costaid

Coping Strategies Against Information Disorder

Smernice

Sofinancira Evropska unija

Financira Evropska unija. Izražena stališča in mnenja so izključno stališča in mnenja avtorjev in ne odražajo nujno stališč in mnenj Evropske unije ali Izvajalske agencije za izobraževanje in kulturo (EACEA). Niti Evropska unija niti EACEA ne moreta biti odgovorna zanje.

Avtorji:

Eliane Smits van Waesberghe in Tim Paulusse - Verwey-Jonker Instituut (glavna urednika)

Leen D'Haenens in Joyce Vissenberg - KU Leuven

Tzvetalina Genova - Mednarodni inštitut za management

Wolfgang Eisenreich - Wissenschaftsinitiative Niederösterreich

Sonja Bercko Eisenreich - Inštitut Integra

Alenka Valjaškova - QUALED

Pantelis Balaouras - Priključki

Izjava o avtorskih pravicah:

Lahko:

- delite - kopirajte in razširjajte gradivo v katerem koli mediju ali formatu.
- prilagajanje - remiks, preoblikovanje in nadgradnja gradiva.

pod naslednjimi pogoji:

- Priznanje avtorstva - Navesti morate ustrezno priznanje, povezavo do licence in morebitne spremembe. To lahko storite na kakršen koli primeren način, vendar ne na način, ki bi nakazoval, da izdajatelj licence podpira vas ali vašo uporabo.
- Nekomercialno - gradiva ne smete uporabljati v komercialne namene.
- ShareAlike - Če gradivo remiksiramo, preoblikujemo ali gradimo na njem, moramo svoje prispevke distribuirati pod isto licenco kot izvirnik.

costaid
Coping Strategies Against Information Disorder

Poglavje 3
Tehnologija in orodja

Ciljna skupina

Te smernice so namenjene tako imenovanim "prvim uporabnikom". "First-liners" je krovni izraz za vse osebe, ki so v neposrednem stiku z osebami, ki so izpostavljene informacijskim motnjam, s poudarkom na skupinah v poklicnem izobraževanju. Primeri oseb, ki spadajo pod ta krovni izraz, so: vzgojitelji, učitelji, trenerji, mladinski svetovalci in svetovalci, socialni delavci in mladinski delavci. Vendar ta seznam ni izčrpen. Področje uporabe tega projekta vključuje tudi druge osebe, ki delajo v izobraževalnem, socialnem ali zdravstvenem sektorju.

3.1 Uvod

Koncept dezinformacij (izmišljenih novic) je bil obravnavan v prejšnjih poglavjih. Novice običajno razširjajo profesionalni ponudniki novic, vključno z javnimi mediji, komercialnimi mediji, neodvisnim profesionalnim novinarstvom ali amaterskimi uporabniki (v primeru platform družbenih medijev). Novice so dostopne in na voljo v različnih oblikah:

- **Digitalne novice**: digitalne novice: novice, ki se razširjajo prek spletnih kanalov v digitalni medijski obliki (besedilo, slike, zvok, video).
- **Tiskani mediji**: Časopisi in revije z besedilno in slikovno vsebino.
- **Oddajanje**: Televizija in radio z video in avdio vsebinami.

V tem poglavju se osredotočamo predvsem na prvo obliko, digitalne novice. Vendar se lahko predstavljene razprave uporabljajo tudi za druge formate, ki uporabljajo video, zvok in slike.

Digitalne novice so večinoma dostopne prek novičarskih platform, ki jih upravljajo profesionalni ponudniki novic. Uporabniki lahko do teh novičarskih platform dostopajo brezplačno ali prek naročniških storitev. Te platforme ponujajo spletne vire v priljubljenih brskalnikih in vire novic na platformah družbenih medijev, kar uporabnikom omogoča prejemanje prilagojenih novic.

Novičnik je spletna stran ali zaslon, ki se pogosto posodablja in prikazuje najnovejše novice ali informacije. Prilagojene novice so storitve, vključene v spletne brskalnike (spletni kanali) in platforme družabnih medijev, ki uporabnikom dostavljajo novice glede na njihove osebne preference. Te preference so določene na podlagi različnih dejavnikov, kot so naročanje na spletne kanale, obiskovanje določenih spletnih strani in drugo.

