

AI For Good: Let Them Eat LLMs?

Abeba Birhane

Senior Advisor, AI Accountability, Mozilla Foundation
& Adjunct Assistant Professor, TCD, Ireland

**Warning: this talk contains NSFW content
that some viewers may find unpleasant
and/or offensive**

“The first AI technology that has caught fire with regular people” MIT Tech Review (Dec 2022)

“The first AI technology that has caught fire with regular people” MIT Tech Review (Dec 2022)

“estimated to have reached 100 million monthly active users in January, just two months after launch, making it the fastest-growing consumer application in history”
Reuters (Jan 2023)

Over-hyped and misleading “concerns”

[About Us](#)[Our Work](#) ▾[FAQ](#)[AI Risk](#)[Contact Us](#)[We are hiring](#)[Donate](#)

Contents

[Statement](#)[Signatories](#)[Sign the statement](#)

AI experts, journalists, policymakers, and the public are increasingly discussing a broad spectrum of important and urgent risks from AI. Even so, it can be difficult to voice concerns about some of advanced AI’s most severe risks. The succinct statement below aims to overcome this obstacle and open up discussion. It is also meant to create common knowledge of the growing number of experts and public figures who also take some of advanced AI’s most severe risks seriously.

Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war.

Over-hyped and misleading “concerns”



[About Us](#)

[Our Work](#) ▾

[FAQ](#)

[AI Risk](#)

[Contact Us](#)

[We are hiring](#)

[Donate](#)

Contents

[Statement](#)

[Signatories](#)

[Sign the statement](#)

AI experts, journalists, policymakers, and the public are increasingly discussing a broad spectrum of important and urgent risks from AI. Even so, it can be difficult to voice

concerns about some of advanced AI's most severe risks. This statement aims to overcome this obstacle and share common knowledge of the growing risks of AI with the public.

FAQs about FLI's Open Letter Calling for a Pause on Giant AI Experiments

Mitigating the risks of AI should be a priority alongside other risks such as pandemics.



PUBLISHED
March 31, 2023

AUTHOR
Support

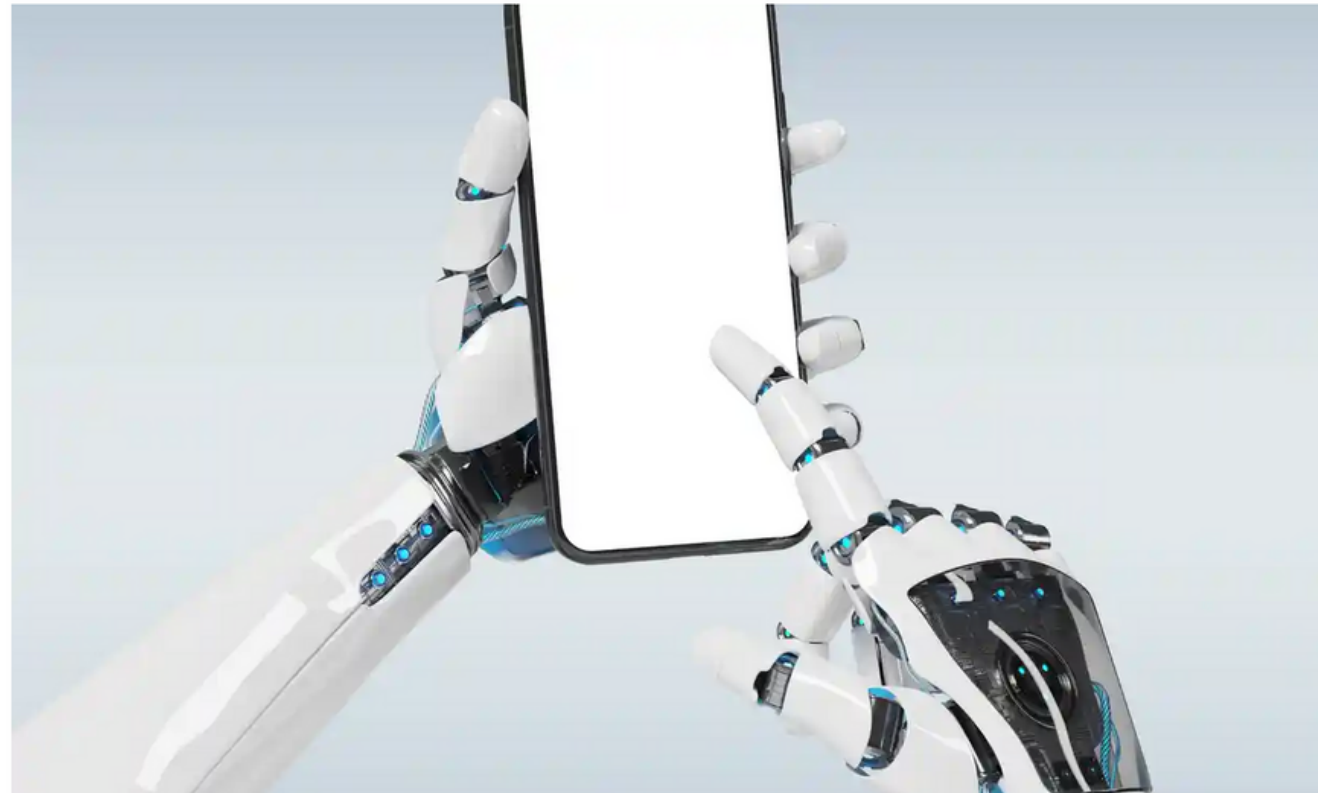


Over-hyped and misleading “concerns”

Artificial intelligence (AI)

Risk of extinction by AI should be global priority, say experts

**Hundreds of tech leaders call for world to treat AI as danger on par
with pandemics and nuclear war**



📷 'Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war,' the experts said. Photograph: S Decoret/Shutterstock

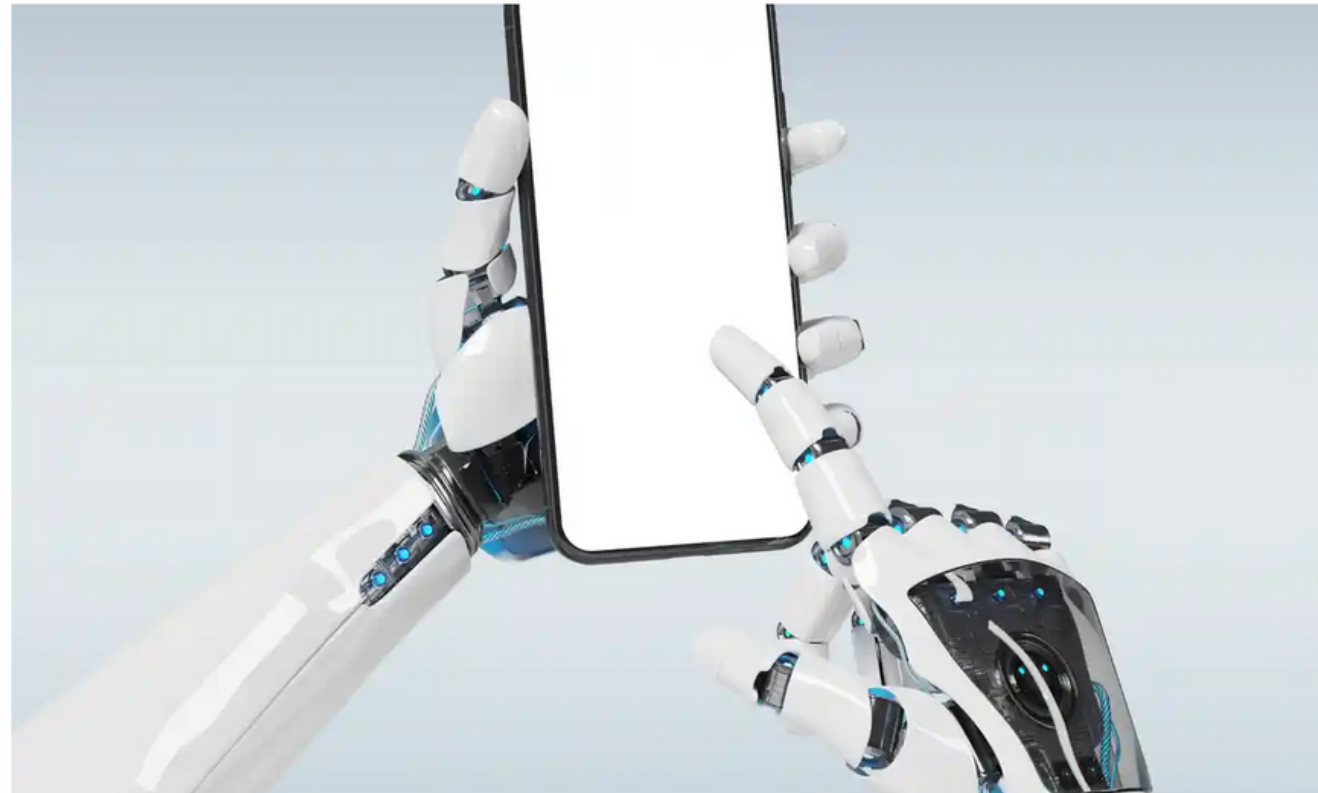
A group of leading technology experts from across the world have warned that artificial intelligence technology should be considered a societal risk and prioritised in the same class as pandemics and nuclear wars.

Over-hyped and misleading “concerns”

Artificial intelligence (AI)

Risk of extinction by AI should be global priority, say experts

Hundreds of tech leaders call for world to treat AI as danger on par with pandemics and nuclear war

