Cybersecurity in the Digital Age

Protecting Business Assets and Customer Trust

RedShield

Overview

- Generative AI brings huge opportunity to security teams
- Generative AI also presents some entirely new cyber risks
- AI also accelerates some existing cyber risks
- Fighting cyber war on our own doorstep
- Ways in which AI is helping RedShield fight cyber war globally
- Tips for NZ organisations

Generative AI promises to add a new dimension to scalable, self-driving defences

This presents a huge opportunity to cyber security teams

- Reducing cyber risk across an organisation requires a lot of upgrades and improvements
    - Auditing, analysing requirements and delivering changes requires both data analysis and code tools, but expertise is scarce
- Effective defence also requires a lot of analysis -making sense of the noise and complexity, so decisions can be made quickly
- Generative AI can be used to accelerate both; building new tools, and building new defences; responding to emerging threats at unprecedented speed

Generative AI presents some entirely new security risks

Generative AI has risks inherent in it, and can also be put to malicious use

- If AI is trained on confidential data, this can potentially be retrieved from the model using carefully designed queries.
- In Adversarial AI, maliciously crafted disinformation is hidden in the training data set, misinforming the model to trigger some nefarious outcome.
- Deep Fakes are potential tools for abuse:
    - Voice recognition based security systems are now easily broken by “deep fake” voice generators
    - There is even the future potential for live interactive video from a fake trusted person

Generative AI also accelerates existing security risks

Generative AI is allowing cybercriminals and state-sponsored threat groups to take control of target systems more efficiently

- Large Language Models (including ChatGPT) can be used by attackers to find and compromise vulnerable systems faster on the internet
- Automated discovery and exploitation of security weaknesses ("fuzzing" and patch analysis)
- Researchers have demonstrated using ChatGPT to write ransomware software, and take over accounts by guessing passwords more efficiently
- Generative AI can write code to help attackers take control of servers and network equipment in larger numbers
    - Security cameras, home wifi routers, home computers, and company IT systems
- This means that denial of service (DDoS) **attacks are also getting bigger**

Fighting cyber war on our own doorstep

How AI is helping protect NZ organisations from geopolitical threats

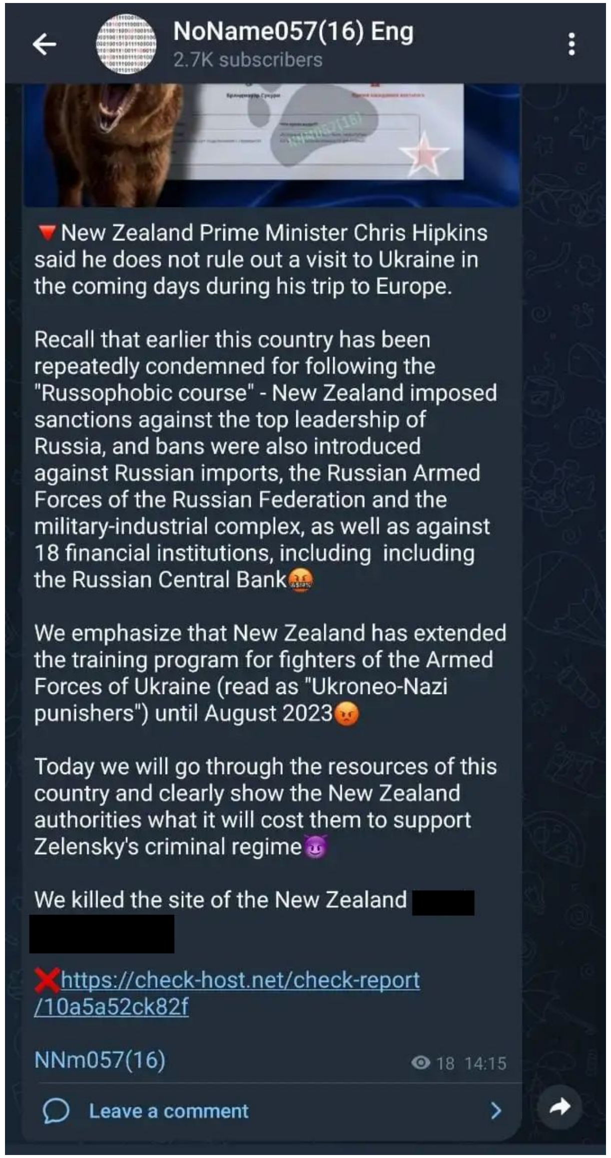
Geopolitical tension brings cyber risk to our doorstep in NZ

In July 2023, NZ Prime Minister Chris Hipkins pledged additional support for Ukraine including equipment, military training, continued sanctions and financial assistance

Pro Russian hacking groups weighed in

“NoName057” threatened to attack NZ government websites

- Specifically calling out sanctions imposed by NZ, and PM Hipkins’ comments in support of Ukraine, as the motivation



🔻New Zealand Prime Minister Chris Hipkins said he does not rule out a visit to Ukraine in the coming days during his trip to Europe.

