National Cyber Security Centre

PLATFORMS RESEARCH

UK National Cyber Security Centre on GenAI

Concerns / Opportunities / Directions

ITU Workshop 19th February 2024

National Cyber Security Centre

PLATFORMS RESEARCH

Whoami and where does GenAI sit at NCSC

Research

Data Science Research

Platforms Research

Policy

Operations

Internal IT / Engineering

Comms

National Cyber Security Centre

PLATFORMS RESEARCH

Prompt: “uk civil servants looking concerned”

Our Concerns

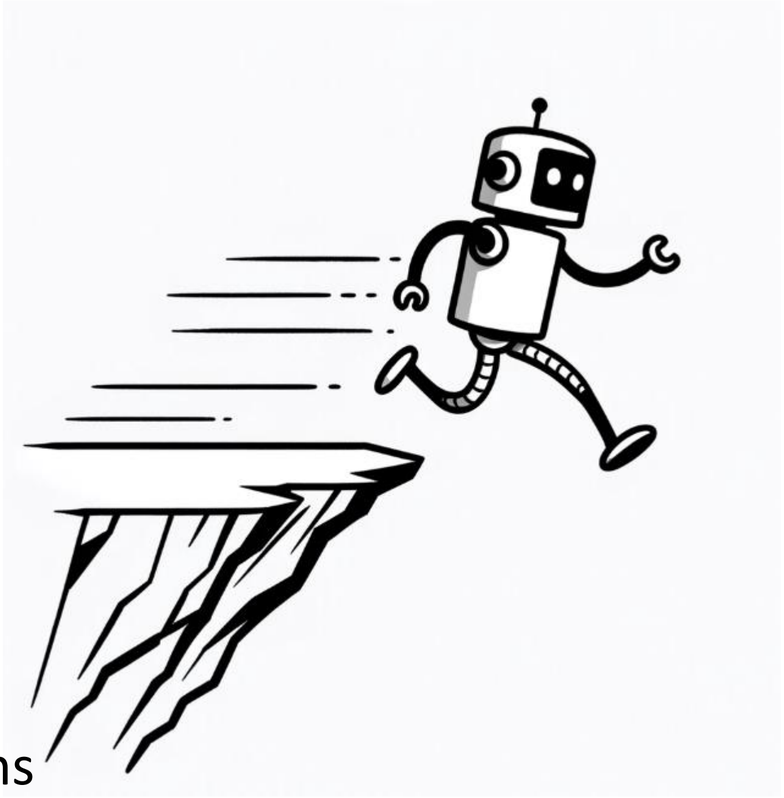
National Cyber Security Centre

PLATFORMS RESEARCH

Our concerns

Rushing and creating risk

- Haste overcoming need for secure by design
- Inadvertently creating dangerous or harmful systems
- Not protecting data when building a system
- Not protecting the system from AI-specific attacks
- Not putting in the controls we'd expect of "normal" systems



National Cyber Security Centre

PLATFORMS RESEARCH

How do we safely go from cruise control to self-driving?

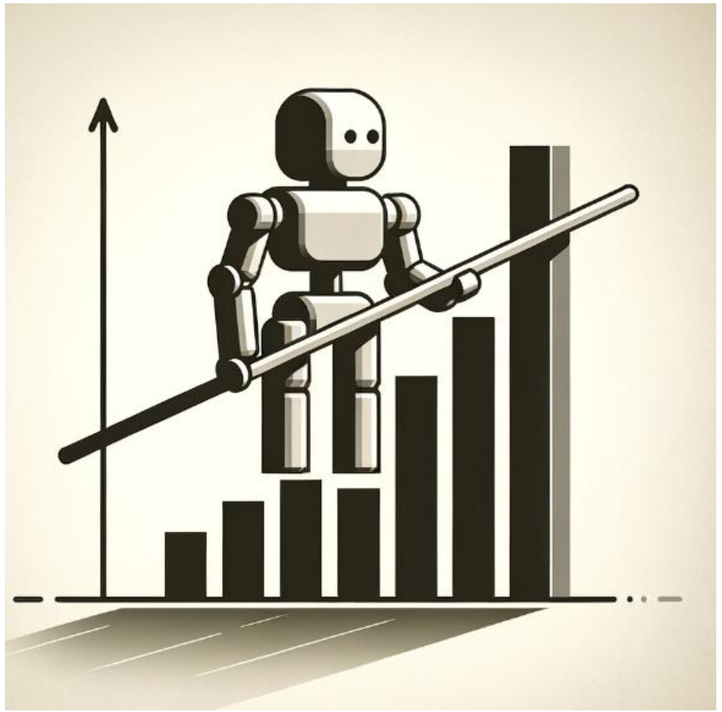
- Currently people don’t trust GenAI to do much
- The potential for what it could do is huge
- How do we learn to evaluate GenAI for specific tasks

