



CENTRUL PENTRU
COMUNICARE STRATEGICĂ
ȘI CONTRACARARE A DEZINFORMĂRII

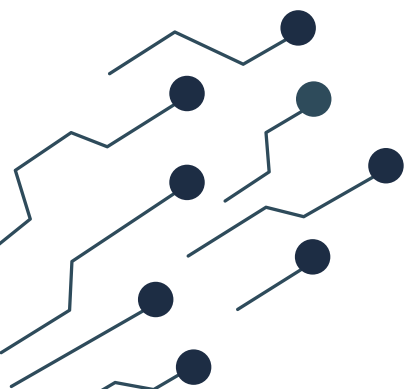
DEEPPFAKE

Ghid pentru utilizatori



CUPRINS:

Cuprins	1
Scurt istoric, evoluție și cadrul legal.....	2
Ce este conținutul generat de IA?.....	3
Cum să identificați materialele false generate de inteligență artificială.....	4
Recomandări pentru cetățeni și societatea civilă.....	6
Concluzii.....	7





SUMAR

Pe măsură ce tehnologia inteligenței artificiale (IA) continuă să avanseze, la fel și metodele utilizate pentru a genera informații înșelătoare sau false. Conținutul generat de inteligență artificială, adesea numit „deepfakes” sau media sintetică, a devenit un instrument pentru campanii de dezinformare care pot înșela cu ușurință publicul. Pentru a vă proteja și a contribui la integritatea ecosistemului informațional din Moldova, este esențial să învățăm cum de identificat potențialele materiale false generate de inteligența artificială de conținutul real.

Acest material furnizat de Centrul pentru Comunicare Strategică și Contracurarea Dezinformării, prezintă câțiva pași și acțiuni pentru a vă ajuta să recunoașteți materialele generate de IA și să le distingeți de conținutul autentic.

CADRUL LEGAL ȘI DE REGLEMENTARE

În Republica Moldova, utilizarea tehnologiilor deepfake poate fi încadrată în mai multe dispoziții legale existente. Crearea sau diseminarea de materiale deepfake care induc în eroare prin atribuirea unei identități false unei persoane poate constitui fals de identitate (art. 177¹ Cod Penal) sau furt de identitate (art. 260⁷ Cod Penal). De asemenea, utilizarea deepfake-urilor în scop de defăimare, răspândirea de informații false sau instigare la ură și discriminare poate fi sancționată ca calomnie (art. 70 Cod Contravențional) sau instigare la discriminare (art. 701 Cod Contravențional).

SCURT ISTORIC ȘI EVOLUȚIE

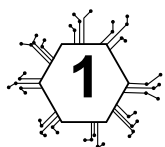
Într-un interval de timp surprinzător de scurt, de la primele apariții în 2017 și până în prezent, tehnologia deepfake a devenit extrem de performantă, permițând manipulări audio-vizuale de un realism izbitor. Factori precum dezvoltarea inteligenței artificiale, puterea de procesare tot mai mare și costurile tot mai mici au făcut ca această tehnologie să devină nu doar accesibilă, ci și o amenințare serioasă pentru securitatea cibernetică și încrederea publică. Acum, când devine tot mai dificil să deosebim ce este adevărat de ce este fals, atât organizațiile, cât și fiecare dintre noi trebuie să acționeze preventiv pentru a contracara riscurile și abuzurile asociate.



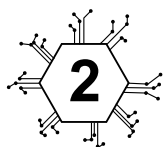


CE ESTE CONȚINUTUL GENERAT DE IA?

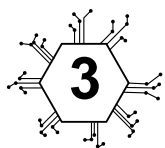
Conținutul generat de IA se referă la text, imagini, audio sau video create sau manipulate de algoritmi de inteligență artificială. Acest tip de conținut este adesea conceput pentru a arăta și a suna realist, ceea ce face dificilă distincția de materialul autentic. Exemplele includ



