

ALEKSANDAR ĐOKIĆ

DEEP FAKE PORN



DUBOKO UČENJE
I STRAH OD LEKCIJA

FAKE
NEWS
tragač

**FAKE
NEWS**
tragač



**Deep Fake Porn –
duboko učenje i strah od lekcija**

Novosadska novinarska škola
FakeNews Tragač
Kosovska 1,
2100 Novi Sad
Telefon: 021/ 424246
Mail: office@novinarska-skola.org.rs

Za izdavača
Milan Nedeljković

Autor
Aleksandar Đokić

Dizajn / ilustracije
Milovan Nikolić
Stefan Janjić

Novi Sad,
Maj 2023.

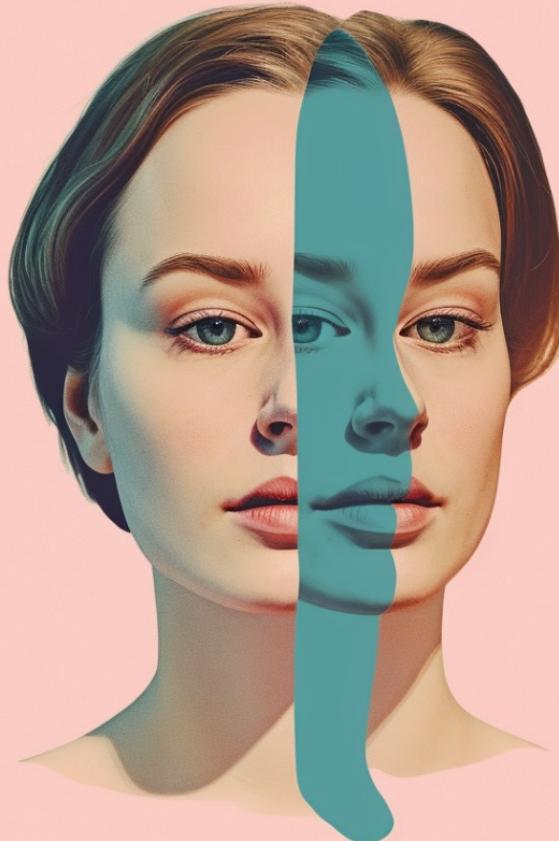
ALEKSANDAR ĐOKIĆ

DEEP FAKE PORN

DUBOKO UČENJE
I STRAH OD LEKCIJA

Ova publikacija je nastala u okviru projekta „Mladi i mediji za demokratski razvoj“ koji Beogradska otvorena škola realizuje uz podršku Švedske. Stavovi i mišljenja autora izneta u ovoj publikaciji ne predstavljaju nužno i mišljenje partnera i donatora.





OD „ZABAVE“ DO HORORA

Muškarac na video-snimku otvara vrata, a u stan ulazi pevačica Arijana Grande. On joj naređuje da se nasmeje, a onda i da se skine. Ona sve to poslušno i čini. To je uvod u dvanaestominutni porno snimak sa popularnom tinejdž zvezdom u glavnoj ulozi. Iako sve izgleda vrlo realistično, na snimku nije zaista ona – u pitanju je video koji je generisala veštačka inteligencija, sa jednog od brojnih pornografskih sajtova namenjenih ljudima koji žele da vide svetski poznate ličnosti u porno snimcima.

Autori Mejson i Holt u svom radu o dipfejku definišu ovaj fenomen kao „proizvod aplikacija veštačke inteligencije (AI) koje spajaju, kombinuju, zamejuju i preklapaju slike i video-snimke da bi se kreirali lažni video-snimci koji izgledaju autentično“. Sama reč „deepfake“ zapravo je neologizam koji obuhvata „duboko učenje“ i „laž“. Sintagmu „dipfejk pornografija“ dobijamo 2017. godine, kada korisnik Redita pod pseudonimom „deepfakes“ stvara zaseban sabredit (stranicu) i poziva druge da mu se pridruže u kreiranju i šerovanju dipfejk pornografije. Sabredit je zabranjen 2018. godine, ali to nije sputalo širenje ovog fenomena. Njegovo vreme je tek počelo.

