Kompetenčné a certifikačné centrum kybernetickej bezpečnosti

Od kybernetických rizík k umelej inteligencii

Bezpečnosť v meniacom sa digitálnom prostredí

Mgr. Marek Zeman, PhD.

Obrázky v prezentácii sú pod licenciou canva.com zakúpené na meno Marek Zeman

Hlas a tvár už nie je dôkaz

AI nemení pravidlá bezpečnosti – zrýchľuje útočníkov aj obrancov.

- Deepfake hovor „riaditeľa“ → urgentný prevod peňazí
- Známy hlas + autorita + časový tlak = úspešný podvod
- Mami, stal som sa súčasťou havárie, potrebujem peniaze
- Ani online meeting nás nezachráni....

- Ochrana: od perimetra k identitám, dátam a procesom



Základy bezpečnosti stále rozhodujú

Klasické kybernetické riziká

Incidenty začínajú pri základoch

Bez základnej hygieny vám ani AI nepomôže.

- Phishing a sociálne inžinierstvo
- Ransomware a vydieranie únikom dát
- Slabé heslá, chýbajúce MFA
- Nezáplatované systémy
- Útoky cez dodávateľov



Aj toto sú základy bezpečnosti

Aktualizácie:

- zapnúť automatické aktualizácie systému aj aplikácií na mobile, počítači a routeri

Dvojfaktorové overenie (2FA):

- hlavne na e-mail, internet banking, sociálne siete a Apple/Google účet

Iné heslo všade: ideálne so správcom hesiel,

Zámok obrazovky: PIN, odtlačok prsta alebo tvár na mobile, heslo na počítači

Aplikácie len z oficiálnych obchodov:

- Google Play a App Store, nie z odkazov v SMS

Overenie cez iný kanál:

- zavolať späť na ina získané číslo

Zálohy: fotky a dokumenty do cloudu alebo na externý disk

Aj toto sú základy bezpečnosti

Zabezpečený domáci Wi-Fi router:

- zmeniť predvolené heslo, používať WPA2/WPA3 a aktualizovať firmvér

Opatrne s verejnou Wi-Fi:

- na kaviarenskej sieti sa nepripájať do banky, radšej použiť mobilné dáta

Kontrola oprávnení aplikácií: prečo potrebuje baterka prístup ku kontaktom a polohe?

Zapnutá ochrana: Windows Defender (stačí), na mobile Google Play Protect

Menej osobných údajov na sociálnych sieťach:

- dátum narodenia, dovolenka či hlasové nahrávky pomáhajú podvodníkom

Rodinné kódové slovo: proti deepfake hovorom „vnúčaťa v núdzi“

Notifikácie z banky: SMS alebo push pri každej platbe, aby si podvod všimli hneď

Vedieť, čo robiť pri probléme: zavolať banke a zablokovať kartu, zmeniť heslá,

AI mení hru

Útočník, terč aj obranca

AI v rukách útočníka

Útočníci už AI používajú – každý deň.

- Bezchybný phishing – personalizovaný, po slovensky
- Deepfake hlas a video – CEO fraud, falošné pohovory
- Rýchlejšie hľadanie zraniteľností
- OSINT

- Čo vidíme: Nižšia vstupná bariéra pre menej zručných



AI ako nový terč útoku

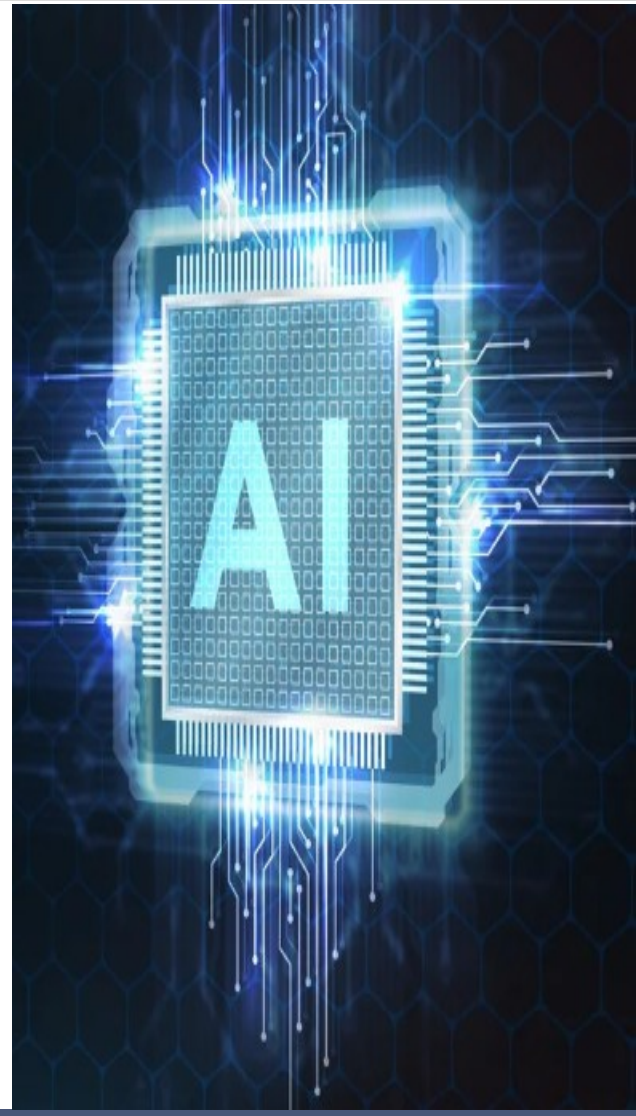
Každý AI systém je zároveň aj nová útočná plocha.

