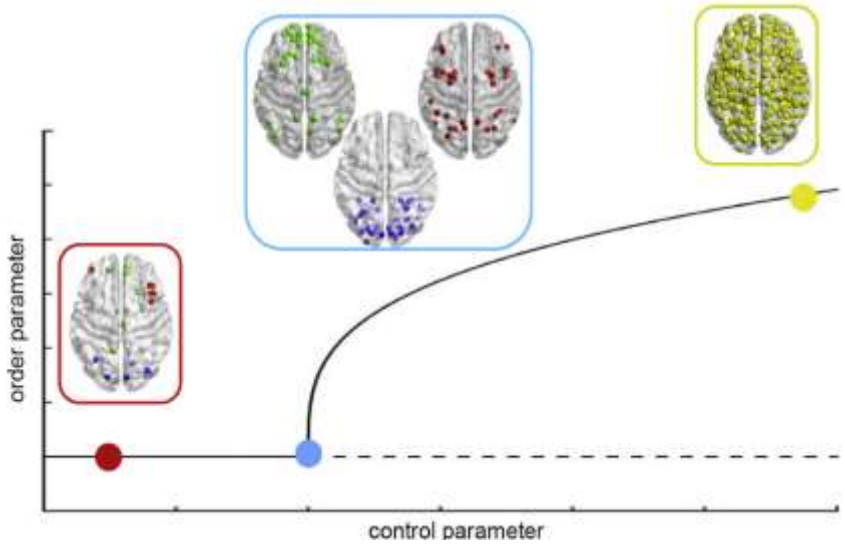
NETWORK SCIENCE OF ONLINE INTERACTIONS

Joao Neto

April 2023

ABOUT ME

- BSc and MSc: Physics
  - Statistical mechanics, network theory
- PhD: Computational Neuroscience
  - Spreading models applied to brain activity



(Cocchi et al, 2017)

- PhD: COVID-19 modeling

Science, *369*(6500), eabb9789 (2020)

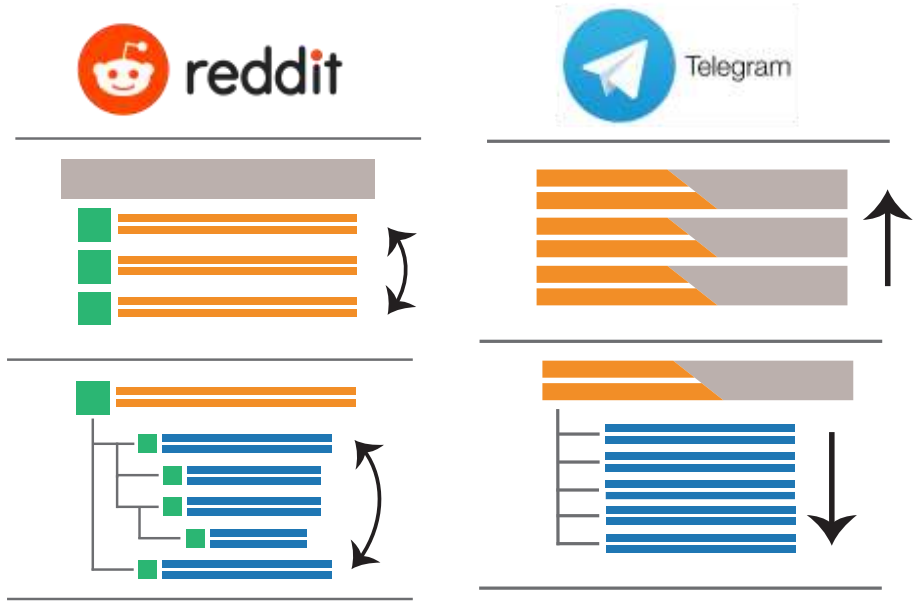
- Post-PhD: computational social science

reddit

Prof. David Garcia

ABOUT ME

- What I'm most interested in:
  - Understanding the structure and dynamics of different social media platforms
  - Enforceable mechanisms to deal with issues (polarization, fake news, echo chambers)
  - Influence policy-making



- How I approach it:
  - Data-driven modeling of social media
  - Toy models to explore mechanisms
  - Large-scale analysis to provide insight into platforms
    - Current data hoard: 11+ platforms, 60+TB

COURSE OBJECTIVES

- You want to understand some very complicated system (say, interactions on a social media platform) and are given some large dataset.

