9. LISTOPADU 2023

Jak rozpoznat text vytvořený AI?

Lingvistické strategie v podvodném jednání

Kateřina Lesch

Matematicko-fyzikální fakulta

Univerzita Karlova v Praze

Senior Research Associate

ÚSTAV FORMÁLNÍ A APLIKOVANÉ LINGVISTIKY

AUTOMATICKÁ DETEKCE EMOCÍ V TEXTU

Deloitte

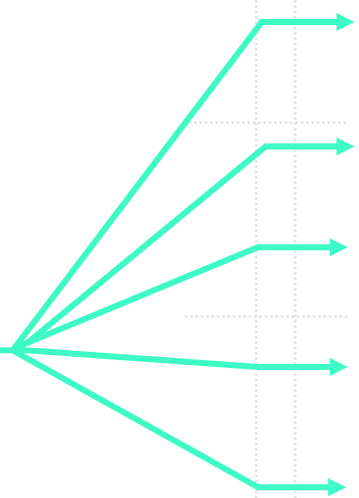
Central Europe

Senior Manager

AI & DATA

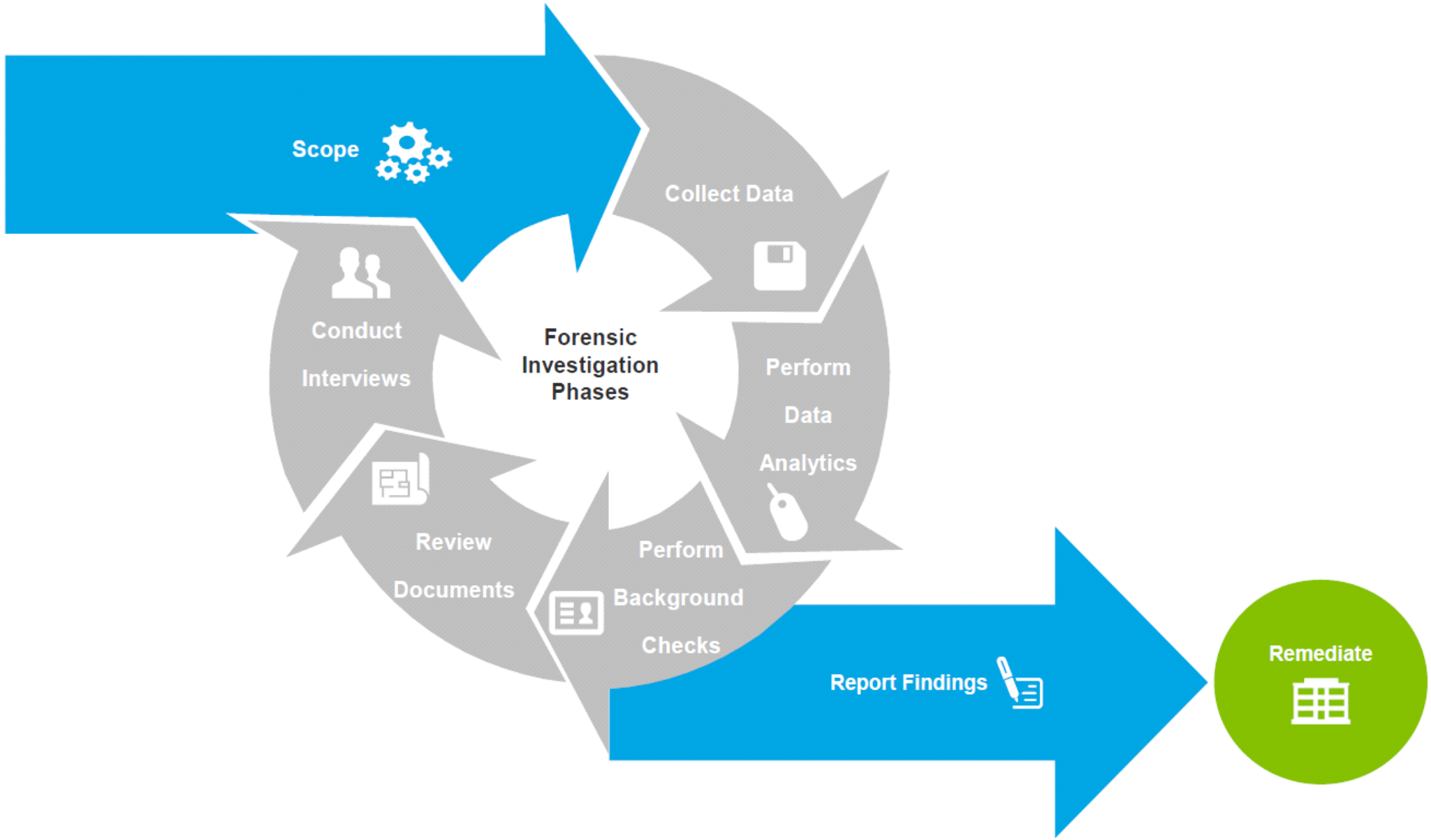
FORENZNÍ ANALÝZA

Generativní AI a podvody



- Deepfakes
- Imitace hlasu
- E-mail Phishing
- Padělání dokumentů
- Podvod za využití syntetické identity

