CIVIC RESILIENCE INITIATIVE

Atsparumo Priemonių gidas

DI proveržis: Kovos su dezinformacija gidas

Publikacijos išleidimą remia Google

Daugiau apie mus!

www.cri.lt

CIVIC RESILIENCE INITIATIVE

@CivicResilience

@CRI

@civicresilienceinitiative

Leidinys sukurtas bendradarbiaujant su dezinformacijos srityje dirbančiomis organizacijomis Baltijos šalyse: „Baltic Security Foundation" (Latvija) ir „National Centre for Defence and Security Awareness" (Estija).

Publikacijos išleidimą remia Google.

DI proveržis: Kovos su dezinformacija gidas

Per pastaruosius dešimt metų skaitmeninių žiniasklaidos priemonių sklaida stipriai išsaugo. Kasdieną mus pasiekia informacijos srautai, plūstantys iš pačių įvairiausių komunikacijos kanalų: socialinių tinklų, tinklaraščių, interneto svetainių, tradicinės žiniasklaidos ar kitų elektroninių leidinių.

Tokia informacinių srautų įvairovė leidžia nesunkiai išsirinkti geriausiai mūsų interesus ir politines ar socialines pažiūras atspindintį šaltinį. Augant įsitraukimui ir socialiniuose tinkluose praleidžiamam laikui, vis daugiau žmonių būtent juos renkasi kaip pagrindinį informacijos šaltinį, dažnai neįvertindami ten egzistuojančių grėsmių. Sparti ir patogi informacijos sklaida socialiniuose tinkluose sukuria tobulas sąlygas greitam dezinformacijos plitimui.

Šiuo metu stipriai išpopuliarėjęs **dirbtinis intelektas (DI)** ir juo paremti **įrankiai suteikia plačias galimybes platinti dezinformaciją ir kenkėjišką informaciją.** DI pagalba sukurtos vaizdo ir garso klastotės gali padaryti didelę žalą visuomenei skleidžiant netikras naujienas ir tokiu būdu mažinant piliečių pasitikėjimą žiniasklaida. Visgi, su dideliu potencialu piktavališkoms provokacijoms ateina ir galimybės jas atremti. Įvaldžius DI, šios technologijos sumanumą galima pasitelkti ir kovoje su dezinformacija.

Lietuva yra nuolatinis melagingos informacijos, kuria aktyviai siekiama formuoti valstybės įvaizdį kaip nepilnavertį, taikinys.

Tokios pastangos kursto nepasitikėjimą vietos valdžia, mūsų partneriais NATO ir Europos Sąjunga bei nuolat bando demotyvuoti piliečius aktyviai dalyvauti valstybės valdyme. Nors profesionalai atlieka didelį darbą paneigdami melagingas istorijas, tačiau užtenka, kad vos viena iš jų pasiektų žiniasklaidą ir žala jau būna padaryta.

Siekdami šviesti visuomenę nesibaigiančioje kovoje su dezinformacija, 2019 metais įkūrėme organizaciją „Pilietinio Atsparumo Iniciatyva“ (angl. Civic Resilience Initiative, CRI). Organizacijoje bendradarbiaujant su dezinformacijos bei medijų ekspertais ir kilo idėja parengti praktinį vadovą „DI Proveržis: Kovos su dezinformacija priemonių gidas“.

Šis dezinformacijos atpažinimo priemonių rinkinys siūlo paprastą informacijos patikrai reikalingų metodų gidą. Tikimės, jog šio leidinio pagalba **nesunkiai išsiugdysite įprotį patikrinti perskaitytų naujienų šaltinius bei juos papildančių nuotraukų ir vaizdo įrašų autentiškumą.**

„Pilietinio atsparumo iniciatyvos“ komanda išsikėlė tikslą tapti pagrindiniu katalizatoriumi Baltijos regione stiprinant visuomenės skaitmeninį atsparumą.

„Pilietinio atsparumo iniciatyvos“ komanda

Šiame praktiniame dezinformacijos vadove rasite pagrindines priemones, kurios jums padės:

- patikrinti, ar informacija internete yra tikra, ar melaginga;
- suprasti, kaip veikia DI ir kaip galima atpažinti jo sugeneruotus vaizdus;
- identifikuoti trolius;
- identifikuoti netikras socialinių tinklų paskyras;
- atpažinti suklastotus vaizdus ir vaizdo įrašus internete;
- imtis veiksmų pastebėjus dezinformaciją;
- įspėti kitus apie socialinėje žiniasklaidoje skleidžiamas melagienas.

Šiuo priemonių gidu siekiame prisidėti prie skaitmeninio atsparumo didinimo ir suvokimo apie saugumą bei reikalingą budrumą informacinėje erdvėje ugdymo.

Šio leidinio tikslas - suteikti informaciją, kuri padėtų moksleiviams, studentams ir plačiajai visuomenei stiprinti skaitmeninį raštingumą ir atsparumą melagienoms.

Leidinys sukurtas bendradarbiaujant su dezinformacijos srityje dirbančiomis organizacijomis Baltijos šalyse: „Baltic Security Foundation“ (Latvija) ir „National Centre for Defence and Security Awareness“ (Estija).

Publikacijos išleidimą remia Google.

Yra daug skirtingų melagingos informacijos skleidimo formų:

Dezinformacija

– informacija, kuri yra melaginga ir sąmoningai sukurta siekiant pakenkti asmeniui, socialinei grupei, organizacijai ar šaliai.

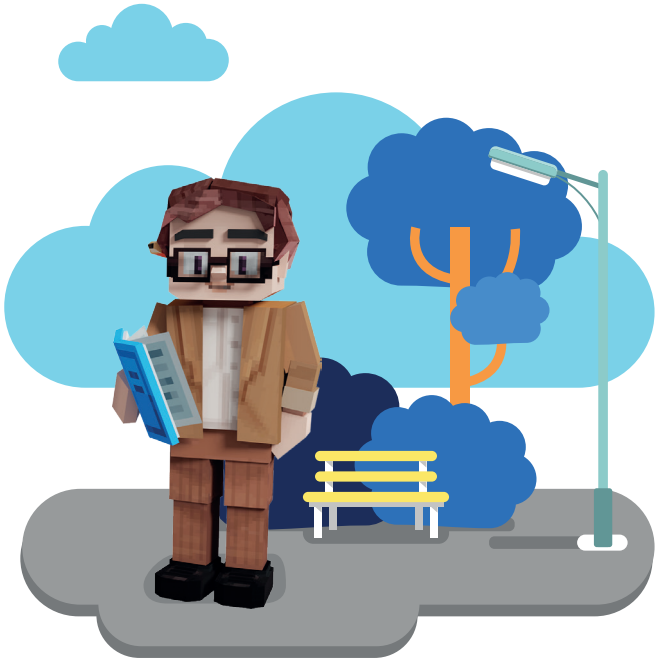
Misinformacija

– informacija, kuri yra melaginga, bet sukurta nesiekiant padaryti žalos.

Malinformacija

– realybe grįsta informacija, naudojama siekiant padaryti žalos asmeniui, organizacijai arba šaliai.

Visi šie informacijos tipai yra pavojingi, nes dėl sensacingo ar tariamai įdomaus turinio jais sparčiai dalinamasi socialiniuose tinkluose. Taip nutinka, kai vartotojai nepatikrindami informacijos ir neapsvarstydami istorijos įtakos ja dalinasi savo paskyroje.



Atpažinimas

Kaip patikrinti naujienas arba įrašus?

Skaitote skandalingą naujieną, kurios tema ar turinys skamba neįtikėtinai? Atsitraukite ir įsitikinkite informacijos patikimumu.

Ką reikėtų daryti?

5 paprasti veiksmai, kuriuos galite atlikti tikrindami informaciją:

1. Įvertinkite šaltinį	Panaršykite svetainėje arba socialinių tinklų paskyroje. Pagalvokite, kas išplatino naujienas ir koks yra pačios istorijos tikslas.
2. Skaitykite ne tik antraštę	Istorijų antraštės gali būti skandalingos, siekiant pritraukti daugiau paspaudimų ir skatinti dalinimąsi. Pasigilinus į istoriją gali paaiškėti, kad antraštėje pateikti teiginiai nėra teisingi.
3. Patikrinkite autorių	Ar nurodytas autorius tikrai egzistuoja? Ar autorius yra patikimas žmogus?
4. Ar šaltiniai patvirtina istoriją?	Dažnai melagingose naujienose nėra nuorodų, kurias paspaudus būtų galima patikrinti faktus. Jeigu jos yra, tačiau manote, kad šaltinis abejotinas, nespauskite ant nuorodų. Gali paaiškėti, kad pradinė žinutė buvo pagražinta arba jos prasmė iškreipta nenurodant jokių šaltinių ir papildomų nuorodų.
5. Patikrinkite datą	Iš naujo paskelbtos senos naujienos nebūtinai yra vis dar aktualios.

Papildomi patarimai:

Rinkitės saugią ir patikimą informaciją

Naujienų skaitymui lietuvių kalba naudokite didžiuosius portalus, tokius kaip LRT, DELFI, 15min, Alfa, TV3 ir panašius į juos. Mažiesiems portalams dėl mažesnių žmogiškųjų ir finansinių resursų gali būti lengviau padaryti įtaką perkant tam tikrą turinį, ten dirba mažiau žurnalistų, galinčių tikrinti informacijos patikimumą, taip pat, į juos gali būti lengviau įsilaužti.

Atsargiai filtruokite informaciją, perskaitytą užsienio portaluose ar socialinių tinklų grupėse:

• Jūsų šaltiniai privalo turėti pakankamą sekėjų ratą. Priešingu atveju ieškokite, kas dar publikavo tokią pačią informaciją ir filtruokite tuos asmenis iš naujo.

• Jei renkate informaciją „X“ (anksčiau žinomas kaip „Twitter“), „TikTok“ ar „YouTube“ platformose, patikimuose kanaluose komentarai po įrašais privalo būti įjungti. Patikimesni įrašai turės ne vieną komentarą, tarp kurių paskelbimo bus solidūs laiko tarpai, jų nesies rašymo stiliaus šablonai. Priešingu atveju komentarų autoriais gali būti troliai ar botai.

• Šaltiniai neturėtų turėti sąsajų su Rusijos valdžiai palankiais portalais (Sputnik, PBK, Rossija 24 ir kiti). Nors šių portalų skelbiamose naujienose ir būna tiesos, tačiau vykstant informaciniam karui nevertėtų rizikuoti ir tikėti rusiškais naujienų šaltiniais, kol jų nepatvirtina Lietuvos ar jos užsienio partnerių portalai.

• Informacija, kuri yra sensacinga, visada turėtų būti paremta patikimais šaltiniais. Jei internete pastebite išskirtinę naujieną, pasižiūrėkite, ar ją skelbiantis šaltinis nurodo ir savo šaltinį, iš kurio šią naujieną sužinojo. Jeigu ne, atlikitę patikrą su Google pagalba:

1) Prie svarbiausių raktažodžių iš jūsų skaitomo įrašo ar straipsnio pridėkite kabutes ir įveskite į Google paiešką.

2) Spauskite mygtuką „Įrankiai“ dešiniajame krašte šalia paieškos lango ir pasirinkite, kad rodytų naujausius.

3) Pateiktų šaltinių patikrą vėl atlikite iš naujo pagal minėtus kriterijus.

• Paklauskite saves, ar šiame šaltinyje tikrai netrūksta įrodymų ir analizės? Iš kurios pusės perspektyvos naujiena yra pateikiama? Ar nėra taip, kad naujienos net nėra ir bandoma manipuliuoti skaitytojų vaizduote ir emocijomis?