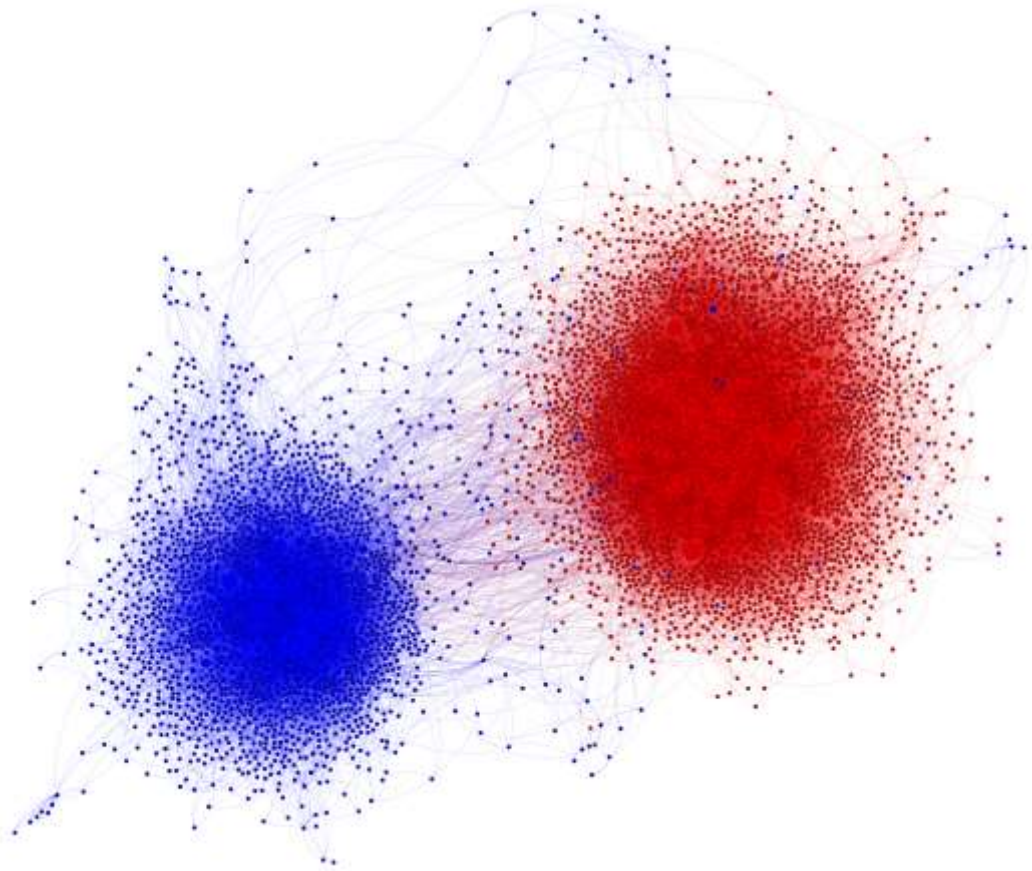
- How do you go about trying to understand it?

(Jan 6 US Capitol attack)

COURSE OBJECTIVES

- Networks are powerful abstractions that can help make sense of large interacting systems
- But:
  - not everything that *can* be represented as a network *should*
  - Not all methods are equally valid for all situations

- **Primary course objectives**
  - learn network theory to analyze interactions between large groups of agents
  - Develop insight of **when** to use network theory and **which parts** of it
  - Apply this new toolset to understand some aspect of Reddit



(a retweet network during the 2010 US election)

