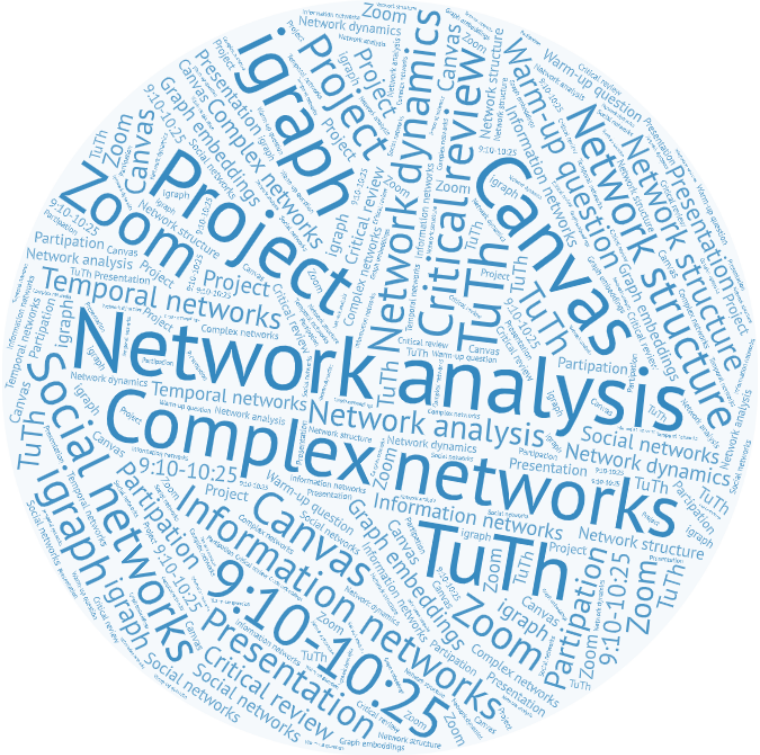
# CptS 591: Elements of Network Science

Assefaw Gebremedhin

assefaw.gebremedhin@wsu.edu

http://www.eecs.wsu.edu/~assefaw

Spring 2021



# About me

- Name: Assefaw Gebremedhin
- Email: assefaw.gebremedhin@wsu.edu
- Webpage: www.eecs.wsu.edu/~assefaw
- Research interests: data science & AI, network science, high performance computing, and applications in bioinformatics, energy systems, and cybersecurity
- Lab: Scalable Algorithms for Data Science Laboratory (https://scads.eecs.wsu.edu)
- NSF CAREER project: Fast and Scalable Combinatorial Algorithms for Data Analytics www.eecs.wsu.edu/~assefaw/fascada
- Teaching at WSU:
  - CptS 475/575: Data Science (Fall 2015--2020)
  - CptS 591: Elements of Network Science (Spring 2015--2020)
  - CptS 317: Automata and Formal Languages (Spring 2020, Spring 2021)
  - CptS/STAT 424: Data Analytics Capstone (Spring 2019)

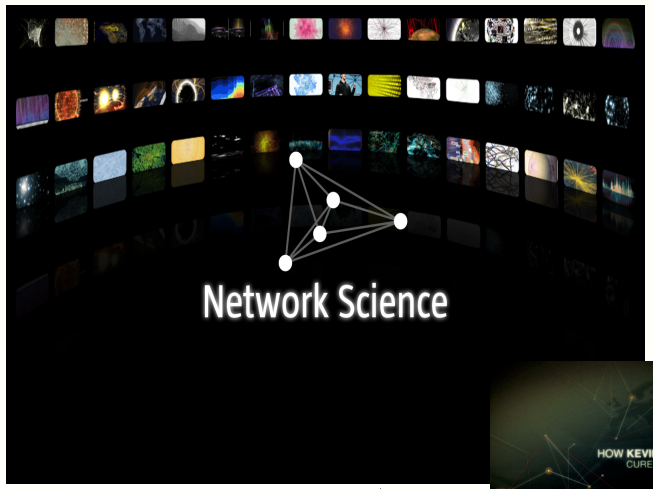
- CptS 591, Spring 2021
  - Lectures: TuTh 9:10—10:25am (via Zoom)
  - Office Hour: Tuesdays 12:00--1pm (via Zoom)

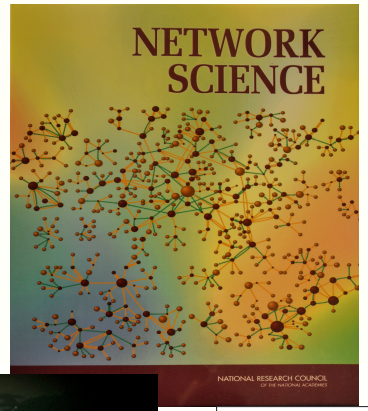
# What I know about the class so far

- 29 enrolled (23 MS students, 6 PhD students)
- Disciplines:
  - Computer Science (22)
  - Electrical Engineering (3)
  - Anthropology (1)
  - Economics (2)
  - Veterinary Science (1)

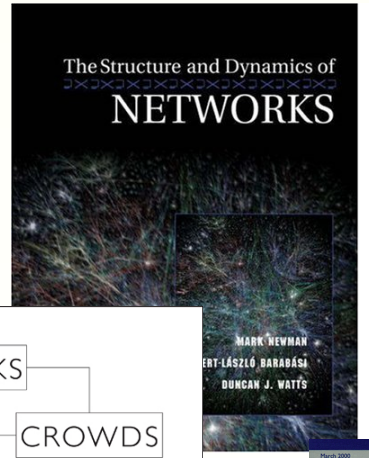
# Big Picture

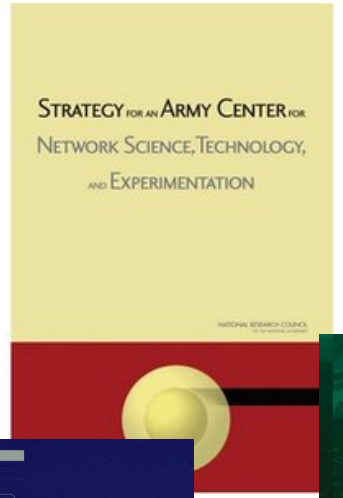
# Who's talking networks?

