The role of trust in sustainability of knowledge-sharing social groups

The case of Stack Exchange Q&A communities

Ana Vranić [1] Aleksandar Tomašević [2] Aleksandra Alorić [1,3] Marija Mitrović Dankulov [1]

[1]Institute of Physics Belgrade, University of Belgrade, Serbia

[2]Faculty of Philosophy, University of Novi Sad, Serbia

[3]Two Desperados, Serbia

Motivation

- Googled a coding question ⇒ Stack Overflow

Motivation

- Googled a coding question ⇒ Stack Overflow
- How **trust** emerges within a community?

Motivation

- Googled a coding question ⇒ Stack Overflow
- How **trust** emerges within a community?
- What is the role of trust in **sustainability** of these communities?

Stack Exchange

- More than just StackOverflow
- The Stack Exchange network has 170+ Q&A communities
- New communities are showing up!
  - Define
  - Commit
  - Beta

191
days in beta

Visit!

Proof Assistants

Beta Q&A site for mathematicians and computer scientists who develop and use proof assistants.

0.6 questions per day

Needs Work – 10 questions per day on average is a healthy beta, 5 questions or fewer per day needs some work. A healthy site generates lots of good content to make sure users keep coming back.

89% answered

Okay – 90% answered is a healthy beta, 80% answered needs some work. In the beta it's especially important that when new visitors ask questions they usually get a good answer.

132 avid users

1,605 total users

Okay – Every site needs a solid group of core users to assist in moderating the site. We recommend:

- 150 users with 200+ rep (on pace for **147 users** at 215 days)
- 10 users with 2,000+ rep (on pace for **11 users** at 215 days)
- 5 users with 3,000+ rep (on pace for **3 users** at 215 days)

1.8 answer ratio

Okay – 2.5 answers per question is good, only 1 answer per question needs some work. On a healthy site, questions receive multiple answers and the best answer is voted to the top.

66 visits/day

Needs Work – 1,500 visits per day is good, 500 visits per day needs some work. A great site benefits people outside the community. Eventually, 90% of a site's traffic should come from search engines.

This site is in **Public Beta**

Anyone can participate!

Sites remain in beta for at least 180 days to build up a critical mass of users, questions, and participation.

proposed by

Andrej Bauer
113k ●14 ●306 ●584

9 months ago

edited by

Henry WH Hack v3.0b
8,942 ●13 ●65 ●157

9 months ago

public beta start

5 months ago

private beta start

6 months ago

What is needed for a sustainable community?

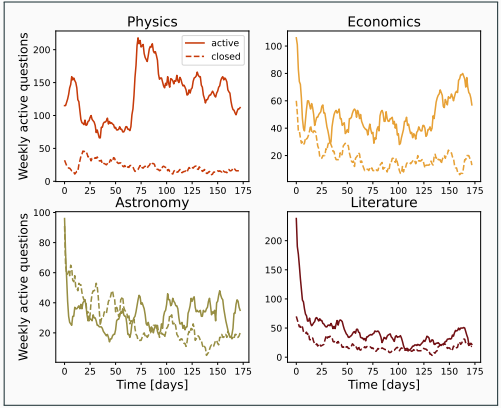
Comparison of *closed* and launched communities centered around similar topics:

- *Theoretical Physics* (233 days) — Physics
- *Economics* (206 days) — Economics
- *Astronomy* (338 days) — Astronomy
- *Literature* (269 days) — Literature

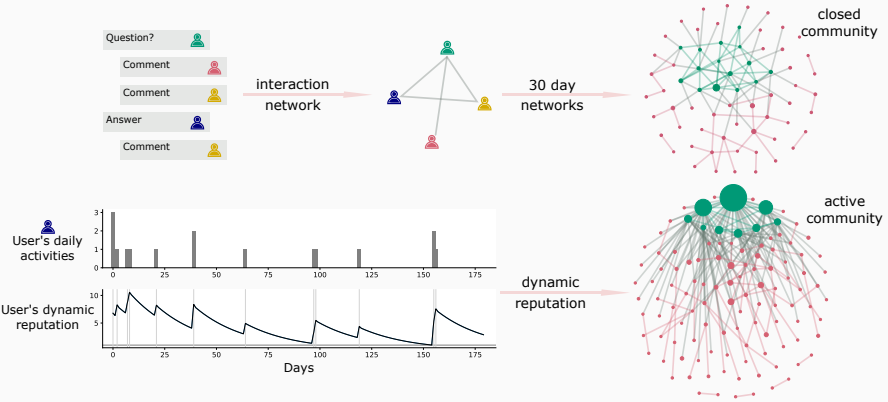
What is needed for a sustainable community?

