12574 970 43 808 82012998 041498186 44 26441517

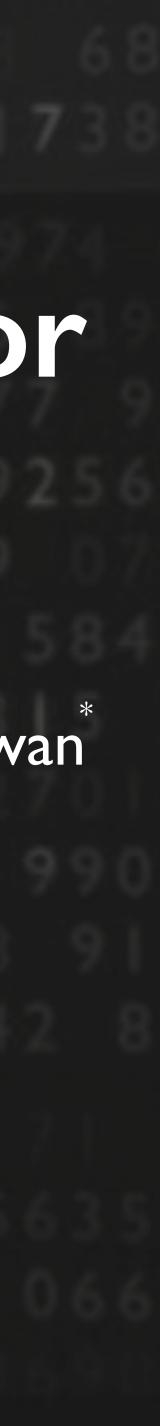
Crowdcomputing and Citizen Science for Large-scale Experiments

Snehalkumar `Neil' S. Gaikwad, Sohan Dsouza, Oana Vuculescu, Andrew Mao, Iyad Rahwan Microsoft Research New York[‡] Aarhus University[†] Massachusetts Institute of Technology

The International Conference on Computational Social Science, 2017, Germany

536 722 60 36 54

uly 10, 2017



A BRIEF HISTORY OF CROWD COMPUTING

CROWD MECHANICS

Crowdcomputing and Citizen Science for Large-scale Experiments

INSIGHTS

Interactive 0 cho 0 Case P 0 Discussions ro P B



Crowdcomputing and Citizen Science for Large-scale Experiments

A Brief History of Crowd Computing



Crowdcomputing and Citizen Science for Large-scale Experiments

In the 9th-20th century



when computers were human

Crowdcomputing and Citizen Science for Large-scale Experiments

WHEN COMPUTERS Were Human

David Alan Grier





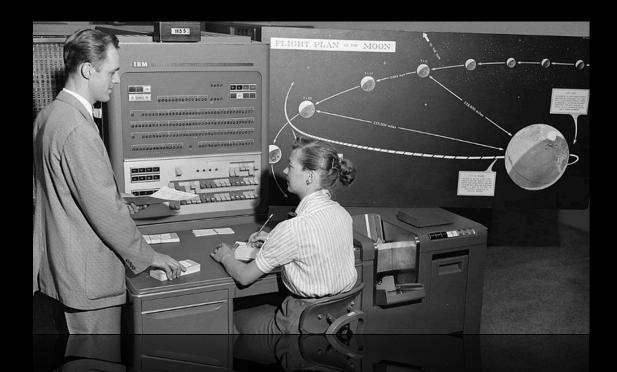
The Mathematical Tables Project 450 human computers







Hidden Figures



Human Computers and IBM 704 source: JPL



Human Computers 1936

BASED ON THE UNTOLD TRUE STORY MEET THE WOMEN YOU DON'T KNOW, BEHIND THE MISSION YOU DO.

Crowdcomputing and Citizen Science for Large-scale Experiments

source: JPL



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



In the 2 st century



Rise of the Connected World

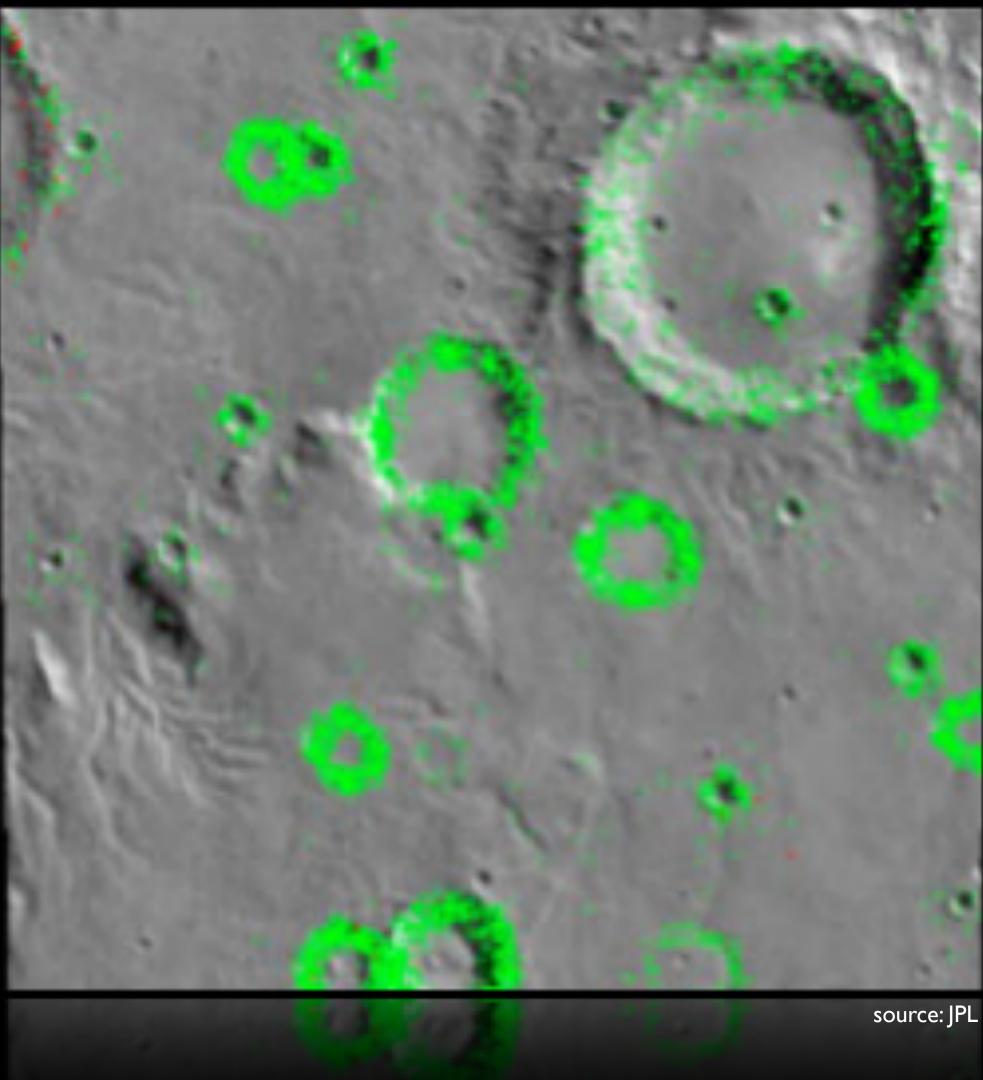


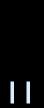


Power of the Crowd crowd as a computational process



Count Craters Clickworkers & Be a Martian





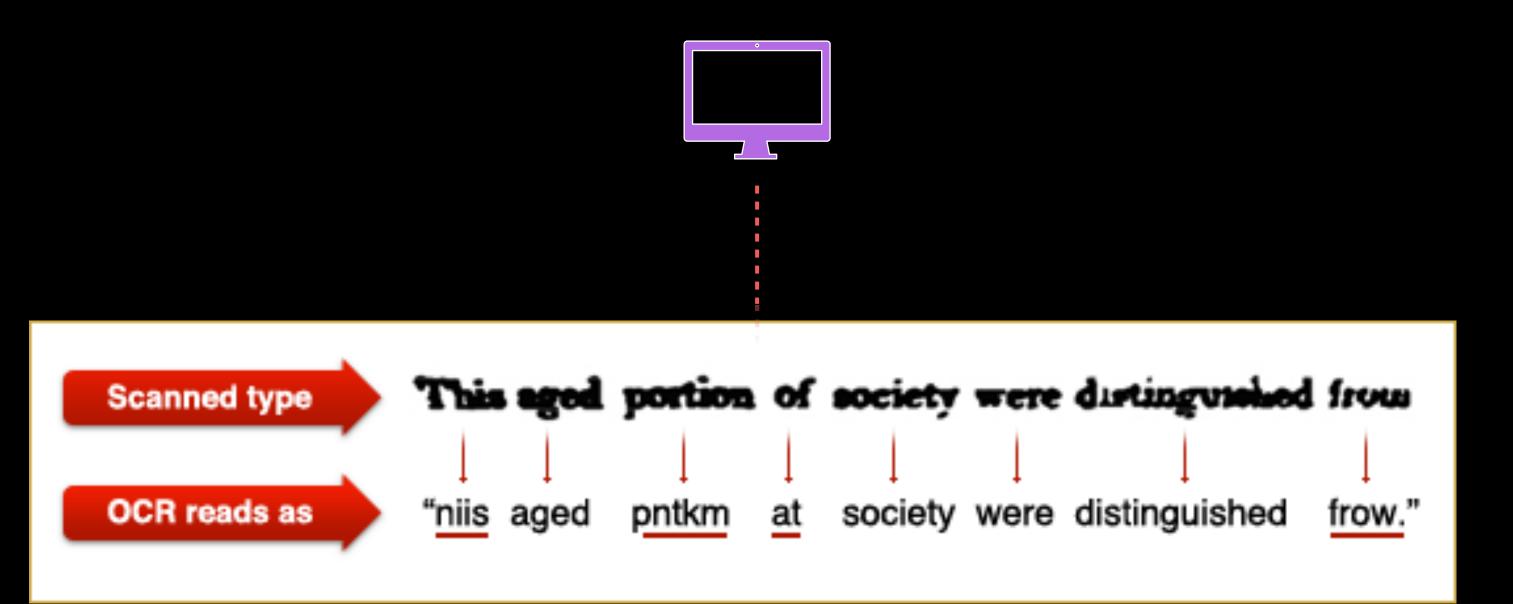
recaptcha



[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
0.2078431	0.1921569	0.1647059	0.1568627	0.1686275
0.2117647	0.1960784	0.1686275	0.1607843	0.1725490
0.2039216	0.1921569	0.1647059	0.1568627	0.1686275
0.1921569	0.1803922	0.1529412	0.1450980	0.1568627
0.1882353	0.1725490	0.1450980	0.1372549	0.1529412
0.1921569	0.1764706	0.1490196	0.1411765	0.1529412
0.2039216	0.1882353	0.1568627	0.1450980	0.1568627
0.2156863	0.2000000	0.1686275	0.1568627	0.1686275
0.2196078	0.2039216	0.1725490	0.1607843	0.1647059
0.2235294	0.2078431	0.1725490	0.1529412	0.1450980
0.2196078	0.2039216	0.1686275	0.1490196	0.1372549
0.2078431	0.1921569	0.1568627	0.1411765	0.1372549
0.2000000	0.1843137	0.1490196	0.1333333	0.1372549

0.1686275 0 0.1843137 0 0.1882353 0.2156863 0.2313725 0.1882353 0.2156863 0.2392157 0.1803922 0.2117647 0.2392157 0.1764706 0.2117647 0.2392157 0.1725490 0.2078431 0.2313725 0.1764706 0.2039216 0.2274510 0.1843137 0.2078431 0.2235294 0.1803922 0.2156863 0.2352941 0.1686275 0.2235294 0.2549020 0.1647059 0.2313725 0.2666667 0.1686275 0.2352941 0.2745098 0.1725490 0.2392157 0.2745098

© NEIL GAIKWAD photography

1 Mar -

© NEIL GAIKWAD photography





0.1921569	0.1803922	0.1529412	0.1411765	0.1529412	0.16
0.2078431	0.1921569	0.1647059	0.1568627	0.1686275	0.18
0.2117647	0.1960784	0.1686275	0.1607843	0.1725490	0.18
0.2039216	0.1921569	0.1647059	0.1568627	0.1686275	0.18
0.1921569	0.1803922	0.1529412	0.1450980	0.1568627	0.18
0.1882353	0.1725490	0.1450980	0.1372549	0.1529412	0.17
0.1921569	0.1764706	0.1490196	0.1411765	0.1529412	0.17
0.2039216	0.1882353	0.1568627	0.1450980	0.1568627	0.17
0.2156863	0.2000000	0.1686275	0.1568627	0.1686275	0.18
0.2196078	0.2039216	0.1725490	0.1607843	0.1647059	0.18
0.2235294	0.2078431	0.1725490	0.1529412	0.1450980	0.16
0.2196078	0.2039216	0.1686275	0.1490196	0.1372549	0.16
0.2078431	0.1921569	0.1568627	0.1411765	0.1372549	0.16
0.200000	0.1843137	0.1490196	0.1333333	0.1372549	0.17

© NEIL GAIKWAD photography

686275 0.1882353 0.2039216 843137 0.2039216 0.2196078 882353 0.2156863 0.2313725 882353 0.2156863 0.2392157 803922 0.2117647 0.2392157 764706 0.2117647 0.2392157 725490 0.2078431 0.2313725 764706 0.2039216 0.2274510 843137 0.2078431 0.2235294 803922 0.2156863 0.2352941 686275 0.2235294 0.2549020 647059 0.2313725 0.2666667 686275 0.2352941 0.2745098 725490 0.2392157 0.2745098

[Von Ahn et al. *CHI*, 2004] ¹⁴





0 1021560	0.1803922	0 1520/12	0 1411765	0 1529/12	0 16
0.1921309	0.1003922	0.1029412	0.1411/00	0.1029412	0.10
0.2078431	0.1921569	0.1647059	0.1568627	0.1686275	0.18
0.2117647	0.1960784	0.1686275	0.1607843	0.1725490	0.18
0.2039216	0.1921569	0.1647059	0.1568627	0.1686275	0.18
0.1921569	0.1803922	0.1529412	0.1450980	0.1568627	0.18
0.1882353	0.1725490	0.1450980	0.1372549	0.1529412	0.17
0.1921569	0.1764706	0.1490196	0.1411765	0.1529412	0.17
0.2039216	0.1882353	0.1568627	0.1450980	0.1568627	0.17
0.2156863	0.2000000	0.1686275	0.1568627	0.1686275	0.18
0.2196078	0.2039216	0.1725490	0.1607843	0.1647059	0.18
0.2235294	0.2078431	0.1725490	0.1529412	0.1450980	0.16
0.2196078	0.2039216	0.1686275	0.1490196	0.1372549	0.16
0.2078431	0.1921569	0.1568627	0.1411765	0.1372549	0.16
0.200000	0.1843137	0.1490196	0.1333333	0.1372549	0.17

© NEIL GAIKWAD photography

686275 0.1882353 0.2039216 843137 0.2039216 0.2196078 882353 0.2156863 0.2313725 882353 0.2156863 0.2392157 803922 0.2117647 0.2392157 764706 0.2117647 0.2392157 725490 0.2078431 0.2313725 764706 0.2039216 0.2274510 843137 0.2078431 0.2235294 803922 0.2156863 0.2352941 686275 0.2235294 0.2549020 647059 0.2313725 0.2666667 686275 0.2352941 0.2745098 725490 0.2392157 0.2745098

[Von Ahn et al. *CHI*, 2004] ¹⁴



New York



0.1921569	0.1803922	0.1529412	0.1411765	0.1529412	0.16
0.2078431	0.1921569	0.1647059	0.1568627	0.1686275	0.18
0.2117647	0.1960784	0.1686275	0.1607843	0.1725490	0.18
0.2039216	0.1921569	0.1647059	0.1568627	0.1686275	0.18
0.1921569	0.1803922	0.1529412	0.1450980	0.1568627	0.18
0.1882353	0.1725490	0.1450980	0.1372549	0.1529412	0.17
0.1921569	0.1764706	0.1490196	0.1411765	0.1529412	0.17
0.2039216	0.1882353	0.1568627	0.1450980	0.1568627	0.17
0.2156863	0.2000000	0.1686275	0.1568627	0.1686275	0.18
0.2196078	0.2039216	0.1725490	0.1607843	0.1647059	0.18
0.2235294	0.2078431	0.1725490	0.1529412	0.1450980	0.16
0.2196078	0.2039216	0.1686275	0.1490196	0.1372549	0.16
0.2078431	0.1921569	0.1568627	0.1411765	0.1372549	0.16
0.2000000	0.1843137	0.1490196	0.1333333	0.1372549	0.17

686275 0.1882353 0.2039216 843137 0.2039216 0.2196078 882353 0.2156863 0.2313725 882353 0.2156863 0.2392157 803922 0.2117647 0.2392157 764706 0.2117647 0.2392157 725490 0.2078431 0.2313725 764706 0.2039216 0.2274510 843137 0.2078431 0.2235294 803922 0.2156863 0.2352941 686275 0.2235294 0.2549020 647059 0.2313725 0.2666667 686275 0.2352941 0.2745098 725490 0.2392157 0.2745098

[Von Ahn et al. *CHI*, 2004] ¹⁴

© NEIL GAIKWAD photography



New York



0.1921569	0.1803922	0.1529412	0.1411765	0.1529412	0.16
0.2078431	0.1921569	0.1647059	0.1568627	0.1686275	0.18
0.2117647	0.1960784	0.1686275	0.1607843	0.1725490	0.18
0.2039216	0.1921569	0.1647059	0.1568627	0.1686275	0.18
0.1921569	0.1803922	0.1529412	0.1450980	0.1568627	0.18
0.1882353	0.1725490	0.1450980	0.1372549	0.1529412	0.17
0.1921569	0.1764706	0.1490196	0.1411765	0.1529412	0.17
0.2039216	0.1882353	0.1568627	0.1450980	0.1568627	0.17
0.2156863	0.2000000	0.1686275	0.1568627	0.1686275	0.18
0.2196078	0.2039216	0.1725490	0.1607843	0.1647059	0.18
0.2235294	0.2078431	0.1725490	0.1529412	0.1450980	0.16
0.2196078	0.2039216	0.1686275	0.1490196	0.1372549	0.16
0.2078431	0.1921569	0.1568627	0.1411765	0.1372549	0.16
0.2000000	0.1843137	0.1490196	0.1333333	0.1372549	0.17

New York

© NEIL GAIKWAD photography

686275 0.1882353 0.2039216 843137 0.2039216 0.2196078 882353 0.2156863 0.2313725 882353 0.2156863 0.2392157 803922 0.2117647 0.2392157 764706 0.2117647 0.2392157 725490 0.2078431 0.2313725 764706 0.2039216 0.2274510 843137 0.2078431 0.2235294 803922 0.2156863 0.2352941 686275 0.2235294 0.2549020 647059 0.2313725 0.2666667 686275 0.2352941 0.2745098 725490 0.2392157 0.2745098

New York

[Von Ahn et al. *CHI*, 2004] ¹⁴



New York



0.1921569	0.1803922	0.1529412	0.1411765	0.1529412	0.16
0.2078431	0.1921569	0.1647059	0.1568627	0.1686275	0.18
0.2117647	0.1960784	0.1686275	0.1607843	0.1725490	0.18
0.2039216	0.1921569	0.1647059	0.1568627	0.1686275	0.18
0.1921569	0.1803922	0.1529412	0.1450980	0.1568627	0.18
0.1882353	0.1725490	0.1450980	0.1372549	0.1529412	0.17
0.1921569	0.1764706	0.1490196	0.1411765	0.1529412	0.17
0.2039216	0.1882353	0.1568627	0.1450980	0.1568627	0.17
0.2156863	0.2000000	0.1686275	0.1568627	0.1686275	0.18
0.2196078	0.2039216	0.1725490	0.1607843	0.1647059	0.18
0.2235294	0.2078431	0.1725490	0.1529412	0.1450980	0.16
0.2196078	0.2039216	0.1686275	0.1490196	0.1372549	0.16
0.2078431	0.1921569	0.1568627	0.1411765	0.1372549	0.16
0.2000000	0.1843137	0.1490196	0.1333333	0.1372549	0.17







Micro Tasks low cognitive overload



Crowdcomputing and Citizen Science for Large-scale Experiments

15 img-source: wired

Metacognition

Creativity

Motivation







Crowdcomputing and Citizen Science for Large-scale Experiments

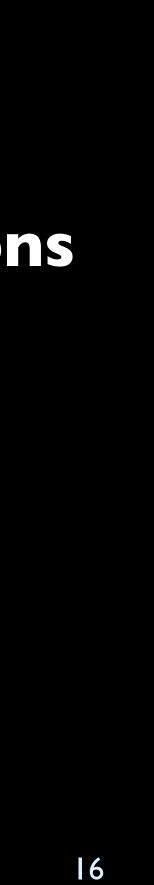


Social Learning



Intuitions

Collaboration



img-source: wired

Macro Tasks high cognitive overload



Thinking at computational and cognitive scale





delving deeper



Crowdcomputing and Citizen Science for Large-scale Experiments

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



Crowdcomputing and Citizen Science for Large-scale Experiments

The Internet as our laboratory and a crowd as collaborator



Research and Solve

the biggest unanswered questions



Crowdcomputing and Citizen Science for Large-scale Experiments

Exploring the life at the cellular level



Understanding RNAs





Dark matter of Biology

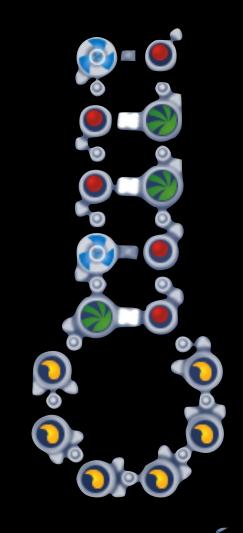
Key to understand life at the cellular level

Manipulate RNAs to help cure Parkinson's disease



Discover the sequence





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

Hairpin RNA Target Shape

Crowdcomputing and Citizen Science for Large-scale Experiments

- GGAGGGUUGAUACGAU
- C
 G
 C
 U

 U
 G
 A
 A
- U •

Nucleotide Sequence



Summer, Gain a New Perspective!

