

Inteligența artificială în scrierea academică

- Punct de vedere al Consiliului Științific al Universității Babeș-Bolyai –

Cuprins

Inteligența artificială în scrierea academică.....	1
Context.....	1
Educația asupra utilizării IAGen	2
IAGen în generarea de cunoaștere - CDI.....	3
IAGen în scrierea academică.....	3
IAGen, GDPR și responsabilitatea	4
Abordarea CS-UBB	4
Recomandări	4
Referințe.....	6
Anexă: recomandări pentru folosirea responsabilă ca IAGen ca instrument copilot în lucrările de tip IAGen-S	8

Context

Instrumentele de inteligență artificială (IIA) sunt în uz și dezvoltare accelerată de la mijlocul secolului XX încoace, odată cu dezvoltarea calculatoarelor. Analizele de date și rezolvarea de probleme complexe cu IIA sunt de multă vreme încetățenite în cercetarea științifică și nu numai. Mai recent, în special din noiembrie 2022 încoace (momentul chatGPT), au intrat de asemenea în uz general aplicații IIA care pot genera elemente anterior asumate apanaj unic al inteligenței umane – text scris, imagini, muzică etc.¹⁻⁴ La aceste IIA generative (IAGen) se referă textul de față. În mediul academic la nivel internațional, IAGen este deja folosit pe scară largă, dar în mare măsură haotic și nereglementat – atât în scopuri legitime cât și ilegite.

Instituții și organizații de referință în lumea academică au găsit de cuviință să reglementeze sau măcar să facă recomandări privind IAGen în activitățile care le privesc. Aceste recomandări/reguli variază mult și se anticipează că vor evolua odată cu domeniul foarte dinamic al IAGen – de la interzicere aproape completă la permisiunea de folosire cu citare sau la altele.⁵

Revista Science⁶ prevede că „Artificial intelligence (AI) AI-assisted technologies [such as large language models (LLMs), chatbots, and image creators] **do not meet the Science journals’ criteria for authorship** and therefore **may not be listed as authors** or coauthors, **nor may sources cited in Science journal content be authored or coauthored by AI tools**. Authors who use AI-assisted technologies as components of their research study or as aids in the writing or presentation of the manuscript should note this in the cover letter and in the acknowledgments section of the manuscript. Detailed information

should be provided in the methods section: The full prompt used in the production of the work, as well as the AI tool and its version, should be disclosed. Authors are accountable for the accuracy of the work and for ensuring that there is no plagiarism. They must also ensure that all sources are appropriately cited and should carefully review the work to guard against bias that may be introduced by AI. Editors may decline to move forward with manuscripts if AI is used inappropriately. Reviewers may not use AI technology in generating or writing their reviews because this could breach the confidentiality of the manuscript. **AI-generated images and other multimedia are not permitted in the Science journals without explicit permission from the editors.** Exceptions may be granted in certain situations—e.g., for images and/or videos in manuscripts specifically about AI and/or machine learning. Such exceptions will be evaluated on a case-by-case basis and should be disclosed at the time of submission. The Science journals recognize that this area is rapidly developing, and our position on AI-generated multimedia may change with the evolution of copyright law and industry standards on ethical use.”

Editura Elsevier⁷ spune că „**Reviewers should not upload a submitted manuscript or any part of it into a generative AI tool** as this may violate the authors’ confidentiality and proprietary rights and, where the paper contains personally identifiable information, may breach data privacy rights. This confidentiality requirement extends to the peer review report, as it may contain confidential information about the manuscript and/or the authors. For this reason, **reviewers should not upload their peer review report into an AI tool**, even if it is just for the purpose of improving language and readability. Peer review is at the heart of the scientific ecosystem and Elsevier abides by the highest standards of integrity in this process. Reviewing a scientific manuscript implies responsibilities that can only be attributed to humans. **Generative AI or AI-assisted technologies should not be used by reviewers to assist in the scientific review** of a paper as the critical thinking and original assessment needed for peer review is outside of the scope of this technology and there is a risk that the technology will generate incorrect, incomplete or biased conclusions about the manuscript. The reviewer is responsible and accountable for the content of the review report. Elsevier’s AI author policy states **that authors are allowed to use generative AI and AI-assisted technologies in the writing process before submission, but only to improve the language and readability of their paper and with the appropriate disclosure**, as per our instructions in Elsevier’s Guide for Authors. Reviewers can find such disclosure at the bottom of the paper in a separate section before the list of references. *Generative AI is a type of artificial intelligence technology that can produce various types of content including text, imagery, audio and synthetic data. Examples include ChatGPT, NovelAI, Jasper AI, Rytr AI, DALL-E, etc.”

