

Tehisaru läbimurre: desinfo avastamise juhend



Juhendi väljaandmist
on toetanud

Google

Rohkem meie kohta!



www.cri.lt



CIVIC RESILIENCE INITIATIVE



@CivicResilience



@CRI



@civicresilienceinitiative

CRI juhend on välja töötatud koostöös organisatsioonidega, mis tegelevad desinfo vastaste meetmete arendamisega: *Baltic Security Foundation* (Läti) ning Riigikaitse ja Julgeoleku Teadmiskeskus (kaitseen.ee) (Eesti).

Juhendi väljaandmist on toetanud Google.

Tehisaru läbimurre: desinfo avastamise juhend

Digimeedia levik on viimase kümne aasta jooksul jätkuvalt suurenenud. Iga päev puutume kokku üha suurema hulga infovoogudega, mis pärinevad väga erinevatest kommunikatsioonikanalitest: sotsiaalvõrgustikud, blogid, veebisaidid, traditsiooniline meedia või muud elektroonilised väljaanded.

Selline infovoogude mitmekesisus muudab lihtsaks valida allikas, mis peegeldab kõige paremini meie huve ja poliitilisi või sotsiaalseid vaateid. Kuna suhtlus ja sotsiaalvõrgustikes vee-detud aeg kasvab, valivad üha enam inimesi neid oma peamiseks teabeallikaks, sageli hindamata neis peituvaid ohte. Teabe kiire ja mugav levik sotsiaalvõrgustikes loob samas ideaalsed tingimused desinformatsiooni kiireks levikuks.

Tehisaru praegune populaarsus ja sellele tuginevad vahendid pakuvad rohkelt võimalusi vale- ja väärinfo levitamiseks. Asjade interneti abil toimivad video- ja helivõltsingud võivad tekitada ühiskonnale märkimisväärt kahju, levitades võltsitud uudiseid ja õõnestades seega kodanike usaldust meedia vastu. Kuid koos suure potentsiaaliga pahatahtlike provokatsioonide

loomiseks suureneb ka võrdne potentsiaal nende vastu võitlemiseks. Kui asjade internet on hästi omandatud, saab selle tehnoloogia leidlikkust kasutada ka desinfo vastu võitlemisel. Balti riigid, sh Eesti on pidevalt sihtmärgiks valeinfokampaaniates, mis püüavad aktiivselt õõnestada riigi usaldusväärset.

Sellised jõupingutused õhutavad ebausaldust kohalike omavalitsuste, meie partnerite NATOs ja Euroopa Liidus ning püüavad pidevalt demotiveerida kodanikke aktiivselt valitsemises osalema.

Kuigi spetsialistid teevad suurepäraselt tööd valeinfo ümberlükkamisel, on kahju sageli tehtud siis, kui desinformatsioon on juba levitatud. Et harida avalikkust lõputus võitluses desinfo vastu, asutati Leedus 2019. aastal kodanikualgatus CRI (*Civic Resilience Initiative*).

Organisatsioon on teinud koostööd desinfo- ja meediaekspertidega ning sündis idee välja töötada see praktiline juhend. See dokument on sisuliselt tööriistakomplekt, mis pakub lihtsat juhendit teabe kontrollimiseks vajalike meetodite kohta.

Loodame, et see väljaanne aitab teil kergesti arendada harjumust kontrollida infoallikate ning neis sisalduvate fotode ja videote autentsust.

CRI meeskond on seadnud enda eesmärgiks olla Baltikumis peamine katalüsaator, mis tugevdab ühiskonna digitaalset vastupanuvõimet.



CRI meeskond

Selle juhendi eesmärk on aidata kaasa digitaalse vastupanuvõime suurendamisele, teadlikkuse tõstmisele turvalisusest ja valvuse vajadusest inforuumis.

Juhend sisaldab palju kasulikke näpunäiteid, mis aitavad õpilastel, üliõpilastel ja teistel huvilistel suurendada oma digi- ja meedia-aitu ning vastupanuvõimet pettuste ja valeteabe suhtes.

Selles juhendis on esitatud rida põhisoo- vitusi, mis võimaldavad tuvastada tõesed teated valeuudistest. Konkreetsemalt sisaldab see juhend vahendeid, et:

- kontrollida, kas veebipõhine teave on tõene või võltsitud
- tuvastada tehisintellekti abil loodud sisu
- tuvastada trolle
- tuvastada võltsitud sotsiaalmeediakontosisid
- tuvastada võltsitud veebipilte
- tuvastada manipuleeritud videoid
- juhendada mida teha siis, kui avastad des-, väär- või kuriinfo
- hoiatada tüüpiliste sotsiaalmeedias esinevate lõksude eest.

CRI juhend on välja töötatud koostöös organisatsioonidega, mis tegelevad desinfo vastaste meetmete arendamisega: *Baltic Security Foundation* (Läti) ning Riigikaitse ja Julgeoleku Teadmiskeskus (kaitsen.ee) (Eesti).

Juhendi väljaandmist on toetanud Google.

Moondatud teavet on palju:

Desinfo:

teave, mis on vale ja teadlikult loodud selleks et sihipäraselt kahjustada isikut, mõnda sotsiaalset rühma, organisatsiooni või riiki.

Eksitav info (väärinfo):

teave, mis on vale, kuid ei ole loodud eesmärgiga tekitada kahju.

Vaenulik info (kuriinfo):

teave, mis on tõene ja põhineb tege-
likkusel ning mida kasutatakse isiku,
organisatsiooni või riigi maine kahjus-
tamiseks.

Kõik need teabetüübid on ohtlikud, sest nad levivad laialt ja kiiresti (nagu viirus), mis juhtub siis, kui mõtlemata tagajärgedele postitavad paljud inimesed või isegi organisatsioonid selliseid lugusid, kuna need tunduvad huvitavad ja sensatsioonilised.



Valeinfo tuvastamine:

kuidas kontrollida uudiseid või postitusi?

Juhul, kui mõni teema on skandaalne või kõlab uskumatult, peaksid võtma aega ja tegema lihtsa kontrolli, et uurida info õigsust ja usaldusväärsust.

Kuidas seda teha?

Siin on viis lihtsat sammu, kuidas saaksid infot kontrollida:

1. Hinda allikat

Uuri lähemalt ja sügavamalt veebilehte või sotsiaalmeediakontot. Mõttele, kes võib olla uudise levitamise taga ja mis oli loo eesmärk.

2. Loe pealkirjast kaugemale

Lugude pealkirjad võivad olla skandaalse sisuga, et meelitada klikke ja soodustada jagamist. Kui süvened loosse, võib selguda, et pealkirjas esitatud väited ei vasta tõele.

3. Tutvu autori taustaga

Kas nimetatud autor on tõesti olemas? Kas autor on usaldusväärne isik?

4. Kas allikad kinnitavad lugu?

Sageli puuduvad vale- ja libauudistes lingid, mille abil saaks fakte kontrollida. Kui loos on viited allikatele, siis klõpsa neid läbi. Võib selguda, et algset sõnumit on ilustatud või selle tähendust moonutatud.

5. Kontrolli kuupäeva

Vanade uudiste uuesti avaldamine ei tähenda, et need on ikka veel ajaja asjakohased.



Lisanipid:

Vali turvaline ja usaldusväärne teabeallikas

Eestikeelseid uudiseid on reeglina turvaline lugeda peamistes veebiportaalides nagu ERR, DELFI jms. Väiksemaid portaaale võib olla lihtsam mõjutada teatud sisu ostmiseks, kuna neil on vähem inim- ja rahalisi ressursse, neil on vähem ajakirjanikke, kes kontrollivad teabe usaldusväärsust, ning nad võivad olla haavatavamad küberrünnakute suhtes.

Filtreeri hoolikalt teavet, mis pärineb välismaistest portaalidest või suhtlusvõrgustikest:

- Allikatel peab olema piisav hulk jälgijaid. Vastasel juhul tuleb täiendavalt uurida, kes veel on sama teavet avaldanud.

- Kui kogud teavet X (endine Twitter), TikTok või YouTube'i platvormidel, pööra tähelepanu kommentaaridele, mis peab olema lubatud. Usaldusväärsetel sissekannetel on rohkem kui üks kommentaar, postituste vahel on kindlad vahed ja need ei ole sama või väga sarnase kirjutamisstiili mallidega loodud. Vastasel juhul võivad need kommentaarid olla trollide või robotite poolt kirjutatud.

- Allikatel ei tohiks olla linke venemeelsetele valitsusportaalidele (Sputnik, Pervõi Kanal, Rossiya 24 jt). Kuigi nende portaalide avaldatud uudistes võib olla mingi osa tõest, ei tasu infosõja kontekstis võtta riski uskuda Venemaa uudisallikaid, kui neid ei kinnita mõni suur usaldusväärne portaal.

- Sensatsiooniline teave peaks alati tuginema usaldusväärsetele allikatele. Kui näed veebis uudiseid, siis veendu, et uudiste allikas mainib ka oma allikaid. Kui ei ole, kontrolli Google abil:

- 1) Lisa jutumärgid kõige olulisematele märksõnadele postitusest või artiklist, mida loed, ja sisesta need Google otsingusse.

- 2) Klõpsa otsingukasti paremas servas oleval nupul „Tööriistad“ ja vali „näita viimaseid“.

- 3) Kontrolli allikad uuesti ülaltoodud kriteeriumide alusel.

- Küsi endalt, kas selles allikas tõesti puuduvad tõendid ja analüüs. Millisest vaatenurgast on uudis esitatud? Kas võib juhtuda, et teatatud sündmused on valed ja lugejate kujutlusvõimega ja emotsioonidega püütakse manipuleerida?

