

JAK NEFINANCOVAT DEZINFORMACE NA VIDEO PLATFORMÁCH?

Anna Dohnalová, Kristína Šefčíková, Prague Security Studies Institute
Červenec 2024



Prague Security
Studies Institute

OBSAH

Úvod	3
Metodologie	3
Aktuální situace a výzvy	4
Moderace obsahu	4
(De)monetizace dezinformací	5
Legislativa	5
Doporučení	7
Demonetizace	7
Moderace obsahu	8

Publikace vznikla ve spolupráci s Adapt Institute
a s podporou Open Information Partnership.

ÚVOD

Video obsah se stává stále populárnějším zdrojem informací a jeho objem v informačním prostředí roste. Lákavost audiovizuálního obsahu si uvědomují i tvůrci problematického obsahu, včetně těch, kteří dlouhodobě a systematicky šíří dezinformace. Jejich aktivita na platformách jako YouTube roste a s tím i jejich sledovanost, která dosahuje desetitisíce až sto tisíc sledovatelů. Kromě platformy pro rozšíření dosahu jim YouTube slouží jako monetizační platforma, ať už pomocí reklam, členských poplatků nebo finančních darů.

Cílem této publikace je poskytnout praktická doporučení k tomu, jak omezit financování dezinformací na video platformách. Výzkumníci PSSI identifikovali aktuální stav a slabá místa opatření YouTube, ale také evropské a národní legislativy, která momentálně umožňují dezinformačním aktérům profitovat ze svých aktivit a která by měla být adresována. Zároveň bychom rádi upozornili na často opomíjenou roli inzerentů v demonetizaci dezinformací.

METODOLOGIE

Závěry této publikace jsou v první řadě založeny na konzultacích s experty v oblasti regulace digitálních platform a fact-checkery, kteří s námi sdíleli jejich zkušenost s komunikací a spoluprací s digitálními platformami. To, jak funguje monetizace obsahu na YouTube bylo konzultováno s marketingovým specialistou. Poznatky z konzultací byly doplněny dodatečnou rešerší aktuální evropské a národní legislativy, pravidel komunity YouTube a možností monetizace na videoplatformách.

V rámci mapování českého dezinformačního ekosystému na YouTube jsme pracovali se vzorkem kanálů, které dlouhodobě a systematicky šíří zavádějící informace, dezinformace a proruskou propagandu. Tento vzorek je založen na stávajících seznamech zdrojů problematického obsahu, u kterých panuje všeobecná expertní shoda.¹ Tyto seznamy jsme aktualizovali o další momentálně populární kanály, které dlouhodobě šíří neověřené a zavádějící informace a zpravidla podporují ruskou propagandu, především o válce na Ukrajině.

¹ Seznam [konspiratori.sk](https://www.konspiratori.sk), seznamy z předchozího výzkumu PSSI zaměřeného na dezinformační aktéry působící na různých [platformách](#) včetně Facebooku, [Telegramu](#) a dezinformačních [webů](#).

AKTUÁLNÍ SITUACE A VÝZVY

Moderace obsahu

YouTube patří k hlavním platformám pro šíření online dezinformací a umožňuje, aby jeho platformu využívali škodliví aktéři k manipulaci a zneužívání druhých a k organizování a získávání finančních prostředků. Napomáhá tomu zejména algoritmus platformy, který funguje primárně na takzvaném principu doporučení. Pokud se uživatel podívá na video kanálu, který šíří manipulativní obsah, YouTube mu automaticky začne doporučovat další podobná videa, čímž posiluje takzvaný efekt bubliny, kdy časo uživatele utvrzuje v jeho názorech. Oficiální a spolehlivé zdroje informací, dodržující novinářské a etické standardy, nejsou zvlášť označovány a často upadají do pozadí.

V seznamu problematických kanálů, které shromáždili výzkumníci PSSI, se primárně objevuje proruský obsah, kde jejich vlastníci sdílejí hlavně ruskou propagandu. To vypadá například tak, že vysílají projev ruského prezidenta Vladimira Putina a překládají je do češtiny. Jsou mezi nimi ale také takzvané pseudo zpravodajské kanály, které vytvářejí svůj vlastní dezinformační obsah nebo překrucují zprávy ze světa ve prospěch válčícího Ruska. Počet sledujících se pohybuje v rozmezí desítek až stovek tisíc.

YouTube ve svých pravidlech komunity uvádí, že jsou na platformě zakázány „určité typy zavádějícího nebo klamavého obsahu s vážným rizikem vzniku závažné škody“, mezi což řadí například propagaci škodlivých léků nebo léčebných postupů či zasahování do voleb.