Recall that earlier this country has been repeatedly condemned for following the "Russophobic course" - New Zealand imposed sanctions against the top leadership of Russia, and bans were also introduced against Russian imports, the Russian Armed Forces of the Russian Federation and the military-industrial complex, as well as against 18 financial institutions, including including the Russian Central Bank🤬

We emphasize that New Zealand has extended the training program for fighters of the Armed Forces of Ukraine (read as "Ukroneo-Nazi punishers") until August 2023😡

Today we will go through the resources of this country and clearly show the New Zealand authorities what it will cost them to support Zelensky's criminal regime😈

We killed the site of the New Zealand [redacted]

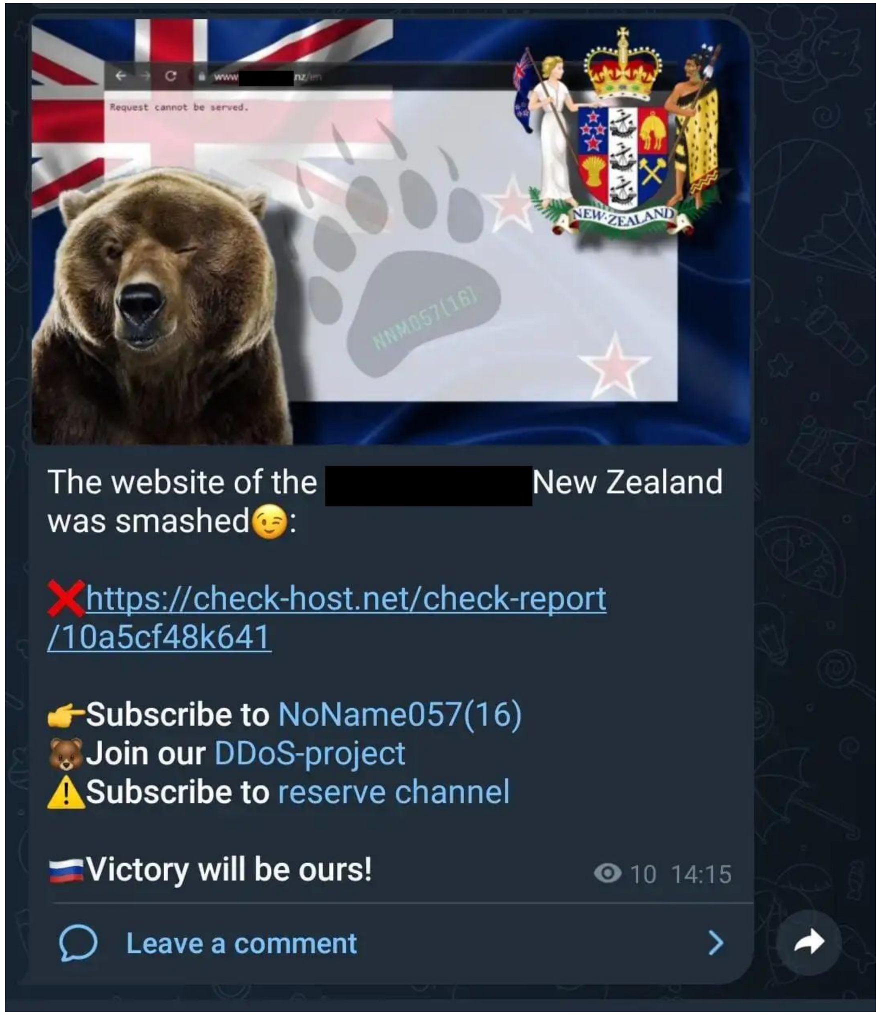
❌https://check-host.net/check-report/10a5a52ck82f

NNm057(16) 18 14:15

Pro Russian hacking groups weighed in

“NoName057” threatened to attack NZ government websites

- Specifically calling out NZ sanctions imposed, and PM Hipkins’ comments in support of Ukraine, as the motivation
- A series of cyber attacks ensued, targeting high profile government websites.
- NoName057 claimed that they had caused disruption to several unprotected websites.



