Applications of Network Science

Social Networks Analysis and Graph Algorithms

Prof. Carlos "ChaTo" Castillo — https://chato.cl/teach

upf. Universitat Pompeu Fabra Barcelona

Sources

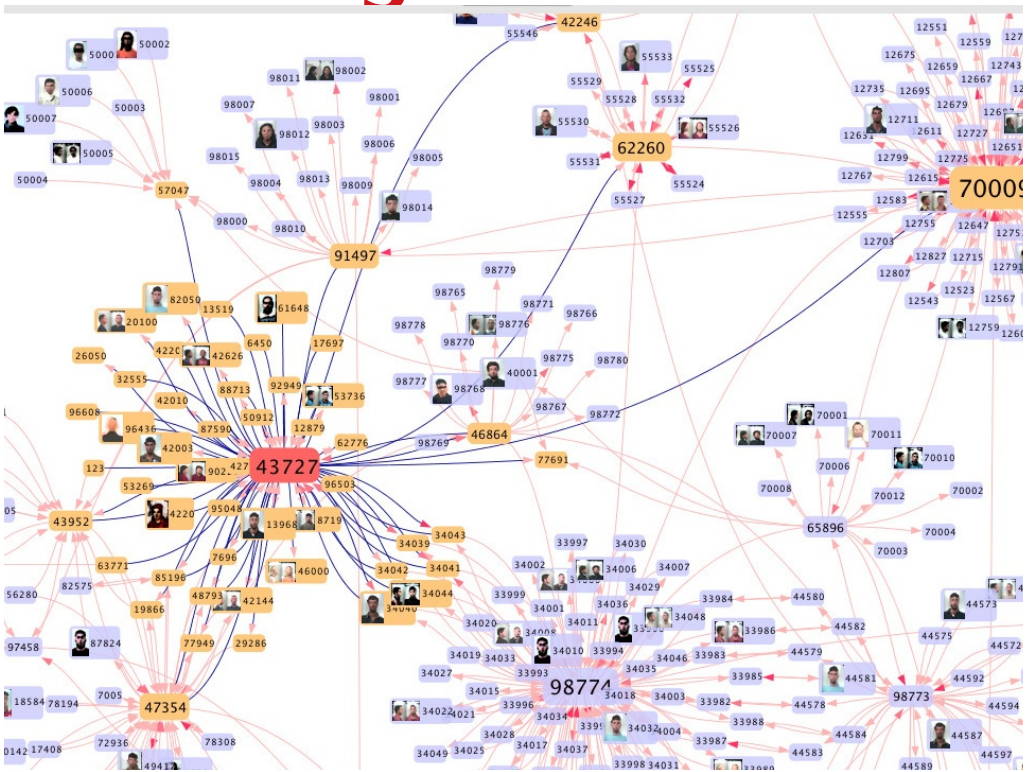
- A. L. Barabási (2016). Network Science – Chapter 01and Chapter 02
- F. Menczer, S. Fortunato, C. A. Davis (2020). A First Course in Network Science – Chapter 00
- URLs cited in the footer of specific slides

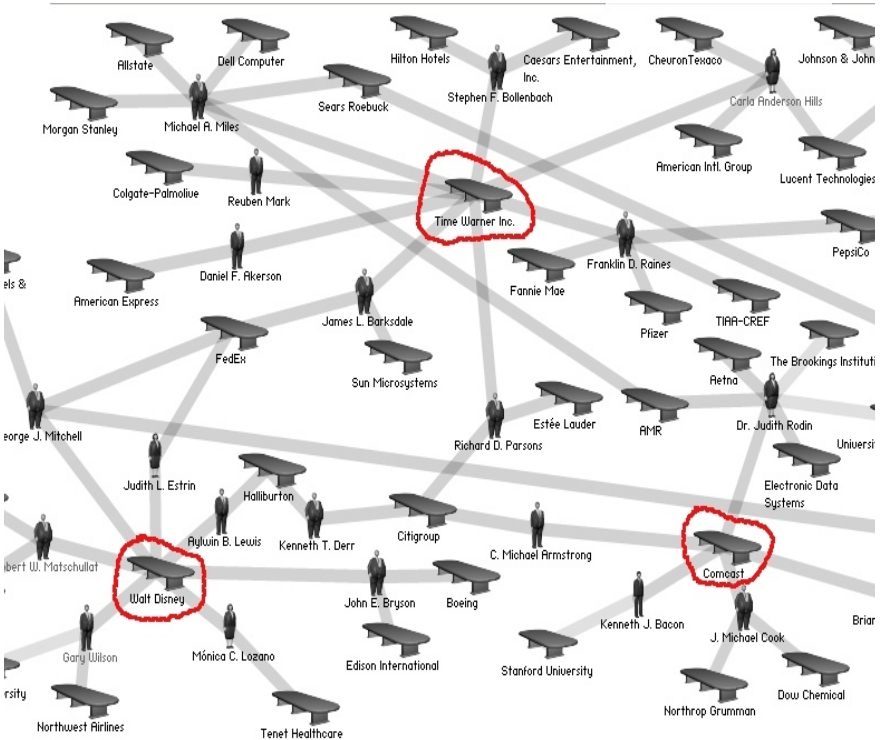
Networks Science

- **Interdisciplinary**; indeed we often address problems from disciplines other than CS
- **Empirical** and data-driven; it is based on the observation of networks
- Quantitative, mathematical, **computational**

