

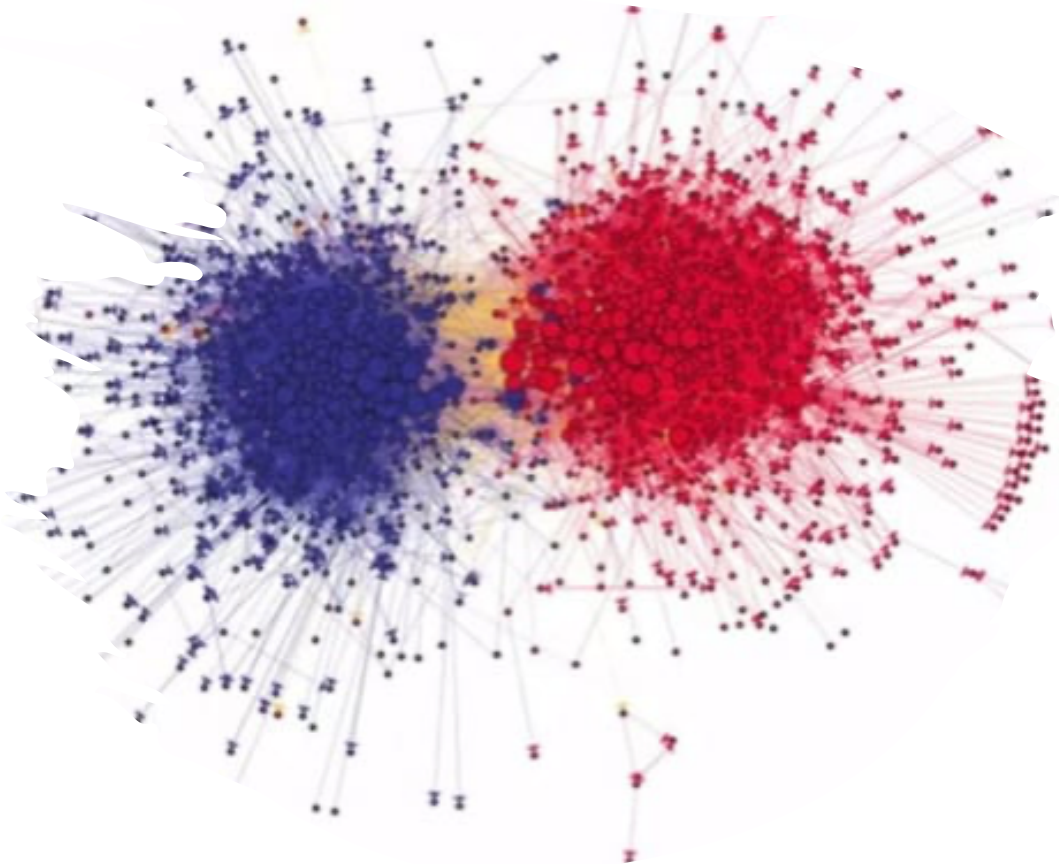
# Social Network Analysis

CLASS I: INTRODUCTION

# What is this module about?

- Social Network Analysis is a field of methodology that's becoming increasingly important in the social and political sciences.
- At a basic level, it's a set of methods that allow us to study networks – of many different varieties, not just online social networks – and discover insights about the human interactions which those networks represent.

