ETH zürich

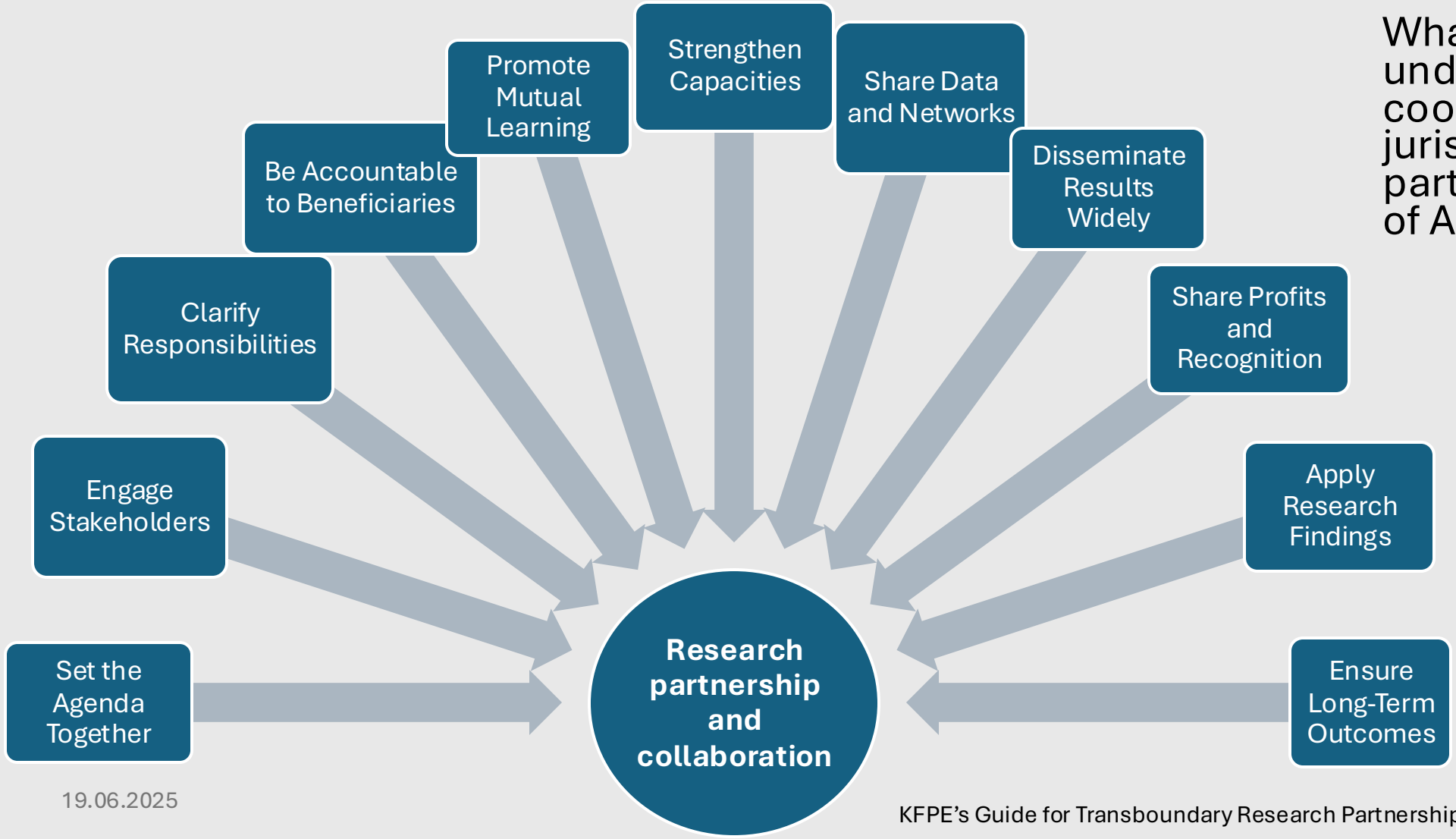
Challenges and Ethical Questions Related to AI Driven Research

Dr. Kebene Wodajo
20 June 2025

Outline

- Background
- What does it mean to undertake and coordinate cross-jurisdictional research partnerships in the age of AI systems?
- Thinking and planning through the AI value chain
- A collaborative, value chain approach
- Why use the AI value chain approach?

