

Introduction to Network Analysis

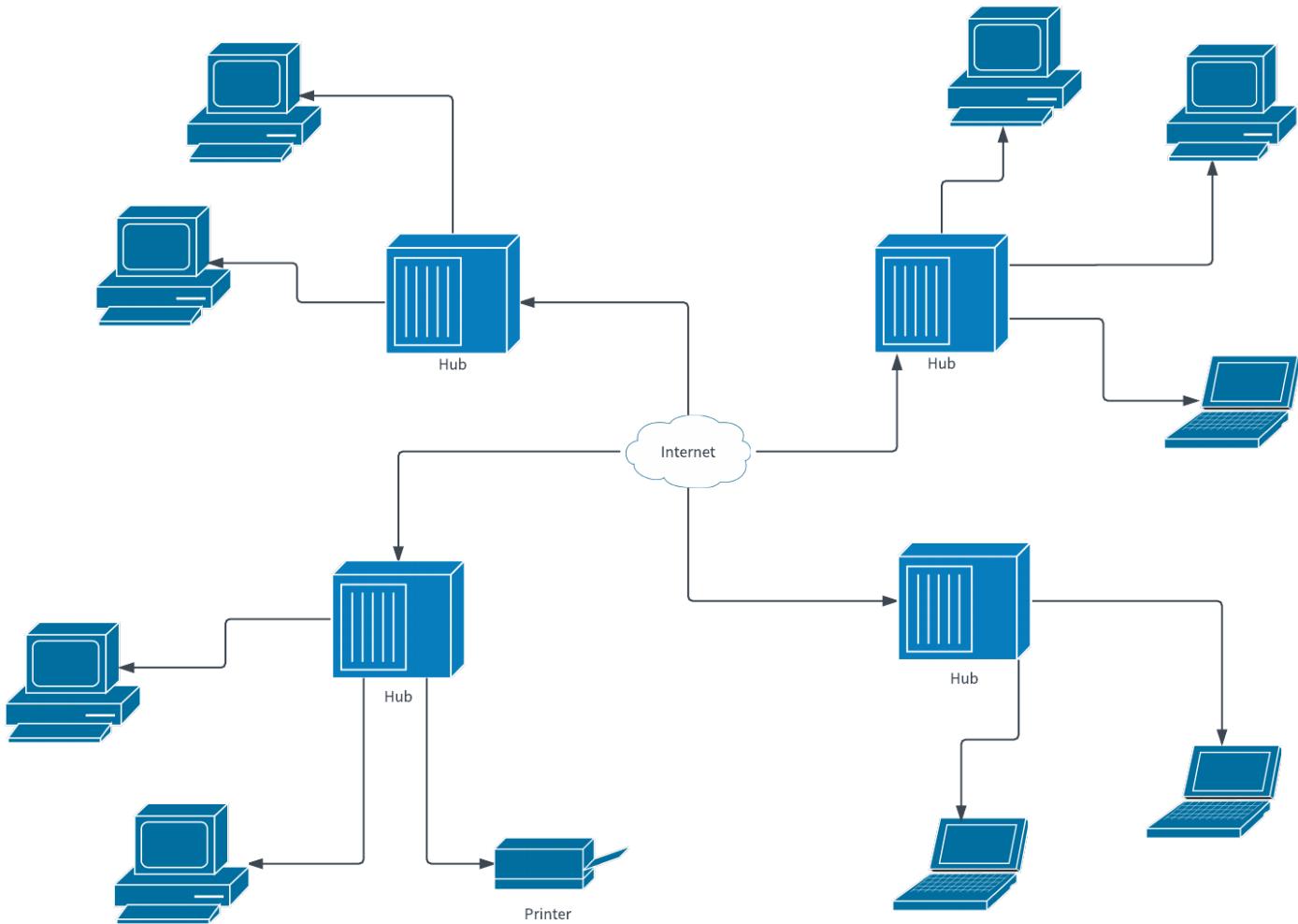
Amy Butler

AGENDA

- Recognize
- Visualize
- Analyze

RECOGNIZE





facebook

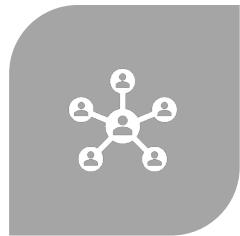
**Facebook helps you connect and share with
the people in your life.**



