



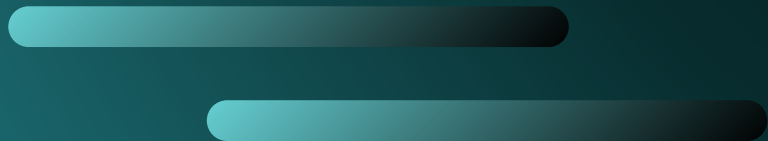
ANALÝZA BEZPEČNOSTNÝCH ÚDAJŮ Z HONEYPOTŮ POMOCÍ VELKÝCH JAZYKOVÝCH MODELOV

Vedúci práce: doc. RNDr. JUDr. Pavol Sokol, PhD. et PhD.

Sophia Petra Krišáková

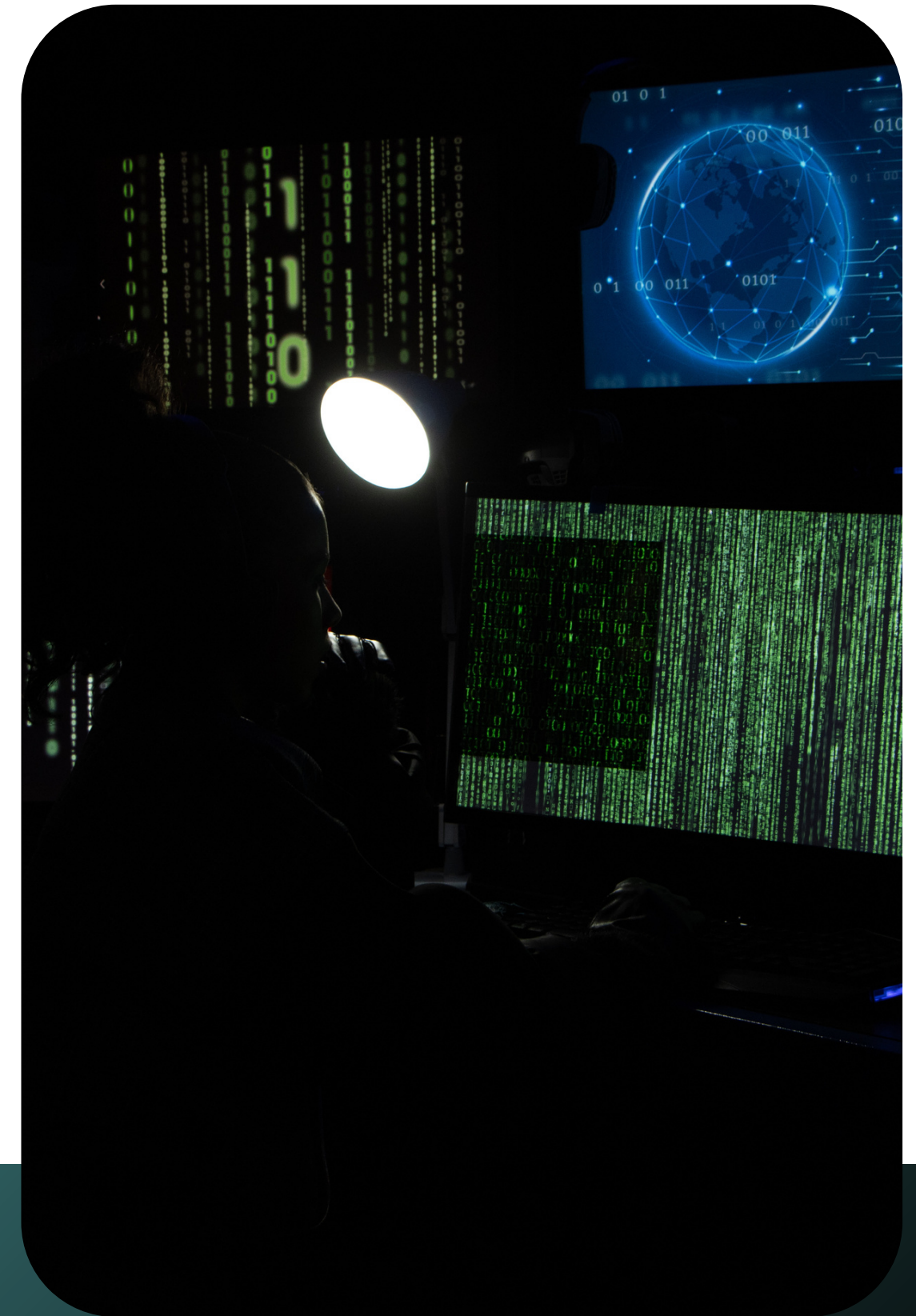
Ústav informatiky, Prírodovedecká fakulta, UPJŠ

