Peer Production

CS 278 | Stanford University | Michael Bernstein

Announcements

Project milestone due Thursday by 4pm

Exam question bank, based on your votes, is coming out soon

Exam is one week from today: 3/4 questions in the question bank, 1/4 staff-written questions that cover through this week

Last time

Crowdsourcing: an open call to a large group of people who self-select to participate

Crowds can be surprisingly intelligent, if opinions are levied with some expertise and without communication, then aggregated intelligently.

Design differently for intrinsically and extrinsically motivated crowds

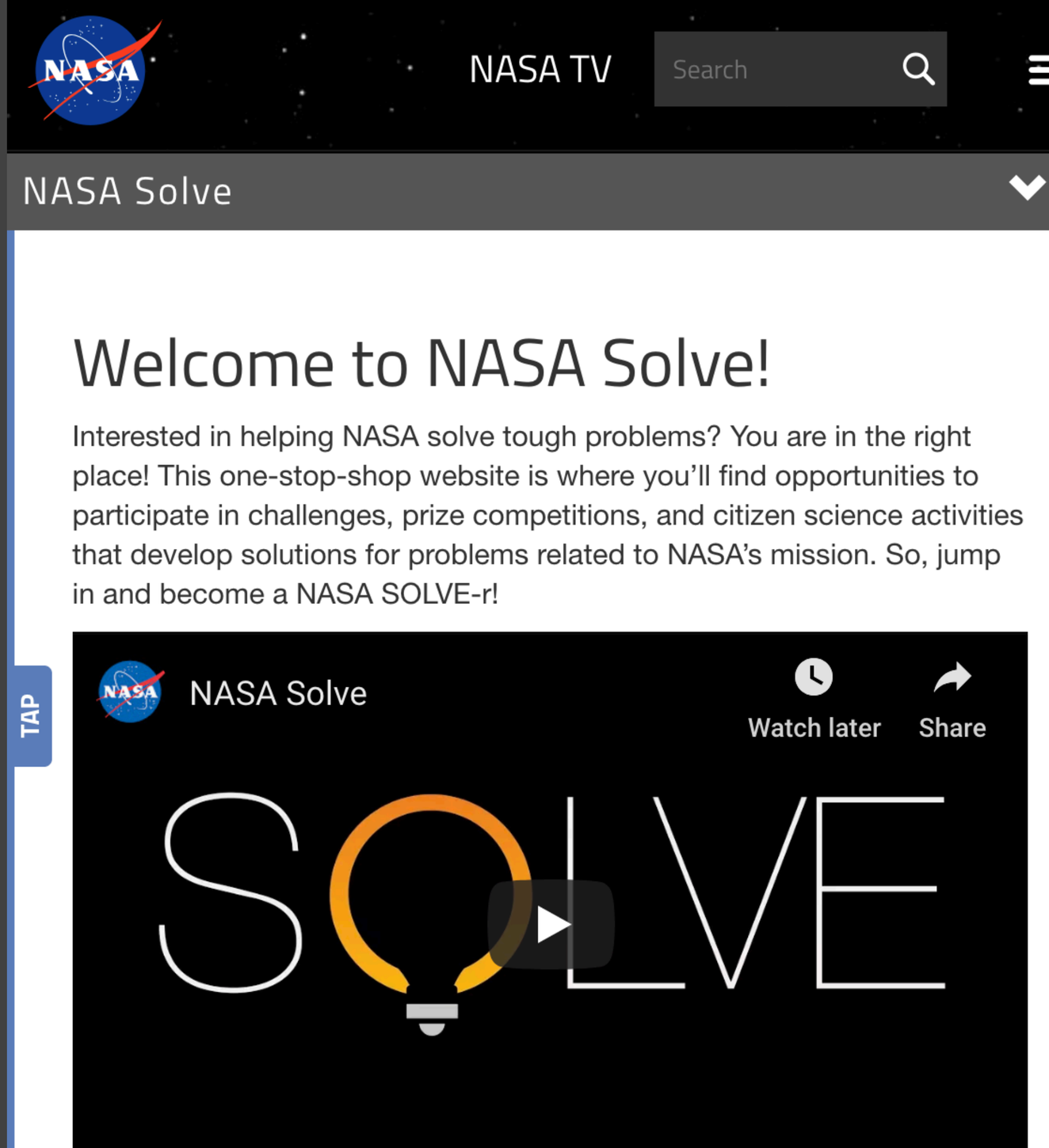
Vandalism—much like other anti-social behavior—is rare, but can happen

Last time

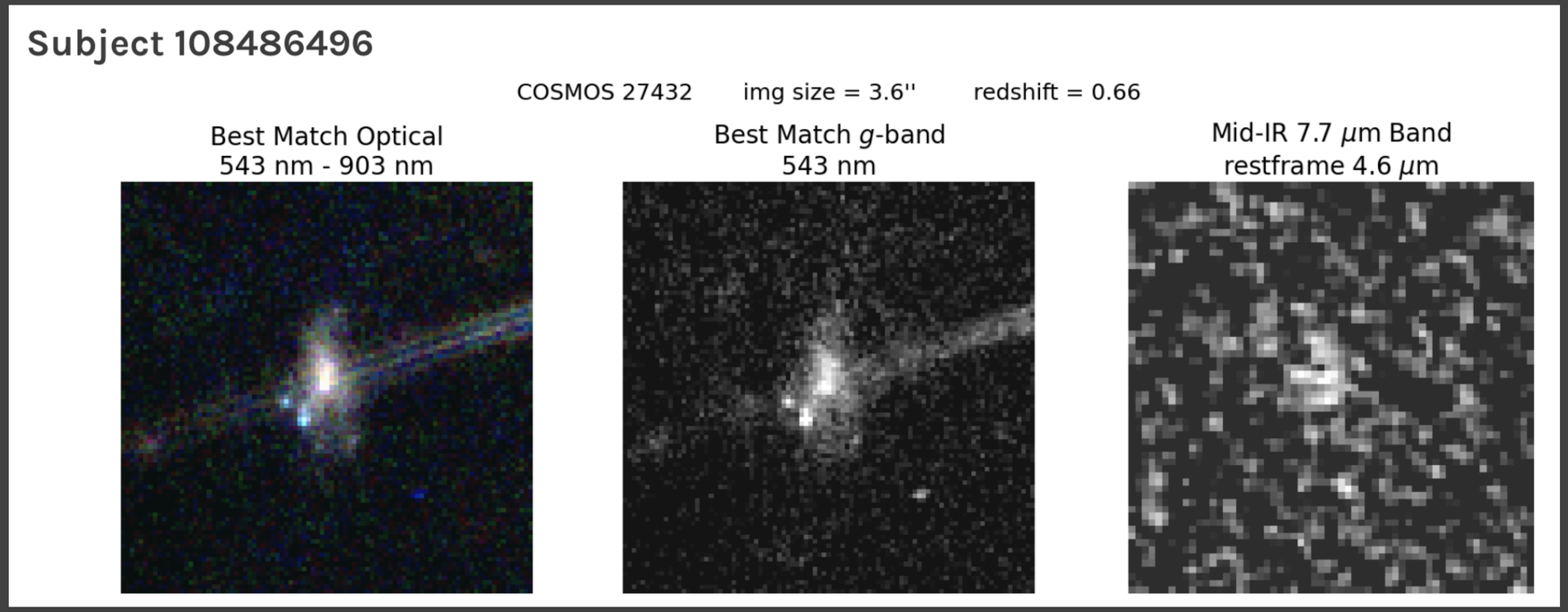
Parallel, independent contributions

But, this only works if the goal can be subdivided into modular components with few or no interdependencies.

Think filling out rows of a spreadsheet or taking argmax



"Crowdsourcing" example submitted by Irene A. Lin



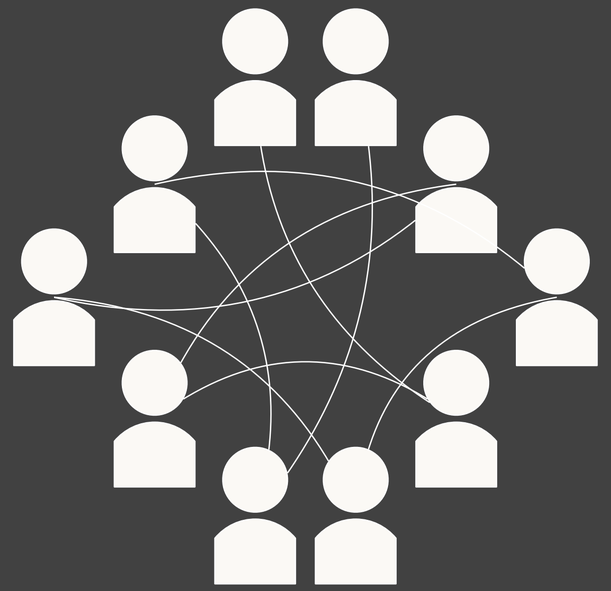
Galaxy Zoo is an online crowdsourcing project where anyone can help classify galaxies. Within 24 hours of launch, volunteers were submitting 70,000 classifications per hour. The platform has led to major discoveries and some users have even become co-authors!