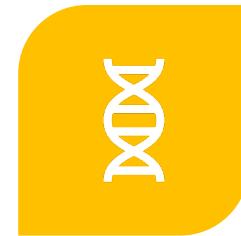
Major Classes of Networks¹



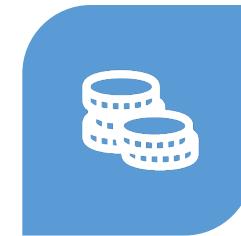
TECHNOLOGICAL



SOCIAL



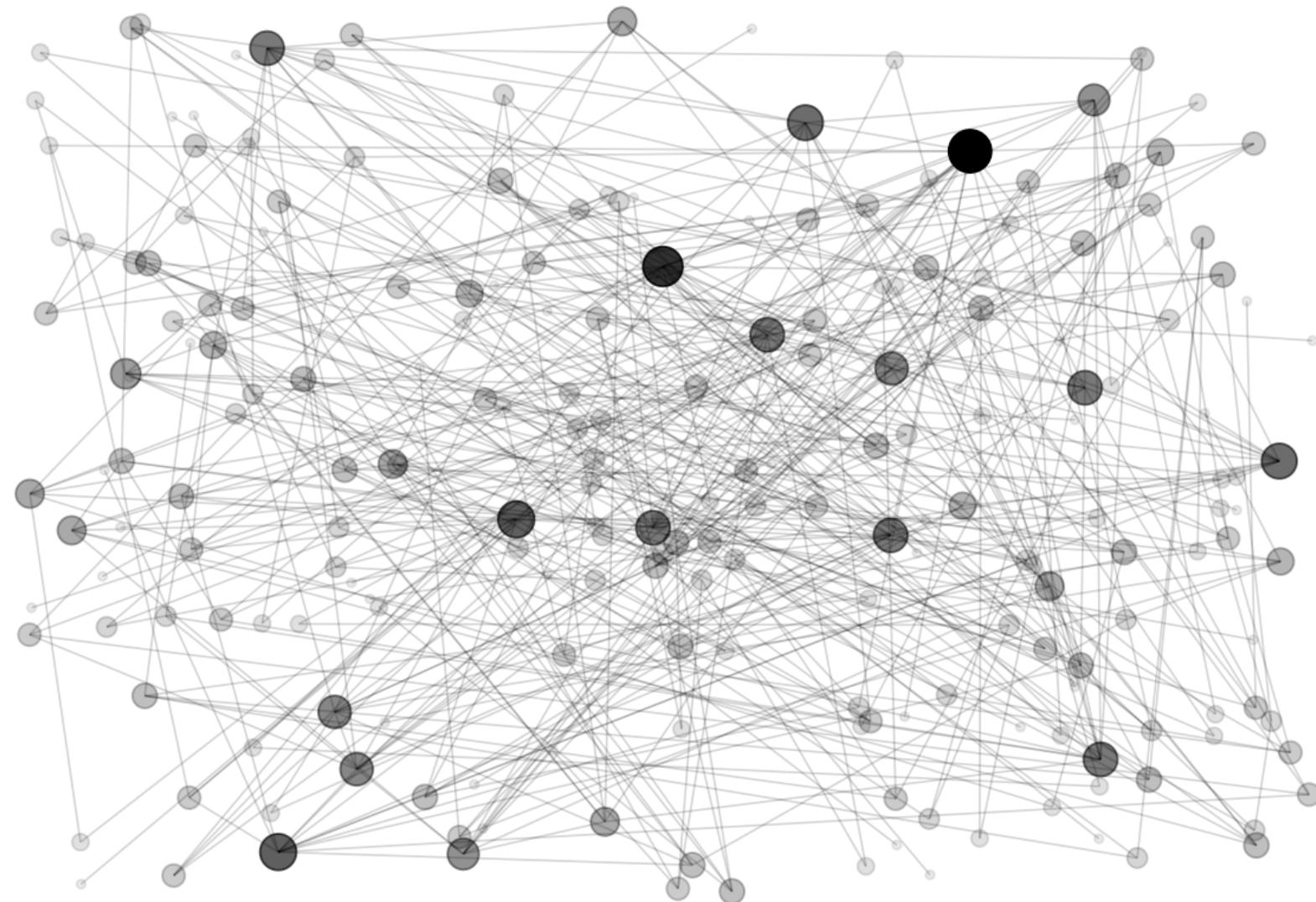