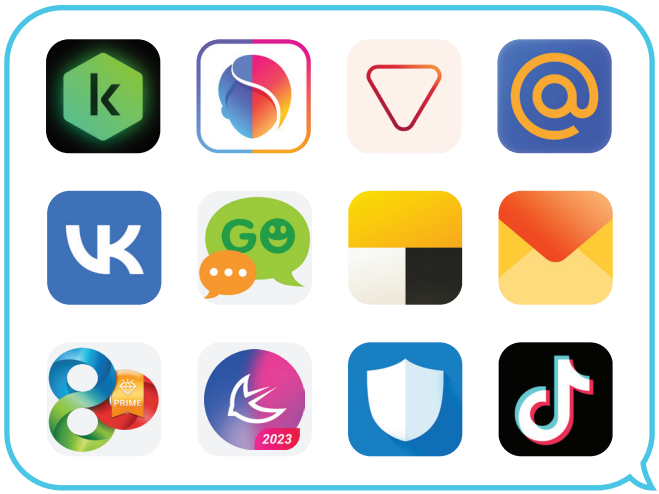
Galiausiai, pamatę sensacingą naujieną vienoje iš mūsų valstybinių institucijų svetainių (pavyzdžiui, Užsienio reikalų ministerijos, urm.lt), apsilankykite patikimuose naujienų portaluose. Viena iš informacinio karo priemonių yra įsilaužimai būtent į mūsų valstybines svetaines bei klaidinančios informacijos platinimas jose. Jeigu į mūsų valstybinių institucijų sistemas bus įsilaužta, medijos neabejotinai apie tai praneš. Tuomet neturėtume tikėti jose skelbiama informacija. Kai kuriais atvejais nėra būtina ir įsilaužti, o gali būti naudojami padirbti svetainių adresai. Pavyzdžiui, vietoje urm.lt atsirastų urm.com ar, dar gudriau, urm.it (i vietoje l). Tokius triukus pastebėti sunku, todėl visada žiūrėkite, ar adreso eilutė tikrai atrodo taip, kaip turėtų.

Pasiūlykite ne tik savo tėvams, bet ir seneliams

Ar jūsų šeimos nariai žiūri rusišką televiziją? Karinė informacija yra viena populiariausių ir dažniausiai rusiškų televizijos kanalų aptarinėjamų temų pastaruosius keletą metų. Suprantama, jog daugelio tėvams ar seneliams rusų kalba yra pagrindinė ir priimtiniausia užsienio kalba. Todėl siūlytume savo šeimos nariams įdiegti „Netflix“ ar kitą namų kino platformą, kurioje būtų ir rusišką vertimą turinčio turinio. Kadangi rusiška televizija yra sukurta taip, kad žiūrovas taptų apatiškas informacijai, perėjimas prie namų kino platformos neturėtų būti sudėtingas.

Pašalinkite rusiškas ir kiniškas programėles

Pašalinkite bet kokias rusų ar jų potencialių sąjungininkų kurtas programėles iš savo telefonų. Dalis jų turi įvairias prieigas prie jūsų duomenų ar lokacijos, o tai gali būti panaudota prieš jus tiek dezinformacijai skleisti, tiek karinių veiksmų metu. Tarp tokių programėlių populiariausios būtų: Kaspersky Antivirus, CheckScan, FaceApp, MyPocket, Mail.ru, Vkontakte, Go SMS Pro, Yandex Taxi, Yandex Mail, Go Launcher, APUS Launcher, Security Master ir net TikTok.



Generatyvinis dirbtinis intelektas (DI)

Generatyvinio DI modeliai gali generuoti įvairų turinį: nuotraukas, tekstą, muziką ar video. Generatyvaus DI pavyzdžiai yra daugeliui gerai žinomi ChatGPT, DALL-E ir kitos programos, kuriose pateikę žodines užklausas galite gauti sugeneruotą tekstinę ar vaizdinę informaciją.

Dideli kalbos modeliai (angl. large language models, LLM) yra generatyvaus DI rūšis, kurie gali suprasti, atpažinti, sukonkretinti ir generuoti tekstą. Norint, kad LLM galėtų atlikti savo užduotį, modelius būtina treniruoti, o tam reikia daugybės tekstinių duomenų. Duomenys dažniausiai gaunami nuskaitant viešai prieinamą informaciją internete ir ją transformuojant.

Neretai kalbos modeliai yra naudojami parašyti tekstą, kuriam neturime laiko ruoštis, ar netgi programavimo kodą, kai gerai nemokame programavimo kalbos. Tai yra puiki priemonė gauti informaciją ar įgyvendinti užduotį, kuri įprastai užimtų nemažai laiko ir pastangų.

Tačiau didieji kalbos modeliai yra naudojami ne tik geriems tikslams. Jie įgalina nusikaltėlius atlikti kibernetines atakas žymiai efektyviau.

1. Socialinė inžinerija

Didieji kalbos modeliai gali sukurti puikų tekstą bet kokia kalba ir šis tekstas bus taisyklingesnis nei išverstas per „Google Translate“ programą. Tekstas bus kokybiškas su mažai klaidų ar visai be jų. Tai leidžia kibernetiniam nusikaltėliui nemokėti pasirinktos aukos kalbos ir vis tiek parašyti tekstą, kuriuo bus bandoma išvilioti pinigus ar priverčiama suvesti prisijungimo duomenis spaudžiant ant nuorodos. Nuo šiol kenksmingi laiškai, žinutės ir pan. bus dar kokybiškesnės ir sunkiau identifikuojamos kaip brukalas. Egzistuoja dedikuoti kalbos modeliai, kurie yra sukurti būtent tokių laiškų generavimui.

2. Klaidingos informacijos skleidimas

Didžiųjų kalbos modelių programos sugeneruotas atsakymas ne visada yra teisingas ir gali neatspindėti tikrovės, todėl vartotojas gali būti suklaidintas, o jo gautas atsakymas pradėti plisti vartotojo socialiniuose ratuose. Tokie modeliai taip pat mokosi iš vartotojų užduodamų klausimų ir pateiktų patikslinimų, todėl jie gali kaupti klaidingą informaciją, mokytis iš jos ir vėliau ją pateikti kaip teisingą kitiems vartotojams.

Generatyvinis dirbtinis intelektas (DI)

Kaip atpažinti didžiųjų kalbos modelių sugeneruotą tekstą?

Šiuo metu nėra įrankio, kuris tiksliai pasakytų, tekstas parašytas dirbtinio intelekto ar ne. Geriausias atpažinimo įrankis yra žmogaus smegenys.

Pasitrenikuok: žemiau esantis tekstas yra parašytas su ChatGPT pagalba. Ar pastebėsi jame esančias klaidas?

Sveiki!

Dėkojame, kad pasirinkote mūsų paslaugas. Norėdami pasinaudoti visais privalumais ir pasirinktiniais funkcionalumais, prašome paspausti [Prisijungti čia], kad galėtumėte prisijungti prie savo paskyros.

Jei dar neturite paskyros, galite ją lengvai sukurti [Registruotis čia].

Ačiū už jūsų pasirinkimą!

Šis trumpas tekstas turi dvi vietas, kurios kelią įtarimą:
• Žodis „pasirinktiniais“ nėra dažnai vartojamas lietuvių kalboje ir turbūt niekada tokiame kontekste.
• Taip pat žodis „jūsų“ yra iš mažosios raidės. Oficialiuose laiškuose lietuvių kalba kreipiniai dažniausiai yra rašomi didžiąja raide.

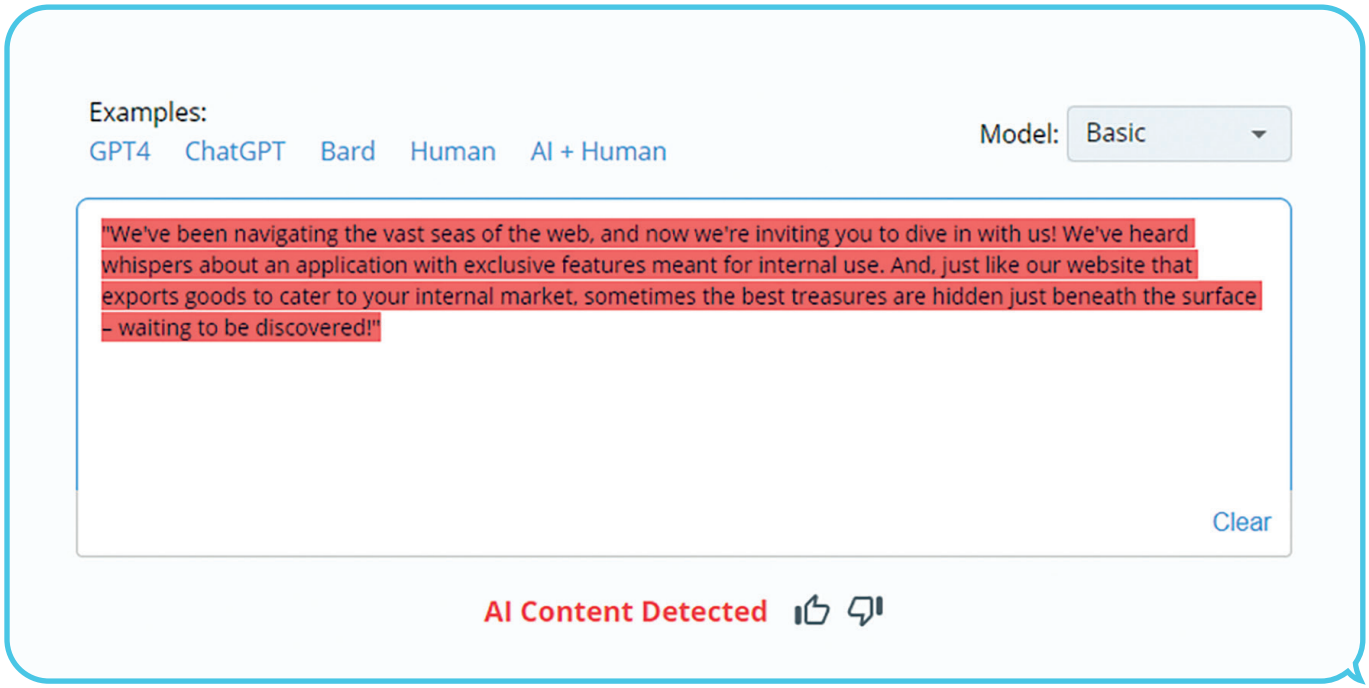
Šis pavyzdys parodo, kad dirbtinis intelektas generuodamas tekstą panaudoja žodžius, kurie nėra dažnai naudojami lietuvių kalboje. Pagalvokite, ar taip kalbate ir ar girdite kasdien kalbant kitus? Kitos užuominos gali būti susijusios su parenkamais neteisingais linksniais ar padaromomis sakinio struktūros klaidomis. Kartais ir pati kalbos kultūra bei viso teksto stilius gali išduoti dirbtinio intelekto darbą nepaisant to, kad klaidų tekste nėra.

Yra keletas internetinių svetainių, kurios gali padėti patikrinti, ar tekstas yra sugeneruotas dirbtinio intelekto. Norint suprasti, kaip automatiniai įrankiai stengiasi atpažinti, ar tekstą parašė DI, reikėtų grįžti prie pradmenų. Pasižiūrėkime, kaip didieji kalbos modeliai generuoja tekstą.

Visiškai supaprastinus šį procesą, tekstą generuojantys modeliai bando atspėti sekantį geriausiai tinkantį žodį sakinyje – tai juos daro nuspėjamais. Kuo daugiau duomenų buvo panaudota treniruojant modelį, tuo DI geriau atspės sekantį žodį. Kai kurie automatiniai aptikimo įrankiai stengiasi nustatyti, ar sakinyje naudojamų žodžių seka yra statistiškai optimaliausia, dar kiti bando atvirkštinį procesą ir stengiasi apskaičiuoti tikimybę, kad tekstą parašė žmogus, analizuodami žmogaus rašymo modelius.

