

Hogy ismerd fel és értsd meg az AI-generált vagy digitálisan manipulált tartalmakat?

A bárki számára elérhető, könnyen használható AI-modellek sokat segíthetnek például a tanulásban, tartalomgyártásban. Azonban **növelhetik a dezinformációt félretájékoztatást is**, ami káros a társadalom és a demokratikus diskurzusok szempontjából. Ezért is fontos megakadályozni, hogy a közös információs terünket elárasszák az AI-generált, digitálisan manipulált dezinformációs tartalmak.

A problémát részben megoldhatják bizonyos új technológiák, mint például a tartalom eredetét meghatározó programok és felderítő szoftverek. **A technológiai megoldások azonban korántsem tökéletesek, mellettük szükségünk van a független tényellenőrök munkájára is, hogy minél több emberhez eljussanak az ellenőrzött, hiteles tények.**

Íme egy gyors áttekintés arról, hogy a hivatásos, független tényellenőrök mit tesznek az AI-generált, félretájékoztató vagy dezinformáló tartalmak azonosítása és cáfolása érdekében, és te mit tanulhatsz tőlük.

Az AI-tartalmak térnyerése

A hivatásos, független tényellenőrök munkájának egyelőre csak kis szeletét teszi ki az AI-generált dezinformáció ellenőrzése. A digitálisan módosított tartalmak nagyobb arányban fordulnak elő a munkájukban.

Az EFCSN tagjai körében végzett belső felmérés szerint azonban **a legtöbb tényellenőr egyetért abban, hogy az AI-generált és a digitálisan manipulált tartalmak jelentősége egyre inkább növekszik** majd a jövőben. Ezt a prognózist az európai választásokhoz kapcsolódó [közelmúltbeli példák](#) is igazolják.

JÓ TUDNI: A digitális módosítás olyan tartalmakra vonatkozik, amelyeket azért változtattak meg jelentősen, hogy az eredetileg közvetített üzenetet manipulálják vagy átalakítsák. Ez igaz az AI-eszközökkel történő szerkesztésre is. Viszont nem vonatkozik az érthetőség vagy a minőség érdekében végzett szerkesztésekre.

Az AI-generált jelzőt azokra a tartalmakra használjuk, amelyeket mesterséges intelligencia hozott létre.



A technológia gyorsan fejlődik, de nem hagyatkozhatunk kizárólag erre

Az AI-szakértők és a hivatásos tényellenőrök egyetértenek abban, hogy **az AI-felismerő eszközök önmagukban nem elegendők a mesterségesen előállított vagy digitálisan módosított tartalmak felismeréséhez, megcáfolásához.**

A szakértők azt javasolják, hogy az AI-felismerő programok használata előtt ismerkedjünk meg a tartalomgeneráló és -felismerő programokkal. A példák és statisztikák megértése az első lépés az eszköz erősségeinek, gyengeségeinek és sikerességének felmérésében. **Ezek az eszközök mindenképpen hasznos kiindulópontot jelentenek.**

A tartalom eredetét vizsgáló programok (például a C2PA Specifications) segíthetnek a tartalmak forrásának és előzményeinek feltérképezésében, ám a vízjelek és más hitelesítő technikák nem mindig megbízhatóak.

Az AI-generált dezinformáció hatása

„Minden egyes zsigeri reakció az elmélyült gondolkodás megkerülésével születik.”

- Christine Dugoin*

PSZICHOLÓGIA: A befolyásolás gyakran az előítéletek és a pszichológiai torzítások kihasználására épít. Önmagunk és közönségünk megértése segíthet a dezinformációval szembeni fellépésben.

CÉLOK: Vajon miért használ valaki AI-eszközöket hamis tartalmak létrehozásához és terjesztéséhez? Milyen szándék állhat a dezinformáció mögött?

- Hatást akarnak gyakorolni egy másik országra vagy közösségre?
- Sok hasonló állítás generálásával el akarják kerülni a lebukást, vagy túl akarják terhelni tényellenőröket?
- Hamis profilok hálózatával, a hitelesség látszatát keltve akarják befolyásolni a közönség gondolkodását és meggyőződését?

*Christine Dugoin az információs befolyásolás témájának kutatója a Sorbonne-on.

A tényellenőrzés sokoldalú megközelítést és komplex megértést igényel.

Mi van akkor, ha a felismerő eszközök nem működnek? Fontos, hogy az állítások kontextusát éppúgy megértsük, mint a tartalmukat. A hivatásos tényellenőrök mesterei a nyomozási és kutatási technikáknak. Íme, pár hasznos tipp tőlük.

