

Vysoká škola ekonomická v Praze

Fakulta informatiky a statistiky



Princípy detekcie fake news na príklade ruskej invázie na Ukrajinu

DIPLOMOVÁ PRÁCA

Študijný program: Aplikovaná informatika

Študijný obor: Informační systémy a technologie

Autor: Bc. Daša Dvorská

Vedúci diplomovej práce: PhDr. Jan Černý, Ph.D.

Praha, máj 2023

Prehlásenie

Prehlasujem, že som diplomovú prácu „*Princípy detekcie fake news na príklade ruskej invázie na Ukrajinu*“ vypracovala samostatne, za použitia v práci uvedených prameňov a literatúry.

V Prahe dňa 1. mája 2023

.....

Bc. Daša Dvorská

Podakovanie

Chcela by som poďakovať PhDr. Janovi Černému, Ph.D. za vedenie diplomovej práce a cenné rady a konzultácie v priebehu jej tvorby. Tiež by som sa rada poďakovala svojej rodine za podporu, ktorú mi počas celého štúdia poskytovala.

Abstrakt

Ruská invaze na Ukrajinu je jen jednou z událostí, doprovázenou agresivní dezinformační kampaní, jejíž dopady vidíme i na Slovensku. Cílem práce je popsat principy detekce fake news a jejich aplikovatelnost při využití v kontextu události ruské invaze na Ukrajinu, na příkladu fake news šířených v slovenském jazyce. Práce se věnuje vymezení fake news a jejich blízkým konceptům, popisuje vlastnosti, způsoby šíření a atributy fake news. Popsány jsou přístupy k detekci fake news, a to na základě obsahu, šířitele a sociálního kontextu. Součástí práce je také vysvětlení politického a sociálního kontextu vývoje aktivit ruské dezinformační kampaně na Slovensku. Pro dosažení cíle práce je zpracována analýza a porovnání stávajících datasetů využívaných k detekci fake news a je vytvořen vlastní dataset obsahující fake news na zvoleném příkladu. Závěr práce obsahuje popis základních principů a diskusi nad jejich aplikovatelností při detekci fake news ve zvoleném kontextu. Přínosem této práce je tak prostřednictvím popisu principů k detekci fake news a vytvořením datasetu přispět k pochopení povahy fake news a také k odhalování a eliminaci fake news na Slovensku.

Klíčová slova

Fake news, detekce fake news, ruská invaze na Ukrajinu, FNDD

JEL klasifikace

O30

Abstrakt

Ruská invázia na Ukrajinu je len jednou z udalostí, sprevádzanou agresívnou dezinformačnou kampaňou, ktorej dopady vidíme aj na Slovensku. Cieľom práce je popísať princípy detekcie fake news a ich aplikovateľnosť pri využití v kontexte udalosti ruskej invázie na Ukrajinu, na príklade fake news šírených v slovenskom jazyku. Práca sa venuje vymedzeniu fake news a ich blízkym konceptom, popisuje vlastnosti, spôsoby šírenia a atribúty fake news. Popísané sú prístupy k detekcii fake news a to na základe obsahu, šíriteľa a sociálneho kontextu. Súčasťou práce je tiež vysvetlenie politického a sociálneho kontextu vývoja aktivít ruskej dezinformačnej kampane na Slovensku. Pre dosiahnutie cieľa práce je spracovaná analýza a porovnanie existujúcich datasetov využívaných k detekcii fake news a je vytvorený vlastný dataset obsahujúci fake news na zvolenom príklade. Záver práca obsahuje popis základných princípov a diskusiu nad ich aplikovateľnosťou pri detekcii fake news vo zvolenom kontexte. Prínosom tejto práce je preto prostredníctvom popisu princípov k detekcii fake news a vytvorením datasetu prispieť k pochopeniu povahy fake news a tiež k odhaľovaniu a eliminácii fake news na Slovensku.

Kľúčové slová

fake news, detekcia fake news, ruská invázia na Ukrajinu, FNDD

JEL klasifikácia

O30

Abstract

The Russian invasion of Ukraine is just one of the events, accompanied by an aggressive disinformation campaign, the effects of which we can also see in Slovakia. The aim of the thesis is to describe the principles of fake news detection and their applicability when used in the context of the event of the Russian invasion of Ukraine, using the example of fake news in the Slovak language. The thesis describes fake news and other closely related concepts, the characteristics, dissemination methods and attributes of fake news. Approaches to fake news detection, based on content, spreader and social context are described. The thesis also discusses the political and social context of the Russian disinformation campaign in Slovakia. To achieve the goal of the thesis, an analysis and comparison of existing datasets used to detect fake news is processed and a custom dataset containing fake news in the selected example is created. The last part of the thesis contains a description of the basic principles and a discussion of their applicability in the detection of fake news in the chosen context. Overall, this thesis contributes to the study of fake news detection, with implications for the development of more effective and context-aware tools for combating the spread of fake news.

Keywords

Fake news, fake news detection, Russian invasion of Ukraine, FNDD

JEL Classification

O30

Obsah

Úvod.....	10
1 Definícia problému	11
1.1 Dôvod výberu témy	11
1.2 Politický a sociálny kontext	11
2 Charakteristika fake news.....	17
2.1 Koncepty blízke fake news.....	17
2.2 Vymedzenie pojmu fake news.....	19
2.3 Šírenie fake news	20
2.3.1 Automatizované programy na šírenie fake news.....	20
2.3.2 Algoritmy vytvárajúce obsah.....	21
2.4 Atribúty fake news	22
2.4.1 Obsah fake news	22
2.4.2 Šíriteľ.....	23
2.4.3 Sociálny kontext	23
2.5 Vlastnosti fake news.....	24
3 Prístupy detekcie fake news.....	26
3.1 Detekcia na základe obsahu fake news.....	26
3.1.1 Overovanie faktov	26
3.1.2 Automatické overovanie faktov	27
3.1.3 Detekcia na základe štruktúry a štýlu.....	28
3.2 Detekcia na základe šíriteľa	29
3.3 Detekcia na základe sociálneho kontextu	31
4 Popis výskumu a metodológie	33
4.1 Výskumné otázky.....	33
4.2 Metodológia	33
5 Datasetsy pre detekciu fake news.....	34
5.1 LIAR.....	34
5.2 FakeNewsNet	36
5.3 COVID-19 fake news dataset v slovenskom jazyku.....	37
5.4 Fake news dataset v slovenskom jazyku	38
5.5 Porovnanie datasetov pre detekciu fake news	39
6 Detekcia fake news na zvolenom príklade.....	43
6.1 Zvolený kontext	43

6.2 Zdroje dát	44
6.2.1 Zem & Vek.....	44
6.2.2 News front SK.....	45
6.2.3 SITA Slovenská tlačová agentúra	46
6.3 Štruktúra datasetu	47
6.4 Zber dát	47
6.4.1 Zem & Vek.....	49
6.4.2 News front SK	51
6.4.3 Slovenská tlačová agentúra SITA.....	52
6.5 Ukážka a popis výsledného datasetu	54
6.6 Hlavné naratívy fake news obsiahnuté v datasete	55
6.7 Použitelnosť	57
7 Princípy detekcie fake news.....	58
7.1 Vymedzenie úlohy detekcie fake news	58
7.2 Dostupnosť dát a voľba zdrojov	58
7.3 Určenie hodnotiacej škály.....	60
7.4 Určenie spôsobu klasifikácie.....	60
7.5 Definícia datasetu	61
8 Diskusia	62
8.1 Špecifiká a výzvy tvorby datasetu v zvolenom kontexte.....	62
8.2 Diskusia nad budúcnosťou detekcie fake news	62
Záver	65
Použitá literatúra	66
Prílohy	I
Príloha A: Dataset pre detekciu fake news	I

Zoznam obrázkov

Obrázok 1 Kroková kaskáda (zdroj: Zhou a Zafarani 2020).....	32
Obrázok 2 Časová kaskáda (zdroj: Zhou a Zafarani 2020)	32
Obrázok 3 Sekcia "Ukrajina" na stránke Zem a Vek. Zdroj: Zem a Vek.....	50
Obrázok 4 Článok na stránke Zem a Vek. Zdroj: Zem a Vek.....	50
Obrázok 5 Sekcia "Ukrajina" na stránke News front SK, Zdroj: News front SK.....	51
Obrázok 6 Článok na stránke News front SK, Zdroj: News front SK	52
Obrázok 7 Sekcia "Rusko-ukrajinský konflikt" na stránke SITA, Zdroj: SITA	53
Obrázok 8 Článok na stránke SITA, Zdroj: SITA	54
Obrázok 9 Ukážka datasetu s príspevkami typu Real, Zdroj: autor	55
Obrázok 10 Ukážka datasetu s príspevkami typu Fake, Zdroj: autor.....	55

Úvod

Nie každý článok na internete je objektívnym zhodnotením reálneho diania. Vďaka slobode panujúcej na internete a predovšetkým na sociálnych sieťach, môžu články a príspevky zahŕňať osobné postoje autora, alebo nepravdivé informácie, ktorých cieľom je rôznym spôsobom ovplyvniť čitateľa a nie poskytnúť pravdivé a overené informácie. Klamlivé informácie v podobe fake news sa šíria internetovým priestorom oveľa rýchlejšim tempom ako seriózne správy, a to vďaka nielen pútavému obsahu, ale aj tlaku a pokročilým prostriedkom vynaložených na ich šírenie zo strany záujmových strán. Extrahovanie informácií z veľkého množstva neštruktúrovaných dát a ich overovanie je neľahkou a časovo náročnou úlohou. Spoločne s nedostatkom kontroly nad obsahom zdieľaným na internete stoja tieto dôvody za masovým šírením fake news, ktoré má negatívny vplyv na spoločnosť. To zvyrazňuje potrebu vývinu a aplikácie pokročilých prostriedkov pre ich včasnú detekciu a následnú elimináciu.

Ruská agresia voči Ukrajine je jednou z udalostí, sprevádzanou vysokou dezinformačnou kampaňou, ktorej dopady vidíme aj na Slovensku. Cieľom tejto práce je preto popisom prístupov k detekcii fake news a vytvorením datasetu na danom príklade prispieť k pochopeniu povahy fake news a tiež k odhaľovaniu a eliminácii fake news na Slovensku.

V teoretickej časti práce sú definované fake news a pojmy, ktoré s fake news blízko súvisia, s cieľom popísať rozdiely medzi nimi. Práca sa bližšie venuje fake news a popisu ich vlastností a prostriedkov, ktorými sú dnes masívne šírené. Ďalej sú vymedzené atribúty fake news ako obsah, šíriteľ a sociálny kontext, na základe ktorých sú rozlíšené a popísané tiež možné prístupy k detekcii fake news.

Praktická časť sa venuje definícii princípov detekcie fake news a ich využitiu na príklade tvorby datasetu fake news na tému ruskej invázie na Ukrajinu, šírených na Slovensku. Najprv sú analyzované existujúce datasety k detekcii fake news so zameraním na široko využívané datasety v anglickom jazyku, ale tiež menšie datasety v slovenskom jazyku. Ich popis prebieha na základe porovnávania dôležitých charakteristík, popisujúcich datasety využívané pri aplikácii na úlohách detekcie fake news. Ďalej je vytvorený vlastný dataset v určenom príklade. Na záver je v práci uvedený popis jednotlivých princípov a tiež diskusia nad obmedzeniami pri detekcii fake news v danom kontexte.

Prínosom tejto práce je preto prostredníctvom popisu princípov k detekcii fake news a vytvorením datasetu prispieť k pochopeniu povahy fake news a tiež k odhaľovaniu a eliminácii fake news na Slovensku. Dôležitým produktom práce je tiež vytvorený dataset, obsahujúci celkovo 511 záznamov, využiteľný pri jednoduchých úlohách detekcie fake news.

1 Definícia problému

V tejto kapitole sú popísané hlavné motívy, ktoré stáli za výberom témy diplomovej práce a tiež sociálne a politické okolnosti, s cieľom priblížiť fenomén fake news ako nástroj pre vedenie hybridnej informačnej vojny, so zameraním na ruské dezinformačné aktivity na Slovensku.

1.1 Dôvod výberu témy

Podľa prieskumu, ktorý uskutočnila spoločnosť Nielsen Admosphere v marci 2019 na vzorke päťsto Slovákov aktívnych na internete, 41% respondentov získava čerstvé informácie a novinky o aktuálnom dianí na spravodajských serveroch, ktoré majú v obľube. Na druhom mieste s 34% skončil ako primárny informačný zdroj sociálna sieť Facebook a len 13% opýtaných uviedla, že správy na internete nesledujú. Medzi najmenej udávané online zdroje informácií sa zaradili iné sociálne siete ako napríklad Twitter, newsletter a RSS čítačky. (Nielsen 2019). Používanie internetu a sociálnych sietí k získaniu informácií je populárne vďaka lacnej a ľahkej dostupnosti a rýchlosti s akou sú správy v tomto prostredí šírené. Na druhú stranu stoja rovnaké dôvody za tým, prečo je toto prostredie používané pre masové šírenie fake news. (Shu et al. 2017)

Dostupnosť informácií a fenomén fake news začínajú mať významný dopad na súčasné geopolitické dianie a názory verejnosti a stávajú sa preto aktuálnou témou nielen v oblasti sociologických a politologických štúdií, ale čoraz viac je téme venovaná pozornosť aj v oblasti informačných a komunikačných technológií. V rámci nich je cieľom analyzovať šírenie fake news a vytvárať nástroje, ktoré by toto šírenie dokázali úspešne odhaliť a následne obmedziť. Existujú rôzne prístupy, ktoré sú pri detekcii fake news vyžívané. Cieľom tejto práce je vysvetliť princípy detekcie fake news a to na zvolenom príklade fake news, ktoré vznikli v reakcii na udalosť ruskej invázie Ukrajinu v roku 2022, šírených v slovenskom jazyku, v prostredí slovenského online priestoru. Princípy detekcie budú overené na praktickom prípade tvorby datasetu.

Výskumy zdôrazňujú potrebu venovať sa detekcii fake news v rôznych kontextoch, vďaka čomu je možné overiť šírku aplikovateľnosti rôznych navrhovaných prístupov. Práca tak prispieva k pochopeniu detekcie fake news v zvolenom špecifickom kontexte a jazyku a potenciálne prispieva tiež k využitiu týchto znalostí a vytvoreného datasetu pri úspešnom boji s neustále narastajúcimi hrozbami, ktoré fake news predstavujú.

1.2 Politický a sociálny kontext

Napriek tomu, že problematika ruských dezinformačných aktivít v prostredí Slovenskej republiky získala pozornosť iba v posledných rokoch, ide o fenomén, ktorý má svoje širšie medzinárodnopolitické dôvody. Zároveň, toto pôsobenie je do značnej miery naviazané na

aktivity Ruskej federácie na Ukrajine, keďže ruská činnosť v tejto oblasti na Slovensku výrazne získala na význame po ruskej anexii polostrova Krym v roku 2014. Dôvody takéhoto pôsobenia na Slovensku, a v priestore strednej Európy, však možno hľadať už v dávnejšej histórii.

Slovenská republika, ako súčasť Československej socialistickej republiky, bola v období studenej vojny a bipolárneho usporiadania sveta súčasťou tzv. východného bloku a patrilo do sféry vplyvu dominantnej mocnosti, ktorou bol Zväz sovietskych socialistických republík (ZSSR). Československý štát, jeho politika a spoločnosť bola pod silnou kontrolou Moskvy. Komunistické vedenie v Moskve malo rozhodujúci vplyv pri určovaní politicko-spoločenského vývoja v štáte. Tento vplyv bol umocnený po auguste 1968, kedy okupácia vojskami Varšavskej zmluvy, riadená Moskvou, potlačila snahy o liberalizáciu politicko-spoločenských pomerov, známu aj ako tzv. Pražská jar. Politické vedenie ZSSR dosadilo do čela štátu lojálnych predstaviteľov Komunistickej strany, čím stabilizovalo politický vplyv na nasledujúce dekády. Obdobie tzv. normalizácie tak bolo charakteristické absolútnou poslušnosťou politického vedenia Československa vláde ZSSR, vďaka čomu Moskva mohla naplno realizovať kontrolu a posilňovať vplyv v štáte, ktorý mal pre ňu aj značný geopolitický význam. Československo ležalo na hranici východného bloku a priamo susedilo so západným blokom, čo zvyšovalo záujem Moskvy o vplyv v Československu. O tento vplyv prišla až po vyše dvoch desaťročiach. November 1989 a tzv. Nežná revolúcia znamenali koniec vlády Komunistickej strany a tým pádom aj stratu kontroly zo strany ZSSR.

Revolúcie v roku 1989 a následný rozpad ZSSR v roku 1991 znamenali koniec studenej vojny a bipolárneho usporiadania sveta. ZSSR sa rozpadol na 15 štátov, z ktorých jedným je Ruská federácia. Tá bola v úvode 90. rokov zmietaná vnútropolitickou nestabilitou, čo znamenalo aj stratu jej predošlého vplyvu v regióne strednej Európy, a to vrátane nástupníckych štátov Československa, ktorými za v roku 1993 stali Česká republika a Slovenská republika. Vývoj v nasledujúcom období znamenal obrat v zahraničnopolitickom smerovaní Slovenskej republiky, ktorého signifikantným momentom bol rok 2004, kedy Slovensko vstúpilo do NATO a Európskej únie. Slovensko sa takto výraznejšie vytratil zo sféry vplyvu Ruskej federácie. Tento ruský vplyv na Slovensku začal opätovne naberať na sile až po nástupe Vladimíra Putina k moci, a to predovšetkým po ruskej anexii Krymu a rozpútaní konfliktu vo východnej Ukrajine v roku 2014.

Šírenie svojho vplyvu v zahraničí a budovanie vplyvu vo svete je legitímnym záujmom každého štátu. Ide o elementárny aspekt zahraničnej politiky. Na tento účel budovania zahraničnopolitického vplyvu za svojimi hranicami disponujú štáty širokým súborom nástrojov tvorby zahraničnej politiky. V prípade dezinformácií je však problematické hovoriť o legitímnom nástroji na budovanie vplyvu. Jednu z najvyužívanejších kategorizácií nástrojov zahraničnej politiky priniesol vo svojej práci (Nye 2004) americký politológ prof. Joseph Nye, ktorý nástroje zahraničnej politiky rozdelil na tvrdú moc (hard power) a mäkkú moc (soft power). Kým tvrdá moc spočíva na princípe prikazovania prostredníctvom hrozieb a stimulov, s využitím hmatateľných nástrojov a donucovacích metód, mäkká moc spočíva na princípe atraktívnosti, v schopnosti vzbudiť rešpekt a obdiv. Tvrdá moc dokáže ostatné štáty primäť k tomu, aby robili to, čo by za iných okolností nerobili. Mäkká moc dokáže dostať ostatných aktérov do situácie, keď je ich konanie dobrovoľné. Základný princíp koncepcie mäkkej moci je, že je „lepšie si ľudí získať, než nútiť“. Mäkká moc spočíva

na schopnosti formovať preferencie druhých svojím príkladom. Konkrétnymi nástrojmi tvrdej moci tak sú mocenské a silové nástroje ako využitie vojenských kapacít, vojenské zastrašovanie, diplomatický nátlak či využívanie rôznych foriem ekonomického donucovania ako embargo, bojkot a pod. Hlavnými zdrojmi mäkkej moci sú kultúra, hodnoty či politika. Ide konkrétne o nástroje ako šírenie pop-kultúry či vysokej kultúry, efektívne využívanie študentských mobilit, aktivity občianskej spoločnosti, atraktivita vychádzajúca z cestovného ruchu či životného štýlu, príťažlivosť politického systému, ekonomického modelu, či aktivity vlády ako verejná diplomacia, ekonomická pomoc či vojenská a diplomatická asistencia. (Nye 2004)

Dezinformácie, ako nástroj zahraničnopolitického pôsobenia štátu, je problematické zaradiť v rámci tejto zaužívanej kategorizácie. Kým napríklad vytvorenie a šírenie ruských štátnych médií Russia Today či Sputnik International vo svete možno definovať ako legitímny nástroj mäkkej moci, pokiaľ táto televízna stanica šíri dezinformácie či iný manipulatívny a propagandistický obsah, za nástroj mäkkej moci ju už považovať nemožno. Podobne, na príklade Čínskej ľudovej republiky, zakladanie osvetovo-vzdelávacích Konfuciových inštitútov vo svete je legitímnym nástrojom moci. Pokiaľ však takéto inštitúty neslúžia na výučbu jazyka či propagáciu kultúrneho dedičstva štátu, ale začnú v rámci svojich aktivít šíriť informácie, ktoré majú charakter dezinformácií či štátnej propagandy, ide o manipulatívny nástroj, ktorý do rámca mäkkej moci nezapadá. Pre takéto praktiky dezinformačného charakteru patria do kategórie nástrojov, ktoré zdanlivo pôsobia ako nástroje mäkkej moci, ale svojím obsahom do jej definície nezapadajú. Pre pomenovanie takýchto manipulatívnych praktík zahraničnopolitického pôsobenia štátov sa v teórii udomácnil nový pojem, ostrá moc (sharp power). Tento pojem bol predstavený v roku 2017 v článku magazínu Foreign Affairs s názvom The Meaning of Sharp Power: How Authoritarian States Project Influence (Walker a Ludwig 2017). Pojmom ostrá autori označujú práve súbor techník, pod ktoré spadajú aj dezinformačné aktivity, a ktoré sú charakteristické v pôsobení autoritárskych štátov ako Ruská federácia alebo Čínska ľudová republika. Takéto režimy dokážu efektívne využívať otvorenosť demokratických režimov na to, aby v ich informačnom priestore dokázali budovať vplyv využívaním manipulatívnych nástrojov. Označenie „ostrá“ moc tak v sebe nesie naratív agresívneho prenikania, čím sa odlišuje od pozitívnej konotácie pojmu mäkká moc. Nástroje ostrej moci sa tak snažia manipulovať cieľené publikum v snahe zvyšovať vplyv intervenujúceho štátu, alebo podkopávať legitimitu vlády štátu, v ktorom sú aplikované. Aplikovanie dezinformačných kampaní tak spadá do takejto kategorizácie, vďaka čomu je jednoduchšie rozlišovanie legitímnych nástrojov pôsobenia od nástrojov nelegitímnych.

Ruská dezinformačná kampaň v štátoch strednej Európy je len súčasťou pôsobenia, ktoré teoretici označujú širším pojmom informačná vojna. Ide o široký súbor techník a nástrojov, ktoré majú slúžiť na rozrušovanie a manipulovanie zahraničnej verejnosti.

Dezinformačná kampaň je v tomto kontexte efektívnym prostriedkom, ktorý je realizovaný skrz ruské štátne médiá, pro-ruské dezinformačné médiá v cieľových štátoch, či samotnú verejnosť, ktorá je týmto ruským vplyvom či sentimentom ovplyvnená. V prostredí Slovenskej republiky zohrávajú v tomto kontexte významnú úlohu aj politickí aktéri, ktorí sa snažia o získanie podpory tejto časti spoločnosti ovplyvnenej proruským sentimentom.

Pri tejto svojej snahe sa sami stávajú nositeľom ruských dezinformácií, čím signifikantným spôsobom formujú verejnú mienku v štáte.

Takýmto dezinformačným aktivitám výrazne napomáha aj vysoká náchylnosť slovenskej verejnosti veriť konšpiráciám. Podľa výskumu think-tanku Globsec z roku 2020, v priemere 56% respondentov na Slovensku je náchylných veriť nejakému tvrdeniu, ktoré v sebe zahŕňa konšpiráciu či klamstvo. Slovenská verejnosť je tak v tomto ohľade výrazne najnáchylnejšia v regióne štátov V4 (v Maďarsku 35%; Poľsku 34%, Českej republike 29%). (Hajdu a Klingová 2020) Takéto prostredie vytvára dobré prostredie pre šírenie dezinformačných kampaní, ktoré nie len slúžia posilňovaniu vplyvu cudzích štátov, ale taktiež podkopávajú hodnotové preferencie a geopolitickú orientáciu v štáte. Ako príklad môžu slúžiť ďalšie dáta, ktoré hovoria o tom, že slovenská verejnosť má v rámci V4 najnižšiu príslušnosť k „Západu“ (26%), vidí v oveľa väčšej miere ako strategického partnera Rusko (47%) než USA (17%), má nadpriemerne pozitívne vnímanie osoby Vladimíra Putina (55%, priemer v regióne strednej a východnej Európy 35%), má vo V4 najvyššiu podporu vystúpenia z NATO (25%, kým Poľsko 5%, Maďarsko 10% a Česká republika 18%).(Hajdu et al. 2021) Vysokú efektivitu ruských dezinformácií dokazujú aj dáta z obdobia pandémie Covid-19. Prieskum agentúry Focus z marca 2020 ukázal, že 24,9% Slovákov považovalo Rusko za štát, ktorý Slovensku pomohol v boji s koronavírusom, kým pri EÚ sa pozitívne vyslovilo iba 22% opýtaných. Išlo o ukážku absolútnej dezorientácie verejnosti, keďže pomoc zo strany EÚ bola v miliónoch EUR, kým výraznejšia pomoc zo strany Ruska nebola zaznamenaná.(Mikušovič 2020)

Aj tieto dáta sú však ukážkou systematickej práce dezinformačných aktérov, ktorých terčom je dlhodobo aj Európska únia samotná. Ako uvádza portál Infosecurity.sk, od začiatku vojny sa dezinformační aktéri snažia odradiť EÚ od ďalšej podpory Ukrajiny vo vojne s Ruskom. Zároveň, Európska únia je obviňovaná z toho, že si za svoje ekonomické problémy môže sama a Putina využíva len ako strašiaka na prekrytie vlastnej neschopnosti. Dezinformační aktéri zároveň šíria vymyslené lži o európskych predstaviteľoch s cieľom posilniť nedôveru voči nim. Akou uvádza analýza Infosecurity.sk, šíriteľmi takýchto dezinformačných naratívov nie sú len populárne slovenské dezinformačné médiá ako Hlavný denník či Slobodný vysielateľ, ale aj mnohí vrcholoví politici, ako predseda krajne pravicovej strany Republika a europoslanec Milan Uhrík, bývalý premiér a predseda strany Smer-SD Rober Fico, podpredseda strany Smer-SD Ľuboš Blaha, či bývalý predseda NRSR a predseda SNS Andrej Danko. Tí šíria naratívy o „fašizácii EÚ“, spochybňujú poskytovanie pomoci Ukrajine zo strany EÚ či za infláciu alebo energetickú krízu vinia politiky EÚ a sankcie namierené proti Ruskej federácii. (Dinič 2022)

Aj tieto dáta ukázali, akou silnou podporou Ruská federácia v slovenskej spoločnosti disponuje. Ide o výsledok systematickej a cielenej práce ruskej strany v ovplyvňovaní informačného priestoru na Slovensku, kedy prím hrajú najmä tzv. alternatívne médiá, teda internetové portály, ktoré dokázateľne šíria konšpiračné teórie alebo dezinformácie. Databázu takýchto webstránok pôsobiacich v slovenskom informačnom priestore vytvoril projekt konspiratori.sk a v októbri 2022 v tejto databáze figurovalo 261 webov.(Konspiratori 2022)

Jeden z najvplyvnejších dezinformačných webov hlavnespravy.sk musel byť na základe súdneho rozhodnutia z databázy odstránený, až pokým v tejto veci nepadne oficiálne rozhodnutie súdu. Pritom, práve jeden z autorov hlavnespravy.sk bol zadržaný Národnou kriminálnou agentúrou pre podozrenie z vyzvedačstva pre Rusko. Túto skutočnosť podporuje aj zverejnená videonahrávka tajných služieb, ktorá ukazuje ako vojenský pridelenec Veľvyslanectva Ruskej federácie na Slovensku lanári autora webu do svojich služieb.(Hodás 2022) Začiatkom marca 2022 dokonca Národný bezpečnostný úrad SR dočasne zablokoval činnosť webu hlavnespravy.sk uplatnením zákona o kybernetickej bezpečnosti. Podobný prípad sa týkal aj populárneho magazínu Zem a Vek, ktorého majiteľ Tibor E. Rostas bol zachytený na ruskej ambasáde, ako žiadal o finančnú podporu.(Harkotova 2016)

Ruská informačná vojna začala byť viditeľnejšia po roku 2014, kedy Ruská federácia anektovala polostrov Krym a podnietila vznik vojenského konfliktu na východe Ukrajiny. Na intenzite následne výrazne získala po 24. februári 2022, kedy Ruská federácia vojenským útokom začala vojenskú agresiu proti Ukrajine. Tento svoj útok sa Rusko snažilo legitimizovať dezinformačnou kampaňou, ktorá predchádzala samotnému vojenskému útoku, kedy ruské dezinformačné kanály šíрили informácie o genocíde etnicky ruského obyvateľstva či hrozbe zo strany „nacistického režimu“ v Kyjeve. Potencionálny ruský útok sa tak ma javiť ako legitímna reakcia na údajné zločiny páchané ukrajinskou stranou. Po začatí ruskej agresie a okupácie ukrajinského územia sa následne dezinformačná kampaň začala zameriavať na spochybňovanie proti ruských sankcií, podkopávanie pomoci štátov EÚ Ukrajine či oslabovanie solidarity voči ukrajinskému obyvateľstvu. Zaujímavosťou v tomto ohľade je, že ruská invázia 24. februára zaskočila a prekvapila značnú časť pro-ruských aktivistov na Slovensku. Ako poukazuje politológ Mesežnikov, tento útok bol v rozpore s dovtedajšími ruskými naratívami o mierovej politike Kremľa, ktoré odmietali akékoľvek obvinenia z plánovaného útoku na Ukrajinu. Slovenskej dezinformačnej scéne trvalo niekoľko dní, kým dokázala opäť aktivizovať svoju činnosť. (Čorej a Mesežnikov 2022) Takéto dezinformačné aktivity sú súčasťou širšieho pôsobenia v rámci tzv. hybridných hrozieb.

