Crowdcomputing and Citizen Science for Large-scale Experiments

Snehalkumar `Neil' S. Gaikwad[*], Sohan Dsouza[*], Oana Vuculescu[†], Andrew Mao[†‡], Iyad Rahwan[*]

Massachusetts Institute of Technology[*] Microsoft Research New York[‡] Aarhus University[†]

The International Conference on Computational Social Science, 2017, Germany

July 10, 2017

A BRIEF HISTORY OF CROWD COMPUTING

CROWD MECHANICS

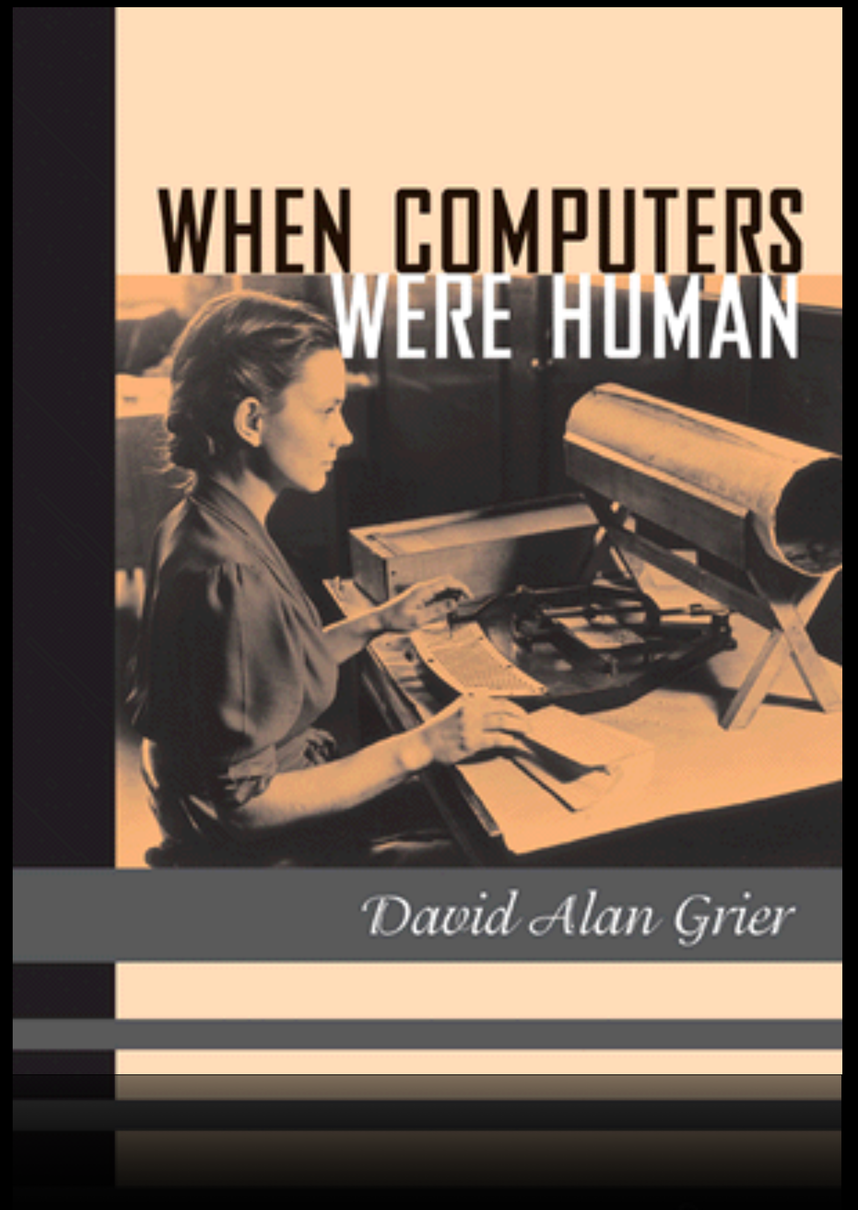
INSIGHTS

Interactive Case Discussions
You choose a problem

I
A Brief History of Crowd Computing

In the 19th-20th century

when computers
were human



The Mathematical Tables Project

450 human computers

[Grier 2007]

Hidden Figures

Human Computers and IBM 704
source: JPL

Human Computers 1936
source: JPL

Ranger 7, the first successful U.S. mission to the moon
source: JPL

In the 21st century

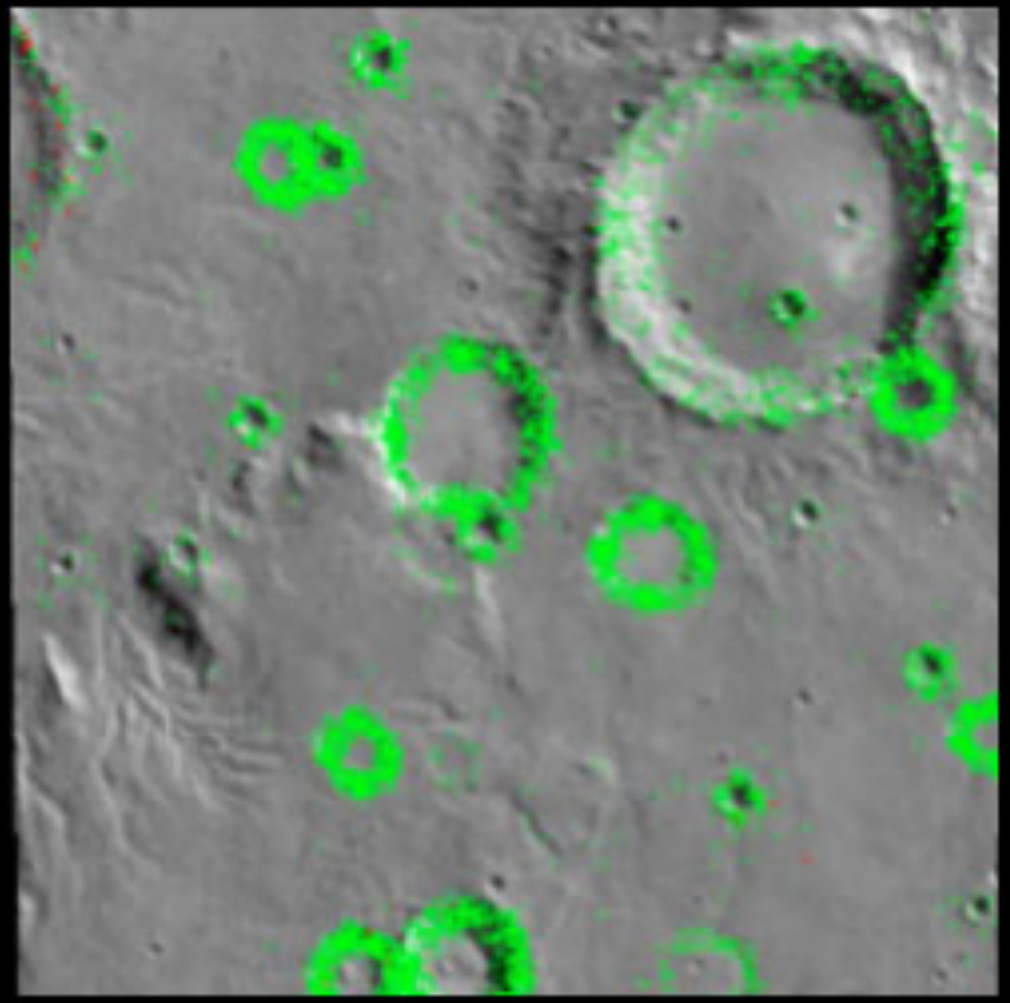

Rise of the Connected World

Power of the Crowd

crowd as a computational process

Count Craters

Clickworkers & Be a Martian

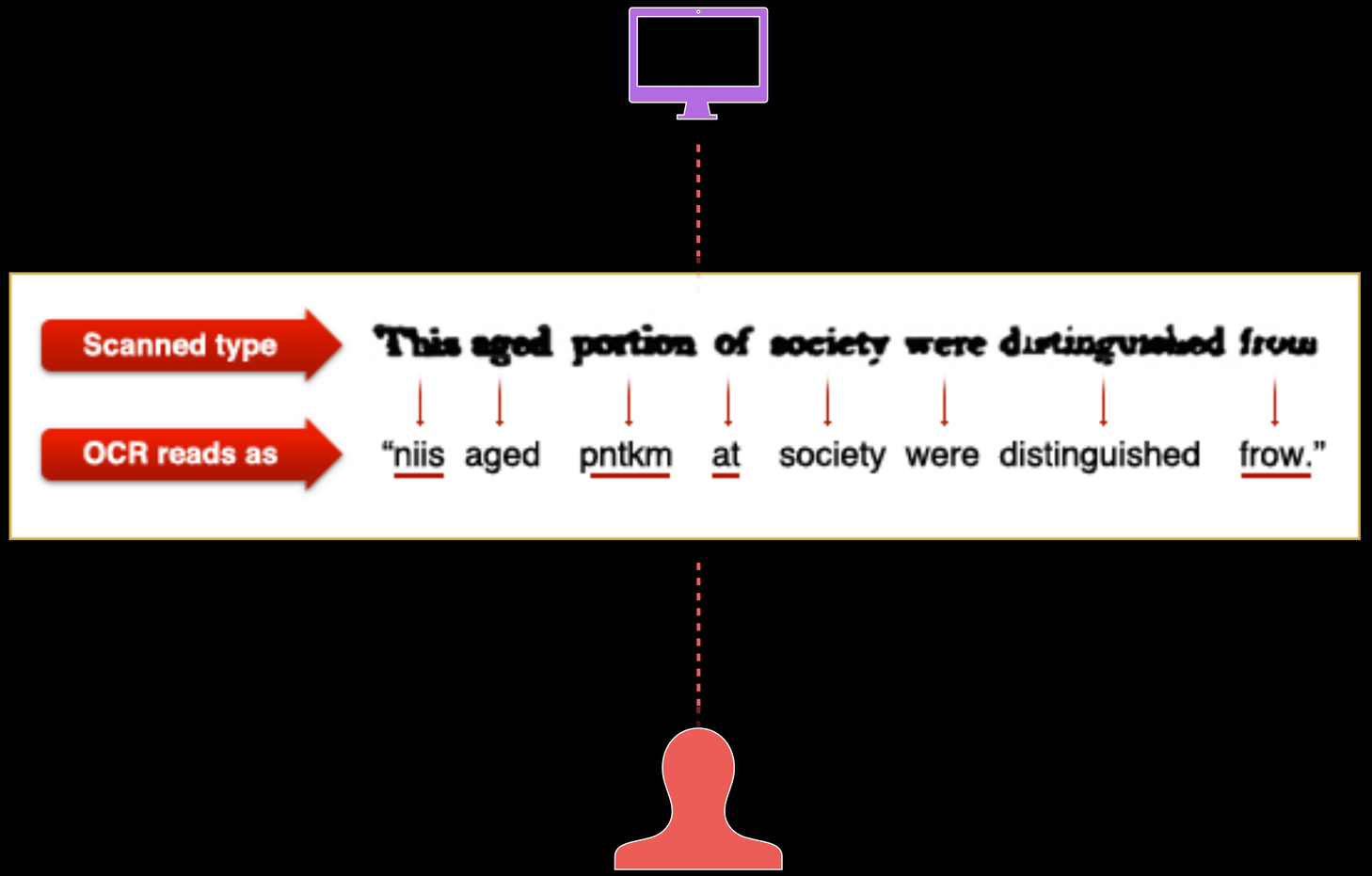


source: JPL

reCAPTCHA

[Luis von Ahn et al. Science, 2008]

reCAPTCHA



[Luis von Ahn et al. Science, 2008]

Machine Vision

Machine Vision


```
0.1921569 0.1803922 0.1529412 0.1411765 0.1529412 0.1686275 0
0.2078431 0.1921569 0.1647059 0.1568627 0.1686275 0.1843137 0
0.2117647 0.1960784 0.1686275 0.1607843 0.1725490 0.1882353 0.2156863 0.2313725
0.2039216 0.1921569 0.1647059 0.1568627 0.1686275 0.1882353 0.2156863 0.2392157
0.1921569 0.1803922 0.1529412 0.1450980 0.1568627 0.1803922 0.2117647 0.2392157
0.1882353 0.1725490 0.1450980 0.1372549 0.1529412 0.1764706 0.2117647 0.2392157
0.1921569 0.1764706 0.1490196 0.1411765 0.1529412 0.1725490 0.2078431 0.2313725
0.2039216 0.1882353 0.1568627 0.1450980 0.1568627 0.1764706 0.2039216 0.2274510
0.2156863 0.2000000 0.1686275 0.1568627 0.1686275 0.1843137 0.2078431 0.2235294
0.2196078 0.2039216 0.1725490 0.1607843 0.1647059 0.1803922 0.2156863 0.2352941
0.2235294 0.2078431 0.1725490 0.1529412 0.1450980 0.1686275 0.2235294 0.2549020
0.2196078 0.2039216 0.1686275 0.1490196 0.1372549 0.1647059 0.2313725 0.2666667
0.2078431 0.1921569 0.1568627 0.1411765 0.1372549 0.1686275 0.2352941 0.2745098
0.2000000 0.1843137 0.1490196 0.1333333 0.1372549 0.1725490 0.2392157 0.2745098
```

Games with Purpose

0.1921569 0.1803922 0.1529412 0.1411765 0.1529412 0.1686275 0.1882353 0.2039216
0.2078431 0.1921569 0.1647059 0.1568627 0.1686275 0.1843137 0.2039216 0.2196078
0.2117647 0.1960784 0.1686275 0.1607843 0.1725490 0.1882353 0.2156863 0.2313725
0.2039216 0.1921569 0.1647059 0.1568627 0.1686275 0.1882353 0.2156863 0.2392157
0.1921569 0.1803922 0.1529412 0.1450980 0.1568627 0.1803922 0.2117647 0.2392157
0.1882353 0.1725490 0.1450980 0.1372549 0.1529412 0.1764706 0.2117647 0.2392157
0.1921569 0.1764706 0.1490196 0.1411765 0.1529412 0.1725490 0.2078431 0.2313725
0.2039216 0.1882353 0.1568627 0.1450980 0.1568627 0.1764706 0.2039216 0.2274510
0.2156863 0.2000000 0.1686275 0.1568627 0.1686275 0.1843137 0.2078431 0.2235294
0.2196078 0.2039216 0.1725490 0.1607843 0.1647059 0.1803922 0.2156863 0.2352941
0.2235294 0.2078431 0.1725490 0.1529412 0.1450980 0.1686275 0.2235294 0.2549020
0.2196078 0.2039216 0.1686275 0.1490196 0.1372549 0.1647059 0.2313725 0.2666667
0.2078431 0.1921569 0.1568627 0.1411765 0.1372549 0.1686275 0.2352941 0.2745098
0.2000000 0.1843137 0.1490196 0.1333333 0.1372549 0.1725490 0.2392157 0.2745098

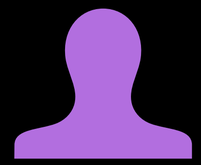
[Von Ahn et al. *CHI*, 2004]

Games with Purpose


```
0.1921569 0.1803922 0.1529412 0.1411765 0.1529412 0.1686275 0.1882353 0.2039216
0.2078431 0.1921569 0.1647059 0.1568627 0.1686275 0.1843137 0.2039216 0.2196078
0.2117647 0.1960784 0.1686275 0.1607843 0.1725490 0.1882353 0.2156863 0.2313725
0.2039216 0.1921569 0.1647059 0.1568627 0.1686275 0.1882353 0.2156863 0.2392157
0.1921569 0.1803922 0.1529412 0.1450980 0.1568627 0.1803922 0.2117647 0.2392157
0.1882353 0.1725490 0.1450980 0.1372549 0.1529412 0.1764706 0.2117647 0.2392157
0.1921569 0.1764706 0.1490196 0.1411765 0.1529412 0.1725490 0.2078431 0.2313725
0.2039216 0.1882353 0.1568627 0.1450980 0.1568627 0.1764706 0.2039216 0.2274510
0.2156863 0.2000000 0.1686275 0.1568627 0.1686275 0.1843137 0.2078431 0.2235294
0.2196078 0.2039216 0.1725490 0.1607843 0.1647059 0.1803922 0.2156863 0.2352941
0.2235294 0.2078431 0.1725490 0.1529412 0.1450980 0.1686275 0.2235294 0.2549020
0.2196078 0.2039216 0.1686275 0.1490196 0.1372549 0.1647059 0.2313725 0.2666667
0.2078431 0.1921569 0.1568627 0.1411765 0.1372549 0.1686275 0.2352941 0.2745098
0.2000000 0.1843137 0.1490196 0.1333333 0.1372549 0.1725490 0.2392157 0.2745098
```

[Von Ahn et al. *CHI*, 2004]

Games with Purpose

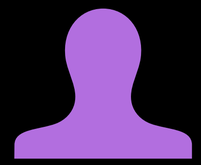


New York

0.1921569 0.1803922 0.1529412 0.1411765 0.1529412 0.1686275 0.1882353 0.2039216
0.2078431 0.1921569 0.1647059 0.1568627 0.1686275 0.1843137 0.2039216 0.2196078
0.2117647 0.1960784 0.1686275 0.1607843 0.1725490 0.1882353 0.2156863 0.2313725
0.2039216 0.1921569 0.1647059 0.1568627 0.1686275 0.1882353 0.2156863 0.2392157
0.1921569 0.1803922 0.1529412 0.1450980 0.1568627 0.1803922 0.2117647 0.2392157
0.1882353 0.1725490 0.1450980 0.1372549 0.1529412 0.1764706 0.2117647 0.2392157
0.1921569 0.1764706 0.1490196 0.1411765 0.1529412 0.1725490 0.2078431 0.2313725
0.2039216 0.1882353 0.1568627 0.1450980 0.1568627 0.1764706 0.2039216 0.2274510
0.2156863 0.2000000 0.1686275 0.1568627 0.1686275 0.1843137 0.2078431 0.2235294
0.2196078 0.2039216 0.1725490 0.1607843 0.1647059 0.1803922 0.2156863 0.2352941
0.2235294 0.2078431 0.1725490 0.1529412 0.1450980 0.1686275 0.2235294 0.2549020
0.2196078 0.2039216 0.1686275 0.1490196 0.1372549 0.1647059 0.2313725 0.2666667
0.2078431 0.1921569 0.1568627 0.1411765 0.1372549 0.1686275 0.2352941 0.2745098
0.2000000 0.1843137 0.1490196 0.1333333 0.1372549 0.1725490 0.2392157 0.2745098

[Von Ahn et al. *CHI*, 2004]

Games with Purpose



New York

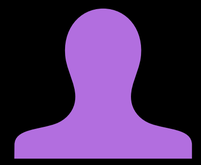
New York

0.1921569 0.1803922 0.1529412 0.1411765 0.1529412 0.1686275 0.1882353 0.2039216
0.2078431 0.1921569 0.1647059 0.1568627 0.1686275 0.1843137 0.2039216 0.2196078
0.2117647 0.1960784 0.1686275 0.1607843 0.1725490 0.1882353 0.2156863 0.2313725
0.2039216 0.1921569 0.1647059 0.1568627 0.1686275 0.1882353 0.2156863 0.2392157
0.1921569 0.1803922 0.1529412 0.1450980 0.1568627 0.1803922 0.2117647 0.2392157
0.1882353 0.1725490 0.1450980 0.1372549 0.1529412 0.1764706 0.2117647 0.2392157
0.1921569 0.1764706 0.1490196 0.1411765 0.1529412 0.1725490 0.2078431 0.2313725
0.2039216 0.1882353 0.1568627 0.1450980 0.1568627 0.1764706 0.2039216 0.2274510
0.2156863 0.2000000 0.1686275 0.1568627 0.1686275 0.1843137 0.2078431 0.2235294
0.2196078 0.2039216 0.1725490 0.1607843 0.1647059 0.1803922 0.2156863 0.2352941
0.2235294 0.2078431 0.1725490 0.1529412 0.1450980 0.1686275 0.2235294 0.2549020
0.2196078 0.2039216 0.1686275 0.1490196 0.1372549 0.1647059 0.2313725 0.2666667
0.2078431 0.1921569 0.1568627 0.1411765 0.1372549 0.1686275 0.2352941 0.2745098
0.2000000 0.1843137 0.1490196 0.1333333 0.1372549 0.1725490 0.2392157 0.2745098

New York

[Von Ahn et al. *CHI*, 2004]

Games with Purpose



New York

0.1921569 0.1803922 0.1529412 0.1411765 0.1529412 0.1686275 0.1882353 0.2039216
0.2078431 0.1921569 0.1647059 0.1568627 0.1686275 0.1843137 0.2039216 0.2196078
0.2117647 0.1960784 0.1686275 0.1607843 0.1725490 0.1882353 0.2156863 0.2313725
0.2039216 0.1921569 0.1647059 0.1568627 0.1686275 0.1882353 0.2156863 0.2392157
0.1921569 0.1803922 0.1529412 0.1450980 0.1568627 0.1803922 0.2117647 0.2392157
0.1882353 0.1725490 0.1450980 0.1372549 0.1529412 0.1764706 0.2117647 0.2392157
0.1921569 0.1764706 0.1490196 0.1411765 0.1529412 0.1725490 0.2078431 0.2313725
0.2039216 0.1882353 0.1568627 0.1450980 0.1568627 0.1764706 0.2039216 0.2274510
0.2156863 0.2000000 0.1686275 0.1568627 0.1686275 0.1843137 0.2078431 0.2235294
0.2196078 0.2039216 0.1725490 0.1607843 0.1647059 0.1803922 0.2156863 0.2352941
0.2235294 0.2078431 0.1725490 0.1529412 0.1450980 0.1686275 0.2235294 0.2549020
0.2196078 0.2039216 0.1686275 0.1490196 0.1372549 0.1647059 0.2313725 0.2666667
0.2078431 0.1921569 0.1568627 0.1411765 0.1372549 0.1686275 0.2352941 0.2745098
0.2000000 0.1843137 0.1490196 0.1333333 0.1372549 0.1725490 0.2392157 0.2745098

New York

[Von Ahn et al. *CHI*, 2004]

Micro Tasks

low cognitive overload

img-source: wired

Metacognition

∞

Creativity

Social Learning

Motivation

Intuitions

Incentives

Collaboration

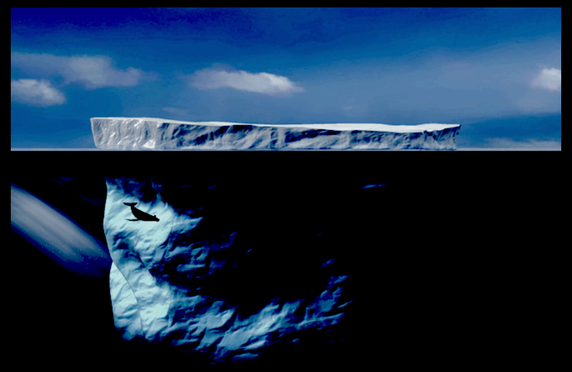
Empathy

img-source: wired

Macro Tasks

high cognitive overload

Thinking at computational and cognitive scale



delving deeper

What can we accomplish if we harness crowd and machine intelligence?

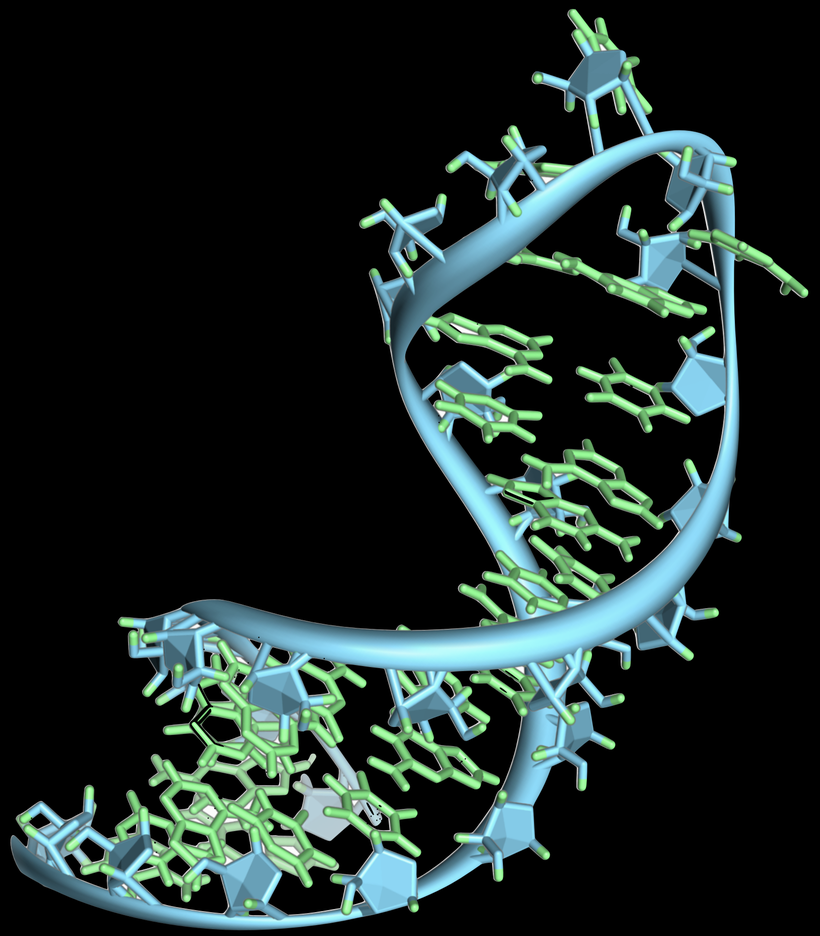
The Internet as our laboratory and
a crowd as collaborator

Research and Solve

the biggest unanswered questions

Exploring the life at the cellular level

Understanding RNAs



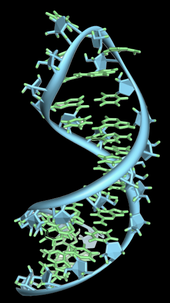
Why RNA?