BIOLOGICAL



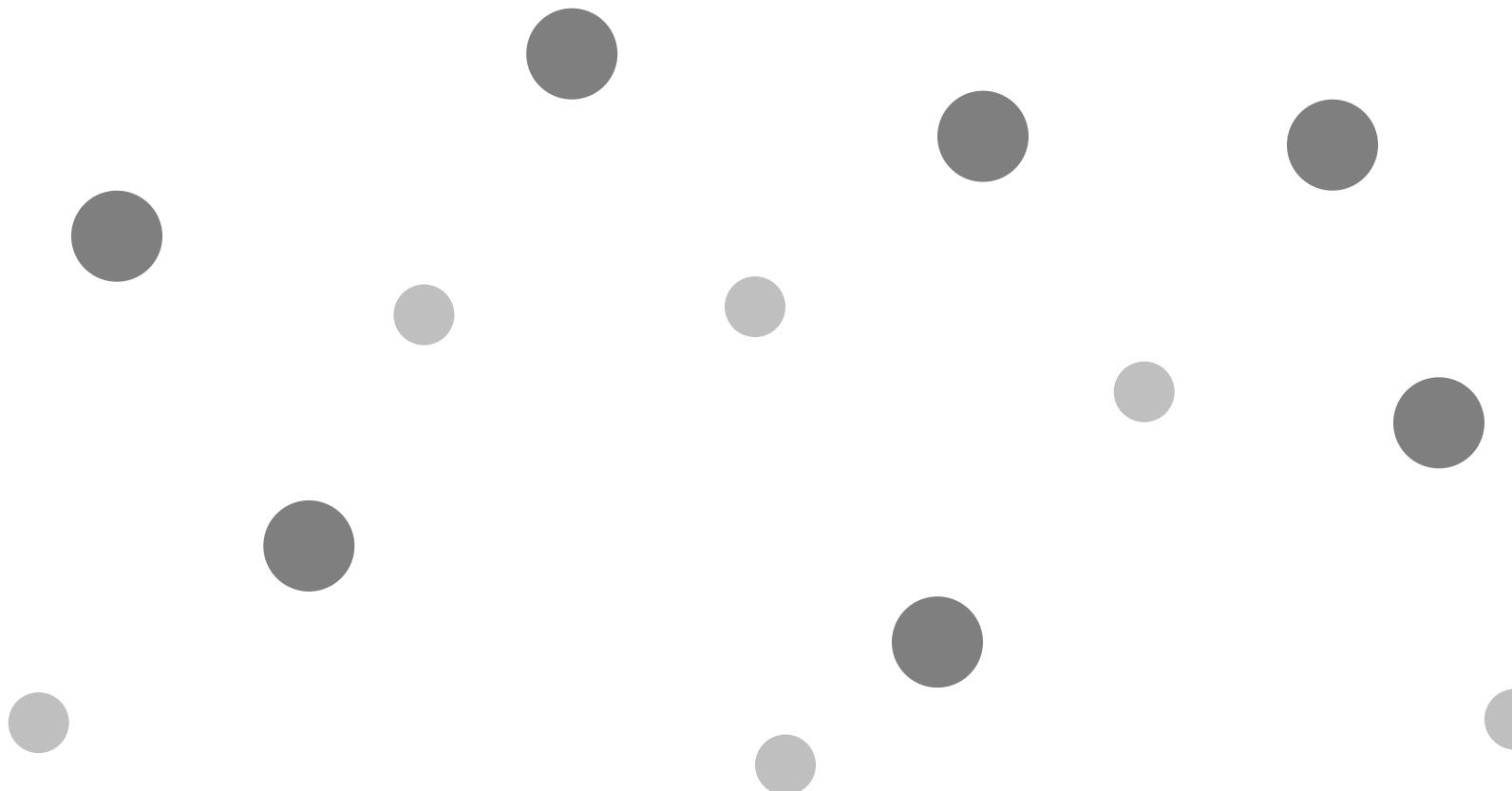
ECONOMIC



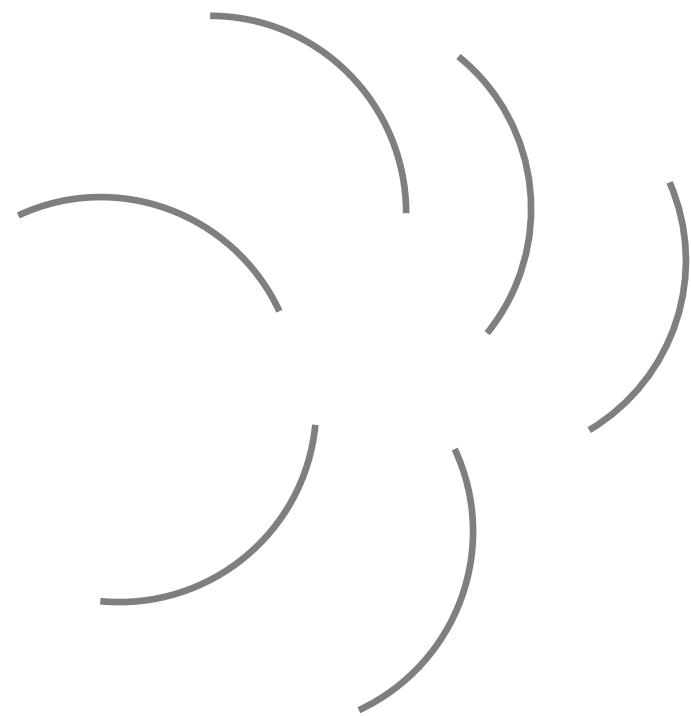
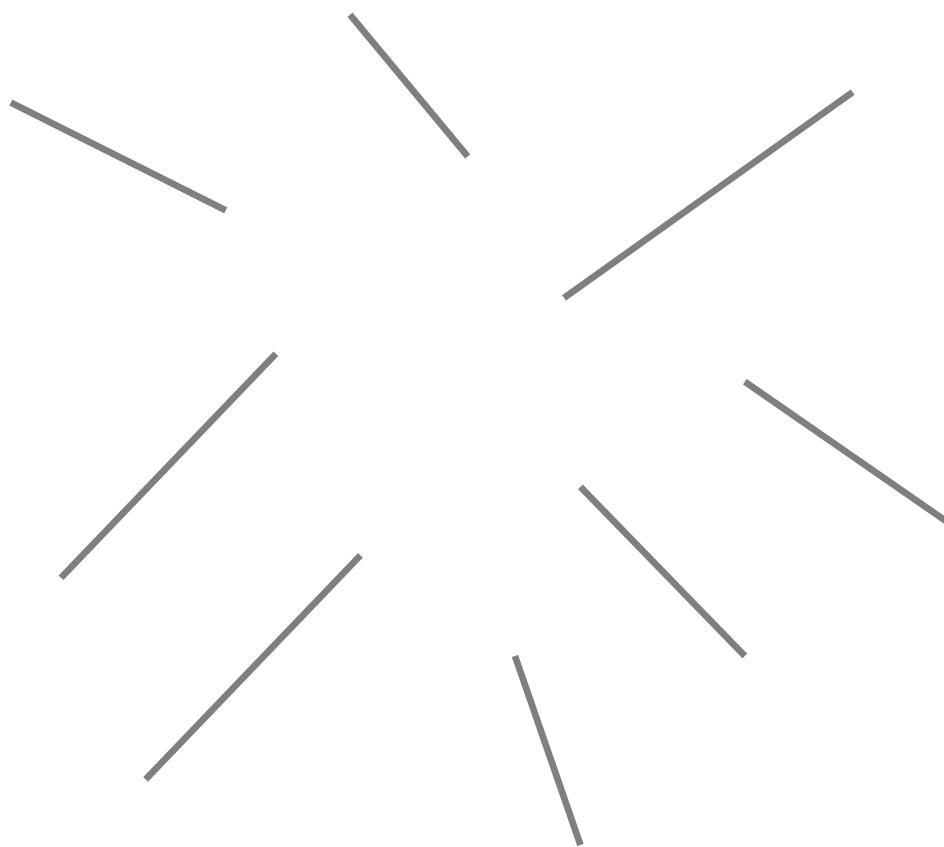
CULTURAL