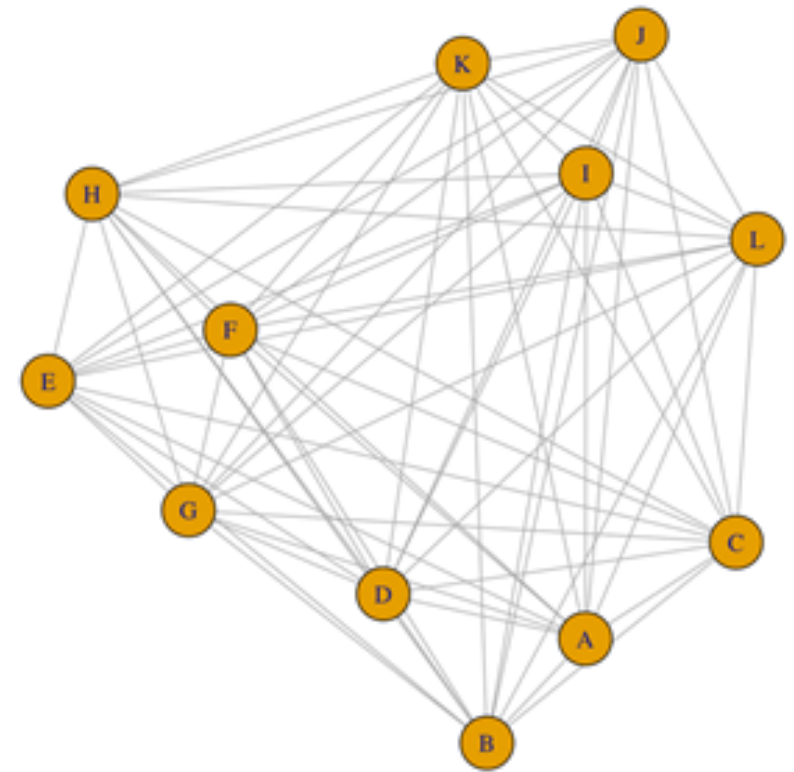
# Adamic & Glance 2004



- This network graph, showing the links between Republican- and Democrat-aligned political blog sites in the early 2000s, became very famous.
- It is now seen as an early warning sign of **polarisation** in online political discourse.

# What is a Network?

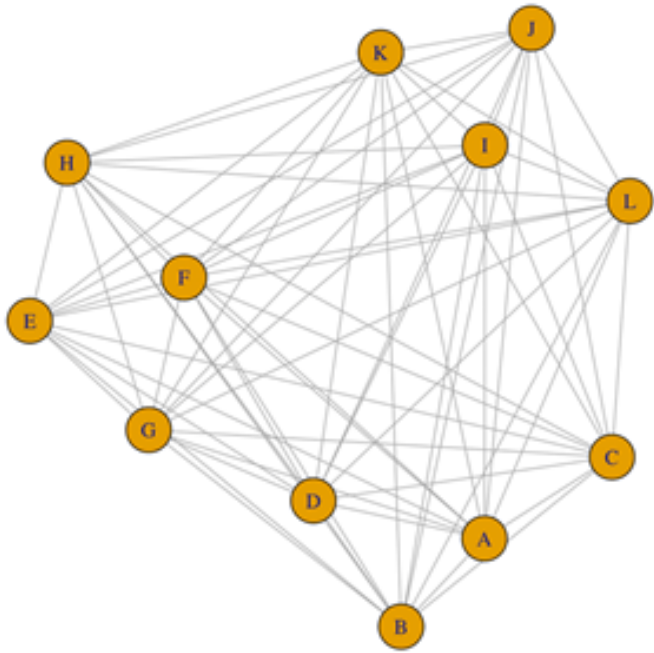
- A **network** is a data structure that is defined by the interactions between the objects being observed.
- It contains two types of information:
  1. Information about the objects being observed;
  2. Information about the interactions between them.