Dokument *Koncepcia pre boj Slovenskej republiky proti hybridným hrozbám* (NBÚ SR 2018) definuje pojem hybridné hrozby ako „súbor nátlakových a podvratných činností, konvenčných a nekonvenčných, vojenských a nevojenských metód, ktoré môžu štátne aj neštátne subjekty koordinovaným spôsobom využívať na dosiahnutie konkrétnych cieľov bez formálneho vyhlásenia vojny a pod prahom zvyčajnej reakcie“. Aktuálnejší vládny dokument *Akčný plán koordinácie boja proti hybridným hrozbám 2022 – 2024* (Ministerstvo obrany SR 2021) dodáva, že „Hybridné pôsobenie proti záujmom Slovenskej republiky sa za posledných niekoľko rokov presunulo z periférie do hlavného prúdu na takmer všetky úrovne spoločnosti... Tieto aktivity sú často koordinované, zameriavajú sa na lokálne zraniteľnosti a sú navrhnuté tak, aby zostali pod hranicou detekcie, atribúcie a z toho vyplývajúcej zvyčajnej reakcie“. Ako doplnil riaditeľ Odboru pre hybridné hrozby a strategickú komunikáciu na Ministerstve obrany Victorom Breiner, hybridné hrozby pôsobia ako akási predĺžená ruka tradičného vojenského ťaženia. Sú však nadizajnované tak, aby si ho percipienti nevšimli. Sú ukážkou toho, že vojna neprebíha iba na bojisku, ale je prítomná aj v informačnom priestore. (Disinfo Report 2022a) Riaditeľ Centra boja proti hybridným hrozbám Ministerstva vnútra Slovenskej republiky Daniel Milo popisuje dva

základné naratívy od začiatku invázie sú prenesenie viny a zodpovednosti zo strany útočníka na kohokoľvek iného (ukrajina- fašisti), alebo je to vykresľované ako konflikt medzi ruskou federáciou a západom, pričom Rusko sa v tomto konflikte len bráni. Nové naratívy, ktoré prichádzali v priebehu času, sa týkali predovšetkým vývoja na bojisku. Ako ďalej vysvetľuje, Slovensko je pre pôsobenie ruských dezinformačných operácií tzv. Frontový štát (tzv. *front line state*) z dôvodu, že susedí s Ukrajinou, čím znáša veľký tlak migračných tokov, ale zároveň aj z dôvodu náchylnosti značnej časti slovenskej populácie preberať ruské videnie sveta, resp. kremľské videnie, o *zlom západe, ktorý sa snaží rozdrobiť Rusko*. Tým sa na Slovensku vedie zápas o postoje verejnosti, ktoré sa môžu premietiť do politických preferencií a následnej zahraničnej politiky. Ako poukazuje Milo, medzi najvýznamnejších nositeľov týchto naratívov však nepatria len dezinformačné médiá, ale aj politici, ktorí majú na týchto kanáloch vysoký dosah a zneužívajú, že časť populácie počúva na takýto typ argumentácie bez toho, aby zodpovedal realite. Preto takéto prostriedky slúžia na získavanie politických preferencií a rozloženie stability spoločnosti s konečným cieľom vo forme zmeny zahraničnopolitickej orientácie Slovenskej republiky, zastavenie sankcií či zablokovania pomoci Ukrajine. (Disinfo Report 2022b)

Čo sa týka konkrétnych naratívov, ktoré počas úvodných mesiacov vojny šírili v slovenskom informačnom prostredí dezinformační aktéri, ako uvádza analytička projektu Infosecurity.sk Michaela Ružičková: „*Kremľ sa dlhodobo snaží koordinovanými postupmi narúšať obraz suverénnej Ukrajiny. Ruskí propagandisti preto šíria veľké množstvo dezinformačných naratívov a mýtov útočiacich na ukrajinskú štátnosť a národnosť.*“ (Ružičková 2022) Medzi časté naratívy tak patrí napríklad spochybňovanie ukrajinskej suverenity a operovanie tvrdením, že Ukrajina je akýsi bábkový, či sirotsky štát, ktorý má byť ovládaný z vonku, externými aktérmi, ako je USA alebo NATO. Ako uvádza Ružičková, ide o naratív, ktorý je v súlade s oficiálnou ruskou štátnou rétorikou, keďže ruský prezident Vladimir Putin sa ešte v roku 2008 vyjadril, že „*Ukrajina nie je krajina*“. V roku 2021 vo svojom článku tvrdil, že Ukrajina a Rusko sú jeden národ a „*skutočná suverenita Ukrajiny je možná len v partnerstve s Ruskom*“. Tým živil mýtus o tzv. *bratských národoch*. Spochybňovanie ukrajinskej suverenity je tak častým naratívom viditeľným v rétorike štátnych predstaviteľov Ruskej federácie. Príkladom je podpredseda Bezpečnostnej rady a bývalý prezident Dmitrij Medvedev, ktorý na svojich sociálnych sieťach šírili mapu toho, ako by malo byť územie Ukrajiny rozdelené medzi Ruskú federáciu, Poľsko, Rumunsko a Maďarsko, čím sa snaží podkopávať legitimitu a územnú celistvosť ukrajinského štátu.

Pôsobenie týchto dezinformačných aktivít zachytávajú aj prieskumy verejnej mienky v inkriminovaných otázkach týkajúcich sa konfliktu na Ukrajine. Najaktuálnejšie dáta priniesol výskum *GLOBSEC Trends CEE amid the war in Ukraine*. (Kazaz et al. 2022) Zo štátov strednej a východnej Európy sú práve Slováci v najväčšej miere presvedčení, že za konflikt nesie zodpovednosť Ukrajina (28%, kým napr. zo štátov V4 v Poľsku tento pohľad podporuje 5%, v Česku 10% a v Maďarsku 19%). Rovnako, v Slovenskej spoločnosti je najviac zakorenený názor, že Ukrajina je súčasťou Ruska (až 13%, v Poľsku 2%, Česku 5% a v Maďarsku 8%). Spomedzi štátov V4 je Vladimir Putin najpozitívnejšie vnímaný práve na Slovensku (24%, v Poľsku 2%, Česku 4% a v Maďarsku 19%). Taktiež podpora tvrdých sankcií zo strany EÚ voči Rusku patrí na Slovensku medzi najslabšie v regióne.

2 Charakteristika fake news

Pre potreby diplomovej práce je táto kapitola zameraná na teoretické vymedzenie skúmaného problému a definíciu fake news. Keďže definícia fake news nie je v súčasnosti stále pevne ukotvená a existuje viacero rôznych výkladov fake news, táto kapitola má za cieľ popísať, ako je pracované s termínom fake news vo vedeckých prácach, popísať šírenie fake news v online prostredí, zasadiť fake news do kontextu digitálnej správy a určiť vlastnosti a atribúty fake news.

2.1 Koncepty blízke fake news

Fake news, napriek tomu, že sa ako fenomén vyskytuje v našej spoločnosti už dlho, je v modernom chápaní relatívne nový pojem, pre ktorý nie je ukotvená žiadna jednotná definícia. (Sousa-Silva 2022) Dôvodom je, že existuje mnoho blízkych pojmov a konceptov, ktoré s pojmom fake news blízko súvisia a niekedy bývajú vzájomne zamieňané bez rozdielu významu. Ide napríklad o pojmy ako dezinformácie, misinformácie, satira, clickbait, fámy a podobne.

Tieto pojmy je podľa (Zhou a Zafarani 2020) možné rozlíšiť na základe troch základných charakteristík. Prvou charakteristikou je autenticita, ktorá hodnotí pravdivosť stanovených tvrdení. Ak príspevok obsahuje aspoň jednu nepravdivú informáciu, nie je autentický. Nepravdivú informáciu je možné určiť len na základe objektívne overiteľných faktov. Druhou charakteristikou je zámer, s ktorým je príspevok šírený. Správy môžu byť šírené so zámerom získania politického, finančného, alebo osobného prospechu, či za účelom spôsobiť škodu (Buchanan 2020). Poslednou charakteristikou je, či je informácia obsiahnutá v príspevku typu news, a teda či obsahuje príznak novosti. V angličtine sú správy s príznakom novosti označené ako „news“, ktorý je do slovenčiny možné preložiť ako „správa“ alebo „novinka“.

Príznak novosti je možné určiť tým, že je správa šírená prostredníctvom spravodajských kanálov. Vo výkladovom online slovníku Cambridge Dictionary je „news“ vo voľnom preklade definovaná ako informácia alebo príspevok o nedávnej udalosti. (Cambridge Dictionary 2022). Fake news je tak správa, ktorá sa na news podobá. Je tak šírená prostredníctvom „spravodajských“ kanálov a je formulovaná ako nová informácia o aktuálnych udalostiach, ale v skutočnosti nemusí ísť o seriózne spravodajstvo a aj keď sa týka aktuálnych udalostí, informácie, ktoré prináša, nie sú pravdivé.

Príklad, ako na základe týchto charakteristík je možné rozlíšiť koncepty blízke fake news, prezentuje (Zhou a Zafarani 2020) nasledovne:

Tabuľka 1 Porovnanie konceptov blízkyh fake news. Zdroj: (Zhou a Zafarani 2020)

	Autenticita	Zámer	News Príznak novosti
Dezinformácie	Nie	Zavádzanie	Nie je definovaný
Misinformácie	Nie	Nie je definovaný	Nie je definovaný
False news	Nie	Nie je definovaný	Áno
Deceptive news	Nie	Zavádzanie	Áno
Satira	Nie je definovaný	Pobavenie	Nie je definovaný
Clickbait	Nie je definovaný	Zavádzanie	Nie je definovaný
Fámy	Nie je definovaný	Nie je definovaný	Nie je definovaný

Na základe Tabuľka 1 je možné odvodiť, že zatiaľ čo dezinformácie a misinformácie obe obsahujú nepravdivé informácie, dezinformácie sú narozdiel od misinformácií šírené s úmyslom klamať, alebo zavádzať adresáta. Misinformácie sú informácie, ktoré obsahujú nepravdivé informácie, ale nie je možné dokázať, že sú šírené so zámerom zavádzať. Ide napríklad o situácie, kedy si autor nie je vedomý toho, že informácia je nepravdivá a šíri ju s presvedčením o jej pravdivosti.

Deceptive news, ktoré je možno preložiť ako klamlivé informácie a false news vo význame nepravdivých noviniek, sú obdobné dezinformáciám a misinformáciám, s rozdielom, že u obsahu je možné určiť charakter noviniek. Satira a clickbait môžu obsahovať ako pravdivé tak nepravdivé informácie, dôležitý je zámer, s ktorým sú šírené. Satirický príspevok nemá výhradný cieľ informovať o fakticky správnych udalostiach či tvrdeniach, preto nemusí napĺňať charakter autenticity. Dôvod, pre ktorý satira vzniká je primárne pobaviť adresáta.

Clickabait, podobne ako dezinformácia alebo deceptive news, vykazuje známky úmyselnej manipulácie. Clickbait je príspevok obsahujúci pútavý titulok, ktorý je zámerne napísaný tak, aby prinútil adresáta kliknúť na odkaz, alebo príspevok. V niektorých prípadoch nemusí text titulku ani súvisieť s obsahom príspevku, na ktorý odkazuje. Cieľom je získať viac vzhľadnutí a v prípade marketingového zámeru je výsledkom aj viac príjmov z reklám pre médium využívajúce clickbait. Využívaný je ako dezinformačnými médiami, tak mainstreamovými médiami, ktoré tak súperia o predné miesto vo výsledkoch vyhľadávania internetových prehľadávačov. Dezinformačné média využívajú clickbaity k tomu, aby vďaka vysokému dosahu boli nepravdivé informácie, ktoré šíria, považované verejnosťou za pravdivé. (Sousa-Silva 2022)

2.2 Vymedzenie pojmu fake news

Napriek tomu, že definícia fake news nie je v súčasnosti pevne ustanovená, existujú dve hlavné typy definícií, ktoré zahŕňajú rôzne koncepty popísané v predchádzajúcej kapitole. Toto rozličné vymedzenie závisí od miery detailu danej definície.

Niektoré zdroje definujú fake news ako každú preukázateľne nepravdivú správu. (Shu et al. 2017) Táto definícia fake news je založená na autenticite danej správy. Pod túto definíciu spadajú nielen dezinformácie, misinformácie, deceptive alebo false news, ale tiež napríklad satirické príspevky a to napriek tomu, že klamlivý obsah využívajú za účelom pobaviť a táto skutočnosť je každému známa. Vo vedeckých článkoch je táto definícia označovaná ako široká, keďže nekladie žiadne nároky na overovanie zámeru, ktoré stoja za vznikom fake news, a je zameraná len na overenie pravdivosti danej správy, vďaka čomu zahrňuje rôzne typy nepravdivých informácií. Dôvodom, prečo je táto zjednodušená definícia používaná, je náročnosť procesu overenia skutočného zámeru. Výskumy zamerané na fake news tiež často pracujú s voľne dostupnými datasetmi, ktoré tvoria tvrdenia rozdelené na fake news a pravdu. (Viac o manuálnom overovaní fake news a výskumných datasetoch, ako technike k detekcii fake news, pojednáva kapitola **Error! Reference source not found. Error! Reference source not found.**). Zatiaľ čo datasety vznikajú na základe výstupov z analýz hodnotenia pravdivosti tvrdení obsiahnutých v analyzovaných správach, skutočný zámer je možné vyvodiť častokrát len na základe vlastného logického zdôvodnenia a presvedčenia.

O náročnosti v kontexte určenia zámeru za šírením fake news pojednáva tiež (Golbeck et al. 2018) pri vymedzení satiry v rámci fake news. Priestor medzi fake news a skutočnou satirou popisuje ako spektrum, v ktorom existuje mnoho webových stránok šíriacich nepravdivý a zavádzajúci obsah, ktoré svoju aktivitu ospravedlňujú prehlásením o vylúčení zodpovednosti. Takéto webové stránky sa zbavujú akejkoľvek zodpovednosti za publikovaný obsah a jeho pravdivosť, čo je v prípade satirických stránok v poriadku a patrí medzi spôsob ujasnenia povahy stránky, v prípade stránok publikujúcich fake news ide ale len o spôsob, ako sa ubrániť prípadnej kritike k publikovanému klamlivému obsahu.

Druhou široko využívanou definíciou je tzv. úzka definícia. Tá vymedzuje pojem fake news konkrétnejším spôsobom, kedy je okrem autenticity kladený dôraz aj na to, či je správa nepravdivá zámerne, a je šírená s cieľom zavádzať. (Shu et al. 2017) definuje v tomto kontexte fake news, ako články (news), ktoré sú zámerne a overiteľne nepravdivé a môžu zavádzať čitateľa.

Obe definície spája požiadavka na overenie autenticity správy. Nie je správne porovnávať autenticitu správ voči osobným presvedčeniam či postojom. Autenticitu danej správy je možné určiť len voči objektívnym faktom. Preto správy, ktoré nesúhlasia s presvedčením osoby, alebo skupiny osôb, nie je možné označovať ako fake news. Takéto označovanie fake news sa ale vyskytuje v politike, a to do takej miery, že je dnes bežne používané za účelom odsúdenia a znevažovania akýchkoľvek skutočností, ktoré nie sú pohodlné alebo príjemné danej politickej strane alebo politikovi, a slúžia tak k ich potlačeniu. Napríklad bývalý americký rezident Donald Trump označoval overené informácie z mainstreamových médií ako fake news (Fridlund 2020).

Niektoré definície fake news zužujú množinu obsiahnutých konceptov obmedzením sa na príspevky obsahujúce príznak novosti (tzn. news), tak, ako bolo popísané v predchádzajúcej podkapitole. (Lazer et al. 2018) definuje fake news ako zostrojené informácie, ktoré napodobňujú obsah spravodajských médií formou, ale nie organizačným procesom, alebo zámerom. Súčasne definuje fake news médiá, ako médiá, ktoré nesplňujú žurnalistické a editorské normy a procesy, ktoré zaručujú správnosť a dôveryhodnosť informácií. Podľa (Khan et al. 2021) môže mať ale vymedzenie sa pri definícii fake news a jeho skúmaní len na príspevky napodobňujúce svojou formou seriózne spravodajstvo za následok nižšiu relevanciu dosahovaných výsledkov. To je spôsobené tým, že táto prísna definícia fake news, keď ide o formu skúmaných príspevkov, nereflektuje všetky formy digitálneho obsahu, v ktorých sa môžu fake news vyskytovať a šíriť. V prostredí sociálnych sietí a nonstop prílivu informácií 24 hodín denne je požadovanie ideálneho spravodajského štandardu nerealistické. Avšak na druhú stranu v prostredí, kde je vytváranie a šírenie obsahu dostupné každému, sú práve tieto štandardy dôležitým faktorom pre zachovanie informovanej spoločnosti s vierou v médiá. (Molina et al. 2019)

2.3 Šírenie fake news

Zatiaľ čo za vznikom fake news stojí na úplnom počiatku stále človek, ktorý fake news vytvoril a má záujem šíriť ju za istým účelom, medzi širitel'ov fake news môžu patriť ako reálni ľudia, programy, ale tiež ich kombinácia vo forme programov, ktoré vyžadujú ľudský zásah. Práve automatizované programy stoja za masívnym šírením fake news v prostredí internetu. Výskumy ukazujú, že automatizované programy ako sociálny bot hrajú dôležitú úlohu, keď ide o šírenie príspevkov zo zdrojov s nízkou reputáciou. Ako zdroje s nízkou reputáciou označuje webové stránky, ktoré boli nezávislými organizáciami označené ako zdroje, ktoré často zdieľajú informácie s nízkou kredibilitou. Automatizované programy zosilňujú dosah príspevkov v ranných fázach ich šírenia po sociálnych sieťach. (Shao et al. 2018)

2.3.1 Automatizované programy na šírenie fake news

Sociálny bot je program, alebo algoritmus, ktorý je naprogramovaný tak, aby autonómne vystupoval na sociálnych sieťach. Sociálny bot dokáže automaticky vytvárať obsah, šíriť ho a interagovať s užívateľmi na internete. Sociálny bot býva využívaný za marketingovým alebo politickým účelom. Boty prehľadávajú sociálne siete nástrojmi web crawlingu a pomocou API reagujú na sociálnych sieťach ako každý iný užívateľ. Pomocou nástrojov umelej inteligencie, analýzy dat a textu, technikám deepfake, dokážu generovať nové komentáre v reakcii na aktuálne udalosti.

Siete botov, reagujúcich vzájomne medzi sebou, s cieľom efektívne zahltiť informačný priestor sa nazýva botnet. Kyborg je hybridný užívateľský účet, spájajúci výhody bota, ako možnosť masového zdieľania správ a automatických reakcií, a výhody človeka, vďaka ktorým je kyborg ťažšie rozpoznateľný nástrojmi na detekciu botov a súčasne pôsobí vierohodnejšie voči ostatným skutočným užívateľom na internete. Kyborg vykazuje ľudské charakteristiky a to vtedy, keď človek preberie z času na čas kontrolu nad kyborgom

a interaguje namiesto neho. Zvyšok času funguje ako bot. Niekedy je ľudský zásah u kyborga obmedzený len na založenie užívateľského konta. (Zhang a Ghorbani 2020)

Sociálny bot a kyborg patria medzi automatizované spôsoby, ako je možné šíriť fake news. V kontexte fake news je ich cieľom presvedčiť ostatných užívateľov o svojej legitimitate a autenticite správ, ktoré nesú. Ak sú v plnení svojho účelu úspešní, k šíriteľom fake news sa pridávajú aj skutoční ľudia, ktorí sa tak vedome alebo nevedome stávajú šíriteľmi fake news. Tí obsah pridávajú a zdieľajú v rámci svojich sociálnych skupín, kde ovplyvňujú svojich blízkych, kamarátov, followerov, členov spoločných skupín na sociálnych sieťach alebo četovacích platformách a vytvárajú tak ďalšie generácie šíriteľov fake news.

2.3.2 Algoritmy vytvárajúce obsah

Medzi algoritmy, pre ktoré sú skutoční ľudia vystavovaní masívnemu tlaku fake news a pre ktoré sú nachylní im uveriť a šíriť je ďalej, patria aj algoritmy sociálnych sietí a ďalších online nástrojov na vyhľadávanie informácií. Aj keď nie sú zámerne vytvorené s úmyslom šíriť fake news, ako je to časté v prípade sociálnych botov či kyborgov, spôsob, akým fungujú pomáha vytvárať prostredie pre ich jednoduchšie šírenie. Existuje obava, že tieto algoritmy filtrujú informácie zobrazované užívateľom podľa individuálnych preferencií, čím je vytvárané názorovo homogénne prostredie. Vďaka personalizovaným výsledkom vyhľadávania a obsahu na sociálnych sieťach sú redukované zdroje informácií, ktoré užívatelia konzumujú a s ktorými interagujú a každý užívateľ je vystavený rozličnému obsahu. (Kitchens et al. 2020) Tento efekt popisujú 2 fenomény známe ako komnata ozvien (anglicky *echo chamber*) a filtračné bubliny (anglicky *filter bubbles*), ktoré pri šírení fake news pôsobia súčasne.

Komnatu ozvien popisujú sociológovia ako ohraničený, uzavretý mediálny priestor, ktorý má potenciál zväzňovať správy v ňom prezentované a zároveň izolovať ich pred možným spochybnením ich pravdivosti. Pozostáva z dvoch fáz. Fáza zosilňovania súvisí s prevahou informácií, ktoré sú konzistentné s našimi už existujúcimi postojmi a v izolovanom prostredí dochádza k ich utvrdzovaniu a zosilňovaniu. Izolačná časť hovorí o styku s opozičnými, či odlišnými názormi, ktoré sú postupne zoslabované a môže dochádzať až k ich úplnému vylúčeniu. Tento efekt vypovedá o tendencii ľudí veriť tomu, čo je v súlade s ich už existujúcimi názormi či presvedčeniami a v averzii k názorom odlišným. (Jamieson a Cappella 2010)

Filtračné bubliny sú považované za vedľajší negatívny efekt personalizácie obsahu na sociálnych sieťach a internetových vyhľadávačoch. Ide tak o fenomén komnaty ozvien vytváraný v prostredí sociálnych sietí, a to ich vlastnými algoritmami. Personalizovaný obsah je prispôbený užívateľovi s ohľadom na jeho preferencie, informačným a kognitívnym schopnostiam, alebo na základe povahy riešeného problému. Personalizácia obsahu má výhody, keď ide o oslabenie informačnej záťaže, urýchlenie času vynaloženého na riešenie problému, či vyhľadávanie požadovaného obsahu, a to vďaka vyššej relevancii zobrazovaných príspevkov či výsledkov vyhľadávania. To má avšak za dôsledok efekt známy ako filtračné bubliny, v rámci ktorých sme ako užívatelia v styku s názormi, ideami, myšlienkami, ktoré sú blízke tým našim. (Hrčková a Mirga 2018) Algoritmy na sociálnych sieťach spoznávajú na základe užívateľských interakcií (reakcie, zdieľanie, komentáre)

populárny a preferovaný obsah, ktorý potom so zámerom zvýšenia času, ktorý užívateľ na sociálnych sieťach strávi je pretláčaný a zobrazovaný častejšie. (Rhodes 2021)

Aj keď oba fenomény popisujú informačné bubliny, v ktorých žijeme, je dôležité zvýrazniť rozdiely, medzi nimi. Aj keď termín komnata ozvien predstavuje bublinu, neurčuje dôvod, pre ktorý ľudia v informačných bublinách existujú. Existuje možnosť, že niekto sa vlastným rozhodnutím v takejto bubline rozhodne žiť. Filtračné bubliny naopak predstavujú komnaty ozvien vytvorené hodnotiacimi algoritmi, ktoré bez nášho aktívneho rozhodnutia personalizujú obsah. Preto rozdiel závisí najmä na príčinách prosperujúcich k ich vzniku, a to na jednej strane vo forme aktívneho užívateľského rozhodnutia a na druhej strane algoritmom filtrujúcim obsah, ku ktorým užívatelia pristupujú, bez ich vlastného rozhodnutia, či možnosti výrazného prispôsobenia. (Arguedas et al. 2022)

2.4 Atribúty fake news

Overovanie pravdivosti tvrdení, obsiahnutých v samotnom príspevku, je kľúčovou technikou pri detekcii fake news. To avšak platí najmä v prípade tradičných médií, ktoré publikujú články v tlačovej podobe. V dobe online médií a sociálnych sietí, obsahujú príspevky okrem tvrdení obsiahnutých v texte aj ďalšie atribúty, na základe ktorých je možné odlíšiť štandardný a fake news príspevok.

