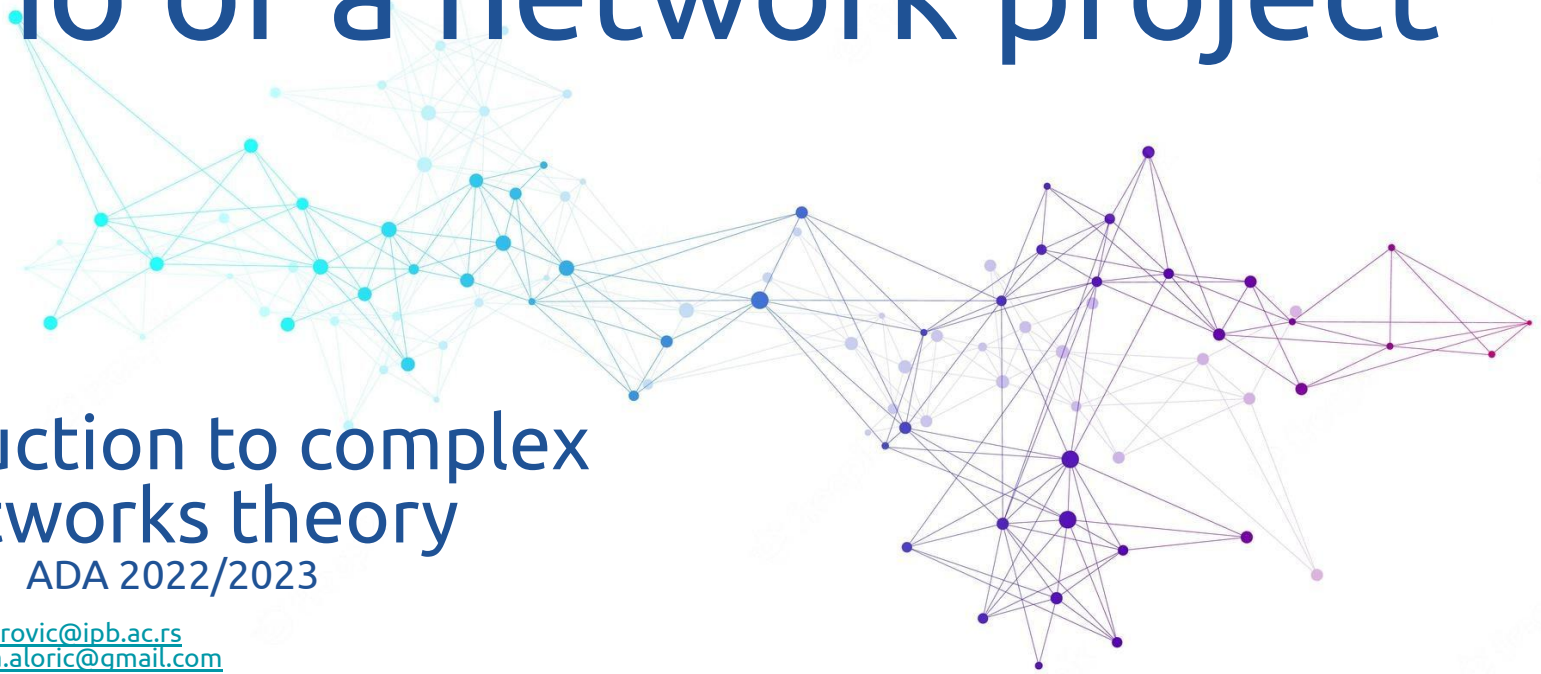
Demo of a network project



Introduction to complex networks theory

ADA 2022/2023

Marija Mitrović Dankulov mitrovic@ipb.ac.rs
Aleksandra Alorić aleksandra.aloric@gmail.com

Aim of this class

- Go through one research project that involves networks from start to ‘finish’

- Particularly:
  - Think about operationalizing research questions into network related measurements
  - Extract networks from data
  - Make research decisions about networks and understand their consequences
  - Discuss results and open new questions

Example project: Sustainability of SE communities

Aleksandar Tomašević

Ana Vranić

Marija Mitrović Dankulov

Aleksandra Alorić

General project motivation

- Googled a coding question, a bug, a syntax check?
  - Chances are the answer lead you to Stack Overflow
  - Did you copy/paste the answer? Why?

General project motivation

- Googled a coding question, a bug, a syntax check?
  - Chances are the answer lead you to Stack Overflow
  - Did you copy/paste the answer? Why?
- Who do we ask for information?
  - How do we decide which information source is trustworthy?
- How new Q&A communities emerge, how they generate interest, how they involve new members and also keep the old ones engaged?
  - Is there something within these growing mechanisms that can tell us a new community might strive or die out?

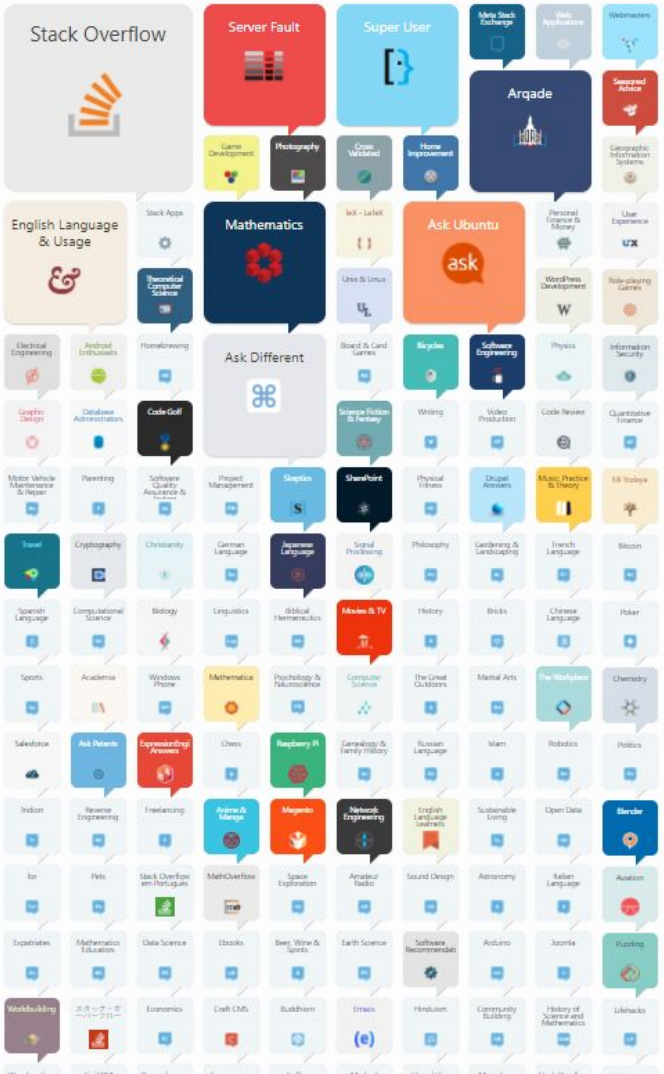
How can we address these questions?

- Theoretically e.g. we dive deep into social studies research on trusting communities and build upon existing theories
- Create questionnaires, organise interviews and focus groups with community members and learn from their experience of community participation
- Analyse digital traces of activities in communities that emerged and died over the years
- …
- Ideally, we want to combine all of these approaches and build on top of it, but today we talk about the data-driven approaches

- Is there some dataset to help us tackle some of the research questions?

StackExchange

- The Stack Exchange network comprises 173 Q&A communities (programming, sciences, language learning, hobies, etc.), 10+ years
- User interaction through: questions, answers, comments, votes…
- Past year on StackExchange:
  - 431.8M monthly visits
  - 3.2M questions asked
  - 3.5M answers submitted
  - 13.5M comments
  - 22.1M upvotes
  - 3M downvotes



define

- 40 “typical” questions
- 60 followers

commit

- 200 formally committed users
- User can commit to up to 3 proposals

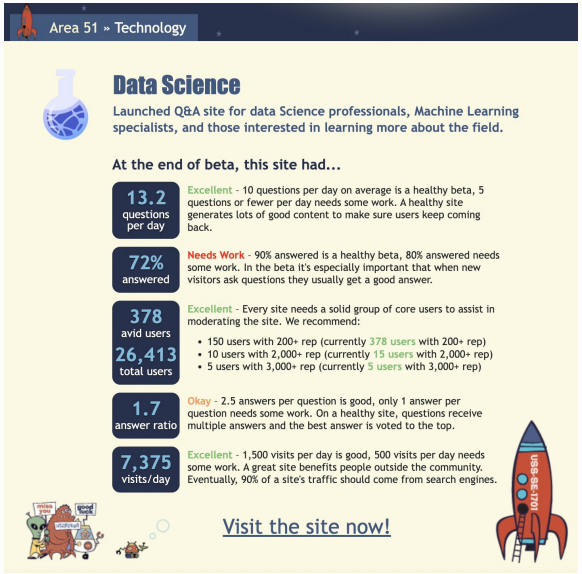
beta

- At least 90 days to reach:
  - High question per day rates
  - High answer rates
  - High visitation rates

Before **graduation**

What is needed for a sustainable community?