Comparison of *closed* and **launched** communities centered around similar topics:

- *Theoretical Physics* (233 days) — **Physics**
- *Economics* (206 days) — **Economics**
- *Astronomy* (338 days) — **Astronomy**
- *Literature* (269 days) — **Literature**

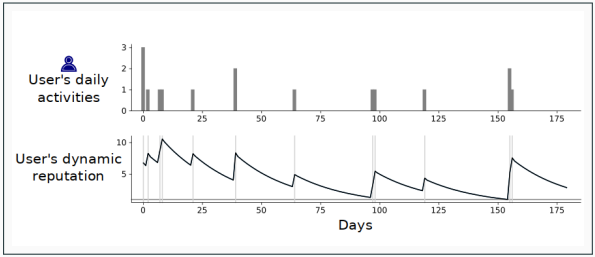


Our approach



Dynamic reputation

- User's reputation changes continuously through time
- Grows with repeated interaction within time window $\Delta_t = 1$ day
- Decays due to inactivity (if $\Delta_n > \Delta_t$)
- All interactions are considered equal and positive



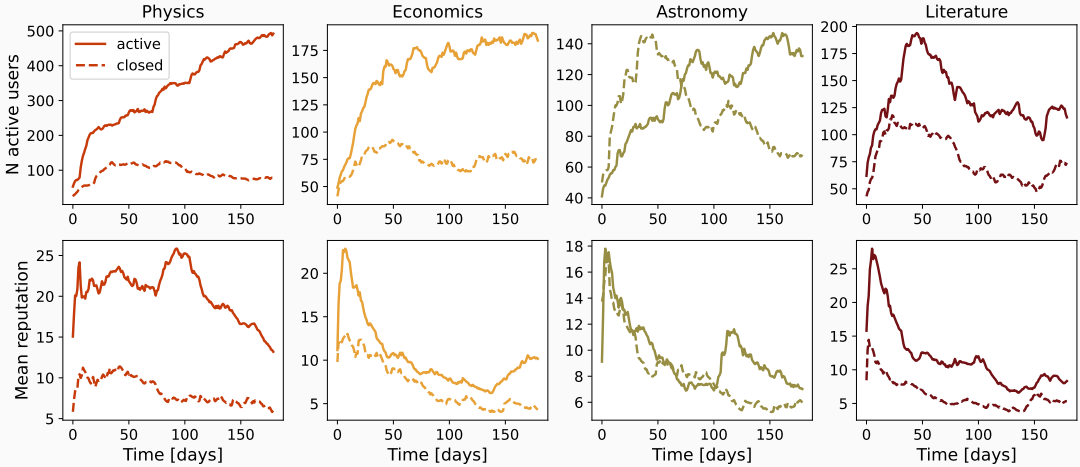
$$T_n = T_{n-1}\beta^{\Delta_n} + I_{b_n}\left(1 + \alpha\left(1 - \frac{1}{A_n + 1}\right)\right)$$

DIBRM model [1]

