Week 01

Introductions

Tuesday, August 24

INFO 5613: Network Science

Brian Keegan
brian.keegan@colorado.edu

University of Colorado Boulder

Agenda

- Introductions
- Course overview
- A brief history of network science
- An information scientist mindset
- Setting up computing environment

Introductions

- Name and pronouns
- Program and year
- Connection to network science
- Greatest achievement over summer

Course overview

What Are We Doing Here?

- Understand the theoretical and methodological implications of relational data
- Apply and interpret metrics for understanding network structure and dynamics
- Develop familiarity with computational tools for analyzing and visualizing networks
- Integrate and explain network methods and theories for general audiences

Schedule

Module	Week	Dates	Topics	Due Date
Fundamentals	1	Aug 24; Aug 26	Fundamentals of networks	*September 17*
	2	Aug 31; Sep 2	Data and ethics of networks	
	3	Sep 7; Sep 9	Visualizing networks	
Structure	4	Sep 14; Sep 16	Node-level structure	*October 15*
	5	Sep 21; Sep 23	Local-level structure	
	6	Sep 28; Sep 30	Network-level structure	
	7	Oct 5; Oct 7	Community structure	
Dynamics	8	Oct 12; Oct 14	Random networks	*November 12*
	9	Oct 19; Oct 21	Network growth	
	10	Oct 26; Oct 28	Diffusion and influence	
	11	Nov 2; Nov 4	Homophily and selection	
Applications	12	Nov 9; Nov 11	Bipartite networks	—
	13	Nov 16; Nov 18	Weighted networks	
	14	Nov 23; Nov 25	**No Class: Fall Break**	
Presentations	15	Nov 30; Dec 2	Presentations	*December 13*
	16	Dec 7; Dec 9		

Course Design

- Four modules
  - **Fundamentals** – data analysis in Python, relational data structures, visualizing networks
  - **Structure** – measuring structure at the node, local, network, and community levels
  - **Dynamics** – changes in network structure and node attributes
  - **Applications** – bipartite/affiliation and weighted+multiplex networks

- **Tuesdays**: lecture + notebook exercises
- **Thursdays**: review notebook + discussion of readings

- In-class and Canvas announcements override syllabus

Evaluation

- **Module Assignments** – 30% – Due the Friday after the module ends
  - 3 total: Sep 17, Oct 15, Nov 12
- **Reading Responses** – 26% – Due Wednesdays by 5pm
  - 13 total: reactions, extensions, explorations, *etc*. to readings to structure Thursday discussion
- **Final Project** – 29% – Due December 14
  - Will be broken up into components with more details and deadlines through semester
- **Attendance** – 15% – Every class
  - 2 classes per week for 15 weeks: Coming to both classes for full credit, see syllabus for absence policy

We will use Python a lot, but all evaluation will be non-technical written submissions.

Computing Requirements

- All code will be delivered using Jupyter Notebooks in Anaconda/Python
  - If you have never used these before or have a new computer: download it
  - If you have used these before, stay tuned
- Readings and lectures will be posted to Canvas: https://canvas.colorado.edu/courses/76045
- Class will meet via Zoom: https://cuboulder.zoom.us/j/96935610716

If you do not have reliable access to a computer to access Anaconda, Canvas, and Zoom: please contact the instructors immediate to work out an accommodation

Class Norms: Zoom Etiquette

- Zoom meeting for each class will be recorded and posted to Canvas
- Participants will be muted by default and should use “Raise Hand” to be called on
- Context switching (Zoom ↔ Jupyter) will be necessary, but please no background browsing
- Use public chat for questions and comments, but don’t use private chat for side conversations
- Disruptive behavior will get a single warning before being kicked and losing day’s participation

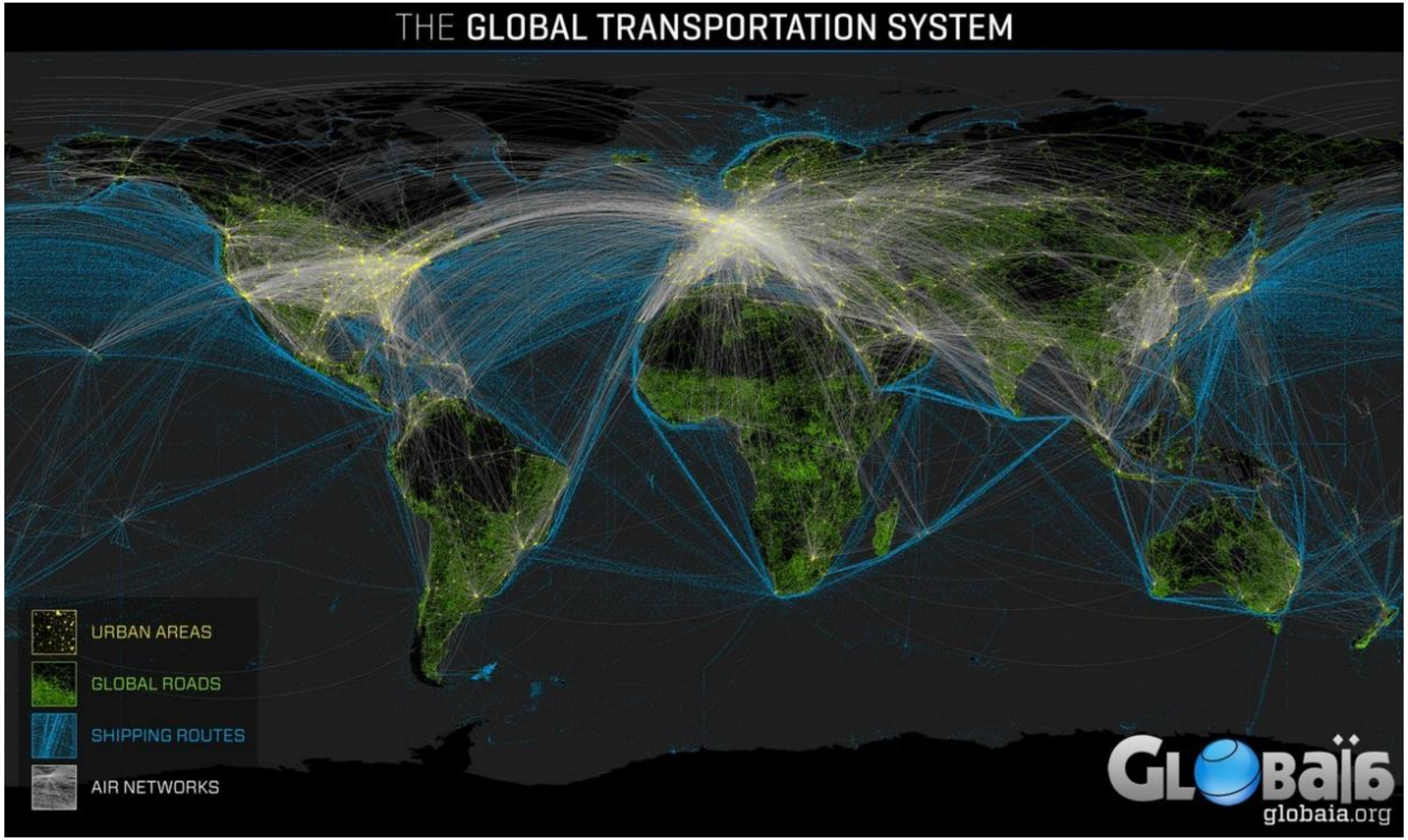
What are other norms and lessons to incorporate from your experiences in other classes?