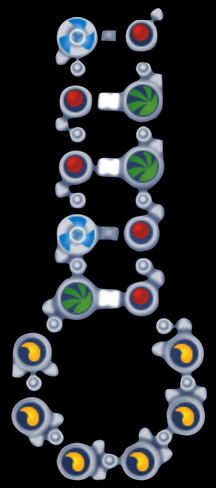
Dark matter of Biology

Key to understand life at the cellular level

Manipulate RNAs to help cure Parkinson's disease

Discover the sequence





Hairpin RNA
Target Shape

RNAs are string of
Nucleotide Base
∈
{ A,C,U,G }

G G A G
G G U U
G A U A
C G A U
C G C U
U G A A
U . . .

Nucleotide Sequence

img-source: Samira Bouaou/Epoch Times

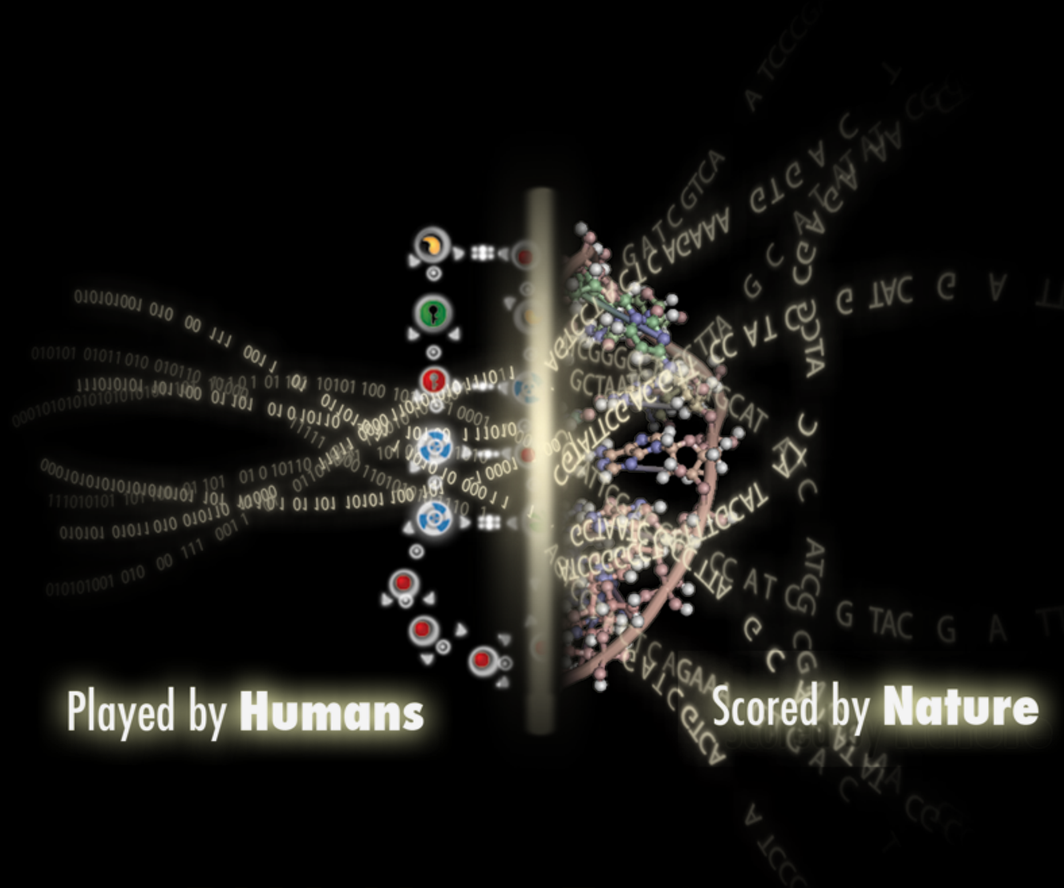
The power of games

Every year millions of human hours are spent solving puzzles in games

img-source: Samira Bouaou/Epoch Times,, Angrybirds.com, Apple App Store

Games for democratizing science

eterna
Make Molecules Advance Science



RNA Nano Engineering

Over 100,000 users from more than 90 Countries

[Lee *et al. PNAS*, 2014]

Pushing the Boundaries of Quantum Physics

Quantum Control Optima

Search for high dimensional complex challenges

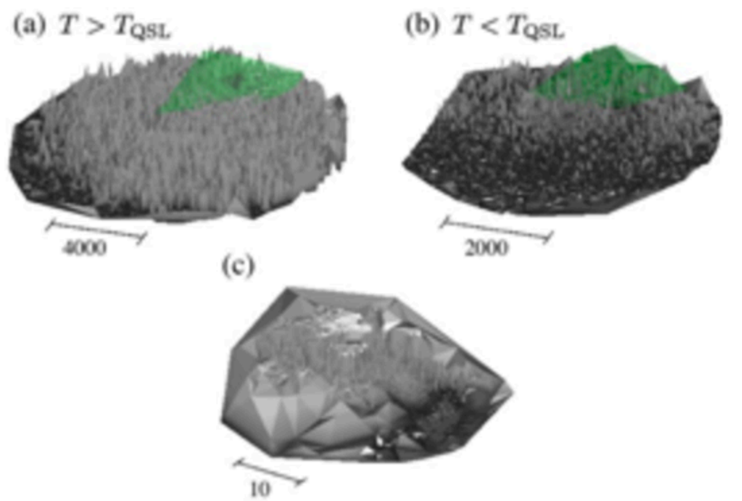
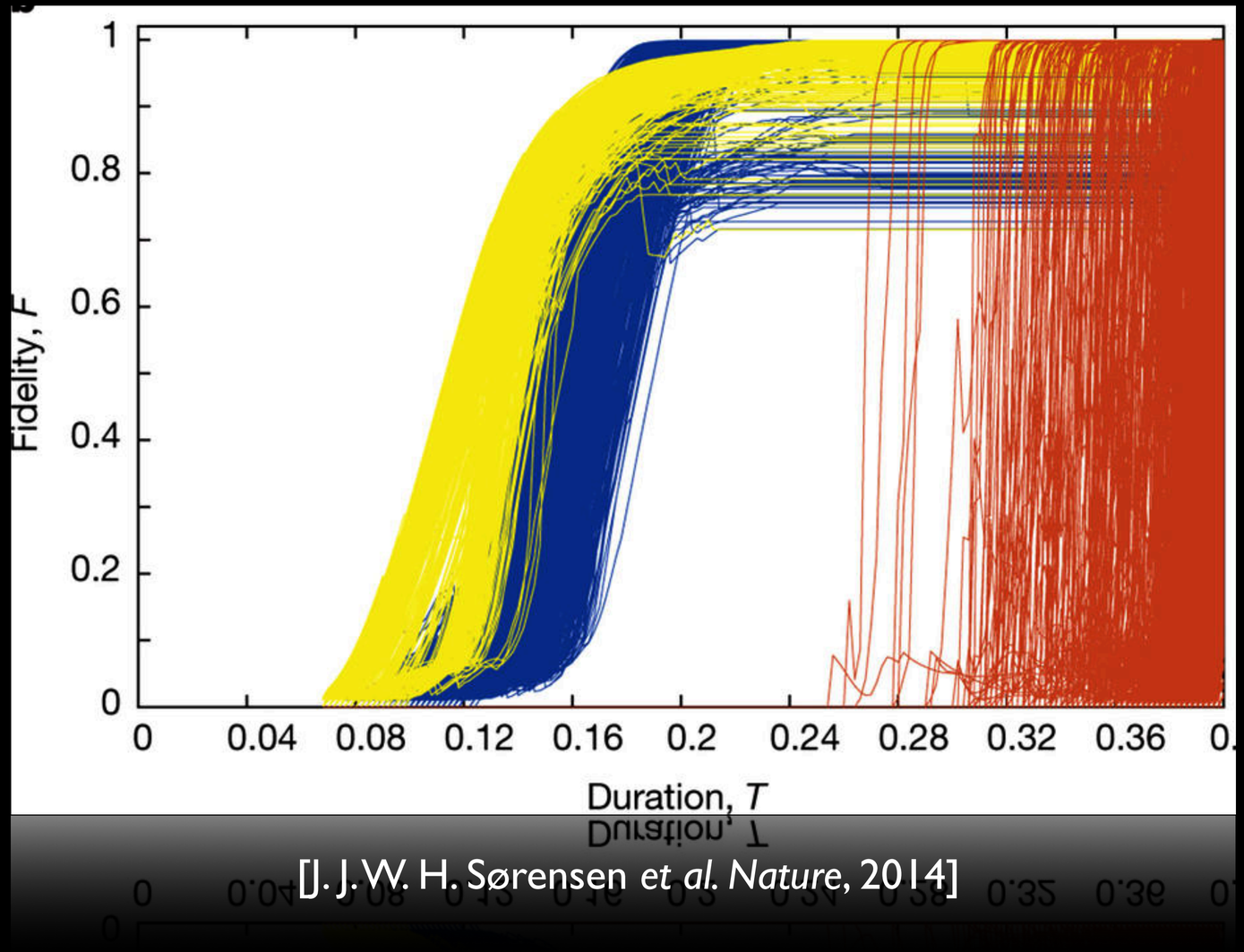


Figure 4 | Optimization landscapes. Panels **a** and **b** show the 2D rendering of the high-dimensional optimization landscape for process durations $T = 0.40$ and $T = 0.17$, respectively. Green areas mark the space probed by CHOP solutions. Panel **b** is the low-dimensional HILO landscape.


QUANTUM MOVES
Play!
Login
Quit
For educational use only

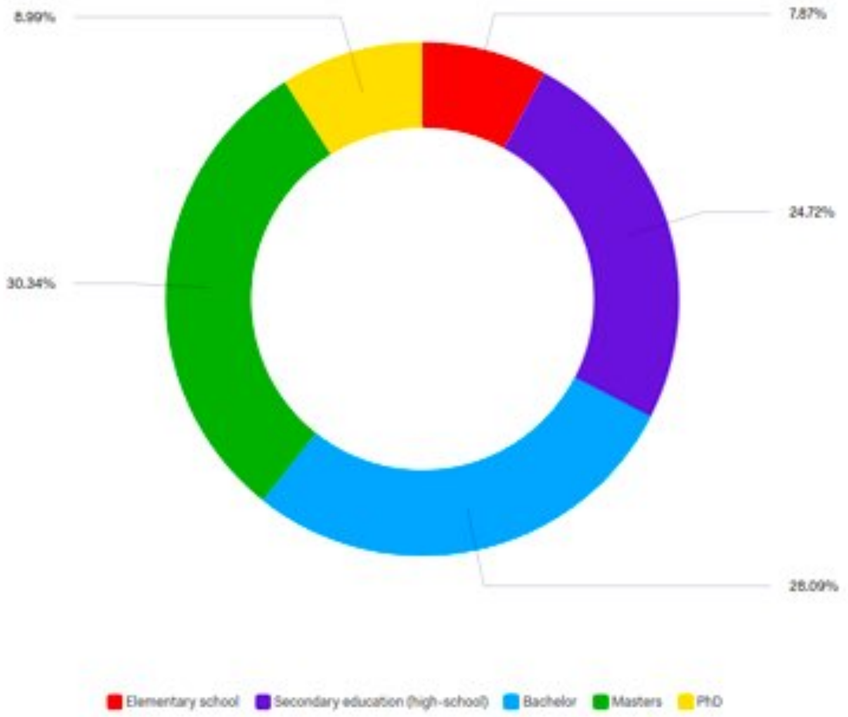
Science@Home



[J. J. W. H. Sørensen *et al. Nature*, 2014]

Players

Highest education level graduated



Number of years (if any) of post-secondary education in physics

#	Answer	%	Count
	0 (I have only studied physics up to/in high-school)	47.54%	29
	one to three	24.59%	15
	three to five	16.39%	10
	more than 5	11.48%	7
	Total	1	61

- United States
- France
- Spain
- Germany
- Romania
- Austria
- United Kingdom
- Russian Federation
- Switzerland
- Italy
- Iceland
- Belgium
- Netherlands
- Sweden
- Australia
- Finland
- Denmark

Democratizing research to increased diversity in the scientific workforce

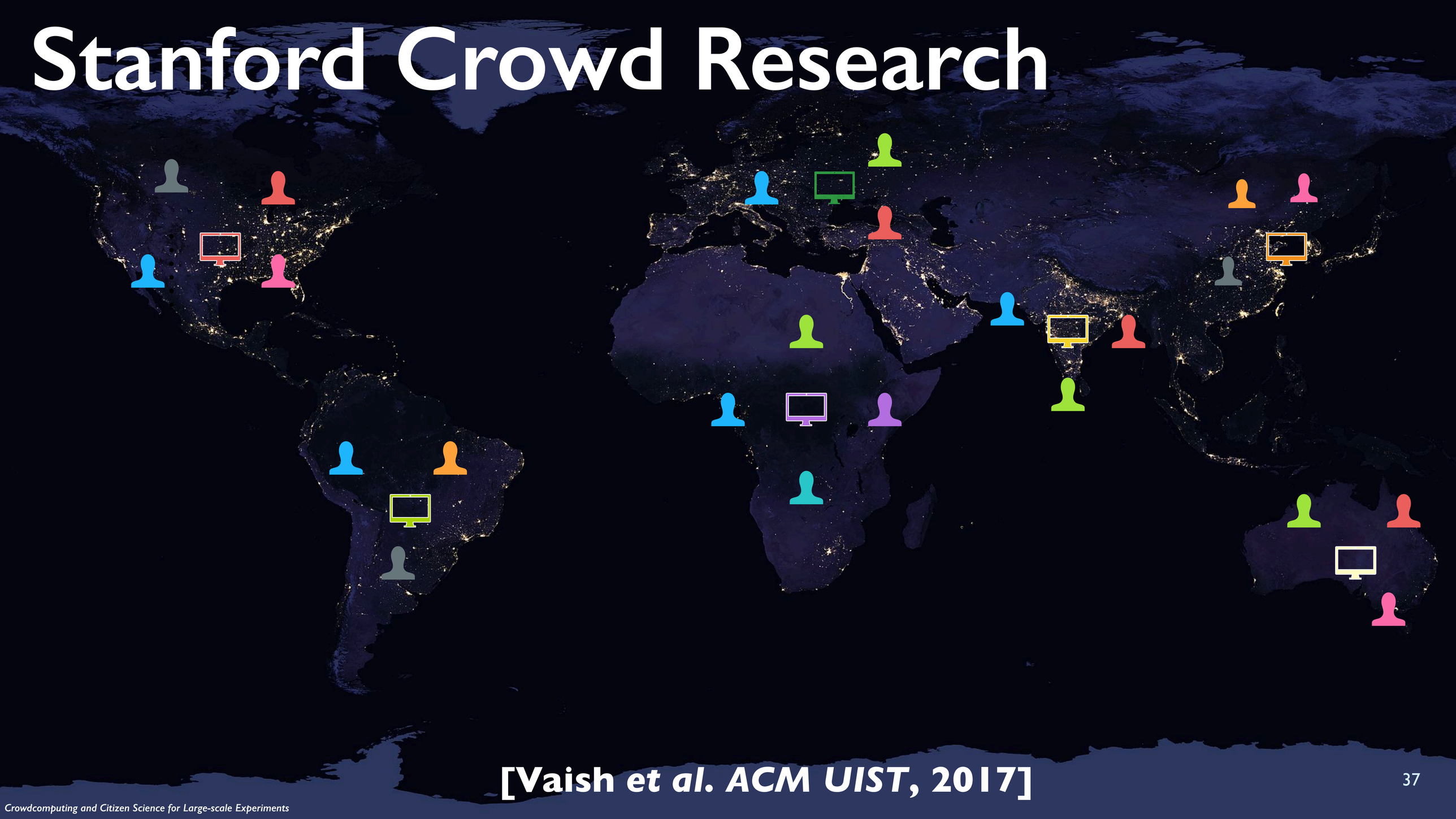
Scientific research remains the domain of the privileged few—
to those blessed with the socioeconomic opportunities

The man who knew ∞

Srinivasa Ramanujan
[1887-1920]

img-source: http://blog.stephenwolfram.com/, Trinity College library, plus.maths.org

Stanford Crowd Research



[Vaish *et al. ACM UIST*, 2017]

Over 1500 Crowd Researchers

DAEMO

[Gaikwad et al. ACM UIST, 2015 2016]

Forum Sign In Get Started

DAEMO

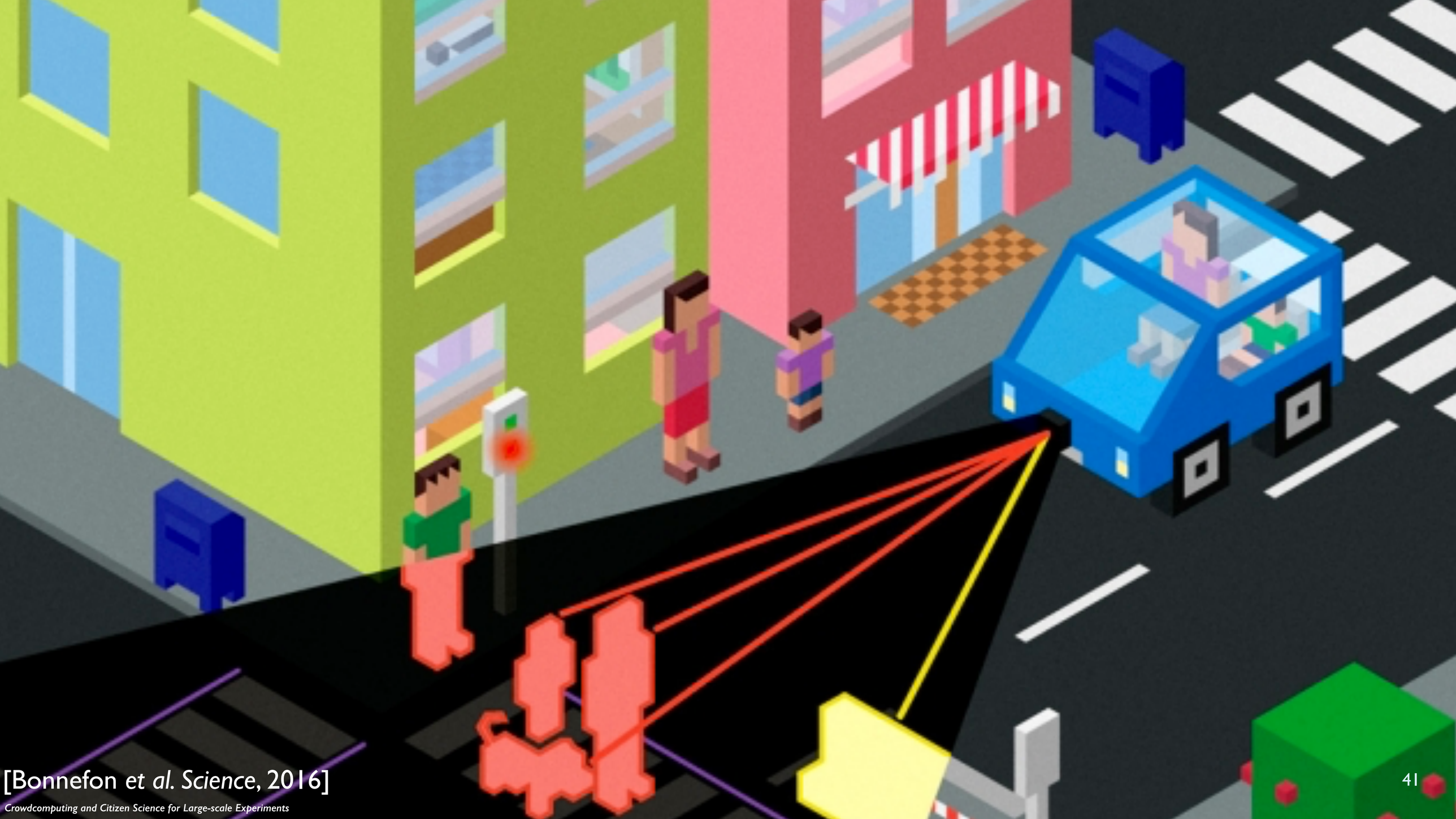
Easier and more equitable crowdsourcing

Stanford University

Your results get better automatically.

Workers who you like will get first access to your tasks. As you give feedback, your favorite workers will do more and more of your work.

Researching AI ethics and algorithmic biases at scale

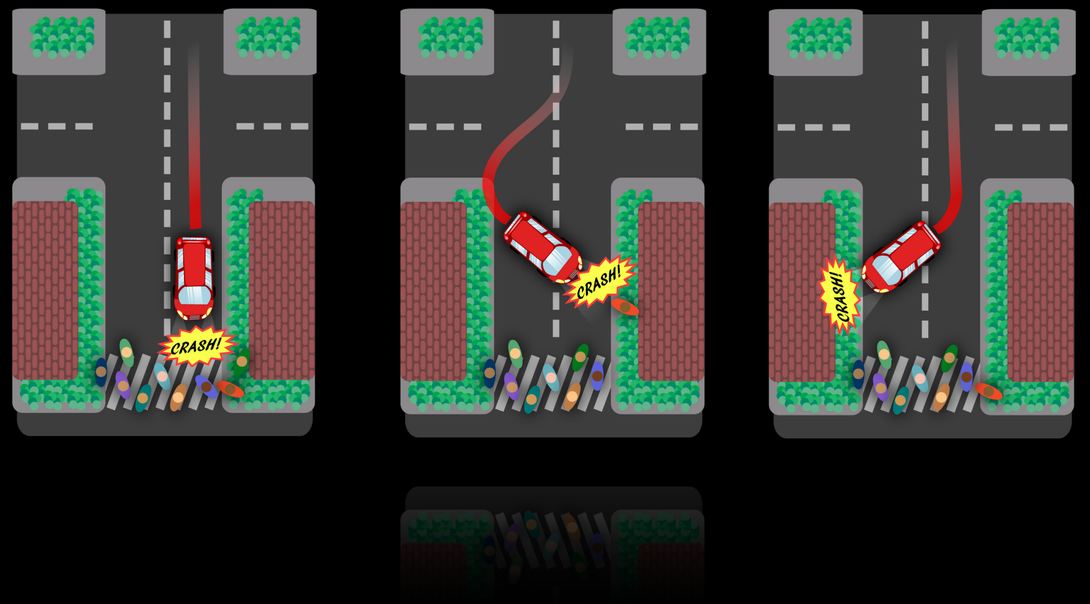


[Bonnefon *et al. Science*, 2016]

Moral Machine

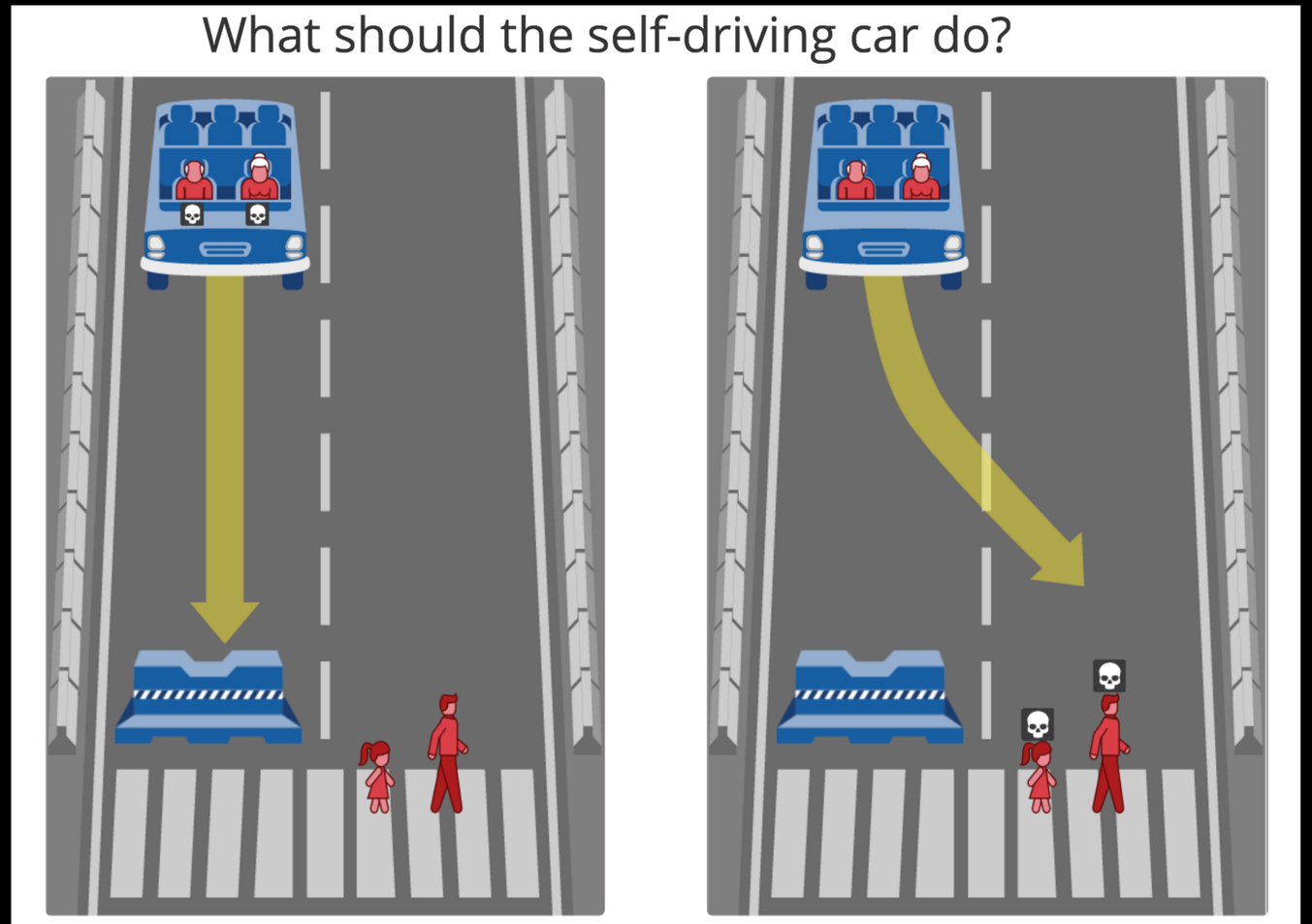
- Human subject data needed on vast number of combinations of factors

- Subscription crowdsourced experiment services would be prohibitive in cost and/or not as flexible



Moral Machine

- Minimalist and direct
- Abstract intuitive interface
- Cross-platform
- Expressing creativity
- Holding mirror to subject
- Sharing results and scenarios
- Discussion and feedback



[www.moralmachine.mit.edu/]

CASE

Let’s choose a problem

II
Crowd Mechanics

Crowd Mechanics

DYNAMICS & LOGISTICS

EXPERIMENT DESIGN & DATA SCIENCE

PLATFORM & INFRASTRUCTURE DESIGN

Interactive Case Discussions about the problem you want to solve

Dynamics & Logistics

When to use the crowd?