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

Attendance

Today

Interdependent, integrated contributions



Think invention, engineering, or design.

How?

There are fundamental differences between parallel and interdependent contribution structures.

We can’t just make a movie or build Linux with parallel contributions.

Johnny Cash Project: crowdsourced music video
One frame per participant — beautiful, slightly anarchic

Star Wars Uncut: crowdsourced movie remake, 2hr long
One scene per participant — style whiplash

How?

If we can't do it with parallel contributions, how do we create complex outcomes with distributed online collaborations?

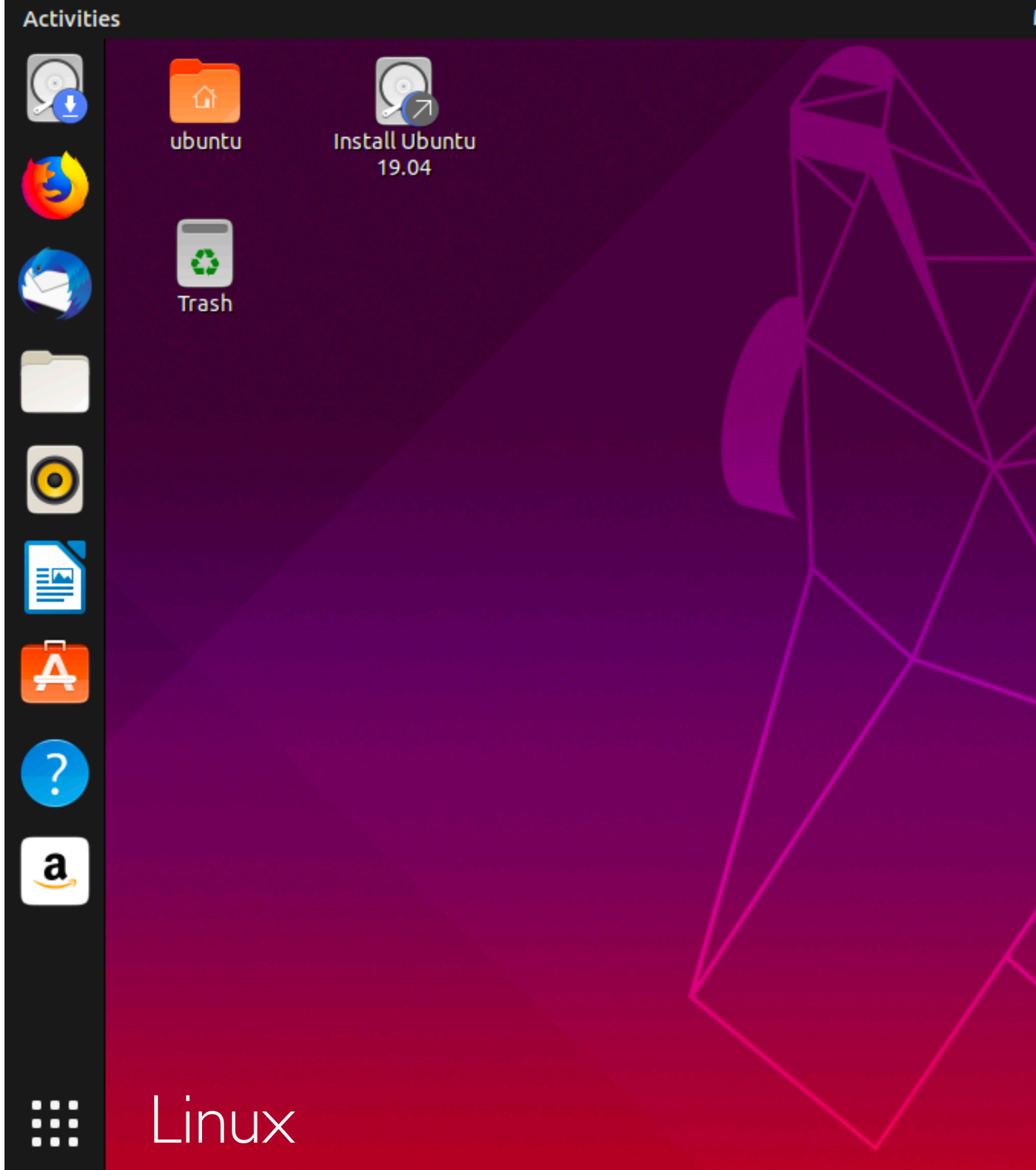
Topics:

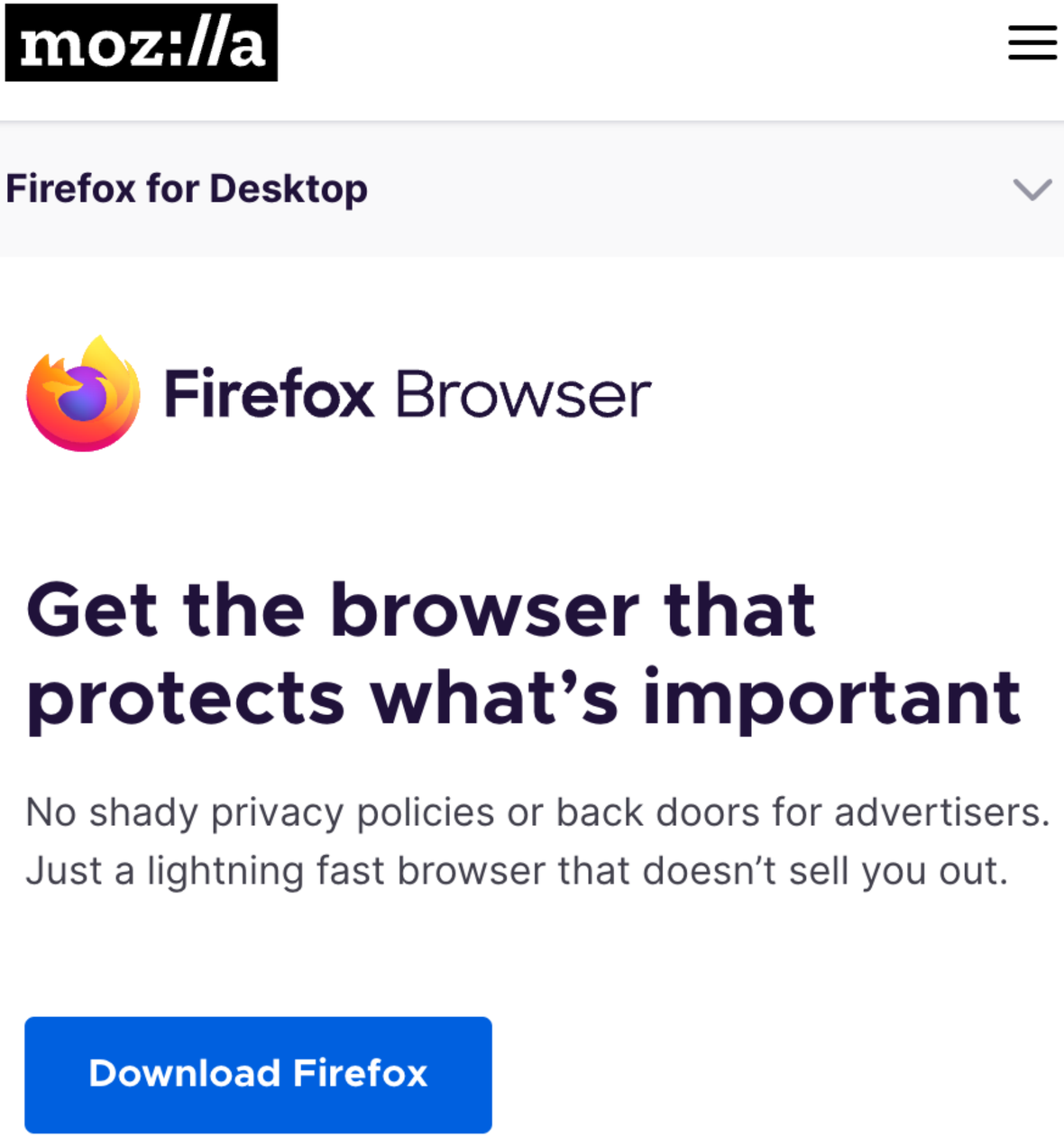
- Peer production
- Workflows
- Convergence and coordinated adaptation