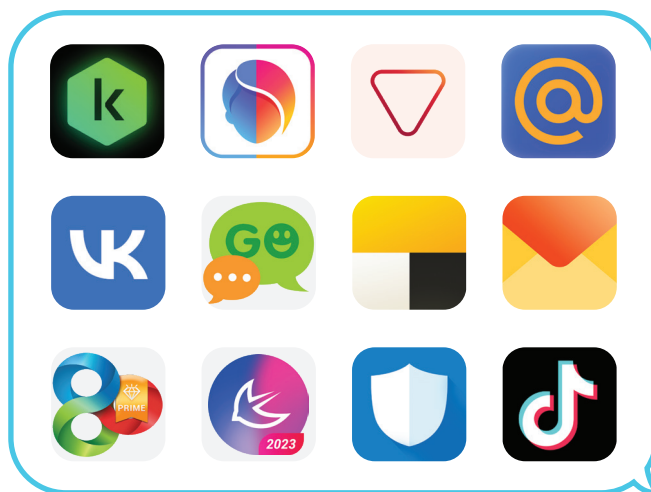
Lõpuks, kui näed mõne avaliku võimu asutuse veebisaidil sensatsioonilist uudist, külasta lisaks ka mõnda mainekat uudisportaaali. Üks infosõja vahenditest on valitsusasutuste veebilehtede häkkimine ning seal eksitava teabe avaldamine ja levitamine. Kui peaks häkkima mõnda asutust, kajastab meedia seda kindlasti. Siis ei tohiks nende veebilehel avaldatud teavet usaldada. Mõnel juhul ei ole häkkimine vajalik ja võidakse kasutada võltsitud veebilehe aadresse. Selliseid trikke on raskem märgata, seega kontrolli alati, et aadressirida näeks tõesti välja nii, nagu peaks.

Räägi usaldusväärsetest teabeallikatest mitte ainult oma vanematele, vaid ka vanavanematele.

Kas sinu pereliikmed vaatavad Venemaa televisiooni? Sõda on viimastel aastatel olnud üks populaarsemaid ja enim arutatud teemasid Venemaa telekanalites. Arusaadavalt võib vene keel olla paljude meie vanemate või eriti vanavanemate jaoks peamine ja eelistatud võõrkeel. Seepärast soovitatakse paigaldada Netflix või mõni muu voogedastusplatvorm, mis sisaldab vene keelde tõlgitud sisu. Kuna Venemaa televisioon on loodud selleks, et muuta vaataja üksikõikseks teabe suhtes, ei tohiks üleminek kodukino platvormile olla keeruline.

Eemalda Vene ja Hiina rakendused.

Eemaldage oma nutitelefoni kõik Venemaa või nende potentsiaalsete liitlaste, näiteks Hiina, toodetud rakendused. Mõnel neist on juurdepääs andmetele või asukohale, mida võidakse kasutada kasutaja vastu kas desinfo levitamiseks või isegi sõjalises tegevuses. Kõige populaarsemad rakendused on järgmised: Kaspersky Antivirus, CheckScan, FaceApp, MyPocket, Mail.ru, Vkontakte, Go SMS Pro, Yandex Taxi, Yandex Mail, Go Launcher, APUS Launcher, Security Master ja isegi TikTok.



Tehisintellekt



TI on seotud tehismudelitega, mis suudavad luua sisu, näiteks pilte, teksti, muusikat jne. Tehisaru näidete hulka kuuluvad sellised tuntud mudelid nagu ChatGPT, DALL-E jt, kus saab hõlpsasti sisestada küsimusi ja saada vastuseks genereeritud teksti või pilte.

Suured keelemudelid on üks tehisintellekti liik, mis suudab teksti mõista, ära tunda, panna konteksti ja genereerida uut sisu. Selleks, et suured keelemudelid saaksid oma ülesandeid täita, tuleb mudeleid koolitada ja see nõuab märkimisväärset hulka tekstiandmeid. Tavaliselt saadakse andmeid, lugedes tekste avalikust veebist ja muutes seda treenimise eesmärgil.

Keelemudeleid kasutatakse sageli selleks, et vastata küsimustele, mille vastust me ei tea, genereerida uut kirjalikku teksti, kui meil ei ole aega seda ette valmistada, või isegi kirjutada programmeerimiskoodi, kui

me ei oska ise programmeerida. See on suurepärase vahend teabe hankimiseks või lühikese sisu loomiseks, mille koostamine nõuaks muidu palju aega ja vaeva.

Kuid keelemudeleid ei kasutata mitte ainult positiivsetel eesmärkidel, vaid see võimaldab ka küberkurjategijatel palju tõhusamalt küberrünnakuid läbi viia.

1. Sotsiaalne insenerlus (manipuleerimine)

Keelemudelite abil saab luua suurepärase teksti mis tahes keeles ja see tekst võib olla täpsem kui "Google Translate" programmi abil tehtud tõlked. Tekst saab olema kvaliteetne ja väheste keelevigadega või üldse ilma. Seetõttu ei pruugi küberkurjategija osata oma valitud ohvri emakeelt, kuid võib siiski kirjutada teksti, millega püüda seda ohvrit petta, et ta saadaks raha või vajutaks mõne pahavara lingile, et saada tema sis-

selogimisandmeid. Tänapäeval on pahatahtlikud e-kirjad, sõnumid jne kvaliteetsemad kui varem ja neid on raskem tuvastada pettustena. On olemas spetsiaalsed keelemudelid, mis on loodud selliste sõnumite genereerimiseks.

2. Desinfo levitamine

Suurte keelemudelite poolt genereeritud vastus ei ole alati täpne ega pruugi kajastada tegelikkust, mis viib kasutajaid eksitusse ja saadud ebatäpne või vale

vastus võib hakata kasutaja suhtlusringkondades levima. Lisaks õpivad keelemudelid kasutajate küsimuste ja kasutajate esitatud selgituste põhjal, mistõttu võib mudel õppida selgeks mõned infotükid, mis on ebaõiged, ja esitada need edasi teistele tõestena, aidates sellega kaasa valeinfo levikule.

Tehisintellekt

Kuidas tuvastada suurte keelemudelite poolt genereeritud teksti?

Praegu ei ole olemas ühtegi 100% kindlat vahendit, mis suudaks lõplikult kindlaks teha, kas tekst on kirjutatud tehisintellekti poolt või mitte. Kõige täpsem

äratundmisvahend on praegu inimese aju. Näiteks alljärgnev tekst on kirjutatud ChatGPT abil. Kas sa märkad selles vigu?

Lugupeetud kasutaja!

Ma loodan, et see e-kiri leiab teid hea tervise juures. Kirjutan teile, et juhtida teie tähelepanu sellele, kui võrd oluline on registreerimise / sisselogimise protsessi lõpuleviimine meie platvormil.

Osana meie pühendumusest pakkuda teile sujuvat kasutuskogemust, palume teil oma registreerimise lõpule viia või oma kontosse sisse logida esimesel võimalusel. See mitte ainult ei taga teie konto turvalisust, vaid annab teile ka juurdepääsu kõikidele meie platvormi funktsioonidele ja eelistele.

Registreerimise / sisselogimise lõpuleviimiseks vajutage palun allpool toodud lingile: [Sisselogimise link](#)

Kui teil tekivad raskused või on küsimusi, võtke julgelt ühendust meie klienditoega.

Täname teid koostöö eest ja loodame teid teenindada edaspidigi.

Selles lühitekstis on mõned kohad, mis tunduvad kahtlased:

- Teksti algus ei kõla loomulikult ja tänapäeval on üsna haruldane saada sellise algusega e-kirju. Tavaliselt viitab selline algus sellele, et tekst on loodud tehisaru poolt.
- Samuti on esimene lõik kirjutatud esimese isiku („mina“) ainsuses, kuid ülejäänud tekst on kirjutatud „meie“ (mitmus) vaatenurgast, mis samuti tundub kahtlane.

Seega näitab see näide, et tehisintellekt kasutab teksti loomisel fraase, mis on automaatselt genereeritud tekstis tavalisemad. Muud vihjed võivad olla seotud valede käändelõppude kasutamisega või vigase lauseehitusega. Lisaks võib mõnikord keele struktuur ja stiil paljastada tehisintellekti töö, isegi kui tekstis ei esine kirjavigu.

Veebis on mitmeid lehti, mis püüavad kindlaks teha, kas tekst on loodud tehisintellekti poolt või mitte. Selleks, et mõista, kuidas automatiseeritud tööriistad üritavad tuvastada tehisaru kirjutatud teksti, peame

pöörduma tagasi selle juurde, kuidas tehisintellekt teksti genereerib.

Lihtsustatult öeldes püüavad teksti genereerivad keelemudelid ennustada lauses järgmist sobivaimat sõna, ja see muudab mudelid prognoositavaks. Mida rohkem andmeid kasutati keelemudeli koolitamiseks, seda paremini oskab tehisaru järgmist sõna pakkuda ja ära arvata. Mõned automaatsed tuvastusvahendid püüavad analüüsida, kas sõnade järjekord lauses on statistiliselt optimaalne, mis näitab, et teksti on kirjutanud tehisaru. Samal ajal tegelevad teised tööriistad vastupidise protsessiga, arvutades inimloodud kirjastruktuuride analüüsi abil tõenäosust, et teksti on kirjutanud inimene.

