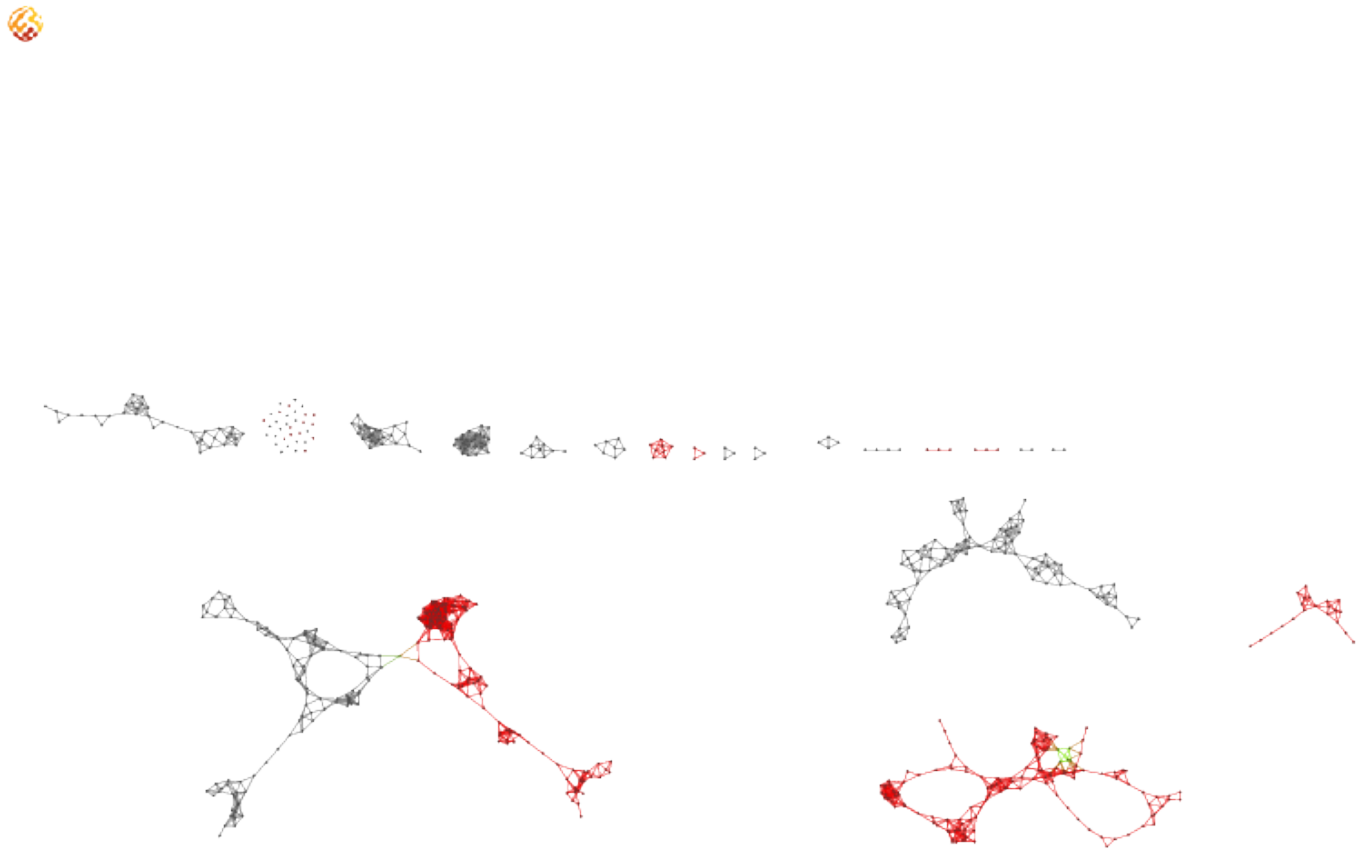
The Topology of Politics

Mikael Vejdemo-Johansson

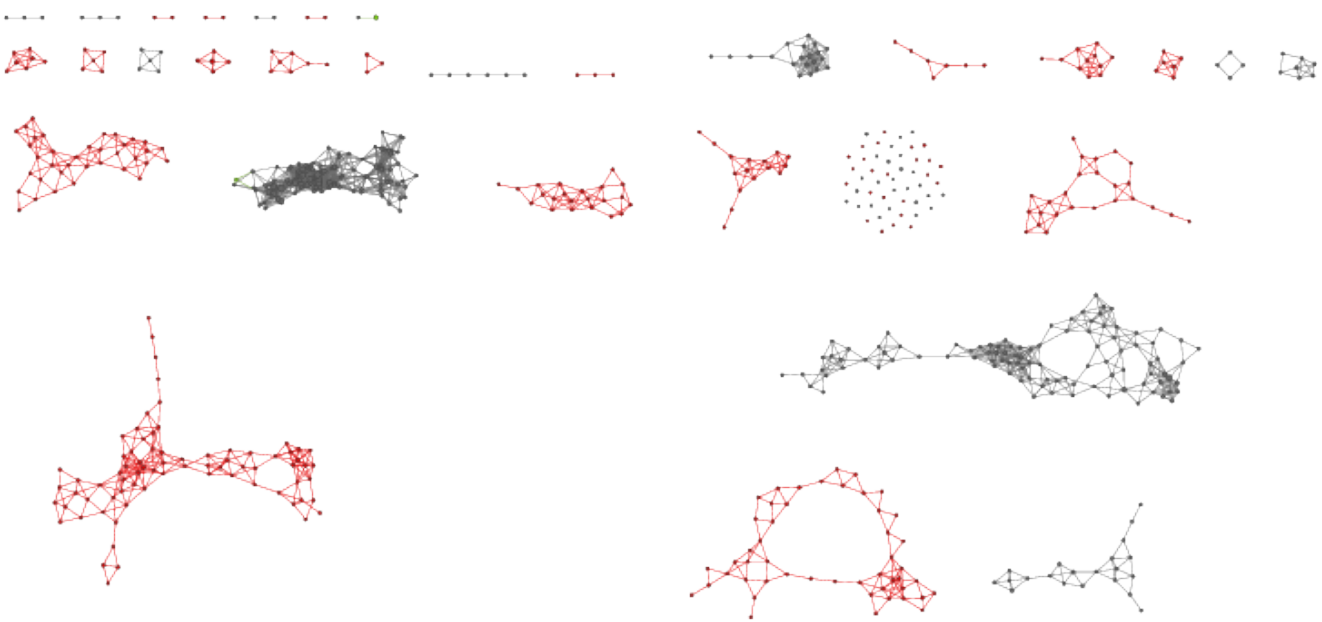
School of Computer Science
University of St Andrews
Scotland

October 26, 2011

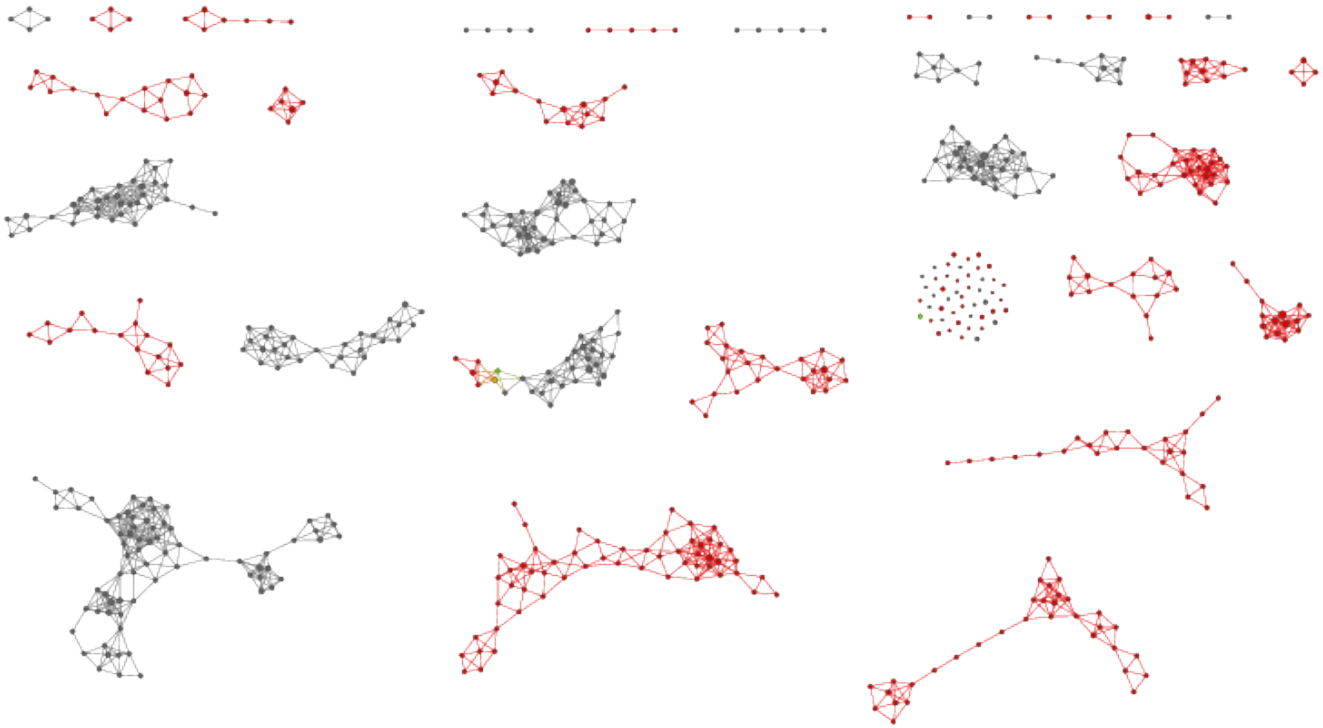
Southern Democrats



Obama wins



Health care reform



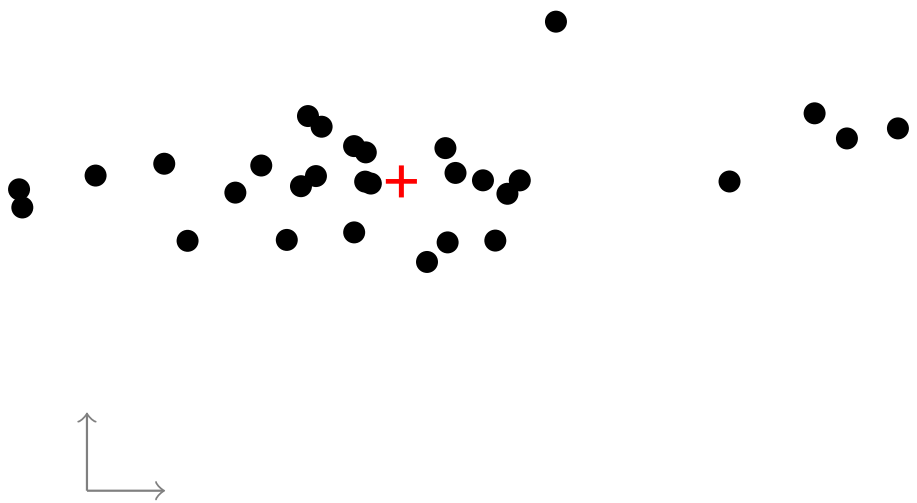
Outline

Shape of data

Fundamentally, data analysis is the task of describing the shape of data:

Fundamental data analysis techniques

Mean (centroid) tells us *where* the data is located.

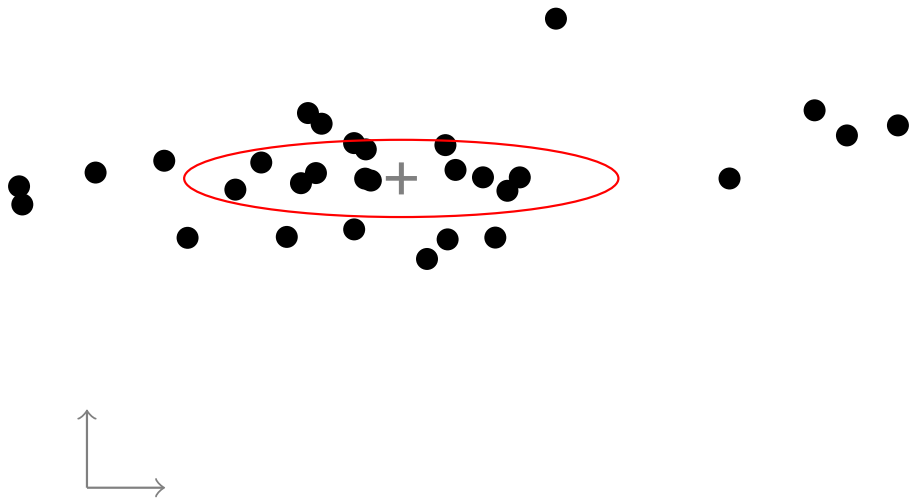


Shape of data

Fundamentally, data analysis is the task of describing the shape of data:

Fundamental data analysis techniques

Standard deviation tells us how spread out the data is.