National Cyber Security Centre

PLATFORMS RESEARCH

Our concerns

Lowering the bar for attackers



- Coaching attackers
- Translation (including domain translation)
- Capabilities that might have not previously been within their ability

National Cyber Security Centre

PLATFORMS RESEARCH

Our concerns

Use of GenAI's strengths for harm

- Harmful content
- Data analysis

National Cyber Security Centre

PLATFORMS RESEARCH

Prompt: “many uk civil servants looking hopeful”

Areas of opportunity

National Cyber Security Centre

PLATFORMS RESEARCH

Areas of opportunity

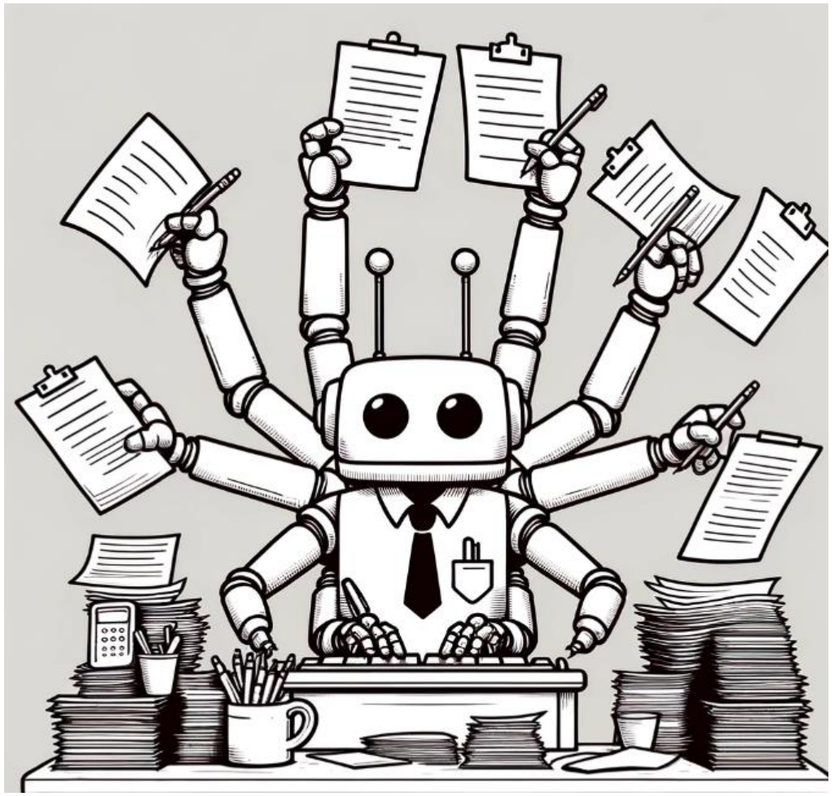
Attackers **are** using GenAI, defenders need to be able to