Background



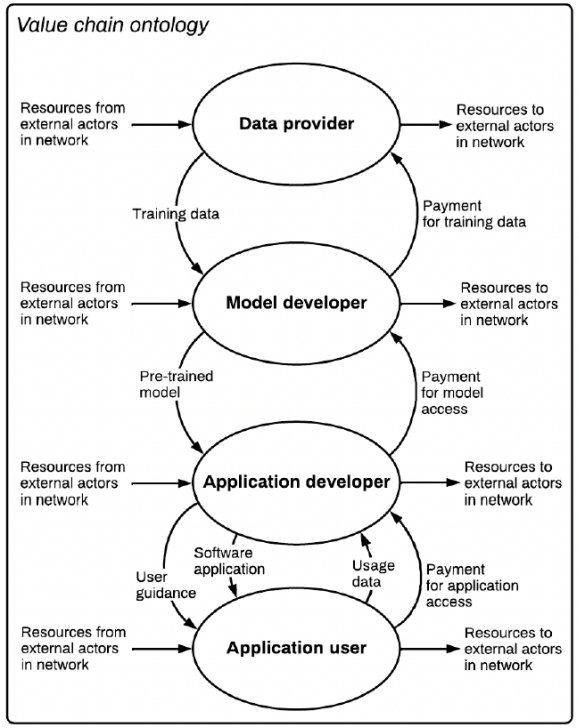
What does it mean to undertake and coordinate cross-jurisdictional research partnerships in the age of AI systems?

What does it mean to undertake and coordinate cross-jurisdictional research partnerships in the age of AI systems?

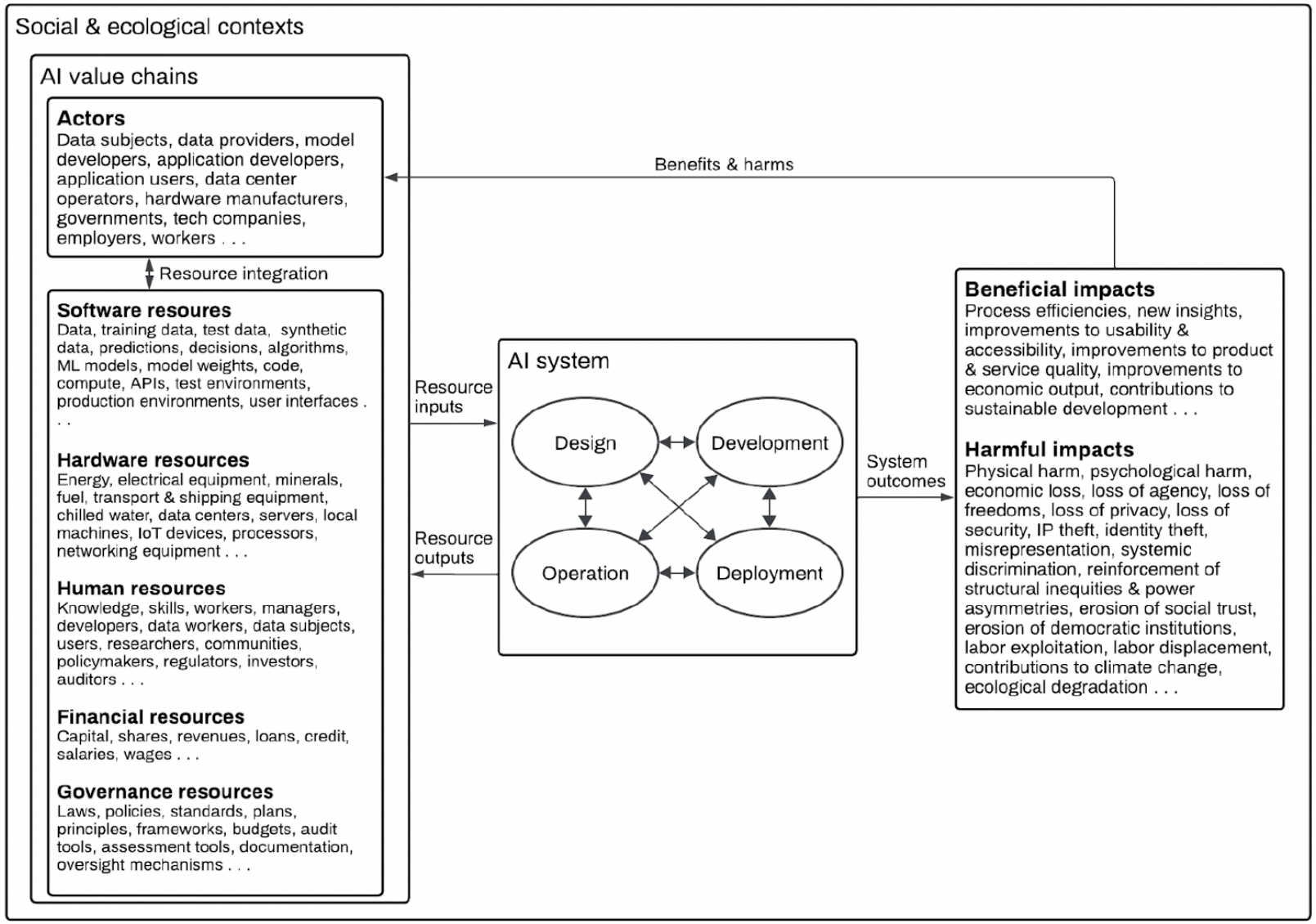
- With the increasing reliance on **AI systems** both as a **“tool” for conducting research** and as a **solution or research output,** unique ethical, practical, social, contextual relevance etc. concerns emerge.