The website of the [redacted] New Zealand was smashed😉:

❌https://check-host.net/check-report/10a5cf48k641

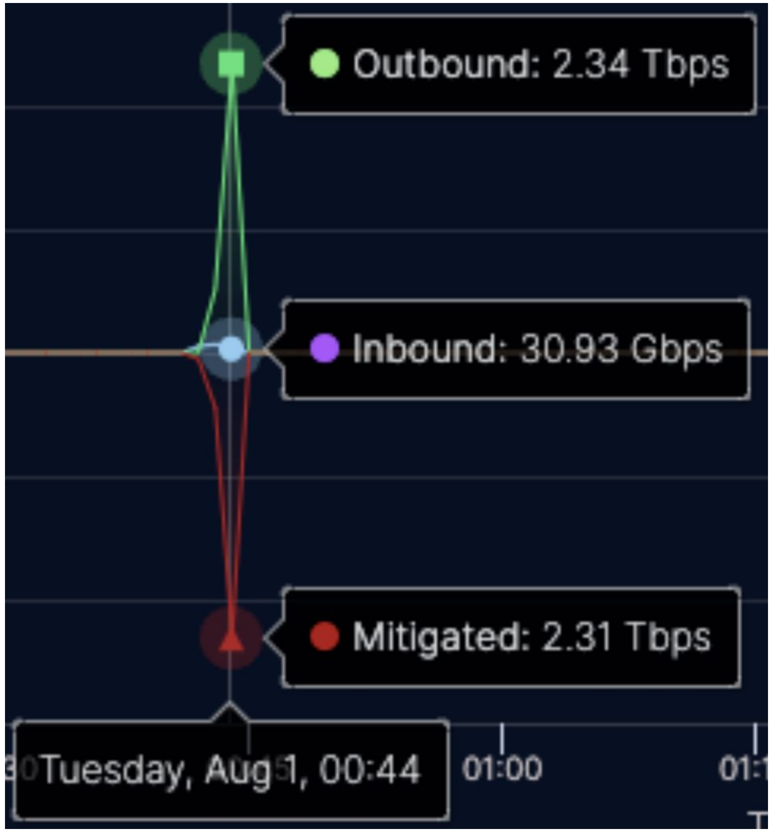
👉Subscribe to NoName057(16)
🐻Join our DDoS-project
⚠️Subscribe to reserve channel

🇷🇺Victory will be ours!