A brief history of network science

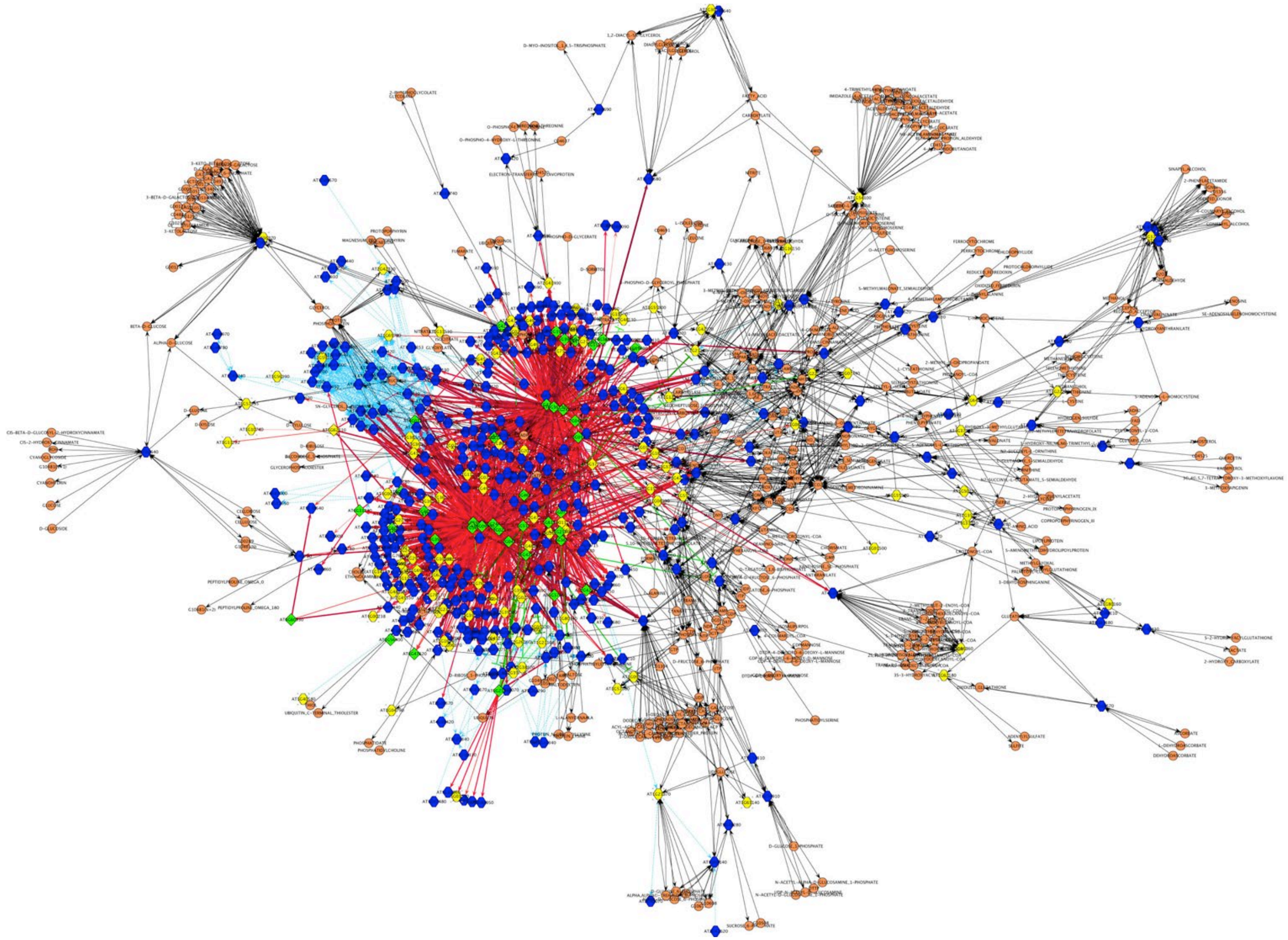
Network science

- Freeman's four defining properties
  - Links among actors are important
  - Collection and analysis of relations linking actors
  - Attention to patterns of those links
  - Use of mathematical and computational models to describe and explain patterns

