Tomas Bata University in Zlín
Faculty of Applied Informatics

Aktuální výzvy v oblastech AI a KB

Výzkumné projekty související s aktuálními výzvami v oblastech AI a KB

Ladislav Dorotík | l_dorotik@utb.cz
David Malaník | dmalanik@utb.cz

Tomas Bata University in Zlín
Faculty of Applied Infromatics
27. 03. 2024

PT LAB

https://ptlab.fai.utb.cz/
https://fai.utb.cz/

AILAB a PTLAB

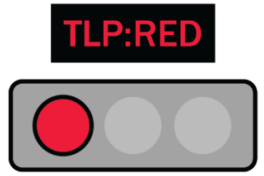
- Penetrační Laboratoř PT LAB
  - Penetrační testování,
  - Školení v oblasti cybersec,
  - Bezpečnostní výzkum,
  - Digital Forensics – zajištění a analýza stop
- Laboratoř umělé inteligence AI LAB
  - Práce s drony – detekce na základě termovize
  - Generativní AI – implementace a fine tunning chat botů
- Spolupráce např. na projektech EU Horizon

Digital Forensics: Proč se zajímat - síla AI?

- https://www.youtube.com/watch?v=F4WZ_k0vUDM&t=5s

Digital Forensics a AI

[2]



Generativní AI v KB – Clipboard Stealer

- Útoky na Clipboard:

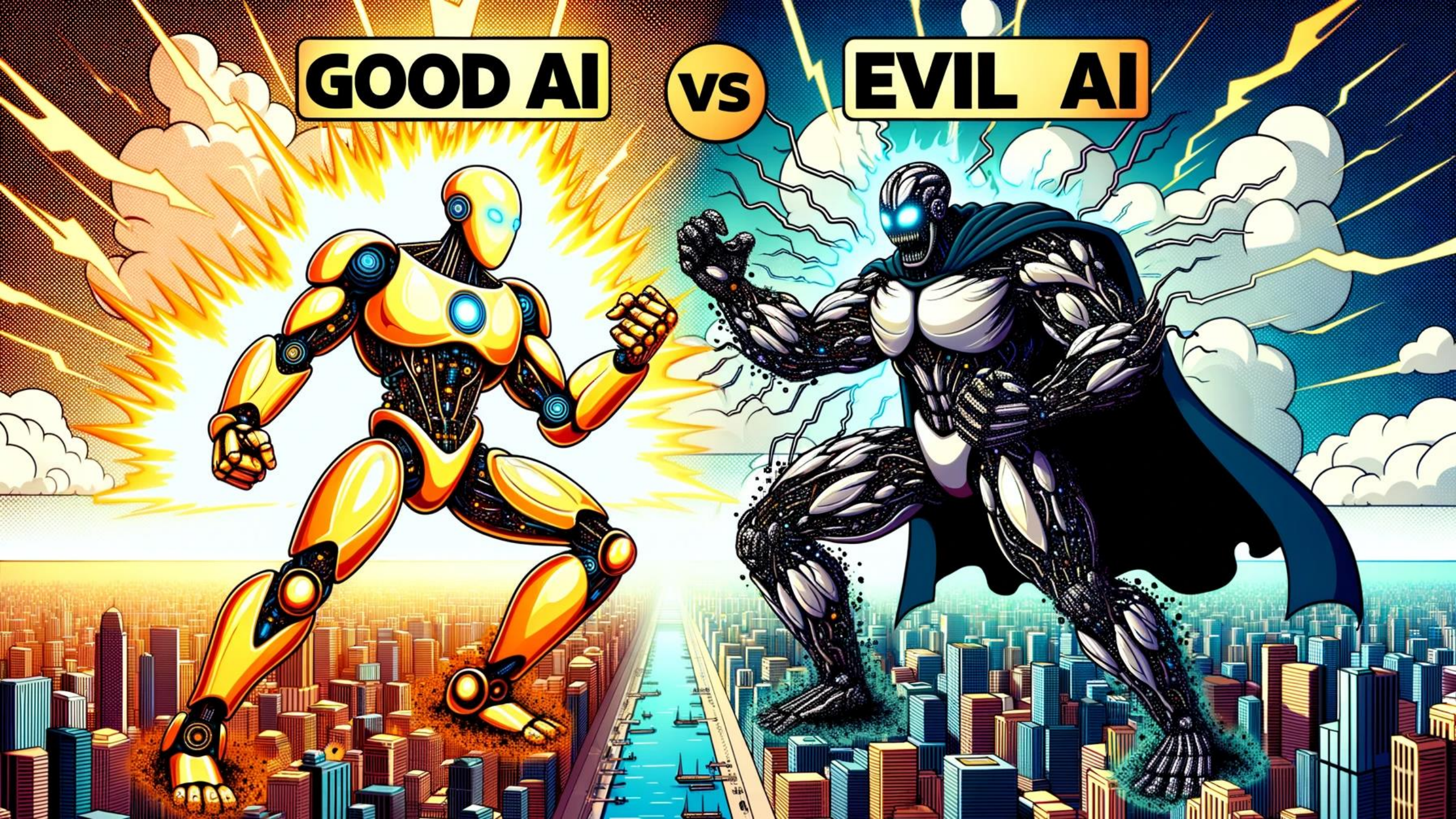
Generativní AI v KB

- AI Chat bot & Malware
  - Zrychluje vývoj malwaru
  - Snižuje náklady
  - Může fungovat jako operátor v hovoru
- Temný Chat bot
  - Stejně jako firmy i útočníci využívají ChatBoty
  - Verze bez restrikcí