Profesor sa univerziteta u Upsali (Švedska) Karl Ehman u svom radu ističe da se dipfejk uglavnom odnosi na „hiperrealistične video zapise u kojima je lice osobe analizirano algoritmom za *duboko učenje*, a zatim postavljeno na lice glumca/glumice u videu. Pošto je algoritam ‘naučio’ crte lica iz različitih uglova i kako se ono kreće u različitim izrazima, može ga replicirati na način koji prati izraze glumca“. Ehman dalje pojašnjava da ovo ne zahteva bilo kakvo kršenje privatnosti ili nedozvoljen pristup informacijama, već da se to može uraditi zahvaljujući javno dostupnim slikama ili video-materijalu koji već postoji na internetu.

„Data Science Srbija“ je organizacija koja se bavi podacima, veštačkom inteligencijom i mašinskim učenjem. Njen osnivač Aleksandar Linc Đorđević za FakeNews Tragač kaže da je kompjuterska obrada fotografija dovela do mnogo toga dobrog, ali otvorila i mogućnosti za zloupotrebu. „Ona se može koristiti u ranoj detekciji raka ili u spasavanju ugroženih od zemljotresa. Međutim, isto tako je dovela i do dipfejka. To je sintetički medij koji u pozadini koristi tehnologiju dubokog učenja, koje trenira neuronske arhitekture poput autoenkodera i/ili generativne adversijalne mreže (GANs)”, dodaje on.

MRDEEP FAKES



Ako ukucate u Gugl „deep fake porn“, među prvim rezultatima naći će se popularni sajтовi sa porno snimcima slavnih ličnosti iz sveta muzike, filma, politike i sporta. Kažemo „slavnih ličnosti“, ali pre svega mislimo „slavnih žena“.

Svoju „radoznalost“ na takvim sajtovima mogu da zadovolje oni koji žele da vide kako u pornografskim scenama izgledaju Andželina Džoli, Dženifer Lopez, Skarlet Johanson, zvezde bolivudskih filmova, ali i političarke, poput done-davne finske premijerke Sane Marin ili bivše prve dame SAD Mišel Obame. U manjoj meri ima i muškaraca – moguće je, na primer, naći gej porno filmove sa kompjuterski generisanim Bredom Pitom, Tomom Hardijem ili Džastinom Biberom u glavnim ulogama.

Među najpopularnijim sajtovima ovog tipa je *MrDeep Fakes*, koji, ako je verovati servisu *SimilarWeb*, mesečno ima više od 17 miliona poseta. Ako se pogledaju demografski podaci, ovaj sajt daleko češće posećuju muškarci (82,52%) nego žene (17,48%). Najviše je korisnika koji imaju od 25 do 34 godine (29,89%), a potom mlađih od 18 do 24 godine (29,54%). Najviše korisnika pristupa ovom sajtu iz SAD (21,42%), Nemačke (6,57%), Velike Britanije (6,56%), Kanade (4,85%) i Južne Koreje (4,36%). Najpopularniji video je onaj sa Dženom Ortegom, glavnom glumicom serije *Sreda*. Video trenutno ima više od šest miliona pregleda.

Prema podacima platforme Whois, sajt *MrDeepFakes* je registrovan 2018. godine. Snimci koji su nastali tada uglavnom se značajno razlikuju od onih svežijih. Stariji snimci su lošije rezolucije, lice deluje „nalepljeno“ jer se razlikuje od ostatka tela po osvetljenosti ili nijansi tena, a mimika je usporena i neprirodna. Današnji video-materijali, posebno kada ih prave veštiji kreatori, izgledaju daleko uverljivije.



ŽRTVE I DISKLEJMERI



Na ovom sajtu objavljene su smernice za kreatore o tome koje poznate ličnosti dolaze u obzir – filmske i muzičke zvezde, influenseri, ljudi iz politike ili biznisa. Targetirana osoba, kako navode, mora imati određeni uticaj na društvenim mrežama, što se određuje prema broju pratilaca – najmanje 100.000 na Jutjubu, Tviteru, Instagramu i Tviču, ili 200.000 na TikToku.

Na sajtovima koji se bave dipfejkom postoje disklejmeri, odnosno odričanje od odgovornosti – navodi se da cilj ove vrste pornografije nije povredljivanje ili ponižavanje bilo koga, već samo zabava. Ovakva hipokrizija deluje prilično bizarno: kao da ženama poručuje „hajde da se zabavimo scenama seksa u koje si ubaćena mimo svoje volje, ali nemoj se zbog toga osećati povređeno niti poniženo“.