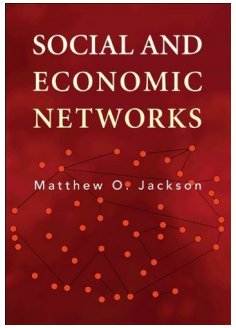


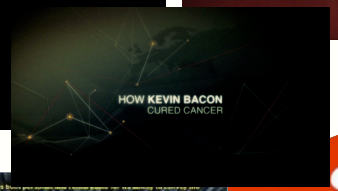



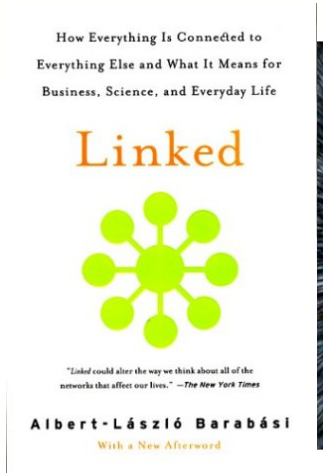

Network Science Center
West Point



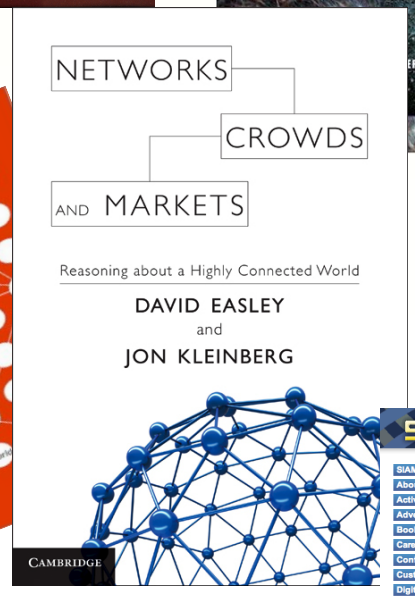



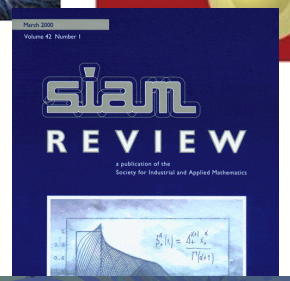



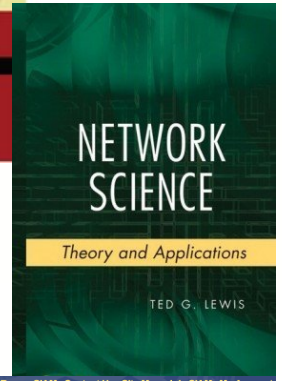



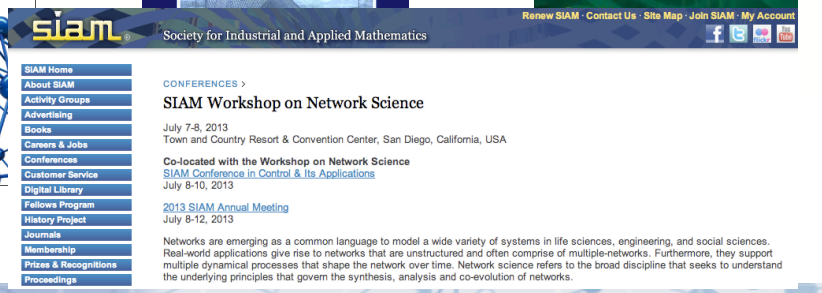







# Complex *connectedness* is everywhere!

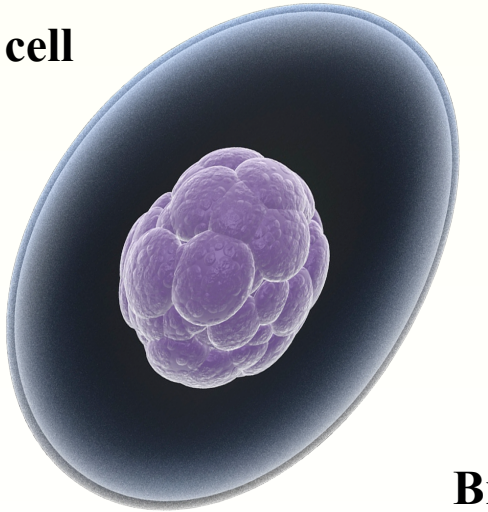
- The *social* interconnections we have
- The *information* we consume
- The *technological* systems we use
- The *economic* systems we live in
- The *political* systems we operate in
- The *organizations* we work at
- The *institutions* we belong to
- The *ecological* systems around us
- *Ourselves* (cell, brain)
- ….

# Complex connectedness is everywhere (in pic)

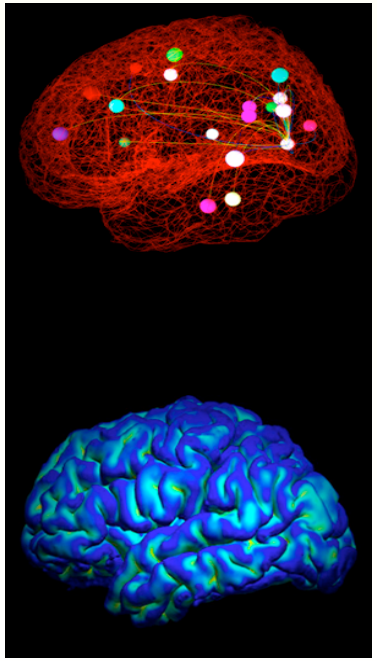
*(Pictures here and elsewhere, unless stated otherwise, are courtesy of Barabasi et al, Network Science Course, NEU,* http://barabasilab.neu.edu/courses/phys5116/.)