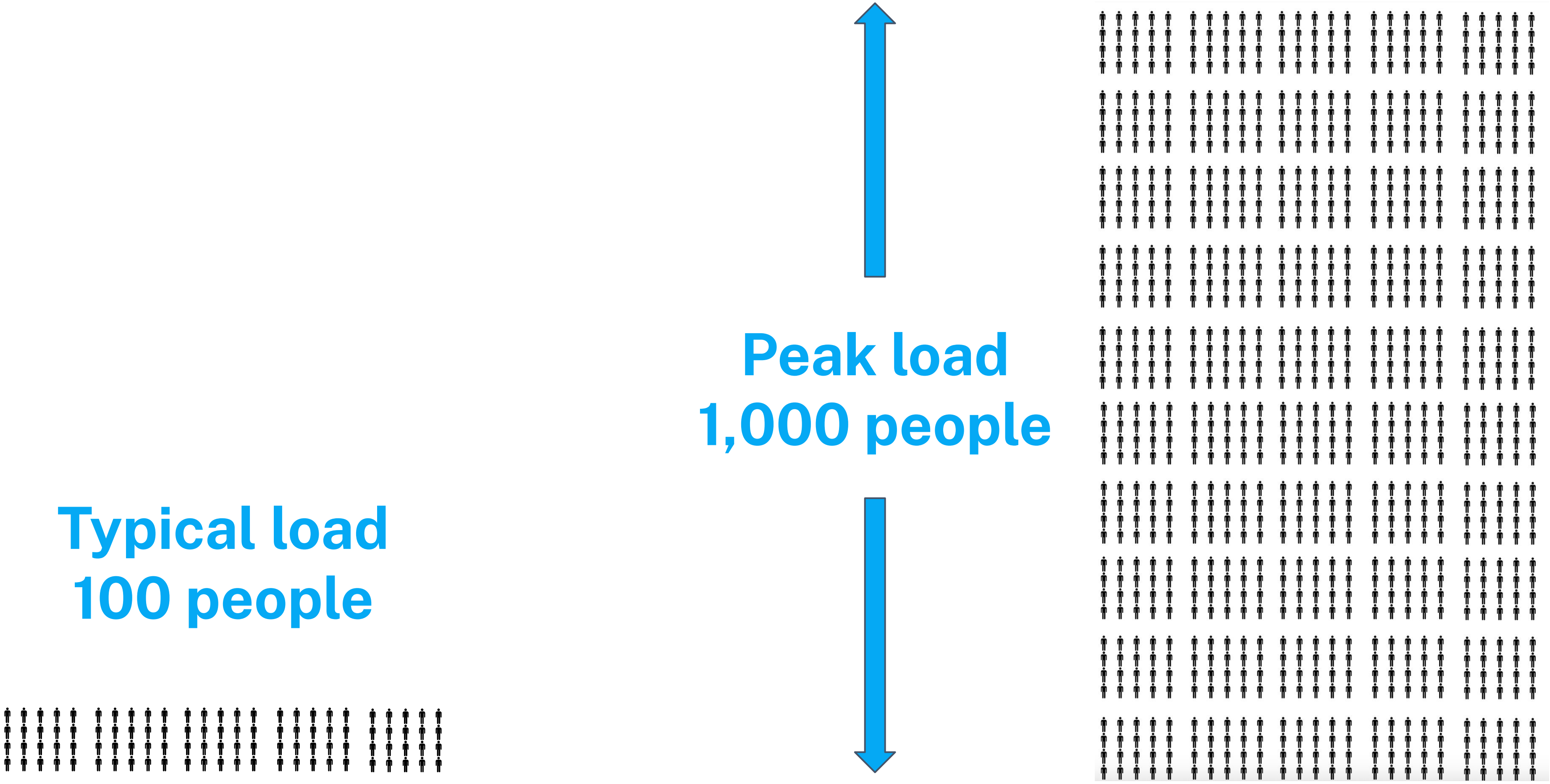
A series of attacks also began targeting websites protected by RedShield

Large scale attacks peaked every ~24 hours

- Large scale attacks occurred throughout the day and night
- Following a repeated pattern targeting particular sites
- Culminating in a 2.34Tbps monster attack, one of the largest DDoS attacks ever reported