Allpool on toodud mõned näited automatiseeritud tööriistadest. Seal on väli, kuhu saab kleepida kahtlase teksti. Pärast analüüsinupule klõpsamist annab tööriist tavaliselt skoori, kui tõenäoliselt on see tekst loodud tehisintellekti poolt. Mõned tööriistad toovad tekstis esile kohad, mis on tõenäoliselt pigem loodud tehisaru poolt kui inimese kirjutatud.

Copyleaks



Examples:

GPT4 ChatGPT Bard Human AI + Human

Model: Basic

"We've been navigating the vast seas of the web, and now we're inviting you to dive in with us! We've heard whispers about an application with exclusive features meant for internal use. And, just like our website that exports goods to cater to your internal market, sometimes the best treasures are hidden just beneath the surface - waiting to be discovered!"

Clear

AI Content Detected



Zerogpt



Your Text is AI/GPT Generated



Dear [Username],

Welcome to [Your Website]! We're thrilled to have you as part of our community.

To complete your registration and unlock the full benefits of your account, please click on the following link:

[Insert Confirmation Link]

By confirming your registration, you'll gain access to exclusive features and updates. If you have any questions or need assistance, feel free to reach out to our support team at [Support Email].

Thank you for choosing [Your Website]! We look forward to providing you with a great experience.

Best regards,
The [Your Website] Team

Highlighted text is suspected to be most likely generated by AI*
572 Characters
91 Words

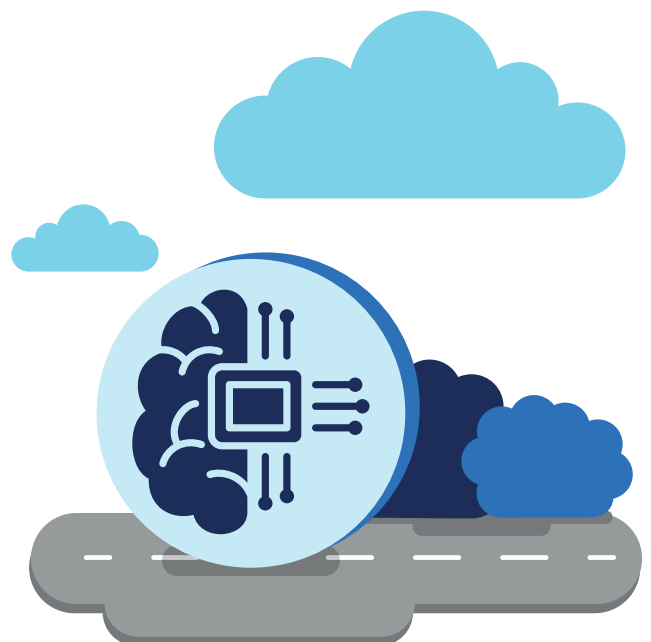
Writer



Gptzero



llm



Teksti kontrollimine

Kuidas kontrollida, kas tekstis sisalduvad andmed on tõesed?

Google, Google, Google!

Lihtsustatult öeldes on otsingumootor Google sinu parim sõber, kui tegemist on desinfo paljastamisega.

Alati saab kontrollida, kas info on tõene või vale. Selleks tuleb kasutada märksõnu, mida saab võtta sind huvitavast uudisest või infotükist, ja vaadata, kas mõni usaldusväärne allikas on sellest juba kirjutanud.

Kui näed midagi kahtlast või emotsionaalselt resoneerivat, siis on suur tõenäosus, et keegi juba räägib sellest.

Kuidas kontrollida?

1. Vali märksõnad, mis kirjeldavad kõige paremini infot, mida otsid;
2. Kasuta mõnda head otsingumootorit (nt Google), et otsida seda infot;
3. Vali allikate filtris „Uudised“ ja viimase tunni või päeva jooksul avaldatud kõige värskemad postitused;
4. Vali usaldusväärsed allikad ja kontrolli teavet.

Peamised reeglid, mida meeles pidada:

- Guugeldamine on esimene samm teabe kontrollimiseks. See on äärmiselt kiire ja tõhus;
- Ükskõik, millist teavet soovid kontrollida; Google'i otsing võib toimida väga hästi nii teksti, foto või video suhtes;
- Oluline on kasutada märksõnu, et leida sama infot usaldusväärse(te)st allika(te)st;

Selleks, et olla guugeldamisel tõhusam, kasuta Google'i põhilisi funktsioone:

1) kui otsid mõnda täpset fraasi, pane see jutumärkide vahele („...“);

2) kui sinu otsingutulemustes on mõni väga populaarne märksõna, kasuta miinust (-) ja seda märksõna, et seda otsingutulemustest välja jätta;

3) kui sa ei tea sõna täpset kirjalpilti või täpset arvu või kuupäeva, võid kasutada * asendusmärgina, et asendada sõnas või numbris või kuupäevas puuduv märk;

• kui valdad mõnda võõrkeelt, saad ka seda kasutada. Otsi, mida teised allikad teistes keeltes selle uudise kohta teatavad.

• kui valdad vaid emakeelt, siis palu kellelgi, keda usaldad, et ta aitaks sul kontrollida seda infot teis(t) es keel(t)es. Aruta nendega, kuidas seda uudist teistes keeleruumides kajastatakse. See võib olla nende jaoks sama väärtuslik kui sulle.



Kasulikud vahendid:

Google



Bing



Yandex



(TÄHELEPANU: Ole ettevaatlik, see on Venemaa veebileht, seega rakenda oma arvutis täiendavaid turvameetmeid. Teadaolevalt on selle veebilehe kasutamine ohutu, kuid soovitatakse mitte kasutada mobiilirakendust - see nõuab paigaldamisel juurdepääsu isikuandmetele).



Pildi kontrollimine

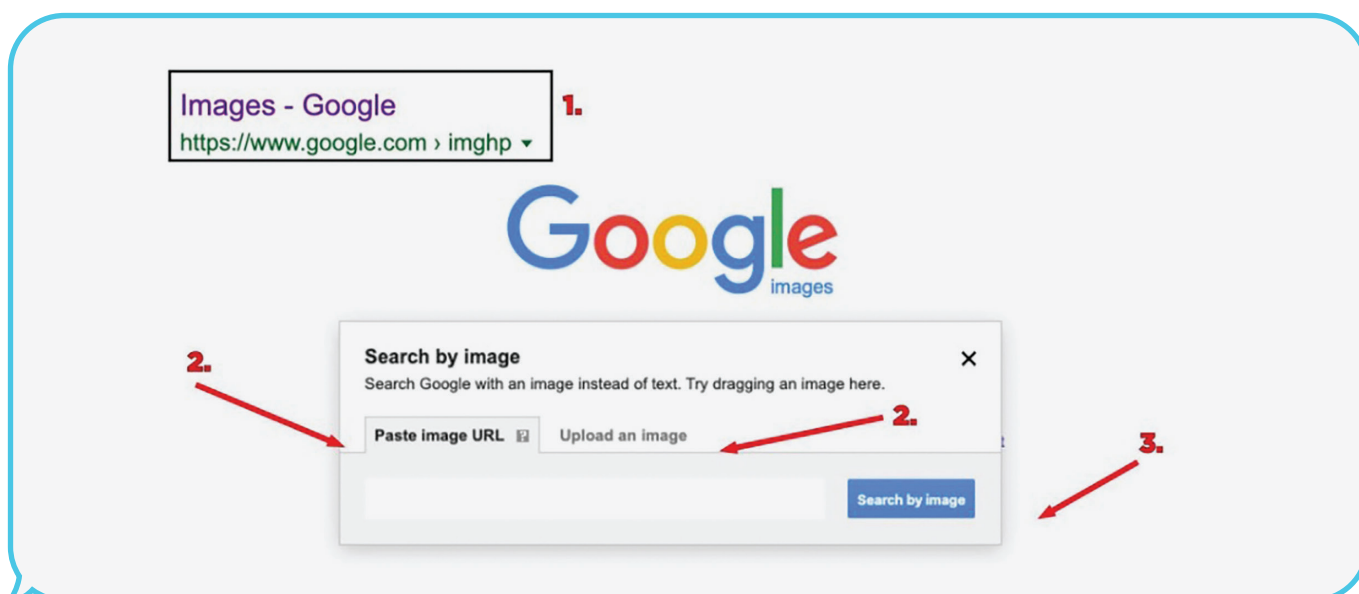
Kuidas kontrollida, kas visuaalne teave on tõene?

Pildiotsing:

Kui tegemist on võltsitud piltidega, on parim kontrollimeetod pöördkujutise (tagurpidi) otsing. Metoodika mõte on kasutada otsingumootorit, kuid märksõnade asemel kasutada pilti. See võimaldab leida kõik identsed või väga sarnased pildid, mis on varem veebis avaldatud. Kuna piltide taaskasutamine (pildi postitamine varasemast ajast ja väitmine, et see on hiljuti tehtud) on endiselt üks peamisi desinfovõtteid, on selle kontrollimine üks parimaid taktikaid selle vastu. Kui leitakse, et mõni pilt oli varem postitatud, on see usaldusväärne viis kinnitada, et tegemist võib olla desinfiga. Muudel juhtudel, kui mõnda pilti on muudetud, aitab pöördotsing leida originaalpildi.

Kuidas kontrollida?

1. Ava üks otsingumootoritest (link kasulikele tööriistadele allpool);
2. Kopeeri link kujutisele või laadi alla pilt ise;
3. Uuri, kas samu või väga sarnaseid pilte on varem veebis avaldatud.