These are NODES aka VERTICES



These are EDGES

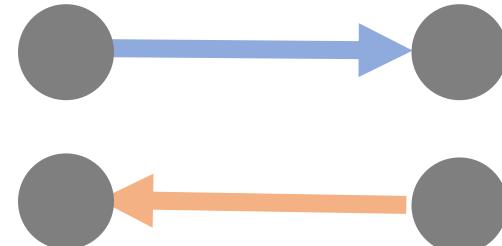


EDGES can have DIRECTION

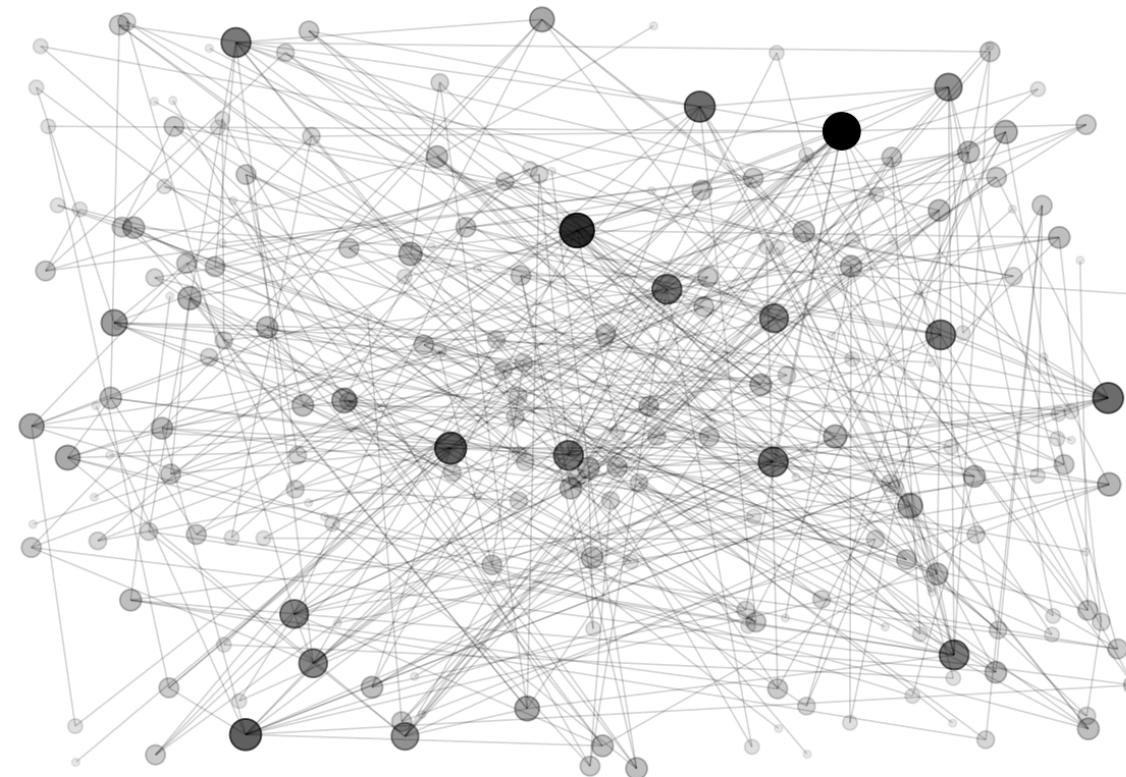
Undirected



Directed



This is an UNDIRECTED GRAPH

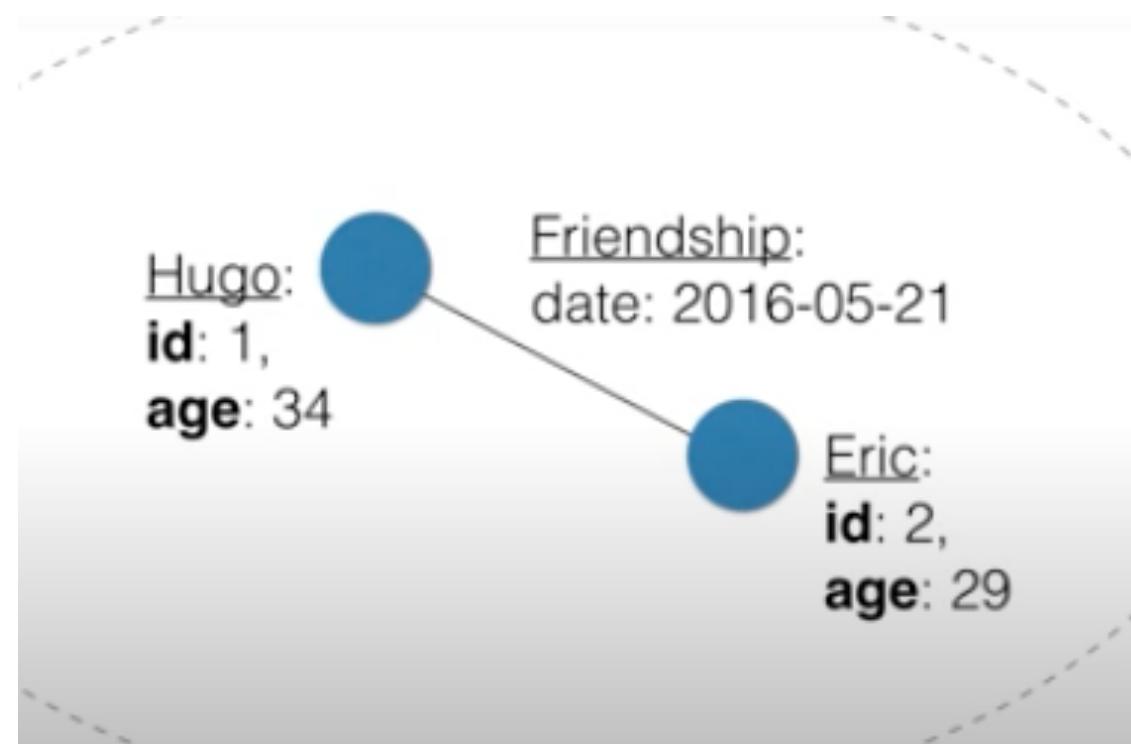


EDGES can have WEIGHT



ATTRIBUTES

- Nodes and edges can have metadata associated with them



VISUALIZE



Network Analysis in Python

- There are options for packages to visualize and analyze.

