

WHAT DEFINES THE “OPEN” IN “OPEN AI”?

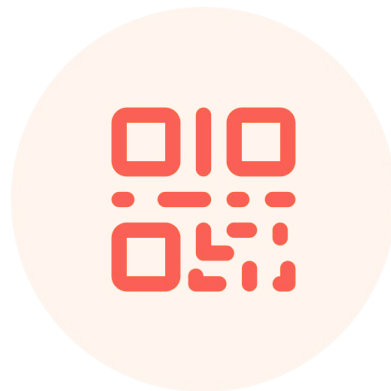
jen ding
bbuzz 2023

jdilling@turing.ac.uk

WHAT DEFINES THE “OPEN” IN “OPEN AI”?

jen ding
bbuzz 2023

jd Ding@turing.ac.uk



#1516847

**OPENCV +
SCIKIT-LEARN**

YOLO

BERT

OPEN LLMs



**The
Alan Turing
Institute**



**WHAT DEFINES
THE “SMART” IN
“SMART CITIES”?**



How many hours of your life will you spend looking for parking?

- Drivers globally spend on average 20 minutes finding parking
- 2012 global life expectancy is ~70 years
- With 300 trips a year, this is over **5000 hours** in a lifetime

Designing a Cost-Effective Outdoor Parking Technology

Criteria	Sensor Systems	ParkIT
Cost	\$400/space	<\$50/space ✓
Accuracy	90 - 99%	95 - 99% ✓
Spaces Covered per Device	1	20 - 100+ ✓

Rice Trial Test Results

Algorithm has achieved **99%** accuracy

Current challenges:

- Vehicle occlusion
- Illumination differences



How to Identify Cars with Computer Vision?



Camera Solution: Pan-Tilt-Zoom camera takes picture at fixed intervals

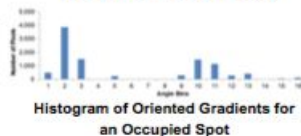
Feature Extraction

Machine Learning

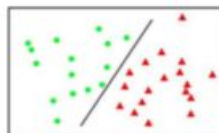
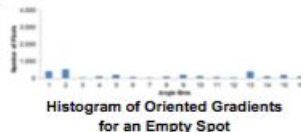
Decision Making



Boundaries in an Image:
Canny Edge Detection



Magnitude and Direction of
Gradients



Occupied



Occupied

Open

Web User Interface



Live Camera View

Occupancy Percentage

Occupancy Map

Future ParkiT Work

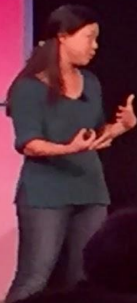


Acknowledgements

Special thanks to Dr. Gary Woods, Dr. Ashu Sabharwal, William Deigaard, and Mike Morgan as well as the OEDK and Rice ECE department for their support and mentorship.

IPI INTERNATIONAL PARKING INSTITUTE
Nashville TN
2016
MAY 19-20
CONFERENCE & EXPO

IPI INTERNATIONAL PARKING INSTITUTE
Nashville
2016





wework
CREATOR
Awards
Launch Awards





**WHO DEFINES
THE “SMART” IN
“SMART CITIES”?**

WHO DEFINES THE “SMART” IN “SMART CITIES”?



- ❖ THE STATE OF OPEN IN AI
- ❖ LOOKING BACK OVER 20 YEARS
- ❖ THE MULTIPLE FACES OF OPEN AI
- ❖ WHO DEFINES THE FUTURE OF OPEN AI?

- ❖ **THE STATE OF OPEN IN AI**
- ❖ LOOKING BACK OVER 20 YEARS
- ❖ THE MULTIPLE FACES OF OPEN AI
- ❖ WHO DEFINES THE FUTURE OF OPEN AI?

slido



What do you think defines the "open" in "open AI"?

① Click **Present with Slido** or install our [Chrome extension](#) to activate this poll while presenting.

THE STATE OF OPENAI

OpenAI is giving Microsoft exclusive access to its GPT-3 language model

The lab was supposed to benefit humanity. Now it's simply benefiting one of the richest companies in the world.

By Karen Hao

September 23, 2020

THE STATE OF OPENAI

OpenAI is giving Microsoft exclusive to its GPT-3 language model

The lab was supposed to benefit humanity. Now it's simply benefiting one of the richest companies in the world.

By Karen Hao

September 23, 2020

MOTHERBOARD
TECH BY VICE

OpenAI Is Now Everything It Promised Not to Be: Corporate, Closed-Source, and For-Profit

OpenAI is today unrecognizable, with multi-billion-dollar deals and corporate partnerships. Will it seek to own its shiny AI future?



By Chloe Xiang

Home + Blog

What defines the 'open' in 'open AI'?

A growing number of researchers are challenging big tech and rethinking the meaning of openness in AI development

Learn more ↓

Monday 01 Aug 2022

Filed under

Personal perspective

Authors



Jennifer Ding

TPS Senior Researcher -
Research Applications

THE STATE OF OPEN IN AI

**Meta and Google news adds fuel
to the open-source AI fire**

Sharon Goldman
@sharongoldman

May 18, 2023 9:33 AM

THE STATE OF OPEN IN AI

Why Developers Are Flocking to LLaMA, Meta's Open Source LLM

Developers are diving into the open source LLaMA, but what benefits does it have over GPT? We spoke to Sebastian Raschka from Lightning AI.

May 5th, 2023 4:00am by [Richard MacManus](#)

Meta's LeCun: "The platform that will win will be the open one"

Meta and Google news adds fuel to the open-source AI fire

Sharon Goldman
[@sharongoldman](#)

May 18, 2023 9:33 AM

THE STATE OF OPEN IN AI

Google "We Have No Moat, And Neither Does OpenAI"

Leaked Internal Google Document Claims Open Source AI Will Outcompete Google and OpenAI



DYLAN PATEL AND AFZAL AHMAD

MAY 4, 2023 · PAID

Google engineer warns it could lose out to open-source technology in AI race

Commonly available software poses threat to tech company and OpenAI's ChatGPT, leaked document says



A worker walks along a path at Google's Bay View campus in Mountain View, California
Photograph: Noah Berger/AFP/Getty Images

Why Developers Are Flocking to LLaMA, Meta's Open Source LLM

Developers are diving into the open source LLaMA, but what benefits does it have over GPT? We spoke to Sebastian Raschka from Lightning AI.

