michele.tizzani@unito.it
lorenzo.dallamico@unito.it

di.unito.it

Lecture 01.ns01

Course description;
Introduction to networks;
Tools and Software

Course: **Complex Networks Analysis and Visualization**
Sub-Module: **NetSci**

Lecture 01.ns01 - Agenda

- Course Description
    - General Information
    - Instructors
    - Class schedule
    - Textbooks
- Introduction to Networks
- Tools and Software

Course description

General Information

- Computer Science's students:
  - INF0007 "Analisi e Visualizzazione delle Reti Complesse", 9 credits
  - MFN0954 "Reti Complesse", 6 credits
- "Fisica dei Sistemi Complessi" students:
  - FIS0127 "Analisi e Visualizzazione di Reti Complesse", 9 credits
- "Stochastics and Data Science" students:
  - MAT0049 "Complex networks", 6 credits

Instructors

- Prof. Giancarlo **Ruffo**, Dipartimento di Informatica,
giancarlo.ruffo@unito.it
twitter: giaruffo
web page: http://www.di.unito.it/~ruffo
groups' web page: http://arcs.di.unito.it
webex room: https://unito.webex.com/meet/giancarlo.ruffo

- Dr. Michele **Tizzani**
michele.tizzani@unito.it
web page: Michele Tizzani (isi.it)

- Dr. Lorenzo **Dall'Amico**
lorenzo.dallamico@unito.it
web page: https://lorenzodallamico@github.io

Course structure

- The course is divided in two sub-modules:
  - NetSci
  - DataViz

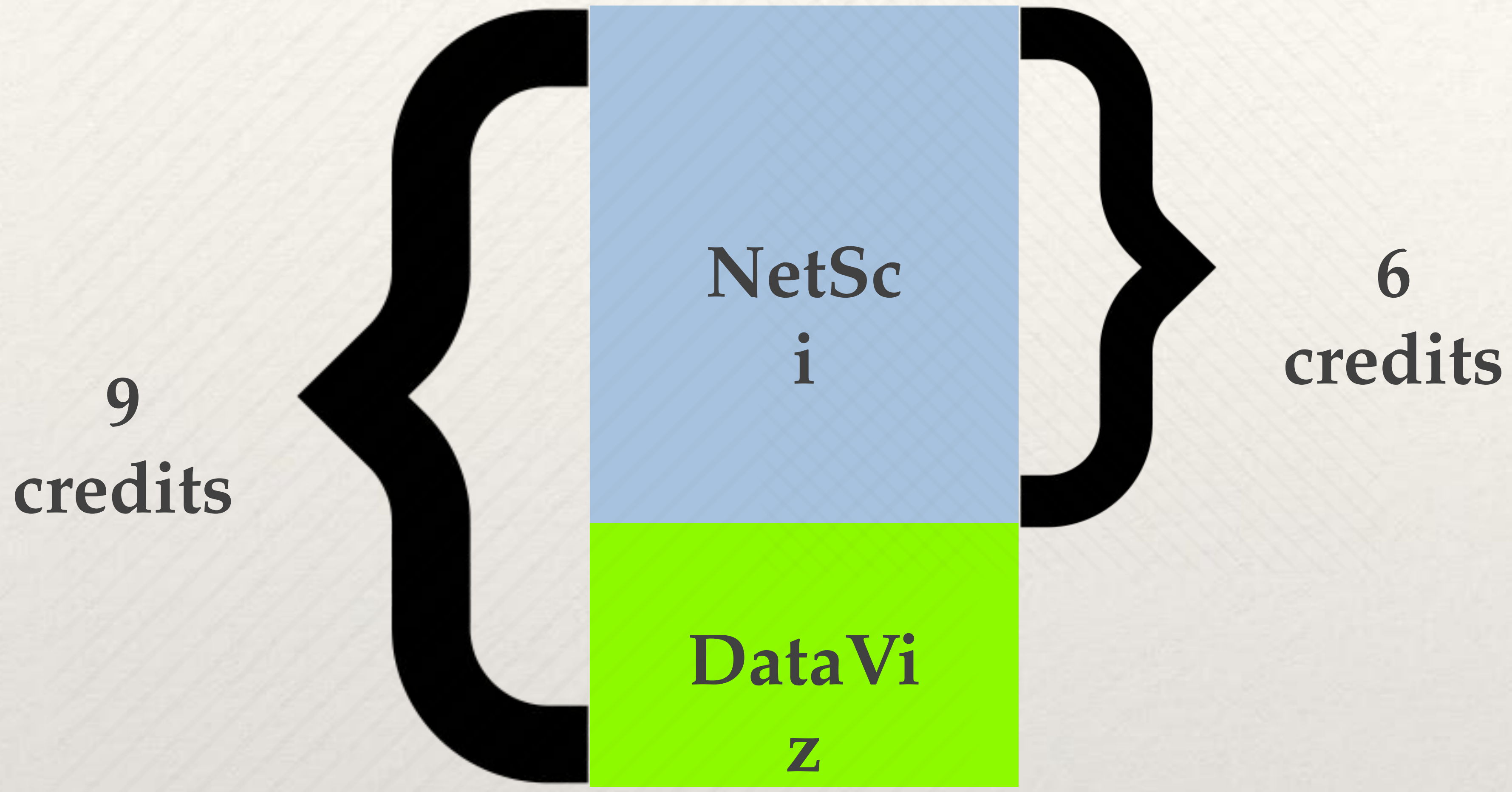
NetSci description and objectives

- Networks are pervasive in our lives: networks of friends, communication, computers, the Web, and transportation are examples we experience, while our brain cells and the proteins in our body form networks that determine our survival and intelligence.
- We introduce the fundamental concepts, principles and methods in the interdisciplinary field of network science, with a particular focus on analysis techniques, modeling, and applications for the World Wide Web and online social media.
- Topics covered include graphic structures of networks, mathematical models of networks, common networks topologies, structure of large scale graphs, community structures, epidemic spreading, PageRank and other centrality measures, dynamic processes in networks, graphs visualization.
- Additionally, students will learn how to apply the basic principles of network science to perform CNA (Complex Network Analysis) tasks on real data, with Python and many different packages/libraries such as networkx, igraph, and so on, as well as advanced graph visualization tools as Gephi.

DataViz description and objectives

- We are exposed continuously to plots, graphics, and visualization of information of any kind, including scientific data visualization. Viz are extremely powerful ("a picture is worth a thousand words"); however, not all the visualizations are designed or implemented correctly, and sometimes they misrepresent underlying data.
- Students will learn basic visualization design and evaluation principles, and learn how to acquire, parse, and analyze large datasets. Data Visualization is just one of the fundamental steps of a more general Data Science process
- Students will also learn techniques for visualizing multivariate, temporal, text-based, geospatial, hierarchical, and (of course) network/graph-based data.
- Additionally, students will utilize Gephi, D3, python and matplotlib, and many other tools to prototype many of these techniques on existing datasets.

Structure and schedule

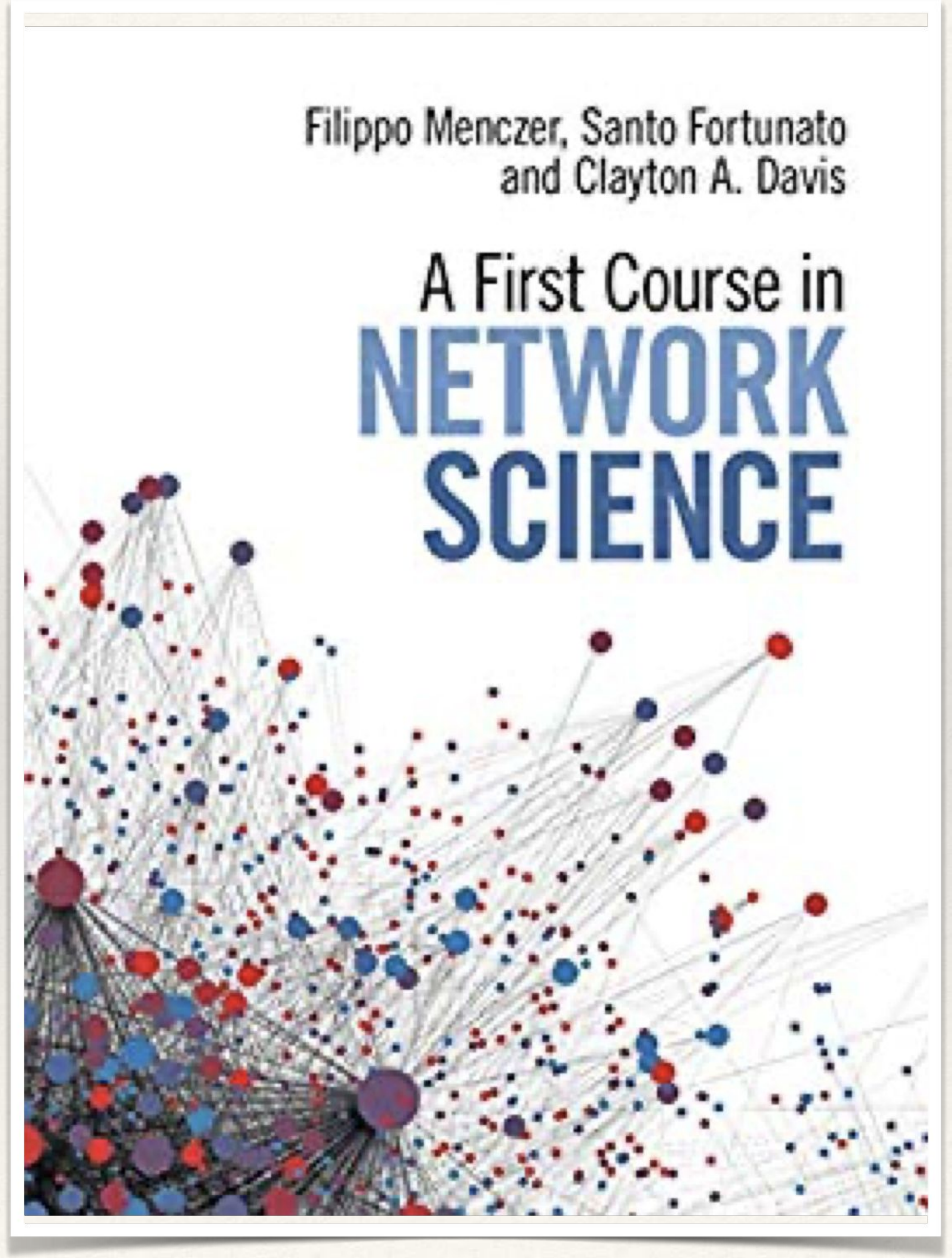


	Monday	Tuesday	Wednesday
Module	**NetSci**	**DataViz**	**NetSci**
When	11am-1pm	4pm-6pm	11am-1pm
Where	Aula E	Aula F	Aula E
Who	Dall'Amico/Tizzani	Ruffo	Dall'Amico/Tizzani
How	English	Italiano	English

