Community Data Science Collective

Prospective graduate student Q&A Session

October 2, 2026

Welcome & Agenda

- Brief introduction & orientation (15 minutes)
- Q&A (20 minutes)
- Breakouts with current student group members (20+ minutes)

What is the Community Data Science Collective?

- Interdisciplinary research group working at multiple campuses

- We:
  - Study (mostly) online communities, participation, and learning...
  - Use (mostly) quantitative and computational methods...
  - Do social science, social computing (HCI/SWE/AI/Cyber), and learning science research.

Who are CDSC faculty? What Ph.D. programs?

Kaylea Champion (University of Washington - Bothell)

Computing and Software Systems

Jeremy Foote (Purdue University)

Brian Lamb School of Communication

Ryan Funkhouser (University of Idaho)

Department of Psychology and Communication

Benjamin Mako Hill (University of Washington)

Communication; Human-Centered Design & Engineering; Computer Science & Engineering; Information School

Aaron Shaw (Northwestern University)

Media, Technology & Society; Technology & Social Behavior; Sociology; Center for Human Computer Interaction + Design

Nathan TeBlunthuis (University of Texas at Austin)

School of Information

What kind of research happens in the CDSC?

What kinds of research happen in the CDSC?

A few illustrative example projects:

(1) Rethinking openness in online communities

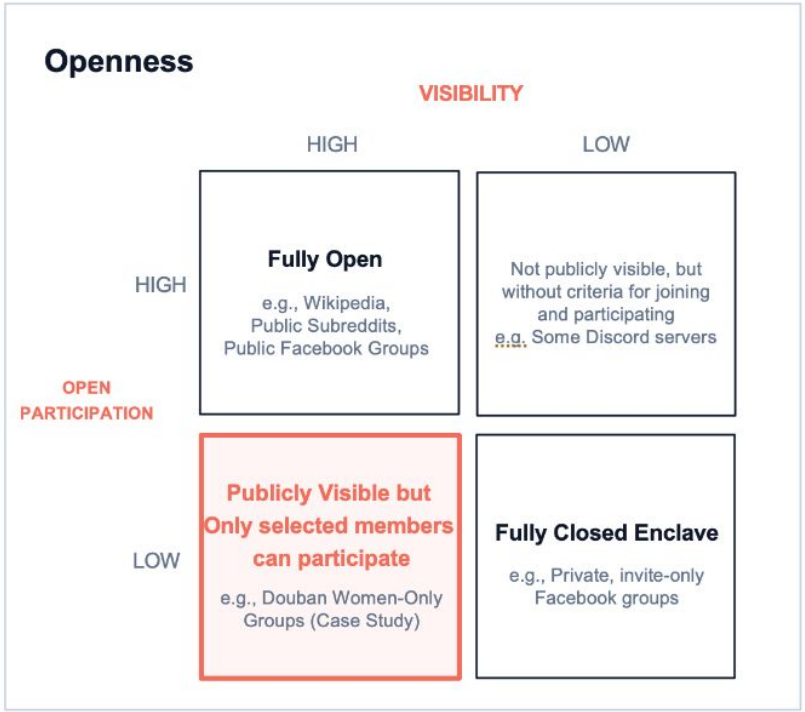
(2) Modeling the introduction of documentation in FLOSS projects

(3) Understanding the StackExchange moderator strike

For each one, we'll tell you a (very) little bit more...

Project #1: Rethinking Openness in Online Communities

Ran Tang, Benjamin Mako Hill, Zilu Tang, and Qinyuan Lei — University of Washington, TU Eindhoven, City University of Hong Kong



Problem: Openness is undertheorized!

Methods: Semi-structured interviews with online community leaders.

We argue that
1. Openness consists of 2 dimensions:
Visibility (who can see group activities);
Open Participation (who can engage and contribute).

2. Openness is a dynamic process rather than a fixed state.

3. The decoupling of visibility and participation leads communities to navigate the tension of the "dual character" of counterpublics (Fraser, 1990) – serving as both inward-facing enclaves for safety and outward-facing bases for public actions – not temporally, but structurally.

Project #2: The Introduction of README and CONTRIBUTING Files in Open Source Software Development

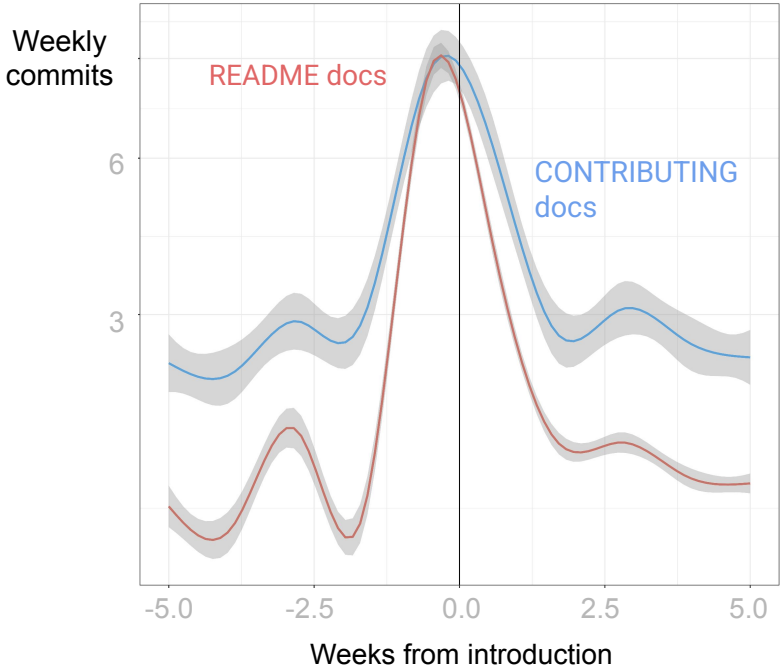
Project team: Matt (Northwestern), Kaylea (UW), Sohyeon (Northwestern; Princeton), Aaron (Northwestern)