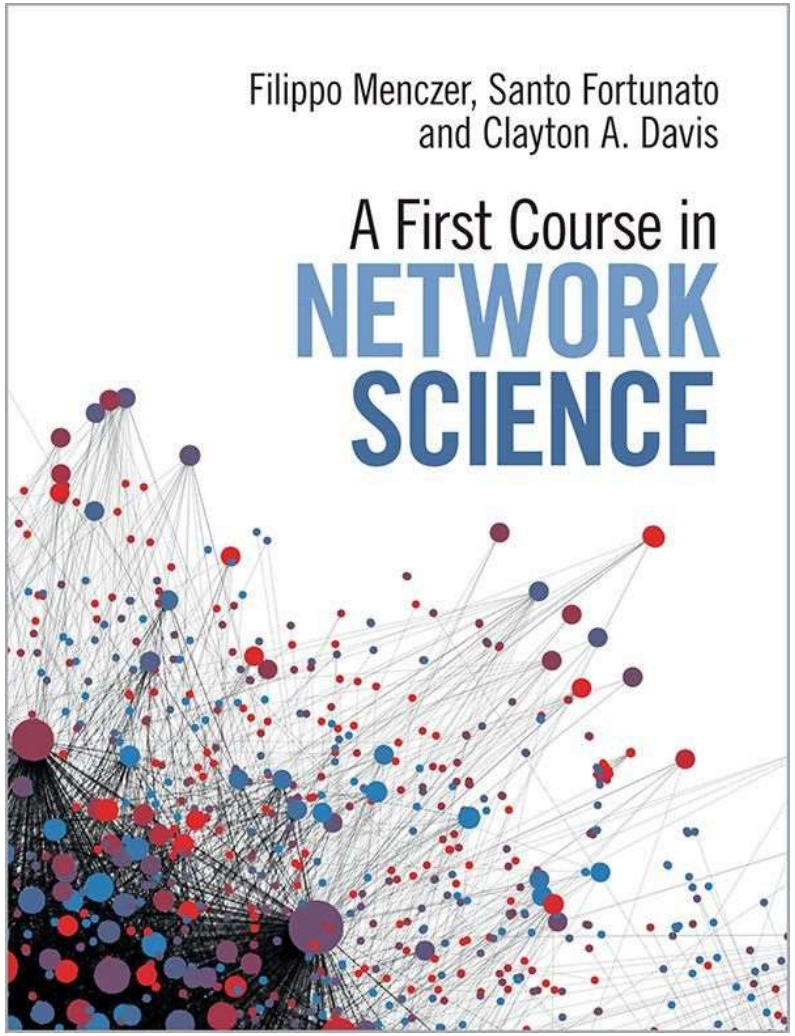
ADMINISTRATIVIA

- Course grading: Reddit project in groups of 3 max
  - 50% from the project presentation
  - 50% from a written report
  - 10% extra by developing the project on Github, publishing a notebook that can generate the figures shown in the report.

- Other information
  - Project needs to be registered by **11/Jun/23**
  - Website: https://github.com/joaopn/teaching_networks_2023
  - Project guideline (with suggestions) available at the website.

SCHEDULE

- Book: "A First Course in Network Science" by Menczer et al
  - Very hands-on, uses Python and NetworkX
- **Friday** lecture schedule
  - Apr 14: Chapter 0 - Introduction
  - Apr 21: Chapter 1 - Network Elements
  - Apr 28: Chapter 2 - Small Worlds
  - May 5: Chapter 3 - Hubs
  - May 12: Chapter 4 - Directions and Weights
  - May 19: Chapter 5 - Network Models
  - May 26: Chapter 6 - Communities
  - Jun 2: Chapter 7 - Dynamics
  - Jun 16: Critical Phenomena on Networks
  - Jun 23: Polarization Dynamics on Networks
  - Jun 30: Scaling laws and Inequality in Social Media



