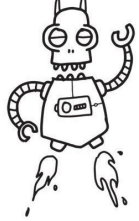
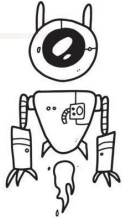
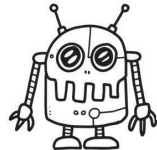
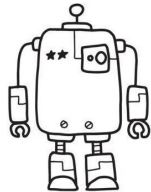
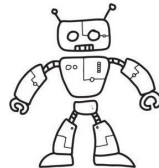
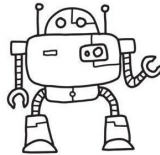
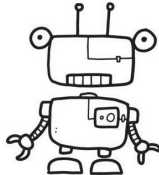
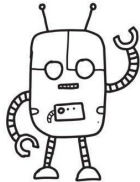
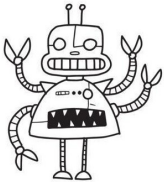


AI: from tech to public policy

Gaël Varoquaux

Inria



What got me here

- I got lucky with machine-learning code and research
 - Co-founded scikit-learn
 - Supervised Arthur Mensch
- Became cool (in 2023)
 - France IA commissionLearning a lot on geo-politico-economics

I don't like ties

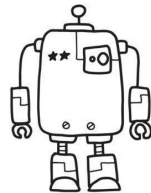
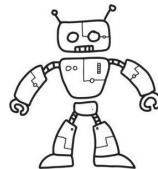
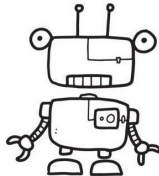
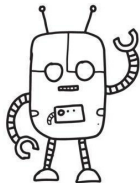
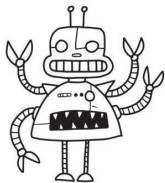
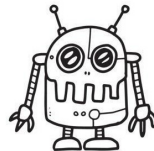
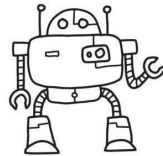
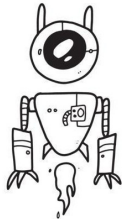
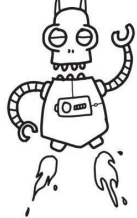


This talk: main messages of collaborative reports + personal views

French AI Commission Report, 2025 [Aghion... 2024]

International Scientific Report on the Safety of Advanced AI, 2025
[Bengio... 2024]

Hype, Sustainability, and the Price of the Bigger-is-Better
Paradigm in AI [Varoquaux... 2024]



The French AI commission

15 experts:

8 tech executives, 4 academics, 3 civil society

Philippe Aghion

Anne Bouverot

Gilles Babinet

Bernard Charlès

Cédric O

Joëlle Barral

Luc Julia

Isabelle Ryl

Alexandra Bensamoun

Yann Le Cun

Arthur Mensch

Nozha Boujemaa

Franca Salis-Madinier

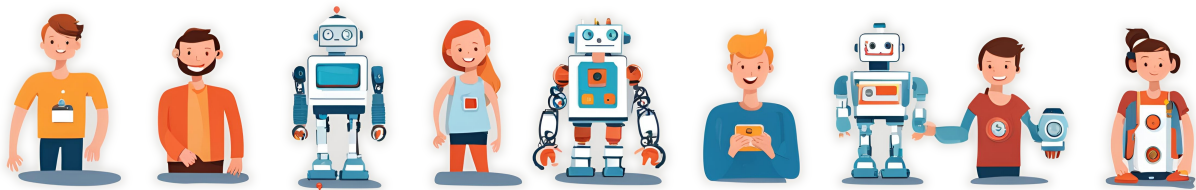
Martin Tisné

Gaël Varoquaux

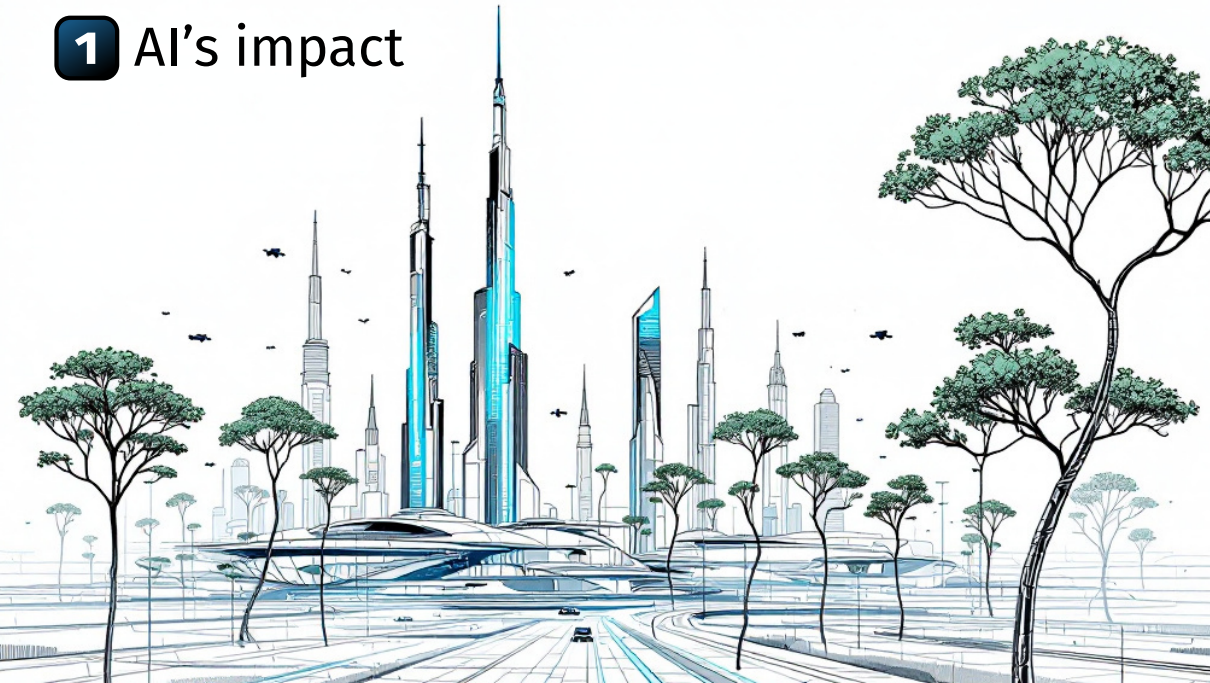


The report

- Briefly assesses AI's impact
- Related dangers
- Gives recommendations for collective appropriation



1 AI's impact



Inventions bring hopes and fears

In modern times, our societies
have been reshaped by
electricity, cars,
digital technologies...

Transformations brought
progress, and new dangers
some fantasized



A driver of a better world

AI is, in itself,
neither good nor bad

It depends what we,
as a society,
do of it



A driver of a better world

Evidence in AI for health

Radiology screening

■ Mammography reading
(FDA-approved^a)

■ Health-economics rational:

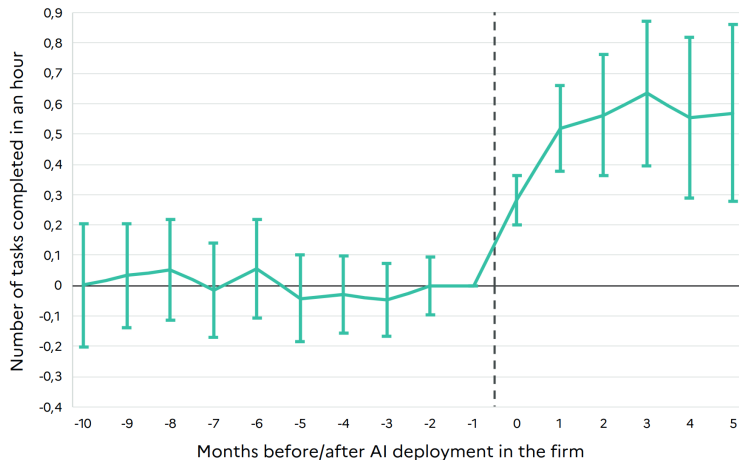
- Early-detected cancer curable
- Ultra-sound is cheap
- Doctor's reading is the bottleneck

^aFDA report on "Mammoscreen", 2020
<https://fda.report/PMN/K192854>



A driver of a better world

Micro-economic evidence in a customer service department



Adoption of GenAI

Brynjolfsson et al, 2023, NBER

But let us be mindful of inflated promises

<https://www.theverge.com/2016/8/23/12603624/delphi-mobileye-self-driving-autonomous-car-2019>

Delphi and Mobileye are teaming up to build a self-driving system by 2019



/ Major auto-suppliers joining forces to create Level 4 autonomous driving

By [Andrew J. Hawkins](#), transportation editor with 10+ years of experience who covers EVs, public transportation, and aviation. His work has appeared in The New York Daily News and City & State.