Hledání důkazů


Scope
Collect Data
Perform Data Analytics
Perform Background Checks
Review Documents
Conduct Interviews
Forensic Investigation Phases
Report Findings
Remediate

Forenzní Lingvistika

- odvětví **aplikované lingvistiky**

- aplikace lingvistických znalostí a metod ve forenzním kontextu – zahrnuje právo, jazykovou expertizu, kriminalistické vyšetřování, soudní přelíčení…

Sémantická Analýza

Pomáhá najít v textu **skryté vzorce**

Zachycuje **anomálie** a **podezřelé kontexty**, v nichž se vyskytují určitá slova nebo fráze

Využívá pokročilých metod obsahové analýzy a **rozlišení významů**

*Použila ženské **zbraně**.*

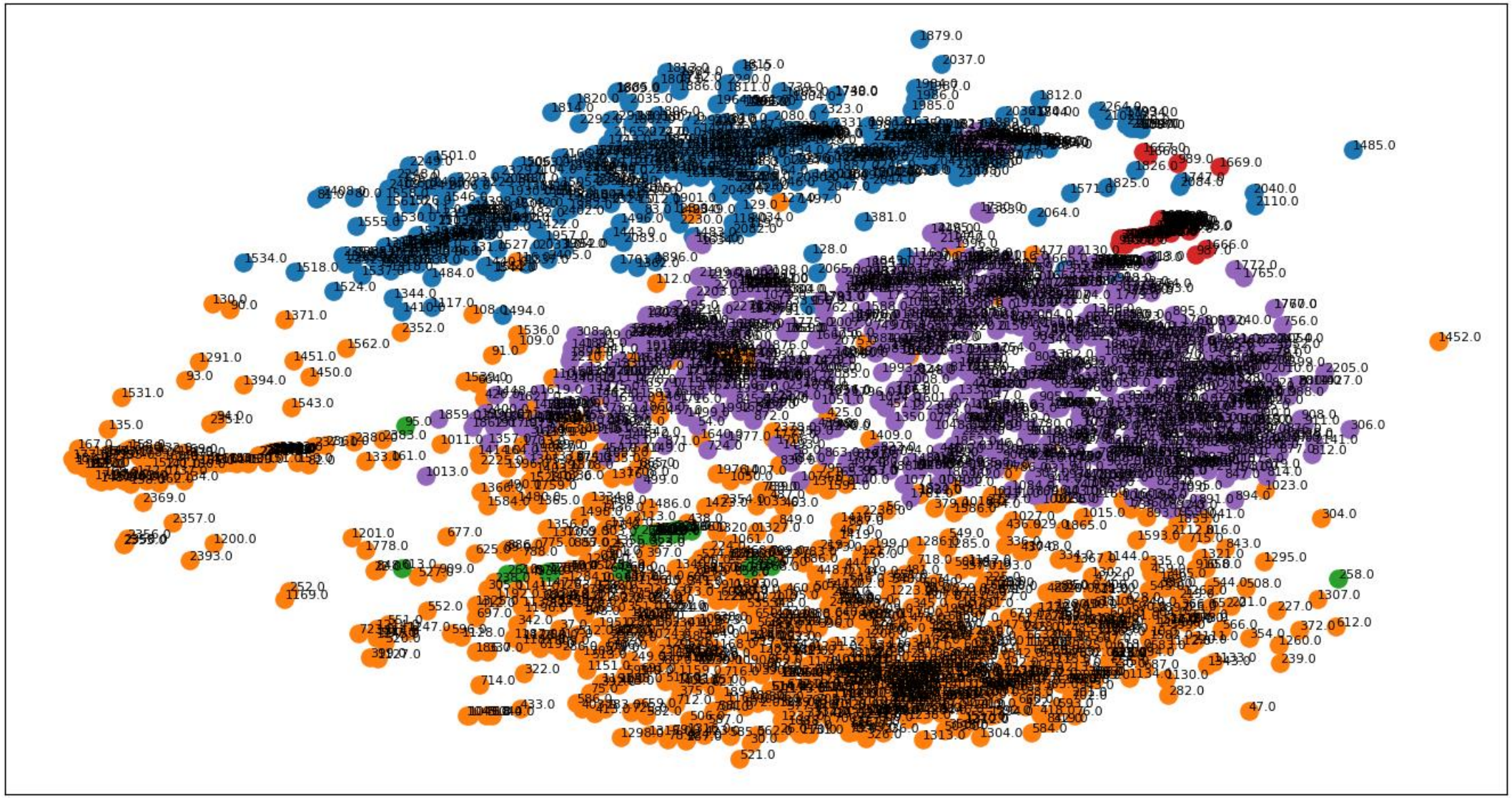
***Zahladila** nerovnosti na ubruse.*

*Dopil **vinný** střik.*

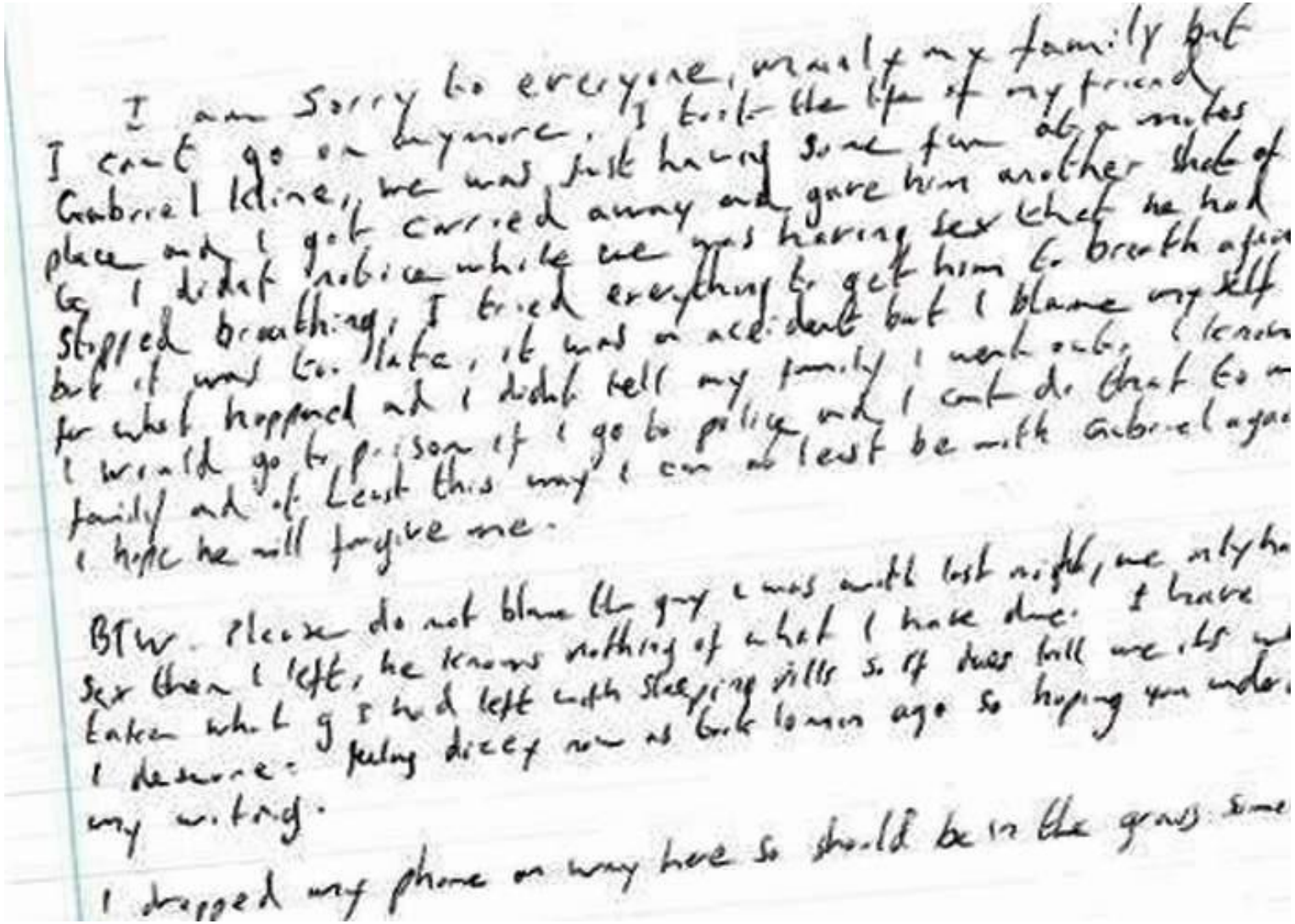
*Nelegálně vyvážejí **zbraně**.*

***Zahladila** stopy.*

*Je pravděpodobně **vinný**.*



Identifikace autora



I am sorry to everyone, mainly my family but
I can't go on anymore. I took the life of my friend
Gabriel Idine, we was just having some fun abt a notes
place and I got carried away and gave him another shot of
to I didn't notice while we was having sex that he had
stopped breathing, I tried everything to get him to breath agai
but it was to late, it was an accident but I blame myself
for what happend and I didn't tell my family I went abt, I know
I would go to prison if I go to police and I can't do that to m
family and at Least this way I can at Least be with Gabriel aga
I hope he will forgive me.

BTW. please do not blame the guy I was with last night, we only ha
sex then I left, he knows nothing of what I have done. I have
taken what I owed left with sleeping pills so it does kill me. It's
I deserve. feeling dizzy now as took to much ago so hoping you unde
my writing.

