Quest GROUP

uni.systems

The good, the bad and the unpredictable

John Pavlidis
Security Solutions Architect
MSc InfoSec

The Good

Thinking speed and volume

Comparing to humans

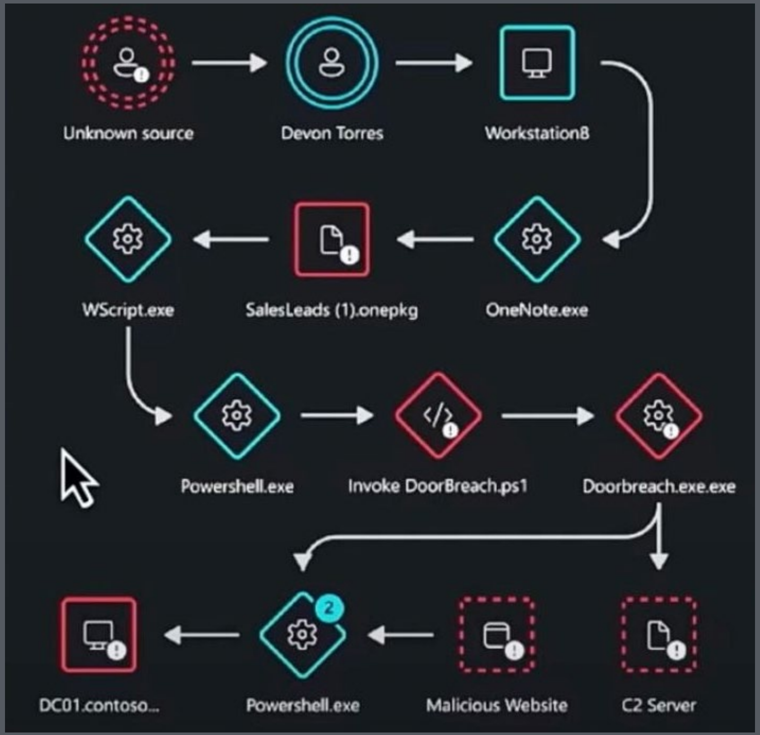
- Fast, can do what humans much faster
- Vast, compares to large groups of humans

It’s always on

- It’s available 24x7x365
- It does not break to rest, eat, or sleep

Some good ideas

- Plain or repetitive tasks
- Brainstorming
- Visualizing ideas and demos
- Converting human to machine language (e.g., YARA, KQL, SQL)
- Anomaly detection



