Problematika deepfake

Roman Chovanec

Bakalářská práce
2023

Univerzita Tomáše Bati ve Zlíně
Fakulta logistiky a krizového řízení

Univerzita Tomáše Bati ve Zlíně
Fakulta logistiky a krizového řízení
Ústav ochrany obyvatelstva

Akademický rok: 2022/2023

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení:	**Roman Chovanec**
Osobní číslo:	**L19559**
Studijní program:	**B2825 Ochrana obyvatelstva**
Studijní obor:	**Ochrana obyvatelstva**
Forma studia:	**Kombinovaná**
Téma práce:	**Problematika deepfake**

Zásady pro vypracování

1. Zpracujte rešerši současného stavu předmětné oblasti.
2. Seznamte se s technikami tvorby deepfake.
3. Proveďte analýzu historického zneužití deepfake.
4. Pojednejte o potenciálu hrozby deepfake do budoucna.

Forma zpracování bakalářské práce: **tištěná/elektronická**

Seznam doporučené literatury:

1. GREGOR, Miloš a Petra VEJVODOVÁ. *Nejlepší kniha o fake news, dezinformacích a manipulacích!!!*. Brno: CPress, 2018. ISBN 978-80-264-1805-4.
2. SHICK, Nina. *Deepfakes: The Coming Infocalypse*. New York: Twelve, 2020. ISBN 1538754304.
3. YOUNG, Nobert. *DeepFake Technology: Complete Guide to DeepFakes, Politics and Social Media*. Spojené státy americké: Nezávisle vydáno, 2019. ISBN 107849469X.

Další odborná literatura dle doporučení vedoucího bakalářské práce.

Vedoucí bakalářské práce: **Ing. Petr Svoboda, Ph.D.**
Ústav ochrany obyvatelstva

Datum zadání bakalářské práce: **1. prosince 2022**
Termín odevzdání bakalářské práce: **5. května 2023**

L.S.

doc. Ing. Zuzana Tučková, Ph.D.
děkanka

prof. Ing. Dušan Vičar, CSc.
ředitel ústavu

V Uherském Hradišti dne 2. prosince 2022

PROHLÁŠENÍ AUTORA BAKALÁŘSKÉ PRÁCE

Beru na vědomí, že:

- bakalářská práce bude uložena v elektronické podobě v univerzitním informačním systému a dostupná k nahlédnutí;
- na moji bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- podle § 60 odst. 1 autorského zákona má Univerzita Tomáše Bati ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- podle § 60 odst. 2 a 3 autorského zákona mohu užít své dílo – bakalářskou práci nebo poskytnout licenci k jejímu využití jen s předchozím písemným souhlasem Univerzity Tomáše Bati ve Zlíně, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše);
- pokud bylo k vypracování bakalářské práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tj. k nekomerčnímu využití), nelze výsledky bakalářské práce využít ke komerčním účelům;
- pokud je výstupem bakalářské práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá. Neodevzdání této součásti může být důvodem k neobhájení práce.

Prohlašuji,

- že jsem na bakalářské práci pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.
- že odevzdaná verze bakalářské práce a verze elektronická nahraná do IS/STAG jsou obsahově totožné.

V Uherském Hradišti, dne: 3.5.2023

Jméno a příjmení studenta: Roman Chovanec

.....
podpis studenta

ABSTRAKT

Tato bakalářská práce se zaměřuje na problematiku technologie deepfake a její potencionální hrozby. V teoretické části jsou představeny základní pojmy, legislativa a používané techniky při vytváření deepfake obsahu. Dále jsou popsány známé případy zneužití této technologie a jaké důsledky můžou nastat. Praktická část je převážně zaměřena na tvorbu možných modelových situací, při kterých došlo ke zneužití deepfake technologie a jsou popsány důsledky, které mohou nastat pro společnost, nebo také potíže v rámci ochrany obyvatelstva. Jsou zde vytvořeny i návrhy opatření jak se proti takovým hrozbám bránit. V další části jsou provedeny řízené rozhory, které má za úkol zjistit jaké mají povědomí běžní lidé o pojmu deepfake.

Klíčová slova: Deepfake, hrozby, technika, umělá inteligence, zneužití

ABSTRACT

This bachelor thesis focuses on the issue of deepfake technology and its potential threats. The theoretical part introduces basic concepts, legislation, and techniques used in creating deepfake content. Furthermore, known cases of misuse of this technology and potential consequences are described. The practical part is mainly focused on creating possible model situations in which deepfake technology has been abused, and the consequences that may arise for society, as well as difficulties in protecting the population. Suggestions for measures to protect against such threats are also proposed. In the next section, structured interviews are conducted to determine the level of awareness of the general public regarding the term deepfake.

Keywords: Artificial intelligence, deepfake, misuse, technique, threats

Na tomto místě bych rád poděkoval za vedení bakalářské práce Ing. Petru Svobodovi, Ph.D. za jeho odborné vedení, cenné rady a připomínky, které mi velmi pomohly při zpracování této práce.

Prohlašuji, že odevzdaná verze bakalářské/diplomové práce a verze elektronická nahraná do IS/STAG jsou totožné.

OBSAH

ÚVOD

V dnešní době se s rostoucí popularitou umělé inteligence a počítačové grafiky setkáváme stále více s fenoménem tzv. "deepfake", což je manipulace obrazového nebo zvukového záznamu pomocí strojového učení, která umožňuje vytvoření realistického podvrhu, který je těžké odhalit. Tento jev se stává stále větším problémem ve společnosti, protože může mít vážné následky v oblasti politiky, obchodu, médií, zabezpečení a osobních vztahů. Z hlediska struktury je práce členěna na teoretickou a praktickou část.

Teoretická část je rozdělena na dvě hlavní kapitoly. Jedná se o docela moderní slovní spojení, které zaznamenává největší ohlasy v současnost. První kapitola teoretické části pojednává o základní rešerši v oblasti právních norem týkajících se technologie deepfake, vymezení základních pojmů a dále také popis, co technologie deepfake je. Další důležitou součástí práce jsou techniky tvorby deepfake materiálů. Bude stručně popsáno, jak takové materiály vznikají a co je k tomu potřeba. Druhá kapitola v teoretické části studie je stručný popis historie vzniku deepfake a následně aplikace takové technologie. Za jakým účelem se tyto materiály vyrábí a příklady využití deepfake technologie.

Praktická část je rozdělena do tří kapitol. V první kapitole je popsán styk široké veřejnosti s pojmem deepfake. Tato kapitola je zaměřena na vytvoření modelových situací, při kterých došlo ke zneužití deepfake technologie a jsou popsány důsledky, které mohou nastat. Jsou zde vytvořeny i návrhy opatření ke zvýšení odolnosti subjektů v kontextu modelových situací. V druhé kapitole jsou vypsány vybrané aplikace pro tvorbu deepfake pomocí mobilní a desktopových aplikací. V poslední kapitole jsou provedeny řízené rozhovory, jejichž úkolem je zjistit, jak tento pojem vnímá veřejnost.

Hlavním cílem bakalářské práce je vytvoření modelových scénářů hrozby deepfake ve vybraných oblastech souvisejících s ochranou obyvatelstva. Stanoveny jsou i dílčí cíle bakalářské práce, a to provést rešerši předmětné problematiky, navrhnout opatření ke zvýšení odolnosti subjektů v kontextu modelových scénářů a realizovat řízené rozhovory s vybranými subjekty za účelem zjištění znalostí problematiky deepfake.

I. TEORETICKÁ ČÁST

1 ZÁKLADNÍ POJMY

S faktem, že ne všechno, co najdeme nebo si přečteme na internetu, je pravda, musí souhlasit snad každý z nás. Internet není regulovaným prostorem, proto se v něm mohou velice rychle šířit falešné informace, které mají různý důvod a dopad. Nejhorší z nich proti sobě dokážou poštvat skupiny obyvatel, ať již jsou či nejsou tvořeny záměrně. Jedním z jevů, který tento fakt potvrzuje, je novinka dnešní doby, tzv. deepfake. Jejich historie není příliš dlouhá, objevily se až okolo roku 2017.

Protože se jedná o relativně nový pojem, akademický výzkum je dosud v této oblasti limitovaný. Z toho důvodu zatím neexistuje jednotná definice, která by přesně popisovala tento fenomén jako celek. Nicméně již existují některé částečné pokusy o zachycení této problematiky. Tato technologie „využívá algoritmy strojového učení k vkládání obličejů a hlasů do obrazových a zvukových záznamů skutečných lidí a umožňuje vytvářet realistickou napodobeninu skutečnosti. Konečným výsledkem je realisticky působící video nebo zvuk, které utváří zdání, že někdo něco řekl nebo udělal.“ (Chesney, Citron 2018)

Deepfake představuje riziko, které dokáže napáchat spoustu škod, a to jak z hlediska kybernetické bezpečnosti, šíření fake news či ochrany osobnosti. Hrozbu pak představují také pro podnikatele, jejich ekonomické cíle. Mezi oběti deepfake patří například bývalý prezident Spojených států amerických Barack Obama, britský premiér Boris Johnson, zakladatel Facebooku Mark Zuckerberg či herečka Emma Watson.

„Pojem znamená doslova hluboké lži postavené na zneužití umělé inteligence k podvodu. Jde o realistické foto nebo video montáže, které jsou možné díky pokroku ve strojovém učení. Zjednodušeně řečeno, tvůrce deepfake zadá do příslušné aplikace obrázky osoby, kterou chce „rozmluvit“, aplikace si zapamatuje mimiku její tváře a další typické projevy vybraného člověka, naučí se je věrohodně zpracovat a prezentovat. Autor falešného videa díky tomu může nechat například předního politika říkat věci, které ve skutečnosti nikdy nevyslovil.“ (Gregor, Vejvodová 2018)

Deepfake využívá algoritmy strojového učení pro vkládání obličejů a hlasů do obrazových a zvukových záznamů skutečných lidí. „Deepfake využívá výkonné techniky strojového učení a umělé inteligence k manipulaci nebo generování vizuálního a zvukového obsahu s vysokým potenciálem uvést příjemce v omyl.“ (Kietzmann, 2020)

Tím se vytvoří realistická napodobenina skutečného člověka. Napodobenina může být ve formě fotografie, videa či zvukové nahrávky. Vytvoření deepfake není nijak složité,

v současné době existuje spousta programů či aplikací, které je možné použít, stejně jaké podrobných návodů na internetu, kterými je možné se řídit. (Reilly,2018)

Výsledkem je obsah, který spojují dvě fakta: je uměle vytvořený a je realistický a přesvědčivý. Tento výsledek může mít audiovizuální obsah, vizuální obsah či obsah auditivní. Vzhledem k pokrokům v oblasti umělé inteligence, tvorby videí a tzv. online trollingu, představuje deepfake skutečnou hrozbu pro mnoho oblastí. Tato krize dezinformací a tvorby deepfake může být nazývána jako „infokalypsa“ (Shick, 2020)

1.1 Legislativa

Jelikož se jedná o poměrně nový fenomén, neexistuje zatím žádná platná legislativa, která by se přímo problematikou deepfake zabývala. Nezbývá tedy než doufat, že se v brzké době dočkáme právní úpravy v této záležitosti. Obecně má však legislativa České republiky v těchto věcech poměrně dlouhou reakční dobu.

V dubnu 2021 představila Evropská komise návrh Nařízení Evropského parlamentu a Rady Evropské unie, který stanoví pravidla využívání umělé inteligence a mění jisté dosavadní právní akty Evropské unie. (Evropská komise, 2021)

Tento návrh nařízení se sice nezmiňuje specificky o deepfake, ale věnuje se transparentnosti systémů umělé inteligence a vyžaduje, aby byl veškerý uměle vytvořený, upravený nebo manipulovaný obsah zveřejněný s informací o svém původu. Článek 52 tohoto Nařízení popisuje povinnosti týkající se transparentnosti systémů umělé inteligence. Tento článek však obsahuje výjimku, která umožňuje použití technologií za účelem prevence, vyšetřování, odhalování nebo stíhání trestných činů, pokud to zákon umožňuje. Stejná výjimka platí i pro případy, kdy je použití technologií nezbytné pro ochranu svobody projevu, umění a vědy zaručené Listinou základních práv EU, s ohledem na záruky práv a svobod třetích stran.

Evropský parlament je jednou z prvních institucí, která formulovala doporučení o pravidlech pro umělou inteligenci. Doporučení Evropského parlamentu se týkají práv duševního vlastnictví, občanskoprávní odpovědnosti i otázky etiky. Cílem pak je, aby se Evropská unie stala lídrem ve vývoji umělé inteligence.