- Prompt injection – skryté pokyny v dokumentoch, e-mailoch, weboch
- Shadow AI – citlivé dáta vo verejných nástrojoch
- AI agenti s oprávneniami – chyba AI = incident
- Otrávené dáta a halucinácie v kritických rozhodnutiach

AI v rukách obrancu

AI posilňuje ľudský úsudok, nenahrádza ho.

- Prompt injection: skryté pokyny v dokumentoch, e-mailoch, weboch
- Jailbreak: obídenie bezpečnostných pravidiel modelu
- Adversariálne vstupy: nepatrná zmena vstupu oklame model
- Únik trénovacích dát: model prezradí, na čom sa učil
- Shadow AI: citlivé dáta vo verejných nástrojoch
- AI agenti s oprávneniami: chyba AI = incident
- Otrávené dáta a halucinácie v kritických rozhodnutiach



Riadenie AI rizík

Od chaosu ku kontrole

Nemôžete riadiť, čo nevidíte

- Inventár:
  - oficiálne nástroje, AI v bežnom softvéri, shadow AI
- Klasifikácia podľa rizika:
  - Aké dáta do systému vstupujú?
  - Čo môže systém sám urobiť?
  - Čo sa stane, keď sa pomýli?
- Vlastník pre každý AI systém



Kontroly, ktoré fungujú

- Výstup AI vždy over: zdroje, fakty, citácie
- AI nedávaj prístup k e-mailu, disku a účtom, ak to nie je nutné
- Žiadne osobné údaje, známky ani citlivé dokumenty do verejnej AI
- Používaj školské alebo overené nástroje, nie prvý chatbot z internetu
- Pozor na cudzie súbory a weby: môžu obsahovať skryté pokyny
- Priznaj, kedy a ako si AI použil

Netreba vymýšľať koleso -

AI riziko je biznis riziko – patrí na stôl vedenia.

- NIST AI RMF
- ISO/IEC 42001 – systém manažérstva AI
- ISO/IEC 23894 – riadenie rizík AI
- Napojenie na existujúce riadenie rizík a ISMS
- NIS2
- AI Act



5 krokov od dnes

1. MFA všade a správa prístupov
2. Inventár AI a vlastník pre každý systém
3. Overenie neobvyklých požiadaviek druhým kanálom
4. Školenia na deepfake a AI phishing
5. Minimálne oprávnenia pre AI

Spolufinancovaný Európskou úniou

Financované Európskou úniou. Vyjadrené názory a postoje sú názormi a vyhláseniami autora(-ov) a nemusia nevyhnutne odrážať názory a stanoviská Európskej únie. Európska únia za ne nepreberá žiadnu zodpovednosť.

NCC-SK SLOVAKIA CYBERSECURITY COORDINATION CENTRE

ECCC EUROPEAN CYBERSECURITY COMPETENCE CENTRE

Kompetenčné a certifikačné centrum kybernetickej bezpečnosti

Spolufinancované Európskou úniou

Vyjadrené názory a postoje sú názormi a vyhláseniami autorov a nemusia nevyhnutne odrážať názory a stanoviská Európskej únie. Európska únia za nich nepreberajú žiadnu zodpovednosť.

Ochrana vlastníckych práv
Všetky autorské práva k tomuto dokumentu sú vyhradené pre Kompetenčné a certifikačné centrum kybernetickej bezpečnosti (ďalej len „KCCKB“). Autorské práva k tomuto dokumentu má a autorské práva k tomuto dokumentu vykonáva KCCKB.

Vlastnícke práva tretích strán
Všetky ochranné známky a iné obdobné právne chránené označenia spomenuté v tomto dokumente sú výhradným vlastníctvom ich vlastníkov, ktorí sú, pokiaľ sa nejedná o KCCKB,
v dokumente označení vo forme príslušnej citácie.

Akékoľvek kopírovanie či iné neoprávnené použitie celého dokumentu alebo jeho časti,
v elektronickej alebo listinnej podobe, bez predchádzajúceho písomného súhlasu jeho autorov, napr. KCCKB, je zakázané. Porušenie vlastníckych a iných súvisiacich práv bude riešené v zmysle právnych predpisov Slovenskej republiky.

Limity informácií
Informácie obsiahnuté v tomto dokumente sú poskytované iba na informačné účely a bez akejkoľvek záruky, výslovnej či implicitnej. Názov KCCKB, logo KCCKB a ďalšie produkty a služby KCCKB sú ochrannými známkami KCCKB. Ostatné názvy spoločností, produktov alebo služieb môžu byť ochrannými známkami, alebo značkami služieb iných organizácií.

www.cybercompetence.sk,

kyberkomunita.sk

www.linkedin.com/company/