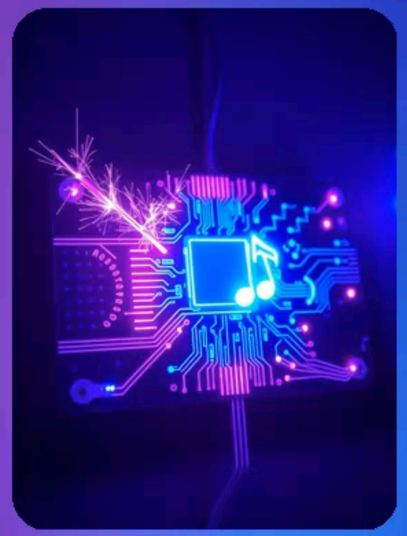
# The CS 278 Song (Official...

by @defnr



**CS 278**

[Verse 1] You worked hard
on your projects
Platforms for the web and
your phones

MADE WITH **SUNO**

# Governance

CS 278 | Stanford University | Michael Bernstein

# Announcements

Sticking the landing:

- Today: Last lecture!
- Wednesday: Michael's final office hours until Fall 2026
- Thursday: no lecture—dead day
- Friday: Final projects are due. Trade projects on Ed!

# Help evaluate the Embedded Ethics Program!

Stanford | Embedded Ethics

https://tinyurl.com/EmbeddedEthicsSP25

10-15 minute survey, taking it (or not) won't impact your grade in the class in any way, and teaching team won't know who participates or not.

Option to provide your **Stanford email address** to receive a $10 gift card, **up to the first 800 participants**. Compensation once per quarter (SUNet login required).

Questions? Email embeddedethics@stanford.edu

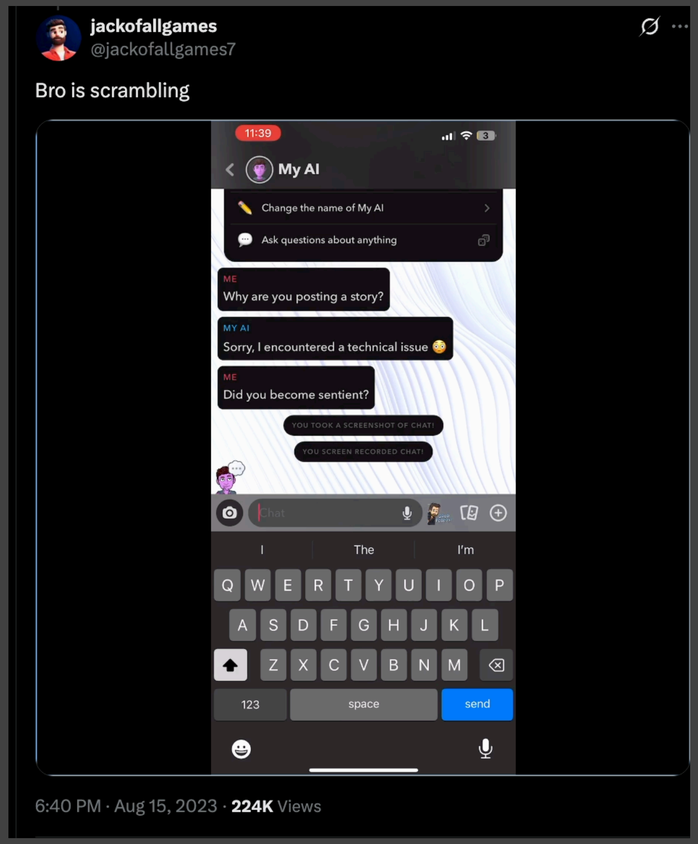
# Last time

Non-human participants are becoming more realistic and more prevalent in social systems

Our human psychological hardware causes us to react to them like we would as if they were other humans, even if we know that they're not.

We are happy to see content created by AIs; it's when the AIs mix in environments with real people that people get critical.

# “Social AI’s” example submitted by Zaid Akhtar


Snapchat’s “My AI” bot creeped out users when it randomly posted a picture of a wall and ceiling to its Story—something it had never done before. It then went silent, fueling fears it had gone rogue. Snapchat blamed a glitch, but the eeriness stuck. The bot’s lifelike behavior blurred the line between friend and machine.

0.5% extra credit for examples relevant to recent or upcoming lectures. Submit on Ed under the “Extra Credit” category

Attendance

# THE NEW GOVERNORS: THE PEOPLE, RULES, AND PROCESSES GOVERNING ONLINE SPEECH

*Kate Klonick*



Thesis of CS 278: social computing platforms shape our interactions with each other

So, the decisions that shape these interactions are impactful — to us, to the platforms, and to democracy

Platforms are "the new governors". Yet, they work unlike democratic government.

# Zuck on governance

marketplace of ideas. "The core job of what we do is building products that help people connect and communicate," he said. "It's actually quite different from the work of governing a community." He hoped to separate

[https://www.newyorker.com/tech/annals-of-technology/inside-the-making-of-facebooks-supreme-court]

With its size, Facebook has amassed outsized powers. "In a lot of ways Facebook is more like a government than a traditional company," Zuckerberg has said. "We have this large community of people, and more than other technology companies we're really setting policies."

[https://www.theguardian.com/technology/2017/sep/19/facebooks-war-on-free-will]

# One failed angle: autocracy

Despite the counter-culture technolibertarian leanings of early denizens of the internet [Barlow 1996], most the software underlying communities derives from a roles-and-permissions model in UNIX

> Mods & administrators: roles that have specific powers

What if you wanted to have more participatory governance? [Zhang, Hugh, and Bernstein 2020]

> (And even if you didn't, many online communities wind up needing to govern themselves regardless.)

# One failed angle: total open participation [Freeman 1970; Rheingold 2022]

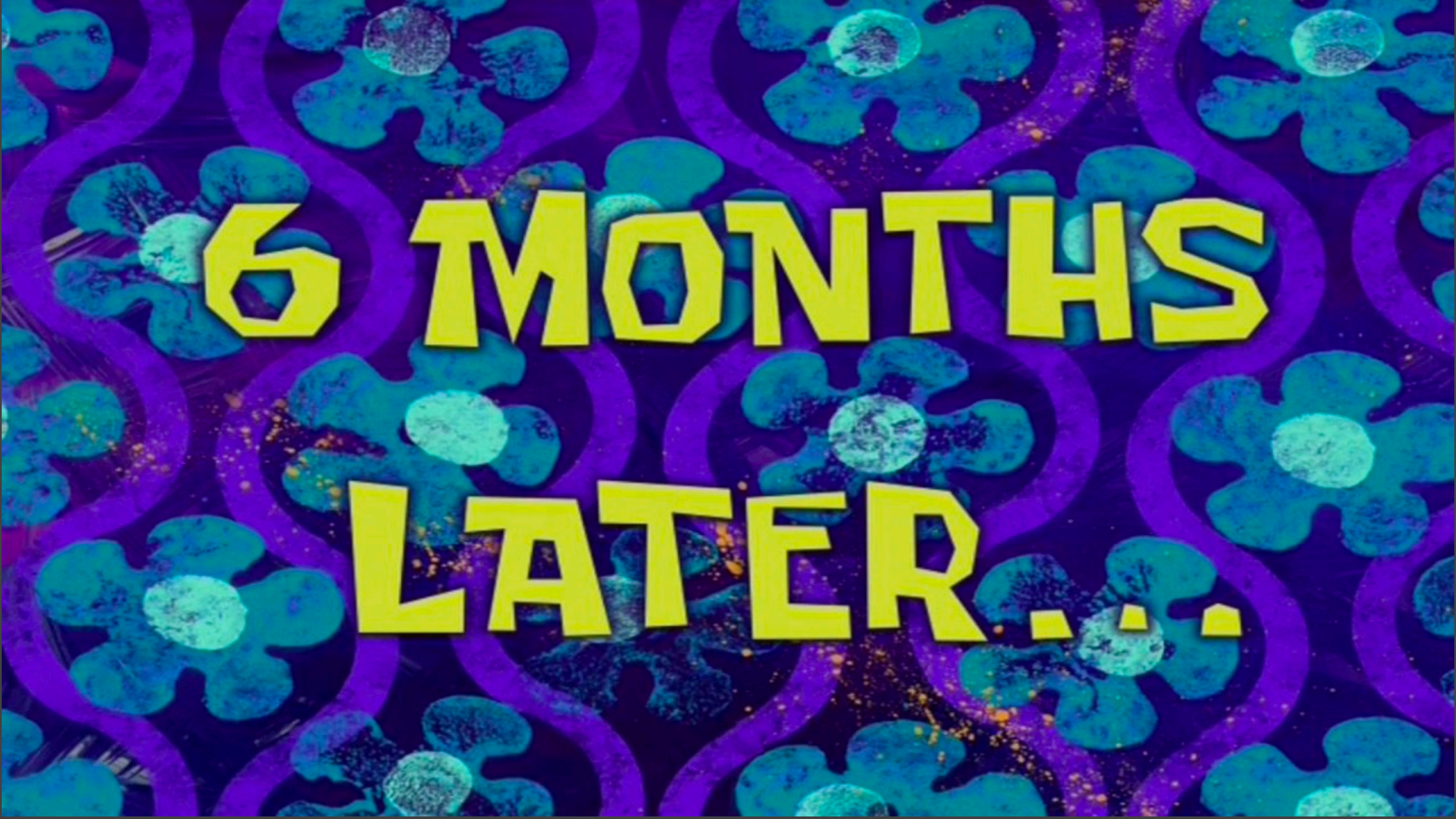
"I joined the network shortly before the Omidyars […] pledged to contribute an initial $25,000 to a group or groups chosen by the community.

Thus began a massive meta thrash. […] Probably around 1000 people checked out the conversation thread, around 200 participated in the conversation, and around 25 people contributed the majority of the posts, which ended up numbering in the thousands.