Help fight

organized crime and collusion





https://itnews.iu.edu/articles/2014/complex-networks-researcher-at-iu-fighting-crime-with-mobile-phone-data.php

https://en.wikipedia.org/wiki/File:Media_corporation_interlocks_-_2004.jpg

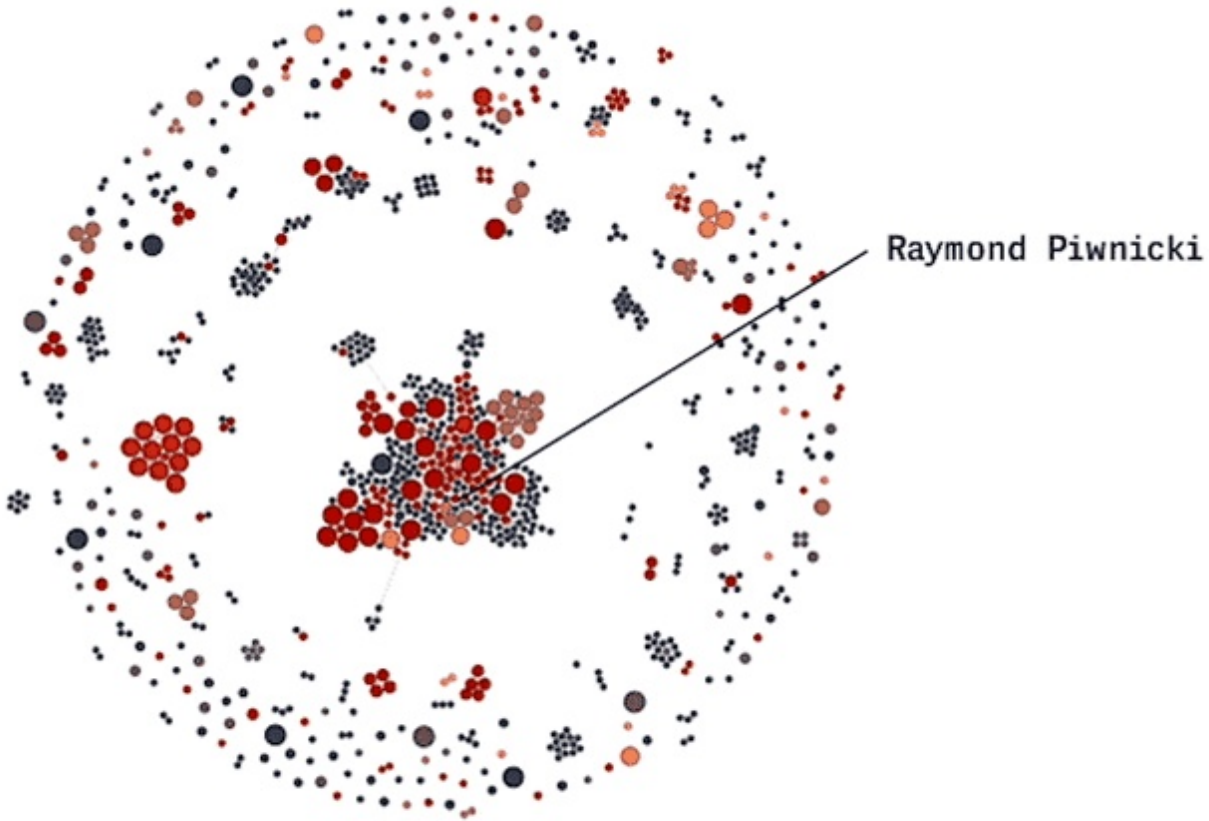
Help fight corruption

The Intercept_

BAD CHICAGO COPS SPREAD THEIR MISCONDUCT LIKE A DISEASE

Rob Arthur

August 16 2018, 3:03 p.m.