Kreatori dipfejk sadržaja moraju, navodno, da poštuju čitav niz procedura, inače će im administratori „automatski odbiti video“. Na primer, videosnimci koji nisu nastali na osnovu dipfejk tehnologije odmah će biti obrisani. Kreatori mogu i da unovče svoj rad tako što će, recimo, ograničiti pristup videu i učiniti ga dostupnim samo onima koji su spremni da plate.

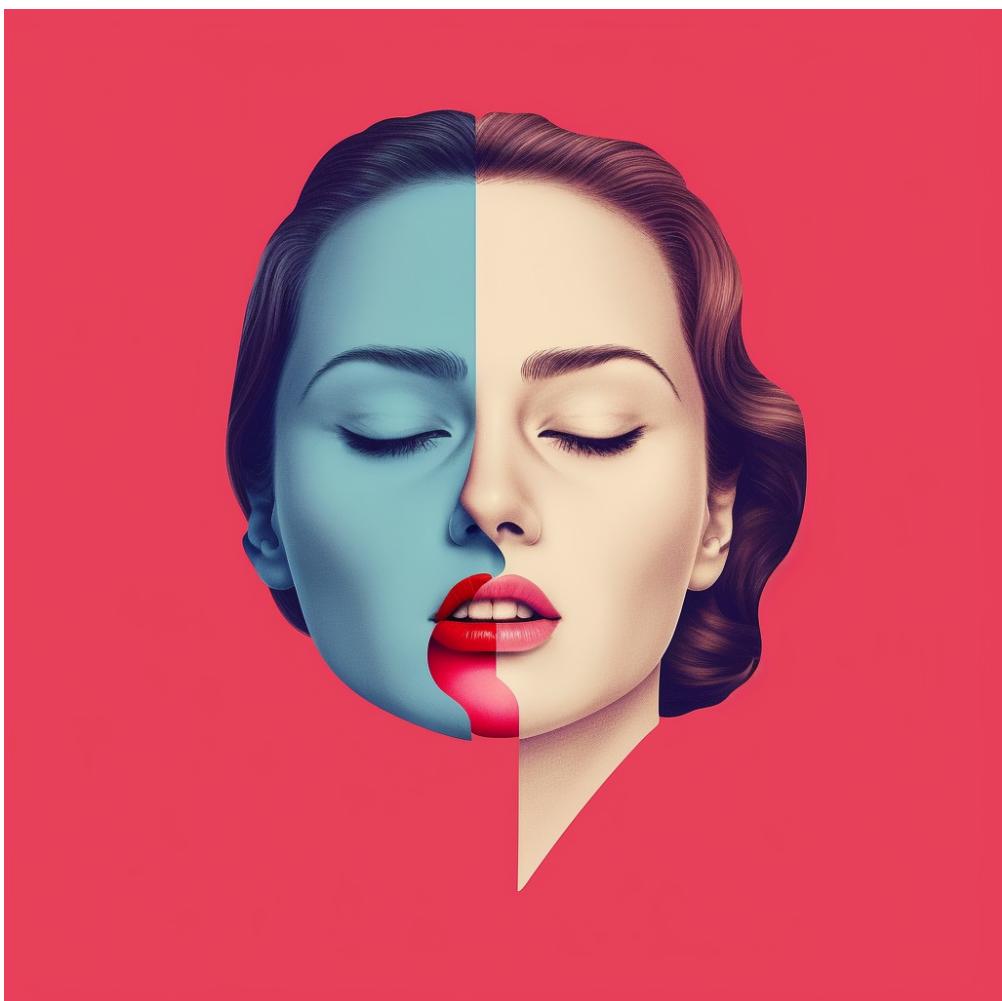
Na sajtu *MrDeepFakes* postoje smernice za kreatore kojima se savetuje da rade u programu *Deep Face Lab* (DFL). Sem smernica, tu su i čitavi tutorijali i uputstva. Na sajtu kompanije Softonic ističe se da je reč o vrhunskom programu za kreiranje dipfejka, koji omogućava da efikasno menjate lica na bilo kojoj slici ili video-snimku. Prema navodima kompanije Softonic, DFL je vodeći program na tržištu, sa preko 95% kreiranih dipfejk video-zapisu.

Osim DFL-a, postoje i drugi softveri za generisanje dipfejk snimaka, pa i pornografskih. Linc Đorđević pominje Lens Ai, Deepfakes Web i Wombo. Malo manje zahtevne ali i manje napredne su mobilne aplikacije u kojima je moguće generisati amaterske dipfejk sadržaje, poput MorphMe i Mivo, koje ukupno broje preko 50 miliona preuzimanja.

