



Mind the AI Gap

Equity, Access & the Risks of Standing Still

Dr Philippa Hardman

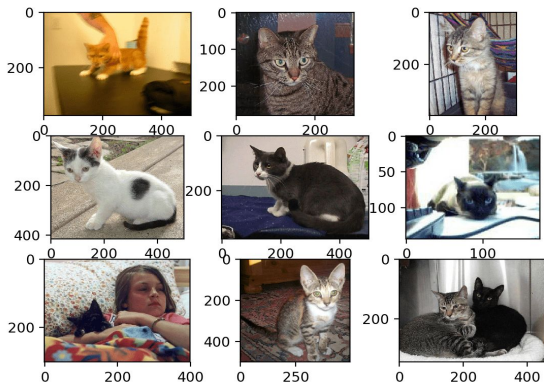


Risk 1: Bias

AI 101

Classification

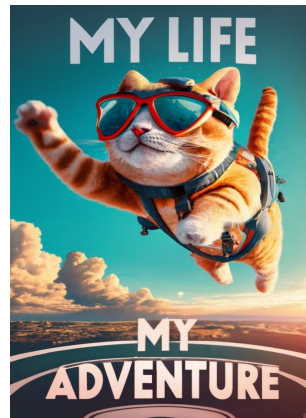
AI is trained to recognise patterns - "machine learning"



"This data pattern = cat"

Generation

AI can reproduce learned patterns on demand






"Generate me a poster for a movie about a cat living its best life"





OPEN

Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations

Laleh Seyyed-Kalantari^{1,2}  , Haoran Zhang³, Matthew B. A. McDermott³, Irene Y. Chen³ and Marzyeh Ghassemi^{2,3} 

Artificial intelligence (AI) systems have increasingly achieved expert-level performance in medical imaging applications. However, there is growing concern that such AI systems may reflect and amplify human bias, and reduce the quality of their performance in historically under-served populations such as female patients, Black patients, or patients of low socioeconomic status. Such biases are especially troubling in the context of underdiagnosis, whereby the AI algorithm would inaccurately label an individual with a disease as healthy, potentially delaying access to care. Here, we examine algorithmic underdiagnosis in chest X-ray pathology classification across three large chest X-ray datasets, as well as one multi-source dataset. We find that classifiers produced using state-of-the-art computer vision techniques consistently and selectively underdiagnosed under-served patient populations and that the underdiagnosis rate was higher for intersectional under-served subpopulations, for example, Hispanic female patients. Deployment of AI systems using medical imaging for disease diagnosis with such biases risks exacerbation of existing care biases and can potentially lead to unequal access to medical treatment, thereby raising ethical concerns for the use of these models in the clinic.

As artificial intelligence (AI) algorithms increasingly affect decision-making in society¹, researchers have raised concerns about algorithms creating or amplifying biases^{2–11}. In this work we define biases as differences in performance against, or in favor of, a subpopulation for a predictive task (for example, different performance on disease diagnosis in Black compared with white patients). Although AI algorithms in specific circumstances can potentially reduce bias¹², direct application of AI has also been shown to systematize biases in a range of settings^{3–7,13,14}. This tension is particularly pressing in healthcare, where AI systems could improve patient health¹ but can also exhibit biases^{2–7}. Motivated by the global radiologist shortage¹⁵ as well as by demonstrations that AI algorithms can match specialist performance particularly in medical imaging¹⁶, AI-based diagnostic tools present a clear incentive for real-world deployment.