Vea taseme analüüs

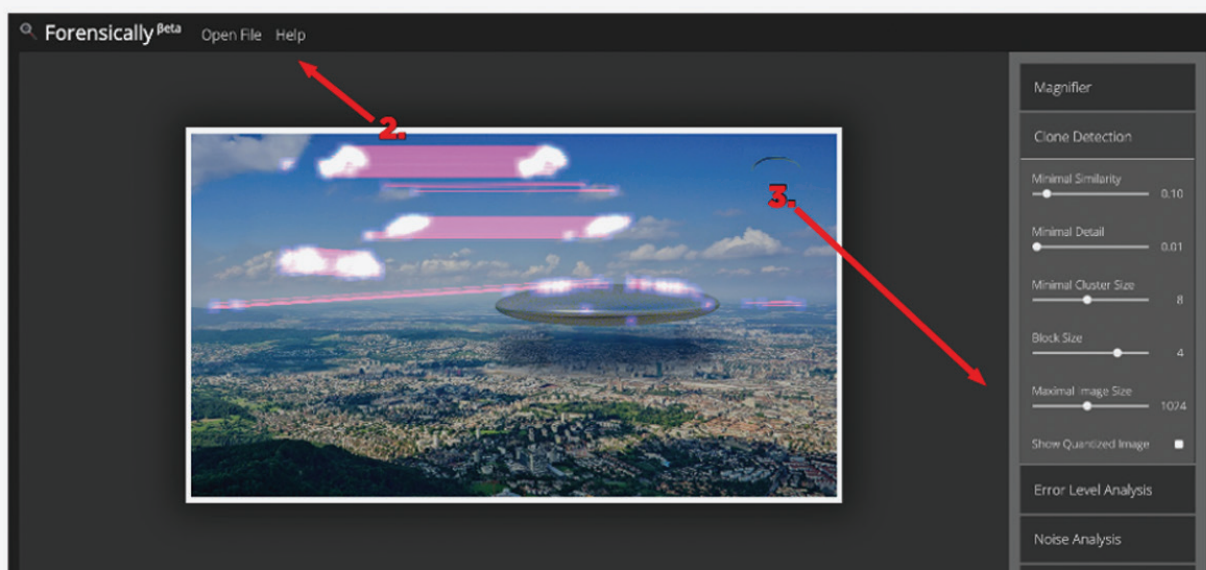
Veetasemete analüüs on täiustatud meetod, mis võimaldab tuvastada pildil eri tihendustasealasid. JPEG-piltide puhul peaks kogu pilt olema ligikaudu sama tihendustasemega. Kui mingi osa pildist on oluliselt erineva veetasemega, siis viitab see tõenäoliselt digitaalsele moondamisele. Praktikas võib vaadata kogu pilti ja tuvastada erinevad kõrge või madala kontrastsusega servad, pinnad ja tekstuurid ning seejärel võrrelda neid alasid analüüsi tulemustega. Kui erinevused on märkimisväärsed, siis tuvastab see analüüs kahtlased piirkonnad, mida on digitaalselt muudetud.

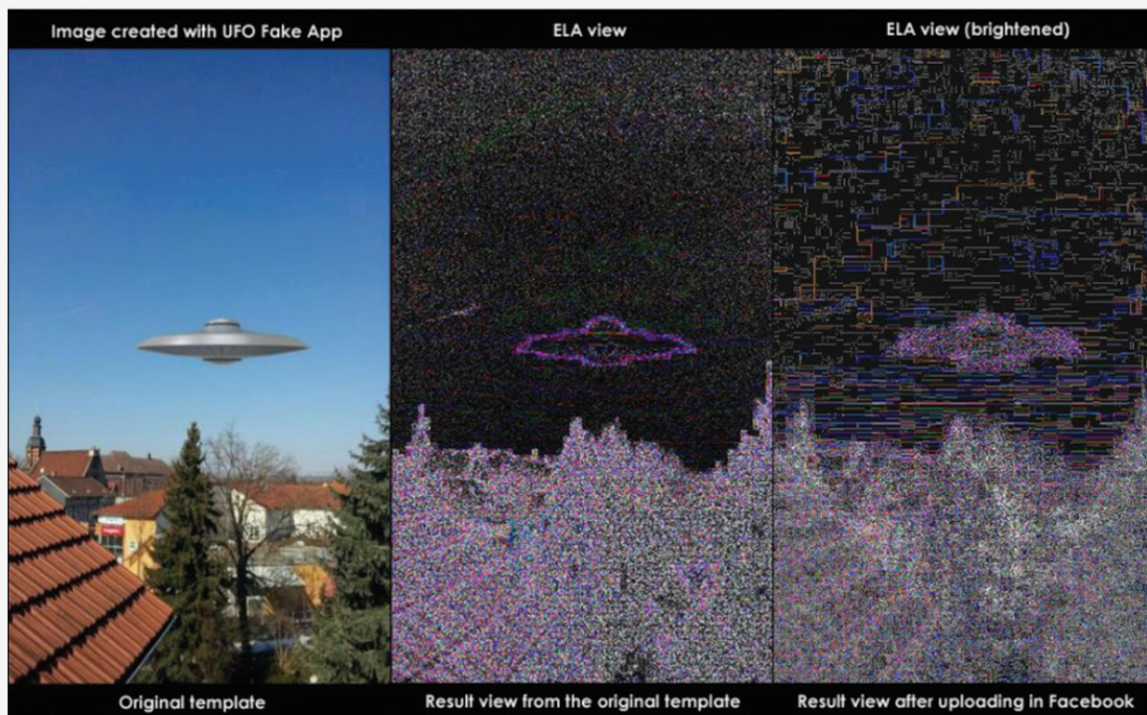
Oluline on lisada, et see meetod ei ole kuulikindel, kuid sellegipoolest on see usaldusväärne esimene samm pildi digimuudatuste tuvastamiseks. Fotoeksperdiis on eraldi haru ja hästi moondatud või võltsitud piltide paljastamine nõuab pikaajalist kogemust, kuid igapäevastes propagandasõnumites ei ole need pildid tavaliselt nii hästi kujundatud ja seega on need kergesti tuvastatavad.

Peamised reeglid, mida meeles pidada:

- Pöördkujutise otsing peaks olema tavapärane viis enne pildisisu usaldamist. Kui pilt ei ole see, mida ta väidab, võib see viidata kogu uudise või sõnumi ebatõesusele;
- Pöördpildiotsingut saab kasutada tundmatute inimeste tuvastamiseks pildil;
- ELA ei ole kuulikindel meetod, kuid see on suurepärane kiire viis kontrollida, kas pilti on muudetud või mitte.

Forensically, free online photo forensics tools - 29a.ch
<https://29a.ch> › photo-forensics ▾ 1.





Kasulikud vahendid:

Google / Pildiotsing



Yandex / Pildiotsing



Google / Pikendamine „RevEye“



Forensically



Foto Forensics



Kuidas tuvastada tehisintellekti poolt genereeritud pilti?

Erinevalt genereeritud tekstist on genereeritud pildi tuvastamine märksa lihtsam, eriti kui püütakse luua realistlikke fotosid. Allpool on toodud üks näide. Kahel fotol on kujutatud inimesi Vilniuse vanalinnas. Esmapilgul tundub, et fotod on ehtsad, kuid vaatame lähemalt.



- Värvid - vasakpoolsel fotol tunduvad värvid ebaloomulikud ja liiga eredad, mõjuvad väsitavalt silmadele.

- Taust - sageli on genereeritud fotodel ebaselge taust või on hoonete / autode jooned taustal kuidagi moonutatud.



- Anomaaliad ja moonutused - fotosid suumides märkad kohe, et teatud kehaosad inimestel tunduvad

ebaloomulikud ja moonutatud: kõrvad näevad eripärased välja, pead tunduvad nagu pintsliiga maalitud jne. Samuti võib tehisaru sageli lisada rohkem kui 5 sõrme või rohkem hambaid kui inimesel oleks. Oluline on pöörata tähelepanu detailidele.


- Vesimärgid - tööriistad, eriti tasuta tööriistad, lisavad fotodele sageli oma vesimärke. Ülaltoodud näidetel võid märgata värvilisi ruudikesi paremas alumises nurgas, mis näitavad, et pildid on kunstlikult loodud DALL-E abil. Loomulikult saab selle vesimärgi käsitsi eemaldada. Kuid platvormid, mis lasevad pilte genereerida, hakkasid lisama ka nähtamatuid vesimärke, mis ei ole palja silmaga tuvastavad. Kui aga foto laaditakse üles kontrollimiseks, siis saab kohe tuvastada, et tegemist on genereeritud pildiga, sest tarkvara tuvastab ka varjatud vesimärgi.



Mitmed automaatsed veebitööriistad aitavad tuvastada genereeritud fotosid. Nende tööriistade kasutamise põhimõte on standardne. Lihtsalt lohista foto või laadi see üles konkreetseesse ossa ja klõpsa nuppu „analüüs“. Mõne sekundi pärast saad tulemuse:

Hivemoderation  

Upload images here to test our model in real-time!
Supports png, jpeg, jpg, webp. Use is subject to this site's [Terms of Service](#)



Upload

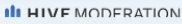
RESULT

The input is: **likely to be AI Generated**

99.9%

BY CLASSES

| Classes | Score |
|---|-------|
| ■ ai_generated | 0.99 |
| ■ dalle | 0.99 |
| ■ not_ai_generated | 0.00 |
| ■ none | 0.00 |
| ■ midjourney | 0.00 |
| ■ stablediffusion | 0.00 |

 HIVE MODERATION

Aiornot  

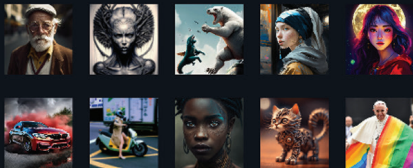
Try AI or Not


IMAGES

AUDIO

AI or Not

Determine whether an image has been generated by artificial intelligence or a human





Drag and drop
or **upload** your image

We support jpeg, png, webp, gif, tiff, bmp.
10Mb of maximum size.
User usage

OR

s

AI OR NOT?

