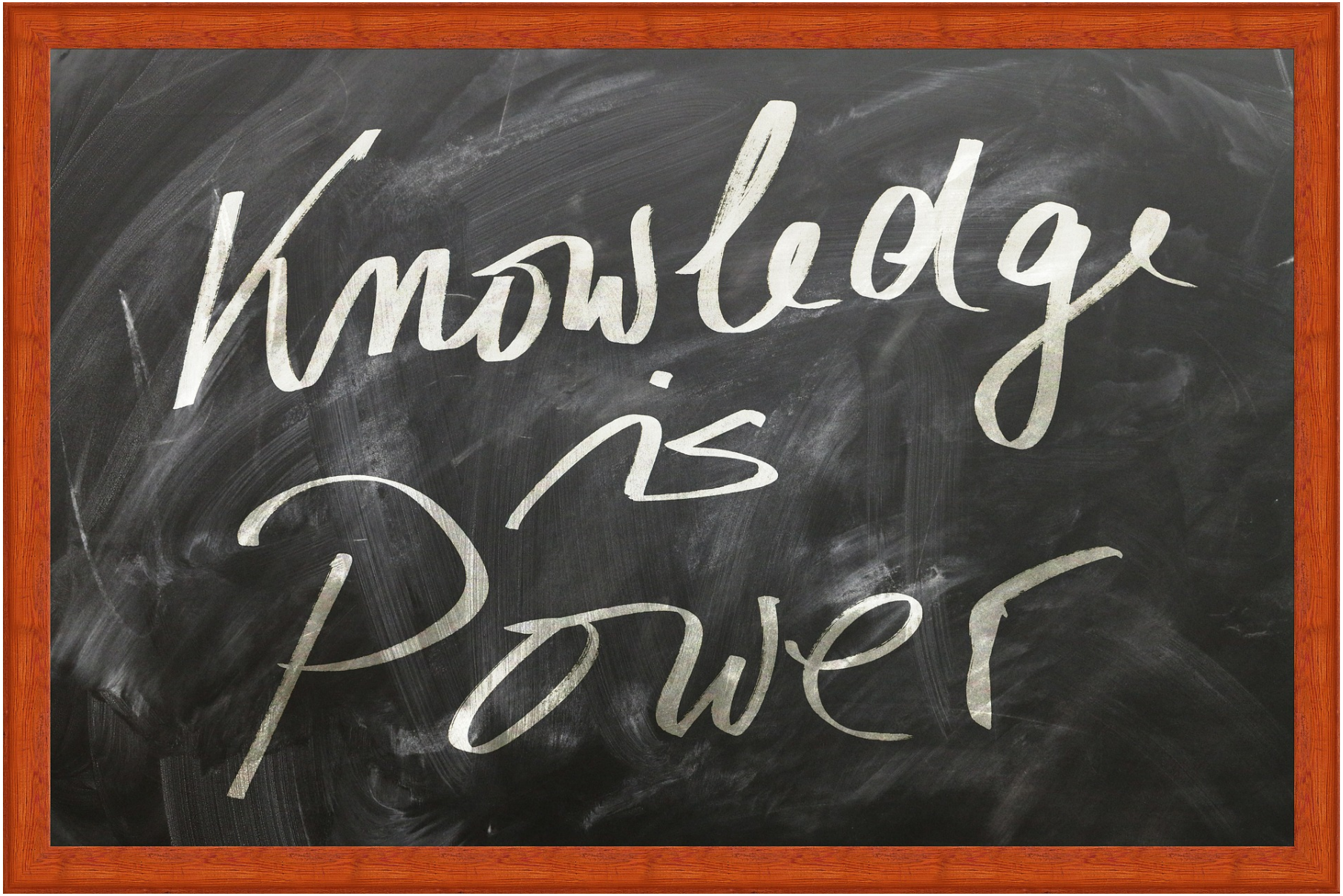
The Use of AI and Epistemic Values in Science

