

Lecture 16

What's Next

Instructor: Chien-Ju (CJ) Ho

Logistics: Project Milestone 2

- Milestone 2: Due Nov 4
 - Summarize your progress
 - Should make enough progress to know **whether the project is feasible**
 - Last chance to convert a research project to an extensive literature survey
- Midterm Pitch: Nov 1 (Tue), in lecture
 - Opportunity to engage other classmates
 - Format
 - Every team gives a **1-min elevator** pitch about their projects
 - Data collection: 3 teams will perform data collection
 - Group discussions
 - Three big groups, with 3~4 teams per group
 - Each team gets **~10 minutes** of time from other teams
 - Everyone is expected to attend

Logistics: Midterm Project Pitch

- Time: Next Tuesday
- What to do beforehand
 - Prepare a 1-page slide and send it to me
 - Send me the slides via email by noon next Tuesday
 - Prepare a 1-min elevator pitch about your project
 - Think about how to utilize the discussion time

Logistics: Presentations and Peer Reviews

- Common suggestions in the peer reviews
 - Engage the audience more
 - eye contact
 - be enthusiastic
 - pace control
 - Succinct slides and visual aids are usually preferred

Lecture Today

Recap of what we have discussed so far

Introduce what we will discuss in the next few lectures

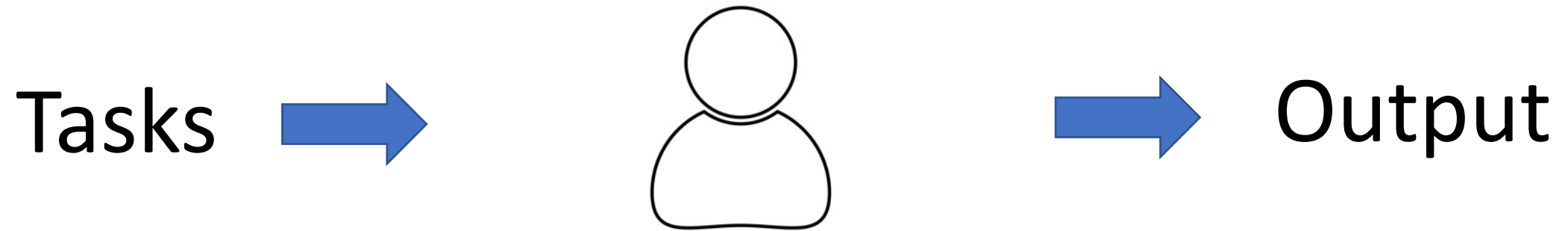
Discussion on related topics that are not covered in this course

What We Have Discussed So far

Human-in-the-Loop Computation:

From Task Solvers' Perspective

From Task Solvers' Perspective



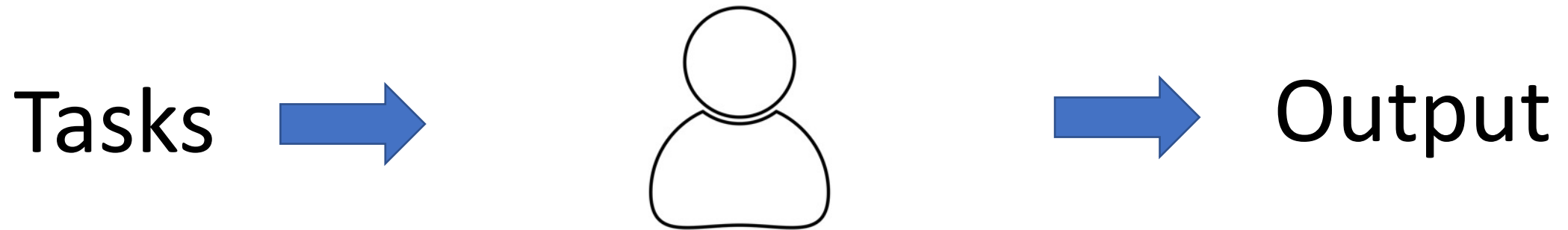
Microtasks



Flower
Dog
Cute

...

From Task Solvers' Perspective



Key Challenge: How To Ensure Output Quality

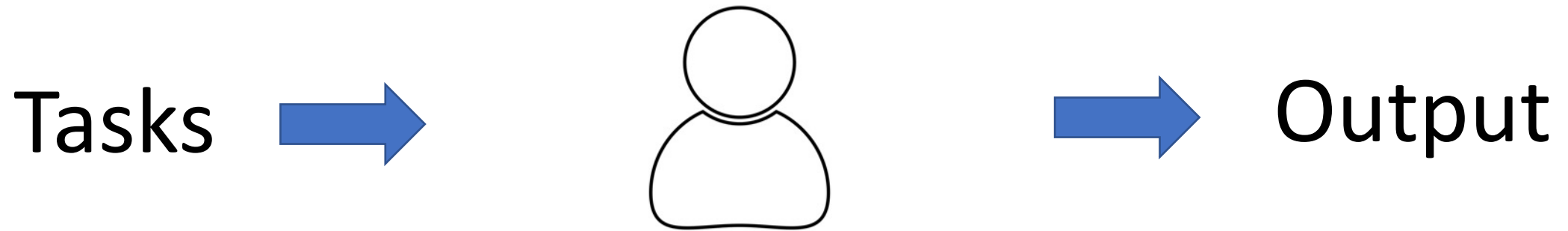
Human as data sources:
Label aggregation

Probabilistic reasoning to
aggregate noisy human data

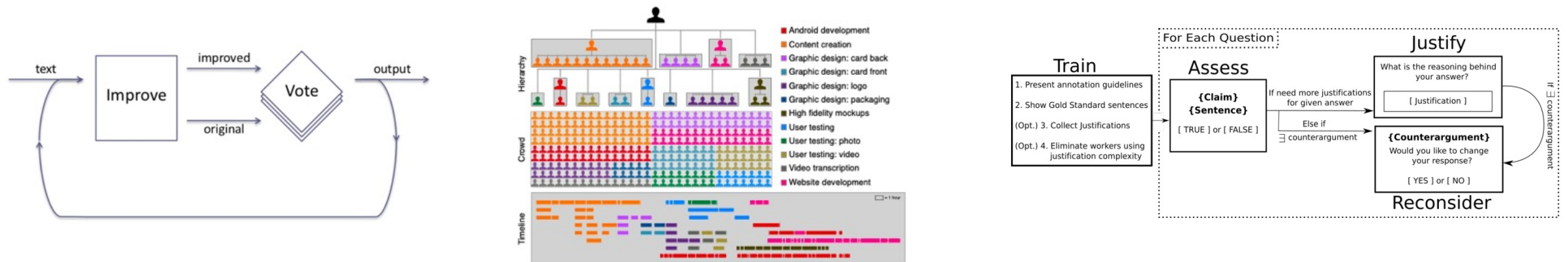
Humans are “Humans”:
Incentive design

Game theoretical modeling of
humans and incentive design

From Task Solvers' Perspective

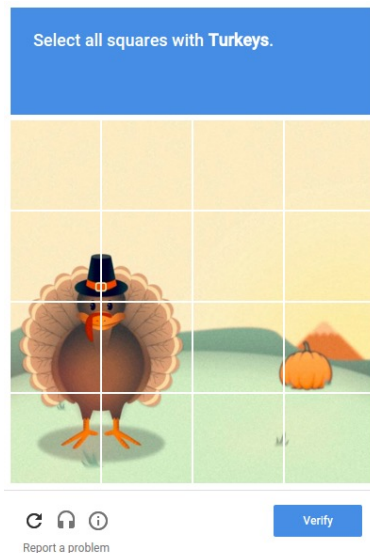


Practical challenges: Real-time, non-independent work, complex tasks



(From Lecture 1) What is this course about?

- Study the design and analysis of human-in-the-loop computation.