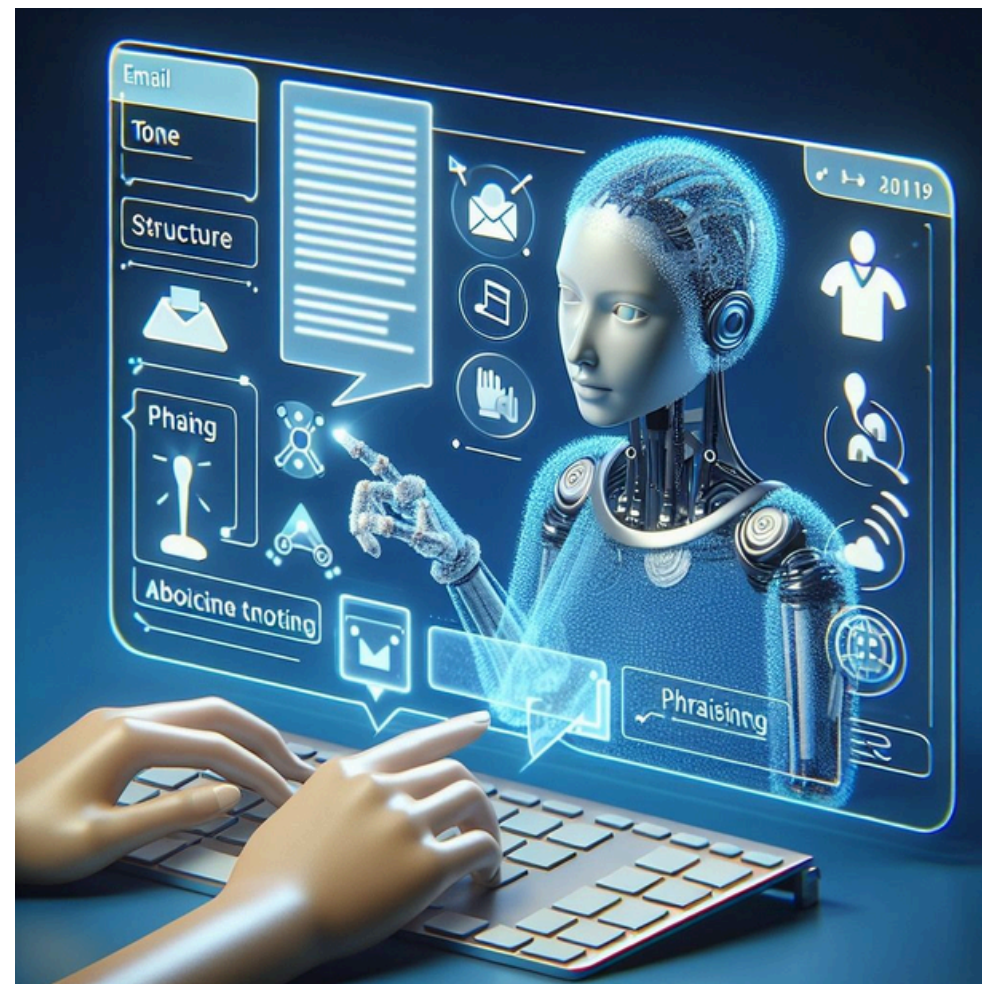
Deepfakes: videoclipuri care arată oameni făcând sau spunând lucruri pe care nu le-au făcut niciodată, generate folosind inteligența artificială.



Text generat de IA: articole sau postări pe rețelele sociale care imită scrisul uman, dar conțin informații înșelătoare sau incorecte.



Audio sintetic: înregistrări vocale generate de inteligență artificială care imită vocile reale.



Aceste forme de conținut pot fi folosite pentru dezinformare, propagandă sau pentru a afecta reputația persoanelor și organizațiilor.

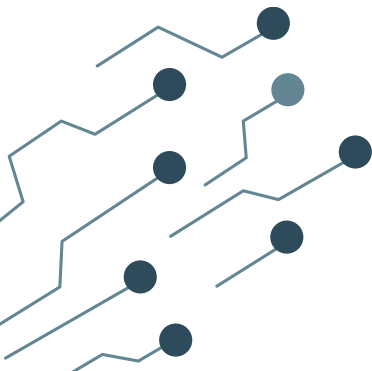


CUM SĂ IDENTIFICAȚI MATERIALELE FALSE GENERATE DE INTELIGENȚĂ ARTIFICIALĂ

Iată pașii și acțiunile concrete pe care le puteți lua pentru a identifica conținutul fals generat de IA:

1. Examinați sursa conținutului

- Verificați sursa: verificați dacă conținutul provine de la o instituție media de renume sau de la un cont oficial. Fiți precauți cu informațiile de pe site-uri web sau conturi de social media necunoscute.
- Simbolul de verificare: sursele de încredere au adesea simboluri de verificare (bife albastre) pe platforme precum Facebook sau Twitter. Dacă conținutul provine dintr-o sursă neverificată sau suspectă, fiți sceptici.
- Verificați cu media de încredere: dacă întâlniți conținut șocant sau neobișnuit, verificați dacă sursele de știri de încredere și consacrate raportează aceleași informații.