Full schedule (with timely updates): on moodle

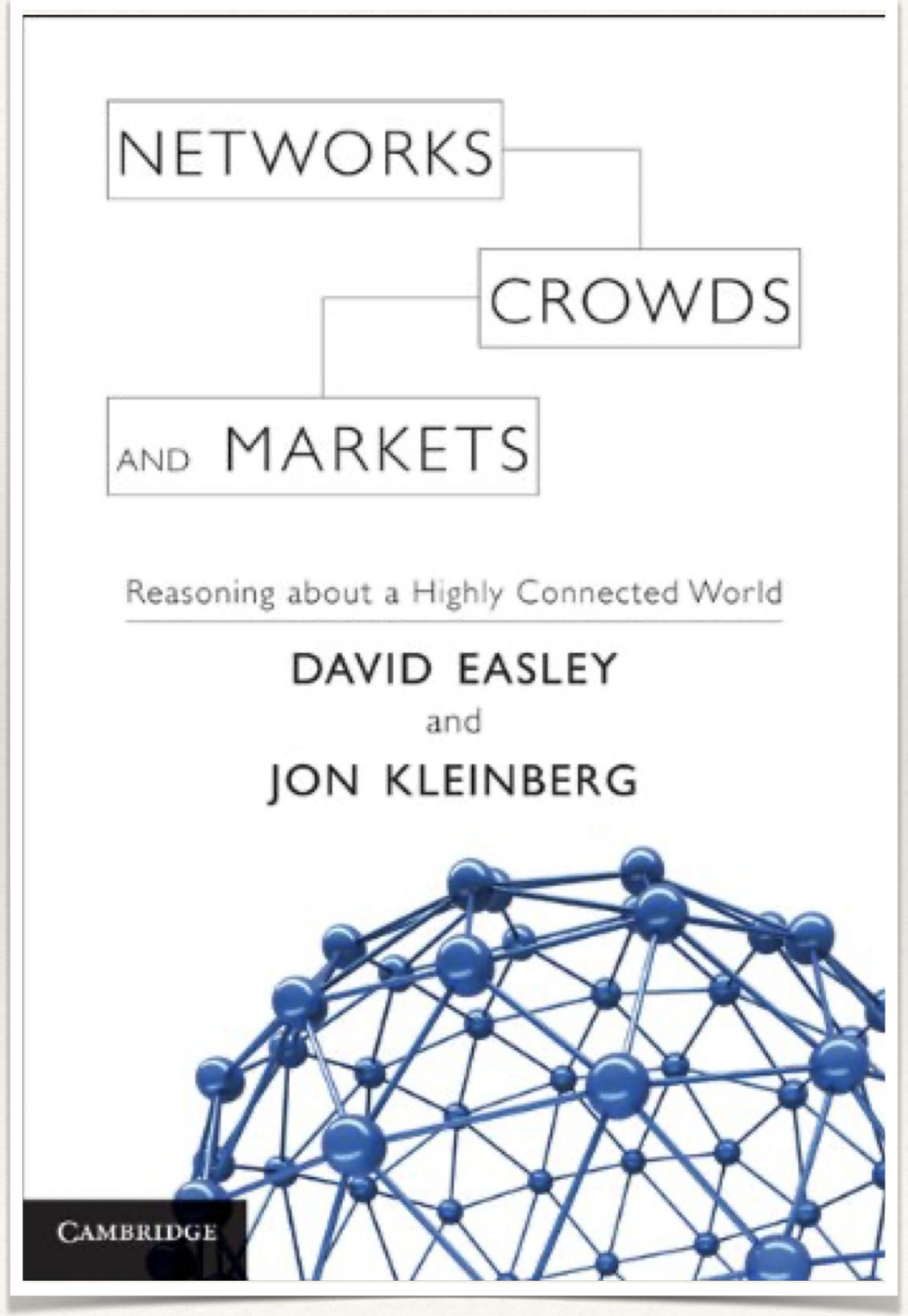
Textbooks

- Filippo **Menczer**, Santo **Fortunato**, Clayton A. **Davis**, *A First Course in Network Science*, Cambridge University Press, 2020
- "A First course": very **basic**, but wide and updated overview
- Few mathematical formalisms, some **illustrative code** (in python)
- Some copies are available in the computer science library
- We will cover **every chapters**



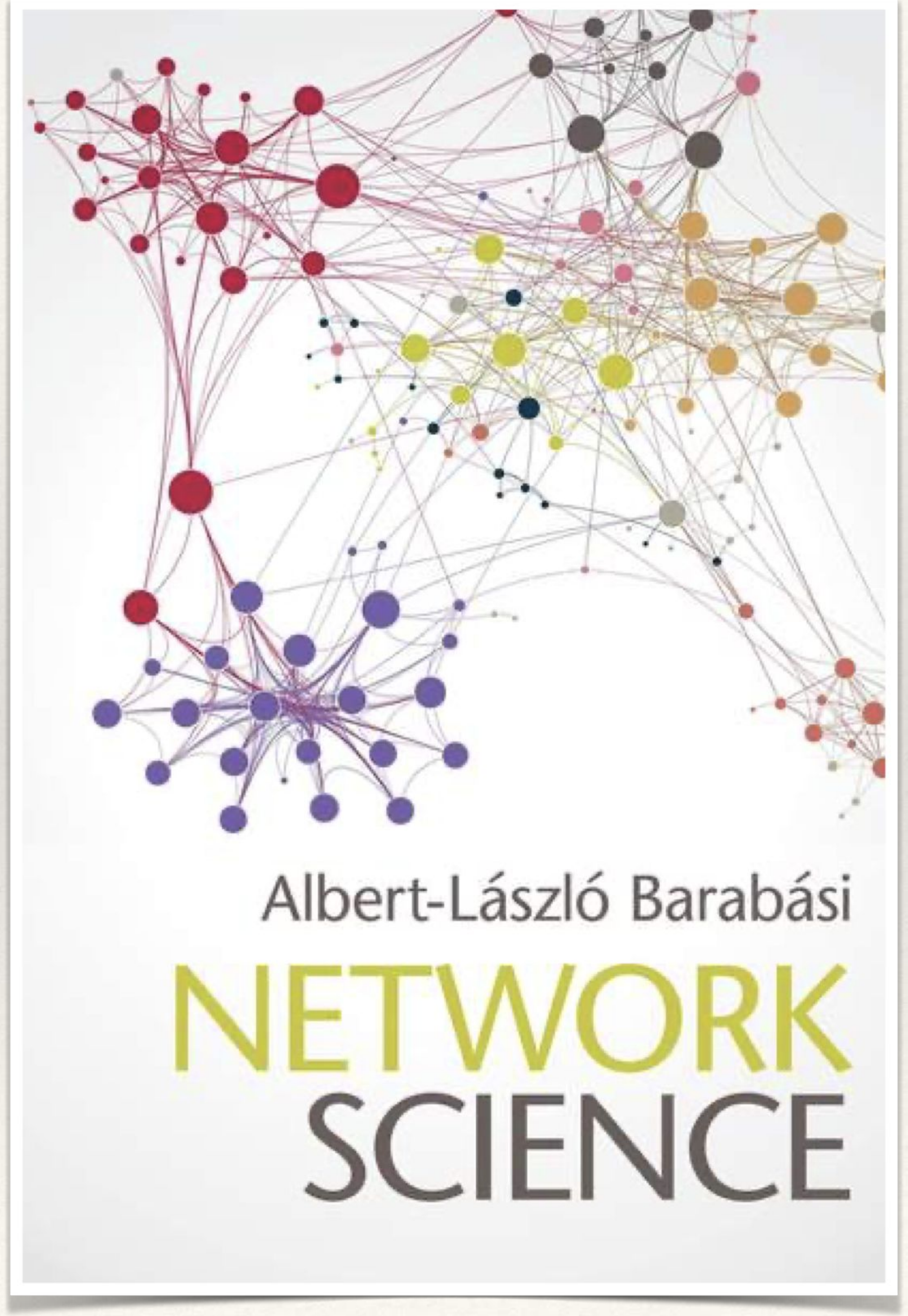
Textbooks

- David **Easley** and Jon **Kleinberg**, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*, Cambridge University Press, 2010
- Very **multidisciplinary**!
- Richer overview (much beyond 'networks'); updated up to 2010
- Few mathematical formalisms, **a lot of examples in many different applicative domains**, no illustrative code
- Available in our libraries but also on-line (pre-print free!): https://www.cs.cornell.edu/home/kleinber/networks-book/
- We will cover **many chapters**



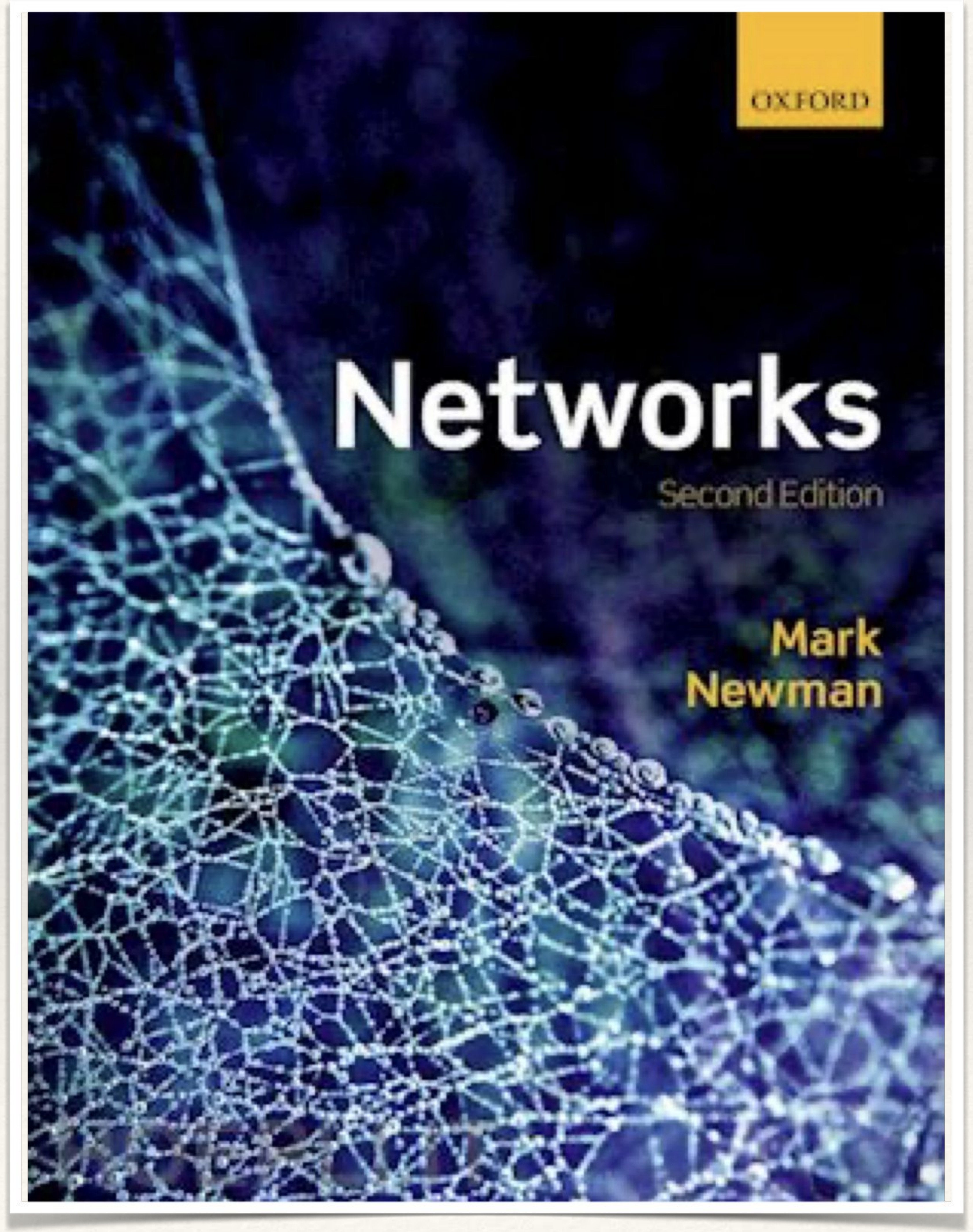
Textbooks - optional

- Albert-László **Barabási**, *Network Science*, Cambridge University Press,
- Read it if interested more to **models**, than to data science
- Much (even if simplified) mathematical formalisms, many plots, some applicative examples, no illustrative code
- Available in our libraries but also on-line (free!): http://networksciencebook.com
- This is mainly **optional,** even if we may adopt some parts