Dr. Vlasta Sikimić

Eindhoven University of Technology

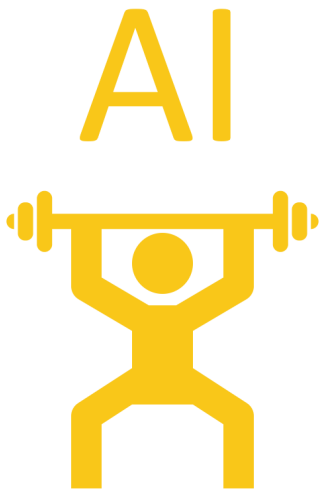
February 21, 2024


Knowledge
is
Power

What can AI do for science?

- AI can assist in making hypothesis.
- A human needs to be in a loop to guarantee quality.
- AI might cause an increase in paper submissions.
- This will put pressure on editors.
- It could increase the difference between researchers who have access to data and who do not.

What else can AI do for science?



- AI can also be used for evaluation of science:
  - self-evaluation (large language models imitating reviewers),
  - evaluating science.

Replication crisis

- Nowadays, we have a replication crisis (often associated with the work by John P. A. Ioannidis).
- One of its roots is the pressing need to publish, both in natural and social sciences.

What could be the relationship between the use of AI and replication crisis?

AI and epistemic values in science

- AI can threaten them,
- but it can also give a platform,
- e.g., linguistic justice in science.

Vučković & Sikimić (2022). How to Fight Linguistic Injustice in Science: Equity Measures and Mitigating Agents, Social Epistemology.

Trustworthy AI

- Fairness of the results,
- Transparency of the procedures,
- Traceability of the people responsible for the results,
- Data protection,
- Motivation and a design that is benefiting humans.

European Commission, Directorate-General for Communications Networks, Content and Technology. 2019. Ethics guidelines for trustworthy AI.

Values corresponding to principles

- Fairness $\rightsquigarrow$ justice $\rightsquigarrow$ epistemic justice.
- Transparency $\rightsquigarrow$ honesty $\rightsquigarrow$ epistemic honesty.
- Accountability $\rightsquigarrow$ responsibility $\rightsquigarrow$ epistemic responsibility and group ep. responsibility.
- Privacy $\rightsquigarrow$ care $\rightsquigarrow$ moderate curiosity.

Hagendorff, T. (2022). A virtue-based framework to support putting AI ethics into practice. Philosophy & Technology, 35(3), 55.

Values in science

- Epistemic values – help us in learning.
- Non-epistemic are e.g., socio-political and economic values.
  - What are the examples of economic values that play a role in science and technology development?

Kuhn's distinctions

1. epistemic/non-epistemic values in science;
2. paradigmatic/preparadigmatic sciences;
3. standard textbooks/diversity.

The role of values in science

- Deciding which topics to research on. Which values can you imagine to be used here?
- The direct use of the new technology yields many value related considerations, ethical, economic, etc.
- Values play a role in choosing a methodology, e.g., ethical drug or AI related testings.

The value free ideal

- How do you understand scientific objectivity?
- Weber's value free ideal of science is connected to the socio-political circumstances and to objectivity.
- Important for science funding and for academic freedom.

How do you understand academic freedom?
How could AI impact it?

