

Viral, dar fals: Cum ne păcălesc imaginile create de inteligența artificială

de Dorin LUCA

Într-un material de anul trecut, vă aduceam în atenție capcanele nostalgiei de pe *Facebook*: imagini care insuflă ideea unui trecut glorios, cu mesaje eurosceptice subliminale, venite din zone obscure.

Între timp, odată cu emergența instrumentelor de inteligență artificială, fenomenul s-a răspândit și a crescut în proporții.

Dacă sunteți utilizatori de *Facebook*, probabil ați observat fenomenul imaginilor create cu ajutorul inteligenței artificiale (AI), cu „opere de artă” din usturoi sau morcovi, care plâng că nu sunt apreciate: *Am făcut asta pentru voi, dar nimeni nu apreciază!* – se lamentează acestea, fiind urmate de mii de aprecieri, distribuiri și comentarii.

Postările de acest gen sunt capcane. O anchetă *Europa Liberă* detaliază felul în care astfel de pagini funcționează și care sunt obiectivele. Postările acestora sunt extrem de atractive și au potențialul de a deveni virale, determinând utilizatorii să aprecieze, să comenteze și să distribuie orbește.

Cum funcționează aceste pagini? Deși în unele cazuri au sute de mii de urmăritori și denumiri în română, paginile sunt administrate din alte țări, iar descrierile postărilor sunt adesea traduse automat în mai multe limbi, indicând o rețea bine organizată, care activează și în alte părți.



Paginile sunt active în mod constant și postează programat, de câteva ori pe zi. *Europa Liberă* notează că unele dintre ele au fost create acum câțiva ani, însă publicarea masivă a imaginilor generate cu AI a început în luna noiembrie 2024.

Aceste pagini promovează un set de mesaje populare și testează în permanență publicul pentru a identifica noi direcții.

Potrivit istoricului Radu Oltean, o categorie importantă de mesaje promovează nostalgia rurală și valorile tradiționale, punând în contrast trecutul „perfect” cu prezentul „superficial”, cu mesaje precum „înainte era mai bine, oamenii erau mai fericiți”.



De asemenea, unele postări aduc sărăcia la rang de virtute, inducând ideea că oamenii nu mai muncesc sau că munca „adevărată” nu mai este apreciată. Scopul este de a induce ideea că prezentul a dus la erodarea valorilor morale, de a crea conflicte între generații și între diferite categorii sociale.

O altă categorie este reprezentată de prostiile generate de AI: imagini cu animale inexistente, natură idealizată, „opere de artă” create din fructe și legume, fantezii vizuale cu trimitere divină. Toate acestea au drept scop reducerea gândirii critice, desensibilizarea audienței, pregătirea pentru acceptarea fără multe întrebări a unor falsuri vizuale.

În cele din urmă, o categorie semnalată de Oltean se referă la mesajele conspiraționiste și ultranaționaliste. Acestea folosesc conspirații uzuale precum *chemtrails*, oculă mondială, radiații etc. Pe lângă ele, vin mesajele despre valori strămoșești și credință: Zalmoxis, spiritualitatea românească, ortodoxia.

Europa Liberă transmite că, din punct de vedere al comunicării digitale, scopul imediat al acestor pagini este creșterea artificială a audienței și a interacțiunilor.



Mesajele sunt ajustate pentru a vedea care dintre ele nasc cele mai puternice reacții și sentimente, care, la rândul lor, creează și cea multă interacțiune. Astfel se formează comunități, în jurul unor valori (credință, tradiție, conspirație), cu implicare emoțională puternică. De aici sunt culese date despre audiență.

În unele cazuri, astfel de pagini sunt crescute până la câteva sute de mii de fani pentru ca ulterior să fie vândute – moment în care sunt folosite în scopuri politice, pentru candidați care nu au nicio legătură cu subiectul inițial al paginii.

Practic, denumirea și scopul paginilor sunt schimbate fără ca utilizatorii să observe, fiind ulterior expuși propagandei politice agresive. În alte cazuri, paginile activează mai subtil, frecvența mesajelor cu tentă politică sporind în perioade de instabilitate politică sau la comanda proprietarilor.



