

Kancelária Rady pre mediálne služby

Pracovný materiál Kancelárie Rady na zasadnutie Rady pre mediálne služby dňa 11. 10. 2023

Predmet: Priebežné výsledky monitorovania funkcionalít platforiem v súvislosti s voľbami do Národnej rady Slovenskej republiky 2023 (monitorované obdobie: 25. 9. – 30. 9. 2023)

Dátum: 4. 10. 2023

NÁVRH UZNESENIA: Rada berie na vedomie predložený materiál.

* * *

Dňa 6.7.2023 Rada pre mediálne služby schválila materiál Monitoring funkcionalít platforiem v súvislosti s voľbami do Národnej rady Slovenskej republiky 2023. Na základe uvedeného materiálu boli v období 25.9. – 30.9.2023 monitorované vybrané platformy na zdieľanie obsahu z pohľadu ich funkcionalít súvisiacich s uverejňovaním informácií o voľbách do Národnej rady Slovenskej republiky dňa 30.9.2023. Monitorované platformy a funkcionality:

- Meta Ads Library – knižnica reklám, subjekt: Meta, platformy: Instagram a Facebook
- Ads transparency center – knižnica reklám, subjekt: Google, platforma: YouTube
- Creative center – knižnica reklám, subjekt: ByteDance, platforma: TikTok
- subjekt: Twitter Ltd., služba: X(Twitter)

Monitoring vybraných funkcionalít

Na základe vykonaného monitoringu v dňoch 25. 9. – 30. 9. 2023 boli zozbierané dáta v tematických okruhoch:

- Umiestňovanie informačných panelov o voľbách počas predvolebnej kampane a v deň konania volieb s funkčným odkazom na relevantný web/podstránku o voľbách MIV SR.
- Progres funkcionalít platforiem, v oblasti transparentnosti politickej reklamy a dostupných funkcionalít platforiem s výsledkami monitoringu z roku 2019

Rada pre mediálne služby (RpMS) v rámci monitorovania volebnej kampane na platformách sociálnych sietí neobdržala žiadny podnet na preverenie od verejnosti. Zároveň však v rámci vlastných monitorovacích kapacít i na základe hlásení od iných inštitúcií štátnej správy RpMS zaznamenala a aktívne s platformami komunikovala v prípade výskytu tzv. deepfake obsahov vytvorených za pomocou využitia technológií AI (artificial intelligence - umelej inteligencie).

Prvé deepfake videá týkajúce sa slovenskej politickej scény, v ktorých boli použité AI technológie, sme zaznamenali v poslednom týždni volebnej kampane t. j. v týždni od 25.9.2023 a tieto sa virálne šírili a cyklicky objavovali na všetkých monitorovaných platformách (Facebook, Instagram, YouTube, TikTok, X/Twitter).

V prípade takto zachytených obsahov boli tieto obsahy zároveň aj súbežne v rámci spravodajského pokrytia volebnej kampane zo strany médií medializované a verejnosť sa teda dozvedela o existencii videí obsahujúcich klamlivé výroky nahovorené hlasmi, ktoré boli pre rádového poslucháča nerozpoznateľné od hlasov skutočných osôb (hlasy ponášajúce sa na hlas prezidentky Slovenskej republiky Zuzany Čaputovej a hlasy ponášajúce sa na hlasy všeobecne verejne známych kandidujúcich politikov a verejne známych akreditovaných novinárov). Charakteristické technické črty týchto videí boli neprirodzená dikcia reči a prízvuk, počuteľné strihy v audiodiostope a pod.. Fakt, že dané videá obsahovali nepravdivé informácie, bol podporený aj následnými vyhláseniami zo strany Polície a nezávislých overovateľov faktov (tzv. 3rd Party Fact Checking), ktorí zmluvne spolupracujú s platformami pri overovaní pravosti publikovaných informácií. Platforma Facebook začala po uverejnení tzv. debunkingu pridávať k takémuto videu tzv. labeling, t.j. informácia pre používateľov, že dané video obsahuje nepravdivé informácie, avšak používatelia aj napriek tejto skutočnosti majú možnosť toto video vzhliadnuť, interagovať (dať tzv. like), komentovať a aj ďalej zdieľať na platforme. Zároveň sa v neskoršej fáze začali objavovať rôzne modifikácie predmetných obsahov (úryvky z videa, pridané titulky do videa, snímky obrazovky - snímky obrazovky - tzv. screenshots z videa s titulkami), pričom na tieto modifikované obsahy nedostala kancelária k dátumu spracovania materiálu od platforiem odpoveď.

Prehľad jednotlivých príspevkov zachytených na platformách Youtube, TikTok a Facebook a X nižšie s doplnením kontextu a s priebehom riešenia týchto obsahov v spolupráci s online platformami. Číselné vyjadrenia (1x a pod.) znamenajú na koľkých rôznych unikátnych adresách (tzv. URLs) sa objavili predmetné obsahy.