Video kontrollimine

Kuidas kontrollida, kas video on ehtne või mitte?

Pildiotsing

Kui tegemist on moondatud või võltsitud videotega, siis sarnaselt piltidega on parim meetod tagurpidiotsing. Kuna videod on tegelikult pildiseeriad, on seega ühe kaadri väljavõtmine ja selle analüüsimine suurepärase võimaluse. Mõlemad tööriistad InVid ja Amnesty Data Viewer võimaldavad leida sarnaseid või identseid videoid, mis on juba veebis varem avaldatud, otsides nii kaadreid kui ka pildiikoone.

Kuidas seda teha?

1. Ava mõni otsingumootor (Amnesty DataViewer või InVid);
2. Sisesta video link;
3. Kontrolli, kas video ilmub varem avaldatute hulgas.



Youtube DataViewer

XI EPICdR + COLPIN / Corrupción judicial / Elber Gutiérrez

Video ID: Neo2Rp87Ifs
Upload Date (YYYY/MM/DD): 2018-11-13
Upload Time (UTC): 18:28:16 (convert to local time)

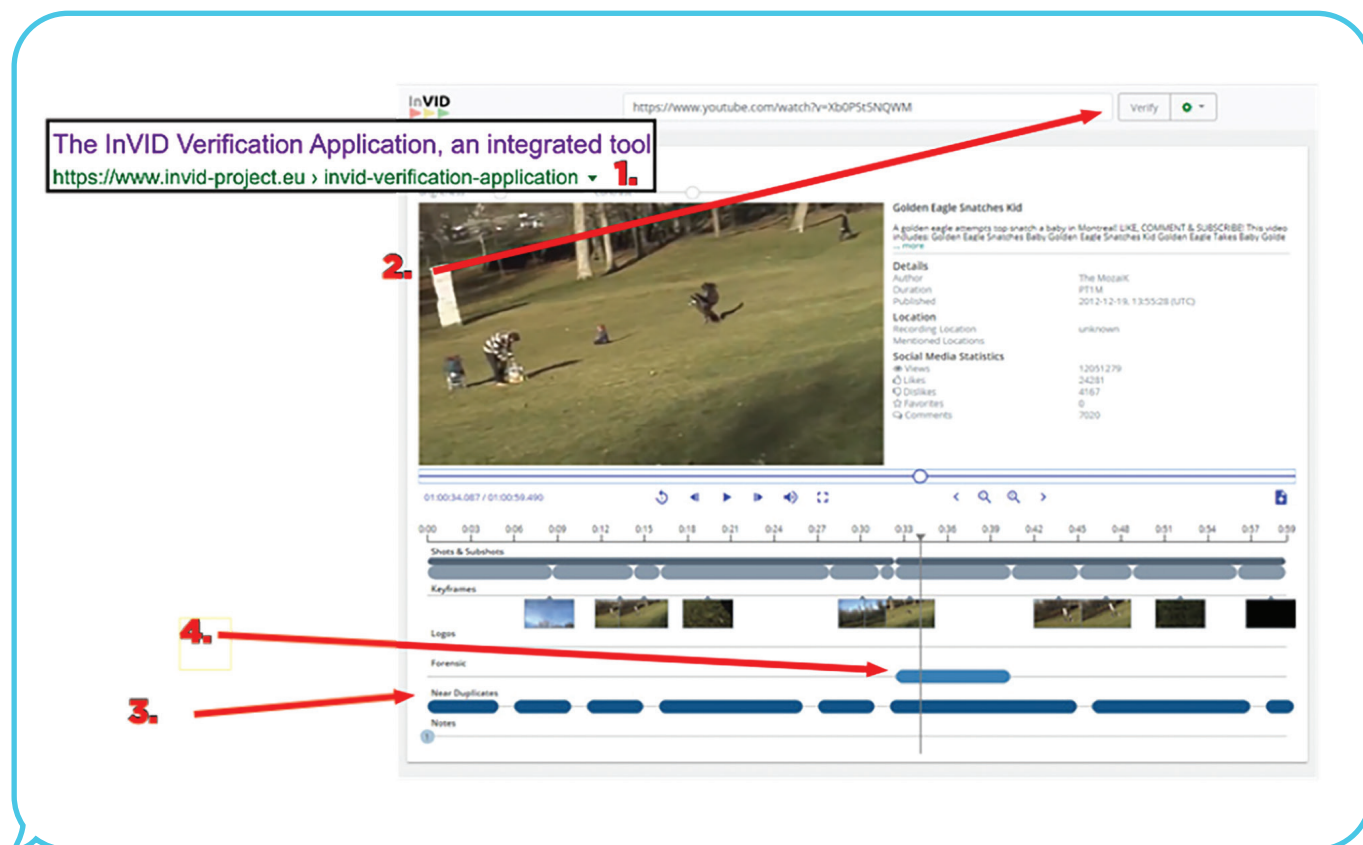
Thumbnails:



[reverse image search](#)

Sügavam analüüs

Mõni tarkvara tuvastab potentsiaalselt muudetud kaadrid – selleks saab kasutada sügavamat analüüsi „Forensic“. Kui InVid tuvastab, et mõni kaader on potentsiaalselt muudetud, on suur tõenäosus, et ka kogu video on võltsitud.



Peamised reeglid, mida tuleb meeles pidada:

- Peamine probleem videote puhul on sama, mis piltide puhul - video taaskasutamine. Videod, mis on võetud enne, postitatakse uuesti ja esitatakse vale- või libauudistes;

- Mainitud vahendid ei pruugi olla samatõhusad kui pildi kontrollimise vahendid, kuid nad suudavad siiski paljusid võltsitud videoid tuvastada.

Kasulikud vahendid:

InVid



Amnesty International
YouTube Viewer



Trollid

Kuidas märgata veebitrolli?

Kes on veebitroll?

Veebitroll on isik, kes tahtlikult algatab konflikti või solvab teisi veebikasutajaid, et segada arutelu ja külvata lahkarvamusi, postitades mõnel lehel või sotsiaalvõrgustikus põletavaid või teemaväliseid kommentaare. Nende eesmärk on meelega provotseerida teisi emotsionaalsele reaktsioonile ja viia arutelu kõrvale. Veebitroll erineb botist, sest troll on inimene, samas kui robotid on automatiseeritud. Need kaks kontotüüpi välistavad teineteist.

Veebitrolli on raskem märgata kui botti, sest need kontod on tavaliselt keerulisemad ja teesklevad aktiivselt, et nad on tavalised inimesed. Allpool on toodud mõned kriteeriumid, mis aitavad veebitrolli tuvastada, kuid need soovitused on pigem alustuseks ja nimekiri ei ole lõplik. Harva on võimalik 100-pro-

sendilise kindlusega öelda, et konkreetne konto kuulub veebitrollile, mitte lihtsalt ei toeta teatud pahahtlikke või vaenulikke narratiive. Enne omapäraste iseloomujoonte uurimist, mis usaldusväärselt osutavad nt Kremli-meelsele trollile, on oluline vaadata liiksaks ühte teist tegurit, mis seda ei tee – poliitiline sisu. Erinevad reaalsed sotsiaalmeediakasutajad kipuvad olema väga parteipoliitilised, eriti kui arutletakse poliitiliste teemade üle.

Järgnevalt on toodud mõned kriteeriumid, mis aitavad trolle tuvastada, kuid need vihjed on pigem soovituslikud ja mitte täiuslikud. Harva on võimalik olla 100% kindel, et konto kuulub trollile ja mitte lihtsalt masendunud või kurjale kasutajale.

1. Vead ingliskeelsete artiklitega: A vs. The

Üks keelelistest tunnustest, mis on iseloomulik paljudele Venemaa kontrollitud veebitrollide kontodele, on võimetus kasutada õigesti grammatilisi artikleid - „A“ ja „The“. Vene keeles ei ole kumbagi.

2. Vead küsimuse sõnastamisel

Teine levinud keeleline märk on suutmatuse küsimust korrektselt sõnastada. Erinevalt inglise, saksa ja prantsuse keelest, ei muutu venekeelses küsimuses sõnade järjekord. Paljud Venemaa veebitrollide kontod postitavad küsimusi, mis säilitavad vene keelele iseloomuliku stilistika.

Pea siiski meeles, et nüüdisaegsed tehisintellekti vahendid tõlgivad tekste üha täpsemalt, seega ei piisa veebitrolli tuvastamiseks ainult keelenüansside märkamisest.

3. Ebaselge või küsitav identiteet

Mõned veebitrollid kasutavad võltsnimesid, mis on mõnes keeles väga levinud, mistõttu on raske eristada konkreetset autorit, või ajatakse tahtlikult segi mõne teise isikuga, näiteks tunnustatud ajakirjanikuga. Veebitrollide poolt kasutatakse nimed on samuti mõeldud selleks, et neid tajutaks traditsiooniliste või „õigesti kõlavate“ nimedena, nii et lugeja kalduks sellist artikli või sotsiaalmeedia kommentaari autorit usaldama ja mitte kahtluse alla seadma. Samuti võib olla kasulik kontrollida profiilipilte, kui need on olemas. Sellised täiendava usalduse saavutamiseks lisatud pildid võivad olla tüüpifotod, mida võib kergesti leida veebist. Need pildid võivad olla ka tahtlikult ebaselged, kui neid lähemalt vaadata (nt fototöödeldud alad, isik kannab päikesepille jne), siis ei ole võimalik isikut selgelt tuvastada.