- Stack Exchange provide some benchmarks:
  - Number of questions per day
  - Percentage of answered questions
  - Answer to question ratio
  - Number of (active) users
  - Number of daily visits
- We were interested to know if there's something that is structurally different between still active and closed communities that promoted/hindered them in achieving these metrics

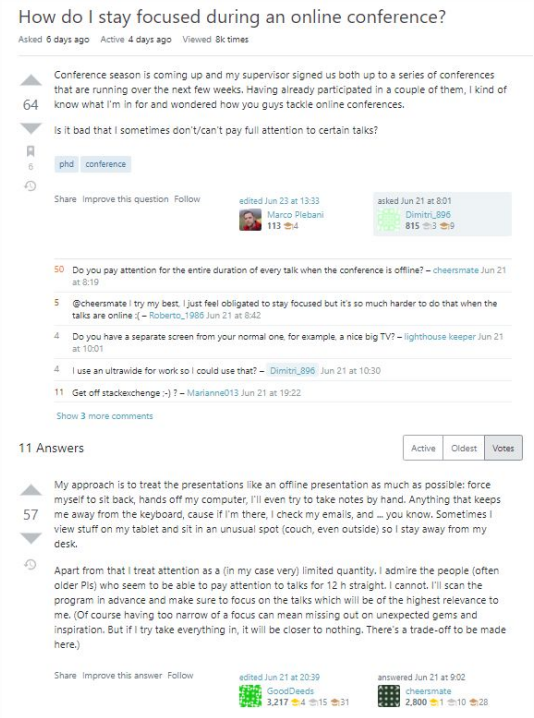


Some hypothesis about Q&A communities

- Let’s get back to our research questions so that we see if the available data can help us somehow
- Communities that survive after initial period are those that:
  - Generate a lot of interest (new activities, but also new members)
    - Have stable or growing number of active members
  - Generate activities that provoke participation of many users
  - Are inclusive towards new members
    - People who are already part of community answer their questions so that newcomers can feel welcome and participate themselves

Methods

Post interactions



How do I stay focused during an online conference?

Asked 6 days ago Active 4 days ago Viewed 8k times

64

Conference season is coming up and my supervisor signed us both up to a series of conferences that are running over the next few weeks. Having already participated in a couple of them, I kind of know what I'm in for and wondered how you guys tackle online conferences.

Is it bad that I sometimes don't/can't pay full attention to certain talks?

6

phd conference

Share Improve this question Follow

edited Jun 23 at 13:33
Marco Plebani
113 4

asked Jun 21 at 8:01
Dimitri_896
815 3 9

50 Do you pay attention for the entire duration of every talk when the conference is offline? – cheersmate Jun 21 at 8:19

5 @cheersmate I try my best. I just feel obligated to stay focused but it's so much harder to do that when the talks are online :(– Roberto_1986 Jun 21 at 8:42

4 Do you have a separate screen from your normal one, for example, a nice big TV? – lighthouse keeper Jun 21 at 10:01

4 I use an ultrawide for work so I could use that? – Dimitri_896 Jun 21 at 10:30

11 Get off stackexchenge :-) ? – Marianne013 Jun 21 at 19:22

Show 3 more comments

11 Answers

Active Oldest Votes

57

My approach is to treat the presentations like an offline presentation as much as possible: force myself to sit back, hands off my computer, I'll even try to take notes by hand. Anything that keeps me away from the keyboard, cause if I'm there, I check my emails, and ... you know. Sometimes I view stuff on my tablet and sit in an unusual spot (couch, even outside) so I stay away from my desk.

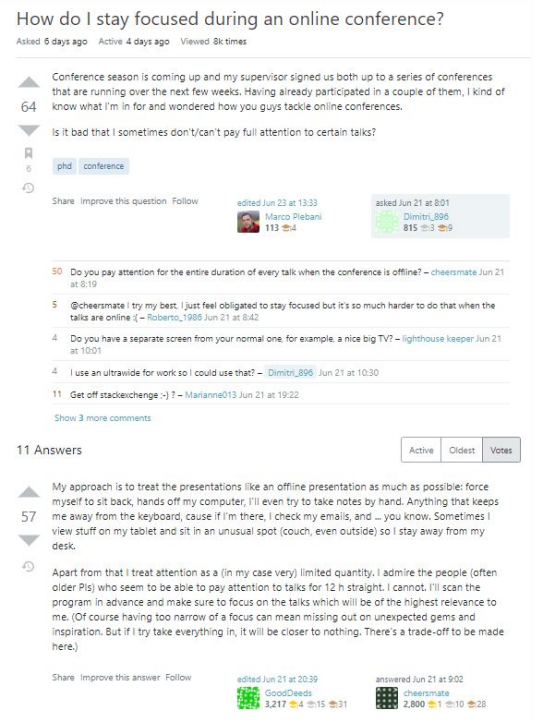
Apart from that I treat attention as a (in my case very) limited quantity. I admire the people (often older PIs) who seem to be able to pay attention to talks for 12 h straight. I cannot. I'll scan the program in advance and make sure to focus on the talks which will be of the highest relevance to me. (Of course having too narrow of a focus can mean missing out on unexpected gems and inspiration. But if I try take everything in, it will be closer to nothing. There's a trade-off to be made here.)

Share Improve this answer Follow

edited Jun 21 at 20:39
GoodDeeds
3,217 4 15 31

answered Jun 21 at 9:02
cheersmate
2,800 1 10 28

Post interactions


How do I stay focused during an online conference?

Asked 6 days ago Active 4 days ago Viewed 8k times

64

Conference season is coming up and my supervisor signed us both up to a series of conferences that are running over the next few weeks. Having already participated in a couple of them, I kind of know what I'm in for and wondered how you guys tackle online conferences.

Is it bad that I sometimes don't/can't pay full attention to certain talks?

phd conference

Share Improve this question Follow

edited Jun 23 at 13:33 Marco Plebani 113 ●4

asked Jun 21 at 8:01 Dimitri_896 815 ●3 ●9

50 Do you pay attention for the entire duration of every talk when the conference is offline? – cheersmate Jun 21 at 8:19

5 @cheersmate I try my best. I just feel obligated to stay focused but it's so much harder to do that when the talks are online :(– Roberto_1986 Jun 21 at 8:42

4 Do you have a separate screen from your normal one, for example, a nice big TV? – lighthouse keeper Jun 21 at 10:01

4 I use an ultrawide for work so I could use that? – Dimitri_896 Jun 21 at 10:30

11 Get off stackexchenge :-) ? – Marianne013 Jun 21 at 19:22

Show 3 more comments

11 Answers

Active Oldest Votes

57

My approach is to treat the presentations like an offline presentation as much as possible: force myself to sit back, hands off my computer, I'll even try to take notes by hand. Anything that keeps me away from the keyboard, cause if I'm there, I check my emails, and ... you know. Sometimes I view stuff on my tablet and sit in an unusual spot (couch, even outside) so I stay away from my desk.

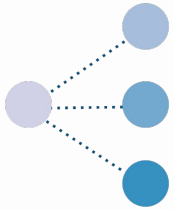
Apart from that I treat attention as a (in my case very) limited quantity. I admire the people (often older PIs) who seem to be able to pay attention to talks for 12 h straight. I cannot. I'll scan the program in advance and make sure to focus on the talks which will be of the highest relevance to me. (Of course having too narrow of a focus can mean missing out on unexpected gems and inspiration. But if I try take everything in, it will be closer to nothing. There's a trade-off to be made here.)

