Přehledová studie

ProInflow: časopis pro informační vědy
Vol. 10, No. 2, 2018, str. 23–51
DOI: https://doi.org/10.5817/ProIn2018-2-3
CC BY 3.0 CZ
http://www.phil.muni.cz/journals/index.php/proinflow

INFORMAČNÍ STRATEGIE BOJE PROTI DEZINFORMACI: PŘEHLEDOVÁ STUDIE VLIVNÝCH TECHNOLOGICKÝCH SPOLEČNOSTÍ PO PREZIDENTSKÝCH VOLBÁCH V USA V ROCE 2016

INFORMATION STRATEGY IN THE WAR AGAINST DISINFORMATION: SURVEY OF INFLUENTIAL TECHNOLOGY COMPANIES AFTER THE PRESIDENTIAL ELECTIONS IN THE USA IN 2016

Petr Ulrich

Masarykova univerzita, Filozofická fakulta

Abstrakt

Účel – Prudký rozvoj informačních technologií spolu se sociálními sítěmi vytvořil bezprecedentní prostředí, ve kterém jsou dezinformace a s nimi spojované informační operace povýšeny na velmi účinné a devastující zbraně hromadného ničení. Účelově vedená dezinformační kampaň v kyberprostoru již dnes dokáže ovlivnit výsledky demokratických voleb na úrovni světové velmoci. Jedná se proto o velmi závažné celospolečenské téma a problém, u kterého dosud neexistuje účinná obrana. Naším cílem je zmapovat aktuální přístupy boje proti dezinformaci z pohledu provozovatele technologických platforem, které jsou pro šíření dezinformace zneužívány, a navrhnout jednoduchý kategorizační model obranných informačních strategií.

Design/metodologie/přístup – Hlavní otázkou, kterou si klademe, je, jak vlivné technologické společnosti ve svých procesech, nástrojích a vizích dalšího produktového směřování reagovaly na událost prezidentských voleb v USA v roce 2016. Svou pozornost zúžíme pouze na oblast sociálních sítí a sledované období uzavřeme březnem 2018. Tato teoretická práce bude založena na následujících metodických postupech:

a) Dokumentová analýza primárních zdrojů publikovaných jednotlivými sledovanými technologickými společnostmi.
b) Mediální analýza zpravodajských zdrojů, které situaci komentují jak z pohledu informační strategie, tak i z pohledu různých sociokulturně-politických „spouštěčů".

Výsledky – Na základě analýzy sledovaných technologických společností a sociokulturně-politických „spouštěčů" jsme stanovili základní kategorizaci informačních strategií boje proti dezinformaci a promítli ji do první verze funkčního modelu, jaký dosud ještě nebyl pro tuto potřebu vytvořen. Z dílčích strategií se jako nejnosnější prokázala strategie postavená na automatizovaných fact-checking systémech s podporou umělé inteligence.

Originalita/hodnota – Mediátorům obsahu, tedy vydavatelům a jejich redakcím, je již v současnosti věnováno mnoho pozornosti v řadě studií a analýz, které navazují na rozsáhlý historický výzkum v oblasti vývoje médií a komunikace. Na druhou stranu rovina informační strategie provozovatele komunikační platformy, na kterou se v naší analýze zaměřujeme, je v tomto kontextu online světa novou a dosud nepříliš probádanou oblastí, která přináší nové otázky a výzvy s ohledem na řízení těchto platforem z perspektivy oborů informačního managementu a informační politiky.

Klíčová slova – dezinformace, informační operace, informační strategie, boj proti dezinformaci, komunikační platformy, analýza informačních strategií, model informačních strategií proti dezinformaci, klasifikace informačních strategií proti dezinformaci, hoax

Abstract

Purpose – The rapid development of information technologies with social networks has created an unprecedented environment in which disinformation and related information operations are recognized as very effective and devastating weapons of mass destruction. A purpose-driven disinformation campaign in cyberspace is able to influence results of democratic elections at the level of the world superpower. This is a very serious social issue and a problem for which there is no effective defense yet.
Our objective is to map current approaches to combating disinformation from the viewpoint of technology platform operators whose platforms are misused to spread misinformation and suggest a simple classification model of defensive information strategies.
Design/methodology/approach – Focus is laid on how influential technology companies reacted to the landmark of the US presidential elections in 2016 in their processes, strategy and product visions. Our attention is limited to the area of social networks and limited by March 2018. This theoretical work is based on the following methodological approaches:

a) Document analyses of primary sources published by individual companies being monitored.
b) Media analyses of news sources covering situation from an information strategy viewpoint as well as the socio-cultural political context and triggers associated with this phenomenon.

Results – Based on the analyses of selected technology companies and socio-cultural political triggers we have defined a simple classification of information strategies for combating disinformation in online platforms. With this classification we derived a first version of a functional model, that hadn't existed before. Strategy based on automated fact-checking systems with the support of artificial intelligence proofed to be the most vital.
Originality/value – Content mediators, publishers and their editors, have already received a significant amount of attention in many studies and analysis following a vast historical research in the area of media and communication. On the other hand, the view of the information strategy of a communications platform operator, which we are focusing on in our work, is in this context of the online world a new and still not very well researched area that brings new questions and challenges regarding the platform management from the perspectives of information management and information policy.

Keywords – disinformation, information operations, information strategy, war against disinformation, communication platforms, analyses of informatic strategies, model of information strategies against disinformation, classification of information strategies against disinformation, hoax

Anotace

Prudký rozvoj informačních technologií spolu se sociálními sítěmi vytvořil bezprecedentní prostředí, ve kterém jsou dezinformace a s nimi spojované informační operace povýšeny na velmi účinné a devastující zbraně hromadného ničení. Účelově vedená dezinformační kampaň v kyberprostoru již dnes dokáže ovlivnit výsledky demokratických voleb na úrovni světové velmoci. Jedná se proto o velmi závažné celospolečenské téma a problém, u kterého dosud neexistuje účinná obrana. Našim cílem je s ohledem na mezník prezidentských voleb v USA v roce 2016 a za použití metody dokumentové a mediální analýzy zmapovat aktuální přístupy boje proti dezinformaci z pohledu provozovatele technologických platforem, které jsou pro šíření dezinformace zneužívány. Zároveň sledujeme sociokulturně-politický kontext spojený s tímto fenoménem a nastiňujeme další možný vývoj. V závěru na základě zjištěných poznatků vyvozujeme vlastní klasifikaci a model informačních strategií boje proti dezinformaci.

Annotation

The rapid development of information technologies with social networks has created an unprecedented environment in which disinformation and related information operations are recognized as very effective and devastating weapons of mass destruction. A purpose-driven disinformation campaign in cyberspace is able to influence results of democratic elections at the level of the world superpower. This is a very serious social issue and a problem for which there is no effective defense yet. Our objective is to map current approaches to combating disinformation from the viewpoint of technology platform operators whose platforms are misused to spread misinformation. Focus is laid on the landmark of the US presidential elections in 2016 using the documentary method and media analysis. At the same time, we follow the socio-cultural political context associated with this phenomenon and outline further possible developments. In conclusion, based on our findings, we draw our own classification and model of information strategies to combat disinformation.

Úvod

Dezinformace v online prostředí je vážným společenským problémem, jehož význam roste. Nejedná se o izolovaný jev a dopad tohoto problému se projevuje na celé škále sociálního spektra a lidské komunikace – od nevinných žertů až po ovlivňování prezidentských voleb na úrovni světové mocnosti ("Facebook's failure", 2017).

Informace jsou bezpochyby velmi účinnou zbraní. Dezinformace je proto efektivní složkou hybridní války (tedy opaku konvenční války[1]) sloužící k destabilizaci celých zemí a k ovlivňování jejich politického směřování, a to jak v měřítku světových velmocí (Giles, 2016), tak i např. u teroristických organizací (Howard, 2015).

Dezinformace jako taková není žádnou novinkou, ale právě prudký rozvoj informačních technologií se sociálními sítěmi, agregátory obsahu a vyhledávači jí díky vysoké efektivitě distribuce dává nový rozměr. Jedná se proto o velmi závažné celospolečenské téma a problém, u kterého dosud neexistuje účinná obrana.

Mediátorům obsahu, tedy vydavatelům a jejich redakcím, je již v současnosti věnováno mnoho pozornosti v řadě studií a analýz, které navazují na rozsáhlý historický výzkum v oblasti vývoje médií a komunikace (Ordway, 2017).

Na druhou stranu rovina informační strategie provozovatele komunikační platformy je v tomto kontextu online světa novou a dosud nepříliš probádanou oblastí, která přináší nové otázky a výzvy s ohledem na řízení těchto platforem z perspektivy oborů informačního managementu a informační politiky.

Mezi určující hráče na tomto poli v současnosti patří společnosti *Facebook*, *Twitter*, *Alphabet (Google)* apod., které tato komunikační prostředí díky svým technologickým platformám umožňují, a nepřímo tak dávají fenoménu dezinformace prostor.

Celá oblast boje proti dezinformaci se nyní bouřlivě vyvíjí, není proto naší ambicí pokrýt ji celou beze zbytku a soustřeďujeme se na její nejdůležitější rysy a události, které ji aktuálně utváří.

Podobná snaha o kategorizaci různých strategií proti dezinformaci z pohledu provozovatele online komunikační platformy dle naší rešerše dosud neexistuje.

Zajímá nás, jak po stanoveném mezníku prezidentských voleb v USA v roce 2016 vypadala situace a trend informačních strategií vlivných technologických společností, které svými internetovými platformami nepřímo šíření dezinformace umožňují.

Našim cílem je zmapovat aktuální přístupy k řešení této problematiky a navrhnout jednoduchý kategorizační model pro strategie boje proti dezinformaci z pohledu provozovatele online komunikační platformy.

Na tomto konkrétním příkladu chceme také ukázat, že problematika dezinformace v online prostředí je globální a týká se nás všech. Téma analyzující situaci u vlivných technologických společností po

[1] S ohledem na rozsah a zaměření článku si dovolujeme velmi zjednodušený pohled na definici hybridní války.

prezidentských volbách v USA v roce 2016 je tudíž zobecnitelné a může informačním pracovníkům ukázat další směr vývoje v této oblasti.