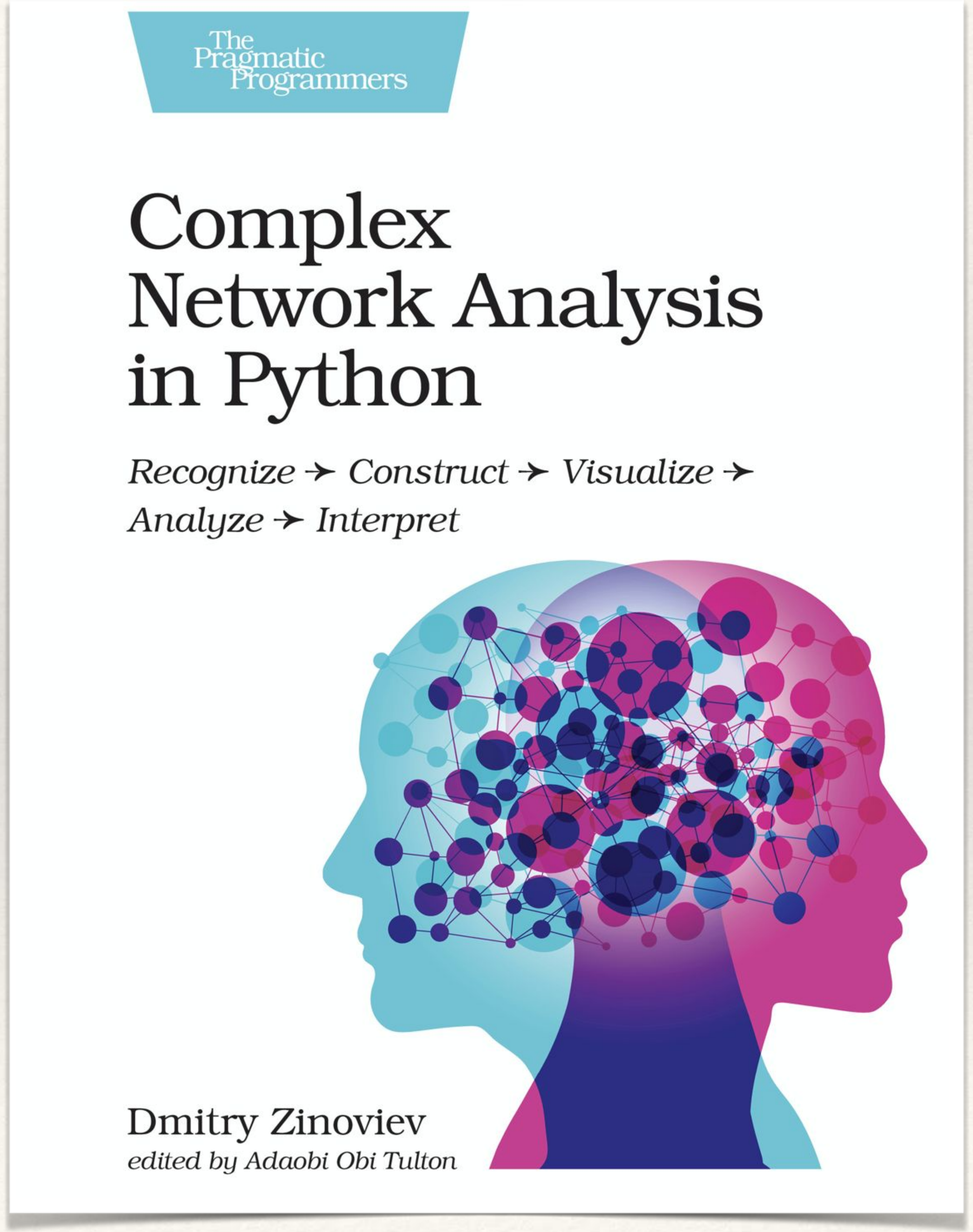
Textbooks- optional

- Mark **Newman**, *Networks*, Oxford University Press
- The **Bible** for network scientists: it is a must if you undergo with graduate studies or with a master's degree thesis on such topics
- Read it if interested to: **models**, deep understanding of **measures**, and **algorithms**
- Very elegant mathematical formalisms, many examples, no illustrative code
- Available in our libraries
- This is mainly **optional,** even if we may adopt some parts



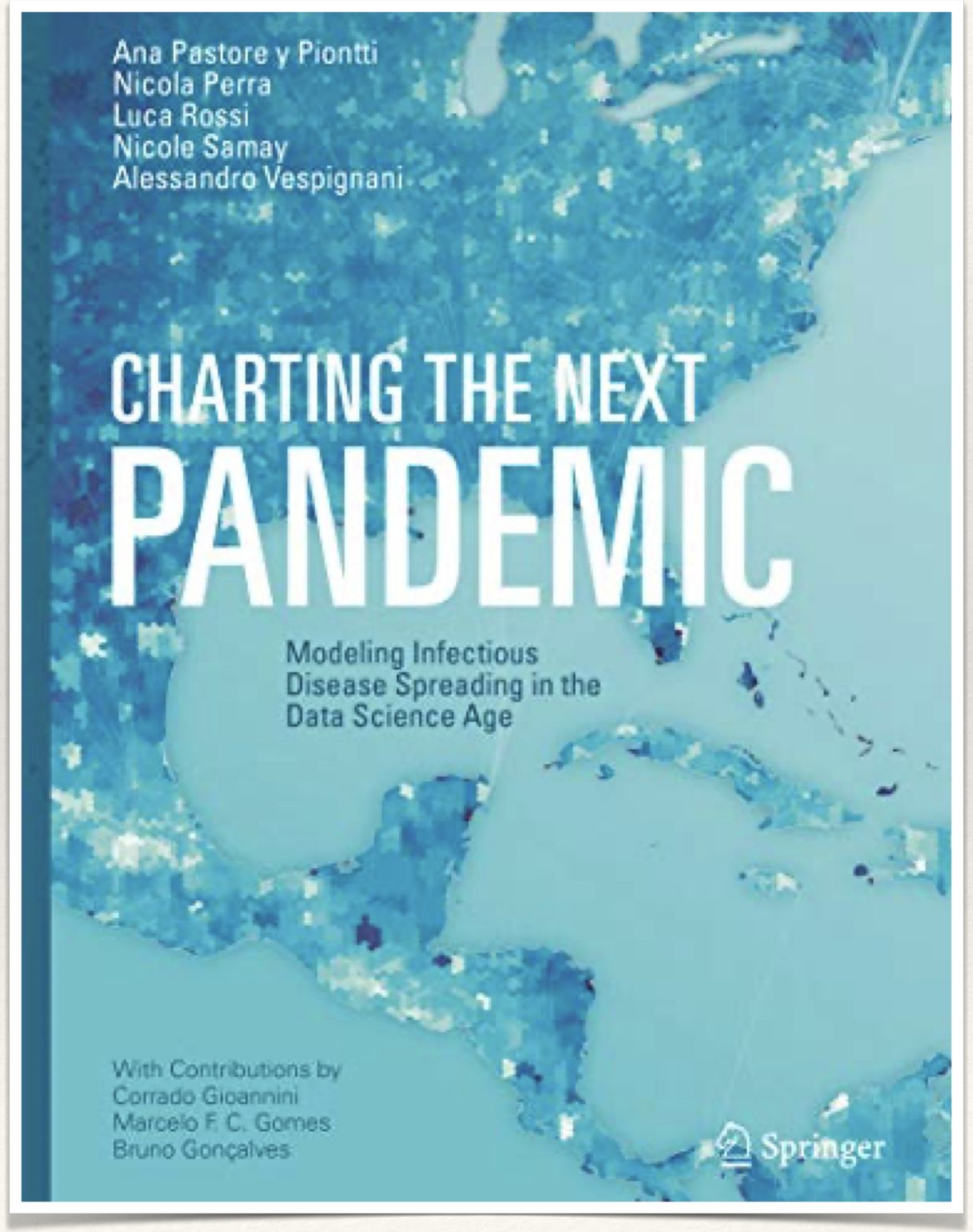
Textbooks- optional

- Dmitry **Zinoviev**, *Complex Network Analysis in Python, Recognize → Construct → Visualize → Analyze → Interpret*, The Pragmatic Bookshelf
- Read it if interested to: **practical data science**, **network analysis**;
- Very **practical:** no much space for ideas, concepts, and limitations behind network measures
- Many illustrative examples, A LOT of illustrative code (python) with many different datasets
- Available in our libraries
- This is mainly **optional,** even if we extract some example from here



Textbooks - optional

- **Pastore y Piontti**, A., **Perra**, N., **Rossi**, L., **Samay**, N., **Vespignani**, A., Charting the Next Pandemic: Modeling Infectious Disease Spreading in the Data Science Age, Springer, 2019
- Read it if interested to: **dynamical processes in networks, epidemics**, **visualization**, **data science**;
- Focused on how to model infectious diseases in networks
- Many illustrative examples, a lot of plots, a great link from ideas to reality
- This is mainly **optional,** BUT we will use part of it for our dataviz module and when introducing epidemics

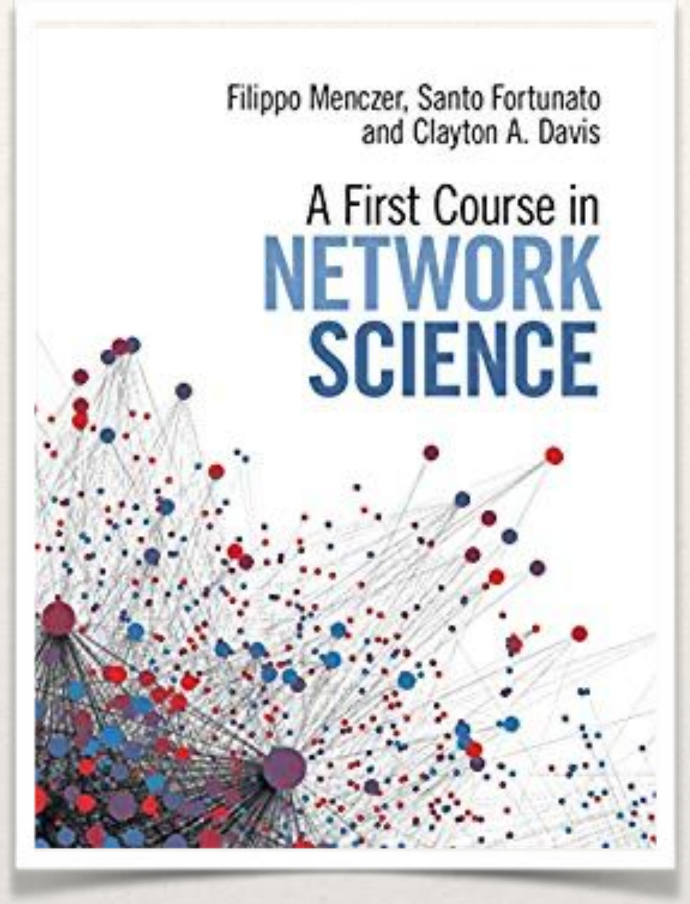


Exam

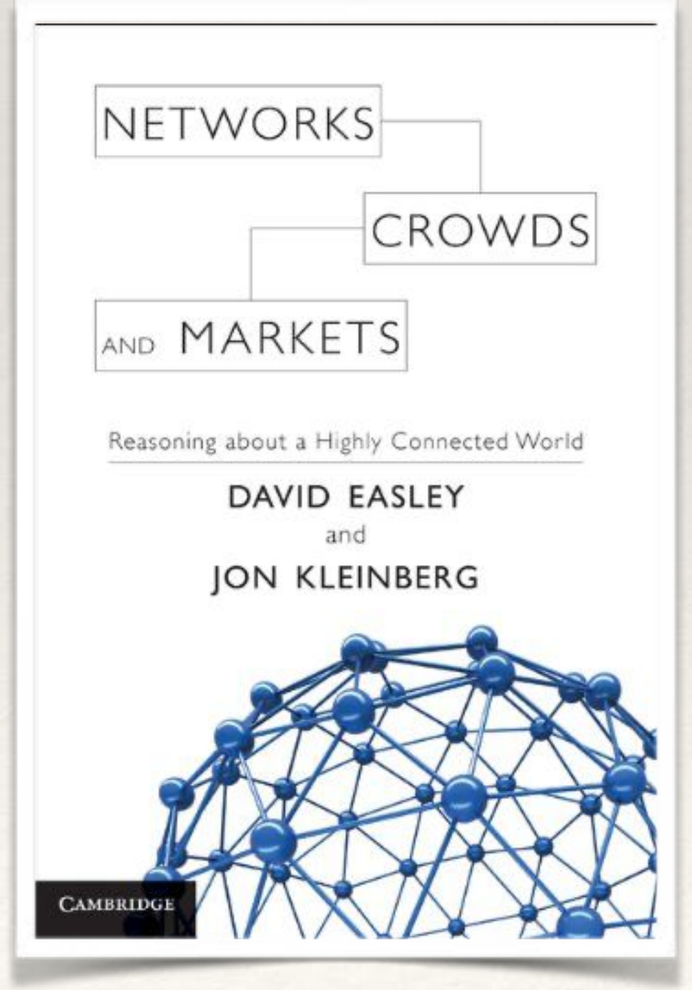
- NetSci:
  - Report: network analysis with real data and interpretation of the results
    Written exam
- DataViz:
  - project: design and implementation of a data visualization platform to explore some real dataset(s)

Introduction to Networks

References



[ns1] Chapter 0



[ns2] Chapter 1 —->

https://www.cs.cornell.edu/home/kleinber/networks-book/networks-book-ch01.pdf

Complex Systems

"for the discovery of the interplay of disorder and fluctuations in physical systems from atomic to planetary scales"

Motivation for Nobel Prize to G.Parisi

Complex Systems

- Complex != Complicated
- composed by many **interacting elements**
- they give rise to emergent **collective phenomena**
- **emergence**: not directly related to individual phenomena
- linearity vs **non linearity**
- **heterogeneous** vs homogeneous

M. E. J. Newman - Complexity survey (2011): [cssurvey.dvi (umich.edu)]

Physical Systems

condensed matter systems such as crystals, magnets, glasses, and superconductors;