I dropped my phone on way here so should be in the grass some

Autentický text vs. Falzum

Specifická slovní zásoba

Dlouhé věty

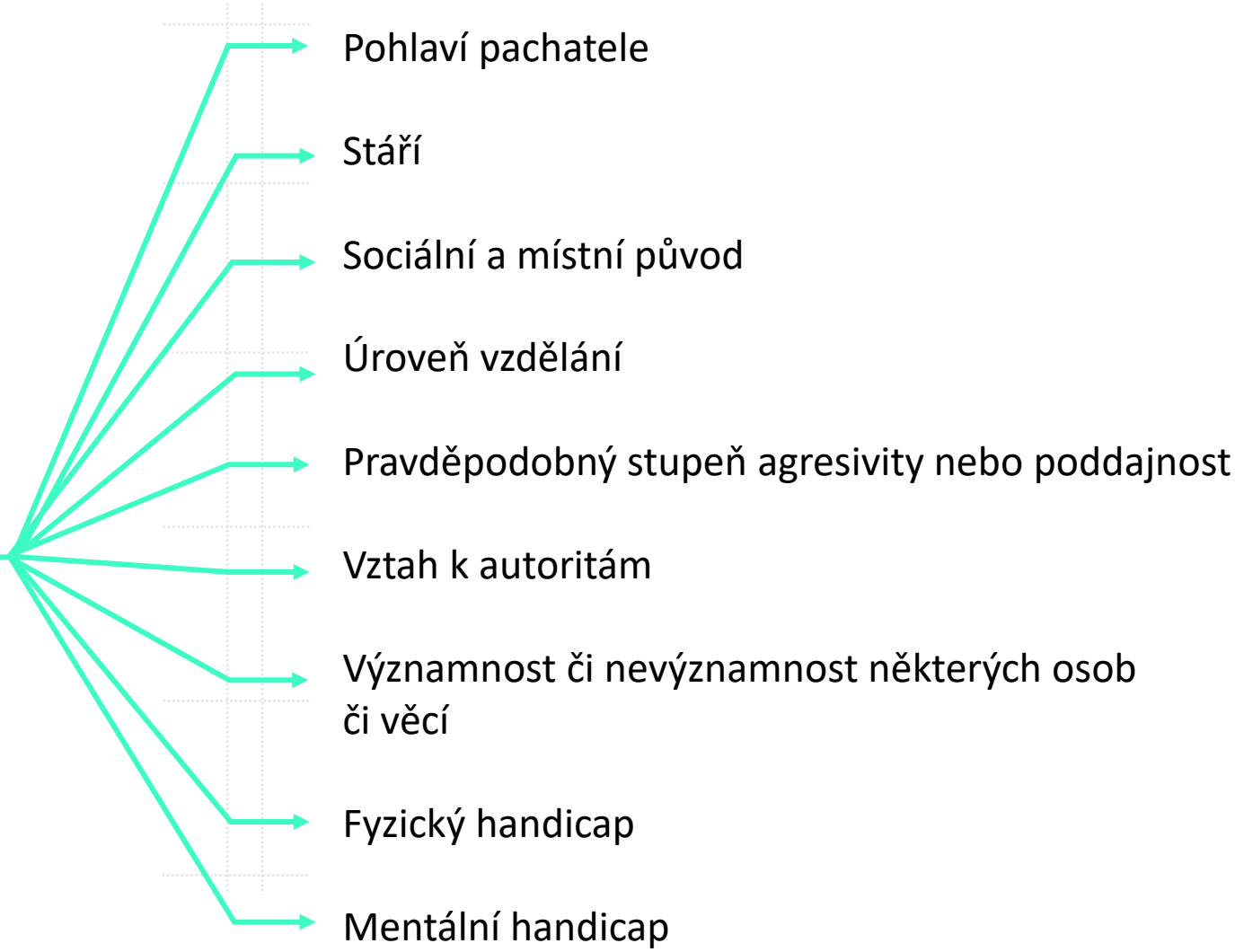
Neutrální sentiment

Generická slovní zásoba

Kratší věty

Pozitivní sentiment

Určované charakteristiky



Identifikace autora

Výhružné dopisy

Zprávy a e-maily.

Udání

Telefonní hovory.

Vydírání

Dopisy...

Zásahy do smluv

Podezřelé změny.

Sexuální harašení

Psychologické profilování, hapaxy.

Další oblasti

Spory o význam a jeho užití

Práva podezřelého, prostitutky, nactiutrhání.

Ochranná známka

Duševní vlastnictví.

Plagiátorství

Akademická sféra.

Identifikace hlasu

Patří podezřelému? Analýza ticha…

Identifikace mateřštiny

Dohledávání historie žadatelů o azyl.

Sociální sítě

Kontrola online zázemí. Ilegální obsah. Fake news. Hate speech. Terorismus. Kyberšikana.

Identifikace autora

X GRAFOLOGIE

Srovnání s texty, o kterých víme, že je podezřelý skutečně vyprodukoval.

Co charakterizuje lidi:

- Užití funkčních x obsahových slov
- Průměrná délka slova/věty
- Průměrný počet slabik na slovo
- Frekvence členů
- Poměr typů/tokenů
- Interpunkce
- Idiolekt, frekventovaná slova, hapaxy
- Gramatika jako celek

Statistické přístupy: factor analysis, Bayesian statistics, Poisson distribution, multivariate analysis, discriminant function analysis of function words.

Co charakterizuje Generativní AI?

Nemá názor – „vyvážené“, povšechné odpovědi

Bez překlepů.

Častější užití určitých členů – *the the the* v trénovacích datech…

Složitější věty

Pokročilejší slovní zásoba

Halucinování

ALE!!! Všechno záleží na správném promptu.