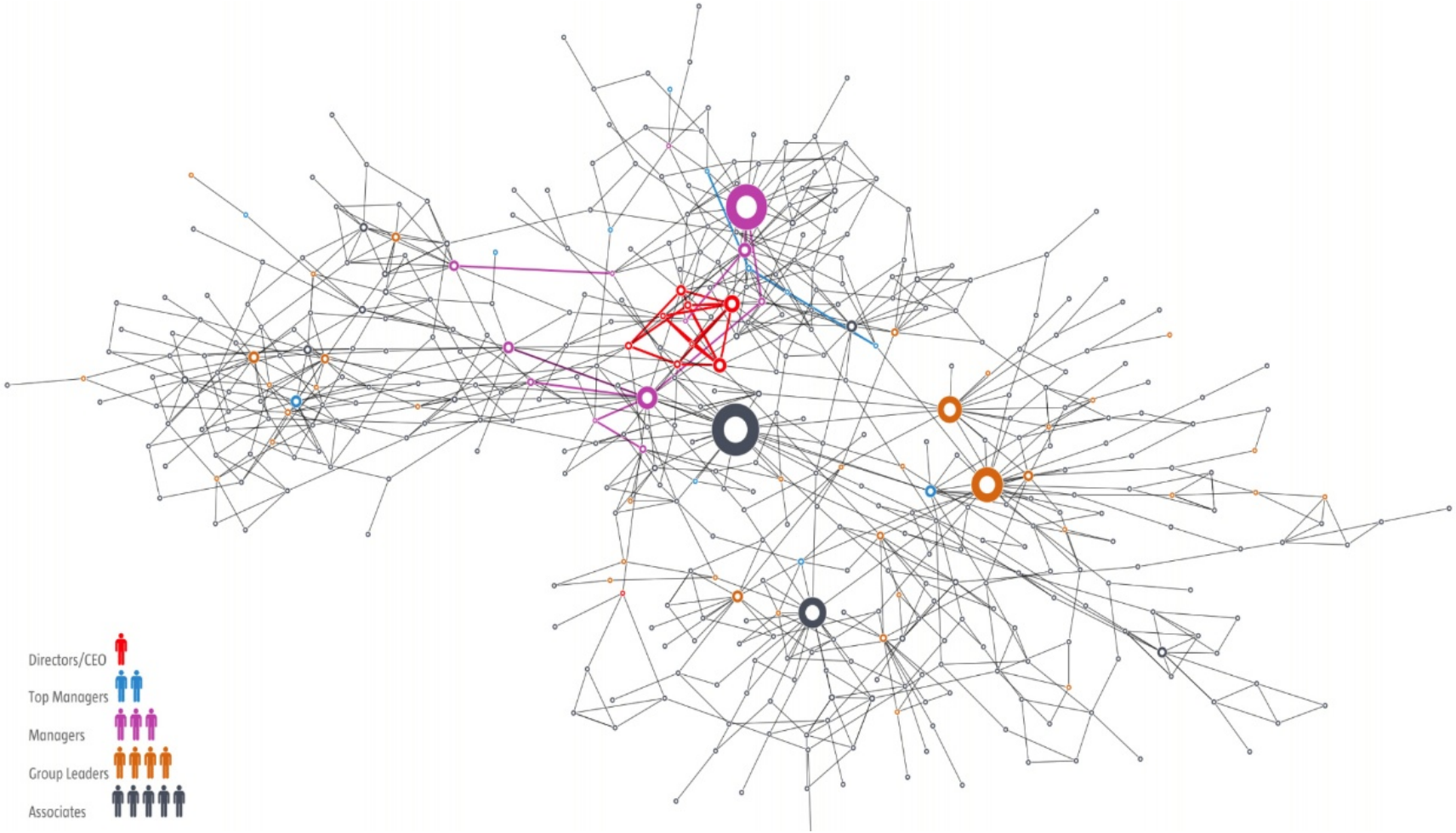


https://theintercept.com/2018/08/16/chicago-police-misconduct-social-network/

Help to forecast epidemics

https://www.youtube.com/watch?v=mm2u9RKwgsY

Help understand organization structures



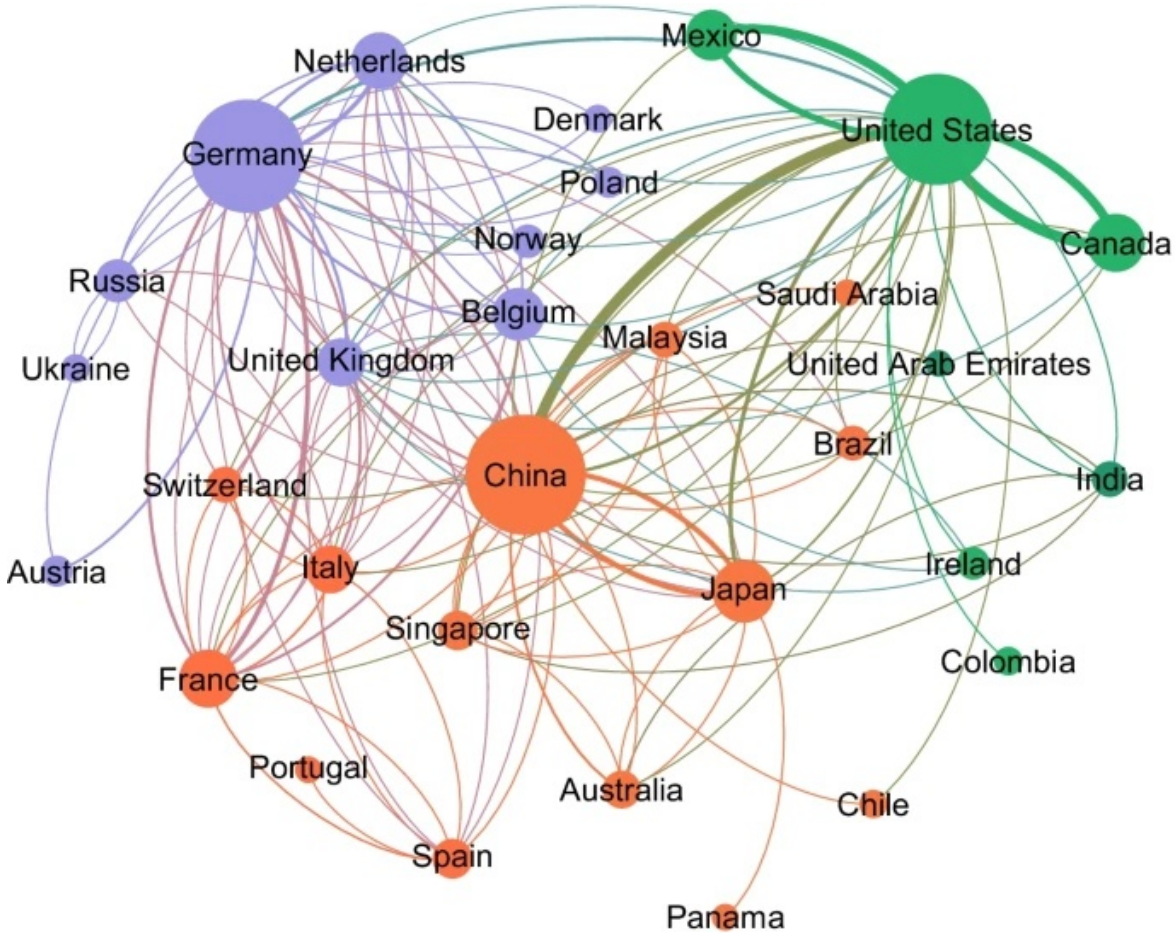
Help improve the communications of an organization