- The advent of **generative AI and large language models (LLMs)** amplifies the need for **critical reflection** on how we design, manage, and evaluate research partnerships.

Thinking and planning through the AI value chain



ıain ontology of AI systems contrasted with a value
ɜms. The actors and resources that appear here are
differences between these two perspectives, not to
·esentative or exhaustive view of the actors and
ɔurces involved in AI systems.

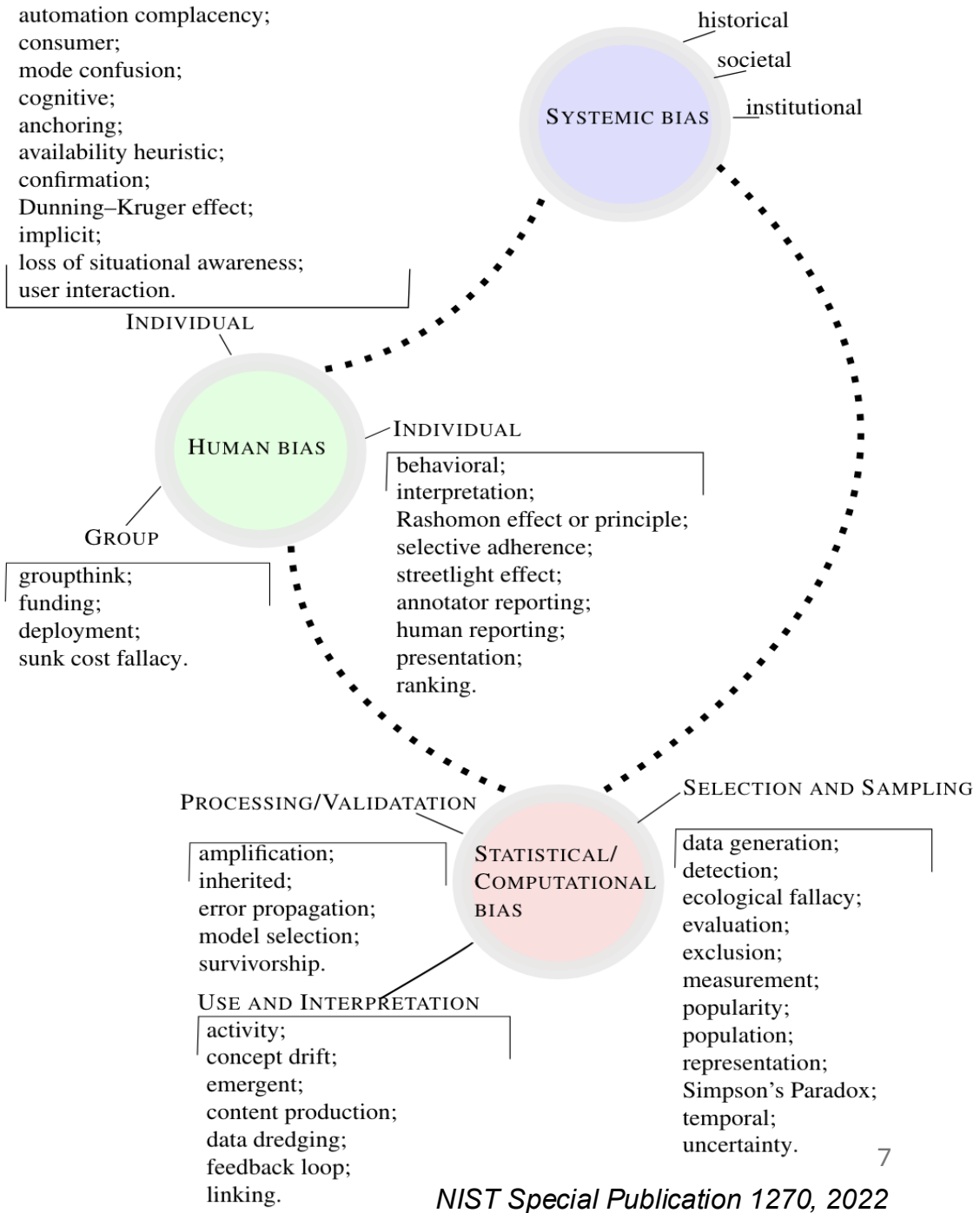


Attard-Frost & Widder, The Ethics of AI Value Chains 2024

Figure 2: The process through which many types of actors situated within AI value chains cause beneficial or harmful impacts by providing resources

Bias and discrimination

- Bias can be introduced *purposefully or inadvertently* into an AI system, or it can emerge as the AI is used in an application:
  - Statistical and computational biases
  - Systemic biases
  - Human biases



