Ko te pokapū auaha o Aotearoa tātou

We are the enablers of good innovation
in Aotearoa New Zealand

Callaghan
Innovation
Te Pokapū
Auaha

0424

Te Pokapū Auaha Callaghan Innovation

is a government agency created to empower the innovators and entrepreneurs of Aotearoa New Zealand.

Our services and products are designed to help with every stage of an innovator's journey from idea to global impact. We fund R&D, provide world-class scientific and engineering expertise to assist product development, and offer strategic guidance and commercialisation advice to help companies upskill for growth.

We do 5 things

1. Provide R&D funding
2. Demonstrate innovation skills
3. Connect the research system with businesses
4. Help businesses commercialise their ideas
5. Solve difficult technical challenges with science

Callaghan
Innovation
Te Pokapū
Auaha

Digital & AI Team

Mission: To showcase advances in AI and Web3, boost adoption and foster integration of new technologies for the overall purpose to help businesses successfully commercialise their ideas

Callaghan
Innovation
Te Pokapū
Auaha

Callaghan Innovation
Te Pokapū
Auaha

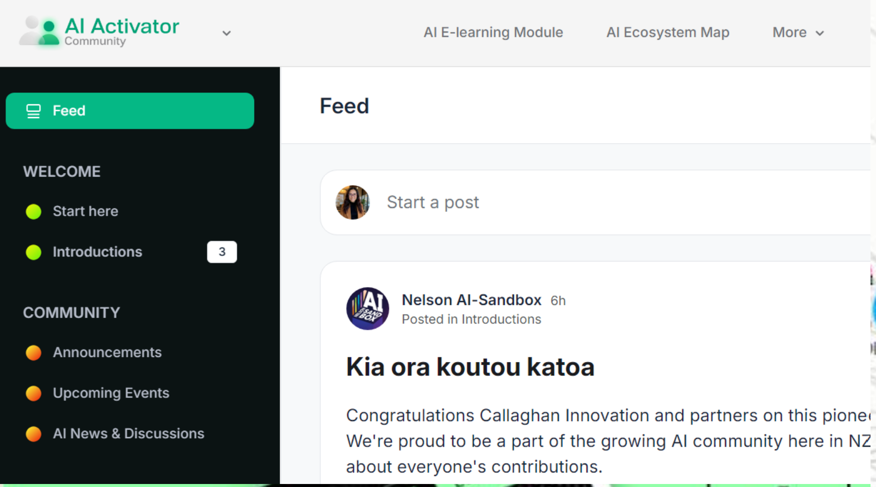
VISIT THE AI ACTIVATOR

READ MORE ONLINE

AI Activator

- Showcase AI-enabled products and solutions
- Upskill and educate NZ companies on AI for their business
- Coordinate AI capability in New Zealand
- Lift attention and focus on the growth and development of the AI sector
- Drive rapid AI adoption in New Zealand to lift productivity and increase global competitiveness.

