CRJM – Centrul de Resurse Juridice din Moldova (old)

Dezinformare ?i deepfakes – posibile implica?ii pentru democra?ie

Jomir Dumitru · Tuesday, May 4th, 2021

Tr?im în epoca post-adev?rului (post-truth), un climat cultural ?i politic în care exist? tot mai pu?in respect pentru adev?r. Acest concept presupune c? oamenilor – ?i mai ales politicienilor – le pas? mai pu?in de ceea ce spun ?i dac? ceea ce spun este adev?rat.[1]

În spatele unei asemenea atitudini formate în timp, nu st? neap?rat ignoran?a, ci eforturile concertate prin care spa?iul public, mai cu seam? cel online, este invadat constant de informa?ii false, p?rtinitoare sau neverificate, a.k.a. Fake news!. Pe lâng? metodele deja devenite „tradi?ionale" precum paginile web de ?tiri false sau trolii, o nou? tehnologie care amenin?? s? consolideze ?i s? înt?reasc? era post-adev?rului sunt tehnologiile ***deepfakes***, numite conven?ional ?i ***synthetic media***.

Ce sunt deepfakes ?i cum sunt create?

Deepfakes sunt videoclipuri false create prin intermediul aplica?iilor digitale ?i sistemelor automate. Imaginile video reale sunt combinate pentru a crea con?inut nou, cu declara?ii sau ac?iuni care nu s-au întâmplat în realitate.

Exemplu de DeepFake: Discrus al fostului pre?edinte american, Barack Obama

Conceptul de baz? din spatele tehnologiei este recunoa?terea facial? (vezi articolul integral aici). Videoclipurile false pot fi create folosind o tehnic? de înv??are automat? ?i inteligen?? artificial? numit? „re?ea contradictorie generativ?" sau acronimul din englez? GANs. În cuvinte simple, modul în care ac?ioneaz? aceast? tehnologie este urm?torul: Un GANs scaneaz? fotografii sau video-uri cu o anumit? persoan?. Tehnologia stabile?te puncte de reper pentru a mapa fa?a persoanei în func?ie de anumite caracteristici specifice precum col?urile ochilor ?i gurii, n?rile ?i conturul liniei maxilarului, printre altele, dup? care poate crea con?inut video nou cu o combina?ie de sunete sau mi?c?ri inexiste.[2] Pe lâng? imaginile video, tehnologia GANs poate fi utilizat? pentru a genera sunete noi din cele existente, sau un text nou dintr-un text existent.

De ce trebuie s? fim precau?i în leg?tur? cu Deepfakes?

Tehnologia **deepfake** nu mai este limitat? doar în cercurile academice ci disponibil? pe internet.

Aplica?ii software precum FakeApp sau FaceSwap, Snapchat sau Facebook Messenger, utilizeaz? aceast? tehnologie în care fa?a real? a persoanei este suprapus? cu fa?a unei alte persoane sau anima?ii.[3] Chiar dac? în prezent, majoritatea video-urilor false sunt de o calitate care u?or poate fi depistat? ca fiind fals?, potrivit exper?ilor, aceast? tehnologie avanseaz? zilnic.[4] Astfel foarte curând ne va fi complicat s? putem discerne dac? un anumit video este sau nu în realitate fals. Rezultatele unui **deepfake** vor fi destul de conving?toare, aceste video-uri fiind foarte dificil de identificat ca fiind false.

Riscul principal de utilizare al tehnologiei deepfake este c? acesta poate permite la un moment dat oric?rui utilizator s? videoclipuri despre oricine, f?când ?i spunând orice.[5] Implica?iile unei asemenea tehnlogii pentru dezinformare sunt majore, mai ales dac? nu doar imaginea video poate fi replicat? dar ?i vocea persoanei, pân? la exactitate. Iat? câteva exemple de utilizare a tehnologiei deepfake:

- Panic? public? indus? prin urgen?e „false” precum calamit??i naturale sau stare de asediu ?i r?zboi;
- Discursuri politice manipulatorii sau false;
    - Atacuri asupra unor persoane publice, ap?r?tori ai drepturilor omului sau jurnali?ti.

Ce putem face în leg?tur? cu Deepfakes?

