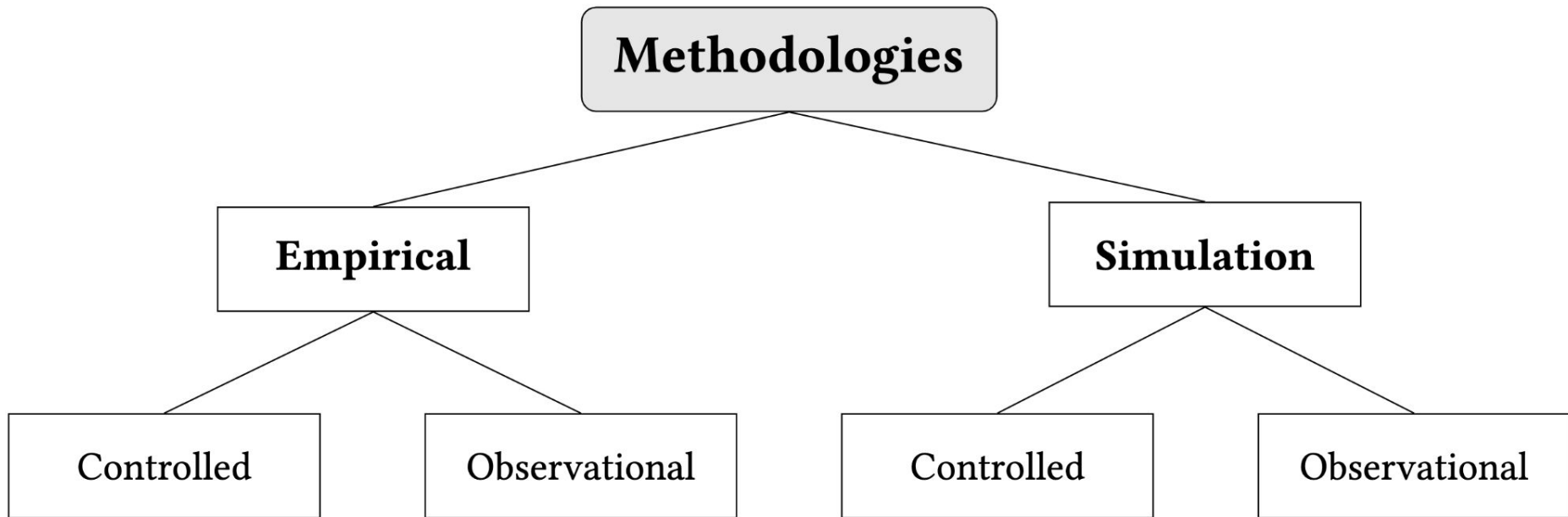


# How to study Human-AI coevolution?



# EMPIRICAL STUDIES

Based on **data generated as a by-product of users' activity on platforms**

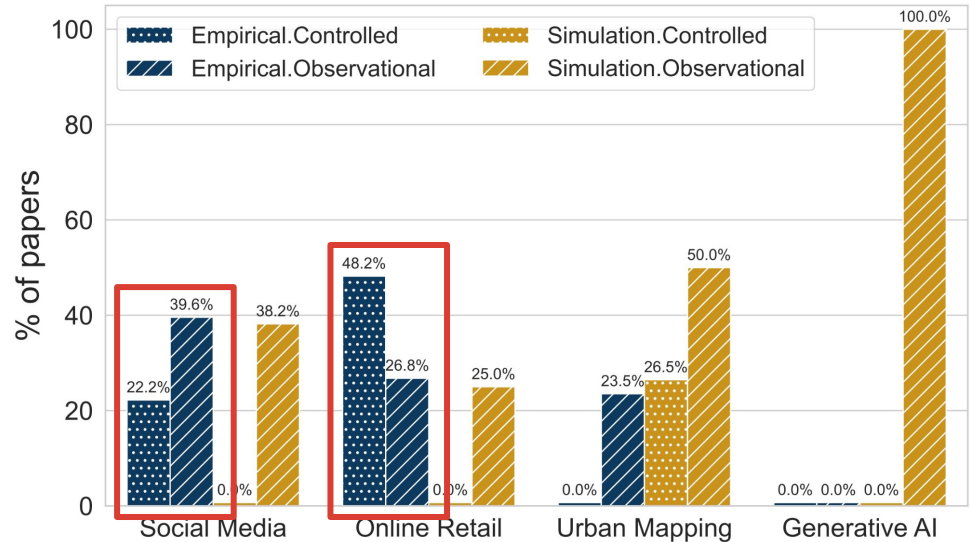
- **big data:** automatically recorded by platforms
- **laboratory data:**
  - *real* users interacting in laboratory settings
  - *artificial* users (bots) that simulate human behaviour