- Reasonably well-defined problem
- Size, diversity, impact, and time
- Can give the crowd something of value
- Scale and reach over selectivity and acquaintance

Making the Crowd

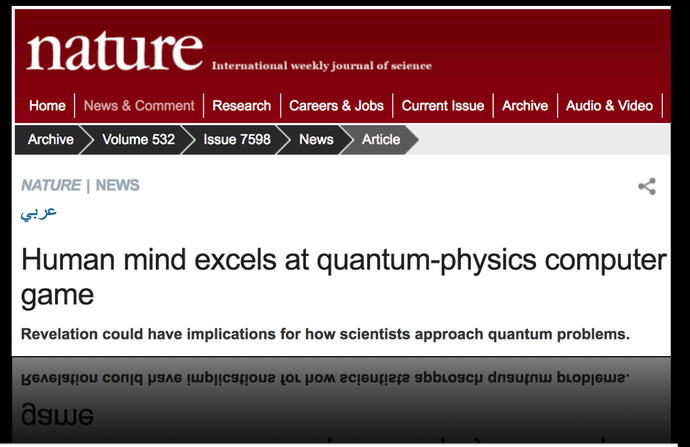
Recruitment and Retention

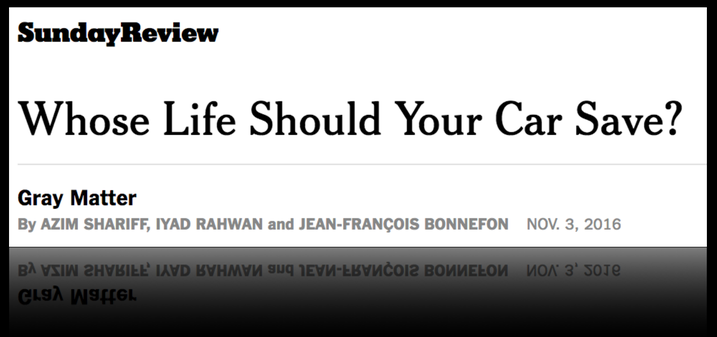
Recruitment

- Social Network
  - Reputation
    - » [Stanford Crowd Research]
  - Volunteer Science / Lab in the wild: Sharing Test-results on Facebook
    - » [Science at Home, EteRNA]
  - “Let’s Play” and commentary videos by YouTube stars
    - » [Moral Machine]

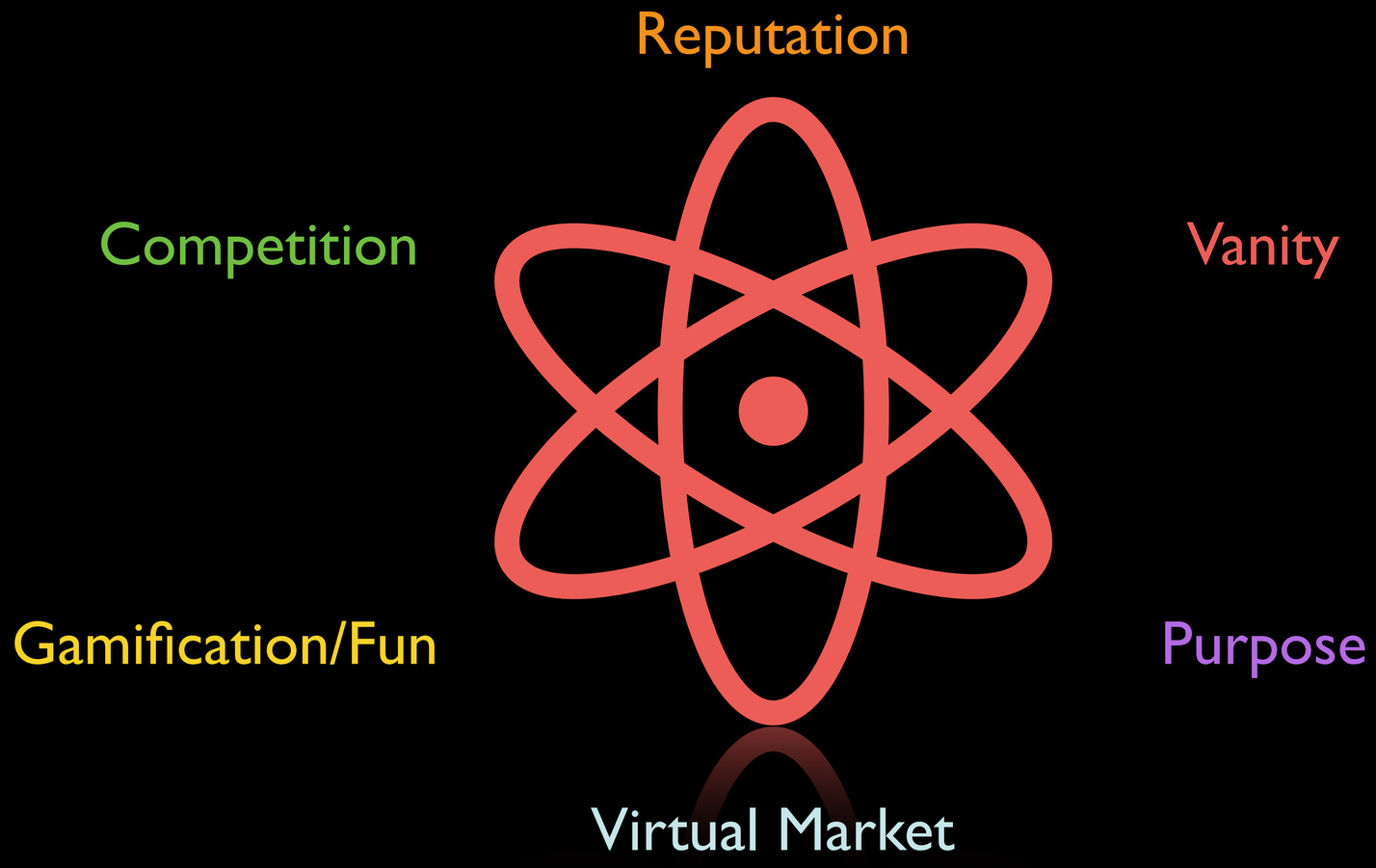
Recruitment

- Press
  - A popular journal (Nature) + the associated press attention
    - » [Science at Home]
  - A popular news paper coverage (e.g. New York Times)
    - » [EteRNA, Moral Machine]
- Challenge
  - Build a quantum computer
  - Help invent Medicine
  - Do ethical thing
  - Do research and become a scientist





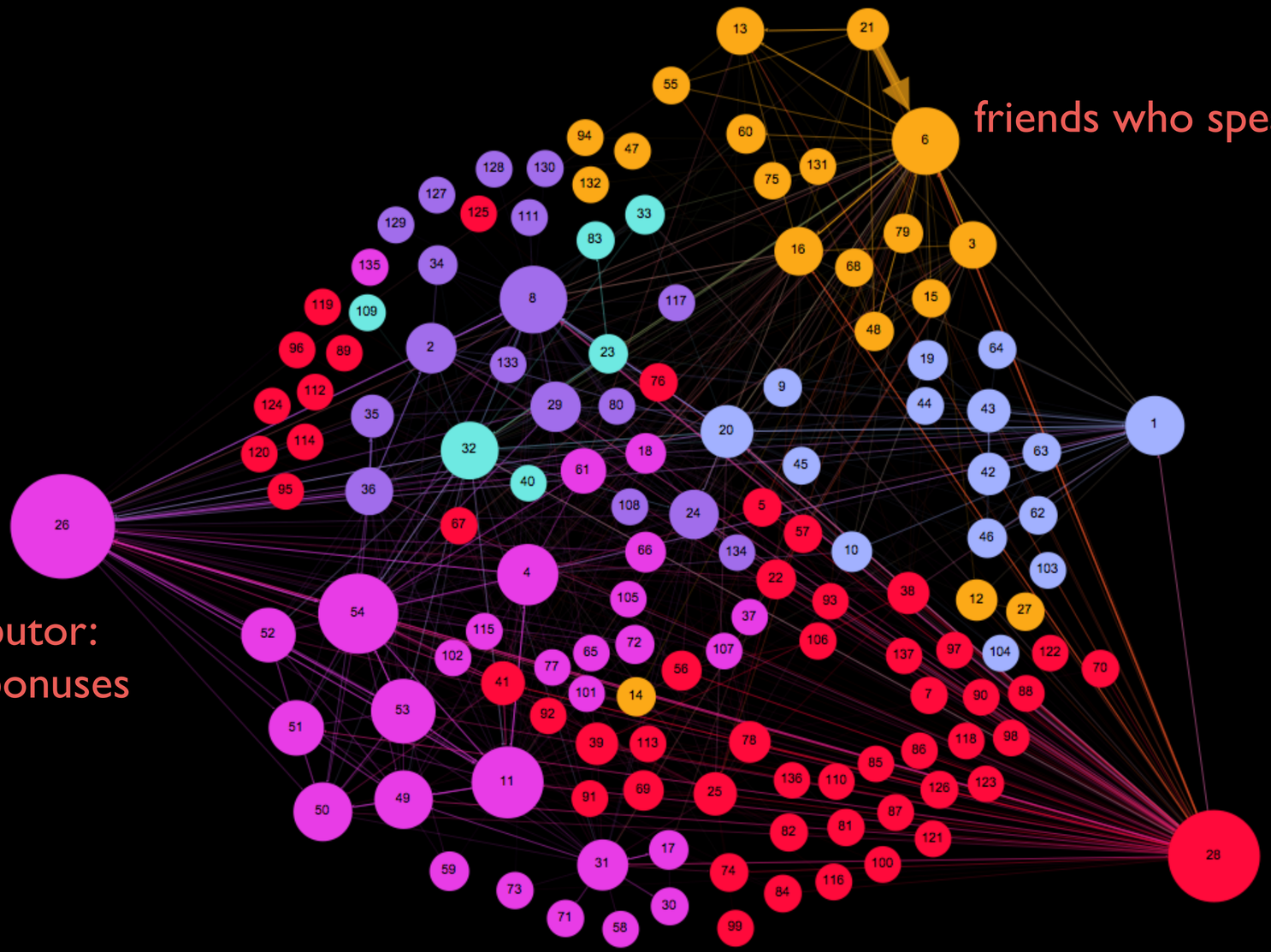
Retention & Incentive Engineering



Motivations: Intrinsic, Extrinsic

Bonuses: Incentives and motivation

Periodic bonuses



friends who speak similar language

lead contributor:
rewarding bonuses

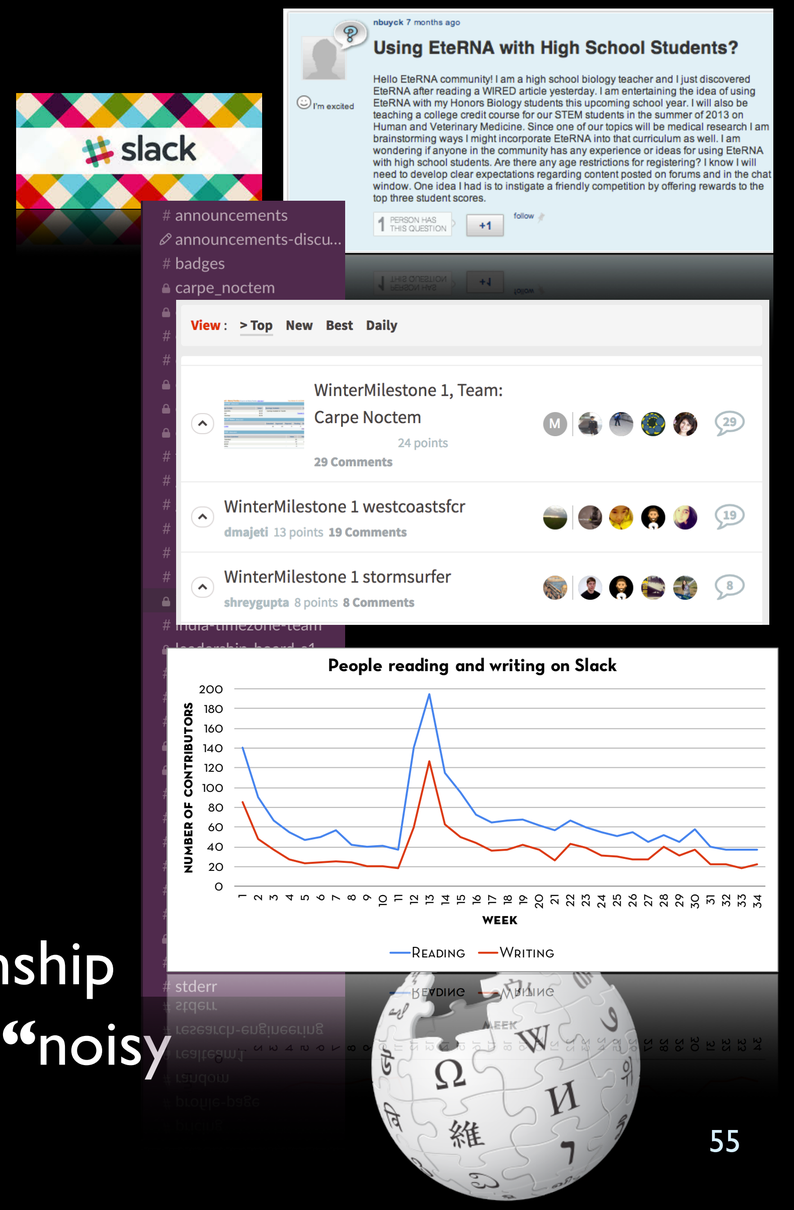
admin:
rewarding bonuses

Organizing the Crowd

Community, Task Distribution, Recognition

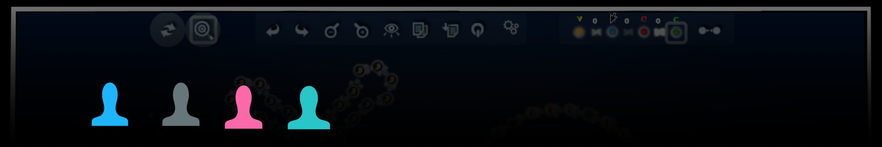
Community

- Task Distributions
  - Various paradigm (e.g., Parallel vs Sequential)
  - Milestones and feedback
- Engagement Channels
  - Asynchronous Communication: Forums, Emails, Wikipedia
  - Synchronous Communication: Slack, Hangouts, Chats
- Protocols
  - Inclusive and constructive environment
  - Empathy —Volunteers' (workers') and requesters' relationship
  - Remember that the crowd is comprised of **people, *not*** "noisy error-prone computers"