4. Kremli-meelsete narratiivide võimendamine

Venemaa valitsus on välja töötanud oma strateegilised narratiivid viimaste aastate peamiste geopoliitiliste sündmuste kohta. Nende loogika järgib Vene Föderatsiooni infojulgeoleku doktriinis (2000. a) kehtestatud põhimõtteid, mis käsitlevad riigi poliitika ja ametliku seisukoha edastamist Venemaa valitsusele olulistes küsimustes. Kuna Kremli-meelsed narratiivid on laialdaselt kättesaadavad veebis, näiteks Venemaa välisministeeriumi või RT Twitteri (X) kontol, on seega lihtne kontrollida, kas samad teemad esinevad ka kahtlustataval kontol. Kontot, mis jagab korduvalt Venemaa valitsuse jutupunkte mõne sündmuse kohta, võib õigustatult pidada Kremli-meelseks.

Kui mõni konto jagab sotsiaalmeedias enamikku või kõiki Kremli narratiive, teeb iseloomulikke keelelisi vigu ja esineb USA või Briti kasutajana, võib tegemist olla Venemaa poolt opereeritava veebitrolliga.



Muud kasulikud vihjed

Veebitrollidel on väljamõeldud, st mitte-toimivad e-posti aadressid:

Kuna veebis kommenteerimiseks nõutakse sageli e-posti aadressi, siis kasutavad veebitrollid selle nõude vältimiseks väljamõeldud aadresse. Enamasti on need juhuslikud ja neid on lihtne märgata, sest need ei sisalda tegelikku nime.

Veebitrollid on selleks, et inimesi ärritada:

Nad ei ole viisakad ega häbene veebis sõimu alustada. Nad ropendavad ja esitavad süüdistusi ning kõlavad tihti vihaselt.

Veebitrollid kasutavad anonüümseid proksisid:

Veebitrollid kasutavad sageli anonüümsust tagavaid proksiservereid, mis näitavad teistsugust internetiprotokolli (IP) aadressi.

Veebitrollid lisavad vestlusele harva midagi väärtuslikku:

Kui veebitrollid reageerivad mõnele arutelule, ei lisa nad midagi sisulist. Selle asemel naljatavad, sõimavad ja solvavad.

Libakontod Facebookis

Kuidas tuvastada, kas Facebooki konto on võltsitud või mitte?

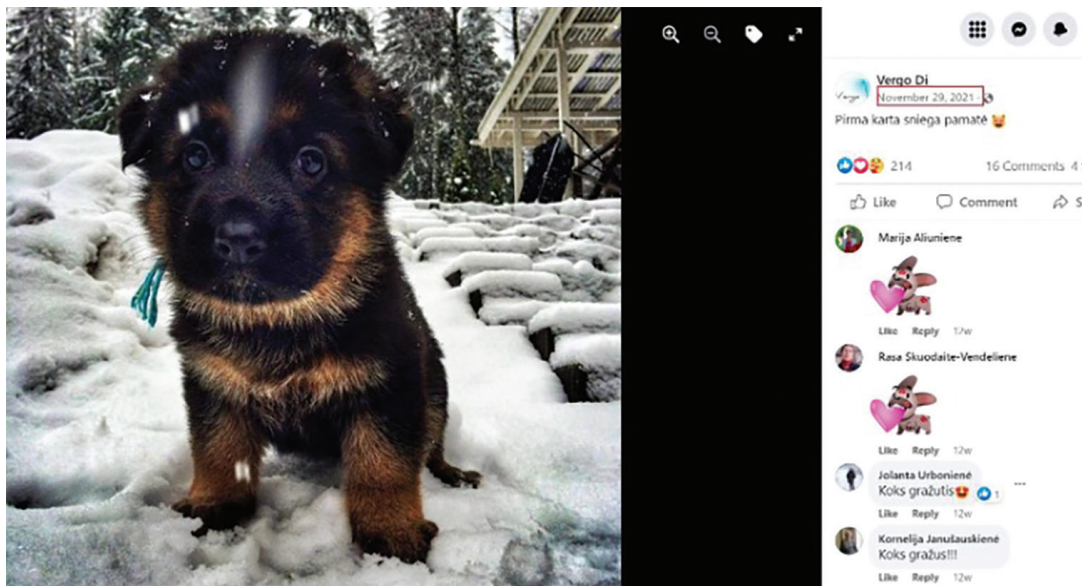
Tavaliselt ei ole libakontod nii aktiivsed kui veebitrollid ja kipuvad olema pigem vaiksed pealtvaatajad. Samasugused kriteeriumid kehtivad libakontode kohta enamikul platvormidel ja sotsiaalmeediakanalites, kuid selles juhendis on näiteks valitud just Facebook. Selle võrgustiku kasutajad kipuvad oma kontodel jagama kõige rohkem isiklikku teavet. Anonüümsed libakontod

püüavad aktiivselt saada Facebooki sõpradeks kahel peamisel põhjusel: et tunda teistele realsemana (kuna neil on sõprade listis hulk päris inimesi), ja et näha rohkem isiklikku infot. Sõltuvalt võltskonto eesmärkidest võib seda kasutada nt mõne organisatsiooni töötajate isikuandmete kogumiseks.



1. Atraktiivsuse tegur

Mida sümpaatsem on pilt, seda suurem on tõenäosus, et nad võetakse sõpradeks vastu ja siis saavad nad isikuandmeid näha.



Saab hõlpsasti kontrollida, kas see on või ei ole internetist kopeeritud pilt, mille on valinud võltsitud konto omanik, et saada rohkem tähelepanu.



1 result

Searched over 52.5 billion images in 0.5 seconds for:
lh3.googleusercontent.com/J3zMaRV__qxt5dSuOpufIhxiMfP39e...

Sort by best match

Filter by website / collection



instagom.com

[explore/tags/germanshepherdpuppies](https://instagom.com/explore/tags/germanshepherdpuppies) - First found on May 28, 2017

Filename: [18722090_1019473188187238_8466430548250722304_n.jpg](#)
(960 x 960, 151.9 kB)

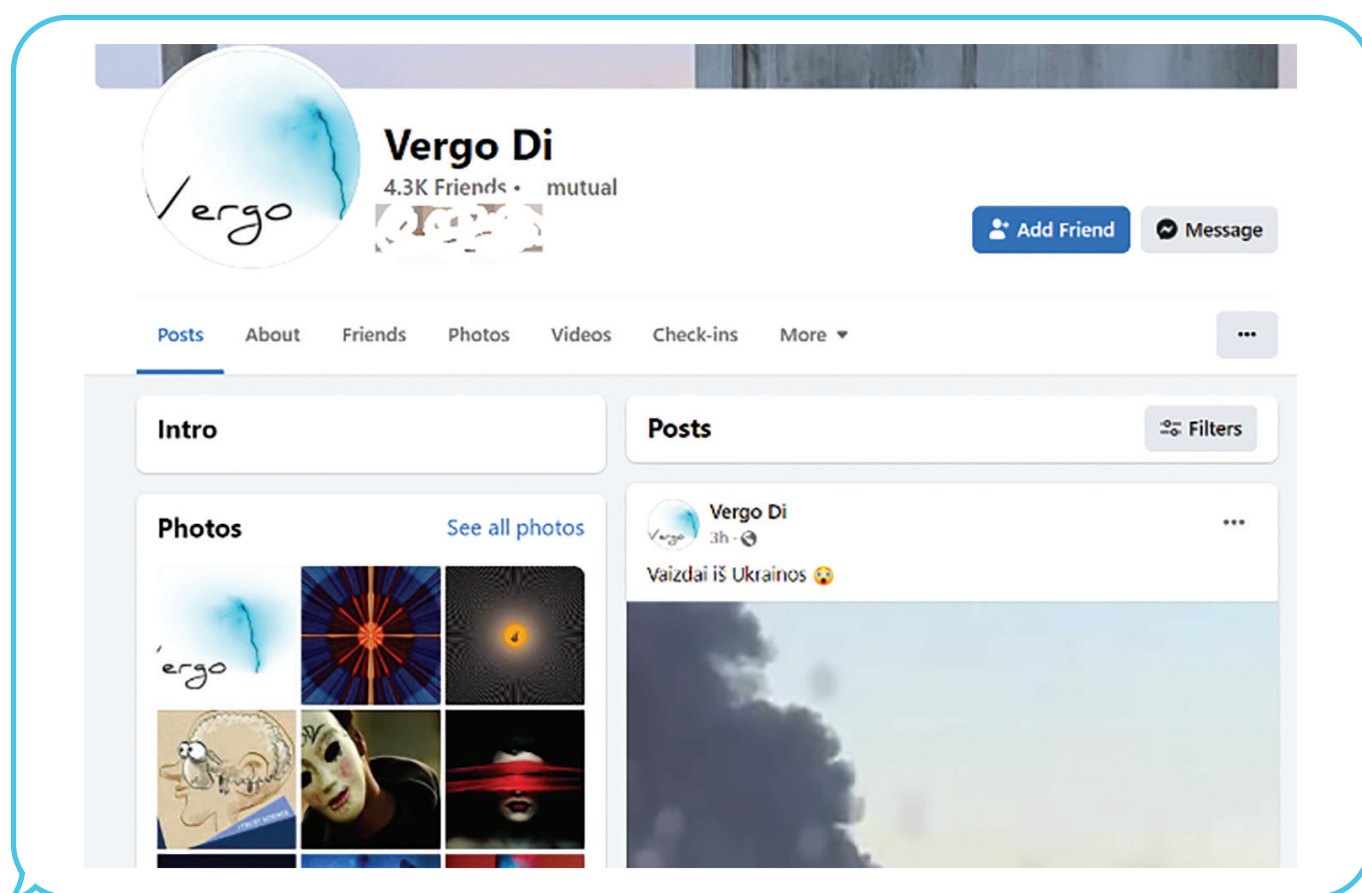
Veebilehe <https://tineye.com/> pilditsing näitab kuupäeva, mil foto esimest korda jagati. Võib järeldada, et tegemist on libakonto ehk publikut köitva profiiliga, mis jagab võltsitud postitusi ja aeg-ajalt propagandasõnumeid.

