INTELLIGENCE ȘI CULTURA DE SECURITATE

CONFERINȚA STIINȚIFICĂ STUDENȚEASCĂ
CONFERENCE PROCEEDINGS

VOLUMUL 4
2025

Editura Academiei Naționale de Informații
„Mihai Viteazul”

INTELLIGENCE
ȘI CULTURA DE SECURITATE

nr. 4 – 2025

– Conferința Științifică Studențească –

Editura Academiei Naţionale de Informaţii „Mihai Viteazul”

Bucureşti, 2025

Editura Academiei Naţionale de Informaţii „Mihai Viteazul”
© ANIMV
Bucureşti, 2025
Telefon: 0377720.000/1216
Fax: 0377721.134; 0377721.125
ISSN 2972 – 1350 ISSN-L 2971 – 8139

CUPRINS

ASTROTURFING ŞI RĂZBOIUL PSIHOLOGIC PE FACEBOOK: GRILE DE VERIFICARE A CONTURILOR FALSE

Cristian HAIDĂU*

Abstract:

Astroturfing is a sophisticated form of psychological warfare and informational manipulation, where coordinated networks of fake accounts create the illusion of genuine public support for specific political, economic, or social narratives. In the context of hybrid warfare, this strategy has become a crucial tool used by both state and non-state actors to influence public perceptions, destabilize societies, and undermine democratic processes. Its impact on national security is significant, as it erodes a population's ability to distinguish between real and fabricated information, thereby reducing resilience to manipulation and propaganda.

This study examines the impact of astroturfing on public opinion, identifies the psychological mechanisms exploited in such campaigns, and proposes a methodological framework for detecting suspicious accounts involved in these activities. Two identification grids for fake Facebook accounts have been developed: an extensive grid with 34 criteria for in-depth analysis and a simplified grid with 10 key indicators, designed for quick and accessible use by regular users. These grids are based on observable behavioral patterns and technical characteristics that may indicate potential manipulative activity. The application of these criteria to a set of suspicious accounts has validated the relevance of each indicator within the current digital ecosystem.

The paper highlights the crucial role of media literacy and appropriate regulations in combating this phenomenon but places particular emphasis on critical thinking as the primary means of resilience against manipulation. In an era where information warfare plays a central role in influence strategies, the ability of individuals to critically analyze and evaluate the information they consume is a key factor in national security and societal stability.

Keywords: *astroturfing, psychological warfare, national security, informational manipulation, fake accounts, social networks, disinformation*

Context şi importanţa subiectului

Fenomenul astroturfing-ului se manifestă într-un context socio-politic complex, marcat de creşterea influenţei tehnologiei şi

* Student doctorand, anul IV, Şcoala Doctorală, SNSPA, Bucureşti, domeniu de cercetare - ştiinţele comunicării, dezinformare şi manipulare în mediul digital, e-mail: Cristian.Haidau.21@drd.snspa.ro

a platformelor digitale asupra opiniei publice. Această metodă de manipulare, utilizată de entități statale și non-statale, exploatează emoțiile colective pentru a genera impresia unui sprijin autentic față de anumite idei, doctrine sau actori politici. România se află într-un context geopolitic extrem de tensionat, fiind aproape de prima linie a conflictului necombativ cinetic dintre Rusia și Ucraina. Această poziție strategică ne expune unor atacuri concertate de psihologie socială, menite să influențeze opinia publică și să servească intereselor Rusiei.

În prezent, informația a devenit o veritabilă armă de război (Roman 2024), utilizată în cadrul conflictelor moderne nu doar pentru propagandă, ci și pentru modelarea percepțiilor colective și influențarea deciziilor politice. Odată cu dezvoltarea rețelelor sociale și a comunicării ultra-rapide și în masă, războiul informațional a devenit un instrument strategic esențial, fiind exploatat intens de actorii ostili pentru a slăbi democrațiile din interior, folosindu-se de vulnerabilitățile lor intrinseci, precum libertatea de exprimare și diversitatea de opinii. Manipularea discursului public prin dezinformare și polarizare face ca societățile democratice să fie mai ușor de influențat, creând confuzie, neîncredere și diviziune.

Mai mult decât atât, succesul unei campanii de dezinformare nu depinde doar de strategia celor care o inițiază, ci și de predispoziția receptorului de a fi dezinformat. Oamenii tind să accepte mai ușor informațiile care le confirmă convingerile și să respingă cele care contravin sistemului lor de valori. Această tendință naturală de a filtra realitatea prin prisma propriilor credințe face ca dezinformarea să fie nu doar posibilă, ci și extrem de eficientă. În esență, un individ expus unui mesaj manipulator este vulnerabil în primul rând pentru că, fie conștient, fie inconștient, dorește să audă și să creadă acel mesaj (Stan 2021, 194).

Astroturfing-ul influențează percepția publică asupra realității, facilitând amplificarea artificială a unor mesaje coordonate strategic pe rețelele sociale. Această practică erodează fundamentul proceselor democratice prin alterarea percepțiilor și fragmentarea spațiului public. Ea contribuie la polarizarea societății, amplificând tensiunile existente și reducând coeziunea socială (Ghiurgiu 2024).

Astroturfing-ul reprezintă o vulnerabilitate strategică în contextul atacurilor hibride, fiind utilizat ca instrument de destabilizare în cadrul operațiunilor de influență desfășurate de actori ostili. În acest context, România este ținta unor operațiuni sofisticate de dezinformare, al căror scop este subminarea stabilității interne și promovarea narațiunilor care avantajează interesele geopolitice ale Rusiei. Aceste atacuri vizează atât

segmentul politic şi electoral, cât şi încrederea populaţiei în instituţiile statului, utilizând tehnici de propagandă şi manipulare psihologică fără precedent. Aceste campanii urmăresc influenţarea percepţiilor colective şi a deciziilor politice prin diseminarea dezinformării, subminarea încrederii în instituţiile statului şi promovarea unui climat de incertitudine, precum şi amplificarea diviziunilor sociale şi radicalizarea opiniilor prin utilizarea conţinutului provocator şi polarizant.

Impactul asupra securităţii naţionale este major, deoarece manipularea opiniei publice poate influenţa direct procesele decizionale critice, inclusiv cele din domeniul electoral, economic şi de apărare. Un exemplu concret este anularea alegerilor din România, o măsură extraordinară adoptată pentru protejarea procesului democratic în faţa unui atac de psihologie socială fără precedent, care ameninţa stabilitatea electorală şi funcţionarea instituţiilor statului. Această decizie a evidenţiat necesitatea unor mecanisme de apărare eficiente împotriva manipulării informaţionale şi a demonstrat vulnerabilitatea proceselor electorale în faţa unor operaţiuni coordonate de influenţă externă.

Scopul şi obiectivele acestui articol

Lucrarea de faţă îşi propune să ofere o înţelegere actualizată a fenomenului de astroturfing, punând accent pe instrumentele prin care utilizatorii obişnuiţi de Facebook pot identifica rapid şi cu o acurateţe rezonabilă conturile suspecte. În acest sens, articolul introduce două grile de analiză adaptate pentru nevoi diferite. Prima este o grilă detaliată, alcătuită din 34 de criterii, destinată unei analize amănunţite a conturilor suspecte, incluzând factori ce ţin de identitate, activitate, reţea socială şi tipare de interacţiune iar cea de-a doua este o grilă simplificată, cu doar 10 criterii esenţiale, concepută pentru a fi utilizată rapid şi intuitiv de către orice utilizator, fără a necesita un proces îndelungat de analiză sau cunoştinţe tehnice.

În 2025, state, organizaţii şi grupuri de interese recurg la tactici avansate de influenţare digitală, utilizând algoritmi de amplificare, conturi automatizate gestionate prin inteligenţă artificială şi conţinut generat automat pentru a crea iluzia unui sprijin popular autentic faţă de anumite idei, doctrine sau personalităţi publice. Aceste metode sofisticate de manipulare afectează capacitatea indivizilor de a distinge între opinia reală şi cea fabricată, având implicaţii majore asupra proceselor democratice şi a climatului social global. În faţa unei opinii prezentate ca fiind a „majorităţii", punctele de vedere individuale, chiar

dacă sunt corecte, își pot pierde din entuziasm și vizibilitate, fenomen descris în literatura de specialitate ca efectul conformismului sub presiunea socială (Dobrescu and Bârgăoanu 2003, 25).

Într-un peisaj digital tot mai influențat de inteligența artificială și de relaxarea regulilor de moderare pe marile platforme sociale, acest articol are ca scop informarea și educarea publicului despre mecanismele actuale ale astroturfing-ului și despre modul în care acesta influențează percepția socială. Obiectivul principal este de a crește nivelul de conștientizare cu privire la manipularea informațională, oferind utilizatorilor metode clare și accesibile pentru a identifica conturile false și activitățile coordonate de dezinformare.

În acest context, articolul analizează principalele resorturi psihologice exploatate în astroturfing și impactul acestora asupra utilizatorilor, alături de metodele tehnologice moderne utilizate pentru a amplifica mesajele artificiale. De asemenea, sunt abordate implicațiile inteligenței artificiale în generarea și gestionarea automată a conținutului, care complică și mai mult eforturile de combatere a dezinformării. Un element esențial al analizei este reprezentat de grila extinsă de identificare a conturilor suspecte, care detaliază fiecare criteriu și relevanța sa în actualul ecosistem digital. În completare, grila simplificată oferă o metodă rapidă prin care utilizatorii pot evalua autenticitatea unui cont de Facebook, fără a fi necesară o verificare complexă.

Prin această lucrare, se urmărește oferirea cititorilor a unor instrumente practice și actualizate pentru recunoașterea și combaterea manipulării informaționale. Totodată, se dorește încurajarea unei atitudini mai critice și mai analitice în evaluarea informațiilor din mediul online.