Enroll in an VJ-SCPS ummer itensive

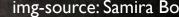
ase your m

gister now:

it scps.nyu.edu/sun call 202-998-7150

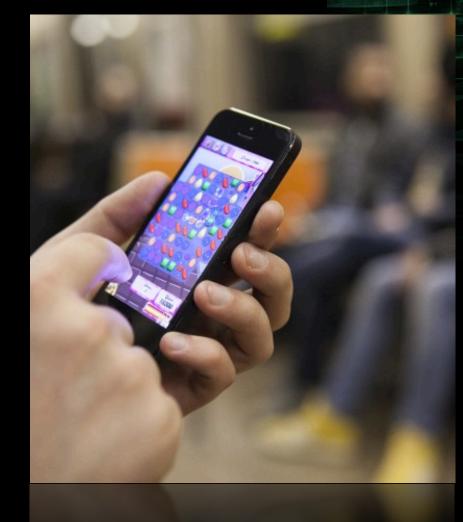


गातामु रू डोकरे रू कह भरखे-खोमाट मुक्त झोडे,





The power of games



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

Crowdcomputing and Citizen Science for Large-scale Experiments

SCI-FI

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

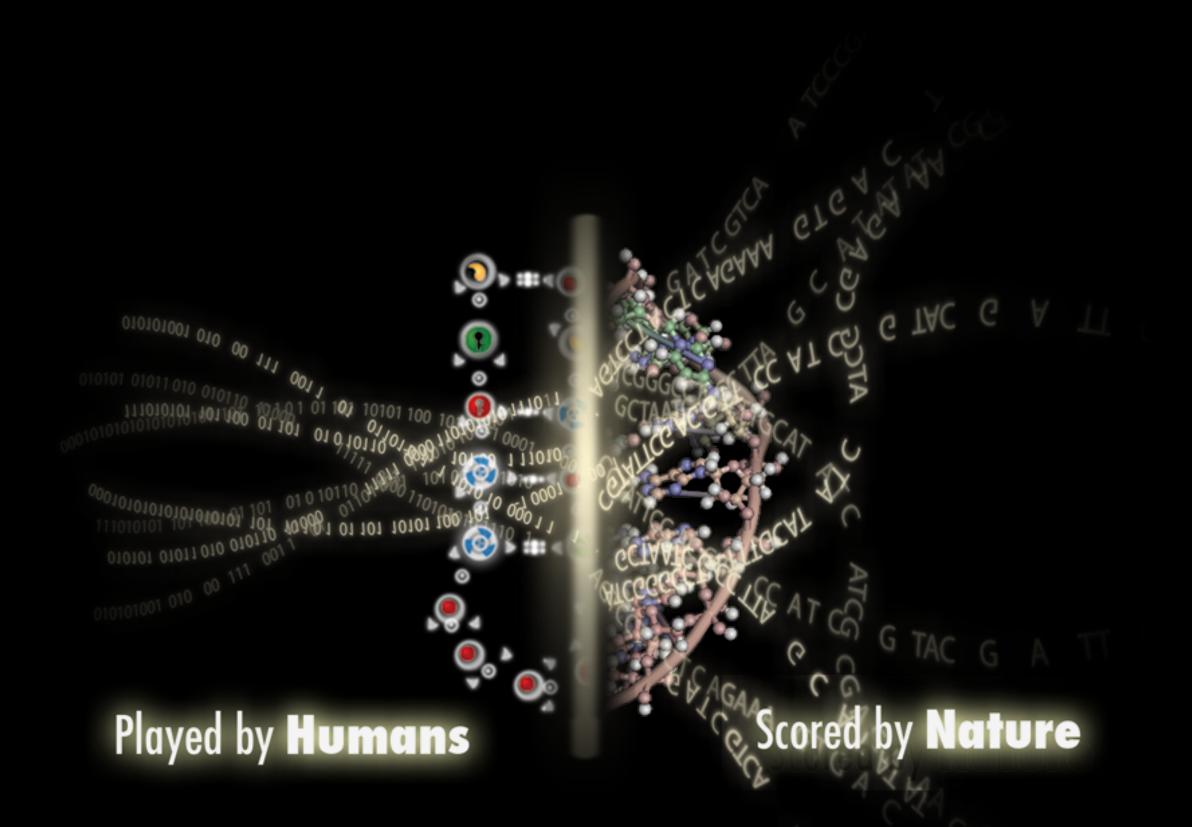


Crowdcomputing and Citizen Science for Large-scale Experiments

Games for democratizing science







RNA Nano Engineering Over 100,000 users from more than 90 Countries

Crowdcomputing and Citizen Science for Large-scale Experiments

[Lee et al. PNAS, 2014]



Pushing the Boundaries of Quantum Physics



Quantum Control Optima

Search for high dimensional complex challenges

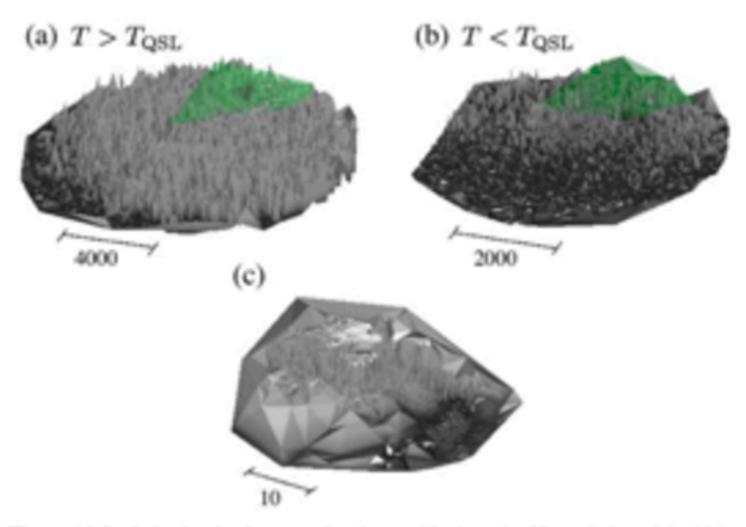


Figure 4 | Optimization landscapes. Panels a and b show the 2D rendering of the highdimensional optimization landscape for process durations T = 0.40 and T = 0.17, respectively. Green areas mark the space probed by CHOP solutions. Panel b is the low-dimensional HILO landscape.

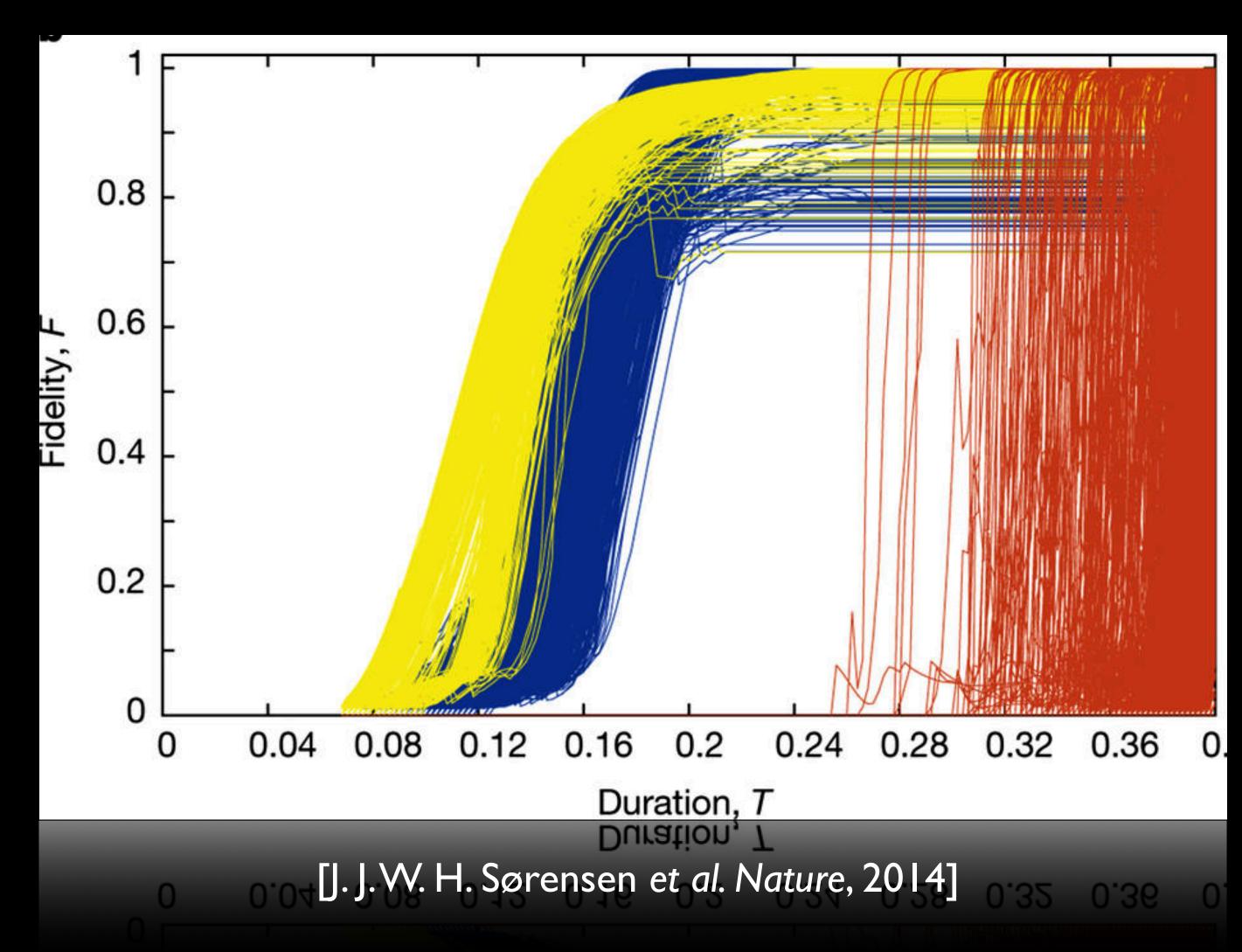




ю



Science@Home

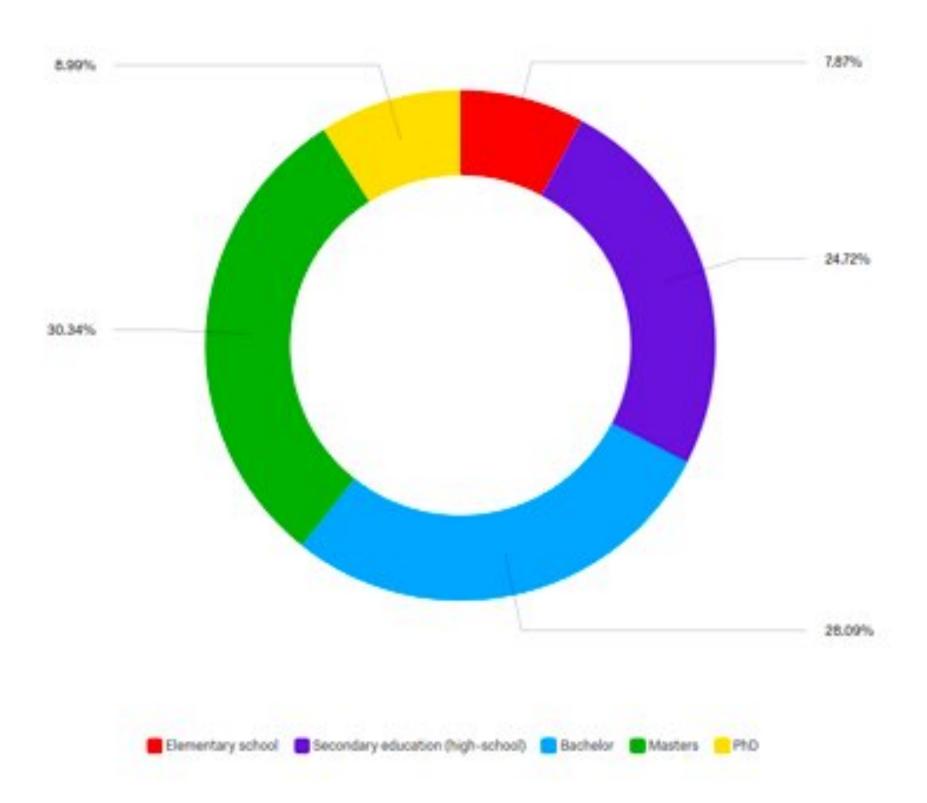






Players

Highest education level graduated



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

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



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



Democratizing research to increased diversity in the scientific workforce

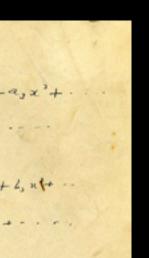


Scientific research remains the domain of the privileged fewto those blessed with the socioeconomic opportunities

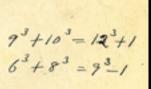


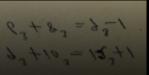
The man who knew oo

 $\frac{1+53x+9x^{2}}{1-81x-88x^{2}+x^{2}} = a_{0} + a_{1}x + a_{2}x^{2} + a_{3}x^{3} + .$ Bruce C. Berndt Ramanujan's or do + d/ + du + ---Notebooks 2-262-122 = 60+6,2+622+4, x+ Part I or he + B, + Bu + $\frac{2 + Px - 10x^{2}}{1 - 8Lx - 8Lx^{2} + x^{3}} = c_{0} + c_{1}x + c_{2}x^{2} + c_{3}x^{3} + \frac{1}{2}x^{2} + \frac{$ or to + the + the + Springer-Verleg New York: Berlin: Heidelberg Tokyo PK+FASTL-R an + 4n = cn + (-1) ~ 3. = 11(+71+3+20 RETABLACONNE 76-10=0 las = 244 374 941 , 19.161 $d_{1}a_{1}^{3} + P_{1}^{3} = f_{1}^{3} + (-1)^{5}$ - 80 cat ill 200 40500 of th BH = EMP3 107 - M(1939 T = 355(1- 9003). 46366091+131266 -496772981-11+548165 to all appropriate miners it - the start and many ples the martine to the second second for about - tabl 135"+ 138" = 172"-1 10105000 和-11-11日本 (会・1) usycate (100) - (massing + 111613 + 114683 = 142583+1 16322 1913 + 8123 = 10103-1 あけますで あかい (み)+で かかい (み) + もいち = 1010-1 10122104/1817 Then an it + e " carge e alcasore is 1+ 3-5671 = 39711-+ 10 oper (" hery + " " heary + " 「(ころかな(祭)+この部に(祭)-4 Trop of to + 0 " Sorry set married They ar a the sy to an good sta 4 - 4 - AC 2. 215715 574177 1+5- 417 1+ 4 + + 14 1 + 1 + m ゆ・いいり いー 学を - いいやういう











Srinivasa Ramanujan [1887-1920]

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



Stanford Crowd Research

Crowdcomputing and Citizen Science for Large-scale Experiments

[Vaish et al. ACM UIST, 2017]









[Gaikwad et al. ACM UIST, 2015 2016]



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.



Crowdcomputing and Citizen Science for Large-scale Experiments

Forum



DAEMO

Easier and more equitable crowdsourcing

Stanford University

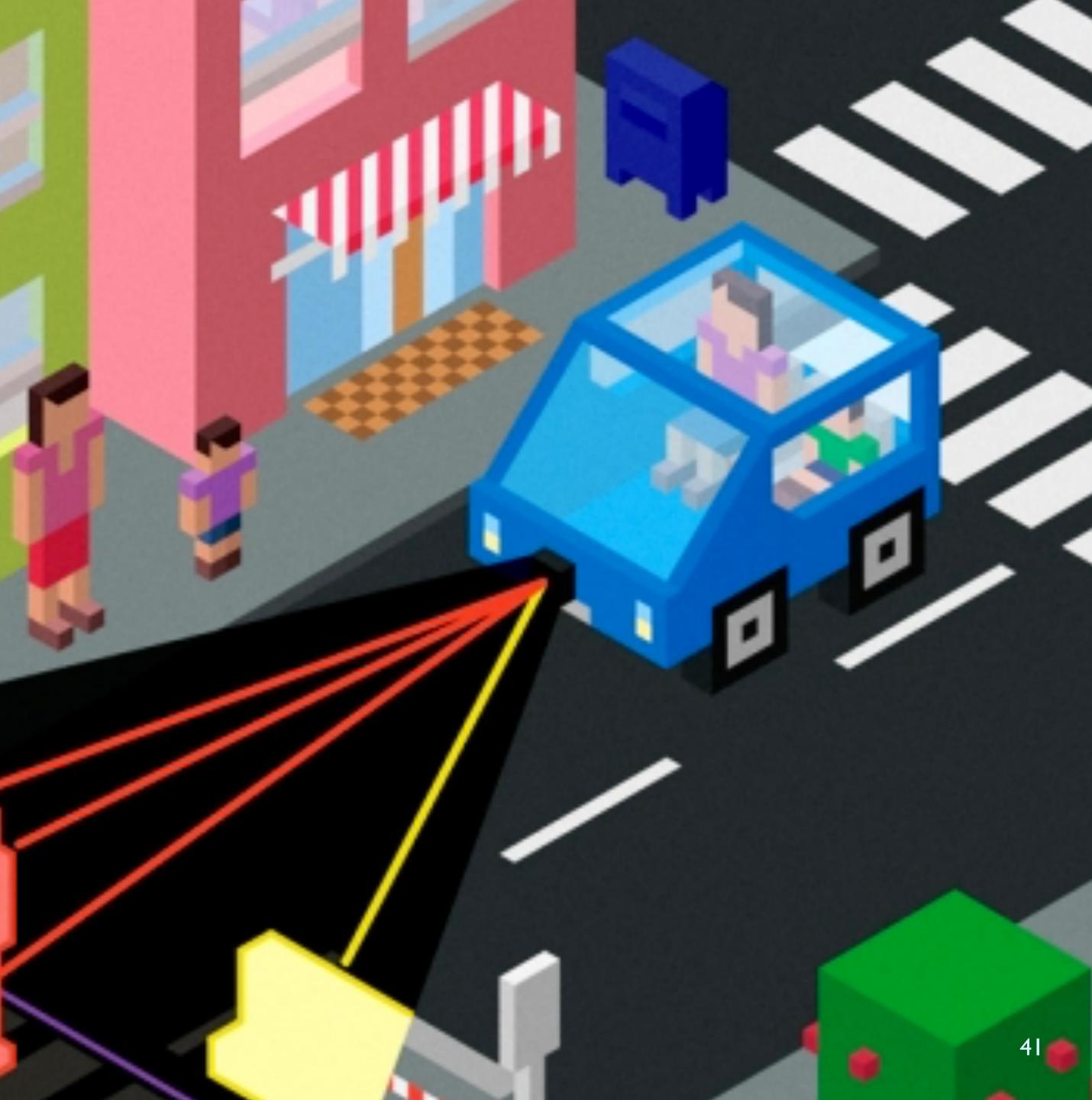




Researching Al ethics and algorithmic biases at scale

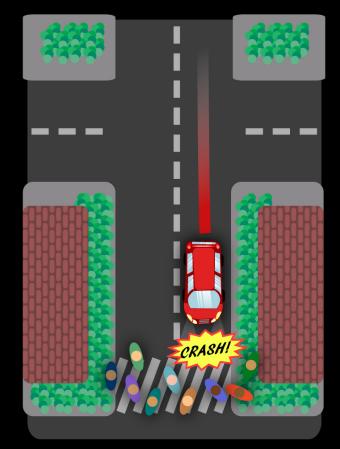
40

[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







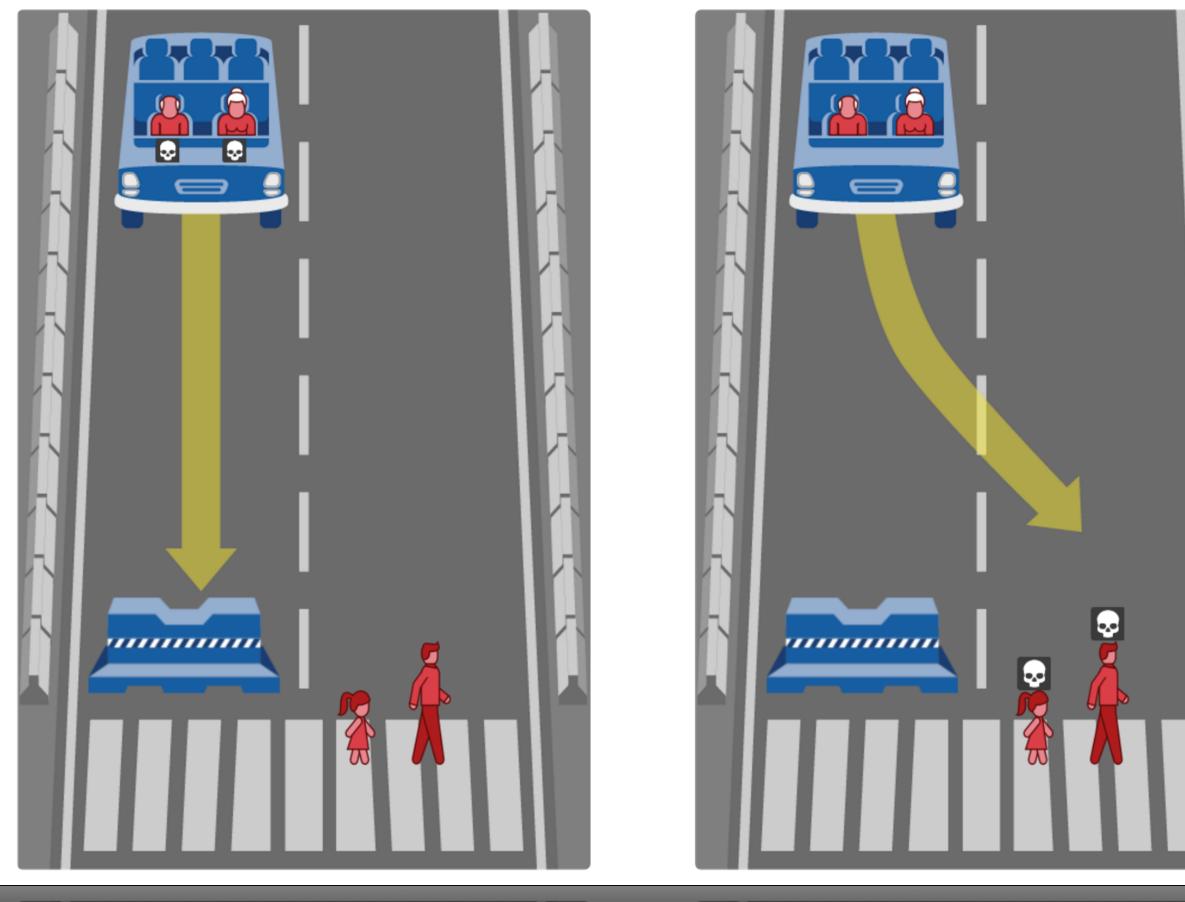


42

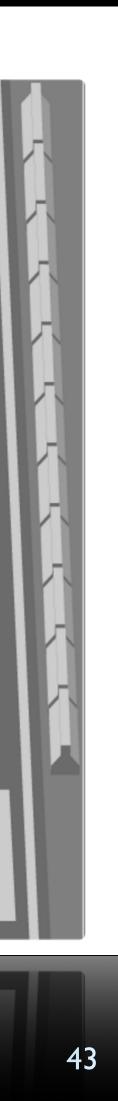
Moral Machine

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

What should the self-driving car do?



