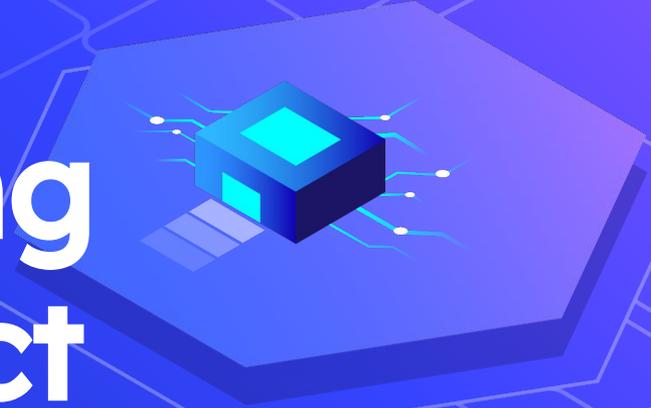# Situating AI on the road from data sharing to societal impact

backchannel 4 questions and comments

Please put questions and comments into the Etherpad at

https://w.wiki/b4q

Structure of the talk

- ⬡ Hello JLab!
- ⬡ Motivation: an explainable title
- ⬡ Perspectives on data sharing, with emphasis on research data
- ⬡ Perspectives on societal impact
- ⬡ Locating AI in the picture
- ⬡ Discussion

Hello JLab!

- I am a data scientist at the University of Virginia's (new) <u>School of Data Science</u>
- My background is in biophysics
- My focus is on
  - open science and collaboration
  - knowledge representation and data quality
- You can find me at <u>@EvoMRI</u>

Motivation: an explainable title

- ⬡ Ingredients
  - · The School's mission
    - · shortened
  - · A quote on AI
  - · Note of caution
  - · Reworking the quote
- ⬡ Title candidates

The School's mission

To be a national and international leader in responsible data science, emphasizing interdisciplinary collaboration which results in furthering discovery, sharing knowledge, and societal benefit

> **The School's mission (shortened)**
>
> To be a leader in responsible data science, emphasizing data sharing and societal impact

" A quote on AI

big data

+ deep learning

+ faster hardware

has been a winning formula

Gary Marcus & Ernest Davis, Rebooting AI

A note of caution

deep learning is

- Greedy [needs lots of data]
- Opaque [hard to explain]
- Brittle [easy to fool]

Gary Marcus & Ernest Davis, Rebooting AI

" Reworking the quote

data sharing

+ ?AI_techniques

+ faster hardware

+ some missing piece?

could lead to societal impact

after Gary Marcus & Ernest Davis, Rebooting AI

Candidate titles

- The societal impact of data sharing seen through the lens of AI
- The societal impact of data sharing in light of AI
- Societal impact of data sharing in light of AI
- Societal impact of data sharing, and what AI has got to do with it
- From data sharing to societal impact via AI
- AI on the road from data sharing to societal impact
- Situating AI on the road from data sharing to societal impact
- The societal impact of data sharing, considered through the lens of AI
- The societal impact of data sharing and what it means for AI
- Artificial intelligence in the context of data sharing, open science and disasters
- Playing the Wikipedia game from AI to Open science and back
- Putting AI on the red team
- Research with AI on the red team
- Research as an adversarial network playing against nature
- What if AI were (also) on science's red team?

Data sharing

- In Biophysics
- As part of open research/ open science
- In disaster contexts
- At scale
- From research infrastructures to ethics
- From cooperation to competition
- Across data life cycle & research cycle

