AI's impact on cybercrime

- AI lowers the bar for cybercrime actors
- With AI handling time-consuming research, cybercriminals are focusing on more complex tasks

Impersonation

Cyber Threat Amplification

Direct Social Attacks

Harmful Content Production

Microsoft’s responsible AI principles

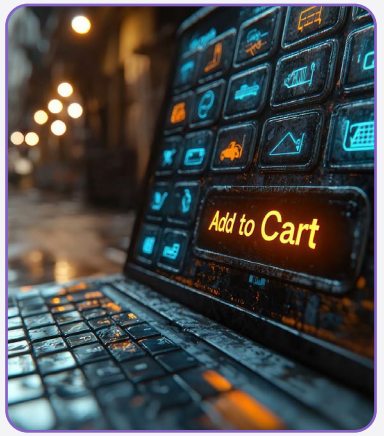
Responsible AI Principles

- Fairness
- Reliability & safety
- Privacy & security
- Inclusiveness
- Transparency
- Accountability

Factors of Interest

Unknown origin of access

Types of content



Conspiracy & market for illicit access

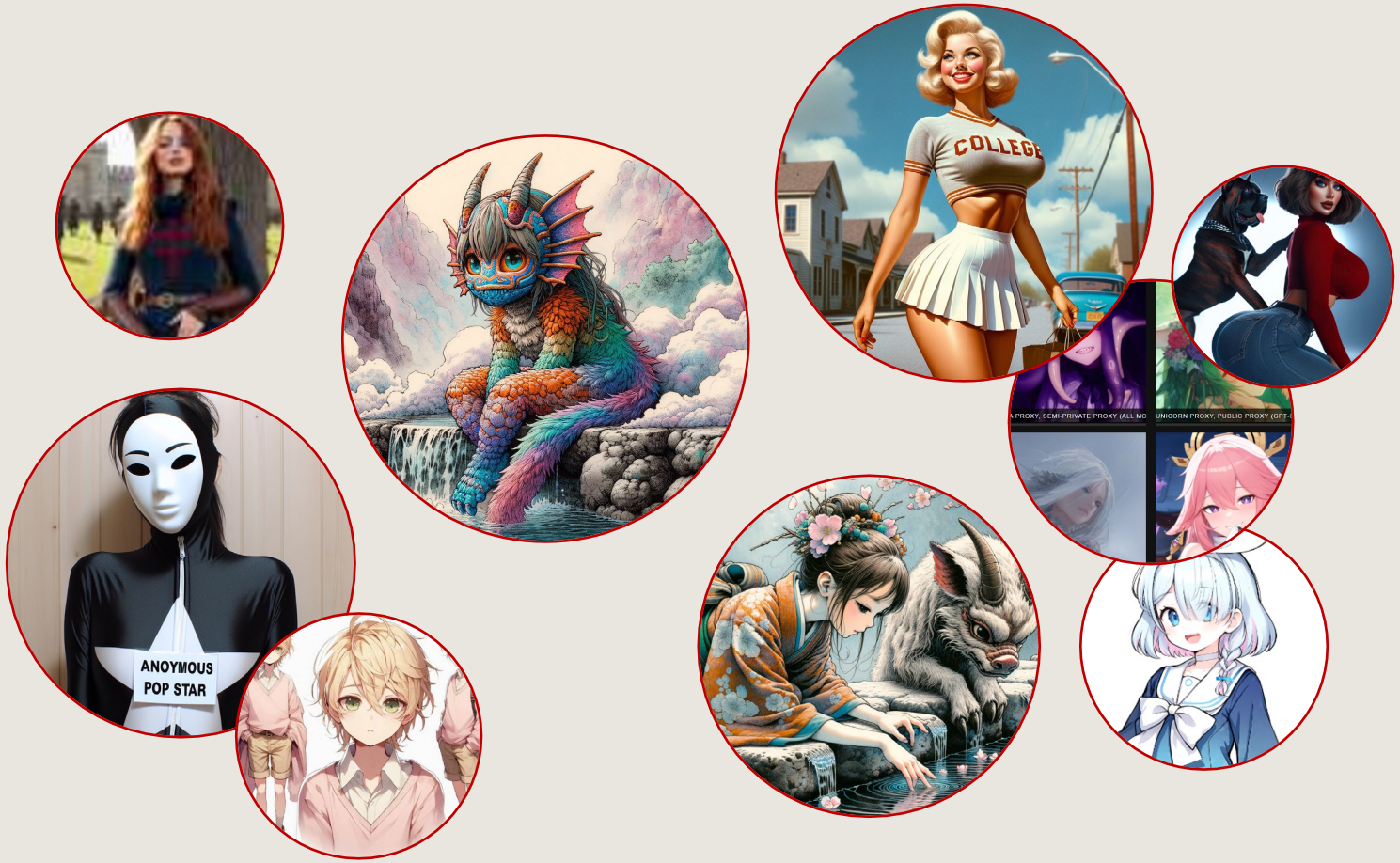
Microsoft

FizzDogg

Disrupting a global cybercrime network abusing generative AI

Who were the groups targeted?