NIST Special Publication 1270, 2022

Carbon footprint

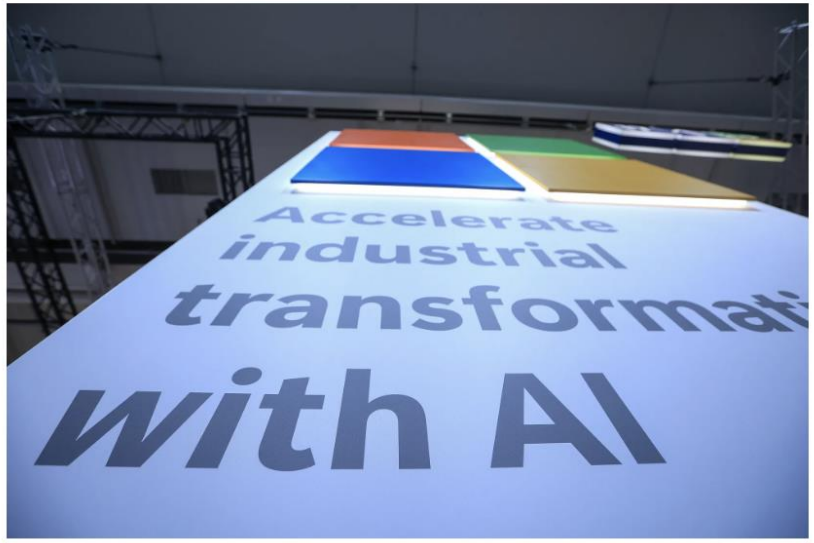
"OpenAI's GPT-3 and Meta's OPT were estimated to emit more than 500 and 75 metric tons of carbon dioxide, respectively, during training. GPT-3's vast emissions can be **partly explained by the fact that it was trained on older, less efficient hardware**. But it is hard to say what the figures are for certain; … and these figures are based on external estimates or, in Meta's case, limited data the company released."

Heikkilä, *We're getting a better idea of AI's true carbon footprint*, MIT Review, Nov. 14, 2022

ch | AI Glossary | Model Collapse Risk | AI's Real Carbon Footprint | Small AI Models

Microsoft's AI Push Imperils Climate Goal as Carbon Emissions Jump 30%

The company's goal to be carbon negative by 2030 is harder to reach, but President Brad Smith says the good AI can do for the world will outweigh its environmental impact.



Microsoft's push for to capitalize on artificial intelligence has made it the world's most valuable company. *Photographer: Krisztian Bocsi/Bloomberg*

By Akshat Rathi and Dina Bass
May 15, 2024 at 6:00 PM GMT+2

Save | Translate | Listen 6:43

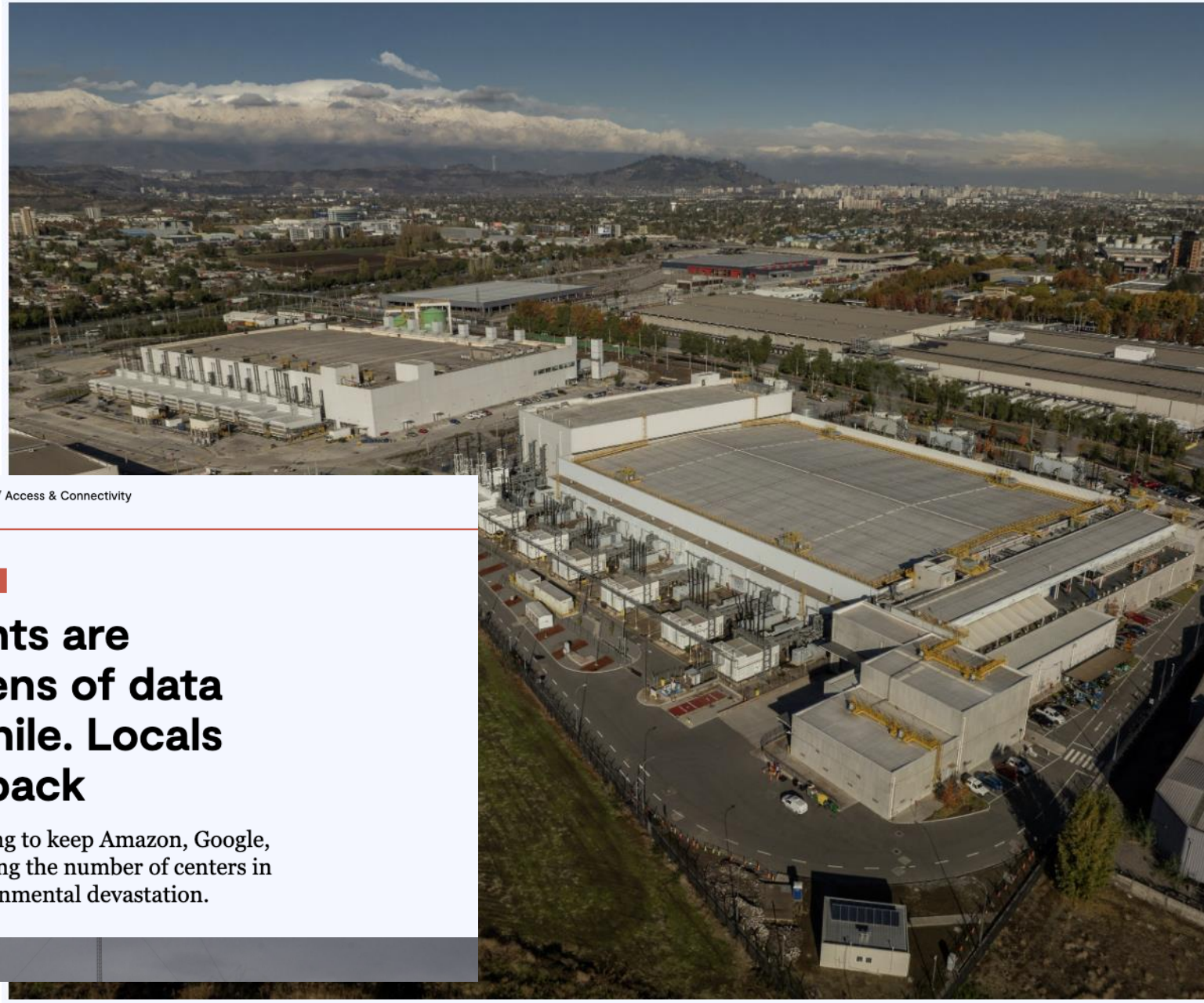
When Microsoft Corp. pledged four years ago to remove more carbon than it emits by the end of the decade, it was one of the most ambitious and comprehensive plans to tackle climate change. Now the software giant's relentless push to be the global leader in artificial intelligence is putting that goal in peril.