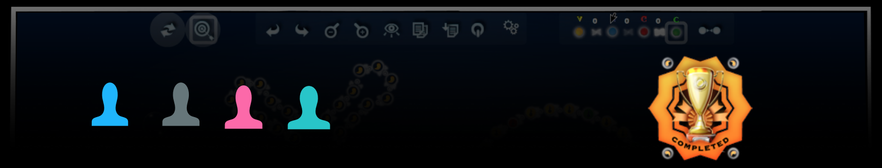
eterna
Make Molecules Advance Science

Process



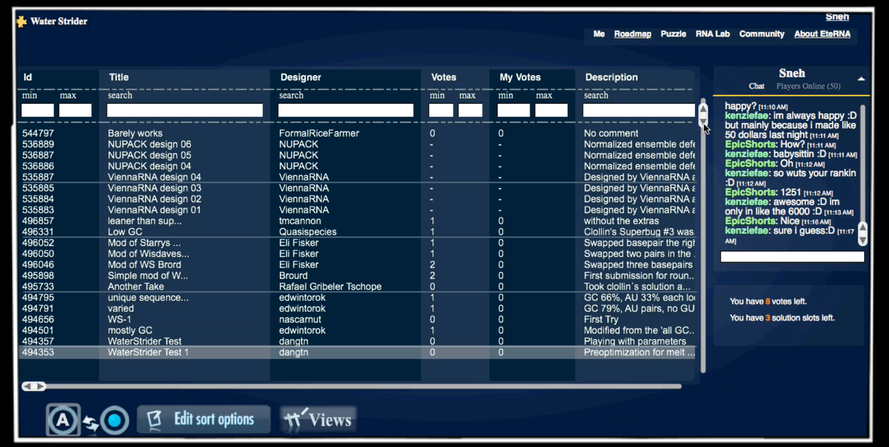
eterna
Make Molecules Advance Science

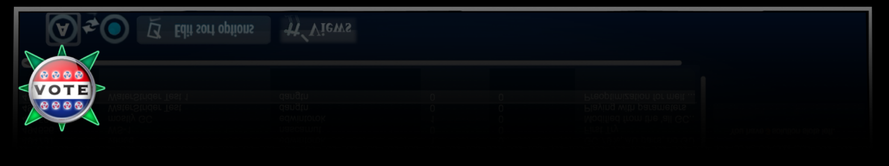
Process



eterna
Make Molecules Advance Science

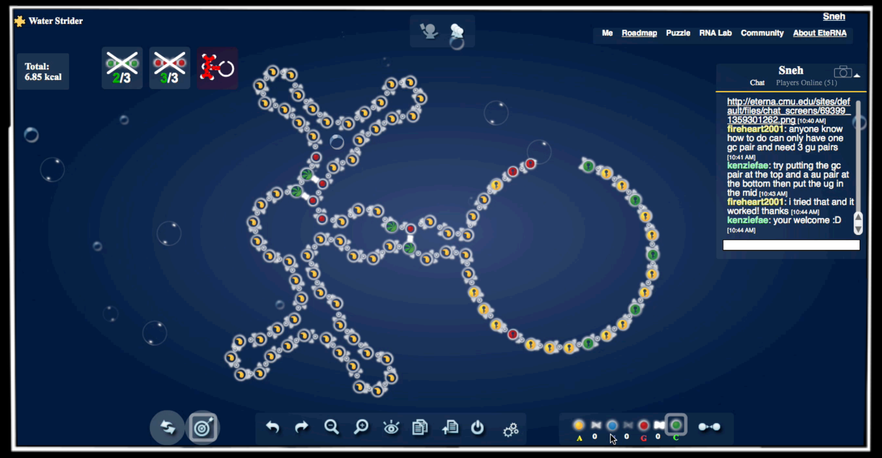
Process

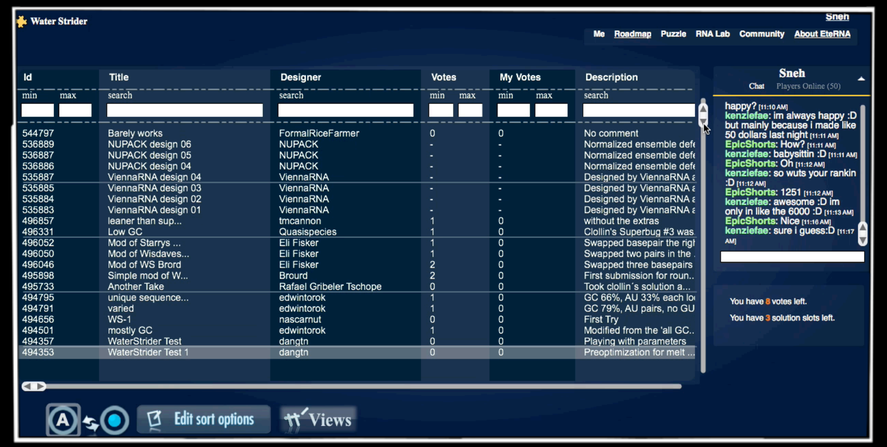




eterna
Make Molecules Advance Science

Process

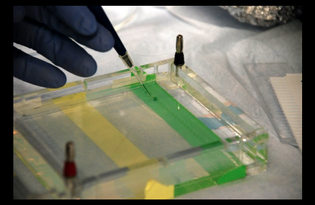


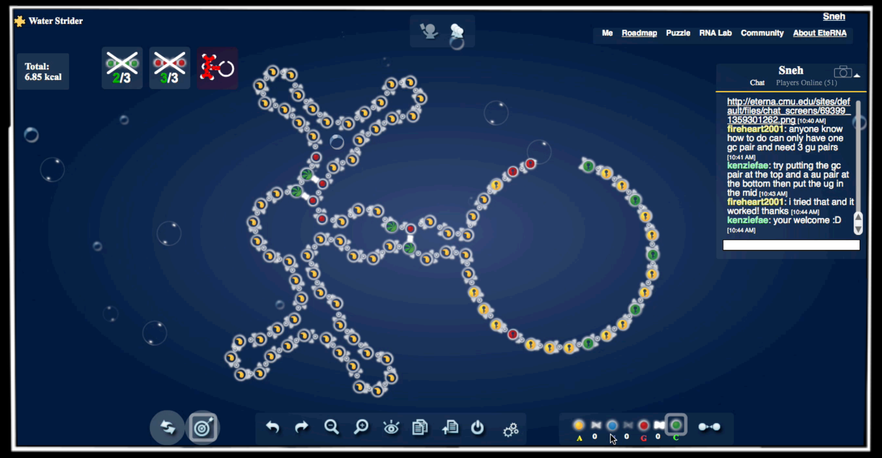


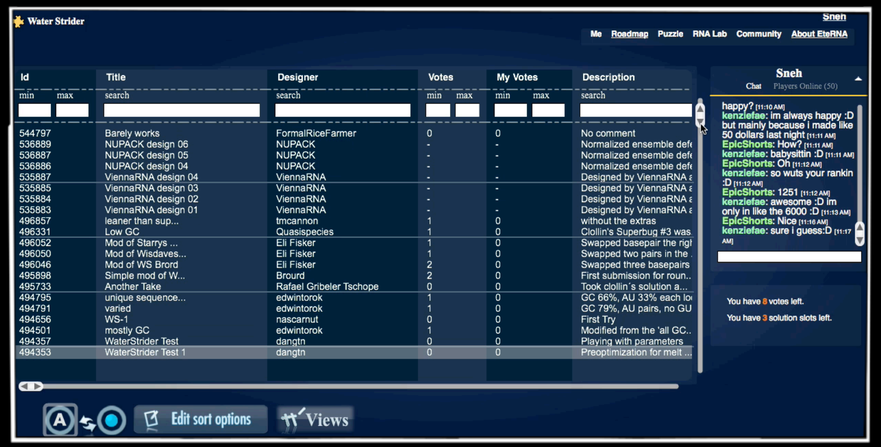
eterna
Make Molecules Advance Science

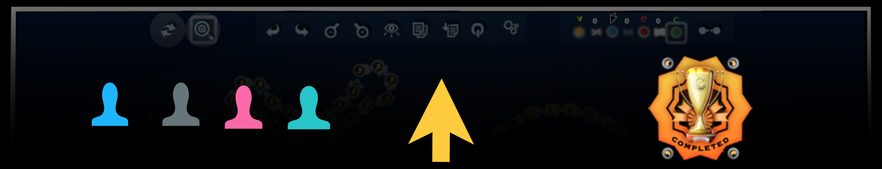
Process

DScore
x/100





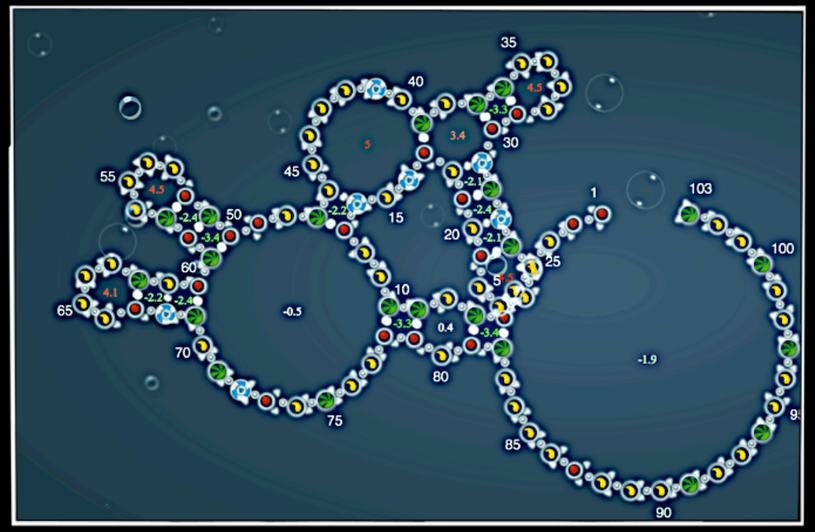




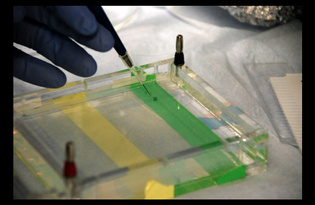
eterna

Make Molecules Advance Science

Process



DScore x/100

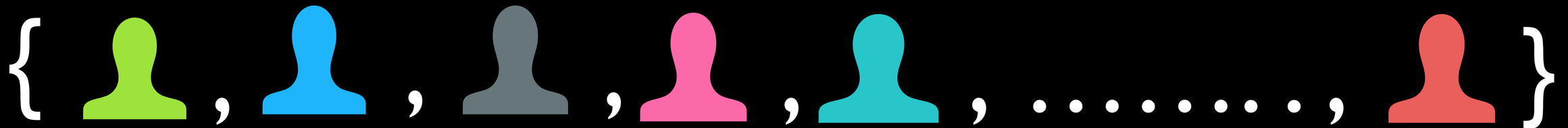


Recognition

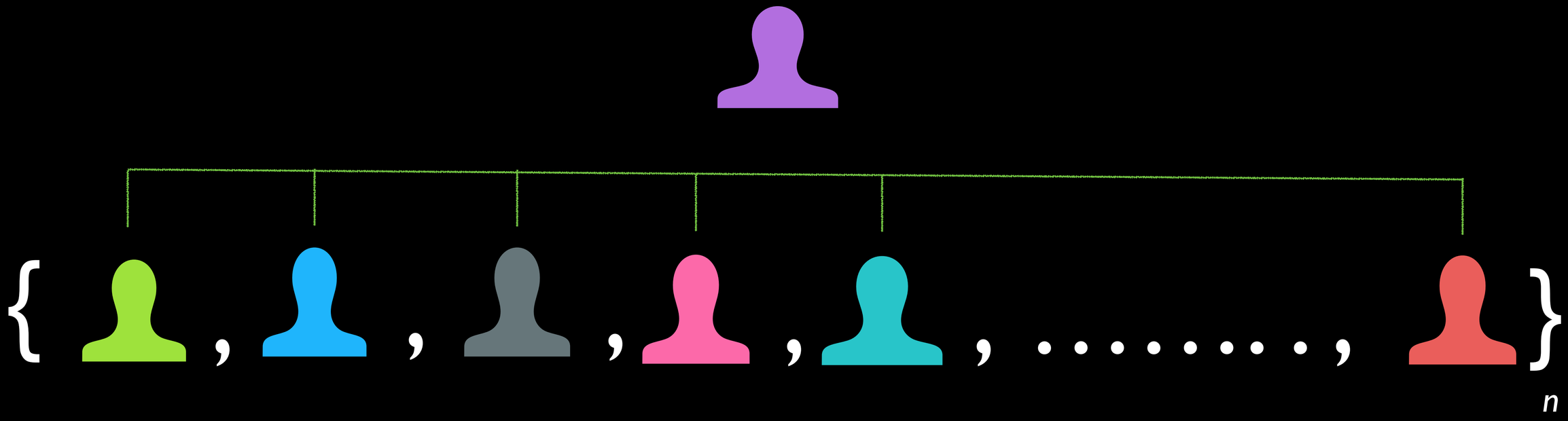
- Admins cannot have a full picture of who is doing what
- Most of the voting rules are confusing and can be gamed
- Voluntary projects—

  how to spot HUBS in a large network and empower them

Nodes in Credit Distribution



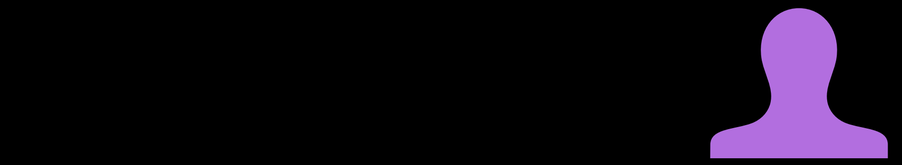
Links in Credit Distribution

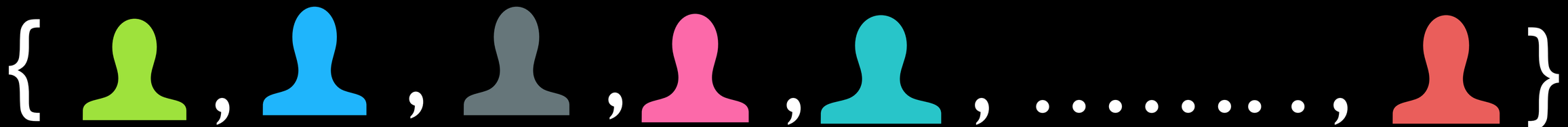


Scores ***s***

{ 10 , 20 , 00 , 40 , 00 , ... , 30 }

Network of Credit Distribution



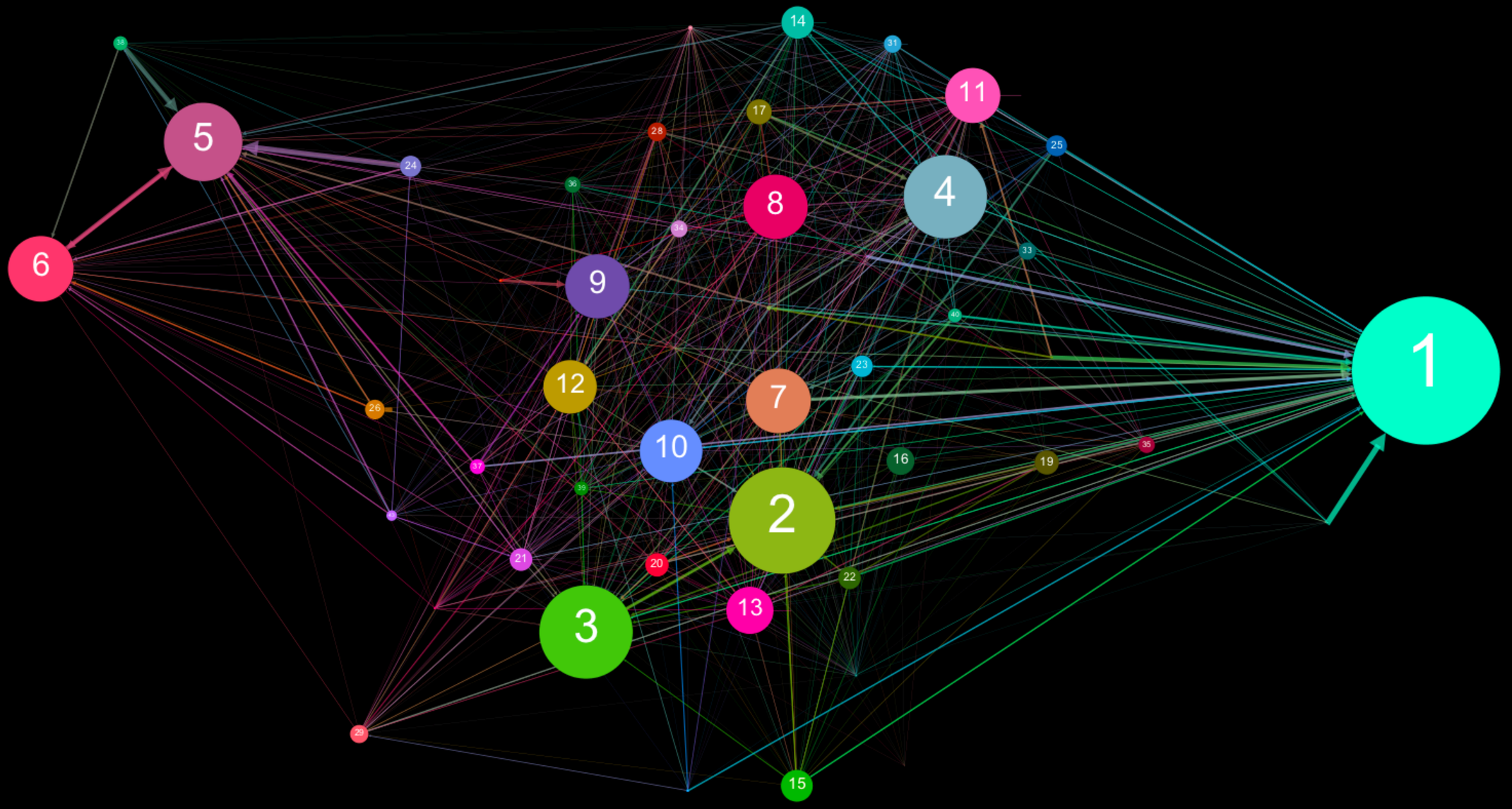


{ 10 , 20 , 00 , 40 , 00 , ……., 30 }

Scores s

$$\sum_{i=1}^{n} s_i = 100$$

Author ranks via page rank



CASE
The Problem and Crowd Dynamics

Recruitment Size, Diversity, Goal/Task

Recognition, Retention, Motivations

Community & Organization

Experiments and Data Science

Data collection

- How to account for Noise
  - Crowd participants are heterogeneous
    - » Long tail of contributions and commitment
    - » Cultural biases and language barriers
    - » Diverse range of expertise

- Data
  - One optimal or “best” solution [Quantum Moves]
  - Hypothesis to discover molecules [EteRNA]
  - Aggregation of data from many individuals [Moral Machine]
  - Scientific interactions [Crowd Research]

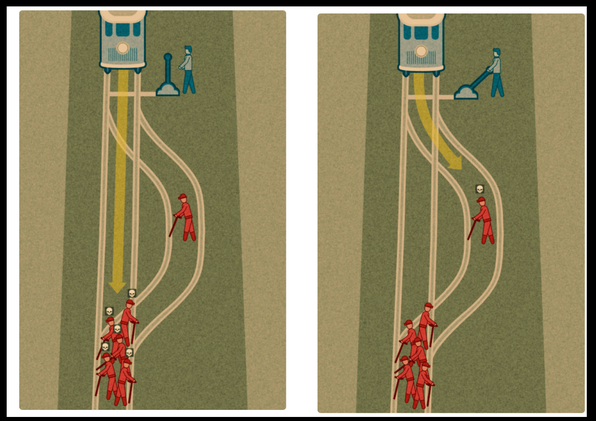
Data collection

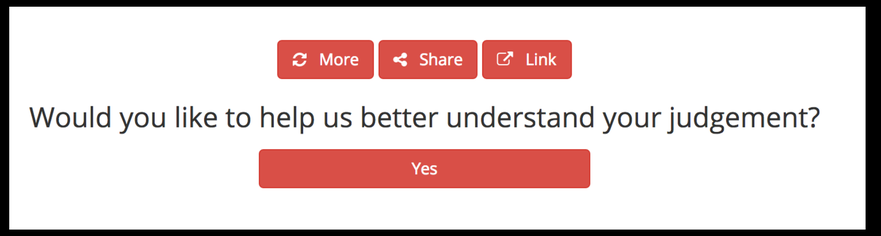
- Barrier to Entry
  - Registration/forms
  - Extra tutorials, milestones
  - Unfamiliar tools or technologies

- Quality Control
  - Organic— crowd filters out bad designs or submissions
  - Capture— Logs, IP, and device IDs (time to action, cursor movement, intermediate activity) to detect and either block or later filter out unthinking click-throughs, spam, or sabotage

Moral Machine - data expansion

- Internationalization (10 languages)
- Addition of post-quiz user survey
- Expansion of experiment scope with classic scenarios





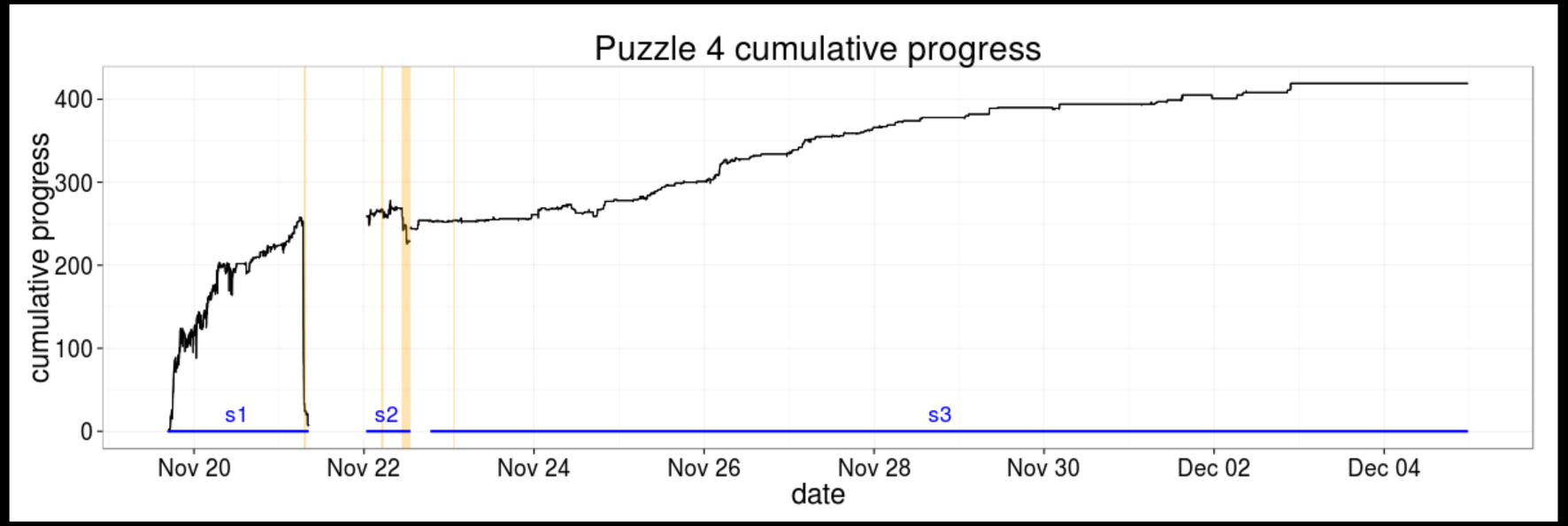
Moral Machine - data collection

3.2M Users

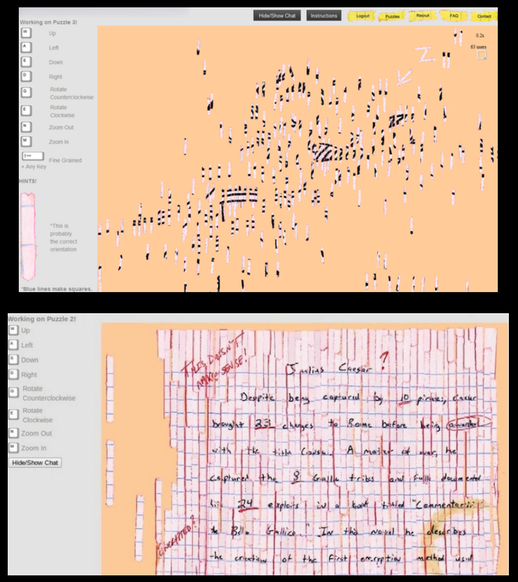
30M Decisions

350k Full Surveys

Sabotage activity logs



[Stefanovitch et al. EPJ Data Science, 2014]



DARPA Shredder Challenge

"*For my **first attack**, […]*
moving all the pieces into a single pile.

[…] However, it seems all you did was lock the pieces together and ban my IP address."

Sabotage activity logs

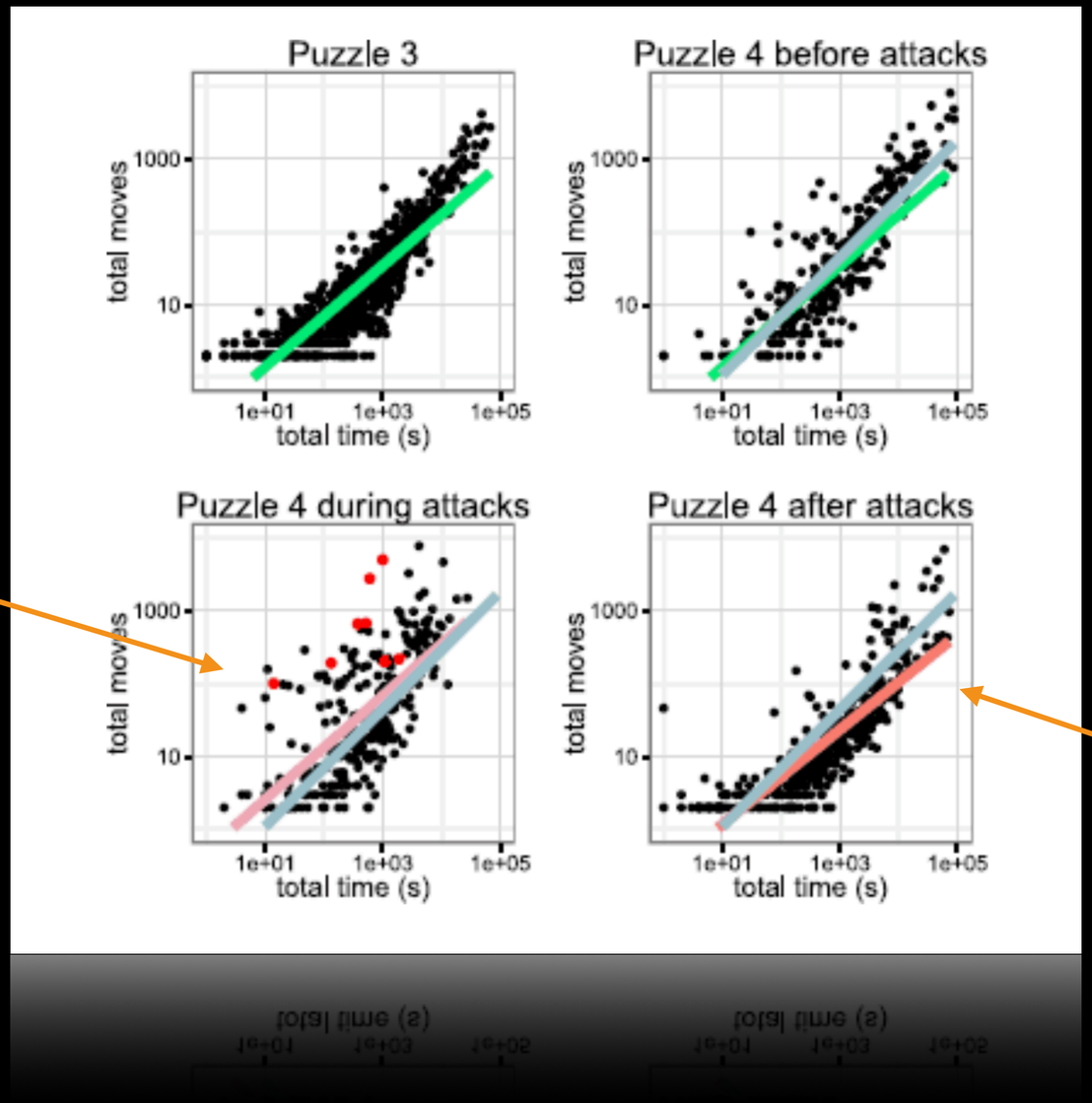
*"Which led me to my **second attack**, using a VPN and a neighbors wireless for some new IP addresses*
[…]
select all the pieces and place them on top of each other, but this got old soon."

Sabotage activity logs

*"Which led me to my **second attack**, using a VPN and a neighbors wireless for some new IP addresses*
[…]
select all the pieces and place them on top of each other, but this got old soon."

*"So I decided to get a bit **more sneaky***
[…]
*I selected a number of pieces, enough to make solving the puzzle difficult and **not so much that people would immediately notice**,*
[…]
moved them off the top of the virtual table"

Building vs policing



dot = 1 user, Red = attacker

Both attackers and responders move fast

Users slowed down after attack

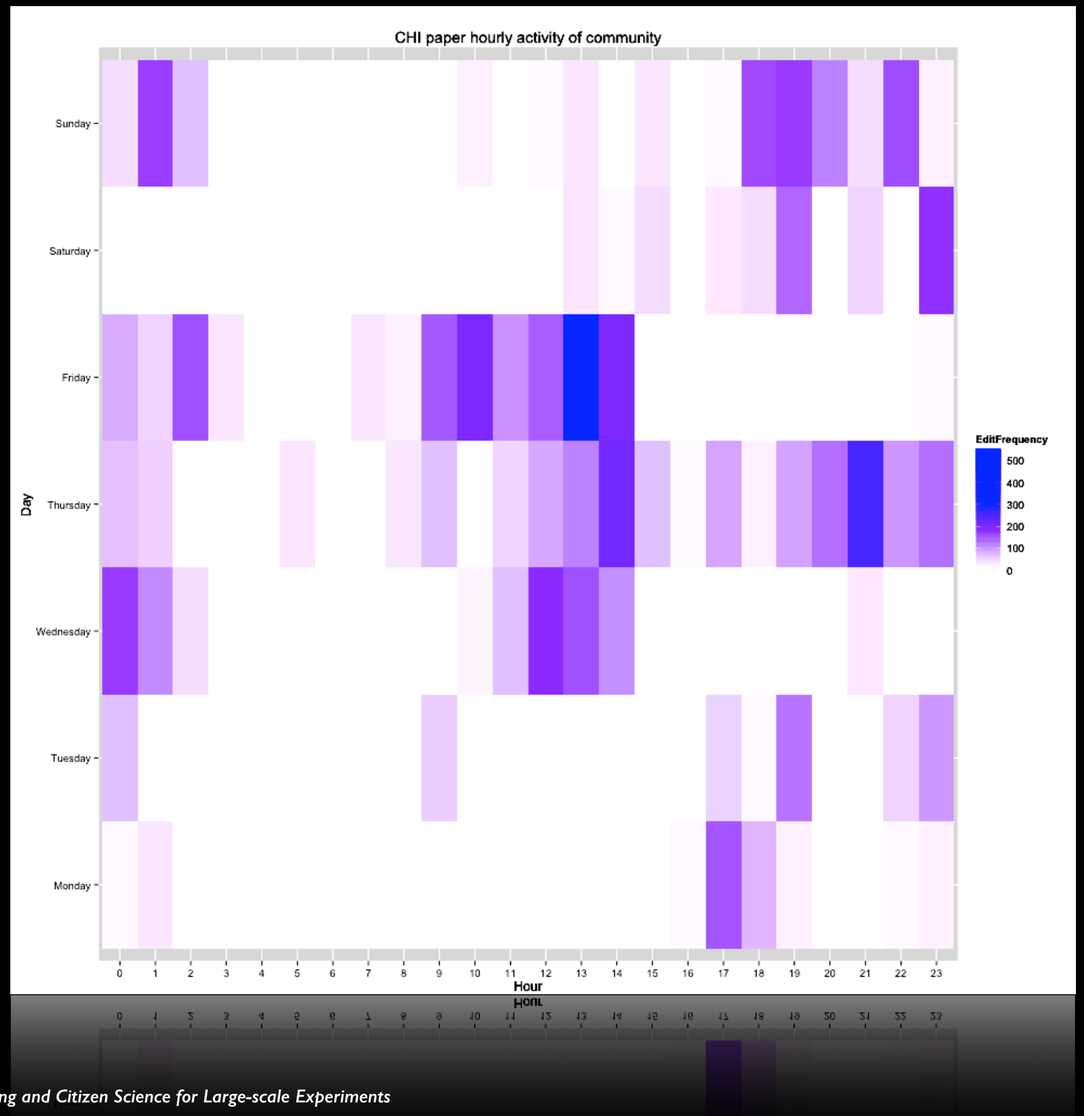
Analysis

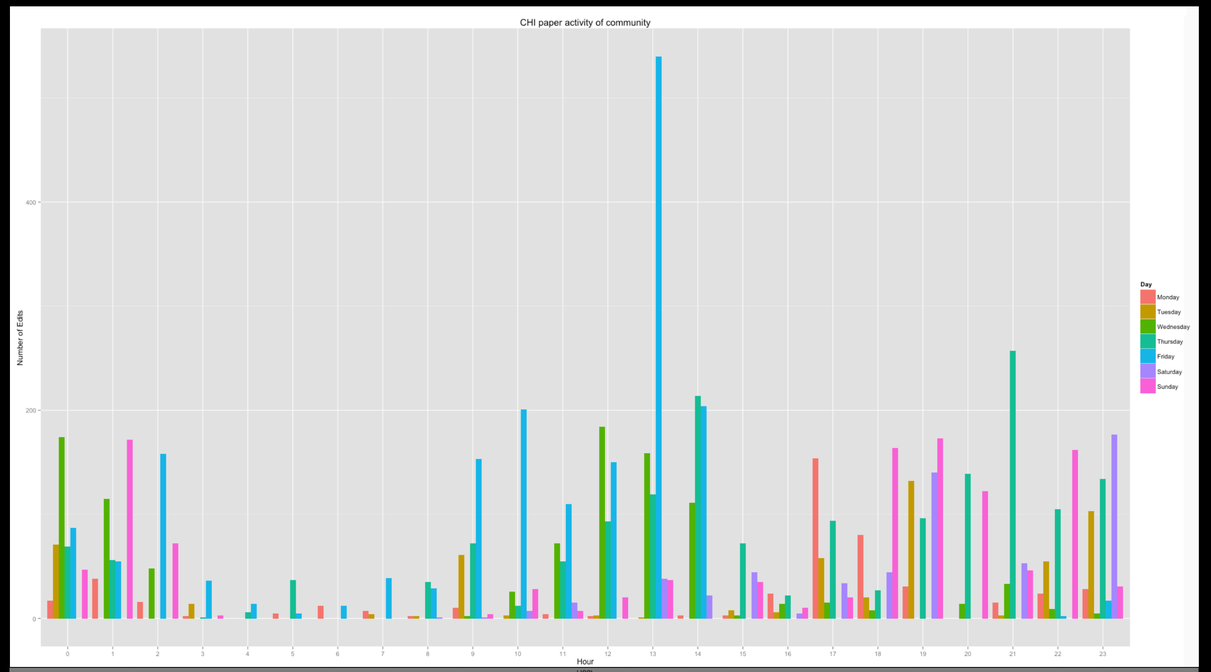
- Plan out data analysis beforehand
- Think about how you will collect your data. What data do you need
  - "Let's collect ALL the data (because we don't know what we will need later)" is a bad strategy
- How will you deal with attrition and dropout, partially observed data?
  - Is attrition correlated with experimental treatments?