Although much work has been done in algorithmic bias¹³ and bias in health^{2–11}, the topic of AI-driven underdiagnosis has been relatively unexplored. Crucially, underdiagnosis, defined as falsely claiming that the patient is healthy, leads to no clinical treatment when a patient needs it most, and could be harmful in radiology specifically^{17,18}. Given that automatic screening tools are actively being developed in research^{19–21} and have been shown to match specialist performance⁶, underdiagnosis in AI-based diagnostic algorithms can be a crucial concern if used in the clinical pipeline for patient triage. Triage is an important diagnostic first step in which patients who are falsely diagnosed as healthy are given lower priority for a

that AI can reduce underdiagnosis in general^{24,25} but these studies do not deeply consider the existing clinical biases in underdiagnosis against under-served subpopulations. For example, Black patients tend to be more underdiagnosed in chronic obstructive pulmonary disease than non-Hispanic white patients⁷.

Here, we perform a systematic study of underdiagnosis bias in the AI-based chest X-ray (CXR) prediction models, designed to predict diagnostic labels from X-ray images, in three large public radiology datasets, MIMIC-CXR (CXR)²⁶, CheXpert (CXP)²⁷ and ChestX-ray14 (US National Institutes of Health (NIH))²⁸, as well as a multi-source dataset combining all three on shared diseases. We focus our underdiagnosis study on individual and intersectional subgroups spanning race, socioeconomic status (as assessed via the proxy of insurance type), sex and age. The choice of these subgroups is motivated by the clear history, in both traditional medicine and AI algorithms, of bias for subgroups on these axes^{6,9,10,11}. An illustration of our model pipeline is presented in Fig. 1.

Results

A standard practice among the AI-based medical image classifiers is to train a model and report the model performance on the overall population regardless of the patient membership to subpopulations^{6,19–23}. Motivated by known differences in disease manifestation in patients by sex⁴, age²⁹, race/ethnicity⁶ and the effect of insurance type in quality of received care¹, we report results for all of these factors. We use insurance type as an imperfect proxy of

Risk 2: Hallucination



Risk 3: Academic Integrity



[GPTZero](#)
[Home](#)
[AI Detection](#)
[Resources](#)
[Team](#)
[Join us!](#)
[CONTACT SALES](#)
[GET STARTED](#)

GPTZero

The **Global Standard** for AI Detection
Humans Deserve the **Truth**

Detect ChatGPT, GPT3, GPT4, Bard, and other AI models. Try it for yourself 🌟

[GPT3](#)
[GPT4](#)
[CHATGPT](#)
[BARD](#)
[HUMAN](#)
[AI + HUMAN](#)

Type in the text you want to check for AI involvement (minimum 250 characters)

[turnitin](#)
Tessa Ruiz
The Goliath of the Sea
Details


How much of this submission has been generated by AI? 🔍

75%

of qualifying text in this submission has been determined to be generated by AI.

How do we detect AI-generated writing?

To learn more about Turnitin's AI writing detection model and how it works, please visit our [AI writing detection page](#).




Educator resources for AI writing

- Find strategies for [approaching AI-generated text in the classroom](#).
- Minimize potential AI misuse with our [AI misuse rubric](#) to review existing writing prompts for AI vulnerability, and our [AI misuse checklist](#) to review options to proactively respond to potential AI misuse in your classroom.
- Stay informed as Turnitin expands its [AI writing capabilities](#).

[Show Disclaimer](#)

The Goliath of the Sea

The majestic blue whale, the goliath of the sea, certainly stands alone within the animal kingdom for its adaptations beyond its massive size. At 30 meters (98 ft) in length and 190 tons (210 short tons) or more in weight, it is the largest existing animal and the heaviest that has ever existed. Despite their incomparable mass, aggressive hunting in the 1900s by whalers seeking whale oil drove them to the brink of extinction. But there are other reasons for why they are now so endangered.



Blue-Whale Balaenoptera Musculus

The blue whale's common name derives from bluish-hue that covers the upper side of its body, while its Latin designation is Balaenoptera musculus. The blue whale belongs to the Mysticeti suborder of cetaceans, also known as baleen whales, which means they have fringed plates of fingernail-like material, called baleen, attached to their upper jaws. Blue whales feed almost exclusively on krill, though they also take small numbers of copepods. An adult blue whale can eat up to 40 million krill in a day.