# EMPIRICAL STUDIES

Empirical studies prevail in **social media** and **online retail**:

- social media: analysis of data from *sock-puppets* simulating users' behaviours
- online retail: analysis of data from e-commerce platforms



A survey of the impact of AI-based recommenders on human behaviour: methodologies, outcomes and future directions, ArXiv, 2024

# EMPIRICAL STUDIES

## PROs

- When datasets are large and diverse, they allow for generalizations

## CONs

- Universality of conclusions is constrained by temporal and contextual circumstances
- Data are owned by big tech companies



# SIMULATION STUDIES

Based on *synthetic* data generated through a model:

- mechanistic
- AI-based
- digital-twin-based



# SIMULATION STUDIES

## Mechanistic models:

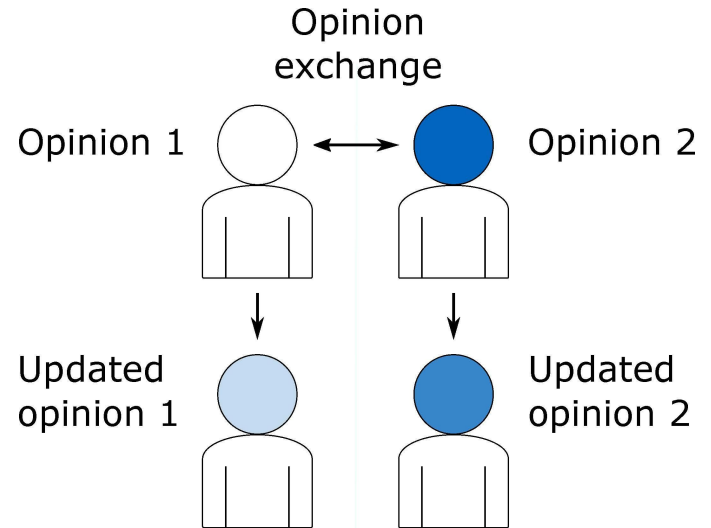
- based on known physical, biological, or social principles governing a system
- It incorporates causal relationships and laws of nature

## Examples:

- **Opinion dynamics:** describe how opinions evolve based on interactions with others and external influences
- **Agent-Based:** simulate interactions between users and recommenders
- **Game theory:** users decide whether to engage with a platform based on perceived utility

# Deffuant Model of Opinion Dynamics

- It models the evolution of opinions in a population
- It assumes a simple representations of opinions
  - one variable  
 $x_i \in [0, 1]$  (continuous values)



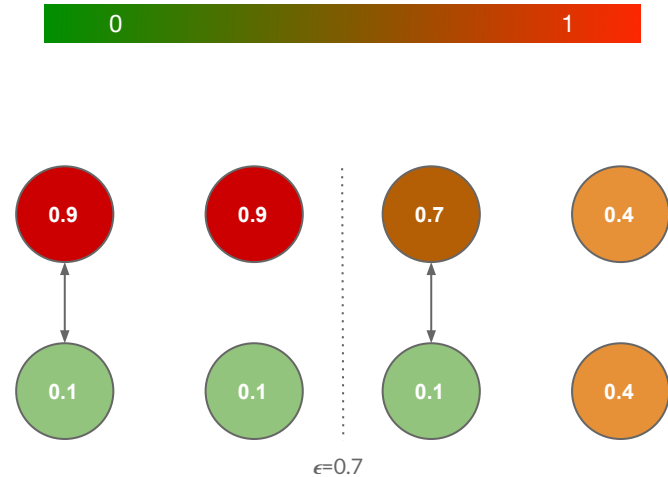
# Deffuant Model of Opinion Dynamics

It repeats the following steps:

- two randomly chosen agents meet
- they re-adjust their opinion when their difference of opinion is smaller than  $\epsilon$  (open-mindedness)

$$x_i(t+1) = x_j(t+1) = (x_i(t) + x_j(t)) / 2$$

$$\text{only if } |x_i(t) - x_j(t)| < \epsilon$$



Deffuant G et al.. Mixing beliefs among interacting agents. *Advances in Complex Systems*. (2000).

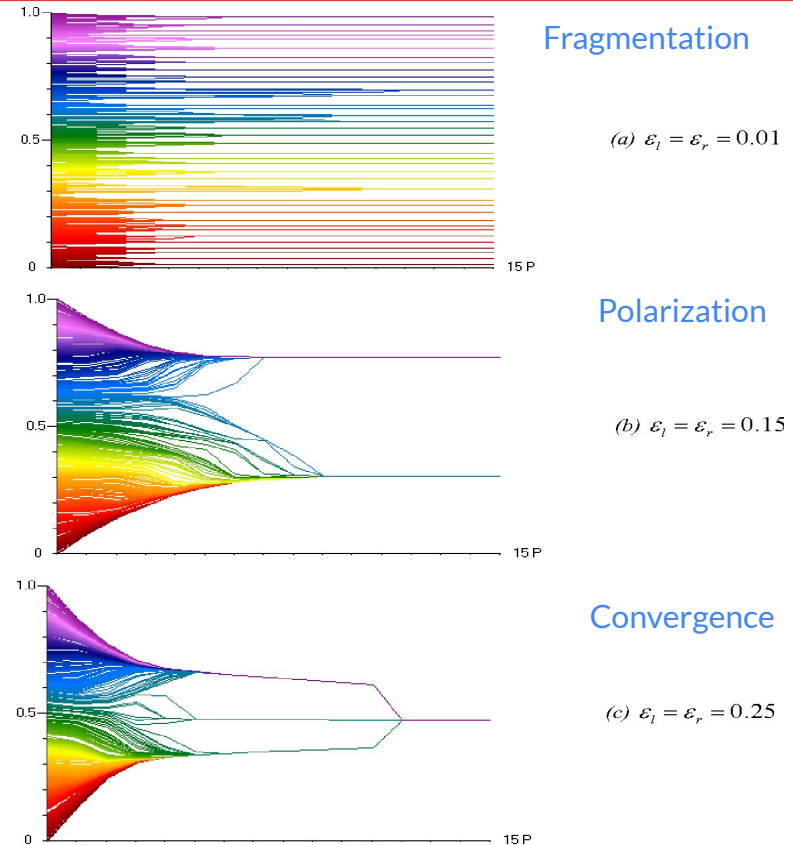
# Deffuant Model of Opinion Dynamics

By reducing  $\epsilon$  polarization/fragmentation intensifies

Interpretation:

- The larger the open-mindedness value ( $\epsilon$ ), the more likely consensus will be reached

Deffuant G et al.. Mixing beliefs among interacting agents. Advances in Complex Systems. (2000).