Definirea și explicarea fenomenului de astroturfing

Astroturfing-ul reprezintă o tactică sofisticată de manipulare informațională, utilizată de grupuri de interese sau entități cu agende ascunse pentru a influența opinia publică cu privire la un subiect, o știre, un personaj sau o inițiativă, fie prin crearea iluziei unui sprijin popular autentic, fie prin discreditarea și denigrarea adversarilor (Merriam-Webster 2023). Deși pare o reacție spontană a cetățenilor, acest fenomen este, în realitate, rezultatul unei strategii bine orchestrate, menite să modifice percepțiile colective și să creeze artificial un curent de opinie favorabil sau defavorabil unei cauze specifice. Astroturfing-ul nu este

folosit doar pentru a crea artificial sprijin în jurul unui mesaj sau al unei cauze, ci şi ca instrument de subminare a credibilităţii unei persoane, instituţii sau idei. Prin răspândirea sistematică a îndoielii şi amplificarea atacurilor orchestrate, aceste campanii reuşesc să erodeze încrederea publicului şi să compromită percepţia asupra legitimităţii unei figuri publice sau a unei iniţiative. Neincrederea este profund contagioasă (Georgescu 2024) şi, odată inoculată într-un grup social, aceasta se propagă rapid, afectând percepţiile colective şi generând o atmosferă de scepticism şi ostilitate. Astfel, astroturfing-ul nu doar modelează sprijinul artificial, ci şi slăbeşte în mod deliberat coeziunea socială prin inducerea neîncrederii generalizate. Acesta se manifestă atât online, prin comentarii, reacţii, distribuiri artificiale, generare de conţinut manipulator şi utilizarea trolilor şi boţilor digitali, cât şi offline, prin organizarea de proteste sau evenimente cu participanţi care fie sunt manipulaţi să participe pe baza unor informaţii înşelătoare, fie sunt recrutaţi şi motivaţi cu anumite beneficii. În unele cazuri, participanţilor li se prezintă un motiv iniţial de protest, care ulterior este deturnat către o altă agendă, fără ca aceştia să fie conştienţi de adevăratul scop al mobilizării.

Termenul "astroturfing" a fost introdus de senatorul texan Lloyd Bentsen în 1985, când a remarcat că o campanie intensă de scrisori adresate Congresului, aparent provenind de la cetăţeni preocupaţi de anumite schimbări legislative, era, de fapt, iniţiată de companii din industria asigurărilor. Denumirea provine de la "AstroTurf", un tip de gazon artificial, ilustrând astfel caracterul prefabricat al acestui tip de influenţă publică (CSMonitor 2009). Fenomenul a evoluat rapid, devenind o componentă cheie a operaţiunilor psihologice moderne şi fiind exploatat atât de actori politici, cât şi de corporaţii sau alte entităţi cu interese specifice.

Prin utilizarea unor metode din ce în ce mai avansate, inclusiv manipularea algoritmilor pentru amplificarea mesajelor pe platformele digitale, crearea automată şi gestionarea conturilor false cu inteligenţa artificială, precum şi utilizarea acesteia pentru generarea de conţinut şi coordonarea activităţilor acestora, astroturfing-ul devine un instrument extrem de eficient şi destul de ieftin în manipularea percepţiilor publice, fiind utilizat tot mai frecvent în conflictele asimetrice şi în cadrul războiului hibrid. Acesta permite actorilor non-statali sau statali să destabilizeze societăţi şi să influenţeze procesele decizionale fără a recurge la mijloace convenţionale de confruntare, folosind manipularea informaţională ca armă strategică. Cercetările arată că informaţiile false

din domeniul politic au un potenţial de viralizare de trei ori mai ridicat decât cele din alte domenii (Bârgăoanu 2018, 146). Această capacitate de diseminare rapidă permite manipularea opiniilor şi crearea unui climat social polarizat, reducând spaţiul pentru dezbaterea raţională.

Astroturfing-ul reprezintă o ameninţare majoră pentru procesele democratice, având potenţialul de a submina încrederea în instituţiile publice şi de a distorsiona mecanismele decizionale. Prin crearea unei percepţii false asupra sprijinului sau opoziţiei faţă de anumite politici sau lideri, acest fenomen poate influenţa alegerile, formularea politicilor publice şi dinamica socială generală.

Implicaţii etice şi legale ale astroturfing-ului

Astroturfing-ul pe Facebook și alte platforme sociale continuă să ridice preocupări etice şi legale, având un impact semnificativ asupra integrităţii discursului public şi a proceselor democratice. În ultimii doi ani, avansul tehnologiilor de inteligenţă artificială şi noile reglementări internaţionale au influenţat radical atât modul în care astroturfing-ul este utilizat, cât şi strategiile de combatere a acestuia.

Implicaţiile etice ale astroturfing-ului sunt multiple şi afectează direct modul în care opinia publică este formată şi influenţată.

Manipularea opiniei publice reprezintă una dintre principalele consecinţe, deoarece astroturfing-ul creează o aparenţă falsă de sprijin popular. Acest fenomen subminează autonomia şi libertatea de exprimare a individului, împiedicând o dezbatere autentică bazată pe fapte. În 2025, utilizarea AI-urilor generative pentru crearea de conturi false şi mesaje hiper-personalizate a amplificat problema, făcând manipularea şi mai greu de detectat sau combătut.

Înşelarea utilizatorilor este o altă problemă majoră, întrucât astroturfing-ul implică folosirea de conturi false, boţi avansaţi şi reţele coordonate pentru a induce iluzia unui sprijin sau a unei opoziţii autentice. Această practică contravine principiului onestităţii şi transparenţei, inducând în eroare publicul cu privire la sursa şi motivaţiile reale ale mesajelor politice. Noii algoritmi de detecţie dezvoltaţi de platformele digitale în 2024 încearcă să limiteze fenomenul, însă eficienţa acestora rămâne sub semnul întrebării.

Dezinformarea şi distorsionarea informaţiei completează tabloul implicaţiilor etice. Astroturfing-ul utilizează tehnici de manipulare a faptelor pentru a promova agende politice sau economice, iar în 2025 AI permite generarea automată de conţinut falsificat, sub forma unor

articole, comentarii şi videoclipuri deepfake. Acest fenomen contravine principiului acurateţei şi corectitudinii informaţiei, afectând capacitatea publicului de a lua decizii informate (Chan 2022). Reglementările Uniunii Europene privind dezinformarea digitală, adoptate în 2024, încearcă să combată această practică, însă aplicarea lor este încă în curs de perfecţionare.

Implicaţiile legale ale astroturfing-ului sunt complexe şi reflectă atât evoluţiile tehnologice recente, cât şi schimbările de reglementare din mediul digital.

Odată cu schimbările politice din 2024 şi reconfigurarea administraţiei americane, s-a conturat o nouă paradigmă legislativă, care privilegiază libertatea de exprimare în detrimentul moderării stricte a conţinutului. În acest context, intervenţia platformelor asupra discursului public, anterior criticată ca posibilă formă de cenzură partizană, a fost descurajată. Ca urmare, marile companii de tehnologie şi-au ajustat strategiile, reducând moderarea umană şi colaborarea cu organizaţiile de fact-checking, în favoarea unei abordări mai permisive, aliniate noilor orientări de reglementare digitală. Deşi tehnologiile AI continuă să fie utilizate pentru identificarea reţelelor de dezinformare, lipsa verificării umane a dus la o creştere a eficienţei campaniilor de manipulare, inclusiv a fenomenului de astroturfing (Meta 2025).

Legislaţia privind protecţia datelor personale reprezintă un alt domeniu afectat de astroturfing, întrucât această practică poate implica colectarea şi utilizarea ilegală a datelor personale ale utilizatorilor, încălcând regulamentele internaţionale în vigoare. Conform noii Directive UE privind confidenţialitatea digitală (2024), utilizarea AI pentru extragerea şi profilarea automată a utilizatorilor în scopuri de manipulare politică sau comercială este interzisă. Sancţiunile au fost înăsprite, incluzând amenzi de până la 4% din cifra de afaceri anuală a companiilor implicate („GDPR" 2025).

Legislaţia electorală şi finanţarea politică au fost, de asemenea, afectate de fenomenul astroturfing-ului, care a devenit o problemă majoră în campaniile electorale. Utilizarea reţelelor false pentru influenţarea alegerilor a fost interzisă prin noile reglementări electorale din SUA şi UE, care prevăd monitorizarea şi sancţionarea entităţilor politice ce beneficiază de astfel de tactici. Noile legi impun şi transparenţă în publicitatea politică digitală, inclusiv dezvăluirea surselor de finanţare şi a metodelor de ţintire. Cu toate acestea, relaxarea reglementărilor privind moderarea conţinutului pe marile platforme sociale a generat un mediu în care entităţile politice şi economice

pot desfășura campanii de influenţă fără un control strict asupra autenticităţii şi veridicităţii informaţiilor distribuite (Autoritatea Electorală Permanentă 2024)

Astroturfing-ul rămâne o ameninţare serioasă la adresa democraţiei, iar evoluţiile tehnologice recente au complicat şi mai mult peisajul manipulării informaţionale. Creşterea utilizării AI şi noile reglementări fac ca detectarea şi prevenirea acestui fenomen să fie o provocare constantă pentru guverne, companii de tehnologie şi societatea civilă. În acelaşi timp, un nou trend promovat de figuri proeminente precum Donald Trump, Elon Musk şi Mark Zuckerberg susţine liberalizarea exprimării pe reţelele sociale, renunţarea la angajaţii responsabili cu moderarea conţinutului şi reducerea colaborării cu organizaţiile de fact-checking. Această tendinţă ridică întrebări legate de vulnerabilitatea crescută a platformelor la campanii de dezinformare şi manipulare, inclusiv la practici precum astroturfing-ul.

Principalele resorturi psihologice exploatate de astroturfing

Principiul mesianismului reprezintă una dintre strategiile eficiente utilizate pentru a influenţa şi manipula masele, bazându-se pe ideea că un lider carismatic poate deveni un mesia politic, un salvator al poporului, capabil să rezolve toate problemele şi să îndeplinească aspiraţiile acestuia (Bichir 2020). Acest principiu nu se limitează doar la un personaj uman, ci poate fi aplicat şi unei ţări, unei doctrine sau unei idei sociale ori politice.