Pateikiame keletą automatinių teksto patikrinimo įrankių pavyzdžių. Įprastai jie turi lauką, į kurį galima įklijuoti įtartiną tekstą. Paspaudus analizavimo mygtuką, įrankis parašys, kokia tikimybė, kad tekstas yra sugeneruotas DI. Kai kurie įrankiai taip pat gali pažymėti teksto vietas, kurios, pasak analizės rezultatų, yra sugeneruotos DI pagalba, o ne parašytos žmogaus.

Copyleaks



Zerogpt

Your Text is AI/GPT Generated

87.91%
AI GPT*

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.

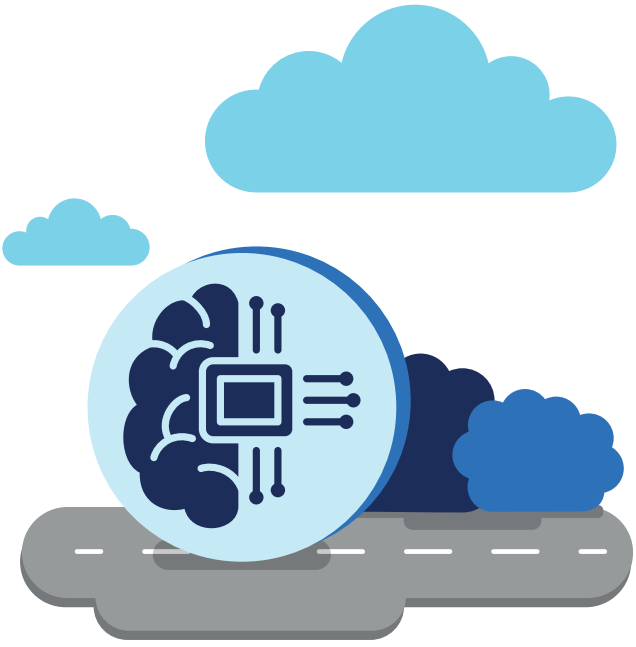
Best regards,
The [Your Website] Team

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

Writer

Gptzero

Ilm



Tikslinės informacijos patikrinimas

Kaip patikrinti, ar tekste skelbiama informacija yra tikra?

Google, Google, Google!

Paieškos sistemos, pavyzdžiui, „Google“, yra geriausias jūsų draugas kovoje su dezinformacija. Jų pagalba visą informaciją galima patikrinti naršyklėje, nustatant, ar ji teisinga, ar melaginga. Paieškoje naudokite raktinius žodžius, kuriuos galite identifikuoti pagal jus dominančią informaciją – tokiu būdu sužinosite, ar apie rūpimą naujieną kalba patikimi šaltiniai. Jei pamatėte įtartino turinio informaciją socialiniuose tinkluose, didelė tikimybė, kad kažkas jau stengiasi demaskuoti melagingos informacijos skleidėjus ir atskleisti tiesą.

Kovoje su dezinformacija turime didelį pranašumą, kadangi sėkmingai identifikavę raktinius žodžius, galime patikrinti ir ne skaitmeninę informaciją bei taip įvertinti iš skirtingų šaltinių mus pasiekiančią informaciją.

Svarbu pabrėžti, kad naudojimasis paieškos sistema leis jums rasti patikimą šaltinį su tuo pačiu turiniu, kurį bandote patikrinti, tačiau pirmiausia įsitikinkite šaltinio patikimumu.

Kaip patikrinti?

1. Nustatykite raktinius žodžius, geriausiai apibūdinančius jūsų ieškomą informaciją.
2. Norėdami rasti tą pačią informaciją, pasinaudokite viena iš prieinamų paieškos sistemų.
3. Pasirinkite šaltinių filtravimą pagal „Naujienas“ ir naujausius įrašus, paskelbtus per paskutinę valandą ar parą.
4. Atpažinkite patikimus šaltinius ir patikrinkite informaciją.

Pagrindinės taisyklės, kurias verta įsidėmėti:

• „Google“ paieška yra geriausias pirmasis informacijos patikrinimo žingsnis. Tai itin greitas ir efektyvus būdas.

• Nesvarbu, kokią informaciją norite patikrinti, ieškant teksto, nuotraukų ar vaizdo įrašų, „Google“ paieška gali veikti labai efektyviai.

• Svarbiausia yra naudoti raktinius žodžius – taip patikimame šaltinyje rasite ieškomą informaciją.

Norėdami efektyviau pasinaudoti „Google“ paieška:

1) Naudokite pagrindines „Google Dorking“ gudrybes, t. y., jei ieškote frazės, rašykite ją tarp kabučių („...“ ar “....”).

2) Jei jūsų paieškos rezultatuose yra labai populiarus raktinis žodis, prirašykite prie jo minusą (**-**), kad jo neįtrauktumėte į paieškos rezultatus.

3) Jei nežinote tikslios žodžio rašybos, tikslaus skaičiaus ar datos, galite prirašyti „*“ kaip pakaitos simbolį, kad pakeistumėte trūkstamą raidę žodyje, visą žodį, skaičių ar datą.

• Jei kalbate užsienio kalbomis, galite tuo pasinaudoti. Ieškokite, ką apie rūpimą naujieną rašo šaltiniai kaimyninėse šalyse ar visame regione.

• Jei puikiai mokate tik vieną kalbą, kreipkitės pagalbos į patikimą žmogų, kuris gebėtų patikrinti svarbią informaciją kitomis kalbomis. Vėliau aptarkite, kaip rūpima informacija pateikiama kitose kalbinėse erdvėse. Tokia praktika gali praturtinti tiek jūsų, tiek padedančiojo žinias.

Naudingos priemonės:

Google

Bing

Yandex

(SVARBU: būkite atsargūs, tai rusiškas puslapis, todėl naudokite papildomas priemones kompiuteryje saugumo užtikrinimui. Kiek žinoma, naudotis internetine svetaine yra saugu, tačiau rekomenduojame nenaudoti mobiliosios programlės - ją įdiegiant prašoma prieigų prie asmeninių duomenų.)

Vaizdų patikrinimas

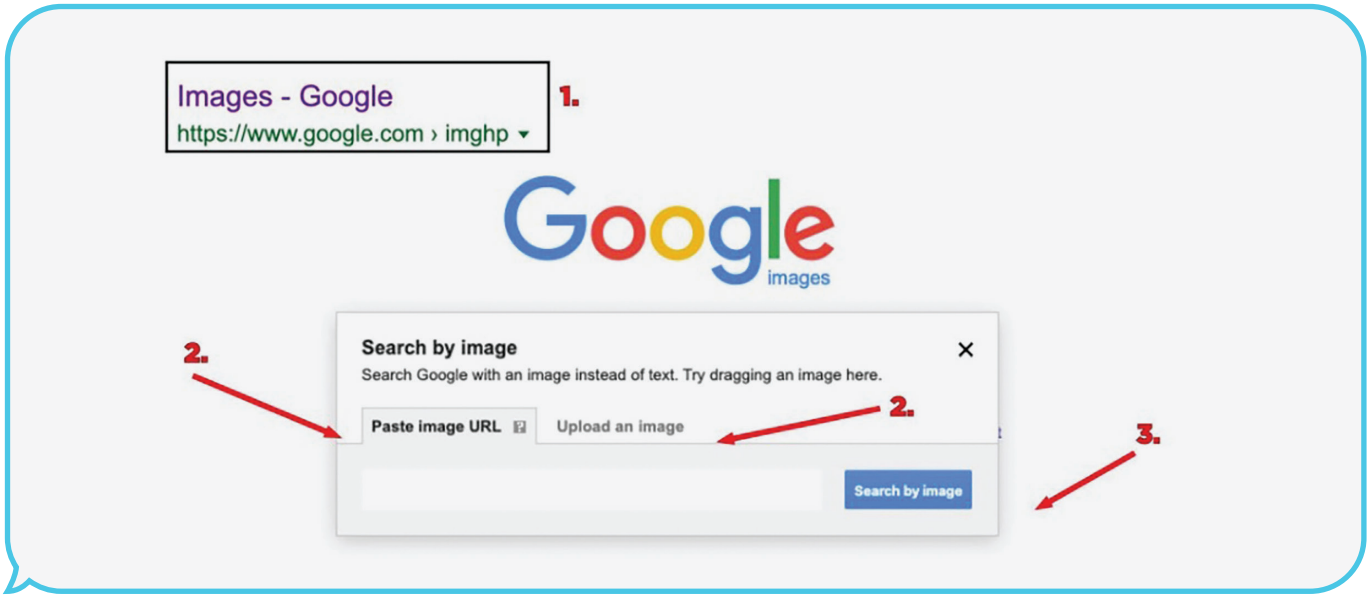
Kaip patikrinti, ar vizuali informacija yra tikra?

Atvirkštinė vaizdo paieška

Vaizdų perdirbimas (angl. image recycling) - prieš tai buvusio vaizdo paskelbimas ir teigimas, kad jis buvo užfiksuotas neseniai - išlieka viena pagrindinių dezinformacijos erdvės problemų. Kalbant apie netikrus ar modifikuotus vaizdus, geriausia jų atpažinimo ir patikrinimo metodika yra atvirkštinė vaizdo paieška. Tai leidžia mums surasti visus anksčiau paviešintus identiškus arba labai panašius vaizdus. Nustatymas, kad vaizdas buvo paskelbtas anksčiau, yra patikimas būdas patvirtinti, jog vaizdas internete patalpintas jau seniai. Kitais atvejais, jei įtarimų sukėlęs vaizdas buvo pakeistas, atvirkštinė vaizdo paieška gali padėti rasti originalų vaizdą.

Kaip patikrinti?

1. Atidarykite vieną iš paieškos sistemų (nuorodos pateikiamos ties „Naudingos priemonės“).
2. Nukopijuokite įtarimų keliančio vaizdo nuorodą arba atsisiųskite ir įkelkite patį vaizdą.
3. Atlikite analizę, ar tie patys arba labai panašūs vaizdai nebuvo paskelbti anksčiau.



Klaidų lygio analizė

Klaidų lygio analizė (angl. error level analysis) yra pažangus metodas, leidžiantis atpažinti vaizdo vietas, kuriose yra skirtingas vaizdo apdorojimo lygis. JPEG formato vaizduose visas vaizdas turėtų būti maždaug tokio paties lygio. Jei vienos vaizdo dalies klaidų lygis žymiai skiriasi, greičiausiai vaizdas buvo pakeistas skaitmeniniu būdu. Palyginkite tokias vietas su klaidų lygio analizės rezultatais – jeigu tam tikros vaizdo dalys ženkliai skiriasi, analizė identifikuoja įtartinas vietas, kurios galėjo būti pakeistos skaitmeniniu būdu.