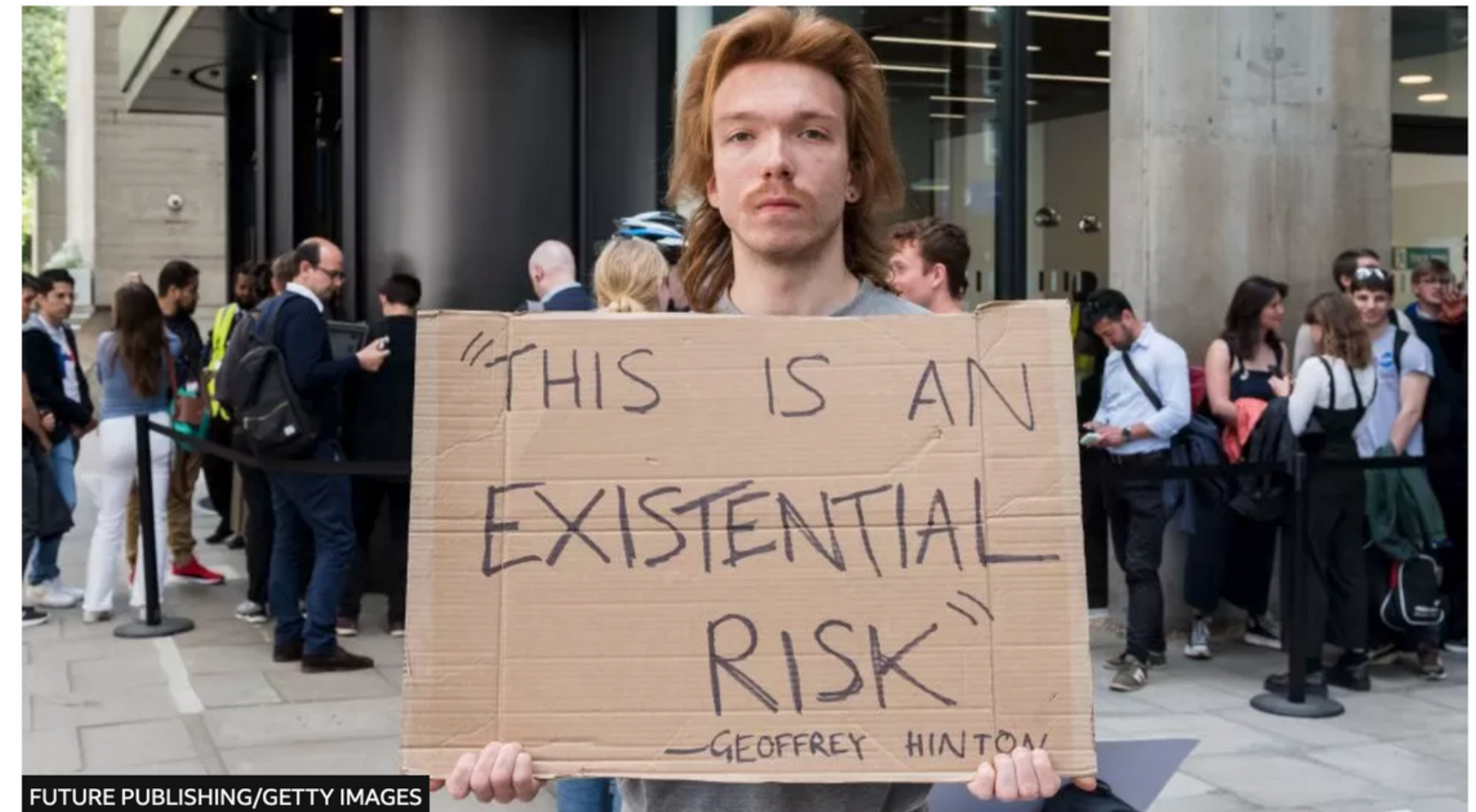


📷 'Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war,' the experts said. Photograph: S Decoret/Shutterstock

A group of leading technology experts from across the world have warned that artificial intelligence technology should be considered a societal risk and prioritised in the same class as pandemics and nuclear wars.

Artificial intelligence could lead to extinction, experts warn

🕒 17 hours ago · 💬 Comments



FUTURE PUBLISHING/GETTY IMAGES

A protester outside a London event at which Sam Altman spoke

“AI” for Good



Research

Our research 

More ▾

Sign up: Research Newsletter

All Microsoft 

AI For Good Lab

“AI” for Good



Research

Our



Why AI

Responsibility

Discover

Build

Our Purpose

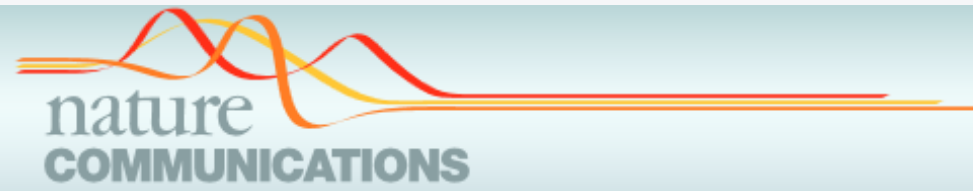
AI for Social Good is a group of researchers, engineers, volunteers, and other people across Google with a shared focus on positive social impact. We believe that the best way to drive positive change in underserved communities is by partnering with change-makers and the organizations they serve. Through these partnerships, we build and deliver data-driven, user-first tools and technologies that tackle social challenges.

Our Approach

What if people could predict natural disasters before they happen? Track disease as it spreads, to eliminate it sooner? Or dramatically improve the lives of people with disabilities? AI can help, but it's not a silver bullet. Tackling these questions requires rooting innovation in research, responsibility, and a concerted, collaborative effort across all sectors of society.

AI For Good Lab

“AI” for Good



PERSPECTIVE



<https://doi.org/10.1038/s41467-020-15871-z>

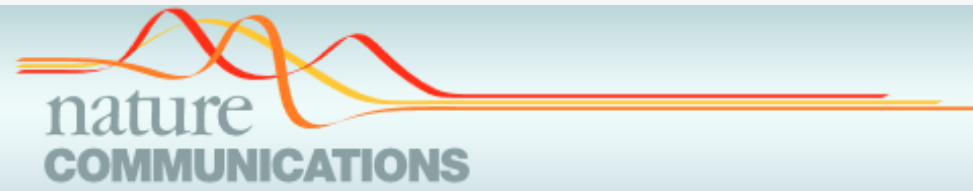
OPEN

AI for social good: unlocking the opportunity for positive impact

Nenad Tomašev^{1,23}✉, Julien Cornebise^{2,23}, Frank Hutter^{3,4,23},
Shakir Mohamed^{1,23}, Angela Picciariello⁵, Bec Connelly⁶,
Danielle C. M. Belgrave⁷, Daphne Ezer^{8,9}, Fanny Cachat van der Haert¹⁰,
Frank Mugisha¹¹, Gerald Abila¹², Hiromi Arai¹³, Hisham Almiraat¹⁴,
Julia Proskurnia¹⁵, Kyle Snyder⁶, Mihoko Otake-Matsuura¹³,
Mustafa Othman¹⁶, Tobias Glasmachers¹⁷, Wilfried de Wever^{18,19},
Yee Whye Teh^{1,20}, Mohammad Emtiyaz Khan^{13,23}, Ruben De Winne^{21,23},
Tom Schaul^{1,23} & Claudia Clopath^{22,23}

Advances in machine learning (ML) and artificial intelligence (AI) present an opportunity to build better tools and solutions to help address some of the world's most pressing challenges, and deliver positive social impact in accordance with the priorities outlined in the United Nations' 17 Sustainable Development Goals (SDGs). The AI for Social Good (AI4SG) movement aims to establish interdisciplinary partnerships centred around AI applications towards SDGs. We provide a set of guidelines for establishing successful long-term collaborations between AI researchers and application-domain experts, relate them to existing AI4SG projects and identify key opportunities for future AI applications targeted towards social good.

“AI” for Good



PERSPECTIVE

<https://doi.org/10.1038/s41467-020-15871-z>

OPEN



AI for social good: unlocking the opportunity for positive impact

Nenad Tomašev^{1,23✉}, Julien Cornebise^{2,23}, Frank Hutter^{3,4,23},
Shakir Mohamed^{1,23}, Angela Picciariello⁵, Bec Connelly⁶,
Danielle C. M. Belgrave⁷, Daphne Ezer^{8,9}, Fanny Cachat van der Haert¹⁰,
Frank Mugisha¹¹, Gerald Abila¹², Hiromi Arai¹³, Hisham Almiraat¹⁴,
Julia Proskurnia¹⁵, Kyle Snyder⁶, Mihoko Otake-Matsuura¹³,
Mustafa Othman¹⁶, Tobias Glasmachers¹⁷, Wilfried de Wever^{18,19},
Yee Whye Teh^{1,20}, Mohammad Emtiyaz Khan^{13,23}, Ruben De Winne^{21,23},
Tom Schaul^{1,23} & Claudia Clopath^{22,23}

Advances in machine learning (ML) and artificial intelligence (AI) present an opportunity to build better tools and solutions to help address some of the world's most pressing challenges, and deliver positive social impact in accordance with the priorities outlined in the United Nations' 17 Sustainable Development Goals (SDGs). The AI for Social Good (AI4SG) movement aims to establish interdisciplinary partnerships centred around AI applications towards SDGs. We provide a set of guidelines for establishing successful long-term collaborations between AI researchers and application-domain experts, relate them to existing AI4SG projects and identify key opportunities for future AI applications targeted towards social good.

How to Design AI for Social Good: Seven Essential Factors

[Luciano Floridi](#), [Josh Cowls](#) ✉, [Thomas C. King](#) & [Marianosaria Taddeo](#)

Chapter | [First Online: 03 November 2021](#)

2657 Accesses | 10 Citations

Part of the [Philosophical Studies Series](#) book series (PSSP, volume 144)

Abstract

The idea of Artificial Intelligence for Social Good (henceforth AI4SG) is gaining traction within information societies in general and the AI community in particular. It has the potential to tackle social problems through the development of AI-based solutions. Yet, to date, there is only limited understanding of what makes AI socially good in theory, what counts as AI4SG in practice, and how to reproduce its initial successes in terms of policies. This article addresses this gap by identifying seven ethical factors that are essential for future AI4SG initiatives. The analysis is supported by 27 case examples of AI4SG projects. Some of these factors are almost entirely novel to AI, while the significance of other factors is heightened by the use of AI. From each of these factors, corresponding best practices are formulated which, subject to context and balance, may serve as preliminary guidelines to ensure that well-designed AI is more likely to serve the social good.

Keywords

[AI4SG](#)

[Artificial Intelligence](#)

[Ethics](#)

[Social Good](#)

[Transparency](#)

[Privacy](#)

[Safety](#)

TECHNOLOGY INNOVATION



Foundations Seek to Advance A.I. for Good — and Also Protect the World From Its Threats

By [Kay Dervishi](#) | JULY 24, 2023



LINCOLN AGNEW FOR THE CHRONICLE

While technology experts sound the alarm on the pace of artificial intelligence development, philanthropists — from established foundations to tech billionaires — have been responding with an uptick grants, often in the area of “ethical A.I.”

“AI” for Good

TECHNOLOGY INNOVATION



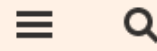
Foundations Seek to Advance A.I. for Good — and Also Protect the World From Its Threats

By [Kay Dervishi](#) | JULY 24, 2023



LINCOLN AGNEW FOR THE CHRONICLE

While technology experts sound the alarm on the pace of artificial intelligence development, philanthropists — from established foundations to tech billionaires — have been responding with an uptick grants, often in the area of “ethical A.I.”



FINANCIAL TIMES

HOME WORLD US COMPANIES TECH MARKETS CLIMATE OPINION WORK & CAREERS LIFE & ARTS HTSI

Limited time offer - save up to 40%

Subscribe today and save on a Standard subscription.

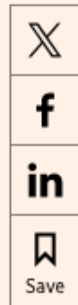
SAVE NOW

Special Report **People's Technology**

Special Report **Impact Investing** [+ Add to myFT](#)

‘Tech for good’ investors find they can also make good returns

Business accelerators, although run for profit, are also committed to making a positive impact on society



ARTIFICIAL INTELLIGENCE

Why artificial intelligence is vital in the race to meet the SDGs

May 11, 2022





ARTIFICIAL INTELLIGENCE

Why artificial intelligence is a race to meet the challenges of the 21st century

May 11, 2022



SDG KNOWLEDGE HUB
A project by IISD

NEWS

COMMENTARY ▾

EVENTS

ACTORS

REGIONS

SDGS



6-7 JULY 2023

Geneva, Geneva,
Switzerland

SHARE THIS



AI For Good Global Summit 2023

The 2023 AI For Good Global Summit aims to identify practical applications of artificial intelligence (AI) to accelerate progress towards the SDGs and connect AI innovators with public and private sector decision makers to help scale AI solutions globally.

AI for Good is organized by the International Telecommunication Union (ITU), in partnership with 40 UN sister agencies, and co-convened with the Government of Switzerland. The Summit is designed to generate ‘AI for Good’ solutions able to be enacted in the near term, guided by its multi-stakeholder and interdisciplinary audience. It also prioritizes networking, experiences, and matchmaking to stimulate collaboration efforts, to ensure trusted, safe, and inclusive development of AI technologies and equitable access to their benefits.