Prin exploatarea acestui principiu, actorii implicaţi în manipulare şi modelare psihologică, inclusiv cei care utilizează tactici de astroturfing, urmăresc să obţină sprijinul şi loialitatea maselor, inducând ideea că liderul lor este un conducător providenţial. Strategia se bazează pe discursuri carismatice, promisiuni grandioase şi manipulare emoţională, având ca scop crearea unei iluzii a puterii şi speranţei în rândul populaţiei. De-a lungul istoriei, lideri precum Hitler, Mussolini, Stalin, Mao Zedong, Perón, Chávez şi Kim Il Sung au folosit acest principiu pentru a mobiliza masele, construindu-şi culturi ale personalităţii şi exploatând vulnerabilităţile sociale. Prin propagandă, retorică populistă şi demonizarea adversarilor, aceştia au obţinut loialitate necondiţionată şi control asupra opiniei publice.

Manipularea prin acest principiu este amplificată de reţelele sociale şi inteligenţa artificială, care permit distribuirea rapidă şi

automatizată a conţinutului propagandistic. Aceste tehnologii creează o imagine idealizată a liderului şi a agendei sale politice, influenţând percepţia publicului. Tendinţa naturală a maselor de a căuta un lider puternic şi carismatic le face susceptibile la astfel de tactici, facilitând instaurarea unui cult al personalităţii şi consolidarea controlului asupra opiniei publice. Acest proces asigură un nivel ridicat de influenţă asupra maselor şi permite exercitarea unui control sporit asupra acestora.

Un alt element esenţial în aplicarea principiului mesianic este identificarea unui inamic comun, perceput ca fiind sursa tuturor problemelor sociale şi economice. Manipulatorii folosesc această strategie pentru a demoniza adversarii politici şi a le atribui responsabilitatea pentru dificultăţile resimţite de populaţie. Crearea unei opoziţii binare între liderul providenţial şi adversarii săi întăreşte loialitatea maselor şi justifică măsuri radicale în numele binelui colectiv (Sturza 2021). Această dihotomie facilitează acceptarea unor acţiuni extreme şi contribuie la consolidarea poziţiei liderului, care devine singura alternativă percepută ca viabilă.

Principiul mesianic este astfel un instrument puternic de manipulare, utilizat pentru a controla opinia publică şi a obţine sprijinul necondiţionat al maselor. În contextul digital actual, reţelele sociale şi inteligenţa artificială facilitează diseminarea rapidă a acestui tip de influenţă, permiţând coordonarea eficientă a campaniilor de astroturfing şi consolidarea unui cadru favorabil războiului psihologic prin inducerea unor percepţii colective controlate artificial. Aplicarea sa în cadrul strategiilor de influenţă informaţională subminează gândirea critică şi reduce capacitatea indivizilor de a analiza obiectiv realitatea politică şi socială. Într-un mediu digital dominat de reţele sociale şi tehnici avansate de comunicare, acest principiu rămâne unul dintre cele mai eficiente mecanisme de mobilizare şi control al opiniei publice.

Principiul maniheismului şi principiul mesianismului sunt două concepte interconectate, utilizate frecvent în strategiile de influenţă şi manipulare a opiniei publice. Maniheismul reprezintă o strategie bazată pe dualitate, în care lumea este împărţită strict între bine şi rău, eliminând orice nuanţe intermediare. Această perspectivă dihotomică creează o delimitare rigidă între grupuri, prezentând unii actori ca fiind moralmente superiori, în timp ce adversarii sunt demonizaţi şi consideraţi principala sursă a problemelor societăţii.

Această diviziune artificială favorizează consolidarea unei imagini mesianice în jurul liderului, care este promovat ca singura soluţie

viabilă pentru redresarea naţiunii sau a grupului vizat. În astfel de cadre ideologice, greşelile liderului sau ale susţinătorilor săi sunt trecute cu vederea sau justificate, în timp ce orice opoziţie este aspru condamnată şi deseori prezentată ca o ameninţare existenţială. Acest tip de narativ rigid restrânge semnificativ capacitatea societăţii de a analiza critic situațiile complexe, reducând dezbaterea democratică și înlocuind-o cu loialitatea necondiționată față de figura centrală (Volkoff 2009, 126–27).

Impactul acestei manipulări prin maniheism este profund, având implicaţii asupra coeziunii sociale şi asupra mecanismelor democratice. Prin eliminarea nuanţelor şi impunerea unei perspective radicale, discursul critic şi dezbaterea raţională sunt marginalizate, conducând la polarizare, excludere şi, în unele cazuri, la radicalizare extremă. Această strategie este adesea utilizată în combinaţie cu tehnici moderne de propagandă, cum ar fi astroturfing-ul şi exploatarea algoritmilor şi grupurilor/camerelor de rezonanţă ale platformelor sociale, facilitând amplificarea mesajelor maniheiste şi crearea unor ecouri informaţionale menite să întărească percepţiile preexistente. Într-un context dominat de reţelele sociale şi de tehnologiile avansate de inteligenţă artificială, impactul maniheismului asupra societăţii contemporane devine o provocare majoră pentru securitatea informaţională şi stabilitatea politică.

Principiul validării sociale joacă un rol fundamental în tactici precum astroturfing-ul politic, unde opinia şi comportamentul sunt manipulate pentru a crea iluzia unei susţineri larg răspândite. Oamenii sunt predispuşi să fie influenţaţi spontan de cei din jur, iar pe măsură ce numărul acestora creşte, tendinţa de a accepta rapid şi fără o analiză critică opinia majorităţii devine mai pronunţată. Acest fenomen se explică prin dorința de a economisi timp și resurse cognitive, evitând procesul laborios al unei evaluări individuale (Chelcea 2006, 247–49).

Astroturferii exploatează această tendinţă prin utilizarea unui număr mare, dar artificial de susţinători, aprecieri şi redistribuiri pentru a sugera popularitatea şi legitimitatea unei anumite agende politice. Prin crearea unei impresii puternice de validare socială, aceştia influenţează utilizatorii să adopte şi să sprijine o anumită idee, bazându-se pe premisa că, dacă mulţi o susţin, aceasta trebuie să fie analizată corespunzător, verificată şi în consecinţă corectă şi/sau adevărată (Oprea 2021, 118).

Manipularea prin astroturfing se bazează astfel pe încrederea excesivă a indivizilor în opiniile celorlalţi, exploatând principiul conformismului şi al validării sociale pentru a distorsiona percepţia

publicului asupra unei cauze sau a unui lider. În era digitală, această strategie este amplificată de algoritmii reţelelor sociale, care favorizează conţinutul cu un nivel ridicat de angajament, indiferent de autenticitatea sa, contribuind astfel la consolidarea unor naraţiuni artificiale.

Principiul autorităţii joacă un rol esenţial în mecanismele de influenţă socială, fiind exploatat frecvent în cadrul strategiilor de astroturfing. Oamenii sunt educaţi să manifeste respect şi supunere faţă de autorităţi, ceea ce le poate diminua tendinţa de a pune sub semnul întrebării informaţiile primite din surse aparent credibile. Astroturferii utilizează această predispoziţie prin crearea de conturi false care par a fi deţinute de experţi sau persoane cu autoritate în domeniu (Chelcea 2006, 253–56).

Printre cele mai frecvent simulate identităţi se numără cadrele militare, medicale şi reprezentanţii bisericii. Aceste personaje sunt alese strategic pentru a spori credibilitatea mesajului şi a reduce scepticismul publicului. Prin prezentarea mesajelor drept fundamentate pe expertiză, cunoştinţe solide sau principii etice incontestabile, manipulatorii reuşesc să câştige încrederea utilizatorilor şi să îi influenţeze în direcţia dorită. Această tehnică este deosebit de eficientă în contextul reţelelor sociale, unde utilizatorii sunt bombardaţi cu informaţii şi rareori dispun de timpul sau resursele necesare pentru a verifica autenticitatea surselor.

În era digitală, inteligenţa artificială şi tehnologiile de deepfake facilitează şi mai mult utilizarea acestui principiu, permiţând generarea de conţinut vizual şi auditiv extrem de convingător. Astfel, fenomenul de astroturfing devine din ce în ce mai sofisticat, exploatând încrederea în autoritate pentru a influenţa percepţiile publice şi a modela opiniile colective.

Principiul fricii şi ameninţării reprezintă o tactică de manipulare eficientă, utilizată frecvent în strategiile de astroturfing pentru a influenţa percepţiile şi comportamentele publicului. Această tehnică se bazează pe exploatarea temerilor legate de aspecte precum securitatea naţională, imigraţia, terorismul, infracţionalitatea, problemele economice şi alte incertitudini, amplificând anxietăţile colective pentru a genera reacţii emoţionale puternice. Prin sublinierea pericolelor şi a consecinţelor negative asociate cu o anumită opţiune, manipulatorii urmăresc să inducă o stare de neliniște şi să stimuleze mecanismele de apărare şi conservare.

Astroturferii folosesc tehnici de amplificare a incertitudinii şi de exagerare a riscurilor pentru a canaliza atenţia publicului către soluţiile

pe care le promovează. Prin crearea unei imagini alarmiste asupra unei alternative, aceştia încearcă să direcţioneze susţinerea şi acceptarea către propria agendă politică sau ideologică. Manipularea emoțională bazată pe frică devine astfel un instrument esenţial pentru influenţarea maselor, determinând indivizii să acţioneze nu pe baza unei analize raţionale, ci ca reacţie la o ameninţare percepută (André 2004).

Această strategie este deosebit de eficientă în era digitală, unde reţelele sociale şi algoritmii platformelor online favorizează conţinutul care generează reacţii emoţionale intense. Astfel, frica şi ameninţarea devin factori determinanţi în modelarea opiniei publice, permiţând celor care folosesc aceste tactici să-şi impună naraţiunile şi să slăbească sprijinul pentru alternativele concurente sau să distragă atenţia de la subiecte incomode pentru propriile interese.

Principiul coeziunii sociale se referă la tendinţa oamenilor de a forma legături puternice şi de a se identifica cu grupurile sociale, generând solidaritate şi susţinere reciprocă (Roşca 2019). Acest principiu este exploatat în strategiile de astroturfing, unde manipulatorii creează comunităţi false (echo-chambers) pe platforme precum Facebook, aparent în sprijinul unei anumite cauze sau candidaţi (Bârgăoanu 2018, 120). Scopul este de a cultiva un sentiment puternic de apartenenţă şi loialitate faţă de aceste grupuri, influenţând utilizatorii să adere şi să susţină cauza promovată, având în vedere puterea şi influenţa pe care grupul le poate exercita asupra membrilor săi.