Na platformah družbenih medijev uporabniki prejemajo tudi skupne medije drugih uporabnikov, ki lahko vključujejo napačne informacije.

Prilagojene novice se v brskalniku ali na platformi družabnih medijev prikažejo prek storitev in algoritmov družabnih medijev. Uporabniki se morda ne zavedajo v celoti, kako ti algoritmi delujejo, saj ostaja nejasno, ali izbira novic temelji izključno na željah uporabnika ali na drugih merilih. Platforme družbenih medijev na primer predstavljajo novice na podlagi ponudnikov novic, ki jim uporabnik sledi, bralnih navad njegovih prijateljev ali nedavno kliknjenih člankov, kar lahko tvori vrsto preferenc. Posledično je seznam novic filtriran, kar pomeni, da niso prikazane vse novice, temveč tiste, za katere algoritem meni, da so za uporabnika najbolj zanimive. Mnogi trdijo, da to ustvarja informacijski mehurček, v katerem novice izbira algoritem in morda ne vključuje novic, ki uporabnika zanimajo, vendar jih algoritem ne vključi. Zato je pomembno, da se ne zanašamo samo na osebne vire novic, temveč obiščemo profesionalne platforme ponudnikov novic, ki jim zaupamo. Vedno je pomembno kritično oceniti verodostojnost in točnost novic, ne glede na obliko ali način dostave.

3.2 Iskalniki in algoritmi

Številna spletna mesta na internetu si prizadevajo, da bi uporabnike čim bolj pritegnila in povečala njihov čas, ki ga preživijo na platformi. To še posebej velja za platforme družbenih medijev, ki uporabljajo različne strategije za izboljšanje uporabniške izkušnje. Ena takšnih strategij je prikazovanje vsebine, ki je v skladu z interesi uporabnikov. Vendar spletna mesta ne morejo brati misli in poznati preferenc uporabnikov. Da bi to rešili, se uporabljajo algoritmi, ki analizirajo podatke o uporabnikih in zagotavljajo prilagojeno vsebino.

Algoritmi so zapletene formule, ki opazujejo, merijo in izračunavajo posameznikove preference glede vsebine. To lahko vključuje dejavnike, kot so čas gledanja določenih vrst videoposnetkov, čas, ki ga uporabnik preživi na določeni objavi, ali dejanja sodelovanja, kot so puščanje všečkov ali komentarjev. Z analizo teh podatkov algoritmi določijo vrsto vsebine, ki uporabnike zadrži dlje časa. Te informacije se nato uporabijo za izbiro in priporočanje podobnih vsebin uporabniku.

Čeprav se ta pristop zdi logičen in neškodljiv, ima algoritemska uporaba tudi potencialne slabosti. Algoritem, ki daje prednost vsebini, ki se zdi uporabniku zanimiva, lahko ustvari filtrirne mehurčke, kjer so uporabniki izpostavljeni le določenim stališčem. Filtrirni mehurčki omejujejo raznolikost vsebine, kar lahko vodi v odmevne komore ali krepi obstoječe.

Poleg filtrirnih mehurčkov lahko algoritmi povečajo ekstremnost vsebine s priporočanjem vedno bolj nišnih, obrobnih in ekstremnih objav. Cilj je, da bi bili uporabniki vključeni, vendar lahko to privede do tega, da so uporabniki usmerjeni v spletne prostore brez različnih stališč, kar povzroči tako imenovano "zajčjo luknjo".

Filtrirni mehurčki in zajčje luknje uporabnike izpostavljajo radikalnim vsebinam in z njimi povezanim skupnostim. Ti spletni prostori so plodna tla za razvoj, rast, izkrivljanje in širjenje napačnih in dezinformacijskih informacij.

Ko uporabniki napredujejo skozi zajčjo luknjo, postanejo skrajne govorne točke in lažne informacije normalne. Ta normalizacija dodatno olajša širjenje in sprejemanje napačnih informacij in dezinformacij iz radikalnih virov.