“A felismerő eszközök soha nem tudnak 100 százalékosan működni – Nem is számítok rá, hogy valaha fognak.”
– Henk van Ess**



VIZSGÁLD MEG A FORRÁST: Meg tudod erősíteni a személyazonosságukat? Miről beszélnek és miket osztanak meg? Milyen hatással lehet az ilyen tartalom az olvasóra?



GYŐZDJ MEG A HITELESSÉGRŐL: Ellenőrizd az információt hiteles forrásokból, például olyan szakértők segítségével, akik gyakorlati tapasztalattal rendelkeznek az adott területen. Van-e értelme a leírtaknak a saját tapasztalataid alapján?



Alkalmazd a **FORENZIKUS MÉDIA** elemzési módszereit, amelyek kiegészíthetik a hagyományos oknyomozó módszereket. Néhány példa: adatgyűjtés, geolokáció, biometrikus azonosítás, mintázatok elemzése és még sok más.



TANULJ ÉS FEJLŐDJ: Azok, akik a mesterséges intelligenciát félrevezető információkat terjesztésére használják, folyamatosan tanulnak, fejlődnek. Alkalmazkodj te is a folyamatosan változó környezethez.

OSZD MEG A MUNKÁDAT

A szakértők azt javasolják, hogy a megcáfolt állítások mellett átláthatóan szerepeljen az elemzés menete és a források megfelelő hivatkozása is. Ez segíthet az olvasóknak a kutatás követésében és a bonyolultabb témák megértésében. Bizonyos esetekben maga a nyomozás fontosabb, mint az, hogy a tartalom mesterséges intelligenciával készült-e.

** Henk van Ess az OSINT és a tényellenőrző technikák szakértője.

Gyors tájékoztató

Mit csinálj és mit ne, mi legyen gyanús?

A következőkben bemutatjuk azokat a gyanús jeleket, amelyek elárulhatják, ha egy tartalom mesterséges intelligencia által generált vagy digitálisan manipulált. Ezek felismerésével és az útmutatóban említett egyéb tippek (kontextus, oknyomozó technikák és azonosítási módszerek) segítségével könnyebben felfedezhetjük a látottak és olvasottak mögött rejlő igazságot.

Szöveg

- Gyakran (de nem mindig) **jobb helyesírással** és nyelvtani eszközökkel él, mint az ember által írt tartalmak.
- Sokszor **túl formális a szöveg megfogalmazása vagy strukturált a nyelvezete**, főleg a közösségi médián fellelhető posztok esetén.
- **Túl sok határozószót vagy melléknevet használ.**
- Feltűnő lehet az emberi érzelmek, humor, szarkazmus és szleng hiánya.
- **Kevés a konkrét részlet** (név, dátum, helyszín) vagy eredeti ötlet.
- És a legfontosabb, hogy mindig gondoljuk végig: helytállóak-e a szövegben szereplő tények?

Videó

- A videókból kivágott állóképeken ne alkalmazzunk olyan programokat, amelyeket az AI-generált képek azonosítására használunk.
- Ellenőrizzük az **arckifejezéseket, az emberi mimikát és mozgást**, például a pislogást, valamint azt, hogy a száj mozgása illeszkedik-e a hanghoz.
- Ellenőrizzük, hogy felfedezhetők-e **éles vágások vagy átmenetek**.

Hang

- **Hasonlítsuk össze a gyanús hanganyagot egy hiteles mintával** olyan eszközök segítségével, amelyek képesek felismerni a beszéd- és légzésmintákat, az intonációt, stb.
- Az azonosító eszközök használatakor kerüljük a statikus vagy háttérzajjal rendelkező, rossz minőségű hangmintákat.
- Mesterséges intelligenciára utalhat a **természetellenes vagy mechanikus beszédmód**, a szünetek vagy a természetes lélegzetvétel hiánya.

Képek

- Keressünk **rendellenes részleteket**: tökéletes bőr, elmosódott háttér, természetellenes szépség vagy fény és olyan furcsaságok, mint például ha valakinek 6 ujjja van.
- **Vizsgáljuk meg, hogy a képeken nem fedezhetők-e fel ismert képgenerátorok vízjelei.**
- Figyeljünk az ábrázolt dolgok részleteire: megfelelnek a valóságnak? Logikusak?
- Ha képfelismerő programokat használunk, **ne a már újra megosztott képeket elemezzük, hanem az először feltöltött verziókat, és a lehető legnagyobb felbontású képeket.**