2. Vähesed fotod

Enamik libakontosid ei postita palju fotosid – kolm-neli on tüüpiline ja vahel on needki erinevate inimeste pildid. Just piisavalt, et luua ajutine illusioon, et selle konto taga on päris inimene.

3. Kummalised elulookirjeldused

Enamiku võltskontode biograafias on väga vähe teavet või tundub esitatud teave kummaline. Näiteks ei ole võimatu, kuid siiski väga ebatõenäoline, et Bronxist pärit noor inimene lõpetas Helsingi ülikooli ja töötab nüüd New Yorgi PR-firmas. Nime kiirkontroll Googles koos profiilipildi tagurpidi pildiotsinguga aitab võltskonto paljastada.



4. Vastamata jätmine

Kui proovida võtta ühendust mõne libakontoga, siis on väga ebatõenäoline, et see vastab isegi mõnele lühiküsimusele. Ideaalis aga on parem isegi mitte proovida ühendust võtta.

5. Enamasti tühi Facebooki sein

Üldiselt on võltskonto seinal ainult uued „meeldimised“, mõne ettevõtte või toote reklaamid ja uued sõbrad.

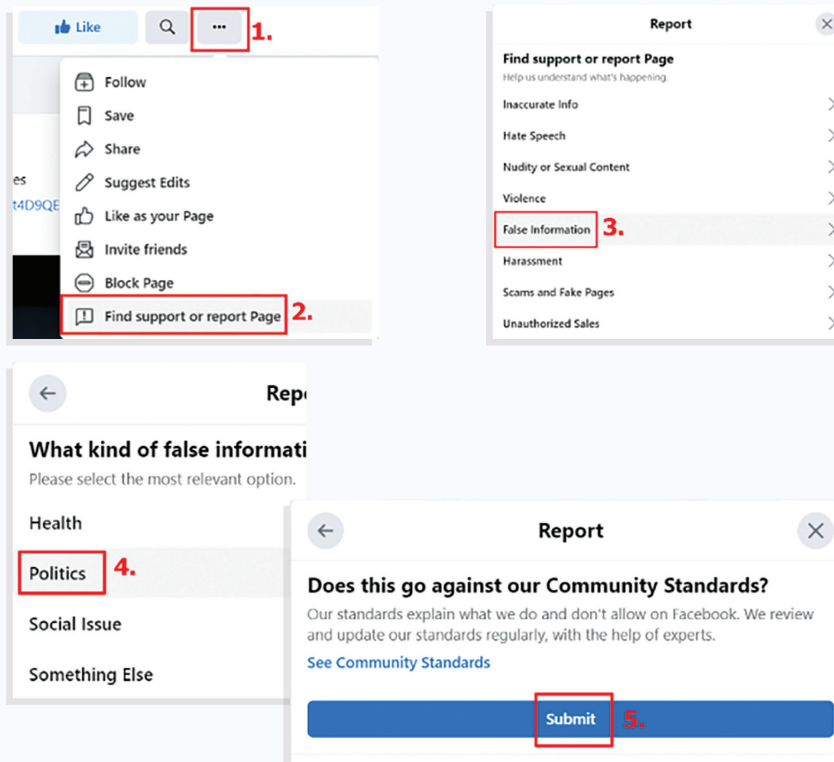
Vastumeetmed

Veebipõhise desinfo tõrjumiseks on vaja kahte protsessi - selle paljastamine ja sellest teatamine.

Facebooki / Twitteri / meediaväljaannete teavitamine

Väga oluline on teavitada oma sõpru, klassikaaslast ja kolleege avastatud desinfost, mis on suunatud mõne organisatsiooni vastu. Igas organisatsioonis peaks olema selge protsess, mille abil töötajad teavad, kuhu saata teateid desinfo loo kohta. Selle sammu peamine eesmärk on teadvustamine, et mõni ringlev sõnum on vale, ning takistada teistel seda lugu jagamast ja uskumast.

Teine samm on teatada sellest sotsiaalmeediaplatformile – neil kõigil on olemas võimalus raporteerida mõnest kahtlasest loost koos konkreetse põhjendusega. Kui sotsiaalmeediaplatform saab kasutajatelt piisavalt palju märkusi või kaebusi, siis võetakse see lugu, kommentaar või sõnum maha. See on meetod, mida kasutavad kodanikuühendused veebipõhise desinfo vastu võitlemiseks. Kui meediaväljaanded levitavad valeuudiseid, tuleks sõltuvalt meediaväljaannete olemusest (kas tegemist on usaldusväärse või propagandaväljaandega) sellest teatada kas neile otse või järelevalveasutustele).



Telegrami kasutamisega seotud riskid ja ohud

Kuna paljud noored kasutavad Telegrami igapäevaseks suhtlemiseks, on nad teadlikud ka selle platvormi teistest funktsioonidest - rühmadest ja kanalitest, millest mõned võivad olla privaatsed ja anonüümsed. Kui avatud suhtlusvood on sotsiaalmeedia puhul tavalised, siis kinnistel gruppidel võib olla spetsiifiline sisu, sealhulgas kallutatud ja moonutatud info, mille päritolu on väga raske kontrollida.

Avalikult on uuritud mitmeid erinevaid desinfokampaaniaid, mis on kasutanud Telegrami kanaleid ja gruppe vale- ja libauudiste levitamiseks, näiteks COVID-19 pandeemia ajal, sh Kremli-meelsed seisukohad Ukraina sõja kohta, paremäärmuslik retoorika ja vandenõuteooriad erinevates keeltes paljudes riikides.

Telegrami kanaleid ja gruppe võidakse kasutada ka kasutajate mobiliseerimiseks, et korraldada mõningaid ideoloogilisi miitinguid või poliitilisi proteste. Ei ole probleemi, kui need üritused on korraldatud demokraatlikult ja läbipaistvalt, kuid mõnikord on nende tegelikud kasusaajad varjatud ja nende isikute kohta, kes on tegevuste taga, on väga vähe teavet. Näiteks kasutati Eestis kampaania „Ma olen venelane“ ajal provokatiivsete üleskutsete levitamiseks mitmeid anonüümseid Telegrami gruppe ja kanaleid.

<https://eng.obozrevatel.com/section-life/news-russians-in-tallinn-threw-a-hysterical-tantrum-because-the-police-forced-them-to-remove-i-am-russian-stickers-from-their-cars-video-27-09-2023.html>

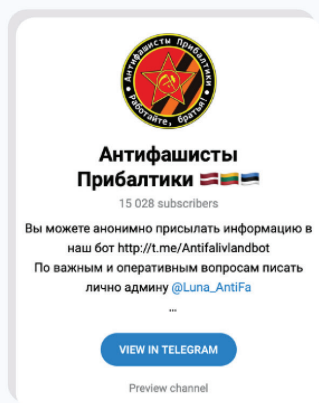


Kuna paljud desinfo levitajad ja propagandistid (ka Kremli-meelsed) peavad ja kasutavad oma Telegram-kanaleid, on lihtne jagada nende sisu teistele rühmadele, mida juhivad anonüümsed administraatorid. Need on üks levinumaid kanaleid, kuidas vabaduse-, lääne- ja Eesti-vastast desinfot digitaalselt levitatakse. Näiteks, Telegrami kanal „Антифашисты Прибалтики“ (Baltikumi antifašistid) on üritanud oma jälgijate arvu kasvatada, tehes mitmeid skandaalseid postitusi. Mõned postitused levitasid nt nn „russofoobia“ narratiivi. Üks postitus käsitles väidetavat russofoobiat Lätis, kus justkui otsustati keelustada „Potsataja“ näitamise mingis nukuteatris. Selles postituses õhutati ka Läti kultuuriministri vastast viha ja laimu, laadides üles üks (võltsitud) foto temast, kus ta poseerib grotesksete esemete kõrval, väites, et „see ongi Läti kultuuri ja rahvusliku identiteedi nägu“.

Teises postituses öeldi, et iga inimene „vabas Läänes“, kes julgeb isegi vihjata Venemaa õigusele kaitseda vene rahvast, satub kohe trellide taha, tema vara konfiskeeritakse ning talle keelatakse igasugune majanduslik ja loominguline tegevus. Sellised postitused näitavad, et Venemaa desinformatsioon üritab aktiivselt propageerida vaenulikku narratiivi, et Balti riikides rünnatakse vene kultuuri ja et venekeelne vähemus ei saa karistusi kartuses nendes küsimustes sõna võtta.