[www.moralmachine.mit.edu/]



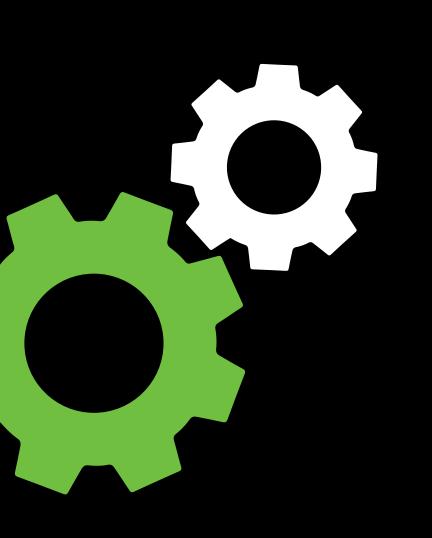
CASE Let's choose a problem

Crowdcomputing and Citizen Science for Large-scale Experiments



44

Crowd Mechanics



45

Crowd Mechanics

DYNAMICS & LOGISTICS

EXPERIMENT DESIGN & DATA SCIENCE

PLATFORM & INFRASTRUCTURE DESIGN

Crowdcomputing and Citizen Science for Large-scale Experiments

solve

46

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

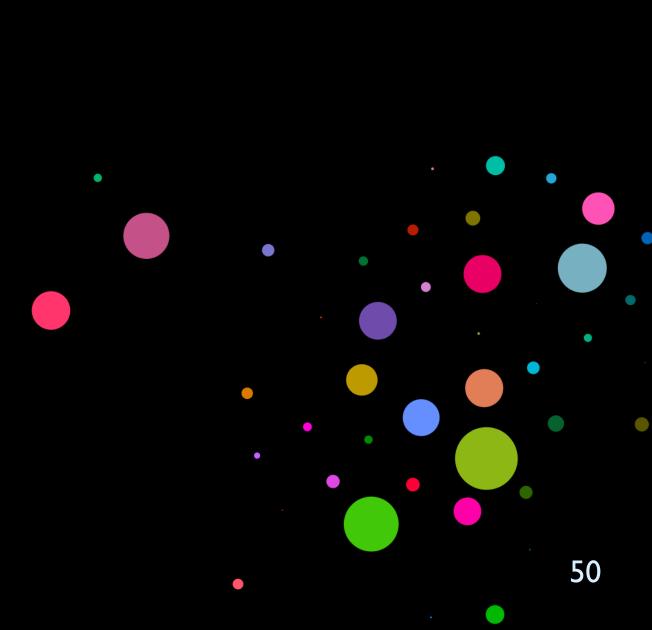
48

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

nature Home News & Comment Research Careers & Jobs Current Issue Archive Audio & Video Archive Volume 532 Issue 7598 News Article **NATURE | NEWS**

عربي

Human mind excels at quantum-physics computer game

Revelation could have implications for how scientists approach quantum problems

Revelation could have implications for how scientists approach quantum problems

game

The New Hork Eimes

Science

WORLD	U.S.	N.Y. / REGION	BUSINESS	TECHNOLOGY	SCIENCE	HEALTH	5		
					ENVIRONMENT				

RNA Game Lets Players Help Find a Biological Prize

RNA Game Lets Players Help Find a Biological Prize

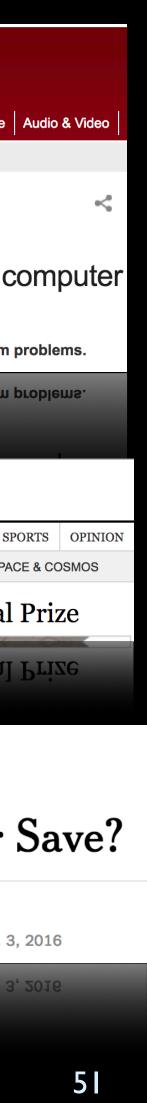
SundayReview

Whose Life Should Your Car Save?

Gray Matter

By AZIM SHARIFF, IYAD RAHWAN and JEAN-FRANCOIS BONNEFON NOV. 3, 2016

By AZIM SHARIFF, IYAD RAHWAN and JEAN-FRANÇOIS BONNEFON NOV. 3, 2016 **Gray Matter**



Retention & Incentive Engineering

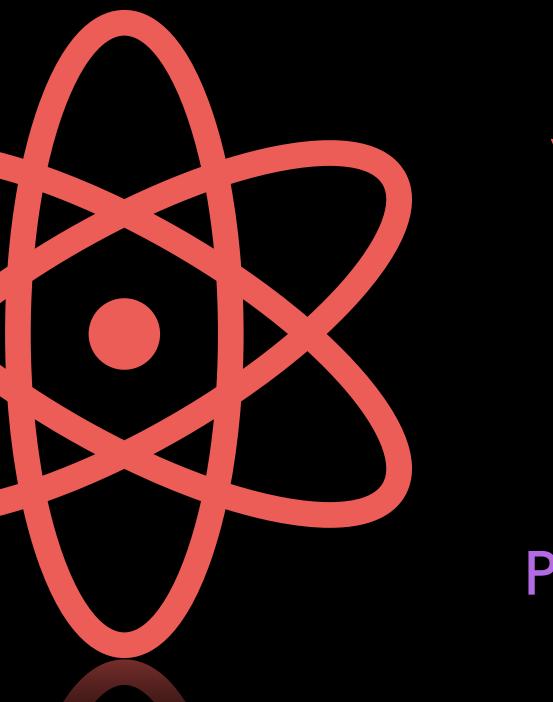
Competition



Motivations: Intrinsic, Extrinsic

Crowdcomputing and Citizen Science for Large-scale Experiments

Reputation



Vanity

Purpose

Virtual Market



52

Bonuses: Incentives and motivation

127

59

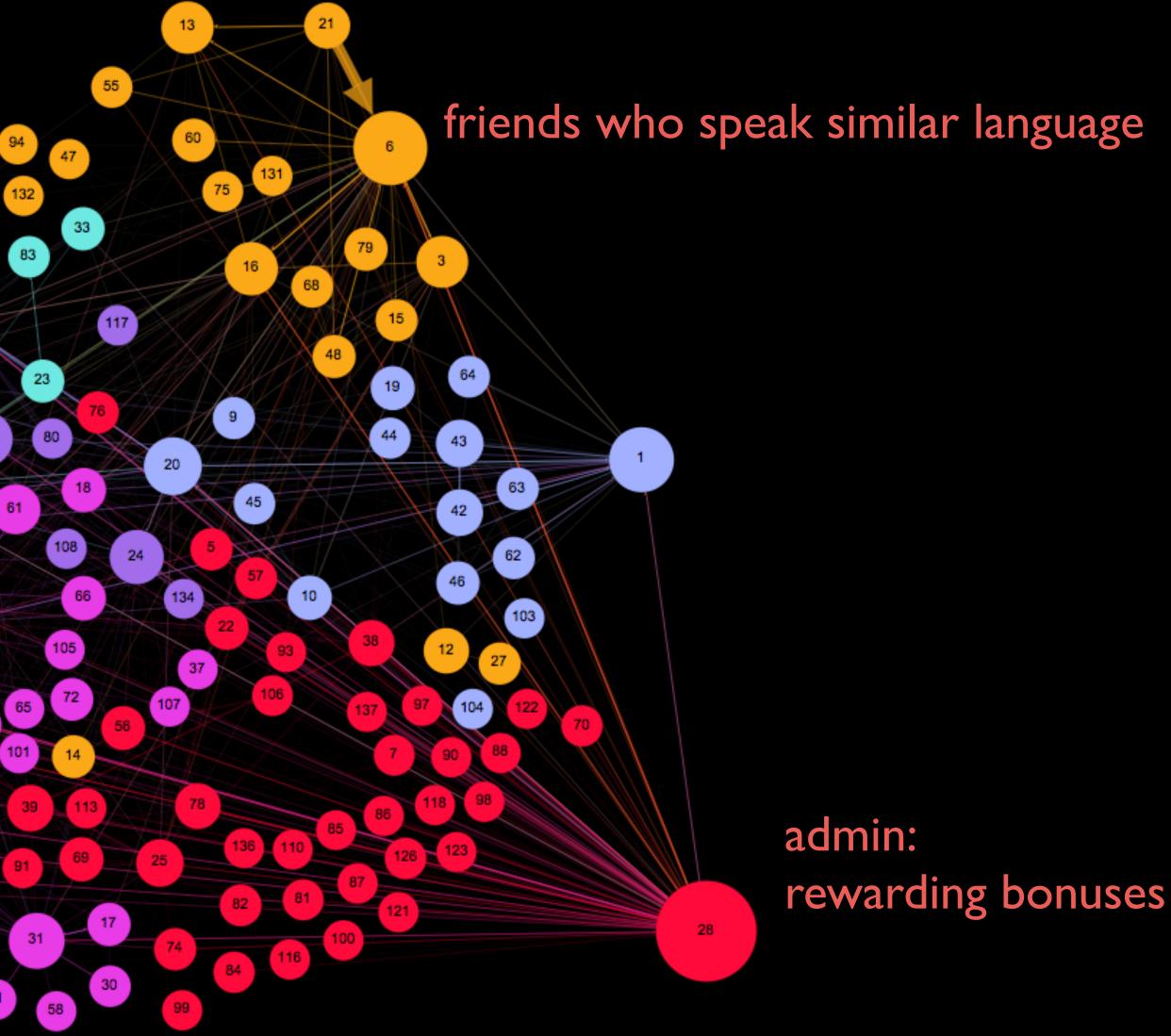
73

109

52

Periodic bonuses

lead contributor: rewarding bonuses











Organizing the Crowd Community, Task Distribution, Recognition



Community

Task Distributions

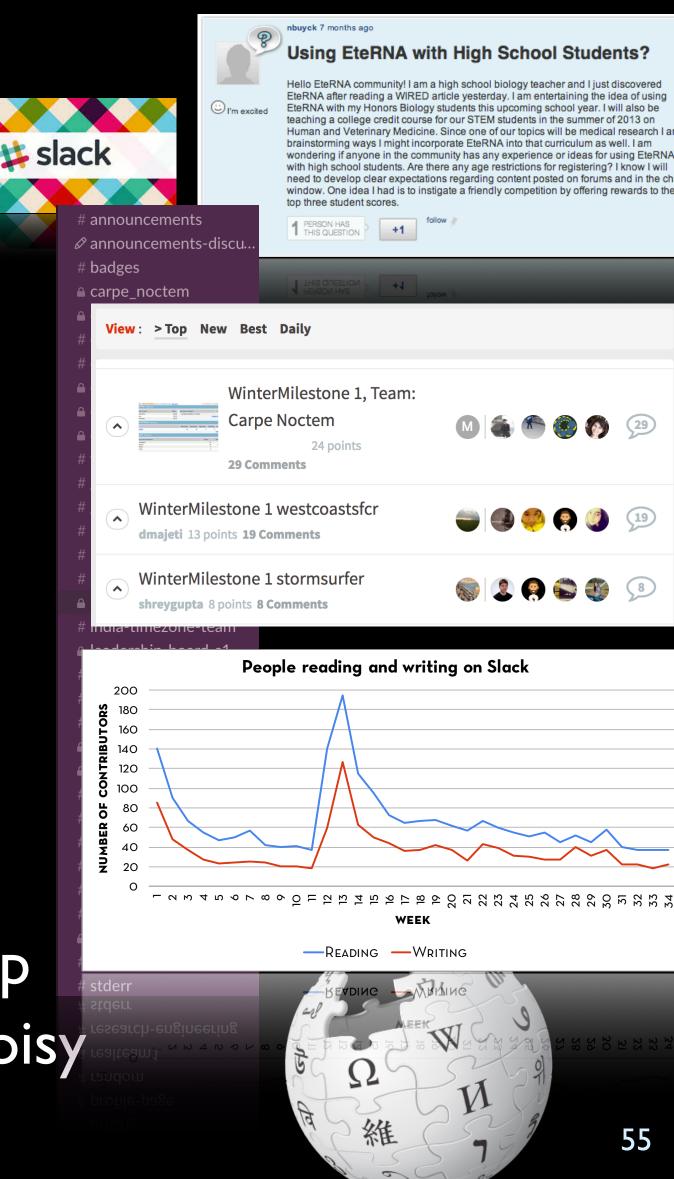
- Various paradigm (e.g., Parallel vs Sequential)
- Milestones and feedback

Engagement Channels

- Asynchronous Communication: Forums, Emails, Wikipedia
- Synchronous Communication: Slack, Hangouts, Chats

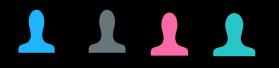
• Protocols

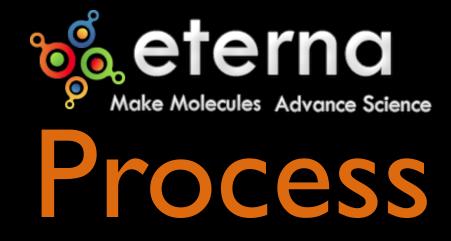
- Inclusive and constructive environment
- Empathy Volunteers' (workers') and requesters' relationship
- Remember that the crowd is comprised of people, not "noisy error-prone computers"



(29) 🛃 🦪 🖪 🚺 🔅 🗐 🌍 🙎 🤗 🚔 🧐 👂





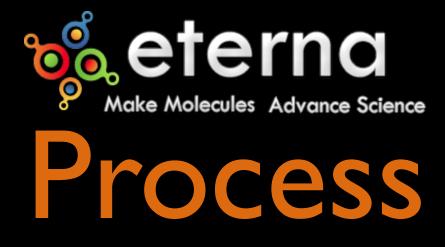










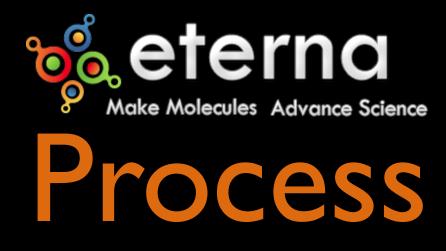










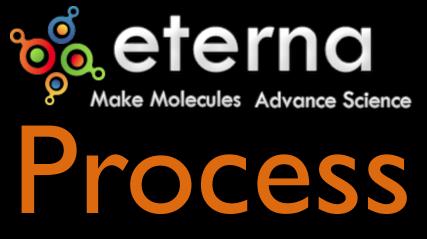








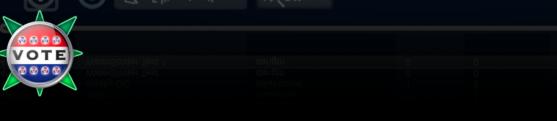








ld	Title	Designer	Votes	My Votes	Description	Chat Players Online
min max	search	search	min max	min max	search	
						happy? [11:10 AM] kenziefae: im always ha
	Barely works	FormalRiceFarmer	0	0	No comment	but mainly because i m 50 dollars last night [11:1
536889	NUPACK design 06	NUPACK	-	-	Normalized ensemble defe	EpicShorts: How? [11:11
536887	NUPACK design 05	NUPACK			Normalized ensemble defe	kenzlefae: babysittin :D
536886	NUPACK design 04	NUPACK			Normalized ensemble defe	EpicShorts: Oh [11:12 AM]
535887	ViennaRNA design 04	ViennaRNA			Designed by ViennaRNA a	kenziefae: so wuts you
535885	ViennaRNA design 03	ViennaRNA		•	Designed by ViennaRNA a	:D [11:12 AM] EpicShorts: 1251 [11:12 A
535884	ViennaRNA design 02	ViennaRNA			Designed by ViennaRNA a	kenzlefae: awesome :D
535883	ViennaRNA design 01	ViennaRNA			Designed by ViennaRNA a	only in like the 6000 :D
496857	leaner than sup	tmcannon	1	0	without the extras	EpicShorts: Nice [11:16 A
496331	Low GC	Quasispecies	1	0	Clollin's Superbug #3 was	kenzlefae: sure i guess
496052	Mod of Starrys	Eli Fisker	1	0	Swapped basepair the righ	AM]
496050	Mod of Wisdaves	Eli Fisker	1	0	Swapped two pairs in the	
496046	Mod of WS Brord	Eli Fisker	2	0	Swapped three basepairs	
495898	Simple mod of W	Brourd	2	0	First submission for roun	
495733	Another Take	Rafael Gribeler Tschope	0	0	Took clollin's solution a	
494795	unique sequence	edwintorok	1	0	GC 66%, AU 33% each lo	You have 8 votes left.
494791	varied	edwintorok	1	0	GC 79%, AU pairs, no GU	You have 3 solution slots l
494656	WS-1	nascarnut	0	0	First Try	Tou have a solution slots in
494501	mostly GC	edwintorok	0	0	Modified from the 'all GC	
494357 494353	WaterStrider Test WaterStrider Test 1	dangtn	0	0	Playing with parameters	
494303	WaterStrider Test 1	dangtn	U	U	Preoptimization for melt	
					•	
	🖌 🕜 🗹 Edit sort optic	ons tt Views				
		DIS TI VIEWS				



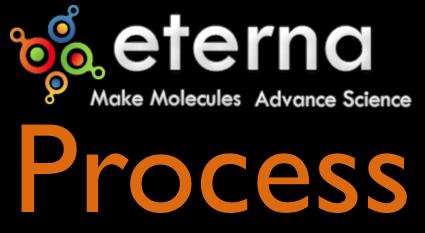












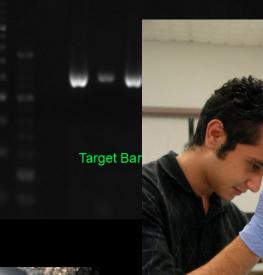




nin max	search	search	min max	<u> </u>		Chat Players On
			min max	min max	search	honny?
						happy? [11:10 AM] kenziefae: im always but mainly because i
	Barely works	FormalRiceFarmer	0	0	No comment	but mainly because i 50 dollars last night [
536889	NUPACK design 06	NUPACK	U I	-	Normalized ensemble defe	EpicShorts: How? It
536887	NUPACK design 05	NUPACK			Normalized ensemble defe	kenzlefae: babysittin
536886	NUPACK design 04	NUPACK			Normalized ensemble defe	EpicShorts: Oh [11:12
535887	ViennaRNA design 04	ViennaRNA			Designed by ViennaRNA a	kenziefae: so wuts y
535885	ViennaRNA design 04 ViennaRNA design 03	ViennaRNA			Designed by ViennaRNA a	:D [11:12 AM]
535884	ViennaRNA design 03	ViennaRNA			Designed by ViennaRNA a	EpicShorts: 1251 [11
535883	ViennaRNA design 02 ViennaRNA design 01	ViennaRNA			Designed by ViennaRNA a	kenziefae: awesome
196857				- 0	without the extras	only in like the 6000
196331	leaner than sup Low GC	tmcannon Quasispecies		0	Clollin's Superbug #3 was	EpicShorts: Nice [11: kenziefae: sure i gu
196052	Mod of Starrys	Eli Fisker		0	Swapped basepair the right	AM]
196052 196050	Mod of Wisdaves	Eli Fisker		0	Swapped basepair the hgr Swapped two pairs in the	
196046	Mod of WS Brord	Eli Fisker	2	0	Swapped two pairs in the . Swapped three basepairs	
195898			2	0	First submission for roun	
195733	Simple mod of W	Brourd		0		
	Another Take	Rafael Gribeler Tschope	0	-	Took clollin's solution a	
194795	unique sequence	edwintorok		0	GC 66%, AU 33% each lo	You have 8 votes left.
94791	varied	edwintorok	1	0	GC 79%, AU pairs, no GU	You have 3 solution slo
194656	WS-1	nascarnut	0	0	First Try	Tou have a solution sit
194501	mostly GC	edwintorok	1	0	Modified from the 'all GC	
194357	WaterStrider Test	dangtn	0	0	Playing with parameters	
194353	WaterStrider Test 1	dangtn	0	0	Preoptimization for melt	