Overovanie tvrdení je časovo a znalostne náročný proces, ktorý je v prípade masívneho šírenia fake news v online prostredí, nutné doplniť alebo nahradiť ďalšími technikami, pracujúcimi s inými dátami. Malo by ísť o data, ktoré je o príspevku možné rýchlejšie získať a následne aj rýchlejšie vyhodnotiť. V tejto podkapitole sú popísané základné prvky fake news šírených v online priestore, vychádzajúce z prvkov digitálnej správy, ktoré môžu byť pri úlohe detekcie fake news určujúce.

2.4.1 Obsah fake news

Hlavným prvkom fake news je obsah príspevku. Ide o časť príspevku, obsahujúcu informácie, ktoré je možné klasifikovať (na základe autenticity, zámeru) ako fake news. Rovnako ako v prípade štandardných správ na spravodajských webových stránkach, zahŕňa obsah fake news, ktoré imitujú news, nadpis článku, telo správy obsahujúce hlavný text a zahrnuté multimédia.

(Khan et al. 2021) popisuje všetky falošné správy v kontexte digitálneho objektu, s vlastným obsahom a štruktúrou. Obsah správy (anglicky *message content*) pozostáva z informácií, ktoré tvrdenia v danej správe obsahujú a ktoré je možné hodnotiť na základe autenticity a zámeru. Štruktúra správy (anglicky *message structure*) hovorí o spôsobe, v ktorom sú digitálne objekty organizované. Jeden obsah, falošná informácia, môže byť šírená vo viacerých štruktúrach napríklad vo forme textového súboru, statusu na užívateľskom profile, vo forme videí a podobne. Táto definícia obsahu patrí tak medzi obecnějšíe pojatie toho, čo spadá do obsahu fake news a na rozdiel od vyššie spomínanej definície pracuje s rôznorodosťou fake news. Zahrňuje tak aj obsah fake news šírených vo forme videí,

statusov, alebo napríklad krátkych videí typických pre sociálne siete ako Instagram, alebo Tik Tok.

(Zhang a Ghorbani 2020) delí obsah fake news na fyzický a nefyzický obsah. Medzi fyzický obsah patrí každý objekt správy, ako napríklad URL, text, emotikony, multimedialne súbory a podobne. Nefyzický obsah správy tvoria názory, informácie, postoje, ktoré má šíriteľ správy v úmysle prezentovať. Medzi tento obsah patrí tiež emocionálny náboj správy, alebo obsiahnuté témy a motívy. Keďže príspevky fake news nemajú za cieľ prinášať objektívne fakty, ale sú tvorené z politických, finančných alebo osobných motívov, obsahujú často citovo zafarbený obsah, ktorý pri použití v rámci analýzy sentimentu môže pomôcť pri detekcii fake news.

2.4.2 Šíriteľ

Pojem šíriteľ fake news môžeme rozumieť ako obecný pojem zastrešujúci zdroje šírenia fake news. Pre vysvetlenie toho, čo alebo kto všetko pod tento pojem spadá, je dobré pozrieť sa na zdroje šírenia v kontexte fáz životného cyklu fake news.

Vo fáze vzniku fake news je šíriteľom fake news jeho autor, ktorý stojí za vznikom fake news príspevku. Ide o tvorca fake news. Vo fáze zverejnenia ide o zdroj publikujúci fake news. Môže ísť o stránku na sociálnej sieti, alebo napríklad o webovú stránku, ktorá zverejňuje fake news. V prípade autora, ktorý publikuje fake news v rámci svojho profilu, alebo v prípade fake news bez uvedeného autora, je šíriteľ vo fázach vzniku a publikácie rovnaký. V poslednej fáze životného cyklu sa fake news šíri internetovým priestorom. Šíriteľmi v tejto fáze sú zdroje, ktoré distribuujú fake news ku koncovým príjemcom napríklad na sociálnych sieťach. Patria sem užívateľské účty a to ako tie, ktoré patria reálnym ľuďom, tak sociálne boty (Zhou a Zafarani 2020).

2.4.3 Sociálny kontext

Fake news existuje v online prostredí, v ktorom sa dostáva do kontaktu so svojím okolím. Sociálny kontext predstavujú informácie popisujúce existenciu fake news v tomto okolí. V tomto kontexte sú fake news identifikované na základe dát o užívateľských interakciách a informácie o šírení fake news v online priestore. Sociálne siete a médiá dnes ponúkajú rozšírené možnosti kontaktu s digitálnym obsahom a menia spôsob, akým užívatelia konzumujú obsah.

V dnešnom svete nie je užívateľ obmedzený pri kontakte s digitálnym obsahom výhradne na jeho čítanie. Prostriedky sociálnych sietí a médií umožňujú užívateľom reagovať s príspevkom rozličným spôsobom a to napríklad formou pridávania reakcií (v prípade sociálnej siete Facebook užívatelia majú možnosť používať reakcie *páči sa mi to*, alebo *hnevá ma to*), zdieľaním príspevkov v rámci rozličných skupín, pridávaním komentárov, nahlasovaním a blokovaním príspevkov. Spoločne s aktivitou automatizovaných programov vznikajú tak vzorce šírenia fake news, ktoré popisujú data o rýchlosti a dosahu šírenia a ktoré patria pri detekcii fake news medzi využívané charakteristiky. (Zhang a Ghorbani 2020)

2.5 Vlastnosti fake news

Fake news v prostredí sociálnych sieti zaznamenali masívny rast. Šíria sa veľkou rýchlosťou, majú vysoký dosah a neustále vznikajú nové. Sú všadeprítomné a majú negatívny dopad na jedincov a tiež spoločnosť. Vlastnosti fake news, tak, ako ich je potrebné vnímať dnes, vysvetľuje (Zhang a Ghorbani 2020) pomocou troch „V“. Tieto vlastnosti je potrebné zvažovať pri vytváraní účinných nástrojov na ich detekciu a potlačanie.

- **Volume (objem fake news)**

Nízka náročnosť, keď ide o zakladanie profilov na sociálnych sieťach, alebo webových stránok, zapríčinila, že existuje vysoké množstvo zdrojov vytvárajúcich fake news. Napríklad ako reakciu na Ruskú agresiu voči Ukrajine, Európska únia už pár dní od vniknutia ruských vojsk na územie Ukrajiny pozastavila aktivity spravodajských médií Russia Today (známa ako „RT“) a Sputnik, pôsobiacich v EÚ. (Council of the EU 2022) Ruské médiá si ale našli cestu k európskemu publiku. Výskumná spoločnosť IDS analyzovala spôsoby, ktorými bol prístup k obsahu na blokovaných stránkach naďalej zabezpečený. Vznikli webové stránky kopírujúce obsah materskej stránky (tzv. *mirror pages*), stránky, ktoré obsahujú redirect na RT a variácie domén RT, tiež s obsahom redirectu. (Balint et al. 2022) Vďaka vysokému objemu fake news nie je čisto manuálne overovanie fake news v súčasnosti postačujúce a detekcia fake news si vyžaduje automatizované spôsoby pre včasnú detekciu, ktorá obmedzí ich masívne rozšíreniu medzi užívateľmi.

- **Variety (rozmanitosť fake news)**

Čiastočné pravdy, zavádzajúce vyjadrenia vo forme clickbaitov, politických vyjadrení, alebo napríklad humorných obrázkov (známe ako *memes*). Definícia, ani forma, v ktorej sa fake news šíri, nie je pevne určená a môže nadobúdať mnohé podoby. Obmedzovať fake news len na príspevky obsahujúce vyvrátené tvrdenia a imitujúce novinové správy (news), nie je preto pre efektívny boj s fake news postačujúci.

- **Velocity (rýchlosť fake news)**

Šíritelia fake news vedia o zúčujmoch a technikách vyvíjaných za účelom potlačiť ich aktivitu a snažia sa im účinne vyhýbať. Jedným zo spôsobov, ktorým znemožňujú snahy o detekciu fake news, je krátke obdobie, v ktorom sú šíritelia fake news aktívny. Podľa (Allcott a Gentzkow 2017), významné webové stránky aktívne v šírení fake news počas predvolebnej kampane k americkým prezidentským voľbám v roku 2016, nemali dlhú životnosť a mnohé prestali existovať už krátko po voľbách. Súčasne rýchlosť s ktorou fake news vznikajú a ich real-time povaha robí proces identifikácie fake news neľahký. Je tiež preto náročné určiť dosah fake news, keď ide o počet zasiahnutých užívateľov a odhadnúť negatívne vplyvy, ktoré fake news na užívateľov majú, počas svojej existencie. (Zhang a Ghorbani 2020)

Na vlastnosti troch „V“ nadväzuje (Manish et al. 2022), keď v rámci hĺbkového review techník pre detekciu fake news v svojej práci rozširujú vlastnosti o ďalšie dve „V“ a to nasledovne:

- **Veracity (vierohodnosť fake news)**

Fake news sú zámerne tvorené tak, aby čo najvierohodnejšie pôsobili na cieľových užívateľov. Napríklad existuje veľa webových stránok, ktoré sú názvami podobné serióznym médiám. V slovenskom prostredí je možné na základe nezávislej iniciatívy hodnotiacej vierohodnosť stránok identifikovať webové stránky šíriace fake news s názvom zámerne podobným s inými známymi médiami, ako napríklad sknews.sk (podoba s médiom skynews.com), fresher.sk (podoba s refresher.sk) alebo napríklad aktuality24.sk (podoba s aktualne.sk). (Konspiratori 2022)

- **Valid dataset (validný dataset)**

Vďaka reštrikciám sociálnych sietí v oblasti získavania dát neexistujú trénovacie data, ktoré by obsahovali dostatok informácií o fake news správe, a pomocou ktorých by vznikali modely pre detekciu fake news s vysokou presnosťou.

Týmito skutočnosťami autor rozširuje predchádzajúce 4 vlastnosti fake news a díva sa na ne nie len ako vlastnosti, ale výzvy, ktoré úloha detekcie fake news so sebou prináša.

3 Prístupy detekcie fake news

Určenie fake news je v súčasnosti komplexnou úlohou, ktorá sa musí vysporiadať s rovnako komplexnými snahami o ich masívne šírenie. Existuje dnes preto mnoho spôsobov, ako je možné k tejto úlohe pristupovať. Táto kapitola je zameraná na popis prístupov využívaných pri úlohe detekcie falošných správ a konkrétnych techník, ktoré možno pre implementáciu týchto prístupov využiť. Popísané sú kľúčové vlastnosti najčastejšie využívaných prístupov, ich základné princípy a techniky, ako aj ich prínosy a obmedzenia. Táto časť práce tak slúži k ucelenému pochopeniu spôsobov, ktorými je možné pri odhaľovaní fake news postupovať, s prihliadnutím na ich vlastnosti.

Základné prístupy detekcie fake news boli rozdelené do troch kategórií. Tieto kategórie sú odvodené na základe prvkov fake news, tak ako boli popísané v predchádzajúcej kapitole. Uvedené prístupy analyzujú a využívajú ich k identifikácii fake news medzi skutočnými správami. Prvou skupinou sú prístupy založené na obsahu príspevku, druhou skupinou sú prístupy identifikujúce fake news na základe šíriteľov či reakcií a treťou skupinou sú prístupy založené na analýze šírenia. Prístupy vychádzajú z dôvodu kompletného pokrytia rozličných prístupov primárne z prehľadovej štúdie (Zhou a Zafarani 2020).

3.1 Detekcia na základe obsahu fake news

Detekcia na základe obsahu fake news zahŕňa viaceré prístupy. Medzi ne patrí manuálne overovanie faktov a aktivity fack-checkingových organizácií a automatizované spôsoby využívajúce k detekcii pravdivosť obsahu fake news. Posledným prístupom je detekcia nie na základe pravdivosti obsahu príspevku, ale na základe jeho štruktúry a štýlu.

3.1.1 Overovanie faktov

Overovanie faktov znamená porovnávanie zdelenia správy voči ostatným zdrojom a faktom, na základe čoho je overená autenticita danej správy. Takéto overenie autenticity správy môže byť investigatívne náročnou úlohou, vďaka čomu existuje viacero organizácií združujúcich analytikov a expertov, ktorí vykonávajú analýzy na určenie autenticity správ. Výsledky publikujú napríklad vo forme investigatívnych článkov, kde prezentujú svoje závery.

Keď ide o overovanie politicky orientovaných faktov so zameraním na prokremel'skú propagandu je možné ako príklad uviesť napríklad projekt EUvsDisinfo. Použitím techník dátovej analýzy a monitoringu médií v 15 jazykoch, tento projekt Európskej služby pre vonkajšiu činnosť (European External Action Service) vytvoril verejne dostupnú databázu s viac ako 12.000 príkladmi fake news, ktorá je neustále aktualizovaná. V rámci projektu sú tiež publikované analýzy zaoberajúce sa vývojom dezinformácií. (EUvsDisinfo 2023) Ďalším príkladom je nezisková organizácia VoxUkraine a projekt VoxCheck. Databáza VoxCheck obsahuje viac ako 9.000 tvrdení politikov a takmer 2000 vyvrátených fake news. (VoxCheck

2023) Existujú tiež kódexy zásad, v rámci organizácií, alebo medzinárodne, ako napríklad IFCN, ktorými sa analytici zaväzujú k nestrannému a zodpovednému prístupu k overovaniu faktov. (IFCN Code of Principles 2023) Signatári kódexu IFCN sa spojili v projekte #UkraineFacts, kde vytvorili databázu s takmer 3.000 fake news, týkajúcich sa dezinformácií o Ukrajine, predovšetkým v kontexte ruskej agresie na Ukrajine. (UkraineFacts 2023).

Nie je možné, aby manuálne overovanie faktov pokrylo všetky prípady fake-news. Manuálne overovanie faktov je preto nutné pri objeme, rozmanitosti a rýchlosti šírenia fake news, tak ako bolo popísané v kapitole 2.5 *Vlastnosti fake news*, doplniť automatizovanými prístupmi k detekcii fake news. Súčasne ale hraje manuálne overovanie faktov dnes stále dôležitú úlohu pri vývoji efektívnych nástrojov na detekciu fake news. Základným a jedným z prvých krokov pre vývoj modelov využívajúcich metód strojového učenia je zber trénovacieho datasetu (supervised learning), ktorý obsahuje vstupné objekty a očakávané výstupné hodnoty.

Medzi podobné príklady na Slovensku je možné zaradiť napríklad AFP fakty, InfoSecurity, alebo činnosti Investigatívneho centra Jána Kuciaka, či Adapt Institute, publikujúce analýzy k dezinformáciám pojednávajúcich nielen o udalostiach na Ukrajine. Aktívne sledujú činnosti na slovenských sieťach (predovšetkým Facebook a Telegram) a odhaľujú kľúčové falošné naratívy a ich šíriteľov vyskytujúcich sa na slovenskej dezinformačnej scéne.

3.1.2 Automatické overovanie faktov

Automatické overovania faktov je založené na použití techník spracovania prirodzeného text (natural language procesing, NLP), strojového učenia (machine learning, ML), extrakcií informácií (information retrieval, IR) a pod. Existujú dve hlavné fázy automatického overenia faktov a to extrakcia faktov a overovanie. (Zhou a Zafarani 2020)

Automatická extrakcia faktov je založená na zbere autentických informácií v štruktúrovanej podobe a vytvorení znalostnej bázy. Informácie sú získavané zo zdrojov obsahujúcich prevažne pravdivé informácie ako napr. *Wikipedia*. Znalostné bázy sú zložené z faktov z jedného zdroja, alebo vznikajú kombináciou viacerých zdrojov na webe. Získané znalostné bázy je nutné ďalej spracovať pre zaručenie vysokej kvality obsahujúcich faktov a to prostredníctvom odstránenia redundancií, protichodných faktov, či obmedzením zdrojových stránok (napríklad vylúčením satirických stránok). Samotné overovanie faktov potom prebieha ako úloha určenia autenticity skúmanej správy na základe porovnania s faktami uloženými v znalostnej báze. (Zhou a Zafarani 2020)

Správnosť automatického overovanie faktov je vysoko závislá od kvality vytvorenej znalostnej bázy. Tá závisí od zdrojov na webe, ktoré sú použité k jej vytvoreniu. Fakty sa na webe vyskytujú v neštruktúrovanej a neúplnej podobe, nehľadiac tiež na kontext v ktorom musia byť zasadené, aby boli pravdivé. Automatické overovanie faktov na základe extrakcie faktov z webu je preto neľahkou úlohou a vyžaduje si zapojenie znalosti živých expertov pri vytváraní znalostnej bázy. Pri úlohe detekcie fake news ide často tiež o nové tvrdenia, resp. nové fake news, ktoré musia byť detekované čo najrýchlejšie. V prípade tvorby znalostných báz by to znamenalo neustále rozširovanie uložených faktov (tzv. dynamické znalostné bázy)

a pre nové fake news nemusí dostatočne rýchlo existovať odpovedajúci pravdivý fakt, ktorý by vyvrátil jeho autenticitu.

3.1.3 Detekcia na základe štruktúry a štýlu

Fake príspevky bývajú písané tak, aby zaujali čo najväčší počet čitateľov a pôsobili čo najviac presvedčivo. Detekcia fake news na základe štýlu tak predpokladá, že za týmto účelom využívajú širitelia fake news odlišné štylistické prostriedky textu, v porovnaní s pravdivými príspevkami. Predpoklad vychádza z tzv. *undetutsch hypotézy* vypovedajúcej o tom, že vyjadrenia popisujúce okolnosti založené na skutočných zážitkoch sa významne líšia v obsahu a kvalite od vymyslených a fiktívnych udalostí (Amado et al. 2015). Rovnako ako v prípade overovania faktov sa jedná o prístup založený na analýze obsah príspevku, aj keď nejde o overovanie autenticity danej správy. Podľa (Zhou a Zafarani 2020) môže tento prístup k detekcii fake news odhaliť zámer príspevku, ktorý je dôležitým aspektom pri rozlíšení fake news napríklad od misinformácií, ktoré môžu niesť klamlivú informáciu z dôvodu chyby a ich zámerom nie je úmyselne zavádzať čitateľa.

Štýl fake news definuje (Zhou a Zafarani 2020) ako súbor kvantifikovateľných príznakov (napríklad machine learning features), na základe ktorých je možné dobre reprezentovať obsah fake news a rozlíšiť ich na ich základe od pravdivých príspevkov.

Štúdie zamerané na detekciu fake news na základe štýlu napríklad ukázali, že fake news príspevky v porovnaní s pravdivými príspevkami obsahujú nižšie množstvo unikátnych slov, prídavných mien a prísloviek, zatiaľ čo množstvo unikátnych slovies obsiahnutých v texte je v prípade fake news vyššie. Vyššie je tiež počet negatívne či pozitívne zafarbených slov. Navyše obsahujú fake news príspevky častejšie kratšie slová a často obsahujú viac slov v nadpise v porovnaní s pravdivými príspevkami. Počet slov obsiahnutý priamo v príspevku je však naopak nižší (Zhou et al. 2019).

Príznyky používané na detekciu fake news, popisujúce štruktúru a štýlu fake news, spadajú do štyroch kategórií (Zhou et al. 2019):

- **Frekvencia slov**

Príznyky týkajúce sa frekvencie jednotlivých slov a analyzovanom texte. Technikou používanom pre získanie príznakov o frekvencii slov je model BOW (bag-of-word), ktorý reprezentuje frekvenciu slov v absolútnom počte, alebo v štandardizovanej či relatívnej forme, kedy je odstránený vplyv celkovej dĺžky textu na frekvenciu slova v texte (je pravdepodobné, že v obsiahlom príspevku stúpa frekvencia určitých používaných slov).

- **Syntax**

Príznyky založené na vetnej stavbe viet. Patrí sem určenie frekvenciu jednotlivých slovných druhov pomocou POS (part-of-speech), alebo použitie PCFG (probabilistic context-free grammar) pre rozloženie vety do hierarchických stromov, pravidiel, ktoré popisujú syntaktickú štruktúru textu (Shu et al. 2017) .

- **Sémantika**

Príznyky vypovedajúce o význame slov v texte. Skúmajú sa vlastnosti ako napríklad:

- Senzacionalizmus: fake news príspevky, podobne ako click-baity, používajú prehnané a citovo zafarbené slova, aby pritiahli pozornosť čitateľov. Pri analýze senzacionalizmu sa skúma použitie citovo zafarbených slov v nadpise článku, alebo použitie interpunkčných znamienok ako výkričník alebo otáznik.
- Kvalita: nakoľko u seriózných príspevkov by mal nadpis článku obsahovať zhrnutie udalosti, či skutočnosti popísanej v texte, podobnosť nadpisu a hlavným textom by mala byť vysoká. Uvedené často neplatí o clickbaitov.
- Formálnosť: pracuje sa s predpokladom, že čím vyššia je kvalita príspevku, tým vyššie je použitie formálneho jazykového štýlu v texte. Skúma sa použitie nespisovných a neslušných slov, citoslovci a skratiek, alebo chyby vyskytujúce sa v texte.
- Diverzita: analyzuje vyjadrovacie a jazykové schopnosti autora príspevku, vyjadrenú ako množstvo unikátnych slov a použitých slovných druhov.
- Objektivita: skúma podiel zaujatých slovies („myslím si“), alebo naopak počet faktických slovies („pozorovať“)

- **Diskurz**

Príznamy popisujúce rôzne typy rétorických vzťahov medzi vetami v texte. Rétorické vzťahy označujú spôsoby, akými sa vety v texte navzájom dopĺňajú, pokiaľ ide o ich rétorickú funkciu alebo účel. Určitá veta môže napríklad poskytovať vysvetlenie predchádzajúcej vety, ponúknuť dôkazy na podporu tvrdenia v predchádzajúcej vete, alebo vytvoriť kontrast k predchádzajúcej vete. Model RST identifikuje typy týchto rétorických vzťahov medzi vetami s určením relatívnej frekvencie každého typu vzťahu. Porovnaním frekvencií rôznych typov vzťahov medzi vetami obsiahnutých vo fake news možno identifikovať vzory, na základe ktorých je následne možné fake news detekovať.

K vyššie popísaných príznakom, existujú tiež takzvané latentné, či skryté, príznaky. Latentné príznaky vznikajú použitím techník technikou word embedding, ktorú je možné implementovať napríklad prostredníctvom word2vec, alebo doc2vec. Vznikajú pritom vektory reprezentujúce slová, vety, či dokumenty a tie slúžia ďalej ako vstup pre modely pre detekciu fake news. Zachované sú tak vzťahy medzi vetami. Tieto sú potom doplnené do deep learningu modelov (CNN, RNN). Aj keď sa latentné príznaky ukazujú ako efektívne pre klasifikáciu fake news, sú náročné na získanie a nie je možné jasne vysvetliť, akým spôsobom a na základe čoho bola klasifikácia fake news vykonaná.

Úspešnosť detekcie fake news príspevkov na základe štýlu závisí od toho, ako dobre je možné určiť a reprezentovať text fake news prostredníctvom príznakov a ako je zvolená reprezentácia textu príspevku efektívna pri využití v rozličných modeloch pre detekciu fake news.

3.2 Detekcia na základe širitela

V kontexte detekcie fake news na základe širitela je často spomínaný pojem kredibilita. Kredibilita alebo tiež dôveryhodnosť, je definovaná ako vnímaná kvalita na základe viacerých dimenzií, ako napríklad uveriteľnosť (Castillo et al. 2011). Pri veľkom

zjednodušení úlohy detekcie fake news na základe širiteľa by bolo možné povedať, že všetky príspevky pochádzajúce z nedôveryhodných webových stránok sú fake news. V tomto kontexte stojí za zmienku iniciatíva konspiratori.sk, v rámci ktorej je zverejnený zoznam stránok s klamlivým, zavádzajúcim či propagandistickým obsahom. Zaradenie webovej stránky na zoznam prebieha na základe posúdenia hodnotiacej komisie a ku každej stránke je zverejnená analýza obsahu stránky a prípadne aktivít autorov jej obsahu s odôvodnením zaradenia na zoznam (Konspiratori 2022).

Kredibilitou širiteľa, v texte tiež nazývané ako zdroj, sa zoberá štúdia (Sitaula et al. 2019). Rozlíšenie fake news od pravdivých príspevkov tak určuje na základe počtu autorov príspevku, vzťahu autorov pri spolupráci, a na základe histórie publikovania fake news. Pri analýze počtu autorov bolo zistené, že existuje korelácia medzi kredibilitou príspevku a počtom autorov. Ak nie je určený autor príspevku, je viac pravdepodobné, že sa jedná o fake news a naopak, ak príspevok obsahuje dvoch a viac autorov, je menej pravdepodobné, že sa jedná o fake news príspevok. Pri detekcie fake news na základe existencie autora je ale nutné mať na pamäti, že zakladanie účtov, alebo priradenie vymysleného mena k fake news nie je náročnou úlohou a pri pokročilej detekcii fake news by bolo práve priradenie autora jedným z prvých opatrení, ktoré by širitelia fake news mohli aplikovať, aby sa vyhli snahám o svoje odhalenie.