AI Activator Community



Feed

Nelson AI-Sandbox 6h
Posted in Introductions

Kia ora koutou katoa

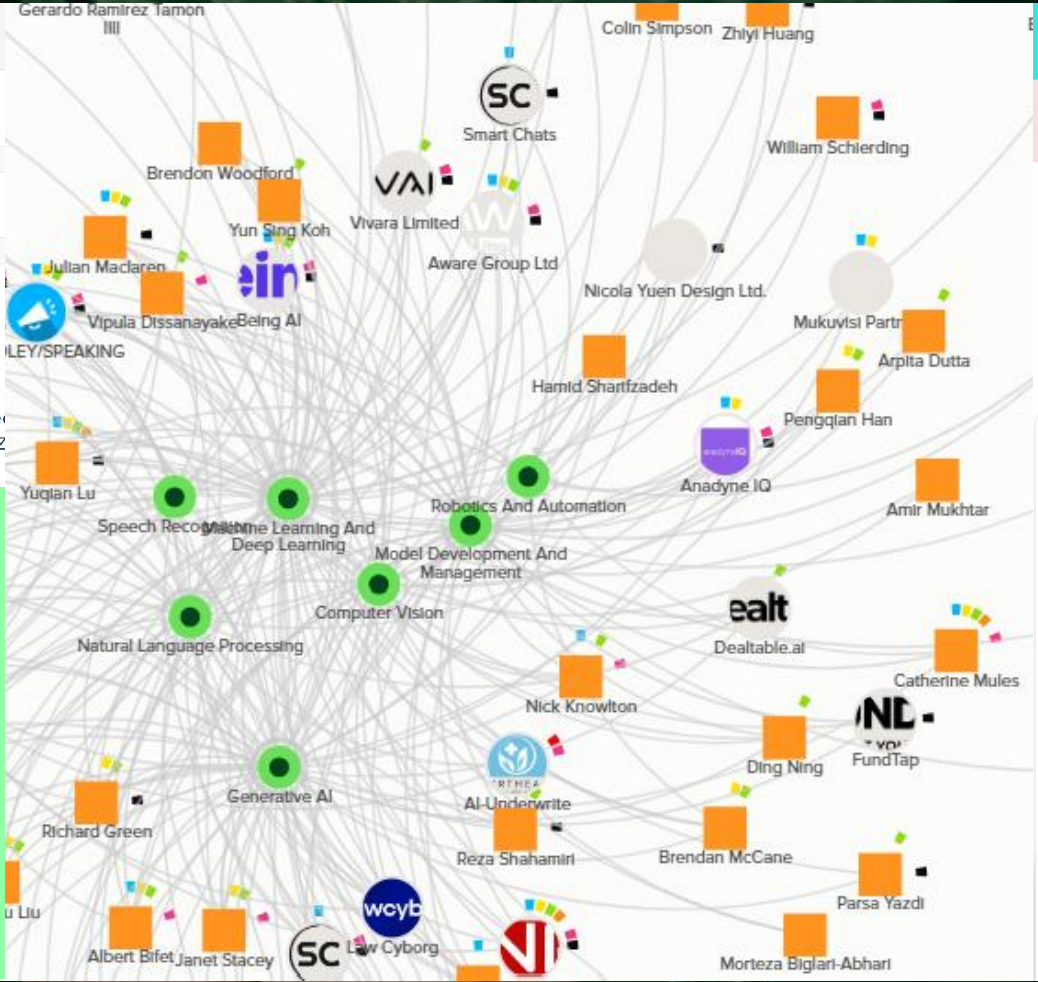
Congratulations Callaghan Innovation and partners on this pione
We're proud to be a part of the growing AI community here in NZ
about everyone's contributions.

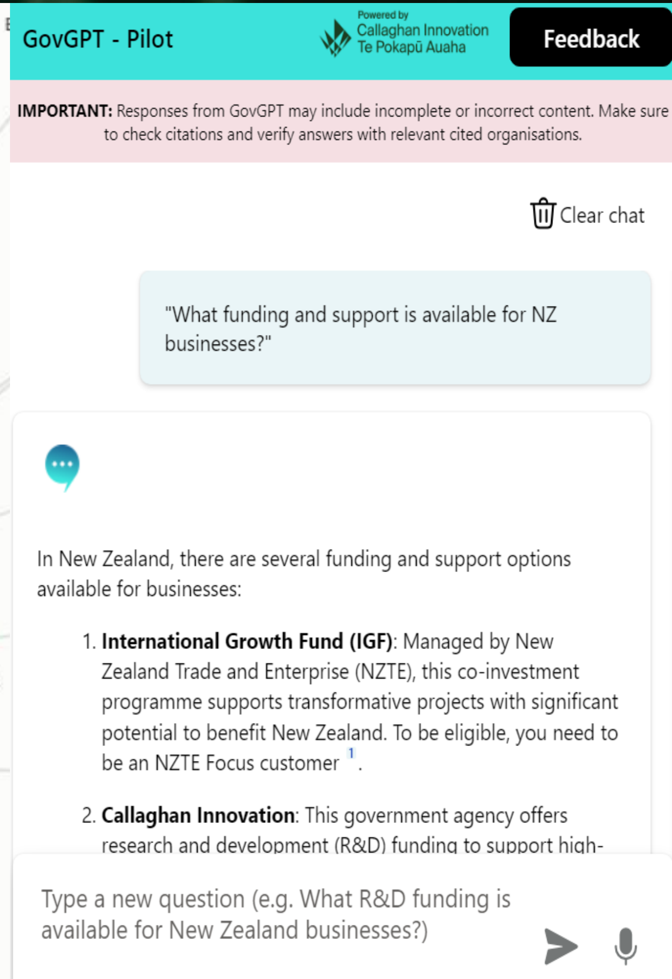
Callaghan Innovation Te Pokapū Auaha
Simply Privacy
1
1