May 5th, 2023 4:00am by [Richard MacManus](#)

Meta's LeCun: "The platform that will win will be the open one"

Meta and Google news adds fuel to the open-source AI fire

Sharon Goldman
[@sharongoldman](#)

May 18, 2023 9:33 AM

THE STATE OF OPEN IN AI

Google "We Have No Moat, And Neither Does OpenAI"

Leaked Internal Google Document Claims OpenAI



DYLAN PATEL AND AFZAL AHMAD
MAY 4, 2023 · PAID

Google engineer warns it could lose to open-source technology in AI race

Commonly available software poses threat to tech company and OpenAI's ChatGPT, leaked document says

Hugging Face launches open-source version of ChatGPT in bid to challenge dominance of closed-source models

A worker walks along a path at Google's Bay View campus in Mountain View, California
Photograph: Noah Berger/AFP/Getty Images

Open Source AI Has a New Champion

Falcon uses a modified Apache licence, meaning the models can be fine-tuned and used for commercial purposes.

By Pritam Bordoloi

Sharon Goldman
@sharongoldman

April 25, 2023 11:44 AM



Why Developers Are Flocking to LLaMA, Meta's Open Source LLM

does it have over GPT? We

What will

Goldman
@sharongoldman

ARTIFICIAL INTELLIGENCE / TECH

Stability AI announces new open-source large language model

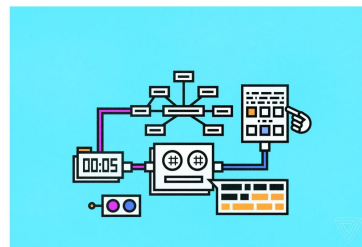


Illustration by Alex Castro / The Verge

Stability AI, the same company behind the AI image generator Stable Diffusion, is now open-sourcing its language model, StableLM.

By Emma Roth, a news writer who covers the streaming wars, consumer tech, crypto, social media, and much more. Previously, she was a writer and editor at MIT.

Apr 19, 2023, 9:21 PM GMT-1 | 2 Comments | 2 May



- ❖ THE STATE OF OPEN IN AI
- ❖ **LOOKING BACK OVER 20 YEARS**
- ❖ THE MULTIPLE FACES OF OPEN AI
- ❖ WHO DEFINES THE FUTURE OF OPEN AI?

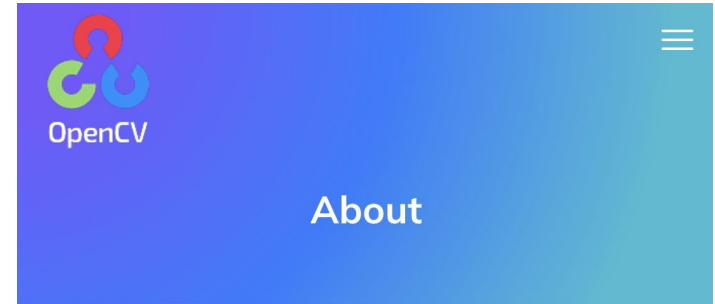
OPEN ML – 2000s



scikit-learn is a Python module for machine learning built on top of SciPy and is distributed under the 3-Clause BSD license.

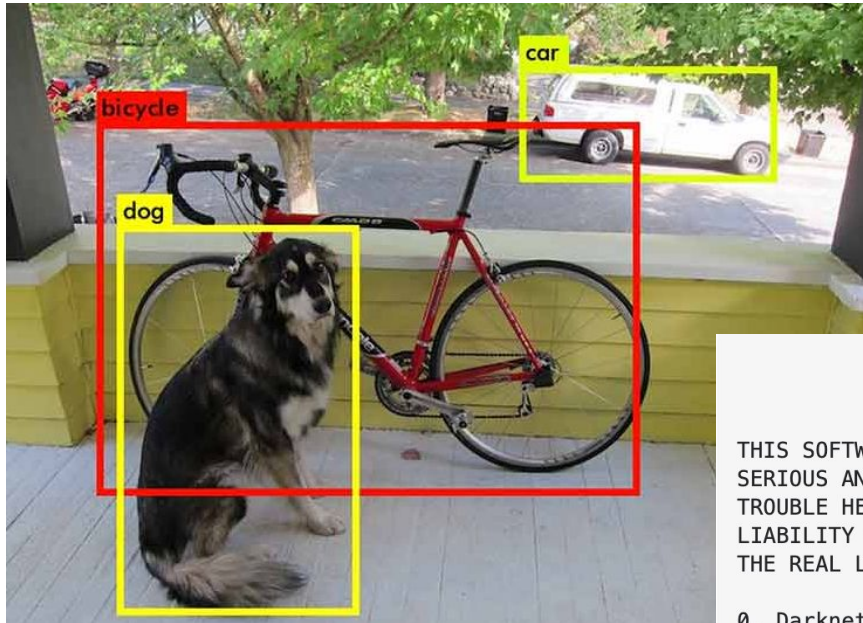
The project was started in 2007 by David Cournapeau as a Google Summer of Code project, and since then many volunteers have contributed. See the [About us](#) page for a list of core contributors.

It is currently maintained by a team of volunteers.



OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products. Being an Apache 2 licensed product, OpenCV makes it easy for businesses to utilize and modify the code.

OPEN ML - 2010s



YOLO LICENSE
Version 2, July 29 2016

THIS SOFTWARE LICENSE IS PROVIDED "ALL CAPS" SO THAT YOU KNOW IT IS SUPER SERIOUS AND YOU DON'T MESS AROUND WITH COPYRIGHT LAW BECAUSE YOU WILL GET IN TROUBLE HERE ARE SOME OTHER BUZZWORDS COMMONLY IN THESE THINGS WARRANTIES LIABILITY CONTRACT TORT LIABLE CLAIMS RESTRICTION MERCHANTABILITY. NOW HERE'S THE REAL LICENSE:

0. Darknet is public domain.
1. Do whatever you want with it.
2. Stop emailing me about it!

OPEN “ML”

- ML artefacts include **data, software, model, trained weights, outputs...**
- For DL, “downloading” a dataset or “running” a model is not accessible for the average person



Max Gruber / Better Images of AI / Ceci n'est pas une banane / CC-BY 4.0

OPEN ML LICENSING



The Turing Way

Search this book...