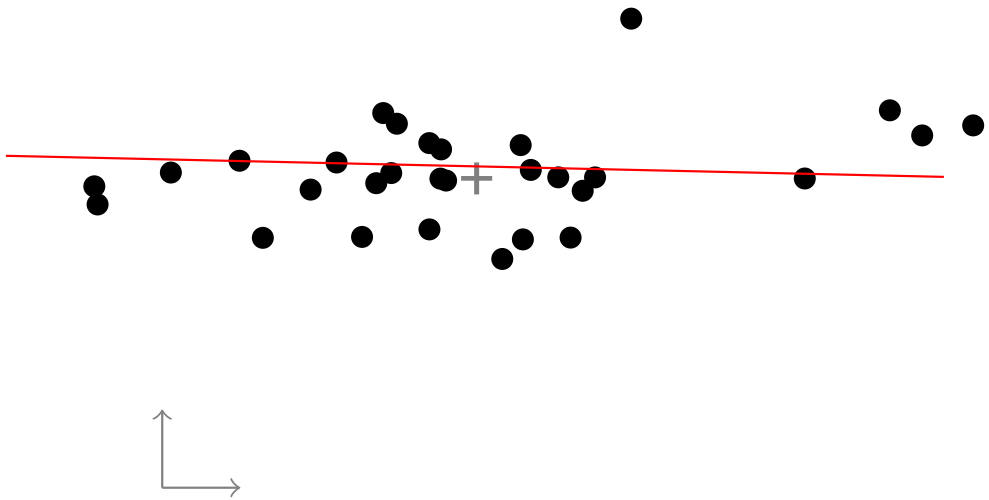


Shape of data

Fundamentally, data analysis is the task of describing the shape of data:

Fundamental data analysis techniques

Regression analyses fit the data to an easy to analyze model.

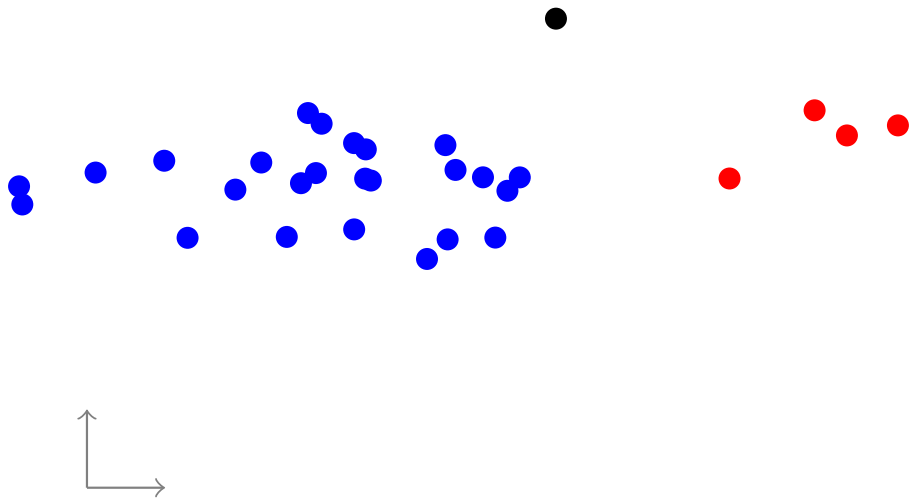


Shape of data

Fundamentally, data analysis is the task of describing the shape of data:

Fundamental data analysis techniques

Cluster analysis divides the data into its connected components.

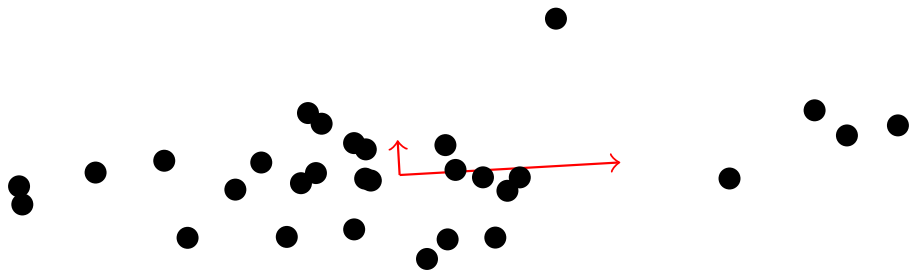


Shape of data

Fundamentally, data analysis is the task of describing the shape of data:

Fundamental data analysis techniques

Principal Component Analysis (and other dimension reduction techniques) give new coordinates that more faithfully represent the data.



Outline

The shape of politics

Data analysis has long pedigree in political science.

- Tufte discusses improvements to techniques in a 1969 paper
- Gallups, polling, and voter prediction leverages more and stronger analysis techniques every year.

Entered into politics in the 90's. Example of the type of analyses done: PCA of votes in House of Congress.

The shape of politics

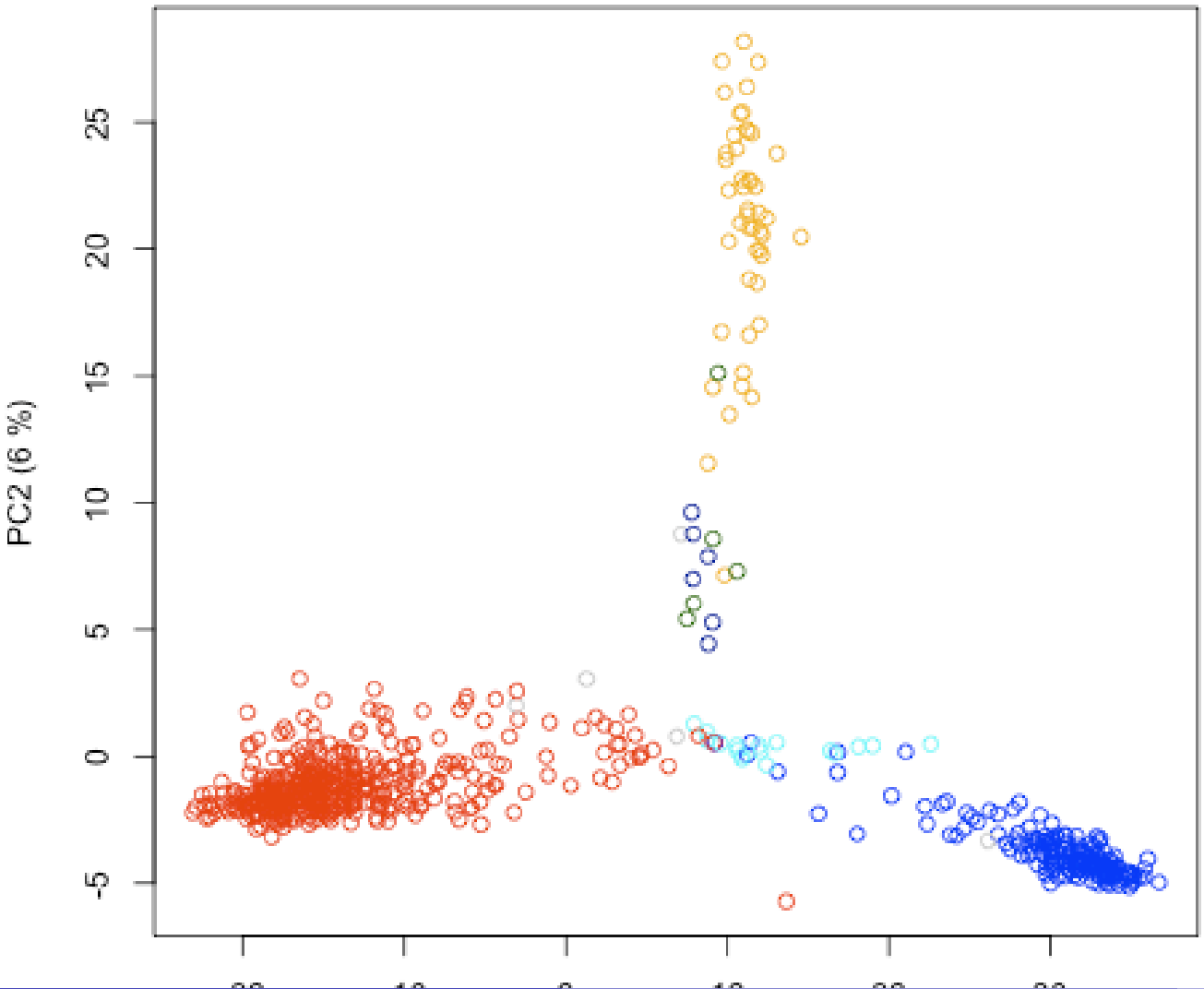
We shall focus on one area of use: the analysis of parliamentary bodies.

Dataset

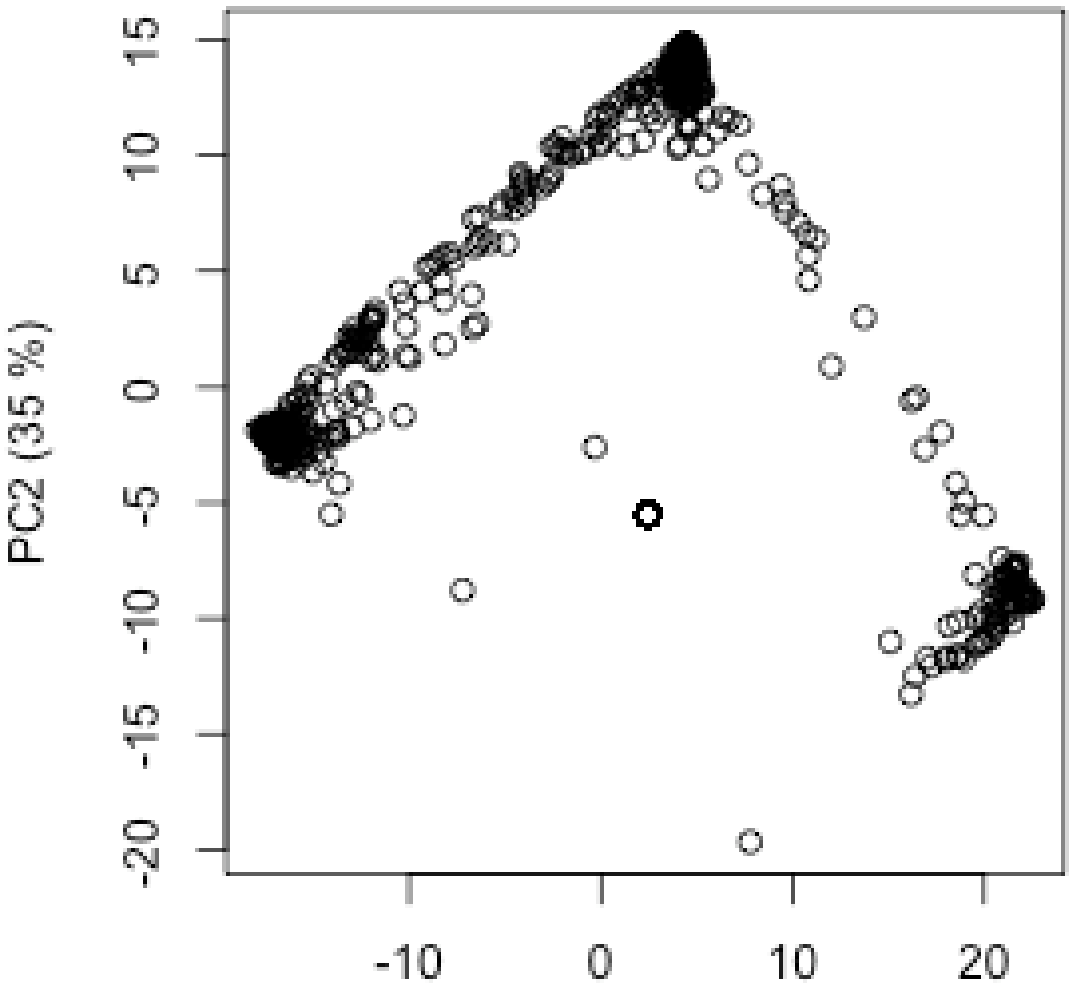
For a parliamentary session, N members of parliament vote on M issues. Votes form an $N \times M$-matrix V of votes cast. Encoded $+1$ for Yea, -1 for Nay and 0 for all other results.

c.f. Dan Rockmore's work using similarly constructed datasets.