„Legislativní iniciativa, (…), vyzývá Evropskou komisi, aby předložila nový právní rámec, který by stanovil etické zásady a právní povinnosti týkající se umělé inteligence. Ten by se pak uplatňoval při vývoji, zavádění a používání umělé inteligence, robotiky a souvisejících technologií v EU, včetně softwaru, algoritmů a dat. (…) Budoucí právní předpisy by se měly

řídit několika principy: umělá inteligence musí být vytvářená člověkem a zaměřená na člověka, je nutné zajistit bezpečnost, transparentnost a odpovědnost, vytvořit záruky proti stereotypům a diskriminaci, zabezpečit právo na nápravu, dbát na sociální a environmentální odpovědnost a chránit soukromí a osobní údaje.“ (Evropský parlament,2020)

V realitě by ovšem i v rámci těchto výjimek mohlo docházet k jejich zneužívání. Realizace povinností zveřejnění informace o umělém vytvoření obsahu či o manipulaci s obsahem by mohla být také problematická, jelikož v návrhu nařízení není specifikováno, jak by tato informace měla vypadat a jaké informace by měla obsahovat. Toto nařízení uvádí v článku 70, že upravený sdílený obsah má být označen a má být uveden jeho umělé vytvoření. V praxi to může znamenat velice nízkou transparentnost, požadavky dané nařízením budou splněny jen do nutné míry. Někteří občané si tak této informace nemusí ani všimnout. Nařízení Evropského parlamentu a Rady Evropské unie nicméně stanovuje také sankce za nedodržení nařízení.

Článek 71 odst. 4 návrhu Nařízení uvádí, že správní pokuty jsou uloženy v případě nesouladu systému umělé inteligence s požadavky nebo povinnostmi uvedenými v tomto nařízení s výjimkou těch, které jsou stanoveny v článcích 5 a 10, a to až do výše 20 000 000 EUR, nebo v případě, že se dopustí porušení společnost, a to až do výše 4 % jejího celkového celosvětového ročního obratu za předchozí finanční rok, dle toho, která z hodnot je vyšší. Sankční hrozba by tak mohla uživatele a tvůrce deepfake a jejich využívání, které by bylo v rozporu s pravidly či zákony, od tohoto jednání odradit.

Stejný článek, odstavec 1 dále uvádí, že členské státy v souladu s podmínkami uvedenými v tomto nařízení stanovují pravidla ukládání sankcí a správních pokut za porušení nařízení a přijímají veškerá opatření nezbytná k zajištění jejich řádného uplatňování. Sankce, které jsou stanoveny, musí být účinné, přiměřené a odrazující. Mají zohledňovat zájmy malých a začínajících podniků a jejich ekonomickou životaschopnost.

Můžeme tedy předpokládat, že unijní úpravu této problematiky, pokud bude členskými státy přijata, bude ve svém právním řádu reflektovat také Česká republika. (Evropská komise,2021)

Ve vztahu k deepfake není zmiňována zásada ochrany osobních údajů, nicméně ve zprávě k návrhu nařízení se uvádí, že má být zajištěn soulad s Listinou základních práv EU a právními předpisy EU o ochraně údajů, konkrétně Nařízení o ochraně osobních údajů 2016/679 – GDPR), ochraně spotřebitele apod.

V americkém státě Virginii existuje zákon, který zakazuje šíření sexuálního nebo pornografického obsahu s cílem vydírání, zastrašování, obtěžování nebo jiného zneužívání. Tento zákon také zahrnuje deepfake videa – falešná videa vytvořená pomocí umělé inteligence. Některé platformy, jako Reddit a PornHub, zakázaly šíření deepfake videí. Twitter a Facebook se snaží bojovat proti dezinformacím a spolupracují s organizacemi zabývajícími se odhalováním nepravdivých informací na internetu. Google se také angažuje při detekci deepfake videí a všechny tyto společnosti usilují o vyvinutí softwaru, který by dokázal identifikovat falešná videa a zabránit jejich šíření.

Vytvoření deepfake pornografického videa je zásahem do soukromí občana a zároveň do jeho práva na informační sebeurčení. Toto právo je garantováno v čl. 10 odst. 3 Listiny základních práv a svobod českého právního řádu, které obsahuje definici toho, jaké informace, komu, za jakých okolností a jakým způsobem jsou šířeny a sdíleny, ale také složku interní, která předpokládá individuální sebeurčení jednotlivce, jeho vnitřní kapacitu pro sebeuvědomění a možnost prožití života dle své vlastní vůle. Z tohoto hlediska pak mohou deepfake představovat zásah do práva jedince na informační sebeurčení, tzn. do jejich práva na respektování jejich lidství.

Vytvoření deepfake pornografického videa může ohrožovat také práva na ochranu osobnosti člověka ve smyslu §81 až §90 zákona č. 89/2012 Sb., občanského zákoníku. Konkrétně se jedná o manipulaci s fotografiemi, které jsou volně dostupné za účelem vytvoření deepfake videa. Dochází tak k nedovolenému použití volně dostupné fotografie, podoby jedince. Jakékoli nakládání s fotografiemi druhých lidí, a to jak za účelem vytvoření pornografického obsahu, tak za účelem vytvoření jakéhokoli obsahu, manipulace s fotografiemi a podobou jedince, je nemorální, a to nehledě na skutečnost, zda je či není sdíleno dále, ačkoli míra morálky se v tomto případě může lišit v ohledu na výsledku či účelovosti deepfake videa či fotografie.

1.2 Používané techniky

S nárůstem a modernizací technologií budeme brzy ovládat technologie jako je strojové učení, automatické hlasové systémy podobné lidem, falešná videa a obrázky tvořené umělou inteligencí, stejně tak bude v nedaleké budoucnosti běžná rozšířená a virtuální realita. Deepfake je technologie, která využívá umělou inteligenci k produkci a úpravě obsahu videa nebo obrázku tak, aby bylo vidět něco, co se nikdy nestalo. (Young, 2020)

Deepfake může být zpracován několika různými technikami, různými formami. Pro účely této práce využijeme rozdělení Kietzmanna a kol. (2020).

Zvuková forma – Deepfake dokáže upravovat či zmanipulovat také zvukové záznamy. Manipulace je možná dvěma způsoby, změnou či imitací hlasu dané osoby či použití hlasu dané osoby pro transformaci textu v řeč. Obě možnosti lze stejně dobře zneužít jako využít ku prospěchu. (Kietzmann, 2020)

Obrazová forma – Jedná se o záměnu obličejů různých osob či jinou manipulaci s obličejem. Obličeje je možné upravovat či je nahrazovat za obličeje jiné.

Video forma – rozlišují se čtyři základní formy využití videozáznamu: výměna či nahrazení obličeje, morfování obličeje, ovládání celého těla a synchronizace rtů. Poslední uvedená možnost zahrnuje kromě vizuálních efektů také auditivní část. Při synchronizaci rtů dochází k úpravě pohybu rtů, slov i výrazů v obličeji tak, aby bylo video i zvukový záznam věrohodné. Výměna a nahrazení obličeje spočívá v nahrazení jednoho obličeje za obličej jiný, při morfování obličeje dochází ve videu k přechodu z jednoho obličeje na druhý, při ovládání celého těla je pak záznam celého těla přenesen na tělo jiného člověka. (Kietzmann, 2020)

Deepfake technologie fungují na několika principech, které si pro snazší porozumění této problematice definujeme. Prvním z těchto termínů je umělá inteligence. Ta je popisována jako „schopnost digitálního počítače či robota ovládaného počítačem plnit úkoly běžně spojované s inteligentními bytostmi.“ (Copeland,2022)

Zpravidla se jedná o činnosti jako jsou učení, ovládání jazyka, rozhodování, vnímání či řešení problémů. Jedná se tak o počítačové systémy s inteligentním chováním. (Copeland, 2022)

Podle vlastností programů se umělá inteligence dělí na tři základní kategorie: slabá, silná a super umělá inteligence. Slabá umělá inteligence se soustředí na provedení jednoho úkonu na vysoké úrovni. Kvalita provedení je pak vyšší než u člověka. Všechny současné systémy spadají právě do této kategorie. (Olckers, 2020)

Silná umělá inteligence dokáže vykonávat více úkolů na vysoké úrovni, dokáže myslet velice komplexně. Super umělá inteligence pak překračuje lidský kognitivní výkon komplexně. Ze super umělé inteligence panují globálně největší obavy. (Olckers, 2020)

Strojové učení je druhým konceptem, který s technologiemi deepfakes souvisí. Jedná se o součást umělé inteligence, zabývá se procesem učení a vyhodnocováním dat za účelem vytvoření přesných rozhodnutí. Umělá inteligence dokáže data analyzovat, vyhodnocovat i dále uzpůsobovat tak, aby došla k lepším výsledkům, aniž by se do procesu jakkoli zapojil člověk. (Olckers, 2020)

Hluboké učení je dalším z konceptů, který se podobá učení strojovému, má však více vrstev, které v datech dokáží rozpoznat vzorce, díky čemuž lépe chápe preference člověka. Je tvořená algoritmy, které se skládají z vrstev, které jsou vzájemně propojeny. Než se data z jednotlivých vrstev předají do jiných vrstev, jsou vyhodnocená a zpracována. (Olckers,2020)

GAN je posledním z termínů, které s deepfake souvisí. Jedná se o systém, v kterém jsou použity dvě neuronové sítě zároveň. První síť je generátor, který data čerpá, druhá síť je diskriminátor, který hodnotí míru využití generátoru. (Chesney, Citron 2018)

Vytvoření deepfake videa je založeno na 3 krocích:

1) **Extrakce** – rozdělí video na snímky a detekuje obličeje, čímž rozsekává video na obrázky. Tento proces trvá pár minut. Poté následuje kontrola snímků a vymazání rozmazaných obličejů, které by kvalitu videa snižovaly.
2) **Trénování** – používá obrázky pro trénování neuronové sítě. Vstupem i výstupem jsou dva obličeje ve stejné pozici, na stejném světle, se stejným výrazem. Tento krok může trvat až několik hodin či dní a je pro výpočetní techniku poměrně náročný.
3) **Konverze** – využije výstup na konkrétním videu. Výsledkem je video se zaměněným obličejem. (Záruba, 2020)

2 HISTORIE A APLIKACE DEEPFAKE TECHNOLOGIE

Popsaná kapitola hovoří o historií deepfake technologie a popíše známý vývoj, který umožnil vzniku uvedené technologie. V další části jsou popsány skutečné případy, ve kterých byla tato technologie zneužita k získání pozornosti nebo za účelem splnění cíle.

2.1 Historie

Historie předchůdců deepfake je díky jejich použití ve filmovém průmyslu poměrně dlouhodobou záležitostí. Ačkoli se o metodách deepfake hovoří až od roku 2017, manipulace se zvukem či obrazem probíhala již o několik desetiletí dříve. Technologie, které využívaly neuronové sítě za účelem úpravy videí se začaly používat koncem devadesátých let minulého století. Můžeme tak konstatovat, že zásadním milníkem byl rok 1997, kdy byla vyvinutá inovativní technologie zvaná Video Rewrite Program. Tato technologie měla za úkol automatizovat práci filmových studií v této oblasti. Již jeho předchůdci dokázali upravovat obličeje, využívat techniky 3D a ze zvukového záznamu oddělovat text. Tento program spojil tyto funkce, dokázal je zautomatizovat a tím dosáhnout významných výsledků. Aplikace dokázala v původním videu upravit pohyb rtů člověka tak, aby vyslovovala slova, která měla v novém videu říkat. Tento software se z původního videa naučit se artikulaci člověka, rozloží řeč člověka na fonémy, změní pořadí jednotlivých úseků, aby korespondovaly s fonémy v nové zvukové stopě. (Song, 2019)

V roce 2001 došlo k inovaci a zefektivnění tohoto procesu, když byl vyvinut model pro detekci obličeje. Díky zlepšování systémů a technologickému postupu vpřed se také proces manipulace s obličeji posouval kupředu. V roce 2016 byl pak vyvinut program Face2Face, který dokáže v živém vysílání upravit výrazy živého jedince na tvář jedince jiného. Tento program zatím nedokázal upravovat zvukový záznam, dokázal však rozeznat výrazy úst podle fotometrické konzistence a v novém videu tak deformovat výraz úst. Běžný uživatel se i přesto mohl snadno naučit upravovat fotografie druhých lidí, tato technologie tak byla předpokladem pro vznik dalších deepfake technologií, navíc se stal velmi snadno dostupným pro běžné uživatele internetu a chytrých technologií. (Song, 2019)

O rok později vznikl program Synthetizing Obama. Ten fungoval na stejném principu jako Video Rewrite Program, byl však mnohem pokročilejší než jeho předchůdce. Byl již schopen vytvářet výraz obličeje na základě zvukového záznamu.