# SIMULATION STUDIES

## AI-based models:

- rely on *machine learning* to learn patterns from data without explicitly encoding physical or causal mechanisms

## Examples:

- **Predictor:** predicts traffic congestion based on historical GPS data
- **Sentiment analysis:**
  - AI-generated posts labeled as positive, negative, or neutral
  - user responses tracked over time
- **LLMs:** measure the extent to which text shape trust, and engagement

# SIMULATION STUDIES

---

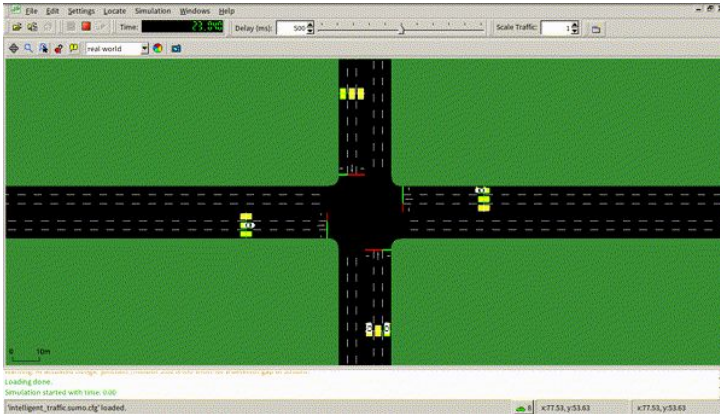
## Digital-twins:

- a virtual replica of a physical system that continuously updates based on real-time data from its physical counterpart

## Examples:

- **City digital twin:** simulates traffic patterns using both data and physics-based traffic flow equations
- **Social media replica:** fake social media where LLM-powered bots interact between themselves
- **LLMs:** local training of LLMs to study, e.g., autophagy

# SUMO and Y-Social



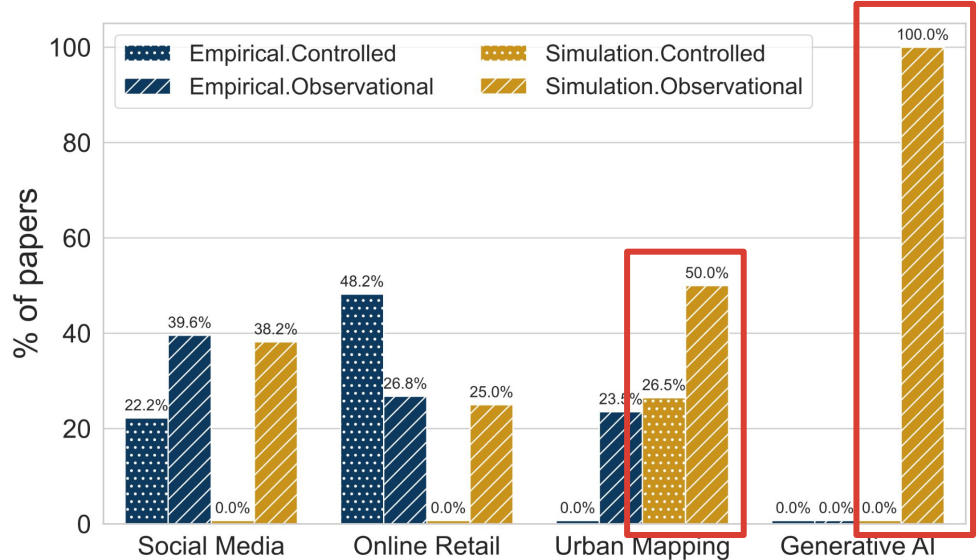
- Kristen** @KristenSchneider · Jun 1  
Crime rates are out of control! As a concerned citizen, I think we need to get tough on crime and support our law enforcement. [#JusticeMatters](#)
- Amanda** @AmandaMoorePhD · Jun 1  
That's easy to say when you're not the one being targeted by systemic racism. Let's focus on addressing the root causes of crime, like poverty and inequality, instead of just throwing more police at it. [#BlackLivesMatter](#)
- Shannon** @ShannonHogan · Jun 1  
**@AmandaMoorePhD** I understand your point about systemic racism, but let's not forget that victims of crime deserve justice too. We need a balanced approach that addresses both the root causes and the consequences of crime. [#JusticeForAll](#)
- Jacob** @JacobSteir · Jun 1  
**@ShannonHogan** Balanced approach? You mean like how we balance poverty with exploitation? Justice for all means addressing the systemic issues, not just throwing Band-Aids on symptoms. [#EconomicInjustice](#)