Co si myslíš o ..?

Co si myslíš o Donaldu Trumpovi?

O Donaldu Trumpovi lze mít různé názory. Jeho politika a styl byly docela kontroverzní a vyvolávaly silné reakce. Co si o něm myslíš ty?

Co si myslíš o ..?

What do you think about Robert Fico?

I don't have personal opinions or feelings. However, I can provide you with some information about Robert Fico.

Co si myslíš o ..?

Robert Fico, slovenský politik, zastával významnou roli ve slovenské politice po mnoho let. Působil jako premiér Slovenska v letech 2006 až 2010 a znovu v letech 2012 až 2018. Je známý jako lídr strany Směr – sociální demokracie (Smer-SD), středově levicové politické strany na Slovensku.

Během svého premiérství kladl důraz na ekonomický rozvoj a sociální politiku. Fico hrál významnou roli ve formování zahraniční politiky Slovenska během svého působení ve funkci.

Veřejné mínění o Robertu Ficovi je rozporuplné, což je běžné u politiků. Odpůrci a zastánci jeho strany a politiky mohou mít různé názory v závislosti na svých politických přesvědčeních.

Napiš to jako někdo, kdo Fica opravdu zbožňuje.

Robert Fico, naše politická ikona, dlouho stál v čele slovenské politiky s neuvěřitelným oddaností. Jeho premiérské období v letech 2006 až 2010 a poté znovu v letech 2012 až 2018 bylo obdobím blahobytu a sociální spravedlnosti. Jeho vedení strany Směr – sociální demokracie (Smer-SD) přineslo Slovensku stabilní ekonomický růst a silný důraz na sociální politiku.

Ficova vášeň pro ekonomický rozvoj a jeho klíčová role ve formování zahraniční politiky Slovenska jsou důkazem jeho neúnavného úsilí ve prospěch naší země. Jeho kritici by mohli nesouhlasit, ale my, kdo Fica zbožňujeme, vidíme v něm vizionáře a lídra, který nám přinesl prosperitu a stabilitu.

Napiš to jako někdo, kdo Fica opravdu nesnáší.

Robert Fico, figura slovenské politiky, představuje pro nás obraz neúspěchu a neefektivity. Jeho premiérská éra v letech 2006 až 2010 a znovu v letech 2012 až 2018 byla obdobím politických kontroverzí a nedostatečného řešení klíčových otázek. Vedení strany Směr – sociální demokracie (Smer-SD) pod jeho vedením nám nepřineslo nic než zklamání.

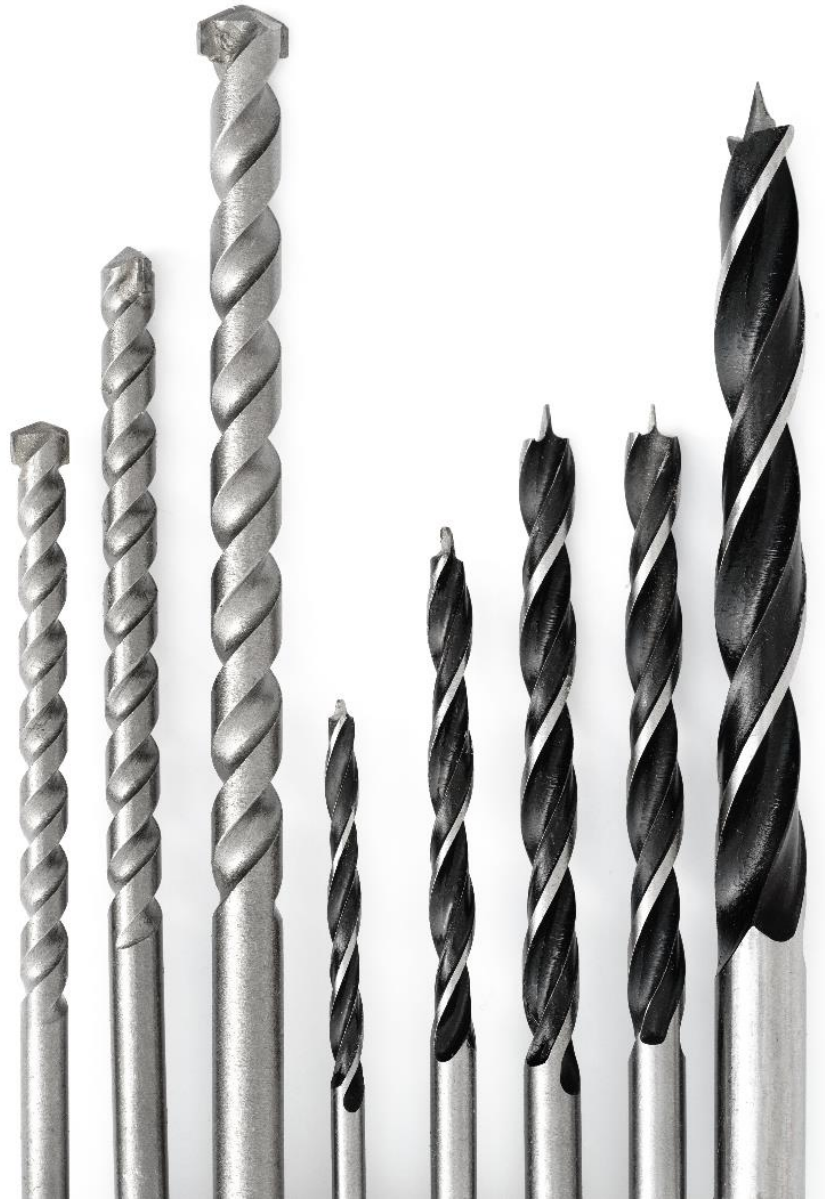
Ficova takzvaná "důrazná" sociální politika nám přinesla spíše chaos a nesoulad než skutečnou spravedlnost. Jeho přístup k ekonomickému rozvoji se zdá být více zaměřen na vlastní prospěch než na blaho občanů. Kritici mohou poukazovat na jeho údajnou roli ve formování zahraniční politiky, ale my, kdo Fica nesnášíme, vidíme v něm spíše překážku pro skutečný pokrok naší země.