Share Improve this answer Follow

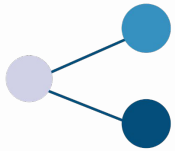
edited Jun 21 at 20:39 GoodDeeds 3,217 ●4 ●15 ●31

answered Jun 21 at 9:02 cheersmate 2,800 ●1 ●10 ●28

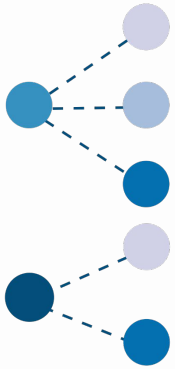
Question - Comments



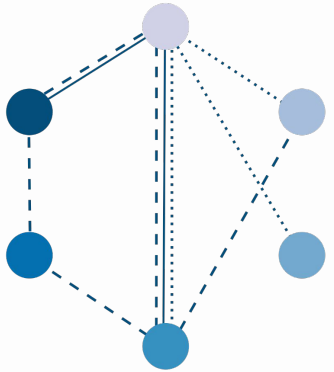
Question - Answers



Answers - Comments

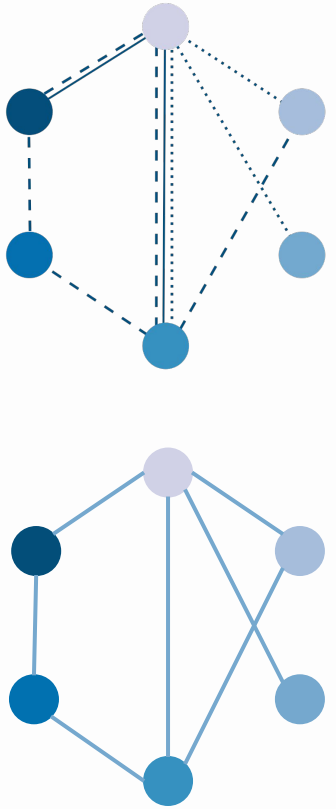


User interactions

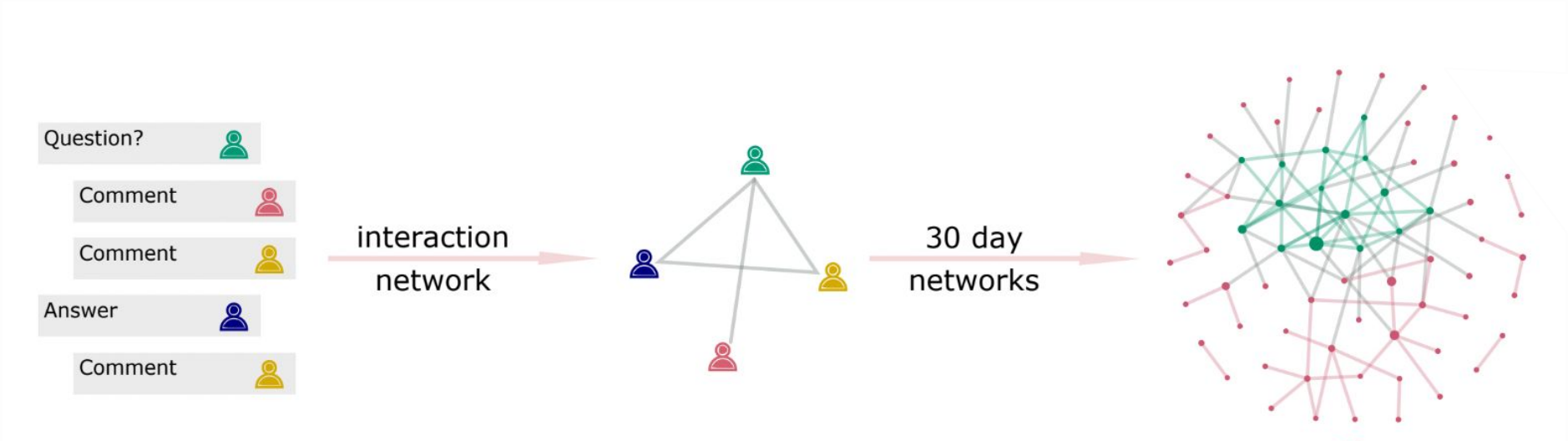


Mapping to a network

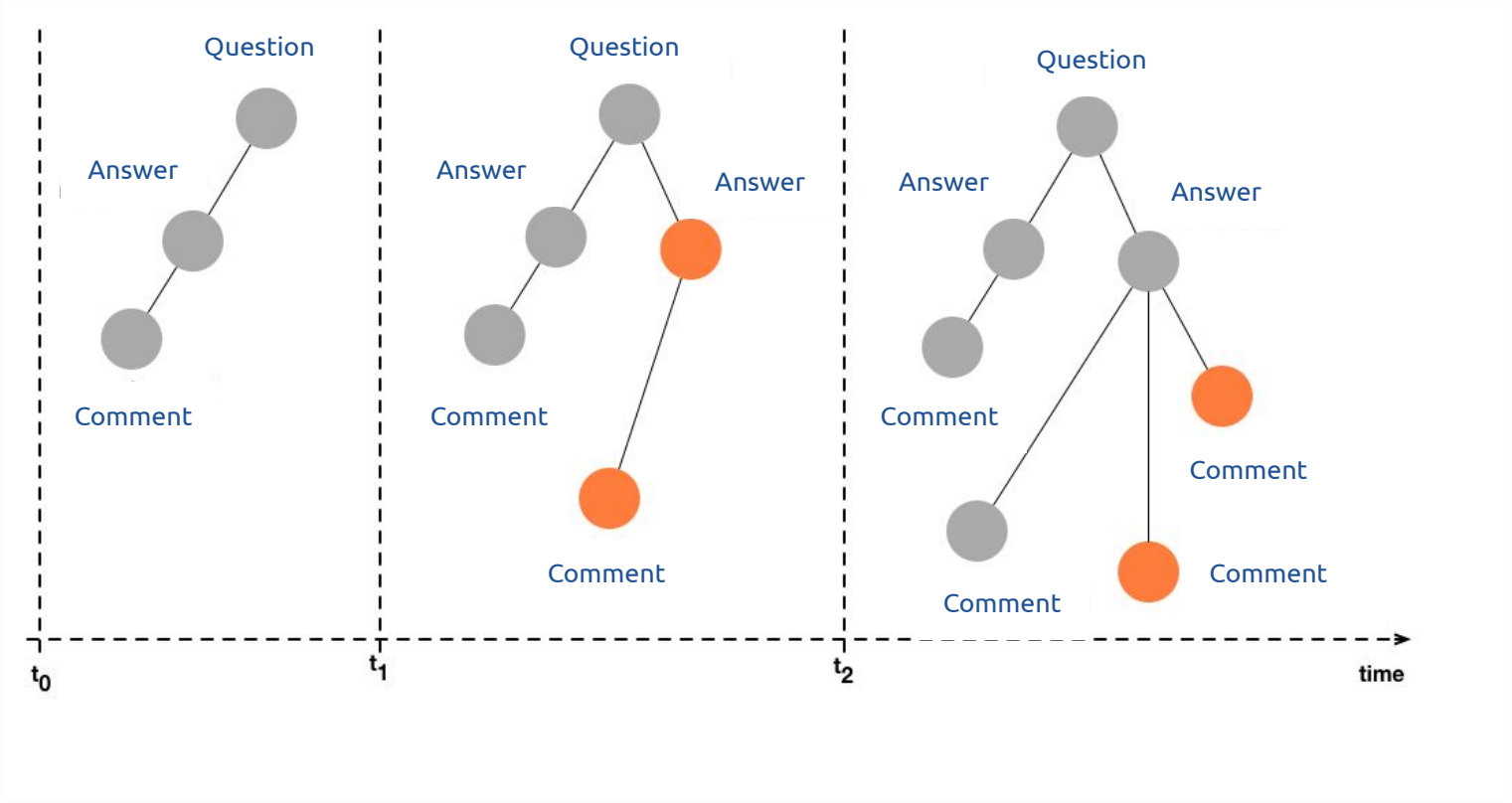
- Decisions:
  - Interactions shown could be mapped to a multiplex network where one layer is question - answers, another is question - comments and answer - comments -> we choose to treat all these interactions as equal
  - Users can repeat interactions - as question and comment on an answer, is that weighted link -> we choose to treat interactions as unweighted network
  - Network shown is only a network of interaction between users at a single post, we aggregate all interactions that happen within defined time window