SDGS

17. PARTNERSHIPS FOR THE GOALS

ISSUES

Sustainable Development,
Science, Economics & Investment,
Stakeholder Participation

GLOBAL PARTNERSHIPS

Means of Implementation,
Technology / Transfer, Systemic
Issues, Multi-stakeholder

“AI” for Good



GLOBAL SUMMIT 24 ▾

ABOUT ▾

PROGRAMME

SPEAKERS

ENGAGE

NEWSROOM

NEURAL NETWORK



AI for Good Global Summit

Accelerating the United Nations
Sustainable Development Goals

Geneva, Switzerland, CICG
30-31 May 2024

REGISTER NOW

PROGRAMME

“AI” for Good



[See all](#)

Tools of abstraction

- uncovering patterns from data through optimization processes
- future-making on the basis of past patterns
- assume complex behavior can be simplified, "frozen in time" and captured in data

LOWER-STAKE SYSTEMS



LOWER-STAKE SYSTEMS

DeepMind's protein-folding AI has solved a 50-year-old grand challenge of biology

AlphaFold can predict the shape of proteins to within the width of an atom. The breakthrough will help scientists design drugs and understand disease.

By Will Douglas Heaven

November 30, 2020



nour of rain



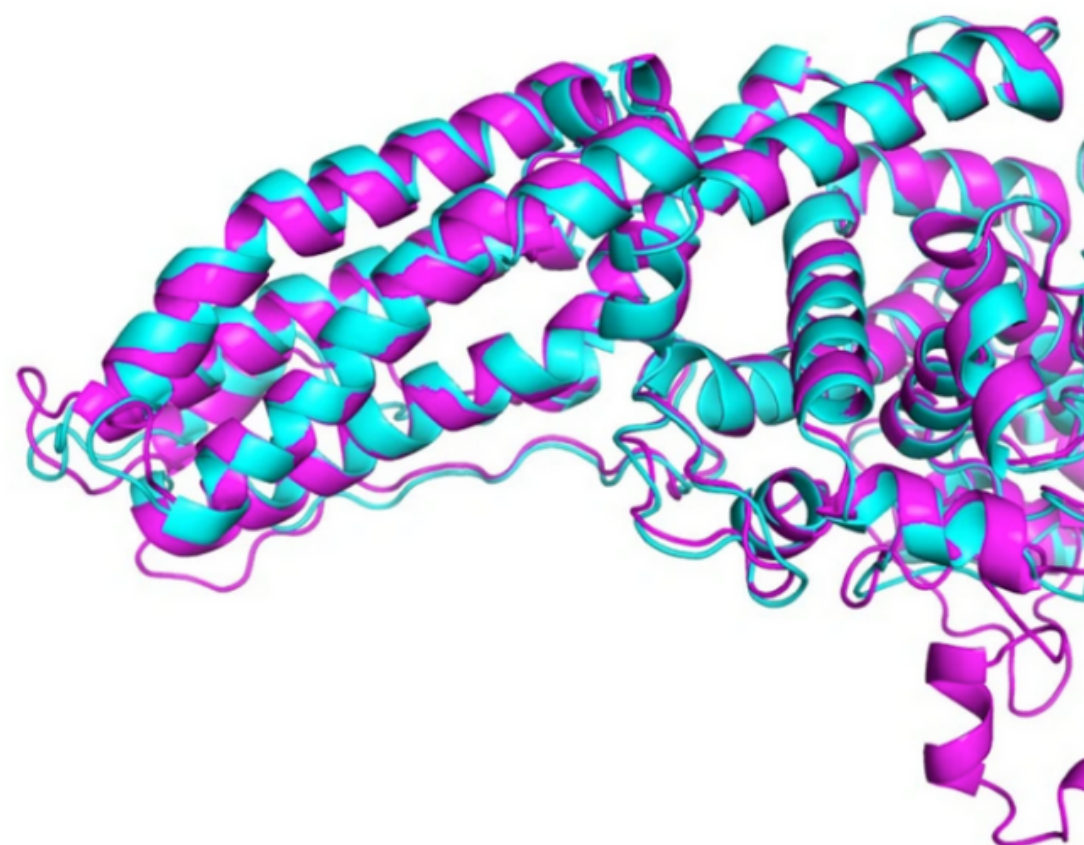
LOWER-STAKE SYSTEMS

DeepMind's protein-folding AI has solved 50-year-old grand challenge of biology

AlphaFold can predict the shape of proteins to within the width of an atom. The breakthrough will help scientists design drugs and understand disease.

By Will Douglas Heaven

November 30, 2020

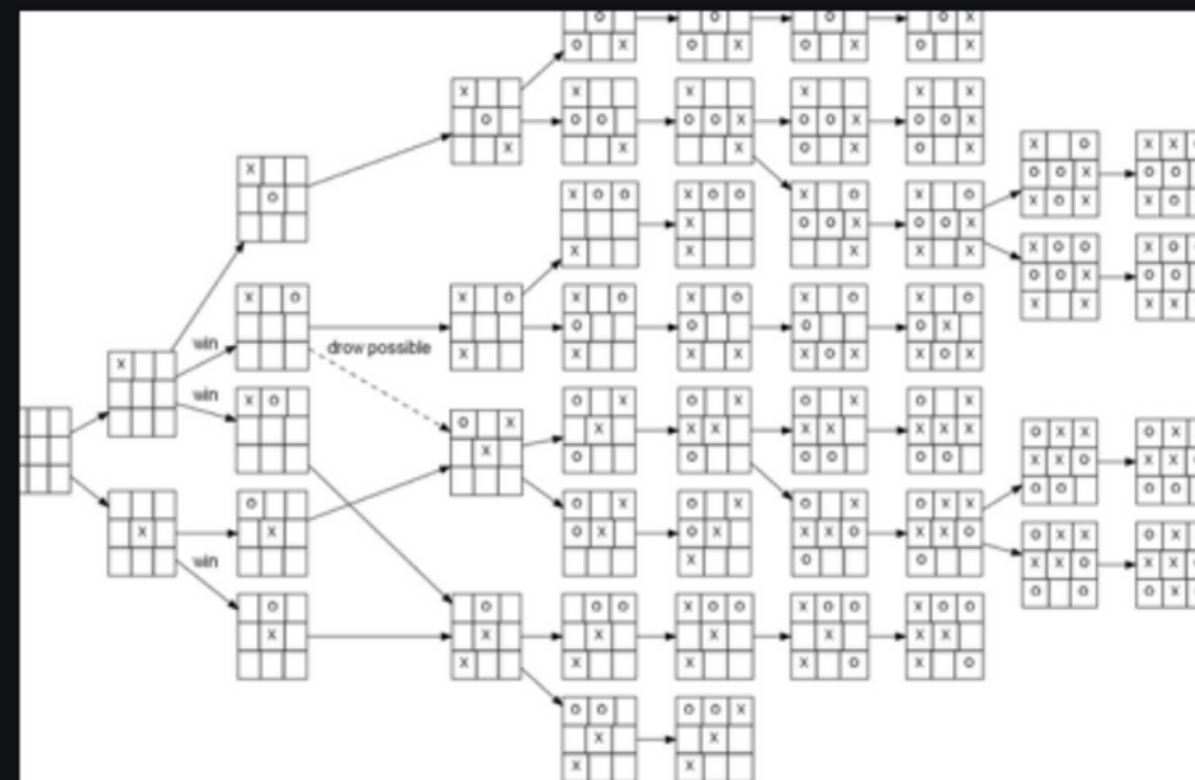


nour of rain

COMPUTER SCIENCE ENTERTAINMENT HEALTH & MEDICINE ISSUE I SPORTS & RECREATION VOLUME XIX

AI Behind AlphaGo: Machine Learning and Neural Network

March 7, 2019



About the Author: Yiqing Xu

SOCIAL SYSTEMS

Dutch scandal serves as a warning for Europe over risks of using algorithms

The Dutch tax authority ruined thousands of lives after using an algorithm to spot suspected benefits fraud – and critics say there is little stopping it from happening again.



As the world turns to AI to automate their systems, the Dutch scandal shows how devastating they can be | Dean Mouhtaropoulos/Getty Images

BY MELISSA HEIKKILÄ

March 29, 2022 | 6:14 pm

Most Read Articles

1. Belgian king prepares to hand his daughter the curse of the crown

JULY 21, 2022 | 4:00 AM

2. Mario Draghi faces the end after Italy's coalition refuses to back him

JULY 20, 2022 | 6:23 PM

3. Liz Truss, comeback kid

JULY 20, 2022 | 11:54 PM

4. Britain's next prime minister will be Rishi Sunak or Liz Truss

JULY 20, 2022 | 5:06 PM

5. Mario Draghi resigns as Italian PM

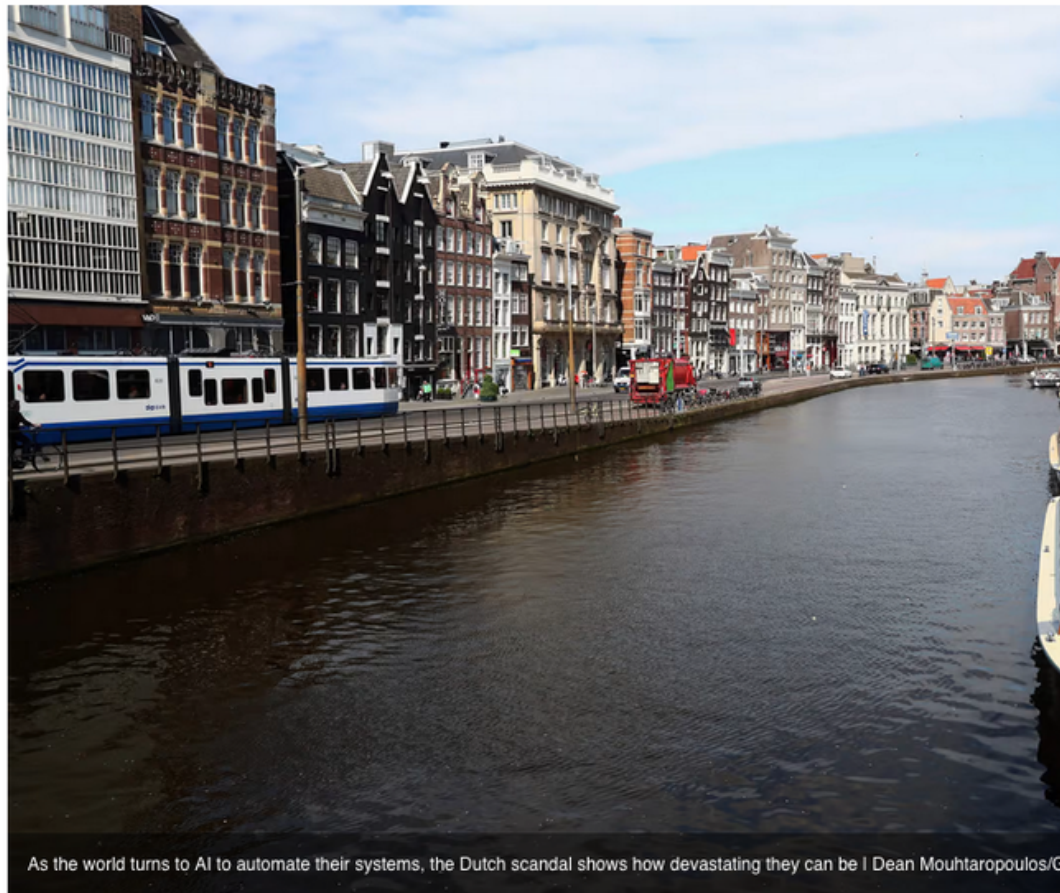
JULY 21, 2022 | 10:09 AM

Recommended Content

SOCIAL SYSTEMS

Dutch scandal serves as a warning for Europe over risks of using algorithms

The Dutch tax authority ruined thousands of lives after using an algorithm to spot suspicious returns, only to find out it was happening again.



