

# **NEGATIVNÍ DOPADY GENERATIVNÍ UMĚLÉ INTELIGENCE BUDOU VIDĚT STÁLE VÍCE**

**Kamil KOPECKÝ**

Generativní umělá inteligence má řadu pozitiv (kterým se věnujeme např. na našich specializovaných stránkách), na druhou stranu nesmíme ignorovat také rizika a problémy, které s sebou tato technologie nese. Ve veřejném prostoru se např. v posledních měsících stále více objevují fotografie a videa, která byla vytvořena s pomocí nástrojů generativní umělé inteligence. Mnohá z nich pak zachycují známé osobnosti - bez jejich souhlasu a s cílem dehonestovat je či zneužít např. k podvodné činnosti. Týkají se však i běžných uživatelů internetu, např. dospívajících dívek, které se staly terčem tzv. svlékacích aplikací.

## **Generovaná pornografie a případ Taylor Swift**

Jedním z vážných problémů spojených s generativní umělou inteligencí je **vytváření sexuálně explicitních** (často pornografických) **fotografií/videí známých osobností** - hereček, zpěvaček, ale také např. političek. Nemusí však jít vůbec o celebrity - terčem se může stát v podstatě kdokoli. Vytvořit tyto materiály je extrémně snadné - aplikací, které toto umožňují, je mnoho a každý den vznikají další.

Nebezpečné jsou pak především tzv. [deep fake technologie](#). Tyto technologie, využívající pokročilé algoritmy strojového učení, umožňují vytvářet extrémně realistická videa a audio nahrávky, ve

---

kterých lze manipulovat s tvářemi, těly a hlasy lidí. Samozřejmostí je pak generování fotografií konkrétních osob v různých situacích (včetně sexuálně explicitních) během několika vteřin.

S pornografií generovanou umělou inteligencí má čerstvou zkušenost např. zpěvačka [Taylor Swift](#). Ve středu 24. ledna 2024 se na síti X (bývalý Twitter) rozšířily její sexuálně explicitní deepfakes fotografie vytvořené pomocí umělé inteligence, které za 19 hodin nasbíraly přes 27 milionů zhlédnutí a více než 260 000 lajků, než byl účet, který snímky zveřejnil, pozastaven. Deepfakes zobrazující Swift nahou a v sexuálních scénách se síti X neustále šíří - nyní již také mimo tuto síť.



Sexuálně explicitní obrázek Taylor Swift generovaný AI (ořezáno)

Tento problém se však netýká pouze známých osobností, problémy s generovanou pornografií zneužívající identitu oznamují dívky z celého světa. Ve Spojených státech např. [nahlásily desítky dívek](#) středoškolského věku, že se staly oběťmi deepfakes. Již dříve došlo k obdobným případům také ve Velké Británii - upravené video, na kterém byly zachyceny pornoherečky, jejichž tváře byly nahrazeny tvářemi dospívajících dívek, přispělo k [sebevraždě 14leté dívky](#). Podobné případy byly hlášeny také ze Španělska či České republiky, o kterých jsme informovali na našem portálu již [v říjnu minulého roku](#). Zcela se tak naplnilo naše varování, že generovaná pornografie zachycující reálné osoby bude v budoucnu vážný problém. Budoucnost je nyní!

## **Generování deep fake videí politiků a narušování fungování demokratické společnosti - včetně voleb**

Jedním z problémů, který deep fake videa přinášejí, je **narušení důvěry ve veřejné instituce a politické procesy a ovlivňování demokratických procesů** (např. voleb). Zfalšovaná videa/audionahrávky, zobrazující politiky v kompromitujících nebo kontroverzních situacích, mohou být snadno šířena prostřednictvím sociálních sítí a dalších médií. Tato videa/audionahrávky jsou často

---

tak přesvědčivá, že je pro průměrného diváka těžké rozpoznat jejich nepravost. V důsledku toho může docházet k šíření dezinformací, které mohou ovlivnit veřejné mínění a volby.

S ovlivněním voleb pomocí těchto technologií má čerstvou zkušenost např. Slovensko - těsně před volbami se začala šířit [podvržená nahrávka telefonického rozhovoru](#) o manipulaci volebních výsledků mezi předsedou Progresivního Slovenska kandidujícího do Národní rady SR a novinářkou Monikou Tódovou. Rozhovor ale ve skutečnosti neproběhl, vše bylo vytvořenou umělou inteligencí. Za podvrh označila nahrávku jak slovenská policie, tak i [fact-checkingová agentura AFP](#). Podvržená nahrávka se pak internetem šířila skutečně virálně a je zřejmé, že část voličů dokázala ovlivnit.

Zkušenost s podvodnými videi máme i v České republice, v průběhu prezidentských voleb v roce 2023 se např. šířilo internetem [upravené video prezidentského kandidáta Petra Pavla](#).

Aktuálně pak koluje internetem např. podvodné deep fake video zaměřené proti [ministru vnitra Vítu Rakušanovi](#).



Deep fake video zaměřené proti ministru Rakušanovi

## Generativní umělá inteligence a podvody

Nástroje generativní umělé inteligence lze zneužít také k **podvodné činnosti** - a to např. na generování phishingových či vyděračských e-mailů, s jejich pomocí můžeme generovat také [dezinformační obsah](#). Na podvody v textové podobě jsme si však již zvykli a nenecháme se jimi tak snadno ovlivnit - nezvykli jsme si však na deep fake videa, která jsou u online podvodů stále častější.

Typickými příklady podvodů, které zneužívají deep fake videa známých osobností, jsou nejrůznější druhy "[podvodných investičních reklam](#)", které slibují za minimální investici ohromný zisk. Podvod je pak doplněn deep fake videi známých osobností (expremiér Andrej Babiš, Žantovský, Tykač apod.), které potvrzují,

---

jak moc je investice skvělá a zaručená. Vše je navíc doplněno o loga skupiny ČEZ, Úřadu vlády ČR či stanice CNN Prima News. Díky těmto videím skutečně může část uživatelů internetu, kteří nedisponují dostatečnou úrovní mediální gramotnosti, sdělení uvěřit a do podvodných produktů investovat. S vysokou pravděpodobností ale budou podvedeni a o své finance přijdou.



Screenshot podvodného deepfake videa zneužívajícího  
Andreje Babiše a značku CNN Prima News

## Zvykejme si, bude hůř

Vzhledem k dostupnosti nástrojů generativní umělé inteligence, které umožňují vytvářet výše uvedený problematický obsah, je třeba počítat s tím, že **podobných případů bude přibývat**. Je třeba se tedy připravit na to, že lidskou podobu lze snadno podvrhnout a že je snadné napodobit jak vzhled člověka (jeho mimiku, gestiku), tak i také lidský hlas. AI nástroje mají potenciál radikálně transformovat způsob, jakým vnímáme digitální obsah, což s sebou nese řadu etických a právních dilemat. Půjde tento typ digitálního obsahu detekovat? Půjde jej regulovat? Co způsobí lidem? Čemu budeme v budoucnu důvěřovat? Jak budeme chránit identitu člověka? To vše jsou otázky, na které musíme velmi rychle najít odpovědi...

Pro E-Bezpečí  
Kamil Kopecký  
Univerzita Palackého v Olomouci

### Zdroje:

<https://www.nbcnews.com/tech/misinformation/taylor-swift-nude-deepfake-goes-viral-x-platform-rules-rcna135669>

---

---

---

---

---

---

---