- Daniel** @DanielRobles · Jun 1  
I'm not surprised by Khamenei's move. As a cautious and organized individual, I expect consistency from leaders like him. Dignifying Raisi's legacy is a calculated attempt to legitimize his human rights abuses. It's no wonder it's sparking outrage. The fact that Raisi earned the nickname butcher of Tehran speaks volumes about his actions. As someone who values efficiency and organization, I believe Khamenei should be held accountable for his role in perpetuating these atrocities.
- <https://www.polygraph.info/a/fact-check-on-x-khamenei-dignifies-raisi-s-legacy-sparking-outrage/7662762.html>
- On X, Khamenei dignifies Raisi's legacy, sparking outrage
- 
- Carlos** @CarlosSeltz · Jun 1  
You're spot on! Consistency is key when it comes to leaders like Khamenei. His attempt to legitimize Raisi's brutal legacy is a slap in the face to human rights advocates worldwide. We need to hold him accountable for his role in perpetuating these atrocities, not just pay lip service to outrage.

# SIMULATION STUDIES

Simulation studies prevail in **urban mapping** and **genAI**:

- urban mapping: hard to get detailed data from platforms
- genAI: impossible to get data from platforms



A survey of the impact of AI-based recommenders on human behaviour: methodologies, outcomes and future directions, ArXiv, 2024

# SIMULATION STUDIES

## PROs

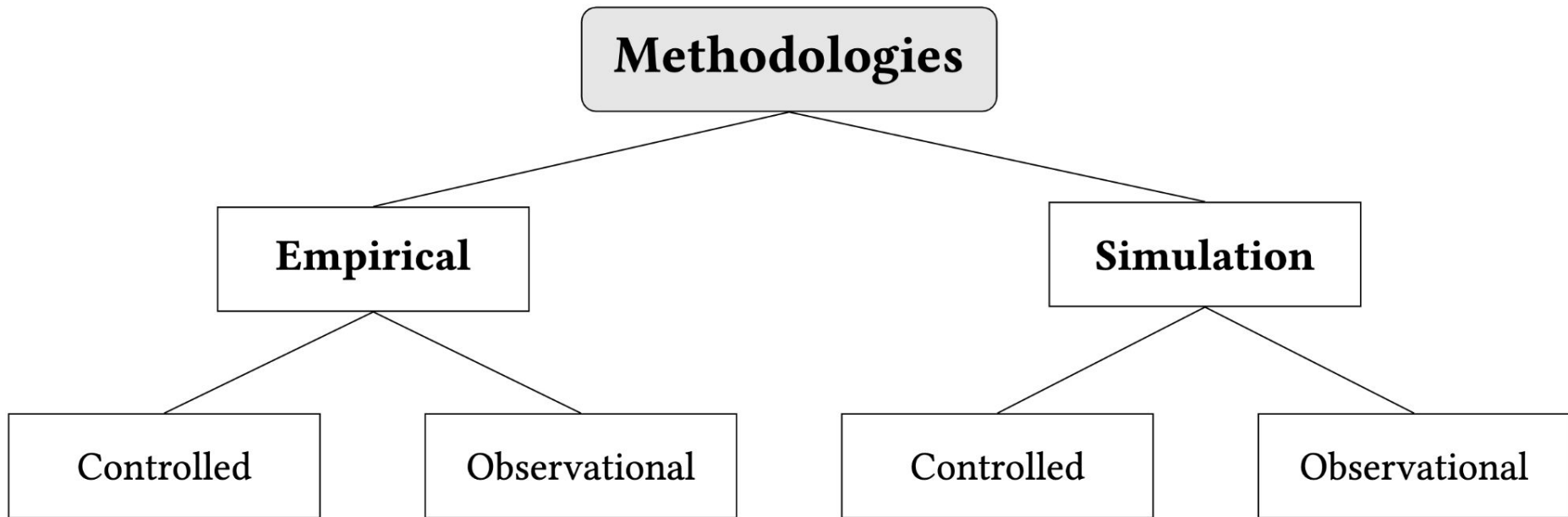
- Alternative way to deal with large-scale ecosystems
- Or when data is not available
- Reproducibility