Aug 23, 2016, 5:07 PM GMT+2

[Link](#) [Facebook](#) [Twitter](#) | [0 Comments \(0 New\)](#)

Delphi and Mobileye, two leading auto-parts suppliers, announced today that they are forming a partnership to develop a near-complete autonomous driving system by 2019. The plan is to create a mass-market, off-the-shelf system that can be plugged into a variety of

But let us be mindful of inflated promises

<https://www.verdict.co.uk/fully-self-driving-cars-unlikely-before-2035-experts-predict/?cf-view&cf-closed>

Fully self-driving cars unlikely before 2035, experts predict

You won't be able to fall asleep at the wheel anytime soon.

Kurt Robson | March 9, 2023

Share this article



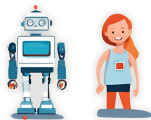
Promises

Generative AI startups =
\$22 billions investments in 2023

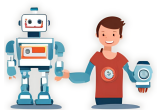
**There is a bubble
in AI**



Between overselling and underestimating



AI = A gradual evolution in technology,
a long term revolution in everyday experience



AI techniques are already ubiquitous



During the industrial revolutions, productivity gains were visible only 20 years after a technology was invented

We must create the best conditions for collective appropriation

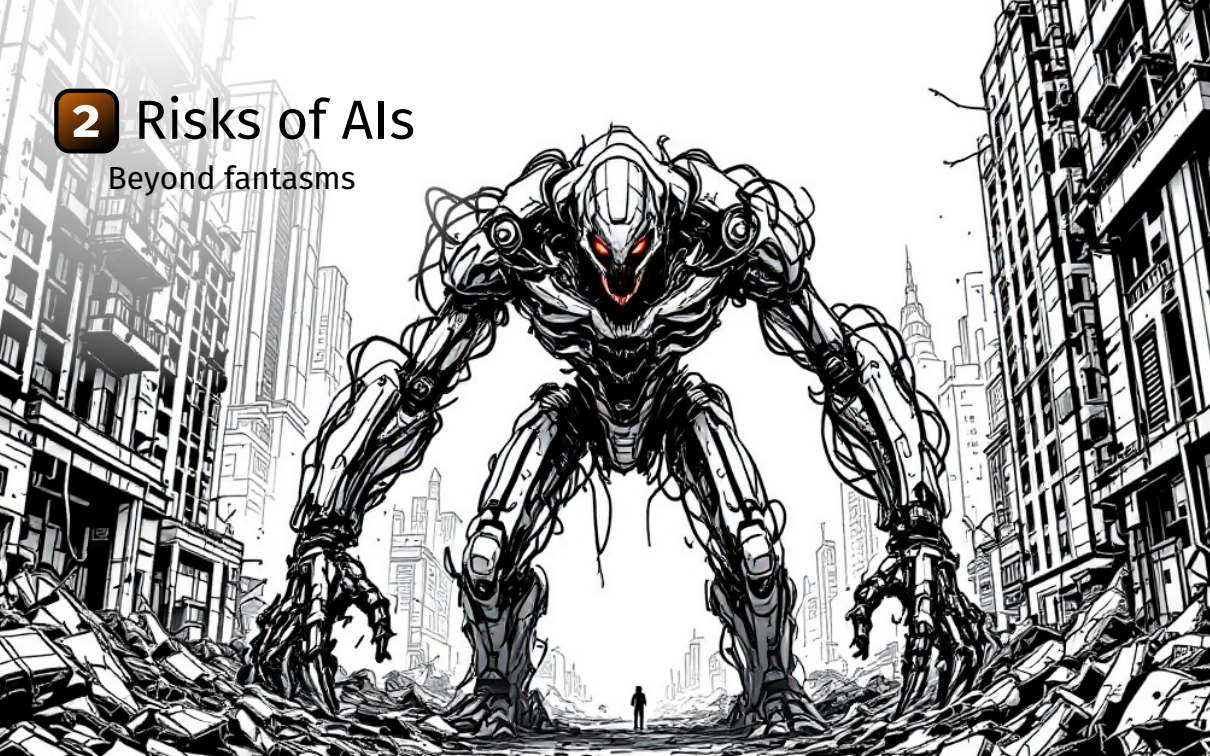
A need for trust

A healthy discussion including risks



2 Risks of AIs

Beyond fantasms



A risk analysis

Risks due to imperfections

Discrimination and stereotyping

Misinformation

Threats to privacy

Production of harmful content

Intellectual property violation

Risks due to malicious use

Cyber criminality

Cyber terrorism

Biosecurity

Mass surveillance

Systemic risks

Concentration of power

Disruption of labor market

Weakening of cultural diversity

Electricity and carbon footprint

Critical emergent behavior

Mis-information

Quality access to information
is the cement of
democratic societies

Threats from AI

- Filter bubbles
- Fake content generation
- Disrupting the economics of news



Threats to privacy

1. Centralized AI scoop up data
2. AI leak training data



Biased AIs discriminate

Biased algorithmic governance

- Many documented cases of harm
eg: Dutch childcare benefits scandal (W)
- Bias is a socio-technical problem
a shortcoming that doesn't annoy all

Perpetuated in language models

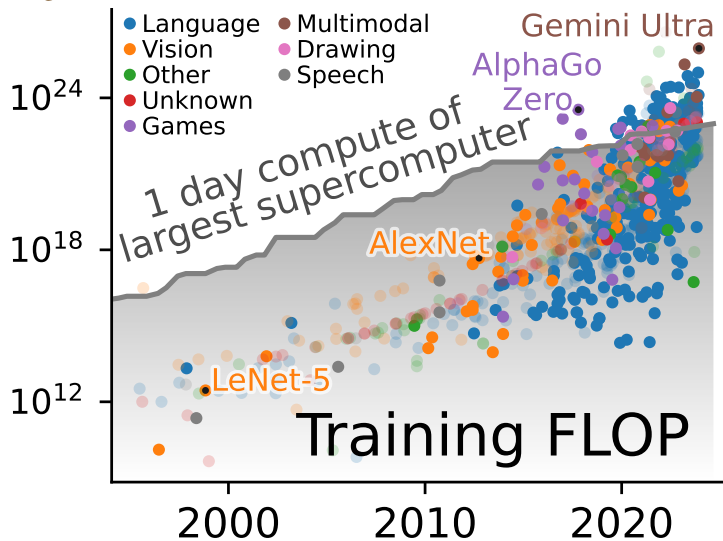
Represent less well South-East Asians

[Chen et al 2024]



