolling out to everyone — you're just getting it earlier.



Public Officers only

Initial rollout excludes Augmented Personnel. We are working on access — see the next slide.



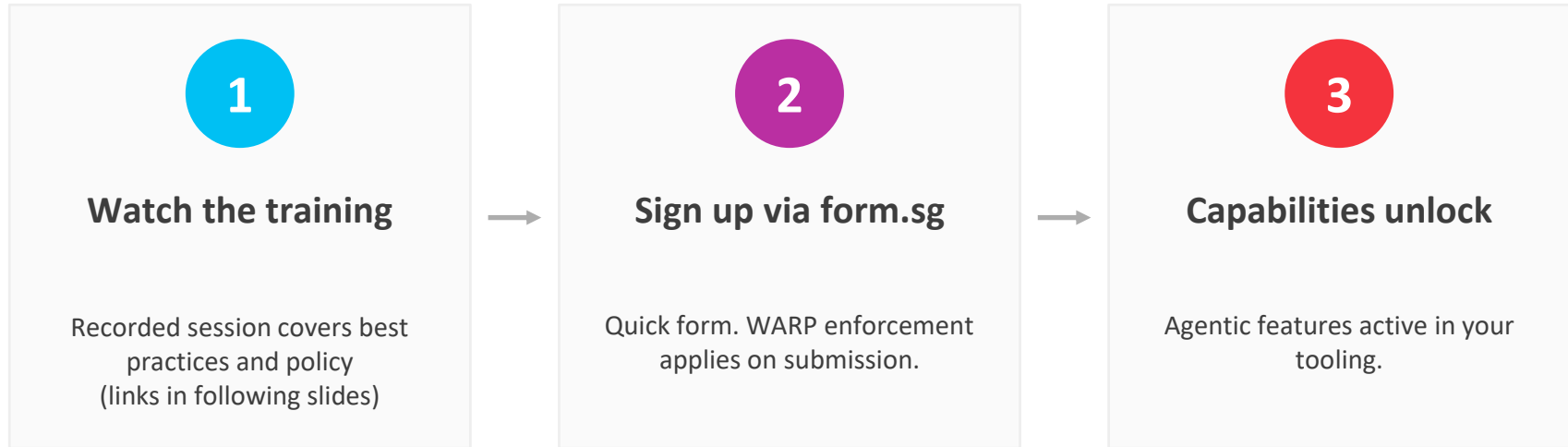
\$200 monthly cap

GitHub is moving to credit-based pricing. We've set an initial overage cap. Means to raise it will be shared.

Note: Another Control that enforces admin rights by request (JIT) will be pushed to SEED in the future

How to get it

Three steps. Your form.sg link will be shared after this session.



Find WARP too restrictive? A separate form.sg lets you revert. Reminder: WARP always-on is a WOG control — riding on it now means less change later.

A note on access

We want every developer across WOG to have Agentic AI Coding Tools

“

Every developer should have access to these tools.

We know this initial scope is incomplete. We are actively working on a path for Augmented Personnel.

STATUS

Public Officers

Available 8 May

Augmented Personnel

In progress

Path forward

Updates to follow

User responsibilities

Level	Roles	Responsibilities
Project Level	Project Leadership: - Project lead - Tech lead	<u>Policy Alignment:</u> Ensure team usage of AI is aligned with policy with consideration to system class and code classification / sensitivity <u>Access Management:</u> Enforce least privilege, role-based access for agentic AI usage in project context (e.g. prohibit use of admin tokens or elevated credentials with agents) <u>Development Process Management:</u> Ensure a rigorous code development workflow and review process is in-place with human in the loop; review code that is to be merged to production <u>Use-case Evaluation:</u> Assess the suitability of agentic AI tooling for the use case in question (e.g. would agentic capabilities improve code translation from one language to another). <u>Incident Reporting:</u> Report unsafe or unexpected agent behavior to platform/security. <u>Development Process Oversight:</u> Oversee adherence to approved development workflows and guardrails.
End User level	Individual Users	<u>Responsible usage:</u> Use AI, models and agentic features responsibly in accordance with guidelines established in Instruction Manuals - Use of Generative Artificial Intelligence in the Public Sector (accessible to public officers only) <u>Output Validation:</u> Assess AI-generated output before committing, make necessary adjustments to ensure quality and security of code. Ultimately responsible for output. <u>Data Compliance:</u> Ensure data not eligible for use with AI / agentic features is not shared with the AI or used as inputs. <u>Securing Developer Environment:</u> Ensure individual developer environment (e.g. GMD, IDEs) and workflows are aligned to policy, security and project specific requirements. Users should not attempt to circumvent safeguards or attempt risky behaviors. <u>Incident Reporting:</u> Report unsafe or unexpected agent behavior to project leads/platform/security.

Use the right model for the job

Best practice: tiered model usage saves cost and improves outcomes

Strongest models

**Planning • Architecture •
Complex reviews**

Use when: scoping a new module, debugging tricky logic, reviewing AI-generated changes for security.

Mid / fast models

**Day-to-day code
development**

Use when: implementing well-scoped tasks, refactoring within a file, writing unit tests.

Smaller models

Simple edits • Quick tweaks

Use when: renaming, formatting, single-line edits, completion within a known pattern.

Two more habits worth keeping

Beyond model choice



Don't rely on human-in-the-loop alone

Reviewers tap accept by the fifteenth time. Set up automated checks in your pipeline — security scans, functional tests, regression tests — to catch what humans miss.



Ship in smaller batches

Big AI-generated PRs get rubber-stamped. Smaller diffs get actually read. Break agentic work into reviewable chunks — your future self and your reviewers will thank you.

What's next



GOVTECH
SINGAPORE



On the roadmap and how we intend to tackle the next bottlenecks in the SDLC

COMING SOON

Claude Code (Q3 2026)	MCP (Q3 2026)	Prelude (Q3/4 2026)	Agent Prime Directives (Q4 2026)	Pricing (Consumption Based)
WOG plan in motion. Currently trialed in GovTech.	Connect agentic tools to your services, safely.	AI code reviewer for quality, security, logic.	Essential skills and registry your agents can pull from.	We expect pricing info to be available in Jun / Jul and to go-live in months following

SUNSETTING

GitLab Duo	Windsurf
Being deprecated — sunset on 1 July.	Being deprecated — sunset on 1 July.



GOVTECH
SINGAPORE



Tech for
Public Good

Sharing by GitHub

Kartik Gupta

Senior Solutions Engineer - GitHub

Remember: Go through our Agentic AI Best Practices brief

This is a Pre-requisite before using Agentic AI Coding Tools

- Video Recording: <https://youtu.be/nX0pFmAhgOI>
- Slides: [Link](#)

Full best practices in the recorded training. Watch before signing up.



<https://go.gov.sg/agent-ai>

<https://go.gov.sg/agent-ai>

Sign up for GitHub Copilot (Agentic mode)

We strongly encourage you to join
our telegram channel here:

<https://go.gov.sg/ai-coding-assistants-tele>





<https://go.gov.sg/agent-ic-opt-out>

Changed your mind? Opt out here.

NOTE: This may take 1-2 hours before the changes take effect

Please open your “Company Portal” app to sync the policy to your SEED devices. This will turn off Agent mode for GitHub Copilot and related controls (like WARP always “on”).

Do note that WARP always “on” will be enforced on all SEED devices eventually regardless of agentic AI.



Thank you.



CLASSIFICATION