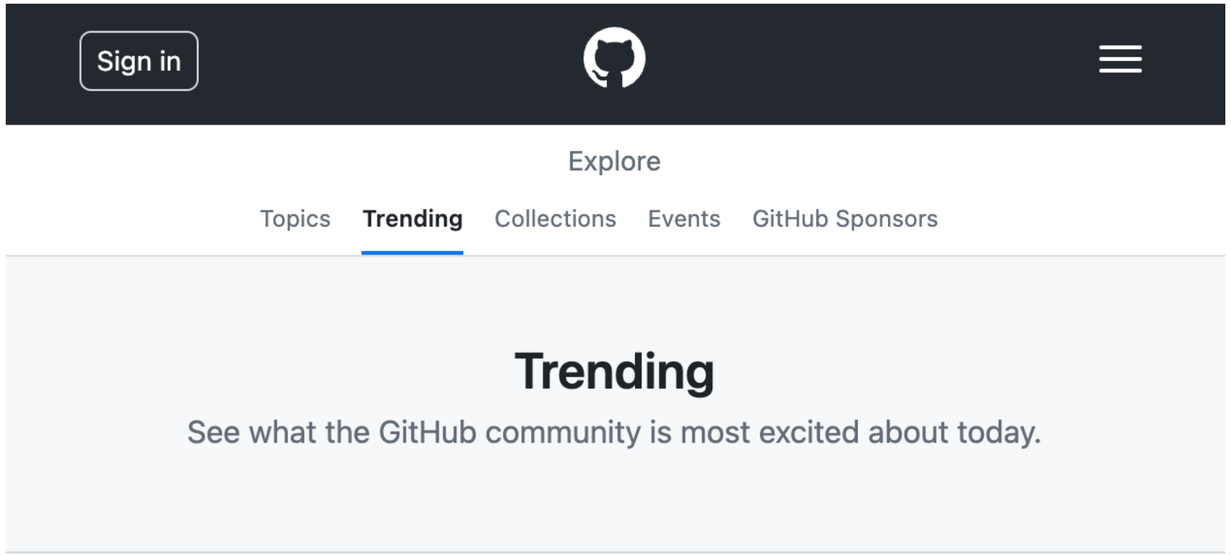
Peer production

Wikipedia

The Free Encyclopedia




moz://a
Firefox for Desktop
Firefox Browser
Get the browser that protects what's important
No shady privacy policies or back doors for advertisers. Just a lightning fast browser that doesn't sell you out.
Download Firefox


Sign in
Explore
Topics
Trending
Collections
Events
GitHub Sponsors
Trending
See what the GitHub community is most excited about today.

What is peer production?

Crowdsourcing: making an open call to a large set of individuals who self-select into tasks

Peer production introduces additional requirements, including… [Benkler 2002]

- Decentralized conception: many control the direction and outcome, not a traditional bureaucracy
- Results treated as a commons: the output is publicly available and generally non-rival (def: when I use it, it doesn't reduce your ability to use it)
- No contracts: governance and work allocation isn't handled through signed contracts

When does peer production work?

Benkler's argument [2002]: peer production outperforms traditional firms (e.g., for-profit companies) when (1) there exists strong intrinsic motivation and (2) work can be broken down into granular, easy-to-integrate tasks.

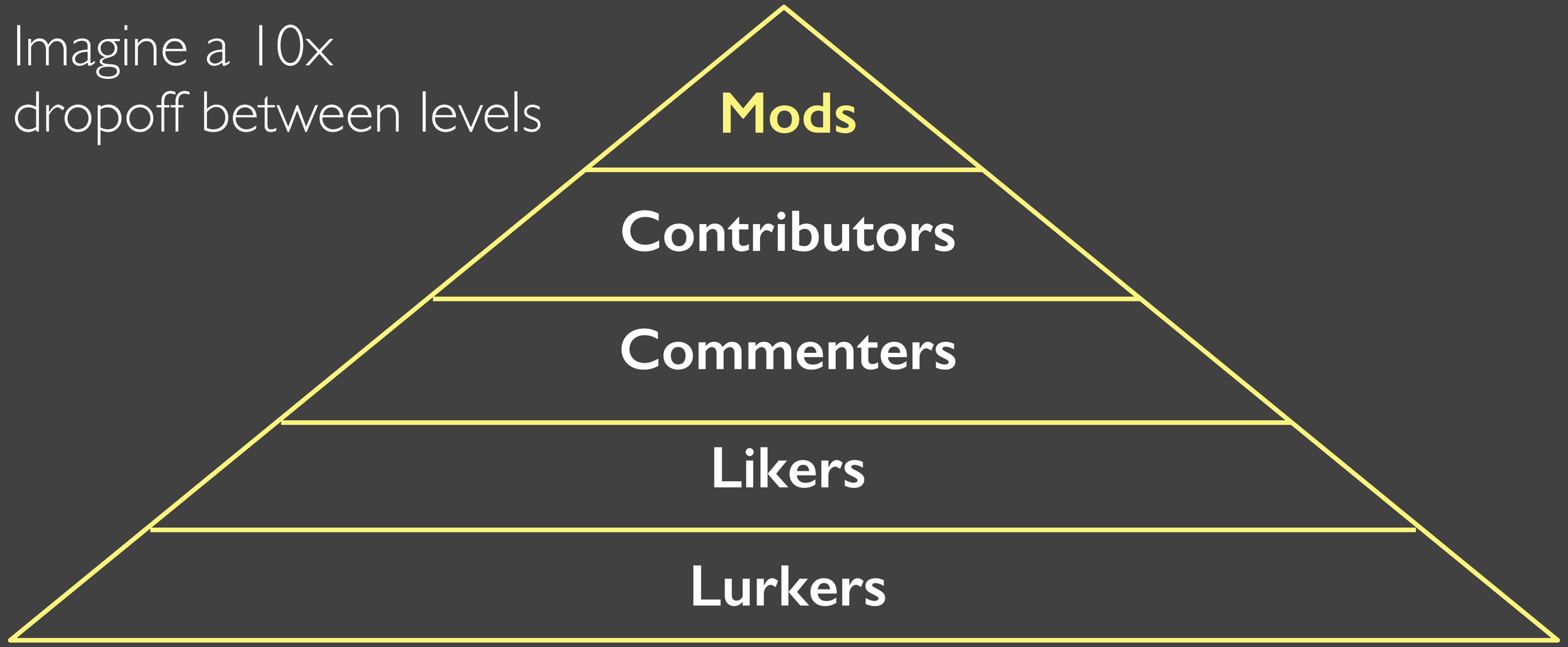
Linux server > Microsoft Windows server

Wikipedia > Encyclopedia Brittanica

Git > Perforce

From before

Imagine a 10x
dropoff between levels

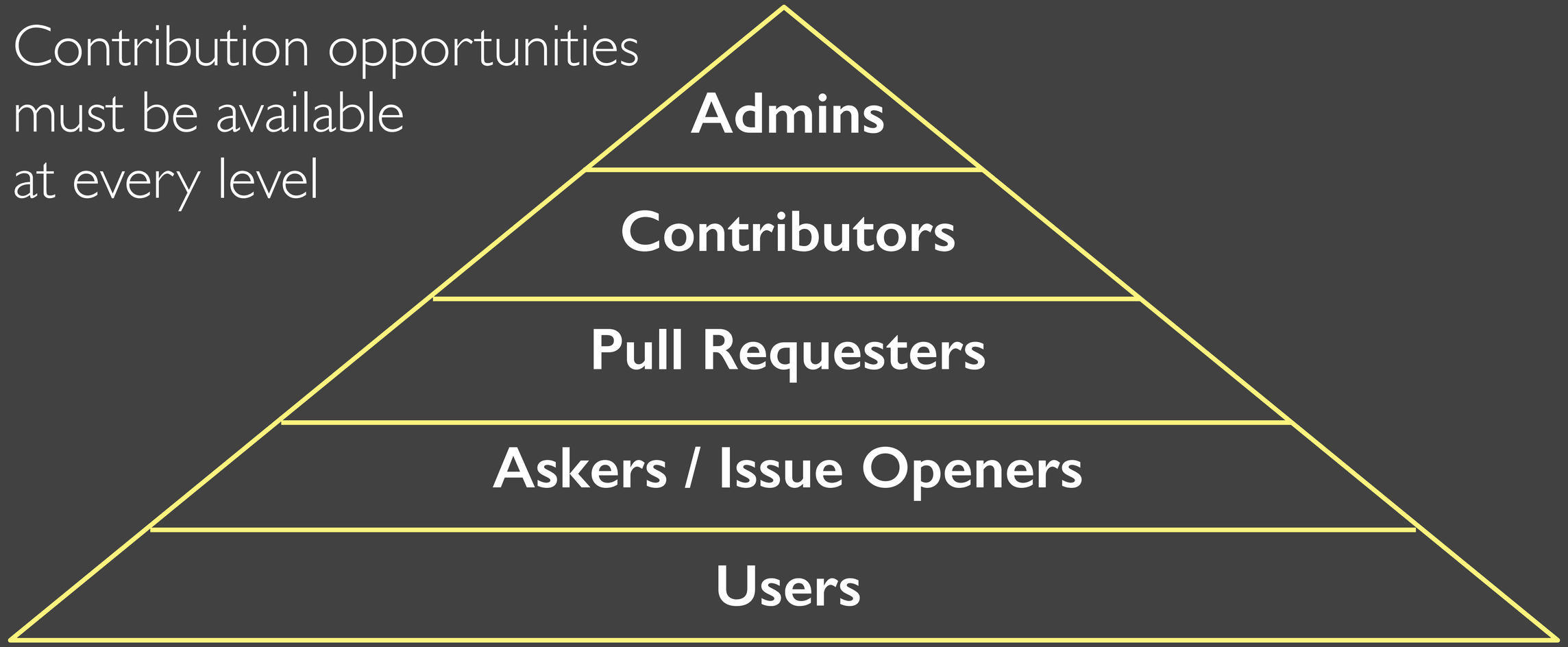


Peer production: wiki

Contribution opportunities
must be available
at every level

Peer production: OSS

Contribution opportunities
must be available
at every level



Why do people do this?