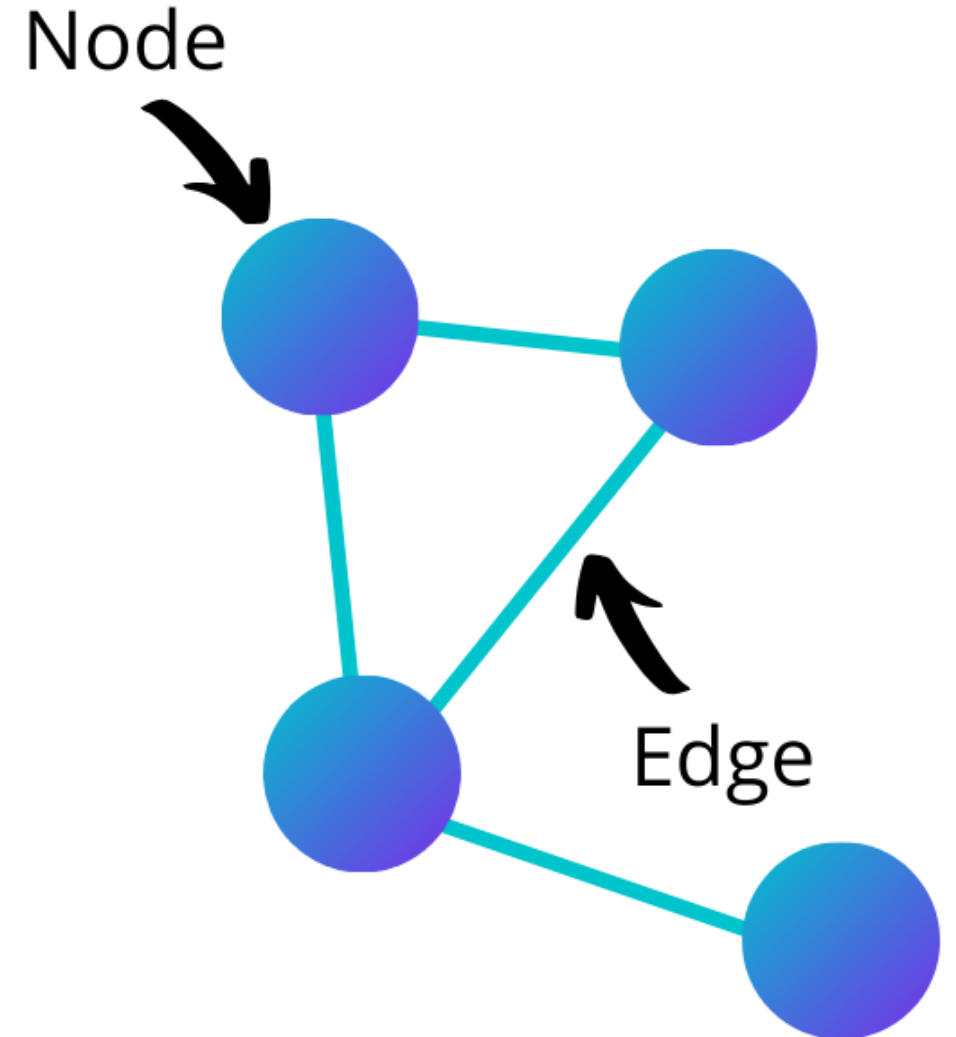
# Networks and Graphs

- This kind of data structure is called a **graph**.
- Technically, "graph" is the abstract, mathematical term for this structure (so you'll hear mathematicians talk about "graph theory" etc.) – once you create one with actual data in it, it's a **network**.



# Nodes and Edges

- A quick word on terminology...
- We've been talking about "objects" and "interactions" or "connections", but in network analysis the technical names for these features are **nodes** and **edges**.
- We'll discuss these more later – for now, just remember the terms.

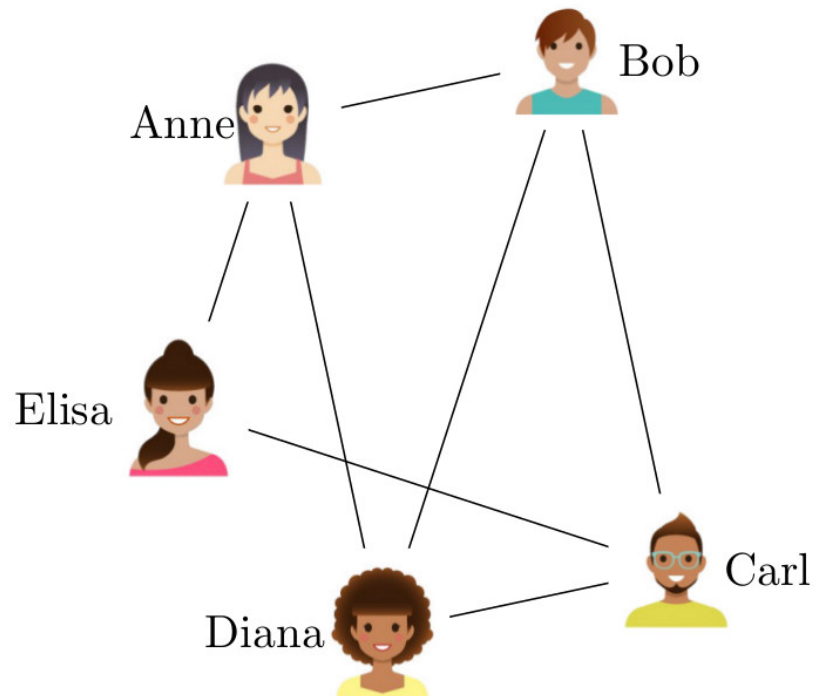


# Graph Theory Applications

- Graph Theory – the study of interconnected networks – is used to study many different kinds of networks. For example:
  - Chemists and physicists use graphs to study molecular bonds and quantum fields.
  - Biologists use graphs to analyse genetics and evolution.
  - Computer scientists use graphs to model computer networks and data flows.
  - Mapping applications use graph theory to calculate optimal transport routes.