25.9.2023

1x

Video vo forme platenej politickej reklamy na účte Milana Uhríka na platforme Instagram obsahovalo hlas, ktorý bol pre bežného poslucháča nerozpoznatelný od hlasu prezidentky Slovenskej republiky Zuzany Čaputovej. Tento hlas v spote uvádzal, že netreba slepo nasledovať progresívne stádo a treba precitnúť a voliť stranu Republika. Hlas ďalej uvádza, že bol vytvorený umelo a môže sa podobať na skutočné hlasy reálnych osôb. Následne vo videu predstavitelia strany Republika Milan Uhrík a Milan Mazurek vyzývajú diváka, aby prišli voliť stranu Republika.

Video spoločne s príslušným kontextom bolo v ten istý deň nahlásené spoločnosti Meta, ktorá prevádzkuje službu Instagram. Dňa 28.9.2023 bola kancelárii doručená odpoveď, že video bolo preskúmané a nezistilo sa žiadne porušenie pravidiel používania služby Instagram a video stále ostáva na službe dostupné.

26.9.2023

2x

Platforma Youtube, v rámci platenej politickej reklamy boli zaznamenané volebné spoty strany Republika. Prvé video bolo totožné s videom, ktoré kancelária zachytila aj v pondelok na platforme Instagram a obsahovalo AI-generovaný hlas podobný hlasu prezidentky Zuzany Čaputovej. Druhé video po formálnej stránke pripomínalo prvé video, avšak hlas v ňom bežnému poslucháčovi pripomínal hlas Michala Šimečku. Hlas vo videu uvádzal, že sa volá Michal a kedysi bol progresívec a veril na sedemdesiat rôznych pohlaví a tehotných mužov, ale nakoniec dostal rozum a bude voliť Republiku. Video opäť obsahovalo oznam o tom, že hlas v ňom bol vytvorený umelo a môže pripomínať hlasy iných osôb a v závere boli rovnako predstavitelia strany Republika Milan Uhrík a Milan Mazurek, ktorí vyzvali diváka voliť stranu Republika.

Obe videá, ktoré boli publikované ako platená politická reklama, spolu s príslušným kontextom nahlásili pracovníci kancelárie Rady pre mediálne služby spoločnosti Google, ktorá prevádzkuje službu YouTube. Zo spoločnosti Google behom pár hodín prišiel oznam, že videá boli zo služby odstránené, nakoľko porušujú pravidlá používania služby pre publikovanie obsahu.

27.9.2023

1x

Video na účte Milana Uhríka na platforme TikTok, to isté už vyššie spomenuté video s hlasom podobným hlasu Zuzany Čaputovej.

Video bolo v ten istý deň nahlásené platforme, dňa 28.9.2023 prišla odpoveď, že video bolo odstránené.

28.9.2023

3x

TikTok video, hlas pripomínajúci hlasu M. Šimečku o zdražovaní piva. Video obsahovalo obrazovú stopu: pohár s pivom. Tieto boli nahlásené platforme a platforma videá následne tieto videá odstránila.

1x

TikTok video, hlas pripomínajúci hlas prezidentky Zuzany Čaputovej, ktorý ako keby znel z rádiového vysielania, ktorý hovorí, že v poslednej dobe počúvala Kulturblog a preto zmenila názor a bude voliť

Milana Mazureka a stranu Republika. Video obsahovalo obrazovú stopu: archívny záber Z. Čaputovej z rozhlasového štúdia v Slovenskom rozhlase.

Video bolo nahlásené platforme, platforma video odstránila.

1x

Platforma TikTok, video obsahujúce hlas pripomínajúci hlas Petra Pellegriniho a k tomu AI generovaná tvár, ktorá má predominantne črty tváre Petra Pellegriniho avšak kombinované s črtami tváre Michala Šimečku. Hlas s náznakom ironického posmešku hovorí o tom, ako treba voliť takéto strany, t.j. strany pod vedením spomínaných politikov, v závere sa zobrazí koláž zostavená z log politických strán PS a Hlas-SD.

Video bolo nahlásené platforme, kancelária čaká na odpoveď, pri spätnej kontrole RpMS zistia, že v čase vypracovania tohto materiálu je video stále dostupné online.

1x

Platforma TikTok, video obsahujúce hlas pripomínajúci hlas Richarda Sulíka a k tomu AI vytvorená tvár, ktorá má predominantne črty tváre Richarda Sulíka avšak kombinované s črtami tváre Michala Šimečku. Hlas s náznakom ironického posmešku hovorí o tom, ako treba voliť takéto strany, t.j. strany pod vedením spomínaných politikov, v závere sa zobrazí koláž zostavená z log politických strán PS a SaS.

Video bolo nahlásené platforme, kancelária čaká na odpoveď, pri spätnej kontrole RpMS zistia, že v čase vypracovania tohto materiálu je video stále dostupné online.

8x

Platforme TikTok, údajná audio nahrávka rozhovoru novinárky Moniky Tódovej s Michalom Šimečkom, predsedom Progresívneho Slovenska.

Jednotlivé URL adresy boli nahlásené platforme TikTok. Platforma následne videá zmazala.

1x

Platforma TikTok, príspevok o tom, že nezáleží na tom, či je nahrávka rozhovoru Michala Šimečka s novinárkou Monikou Tódovou zmanipulovaná, alebo nie. Aj tak údaje to svedčí o tom, že strana Progresívne Slovensko sa bude snažiť zmanipulovať voľby.