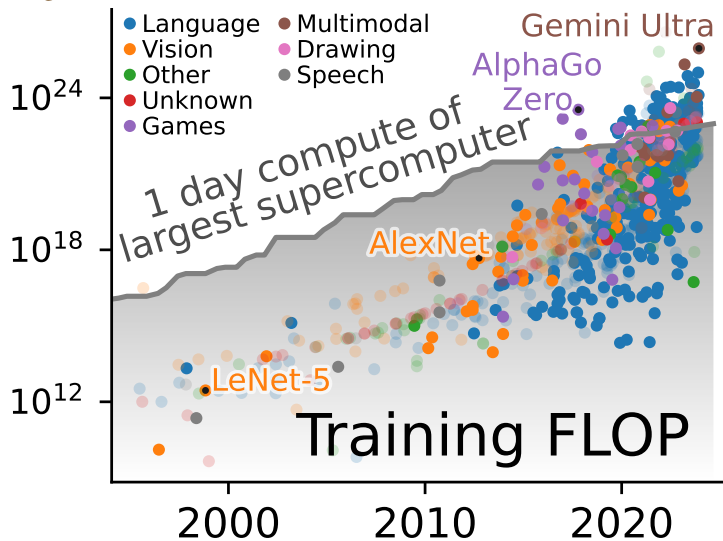
An ever-rising footprint

AI compute
overtook largest
supercomputers



An ever-rising footprint

AI compute
overtook largest
supercomputers



Local electricity supply is a bottleneck for datacenters

Concentration of power

Big gets bigger

- More data, more compute
- Automation centralizes choices

Couples to the other risks

- Mis-information
- Privacy
- Bias



Need for multi-stakeholder AI

Open source

- Openness enables innovation
- Helps mitigate bias
- Important notion: commons

Challenges to open source

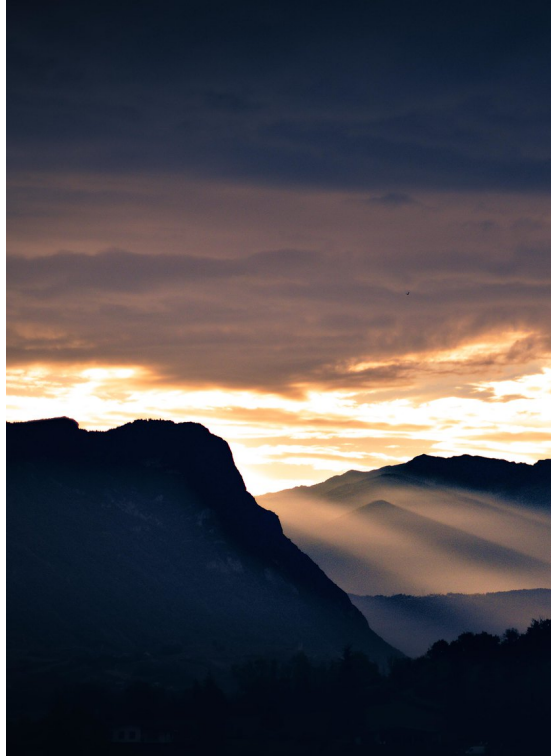
- Liabilities
 - Can kill open source
- Lack of transparency
 - What data went in the model?
- Open washing
 - Source \neq non-commercial open weights



AI is, in itself,
neither good nor bad

It depends what we,
as a society,
do of it

We need
collective appropriation
and open source



3 Reaching a balance?

Some personal thoughts





Technology



makes a better world



makes a better world?

Are we enabling good?

Technological systems
embody governance choices

- Implicit choices (eg bias)
- Explicit choices (eg buying platforms)

Selling influence to the best bidder
Mind control of the rich on the poor



Beyond individual decisions

Norms and regulations
define what's acceptable

informed by
academic and government papers

Everybody is a lobbyist
The influence game



The narrative of AI research: big AI

[Varoquaux... 2024]

Vision and language models have shown

bigger is better

*“our results can be improved simply by **waiting for faster GPUs and bigger datasets**”*

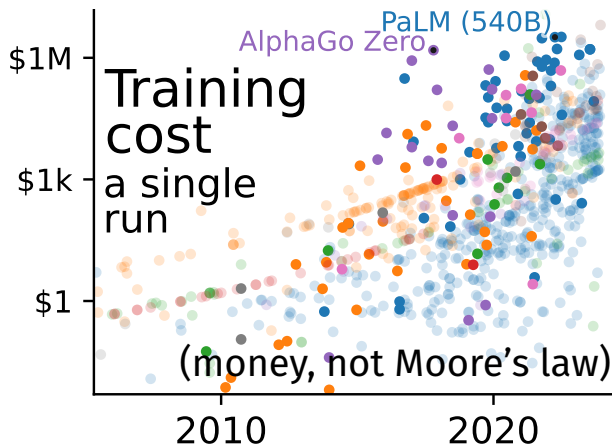
[Krizhevsky... 2012]

*“leverag[ing] **computation** [is] ultimately the **most effective** [...] The ultimate reason for this is **Moore’s law**”*

[Sutton 2019]

Downplays all the architecture & training progress

Big AI gets too big



The economics of capture
Big tech concentrates power

Open source no longer enables
setting the agenda



Realpolitik 101

Ethics is for losers

At war, the nastiest wins

Respecting ecology, privacy
these are losing games

Regulations could level the field
but fail because of competitions
across countries



Economic “efficiency” dreams

Tech says: democracy is too slow,
inefficient for the tech revolution
“too much regulations”

Elites are self-serving [\[Gilens 2012\]](#)
as we all are

Nightmare societies

As oligarchies concentrate power
they discard redundant elites

Putin modernized Russia
before his nazi killing frenzy



Democracy is
the worst form of government
– except for all the others that
have been tried.