Normal user load of 100 - 1,000 people using the site

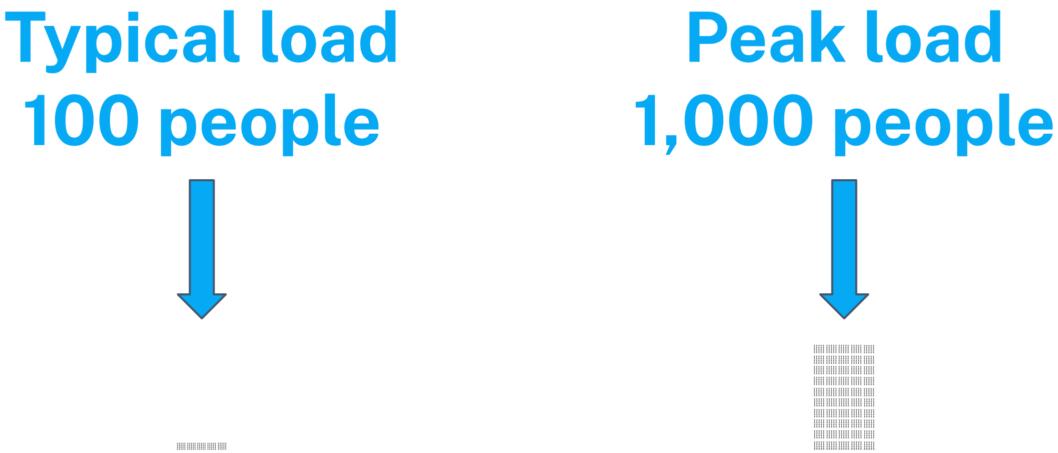


Peak load
1,000 people

Typical load
100 people

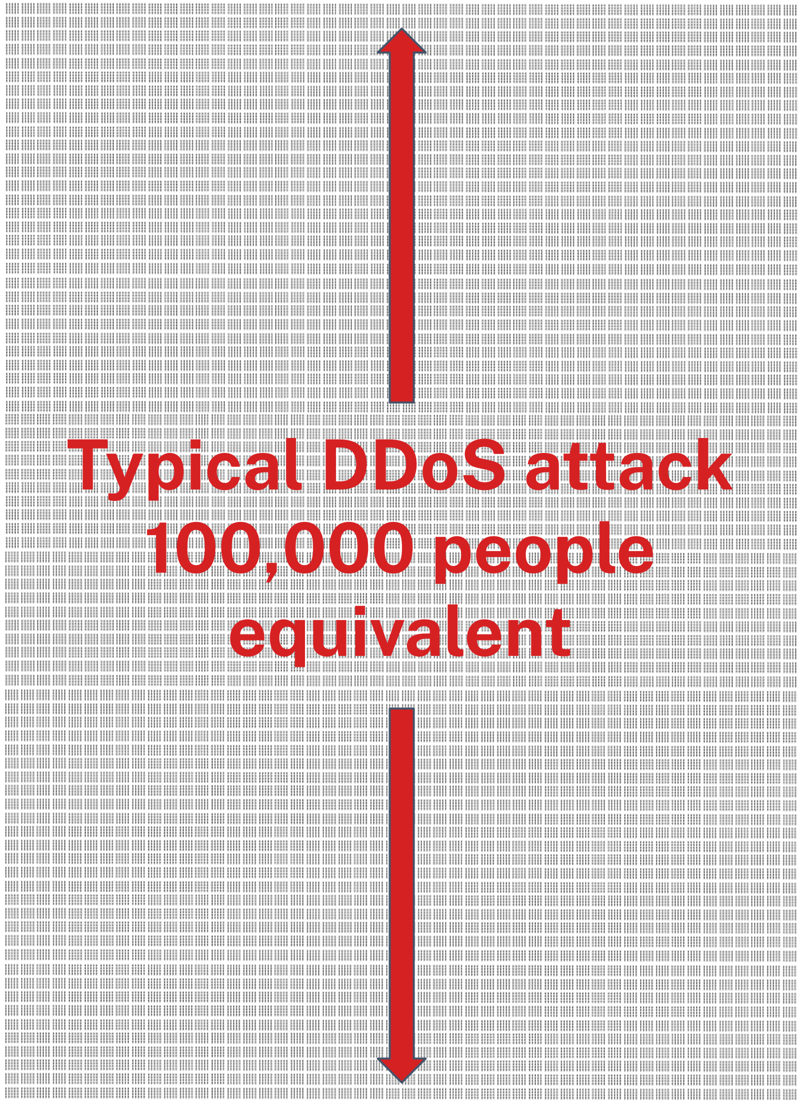
A typical large “DDoS” attack

- On a big day, the website serves 1,000 people concurrently.
- A common DDoS attack would see 100x this traffic, equivalent to 100,000 normal users.
- DDoS attacks like this are designed to massively overwhelm networks and servers.



