BLACK CELL
Protecting critical infrastructures

Amikor az alert gondolkodik

Date
2026

Created by
Gyebnár Gergő

Magamról

Technológiai szemléletű vállalkozó és kutató vagyok

A Black Cell alapítójaként a gyakorlati és innovatív biztonsági megoldások fejlesztésére fókuszálok.

Jelenleg PhD-hallgató vagyok az NKE KMDI-n, ahol kutatásom fókusza a gépi tanulási módszerek alkalmazása a kritikus infrastruktúrák kibervédelmi eseménykezelésében.

2010 óta foglalkozom kiberbiztonsággal.

A compliance kifejezetten idegesít. 🙂

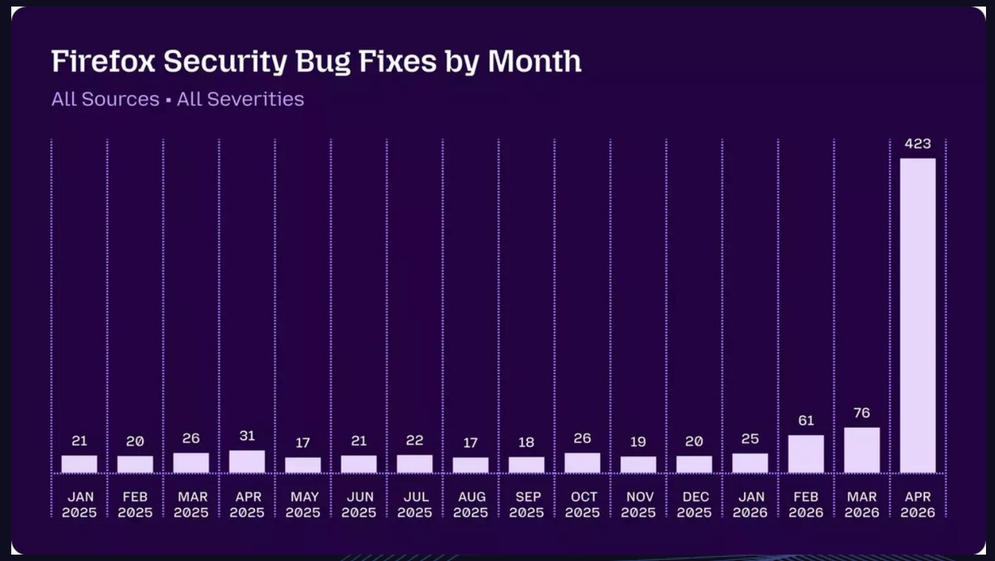
Az AI 💀

Warm up

Amivel szembe kell nézni

Project Glasswing

A Project Glasswing egy olyan kísérleti kezdeményezés, amely a kiberfenyegetések feltérképezésére és a sérülékenységek gyors elemzésére összpontosít.



Abliteráció

Az abliteráció egy olyan technika a mesterséges intelligencia világában, amellyel eltávolítják a modellek cenzúráját vagy biztonsági korlátait anélkül, hogy azokat újra kellene tanítani.
G0DM0D3; Nytheon AI (Tor); WormGPT