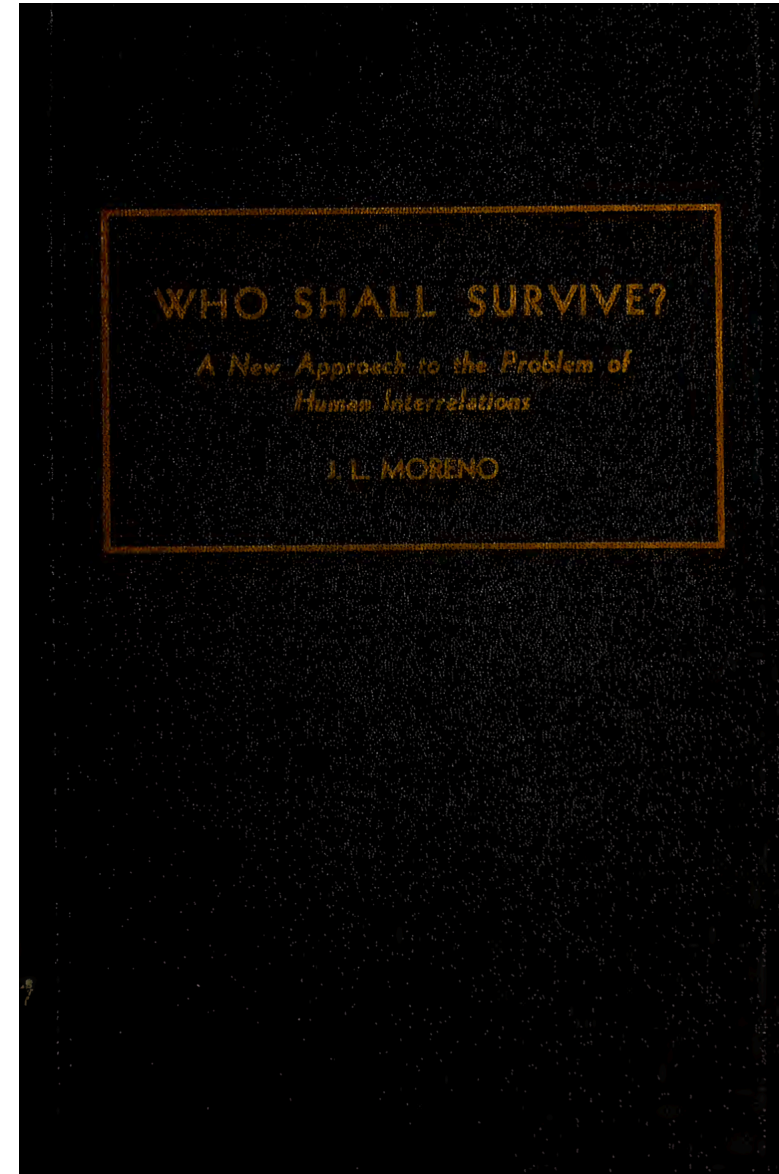
# Social Networks

- We're social scientists – we're interested in using graphs to study people.
  - In this case, the **objects** on the graph become individuals (or some representation of them, like an account on a website).
  - The **interactions** could be anything – friendships, encounters, shared interests, financial transactions...



# Social Network Analysis: Origins

- When we say “social network analysis”, or even just mention a “social network”, you probably have a mental image of something very modern – at least something from the Internet age.
- You probably don’t imagine the first book on social network analysis being written in 1934!