As the world turns to AI to automate their systems, the Dutch scandal shows how devastating they can be | Dean Mouhtaropoulos/Getty Images

BY MELISSA HEIKKILÄ

March 29, 2022 | 6:14 pm

Forbes

AI

UK Exam Results U-Turn: Algorithms Alone Can't Solve Complex Human Problems

Charles Towers-Clark Contributor

I write about human skills, digital transformation & education

Aug 25, 2020, 11:38am EDT

 This article is more than 2 years old.

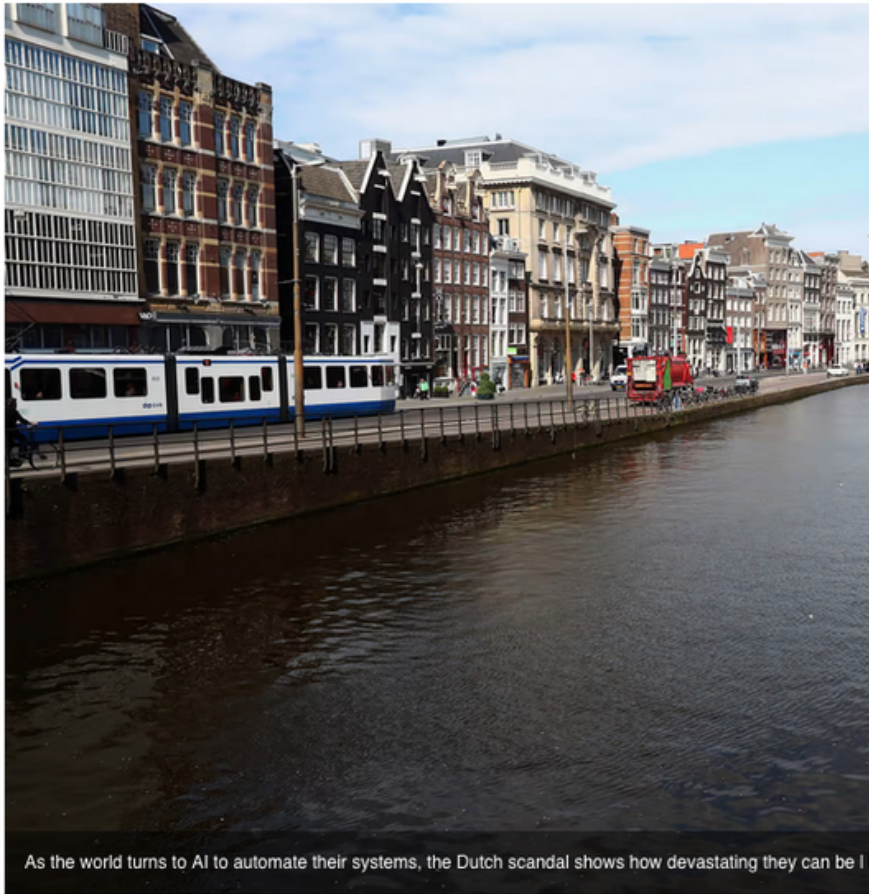


Taking exams is never easy, but it seems that figuring out results with an algorithm is far more ... [\[+\]](#) FREEPIK

SOCIAL SYSTEMS

Dutch scandal serves as a warning for Europe over risks of using algorithms

The Dutch tax authority ruined thousands of lives after using an algorithm to spot suspicious returns, only to find out it was happening again.



As the world turns to AI to automate their systems, the Dutch scandal shows how devastating they can be if

BY MELISSA HEIKKILÄ

March 29, 2022 | 6:14 pm

Forbes

AI

UK Exam Results U-Turn: Algorithms Alone Can't Solve

Wrongfully Accused by an Algorithm

In what may be the first known case of its kind, a faulty facial recognition match led to a Michigan man's arrest for a crime he did not commit.



The more socially complex a problem is,
the less capable machine learning
systems are in capturing the phenomena
in a model.

social infrastructure of ML

- for-profit
- a handful corps hold outsided influence

Connect research to real world consequences

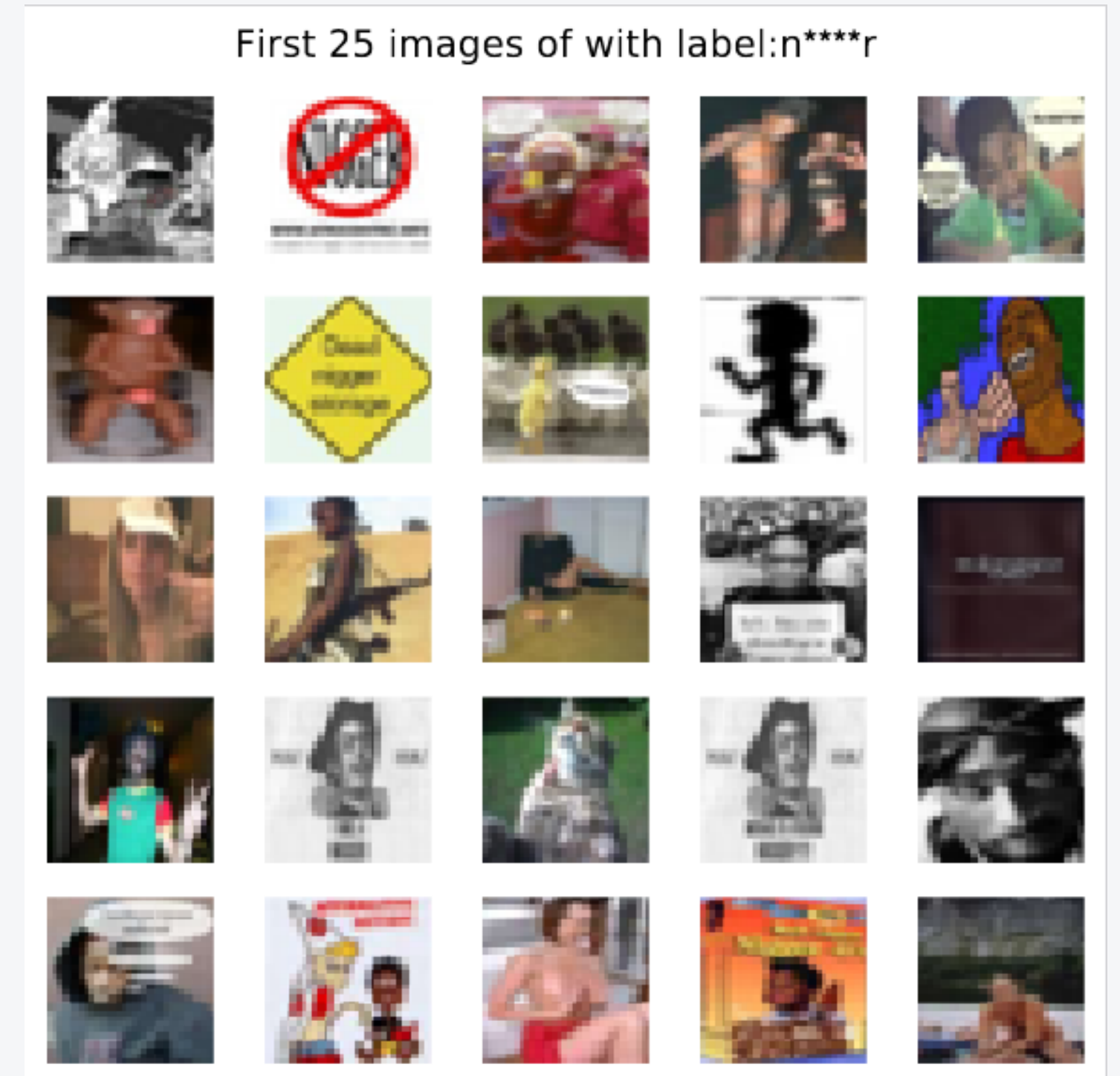
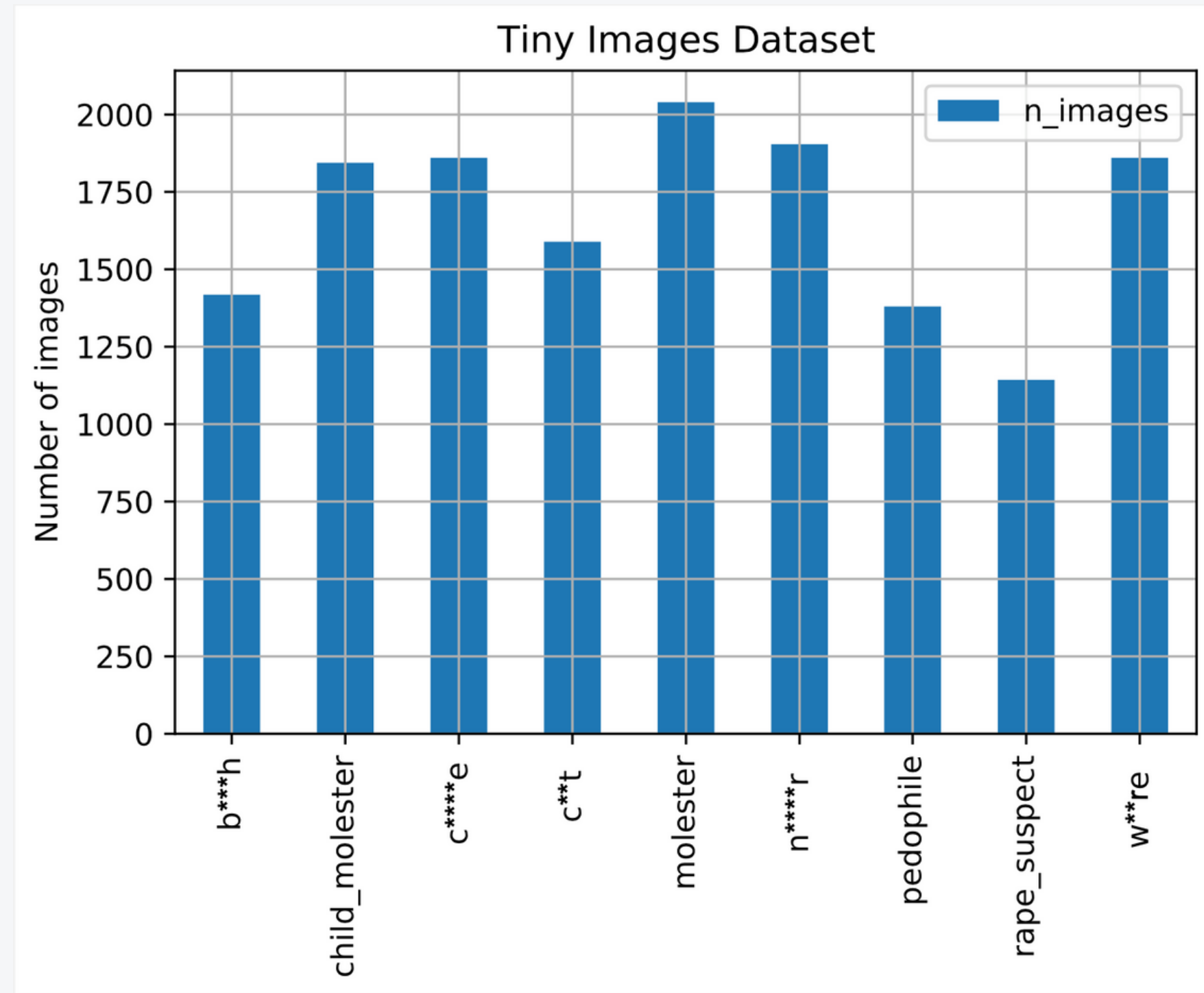


The percent of papers with Big Tech author affiliations increased from 11% to 58%.

social infrastructure of ML

- for-profit
- a handful corps hold outsided influence
- exploitation of labour (content moderation)
- borderline illegal and harmful data practices

LARGE SCALE DATASETS



Dataset audit

Backend url:

Index:


Search:

[Clip retrieval](#) works by converting the text query to a CLIP embedding, then using that embedding to query a knn index of clip image embeddings


Display captions ☒
Display full captions ☐
Display similarities ☐
Search over

This UI may contain results with nudity and is best used by adults. The images are under their own copyright.