2. Analizați calitatea vizuală și audio

- Anomalii video: în videoclipurile deepfake, acordați atenție detaliilor precum iluminarea inconsecventă, mișcările nenaturale sau expresiile faciale nepotrivite. Este posibil ca ochii să nu clipească în mod natural, iar vorbirea poate să nu se sincronizeze perfect cu mișcările buzelor.
- Inconsecvențe audio: dacă conținutul este audio, auziți pauze ciudate, tonuri nepotrivite sau inflexiuni nenaturale care nu sună ca un om real.
- Inconsecvențe ale imaginii: imaginile generate de AI pot afișa nereguli, cum ar fi zone neclare sau pixelate, reflexii nerealiste sau obiecte distorsionate, în special în fundal sau detalii mai fine (cum ar fi mâinile sau ochii).



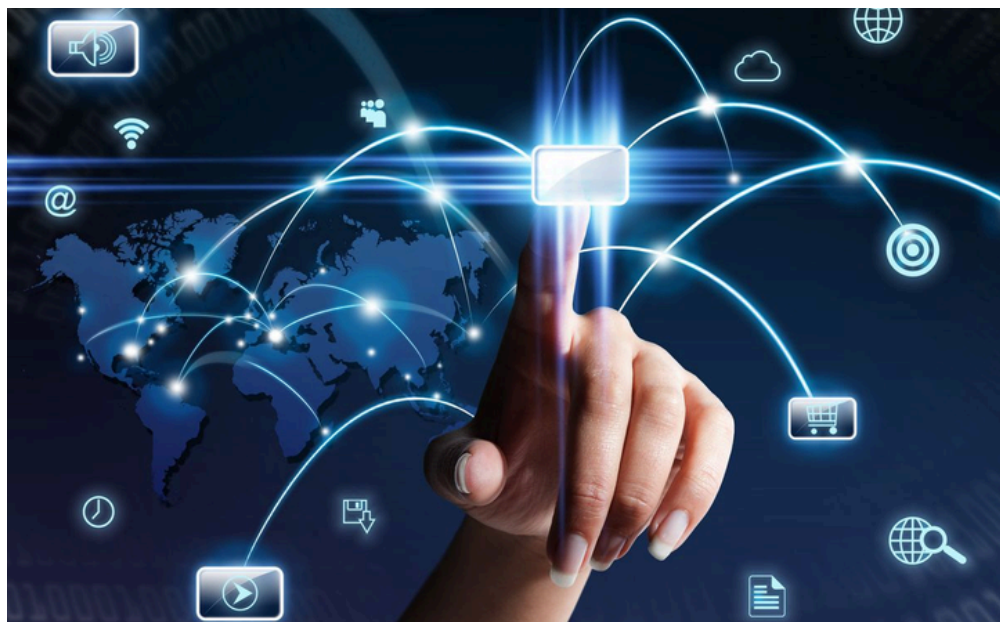


3. Inspectați conținutul pentru defecte logice

Afirmații ciudate: fiți sceptici față de afirmațiile senzaționale sau scandaloase care par concepute pentru a provoca o reacție emoțională. Textul generat de IA exagerează adesea pentru a manipula emoțiile.

Limbaj inconsecvent: textul generat de IA poate fi lipsit de coerență sau poate schimba brusc tonul. Căutați fraze incomode, repetiții sau greșeli gramaticale care nu se potrivesc cu stilul general.

Prea frumos pentru a fi adevărat: unele imagini sau videoclipuri generate de inteligență artificială sunt concepute pentru a părea prea perfecte. Dacă ceva pare prea înscenat sau lustruit, ar putea fi medii sintetice.