Documentul de față oferă punctul de vedere al Consiliului Științific al Universității Babeș-Bolyai (CS-UBB) și servește drept (1) normă internă în activitățile de evaluare și analiză specifice CS-UBB, dar și (2) recomandare colegială pentru uz în afara CS-UBB.

Educația asupra utilizării IAGen

În societățile dezvoltate tehnologic cu care România este în parteneriat și/sau competiție, IAGen își consolidează de ceva vreme statutul de unealtă de rutină. Cunoașterea IAGen trebuie azi privită pe treaptă similară cu alte instrumente digitale precum cele din pachetele de software de tip Office. O persoană instruită în spiritul modern al înțelegerii generării procesului de cunoaștere (așadar, familiară cu elementele de bază ale proceselor de tip cercetare-dezvoltare-inovare, CDI) va avea fie de utilizat IAGen fie de lucrat în colaborare sau competiție cu utilizatori IAGen. Este esențial ca studenții și cercetătorii / cadrele didactice să fie familiari cu elementele de bază ale acestor tehnologii – respectiv cu capacitățile (în continuă evoluție accelerată), limitările și standardele de uz responsabil specifice domeniului științific

în care aceștia/acestea activează. Module de instruire/ seminar / laborator sunt recomandate astfel în toate domeniile, la toate facultățile / institutele. Amploarea și substanța acestor module trebuie adaptată specificului domeniului către care se adresează - de la domenii vocaționale în care IAGen poate fi un element central al aptitudinilor profesionale la unele zone de științe exacte experimentale unde IAGen poate fi o unealtă ocazională de corectare de text (de tip „spellchecker”).

Cu viteze diferite în funcție de domeniu/facultate, IAGen înlocuiește sau face redundante etape sau elemente / procese anterior inerente activității în domeniu – procese care în unele cazuri aveau chiar specialiști dedicați. În paralel, IAGen poate aduce apariția de sub-specializări noi. Atât instructorii cât și cei instruiți trebuie să fie la curent și pregătiți pentru aceste schimbări.

IAGen în generarea de cunoaștere- CDI

În domeniile vocaționale din zona artelor, există deja o direcție de activitate în care IAGen este un instrument *principal* de activitate – fie că e vorba de generarea de muzică, de text sau de imagine, fie de dezvoltarea algoritmilor care le fac pe acestea posibile. Diverse alte domenii umaniste sau sociale au deja ca standard de activitate IAGen nu doar în mediul academic ci mai ales în afara lui. În unele domenii experimentale și exacte, în special acolo unde se operează cu volume mari de date, sau la generarea de cod în limbaje de programare, IAGen este de asemenea un instrument standard. Dimpotrivă, în alte domenii IAGen este privită ca o unealtă opțională de tip „autocorrect” sau „spellchecker”, dar nu una centrală sau măcar inerentă generării de cunoaștere / proceselor de tip CDI. Este esențial de reținut că IAGen au fost demonstrate a furniza uneori afirmații, teorii, date și referințe bibliografice complet ficționale. De aceea, documentarea pe subiecte științifice cu ajutorul IAGen poate fi riscantă dacă se utilizează orbește, fără cunoașterea răspunsurilor în prealabil și/sau fără documentarea și verificarea lor ulterioară.

IAGen în scrierea academică

Cercetători și instituții din primele rânduri de la frontiera cunoașterii folosesc azi IAGen în scrierea de propuneri de granturi, rapoarte, lucrări științifice etc – în mod responsabil și simplificând / eficientizând munca. O astfel de utilizare responsabilă implică două coordonate: (1) evitarea/prevenirea generării de informații/cunoștințe false prin abuzarea capacității generative a IAGen și (2) recunoașterea transparentă a utilizării IAGen. Pe de altă parte, deficiențele de educație sau de standarde morale / profesionale duc de asemenea la utilizarea pe scară largă a IAGen pentru fraudă academică – acolo unde autorii nu dau credit utilizării IAGen sau chiar folosesc IAGen pentru a acoperi alte elemente de fraudă, precum generarea de date false sau plagiatul. Această problemă este în special stringentă în situațiile în care obiectivul activității respective de scriere era demonstrarea abilităților de raționament / sinteză / critică /etc în cadrul unui proces de instruire sau examinare. Utilizarea responsabilă și eficientă a IAGen e imposibilă dacă utilizatorii nu știu ce anume încearcă IAGen să mimeze/simplifice.