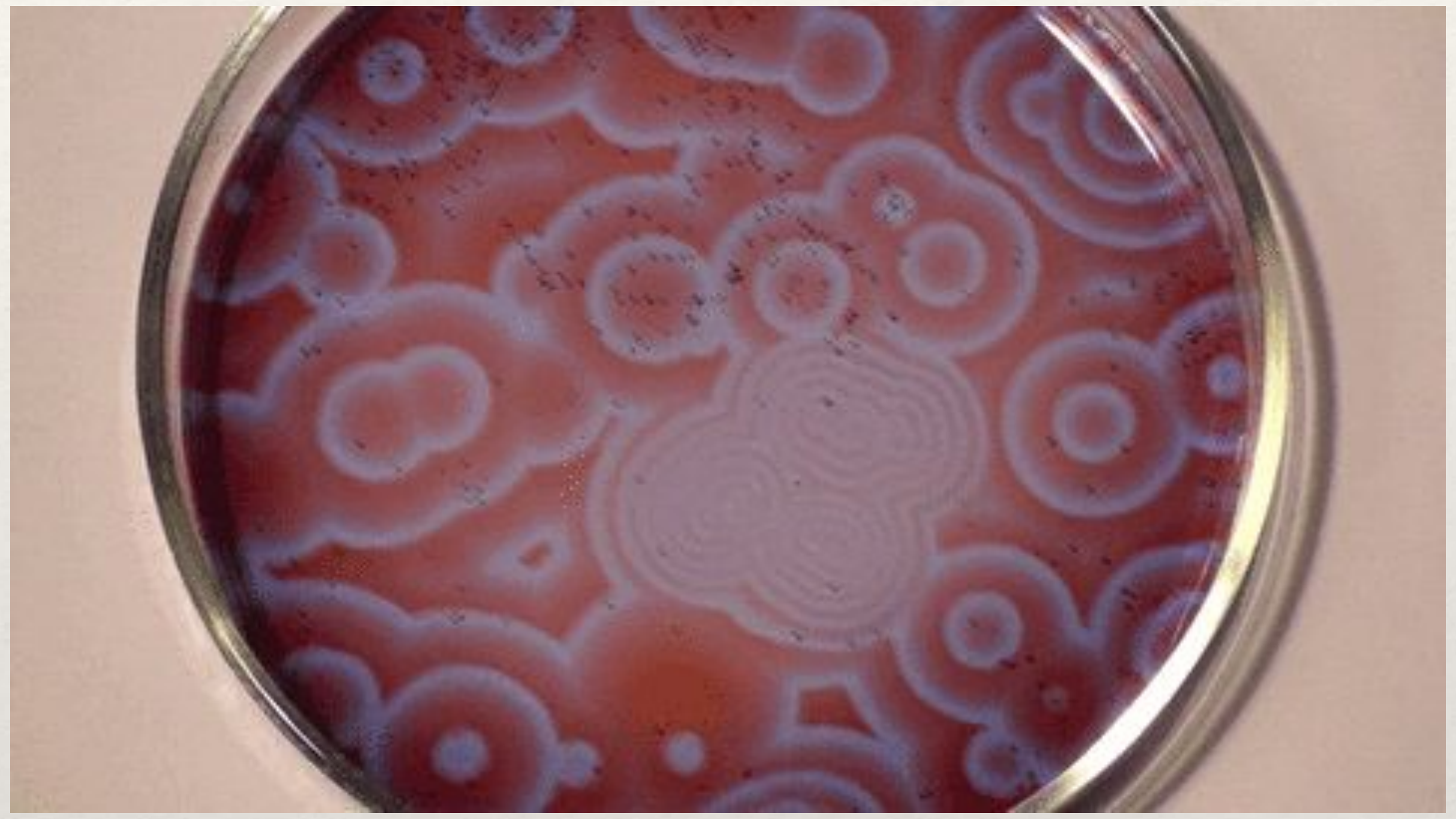
hydrodynamical systems;

spatiotemporal pattern formation in systems like chemical oscillators and excitable media;

molecular self-assembly, including tiling models, biomolecules, and nanotechnological examples;

biophysical systems such as protein folding and the physical properties of macromolecules;

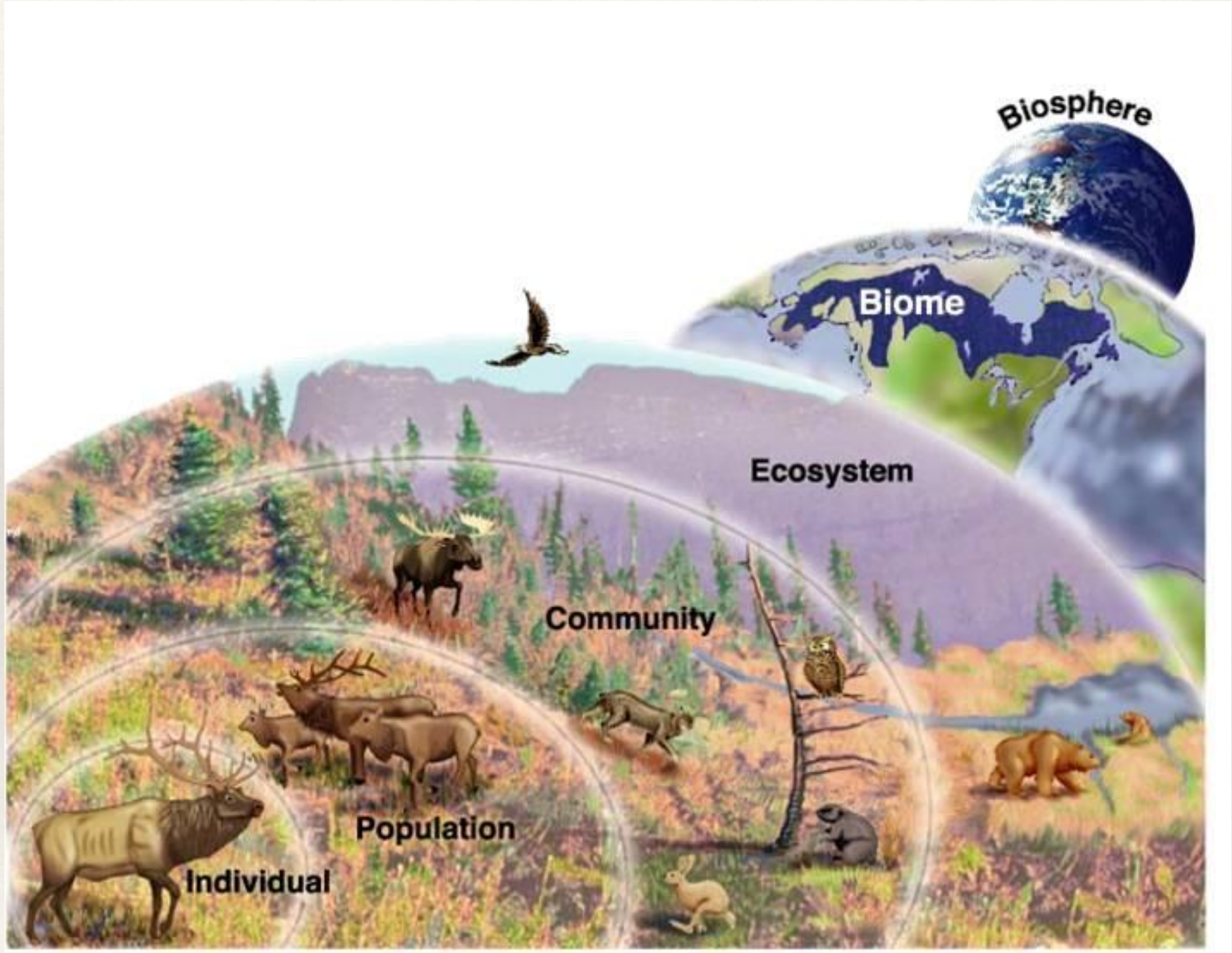
computation performing systems, including analog and quantum computers.



Example of emergent phenomenon, non-linear chemical oscillator. (Belousov-Zhabotinsky)

Ecosystems and biological evolution

The biosphere, both in its present state and over evolutionary history.



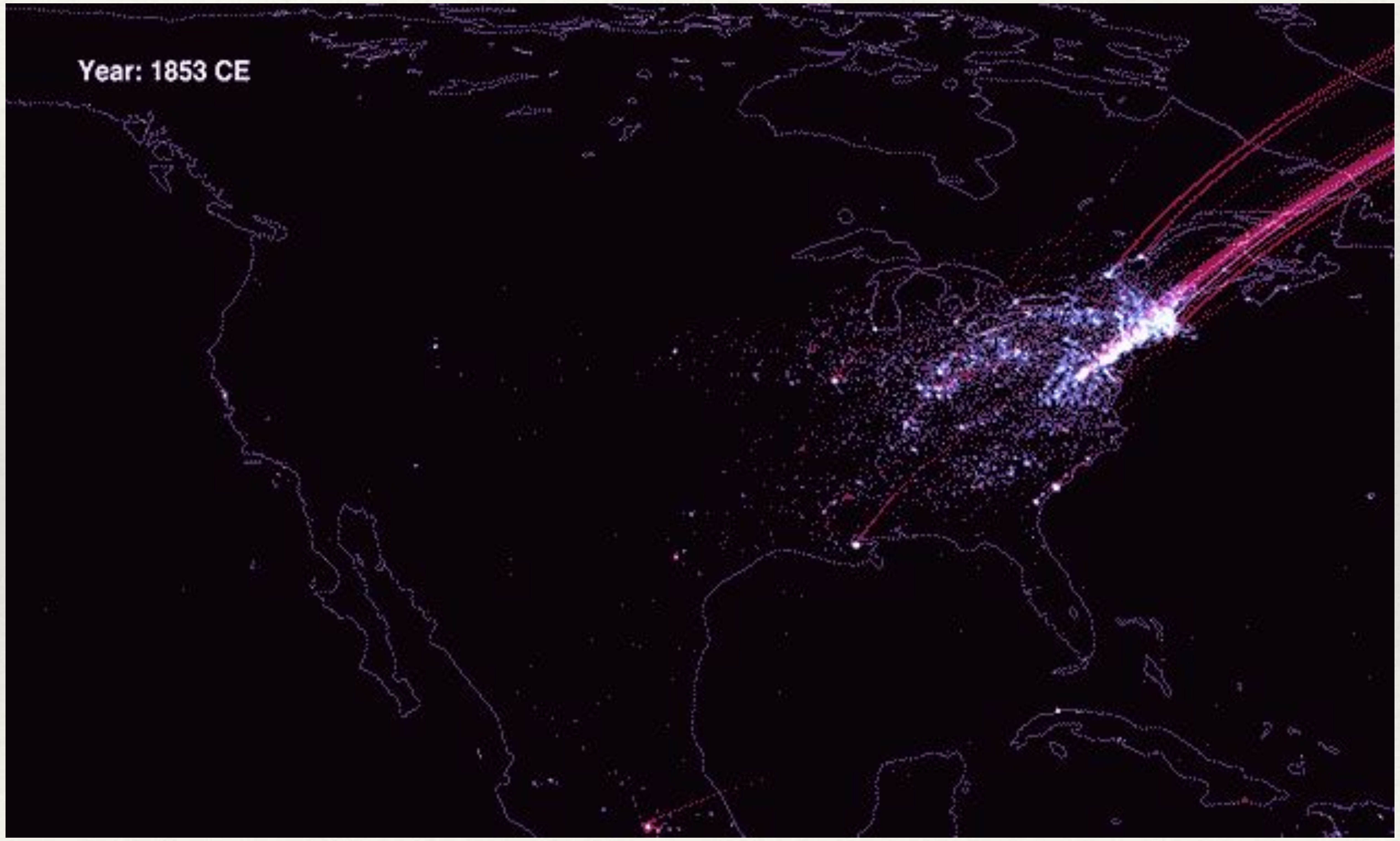
Human societies

Three aspects of human societies are particularly interesting as complex systems:

(1) urban planning and the physical structure of society;

(2) the social structure of society and social networks;