The shape of politics



The shape of politics



Outline

Wherefore topology?

Topology gives qualitative information

Topology is about geometric properties not dependent on a metric: How many pieces? Are there holes? Bubbles? Can you get turned around while moving?

Topology captures continuity, connectedness, nearness

In particular, the non-dependence on a metric helps if metrics are ill-motivated, eg phylo-genetics.

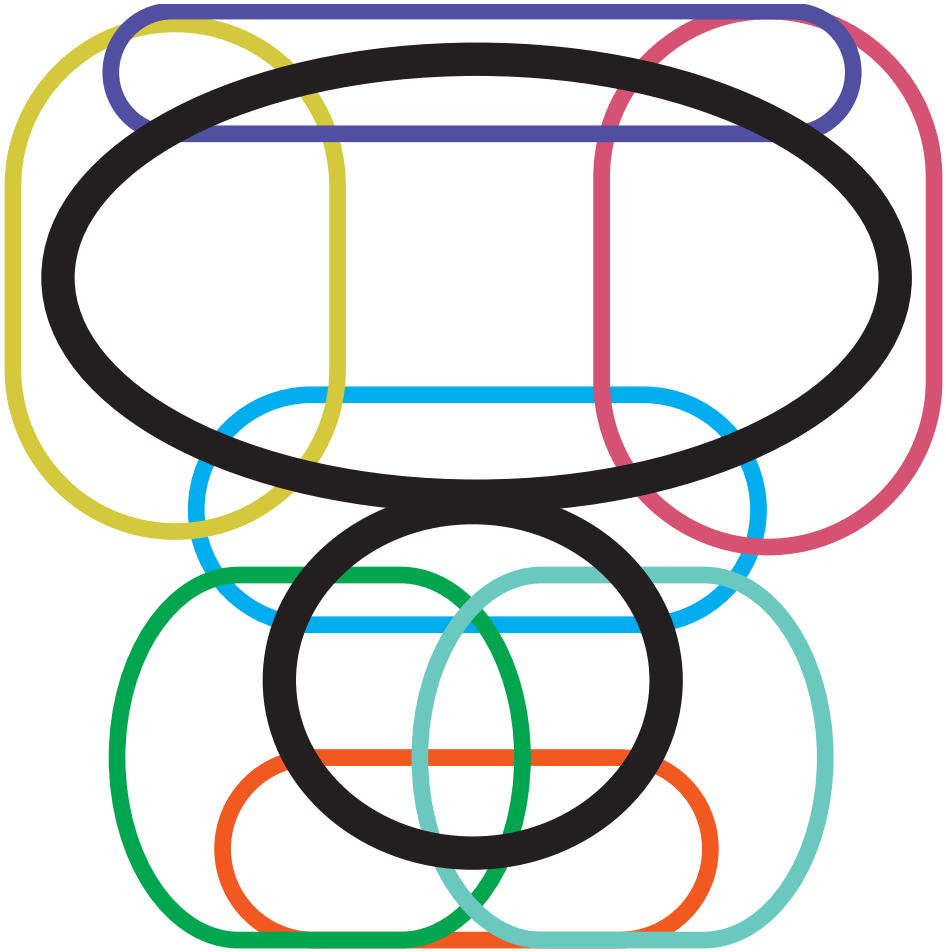
A topological analysis method

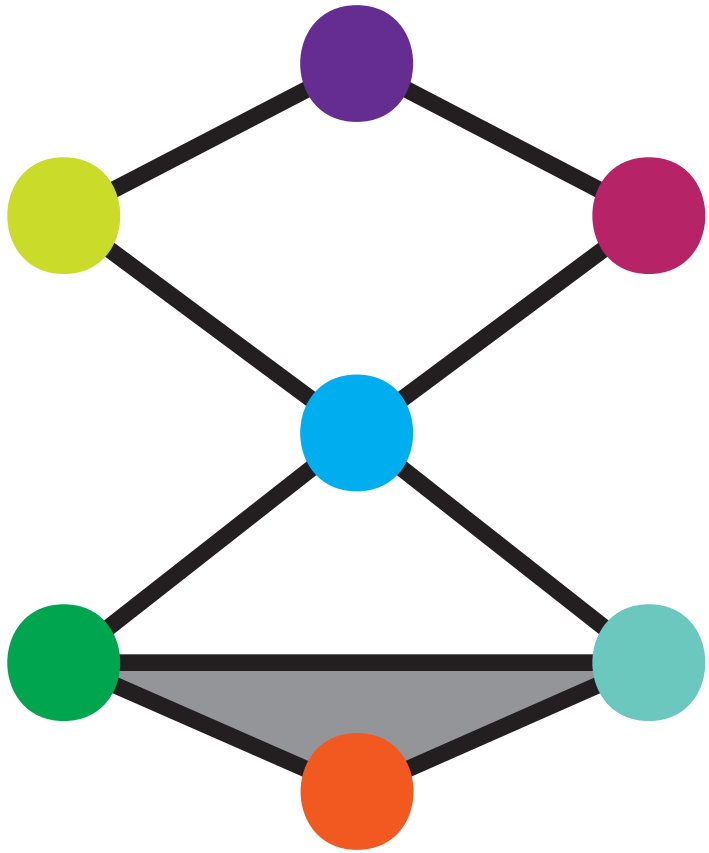
In a recent PhD thesis at Stanford [Singh, '08], a topological method for data analysis was introduced.

Fundamental topological result: Nerve lemma

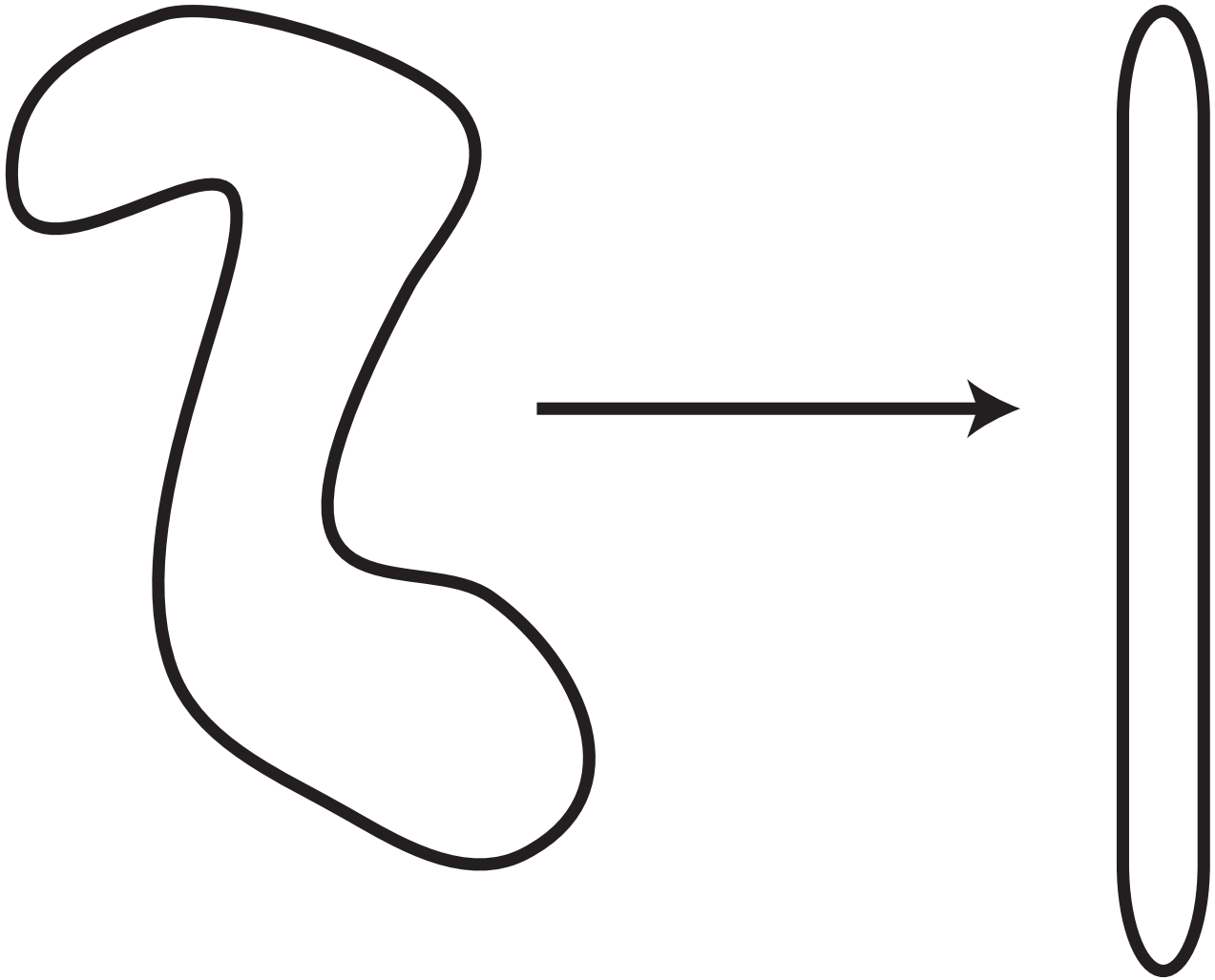
Suppose a space X is subdivided $X = \bigcup_i X_i$ into *contractible* (read simple) components.
Then X is equivalent to the *nerve* of the covering.