Attempts at consensus continued to fail on minority objections. My contributions to the conversation thread consisted largely of unheeded warnings that without a clear decision-making procedure, this conversation was doomed to be an infinite meta rathole."

**No governance is still governance — and it's bad governance.**

| | |
|---|---|
| Michael Bernstein | **Ugrad requirement proposal** |
| Mehran Sahami | **Re: Ugrad requirement proposal** |
| James Landay | **Re: Re: Ugrad requirement proposal** |
| Michael Bernstein | **Re: Re: Re: Ugrad requirement proposal** |
| Fei-Fei Li | **Re: Re: Re: Re: Ugrad requirement proposal** |
| Dorsa Sadigh | **Re: Re: Re: Re: Re: Ugrad requirement proposal** |
| Michael Bernstein | **Re: Re: Re: Re: Re: Re: Ugrad requirement proposal** |


6 MONTHS
LATER...

Michael Bernstein

Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:
Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:

Mehran Sahami

Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:
Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:

James Landay

Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:
Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:

Michael Bernstein

Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:
Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:

Fei-Fei Li

Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:
Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:

Dorsa Sadigh

Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:
Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:

Michael Bernstein

Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:
Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:

# Today: how do we govern?

| | |
|---|---|
| Mehran Sahami | **Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:** |
| James Landay | **Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:** |
| Michael Bernstein | **Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:** |
| Fei-Fei Li | **Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:** |
| Dorsa Sadigh | **Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:** |
| Michael Bernstein | **Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:** |

# Today: how do we govern?

Outline:

- Governance
- Decision-making

Not in scope for today: government regulation

| | |
|---|---|
| Fei-Fei Li | **Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:** |
| Dorsa Sadigh | **Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:** |
| Michael Bernstein | **Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:** |

# Governance

*"In democratic countries the science of association is the mother science; the progress of all the others depends on the progress of that one."*
–Alexis de Tocqueville, 1835

# Social media: the public sphere?

Social media as public sphere [Habermas 1989]: "Social media should be a place where we all gather publicly to rationally debate the issues that matter to us."

But not all communities have equal access or buy-in to the rational debate of the public sphere [Fraser 1990]

So how do we discuss and govern if the "town square" isn't practically achievable?

# Ostrom: governing the commons [Ostrom 1990]

When a resource is shared, the most effective groups require that "those affected by the rules can participate in modifying the rules"

In other words, successful collective decision-making requires:

1) Operational and collective rules, defining what we're allowed to do

2) Constitutional rules, the meta rules of how we change our governance

[Frey, Krafft, and Keegan 2019]

# Popular governance models

[https://opensource.guide/leadership-and-governance]

BDFL: "Benevolent Dictator for Life" who makes all final decisions.

Examples: Ethereum, Django, Swift, Ruby, Pandas, Ubuntu, Linux, SciPy, Perl

Reputation ladder: top contributors are granted decision making rights. Policy decisions via committee vote.

Issue: outspoken people are the ones who get credit

Examples: Red Hat, StackOverflow, Apache

Liberal contribution: allow as many contributors as possible, and use consensus-seeking for policy decisions

Examples: node.js and Rust

# Popular governance models

[https://opensource.guide/leadership-and-governance]

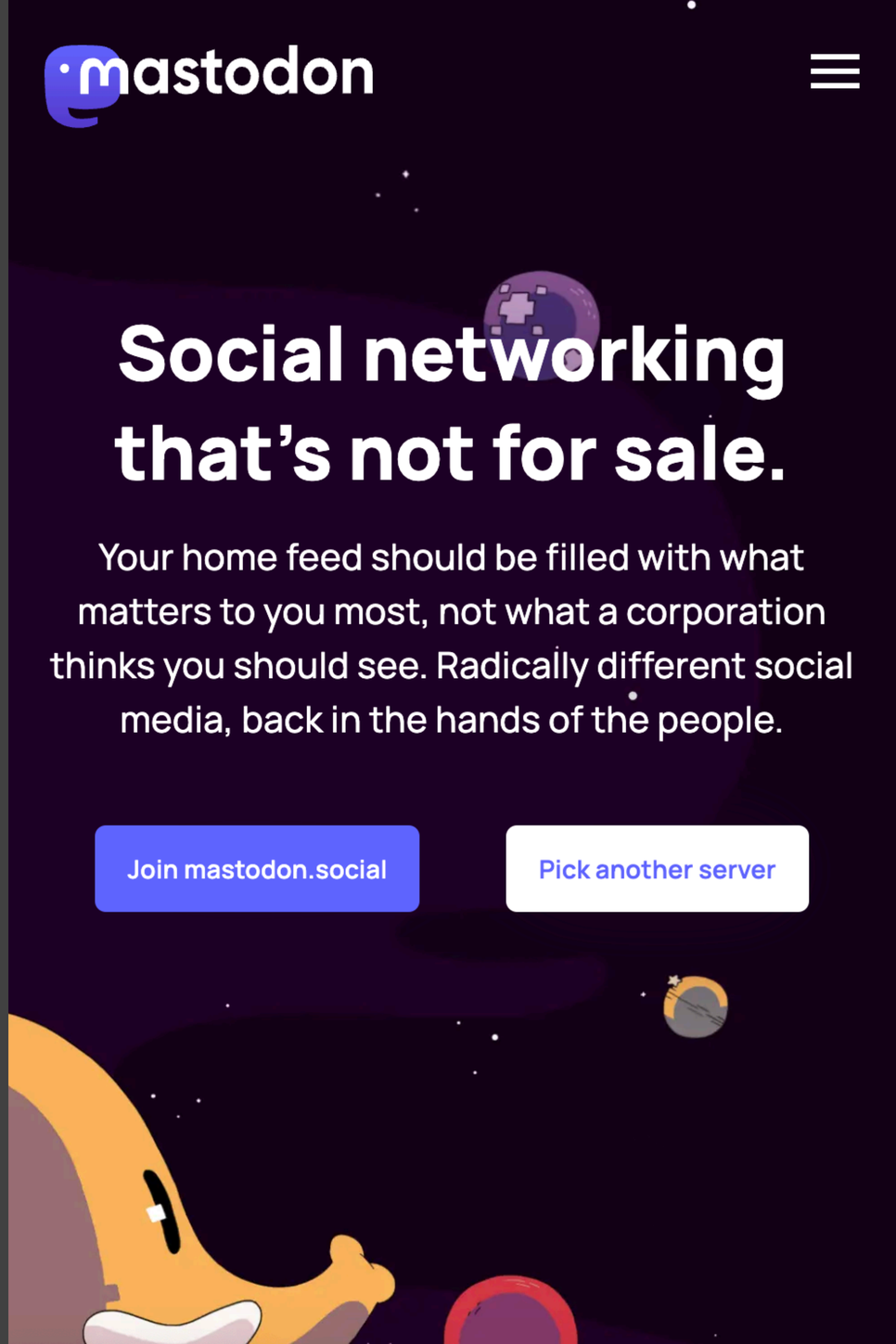
Steering committee: regular elections from active contributors produce a small committee empowered to make decisions when consensus isn't working

Examples: Python (post-Guido), node.js (for resolving technical disputes)

# Decentralization: is it an answer?

Mastodon and others stand for decentralization: individual servers run their own affairs, and pass content between each other via a protocol

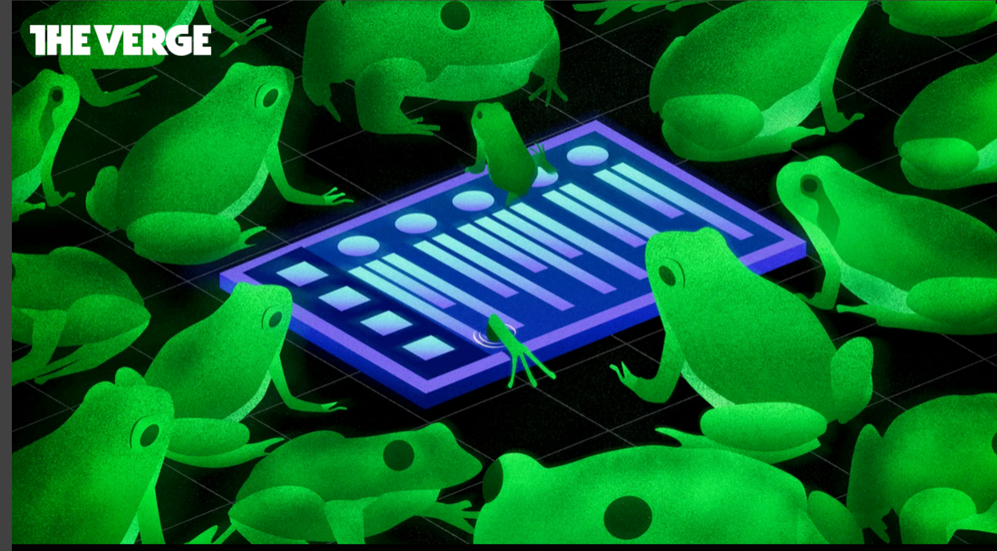
For example, a stanford.social Mastodon server could verify all accounts, and publish posts to other servers



# Decentralization: is it an answer?

The issue: when hate groups start a server, you can only block them from publishing to your server, not create any broader consequences for them

This is where federal governance is useful: the ability of a central government to maintain overall standards




WEB

## HOW THE BIGGEST DECENTRALIZED SOCIAL NETWORK IS DEALING WITH ITS NAZI PROBLEM

*Mastodon was built to be a kinder, more decentralized version of Twitter — then Gab showed up*

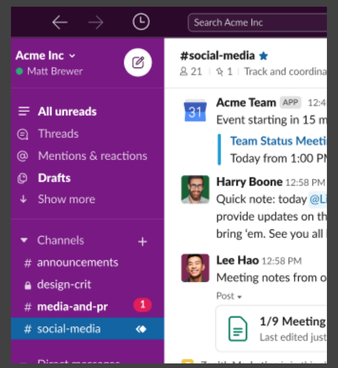
By Adi Robertson | @thedextriarchy | Jul 12, 2019, 2:51pm EDT

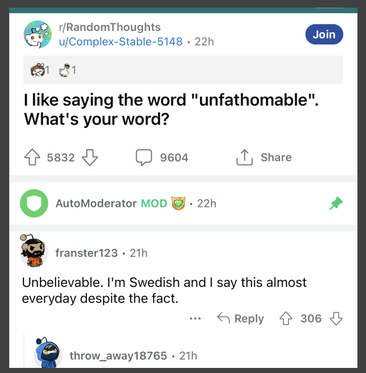
*Illustration by Alex Castro*

SHARE

Over the past few years, Mastodon has become the model for a friendlier kind of social network, promising to keep out the hateful or ugly content that proliferates on larger and more centralized networks. Journalists hailed it as "Twitter without Nazis" and for years, it's generally lived up to that promise. But last week, the social

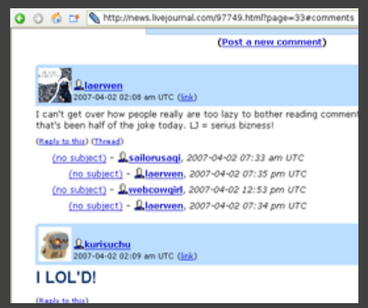
# Recall: From How is content delivered?