Biophysics
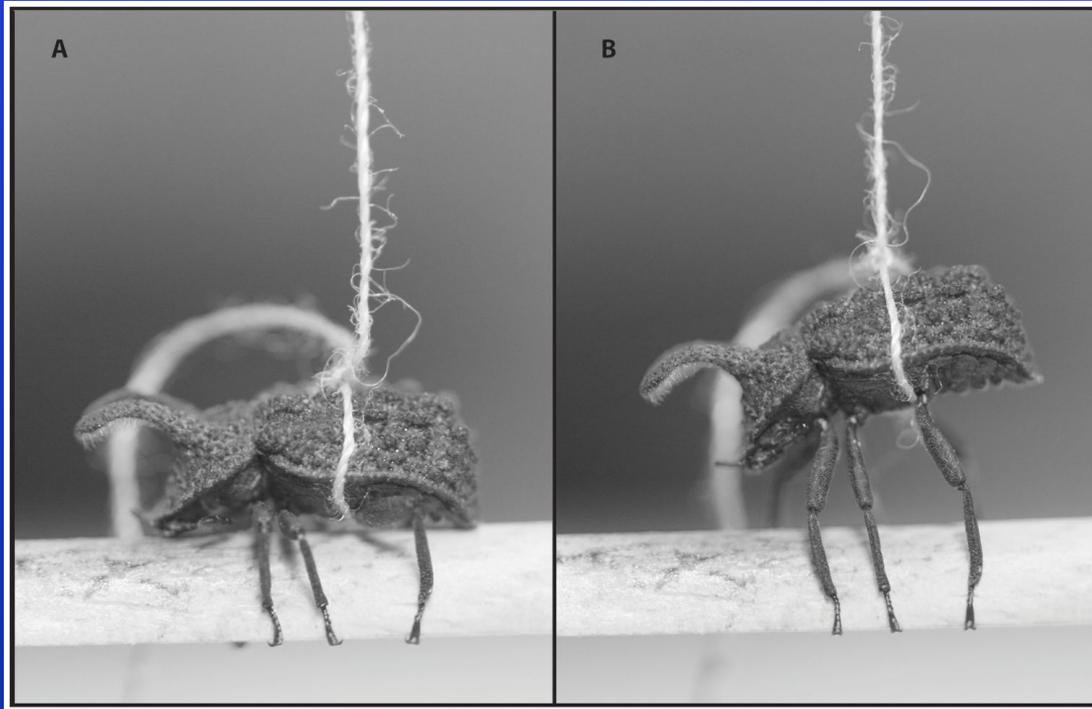
Grip strength trial with a _Bolitotherus cornutus_ beetle

Source:
Benowitz et al. (2012), CC BY 2.5

Data sharing in biophysics

- Across spatial scales of biological organization
  - From the molecular level to the biosphere
- Across temporal scales of biological organization
  - From femtoseconds to millions of years
- Across biological species and engineered variants
- Across methodological approaches in field, lab or elsewhere
- Between experimentalists, theorists, empiricists and others
- With biology, physics and beyond
- With academia, industry, government and others, globally

A biophysical note on the "faster hardware"

Faster hardware does not necessarily imply Moore's law - it could also involve unconventional computing.

For instance, some reaction-diffusion systems or organisms like the slime mold *Physarum polycephalum* can be used for specific kinds of computations, e.g. network optimization tasks.

Background: Advances in Physarum Machines

Image source: frankenstoen, CC BY 2.5

The open landscape

Data sharing as part of open science

- Open science is about opening the research process, to facilitate
  - Collaboration - including with AI
  - Scrutiny - including from AI
  - Uptake, reuse and refinement - including through AI
- Papers
  - Publishing the research process
  - The Transformative Nature of Transparency in Research Funding
- Video

Disasters

"Open data matters most when the stakes are high

[...]

One of the great lessons we learned through the experience was the power of data to illuminate our path to recovery."

—Denice Ross (2015), Presidential Innovation Fellow "*Ten Years After Katrina: New Orleans' Recovery, and What Data Had to Do with it*"

New Orleans in 2005, flooded by Hurricane Katrina (by Jocelyn Augustino, public domain)

Data sharing in disaster contexts (i)

- Humanity has a long history of sharing in disaster contexts
- Data sharing is a relatively new element of disaster response
  - E.g. many COVID-19-related dashboards and databases
- Disaster-related data typically has low veracity (especially at the beginning), but high velocity/ volume/ variety
  - AI could be deployed immediately for known disaster types
- Affected populations and environments are especially vulnerable
- Decisions in disaster contexts - right or wrong, timely or not - can have profound societal impact in either direction
  - How can AI assist with that?