15.05.2025



Motivácia

- LLM + Cybersecurity



Honeypots

- lákadlo pre útočníkov, simulovaný systém
- zhromažďovanie informácií o metódach útoku
- identifikácia zraniteľností
- maskovaný ako zraniteľný cieľ



LLM

- veľký jazykový model
- neurónové siete - veľa parametrov
- tréning
 - tokenizácia
 - vektorizácia
 - transformácia
 - dekódovanie,
- predikcia a generovanie textu
- typy
 - offline
 - online (problém pri bezpečnostných dátach)

Prompt engineering

- proces navrhovania, optimalizácie a testovania promptov pre generatívne modely, ako sú LLM, aby sme dosiahli požadované výsledky
- Cieľ: Zlepšiť kvalitu a relevantnosť odpovedí generovaných modelom
- Príklad:
 - Zlá otázka: 'Povedz mi niečo o histórii.'
 - Dobrá otázka: 'Aké dôležité udalosti sa stali počas 2. svetovej vojny v Európe?'

Návrh postupu riešenia

**Analýza
podobných
prác**

**Predspracovanie
a analýza dát**

**Možnosť
použitia iných
dát, alebo
požiadaviek**

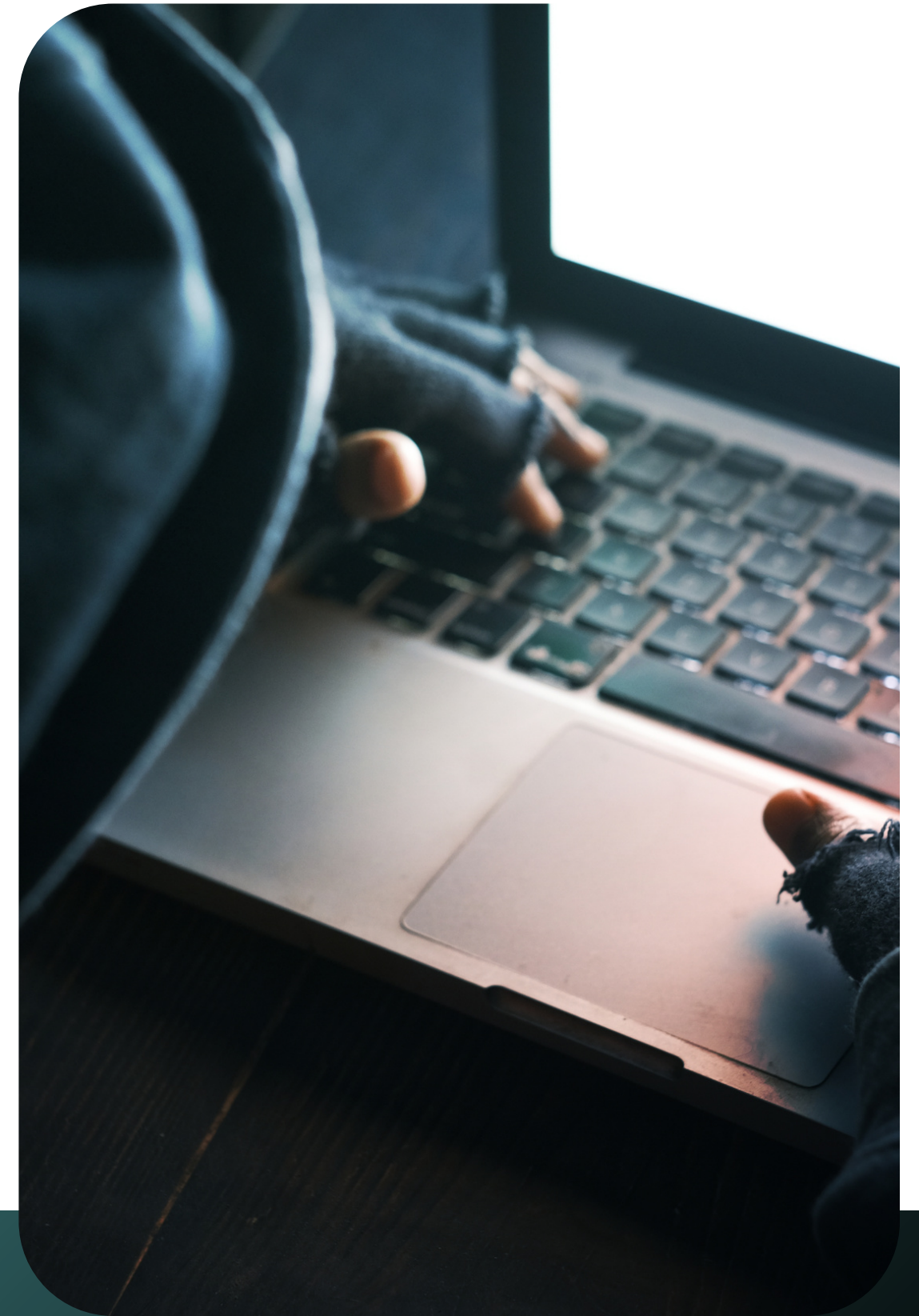
**Analýza
možných
prípadov**

Analýza podobných prác

- Vplyv ChatGPT na digitálnu forenznú analýzu,
- experimenty: porozumenie artefaktov, vyhľadávanie dôkazov, generovanie kódu, detekcia anomálii, riešenie incidentov,
- vplyv veľkých jazykových modelov na kybernetickú bezpečnosť,
- riziká LLM v kybernetickej bezpečnosti a ich zmiernenie
- predpovede vývoja LLM a regulačné možnosti
- techniky na bezpečný vývoj a nasadenie LLM,
- bezpečnosť a súkromie pri používaní LLM,