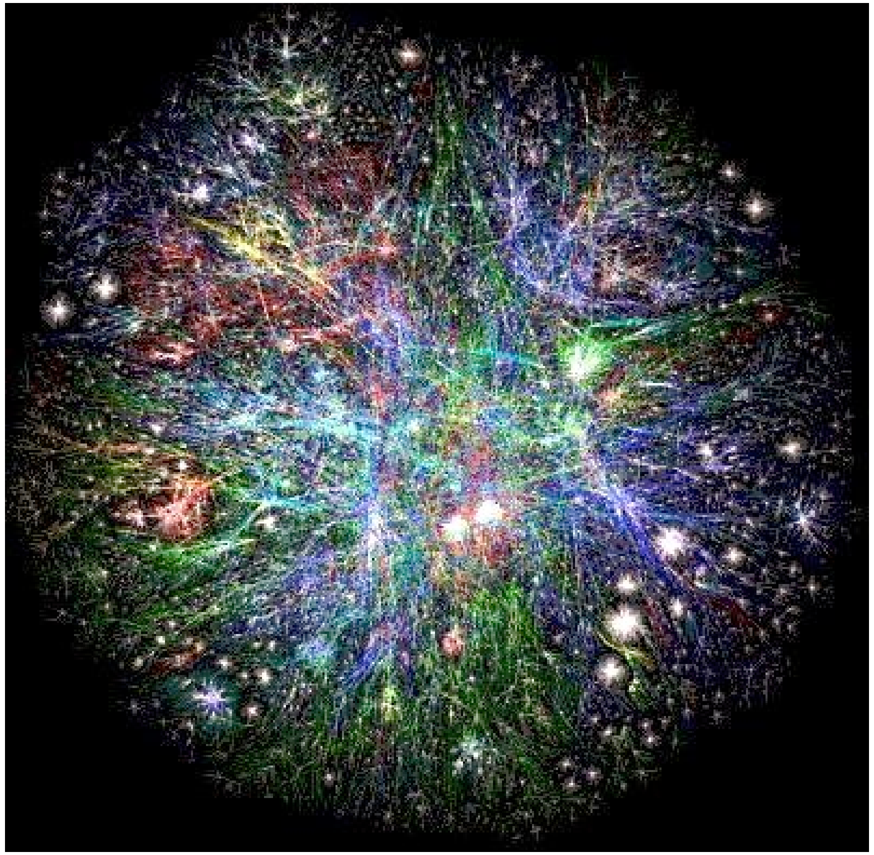
Transportation networks



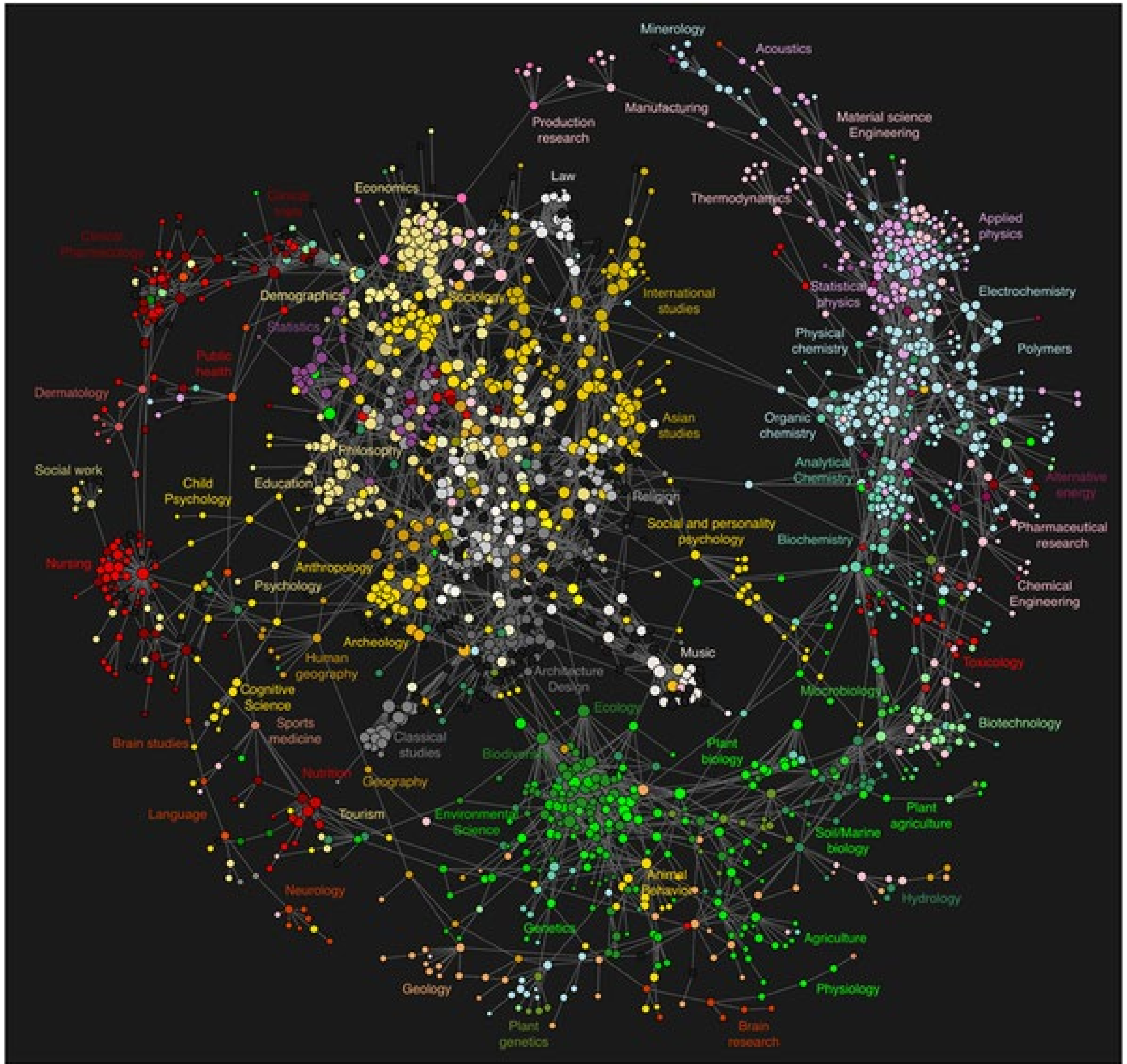
Biological networks



Information networks

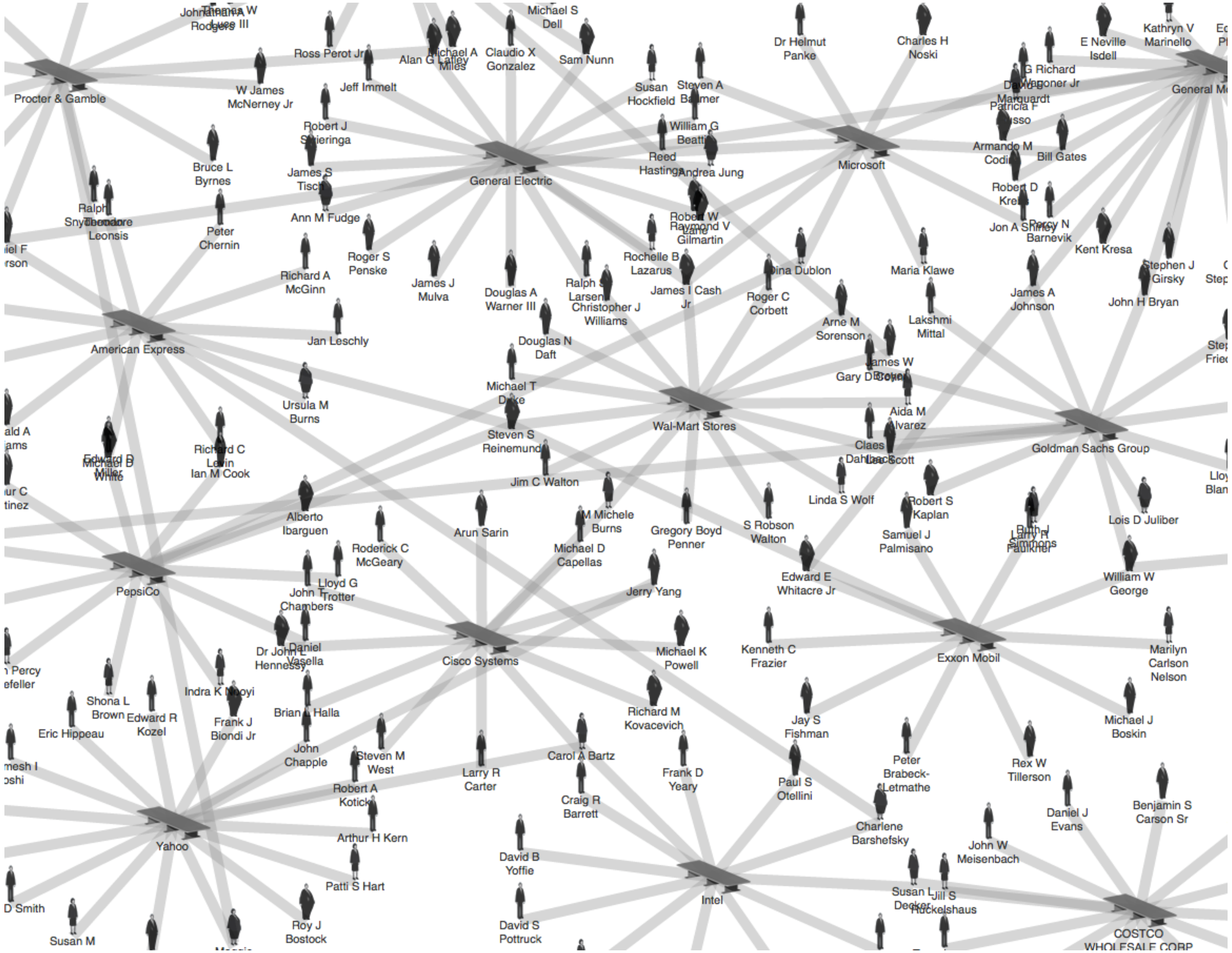


Citation networks



http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0004803

Organizational networks

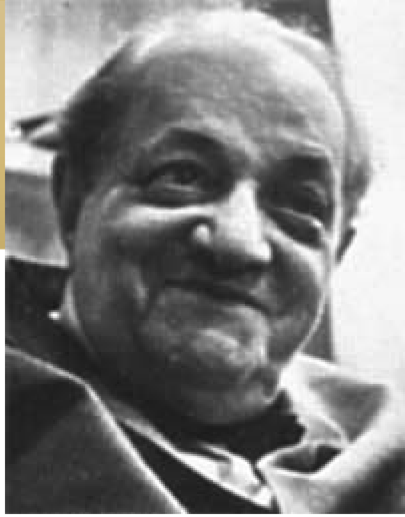


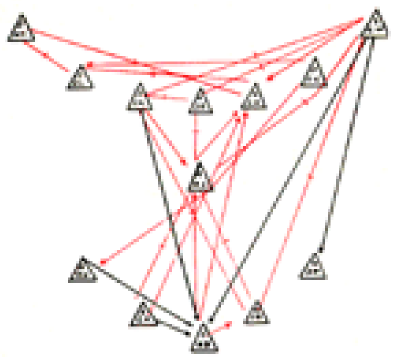
theyrule.net

Social networks

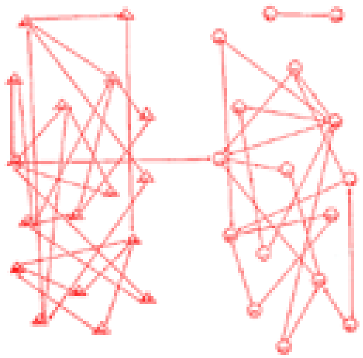
1930s: Jacob Moreno

- Jacob Moreno introduced the ideas and tools of *sociometry*
- *Sociograms* combine network data and data visualization