**World economy**

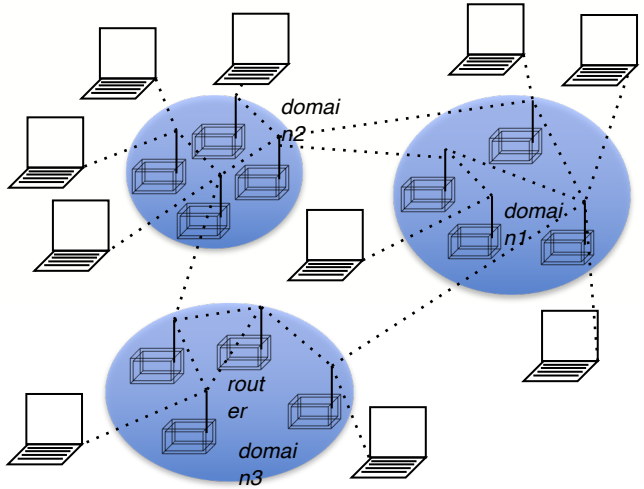
**Human cell**



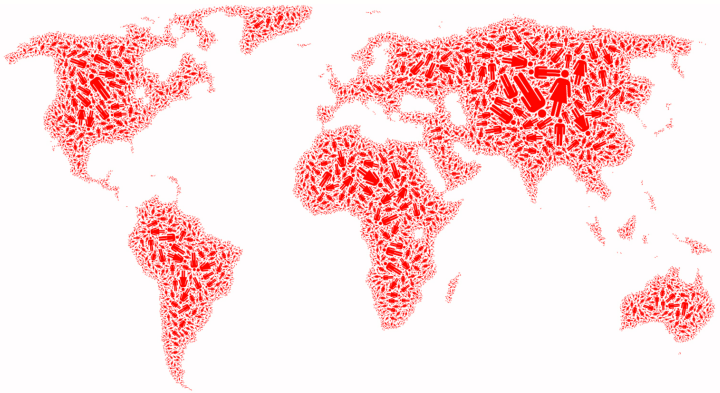
**Brain**



**Internet**



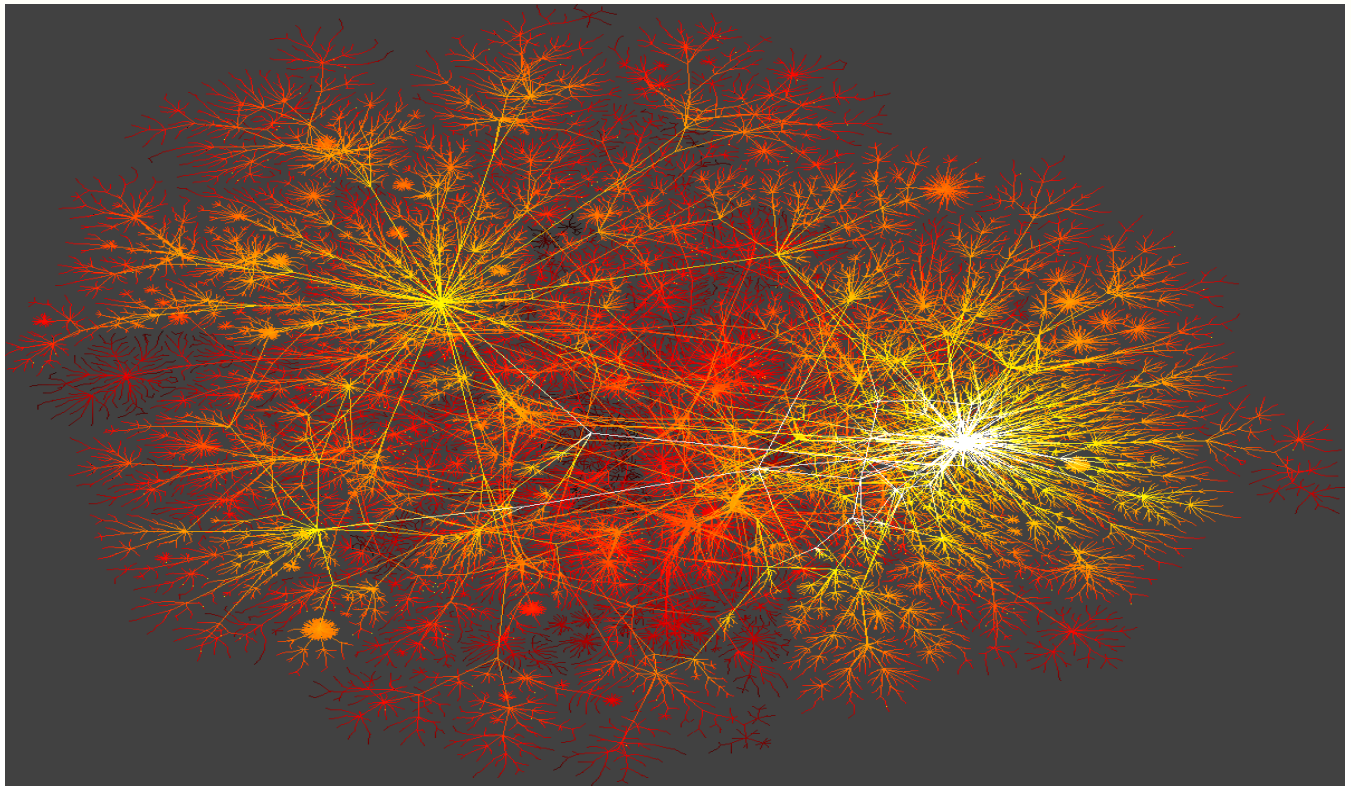

**Society**



...and many more

# An underlying feature:

Behind each such system there is an intricate wiring diagram, *network*, that encodes the interactions between the components.



And to understand the systems, we must understand the networks behind!