4. Utilizați instrumente online pentru a verifica conținutul

Instrumente de detectare a deepfake-urilor: utilizați platforme online sau extensii de browser special concepute pentru a detecta deepfakes sau conținut manipulat, cum ar fi Deepware Scanner sau Sensity.

Căutare inversă de imagini: dacă bănuiți că o imagine ar putea fi generată sau modificată de inteligență artificială, efectuați o căutare inversă de imagini folosind platforme precum Google sau TinEye. Acest lucru vă poate ajuta să găsiți sursa originală a imaginii și să vedeți dacă a fost modificată.

Verificarea metadatelor: utilizați instrumente online pentru a verifica metadatele videoclipurilor sau imaginilor. Metadatele pot oferi indicii despre când, unde și cum a fost creat conținutul.

5. Fiți conștienți de context și de timp

Referință încrucișată cu evenimente de știri: comparați conținutul în cauză cu evenimente din lumea reală. Dacă conținutul nu se aliniază cu cronologii sau evenimente confirmate, probabil că este fals.

Momentul și intenția: campaniile de dezinformare lansează adesea conținut generat de inteligență artificială în perioade sensibile, cum ar fi alegerile, crizele sau tulburările politice. Fiți foarte precauți cu conținutul în aceste perioade.



RECOMANDĂRI PENTRU CETĂȚENI ȘI SOCIETATEA CIVILĂ

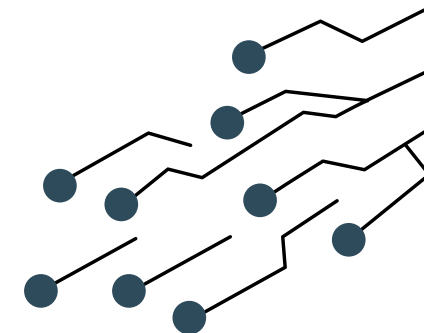
1. **Fii critic și informat:** dezvoltă obiceiul de a pune la îndoială tot ceea ce vezi online. Nu distribuți conținut înainte de a-i verifica autenticitatea, mai ales dacă evocă răspunsuri emoționale puternice.

3. **Urmăriți surse de încredere:** rămâneți conectat la instituții media de încredere, canale guvernamentale și organizații de renume. Canalele oficiale vor oferi actualizări exacte în timpul crizelor sau evenimentelor.

5. **Raportați conținutul suspect:** dacă întâlniți materiale suspecte generate de inteligență artificială, raportați-le platformelor de fact-checking sau împărtășiți-vă îngrijorările cu instituțiile abilitate (Centrul) sau supraveghetorilor mass-media. Centrul pentru Comunicare Strategică și Combaterea Dezinformării oferă cetățenilor o platformă pentru a raporta campaniile de dezinformare și conținut dăunător la adresa de email dialog@stratcom.md.

2. **Educați-i pe ceilalți:** răspândiți conștientizarea conținutului generat de IA și despre cum să-l identificați în rândul prietenilor, familiei și comunității. Dezinformarea se răspândește ușor, dar la fel se răspândește și adevărul atunci când oamenii sunt informați.

4. **Participați la programe de educație media:** participați la ateliere sau cursuri de educație media oferite de Centrul pentru Comunicare Strategică și Combatere a Dezinformării pentru a vă îmbunătăți abilitățile de identificare a conținutului fals.



CONCLUZIE

Într-o lume în care conținutul generat de IA devine din ce în ce mai răspândit, identificarea materialelor false este crucială pentru protejarea ecosistemului informațional și a democrației din Moldova. Rămânând vigilenți, folosind instrumentele potrivite și făcând schimb de conștientizare, cetățenii și societatea civilă pot juca un rol esențial în protejarea adevărului și a transparenței în lumea noastră digitală.

Centrul pentru Comunicare Strategică și Combatere a Dezinformării se angajează să vă ajute să navigați prin aceste provocări și încurajează toți cetățenii Republicii Moldova să contribuie la reziliența mediului nostru informațional.

Acest articol servește drept bază pentru a spori gradul de conștientizare și capacitatea publicului de a detecta dezinformarea generată de IA, contribuind la protejarea securității naționale a Moldovei și la promovarea unei societăți informate.