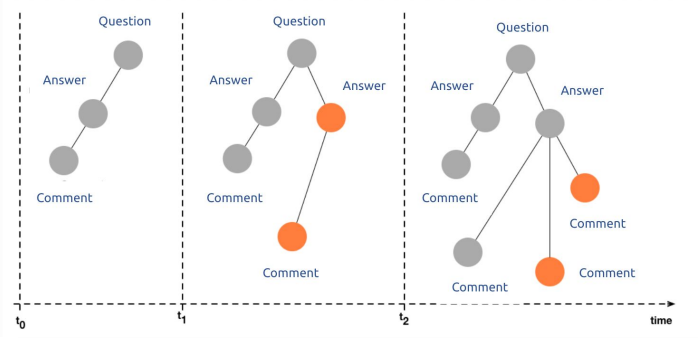
Mapping to a network

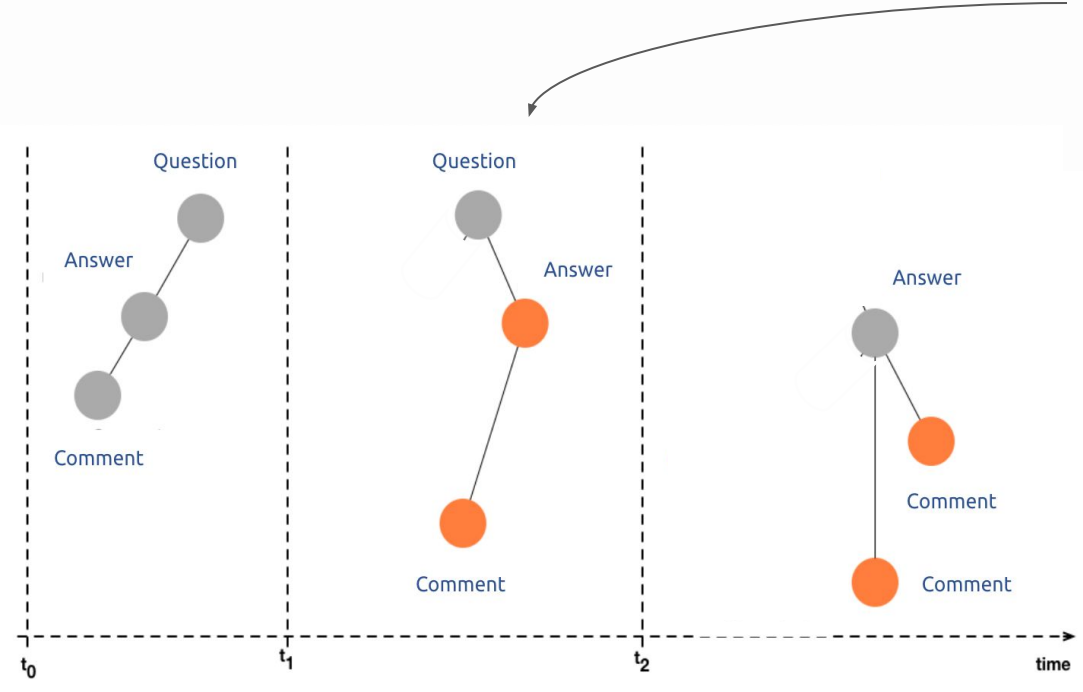


Temporal aspect of interactions



Interactions within a time window





Top choice could be useful when your main point of your research is network grow, but if you are interested in understanding some sort of current activity of user within some community, then the left choice could be better.

How network methods help us address hypothesis?

Hypothesis: Sustainable communities have stable or growing number of active members

- Measure size of interaction networks over time

Hypothesis: Sustainable communities are cohesive and generate activities that provoke participation of many users

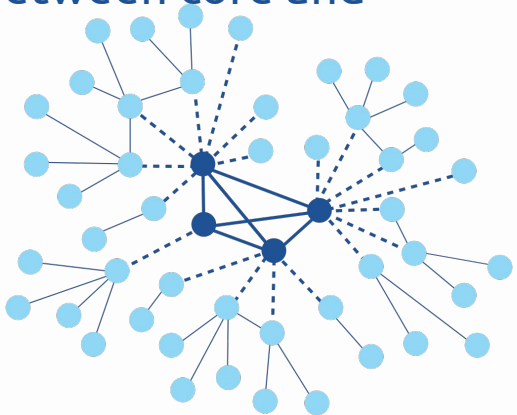
- Measure network connectedness, density, clustering coefficient

Hypothesis: Sustainable communities include new members

- Identify important users, measure density of connections between them and between them and less ‘important users’

How to operationalize measurement of community inclusion?

- We aim to investigate whether popular users interact more among themselves or they engage with newbies and less popular users
- Popular users could be users with the most posts, most votes, most accepted answers, etc.
- But we utilize networks for this too and find the core members of interaction network (more on core in the following lessons)
- We investigate how interactions within the core and between core and periphery change over time



Data

What kind of data is available

- Roughly once in 3 months, new data is available
- Where to find it and more about the data scheme you can read here

View of single data release:

Files for stackexchange_20220307

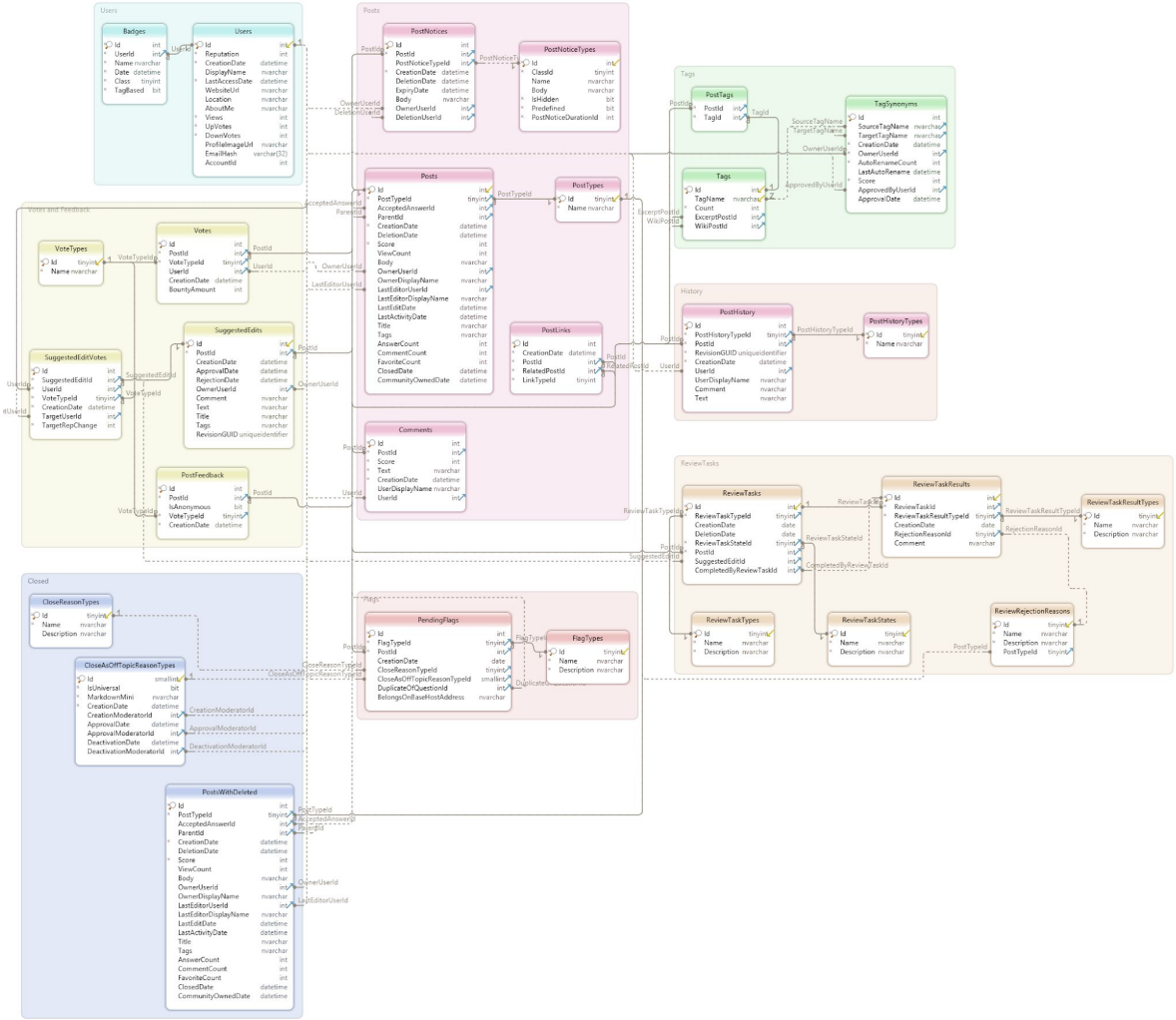
Name	Last modified	Size
Go to parent directory		
3dprinting.meta.stackexchange.com.7z (View Contents)	28-Apr-2022 18:35	673.6K
3dprinting.stackexchange.com.7z (View Contents)	28-Apr-2022 18:35	15.0M
Sites.xml	28-Apr-2022 18:56	340.9K
__ia_thumb.jpg	28-Apr-2022 18:56	189.0B
academia.meta.stackexchange.com.7z (View Contents)	28-Apr-2022 18:35	4.9M
academia.stackexchange.com.7z (View Contents)	28-Apr-2022 18:36	143.3M
ai.meta.stackexchange.com.7z (View Contents)	28-Apr-2022 18:36	947.0K
ai.stackexchange.com.7z (View Contents)	28-Apr-2022 18:36	26.1M
android.meta.stackexchange.com.7z (View Contents)	28-Apr-2022 18:36	3.0M
android.stackexchange.com.7z (View Contents)	28-Apr-2022 18:36	108.5M
anime.meta.stackexchange.com.7z (View Contents)	28-Apr-2022 18:36	4.0M
anime.stackexchange.com.7z (View Contents)	28-Apr-2022 18:36	32.1M
apple.meta.stackexchange.com.7z (View Contents)	28-Apr-2022 18:36	4.3M
apple.stackexchange.com.7z (View Contents)	28-Apr-2022 18:36	248.2M
arduino.meta.stackexchange.com.7z (View Contents)	28-Apr-2022 18:36	956.9K
arduino.stackexchange.com.7z (View Contents)	28-Apr-2022 18:36	66.9M
askubuntu.com.7z (View Contents)	28-Apr-2022 18:36	915.0M
astronomy.meta.stackexchange.com.7z (View Contents)	28-Apr-2022 18:36	1,006.0K
astronomy.stackexchange.com.7z (View Contents)	28-Apr-2022 18:36	37.5M
aviation.meta.stackexchange.com.7z (View Contents)	28-Apr-2022 18:37	2.3M

