

Bias in AI and the implications on diversity, equity and inclusion

Leo Anthony Celi MD MS MPH
Massachusetts Institute of Technology
Beth Israel Deaconess Medical Center
Harvard T.H. Chan School of Public Health



Diversity, equity and inclusion: its implications on bias in AI

Leo Anthony Celi MD MS MPH
Massachusetts Institute of Technology
Beth Israel Deaconess Medical Center
Harvard T.H. Chan School of Public Health



Presenter Disclosure

- National Institute of Health
- Philips
- Google
- Amazon
- Massachusetts Life Sciences Center



OpenAI announces ChatGPT successor GPT-4

🕒 1 day ago





GPT-4 is OpenAI's most advanced system, producing safer and more useful responses

We're excited to see how people use GPT-4 as we work towards developing technologies that empower everyone.

2.4 Harms of representation, allocation, and quality of service

Language models can amplify biases and perpetuate stereotypes.[40, 41, 42, 43, 44, 45, 46, 6] Like earlier GPT models and other common language models, both GPT-4-early and GPT-4-launch continue to reinforce social biases and worldviews.

As GPT-4 and AI systems like it are adopted more widely in domains central to knowledge discovery and learning, and as use data influences the world it is trained on, AI systems will have even greater potential to reinforce entire ideologies, worldviews, truths and untruths, and to cement them or lock them in, foreclosing future contestation, reflection, and improvement.[47, 48, 45, 49] In

then fine-tuned using Reinforcement Learning from Human Feedback (RLHF) [34]. Given both the competitive landscape and the safety implications of large-scale models like GPT-4, this report contains no further details about the architecture (including model size), hardware, training compute, dataset construction, training method, or similar.