Z razumevanjem vpliva algoritmov na uporabniško izkušnjo lahko bolje razumemo tveganja, povezana s filtrirnimi mehurčki, komorami odmevov in zajčjimi luknjami. To znanje je bistvenega pomena za krmarjenje po spletni pokrajini in reševanje izzivov, ki jih prinaša informacijska neurejenost.

3.3 Spletne strategije dezinformacijskih in ekstremističnih organizacij

V prejšnjih razdelkih smo raziskali, kako se informacijski nered ustvarja in ohranja v komorah odmevov in filtrirnih mehurčkih. Kako pa napačne informacije dosežejo ljudi zunaj teh prostorov?

Najpogostejše metode

Napačne in dezinformativne informacije se lahko širijo z neštetimi oblikami komunikacije. Vendar se najpogosteje širijo prek različnih kanalov, kot so platforme družbenih medijev, spletna mesta, elektronska pošta in ustno izročilo. Najpogostejši načini, s katerimi se širijo dezinformacije, so tesno povezani s sedmimi kategorijami problematičnih vsebin, ki so obravnavane v *poglavju 1: Razumevanje "lažnih* novic":

- **Satira ali parodija**: Nekatere napačne in dezinformativne informacije so ustvarjene v zabavne namene ali kot satira, vendar jih je mogoče napačno razumeti kot prave novice. Satirična spletna mesta ali računi v družabnih medijih lahko objavljajo humorne ali pretirane zgodbe, vendar jih lahko bralci, ki se ne zavedajo njihove satirične narave, zamenjajo za dejanske informacije.
- **Clickbait**: članki, ki vsebujejo napačne ali dezinformativne informacije, pogosto uporabljajo senzacionalne ali zavajajoče naslove, da bi pritegnili pozornost in dosegli več klikov ali ogledov. Njihov namen je izkoristiti radovednost ali čustva ljudi, da bi spodbudili promet na spletno mesto in ustvarili prihodek z oglaševanjem.
- **Zavajanje**: Pri tem gre za izkrivljanje ali napačno predstavljanje dejanskih novic s selektivnim predstavljanjem dejstev ali izpuščanjem ključnih informacij. Lahko gre za jemanje izjav iz konteksta, spreminjanje slik ali videoposnetkov ali sprevračanje pomena zgodbe, da ustreza določeni zgodbi.
- **Upodabljanje**: napačne in dezinformativne informacije lahko vključujejo tudi upodabljanje uglednih virov novic ali javnih osebnosti, da se lažnim informacijam doda verodostojnost. To je mogoče storiti z ustvarjanjem lažnih spletnih strani ali računov v družabnih medijih, ki posnemajo legitimne vire in bralce zavedejo, da so informacije vredne zaupanja.
- **Politična manipulacija**: Včasih se informacijske motnje ustvarjajo ali širijo z namenom vplivanja na javno mnenje ali volitve. To lahko vključuje širjenje lažnih informacij o političnih kandidatih, manipuliranje z javnim razpoloženjem ali izkoriščanje obstoječih predsodkov in delitev v družbi.
- **Fabrikacija**: napačne in dezinformativne informacije so lahko v celoti izmišljene in nimajo nobene podlage v resničnosti. Gre za ustvarjanje lažnih zgodb, citatov ali dogodkov za zavajanje bralcev ali gledalcev.

- **Povečanje prek družabnih medijev**: Pri širjenju napačnih in dezinformacijskih informacij imajo pomembno vlogo platforme družbenih medijev. Lažne zgodbe lahko hitro postanejo viralne, saj jih uporabniki delijo in ponovno objavljajo, pogosto brez preverjanja točnosti informacij. Algoritmi, ki jih uporabljajo platforme družbenih medijev, lahko prav tako prispevajo k širjenju, saj spodbujajo vsebino, ki temelji na vključenosti in ne na točnosti.

Za širjenje napačnih in dezinformacijskih informacij se pogosto uporabljajo tudi druge metode, ki jih je treba podrobneje obravnavati: trolling, doxxing in mainstreaming. To je opisano v naslednjih dveh podpoglavjih.