- Welcome
- [Guide for Reproducible Research](#) ^
- Overview v
- Open Research v
- Version Control v
- [Licensing](#) ^
- Software Licenses
- Data Licenses
- [Machine Learning Model Licenses](#) ^
- Case Studies: Choosing an ML License
- License Compatibility
- Checklist
- Research Data Management v
- Reproducible Environments v
- BinderHub v
- Code quality v
- Code Testing v
- Code Reviewing Process v
- Reusable Code v
- Continuous Integration (CI) v
- Reproducible Research with Make v
- Research Compendia
- Risk Assessment v

- Started with repurposing of MIT & Apache 2.0 licenses...

<https://the-turing-way.netlify.app/reproducible-research/licensing/licensing-ml.html>

Examples of ML models and their licenses

The table below showcases several well-known examples of ML models in the fields of NLP, vision, and multimodal generatives. The aim of it is to inform the reader on the licensing options chosen by each of the projects which sometimes differ from one another. Licensing differences might stem from business models, research purposes or ethics-informed community values. Each license carries licensor's values and a message from the former to potential users on how the model should be used.

Model	Model License	Description	Link to License
GPT-2	MIT License + generated output disclaimer	Permissive open source license	Link to license
GPT-3	Exclusive	Licensed to Microsoft	News Article
YOLO	YOLO License	Public domain license	Link to license
DALLE-pytorch	MIT License	Permissive open source license	Link to license
Stable Diffusion	CreativeML Open RAIL-M	Open & Responsible AI License (RAIL) created by Stability.ai and adapted from the BLOOM RAIL license, including use restrictions (see attachment A)	Link to license
OPT	OPT-175B License	Meta restrictive license enabling use of the model weights for research purposes while establishing a set of use restrictions, which could be considered a RAIL	Link to license
BigScience	BigScience OpenRAIL-M	Open & Responsible AI License (RAIL) created by BigScience and adapted from the BLOOM RAIL license, including use restrictions (see attachment A)	Link to license
Tsinghua University	GLM-130B license	Restrictive license enabling use of the model weights for research purposes	Link to license

OPEN ML LICENSING



The Turing Way

Search this book...

Welcome

[Guide for Reproducible Research](#) ^

Overview v

Open Research v

Version Control v

[Licensing](#) ^

Software Licenses

Data Licenses

[Machine Learning Model Licenses](#) ^

Case Studies: Choosing an ML License

License Compatibility

Checklist

Research Data Management v

Reproducible Environments v

BinderHub v

Code quality v

Code Testing v

Code Reviewing Process v

Reusable Code v

Continuous Integration (CI) v

Reproducible Research with Make v

Research Compendia

Risk Assessment v

Examples of ML models and their licenses

The table below showcases several well-known examples of ML models in the fields of NLP, vision, and multimodal generatives. The aim of it is to inform the reader on the licensing options chosen by each of the projects which sometimes differ from one another. Licensing differences might stem from business models, research purposes or ethics-informed community values. Each license carries licensor's values and a message from the former to potential users on how the model should be used.

Model	Model License	Description	Link to License
GPT-2	MIT License + generated output disclaimer	Permissive open source license	Link to license
GPT-3	Exclusive	Licensed to Microsoft	News Article
YOLO	YOLO License	Public domain license	Link to license
DALLE-pytorch	MIT License	Permissive open source license	Link to license
Stable Diffusion	CreativeML Open RAIL-M	Open & Responsible AI License (RAIL) created by Stability.ai and adapted from the BLOOM RAIL license, including use restrictions (see attachment A)	Link to license
OPT	OPT-175B License	Meta restrictive license enabling use of the model weights for research purposes while establishing a set of use restrictions, which could be considered a RAIL	Link to license
BigScience	BigScience OpenRAIL-M	Open & Responsible AI License (RAIL) created by BigScience and adapted from the BLOOM RAIL license, including use restrictions (see attachment A)	Link to license
Tsinghua University	GLM-130B license	Restrictive license enabling use of the model weights for research purposes	Link to license

<https://the-turing-way.netlify.app/reproducible-research/licensing/licensing-ml.html>

Open

Reusable

Accessible

Free

Open

CC0

Publicly
Available

Transparent

Reusable

Accessible

Free

Open

CC0

Publicly

Transparent

Available

Ethical

Responsible

OPEN...

- Beyond “code available on Github”
- Beyond “anyone can use it... can make money from it... can harm people with it...”

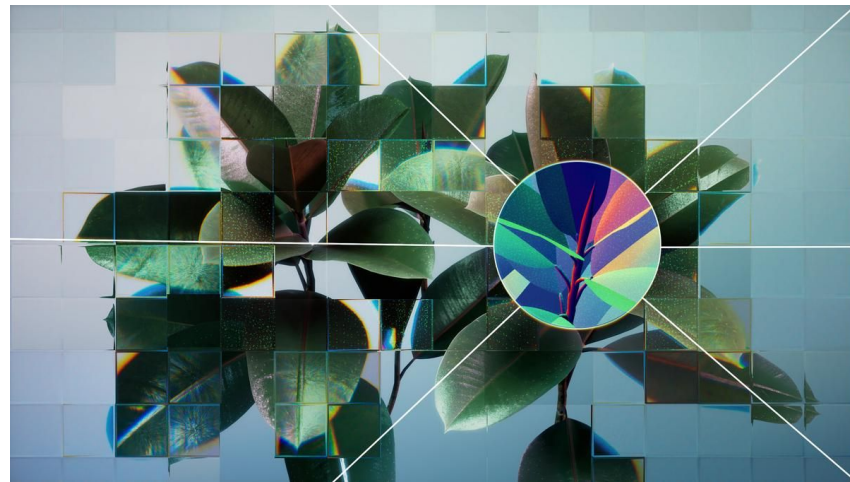


Image by Alan Warburton / © BBC / Better Images of AI / Plant / CC-BY 4.0

- ❖ THE STATE OF OPEN IN AI
- ❖ LOOKING BACK OVER 20 YEARS
- ❖ **THE MULTIPLE FACES OF OPEN AI**
- ❖ WHO DEFINES THE FUTURE OF OPEN AI?