În concluzie, este inevitabil să avem două contexte de utilizare a IAGen: (1) cele în care sarcina centrală este testarea/instruirea unor elemente umane ce țin de originalitate, raționament, analiză critică etc (și unde IAGen poate fi un instrument secundar dar nu unealta principală – deci folosirea sa prezintă potențial de fraudă din partea persoanelor examinate) și (2) cel unde persoanele deja instruite folosesc IAGen pentru a simplifica sarcini ce pentru ele au devenit deja de rutină.

Cuvântul „rutină” din fraza anterioară este unul esențial. Cercetarea științifică responsabilă știe diferența între rutină și originalitate – și să le dea fiecareia statutul potrivit. Dacă azi IAGen poate scrie o

sinteză pe un subiect științific, atunci fie (1) sinteza aceea e relativ de rutină și nu e cazul ca un om să facă efortul de a o scrie în manieră clasică, fie (2) e o sinteză incomplet instruită și prin urmare greșit făcută. În ambele situații, nu este acceptabil ca textele respective să fie puse pe trepte egale cu texte care îndeplinesc cerința esențială a științei – de a împinge frontierele cunoașterii dincolo de rutină.

IIAGen, GDPR și responsabilitatea

IIAGen folosesc în general platforme de calcul și stocare a datelor/informațiilor localizate în afara jurisdicției în care operează de obicei utilizatorii standard din mediul academic. Cerințele, documentele și datele utilizatorilor sunt transferate așadar către entități externe cu limitări legislative de multe ori neclare utilizatorilor. De acolo, acele informații pot circula fie în spatele unor uși închise, fie indirect către alți utilizatori în texte/creații ulterioare ale aceleiași IIAGen. Pentru majoritatea aplicațiilor, acest lucru este irelevant, fiind vorba de date/informații triviale. Totuși, când vine vorba de date cu caracter reglementat GDPR, sau cu caracter confidențial/secret de altă natură, în special altele decât cele ale utilizatorului/operatorului, acest element devine problematic – utilizatorul riscând să încalce regulile de confidențialitate sau chiar legile locale.

Abordarea CS-UBB

CS-UBB gestionează ocazional competiții de proiecte pentru finanțarea cercetării. În aceste competiții, CS-UBB caută excelența și generarea de cunoaștere de frontieră într-o manieră care să evite birocrăția excesivă – cu alte cuvinte să recompenseze generatorii de cunoaștere și să le răpească în acest proces cât mai puțin din timp. Cererile de finanțare la CS-UBB sunt gândite deci pe coordonata excelenței și eficienței – concise și bogate în *substanță*. În acest spirit, contribuțiile de rutină ale IIAGen nu își au locul. **CS-UBB își asumă să nu ceară și să nu stimuleze cereri de grant unde IIAGen este folosită la scriere.** Desigur, IIAGen rămâne legitimă în aceste granturi ca unealtă sau subiect de cercetare în sine – iar capacitatea de a folosi IIAGen în scrierea academică nu este în niciun fel descurajată. De asemenea, folosirea IIAGen în texte generice sau rezumative – inclusiv de popularizare a științei - poate fi acceptabilă. Utilitatea IIAGen pentru subiecte științifice de nișă poate varia foarte mult în funcție de sursă/program. Astfel, cele mai general utilizate IIAGen sunt antrenate/calibrate pe seturi de date/noțiuni foarte generale la nivel global, astfel pot fi problematice sau insuficient adaptate folosirea într-un domeniu îngust de cunoaștere sau către un sector îngust/local de public. Pe de altă parte, unele IIAGen sunt bine antrenate pe domenii înguste. Mai mult decât atât, există deja implementată capacitatea unor IIAGen (de pilda chatGPT) de a genera text pe baza unor documente integrate de către utilizator, deci capacitatea de a furniza afirmații corecte este mult mai ridicată. În oricare variantă, indiferent de IIAGen folosită, afirmațiile trebuie trecute / verificate prin lentila cercetătorilor care decid să folosească acele afirmații.