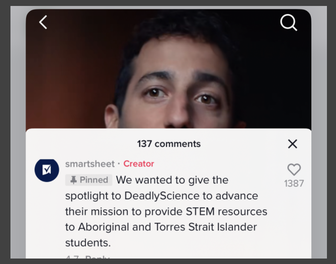

## Spaces

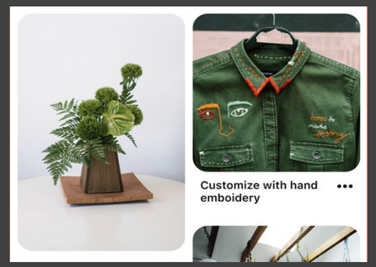
Joinable locations


## Network

Friend or follow relationships

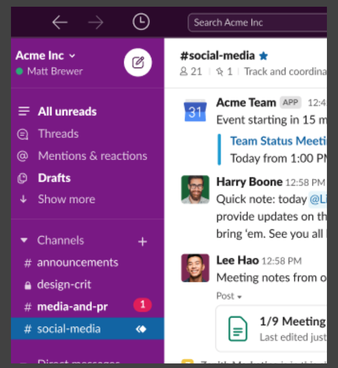


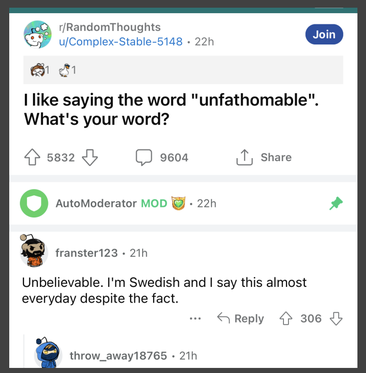




## Commons

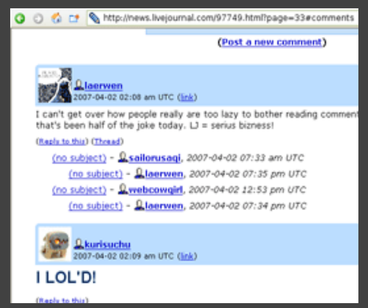
Drawn from entire platform

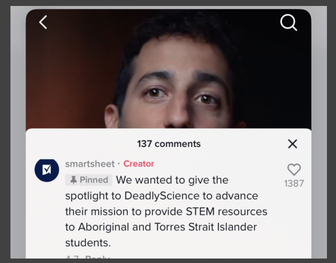
# Recall: From How is content delivered?

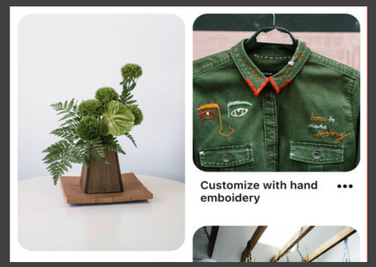







## Spaces

Local government and platform governance

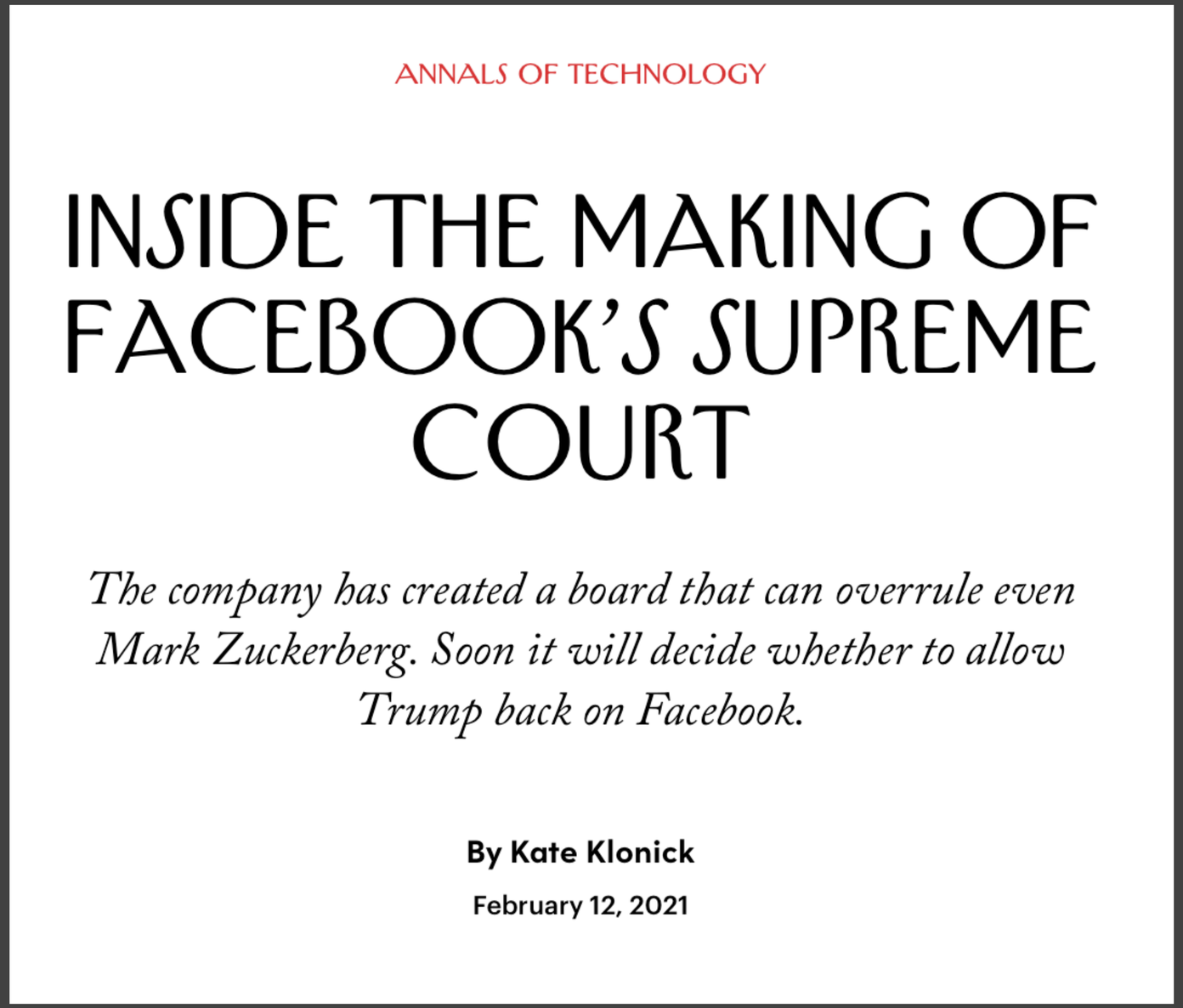
## Network

Platform governance only

## Commons

Platform governance only

One experiment: an independent judiciary body that reviews cases from Facebook and is intended to establish precedent in its moderation policies.



ANNALS OF TECHNOLOGY

# INSIDE THE MAKING OF FACEBOOK'S SUPREME COURT

*The company has created a board that can overrule even Mark Zuckerberg. Soon it will decide whether to allow Trump back on Facebook.*

**By Kate Klonick**

**February 12, 2021**

# Governance evolves over time

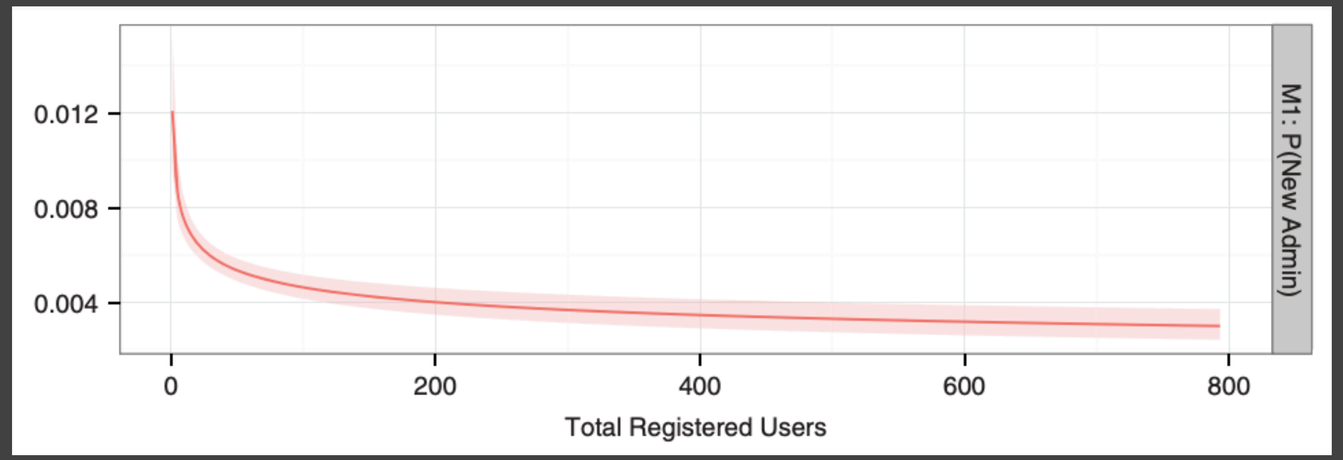
Self-governing communities often start out with very few rules and then grow in the number and complexity of their rules over time [Frey and Sumner 2019]

It's not clear that this is optimal: communities may be reacting to drama rather than planning

# Governance ossifies over time

While open-source projects and collaborative wikis sound very decentralized, in practice, leadership hierarchies emerge. [Benkler, Shaw and Hill 2015]

As a system grows, it's harder to become an admin [Shaw & Hill 2014]



# Resolving conflict: juries

When there is bad behavior, must we rely on mods? Can we empower a jury of your peers?