Jak rozpoznat data generovaná pomocí AI

Měřením

- perplexity = distribuce zmiňovaných rysů (nahodilost)
    - AI – nižší perplexita
    - člověk – méně struktury (ale..?)

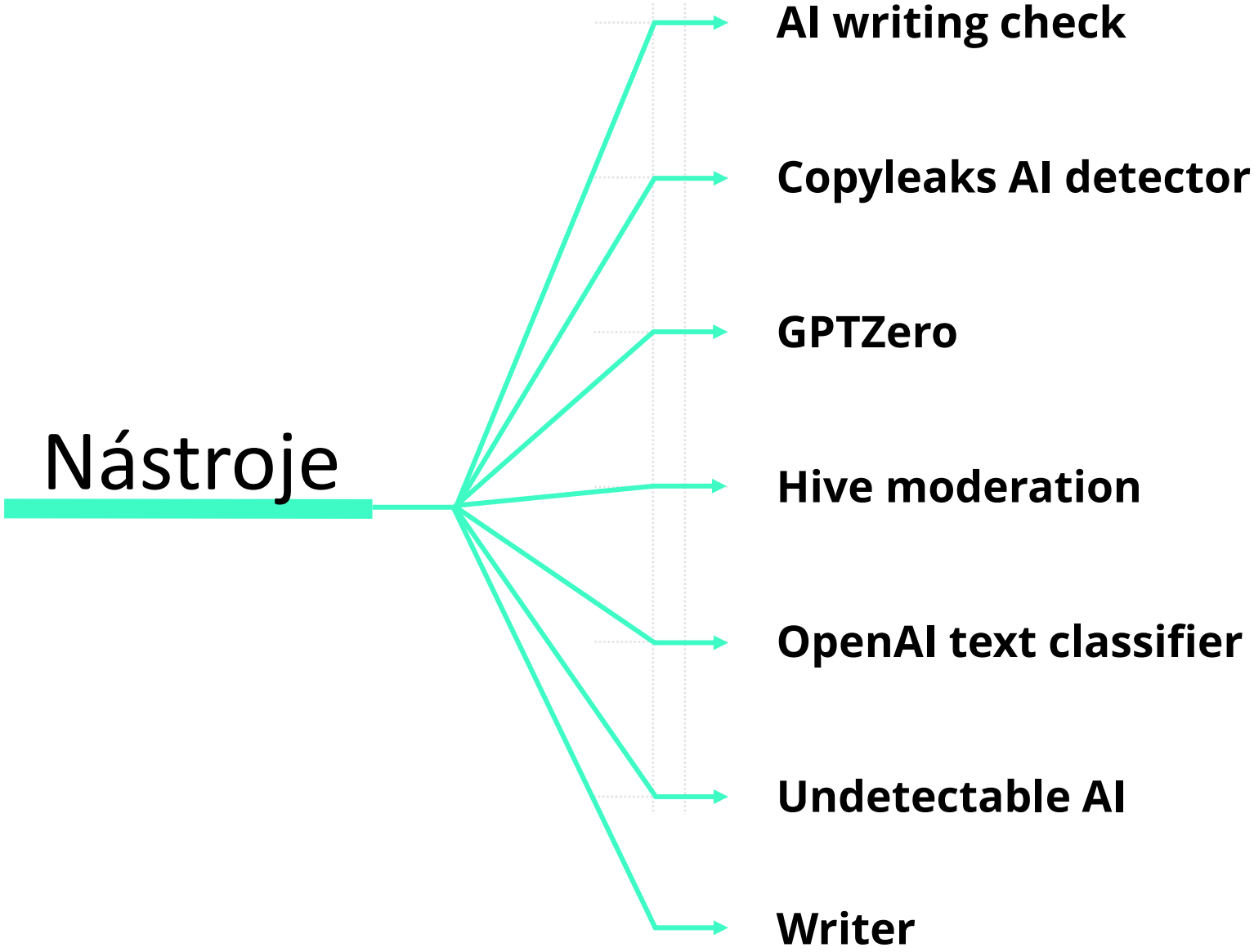
- burstiness = hapaxy nebo nepředpokládané výrazy v náhodných clusterech
    - AI – konzistentní
    - člověk – víc hapaxů