Generarea de text pentru procesele de evaluare este azi fezabilă azi; CS-UBB va aplica aici aceleași principii ca pentru scrierea de granturi.

Recomandări

1. În opinia CS-UBB, **publicațiile științifice (și nu numai) trebuie să diferențieze clar între conținutul generat nemijlocit de către om și conținutul generat cu IIAGen.** Ambele trebuie ținute la standarde serioase și în niciun caz confundate. Ambele pot fi meritorii, dar fiecare cu utilitatea ei. În decizia de a scrie / publica, autorii trebuie să își asume conștient această

diferențiere. Nu trebuie lăsată să persiste iluzia că articole științifice scrise cu IAGen să fie apreciate la aceeași valoare și cu aceleași standarde ca și cele scrise înainte de era IAGen.

2. **Dacă un text de articol științific (în special la secțiunea introductivă sau de discuții) poate fi scris de către IAGen la fel de bine ca de un om, atunci probabil acel text nu este unul care generează cunoaștere și nu ar trebui prezentat drept „publicație științifică” ci mai degrabă „text de popularizare”.** Atunci când scriem un text științific destinat specialiștilor, fie el propunere de proiect/grant, articol, sau carte, regula trebuie să fie aceasta: la frontiera cunoașterii IAGen nu este antrenat să opereze, deci nu ar putea scrie acest text; dacă am scris însă acel text cu IAGen, înseamnă că am renunțat la ambiția de a ne plasa la frontierele cunoașterii, adică de a face CDI cu impact serios.
3. **IAGen este un subiect ce își poate găsi locul în cadrul educației pentru competențe digitale în toate domeniile științei.**
4. În epoca IAGen originalitatea și autoratul eseurilor și referatelor ca „teme pentru acasă” devin imposibil sau nesustenabil de controlat. Instrumentele de verificare a originalității textelor verifică azi nu doar plagiatul ci și contribuțiile IAGen – însă gradul de succes al acestor verificări este (și e de așteptat să rămână) nesatisfăcător. De aceea, recomandarea noastră este **să se renunțe la astfel de eseuri ca modalitate de evaluare** – sau, dacă nu se poate renunța, atunci să se recurgă la modalități suplimentare de examinare pe marginea eseurilor, precum (1) **discuție directă și interactivă** (nu o simplă „prezentare” unilaterală) pe tema/ideile eseului sau (2) **redactarea eseului ad hoc, în condiții de examen scris.**
5. **Lucrările de finalizare a studiilor** includ în multe domenii o secțiune de rezultate originale generate de către studenții autori ai acestor lucrări. Dacă ele sunt cu adevărat originale, atunci IAGen va avea un rol minim. Pe de altă parte toate lucrările includ o secțiune de „studiu de literatură” unde acoperirea/sumarizarea unor cunoștințe deja existente este ținta ideală pentru IAGen. Cu îndrumarea potrivită și efortul adecvat, se poate și aici redacta un text cu adevărat de frontieră, adică unul inaccesibil IAGen – însă o asemenea ambiție este nesustenabilă pentru învățământul în masă. Și aici, ca și în cazul eseurilor, soluția poate fi de a accepta posibilitatea utilizării IAGen și de a **muta accentul pe examenele de tip discuție care cer candidaților să demonstreze cunoașterea lucrurilor scrise în lucrarea de finalizare a studiilor.** Pentru nivelul licență, o soluție rezonabilă în unele domenii poate fi înlocuirea lucrării de finalizare a studiilor cu examene propriu-zise – cu observația că aptitudinile inerente unei lucrări de licență tradiționale (documentarea, redactare de texte sintetice, aplicarea noțiunilor teoretice la o problemă concretă complexă) ar trebui și atunci să se regăsească între țintele atinse și evaluate pe parcursul procesului educațional.
6. Propunem ca documentele academice de tip CDI (eseu, articol, disertație, propunere de grant etc) să fie clasificate în funcție de gradul de utilizare a IAGen:
 - a) **IAGen-S**, tipic pentru standardele revistei Science sau cele pentru evaluatorii Elsevier citate la începutul documentului de față, adică fără utilizarea IAGen sau cel mult cu instrumente de rutină de tip corector de gramatică sau auto-completare din programele de procesare de text (ex., Word),
 - b) **IAGen-copilot** unde IAGen este folosit ca auxiliar pentru reformularea substanțială a textului sau unde o parte a textului este generată de la zero cu IAGen. Anexa 1 include recomandări suplimentare pentru IAGEN-S și IAGen-copilot.
 - c) **IAGen-1** unde textul este generat esențialmente cu IAGen.