Two communities that use this approach:

- Sina Weibo: estimated 20,000–60,000 judges recruited from the user base who review cases of verbal abuse and personal attacks. About 2,000 expert judges review more complex cases such as rumor propagation.

- League of Legends: judges at The Tribunal (now defunct) reviewed cases of AFK flaming, harassment, racial slurs, and more

# Peer juries: complications

[Kou et al. 2017; Fan and Zhang 2020; Hu, Whiting and Bernstein 2021]

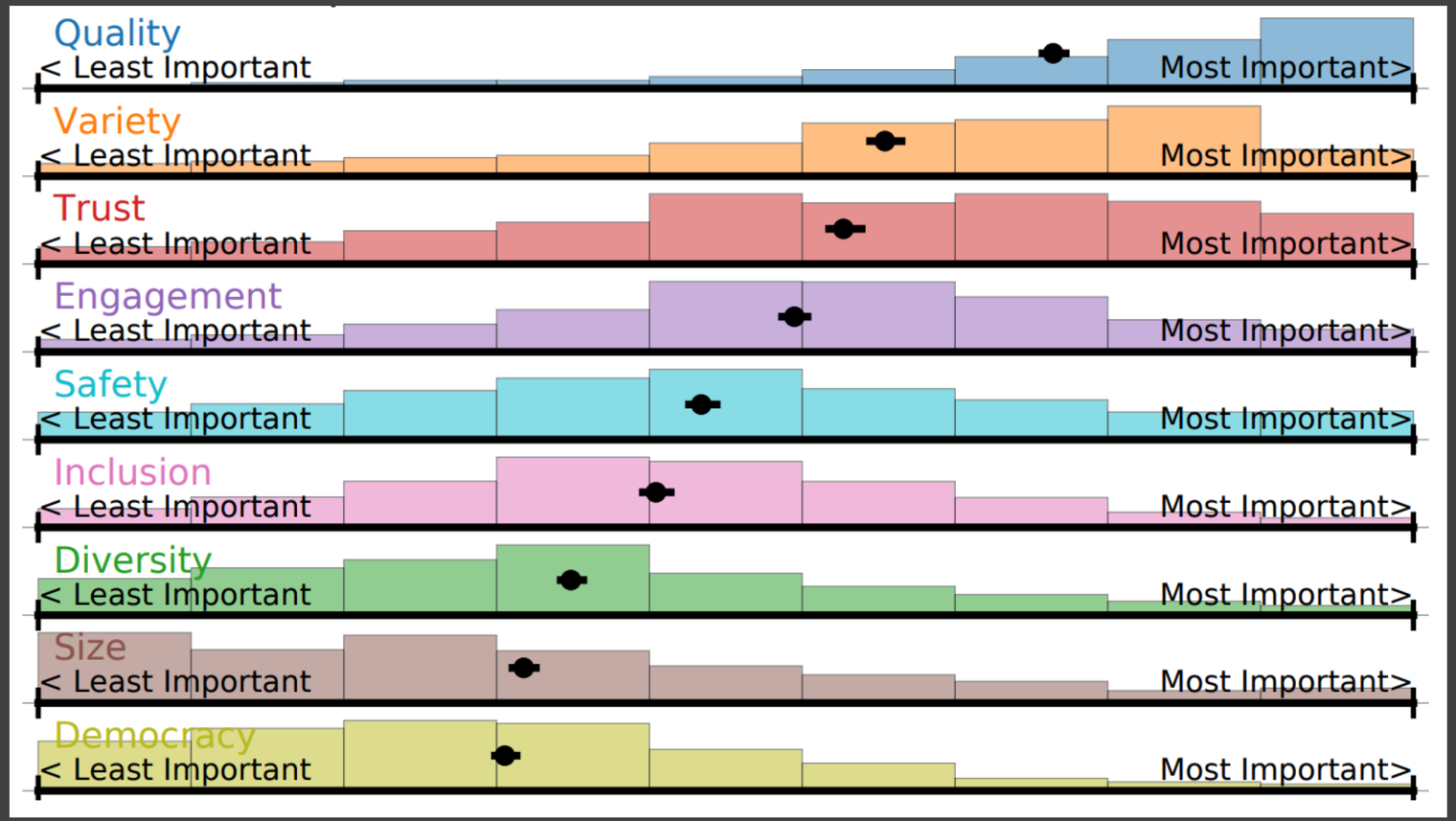
Users find the human-driven system more procedurally just than the platforms' decisions or than algorithmic systems, but still have limited trust in each other:

> "But why should I be judged by other ordinary Weibo users?"
>
> "As far as I know they just let random players make random decisions over whether a player can continue to play [League of Legends] or not."

Why is there less trust in these systems than in local, offline juries? What could be done about it? [1 min]

# However…

Surveying many Reddit communities about the most and least important values their community might express:




[Weld, Zhang, and Althoff 2022]

# And equally concerning…

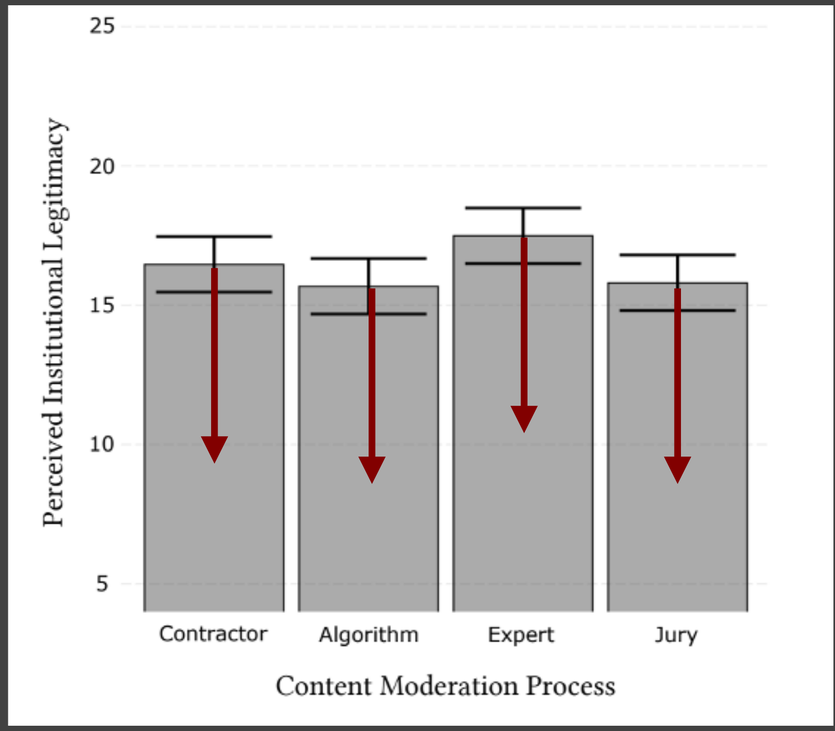
We ran a survey experiment comparing the legitimacy attached to moderation decisions made by AI vs. paid contractors vs. expert panels vs. digital juries… [Pan et al. 2022]

Which do you think was the most legitimate?

Our study result: expert panels (by ~1pt of 25)

But, whether people agreed with the result had much more influence (~7pt out of 25)

∴ It will be a severe challenge for any process to build legitimacy in a polarized media environment




7pt decrease if I disagree with the result

# Envisioning opportunities

Could we envision platforms that are…

- Constitutionally governed?
- Locally owned and governed?
- Nonprofits?
- B Corps?
- Public utilities?