Da bi jedan video izgledao što uverljivije, kreator mora da sakupi što više materijala o određenoj poznatoj ličnosti koju želi da „smesti“ u porno film – preporuka je između 3.000 i 8.000 fotografija i snimaka.

Kako piše u smernicama za korišćenje softvera DFL, poželjno je da video snimci prikazuju slavnu ličnost iz raznih uglova, ali i sa različitim facijalnim eks-

presijama, kako bi pokrili sve moguće izraze lica prilikom montaže. Takođe se predlaže i kombinacija video zapisa i fotografija. Video-snimci i fotografije treba da pokrivaju sve ili barem većinu mogućih pozicija lica ili glave – pogled gore, dole, levo, desno, pravo u kameru, a najbolji način da se to postigne jeste korišćenje više intervjeta i/ili filmova. Kada je reč i o različitim izrazima lica, to uključuje otvorena i zatvorena usta, otvorene i zatvorene oči, osmehe, mrštenje, pogled u različitim pravcima: što više različitih izraza lica pronađete, to će rezultati biti bolji, navodi se u smernicama. Svim tim raznolikim snimcima kreator „puni“ softver – na primer, pomoću 3.000 fotografija i video-snimačke Andeline Džoli softver „uči“ kako izgleda njeno lice sa svim mogućim ekspresijama i iz svih mogućih uglova. Kreator takođe u softver dodaje i već postojeći porno-snimanak, i onda veštačka inteligencija lice porno glumice zamjenjuje licem Andeline Džoli.



Sadržaj na ovom sajtu je besplatan, ali je od kreatora moguće naručiti video-snimanak sa konkretnom poznatom ličnošću. Plaća se u kriptovalutama, te svi, i kreator i naručilac u ovom slučaju, ostaju potpuno anonimni, a cena se dogovara preko privatnih poruka. Neki od kreatora zarađuju i od obučavanja drugih da rade u DFL programu.

Linc Đorđević za Tragač ističe savet koji se može koristiti u prepoznavanju dipfejka, a to je da pažljivo gledamo sadržaj jer postoje nepravilnosti na licima koje možemo uočiti.

„AKO JE LEGALNO MAŠTATI O SEKSU...“



Kako na dipfejk industriju gledaju oni koji prave ovakve sadržaje, pokušali smo da saznamo direktno od njih. Poslali smo pitanja na desetak profila sa sajta *MrDeepFakes*, a odgovorio je samo jedan, čiji je nadimak na ovom sajtu TMBDF (*Truly MindBlowing DeepFakes*).

Najpre je odbio da razgovara zbog bojazni da će njegovi odgovori biti istrgnuti iz konteksta. „Nemam šta da kažem. Sve što kažem vi ćete interpretirati prema svom narativu, a to je da smo svi mi silovatelji koji mrze žene“. Onda se, ipak, predomislio. „Najveći broj kreatora ovo radi jer je to zanimljiv način da se porno-filmovi unaprede pomoću atraktivnih poznatih ličnosti, nema agende niti mržnje prema ženama“.