DScore x/100



L20 L10 A B C D E F G H



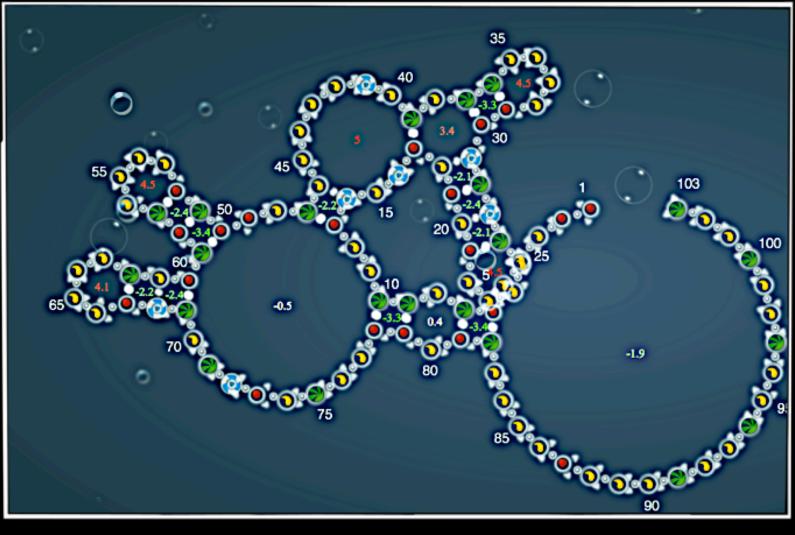


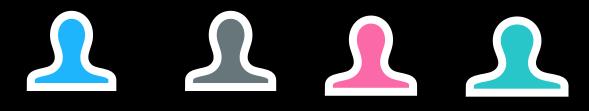






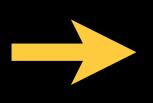








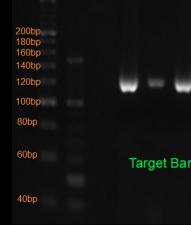






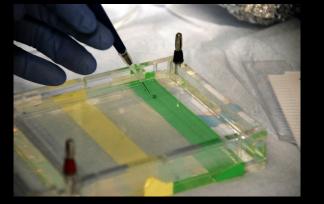


DScore x/100



L20 L10 A B C D E F G H











Recognition

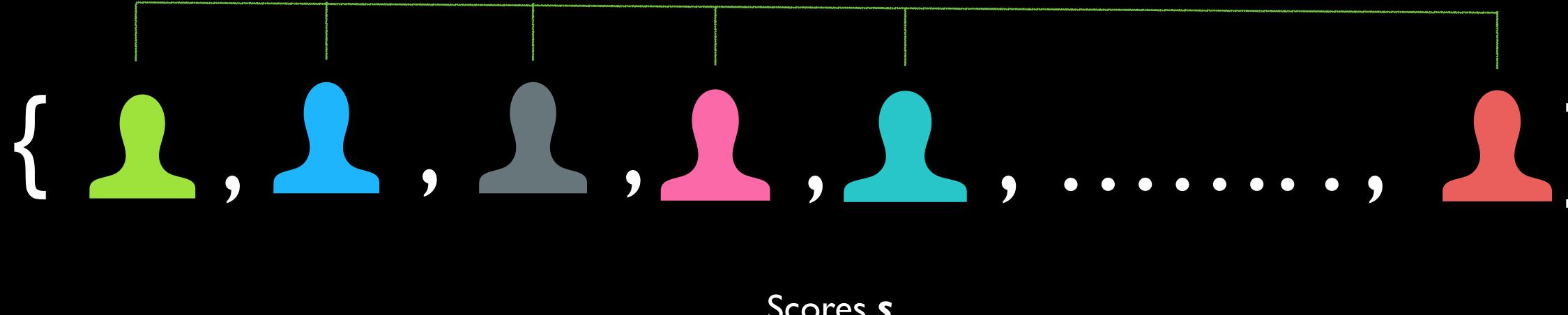
- Admins cannot have a full picture of who is doing what
- Most of the voting rules are confusing and can be gamed
- Voluntary projects how to spot HUBS in a large network and empower them



Nodes in Credit Distribution



Links in Credit Distribution



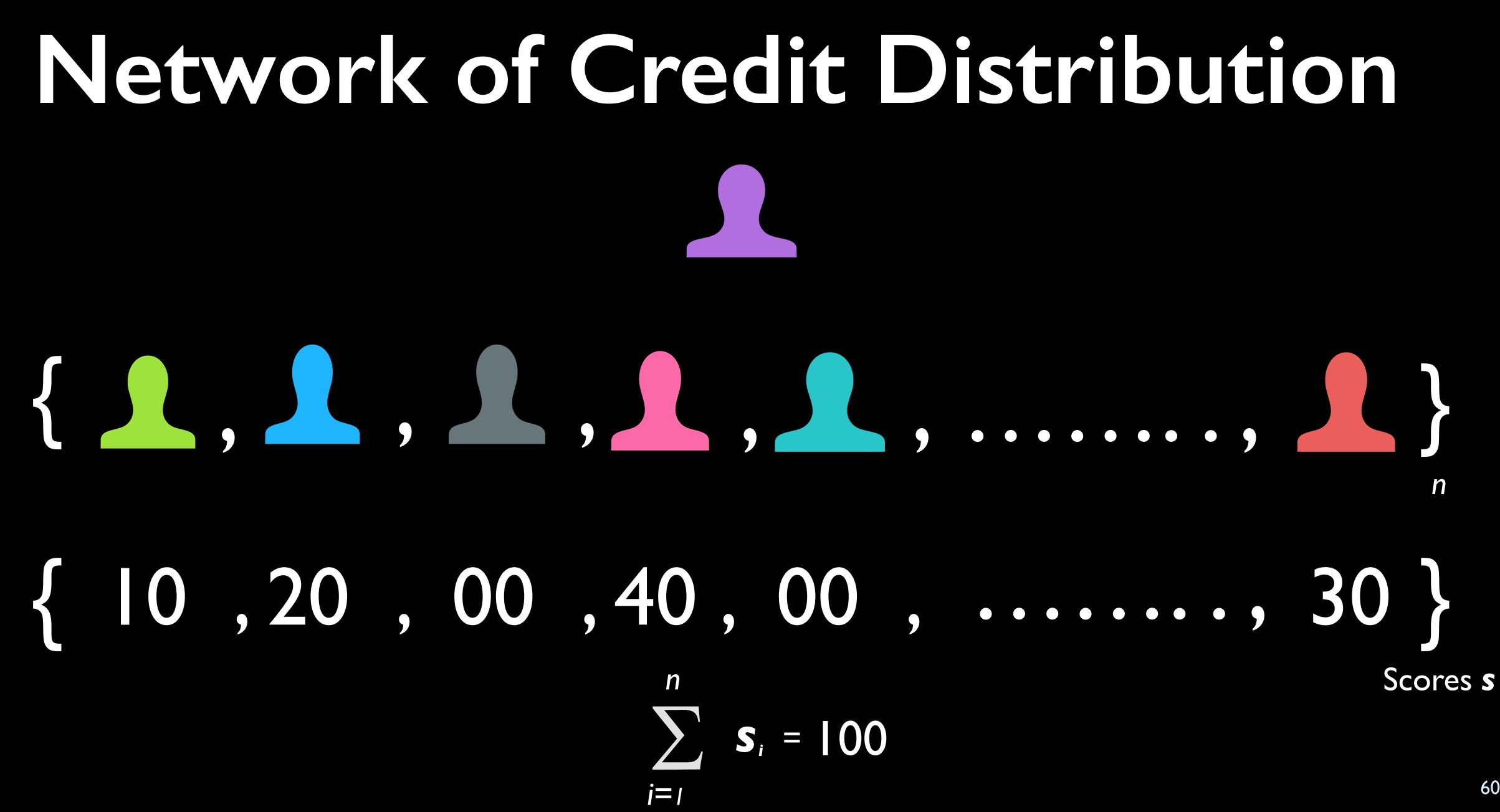
Crowdcomputing and Citizen Science for Large-scale Experiments

Scores s

 $\{10, 20, 00, 40, 00, \dots, 30\}$

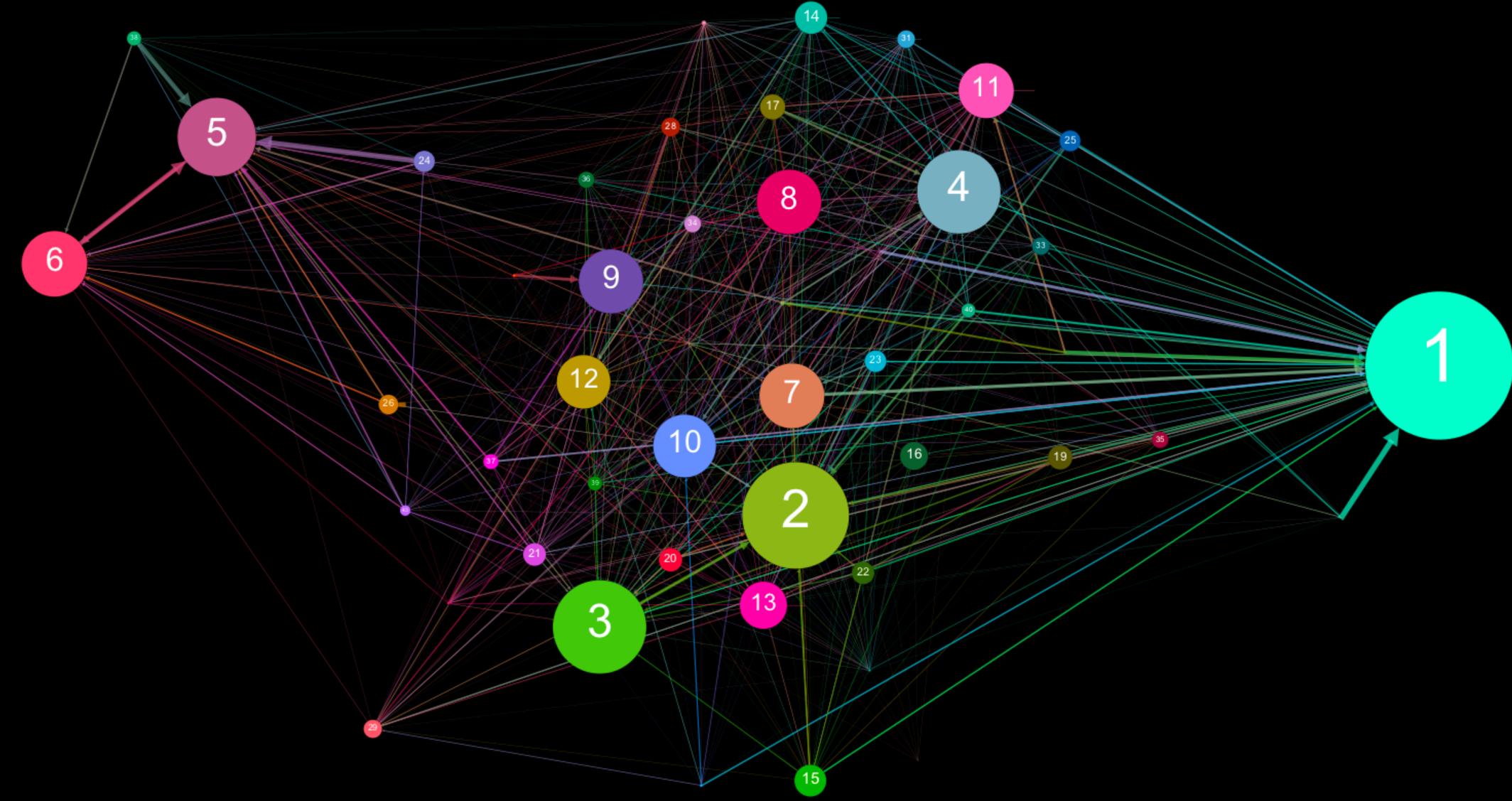


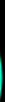
n

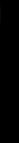


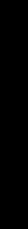


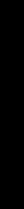
Author ranks via page rank

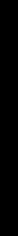


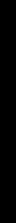


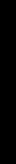














CASE The Problem and Crowd Dynamics

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 D
 - Organic— crowd filters out bad designs or submissions
 - spam, or sabotage

• 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,



Moral Machine - data expansion

Internationalization (10 languages)

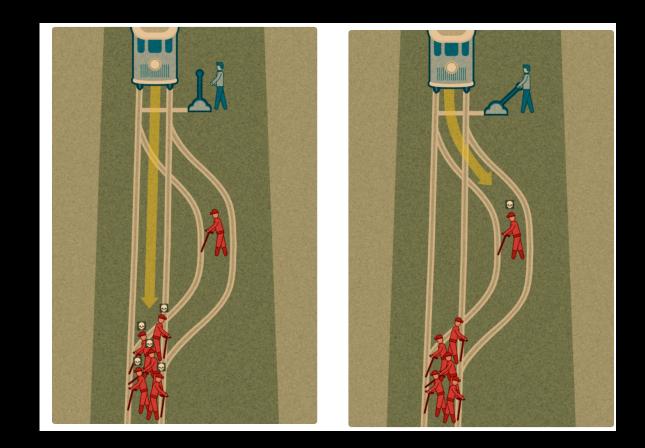
Addition of post-quiz user survey

Expansion of experiment scope with classic scenarios



Crowdcomputing and Citizen Science for Large-scale Experiments





😂 More < Share 🖸 Link Would you like to help us better understand your judgement? Yes



30M Decisions

350k Full Surveys

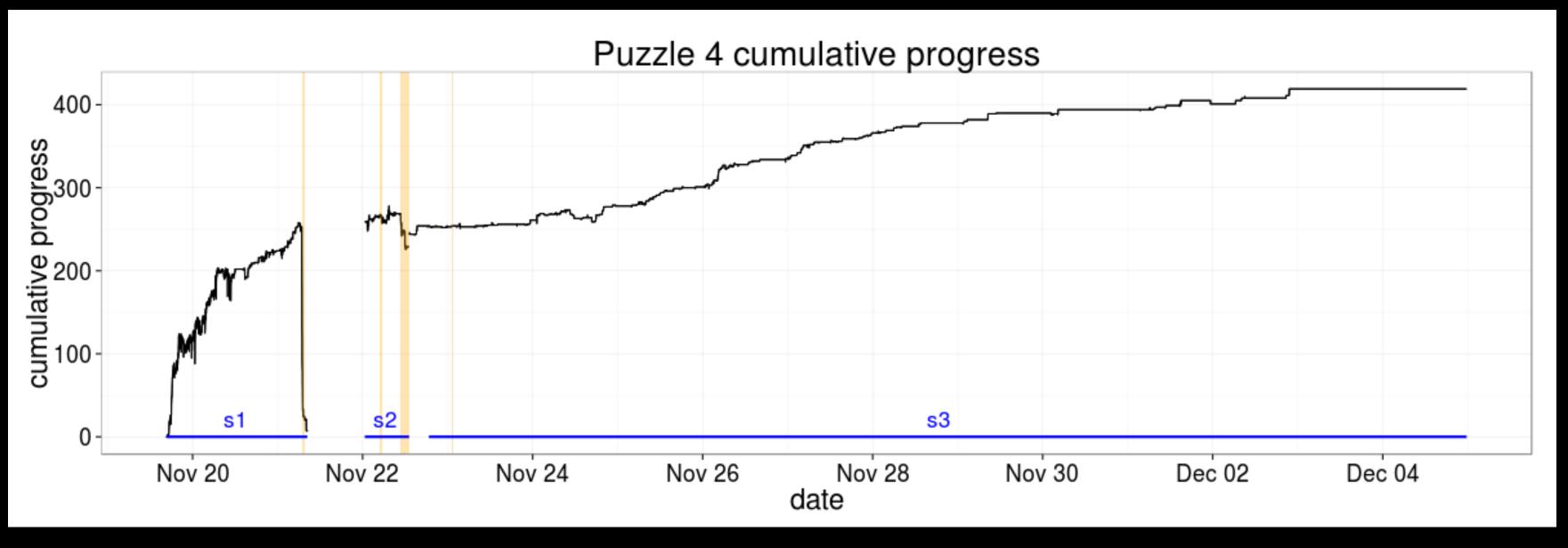
Crowdcomputing and Citizen Science for Large-scale Experiments

Moral Machine - data collection

3.2 Musers



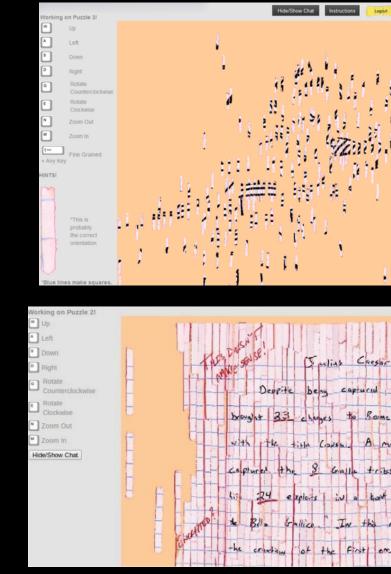
Sabotage activity logs



[Stefanovitch et al. EP] Data Science, 2014]

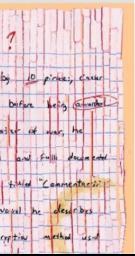
"For my first attack, [...] moving all the pieces into a single pile. [...] However, it seems all you did was lock the pieces together and ban my IP address."





DARPA Shredder Challenge









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"

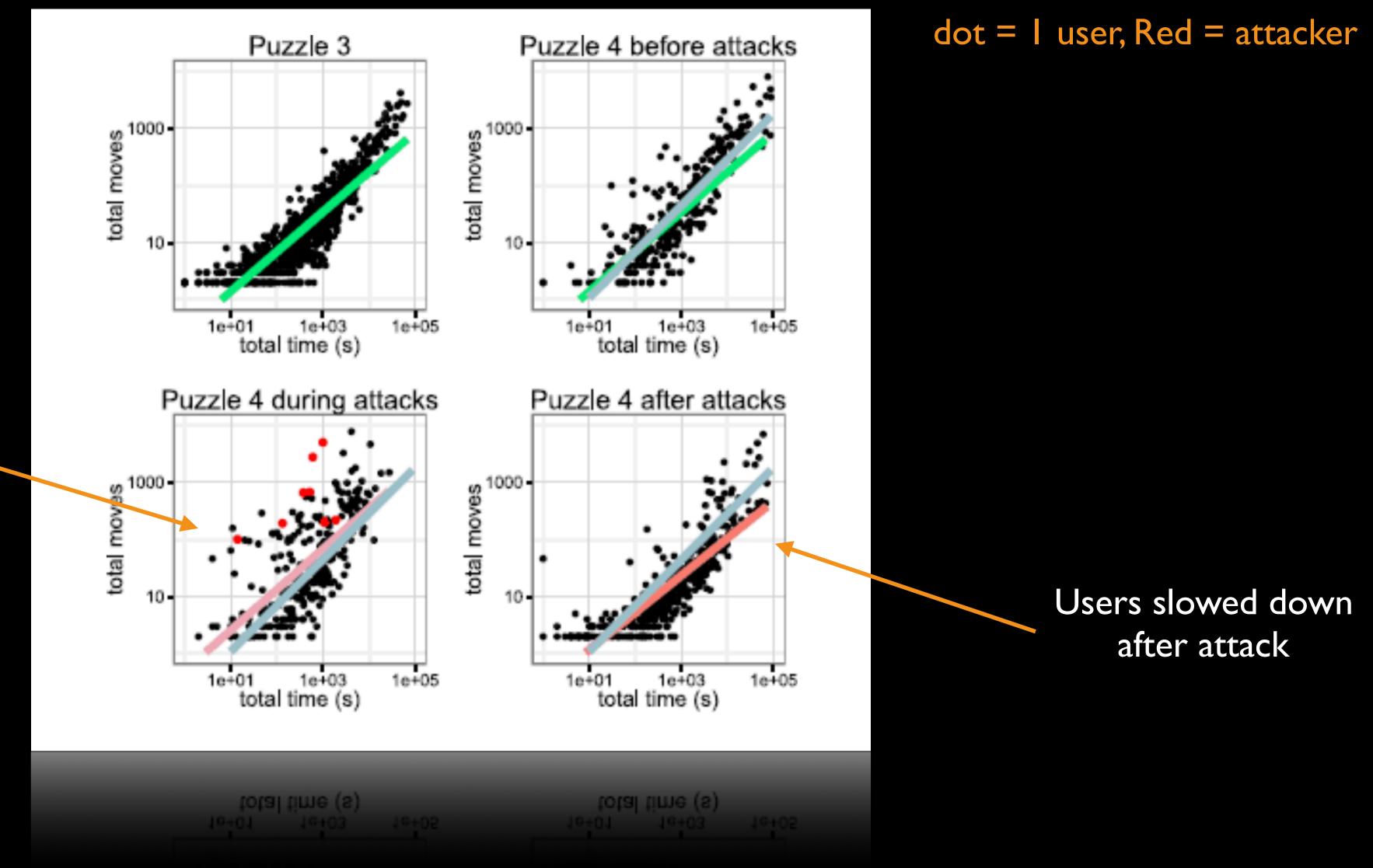






Building vs policing

Both attackers and responders move fast







Analysis

- Plan out data analysis beforehand
- need
 - bad strategy
- How will you deal with attrition and dropout, partially observed data?
 - Is attrition correlated with experimental treatments?