- About 3M e-mails sent or received by an EU research organization address
- Nodes are e-mail addresses (~1K internal, ~250K external)
- Edges are e-mails

https://www.youtube.com/watch?v=4JS-30dglqg

Help to understand international trade

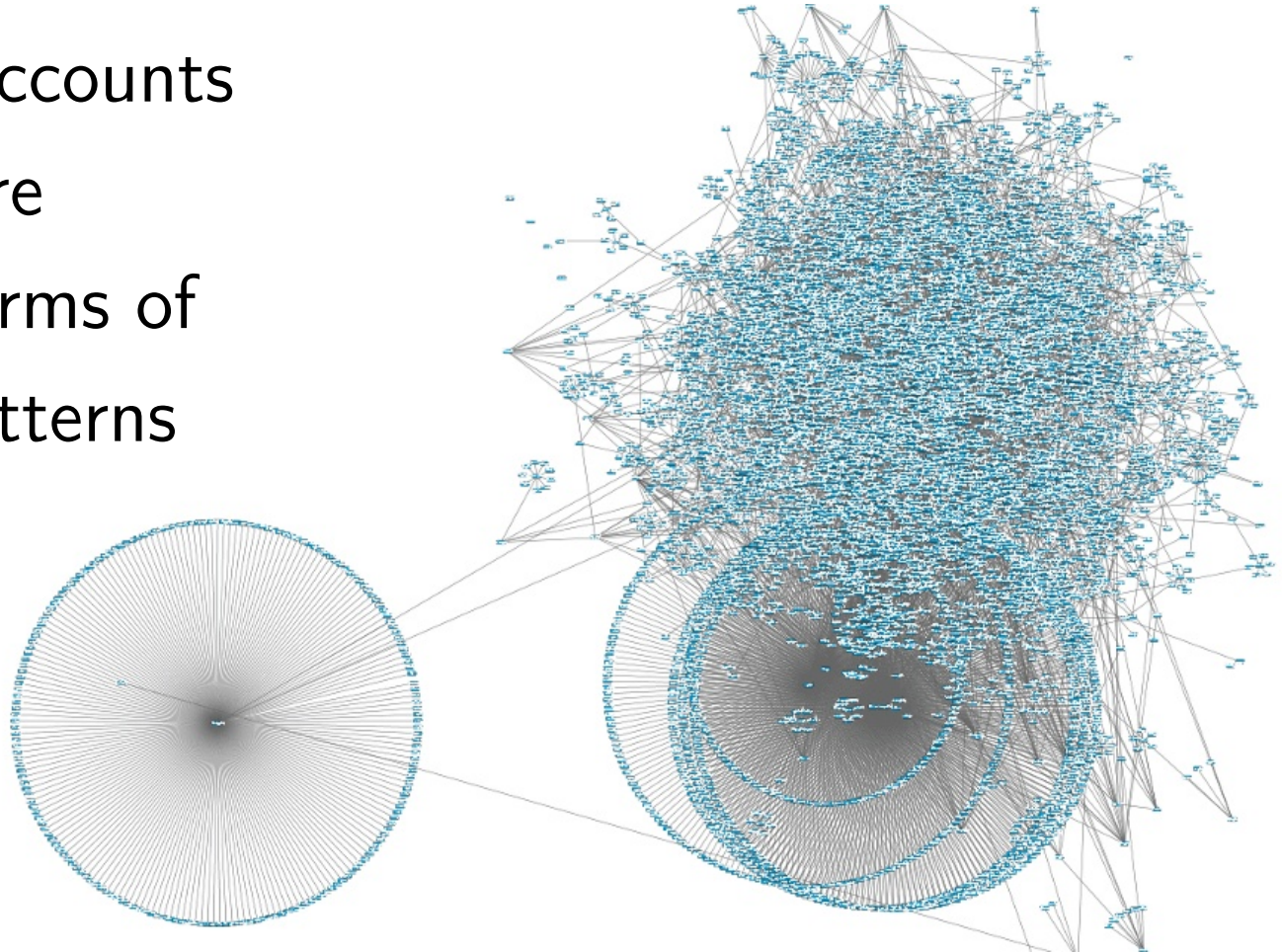
Multiple structural, economic, geographical, and political factors affect the global trade network structure.