Bloomberg May 15, 2024

Emissions from infrastructure

rĕst ŏf wo̊rld

Latin America / U.S tech giants are building dozens of data centers in Chile. Locals are fighting back



The Google data center in Quilicura opened in 2015. More than 10 data centers have since been approved to operate in the Santiago metropolitan area.

rĕst ŏf wo̊rld
Reporting Global Tech Stories

Latin America / Access & Connectivity

SILICON VALLEY IN REST OF WORLD

U.S tech giants are building dozens of data centers in Chile. Locals are fighting back

Multiple groups are working to keep Amazon, Google, and Microsoft from doubling the number of centers in the country, fearing environmental devastation.

Rest of world, May 2024

Outsourcing of data labeling work/activities

- Case:
  - OpenAI's outsourcing to workers employed by Sama AI in Kenya

TIME SIGN UP FOR OUR IDEAS NEWSLETTER POV

To build that safety system, OpenAI took a leaf out of the playbook of social media companies like Facebook, who had already shown it was possible to build AIs that could detect toxic language like hate speech to help remove it from their platforms. The premise was simple: feed an AI with labeled examples of violence, hate speech, and sexual abuse, and that tool could learn to detect those forms of toxicity in the wild. That detector would be built into ChatGPT to check whether it was echoing the toxicity of its training data, and filter it out before it ever reached the user. It could also help scrub toxic text from the training datasets of future AI models.

To get those labels, OpenAI sent tens of thousands of snippets of text to an outsourcing firm in Kenya, beginning in November 2021. Much of that text appeared to have been pulled from the darkest recesses of the internet. Some of it described situations in graphic detail like child sexual abuse, bestiality, murder, suicide, torture, self harm, and incest.

OpenAI's outsourcing partner in Kenya was Sama, a San Francisco-based firm that employs workers in Kenya, Uganda and India to label data for Silicon Valley clients like Google, Meta and Microsoft. Sama markets itself as an "ethical AI" company and claims to have helped lift more than 50,000 people out of poverty.

Bluesky Gains Million Users Since Election as People Flock From X

BUSINESS

Wanjira Mathai: TIME100 Climate 2024

TIME SIGN UP FOR OUR IDEAS NEWSLETTER POV

BUSINESS + TECHNOLOGY

Exclusive: OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic

15 MINUTE READ

TIME SIGN UP FOR OUR IDEAS NEWSLETTER POV

Sama's office in Nairobi, Kenya, on Feb. 10, 2022. Khadija Farah for TIME

The data labelers employed by Sama on behalf of OpenAI were paid a take-home wage of between around $1.32 and $2 per hour depending on seniority and performance. For this story, TIME reviewed hundreds of pages of internal Sama and OpenAI documents, including workers' payslips, and interviewed four Sama employees who worked on the project. All the employees spoke on condition of anonymity out of concern for their livelihoods.