The nerve of a covering

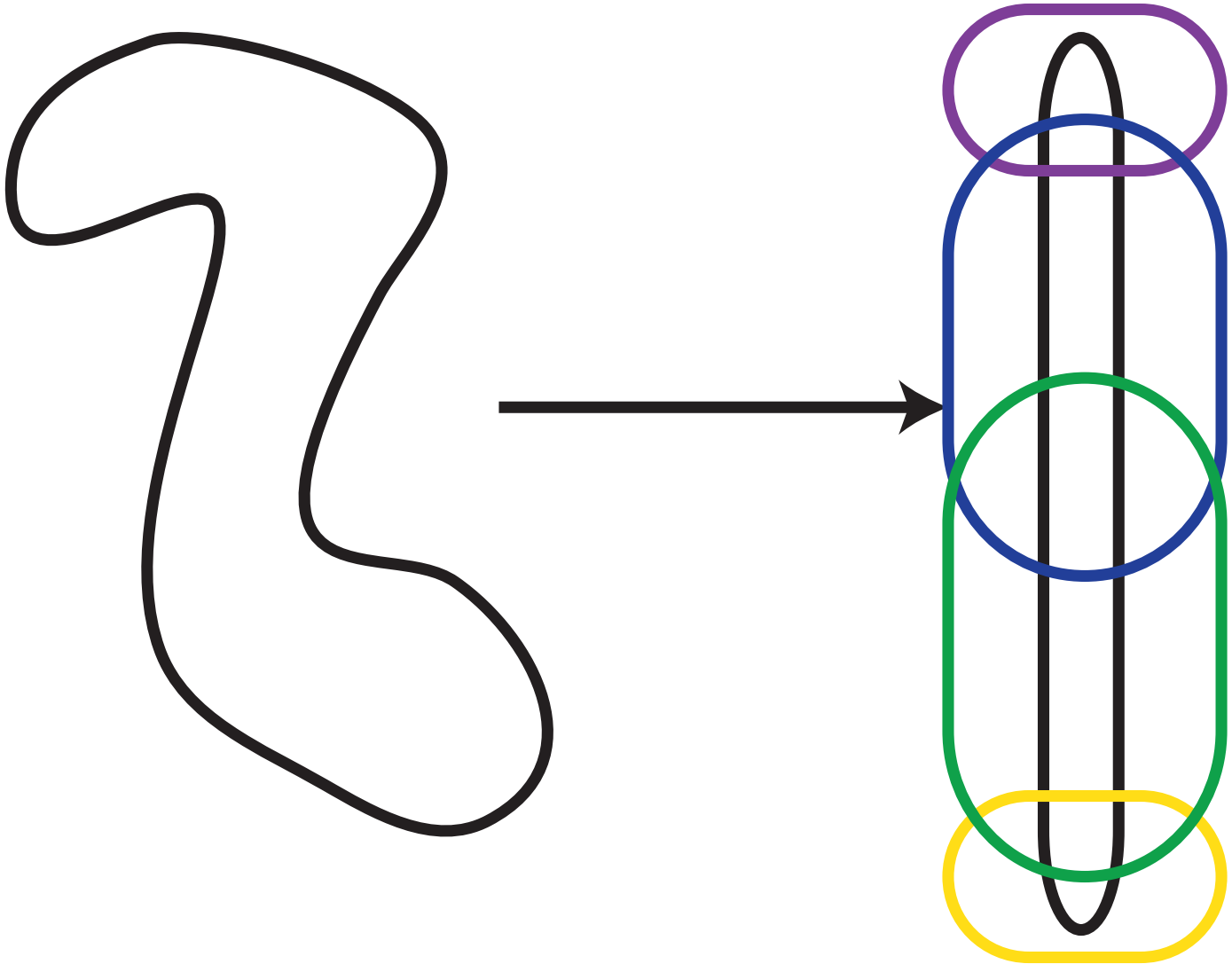




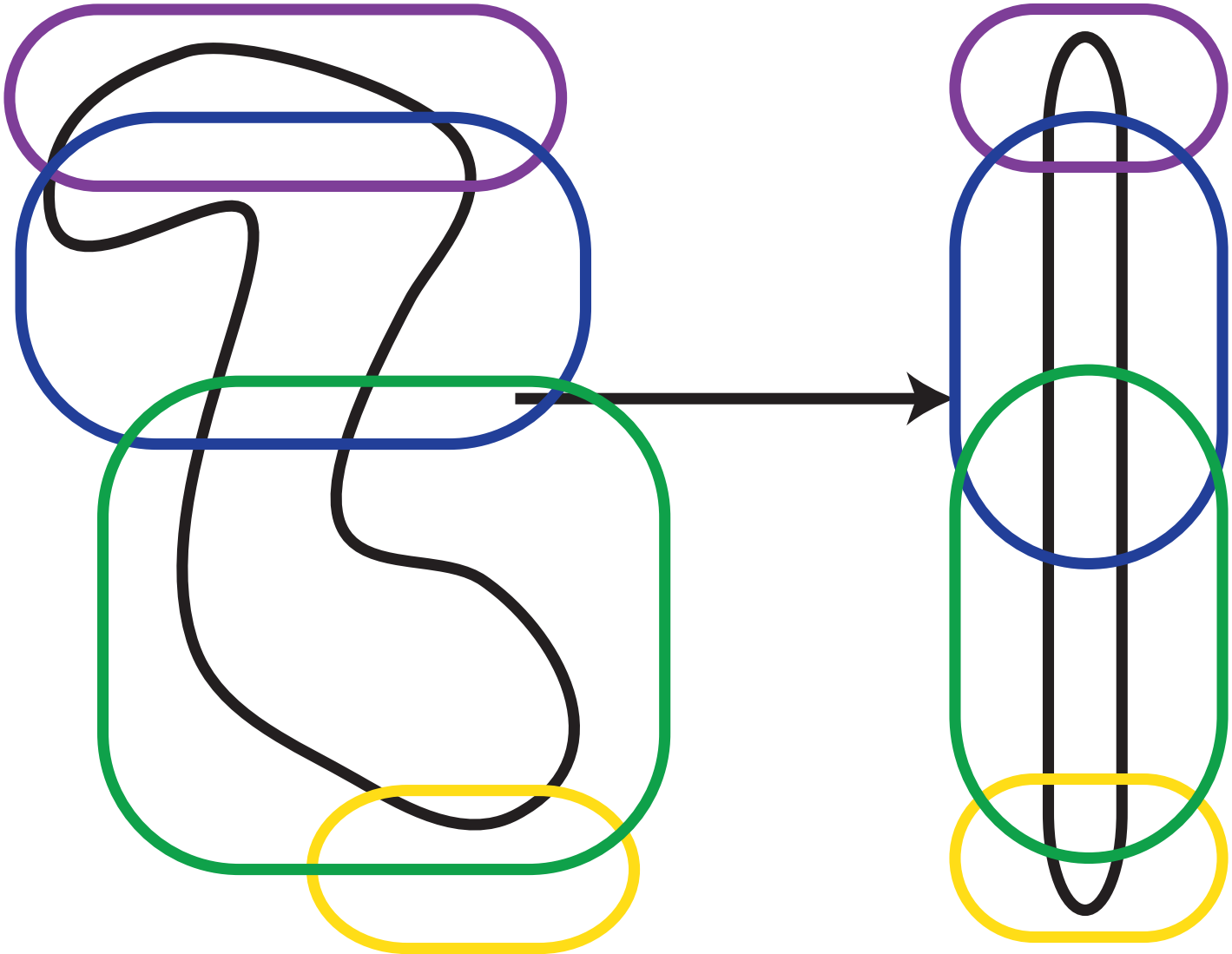
Topological application



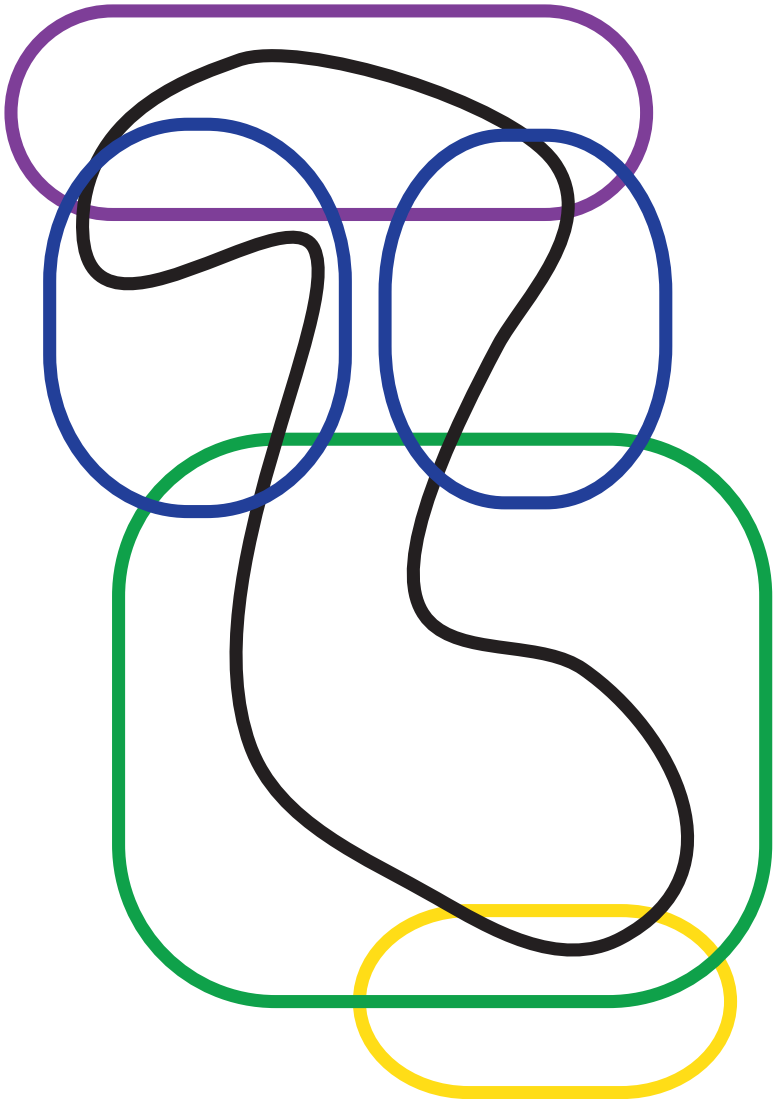
Topological application



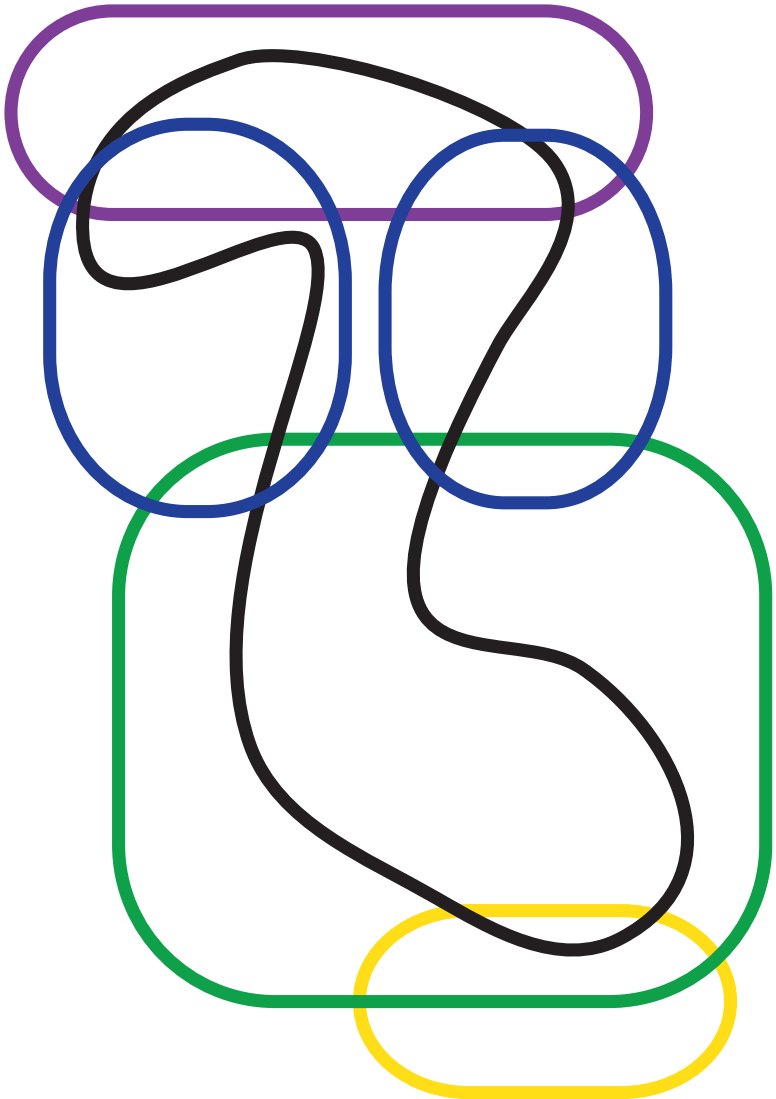
Topological application

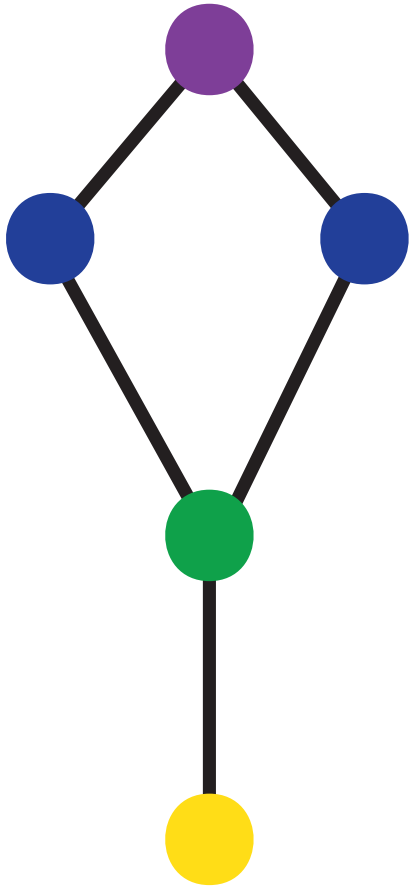


Topological application



