

## Extremely productive authors (EPA): alcune riflessioni sulla crescita dei ricercatori “estremamente produttivi”.

mauriziolana

The number of researchers pumping out more than 60 papers a year has almost quadrupled.

Con questa frase si apre un articolo pubblicato in Nature in dicembre scorso:

Conroy, Gemma. «Surge in Number of ‘Extremely Productive’ Authors Concerns Scientists». Nature 625, fasc. 7993 (4 gennaio 2024): 14–15. <https://doi.org/10.1038/d41586-023-03865-y>.

che riferisce di una ricerca pubblicata in bioRxiv:

Ioannidis, John P.A., Thomas A. Collins, e Jeroen Baas. «Evolving Patterns of Extremely Productive Publishing Behavior across Science». Preprint. Scientific Communication and Education, 24 novembre 2023. <https://doi.org/10.1101/2023.11.23.568476>.

Che cosa si intende con extremely productive author (EPA)? Un ricercatore “pumping out more than 60 papers a year”, cioè un ricercatore che pubblica 1 articolo ogni 6 giorni, weekend compresi (la ricerca è stata fatta nel database di Scopus).

Nell'articolo citato di Nature la discussione si sviluppa intorno agli extremely productive authors (EPA) della Thailandia perché è il paese che ne ha avuto il più alto tasso di crescita. Ma in numeri assoluti in testa c'è l'Arabia Saudita, seguita da vicino dall'Italia e poi dall'India:

### EXTREME GROWTH

Saudi Arabia had the highest number of extremely productive authors among the countries that have seen the fastest growth in the phenomenon. However, Thailand had the sharpest increase between 2016 and 2022.