The usefulness of the outcome to the contributor; hedonic pleasure of contributing (e.g., writing software); increased social capital, reputation, and status [von Hippel and von Krogh 2003, von Krogh et al. 2003, Benkler, Shaw and Hill 2015]

People self-select into communities that match their motivations [Glott et al. 2010, Ghosh and Prakash 2000; Belenzon and Schankerman 2008, Benkler, Shaw and Hill 2015]

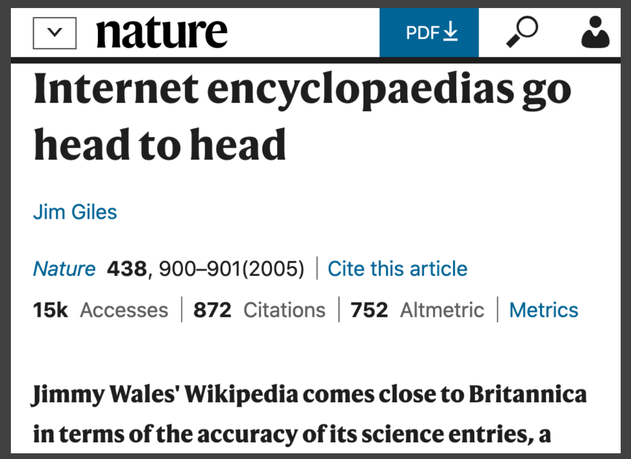
- Those extrinsically motivated by reputation and employment will contribute more to industry-sponsored projects
- Those more intrinsically motivated will contribute more to free communities

But does it really work?

Pros

Linus's Law: "With enough eyes, all bugs are shallow" [Raymond 1999]

Wikipedia used to be disallowed as a citable source because it could not be trusted. But then:


nature

PDF

Internet encyclopaedias go head to head

Jim Giles

Nature **438**, 900–901(2005) | Cite this article

15k Accesses | **872** Citations | **752** Altmetric | Metrics

Jimmy Wales' Wikipedia comes close to Britannica in terms of the accuracy of its science entries, a

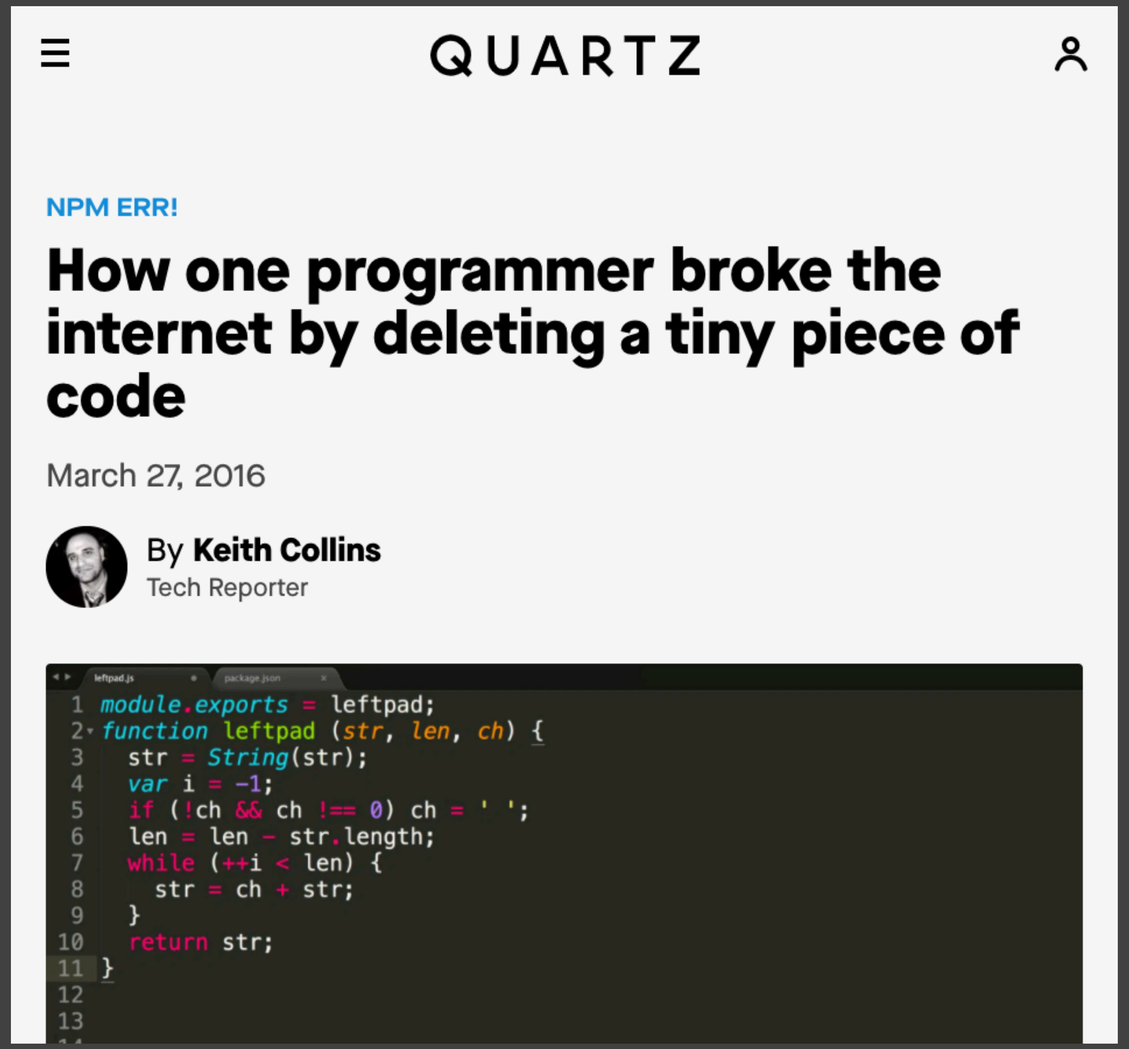
Cons

Many efforts do not achieve critical mass needed for quality [Ghost Town lecture]

Peer production appears better at creating functional artifacts (e.g., code) than creative artifacts (e.g., movies) [Benkler 2006]

No direct control of peoples' efforts: 1.5B monthly Wikipedia views go to articles that would be higher quality if editors optimally distributed their work to meet reader demand. [Warncke-Wang et al. 2015]

And errors do occur…



QUARTZ

NPM ERR!

How one programmer broke the internet by deleting a tiny piece of code

March 27, 2016

By **Keith Collins**
Tech Reporter

leftpad.js package.json

```
module.exports = leftpad;
function leftpad (str, len, ch) {
  str = String(str);
  var i = -1;
  if (!ch && ch !== 0) ch = ' ';
  len = len - str.length;
  while (++i < len) {
    str = ch + str;
  }
  return str;
}
```