Human as data sources:

Label aggregation

Probabilistic reasoning to aggregate noisy human data

Practical challenges:

Real-time and complex tasks

Studies on workflow and team designs from HCI perspective

Humans are “Humans”:

Incentive design

Game theoretical modeling of humans and incentive design

Crowdsourcing and Human Computation



Human-in-the-Loop Computation

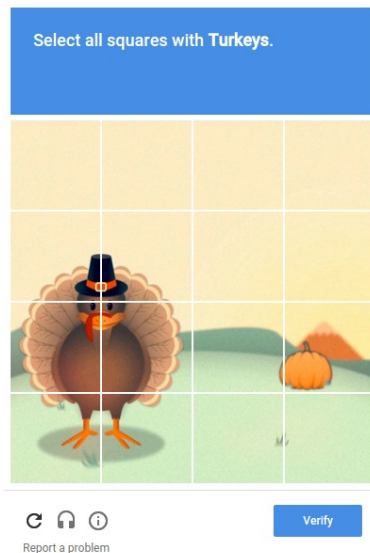
Remaining Lectures:

Focus on Human-AI Interactions

- Fairness in AI (Nov 3, 8)
 - Are the decisions made by AI *fair* to humans involved?
- Ethical decision making and participatory design (Nov 10)
 - Can we, and how can we, ask humans to help AI make "ethical" decisions
- Strategic machine learning (Nov 15)
 - What if humans who generated data are influenced by the outcome of AI?
- Interpretable machine learning (Nov 17)
 - How to make the predictions by AI understandable by humans
- Human-AI Collaboration (Nov 29 and Dec 1)
 - Can we team up humans and AI to solve tasks together?

(From Lecture 1) What is this course about?

- Study the design and analysis of human-in-the-loop computation.



Human as data sources:

Label aggregation

Probabilistic reasoning to aggregate noisy human data

Practical challenges:

Real-time and complex tasks

Studies on workflow and team designs from HCI perspective

Humans are “Humans”:

Incentive design

Game theoretical modeling of humans and incentive design

Selected recent topics:

Ethical issues of AI/ML, learning with strategic behavior, Human-AI collaborations.

The focus of the next seven lectures!

An Emerging Research Agenda on AI/ML + Humans/Society

- WashU Division of Computational and Data Sciences
 - PhD program hosted by CSE, Political Science, Social Work, Psychology and Brain Science
- Stanford Institute for Human-Centered Artificial Intelligence
- MIT Institute for Data, Systems, and Society
- CMU Societal Computing
- USC Center for AI in Society

Before we move on...

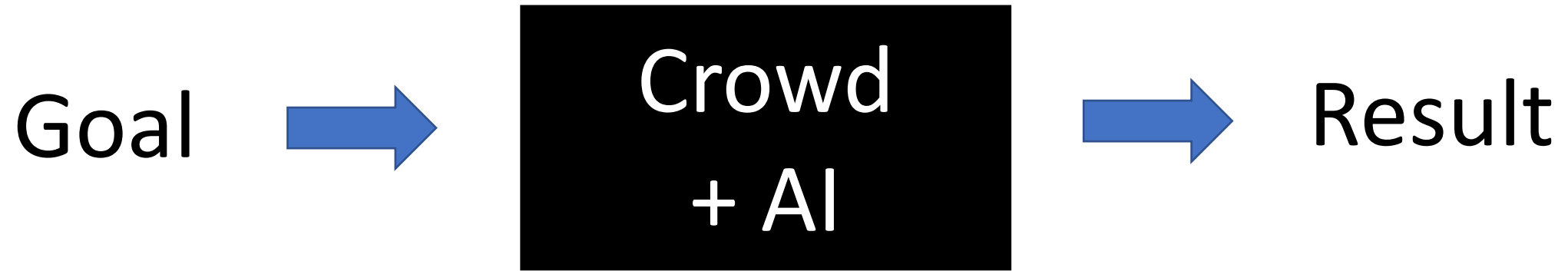
Required Reading:

Mathematical Foundations For Social Computing.

Yiling Chen, Arpita Ghosh, Michael Kearns, Tim Roughgarden, Jennifer Wortman Vaughan.

Communications of the ACM. 2016

Crowdsourcing Compiler



- Understanding humans, developing realistic *human models*, and incorporating them into the *computation framework*.
- Multidisciplinary in nature

Machine
Learning

Economics
Theory

Computational
Social Science

Optimization

Human-Computer
Interaction

and more...

Warm-Up Discussion

- What do you think are the top challenge(s) in designing/implementing the crowdsourcing compiler?
- For crowdsourcing to be sustainable, we need people to be willing to be part of it. What do you think are the important components we can/should/might add for crowdsourcing to enable a *real* career (e.g., adding career development, skill training, social support, etc)?

Brief Discussions on Topics that We Won't Cover

Beyond Solving Objective Tasks

- Fair division among the crowd
- Crowd research: open and scalable lab
- Crowdsourcing democracy
- Incentives in Blockchain

Fair Division (Resource Allocation)

- Classical example:
 - How to fairly split (fair: envy-free) the cake among two people?



- General research question:
 - How to design mechanisms to allocate resources with “good” properties
 - participants truthfully report their preferences
 - no one is envy of others

Fair Division

- Who should do the household chores



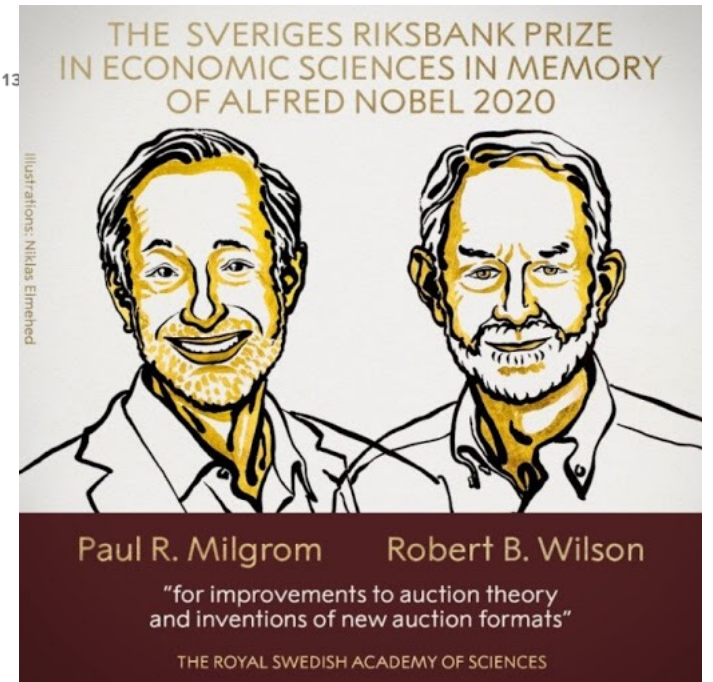
PERSONAL FINANCE

The Couple That Pays Each Other to Put Kids to Bed

PUBLISHED THU, FEB 13 2014 • 11:41 AM EST | UPDATED THU, FEB 13

Essentially a **second-price auction**:

- Each “bid” how much she/he thinks the work is worth
- High bidder pays the amount the low bidder bids
- Low bidder does the work and gets paid