Pro informační pracovníky je toto téma obzvláště důležité, protože to jsou právě oni, kdo vytváří pravidla, nástroje a služby v online prostředí.

Základní definice

Informační operace

Pro naše účely si termín informační operace zúžíme na oblast vlivových operací. Informační, neboli také vlivové operace (*Information operations*) (Armistead, 2010), jsou informační kampaně zatím primárně politického rázu, často organizované na vládní úrovni, ale také soukromými vlivovými organizacemi, které si kladou za cíl destabilizovat společnost za daným účelem. Nejčastějším záměrem informačních operací je dosažení strategického anebo geopolitického cíle pomocí manipulace veřejného mínění – typicky před volbami. Jejich hlavní municí jsou dezinformace, falešné zprávy či falešné zesilovače a jejich vojáky jsou tzv. internetoví trollové.

Dle posledních analýz StratCom je hlavním hráčem na poli informačních operací Rusko, kde se tento typ invazivních aktivit stal součástí ruské strategie hybridní války a také obrany (Giles, 2016). Ruské informační operace jsou zaměřené jak směrem dovnitř na vlastní občany, tak zejména do tzv. nárazníkových zón ohnisek konfliktů (Friedman, 2015), které leží na pomezí sféry Východu a Západu. Do těchto zón patří například Ukrajina, Balkán, Finsko ("Hybrid influence", 2017), ale také Česká republika.

Působení informačních operací, které jsou v médiích připisovány Rusku, bylo již také prokázáno mimo tradiční zóny ohnisek konfliktů, a to přímo v srdci „nepřítele", např. v námi sledovaném období právě při ovlivňování prezidentských voleb v USA v roce 2016 (Allcott & Gentzkow, 2017) nebo při francouzských prezidentských volbách v roce 2017 (jak ukázala studie tzv. MacronLeaks) (Ferrara, 2017), tak i při parlamentních a prezidentských volbách v ČR v roce 2018.

S nadsázkou lze říct, že informační operace jsou ve srovnání s tradiční propagandou vysoce efektivní a řádově levnější než jeden Mig vybavený hlavicí s plochou dráhou letu.

Falešné zprávy

Falešné zprávy (*fake news*) představují zpravodajství, které je záměrně a ověřitelně nepravdivé a které může být pro čtenáře zavádějící. Falešné zprávy mohou být součástí informačních (vlivových) operací (MacFarquhar, 2016), stejně tak na nich může být postaven obchodní záměr klikacích farem, jejichž jediným cílem je díky atraktivnímu – byť smyšlenému – obsahu profitovat z prodeje reklamy (Tynan, 2016).

Server BuzzFeed provedl v roce 2016 analýzu dvaceti nejúspěšnějších falešných zpráv šířících se na sociální síti Facebook, které porovnal s dvaceti nejúspěšnějšími legitimními zprávami v tradičních médiích, jakými jsou např. New York Times, Washington Post, Huffington Post, NBC News a další. Výsledkem této analýzy bylo, že byť legitimní zprávy měly lepší nástup, tak falešné zprávy na konci sledovaného tříměsíčního období

skončily s výrazně lepším výsledkem dopadu co do sdílení, reakcí a komentářů na Facebooku (Silverman, 2016).

Internetoví trollové

Pro naše účely termín internetový troll zúžíme na označení buď pro daného jedince nebo častěji pro celou koordinovanou síť často falešných identit působících zejména na sociálních sítích, internetových diskuzních fórech, blozích a v diskuzích u publikovaných mediálních zpráv, které mají za úkol zaplňovat informační prostor dezinformačním obsahem, polarizovat a prosazovat konkrétní narativ odpovídající dané informační operaci, jejíž jsou součástí.

Trollové se často vyskytují v organizovaných skupinách v tzv. *trollích farmách.*

Termín *trollí farma* zavedla ruská investigativní novinářka A. Garmazhapova, která v roce 2013 v utajení infiltrovala jednu z trollích farem v Petrohradě a důkladně popsala jejich modus operandi (Garmazhapova, 2013).

Trollí farmy zaměstnávají armádu blogerů, kteří dnem i nocí ve 12 hodinových směnách na internetu produkují dezinformace, přičemž za částí některých trollích sítí stojí tzv. boti, tedy specializované počítačové skripty.

Pro naše účely si označení *bot* zúžíme na počítačové programy specializované pro šíření dezinformace. Tyto automatizované systémy se pokoušejí maskovat a tvářit se jako legitimní lidští uživatelé.

Rozsáhlá studie specializovaného centra pro strategickou komunikaci StratCom ve svém výzkumu odhalila, že *„dva ze tří uživatelských účtů na síti Twitter, které v ruštině masivně šíří zprávy včetně dezinformace o posílení vojenské přítomnosti NATO v Pobaltí a v Polsku, ovládají boti"* (Fredheim, 2018).

Sítě automatizovaných botích systémů na síti Twitter již sleduje a jejich aktivitu veřejně zpřístupňuje např. projekt Hamilton 68 americké nadace German Marchall Fund. Hamilton 68 je online nástroj, který v reálném čase zpřístupňuje přehled ruské propagandy a dezinformačních aktivit.

Trollí farmy jsou zakládány i mimo Rusko (Benedictus, 2016) a důkazy o jejich existenci jsou například i v sousedním Polsku (Gorwa, 2017) nebo na Slovensku (Goda, 2018).

Metodika trollích farem se neustále zlepšuje, je již poměrně sofistikovaná a jejich úspěšnost v manipulaci s veřejným míněním je vysoká, a to dokonce u tak informačně vzdělaného národa, jakým jsou např. Finové ("Why Is Finland Able to Fend Off Putin's Information War?", 2017). Díky škálovatelnosti sociálních sítí ji proto v kombinaci s *falešnými zesilovači* považujeme za největší hrozbu pro demokracii.

Falešné zesilovače

Tzv. falešné zesilovače (*False amplifiers*) představují termín specifický pro oblast sociálních sítí. Je to koordinovaná aktivita využívající sítě falešných identit trollů, kterým se podařilo infiltrovat doménu reálných

uživatelů. Tito reální uživatelé jim pak nevědomky propůjčují svou důvěryhodnost ve vztahu ke svým „přátelům“ na sociální síti, a rozsah dezinformace tak neúmyslně ještě dále zesilují.

Typickým příkladem jsou zde tzv. *bikini trollové*, tedy falešné identity s profilovými fotografiemi mladých žen, které buď slibují své „oběti“ v žádostech o přátelství sexuální podtext, nebo jen v online diskuzích jemně žádají diskutující o přehodnocení názoru na dané téma, které je předmětem informační operace (Boffey, 2016). V prvním případě se často po uplynutí určité doby z falešných profilů změnou profilové fotografie nepozorovaně vyklube nová identita. Ta pak v online komentářích díky viditelné vazbě nastaveného „přátelství“ se svou obětí působí důvěryhodně, a může tak velmi efektivně indoktrinovat širší sociální okolí.

Dezinformace

Termínem dezinformace označujeme zavádějící či jinak nepřesný a zmanipulovaný obsah, který je šířen záměrně v rámci informační operace. Dezinformace může zahrnovat jak falešné zprávy, tak mnohem jemnější metody, jako např. podsouvání zavádějících citací a příběhů nic netušícím prostředníkům, kteří pak z titulu své důvěryhodnosti vůči svým přátelům fungují jako zesilovače dezinformace ve své infosféře.

Sociokulturně-politický kontext a spouštěče

Veřejné mínění a s ním utvářený sociokulturně-politický kontext je důležitým hybatelem, který v pravou chvíli dokáže posunout pokrok o notný kus dopředu, nebo jej také na dlouho zadržet.

Co a proč se změnilo, že najednou musí zástupci sociálních sítí svědčit v americkém Kongresu kvůli podezření ze zneužití jejich platforem k informačním operacím při prezidentských volbách? Proč právě probíhající skandál *Cambridge Analytica* sráží Facebook na kolena, a to navíc kvůli něčemu, co se stalo před osmi lety a už před třemi lety bylo napraveno a medializováno? Přičemž velmi podobné využití citlivých osobních dat pomocí *Open Graph API* ze sociální sítě Facebook pro politické účely již bylo zaznamenáno u volební kampaně bývalého amerického prezidenta Baracka Obamy v roce 2012 (McCullagh, 2018), ale bez negativního mediálního ohlasu a spíše jako pozitivní případová studie přelomového politického marketingu.

Proč se dá očekávat, že právě nyní bude mezi zodpovědnými technologickými giganty skutečná vůle odložit technokratické růžové brýle a věci řešit? A hlavně jakým způsobem?

Medializace a vlna vyšetřování zneužití sociálních sítí pro informační operace se dle naší mediální analýzy novinových článků nejzřetelněji týká Facebooku, přičemž i další komunikační kanály, jako je Twitter a ostatní blogovací platformy, jsou pro účely šíření dezinformace zneužívány také. Za daných okolností jsou však ve zpravodajství zastíněny Facebookem, který je hlavním terčem investigativních novinářů a politických komentátorů. Ti s sebou strhávají vlnu nevole uživatelů těchto platforem.

Odpovědi na tyto otázky poměrně dobře zmapoval americký časopis Wired, který ve svém článku *Inside the two years that shook Facebook – and the World* (Thompson & Vogelstein, 2018) rozkryl sérii událostí a spouštěčů, které vedly až k současnému „křižáckému tažení“ na Facebook, a s ním i na celou oblast sociálních sítí.

Na hraně zákona

Nově nastupující technologické online komunikační platformy velmi záhy změnily způsob, jak lidé vyhledávají a konzumují informace. Na počátku byl idealisticko-technokratický pohled, kdy demokratizované informace ("What is the future of news?", 2018) změní svět – to se také stalo, otázkou ale je, zda k lepšímu.

