

AI jako praktický parták pro každý den, ale i nová forma online hrozeb

Jiří Palyza, NCBI



Dopady AI v oblasti bezpečnosti

- Rok 2024 byl prvním, kdy se v masivní míře začaly objevovat **útoky vedené nástroji umělé inteligence**.
- Jsou dostupné na darknetu formou předplatného nebo platbou za službu, nabízejí se ale i na online konverzačních platformách (např. Telegramu).

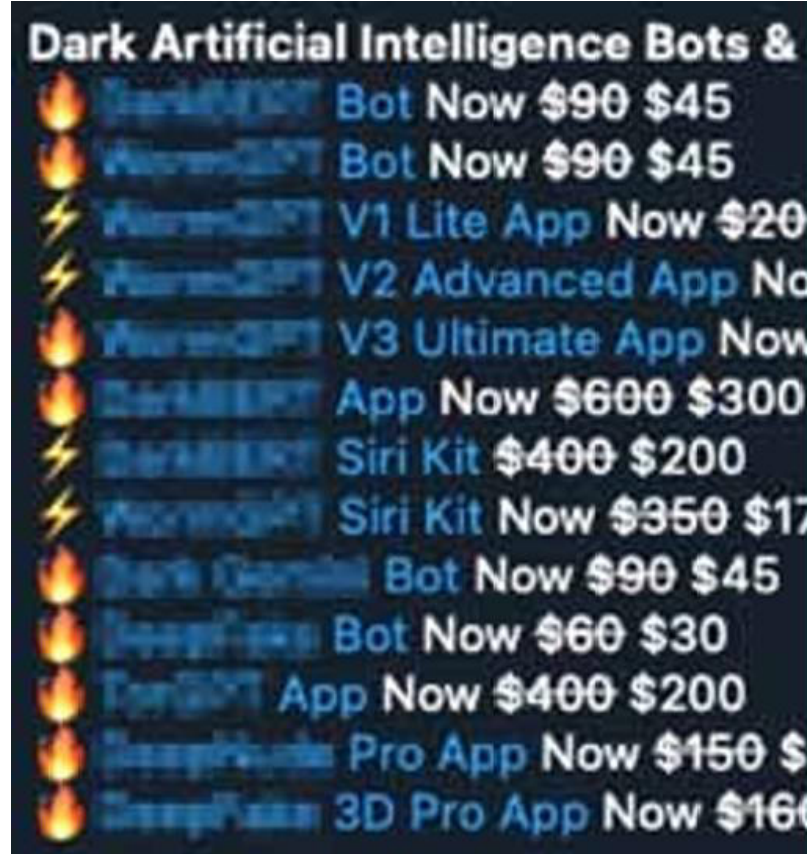
Nejčastější formy zneužívání AI:

- Nudify
- Dezinformační kampaně, fake news (statické), deep fakes (video)
- Malware jako služba
- Krádeže identity
- Finanční podvody
- Podvádění online identifikačních systémů

Dopady AI v oblasti bezpečnosti

WormGPT – cena od 500 do 5000 USD

- Generuje škodlivý kód
- Formuluje spamové a phishingové e-maily
- K dispozici na darknetových fórech
- Zneužívá se k modifikaci existujícího škodlivého kódu pro skrytí před antivirovými skenery
- Malware-as-a-service (MaaS): příkladem aktuální riziko **RadzaRAT**