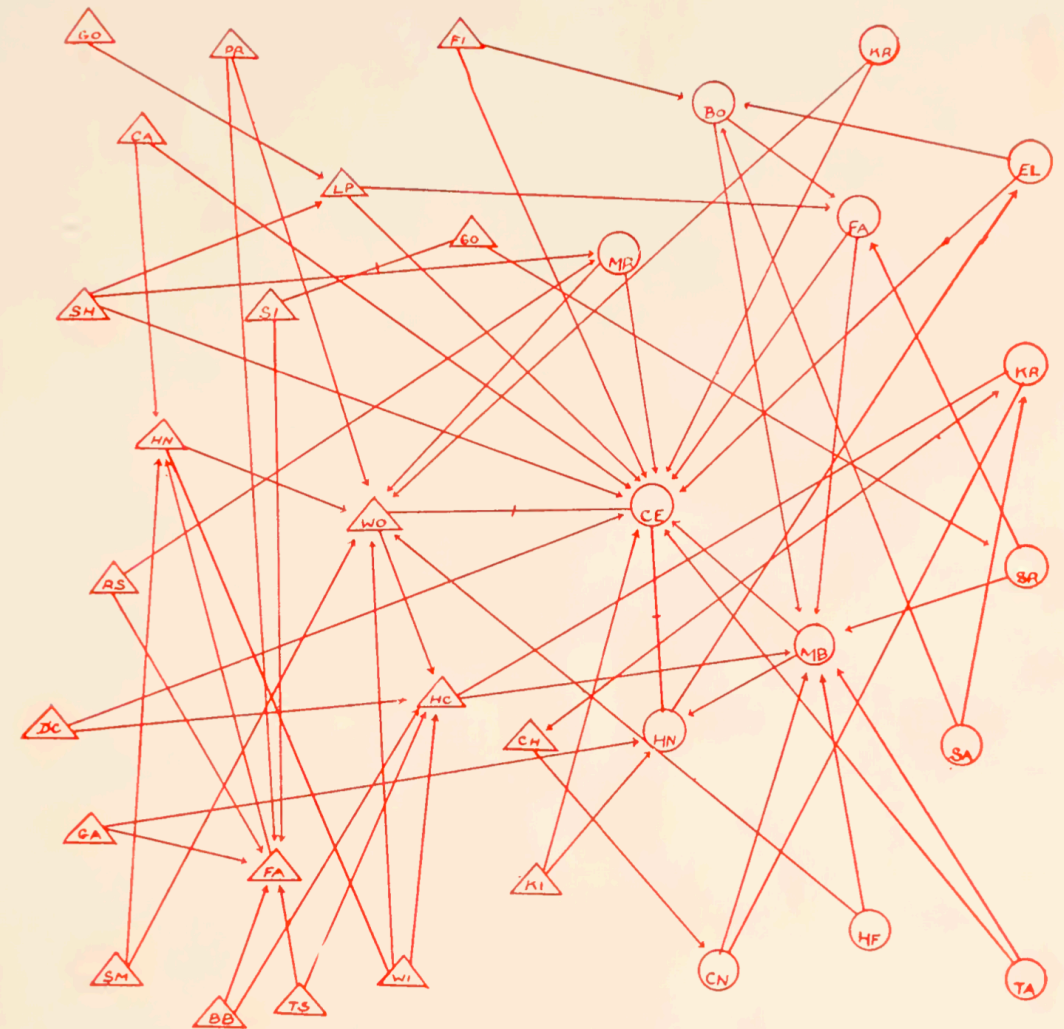
# JL Moreno's "Who Shall Survive?"

- Jacob Moreno (1889-1974) was a Romanian psychiatrist and sociologist.
- He was interested in studying human psychology in groups – believing that observing group interactions could provide insights that you'd miss if you only considered individuals.

- Working in New York in the 1930s, Moreno devised a series of experiments aiming at understanding group dynamics.
- Observing children in a schoolyard over multiple days, he saw how sub-groups would form.
  - These groups were semi-stable but also dynamic; they formed and re-formed in slightly different configurations on different days.
  - The groups interacted with each other and had different types of internal dynamics.



- To better understand these groups, Moreno devised an experiment.
- Each child was asked to (privately) write down the name of the two people they'd like to sit next to in class.
- When Moreno drew a graph of the resulting data, he noticed clear patterns emerging.