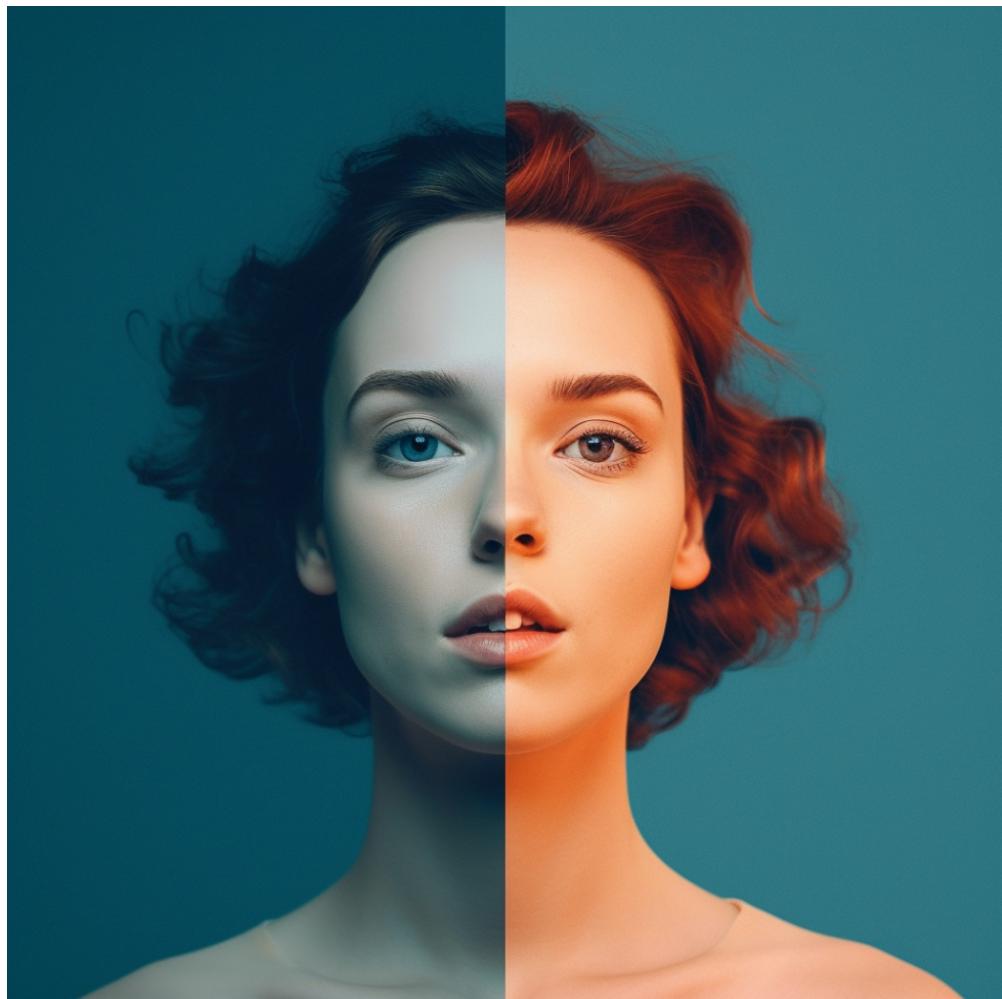
Za Tragač kaže da pomoću dipfejka fantazija postaje korak bliža stvarnosti. „Gledati poznate ličnosti tokom seksa umesto samo ih zamišljati, zameniti ružna ili makar prosečna lica porno zvezda licima nekih atraktivnijih poznatih ličnosti – to generalno može biti zabavno čak i kada je nevešto napravljeno“, kaže on.

Sve ovo je počeo da radi zato što mu je, kako kaže, bilo dosadno, pa je odlučio da se zabavi. „Mnogi muškarci to rade, kao i neke devojke, štaviše ponekad vidimo ljude koji prave gej/ lezbejski sadržaj, pa čak i trans, tako da čak i

pripadnici LGBT populacije prave dipfejk pornografiju, to nije rezervisano samo za muške kreatore“, priča TMBDF.

Novinarki Andželi Milivojević, međutim, ovde ništa nije zabavno – ona dipfejk pornografiju smatra digitalnim nasiljem, jer je najveći broj ovih snimaka napravljen bez saglasnosti žene, odnosno muškarca.

„Ako znamo da je dipfejk manipulacija pornografskih snimaka u kojima je originalno lice glumice zamenjeno licem neke druge žene, koja za to nije dala suglasnost, a taj snimak je naknadno učinjen dostupnim hiljadama ljudi, sve to govori o tome da osoba, žrtva dipfejka prolazi kroz traumu“, kaže ona.



TMBDF, međutim, smatra da „ako je legalno fantazirati o seksu sa poznatom ličnošću, onda je u redu kreirati i dipfejk video“. Za Tragač objašnjava da kreatori tog sadržaja ne upadaju ni u kuće slavnih ličnosti da ih fotografišu,

niti ih kidnapuju da bi snimili porno film sa njima. „Poznate ličnosti već defiluju polugole na crvenim tepisima i imaju scene seksa u svakom svom filmu, da ne spominjem da se fotografiju obnažene – one žele da ih ljudi vide gole, tako da ne vidim u čemu je zaista problem sa dipfejkom“, dodaje. On ističe da je sadržaj koji koriste da prave dipfejk pornografiju sa slavnima već dostupan – to su filmovi u kojima oni glume, njihovi intervjui, fotografije, reklame u kojima se pojavljuju, njihove društvene mreže. „Oni biraju da nemaju privatnost time što su slavne ličnosti“, priča on.