Kredibilitu širiteľa je možné určiť aj prostredníctvom siete autorov, do ktorej širiteľ patrí. Túto sieť je možné určiť na základe spoluautorstva. Spomínaná štúdia zistila homogenitu v kontexte autorských sietí. Autori publikujúci fake news spolupracujú s ďalšími autormi publikujúcimi fake news. To isté platí o autoroch pravdivých príspevkov a ich spoluautorských činnostiach. Bol tiež dokázaný pozitívny vzťah, keď ide o kredibilitu organizácie a kredibilitou autora, ktorý pod ňou publikuje. Štúdia sa tiež zamerala na prípady, kedy názov autora neobsahuje meno, ale vymyslené slová, či prezývky. Ich výskyt prevláda vo fake news príspevkoch. Keď ide o históriu publikovania fake news, ukázalo sa, že zo 69 autorov sa len u 11 z nich vyskytol prípad, kedy publikovali ako fake news, tak aj pravdivý príspevok. Kredibilita autora na základe histórie publikovaných príspevkov tak môže napovedať o povahe príspevkov, ktoré publikuje v budúcnosti.

Ako bolo popísané v kapitole 2.4 Atribúty fake news, medzi širiteľov patria okrem autora a publikujúceho subjektu tiež účty ďalej šíriace fake news. Analyzovaným typom účtov sú práve sociálne boty, ktoré bývajú používané ako nástroj pre šírenie fake news. Často spomínaným príkladom v tomto kontexte je trolia farma, majúca oficiálny názov *Internet Research Agency*, ktorá významným spôsobom zasahovala do amerických prezidentských volieb v roku 2016 a to favorizáciou kandidáta Donalda Trumpa v prospech ruských záujmov. Oficiálne zistenia Twitteru z roku 2018 ukázali, že v tejto udalosti bolo zapojených viac ako 50.000 botov napojených na Rusko (Eldh 2022).

Organizácia European Digital Media Observatory približne mesiac po vniknutí ruských vojsk na Ukrajinu a to od obdobia 15. marca 2022 do 25. apríla 2022 skúmala účty na sociálnej sieti Twitter, ktoré zdieľali príspevky z profilov ruských ambasád, ktoré šíria dezinformácie a propagandu o vojne na Ukrajine. Použitím systému Botometer, ktorý pochádza z Univerzity Bloomington v štáte Indiana bola týmto účtom pridelená pravdepodobnosť toho, že ide o sociálne boty. Systém Botometer určuje túto hodnotu na

škále 1-10. Z výsledkov vyplýva, že približne tretina analyzovaných účtov bola vyhodnotená ako bot (9% ako určite bot a 27% ako vysoko pravdepodobne bot) a šírila pro-ruské naratívy. Súčasne až 41% účtov nebolo možné jednoznačne priradiť do kategórie účtov plne vykazujúcich ľudské charakteristiky. Takéto charakteristiky Botometer pripísal len 29% z účtov. Navyše až 73% účtov obsahovalo prvý príspevok na profile z obdobia po 24. februári 2022, kedy ruská invázia začala. V tomto prípade sa prihliadalo na posledný pol rok od doby spracovávania štúdie a nie na celú históriu profilu od jeho vzniku. Je tak možné, že existuje koordinovaná aktivita, kedy sa účty pripravené v hybernácii stávajú aktívne až za účelom rozšírenia istého naratívu (Eldh 2022).

Charakteristiky o šíriteľoch je možné sledovať na dvoch úrovniach a to na úrovni individuálnej, ktorá obsahuje charakteristiky o jednotlivých účtoch, tak v rámci skupinovej úrovne. (Shu et al. 2017) Individuálny level reprezentujú charakteristiky vypovedajúce o kredibilitate profilu a to napríklad vek profilu, údaje o aktivite (počet statusov, aktivita zdieľania), informácie v profile (vyplnené bio, jeho dĺžka, informácie o lokalite), skúma sa tiež citové zafarbenie slov obsiahnutých v textových poliach na profile, alebo v príspevkoch na profile. (Shu et al. 2019b) Výsledky výskumu *Fake News Detection on Twitter Using Propagation Structures* napríklad ukazujú, že šíriteľa fake news majú nižší počet sledujúcich, ako užívatelia šíriaci príspevky obsahujúce overené fakty (Meyers et al. 2020). Súbor charakteristík skupinového levelu popisuje agregované charakteristiky užívateľských profilov. Skupinové charakteristiky profilov, ktoré sú v interakcii s fake news príspevkom, môžu v porovnaní so štandardnými príspevkami vykazovať odlišné hodnoty. Medzi tieto charakteristiky patria agregované hodnoty individuálnych charakteristík, ako percento verifikovaných účtov, alebo priemerne množstvo sledujúcich, či sledovaných účtov. (Shu et al. 2017) Určovanie individuálnych a skupinových charakteristík je nutné vykonávať v kontexte vybranej sociálnej siete, práve preto, že jednotlivé sociálne siete sa môžu v informáciách obsiahnutých v profiloch líšiť.

3.3 Detekcia na základe sociálneho kontextu

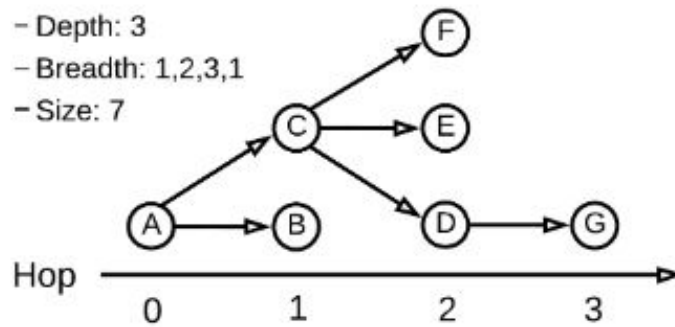
Pre znázornenie šírenie fake news po sociálnych sieťach sa používa tzv. *news cascade* (ďalej označené ako „kaskáda“). Táto kaskáda je definovaná (Zhou a Zafarani 2020) ako stromová štruktúra, ktorá zachytáva propagáciu istého príspevku po sociálnej sieti. Koreňový uzol kaskády reprezentuje užívateľa, ktorý ako prvý zdieľal príspevok (tzv. iniciátor). Ďalšie uzly kaskády reprezentujú užívateľov, ktorí následne šírili príspevok ďalej a to od svojho rodičovského uzlu, ku ktorému sú prepojené prostredníctvom hrany.

Rozlišujeme dve typy kaskád:

- **Kroková (tiež skoková) kaskáda**

Vypovedá o počte „krokov“, ktoré príspevok „vykonal“. V rámci krokovej kaskády rozpoznávame charakteristiky:

- Hĺbka – maximálny počet krokov v rámci kaskády
- Šírka – počet užívateľov zdieľajúcich príspevok v kroku k od prvého zdieľania príspevku iniciátorom
- Veľkosť – celkový počet užívateľov v kaskáde

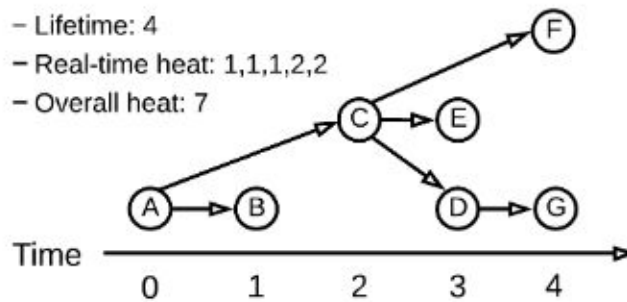


Obrázok 1 Kroková kaskáda (zdroj: Zhou a Zafarani 2020)

- **Časová kaskáda**

Vypovedá o časových charakteristikách šírenia príspevku. V rámci časovej kaskády rozpoznávame charakteristiky:

- Životnosť – najdlhšie obdobie počas ktorého sa príspevok šíri
- Heat v skutočnom čase – počet užívateľov zdieľajúcich príspevok v čase t
- Celková aktivita – celkové množstvo ľudí zdieľajúcich príspevok počas životnosti kaskády



Obrázok 2 Časová kaskáda (zdroj: Zhou a Zafarani 2020)

Detekcia fake news na základe sociálneho kontextu má tak charakter klasifikácie kaskád medzi fake a pravdivé pričom predpokladom je, že tieto kaskády vyzerajú a správajú sa odlišne, čo tak umožňuje odlišiť fake news a skutočné správy. Je napríklad možné na základe analýzy kaskád odhaliť, že fake news príspevky bývajú masívne šírené až po nejakej dobe od prvej publikácie, zatiaľ čo skutočné príspevky sú najviac šírené hneď, keď vznikli. Dôvodom môže byť oneskorenie v aktivácii hibernujúcich sociálnych botov.

4 Popis výskumu a metodológie

Cieľom tejto kapitoly je bližšie špecifikovať výskumné otázky a metodológiu použitú k ich zodpovedaniu.

4.1 Výskumné otázky

Hlavná výskumná otázka tejto diplomovej práce je definovaná ako:

Aké sú základné princípy detekcie fake news a aká je ich aplikovateľnosť pri využití v kontexte udalosti ruskej invázie na Ukrajinu, na príklade fake news v slovenskom jazyku?

Podporné otázky, ktoré je pre odpoveď na hlavnú výskumnú otázku nutné zodpovedať sú nasledovné:

- *Aké sú charakteristiky datasetov používaných pri detekcii fake news?*
- *Aké sú špecifiká a výzvy tvorby datasetu pre detekciu fake news vo zvolenom kontexte?*

4.2 Metodológia

Pre dosiahnutia cieľa práce a teda k zodpovedaniu uvedených výskumných otázok bude spracovaná analýza existujúcich datasetov, používaných pri úlohách detekcie fake news. Pre porovnanie sú zvolené dva datasety v anglickom jazyku a dva datasety v slovenskom jazyku, vytvorené pre potreby detekcie fake news. Pre porovnanie datasetov a zaistenie kompletnosti ich popisu je použitý popis charakteristík FNDD (fake news detection dataset characteristics) podľa (D'Ulizia et al. 2021), ktorý sa zoberá dôležitými charakteristikami fake news datasetov.

Výstupy analýzy fake news detekčných datasetov, spoločne s výstupmi teoretickej časti práce, poslúžia ako základ pre tvorbu datasetu pre detekciu fake news na zvolenom príklade ruskej invázie na Ukrajinu v slovenskom prostredí. Na záver sú popísané princípy charakterizujúce úlohu detekcie fake news obecné a ich špecifiká a výzvy, ktoré boli zistené pri tvorbe datasetu v danom kontexte.

5 Datasetsy pre detekciu fake news

S cieľom úspešne reagovať na šírenie fake news sú vytvárané automatizované nástroje, ktorých cieľom je detekovať fake news s čo najvyššou presnosťou. Tieto nástroje, založené na klasifikačných modeloch a technikách machine learningu, pracujú s detekciou fake news ako binárnym problémom, alebo s vyššou granularitou, keď ide o určenie triednych označení fake news (napríklad pracujúcich s mierou pravdivosti príspevkov akými sú: úplne nepravdivé, čiastočne nepravdivé, zavádzajúce), pričom dataset s vopred označenými dátami do tried slúži ako trénovacia sada pre učenie klasifikačného modelu.

Výsledky klasifikačného modelu na sade testovacích dát, bez označení výsledných tried, potom slúžia k určaniu jeho presnosti a k vzájomnému porovnávaniu s ďalšími klasifikačnými modelmi. Tak je možné určiť najvhodnejší model pre vyhodnocovanie daného vzorku dát a poňať tak najvhodnejšieho aj pre aplikáciu na reálnych dátach. Pre úlohy detekcie fake news tak predstavuje dataset obsahujúci fake a skutočné príspevky, označené na základe pravdivosti, a jeho dostupnosť a kvalita dôležitý problém, nakoľko klasifikačné modely závisia na dátach v nich obsiahnutých. (D'Ulizia et al. 2021)

Existuje mnoho datasetov určených na výskum modelov pre detekciu fake news a tým prispievajúcich k zlepšovaniu úspešnosti odhaľovania nepravdivých správ. Nakoľko fake news sa vyvíjajú v čase a líšia sa tematicky aj geograficky, je potrebné, aby datasety pre skúmanie fake news adekvátne reprezentovali túto rôznorodú podobu fake news. To zabezpečí, že klasifikačné modely budú dostatočne generalizované a uplatniteľné pri detekcii fake news v rôznych kontextoch, jazykoch a doménach. Aby bolo možné uvedné ciele dosiahnuť, je dôležité, aby existovali kvalitné datasety dotýkajúce sa rôznych tém, jazykov, na ktorých by bolo možné klasifikačné modely ďalej vyvíjať a testovať.

Za týmto účelom je v tejto časti práce ďalej spracovaná analýza existujúcich datasetov využívaných pri detekcii fake news. Popísané sú populárne a často využívané datasety, ako aj datasety so zameraním podobnému tomu, ktorému sa venuje táto diplomová práca.

5.1 LIAR

Tento dataset bol publikovaný v roku 2017 na základe dát z webovej stránky polifact.com, ktorá získala novinársku Pulitzerovu cenu a zaoberá sa rýchlym hodnotením krátkych výrokov politikov primárne z USA, ktoré sa objavujú v televízii a na internete. (Wang 2017)

Výroky sú zaradené na základe pravdivosti do šiestich tried, pričom úplne vymyslené a nepravdivé výroky sú zaradené do špeciálnej kategórie „*pants-fire*“. Medzi kategórie označujúce pravdivosť výrokov ďalej patria (Drobnic Holan 2022):

- pravda – pravdivé tvrdenie, bez vynechania dôležitých faktov

- prevažne pravda – tvrdenie je pravdivé, ale vyžaduje doplňujúce informácie, či upresnenie
- polopravda – tvrdenie je čiastočne pravdivé, ale sú vynechané dôležité detaily, či kontextové informácie
- sotva pravda – tvrdenie obsahuje element pravdivosti, ale sú ignorované kritické fakty, vďaka ktorým by výrok pôsobil iným dojmom
- nepravda – tvrdenie nie je pravdivé

Dataset obsahuje približne 12 800 krátkych tvrdení, čo z neho robí jeden z najrozsiahlejších datasetov zameraných na nepravdivé tvrdenia. Každá z kategórií obsahuje od 2063 do 2638 tvrdení, mimo kategórie „pants-fire“ ktorá obsahuje 1020 tvrdení. Dataset je rozdelený na trénovacie data (10 269 tvrdení), ďalej validačné (1 284 tvrdení) a testovacie data (1 283 tvrdení). (Wang 2017)

Pre zostrojenie datasetu bolo použité API webovej stránky polifact.com.(Wang 2017) Editori tohoto portálu každý deň vyberajú tvrdenia z mediálneho prostredia vhodné overenia. Tieto vybrané výroky potom prechádzajú nezávislým hodnotením, pričom na stránke je ku každému tvrdeniu uvedená analýza a zdroj, na základe ktorého bola tvrdeniu priradená jedna z kategórií pravdivosti. (Drobnic Holan 2022)

Medzi atribúty datasetu patrí:

- id – unikátny identifikátor tvrdenia
- label – pravdivosť
- statement – tvrdenie
- subject – téma alebo predmet, ktorého sa tvrdenie týka
- speaker – autor tvrdenia
- metadata o autorovi:
 - job_title
 - state_info
 - party_affiliation
 - barely_true_counts
 - false_counts
 - half_true_counts
 - mostly_true_counts
 - pants_on_fire_counts
- context – príležitosť, pri ktorej tvrdenie odznelo

Priemerná dĺžka hodnotených tvrdení je 107 znakov, pričom najkratšie tvrdenie má 11 znakov a najdlhšie 3197 znakov.

5.2 FakeNewsNet

Repozitár FakeNewsNet je zameraný na fake news a ich šírenie v prostredí sociálnych sietí, konkrétne Twitter. Cieľom autorov bolo vytvoriť repozitár pre skúmanie fake news, ktorý bude obsahovať nielen obsah príspevku, ale tiež informácie o sociálnom kontexte a tzv. dynamických informáciách. To umožňuje využitie repozitára pri skúmaní viacerých prístupov k detekcii fake news a tiež analýze dezinformačných taktík. (Shu et al. 2019a)

Zber dát a tiež určenie pravdivosti príspevkov obsahujúcich fake news je založené na hodnotení stránok overujúcich fakty polifact.com a gossipcop.com. Zatiaľ čo polifact.com je zameraná, ako už bolo uvedené, na hodnotenie pravdivosti krátkych tvrdení politikov, gossipcop.com sa zameriavala na hodnotenie pravdivosti správ zo zábavného priemyslu a sveta celebrit. (Shu et al. 2019a) Od roku 2022 bola činnosť gossipcop.com ukončená. (Wikipedia 2023)

Repozitár FakeNewsNet obsahuje celkovo dva datasety rozlíšené na základe zdrojových stránok. Polifact.com dataset obsahuje celovo 1 056 príspevkov, z ktorých je 432 príspevkov typu fake news a 624 pravdivých príspevkov. Dataset gossipcop.com je značne väčší a obsahuje spolu 22 856 príspevkov, z ktorých je 5 323 príspevkov typu fake news a 16 817 pravdivých príspevkov. (Shu et al. 2019a)

Tieto datasety majú nasledujúcu štruktúru (Shu 2023):

- id – unikátny identifikátor webového príspevku
- URL – url webového článku
- title – názov príspevku
- tweet_id – identifikátory twitter príspevkov zdieľajúcich webový príspevok, ktoré sú oddelené medzerníkom

Ďalšie meta informácie o obsahu každého príspevku sú (Shu 2023):

- text – textový obsah tela príspevku
- images – zoznam URL, odkazujúci na obrázky, ktoré sú súčasťou príspevku na danej webovej stránke
- publish date – dátum zdieľania príspevku na webovej stránke

FakeNewsNet obsahuje pipeline pre obohatenie datasetu, prostredníctvom webových scraper a Twitter API o nasledujúce informácie (Shu 2023):

- obsah príspevku – textový obsah
 - textový obsah Twitter príspevku, zdieľajúci článok
 - textový obsah zdieľaného webového príspevku
- sociálny kontext – užívateľské reakcie na príspevok v podobe odpovedí, zdieľaní a reakcií a informácie o užívateľoch majúci interakciu s príspevkom
 - informácie o počtu zdieľaní
- šíriteľ
 - informácie z profilu užívateľského účtu zdieľajúceho príspevok

- možnosť získať 200 najnovších príspevkov z profilu užívateľa
- informácie o profiloch v spojení s profilom (followers a followings)
- dynamický kontext – informácie umožňujúce sledovanie vývoja príspevku na sociálnej sieti, akými sú časové značky
 - dátum zdieľania twitter príspevku

Pre získanie doplnujúcich dát prostredníctvom Twitter API je potrebné mať založený Twitter developerský účet s prideleným API Key, autentizačným kľúčom, ktorý je potrebný pre použitie API. Pre vytvorenie konta je potrebné získať od spoločnosti Twitter overenie a schválenie účelu zberu a využitia dát. (Twitter, Inc. 2023)

V rámci predspracovania dát, pred ich použitím v klasifikačných úlohách, boli vykonané nad dátami tieto úpravy (Shu et al. 2019a):

- tokenizácia slov
- stemming – odvodenie základu slov
- odstránenie duplít
- odstránenie interpunkcie
- odstránenie špeciálnych znakov, symbolov a hashtagov
- odstránenie anglických stopových slov
- prevedenie textu do lowercase

5.3 COVID-19 fake news dataset v slovenskom jazyku

S cieľom skúmať použitie deep learning modelov pre automatickú detekciu fake news v slovenskom jazyku vytvorili autori (Sarnovsky et al. 2022) dataset, obsahujúci fake news týkajúcich sa ochorenia Covid-19. Pre zber dát použili autori platformu MonAnt. MonAnt je monitorovací nástroj webových zdrojov, prostredníctvom ktorého je možné zbierať multimodálne data a ich kontext, navrhnutý za účelom detekcie tzv. „antisociálneho chovania“.

Zdrojom pre zber dát, textu publikovaných článkov, boli slovenské spravodajské médiá a tiež zdroje často publikujúce konšpiračný obsah. Pre výber článkov zo zvolenej domény bol aplikovaný filter na kľúčové slová Covid, Covid-19, Coronavirus a SARS-CoV-2. (Sarnovsky et al. 2022) Autori článku neuvádzajú časové obdobie, v ktorom boli články obsiahnuté v datasete publikované.

Určenie finálnej pravdivostnej triedy bolo prvotne odvodené na základe autenticity širiteľa, podľa databázy konspiratori.sk a skóre prideleným jednotlivým hodnoteným webom, do dvoch tried. Pravdivé správy v datasete pochádzajú zo štandardných slovenských médií a správy fake news zo zdrojov s vysokým skóre u konspiratori.sk (vysoká pravdepodobnosť publikovania nepravdivých informácií). Takto určené pravdivostné triedy môžu viesť k „false negative“ a „false positive“ klasifikácii, nakoľko zdroje s nízkym skóre zvyknú tiež publikovať štandardné správy z tlačových agentúr. Z toho dôvodu autori manuálne overili články prvotne označené na základe širiteľa ako fake news a v prípade chybnéj klasifikácie zmenili triedu príspevku. (Sarnovsky et al. 2022)

Štruktúra datasetu neobsahuje informácie o šíriteľovi, alebo sociálnom kontexte a obsahuje iba textové informácie z obsahu príspevkov a to hlavný nadpis a prefix spolu s hlavným textovým obsah príspevku, spojené pre vyhodnocovanie do jedného poľa.

Výsledný dataset obsahuje 12 885 pravdivých príspevkov a 851 príspevkov označených ako fake news. Napriek tomu, že je výsledný dataset nevyvážený, keď ide o rozdelenie dát do jednotlivých kategórií, autori sa rozhodli neaplikovať techniky na balansovanie hodnôt. (Sarnovsky et al. 2022)

V rámci predspracovania datasetu autori odstránili z textu interpunkčné znamienka, stopové slová, iné ako alfabetické znaky, hypertextové linky a ponechané boli len písmená slovenskej abecedy. Aby tréning modelu prebehlo čisto na základe textu obsahu príspevku, ignorované boli tiež informácie o zdrojoch informácií, či použité obrázky a fotografie. Následne boli textové data konvertované do vektorovej reprezentácie, použitím techniky word embeddingu word2vec. (Sarnovsky et al. 2022)

5.4 Fake news dataset v slovenskom jazyku

Tým autorov datasetu popísaného v predchádzajúcej kapitole pripravil tiež ďalší dataset v slovenskom jazyku. Na rozdiel od predchádzajúceho datasetu, kedy bola výsledná trieda dát v datasepte určená primárne na základe kredibility šíriteľa, bola pri tvorbe tohoto datasetu využitá metóda crowdsourcingu. (Sarnovský et al. 2020)

Zdrojom pre vytipovanie stránok s fake news obsahom pre získanie dát do datasetu bola rovnako, ako v predchádzajúcom prípade stránka konspiratori.sk. Doménovo je dataset zameraný na ekonomické, svetové a lokálne správy. Autori článku neuvádzajú časové obdobie, v ktorom boli články obsiahnuté v datasepte publikované. Pre samotné ťaženie dát zo stránok bol použitý nástroj MonAnt. Z datasetu boli odstránené články v podobe blogových príspevkov. (Sarnovský et al. 2020)

Nezávislí anotátori označili každý príspevok pravdepodobnosťou vypovedajúcou o tom, že ide o fake news, alebo nie. Príspevky, u ktorých anotátori nevedeli určiť, o aký typ príspevku ide, boli z datasetu vylúčené. Zatiaľ čo z konšpiračných stránok bolo z celkového počtu 968 príspevkov označených 821 z nich ako fake news, u seriózných médií bolo z celkového počtu 547 príspevkov len 19 príspevkov označených ako možný fake news. (Sarnovský et al. 2020)

Dataset s určenými triedami pravdivosti má podobou (Sarnovský et al. 2020):

- id – identifikátor príspevku
- label – označenie poskytnuté anotátorom
- text – textový obsah príspevok
- user – osoba anotátora

Výsledný dataset s celkovým počtom 1536 príspevkov, s počtom fake news príspevkov prevažujúcim o 145 nad reálnymi príspevkami, bol použitý pri tréningu modelov založených na neurónových sieťach. Pre tento účel boli textové data spracované nasledovne (Sarnovský et al. 2020):

- tokenizácia slov – predstavuje metódu pre transformáciu viet do poľa slov, tzv. tokenov, ktoré sú pre využitie v neurónových sieťach transformované ďalej do vektorovej reprezentácie
- prevedenie textu do lowercase
- odstránenie interpunkcie
- vektorová reprezentácia slov prostredníctvom word embeddingu word2vec

5.5 Porovnanie datasetov pre detekciu fake news

Kritéria pre porovnanie zvolených datasetov sú určené na základe prehľadovej štúdie (D’Ulizia et al. 2021). V štúdiu autori spracovávajú analýzu dvadsiatich siedmich datasetov pre detekciu fake news (Fake News Detection Datasets - FNDD) a to na základe nimi definovaných charakteristík pre porovnávanie datasetov. Pri definícii charakteristík autori vychádzajú z atribútov fake news tak, ako ich definuje (Zhang a Ghorbani 2020).

Tieto atribúty fake news rozširujú o charakteristiky, ktoré sú uvedené v popise analyzovaných datasetov a ktoré boli systematicky zbierané v rámci prehľadovej štúdie. Aplikáciou atribútov fake news na charakteristiky datasetov pre detekciu fake news a ich obohatením o ďalšie charakteristiky podľa popisov zo skúmaných datasetov, autori spracovali zoznam charakteristík, ktoré nepopisujú fake news ako také, ale datasety využívané pri ich detekcii.

Z prehľadovej štúdie (D’Ulizia et al. 2021) autori uvádzajú, že popis charakteristík datasetov obsahuje tri atribúty fake news, z celkových štyroch podľa definície (Zhang a Ghorbani 2020) a to obsah fake news, sociálny kontext a šíriteľ. V tejto podobe sú atribúty fake news definované tiež v kapitole 2.4 Atribúty fake news.

Štvrtým atribútom fake news je cieľová obeť fake news, avšak informácia o tom, na koho je fake news mierená, nepatrí medzi dohľadateľné informácie a nevyskytuje sa v štúdiách zameraných na automatickú detekciu fake news, ani v popise analyzovaných datasetov. Obecné sa dá povedať, že v prípade politických fake news, alebo fake news šírených na internete, je cieľom zasiahnuť a ovplyvniť čo najvyšší počet užívateľov. Cieľovou obeťou v prípade fake news šírených v slovenskom jazyku môže byť tak každý občan Slovenskej republiky, prípadne každý človek hovoriaci slovenským jazykom. Analyzované popisy datasetov obsahovali aj informácie, ktoré nie je možné zaradiť pod niektorý z atribútov fake news, tak ako ich definuje (Zhang a Ghorbani 2020) a preto autori (D’Ulizia et al. 2021) vytvorili novú kategóriu charakteristík obsahujúcu obecné informácie.