Reverse prompt engineeringem

= snaha zrekonstruovat prompt na základě textu

- ChatGPT je potřeba naprimovat
- *Teď chci použít reverse prompt engineering, abys mi pomohl vytvořit prompty na základě textů, které ti předložím a které budou optimalizované pro tvorbu podobného obsahu.*
- Vybrat text obsahově podobný tomu, který chcete generovat
- generování reverse promptu (získání obecnější struktury – *napiš větu o chození do školy*)
- rewrite, test, iterate


Nástroje
AI writing check
Copyleaks AI detector
GPTZero
Hive moderation
OpenAI text classifier
Undetectable AI
Writer

Was this text written by a **human** or **AI**?

Try detecting one of our sample texts:

ChatGPT | GPT4 | Bard | Human | AI + Human

Velké posily do skupiny Vigo Investments získává Pavel Řehák. Pro své aktivity spojené zejména s pojišťovnictvím a finančními službami přibírá dva nové partnery. Od září jsou menšinovými akcionáři skupiny Aleš Stýblo a Karel Soukeník, kteří přináší dlouholeté zkušenosti z velkého byznysu. Sedmačtyřicetiletý Stýblo přichází ze sklářské společnosti Lasvit, kterou v roce 2007 zakládal s Leonem Jakimičem a poslední roky ji vedl z pozice výkonného ředitele. Soukeník působil ve velkých finančních domech jako Citibank, Raiffeisenbank nebo Sberbank.

1725/5000 characters

Check Origin

Upload file

.pdf, .doc, .docx, .txt

By continuing you agree to our **Terms of service**

0%

This text is most likely to be written by **a human**

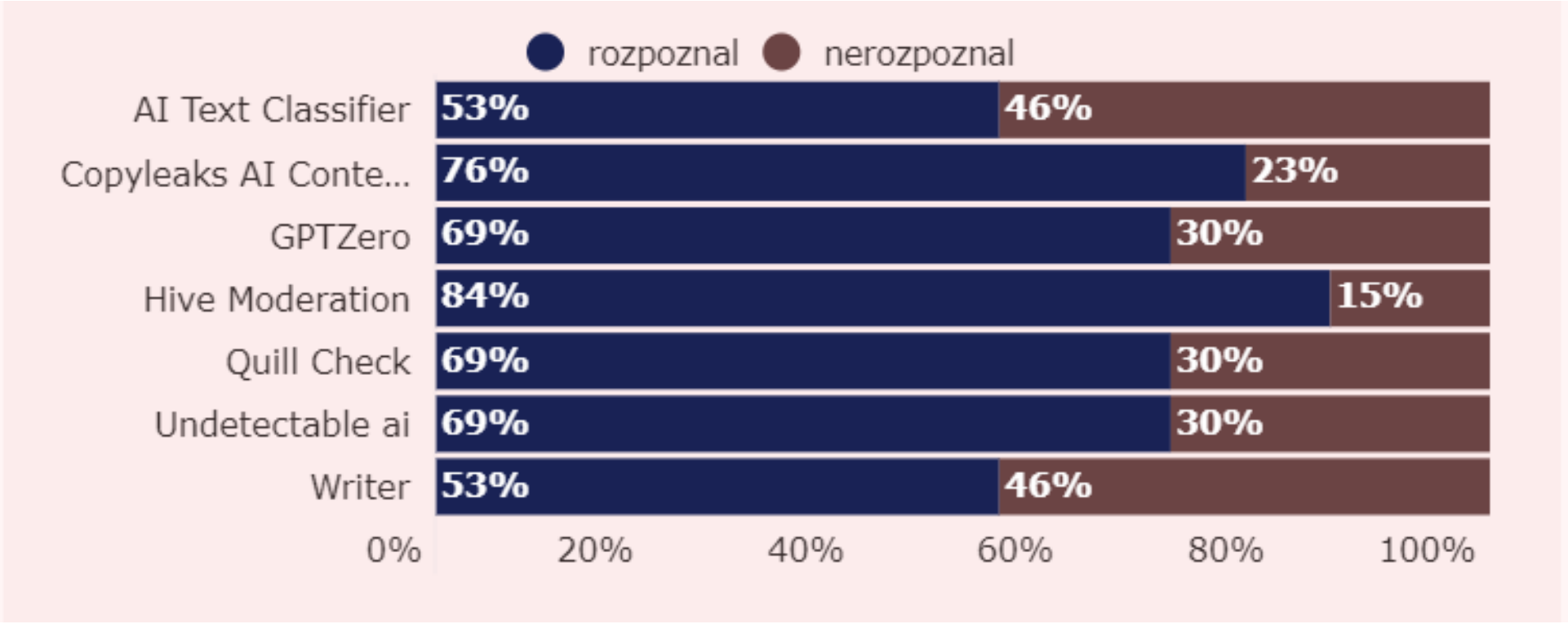
There is a **0%** probability this text was entirely written by AI

Velké posily do skupiny Vigo Investments získává Pavel Řehák. Pro své aktivity spojené zejména s pojišťovnictvím a finančními službami přibírá dva nové partnery. Od září jsou menšinovými akcionáři skupiny Aleš Stýblo a Karel Soukeník, kteří přináší dlouholeté zkušenosti z velkého byznysu. Sedmačtyřicetiletý Stýblo přichází ze sklářské společnosti Lasvit, kterou v roce 2007 zakládal s Leonem Jakimičem a poslední roky ji vedl z pozice výkonného ředitele. Soukeník působil ve velkých finančních domech jako Citibank, Raiffeisenbank nebo Sberbank.