Data poisoning

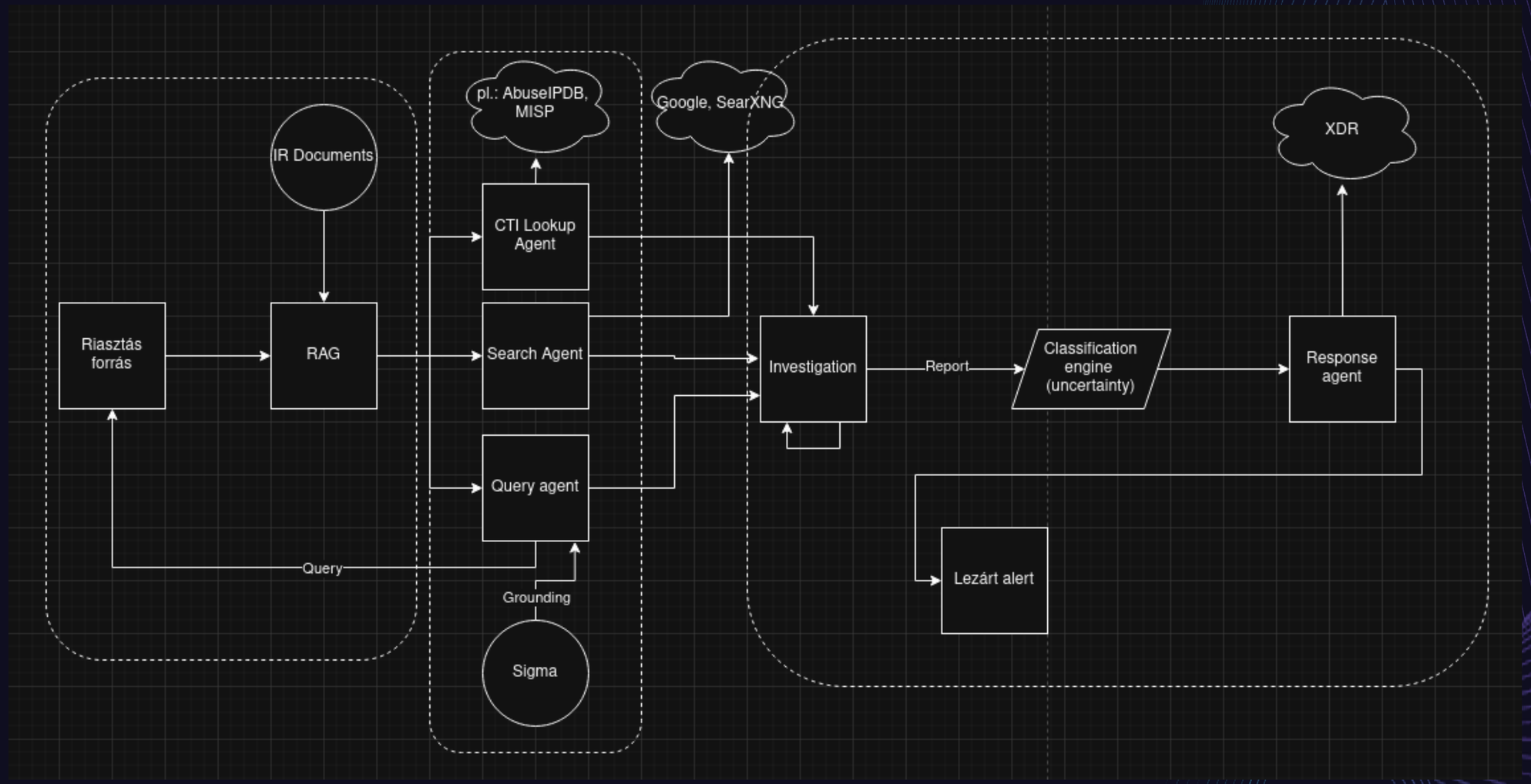
A tanítóadatokban lévő minimális manipuláció (akár 0,001%) elegendő backdoor beépítéséhez, amely a nyilvános forráskód-repo-kból (pl. ~2000 GitHub tároló) a nagy alapmodelleken, majd az abból származtatott újabb rendszereken keresztül terjed.

AI alapú malware-ek

Új generációs kártevőcsaládok, amelyek a végrehajtás során nagy nyelvi modelleket integrálnak. Ez a technika a „just-in-time" ön-módosítás, amelynek segítségével dinamikusan generálnak malicios szkripteket, obfuscálják saját kódjukat, és igény szerint hoznak létre funkciókat hardkódolt payload helyett.
(LameHug / PromptSteal, PromptFlux)

Strike back

A mi koncepciónk



Roadmap

Munkafolyamatok

2: Környezet
Os, DB, dependenciák és könytárak, mentés, orchestration

4: Foundation agent
Writer agent, szerepkörök meghatározása

6: Investigation agent
Elvégzi a kivizsgálást a kapott eredmények alapján

Step 7: Response agent
XDR, reporting, Lesson-what-we-learned








1: Research
Hardware, HLD, LLD, Modellek, Inference endpoint

Step 3: RAG
Index tervezés, dokumentumok újrarendezése relevancia szerint, scoring algoritmus finomhangolása, llm reranking

5: Query, CTI és search agent
Query agent prompt, investigation state tracking, eredmények kiértékelése, MISP, Kagi SearX, Google