CLASS STRUCTURE, 1ST GRADE

21 boys and 14 girls. *Unchosen*, 18, GO, PR, CA, SH, FI, RS, DC, GA, SM, BB, TS, WI, KI, TA, HF, SA, SR, KR; *Pairs*, 3, EI-GO, WO-CE, CE-HN; *Stars*, 5, CE, WO, HC, FA, MB; *Chains*, 0; *Triangles*, 0; *Inter-sexual Attractions*, 22.





- Moreno's studies of the schoolchildren (which he also repeated in other settings with adults) had uncovered something vitally important.
- Analysing their community as a network allowed him to see very different roles played by different people.
- The shape and pattern of the network – known as its **topography** – could be interpreted in order to understand the community not merely as a collection of individuals, but as a dynamic social system.



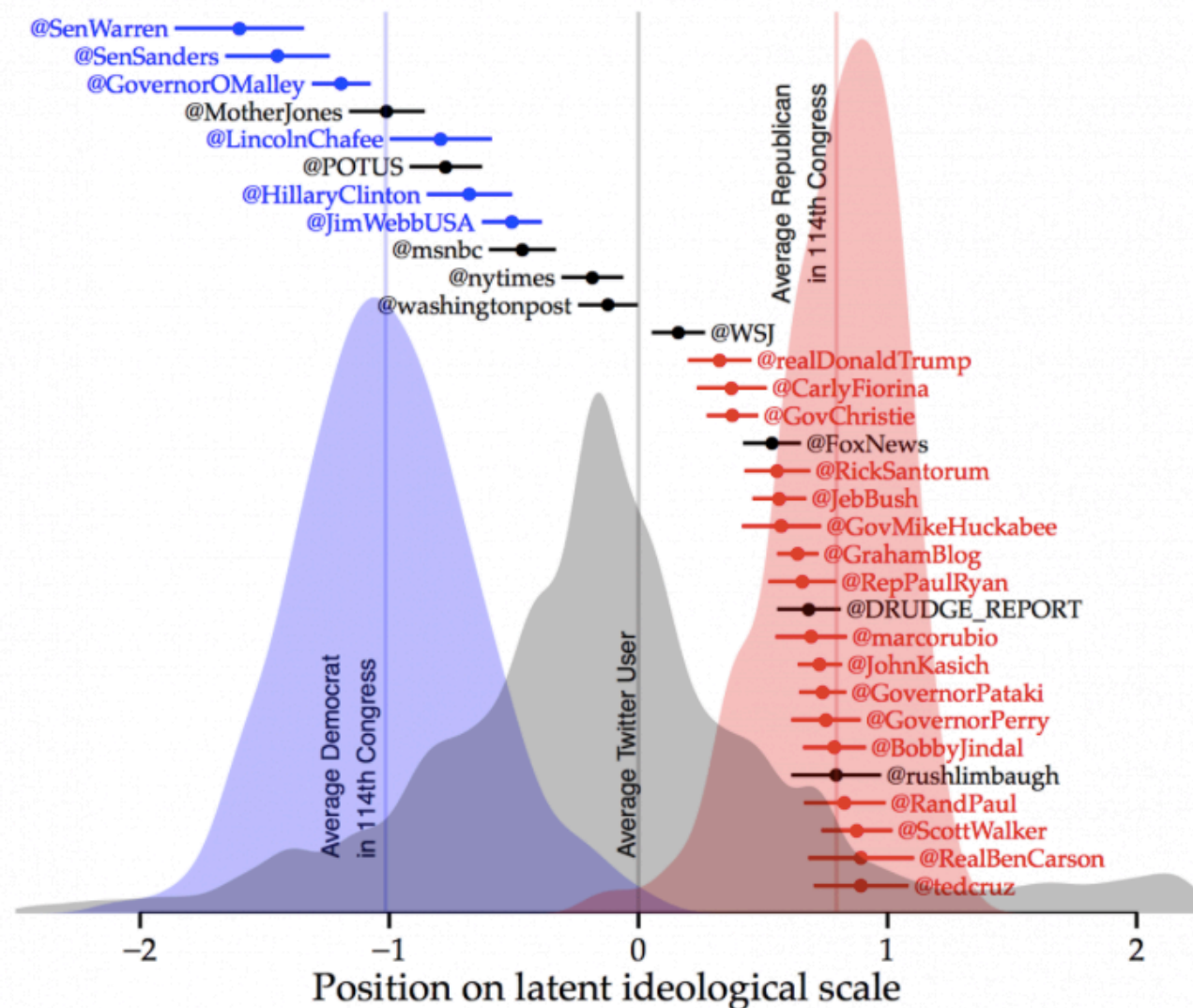
# Online and Offline Social Networks

- Of course, 86 years later, our use of the word “social network” has evolved; it now commonly refers to sites like **Facebook**, **Twitter** and **Instagram**.
- In social science, the definition remains broader – those services **are** social networks, but so is a town, a classroom, a parliament, or any other group of human beings interacting.