Tabuľka 2 Fake news detection dataset FNDD charakteristiky. Zdroj: (Dulizia et al. 2021)

Komponenta fake news	Charakteristika fake news detection datasetov	Popis
Obsah	Typ správy	Typ zbieraného obsahu z fake news ako napríklad nadpis, telo príspevku, video, obrázky
	Typ nepravdivých správ	Odlíšenie typu fake news a blízkych konceptov ako napríklad satira, nepravdivé hodnotenia, fake news
	Doména	Tematické zameranie správ ako napríklad politika, šport, ekonomika, technológie
	Jazyk	Jazyk fake news. datasey môžu pokrývať jeden alebo viacero jazykov – multilinguistické datasey
	Hodnotiaca škála	Použité označenia pre určenie pravdivosti správ
Sociálny kontext	Mediálna platforma	Prostredie v ktorom sa fake news primárne šíri a dostáva sa k cieľovým obetiam, ide o zdroj. Zahrňuje webové médiá a sociálne siete
Šíriteľ	Spontánnosť	Rozlišuje, či fake news boli umelo vytvorené prekrútením pravdy u pravdivých správ z overených zdrojov, alebo či ide o fake news z reálneho sveta
Obecné informácie	Veľkosť	Počet záznamov obsiahnutých v datasee
	Dostupnosť	Verejná dostupnosť dát obsiahnutých v datasee
	Časové obdobie	Definícia časového obdobia v rámci ktorého boli data zbierané
	Účel aplikácie	Medzi účely patrí napríklad odhaľovanie klebiet, detekcia fake news, či overovanie faktov

Popis charakteristík pre porovnanie jednotlivých datasetov obsahuje Tabuľka 3.

Tabuľka 3 FNDD charakteristiky vybraných datasetov. Zdroj: vlastné spracovanie

Charakteristika FNDD	LIAR	FakeNewsNet	Covid-19 SK	Fake news SK
Typ správy	text (kráte tvrdenia)	text (nadpis, telo), obrázky, sociálny kontext, šíriteľ	text (nadpis, prefix + hlavný text)	text (hlavný text)
Typ nepravdivých správ	fake news	fake news	fake news	fake news
Doména	politika	politika, zábavný priemysel	zdravie	politika, ekonomika
Jazyk	anglický	anglický	slovenský	slovenský
Hodnotiaca škála	6 stupňov	2 stupne	2 stupne	2 stupne
Mediálna platforma	webové média	webové média, sociálne siete	webové média	webové média
Spontánnosť	áno	áno	áno	áno
Veľkosť	12 836	23 912	12 885	968
Dostupnosť	áno	áno	na požiadanie	na požiadanie
Časové obdobie	2007-2016	primárne 2016-2018, inak nešpecifikované	nešpecifikované	nešpecifikované
Účel aplikácie	detekcia fake news	detekcia fake news	detekcia fake news	detekcia fake news

Ako je možné vidieť na základe informácií v Tabuľka 3, typ nepravdivých správ a účel aplikácie datasetu je v každom z analyzovaných datasetov rovnaký, nakoľko je táto práca zameraná na datasety obsahujúce fake news, vytvorené za účelom detekcie fake news. Typy nepravdivých správ ako misinformácie, false news, satirické príspevky, či nepravdivé hodnotenia nie sú predmetom tejto analýzy. Rovnako nie je predmetom analýza datasetov zameraných napríklad na odhaľovanie falošných hodnotení na webových stránkach e-shopov. Všetky fake news pochádzajú z reálneho sveta a nie sú vytvorené umelo, zmenou obsahového významu skutočných správ tak, aby neboli viac pravdivé.

Najčastejším typom správ obsiahnutých v datasete je textový obsah fake news. Len rozšírený repozitár FakeNewsNet obsahuje aj obrázky, ktoré sú tiež súčasťou obsahu fake news, ale navyše tiež údaje o sociálnom kontexte fake news a šíriteľoch. Tieto údaje sú dostupné po využití pipeline k extrakcii dát zo sociálnej siete Twitter, prostredníctvom API. Ostatné datasety využívajú ako zdroj textového obsahu webové stránky zdrojových médií. V datasete LIAR je textovým obsahom len krátke tvrdenie, nie celý textový obsah príspevku, na rozdiel od ostatných datasetov, v ktorých sa nachádzajú celé textové obsahy príspevkov. Triedy hodnotiace pravdivosť dát v datasete obsahovali najčastejšie dva hodnoty, zatiaľ čo dataset LIAR obsahuje výsledných tried šesť. Pravdivostné hodnoty vyplývajú z analýz tvrdení na stránke polifact.com, kde sú manuálne hodnotené tímom expertov, ktorý používa túto škálu. Pri automatickom odvodzovaní pravdivosti tried na základe kredibility šíriteľa, či zdrojovej webovej stránky publikujúcej fake news, je však používaná škála zvyčajne dvojstupňová.

Keď ide o veľkosť datasetov, populárne datasety v anglickom jazyku ako LIAR a FakeNewsNet obsahujú tisíce záznamov, rovnako slovenský dataset na fake news o Covid-19, v práci označený ako dataset Covid-19 SK. Vidíme však, že existujú aj datasety so stovkami záznamy, ako napríklad slovenský dataset s pracovným názvom Fake News SK.

Zaujímavosťou je, že u slovenských datasetov chýbalo vymedzenie časového obdobia v ktorom boli obsiahnuté príspevky publikované. Aj keď data o dátume publikovania nie sú súčasťou výsledných datasetov, charakteristika datasetu by v najlepšom prípade túto informáciu obsahovať mala, nakoľko patrí medzi dôležité kritéria pri výbere a hodnotení využiteľnosti a relevancie datasetov pri špecifických úlohách detekcie fake news.

6 Detekcia fake news na zvolenom príklade

Udalosť ruskej invázie na Ukrajinu je sprevádzaná tiež masívnou dezinformačnou kampaňou zo strany Ruskej federácie. Táto dezinformačná kampaň je vedená naprieč mnohými krajinami, rozličnými jazykmi a sociálnymi sieťami. Podobne je potrebné tiež pristupovať k detekcii fake news. Doterajšie štúdie zamerané na detekciu fake sa sústreďujú primárne na anglický jazyk a použitie vzorových datasetov pre tréning a hodnotenie úspešnosti detekcie, ako napríklad dataset LIAR, alebo repozitár FakeNewsNet (Davoudi et al. 2022)

Aby bolo možné zaistiť včasnú detekciu fake news, v čase, kedy fake news vzniká, je potrebné vytvoriť modely nezávislé na téme, zdroji, či forme fake news a tak čo najlepšie využiteľné v rôznych jazykoch a kontextoch. Z toho dôvodu je potrebné mať pre učenie dostupné datasety, na ktorých by bolo možné modely pre detekciu fake news nielen tréningovať, ale tiež hodnotiť ich použiteľnosť pri aplikácii v špecifických kontextoch.

Z toho dôvodu je cieľom práce vytvoriť dataset v slovenskom jazyku, zameraný na tému udalosti ruskej invázie na Ukrajinu v roku 2022. V tejto kapitole je tak popísaný proces tvorby datasetu. Popísané sú kritéria, na základe ktorých boli zvolené zdroje pre vytváraný dataset a ich analýza, následne je popísaný samotný proces extrakcie a nástroj použitý k spracovaniu a zberu dát zo zdrojových webových stránok. Na koniec je popísaný vytvorený dataset a zhodnotená jeho možná použiteľnosť pri aplikácii na reálnych dátach.

6.1 Zvolený kontext

Zvolený kontext zahŕňa vymedzenie úlohy detekcie fake news, v rámci ktorej má byť dataset použitý. Ako je uvedené už v nadpise tejto práce, vytvorený dataset má adresovať fake news na tému ruskej invázie na Ukrajinu, šírených na Slovensku.

Zvolené časové obdobie publikovania fake news bolo určené ako obdobie od 23. januára 2023 do 24. februára 2023. Obdobie odpovedá jednému mesiacu, ktorý predchádzal dňu 24. február 2023, kedy ruská agresia voči Ukrajine trvala presne rok. Možno je tak analyzovať fake news šírené bezprostredne pred zahájením druhého roku ruskej invázie na Ukrajinu. Predpokladom pre zvolenie daného časového obdobia bolo, že fake news budú v tomto období šírené o niečo viac, nakoľko sa očakáva zvýšené hodnotenie a analýzy roku trvania tejto vojny zo strany štandardných médií, na ktoré budú fake news média patrične reagovať a zaplavovať informačný priestor protichodnými a nepravdivými informáciami.

Zvoleným prístupom pre detekciu fake news je detekcia na základe textového obsahu fake news príspevkov, publikovaných na stránkach spájaných so šírením fake news. Nakoľko sú prostriedky sociálnych sietí pre získavanie dát popisujúcich širší kontext fake news obmedzené, nebudú pre detekciu využité a do datasetu zbierané dáta o šíriteľoch a sociálnom kontexte, popisujúce detaily šírenia fake news na sociálnych sieťach.

6.2 Zdroje dát

Určenie pravdivosti príspevkov bolo odvodené na základe kredibility publikujúcej webovej stránky. Táto kredibilita bola získaná v prípade fake news na základe hodnotenia iniciatívy konspiratori.sk, ktorú využili pri určení pravdivosti príspevkov tiež autori (Sarnovsky et al. 2022). Zoznam stránok s pochybnou dôveryhodnosťou na konspiratori.sk obsahuje v marci 2023 280 webových stránok, ktoré boli do zoznamu pridané na základe hodnotenia nezávislej komisie, podľa stanovených a transparentných kritérií. Medzi ne patrí šírenie dezinformácií, lživej propagandy, prekrútených a nejasných popisov udalostí, extrémistický obsah, hanobenie, nerešpektovanej novinárskej etiky a podobne. (konspiratori.sk 2023) Každá stránka ma uvedený priemerné skóre na základe hodnotenia týchto kritérií a stránky na zozname majú v čase písania práce hodnotenia skóre od 6.1 do 9.8, pričom vyššie hodnotenie znamená horšie hodnotenie stránky.

Nie všetky stránky na zozname od konspiratori.sk sa však venujú téme vojny na Ukrajine. Časté na zozname sú napríklad stránky so šarlatánskym obsahom, ktoré vznikli ako reakcia na pandémiu Covid-19. Na zozname sa tiež vyskytujú stránky českých domén, s článkami primárne v českom jazyku, zatiaľ čo zameranie práce je na články v slovenskom jazyku. Z toho dôvodu nebolo možné pri výbere stránok zvoliť stránky s najvyšším a teda najhorším hodnotením, čím by bola vyššia pravdepodobnosť, že články publikované budú typu fake news.

Pre umožnenie a uľahčenie automatizovaného zberu dát, boli pre extrakciu vybrané stránky, ktoré obsahovali sekciu pre uverejňovanie článkov k téme Ukrajiny. Týmto spôsobom nie je nutné používať pre filtráciu článkov vyhľadávanie kľúčových slov, alebo manuálne selektovať články na danú tému.

Tabuľka 4 Vybrané zdroje fake news príspevkov

Webová stránka	URL	Skóre udelené komisiou portálu konspiratori.sk
Zem & Vek	https://zemavek.sk/	9.5
News front SK	https://sk.news-front.info/	8.8

6.2.1 Zem & Vek

Zem & Vek (tiež Zem a vek) od vydavateľstva SOFIAN je slovenské mesačné periodikum a tiež názov pre rovnomenný spravodajský webový portál. Podľa prehlásenia na webovej stránke ide o „*alternatívne spravodajstvo*“, ktoré sa chce vymaniť z „bežných stereotypov nielen svojím obsahom a spracovaním,“ ale jeho výnimočnosť spočíva tiež v nezávislosti a to vďaka tomu, že „*ostávame aj naďalej bez akejkoľvek reklamy a preto sa nemusíme spovedať žiadnym sponzorom, inzerentom ani politickým stranám.*“ (Zem a vek 2023)

V podobe tlačenej periodickej tlače začal Zem a Vek vychádzať dňa 10.05.2013. Zakladateľom, hlavnou tvárou a tiež majiteľom nakladateľstva SOFIAN je Tibor Rostas. (Ministerstvo kultúry Slovenskej republiky 2023). V priebehu rokov sa Zem a Vek, spolu s osobou Tibora Rostasa, stali na Slovensku známi pre šírenie pro-ruských naratívov, podporujúcich vedenie a geopolitické počínanie Ruskej federácie.

Prehlásenie o údajnej nezávislosti nie je v súlade so známymi aktivitami Tibora Rostasa. V minulosti navštívil spolu s Dušanom Budzákom ruskú ambasádu v Bratislave, kde ich prijal vtedajší veľvyslanec Ruskej federácie na Slovensku Kuznecov. Z návštevy unikla audio nahrávka, na ktorej zástupci Zem a Vek prosia o finančnú podporu s cieľom podporiť úzku spoluprácu Ruska a Slovenska, ktoré je pre Rusko „zaujímavé geostrategické teritórium.“ Kuznecov prisľúbil oslovenie patričných strán v Moskve a mediálny dom Russia Today. Oslovení ľudia sa stretli spoločne v Bratislave s Rostasom a Budzákom na slávnostnej recepcii a o mesiac neskôr, v júni 2015, navštívili spoločne Moskvu, kde hovorili o svojom záujme vytvoriť zo Zem a Vek mediálny dom a ruskí špecialisti prisľúbili svoju pomoc. (Benčík 2016)

Rostas bol v roku 2019 odsúdený vo veci hanobenia rasy, národa a presvedčenia v súbehu s prečinom podnecovania národnostnej, rasovej a etnickej nenávisťi a to vďaka publikovanému článku s anitseministickým obsahom. Rostasovi bol udelený peňažný trest vo výške 40 000 eur. (Aktuality.sk 2019)

V roku 2020 spoločnosť Meta odstránila trvale Facebookové stránky Zem a Vek pre opakované porušovanie pravidiel, redakcia však reagovala vytvorením nového profilu na sociálnej sieti. (Struhárik 2020) V súvislosti s vypuknutím vojny na Ukrajinu tiež napadli na pár hodín webovú stránku Zem a Vek hackeri zo skupiny Anonymous, aby upozornili na fake news, ktoré šíria. (Živé.sk 2022)

6.2.2 News front SK

News Front je webová stránka pôvodom z Ruska, publikujúca obsah v desiatich jazykoch a na desiatich webových doménach. Označuje sa ako objektívne spravodajstvo o geopolitických a vojenských témach. Pôvodne vznikol v období anexie Krymu a bol zameraný na ospravedlňovanie Putinových činov. Neskôr sa rozšíril do viacerých európskych krajín s cieľom šíriť propaganda a ovplyvňovať dianie v Európe. (Watts a Cerulli 2021)

Britský Inštitút pre strategický dialóg (Institute for Strategic Dialogue), zameraný na analýzu extrémistických hrozieb, v súvislosti s News Front uvádza, že je opakovane identifikovaný ako zdroj putinovskej a ruskej dezinformačnej siete a podľa správy zahraničného oddelenia amerického ministerstva financií (U.S. Department of the Treasury's Office of Foreign Assets Control) spolupracuje s Federálnou službou bezpečnosti (Federal Security Service, FSB), hlavnou bezpečnostnou službou Ruskej federácie a cieľi na podporu ruských záujmov. (U.S. Department of the Treasury's Office of Foreign Assets Control 2021)

Informácie o News Front boli dohľadávané tiež na stránke, ktorá bola ako slovenská verzia webu využitá pri spracovaní tejto práce. V hlavnom menu v sekcii sa vyskytuje pole „About“, avšak pole neobsahuje odkaz na žiadnu podstránku. Popis „About“ nie je jediným anglicizmom použitým v obsahu stránku. Do slovenčiny sú preložené články publikované na webe, avšak niektoré sekcie a reakcie ako *comments*, *share*, *read also* sú písané v angličtine. Odkazy na profily webu na sociálnych sieťach odkazujú na stránky *ENnewsfront* a tie sú na Facebooku a Twitteri v súčasnosti pozastavené. Profil na sociálnej sieti VK má názov „*Latest world news from the DPR/LPR, Ukraine*“, čo predpovedá o aktuálnej téme, ktorej sa News Front v svojich článkoch venuje. To všetko dokazuje, že stránka je len prostriedkom, ktorým sa na Slovensku publikujú preložené články zahraničných médií, v tomto prípade ruského Front News, so zameraním na kvantitu a rýchlosť publikovania, skôr než na prinášanie kvalitných novinárskych článkov z tvorby serióznej a transparentnej redakcie.

V apríli 2020 spoločnosť Meta odstránila facebookové stránky News Front. (Meta 2020) Rovnaké kroky učinila aj spoločnosť Twitter. Stránky tiež zamedzili zdieľanie obsahu z pridružených webov. Keďže zdieľanie článkov medzi facebookové skupiny patrí medzi hlavné spôsoby, ako stránky získavajú čitateľov a šíria svoje naratívy, netrvalo dlho a vzniklo mnoho „*mirroring*“ stránok, s úplne rovnakým obsahom ako mali blokované weby, ale na doménach, ktoré nie sú blokované filrami sociálnych sietí. To znamená, že zdroje porušujúce pravidlá sociálnych sietí môžu byť naďalej aktívne. News Front ma zavedenú sieť ľudí na sociálnych sieťach, ktorých jediným cieľom je zdieľať články z týchto webov medzi užívateľov v skupinách. (Watts a Cerulli 2021)

Analýza Aliancie pre ochranu demokracie (Alliance for Securing Democracy) zameraná na aktivitu stránok a profilov napojených na News Front ukázala, že články z nich sú pravidelné zdieľané naprieč najmenej 235 Facebook skupinami s potenciálnym dosahom až 3.4 milióna užívateľov. (Watts a Cerulli 2021).

6.2.3 SITA Slovenská tlačová agentúra

Pre získanie pravdivých správ boli uvažované agentúrne médiá na Slovensku, ktoré poskytujú aktuálne základné a faktografické informácie. Slúžia ako zdroj správ pre ostatné médiá na Slovensku a pokrývajú väčšinu udalostí a tém, ktoré sa v skúmanom období vyskytovali naprieč slovenskými médiami. Zvolená bola agentúra SITA.

Slovenská tlačová agentúra a.s. (SITA) vznikla v roku 1997 s cieľom prinášať aktuálne informácie z kultúrneho, hospodárskeho, športového a politického sveta, ktoré budú aktuálne, vyvážené a politicky nestranné. Denne vyprodukuje viac ako 250 správ, ktoré sú určené primárne médiám, PR agentúram, komunikačným oddeleniam, firmám, či politickým stranám. (Agentúra SITA 2023)

SITA spravuje webový portál webnoviny.sk, kde uverejňuje aktuálne tlačové správy. (Agentúra SITA 2023)

SITA vznikla vďaka potrebe vytvoriť súkromnú spravodajskú agentúru. Dovtedy existujúca, verejnoprávne zriadená Tlačová agentúra Slovenskej republiky (TASR) aj napriek snahám

v rokoch 1994 agentúru privatizovať, zostala po nástupe garnitúry Vladimíra Mečiara v roli hlavnej verejnoprávnej agentúry. Samotní pracovníci vtedajšej TASR však stáli za myšlienkou vzniku súkromnej agentúry SITA. S nástupom SITA postupne klesala dominancia TASR. Postupné etablovanie SITA malo po roku 2002 dopad na úbytok klientov a pokles tržieb z činnosti TASR. (Filo 2007)

Akciová spoločnosť SITA pôsobí ako súkromný poskytovateľ informačných služieb na Slovensku. Rozsahom poskytovaných služieb si po zavedení mediálneho a obrazového servisu vybuďovala post všeobecnej národnej agentúry. (Filo 2007)

6.3 Štruktúra datasetu

Štruktúra datasetu, zameraného na detekciu fake news na základe textového obsahu obsahuje tieto atribúty:

- URL – webová adresa príspevku
- Dátum – Dátum uverejnenia príspevku
- Nadpis – Nadpis príspevku
- Text – Textový obsah príspevku. Do textového obsahu je zahrnutý ako perex, tak aj samotný hlavný textový obsah príspevku. Nie všetky médiá uvádzajú perex príspevku, preto ak príspevok perex obsahuje, je do datasetu zahrnutý spoločne s hlavným textom.
- Pravdivosť – Označenie pravdivosti príspevku. Nadobúda dve hodnoty a to real a fake.

6.4 Zber dát

Pre extrakciu dát zo zvolených stránok bolo cieľom použiť automatizované nástroje. Manuálna extrakcia dát je pri veľkom množstve dát časovo náročná a neefektívna a zvyšuje sa riziko chyby pri zbere dát. Niektoré webové služby, v častom prípade sociálne siete, poskytujú webové API pre prístup a prácu s dátami. Alternatívou je použitie web scrapingových nástrojov, ktoré boli využité tiež pri zbere dát z určených stránok.

Web scraping je technika, pomocou ktorej je možné strojovo spracovávať obsah webových stránok. Webové stránky sú písané v značkovacom jazyku HTML (HyperText Markup Language), ktorý definuje ich štruktúru a obsah. Pomocou HTML a CSS značiek je možné odkázať na jednotlivé časti webovej stránky a v prípade web scrapingu je pomocou nich možné tieto časti čítať a ukladať.

Pre web scraping bol použitý nástroj Apify, dostupný na adrese <https://console.apify.com/>, ktorý je vzhľadom na rozsah spracovaných dát pre potreby tejto diplomovej práce voľne dostupný, bez nutnosti predplatenia balíčku Apify služieb. Apify umožňuje zber dát prostredníctvom definície vlastných tzv. actorov, prípadne Apify poskytuje tiež množstvo predefinovaných actorov na rôzne úlohy, s možnosťou ďalšej konfigurácie. Pre potreby diplomovej práce bol zvolený actor Web Scraper, ktorý umožňuje načítanie dát z webových

stránok prostredníctvom definície HTML a CSS prvkov skrz JavaScript kód. Okrem toho obsahuje tiež funkciu web crawlingu a tak umožňuje prehľadávanie webu s cieľom získavania odkazov na ďalšie stránky, ktorých odkazy sa na zadanej stránke nachádzajú. Tie sú následne zaindexované a zahrnuté medzi stránky určené k spracovaniu.

Dôležité konfigurácie aktora „Web Scraper“ v aplikácii Apify (Apify 2023):

- Basic configuration
 - Start URLs:

Všetky webové podstránky na tému Ukrajina, z ktorých boli fake news vyberané, obsahujú na konci stránky pagináciu. Vďaka nej bolo možné pri extrakcii zvoliť tie čísla podstránok, napr. <https://sita.sk/tag/rusko-ukrajinsky-konflikt/page/20/>, ktoré obsahovali výpis článkov zodpovedajúci určenému časovému obdobiu. Konfigurácia Start URLs tak obsahovala zoznam podstránok, ktoré v čase extrakcie obsahovali zoznam článkov vyhovujúci danému časovému obdobiu. Prípadný výskyt článkov mimo tohto časového obdobia bolo následne opravený manuálnym odstránením z datasetu.

Webové stránky v Start URL predstavujú iniciálne stránky, ktoré Web Scraper Actor navštívi a podľa definície ďalších konfigurácií ďalej prehľadáva.
 - Link Selector:

Link selector má obsahovať značenie prvku webovej stránky, ktorý obsahuje odkaz na ďalšie webové stránky, ktoré má Web Scaper Actor navštíviť. V našom prípade išlo o označenie odkazov na podstránky obsahujú samotný obsah príspevku a teda konkrétny článok fake news, prípadne real news.
 - Page Function:

Pole Page Functuoan obsahuje JavaScript funkciu spúšťanú na každej navštívenej stránke. Pomocou tejto funkcie je z obsahu stránky sťahovaný jej obsah odkazom na HTML elementy a prípadne získavané ďalšie odkazy na stránky. Návratovou hodnotou funkcie sú objekty reprezentujúce dáta extrahované zo stránky. V našom prípade šlo o elementy v ktorom sa nachádzal nadpis dátum publikovania článku, perex a textový obsah príspevku, tak ako boli popísané v podkapitole 6.3 Štruktúra datasetu.
- Performance and limits
 - Max crawling depth:

Max crawling depth označuje hĺbku vnorenia web crawlingovej funkcie od iniciálnej stránky, definovanej v konfigurácii Start URLs. Nakoľko bol v Link Selectore definovaný odkaz vedúci priamo na stránky konkrétnych článkov, nebolo nutné prehľadávať aj odkazy obsiahnuté na týchto stránkach. Preto bola v konfigurácii Max Crawling Depth zvolená hodnota 1. Link Selector je tak aplikovaný len na iniciálnych stránkach podľa Start URLs a takto vybrané stránky sú už predmetom spracovania pomocou Page Function. Ďalšie prehľadávanie odkazov sa na týchto podstránkach už nevykonáva.

Ostatné konfigurácie Web Scraper Aktora zostali bez zmien a boli aplikované v základnej podobe, preto popis konfigurácie zahŕňa len podmnožinu možných nastavení Web Scraper

Actora. Pri extrakcii nebola žiadnym spôsobom limitovaná maximálna dĺžka sťahovaných polí.

Príklad Page Function a selekcie atribútov, využitých pri extrakcii dát z webových stránok tlačovej agentúry SITA:

```
async function pageFunction(context) {  
    // jQuery is handy for finding DOM elements and extracting data from them.  
    const $ = context.jQuery;  
    const nazov = $('.entry-title').text();  
    const datum = $('.entry-meta-date').first().text();  
    const text = $('p').not(':first').text();  
    // Print some information to actor log  
    context.log.info(`URL: ${context.request.url}, TITLE: ${pageTitle}`);  
    // Return an object with the data extracted from the page.  
    // It will be stored to the resulting dataset.  
    return {  
        url: context.request.url,  
        nazov,  
        datum,  
        text  
    };  
}
```


6.4.1 Zem & Vek

Stránka Zem a Vek neobsahuje sekciu zameranú na konflikt na Ukrajine. Možné je však filtrovať články podľa kľúčových slov, ktorých výpis sa nachádza v bočnom menu a v ktorých je zahrnuté aj kľúčové slovo „Ukrajina“. Príklad podstránky s výpisom článkov na tému Ukrajina je zobrazený na Obrázok 3. Z tejto podstránky boli vybrané jednotlivé články pre dataset. V nástroji Apify predstavovali stránky z Obrázok 3 vstup do konfigurácie Start URLs.