- iGraph
- Graph-tool
- NetworkX
- NetworKit



NetworkX

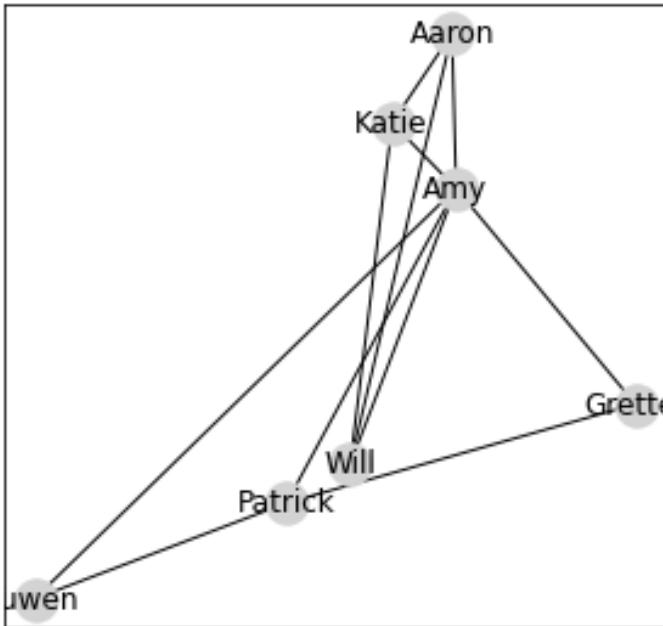
```
import networkx as nx
```

NetworkX

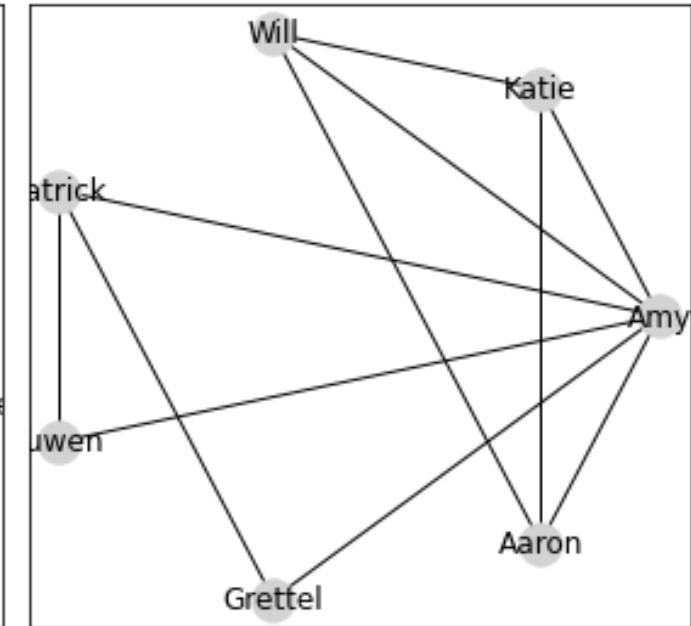
Matplotlib

Layout Algorithms

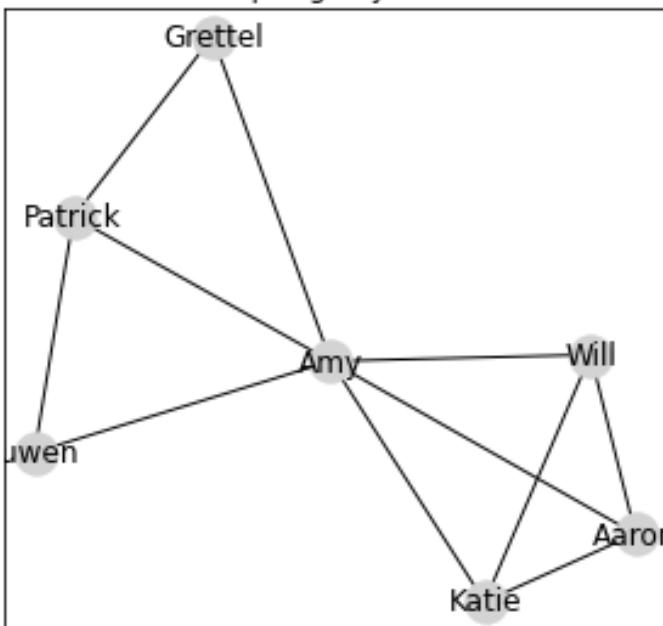
Random Layout



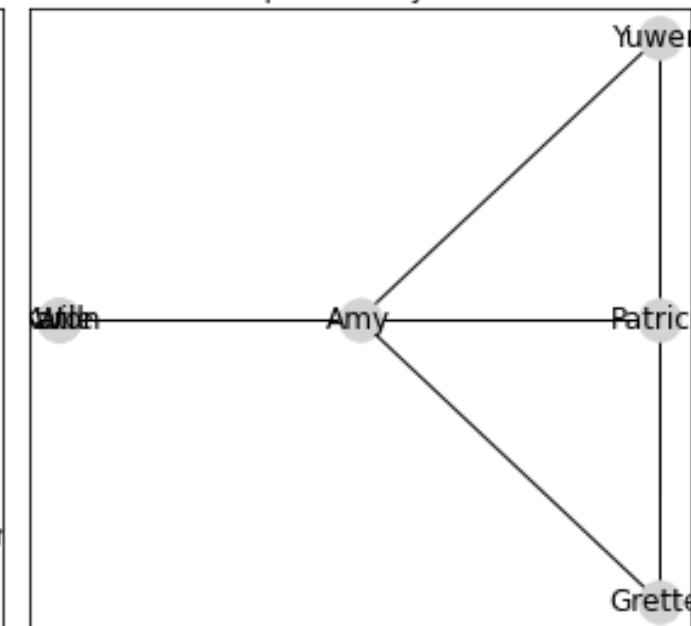
Circular Layout



Spring Layout



Spectral Layout



Eurovision 2018 Final Votes



<https://towardsdatascience.com/social-network-analysis-from-theory-to-applications-with-python-d12e9a34c2c7>

There are numerous Graph Visualization Tools!



[Download](#) [Blog](#) [Wiki](#) [Forum](#) [Support](#) [Bug tracker](#)[Home](#) [Features](#) [Learn](#) [Develop](#) [Plugins](#) [Services](#) [Consortium](#)

The Open Graph Viz Platform

Gephi is the leading visualization and exploration software for all kinds of graphs and networks. Gephi is open-source and free.

Runs on Windows, Mac OS X and Linux.

[Learn More on Gephi Platform »](#)