Everybody Dance Now je software, který aplikuje pohyb celého těla z jednoho videa na druhé. Nejdříve detektorem detekuje pohyby postav na videích, výšku kostry, pozici kotníků a ty pak v jednotlivých snímcích nového videa upraví.

Rok 2017 byl také označen jako počátek technologie deepfake a to, když uživatel jménem deepfakes zveřejnil na platformě Reddit pornografické video herečky Gal Gadot, které bylo velmi propracované, její výrazy v obličeji, mimika i pohyby rtů vypadaly jako by se jednalo o skutečné video. Celý tento fenomén je tedy postaven na vzniku pornografických nahrávek známých osobností. (Chesney, Citron, 2018)

Koncem roku 2017 se objevila další pornografická videa, která znázorňovala další známé osobnosti jako např. Scarlett Johannson, Emma Watson, Michelle Obama či Ariana Grande jejichž obličeje byly přidány do pornografických nahrávek. Autor těchto videí zveřejnil také kódování, kterým byla videa upravena, díky čemuž se technologie deepfake staly dostupné i pro běžné uživatele. Pornografická videa tvoří převážnou většinu deepfake tvorby, přesto se užívají také k jiným účelům, a to jak k zábavným účelům, tak k účelům filmového průmyslu. (Chesney, Citron, 2018)

V následujícím roce se na trhu objevila aplikace, která deepfake technologii přímo využívala: aplikace FakeApp. Tato aplikace funguje na intuitivním základě, díky kterému běžný uživatel nepotřebuje žádné speciální dovednosti, aby deepfake video vytvořil. K této aplikaci se záhy přidružily také aplikace FaceSwab a DeepFaceLab. Ve videu je obličej jedné osoby nahrazen novým zadaným obličejem. Podobnou aplikací je Zao, která se objevila krátce po aplikaci FakeApp. (Song, 2019)

Výstupy z těchto aplikací nejsou dokonalé, pozorný divák odhalí nepřesnosti, přesto je nutné uznat vývoj, který tyto technologie mají a do budoucna je možné předpokládat, že kvalita deepfake materiálů vytvořených přes aplikace se bude zlepšovat a dále vyvíjet. Díky aplikacím se deepfake technologie staly dostupné takřka všem. Stejně tak může být v současné době cílem deepfake videa každý z nás. Tato videa či fotografie mohou být užity pro vydírání, krádeže, šíření pomluv si jiné trestné činnosti.

2.2 Aplikace využití deepfake technologie

Jak již bylo uvedeno výše, deepfake představuje pro společnosti několik hrozeb. Jedna z nich spočívá v ohrožení práv na ochranu osobnosti. Deepfake se ve společnosti vyskytuje teprve několik let, přesto se objevilo již několik případů, kdy byla práva člověka omezena či ohrožena. Nejčastěji diskutovanou problematikou je využití této technologie pro vytváření pornografických videí, ve kterých figurují známé osobnosti i běžní občané. Statistiky

uvádějí, že pornografická videa tvoří až 96 % veškerých deepfake videonahrávek. (Dvořáková, 2020)

Technologie deepfake spočívá ve vytvoření ultrarealistických videí osob, jejichž hlavy jsou nahrazeny hlavami jiných osob, takové metodě se říká face swap. Strojové učení umožňuje velmi uvěřitelné splynutí obrazu, ale i převzetí pohybů, drobné mimiky a gest vyobrazené osoby. Ve výsledku tak stačí nakombinovat cílové video se snímky a videozáznamy (tzv. face-setem“) osoby, kterou do cílového videa chceme zakomponovat. (Harris, 2019)

Vznikne tak cílová pornografická nahrávka, která u sledujícího vytvoří dojem, že se jedná o někoho, u koho bychom natočení takového typu videí neočekávali. „Deepfakes jsou přesvědčivá videa a obrázky lidí, kteří na nich dělají nebo říkají věci, které nikdy neudělali ani neřekli.“ (Silbey, Woodrow, 2019)

Obětí deepfake pornografických videí se z převážné většiny stávají ženy, technologicky zdatní uživatelé internetu vytvoří pornografické nahrávky žen, které by si přáli vidět v těchto situacích a ve většině případů nahé. Nejčastěji se jedná o celebrity, jejichž face sety jsou velmi snadno dohledatelné na internetu. Výjimkami ovšem nejsou ani ženy, které slavné nejsou. Jedná se zpravidla o kamarádky, známé či bývalé přítelkyně tvůrců těchto videí. Vytvoření takového videa je, jak bylo uvedeno výše, jen otázkou dostupnosti materiálů, face setů a aplikací či programů, ke kterým jsou návody dostupné na internetu. Ne všechna takto vytvořená videa jsou ostrá či dokonalá, dá se však předpokládat, že s rozvojem technologie v budoucnu bude docházet také k vylepšení těchto videí. (Dvořáková, 2020)

Nelze však než konstatovat, že ať je účel deepfake videí s pornografickým zaměřením jakýkoli, jedná se o nedobrovolnou pornografii šířenou po internetu bez vědomí či svolení ať už původních tvůrců či osob, které jsou do videí přidány umělou inteligencí. Výjimkou nejsou ani tzv. Revenge-Porn, které jsou vypouštěny za účelem někoho poškodit.

V této souvislosti je pak třeba stručně zmínit také video, ve kterém Barack Obama kritizuje svého nástupce v americkém prezidentském úřadu Donalda Trumpa. Ve druhé polovině této nahrávky je pak uvedeno, že se jedná o podvrh a jeho tvůrci upozorňují na hrozbu deepfake pro společnost. (Fagan, 2018)

Deepfake videa mají potenciál stát se hrozbou pro celou společnost. Proti zneužívání této technologie se dá bojovat nejrůznějšími legislativními opatřeními, vzděláváním, ale i systémy pro kontrolu multimediálních záznamů a detekci falešných nahrávek. Tato videa dokážou ohrozit nejen kariéry známých osobností, jak by se dalo domýšlet z příkladů uvedených výše, ale může být ohrožená také reputace a majetek ekonomických subjektů, politické zájmy či zájmy jednotlivých států. Za pomocí deepfake může dojít k odcizení

tajných firemních dat či majetku. Autoři deepfake si mohou najít skulinku v bezpečnostních postupech firmy, a tak jim jsou různá data naservírována takřka pod nos.
Dá se proto očekávat, že takových a podobných případů bude s rozvojem technologií přibývat, ať již se bude jednat o fake news, či zneužití osobnostních práv jednotlivců, známých osobností, politiků, či další podvody. Je proto třeba dbát zvýšené pozornosti a hrozbu deepfake nepodceňovat. Rostoucí kvalita deepfake videí znamená, že v budoucnu bude čím dál tím obtížnější oddělit falešné zdroje informací od skutečných, a to jak pro běžné občany, tak na příklad pro novináře či badatele.

Deepfake ovšem nemají jen negativa a nedochází pouze k podvodnému jednání, technologie deepfake může totiž i být velmi nápomocná. Tato technologie se dá využít například pro lidi, kteří z různých důvodů přišli o hlas, a tím o schopnost mluvit. Díky těmto technologiím by mohli svůj hlas, ačkoli v umělé podobě, získat zpět. Dalším příkladem této technologie je na příklad Siri, virtuální asistentka. Podobná technologie by se dala využít na příklad při dabování filmů či čtení audioknih. V současné době pracuje britská společnost Synthesia na vytvoření takové umělé inteligence, která by synchronizovala pohyb rtů s tím, co říká dabér. Diváci filmů by tak měli příjemnější zážitek. Podobně pak pracují společnosti na zvýšení rozlišení záběrů s nízkým rozlišením. Tím by se mohly vylepšit starší filmy či videohry. (Weghe, 2019)
Dalším příkladem využití deepfake technologií je reklamní průmysl, kde mohou být použity pro marketingové účely. Pozitivním příkladem je na příklad vytvoření Salvadora Dalího za účely přivítání hostů v Muzeu Salvadora Dalího na Floridě. Většina deepfake fotografií či videí slouží však k účelu pobavení, když ukazují známé osobnosti v situacích, ve kterých se nikdy neobjevily.
Na těchto příkladech je tak dobře vidět, že deepfake může sloužit také k pozitivním účelům a nemusí jen škodit.

2.3 Příklady využití deepfake technologie

V současné době už nemůžeme mluvit o možnosti zneužití technologie deepfake jako pouze o teoretické hrozbě. Během posledních let se totiž stalo několik událostí, u kterých byla snaha využitím deepfake dehonestovat a zesměšnit vysoce postavené politiky, reportéry různých novin nebo taky ohrožovat integritu státu jako takového. Proto zde bude uvedeno několik příkladů z historie, které se opravdu staly.

1. Cílem několika útoků se stala americká demokratická politička a poslankyně Sněmovny reprezentantu Spojených Státu Nancy Patricia Pelosiová. Jeden takový útok se stal v roce 2019, kdy na sociální sítě Facebook a Youtube bylo zveřejněno video, ve kterém je uvedená poslankyně pod vlivem alkoholu. Podobné video sdílel i bývalý prezident Spojených státu amerických Donald Trump. Obě videa byla vyhodnocena jako falešná, kdy u jednoho bylo video zpomaleno o 75 %, a u druhého byl nahrazen obličej poslankyně za jinou osobu. Za krátkou dobu, co bylo video vystaveno na uvedených sítích si ho stihlo přehrát až 2 500 000 lidí. (CBS, 2019)

2. Oblíbeným cílem útoku je jeden z nejznámějších herců dnešní doby Tom Cruise. V roce 2021 bylo na platformě TikTok zveřejněno několik videí zobrazující tohoto hollywoodského herce. V únoru roku 2021 bylo zveřejněno jedno z prvních videí, na kterém je vidět herec jako hráč golfu. Na dalším videu je herec předvádějící trik s mincí. Jako méně neškodné se dá považovat video, ve kterém je herec ukázán jako kandidát na prezidenta Spojených státu amerických ve volbách v roce 2020. Uvedená videa byla mnohem kvalitnější než příklad uvedený jako první. Během několika dní tyto videa shlédlo několik miliónu lidí. (Marr, 2022)

3. Jednou z obětí, kdy byla tato technologie použita k zastrašování a vydíraní je indická novinářka píšící pro Washington Post, a to Rana Ayyub. Jedná se o novinářku, která vždy ráda psala o kontroverzních tématech, které byly směřovány i na indickou vládu. Takové vyjadřování se jí ale v dubnu roku 2018 nevyplatilo. Ayyub se totiž ostře ohradila proti jedné z indických politických stran, která měla vyjadřovat podporu podezřelému v případu znásilnění 8leté holčičky. Den na to se objevilo několik příspěvků ze sociální sítě Twitter, ve kterých je vidět účet Rany Ayyub a zprávy ve kterých je napsáno – „Miluji Pákistán, Nenávidím Indii a Indy“ a další obdobné příspěvky. Následující dny situace eskalovala, kdy v určitých politických kruzích začalo kolovat video, ve kterém měla Ayyub být v sexuálním aktu. Následně dokonce na veřejnost uniklo její telefonní číslo a začala dostávat výhružné a znepokojivé zprávy přes aplikaci Whatsapp. Vše výše uvedené dohnalo žurnalistku až ke zdravotním problému, kdy musela být následně hospitalizovaná v nemocnici. I přes to, že reportérka byla podat trestní oznámení a příslušné úřady se takovým

útokem zabývaly, tak podobné návrhy, osočování a zastrašování nepřestalo. Sama uvádí „*Od té doby, co bylo zveřejněno takové video, nejsem stejná osoba. Dříve jsem měla velmi silné názory, nyní jsem mnohem opatrnější ohledně toho, co zveřejňují online. Musela jsem se sama z nutnosti hodně cenzurovat. Neustále přemýšlím, co když se mi něco podobného stane znovu.“* (Ayyub, 2018)

Výše uvedené příklady jsou jen výběr z mnoha případů zneužití deepfake technologie. Rychlý technologický pokrok a společenské změny vedou ke vzniku nových situací, na které je třeba reagovat. Je důležité být otevřený novým myšlenkám a schopnost přizpůsobit se novým hrozbám, které přináší moderní svět.

DÍLČÍ ZÁVĚR

Teoretická část bakalářské práce se zabývá vysvětlením základních pojmů spojených s technologií, a hlavně problematikou deepfake. V další části je popsán legislativní rámec ohledně problematiky deepfake. Tato kapitola je důležitá z důvodu toho, že jsou zde vypsány nařízení Evropské Unie, které tuto problematiku definují. V České legislativě tato problematika prozatím není výslovně zakotvena. Dále v této kapitole je podstatné to, že zneužití této technologie může mít vážné následky a právní rámec v České republice není prozatím připraven se řádně vypořádat s tímto problémem. V další časti je popsána technická část tvorby deepfake médií. Jsou zde popsány různé části vytváření deepfake videí nebo fotografie jako třeba strojové učení nebo taky GAN. V poslední části teoretické části bakalářské práce je popsána historie technologie deepfake, převážně v jakém období taková technologie vznikla a jakými událostmi se dostala do podvědomí široké veřejnosti. Dále jsou stručně sepsány některé příklady událostí, ve kterých byla technologie deepfake použita, či zneužita za účelem zviditelnění se, získání pozornosti nebo také dosažení určitého cíle.