Social Epistemology

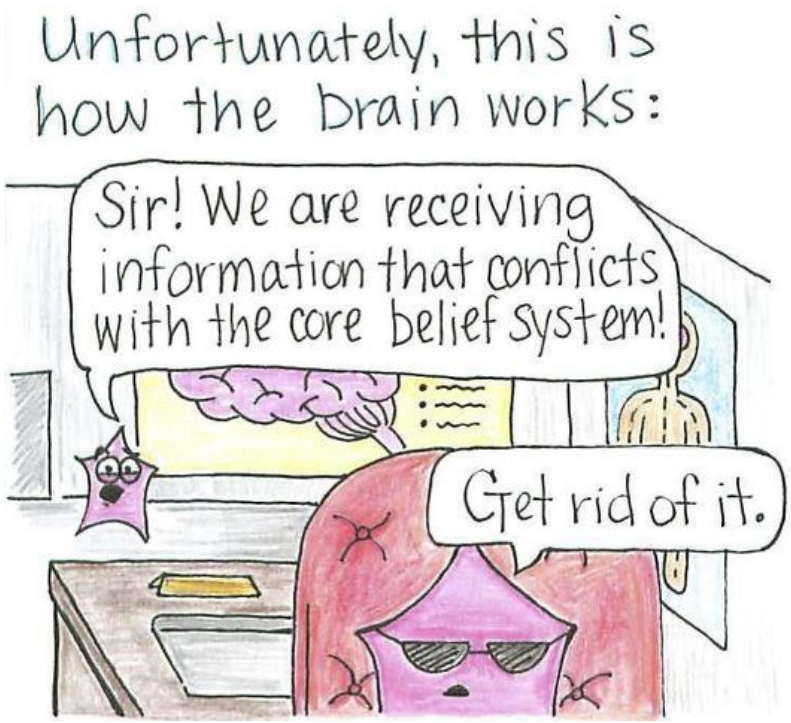
- **Traditional epistemology:** focus on the analysis of the notion of individual knowledge.

- **Social epistemology:** focus on the group knowledge and gaining knowledge (learning) within a group.
  - **Goal:** maximization of group knowledge.

Group knowledge

Motivation

- Maximizing scientific knowledge as the whole.
- Optimal distribution of resources.
- Examining the role of cognitive diversity.



Group behavior: epistemic tolerance

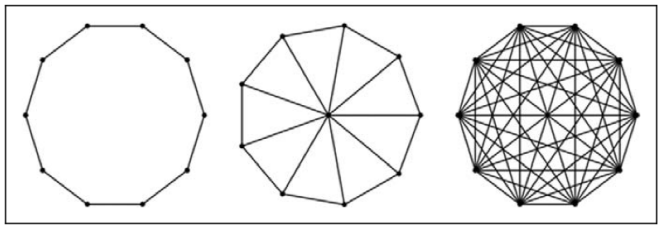
Epistemic authoritarianism

Trust me I know it

- Epistemic authoritarianism;
- Questioning of the scientific method.

Cognitive diversity: Zollman's optimization

Epistemic network – a ten-person cycle, wheel and complete graph:



Results: less connectivity among the researchers contributes to the truthfulness of their findings; but (in some cases) slows down the learning process.

- ✓ It is a hypothesis-driven simulation;
- ✓ It is a good general argument;
- ✓ A quantitative result.

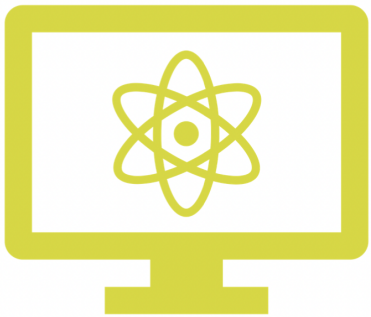
Isaac Newton struggles to write the economic impact section of his 'gravity' proposal.

cartoon by Eoin O'Sullivan

Methods of grant review

- Grant peer review is time-consuming, costly, and sometimes criticized as unreliable.
  - The correlation between the results of peer review grant assessments and the citations is weak (e.g., Doyle et al. 2015, Fang et al. 2016).
- A lottery approach reduces biases and the workload of the reviewers, but it introduces randomness in the decisions about science funding.

Machine learning in grant review



- Machine learning has already been used for prescreening of job applications in industry, finding reviewers, etc.
- It is a fast and cheap method.
- But how reliable is it for predicting scientific success?
  - This strongly depends on the data at hand.
- Yet, what do we define as scientific success?
  - Inclusion is one candidate.