- Trained as a psychiatrist, notorious womanizer, not a methodologist
- Megalomaniacal tendencies, abandoned sociometry to focus on group therapy but returned with a second edition in 1950s to claim credit for developments
- Still credited as "father" of social network analysis… more on this in a bit

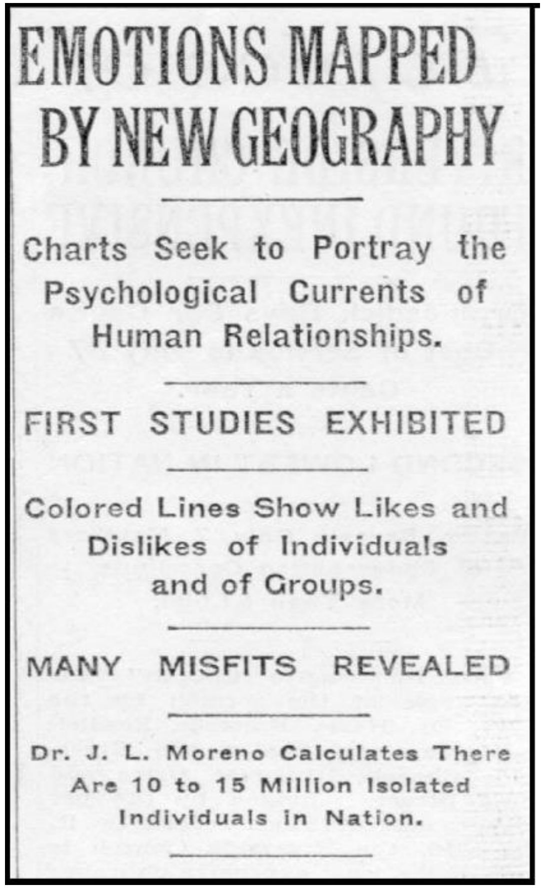




Friendship choices among 4th graders (Moreno 1934)



Positive & negative affect in a football team (Moreno 1934)



EMOTIONS MAPPED BY NEW GEOGRAPHY

Charts Seek to Portray the Psychological Currents of Human Relationships.