Problem: Projects receive a lot of recommendations about documentation as a way to recruit contributors to FLOSS projects. Does it work? (Answer: Not really)

Data & Methods: 4200+ FLOSS projects packaged in Debian; documentation and commit history. (*Incl. data Kaylea collected for earlier work!*). Quantitative modeling.

A finding: Projects introduce docs right **after** spikes of commits (not before).

Published!
https://ieeexplore.ieee.org/abstract/document/11024374



Project #3: Understanding the 2023 Stack Exchange Strike

Project Team: Yiwei Wu(UT Austin), Leah Ajmani(U of Minnesota), Nathan TeBlunthuis(UT Austin), Hanlin Li(UT Austin)

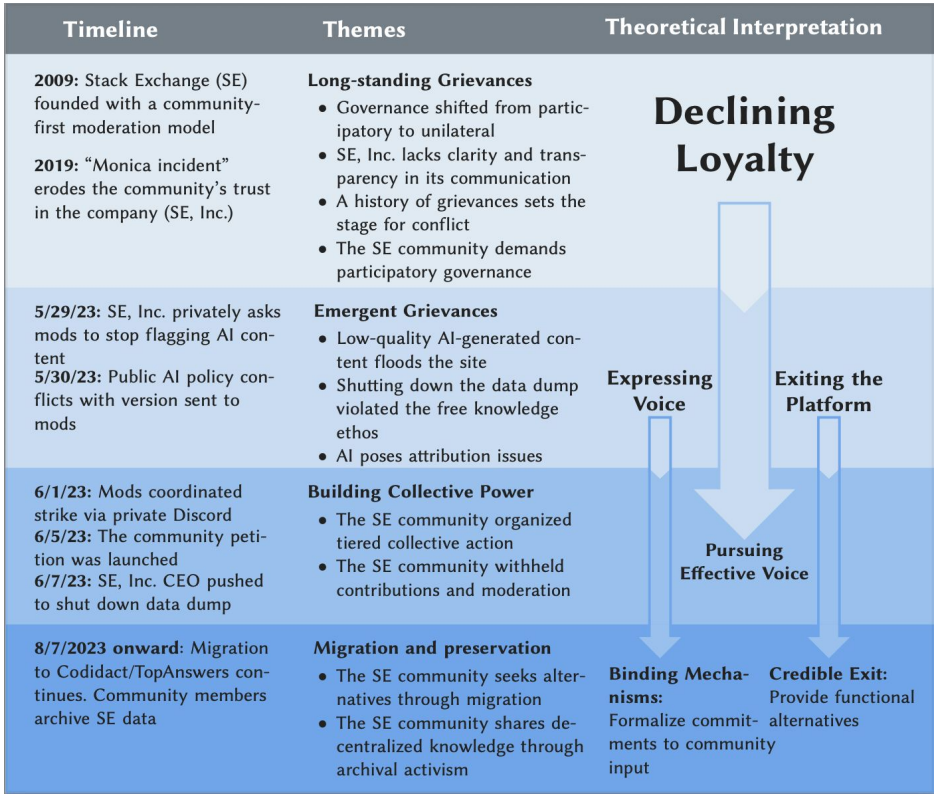
Problem: When Stack Exchange moderators and contributors went on strike in 2023, it looked like a fight over AI. Was it? (Answer: Not really.)

Method: A **qualitative study** by collecting data from online discussion threads and conducting interviews with 14 community members followed by thematic analysis

We adopted Albert O. Hirschman's **exit, voice, and loyalty** as the main theoretical framework.

Participants described **a long-term** deterioration of their relationship with SE platform, which set the stage for the strike.

Published! (CSCW'26)
https://dl.acm.org/doi/pdf/10.1145/3788040



Where can you learn and find more?

This was a wildly incomplete set. There are many more people and projects!

- CDSC wiki (especially the people page and the publications page)
- CDSC research blog
- CDSC Zenodo

Q&A

Start with some questions submitted ahead of time.

Use the Q+A system to write / vote on questions

Also, current student members of the group wrote a “not-so-obvious FAQ” and we responded to some of this year’s questions at https://wiki.communitydata.science/Prospective_PhD_Student_Q%26A_(October_2026)

Breakouts!

Find people you'd like to talk to, and join the conversation!

Breakout rooms are set up with current students from each campus.

In ~12 minutes, we will alert those who would like to switch rooms

Backup/Annex Slides….