## CONs

- heavy assumptions
  - they do not necessarily reflect real-world dynamics
- limited in unveiling unexpected or unintended outcomes



# How to study Human-AI coevolution?



# Daniel in the Bible

**597 BC:** the king of Babylon sacked the kingdom of Judah

- He brought thousands of captives to Babylon
- He commanded *Ashpenaz* to reeducate children in the language and culture of Babylon,
  - to serve in his court
- As part of the education, they would get to eat royal meat and drink royal wine



# Daniel in the Bible

Daniel, refused to touch royal meet.

- He proposed a 10-days experiment
  - to convince Ashpenaz that vegetarian diet is good as well
- Four children will be feed with vegetarian diet (treatment group)
- Four children will be feed with carnivore diet (control group)



# CONTROLLED STUDIES

Users are split into:

- **experimental group(s)**  
users do receive a recommendation
- **control group**  
users do not receive a recommendation

Differences in behaviour are analyzed



# CONTROLLED STUDIES

---

Users are split into:

- **experimental group(s)**  
users do receive a recommendation
- **control group**  
users do not receive a recommendation

Analogy with medical experiments:

- patients in experimental group(s) receive a drug
- patients in the control group receive a placebo or nothing

Differences in behaviour are analyzed

# A study on Social Media

Twitter made a notable controlled study (Huszar et al. 2021, PNAS)

In 2016, a new *personalised recommender* was introduced on the platform:

- **treatment group:**

4% of users → personalised recommendations

- **control group:**

1% of all global Twitter users → reverse chronological order

Personalised recommendations amplify political messages

# A study in Urban Mapping

Cornacchia et al. 2022 used a digital twin to simulate the impact of navigation services

Given an adoption rate  $r$ :

- $r\%$  of drivers (chosen at random) follow (blindly) the suggestions by a navigation service
- $(100-r)\%$  of drivers follow a noisy fastest route

The impact on CO<sub>2</sub> emissions is studied varying  $r$  from 0% to 100%.

# Question

---

There is crucial difference between controlled experiments in online platforms (e.g., social media) and controlled experiments in medicine.

Which one?

# CONTROLLED STUDIES

"the [potential outcome] observation on one unit should be unaffected by the particular assignment of treatments to the other units"

**D. R. Cox, Planning of Experiments, 1992**

- Known as the Stable Unit Treatment Value Assumption (SUTVA)

On online platforms, users in the control group can never be isolated from the indirect effects of recommendations: they are influenced by choices by users in the treatment group

- This violates SUTVA

# CONTROLLED STUDIES

## PROs

- Sample randomization reduces selection biases
- Allow isolating a variable's effect
- Allow causal inference \*

## CONs

- inclusion and exclusion criteria might limit the generalisability
- hard to design: require direct access to platforms' users and recommenders



# OBSERVATIONAL STUDIES

Data describe the behaviour of users **under a single recommendation principle**, without any control

- users are not split into separate groups *at the same time*
- in this definition, *quasi-experiments* are not controlled



# OBSERVATIONAL STUDIES

Data describe the behaviour of users **under a single recommendation principle**, without any control