Uživatelé mají možnost takový obsah nahlásit, videa s nežádoucím obsahem platforma následně vyhodnotí a odstraňuje, avšak o tom, zda obsah spadá do takové kategorie, si rozhoduje sama, nemá to na starost nezávislý subjekt. YouTube se navíc při moderaci obsahu řídí metrikou času (minut) stráveného uživateli na problematickém obsahu. Tím se často dostává do konfliktu s otázkou, jestli se jedná o populární obsah, který YouTube odstranit nechce, protože na něm uživatelé tráví čas a z toho má platforma užitek.

V rámci tohoto procesu YouTube využívá externí lidské hodnotitele a odborníky, aby zhodnotili, jestli obsah propaguje nepodložené konspirační teorie nebo nepřesné informace. Tento přístup však představuje další problém. K tomu, aby lidé dokázali jednoznačně identifikovat manipulativní a nepravdivý obsah, potřebují jasný soubor faktů. U dezinformací v době pandemie COVIDU-19 se YouTube spoléhal na konsenzus odborníků ze zdravotnických organizací, jako jsou CDC a WHO, avšak ve většině ostatních případů jsou dezinformace méně jednoznačné a nastává problém v tom, že je YouTube nedokáže definovat. Pokud YouTube ale obsah smaže, dezinformačním aktérům to často nezabrání v tom, aby si vlastníci založil kanál další, kde bude šířit podobný či stejný obsah. Každopádně případ pandemie ukázal, že prioritizovat důvěryhodný obsah je možné, při dostatečném tlaku ze strany tvůrců politik. V tomto případě na začátku pandemie proběhly vůči platformám ze strany Evropské komise a eurokomisařky pro hodnoty a transparentnost Věry Jourové intervence za účelem prioritizování důvěryhodného obsahu.

K řešení problematického obsahu má YouTube i Program prioritních nahlašovatelů. Tento program má schváleným organizacím poskytovat prioritní komunikační kanál s YouTube pro nahlašování problematického obsahu. Členové programu však konstatují minimální reaktivnost platformy a z toho plynoucí demotivaci využívat status prioritního nahlašovatele.² Tento program tak představuje teoreticky užitečný, ale prozatím vysoce dysfunkční koncept, který zatím existuje sepiše jako formalita.

² Zdroj: Konzultace s organizací, která je členem Programu prioritních nahlašovatelů.

Region střední Evropy je ze strany platform také dlouhodobě opomíjen co se týče spolupráce s místními moderátory obsahu a fact-checkerů. Jejich počty (v poměru k velikosti země) jsou dlouhodobě nedostatečné, například ve srovnání se zeměmi západní Evropy a čelí nezvládatelnému množství obsahu.

(De)monetizace dezinformací

Dezinformační aktéři si mohou na YouTube vydělat pomocí reklam, členských příspěvků nebo finančních darů. Co se týče příjmu z reklamy, tvůrcům video obsahu jde až 60 % příjmu z reklamy, zbytek si nechává platforma. Je však důležité podotknout, že ne každý tvůrce video obsahu má z reklam příjem. I v případech kanálů, které nejsou součástí partnerského programu YouTube a tudíž neprofitují z reklam, platforma k jejich videím reklamy umísťuje. Celý zisk tak jde ve prospěch YouTube. Z tohoto důvodu je například obtížné odhadovat příjem YouTube kanálů. Pomocí partnerského programu AdSense je navíc možné umístit video reklamu i na externí web a dále si tak zvýšit příjem. Když totiž uživatel klikne na takový web, automaticky spustí také video reklamu.

Součástí problému je i dlouhodobě opomíjený potenciál inzerentů přispět k zastavení financování dezinformací. V první řadě si inzerenti často neuvědomují, že jejich reklamy se mohou prostřednictvím automatizovaných systémů objevit u dezinformačního a konspiračního obsahu. Přitom to představuje reputační riziko pro jejich značku. Tímto způsobem také nepřímě, a často nevědomě, svým reklamním rozpočtem podporují tvůrce problematického obsahu.

K této skutečnosti přispívá často nízké povědomí inzerentů o informačním prostředí a nízká schopnost rozlišovat kvalitní a problematické zdroje informací.³ Nedílnou součástí problému jsou i zdroje a čas. Větší inzerenti v podobě velkých firem mají dedikované zaměstnance, kteří mohou věnovat pozornost zodpovědnému umístování reklam, případně mají možnost investovat do softwarů, které jim pomáhají situaci monitorovat. Velkou mezeru tak představuje zástup malých inzerentů, kteří reklamy na YouTube využívají, ale nemají čas a zdroje zajímat se o etickou dimenzi umístování reklam, aktivně se tomu věnovat nebo platit drahé kontrolní softwary. YouTube přitom poskytuje jednoduché možnosti regulace umístování reklam, ale často zůstávají nevyužité.