Fair Division

- Spliddit



Share Rent



Split Fare



Assign Credit



Divide Goods

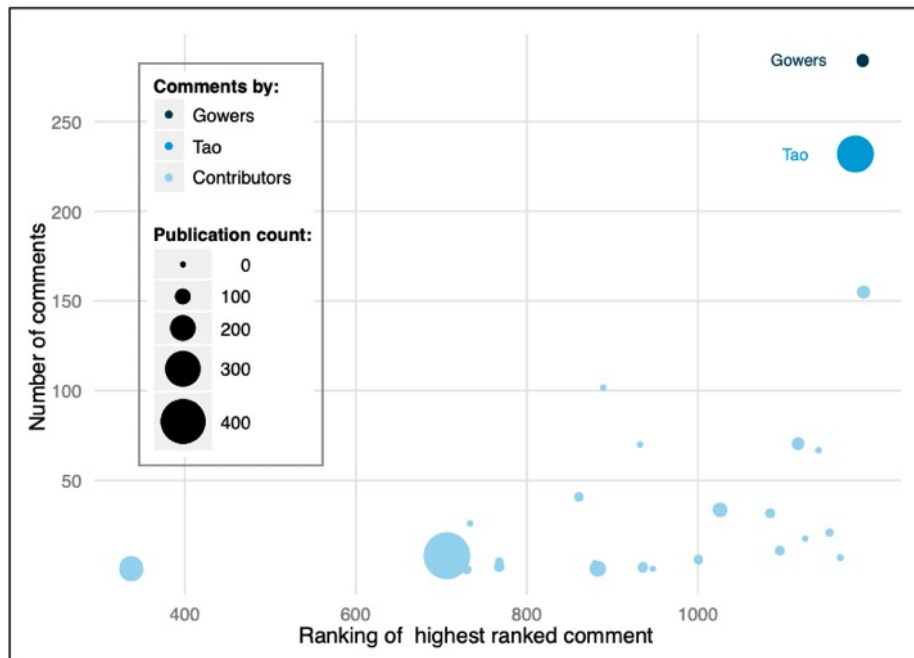


Distribute Tasks

- How do we allocate payments / credits to workers who work together for a complex task?
 - What are the "good" properties we want to achieve
 - How to design algorithms to achieve that?
- Fair division / resource allocation on societal issues
 - How do we allocate "donated organs" to patients who need them?
 - How do we allocate government resources to homeless people?
 - How do we allocate COVID vaccines to people.

Crowd Research: Scalable and Open Lab

- Most research projects are done by small groups of researchers
- Can we scale up research as well?
- Success story:
 - Polymath project: Collaborative Math Problem Solving
 - Published papers under the pseudonym **D.H.J. Polymath**.



Majority of contributions are done by a few

- Timothy Gowers (U Cambridge)
- Terence Tao (UCLA)

Many have made solid contributions

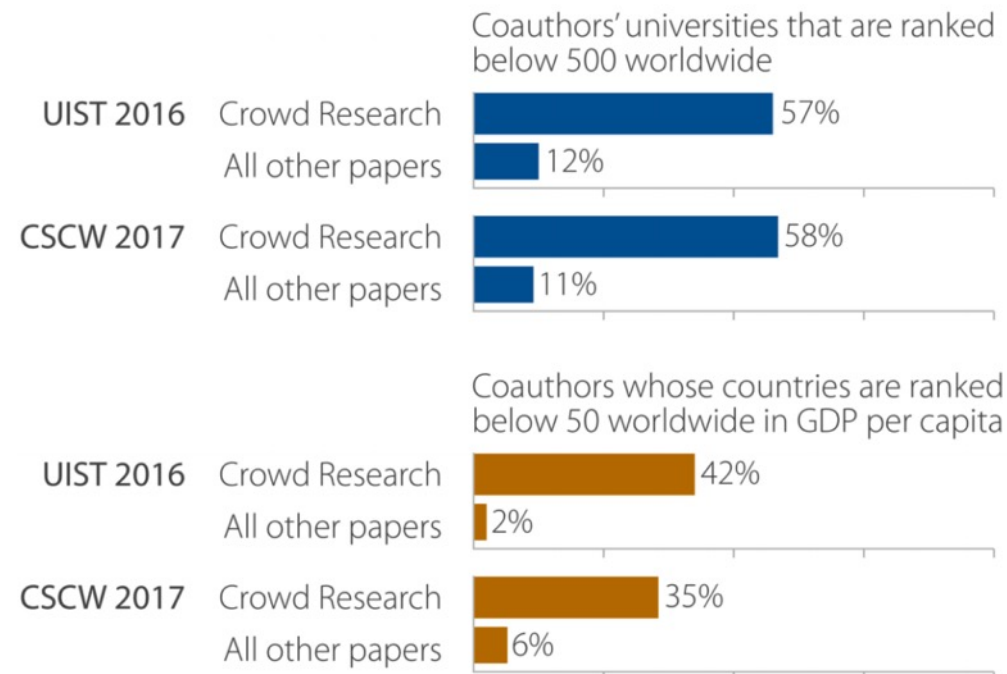
Crowd Research: Scalable and Open Lab

- Incorporating diverse thoughts / skills
 - 1500 participants from 62 countries for 3 research projects



Crowd Research: Scalable and Open Lab

- Incorporating diverse thoughts / skills
 - 1500 participants from 62 countries for 3 research projects
- Enabling research opportunities to students in less resourceful institutions



Crowd Research: Scalable and Open Lab

- Challenges
 - How to maintain the research progress?
 - How to distribute the credits?



Crowd Research: Scalable and Open Lab

- Outcome:
 - A couple of work-in-progress reports
 - Two top-tier conference papers (UIST'16, CSCW'17)
 - Recommendation letters are provided to participants with significant contributions for graduate school applications

Boomerang: Rebounding the Consequences of Reputation Feedback on Crowdsourcing Platforms

Snehalkumar (Neil) S. Gaikwad, Durim Morina, Adam Ginzberg, Catherine Mullings, Shirish Goyal, Dilrukshi Gamage, Christopher Diemert, Mathias Burton, Sharon Zhou, Mark Whiting, Karolina Ziulkoski, Aipta Ballav, Aaron Gilbee, Senadhipathige S. Niranga, Vibhor Sehgal, Jasmine Lin, Leonardy Kristianto, Angela Richmond-Fuller, Jeff Regino, Nalin Chhibber, Dinesh Majeti, Sachin Sharma, Kamila Mananova, Dinesh Dhakal, William Dai, Victoria Purynova, Samarth Sandeep, Varshine Chandrakanthan, Tejas Sarma, Sekandar Matin, Ahmed Nasser, Rohit Nistala, Alexander Stolzoff, Kristy Milland, Vinayak Mathur, Rajan Vaish, Michael S. Bernstein

Stanford Crowd Research Collective
Stanford University
daemo@cs.stanford.edu

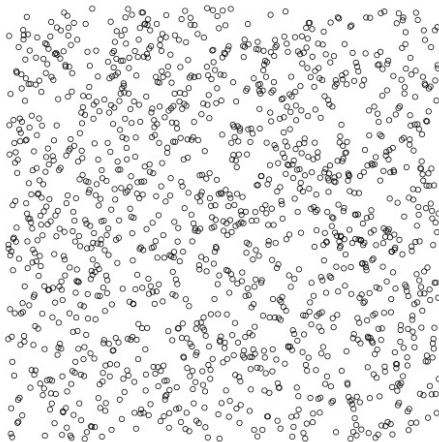
Crowd Guilds: Worker-led Reputation and Feedback on Crowdsourcing Platforms

Mark E. Whiting, Dilrukshi Gamage, Snehalkumar (Neil) S. Gaikwad, Aaron Gilbee, Shirish Goyal, Aipta Ballav, Dinesh Majeti, Nalin Chhibber, Angela Richmond-Fuller, Freddie Vargus, Tejas Seshadri Sarma, Varshine Chandrakanthan, Teogenes Moura, Mohamed Hashim Salih, Gabriel Bayomi Tinoco Kalejaiye, Adam Ginzberg, Catherine A. Mullings, Yoni Dayan, Kristy Milland, Henrique Orefice, Jeff Regino, Sayna Parsi, Kunz Mainali, Vibhor Sehgal, Sekandar Matin, Akshansh Sinha, Rajan Vaish, Michael S. Bernstein

Stanford Crowd Research Collective
daemo@cs.stanford.edu

Crowdsourcing Democracy

- Democracy is a crowdsourcing process
 - Vote for a leader to make decisions
 - Vote to determine the policy through referendum
 - And more
- Is the crowd always wise?