https://doi.org/10.1007/s41109-022-00479-7

Fight misinformation and hate online

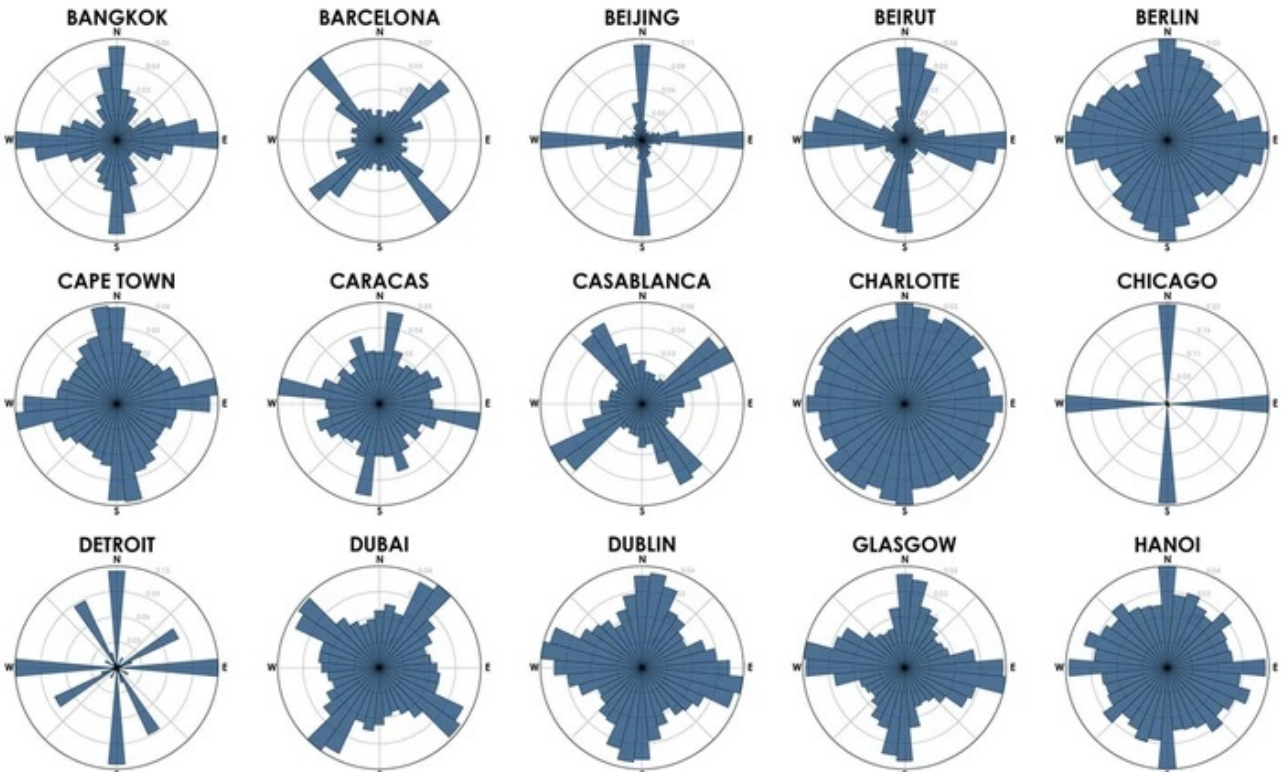
Inauthetentic accounts
a.k.a. "**bots**" are
anomalies in terms of
connectivity patterns



https://gigazine.net/gsc_news/en/20150310-graph-8000-fake-twitter-account/

Improve mobility within cities

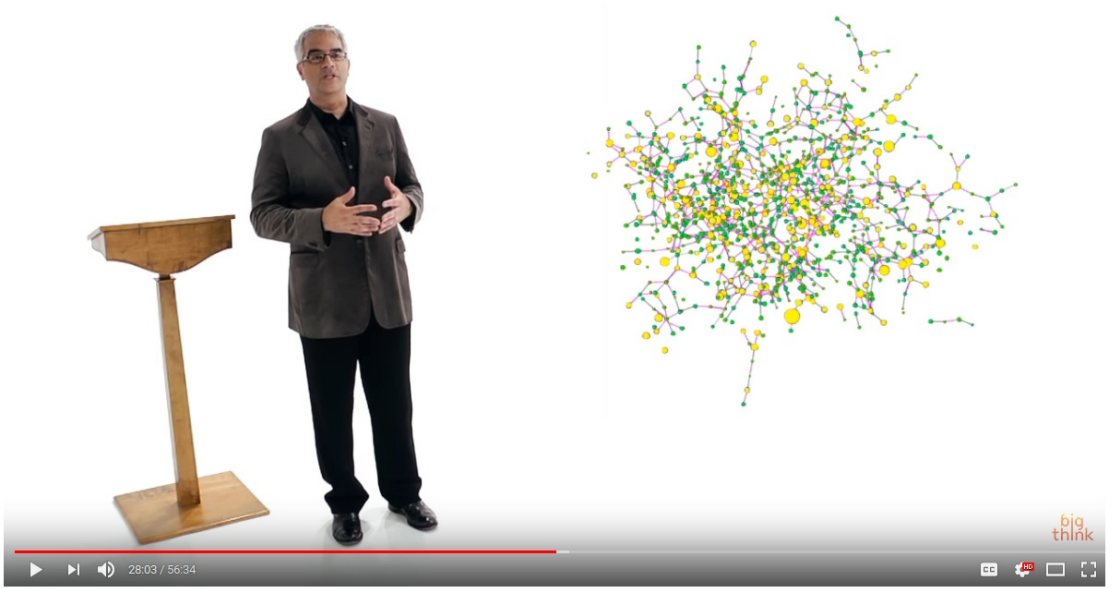
City grids have polarities that can be seen through networks analysis.