These gargantuan beasts used to dominate all the oceans of the Earth up until the late nineteenth century, when the technology was developed to effectively hunt and harvest them. In 1864, the Norwegian Svend Foyn equipped a steamboat with harpoons specifically designed for catching large whales. This led to the killing of hundreds of thousands of whales up until 1966, when the International Whaling Commission banned the practice.

The blue whale certainly appears grand in size and beauty, but the sounds it produces and how it communicates are also sublime. Amazingly, their vocalizations can reach 155 and 188 decibels and have a frequency range of 10 to 40Hz. Though they typically do not "sing" in the same way that Humpback whales do, some subspecies have been observed producing songs that consist of up to four notes. Blue whale calls are still not fully understood, and scientists are currently working on determining their purposes. Among the hypotheses, researchers believe the calls could serve to

“The expectation that teachers & students act with honesty, trust, fairness, respect & responsibility.”

AI Anti-Detection Technologies

The screenshot displays the Originality.ai interface. The top navigation bar includes the logo, 'Start New Scan', and a balance of 1312. The left sidebar lists various tools: Dashboard, Content Scanner, Start New Scan, My Content Scans, Tags, Website Scanner, My Team, and API Access. A 'Submit beta feedback' button is also present. The main content area shows a scan result for the text 'As individuals, we deify happiness.' with a 'Words: 375' count. The scan cost is 4 credits. The detection settings are: Detect AI (checked), Check Plagiarism & Readability (unchecked), and Fact Checking Aid (unchecked). The AI Detection Score is 0% Original and 100% AI. A large red circle highlights the score. The text being scanned is highlighted in red, indicating it was detected as AI-generated. The interface also includes a 'Scan Again' button and a 'Support' section with contact information.

Originality.ai Start New Scan Balance: 1312

Dashboard Content Scanner Start New Scan My Content Scans Tags Website Scanner My Team API Access

Submit beta feedback

Support: Help: help@originality.ai Email: support@originality.ai Logout

Scan Complete. Run Scan Again? Start New Scan Previous Scans

As individuals, we deify happiness. Highlight AI Text ✓ AI RESULTS PLAGIARISM

Paste from URL Paste from File Words: 375

Scan will cost 4 credits.

Detect AI Check Plagiarism & Readability Fact Checking Aid

Scan Again

While joy may be fleeting, true happiness acts as a profound catalyst stimulating positive transformation throughout our lives, for it is an emotion universally pursued as more than a moment but less than a permanent state. The benefits of being happy extend far beyond the mere feeling of joy; they permeate our mental, physical, and social well-being, making it a pivotal aspect of a fulfilling life. Mentally, happiness can significantly improve our cognitive functioning. Research has demonstrated that those who are cheerful frequently exhibit improved focus, more inventive thinking, and enhanced abilities to solve difficulties. Happiness, by reducing stress levels, enables the brain to work at its best by fostering a calm environment where mental dexterity can truly shine through. Happiness, by cultivating within us a positive outlook, sows the seeds for resilience and hope to take root and flourish throughout our days.

AI Detection Score

0% Original 100% AI

Originality.ai

This score reflects our AI's confidence in predicting that the content scanned was produced by any popular AI tool such as ChatGPT, GPT-4, Bard, Claude 2 etc. A score of 90% original and 10% AI should be thought of as "We are 90% confident that this content was created by a human" and NOT that 90% of the article is Human and 10% AI. Learn more about the score. [more](#) about the score.

Looking for our multi-language version? Select it from the dropdown below.



Ethan Mollick · Following

Associate Professor at The Wharton School

10mo ·



Teachers should be wary of AI detection tools for five reasons:

- 1) They aren't very good. For example, OpenAI's detector only identifies GPT output 26% of the time. (Also, never ask ChatGPT if something was created by ChatGPT - it will make up an answer, it doesn't know)
- 2) All the current tools were made for GPT-3.5 or earlier, not the newer GPT-4 models, which are much more capable
- 3) Even small changes to the text (including asking the AI to revise its own text) can break detectors. I actually had students in my class "cheat" on an assignment, and those who did any iteration at all fooled detectors.
- 4) There often are false positives. Even the OpenAI tool has a 9% false positive rate.
- 5) A new paper shows that, in the end, AIs can always beat detectors

Can AI-Generated Text be Reliably Detected?