Or do you think that the current model is the only sustainable one? [2min]

Metagov

# If you want to go deeper…

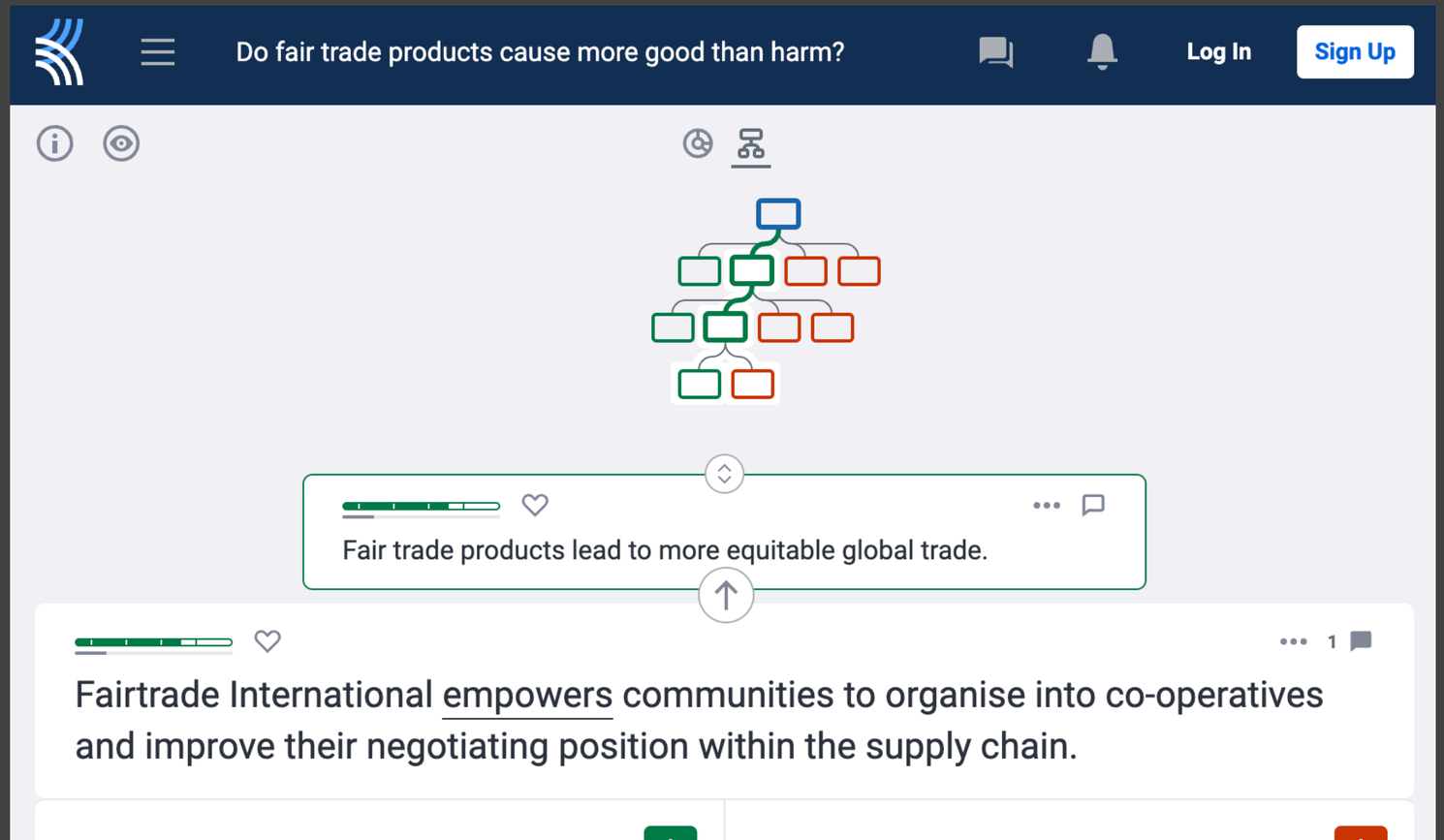
# Decision-making

# Why is this hard?

The same features that make it easy to gather online also make it easy to disperse. [Salehi et al. 2015]

# Structured debate

Deliberation: add metadata so that similar arguments get merged and replies get connected to the original argument




Kialo tree-structured debates

[Kriplean et al. 2012]

Statewide measures

< prev next >

# Should we License marijuana production, distribution, and possession?

## Top Cons


Dispensary growers and employees are still at risk from the Federal Government.

10/7/2012, 0 comments


Any revenue raised can be seized by the Federal Government.

10/7/2012, 0 comments


Pointless to legalize, the fed's won't let you have it (read more)

## Top Pros

Legalizing marijuana and taxing sales would generate more revenue for the state.

9/6/2012, 1 comment


"Creating" criminals for something that's no different than alcohol use, which is legal, ruins lives unnecessarily and burdens society

10/5/2012, 2 comments

# pol.is


I feel that as a creator, the content I generate using AI should be regarded as my copyright.

I think AI should label training data before model training to reduce model copying and social bias.

I think AI should prevent the presentation of hate speech

debatewise.org

Search Debates Go

Browse Our Categories


# GMOs Are Good

Last updated: March 7, 2019

In its broadest sense, a genetically modified organism (GMO) refers to an organism whose genetic material has been altered in a way that does not occur naturally. Genes change naturally, either by mating or by natural recombination. But in this case, genetic fragments are scientifically inserted into the DNA of another organism to transform its collective genetic makeup, a process known as gene splicing. In plant and animal biotechnology, there are three fundamental areas that genetic modification is concerned with; the quality of

**Yes because...**

### *Production of Edible Medicine and Vaccines*

This is perhaps the most innovative application of GMO. This process involves the modification of animal and plant genes in order to yield edible output with preventive molecules, for example milk, eggs and fruit. Edible vaccines, produced in milk or fruit could ease manufacturing and distribution costs by making it globally accessible to people. Vaccination through injection has many disadvantages, including the need for medically trained staff, high costs, not to mention constant cooling during transportation and storage. Use of needles also increases the risk of infections. In these cases, an edible vaccine comes in handy. In recent years, there have been examples of transgenic plants developed by researchers to help developing countries. Transgenic potatoes which contain cholera toxins have been developed to immunize against diseases. In 2004, the Pharma-Planta Programme was granted 12 million euros to develop genetically modified plants to help grow vaccines against tuberculosis and rabies.

**No because...**

It is important to note that development of edible vaccines is still at very early stages. So far, the benefits outlined are only human-centric. It has not been articulated clearly how implementation of this

Are these designs enough to craft decisions? If not, what would it take?
[1 min]

# No. Back to this situation…

| | |
|---|---|
| Michael Bernstein | **Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:** |
| Mehran Sahami | **Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:** |
| James Landay | **Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:** |
| Michael Bernstein | **Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:** |
| Fei-Fei Li | **Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re: Re:** |


Losing momentum, no viable path

Outright flaming or violent disagreement

between scylla and charybdis…

[Salehi et al. 2015]

# Michael's take

Adding metadata to discussion is helpful usability-wise, but is no panacea.

In contrast, structuring the rules and roles by which we're able to engage with each other is much more likely to produce productive deliberation.

Most online communication tools such as email fail at deliberation because they don't structure those rules and roles. We just continue to ricochet from stalling to friction and back.

# Voting

Idea 1

Idea 2

Idea 3

Idea 4

Idea 5

"Vote on your top two ideas"

Strengths: simple user model, useful for selecting a single best option

Weaknesses: known pathological cases (instant runoff voting improves), not great for producing a ranking

# Liquid democracy

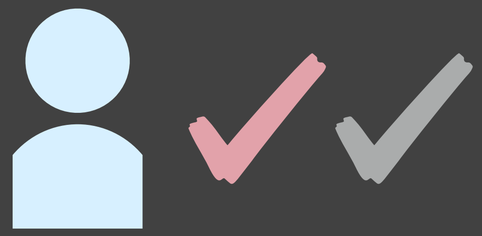
Idea 1

Idea 2 ✓

Idea 3 ✓✓✓

Idea 4

Idea 5



I can vote directly, or delegate my vote to a person or institution who I think knows more about the issue. They can then either vote or delegate their own votes.

# Liquid democracy

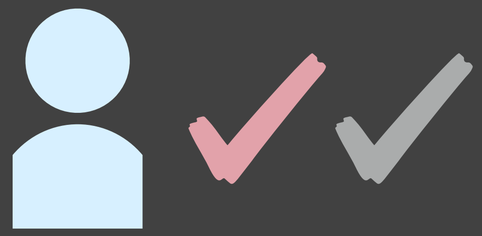
Idea 1

Idea 2 ✓

Idea 3 ✓✓✓

Idea 4

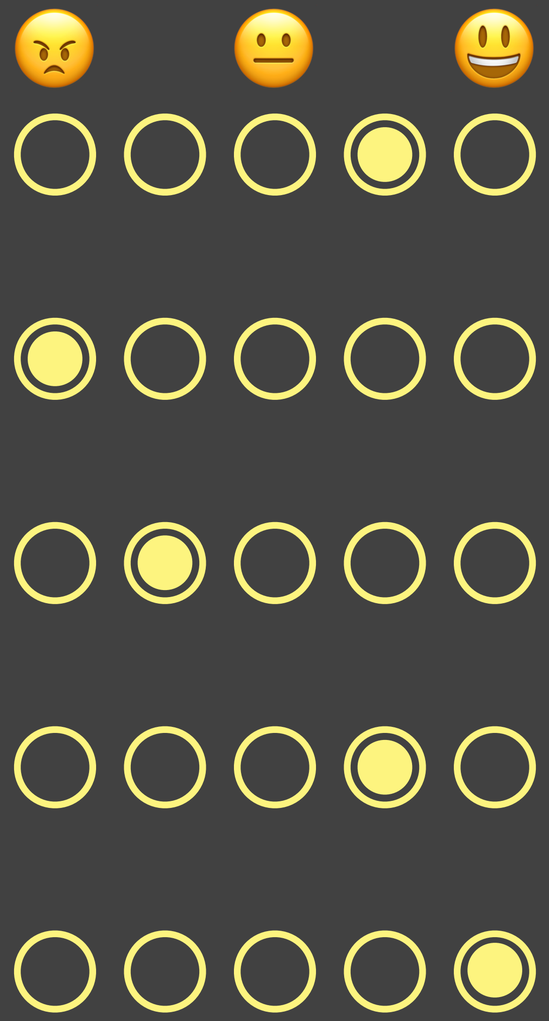
Idea 5



Benefits: compromise between direct and representative democracy; made feasible by the web.