256	649	900	1296
350	650	978	1500
375	700	1000	1700
387	720	1008	2000
455	730	1024	2100
494	739	1028	2500
519	800	1200	2500
550	800	1200	3000
625	847	1232	10,000
625	899	1250	102000

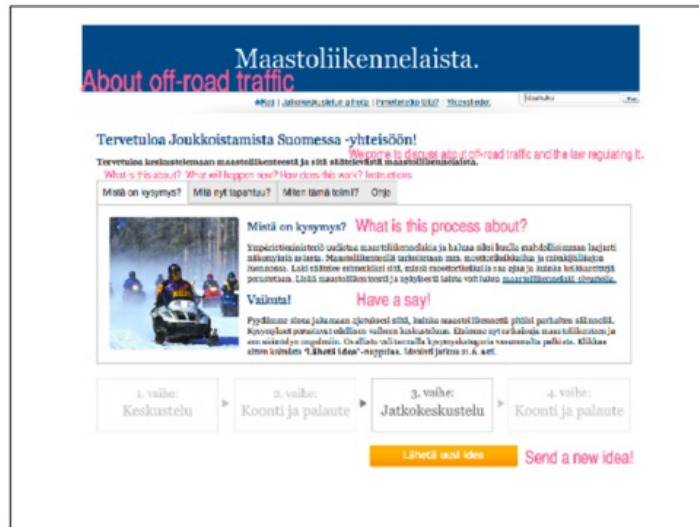
Mean: 3789.65

Median: 899.5

True Answer: 721

Crowdsourcing Democracy

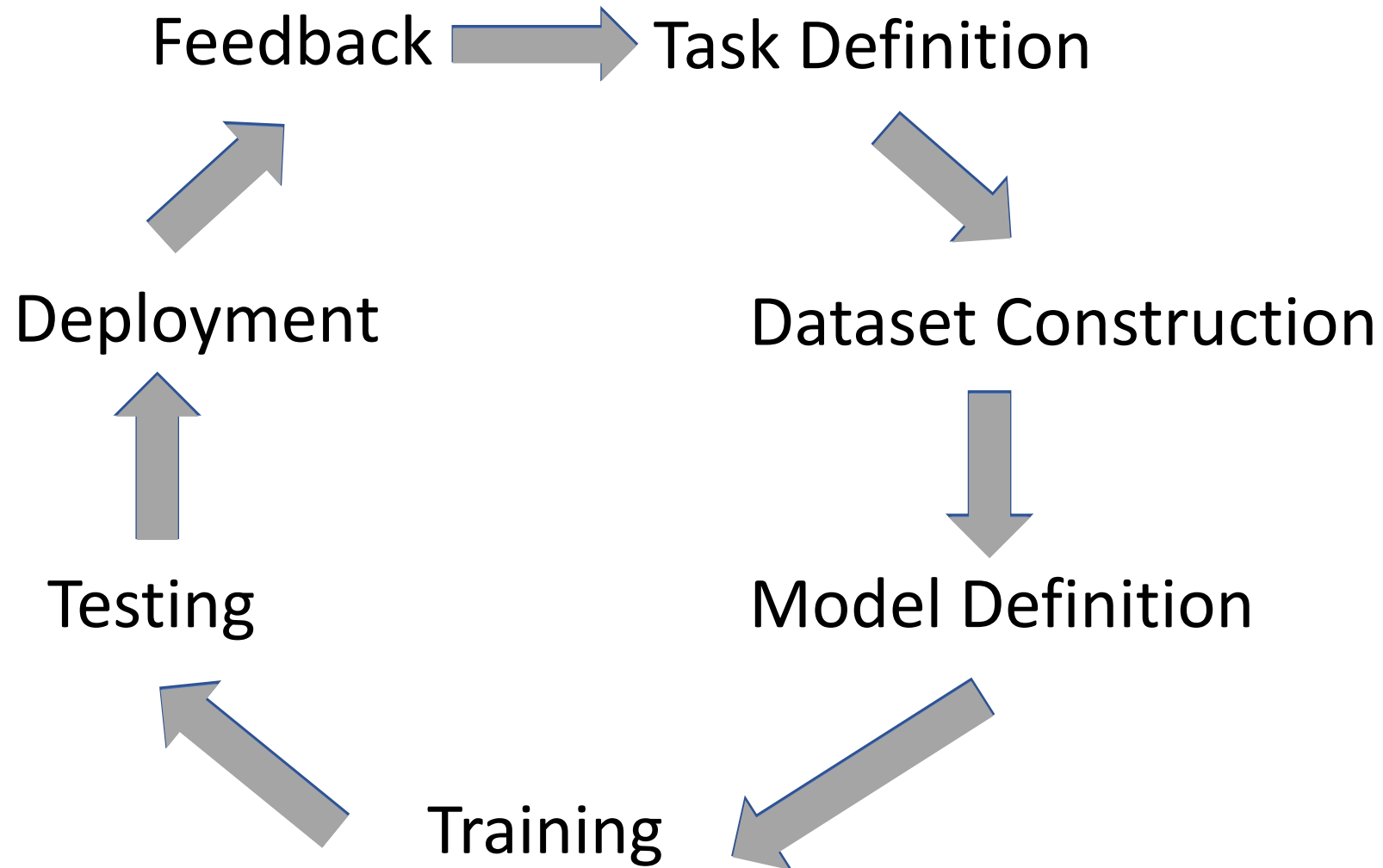
- Helping the crowd to make more informed decisions
 - E.g., enabling information exchange, deliberation



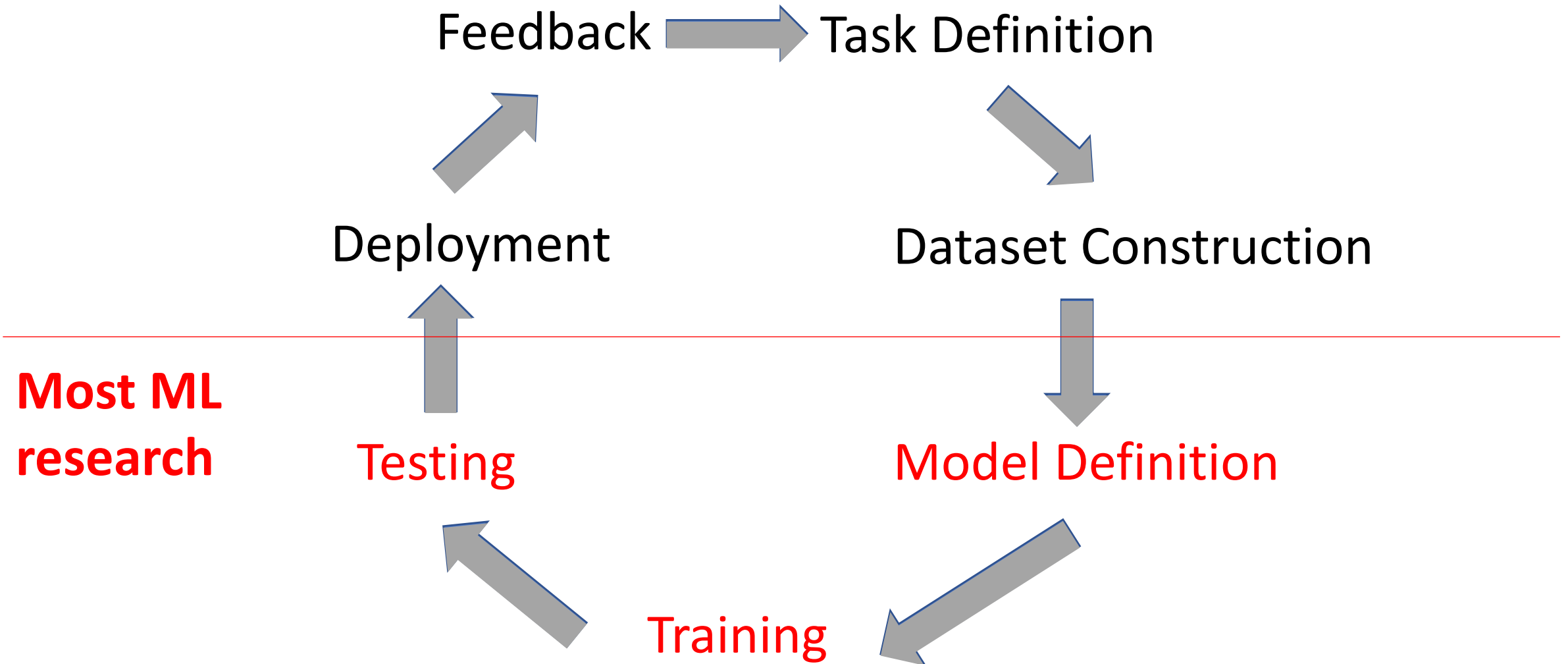
- Potential issues to be careful about
 - Fake news – Misinformation
 - Polarization in social networks