Topological application

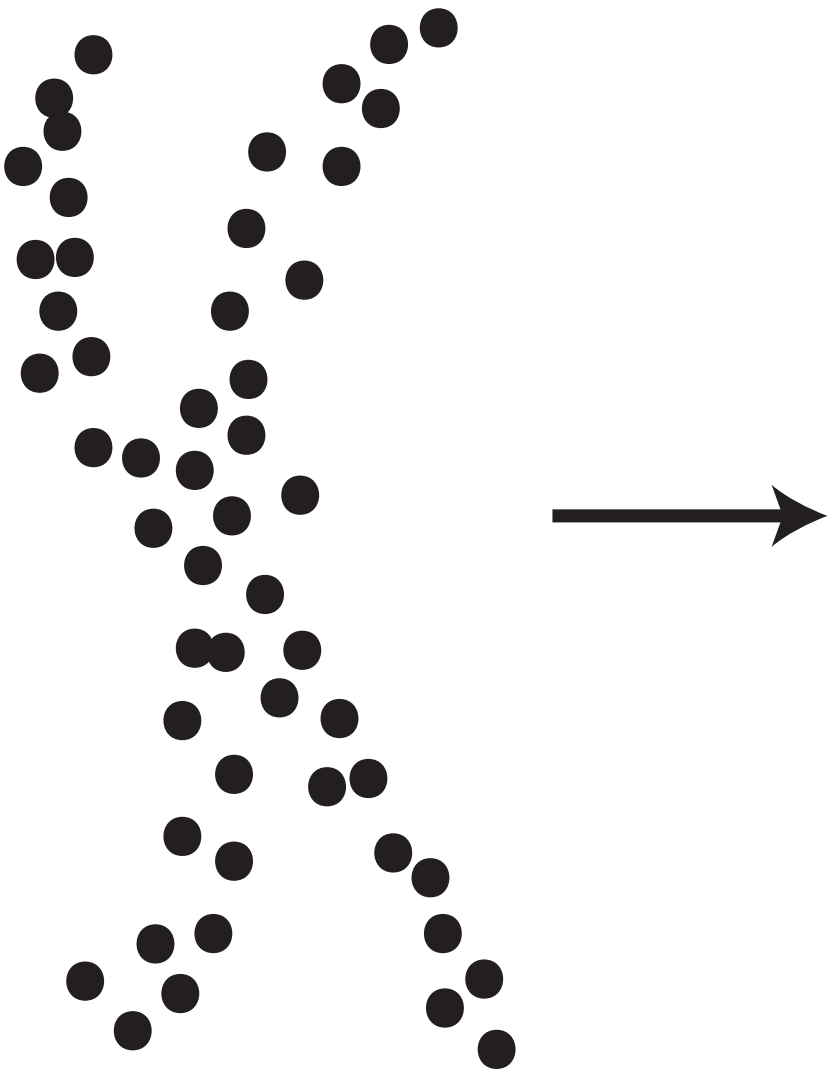




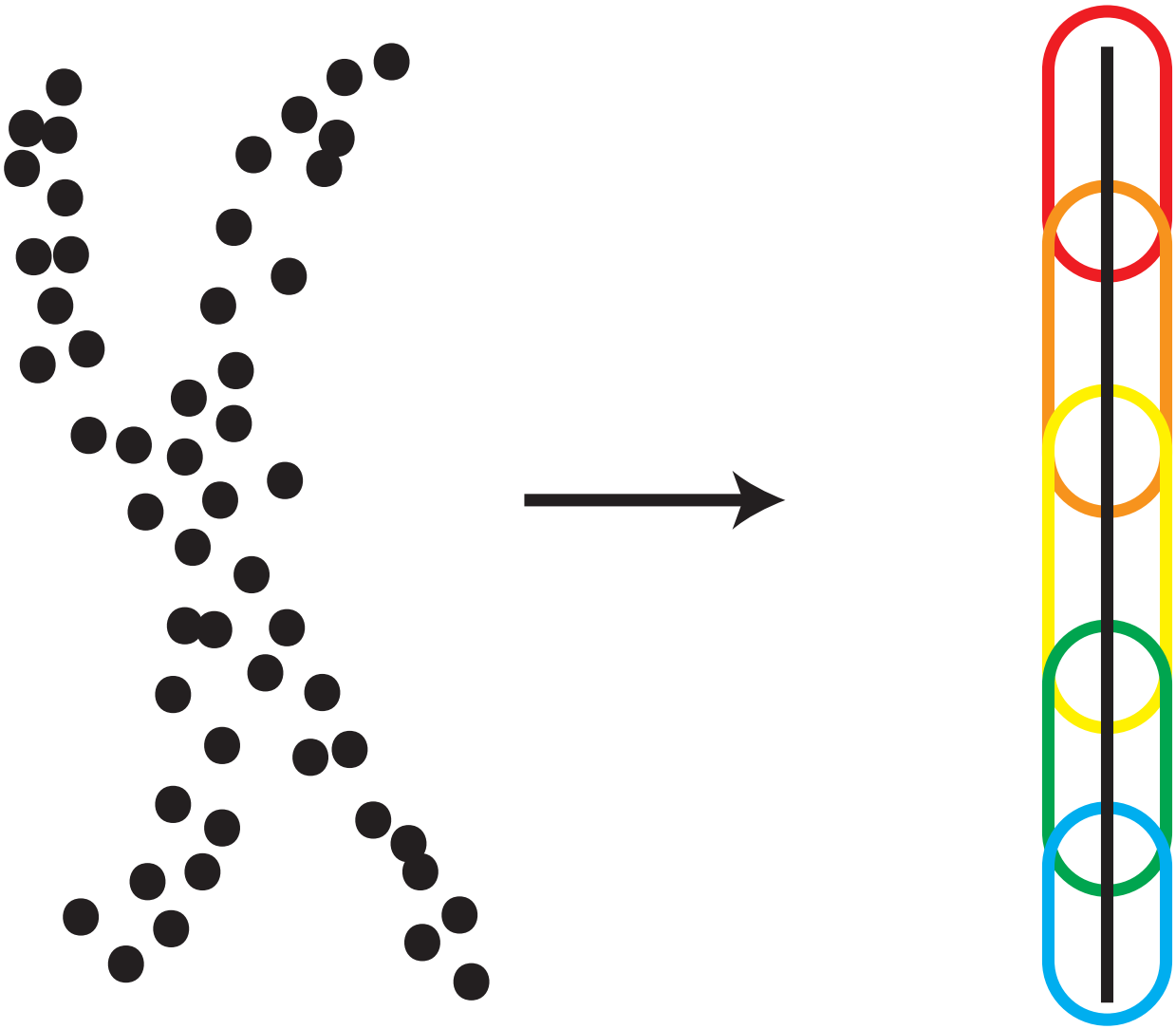
Translate topology to statistics

Continuous function	Measurement function on datapoints
Covering of target space	Covering of datapoints
Preimages	Preimages
Connected components	Clusters
Nerve complex	Mapper diagram