The story of the workers who made ChatGPT possible offers a glimpse into the conditions in this little-known part of the AI industry, which nevertheless plays an essential role in the effort to make AI systems safe for public consumption. "Despite the foundational role played by these data enrichment professionals, a growing body of research reveals the precarious working conditions these workers face," says the Partnership on AI, a coalition of AI organizations to which OpenAI belongs. "This may be the result of efforts to hide AI's dependence on this large labor force when celebrating the efficiency gains of technology. Out of sight is also out of mind." (OpenAI does not disclose the names of the outsourcers it partners with, and it is not clear whether OpenAI worked with other data labeling firms in addition to Sama on this project.)

E-waste - how does it relate to the use of AI in research?

unitar
United Nations Institute for Training and Research

SEARCH | ACCESS YOUR COURSES | OUR MAIN OFFICES | ENGLISH

INFORMAL E-WASTE RECYCLING, GHANA, 2023 © MUNTAKA CHASANT / FONDATION CARMIGNAC.

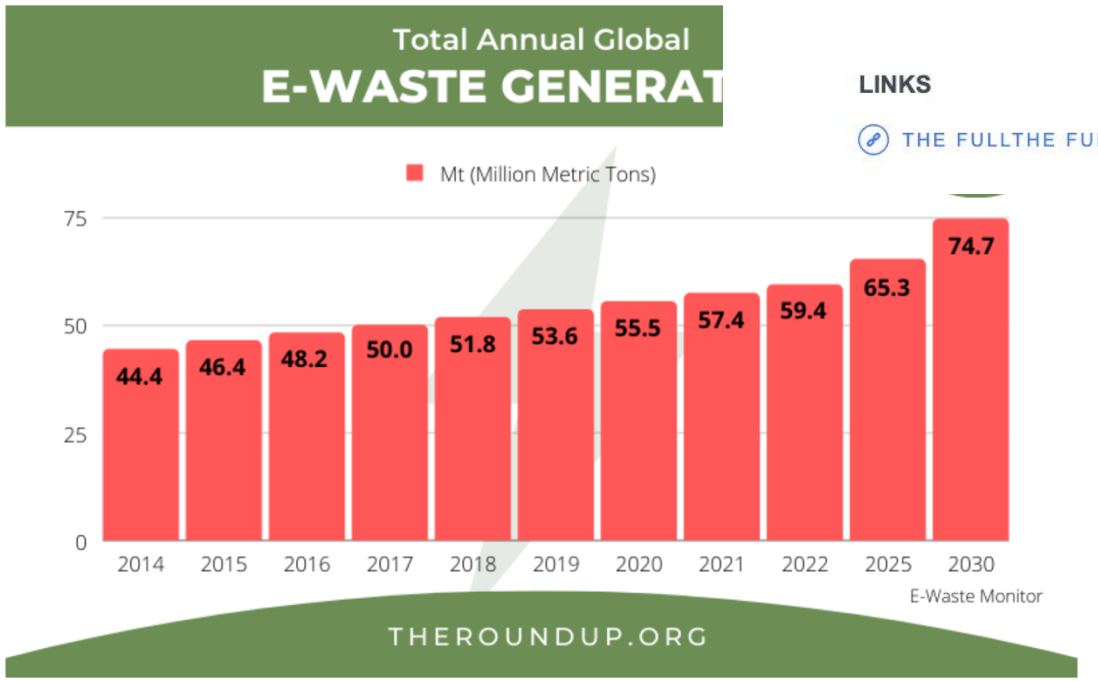

© Muntaka Chasant / Fondation Carmignac.