$$0 < \beta < 1$$

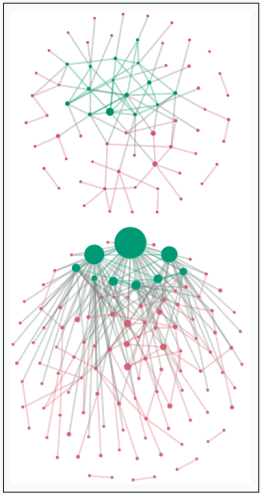
$$\alpha \geq 1$$

Active users and mean reputation

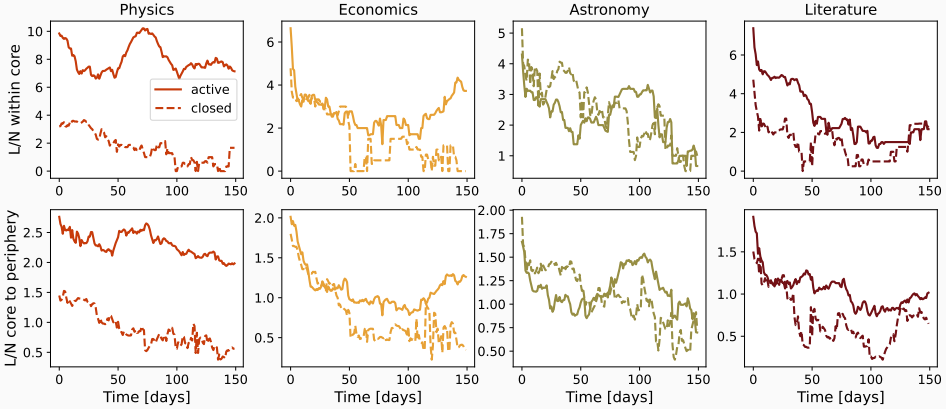


Networks of interactions

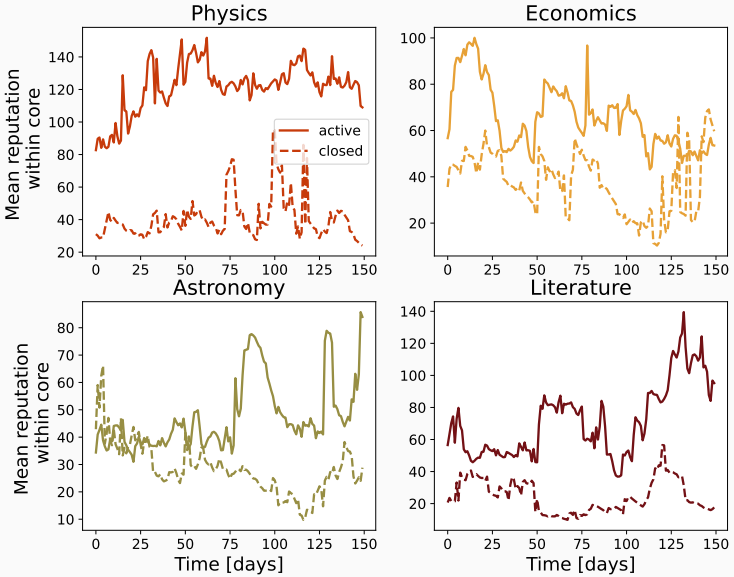
- Do popular users interact more among themselves?
- From first 180 days, capture 30-day user interaction networks with sliding window (+1 day)
- For each network, detect core-periphery structure using Bayesian stochastic block model [2]



Core-periphery link density

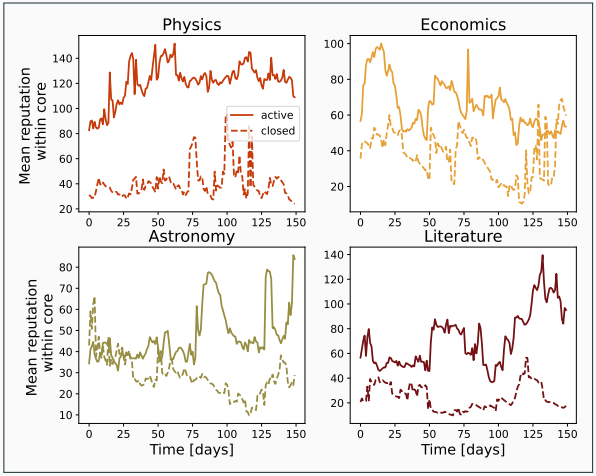


Reputation within the core



Summary

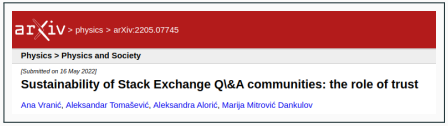
- Launched communities are more cohesive & inclusive
- Mean reputation is higher in active communities
- Outlier: Astronomy



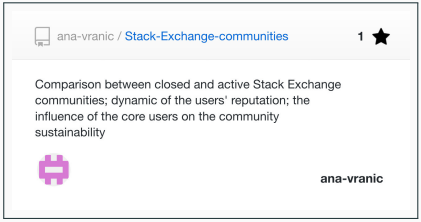
Thank you!

Thank you!

ArXiv preprint [3]



GitHub repository



Overview

Table 1: Community overview for the first 180 days.