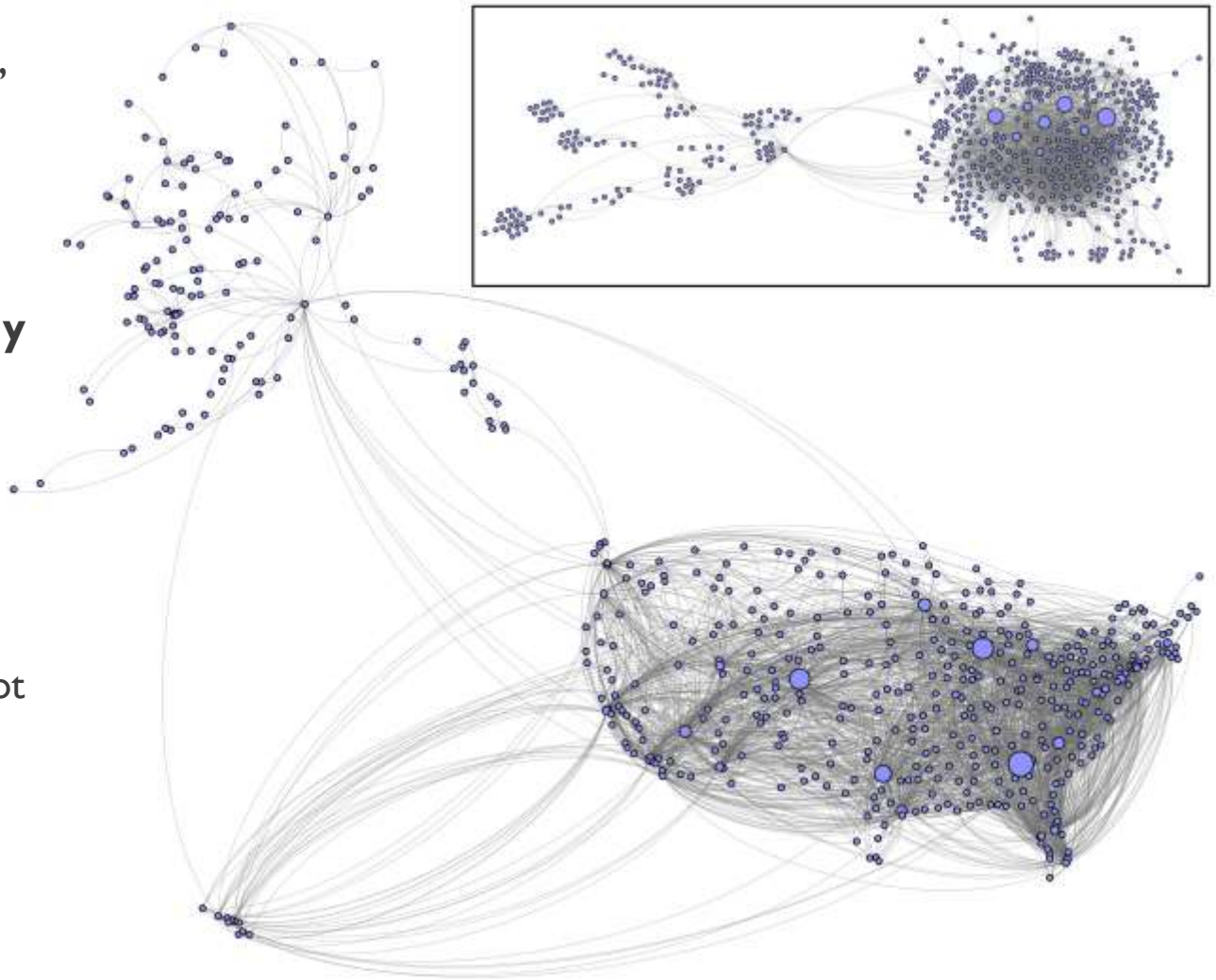
SCHEDULE

- **Wednesdays**: short (<30 min) talk on some related topic that may be useful for your project + exercise discussion
  - Apr 19: Python refresh + coding tools (Github, Copilot, chatGPT)
  - Apr 26: Software carpentry
  - May 3: Data gathering: sites, APIs
  - **May 10: A primer on Reddit**
  - **May 17: A primer on Reddit part II**
  - May 24: Detecting inorganic dynamics on Social Media
  - May 31: Descriptive vs Inference methods
  - Jun 14: Handling large datasets/databases
  - Jun 21: Sampling bias and social media data
  - Jun 28: Depolarization mechanisms

QUESTIONS SO FAR?