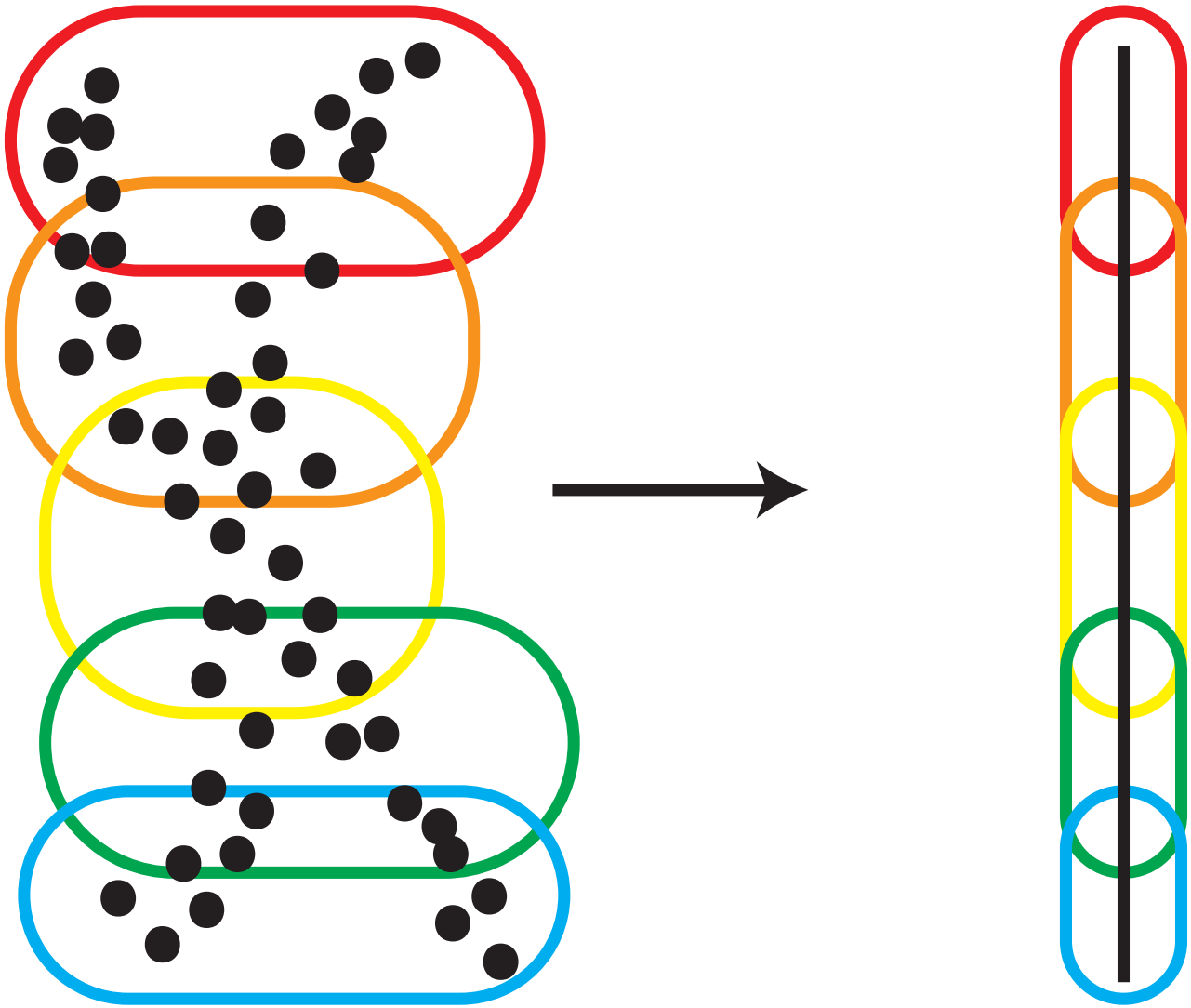
Mapper algorithm



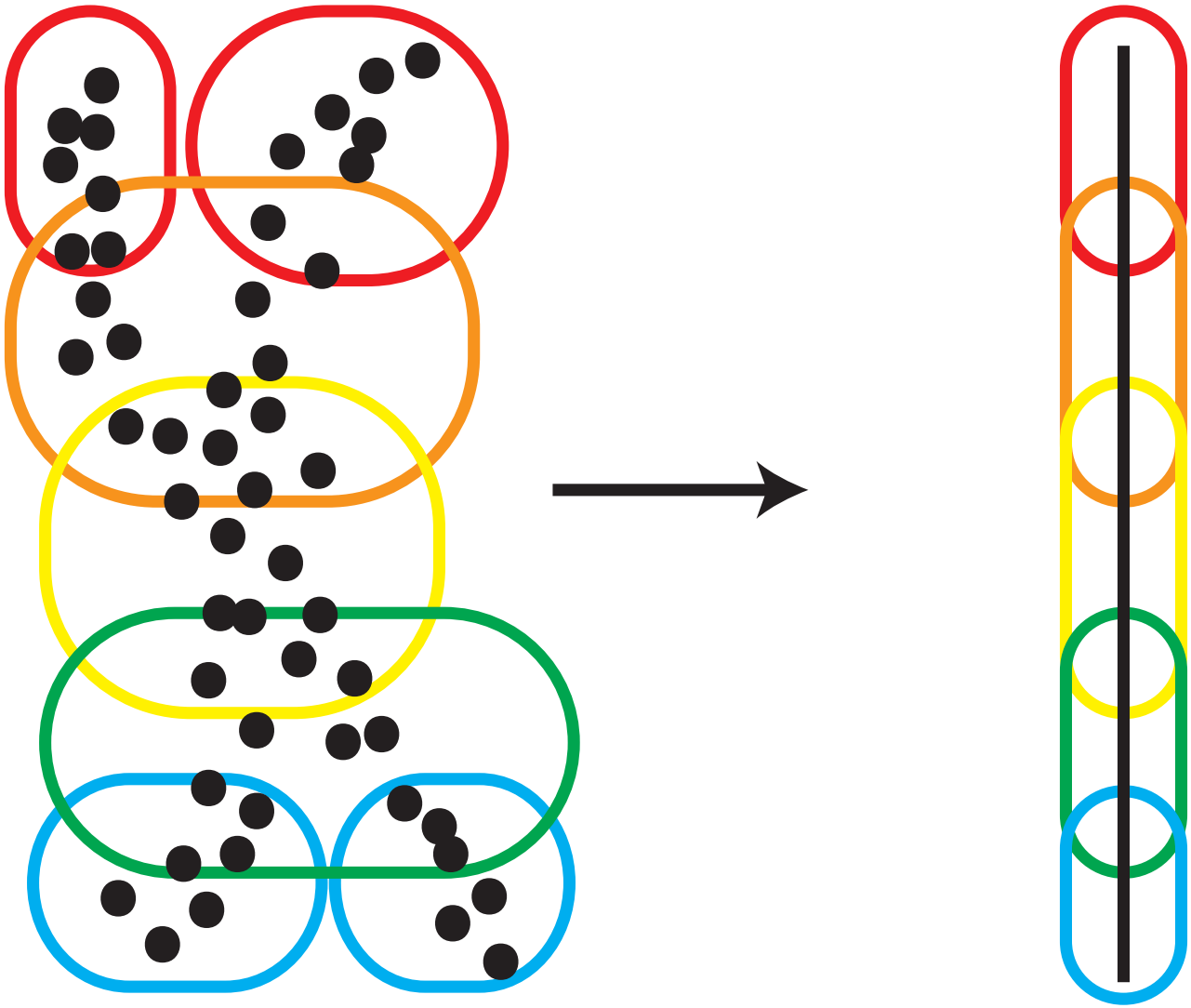
Mapper algorithm



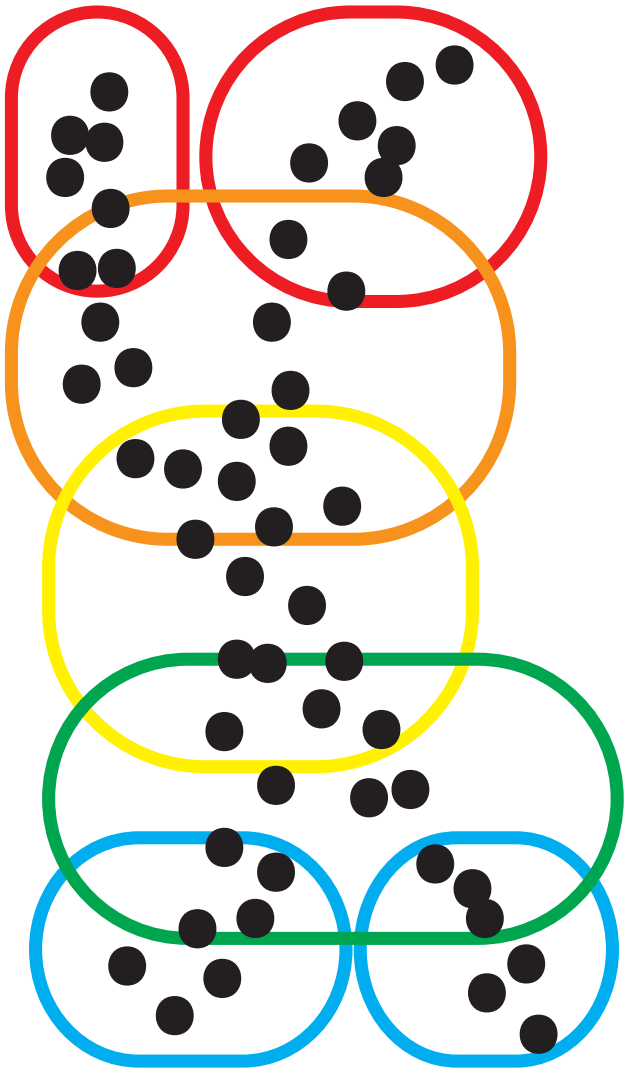
Mapper algorithm

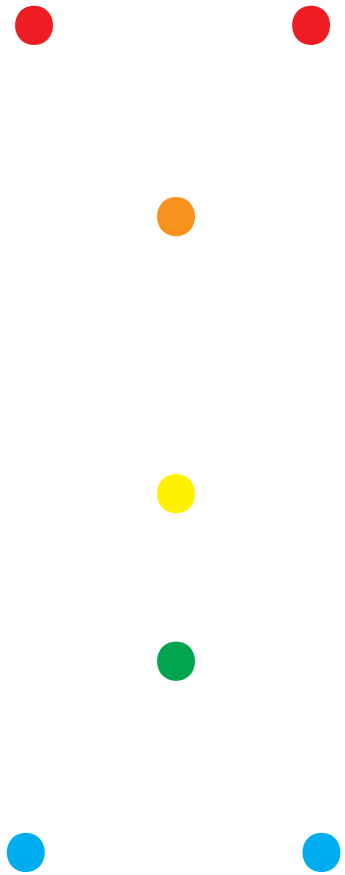


Mapper algorithm

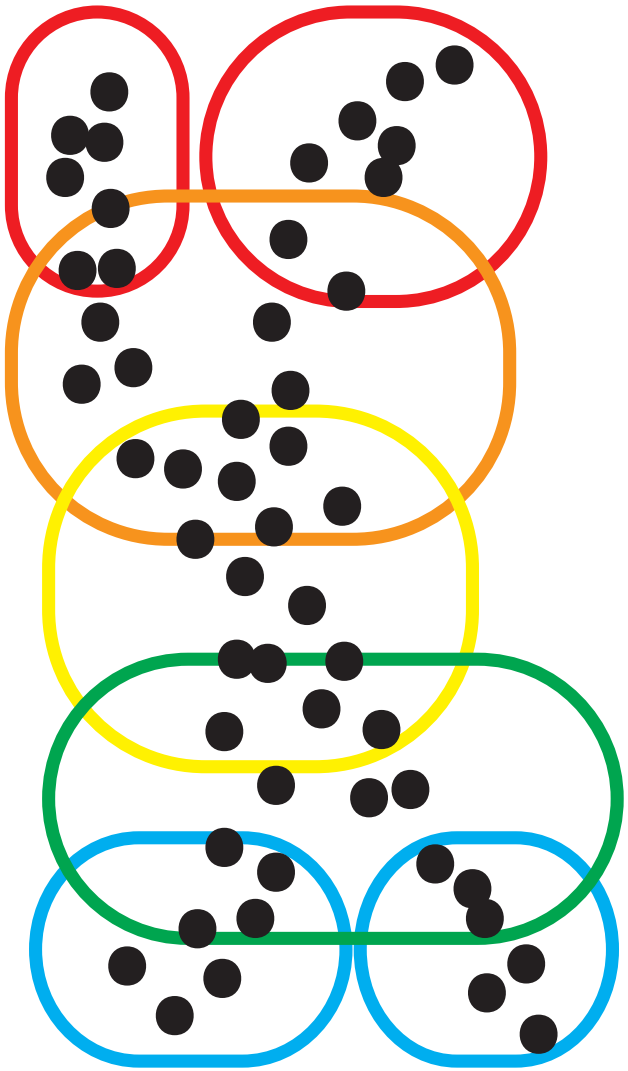


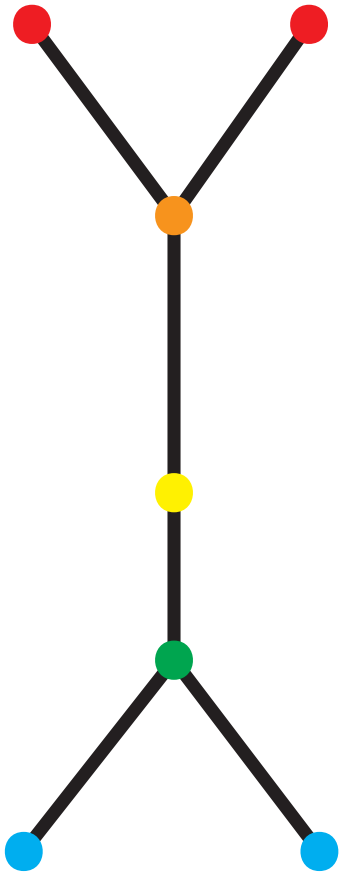
Mapper algorithm





Mapper algorithm





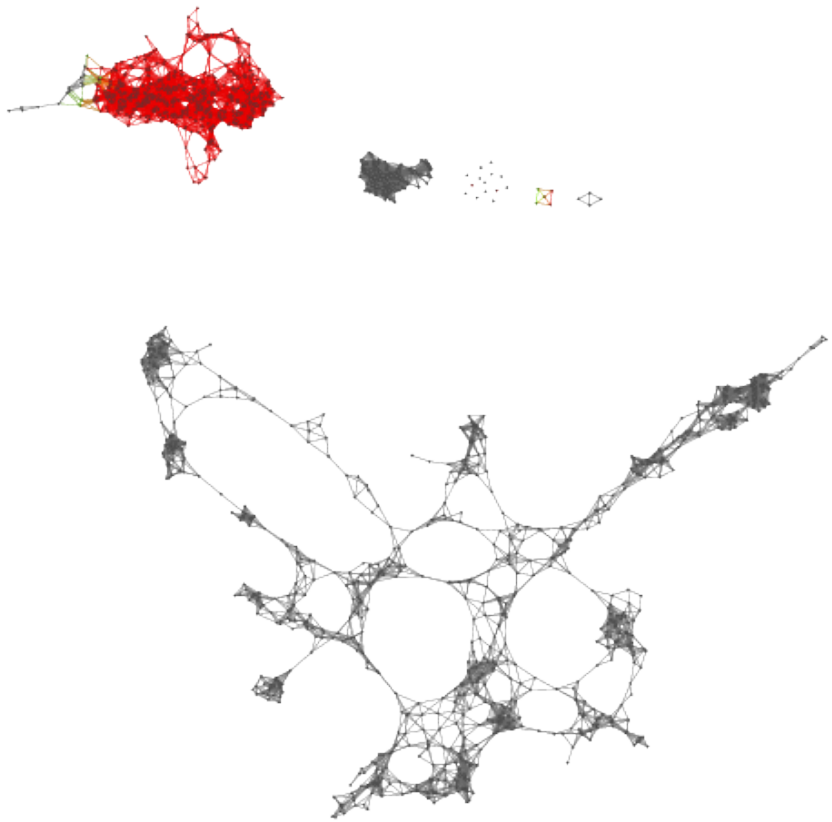
Mapper algorithm

Implementation

This method is provided in a software package currently marketed by Ayasdi.
Startup company founded by Gurjeet Singh (original thesis on Mapper) and Gunnar Carlsson (thesis advisor).