Think about how you will collect your data. What data do you

• "Let's collect ALL the data (because we don't know what we will need later)" is a



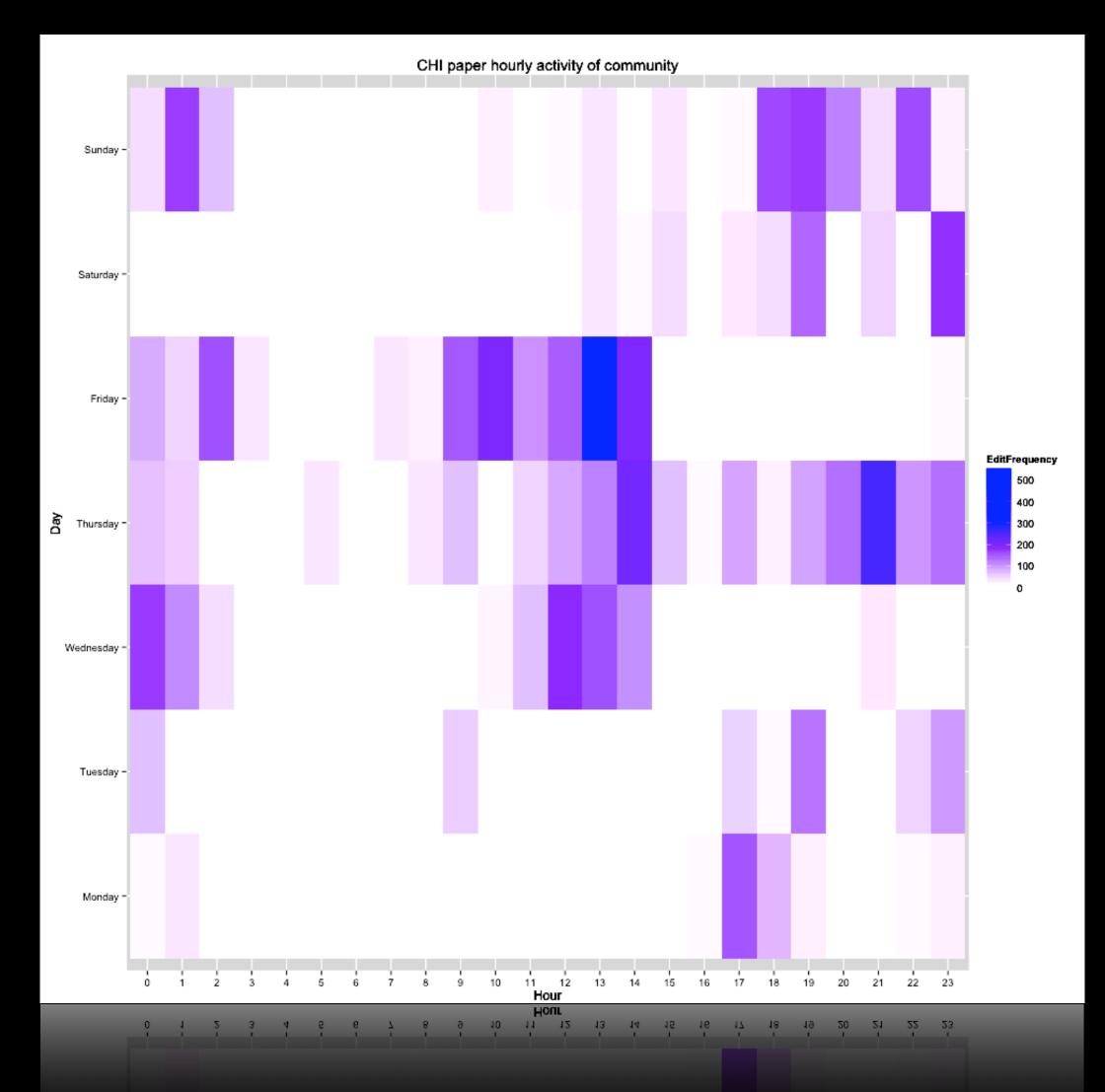
Analysis

- What analyses will you run if you had your data?
 - Do a dry run before going and collecting all the data
- them!
 - Qualitative complements Quantitative

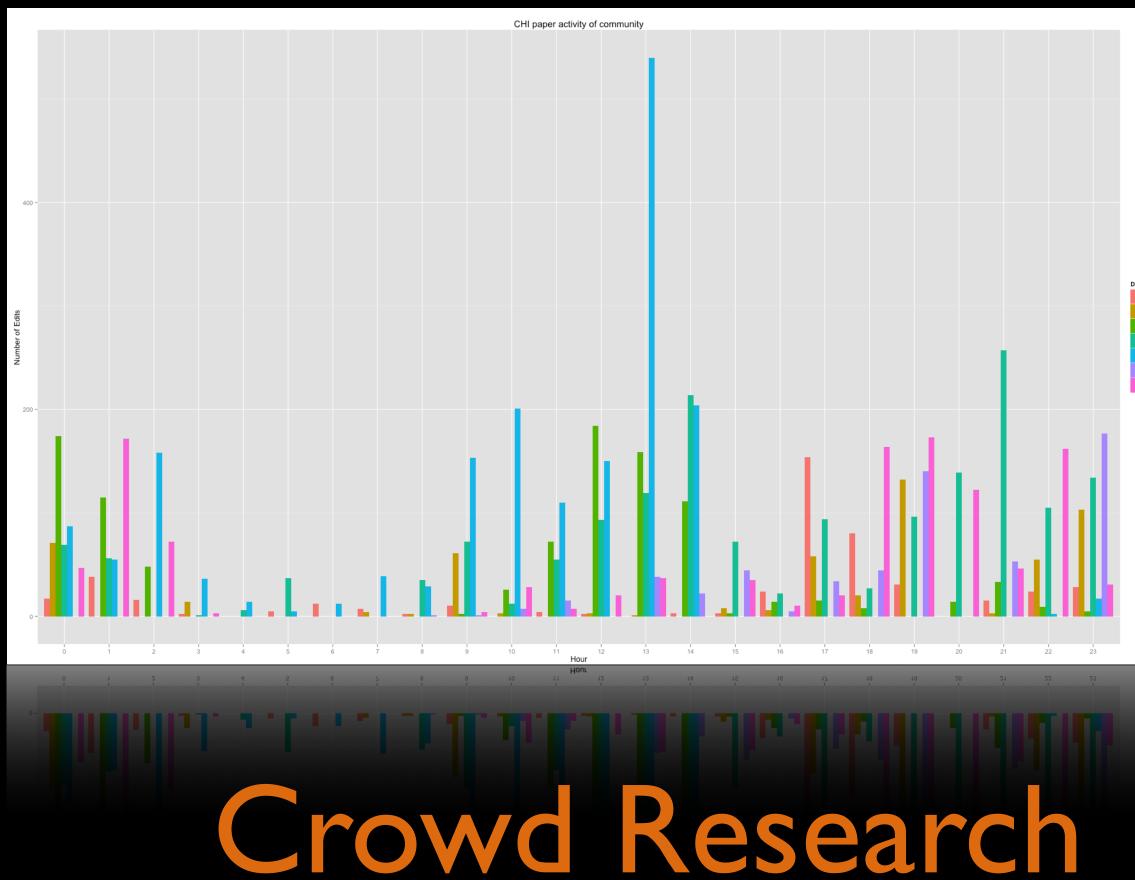
The best way to find out what people are thinking/doing: ask



Activity Scheduling [time, day]



Crowdcomputing and Citizen Science for Large-scale Experiments





Monday Tuesday Wednesda Thursday Friday Saturday

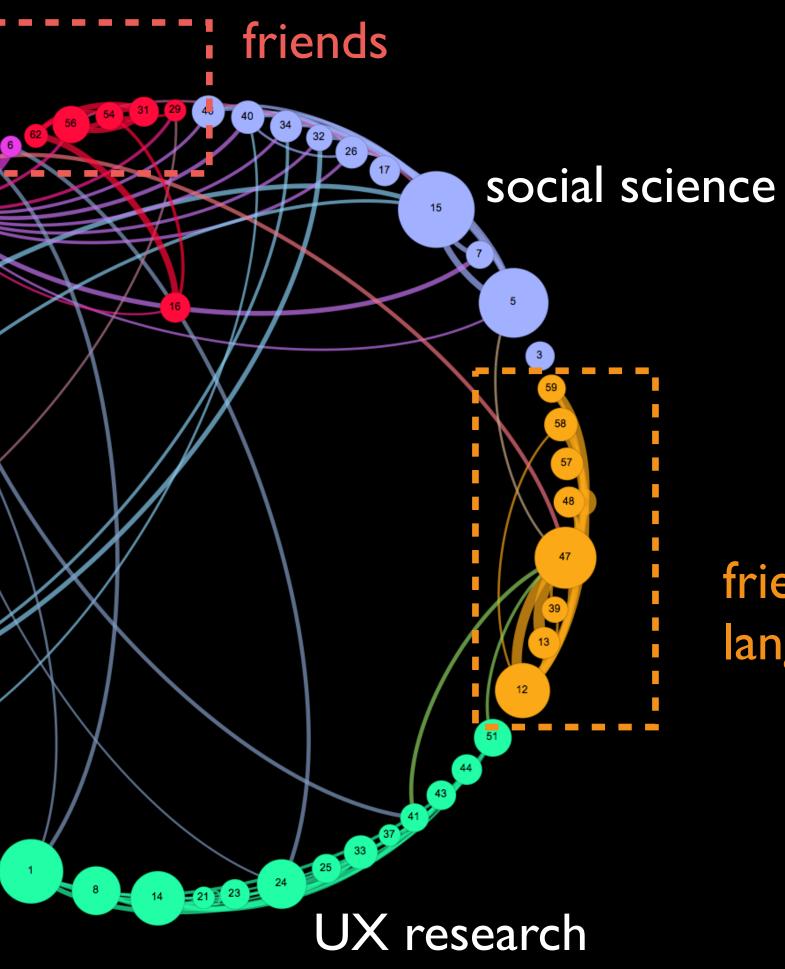


Community Formation

engineering & programming

design

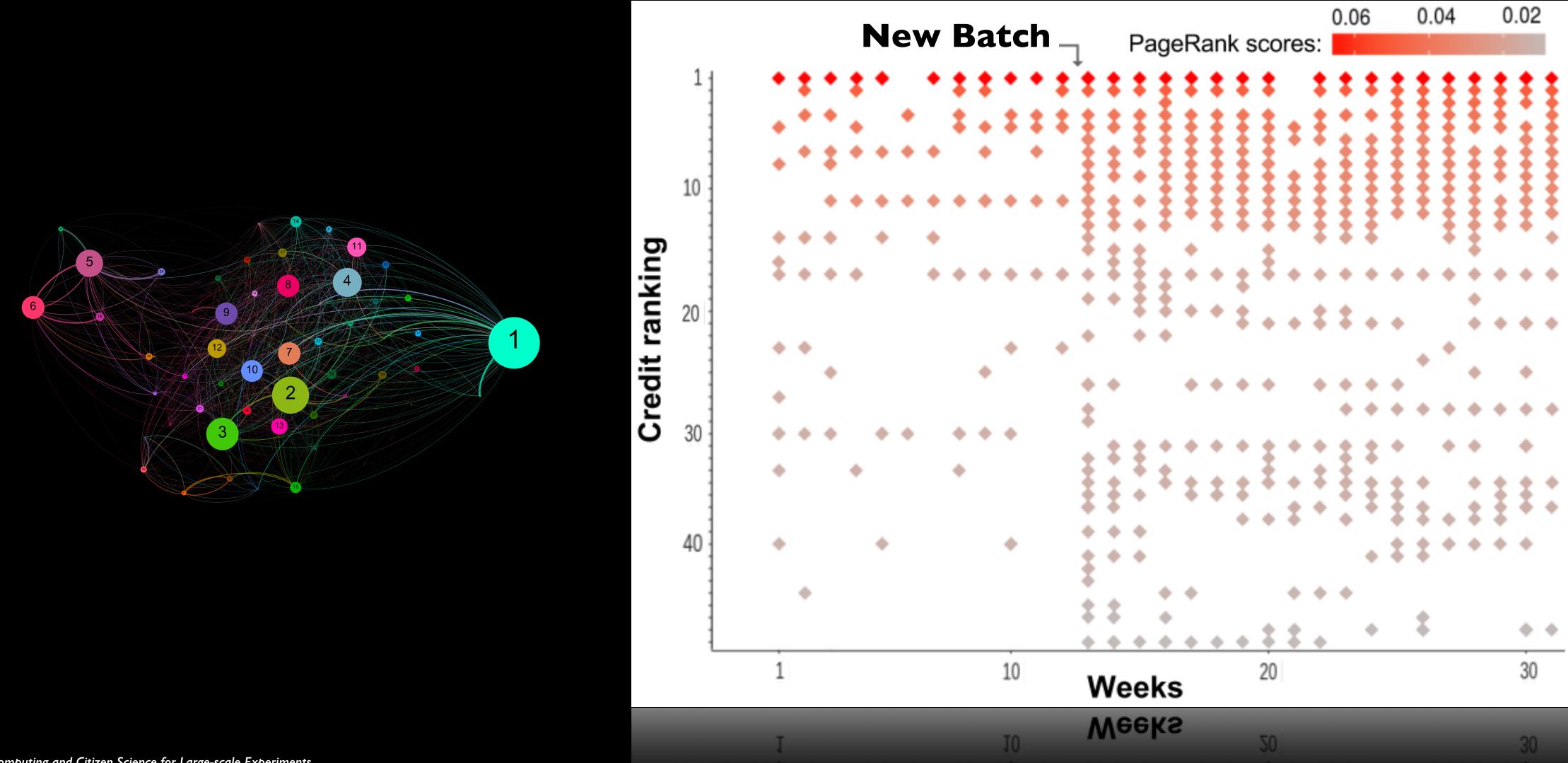
Crowdcomputing and Citizen Science for Large-scale Experiments



friends who speak similar languages



Characteristics of leaders



75

CASE The Problem and Data Science

Measurements, Hypotheses, Analysis

Crowdcomputing and Citizen Science for Large-scale Experiments

Barriers to Entry & Task Logistics

Data Collection



76

Platform and Infrastructure

Crowdcomputing and Citizen Science for Large-scale Experiments

Design



No Infrastructure

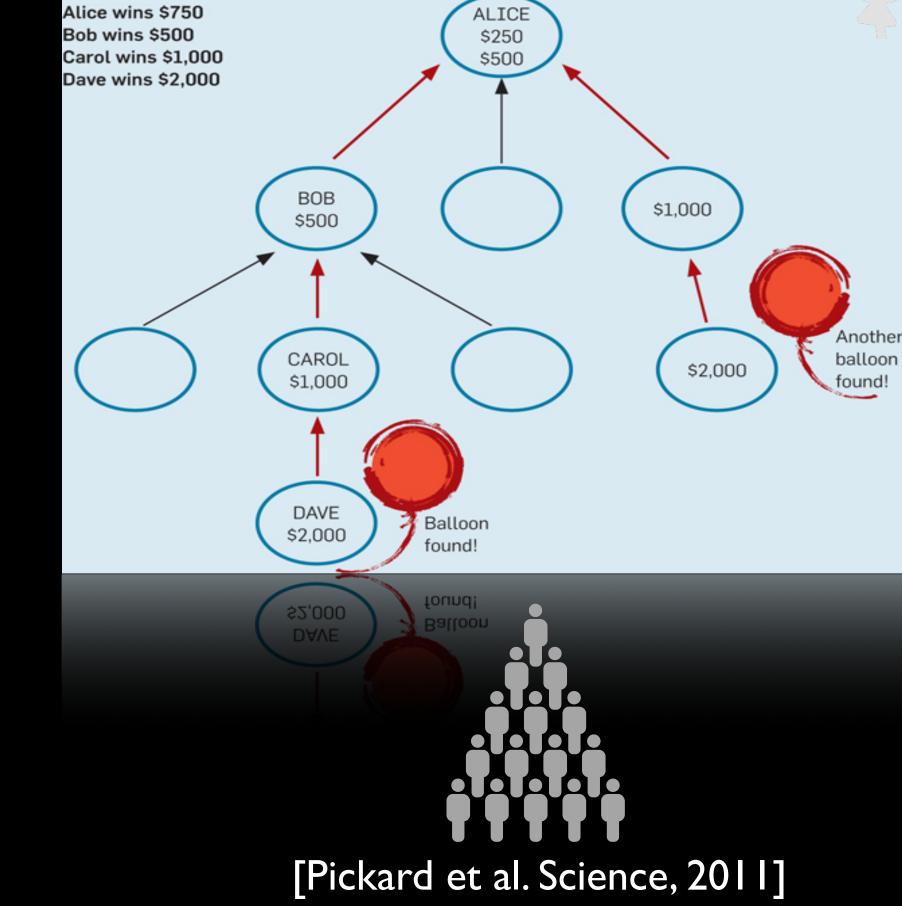


DARPA NETWORK GHALLENGE













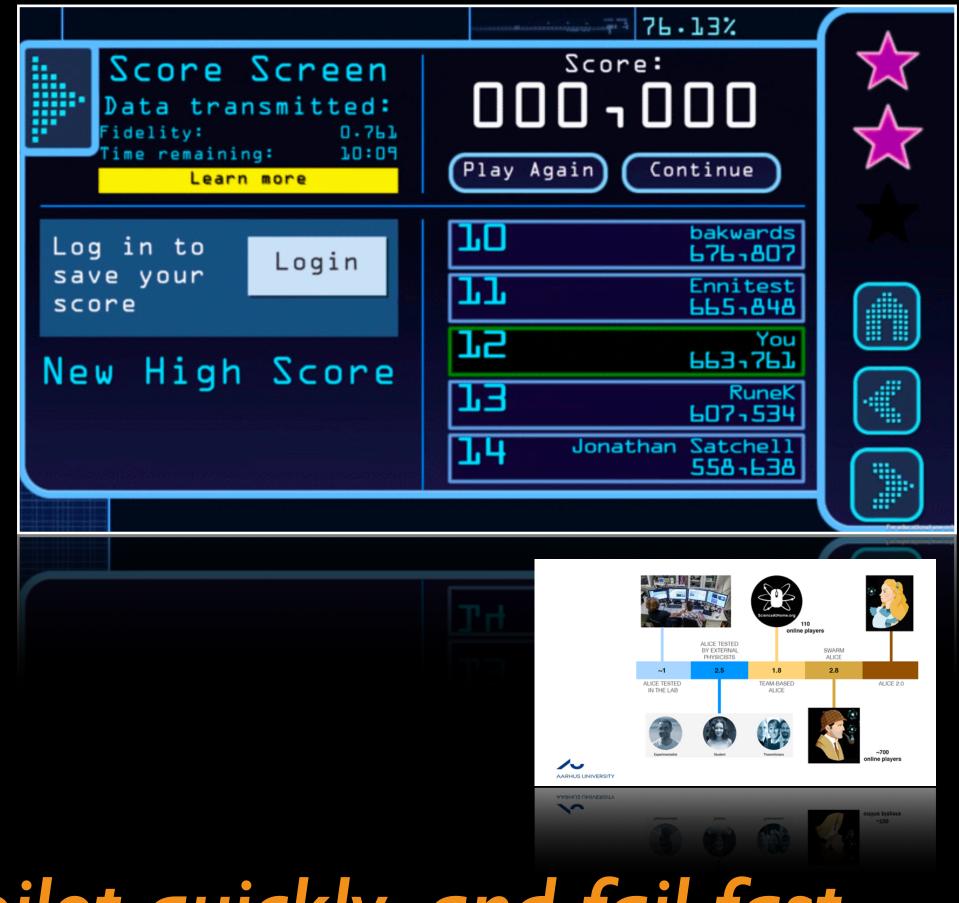


Game Mechanics & Engineering





Iterative Development— Prototype, pilot quickly, and fail fast





Web Engineering



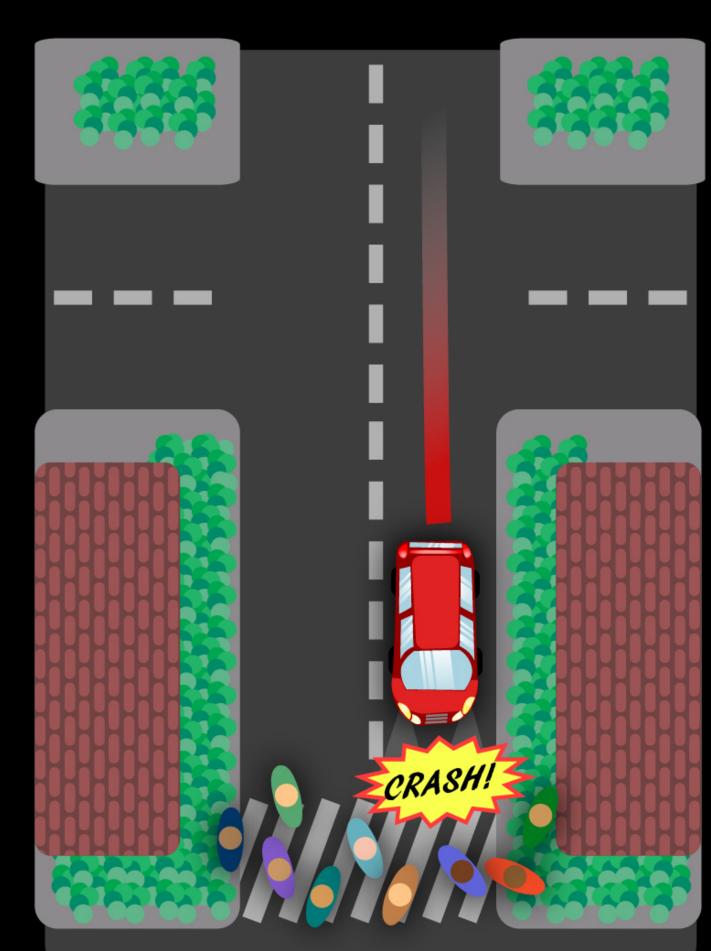
Home Judge Classic Design



Crowdcomputing and Citizen Science for Large-scale Experiments

About Feedback 🔇 En Browse







UU

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



Crowdcomputing and Citizen Science for Large-scale Experiments

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





Crowdcomputing and Citizen Science for Large-scale Experiments

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

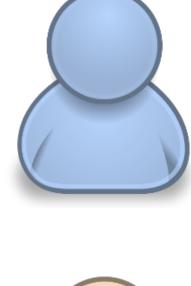
- Incentives: financial, via MTurk
- Retention: combination of social and financial incentives
- Would people come back each day? • We didn't know