7. Este de așteptat ca documentele de tip IIAGen-S să fie dificil de generat și gestionat în învățământul de masă, pentru că verificarea îndeplinirii cerințelor este aproape imposibil de făcut post-factum fără eroare; eseuri scrise ad-hoc în sala de examen, eventual de mână, se pot încadra în această categorie. Categoriile IIAGen-copilot și IIAGen-1 pot fi utile pentru anume discipline științifice. În funcție de specificul domeniului, recomandăm pentru lucrările de finalizare a studiilor categoria IIAGen-S, dar notăm de asemenea că în anumite domenii poate fi productiv să de folosească IIAGen-copilot. Insistăm că impunerea unui anumit standard trebuie adaptată specificului domeniului, experienței candidaților/autorilor și capacității de verificare din partea examinatorilor.
8. Acolo unde este posibil, regulamentele de examinare ar trebui să includă explicit cerința de identificare și/sau de încadrare într-una dintre cele 4 categorii descrise mai sus (S, copilot, sau 1) – și să aibă modalități realiste de verificare/implementare a oricăror limitări impuse autorilor în privința utilizării IIAGen. Similar și pentru alte contexte instituționale sub controlul UBB – spre exemplu revistele editate la UBB.
9. Permitearea accesului IIAGen la un document implică în general transferarea aceluși document către servere externe și accesul la document al (1) proprietarilor/creatorilor IIAGen și (2) indirect al potențialilor viitori utilizatori ai acelei IIAGen. De aceea, folosirea IIAGen cu documente confidențiale sau sensibile (ex., date și/sau de cercetare nepublicate) este de evitat sau interzis.

Referințe

- (1) Buriak, J. M.; Hersam, M. C.; Kamat, P. V. Can ChatGPT and Other AI Bots Serve as Peer Reviewers? *ACS Energy Lett.* **2023**, 191–192. <https://doi.org/10.1021/acseenergylett.3c02586>.
- (2) Shaw, C.; Yuan, L.; Brennan, D.; Martin, S.; Janson, N.; Fox, K.; Bryant, G. GenAI in Higher Education: Fall 2023 Update Time for Class Study. 2023. tytonpartners.com/time-for-class-2023/GenAI-Update.
- (3) Conroy, G. How ChatGPT and Other AI Tools Could Disrupt Scientific Publishing. *Nature* **2023**, 622 (7982), 234–236. <https://doi.org/10.1038/d41586-023-03144-w>.
- (4) Van Noorden, R. ChatGPT-like AIs Are Coming to Major Science Search Engines. *Nature* **2023**, 620 (7973), 258. <https://doi.org/10.1038/D41586-023-02470-3>.
- (5) *Using Generative AI for Your Scientific Writing? Be Aware of Journal Policies : Civil and Environmental Engineering Communication Lab.* <https://mitcommlab.mit.edu/cee/2023/08/27/using-generative-ai-for-your-scientific-writing-be-aware-of-journal-policies/> (accessed 2024-01-22).
- (6) *Science Journals: Editorial Policies | Science | AAAS.* <https://www.science.org/content/page/science-journals-editorial-policies#image-and-text-integrity> (accessed 2024-01-22).
- (7) *Publishing ethics | Elsevier policy.* <https://www.elsevier.com/about/policies-and-standards/publishing-ethics#0-publishing-ethics> (accessed 2024-01-22).