Príspevok bol nahlásený platforme. Platforma následne obsah zmazala.

6x

Na platforme Facebook sa rozšírila rovnaká nahrávka rozhovoru Michala Šimečku s novinárkou Monikou Tódovou.

Príspevky boli nahlásené platforme, v rámci odpovedí zo strany spoločnosti Meta bolo možné sledovať určitú nekonzistentnosť: 4 príspevky boli zmazané, 1 (identický s predošlými) bol ponechaný online, pretože neporušuje zásady používania služby, 1 príspevok bol odoslaný na kontrolu nezávislej organizácii pre overovanie faktov (tzv. third party fact-checking organisation - 3PFC). Pri spätnej kontrole kancelária RpMS zistila, že momentálne už je všetkých 6 linkov nedostupných)

29.9.2023

27x

Platforma Facebook, nahrávka rozhovoru novinárky Moniky Tódovej s Michalom Šimečkom.

Príspevky boli nahlásené platforme, odpoveď prišla v rovnaký deň (v 9 prípadoch prišla odpoveď 30.9.)
Sumár odpovedí platformy, spoločnosti Meta:

- 16x odoslané na kontrolu nezávislej organizácii pre overovanie faktov (tzv. third party fact-checking organisation - 3PFC)

- 2x príspevky boli zmazané

- 9x platforma v odpovedi uviedla, že nie je zrejmé porušenie komunitných štandardov používania ich služieb (v angličtine "it isn't clear to us that it violates Facebook's Community Standards or Instagram's Community Guidelines")

Pri spätnej kontrole 2.10.2023 konštatuje kancelária RpMS pre týchto 27 nahlásení stav: 10x obsah stále dostupný online, 10x označenie (tzv. label) o tom, že informácie vo videu sú nepravdivé (video je prekryté daným oznámením, používateľ ho však stále môže prehrávať, zdieľať, komentovať), 7x príspevky odstránené.

1x

Platforma Instagram, príspevok obsahujúci AI generovaný hlas prezidentky Zuzany Čaputovej. Príspevok bol nahlásený platforme, v odpovedi z dňa 30.9. 2023 platforma uviedla, že nie je zrejmé porušenie komunitných štandardov používania ich služieb a príspevok je stále dostupný na platforme.

1x

Platforma X (predtým známa ako Twitter), nahrávka rozhovoru Moniky Tódovej s Michalom Šimečkom. Príspevok bol nahlásený platforme. Ostal bez odpovede platformy. Pri spätnej kontrole pribudol k príspevku oznam na spodnej lište, že dané informácie nie sú pravdivé a zároveň pribudol odkaz na stránku v anglickom jazyku, kde sú popísané pravidlá používania platformy X.

1x

Platforma YouTube, nahrávka údajného rozhovoru Moniky Tódovej s Michalom Šimečkom. Príspevok bol nahlásený platforme. Odpoveď v rovnaký deň potvrdila vymazanie obsahu z platformy.

2x

Platforma TikTok, nahrávka údajného rozhovoru Moniky Tódovej s Michalom Šimečkom. Príspevky nahlásené platforme. Odpoveď v rovnaký deň, platforma potvrdila vymazanie obsahu z platformy.

1x

Platforma TikTok, AI generovaný hlas Michala Šimečku, ktorý hovorí o tom, že bude voliť stranu Republika. Príspevok nahlásený platforme. Odpoveď v rovnaký deň, platforma potvrdila vymazanie obsahu z platformy.

30.10.2023

15x

Platforma Facebook. Z daného počtu 2x údajný rozhovor Moniky Tódovej s Michalom Šimečkom a 13x modifikácia tohto obsahu v zmysle vystrihnutý fragment predmetného videa, alebo snímky obrazovky (tzv. screenshots) z videa s prepisom rozhovoru. Obsahy boli nahlásené platforme dňa 02.10. 2023. V rámci spätnej kontroly kancelária RpMS zistila, že tri videá (2x kompletne video a 1x pridané titulky do videa) boli označené labelom, zvyšné fragmenty videa a screenshots s prepisom zostávajú zatiaľ stále online. 04.10.2023 odpoveď od platformy že všetky dané obsahy boli odoslané na kontrolu nezávislej organizácii pre overovanie faktov (tzv. third party fact-checking organisation - 3PFC).

SUMÁR:

Platforma	počet eskalovaných obsahov	konečný stav (online / labeling / odstránené)		
Instagram	2	2	-	-
Facebook	48	22	13	13
TikTok	19	2	-	17
YouTube	3	-	-	3
X (Twitter)	1	-	1	-

SPOLU:	73	26	14	33

* * *

ZÁVER:

Výsledky a následné zistenia vyplývajúce z vykonaného monitoringu funkcionalít platforiem na zdieľanie obsahu budú obsiahnuté v materiáli, ktorý bude vypracovaný a predložený na zasadnutie Rady podľa schváleného materiálu Monitoring funkcionalít platforiem v súvislosti s voľbami do Národnej rady Slovenskej republiky 2023.