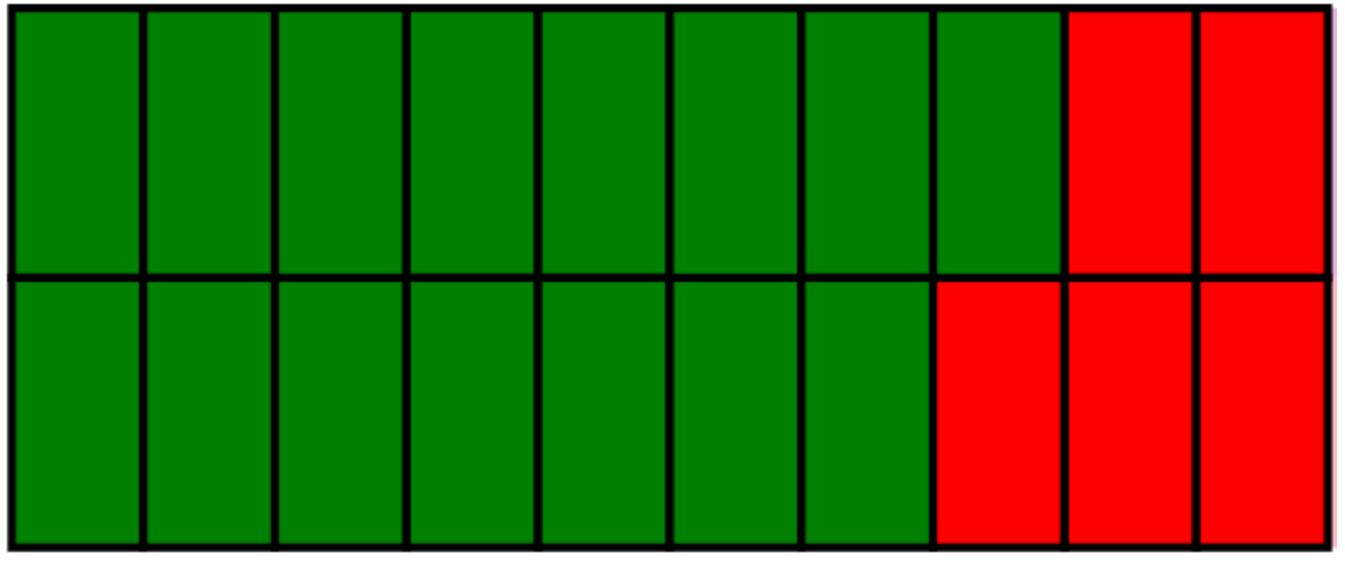


Our experiment



anonymous partners





Round 1

Cooperate

Defect

Round 10

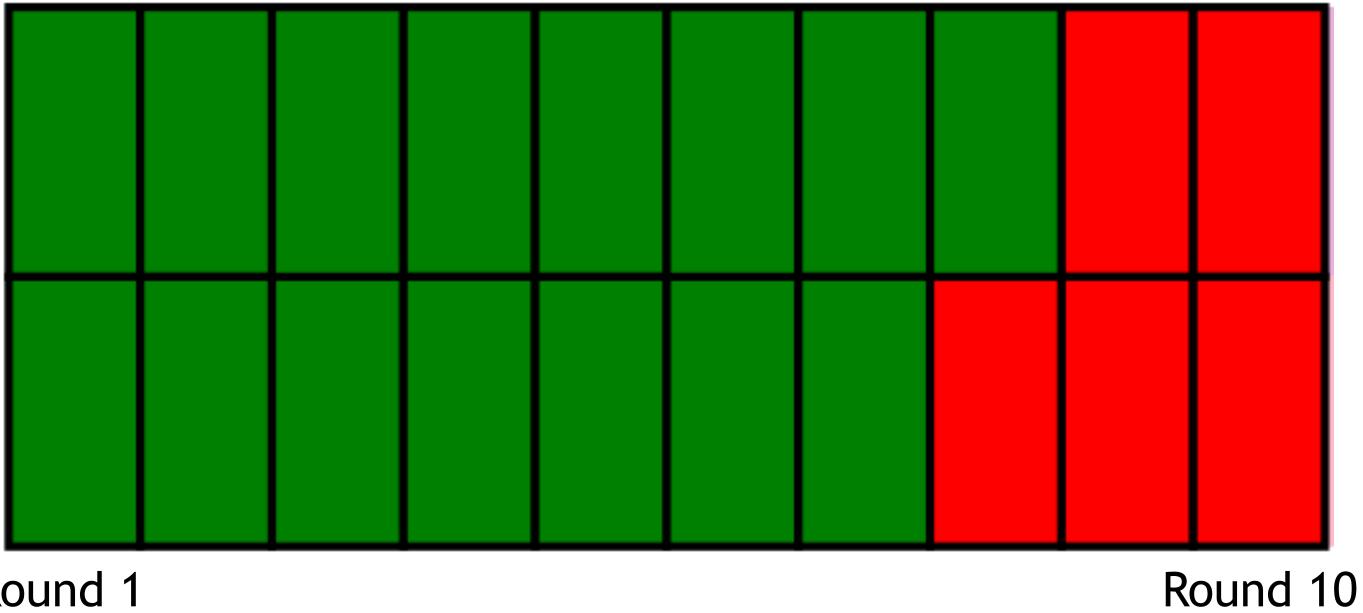


Our experiment



anonymous partners





Round 1

Cooperate Defect

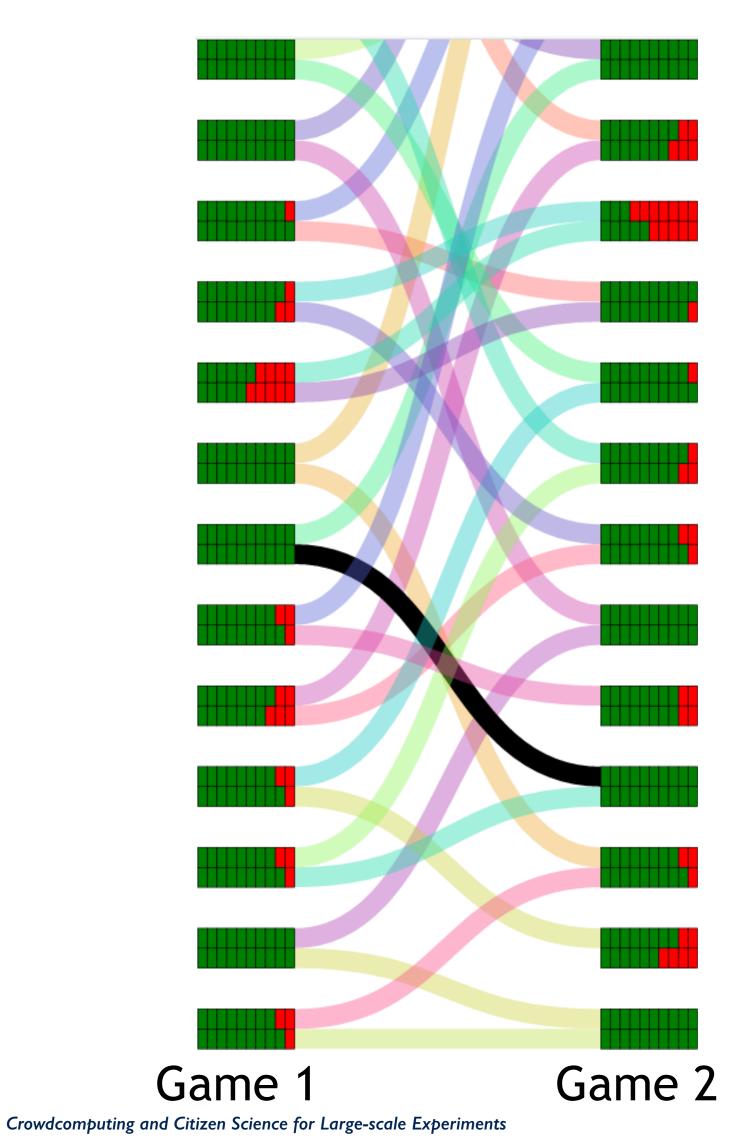
Game



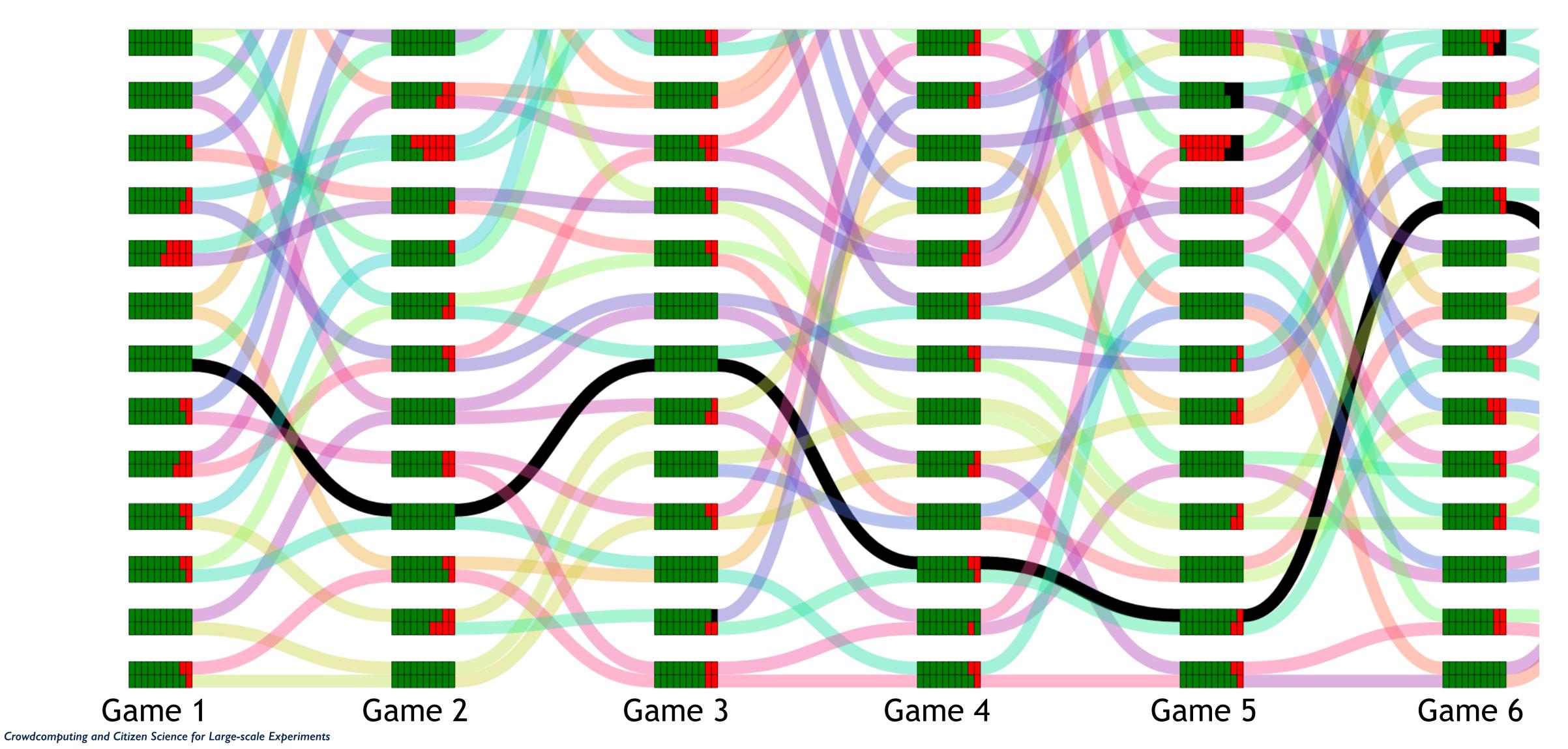
_	 	



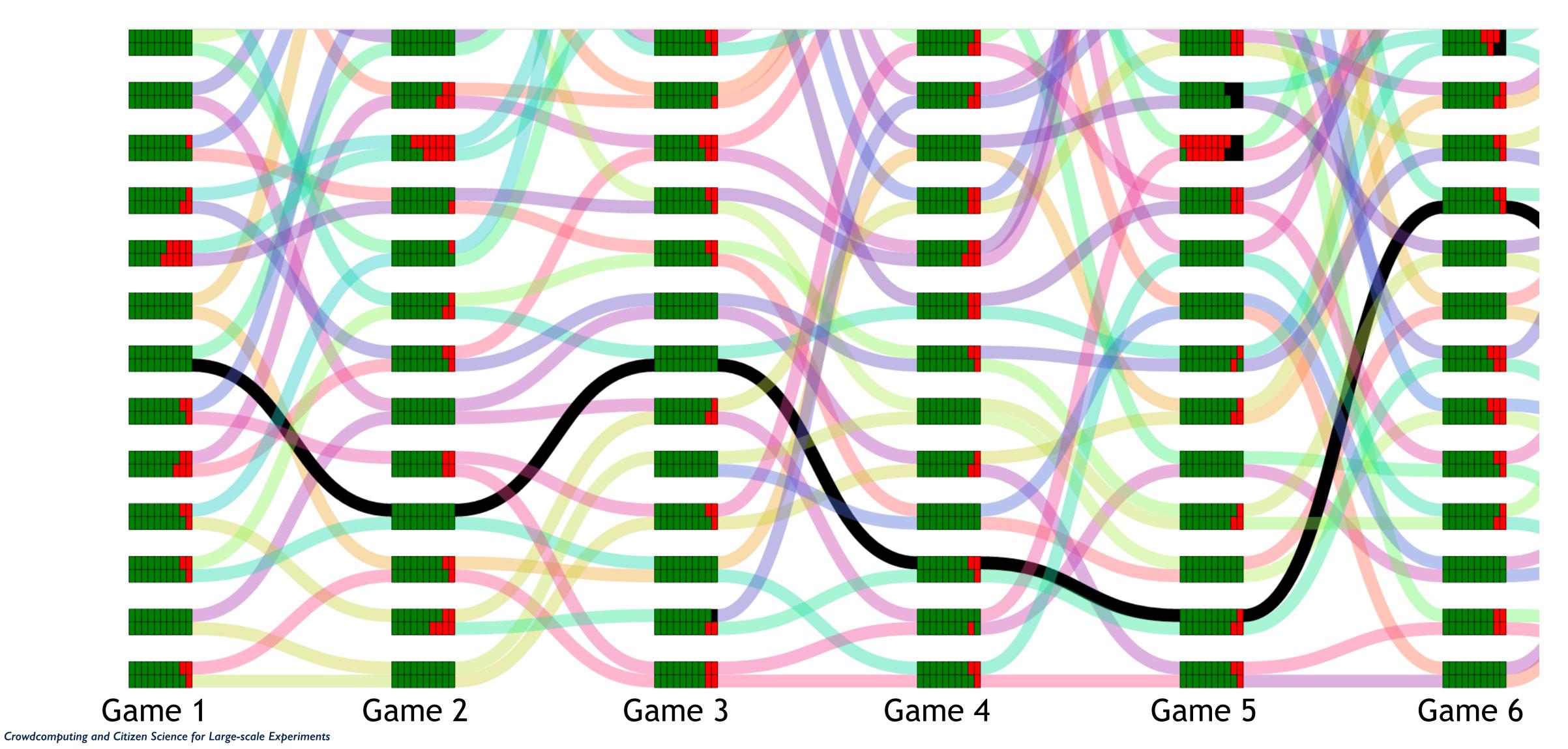






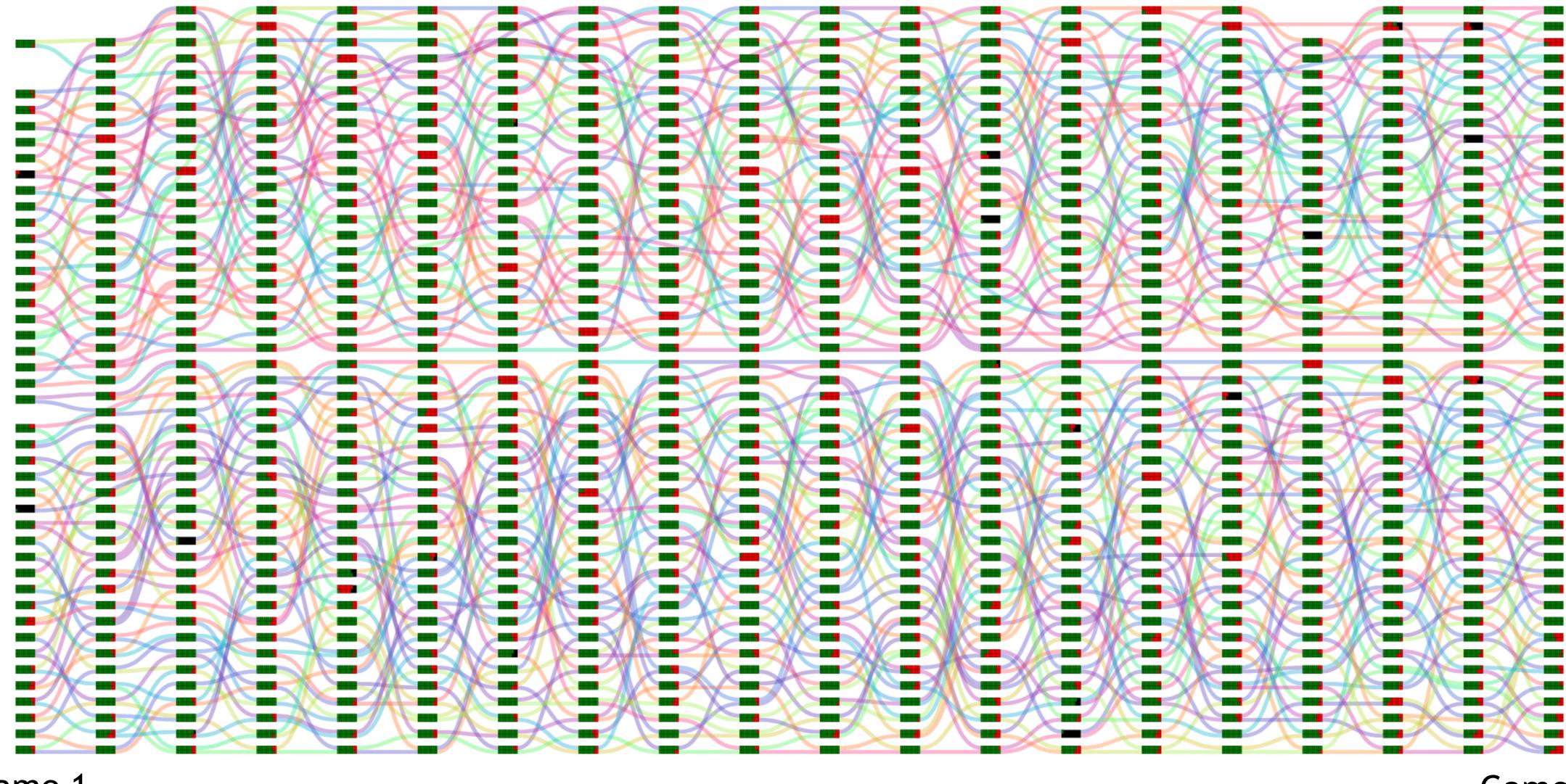






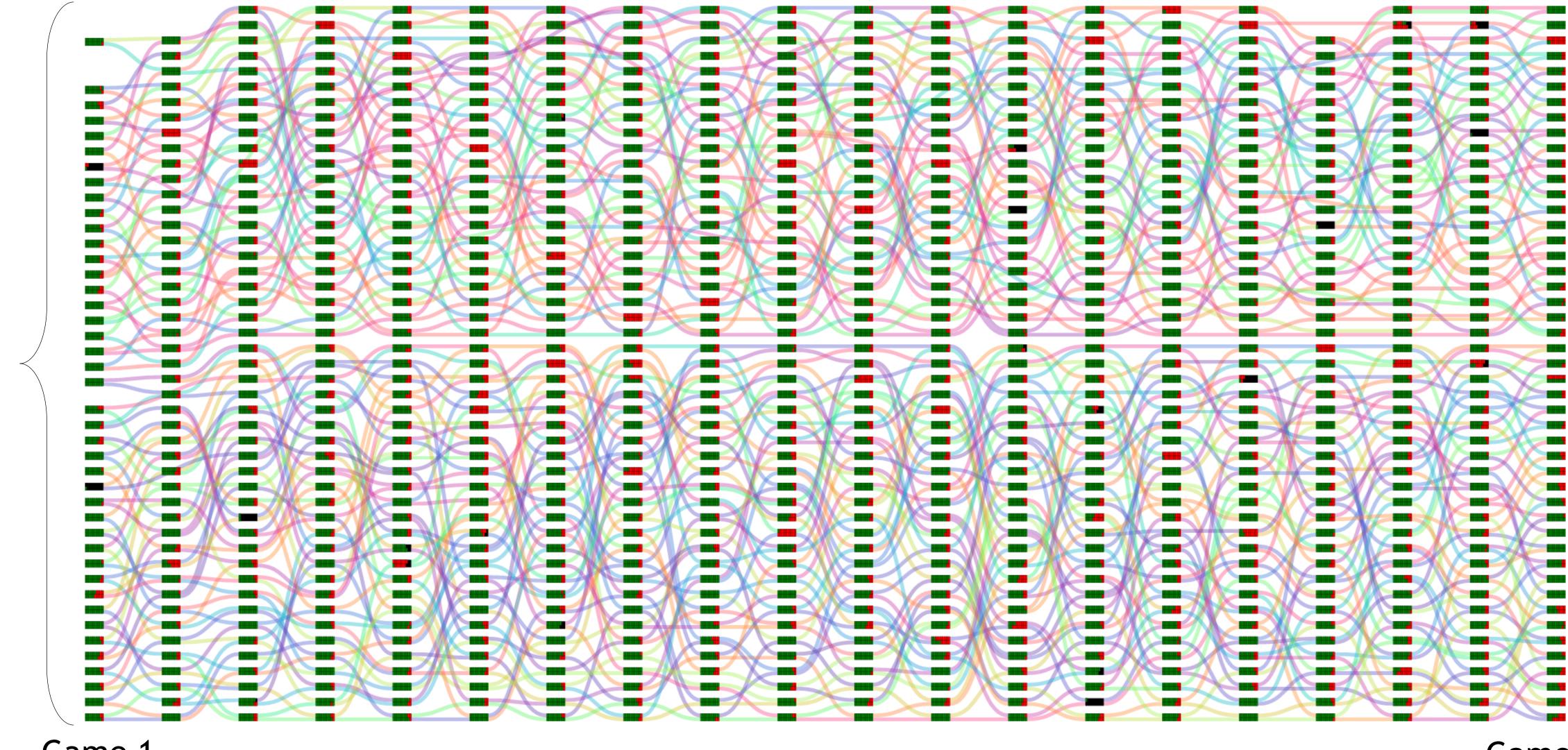


One experiment session – 20 games





One experiment session – 20 games



50 pairs



Aug 4, 2015 - Day 1	

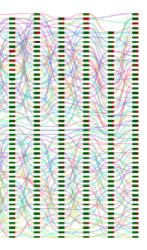
Z
X
X

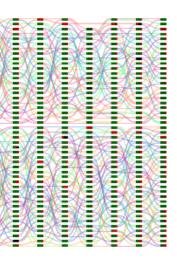
Crowdcomputing and Citizen Science for Large-scale Experiments

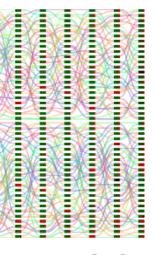
X NUMBER OF STREET

X + + + 3

Aug 31, 2015 - Day 20











Logistics

- 20 days
- 20 games per day
- I0 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