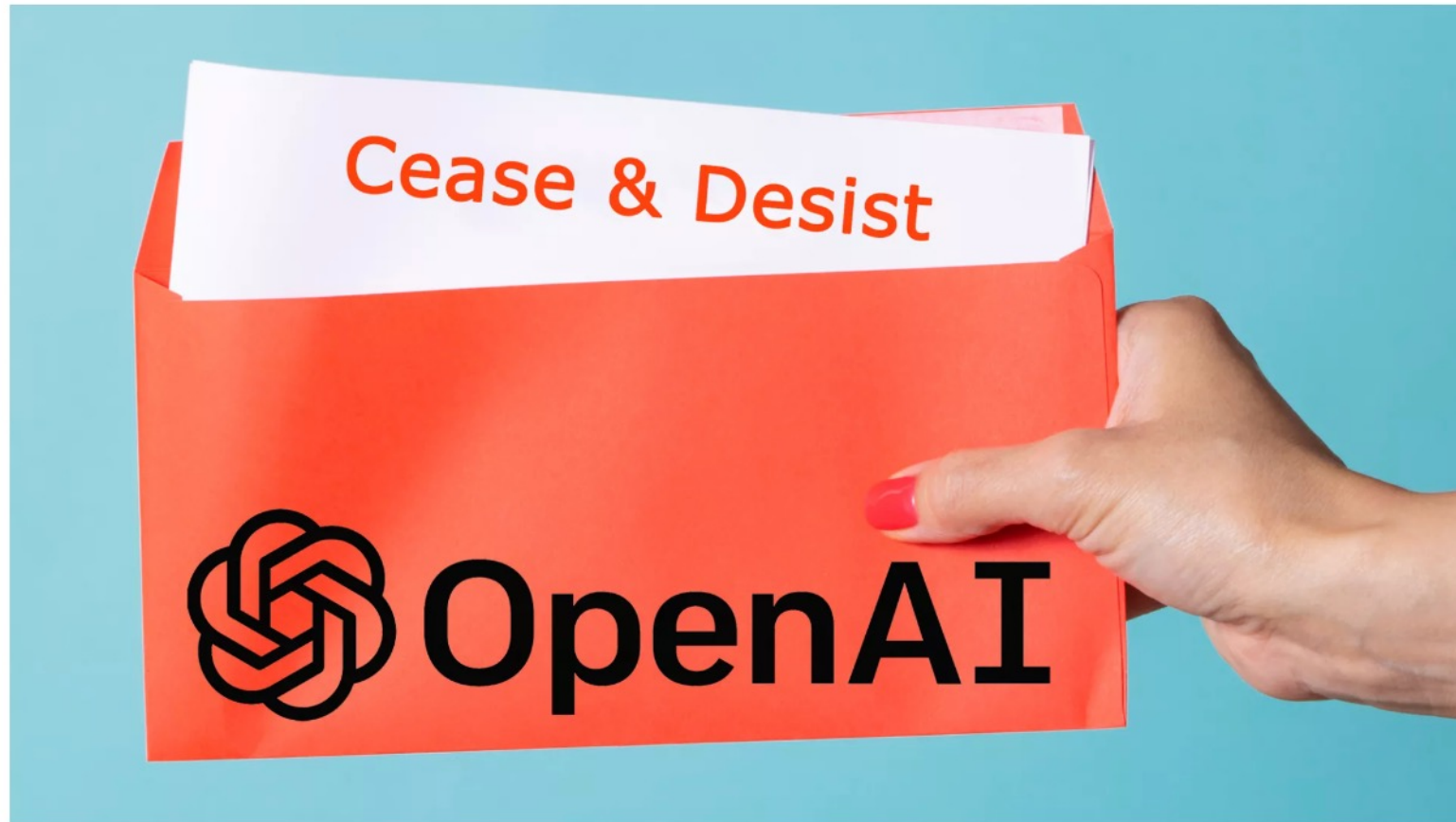


OpenAI Threatens Popular GitHub Project With Lawsuit Over API Use

By [Avram Piltch](#) last updated 4 days ago

GPT4Free uses other sites' connections to OpenAI.

       Comments (20)



DR. GPT WILL SEE YOU NOW—

GPT-4 will hunt for trends in medical records thanks to Microsoft and Epic

Generative AI promises to streamline health care, but critics say not so fast.

BENJ EDWARDS - 4/18/2023, 4:14 PM



TRENDING TOPICS

What is Auto-GPT



Microsoft plans to sell 'private' ChatGPT to businesses

TIMESOFINDIA.COM / May 3, 2023, 18:09 IST

 SHARE



AA



MIT News

ON CAMPUS AND AROUND THE WORLD

 [SUBSCRIBE](#)

An artificial intelligence tool that can help detect melanoma

Using deep convolutional neural networks, researchers devise a system that quickly analyzes wide-field images of patients' skin in order to more efficiently detect cancer.



The Markup

Big Tech Is Watching You. We're Watching Big Tech.

Machine Learning

**Major Universities Are Using Race as
a “High Impact Predictor” of
Student Success**



Appriss, Inc. Closes Sale of Appriss Insights, LLC to Equifax in \$1.825 Billion Deal

Appriss continues to focus on delivering *Knowledge For Good*® through innovative data and analytics solutions in healthcare and retail verticals.



Economy | Technology

White House tells tech CEOs they have 'moral duty' on AI



Section 1557 of the Patient Protection and Affordable Care Act

The Office for Civil Rights (OCR) enforces Section 1557 of the Affordable Care Act (Section 1557), which prohibits discrimination on the basis of race, color, national origin, age, disability, or sex (including pregnancy, sexual orientation, gender identity, and sex characteristics), in covered health programs or activities. 42 U.S.C. 18116.




JOINT STATEMENT ON ENFORCEMENT EFFORTS AGAINST DISCRIMINATION AND BIAS IN AUTOMATED SYSTEMS

*Rohit Chopra, Director of the Consumer Financial Protection Bureau,
Kristen Clarke, Assistant Attorney General for the Justice Department's Civil Rights Division,
Charlotte A. Burrows, Chair of the Equal Employment Opportunity Commission, and
Lina M. Khan, Chair of the Federal Trade Commission
issued the following joint statement about enforcement efforts to protect the public
from bias in automated systems and artificial intelligence:*