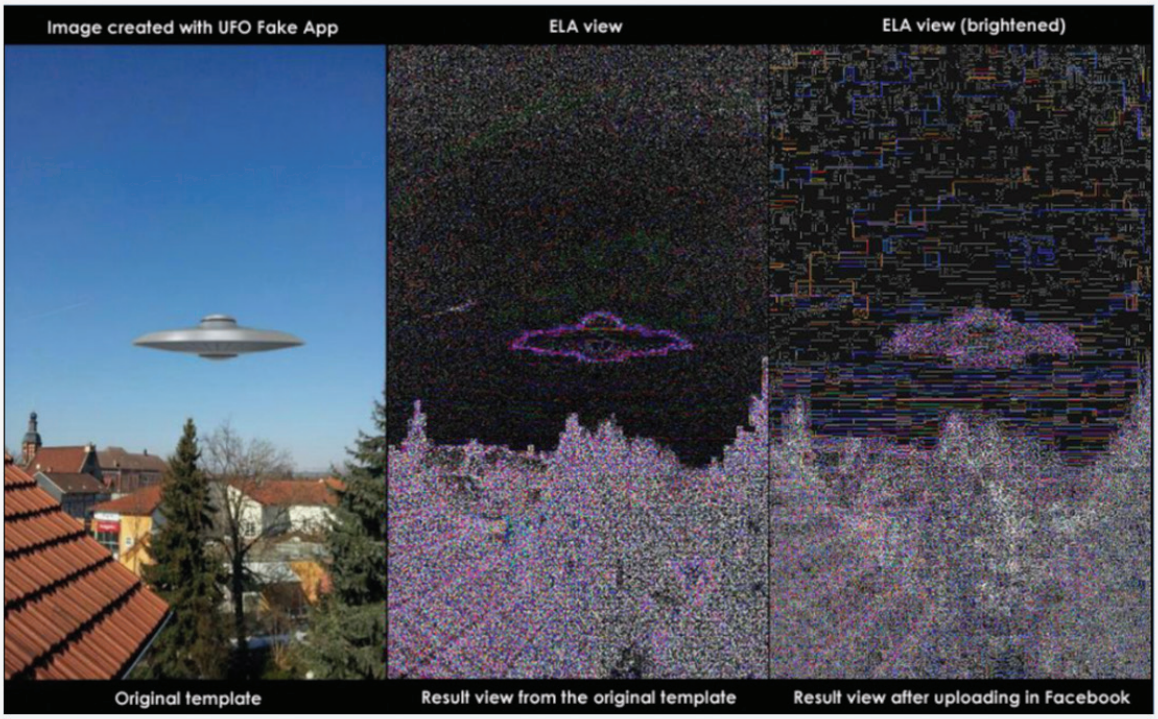
Svarbu pažymėti, kad šis metodas turi trūkumų ir nėra šimtu procentų patikimas, tačiau tai yra pirmas žingsnis, leidžiantis identifikuoti skaitmeninius vaizdo pakeitimus. Nuotraukų kriminalistika yra atskira mokslo šaka, o norint demaskuoti gerai perdirbtus vaizdus, reikalinga ilgametė patirtis. Kasdienėms propagandinėms žinutėms dažniausiai naudojami vaizdai nėra gerai perdirbti ir juos gana lengva atpažinti.

Pagrindinės taisyklės, kurias verta įsidėmėti:

• Prieš pasitikint vaizdu, atvirkštinė vaizdo paieška turėtų būti įprasta praktika. Realybės ar dabarties laiko neatitinkantys vaizdai gali iškreipti tikrąją žinutę.

• Klaidų lygio analizė nėra metodas, kuriuo vertėtų aklai pasitikėti, tačiau tai efektyvus būdas patikrinti vaizdo originalumą.

• Atvirkštine vaizdo paieška galite pasinaudoti norėdami paveikslėlyje atpažinti nežinomus žmones, vietas, pastatus ir kitą informaciją.



Naudingos priemonės:

Google /atvirkštinė vaizdo paieška

Yandex /atvirkštinė vaizdo paieška

Google /plėtinys „RevEye“

Forensically

Foto Forensics

Generatyvinis dirbtinis intelektas (DI)

Kaip atpažinti sugeneruotą vaizdą?

Sugeneruotą vaizdą atpažinti yra lengviau nei sugeneruotą tekstą, ypač jei yra bandoma generuoti tikroviškas nuotraukas. Keletas pavyzdžių pateikiama žemiau.
Nuotraukose yra pavaizduota žmonių minia Vilniaus senamiestyje. Iš pirmo žvilgsnio atrodo, kad nuotraukos yra tikros, tačiau įsižiūrėkime labiau:

• Spalvos – kairėje nuotraukoje, spalvos atrodo nenatūralios ir per daug ryškios, rėžiančios akį.

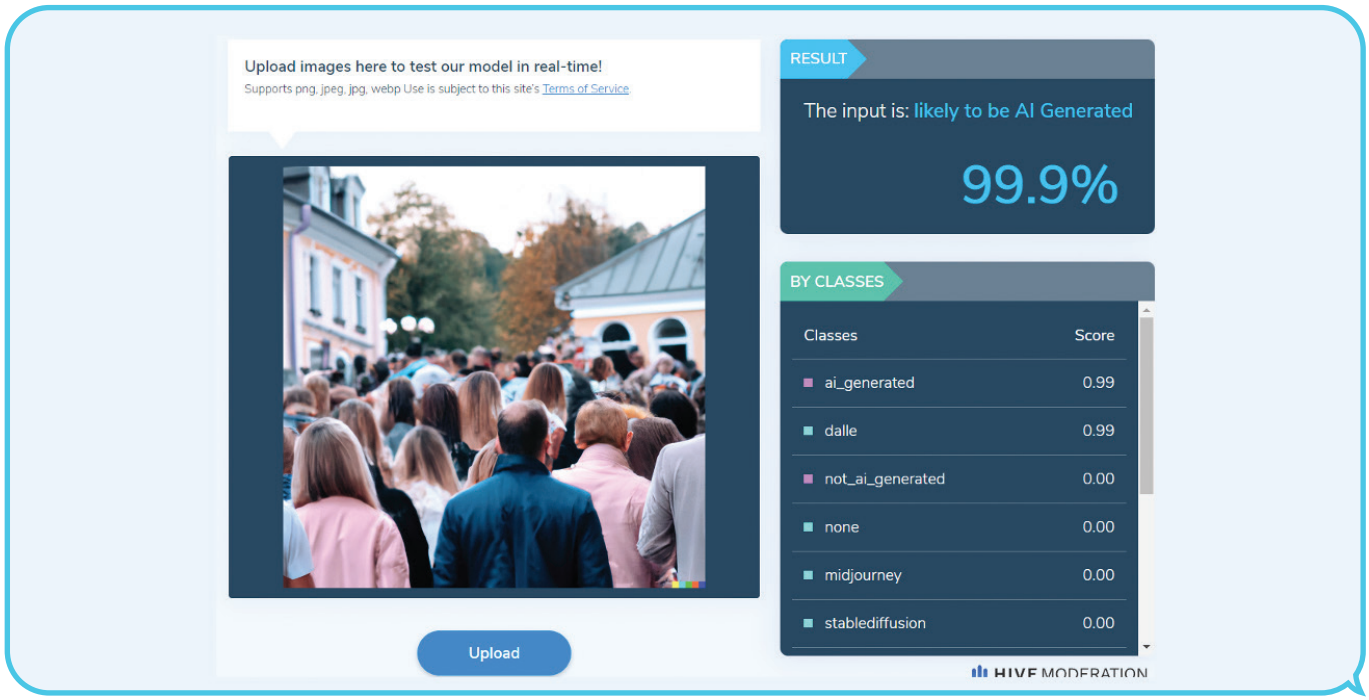
• Fonas – dažnai sugeneruotos nuotraukos turi neryškų foną arba jame esančių pastatų ar automobilių linijos yra iškraipomos.

• Anomalijos ir iškraipymai – priartinus nuotraukas iškart pastebėsite, kad žmonių kūno dalys yra nenatūralios ir iškraipytos: ausys atrodo keistai, galvos lyg tapytos teptuku. Taip pat, DI intelektas dažnai prideda daugiau nei 5 pirštus ar daugiau dantų, negu žmogus gali turėti. Visada atkreipkite dėmesį į smulkmenas.

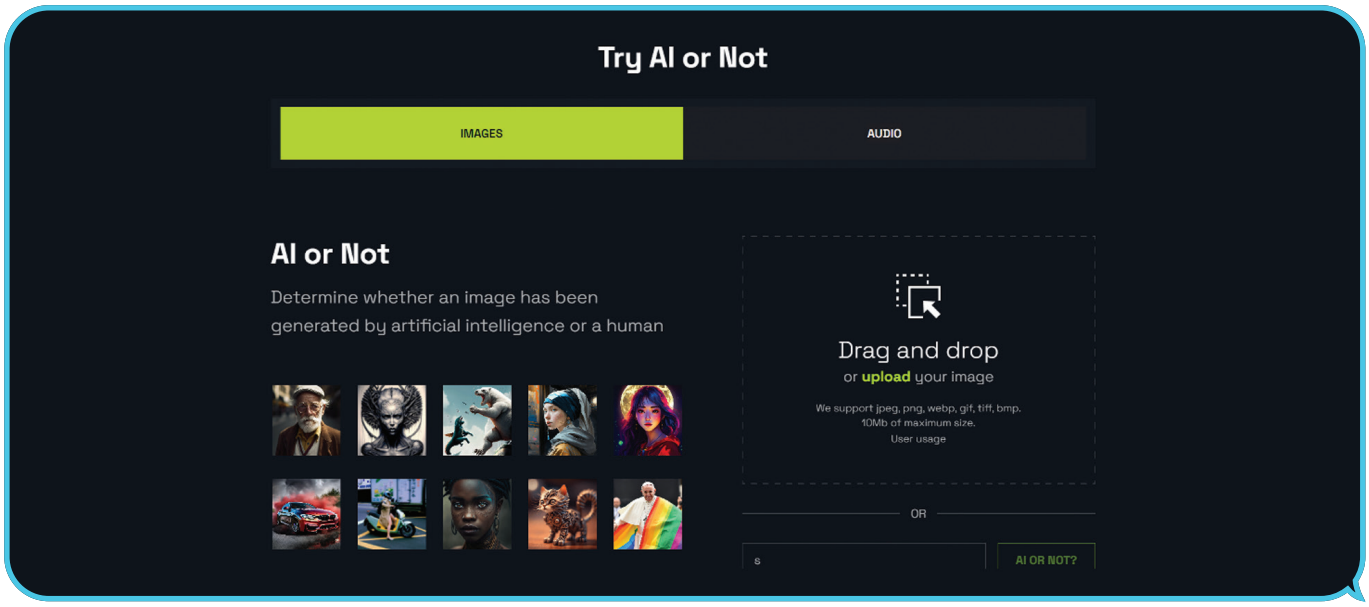
• Vandens ženklai (angl. watermarks) – įrankiai, ypač nemokami, dažnai prideda savo vandens ženklus nuotraukose. Pavyzdžiuose galime pastebėti dešiniam apatiniam kampe esančius spalvų kvadratus – tai ženklas, kad vaizdai buvo sugeneruoti Dall-E pagalba. Aiškiai matomą vandens ženklą galima nesunkiai panaikinti rankiniu būdu, tačiau atsiranda vis daugiau plika akimi nematomų vandens ženklų. Jei tokią nuotrauką įkelsime į tikrinimo įrankį, jis iškart nustatys, kad tai nėra tikras vaizdas.

Štai keletas automatinių įrankių, kurie padės atpažinti sugeneruotas nuotraukas. Jais naudotis labai paprasta: užveskite ar įkelkite nuotrauką ant specialios puslapio dalies ir paspauskite „įkelti“. Po kelių sekundžių gausite rezultatą:

Hivemoderation



Aiornot



Vaizdo įrašų patikrinimas

Kaip patikrinti, ar vaizdo įrašas yra tikras?