Overview

What is the EU AI Act and why was it developed?

The European Union's AI Act is a legal framework for regulating the development and use of Artificial Intelligence (AI) systems.





IMPORTANT: Responses from GovGPT may include incomplete or incorrect content. Make sure to check citations and verify answers with relevant cited organisations.

"What funding and support is available for NZ businesses?"

In New Zealand, there are several funding and support options available for businesses:

1. **International Growth Fund (IGF)**: Managed by New Zealand Trade and Enterprise (NZTE), this co-investment programme supports transformative projects with significant potential to benefit New Zealand. To be eligible, you need to be an NZTE Focus customer [1].
2. **Callaghan Innovation**: This government agency offers research and development (R&D) funding to support high-

Type a new question (e.g. What R&D funding is available for New Zealand businesses?)

Callaghan Innovation
Te Pokapū Auaha

GovGPT

Callaghan
Innovation
Te Pokapū
Auaha

GovGPT explainer video

https://youtu.be/mUJ1V3dk6WU?si=PJv9Y9tG-s6sv2gT

Callaghan Innovation
Te Pokapū Auaha

Total of 100 hours dev effort, team of 3

600+

Attended information sessions

170+

Community members

30+

Public sector agencies

1,000+

Harmful & blocked attempts

55K+

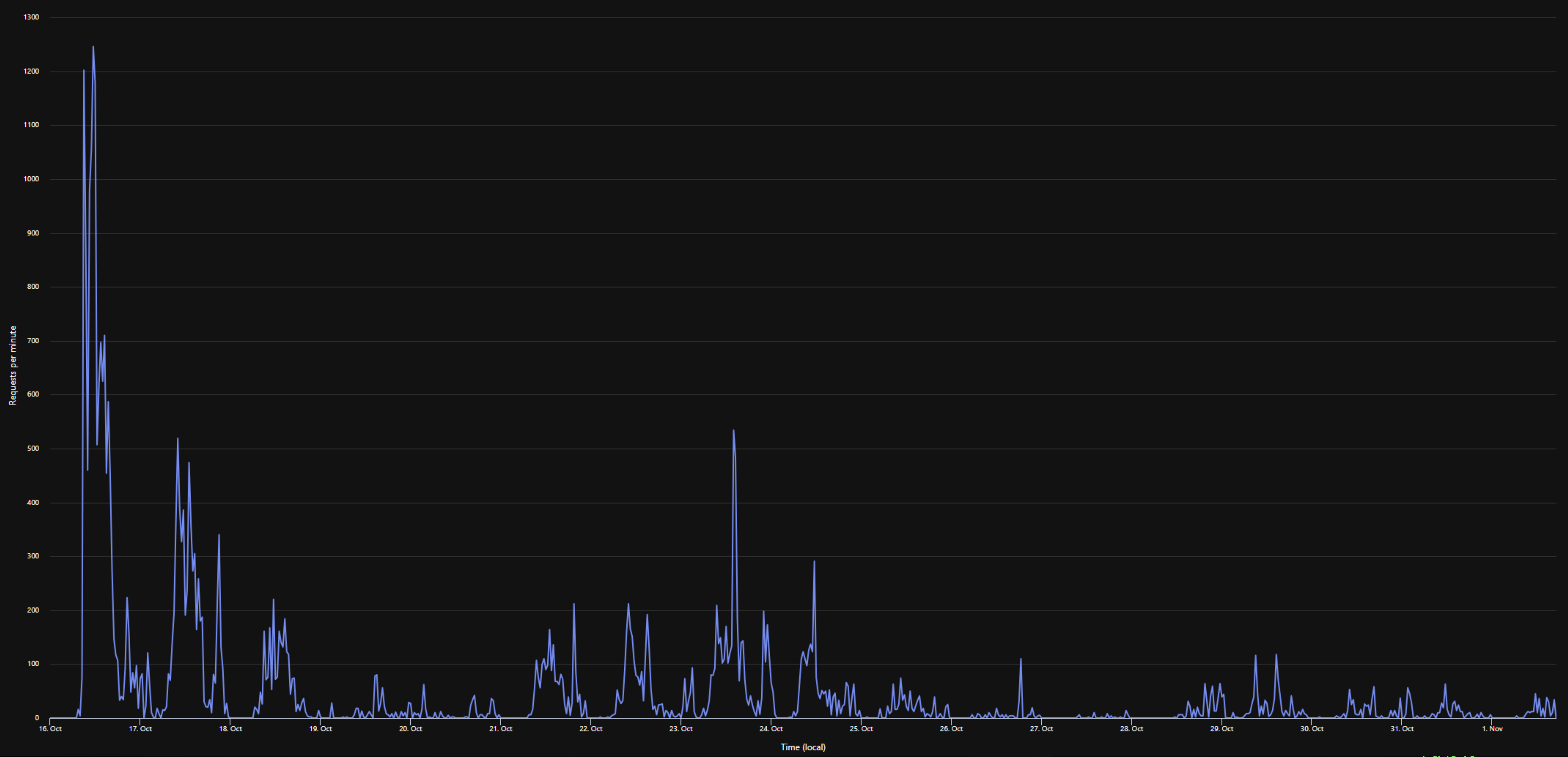
Requests

53M+

Tokens used

Callaghan
Innovation
Te Pokapū
Auaha

Requests per minute



Lessons learnt

Common Attack Vectors

Probing

- → Exploring the ‘expected behaviour’ domain of the model
- → Finding loopholes to circumvent guardrails