Există și cazuri în care paginile, după ce adună un număr considerabil de fani, pur și simplu vând spațiu de publicitate, în scop comercial.

Efectele pe termen lung sunt multiple. Pe de o parte, mulți utilizatori sunt atrași de imaginile impresionante și se implică emoțional în comentarii și distribuiri, fără să își dea seama că interacționează cu conținut generat artificial.

Pe de altă parte, această tendință contribuie la dezinformare și la diminuarea capacității utilizatorilor de a distinge între realitate și conținut creat artificial. Fiecare postare activează emoții negative, de la frică, resentimente la nostalgie pentru un trecut idealizat, în final publicul devenind mai vulnerabil la manipulare.

În plus, fenomenul ridică probleme legate de etică și de impactul AI asupra ecosistemului digital. Deși imaginile generate de AI pot părea inofensive sau distractive, ele fac parte din strategii bine-gândite, mult mai ample. *Facebook* nu pare preocupat de fenomen.



Recent, utilizatorii *Reddit* au semnalat faptul că însuși creatorul rețelei, Mark Zuckerberg, a reacționat la o astfel de postare, ceea ce ne arată o slabă înțelegere a fenomenului chiar și din partea celor care ar trebui să gestioneze situația, precum și faptul că această strategie de manipulare nu este limitată doar la cazul României.

Sectă religioasă chineză, activă pe același model

În decembrie anul trecut, o anchetă jurnalistică din Ungaria mi-a atras atenția, pentru că părea să vorbească despre un fenomen pe care îl observasem și în România. Articolul de pe *Telex.hu* investiga proliferarea paginilor de *Facebook* cu tematică religioasă, care utilizau imagini generate de inteligența artificială pentru a-și crește interacțiunile.

Autorii subliniau că astfel de pagini publică frecvent imagini kitsch cu Iisus, îngeri sau alte simboluri religioase, însoțite de mesaje ce îndeamnă utilizatorii să comenteze cu „Amin” sau să interacționeze în alte moduri, promițând, în schimb... binecuvântări.

Investigația a relevat că multe dintre aceste pagini sunt administrate de membri ai sectei chineze Biserica Dumnezeului Atotputernic (BDA / *The Church of Almighty God*, CAG), care utilizează rețelele sociale pentru a răspândi învățăturile sale și pentru a recruta noi adepți la nivel internațional.

Această sectă este cunoscută pentru metodele sale controversate și pentru folosirea tehnologiei în scopuri de prozelitism.

În România, fenomenul a fost investigat recent de jurnaliștii de la *Europa Liberă*, care arătau că același strategii sunt folosite și la noi. BDA a creat numeroase pagini de *Facebook* cu tematică religioasă, unde distribuie mesaje spirituale atrăgătoare, uneori ilustrate cu imagini generate de AI.



Fanii activi ai acestor pagini sunt recrutați prin intermediul *Messenger* pentru a participa la predici online. Persoanele care contactează utilizatorii se ascund în spatele unor identități false, pretinzând că sunt din România, deși indicii despre identitatea lor străină există chiar în profilurile pe care le utilizează.

Astfel, BDA utilizează mediul online pentru a se infiltra în comunități religioase din România, profitând de utilizatorii susceptibili la manipulare.

Nu există soluții miracol pentru a scăpa de acest tip de manipulare. În lipsa unei reglementări a rețelelor de socializare, rămâne la latitudinea fiecăruia să fie vigilent și să pună sub semnul întrebării orice mesaj fantezist peste care poate da din întâmplare.

Un grup de *Facebook* – *AI de noi* – a identificat peste o mie de pagini și profiluri, începând cu 22 decembrie 2024, care distribuie conținut manipulator generat de AI. Fiecare dintre noi poate semna astfel de pagini, deși *Facebook*, în general, nu vede nimic greșit la ele, având în vedere că acestea generează interacțiuni profitabilă pentru platformă. Raportările, de multe ori, nu dau niciun rezultat.