Atvirkštinė vaizdo paieška

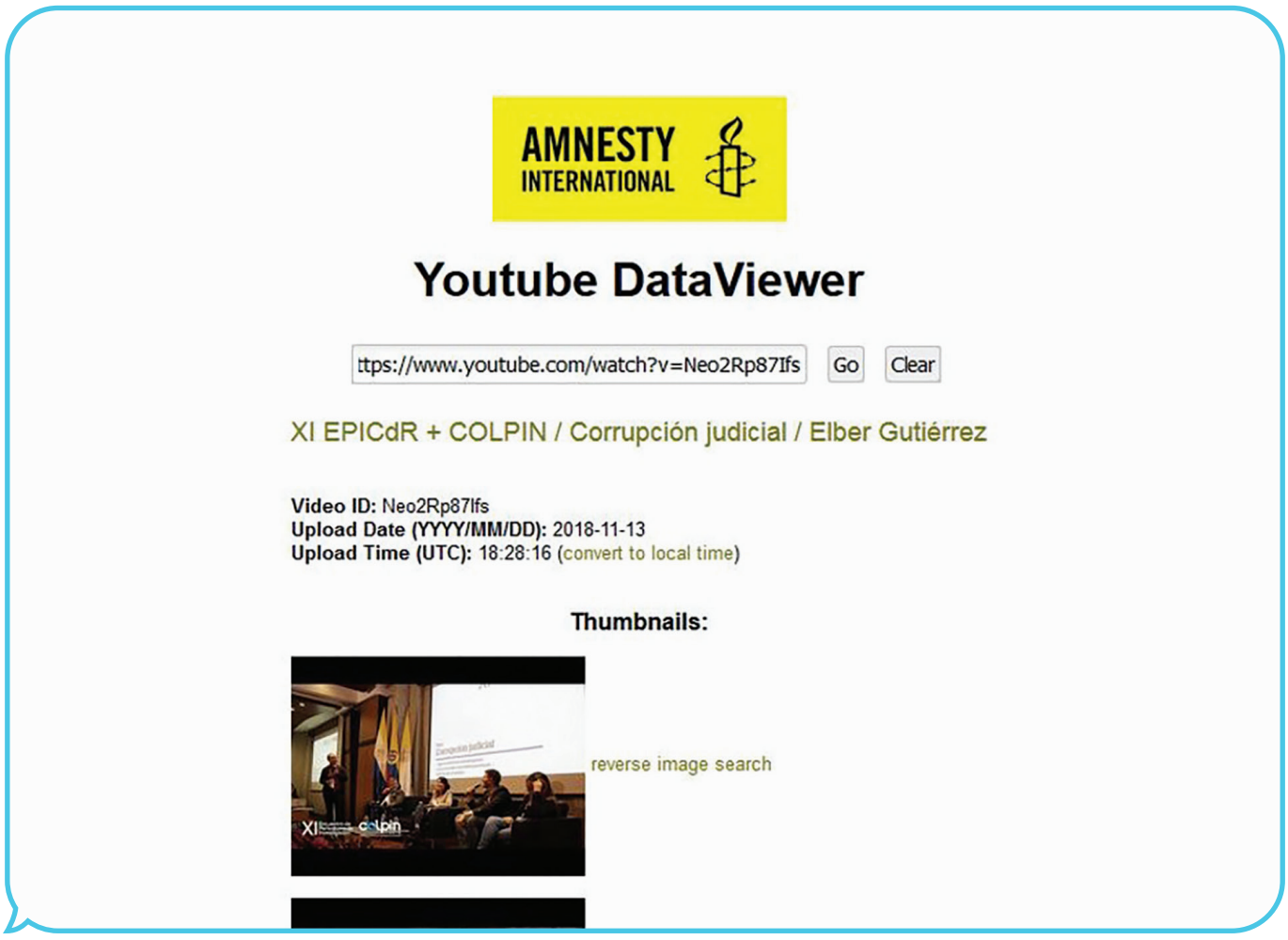
Tikrinant vaizdo įrašus, geriausia pasinaudoti atvirkštine vaizdo paieška, kuri padeda atpažinti ir jau aptartus sugeneruotus statiškus vaizdus. Kadangi vaizdo įrašai yra tik vaizdų serija, paieškai puikiai tinka atskiro kadro išėmimas ir individuali jo paieška. „InVid“ ir „Amnesty DataViewer“ įrankiai leis atrasti jau anksčiau internete paskelbtus panašius arba identiškus vaizdo įrašus, ieškant tiek kadrų, tiek miniatiūrų (angl. thumbnails).

Kaip patikrinti?

1. Atidarykite vieną iš paieškos sistemų („Amnesty DataViewer“ arba „InVid“).

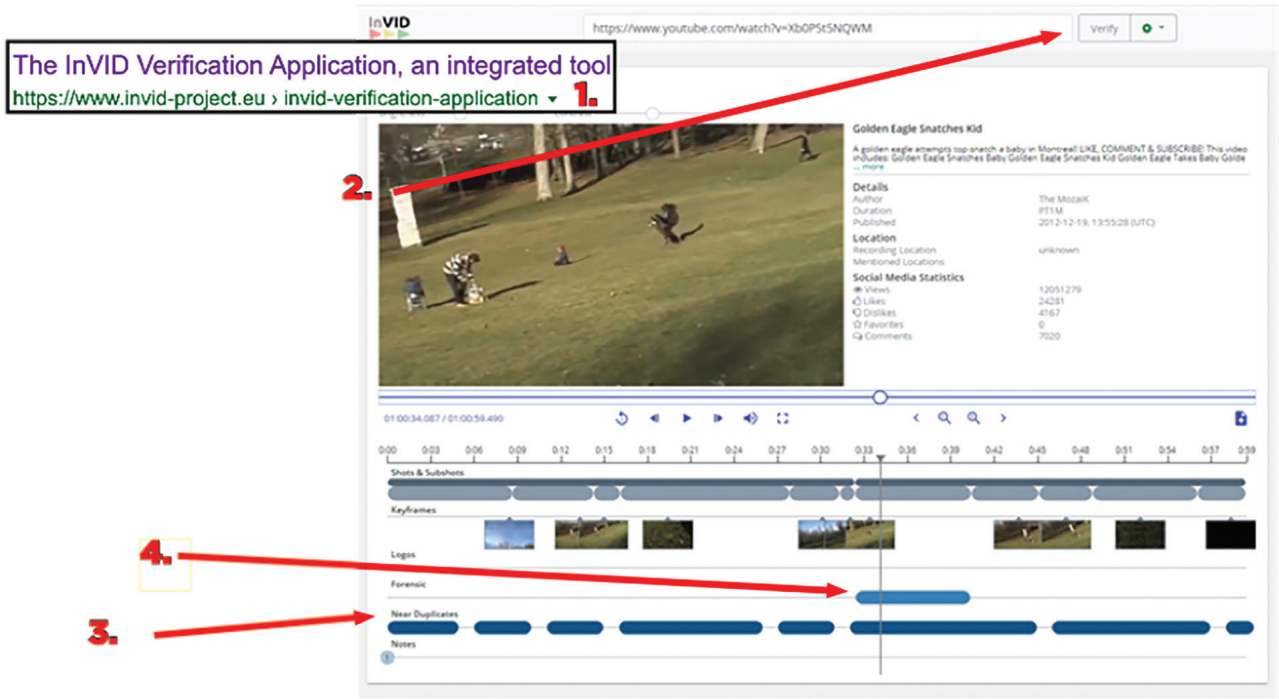
2. Įklijuokite vaizdo įrašo nuorodą.

3. Patikrinkite, ar vaizdo įrašas rodomas tarp dublikatų.



Mokslinė analizė

„InVid“ taip pat turi mokslinės analizės funkciją, kai programinė įranga nustato galimai pakeistus kadrus – jie pasirodo analizės lange šalia ženklo „Forensic“. Jei „InVid“ nustato kadrus kaip galimai pakeistus, tikimybė, kad vaizdo įrašas yra melagingas, išauga.



Pagrindinės taisyklės, kurias verta įsidėmėti:

• Kaip ir su statiniais vaizdais, vaizdo įrašų pernaudojimas yra pagrindinė dezinformatorių taktika. Vaizdo įrašai, kurie buvo nufilmuoti anksčiau, yra paskelbiami iš naujo ir pateikiami kaip melagingos šiandienos naujienos.

• Šios priemonės nėra tokios veiksmingos kaip statinių vaizdų patikrinimo priemonės, todėl verta išbandyti jas abi – “Amnesty DataViewer” ir “InVid” pateikia šiek tiek skirtingus rezultatus.

Naudingos priemonės:

InVid

Amnesty International YouTube Viewer

Troliai

Kaip atpažinti trolį internete?

Kas yra trolis?

Trolis yra asmuo, kuris tyčia inicijuoja konfliktą internete arba įžeidžia kitus naudotojus, taip siekdamas išprovokuoti emocinį atsaką ir/ar nutraukti diskusijas. Jo tikslas išblaškyti ir sėti susiskaldymą, skelbiant kurstančius ar su tema nesusijusius įrašus internetinėje bendruomenėje arba socialiniame tinkle. Esminis trolio skirtumas nuo boto yra toks, kad botai yra automatizuoti, o trolis yra tikras asmuo. Atpažinti trolius yra sunkiau nei botus, nes šios paskyros paprastai yra sudėtingesnės ir aktyviai apsimeta tikrais žmonėmis. Prieš pradedant nagrinėti veiksnius, kurie padeda patikimai atpažinti Kremliui palankų trolį, svarbu atkreipti dėmesį į vieną veiksnį, kuris trolio neapibūdina. Realūs šių laikų socialinių tinklų naudotojai yra labai šališki, ypač kai kalbama politinėmis temomis, o troliai – ne.

Toliau pateikiame keletą kriterijų, kurie padės atpažinti trolius, tačiau perspėjame – šios užuominos yra orientacinės, o ne galutinės. Retai kada galima šimtu procentų patvirtinti, kad tam tikra paskyra priklauso troliui, o ne neigiamai nusiteikusiam vartotojui.

1. Klaidos tekstuose: anglų kalbos artikeliai „a“ ir „the“

Daugeliui žinomų rusiškų trolių paskyrų yra būdinga nemokėti tinkamai naudoti gramatinių artikelių „a“ ir „the“. Rusų kalboje nėra nei vieno iš šių kalbinių ženklų.

2. Klaidos formuojant klausimą

Troliai nesugeba suformuluoti gramatiškai taisyklingo klausimo. Skirtingai nei anglų, vokiečių ar prancūzų kalboje, rusų kalboje žodžių tvarka sudarant klausimus nesikeičia. Daugelis žinomų trolių paskyrų užduoda klausimus, kuriuose žodžių tvarka tokia pati kaip ir paprastuose sakiniuose.

3. Neaiški arba abejotina tapatybė

Kai kurie troliai pasivadina netikrais vardais, kurie yra labai paplitę tam tikroje kalboje. Dėl to gali būti sunku atskirti konkretų autorių arba lengva sąmoningai supainioti jį su kitu žmogumi, pavyzdžiui, žinomu žurnalistu. Troliai taip pat siekia, kad jų vardai būtų suvokiami kaip tradiciniai arba „teisingi". Tokiu būdu kiekvienas skaitytojas būtų linkęs jais pasitikėti ar jam nekiltų abejonių dėl straipsnio ar komentaro socialiniuose tinkluose autoriaus tapatybės.

Taip pat, gali būti naudinga patikrinti trolių profilio nuotraukas, jei tokių yra. Nuotraukos įkeliamos siekiant sukurti didesnį pasitikėjimą. Jos gali būti paimtos iš nuotraukų duombazių, todėl jas galite lengvai rasti internete. Tokios nuotraukos taip pat gali būti sąmoningai redaguotos (pavyzdžiui, tariamas asmuo dėvi akinius nuo saulės ir pan.), kad būtų sudėtingiau identifikuoti asmenį.