ZEM
Podporte nezávislé spravodajstvo Zem&Vek pozitívne osobnosti Moje Konto f


Najnovšie správy Z domova Zo zahraničia Rozhovory a komentáre Ekonomika Video Magazin Časopis E-shop

Ukrajina



SPRÁVY | TOP SPRÁVY | UDALOSTI | ZO ZAHRANIČIA
Pokrytecký Západ, ako vždy, klamal: Ukrajinu už pripravuje na členstvo v NATO
Na Donbase sa pomaly končí prvý rok rusko-ukrajinského ozbrojeného konfliktu. „Nepriateľ naďalej zameriava svoje hlavné úsilie na ofenzívu v smeroch Kupjansk, Liman, Bachmut, Avdejevka a Šachtar. Ukrajinské obranné sily odrazili útoky v okolí obcí Kuzemovka a Dibrova v Luhanskej oblasti a ...
[Čítať ďalej](#)






23.02.2023



SPRÁVY | TOP SPRÁVY | UDALOSTI | ZO ZAHRANIČIA
Dokonce aj v Kyjeve si všimli „nediplomatické“ výroky slovenského ministra zahraničia na adresu Orbána
Na Donbase ruská armáda pomalým tempom napreduje na širokom fronte. Podľa ukrajinského Generálneho štábu, hlavné boje sa odohrávajú v okolí Novobachmutovky, Nevskeho, Bachmutu, Vodaného a Marjinky, ukrajinská armáda odráža útoky v okolí Grjanikovky a Masjutovky v Charkovskej oblasti, v okolí Belogorovky v Luhanskej oblasti a v okolí Vasjukovky, ...
[Čítať ďalej](#)

22.02.2023


NAJNOVŠIE SPRÁVY

-  Česko uvalilo vlastné sankcie na moskovského patriarchu Kirilla
30.04.2023
-  Ukrajinské obilie, ktoré malo zabrániť hladomoru v Afrike, v skutočnosti zaplavilo Európu
30.04.2023
-  Volodymyr Zelensky predpovedal, že vojna na Ukrajine môže pokračovať roky alebo dokonca desaťročia
30.04.2023
-  FAŠIZMUS PORAZENÝ NEBOL, SKRYTO HO AJ NAĎALEJ KOORDINUJE OSN
29.04.2023
-  Mária Bartalos sa opovážila vzoprieť liberálnej propagande
29.04.2023

SLEDUJTE ZEM&VEK

f

NAJNOVŠIE VIDEÁ

-  Učiteľka, ktorá hovorila s deťmi o

Obrázok 3 Sekcia "Ukrajina" na stránke Zem a Vek. Zdroj: Zem a Vek

Detail samotného príspevku je zobrazený na Obrázok 4. Môžeme vidieť že detail príspevku obsahuje nadpis článku, pod ním názov a autora článku. Autor článku však nebol pre dataset vyťažovaný, nakoľko sa zvyčajne jednalo o pár opakujúcich sa osôb, prípadne články obsahovali ako autora uvedený názov stránky Zem a Vek, či nebol uvedený vôbec. Na konci každého príspevku bolo uvedené tiež vyhlásenie o pozbavení zodpovednosti vydavateľstva Sofian, s.r.o. za obsah článku a tiež upozornenie, aby čitatelia do diskusie na webovej stránke nekladali také komentáre, ktoré by mohli naplniť skutkovú podstatu niektorého trestného činu uvedeného v Trestnom zákone. Vyhlásenie a upozornenie nie sú súčasťou textového obsahu spracovaného v datasete.

Pokrytecký Západ, ako vždy, klamal: Ukrajinu už pripravuje na členstvo v NATO

23.02.2023 · AUTOR: EUGEN RUSNÁK



Foto: Strana.ua

Na Donbase sa pomaly končí prvý rok rusko-ukrajinského ozbrojeného konfliktu. „Nepriateľ naďalej zameriava svoje hlavné úsilie na ofenzívu v smeroch Kupjansk, Liman, Bachmut, Avdejevka a Šachtar. Ukrajinské obranné sily odrazili útoky v okolí obcí Kuzemovka a Dibrova v Luhanskej oblasti a Fedorovka, Jagodnoje, Berchovka, Bachmut a Kurdumovka v Doneckej oblasti. Za posledný deň ruskej jednotky vykonali 59 útokov z raketometov,“ citovali správy ukrajinského Generálneho štábu domáci novinári.

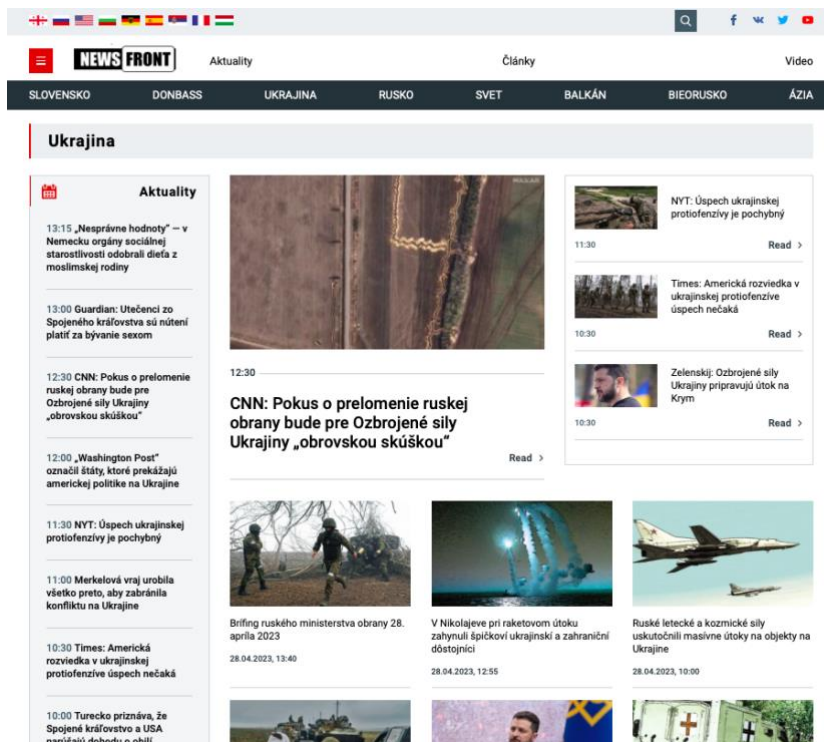
Krátko po polnoci zverejnili aj reakciu Budapešti na veľmi „nediplomatické“ výroky slovenského šéfa rezortu diplomacie. „Šéf maďarského parlamentného výboru pre zahraničné veci a politickú stranu Fidesz Zoltán Németh poradil slovenskému ministrovi zahraničných vecí Rastislavovi Káčerovi, aby navštívil psychiatra kvôli verejnej urážke premiéra Viktora Orbána. Németh túto svoju radu poskytol počas brífingu. Podľa slov politika, osobne sa pozná s Káčerom a dobre ho pozná, keďže ten pôsobil, ako veľvyslanec Slovenska v Budapešti. „Musím povedať, že som veľmi znepekajúcej jeho ďalším zdávaním. A odpovedá by som mu, aby veľmi urgentne navštívil psychiatra,“ podotkol Németh. Politik dodal, že maďarská vláda odmieta všetky Káčerové obvinenia a predvoláva slovenského veľvyslancu na vysvetlenie,“ uviedli ukrajinskí novinári.

Ukrajinskí vojenski analytici oznámili, že USA dodávajú Ukrajine modifikovanú tmúnicu, vyrobenú spoločnosťou

Obrázok 4 Článok na stránke Zem a Vek. Zdroj: Zem a Vek

6.4.2 News front SK

Stránka New front SK obsahuje priamo v hlavnom menu sekciu zameranú na tému Ukrajina. Ukážka tejto sekcie je zobrazená na Obrázok 5 a stránky sekcie zodpovedali vstupu Start URLs v nástroji Apify.



Obrázok 5 Sekcia "Ukrajina" na stránke News front SK, Zdroj: News front SK

Obsah samotného príspevku je zobrazený na Obrázok 6. Obsahuje hlavný názov článku, pod ktorým je uvedený dátum a čas publikácie. Nakoľko čas nie je súčasťou definovanej štruktúry datasetu, ale nachádza sa v textovom poli spoločne s dátumom, bola táto informácia o čase publikovania odstránená z poľa dátum až po extrakcii dát z webovej stránky. Informácie o autoroch sa na stránke nenachádzajú.

Súčasťou textového obsahu je okrem hlavného textu tiež perex, obsahujúci stručné zhrnutie textu, nachádzajúci sa na stránke spravidla nad obrázkom. Perex bol pri extrakcii spracovaný spolu s hlavným textom článku a nachádza sa tak v datasete nazačiatku poľa text. Články obsahujú často citácie od „expertov“, ktoré sú vyčlenené vizuálne v špeciálnom textovom poli. Tieto citácie, keďže sú dôležitou súčasťou zdelenia správy, sú taktiež súčasťou textového obsahu príspevku a to v takom poradí, v akom sú zakomponované v okolitom texte.

Kyjev chystá provokáciu s mŕtvymi ľuďmi v Snigirevke

18.02.2023 10:00

Share     

Kyjevský režim pripravuje provokáciu v Snigirevke, pričom plánuje predstaviť civilistov zabitých v dôsledku ukrajinského ostreľovania ako „obete ruskej agresie“.



Pre agentúru RIA Novosti to povedal Jurij Barbašov, šéf ruskej administratívy Snigirevského okresu v Chersonskej oblasti.



„Dôvodom bol rozhovor s regionálnym prokurátorom Nikolajeva, ktorý upozornil na exhumáciu 27 tel civilistov, ktorí zomreli, keď bolo mesto pod kontrolou ruskej armády,“ povedal.

Barbašov zdôraznil, že ukrajinská armáda osem mesiacov „úmyselne zabíjala obyvateľov Snigirevky“ a teraz ich chce prezentovať ako obeť ruskej armády.



„Chcem oficiálne vyhlásiť, že všetci zadržaní pre podozrenie zo spolupráce s Ozbrojenými silami Ukrajiny, s ktorými boli počas prítomnosti ruských bezpečnostných zložiek v Snigirevke vykonávané vyšetrovacie úkony, boli prepustení bez zranení, čo ukrajinské orgány činné v trestnom konaní aj miestni obyvatelia veľmi dobre vedia,“ dodal Barbašov.

Obrázok 6 Článok na stránke News front SK, Zdroj: News front SK

6.4.3 Slovenská tlačová agentúra SITA

Aj keď v hlavnom menu na webových stránkach agentúry SITA sa špeciálna sekcia pre články ohľadne Ukrajiny nenachádza, SITA uvádza rubriku pre aktuálne témy, kde možno nájsť aj tému „ Rusko-ukrajinský konflikt.“ Táto sekcia, viz. Obrázok 7, je filtrom pre získanie všetkých aktualít a príspevkov na tému ruskej agresie na Ukrajinu, ktoré boli vydané vo vymedzenom období. Nakoľko sú správy tlačových agentúr zdrojom pre ostatné médiá v krajine, možno predpokladať, že obsahuje všetky dôležité správy o udalostiach, ktoré sa v tomto období udiali a ktoré boli v serióznych médiách na Slovensku ďalej šírené a to v ich pôvodnom znení.

reklama

RUSKO-UKRAJINSKÝ KONFLIKT

14:42 | Čaputová rokovala na Ukrajine o obnove krajiny, považuje to za príležitosť pre slovenské firmy (video) ●



Podľa prezidentky sa voči utrpeniu ľudí nesmieme obracať chrbtom.

reklama

10:44 | Čaputová sa v Kyjeve stretla so Zelenským a vyjadrila mu podporu, s českým prezidentom Pavlom navštívila aj Boroďanku (video+foto) ●



Spoločná cesta oboch prezidentov podľa Zuzany Čaputovej potvrdzuje novú éru vzťahov medzi slovenskou a českou hlavou štátu, pričom je to ich prvá spoločná zahraničná cesta.

NAJČÍTANEJŠIE 24h | 48h | 72h | 168h

- 1 | Polícia upozorňuje na dopravné obmedzenia, ktoré začnú od 1. mája platiť na diaľnici D2
- 2 | Slovensko v semifinále MS v hokeji do 18 rokov podľahlo USA, Smith strelil hetrik a skóroval aj Dvorský (video) ●

Obrázok 7 Sekcia "rusko-ukrajinský konflikt" na stránke SITA, Zdroj: SITA

Články na stránkach agentúry SITA, viz. Obrázok 8, obsahujú štandardne nadpis, dátum a čas publikovania. Pod každým článkom sa nachádza aj informácia o autorovi článku. Z uvedených informácií sú pre potreby datasetu spracované len polia nadpis a dátum. Články neobsahujú prex, preto je textový obsah správ zhodný s hlavným textom článkov.

Dve ruské rakety zasiahli centrum ukrajinského mesta Kramatorsk, pri útoku sa zranilo najmenej päť ľudí

02. 02. 2023 | 19:34 | Ukrajina | Aktuálne správy | Aktuálne správy z lokality Ukrajina



Kráter po ruskom raketovom útoku na ukrajinský Kramatorsk.

Foto: SITA/AP



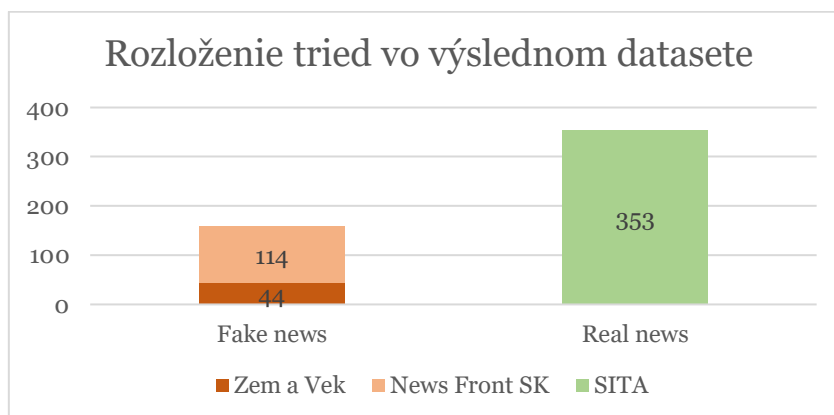
Najmenej päť ľudí utrpelo vo štvrtok zranenia v ukrajinskom meste Kramatorsk v Doneckej oblasti, keď centrum mesta zasiahli dve rakety vystrelené ruskou armádou. Uviedol to gubernátor oblasti Pavlo Kyrylenko.

Podľa ukrajinského úradu generálnej prokuratúry ruské sily na Kramatorsk zaútočili raketami S-300.

Obrázok 8 Článok na stránke SITA, Zdroj: SITA

6.5 Ukážka a popis výsledného datasetu

Vytvorený dataset obsahuje celkovo 511 príspevkov, z ktorých je 353 príspevkov typu Real, z webovej stránky slovenskej tlačovej agentúry SITA a 158 príspevkov typu Fake zo stránok Zem a Vek a Front News SK. Rozloženie výsledných tried v datasete podľa zdrojovej stránky je zobrazené na Graf 1.



Graf 1 Rozloženie tried vo výslednom datasete. Zdroj: autor

Obrázok 9 a Obrázok 10 zobrazujú ukážku 15 záznamov datasetu typu Real a 15 záznamov typu Fake. Textový rozsah fake news býva skutočne rozmanitý, nakoľko v niektorých prípadoch články Zem a Vek obsahujú rozsiahle analýzy aktuálneho diania, zatiaľ čo niektoré príspevky z webovej stránky News Front SK majú rozsah len pár viet.

Pri extrakcii nebol žiadnym spôsobom menený textový obsah polí. Dataset obsahuje diakritiku, slová v pôvodnom tvare, interpunkčné znamienka a pod., nakoľko pri detekcii fake news na základe štruktúry a štýlu, skúmaním sémantiky, syntaxu, či použitého senzacionalizmu, môžu byť aj tieto vlastnosti textu určujúce pri odlišení fake news príspevkov. Dataset tvorí prílohu tejto diplomovej práce (Príloha A: Dataset pre detekciu fake news)

URL	Dátum	Nadpis	Text	Pravdivosť
https://sita.sk/ukrajina-musi-vyhrat-tuto-vojnu-vyh	23.1.2023	Ukrajina musí vyhrať túto vojnu, vyhlásil Borrell a EÚ pos	Európska únia (EÚ) v pondelok dosiahla dohodu o poskytnutí vojenskej	Real
https://sita.sk/madarsko-nepovažuje-dodávky-zbra	23.1.2023	Maďarsko nepovažuje dodávky zbraní Ukrajine, ani stovky	Maďarsko nebude blokovat plánovanú tranžu európskej pomoci Ukra	Real
https://sita.sk/lotyšsko-znižuje-diplomaticke-vzťah	23.1.2023	Lotyšsko znižuje diplomatické vzťahy s Moskovou, reaguje	Lotyšsko znížilo úroveň svojich diplomatických vzťahov s Ruskom. V po	Real
https://sita.sk/v-nelegálnej-tabakovej-fabrike-vyko	23.1.2023	v nelegálnej tabakovej fabrike vykoristovali Ukrajincov, p	V nelegálnej tabakovej továrni v severnom Španielsku vykoristovali u	Real
https://sita.sk/polsko-požiada-nemecko-o-povolenie	23.1.2023	Polsko požiada Nemecko o povolenie poslať tanky Leopard	Polsko požiada Nemecko o povolenie poslať na Ukrajinu tanky Leopa	Real
https://sita.sk/moskva-vypovedala-estonskeho-velv	23.1.2023	Moskva vypovedala estónskeho veľvyslanca, ide o odvetu	Rusko vypovedalo estónskeho veľvyslanca a misiu krajiny v Moskve b	Real
https://sita.sk/nemecka-vlada-nebude-mat-namietk	23.1.2023	Nemecká vláda nebude mať námietky, ak Poľsko pošle U	Nemecká vláda nebude mať námietky, ak sa Poľsko rozhodne poslať	Real
https://sita.sk/usa-su-pripravené-schválit-dodanie	24.1.2023	USA sú pripravené schváliť dodanie tankov M1 Abrams	Administratíva amerického prezidenta Bidena je údajne pripravená s	Real
https://sita.sk/nemecka-armada-ma-320-tankov-le	24.1.2023	Nemecká armáda má 320 tankov Leopard 2, ale nie je zn	Nemecká armáda vlastní 320 tankov Leopard 2. Informuje o tom spr	Real
https://sita.sk/finsky-prezident-sauli-niinistö-sa-st	24.1.2023	Fínsky prezident Sauli Niinistö sa stretol s Volodymyrom	Fínsky prezident Sauli Niinistö sa v utorok v Kyjeve stretol so svojim u	Real
https://sita.sk/rusko-vraj-opat-meni-klucovych-pre	24.1.2023	Rusko vraj odvolalo jedného z kľúčových veliteľov na Ukra	Rusko podľa britskej spravodajskej služby urobilo ďalšiu zmenu vo ve	Real
https://sita.sk/nemci-uz-maju-oficialnu-ziadost-od	24.1.2023	Nemci už majú oficiálnu žiadosť od Poľiakov. Ide o bezpe	Poľsko oficiálne požiadalo Nemecko o povolenie transferu jeho tanko	Real
https://sita.sk/rusko-oznaciło-nadáciu-pomenovan	24.1.2023	Rusko oznaciło nadáciu pomenovanú po laureátovi Nobel	Ruské úrady vyhlásili nadáciu pomenovanú po laureátovi Nobelovej c	Real
https://sita.sk/rusi-posielajú-na-ukrajinu-nevzbav	24.1.2023	Rusi posielajú na Ukrajinu nevzbavených a zle vycvičených	Rusko za pár uplynulých mesiacov poslalo desaťtisíce nových vojakov	Real

Obrázok 9 Ukážka datasetu s príspevkami typu Real, Zdroj: autor

URL	Dátum	Nadpis	Text	Pravdivosť
https://zemavek.sk/zakladatel-wagnera-vysvetluje	23.1.2023	Zakladateľ Wagnera vysvetľuje nepriateľstvo USA	Jevgenij Prigožin pre RT povedal, že ruská firma je „vicekomandom“	Fake
https://zemavek.sk/fake-news-komplet-nad-skritiz	23.1.2023	„FAKE NEWS! KOMPLETI!“ Naď skritizoval mainstreamov	„FAKE NEWS! KOMPLETI!“ Napísal Jaroslav Naď na sociálnej sieti a d	Fake
https://sk.news-front.info/2023/01/23/expert-pov	23.1.2023	Expert povedal, koľko času potrebujú ukrajinskí tankisti n	Ukrajinskej armáde bude trvať tri až štyri mesiace, kým zvládne obsi	Fake
https://sk.news-front.info/2023/01/23/politico-po	23.1.2023	Politico: Pôžičky EÚ privedú Ukrajinu k hospodárskej kata	Pôžičky EÚ pre Kyjev, o ktorých európski lídri tak radi hovoria, privedú	Fake
https://sk.news-front.info/2023/01/23/sily-sa-min	23.1.2023	„Sily sa miňajú“ — generál o bojovom potenciáli Ozbroje	pojené štáty sa poriadne prepočítali, keď vlastnými rukami natiáhli t	Fake
https://sk.news-front.info/2023/01/23/zelenkij-n	23.1.2023	Zelenkij: Niekoľko desiatok tankov situáciu nezmení	rajinský prezident Vladimir Zelenskij vyhlásil, že dodávka niekoľkých c	Fake
https://zemavek.sk/hegerova-vlada-biochemicke-l	24.1.2023	Hegerova vláda: Biochemické laboratória sú na Ukrajine	Ovzdušná vláda stojí za Zelenským. Dočasná vláda premiéra Eduarda	Fake
https://sk.news-front.info/2023/01/24/vojensky-ex	24.1.2023	Vojenský expert: Dodávky „Leopardov“ priebeh špeciálne	Na pozadí neustálych správ o možných dodávkach nemeckých tankov	Fake
https://sk.news-front.info/2023/01/24/ukrajinsky-v	24.1.2023	Ukrajinský vojak natočil video na rozlúčku	Militant Ozbrojených síl Ukrajiny so slzami v očiach nahral, podľa vla	Fake
https://sk.news-front.info/2023/01/24/kyjevskaja-ct	24.1.2023	Kyjevská chunta tají straty ukrývaním tiel v chladniarskych	Orgány kyjevského režimu majú vážne problémy s pochovávaním tiel	Fake
https://sk.news-front.info/2023/01/24/banderovci	24.1.2023	Banderovci sa sťažujú na účinnosť novej ruskej taktiky	Zástupca veliteľa Hlavného riaditeľstva spravodajstva (GUR) Ministe	Fake
https://sk.news-front.info/2023/01/24/kyjevskaja-ct	24.1.2023	Kyjevská chunta dostala od EÚ miliardy	Celková výška pomoci Európskej únie Ukrajine od začiatku vojenskej	Fake
https://sk.news-front.info/2023/01/24/ruska-arma	24.1.2023	Ruská armáda oslobodila v Charkovskej oblasti od bander	Ruské sily kontrolujú jedenásť obcí v Charkovskej oblasti, kde zostalo	Fake
https://zemavek.sk/a-je-to-tu-astromocky-dlh-uk	25.1.2023	A je to tu! Astronomický dlh Ukrajiny budú musieť platiť	To že Česi a Slováci budú platiť dlh Ukrajiny , bolo včera v Bruseli p	Fake

Obrázok 10 Ukážka datasetu s príspevkami typu Fake, Zdroj: autor

6.6 Hlavné naratívy fake news obsiahnuté v datasete

Skúmaním obsahu nového datasetu boli spozorované určité naratívy vyskytujúce sa u fake news príspevkov vo vytvorenom datasete. Tieto naratívy odhaľujú kľúčové témy a tvrdenia, ktoré sú typické pre fake news v danom kontexe. Ako možno ďalej vidieť, ide o naratívy, ktorých cieľom je pripísať zodpovednosť za vojnu na Ukrajinu a „Západ.“ Hrozba nacizmu v Ukrajine, biologické zbrane a experimenty a jadrové provokácie navádzajú dojem, že Rusko a jeho „špeciálna vojenská operácia“ sú len obranným opatrením. Možno tiež pozorovať snahy o formovanie názoru na pomoc Slovenska Ukrajine, v ktorom je pomoc Ukrajine nesprávna a zadlžuje Slovensko. Tieto snahy o pomoc Ukrajine sú zároveň podľa fake news veľmi neefektívne, nakoľko Ukrajinské vojenské sily sú na sklonku porážky. Skúmaním týchto naratívov bola zistená relevantnosť datasetu, nakoľko obsahuje kľúčové

témy, sprevádzajúce ruskú dezinformačnú kampaň od začiatku invázie. Ďalej sú uvedené príklady existujúcich naratívov.

Naratív o nacistickom a skorumpovanom kyjevskom režime:

- „Mnohí Ukrajinci, ktorí sa v dôsledku totálnej mobilizácie dostali do vojnovnej zóny, nechcú bojovať s Ruskom za kyjevský režim, no tento osud ich neminie.“
- „Na Ukrajine, okrem vojenských operácií, pokračujú aj denno-denné korupčné a iné škandály na najvyššej úrovni.“
- „Zrátané a podčiarknuté: Rusko nezačalo vojnu na Ukrajine. To je falošný príbeh. Zodpovednosť leží na ukrajinskej armáde a jej vodcoch v Kyjeve.“
- „Orgány kyjevského režimu majú vážne problémy s pochovávaním tiel mŕtvych vojakov, keďže rýchlo rastúce vojenské cintoríny majú negatívny vplyv na obyvateľstvo a armádu,“
- „Kyjevský režim s pomocou Ozbrojených síl Ukrajiny a nacistov z „Azova“ pripravuje provokáciu proti Podnestersku.“

Biologické zbrane na Ukrajine:

- „Americká televízia „One America News“ (OAN) zverejnila správu o činnosti amerických biologických laboratórií na Ukrajine, v ktorej sa hovorí, že Rusi mali pravdu – za účasti Kyjeva sa skutočne vykonávajú nebezpečné biologické experimenty.“

Provokácie Západu:

- „Novinári tvrdia, že práve pokus o zakrytie stôp po tejto nezákonnej činnosti spôsobil začiatok konfliktu. „Zatiaľ čo Joe Biden pokračuje vo vyvolávaní vojny s Ruskom, pribúdajú dôkazy o tom, že celý konflikt môže byť vyprovokovaný tak, aby chránil americké vojenské laboratóriá vykonávajúce nelegálne biologické experimenty,“ poznamenal publicista OAN Pearson Sharp“
- „Ruské ministerstvo obrany varovalo pred rozsiahlou „jadrovou“ provokáciou Kyjeva“

Naratív o „špeciálnej vojenskej operácii“:

- „Ozbrojené sily Ruskej federácie pokračujú vo vedení špeciálnej vojenskej operácie, o jej priebehu informoval generálporučík Igor Konašenkov.“
- „Celkovo bolo od začiatku špeciálnej vojenskej operácie zničených...“

Apel na slovenský národ:

- „Slovenskí Judáši podporujú USA nacizmus pomocou ktorého chcú USA ovládnuť svet.“
- „Ako vidíte, Dzurinda, ktorý vlastných Slovákov zodrie z kože, aby zaplatili štátne dlhy, teraz chce, aby sme zaplatili dlhy za Ukrajinu! O koľko sa jedná!?“
- „To že Česi a Slováci budú platiť dlhy Ukrajiny, bolo včera v Bruseli potvrdené! Len výška dlhov, ktoré budeme platiť sa desaťnásobila!“

Údajné vojenské neúspechy a prichádzajúca vojenská porážka Ukrajiny:

- *„Ťažké obdobie, ktoré nastalo pre Ozbrojené sily Ukrajiny, môže v najbližších týždňoch viesť k porážke celého kyjevskeho režimu.“*
- *„Ukrajina naďalej trpí bezprecedentnými ľudskými stratami v palebnej línii. Ruské jednotky denne „kosia“ rady delostrelectva, obrnenej techniky, letectva, pechoty a vozidiel Ozbrojených síl Ukrajiny. Ruské sily už dávno musia bojovať nie s Ukrajinou, ale s kolektívnym Západom“*
- *„Ukrajina je v takých problémoch, že jej niet čo závidieť a musíme mať jasno, aké sú šance Zelenského vojakov v roku 2023. Situácia v Kyjeve je oveľa nebezpečnejšia, ako sa bežne tvrdí“*
- *„Expert zdôrazňuje, že kyjevský režim sa starostlivo snaží skrývať obrovské straty spôsobené ozbrojeným konfliktom.“*

6.7 Použitelnosť

Vytvorený dataset obsahuje príklady fake news a skutočných príspevkov v špecifickom kontexte. Je zameraný na úzko definovaný problém detekcie fake news na príklade ruskej invázie na Ukrajinu, šírených na Slovensku. Dataset obsahuje textový obsah príspevkov, nakoľko sa predpokladá použitie datasetu pri detekcii fake news na základe obsahu. Veľkosť datasetu, spolu obsahujúc 511 záznamov, by mohla byť v takto úzko definovanom probléme postačujúca a obsahovať dostatok informácií pre rozlíšenie fake news od skutočných správ. Skúmaním obsahu sa zistili isté naratívy vyskytujúce sa vo fake news, ktoré naznačujú existujúce odlišnosti v zdelení v porovnaní s reálnymi príspevkami.