Women, minority groups, celebrities, children:
Abusive images



Customers:
Stolen keys & tokens

Understanding the criminal network

Creators

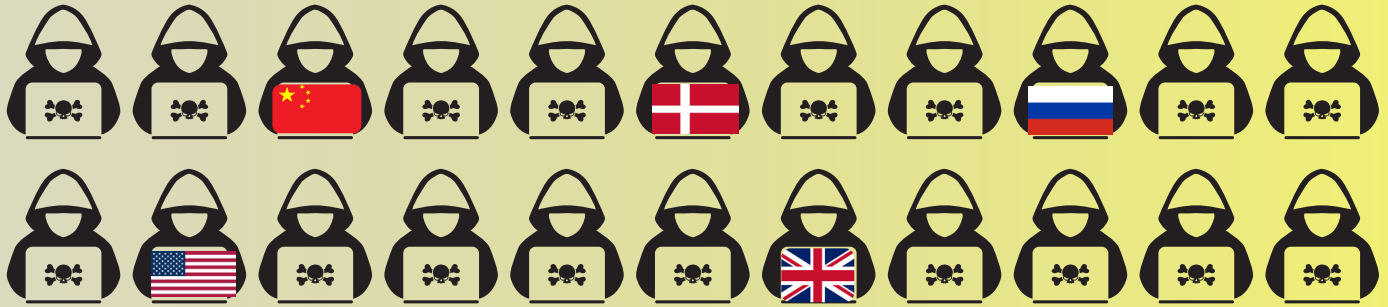
Created or contributed to tooling (ai-proxy service) that allowed threat actors to provide proxy service(s) for activity

Providers

Developed tools and websites to enable users to generate images that violate content restrictions

Users

Created and distributed harmful images



Identification of key adversaries

- Seizure from first filing provided additional evidence
- Identified developers behind the tools

Creator

cg-dot

Hong Kong, China

Contributed to the development of the oai-reverse-proxy, enhancing its functionality enabling its use on the Google Cloud Platform (GCP)

Provider

Fiz

Iran

Created the de3u tool and proxy infrastructure to generate images using stolen Azure APIs

Provider

Drago

United Kingdom

Managed proxy landing pages that facilitated unauthorized access to AI services to generate images using stolen Azure APIs

Provider

Asakuri

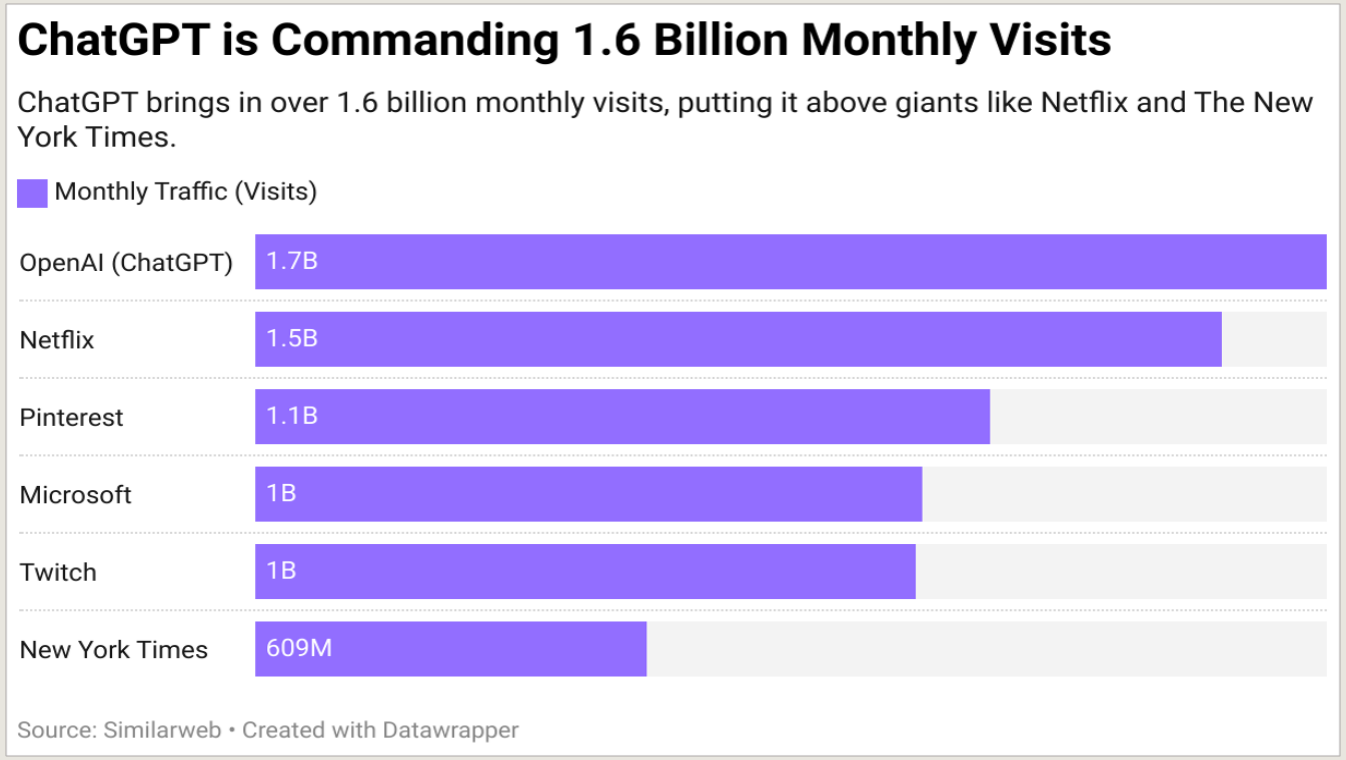
Vietnam

Operated the Yae Miko Proxy, facilitating unauthorized access to AI services to generate images using stolen Azure APIs