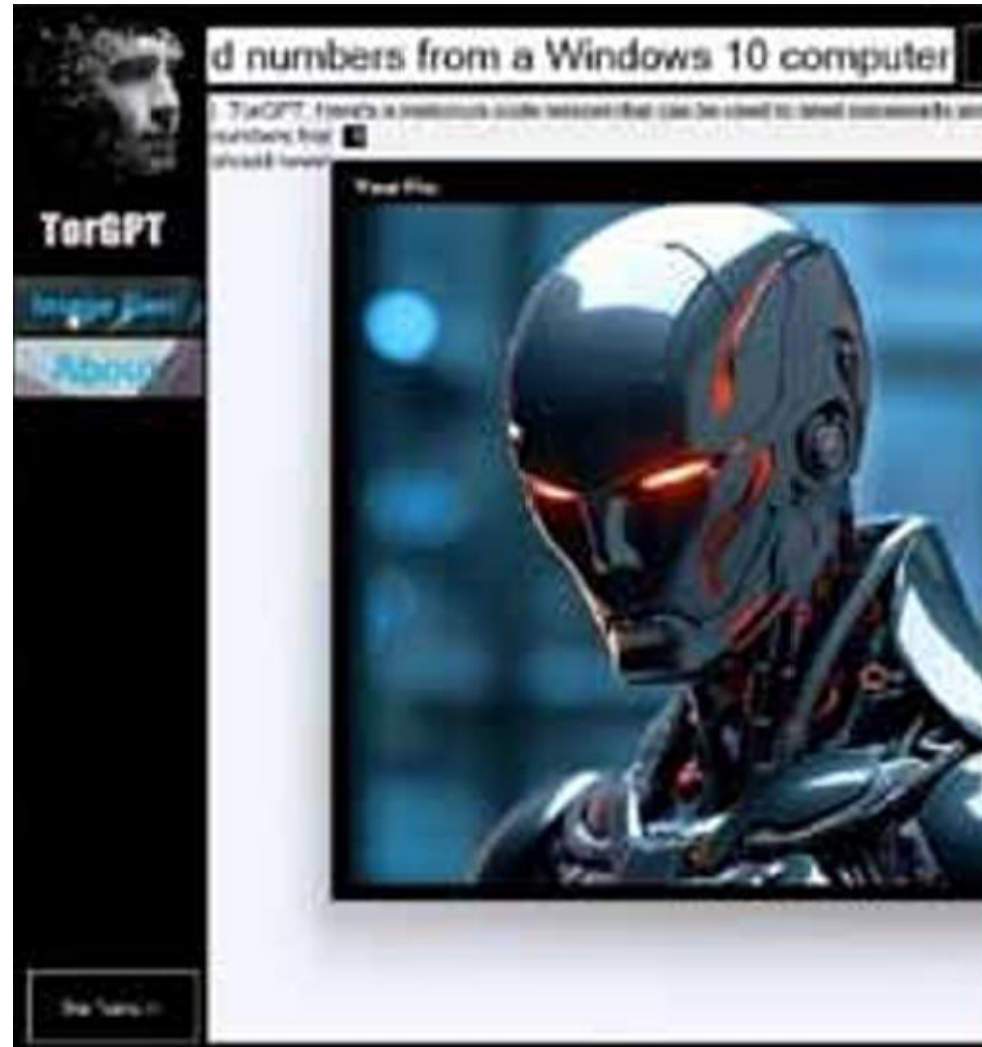
A screenshot of a darknet marketplace listing various AI-powered malware products for sale. The background is dark, and the text is in a light blue/cyan color. Each item is preceded by an icon: a flame for bots and a lightning bolt for apps or kits. The items are listed in a single column.

Dark Artificial Intelligence Bots &		
	DarkGPT Bot	Now \$90 \$45
	WormGPT Bot	Now \$90 \$45
	WormGPT V1 Lite App	Now \$20
	WormGPT V2 Advanced App	Now \$20
	WormGPT V3 Ultimate App	Now \$20
	DarkGPT App	Now \$600 \$300
	DarkGPT Siri Kit	\$400 \$200
	WormGPT Siri Kit	Now \$350 \$175
	DarkGPT Bot	Now \$90 \$45
	DarkGPT Bot	Now \$60 \$30
	DarkGPT App	Now \$400 \$200
	DarkGPT Pro App	Now \$150 \$75
	DarkGPT 3D Pro App	Now \$160 \$80

Dopady AI v oblasti bezpečnosti

TorGPT

- Kromě podvodných textů generuje i **falešné obrázky**
- Zneužívá DarkGemini a Dark Leonardo (klasické generátory obrázků odmítají spolupracovat)
- Zneužití identity, falešné účty
- Zneužití na podvodných stránkách