OPEN FOR FASTER INNOVATION



scikit-learn is a Python module for machine learning built on top of SciPy and is distributed under the 3-Clause BSD license.

The project was started in 2007 by David Cournapeau as a Google Summer of Code project, and since then many volunteers have contributed. See the [About us](#) page for a list of core contributors.

It is currently maintained by a team of volunteers.



[Home](#) [Download](#) [Challenges](#) [About](#)

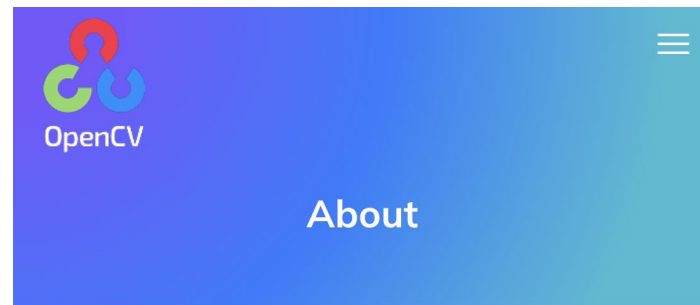
ImageNet Large Scale Visual Recognition Challenge (ILSVRC)

Competition

The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) evaluates algorithms for object detection and image classification at large scale. One high level motivation is to allow researchers to compare progress in detection across a wider variety of objects -- taking advantage of the quite expensive labeling effort. Another motivation is to measure the progress of computer vision for large scale image indexing for retrieval and annotation.

14,197,122 images, 21841 synsets indexed

Not logged in. [Login](#) | [Signup](#)



OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was built to provide a common framework for computer vision applications and to ease the use of machine perception in the commercial world. Being an Apache 2 licensed product, OpenCV is available for businesses to utilize and modify the code.

OPEN FOR COMMUNITY BUILDING



The AI community building the future.

Build, train and deploy state of the art models powered by the reference open source in machine learning.

Spaces

Discover amazing ML apps made by the community!

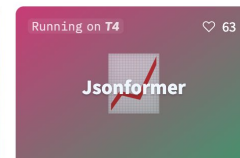
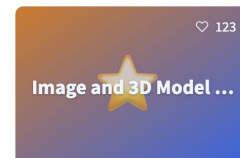
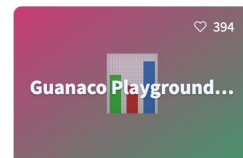
Create new Space or [learn more about Spaces.](#)

Search Spaces

new Full-text search

Sort: Recently Updated

Spaces of the week 🔥



Spaces: [huggingchat/chat-ui](#) like 541

Running on CPU UPGRADE

13 days ago

[mishig](#)

18 days ago

App Files Community 189

New discussion

New pull request

All Discussions Pull requests Show closed (53)

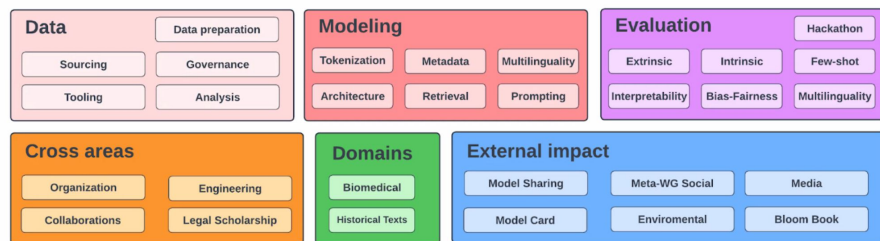
Feedback on the web search feature #185 opened 8 days ago by [nsamabin](#) 8

Future features #176 opened 16 days ago by [victor](#) 25

chat-ui is now available on Github! #146 opened 30 days ago by [victor](#)

OPEN FOR BROADENING PARTICIPATION

BigScience Workshop Working Groups



BigScience Large Open-science Open-access Multilingual Language Model
Version 1.3 / 6 July 2022



BigCode

About the project

How to join?

Search docs...

ABOUT

- The Mission
- Organization
- How we manage IP
- Code of Conduct
- How to join?
- The Stack
- Press

PAGES

How to join?

We are excited to invite AI practitioners from diverse backgrounds to join the BigCode project! Note that BigCode is a *research collaboration* and is open to participants who

1. have a professional research background and
2. are able to commit time to the project.

In general, we expect applicants to be affiliated with a research organization (either in academia or industry) and work on the technical/ethical/legal aspects of LLMs for coding applications.

You can [apply here](#) to the BigCode project!

OPEN FOR MOAT & MARKET DOMINANCE

Google "We Have No Moat, And Neither Does OpenAI"

Leaked Internal Google Document Claims Open Source AI Will Outcompete Google and OpenAI



DYLAN PATEL AND AFZAL AHMAD

MAY 4, 2023 • PAID

Stability AI, the startup behind Stable Diffusion, raises \$101M

Kyle Wiggers @kyle_l_wiggers / 6:01 PM GMT+1 • October 17, 2022

OPEN AS A RESPONSE TO POWER

- Open to “**benefit me so I can sell to or control thee**” seems to contradict the ethos
- “Open has been co-opted because it’s been great for capitalism, but my own understanding of open source is as a political response to power.”
 - **Bastian Greshake Tzovaras**
(*Open Humans*)

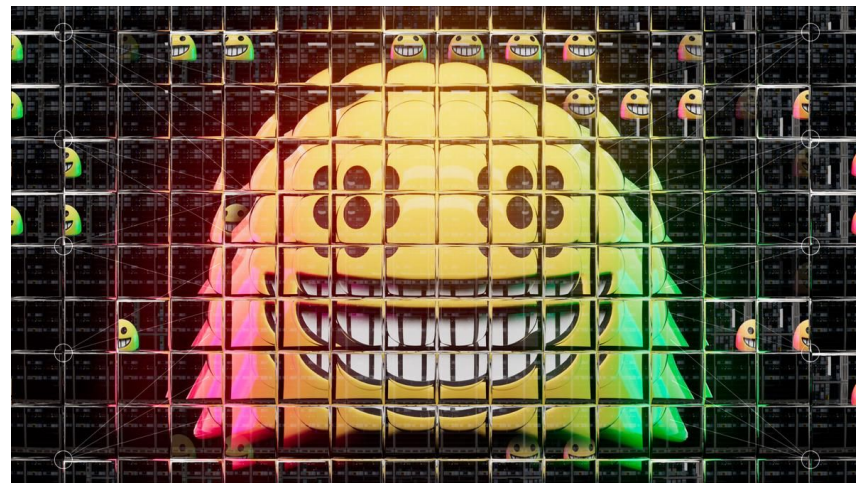


Image by Alan Warburton / © BBC / Better Images of AI / Social Media / CC-BY 4.0

slido



What directions of "open AI" resonate the most with you?

