

KLONOVÁNÍ LIDSKÉHO HLASU S POMOCÍ AI OTEVÍRÁ PROSTOR PRO NOVÉ DRUHY DEZINFORMACÍ. NAKLONOVAT HLAS KOHOKOLI SE STALO EXTRÉMNĚ SNADNÉ.

Kamil KOPECKÝ

V posledním půl roce razantně narostl počet veřejně dostupných nástrojů, které uživatelům s pomocí umělé inteligence umožňují “naklonovat” lidský hlas. Samotný proces trénování a generování autentického lidského hlasu s pomocí umělé inteligence trvá pouze několik minut. Výsledek je pak skutečně od původního hlasu téměř k nerozeznání, což otevírá celou řadu možností, jak lze tuto technologii využít, ale třeba také zneužít.

Přednedávnem Český rozhlas informoval své čtenáře o tom, že si nechal experty ze Západočeské univerzity vytvořit klon hlasu Karla Gotta, pomocí kterého pak dnes již zesnulý umělec virtuálně “načetl” část své autobiografie (tzv. projekt [Gott navždy](#)). Tvůrci se při klonování hlasu drželi přísných etických principů, vše vzniklo se souhlasem rodiny zesnulého, zároveň celá situace představovala i zajímavý právní problém. Tvůrci také garantovali, že hlas bude syntetizovat pouze to, co zesnulý umělec skutečně řekl či napsal.

Problémem je, že vytvořit podobný záznam s využitím umělé inteligence je v současnosti velmi jednoduché a nástrojů, které umožňují naklonovat lidský hlas je celá řada. Jejich používání je pak často zdarma, případně za velmi mírný poplatek. K tomu, abyste dokázali lidský hlas naklonovat, stačí jen několikaminutový zvukový záznam, který obsahuje mluvené slovo konkrétního

člověka. Extrahovat lidský hlas však můžete také např. z videa. Umělá inteligence záznam lidského hlasu zanalyzuje a vygeneruje zvukový model. Poté stačí umělé inteligenci napsat, co přesně má daný syntetizovaný hlas říkat, a výsledek máme během chvíle k dispozici a ke stažení.

V praxi pak můžeme velmi **rychle vytvářet syntetizované hlasy umělců, politiků či jiných významných osobností, vkládat jim do úst slova, která nikdy nepronесли, a vydávat je za skutečnost**. To vše bude opět klást zvýšené nároky na uživatele online médií, kteří si musí uvědomit, že napodobit konkrétní lidský hlas je velmi snadné. Problém nastane také v případě odposlechů – jak prokázat, že odposlech skutečně zachycuje konkrétního člověka, a ne pouze jeho syntetizovaný hlas?

Ukázka naklonovaného hlasu vytvořená za 6 minut

Problém mohou představovat také telefonáty – např. Jak například poznat, že nám zprávu v hlasové schránce zanechalo právě naše dítě, a nejde o jeho klonovaný hlas? Jak poznat, že nám skutečně volá naše známá paní Nováková z České pojišťovny a ne její klon?

V současnosti není automatizovaná syntéza nahraného lidského hlasu ještě zcela dokonalá, AI chybí především v případě frázování či intonace, na druhou stranu vývoj jde stále kupředu a můžeme předpokládat, že za rok budou volně dostupné nástroje a jejich výstupy daleko lepší než nyní. A rozpoznat autentický hlas živého člověka od hlasu syntetizovaného se stane stále obtížnější.

Tím se dostáváme k problému mediální a informační gramotnosti – **jsme jako společnost připraveni na nárůst tohoto typu (dez)informací ve veřejném prostoru? Dokážeme odhalit, že s námi jejich prostřednictvím někdo manipuluje? Dokážeme vyvinout nástroje, které by umožňovaly detekci naklonovaných hlasů? Jak se přizpůsobíme nové situaci?**

Pro E-Bezpečí
Kamil Kopecký
Univerzita Palackého v Olomouci