Human-in-the-Loop Computation: From Machine Learning Perspective

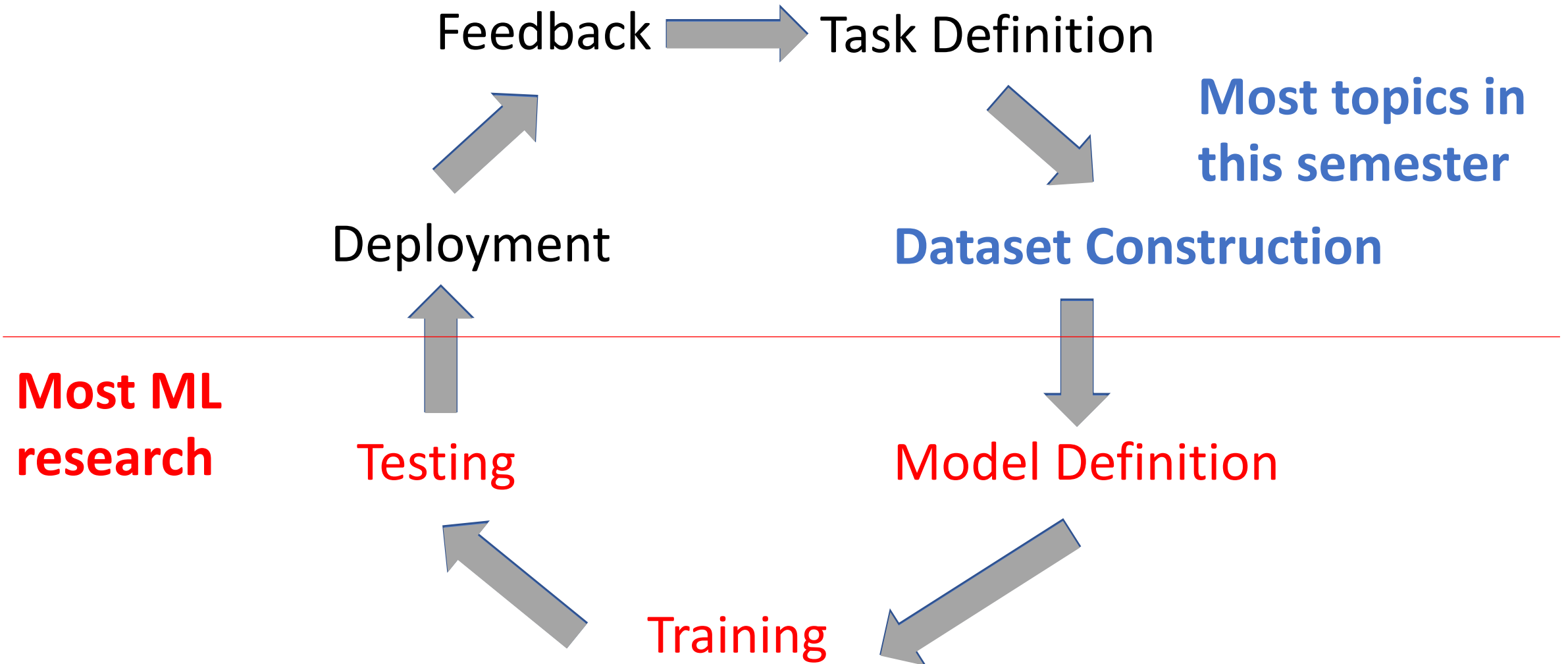
Machine Learning Lifecycle



Machine Learning Lifecycle

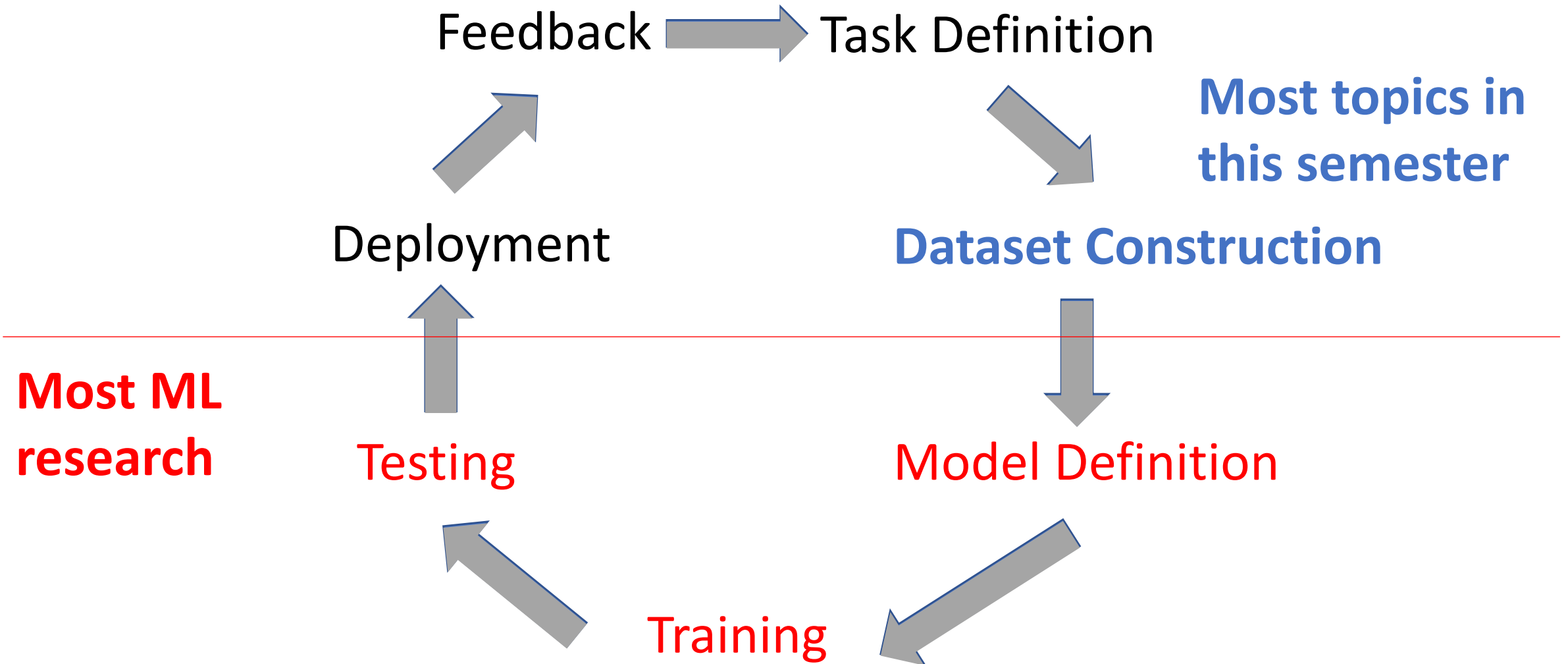


Machine Learning Lifecycle

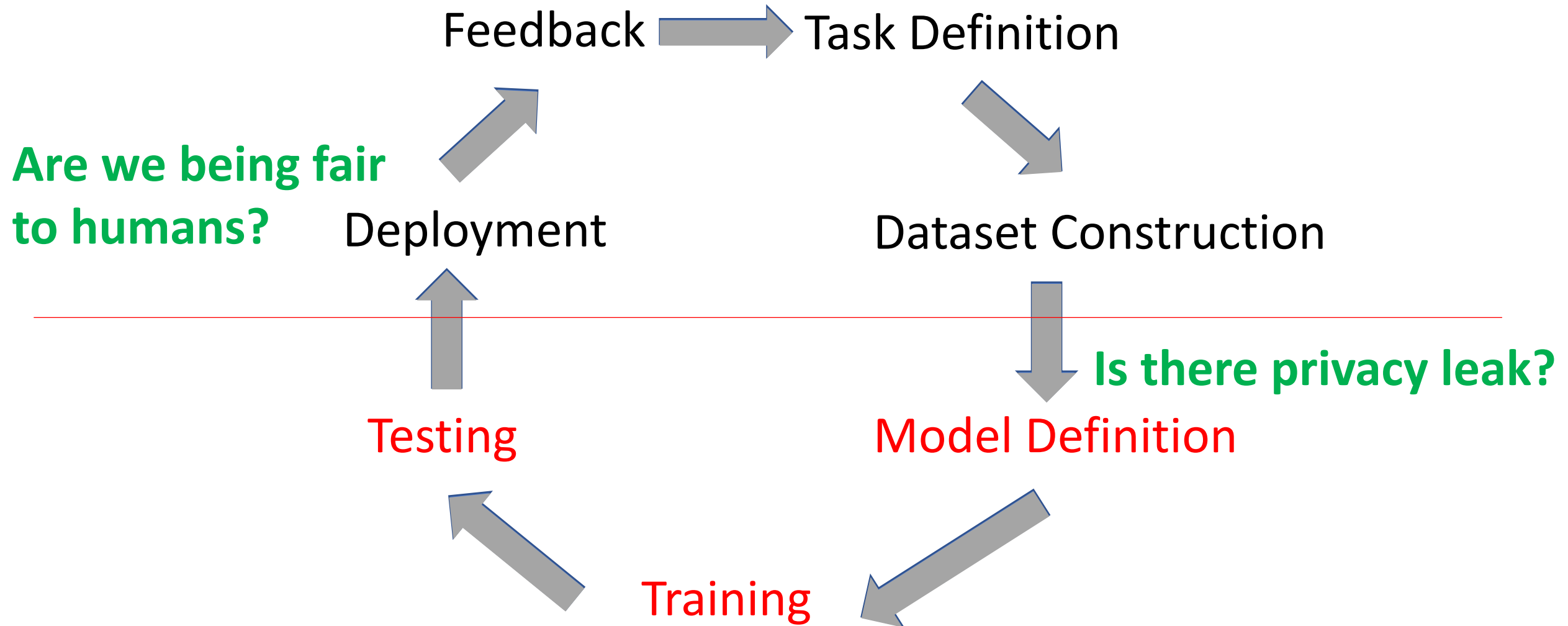


Machine Learning Lifecycle

Humans can be involved in every aspect of the process



Machine Learning Lifecycle



Discussion on Privacy

Netflix Challenges

- First Netflix challenge
 - Announced in 2006
 - Released a dataset of 100,480,507 ratings that 480,189 users gave to 17,770 movies.
 - Award \$1 million to first team beating their algorithm by 10%
 - Data format: <user, movie, date of grade, grade>
 - User and movie names are replaced with integers
- Is there a second Netflix challenge?
 - Announced in August 2009
 - Cancelled in March 2010
 - Why?
 - Privacy lawsuits and FTC involvements

Robust De-anonymization of Large Sparse Datasets

Arvind Narayanan and Vitaly Shmatikov

The University of Texas at Austin

Netflix Dataset



IMDB Data

Why is Anonymization Hard?