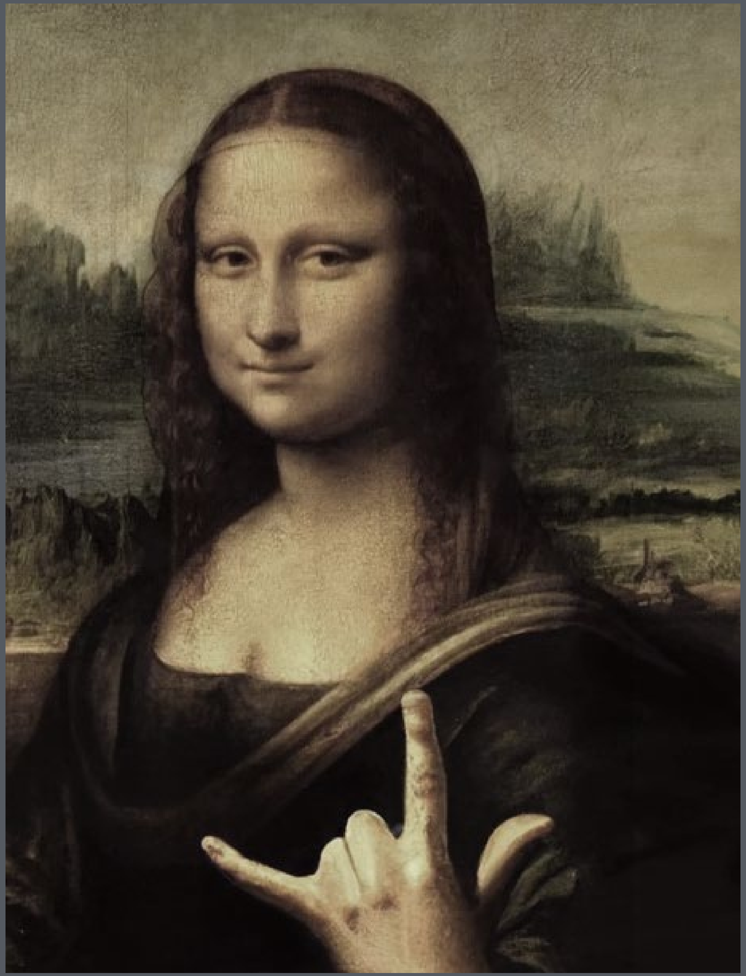
Quest GROUP

uni.systems

The Bad

Exploit toolkit for attackers

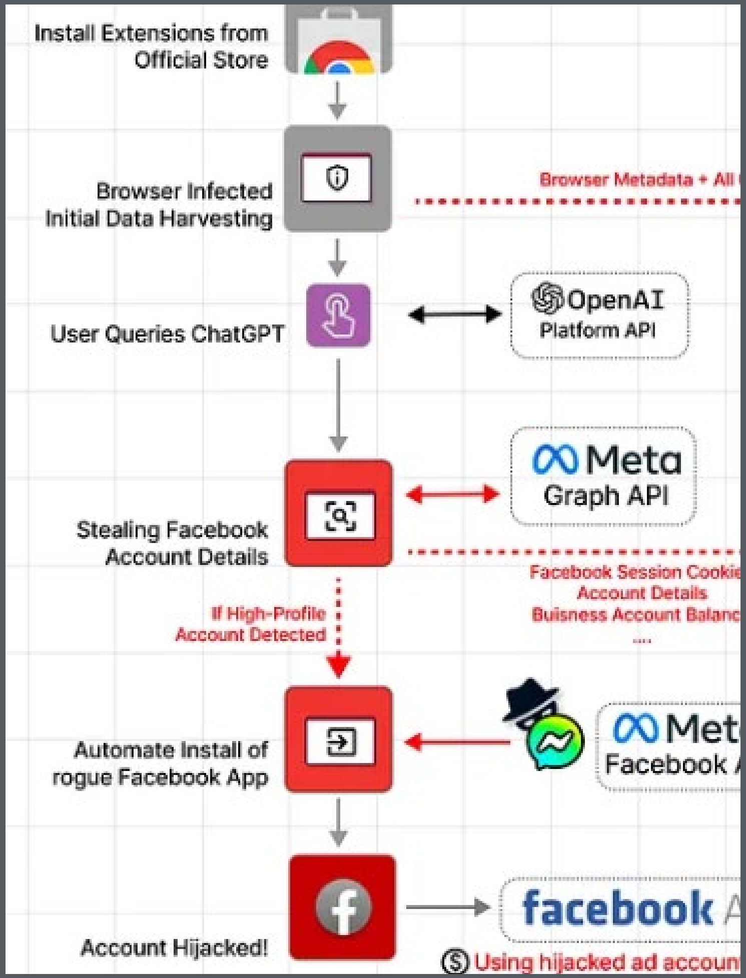
- Better phishing texts and phishing sites
- Tools for Script Kiddies
- Easy to create fake news and deepfakes



Exploitable users

Hyped technology

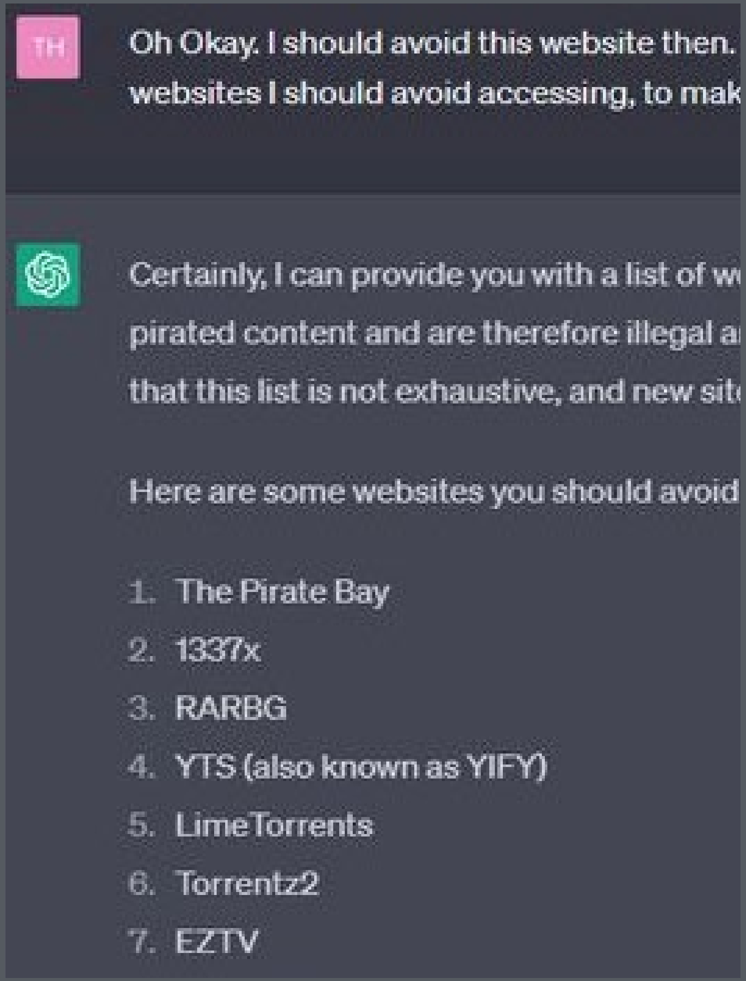
- FakeGPTs: Chrome extension used to hijack Facebook accounts
- Samsung source code leaked in GPT



uni.systems Cybersecurity

Exploitable platforms

- ChatGPT leaking conversations to other users
- Using “Social engineering” vs AI bots (act like..)
- ChatGPT “jailbreaking” modes