FIRST STUDIES EXHIBITED

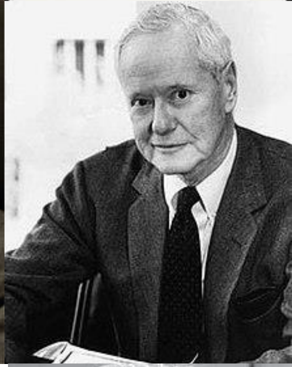
Colored Lines Show Likes and Dislikes of Individuals and of Groups.

MANY MISFITS REVEALED

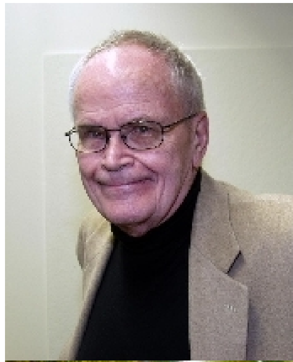
Dr. J. L. Moreno Calculates There Are 10 to 15 Million Isolated Individuals in Nation.

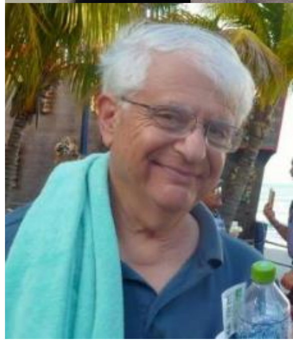
1950s: Group networks

- Kurt Lewin (top), “father” of social psychology, in 1930s is interested in how individual behavior influenced by others’ behaviors, “group dynamics”
- Alex Bavelas (2nd), a Lewin student, pioneered ideas to use experiments to analyze how social structure influences group performance
- Paul Lazarsfeld & Robert Merton (3rd), sociologists who had been exposed to Moreno & Jennings, interested in mass communication and influence
- Everett Rogers (bottom), a rural sociologist interested in agricultural inventions, discovers sociometry to explain “diffusion of innovation”



1960s: Methodological acceleration

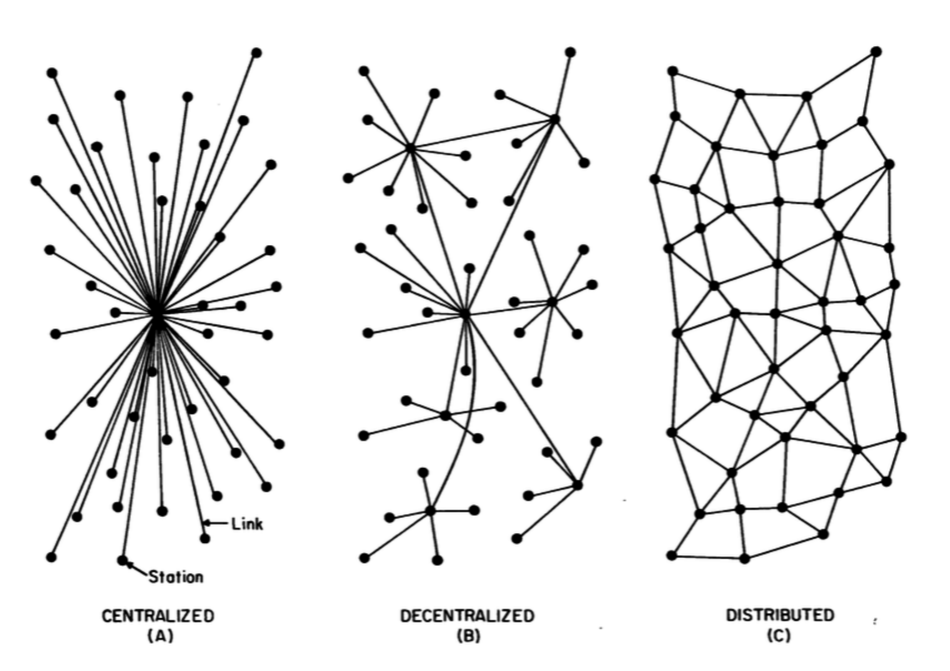


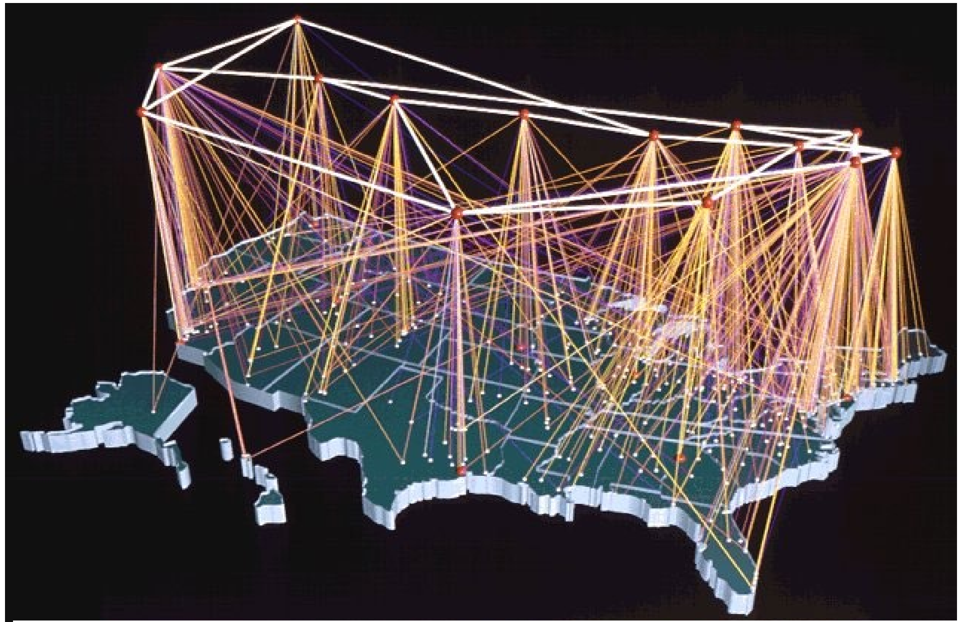


- Stanley Milgram (top), the (in)famous social psychologist, introduces the “small world experiment” popularized by the “six degrees of separation”
- Paul Erdős (2nd), the famous mathematician, develops models of “random graphs” in 1959 that become the foundation of statistical inference on networks
- Harrison White (3rd), a theoretical physicist, worked at Bavelas’ lab in 1950s, and starts a second dissertation in sociology, moves to Harvard in 1963 and starts the “Harvard Revolution” in sociology, advises many influential SNA researchers
- Granovetter (bottom), a White student, publishes “The Strength of Weak Ties” rejected in ’69, accepted in ‘73, now has >45,000 citations making it the most cited paper in sociology; also known for work on embeddedness