Algorithmically predicting project efficiency in physics

Fermilab

- The results of the predictive analyses show moderately high accuracy.

Sikimić & Radovanović (2022). Machine learning in scientific grant review: algorithmically predicting project efficiency in high energy physics. European Journal for Philosophy of Science.

Our research: a data-driven approach

1. Theoretic analysis (application of notions from social epistemology to high energy physics);
2. Empirical analysis of the data from Fermilab:
   - qualitative analysis;
   - quantitative analysis – Data Envelopment Analysis **DEA**;

   Perović, Radovanović, Sikimić & Berber. Optimal Research Team Composition: Data Envelopment Analysis of Fermilab Experiments, *Scientometrics*, 2016.
3. predictive analysis (predicting optimal distribution of resources).

Sikimić, Radovanović, Perović. Predictive Analysis of Experiment Efficiency in High Energy Physics: A Machine Learning Approach, in preparation.

Algorithmic grant review: can we do it?

Epistemic question

- We should be careful when applying data mining techniques to grant review.
- Only similar experiments are comparable.
- Field-wise organization of data repositories.
- We need to keep in mind that scientific progress is difficult to predict.

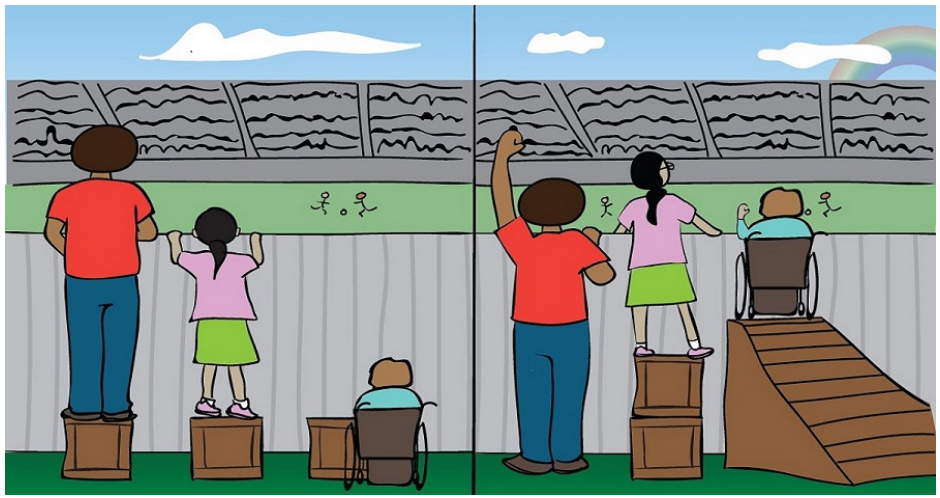
Algorithmic grant review: should we do it?

Ethical concerns

- Adversarial behavior;
- Algorithmic fairness;
- Data privacy and profiling of researchers.

Affirmative action

- Affirmative action is intended to improve the situation of a group that has previously suffered injustice.
- In this way we interfere hoping that in the future injustice will be overcome.
- Still members of the privileged group might feel discriminated because of affirmative actions.