https://doi.org/10.1007/s41109-019-0189-1

Help understand society, diseases, and design new treatments and drugs

Must watch:
Nicholas Christakis (1 hour)



https://www.youtube.com/watch?v=wadBvDPeE4E

What we will learn

- To describe a network in formal terms
- To identify it as such and characterize it
- To visualize different networks
- To operate with networks programmatically
- To find important nodes and communities
- To make discoveries or help others make them
- **Much more (to a large extent, it's up to you!)**

How we will learn

- Theory sessions:
  - Help you understand how to model complex networks
  - Help you find important nodes, communities, and track influence
  - Do some simple (and not so simple) exercises to check that you understood correctly each concept, and to help you remember
- Practice sessions:
  - Help you work with complex networks
  - Manage and analyze graphs in Python
- **My focus is on what I think has value for you as a data scientist**

Summary

Things to remember

- Applications of complex networks analysis

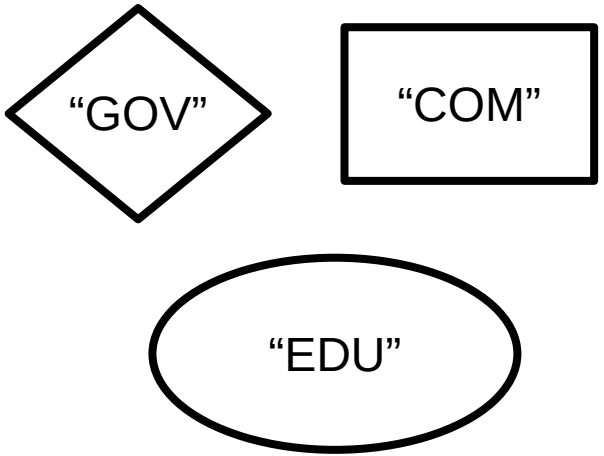
Additional contents (not included in exams)

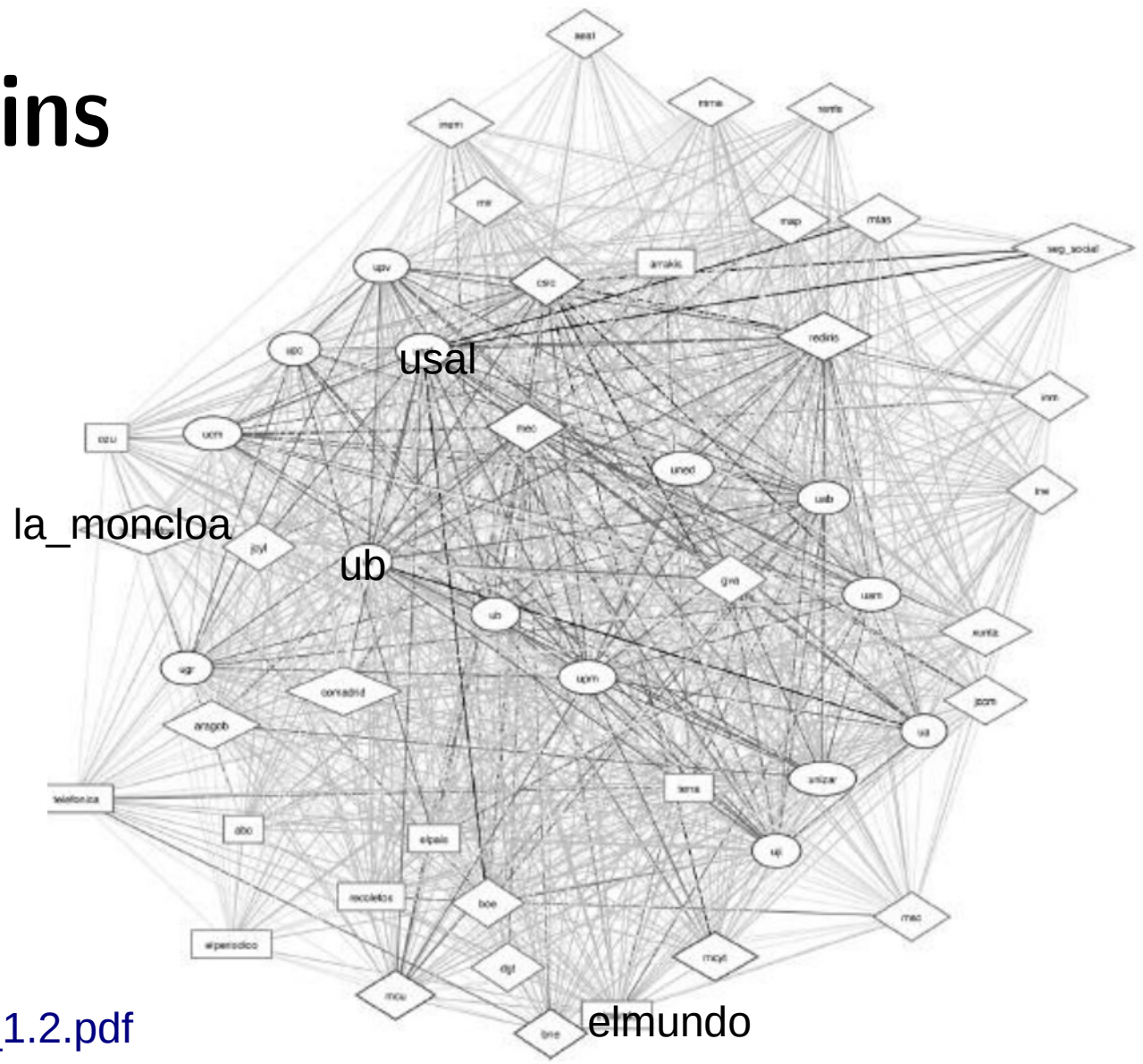
Why network science is important **to me**

PhD work (2000-2004)

- Collecting web pages
- Characterizing national web domains
  - Chile, Korea, Greece, Spain ...