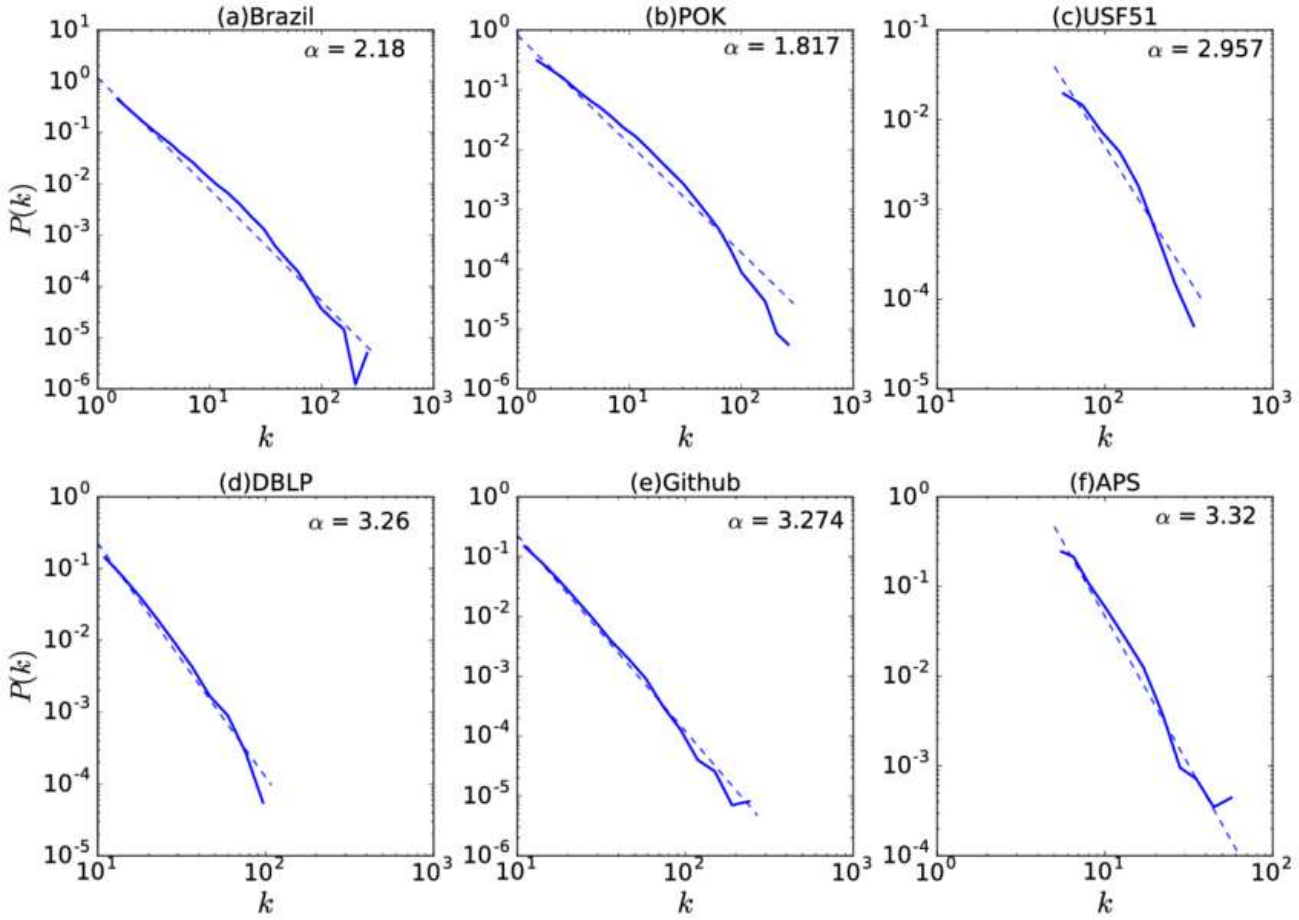
CHAPTER 0 - INTRODUCTION

- Networks are abstract representations of systems, made of links and nodes
- What the links and nodes represent depends on the system, and there is no unique way
- Results of your network analysis depends **crucially** on how you build it
- Example: US flight network
  - Straightforward: nodes are airports, links are routes
  - But what if the layout is not spatial?
  - What if we are interested in the amount of traffic, not only if a route exists?