Winston Churchill

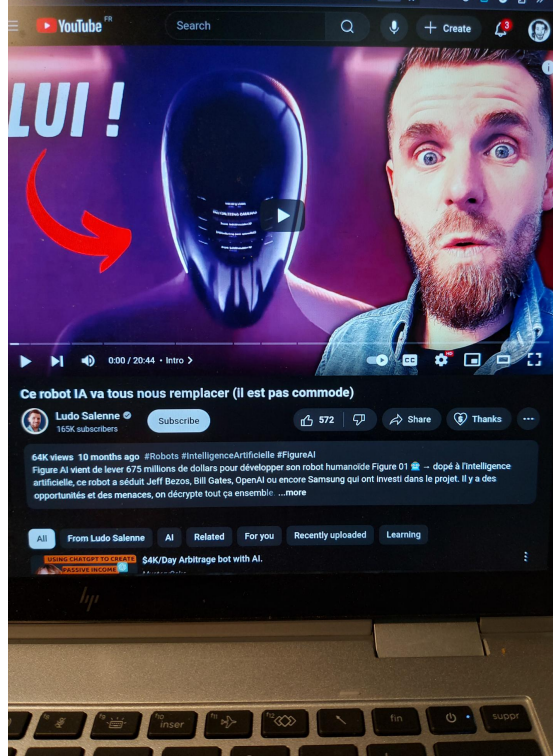


Rising anti-AI sentiment

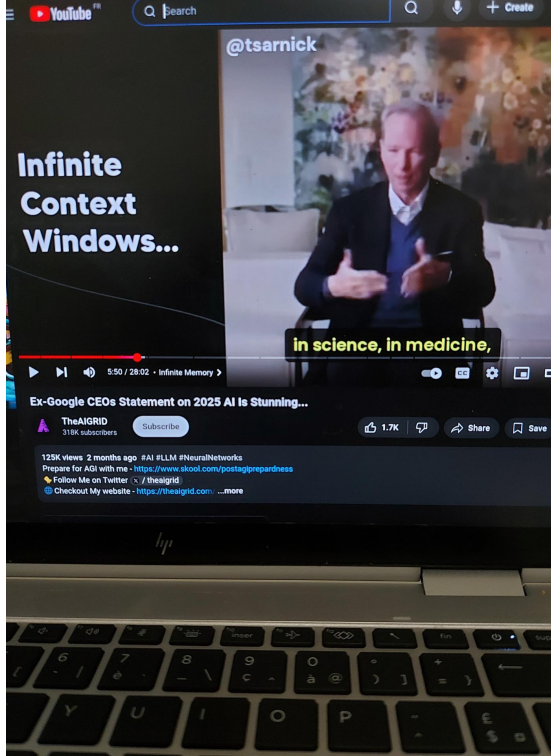
Endangered democracy

Polarization

Loss of trust in science



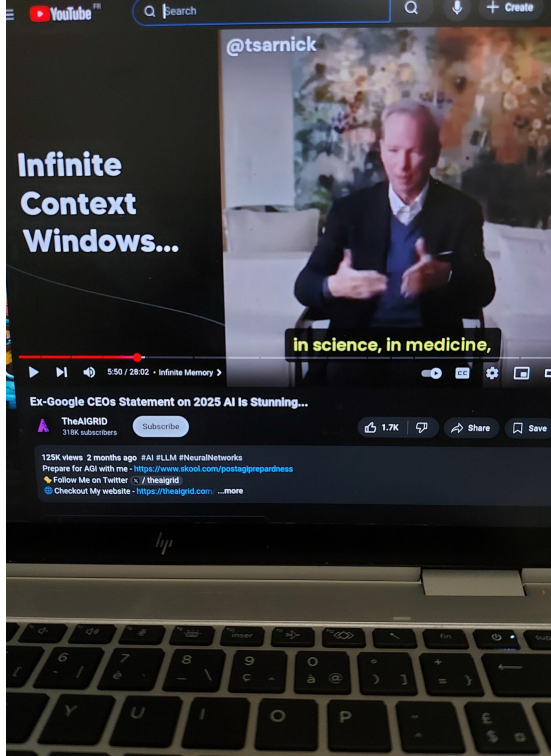
Others are more positive



Others are more positive

“It is difficult to get a man to understand something, when his salary depends on his not understanding it.”

— Upton Sinclair



Are we the
baddies?