În paralel, manipulatorii pot utiliza teoria identităţii sociale pentru a cultiva sentimente de separare şi antagonism faţă de alte grupuri sociale, alimentând astfel conflicte şi diviziuni. Identitatea socială este un factor determinant în formarea opiniilor şi comportamentelor, întrucât indivizii îşi construiesc percepţia de sine în raport cu grupurile de apartenenţă (Boncu 2014). Astroturferii exploatează această tendinţă pentru a crea şi amplifica tensiuni, diminuând solidaritatea şi sprijinul faţă de grupurile sau cauzele opuse.

Acest principiu este strâns legat de **teoria conformismului social**, care evidenţiază tendinţa oamenilor de a se conforma normelor şi aşteptărilor sociale pentru a evita conflictele şi a obţine aprobarea celorlalţi (Lazarsfeld, Berelson şi Gaudet 2004, 205–6). De asemenea, **teoria autoclasificării** subliniază că indivizii tind să se categorizeze în anumite grupuri sociale şi să îşi construiască identitatea personală prin asociere cu acestea (Turner and Reynolds 2012).

Aceste teorii se completează şi oferă un cadru explicativ asupra modului în care identitatea socială poate fi influenţată şi manipulată pentru promovarea anumitor agende. Prin astroturfing, aceste efecte sunt amplificate, facilitând controlul naraţiunilor dominante şi slăbirea sprijinului pentru grupurile sau cauzele considerate opozante. Astfel, coeziunea socială devine un instrument puternic în strategii de influenţă informaţională, având implicaţii semnificative asupra comportamentului colectiv şi asupra stabilităţii sociale.

Metodologia de cercetare şi validarea instrumentelor propuse

Procesul de elaborare şi validare a grilelor de identificare a conturilor false s-a desfăşurat pe parcursul a aproximativ două luni, in intervalul ianuarie–februarie 2025. Demersul a urmat o abordare inductivă, construită treptat pe baza observaţiei directe şi a corelării tiparelor recurente identificate în mediul online. Punctul de plecare metodologic l-a constituit cadrul teoretic propus de studiul (McAfee 2022), care defineşte şase criterii de bază pentru detectarea profilurilor neautentice. Aceste repere au fost utilizate ca filtru iniţial, cercetarea având ca obiectiv adaptarea şi extinderea lor pentru a surprinde mai adecvat complexitatea actuală a fenomenului astroturfing.

În acest context, conturile false sau deturnate sunt instrumente deliberate, folosite fie pentru beneficii economice (inclusiv prin activităţi infracţionale), fie pentru campanii de influenţă precum astroturfing-ul. Ele contribuie la crearea artificială a validităţii sociale, la amplificarea mesajelor şi la polarizarea discursului public, inclusiv prin inducerea fricii, urii şi neîncrederii (Catrina 2024). Din această perspectivă, identificarea acestor conturi devine o etapă necesară pentru înţelegerea modului în care sunt construite şi propagate naraţiunile manipulative în spaţiul public digital.

În prima etapă a studiului, criteriile McAfee au fost utilizate pentru selectarea unui eşantion relevant de analiză. Procesul a presupus monitorizarea a patru pagini de Facebook active în spaţiul digital românesc, alese pe baza vizibilităţii ridicate şi a tendinţei de a disemina naraţiuni populiste şi polarizante. Din interacţiunile generate în jurul acestor pagini a fost extras un lot de 20 de conturi care îndeplineau cel puţin patru dintre cele şase criterii de bază. Acest prag a fost stabilit pentru a reduce riscul includerii accidentale a unor utilizatori legitimi, orientând analiza către profiluri cu o probabilitate mai ridicată de neautenticitate şi cu o prezenţă constantă în amplificarea mesajelor respective.

Etapa următoare a constat într-o examinare detaliată, realizată manual, a celor 20 de conturi selectate, vizând istoricul activității, structura rețelei de conexiuni și elementele tehnice vizibile. Această analiză a permis identificarea unor tipare comportamentale suplimentare, care nu erau acoperite de modelul inițial. Prin sistematizarea acestor observații, lista inițială de indicatori a fost extinsă la un total de 34 de criterii care stau la baza grilei extinse de analiză propusă în această lucrare.

Pornind de la grila extinsă, a fost elaborată ulterior și o versiune simplificată, limitată la 10 criterii esențiale. Selecția acestora a avut în vedere două principii: frecvența apariției în cadrul eșantionului analizat și gradul de accesibilitate pentru un utilizator normal. În timp ce frecvența a putut fi evaluată în mod obiectiv, criteriul accesibilității a fost stabilit printr-o apreciere subiectivă și asumată, din perspectiva unui utilizator obișnuit care are nevoie de o evaluare rapidă. Această opțiune metodologică reflectă intenția de a oferi un instrument practic, utilizabil în afara mediului academic, fără a impune o rigoare cantitativă excesivă într-un demers cu finalitate predominant educațională.

Pentru a verifica funcționarea grilelor, acestea au fost aplicate comparativ pe două eșantioane distincte. Primul a inclus cele 20 de conturi suspecte inițiale, reevaluate prin prisma grilei extinse pentru a observa consistența rezultatelor, iar al doilea eșantion, utilizat cu rol orientativ de control, a fost alcătuit din 20 de conturi autentice, selectate din cercul extins de cunoscuți, urmărindu-se o diversitate rezonabilă din punctul de vedere al vârstei, sexului, nivelului de educație și domiciliului. Această etapă a avut rolul de a observa potențialul de rezultate fals-pozitive și de a evalua măsura în care grilele pot diferenția între comportamente artificiale și activitatea obișnuită a utilizatorilor reali.

Din perspectiva resurselor implicate, testarea a evidențiat diferențe clare de timp între cele două instrumente. Aplicarea grilei extinse a necesitat, între 15 și 30 de minute per cont, în funcție de volumul și complexitatea activității analizate, în timp ce grila simplificată a permis o evaluare mai rapidă, cuprinsă între 3 și 7 minute. Această diferență susține utilizarea complementară a celor două instrumente: grila extinsă este adecvată pentru analiză aprofundată și cercetare, iar varianta restrânsă răspunde nevoii de orientare rapidă și igienă informațională cotidiană. Deși metodologia are limitele inerente unei abordări calitative, rezultatele oferă un cadru operațional coerent, adaptat realităților ecosistemului digital din anul 2025.

Tabelul nr. 1: Grilă extinsă de analiză a conturilor suspecte de Facebook.

	Grilă extinsă de analiză a conturilor suspecte de Facebook	
Nr.	**Criteriu**	**Bifat**
1.	**Nume atipic** - Combinarea de nume din culturi diferite sau alăturarea a două nume de familie prin generare aleatorie nepotrivită.	
2.	**Nume schimbat recent** - Link-ul contului nu corespunde cu numele actual, indicând posibilitatea unei reutilizări.	
3.	**Fără poză de profil/cover** - Lipsa unei imagini de profil poate indica un cont artificial.	
4.	**Poză de profil generică, generată cu AI, furată de pe internet sau de calitate foarte proastă** - Astfel, conturile false evită identificarea, mascându-şi identitatea reală şi creând o aparenţă de autenticitate fără a expune informaţii personale.	
5.	**Schimbări frecvente de poza de profil** - Indică încercări de a evita detectarea automată a contului fals.	
6.	**Cont fără informaţii personale –** Generarea automată de conturi false se îngreunează cu fiecare detaliu în plus şi din acest motiv de multe ori se alege cantitatea în detrimentul calităţii.	
7.	**Cont nou (<6 luni) sau schimbări bruşte de identitate**. - Conturile false sunt adesea create rapid pentru scopuri specifice.	
8.	**Cont fără activitate/conţinut** - Un cont real are postări şi interacţiuni, în timp ce unul fals poate fi gol sau cu puţine informaţii.	
9.	**Link cont nepersonalizat -** Link generat automat pentru a putea reutiliza mai uşor contul în alte campanii.	
10.	**Frecvenţă neregulată a postărilor** - Alternanţă între perioade de inactivitate şi activitate intensă, specifică conturilor automate.	
11.	**Activitate anormală** - Volum mare de postări fără interacţiuni reale. Poate indica un cont automatizat care distribuie conţinut.	
12.	**Postări 24/24** - Activitate constantă, inclusiv în intervale neobişnuite pentru fusul orar. Un utilizator real are pauze naturale de activitate.	
13.	**Postări fără descriere -** Distribuirea de conţinut fără explicaţii poate fi un semnal de automatizare.	
14.	**Trecerea bruscă de la conţinut personal la activitate suspectă -** Un cont furat încetează brusc să mai posteze conţinut personal şi începe să distribuie materiale specifice astroturfing-ului sau manipulării, indicând o posibilă preluare şi reutilizare.	
15.	**Descrieri scrise greşit/incoerent -** Textele pot fi traduse automat sau generate cu inteligenţa artificială.	

16.	**Conținut intens pe o temă -** Conturile false sunt adesea concentrate doar pe un subiect (ex: politică, conspirații etc).	
17.	**Conținut preponderent instigator -** Mesaje care promovează ură, divizare socială sau polarizare politică.	
18.	**Raport ciudat între numărul de prieteni și cel de interacțiuni -** Un cont cu mii de prieteni, dar fără interacțiuni, poate fi suspect.	
19.	**Prieteni cu conturi suspecte -** Conexiuni cu alte conturi care prezintă aceleași caracteristici suspecte.	
20.	**Prieteni dispersați geografic -** Conturile false au adesea conexiuni internaționale fără logică aparentă.	
21.	**Toți sau majoritatea prietenilor sunt adăugați recent –** Acest aspect poate indică o rețea artificială construită rapid pentru propagandă.	
22.	**Adaugă prieteni doar dintr-o anumită nișă -** Conturile false se infiltrează adesea în comunități specifice.	
23.	**Cerc social incoerent** - Discrepanțe sociale și culturale între profil și prieteni.	
24.	**Număr mic de prieteni (<250) -** Conturile false au adesea foarte puțini prieteni, deoarece este dificil să creeze rapid o rețea autentică de conexiuni.	
25.	**Tipar de distribuire în masă**- Distribuire masivă de conținut, fără comentarii proprii.	
26.	**Lipsa reacțiilor la propriile postări -** Poate indica lipsa unei audiențe reale.	
27.	**Nu răspunde la mesaje private/comentarii -** Un cont real în general interacționează cu ceilalți.	
28.	**Comentarii generice, scurte și repetitive** - Ex: „Adevărat!", „Fake news!", „Toată lumea știe", „Minciuni! Propagandă!" etc.	
29.	**Distribuire excesivă de link-uri suspecte -** Link-uri către site-uri obscure sau de propagandă.	
30.	**Legături cu alte conturi suspecte -** Își primește like-urile și comentariile doar de la aceleași conturi.	
31.	**Postări codificate -** Utilizarea deliberată a simbolurilor sau greșelilor ortografice pentru a evita detectarea automatizată a platformei.	
32.	**Comentarii cu link-uri ascunse –** Plasarea strategică a link-urilor in partea de text care nu se vede, *„blind spotul"* („punctul mort") - „Vezi mai mult".	
33.	**Schimbare bruscă a limbii utilizate în conținut –** Posibil semn de cont furat prin strategii de phishing.	
34.	**Diferențe între locația profilului și istoricul activității**- Contul pretinde că este din România, dar toate check-in-urile, evenimentele sau postările vechi sunt din altă țară. Acest lucru poate indica un cont furat.	
	TOTAL criterii îndeplinite	

Tabelul nr. 2: Niveluri de alarmare pentru un cont suspect pe grila extinsă.