- Even without explicit identifiable information (e.g., ID, name), other detailed information about you might still reveal who you are

<i>office</i>	<i>department</i>	<i>date joined</i>	<i>salary</i>	<i>d.o.b.</i>	<i>nationality</i>	<i>gender</i>
London	IT	Apr 2015	£####	May 1985	Portuguese	Female

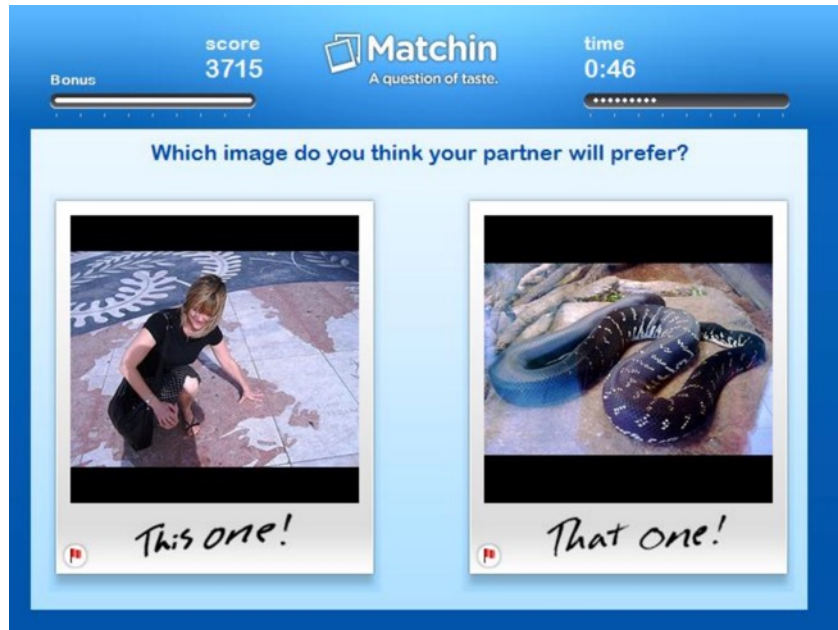
- What can we do?
 - Adding noises

<i>office</i>	<i>department</i>	<i>date joined</i>	<i>salary</i>	<i>d.o.b.</i>	<i>nationality</i>	<i>gender</i>
UK	IT	2015	£####	1980-1985	—	Female

Tradeoff between **privacy** and **utility**

Another Example

- Matchin: A Game for Collecting User Preferences on Images



- Building gender models using user labels
- Ask MTurk workers to compare 10 pairs of images.
 - Accuracy for guessing the gender: 78.3%

Unreasonable Privacy Expectations

- Can we get privacy for free?
 - No, privatizing means information loss (\Rightarrow accuracy loss)
- Absolute privacy is not likely.
 - Who you are friends with might reveal who you are

September 22, 2009 by [Ben Terris](#)