Site	Status	First Date	n_u	n_q	n_a	n_c
Physics	Closed	09/14/11	281	349	564	2213
	Active	08/24/10	1176	2124	4802	15403
Economics	Closed	10/11/10	275	368	458	1253
	Active	11/18/14	648	1024	1410	3553
Astronomy	Closed	09/22/10	336	474	953	1444
	Active	09/24/13	405	644	959	2170
Literature	Closed	02/10/10	284	318	523	1097
	Active	01/18/17	478	910	907	3301

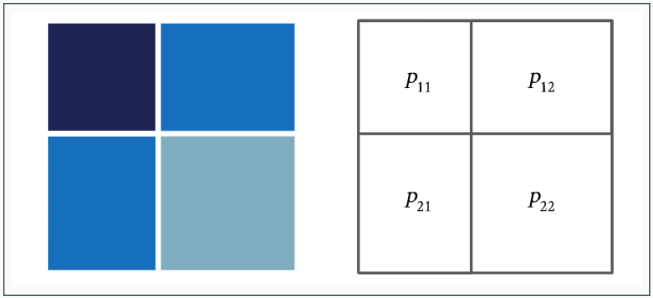
Note: Number of users n_u, number of questions n_q, number of answers n_a, number of comments n_c

Core-periphery detection

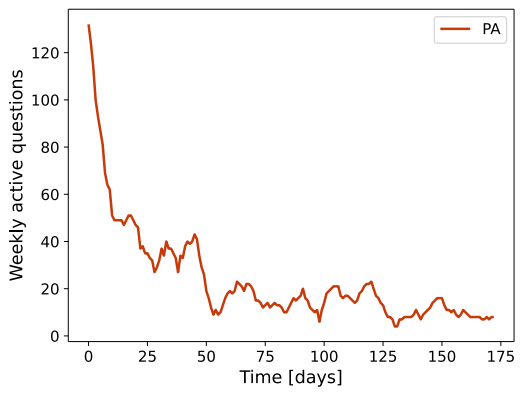
Bayesian Stochastic Block Model

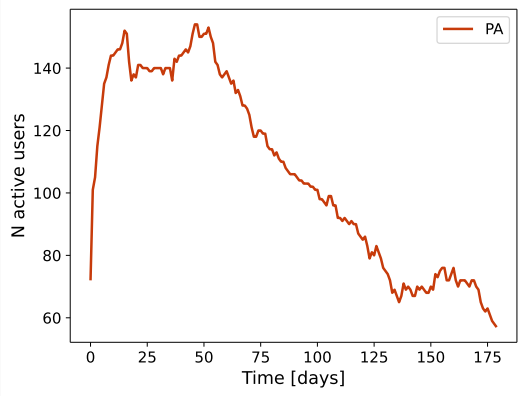
$$P(\theta, \mathbf{p} \mid \mathbf{A}) \propto P(\mathbf{A} \mid \theta, \mathbf{p}) P(\theta) P(\mathbf{p})$$