Vinu Sankar Sadasivan
vinu@umd.edu

Aounon Kumar
aounon@umd.edu

Sriram Balasubramanian
sriramb@umd.edu

Wenxiao Wang
wxw@umd.edu

Soheil Feizi
sfeizi@umd.edu

Department of Computer Science
University of Maryland

Abstract

The rapid progress of Large Language Models (LLMs) has made them capable of performing astonishingly well on various tasks including document completion and question answering. The unregulated use of these models, however, can potentially lead to malicious consequences such as plagiarism, generating fake news, spamming, etc. Therefore, reliable detection of AI-generated text can be critical to ensure the responsible use of LLMs. Recent works attempt to tackle this problem either using certain model signatures present in the generated text outputs or by applying watermarking techniques that imprint specific patterns onto them. In this paper, both empirically and theoretically, we show that these detectors are not reliable in practical scenarios. Empirically, we show that *paraphrasing attacks*,

The AI-Education Divide



***“Longstanding failure to
create effective
frameworks for learning
technologies in state
funded education
systems.”***

Public school bans on AI tools like ChatGPT raise fears private school kids are gaining an unfair edge and widening a digital divide

By national education and parenting reporter [Conor Duffy](#)

Posted Thu 25 May 2023 at 10:17pm



Matt Esterman believes it is important to teach students what to watch out for when using AI. (ABC News: [Conor Duffy](#))

Top Stories

'You know there's a camera up there?': CCTV assault on handcuffed teen reveals youth justice problems

LIVE

Kokkinakis steps up looking for an upset, as Tomljanović resumes rivalry

Five quick hits — Travis's ton and Hazlewood's ridiculous spell on day two

Former ABC presenter Antoinette Lattouf vows to fight unlawful termination case 'as long as it takes'

Seven people killed as Pakistan launches retaliatory air strikes against Iran

'A special relationship': The 'old friend' Djokovic visits each year in Melbourne's Royal Botanic Gardens

Shock loss of 65,000 jobs before Christmas the biggest outside of

A Vision for AI in Education

**Equity of
Access
to AI Tools**

**Equity of Access to AI
Knowledge & Skills**

**Address Bias in Data
Collection & Ingestion**

Surface & Manage Data Bias



About ▼ Research Media ▼ Support

About us

We are an interdisciplinary and globally distributed AI research institute rooted in the belief that AI is not inevitable, its harms are preventable, and when its production and deployment include diverse perspectives and deliberate processes it can be beneficial. Our research reflects our lived experiences and centers our communities.



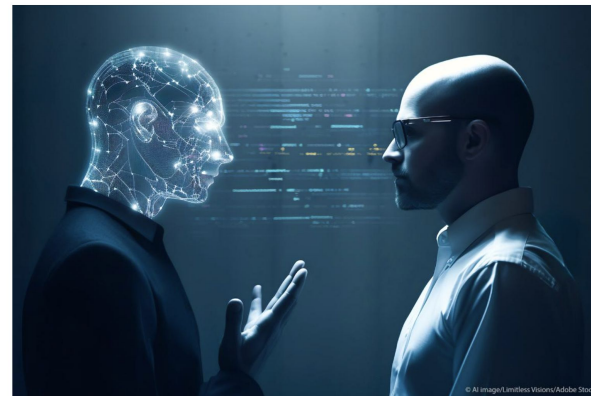
European Parliament

EU AI Act: first regulation on artificial intelligence

Society Updated: 19-12-2023 - 11:45
Created: 08-06-2023 - 11:40



The use of artificial intelligence in the EU will be regulated by the AI Act, the world's first comprehensive AI law. Find out how it will protect you.