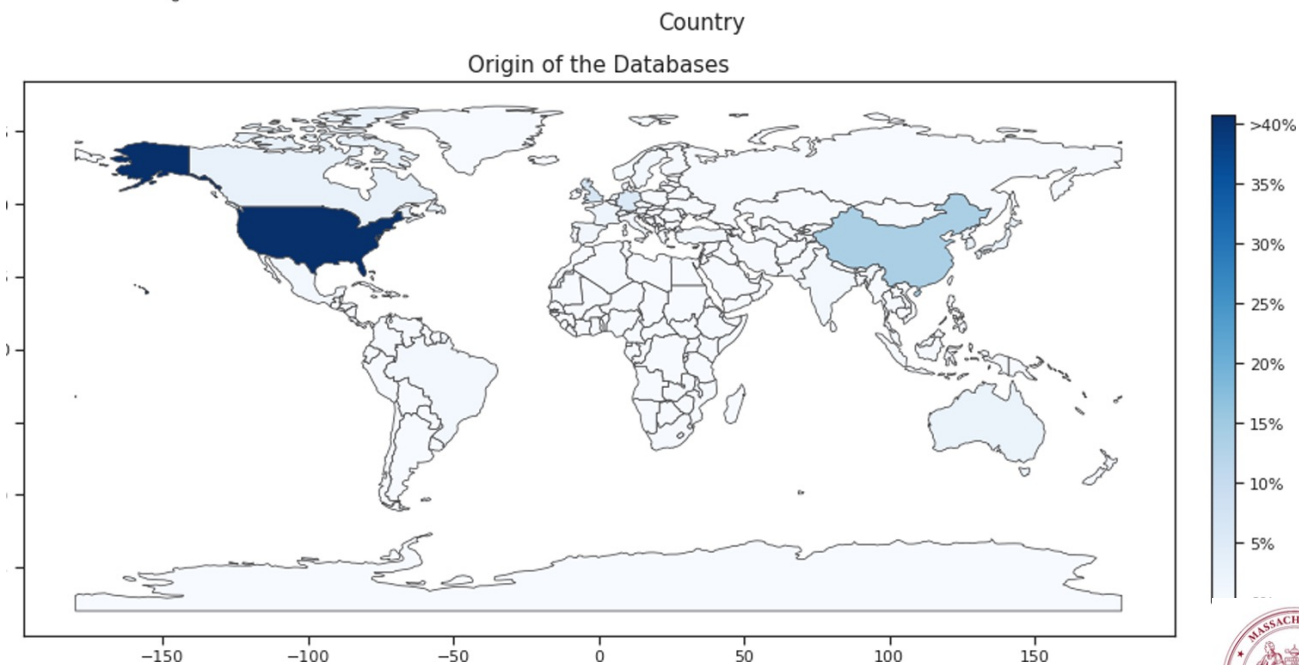
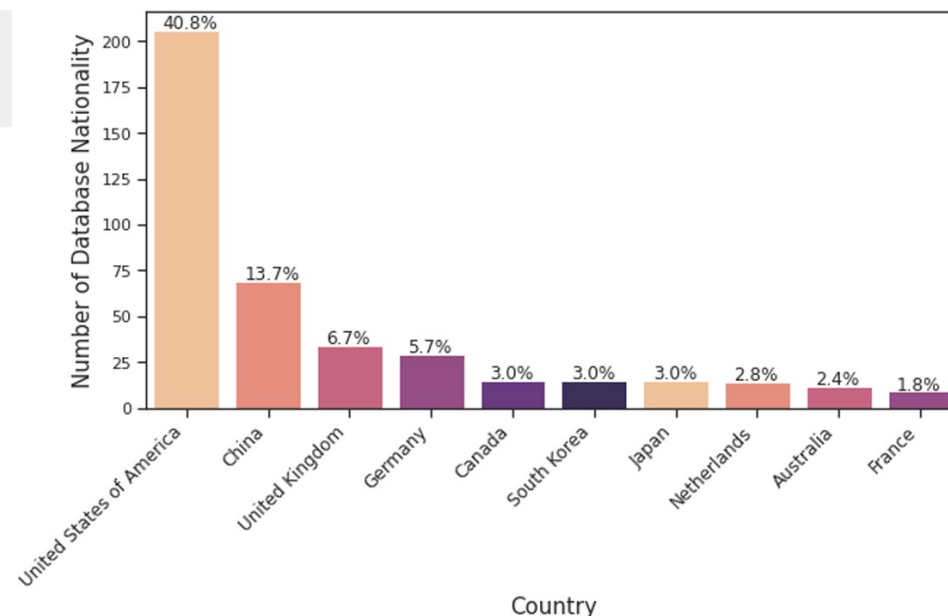




Sources of bias in artificial intelligence that perpetuate healthcare disparities—A global review