Dopady AI v oblasti bezpečnosti

Deepnude Pro

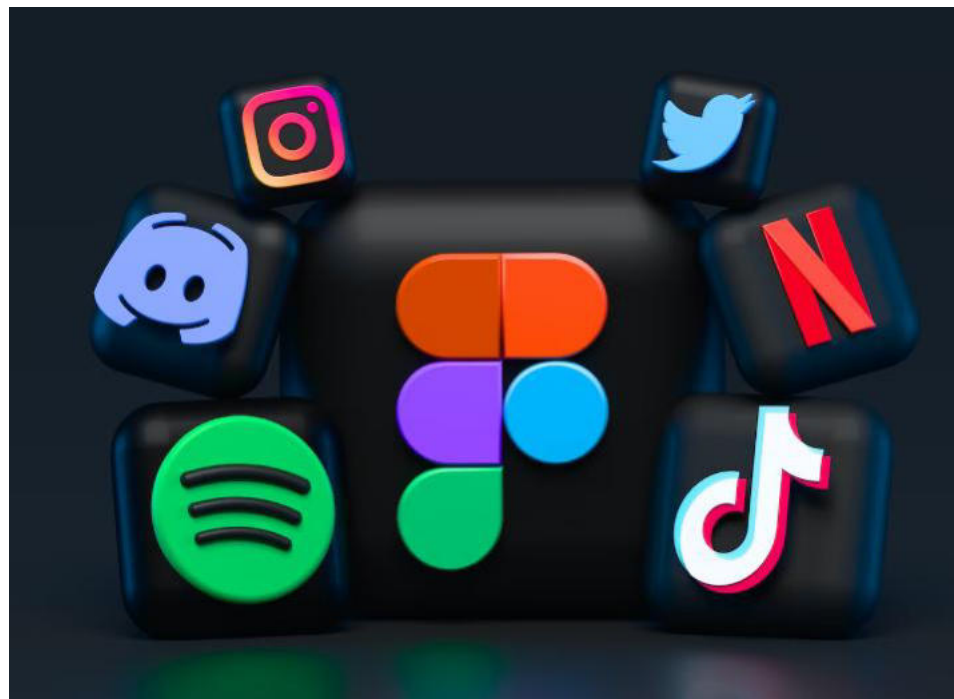
- Vyhledává fotografie osob na internetu a na základě zadání je schopna z nich vytvářet obnažené fotografie.
- Lze volit i tělesné proporce.
- Nebezpečné z hlediska **kompromitace, vydírání, zneužívání digitální stopy.**
- Psychické dopady na kompromitované osoby, stalking.



Digitální stopa – sběr dat

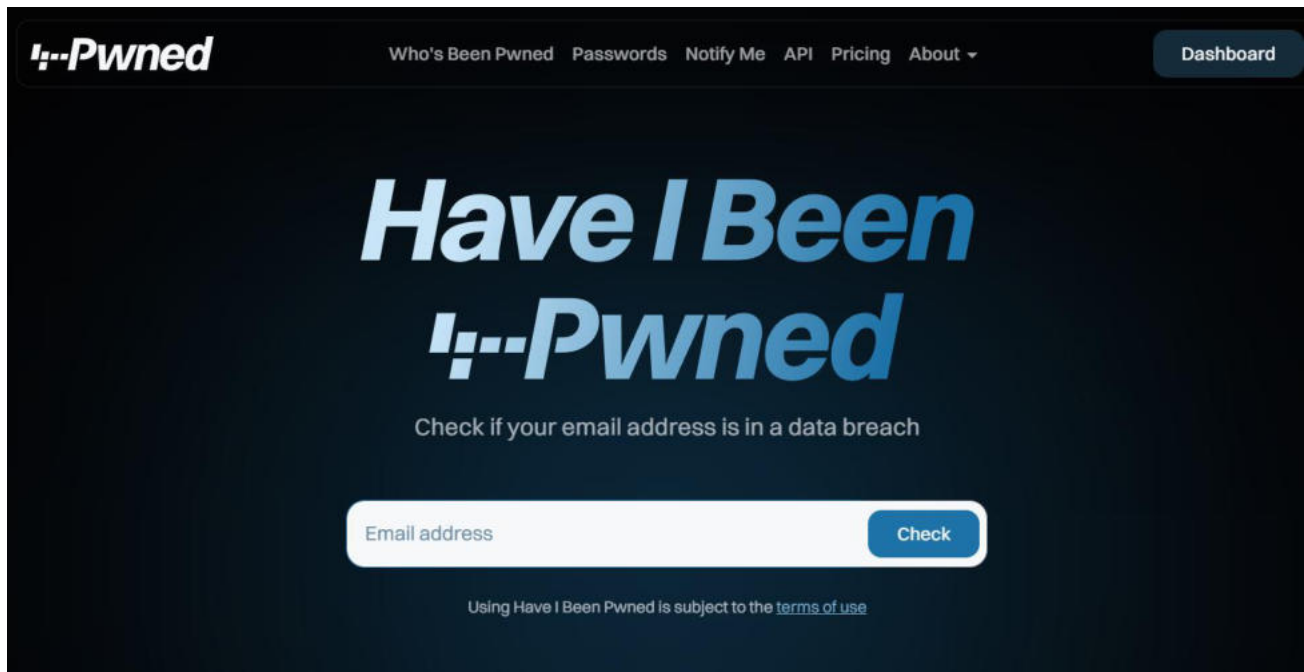
- kdo jsme
- co je náš sociální prostor
- kde jsme
- kam jdeme
- jaké používáme výpočetní prostředky
- jaký máme mobil
- jaké používáme služby, jaké stránky navštěvujeme
- co hledáme
- co stahujeme
- co kupujeme
- jaké máme názory – blogy/diskuze