View of data for single community:

file	as jpgtimestamp	size
	2022-03-07 01:49:44	90288513
Badges.xml	2022-03-07 01:48:17	55161957
Comments.xml	2022-03-07 01:48:29	355187144
PostHistory.xml	2022-03-07 01:49:16	1625893797
PostLinks.xml	2022-03-07 01:49:41	13096971
Posts.xml	2022-03-07 01:49:40	870257697
Tags.xml	2022-03-07 01:49:41	77297
Votes.xml	2022-03-07 01:49:53	213243290

Data structure

- Rich data source for various research questions
- But pay attention it is optimized for its main purpose (functioning of SE) not research, so for your research question you will need to make it useful and understand its limits



Content of posts.xml file:

```
<?xml version="1.0" encoding="utf-8"?>
<posts>
  <row Id="1" PostTypeId="1" AcceptedAnswerId="51" CreationDate="2016-01-12T18:45:19.963" Score="10" ViewCount="389" Body="&lt;p&gt;When I've printed an object I've had to choose between high resolution and quick prints.  What techniques or technologies can I use or deploy to speed up my high resolution prints?&lt;/p&gt;&#xA;" OwnerUserId="16" LastActivityDate="2017-10-31T02:31:08.560" Title="How to obtain high resolution prints in a shorter period of time?" Tags="&lt;resolution&gt;&lt;speed&gt;&lt;quality&gt;" AnswerCount="2" CommentCount="6" ContentLicense="CC BY-SA 3.0" />
  <row Id="2" PostTypeId="1" AcceptedAnswerId="12" CreationDate="2016-01-12T18:45:51.287" Score="31" ViewCount="6156" Body="&lt;p&gt;I would like to buy a 3D printer, but I'm concerned about the health risks that are associated with its operation. Some groups of scientists say it can be &lt;a href=&quot;http://www.techworld.com/news/personal-tech/scientists-warn-of-3d-printing-health-effects-as-tech-hits-high-street-3460992/&quot;&gt;harmful&lt;/a&gt; for humans.&lt;/p&gt;&#xA;&#xA;&lt;p&gt;What do I need to consider before buying a 3D printer if I care about my health? Are there any safe printers?&lt;/p&gt;&#xA;" OwnerUserId="20" LastEditorUserId="334" LastEditDate="2016-11-15T16:16:11.163" LastActivityDate="2019-06-10T23:18:34.190" Title="Is 3D printing safe for your health?" Tags="&lt;print-material&gt;&lt;safety&gt;&lt;health&gt;" AnswerCount="4" CommentCount="1" FavoriteCount="4" ContentLicense="CC BY-SA 3.0" />
  <row Id="3" PostTypeId="1" AcceptedAnswerId="152" CreationDate="2016-01-12T18:46:22.083" Score="18" ViewCount="2625" Body="&lt;p&gt;I know the minimum layer height will effect how detailed of an item you can print and the amount of time it takes to print something, but is it necessary to have an extremely low minimum layer height if you plan to print only larger objects?&lt;/p&gt;&#xA;" OwnerUserId="11" LastEditorUserId="11" LastEditDate="2016-01-12T22:00:36.347" LastActivityDate="2016-09-19T15:41:06.537" Title="How important is the minimum layer height on a 3d printer?" Tags="&lt;quality&gt;&lt;resolution&gt;" AnswerCount="3" CommentCount="5" FavoriteCount="2" ContentLicense="CC BY-SA 3.0" />
  <row Id="4" PostTypeId="1" AcceptedAnswerId="1289" CreationDate="2016-01-12T18:50:55.973" Score="17" ViewCount="360" Body="&lt;p&gt;Plastic is used in 3D FDM/FFF printing partly because it had a wide temperature range for its glass state - where it can be flowed with some force, but won't flow due only to gravity.&lt;/p&gt;&#xA;&#xA;&lt;p&gt;Most metals have a very narrow, or non-existant, glass state.  They transition from solid to liquid with almost no flowable-but-not-liquid state.&lt;/p&gt;&#xA;&#xA;&lt;p&gt;Are there any metals or alloys that display a glass transition state?&lt;/p&gt;&#xA;" OwnerUserId="16" LastEditorUserId="98" LastEditDate="2016-06-09T02:10:35.890" LastActivityDate="2016-06-10T13:32:20.493" Title="Are there any metals that exhibit a large glass state?" Tags="&lt;fdm&gt;&lt;material&gt;&lt;print-material&gt;&lt;metal-parts&gt;" AnswerCount="4" CommentCount="0" FavoriteCount="2" ContentLicense="CC BY-SA 3.0" />
  <row Id="5" PostTypeId="1" AcceptedAnswerId="77" CreationDate="2016-01-12T18:53:53.623" Score="38" ViewCount="3713" Body="&lt;p&gt;What are the main differences when using ABS over PLA and vice versa?&lt;/p&gt;&#xA;" OwnerUserId="11" LastEditorUserId="20" LastEditDate="2016-01-15T17:02:37.707" LastActivityDate="2017-08-02T09:49:07.263" Title="How is PLA different from ABS material?" Tags="&lt;filament&gt;&lt;abs&gt;&lt;fdm&gt;&lt;pla&gt;" AnswerCount="5" CommentCount="5" FavoriteCount="6" ContentLicense="CC BY-SA 3.0" />
  <row Id="6" PostTypeId="1" AcceptedAnswerId="27" CreationDate="2016-01-12T18:57:13.350" Score="11" ViewCount="609" Body="&lt;p&gt;My MakerBot printer supports only two filaments at the same time.&lt;/p&gt;&#xA;&#xA;&lt;p&gt;What are techniques to print objects with more than two colors for one object?&lt;/p&gt;&#xA;" OwnerUserId="20" LastEditorUserId="8884" LastEditDate="2018-09-16T12:35:19.097" LastActivityDate="2018-09-16T12:35:19.097" Title="Multi-color printing with desktop 3D printer?" Tags="&lt;filament&gt;&lt;makerbot&gt;&lt;dual-nozzle&gt;&lt;color&gt;" AnswerCount="5" CommentCount="0" FavoriteCount="1" ContentLicense="CC BY-SA 3.0" />
  <row Id="7" PostTypeId="5" CreationDate="2016-01-12T18:57:48.103" Score="0" Body="&lt;p&gt;Filament is the plastic strands used as the print material for 3d printers. The most common types are PLA and ABS. There are also other types of filament that used other types of plastics as well as other materials.&lt;/p&gt;&#xA;" OwnerUserId="11" LastEditorUserId="11" LastEditDate="2016-01-15T17:04:10.283" LastActivityDate="2016-01-15T17:04:10.283" CommentCount="0" ContentLicense="CC BY-SA 3.0" />
  <row Id="8" PostTypeId="4" CreationDate="2016-01-12T18:57:48.103" Score="0" Body="For questions related to different filaments used as the print material." OwnerUserId="11" LastEditorUserId="11" LastEditDate="2016-01-15T17:04:07.180" LastActivityDate="2016-01-15T17:04:07.180" CommentCount="0" ContentLicense="CC BY-SA 3.0" />
```