# Barberá 2015

- This may not look like a network – but it's based on network analysis techniques.
- Barberá used the overlap between politicians' followers on social media to calculate the distances between them in a network.

Twitter ideology scores of potential Democratic and Republican presidential primary candidates



Source: author's elaboration from Twitter data. Figure for The Monkey Cage / Washington Post by Pablo Barberá, NYU Data Science

# Examples of Networks

- In fact, this very broad definition helps to show why network analysis has become so important – because a great many things in the world are best explained through the study of **interactions**.
- Let's take a look at a few relevant examples of data around us that could be effectively represented and analysed as a network.

# Small Groups

- **Parliaments**

Edges: interactions between politicians (participation in events, co-sponsoring bills, etc.)

- **Companies / Organisations**

Edges: emails between individuals; participation in meetings or projects

- **Local Communities**

Edges: physical interactions; shared memberships of clubs or organisations; family relationships.

# Large Groups

- **Online Communities**

Edges: shares / retweets; follow / follower relationships; mutual following of third parties

- **Large Towns or Cities**

Edges: living / working in the same districts; shared attendance at major events (sports etc.) or shared habits (drinking in the same bars)

- **Transport Networks**

Edges: Passenger movements through the network



# Academic Literature

- Academic / scientific literature can be modelled as a network in two key ways:
  - Use citations as edges – linking publications / papers according to the research they cite, and are cited by.
  - Use co-authoring as edges – linking researchers according to the people they have collaborated with.

# (Mis)information Flows

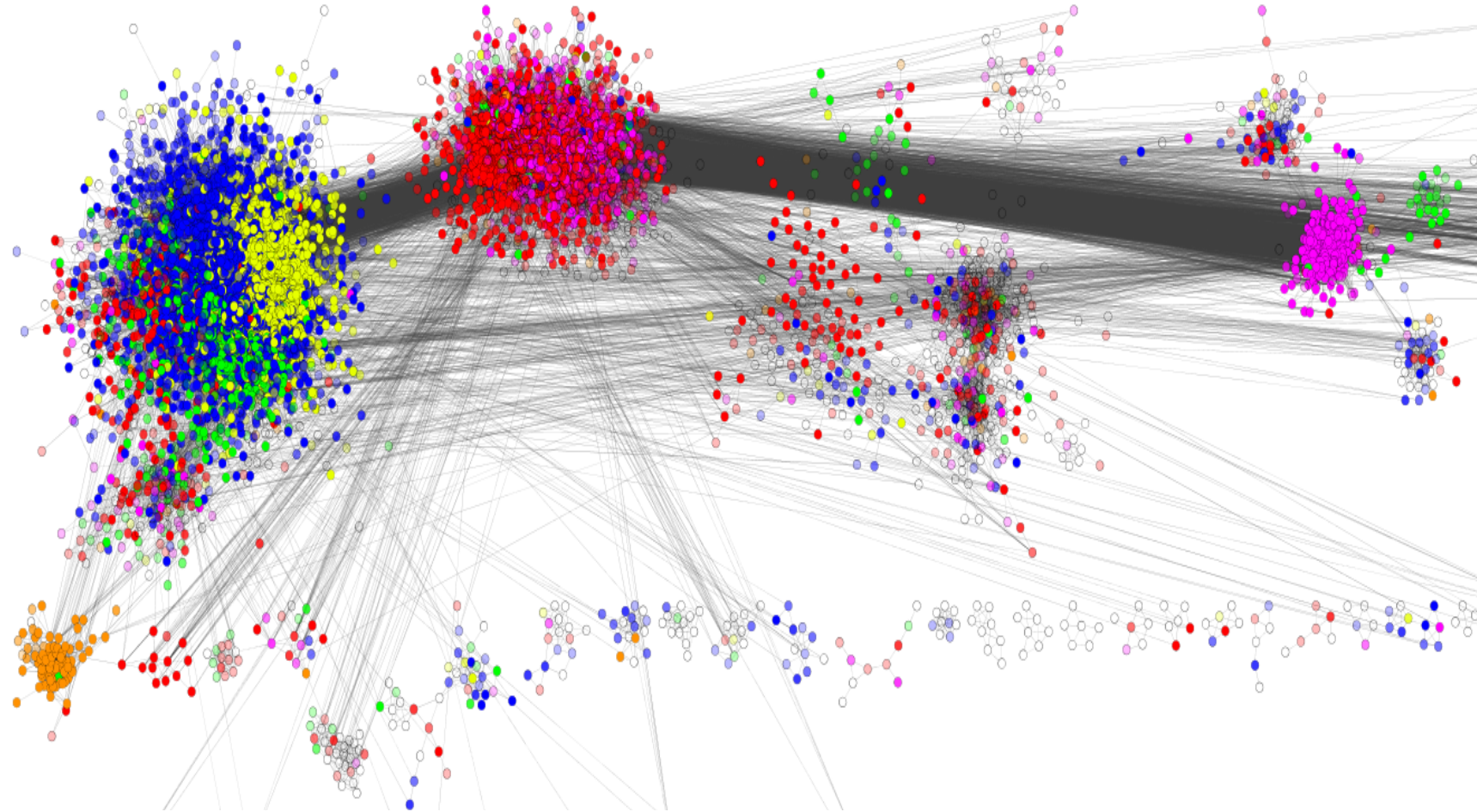
- The flow of information through a community can be traced by modelling how it propagated through a network.
- For example, by constructing a network from Twitter data, we could trace back misinformation / propaganda stories to see:
  - which accounts created the information initially;
  - which accounts were responsible for spreading it into new communities.