Niveluri de alarmare pentru un cont suspect pe grila extinsă	
0-5 criterii -» **Cont probabil autentic**	Contul nu prezintă semne evidente de activitate suspectă.
6-10 criterii -» **Cont puţin suspect**	Există semnale de alarmă, dar nu suficient de multe pentru a bănui serios că este un cont fals sau parte dintr-o reţea de manipulare. Poate fi un cont nou sau al unei persoane care nu prea foloseste platforma. Se interpretează şi în funcţie de ce criterii bifează.
11-17 criterii -» **Cont suspect**	Contul prezintă suficiente semne de activitate problematică. Există o probabilitate mare ca acesta să fie un cont fals, utilizat pentru manipulare, propagandă şi/sau astroturfing.
18+ criterii -» **Cont foarte suspect**	Contul bifează multe criterii pentru a fi considerat fals sau furat, parte a unei reţele de dezinformare sau un cont utilizat pentru propagandă. Se recomandă raportarea la Facebook.

Tabelul nr. 3: Grilă restrânsă de analiză a conturilor suspecte de Facebook.

	Grilă restrânsă de analiză a conturilor suspecte de Facebook	
Nr.	**Criteriu**	**Bifat**
1.	**Fără poză de profil/cover** - Lipsa unei imagini de profil poate indica un cont artificial.	
2.	**Poză de profil generică, generată cu AI, furată de pe internet sau de calitate foarte proastă** - Astfel, conturile false evită identificarea, mascându-şi identitatea reală şi creând o aparenţă de autenticitate fără a expune informaţii personale.	

3.	**Cont nou (<6 luni) sau schimbări bruște de identitate**. - Conturile false sunt adesea create rapid pentru scopuri specifice.	
4.	**Tipar de distribuire în masa -** Distribuire masivă de conținut, fără comentarii proprii.	
5.	**Comentarii generice, scurte și repetitive** - Ex: „Adevărat!", „Fake news!", „Toată lumea știe", „Minciuni! Propagandă!" etc.	
6.	**Prieteni cu conturi suspecte -** Conexiuni cu alte conturi care prezintă aceleași caracteristici suspecte.	
7.	**Număr mic de prieteni (<250) -** Conturile false au adesea foarte puțini prieteni, deoarece este dificil să creeze rapid o rețea autentică de conexiuni.	
8.	**Conținut intens pe o temă -** Conturile false sunt adesea concentrate doar pe un subiect (ex: politică, conspirații etc).	
9.	**Legături cu alte conturi suspecte -** Își primește like-urile și comentariile doar de la aceleași conturi.	
10.	**Postări codificate -** Utilizarea deliberată a simbolurilor sau greșelilor ortografice pentru a evita detectarea automatizată a platformei.	
	TOTAL criterii îndeplinite	

Tabelul nr. 4: Niveluri de alarmare pentru un cont suspect pe grila restrânsă

Niveluri de alarmare pentru un cont suspect pe grila restrânsă	
0-3 criterii -» Cont probabil autentic	Contul nu prezintă suficiente semne evidente de activitate suspectă. Poate fi cont nou sau al unei persoane fără prea multă activitate.
4-6 criterii -» Cont suspect	Contul prezintă suficiente semne de activitate problematică. Există o probabilitate mare ca acesta să fie un cont fals, utilizat pentru manipulare, propagandă și/sau astroturfing.
7+ criterii -» Cont foarte suspect	Contul bifează multe criterii pentru a fi considerat fals sau furat, parte a unei rețele de dezinformare sau un cont utilizat pentru propagandă. Se recomandă raportarea la Facebook.

Explicarea criteriilor din grila de identificare a conturilor false din Facebook:

1. Nume atipic – Generarea automată a numelor pentru conturi false poate duce la discrepanțe evidente, cum ar fi combinarea a două

nume din culturi diferite sau formatarea incoerentă a prenumelui şi numelui. Un semnal puternic de alarmă este nepotrivirea dintre nume şi alte elemente ale contului, cum ar fi poza de profil sau regiunea geografică indicată. În 2025, algoritmii de generare sunt mai avansaţi, dar multe reţele de boţi continuă să folosească baze de date slabe, ceea ce face ca numele incoerente să rămână un indicator valid.

2. Nume schimbat recent – Poate indica reutilizarea unui cont, adesea parte dintr-o reţea de boţi sau pregătit pentru revânzare şi reutilizare în alte scopuri. Conturile vechi sunt mai valoroase, deoarece trec mai uşor de filtrele Facebook, iar schimbarea numelui este un prim pas în procesul de ascundere a activităţii anterioare. Deşi Facebook permite modificarea link-ului profilului, multe conturi suspecte nu o fac, din motive precum restricţiile platformei, graba operatorilor sau limitările de diverse resurse. Astfel, dacă link-ul contului conţine un nume diferit de cel actual, acesta poate fi un indiciu puternic al unui cont reciclat sau compromis.

3. Fără poză de profil/cover – Lipsa unei poze de profil sau a unui cover este frecventă la conturile false, în special cele generate rapid pentru scopuri de propagandă sau spam. Conturile automate simple sunt mai uşor şi mai rapide de creat fără a încărca imagini, evitând astfel verificările suplimentare şi economisind resurse. Această tactică ajută conturile false să evite atragerea atenţiei, deoarece utilizatorii reali tind să ignore sau să nu interacţioneze cu profiluri goale. În plus, boţii care generează şi controlează un număr mare de conturi pot avea limitări hardware, iar simplificarea conturilor reduce costurile.

4. Poză de profil generică, generată cu inteligenţă artificială, furată de pe internet sau de calitate foarte proastă – Conturile false folosesc frecvent poze de profil generice, fie imagini cu peisaje, flori sau animale, fie fotografii generate cu inteligenţă artificială, furate de pe internet sau de calitate foarte slabă. Scopul acestora este de a evita detectarea automată şi de a crea o aparenţă de autenticitate fără a expune informaţii personale. Unele conturi aleg imagini prea artistice pentru a părea atractive, iar altele folosesc poze reale, dar furate, ceea ce poate fi verificat prin căutări inverse. De asemenea, aceste conturi evită postarea de fotografii cu alte persoane reale, deoarece menţinerea unei identităţi credibile ar necesita interacţiuni constante.

5. Schimbări frecvente ale pozei de profil – Conturile false îşi pot schimbă frecvent poza de profil pentru a evita detectarea automată şi pentru a crea impresia unei activităţi autentice. Această practică este folosită de reţelele de boţi pentru a înşela algoritmii de moderare şi

pentru a împiedica utilizatorii să le recunoască. Uneori, schimbările sunt realizate între imagini generice, poze furate sau chiar fotografii generate cu AI. Dacă un cont își modifică imaginea neobișnuit de des, fără un motiv clar, poate fi un indiciu că încearcă să ascundă o identitate falsă sau că este parte dintr-o rețea artificială.

6. Cont fără informații personale – Cu cât se dorește crearea automată a unor conturi de Facebook mai complete și mai realiste, cu atât procesul devine mai dificil și mai expus detectării de către algoritmii platformei. Din acest motiv, administratorii rețelelor de boți preferă să genereze profiluri minimale, completând doar datele strict necesare pentru a trece de filtrele inițiale ale Facebook.

Atunci când se încearcă adăugarea unor informații mai detaliate, pot apărea incongruențe, precum nepotrivirea dintre vârstă și fotografia de profil, disonanțe între genul numelui și cel al pozelor sau contradicții între locul de naștere, nume și poza de profil. Pentru a evita aceste erori și pentru a minimiza riscul de identificare, conturile false rămân în general lipsite de detalii personale, permițând și reutilizarea lor în diverse campanii de manipulare.

7. Cont nou (<6 luni) sau schimbări bruște de identitate – Facebook își îmbunătățește constant algoritmii de detectare a conturilor false, ceea ce face ca durata de viață a acestora să fie, în general, scurtă. De aceea, majoritatea conturilor false utilizate în campaniile de manipulare sunt conturi noi, create rapid și în număr mare prin botnet-uri, automatizări și rețele proxy. Deși conturile furate sunt mai greu de detectat, utilizarea acestora este limitată, deoarece securizarea cu doi factori și procedurile de recuperare le fac mai greu de exploatat la scară largă. Astfel, schimbările bruște de identitate sau activitatea ridicată a unui cont recent creat pot fi indicatori eficienți ai unui profil suspect.

8. Cont fără activitate/conținut – Întrucât generarea automată de conținut autentic și convingător la scară largă este dificilă, multe conturi false aleg să nu posteze deloc sau să aibă un conținut minim, pentru a evita detectarea de către algoritmii Facebook. Lipsa activității este o strategie de camuflare folosită de rețelele de boți, care preferă să interacționeze prin like-uri, comentarii generice sau distribuiri automate, fără a publica postări proprii. Totuși, manipularea online este un fenomen dinamic, iar calitatea și strategia utilizării conturilor false variază în funcție de experiența și resursele celor care le gestionează.