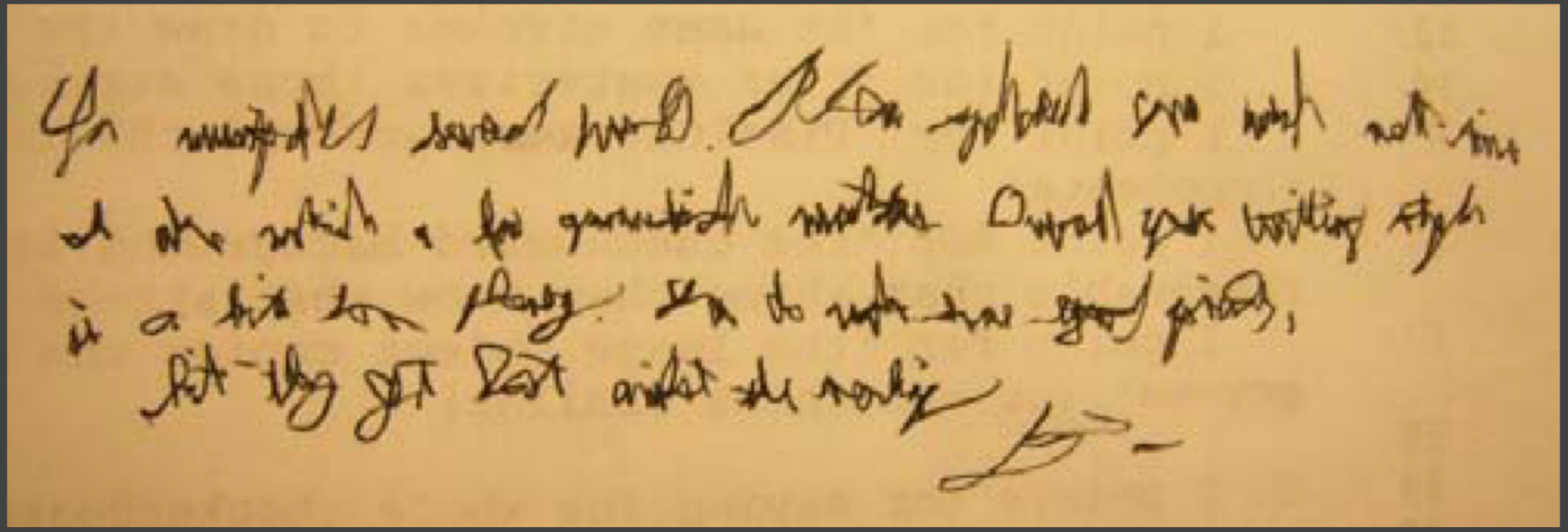
node.js `leftpad` module incident

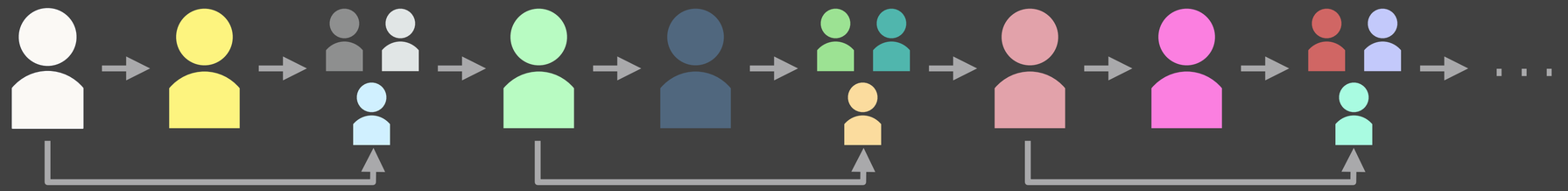
So given these tradeoffs, when would you opt for peer production over firm-based production, assuming you had moderate but not infinite funds? [2min]

Workflows

Iterative crowd algorithm

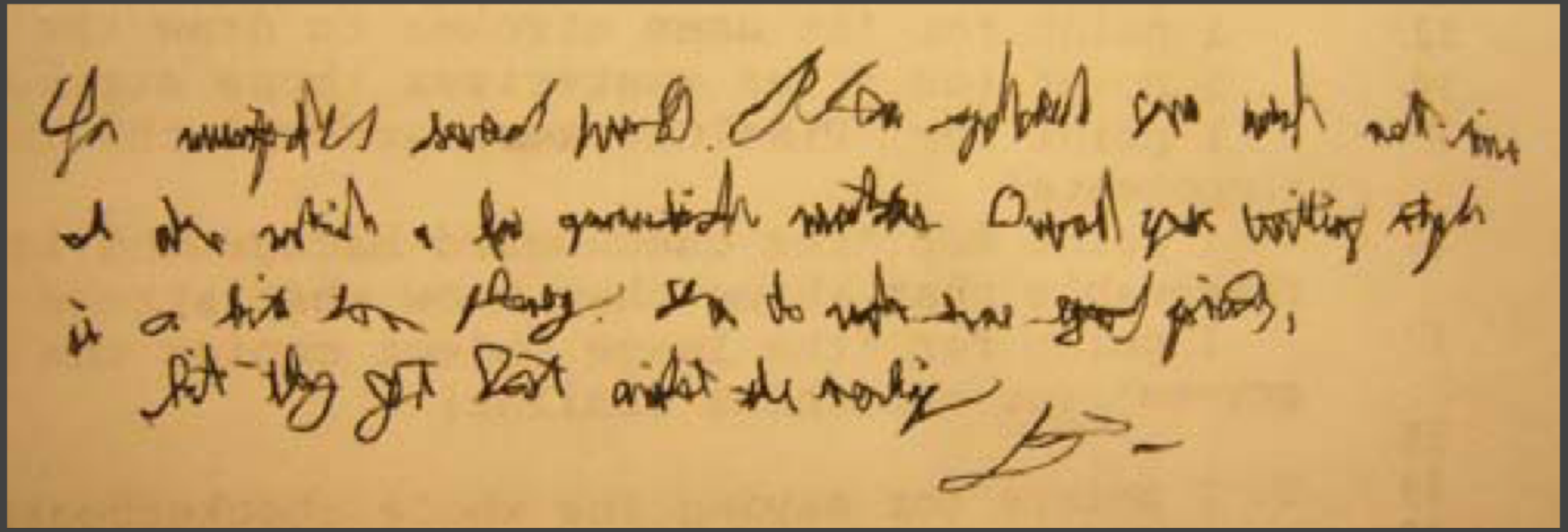
[Little et al. 2009]



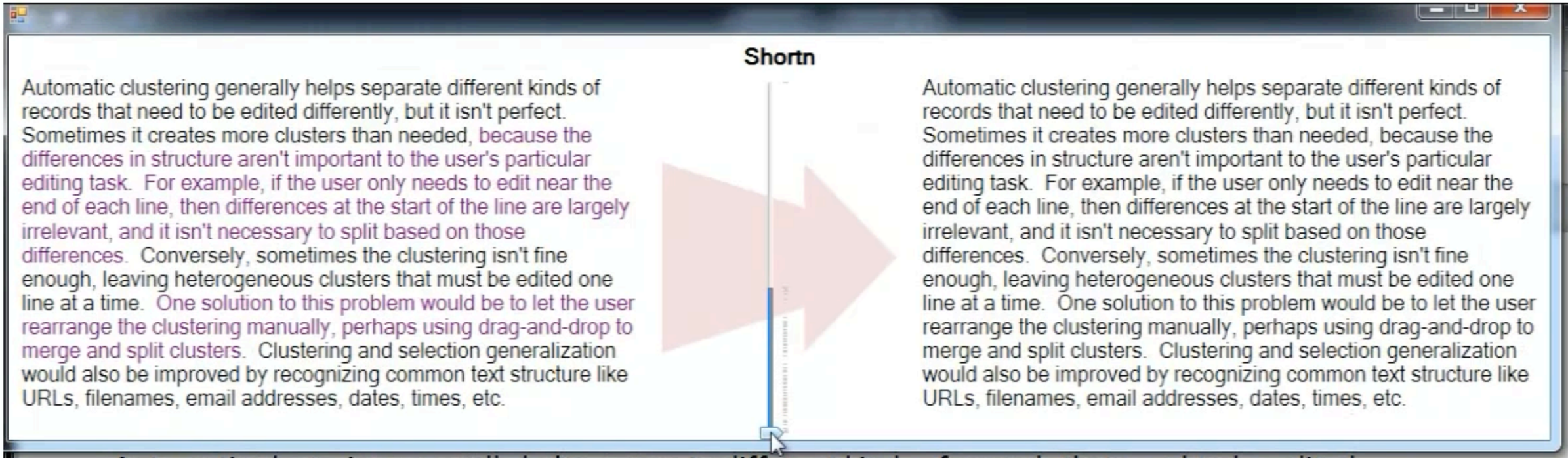


Iterative crowd algorithm

[Little et al. 2009]



You (misspelled) (several) (words). Please spellcheck your work next time. I also notice a few grammatical mistakes. Overall your writing style is a bit too phoney. You do make some good (points), but they got lost amidst the (writing). (signature)



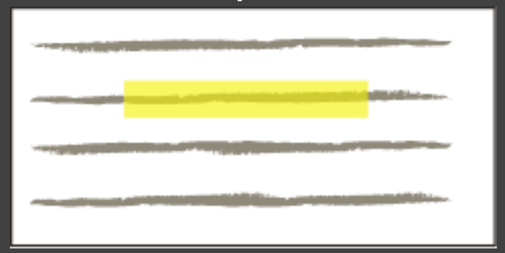
Automatic clustering generally helps separate different kinds of records that need to be edited differently, but it isn't perfect. Sometimes it creates more clusters than needed, because the differences in structure aren't important to the user's particular editing task. For example, if the user only needs to edit near the end of each line, then differences at the start of the line are largely irrelevant, and it isn't necessary to split based on those differences. Conversely, sometimes the clustering isn't fine enough, leaving heterogeneous clusters that must be edited one line at a time. One solution to this problem would be to let the user rearrange the clustering manually, perhaps using drag-and-drop to merge and split clusters. Clustering and selection generalization would also be improved by recognizing common text structure like URLs, filenames, email addresses, dates, times, etc.