Facebooku se podařilo bitvu o demokratizaci informací vyhrát. Jak ve svém přehledovém článku v časopisu Wired dále uvádí Thompson a Vogelstein, Facebook *„v roce 2013 nahradil Twitter jakožto hlavní zdroj online distribuce zpráv a v roce 2015 nahradil Google, který do té doby ovládal trh směrování čtenářů na stránky vydavatelů."*

Tímto svým strategickým krokem se ale Facebook dostal do přímé konkurence s etablovanými vydavateli. Zavedením své funkce *Instant Articles* je de facto odsunul na vedlejší kolej a upevnil tak svou pozici hlavního zdroje pro přístup k informacím a zároveň konzumaci zpráv.

Facebook tak ovládl zpravodajství, ovšem jakožto otevřená technologická platforma se tím proti své vůli začal vzdalovat ustanovení sekce 230 amerického zákona *Communications Decency Act* z roku 1996, který jej dosud chránil ("Section 230", 2018).

Podle tohoto zákona je provozovatel komunikační platformy pouhý zprostředkovatel, a nenese tak odpovědnost za obsah tvořený jeho uživateli ("The Law that Gave Us the Modern Internet – and the Campaign to Kill It", 2013). Pokud by Facebook nebo jakýkoliv jiný provozovatel sociální sítě či blogovací platformy začal editorsky upravovat uživatelský obsah či jinak editorsky rozhodoval o jeho zobrazení, pak by přestal být pouhou platformou a stal by se vydavatelem. Tím by o tuto svou imunitu přišel, což by pro jeho koncept otevřené technologické platformy mohlo být devastující. Nemůžeme se proto divit, že se společně s ostatními provozovateli komunikačních platforem do poslední chvíle Facebook snažil držet své technokratické vize a obsah protékajících zpráv příliš neřešil.

Navíc stejně jako pro Facebook, tak i pro ostatní hráče v oblasti obsahu na síti Internet je hlavním zdrojem příjmů reklama a angažovanost uživatelů k obsahu. Čistě pragmaticky řečeno, čím více zajímavých zpráv, které dokážou aktivovat uživatele a prodat inzerci, tím lépe – bez ohledu na původ, charakter a autenticitu těchto zpráv.[2] To není příliš dobré výchozí stanovisko pro boj proti dezinformaci.

Aféra Trending Topics

Pro porozumění aféry *Trending Topics* (Bowles & Thielman, 2016) je zapotřebí nejdříve zmínit princip fungování tzv. *News Feed*, neboli kanálu vybraných příspěvků.

News Feed je základem sociální sítě a online médií. Jedná se o jednu konkrétní stránku, jejímž prostřednictvím uživatelé konzumují zprávy dané online služby. Obsah News Feedu je zpravidla

[2] S výjimkou jednoznačně protizákonného obsahu, který je jasně specifikován v tzv. Community Guidelines.

personalizovaný dle preferencí konkrétního uživatele a znalostní báze jeho chování při užívání dané online služby. Jako příklad si zde uveďme News Feed sociální sítě Facebook.

Z principu News Feedu je patrné, že v daný okamžik o pozornost konkrétního uživatele soupeří větší množství zpráv, které jsou následně filtrovány dle algoritmů specifických pro danou online službu. Na základě těchto algoritmů se tak některé algoritmicky vybrané zprávy prostřednictvím News Feedu k uživateli dostanou dříve, jiné později a většina z celkového množství v danou chvíli dostupných zpráv se k uživateli nedostane vůbec.

News Feed je tak nejcennější a pro producenty zpráv nejdůležitější částí sociální sítě, resp. online služby, která ovlivňuje, jaké informace se k uživateli dostanou, a jaké již ne.

Právě ve spojení s News Feedem sociální sítě Facebook se jedním ze stěžejních milníků na cestě od techno-optimismu k aktivnímu hledání řešení boje proti dezinformaci stala aféra Trending Topics, kdy se ukázalo, že místo deklarovaných algoritmů rozhodují o zobrazování informací lidští kurátoři Facebooku.

Když pomineme poměrně výbušný politický rozměr této aféry, kdy byly upřednostňovány zprávy jedné politické strany, tak již zde stál Facebook na hraně sekce 230 onoho komunikačního zákona, který jej dosud chránil.

Této chyby ve svém tažení proti Facebooku zřejmě využila tradiční média a jejich velcí vydavatelé, jako např. Murdochova *News Corp*, které si Facebook svým technokratickým přístupem a postupným upevňováním moci na poli zpravodajství znepřátelil.

Pro Facebook to byla první zatěžkávací zkouška, ze které se mu nakonec podařilo vyjít relativně úspěšně s odkazem na to, že Facebook zůstává otevřenou platformou a práce lidských kurátorů byla pouze dočasná, s cílem podpořit učící se algoritmy. Dílčím výsledkem bylo vydání manifestu *Building a Better News Feed for You* (Mosseri, 2016), ve kterém Facebook poprvé veřejně nastavil pravidla svého News Feedu, tedy nejcennějšího místa své aplikace, kde dochází k prezentaci a konzumaci zpráv.

Po této negativní zkušenosti se však Facebook o to víc snažil držet své role otevřené technologické platformy a cíleně popíral a bagatelizoval projevy informačních operací na své platformě (Romm & Wagner, 2017).

Thompson a Vogelstein vysvětlují, že *„kdyby Facebook převzal odpovědnost za falešné zprávy na své platformě, musel by pak převzít zodpovědnost za mnohem více. Facebook tak měl mnoho důvodů ponechat hlavu v písku."*

Do soukolí osudu však v tuto chvíli vstupuje zásadní kolečko amerických prezidentských voleb v roce 2016, které maximálně využilo pákový efekt sociální platformy Facebooku pro své informační operace a s nimi spojené šíření dezinformací.

Facebook toto zneužití své platformy ve světle nedávných zkušeností zcela programově a trestuhodně zaspal. Svými následnými aktivitami se pak přirozeně snažil dohnat ztracený čas, uklidnit média, politiky a samozřejmě také znepřátelené vydavatele médií.

Součástí těchto aktivity Facebooku bylo oznámení o zpřístupnění platformy externím fact-checking službám (viz níže) pro kontrolu zpráv (Newton, 2016), následované vznikem speciální organizační jednotky (květen 2017), která měla za úkol zlepšit integritu News Feedu a nastavit principy pro vydavatele (Mosseri, 2017).

Tato nová organizační jednotka se následně transformovala do nového projektu *Facebook Journalism Projec*, jehož interním cílem bylo zlepšit vztahy se znepřátelenými vydavateli a vylepšit mediální obraz Facebooku. Proklamovaným cílem však bylo nastavit platformu pro podporu novinářské práce a kritického myšlení.

Vyšetřování amerického Kongresu k prezidentským volbám

V průběhu roku 2017 vyšlo najevo, že Facebook byl „napaden" zahraniční informační operací jednoznačně zacílenou na manipulování amerického veřejného mínění v průběhu prezidentských voleb. Zde se již jednalo o otázku národní bezpečnosti, a proto Kongres Spojených států zahájil v této věci vlastní vyšetřování, které zahrnovalo také tři technologické společnosti: Facebook, Twitter a Google (dnešní Alphabet).

Přístup Facebooku byl dle série vyjádření v médiích opět velmi vlažný, s cílem nenechat se coby technologická firma vtáhnout do politického boje a znovu neotvírat staré rány po aféře s Trending Topics. Interně však pracoval, snažil se pochopit novou situaci a najít řešení.

V dubnu 2017 Facebook na základě svého interního šetření vydal manuál *Information Operations and Facebook v1.0* (Weedon, Nuland, & Stamos, 2017), který rámcově popisuje možnosti zneužití platformy pro informační operace a možné způsoby obrany.

Dále pak pod tlakem vyšetřování Facebook podle dat, která byl s předmětem vyšetřování schopný ve své platformě dohledat, zveřejnil, že prostřednictvím prokazatelně ruských falešných účtů bylo na jeho platformě zakoupeno 3.000 reklamních jednotek v celkové částce pouhých 100.000 USD (Shane & Goel, 2017).

V tomto bodě se Facebook opět snažil bagatelizovat skutečný význam a rozsah problému, protože podle studie (Albright, 2017) zveřejněné Jonathanem Albrightem z Tow Center for Digital Journalism na Kolumbijské univerzitě byla ruská informační operace sdílená více než 340 milionkrát, a mohla tak ovlivnit více než 10 milionů Američanů (Glaser, 2017).

Během slyšení Facebooku, Twitteru a Google v americkém Kongresu Dianne Feinsteinn, senátorka za stát Kalifornie, pohrozila směrem k technologickým firmám zhmotněním jejich noční můry, které se již dlouho obávaly: *„Vytvořili jste tyto platformy, které jsou nyní zneužívány. Jste to proto vy, kdo musí najít řešení – jinak to uděláme my"* (Newton, 2017).

Souhra těchto spouštěčů měla za následek mediální vlnu volající po státní regulaci sociálních sítí. Toto mediální tažení bylo navíc podporováno některými akcionáři Facebooku a také jeho bývalými klíčovými zaměstnanci, jako byl např. bývalý Privacy manažer Sandy Parkilas, který v NY Times zveřejnil příspěvek pochybující o schopnosti Facebooku v této věci zasáhnout: *We Can't Trust Facebook to Regulate Itself* (Parakilas, 2017).

Stát se předmětem státní regulace je přitom pro úspěšnou komerční firmu často devastující.

V následujících měsících vlna kritiky sociálních sítí a zejména Facebooku zesílila, doprovázena po celém světě konkrétními případy, kdy informační operace prováděné v rámci jeho sociální platformy negativně zasáhly do demokratického procesu ostatních zemí.

Aféra Cambridge Analytica

V březnu 2018 otřásl světem další skandál přímo spojený s Facebookem, kdy se ukázalo, že společnost *Cambridge Analytica* podvodně zneužila osobní data uživatelů Facebooku k cílené politické reklamní kampani Donalda Trumpa při amerických prezidentských volbách v roce 2016 a podobně tak učinila i v případech voleb v dalších částech světa ("Revealed ", 2018).

Tento skandál měl za následek další vlnu vyšetřování a také dokonce Facebookovou anti-kampaň #DELETEFACEBOOK (Griffin, 2018), kdy organizace i jednotliví uživatelé na protest mazali své profily a facebookové stránky a opouštěli sociální platformu Facebooku. Mezi účastníky kampaně byl i Elon Musk, který smazal facebookové stránky svých ikonických firem Tesla a SpaceX (Grush, 2018). Tak silná antikampaň znamenala pro Facebook těžkou ránu.