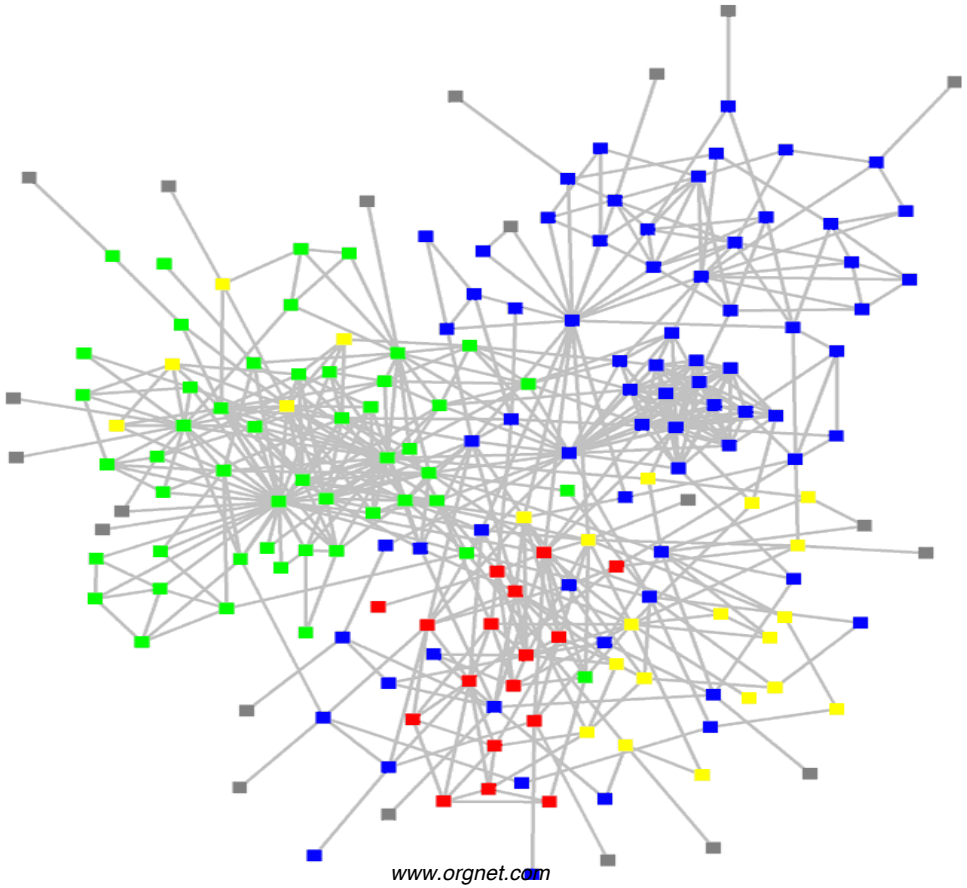
# Networks: Social

# Networks: structure of an organization

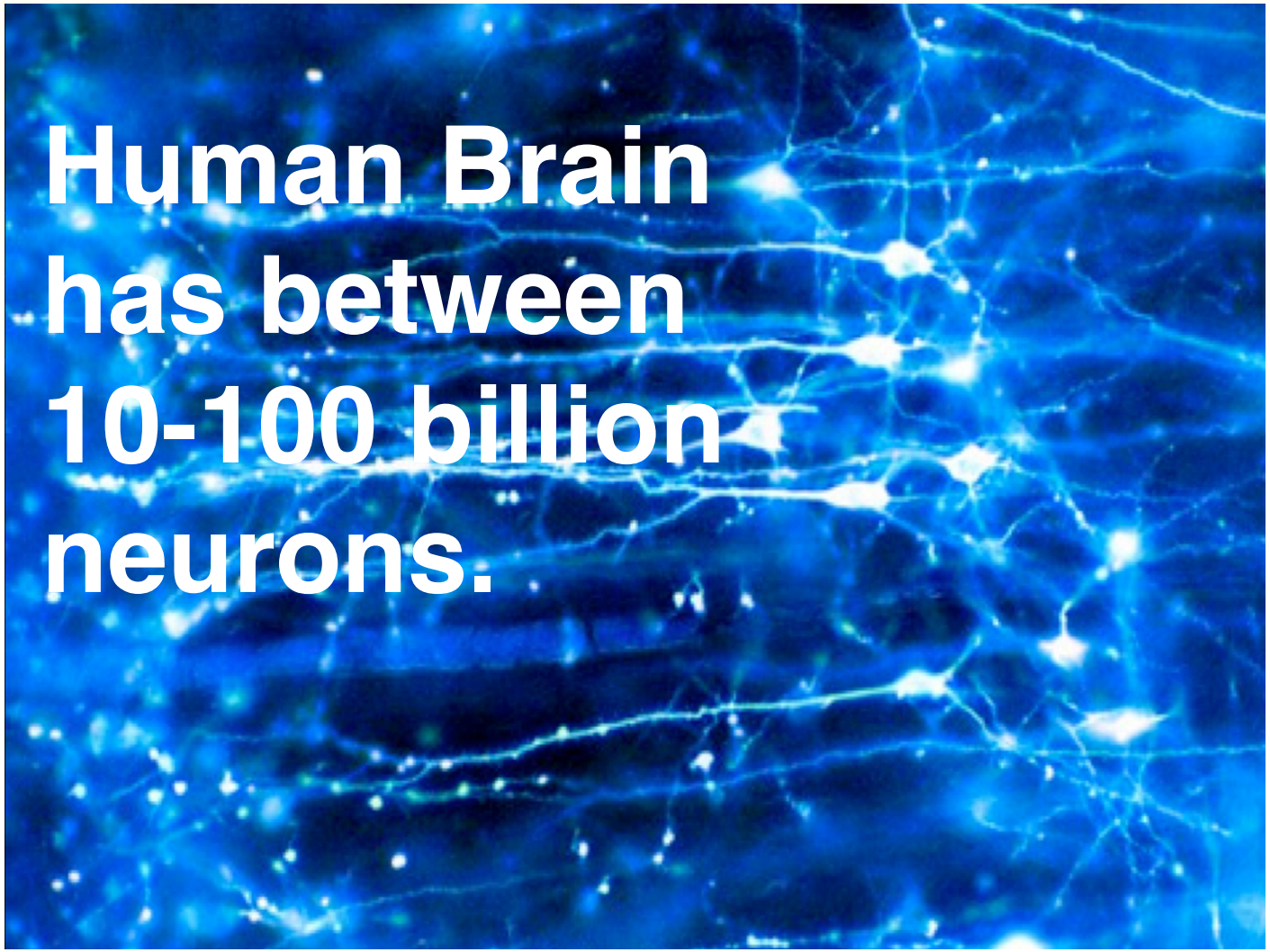
: departments

: consultants

: external experts

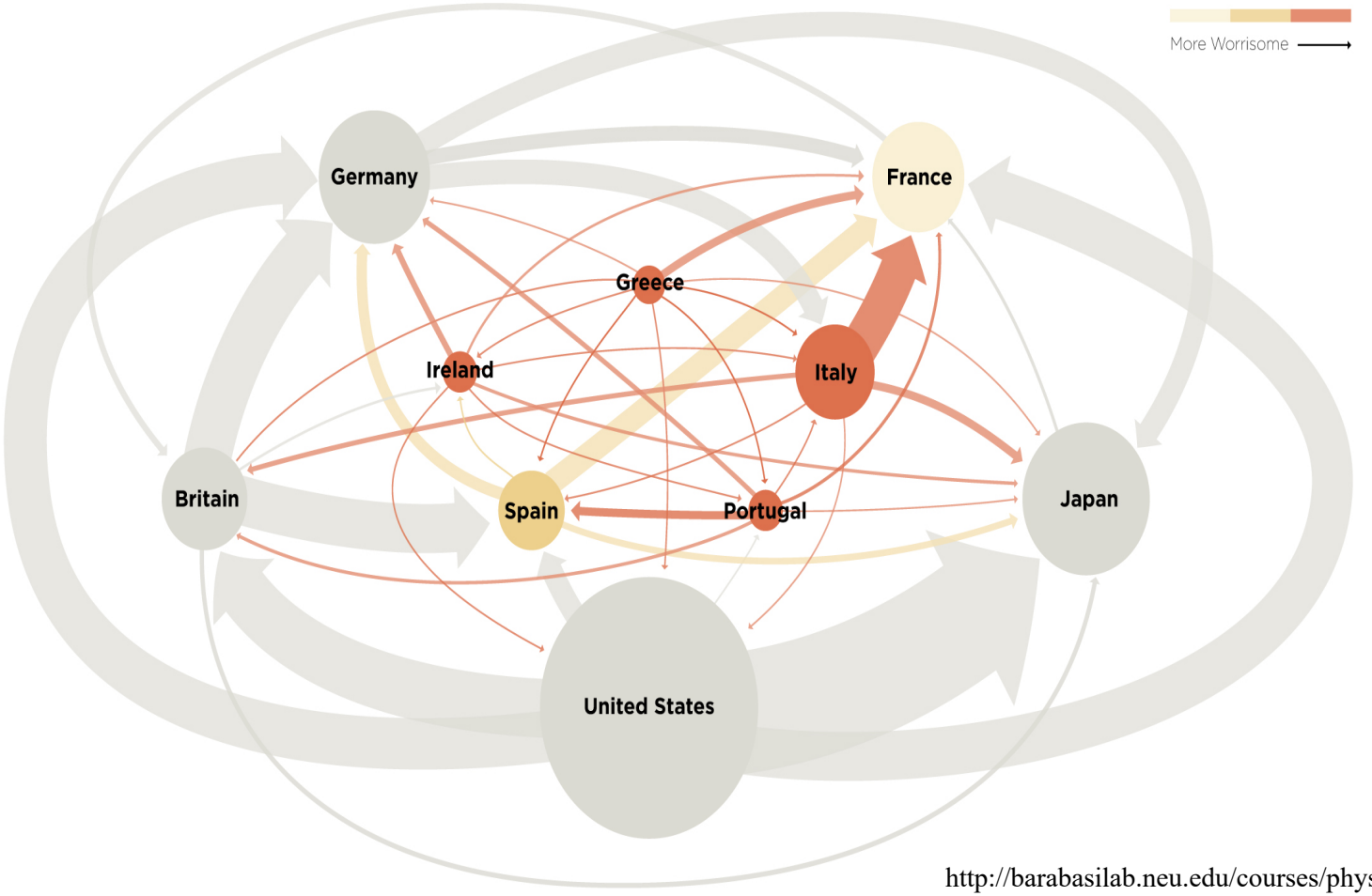


# Networks: Brain




http://barabasilab.neu.edu/courses/phys5116/

# Networks: Financial

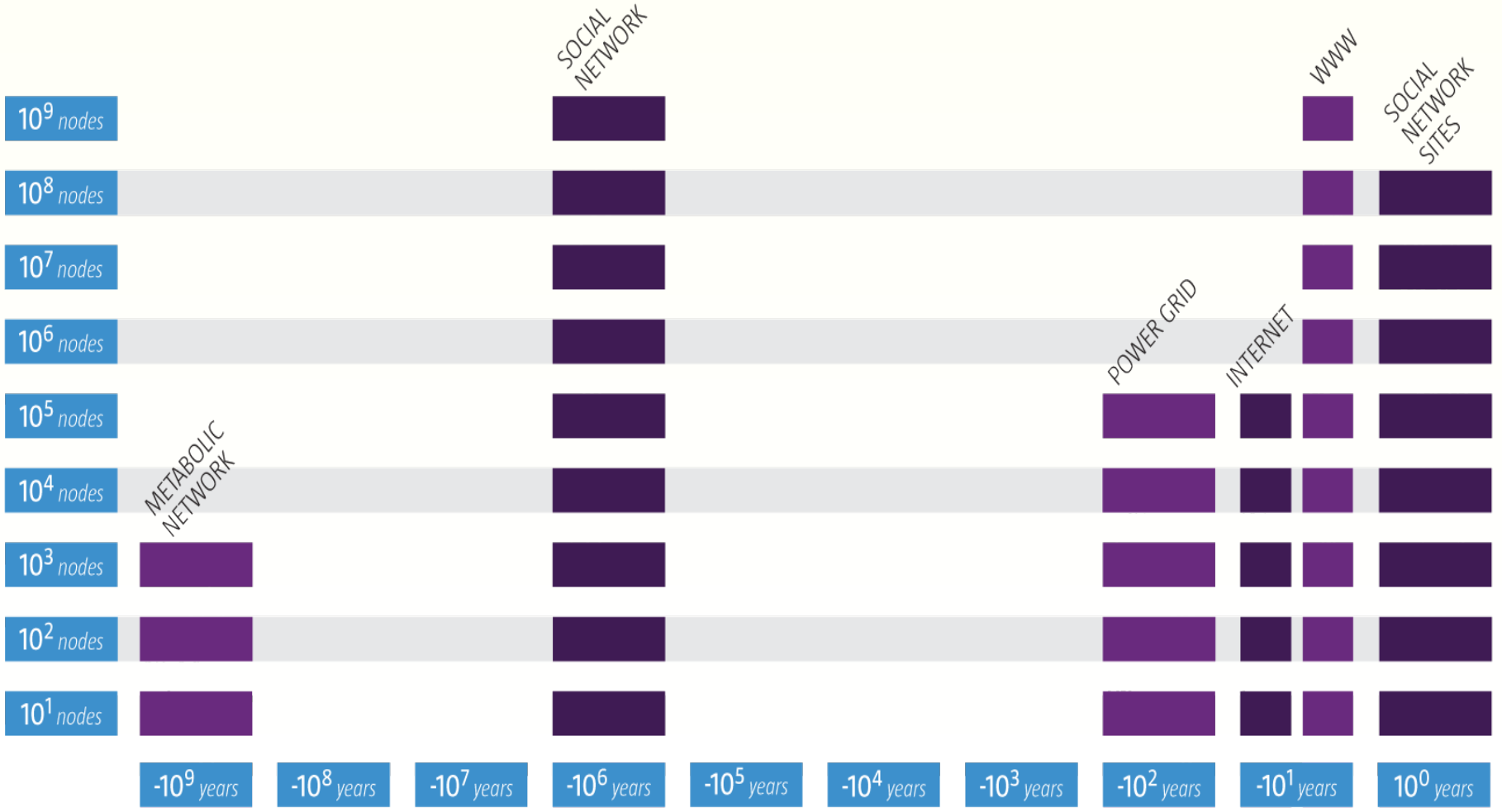



http://barabasilab.neu.edu/courses/phys5116/

# Reasoning about networks

- Study aspects
  - Structure and Evolution
  - Behavior and Dynamics
- Full understanding requires synthesis of ideas from various disciplines, including
  - Computer science
  - Applied mathematics
  - Natural sciences
  - Statistics
  - Economics
  - Sociology

# Networks, why now?




http://barabasilab.neu.edu/courses/phys5116/