Posts structure

- Always go to the data documentation, image on the right is taken from here
- For the methods we've described, we need to know:
  - if post is a Question or Answer (so we need PostTypeId)
  - when the post was created (CreationDate)
  - who created it (OwnerUserId)
  - etc..

Posts / PostsWithDeleted

You find in `Posts` all non-deleted posts. `PostsWithDeleted` includes rows with deleted posts while sharing the same columns with `Posts` but for deleted posts only a few fields populated which are marked with a [1] below.

- `Id`[1]
- `PostTypeId`[1] *(listed in the `PostTypes` table)*

  1 = Question
  2 = Answer
  3 = Orphaned tag wiki
  4 = Tag wiki excerpt
  5 = Tag wiki
  6 = Moderator nomination
  7 = "Wiki placeholder" *(seems to only be the election description)*
  8 = Privilege wiki
- `AcceptedAnswerId` *(only present if `PostTypeId = 1`)*
- `ParentId`[1] *(only present if `PostTypeId = 2`)*
- `CreationDate`[1]
- `DeletionDate`[1] *(only non-null for the SEDE `PostsWithDeleted` table. Deleted posts are not present on `Posts`. Column not present on data dump.)*
- `Score`[1]
- `ViewCount` *(nullable)*
- `Body` *(as rendered HTML, not Markdown)*
- `OwnerUserId` *(only present if user has not been deleted; always -1 for tag wiki entries, i.e. the community user owns them)*
- `OwnerDisplayName` *(nullable)*
- `LastEditorUserId` *(nullable)*
- `LastEditorDisplayName` *(nullable)*
- `LastEditDate` (e.g. `2009-03-05T22:28:34.823`) - *the date and time of the most recent*

Custom scripts to extract needed data

- Scripts used for this project you can find here

```
#filter questions
#file = '../astronomy/Posts.xml'

xml_file = sys.argv[1] #'../data/launched/astronomy/Posts.xml'
csv_file = sys.argv[2] #'./launched/astronomy/astronomy_questions.csv
tree = etree.parse(xml_file)
root = tree.getroot()

columns = ["QId", "QUserId", "QTags", "QTime", "QAcceptedAnswerId", "
#dataframe = pd.DataFrame(columns = columns)
dictn = {}
i=0
for node in root:
    OID=node.get("OwnerUserId")
    PST = node.get("PostTypeId")
    if PST=="1":
        ID = node.get("Id")
        Tags = node.get("Tags")
        Time=node.get("CreationDate")
        AAId=node.get("AcceptedAnswerId")
        Score=node.get("Score")
        VCount=node.get("ViewCount")
        ACount=node.get("AnswerCount")
        CCount=node.get("CommentCount")
        dictn[i] = {"QId": ID, "QUserId": OID, "QTags": Tags, "QTime"
"QAnswerCount":ACount, "QCommentCount":CCount}
        i+=1


dataframe = pd.DataFrame.from_dict(dictn, "index")
```

Data analysis steps

Raw data extraction

- We get useful data from raw files
  - Time of interaction
  - Type of interaction
  - Participants who interact
- xml -> csv

Create networks from interaction data

- Select:
  - Data within certain timeframe
  - Types of interactions based on which network is constructed
  - Type of network representation adequate for the problem

Calculate network properties

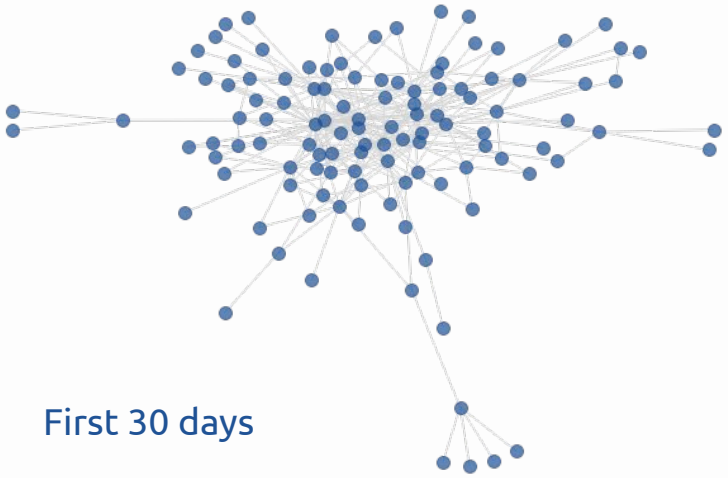
- Measure network properties:
  - Size
  - Density
  - Degrees
  - Clustering coefficient

Stack Exchange website on theoretical physics

Single network data

- For different time window sizes we get networks of different size, density, etc
- In what follows, we will focus on networks constructed based on interactions within 30 days and we will look how features of networks change over time
  - Size
  - Connectedness
  - Density
  - Clustering coefficient

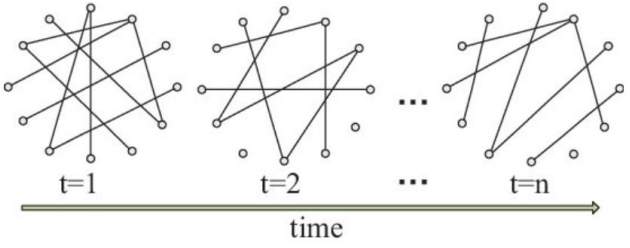
First 10 days



First 30 days

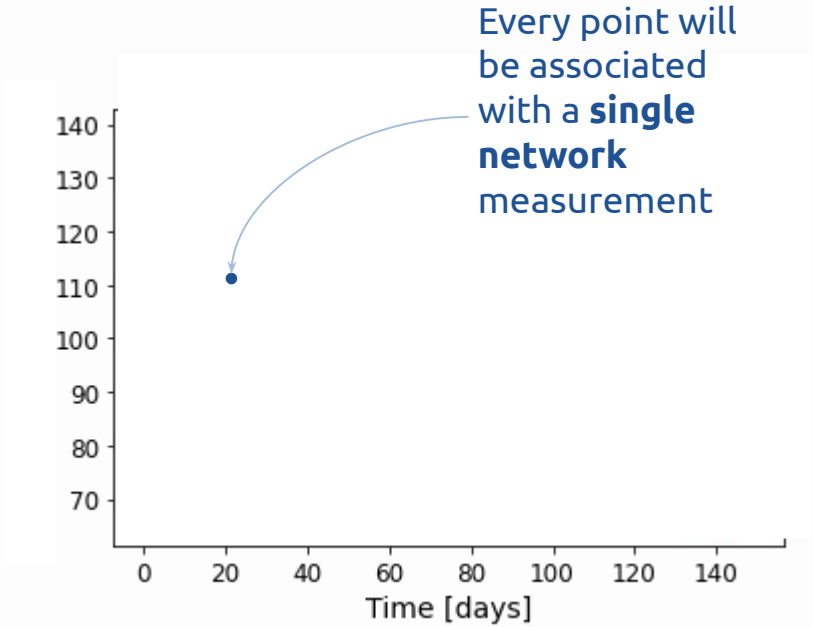
Sliding window network

- Every day, for the period of 6 months, we create a new network based on 30 days of interactions
  - E.g. interactions from day 0 to day 29 (included), from day 1 to day 30, … from day 149 to 178
- These 150 different networks we analyse and try to learn about SE community based on changes in network structure