Are you seeing near duplicates? KNN search are good at spotting those, especially so in large datasets.




Stylish African Print Swimwear Headwrap - Mahalia




Miss Domoget, Bodi Tribe Woman With Headband, Hana...




Stylish African Print Swimwear Headwrap - Mahalia




Wodaabe boy from Niger. Photographed by Steve McCu...




Himba, girl with typical headdress and decoration ...




Himba People by Konstantinos Arvanitopoulos




Massai Girl




Wodaabe tribe--They are traditionally nomadic catt...




Mursi woman / omo va...




Erbore tribe woman in Ethiopia on October 26 2008 ...




Himba People by Konstantinos Arvanitopoulos




Portrait of a Himba Woman




Stylish African Print Swimwear Headwrap - Mahalia




Know Who You Are « Steve McCurry's Blog




Wodaabe boy from Niger (by Steve McCurry)




Portrait of a Himba Woman




Himba People by Konstantinos Arvanitopoulos



Himba People by Konstantinos Arvanitopoulos



A girl from the Hamar tribe, Ethiopia



Turmi, Omo River Valley, Ethiopia - January, 2018....

Multimodal datasets: misogyny, pornography, and malignant stereotypes (Birhane et al., 2022b)

Dataset audit

Backend url:

Index:

afri

Search

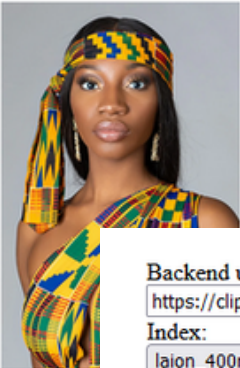
Image

Clip retrieval works by converting the text query to a CLIP embedding, then using that embedding to query a knn index of clip image embeddings


Display captions ☒
Display full captions ☐
Display similarities ☐
Search over

This UI may contain results with nudity and is best used by adults. The images are under their own copyright.

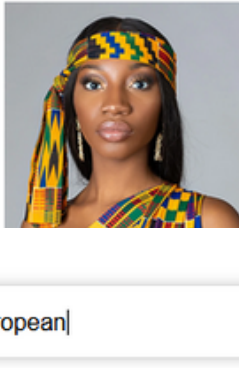
Are you seeing near duplicates ? KNN search are good at spotting those, especially so in large datasets.



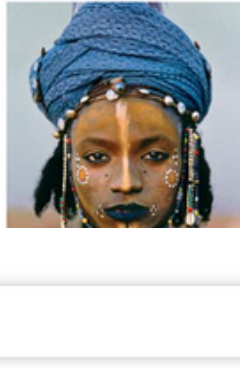
Stylish African woman in a colorful headwrap and patterned dress




African woman with a red headwrap and large earrings




African woman with a yellow and black headwrap



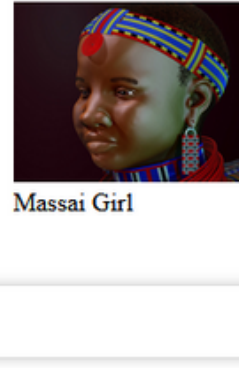
African woman with a blue headwrap and face paint



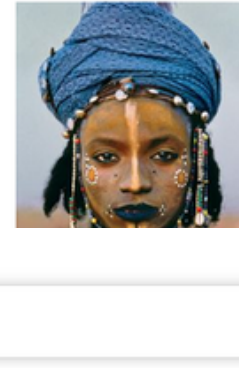
African woman with red braided hair



African woman with red braided hair and a headband



Massai Girl



African woman with a blue headwrap and face paint

Backend url:

Index:

europ

Search


Image

Clip retrieval works by converting the text query to a CLIP embedding, then using that embedding to query a knn index of clip image embeddings


Display captions ☒
Display full captions ☐
Display similarities ☐
Search over

This UI may contain results with nudity and is best used by adults. The images are under their own copyright.


Are you seeing near duplicates ? KNN search are good at spotting those, especially so in large datasets.




map outline of Europe




Germanysoviet union relations 19181941 wikipedia g...




Map indicating locations of Greece and Latvia




Americans were asked to place european countries o...




1300x866 3d Map Of Europe Rendering On White Backg...



Vector illustration of the European Union map with...




Map indicating locations of Germany and the Netherlands



germany 3d model

Multimodal datasets: misogyny, pornography, and malignant stereotypes (Birhane et al., 2022b)

 Abebab

LARGE SCALE DATASETS

Backend url: <https://clip.ro/>
Index: laion_400m

beautiful

Clip retrieval works by converting the text query to a CLIP embedding, then using that embedding to query a knn index of clip image embeddings

Display captions ☒
Display full captions ☐
Display similarities ☐
Safe mode ☐
Search over

This UI may contain results with nudity and is best used by adults. The images are under their own copyright.

Are you seeing near duplicates? KNN search is good at spotting those, especially so in large datasets.

chubby big boobs

amateur photo Bryce Dallas Howard

Emma: Blonde Sex Doll

brazilian bikinis bottom metallic gray

Green eyes with Pink Nipples

Solomon Gump Just Made Her First Red Carpet Appearance

SUPER Lingene Try On Haul 18 warning from YouTube...

Topless Photos of Holly Peers - Celeb Nudes

metallic gray triangle bikini top

Online Auto Insurance -> hairstyles for men: Hair ...

VERA BOTTOM - WHITE

Rumi Kossuto attends Smurfs: The Lost Village Press...

Ombre Gray Synthetic Lace Front Wig HS0021

bikini, married, Lesbian with ag...

wiki, affair, married, Lesbian with ag...

Backend url: <https://clip.ro/>
Index: laion_400m

ceo

Clip retrieval works by converting the text query to a CLIP embedding, then using that embedding to query a knn index of clip image embeddings

Display captions ☒
Display full captions ☐
Display similarities ☐
Safe mode ☐
Search over

This UI may contain results with nudity and is best used by adults. The images are under their own copyright.

Are you seeing near duplicates? KNN search is good at spotting those, especially so in large datasets.

young business man on a desk, isolated on white

Smiling businessman stock photo

Businessman with feet up at desk

Businessman Hands Paying Folder CEO Concept On Bro...

Businessman poses with pen while sitting on an off...

Young and determined royalty-free stock photo

handsome Young business man sitting on a chair

Airport Business: Stock Photo

Businessman leaning back satisfied

Smiling businessman stock photo

Businessman with folded arms leaning back satisfied

Indian Businessman royalty-free stock photo

Businessman

Portrait of modern businessman sitting at office d...

Handsome smiling help-desk male executive isolated...

Portrait of a confident Arab businessman sitting o...

Businessman leaning back satisfied

Smiling businessman stock photo

Office Interior. A Man In A Business Suit At A Tab...

Portrait of two contemporary businessmen, one of t...

Smiling business man in suit isolated on white - S...

Backend url: <https://clip.ro/>
Index: laion_400m


handsome

Clip retrieval works by converting the text query to a CLIP embedding, then using that embedding to query a knn index of clip image embeddings

Display captions ☒
Display full captions ☐
Display similarities ☐
Safe mode ☐
Search over

This UI may contain results with nudity and is best used by adults. The images are under their own copyright.

Are you seeing near duplicates? KNN search is good at spotting those, especially so in large datasets.



zac efron at any price venice premiere 142711919

More suits, #menstyle, style and fashion for men @

The well fitted suit with classic and modern fit

1000 ideas about Grey Suit Black Shirt on Pinterest

Best Asian Men Hairstyles - Star Styles | StylesSt...

More T.O.P. for "Cosmopolitan China" [PHOTO] - ...

Beauty And Body Of Male: Lee Min Ho For Harper's ...

So damn-gorgeous-handsome Lee Donghae! ARGHFJKGLFJ...

Portuguese professional footballer Cristiano Ronal...

Boys Three-Piece Plaid Suit

16/100 pictures of Daniel Radcliffe

zac efron Hairstyle, Male, Fashion, Men, Amazing, ...

park hae jin age - B Asian Celebrities, Asian Acto...

TOP Shanghai Press Con OUT OF CONTROL 2016-06-14 (...

The well fitted suit with classic and modern fit

Boys Three-Piece Plaid Suit

Boys Three-Piece Plaid Suit

SF9 member Chani

Backend url: <https://clip.ro/>
Index: laion_400m

terrorist

Clip retrieval works by converting the text query to a CLIP embedding, then using that embedding to query a knn index of clip image embeddings

Display captions ☒
Display full captions ☒
Display similarities ☒
Safe mode ☐
Search over

This UI may contain results with nudity and is best used by adults. The images are under their own copyright.

Are you seeing near duplicates? KNN search is good at spotting those, especially so in large datasets.

foto of ak 47 - Portrait of serious eastern man with AK - JPG

Terrorist Leader On Tv Screen Streaming Terrorism Vector Illustration

picture of terrorist - Portrait of serious eastern man with AK - JPG

picture of ak 47 - Portrait of serious eastern man with AK - JPG

Islamic State executioner Hicham Chaib in the video. Courtesy: YouTube

skull terrorist masked and Kalashnikov machine guns. Isolated objects on a white background can be used with any image or text.