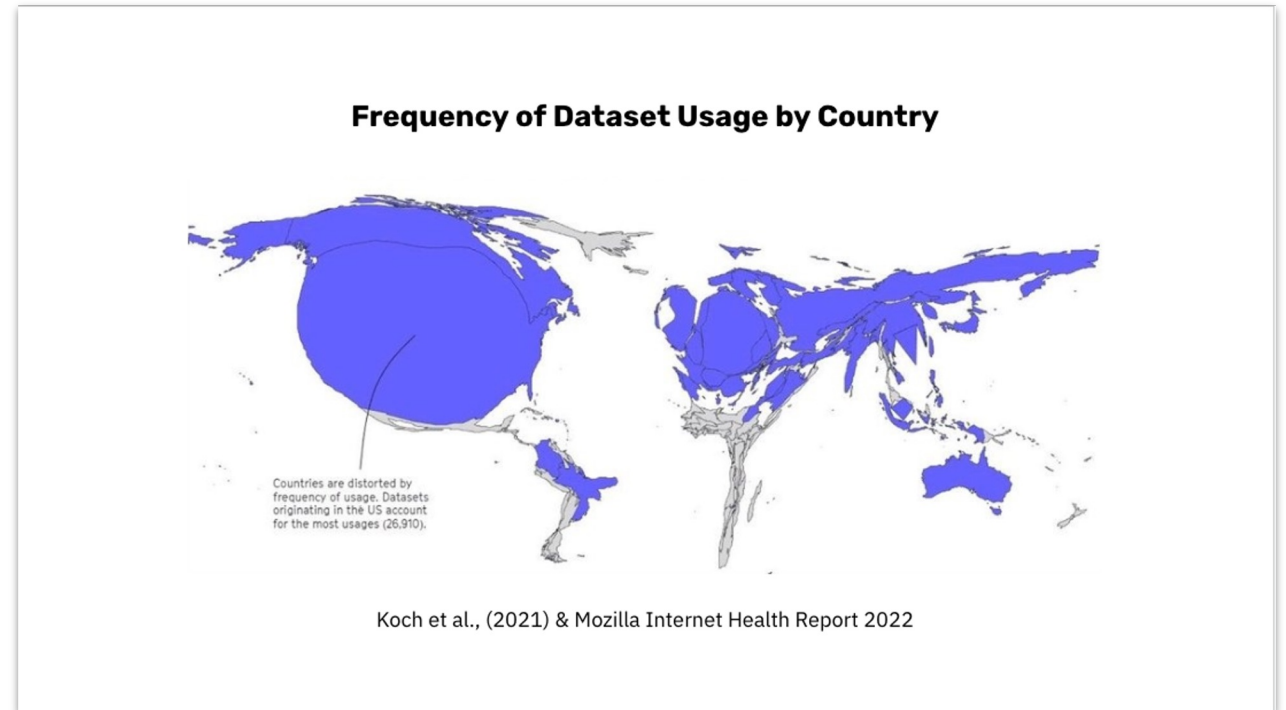
Leo Anthony Celi, Jacqueline Cellini, Marie-Laure Charpignon, Edward Christopher Dee, Franck Deroncourt, Rene Eber, William Greig Mitchell , Lama Moukheiber, Julian Schirmer, Julia Situ, Joseph Paguio, Joel Park, Judy Gichoya Wawira, Seth Yao, for MIT Critical Data

Published: March 31, 2022 • <https://doi.org/10.1371/journal.pdig.0000022>



The Medical Knowledge Landscape

- Where does the research informing clinical guidelines come from?
- How representative are cohorts used to train models?
- Who are given the platform to impact health research?
- What groups are being funded?
- **Can we measure and track this knowledge landscape?**



Gatekeepers of Medical Knowledge

- Researchers decide the methodology of knowledge collection
- Journals decide what research reaches the masses
- Funders & Industry decide what research are conducted
- Universities determine who have a seat at the table
- **Key players - journals, investigators, funders - should be evaluated for their impact on the knowledge ecosystem.**



Countries contributing to clinical-trial data

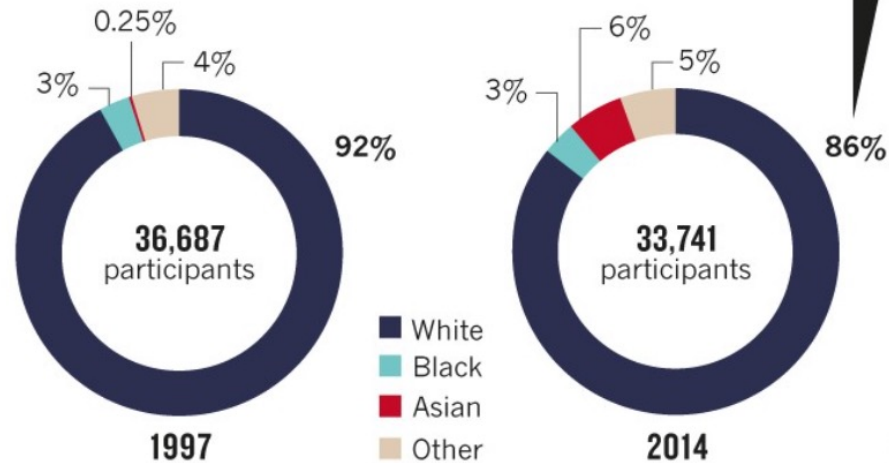
■ Enrolled participants in both years*
 ■ Enrolled participants in 2014
 ■ No enrolment



Asia has seen the most significant increase in numbers of participants enrolling in trials over this time period.