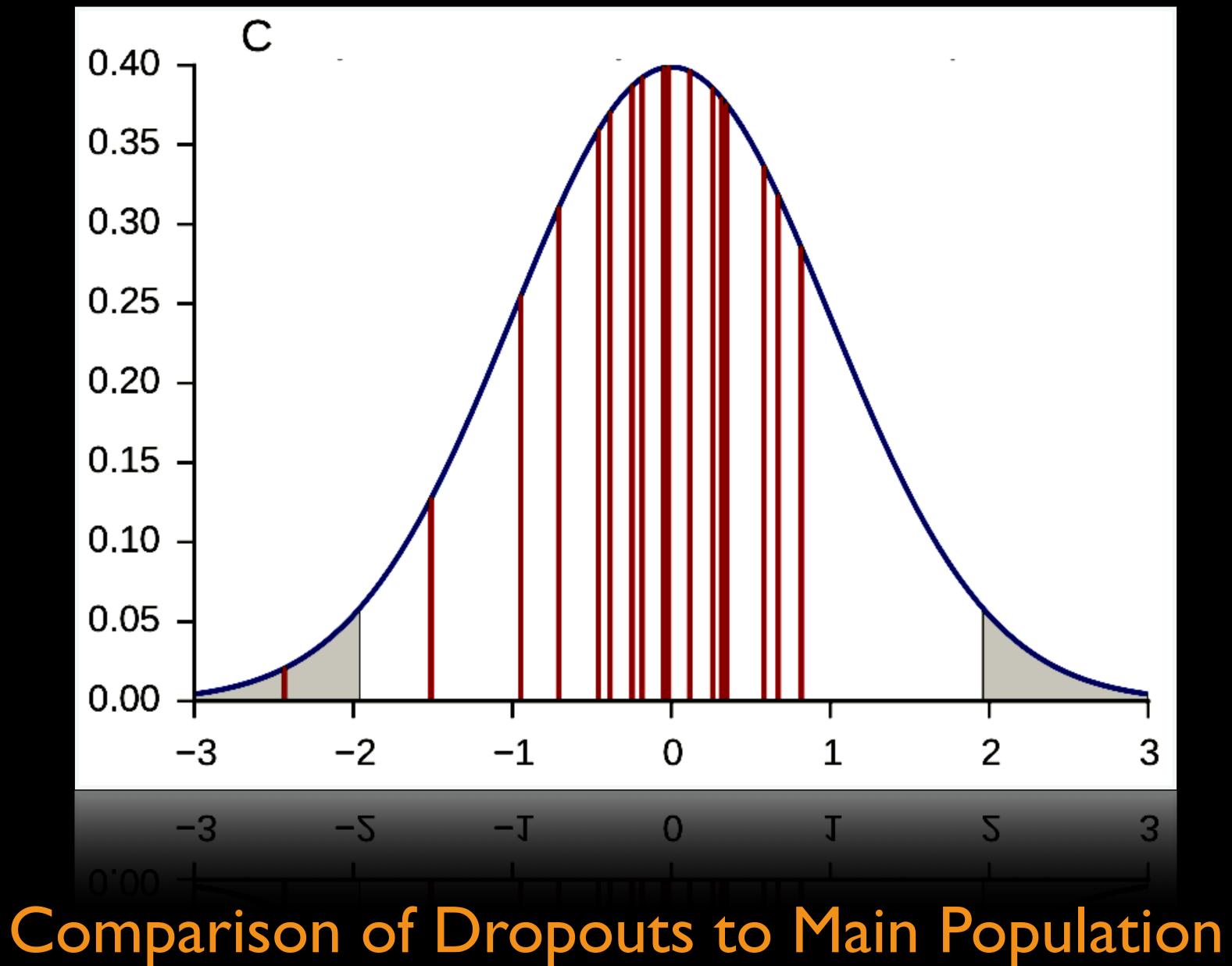
Social incentives

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

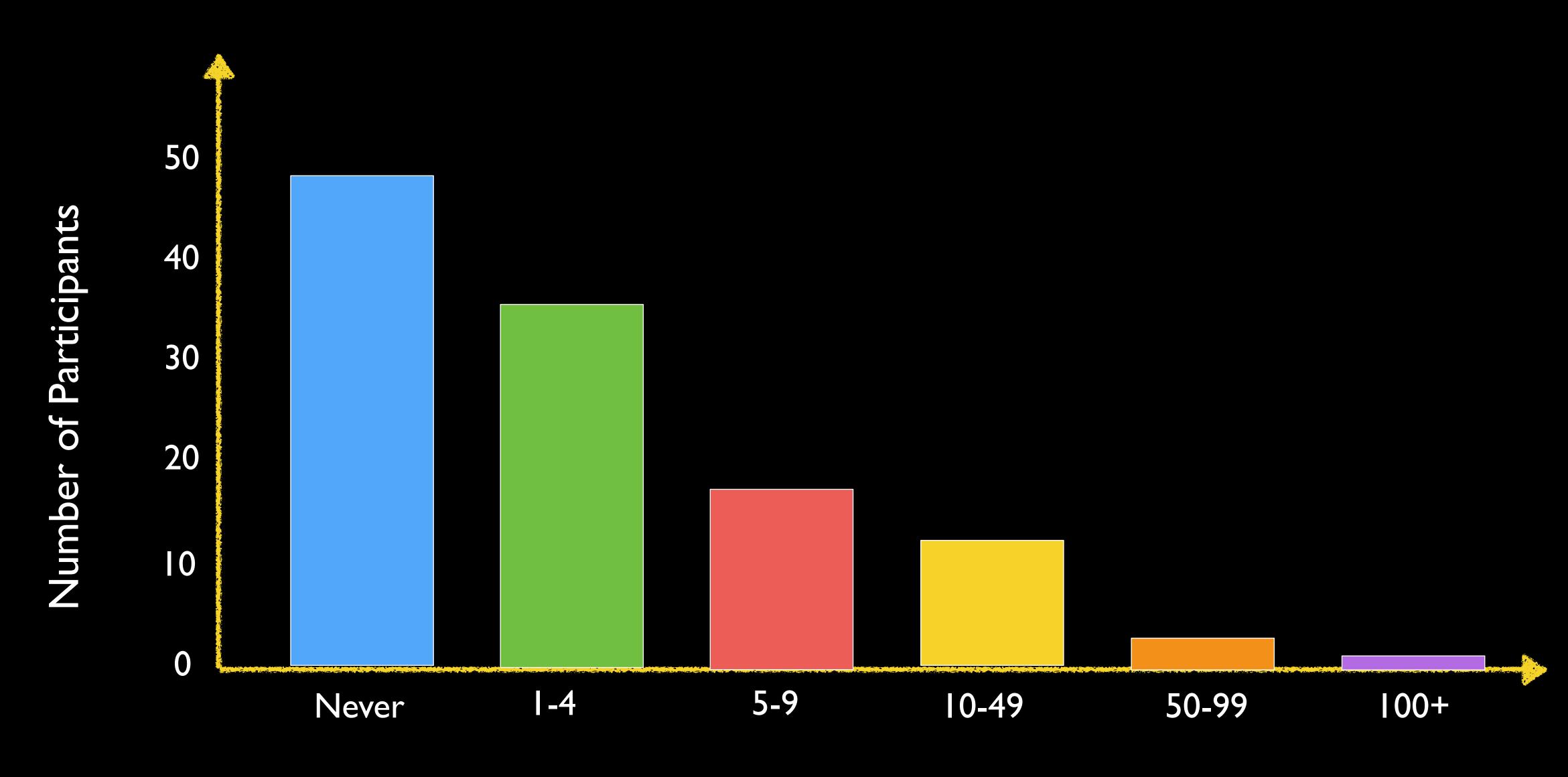
• Financial incentives

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









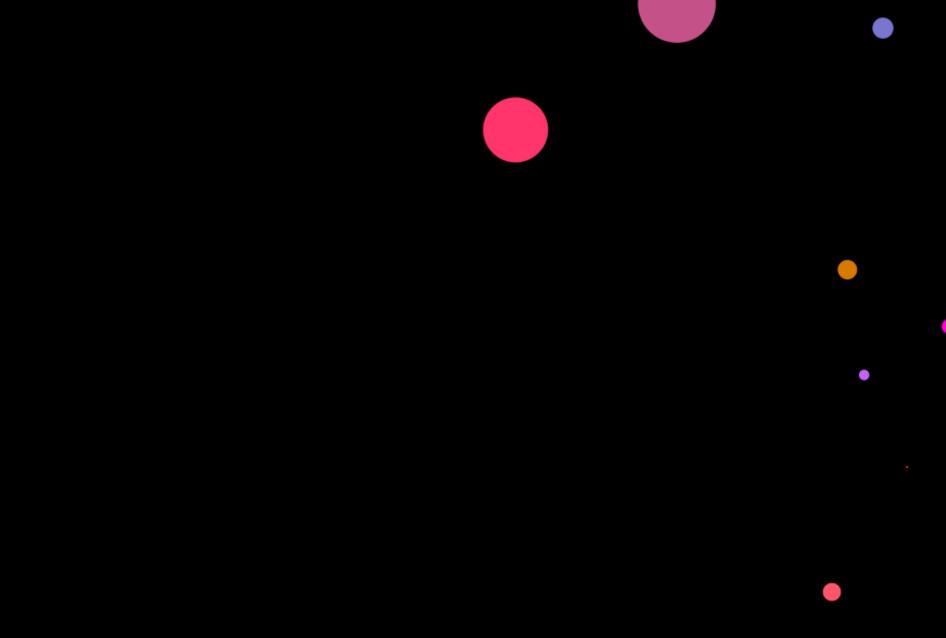
Past Prisoner's Dilemma Experiments



Strategic Behavior in the Stanford Crowd Research

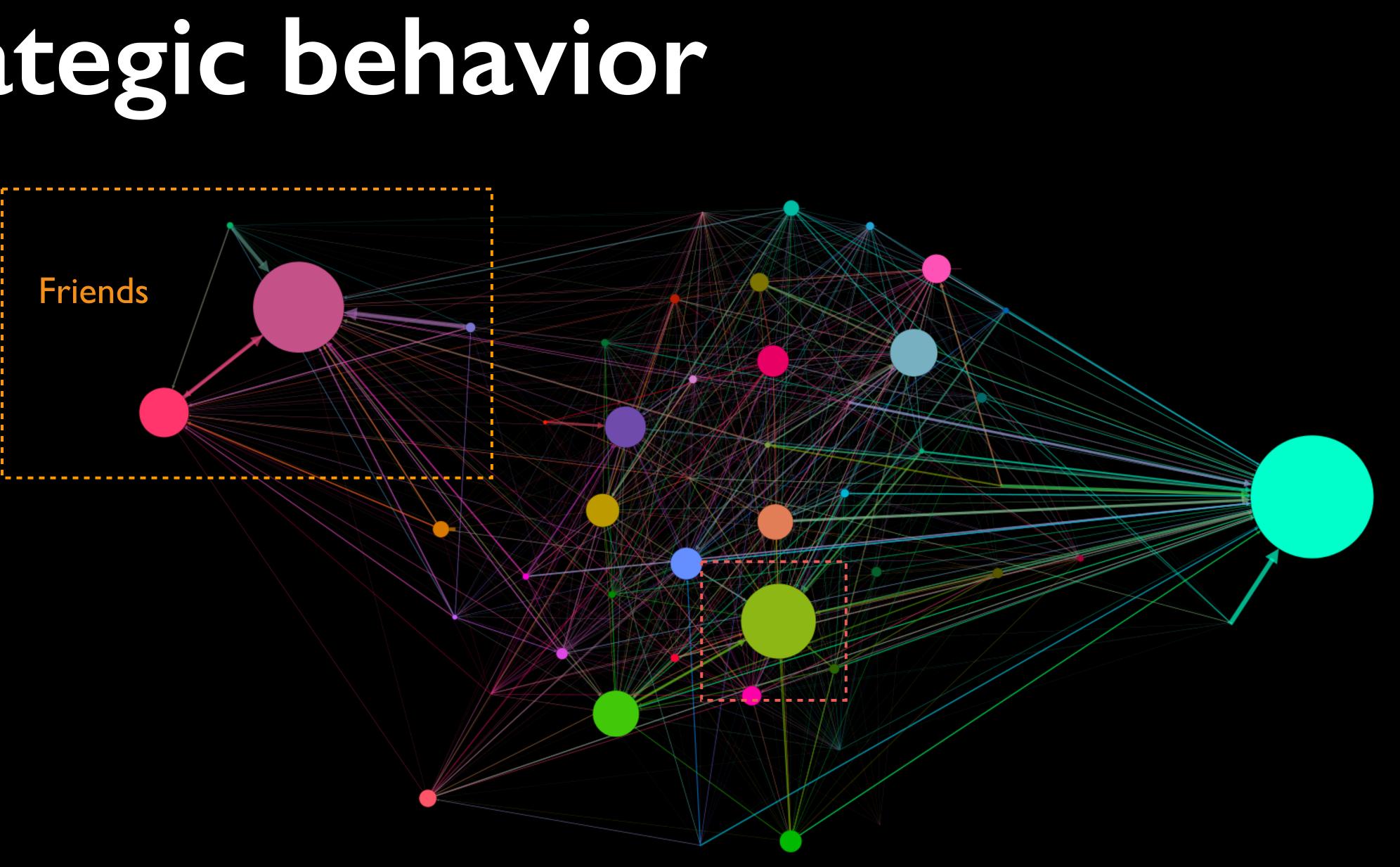


Crowd Dynamics





Strategic behavior



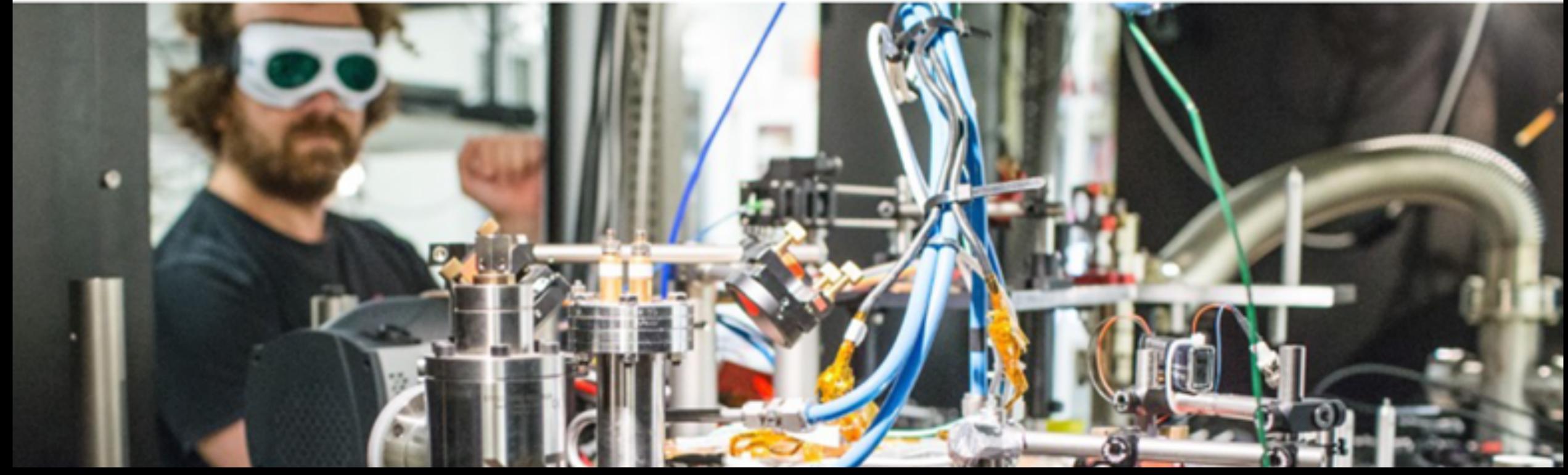


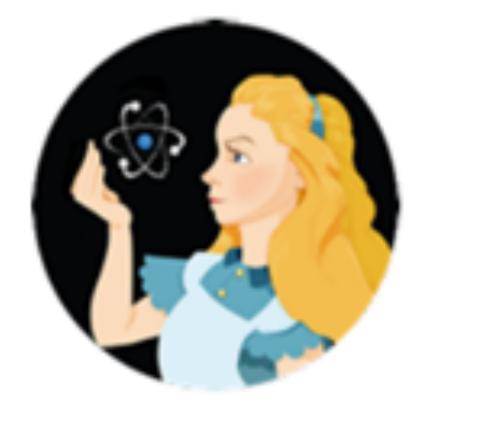


Crowdcomputing and Citizen Science for Large-scale Experiments

Alice Remote Control







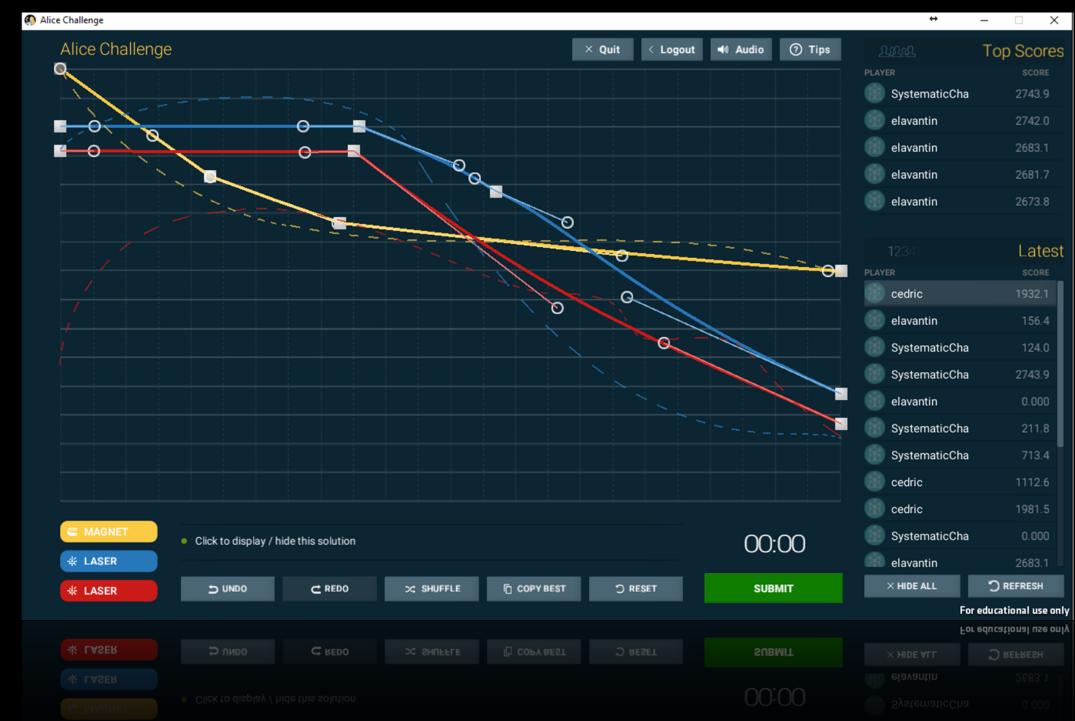


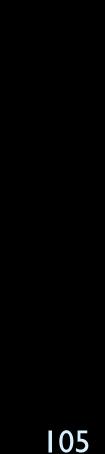
Chalenges

- Unprecedented experimental setup leads to unprecedented problems

 - Recruitment, team allocation, waiting room?
 - Data backlog
 - Gamification matters