This illustration of artificial intelligence has in fact been generated by AI

As part of its [digital strategy](#), the EU wants to regulate artificial intelligence (AI) to ensure better conditions for the development and use of this innovative technology. [AI can create many benefits](#), such as better healthcare; safer and cleaner transport; more efficient manufacturing; and cheaper and more sustainable energy.

In April 2021, the European Commission proposed the first EU regulatory framework for AI. It says that AI systems that can

AI Policy & Training

EducationWeek®

SIGN IN SUBSCRIBE

Educators Want to Know How to Use AI Effectively in Academic Settings

What kind of AI-related professional development topics or approaches would be VERY useful right now to you or—if you are an administrator—to the teachers in your district or school? Select all that apply.

Topic	Percentage
How to teach students how to use AI responsibly and effectively in academic settings	62%
A basic introduction to AI	57%
Ways to detect improper AI use by students	54%
How to prepare students for AI use in the workforce	42%
Ways to use AI to save time on administrative tasks/paperwork	40%
Instruction on incorporating AI into learning across the curriculum	39%
Instruction on incorporating AI into teaching specific subjects	36%
How to teach students about AI's technical features	25%
Professional development that teaches subjects using AI	20%
I can't think of any AI-related professional development topics or approaches that would be very useful right now	8%
Other	4%

Virtual Career Fair
Jan. 25, 2024
2-3 p.m. ET
11 a.m.-3 p.m. PT
SIGN UP NOW!

Top School Jobs

PACE

Topics Initiatives Publications

The Urgent Need to Update District Policies on Student Use of Artificial Intelligence in Education

COMMENTARY

HOME | NEWSROOM | The Urgent Need to Update District Policies on Student Use of Artificial Intelligence in Education

SHARE

AUTHORS
[H. Alex Gallagher](#) | Policy Analysis for California Education, Stanford University
[Benjamin W. Cottingham](#) | Policy Analysis for California Education, Stanford University

PUBLISHED: June 15, 2023

During the 2022–23 school year, artificial intelligence (AI) evolved from an experimental technology few had heard of into readily available technology that has become widely used by educators and students. There are many ways educators can use AI that may positively revolutionize education to benefit classroom instruction, to support data use and analysis, and to aid in decision-making. The biggest potential upsidess of AI for education will be accompanied by major disruptions, however, and districts will need time for thoughtful consideration to avoid some of the worst possible pitfalls. This commentary focuses not on how best to harness the potential of AI in education over the long term but instead on the urgent need for districts to respond to student use of AI. We argue that during summer 2023, districts should adopt policies for the 2023–24 school year that help students to engage with AI in productive ways and decrease the risk of AI-related chaos due to society's inability to detect inappropriate AI use.

The rapid development of AI is causing great concern in education, especially around the potential for widespread misuse of leading-edge products like ChatGPT. ChatGPT, a generative AI chatbot with never-before-seen capabilities, has the power (along with other AI tools) to reshape education because of its ability to mimic human

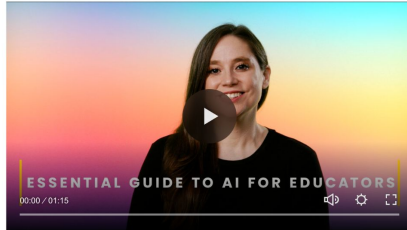
RELATED TOPICS
[Continuous school improvement & success](#)
[Educational governance & policy](#)
[Educator workforce & effectiveness](#)

RELATED PUBLICATION
[AI Policy Guidance for Schools: A TeachAI Toolkit](#)

RELATED COMMENTARY
[From Reactive to Proactive: Putting Districts in the AI Driver's Seat](#)

FROM THE NEWSROOM

Ready to unlock the power of AI in your teaching practice?