*Ireland didn't enrol patients in 2014.

Ethnicity of participants in clinical trials



Despite the fraction of Asian participants increasing, most participants are still white.

©nature

Source: US Food and Drug Administration



Case Study: Sepsis Literature



48.9 MILLION CASES
11 MILLION DEATHS

1 IN EVERY 5
DEATHS WORLDWIDE
ARE ASSOCIATED
WITH SEPSIS

85% OCCUR IN
LOW- OR
MIDDLE-INCOME
COUNTRIES

2 OUT OF
EVERY 5 CASES
ARE IN CHILDREN
UNDER 5



Recommendation	Studies (n)	High n (%)	Upper Middle n (%)	Lower-Middle n (%)	Low n (%)
Screening & Initial Treatment	30*	23 (77)	5 (17)	3 (10)	4 (13)
Infection	126*	114 (90)	16 (13)	10 (8)	8 (6)
Hemodynamic Management	47*	42 (89)	3 (6)	4 (9)	2 (4)
Ventilation	41*	39 (95)	9 (22)	0 (0)	0 (0)
Additional Therapies	36*	33 (92)	3 (8)	2 (6)	0 (0)
Long Term Outcomes/ Goals of Care	83	82 (99)	1 (1)	0 (0)	0 (0)



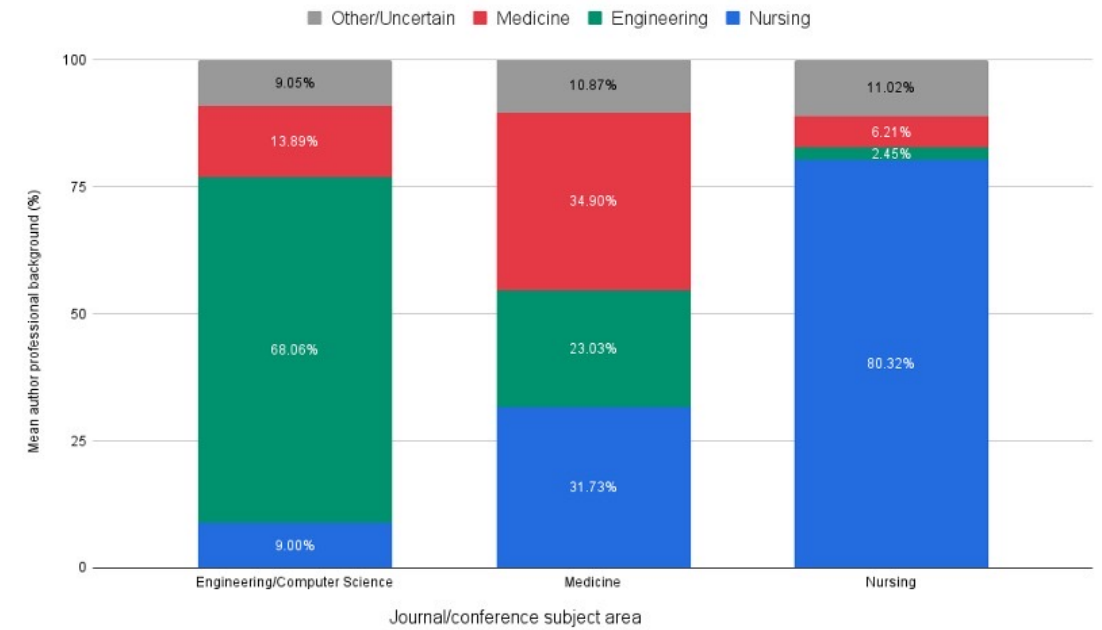
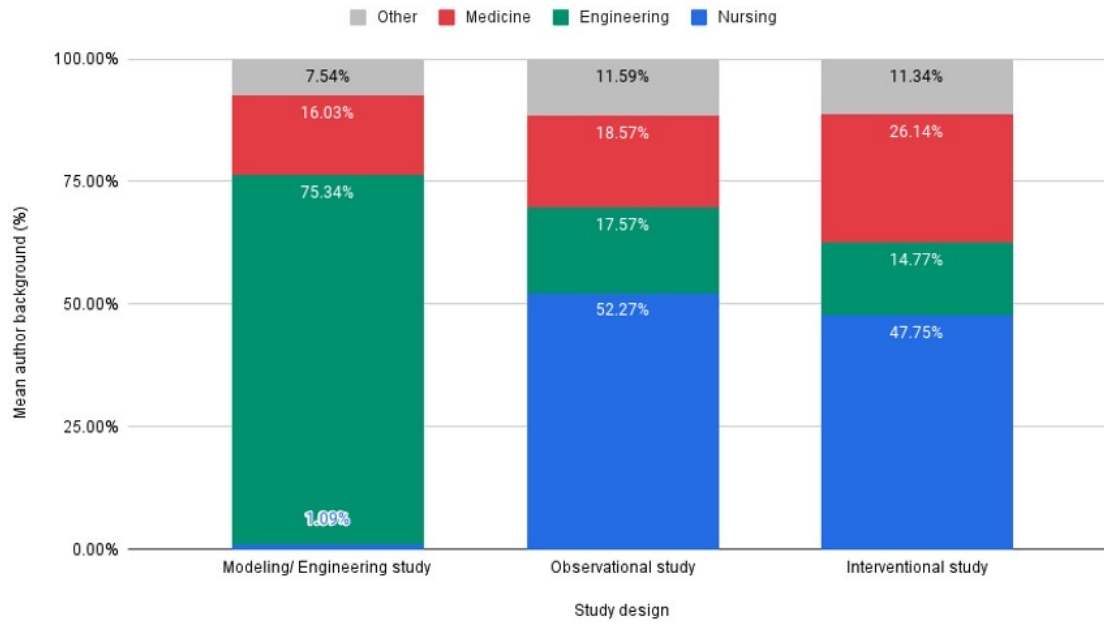
Recommendation	Studies (n)	East Asia Pacific n (%)	Europe Central Asia n (%)	Latin America Caribbean n (%)	Middle East North Africa n (%)	North America n (%)	South Asia n (%)	Sub-Saharan Africa n (%)
Screening & Initial Treatment	30*	2 (7)	10 (33)	4 (13)	1 (3)	14 (47)	0 (0)	4 (13)
Infection	126*	25 (20)	55 (44)	12 (10)	9 (7)	54 (43)	3 (2)	8 (6)
Hemodynamic Management	47*	9 (19)	16 (34)	1 (2)	3 (6)	25 (53)	0 (0)	1 (2)
Ventilation	41*	8 (20)	27 (66)	10 (24)	5 (12)	19 (46)	0 (0)	1 (2)
Additional Therapies	36*	3 (8)	19 (53)	2 (6)	2 (6)	13 (36)	2 (6)	0 (0)
Long Term Outcomes/ Goals of Care	83	7 (8)	31 (37)	0 (0)	0 (0)	46 (55)	0 (0)	0 (0)