Trolanje in doxxing

Ena od pogosto uporabljenih strategij je trolanje. Trolanje je opredeljeno kot namerna uporaba nevljudnosti, agresivnosti, zavajanja in manipulacije v spletni komunikaciji, da bi izzvala konflikt ali zabavo. Trolanje z zavajanjem, manipuliranjem ali agresivnostjo sproža spletne konflikte. Za lastno zabavo izrivajo pogovore in se v bistvu ukvarjajo z digitalnim ustrahovanjem.

V majhnem obsegu se lahko zdi, da je trolanje razmeroma neškodljivo in da je zgolj nadležno. Vendar pa se lahko, ko imajo organizirane skupine trolov določen cilj, ta nadloga hitro spremeni v epidemijo dezinformacij. Primer tega je ruska uporaba trolov v družbenih medijih kot "orožja". Rusija je uporabila veliko mrežo trolov, ki so po vsem svetu širili dezinformacije v več jezikih, da bi nadzorovali spletni diskurz o Rusiji. Ti troli niso samo širili lažnih informacij, ampak so tudi ciljali na uporabnike družbenih medijev z objavami, ki so odstopale od pripovedi, za katero so bili organizirani. Posledično so se številni uporabniki družbenih medijev vzdržali razpravljanja o Rusiji, kar je trolom omogočilo, da so s svojimi napačnimi informacijami nadzorovali pripoved.

Doxxing, druga oblika spletnega ustrahovanja, vključuje razkrivanje osebnih podatkov ali identitete posameznikov na spletu brez njihove privolitve. Čeprav ta taktika ne vključuje posebej širjenja dezinformacij, je še ena od strategij, ki jih uporabljajo spletni troli, da bi nadzorovali pripoved o določeni temi, podobno kot trolanje. Doxxing se lahko uporablja za ustrahovanje uporabnikov družbenih medijev, s čimer se zmanjša njihova pripravljenost za objavo vsebine, ki je v nasprotju z želeno pripovedjo trola.

Razumevanje vpliva trolanja in doxxinga je ključno za prepoznavanje različnih taktik, ki se uporabljajo za manipulacijo in nadzor spletnih pripovedi. Te strategije ne prispevajo le k širjenju lažnih informacij, temveč predstavljajo tudi izziv pri spodbujanju odprtega in informiranega digitalnega okolja.

Vključevanje

Pomembna strategija širjenja dezinformacij in ekstremističnih vsebin je njihova normalizacija ali "vključevanje". Pri tem ima ključno vlogo izpostavljanje. Izpostavljenost napačnim in dezinformacijam lahko privede do vztrajnih napačnih prepričanj ljudi v zvezi z določenimi temami, kar normalizira napačne ideje v njihovih glavah. To izpostavljanje se lahko zgodi v različnih oblikah.

Vplivneži

Ena od pogostih oblik izpostavljenosti je razširjanje informacij med posamezniki, ko posamezniki delijo informacije z drugimi. To se lahko zgodi z osebnimi stiki ali v večjem obsegu z vplivneži v družbenih medijih. Vplivneži, ki imajo velik doseg v različnih skupinah, lahko nevede ali namerno delijo napačne informacije, kar vpliva na veliko število posameznikov. Takšna široka izpostavljenost vodi v normalizacijo napačnih informacij med različnimi občinstvi.

Ironija, satira in memi

Ekstremistični posamezniki in organizacije pri širjenju svojih idej pogosto uporabljajo humor, satiro in ironijo.

Satira je lahko močno orodje za izpodbijanje zatiralskih ideologij, spreminjanje pripovedi ali normalizacijo nišnih pogledov v glavnem toku. Na področju dezinformacij se satira uporablja na različnih področjih. Parodična spletna mesta, kot sta The Onion ali De Speld, objavljajo neresnične vsebine za namene humorja, ne da bi nameravala zavajati javnost. Nekateri posamezniki in skupine pa satiro in ironijo uporabljajo z zlonamernim namenom, da bi diskreditirali glavno novinarstvo, znanost ali spodbujali ekstremistične ideje in dezinformacije. Z uporabo satire in humorja postane takšna vsebina v političnem diskurzu dostopnejša in sprejemljivejša, s čimer je izpostavljena širšemu občinstvu.