Na základě této další vlny událostí začal Facebook (Wong, 2018) i ostatní hráči jednat a hledat řešení. Vyvíjený společenský tlak a obava ze státních regulací se konečně staly velmi silnou motivací pro hledání řešení na ochranu soukromí a také boje proti dezinformaci.

Otázkou nyní zůstává, zda výsledkem tohoto sociokulturně-politického tlaku bude skutečný pokrok v boji proti dezinformaci, nebo se sledované technologické společnosti v čele s Facebookem vrátí ke své dosavadní krátkozraké strategii odvádění pozornosti, kterou jsme u nich dosud sledovali. I dílčí pokrok je však vítaný – a ten již můžeme vidět.

Doplňovací volby v USA v roce 2018

Dle našeho soudu posledním klíčovým hybatelem jsou ve sledovaném období chystané doplňovací volby v USA, které pod tlakem dosavadního vývoje a souhry ostatních spouštěčů budou mít velký dopad na konkrétní kroky vlivných technologických firem v boji proti dezinformaci.

Facebook již v této souvislosti vydal svůj čtyřbodový obranný plán systémových opatření, kterým hodlá zamezit dalšímu zneužití své platformy nejen při těchto amerických doplňovacích volbách, ale také u demokratických procesů v jiných zemích.

Je to sice teprve začátek na dlouhé cestě při řešení tak závažného a složitého problému, jako je dezinformace a informační operace, ale už samotný začátek lze považovat za pozitivum.

Analýza informačních strategií sledovaných organizací

Je nasnadě, že kromě námi zvolených tří významných technologických společností působí na trhu celá řada dalších subjektů, jejichž platformy jsou také velmi často zneužívány pro šíření dezinformace. Jsou to např. oblíbené blogovací platformy, messaging platformy, hostingové platformy apod.

Důležité je si uvědomit, že dezinformace působí v rámci informačních operací jako bájná Hydra s mnoha hlavami, z nichž každá hlava plně využije prostor a zranitelnost systému, který v rámci infosféry v dobré víře vytvoříme. Není to proto jen záležitost námi sledovaných technologických společností, ale v podstatě jakákoliv komunikační platforma, byť by se jednalo jen o nástěnku v knihovně – může být a bez náležité péče a ochrany s vysokou pravděpodobností také bude využita pro šíření dezinformace. Právě široký průnik dezinformace a konkrétního narativu do okrajových komunikačních kanálů a lokálních domén dává informační operaci potřebnou sílu.

Výchozí pozice

Jak již vyplynulo z předchozí kapitoly, vzestup online světa a s ním spojených technologických společností proběhl relativně velmi rychle. Celý proces byl hnán technologicko-optimistickou vizí demokratizace informací a lepšího světa, který nové technologie lidem pomohou vytvořit.

Tato doba by se dala shrnout do mantry „move fast and break things" (Taplin, 2017), podle níž rychle rostoucí technologické startupy neměly příliš prostor dívat se do zpětného zrcátka, zda se jim náhodou jejich platformy nezačaly vymykat z rukou.

Z analýzy celé řady firemních prohlášení, produktových kroků a mediálních komentářů lze vyvodit, že si dnešní technologičtí giganti buď nebyli, nebo nechtěli být dlouho vůbec vědomi svého skutečného vlivu na svět a zejména s ním spojené odpovědnosti.

Sebereflexi nepomohl ani již zmíněný článek 230 amerického komunikačního zákona, který jim coby platformám dával imunitu na zodpovědnost za obsah. Pokud by tuto dělicí čáru překročili a začali by problematiku dezinformace skutečně řešit, pak by o svou imunitu mohli přijít, a tím otevřít pověstnou Pandořinu skříňku špatných zpráv pro své akcionáře.

Článek 230 komunikačního zákona, který byl původně pojistkou svobody slova na internetu a také hnací silou pro vznik Internetu s plejádou komunikačních služeb tak, jak je nyní známe, se paradoxně na dlouhou dobu stal zároveň obávaným strašákem, který zmrazil snahy v boji proti dezinformaci. Teprve nyní se ve světle měnícího se sociokulturně-politického kontextu a s ním spojených spouštěčů zdá, že se věci daly konečně do pohybu.

Společné prvky informační strategie

Community Guidelines

Základním a společným prvkem informační strategie všech námi sledovaných hráčů jsou *Community Guidelines, Rules, Terms and Policies* či *Community Standards*, tedy zveřejněná kodifikace pravidel pro používání platformy jejími uživateli (dále jen *Community Guidelines*).

Community Guidelines se průběžně vyvíjí, a reflektují tak měnící se svět a s ním spojenou zkušenost provozovatelů těchto platforem. Ta je prostřednictvím průběžně vznikajících pravidel jednotně komunikována všem zúčastněným stranám.

Kromě základních pravidel akceptovaného chování jsou Community Guidelines a s nimi spojené monitorovací nástroje primárně zaměřené na dodržování snadno definovatelných, a proto také uplatnitelných zákonných a společenských norem, např. ve vztahu k sexualitě, propagaci zakázaných ideologií, spamu, malware, harassementu apod.

Algoritmy, ne lidé

S ohledem na podstatu námi sledovaných technologických firem, odstavec 230 komunikačního zákona a také s přihlédnutím na enormní množství obsahu, který jejich platformami každou minutu na celém světě protéká, je v rámci informační strategie těchto společností kladen velký důraz na technologická řešení jakéhokoliv problému.

Armáda lidí totiž nikdy nebude tak výkonná, aby mohla zkontrolovat veškerý obsah, který právě teď na světě vznikl. Technologická řešení s podporou algoritmů a rozvíjející se umělé inteligence jsou naopak dobře škálovatelná.

Navíc pro technologické organizace je již z principu věci mnohem bližší algoritmický (tedy technologický) přístup k řešení problému. Editorské zásahy lidských kurátorů do obsahu a jeho zobrazování tak jdou proti této filosofii. Připomeňme si, že porušení tohoto pravidla bylo spouštěčem aféry Trending Topics.

Automatizovaný monitoring je ale zatím vhodný pouze pro jednoduché a snadno definovatelné situace, jako např. strojově rozpoznatelný obrázek bradavky, porušení autorských práv (např. u hudby na YouTube nebo na Facebooku) nebo strojová identifikace falešných účtů pomocí kognitivních heuristik.

Algoritmům ale vždy budou muset asistovat školené týmy kurátorů, které budou práci strojů kontrolovat a rozhodovat u sporných situací.

Algoritmy se z rozhodování lidských specialistů učí, a postupně se tak zlepšují a posouvají hranici mezi strojem a člověkem v kurátorském procesu. Lidé zde ale budou zapotřebí vždy. Jen např. Facebook v tuto chvíli zaměstnává kolem 10.000 takových kurátorů a na tento rok v souvislosti s doplňovacími volbami v USA plánuje v očekávání dalších informačních operací zdvojnásobení jejich počtu.

Crowdsourcing

Dalším důležitým prvkem vysledované informační strategie jednotlivých hráčů je tzv. crowdsourcing indikátorů, které jdou nad rámec možností strojového učení dnešní doby. Uživatelé tak mají možnost provozovateli platformy jednoduše nahlásit závadný obsah, který je následně zkontrolován lidským kurátorem.

Z jednotlivých signálů generovaných uživateli a následných rozhodnutí lidských kurátorů opět vznikají data, která slouží pro učení umělé inteligence stojící za algoritmy platformy.

Příkladem zde může být tlačítko *Downvote*, které Facebook v současné době testuje na americkém trhu. Toto tlačítko bude na rozdíl od tlačítka „palce nahoru" do systému Facebooku signalizovat nevhodný, urážlivý, zavádějící nebo jinak závadný obsah (Roy, 2018).

Převedení odpovědnosti na koncového uživatele

Kromě základní bible v podobě Community Guidelines, automatizovaných řešení v kombinaci s doplňkovými lidskými kurátory a crowdsourcingových indikátorů, je v informační strategii jednotlivých firem velmi patrná stopa převedení odpovědnosti na koncového uživatele platformy, který má možnost jednoduše nežádoucí obsah ve svém News Feedu umlčet a který je také v rámci různých nastavení sám zodpovědný za ochranu svého soukromí.

Ze zkušenosti s dezinformací však víme, že toto přenesení odpovědnosti na koncového uživatele je samo o sobě zcela nedostačující.

Nastavení jednotlivých uživatelů v rámci dané kulturní domény však opět může sloužit jako základní linie pro zobecnitelné nastavení parametrů všech podobných uživatelů v dané doméně, kteří si své nastavení sami nepřizpůsobili.

Platforma zde algoritmicky pracuje s velkými daty a vyvozuje z nich závěry na způsob místního referenda, tedy *techno-referenda*.

Techno-referendum díky automatizované práci s velkými daty sleduje nejčastější kombinace nastavení v dané komunitě a podobně jako v referendu zjištěný výsledek povýší na standard a výchozí nastavení pro všechny ostatní členy komunity, kteří si své vlastní nastavení dosud neupravili.

Výsledkům techno-referenda se však na rozdíl od běžného referenda nemusí uživatel podrobit a jednoduchou manuální úpravou svého nastavení z jeho působení může kdykoliv snadno vystoupit. Tato jeho akce ovlivní výsledky techno-referenda pro další uživatele, pokud množství podobně jednajících uživatelů překročí algoritmicky danou kritickou masu v dané komunitě.

Osvěta a vzdělávání

Se snahou o převedení části odpovědnosti na koncového uživatele nutně přichází také osvěta, která uživatele postupně vzdělá v problematice ochrany soukromí a kritického myšlení. Ty jsou nezbytné nejen pro život na internetové sociální síti.

Jednotlivé tři sledované společnosti tak buď samy vytváří, nebo se podílí a podporují třetí strany při osvětových a vzdělávacích kampaních – jedná se však o běh na velmi dlouhou trať, protože dle zkušenosti z již zmíněného Finska i vzdělaný uživatel snadno podlehne sofistikované informační operaci.