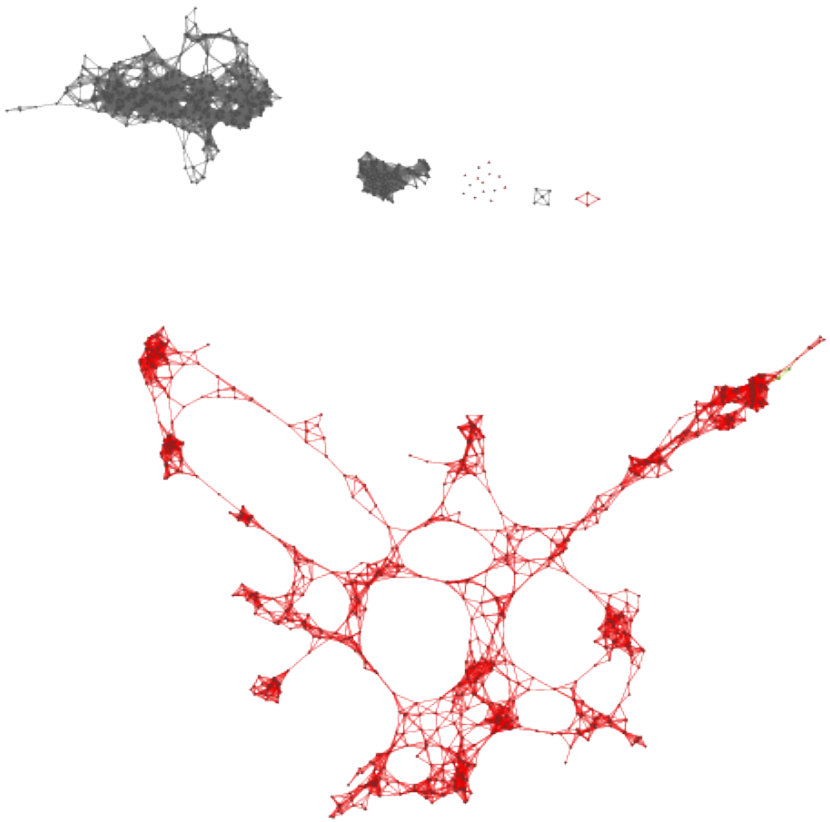
UK Parliament

Parliamentary period 2001–2005. Conservative



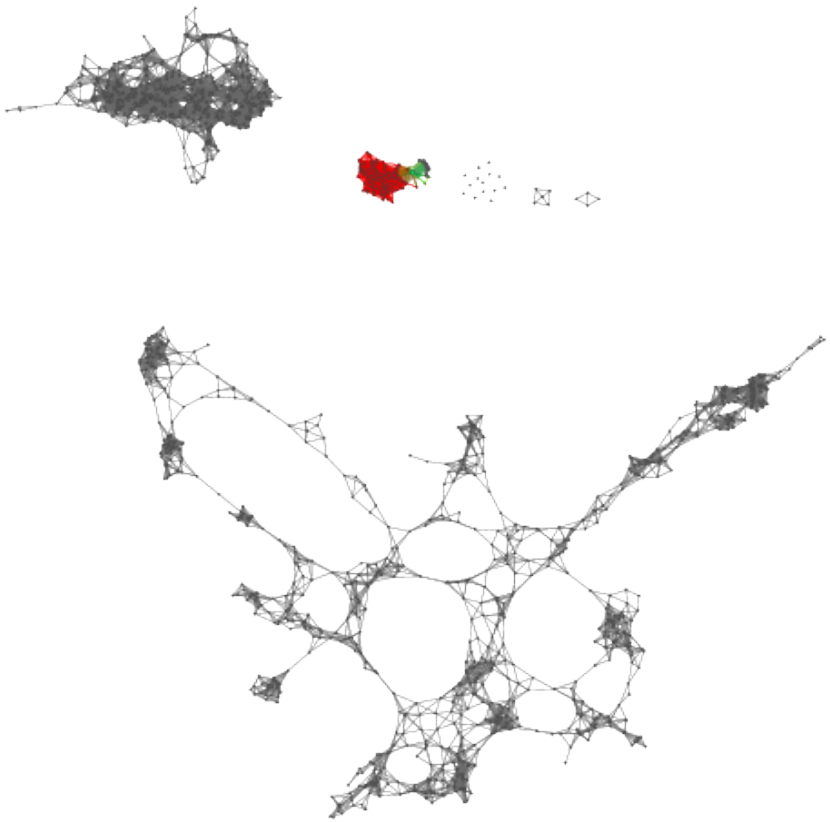
UK Parliament

Parliamentary period 2001–2005. Labour



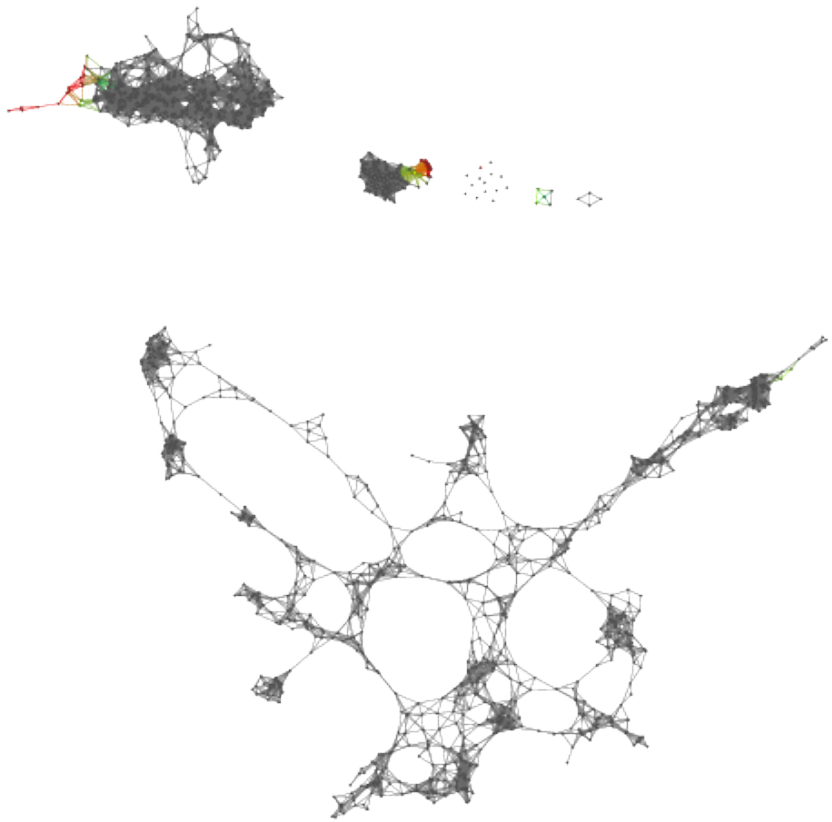
UK Parliament

Parliamentary period 2001–2005. Liberal Democrats



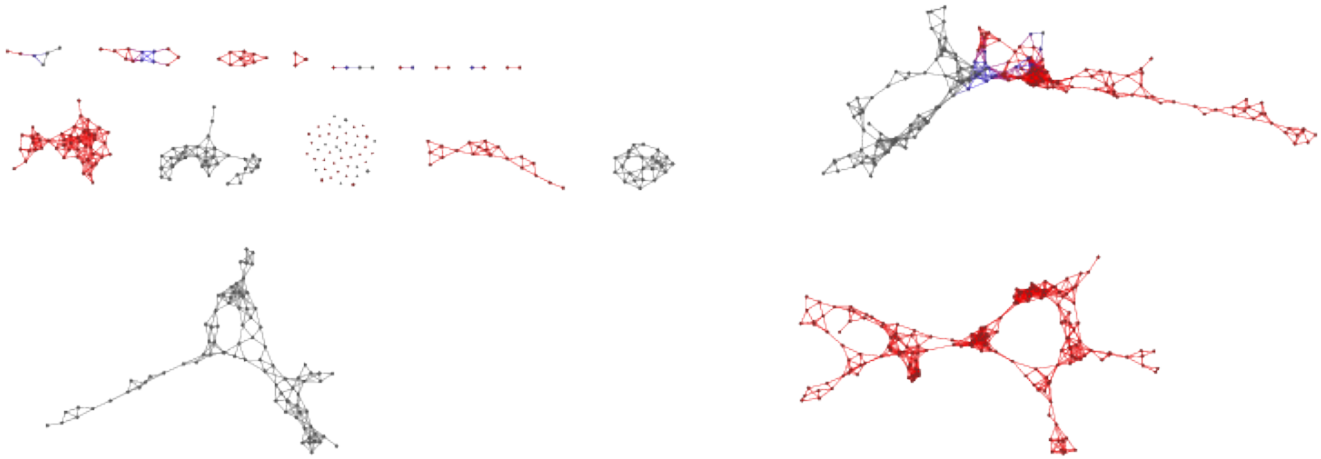
UK Parliament

Parliamentary period 2001–2005. Regional parties



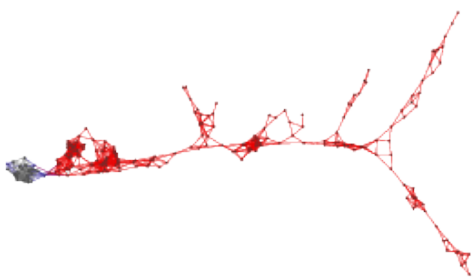
House of Representative.

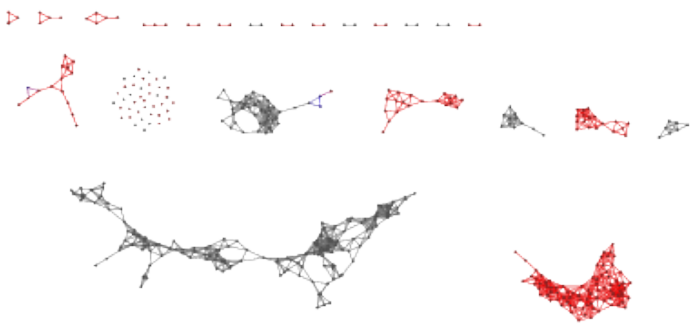
1991



House of Representative.