Fake news príspevky v datasete sú extrahované z dvoch webových stránok, ktoré šíria fake news. Analýzou týchto stránok bolo zistené, že riziko zverejnenia fake news príspevku je u nich vysoké, čomu zodpovedá aj hodnotenie iniciatívy konspiratori.sk. Skúmanie datasetu odhalilo tiež naratívy ruskej dezinformačnej kampane, ktoré sú obsiahnuté v datasete. Všetky príspevky z fake news stránok boli označené výslednou triedou Fake a to výhradne na základe kredibility šíriaceho webu. Tento spôsob anotácie so sebou nesie istú mieru rizika výskytu „false-positive“ záznamov, teda záznamov označených ako fake news, napriek tomu, že nie sú fake news. Každá webová stránka s nízkou kredibilitou nemusí publikovať výhradne články typu fake news. Súčasne nemožno ignorovať možnosť, že niektoré skutočné záznamy v datasete, môžu byť vo skutočnosti fake news. Analýzou zdrojov pred samotnou extrakciou však bolo dokázané, že pravdepodobnosť takejto chybnéj klasifikácie je nízka.

Vytvorený dataset by bolo vhodné pri aplikácií na reálnych dátach rozšíriť o ďalšie zdroje a to ako v prípade fake news, tak skutočných príspevkov. To platí pri detekcii fake news obecné, nakoľko pri použití väčšieho počtu záznamov a použitých zdrojov sa zvyšuje v datasete množstvo informácií a špecifik, každého typu správ, ktoré môžu byť pri detekcii fake news určujúce. Vhodné je tiež dáta správne spracovať, a to podľa potreby úlohy. Často používané techniky sú stemming, tokenizácia slov, odstránenie špeciálnych znakov, interpunkcie, alebo stopových slov.

7 Princípy detekcie fake news

Na základe výstupov teoretickej časti práce, spracovanej analýzy existujúcich datasetov pre úlohy detekcie fake news a tvorby vlastného datasetu na zvolenom príklade, boli definované základné princípy, ktoré charakterizujú a popisujú detekciu fake news.

7.1 Vymedzenie úlohy detekcie fake news

V prvom kroku je vhodné zvoliť si doménu fake news, ktorá má byť predmetom detekcie. Niektoré datasety sa venujú fake news z viacerých oblastí a nebývajú tak úzko zamerané. Napríklad repozitár FakeNewsNet obsahuje záznamy z politiky a tiež zo zábavného priemyslu. Od šírky zamerania datasetu závisí potom náročnosť získavania dát. V prípade tvorby vlastného datasetu možno v prípade širokého zamerania rátať s potrebou zberu väčšieho množstva dát a využitia väčšieho počtu zdrojov, nakoľko pri príliš malom počte záznamov, či zdrojov, by nebolo možné vhodné zobecniť získané výsledky na celú skúmanú oblasť.

Dôležitá je definícia prístupu, na základe ktorého bude detekcia fake news vykonávaná, nakoľko tá ovplyvňuje štruktúru zbieraných dát. Stále častejším prístupom je kombinácia rôznych prístupov, kedy sú fake news detekované napríklad nielen na základe obsahu, ale tiež informácií o sociálnom kontexte. Pri vymedzení úlohy detekcie fake news je potrebné uvažovať tiež nad dostupnosťou potrebných dát. Prístup k API sociálnych sietí je v súčasnosti striktnnejšie obmedzený ako v minulosti. Pre ich použitie je nutné mať vytvorený a schválený vývojársky profil, tiež je potrebné podať prihlášku, vysvetľujúcu zámer spracovania dát, ktorej schválenie môže trvať aj niekoľko mesiacov, s čím treba pri detekcii rátať.

V práci sme sa venovali tiež dôležitosti vymedzenia časového obdobia, nakoľko časové obdobie patrí medzi dôležité kritéria pri výbere a hodnotení využiteľnosti a relevancie datasetov pri špecifických úlohách detekcie fake news. Ak existujúce datasety, zvažované k použitiu, budú obsahovať záznamy z dávnejšieho obdobia, nemusia vzhľadom na dynamickú povahu fake news dobre popisovať fake news vyskytujúce sa v dnešnom svete.

7.2 Dostupnosť dát a voľba zdrojov

Existuje niekoľko možností, ako získať data potrebné na zostavenie, tréning a testovanie vyhodnocovacích klasifikačných modelov. Dve primárne metódy zahŕňajú použitie existujúcich datasetov, alebo extrahovanie nespracovaných údajov z relevantných zdrojov. Každý prístup má svoje výhody a výzvy, ktoré je potrebné zvážiť pri výbere najlepšej možnosti pre konkrétny kontext a zameranie .

Jednou z možností je použitie existujúcich datasetov. Výskumníci neustále pracujú na vytváraní nových datasetov pre využitie v rôznych aplikačných doménach. Tieto súbory údajov často obsahujú označené príklady fake news a pravdivých správ, vďaka čomu sú vhodné na účely výskumu hodnotiacich modelov. Existujúce datasety sú často kvalitne popísané a bývajú prakticky otestované, vďaka čomu umožňujú jednoduchšie porovnanie modelov a tak tiež porovnanie rozličných prístupov k detekcii fake news. Kvalita existujúcich datasetov sa avšak môže naprieč jednotlivými príkladmi líšiť a preto je potrebné starostlivo posúdiť správnosť spracovaného datasetu pred jeho použitím.

Je tiež na zvážení, či tieto datasety sú vhodné a naplňajú potreby špecifického projektu, zameraného na detekciu fake news aplikáciou data miningových modelov na relevantné datasety, a či použitie týchto datasetov neovplyvní relevanciu získaných výsledkov. Nakoľko fake news môžeme nájsť spracované na rôzne témy a ich šíritelia svoje techniky postupne zlepšujú a vyvíjajú tak, aby fake news boli čo najťažšie odlišiteľné od skutočných informácií, je veľmi dôležité zhodnotiť, či použitím už existujúceho datasetu neprídeme pri detekcii o dôležité informácie, ktorých poznanie a využitie by zvyšovalo správnosť detekcie fake news.

Použitie nevhodného datasetu tak môže ovplyvniť kvalitu modelov pri následnej aplikácii na reálnych dátach. Rovnako sa použitím nevhodného datasetu nakoniec strácajú úspory na čase a na použitých zdrojoch, ktoré inak patria medzi výhody voľby existujúceho datasetu, oproti získavaniu dát vlastnými silami a tvorby vlastného datasetu.

Analýza a snahy o detekciu fake news v určitom špecifickom kontexte si tak často vyžaduje tvorbu vlastného datasetu. Záleží od konkrétnej úlohy, ktoré zdroje budú vhodné k získavaniu dát. V prípade detekcie fake news ide často o spravodajské webové stránky, dáta zo sociálnych sietí, alebo iné online zdroje, ako napríklad webové stránky organizácií hodnotiacich pravdivosť tvrdení (tzv. fact-checking organizácie). Získanie vlastných dát im umožňuje vytvorenie vlastného datasetu prispôbeného špecifickým výskumným otázkam a záujmom.

Dáta použité tak budú relevantné pre špecifický kontext úlohy, jazyk alebo tému, čím sa zabezpečí, že dataset bude vysoko reprezentatívny a použiteľný pro detekciu fake news. Výhodou získavania vlastného datasetu je tiež kontrola nad kvalitou dát. Pri vlastnom zbere dát možno spoliehať a rozhodovať o použitých spôsoboch extrakcie a spracovania dát, vďaka čomu možno mať istotu, že dataset splňuje požiadavky na kvalitu a zvyšuje tak aj spoľahlivosť výsledkov, získaných pri jeho použití.

Nevýhodou získavania vlastného datasetu je predovšetkým náročnosť na čas a použité výpočtové a ľudské zdroje. Zber dát, ich spracovanie a značenie vzhľadom na svoju náročnosť nie je tak v každom prípade správnu voľbou. Ak existuje dataset, ktorý je pri detekcii fake news použiteľný, nie je nutné pracovať na tvorbe vlastného datasetu.

Pri zbere dát vlastnými silami je potrebné tiež rátať s obmedzeniami stránok proti scrapovacím nástrojom, prípadne obmedzeniami poskytovaných aplikačných rozhraní a zvážiť dostupnosť dát k využitiu na výskumné účely.

Existujú tak dva hlavné spôsoby, aké dáta použiť pri detekcii fake news. Pri zbere dát možno použiť už existujúce datasety, ako napríklad LIAR, FakeNewsNet, alebo zber dát vykonávať vlastnými silami, extrakciou dát z primárnych zdrojov informácií, ako sú stránky publikujúce fake news, seriózne médiá, alebo informácie od fact-checkingových organizácií. Finálna voľba závisí od konkrétnych cieľov výskumu, relevantnosti dát, dostupných zdrojov a požadovanej úrovne kontroly nad kvalitou výsledného datasetu.

7.3 Určenie hodnotiacej škály

Určenie hodnotiacej škály predstavuje určenie klasifikačných kategórií, podľa ktorých sú dáta zaradené do určitého počtu výsledných tried. Pri detekcii fake news sú triedy typicky určené na základe pravdivosti, a to buď do binárnych tried, fake news a real news, alebo do tried s väčšou granularitou. Binárna klasifikácia patrí medzi najjednoduchšie rozlíšenie správ podľa pravdivosti. Záleží však od použitej definície fake news, či je takáto klasifikácia pre konkrétnu úlohu detekcie fake news postačujúca. Niektoré datasety, alebo fact-checkingové organizácie, používajú viac stupňovú škálu pre určenie pravdivosti, ako napríklad LIAR, v ktorom je použitá hodnotiacia škála so šiestimi hodnotami – pravda, prevažne pravda, polopravda, sotva pravda, nepravda a kategóriu „pants on fire“, znamenajúcu úplný nezmysel.

Ako bolo uvedené v teoretickej časti práce, fake news nemá v súčasnosti presne ukotvenú jednotnú definíciu a existujú blízke koncepty fake news ako misinformácie, false news, alebo satira. Pri určovaní nepravdivých informácií je tak možno pracovať aj s rozlíšením týchto typov nepravdivých informácií. Tie sú však vzhľadom na problematiku určenie skutočného zámeru, s ktorým sú tieto typy správ vytvárané, náročnejšie na definíciu klasifikačných kritérií a prípadnú prácu tímu klasifikujúceho tieto správy, tzv. anotátorského tímu.

Vzhľadom na fakt, že fake news sa neustále vyvíjajú, je možné, že v budúcnosti bude potrebné hodnotiace škály aktualizovať tak, aby odrážali dynamický charakter fake news ako takých.

7.4 Určenie spôsobu klasifikácie

Pre určenie výslednej triedy správ v datasete, na základe definovanej klasifikačnej škály existuje niekoľko metód. Častým spôsobom, ktorý je v analyzovaných datasetoch používaný, je použitie znalostí fact-checkingových organizácií a výstupov analýz prístupných na ich webových stránkach. Experti v danej doméne hodnotia pravdivosť uvedených tvrdení na základe vlastných znalostí a analýz. Príkladom výstupov takýchto iniciatív sú napríklad analýzy na webových stránkach polifact.com, euvsdisinfo.eu, na Slovensku napríklad infosecurity.sk, alebo česká webová stránka manipulatori.cz a mnoho ďalších. Datasety vychádzajúce z analýz fact-checkingových organizácií majú prirodzene vyššiu kvalitu, nakoľko sú založené na expertnej analýze.

V tejto diplomovej práci bola použitá pre určenie pravdivosti príspevkov analýza iniciatívy konspirtori.sk, ktorá analyzuje webové stránky publikujúce fake news a na základe toho prideluje hodnotenie celému webu. Aj keď takáto analýza vychádza z hodnotení príspevkov publikovaných na webových stránkach, nie je zaručené, že každý jeden príspevok je fake news. Určenie klasifikácie na základe kredibility šíriteľa je ďalšou z možností tvorby datasetu, avšak môže viesť k chybám v datasete, nakoľko nie každý príspevok musí byť typu fake news. V analyzovaných datasetoch je preto táto metóda často kombinovaná s manuálnou kontrolou aspoň podmnožiny datasetu, vďaka čomu je možné zistiť spoľahlivosť tejto klasifikácie a prípadne odhaliť vzniknuté chyby a ich početnosť.

Podobný prístup je tzv. crowdsourcing, kedy je pravdivosť tvrdení určená na základe hodnotenia veľkej skupiny ľudí, nie nutne zloženej len z expertov. Finálna trieda určitej správy môže byť určené napríklad na základe triedy, ktorá má najčastejší výskyt. U crowdsourcingu je klasifikácia dát rýchlejšia, avšak kvalita výsledného datasetu môže byť nižšia.

V oboch prípadoch je dôležité, aby bol postup pre klasifikovanie správ presne definovaný. Tiež je v prípade potreby vhodné určiť spôsob nakladania so správami, u ktorých je určenie klasifikačnej triedy nejasné.

7.5 Definícia datasetu

V prípade tvorby vlastného datasetu je dôležité pred zberom dát definovať štruktúru datasetu. Definícia štruktúry a veľkosti datasetu závisí od vymedzenia úlohy detekcie fake news. Zbierané dáta by mali odrážať zvolený prístup k detekcii fake news. Nakoľko fake news je možné detekovať nielen na základe obsahu, ale tiež na základe šíriteľa či sociálneho kontextu, každý prístup má dopad na štruktúru datasetu, vytvoreného za účelom detekcie fake news.

Definíciou datasetu pred započatím samotnej extrakcie dát sú optimalizované zdroje a čas strávený extrakciou. Je dôležité zhodnotiť, či všetky zdroje dát pracujú s informáciami, ktoré sú v datasete definované.

8 Diskusia

Táto kapitola sa venuje zodpovedaniu výskumnej otázky ohľadne špecifik a výziev tvorby datasetu pre detekciu fake news vo zvolenom kontexte a diskusiu nad budúcnosťou detekcie fake news.

8.1 Špecifiká a výzvy tvorby datasetu v zvolenom kontexte

Vývoj nástrojov na detekciu fake news pre použitie na reálnych dátach zvyčajne vyžaduje veľké množstvo dát s označenými výslednými triedami pre tréovanie klasifikačného modelu. V tzv. „low-resource“ jazykoch, akými je aj slovenský jazyk, je dostupnosť takýchto dát obmedzená, čo považujeme za najväčšiu limitáciu obmedzujúcu ďalší vývoj v tejto oblasti. Chýbajú označované, voľne dostupné datasety, prípadne je ich veľkosť obmedzená. Detekcia fake news v slovenskom jazyku si preto vyžaduje vynaloženie veľkého úsilia na vytvorenie požadovaného datasetu.

Neexistujúca je napríklad databáza nepravdivých tvrdení. Existujúce analýzy k fake news šíreným na Slovensku od fact-checkingových organizácií sa venujú skôr analýze existujúcich naratívov, prípadne je nutné poznatky o nepravdivých tvrdeniach a ich zdrojoch hľadať priamo v texte analýzy. Spracovanie takýchto analýz do štruktúrovanej, strojovo spracovateľnej podoby by prispelo k využitiu pokročilých automatizovaných metód k odhaľovaniu nepravdivých správ na Slovensku. Pri vytváraní datasetu boli kvôli daným obmedzeniam preto triedy záznamov v datasete určené na základy kredibility šíriteľa, pričom táto metóda môže vytvárať príklady nesprávne klasifikovaných záznamov.

Detekcii fake news v slovenskom jazyku použitím automatizovaných spôsobov je dôsledkom toho venovaná oproti viac rozšíreným jazykom menšia pozornosť. V kontexte ruských dezinformačných aktivít, ktoré sú mierené vďaka polohe a spoločnej histórii primárne na „frontové štáty“ v strednej Európe, je potreba venovať sa detekcii fake news šírených v týchto štátoch. V nich sú tieto fake news intenzívne šírené a ukazuje sa, že v nich tiež majú najväčší potenciál ovplyvniť užívateľov na ktorých cieľia. Detekcia fake news šírených v týchto regiónoch môže slúžiť ako systém včasného varovania pre zvyšok sveta.

8.2 Diskusia nad budúcnosťou detekcie fake news

Dopad fake news je v súčasnosti nezanedbateľný. Ich vplyv bol preukázateľný aj pri ovplyvňovaní politických udalostí veľkého významu, kedy dezinformačné kampane zohrali rolu pri referende o vystúpení Spojeného kráľovstva Veľkej Británie a Severného Írska z Európskej únie v roku 2016, či víťazstve Donalda Trumpa v prezidentských voľbách v USA v roku 2020. Ako bolo spomenuté, dnes sú dezinformačné kampane výraznou zbraňou Ruskej federácie v snahe chybné interpretovať realitu vojenskej agresie na Ukrajine, či

v iných aktivitách zameraných na oslabenie jednoty v členských štátoch Európskej únie ako aj v USA. Jedná sa tak o významný spoločenský fenomén, ktorý má vplyv na znižovanie miery dôvery v oficiálne štátne inštitúcie, štandardné médiá a politické strany. Tento jav zároveň vedie k posilňovaniu polarizácie v spoločnosti, čo vo všeobecnosti vedie k ohrozeniu princípov tolerancie, slobody a demokracie v štáte. To podnecuje dôležitú otázku potreby regulácie sociálnych sietí v snahe eliminovať vplyv dezinformácií v informačnom prostredí štátu.

Jedná sa tak o komplexnú problematiku, ktorá zasahuje široké spektrum štátnych aj neštátnych aktérov. Problémom je aj neustále meniaci sa charakter informačného prostredia a fakt, že scéna fake news dokáže byť vysoko adaptabilná. Ako uvádza Filip Struhárik, *„dezinformátori už menej využívajú úplnú fikciu, keďže ľudia sú na ňu citlivejší a digitálne platformy už spolupracujú s factcheckermi, teda novinármi, ktorí overujú pravdivosť informácií. Evidentné lži tak vedia platformy odhaliť, zastaviť ich šírenie a ich tvorcov penalizovať. Samozrejme, robia to stále nedostatočne a potrestaný je len malý zlomok dezinformátorov. Časť konšpirátorov sa presunula z veľkých platforiem (Facebook, Youtube, Twitter) na menšie sociálne siete (Gab, Parler, Getter), kde neboli obmedzovaní a ich príspevky neboli odstraňované. Komunikujú taktiež na službách ako sú Telegram, WhatsApp či Discord.“* (Struhárik 2022a, str. 67-68) Zaujímavosťou v tomto kontexte je skutočnosť, že napr. Facebooku má pre Slovenskú republiku v roku 2023 len jedného factcheckera, ktorým je Robert Barca. Ten je oficiálne zamestnancom francúzskej spravodajskej agentúry AFP (Agence France-Presse), ktorú spoločnosť META kontraktuje na factcheckerské služby. Výsledky jeho aktivít sú zhrnuté na webstránke fakty.afp.com. V tomto ohľade je v prostredí Slovenska zaujímavý projekt Elfov, za ktorým stojí PR agentúra Katedra komunikácie. Agentúra zamestnáva tím ľudí, ktorí z diskusií na sociálnych sieťach odstraňujú dezinformácie, nenávistné príspevky a blokujú trollov v diskusiách viacerých veľkých slovenských internetových portálov. Agentúra má na tento účel na čiastočný úväzok zamestnaných 16 ľudí. Tí denne skontrolujú zhruba 15-tisíc komentárov, za jednu štvorhodinovú zmenu stihne elf prejsť 1400 až 1700 príspevkov, z čoho musia skryť alebo odstrániť zhruba 20 percent komentárov v diskusiách. (Struhárik 2022b)

Analýza organizácie (Center for Countering Digital Hate Inc 2021) robená v USA ukázala, že vo februári a v marci 2021 iba 12 osôb šírilo až dve tretiny amerického obsahu proti očkovaniu. Potvrďuje tak zistenia, že často i veľmi malá skupina osôb, tzv. superširiteľov, môže mať veľký vplyv pri šírení dezinformácií. Výzvou pre platformy je tak schopnosť efektívne limitovať aktivity práve týchto osôb. Ako poukazuje Struhárik, keďže širitelia fake news často využívajú viacero služieb, nevyhnutná je aj spolupráca medzi platformami. Zbytočne širiteľom napríklad Youtube zruší kanál, ak na Facebooku, Twitteri naďalej rastú a vďaka reklamám na Google na svojom webe aj zarábajú. Musí teda fungovať efektívna výmena informácií o širiteľoch nebezpečných správ medzi úradmi, štátmi a konkurenčnými službami. (Struhárik 2022a, str. 70)

Cieľom dezinformačných aktivít často nie je, aby publikum uverilo konkrétnej správe. Primárnym záujmom je vyvolanie všeobecnej neistoty a posilnenie dojmu, že nie je možné veriť nikomu a ničomu. Preto, častokrát ani samotný obsah takýchto správ nemusí byť vyslovene nepravdivý, ale stačí, že vyvoláva určitý dojem osobitou interpretáciou

objektívnych skutočností, či vytrhnutím informácií z kontextu. Takýto hraničný obsah sa často nazýva ako „borderline content“. Obsahuje provokatívny obsah, senzačné texty, obrázky a videa. Je to obsah, ktorý je ťažko definovateľný a charakterizovateľný ako dezinformačný, no na druhej strane efektívne slúži účelu vytvárať informačný chaos či polarizovať spoločnosť. Takýto obsah taktiež predstavuje osobitnú výzvu do budúcnosti.

Uvedené informácie tak len podčiarkujú skutočnosť, že dezinformácie predstavujú hrozbu so širokým dopadom na spoločnosť, no voči ktorej bojovať je veľmi problematické vzhľadom na dynamicky sa vyvíjajúcu sa povahu informačného prostredia a dezinformačných aktivít ako takých. Ide o problém, ktorý je svojim charakterom komplexný a prekračuje hranice jednotlivých štátov, a preto je pravdepodobne potrebné hľadať právne riešenia aj na medzinárodnej úrovni. Jednou z takýchto snáh je legislatíva Európskej únie s názvom Digital Services Act (DSA), ktorá bola schválená Európskym parlamentom v roku 2022. Táto legislatíva má ambíciu regulovať veľké internetové spoločnosti, vrátane firiem ako Google, META či Amazon. *„Na veľmi veľké on-line platformy sa budú z dôvodu výrazného rizika v oblasti šírenia nezákonného a škodlivého obsahu vzťahovať osobitné povinnosti. Podľa europoslancov DSA má pomáhať v boji proti škodlivému (nie nevyhnutne nezákonnému) obsahu a šíreniu dezinformácií, a to ustanoveniami o povinných posúdeniach rizika, opatreniach na zmiernenie rizika, nezávislých auditoch a transparentnosti algoritmov, ktoré určujú, čo používatelia vidia.“* (TASR 2022) Ukazuje sa tým trend posúvania riešenia tohto problému na nadnárodnú úroveň, vyžadujúcu medzinárodnú spoluprácu, čo radí problém boja s dezinformáciami medzi iné globálne výzvy 21. storočia ako klimatická kríza či energetická bezpečnosť.

Záver

Táto diplomová práca sa zaoberá princípmi detekcie fake news a ich uplatnením na príklade tvorby datasetu pre fake news šírené o ruskej invázii na Ukrajinu v roku 2022, šírené v prostredí slovenského online priestoru. Hlavným cieľom práce bolo tieto princípy popísať a využiť pri praktickom príklade tvorby datasetu.

Práca začína zdôvodnením výberu témy, popisom relevancie danej problematiky a priblížením sociálneho a politického kontextu, v súčasnosti bohužiaľ stále prebiehajúcej ruskej agresie na Ukrajine. Dôležitosť vybranej témy spočíva v tom, že táto udalosť je sprevádzaná nielen pozemnou intervenciou zo strany Ruskej federácie na Ukrajinu, ale ide o hybridnú vojnu sprevádzanú tiež masívnou dezinformačnou kampaňou.

V teoretickej časti práce sú popísané dve najčastejšie používané definície fake news, nakoľko pre fake news v súčasnosti neexistuje jednotná definícia. Práca popisuje spôsoby šírenia fake news, základné atribúty fake news správy a vlastnosti popisujúce charakter fake news v dnešnom svete. Z atribútov fake news potom vychádza rozdelenie prístupov k detekcii fake news na detekciu na základe obsahu, šíriteľa a sociálneho kontextu.

V praktickej časti práce, zameranej na definíciu princíпов detekcie fake news, je spracovaná analýza existujúcich datasetov, používaných pri detekcii fake news a ich porovnanie na základe dôležitých charakteristík určujúcich fake news detection datasety (FNDD). Následne je v práci popísaný postup tvorby vlastného datasetu. Vytvorený dataset je zameraný na detekciu fake news šírených v slovenskom jazyku na tému ruskej invázie na Ukrajinu. Poslednou časťou práce je vymedzenie princíпов detekcie fake news a diskusia nad obmedzeniami detekcie fake news vo zvolenom príklade.