4. Prokremliškų pasakojimų platinimas

Rusijos vyriausybė sukūrė savitą pasakojimą apie pagrindinius pastarųjų penkerių metų geopolitinius įvykius. Tai grindžiama principais, įtvirtintais Rusijos Federacijos informacijos saugumo doktrinoje „Dėl valstybės politikos ir oficialios pozicijos Rusijos vyriausybei svarbiais klausimais perteikimo" (2000 m.). Kadangi Kremliui palankūs pasakojimai yra plačiai prieinami internetiniuose šaltiniuose, pavyzdžiui, Rusijos užsienio reikalų ministerijos arba RT „X" paskyrose, nesunku patikrinti, ar tos pačios temos atsikartoja įtariamoje paskyroje. Paskyra, kurioje nuolat dalijamasi Rusijos vyriausybės skelbiama informacija, gali būti pagrįstai laikoma Kremliui palankių pažiūrų.

Jei tarp paskyrai priklausančių įrašų ar pasidalijimų vyrauja Kremliaus pasakojimai, yra daromos būdingos kalbinės klaidos, o žmogus apsimeta amerikiečių ar britų kilmės naudotoju, tai gali būti Rusijos valdomas trolis.

Kitos galimos užuominos:

Troliai turi vienkartinius el. pašto adresus

Kadangi daugelyje puslapių ar tinklaraščių, kuriose leidžiama komentuoti straipsnius, reikia įvesti ir el. pašto adresą, troliai apeina šį punktą įvesdami išgalvotus. Dauguma vienkartinių el. pašto adresų yra atsitiktiniai ir lengvai pastebimi, nes jie neatspindi tikrojo asmens vardo.

Trolių tikslas yra sukelti žmonių pasipiktinimą

Jie nėra mandagūs ir nesigėdija veltis į atvirą kovą. Jie mėtosi kaltinimais ir dažniausiai skamba piktai.

Troliai naudojasi anonimiškais tarpiniais serveriais

Troliai dažnai naudojasi anoniminėmis programomis arba tarpiniais serveriais, kurie rodo kitokį interneto protokolo (IP) adresą nei esate įpratę.

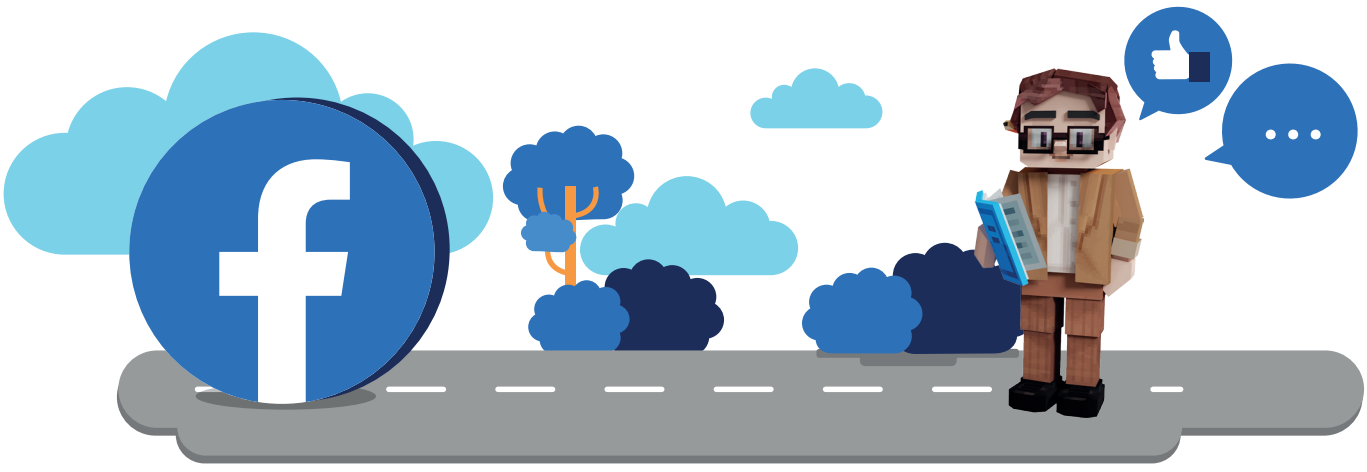
Troliai retai pasako ką nors vertingo viso pokalbio metu

Kai troliai įsiterpia į bendruomenės diskusiją, jie jai neprideda nieko prasmingo. Vietoje to, jie juokauja, priekaištauja ir įžeidinėja.

Netikros „Facebook“ paskyros

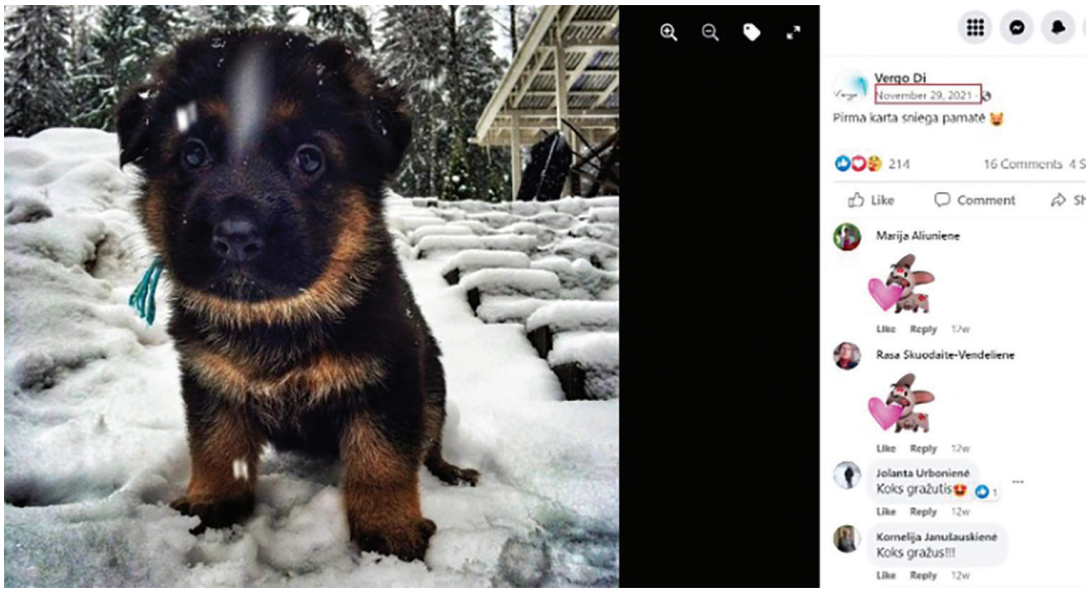
Kaip pastebėti, kad „Facebook“ paskyra yra netikra?

Netikros socialinių tinklų paskyros nėra tokios aktyvios kaip troliai ir dažniausiai atlieka tylių žiūrovų vaidmenį. Panašūs kriterijai netikroms paskyroms yra taikomi daugumoje platformų, tačiau šįkart „Facebook“ pasirinkome kaip pagrindinį pavyzdį. Tikri „Facebook“ vartotojai savo paskyrose dažnai dalijasi asmenine informacija, todėl šios paskyros aktyviai bando tapti jūsų draugais dėl dviejų pagrindinių priežasčių: siekdamos atrodyti tikresnės, nes draugų sąraše turi daugybę tikrų žmonių, ir būti draugais tam, kad pamatytų daugiau asmeninės informacijos. Priklausomai nuo netikros paskyros tikslų, ji gali būti naudojama renkant asmeninę įvairių organizacijų darbuotojų informaciją.

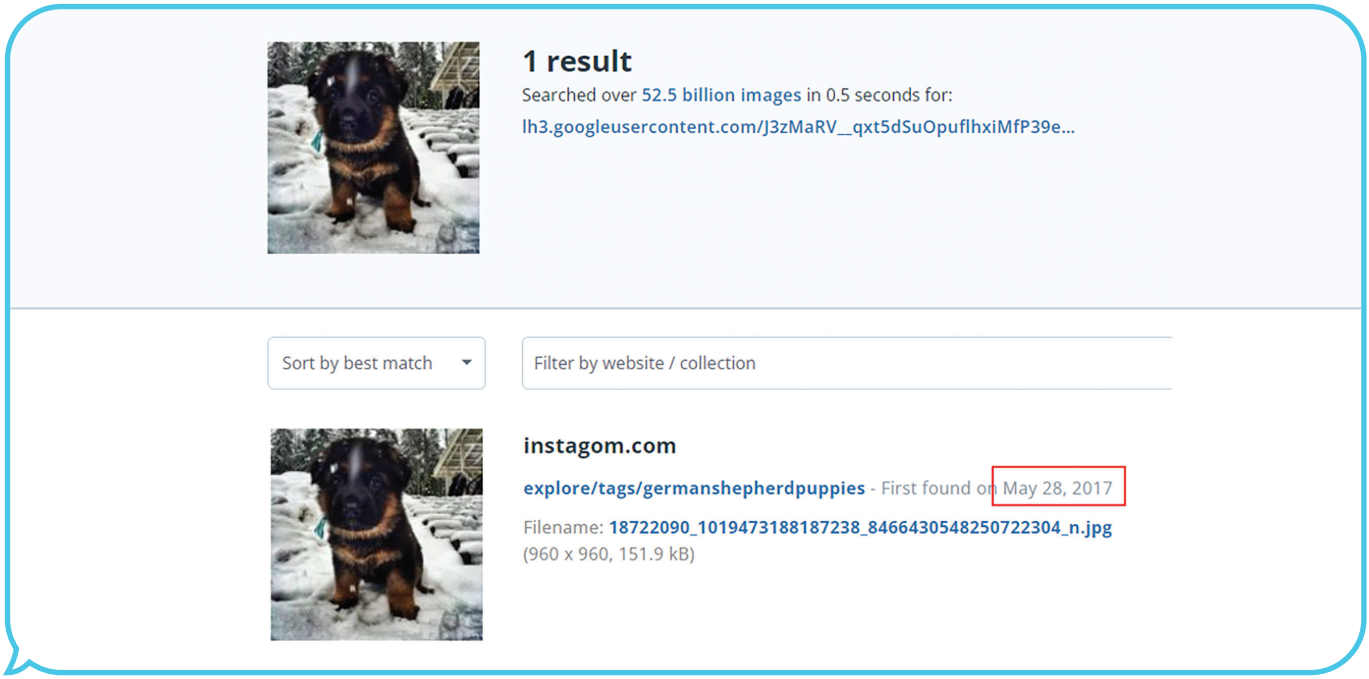


1. Patrauklumo veiksnys

Nepažįstamų vizualiai patrauklių vartotojų paskyros, kurios kviečia jus tapti draugais, gali būti netikros.



Galime nesunkiai patikrinti, ar tai nėra iš interneto nukopijuotas vaizdas, kurį netikros paskyros savininkas pasirinko siekdamas pritraukti jūsų dėmesį ir priimti į draugus.



