

Facebook hodlá mazat videa s „deepfake“, jiná manipulace však nevadí

- CZ24 News | 13. ledna 2020

Společnost Facebook bude ze své stejnojmenné sociální sítě odstraňovat sofistikovaná podvodná videa, jejichž cílem je zmanipulovat diváky. Zákaz se podle sdělení firmy týká zejména obsahu vytvářeného prostřednictvím metod strojového učení, který se označuje termínem „deepfake“. Povoleny naopak podle všeho zůstávají jednodušší formy manipulace, které v poslední době opakovaně vyvolávají rozruch na americké politické scéně.

Deepfakes podle Facebooku představují pro digitální platformy a společnost obecně vážný problém. Na internetu jsou stále relativně vzácností, jejich užívání ale sílí, napsala na firemním webu viceprezidentka Monika Bickertová. Nedávno se objektem falešného videa tohoto typu stal i šéf Facebooku Mark Zuckerberg, když jeden britský umělec upravil úryvek z jeho televizního interview. Zuckerberg se pak ve věrohodně vyhlížejícím klipu mimo jiné chlubil tím, že má absolutní moc nad „ukradenými daty miliard lidí“.

Z instagramu, jehož vlastníkem je už několik let Facebook, příspěvek dosud stažen nebyl. Na největší sociální síti světa už by ale nyní zřejmě byl zakázán. „Budeme odstraňovat zavádějící upravený obsah,... půjde-li o produkt umělé inteligence či strojového učení vytvořený spojením, nahrazením či překrytím obsahu ve videu, čímž je navozen dojem, že je (to video) autentické,“ vysvětluje Bickertová.

Zákaz se vztahuje i na jinak upravená videa, která by mohla uživatele přesvědčit, že osoba na obrazovce „pronesla slova, která ve skutečnosti neřekla“. Editace obsahu přitom musí být za hranicí pouhého vylepšení kvality a musela by být provedena tak, že pro „průměrného člověka“ by nebyla zjevná.

„Opilá“ Pelosiová i „rasista“ Biden

Mnozí analytici nicméně mají za to, že alespoň v nejbližší době mohou větší hrozbu než „deepfakes“ představovat mnohem jednodušší formy manipulace. Média v souvislosti s novým opatřením Facebooku připomínají zejména loňskou kauzu s předsedkyní americké Sněmovny reprezentantů Nancy Pelosiovou. Na facebooku a dalších platformách se loni šířily zpomalené záběry z jejího projevu, které byly prezentovány jako důkaz, že Pelosiová je opilá nebo trpí nějakými zdravotními potížemi.

„Upravené video předsedkyně Pelosiové nesplňuje měřítko tohoto opatření a nebylo by smazáno. Blokována budou jen videa vytvořená umělou inteligencí, která ukazují lidi, jak říkají fiktivní výroky,“ sdělil Facebook agentuře Reuters.

Bickertová v blogovém příspěvku doplňuje, že ohlášený zákaz se netýká ani parodií a satiry či videí, v nichž chybějí části vyjádření mluvčího, nebo v nichž je změněno pořadí výroků. Sítem by tak prošel i hojně sdílený klip z minulého týdne, v němž byla vytržena z kontextu slova bývalého amerického viceprezidenta Joea Bidena ve snaze zobrazit ho jako rasistu.

Výjimka pro parodie a satiru by podle deníku The Washington Post mohla vést k ošemetným polemikám při posuzování manipulativního obsahu. Mezery v novém opatření Facebooku vzbuzují

obavy také u profesora Kalifornské univerzity Hanyho Farida, který se specializuje na analýzu digitálních záběrů. Zákaz je podle něj příliš „úzce vymezen“.

„Tato zavádějící videa (s Pelosiovou a Bidenem) byla vytvořena jednoduchými technologiemi a neopírala se o metody založené na umělé inteligenci, přesto byla přinejmenším tak zavádějící jako deepfake,“ řekl Farid. „Proč se zaměřovat pouze na deepfakes a ne na širší problém s úmyslně manipulativními videi?“ tázal se.

Facebook se hájí tím, že v případech, kdy daný příspěvek neporušuje pravidla platformy, může být stále video po analýze označeno za „falešné nebo částečně falešné“. Provozovatel této sociální sítě následně omezí distribuci takového příspěvku a připojí k němu upozornění nezávislých kontrolorů. Podle Bickertové je uvedený postup lepší, než kdyby se videa mazala, protože „by stále byla dostupná jinde na internetu“.