• Social science investigations were dependent on having a team online simultaneously





Crowdcomputing and Citizen Science for Large-scale Experiments

Discussion



Crowd = Richer Research at Scale

The Internet as a lab

Size. Scale

More Samples 🞽 Large Social Interactions

Crowdcomputing and Citizen Science for Large-scale Experiments

Complexity, Realism

Realistic vs Abstract More precise instrumentation

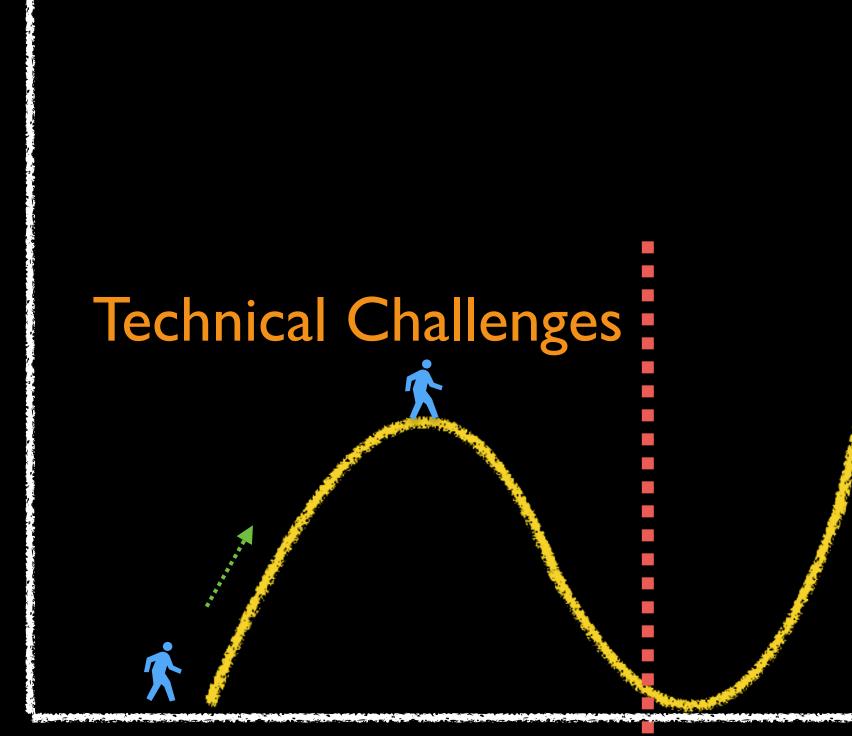


Duration, Participation Longer period of time Fewer resource constraints



107

Crowd is synergistic

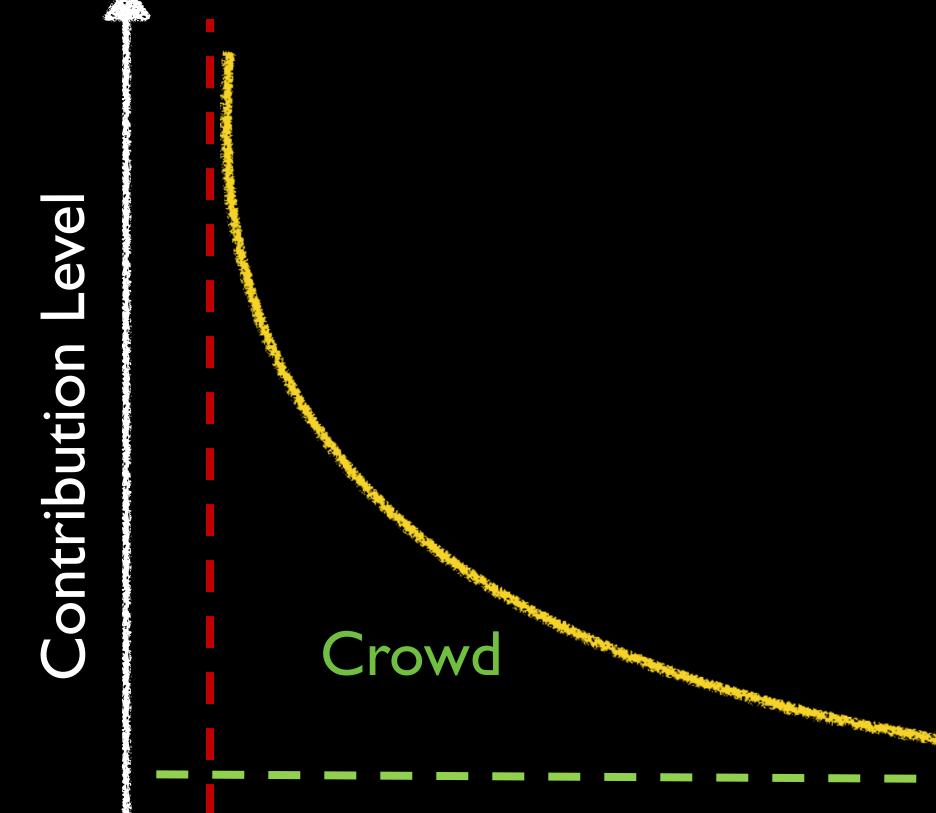


Logistical Challenges



Make use of the long tail

Scientists, Experts



% of Participants

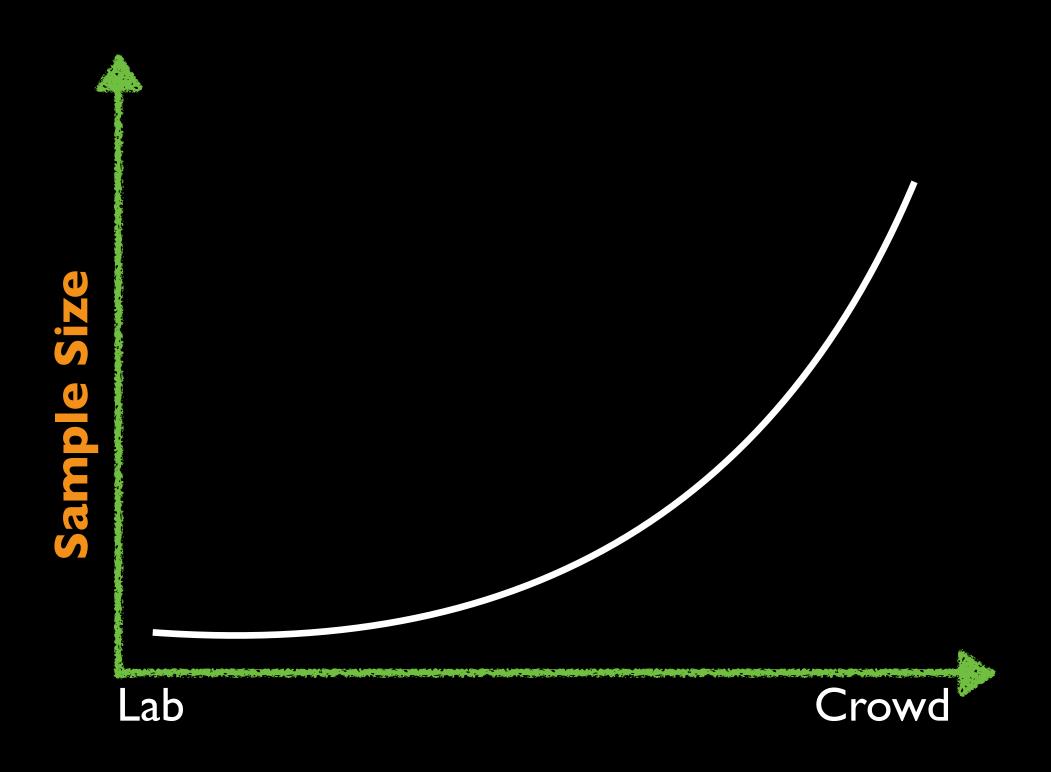
Lab experiments

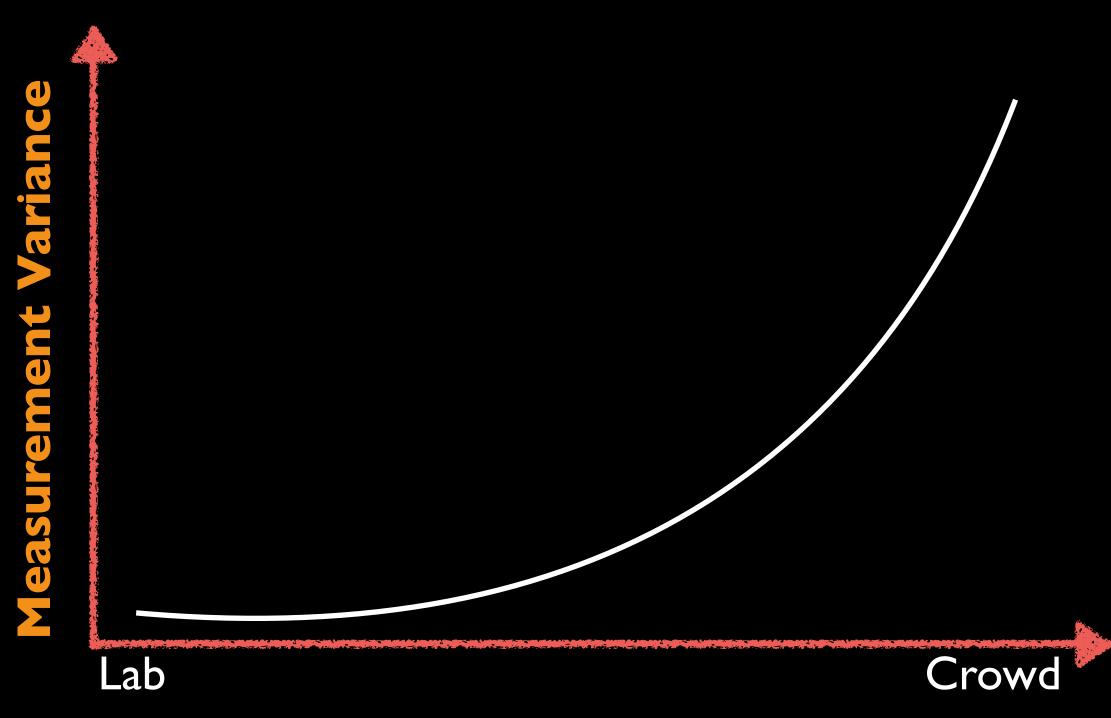
the second s

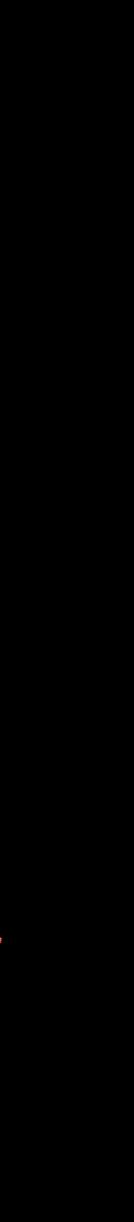




Tradeoffs in increasing size







10

Global Ecosystem

Education, Partnership

Scientist

Data, Insights

Fun, Engagement

Crowd

Research Ideas, Feedback



Crowdcomputing and Citizen Science for Large-scale Experiments

Non-traditional impact





Democratize research Stanford Crowd Research



Science at the global scale EteRNA, Science Home—Quantum Moves



Rethink about Al and society Moral Machine





12574 970 43 808 82012998 041498186 44 2644151

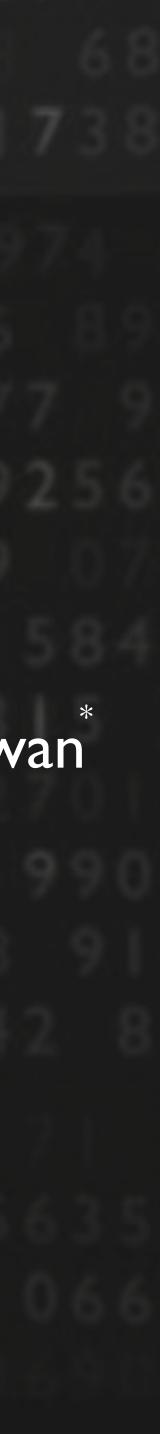
Crowdcomputing and Citizen Science for Large-scale Experiments

Snehalkumar `Neil' S. Gaikwad, Sohan Dsouza, Oana Vuculescu, Andrew Mao, Iyad Rahwan Microsoft Research New York[‡] Aarhus University[†] Massachusetts Institute of Technology

The International Conference on Computational Social Science 2017, Germany

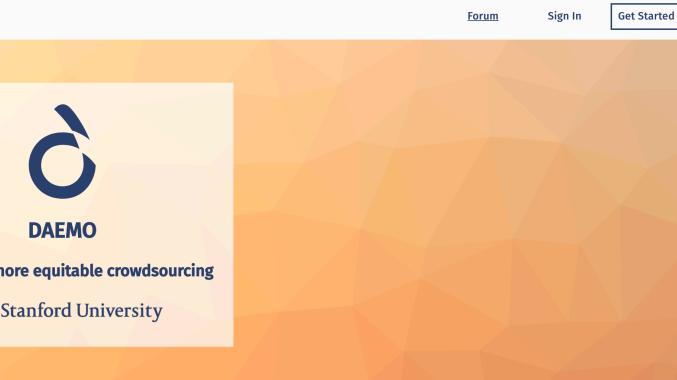
536 722 60 36 54

|uly 10, 2017



	Announcement
DAEMO	57 555024504
41044 14	6
10295241	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 more and more of your work. Post Tasks Get Results Rate Workers Grant Early Access
	Prost rasks Get Results Rate Workers Grant Larry Access Which Reddit thread will be most popular? Prediction of popular Reddit threads A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
	rnost popular? Fredit: threads P P P P P P P P P P P P P P P P P P P
send an email	Daemo is live! to gaikwad@mit.edu or da

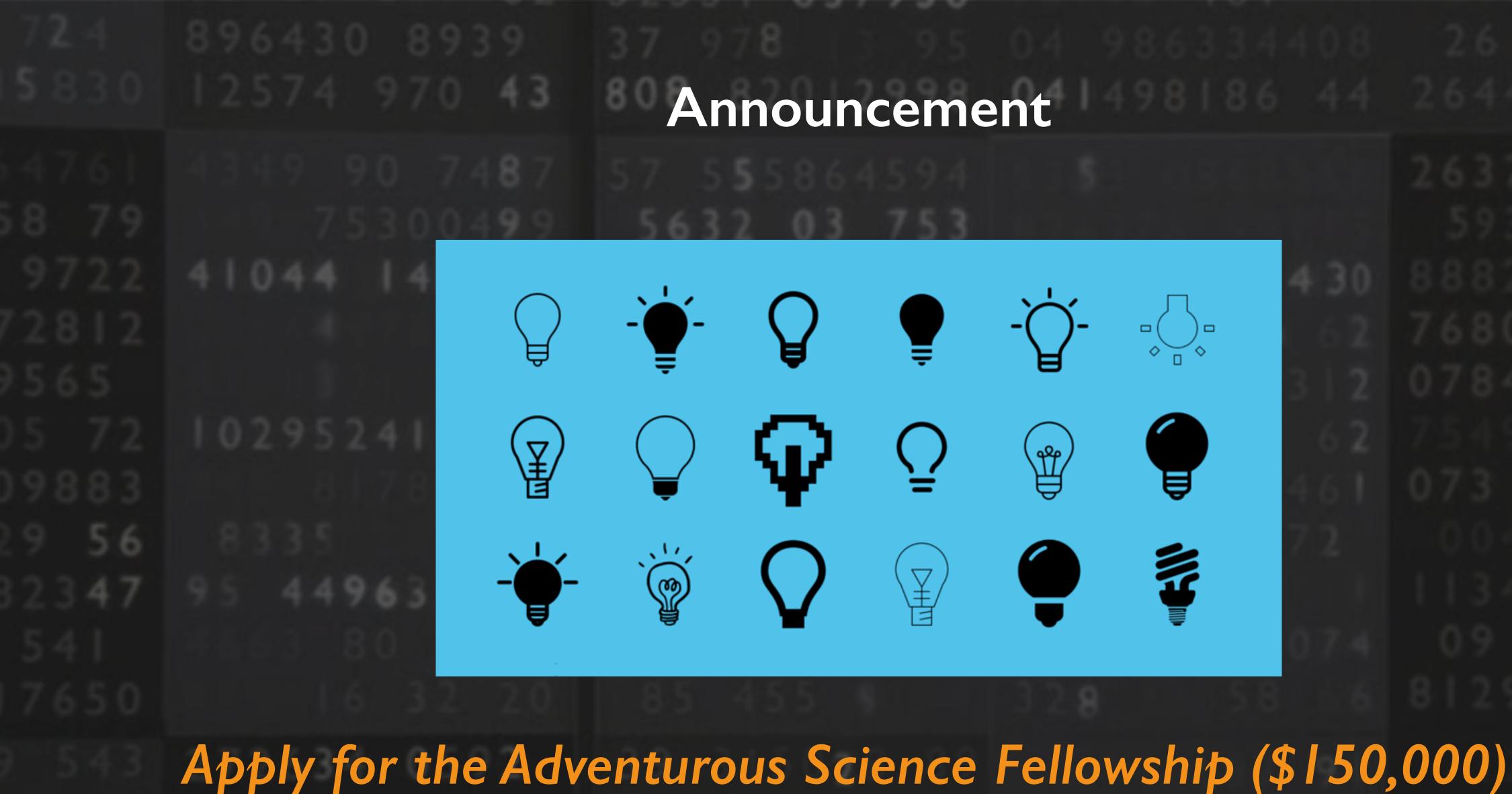
722 60 36 54



s will do

aemo@stanford.edu





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



- » [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







- » [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
- » [peirotis et al. ACM SIGKDD, 2010] Quality management on amazon mechanical turk
- » [peirotis 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



- » [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

2

- » [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. 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, 207] A solution to the single-question crowd wisdom problem
- » [Prelec et al. Science, 2004] A Bayesian truth serum for subjective data
- » [Rutherford et al. PNAS, 2012] Limits of social mobilization

» [Mao et al. Nature Comm, 2016] Resilient cooperators stabilize long-run cooperation in the finitely repeated prisoners dilemma

» [Procaccia et al. ACM Transactions on Economics and Computation, 2013] Approximate mechanism design without money

» [Rzeszotarski et al. UIST, 2011] Instrumenting the crowd: using implicit behavioral measures to predict task performance



- » [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. EP] 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