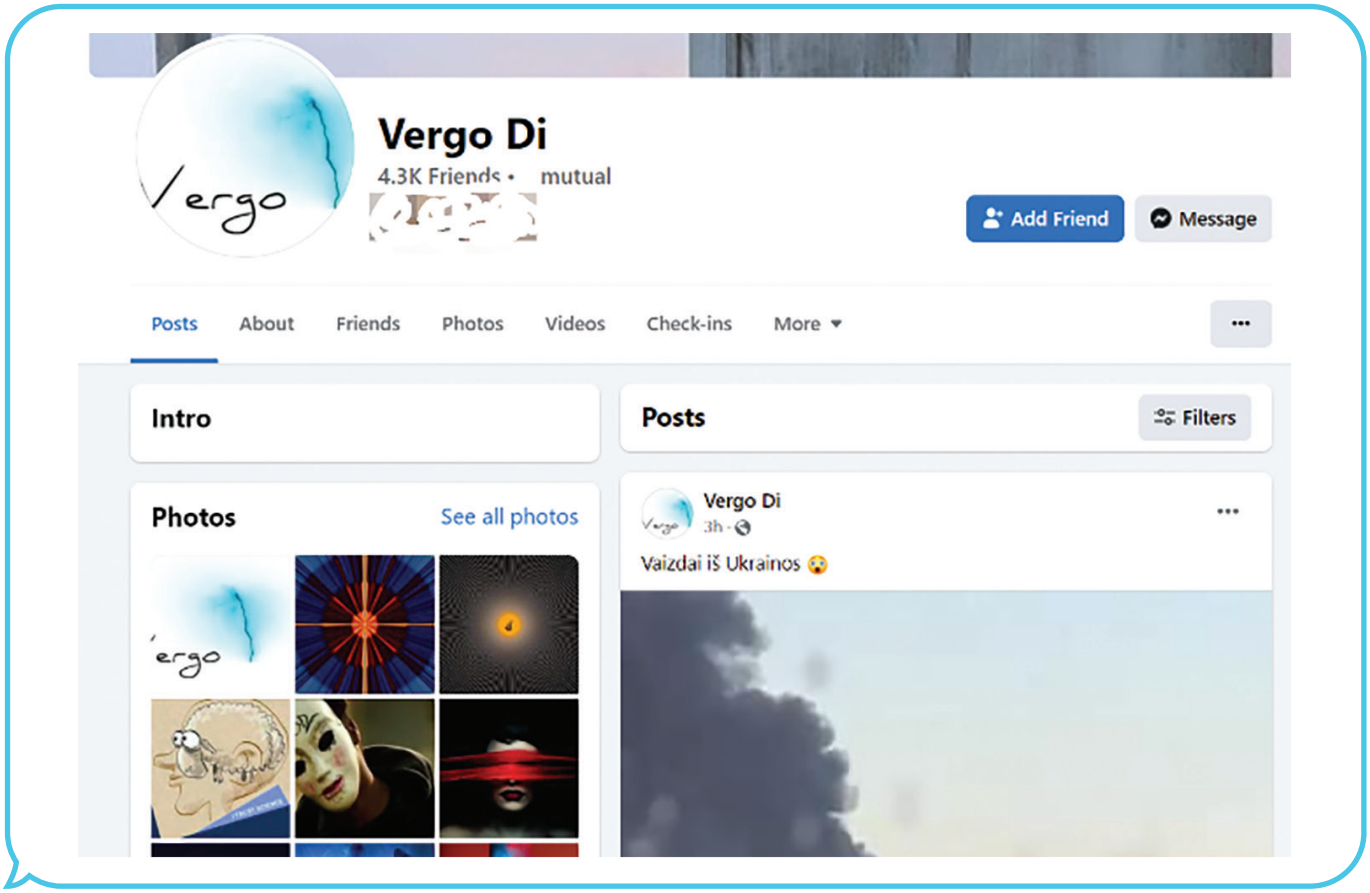
https://tineye.com/ paveiksliukų paieškoje randame datą, kada nuotrauka pasidalinta pirmąkart. Galime daryti išvadą, kad profilis yra netikra, auditoriją sutraukusi anketa, kurioje dalinamasi netikrais įrašais, o retkarčiais – ir propagandinėmis žinutėmis.

2. Nedaug įkeltų nuotraukų

Dauguma netikrų paskyrų neskelbia daug nuotraukų – dažniausiai tris arba keturias, kartais tai yra skirtingų žmonių nuotraukos. Tiek nuotraukų pakanka siekiant sukurti laikiną iliuziją, kad už paskyros slypi tikras žmogus.

3. Keistos biografijos

Daugumos netikrų paskyrų biografijose yra labai mažai informacijos arba pateikta informacija atrodo keista. Pavyzdžiui, tai nėra neįmanoma, bet labai mažai tikėtina, kad žmogus, gyvenantis Bronkse, lankė Helsinkio universitetą, taip pat yra labai jaunas, o jau dirba Niujorko viešųjų ryšių įmonėje. Greitai patikrinę jo vardą „Google“ paieškoje ir atlikę atvirkštinę jo profilio nuotraukos paiešką galite įsitikinti, kad paskyra yra netikra.



4. Nereagavimas į žinutes

Jei parašysite netikrai paskyrai žinutę, mažai tikėtina, kad sulauksite atsakymo net ir į trumpą klausimą. Geriausia net nebandyti susisiekti ar kitais būdais užmegzti pokalbį.

5. Dažniausiai tuščia „Facebook“ siena

Vieninteliai dalykai, kuriuos rasite ant vienos iš šių netikrų „Facebook“ paskyrų sienų, yra nauji paspaudimai „patinka“ „Facebook“ įmonių arba produktų puslapiuose ir nauji draugai.

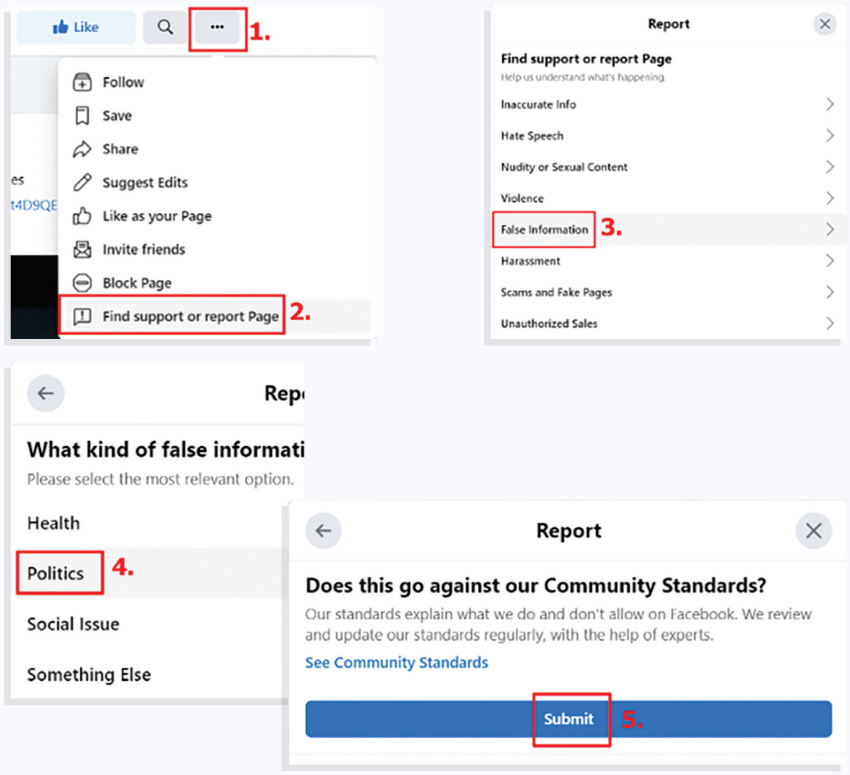
Reagavimas

Norint aktyviai kovoti su dezinformacija internete, būtina atlikti du veiksmus – ją atskleisti ir apie ją pranešti.

„Facebook“ / „X“ / žiniasklaidos pranešimai

Labai svarbu, kad jūsų draugai, bendraklasiai ar kolegos žinotų apie skleidžiamą dezinformaciją, nukreiptą prieš jūsų organizaciją. Kiekviena organizacija privalo turėti aiškią tvarką, kuri užtikrintų, jog jos nariai žino, kur siųsti pranešimą apie pastebėtą netikrą istoriją. Pagrindinis to tikslas yra vieningai informuoti organizacijos narius, kad tam tikras skleidžiamas pranešimas yra netikras, ir užkirsti jiems kelią dalytis istorija bei ja tikėti.

Antras žingsnis – pranešti apie tai socialinių tinklų platformoje. Visos socialinių tinklų platformos turi pranešimo apie naujieną ar bet kokio formato įrašą funkciją, nurodant konkrečią priežastį, kodėl apie tai yra pranešama. Jei socialinio tinklo platforma gaus pakankamai naudotojų pranešimų, sukurta istorija arba netikras pranešimas bus panaikinti. Tai yra metodas, kurį pilietinės organizacijos naudoja kovai su dezinformacija internete. Jei žiniasklaidos priemonės skleidžia netikras istorijas, atsižvelgiant į žiniasklaidos priemonių pobūdį, apie tai reikėtų pranešti arba joms pačioms arba nacionalinėms žiniasklaidos kontrolės institucijoms. Pavyzdžiui, žurnalistų etikos inspektoriui, kuriam galima nusiųsti e.skundą: https://www.zeit.lt/lt/e.skundas/560?fbclid=IwAR3d1M6DGE-noolxRLHLqGLP7v_oNVb_tEIlC2WhzMuI2piC7WE-46dwYtj-



„Telegram“ naudojimo rizikos ir pavojai

Karas Ukrainoje stipriai išpopuliarino socialinį tinklą „Telegram“. Jis yra patrauklus vartotojams, kadangi jame neveikia algoritmai, kurie galėtų daryti įtaką rodomam turiniui – klientai patys renkasi, ką žiūrėti ir platinti. Džiaugtis tikrai nereikia, kadangi čia gali slypėti didžiausi pavojai – esant didelei laisvei, atsiranda rizika gauti ir kartu skleisti dezinformaciją. Šis socialinis kanalas labai patrauklus norint kurti privačias grupes, kurios tampa dar viena terpe dezinformacijai skleisti.

Kadangi daugelis jaunų žmonių naudoja „Telegram“ kasdieniams susirašinėjimams, jie taip pat žino apie kitas šios platformos funkcijas – grupes ir kanalus, kurie iš jų gali būti privatūs ir anoniminiai. Nors atviri komunikacijos srautai yra įprasti socialinėje žiniasklaidoje, uždarose bendruomenėse gali būti platinamas specifinis turinys, įskaitant šališką ir iškraipytą informaciją, kurios kilmę labai sunku patikrinti.

Atlikti tyrimai atskleidžia, jog dalyje dezinformacijos kampanijų, klastotėms platinti buvo naudojami „Telegram“ kanalai ir grupės. Pavyzdžiui, tarp skleidžiamų žinučių įvairiomis kalbomis netrūko informacijos apie COVID-19 pandemiją, taip pat Kremliui palankių pažiūrų apie karą prieš Ukrainą, kraštutinių dešiniųjų retorikos ir sąmokslo teorijų.

„Telegram“ kanalai ir grupės taip pat gali būti naudojami siekiant sutelkti vartotojus į ideologinius susibūrimus ar politinius protestus. Jei šie renginiai organizuojami demokratiškai ir skaidriai, dėl jų abejonių ar problemų nekyla, tačiau kartais tikrieji naudos gavėjai yra slepiami, o informacijos apie šiuos realius asmenis yra gana mažai arba jos nėra visai. Pavyzdžiui, kelios anoniminės „Telegram“ grupės ir kanalai buvo panaudoti provokuojantiems skambučiams skleisti Estijoje vykusios kampanijos „Aš – rusas“ metu (nuoroda: https://eng.obozrevatel.com/ section-life/news-russians-in-tallinn-threw-a-hysterical-tantrum-because-the-police-forced-themto-remove-i-am-russian-stickers-from-their-carsvideo-27-09-2023.html)

Kadangi daugelis dezinformacijos platintojų ir propagandistų (taip pat ir Kremliui palankių) naudojasi „Telegram“, jų turiniu lengva dalytis kitoms grupėms ir bendruomenėms, kurioms vadovauja anoniminiai administratoriai. Tai vienas iš labiausiai paplitusių kanalų, kuriame skaitmeniniu būdu skleidžiama antivakarietiška ar antiliberali dezinformacija.