- návrh bezpečnej integrácie LLM a zhrnutie ich dvojznačnosti v kybernetickej bezpečnosti,
- nový prístup k honeypot systémom, ktoré využívajú LLM na interakciu s útočníkmi,
- zlepšiť detekciu hrozieb a analýzu útokov prostredníctvom pokročilej interakcie, ktorá napodobňuje ľudské správanie.

Use case

Predikcia správania útočníka



- <https://haas.nic.cz/stats/>



	time	time_closed	ip	country	login_successful	commands
0	2024-09-01T00:00:07.720868+00:00	2024-09-01T00:00:08.726674+00:00	112.72.235.19	KR	False	[]
1	2024-09-01T00:00:07.785101+00:00	2024-09-01T00:00:08.789687+00:00	112.72.235.19	KR	False	[]
2	2024-09-01T00:00:07.808650+00:00	2024-09-01T00:00:08.813711+00:00	112.72.235.19	KR	False	[]
3	2024-09-01T00:00:07.865543+00:00	2024-09-01T00:00:08.870439+00:00	112.72.235.19	KR	False	[]
4	2024-09-01T00:00:07.899126+00:00	2024-09-01T00:00:08.903224+00:00	112.72.235.19	KR	False	[]
...	...	...	...	...	...	...
22026	2024-09-01T23:59:42.472663+00:00	2024-09-01T23:59:42.981977+00:00	122.46.16.230	KR	True	[ps grep '[Mm]iner']
22027	2024-09-01T23:59:44.649059+00:00	2024-09-01T23:59:45.102192+00:00	122.46.16.230	KR	True	[ps -ef grep '[Mm]iner']
22028	2024-09-01T23:59:46.752907+00:00	2024-09-01T23:59:47.258398+00:00	122.46.16.230	KR	True	[ls -la /dev/ttyGSM* /dev/ttyUSB-mod* /var/spo...
22029	2024-09-01T23:59:49.092738+00:00	2024-09-01T23:59:49.586087+00:00	122.46.16.230	KR	True	[echo Hi cat -n]
22030	2024-09-01T23:59:50.550074+00:00	2024-09-02T00:00:01.413132+00:00	180.117.252.74	CN	True	[/ip cloud print]