$$p_{11} > p_{12} > p_{22}$$

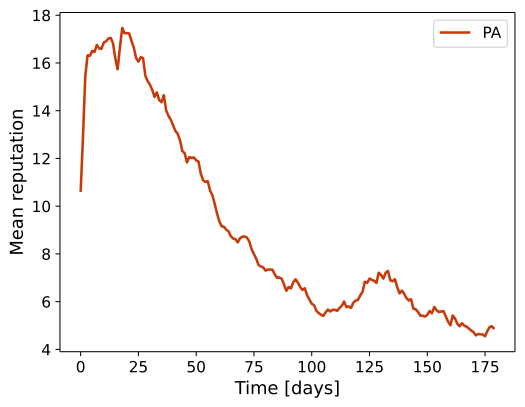


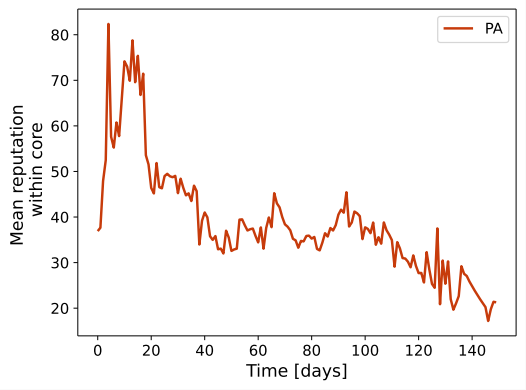
Proof Assistants



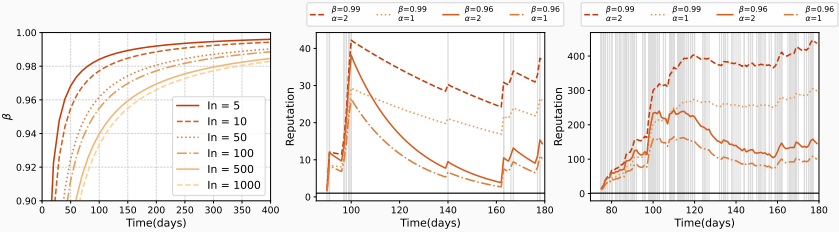


Proof Assistants

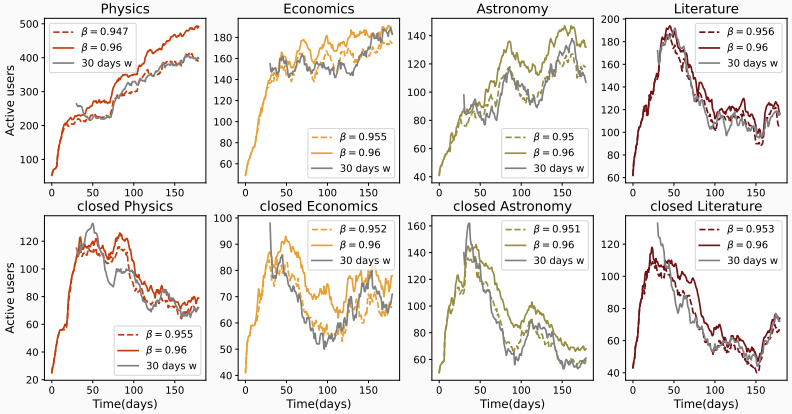




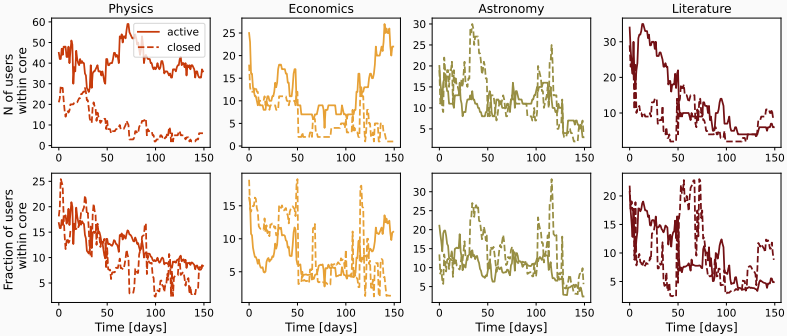
Selecting β



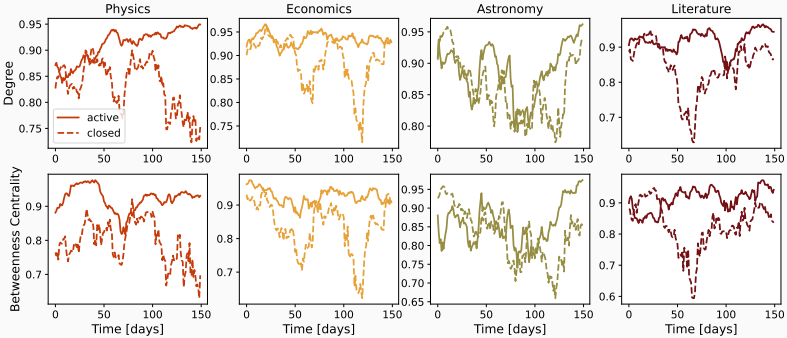
β and number of active users



Users within the core



Centrality


Physics
Economics
Astronomy
Literature
Degree
Betweenness Centrality
active
closed
Time [days]
Time [days]
Time [days]
Time [days]

References

[1] E. Yashkina, A. Pinigin, J. Lee, *et al.*, "Expressing Trust with Temporal Frequency of User Interaction in Online Communities," in *Advanced Information Networking and Applications*, L. Barolli, M. Takizawa, F. Xhafa, and T. Enokido, Eds., vol. 926, Cham: Springer International Publishing, 2020, pp. 1133–1146, ISBN: 978-3-030-15031-0 978-3-030-15032-7. DOI: 10.1007/978-3-030-15032-7_95.

[2] R. J. Gallagher, J.-G. Young, and B. F. Welles, "A clarified typology of core-periphery structure in networks," *Science Advances*, vol. 7, no. 12, eabc9800, 2021. DOI: 10.1126/sciadv.abc9800.

[3] A. Vranić, A. Tomašević, A. Alorić, and M. M. Dankulov, *Sustainability of Stack Exchange Q\&A communities: The role of trust*, 2022. DOI: 10.48550/arXiv.2205.07745. arXiv: 2205.07745 [physics].