National Cyber Security Centre

PLATFORMS RESEARCH

Areas of opportunity

Automating tasks that currently take defender time

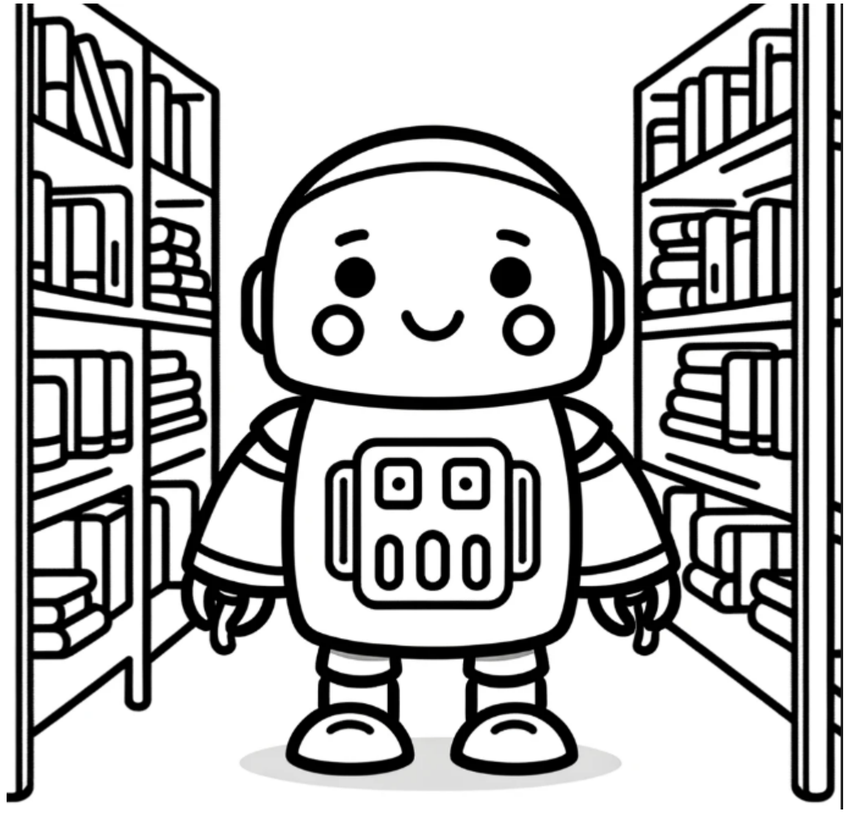


National Cyber Security Centre

PLATFORMS RESEARCH

Areas of opportunity

Able to look at the data defenders don’t have time for



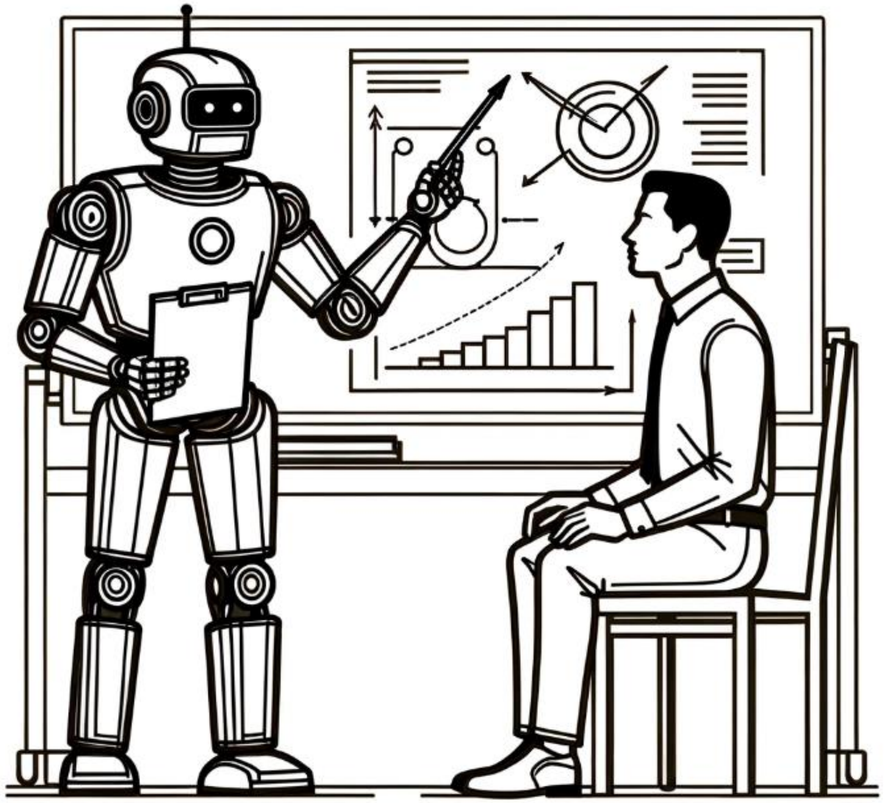
National Cyber Security Centre

PLATFORMS RESEARCH

Areas of opportunity

As "always available" coaches

- For defenders, coders etc
- For everyone to have access when required to guidance – help security make sense for them



National Cyber Security Centre

What are NCSC doing?

Research
- internally
- in academia
- in industry

Publications such as the Guidelines for secure development of AI systems

Working with ETSI SAI Group (est 2019)

Working with the AISI

Co-Published with 23 international governmental cyber security organisations

Building our policy position

Using them ourselves

National Cyber Security Centre

PLATFORMS RESEARCH

In Summary

- NCSC are concerned both with people making insecure systems as well as actors using systems in a way that increases their capability
- We feel there are huge advantages for defenders
- Whilst there are a lot of research questions still needing answering, international collaboration will be key