An Essential Guide to AI for Educators

This **FREE** 2-hour, hands-on course is designed to help educators get started using ChatGPT to save time, engage students, and implement AI responsibly.

Enroll Now

COURSE

AI for Everyone



- 1 Course
- Introductory
- 6 hours
- Andrew Ng

Enroll Now

Blog

Teaching with AI

We're releasing a guide for teachers using ChatGPT in their classroom—including suggested prompts, an explanation of how ChatGPT works and its limitations, the efficacy of AI detectors, and bias.

Find additional resources in our new FAQ >

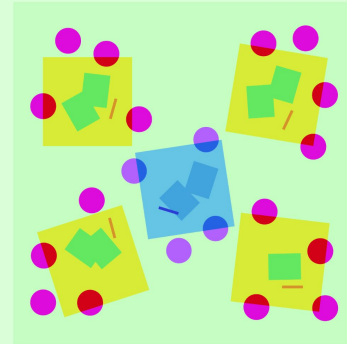



Illustration: Ruby Chen

Build Inclusive & Accessible AI Tools

 Khanmigo
AI power

This conversation is recorded and viewable by your parent and teacher.


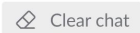

Can you just tell me the answer?

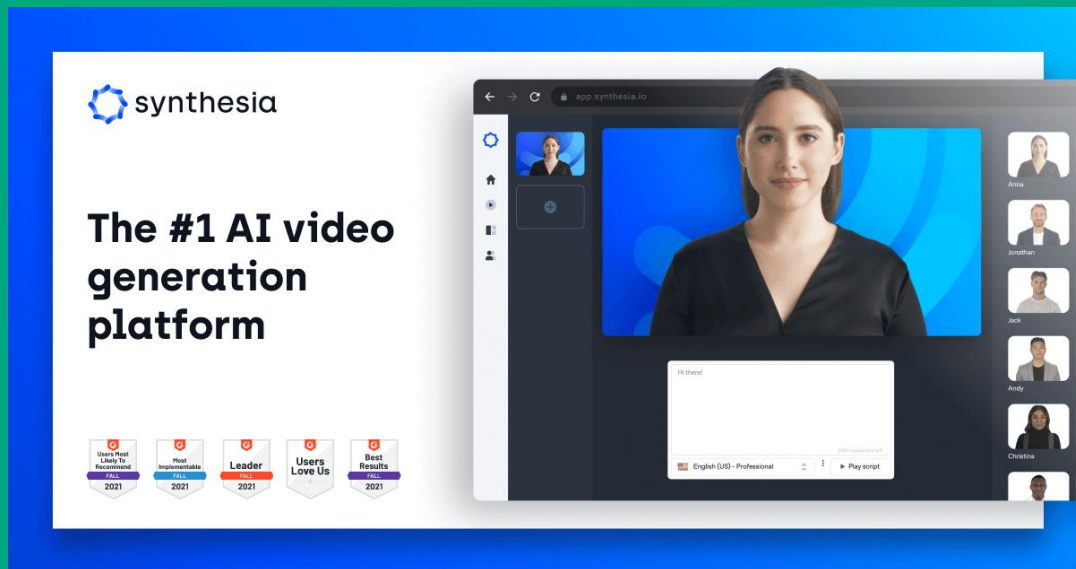
It's important that you learn how to do this yourself! What do you think you need to do to multiply 2 by $5/12$?


They need the same denominator.

That's a good thought, but in this case, you don't need to find a common denominator because you're multiplying, not adding or subtracting. What else could you try?


Type message (do NOT share any personal data)




 synthesia

The #1 AI video generation platform



Hi there!

English (US) - Professional 

Video player interface showing a woman speaking, with a sidebar of other video thumbnails and a list of names (Anna, Jonathan, Jack, Andy, Christina) on the right.



Thank You!

LinkedIn



Dr Philippa Hardman

AI + education. Creator of the DOMSTM
learning science design engine | TEDX Speake...