# Catalysts for emergence of network science

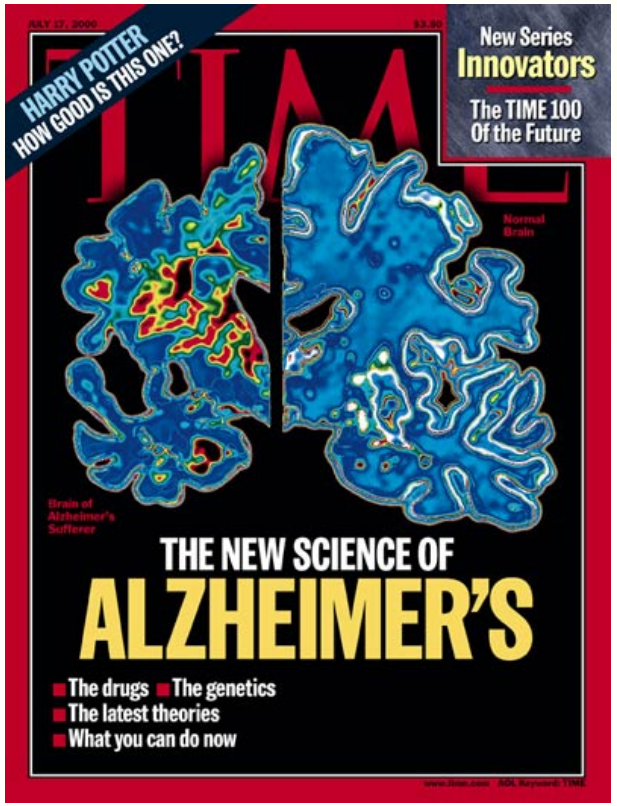
- Availability of network "maps"
  - The Internet, cheap digital storage, and computational technologies made it possible to collect, assemble, share, and analyze data pertaining to real networks
- Recurring similarity
  - Networks from science, nature, and technology are more similar than one would expect
- Confluence of ideas and tools
  - Newer ways of reasoning about interconnectedness are being born by integration of ideas and tools from various disciplines

# Characteristics of Network Science

- Interdisciplinary
  - Common language for interaction
  - Cross-fertilization of ideas and tools
- Empirical, data driven
  - Focuses on data and utility
- Quantitative and Mathematical
  - Graph theory (to deal with graphs)
  - Statistical physics (to deal with randomness and universal organizing principles)
  - Engineering + control + information theory + statistics + data mining (to deal with extracting information from incomplete and noisy data)
- Computational
  - Size of networks and nature of data result in formidable computational challenges
  - Algorithms, database management, data mining

# Impact of network science

- **Economic**
  - Web search
  - Social networking
- **Health**
  - Drug design
  - Metabolic engineering
- **Security**
  - Fighting terrorism (net-war)
- **Epidemics**
  - Epidemic prediction (biological, electronic viruses)
  - Halting spread
- **Brain Science**
  - In 2010 NIH initiated the Connectcome project, aimed at developing a neuron-level map of mammalian brains
- **Management**
  - Uncovering the internal structure of an organization