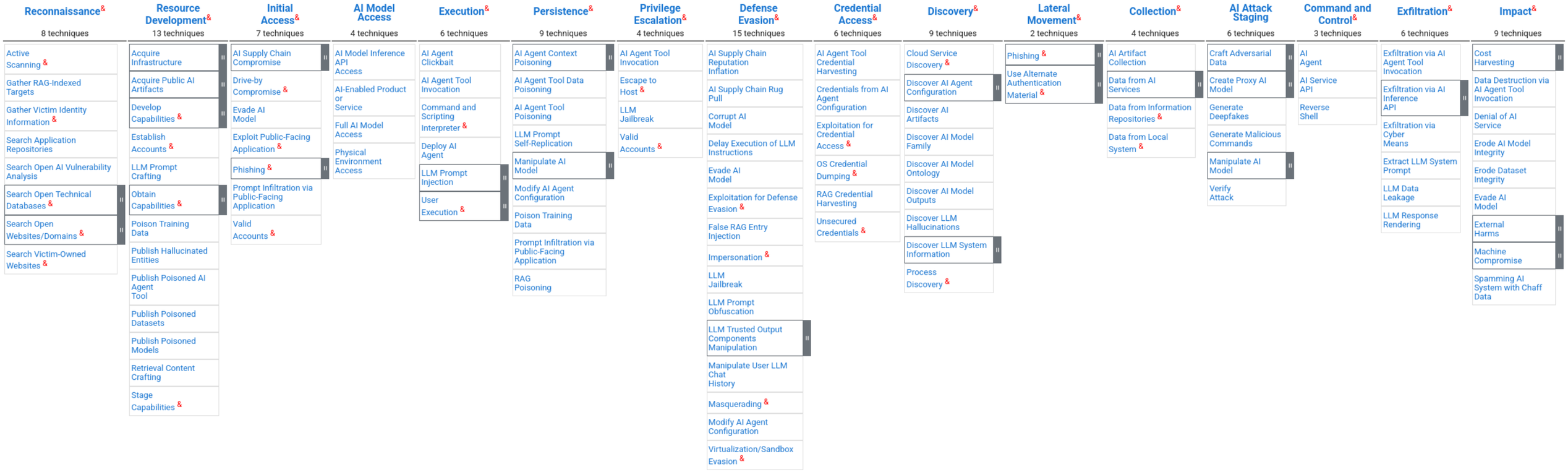
7: Classification engine
BERT

Dehogy van

Akkor kész?!

Strike back

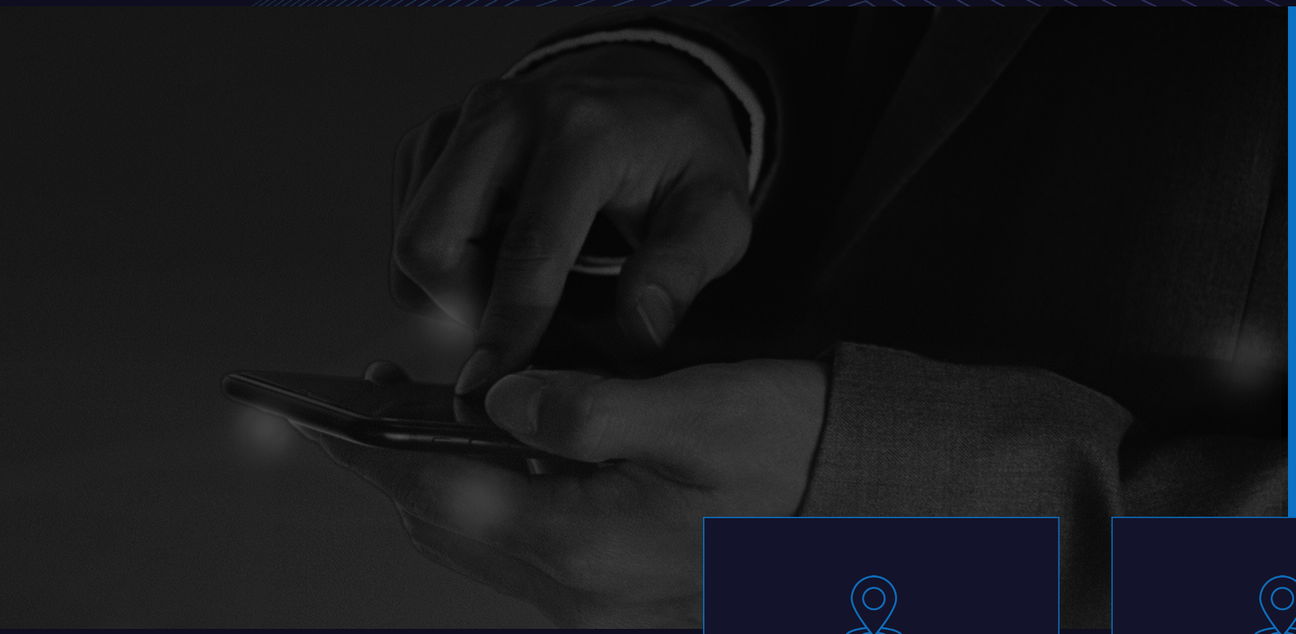
AI Security



Contact

Keep in touch

Visit our website, drop us a mail or contact your Account Executive!



Contact US

+36 1 605 0302

info@blackcell.io

www.blackcell.io

Frankfurt am Main
Germany

Budapest
Hungary

Dubai
UAE

BLACK CELL
Protecting critical infrastructures

Thank you!

www.blackcell.io