Analysis

- What analyses will you run if you had your data?
  - Do a dry run before going and collecting all the data
- The best way to find out what people are thinking/doing: ask them!
  - *Qualitative complements Quantitative*

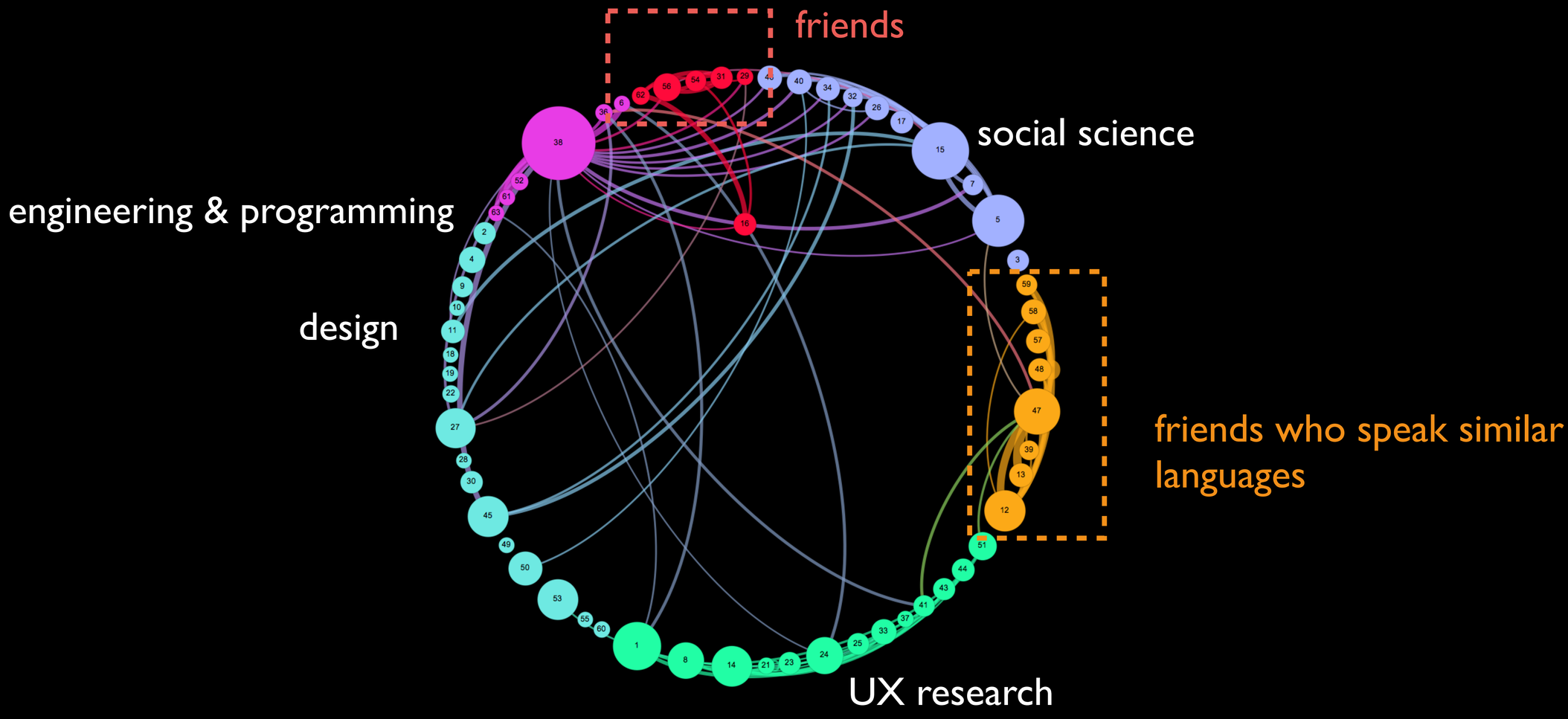
Activity Scheduling [time, day]



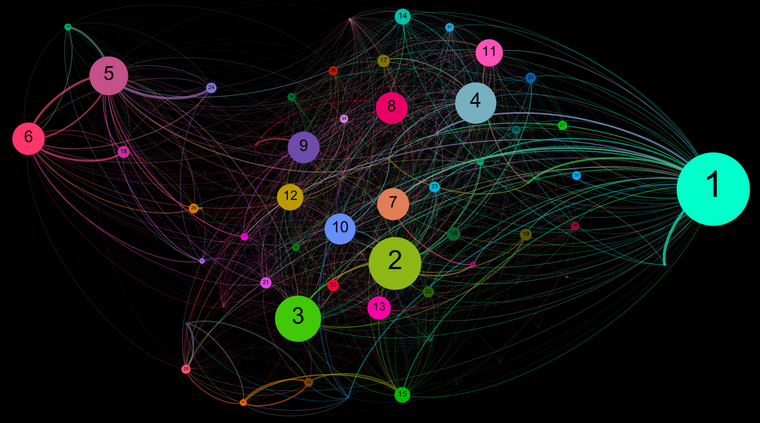


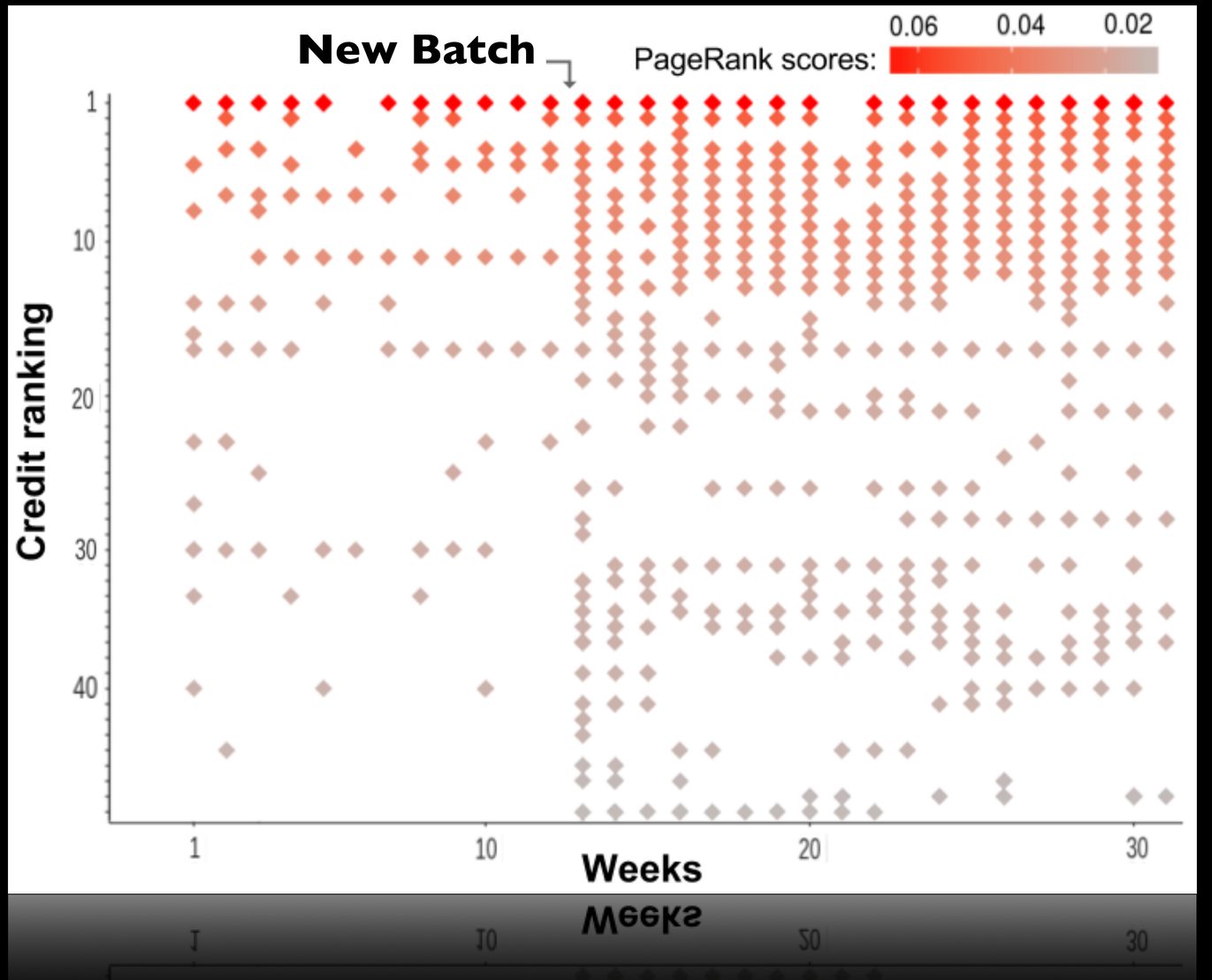
Crowd Research

Community Formation



Characteristics of leaders





CASE
The Problem and Data Science

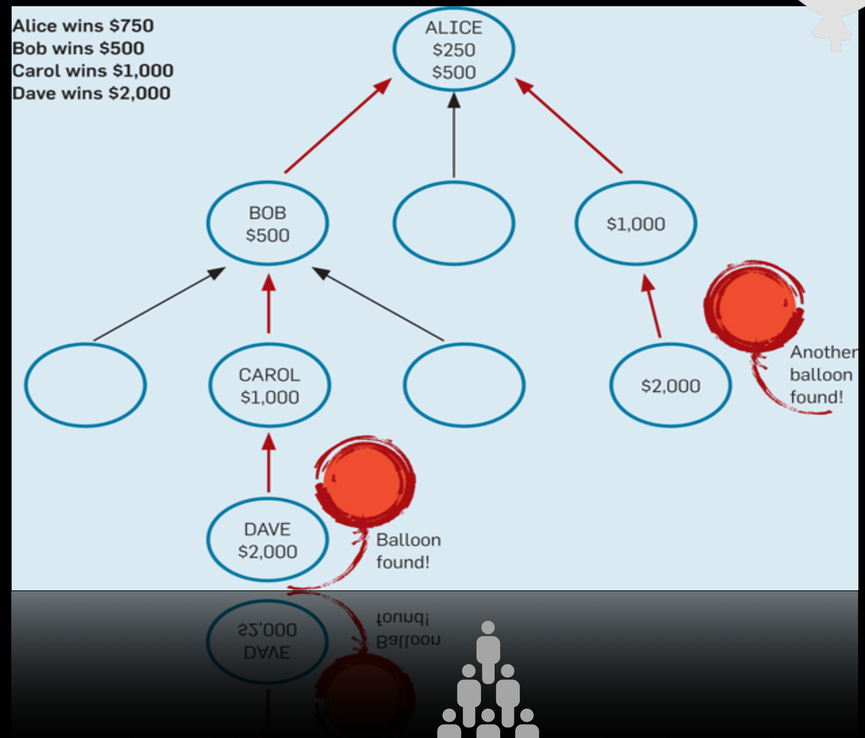
Barriers to Entry & Task Logistics

Data Collection

Measurements, Hypotheses, Analysis

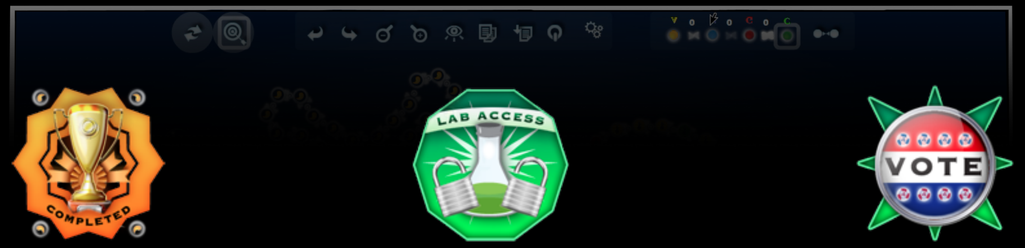
Platform and Infrastructure Design

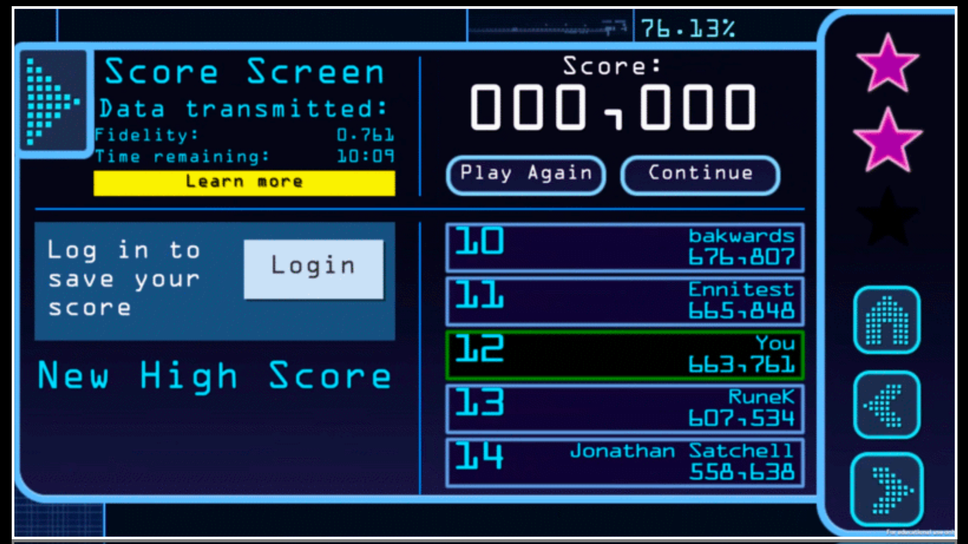
No Infrastructure

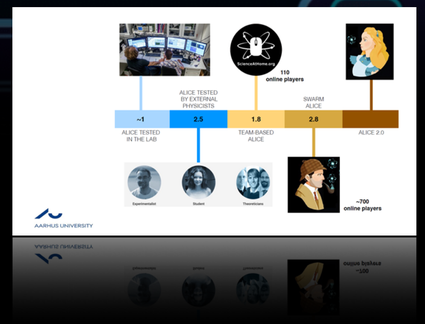


[Pickard et al. Science, 2011]

Game Mechanics & Engineering

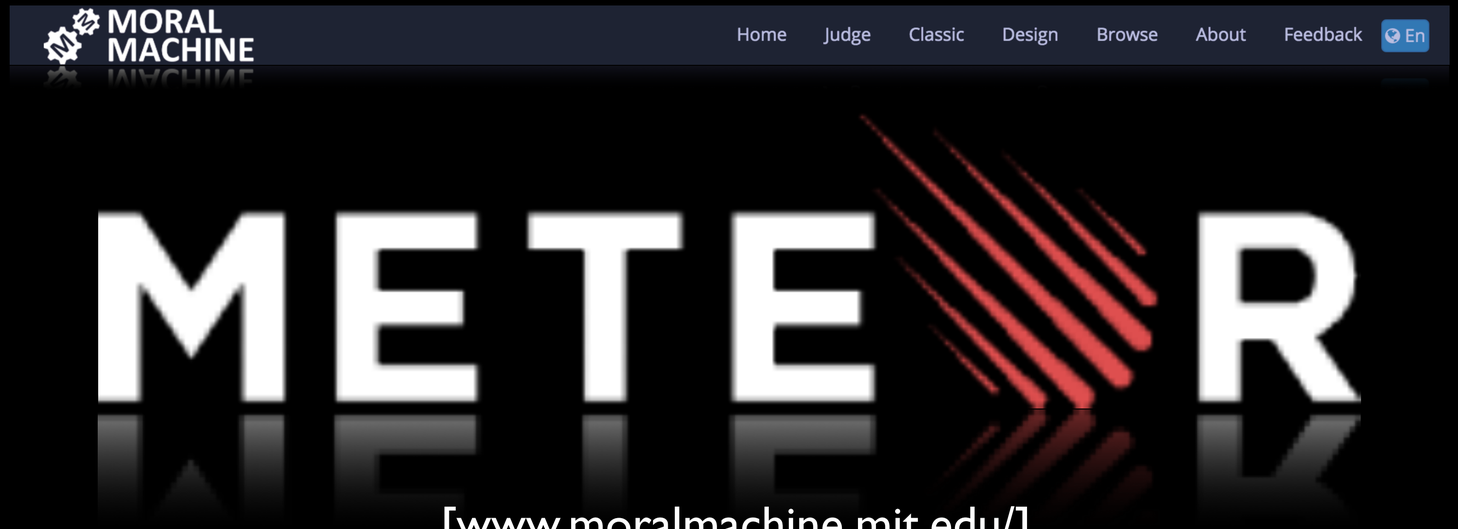




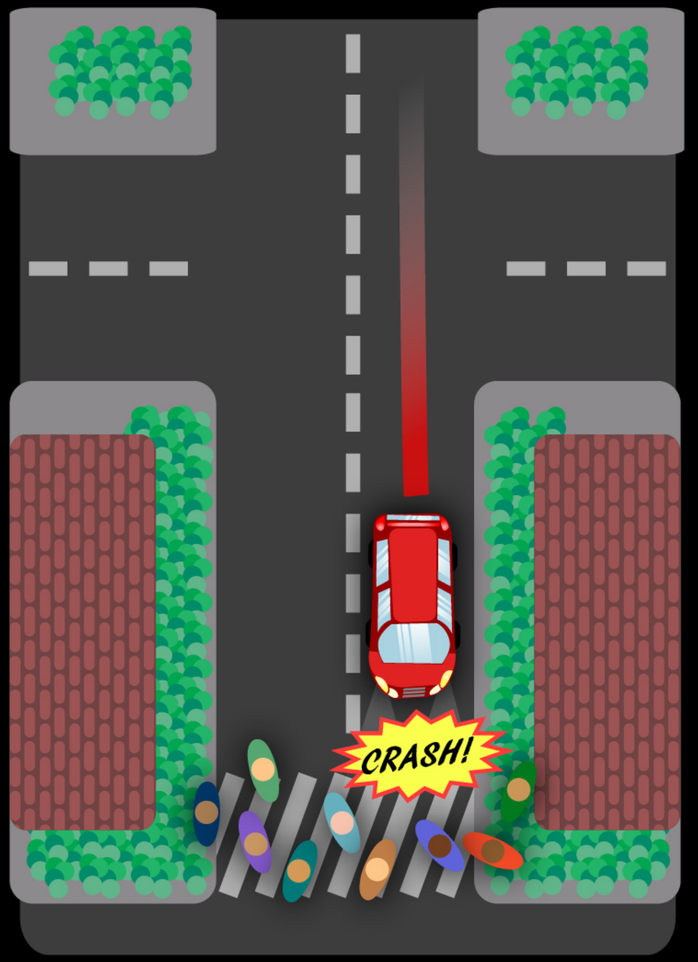


Iterative Development— Prototype, pilot quickly, and fail fast

Web Engineering

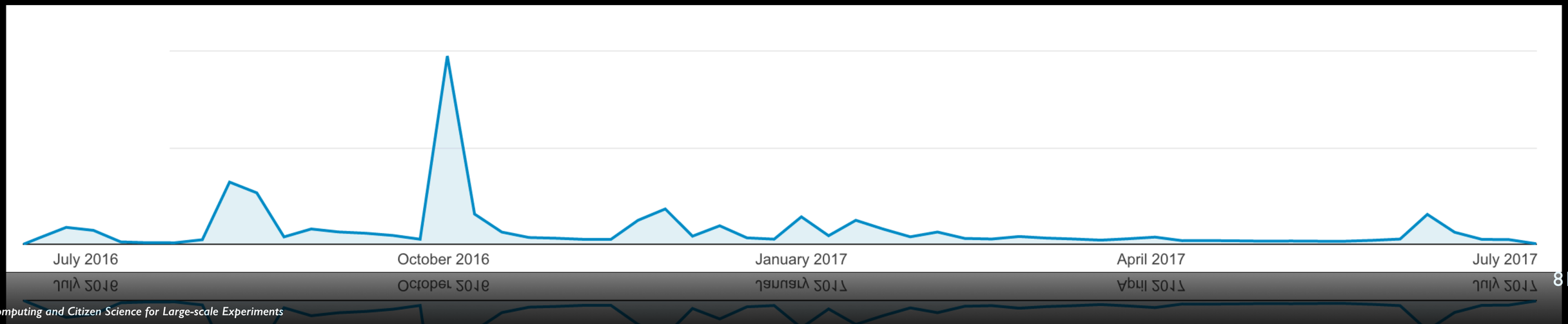


[www.moralmachine.mit.edu/]



Moral Machine - Scaling Up

- Demo run and press-only release
- Initial launch at small scale on cloud server
- Optimization and caching for growing audience
- Shift to scalable pay-by-usage services



CASE
The Problem and Platform Design

Design Thinking & Human Factors

Logistics
(e.g., Open Sourced/ Closed Sourced)

System Engineering

III
Insights

At the end of crowd projects we typically show a nice bundled-up presentation, which looks like everything went smoothly

It didn't!