Top .es domains (~2006)



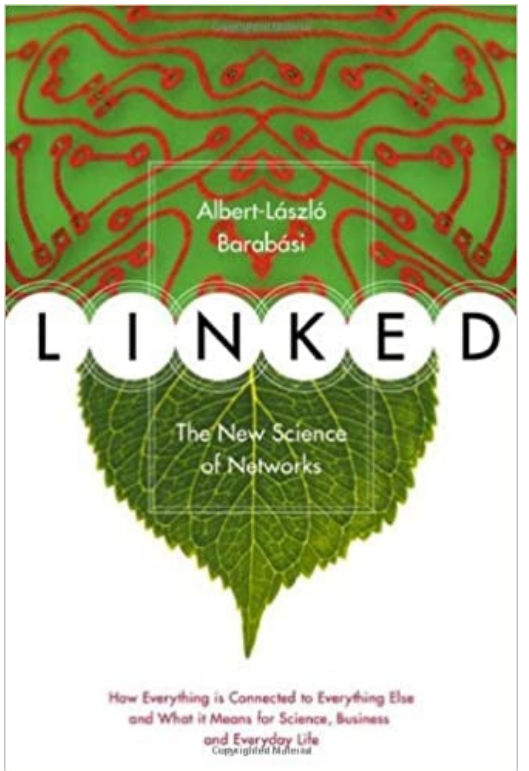


http://eprints.rclis.org/9292/1/vol15_1.2.pdf

An influential book (to me)

This book came out in 2002 and made me see networks everywhere; it's an easy read, written for the general public, highly recommended

Its author, Albert-László Barabási visited my university **in Chile** while I was a PhD student :-)



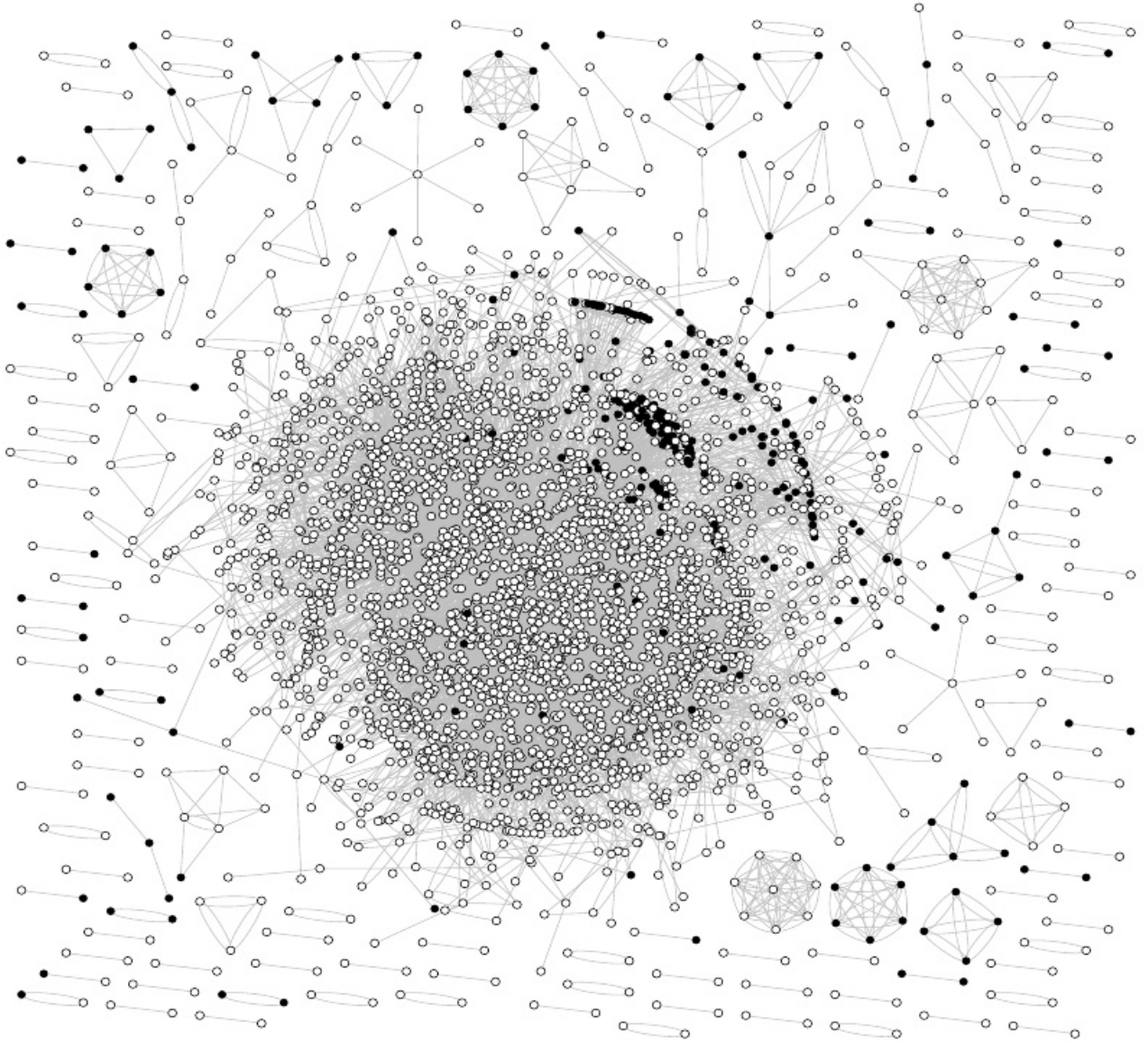
Early post-doctoral work (~2006-2009)