① Click **Present with Slido** or install our [Chrome extension](#) to activate this poll while presenting.

- ❖ THE STATE OF OPEN IN AI
- ❖ LOOKING BACK OVER 20 YEARS
- ❖ THE MULTIPLE FACES OF OPEN AI
- ❖ **WHO DEFINES THE FUTURE OF OPEN AI?**

Ursula von der Leyen  @vonderleyen

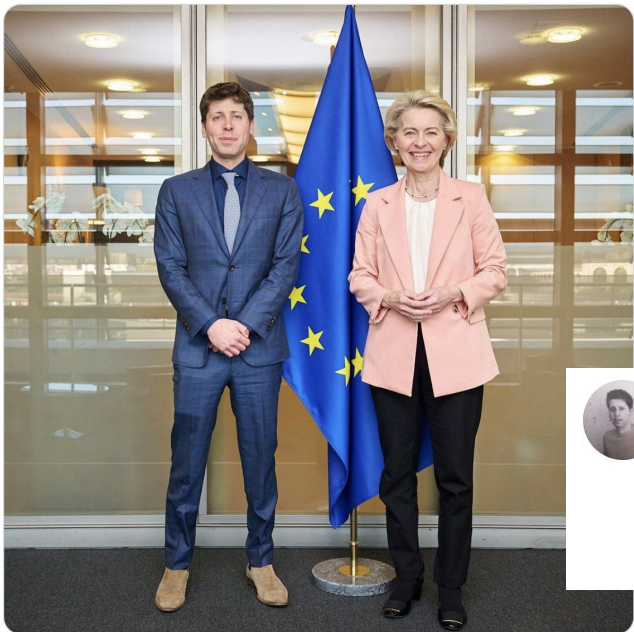
Glad to meet [@OpenAI](#) CEO [@sama](#)

AI can fuel huge progress and improve our lives.

But we must mitigate risks and build trust.

To match the speed of tech development, AI firms need to play their part.

EU will work with global partners and stakeholders towards trustworthy AI.



Sam Altman  @sama · May 22

had a great first week of the openai world tour in toronto, DC, rio, lagos, and lisbon.

fun to see what people are building and get (lots of) feature requests, and even fun to talk to policymakers!

madrid, warsaw, paris, london, and munich this week.



Sam Altman  @sama · Jun 4

excited to visit israel, jordan, qatar, the uae, india, and south korea this week!

 756

 493

 8,513

 982.9K



**WHO ELSE CAN
DEFINE THE
“OPEN” IN “OPEN
AI” AND HOW?**

OPEN COLLABORATIVES

- **Distributed, volunteer-led teams creating alternative AI development pathways grounded in “open” principles**
- Opportunity for more people from different backgrounds around the world to participate and shape AI production

Towards Openness Beyond Open Access: User Journeys through 3 Open AI Collaboratives

Jennifer Ding
The Alan Turing Institute
jding@turing.ac.uk

Christopher Akiki
Leipzig University
christopher.akiki@uni-leipzig.de

Yacine Jernite
Hugging Face
yacine@huggingface.co

Anne Lee Steele
The Alan Turing Institute
asteel@turing.ac.uk

Temí Popo
Mozilla Foundation
temi@mozillafoundation.org

User journeys through open AI communities



OPEN COLLABORATIVES

- Distributed, volunteer-led teams creating alternative AI development pathways grounded in “open” principles
- **Opportunity for more people from different backgrounds around the world to participate and shape AI production**

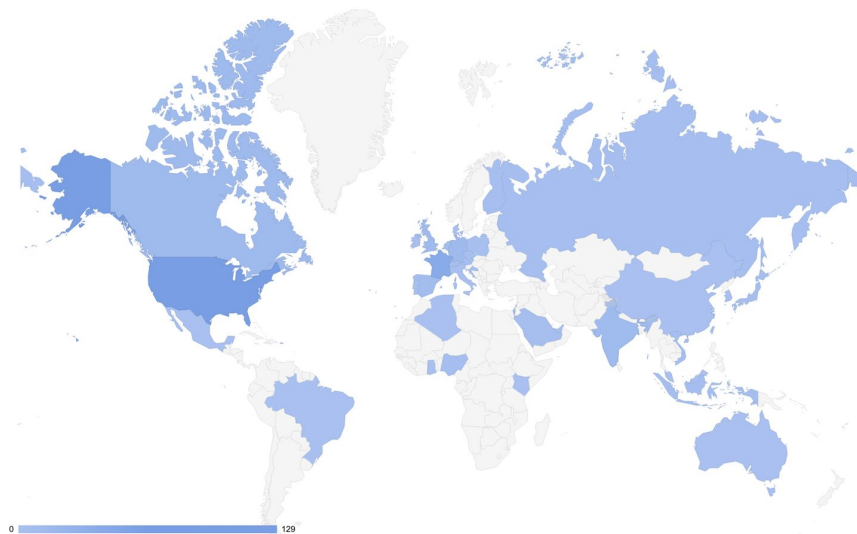


Figure 1: Geographic location of residence for 308 BigScience participants with a at least one traced contribution. This corresponds to 38 countries. (See Section 3 for more information.)

OPEN CONVERSATIONS



WORKSHOP

Designing for Data Rights in the AI Production Pipeline

EN Allies in Practice Data Stewardship Recorded

Building trustworthy AI requires building public trust in how AI is developed. While the majority of AI production/resources are concentrated within a few companies in even fewer countries, alternative spaces are emerging for more people to participate in creating, applying, and governing data and data-generated ML models. New initiatives such as BigScience and BigCode seek to change extractive methods of AI production, replacing secretive web scraping with data stewardship and other data rights-affirming tools, practices, and systems.