- e.g., behaviour of Facebook users, routes followed by drivers

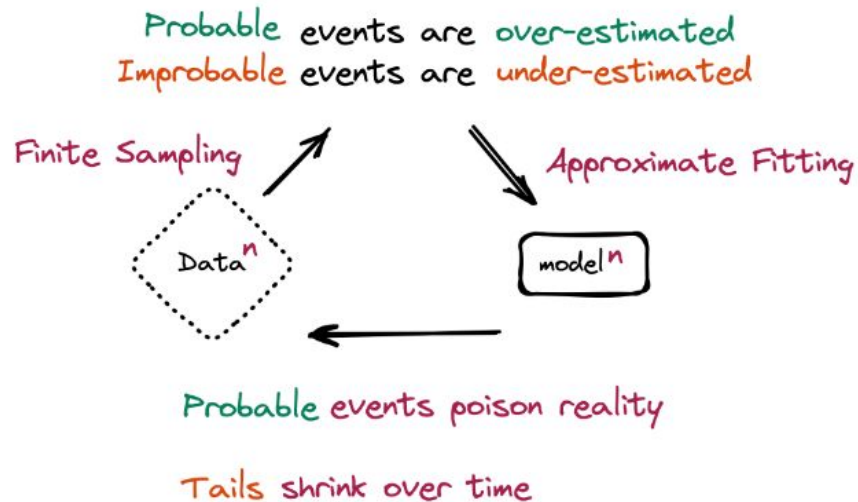


# A study on LLMs

Simulation controlled

What happens when LLMs are recursively trained on the synthetic data (**self-consuming loop**)?

- Text is generated, and it is used to fine-tune LLMs
- This is repeated for 10 times



# OBSERVATIONAL STUDIES

## PROs

- Offer broad insights when data is large and representative

## CONs

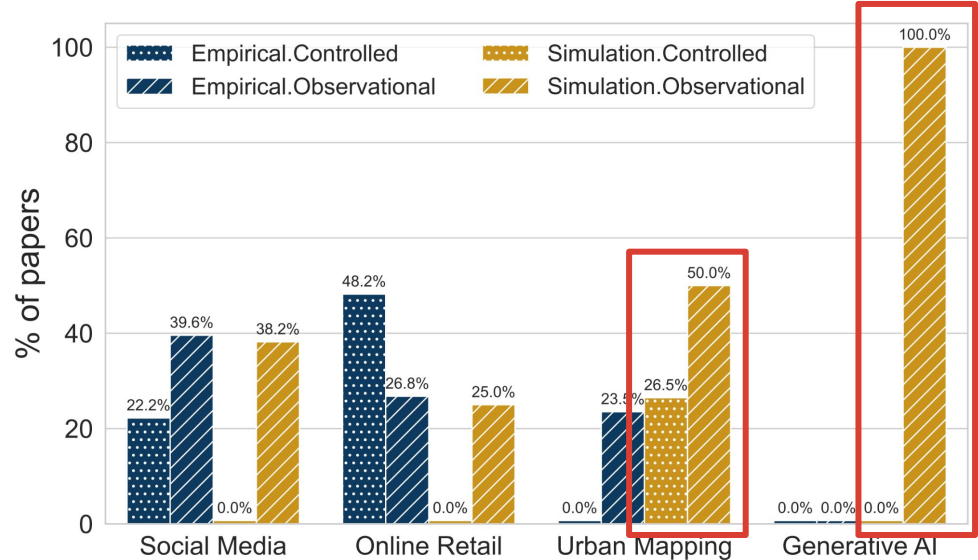
- Struggle to establish causal relationships
- Susceptible to biases, measurement errors, and confounding variables



# CONTROLLED vs OBSERVATIONAL

Observational studies are typically more common than controlled ones:

- they are easier to perform
- online retail is an exception



A survey of the impact of AI-based recommenders on human behaviour: methodologies, outcomes and future directions, ArXiv, 2024

# Quiz session

Experiment:

- A study selects 10 users on a social media platform
- On day 1, the users are exposed to recommender R1
- On day 2, the users are exposed to a recommender R2
- The number of new likes is computed for each user

Is this an empirical or a simulation study?

 A. Empirical

 B. Simulation

# Quiz session

Experiment:

- A study selects 10 users on a social media platform
- On day 1, the users are exposed to recommender R1
- On day 2, the users are exposed to a recommender R2
- The number of new likes is computed for each user

Is this an observational or controlled study?

