Opolnomočenje uporabnikov proti tehnologiji AI Deepfake

Nevidna grožnja za podjetja

Leta 2024 so se “deepfake” napadi zgodili vsakih

5 minut

244%

letno povečanje ponarejanja digitalnih dokumentov[1]

49%

podjetij je med novembrom 2023 in novembrom 2024 doživelo tako avdio kot video “deepfake” napade[2]

61%

vodstvenih delavcev v letu 2024 ni vzpostavilo nobenih protokolov za obvladovanje tveganj tehnologije “deepfake”[3]

Letna porast tehnologije “deepfake” (Q1 2023 – Q1 2024) je prisotna v digitalno intenzivnih panogah:

iGaming	Tržnice	Fintech
1520%	**900%**	**533%**

Kripto	Svetovanje
217%	**138%**[4]

Tehnologija AI Deepfake postaja čedalje bolj sofisticirana

59%

anketirancev je poročalo o težavah pri ločevanju med vsebino, ki jo ustvarijo ljudje, in tisto, ki jo ustvari umetna inteligenca

84%

vprašanih, ki poznajo generativno UI, meni, da bi morala biti vsebina, ustvarjena z umetno inteligenco, vedno jasno označena[5]

Protiukrepi blagovnih znamk

Za boj proti vse večji grožnji tehnologije “deepfake” so podjetja, kot je HONOR, razvila tehnologijo za zaznavanje “deepfakov” in sodelujejo z različnimi partnerji pri vzpostavljanju standardov in krepitvi regulacije. Globalni trg zaznavanja “deepfakov” naj bi rasel za 42 % letno, s 5,5 milijarde USD v letu 2023 na 15,7 milijarde USD v letu 2026.[6]

HONOR-jeva rešitev

- S poudarkom na človeku prijaznih inovacijah HONOR uporablja umetno inteligenco za izboljšanje uporabniške izkušnje. Zaradi naraščajočega vpliva in tveganj tehnologije “deepfake “ je HONOR leta 2024 predstavil prvo funkcijo UI za zaznavanje posnetkov “deepfake” v industriji.
- Ta funkcija zaznava napake, ki so nevidne človeškemu očesu, kot so očesni stik, osvetlitev, jasnost slike in predvajanje videoposnetkov.
- Če funkcija zazna sintetično ali spremenjeno vsebino, bo uporabnik takoj prejel opozorilo o tveganju, kar ga bo zaščitilo pred nadaljnjo izpostavljenostjo potencialnim posnetkom “deepfake”.

Drugi protiukrepi v industriji

Adobe, Arm, Intel, Microsoft in Truepic so ustanovili organizacijo C2PA za razvoj tehničnih standardov za preverjanje izvora in zgodovine digitalne vsebine, vključno z UI-ustvarjenimi vsebinami.[7]

Microsoft uporablja zaščite UI za preprečevanje zlorab tehnologije “deepfake”, vključno z zamegljevanjem obrazov na fotografijah, naloženih v Copilot.[8]

Truepic tehnologija je vgrajena v Snapdragonov Trusted Execution Environment, ki zagotavlja, da so izvirne podrobnosti medijev preverjene in kriptografsko zapečatene že ob njihovi ustvaritvi.[9]

McAfee uporablja Qualcommov Snapdragon X Elite NPU za lokalno analizo in zaznavanje “deepfakov”, kar omogoča hitrejšo detekcijo in ohranja zasebne uporabniške podatke izven oblaka.[10]

1 Entrust Cybersecurity Institute – 2025 Identity Fraud Report
2 Regula – The Deepfake Trends 2024
3 Business.com - Deepfake Threats: Survey Reveals Business Impacts
4 KPMG – Deepfake – How real is it?
5 Deloitte - Deepfake disruption: A cybersecurity-scale challenge and its far-reaching consequences
6 Deloitte - Deepfake disruption: A cybersecurity-scale challenge and its far-reaching consequences
7 Coalition for Content Provenance and Authenticity
8 Microsoft Transparency Note for Microsoft Copilot
9 Truepic - Future of Transparency in Gen AI Starts with Smartphones
10 Qualcomm - AI-powered productivity: Windows on Snapdragon X Elite welcomes game-changing apps