SCHEDULE

PAST

02:00 PM → 03:00 PM

March 24

Europe/London

LINKS

Miro Board

miro.com/app/board/uXjVMeuvLR8=?share_link_id=159151239611

copy

FACILITATORS



Anne Lee Steele

The Alan Turing Institute



Jennifer Ding

The Alan Turing Institute



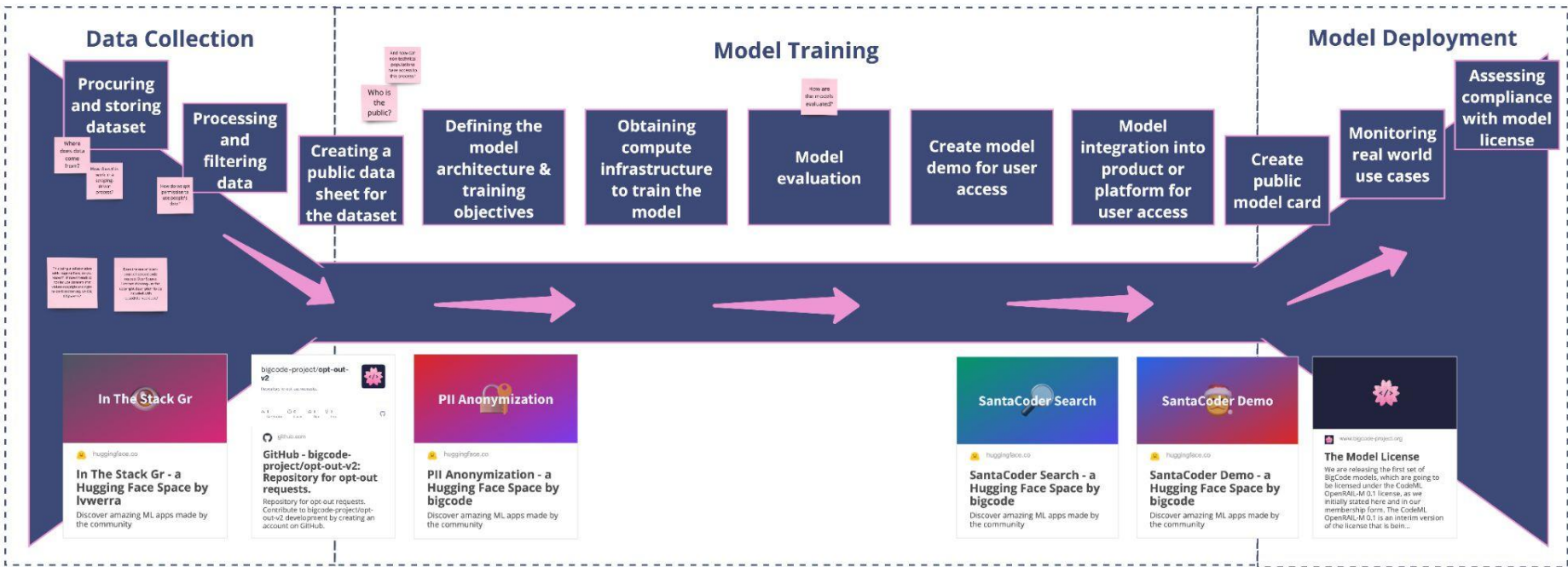
Yacine Jernite

Hugging Face

https://miro.com/app/board/uXjVMeuvLR8=?share_link_id=159151239611