Weaknesses: a small number of people can gain massive influence [Kahng, Mackenzie, and Procaccia 2018]

# Likert Scale Rating

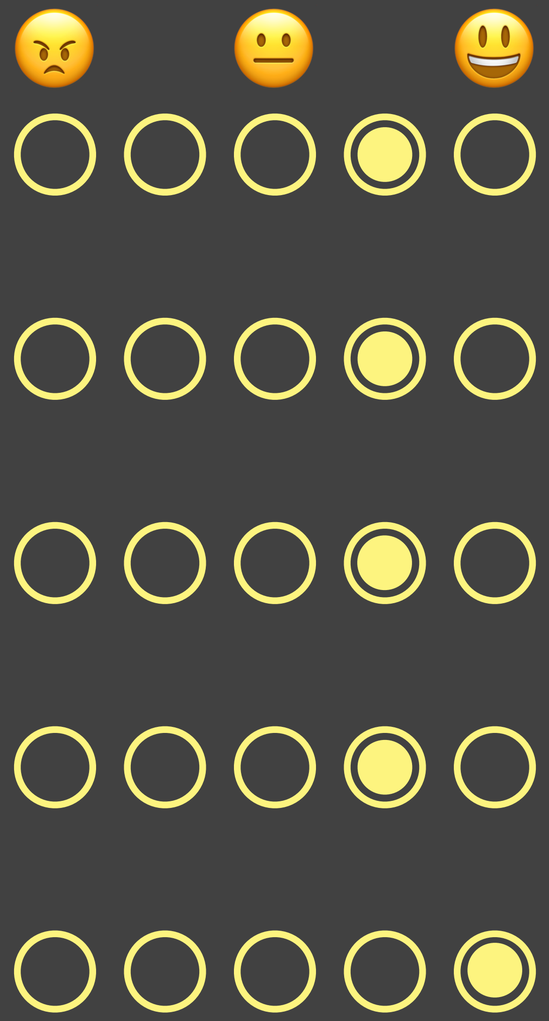



"Rate each idea"

Strengths: gets more information per idea, allows ranking

Weaknesses: people tend to use the scale differently

# Likert Scale Rating

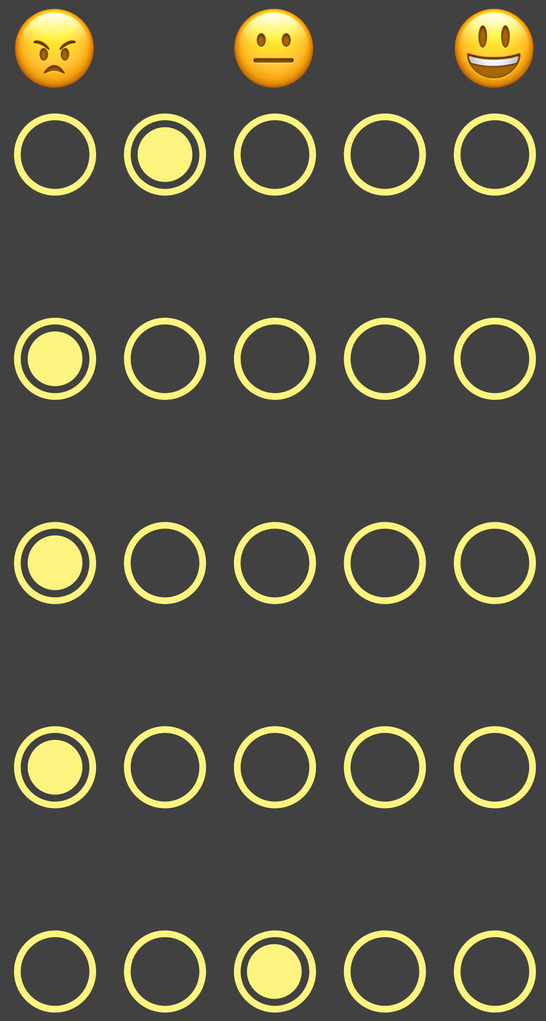



"Rate each idea"

Strengths: gets more information per idea, allows ranking

Weaknesses: people tend to use the scale differently (some are nice)

# Likert Scale Rating

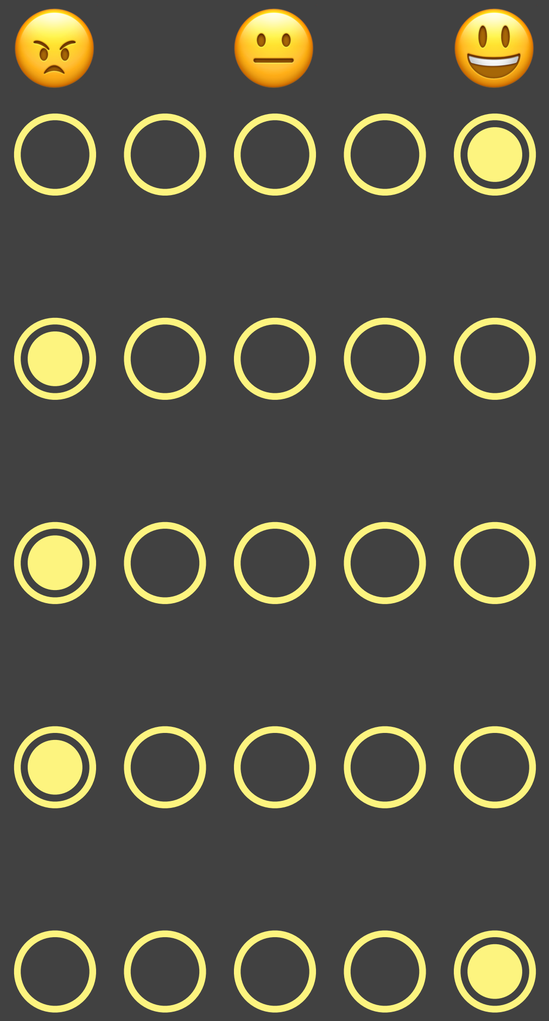



"Rate each idea"

Strengths: gets more information per idea, allows ranking

Weaknesses: people tend to use the scale differently (some are nice, some are mean)

# Likert Scale Rating

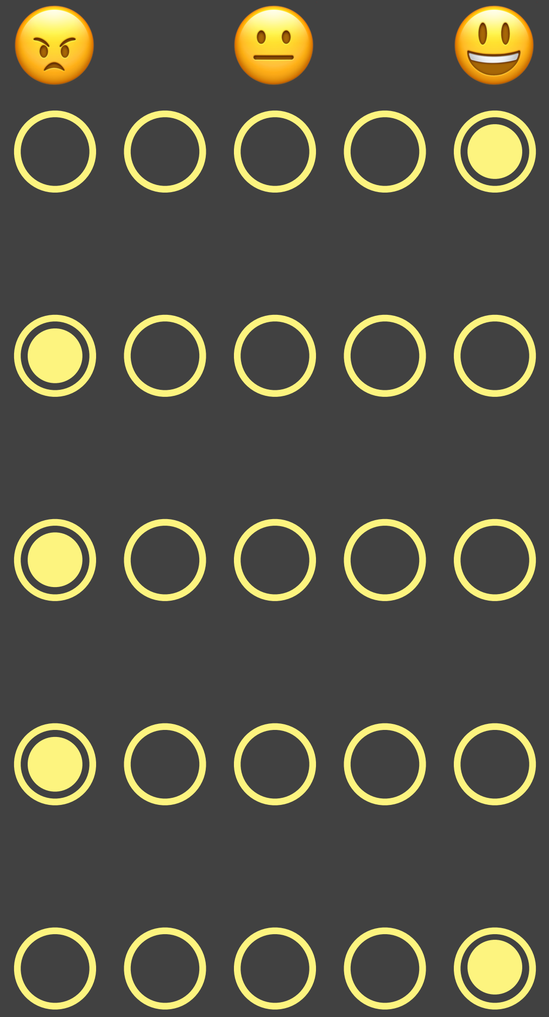



"Rate each idea"

Strengths: gets more information per idea, allows ranking

Weaknesses: people tend to use the scale differently (some are nice, some are mean, many are extreme)

# Likert Scale Rating

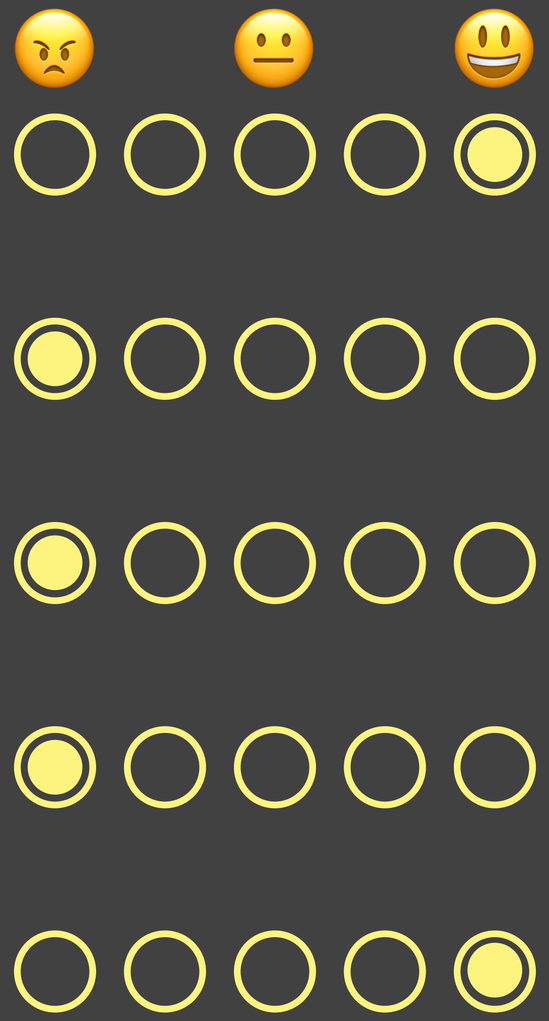



"Rate each idea"

Strengths: gets more information per idea, allows ranking

Weaknesses: people tend to use the scale differently (some are nice, some are mean, many are extreme), we have limited resolution into the differences between the 5s