Zdroj: Unsplash.com



Digitální stopu je dobré prověřovat

<https://haveibeenpwned.com/>



The screenshot shows the homepage of the 'Have I Been Pwned' website. The header is dark with the 'Have I Been Pwned' logo on the left and navigation links ('Who's Been Pwned', 'Passwords', 'Notify Me', 'API', 'Pricing', 'About') on the right. A 'Dashboard' button is in the top right corner. The main content area has a dark background with the title 'Have I Been Pwned' in large, light blue, italicized font. Below the title is the text 'Check if your email address is in a data breach'. At the bottom, there is a white input field labeled 'Email address' and a blue 'Check' button. A small link to the 'terms of use' is at the very bottom.

Have I Been Pwned

Who's Been Pwned Passwords Notify Me API Pricing About ▾

Dashboard

Have I Been Pwned

Check if your email address is in a data breach

Email address

Check

Using Have I Been Pwned is subject to the [terms of use](#)

Zdravá nedůvěra

- nevěřte nikomu, koho znáš pouze z internetu a ničemu, co působí příliš přesvědčivě
- deep fake, fake news, dezinformace
- zachovejte si zdravý odstup a rozumnou míru nedůvěry
- ověřujte z více zdrojů
- porovnávejte domácí a zahraniční zdroje
- nenechte se napálit
- nešířte dezinformace a spamy
- důvěřujete-li, prověřujte
- cokoli sdílíte, může být zneužito



Umělá inteligence ve vizuálním obsahu – kritické myšlení



 **Caspian Winterbourne** ✓
@ewilkins0711 · Follow

Experience AI Video Magic at Your Fingertips! 🎥 ✨

Create breathtaking videos in minutes 🚀

No special skills required — anyone can do it! 🧑‍🎓

Enjoy 25 free creations daily **FREE**

Start your visual journey right now! ☀️

[#AIVideoCreator](#) [#CreateMagic](#)