Jisté zlepšení by mohlo nastat s rozvojem umělé inteligence (AI) a její schopnosti rozeznávat problematický obsah, čímž by se mohla stát stále dostupnějším nástrojem i pro malé inzerenty. AI může fungovat jako podpůrný nástroj pro prvotní vytipování problematického obsahu, nemělo by však plně nahradit moderátory obsahu se znalostí místního jazyka a kontextu.

Legislativa

Evropská komise v únoru 2024 zavedla do praxe nové nařízení, takzvaný Akt o digitálních službách (DSA), který požaduje po provozovatelích sítí, včetně YouTube, aby zveřejňovali, jak s dezinformacemi a propagandou nakládají. Při porušení nařízení hrozí takzvaným velmi velkým online platformám a vyhledávačům (VLOPs) pokuta až do šesti procent jejich ročního obrátu, což v praxi znamená, že se pro provozovatele

³ Zdroj: Diskuse s představiteli firem na seminářích o brand safety, které PSSI zorganizovalo společně s obchodními komorami aktivními v České republice.

z dosavadních doporučení EU staly závazky. Platformy, které by i přes to trvale odmítaly dodržovat předpisy, EU může dočasně suspendovat.

Doposud prostředí YouTube fungovalo na principu dobrovolné seberegulace, teď už je řešení propagandy a dezinformací povinné. Změna konkrétně nastala ve statusu takzvaného Kodexu zásad boje proti dezinformacím, který byl před zavedením DSA dobrovolný, případně se platformy mohly zavázat jen k některým částem Kodexu. S příchodem DSA se Kodex stal povinným, případně platformy mohou přijít s vlastním kodexem. Adresování problému dezinformací se každopádně stalo povinností.

Po změně legislativy dlouho volali experti, její vymahatelnost je však i přes opakované napomínání platformem velmi složitá. Podle EU totiž zodpovědnost za obsah přebírají platformy, provozovatelé ale větší váhu přikládají samotným uživatelům, kteří by podle nich měli manipulativní obsah aktivně nahlašovat. V České republice navíc implementace DSA zaostává, neboť ještě nebyl přijat implementační Zákon o digitální ekonomice, který by například formálně předal pravomoci národnímu koordinátorovi digitálních služeb, kterým má být na základě stávající dohody Český telekomunikační úřad.

Platformy se na přijetí DSA nicméně připravovaly už dlouho dopředu, například zavedením zmíněného Programu prioritních nahlašovatelů. Přesto však v souladu s novou směrnicí provedla společnost YouTube jen minimální změny. Provozovatelé například rozšířili transparentnost reklam, čímž chtějí uniijním orgánům lépe poskytnout informace o cílení reklam zobrazovaných na území Evropské unie. Zároveň v posledních dnech spustili experimentální pilotní program, který má uživatelům poskytovat lepší kontext a řešit problém manipulativního obsahu na platformě.⁴ Nová funkce Poznámek může pomoci například objasnit, zda se u videí jedná o parodii, nebo dát divákům vědět, když jsou starší záběry mylně prezentovány jako aktuální.

⁴ Zatím v USA a Spojeném království.

DOPORUČENÍ

Demonetizace

Zapojit inzerenty do demonetizace dezinformací: Pro účely omezení financování dezinformací je důležité zapojit samotné inzerenty a budovat standard odpovědného umísťování reklam, které se bude vyhýbat dezinformačnímu obsahu. YouTube k tomu poskytuje jednoduché nástroje, které ovšem často zůstávají nevyužívané. Nejjednodušší cestou pro inzerenty je využívat možnost vyloučení konkrétních YouTube kanálů při umísťování reklam (takzvaný blacklisting). Funguje na principu zadání ID čísel konkrétních kanálů, na kterých inzerent svou reklamu nechce. Blacklisting je obzvláště praktické řešení pro inzerenty, kteří mají nízké povědomí o dezinformačním ekosystému.

Aktivně podporovat inzerenty v ochraně jejich značky: Občanská společnost se už teď snaží aktivně zvyšovat povědomí o možnostech inzerentů a podporovat je v ochraně značky před dezinformačním obsahem. PSSI vytvořilo seznam YouTube kanálů s problematickým obsahem, na kterých nedoporučuje inzerovat, a nabízí firmám pomoc s nastavením jejich reklamních kampaní. Dalším partnerem v těchto aktivitách je spolek NELEŽ, který se věnuje i problematice inzerce na dezinformačních webech. Podpora a šíření povědomí tak může jít ze strany občanské společnosti, ale i ze strany profesních organizací.