Anexă: recomandări pentru folosirea responsabilă ca IIAGen ca instrument copilot în lucrările de tip IIAGen-S ¹

PRINCIPII

1. Utilizarea responsabilă a IIAGen: se va folosi IIAGen doar pentru a îmbunătăți lucrărilor lor, dar nu pentru a înlocui sarcinile esențiale de autor, precum analiza conceptuală științifică
2. Promovarea originalității: se va evita copierea textelor generate exclusiv de către IIAGen și mai ales fără un input de structură în prealabil și verificare a materialului generat, asigurând originalitatea și autenticitatea lucrărilor astfel generate.
3. Responsabilitatea personală: autorii (ex., studenții și îndrumătorii în cazul lucrărilor de licență) vor purta întreaga responsabilitate pentru conținutul lucrărilor lor academice, asigurându-se că folosirea IIAGen este efectuată cu discernământ și integritate.
4. Control uman: tehnologia va fi aplicată cu supraveghere și control uman, iar toate lucrările vor fi revizuite și editate cu atenție, deoarece IIAGen poate genera rezultate cu formulări ce sugerează autoritatea derivată din cunoaștere, dar pot fi incorecte, incomplete sau părtinitoare. Studenții sunt în cele din urmă responsabili și răspunzători pentru conținutul lucrărilor lor.

CERINȚE

5. Citare: utilizarea IIAGen va fi declarată în lucrare. Această cerință este din aceeași categorie cu cele legate de explicitarea tehnologiilor și instrumente folosite în cercetare, precum instrumentele de analiză statistică, de codarea a interviurilor sau analiză de conținut. Toate instrumente vor fi citate, iar acolo unde acest lucru este relevant folosirea lor va fi discutată în secțiunea de metodologie. Autorii au responsabilitatea de a cita în lucrarea lor instrumentele folosite și deținătorii/creatorii acestora, fie umani, fie corporații (precum: Bard, Claude, LLaMA, PaLM, Jasp, Nvivo, Atlas-TI, plugin-uri de ChatGPT, MyGPTs, ChatGPT Data Analytics).
6. Responsabilitate auctorială umană: nu se vor lista IIAGen ca autori sau co-autori, nici nu se vor cita IIAGen ca autori. Vocea autorului implică responsabilități și sarcini care pot fi atribuite și efectuate doar de către oameni.
7. Interzicerea fabricării datelor: autorii vor evita folosirea IIAGen pentru generarea de date / conținuturi false, cu excepția cazurilor unde IIAGen este subiectul principal al cercetării. Aceasta este o cerință mai largă legată de probitatea profesională - de a nu inventa date în cercetare și de responsabilitatea autorilor de a-și face tot procesul de cercetare repetabil, adică științific.
8. Respectarea eticii și deontologiei: autorii vor respecta normele de etică și deontologie profesională în folosirea IIAGen, inclusiv protecția datelor personale. Cerința este mai largă, cea de a proteja subiecții și datele personale în cercetare.

RECOMANDĂRI

9. Informare și înțelegere: autorii vor căuta să se informeze amănunțit despre cum funcționează IIAGen și vor înțelege avantajele și limitele acestora înainte de utilizare.
10. Verificarea informațiilor: autorii vor asigura acuratețea datelor și informațiilor din lucrările lor, folosind surse verificabile și credibile. Acesta este un principiu mai larg, ce ține de utilizarea în argumentația științifică a altor surse științifice.

¹ Anexă preluată din practicile Facultății de Sociologie și Asistență Socială a UBB

11. Complementaritatea IIAgen: autorii pot folosi IIAgen pentru a obține înțelegerea unor concepte, sau o viziune de ansamblu, sau/și pentru a-și organiza ideile, dar vor parcurge și utiliza literatura de specialitate, argumentare logică și gândire critică pentru fundamentarea lucrării.
12. IIAgen ca instrument de descoperire a literaturii științifice: autorii pot folosi AI-ul pentru a identifica surse și autori relevanți, dar vor consulta și cita aceste surse în mod adecvat în cadrul lucrării lor.
13. Reformulare: autorii pot folosi IIAgen ca un instrument de îmbunătățire stilistică a lizibilității textului formulat de către ei și a limbajului academic folosit de aceștia.
14. AI ca sursă de inspirație: autorii pot utiliza formulările și stilul AI-ului ca punct de plecare pentru redactarea proprie, însă se vor asigura că lucrarea finală este rezultatul propriilor cunoștințe și cercetări. Preluarea textului formulat integral de către AI fără a marca clar acest lucru este o fraudă academică.