Pavyzdžiui, „Telegram“ kanalas „Антифашисты Прибалтики“ (Antifašistinis Pabaltijis), galima pastebėti, yra atsakingas už „rusofobijos“ naratyvo peržiūrų skaičiaus išpūtimą – jo įrašai surenka net po kelis šimtus tūkstančių peržiūrų.

Viename iš kanale patalpintų įrašų buvo kalbama apie tariamą Latvijos rusofobiją, kadangi viename lėlių teatre nuspręsta uždrausti rodyti „čeburaškas“ (liet. Kulverstukas, sovietinių filmukų personažas). Šiuo įrašu taip pat buvo kurstoma neapykanta prieš Latvijos kultūros ministrą, įkeliant (netikrą) jo nuotrauką, kurioje jis pozuoja prie groteskiškų objektų, teigiant, kad „tai Latvijos kultūros ir tautinės tapatybės veidas“.

Kitame įraše buvo teigiama, kad bet kuris „laisvų Vakarų“ žmogus, išdrįsęs net užsiminti apie Rusijos teisę ginti rusus, akimirksniu atsiduria už grotų su konfiskuotu turtu ir draudimu vykdyti bet kokią ūkinę ir kūrybinę veiklą. Šie įrašai rodo, kad Rusija aktyviai bandė skleisti naratyvą, kad Baltijos šalyse yra puolama rusų kultūra, o rusakalbių mažuma šalyse negali išreikšti savo nuomonės, bijodama represijų.

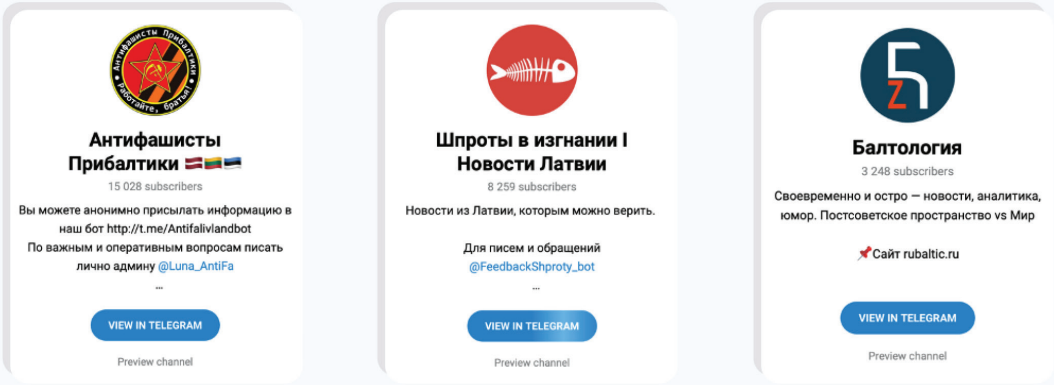
Keletas paprastų patarimų, kurie gali padėti sustiprinti jūsų skaitmeninę higieną „Telegram“ kanale:

• Prieš prisijungdami prie grupės ar kanalo įsitikinkite, kad turinys 100% jus domina, ieškokite ir teiraukitės daugiau informacijos apie internetinės bendruomenės administratorių.

• Būkite budrūs – tikimybė anoniminėse grupėse ir kanaluose gauti nepatikrintą ir/ar iš neaiškių šaltinių atėjusią informaciją, jei tema yra aktuali ir svarbi, be galo didelė. Turinys turi turėti patikimus šaltinius ir negali pasižymėti spekuliatyviomis nuomonėmis, „alternatyviais faktais“ ar supaprastintomis propagandinėmis klišėmis.

• Jei kokia nors informacija „Telegram“ grupėje ar kanale jus emociškai išprovokavo, paklauskite savęs, kodėl taip atsitiko ir kam tai naudinga – neskubėkite reaguoti ar dalintis informacija, nes ji gali būti šališka, poliarizuojanti, įžeidžianti ar tiesiog netikra.

• Apie bet kokį įtarimą keliantį įrašą „Telegram“ grupėje ar kanale galima pranešti administratoriams, žiniatinklio policijai, taip pat skaitmeninei faktų tikrinimo bendruomenei (pavyzdžiui, „CRI“ Lietuvoje ar „Propastop“ Estijoje). Prieš tai atlikdami įsitikinkite, jog išsaugojote kiek įmanoma originalių įrašų, pavyzdžiui, ekrano kopiją su tekstine ar vaizdine informacija.



„Tik Tok“ naudojimo rizikos ir pavojai

Jaunimo tarpe neseniai išpopuliarėjęs kiniškos kilmės „TikTok“ pasižymi dideliu kiekiu tinkle platinamos dezinformacijos. Patys programėlės kūrėjai akcentuoja, jog deda dideles pastangas kovai su dezinformacija, radikaliu ekstremizmu ir neapykantą kurstančiu elgesiu, tačiau kaip jiems iš tikrųjų sekasi?

Nors iš pat pradžių šis socialinis tinklas neatrodė pavojingas, ilgainiui daugėjant vartotojų, ėmė keistis ir turinys. Dabar tinkle netrūksta įrašų, kuriuose skleidžiami Kremliui palankūs **naratyvai***.

***Naratyvas** tai sistemiškai ir nuosekliai formuojamas pasakojimas, kuris sukuriamas kartojant žinutes tam tikra tema, papildant šias žinutes naujais faktais, kontekstu.

Naratyvas – tai pasakojimas, įtikinamai perduodantis pagrindinę žinią, fomuojantis nuomonę.

Platforma „TikTok“ pasižymi labai neaiškiais algoritmais ir atsižvelgimu į autoritarinių režimų įstatymus. Šis socialinis tinklas taip pat išsiskiria iš kitų tuo, jog turi priklausomybės efektą. Jis pasireiškia nuolatiniu trumpų vaizdo įrašų, kurie turi daug emocinių elementų ir įsimintiną muzikinį takelį, žiūrėjimu. Kuo daugiau laiko praleidžiama „TikTok“, tuo geriau pradeda veikti algoritminis informacijos pateikimas, kada bendrame vartojamo turinio sraute atsiranda propagandinės žinutės.

Galima išskirti, jog yra du esminiai iššūkiai, susiję su „TikTok“ naudojimu:

• **Agresyvus duomenų rinkimas.** Diegiant programėlę į telefoną ar kitą išmanųjį įrenginį, ji paprašo daugiau prieigos prie duomenų ir vėliau juos renka. Riziką naudojantis „TikTok“ kelia tai, jog programėlei galima matyti vartotojo kontaktus, taip pat, kokios kitos programėlės naudojamos įrenginyje, sužinoti daugiau detalių apie lokaciją ir identifikuoti, kur yra konkretus įrenginys. Pavojus kyla ir dėl to, jog susirašinėjimai, kurie vyksta programėlėje, gali būti stebimi ir pagal tam tikrus raktažodžius ir patekti į pačios kompanijos radarą.

• **Šešėlinis įrašų draudimas** (angl. shadow banning). Jeigu vartotojas paskelbia įrašą, kurio turinys nepatinka „TikTok“ kūrėjams, įrašas gali būti paslėptas, t. y. bus įvykdytas šešėlinis draudimas.

„TikTok“ naudoja algoritmą kaip įrankį dėmesiui pritraukti ir jį išlaikyti bei stengiasi kiekvienam vartotojui suteikti individualizuotą patirtį. Algoritmas naudoja iš vartotojų surinktus duomenis tam, kad nustatytų, koks turinys galėtų juos sudominti. Pavyzdžiui, kuo ilgiau žiūrėsite vaizdo įrašą, tuo daugiau panašių vaizdo įrašų ateityje pamatysite „TikTok“. Ši programėlė taip pat prisimena jūsų paieškų raktažodžius, kad ateityje galėtų pasiūlyti tokio paties stiliaus vaizdo įrašų. Be to, „TikTok“ algoritmas sujungia panašių pomėgių vartotojus rodydamas jiems panašų turinį. Jei matote tuos pačius vaizdo įrašus kaip ir kai kurie jūsų draugai, tai nėra atsitiktinumas.

Atminkite, kad „TikTok“ nebuvo sukurta naujienoms perduoti – trumpi vaizdo įrašai yra labiau pritaikyti vartotojo pramogoms. Slenkant ir žiūrint vaizdo įrašus „TikTok“, suaktyvinamos smegenų dalys, atsakingos už sėkmės pojūtį - kaip ir lošiant azartinius žaidimus su viltimi pakelti nuotaiką, tačiau vietoje tikro pelno sugaištama daug laiko. Kadangi „TikTok“ vaizdo įrašai yra trumpi, linksmi ir lengvai pasiekiami, tai sukuria tam tikrą priklausomybę, dėl kurios prarandamas dėmesys ir laiko nuovoka.

„TikTok“ taip pat išnaudoja jūsų smalsumą ir baimę, praleisti ką nors svarbaus. Ar girdėjote sakinį „Jei nesate „TikTok“, jūsų iš viso nėra“? Tai vadinama manipuliacija, kuri siekia priversti jaunus žmones nuolat palaikyti ryšį.

Kitas „TikTok“ naudojamas triukas – pykčio auginimas. Programa siūlo vaizdo įrašus, kurie įžeidžia žmonių grupę, ideologiją ar judėjimą, tikintis, kad toks turinys įžeis žiūrovus ir pradės emociškai karštas diskusijas, kurios pritraukia daugiau dėmesio ir vaizdo įrašai tampa virusiniai (angl. viral). Šie metodai naudojami norint padidinti paspaudimų ir sekėjų skaičių.

„TikTok“ apstu kibernetinių patyčių atvejų, todėl kyla didelė rizika tapti priekabiavimo ar neapykantos kurstymo aukomis. Be to, „TikTok“ turinį gali kurti bet kas, tad jame neabejotinai bus neobjektyvios arba manipuliuojamos informacijos. Kadangi „TikTok“ yra daug suasmeninto turinio, jauniesiems vartotojams dar sunkiau atskirti tam tikrų vartotojų nuomonę nuo faktų.

„TikTok“ gali būti tinkamas vartojimui socialinis tinklas, jei vadovausitės paprastais jo naudojimo patarimais:

• Apribokite savo kasdienį laiką, kurį praleidžiate „TikTok“ – susitikti su draugais realiame gyvenime visada smagiau nei internete.

• Nepamirškite kritiškai vertinti į naujienas panašų turinį „TikTok“ – dar kartą patikrinkite jus dominančią informaciją ir kituose šaltiniuose (ne socialinės žiniasklaidos priemonėse).

• Praneškite apie kibernetinį priekabiavimą ir neapykantą kurstančią kalbą – neplatinkite įžeidžiančių vaizdo įrašų ar emociškai jautraus turinio.

• Jei jaučiatės apgauti ar manipuliuojami, pasitarkite su tėvais arba e-policija.

Išplėstinės priemonės ir kreipimasis pagalbos

Kalbant apie sunkesnių dezinformacijos atvejų atskleidimą, galimi du pagrindiniai būdai: pasinaudojant sudėtingesniais atvirojo kodo metodais arba kreipiantis pagalbos į internetinę tyrimų bendruomenę.

Dauguma internetinių priemonių gana lengva naudotis ir jose pateikiamos nuoseklios instrukcijos, kaip tai padaryti. Toliau pateikiamos dvi didelės ir naudingiausios priemonės:

Bellingcat's Online Investigation Toolkit

Online Open Source Tool Box

Jeigu klausimų vis dar daugiau negu atsakymų, susisiekite su internetinių tyrimų bendruomene ir pateikite jiems informaciją apie jūsų pastebėtą melagingą istoriją. Dauguma tyrėjų mielai demaskuos netikrą istoriją ir pasidalins ją paneigiančiu turiniu internete.

Vilnius, Lietuva
Tiražas 500

Publikacijos išleidimą
remia:

Google