22031 rows × 6 columns

Predspracovanie a analýza dát

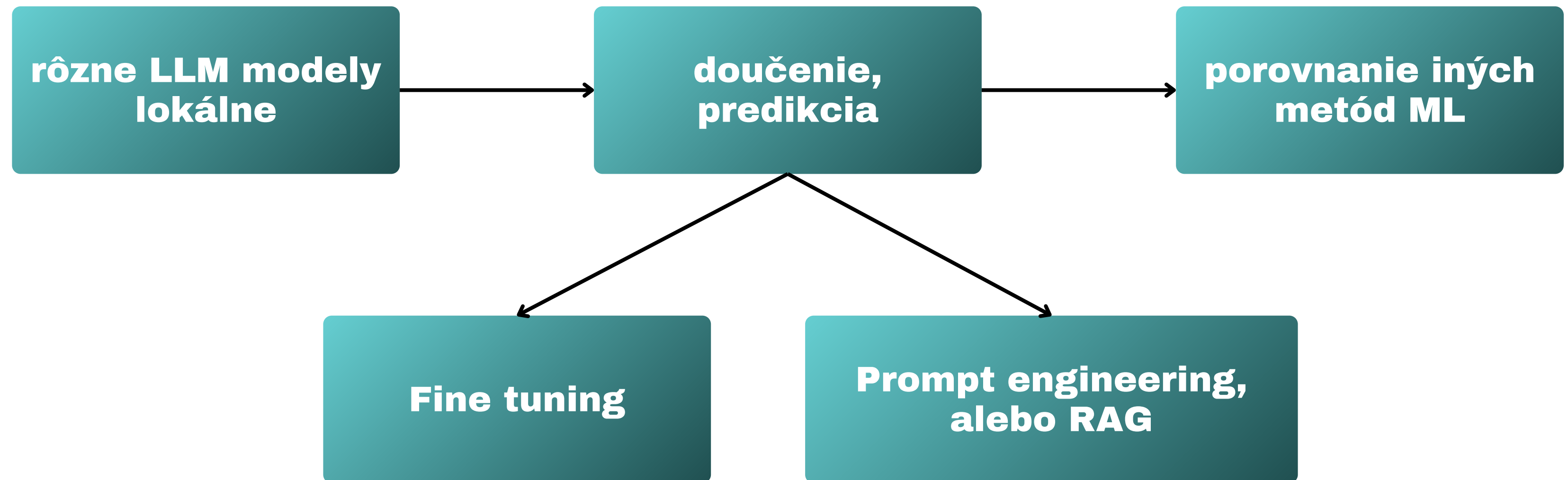
**veľké množstvo
JSON súborov**

zlúčenie

predspracovanie

☐	 2017-01-27.json	pred 7 rokmi	145 B
☐	 2017-01-28.json	pred 7 rokmi	8.72 kB
☐	 2017-01-29.json	pred 7 rokmi	5.97 kB
☐	 2017-01-30.json	pred 7 rokmi	1.33 kB
☐	 2017-01-31.json	pred 7 rokmi	263 B
☐	 2018-01-01.json	pred 7 rokmi	43.7 kB
☐	 2018-01-02.json	pred 7 rokmi	43.6 kB
☐	 2018-01-03.json	pred 7 rokmi	49.9 kB
☐	 2018-01-04.json	pred 7 rokmi	18.9 kB
☐	 2018-01-05.json	pred 7 rokmi	11.9 kB
☐	 2018-01-06.json	pred 7 rokmi	10.6 kB
☐	 2018-01-07.json	pred 7 rokmi	5.52 kB
☐	 2018-01-08.json	pred 7 rokmi	11.6 kB
☐	 2018-01-09.json	pred 7 rokmi	53.6 kB
☐	 2018-01-10.json	pred 7 rokmi	39.9 kB
☐	 2018-01-11.json	pred 7 rokmi	52.9 kB
☐	 2018-01-12.json	pred 7 rokmi	54.8 kB

Postup



Ďakujem za pozornosť