Budovat partnerství pro oslovení inzerentů: Při šíření povědomí o ochraně značky jsou klíčová partnerství, která pomohou dostat se přímo k cílové skupině inzerentů. Partnery mohou být obchodní komory, svazy nebo sítě komunikačních agentur, které sdružují marketingové specialisty. Zvláště vhodnými partnery jsou asociace podporující zodpovědný, etický byznys a cíle udržitelnosti (ESG).

Školení marketingových specialistů: Pro budování lepší odolnosti inzerentů bude důležité šířit povědomí o reputačním riziku v souvislosti s dezinformačním obsahem již při školení budoucích marketingových specialistů a odborníků na digitální reklamu. Nejlepší je začít co nejdříve, v rámci kurzů na středních školách, obchodních akademiích, případně vysokých školách.

Spolupráce s finančním úřadem: V případech YouTube kanálů, které monetizují dezinformační, konspirační nebo extremistický obsah, může být v případě nespolupracující platformy vhodným nástrojem pro jejich demonetizaci spolupráce s finančním úřadem. Fyzické osoby, které profitují z inzerce na jejich kanálu, mají povinnost tyto příjmy danit ve speciálním režimu jako příjmy ze zahraničí. Alternativou tedy může být prověření dodržování těchto povinností ze strany finančních úřadů.

Moderace obsahu

Informovat uživatele o tom, že video, které sledují, může obsahovat manipulativní obsah: YouTube by měl být schopen rozpoznat a informovat své uživatele, že sledují dezinformační obsah. Uživatelé by měli přinejmenším dostat možnost ke konkrétním videím dodat kontext, pokud se jedná o zavádějící informaci. Podobnou funkcí umožňuje například platforma X. Algoritmus platformy například nerozpozná Rusko jako agresora, zatímco doporučuje videa týkající se rusko-ukrajinské války. S tím souvisí potřeba tlaku na zvyšování počtu místních moderátorů obsahu a fact-checkerů s potřebnými znalostmi jazyka a místního kontextu. Na podobném principu má potenciál fungovat již výše zmíněná pilotní funkce Poznámek.

Zaměřit se na možnost uvalit sankce na osobu a konkrétního vlastníka kanálu: Jeden ze zásadních problémů systému regulace online prostředí na YouTube je, že když provozovatelé smažou jedno video, nebo kanál, jeho vlastníkovvi to nebrání v tom si krátce na to založit několik dalších. Zakázat činnost konkrétního jednotlivce, nikoli kanálu, by tak mohlo zajistit úspěšnější pokrývání manipulativního obsahu a zabránit jeho dalšímu vzniku.⁵ Tento druh zásahu platformy by mohl fungovat na principu opakovaného přestupku.

Využít potenciálu umělé inteligence: Jelikož je prostředí YouTube tak rozsáhlé a v každé vteřině vznikne několik tisíc videí napříč platformou, schopnosti regulace prostřednictvím lidských zdrojů jsou a vždy budou omezené. Umělá inteligence do budoucna představuje potenciál přinejmenším v lepší kategorizaci obsahu a identifikaci manipulativních taktik. AI může fungovat jako podpůrný nástroj pro prvotní vytipování problematického obsahu, nemělo by však plně nahradit moderátory obsahu se znalostí místního jazyka a reálií.

Využívat data ze zpráv o transparentnosti v advokační činnosti: DSA vyžaduje od platforem včetně YouTube pravidelné [zprávy o transparentnosti](#) o jejich aktivitách v oblasti moderace obsahu. Tyto zprávy obsahují konkrétní data o aktivitě, případně neaktivitě platforem v této oblasti. Státní instituce ale i občanská společnost by proto mohly více pracovat s těmito daty a využívat je při advokační činnosti a vyvíjení tlaku na platformy, zejména v případech, kdy zprávy o transparentnosti nabízejí důkazy o nedostatečných a pomalých reakcích na problematický obsah. Rovněž by s těmito daty mohla více pracovat média a informovat veřejnost o pasivitě platforem.

Vymáhat zvýšení efektivity Programu prioritních nahlašovatelů: Vymáhání jeho efektivity by mohlo být zapojeno například do DSA nebo budoucí evropské legislativy. Zhodnocení výsledků a dopadu tohoto programu by mohlo být zahrnuto do reportovacích povinností YouTube a motivovat tak k obecně většímu využití programu.

Motivovat širokou veřejnost k nahlašování problematického obsahu: Větší počet nahlášení problematického obsahu běžnými uživateli je jednou z možností, jak upoutat pozornost digitálních platforem a motivovat je k zásahu. Odborníci v oblasti digitálních platforem upozorňují, že běžní uživatelé možnost ohlašování alespoň prozatím využívají minimálně.

⁵ Účinnost takového kroku se ukázala v případě extremisty Daniela Bombice na Slovensku. Rovněž Instagram používá tento druh sankce.