source: equitytool.org

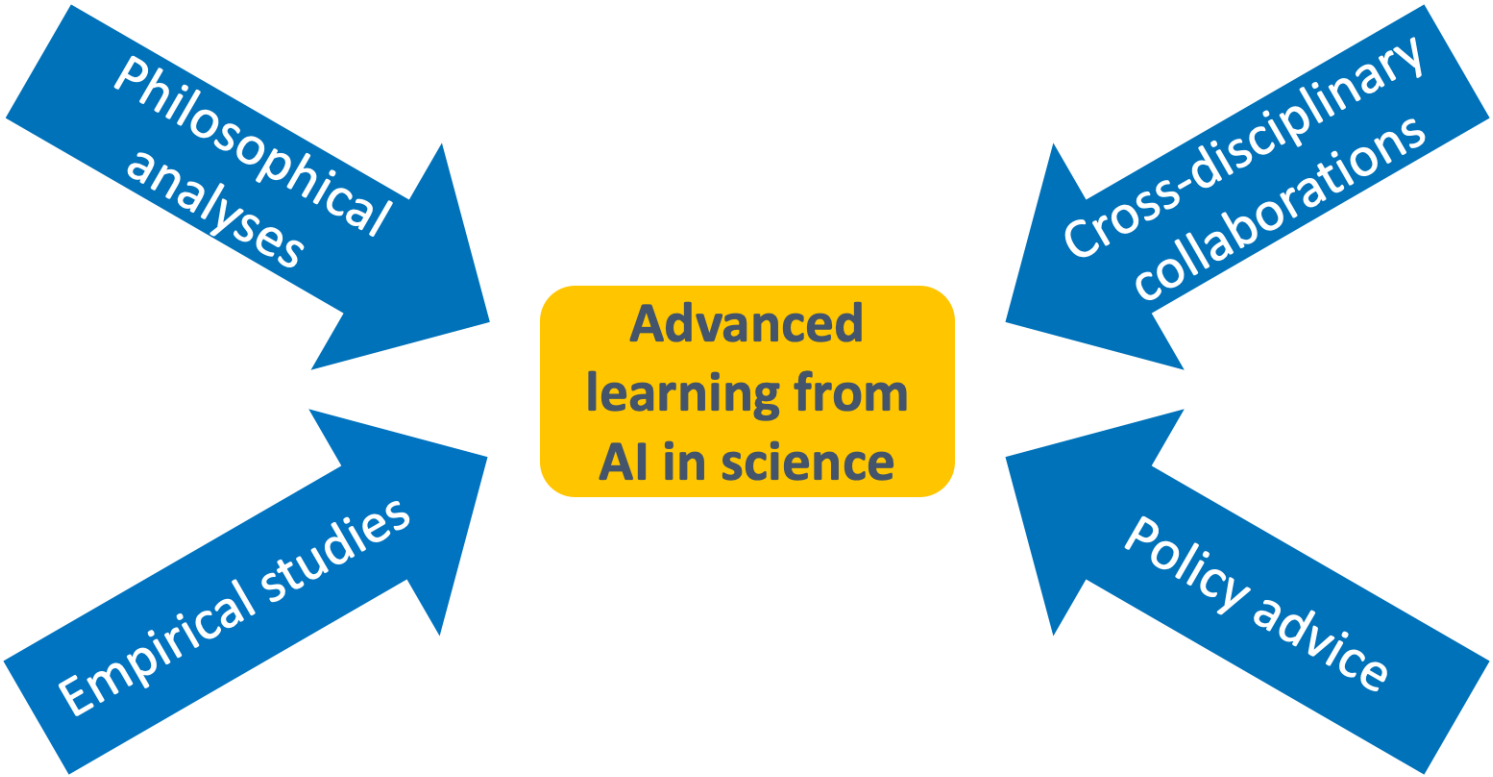
equality vs. equity

Values

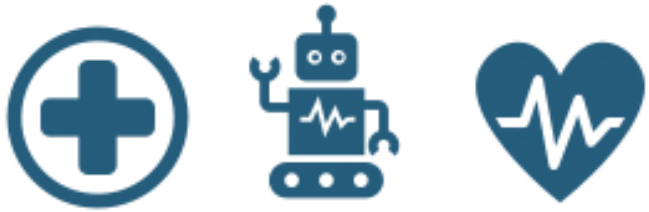
- Values, such as epistemic inclusion and equity, can be differently implemented:
  - Within the algorithm;
  - Data curation level;
  - Human in the loop taking care of it.

Sikimić, V. (2023). Epistemic Inclusion as the Key to Benefiting from Cognitive Diversity in Science. Social Epistemology, 37(6), 753-765.