Devin [https://www.youtube.com/watch?v=fjHtjT7GO1c]

Devin – Kritický pohled

- Enormní snížení nákladů
- Nižší kontrola nad procesem
- Pakliže zvládne v rámci etického hackingu penetrační test, zvládne i …?
  - Prohlížeč
  - IDE
  - CLI


GOOD AI
VS
EVIL AI

AI v KB

- Enormní nárust malwaru ve všech odvětví.
- Nemožné analyzovat vše manuálně a následně detekovat dle signatur
  - Variace APK – 1 Malware -> 10-100 aplikací.
- Nasazení AI do procesu detekce hrozeb, malwaru, zranitelností, zabezpečení.

AI vs AI ve službách KB

- Téměř žádný rozdíl …

- Odborník na rajčata vs. Odborník na síťový provoz
- Odborníci z AI

- Data jsou základ

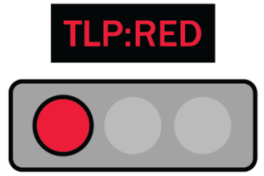
AI v KB - DATA

- Kvalita umělé inteligence je přímo závislá na **kvalitě datové sady**.
- Faktory ovlivňující kvalitu dat v KB:
  - **Stáří** -> OS, Obecně používané technologie
  - Platforma -> Windows, Linux, Android
    - Obtížně kombinovatelné s např. Android platformou
  - **Vývoj** -> Měnící se povaha systémů (API Levely), DS nutné aktualizovat
  - Zpožděné zachycení
    - Malware (starší) vs Benigní (novější) -> zero day problémy
  - Zkreslení výsledků -> false positive/negative
  - **Reproducibilita -> Ano, kritické pro digital forensic**

Digital Forensics a AI

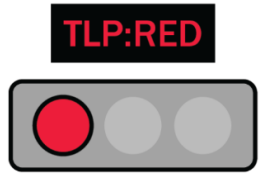
- MobilEdit (Image and Video Processing):
  - detekce zbraní, drog, pornografie

- Obecně velký prostor pro výzkum
  - Detekce evidencí (AI se v pátek večer netěší domů),
  - Automatická analýza logů,
  - NLP – Kontextuální analýza textového obsahu (do jisté míry nasazeno -> Keywords)
  - Forensic Triage (preprocessing dat – zrychlení prvotní analýzy)
  - Devin [3]-> Automatizované AI penetrační testy?



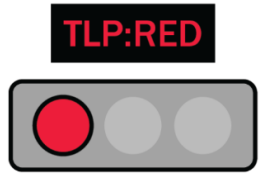
Výzkumné možnosti GRAY ZONE

- Výhody DarkAI nemusí nutně využívat jen „špatná strana“
- ….



Výzkumné možnosti GRAY ZONE

- A co jít trochu dále



Výzkumné možnosti GRAY ZONE

- **Nástroje „špatné“ strany….**

Výzkumné možnosti

- Preprocessing dat pro využití v AI
- Faktory ovlivňující kvalitu výstupů
- Pravdivost výstupů – má to AI fakt dobře?
- Etické faktory AI
- Autonomní AI – kde je hranice našeho vlivu?

Tomas Bata University in Zlín
Faculty of Applied Informatics

Děkuji za pozornost.

Aktuální výzvy v oblastech AI a KB

Výzkumné projekty související s aktuálními výzvami v oblastech AI a KB

Ladislav Dorotík | l_dorotik@utb.cz
David Malaník | dmalanik@utb.cz

Tomas Bata University in Zlín
Faculty of Applied Infromatics
27. 03. 2024

PT LAB

https://ptlab.fai.utb.cz/
https://fai.utb.cz/

Reference

[1] Deutsche Telekom. (2023, July 3). *Nachricht von Ella* | *Without Consent* [Video file]. YouTube. https://www.youtube.com/watch?v=F4WZ_k0vUDM&t=5s

[2] Geng, J., Huang, D., & Torre, F.D. (2022). *DensePose From WiFi*. ArXiv, abs/2301.00250. doi: 10.48550/arXiv.2301.00250

[3] Cognition. (2024, March 12). *Introducing Devin, the first AI software engineer* [Video file]. YouTube. https://www.youtube.com/watch?v=fjHtjT7GO1c

[4] OpenAI. (2024). ChatGPT, version GPT-4, as of April 2023. (Mar 14 version) [Large language model] https://chat.openai.com/chat