- Web spam pages
  - Pages created to deceive search engines
  - Attract traffic by stuffing themselves with keywords
  - Increase link score of other pages
  - Methods evolve all the time, how to catch them?

An Eureka! Moment

2006

Visualization of a web spam dataset using gnuplot; spam nodes (in black) cluster together!



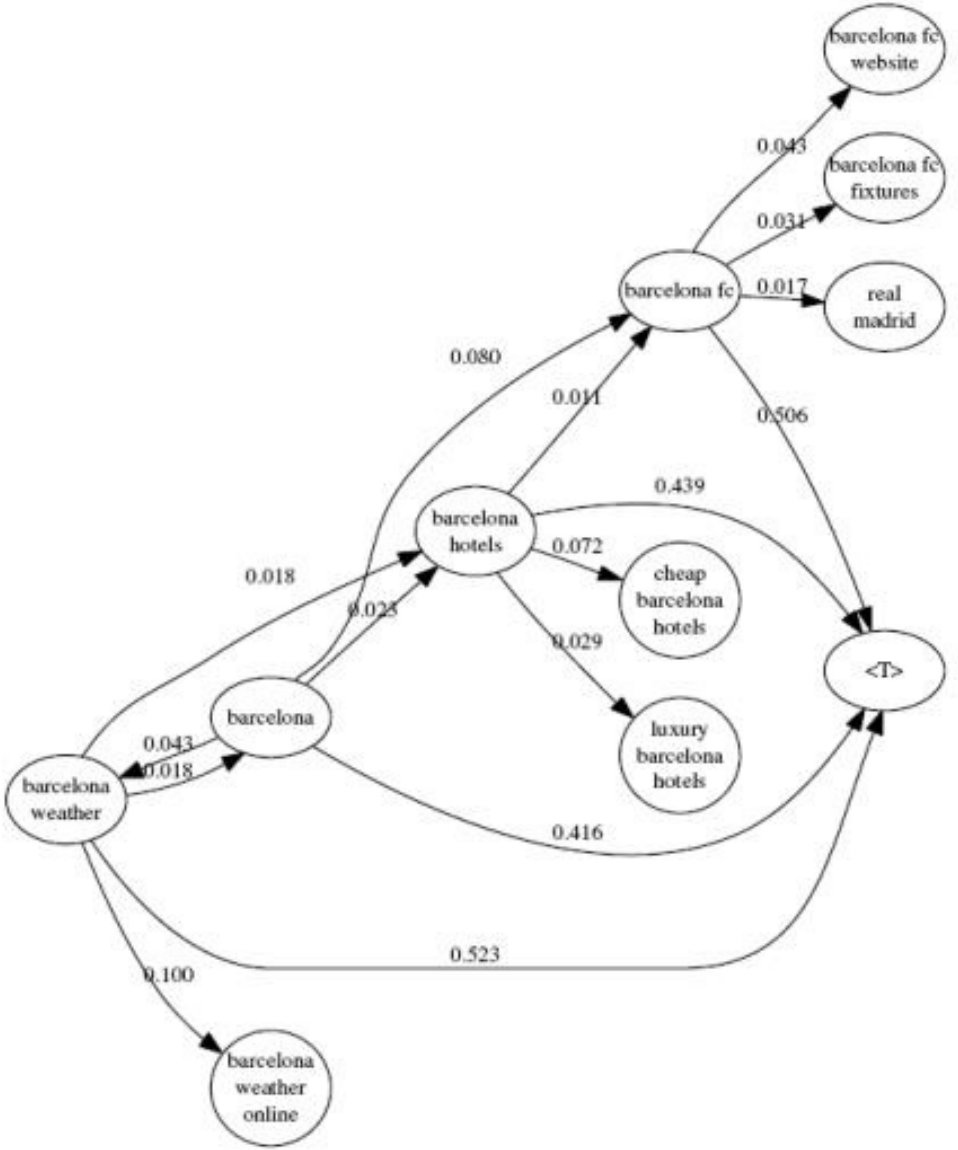
Paper: https://doi.org/10.1145/1277741.1277814

Query flows

2008

We wonder what is the most likely query before or after another query?

How are they connected? This is how we developed **query flow graphs**



Paper: https://doi.org/10.1145/1458082.1458163

Graphs in my own work

- Everywhere! — See https://chato.cl/research/
- Currently:
  - part of a larger toolbox
  - skeptical about structural-only conclusions