Murphy's law

think of as much as you can beforehand,
but don't be surprised if something goes wrong

Here are some of our stories showing mistakes and recovery

Launch of Long-run Prisoner's Dilemma Experiment

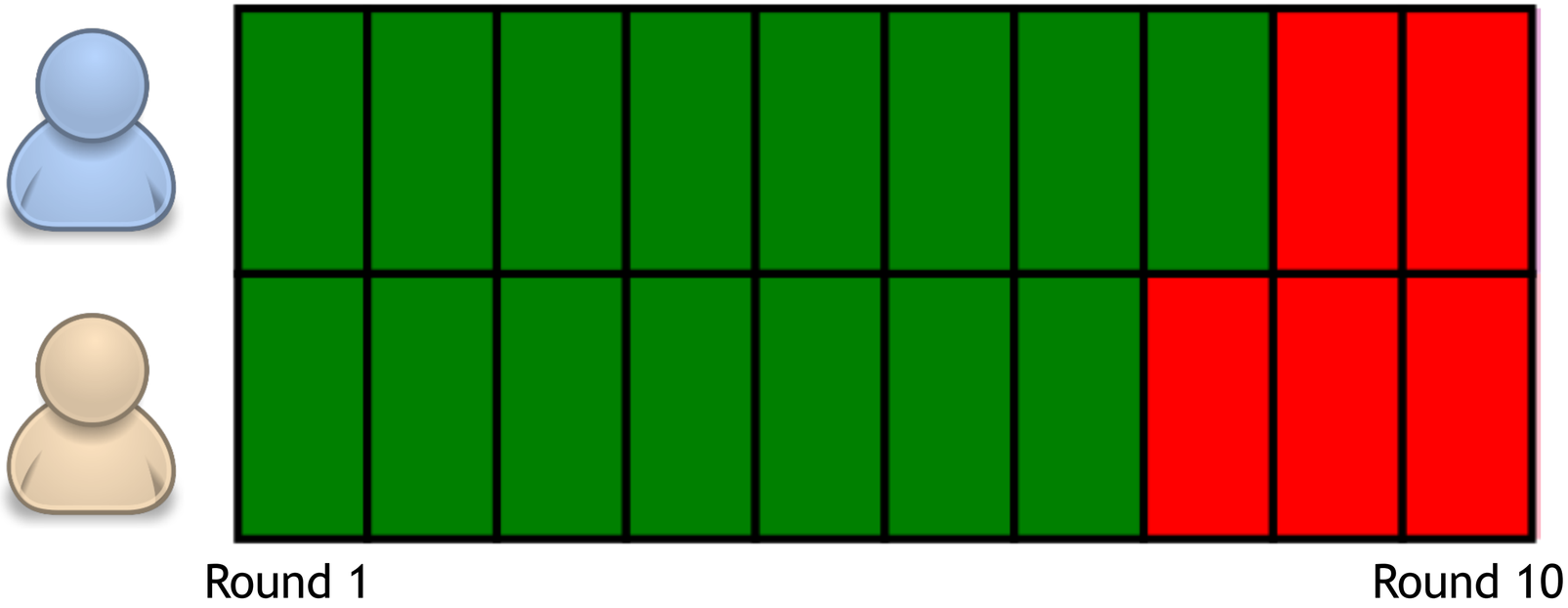
Long-run PD experiment

- Incentives: financial, via MTurk
- Retention: combination of social and financial incentives
- Would people come back each day?
  - We didn't know

Our experiment

amazon mechanical turk™ Artificial Artificial Intelligence

anonymous partners



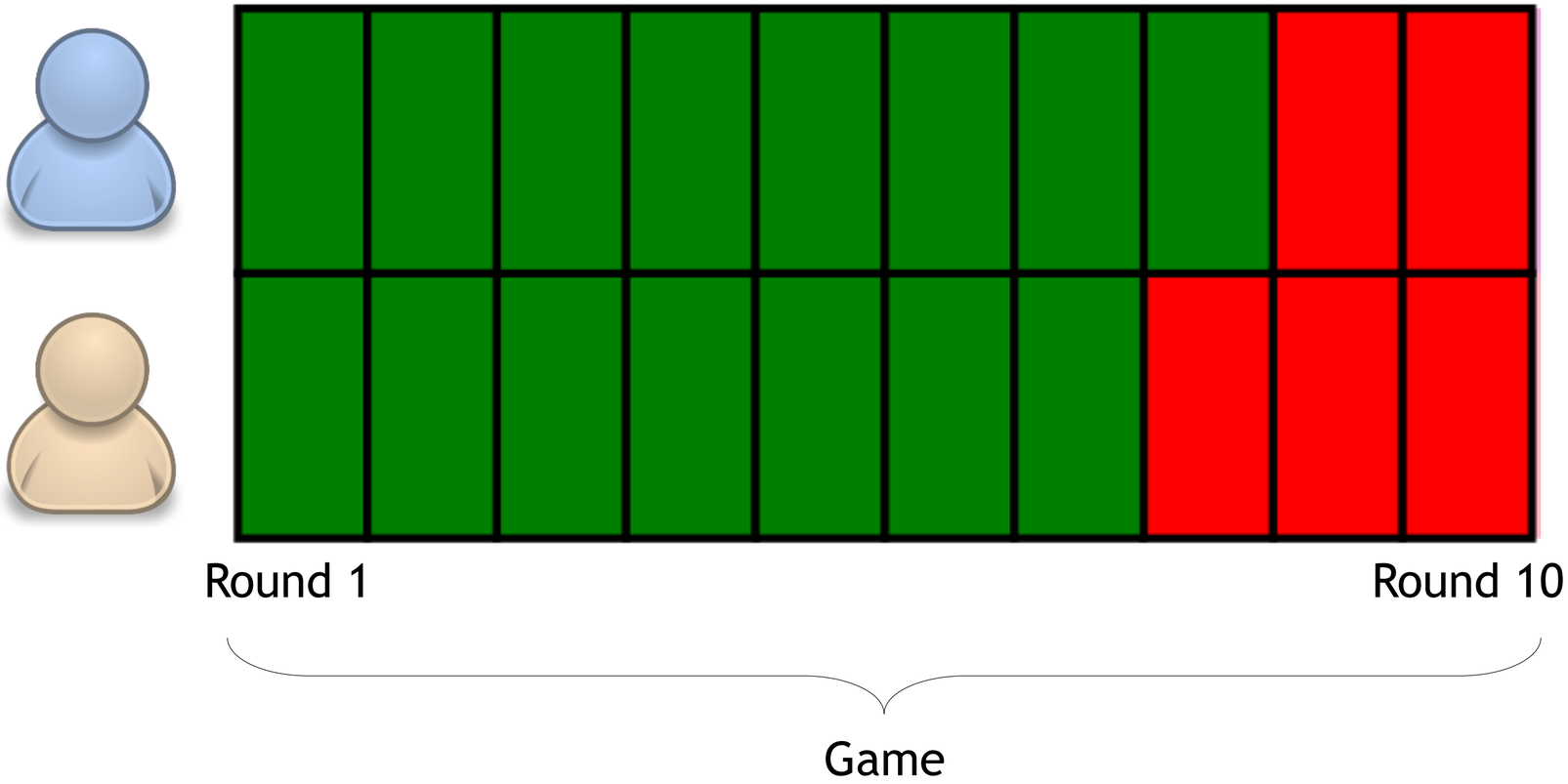
Cooperate

Defect

Our experiment

amazon mechanical turk™ Artificial Artificial Intelligence

anonymous partners



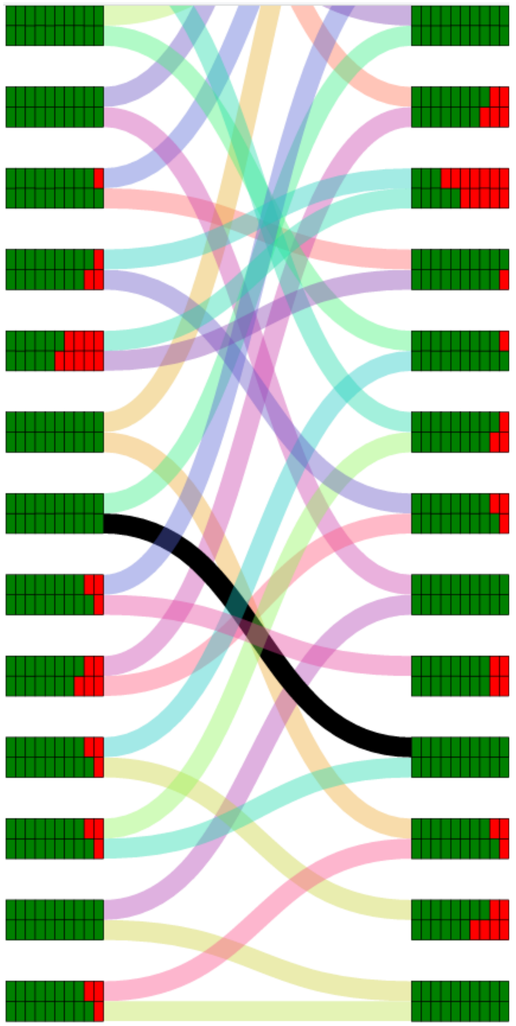
Cooperate

Defect

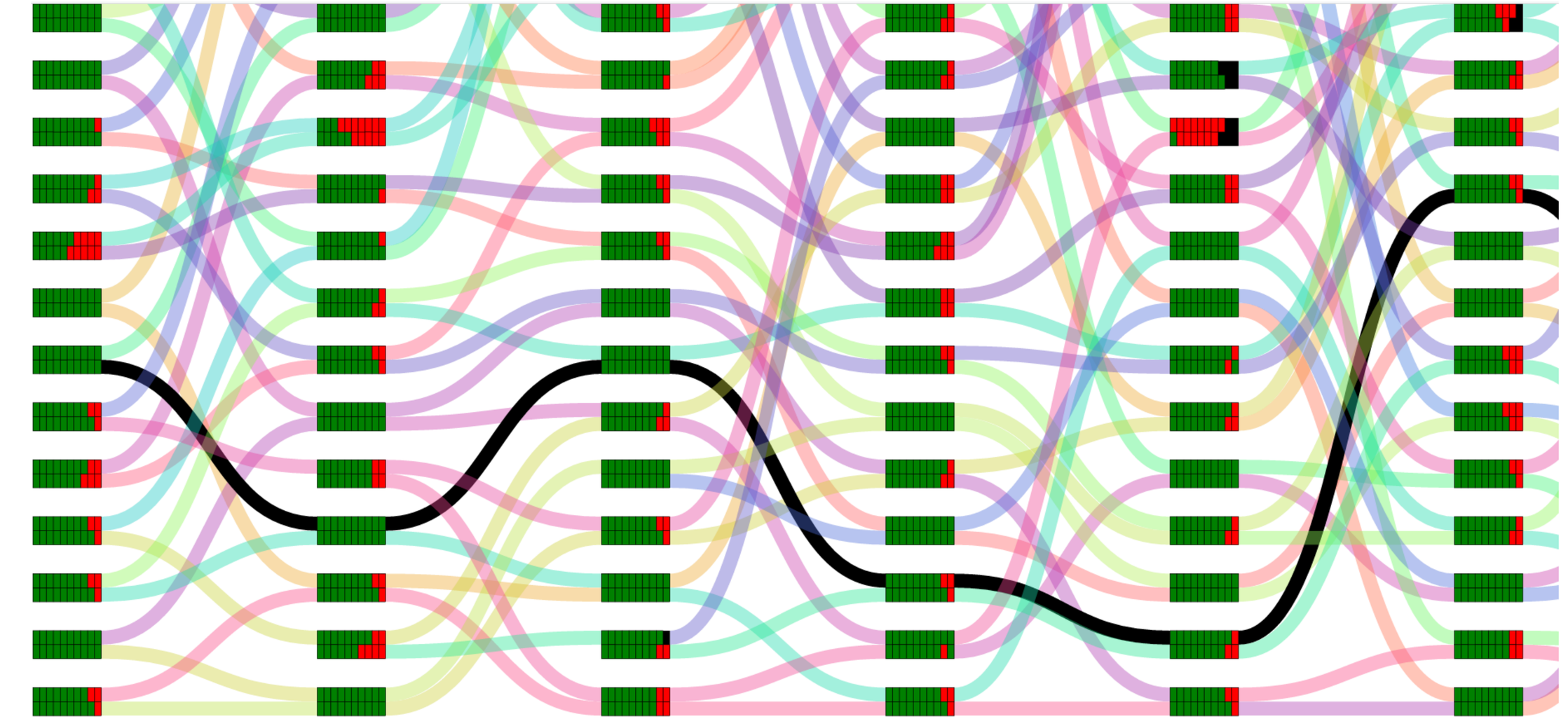
Random rematching across games

Game 1

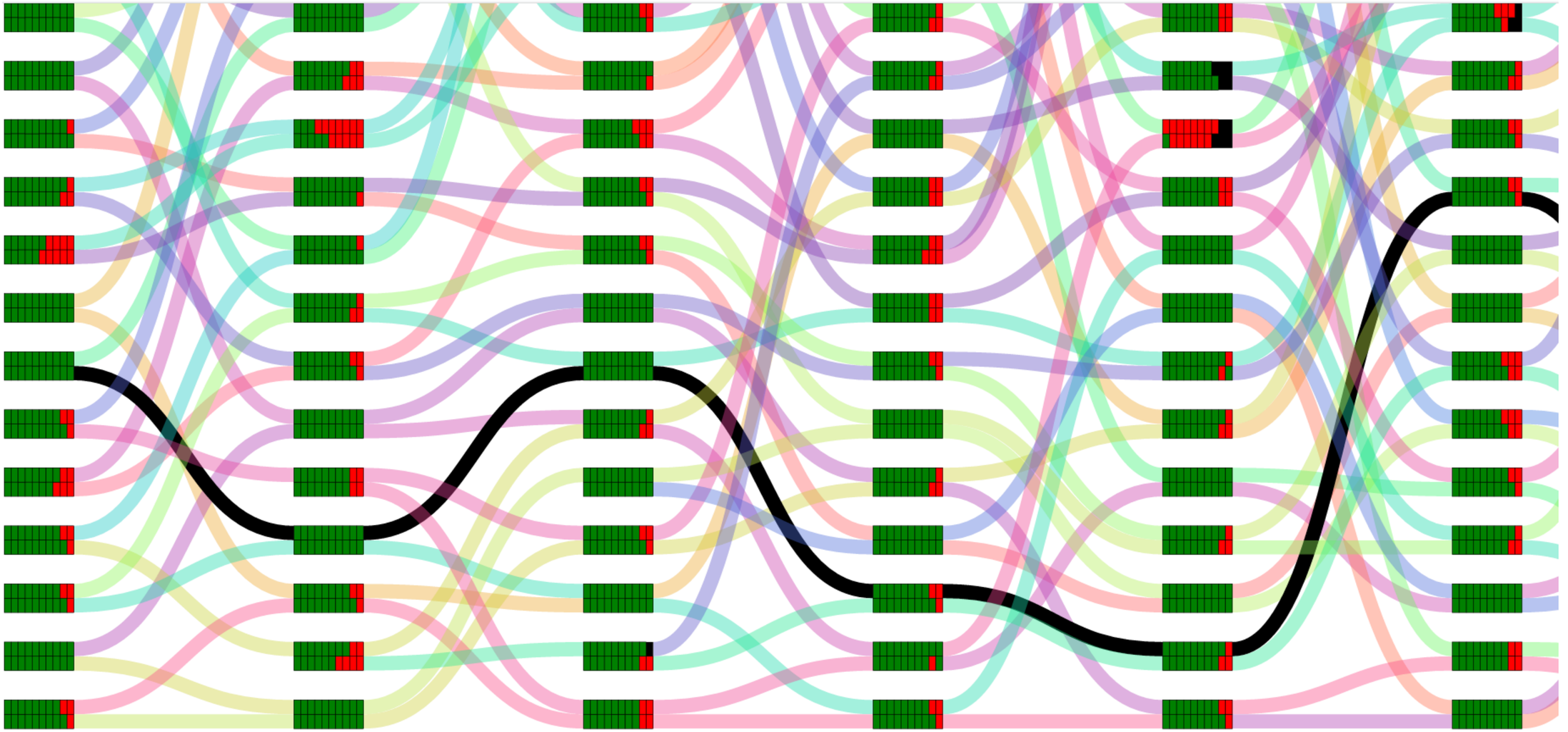
Random rematching across games



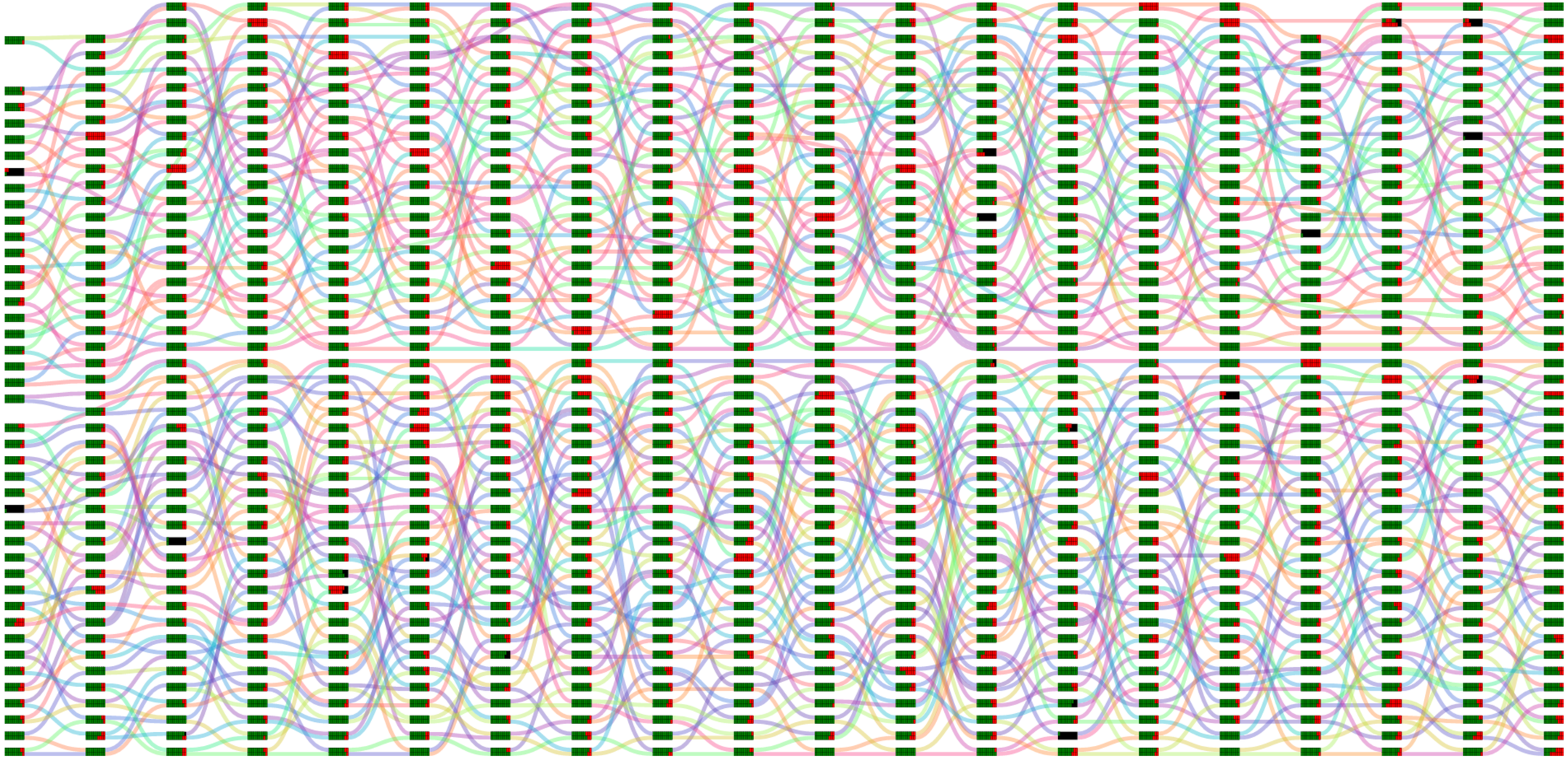
Random rematching across games



Random rematching across games



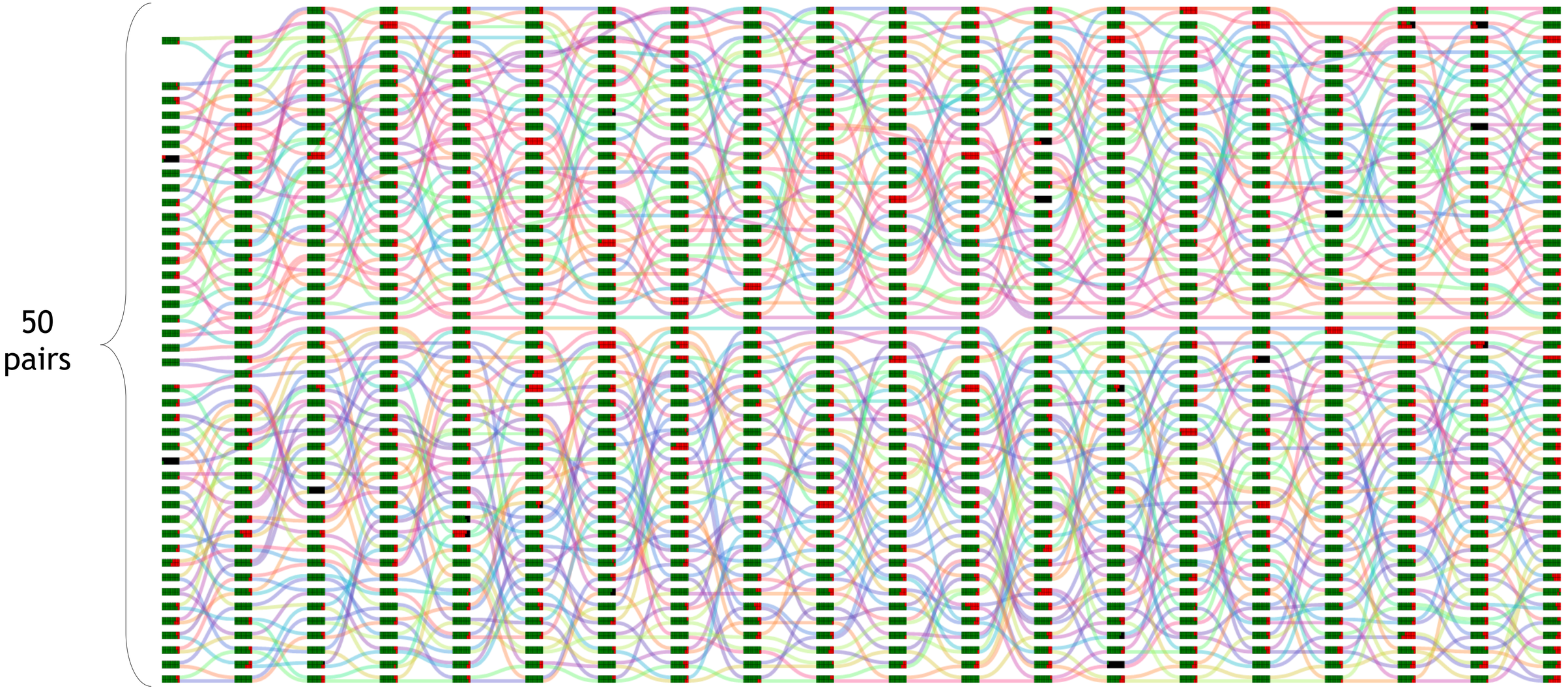
One experiment session – 20 games

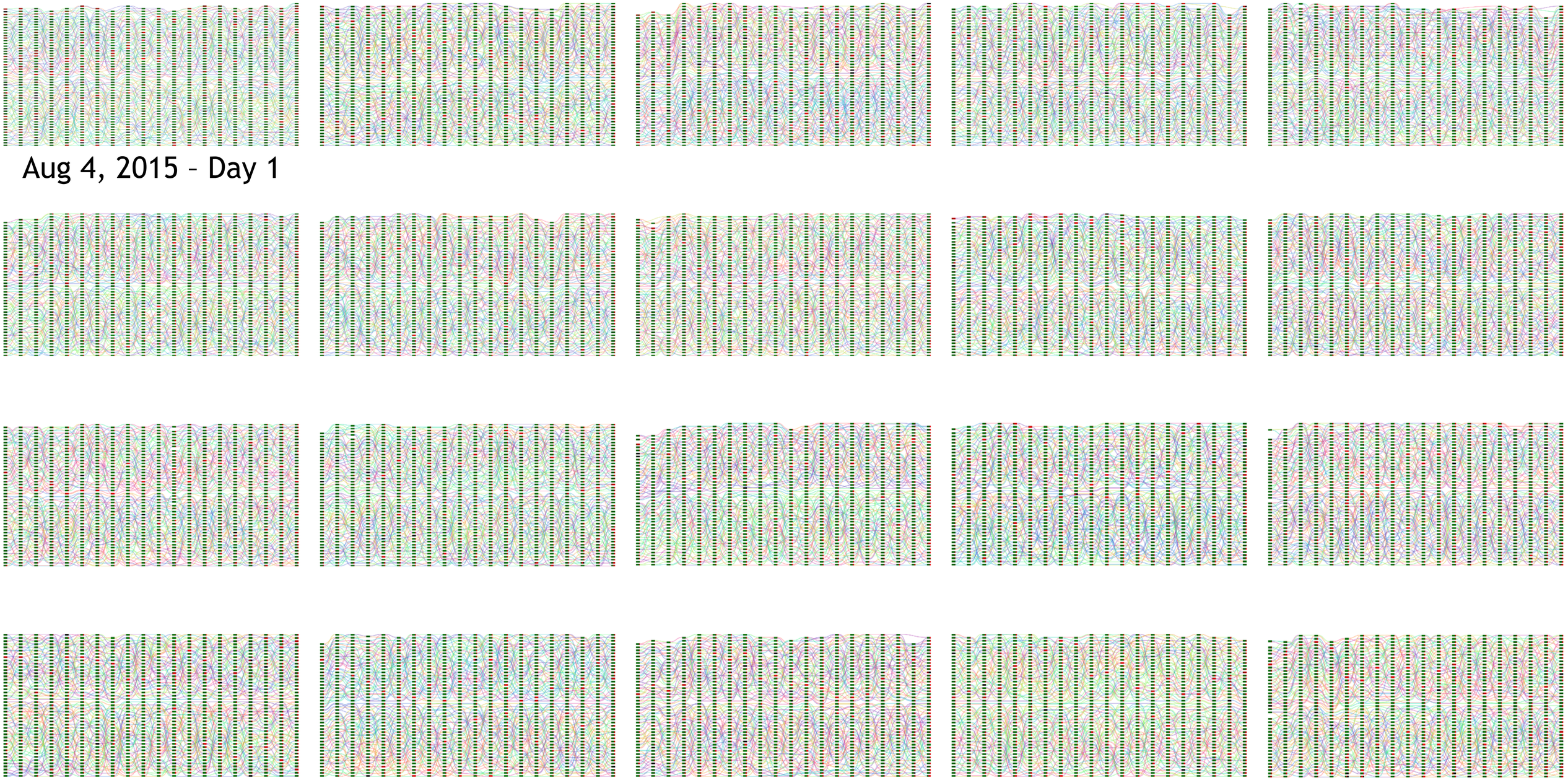


Game 1

Game 20

One experiment session – 20 games





Aug 4, 2015 – Day 1

Aug 31, 2015 – Day 20

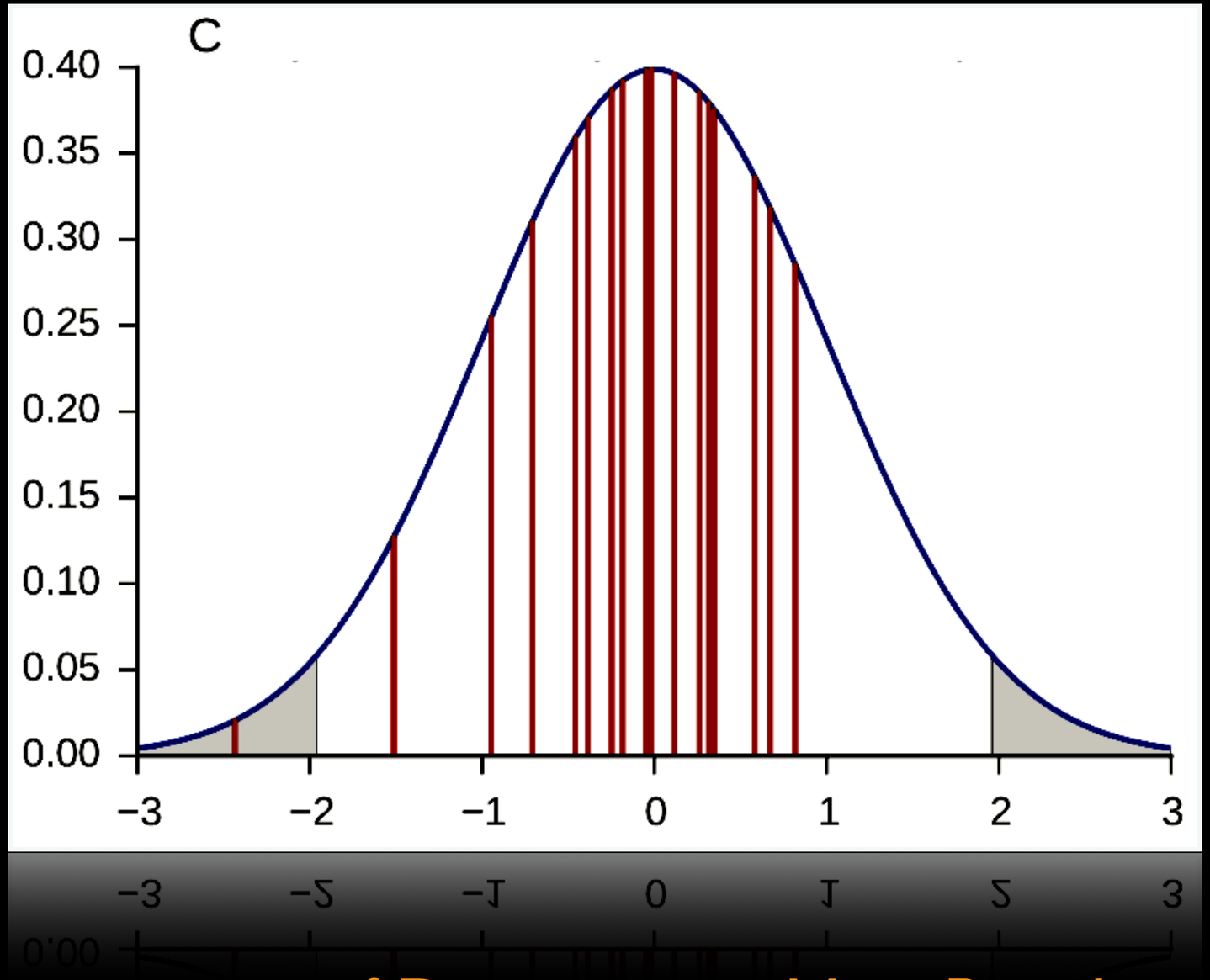
Logistics

- 20 days
- 20 games per day
- 10 rounds per game
- 113 players to start, 94 completed after attrition
- 375,000 decisions: ~20 times longer than previous experiments