9. Link cont nepersonalizat – Multe conturi false de Facebook folosesc link-uri generate automat, fie pentru că nu îndeplinesc cerinţele platformei pentru personalizare, fie pentru a fi mai uşor reutilizate în alte campanii, fie pentru că procesul de personalizare a URL-ului este omis intenţionat, fiind considerat inutil. Resursele sunt concentrate pe crearea unui număr mare de profiluri, iar schimbarea link-ului ar necesita programare suplimentară, consum de timp şi un risc mai mare de erori, ceea ce face ca reţelele de conturi false să evite acest pas şi să aleagă variantele cele mai simple şi rapide. În plus, link-urile personalizate pot servi drept indiciu pentru identificarea şi monitorizarea conturilor suspecte, motiv pentru care reţelele de dezinformare preferă să le lase generice, făcând conturile mai greu de urmărit şi detectat.

10. Frecvenţă neregulată a postărilor – Conturile false prezintă adesea o frecvenţă neregulată a postărilor, alternând între perioade lungi de inactivitate şi episoade de activitate intensă. Această oscilaţie este un indiciu al automatizării, deoarece boţii sau operatorii umani care le controlează acţionează în valuri organizate, în funcţie de necesităţile campaniilor de dezinformare, spam sau influenţă. Un cont real are, de obicei, o activitate constantă şi organică, în timp ce unul fals poate posta masiv într-un interval scurt, distribuind conţinut repetitiv, pentru ca apoi să dispară. Acest comportament este frecvent întâlnit în reţelele coordonate, unde conturile sunt activate doar în momente strategice, evitând detectarea şi restricţiile algoritmilor Facebook.

11. Activitate anormală – Conturile false sau automatizate pot prezenta activitate anormală, caracterizată printr-un volum mare de postări zilnice, fără a genera interacţiuni reale. Studiile arată că utilizatorii obişnuiţi postează, în medie, 1-2 postări pe zi, iar frecvenţe mai mari sunt specifice paginilor oficiale sau influencerilor. În schimb, un cont suspect poate posta de peste 4-5 ori pe zi, adesea conţinut distribuit automat, fără comentarii personalizate sau răspunsuri la interacţiuni.

Această strategie este utilizată de reţelele de astroturfing şi propagandă, care se bazează pe cantitate, nu pe engagement autentic. Lipsa reacţiilor naturale, cum ar fi like-uri sau comentarii de la prieteni reali, poate fi un indicator clar al unui cont fals.

12. Postări 24/24 – Un utilizator real are pauze naturale de activitate, reflectând orele de somn, muncă şi alte activităţi offline. În schimb, conturile false sau automatizate pot avea postări constante, 24/24, indiferent de fusul orar al ţării în care pretind că activează. Acest

tip de activitate nefirească este un indiciu puternic al utilizării de boți, al conturilor gestionate de echipe din alte zone geografice sau al rețelelor coordonate de dezinformare. Adesea, coordonatorii boților nu setează un ritm realist de postare, ceea ce face ca aceste conturi să fie active în timpul nopții, fără variații naturale. Deși utilizatorii reali pot avea activitate nocturnă ocazională, postările constante la ore atipice pot ridica suspiciuni, mai ales atunci când sunt combinate cu alte semnale de alarmă.

13. Postări fără descriere – Conturile false distribuie adesea postări fără descriere, deoarece generarea automată de texte coerente, relevante și corect scrise rămâne încă o provocare pentru rețelele de boți. În plus, Facebook utilizează algoritmi avansați care analizează conținutul textual pentru a detecta tipare suspecte. Prin urmare, lipsa unei descrieri poate fi o strategie folosită pentru a evita analiza lingvistică și a reduce riscul de detectare. Deși utilizatorii reali pot uneori să partajeze conținut fără comentarii, conturile false fac acest lucru constant, postând sau redistribuind materiale fără nicio explicație personalizată. Această lipsă de context și de interacțiune autentică este un indiciu bun de automatizare sau de activitate coordonată.

14. Trecerea bruscă de la conținut personal la activitate suspectă – Un cont real de Facebook are un tipar de activitate constant, postând ocazional conținut personal, actualizări despre viața utilizatorului sau interacțiuni autentice. În schimb, un cont furat încetează brusc să mai posteze astfel de informații și trece la distribuirea conținutului standardizat, adesea axat pe propagandă, dezinformare sau astroturfing. Această schimbare bruscă poate fi un semnal clar al reutilizării contului într-o rețea coordonată. Administratorii conturilor compromise evită uneori să șteargă istoricul postărilor pentru a menține o aparentă autenticitate, însă modifică tipul de conținut promovat. Dacă un profil își schimbă brusc stilul de postare, renunțând la interacțiuni personale în favoarea unui flux suspect de distribuiri, acesta trebuie analizat atent.

15. Descrieri scrise greșit/incoerent – Textele incoerente sau cu greșeli gramaticale pot indica utilizarea traducerilor automate, a generatoarelor de text slab optimizate sau a unor algoritmi care nu sunt adaptați la nuanțele limbii române. Deși inteligența artificială a evoluat semnificativ, unele rețele de boți încă folosesc sisteme mai puțin avansate, ceea ce duce la formulări rigide, structuri ciudate sau greșeli frecvente. Totuși, nu toate textele incorecte aparțin conturilor false și nu toate textele corect scrise sunt autentice. Un cont devine suspect atunci

când postează în mod constant mesaje cu erori recurente, traduceri nefireşti sau fraze care nu se potrivesc contextului conversaţional.

16. Conţinut intens pe o temă – Conturile false sunt adesea concentrate pe un singur subiect, reflectând agenda reţelei care le controlează. Acestea pot promova teme precum politică, conspiraţii, activism extremist sau dezinformare. Spre deosebire de utilizatorii reali, care abordează subiecte diverse şi au interacţiuni variate, conturile suspecte postează exclusiv conţinut legat de un singur domeniu, fără variaţii naturale. Acest tipar indică o posibilă implicare într-o campanie coordonată de astroturfing, menită să amplifice artificial un mesaj, să polarizeze opinia publică sau să creeze percepţia falsă că există un interes masiv pentru o anumită idee.

17. Conţinut preponderent instigator – Conturile false pot fi utilizate pentru a polariza discursul public, promovând mesaje care instigă la ură, divizare socială şi conflict. Aceste tactici sunt frecvent folosite în campanii de manipulare şi destabilizare, deoarece tensiunile sociale facilitează influenţarea opiniilor şi radicalizarea grupurilor. Astfel de conturi distribuie conţinut menit să exacerbeze diferenţele politice, culturale sau ideologice, amplificând teme controversate precum teorii ale conspiraţiei, naţionalism extrem, conflicte etnice sau sociale. De multe ori, postările sunt construite pentru a stârni emoţii puternice, cum ar fi furie sau indignare, generând reacţii rapide şi diminuând gândirea critică. Această strategie permite manipularea percepţiei colective şi crearea iluziei unei susţineri largi pentru idei extreme.

18. Raport ciudat între numărul de prieteni şi cel de interacţiuni – Un cont care are mii de prieteni, dar prezintă foarte puţine interacţiuni reale poate fi un indiciu al unui profil fals sau automatizat. În mod natural, un utilizator activ pe Facebook primeşte like-uri, comentarii şi reacţii la postările sale, în special de la prietenii apropiaţi. În schimb, conturile suspecte pot avea liste extinse de prieteni, dar fără ca aceştia să interacţioneze real cu postările lor. Această discrepanţă poate apărea deoarece mulţi boţi şi conturi cumpărate adaugă prieteni în masă, dar nu au un istoric autentic de conversaţii sau engagement.

19. Prieteni cu conturi suspecte – Conturile false de pe Facebook adoptă uneori strategii de interconectare între ele pentru a crea o aparentă autenticitate. Boţii sau conturile controlate manual din reţele de dezinformare se adaugă reciproc ca prieteni, îşi distribuie conţinutul, îşi dau like-uri şi comentarii între ele, încercând să genereze o impresie falsă de popularitate şi legitimitate.

Un indiciu al unui cont suspect este o listă de prieteni formată majoritar din alte conturi cu caracteristici similare. Acest tipar este des întâlnit în rețelele de propagandă, unde un grup de conturi suspecte se validează reciproc pentru a amplifica artificial mesajele și a crește vizibilitatea conținutului distribuit.

20. Prieteni dispersați geografic – Un utilizator real are, de obicei, o rețea de prieteni formată în mare parte din persoane din aceeași țară sau cu care împărtășește un context comun, precum școala, locul de muncă sau evenimente sociale. În schimb, conturile false prezintă frecvent prieteni dispersați geografic, adăugând utilizatori din zone fără nicio conexiune logică.

Acest tipar poate apărea în cazul conturilor automatizate care se adaugă între ele pentru a crea o aparență de autenticitate, dar și în cazul conturilor cumpărate, care sunt reutilizate pentru diverse scopuri. De exemplu, un cont aparent românesc, dar cu prieteni preponderent din Africa, Asia sau America de Sud, poate fi suspect. Rețelele de boți sau de propagandă globală folosesc această strategie pentru a-și întări artificial credibilitatea și pentru a evita detectarea rapidă de către algoritmii Facebook, însă această lipsă de coerență geografică rămâne un indicator bun al unui cont suspect.

21. Toți sau majoritatea prietenilor sunt adăugați recent – Un cont real își formează rețeaua de prieteni treptat, prin interacțiuni naturale. În schimb, un cont fals adaugă zeci sau sute de prieteni într-un timp foarte scurt, semn că face parte dintr-o rețea artificială. Unele conturi sunt create de la zero, iar altele sunt conturi furate prin phishing. Acestea își schimbă numele, poza de profil și lista de prieteni, ștergând conexiunile originale și adăugând rapid noi prieteni specifici scopului urmărit. Dacă un cont are multe conexiuni adăugate brusc, fără interacțiuni reale, este probabil un profil fals folosit într-o campanie coordonată.