Find-Fix-Verify

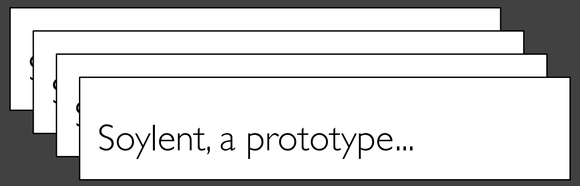
[Bernstein et al. 2010]

Find-Fix-Verify is a design pattern for open-ended tasks.

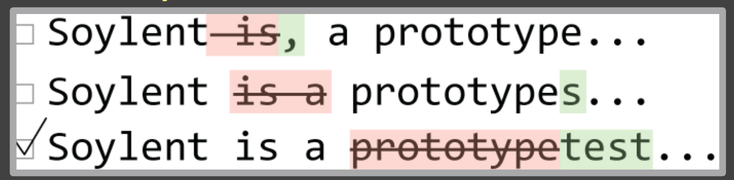
Find a problem



Fix the problem

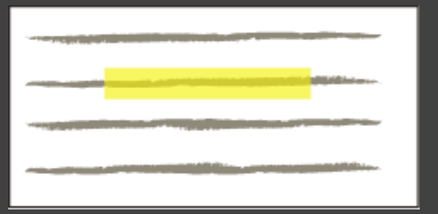


Verify each fix



Find

"Identify at least one area that can be shortened without changing the meaning of the paragraph."



Independent agreement to identify patches

Fix

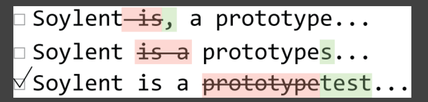
"Edit the highlighted section to shorten its length without changing the meaning of the paragraph."

Soylent, a prototype...

Randomize order of suggestions

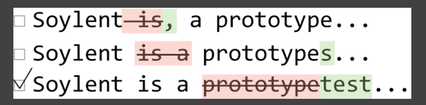
Verify

"Choose at least one rewrite that has style errors, and at least one rewrite that changes the meaning of the sentence."

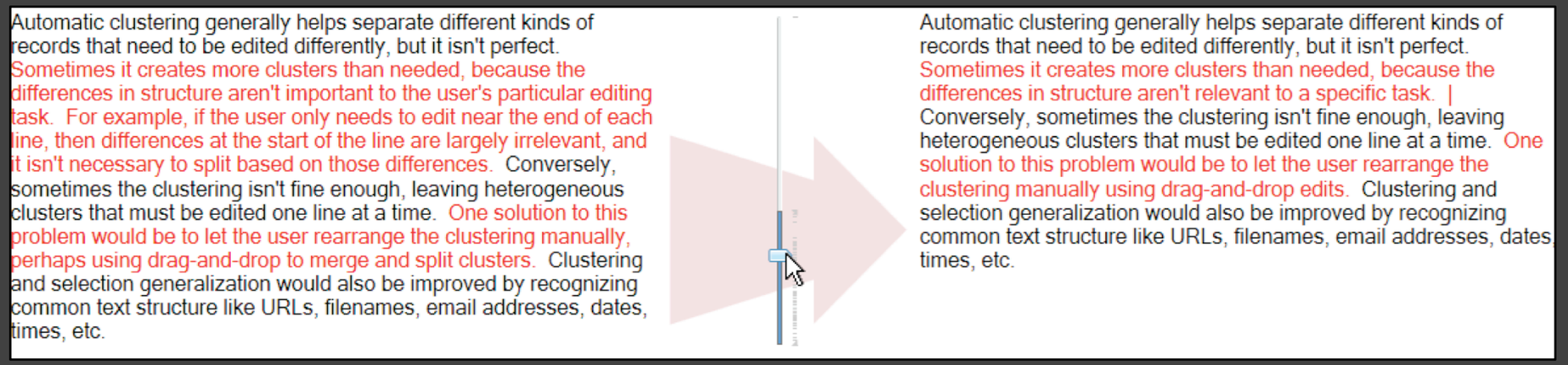


Verify

"Choose at least one rewrite that has style errors, and at least one rewrite that changes the meaning of the sentence."



Keep suggestions that do not get voted out



Mechanical Novel

[Kim et al. 2017]

Could we crowdsource the Great American Novel?

Unlike most crowdsourcing workflows, creative work requires tight interconnections between different parts of a story, and between the high-level goal and low-level text

Reflect

choose a high-level goal

Revise

break into tasks and edit

The Hot Air Balloon

A young boy named Malcolm finds himself alone in a runaway hot air balloon and accidentally travels to a city in the sky.

Malcolm found himself fascinated by the balloon. He imagined what it would be like to fly to some exotic location, soaring above the clouds. He closed his eyes and saw himself flying through the clouds. "Hey you!" he heard a voice shout. "What are you doing in my balloon!" The operator yelled as he dropped his snack and ran toward the balloon. Malcolm, startled by the man stumbled. Trying not to fall he reached out and grabbed a rope hanging next to him. As he pulled it, it released the ballast and the balloon leapt into the air. Before he knew it he was rising up

#1

+

relief as he felt the balloon descend a bit. Then suddenly a great gust of wind took the balloon and sent it higher and higher. Fell to the floor, grabbing a lever on the way down that made the balloon rise even faster.

Answer the following questions about the story:

What did you like about the story above?

I liked the sen

What do you wish you could change about the story above?

I wish...

Is the wish you wrote above mostly about the plot, the characters, or the writing? (select one)

Choose...

In one sentence, suggest something for the next revision of the story that could change to address your comments above.

Start your sentence with the words "What if...". Try to write something specific that fits the story.

What if...

Submit

e story above! Which change should be the main change for the next revision?

ut the lo you	**Think about what you wish you could change about the story. Which statement below do you agree with most?**	**Which new idea below do you think should be the main change for a revision of the story above?**
detail and imagine oughout.	I wish there would be a bit more exposition at the beginning of the story. The events all start very quickly and it's easy to not understand what's going on.	What if instead of beginning so abruptly, we learned a little more about Malcolm first?
ot air ballon	The order of the paragraphs. It jumped all over the place.	What if... We wrote a new story about hot air balloons and a mistaken escape?
he previous used and it's iting than	I want to see more exciting things happen to Malcom in Cloud City, and I want to know about his return to the ground.	What if Malcom investigates the Cloud City and then tries to get home, maybe bringing along a friend from Cloud City.

Make the following change to the text on the left:

First, there is a typo ("she he") should be "should he". Finish paragraph with: Perhaps he should take a little gander and see if he can get some help with the balloon.

When he woke up, he found himself staring at glow of big, beautiful crystal-like glass buildings, shimmering in the brilliant sunshine. One was emerald green, another ruby red, still another sparkling like diamonds, others radiating spectacular colors he had never seen. He was a little scared, to be in an unfamiliar place, but he was mostly curious. He could sense that something interesting

He suddenly found himself staring at gargantuan glass buildings, their exteriors shimmering under the sun's rays. These buildings were like no others he'd seen before - one was emerald colored, another ruby red, and a third sparkling like a multifaceted diamond. This scenery was frightening, yet, he was curious. Something was going on in this city, and it could be something wonderful. Maybe there were helpful people down below? Perhaps he should take a little gander and see if he can get some help with the balloon.

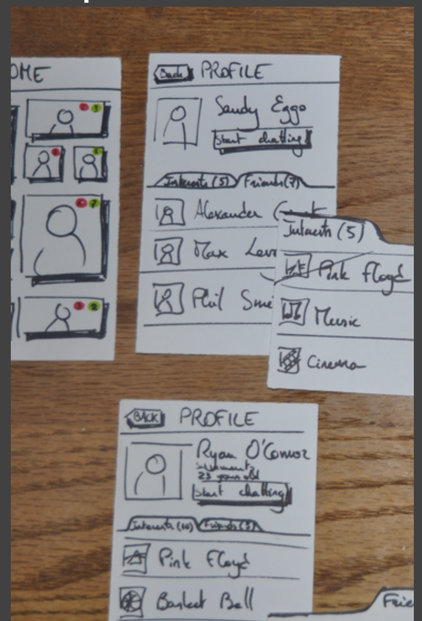
#1
#2
#3
#4
#5
#6
#7
+

Flash Teams

[Retelny et al., UIST '14]

Computationally-guided teams of crowd experts supported by lightweight team structures.