Data sharing in disaster contexts (ii)

- Project:
    - Quantifying the Impact of Data Sharing on Outbreak Dynamics
- Preprint:
    - Using logical constraints to validate information in collaborative knowledge graphs: a study of COVID-19 on Wikidata
- Papers:
    - Open drug discovery for the Zika virus
    - Strategies and guidelines for scholarly publishing of biodiversity data

Data sharing at scale

- Scale of the data
  - Volume - AI usually needs lots of data
  - Velocity - AI can handle a range of velocities
  - Variety - AI can benefit and contribute, though not mature yet
  - Veracity - AI often not robust enough
- Scale of data users and usage
  - Citizen science - AI can assist
  - Open APIs - AI can use and contribute
  - FAIR data - AI can use and contribute

Data sharing - other perspectives

- Through research infrastructures - AI can use and enhance them
- From an ethical perspective - a minefield for AI
- In the context of cooperation - AI can assist and contribute
- In the context of competition - AI can monitor and compete
- Across the data life cycle - AI is strong in data analysis and idea generation, weak in data documentation and preservation
- Across the research cycle - "robot scientists" can cover it all for well-defined areas with machine-friendly infrastructure

Open sharing maximizes societal impact

- Enhances research & education
- Engages diverse communities (which could include AI) & leverages their distributed expertise
- Stimulates innovation & quality control
- Drives adoption & collaboration & reuse
- Saves resources & speeds things up

Societal impact

- Addressing the climate emergency
- Addressing the COVID-19 pandemic
- Addressing (other) Sustainable Development Goals
  - …
- Addressing the needs of specific communities
  - …
- Addressing the needs of specific organizations
  - …
- Addressing the needs of individuals
  - …
- How can AI assist or benefit?
- What are the side effects of AI?

Addressing the climate emergency (i)

We must abandon the conceit that individual, isolated, private actions are the answer. They can and do help. But they will not take us far enough without <u>collective action</u>.

<u>Al Gore, Nobel Peace Prize acceptance speech (2007)</u>

Addressing the climate emergency (ii)

We must understand the connections between the climate crisis and the afflictions of poverty, hunger, HIV-Aids and other pandemics. As these problems are linked, so too must be their solutions. We must begin by making the common rescue of the global environment the central organizing principle of the world community.

Al Gore, Nobel Peace Prize acceptance speech (2007)

Addressing the climate emergency (iii)

- Here are 10 ways AI could help fight climate change
- Climate change and machine learning
- How AI Is Helping Solve Climate Change
- Fighting climate change with AI
- Machine learning and artificial intelligence to aid climate change research and preparedness
- The Amazing Ways We Can Use AI To Tackle Climate Change
- Artificial Intelligence—A Game Changer for Climate Change and the Environment
- AI can fight climate change but there's a catch: Optimization doesn't automatically equal emissions reduction
- AI and Climate Change: How they're connected, and what we can do about it
- Tackling Climate Change with Machine Learning
- How artificial intelligence can tackle climate change
- Climate Change AI

" **Reworking the quote**

data sharing

+ ?AI_techniques

+ faster hardware

+ some missing piece?

could lead to societal impact

after Gary Marcus & Ernest Davis, Rebooting AI

What about that missing ingredient?

- Collective action problems
  - Examples:
    - Climate action, Open science, Data governance, Ethical AI
  - Solutions require aligning individual and collective interests
- Collective intelligence?
  - The Future of Minds and Machines: How artificial intelligence can enhance collective intelligence
  - The research frontier: where next for AI and collective intelligence?

Credits

Special thanks to all the people who made and released these awesome resources for free:

◇　Presentation template by SlidesCarnival, under CC BY 4.0

Thanks!

Any questions?

You can find me at:

@EvoMRI

dm7gn@virginia.edu

backchannel 4 questions and comments

Please put questions and comments into the Etherpad at

https://w.wiki/b4q