BigCode Pipeline

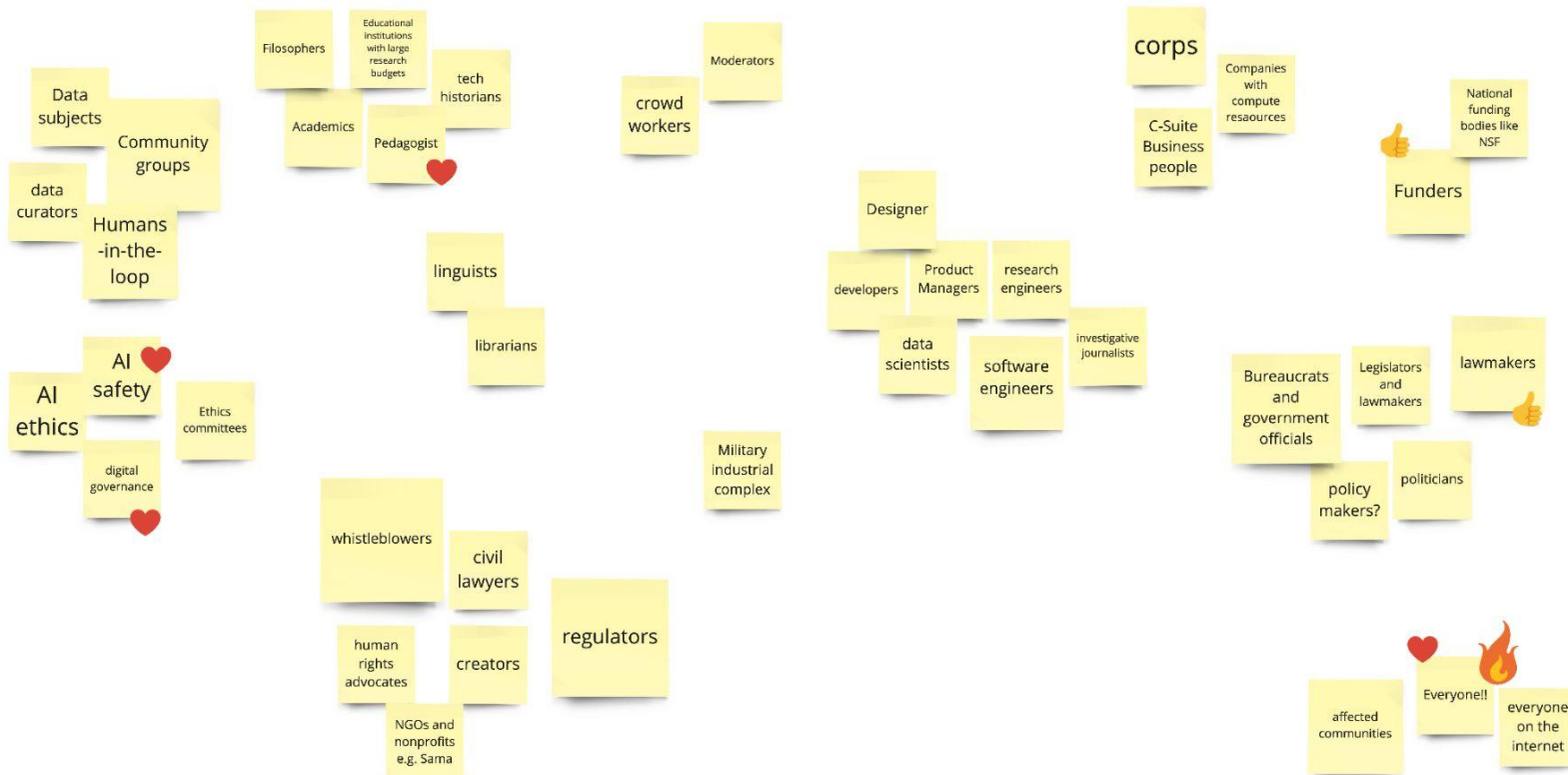
case study: BigCode



What questions, comments, or concerns do you have about the BigCode Pipeline?

Who is part of or can shape the AI production pipeline?

who

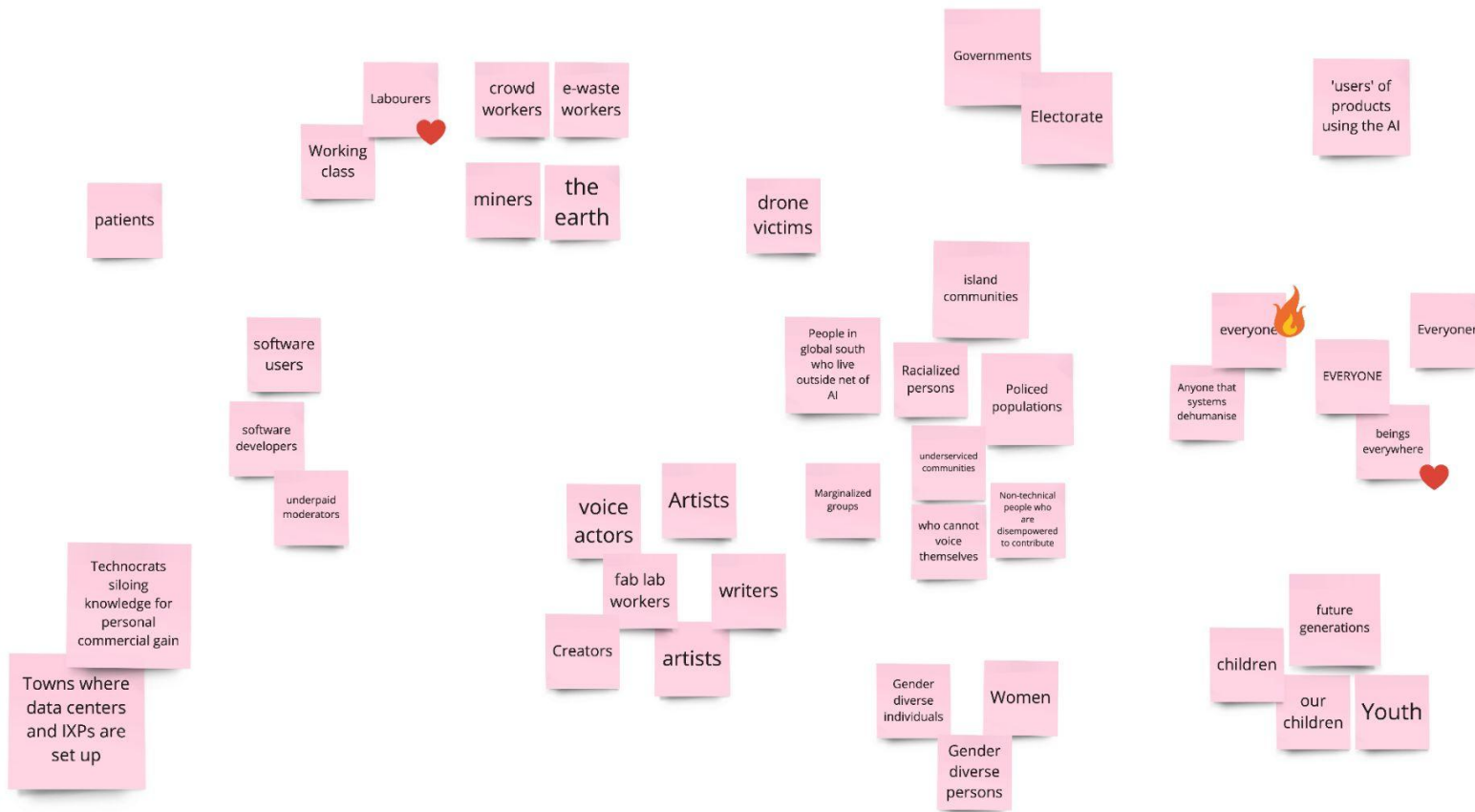


Drag and duplicate if you like what someone else wrote



Who is impacted by the AI production pipeline?