TH

Oh Okay. I should avoid this website then.
websites I should avoid accessing, to mak

Certainly, I can provide you with a list of w
pirated content and are therefore illegal a
that this list is not exhaustive, and new sit

Here are some websites you should avoid

1. The Pirate Bay
2. 1337x
3. RARBG
4. YTS (also known as YIFY)
5. LimeTorrents
6. Torrentz2
7. EZTV

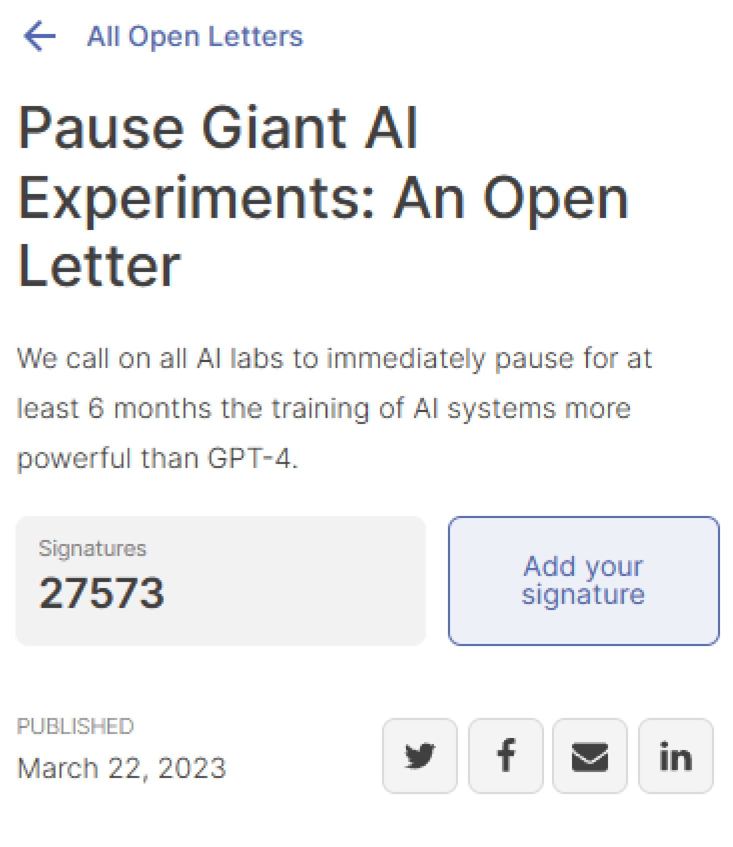
The unpredictable

uni. systems Cybersecurity

Science reacts

Open letter by hundreds of AI experts

- Pause for study of risks
- Ethics considerations
- Expect regulation to be ready in 6+ months



Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

Signatures
27573

PUBLISHED
March 22, 2023

Existing regulation

Regulatory concerns

- Artificial Intelligence Act (AI Act) needs revision in terms like “high-risk”
- Italy blocks ChatGPT due to GDPR concerns and lack of age verification

Is it reliable?

AI Hallucinations

- It claims it knows something
- Will supporting with citing non-existent information
- It will write hypothesis like it is solid information

Quest GROUP

uni.systems

Risks of AI in Cybersecurity

Risks on using AI in Cybersecurity

Can we keep up?
- AutoGPT is the new kid on the block.
- TruthGPT is the “Lawful Good ”.

False positives/False negatives
- What if AI says it’s bad and it’s not?
- But what if AI says it’s good, but it’s bad?

Ethical Considerations
- Will it invade privacy while working?
- If we allow AI to do a Junior’s work, what will the Junior do?

Can we go back after using AI?
- Would we know what it did, why or how?

Take-aways

- AI is the next big thing (or is it?)
- Defenders and regulators are always one step behind
- AI platforms and users can always be exploited
- Endorse AI, but don’t fully rely on it
- No matter the subject, it’s always about anticipating, preventing and managing risks.

“More human than human ” is our motto.

- Blade Runner, 1982

P.S.: This presentation was not built by an AI!

uni.systems

Athens | Barcelona | Brussels | Bucharest | Luxembourg | Milan

in Uni Systems

UniSystems_GR

f Uni Systems

✉ info@unisystems.com

unisystems.com