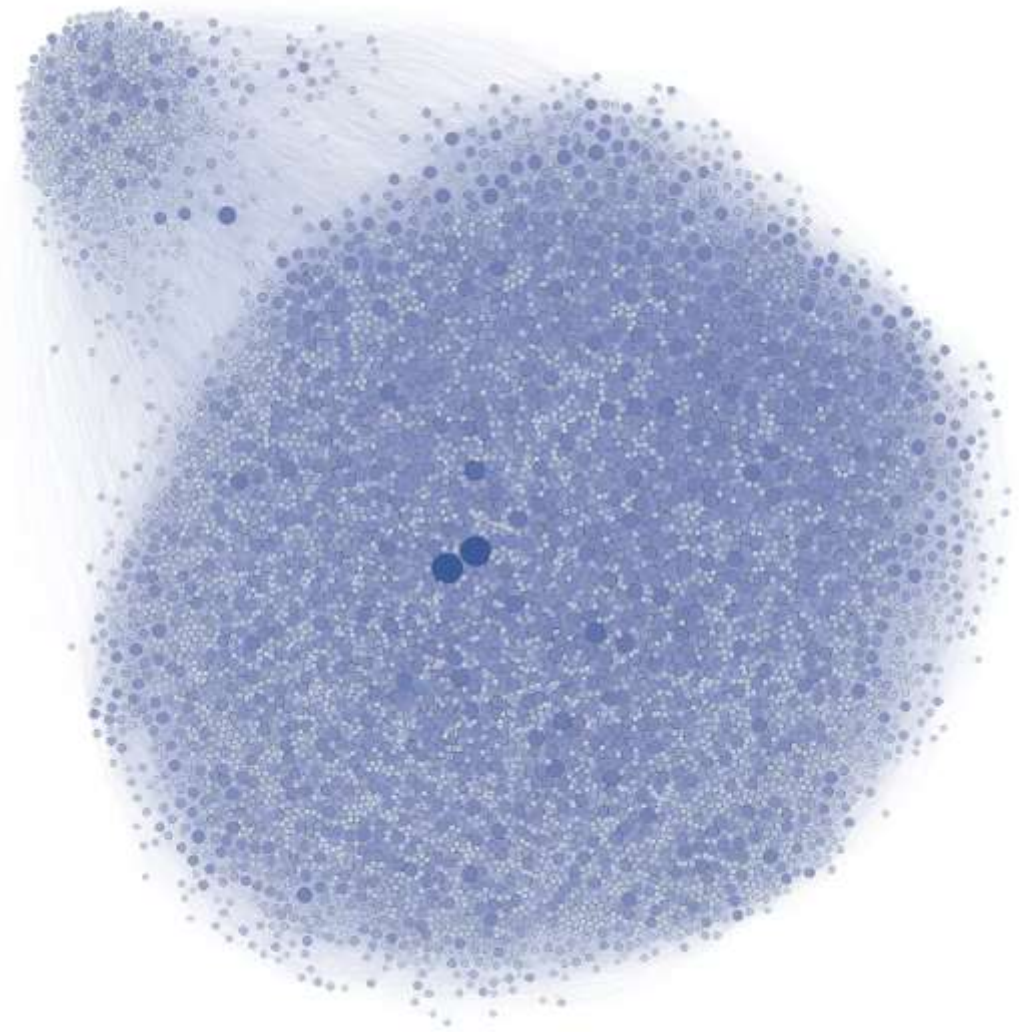
CHAPTER 0 - INTRODUCTION

- Networks can be grouped by type, e.g.
  - Social networks
  - Communication networks
  - Information networks (e.g. the Web, Wikipedia)
  - Transportation networks
  - Biological networks
- Benefits
  - Networks of the same type tend to have more similar properties
  - Example: degree distribution tends to be lighter for social and biological networks