Ekstremistične vsebine so pogosto zanimive za mlade kot oblika zabave ali iskanja senzacij. Mladi posamezniki, ki iščejo smisel, težijo k intenzivnim in novim izkušnjam, zato so bolj dovzetni za ekstremistične ideje in z njimi povezane dezinformacije.

Memi, ki so široko deljeni kosi humornih kulturnih vsebin, so še ena možnost za širjenje ekstremističnih ideologij. Memi so na voljo v različnih oblikah, vključno s slikami, videoposnetki, zvočnimi posnetki, emojiji in simboli. Čeprav memi sami po sebi niso škodljivi, jih skrajneži uporabljajo za normalizacijo svojih idej. Igriva narava memov ekstremistom omogoča, da prikrijejo, ovržejo ali zanikajo škodljivost svojih sporočil. Takšna "ostra" ali provokativna vsebina postane sprejemljivejša, ob obtožbah o seksizmu, rasizmu ali ksenofobiji pa jo lahko ustvarjalci zlahka zavrnejo kot "samo šalo". To brisanje meja med igrivimi potegavščinami in problematično vsebino ustvarja dvoumnost, zaradi česar je težko ločiti nedolžne šale od ekstremističnih sporočil. Žabec Pepe, lik iz spletne risanke, ki je bil sprva ustvarjen kot neškodljiva šala, so si ga prisvojili spletni belski supremacisti. To je povzročilo zmedo pri internetnih uporabnikih, saj so se ekstremistične različice tega mema

pomešale z neškodljivimi. Do normalizacije ekstremističnih vsebin pride, ko je tem sporočilom izpostavljenih več posameznikov, kar zabriše meje med tem, kaj je sprejemljivo in kaj ne.

3.4 Manipulirana vsebina

S tehničnega vidika so vse informacije ali "novice" v medijih kombinacija besedila, slike, zvoka in videa. Vendar pa je treba ugotoviti, ali so informacije verodostojne ali ne. Opozoriti velja, da lahko napačne informacije uporabljajo pristne slike, vendar zgodbo manipulirajo in izkrivljajo dejanska dejstva.

V preteklosti je veljalo, da lahko besedilo napiše vsakdo, medtem ko se je za slike, zvok in videoposnetke domnevalo, da so bolj ali manj avtentični in da je za njihovo spreminjanje potrebno strokovno znanje. Vendar je z nedavnim tehnološkim napredkom mogoče manipulirati tudi s slikami, zvokom in videom. To lahko dosežejo strokovnjaki ali aplikacije, ki uporabljajo sisteme umetne inteligence, kot je tehnologija deepfake. Zato je treba razlikovati, ali je bil zvok ali video posnetek resnično posnet z mikrofonom ali videokamero ali pa je rezultat strokovnega urejanja ali sistemov umetne inteligence (generativna umetna inteligenca in sintetični mediji: glasovni kloni, videoposnetki **deepfake).** Poleg tega bi moralo biti tehnično izvedljivo ugotoviti prvotni vir, proizvajalca ali izdajatelja slike, zvočnega ali video vira. To pa zato, ker je vire mogoče večkrat deliti, kopirati ali razširjati po svetovnem spletu in družabnih medijih. Zato je lahko za običajne uporabnike težko določiti prvotni vir in proizvajalca, tudi če sumijo, da je novica morda napačna informacija.