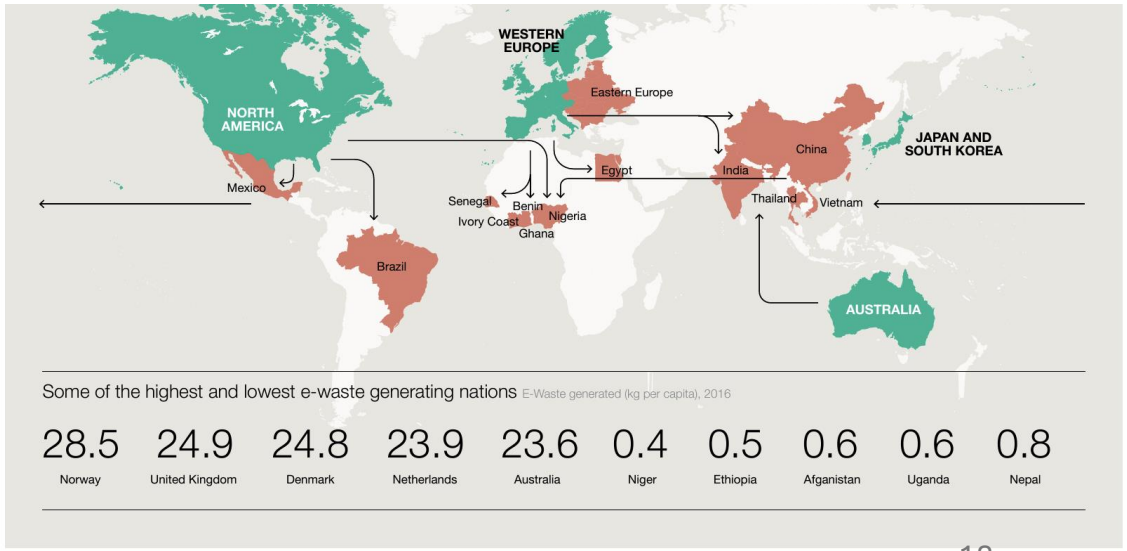

© Muntaka Chasant for Fondation Carmignac


© Muntaka Chasant for Fondation Carmignac

LINKS

THE FULLTHE FULL GLOBAL E-WASTE MONITOR REPORT IS AVAILABLE TO DOWNLOAD.





TheRoundup.org, Apr. 14, 2024

Source: Lewis 2011, The Global E-waste Statistics Partnership, 2018

Critical minerals and raw materials

AMNESTY INTERNATIONAL ENGLISH WHO WE ARE WHAT WE DO

< NEWS

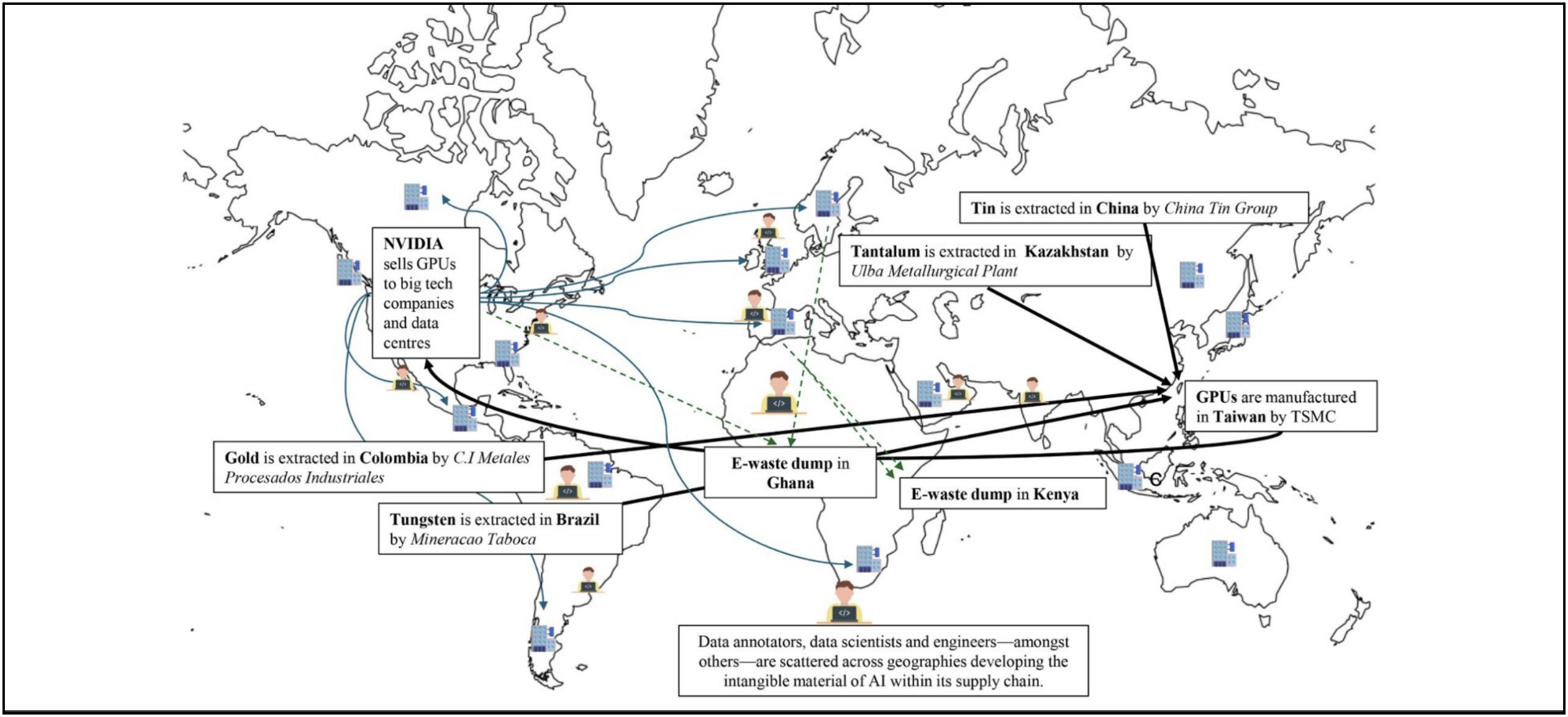
January 19, 2016

Exposed: Child labour behind smart phone and electric car batteries

Major electronics brands, including Apple, Samsung and Sony, are failing to do basic checks to ensure that cobalt mined by child labourers has not been used in their products, said Amnesty International and Afrewatch in a report published today.