II. PRAKTICKÁ ČÁST

3 PŘÍKLADY ZNEUŽITÍ A NÁVRHY OPATŘENÍ

V kapitole „Příklady zneužití a návrhy opatření“ se budeme věnovat možnostem kontaktu běžného uživatele s technologií deepfake a hlavně budou projednány různé oblasti společnosti, ve kterých je velké riziko zneužití technologie deepfake.

3.1 Kontakt běžného uživatele s deepfake technologií

Fenomén deepfake v posledních letech získává více pozornosti. Zcela jistě se tak dá říct, že má potenciál, že bude využit v různých průmyslových odvětvích, a to především v zábavním průmyslu, ve filmovém průmyslu, vzdělávacích videích, hrách, sociálních médiích, v digitální komunikaci, ale také ve zdravotnictví, v módě, elektronickém obchodě apod. Deepfake na sebe poprvé upozornil teprve před pár lety, kdy byla na sociální síti Reddit zveřejněná pornografická videa s tvářemi známých osobností.

Deepfake dokáže manipulovat s veřejným míněním. Ve Spojených státech dokonce FBI varovala v souvislosti s prezidentskými volbami v roce 2020, že dojde k akcím, které mají za úkol ovlivnit veřejné mínění pomocí deepfake videí. Tato videa měla sloužit pro zkompromitování jednoho z kandidátů, zmatení veřejnosti, voličů. Tyto obavy se ve výsledku ukázaly jako plané, v souvislosti s volbami se objevila pouze dvě transparentní videa, která měla navíc za účel varovat veřejnost před šířenými dezinformacemi. Přesto se objevily ve značné míře také fake news, u nich ale nedošlo k použití umělé inteligence pro tvorbu deepfake videí.

V roce 2018 se v Gabonu na veřejnost dostalo video, na kterém se objevuje prezident země Ali Bongo, který se dlouho neukázal na veřejnosti a předpokládalo se, že je mrtvý. Tato skutečnost byla předpokladem pro pokus o vojenský převrat v zemi. Nepravost tohoto videa však nebyla potvrzena.

V roce 2022 se objevilo video s ukrajinským prezidentem Volodymyrem Zelenským, který ukrajinské vojáky vyzývá, aby složili zbraně. Sám Zelenskyj pravost tohoto videa velice brzy popřel, sociální sítě ve velkém tato videa mazala. Toto video není kvalitně provedené, posoudit jeho pravost nebylo příliš složité – Zelenskyj je ve videu nepohyblivý, pohybují se jen jeho mimické svaly. Rychlá reakce sociálních sítí však byla ukázková, neboť toto deepfake video mohlo mír nebezpečné následky.

Mínění běžných uživatelů není ohroženo použitím deepfake, ale panují zde obavy z jejich zneužití. Technologie pro tvorbu deepfake v současné době neprodukuje natolik autentická

videa, která by vedla ke zmatení veřejnosti. Deepfake videa nemají tu moc, aby ovlivnila veřejné mínění.

V rámci politické agendy nemají momentálně deepfake videa v podstatě smysl. Nejsou natolik propracovaná, aby se nezjistilo, že se jedná o videa nepravá. Sloužit mohou k případům, kdy ukazují politika, který se delší dobu neobjevil na veřejnosti a panují obavy o jeho zdraví, či v případě krátkých, neostrých videích, kde se politik ukazuje na akci, na které nebyl. Tuto funkci však stejně dobře mohou zaručit i falešné zprávy, které veřejnost informují, přesto nepřikládají žádný důkaz, že se jedná o pravdivou informaci. Vždy se najdou čtenáři, kteří takovým falešným zprávám věří a nepotřebují k tomu deepfake videa. Jejich mínění je na základě těchto falešných zpráv však ovlivněno.

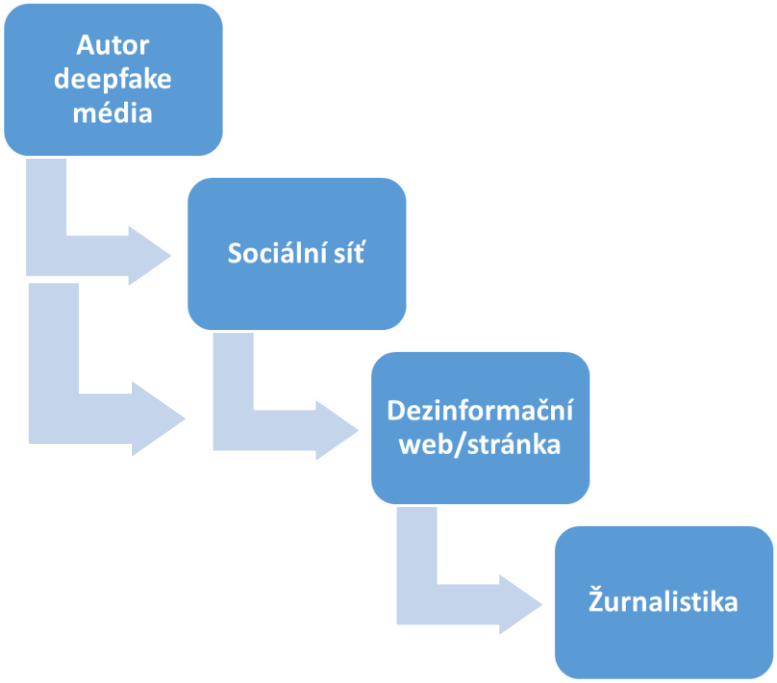
Více je falešnými videi a zprávami ohrožená reputace žen, a to kvůli využití těchto technologií pro pornografické účely. Ačkoli se první příklady deepfake zaměřovaly na pornografii či měly zábavní účel, nejde však zapomínat na jejich potenciál druhým ublížit, pomstít se, šikanovat je. Mohou také sloužit pro politickou propagandu, sabotáž, diskreditaci jednotlivců, jejich vydírání, použití falešných videí u právních sporů či manipulaci s trhem. Musíme brát v potaz rychlý rozvoj technologií, být si vědomi možných dopadů na jednotlivce i společnost a být připraveni s deepfake bojovat. Velkým problémem je využití deepfake technologií dezinformačními zdroji. Tomuto problému bude společnost v příštích letech muset čelit, videa odhalovat a mazat je, ačkoli to bude technologicky náročné.

Většinová veřejnost se tak dostane do kontaktu pouze s deepfake, který je vytvořen za účelem pobavení druhých. Běžně je možné si stáhnout některou z aplikací, které dokáží zaměnit tváře známých osobností za tváře běžných lidí. Některé z těchto aplikací jsou velice propracované a výsledky jsou realistické. Tyto výsledky poté uživatelé sdílí se svou rodinou, přáteli, známými či kolegy.

Mezi tyto aplikace, které může široká veřejnost používat, patří na příklad FakeApp, FaceSwap či DeepFaceLab. Pro použití těchto aplikací není zapotřebí mít žádné odborné znalosti. Aplikace FakeApp tak poskytovala všem prohodit si tváře s tváří nějaké slavné osobnosti či svého rodinného příslušníka, a tím vytvářet vtipné obrázky. Užívání těchto aplikací může přinášet jistá rizika. Jedná se o případné zneužití dat, jelikož uživatelé povolují přístup ke svým osobním údajům.

Zájem o deepfake technologie projevily také různé obchodní společnosti. Ty je využívají pro svá edukační videa. V roce 2020 se na příklad objevila kampaň, která cílila na bezpečnost ve školách. Za pomocí deepfake technologií v ní vystoupil student, který byl zastřelen při masové střelbě ve škole v Parklandu na Floridě.

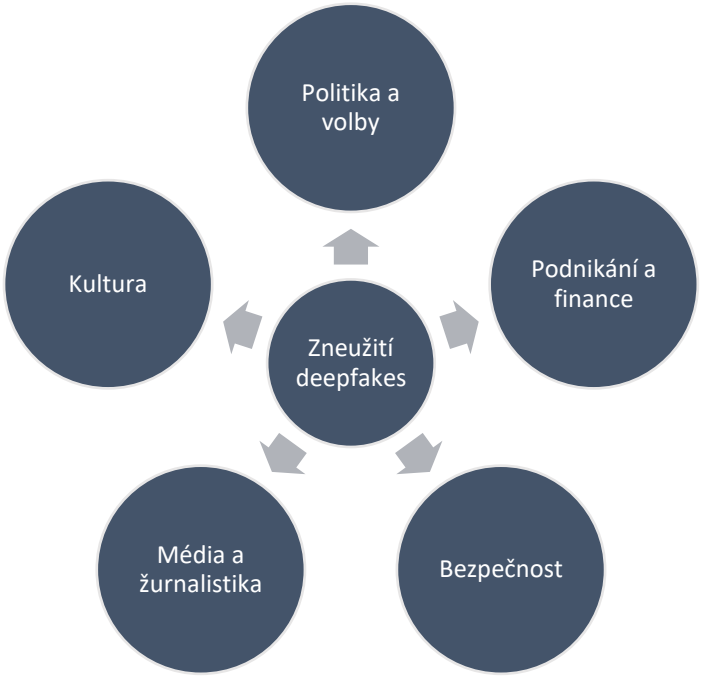
Významný potenciál mají deepfake technologie v zábavním průmyslu. Ve filmech se tak mohou objevovat i herci, kteří už nejsou mezi živými. Přínos mají Deep fake technologie i v oblasti her, obchodování, reklamě, vzdělávacích médiích aj.



Obrázek 1- Možná cesta šíření deepfake videa či fotografie

3.2 Tvorba modelových situací

Pro přechod ke tvorbě modelových situací zneužití technologie Deepfake v této kapitole byla provedena důkladná literární rešerše popsaná v teoretické části bakalářské práce. Tato kapitola se zaměřuje na to ukázat, že využití deepfake technologie není pouze abstraktním konceptem, ale že se jedná o reálnou hrozbu, která může mít vážné následky. Bude popsáno několik teoretických scénářů, které by mohly ohrozit různé odvětví společnosti a převážně zaměření na oblast ochrany obyvatelstva. Kromě toho budou popsány i možné opatření k ochraně těchto oblastí před negativními dopady.



Obrázek 2 - Oblasti zneužití deepfakes

Níže popsané modelové situace vychází hlavně z toho, že dochází k úmyslnému šíření deepfake médií k dosažení určitého negativního efektu za účelem zisku, ovlivňování mysli obyvatel, oslabení morálky či dosažení určitého stavu strachu. Vycházím z toho, jak je i popsáno v této bakalářské práci, že tvorba deepfake médií již není pouze v oblasti odborníků na výpočetní techniku, ale může ji provádět prakticky kdokoliv. V roce 2023 je velmi snadné, že určitá kontroverzní videa nebo fotografie, které budou sdíleny na sociálních sítích, nasbírají tisíce nebo dokonce stotisíce zhlédnutí za velmi krátkou dobu. Díky široké dostupnosti technologií a jejich snadnému použití se tyto obsahy mohou rychle šířit po internetu a získávat velkou pozornost uživatelů. Tímto způsobem se může kontroverzní obsah rychle a masivně šířit, což může mít značný dopad na veřejné mínění, vnímání událostí nebo chování lidí.

3.2.1 Politika a volby

Před volebním obdobím do poslanecké sněmovny České republiky se na různých sociálních sítích objeví video nebo fotografie, která údajně ukazuje kandidáty, jak říkají nějakou kontroverzní věc nebo ukazují něco naprosto nevhodného. Tyto média by se mohla šířit tak rychle po všech sociálních sítích pomocí různých známých konspiračních teoretiků, že mnoho voličů je mohlo vidět dříve, než by zobrazený kandidát měl možnost zveřejnit vysvětlení či vyvrácení. To by mohlo mít významný vliv na to, jak budou voliči vnímat takového kandidáta, jeho politickou stranu, a nakonec by to mohlo mít i vliv na výsledek voleb.
Další možností zneužití deepfake technologie by mohlo být vytvoření falešného videa, které ukazuje, jak určitý kandidát přijímá úplatky, čím se dopouští protiprávního jednání dle § 331 zákona č. 40/2009 Sb. Trestní zákoník. Pokud by se takové video rozšířilo dostatečně rychle, mohlo by mít fatální následky pro reputaci kandidáta, a nakonec i na výsledek voleb.