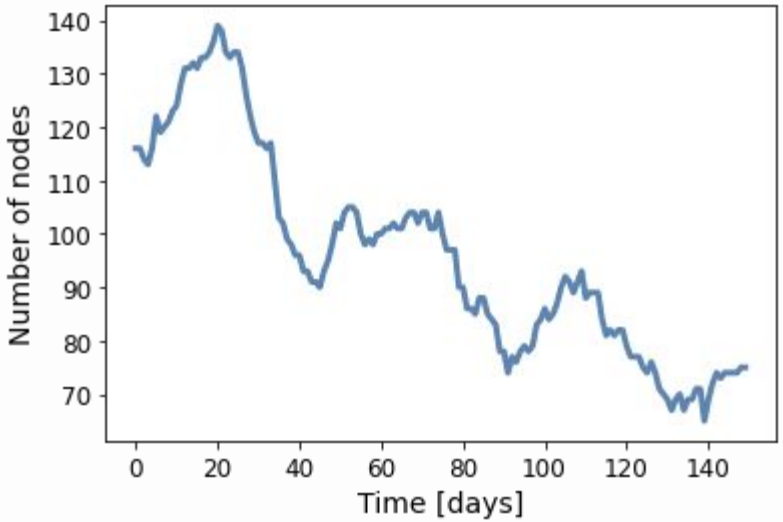
Data presentation

- In the following slides, we will look at time series of network changes
- The x axis will always be time (initial day of network time window) and on the y axis we will show network characteristic we measured



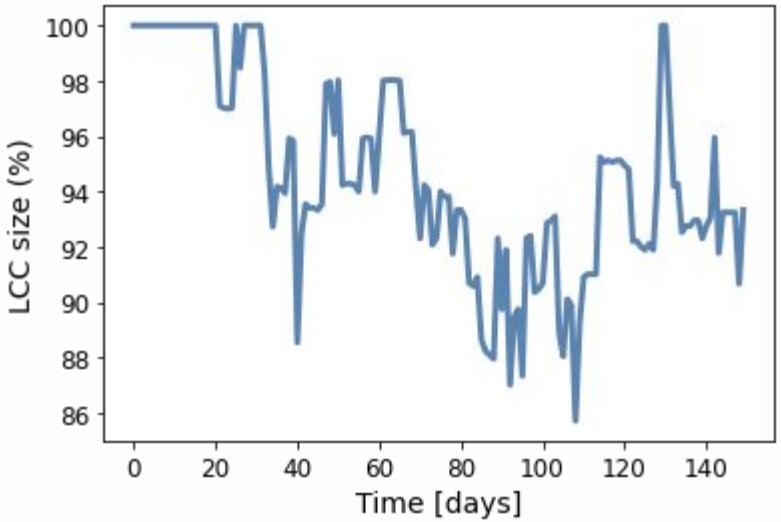
Network size

- We see that initially (first 20 days) number of nodes in the network was growing
  - New users were engaging in question/answer/comment interactions
- After the initial period of growth, we see overall decaying trend, with a few short term increase in participation, but network size never reaches initial maximum



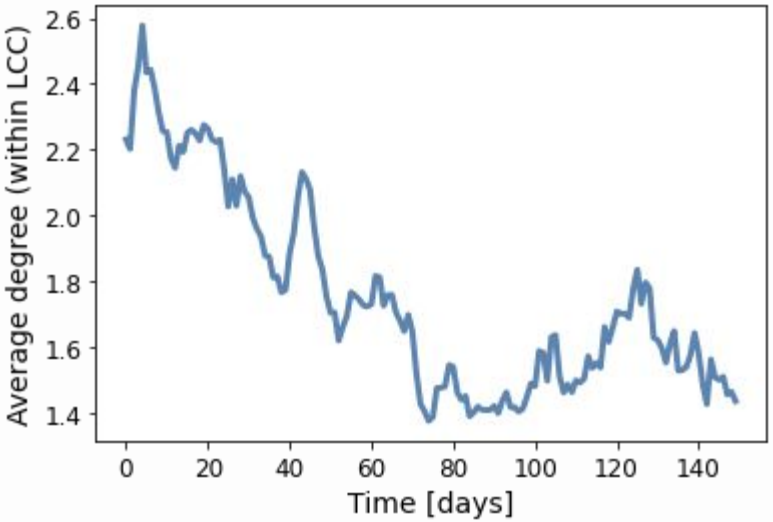
Network connectedness

- In the previous slide we only counted number of nodes (users who participated in discussions) without considering how many of them are part of a single connected component
- The plot on the right shows how big fraction (of all nodes) is within the largest connected component
- We see that initially (first ~20 days) the network was connected
- Later, largest connected component had between 85% and 98% of all nodes



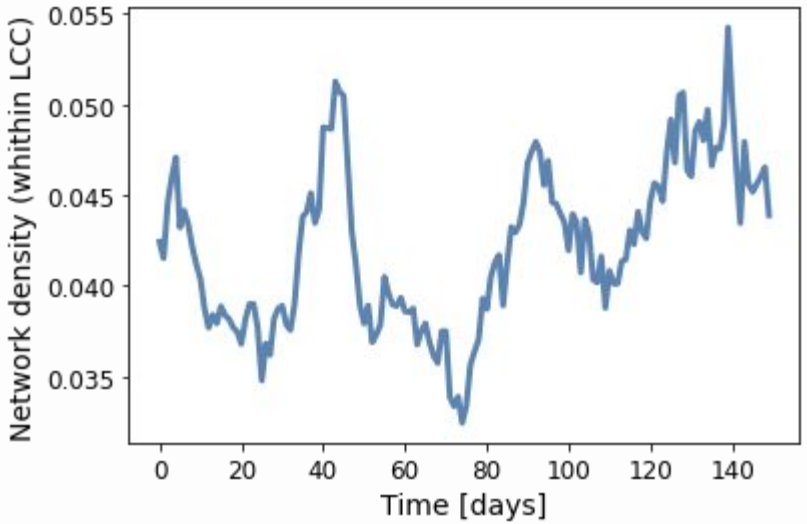
Network average degree

- Not only that the number of people who engage in the discussion (total number of nodes) decreases over time, but the average number of users single user interact with (average degree)
- We that initially users on average had 2.2 to 2.6 nearest neighbours, but as the time passes that number decays to 1.4



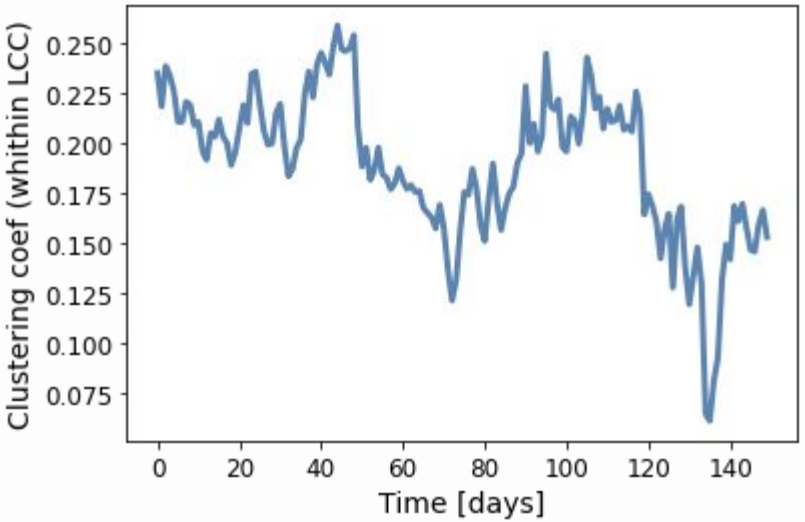
Network density

- When studying network density, we don’t see a trend that persists over time
- There is around 4% of all possible links present in the network