22. Adaugă prieteni doar dintr-o anumită nișă – Conturile false sunt adesea create pentru a se infiltra în comunități specifice, precum grupuri politice, conspiraționiste, economice sau activiste. În loc să adauge prieteni în mod diversificat, acestea își formează rețeaua preponderent sau chiar exclusiv în jurul unei teme de interes, conectându-se cu utilizatori care împărtășesc aceeași ideologie sau preocupare. Această strategie ajută conturile suspecte să capete legitimitate, să amplifice mesaje coordonate și să influențeze mai ușor conversațiile din acea nișă. De multe ori, aceste profiluri sunt utilizate pentru propagandă, manipulare sau escrocherii, fiind concepute pentru

a răspândi dezinformări într-un anumit mediu. Dacă un cont are prieteni preponderent sau exclusiv dintr-o singură categorie (ex. activişti radicali, promotori ai unei/unor conspiraţii etc), fără variaţie socială, acest tipar poate fi un semnal de alarmă.

23. Cerc social incoerent – Un cont real are o reţea de prieteni logică, formată în mod natural. În schimb, conturile false pot prezenta discrepanţe evidente, fie din cauza unei construcţii artificiale, fie pentru că încearcă să exploateze principiul autorităţii, pretinzând că aparţin unor profesii de încredere, precum militari, doctori, profesori sau preoţi. Totuşi, astfel de profiluri ar trebui să aibă prieteni din acelaşi domeniu. Un medic fără conexiuni cu alţi doctori, un militar fără prieteni din armată sau un bancher bogat cu o reţea formată preponderent din persoane cu venituri mici pot ridica suspiciuni.

24. Număr mic de prieteni (<250) – În 2013, un studiu arăta că utilizatorii de Facebook aveau, în medie, 338 de prieteni, într-o perioadă în care platforma număra 1,23 miliarde de utilizatori activi lunar. De atunci, Facebook a crescut considerabil, ajungând în 2025 la peste 3 miliarde de utilizatori. Deoarece nu există studii recente care să ofere o medie actualizată a numărului de prieteni pe Facebook, estimez că această valoare a crescut proporţional cu expansiunea platformei. Dacă numărul de utilizatori s-a extins de aproximativ 2,5 ori, consider că şi numărul mediu de prieteni a crescut în mod similar, situându-se acum între 500 şi 700 de prieteni per utilizator activ.

Pentru a stabili pragul minim sub care un cont devine suspect, mă bazez pe analiza distribuţiei sociale. În orice reţea socială, utilizatorii tind să se încadreze în jurul mediei, iar cei care cad semnificativ sub acest prag sunt fie utilizatori ocazionali, fie conturi suspecte. Consider că un prag de 250 de prieteni este rezonabil, fiind situat între 30% şi 40% din media estimată, ceea ce înseamnă că un cont real ar trebui să depăşească acest număr pentru a nu ridica suspiciuni.

Conturile false au dificultăţi în acumularea prietenilor, fie din cauza lipsei unei identităţi credibile, fie din cauza restricţiilor Facebook, care limitează cererile de prietenie şi elimină automat conturile suspecte. Desigur, un număr mic de prieteni nu este singurul criteriu care defineşte un cont fals. Însă, pe baza acestei analize, consider că un profil cu mai puţin de 250 de prieteni poate ridica suspiciuni şi trebuie analizat în combinaţie cu alţi factori.

25. Tipar de distribuire în masa – Un utilizator autentic de Facebook îşi exprimă, de regulă, opiniile şi prin comentarii, reacţii şi

postări proprii. În schimb, conturile suspecte afișează un tipar de distribuire excesivă, în care singura sau preponderenta activitate este cea de distribuire masivă de conținut, fără comentarii sau explicații. Acest comportament sugerează că respectivul cont este folosit pentru propagare, nu pentru interacțiune, fiind un instrument folosit în campaniile de astroturfing. Astfel de conturi distribuie frecvent postări din aceleași surse, preponderent îndoielnice și concentrate pe un singur subiect. În contextul astroturfing-ului, acest model de distribuție este utilizat pentru a crea iluzia unui sprijin popular, amplificând artificial anumite mesaje, fie prin conținut politic, fie prin dezinformare strategică. Mai mult, distribuirea excesivă în grupuri multiple, într-un timp scurt, poate indica o tentativă de manipulare sau automatizare, caracteristici esențiale ale rețelelor de influență artificială.

26. Lipsa reacțiilor la propriile postări – Un utilizator real de Facebook primește, de regulă, reacții și comentarii la postările sale, fie de la prieteni, fie de la urmăritori cu interese comune. În schimb, conturile suspecte au o lipsă totală sau aproape totală de interacțiuni, ceea ce poate indica o audiență reală inexistentă. Acest tipar este comun în cazul conturilor automatizate folosite în campanii de astroturfing, unde scopul principal este de distribuirea a unui mesaj.

27. Nu răspunde la mesaje private și comentarii – Un cont real de Facebook interacționează în mod natural cu ceilalți utilizatori, răspunzând la comentarii și mesaje private, fie și ocazional. În schimb, conturile false sunt create în număr mare și gestionate automat sau parțial, ceea ce face dificilă monitorizarea și răspunsul personalizat la interacțiuni. Deoarece aceste conturi sunt utilizate pentru astroturfing, ele nu sunt concepute pentru conversații directe. Răspunsurile ar necesita resurse tehnologice și umane semnificative, ceea ce nu este eficient pentru administratorii acestor rețele. În plus, un răspuns neinspirat sau incoerent ar putea crește suspiciunile utilizatorilor reali.

28. Comentarii generice, scurte și repetitive – Un utilizator real tinde să lase comentarii variate, cu argumente, opinii sau reacții personalizate. În schimb, conturile suspecte folosesc mesaje scurte, generice și repetitive, precum „Adevărat!", „Minciuni!", „Fake news!", „Toată lumea știe!", „Trezirea!" etc. Acest tipar este comun în campaniile de astroturfing și manipulare, unde conturile false sunt folosite pentru a amplifica artificial anumite mesaje. Comentariile generice sunt preferate pentru că sunt simple, permit ușoare erori în interpretarea contextului și sunt ușor de generat automat.

29. Distribuire excesivă de link-uri suspecte – Conturile false distribuie constant link-uri de pe site-uri special create pentru manipulare, concepute să pară publicaţii autentice. Aceste platforme imită site-uri de ştiri cunoscute, folosind nume asemănătoare sau chiar denumiri ale unor publicaţii care nu mai există, pentru a părea surse legitime. În loc să partajeze conţinut diversificat, asemenea unui comportament uman normal, aceste conturi promovează în mod repetitiv aceleaşi surse obscure, cu informaţii manipulate, conspiraţii sau propagandă.

31. Legături cu alte conturi suspecte – Un cont autentic de Facebook primeşte interacţiuni diverse, de la prieteni reali cu opinii şi comportamente variate. În schimb, conturile suspecte funcţionează în reţele coordonate, unde like-urile, comentariile şi distribuirile provin mereu din aceleaşi surse, adesea alte conturi suspecte. Acest tipar poate indica o infrastructură artificială de amplificare a mesajelor, folosită în campanii de astroturfing, propagandă sau dezinformare. De multe ori, aceste conturi se susţin reciproc prin comentarii generice sau reacţii automate, fără o interacţiune autentică. Analizând profilurile care interacţionează frecvent cu un cont suspect, putem observa că multe dintre ele prezintă aceleaşi caracteristici artificiale, au activitate restrânsă la un singur subiect şi conexiuni limitate la alte conturi suspecte. Acest comportament poate indica o reţea organizată, menită să creeze falsa impresie de popularitate şi să manipuleze percepţiile publicului.

31. Postări codificate – Pentru a evita detectarea de către algoritmii Facebook, conturile suspecte folosesc uneori postări codificate, în care anumite cuvinte cheie sunt modificate deliberat prin simboluri, caractere speciale sau greşeli ortografice. Această tehnică încearcă să păcălească algoritmii, îngreunând identificarea şi ştergerea conţinutului care încalcă regulile platformei. De exemplu, în loc de „vaccin", se poate scrie „v@cc1n", iar în loc de „guvern", se foloseşte „g0vern" sau „gu vern min ciu nă". Unele postări încearcă să păcălească algoritmii şi prin scrierea fonetică greşită „guvărnul minchinos". Această metodă este des utilizată în campaniile de astroturfing şi dezinformare, unde este crucial ca mesajele să circule fără a fi blocate. Un cont care recurge frecvent la astfel de tactici, mai ales pe subiecte controversate, poate indica o încercare deliberată de manipulare sau evitare a moderării platformei.

32. Comentarii cu link-uri ascunse – O tehnică frecvent utilizată de conturile suspecte pentru a disemina conţinut manipulat

sau propagandistic este plasarea strategică a link-urilor în *„blind spotul”* („punctul mort”) Facebook, ascunzându-le sub butonul „Vezi mai mult” din comentarii. Această metodă este folosită în încercarea de a evita detectarea automată a link-urilor suspecte, deoarece Facebook analizează cu prioritate partea vizibilă a unui comentariu, din considerente de optimizare a resurselor hardware. Pentru a determina utilizatorul să dea click pe „Vezi mai mult” și ulterior să acceseze linkul, prima parte a mesajului este concepută în stil clickbait – un text scurt, care captează atenția, creează curiozitate și îl impulsionează să dea clic si pe link.

33. Schimbare bruscă a limbii utilizate în conținut – Un utilizator real își păstrează, de regulă, coerența în utilizarea limbii, chiar dacă poate interacționa ocazional în alte limbi. În schimb, un cont suspect poate prezenta o schimbare bruscă și completă a limbii în care postează și comentează, fără o explicație logică. Acesta este un posibil semn că respectivul cont a fost furat printr-o strategie de phishing și ulterior reutilizat pentru un alt scop, fără a fi resetat complet. Hackerii care operează astfel de conturi preferă uneori să păstreze istoricul profilului, deoarece acest lucru îl face mai credibil. De exemplu, un cont care anterior posta doar în sârbă și brusc, începe să posteze exclusiv în română, poate fi parte dintr-o rețea de astroturfing sau dezinformare.

34. Diferențe între locația profilului și istoricul activității – Un utilizator autentic are, de obicei, o activitate coerentă cu locația sa declarată, reflectată în check-in-uri, evenimente, postări și interacțiuni. În schimb, un cont suspect poate prezenta discrepanțe între țara pe care o afișează și istoricul activității sale. De exemplu, un profil care pretinde că este din România, dar are check-in-uri și postări vechi dintr-o altă țară, fără nicio explicație logică (ex. mutare, călătorii frecvente), poate fi un indiciu al furtului și reutilizării contului. Acest lucru este frecvent întâlnit în cazul conturilor compromise prin phishing, care sunt preluate și integrate în rețele de astroturfing și manipulare digitală.