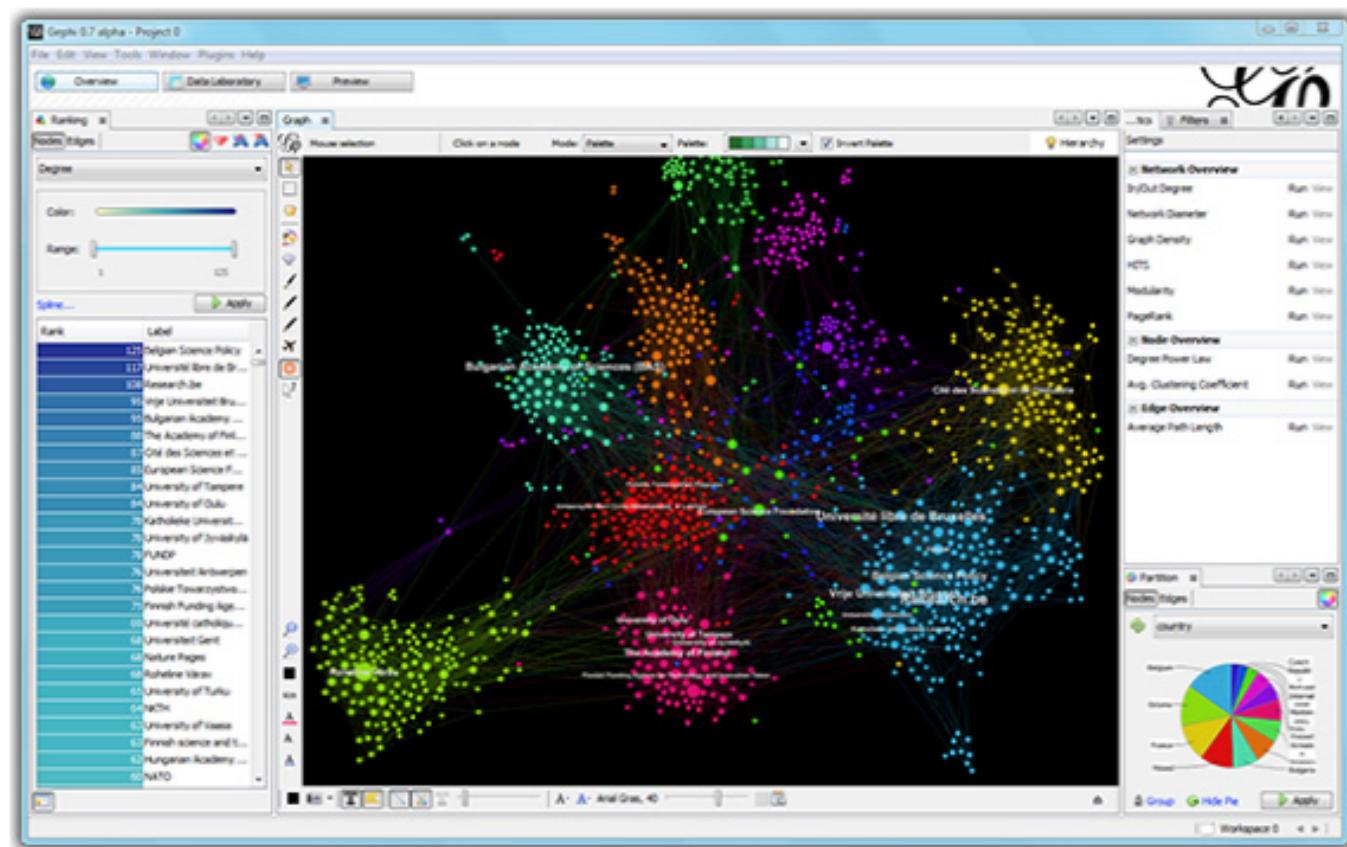


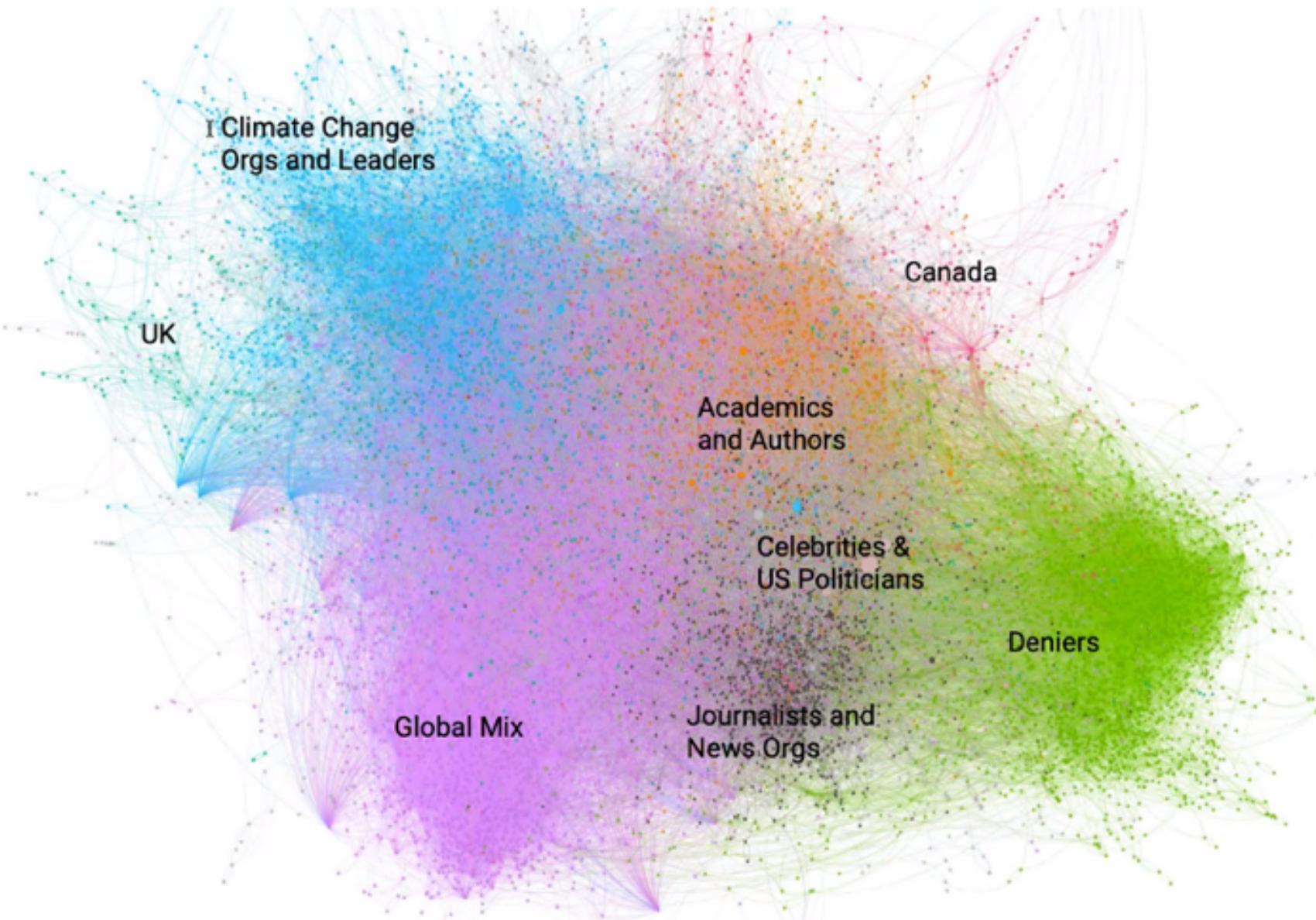
Download FREE
Gephi 0.9.2

[Release Notes](#) | [System Requirements](#)

► [Features](#)
► [Quick start](#)

► [Screenshots](#)
► [Videos](#)





https://zeromh.github.io/climate_change_conversations/

NetworkX to Gephi

- Export from NetworkX as a GraphML file to import in Gephi
- Gephi can import a spreadsheet, needs to be structured a certain way
- Blogs and videos available for more instructions on Gephi

ANALYZE



Jupyter Notebook

NetworkX starter

This notebook contains examples of creating and analyzing network graphs with the NetworkX package. The NetworkX package is typically imported as nx.

There are 4 types of network graphs.

1. **Undirected Graph** has only undirected edges. Create with nx.Graph()
2. **Directed Graph** has at least one directed edge. Create with nx.DiGraph()
3. **MultiGraph** is an undirected graph with multiple edges between the same nodes. Create with nx.MultiGraph()
4. **Directed MultiGraph** is a directed graph with parallel edges. Create with nx.MultiDiGraph()

```
In [2]: import networkx as nx  
import pandas as pd  
import matplotlib.pyplot as plt
```

Build the network

A network can be built by adding nodes and edges, one or several at a time. NetworkX can also take a list of tuples or a dictionary to create the graph.

What is the size of our network?

- How many nodes?
- How many edges?
- Example: *This network has 205 nodes and 180 edges.*

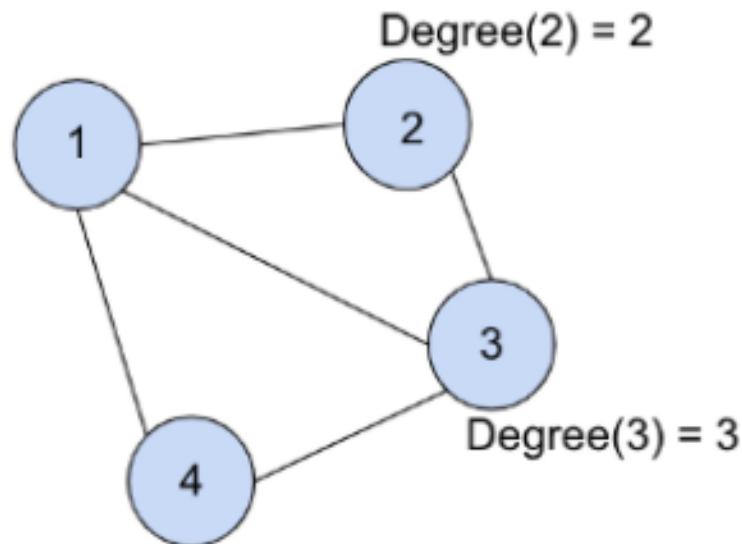
How many connections does each entity have?