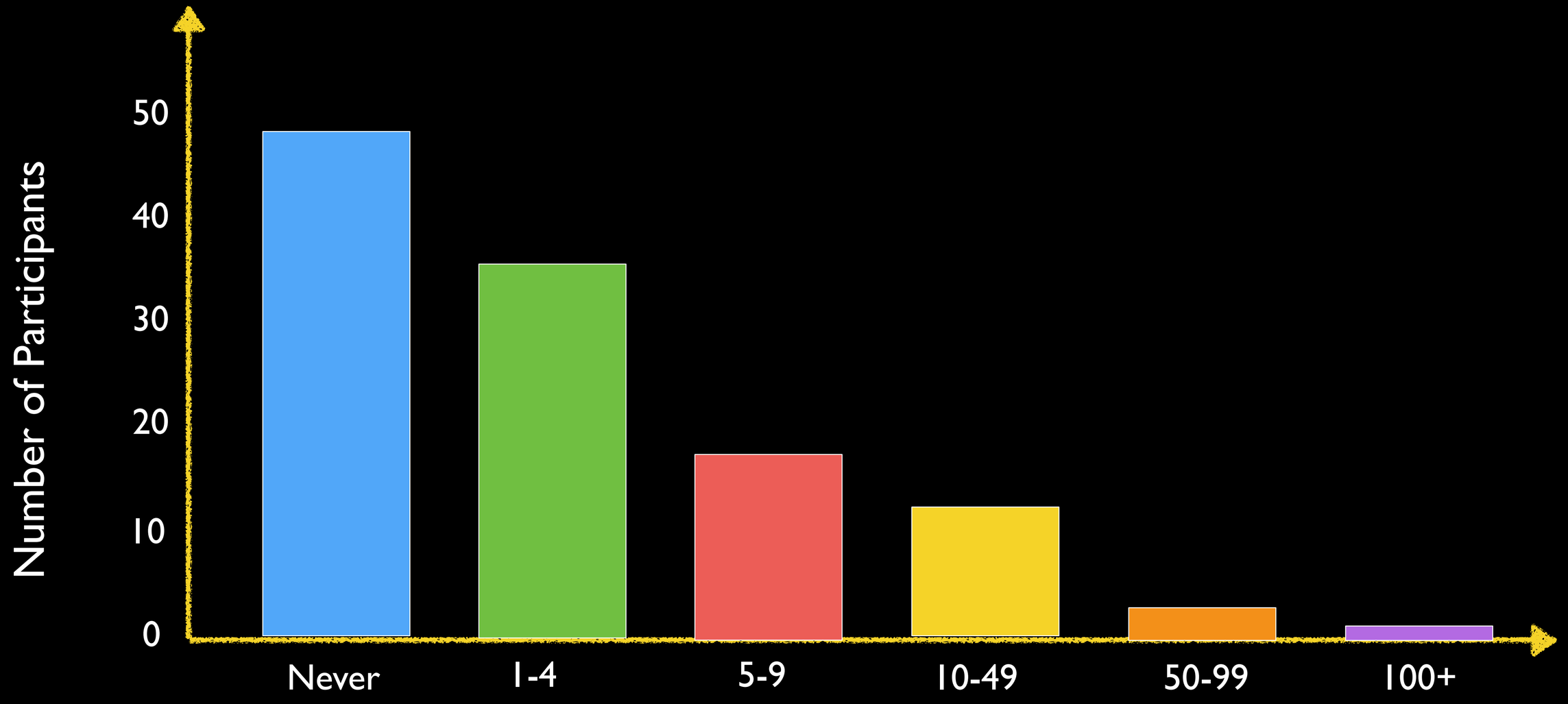
Launch of long-run PD experiment

- Social incentives
  - Expectation of month-long experiment
  - Daily reminders of commitment

- Financial incentives
  - Players kicked out for missing more than two days
  - Daily cash payments proportional to payoffs
  - Lump sum bonus for completion



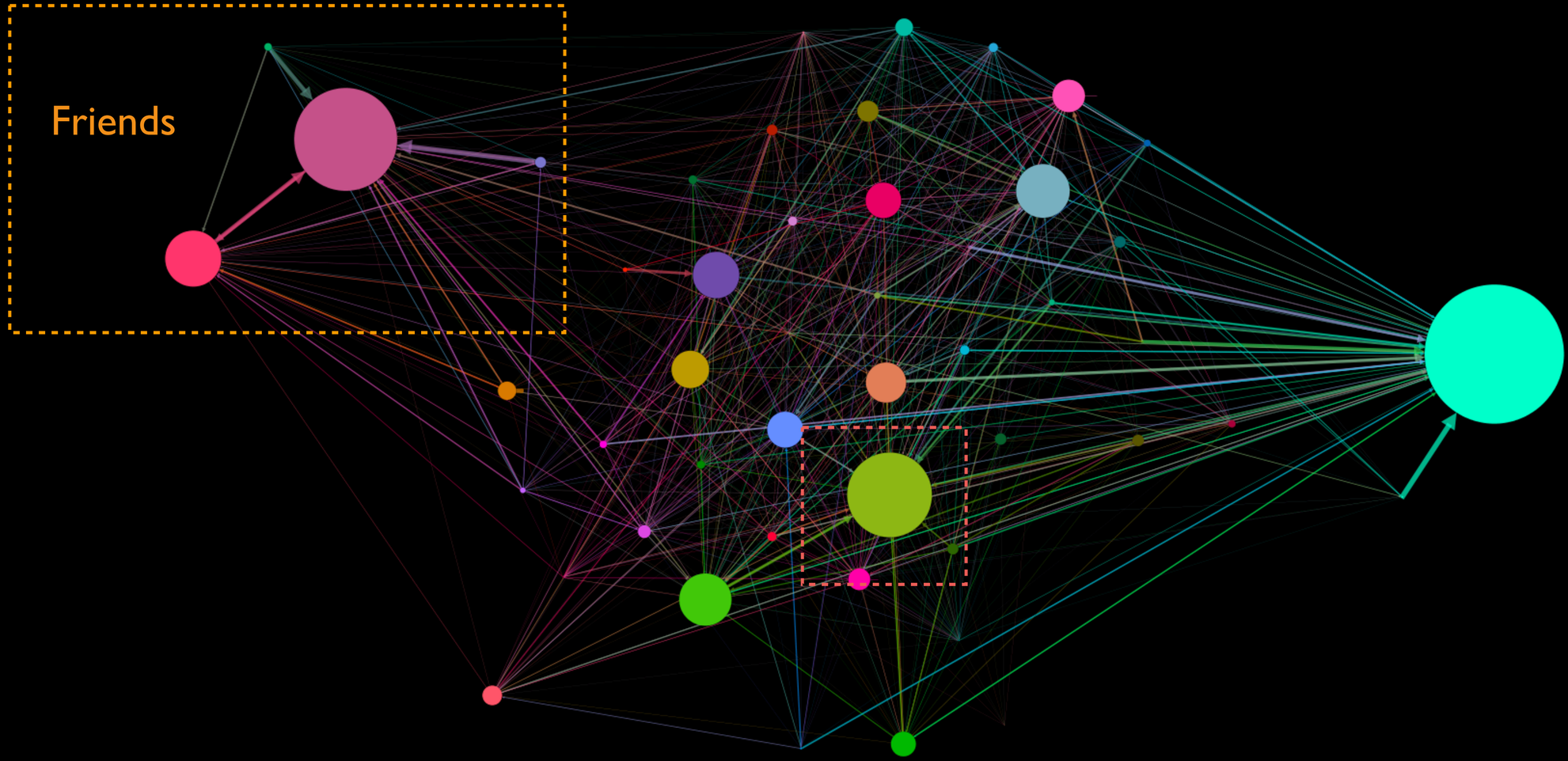
Comparison of Dropouts to Main Population


Number of Participants
50
40
30
20
10
0
Never
1-4
5-9
10-49
50-99
100+
Past Prisoner's Dilemma Experiments

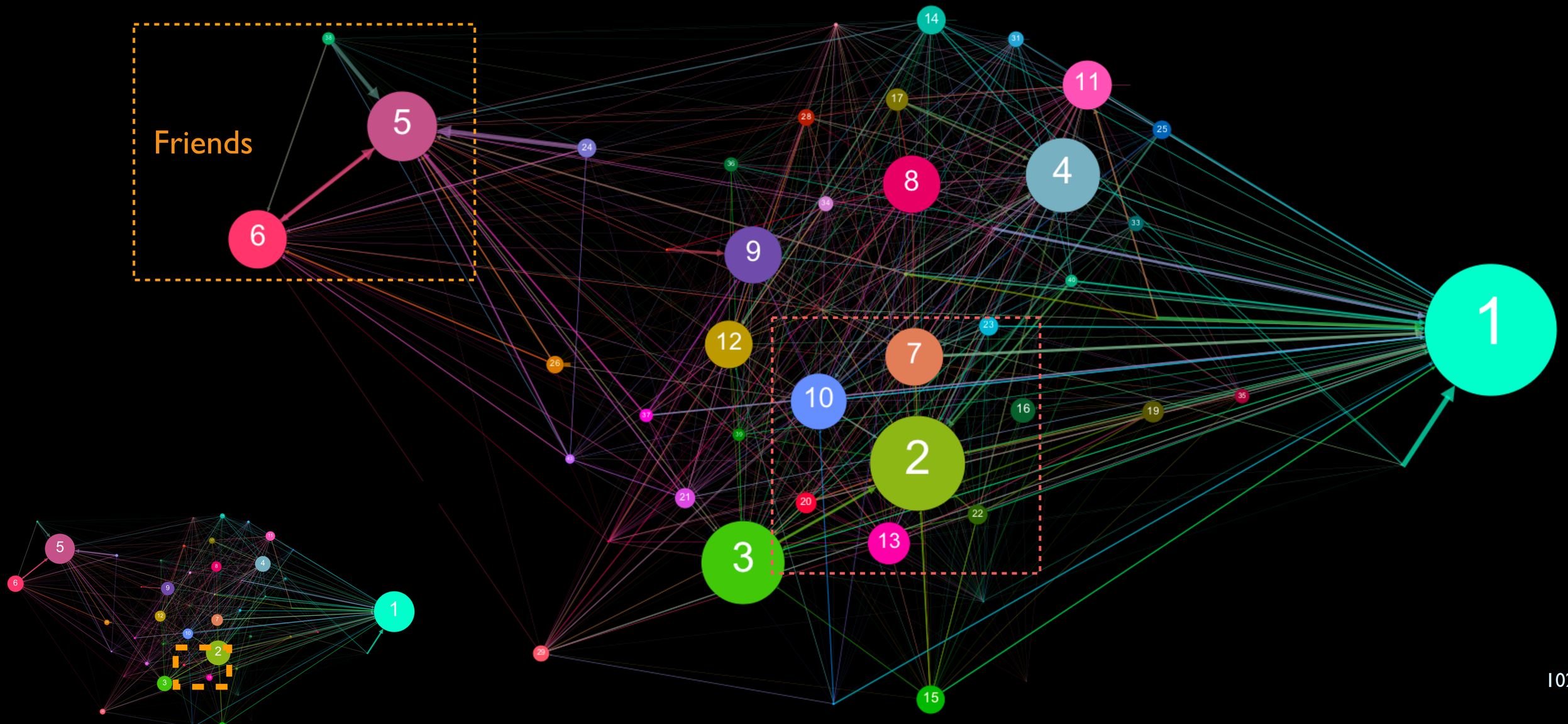
Strategic Behavior in the Stanford Crowd Research

Crowd Dynamics

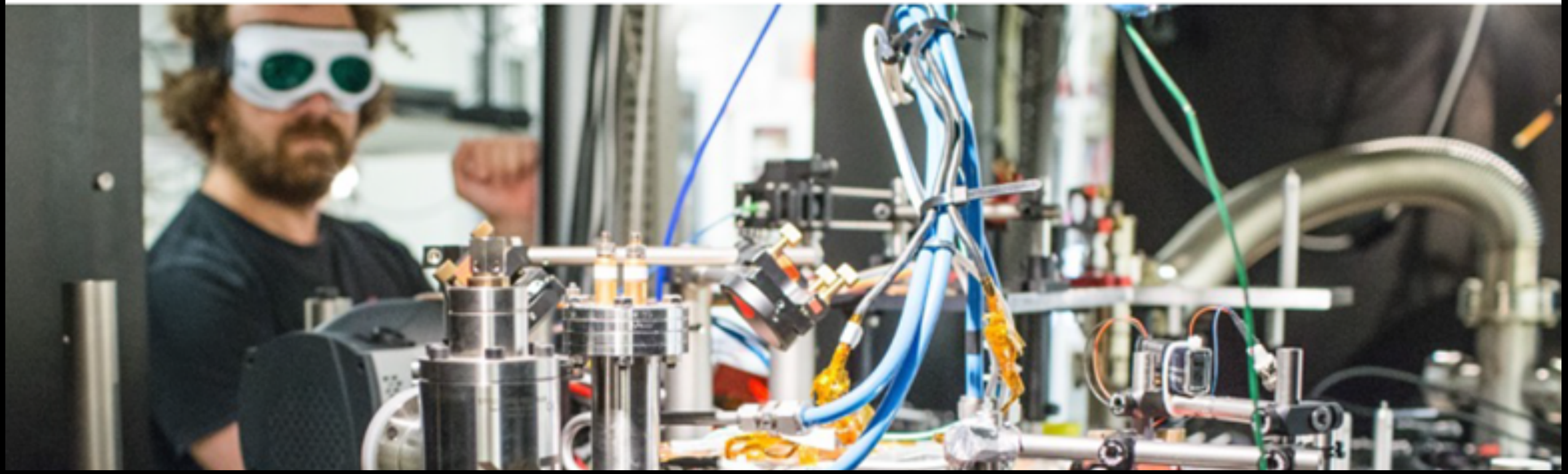
Strategic behavior



Fairness via page rank



Alice Remote Control



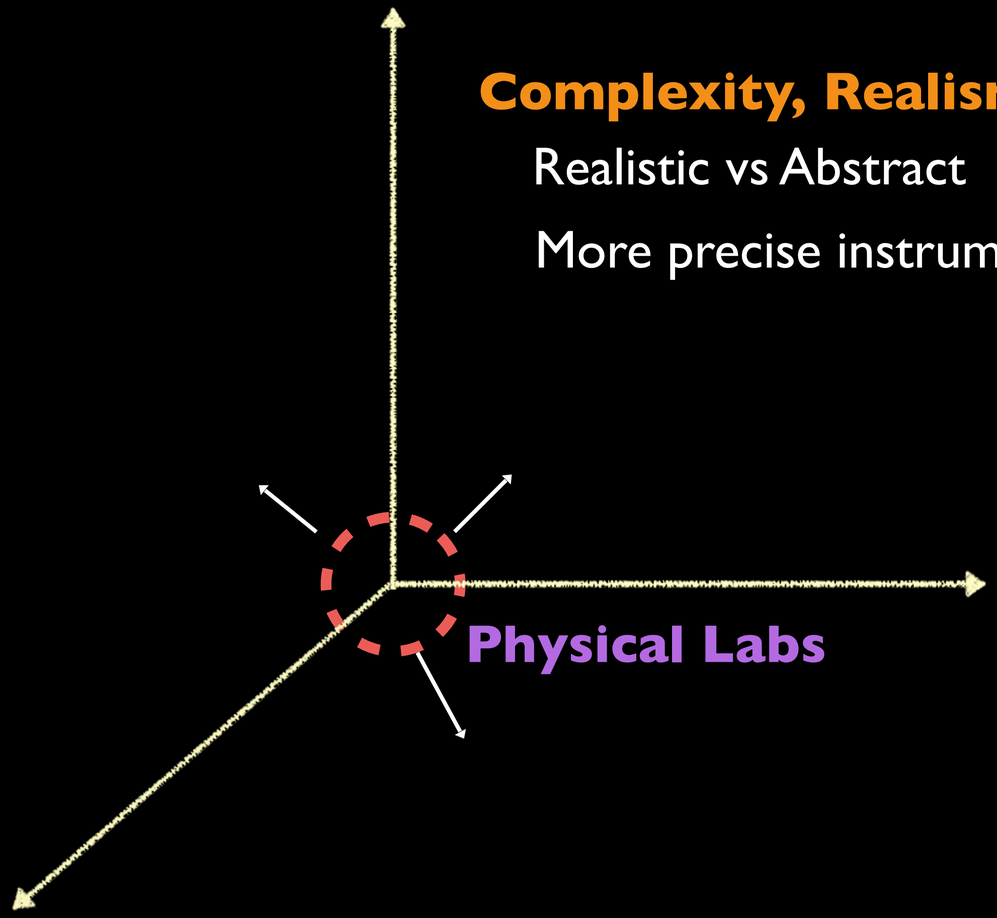
Challenges

- **Unprecedented experimental setup leads to unprecedented problems**
  - Social science investigations were dependent on having a team online simultaneously
  - Recruitment, team allocation, waiting room?
  - Data backlog
  - Gamification matters