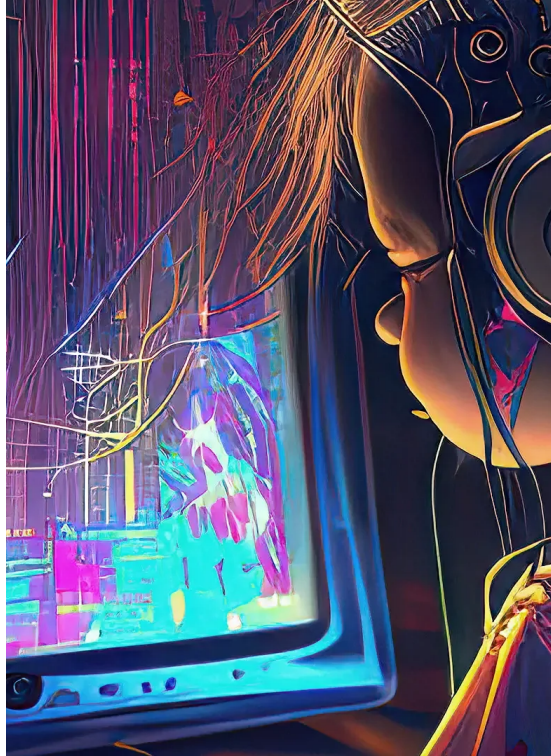


What can we, in tech, do?

Develop measures, document

- “If you can not measure it, you can not improve it.” Lord Kelvin
- It’s all about controlling the narrative, people’s worldview

Build technologies thinking about who benefits and who loses

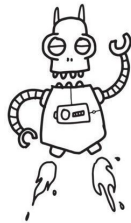


Yeah tech is cool
But have you tried people?

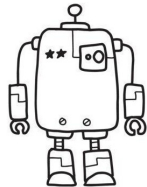
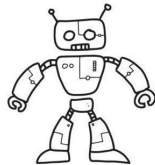
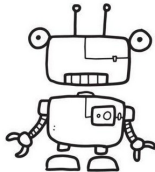
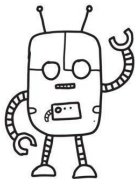
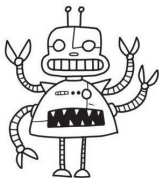
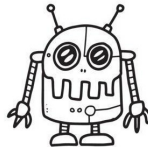
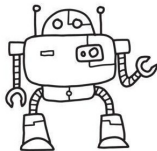
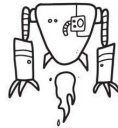
Tech, by itself, is not enough

Technosolutionism
won't save us





Between tech and public policy

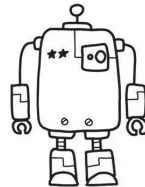
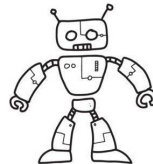
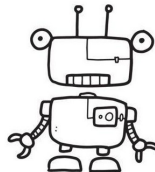
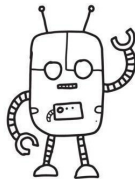
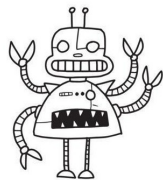
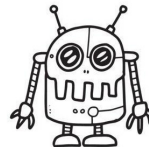
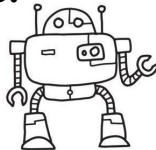
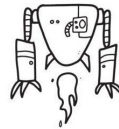
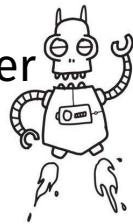


Battle of narrative techno-optimism / doomer

AI is, in itself, neither good nor bad

Need openness and collective appropriation

Also large-scale human factors:
society, power



References I

- P. Aghion, A. Bouverot, A. Amabile, C. Canivenc, G. Babinet, J. Barral, A. Bensamoun, N. Boujemaa, B. Charlès, L. Julia, ... IA: Notre ambition pour la france, 2024.
- Y. Bengio, S. Mindermann, D. Privitera, T. Besiroglu, R. Bommasani, S. Casper, Y. Choi, D. Goldfarb, H. Heidari, L. Khalatbari, ... International scientific report on the safety of advanced ai (interim report). *arXiv preprint arXiv:2412.05282*, 2024.
- M. Gilens. *Affluence and influence: Economic inequality and political power in America*. Princeton University Press, 2012.
- A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012.
- R. Sutton. The bitter lesson. *Incomplete Ideas (blog)*, 13(1), 2019.
- G. Varoquaux, A. S. Luccioni, and M. Whittaker. Hype, sustainability, and the price of the bigger-is-better paradigm in ai. *arXiv preprint arXiv:2409.14160*, 2024.