Jedním z hlavních cílů informačních operací totiž je polarizací veřejného diskursu a falešnými zprávami rozmazat hranici mezi pravdou a lží a dosáhnout tzv. stavu post-truth, kdy je i pro vzdělaného jedince velmi obtížné se zorientovat v tom, komu a čemu věřit.

Fyzická bezpečnost

Nedílnou součástí informační strategie, která je společná všem hráčům, je samozřejmě fyzická bezpečnost dat a jednotlivých uživatelských účtů, včetně snahy zamezit vzniku a projevům falešných účtů.

Právě falešné účty tvoří základ trollích farem, proto je tato oblast ve strategii boje proti dezinformaci velmi důležitá.

Umělá inteligence

Přes různé experimenty a hrátky s přenosem odpovědnosti a fyzické bezpečnosti jsou naprosto klíčovým a dominantním prvkem informační strategie sledovaných organizací právě automatizované algoritmy v kombinaci se strojovým učením a Community Guidelines.

V boji proti dezinformaci jsou však automatizovaná řešení zatím stále velmi neobratná, a proto také nepříliš použitelná. To ostatně ukazuje případová studie Facebooku s Trending Topics, která odstartovala tok událostí a spouštěčů, který technologické společnosti dovedl až k vynucené sebereflexi.

Organizační struktura

Nelze však říci, že by se na poli boje proti dezinformaci za celou dobu nic konkrétního nedělo a že by v jednotlivých sledovaných společnostech i mimo ně neexistovali lidé, kteří by neviděli dále než jen za nejbližší kvartál obchodních výsledků, a kteří by se nesnažili nalézt řešení.

Vznikla celá řada projektů a iniciativ, které se však zatím nedočkaly potřebného průlomu, např. proto, že zatím nedostaly dostatečný prostor a zdroje – chyběla jasná motivace a informační strategie, která by měla boj proti dezinformaci jako svou prioritu.

Samotné technologické opatření je totiž bezzubé, pokud společně s ním není zavedeno také odpovídající organizační opatření, které demonstruje skutečnou vůli k řešení. Bez tohoto opatření pak všechny ostatní kroky zůstávají spíše jen na úrovni alibi.

Příklady dalších informačních strategií mimo sledované organizace

Kromě sledovaných technologických společností považujeme za důležité nahlédnout také na další zajímavé aktivity na poli boje proti dezinformaci, které buď byly přímo reakcí na mezník amerických prezidentských voleb v roce 2016, nebo měly do tohoto období přesah.

Mezi tyto aktivity dle našeho mínění patří např. osvěta, strategie zaměřené proti monetizaci dezinformace, monitorovací nástroje, fact-checking a s ním spojené automatizované systémy nebo vládní opatření.

Dá se očekávat, že některé z těchto principů budou přijaty a dále rozvíjeny sledovanými hráči.

Kritické myšlení

Osvěta a vzdělávání veřejnosti v kritickém myšlení je základní stavební blok boje proti dezinformaci. Je to však běh na velmi dlouhou trať a riziko, že s ohledem na prudký rozvoj technologií a sofistikovaných postupů na straně útočníků také s nejistým výsledkem.

Studie deseti strategických předpovědí na období po roce 2018 z dílny analytické společnosti Gartner v předpovědi číslo čtyři očekává výrazný nárůst falešných zpráv.

Gartner ve své studii přímo říká, že *„kolem roku 2022 bude většina obyvatel vyspělých ekonomik konzumovat více falešných než pravdivých zpráv"* (Panetta, 2017). Jinými slovy, bez ohledu na naše kritické myšlení nebudeme v tomto návalu dezinformace vědět, čemu věřit, a čemu ne. Lze očekávat, že v tomto post-truth období, kde dezinformace bude téměř všude kolem nás, bude velmi obtížné odolávat pouze díky osvětě a kritickému myšlení, protože i osvěta bude zmanipulovaná (viz případ factchecking služby demagog.cz, jejíž výsledky byly manipulovány pomocí konkurenčního factchecking webu mujdemagog.cz při parlamentních volbách v ČR v roce 2017) (Břešťan, 2017) – nebude na to čas a naše vrozené kognitivní procesy nás snadno zradí (Cialdini, 2007).

My lidé jsme vizuálně orientovaní tvorové, a ne nadarmo se říká, že „vidět znamená věřit". Více jak 90 % informací, které náš mozek zpracovává, jsou vizuálního charakteru. Proto s pokročilými technologiemi vizualizace (Perry, 2017) a v kombinaci s onou Hydrou informační operace je a bude velmi snadné nás obelstít.

Technologické možnosti dramaticky rostou nejen na straně platforem, které se snaží útokům dezinformací a informačních operací bránit, ale také na straně útočníků, kterým nové technologie s podporou umělé inteligence (AI) mohou pomáhat produkovat stále uvěřitelnější dezinformační artefakty, pro které se ujalo označení *deep-fake* (Snow, 2017).

Příkladem může být realisticky zmanipulované video bývalého amerického prezidenta Baracka Obamy nebo současného ruského prezidenta Vladimíra Putina, kteří tak jsou díky své vizuální reprezentaci s podporou AI pouhými loutkami v rukou uživatele technologie ("A.I. could fabricate fake news", 2017).

Dá se proto očekávat, že závod s útočníky v aréně dezinformace bude hnací silou rozvoje technologií. Bude to dlouhý závod „o prsa" s mnoha oběťmi a pravděpodobně bez konce.

Přes toto všechno se však nelze vzdávat a v oblasti osvěty a vzdělávání v kritickém myšlení je zapotřebí pokračovat.

De-monetizace dezinformace

Zkušenost s americkými prezidentskými volbami v roce 2016 ukázala, že čistě ekonomické cíle a s nimi spojená schémata mohou vážně narušit demokratický proces. Za příklad si zde můžeme vzít skupinu makedonských teenagerů, kteří byli schopni čistě na bázi klikací farmy a analytických nástrojů sociální sítě vybudovat na dezinformaci dobře fungující byznys.

Díky analytickým nástrojům, které sociální síť standardně poskytuje pro podporu návratnosti investic (ROI) inzerentů, skupina rychle zjistila, že téma prezidentských voleb přitahuje zájem a angažovanost čtenářů, přičemž senzacechtivé a dezinformační příspěvky o prezidentském kandidátovi Donaldu Trumpovi přitahovaly více kliků, a tím i výnosů z inzerce, než falešné zprávy o druhém kandidátovi, Hillary Clinton.

Algoritmy identifikující a znevýhodňující virální šíření dezinformace klikacích farem (tedy webových serverů a stránek s dezinformačním obsahem), a tudíž i celou efektivitu jejich obchodního modelu, tak mají šanci na úspěch při potlačování tohoto nežádoucího fenoménu.

Na druhou stranu státem sponzorované informační operace tento ekonomický problém příliš nepostihuje. Dá se ale očekávat, že i v tomto šedém prostředí budou modelově fungovat stejné ekonomické zákony, jako např. u grantů. Ne každý na ně dosáhne, a ne každý má to štěstí mít svůj provoz profinancován ze sta procent, proto se odstřihnutí těchto dezinformačních serverů od jejich přirozeného zdroje příjmů z reklamy jeví jako dobrý směr pro oslabení jejich provozu.

Monitorovací nástroje

Monitorovací nástroje vychází jednak přímo z veřejných dat, poskytovaných jednotlivými technologickými platformami, a pak také z výzkumů specializovaných, tzv. watchdog organizací.

Zmínit např. můžeme databázi Cyber Operations Tracker, která obsahuje záznamy veřejně známých incidentů státem sponzorovaných informačních operací od roku 2005.

Do této kategorie např. patří již výše zmíněný projekt Hamilton 68, Twitter Trails nebo již ukončený opensource projekt Truth Goggles, který byl aktivní od ledna 2011 do září 2013.

Fact-checking

Fact-checking služby představují nezávislé platformy pro ověřování zejména politicky orientovaných zpráv ve veřejném prostoru. Za těmito službami stojí zpravidla neziskové organizace s týmy informačních specialistů a transparentními procesy pro ověřování zpráv a jejich dílčích částí.

Příkladem fact-checking služeb může být např. Poynter.org, Snopes.org nebo český Demagog.cz či HlídacíPes.org.

Tyto služby jsou využívány zejména novináři, kterým usnadňují práci s fakty a pomáhají jim minimalizovat riziko šíření podstrčené a již vyvrácené dezinformace. Fact-checking služby proto úzce spolupracují s vydavateli, novináři a také se zpravodajskými distribučními systémy, jako např. Google News, kterým poskytují svá hodnocení.

Google News společně s dalšími hráči využívá pro ověřování fakt otevřenou technologii protokolu schema.org ClaimReview a s ní spojený proces fact-checkingu, která ve svém značkovacím jazyce umožňuje vydavatelům i fact-checking organizacím doplňovat a verifikovat fakta, řešit zpětnou vazbu s ohledem na revizi udělených hodnocení a tyto informace sdílet.

Otevřená technologie protokolu ClaimReview tak umožňuje efektivní sdílení fact-checking metadat mezi tvůrci obsahu, kurátory, fact-checking specialisty a distribučními kanály.

Od dubna 2017 je protokol ClaimReiew integrován také do výsledků vyhledávání Google ("Google launches Fact Check in search results worldwide", 2017).

Časté jsou již i online fact-checking seance u příležitosti důležitých politických debat, které jsou živě vysílány např. v televizi, a k jednotlivým výrokům je zároveň v reálném čase doplňováno ověření pravdivosti.

Manuální ověřování zpráv je však velmi náročný proces, který se obtížně škáluje. Není tak možné v reálném čase prověřovat všechny publikované zprávy a zajistit pružnou reakci na zpětnou vazbu, která by měla vydané hodnocení ovlivnit. Lidské síly nelze škálovat, proto vše ve srovnání s hektickým tempem internetu trvá dlouho.

Další nevýhodou fact-checkingu je role lidského faktoru fact-checkera, který je dále umocněn nepříliš pružnou reakcí na prověřování zpětných vazeb a úpravy hodnocení. Dá se také očekávat, že u části konzumentů bude vůči konkrétní fact-checking službě existovat předsudek, protože jak již víme z případu dezinformačního serveru Můjdemagog.cz, tak i fact-checking služby mohou být na určité úrovni zmanipulované.