# Likert Scale Rating



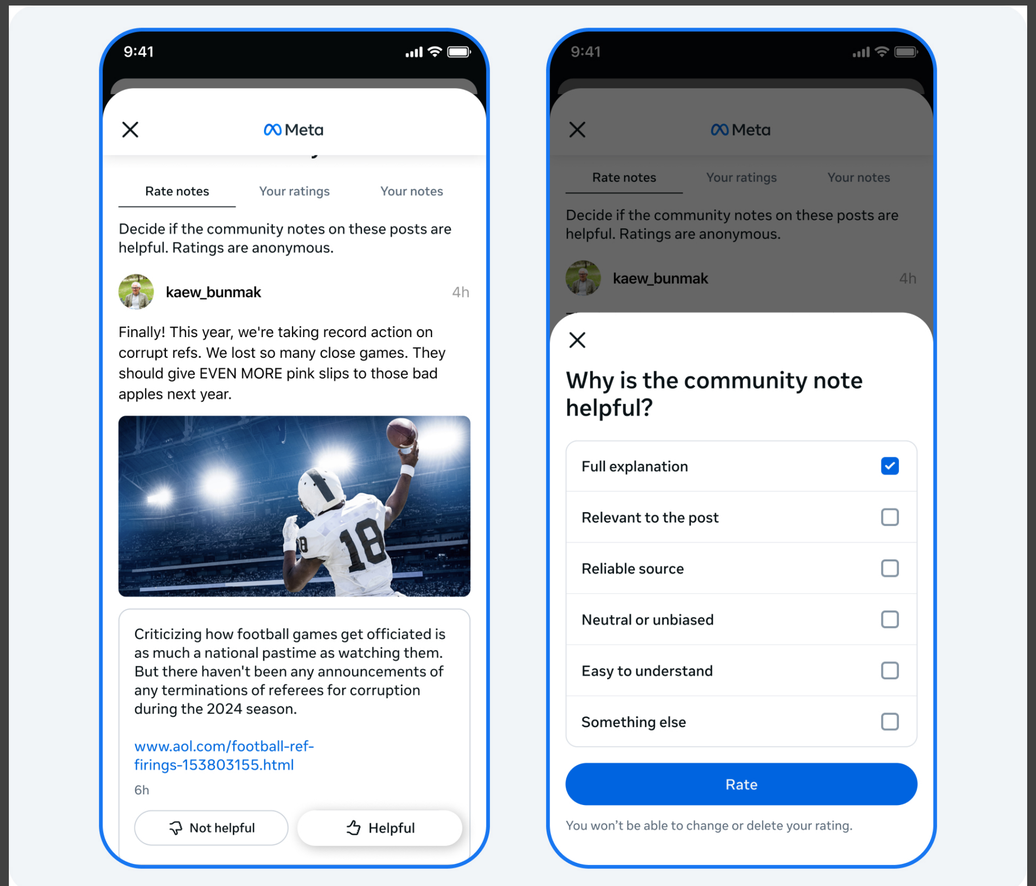

**6. Pho Avenue**
375 reviews
$$ · Vietnamese

**3. Pho Vi Hoa**
1073 reviews
$$ · Vietnamese

**9. Hometown Noodle**
564 reviews
$ · Vietnamese, Noodles

As a result, not a ton of signal to use to tell these restaurants apart on Yelp.

# One approach: bridging algorithms




Anyone can author a note, and anyone can vote.

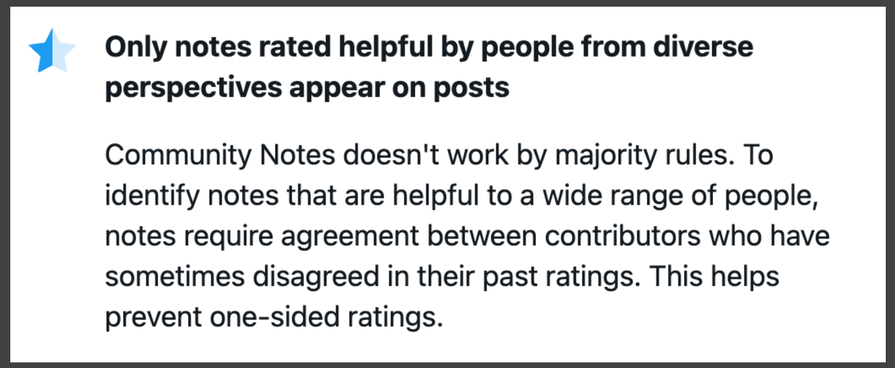
Helps platforms dodge bias claims:

> **X doesn't choose what shows up, the people do**
>
> X doesn't write, rate or moderate notes (unless they break X's Rules.) We believe giving people a voice to make these choices together is a fair and effective way to add information that helps people stay better informed.

But how do we avoid communities just brigading each others' posts? [Allen, Martel, Rand 2022]

# One approach: bridging algorithms

Community Notes collects votes from different people on different topics




They run a matrix factorization algorithm (e.g., PCA) to "flatten" all these votes into a small number of dimensions that capture most disagreement. The first dimension that arises is typically politics. The ootes that appear

live have to appeal broadly across that first dimension [Wojcik et al. 2022]: this is called a bridging algorithm because it identifies content

# Comparison ranking

Which of these two ideas
do you prefer?

**Idea 1**

**Idea 2**

# Comparison ranking

Which of these two ideas
do you prefer?

**Idea 4**

**Idea 3**

# Comparison ranking

Which of these two ideas
do you prefer?

**Idea 1**

**Idea 3**

# Comparison ranking

## Which of these hard difficulty questions is a better midterm question?

*NOTE: Some submissions have links. Please follow them and give each question the time and respect you would want given to your own.*

Some messaging apps show you when the other person is typing. What is the name of this design pattern? What are the two requirements needed to make it successful?

The left one

Briefly explain the concept of honest signals. Provide a design example of a social computing system that maintains and deepens strong ties through fostering honest signals from its users. In your reply, outline the type of user behaviours that would take place within the system.

The right one

# Comparison ranking

RLHF pipeline

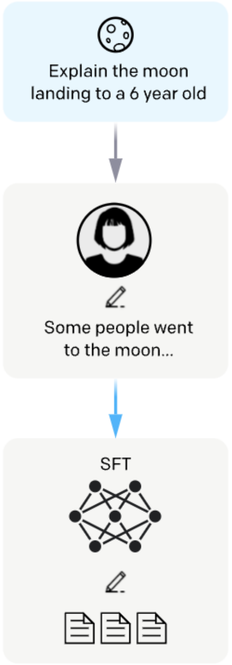
Step 1

**Collect demonstration data, and train a supervised policy.**

A prompt is sampled from our prompt dataset.

A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT-3 with supervised learning.



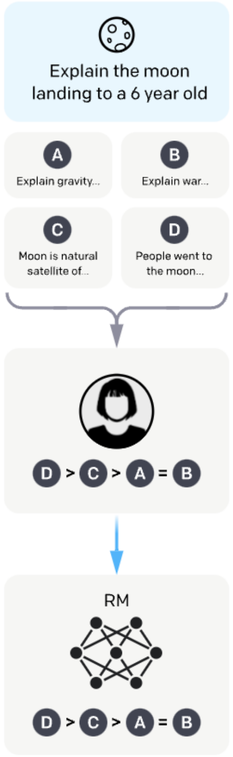

Step 2

**Collect comparison data, and train a reward model.**

A prompt and several model outputs are sampled.

A labeler ranks the outputs from best to worst.

This data is used to train our reward model.




Step 3

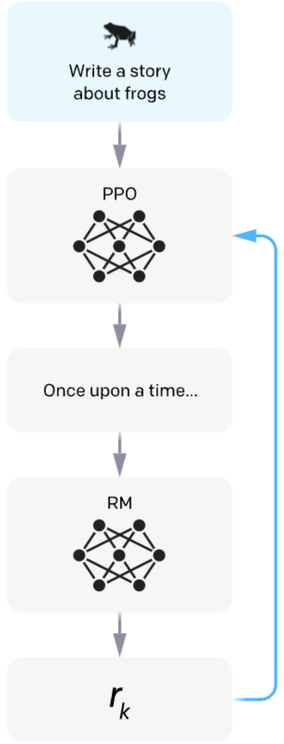
**Optimize a policy against the reward model using reinforcement learning.**

A new prompt is sampled from the dataset.

The policy generates an output.

The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.



# Comparison ranking

But how do we turn a bunch of comparisons into a score or ranking per item?

Intuition:

- If I beat something that's known to be low ranked, I must not be terrible.
- If I beat something that's known to be high ranked, I must be really good.

But how do I know what's low ranked and what's high ranked?

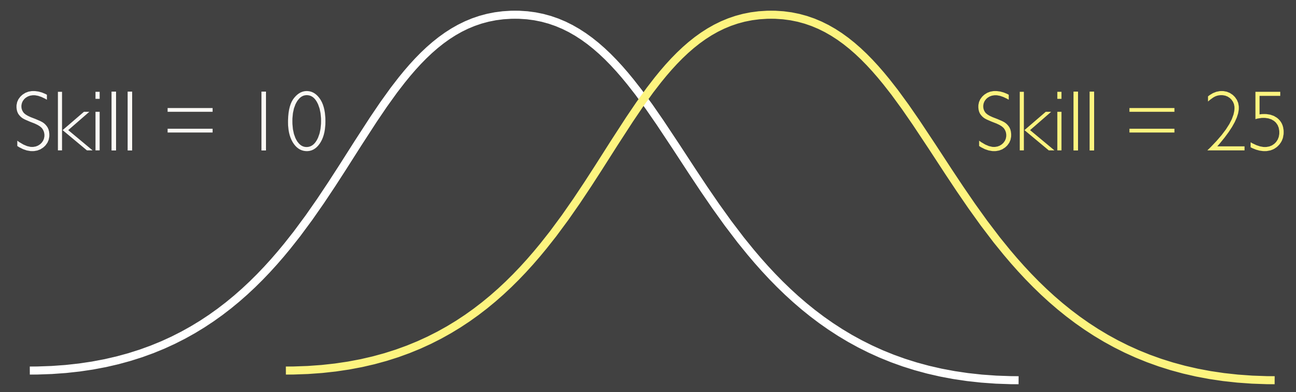
# TrueSkill and Elo

Elo is the system that was developed to rank chess players based on their win-loss records against each other.