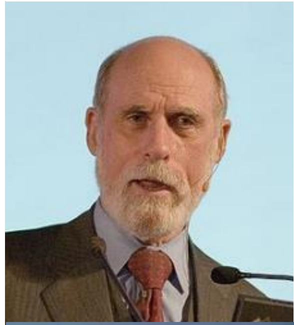
70s-80s: Internet

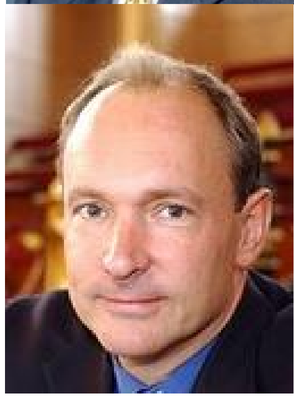
- Defense Advanced Research Projects Agency (DARPA) funds research to support “interconnecting” big, expensive mainframes
- Paul Baran (top) hired to build computer networks that would remain reliable even if central nodes were destroyed (in a nuclear war)
- Vincent Cerf and Robert Kahn (middle) invent protocols like TCP/IP for routing information in a distributed network assuming unreliable nodes
- Tim Berners Lee (bottom) builds on TCP/IP and develops HTML, URLs, web browsers, web servers



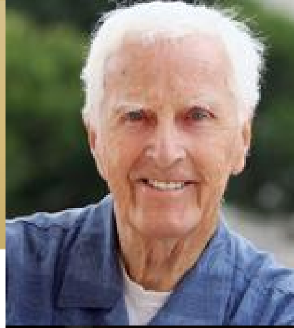


Donna Cox & Robert Patterson, 1992





80s-90s: Social capital

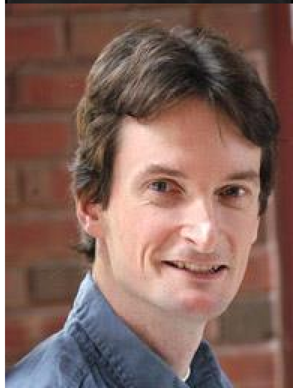


- Linton Freeman (top) integrates mathematical graph theory and sociological role theory to propose new definitions of centrality, including betweenness centrality
- Barry Wellman (2nd) organizes SNA community, studies role of networks in offline communities and in mid 1990s becomes interested in potential of "online communities"
- Ron Burt (3rd) uses Freeman's betweenness centrality to argue that brokers between disconnected groups have more power than people with many connections
- Robert Putnam (bottom) publishes *Bowling Alone* in 1995 arguing the degradation in American social and political life is a lack of meaningful social relations ("social capital")

90s-00s: Network science

- Jon Kleinberg (top), a theoretical computer scientist, begins to analyze the structure of hyperlinks on WWW and develops HITS algorithm, which influences Google's PageRank
- László Barabási (2nd), a statistical physicist, develops the "preferential attachment" model to explain the extremely skewed connection patterns in vastly different datasets
- Duncan Watts (3rd), an applied physicist, develops a "small world model" with advisor Strogatz that explains how a few long-distance ties can create Milgram's small worlds
- Mark Newman (bottom), a statistical physicist, develops efficient statistical methods for finding community structure, assortative degree mixing, and modeling connectivity



00s-10s: Social computing





- Joe Konstan & John Riedl (top) developed “recommender systems” from user-movie networks that now power algorithms behind Netflix, Amazon, Facebook, *etc*.
- Alex Vespignani (2nd) developed models of dynamic networks with applications for predicting and controlling epidemic processes (diseases, memes, *etc*.)
- David Lazer (3rd) proposed that large digital data can enable new forms of social science not reliant on traditional methods like surveys, ethnography, interviews, lab experiments, *etc*.
- Matt Salganik (bottom), a Watts student, did first large-scale online social experiments showing how popularity dynamics can be biased and manipulated

Network science is not all white & male

- **Beth Wellman** (1926), **Helen Bott** (1928), **Elizabeth Hagman** (1933) pioneered network data collection by examining relationships among children. They combined interview and observational data, developed matrix notation for relationships, and measured biases in self-reported data
- **Helen Hall Jennings** (1934), a psychology PhD student at Columbia specializing in statistics, almost certainly drove most of the methods and data collection of Moreno's "sociograms" but was only partially credited by Moreno until 1980s → mother and co-founder of SNA
- **Allison Davis** and **St. Clair Drake** analyzed relations in their 1940s African American communities, their Southern Women network dataset is a classic on social stratification
- **Elizabeth Bott** analyzed how family roles structured urban families' social networks in 1950s
- **Kathleen Carley**, **Bonnie Erickson**, and **Karen Cook** emerged in 1980s applying networks to study textual data, deviance, and exchange/trust (respectively)
- **Hermina Ibarra, Wendy Hall, Lada Adamic**, **Ginestra Bianconi**, **Jennifer Neville**, **Leman Akoglu**, **Kristina Lerman**, **Lise Getoor**, **Karrie Karahalios**, **Raissa D'Souza**, **Marta Gonzalez**, **Sandra Gonzalez-Bailon**, **Meeyoung Cha**, **Vittoria Colizza**, **Danielle Bassett**

An information science mindset

A Modest Proposal

Growth Mindset

Computational Thinking

Hacker Ethic

Scientific Norms

A Modest Proposal

Growth Mindset

Computational Thinking

Hacker Ethic

Scientific Norms

Growth Mindset

- Pre-Renaissance: fixed endowments, predetermined positions, ability is heritable, mobility is a threat to social order
  - Not gone: high-stakes testing used to “discover” ability and to exclude
- Early cognitive psychology
  - Learned helplessness: animals exposed to uncontrollable shocks made little effort to prevent shocks even after they were controllable
  - Attribution theory: people’s explanations for what happens to them shapes their reactions to future events
- Fixed mindset: fears lack of ability, need to prove ability, challenge-avoiding, stereotype belief/threat, mind as a mold
- Growth mindset: fears lack of effort, need to improve ability, challenge-seeking, resilience, mind as a muscle
- We make these judgments about ourselves and others

Dweck, C. S., & Yeager, D. S. (2019). Mindsets: A View From Two Eras. *Perspectives on Psychological Science*; Buchanan, A. (2017). The nature of mindsets.