Typical load
100 people

Peak load
1,000 people

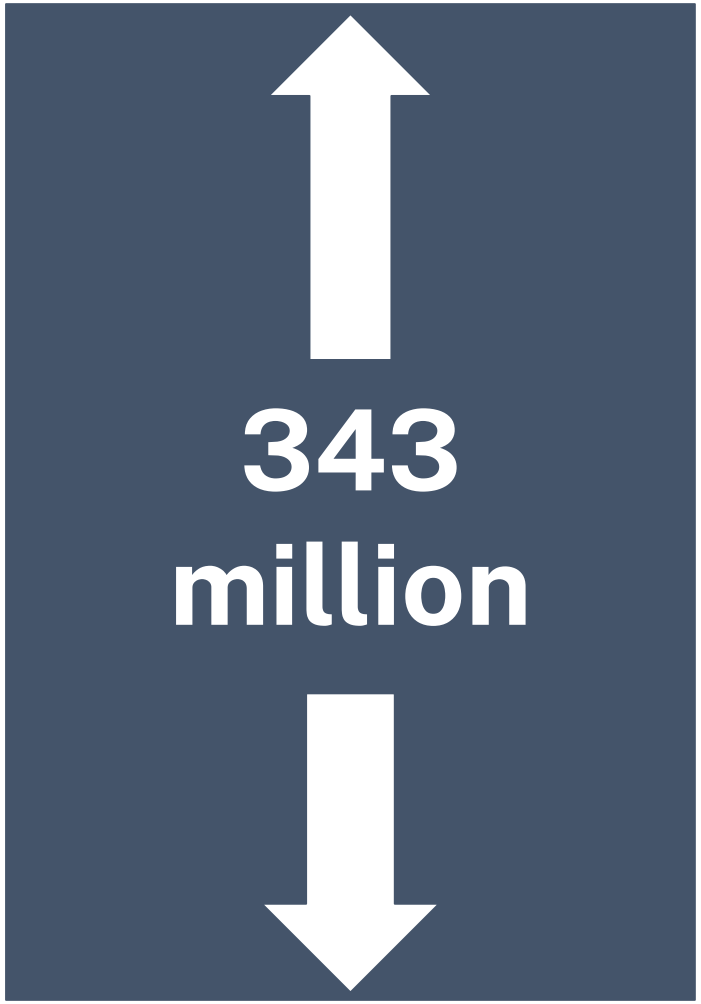


Typical DDoS attack
100,000 people equivalent

Colossal attack, equivalent to 343 million simultaneous web visitors

- This attack consumed network bandwidth equivalent to the entire population of the United States all visiting the website at once.
- Identifying and blocking the bad traffic throughout an attack at this scale, whilst providing uninterrupted service to good users, requires a massive globally distributed infrastructure, with autonomous machine-driven controls

1000
100,000



Meanwhile, if DDoS is a battering ram, hacking attempts are like lockpicks

Separate to DDoS attacks, website hacking attempts continue to hide within in the good traffic:

12.6 billion	Legitimate website transactions delivered by RedShield per month
100 trillion	Security inspection point checks performed on website transactions per month - checking for good vs bad
60 million	Security checks performed per second at normal peak, across all customer website traffic
40 million	Malicious hacking attempts blocked in August
>80,000	Website vulnerabilities under management

- During August 2023:
  - Across the network, we routinely performed 60 million security checks per second, or over 100 trillion checks during the month, to determine the good from the bad.
  - All of this security analysis must be performed in real time, without interrupting any good users interacting with the websites, and without ever letting anything bad through which could lead to a compromise.
- Human security teams could never keep up with this task without autonomous defences, governed by layers of intelligent automation.

Ways in which AI is helping us fight the cyber war

Unsupervised learning

- Autonomous network blocks obvious illegal activity
- Suspicious interactions with apps (eg, excessive login attempts) trigger defences
- Attack data is analysed for intelligence sharing and reused in defence
- Smart scanners learn and crawl applications looking for flaws
- Automated agents continuously update security policies to keep defences up to date with new threats and vulnerabilities

Supervised learning

- Consultants train and manage models to recognise false positives, and finetune policies per website

Generative AI

- “Everyone has a pocket developer”
- Rapid analysis of vulnerabilities and exploit code
- Producing code for shield templates and reporting

Future development

- End-to-end toolchains (langchain) will offload complete security processes
  - *Vulnerability announcement -> analysis -> shield development and deployment -> change control and customer service announcements in minutes*

Tips for NZ organisations - we need a sense of urgency

The NZ economy stands to benefit tremendously from the incorporation of AI tools into daily business processes. Much advice around AI usage advocates designing security in from the start, and ensuring that security requirements and policies are well documented before permitting access to AI toolsets.

Unfortunately, adoption is far ahead of policy, due to strong demand.

Companies which are slow to allow LLMs may struggle to enforce the ban, and also risk being left behind by industry disruptions already underway.

80% of the Fortune 500 has employees using ChatGPT for work already

Tips for NZ organisations - we need a sense of urgency

- Ensure that your organisation has a pragmatic policy relating to the use of cloud-hosted AI tools such as ChatGPT.
    - Many free resources exist on the internet to inspire the strategy and details of this policy.
    - Ensure that this process is moving quickly -even conservative organisations should be well along the path.
    - “If you can put it into a Google search box, you can put it into ChatGPT”.
- Use corporate credentials and MFA to authenticate to AI systems such as ChatGPT -chat history and file uploads are visible.
- Address the early adopters who set up their personal accounts before it was permitted at your organisation, and may have paid to upgrade to ChatGPT+ and API access.
- Ensure all users are well advised about sharing sensitive information which may effectively enter the public domain.
- Set up discussion areas in Slack and Teams, share resources and updates about creative and useful ways to benefit from generative AI. Provide positive and negative examples of acceptable usage.

https://www.kordia.co.nz/ai-usage-policy-checklist

RedShield

www.redshield.co

sales@redshield.co