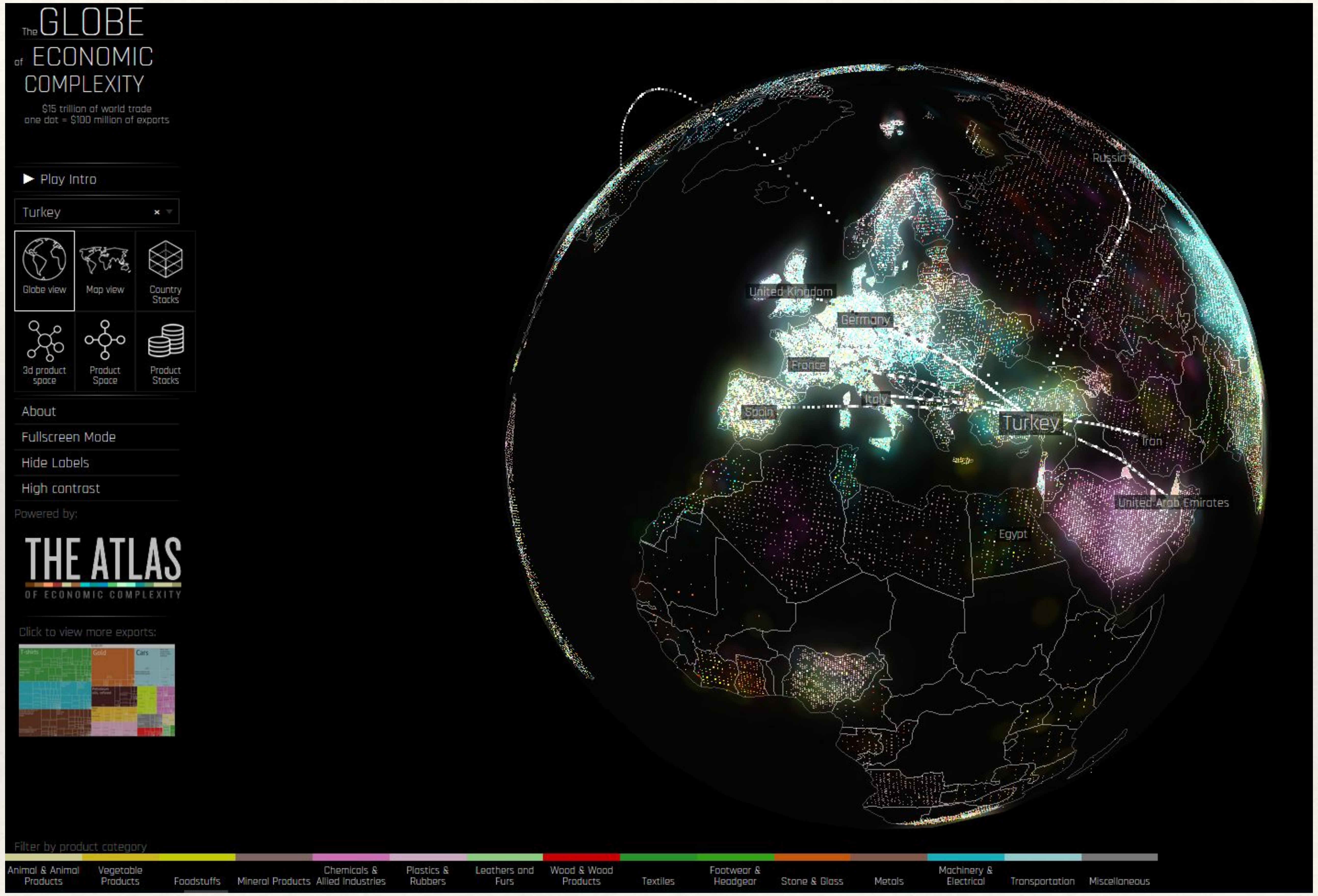
(3) differences between societies as revealed by sociological experiments



Mobility patterns in US (by Maximilian Schich)

Economics and markets

Manufacturers, traders, and consumers interacting to produce the emergent phenomenon we call the economy.



[The Atlas of Economic Complexity (harvard.edu)]

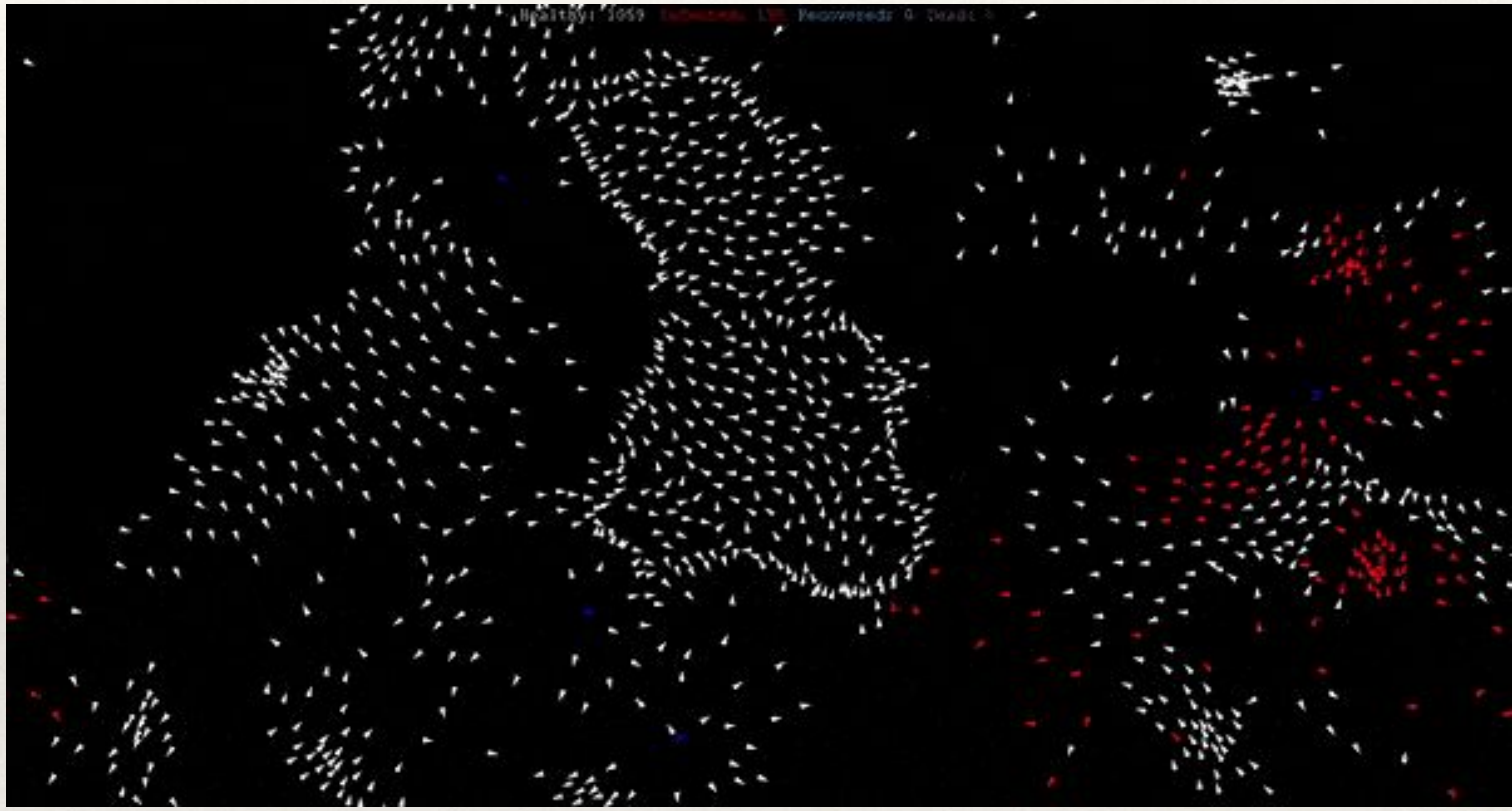
Pattern formation and collective motion

Interactions of agents in a complex system can produce spatial patterns of many kinds and systems that do this are seen in many branches of science, including physics and biology.

Flocking or schooling in birds or fish is a cooperative phenomenon in which the animals in fly or swim in roughly the same direction, possibly turning as a unit. It's believed that animals achieve this by simple self-enforced rules that involve copying the actions of their nearby neighbors while at the same time keeping a safe distance.

From local to a global - Epidemics

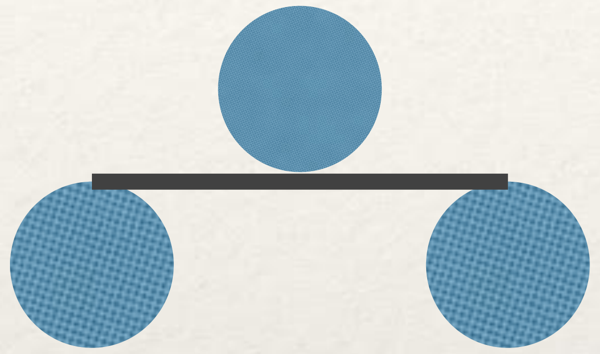
Epidemics is a collective phenomena that emerges from the transmission of a pathogen to a population through interaction among individuals.



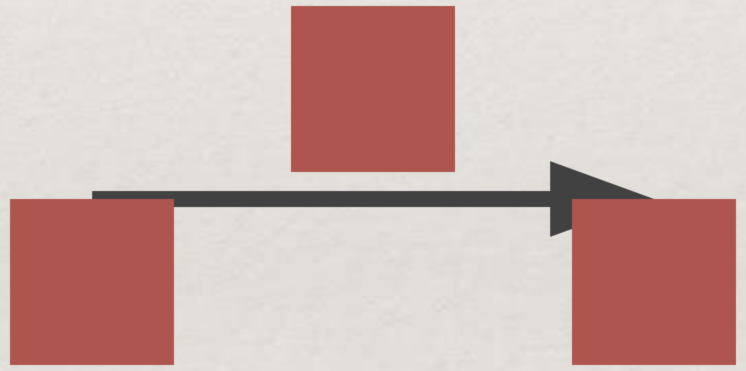
epidemic-simulations · GitHub Topics · GitHub

Networks are "everywhere"

- Social Networks
    - actors (individuals, also agents)
    - social ties



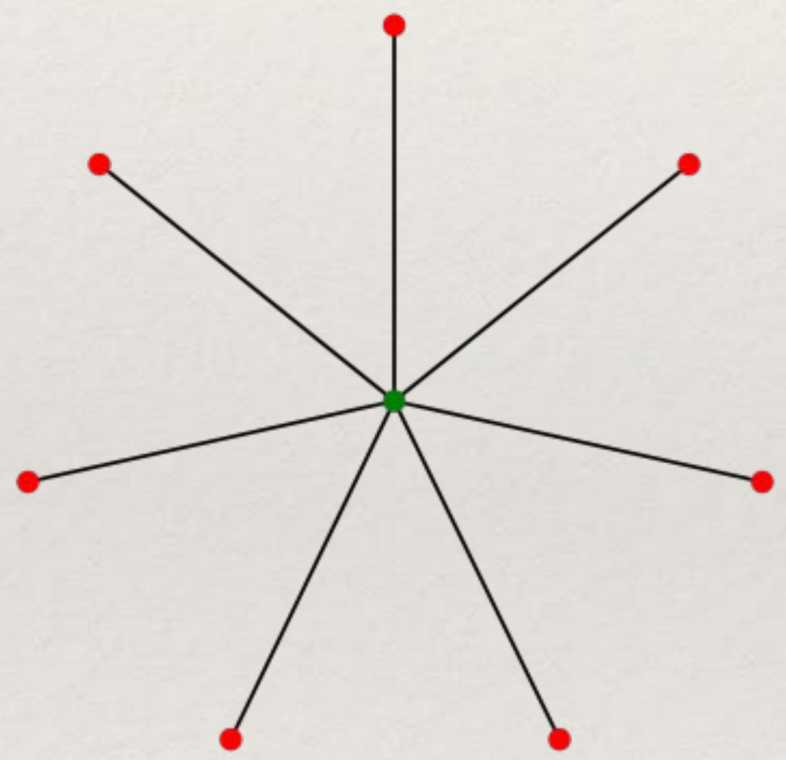
- Information systems
    - book, web page
    - citation, link



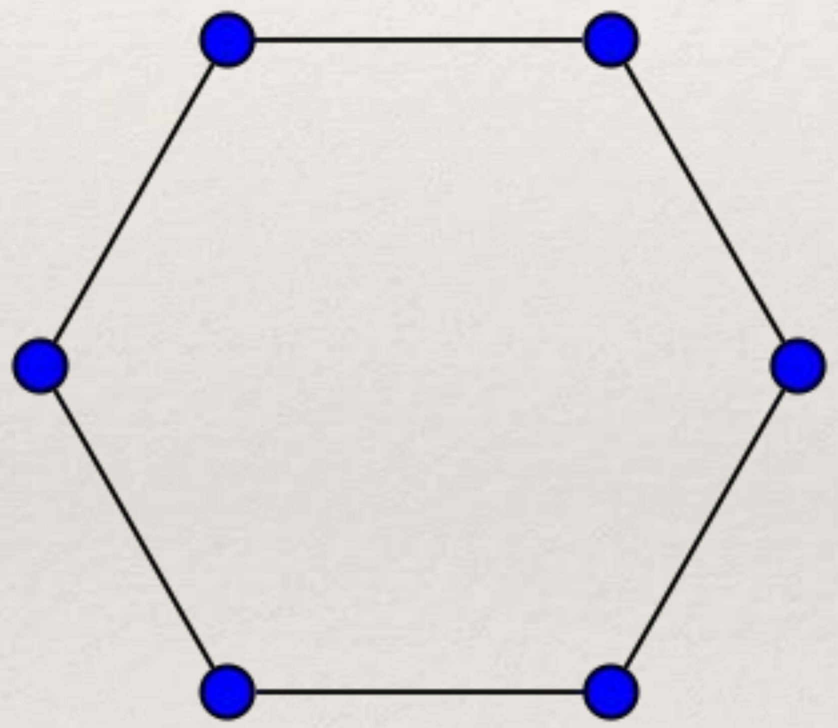
- Protein interaction networks, food web
- Technical and economic systems
- Can structural crisis be predicted?