Da lahko uporabniki razlikujejo med resničnimi in izmišljenimi novicami, je potrebnih več korakov. Več informacij o tem najdete v *poglavju 2: Ukrepi in veščine*. Priporočamo, da to poglavje preberete, da se seznanite z zapletenimi podrobnostmi za prepoznavanje lažnih informacij. Vendar je tu kratek, zelo poenostavljen povzetek:

- **Korak 1: Ozaveščanje uporabnikov:** Uporabniki se morajo zavedati, da so novice lahko izmišljene. Izvajanje dejavnosti ozaveščanja je ključnega pomena za obveščanje uporabnikov o tem, kaj so izmišljene novice in kako se lahko zaščitijo pred njihovimi posledicami.
- **Korak 2: Preverite zanesljivost založnika**: Večja ozaveščenost o izmišljenih novicah spodbuja uporabnike, da se sprašujejo o zanesljivosti virov novic in izdajateljev. Pri tem je treba upoštevati obliko medija, ne glede na to, ali gre za televizijski kanal, revijo, časopis (spletni ali tiskani) ali platformo družbenih medijev. Medijski kanali, ki omogočajo enostavno deljenje ali razširjanje novic, so lahko manj zanesljivi. Nasprotno pa so medijski kanali, ki olajšajo identifikacijo in preverjanje vira, običajno bolj zanesljivi.

V zvezi z novicami, ki se razširjajo prek interneta, bi morali ponudniki storitev, kot so novičarski servisi in omrežja družbenih medijev, uporabiti nove tehnologije za preverjanje zanesljivosti virov ter sledenje prvotnemu izdajatelju in viru. Tehnologija veriženja blokov je ena od takšnih tehnologij, ki lahko olajša ta prizadevanja.

Z upoštevanjem teh korakov in izkoriščanjem tehnologije lahko uporabniki pridobijo moč, da se znajdejo v digitalnem okolju, ločijo prave novice od napačnih informacij in sprejemajo informirane odločitve o informacijah, ki jih najdejo.

Razumevanje deepfakes: sintetične medijske manipulacije

Deepfakes so po definiciji Cambridge Dictionary "*video ali zvočni posnetki, ki zamenjajo obraz ali glas nekoga drugega tako, da je videti resničen*".

V članku "Deepfake explained" iz leta 2020 pisateljica Meredith Somers navaja, da "*(a) deepfake pomeni posebno vrsto sintetičnih medijev, kjer je oseba na sliki ali videoposnetku zamenjana s podobo druge osebe*". Poleg tega je pojasnjeno, da je "*izraz 'deepfake' konec leta 2017 prvič skoval istoimenski uporabnik Reddita. Ta uporabnik je na spletnem mestu za novice in združevanje ustvaril prostor, kjer je delil pornografske videoposnetke, ki so uporabljali odprtokodno tehnologijo zamenjave obrazov.*"

Globoki ponaredki se uporabljajo v različnih sektorjih in za različne namene. Nekateri pomembni primeri vključujejo:

- **izsiljevanje:** Globoki ponaredki se lahko uporabijo za ustvarjanje lažnega obremenilnega gradiva, kar lahko vodi v izsiljevanje. Poleg tega lahko žrtve dejanskega izsiljevanja trdijo, da so dokazi ponarejeni, kar jim omogoča verodostojno zanikanje, saj je globoke ponaredke vedno težje razlikovati od pristne vsebine.
- **Pornografija:** Na internetu je postala zelo razširjena globoka pornografija. Poročilo nizozemskega zagonskega podjetja za kibernetsko varnost Deeptrace ocenjuje, da je bilo približno 96 % vseh spletnih globokih ponaredkov pornografskih.
- **Politika:** Globoki ponaredki so bili uporabljeni za napačno predstavitev znanih politikov v videoposnetkih, širjenje dezinformacij in manipuliranje z javnim dojemanjem. Primeri vključujejo globoke ponaredke, v katerih nastopajo Barack Obama, Donald Trump, Volodymyr Zelensky in Vladimir Putin.
- **Igranje/filmi:** Obstajajo ugibanja o uporabi deepfakes za ustvarjanje digitalnih igralcev v prihodnjih filmih. Čeprav so bili digitalno ustvarjeni ali spremenjeni ljudje v filmih prikazani že prej, bi lahko deepfakes prispevali k novim dosežkom na tem področju.
- **Družbeni mediji:** Uporabniki so globoke ponaredke uporabljali na različnih platformah družbenih medijev. Posamezniki zamenjajo obraze v prizorih priljubljenih filmov ali serij s svojimi in tako ustvarijo osebne videoposnetke. Platforme, kot je Facebook, so sprejele ukrepe za odkrivanje in označevanje deepfakes kot lažnih, s čimer so zmanjšale njihovo prednost v kanalih uporabnikov.

