

# Ako odhaliť a pochopiť obsah vytvorený AI a digitálne upravený obsah

AI modely, ktoré sa jednoducho používajú a je k nim ľahký prístup, pomáhajú ľuďom v učení aj tvorbe obsahu. Zároveň však **znásobujú aj riziká**, ktoré **nepravdivé informácie** nesú pre otvorenú spoločnosť a demokraciu. Je dôležité, aby sme ochránili svoj informačný priestor od zahltenia nepravdivými informáciami, ktoré vytvorila umelá inteligencia alebo boli digitálne upravené.

Súčasťou riešenia sú nové technológie, ako softvéry na detekovanie AI alebo iniciatívy preukázania pôvodu obsahu. **Ale technologické riešenia nie sú dokonalé a potrebujeme prácu nezávislých fact-checkerov, ktorí spoločnosti poskytujú overené informácie.**

Ponúkame preto krátky prehľad, ako profesionálni fact-checkeri odhaľujú a vyvracajú nepravdivé informácie vytvorené umelou inteligenciou a čo sa od nich môžete naučiť.

## Obsah vytvorený AI je na vzostupe

Dnes tvoria nepravdivé informácie vytvorené umelou inteligenciou len malú časť obsahu, ktorý overujú profesionálni a nezávislí fact-checkeri. Častejšie sa venujú digitálne upravenému obsahu.

Podľa interného prieskumu členov EFCSN sa však **väčšina fact-checkerov zhodla, že význam digitálne upraveného obsahu alebo obsahu vytvoreného AI bude v budúcnosti rásť**. Nedávne príklady pred voľbami do Európarlamentu túto predpoveď podporujú.

**VYSVETLENIE:** *Digitálne upravený* môže byť akýkoľvek obsah, ktorý bol natoľko upravený, že správa, ktorú pôvodne niesol, sa zmenila alebo ňou bolo manipulované – vrátane úprav nástrojmi AI. Nezahŕňa vylepšenia kvality alebo zrozumiteľnosti.

*Vytvorený AI* je akýkoľvek obsah, ktorý bol vytvorený pomocou nástrojov umelej inteligencie.



# Technológia postupuje rýchlo, ale nemôžeme sa na ňu spoliehať.

Odborníci na AI a fact-checkeri sa zhodujú: **nástroje na detekciu AI zatiaľ nestačia na to, aby odhalili obsah vytvorený AI a digitálne upravený obsah.**

Skôr, než použijete AI detektor, odborníci odporúčajú, aby ste sa zoznámili s nástrojmi na tvorbu a detekciu AI. Keď fact-checkeri rozumejú, akým spôsobom sa modely učia a prepoja to s trochu štatistiky, vedia lepšie posúdiť silné a slabé stránky daného nástroja, aj pravdepodobnosť úspechu, či umelú inteligenciu odhalí. **Napriek tomu môžu byť tieto nástroje užitočným prvým krokom.**

## Pôvod obsahu

Iniciatívy, ako sú napríklad špecifikácie C2PA, môžu pomôcť overiť zdroj a pôvod mediálneho obsahu, ale vodoznaky a označovanie nie sú nepriestrelné.

## Ako nás ovplyvňujú AI dezinformácie

*„Vždy, keď zareagujete inštinktívne, vynechali ste rozum.“*

– Christine Dugoin\*

**PSYCHOLÓGIA:** Vplyvové operácie často využívajú kognitívne skreslenia.

Je dôležité pochopiť svoje vlastné skreslenia a aj tie, ktorým podliehajú naše cieľové skupiny. Môže nám to v boji proti dezinformáciám pomôcť.

**CIELE:** Prečo by aktér so zlými úmyslami mohol využiť na šírenie dezinformácií umelú inteligenciu? Čo chce v skutočnom svete dosiahnuť?

- Rozšíriť svoj dosah na inú krajinu či komunitu?
- Vyhnúť odhaleniu alebo zaplaviť fact-checkerov vytvorením množstva variácií podobného obsahu?
- Ovplyvniť myslenie a presvedčenie ľudí tým, že vytvorí sieť falošných účtov a vyvolá dojem dôveryhodnosti?

\* Christine Dugoin je výskumníčkou v oblasti informačného vplyvu na La Sorbonne.