Ovakva vrsta opravdavanja može zvučati uverljivo, ali je suštinski neutemeljena i štetna na nekoliko nivoa. Javne ličnosti pristaju odabirom profesije na manji stepen privatnosti, ali se ne odriču privatnosti u celosti. Kada poznata glumica snima erotske scene za potrebe filma, ona ima svest o tome šta ta scena podrazumeva i na koji način će biti distribuirana. Kreiranje dipfejk pornografije ne može se poistovetiti s fantaziranjem o seksu s poznatim ličnostima, jer je film jedan oblik materijalizacije, dokumenta, čak i onda kada je lažiran. I profesorka prava sa Univerziteta u Daramu (Velika Britanija) Kler MekGlin ističe za Tragač da odricanja od odgovornosti nisu dovoljna, te da je distribucija dipfejk pornografije neke osobe bez njene saglasnosti nedvosmisleno pogrešna i štetna.

PET PO PET



Kada je reč o unovčavanju svog rada, TMBDF kaže da ovaj posao nije profitabilan i da od njega ne može da se živi. Ipak, na sajtu *MrDeepFakes* on je jedan od popularnijih kreatora, čiji portfolio obuhvata gotovo 1.000 dipfejk porno snimaka. Od toga, gotovo polovina je nedostupna „običnim“ korisnicima, budući da se pristup naplaćuje. Cena pristupa varira, ali u proseku iznosi oko pet dolara. Jedan od najpopularnijih klipova, onaj sa pevačicom Bili Ajliš, ima oko 300.000 pregleda, a da biste mu pristupili morate da platite 500 tokena, što je oko šest dolara.

Osim toga, ovaj kreator zarađuje i od pravljenja dipfejk video-snimaka po narudžbini, a svojim pratiocima je dao mogućnost da se pretplate na njegov

kanal po ceni od 13 dolara mesečno i time dobiju pristup svim videima. Za 100 dolara zainteresovani mogu dobiti dvonedeljni kurs DFL softvera. Kako navodi u svom portfoliju, napredni nivo obuke – „ako želite da znate sve tajne“ – naplaćuje 1.000 dolara.

Dok obilato eksploratiše lica poznatih ličnosti bez njihove saglasnosti, TMBDF ima svoja „etička pravila“ – ne pravi gej kontent i nasilne sadržaje, a potencijalnim klijentima poručuje i da im neće praviti lažne klipove sa njihovim bivšim partnerima.