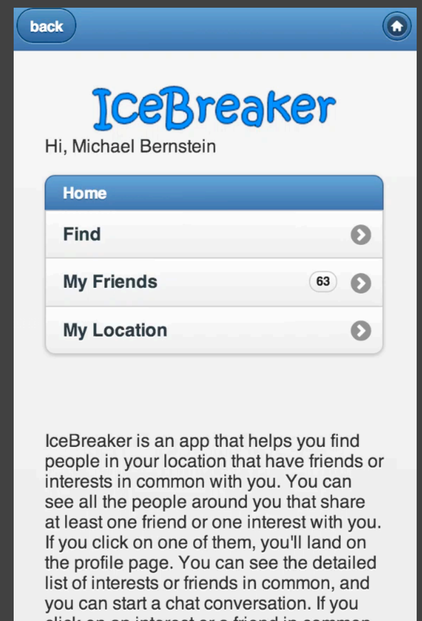
Input



Flash Team

Design workflow

Output



What crowdsourcing can't do

Limits of algorithmic coordination

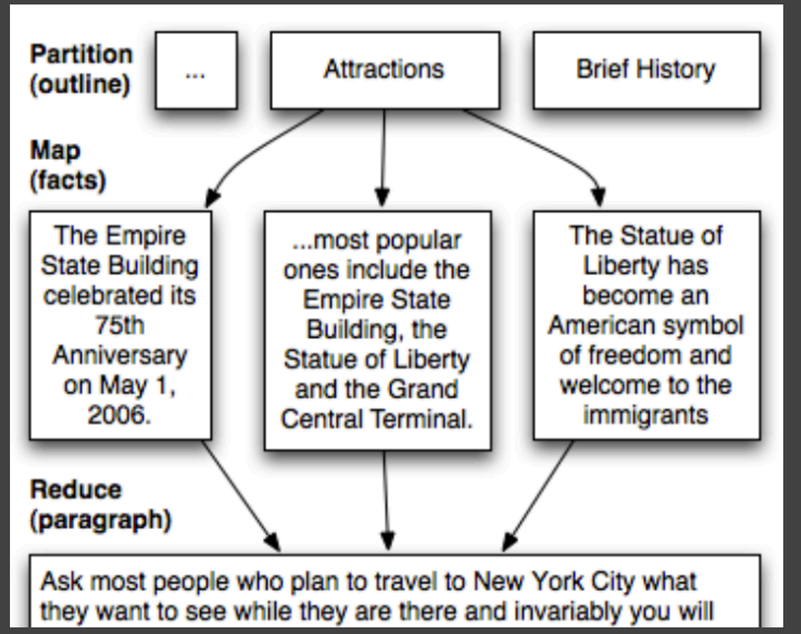
So far, goals such as invention, production, and engineering have remained largely out of reach [Kittur et al. 2013]

Why? [1 min]

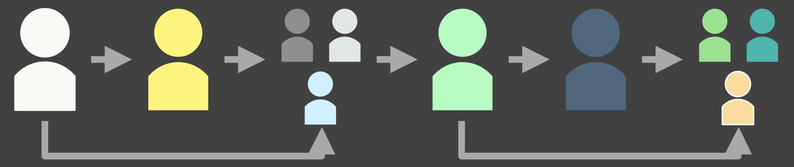
Dominant architecture: algorithms

Modularize and pre-define all possible behaviors into workflows

Computation decides which behaviors are taken, when, and by whom; optimizes, error-checks, and combines submissions



[Kittur 2011]



[Little 2009; Little 2010]
[Dai and Weld 2010]

Limits of algorithmic coordination

Returning to the question: why have complex goals remained largely out of reach?

Open-ended, complex goals are fundamentally incompatible with a requirement to modularize and pre-define every behavior [Van de Ven, Delbecq, and Koenig 1976; Rittel and Webber 1973; Schön 1984]

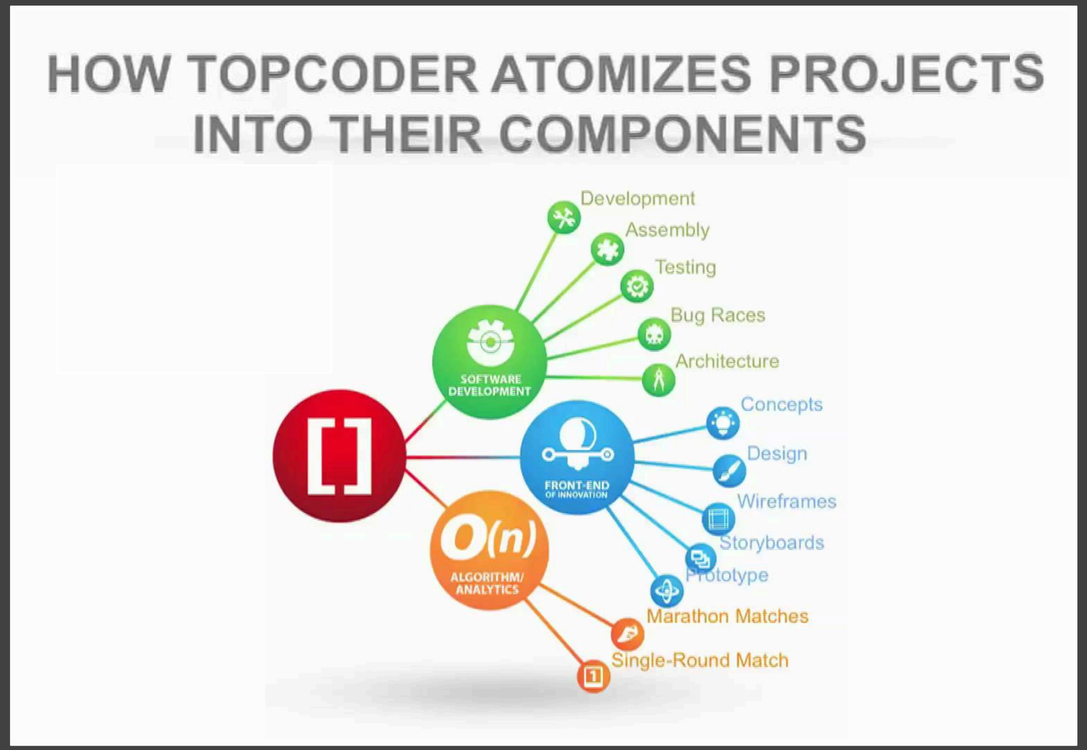
Of course, workflows cause issues… [Retelny et al. 2017]

Workflows are inflexible and can become games of telephone, where intent gets lost and errors can propagate over time

Limits of crowdsourcing and peer production

"Peer production is limited not by the total cost or complexity of a project, but by its modularity." [Benkler 2002]

"With the Linux kernel […] we want to have a system which is as modular as possible. The open–source development model really requires this, because otherwise you can't easily have people working in parallel." [Torvalds 1999]



[Boudreau, Lacetera, and Lakhani 2011]

Interdependence and collective action remain challenging

The result: algorithmic, workflow-based architecture confines collaborations to goals so predictable that they can be entirely modularized and pre-defined.

But many valuable collective activities do not fit this criteria.

Why are these challenging?

Convergence: crowds are excellent at generating ideas and at spreading awareness, but it's much more challenging for them to build consensus toward a single action.

Convergence

[Example via Niloufar Salehi]

The New York Times

Twitter Users Split on Boycott Over Platform's Move Against Rose McGowan

VOGUE

CULTURE > OPINION

Why I'm Not Joining the Women Boycotting Twitter Today

The Washington Post

#WomenBoycottTwitter revives post-election conversations about the lack of solidarity among women

MADAMENOIRE

Black Women Commandeer #WomenBoycottTwitter And Turn It Into A Celebration Of Ourselves With #WOCAffirmation

Convergence

[Example via Niloufar Salehi]

Why are these challenging?

Coordinated adaptation: changing direction in sync with each other.

Crowds are excellent at executing pre-defined tasks, but it's much more challenging for them to continually re-evaluate goals and adapt in sync.

colnick101: !!yes after all these years

neikrodent: hard right

creepplosion: 420

richie05d: SPAM 420!

Bunneo: drive

Barn16: LET'S DO THIS

Xcizer7911: reverse

Tomnyan6803: 3

jokogaming1: Drive

zepinktacos: MUM IM A PILOT!!

panseysandpaws: drive

xxHulkermanxx: drive

Khaely_: Dawn

richie05d: 420

ajoconnor64: hi YouTube?

Q2Vertigo: drive

0.00

HYPERAVA FOLLOW

SNIPPYLLAMA 100 BITS

ROLFER FOLLOW

Can Twitch Chat land an Airplane without dying?

Rules: fly for at least 30 seconds, then land the plane without dying. !GtaCommands for controls.

Hybrid peer production

Why is it that many successful peer production projects form either nonprofits or for-profit organizations to support their efforts?