Opatření

- Edukace veřejnosti o existenci deepfake technologie, jejích možnostech a rizicích, která s sebou nesou. Informovat o tom, že během volebního období je velké riziko zneužití této technologie a je třeba si lépe ověřovat zdroje toho co je na internetu vidět.
- Vytvoření specializovaného týmu za politickou stranu, který se bude věnovat dané problematice. Převážně by měl být tým složený z odborníku na výpočetní techniku, dezinformace a deepfake. Tito odborníci by měli být zdatní v nástrojích pro ověřování pravosti videí a obrázku, které se šíří během volební kampaně. Takové nástroje by měly pomoci rychleji identifikovat potenciální deepfake obsah a snížit, tak šíření falešných informací.
- Spolehlivá spolupráce s bezpečnostními orgány, které by měly být obeznámeny s touto problematikou. V rámci bezpečnostních orgánů by měl být podporován výzkum a vývoj různých forenzních technologií, které by rychleji a spolehlivěji identifikovaly deepfake obsah. Bezpečnostní orgány by měly také spolupracovat s technologickými firmami, jako jsou Meta nebo Google, které se již nějakou dobu zabývají touto problematikou.

3.2.2 Podnikání a finance

Byl vytvořen falešný rozhovor s významným finančním analytikem, který byl publikován na internetu, taky se rychle šířil na sociálních sítích a různých finančních fórech. Finanční analytik, se kterým měl být veden rozhovor vydává určité doporučení k investicím. Analytik se zdál být důvěryhodný a jeho předpovědi byly prezentovány jako zaručené zisky. Rozhovor byl publikován na stránkách, které se tvářily jako oficiální finanční zpravodajské portály. Falešný rozhovor získal velkou pozornost, což vedlo k tomu, že se začal šířit i mezi mnohými potenciálními investory. Podvodníci začali využívat rozhovoru k navození dojmu vysokých zisků z investic do konkrétních akcií a komodit. Využili také různých online marketingových technik, jako jsou e-maily, reklamy a sociální média, kde se odesílatelé tvářili jako renomovaní investiční experti, kteří doporučují nákup akcií a komodit z rozhovoru. Mnoho investorů bylo nalákáno na falešné doporučení a investovalo své peníze na základě informací z tohoto falešného rozhovoru. Podvodníci takto získali finanční prostředky od mnoha lidí, kteří spoléhali na pravost rozhovoru.

Rozhovor obsahoval hluboké falzy a byl vytvořen pomocí nejmodernější deepfake technologie, která dokázala věrohodně napodobit hlas, vzhled a mimiku analytika. Finanční podvodníci, kteří za tímto podvodem stáli, tak chtěli dosáhnout zisku finančních prostředků od nevědomých investorů, čímž se mohli dopustit trestného činu dle § 209 zákona č. 40/2009 Sb. Trestní zákoník. Nicméně po nějakém čase pravost rozhovoru byla zpochybněna a objevily se pochybnosti ohledně použití deepfake technologie v rozhovoru. Když se podvod následně odhalil, mnoho investorů utrpělo velké finanční ztráty a byla celkově narušena důvěryhodnost v online zdrojích informací. Takovýto scénář zneužití deepfake technologie by mohl mít v životě vážné důsledky pro celý finanční trh.

Opatření

- Žurnalisté, média finančních portálů a investoři by měli důkladně verifikovat zdroje informací a ověřovat pravost obsahu, který je publikován na internetu nebo jiných médiích. Měli by být obeznámeni s různými druhy technických metod ověřování, jako jsou například digitální podpisy, watermarky nebo analýzy digitálních stop. Tyto metody mohou být využity k ověření, zda je obsah originální nebo byl vytvořen

za pomoci deepfake technologie. Důvěryhodnost zdrojů je klíčovým prvkem při hodnocení pravosti informací.

- Investoři by měli být vzděláváni o rizicích spojených s deepfake technologií a finančními podvody, kterých bylo v roce 2022 přes 10 000. Takové číslo se ani v roce 2023 prozatím nezmenšuje. Dále by měli být investoři vybíraví při přijímání investičního rozhodnutí a provádění finančních transakcí. Důkladná analýza a ověření informací od různých zdrojů může pomoci minimalizovat riziko padělání a zneužití pomocí technologie deepfake. Investoři by měli případné pochyby o účtech, na které chtějí poslat své finanční prostředky nebo akcie, které by chtěli nakoupit konzultovat přes svou banku.

3.2.3 Bezpečnost

V roce 2023 se objevuje nový případ zneužití deepfake technologie ve společnosti, a to ve spojitosti s Policií České republiky. Skupina technicky zručných zločinců začíná vytvářet falešná videa, ve kterých jsou fiktivně znázorněni policisté páchající různé protiprávní jednání. Tato deepfake videa jsou šířena prostřednictvím sociálních sítí a dalších online platforem pomocí sociálních botů s cílem diskreditovat Policii České republiky a zpochybnit její vážnost, autoritu a důvěryhodnost. Je vytvořeno několik různých videí za použití různých technik deepfakes. Na jednom je například osoba, která má pozměněnou tvář a mimiku, aby znázorňovala nějakého známého příslušníka Policie České republiky. Tento údajný policista se na videu chová neeticky, agresivně za použití fyzického násilí nebo nějak jinak zasahuje do osobních práv člověka. Na dalším videu je upravené audio, kdy se policista vyjadřuje hrubě a nezdvořile. Takové zneužití technologie deepfake má vážné následky. Důvěra v Policii České republiky je vážně narušena, což má za následek snížení respektu vůči práci policistů a narušuje jejich autoritu. Navíc se objevují protesty a obvinění z neetického chování policie, která je nyní na veřejnosti považována za nedůvěryhodnou a nelegitimní. Na takové chování občanů mohou pachatelé, pokud nebudou již vypátrání, čekat a důsledky mohou být několikanásobně horší.

Opatření

- Posílení speciálního útvaru, který se problematikou deepfake zabývá. Vedení Policie České republiky by mělo naslouchat příslušníkům z takového útvaru a na základě jejich poznatků a zjištění zavést nové protokoly pro ověřování autentičnosti videonahrávek a zvýšení kybernetické odolnosti policie.
- Členové speciálního útvaru se zaměřením na kybernetickou bezpečnost by měli poskytnout určitým policejním útvarům technickou podporu a školení, hlavně co deepfake je, jak případně rozpoznat deepfake videa nebo fotky. Z vlastí zkušenosti bylo zjištěno, že značné množství příslušníků Policie České republiky není o této problematice obeznámeno nebo se s tímto pojmem nikdy nesetkalo.
- Policie České republiky by se měla účastnit nebo sama vést mediální kampaň zaměřenou na informovanost a osvětu veřejnosti ohledně deepfake technologie, možného zneužití a dalších rizik spojených s takovou problematikou.

3.2.4 Média a žurnalistika

V letošním roce se v České republice objevuje skandál ohledně zneužití technologie deepfake v oblasti žurnalistiky. Jeden z renomovaných novinářů je obviněn z publikace falešného videa, které mělo záměrně zdiskreditovat politického kandidáta před blížícími se volbami do senátu.

Uvedený novinář je známý svou kritikou politických elit a byl vždy považován za důvěryhodný zdroj informací s velkým počtem sledujících na všech sociálních sítích. I z tohoto důvodu zveřejněné video, které ukazuje výše zmíněného kandidáta v kompromitující situaci nabralo na věrohodnosti. Toto video rychle získává velkou pozornost na sociálních sítích a stává se virálním. Média ho začínají šířit a zveřejňovat jako jasný důkaz korupce na politické scéně v České republice. Video rozmíchá během pár dní širokou diskusi a vyvolává určité kontroverze. Významné politické strany a skupiny volají po okamžitém vyšetřování a odstoupení onoho kandidáta.

Nicméně později se zjistilo, že video bylo deepfake. Bližším šetřením od Policie České republiky bylo odhaleno, že video bylo digitálně upraveno, tak aby vypadalo autenticky, ale ve skutečnosti bylo kompletně sestaveno z falešných obrazových a zvukových stop. Žurnalista tvrdí, že si nebyl vědom toho, že se jedná o falešné video. Tvrdí, že mu bylo posláno z anonymní adresy na e-mail.

Opatření

- Žurnalisté by měli věnovat zvýšenou pozornost ověřování zdrojů a autenticitě videa a zvuku, které používají nebo které plánují zveřejnit. Je důležité, aby si byli jistí pravostí zdroje před jeho použitím. K prověření o autenticitě videa by měli být obeznámeni s využitím různých detekčních programu na deepfake jako je například Deepware, Reality Defender nebo DeepFake Detect.
- Vytvoření certifikovaných databází s autentickými videi nebo fotografiemi, které by mohli novináři vkládat do výše zmíněných programu, a tak je využít jako zdrojový materiál pro ověření autenticity videa.

3.2.5 Zábava a kultura

Technologie vytváření deepfake videí a fotografií se za posledních pár let posunula rychle dopředu. Takový posun umožnil využití deepfake pro tvorbu audiovizuálního obsahu v kultuře. Tato technologie umožňovala režisérům a umělcům vytvářet realistické deepfake filmy, seriály, hudební video klipy a další kreativní díla, která si rychle získala obrovskou popularitu mezi diváky po celém světě.

Nicméně, rychle se objevila skupina lidí, která začala zneužívat deepfake technologií v kultuře k nekalým účelům. Jednou z nejvýraznějších situací takového zneužití se stal případ, kdy byla tato technologie použita k pošpinění reputace jedné slavné herečky. Tato herečka byla nově vycházející hvězda kinematografie, která již měla velkou skupinu fanoušků po celém světě. Jednalo se o mladou, půvabnou slečnu, která v mnohých mohla vyvolávat pocit žárlivosti. Takoví lidé se proto rozhodli použít technologií deepfake k vytvoření upravených videí, ve kterých byla znázorněna jako agresivní, arogantní a nevyzpytatelná osoba. Tato falešná videa začala rychle kolovat po všech sociálních sítích,

ačkoli byla zcela nepravdivá. Upravená videa ukazovala zmíněnou herečku ve fiktivních nebo sexuálních situacích, což mělo za cíl herečku zdiskreditovat a pošpinit tak její osobu. Herečka se snažila bránit svoji reputaci veřejným vystoupením a různými charitativními akcemi, ale bylo již pozdě. Videa se šířila velkou rychlostí a začala mít dopad na její osobní i kariérní život. Nakonec uvedla, že tyto videa jí zničila celý život.

Opatření

- Zavedení lepšího označení deepfake obsahu, kdy by mohlo být povinné označení deepfake obsahu u veškerých audiovizuálních děl v programu, ve kterém je tento obsah vytvářen. Takové označení by mohlo obsahovat zřetelné a viditelné označení v samostatném díle, které by mohlo ihned informovat diváky, že se jedná o deepfake obsah. Diváci by se následně mohli snáze rozhodnout, zda je obsah pravý nebo se jedná o deepfake obsah, což by pomohlo předejít případným omylům nebo různým výhružkám ze strany diváků.
- Potřeba získat licenci nebo povolení pro využití deepfake technologie ve výrobním procesu audiovizuálních děl. Taková licence nebo povolení by byla potřeba jen u programů, které jsou schopny vytvářet kvalitní a věrohodný obsah.

3.2.6 Zdravotnictví

Rok 2021 byl pro celý svět rokem pandemie Covid-19, což přinášelo mnoho výzev a problémů. Mezi ně patřilo i různé šíření dezinformací a falešných informací týkajících se nemoci a následné léčby. Vláda České republiky společně s lékařskou komorou se snažily situaci korigovat mnohými televizními vystoupeními, což ale napomohlo dezinformátorům získat materiál vhodný pro vytvoření deepfake videí, které by zpochybnily kvalifikaci, schopnosti a úsudek vlády České republiky i lékařů v boji proti tak nečekané pandemií. Tato skupina lidí vytvořila falešné deepfake videa, na kterých jsou fiktivně znázornění lékaři, kteří tvrdí, že Covid-19 je falešná pandemie a léčba proti nemoci je neúčinná a zbytečná. Tyto deepfake videa jsou rychle šířena pomocí skupin dezinformátorů prostřednictvím sociálních sítí a dalších platforem, čímž způsobují zmatek a nespokojenost mezi obyvateli České republiky. Těchto videí si všimnou i lidé, kteří jsou v osobním životě opatrní a v běžné situaci by takovým informacím nevěřili. Bohužel vzhledem k tomu, že se jednalo o zcela novou situaci, se kterou se většina společnosti nikdy nesetkala, tak mohli i opatrní občané

takovým videím rychle uvěřit. To způsobuje velký problém pro zdravotnické pracovníky, kteří se snaží ze všech sil informovat a ochránit své pacienty. Tyto deepfake videa mohou vést k nesprávnému podávání léku a k odmítnutí vakcíny proti nemoci Covid-19, což by mohlo mít vážné a někdy i fatální následky.