Growth Mindset Questions

- How will you challenge yourself today?
- What can you learn from a mistake?
- Who can you ask for help?
- How could you make this more interesting?
- What can you do to focus better?
- What’s the next challenge to tackle?
- What support resources can you find?
- What else do you want to learn?

A Modest Proposal

Growth Mindset

Computational Thinking

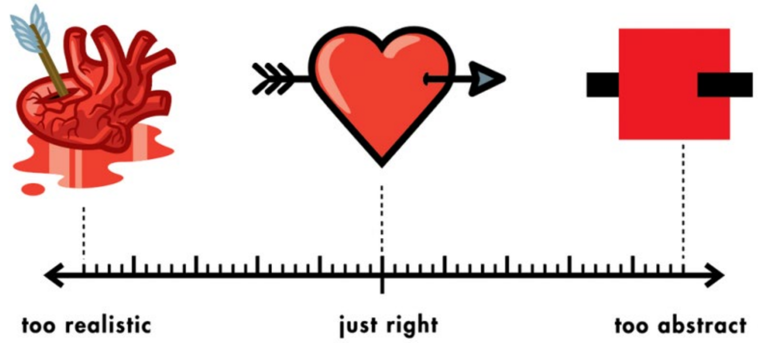
Hacker Ethic

Scientific Norms

Computational Thinking (CT)

- Formulating problems and their solutions that can be executed computationally
- Reformulation, recursion, decomposition, abstraction, testing **[Wing 2006, Grover & Pea 2013]**
- Creativity, algorithmic thinking, cooperativity, critical thinking, problem solving **[Korkmaz, et al. 2017]**
- Data practices, simulation, computational problem solving, systems thinking **[Weintrop, et al. 2016]**
- Decomposition, abstraction, algorithms, debugging, iteration, generalization **[Shute, et al. 2017]**

THE ABSTRACT-O-METER



https://computersciencewiki.org/index.php/File:Abstract_heart.png

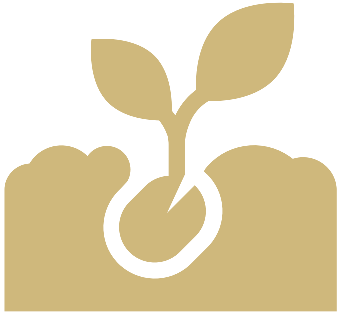
Wing, J.M. (2006). Computational thinking. *CACM*; Grover, S., & Pea, R. (2013). Computational Thinking in K–12. *Educational Researcher*; Korkmaz, Ö., et al. (2017). A validity and reliability study of the computational thinking scales. *Computers in Human Behavior*; Shute, V., et al. (2017). Demystifying computational thinking. *Educational Research Review*; Weintrop, D., et al. (2016). Defining Computational Thinking for Mathematics and Science Classrooms. *Journal of Science Education and Technology*.

CT Isn't Just Technical Skills

- Concepts: sequences, loops, parallelism, events, conditionals, operators, data
  - This is what INFO 1201 and 2201 are designed for and INFO 3401 and 3402 will continue to reinforce
- Practices
  - <u>Experimenting and iterating</u>: developing, experimenting, and developing some more
  - <u>Testing and debugging</u>: making sure things work and solving problems when they arise
  - <u>Reusing and remixing</u>: building on existing projects or ideas and sharing your own work
  - <u>Abstracting and modularizing</u>: building something complex by putting together smaller parts
- Perspectives
  - <u>Expressing</u>: computation as a medium for creative and critical expression
  - <u>Connecting</u>: computation as a tool for of creating for and interacting with others
  - <u>Questioning</u>: computation as a tool for investigating how the world works

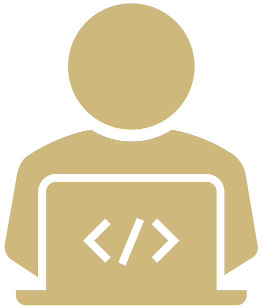
Brennan, K. & Resnick, M. (2012). *New Frameworks for Studying and Assessing the Development of Computational Thinking*.

A Modest Proposal

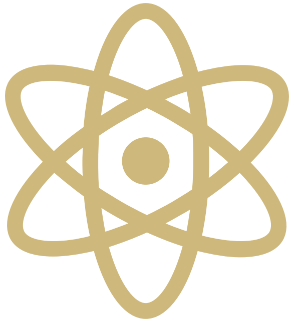


Growth Mindset

Computational Thinking



Hacker Ethic



Scientific Norms

Hacker Ethic

- "Hacking" as a culture of subverting technical systems not cyber-crime
- Common values within hacker ethic **[Levy 1984]**
  - Sharing & openness – access to computers/software/data/information should be free and open
  - Autonomy & meritocracy – hackers should be free to pursue individual goals and judged by their work
  - Creativity & playfulness – novelty and efficiency; irreverence and references to geek culture
  - Decentralization & anti-authoritarianism – build technologies to resist bureaucracy and support freedom
- Exploring then exploiting opportunities by gaining then using know-how
- Moving from curiosity and observation to execution and efficiency
- These values emerged from a culture that remains overwhelmingly white and male **[Coleman 2012]**

Levy, S. (1984). *Hackers: Heroes of the computer revolution*; Coleman, E.G. (2012). *Coding freedom: The ethics and aesthetics of hacking*.