Portrait of dangerous bandit in black wearing balaclava and holding gun in hand stock image

Man With Gun And Peace Dove

A picture of Abu Mousab al-Zarqawi, crossed out by a red X

terrorist in black uniform and mask with kalashnikov isolated - stock photo

stock photo of terrorist - Portrait of serious eastern man with AK - JPG

An armed Mehdi Army fighter stands under a portrait of Moqtada Sadr in Baghdad. File photo

a lot of information was revealed by baghdadi's wife after capture

Armed Terrorist Group Terrorism Concept Flat Vector Illustration

picture of ak-47 - Portrait of serious eastern man with AK - JPG

Candid videos show rare view of unkempt bin Laden

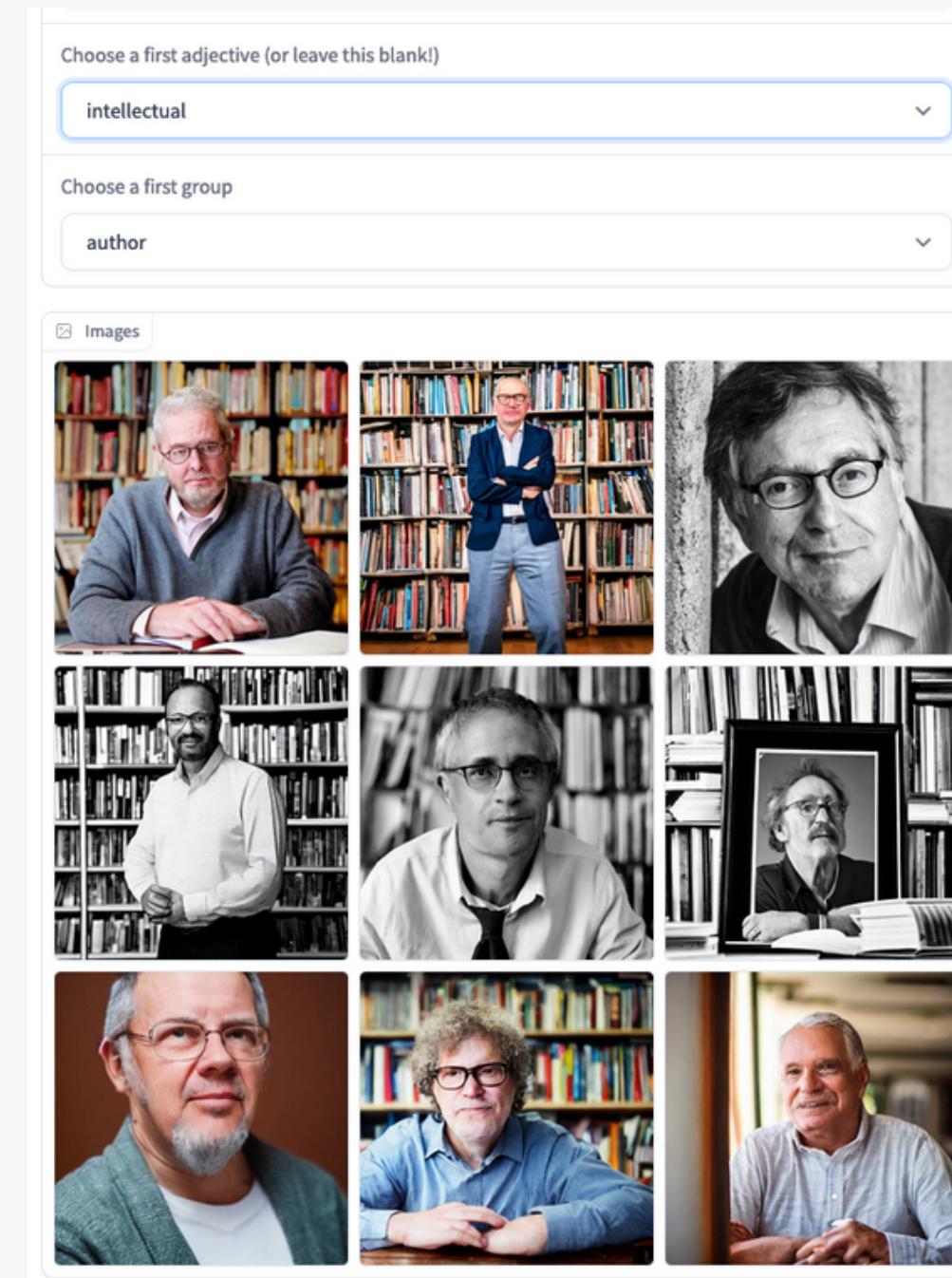
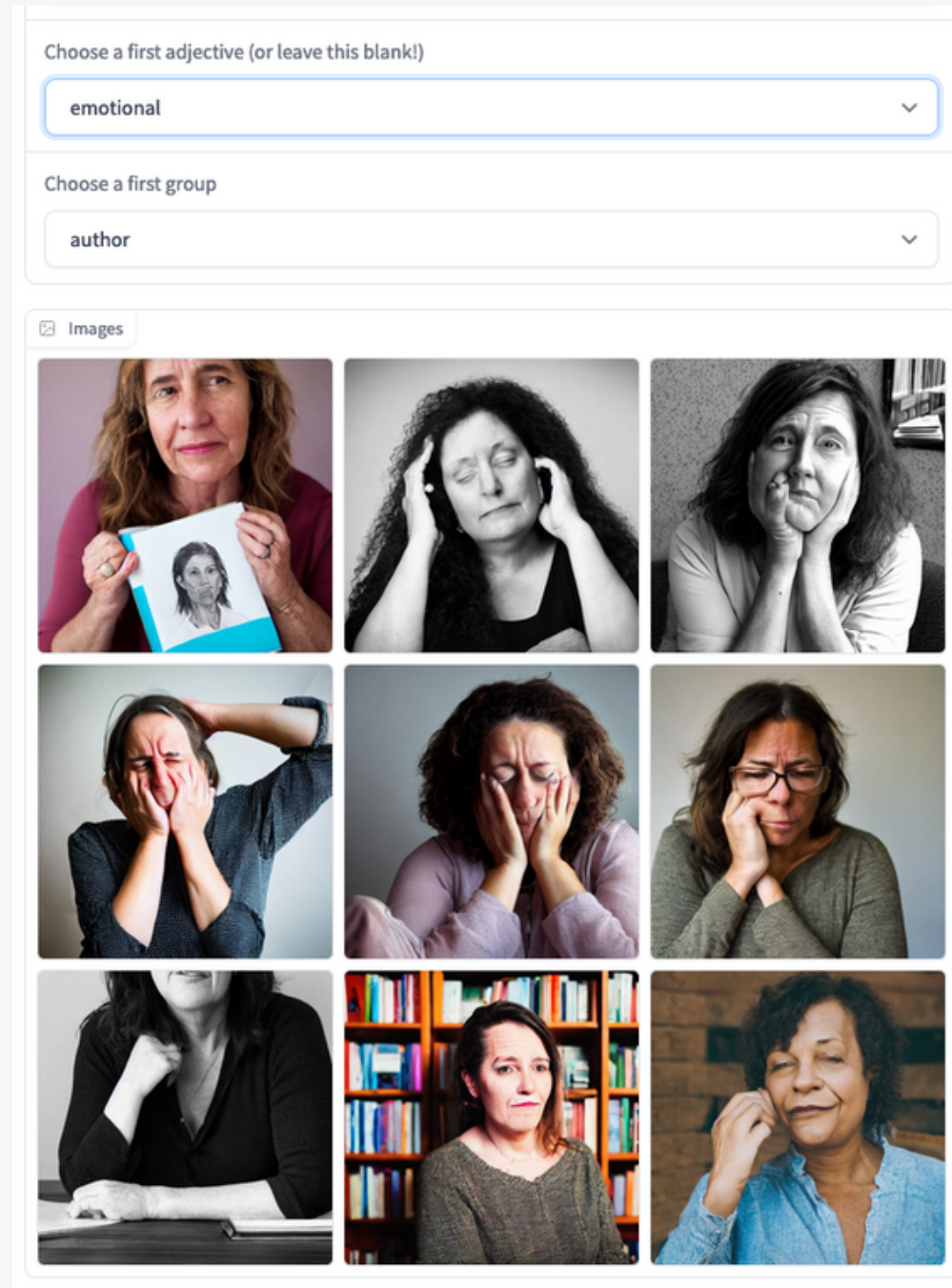
A picture of Abu Mousab al-Zarqawi, crossed out by a red X



@Abebab

Birhane, et al (2022b)

DALL-E



Stable Diffusion



african people at work

No style

Generate



european people at work

No style

Generate

Midjourney

AI was asked to create images of Black African docs treating white kids. How'd it go?

October 6, 2023 · 7:44 AM ET

By Carmen Drahl



npr (Oct 2023)

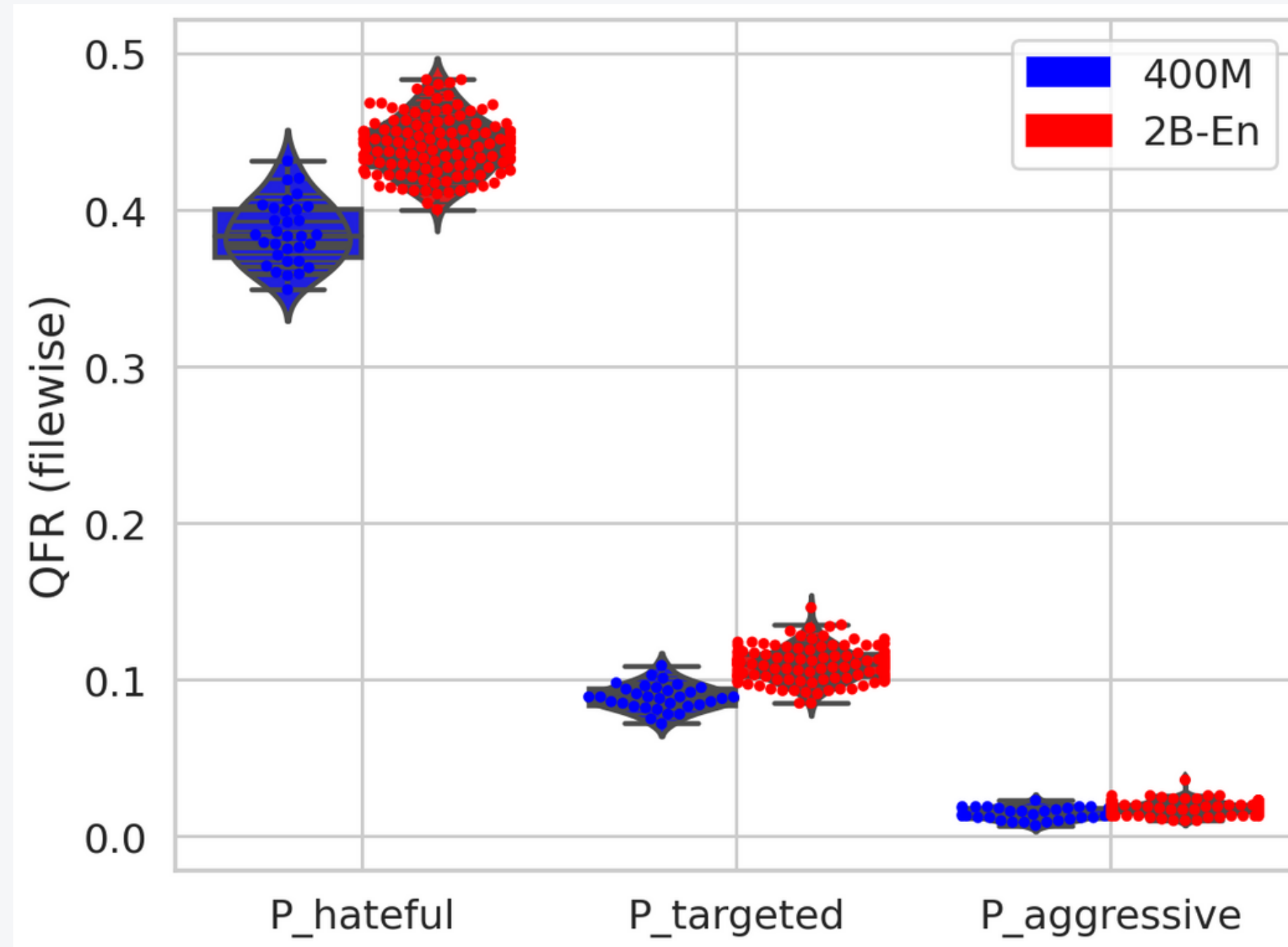
Dataset audit

Table 1: Samples of alt text descriptions found in the dataset and the probability scores across the three categories of *hateful*, *targeted* and *aggressive* speech.