Opatření

- Lékaři by se měli snažit o větší informovanost svých pacientů o možných rizicích spojených s deepfake technologií. Pacienti by měli být vedeni k tomu, aby nedůvěřovali neověřeným zdrojům informací, ale aby hledali informace od důvěryhodných zdrojů, jako jsou oficiální webové stránky zdravotnických organizací, nebo se při jakémkoli podezření obrátili na svého ošetřujícího lékaře pro ověřené informace o zdraví.
- Lékaři a další pracovníci ve zdravotnictví by měli být opatrní při sdílení svých videí na sociálních sítích nebo jiných online platformách. Měli by umět používat některé technické možnosti, jako jsou digitální podpisy a watermarky, aby zajistili autenticitu a integritu svých videí.

3.2.7 Tísňová linka Policie České republiky

S rozvojem technologie deepfake se objevuje nová možnost narušení provozuschopnosti tísňových linek v České republice. Neznámé osoby začínají volat na tísňovou linku a vydávají se za občany, kteří se právě nachází v nebezpečné a tísňové situaci, kdy potřebují okamžitou pomoc. Tato skupina zločinců využívá nejmodernější technologie deepfake k tomu, aby se zdál být legitimním volajícím a zjistili tak informace o policejních zdrojích a způsobu, jakým policie v určitých případech pracuje. S cílem maximalizovat škody, tyto osoby také udávají nesprávné informace, aby vedli policisty k nesprávným zásahům a tím narušili jejich práci. Takové zneužití deepfake technologie má vážné důsledky, nejenže zpochybňuje integritu a důvěryhodnost tísňového volání, ale může také vést k nebezpečným situacím, kdy policisté nejsou schopni včas reagovat na skutečné situace a nebezpečí.

Opatření

- Tísňová linka by měla mít nejnovější možné technologie, které by napomohly rozeznat deepfake audio, což by mohlo minimalizovat riziko toho, že by nějaká neznámá osoba mohla pomocí deepfake technologie napodobit hlas jiné osoby.
- Operátoři tísňové linky by měli absolvovat speciální školení, které by je informovalo o možnostech a rizicích deepfake technologie. Mohou se dozvědět, jak rozpoznat potencionální deepfake hovory. Důkladná osvěta by měla zahrnovat i technologie a nové metody detekce.

Vlastní zkušeností bylo zjištěno, že technické vybavení tísňové linky Policie České republiky neodpovídá nejnovějším možnostem v technologií a panuje poměrně velká nevědomost na operačním středisku Policie České republiky ohledně této problematiky.

4 APLIKACE POUŽÍVANÉ K TVORBĚ DEEPFAKE

V této kapitole se budeme věnovat možnostem tvorby deepfake obsahu pomocí aplikací pro stolní počítače a mobilní telefony. Nejprve se zaměříme na aplikace, které jsou používané v rámci mobilních telefonů. Tyto aplikace jsou rozšířenější, ale úroveň kvality vytvořeného videa či fotografie není vysoká. Dále budou popsány desktopové aplikace. Tyto aplikace jsou méně rozšířené, ale garantují vyšší kvalitu vytvořeného obsahu.

4.1 Mobilní aplikace

V dnešní době existuje mnoho aplikací, které jsou používané na mobilních telefonech. Většina těchto aplikaci je zdarma a k dispozici ke stažení na různých platformách. Tyto aplikace jsou navrženy tak, aby byly co nejjednodušší na použití a aby jejich uživatelé nemuseli mít obsáhlé znalosti v oblasti počítačové grafiky nebo programování. Aplikace s názvem FaceAapp A Wombo jsou velmi oblíbené u uživatelů kvůli možnostem vytváření vtipných videí, které jsou následně velmi oblíbené na sociálních sítích.

FaceApp

Je aplikace vyvinuta ruskou společností Wireless Lab12. Ta se proslavila tím, že dokázala v jedné z funkcí vytvořit z fotografie jedince jeho zestárlejší podobu. Nabízí ale i další funkce a filtry, dokáže na tvář přidat úsměv, změnit barvu pleti, pohlaví jedince na fotografii a další. Tuto aplikaci si mohou stáhnout uživatelé systému Android i iOS.

Deepfake můžeme díky této aplikaci vytvářet, aniž bychom byli odborníky či za to někomu platili. Přesto většina takto vytvořených fotografií či videí slouží jen pro pobavení druhých, nemají za cíl ovlivňovat veřejné mínění společnosti. Většina z funkcí je bezplatná, jen některé z nich si musí uživatel zaplatit.

FaceApp nepřináší pouze výhody, přináší s sebou také různá rizika. Nejzásadnějším a nejnebezpečnějším z nich je nakládání s osobními údaji svých uživatelů. Jak uvádí ve svých obchodních podmínkách, vyhrazuje si širší práva při zpracovávání osobních informací a jejich užití pro obchodní účely. Jak je s jeho daty nakládáno, uživatel nemá šanci zjistit.

Zao

Jedná se o čínskou aplikaci, která dokáže vytvořit deepfake video na základě nahrání jedné fotografie. Obličej z nahrané fotografie dokáže nahradit obličej známého herce ve filmech.

Za aplikací Zao stojí čínský programátor a vývojář MoMo. Vývoj této aplikace probíhal pod taktovkou agentury, která vlastní na příklad i známou čínskou seznamku Tantan. Při porovnání reality s verzí vytvořenou aplikací Zao je však vidět rozdíly, je snadno rozpoznatelné, že se jedná o upravené video. Algoritmus umělé inteligence nemá v této aplikaci dostatek podkladů pro vytvoření kvalitnějšího videa. Tato aplikace je určená pro čínský trh.

Wombo

Je jedním z nejnovějších programů, který na mobilních zařízeních dokáže upravit videa. Stačí jen nahrát fotografii a aplikace z něj udělá video, ve kterém se jedinec na fotografii bude hýbat a zpívat světové hity. Za vším stojí umělá inteligence a špičková deepfake technologie. Výsledek slouží výhradně k pobavení druhých. Aplikace je k dostání zdarma na Androidu i iOSu, kvalita je však špatná a obsahuje logo aplikace. Pro lepší kvalitu bez loga je nutné si zaplatit rozšířenou verzi aplikace.

FaceSwap

FaceSwap je nová revoluční technologie umělé inteligence, která vám umožňuje změnit jakoukoli tvář na jakémkoli videu nebo obrázku, což vám dává neomezené možnosti pro obsah videa a obrázků. FaceSwap je cloudová aplikace, která využívá výkonnou technologii umělé inteligence, která vám umožňuje změnit libovolnou tvář na jakémkoli videu nebo obrázku, zaměnit je za různé herce, skutečné, animované. Výhodou je, že se dá snadno odstranit pozadí na fotografii i videu.

4.2 Desktopové aplikace

Aplikace na stolním počítači, které jsou popisované níže poskytují uživatelům mnohem více pokročilých funkcí a možností než mobilní aplikace. Díky těmto možnostem vytvářejí mnohem kvalitnější a realističtější deepfake tvorbu. Tyto aplikace také umožnují uživatelům větší kontrolu nad výsledným produktem, což znamená, že mohou upravovat a vylepšovat svá díla do mnohem podrobnějších detailů. Tyto aplikace však vyžadují mnohem vyšší technické znalosti než aplikace užívané na mobilních telefonech.

DeepFaceLab

Jedná se o grafický a designový nástroj, který vám umožní efektivně zaměnit obličeje na jakémkoli obrázku nebo videu. Tento open-source deepfake systém, vyvinutý společností sf-editor1, vede na trhu s více než 95 % samotných vytvořených deepfake videí. Tento program je snadno použitelný i pro uživatele, kteří nemají žádné komplexní znalosti či zkušenosti v oblasti hlubokého učení. Poskytuje poměrně flexibilní a volnou spojovací strukturu pro posílení uživatelského kanálu těmi nejjednoduššími metodami.

Kromě nahrazení tváře na obrázku nebo videu vám DeepFaceLab umožňuje také změnit hlavu, omladit obličej, či dokonce manipulovat se rty během projevů. Tato konkrétní funkce však vyžaduje dovednosti v softwaru pro úpravu videa, jako je Adobe After Effects nebo Davinci Resolve.

DeepFaceLab je efektivní platforma pro výrobu různých deepfake obsahů. Jeho poskytovaný kanál tak jednoduchý, aby vyhovoval všem úrovním technických znalostí, přesto však nabízí flexibilní množství přizpůsobení, které mohou uživatelé, kteří jsou dostatečně obeznámeni s programem, snadno upravit.

FakeApp

FakeApp je pokročilá aplikace pro úpravu videa, která uživatelům umožňuje měnit tváře lidí ve svých videích pomocí síly strojového učení a zpracování umělé inteligence. S využitím všech nejnovějších pokroků ve vývoji a deepfake algoritmů, FakeApp představuje komplexní balíček všeho v jednom, který komukoli umožňuje úspěšně nahradit tvář osoby ve videu tváří zcela jiné osoby. Tento typ aplikace byl původně schopen přidávat pouze počítačem generované nebo statické 2D obrázky kolem obličejů lidí (falešné brýle, zaječí uši, kníry a další jednoduché prvky), neuvěřitelný pokrok v této oblasti umožnil moderním vývojářům softwaru přizpůsobil obličeje z široké databáze na cílové obličeje ve videích.

Původně se používala pro zábavní účely a profesionální výměnu tváří pro televizní reklamy nebo filmy, v poslední době se ale různá softwarová řešení schopná výměny tváří využívá i pro nekalé činy.

I když tato aplikace může vytvářet realistické výsledky výměny obličejů, uživatelé k tomu budou muset investovat značné množství času, úsilí a poskytnout dostatek referenčních dat. Tato referenční data musí přijít ve formě široké škály fotografií obličeje, které jsou posléze

analyzovány, spárovány a přeměněny do požadované podoby. Čím více dat bude shromážděno, tím lepší budou konečné výsledky.

FakeApp nejprve analyzuje video, které uživatelé poskytnou, a pokusí se izolovat nejen všechny pohyby hlavy, ale také polohy a pohyby očí, úst a dalších struktur obličeje. Po shromáždění těchto dat se aplikace pokusí přiřadit vaše referenční fotografie k videu a pečlivě se bude snažit zachovat prvky, jako je pohyb očí a úst. Konečný výsledek bude záviset nejen na rozsahu a kvalitě referenčních fotografií, ale také na čase, který aplikace využije ke správnému spojení všeho dohromady.

FakeApp není uživatelsky přívětivá aplikace a bude vyžadovat, aby uživatelé zpracovávali nepraktické a trochu matoucí rozhraní a pracovní postupy. Kromě těchto problémů je aplikace také náročná na zdroje a někdy bude vyžadovat vícehodinové zpracování, při kterém bude váš CPU, GPU a RAM silně zatěžován. Navíc samotná aplikace zabírá více než 1,5 GB a vyžaduje mnohem více volného místa na vašem místním úložišti a paměti RAM, což téměř nebo zcela znemožňuje spuštění na starších/slabších stolních počítačích a mnoha noteboocích.

5 ŘÍZENÉ ROZHOVORY

Řízené rozhovory byly realizovány formou krátkého dotazování 5 respondentů, které jsem se snažil vybírat tak, aby jejich věkové kategorie a pohlaví byly různé. Věřím, že tato různorodost může znamenat rozdílné povědomí o fenoménu deepfake, názorech a znalostech o této problematice. Mezi 5 respondenty tak můžeme najít 2 ženy ve věku 35 a 22 let, a 3 muže ve věku 39, 27 a 53 let. Míním, že s přibývajícím věkem bude povědomí o prostředcích moderní technologie nižší.
Předběžně byly proto stanoveny tyto dvě hypotézy:

Hypotéza 1: Mladší generace budou mít větší znalosti o problematice deepfake než generace starší.
Hypotéza 2: Respondenti budou v problematice spatřovat více negativ než pozitiv.