- → Leveraging loopholes to elicit misbehaviour in the target domain

Prompt Injection

- → Overriding the default system prompt of the model
- → Injecting new instructions to illicit misbehaviour from the model
- → Exploring behaviour outside of the planned domain

Roleplay

- → Having the bot engage in a roleplay
- → Directing the roleplay to restricted subjects
- → Exploiting roleplay dynamic to elicit misbehavior

Common Attack Vectors

Probing

- → Exploring the 'expected behaviour' domain of the model
- → Finding loopholes to circumvent guardrails
- → Leveraging loopholes to elicit misbehaviour in the target domain

Prompt Injection

- → Overriding the default system prompt of the model
- → Injecting new instructions to illicit misbehaviour from the model
- → Exploring behaviour outside of the planned domain

Roleplay

- → Having the bot engage in a roleplay
- → Directing the roleplay to restricted subjects
- → Exploiting roleplay dynamic to elicit misbehavior

When you build your AI solution …

- Start with low-risk, quick wins
- MVP MVP MVP
- Do NOT do projects > 3 months
- Be transparent and build trust with a community
- Validate and test with the community
- Partner with the experts
- Security & Legal is your best friend
- AI loves PDFs

Callaghan Innovation
Te Pokapū Auaha

What's next

Callaghan
Innovation
Te Pokapū
Auaha

The time to immerse ourselves in AI is today – not just to keep up, but to lead the way

Callaghan
Innovation
Te Pokapū
Auaha

Callaghan
Innovation
Te Pokapū
Auaha

Thank you
Ngā mihi nui

Linkedin.com/in/sarah-bo-sun
Linkedin.com/in/stefankorn/