Limitările studiului

Deși prezenta cercetare propune instrumente operaționale validate pentru identificarea conturilor implicate în campanii de astroturfing, aceasta prezintă anumite limitări metodologice. În primul rând, validarea grilelor a fost realizată pe un eșantion calitativ restrâns, format din 20 de conturi suspecte. Deși rezultatele susțin ipotezele

teoretice formulate, dimensiunea eşantionului limitează generalizarea statistică la nivelul întregului ecosistem digital.

În al doilea rând, analiza s-a concentrat exclusiv asupra platformei Facebook, cu particularităţile sale structurale şi algoritmice. Indicatorii de manipulare şi tipologiile de conturi false pot diferi semnificativ pe alte platforme, unde nivelul de filtrare, gradul de permisivitate şi formatele de conţinut sunt diferite.

Nu în ultimul rând, dinamica accelerată a tehnologiilor de inteligenţă artificială generativă (AI) reprezintă o provocare majoră. Capacitatea tot mai avansată a modelelor AI de a produce imagini hiperrealiste, texte corecte, coerente şi interacţiuni aparent autentice riscă să reducă în curând relevanţa unor indicatori vizuali sau lingvistici utilizaţi în prezent, ceea ce impune actualizarea periodică a criteriilor de detecţie.

Concluzii

Studiul de faţă arată că astroturfing-ul nu constituie o simplă problemă tehnologică, ci o formă complexă de influenţă psihologică, care exploatează vulnerabilităţi cognitive fundamentale pentru a distorsiona realitatea percepută. Analiza teoretică evidenţiază faptul că eficienţa acestor campanii este strâns legată de activarea unor mecanisme de psihologie socială: validitatea socială, care creează iluzia consensului, apelul la frică şi ameninţare, menit să inhibe gândirea critică, precum şi utilizarea schemelor maniheiste şi mesianice pentru polarizarea discursului public. În acest cadru, utilizatorul obişnuit este transformat din receptor pasiv într-un agent involuntar de diseminare a dezinformării.

Răspunsul operaţional propus şi testat în această cercetare constă în formalizarea procesului de detecţie prin cele două grile de evaluare a conturilor suspecte. Validarea empirică a demonstrat că, în pofida progresului tehnologic, conturile false continuă încă să lase urme comportamentale şi structurale identificabile. Grila extinsă s-a dovedit adecvată analizei aprofundate, în timp ce varianta simplificată oferă un instrument rapid, util pentru menţinerea unei minim scut în mediul digital.

Cu toate acestea, nici cele mai performante instrumente nu pot substitui complet factorul uman. Aşa cum a fost argumentat anterior, educaţia digitală şi alfabetizarea media reprezintă în contextul anului 2025, elemente centrale ale rezilienţei societăţii. Utilizatorii trebuie să fie

capabili nu doar să identifice conținutul manipulator, ci și să înțeleagă logica algoritmilor ce amplifică artificial anumite narațiuni. În mod complementar, securitatea cibernetică personală devine o componentă esențială, întrucât compromiterea conturilor legitime facilitează integrarea acestora în rețele de influență prin furt de identitate.

Un element suplimentar, preluat din practica jurnalistică clasică, rămâne esențial în procesul de evaluare critică a discursului public: întrebarea privind beneficiarul unei narațiuni. Analiza oricărui mesaj ar trebui să includă constant o reflecție asupra intereselor pe care acesta le servește, a actorilor care ar putea sta în spatele său și a efectelor urmărite la nivel social, economic sau politic. O astfel de abordare presupune depășirea reacțiilor emoționale imediate și plasarea informației într-un cadru mai larg, care permite o înțelegere mai lucidă a dinamicilor de influență.

În final, deși responsabilitatea individuală și cultivarea gândirii critice sunt indispensabile, combaterea eficientă a astroturfing-ului nu poate rămâne exclusiv la nivel individual. Cercetarea subliniază necesitatea unei cooperări între platformele digitale și autorități, în vederea dezvoltării unor mecanisme automatizate de detecție, fundamentate pe indicatorii identificați și a adaptării cadrului normativ la provocările generate de inteligența artificială. Doar prin convergența dintre utilizatorul informat, capabil de analiză calmă și critică și mecanismele instituționale de protecție poate fi menținută in limite rezonabile siguranța spațiului public digital în fața operațiunilor de influență hibridă.

Bibliografie

1. André, Christophe. 2004. *Psihologia fricii: temeri, angoase și fobii*. București: Editura Trei.

2. Autoritatea Electorală Permanentă. 2024. „Proiect de hotărâre privind aprobarea Ghidului finanțării campaniei electorale”. *Roaep.ro*. Accesat la 20 februarie 2025. https://www.roaep.ro/legislatie/wp-content/uploads/2025/02/PROIECT-HOTARARE-MATERIAL-PUBLICITATE-POLITICA.pdf.

3. Bârgăoanu, Alina. 2018. *#FAKENEWS. O nouă cursă a înarmării*. București: Evrika Publishing.

4. Bichir, Florian. 2020. „Conceptul de mesianism în geopolitică". *Geopolitica*, 2020. https://www.geopolitic.ro/wp-content/uploads/2020/02/GEOPOLITICA1.html.

5. Boncu, Ştefan. 2014. *Psihologie socială*. Iaşi: Editura Polirom.

6. Catrina, Geanina. 2024. „Legile internetului. Reacţii la ameninţările informaţionale". *Revista Intelligence*, 20 august 2024. https://intelligence.sri.ro/legile-internetului-reactii-la-amenintarile-informationale/.

7. Chan, Jovy. 2022. „Online Astroturfing: A Problem beyond Disinformation". *Philosophy & Social Criticism*, iunie, 01914537221108467. https://doi.org/10.1177/01914537221108467.

8. Chelcea, Septimiu. 2006. *Opinia publică. Strategii de persuasiune şi manipulare*. Bucureşti: Editura Economică.

9. CSMonitor. 2009. „The Time-Honored Practice of Astroturf Lobbying". *Christian Science Monitor*, 2009. https://www.csmonitor.com/USA/Politics/Decoder/2009/0903/The-time-honored-practice-of-Astroturf-lobbying.

10. Dobrescu, Paul şi Bârgăoanu, Alina. 2003. *Mass media şi societatea*. Bucureşti: Comunicare.ro.

11. GDPR. 2025. „General Data Protection Regulation in 2025". *ComplyDog*, 2025. https://complydog.com/blog/gdpr-in-2025.

12. Georgescu, Paul. 2024. „Vinovaţii fără vină. Diseminarea neintenţionată a mesajelor false". *Revista Intelligence*, 2024. https://intelligence.sri.ro/vinovatii-fara-vina-diseminarea-neintentionata-mesajelor-false/.

13. Ghiurgiu, Florin Mitruţ. 2024. „Infosfera şi ecourile realităţii virtuale". *Revista Intelligence*, 2024. https://intelligence.sri.ro/infosfera-si-ecourile-realitatii-virtuale/.

14. Lazarsfeld, Paul; Berelson, Bernard şi Gaudet, Hazel. 2004. *Mecanismul votului. Cum se decid alegătorii într-o campanie prezidenţială*. Bucureşti: Comunicare.ro.

15. McAfee. 2022. „How To Spot A Fake Facebook Account". *McAfee Blog*, 9 decembrie 2022. https://www.mcafee.com/learn/spot-fake-facebook-account/.

16. Merriam-Webster. 2023. „Definition of ASTROTURFING". *Merriam-Webster.com*. Accesat la 14 iulie 2023. https://www.merriam-webster.com/dictionary/astroturfing.

17. Meta. 2025. „Condiţiile de utilizare Meta". *Facebook.com*. Accesat în 2025. https://ro-ro.facebook.com/terms/.

18. Oprea, Bogdan. 2021. *Fake news şi dezinformare online: recunoaşte şi verifică*. Iaşi: Editura Polirom.

19. Roman, Dan. 2024. „Internetul, teatru de război. Militarizarea informaţiei". *Revista Intelligence*, 6 iunie 2024. https://intelligence.sri.ro/internetul-teatru-de-razboi-militarizarea-informatiei/.

20. Roşca, Tatiana. 2019. „Dinamica componentelor identităţii sociale în psihologia modernă". *Studia Universitatis Moldaviae*, nr. 9 (129).

21. Stan, Mircea. 2021. *Programul de măsuri active al KGB-GRU împotriva României (1964-1989)*. Bucureşti: Editura Militară.

22. Sturza, Cătălin. 2021. „Pericolele mesianismului politic". *Adevărul*, 27 august 2021. https://adevarul.ro/blogurile-adevarul/pericolele-mesianismului-politic-2116291.html.

23. Turner, J.C. şi Reynolds, Katherine. 2012. „Self-Categorization Theory". În *Handbook of Theories in Social Psychology*, 399-417. Londra: SAGE Publications.

24. Volkoff, Vladimir. 2009. *Tratat de dezinformare*. Bucureşti: Editura Antet.

Aceasta este al patrulea volum de proceedings al Conferinței Științifice Intelligence și Cultura de Securitate (ICS), care cuprinde lucrările prezentate în cadrul ediției din 2025 - ICS 2025, publicat de Academia Națională de Informații „Mihai Viteazul" (ANIMV). ICS continuă să ofere studenților o platformă pentru dialog academic și pentru a împărtăși realizările lor științifice.

Ediția actuală își extinde participarea la un spectru mai larg de contributori, incluzând atât doctoranzi, cât și studenți din programele de master, cu un interes crescut pentru domenii precum intelligence, securitate națională, istorie și relații internaționale.

Organizarea conferinței a fost posibilă datorită eforturilor continue ale doctoranzilor și ale conducătorilor de doctorat din cadrul Școlii Doctorale Intelligence și Securitate a ANIMV.

Anticipăm cu entuziasm noi discuții și schimburi de idei în viitoarea ediție a conferinței.

EDITURA ANIMV

ISSN 2972-1350
ISSN-L 2971-8139