Výběr této metody dotazování se pro účely této bakalářské práce jevil jako nejvhodnější. Forma rozhovoru nepřináší takovou anonymitu jako šetření za pomoci dotazníku, který je navíc distribuován mezi větší počet respondentů – čímž je zaručen větší výzkumný vzorek. Největší pozitivum v rozhovorovém šetření však spatřuji ve faktu, že pokud si respondent není jist svou odpovědí, dotazující mu může pomoci – jako v případě této bakalářské práce – ne každý z respondentů věděl, co to deepfake je a když jim bylo v průběhu rozhovoru vysvětleno, o co se jedná, zdál se jim termín povědomý či se jim rovnou vybavily konkrétní příklady. To byl hlavní důvod, proč byla nakonec vybrána forma přímá, prezenční forma rozhovoru. V případě dotazníkového šetření bez účasti dotazujícího by šetření skončilo u třetí otázky. V následující kapitole se budeme zabývat přímo řízených rozhovorů s jednotlivými respondenty. Všichni respondenti byli před samotným dotazováním ujištěni o anonymitě těchto řízených rozhovorů, souhlasili, že jejich odpovědi budou sloužit výhradně pro účely této bakalářské práce. Všichni respondenti nicméně souhlasili, že bude uveřejněno jejich křestní jméno, věk a pracovní pozice.
Rozhovory byly se souhlasem respondentů zaznamenány nahrávací zařízení a později přepsány mnou do textového souboru.

5.1 Výsledky řízených rozhovorů

MAREK, 39 let, administrativní pracovník

Jaký je Váš vztah k moderním technologiím?

Docela se o ně zajímám, aktivně používám pouze Messenger, YouTube, ale znám i jiné, s některými přímé zkušenosti nemám, ale když se řekne TikTok, vím, o čem je řeč. Zajímají mě spíše nové mobilní telefony, jejich vývoj, zkoumám, jakým směrem se vyvíjí jejich parametry a zajímám se také o jejich budoucnost.

Používáte aktivně sociální sítě? Jak často?

Jak jsem řekl dříve, jen Messenger.

Víte, co znamená pojem „deepfake“?

Vím, že je to technologie, která mění ve videích obličeje, zaměňuje hlasy, mimiku. Některé jsou docela vypracované.

Setkali jste se někdy s deepfake?

Možná ano, ale neuvědomuji si konkrétní podobu videa.

Jaké pozitiva či negativa podle Vás deepfake přináší?

Myslím, že spíše je zajímavé sledovat ten vývoj techniky a jistě se i v budoucnu najde nějaké dobré využití. Zatím jsem ale spíše slyšel jen o jejich využití pro zábavu jiných.

VERONIKA, 35 LET, HR konzultantka

Jaký je Váš vztah k moderním technologiím?

Tak v tomto docela pokulhávám, tak nějak sleduji, co se děje, ale více se o to nezajímám. Nevím, jak se propojují různé aplikace s televizí, například, nevím, jaké parametry má nová televize mít, ať má kvalitní signál, obraz, zvuk. To je pro mě veliká neznámá, ale nijak mne to ani netrápí.

Používáte aktivně sociální sítě? Jak často?

Používám je v podstatě pořád, pořád sleduji Instagram, Facebook, Twitter, Pinterest, WhatsApp. Snažím se ale taky držet 1 den týdně offline, zpravidla to bývají neděle, kdy nechodím do práce, trávím více času v přírodě, s přáteli či rodinou.

Víte, co znamená pojem „deepfake“?

Ano, již jsem o něm slyšela.

Setkali jste se někdy s deepfake?

Viděla jsem jednou video s nějakým politikem a stejně jsem i slyšela o případech, kdy byly zneužity tváře hereček u porno videí.

Jaké pozitiva či negativa podle Vás deepfake přináší?

U pozitiv si nejsem úplně jistá, bojím se spíš zneužití, protože tyhle technologie půjdou asi dopředu, budou se vyvíjet, lidi se budou více učit s nimi pracovat, budou propracovanější a bude těžké oddělit realitu od „fejku".

LUCIE, 22 let, student

Jaký je Váš vztah k moderním technologiím?

Jsem spíš v tomhle „divák", ale přítel je strašný nadšenec. Také můj otec se o moderní technologie zajímá, doma máme všichni věci od Apple a řekla bych, že se jim daří držet krok s dobou, vyznají se ve všech různých zkratkách, dokážou si všechno mezi sebou spojit, mají na tohle hlavu. Občas jim závidím.

Používáte aktivně sociální sítě? Jak často?

Jsem online pořád, jsem aktivní na Instagramu a TikToku, točím i krátká videa, reels.

Víte, co znamená pojem „deepfake"?

Ten pojem znám spíš povrchně.

Setkali jste se někdy s deepfake?

Už si vybavuju, deepfake znám, někdy jsou fakt dobře udělané a je těžko je rozeznat od reality.

Jaké pozitiva či negativa podle Vás deepfake přináší?

Pro mě znamenají postup v technologiích dopředu, což je super, je fajn vidět, jaký dělá věda a IT pokroky, jaké dneska existují vymoženosti. Pro mě jsou určitě přínosem.

JAKUB, 27 let, programátor

Jaký je Váš vztah k moderním technologiím?

Velice úzký, potřebuji je k práci, ale trávím s nimi i dost volného času.

Používáte aktivně sociální sítě? Jak často?

Pořád. V práci, doma. Vím, že to není úplně OK, ale doba si je žádá.

Víte, co znamená pojem „deepfake"?

No jasně. Fotky, videa, zvuky, které někdo upravil, ale působí jako opravdové. Sám jsem si pár zkusil vytvořit, nikde jsem je ale nesdílel.

Setkali jste se někdy s deepfake?

Ano, pasivně i aktivně.

Jaké pozitiva či negativa podle Vás deepfake přináší?

Jako pozitivum možná zábava – vytvoříte nějakou úpravu fotek, zasmějete se s kamarády. Horší jsou případy, že to fakt někomu uškodí, ublíží – a to jak psychicky, tak finančně. Některé celebrity s nimi docela bojují a není se asi čemu divit, některé kompromitující fotky nebo videa jim můžou zničit pověst, kariéru, život.

ALEŠ, 53 let, lékař

Jaký je Váš vztah k moderním technologiím?

Snažím se s nimi držet krok, občas jsou nutné i k výkonu mé práce. Doma ale tápu a většinou mi musí poradit některý ze sousedů.

Používáte aktivně sociální sítě? Jak často?

Nepoužívám, jsem jiná generace, nic mi to neříká, ale moje tři děti jsou na nich přímo přilepené.

Víte, co znamená pojem „deepfake“?

Nikdy jsem o něm neslyšel.

Setkali jste se někdy s deepfake?

Asi ano, ale konkrétní případ si nevybavím.

Jaké pozitiva či negativa podle Vás deepfake přináší?

Pozitiva nevidím, vždyť to je postavené na hlavu, nechápu, k čemu to má sloužit. Jedině to může lidem škodit, určitě by se to mělo nějak regulovat či zakázat. Už tak je tahle doba v některých věcech dost zvrácená.

5.2 Shrnutí a zhodnocení výsledků řízených rozhovorů a následná doporučení

Řízené rozhovory pro účely této bakalářské práce probíhali počátkem roku 2023. Pro účely této práce jsem oslovil 5 pro mě neznámých osob, které jsem požádal, zda by mohli na diktafon odpovědět na 5 otázek týkajících se moderních technologií a deepfake problematiky. Všichni respondenti souhlasili s nahráváním rozhovoru a jeho následným využitím pro potřeby této závěrečné práce. Zároveň byli před samotným dotazováním ujištěni, že rozhovor je zcela anonymní, přesto, pokud chtějí, mohou poskytnout základní údaje o sobě, křestní jméno, věk, pracovní zařazení či město, ve kterém bydlí. S posledním

údajem nicméně souhlasili pouze dva respondenti, proto jsem se rozhodl nezmiňovat jej ani u jedné z dotazovaných osob.

Respondentům bylo položeno těchto 5 otázek:

Jaký je Váš vztah k moderním technologiím?

První otázka se zabývala vztahem respondentů k moderní technologií, jak jsou technologicky zdatní, zda mají přehled o novinkách v technologickém světě a zda si poradí, když je třeba udělat něco, co je spojené s technikou.

V odpovědích respondentů je tak zřetelně vidět, že vztah respondentů k technologickým pokrokům moderní doby je tím užší, čím nižší je jejich věk. Ačkoli nelze generalizovat, jelikož je výzkumný vzorek příliš úzký, při pohledu na dnešní společnost je možné názory v těchto rozhovorech a výsledky průzkumu brát jako směrodatný. Mladší generace má k moderním technologiím provázanější vztah, stejně tak může vykazovat vyšší znaky závislosti na moderních technologiích. Moderní technologie jsou skvělý prostředek pro zpestření běžného dne, usnadňuje nám životy – komunikace je snazší, jelikož můžeme komunikovat neustále. Potřebujeme-li cokoli zjistit, zařídit, můžeme jednoduše vzít do ruky telefon a všechno vyřídit. Lidé jsou si díky moderním technologiím blíže a zároveň dál – vytrácí se schopnost komunikovat naživo, jednodušší je sdělit novinky přes komunikační kanály. Výjimkou pak nejsou ani zásadní rozhodnutí či události – svým rodinám oznamujeme smskami zásnuby, úmrtí druhých i informujeme o rozchodu či rozvodu.

Naši respondenti potřebují moderní technologie pro výkon svého zaměstnání, uznávají, že jim usnadňují život. Někteří z nich však také uznávají, že krok s moderními technologiemi nedrží – či jen do té míry, jakou sami potřebují.

Používáte aktivně sociální sítě? Jak často?

Druhá otázka se zabývala vztahem respondentů k sociálním sítím, zda je používají, zda jsou aktivní či pasivní uživatelé, či jde tento svět mimo ně. V současné době důležitost sociálních sítí, komunikačních kanálů prudce stoupá a dá se předpokládat, že v tomto trendu budou sítě pokračovat také v příštích letech.

Také v této otázce se odráží hledisko věku, zatímco generace mladší je na sítích velice aktivní, používá je denně, aktivně sdílí obsah, generace třicátníků od sociálních sítí pomalu upouští (či naopak si k nim nenašla cestu, sociální sítě nejsou dlouhotrvajícím trendem), generace dnešních padesátníků k životu sociální sítě zpravidla nepotřebuje.

Sociální sítě v současnosti odráží úspěch mladého člověka či dítěte v kolektivu. Aby zapadl do party, musí držet krok s novinkami, být stále online. Kdo je stále k dispozici, má

zvýšenou podporu svých vrstevníků. Je nicméně velmi snadné dostat se skrze své virtuální aktivity do potíží. Téměř desetina mladých lidí se setkala s negativními aspekty sociálních sítí jako jsou kyberšikana, podvody, sexting aj.

Sociální sítě mohou mít také negativní vliv na duševní zdraví svých uživatelů – nízké sebevědomí, úzkosti, deprese, nedostatečné sebeurčení apod.

Víte, co znamená pojem „deepfake“?

Třetí otázka se zabývala samotným deepfake. Dalo se očekávat, že tento pojem nebude znít povědomě všem – tento důvod byl považován za hlavní při výběru formy řízených rozhovorů. Tato domněnka se naplnila, většina z respondentů si nebyla jistá, o co se jedná, po mém vysvětlení pojmu deepfake již věděli, o co se jedná a někteří si také vzpomněli na konkrétní příklady. Deepfake mají spoustu kladných i záporných stránek, někdy je velice těžké rozlišit deepfake video od originálu. Veřejnost, která nemá zkušenosti s touto technologií, či nemá ani ponětí, že „něco takového jde udělat“, může být zmatená, zklamaná či rozčarovaná tím, čeho se od své oblíbené celebrity dočkala.

Rozhovory s respondenty dokazují, že široká veřejnost zatím nemá větší znalosti o této problematice – nicméně pokud jim bylo následně vysvětleno, o co se jedná, uvědomili si, že deepfake znají či s nimi mají nějakou zkušenost.

Nabízí se proto poskytnout veřejnosti širší informace o této problematice, informovat je prostřednictvím kampaní, stejně jako se informuje o různých podvodných taktikách.

Po osvětlení, co tento pojem obnáší, si respondenti vzpomněli, o co se jedná a že se s tímto moderním fenoménem již setkali. Vzpomněli si také na konkrétní příklady, se kterými se již v minulosti setkali.

Setkali jste se někdy s deepfake?

V předposlední otázce jsem se respondentů přímo ptal, zda se s deepfake již někdy konkrétně setkali a většina z nich souhlasila, že se s upravenými fotkami či videi již někdy setkala. Ačkoli v předchozí otázce váhali, co to deepfake vlastně je, po vysvětlení pojmu zjistili, že se s takovým obsahem již setkali. Některé deepfake videa jsou provedená tak dobře, že je velice těžké rozlišit, zda se jedná o pravé či zfalšované video. Jedna z respondentek si dokonce vzpomněla na konkrétní případy, kdy byly obličeje ženských hereček zneužity v pornografickém průmyslu. Jak však již bylo řečeno, někdy je velmi obtížné rozeznat, že se nejedná o pravé video, které mohlo na veřejnost uniknout. Tyto případy mohou mít nepříznivý vliv na kariéru i celkový život postižené známé osobnosti.