Mõned lihtsad soovitusused aitavad tugevdada digihügieeni Telegramis:

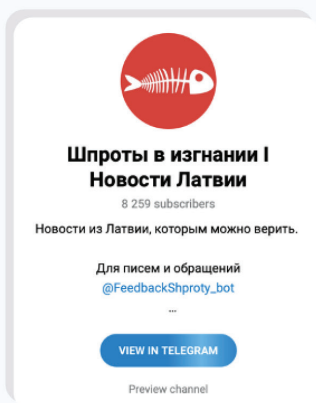
- enne grupiga või kanaliga liitumist veendu, et sisu on 100% see, millest oled huvitatud, sest nimi ja kindlastatud postitused võivad olla eksitavad, vaata pealkirjast sügavamale edasi ja otsi rohkem teavet selle kanali ja grupi administraatori(te) kohta.
- ole teadlik des-, väär- või kuriinfost anonüümsetes rühmades ja kanalites - kui mingi teema on tõsine, peab tegelikkusele vastav sisu olema tõendatud usaldusväärsete allikatega ega tohi sisaldada spekulatiivseid arvamusi, „alternatiivseid fakte“ või lihtsustatud propagandistlikke klišeesisid.
- kui mingi teave Telegrami grupis või kanalisis prooviseeris sind emotsionaalselt, küsi endalt, miks see juhtus ja kellele see kasulik on – ära kiirusta reageerima või jagama seda infot, sest see võib olla erapoolik, polariseeriv, solvav või lihtsalt võltsitud.
- igast häirivast sisust Telegrami grupis või kanalisis võib teatada nii administraatori(te)le, veebipolitseile kui ka faktikontrollijatele (nt CRI Leedus või Propastop Eestis). Kindlasti salvesta originaalsisu nii hästi kui võimalik (näiteks ekraanipilt, kopeeritud tekstid, kujutised, lingid jne).



Антифашисты Прибалтики
15 028 subscribers

Вы можете анонимно присылать информацию в наш бот <http://t.me/Antifalivlandbot>
По важным и оперативным вопросам писать лично админу @Luna_AntiFa

VIEW IN TELEGRAM
Preview channel

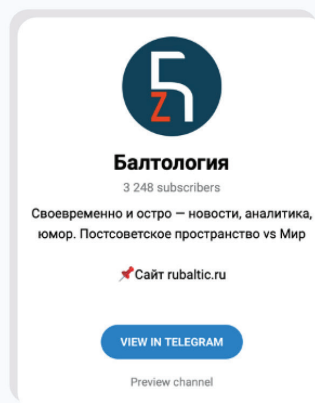


Шпроты в изгнании | Новости Латвии
8 259 subscribers

Новости из Латвии, которым можно верить.

Для писем и обращений @FeedbackShproty_bot

VIEW IN TELEGRAM
Preview channel



Балтология
3 248 subscribers

Своевременно и остро — новости, аналитика, юмор. Постсоветское пространство vs Мир

Сайт rubaltic.ru

VIEW IN TELEGRAM
Preview channel



TikToki kasutamise riskid ja ohud

Hiina päritoluga TikTok on sotsiaalvõrgustike seas n-ö tõusev täht, mis kogub populaarsust noorte seas vaatamata sellele, et selles võrgustikus liigub palju desinfot. TikTok'i asutajad ise küll rõhutavad, et nad teevad suuri jõupingutusi desinfo, radikaalse ekstremismi ja vihakõne levitamise piiramiseks. Kas see tegelikult ka nii?

Kuigi alguses ei tundunud see sotsiaalvõrgustik ohtlik, hakkas sinna ilmuma palju Kremlile soodsaid **narratiive***

***Narratiiv** on süstemaatiliselt ja järjepidevalt moodustatud temaatiline lugu, mis luuakse korduvatest sõnumitest, täiendades neid uute faktidega ja laiema kontekstiga.

Narratiiv on suurlugu, mis veenvalt edastab põhisõnumit ja kujundab arvamust.

Sotsiaalvõrgustik TikTok eristub teistest selle poolest, et sellel on sõltuvust tekitav mõju, kui inimesed vaatavad pidevalt oma nutitelefoni lühivideoid, millel on palju emotsionaalseid elemente ja lahe muusika. Siis hakkab tööle teabe algoritmiline esitamine, kuhu sekka lisatakse sageli ka propagandasõnumeid. Seetõttu iseloomustavad TikTok'i platvormi väga ebamäärased algoritmid ja üldiselt arvestab see ka autoritaarsete riigirežiimide seadusi.

TikToki kasutamisega on seotud kaks põhilist riski:

- **jõuline andmete kogumine.** Kui see rakendus paigaldatakse nutiseadmesse, nõuab see suuremat juurdepääsu andmetele, mida see hakkab koguma. Üks TikTok'i kasutamisega seotud suuri riske on see, hakkab koguma. Üks TikTok'i kasutamisega seotud suuri riske on see, et rakendus näeb kasutaja kontakte, samuti näeb see muid seadmes kasutatavaid rakendusi ja saab tuvastada nutiseadme asukohta. Samuti on suureks ohuks, et TikTok'i kirjavahetust saab jälgida ja teatud märksõnade põhjal võib see sattuda ettevõtte enda seireradarile.

- **kirjete varjatud keelustamine** (*shadow banning* - inglise keeles). Kui kasutaja teeb postituse, mille sisu ei meeldi TikTok'i võimudele, peidetakse see postitus ära, st rakendatakse tsensuuri.

TikTok kasutab algoritmi kui vahendit, et püüda ja hoida tähelepanu ning pakkuda igale kasutajale isikupärasest kogemust. Algoritm kasutab kasutajatelt kogutud andmeid, et teha kindlaks, milline sisu võiks sind konkreetselt huvitada. Näiteks, mida kauem vaatad mõnda videot, seda rohkem sarnaseid videoid näed tulevikus oma TikToki. Samuti jätab see meelde sinu otsingud huvide kohta, et pakkuda sulle tulevikus sarnase stiili ja sisuga videoid. Lisaks toob TikTok'i algoritm sarnaste huvidega kasutajaid kokku, näidates neile sarnast sisu. See ei ole pelgalt juhus, kui näed samu videoid, mis ka sinu sõbrad.

Pea meeles, et TikTok ei ole loodud tõsiste uudiste edastamiseks, sest seal olevad lühivideod on kohandatud kasutaja meelelahutuslikele huvidele. TikToki videote kerimine ja vaatamine aktiveerib inimaju neid osi, mis vastutavad eduelamuse eest (nagu hartsartmängudes), lootusega meeleolu tõsta. Kuid tegeliku kasumi asemel raisatakse tohutult aega. Kuna TikToki olevad videod on lühikesed, meelelahutuslikud ja kergesti kättesaadavad, tekitab see sõltuvust, mis viib tähelepanu kadumiseni.

TikTok kasutab ära nii uudishimu kui ka hirmu jääda millestki olulisest ilma. Kas oled kuulnud lauset „Kui sind ei ole TikTakis, siis ei ole sind olemas“? See on tungiv manipulatsioon, et panna noori inimesi olema pidevas kontaktis. Teine trikk, mida TikTok kasutab, on vihakõne: see pakub videoid, mis solvavad mingit inimgruppi, ideoloogiat või liikumist, lootuses, et selle sisse puudutab vaatajaid ja käivitab emotsionaalselt tulise arutelu, mis omakorda tõmbab rohkem tähelepanu ja muudab seeläbi videod viiruslikuks. Neid meetodeid kasutatakse ära klikkide ja jälgijate arvu suurendamiseks. Kuna TikTakis esineb palju veebikiusamise juhtumeid, on suur oht sattuda ahistamise või vihakõne ohvriks. Lisaks, kuna igaüks võib TikTakis sisse luua, võib see sisaldada kallutatud või manipuleeritud infot. Kuna TikTakis liigub palju isikustatud sisse, teeb see noortel veelgi raskemaks eristada kellegi suvalist arvamust tõsielude faktidest.

TikTok võib olla suurepärane platvorm meelelahutuseks, kui järgid lihtsaid soovitusi selle kasutamiseks:

- piira oma igapäevast aega, mida sa TikTakis veedad - sõpradega kohtumine päriselus on alati lõbusam kui veebis
- jää kriitiliseks TikTakis uudislaadse sisse suhtes – alati kontrolli täiendavalt teistest allikatest (mitte sotsiaalmeediast) teavet, mis sind huvitab
- teata veebikiusamisest ja vihakõnest, kui näed seda
- väldi solvavate videote või väga emotsionaalse ja häiriva sisse levitamist
- kui tunned, et oled segaduses või sinuga manipuleeritakse, pea nõu vanemate või veebipolitseinikuga.

Muud abivahendid

Kui arvad, et tegemist võib olla keerulisemate desinforjuhtumitega, võib kaaluda 2 järgmist lähenemisviisi: kasutada professionaalseid avatud lähtekoodiga meetodeid või pöörduda abi saamiseks faktikontrollijatest ekspertide poole. Enamik veebipõhiseid vahendeid on suhteliselt hõlpsasti kasutatavad ja annavad samm-sammult juhiseid. Allpool on esitatud kaks suurt ja kõige kasulikumat tööriista:

Bellingcat's Online Investigation Toolkit



Online Open Source Tool Box



Teine võimalus on pöörduda desinfo palistamisega tegelevate ekspertide poole ja anda neile vihje kahtlase loo / uudise / postituse kohta. Enamik veebiuurijaid paljastab valeuudise hea meelega ja jagab seda tulemust teiste kolleegidega.



Vilnius, Leedu
Tiraaž 500
©CRI, 2024



Juhendi väljaandmist
on toetanud

Google