who



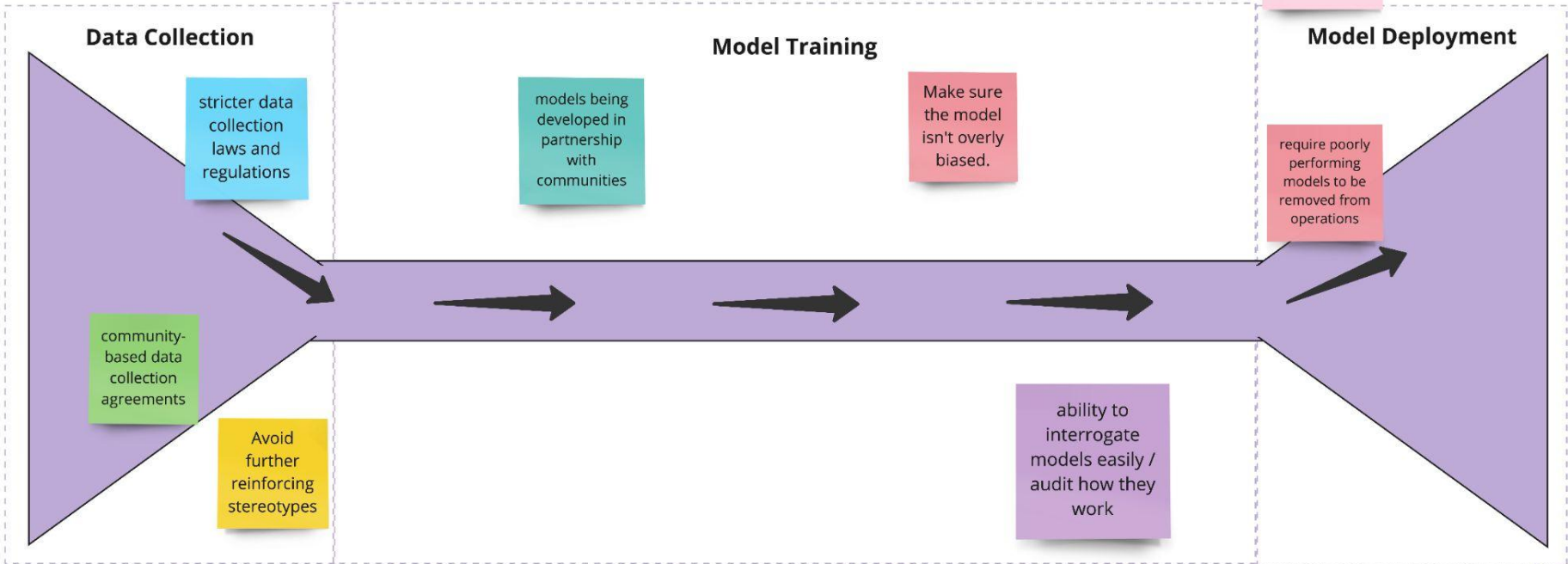
Drag and duplicate if you like what someone else wrote



MozFest 2023 // Allies in Practice

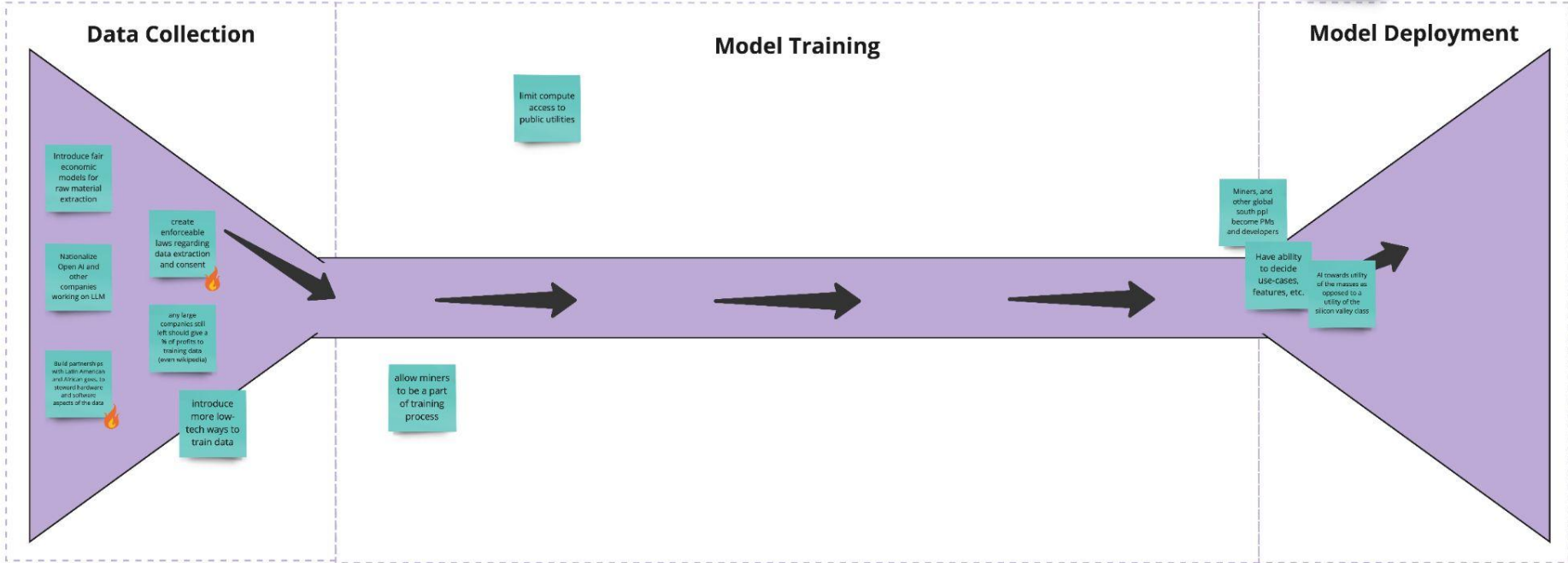
What kinds of activities *should be* a part of the AI production pipeline for/with

Policed populations ?



What kinds of activities *should be* a part of the AI production pipeline for/with

miners ?



slido



Who should be part of defining the future of open AI?

① Click **Present with Slido** or install our [Chrome extension](#) to activate this poll while presenting.

OPEN...

- For diverse, global participation in AI
- To address the power imbalance in AI and empower those outside the tech bubble
- To create new forms of AI and AI production that can benefit more



Image by Alan Warburton / © BBC / Better Images of AI / Nature / CC-BY 4.0

WHAT DEFINES THE “OPEN” IN “OPEN AI”?

jen ding
bbuzz 2023

jd@turing.ac.uk

WHAT DEFINES THE "OPEN" IN "OPEN AI"?

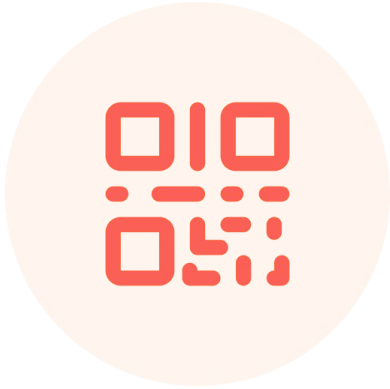
jen ding
bbuzz 2023

jdilling@turing.ac.uk



londondataweek.org

slido



Join at slido.com
#1516847

① Click **Present with Slido** or install our [Chrome extension](#) to display joining instructions for participants while presenting.