Discussion

Crowd = Richer Research at Scale

The Internet as a lab



Complexity, Realism

Realistic vs Abstract

More precise instrumentation

Physical Labs

Duration, Participation

Longer period of time

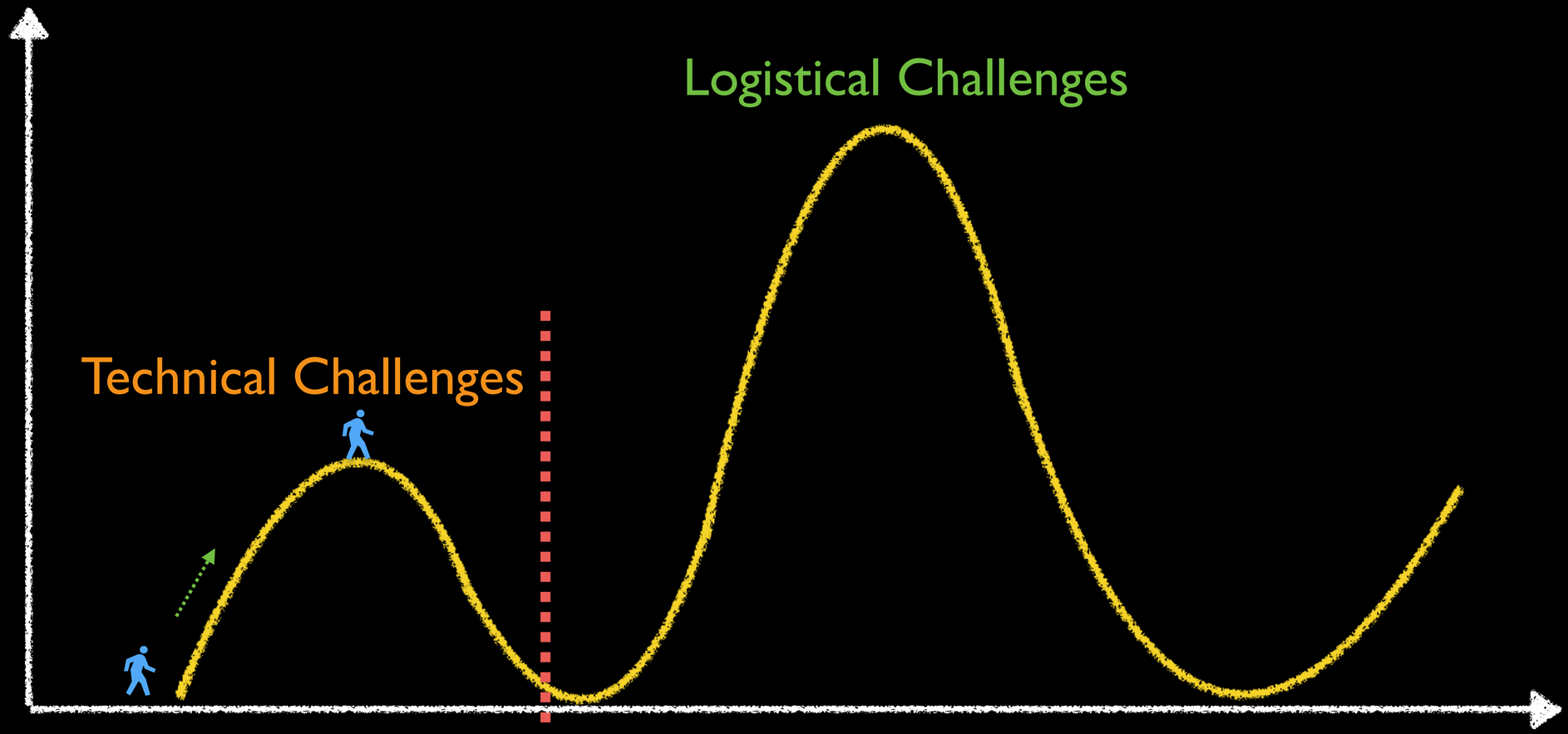
Fewer resource constraints

Size, Scale

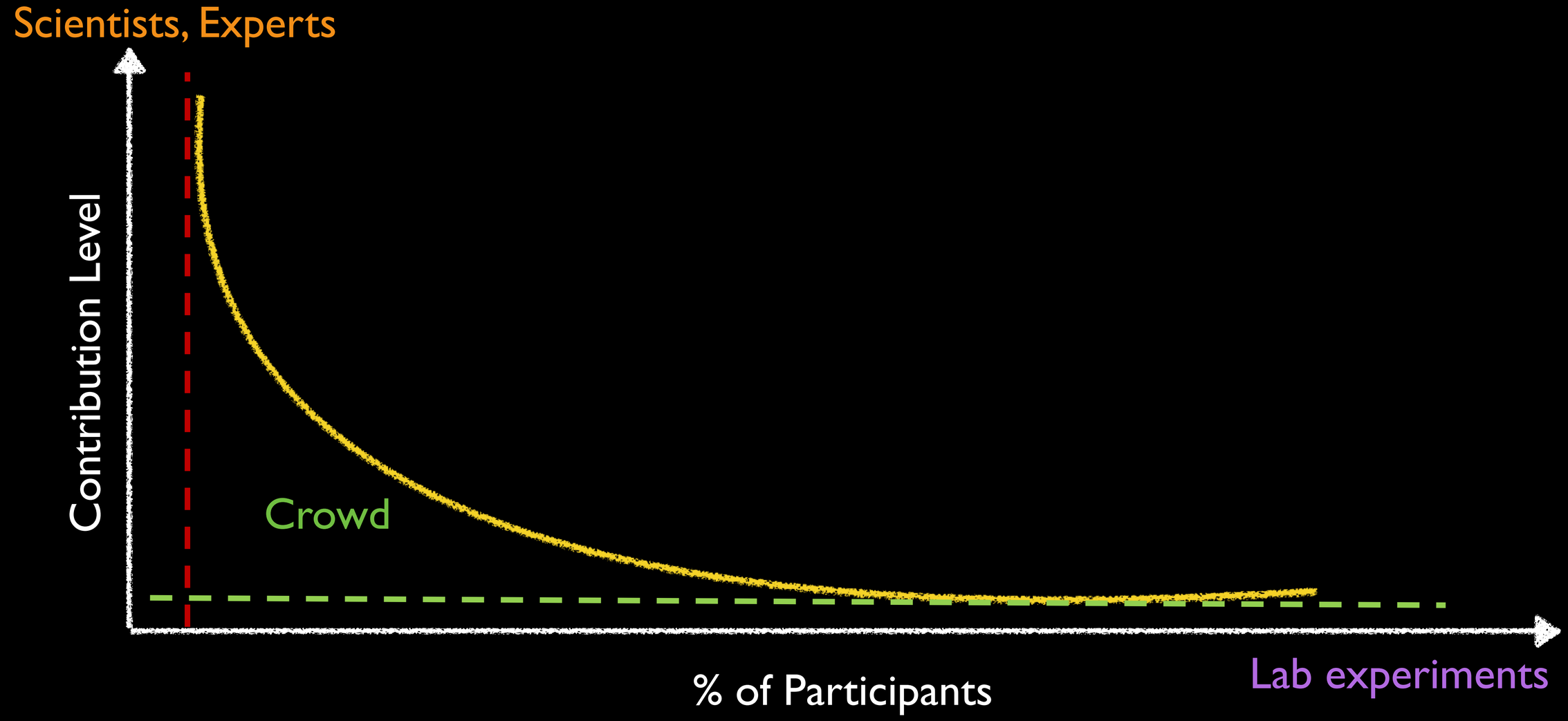
More Samples

Large Social Interactions

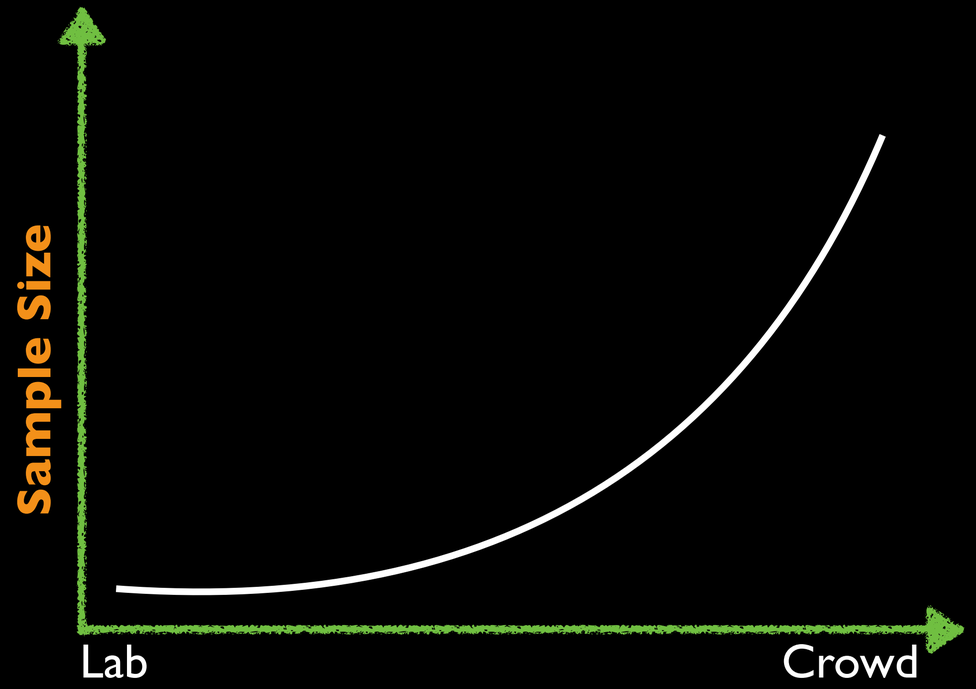
Crowd is synergistic

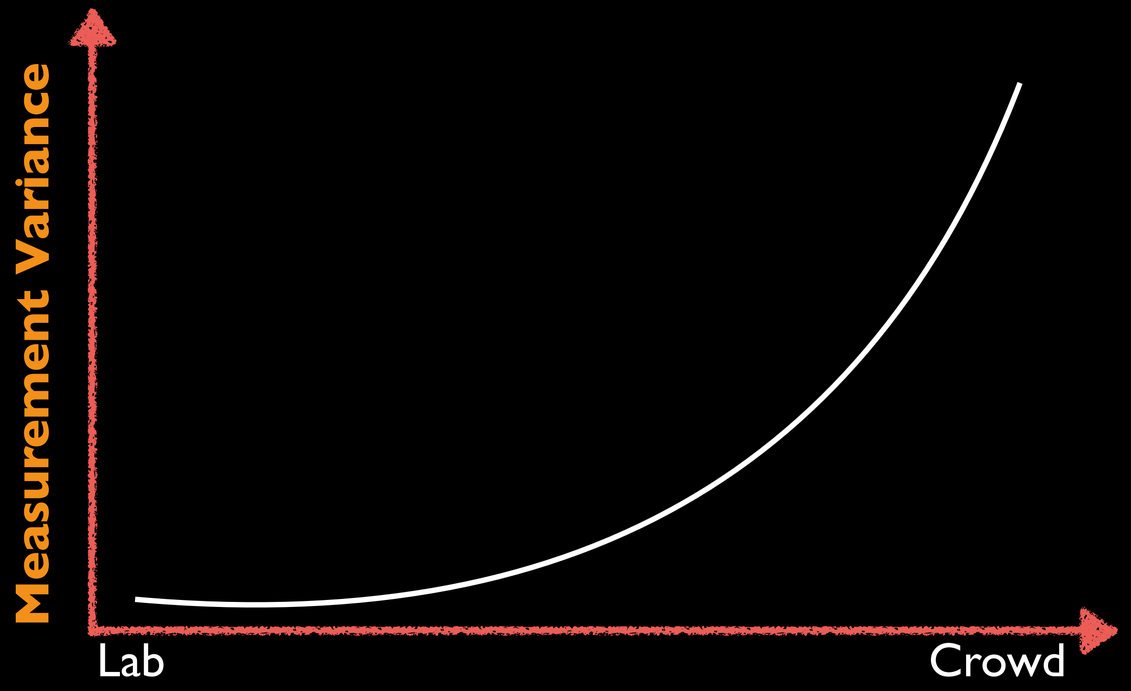


Make use of the long tail

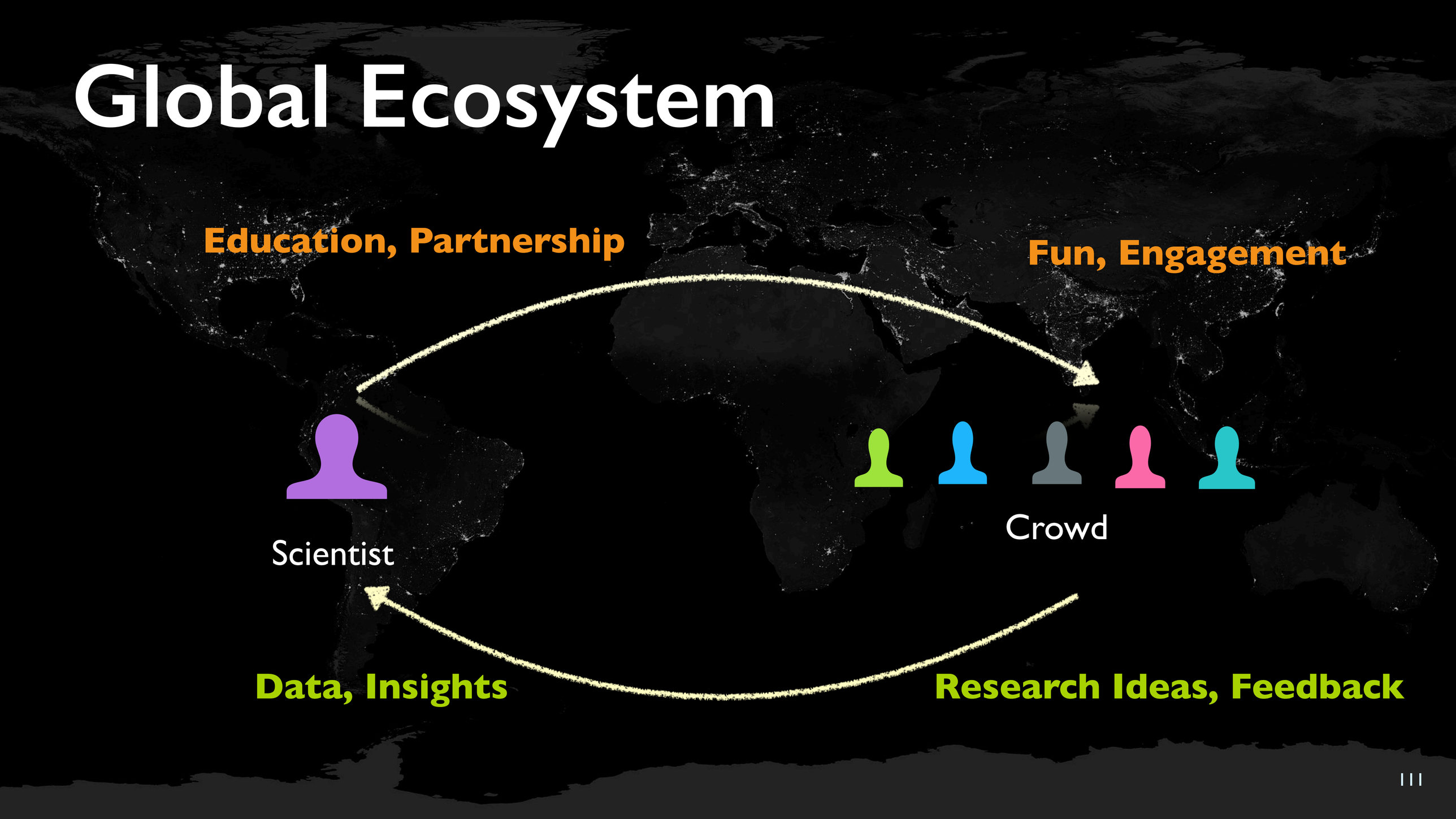


Tradeoffs in increasing size





Global Ecosystem



Non-traditional impact

Democratize research

Stanford Crowd Research

Science at the global scale

EteRNA, Science Home—Quantum Moves

Rethink about AI and society

Moral Machine

Crowdcomputing and Citizen Science for Large-scale Experiments

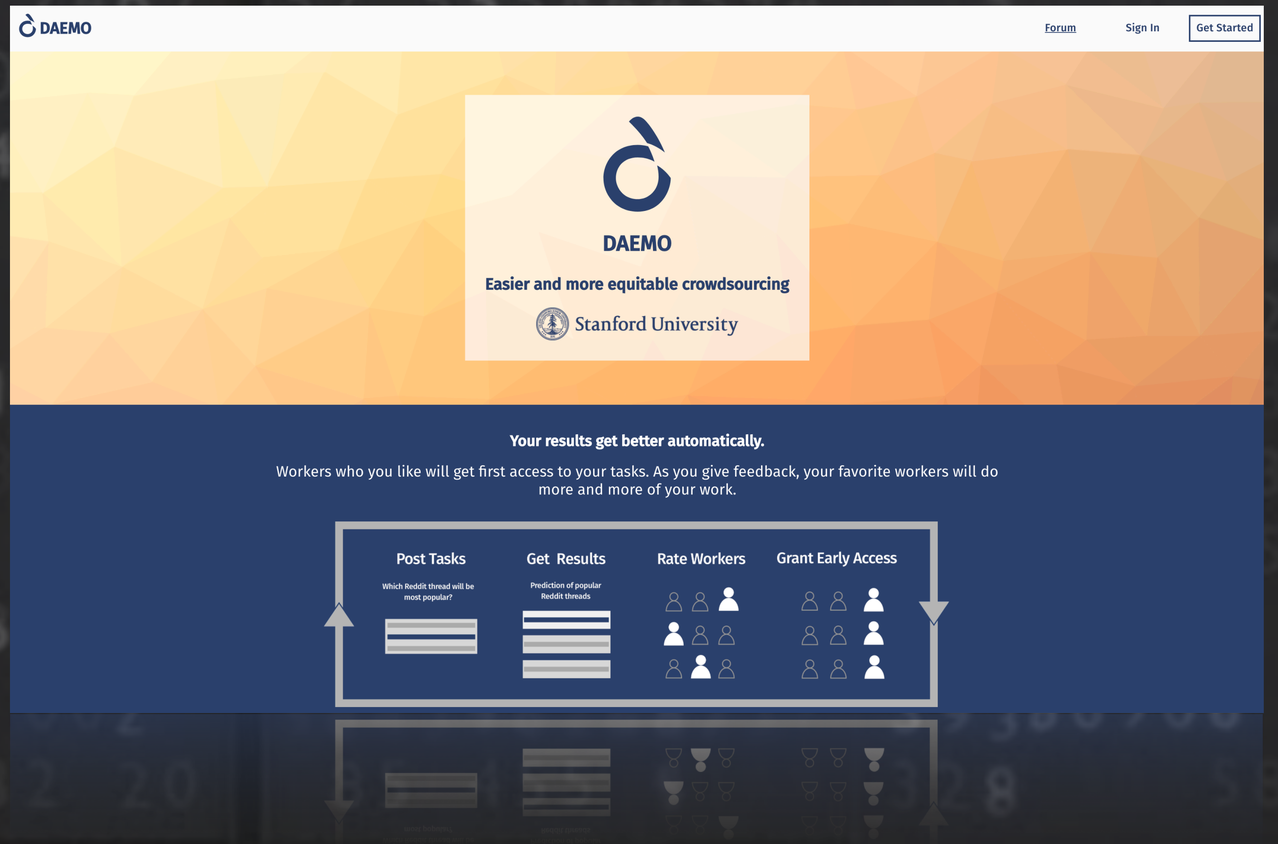
Snehalkumar `Neil' S. Gaikwad[*], Sohan Dsouza[*], Oana Vuculescu[†], Andrew Mao[†‡], Iyad Rahwan[*]

Massachusetts Institute of Technology[*] Microsoft Research New York[‡] Aarhus University[†]

The International Conference on Computational Social Science 2017, Germany

July 10, 2017

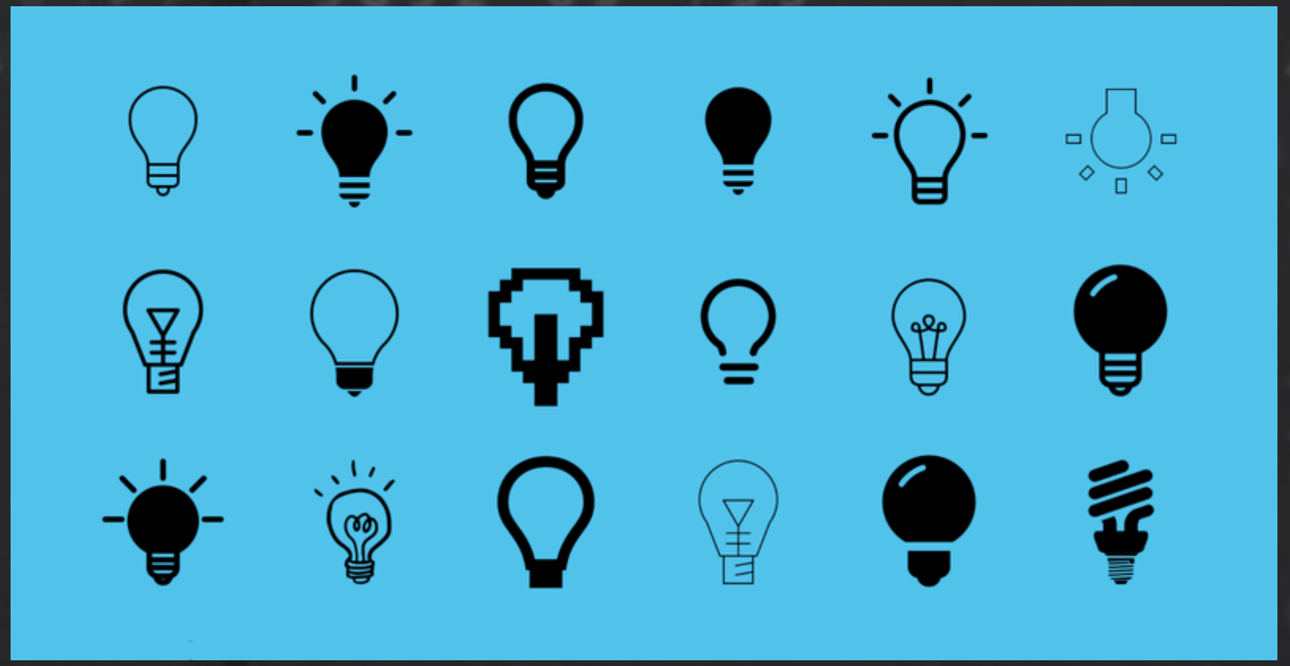
Announcement



Daemo is live!

send an email to gaikwad@mit.edu or daemo@stanford.edu

Announcement



Apply for the Adventurous Science Fellowship ($150,000)

https://www.media.mit.edu/scifellows/

References & Recommended Readings

» [Akerlof *et al. Quarterly Journal of Economics*, 1970] *The market for lemons: Quality uncertainty and the market mechanism*
» *[Antin et al. CHI*, 2012*]* *Social desirability bias & self-reports of motivation: a study of amazon mechanical turk in the US & India*
» [Bernstein *et al. UIST*, 2010] *Soylent: A Word Processor with a Crowd Inside*
» [Bernstein *et al. UIST*, 2011] *Crowds in two seconds: Enabling realtime crowd-powered interfaces*
» [Bigham *et al. UIST*, 2010] VizWiz: nearly real-time answers to visual questions
» [Bigham *et al. ASSETS*, 2011] *The design of human-powered access technology*
» [Bonnefon *et al. Science*, 2016] *The social dilemma of autonomous vehicles, Science*
» [Boudreau *et al. HBR*, 2013] *Using the crowd as an innovation partner*
» [Boudreau *et al. Management science*, 2011] *Incentives and problem uncertainty in innovation contests: An empirical analysis*
» [Cheng *et al. CHI*, 2015] *Measuring crowdsourcing effort with error-time curves*
» [Cooper *et al. Nature*, 2010] *Predicting protein structures with a multiplayer online game*
» [Dow *et al. CSCW*, 2012] *Shepherding the crowd yields better work*
» [Horton *et al. ACM EC*, 2010] *The labor economics of paid crowdsourcing*
» [Horton *et al. Experimental Economics*, 2010] *The online laboratory: Conducting experiments in a real labor market*
» [Gaikwad *et al. UIST*, 2015] *Daemo: A self-governed crowd- sourcing marketplace*
» [Gaikwad *et al. UIST*, 2016] *Boomerang: Rebounding the Consequences of Reputation Feedback on Crowdsourcing Platforms*

References & Recommended Readings

» [Ghosh *et al. ACM EC*, 2011] *Who moderates the moderators?: crowdsourcing abuse detection in user-generated content*
» [Gomes *et al. NIPS*, 2011] *Crowdclustering*
» [Grier, *Princeton University Press*, 2007] *When Computers Were Human*
» [Gupta *et al. CHI*, 2012] *mclerk: enabling mobile crowdsourcing in developing regions*
» [Ho *et al. ICML*, 2013] *Adaptive task assignment for crowdsourced classification*
» [Ho *et al. WWW*, 2015] *Incentivizing high quality crowdwork*
» [Ipeirotis *et al. ACM SIGKDD*, 2010] *Quality management on amazon mechanical turk*
» [Ipeirotis *et al. XRDS*, 2010] *Analyzing the amazon mechanical turk marketplace*
» [Kamar *et al. AAMS*, 2012] *Incentives for truthful reporting in crowdsourcing*
» [Kim *et al. UIST*, 2013] *Cobi: A community-informed conference scheduling tool*
» [Kim *et al. CSCW*, 2017] *Mechanical novel: Crowdsourcing complex work through reflection and revision*
» [Kittur *et al. CHI*, 2008] *Crowdsourcing user studies with mechanical turk*
» [Kittur *et al. CSCW*, 2008] *Harnessing the wisdom of crowds in wikipedia: quality through coordination*
» [Kittur *et al. UIST*, 2011] *Crowdforge: Crowdsourcing complex work*
» [Kittur *et al. CSCW*, 2013] *The future of crowd work*
» [Kraut *et al. MIT Press*, 2011] *Building successful online communities: Evidence-based social design*

References & Recommended Readings

» [Krogh *et al. Research Policy*, 2003] *Community, joining, and specialization in open source software innovation: a case study*
» [Kulkarni *et al. TOCHI*, 2013] *Peer and self assessment in massive online classes*
» [Marcus *et al. Foundations and Trends in Databases*, 2015] *Crowd- sourced data management industry and academic perspectives*
» [McInnis *et al. CHI*, 2016] *Taking a hit: Designing around rejection, mistrust, risk, & workers' experiences in amazon mechanical turk*
» [Nebeling *et al. UIST*, 2015] *WearWrite: Orchestrating the Crowd to Complete Complex Tasks from Wearables (…)*
» [Lasecki *et al. UIST*, 2012] *Real-time captioning by groups of non-experts*
» [Lasecki *et al. CSCW*, 2015] *Information extraction and manipulation threats in crowd-powered systems*
» [Lee *et al. PNAS*, 2014] *RNA design rules from a massive open laboratory*
» [Lee *et al. AAAI HCOMP*, 2014] *Crowdsourcing for Participatory Democracies: Efficient Elicitation of Social Choice Functions*
» [Little *et al. ACM SIGKDD workshop on human computation*, 2010] *Exploring iterative and parallel human computation processes*
» [Luther *et al. ACM CSCW*, 2015] *Structuring, aggregating, and evaluating crowdsourced design critique*
» [Malone *et al. ACM CSCW*, 1990] *What is coordination theory and how can it help design cooperative work systems?*
» [Malone *et al. ACM SIGSMALL*, 1980] *What makes things fun to learn? Heuristics for designing instructional computer games.*
» [Malone *et al. MIT Sloan Management Review*, 2010] *The Collective Intelligence Genome*

References & Recommended Readings

» [Mao *et al. AAAI HCOMP*, 2013] *Volunteering versus work for pay: Incentives and tradeoffs in crowdsourcing*
» [Mao *et al. AAAI HCOMP*, 2012] *Turkserver: Enabling synchronous and longitudinal online experiments*
» [Mao *et al. Nature Comm*, 2016] *Resilient cooperators stabilize long-run cooperation in the finitely repeated prisoners dilemma*
» [Mao *et al. PloS One*, 2016] *An experimental study of team size and performance on a complex task*
» [Mason *et al. ACM SigKDD Explorations Newsletter*, 2010] *Financial incentives and the performance of crowds*
» [Mason *et al. Behavior research methods*, 2012] *Conducting behavioral research on amazon's mechanical turk*
» [Martin *et al. CSCW*, 2014] *Being a Turker*
» [Parameswaran *et al. VLDB*, 2011] *Human-assisted graph search: it's okay to ask questions*
» [Peng *et al. AAAI*, 2010] *Decision-theoretic control of crowd-sourced workflows*
» [Pickard *et al. Science*, 2011] *Time-Critical Social Mobilization*
» [Prelec *et al. Nature*, 2017] *A solution to the single-question crowd wisdom problem*
» [Prelec *et al. Science*, 2004] *A Bayesian truth serum for subjective data*
» [Procaccia *et al. ACM Transactions on Economics and Computation*, 2013] *Approximate mechanism design without money*
» [Rutherford *et al. PNAS*, 2012] *Limits of social mobilization*
» [Rzeszotarski *et al. UIST*, 2011] *Instrumenting the crowd: using implicit behavioral measures to predict task performance*

References & Recommended Readings

» [Salehi *et al.* CHI, 2015] *We are dynamo: Overcoming stalling and friction in collective action for crowd workers*
» [Shah *et al. HRI*, 2011] *Improved human-robot team performance using chaski, a human-inspired plan execution system*
» [Shaw *et al. CSCW*, 2011] *Designing incentives for inexpert human raters*
» *[Sheng et al. ACM SIGKDD*, 2008] *Get another label? Improving data quality and data mining using multiple, noisy labelers*
» [Silberman *et al. ACM SIGKDD workshop on Human Computation*, 2010] *Sellers' problems in human computation markets*
» [Singer *et al. WWW*, 2013] *Pricing mechanisms for crowdsourcing markets*
» [Sørensen *et al. Nature*, 2016] *Exploring the quantum speed limit with computer games*
» [Stephane *et al. Upwork*, 2015] *Hire fast & build things: How to recruit and manage a top-notch team of distributed engineers*
» [Stefanovitch *et al. EPJ Data Science*, 2014] *Error and attack tolerance of collective problem solving: The darpa shredder challenge*
» [Vaish *et al. CHI*, 2014] *Twitch crowdsourcing: crowd contributions in short bursts of time*
» [Vaish *et al. UIST*, 2017] *Crowd Research: Open and Scalable University Laboratories*
» [*Valentine et al. CHI*, 2017] *Flash Organizations: Crowdsourcing Complex Work by Structuring Crowds As Organizations*
» [Von Ahn *et al. CHI*, 2004] *Labeling images with a computer game*
» [Von Ahn *et al. Communications of the ACM*, 2008] *Designing Games With a Purpose*
» *[Watts et al. Journal of consumer research*, 2007] *Networks, dynamics, and the small-world phenomenon*
» *[Woolley et al. Science*, 2010] *Evidence for a collective intelligence factor in the performance of human groups*
» *[Werfel et al. Science*, 2014] *Designing Collective Behavior in a Termite-Inspired Robot Construction Team*