Conclusions



Medical assessment and AI



- AI based medical risk prediction tools are currently applied to approximately 200 million Americans.
- They should identify the people that require additional medical care because of their high health risk.
- Algorithm allocated less medical resources to people of color because there was less money spent on them in the past.

Obermeyer, Z. et al. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366, 447–453.

Biases

Algorithms mirror human biases.

Cognitive biases:

- A form of heuristic reasoning that is not perfectly fitting for the relevant case and has a suboptimal result.
- Examples: sunk-cost bias, simplicity bias, etc.

Biases with a social component:

- Examples: authority bias, stereotyping, groupthink, etc.

Debiasing strategies

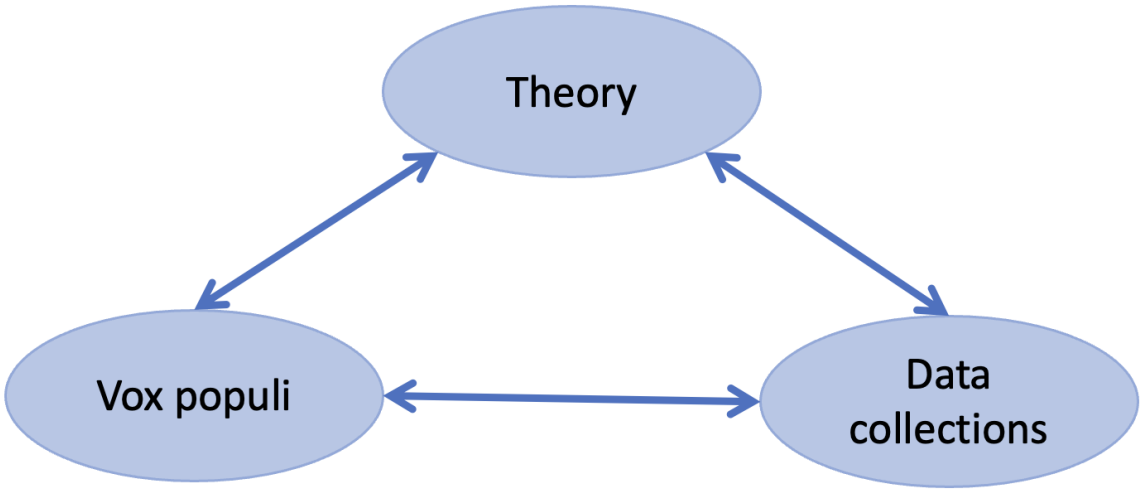
- Modifying the algorithms.
- Improving the training data.
- Using relevant parameters, not proxies:
  - e.g., making medical decisions based on health benefits instead on money spent in the past.
- Increasing fairness in the society.

No raw data

- In her work, Sabina Leonelli points out that data are always curated and interpreted.
- Moreover, according to Leonelli, data 'travel' and get re-interpreted.
- The need for data democratization in the context of inclusive science.

Leonelli (2014). What difference does quantity make? On the epistemology of Big Data in biology. Big data & society.