(edges)

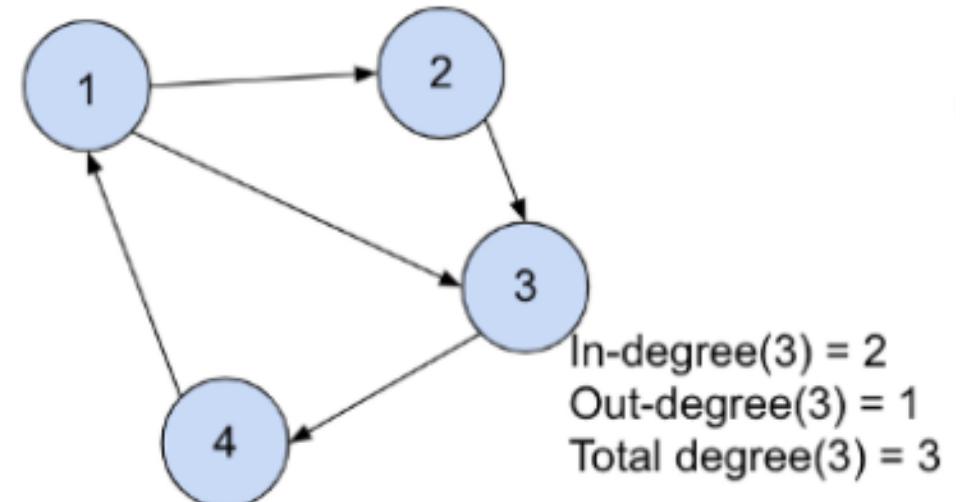
(node)

NODE DEGREE

Undirected Graph²



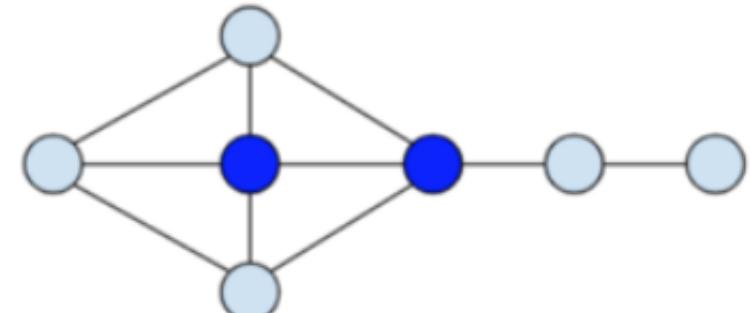
Directed Graph²



Who/What is most important/influential?²

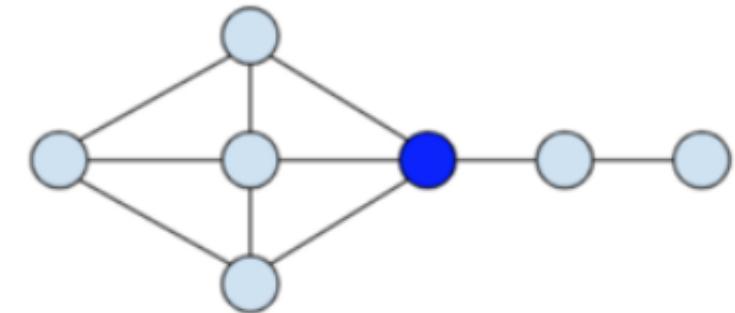
- **DEGREE CENTRALITY**

Important nodes have many connections



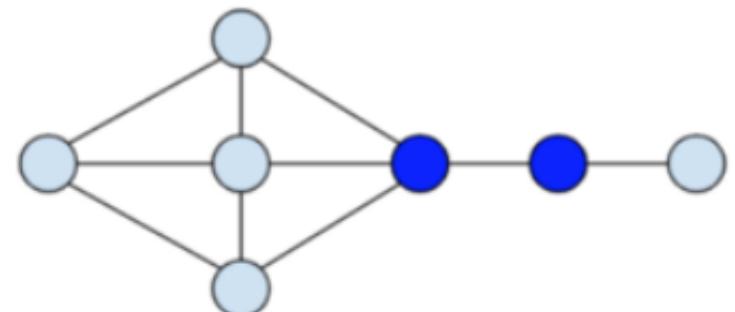
- **BETWEENNESS CENTRALITY**

Important nodes connect other nodes



- **CLOSENESS CENTRALITY**

Important nodes are close to other nodes



Who/What is most important/influential?

- **EIGENVECTOR CENTRALITY**

Important nodes have many connections to other important nodes

- **PAGE RANK**

A variant of the Eigenvector Centrality that is used to analyze directed networks. Important nodes have many in-coming edges.