Imagine that each player's performance across a number of games is normally distributed. Sometimes they play amazingly, sometimes less so. Our goal is to estimate the mean of each player's distribution. Each game is a draw from the players' distributions.

# TrueSkill and Elo

Intuitively, in Elo, we have some belief in the skill of each player before they play each other, and we update that belief based on the result of the game.




If white beats yellow, white's skill score is updated by a multiplier $\alpha$ of

$$\alpha(25 - 10) = \alpha \cdot 15$$

$\alpha$ is tuned on how quickly the score should adapt based on recent games

# TrueSkill and Elo

In TrueSkill [Herbrich, Minka, and Graepel 2006], the same general idea holds, except the entire algorithm is done by performing Bayesian inference on a generative model

$$p(skill \mid result) = \frac{p(result \mid skill) \cdot p(skill)}{p(result)}$$

← Bayes' rule

# Summary

Social computing systems are great at eliciting a lot of opinions, but generally terrible at helping produce consensus toward a decision.

Deliberation is challenging because there are no stopping criteria. Structuring the rules of the debate can help overcome stalling and friction.

There are many popular governance models — but the most important thing is to pick one from day one

Different elicitation methods such as voting, liquid democracy, rating, and comparison ranking provide possible solutions for single choices

# Pulling it all together

# Thank you, TAs!

# Congrats, seniors

(About one third of you are seniors)

# What I hope you take away

That every social system is designed, either explicitly or by default.

That designs can have substantial influence over the behaviors in that system.

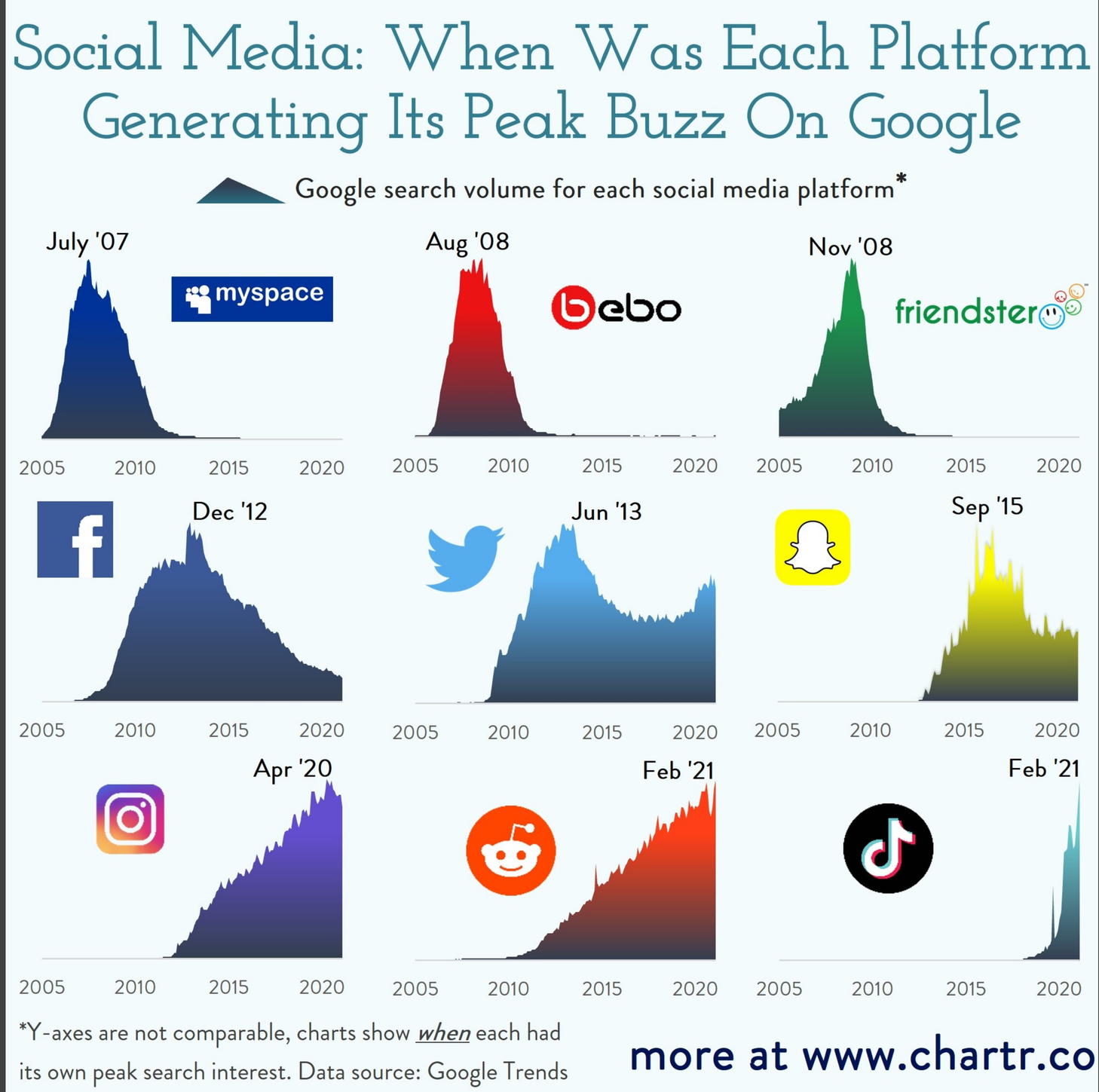
That, as socio-technical systems, those designs require a combination of computation and of structured human behavior to succeed.

That we have many tools in our toolbox to help us create enlightening, fun, and meaningful spaces.

The more things change, the more they stay the same.

I, a geriatric millennial, am not cool. Ergo, systems I liked are no longer cool. Ergo, systems you like will soon no longer be cool.

But the fundamentals will remain.



# How can we design the social systems that we inhabit?

# fin.

have a great summer!

# References


Allen, Jennifer, Cameron Martel, and David G. Rand. "Birds of a feather don't fact-check each other: Partisanship and the evaluation of news in Twitter's Birdwatch crowdsourced fact-checking program." Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022.

Barlow, John Perry. "A Declaration of the Independence of Cyberspace." Electronic Frontier Foundation. 1996.

Benkler, Yochai, Aaron Shaw, and Benjamin Mako Hill. "Peer production: A form of collective intelligence." Handbook of collective intelligence 175 (2015).

De Tocqueville, Alexis. Democracy in america. 1835.

Fan, Jenny, and Amy X. Zhang. "Digital juries: A civics-oriented approach to platform governance." Proceedings of the 2020 CHI conference on human factors in computing systems. 2020.

Fraser, Nancy. "Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy." Social Text, no. 25/26, 1990, pp. 56–80. JSTOR, https://doi.org/10.2307/466240. Accessed 26 May 2024.

Freeman, Jo. "The tyranny of structurelessness." Berkeley Journal of Sociology (1972): 151-164.

Frey, Seth, and Robert W. Sumner. "Emergence of integrated institutions in a large population of self-governing communities." PloS one 14.7 (2019): e0216335.

# References


Frey, Seth, P. M. Krafft, and Brian C. Keegan. "" This Place Does What It Was Built For" Designing Digital Institutions for Participatory Change." Proceedings of the ACM on Human-Computer Interaction 3.CSCW (2019): 1-31.

Goelz, Paul, et al. "The fluid dynamic of liquid democracy." WINE 2018 (2018).

Habermas, Jurgen. The structural transformation of the public sphere: An inquiry into a category of bourgeois society. MIT press, 1991.

Herbrich, Ralf, Tom Minka, and Thore Graepel. "TrueSkill™: a Bayesian skill rating system." Advances in neural information processing systems 19 (2006).

Hu, Xinlan Emily, Mark E. Whiting, and Michael S. Bernstein. "Can Online Juries Make Consistent, Repeatable Decisions?." Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021.

Klonick, Kate. "The new governors: The people, rules, and processes governing online speech." Harv. L. Rev. 131 (2017): 1598.

Kou, Yubo, et al. "Managing disruptive behavior through non-hierarchical governance: Crowdsourcing in League of Legends and Weibo." Proceedings of the ACM on Human-Computer Interaction 1.CSCW (2017): 1-17.

Kriplean, Travis, et al. "Supporting reflective public thought with considerit." Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work. 2012.

# References


Ostrom, Elinor. Governing the commons: The evolution of institutions for collective action. Cambridge university press, 1990.

Pan, Christina A., et al. "Comparing the Perceived Legitimacy of Content Moderation Processes: Contractors, Algorithms, Expert Panels, and Digital Juries." Proceedings of the ACM on Human-Computer Interaction 6.CSCW1 (2022): 1-31.

Rheingold. Howard. "Online Governance: Six Case Studies". https://hrheingold.medium.com/online-governance-six-case-studies-177bfbb3f541

Salehi, Niloufar, et al. "We are dynamo: Overcoming stalling and friction in collective action for crowd workers." Proceedings of the 33rd annual ACM conference on human factors in computing systems. 2015.

Shaw, Aaron, and Benjamin M. Hill. "Laboratories of oligarchy? How the iron law extends to peer production." Journal of Communication 64.2 (2014): 215-238.

Weld, Galen, Amy X. Zhang, and Tim Althoff. "What Makes Online Communities 'Better'? Measuring Values, Consensus, and Conflict across Thousands of Subreddits." Proceedings of the International AAAI Conference on Web and Social Media. Vol. 16. 2022.

Wojcik, Stefan, et al. "Birdwatch: Crowd wisdom and bridging algorithms can inform understanding and reduce the spread of misinformation." arXiv preprint arXiv:2210.15723 (2022).

Zhang, Amy X., Grant Hugh, and Michael S. Bernstein. "PolicyKit: building governance in online communities." Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology. 2020.

# Social Computing

CS 278 | Stanford University | Michael Bernstein

Creative Commons images thanks to Kamau Akabueze, Eric Parker, Chris Goldberg, Dick Vos, Wikimedia, MaxPixel.net, Mescon, and Andrew Taylor.