Jaké pozitiva či negativa podle Vás deepfake přináší?

Poslední otázka byla pro respondenty takzvanou otázkou na tělo. Dá se však shrnout, že negativa deepfake v názorech společnosti převládají. Jediným pozitivem, jaké se dalo u deepfake najít, byla jeho zábavní funkce, respondenti se však shodli na tom, že deepfake mohou představovat do budoucna hrozbu pro společnost, pro kariéry i životy jednotlivců. Jeden z respondentů vyjádřil svůj nesouhlas s touto technologií také názorem, že by se deepfake měl redukovat, cenzurovat či přímo zakázat.

Deepfake však nemá pouze jedno pozitivum, což dokládá má přímá zkušenost z výstavy Alfonse Muchy, který za pomocí moderní technologie návštěvníky galerie vítal a promlouval k nim. Jednalo se o velmi zdařilé video, přesto bylo vidět, že je upravené, synchronizace rtů s mluveným slovem nebyla přesná. Věřím, že pro tyto účely je vhodné deepfake technologie používat a videa tvořit. Pokud má mít video edukační formát, deepfake mají smysl.

Pokud však mají škodit, bude lepší najít způsob, jak jejich tvorbu regulovat či cenzurovat. To je možné vidět na teoretickém příkladu mladé herečky, která je v začátcích své kariéry a se slávou se těžce vyrovnává. Pokud se tato herečka objeví v deepfake porno videu, může tento jev zanechat hluboké šrámy na její psychice. Žena může být touto nálepkou poznamenaná na celý život, mít psychické problémy, propadnout závislostem či se rozhodnout svůj život ukončit. Nabízí se proto otázka účelu tvorby těchto videí, pokud jsou tvořená za účelem někomu ublížit, jednoznačně by měla být zakázána.

S dalším rozvojem technologií se problematika deepfake dostane více na veřejnost, to je zaručené. Otázka regulace deepfake se bude také řešit a evidentně nebude jednoduché řešit hranici, kde jsou deepfake videa ještě v pořádku, legální a kde již v pořádku nejsou. O této problematice by se tak mělo informovat, stejně jako se informuje o nežádoucích účincích návykových látek, o phisingu, me too, bossingu či podvodných e-mailech z Nigérie. Všechny tyto aspekty deepfake se tak dostanou do souvislostí, do společenských či dokonce právních norem. Jistá regulace deepfake obsahu, který se dostane na veřejnost, je tak prakticky nutná.

5.3 Vyhodnocení hypotéz

Závěrem lze tedy říct, že výsledky dotazování potvrdily obě hypotézy stanovené před samotných průzkumem.

- Hypotéza 1: Mladší generace budou mít větší znalosti o problematice deepfake než generace starší.

Hypotéza 1 byla potvrzena.

Výsledky řízených rozhovorů skutečně ukazují na to, že mladší generace má větší znalosti o problematice deepfake než starší generace. Mladší generace je od ranného věku více vystavena digitálním technologiím a médiím, než byla generace starších občanů. Mladší lidé mají často větší povědomí o rizicích, která hrozí v digitálním světě, včetně hrozeb spojených s deepfake technologií. Díky takovému povědomí jsou mladší lidé schopni se lépe vyhnout nebezpečím spojených s technologií deepfake.

- Hypotéza 2: Respondenti budou v problematice spatřovat více negativ než pozitiv.

Výsledky řízených rozhovorů prokázaly, že většina respondentů skutečně spatřuje v problematice deepfake technologií více negativ než pozitiv. Tázaní respondenti si uvědomují rizika spojená s tázanými technologiemi, jako jsou například možnosti zneužití pro vydíraní nebo manipulaci s informacemi. Hrstka respondentů však vidí také určité pozitivní aspekty deepfake technologie, například v oblasti kultury a zábavy. Celkově lze však konstatovat, že většina respondentů vyjadřuje negativní pohled na deepfake technologií a zdůrazňuje důležitost prevence.

Je však nutné brát v potaz, že dotazovaný vzorek byl velice úzký, aby se daly vyvozovat závěry, vždy se mohou najít výjimky – můžeme najít technicky nezdatnou mladou ženu a naopak technicky znalého jedince v seniorském věku. Přesto, pokud vezmeme v potaz pohled na moderní společnost, vidíme, že znalost technologií, sociálních sítí a komunikačních kanálů či práce s mobily, počítači je u mladší generace na vyšší úrovni než u generace jejich rodičů či prarodičů.

ZÁVĚR

Bakalářská práce se zabývala problematikou deepfake. V teoretické části byl popsán známý význam pojmu deepfake. Dále byla popsána dostupná legislativa, zejména tedy různá nařízení a doporučení Evropské unie. Bylo zjištěno, že v České republice není problematika deepfake řádně zákoně ošetřena. Dále byla popsána technická stránka problematiky deepfake, kdy byly vysvětleny formy deepfake a to zvuková, obrazová a video forma. Největší pozornost byla věnována kapitole historie a aplikace deepfake technologie. Bylo popsáno, v jakém období taková technologie vznikla a jakou událostí se dostala do podvědomí široké veřejnosti. Dále jsou stručně sepsány některé příklady událostí, ve kterých byla technologie deepfake použita, či zneužita za účelem zviditelnění se, získání pozornosti nebo také dosažení určitého cíle.

V praktické části je stručně popsán styk široké veřejnosti s pojmem deepfake. Hlavní zaměření praktické části je tvorba modelových situací, při kterých došlo ke zneužití deepfake technologie a jsou popsány důsledky, které mohou nastat pro společnost, nebo také potíže v rámci ochrany obyvatelstva. Jsou zde vytvořeny i návrhy opatření, jak se proti takovým hrozbám bránit. Na základě těchto vytvořených modelových situací, lze předpokládat, že v případě kvalitního využití deepfake technologie za účelem dosažení vytyčeného cíle, tak hrozí velké nebezpečí pro většinu oblastí v rámci ochrany obyvatelstva, ale i dalším vrstvám společnosti. Dále jsou popsány běžně dostupné aplikace na mobilních telefonech a stolních počítačích, které se využívají ke tvorbě deepfake.

V další části byly provedeny řízené rozhovory s pěti respondenty, kdy zde byla snaha o vyvrácení nebo potvrzení vytvořených hypotéz. První hypotéza předpokládala, že mladší generace bude mít větší znalosti o problematice deepfake než generace starší. Tato hypotéza byla potvrzena. Druhá hypotéza předpokládala, že respondenti budou v problematice spatřovat více negativ než pozitiv a byla taktéž potvrzena. Tímto byli hlavní a dílčí cíle bakalářské práce zpracovány a splněny.

SEZNAM POUŽITÉ LITERATURY

AYYUB, Rana. I Was The Victim Of A Deepfake Porn Plot Intended To Silence Me. In: Huffingtonpost [online]. 2018 [cit. 2023-05-01]. Dostupné z: https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316

CBS Interactive. Doctored Nancy Pelosi video highlights threat of "deepfake" tech. In: Cbsnews [online]. 2019 [cit. 2023-05-01]. Dostupné z: https://www.cbsnews.com/news/doctored-nancy-pelosi-video-highlights-threat-of-deepfake-tech-2019-05-25/

COPELAND, B. (2022, March 18). artificial intelligence. Encyclopedia Britannica. https://www.britannica.com/technology/artificial-intelligence

DVOŘÁKOVÁ, M. Revenge porn a deepfakes: ochrana soukromí v éře moderních technologií [online]. Revue pro právo a technologie [online]. 2020, č. 22, s. 51-89 [cit. 11. 7. 2022]. Dostupné z: https://journals.muni.cz/revue/article/view/13416/pdf

FAGAN, K. A viral video that appeared to show Obama calling Trump a 'dips---' shows a disturbing new trend called 'deepfakes' [online]. bussinesinsider.com [online]. 17. 4. 2018 [cit. 11. 7. 2022]. Dostupné z: https://www.businessinsider.com/obama-deepfake-video-insulting-trump-2018-4

GREGOR, Miloš a Petra VEJVODOVÁ. Nejlepší kniha o fake news, dezinformacích a manipulacích!!!. Brno: CPress, 2018. ISBN 978-80-264-1805-4

HARRIS, Douglas. Deepfakes: False Pornography Is Here and the Law Cannot Protect You. Duke Law & Technology review [online]. 2019, roč. 17, 1 [cit.11.7.2022], s.99. Dostupné z: https://scholarship.law.duke.edu/dltr/vol17/iss1/4

CHESNEY, Robert, a Danielle Keats Citron. 2018. „Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security.“ California Law Review 107: 1753-1813. https://doi.org/10.2139/ssrn.3213954.

KIETZMANN, Jan, Linda W. Lee, Ian P. Mccarthy, a Tim C. Kietzmann. 2020. „Deepfakes: Trick or Treat?“ Business Horizons 63 (2): 135–46. https://doi.org/10.1016/j.bushor.2019.11.006.

MARR, Bernard. Poznáte, že to není Tom Cruise? Deepfakes se přibližují realitě a stávají se skutečnou hrozbou. In: Forbes [online]. 2022 [cit. 2023-05-01]. Dostupné z: https://forbes.cz/poznate-ze-to-neni-tom-cruise-deepfakes-se-priblizuji-realite-a-stavaji-se-skutecnou-hrozbou/

NÁVRH NAŘÍZENÍ EVROPSKÉHO PARLAMENTU a Rady Evropské unie [online]. eur-lex.europa.eu [online], 2021 [cit. 10. 7. 2022]. Dostupné z: https://eur-lex.europa.eu/legal-content/CS/TXT/?uri=CELEX:52021PC0206

OLCKERS, AYRAN. 2020. „Artificial Intelligence: AI Terms Simply Explained.“ Towards Data Science. Medium. 12.7.2022. https://towardsdatascience.com/artificial-intelligence-ai-terms-simply-explained-745c4734dc6c

PARLAMENT definuje první pravidla EU pro umělou inteligenci | Zpravodajství | Evropský parlament. [online]. Copyright ©AdobeStock [cit. 20.07.2022]. Dostupné z:https://www.europarl.europa.eu/news/cs/pressroom/20201016IPR89544/parlament-definuje-prvni-pravidla-eu-pro-umelou-inteligenci

REILLY, K., KOVACH, S. Face-swapping videos could lead to more 'fake news' [online]. bussinesinsider.com [online]. 13. 4. 2018 [cit. 11. 7. 2022]. Dostupné z: https://www.businessinsider.com/fakeapp-lets-people-make-fake-videos-deepfakes-2018-4.

SHICK, Nina. Deepfakes: The Coming Infocalypse. New York: Twelve,2020. ISBN 1538754304.

SILBEY, Jessica a Woodrow Hartzog. 2018. „The Upside of Deep Fakes.“ Maryland Law Review 78 (4): 960-66. https://digitalcommons.law.umaryland.edu/cgi/viewcontent.cgi?article=3837&context=mlr.

SONG, David. 2019. „A Short History of Deepfakes.“ Medium. 25. 9. 2019 [cit. 11. 7. 2022]. Dostupné z: https://medium.com/@songda/a-short-history-of-deepfakes-604ac7be6016.

USNESENÍ PŘECEDNICTVA České národní rady č. 2/1993 Sb., o vyhlášení Listiny základních práv a svobod jako součásti ústavního pořádku České republiky, ve znění ústavního zákona 162/1998 Sb.

WEGHE, Tom Van de. Six lessons from my deepfakes research at Stanford: How should journalists address the growing problem of synthetic media. Medium [online]. 29.5.2019 [cit. 12.7.2022]. Dostupné z: https://medium.com/jsk-class-of-2019/six-lessons-from-my-deepfake-research-at-stanford-1666594a8e50

YOUNG, Nobert. DeepFake Technology: Complete Guide to DeepFakes, Politics and Social Media. Spojené státy americké: Nezávisle vydáno, 2019. ISBN 107849469X..

ZÁRUBA, T. Deepfakes [online]. vskp.vse.cz/ [online]. Praha, 2020 [cit. 12. 7. 2022]. Dostupné z: https://vskp.vse.cz/80219_deepfake

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

2D dvoudimenzionální

3D trojdimenzionální

CPU procesor (anglicky Central Processing Unit)

GAN Generátorová sít (anglicky Generative Adversarial Network)

GB gigabajt

GDPR ochrana osobních údajů (anglicky General Data Protection Regulation)

GPU grafický procesor (anglicky Graphics Processing Unit)

FBI Federální úřad pro vyšetřování (anglicky Federal Bureau of Investigation)

iOS operační systém iPhone (anglicky iPhone Operating System)

RAM paměť počítače (anglicky Random Access Memory)

SEZNAM OBRÁZKŮ