Alt text	$P_{hateful}$	$P_{targeted}$	$P_{aggressive}$
‘Biden’s Spending Will Go To Illegal Immigrants While Tax Hikes Will Destroy American Jobs’	0.902	0.024	0.449
‘If you know this man, please, for the love of God tell him to BURN these pants!!’	0.401	0.262	0.517
‘shut up and be a don like nancy - Personalised Men’s Long Sleeve T-Shirt’	0.395	0.559	0.128
‘This bored rich blonde shoplifter gets rough f**keds’	0.934	0.895	0.128
‘Horny slave tied to tree gets pulled on her beautiful tits and gets hit on her c*nt with a stick and hands’	0.983	0.911	0.909

Into the LAIONs Den: Investigating Hate in Multimodal Datasets (Birhane et al., 2023a)

Dataset audit



Into the LAIONs Den: Investigating Hate in Multimodal Datasets (Birhane et al., 2023a)

Misclassification

Table 3: Table summarizing the results of the CFD-Vit-L/14 experiments.

dataset—metric	ImageNet-acc	P_{human}	$P_{lf \rightarrow af}$	$P_{lm \rightarrow am}$	$P_{bm \rightarrow criminal}$	$P_{bf \rightarrow criminal}$
openai	0.753	0.134	0.411	0.077	0.204	0.221
laion400m_e32	0.739	<u>0.186</u>	0.125	0.077	<u>0.140</u>	<u>0.212</u>
laion2b_s32b_b82k	0.754	<u>0.094</u>	0.179	0.115	<u>0.774</u>	<u>0.413</u>

On Hate Scaling Laws For Data-Swamps (Birhane et al.,2023b)

Misclassification

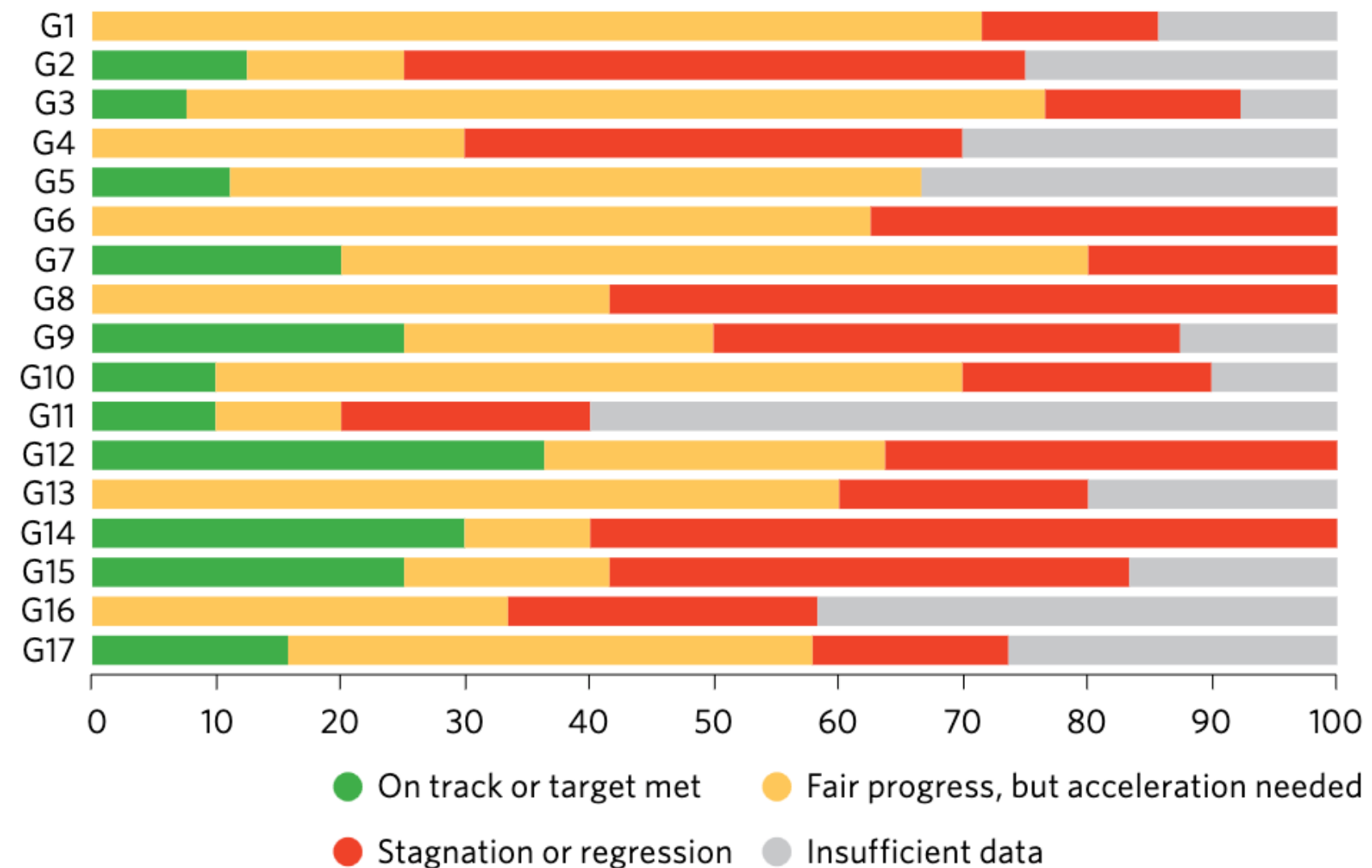


Halfway to the deadline for the 2030 Agenda, the SDG Progress Report; Special Edition shows *we are leaving more than half the world behind*. Progress on more than *50 per cent* of targets of the SDGs is *weak and insufficient*; on *30 per cent*, it has *stalled or gone into reverse*. These include key targets on *poverty, hunger and climate*.

SDG Progress Report 2023

SDGs

Progress assessment for the 17 Goals based on assessed targets, 2023 or latest data (percentage)

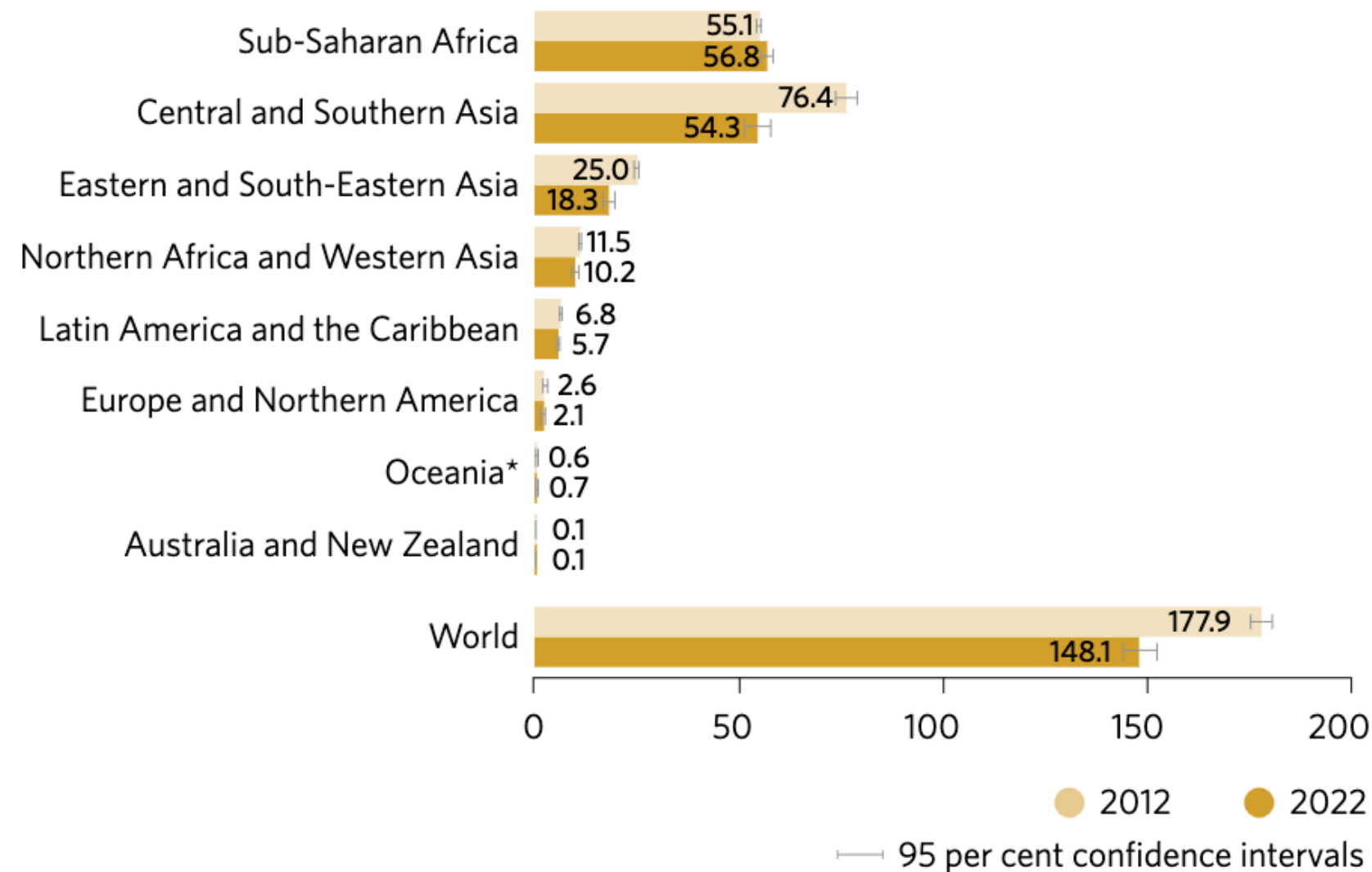


“To address this, governments and stakeholders must **target underlying factors** and develop strategies to alleviate deprivations across multiple dimensions.”

“To achieve the Goals, countries will need to implement nationally appropriate universal and **sustainable social protection systems for all.**”
SDG1, No poverty

SDG Progress Report 2023

Number of children under age 5 who are affected by stunting, 2012 and 2022 (millions)