# Economic Impact



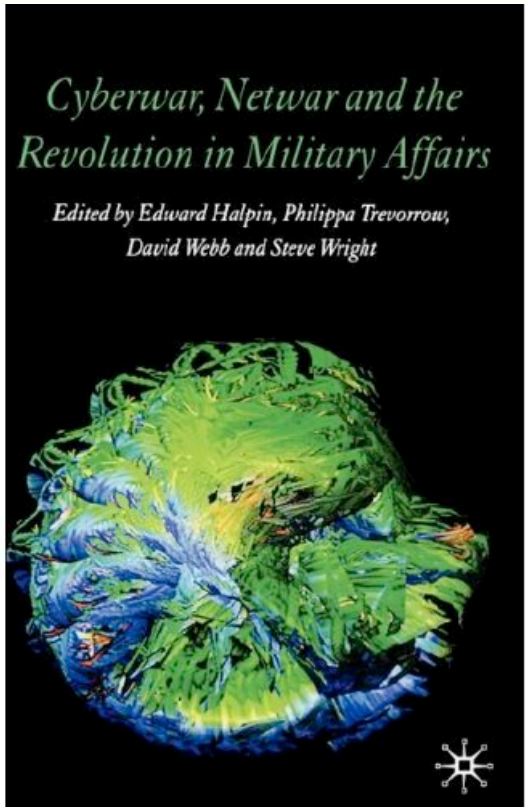

**Google**
Market Cap(2010 Jan 1):
*$189 billion*

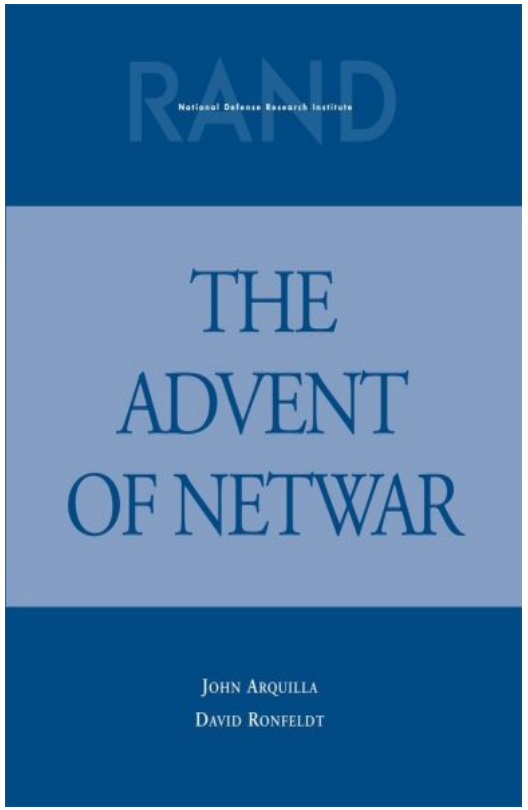
**Cisco Systems**
networking gear Market cap (Jan 1, 2919):
*$112 billion*

**Facebook**
market cap:
*$50 billion*

*www.bizjournals.com/austin/news/2010/11/15/facebooks... - Cached*

# Military impact

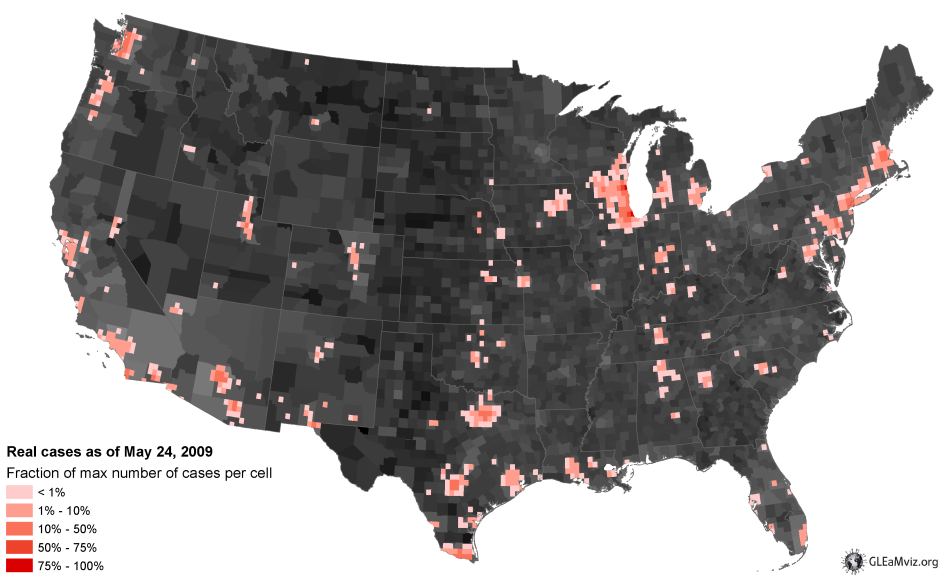


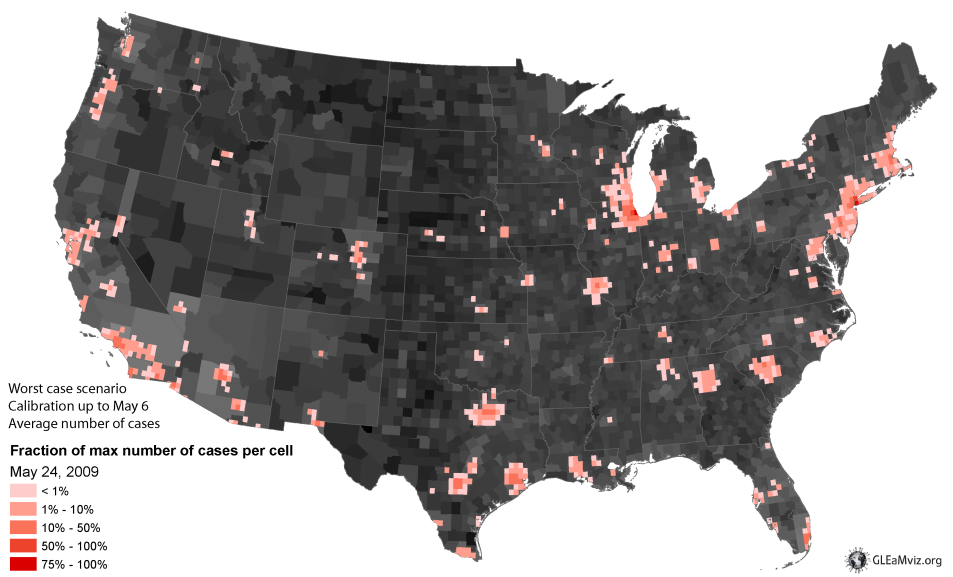

*http://www.slate.com/id/2245232*

# EPIDEMIC FORECAST
## Predicting the H1N1 pandemic

**Real**




**Projected**




http://barabasilab.neu.edu/courses/phys5116/

# This course in focus

# Goals

Students will be introduced to select

- mathematical and computational methods used to *analyze* networks
- models used to understand and *predict* behavior of networked systems
- theories used to reason about network *dynamics*