Nurse-engineer collaboration in critical care alarm research: a systematic bibliometric analysis

Louis Agha-Mir-Salim, Lucas McCullum, Enrico Dähnert,
Yanick-Daniel Scheel, Ainsley Wilso4, Marianne Carpio,
Carmen Chan, Claudia Lo, Lindsay Maher, Corinna Dressler,
Felix Balzer, Leo Anthony Celi, Akira-Sebastian Poncette,
Michele M. Pelter





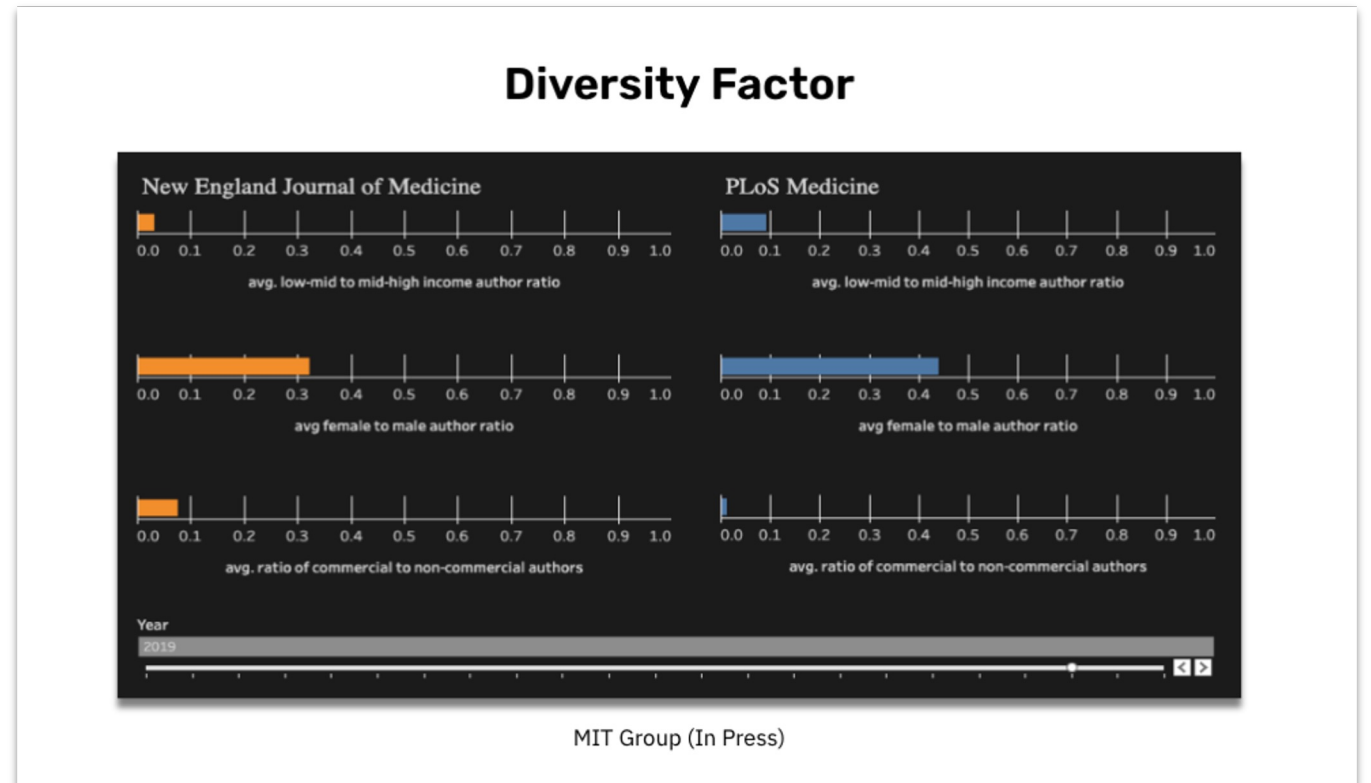
Models & Guidelines

- What patients are represented in studies that inform guidelines?
- What data has a model been trained and validated on?
- What is the gender and race-ethnicity balance of the cohorts?
- What are the demographics of the investigators? What communities do they represent?