Odkrivanje in blaženje globokih ponaredkov: tehnološki pristopi

Raziskovalci dejavno raziskujejo metode za odkrivanje in obravnavanje problema lažnih zvočnih in video posnetkov. Uporabljajo različne pristope:

- **Algoritemsko odkrivanje:** Eden od pristopov vključuje razvoj algoritmov, ki lahko prepoznajo manipulirano vsebino. Ti algoritmi analizirajo različne vizualne in zvočne znake, da bi odkrili neskladnosti ali nepravilnosti, ki kažejo na prisotnost globokega ponaredka. Z uporabo tehnik strojnega učenja in umetne inteligence lahko ti algoritmi sčasoma izboljšajo svoje zmožnosti odkrivanja.
- **Tehnologija veriženja blokov:** Druga tehnika predlaga uporabo tehnologije veriženja blokov za preverjanje vira medija. Veriga blokov je digitalna glavna knjiga, ki na varen, pregleden in proti ponarejanju odporen način beleži transakcije v omrežju računalnikov. Uporablja decentralizacijo in kriptografijo za zagotavljanje zaupanja brez potrebe po osrednjem organu. V tem scenariju bi bilo treba videoposnetke pred prikazom na platformah družbenih medijev preveriti prek blokovne verige. Z zagotavljanjem, da so odobreni samo videoposnetki iz zaupanja vrednih virov, bi se lahko zmanjšalo širjenje potencialno škodljivih deepfake medijev.
- **Digitalni podpisi:** Nekateri predlagajo digitalno podpisovanje vseh videoposnetkov in slik, posnetih s fotoaparati, vključno s fotoaparati pametnih telefonov, kot sredstvo za boj proti globokim ponaredkom. To bi pomenilo dodeljevanje edinstvenih digitalnih podpisov vsakemu mediju, kar bi omogočilo izsledite vsake fotografije ali videoposnetka do njegovega prvotnega lastnika. Čeprav lahko ta pristop pomaga pri sledenju izvora vsebine, obstajajo pomisleki glede morebitne zlorabe, na primer pri usmerjanju na disidente ali kršenju zasebnosti.

3.5 Reference

Aro, J. (2016). Vojna v kibernetskem prostoru: Propaganda in trolanje kot orodji vojskovanja. *European View*, *15*(1), 121-132. https://doi.org/10.1007/s12290-016-0395-5

Angleški slovar Cambridge: V tem primeru je treba upoštevati tudi *pomen in definicije.* (2023). https://dictionary.cambridge.org/dictionary/english

Daniels, J. (2018). Algoritmični vzpon "alt-pravice". *Contexts*, *17*(1), 60-65. https://doi.org/10.1177/1536504218766547

Egelhofer, J. L., & Lecheler, S. (2019). Lažne novice kot dvodimenzionalni pojav: okvir in raziskovalni načrt. *Annals of the International Communication Association*, *43*(2), 97-116. https://doi.org/10.1080/23808985.2019.1602782

Greene. (2019). Satira "Deplorable": (Alt-Right Memes, White Genocide Tweets, and Redpilling Normies): "Alt-Right Memes, White Genocide Tweets, and Redpilling Normies". *Studies in American Humor*, *5*(1), 31-69. https://doi.org/10.5325/studamerhumor.5.1.0031

Hardaker, C. (2013). "Uh. ... ne da bi bili pikolovski,,,,,brez tega... .v pretekliku je drag dragged in ne drug." *Journal of Language Aggression and Conflict*, *1*(1), 58-86. https://doi.org/10.1075/jlac.1.1.04har

IED. (2022, 23. avgust). *Kako delujejo algoritmi družbenih medijev*. Inštitut za razvoj podjetništva. https://ied.eu/blog/technology-blog/how-do-social-media-algorithms-work/