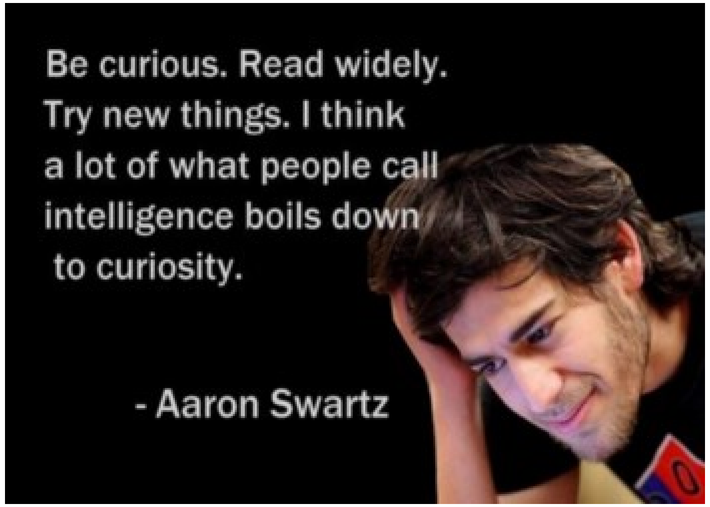
The Hacker Attitude

1. **The world is full of fascinating problems waiting to be solved.**
   - “to be a hacker you have to get a basic thrill from solving problems, sharpening your skills, and exercising your intelligence.”
2. **No problem should ever have to be solved twice.**
   - “it's almost a moral duty for you to share information, solve problems and then give the solutions away just so other hackers can solve *new* problems instead of having to perpetually re-address old ones.”
3. **Boredom and drudgery are evil.**
   - “Hackers should never be bored or have to drudge at stupid repetitive work, because when this happens it means they aren't doing what only they can do — solve new problems.”
4. **Freedom is good.**
   - “[Authoritarians] distrust voluntary cooperation and information-sharing. [Hackers] have to develop an instinctive hostility to censorship, secrecy, and the use of force or deception to compel responsible adults.“
5. **Attitude is no substitute for competence.**
   - “Hackers won't let posers waste their time, but they worship competence... at demanding skills that few can master and skills that involve mental acuteness, craft, and concentration is best.”

Raymond, E.S. (2001). How To Become A Hacker.

Ten Commandments of Hacker Culture

1. **Give before you get**. Horizontal groups, judged on contributions, respect is earned.
2. **Don’t ask for permission**. Do things without permission, then evaluate the outcome.
3. **Doing > Talking**. Doing creates new solutions and is sign of respect.
4. **No excuses**. Hackers are responsible for their actions.
5. **Solve problems**. Use creativity to solve problems in new ways.
6. **Follow your curiosity**. Hunger for learning and knowing more about topics.
    1. **Failing is growing**. Grow from overcoming adversity and uncertainty.
7. **Know tools/communities**. Don’t reinvent the wheel.
8. **Always be learning**. Explore new ideas and experiences.
9. **Get involved**. Fix things wrong in the world.
10. **Have fun**. Work on projects you care about.



Morato, M. (2015). The Ten Commandments of Hacker Culture.

A Modest Proposal

Growth Mindset

Computational Thinking

Hacker Ethic

Scientific Norms

Scientific Norms

- Merton (1942) proposed four norms that guided scientific culture (CUDOS)
  - Communalism: common ownership of scientific results and methods, duty to share
  - Universalism: scientific work should be evaluated objectively and impersonally
  - Disinterestedness: scientific work should be altruistic and avoid self-enrichment and promotion
  - Organized skepticism: transparency of results, consideration of all evidence, scrutiny by peer review
- Widely-criticized and many counter-examples exist, but they remain influential with new norms
  - Governance: scientists are responsible for direction, control, and self-regulation of science
  - Quality: scientists judge contributions on the basic of quality not quantity
  - Calling: science is a higher purpose worthy of sacrificing material benefits
  - Breadth: scientists fulfill responsibilities like teaching and service in addition to research

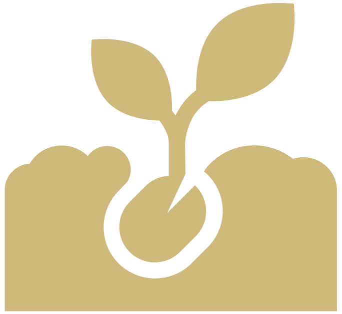
Merton, R. K. (1942). A note on science and democracy. *J. Legal & Pol. Soc.;* Anderson, M. S., et al. (2010). Extending the Mertonian Norms. *Journal of Higher Education*.

Science as a Profession

- Understanding issues of ethics, social responsibility, and awareness of regulations and policies
  - Privacy, property, accessibility, security, human rights
- Communicating results of analyses back to both expert and general audiences
  - Visualization, evaluation, explanation, trust, persuasion, storytelling
- Collaborating to manage complexity and accomplish larger goals
  - Teamwork, brokerage, delegation, translation, infrastructuring
- Management of people and projects
  - Planning, documentation, leadership, conflict resolution, mentorship, consciousness

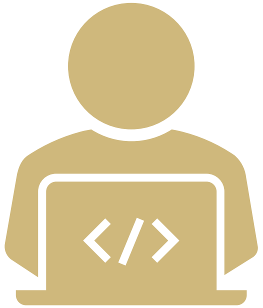
Hall, K.L., et al. (2018). The science of team science. *American Psychologist*; National Academies. (2017). Communicating Science Effectively.

Data Science Mindset Components

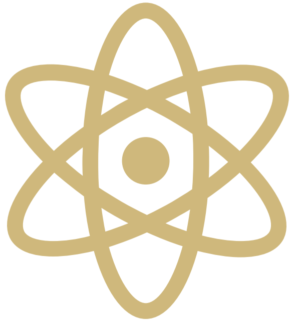


Growth Mindset

Computational Thinking



Hacker Ethic



Scientific Norms

- Growth Mindset: effort not ability, continual improvement, challenge-seeking, resilience
- Computational Thinking: concepts, practices, perspectives of applying computing technologies
- Hacker Ethic: sharing, openness, creativity, autonomy, curiosity, bias towards action
- Scientific Norms: communalism, skepticism, responsibility, communication, collaboration

Setting up computing environment

Next class

Next class

- Readings
  - Borgatti, S. P., Mehra, A., Brass, D. J., and Labianca, G. (2009). Network analysis in the social sciences. *Science*, 323(5916):892–895
  - Butts, C. T. (2009). Revisiting the foundations of network analysis. Science, 325(5939):414–416
  - Brandes, U., Robins, G., McCranie, A., and Wasserman, S. (2013). What is network science? *Network Science*, 1(1).
- Post a 500-ish word discussion on Canvas summarizing your previous exposure to and current interests in network analysis and theories and your learning goals for our class
- On Thursday, we'll discuss readings, themes from discussions, debug any issues with configuring computing environment