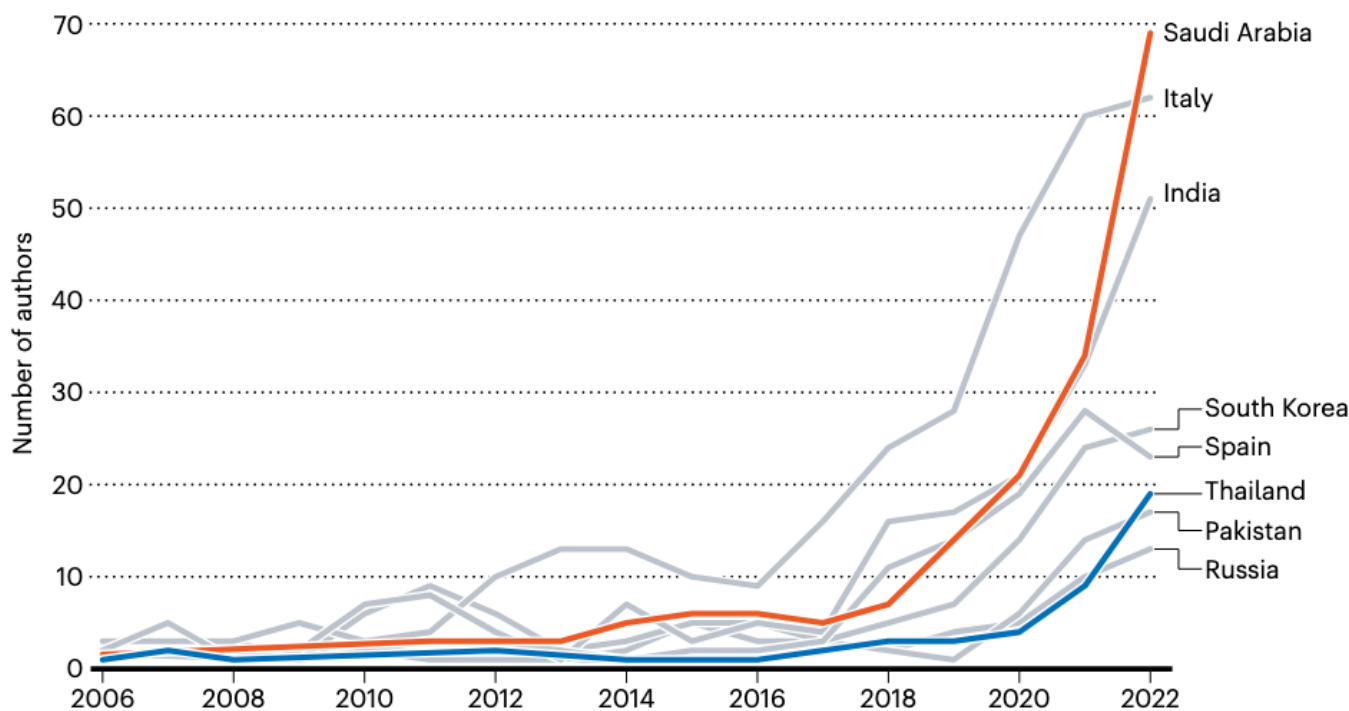


Grafico degli “Extremely Productive Authors”. Copyright Nature – doi: <https://doi.org/10.1038/d41586-023-03865-y>

Gli autori della ricerca pubblicata in bioRxiv scrivono che

most EPA authors were in Clinical Medicine, but from 2016 to 2022 the largest relative increases were 46 seen in Agriculture, Fisheries & Forestry (14.6-fold), Biology (13-fold), and Mathematics 47 and Statistics (6.1-fold)”.

Si potrebbe pensare di guardare con un certo distacco questi dati, anche se siamo uno dei paesi che hanno il più alto numero di EPA. Ma la questione sottesa, che riguarda tutti, è: come si fa a produrre ogni 6 giorni un articolo indicizzato? (cioè un articolo dignitosamente apprezzabile – dico “apprezzabile”, non “valido”)

Nell'articolo in Nature, Ioannidis (uno degli autori della ricerca) afferma: “I suspect that questionable research practices and fraud may underlie some of the most extreme behaviours” ma si deve ricordare che esistono campi disciplinari in cui sono frequenti articoli con decine di autori il che potrebbe parzialmente spiegare l'alta produttività.

Il fatto è che i sistemi di IA non solo potranno aiutare gli EPA a raggiungere il livello di hyperprolific authors cioè autori da 72 articoli all'anno, 1 ogni 5 giorni,

(cfr. Ioannidis, John P. A., Richard Klavans, e Kevin W. Boyack. «Thousands of Scientists Publish a Paper Every Five Days». Nature 561,

fasc. 7722, settembre 2018: 167–69. <https://doi.org/10.1038/d41586-018-06185-8>)

ma anche potrebbero aiutare tutti noi a scrivere di più.

Li useremmo, non li useremmo? La scelta costituirebbe un vantaggio o uno svantaggio netto e totale? Oppure potrebbe portare sia vantaggi sia svantaggi? In quale ambito?

La differenza sottile ma netta è tra l'uso di sistemi che potenziano le nostre capacità intellettuali, sistemi che tutti noi usiamo da tempo grati a Douglas Engelbart (Engelbart, Douglas, e William K. English. «A research center for augmenting human intellect». In AFIPS Conference Proceedings of the 1968 Fall Joint Computer Conference, 33:395–410. San Francisco, 1968 <https://web.stanford.edu/dept/SUL/library/extra4/sloan/mousesite/Archive/ResearchCenter1968/ResearchCenter1968.html>) e sistemi che potenziano la capacità produttiva di scrittura scientifica, fino a sostituire l'autore.

A mia conoscenza l'unico vero esperimento dichiarato in questo ambito è ancora quello descritto in Osmanovic Thunström, Almira. «We Asked GPT-3 to Write an Academic Paper about Itself – Then We Tried to Get It Published». Scientific American, giugno 2022. <https://www.scientificamerican.com/article/we-asked-gpt-3-to-write-an-academic-paper-about-itself-then-we-tried-to-get-it-published/>, su cui poi gli autori hanno riflettuto in Osmanovic Thunström, Almira, e Steinn Steingrímsson. «Does GPT-3 Qualify as a Co-Author of a Scientific Paper Publishable in Peer-Review Journals According to the ICMJE Criteria? A Case Study». Discover Artificial Intelligence 3, fasc. 1 (6 aprile 2023): 12. <https://doi.org/10.1007/s44163-023-00055-7>.

Si pongono qui 2 metaquestioni:

- 1) se una tecnologia esiste, la si usa? senza limitazioni o con limitazioni? (l'esempio estremo è la bomba atomica);
- 2) fino a che punto un ricercatore può lecitamente/eticamente/(legalmente?) potenziare le sue capacità produttive? l'uso di sistemi di IA nella scrittura degli articoli scientifici è il doping della ricerca? sì, no, in quali forme, sempre, mai?