Networks structural aspects

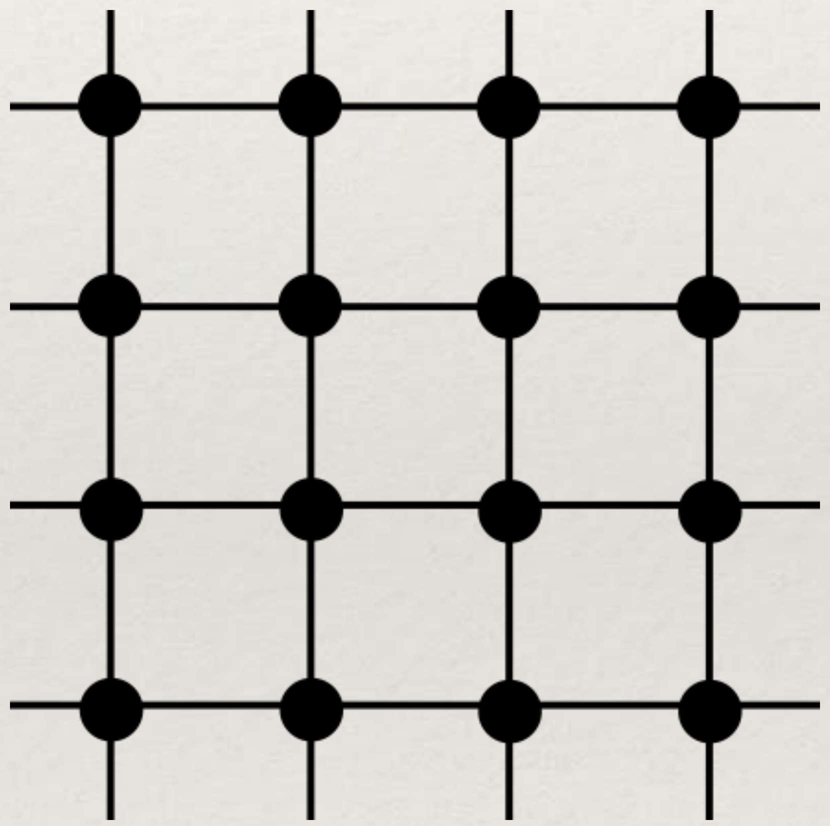
- "trivial" **representation** of a complex system
- **Simple** networks: few characteristics describe the network



Star



Ring

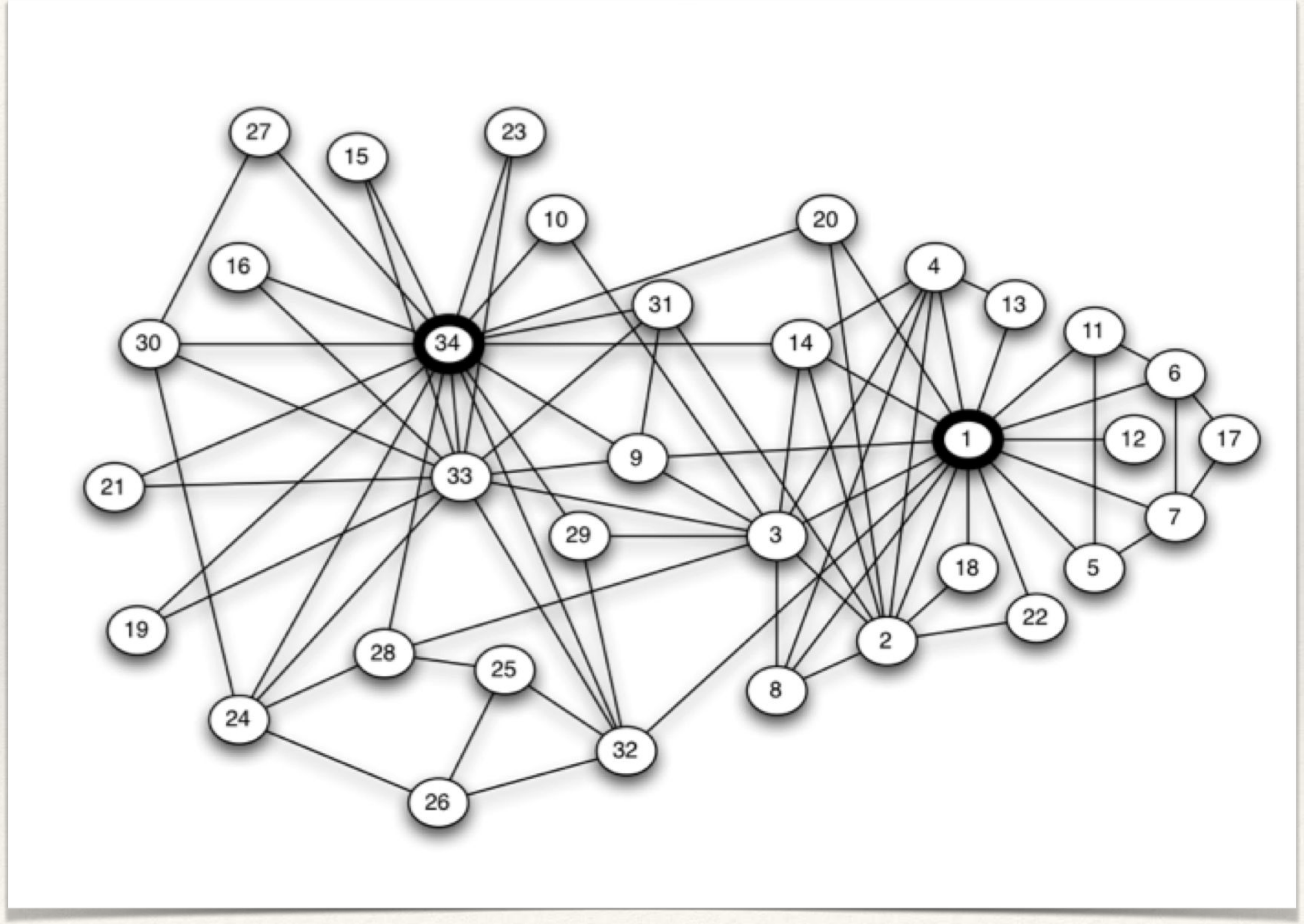


Grid

- We need a **language** and a **framework** to describe complex networks

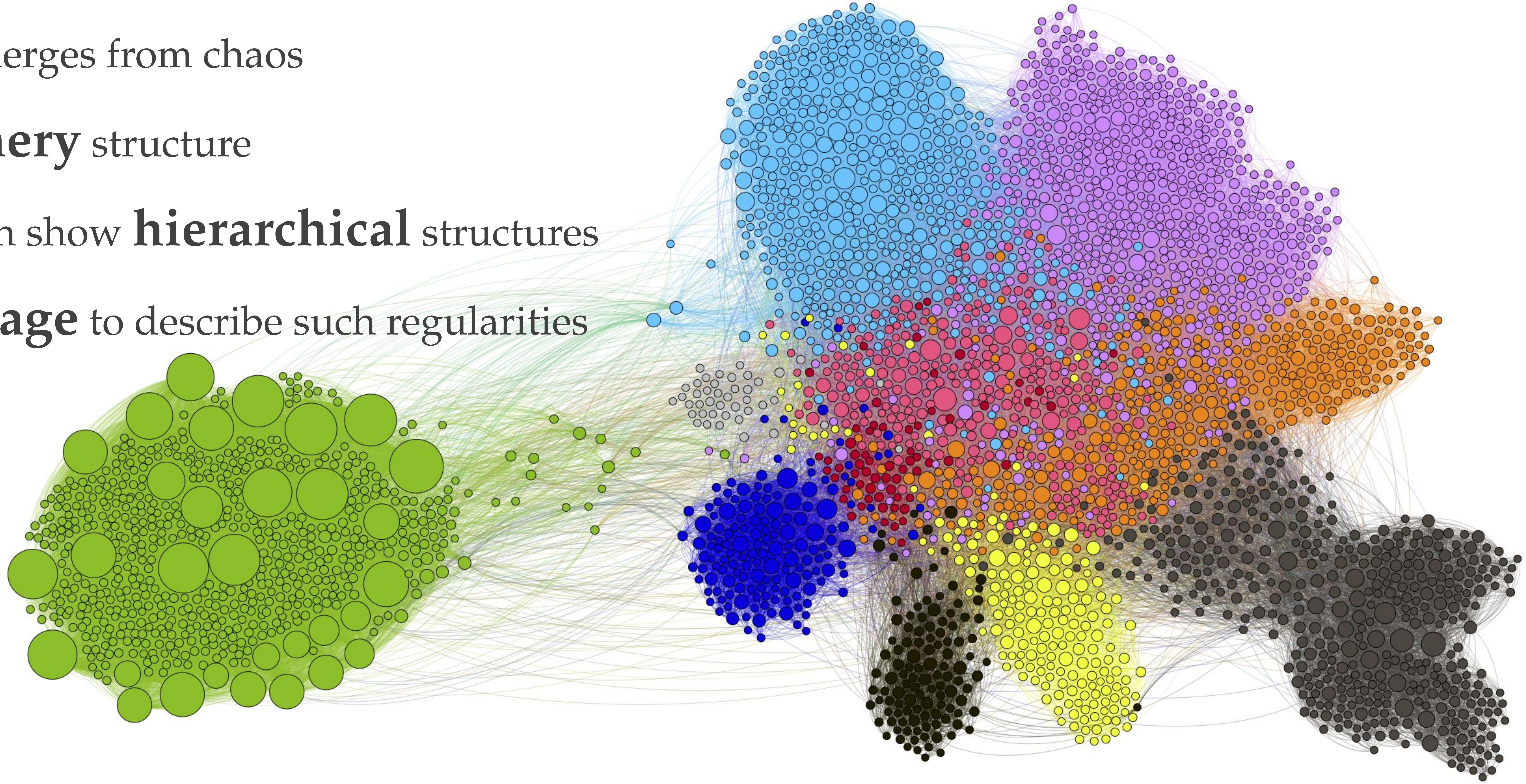
Example: Zachary's Karate Club

- apparently no simple description is possible
- however, there are some **regularities**
- some individuals are more "**central**" than others
- does it resemble some kind of **randomness**?