UKLANJANJE SADRŽAJA



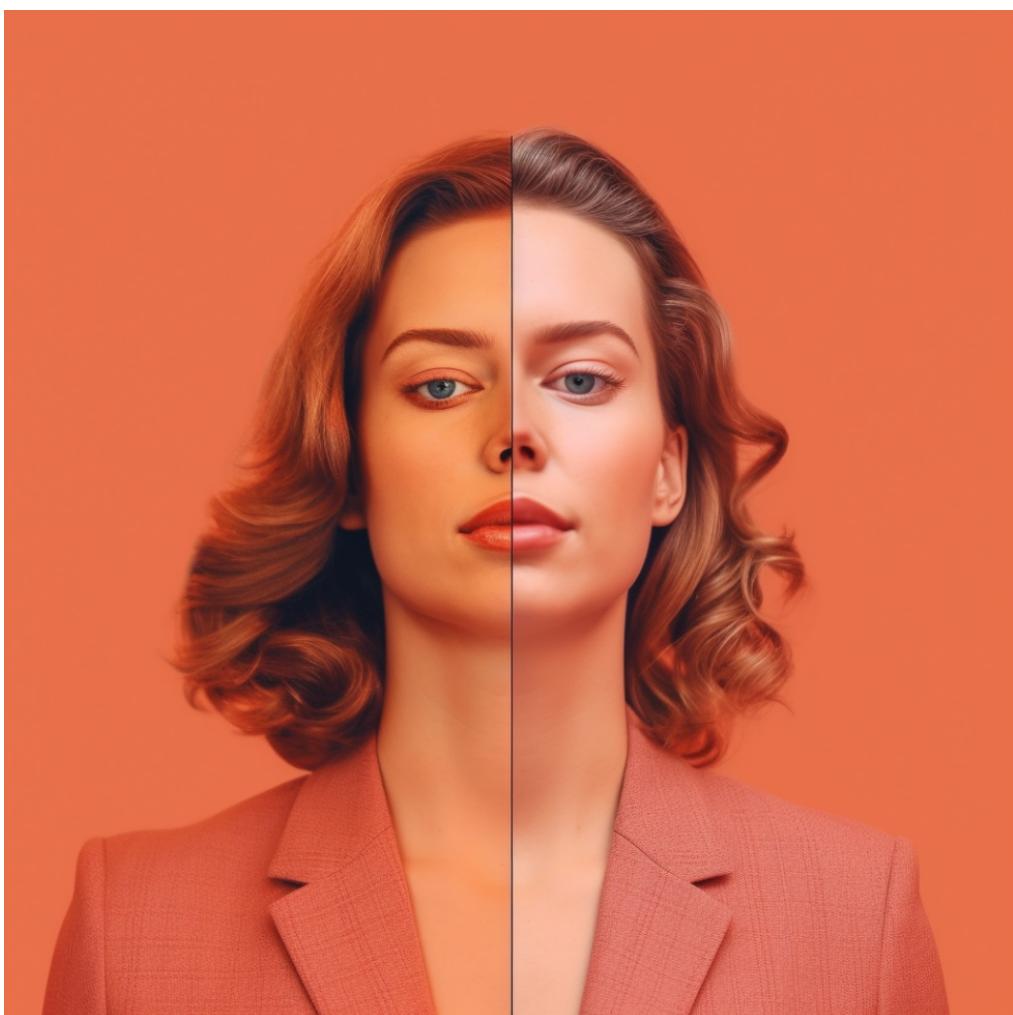
Video-snimci sa anonimnim osobama su na ovakvima sajtovima zabranjeni, upravo zbog sprečavanja osvetničke pornografije. Ipak, neki drugi pornografski sajtovi dipfejk pornografiju i osvetničku pornografiju tretiraju isto – kao pornografiju „objavljenu bez pristanka“. Sajt Pornhab sve ovo podvodi pod „non-consensual content“, odnosno sadržaj bez saglasnosti, gde svrstava osvetničku pornografiju, tajno snimanje, hakovane snimke iz telefona, pa i dipfejk pornografiju.

Prema statistici iz 2022. godine Pornhab je uklonio oko 13.000 videa i fotografija koje su bile objavljene bez saglasnosti (uključujući dipfejk pornografiju). Ipak, da li je to dovoljno? Pre dve godine više od 30 žena tužilo je Pornhab, jer je, kako navode, koristio sadržaj bez saglasnosti da bi „postao dominantna onlajn pornografska kompanija u svetu“, prenosi NBCNews. Žene koje su podnele tužbu istakle su da su bile žrtve seksualne eksploracije, silovanja i trgovine ljudima. Serena Flajts i više od 30 žena navele su u tužbi da su bile žrtve neovlašćenih video-snimaka koje je postavio i distribuirao Pornhab. Serena je tada pred Donjim domom Kanade istakla da je bila jedna od ljudi koji su završili kao beskućnici, napustili školu i završili na drogama, potpuno odvojeni od svoje

porodice. „Na kraju sam pokušala da se ubijem mnogo puta. Završila sam u duševnim bolnicama“, istakla je Flajts, a prenosi NBCNews.

Milivojević ističe da je upravo Pornhab pre oko dve godine uveo obveznu saglasnost osoba na snimku, kao i proveru identiteta pre nego što se dozvoli postavljanje snimka. „To je jedan od načina da se dipfejk zaustavi, jer bez dozvole nema objavljivanja“, ističe ona.

Kad ukucate „deepfake“ na Pornhab pretraživač, sajt odmah izbací upozorenje da to što hoćete možda nije legalno. Potencijalne žrtve ovakvih praksi Pornhab upućuje, između ostalog, na sajt „Cyber Civil Rights“, gde građani mogu, u zavisnosti od države u SAD u kojoj žive, pokušati da zaštite svoja prava i sudskim putem. Zakone protiv dipfejka zasad imaju samo četiri američke države: Kalifornija, Njujork, Džordžija i Virdžinija.



U Velikoj Britaniji, kaže novinarka Milivojević, postoji inicijativa da dipfejk slike i snimci budu kriminalizovani, a to potvrđuje i profesorka Kler Mek-Glin, koja kaže da su u Engleskoj i Velsu predlozi pred parlamentom, dok u Škotskoj već postoji dipfejk kao krivično delo.