✓ A. Observational

✗ B. Controlled

# Quiz session

Experiment:

- Two groups of users (G1-G2) are selected on an online retail platform
- G1 is exposed to recommender R1
- G2 is exposed to recommender R2
- The diversity of purchased products is measured

Is this an empirical or a simulation study?

✓ A. Empirical

✗ B. Simulation

# Quiz session

Experiment:

- Two groups of users (G1-G2) are selected on an online retail platform
- G1 is exposed to recommender R1
- G2 is exposed to recommender R2
- The diversity of purchased products is measured

Is this an observational or controlled study?

 A. Observational

 B. Controlled

# Quiz session

Experiment:

- Two groups of users (G1-G2) are selected on a platform
- G1 is exposed to recommender R1
- G2 is exposed to recommender R2
- The accuracy of R1 and R2 is evaluated

Is this an empirical or a simulation study?

**X** A. Empirical

**X** B. Simulation

# References

## Articles (useful for the project):

- L. Pappalardo et al. **A survey on the impact of AI-based recommenders on human behaviours: methodologies, outcomes and future directions**, 2024, <https://doi.org/10.48550/arXiv.2407.01630>
  - Section 2.3 Methodologies
  - Section 7.1 Methodologies
  - Section 7.2 Methodologies
- H. O. Stolberg et al. **Randomized controlled trials**. American Journal of Roentgenology 2004, <https://doi.org/10.2214/ajr.183.6.01831539>

# Books & articles

---

## To learn more:

- J. Pearl, **The Book of Why: The New Sciences of Cause and Effect**, Basic Books, 2018
- D. R. Cox, **Planning of Experiments**, John Wiley & Sons, 1992

## Intellectually stimulating:

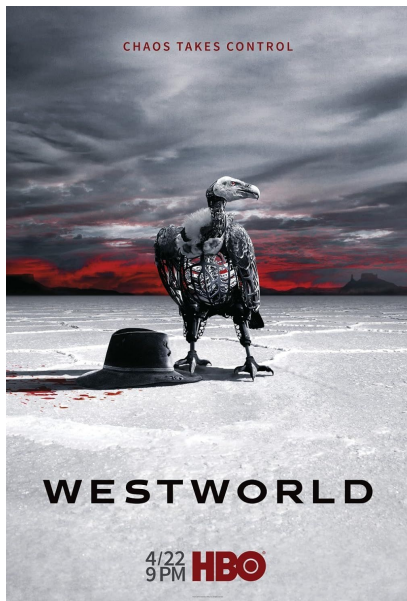
- C. Haney, W. C. Banks, P. G. Zimbardo, **A study of prisoners and guards in a simulated prison**, Naval Research Review 30, 1973
- J. W. Treece Jr., **Daniel and the Classic Experimental Design**, ICR.org
- A. Huxley, **Brave New World**, Chatto & Windus, 1932

# Movies

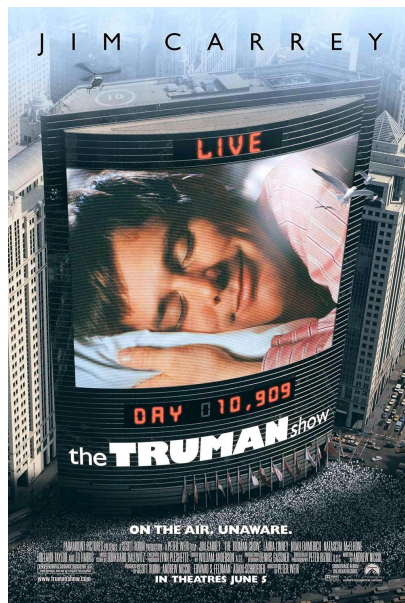
The Experiment  
2010



Westworld  
2016-2022



The Truman Show  
1998



Black Mirror  
2011-