# Fact-checking vyžaduje komplexný prístup a dôsledné pochopenie témy

Ak nástroje na detekciu nefungujú, čo teda funguje? Je kľúčové pochopiť nielen obsah tvrdenia, ale aj kontext, v ktorom sa vyskytlo. Profesionálni fact-checkeri sú odborníci v týchto oblastiach výskumu. Tu je zopár tipov.

*„Nástroje detekcie nikdy nefungujú na 100% - pravdepodobne ani nikdy nebudú.“*

*- Henk van Ess\*\**



**POSÚDTE ZDROJ:** Viete overiť jeho identitu? O čom rozpráva a čo zdieľa? Kto reaguje na vytvorený obsah? Aký účinok môže mať na čitateľov?



**ZHODNOŤTE DÔVERYHODNOSŤ:** Overte si informácie v dôveryhodných zdrojoch, ako sú napríklad odborníci s praktickými skúsenosťami v oblasti. Dáva obsah na základe vašich vedomostí zmysel?



Použite nástroje **FORENZNEJ ANALÝZY MÉDIÍ** na doplnenie tradičnej investigatívnej žurnalistiky a výskumu dokumentov. Napríklad hromadné získavanie dát (tzv. data scraping), geolokáciu, biometrické rozpoznávanie alebo analýzu vzorcov.



**UČTE SA & PRISPÔSOBTE SA:** Tvorcovia nepravdivých informácií vytvorených AI sa sústavne rozvíjajú a zlepšujú. Prispôbte svoj prístup meniacim sa podmienkam.

## ZDIEĽAJTE SVOJU PRÁCU

Odborníci odporúčajú, aby ste spolu s overením dezinformácie zdieľali aj analýzu a odkazy na zdroje. Vďaka tomu môžu čitatelia sledovať postup investigatívy a porozumieť aj zložitejším naratívom. V niektorých prípadoch je investigatíva dôležitejšia, než samotný fakt, že obsah vytvorila umelá inteligencia.

\*\* Henk van Ess je expertom v oblasti OSINT a fact-checkingových techník.

# Rýchly návod: Čo robiť, čo nerobiť a na čo sa zamerať

Toto sú indície, ktoré môžu naznačiť, či je príspevok generovaný umelou inteligenciou alebo digitálne upravený. Spolu s ostatnými tipmi v tomto návode (kontext, techniky a nástroje na detekciu) vám pomôžu odhaliť pravdu za tým, čo vidíte.

## Text

- Často (ale nie vždy) má **lepšiu gramatiku** než človek.
- Môže použiť **príliš formálny jazyk** na prostredie sociálnych sietí.
- **Priveľa prísloviak alebo prídavných mien.**
- Nedostatok emócií, humoru, sarkazmu a zaužívaných fráz.
- Môžu **chýbať konkrétne detaily** (mená, dátumy, miesta) a originálne myšlienky.
- Najdôležitejšie: sú informácie v texte pravdivé?

## Video

- Nepoužívajte detektor AI obrázkov na zábery z videa.
- Sledujte **výrazy tváre a pohyb**, ako napríklad žmurkanie alebo či pohyb úst sedí so zvukom.
- Znakom môžu byť aj **ostré prechody a strihy**.

## Zvuk

- **Porovnajte podozrivé audio s autentickou hlasovou nahrávkou** pomocou nástrojov, ktoré rozoznávajú rozdiely v reči, v dychu alebo v intonácii.
- Detektory nepoužívajte na nekvalitné zvukové nahrávky s hlukom a šumom v pozadí.
- Varovným znakom je **neprirodzená alebo mechanicky znejúca reč**, málo prestávok a nádychov.

## Obrázky

- Hľadajte **neprirodzené detaily**: dokonalá pleť, rozmazané pozadie, neprirodzená krása alebo osvetlenie, či prsty navyše.
- **Hľadajte vodoznak** známych generátorov obrázkov.
- Venujte pozornosť detailom: dáva obrázok zmysel? Súvisí s kontextom?
- Detektor používajte na **obrázky s vysokým rozlíšením alebo pôvodné verzie** obrázkov, nie verzie, ktoré boli opakovane prezdieľané.