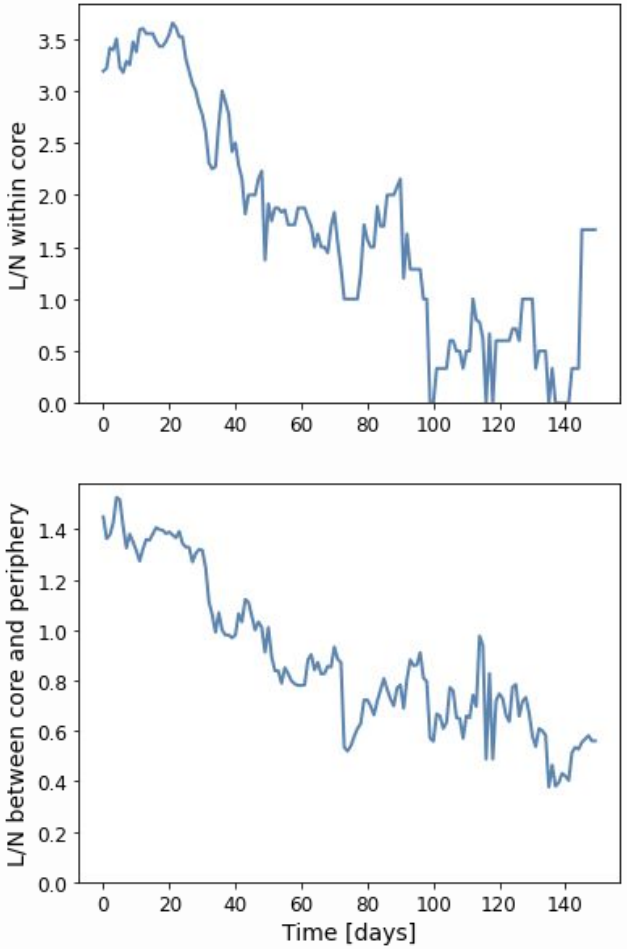
Network clustering coefficient

- We see that previously observed drop in overall density (around day 70) is also present as one of minimal values in the local clustering coefficient
  - This is interesting because this is not the low point in terms of number of users, this only means that user interactions are sparser
- Lowest clustering value (when on average only 7% of triadic closures are present in nodes neighbourhoods) happens near the end of studied period, and at that point overall density si not that low



Network mixing

- We see that near the end the most important users (core members) are not connected among themselves (see time>100 in top plot) although they do connect with peripheral nodes



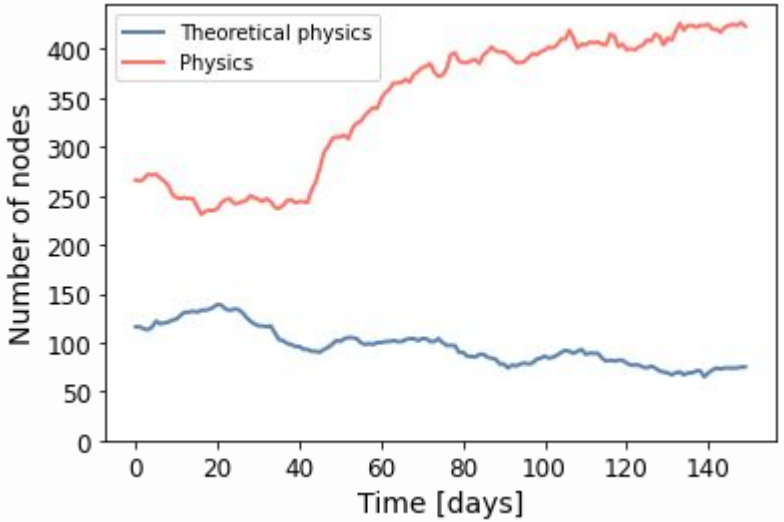
How can we interpret these measurements?

- When there is a clear trend we might be able to conclude that over time size of network decreases or connectivity is stable, etc.
- But questions like ‘is the growth fast enough, or is the decay dangerously close to community diminish’ are hard to answer based on a single measurement (although it is many time measurements, it is still a single community)
- We need some point of reference, or benchmark

Comparison with physics community

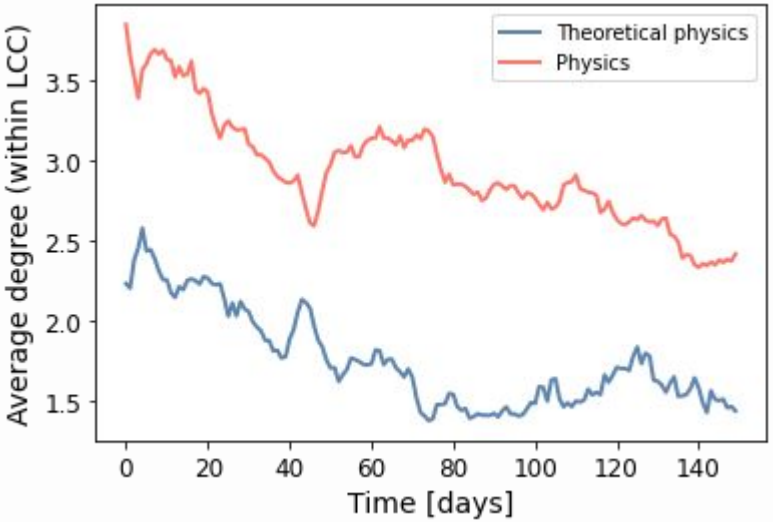
Network size

- Physics community already at start has more users in the network and their number quickly starts growing



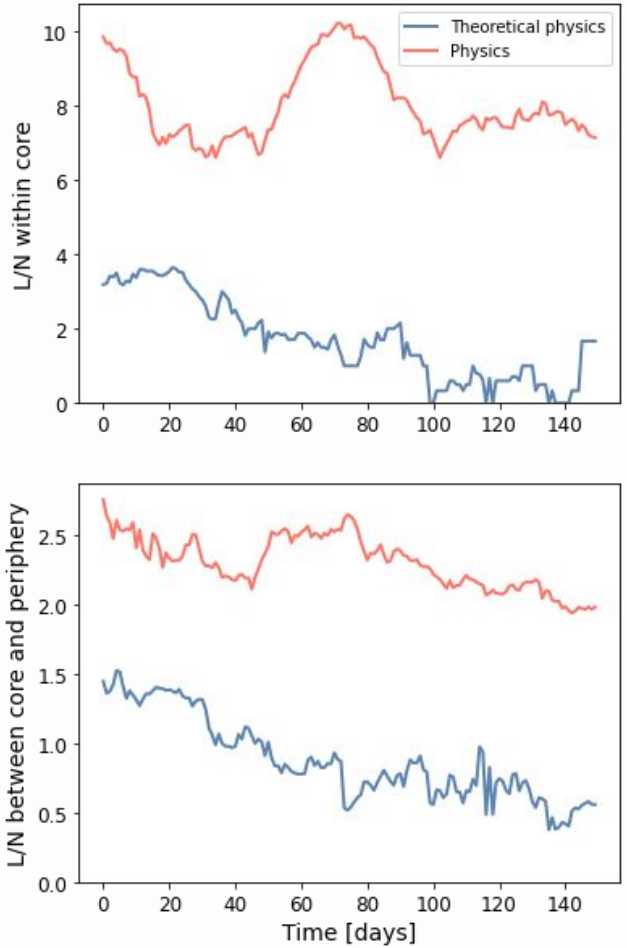
Network average degree

- Similar to the theoretical physics, physics community average degree also decays over time
- But average degree on start and also at its lowest is much higher than in theoretical physics community, signalling better connectivity among members



Network mixing

- Also when we look at mixing among important users and between them and the rest of the network we see that the physics community has higher metrics:
  - Core itself is better connected with average around 8 connections per node (compared to around 2 and less)
  - Core connects well to the periphery with two to three more times connections compared to theoretical physics



Other reference points

- Is physics community the best for comparison (what about differences in users, topics, etc)
- Is this way of growth/activity the only one, maybe low levels of activity in Theoretical physics are fine
- What about network measurement, is clustering coefficient (or any other measure we've got) too high, too low, can we create some expectations for a network with same or similar characteristics (say N nodes and L links)

- Those are our next topics - **random models of networks**

Research discussion

- Community survival
  - Physics community was fully launched
  - Theoretical physics community was in beta phase for 233 days until it was discounted
- Our conclusion (based on other pairs of communities) is that active communities are more cohesive & inclusive
- What can we change and investigate effects of our research decisions:
  - type of network
  - network window size
  - choice of interactions (include comments or not)
- How generalizable is this research? Does it tell us only about the stack exchange or there is something more general about online question and answer platforms (we could do something similar investigating for example Quora)

Further reading

- The preprint about the research I've been talking about is here and all codes to reproduce results are here
- A book that could be interesting to those of you who aim to work in computational social science is Bit by bit (not related to networks, just general social science in digital world)

Thanks