And students will apply what they learn by completing a semester project and a set of assignments

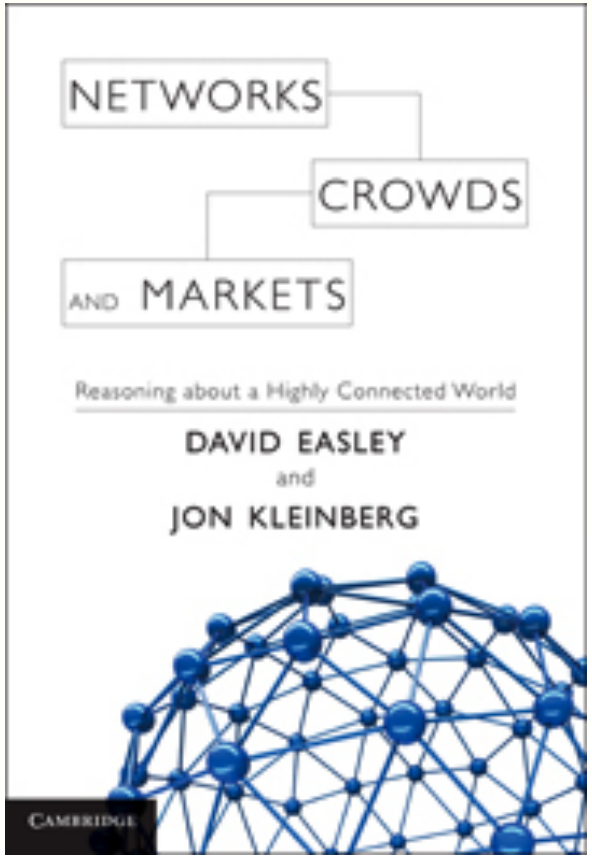
# (Tentative) list of topics

- Network structure, modeling and algorithms
  - Graph theory refresher
  - Basic network properties
  - Random graphs
  - Spectral analysis
  - Centrality
  - PageRank, Hubs and Authorities
  - Graph similarity
  - Community detection
  - Signed networks
- Graph embeddings and representation learning
- Network dynamics
  - Cascading behaviors
  - Information diffusion
  - Epidemic models
  - Influence maximization
- Temporal networks
  - Models
  - Algorithms
  - Applications

*List may be updated later*

# Books

- Frequent reference:
  - Easley and Kleinberg, Networks, Crowds and Markets, Cambridge Univ. Press, 2010
- Other/related references
  - M.E. J. Newman, *Networks: An Introduction*, Oxford University Press, 2010
  - U. Brandes and T. Erlebach (Eds.), *Network Analysis: Methodological Foundations,* Springer 2005
  - A. Barabasi, *Network Science*, e-book



# Software

We will use igraph as the primary software tool for network analysis:

- *igraph* : *http://igraph.org*

Other related tools

- *networkX* : *http://networkx.github.io*
- *networKit* : *https://networkit.iti.kit.edu/*
- *SNAP* : *http://snap.stanford.edu/*

# Expectation

Basic knowledge of:

- Algorithms
- (Graph theory)
- Linear algebra
- Probability and Statistics

Reasonable programming experience:

Python, R, C/C++, Java

# Course work

- Three assignments + a few exercises (30%)
  - Individual
- One semester project (50%)
  - Collaborative (a team of two or three)
- Mid-term exam (18%)
- Class participation (2%)
  - Discussions in class, in-class exercise

- Project breakdown: 50%
  - Reaction paper: 7%
  - Project proposal: 7%
  - Presentation: 8%
  - Final report: 28%

# Project

- Could take one of several forms:
  - *Experimental analysis* of an interesting dataset using existing methods and software
  - *(Experimental) comparison* of existing methods and software tools in the context of a specific application
  - *Theoretical analysis* of a model/an algorithm in a specific application
  - *Implementation* of a new method
  - *In-depth survey* of a research topic
- Students required to work in teams of two or three
  (solo projects allowed if there are valid reasons)

# Lecture material and resources

- Course website: everything will be on Canvas
  - Slides, reading materials, announcements, and other resources
- Canvas will also be used to handle assignment and project submissions
- The Easley & Kleinberg reference book is available on-line
- Check the Canvas page of the course regularly for info and updates

# Related courses elsewhere

- Cornell (Jon Kleinberg and Eva Tardos, Networks)
  - https://courses.cit.cornell.edu/cs2850_2016fa/
- Stanford (Jure Leskovek, Social and Information Network Analysis)
  - http://web.stanford.edu/class/cs224w/
- Northeastern (Barabasi lab, Complex Networks, Fall 2018)
  - https://www.barabasilab.com/course
- Yale (Dan Speilman, Spectral Graph Theory, Fall 2015)
  - http://www.cs.yale.edu/homes/spielman/561/

# A few words on policies

- Classroom conduct
  - Mute when not speaking
  - Arrive on time and remain throughout the class
- Correspondence
  - All class related correspondence should be made via Canvas
- Class participation
  - Required
- Late work or missing work
  - Submissions should happen by due dates and times
  - Late submissions up to 48 hours accepted with 10% grade penalty per 24 hours late
  - Missing work or work submitted  more than 48 hours late will be counted as zero
- Academic Integrity
  - Strictly enforced
- Read syllabus for COVID-19 policy and other statements

# Thanks!

- Welcome, once again
- It is going to be a fun semester
- Put your best effort
- You will be rewarded