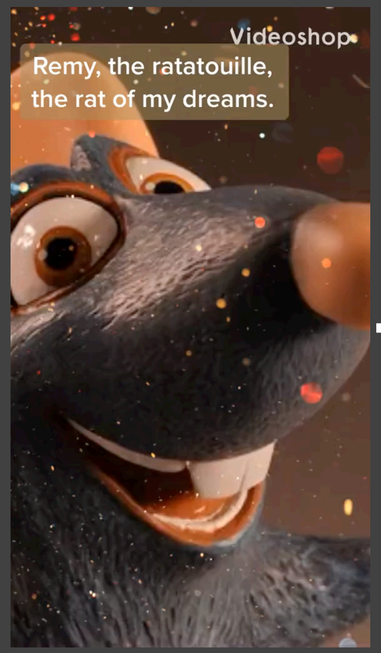
- Nonprofit: PyTorch (PyTorch Found.), Kubernetes (CNCF), Django (DSF)
- For profit: Docker (Docker, Inc.), Ubuntu (Canonical), MySQL (Oracle)

Peer production struggles with tasks that traditional contract-based firms achieve (e.g., keeping release schedules, integrated contributions). So, they form hybrid models to compensate.

- Example: supposedly simple goal for Ubuntu—plugging a USB drive and the drive appearing on your desktop

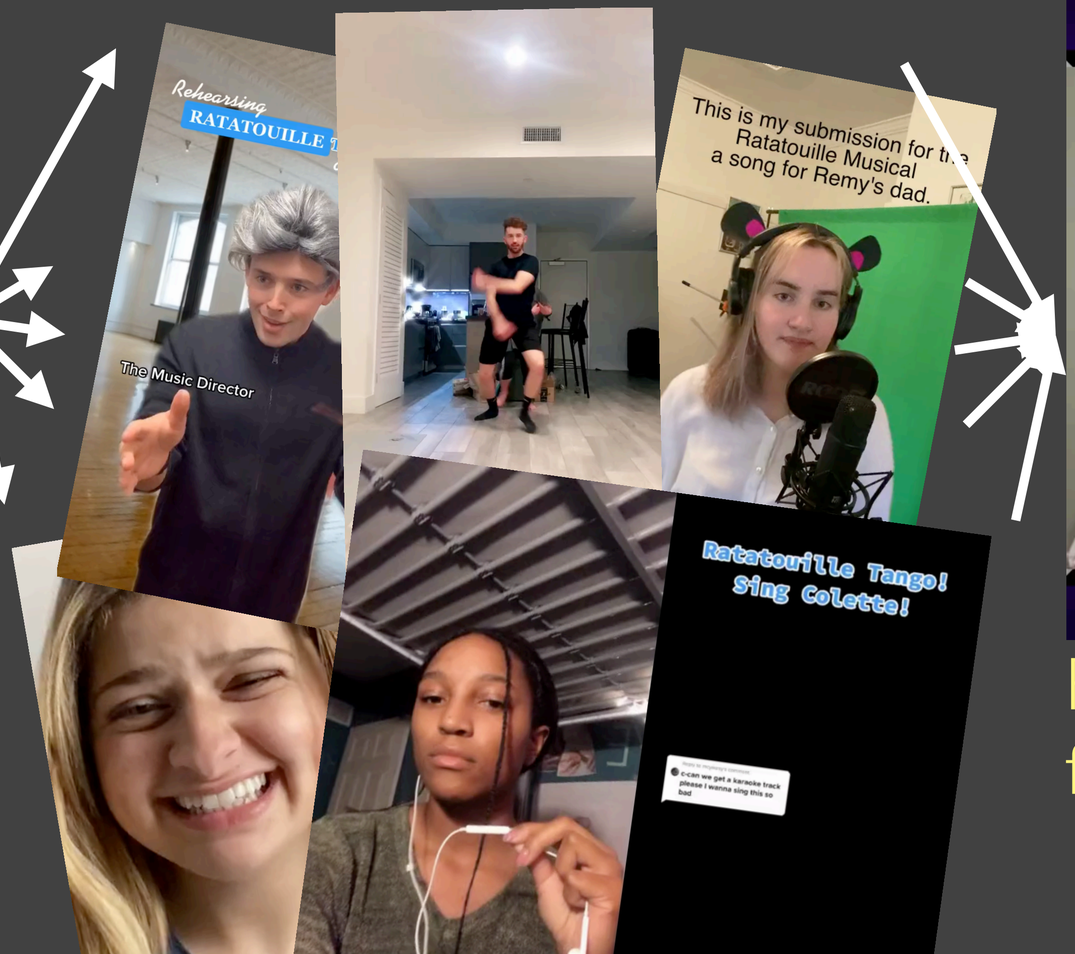
Ratatouille TikTok musical

Seed idea



Structure

Divergence



Convergence

Required a focused team

Has your opinion changed?

When would you opt for peer production over firm-based production, assuming you had moderate but not infinite funds?

Which would you use if the goal were to:

- Write a lecture for CS 278?

- Redesign the requirements for your major?

- Decide on the future of Stanford neighborhoods?

[2min]

Exam

Closed-book exam on paper

Questions sampled from the question bank of top questions from voting. Question bank posted in advance.

- Roughly 1/4 Easy questions, 1/4 Medium questions, 1/4 Hard questions…
- And 1/4 staff-written questions

Study groups OK, but no collaboration on or sharing notes or answers

Details on the website

Summary

Shifting from simple wisdom-of-the-crowd tasks requires much more than just a scaling up of ambition: it requires designing for interdependence.

Peer production — the term encompassing shared open work (e.g., Wikipedia, open source) is one powerful method for volunteer coordination. Workflows and algorithms offer another approach. Both have their issues.

Aiming higher means we will need to solve issues of convergence and coordinated adapatation.

References

Benkler, Yochai. "Coase's penguin, or, linux and the nature of the firm"." Yale law journal (2002): 369-446.

Benkler, Y. (2006). The Wealth of Networks: How Social Production Transforms Markets and Freedom. New Haven and London: Yale University Press.

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

Belenzon, Sharon, and Mark A. Schankerman. "Motivation and sorting in open source software innovation." (2008).

Bernstein, Michael S., et al. "Soylent: a word processor with a crowd inside." Proceedings of the 23nd annual ACM symposium on User interface software and technology. 2010.

Boudreau, Kevin J., Nicola Lacetera, and Karim R. Lakhani. "Incentives and problem uncertainty in innovation contests: An empirical analysis." Management science 57.5 (2011): 843-863.

Dai, Peng, and Daniel Weld. "Decision-theoretic control of crowd-sourced workflows." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 24. No. 1. 2010.

Ghosh, Rishab Aiyer, and Vipul Ved Prakash. "The orbiten free software survey." First Monday (2000).

Glott, Ruediger, Philipp Schmidt, and Rishab Ghosh. "Wikipedia survey–overview of results." United Nations University: Collaborative Creativity Group 8 (2010): 1158-1178.

References

Hippel, Eric von, and Georg von Krogh. "Open source software and the "private-collective" innovation model: Issues for organization science." Organization science 14.2 (2003): 209-223.

Kim, Joy, et al. "Mechanical novel: Crowdsourcing complex work through reflection and revision." Proceedings of the 2017 acm conference on computer supported cooperative work and social computing. 2017.

Kittur, Aniket, et al. "The future of crowd work." Proceedings of the 2013 conference on Computer supported cooperative work. 2013.

Kittur, Aniket, et al. "Crowdforge: Crowdsourcing complex work." Proceedings of the 24th annual ACM symposium on User interface software and technology. 2011.

Little, Greg, et al. "Turkit: tools for iterative tasks on mechanical turk." Proceedings of the ACM SIGKDD workshop on human computation. 2009.

Little, Greg, et al. "Turkit: human computation algorithms on mechanical turk." Proceedings of the 23nd annual ACM symposium on User interface software and technology. 2010.

Retelny, Daniela, et al. "Expert crowdsourcing with flash teams." Proceedings of the 27th annual ACM symposium on User interface software and technology. 2014.

Retelny, Daniela, Michael S. Bernstein, and Melissa A. Valentine. "No workflow can ever be enough: How crowdsourcing workflows constrain complex work." Proceedings of the ACM on Human-Computer Interaction 1.CSCW (2017): 1-23.

References

Raymond, Eric. "The cathedral and the bazaar." Knowledge, Technology & Policy 12.3 (1999): 23-49.

Rittel, Horst WJ, and Melvin M. Webber. "Dilemmas in a general theory of planning." Policy sciences 4.2 (1973): 155-169.

Schön, Donald A. The reflective practitioner: How professionals think in action. Routledge, 2017.

Torvalds, L. 1999. "The Linux Edge". In DiBona, Chris, and Sam Ockman, Eds. "Open sources: Voices from the open source revolution. " O'Reilly Media, Inc.", 1999.

Van de Ven, Andrew H., Andre L. Delbecq, and Richard Koenig Jr. "Determinants of coordination modes within organizations." American sociological review (1976): 322-338.

Von Krogh, Georg. "Open-source software development." MIT Sloan Management Review (2003).

Warncke-Wang, Morten, et al. "Misalignment between supply and demand of quality content in peer production communities." Proceedings of the International AAAI Conference on Web and Social Media. Vol. 9. No. 1. 2015.

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.