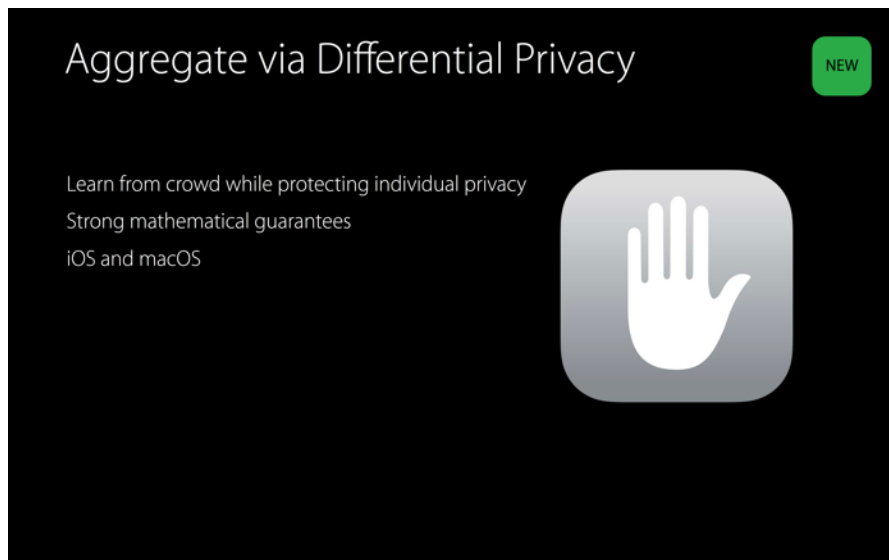
MIT Students' Facebook 'Gaydar' Raises Privacy Issues

(Maybe) More Reasonable Expectations

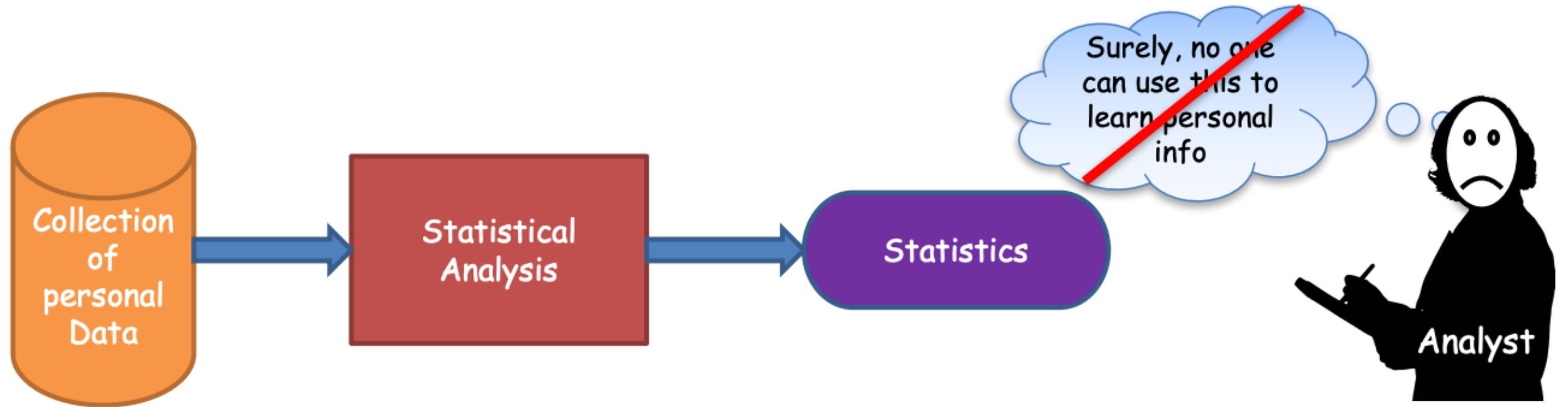
- Quantitative
 - Want a knob to tune the tradeoff between accuracy and privacy loss
- Plausible deniability
 - Your presence in a database cannot be ascertained
- Prevent targeted attacks
 - Limit information leaked even with side knowledge

Differential Privacy

- A formal notion to characterize privacy.
- History
 - Proposed by Dwork et al. 2006
 - Win the Gödel Prize in 2017
 - Apple announced to adopt the notion of differential privacy in iOS 10 in 2016