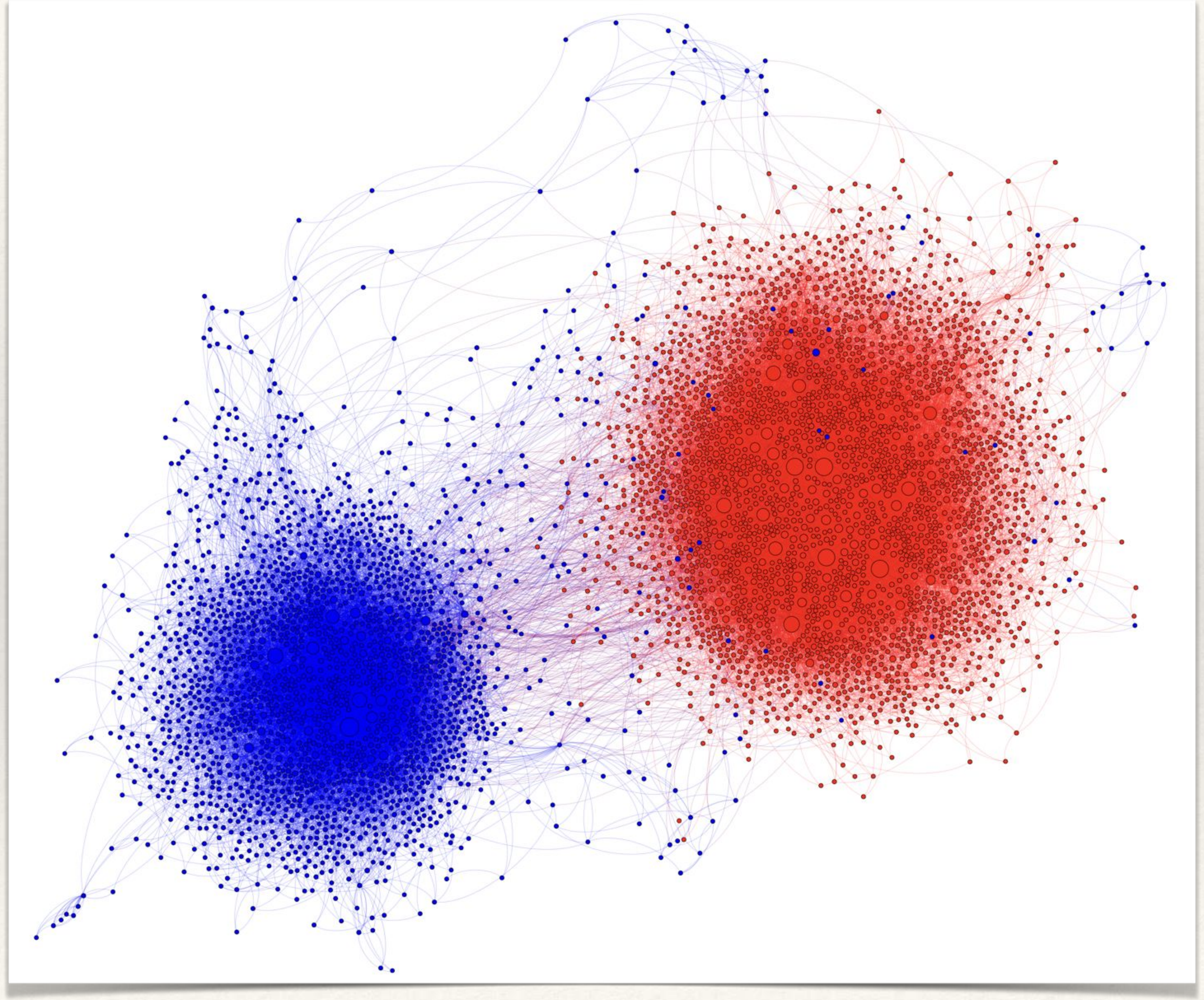
Example: the Internet

- Some "**order**" emerges from chaos
- **core vs periphery** structure
- other networks can show **hierarchical** structures
- we need a **language** to describe such regularities



example: retweet networks

- communities or clusters (**partition**)
- "**echo chambers**"
- **homophily**: tendency of individuals to link with similar ones
- warning: no trivial linear relationships but **interplay**



Pros and cons of visualization

- Visualizing a network suggests some **inherent** complexity
- this says a lot about the **structure**, but almost nothing about the **reasons**
- It is difficult to summarize succinctly the whole network
  - core vs periphery structure
  - hierarchical structure
  - tightly-linked regions
  - …
- we need a **language**!

Behavior and Dynamics

- Connectedness
  - structure
  - behavior
- each **individual's actions have consequences** for the outcomes of everyone in the system
- we need a **framework for reasoning** about behavior and interactions in networks
- we need to take into account a **rational behavior** that sometimes lead to **strategies**
- actions are not evaluated at individual level (in isolation), but at a **network level**

From data to networks

- We have **massive network datasets**
- data-driven models allow you to make predictions BUT sometimes these predictions cannot be generalized: networks can be **different**

individual vs population level

- when a large group of people is tightly interconnected, they often respond in complex ways that can be observable only at the **population level**:
  - rich get richer
  - winner takes it all
  - viral ideas
  - …
- At **individual level**: sometimes an information goes viral or one person becomes very popular, some other times this does not happen!
  - **Why**? no prediction at this level yet

Main topics in this course

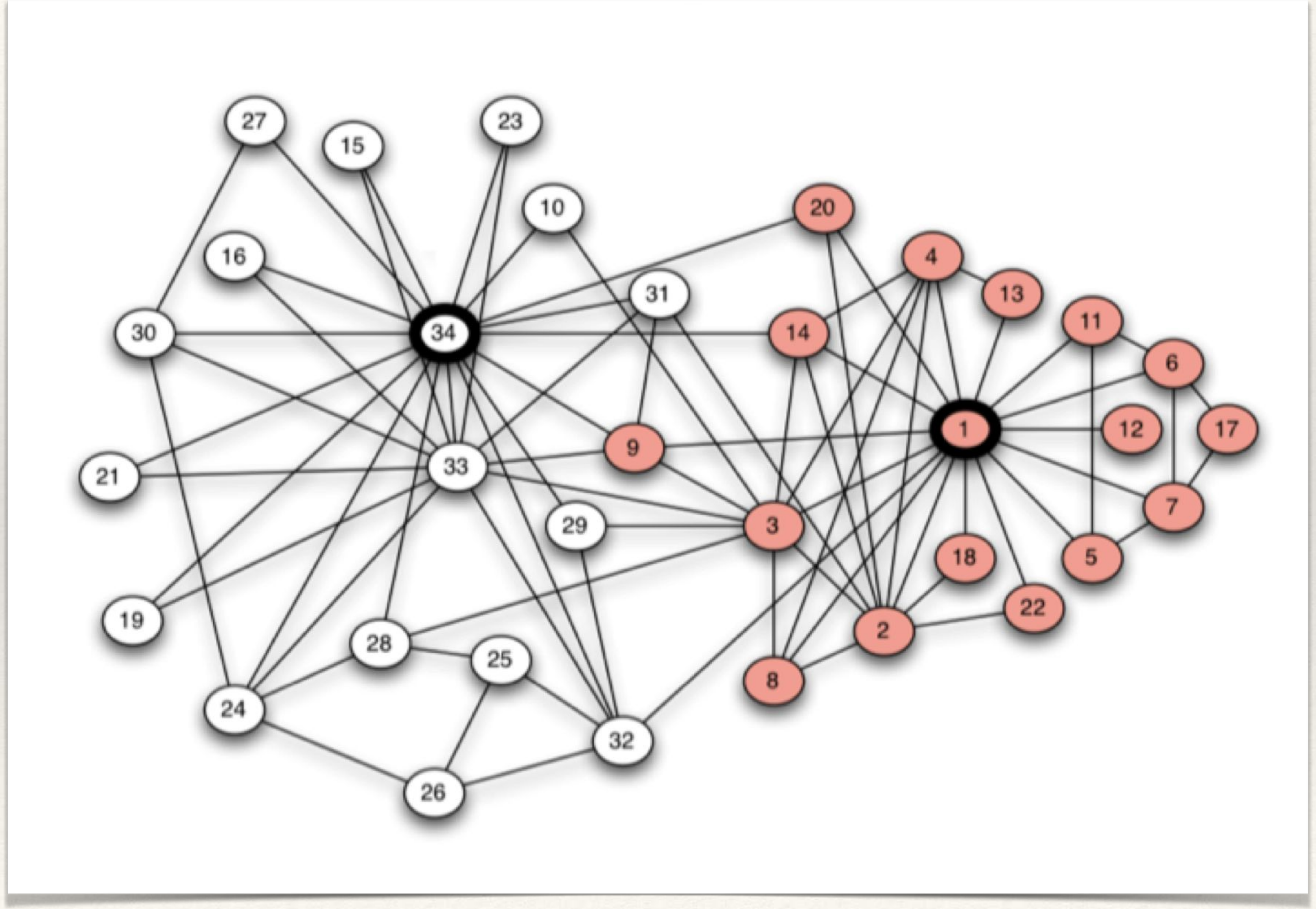
- Graph Theory (from next lecture) and Complex Network Analysis
- Game Theory (only basic concepts)
- Information Networks
- Network Dynamics

Graph Theory

- paths/distances
- clustering coefficient
- degree and degree distribution
- centralities
- groups/clusters/partitions/communities
- structural balance
- homophily and spatial segregation

Individuals vs groups

- we can find clues to the latent schism that eventually split the group into two separate clubs
- we can have different roles



Game Theory

- framework where our decisions depend on others' decisions
- **strategy** -> **pay-off**
- ex. in transportation networks, choice of a route can bring to **congestion** (Braess' Paradox)
- **Equilibrium**: a state that is self-reinforcing.
  - no individual have an incentive to unilaterally change their strategy, even if those individuals know how others will behave

Information Networks

- Ex.: the Web

React to maximize rank

Search Engines

Web pages authors

Rank

- Interplay: if you define a measure (e.g., PageRank), the application of such measure may change the system

Network dynamics: population effects

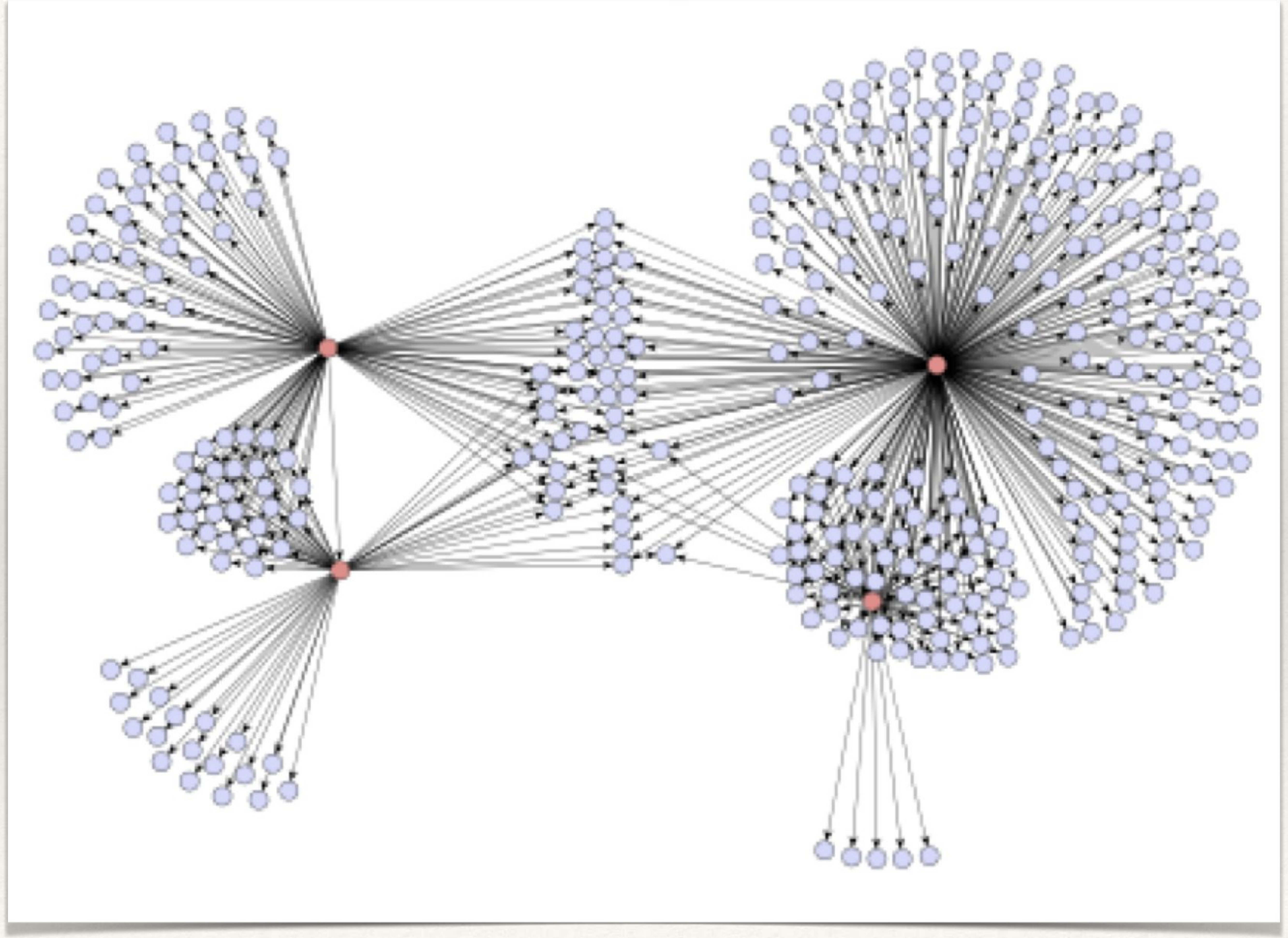
- In large populations we have the emergence of new ideas, beliefs, opions, innovations, technologies, products, and social conventions
- When and how they establish as **social practices**?
- Individuals can **influence** or being influenced by others
- At the surface: humans have the tendency to **conform**
- **Why**?