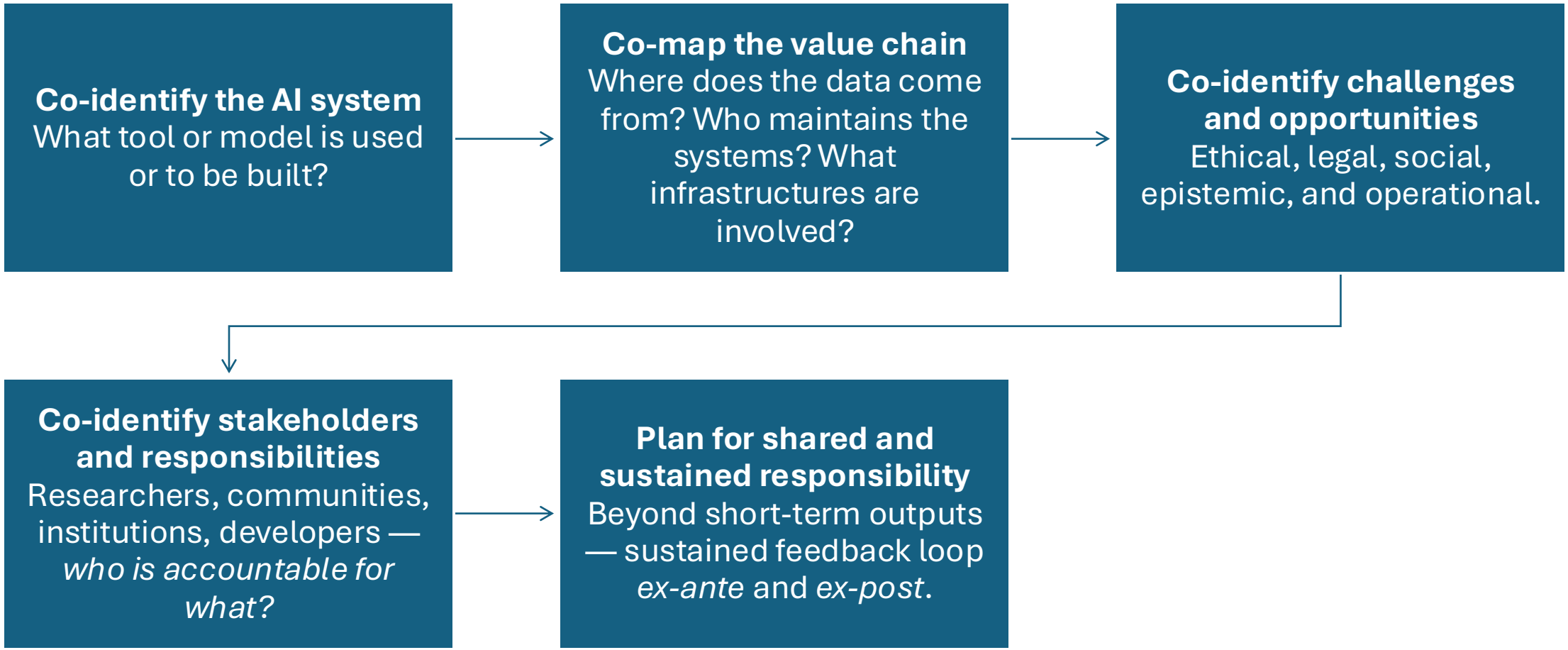
The report, This is what we die for: Human rights abuses in the Democratic Republic of the Congo power the global trade in cobalt, traces the sale of cobalt, used in lithium-ion batteries, from mines where children as young as seven and adults work in perilous conditions.

> **“The glamourous shop displays and marketing of state of the art technologies are a stark contrast to the children carrying bags of rocks, and miners in narrow manmade tunnels risking permanent lung damage.**
>
> **Mark Dummett, Business & Human Rights Researcher at Amnesty International**



Valdivia, A. (2024). The *supply chain capitalism of AI*: a call to (re)think algorithmic harms and resistance through environmental lens. *Information, Communication & Society*, 1–17

A Collaborative, Value Chain Approach



Why Use the AI Value Chain Approach?

- Reveal areas for meaningful collaboration
  - Helps researchers to be context specific and consider issues that must be:
    - Addressed within the current project
    - Taken up in future research agendas

- Identifying key actors and collaboration strategies
  - Inform stakeholder-specific approaches to responsible research collaboration/partnership

Thank you.