„Dipfejk bi trebalo da bude predmet krivičnih sankcija. Odricanje od odgovornosti na sajtovima nije dovoljno. Svako ko ima sliku na internetu mogao bi biti žrtva dipfejk pornografije“, kaže profesorka iz Velike Britanije za Tragač.

Novinarka Milivojević ističe da su prve dve adrese kojima žrtve treba da se obrate u Srbiji policija i Tužilaštvo za visokotehnološki kriminal.

Ona je ove godine objavila tekst za koji je više od 25 žena i devojaka iz Srbije govorilo o šoku, strahu i stidu koje su doživele kao žrtve osvetničke pornografije anonymnih profila koji su bez njihove dozvole svakodnevno delili intimne slike i snimke u Telegram grupama.

Sličnost dipfejk i osvetničke pornografije Milivojević vidi u viralnosti kroz objavljuvanje na internetu, kao i u posledicama koje ostavlja na žrtve. „Generalno, dipfejk se često smatra sivom zonom, ali potencijalno može biti procesuiran kao tužba zbog klevete ili kršenje autorskih prava na lik ili portret osobe koja je zloupotrebljena“, kaže novinarka Milivojević.

S druge strane, Bojan Perkov iz SHARE fondacije ističe da Srbija nije „veliko tržište“ kada je u pitanju dipfejk pornografija i da će dipfejk ostati fokusiran na poznate ličnosti. On dodaje da su takve vrste generisanih sadržaja karakteristične za svetske zvezde iz sveta šoubiznisa ili svetske lidere, zbog toga što postoji ogroman broj fotografija i snimaka na kojima se oni nalaze.

„Male su šanse da će neko koristiti takav softver da kreira osvetničku pornografiju da bi se svetio, na primer, bivšoj devojci. Ima jednostavnijih načina od generisanja dipfejk materijala – na primer objavljuvanje stvarnih snimaka bez pristanka te osobe“, kaže Perkov.

Reč je o nasilju, kaže on, zbog čega je SHARE fondacija napravila platformu Cyber intima, preko koje žrtve mogu da prijave ukoliko je neko objavio na internetu njihove intimne sadržaje. „Pomoći ovog alata oni koji pretrpe tu vrstu zloupotrebe mogu da zahtevaju uklanjanje materijala ili da se jave nekoj od organizacija za podršku ukoliko im je potrebna pomoć u tom procesu“, priča Perkov.

Novinarka Milivojević za Tragač kaže da dipfejk tehnologiju može da zloupotrebi bilo ko, pa i korisnici iz Srbije.

„Da li će to biti način osvete ili nečega drugog, to je pitanje svakog individualnog slučaja. Za razliku od osvetničke pornografije ili nasilja nad ženama kroz objavljuvanje privatnih slika ili snimaka bez dozvole, što često ima svrhu da

ženu ponizi ili osramoti, dipfejk može da bude napravljen od potpuno nepoznate osobe koja želi da ispuni neku svoju fantaziju“, ističe ona.

ŠTA ŽRTVA MOŽE DA UČINI



Kejt Ajzaks, koja je žrtva dipfejk pornografije, za [BBC](#) je pričala o svom slučaju i kako je za njega saznala. „Panika me je obuzela, neko mi je uzeo lice i stavio ga na porno video i učinio da izgleda da sam to ja“, rekla je ona. „Srce mi se steglo, nisam mogla jasno da razmišljam. Sećam se samo osećaja da će se ovaj video pojavljivati svuda“.

Ubrzo nakon toga, pokrenula je kampanju [#NotYourPorn](#) na kojoj se mogu pronaći saveti šta uraditi ako se pronađete na nekom od dipfejk videa ili ukoliko vam neko dojavi da je video takav sadržaj.

Bitno je sakupiti dokaze – važno je preuzeti snimke i skrinštovati datume, kao i vreme, korisnička imena i URL-ove i smestiti ih u bezbedan folder, a potom i zaštititi lozinkom. Dalje je neophodno prijaviti naloge – ukoliko je to društvena mreža, bitno je prijaviti fotografiju/video na toj istoj platformi. Potom, pozvati policiju i zvanično prijaviti šta se desilo i sa njima podeliti dokaze koje ste prikupili. Poslednja, i ne najmanje važna stvar: tražite pomoć i savet od stručnjaka. ●