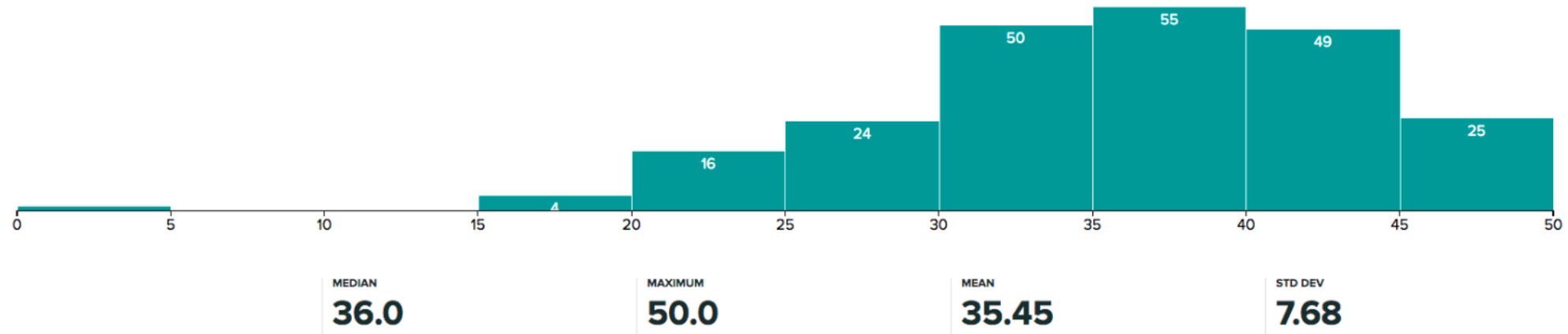


Differential Privacy



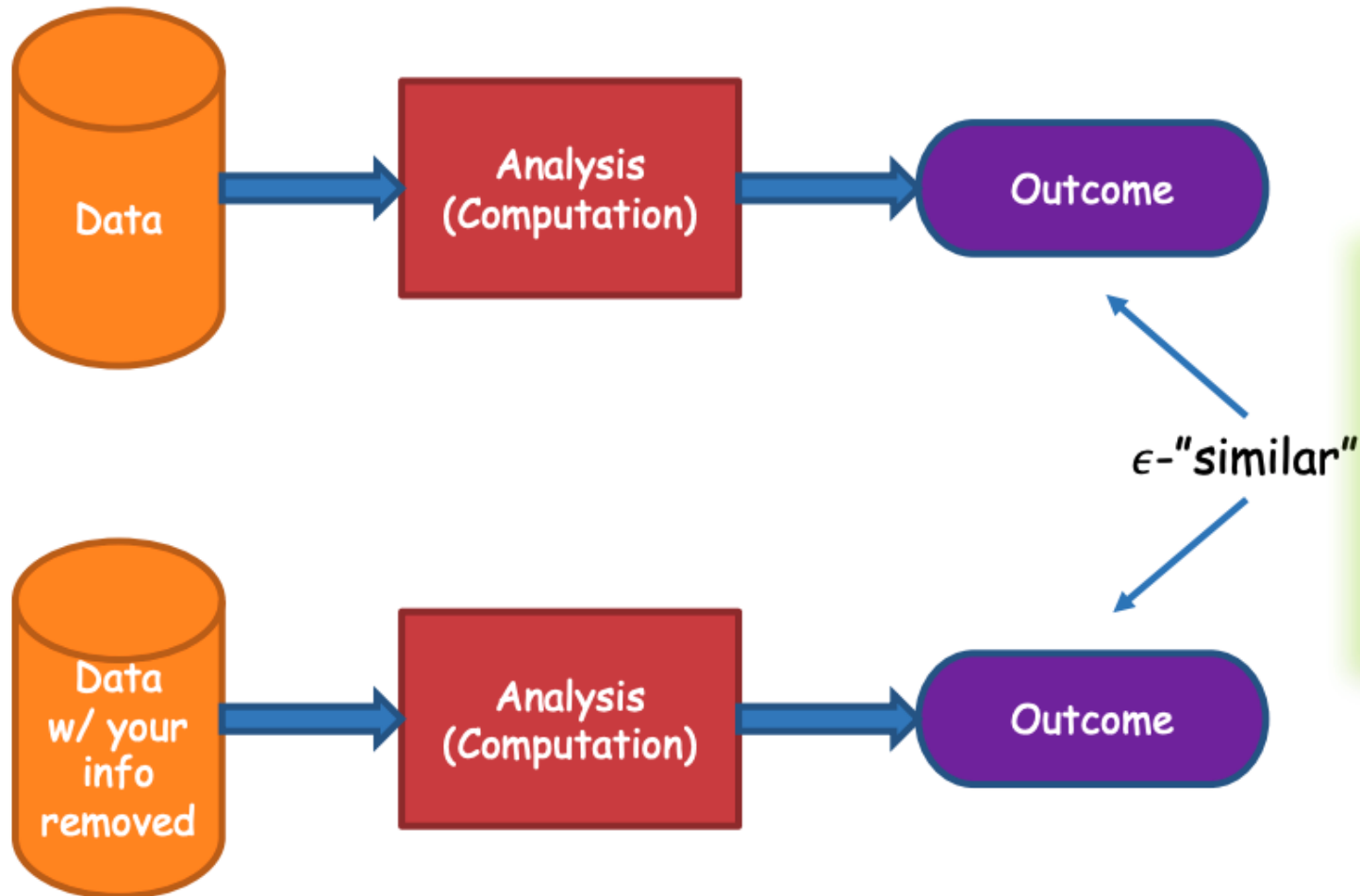
Differential Privacy

- Assume we have an exam in this course. And I have distributed this score distribution.



- How much of the private information (your individual grades) do I reveal?
- What if there are only 2 students in the class?

Differential Privacy



Differential Privacy

- Notations

- A : a randomized algorithm.
- D_1, D_2 : two “neighboring” database (with only one-entry difference)
- ϵ : privacy budget

- ϵ -differentially private

- A is ϵ -differentially private if for any neighboring databases D_1 and D_2 , and for any algorithm output Y , we have

$$\Pr[A(D_1) \in Y] \leq e^\epsilon \Pr[A(D_2) \in Y]$$

$$e^\epsilon \approx 1 + \epsilon \text{ when } \epsilon \text{ is small}$$

Intuition:







































































































The change of output is small
if the change of data is small

How to Be Differentially Private

- Let the output of A be the average of users' ages
- Consider two extreme cases
 - If the size of the database is 1
 - If the size of the database is infinity
- Add noise
 - We can tune the amount of noise to tradeoff privacy and accuracy
- A majority of the differentially private algorithms use a similar approach

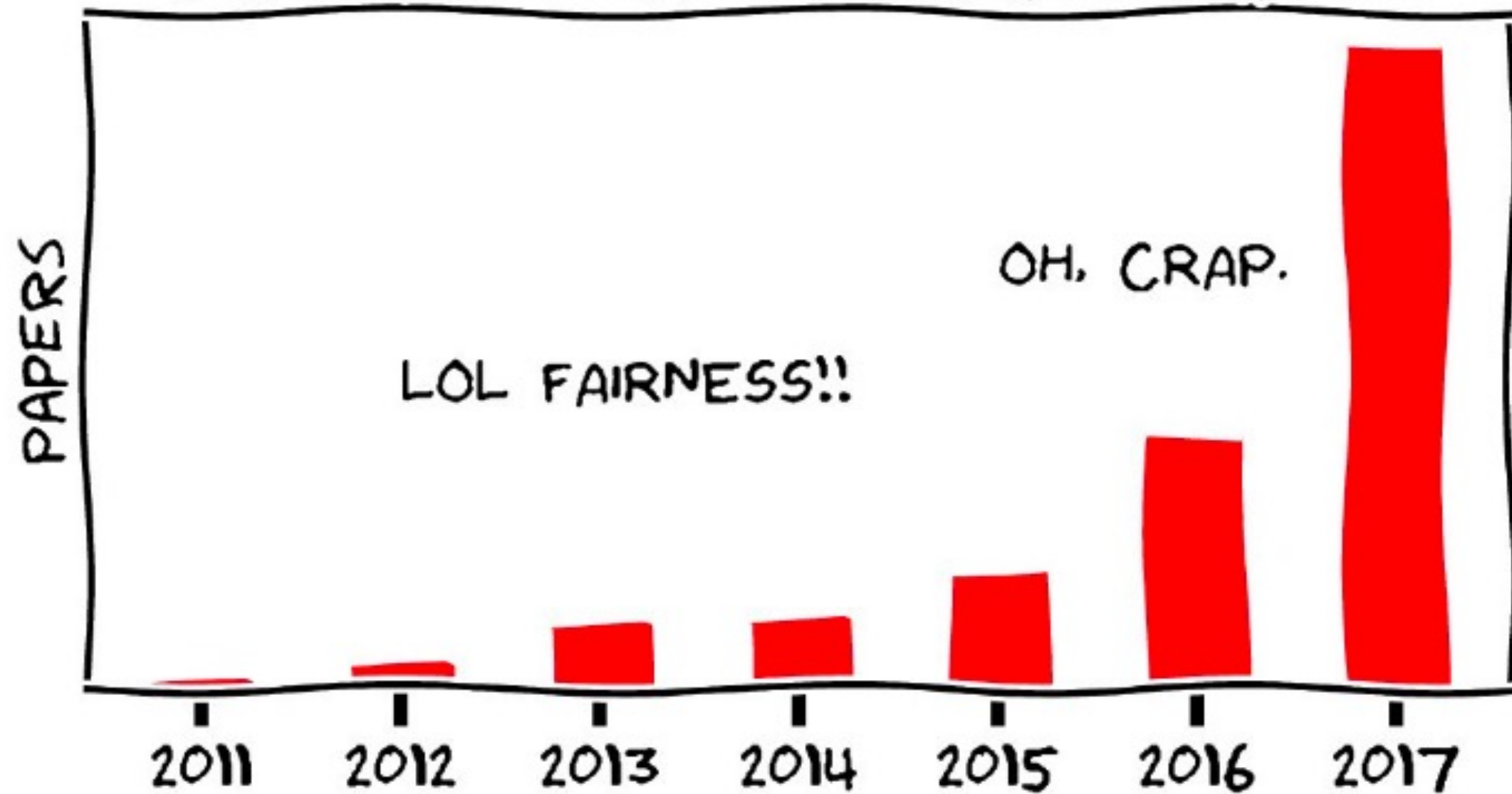
Discussion

- Differential privacy is a formal tool that we can tune the privacy budget to tradeoff privacy and utility/accuracy.
- We have been giving the big tech companies a lot of information. Have you been worried about any of the privacy issues? What's the line you will choose privacy or utility?

	Google	Facebook	Apple	Twitter	Amazon	Microsoft
Name						
Gender						
Birthday						
Phone Number						
Email Address						
Location						
Relationship Status						
Work						
Income Level						
Education						
Race/Ethnicity						
Religious Views						
Physical Address						
Facial Recognition Data						
Political Views						
Credit Cards						
Government IDs (Such as Social Security)						

Fairness

BRIEF HISTORY OF FAIRNESS IN ML



Isn't the point of ML to discriminate?

Want to avoid “unjustified” discrimination.

Example: Loan Applications

- By law, the banks can't discriminate people according to their race.
- First natural approach (fairness through blindness)
 - remove the race attribute from the data
- Guess what happened?
 - Redlining



What should we do?

- From computer scientists / engineers' point of view....
- Give me an operational definition of fairness, I'll implement a system that satisfy it!
- How should we define fairness? Is it even possible to define an universal fairness notion?
- More in the next few lectures.