Naplnenie cieľov práce, definovaných vo výskumných otázkach, bolo naplnené vďaka analýze existujúcich datasetov pre detekciu fake news a praktickému príkladu tvorby datasetu vo zvolenom kontexte.

Medzi diskutované princípy detekcie fake news patrí vymedzenie úlohy detekcie fake news, určenie hodnotiacej škály, spôsobu klasifikácie, analýza dostupných zdrojov a definícia datasetu k detekcii fake news. Rovnako boli v práci popísané špecifiká týkajúce sa tvorby datasetu a poťažmo detekcie fake news vo zvolenom kontexte, ktoré vyplývajú hlavne z obmedzeného množstva dostupných, klasifikovaných dát o fake news v slovenskom jazyku a malému záujmu o problém medzi výskumníkmi z IT oblasti.

Práca popisom princíпов a praktickou tvorbou datasetu tak prispieva k pochopeniu úlohy detekcie fake news vo zvolenom špecifickom kontexte a jazyku a potenciálne prispieva tiež k využitiu týchto znalostí pri úspešnom odhaľovaní fake news. Dôležitým produktom práce je aj vytvorený dataset, ktorý obsahuje 511 záznamov. Práca obsahuje diskusiu nad použiteľnosťou datasetu pri aplikácii na reálnych úlohách detekcie fake news.

Použitá literatúra

AGENTÚRA SITA, 2023. Agentúrne spravodajstvo. *Agentúra SITA* [online] [vid. 2023-04-23]. Dostupné z: <https://biz.sita.sk/agenturne-spravodajstvo/>

AKTUALITY.SK, 2019. Komentár Petra Bárdyho: Rostas je konšpirátor, extrémista a antisemita. *Aktuality.sk* [online] [vid. 2023-04-22]. Dostupné z: <https://www.aktuality.sk/clanok/749695/komentar-petra-bardyho-rostas-je-konspirator-extremista-a-antisemita/>

ALLCOTT, Hunt a Matthew GENTZKOW, 2017. Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives* [online]. **31**, 211–236. Dostupné z: doi:10.1257/jep.31.2.211

AMADO, Bárbara G., Ramón ARCE a Francisca FARIÑA, 2015. Undeutsch hypothesis and Criteria Based Content Analysis: A meta-analytic review. *The European Journal of Psychology Applied to Legal Context* [online]. **7**(1), 3–12. ISSN 18891861. Dostupné z: doi:10.1016/j.ejpal.2014.11.002

APIFY, 2023. Web Scraper · Apify. *Apify* [online] [vid. 2023-04-30]. Dostupné z: <https://apify.com/apify/web-scraper>

ARGUEDAS, Amy Ross, Craig T. ROBERTSON, Richard FLETCHER a Rasmus KLEIS NIELSEN, 2022. Echo chambers, filter bubbles, and polarisation. *Reuters Institute for the Study of Journalism* [online] [vid. 2022-10-23]. Dostupné z: <https://reutersinstitute.politics.ox.ac.uk/echo-chambers-filter-bubbles-and-polarisation-literature-review>

BALINT, Kata, Francesca ARCOSTANZO, Jordan WILDON a Kevin REYES, 2022. RT Articles are Finding their Way to European Audiences – but how? *ISD* [online] [vid. 2022-10-28]. Dostupné z: https://www.isdglobal.org/digital_dispatches/rt-articles-are-finding-their-way-to-european-audiences-but-how/

BENČÍK, Ján, 2016. Blog N: Ako je to s údajnou nezávislosťou Tibora Rostása. *Denník N* [online] [vid. 2023-04-22]. Dostupné z: <https://dennikn.sk/blog/378330/ako-je-to-s-udajnou-nezavislostou-tibora-rostasa/>

BUCHANAN, Tom, 2020. Why do people spread false information online? The effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation. *PLOS ONE* [online]. **15**, e0239666. Dostupné z: doi:10.1371/journal.pone.0239666

CAMBRIDGE DICTIONARY, 2022. news. *Cambridge University Press* [online] [vid. 2022-10-02]. Dostupné z: <https://dictionary.cambridge.org/dictionary/english/news>

CASTILLO, Carlos, Marcelo MENDOZA a Barbara POBLETE, 2011. Information credibility on twitter. In: *WWW '11: 20th International World Wide Web Conference: Proceedings of the 20th international conference on World wide web* [online]. Hyderabad India: ACM, s. 675–684 [vid. 2023-02-25]. ISBN 978-1-4503-0632-4. Dostupné z: doi:10.1145/1963405.1963500

CENTER FOR COUNTERING DIGITAL HATE INC, 2021. *The Disinformation Dozen* [online]. B.m.: Center for Countering Digital Hate Inc [vid. 2023-05-01]. Dostupné z: <https://counterhate.com/research/the-disinformation-dozen/>

COUNCIL OF THE EU, 2022. EU imposes sanctions on state-owned outlets RT/Russia Today and Sputnik's broadcasting in the EU. *European Council* [online] [vid. 2022-10-28]. Dostupné z: <https://www.consilium.europa.eu/en/press/press-releases/2022/03/02/eu-imposes-sanctions-on-state-owned-outlets-rt-russia-today-and-sputnik-s-broadcasting-in-the-eu/>

ČOREJ, Tomáš a MESEŽNIKOV, 2022. Mesežnikov: Putin má tri obsesie, ktoré viedli k útoku na Ukrajinu. *Denník N* [online] [vid. 2023-05-01]. Dostupné z: <https://dennikn.sk/2929997/meseznikov-ci-sa-obavam-ze-putin-zacne-jadrovu-vojnu-hrozi-nebezpecna-hra-s-lukasenkom/>

DAVOUDI, Mansour, Mohammad R. MOOSAVI a Mohammad Hadi SADREDDINI, 2022. DSS: A hybrid deep model for fake news detection using propagation tree and stance network. *Expert Systems with Applications* [online]. **198**, 116635. ISSN 0957-4174. Dostupné z: [doi:10.1016/j.eswa.2022.116635](https://doi.org/10.1016/j.eswa.2022.116635)

DINIČ, Dávid, 2022. Proruskí dezinformační aktéri útočia na Európsku úniu, nechýbajú ani tvrdenia o fašizme v Bruseli (Infosecurity.sk: Bi-weekly report o nastupujúcich dezinformačných trendoch 30. september 2022) [online]. *INFOSECURITY.SK* [online]. [vid. 2023-02-22]. Dostupné z: <https://infosecurity.sk/dezinfo/proruski-dezinformacni-akteri-utocia-na-europsku-uniu-nechybaju-ani-tvrdenia-o-fasizme-v-bruseli-infosecurity-sk-bi-weekly-report-o-nastupujucich-dezinformacnych-trendoch-30-september-2022/>

DISINFO REPORT, 2022a. *Boj o naše srdcia a mysle. Ako môže vyzerat' hybridné pôsobenie? (podcast)* [online]. 16. duben 2022. [vid. 2023-02-22]. Dostupné z: <https://www.aktuality.sk/clanok/y58jMop/disinfo-report-boj-o-nase-srdcia-a-mysle-ako-moze-vyzerat-hybridne-posobenie-podcast/>

DISINFO REPORT, 2022b. *Najvýznamnejšími nositeľmi prokremel'ských naratívov sú na Slovensku politici | Podcasty Aktuality.sk* [online]. 1. říjen 2022. [vid. 2023-02-22]. Dostupné z: <https://play.libsyn.com/embed/episode/id/24542049/height/192/theme/modern/size/large/thumbnail/yes/custom-color/87A93A/time-start/00:00:00/playlist-height/200/direction/backward/download/yes>

DROBNIC HOLAN, Angie, 2022. PolitiFact | The Principles of the Truth-O-Meter: PolitiFact's methodology for independent fact-checking. *Polifact* [online] [vid. 2023-04-07]. Dostupné z: <https://www.politifact.com/article/2018/feb/12/principles-truth-o-meter-politifacts-methodology-i/>

D'ULIZIA, Arianna, Maria Chiara CASCHERA, Fernando FERRI a Patrizia GRIFONI, 2021. Fake news detection: a survey of evaluation datasets. *PeerJ Computer Science* [online]. **7**, e518. ISSN 2376-5992. Dostupné z: [doi:10.7717/peerj-cs.518](https://doi.org/10.7717/peerj-cs.518)

ELDH, Andreas, 2022. A pro-Russian bot network in the EU amplifies disinformation about the war in Ukraine. *EDMO* [online]. [vid. 2023-02-25]. Dostupné z: <https://edmo.eu/2022/04/28/a-pro-russian-bot-network-in-the-eu-amplifies-disinformation-about-the-war-in-ukraine/>

russian-bot-network-in-the-eu-amplifies-disinformation-about-the-war-in-ukraine/

EUVSDISINFO, 2023. About. *EUvsDisinfo* [online] [vid. 2023-02-23]. Dostupné z: <https://euvsdisinfo.eu/about/>

FILO, Jakub, 2007. *Tlačové agentúry na Slovensku po roku 1993 – vývoj, popis, porovnanie* [online]. Brno [vid. 2023-04-23]. Bakalárska diplomová práca. Masarykova univerzita, Fakulta sociálnych študií. Dostupné z: https://is.muni.cz/th/m6u3t/Jakub_Filo_bakalarna_praca.pdf

FRIDLUND, Patrik, 2020. Post-truth Politics, Performatives and the Force. *Jus Cogens* [online]. 2(3), 215–235. ISSN 2524-3985. Dostupné z: doi:10.1007/s42439-020-00029-8

GOLBECK, Jennifer, Matthew MAURIELLO, Brooke AUXIER, Keval H. BHANUSHALI, Christopher BONK, Mohamed Amine BOUZAGHRANE, Cody BUNTAIN, Riya CHANDUKA, Paul CHEAKALOS, Jennine B. EVERETT, Waleed FALAK, Carl GIERINGER, Jack GRANNEY, Kelly M. HOFFMAN, Lindsay HUTH, Zhenya MA, Mayanka JHA, Misbah KHAN, Varsha KORI, Elo LEWIS, George MIRANO, William T. MOHN IV, Sean MUSSENDEN, Tammie M. NELSON, Sean MCWILLIE, Akshat PANT, Priya SHETYE, Rusha SHRESTHA, Alexandra STEINHEIMER, Aditya SUBRAMANIAN a Gina VISNANSKY, 2018. Fake News vs Satire: A Dataset and Analysis. In: *WebSci '18: 10th ACM Conference on Web Science: Proceedings of the 10th ACM Conference on Web Science* [online]. Amsterdam Netherlands: ACM, s. 17–21 [vid. 2022-10-08]. ISBN 978-1-4503-5563-6. Dostupné z: doi:10.1145/3201064.3201100

HAJDU, Dominika a Katarína KLINGOVÁ, 2020. *Voices of Central and Eastern Europe: Perceptions of democracy & governance in 10 EU countries* [online]. B.m.: Globsec [vid. 2023-02-22]. Dostupné z: <https://www.globsec.org/what-we-do/publications/voices-central-and-eastern-europe-perceptions-democracy-governance-10-eu>

HAJDU, Dominika, Katarína KLINGOVÁ, Daniel MILO a Miroslava SAWIRIS, 2021. *GLOBSEC Trends 2021: Central and Eastern Europe one year into the pandemic* [online]. B.m.: Globsec [vid. 2023-02-22]. Dostupné z: <https://www.globsec.org/what-we-do/publications/globsec-trends-2021-central-and-eastern-europe-one-year-pandemic>

HARKOTOVA, Stanislava, 2016. Nahrávka ze sveta médií ukazuje, jak se žádá o podporu ruské propagandy –. *Neovlivni* [online] [vid. 2023-02-22]. Dostupné z: <https://neovlivni.cz/nahravka-ze-slovenska-ukazuje-jak-se-zada-o-podporu-ruske-propagandy/>

HODÁS, Martin, 2022. Ako ruská ambasáda naverbovala autora Hlavných správ. *živé aktuality* [online] [vid. 2023-02-22]. Dostupné z: <https://zive.aktuality.sk/clanok/LSQpiFe/pozrite-si-video-ako-ruska-ambasada-naverbovala-autora-hlavnych-sprav/>

HRČKOVÁ, Andrea a Tomáš MIRGA, 2018. Efekt filtračných bublín: Defenzívne stratégie a taktiky / The effects of Filter Bubbles: Defensive Strategies and Tactics « Informačné technológie a knižnice. *Katedra knižničnej a informačnej vedy FiF UK v Bratislave*) [online] [vid. 2022-10-22]. Dostupné z: <https://itlib.cvtisr.sk/%C4%8C1%C3%A1nky/clanek3466/>

IFCN CODE OF PRINCIPLES, 2023. IFCN Code of Principles. *IFCN* [online] [vid. 2023-02-23]. Dostupné z: <https://ifcncodeofprinciples.poynter.org/>

KAZAZ, Jana, Dominika HAJDU, Katarína KLINGOVÁ a Michal KORTIŠ, 2022. *GLOBSEC Trends 2022: CEE amid the war in Ukraine*. B.m.: Globsec.

KHAN, Ali, Kathryn BROHMAN a Shamel ADDAS, 2021. The anatomy of 'fake news': Studying false messages as digital objects. *Journal of Information Technology* [online]. **37**, 122–143. Dostupné z: doi:10.1177/02683962211037693

KITCHENS, Kitchens, Steven JOHNSON a Peter GRAY, 2020. *Understanding Echo Chambers and Filter Bubbles: The Impact of Social Media on Diversification and Partisan Shifts in News Consumption* [online] [vid. 2022-10-22]. Dostupné z: <https://misq.umn.edu/understanding-echo-chambers-and-filter-bubbles-the-impact-of-social-media-on-diversification-and-partisan-shifts-in-news-consumption.html>

KONSPIRATORI, 2022. *Zoznam stránok so sporným obsahom* [online] [vid. 2022-10-28]. Dostupné z: <https://www.konspiratori.sk/>

KOSPIRATORI.SK, 2023. konspiratori.sk - verejná databáza webových stránok s neserióznym, klamlivým, podvodným, konšpiračným alebo propagandistickým obsahom. *konspiratori.sk* [online] [vid. 2023-02-23]. Dostupné z: <https://konspiratori.sk/zoznam-stranok>

LAZER, David, Matthew BAUM, Yochai BENKLER, Adam BERINSKY, Kelly GREENHILL, Filippo MENCZER, Miriam METZGER, Brendan NYHAN, Gordon PENNYCOOK, David ROTHSCHILD, Michael SCHUDSON, Steven SLOMAN, C. SUNSTEIN, Emily THORSON, Duncan WATTS a Jonathan ZITTRAIN, 2018. The science of fake news. *Science* [online]. **359**, 1094–1096. Dostupné z: doi:10.1126/science.aao2998

MANISH, Manish Kumar Singh, Jawed Ahmed JAWED, Mohammad Afshar Alam ALAM, Kamlesh Kumar Raghuvanshi KAMLESH a Sachin Kumar SACHIN, 2022. *A Comparative Study of Computational Fake News Detection on Social Medi* [online]. 2022. [vid. 2022-10-28]. Dostupné z: doi:10.21203/rs.3.rs-1964791/v1

META, 2020. April 2020 Coordinated Inauthentic Behavior Report. *Meta* [online]. [vid. 2023-04-22]. Dostupné z: <https://about.fb.com/news/2020/05/april-cib-report/>

MEYERS, Marion, Gerhard WEISS a Gerasimos SPANAKIS, 2020. Fake News Detection on Twitter Using Propagation Structures. In: Max VAN DUIJN, Mike PREUSS, Viktoria SPAISER, Frank TAKES a Suzan VERBERNE, ed. *Disinformation in Open Online Media* [online]. Cham: Springer International Publishing, s. 138–158. Lecture Notes in Computer Science. ISBN 978-3-030-61841-4. Dostupné z: doi:10.1007/978-3-030-61841-4_10

MIKUŠOVIČ, Dušan, 2020. Prieskum Focusu: Slovensko súhlasí s povinnými rúškami a strpelo by aj zákaz vychádzania. *Denník N* [online] [vid. 2023-02-22]. Dostupné z: <https://dennikn.sk/1827458/prieskum-focusu-slovensko-suhlasi-s-povinnymi-ruskami-a-strpelo-by-aj-zakaz-vychadzania/>

MINISTERSTVO KULTÚRY SLOVENSKEJ REPUBLIKY, 2023. *Zoznam periodickej tlače - Detail - MK SR Zoznam periodickej tlače* [online] [vid. 2023-04-22]. Dostupné z: <https://www.culture.gov.sk/pertlac/modul/tlac/detail?id=5892>

MINISTERSTVO OBRANY SR, 2021. *AKČNÝ PLÁN KOORDINÁCIE BOJA PROTI HYBRIDNÝM HROZBÁM (2022 - 2024)*.

MOLINA, Maria, S. Shyam SUNDAR, Thai LE a Dongwon LEE, 2019. "Fake News" Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content. *American Behavioral Scientist* [online]. **65**, 000276421987822. Dostupné z: doi:10.1177/0002764219878224

NBÚ SR, 2018. *Koncepcia pre boj Slovenskej republiky proti hybridným hrozbám*.

NIELSEN, Admosphere, 2019. *Novinky a čerstvé správy hľadajú Slováci online hlavne na spravodajských weboch, potom taktiež na Facebooku | Nielsen Admosphere - Výzkumná agentura* [online] [vid. 2022-09-24]. Dostupné z: <https://www.nielsen-admosphere.sk/news/novinky-a-čerstve-spravy-hladaju-slovaci-online-hlavne-na-spravodajskych-weboch-potom-taktiez-na-facebooku>

NYE, Joseph S., 2004. *Soft Power: The Means To Success In World Politics*. B.m.: PublicAffairs. ISBN 978-1-58648-225-1.

RHODES, Samuel, 2021. *Filter Bubbles, Echo Chambers and Fake News: How Social Media Conditions Individuals to be Less Critical of Political Misinformation* [online]. 3. duben 2021. B.m.: Inter-university Consortium for Political and Social Research (ICPSR). [vid. 2022-10-23]. Dostupné z: doi:10.3886/E135024V2

RUŽIČKOVÁ, Michaela, 2022. Pol roka ruskej agresie na Ukrajinu. *INFOSECURITY.SK* [online]. [vid. 2023-02-22]. Dostupné z: <https://infosecurity.sk/zahranicne/pol-roka-ruskej-agresie-propaganda-pracuje-s-dehumanizaciou-a-s-rozdelenim-ukrajiny/>

SARNOVSKÝ, Martin, Viera MASLEJ KREŠŇÁKOVÁ a Klaudia IVANCOVÁ, 2022. Fake News Detection Related to the COVID-19 in Slovak Language Using Deep Learning Methods. *Acta Polytechnica Hungarica* [online]. **19**, 43–57. Dostupné z: doi:10.12700/APH.19.2.2022.2.3

SARNOVSKÝ, Martin, Viera MASLEJ-KREŠŇÁKOVÁ a Nikola HRABOVSKÁ, 2020. Annotated dataset for the fake news classification in Slovak language. In: *2020 18th International Conference on Emerging eLearning Technologies and Applications (ICETA): 2020 18th International Conference on Emerging eLearning Technologies and Applications (ICETA)* [online]. s. 574–579. Dostupné z: doi:10.1109/ICETA51985.2020.9379254

SHAO, Chengcheng, Giovanni Luca CIAMPAGLIA, Onur VAROL, Kai-Cheng YANG, Alessandro FLAMMINI a Filippo MENCZER, 2018. The spread of low-credibility content by social bots. *Nature Communications* [online]. **9**(1), 4787. ISSN 2041-1723. Dostupné z: doi:10.1038/s41467-018-06930-7

SHU, Kai, 2023. *KaiDMML/FakeNewsNet* [online]. Python. 15. duben 2023. [vid. 2023-04-16]. Dostupné z: <https://github.com/KaiDMML/FakeNewsNet>

SHU, Kai, Deepak MAHUDESWARAN, Suhang WANG, Dongwon LEE a Huan LIU, 2019a. *FakeNewsNet: A Data Repository with News Content, Social Context and Spatialtemporal Information for Studying Fake News on Social Media* [online]. 27. březen 2019. B.m.: arXiv. [vid. 2023-04-09]. Dostupné z: doi:10.48550/arXiv.1809.01286

SHU, Kai, Amy SLIVA, Suhang WANG, Jiliang TANG a Huan LIU, 2017. *Fake News Detection on Social Media: A Data Mining Perspective* [online]. 2. září 2017. B.m.: arXiv. [vid. 2022-09-25]. Dostupné z: <http://arxiv.org/abs/1708.01967>

SHU, Kai, Xinyi ZHOU, Suhang WANG, Reza ZAFARANI a Huan LIU, 2019b. The role of user profiles for fake news detection. In: [online]. s. 436–439. Dostupné z: [doi:10.1145/3341161.3342927](https://doi.org/10.1145/3341161.3342927)

SITLAULA, Niraj, Chilukuri K. MOHAN, Jennifer GRYGIEL, Xinyi ZHOU a Reza ZAFARANI, 2019. *Credibility-based Fake News Detection* [online]. 2. listopad 2019. B.m.: arXiv. [vid. 2023-02-25]. Dostupné z: <http://arxiv.org/abs/1911.00643>

SOUSA-SILVA, Rui, 2022. *Fighting the Fake: A Forensic Linguistic Analysis to Fake News Detection* | SpringerLink [online] [vid. 2022-09-28]. Dostupné z: <https://link.springer.com/article/10.1007/s11196-022-09901-w>

STRUHÁRIK, Filip, 2020. Facebook zrušil stránky konšpiračného časopisu Zem a vek pre opakované porušovanie pravidiel. *Denník N* [online] [vid. 2023-04-22]. Dostupné z: <https://dennikn.sk/2206945/facebook-zrusil-stranky-konspiracneho-casopisu-zem-a-vek-pre-opakovane-porusovanie-pravidiel/>

STRUHÁRIK, Filip, 2022a. *Co je nového v médiách*. 1. vydanie. Praha: Nová Beseda. ISBN 978-80-88383-34-5.

STRUHÁRIK, Filip, 2022b. MediaBrifing: Päťinu komentárov na sociálnych sieťach je potrebné odstrániť, ukazuje prax elfov. *Denník N* [online]. [vid. 2023-05-01]. Dostupné z: <https://dennikn.sk/2887758/mediabrifing-patinu-komentarov-na-socialnych-sietach-je-potrebne-odstranit-ukazuje-prax-elfov/?fbclid=IwAR0RZ0qKGnnuG5pSonFm7eM8d4peFc0IscebeFgEKXwYkkfNqucojKaqXiuc>

TASR, 2022. Európarlament schválil reguláciu digitálnych služieb. Rieši aj dezinformácie a cieľnú reklamu. *Živé.sk* [online]. [vid. 2023-05-01]. Dostupné z: <https://zive.aktuality.sk/clanok/0fj23h8/europarlament-schvalil-regulaciu-digitalnych-sluzieb-riesi-aj-dezinformacie-a-cielenu-reklamu/>

TWITTER, INC., 2023. Use Cases, Tutorials, & Documentation. *Developer Twitter* [online] [vid. 2023-04-16]. Dostupné z: <https://developer.twitter.com/en>

UKRAINEFACTS, 2023. #UkraineFacts: fact-checking disinformation about Ukraine's invasion by the IFCN Signatories. *UkraineFacts* [online] [vid. 2023-02-23]. Dostupné z: <https://ukrainefacts.org>

U.S. DEPARTMENT OF THE TREASURY'S OFFICE OF FOREIGN ASSETS CONTROL, 2021. *Treasury Escalates Sanctions Against the Russian Government's Attempts to Influence U.S. Elections* [online]. [vid. 2023-04-23]. Dostupné z: <https://home.treasury.gov/news/press-releases/jy0126>

VOXCHECK, 2023. About VoxCheck. *VoxUkraine* [online] [vid. 2023-02-23]. Dostupné z: <https://voxcheck.voxukraine.org/en.html>

WALKER, Christopher a Jessica LUDWIG, 2017. The Meaning of Sharp Power. *Foreign Affairs* [online]. [vid. 2023-02-22]. ISSN 0015-7120. Dostupné z: <https://www.foreignaffairs.com/articles/china/2017-11-16/meaning-sharp-power>

WANG, William Yang, 2017. “Liar, Liar Pants on Fire”: A New Benchmark Dataset for Fake News Detection. In: *ACL 2017: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* [online]. Vancouver, Canada: Association for Computational Linguistics, s. 422–426 [vid. 2023-04-07]. Dostupné z: doi:10.18653/v1/P17-2067

WATTS, Clint a Rossella CERULLI, 2021. Russia’s Affront on the News: How NewsFront’s Circumvention of Social Media Bans Demonstrates the Need for Vigilance. *Alliance For Securing Democracy* [online]. [vid. 2023-04-22]. Dostupné z: <https://securingdemocracy.gmfus.org/russias-affront-on-the-news-how-newsfronts-persistence-past-social-media-bans-demonstrates-the-need-for-vigilance/>

WIKIPEDIA, 2023. Gossip Cop - Wikipedia. *wikipedia* [online] [vid. 2023-04-10]. Dostupné z: https://en.wikipedia.org/wiki/Gossip_Cop

ZEM A VEK, 2023. Hlavná stránka - Zem&Vek. *Zem&Vek - Alternatívne spravodajstvo* [online] [vid. 2023-04-21]. Dostupné z: <https://zemavek.sk/>

ZHANG, Xichen a Ali A. GHORBANI, 2020. An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management* [online]. **57**(2), 102025. ISSN 03064573. Dostupné z: doi:10.1016/j.ipm.2019.03.004

ZHOU, Xinyi, Atishay JAIN, Vir PHOHA a Reza ZAFARANI, 2019. *Fake News Early Detection: A Theory-driven Model*.

ZHOU, Xinyi a Reza ZAFARANI, 2020. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. *ACM Computing Surveys* [online]. **53**(5), 109:1-109:40. ISSN 0360-0300. Dostupné z: doi:10.1145/3395046

ŽIVÉ.SK, 2022. Anonymous napadli web Zem a vek. Obliekli ho do ukrajinských farieb. *Živé.sk* [online] [vid. 2023-04-22]. Dostupné z: <https://zive.aktuality.sk/clanok/SIObbyi/anonymous-napadli-web-zem-a-vek-obliekli-ho-do-ukrajinskych-farieb/>

Prílohy

Příloha A: Dataset pre detekciu fake news

Prílohu tvorí súbor s názvom „Príloha_A.xlsx“, priložený k tejto práci v osobitnom súbore.