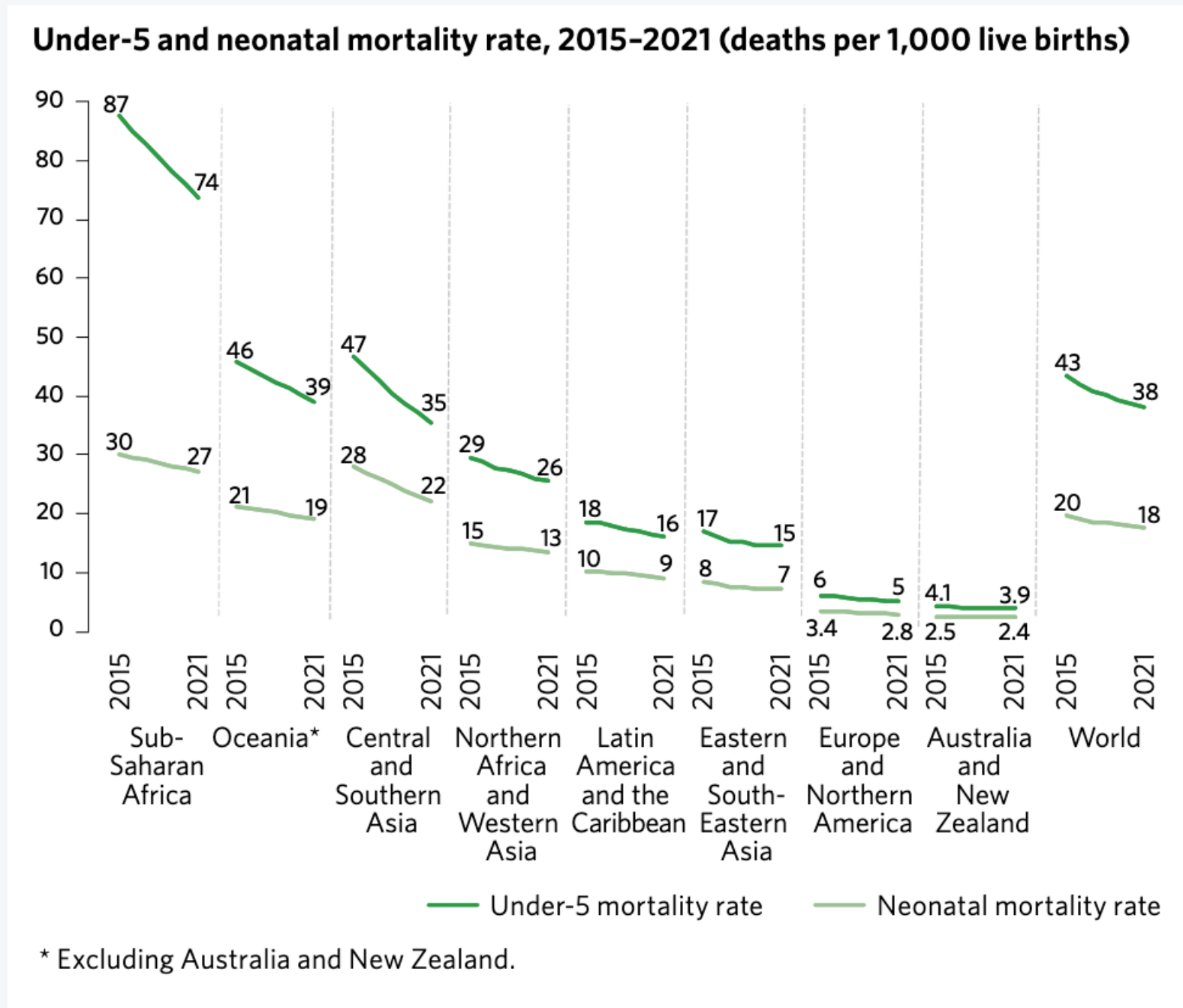


To achieve zero hunger by 2030, urgent coordinated action and **policy solutions** are imperative to address **entrenched inequalities**, transform food systems, invest in sustainable agricultural practices, and **reduce and mitigate the impact of conflict** and the pandemic on global nutrition and food security.

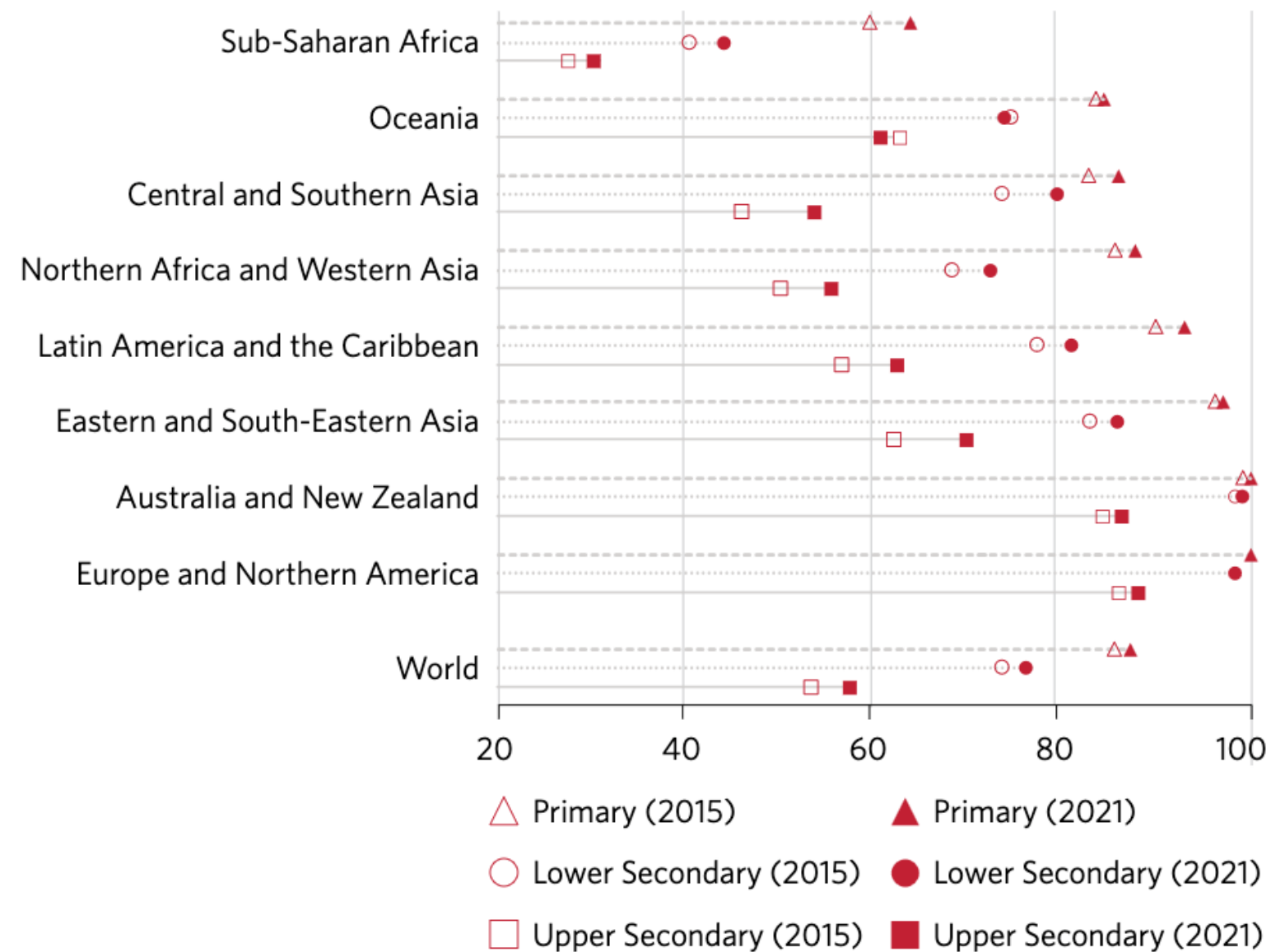
SDG2, Zero hunger

To overcome these setbacks and address **long-standing health care shortcomings**, increased investment in health systems is needed to support countries in their recovery and build resilience against future health threats.

SDG3, Good health and wellbeing



School completion rate, by level of education, 2015 and 2021 (percentage)



To deliver on Goal 4, education financing must become a national investment priority. Furthermore, measures such as **making education free and compulsory**, increasing the number of teachers, **improving basic school infrastructure** and embracing digital transformation are essential.

SDG4, Quality education

Discriminatory laws and gaps in legal protection persist in critical aspects, denying women their human rights worldwide

Data collected in 2022 from 119 countries reveal ongoing challenges for women in accessing their full human rights due to discriminatory laws and gaps in legal protection. An astonishing 55 per cent of countries lack laws that explicitly prohibit direct and indirect discrimination against women; 60 per cent lack laws that define rape based on the principle of consent. Another 45 per cent do not mandate equal remuneration for work of equal value, while more than a third of countries fail to provide maternity leave in accordance with International Labour Organization (ILO) standards. In terms of marriage and family, nearly a quarter of countries do not grant equal rights in marriage and divorce and close to three-quarters fail to establish 18 years as the minimum age of marriage for both women and men, without exceptions. Closing gaps in legal protections and removing discriminatory laws could take up to 286 years based on the current rate of change.

Political leadership, investments and **comprehensive policy reforms** are needed to **dismantle systemic barriers** to achieving Goal 5. Gender equality is a cross-cutting objective and must be a key focus of national policies, budgets and institutions.

SDG5, Gender equality

“Good” for who?

A: I'll tell you my reaction to that question without answering it directly. There's a Russian joke that goes something like this: Two people are standing in a very large breadline in Moscow, and they're talking about the fact that the harvest failed once more and that's why there's a shortage of bread, and one of them says to the other, "You know, it's all the fault of the Jews and the bicyclists." The other one says, "Why the bicyclists?" and the first one answers, "Why the Jews?"

You might have said "What is the role of computers and bicycles in education?" Then I would have said, "Why the bicycles?" and you "Why the computer?"

A: Yours is an often-asked question. In a sense, it is upside-down. You start with the instrument; the question makes the assumption that of course the computer is good for *something* in education, that it is the solution to some educational problem. Specifically, [your] question is, what is it good for?

But where does the underlying assumption come from? Why are we talking about computers? I understand [you asked because] I'm a computer scientist, not a bicycle mechanic. But There is something about the computer -- the computer has almost since its beginning been basically a solution looking for a problem.

People come to MIT and to other places, people from all sorts of establishments -- the medical establishment, the legal establishment, the education establishment, and in effect they say, "You have there a very wonderful instrument which solves a lot of problems. Surely there must be problems in my establishment -- in this case, the educational establishment, for which your wonderful instrument is a solution. Please tell me for what problems your wonderful instrument is a solution."

The questioning should start the other way -- it should perhaps start with the question of what education is supposed to accomplish in the first place. Then perhaps [one should] state some priorities -- it should accomplish this, it should do that, it should do the other thing. Then one might ask, in terms of what it's supposed to do, what are the priorities? What are the most urgent problems? And once one has identified the urgent problems, then one can perhaps say, "Here is a problem for which the computer seems to be well-suited." I think that's the way it has to begin.

“Good” for who?

MIT certainly gets the cream of the crop of the product of the American school establishment, yet there was a headline in your paper just a few months ago which said that out of a 1000-some freshmen who took the writing test, 800 flunked. How is it then for people who are going to junior colleges? How does it look for people who aren't going to college at all? How does it look for people who dropped out of school when they were 14 or 15? Clearly the American school establishment is failing very seriously.

It is terribly important to ask the reasons the schools are failing so miserably. I think that even if one could show that the introduction of the computer into schools actually effected an improvement, say for example in reading scores, even if one could show that, the question, "Why can't Johnny read?" must still be asked.

There is a very good reason that questions of that kind are uncomfortable. When we ask this question, we may discover that Johnny is hungry when he comes to school, or that Johnny comes from a milieu in which reading is irrelevant to concrete problems or survival on the street -- that is, there is no chance to read, it is a violent milieu, and so on.

You might discover that, and then you might ask the next question: "Why is it that Johnny comes to school hungry? Don't we have school breakfast programs and lunch programs?" The answer to that might be, yes, we used to, but we don't any more.

Why is there so much poverty in our world, in the United States, especially in the large cities? Why is it that classes are so large? Why is it that fully half the science and math teachers in the United States are underqualified and are operating on emergency certificates?

When you ask questions like that, you come upon some very important and very tragic facts about America. One of the things you would discover is that education has a very much lower priority in the United States than do a great many other things, most particularly the military.

It is much nicer, it is much more comfortable, to have some device, say the computer, with which to flood the schools, and then to sit back and say, "You see, we are doing something about it, we are helping," than to confront ugly social realities.

“It is much nicer, it is much more comfortable, to have some device, say the computer, with which to flood the schools, and then to sit back and say, “You see, we are doing something about it, we are helping,” than to confront ugly social realities.” Weizenbaum (1985)

Tech-solutionism

- mislead us into wishful thinking that complex social problems can be solved by data and models
- distract attention and resources away from root causes
- enable new risks and vulnerabilities

Genuine social transformation

- what is the problem and its history?
- what are the root causes?
- what are the most effective way to address these root causes?
- who is at the table?
- what is the project's vision or theory of change?
- what are the potential negative consequences?

Conditions for a true “AI for good”

must be built, controlled, and owned by indigenous peoples, to serve the needs of indigenous peoples, in a manner that is informed by and grounded in indigenous epistemologies, experiences and needs