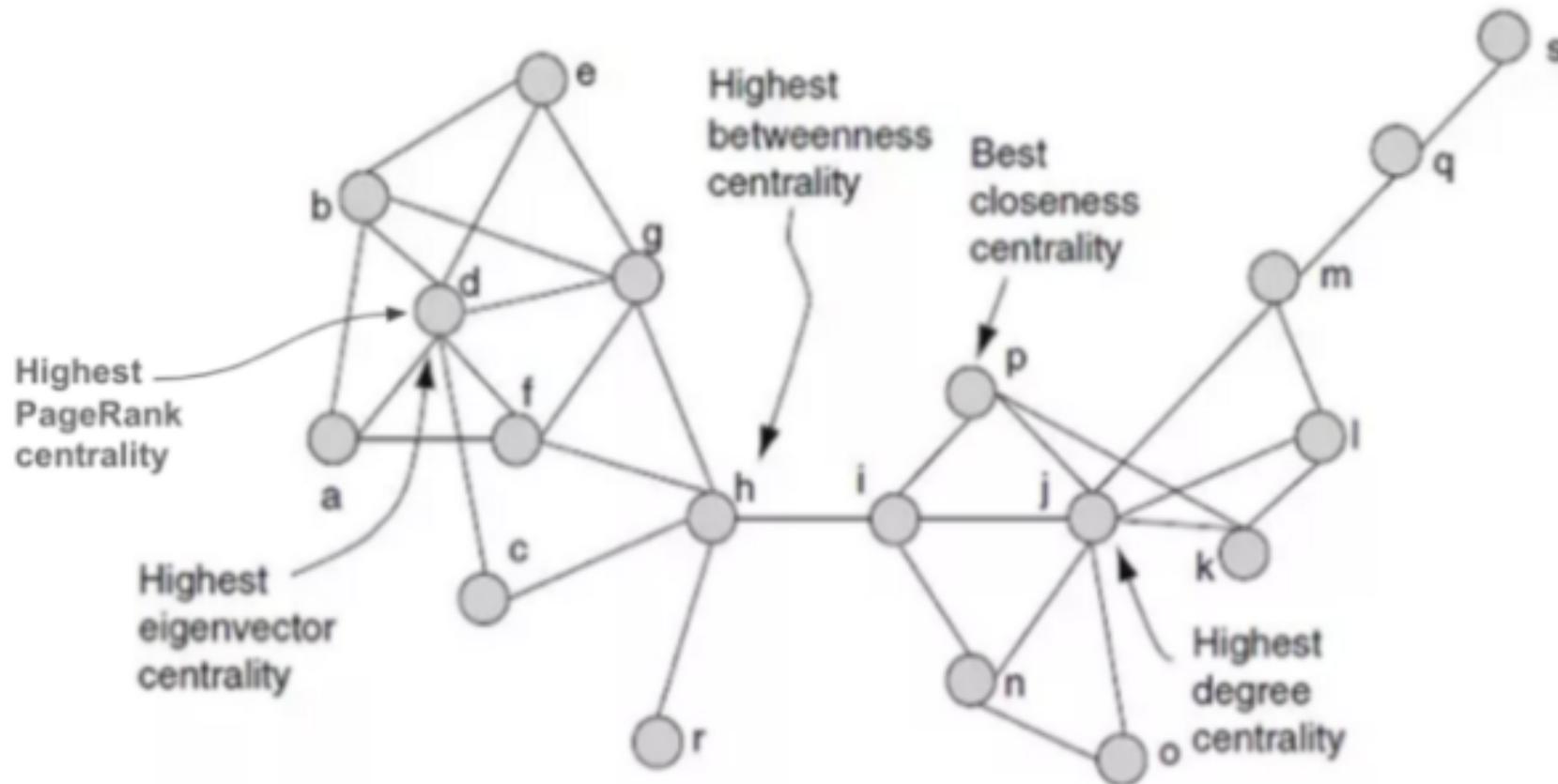


Illustration of various centrality measures. Source — [Arroyo, D. "Discovering Sets of Key Players in Social Networks."](#) Computational Social Network Analysis (2010)

SUMMARY



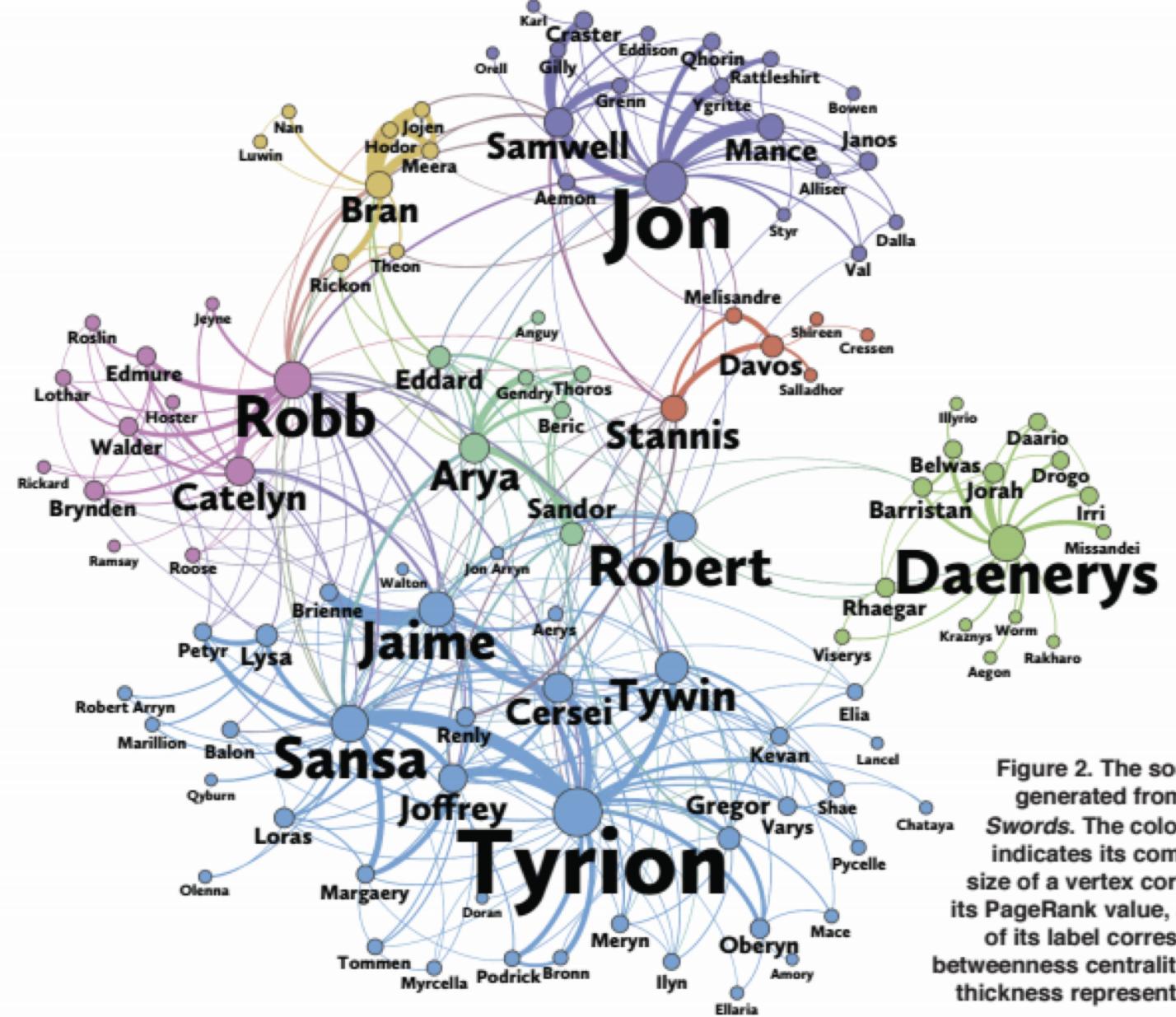
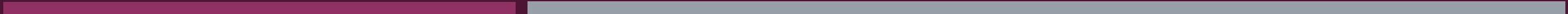


Figure 2. The social network generated from *A Storm of Swords*. The color of a vertex indicates its community. The size of a vertex corresponds to its PageRank value, and the size of its label corresponds to its betweenness centrality. An edge's thickness represents its weight.



THANK YOU!



Citations

1. *Complex Network Analysis in Python*, by Dmitry Zinoviev, p5
2. An Introduction to Network Analysis (part 1) | by Kristina Khvatova
<https://medium.com/@kristinakhvatova/an-introduction-to-network-analysis-part-1-fa2a92455e34>

Resources

- *Complex Network Analysis in Python*, by Dmitry Zinoviev
- *Data Science from Scratch, First Principles With Python*, by Joel Grus
- *Networks*, 2nd Edition, by Mark Newman