Speciali?tii recomand? câ?iva pa?i care merit? urma?i atunci când credem c? avem de afacere cu un video care portretizeaz? o declara?ie neobi?nuit? sau informa?ie posibil fals?. Ca ?i în cazul informa?iei scrise, pretins false, trebuie s? urm?m câ?iva pa?i de control[6]:

1. **S? analiz?m cu aten?ie de fiecare dat? sursa** – speciali?tii recomand? de fiecare dat? s? atragem aten?ie originii con?inutului video. Câteva întreb?ri care neap?rat trebuie adresate sunt: A fost înc?rcat de c?tre un utilizator cunoscut sau într-un cont aleatoriu pe o re?ea social?? Este sau nu con?inut sponsorizat? Cine pretinde c? este proprietarul acesteia?. A fost publicat de c?tre un portal de ?tiri cu reputa?ie bun? sau un site/portal necunoscut?
2. **S? verific?m dac? video-ul mai poate fi g?sit online ?i în alt? parte**: Verific?m unde mai poate fi g?sit videoclipul, dac? mai este (sau nu) de g?sit online, pentru a vedea dac? sunt mai multe ?tiri cu acest video din partea unor publica?ii cu reputa?ie bun? sau persoane credibile (jurnali?ti, formatori de opinie);
3. **S? ne ab?inem de la concluzii** pân? nu primim confirm?ri suplimentare. La aceast? etap? este foarte important s? nu distribuim mai departe video-ul. Acest lucru este valabil mai ales în situa?iile de ultim? or?, în care informa?iile se mi?c? rapid ?i sunt adesea gre?ite sau interpretate gre?it în primele ore de la difuzare, pân? ce sunt verificate.

Suplimentar pa?ilor descri?i, putem apela la anumite programe sau organiza?ii media care verific? în regim online dac? im con?inutul video are semne a fi fals, sau dac? este acesta este original.

De exemplu, programul **Reality Defender**[7] scaneaz? ?i verific? dac? con?inutul este fals, manipulatorii sau original. Un raport asupra fiec?rui videoclip trimis rezum? concluziile f?r? a specula inten?ia manipul?rilor.

Un alt mecanism, **Platforma Stop-fals**[8], o campanie împotriva informa?iei false ?i tenden?ioase, desf??urat? de Asocia?ia Presei Independente (API) din Republica Moldova stocheaz? o baz? de date ?i cu ajutorul speciali?tilor verific? autenticitatea con?inutului video. Oricine poate sesiza

platforma stop-fals dac? întâlne?te con?inut posibil fals, inclusiv video.

R?mâne?i al?turi de CRJM pentru a afla mai multe detalii despre noile tehnologii ?i drepturile omului în era digital?. Ne pute?i urm?ri pe pagina noastr? web CRJM.org ?i re?elele de socializare Facebook, OK.ru, Twitter, Linkedin.

Acest articol face parte dintr-o serie de publica?ii non-academice realizate de Centrul de Resurse Juridice din Moldova (CRJM) în cadrul proiectului „Program de capacitare în drepturi digitale" sus?inut de Centrul Interna?ional pentru Drept non-profit (ICNL). Opiniile exprimate apar?in CRJM ?i nu reflect? în mod necesar pozi?ia ICNL.

[1] https://dictionary.cambridge.org/dictionary/english/post-truth.

[2] How to Spot a Deepfake Like the Barack Obama?Jordan Peele Video", Craig Silverman,

BuzzFeed News: https://www.buzzfeed.com/craigsilverman/obama-jordan-peele-deepfake-video-debunk-buzzfeed; Mal-uses of AI-generated Synthetic Media and Deepfakes: Pragmatic Solutions Discovery Convening (2018).

[3] Deepfakes: It's Not What it Looks Like", Sam Gregory, Atlantic Council:

https://www.youtube.com/watch?v=Qh_6cHw50l0&feature=youtu.be; "Mal-uses of AI-generated Synthetic Media and Deepfakes: Pragmatic Solutions Discovery Convening" (2018).

[4] Fake Videos of Real People and How to Spot Them", Supasorn Suwajanakorn, TED:

https://www.youtube.com/watch?v=o2DDU4g0PRo.

[5] https://www.reuters.com/article/us-chertoffdonahoe-hacking-commentary-idUSKCN1NH1W2

[6] Commentary: For Election Hackers a New and More Dangerous Tool", Michael Chertoff and

Eileen Donahoe, Reuters: https://www.reuters.com/article/us%E2%80%90chertoffdonahoe%E2%80%90hackingcommentary/commentary%E2%80%90for%E2%80%90election%E2%80%90hackers%E2%80%90a%E2%80%90new%E2%80%90and%E2%80%90more%E2%80%90dangerous%E2%80%90tooldUSKCN1NH1W2

[7] Reality Defender: https://rd2020.org/index.html.

[8] Platforma Stop-Fals (Asocia?ia Presei Independente): https://semnale.stopfals.md/ro/about.

This entry was posted on Tuesday, May 4th, 2021 at 4:28 pm and is filed under Blog
You can follow any responses to this entry through the Comments (RSS) feed. Both comments and pings are currently closed.