I přes výše uvedená negativa je dle našeho názoru fact-checking v kombinaci s protokolem ClaimReview zatím to nejlepší, co máme v boji proti dezinformaci k dispozici.

Proces fact-checkingu se skládá ze čtyř vzájemně navazujících bloků: Monitoring, Identifikace tvrzení, Ověření tvrzení a Publikování výsledků. Strojově zatím nelze osobu zkušeného fact-checkera plně nahradit, ale již s dnešními technologiemi lze jeho práci výrazně zefektivnit automatizací částí jednotlivých bloků.

S rozvojem strojového učení a profesionalizací fact-checkingu se již nyní rýsují automatizované systémy pro podporu ověřování zpráv. Celá problematika automatizovaného prověřování zpráv je poměrně dobře popsána ve white paperu *The State of Automated Factchecking* (Babakar & Moy, 2016) organizace Full Fact ze srpna 2016.

Další zajímavou aktivitou je soutěž *Fake News Challenge* ("Fake News Challenge", 2017), kde více jak sto dobrovolníků v 71 týmech napříč akademickou a komerční sférou kolem světa soupeří o nalezení technologického řešení pro zefektivnění procesu fact-checkingu pomocí umělé inteligence.

Automatizované systémy

V návaznosti na osvětu, vzdělávání v kritickém myšlení a fact-checking služby je zapotřebí si přiznat a dobře uvědomit, že hlavní bitvy v boji proti dezinformaci budou probíhat na úrovni umělé inteligence, a to na obou stranách.

Útočníci budou díky AI fabrikovat mnohem sofistikovanější, a tudíž i uvěřitelnější vizuální artefakty (Snow, 2017), zasazené do reálně zfabrikovaného kontextu a za výrazné podpory autonomních armád botích trollů a astroturfing skupin a stránek.

Proti nim bude stát opět umělá inteligence, která se bude snažit tyto projevy identifikovat a eliminovat.

Dle našeho názoru jsme na pokraji největšího válečného konfliktu v dějinách lidstva, který bude probíhat bez jediného výstřelu, a navíc v době pomyslného míru. Bez ohledu na počet lidských obětí zde bude nejvíce hmatatelným a snad i jediným výsledkem opravdu silná a autonomní umělá inteligence, která se od nás lidí vycvičila v tom nejhorším. Pandořina skříňka již byla otevřena.

Když odhlédneme od oblasti útočníků, tak rozvoj automatizovaných systémů a strojového učení je nyní hlavně poháněn snahou sledovaných technologických firem o ochranu jejich platforem (např. identifikace falešných účtů a falešných zesílení), kterou pro naše účely navrhujeme označit za rozvoj *taktických obranných systémů pro eliminaci přímých hrozeb*.

Jak jsme si již výše uvedli, vedle těchto taktických obranných systémů pro eliminaci přímých hrozeb vznikají na poli fact-checkingu automatizované systémy pro ověřování pravdivosti jednotlivých tvrzení a sdílení výsledků. Tyto aktivity pro naše potřeby navrhujeme kategorizovat jako *zpravodajské strategie*.

Mezi těmito automatizovanými nástroji vycházejícími z fact-checkingu můžeme vedle již zmíněné Fake News Challenge zmínit např. aktivity technologického startupu AdVerif.ai, který se specializuje na identifikaci manipulativní inzerce a falešných zpráv.

Pro představu si uvedeme také několik příkladů technologií, které mohou být zneužity pro fabrikování dezinformačních artefaktů.

Jako první příklad uveďme aplikaci pro chytré telefony *FaceApp*, která díky neuronové síti dokáže lidskou tvář libovolně transformovat (přidat emoce jako úsměv, zamračení, ubrat nebo přidat roky, změnit pohlaví apod.).

Další zajímavou technologií je projekt *Lyrebird* z Montrealské univerzity. Tato technologie umožní na základě minutového záznamu řeči konkrétní osoby napodobit hlas daného člověka a vyprodukovat tak libovolný audio záznam jeho fiktivního prohlášení.

Poslední ukázkou sofistikovaných technologií na bázi hlubokého strojového učení je technologie *Face2Face* ze Standfordské univerzity, která dokáže mapovat mimiku obličeje jedné osoby do velmi realistického chování videozáznamu obličeje druhé osoby.

Již jen kombinace možností technologie Lyrebird a Face2Face je v rukou útočníka při informační operaci přímo děsivá a naznačuje nám obrysy a rozsah problému, kterému budeme muset v boji proti dezinformaci čelit.

Legislativní opatření

Informační operace přirozeně neprobíhají v legislativním vzduchoprázdnu a lze s nimi bojovat také právní cestou. Zde však není výzvou ani tak nedostatečnost aktuální legislativy, protože problém šíření dezinformace je řešitelný již stávající právní normou řady států, ale pomalost soudního řízení, které může trvat roky a je tak v aktivním boji proti dezinformaci účinné jen velmi málo.

Dobrým příkladem zde může být snaha finské novinářky Jessicy Aro, která se sama stala obětí informační operace a část strategie své obrany postavila na soudní cestě (Aro, 2016).

Dalším aspektem zde jsou vládní opatření, která zákonnou normou v kombinaci s represivními složkami tvrdě potírají případy výskytu dezinformací a s nimi spojených informačních operací.

Příkladem takového přístupu je aktuálně diskutovaná strategie malajské vlády, která za případ dezinformace navrhuje deset let vězení (Robertson, 2018).

Tato opatření však s sebou zároveň přináší pochybnosti o jejich demokratičnosti a etičnosti a vzniká otázka, kdo v tomto případě může kontrolovat nejvyšší státní kontrolní orgány.

Dalším příkladem vládního opatření mohou být rodící se snahy o regulaci sociálních sítí a online služeb.

Strategická komunikace

Strategická komunikace je způsob využívání komunikace státem nebo organizací za účelem podpory nebo dosažení vlastních dlouhodobých cílů (Giampaolo di Paola & Panizzi, 2011). Totéž lze však říci i o informačních operacích, pod které spadají vlivové operace a dezinformace. Hranice mezi strategickou komunikací a propagandou je proto velmi tenká. Strategická komunikace je často vnímána také jako tzv. bílá propaganda (Jowett, O'Donnell, & Jowett, 2012).

Příkladem strategické komunikace v praxi je např. vytvoření již zmíněné specializované organizace StratCom na platformě Evropské unie.

Strategická komunikace je v boji proti dezinformaci velmi důležitá, podobně jako potřeba kritického myšlení.

Bohužel podobně jako u kritického myšlení je i strategická komunikace velmi obtížně škálovatelná vůči neustále se přizpůsobujícím strategiím útočníků a je velmi snadné ji obrátit ve prospěch útočníka (Wanless & Berk, 2018). To můžeme ostatně velmi dobře vidět ve stávající uprchlické krizi, kdy snaha o pomoc potřebným je okamžitě přetavena v nálepkování „vítači“, „sluníčkáři“, „islamizace“ apod.

Model informačních strategií boje proti dezinformaci

Na základě našich dosavadních zjištění se nyní pokusíme jednotlivé identifikované informační strategie boje proti dezinformaci kategorizovat a převést do použitelného modelu.

V rámci základní kategorizace informačních strategií považujeme za užitečné rozdělení dle charakteru přístupu k informačním příležitostem a rizikům na dvě hlavní skupiny:

- Taktické obranné strategie pro eliminaci přímých hrozeb
  (fyzická bezpečnost, crowdsourcing signály, kognitivní analýza apod.)
- Zpravodajské strategie
  (fact-checking, spolupráce s vládními zpravodajskými agenturami…)

Samotné podkategorie informačních strategií ve výše uvedeném specializovaném členění pak dle našeho názoru můžeme rozdělit na:

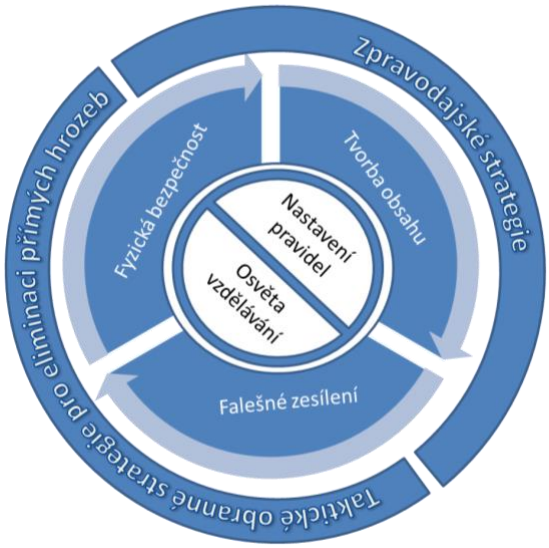
- Nastavení pravidel
  (Community Guidelines, organizační opatření, …)
- Vzdělávání a osvěta
  (Vzdělávání v kritickém myšlení a ochraně soukromí, …)
- Fyzická bezpečnost
  (Cílené a neoprávněné získávání dat, krádeže identit, …)
- Tvorba obsahu
  (Tvorba dezinformačních artefaktů a jejích rozšíření do informačních sídel, podsouvání dezinformačních příběhů a narativu do médií, veřejnosti a astroturfing skupin, tvorba falešných účtů a person apod.)
- Falešné zesílení
  (Falešné účty pro šíření dezinformace sestávající z lidských i botích trollích farem, manipulace pomocí tzv. astroturfing skupin, komentářový spam atd.)

Informační strategie by měla jít napříč celým životním cyklem informace: od vytvoření nebo generování informace přes její shromažďování, zaznamenávání a ukládání, zpracování, distribuci a přenos, spotřebu a užití, recyklaci či vymazání až po nový začátek. Jedná se tedy o cyklický proces, kterému nejlépe odpovídá představa kruhu, nebo spirály.

Náš návrh kategorizace strategií boje proti dezinformaci nejlépe znázorníme jako kružnicový model s jádrem, středovou a obvodovou kružnicí, v rámci kterých definujeme klasifikací jednotlivých přístupů a nástrojů informačních strategií boje proti dezinformaci.