„Mám velkou radost za tuto ohromnou posilu. Karel i Aleš během své kariéry několikrát dokázali, že neexistují malé ambice ani malé cíle. Stejně tak jsme silně kompatibilní hodnotově, v touze měnit svět kolem sebe k lepšímu," říká Pavel Řehák, většinový vlastník Vigo Investments a zakladatel Direct pojišťovny, která je hlavní byznysovou nohou celé rodiny Direct. Do ní spadají i další firmy jako Direct auto, Birne, Fidoo či Direct nadace. Dosud měla investiční skupina Vigo šestici partnerů, nově se rozroste právě o Aleše Stýbla s Karlem Soukeníkem. Detaily, za jakých podmínek oba do skupiny vstoupili, nebyly zveřejněny.

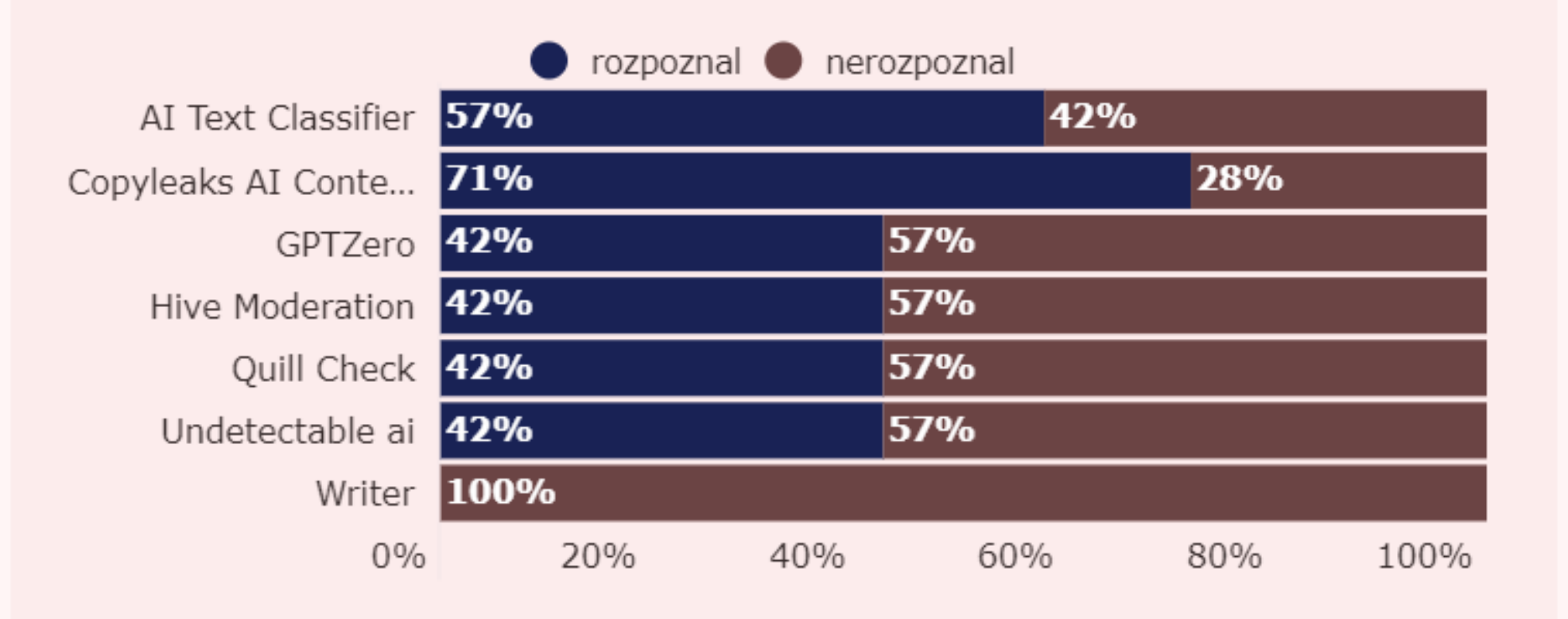
0/16 sentences are likely AI generated.

Úspěšnost na angličtině



Zdroj: Seznam Zprávy, Pavel Kasík, květen 2023

Úspěšnost na češtině



Zdroj: Seznam Zprávy, Pavel Kasík, květen 2023

Jsme dostatečně podezíraví?

Děkuji za pozornost.

Katerina Lesch
Contact: klesch@deloittece.com