Un recente post di Gary Marcus, pubblicato nel suo blog, «The exponential enshittification of science» (Substack newsletter. Marcus on AI (blog), 15 marzo 2024. <https://garymarcus.substack.com/p/the-exponential-enshittification>) parla della presenza di “frammenti di un discorso GPToso” (spero che Barthes mi perdoni!) in articoli scientifici di recente pubblicazione, cioè articoli i cui autori hanno usato GPT per scrivere l'articolo e nel testo sono rimaste formulazioni come “I'm very sorry, but I don't have access to real-time information or patient-specific data, as I am an AI language model. I can provide general information about managing hepatic artery, portal vein, and ...” oppure “Certainly, here is a list of both Geo-natural and geosynthetic materials commonly used in civil engineering and environmental applications.” “Certainly, here is information about the process of a clinical trial application presented in a point-wise format:” “Certainly, here is a concise summary of the provided sections:”

Se questi autori fossero stati più attenti, avrebbero cancellato le tracce visibili dell'uso di un sistema di IA generativa nella scrittura dell'articolo (non sappiamo se per scrivere l'articolo) , e certamente altri lo sono stati. Sicché è ragionevole pensare che in tutti gli ambiti disciplinari ci siano articoli scritti con/da un sistema di IA senza che tale intervento sia dichiarato.

In Gennaio 2023 si erano espresse su “GPT autore” due riviste top level:

Nature aveva pubblicato un articolo dal titolo chiarissimo:

Stokel-Walker, Chris. «ChatGPT Listed as Author on Research Papers: Many Scientists Disapprove». Nature, 18 gennaio 2023. <https://doi.org/10.1038/d41586-023-00107-z>.

Science aveva aggiornato la sua policy editoriale

Thorp, H. Holden. «ChatGPT Is Fun, but Not an Author». Science 379, fasc. 6630 (27 gennaio 2023): 313–313. <https://doi.org/10.1126/science.adg7879>.

dichiarando che non avrebbe accettato articoli con GPT autore:

For years, authors at the Science family of journals have signed a license certifying that “the Work is an original” (italics added). For the Science journals, the word “original” is enough to signal that text written by ChatGPT is not acceptable: It is, after all, plagiarized from ChatGPT. Further, our authors certify that they themselves are accountable for the research in the paper. ...

We are now updating our license and Editorial Policies to specify that text generated by ChatGPT (or any other AI tools) cannot be used in the work, nor can figures, images, or graphics be the products of such tools. And an AI program cannot be an author. A violation of these policies will constitute scientific misconduct no different from altered images or plagiarism of existing works.

Sempre in Science il tema era stato ripreso in Settembre 2023 (Van Noorden, Richard, e Jeffrey M. Perkel. «AI and Science: What 1,600 Researchers Think». Nature 621, fasc. 7980, 27 settembre 2023: 672–75. <https://doi.org/10.1038/d41586-023-02980-0>)

Le posizioni espresse sono sfaccettate e complesse e sarebbe troppo lungo esporle qui: è meglio leggere direttamente l'articolo.

La policy attualmente in vigore in Science (<https://www.science.org/content/page/science-journals-editorial-policies#authorship>) si apre con queste righe:

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.

E poi ricorda che l'uso di sistemi di IA rompe l'accordo di riservatezza ed esclusiva che l'autore stipula con l'editore. Sono giusto alcuni spunti di riflessione sul fatto che c'è già chi ha messo chatGPT al lavoro per i suoi articoli scientifici. Questo ovviamente aumenta la produttività, e aiuta a diventare extremely productive authors.

Molto, ovviamente, dipende dai reviewer. Ma i reviewer che cosa possono fare? Quali strumenti o criteri stabilmente affidabili si possono adottare per individuare testi artificiali? Attualmente i tool per riconoscere testo prodotto da sistemi di AI non hanno performance soddisfacenti.

---

OpenEdition vi suggerisce di citare questo post nel modo seguente:

mauriziolana (30 Marzo 2024). Extremely productive authors (EPA): alcune riflessioni sulla crescita dei ricercatori “estremamente produttivi”. *Informatica Umanistica e Cultura Digitale: il blog dell' AIUCD*. Recuperato il 12 Dicembre 2024 da <https://doi.org/10.58079/w4ye>

---

- 
- 
- 
-