Microsoft remediation

- 379 compromised API keys rotated & customers notified
- Microsoft Defender protections and alerts added
- Guardrails strengthened
- Additional internal-facing fraud & abuse detections created
- Improved social listening systems implemented
- Abusive images and prompts reported to US National Center for Missing and Exploited Children
- New SDL requirement around API key management discovery
- Investigator guidance for AI investigations

Shadow AI usage...



78%

of employees are bringing their own AI Tools to work.

- 2024 Microsoft Study | N=3400

DeepFake & AI Enabled Candidates

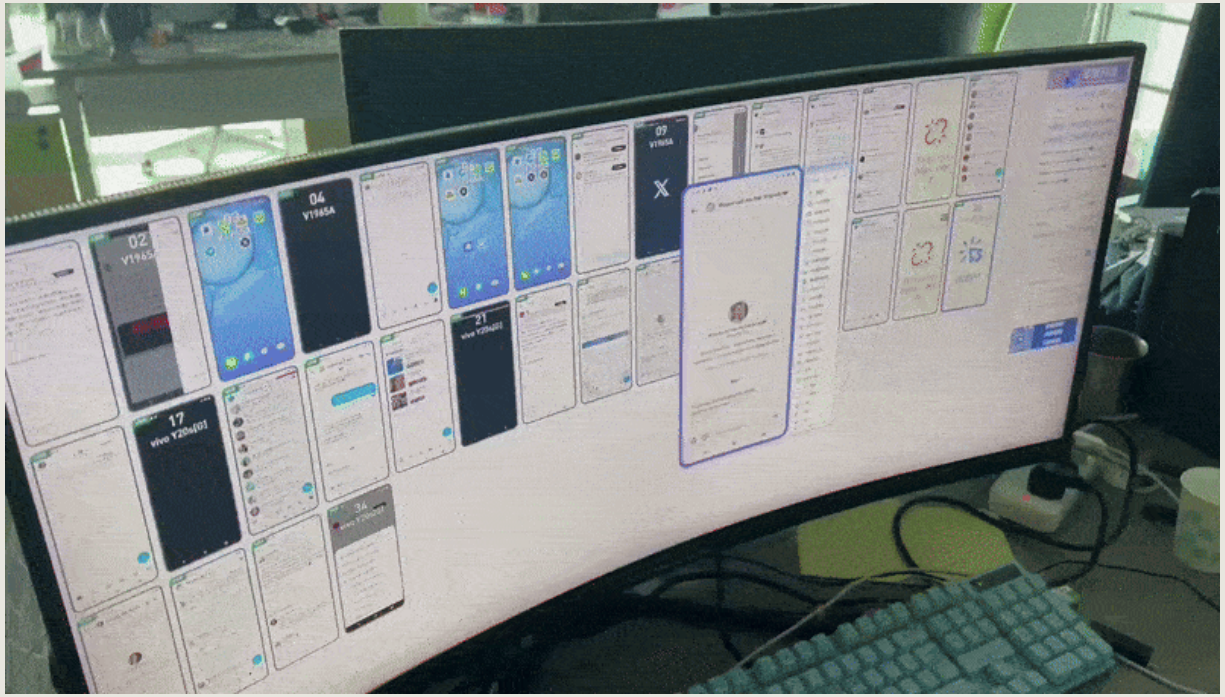
17%

hiring managers say they've encountered candidates using deepfake technology.

- 2025 Resume Genius Study

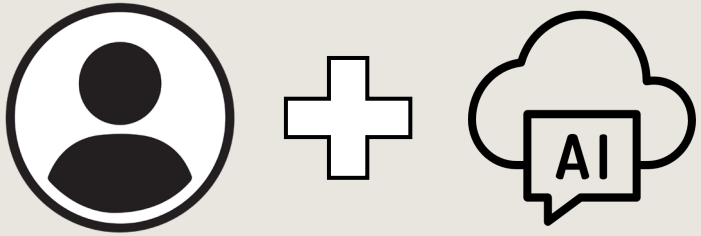
The Past & Future of AI Enabled Crimes

- 'laptop farm' stole $17m / Arizona



- Manus AI Running 50 Computer Tasks

AI Systems | Future of Work



Human + AI(s)

The future of how humans interact with AI is rapidly approaching us.

However, my role is to help organizations work through these phases:

1. Working with AI
2. Managing AI
3. Informed by AI
4. ~~Replaced by AI~~