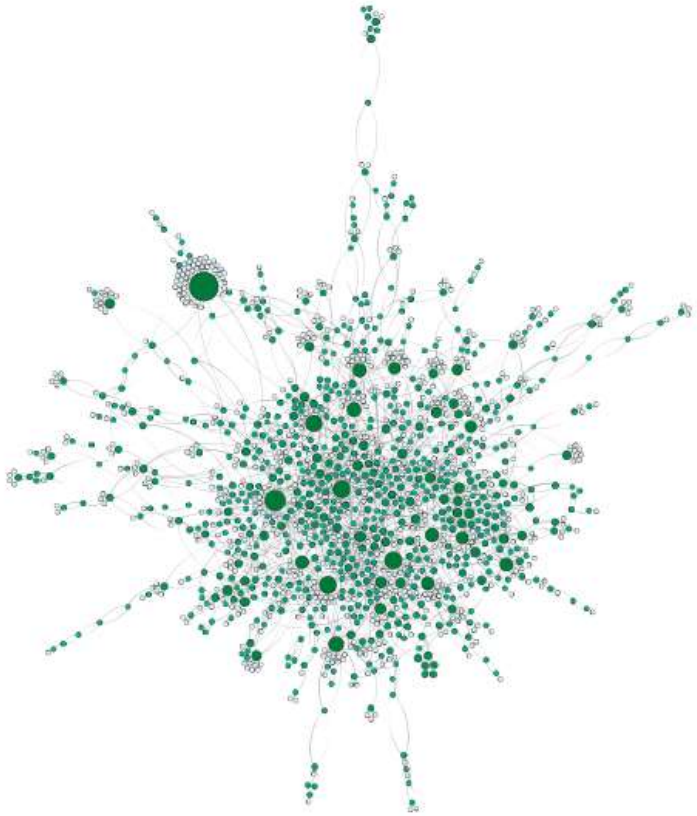
CHAPTER 0 - INTRODUCTION

- Social networks
  - Facebook users at Northwestern University
  - Nodes are users
  - Should links have direction and/or weights?

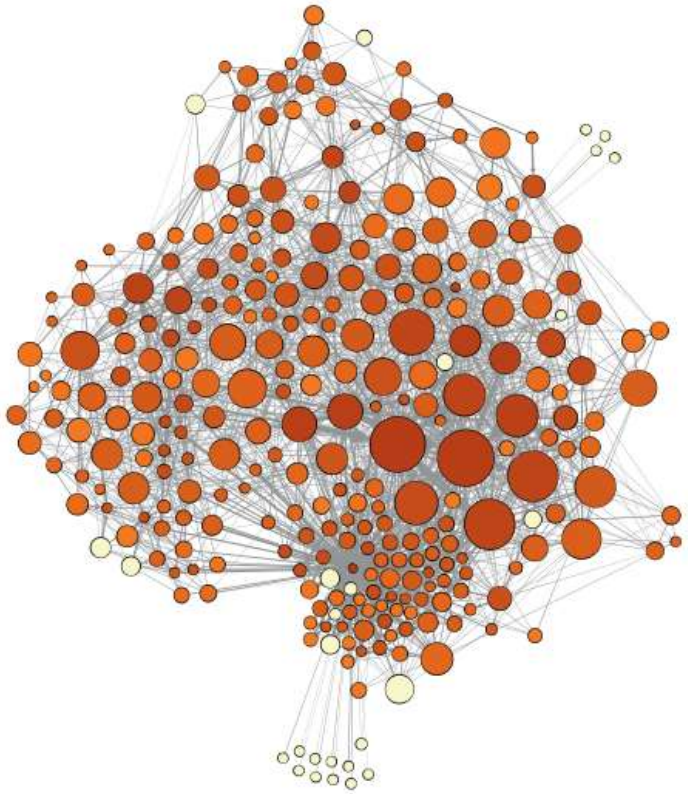


CHAPTER 0 - INTRODUCTION

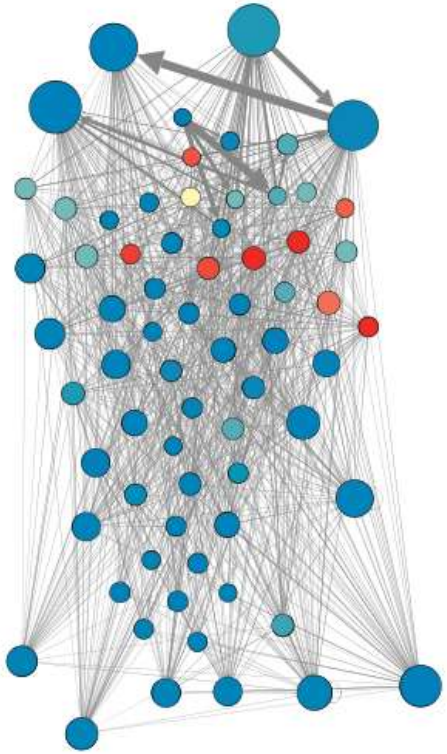
- Biological networks



Protein interaction network



C. elegans neural network



Food web in Florida

SUMMARY

- Using networks to understand complex phenomena is a type of modelling, with many (implicit or explicit) design choices
- Applying it as a black box is unlikely to tell you much
- At a large scale, it becomes increasingly useful
- In this course you will
  - Learn the basic tools of networks
  - Use them to understand some slice of Reddit

- Links:
  - joao.pinheiro-neto@uni-konstanz.de
  - https://github.com/joaopn/teaching_networks_2023

Resources

Mason Porter on Community Detection and Data Topology

Santa Fe Institute Subscribe