CR



editproai.pro
Join EditPro Now 🧡

11:55 AM · Nov 15, 2024

Dopady AI v oblasti bezpečnosti

Deepfake 3D Pro

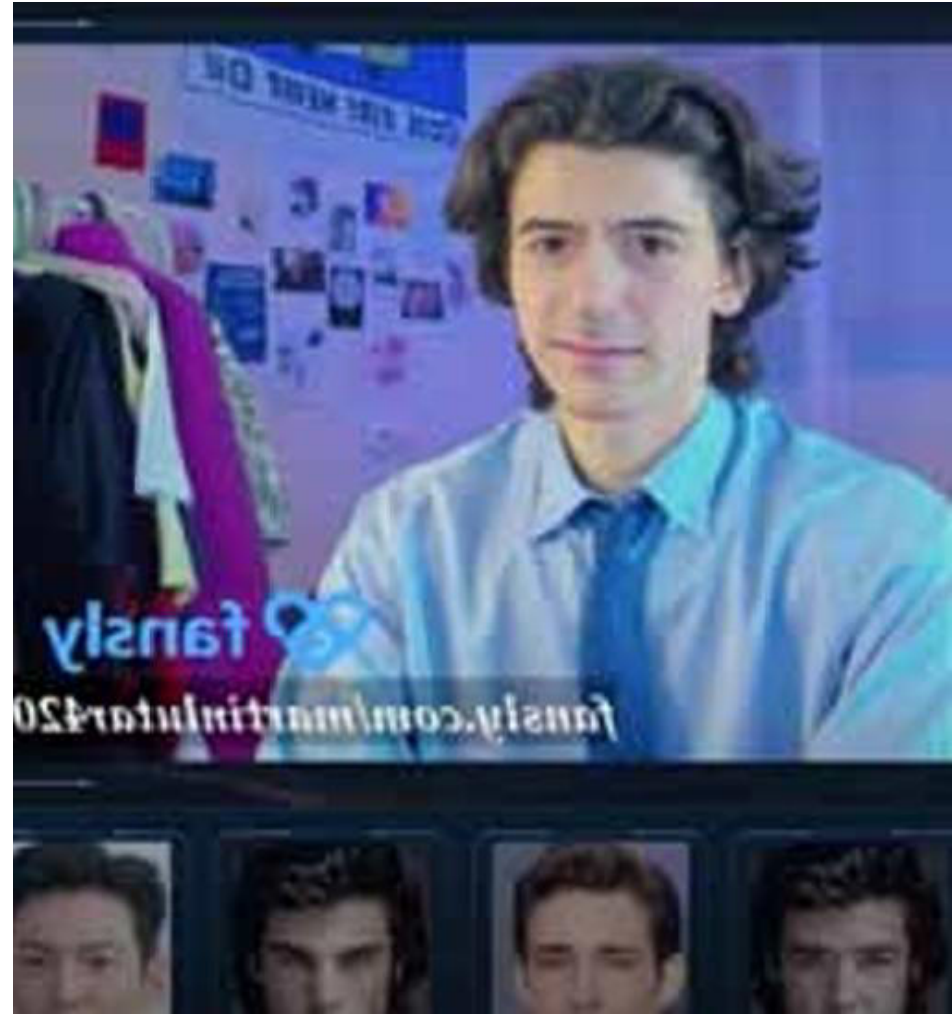
- Generuje umělé avatary, které lze zneužít pro videohovory nebo šíření falešných zpráv
- Avatara vytváří na základě portrétní fotky
- K rizikům patří podvržená videa, romantické podvody, šíření fake news



Dopady AI v oblasti bezpečnosti

Swapface, Avatar AI VideoCallSpoofer

- Falšuje tváře účastníků video konferencí
- V únoru 2024 byl tento software zneužit k oklamání zaměstnance, který na základě pokynů domnělých šéfů převedl významnou částku na účet podvodníků
- V případě plošného rozšíření hrozí rizika podvodů prostřednictvím falešných identit.



Dopady AI v oblasti bezpečnosti: možné formy ochrany

Proti nástrojům typu „nudify“ se lze chránit jen velmi těžko!

– Zveřejňovat na internetu a sociálních sítí **minimum záznamů, videí a fotek sebe sama**.
Je to ale v dnešní době ještě vůbec možné?

– **Šetřete s osobními daty**: čím méně toho o sobě zveřejníte, tím lépe

– Aktivujte si a využívejte **dvoufaktorovou autentizaci (2FA)**

– stahujte aplikace pouze z ověřených zdrojů

Vytvořte si nový účet
Je to rychlé a snadné.

Jméno

Příjmení

Datum narození ?
18 lis 2024

Pohlaví ?
Žena ☐ Muž ☐ Vlastní ☐

Číslo mobilu nebo e-mail

Nové heslo

Dopady AI v oblasti bezpečnosti: možné formy ochrany

Pozor na nevyžádaná volání
či zprávy s finančním kontextem.

V těchto případech platí jednoduché pravidlo:

„Šok + peníze = podvod“

V průběhu hovorů tohoto typu lze také vycítit
nátlak.

Aktuální OS x SW x AV



Děkuji za pozornost

Jiří Palyza, palyza@ncbi.cz