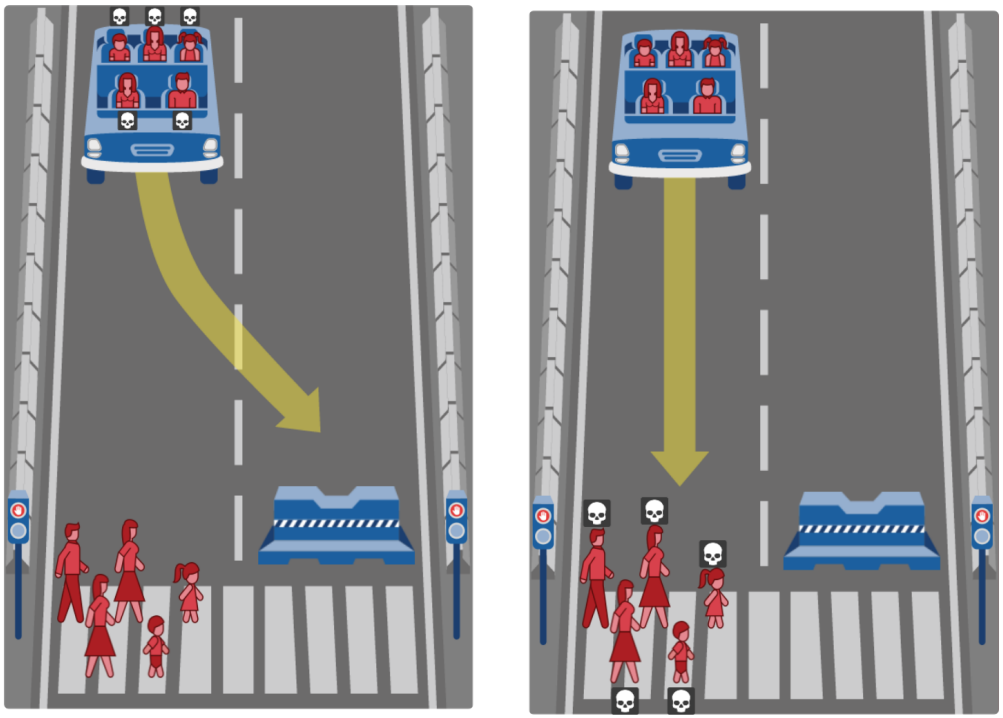
From data to theory and back



- Even synthetic data need to be representative of the real world.
- We have the constant back and forth (circular) interaction between theory and empiricism.
- For the wisdom of the crowds effect we need high quality information.

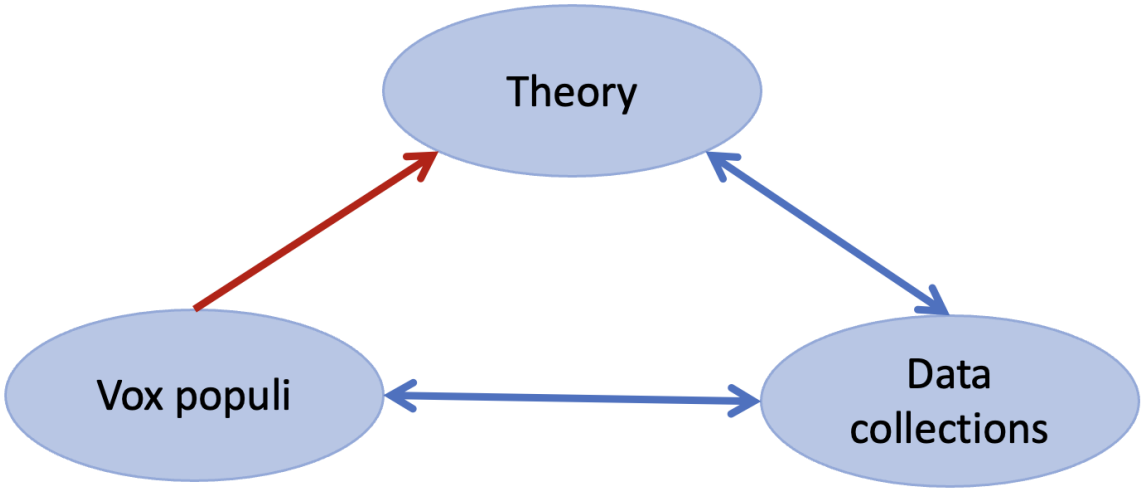
Irreducibility of the normative aspect to vox populi

- Does asking people about their moral intuitions create a moral theory?
- The Moral Machine experiment.



Awad, E., Dsouza, S., Kim, R. et al. The Moral Machine experiment. Nature 563, 59–64 (2018).

Normative theory and the public opinion



- If the flow would go only in one direction we would be in a reenforcing bubble.
- We want interaction and correction – dynamic updating of all three components.