Jádro modelu je tvořeno podkategoriemi *„Nastavení pravidel“* a *„Vzdělávání a osvěta“*. Kolem jádra (ve středové kružnici), se obtáčí podkategorie *„Fyzická bezpečnost“*, *„Tvorba obsahu“* a *„Falešná zesílení“*. Kolem nich (ve vnější obvodové kružnici) si pak představme dvě hlavní kategorie *„Taktické obranné strategie pro eliminaci přímých hrozeb“* a *„Zpravodajské strategie“*.

Pokud bychom do modelu chtěli zahrnout také strategickou komunikaci, pak by tato byla reprezentována prostorem vně kružnice.



Obr. 1 Základní půdorys modelu kategorizace informačních strategií boje proti dezinformaci

Z výše uvedeného je patrná silná provázanost jednotlivých dílčích strategií, kdy v rámci boje proti dezinformaci nelze spoléhat pouze na jednu dílčí strategii. Každé dílčí opatření do jisté míry prochází více kategoriemi, kterých se dotýká a které by měly mít schopnost na toto dílčí opatření navazovat.

Stejně jako v konvenční válce, tak i v kyberprostoru, kde se dezinformační válka odehrává, je pro úspěšné tažení či obranu zapotřebí kombinovat více vzájemně se doplňujících strategických oblastí.[3]

Když si ještě na chvíli vypůjčíme paralelu s konvenčním válečným konfliktem, tak při účinné obraně hranice si nevystačíme pouze se závorou, systémem celní a pasové kontroly a ostnatým drátem, ale musíme mít také adekvátně nastavenou vzdušnou obranu, protileteckou obranu, pozemní a vodní síly, občany vycvičené pro případ konfliktu a také schopnost agresora odstavit protiútokem. V kyberprostoru je to obdobné.

Z této silné provázanosti vyplývá, že pro úspěšný boj proti dezinformaci v rámci komunikační platformy nelze spoléhat jen na jeden přístup, ale je zapotřebí kvalitní mix všech výše uvedených kategorií.

Závěr

Ověřili jsme si, že problematika dezinformace právě díky novým technologickým komunikačním platformám na principu many-to-many přerostla ve zcela bezprecedentní ohrožení společnosti, které oproti původnímu očekávání techno-optimistických vizionářů podkopává samotné principy demokracie.

Společnost zasažená tímto dezinformačním chaosem se polarizuje a fragmentuje do vyhraněných skupin, které se vzájemně netolerují. Tato vzájemná nevraživost skupin se přenáší z online do offline prostředí,

[3] Boj proti dezinformaci zde přirovnáváme ke konvenčnímu pojetí války, protože stejně jako ve válečném konfliktu i zde jsou ohroženy zdroje a uplatněny sofistikované útočné strategie agresora, které ohrožují naši společnost, a je zapotřebí se s nimi adekvátně vyrovnat.

a přerůstá tak až v násilné excesy a ozbrojené konflikty. Je to velmi nebezpečná situace, která se může velmi rychle zvrhnout do dystopické a sebedestruktivní post-truth společnosti, apatické k realitě. Realita je v takové společnosti již jen prázdný pojem.

V takové společnosti vyhrávají demagogové ovládající principy informačních operací a disponující potřebnou technologií. Ukázali jsme si, že jedna takováto organizovaná skupina dokáže ovlivnit demokratické volby na úrovni světové mocnosti, a to při relativně zanedbatelných nákladech.

S prudkým rozvojem tzv. deep-fake technologií na bázi neuronových sítí již u sofistikované informační operace v online komunikační platformě a na jejích periferiích nebude možné rozeznat realitu od dezinformace – bez ohledu na míru kritického myšlení běžného jedince.

Započala tak permanentní hybridní válka v kyberprostoru, vůči které dosud neexistuje efektivní obrana.

Technologické společnosti působící v oblasti sociálních sítí nemají samy o sobě mnoho motivace svůj techno-optimistický přístup měnit. Bez externích motivátorů, které může znamenat např. dobře nastavený vliv státního regulátora, proto není příliš naděje na změnu.

Na základě analýzy sledovaných technologických společností jsme stanovili základní kategorizaci informačních strategií boje proti dezinformaci a promítli ji do první verze funkčního modelu, jaký dosud ještě nebyl pro tuto potřebu vytvořen.

Ukázali jsme si, že v rámci informační strategie boje proti dezinformaci nelze spoléhat pouze na jednu dílčí strategii, ale je nutné pracovat s jejich vyváženým souborem – právě k tomu by měl pomoci i námi vytvořený model a kategorizace.

Z dílčích strategií se jako nejnosnější prokázala strategie postavená na automatizovaných fact-checking systémech s podporou umělé inteligence. Jedná se ovšem zatím jen o oblast velmi mladou, která trpí řadou neduhů – funkčním omezením počínaje a etikou algoritmů konče. Lze však očekávat, že je jen otázkou času a působení externích motivátorů, aby tyto nedostatky byly ve střednědobém horizontu vyřešeny.

Otázkou zůstává, zda je vůbec v moci jedné konkrétní platformy se vlastními silami vypořádat s problematikou dezinformace a informačních operací. Je pravděpodobné, že pro efektivní řešení této problematiky bude zapotřebí spojit síly napříč různými platformami v rámci celého životního cyklu informace a spolupracovat. Zde lze očekávat působení role regulátora.

Uvědomili jsme si také, že rozvoj systémů umělé inteligence v rámci boje proti dezinformaci přirozeně povede k efektu Červené královny, který biologové povýšili na evoluční princip (Morris, 2017). V praxi to bude znamenat nekonečný inovační cyklus, jehož výsledkem bez cílené podpory kritického myšlení bude pouze velmi sofistikovaná umělá inteligence vycvičená v tom nejhorším, co jako lidé máme v sobě.

Použité zdroje

A.I. could fabricate fake news: Artificial intelligence could make fake news even harder to spot. [Online]. (2017). Retrieved April 02, 2018, from https://www.facebook.com/verge/videos/1618100861559584/

Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal Of Economic Perspectives*, Spring 2017(Volume 31, Number 2), 26.

Albright, J. (2017). Total Reach by Page [Online]. Retrieved March 30, 2018, from https://public.tableau.com/profile/d1gi#!/vizhome/FB4/TotalReachbyPage

Armistead, L. (2010). *Information operations matters: best practices*. Washington, D.C.: Potomac Books.

Aro, J. (2016). The cyberspace war: propaganda and trolling as warfare tools, 12.

Babakar, M., & Moy, W. (2016). The State of Automated Factchecking: How to make factchecking dramatically more effective with technology we have now, 36.

Benedictus, L. (2016). Invasion of the troll armies: from Russian Trump supporters to Turkish state stooges [Online]. Retrieved March 31, 2018, from https://www.theguardian.com/media/2016/nov/06/troll-armies-social-media-trump-russian

Boffey, D. Europe's new cold war turns digital as Vladimir Putin expands media offensive: Russia is deploying social media trolls in an attempt to effect political change, and Britain is not immune [Online]. Retrieved March 24, 2018, from https://www.theguardian.com/world/2016/mar/05/europe-vladimir-putin-russia-social-media-trolls

Bowles, N., & Thielman, S. Facebook accused of censoring conservatives, report says [Online]. Retrieved March 25, 2018, from https://www.theguardian.com/technology/2016/may/09/facebook-newsfeed-censor-conservative-news

Břešťan, R. (2017). Hnutí ANO se inspirovalo projektem Demagog, na webu Můj demagog vyvrací „lži o Babišovi" [Online]. Retrieved April 02, 2018, from https://hlidacipes.org/hnuti-ano-se-inspirovalo-projektem-demagog-webu-muj-demagog-vyvraci-lzi-babisovi/

Facebook's failure: did fake news and polarized politics get Trump elected? [Online]. Retrieved November 05, 2017, from https://www.theguardian.com/technology/2016/nov/10/facebook-fake-news-election-conspiracy-theories

Fake Obama created using AI video tool - BBC News [Online]. (2017). Retrieved April 02, 2018, from https://www.youtube.com/watch?v=AmUC4m6w1wo

Fake News Challenge: Exploring how artificial intelligence technologies could be leveraged to combat fake news. [Online]. Retrieved April 01, 2018, from http://www.fakenewschallenge.org/

Ferrara, E. (2017). DISINFORMATION AND SOCIAL BOT OPERATIONS IN THE RUN UP TO THE 2017 FRENCH PRESIDENTIAL ELECTION. *Information Sciences Institute*, 33.

Fredheim, R. (2018). *ROBOTROLLING*. Litva: NATO STRATEGIC COMMUNICATIONS CENTRE OF EXCELLENCE.

Friedman, G. (2015). *Ohrožená Evropa: rodící se krize*. Praha: Aligier s.r.o. společně s Nakladatelstvím Tomáš Krsek.

Garmazhapova, A. Где живут тролли. И кто их кормит: Специальный репортаж из офиса, в котором вешают лапшу в три смены [Online]. Retrieved March 24, 2018, from https://www.novayagazeta.ru/articles/2013/09/07/56253-gde-zhivut-trolli-i-kto-ih-kormit

Giampaolo di Paola, & Panizzi, M. (2011). NATO Military Public Affairs Policy [Online]. Retrieved October 20, 2018, from https://www.nato.int/ims/docu/mil-pol-pub-affairs-en.pdf

Giles, K. (2016). THE NEXT PHASE OF RUSSIAN INFORMATION WARFARE, 16.

Giles, K. (2016). Russia's 'New' Tools for Confronting the West: Continuity and Innovation in Moscow's Exercise of Power. *Russia And Eurasia Programme*, 73.

Giles, K. (2016). HANDBOOK OF RUSSIAN INFORMATION WARFARE. *Fellowship Monograph*, 90.