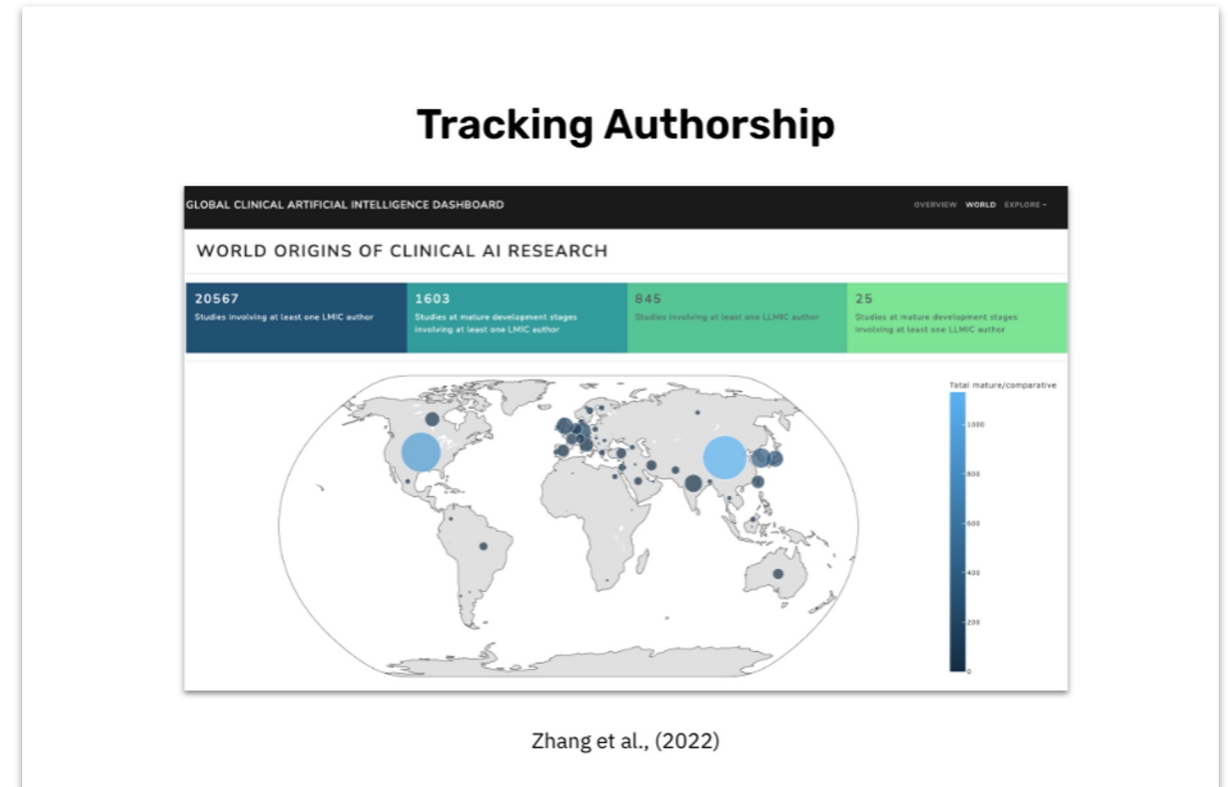
Journals

- What are the proportion of LMIC authors?
- Where is the data used in studies coming from?
- Is there a diversity of perspectives being platformed?
- What institutions are represented?
- How are they contributing to Equitable Science?²



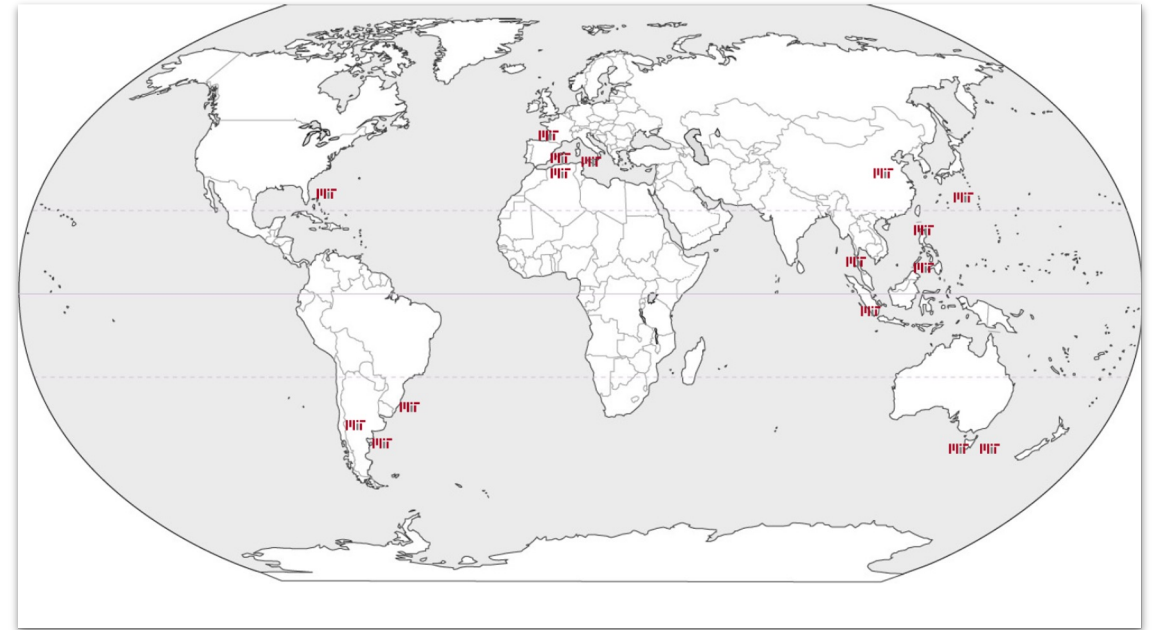
Authors

- Who do they work with in terms of co-authorship?
- Whose perspectives are represented in the papers they co-author?
- Is their research transparent and reproducible?
- Is research performed on a diverse cohort?
- Do they contribute resources such as datasets and repositories to the research community?



Start of the Journey: Who else should be on this table?

- Data used to inform models and guidelines need characterizing.
- The knowledge landscape needs mapping.
- This is an emerging field whose importance is only growing.
- **We need a diverse group passionate about revamping the knowledge ecosystem that will benefit everyone.**





**Every system
is perfectly
designed to
achieve
exactly the
results it gets**



[iScience](#). 2021 Jun 25; 24(6): 102656.

PMCID: PMC8209268

Published online 2021 Jun 10. doi: [10.1016/j.isci.2021.102656](https://doi.org/10.1016/j.isci.2021.102656)

PMID: [34169236](https://pubmed.ncbi.nlm.nih.gov/34169236/)

Village mentoring and hive learning: The MIT Critical Data experience

[Christopher V. Cosgriff](#),¹ [Marie Charpignon](#),² [Dana Moukheiber](#),³ [Mary E. Lough](#),⁴ [Judy Gichoya](#),⁵ [David J. Stone](#),⁶ and [Leo Anthony Celi](#)^{7,8}

▶ [Author information](#) ▶ [Copyright and License information](#) [Disclaimer](#)