Johnson, D. in Johnson, A. (2023, 15. junij). Kaj so globoki ponaredki? Kako ponarejeni zvočni in video posnetki, ki jih poganja umetna inteligenca, izkrivljajo naše dojemanje resničnosti. *Business* Insider. https://www.businessinsider.com/guides/tech/what-is-deepfake?international=true&r=US&IR=T

Levy, G. in Razin, R. (2019). Echo Chambers and Their Effects on Economic and Political Outcomes (Odmevne komore in njihovi učinki na gospodarske in politične rezultate). *Annual Review of Economics*, *11*, 303-328. https://doi.org/10.1146/annurev-economics-080218-030343

Lewis, B., & Marwick, A. E. (2017). Medijska manipulacija in dezinformacije na spletu. *Data & Society Research Institute.* https://www.posiel.com/wp-content/uploads/2016/08/Media-Manipulation-and-Disinformation-Online-1.pdf

McNealy, J. (2015). Bralci se negativno odzovejo na razkritje identitete plakata. *Newspaper Research Journal*, *38*(3). https://doi.org/10.1177/0739532917722977

Munn, L. (2019). Alt-right pipeline: (1): Individualne poti do spletnega ekstremizma. *First* Monday. https://doi.org/10.5210/fm.v24i6.10108

Vzorec, I. (2020, 13. januar). Kaj so globoki ponaredki in kako jih lahko prepoznate? *The* Guardian. https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them

Schumpe, B. M., Bélanger, J. J., Moyano, M. in Nisa, C. F. (2020). The role of sensation seeking in political violence (Vloga iskanja občutkov pri političnem nasilju): (1): Razširitev teorije o iskanju pomembnosti (Significance Quest Theory). *Journal of Personality and Social Psychology*, *118*(4), 743-761. https://doi.org/10.1037/pspp0000223

Seth, S. (2023, 11. september). Deset največjih svetovnih podjetij na področju novičarskih medijev. Investopedia. https://www.investopedia.com/stock-analysis/021815/worlds-top-ten-news-companies-nws-gci-trco-nyt.aspx

Somers, M. (2020, 21. julij). Deepfakes, explained. *MIT Sloan.* https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained

Tandoc, E. C., Lim, Z. W., & Ling, R. (2017). Opredelitev "lažnih novic": A Typology of Scholarly Definitions: A Typology of Scholarly Definitions. *Digital Journalism*, *6*(2), 137-153. https://doi.org/10.1080/21670811.2017.1360143

Van Puffelen, M. (2021). Rechtsextremisme: Geweld met een rechtsextremistisch motie. In *DSP-groep*. DSP-groep. https://www.dsp-groep.nl/wp-content/uploads/18MP-Rechtsextremisme-DSP-2021.pdf

Van Wonderen, R. (2023). *Rechts-extremistische Radicalisering op Sociale Media Platformen*. Verwey-Jonker Instituut.

Van Wonderen, R. (2023). *Richtlijn / onderbouwing Radicalisering*. Verwey-Jonker Instituut.

Van Wonderen, R. in Peeters, M. (2021). *Werken aan weerbaarheid tegen desinformatie en eenzijdige meningsvorming. Evaluatie lesprogramma Under Pressure.* Utrecht: Verwey-Jonker Instituut. https://www.verwey-jonker.nl/wp-content/uploads/2022/04/120550_Werken-aan-weerbaarheid-tegen-desinformatie-eenzijdige-meningsvorming.pdf.

Wasike, B. (2022). Ko vplivnež reče skoči! Kako signaliziranje vplivneža vpliva na sodelovanje z dezinformacijami COVID-19. *Social Science & Medicine*, *315*, 115497. https://doi.org/10.1016/j.socscimed.2022.115497

Wolfowicz, M., Weisburd, D. in Hasisi, B. (2021). Proučevanje interaktivnih učinkov filtrirnega mehurčka in odmevne komore na radikalizacijo. *Journal of Experimental Criminology*, *19*(1), 119-141. https://doi.org/10.1007/s11292-021-09471-0