Networks Dynamics: Structural Effects

- Populations vs local behaviors
- Individuals may have incentive to adopt the behavior of their neighbors: **cascading behavior**
  - social contagion
  - epidemics
- The structure of the network has a role

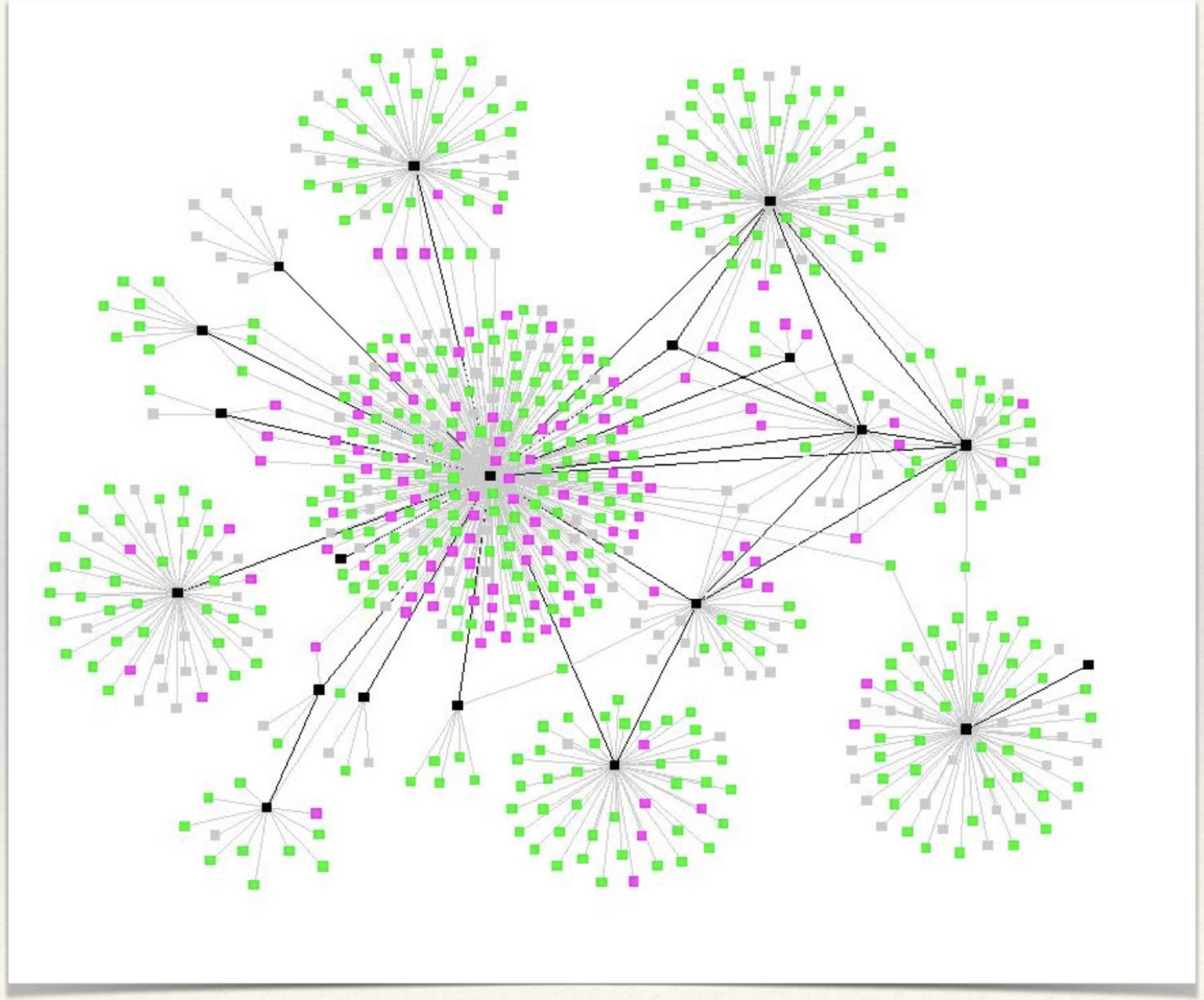
Social contagion

- Influence: the adoption of a new product or innovation can cascade through the network structure
- ex.: e-mail recommendations for a Japanese graphic novel
- We have:
  - spreaders
  - closed communities that can stop the contagion



Epidemics

- Epidemic disease as another form of cascading behavior in a network
- ex.: tuberculosis outbreak
- network level dynamics are similar and insights from the study of biological epidemics are also useful in thinking about the process by which things spread in other networks (e.g., rumors and fake-news)



Tools and Software

Network analysis software

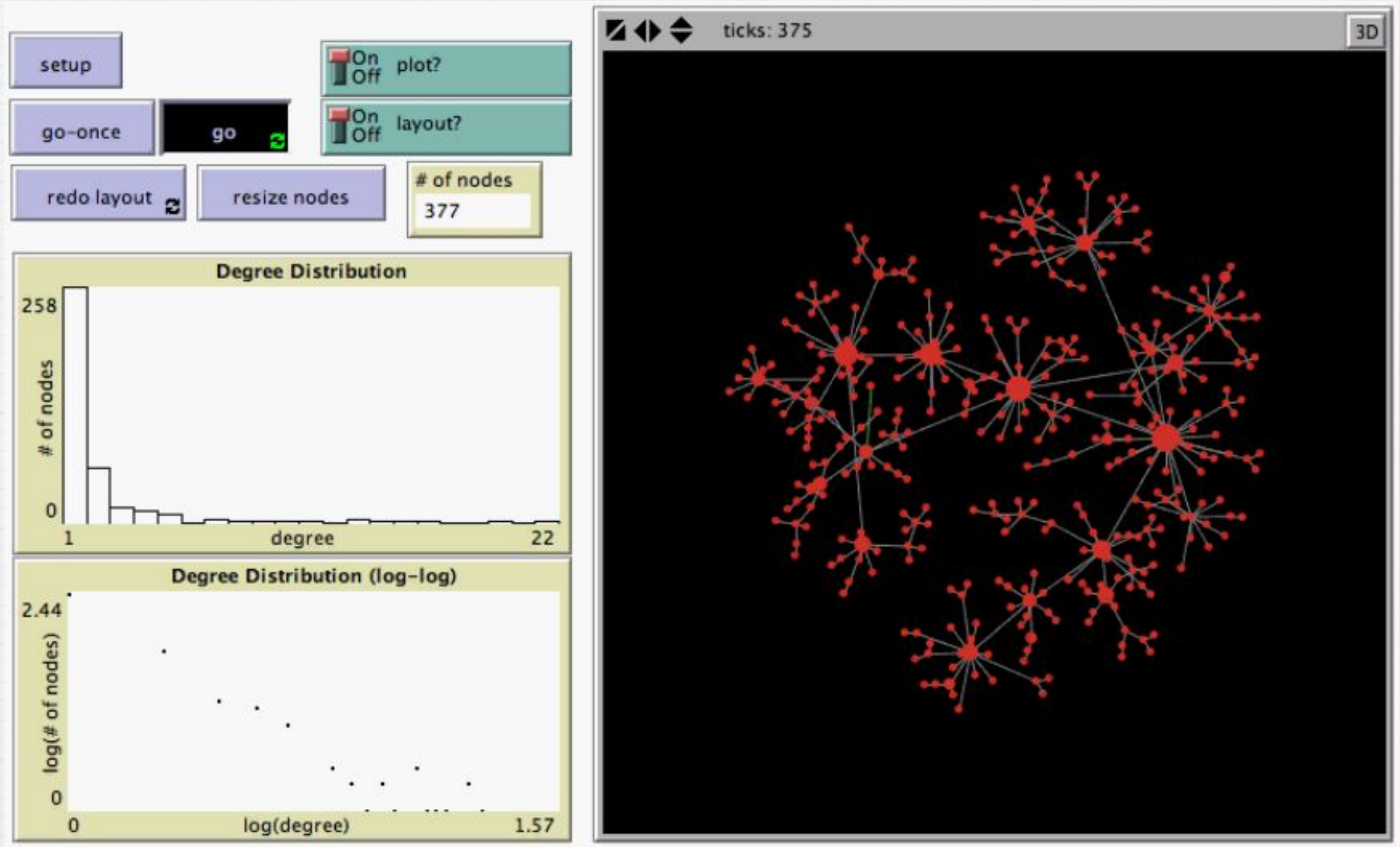
- Python 3+Jupyter/IPython
- alternatively: R
- (We can show examples with both languages)
- local distribution: Anaconda (recommended)
- cloud distributions (free): many
  - ask the instructors for information

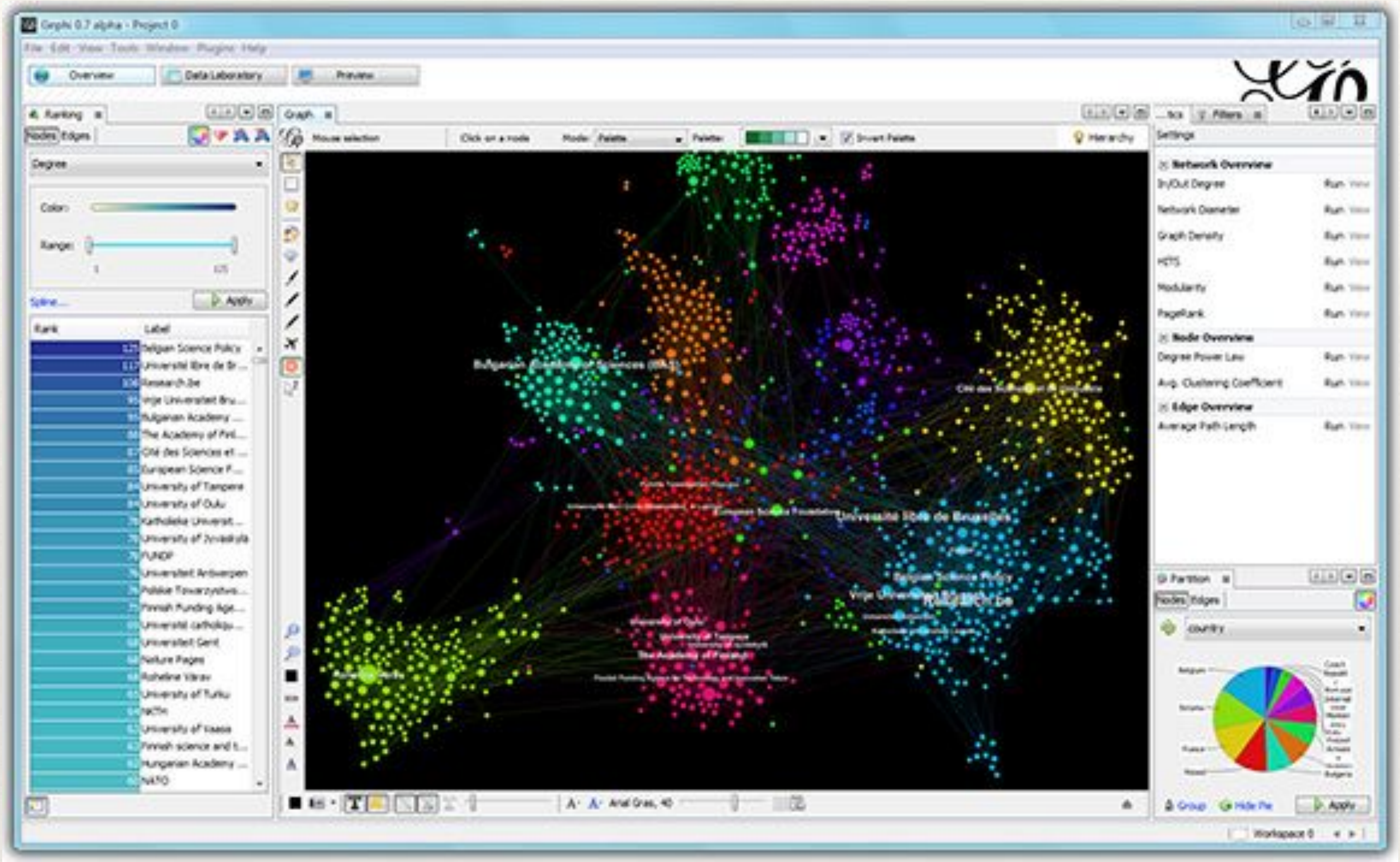
Libraries/packages

	graph-tool	iGraph	Net-workX	NetworKit
Implementation language	C/C++	C/C++	Python	C/C++
Language bindings	Python	C, Python, R	Python	C++, Python
Installation effort	Hard	Medium	Easy	Medium
OpenMP support	Yes	No	No	Yes
Relative slowdown[17]	1	1--4	40--135	N/A
Built-in community detection	Yes	Yes	No	Yes
Built-in advanced layouts	Yes	Yes	No	Yes

Other software and tools

- Gephi (gephi.org)
- NetLogo (ccl.northwestern.edu/netlogo)





Q&A