Glaser, A. (2017). What Was Russia Up To?: Here's what we know about how Russia tried to use Facebook, Google, and Twitter to sway the 2016 election. [Online]. Retrieved March 30, 2018, from http://www.slate.com/articles/technology/future_tense/2017/10/what_we_know_about_russia_s_use_of_american_facebook_twitter_and_google.html

Google launches Fact Check in search results worldwide: After a tentative launch in October 2016, Google has released its Fact Check feature in search results worldwide. [Online]. (2017). Retrieved April 01, 2018, from https://searchenginewatch.com/2017/04/10/google-launches-fact-check-in-search-results-worldwide/

Gorwa, R. (2017). *Computational Propaganda in Poland: False Amplifiers and the Digital Public Sphere: Computantional Propaganda Research Project* (1st ed.). UK: University of Oxford.

Goda, J. Pod falošným menom som písal hoaxy pre Hlavné správy [Online]. Retrieved March 25, 2018, from https://dennikn.sk/995647/pod-falosnym-menom-som-pisal-hoaxy-pre-hlavne-spravy/

Griffin, A. (2018). DELETE FACEBOOK CAMPAIGN TAKES OFF – BUT ACTUALLY REMOVING YOUR DATA MIGHT PROVE MORE DIFFICULT THAN IT SEEMS [Online]. Retrieved March 30, 2018, from https://www.independent.co.uk/life-style/gadgets-and-tech/news/delete-facebook-cambridge-analytica-campaign-deactivate-data-remove-hide-privacy-a8266671.html

Grush, L. (2018). Elon Musk has removed Tesla and SpaceX's Facebook pages after Twitter challenge [Online]. Retrieved March 30, 2018, from https://www.theverge.com/2018/3/23/17156402/elon-musk-deleted-tesla-and-spacex-facebook-pages-twitter-challenge

Howard, J. G. (2015). *Information Operations and the Islamic State* (Master's Capstone Theses). Charles Town, WV.

Hybrid influence – lessons from Finland [Online]. (2017). Retrieved November 05, 2017, from https://www.nato.int/docu/review/2017/Also-in-2017/lessons-from-finland-influence-russia-policty-security/EN/index.html

Jowett, G., O'Donnell, V., & Jowett, G. (c2012). *Propaganda & Persuasion* (5th ed). Thousand Oaks, Calif.: SAGE.

MacFarquhar, N. (2016). A Powerful Russian Weapon: The Spread of False Stories [Online]. Retrieved April 02, 2018, from https://www.nytimes.com/2016/08/29/world/europe/russia-sweden-disinformation.html

McCullagh, D. Obama Harvested Data from Facebook and Bragged About It. Why Are We Only Freaking Out About This Now?: Why are politicians now freaking out about a feature that has been publicly documented since its inception and that was discontinued three years ago? [Online]. Retrieved March 25, 2018, from http://reason.com/archives/2018/03/23/cambridge-analytics-dust-up-reveals-lawm

Mosseri, A. (2016). Building a Better News Feed for You [Online]. Retrieved March 30, 2018, from https://newsroom.fb.com/news/2016/06/building-a-better-news-feed-for-you/

Mosseri, A. (2017). Improving the Integrity of News Feed: Principles for Publishers [Online]. Retrieved April 07, 2018, from https://media.fb.com/2017/05/19/improving-the-integrity-of-news-feed-principles-for-publishers/

Morris, I. (2017). *K čemu je dobrá válka?: konflikty a pokrok civilizace*. Praha: Argo.

Newton, C. (2016). Facebook partners with fact-checking organizations to begin flagging fake news: Plus new tools for reporting hoaxes [Online]. Retrieved April 07, 2018, from https://www.theverge.com/2016/12/15/13960062/facebook-fact-check-partnerships-fake-news

Newton, C. (2017). Senators blast tech companies over Russian meddling: 'Do something about it — or we will': Sen. Dianne Feinstein leads charge against Facebook, Google, and Twitter [Online]. Retrieved March 30, 2018, from https://www.theverge.com/2017/11/1/16591646/facebook-senate-hearing-feinstein-russia-google-twitter

Nothing is real: How German scientists control Putin's face [Online]. (2016). Retrieved April 02, 2018, from https://www.youtube.com/watch?v=ttGUiwfIYvg

Panetta, K. (2017). Gartner Top Strategic Predictions for 2018 and Beyond: From bots and AI to counterfeit reality and fake news, these predictions require IT leaders to pace their adoption. [Online]. Retrieved April 02, 2018, from https://www.gartner.com/smarterwithgartner/gartner-top-strategic-predictions-for-2018-and-beyond/

Parakilas, S. (2017). We Can't Trust Facebook to Regulate Itself [Online]. Retrieved March 30, 2018, from https://www.nytimes.com/2017/11/19/opinion/facebook-regulation-incentive.html

Perry, P. (2017). These A.I. tools could lead to the next generation of fake news [Online]. Retrieved April 02, 2018, from http://bigthink.com/philip-perry/these-ai-tools-could-lead-to-the-next-generation-of-fake-news

Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach: Whistleblower describes how firm linked to former Trump adviser Steve Bannon compiled user data to target American voters [Online]. (2018). Retrieved March 30, 2018, from https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

Robert B. Cialdini. (2007). *Influence: the psychology of persuasion* (Rev. ed.). New York: Collins.

Robertson, A. (2018). Malaysian government threatens 10-year prison sentences for pushing fake news [Online]. Retrieved April 01, 2018, from https://www.theverge.com/2018/3/26/17163920/malaysia-fake-news-law-proposal-election-najib-razak

Romm, T., & Wagner, K. (2017). Facebook admits 'malicious actors' spread misinformation during the 2016 U.S. election: It also cites a government report that found Russia played a major role in the presidential race. [Online]. Retrieved April 22, 2018, from https://www.recode.net/2017/4/28/15476142/facebook-report-trump-clinton-russia-us-presidential-election

Roy, E. A. (2018). Facebook rolls out trial of 'dislike' button for downvoting comments [Online]. Retrieved July 08, 2018, from https://www.theguardian.com/technology/2018/may/01/facebook-rolls-out-trial-of-dislike-button-for-downvoting-comments

Shane, S., & Goel, V. (2017). Fake Russian Facebook Accounts Bought $100,000 in Political Ads [Online]. Retrieved March 30, 2018, from https://www.nytimes.com/2017/09/06/technology/facebook-russian-political-ads.html

Section 230: A Key Legal Shield For Facebook, Google Is About To Change: THE NEW CLASH BETWEEN FREE SPEECH VS. PRIVACY [Online]. (2018). Retrieved March 31, 2018, from https://www.npr.org/sections/alltechconsidered/2018/03/21/591622450/section-230-a-key-legal-shield-for-facebook-google-is-about-to-change

Silverman, C. (2016). This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook [Online]. Retrieved April 02, 2018, from https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook

Snow, J. (2017). AI Could Set Us Back 100 Years When It Comes to How We Consume News [Online]. Retrieved April 02, 2018, from https://www.technologyreview.com/s/609358/ai-could-send-us-back-100-years-when-it-comes-to-how-we-consume-news/

Snow, J. (2017). Can AI Win the War Against Fake News?: Developers are working on tools that can help spot suspect stories and call them out, but it may be the beginning of an automated arms race. [Online]. Retrieved April 02, 2018, from https://www.technologyreview.com/s/609717/can-ai-win-the-war-against-fake-news/

Taplin, J. T. (2017). *Move fast and break things: how Facebook, Google, and Amazon cornered culture and undermined democracy*. New York: Little, Brown and Company.

The Law that Gave Us the Modern Internet—and the Campaign to Kill It: Ever heard of Section 230 of the Communications Decency Act? It gave birth to the social web. Here's why we need more laws just like it. [Online]. (2013). Retrieved March 31, 2018, from

https://www.theatlantic.com/business/archive/2013/09/the-law-that-gave-us-the-modern-internet-and-the-campaign-to-kill-it/279588/

Thompson, N., & Vogelstein, F (2018). Inside the Two Years That Shook Facebook—and the World: How a confused, defensive social media giant steered itself into a disaster, and how Mark Zuckerberg is trying to fix it all. [Online]. Retrieved March 25, 2018, from https://www.wired.com/story/inside-facebook-mark-zuckerberg-2-years-of-hell/

Tynan, D. (2016). How Facebook powers money machines for obscure political 'news' sites: From Macedonia to the San Francisco Bay, clickbait political sites are cashing in on Trumpmania – and they're getting a big boost from Facebook [Online]. Retrieved April 02, 2018, from

https://www.theguardian.com/technology/2016/aug/24/facebook-clickbait-political-news-sites-us-election-trump

Ordway, D. -M. (2017). Fake news and the spread of misinformation [Online]. Retrieved April 29, 2018, from https://journalistsresource.org/studies/society/internet/fake-news-conspiracy-theories-journalism-research

Wanless, A., & Berk, M. (2018). The Strategic Communication Ricochet: Planning Ahead for Greater Resiliency [Online]. Retrieved October 20, 2018, from https://thestrategybridge.org/the-bridge/2018/3/7/the-strategic-communication-ricochet-planning-ahead-for-greater-resiliency

Weedon, J., Nuland, W., & Stamos, A. (2017). Information Operations and Facebook: Verze 1.0 [Online], 13. Retrieved from https://fbnewsroomus.files.wordpress.com/2017/04/facebook-and-information-operations-v1.pdf

What is the future of news? Bleak, probably.: A professional fact-checker explains why our fake news problem isn't going away. [Online]. Retrieved March 25, 2018, from https://www.vox.com/conversations/2016/11/28/13714596/media-democracy-donald-trump-fake-news-internet-journalism-social-media

Why Is Finland Able to Fend Off Putin's Information War?: Helsinki has emerged as a resilient front against Kremlin spin. But can its successes be translated to the rest of Europe? [Online]. Retrieved November 05, 2017, from http://foreignpolicy.com/2017/03/01/why-is-finland-able-to-fend-off-putins-information-war/

Wong, J. C. (2018). Facebook The Cambridge Analytica Files Mark Zuckerberg apologises for Facebook's 'mistakes' over Cambridge Analytica: Following days of silence, CEO announces Facebook will change how it shares data with third-party apps and admits 'we made mistakes' [Online]. Retrieved March 30, 2018, from https://www.theguardian.com/technology/2018/mar/21/mark-zuckerberg-response-facebook-cambridge-analytica