# International Relations

- Trade flows between countries and regions can be represented and analysed as networks.
- Similarly, network analysis can help to clarify the complex web of security treaties and relationships which bind many of the world's countries together.

# Fahey 2018

- A “home-grown” one, this time...
- This network shows the similarity between Japanese Twitter users, colour-coded by their political alignment.



# Objectives of this Module

- Our objective by the end of the module is to be able to:
  - Download network data from an online social network;
  - Turn that data into a network in R;
  - Analyse that network to discover what sub-groups or communities it contains, or how information spreads over it;
  - Visualise that network with a publication-quality image which shows the results of our analysis.



# But first.... Small steps.

- Online social networks are **big and complex**.
- Before we start working with that kind of data, we need to practice with smaller-scale networks.
- This is also useful for seeing how you might apply social network analysis to things other than online networks!
  - Lots of really good social science work has been done using network theory to study international relations, relationships between politicians and lobbyists, political donations...

# Course Structure

|           | Week One                          | Week Two                            |
|-----------|-----------------------------------|-------------------------------------|
| Monday    | Introduction Lecture              | Building Networks from Twitter Data |
| Tuesday   | Network Analysis Overview         | Analysing Twitter Networks          |
| Wednesday | Creating Networks in R            | Combining Text and Network Analysis |
| Thursday  | Simple Analysis and Visualisation | Visualising Large-Scale Networks    |
| Friday    | More Advanced Analysis            | New Frontiers and Next Steps        |

# Assessment

- There will be an assignment for this course every day. It won't usually be long or complex – but you will have something to submit every single day.
  - These assignments will make up 60% of your grade.
- In addition, there will be a couple of longer assignments – one mid-way through the course and one final project – which make up 15% and 25% respectively.

# Today's Assignment

- For your first assignment, I want you to think about your own research field – this can be the field you're working on for your final thesis or just the field you're most interested in.
- What are some networks that exist and could be analysed in that field?
  - Don't just say "Twitter" –write briefly about the actual actors (i.e. nodes) you're interested in (politicians? Activists? Media outlets?) and how their interactions could be turned into a network.

# Assignment Details

- Full details of each class' assignment will be posted on my website each day – [www.robfahey.net](http://www.robfahey.net)
- The slides and any class materials (R scripts, data files etc.) for each class will also be available there.