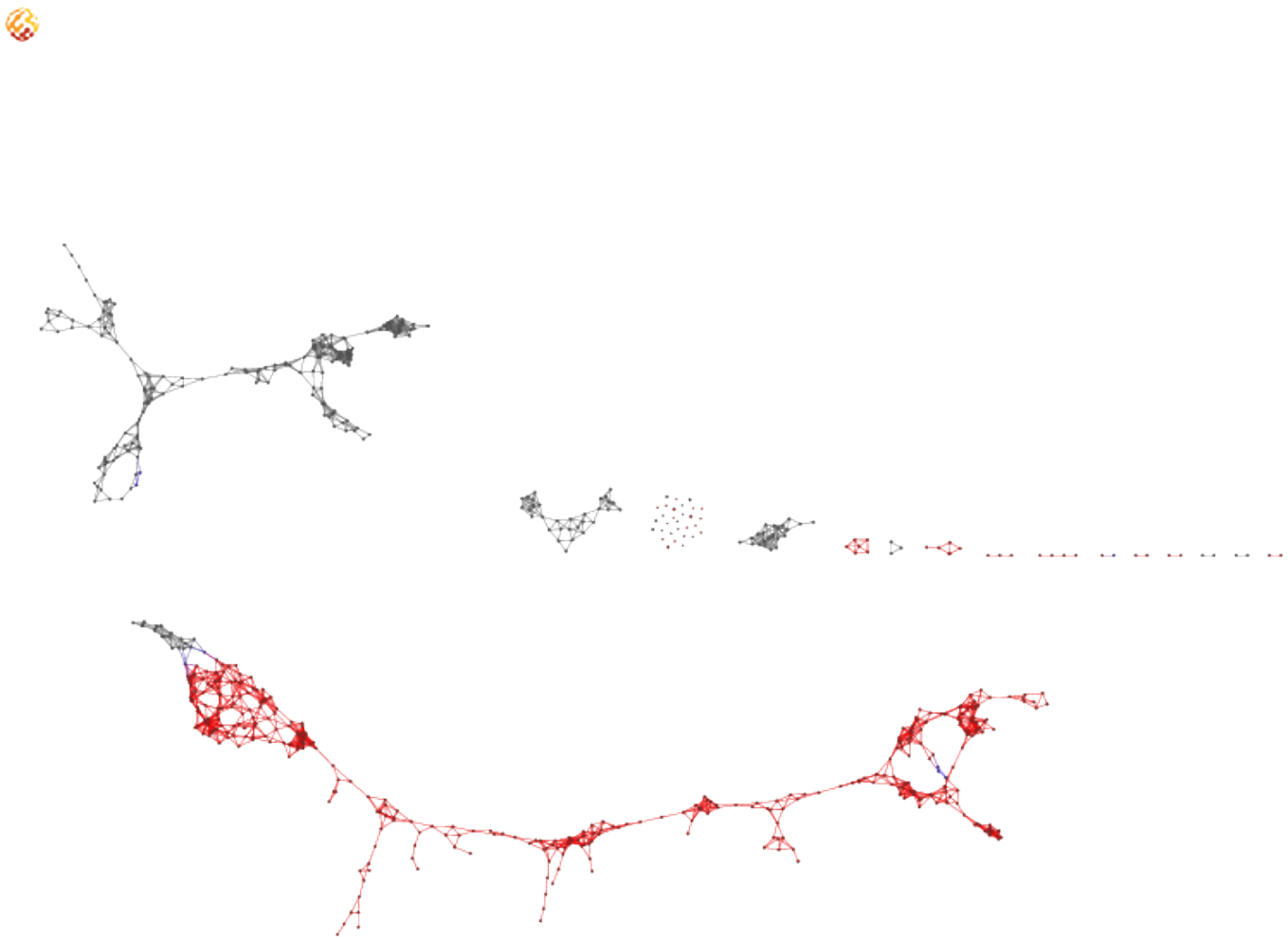
1992





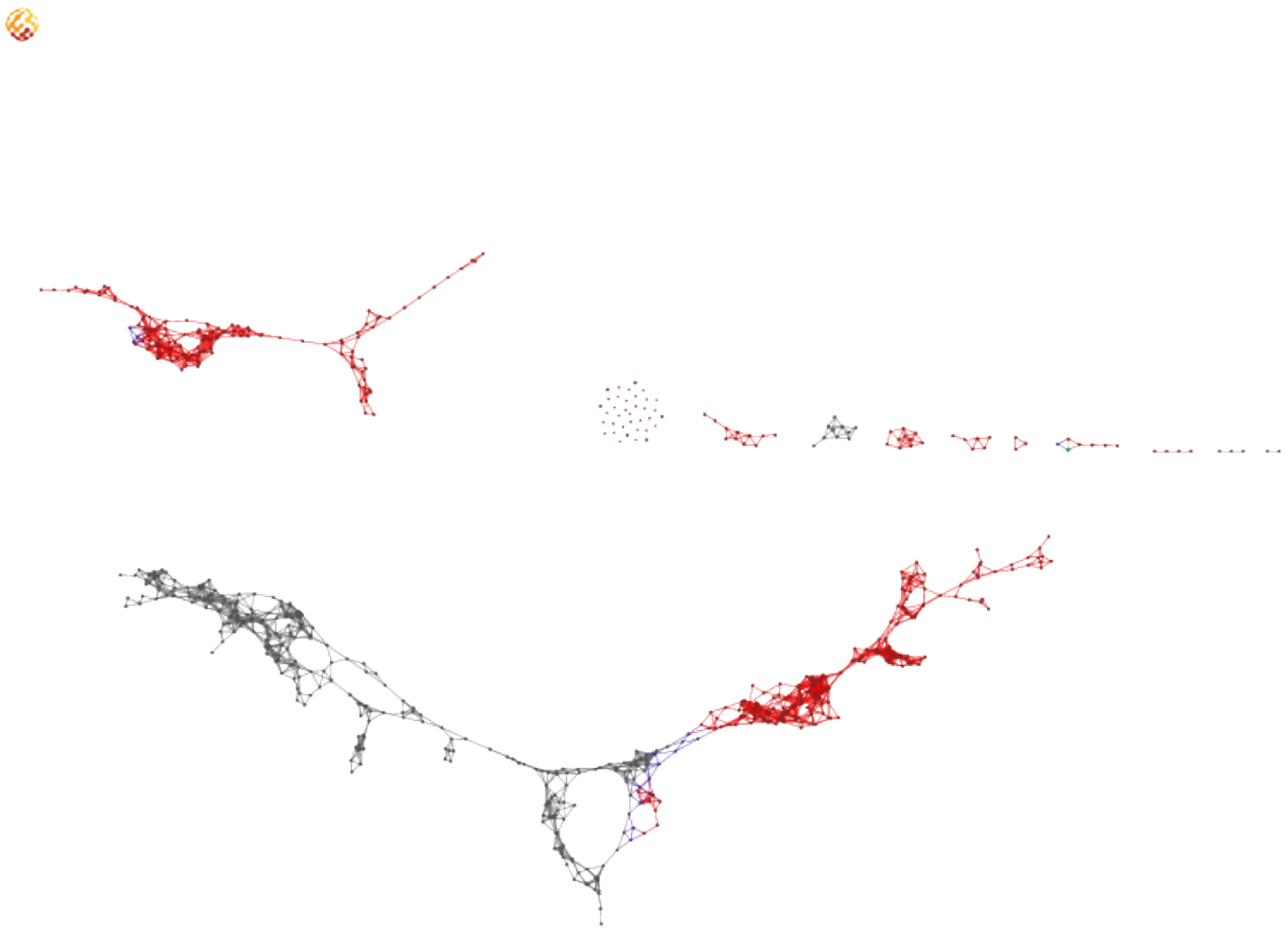
House of Representative.

1993



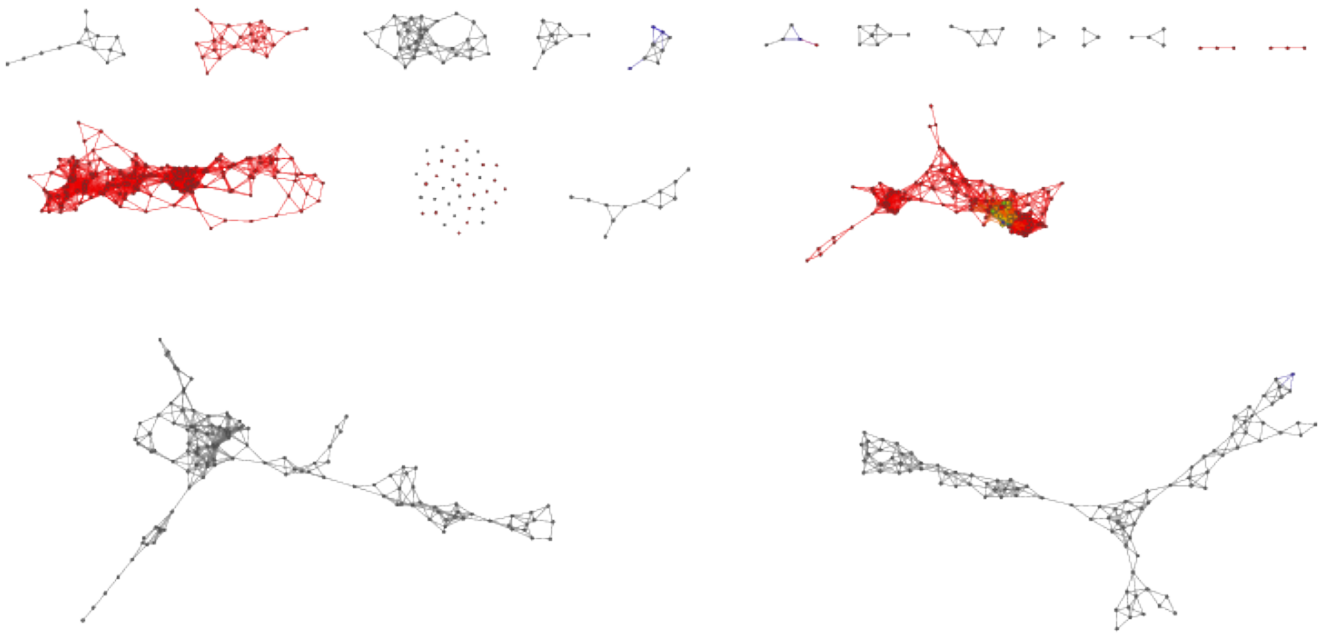
House of Representative.

1994



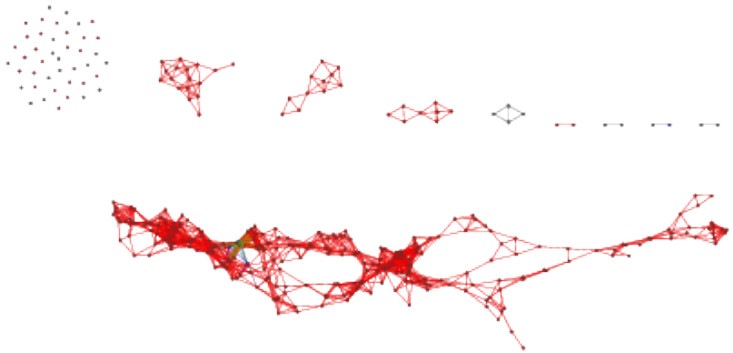
House of Representative.

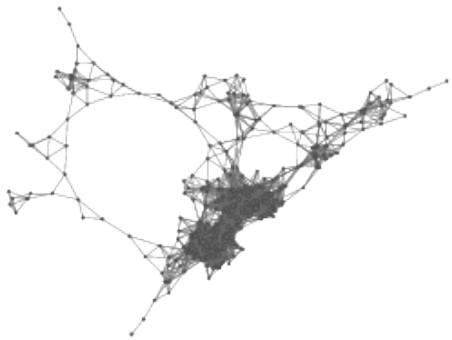
1995



House of Representative.

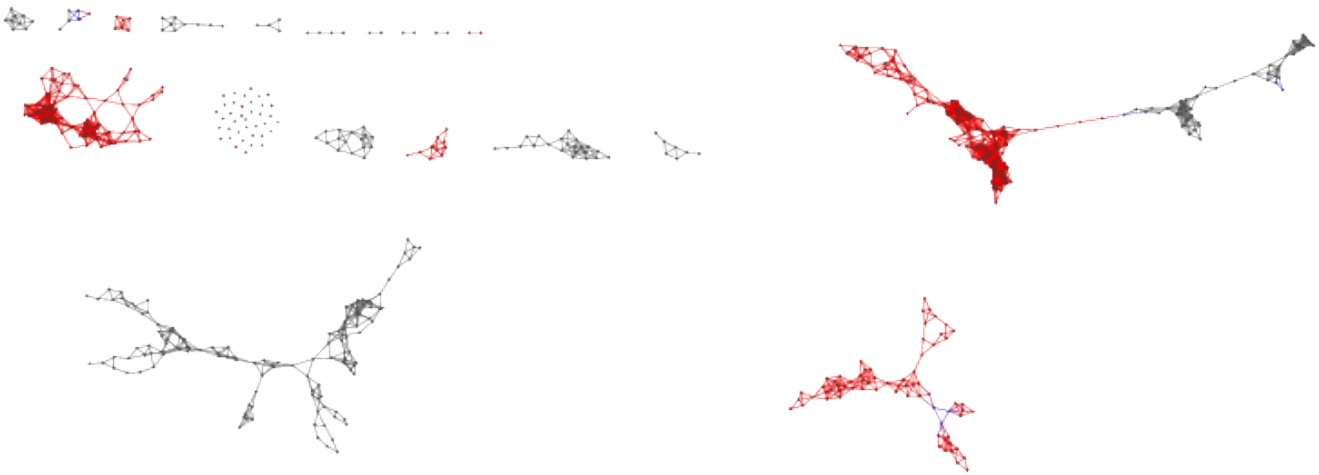
1996





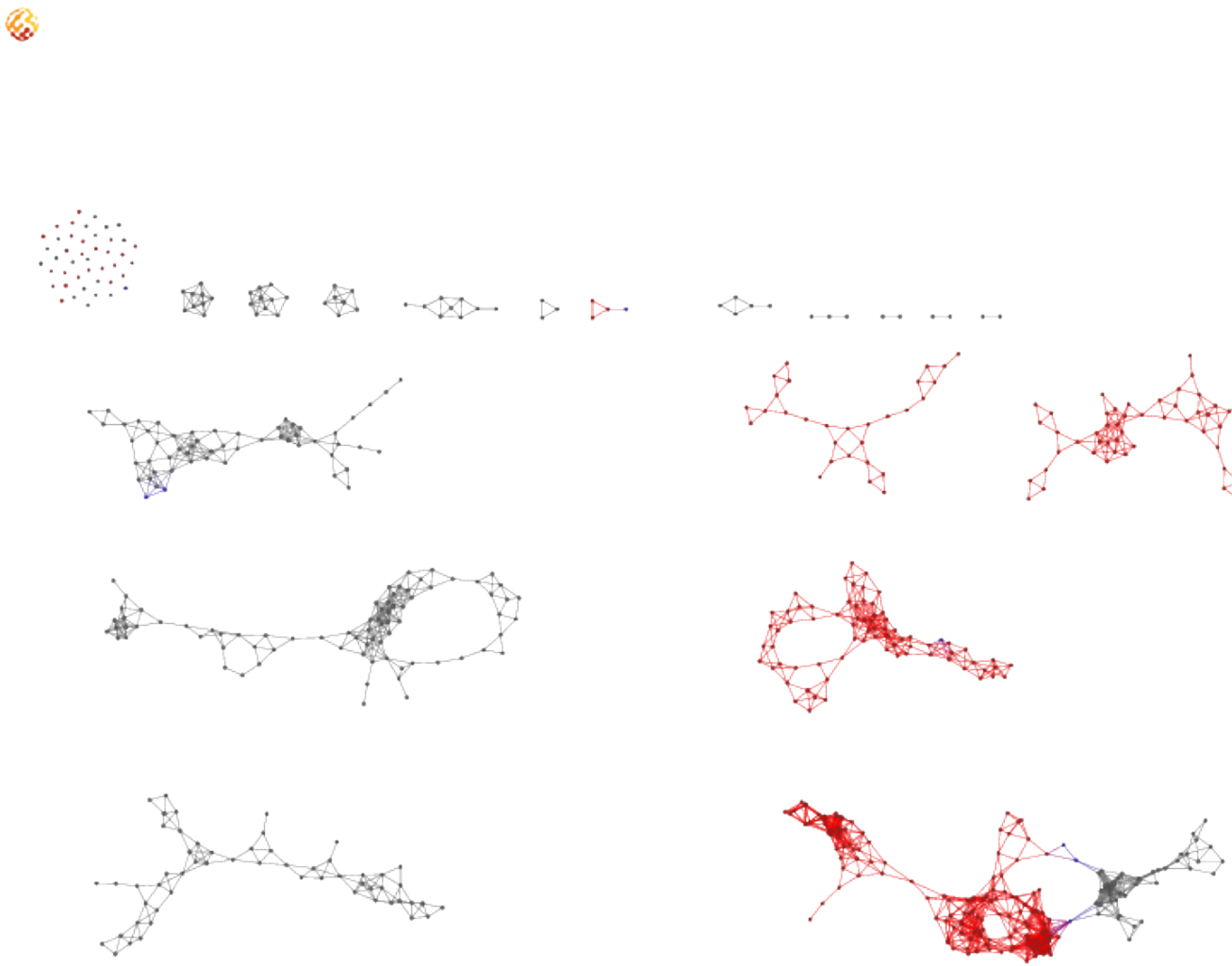
House of Representative.

1997



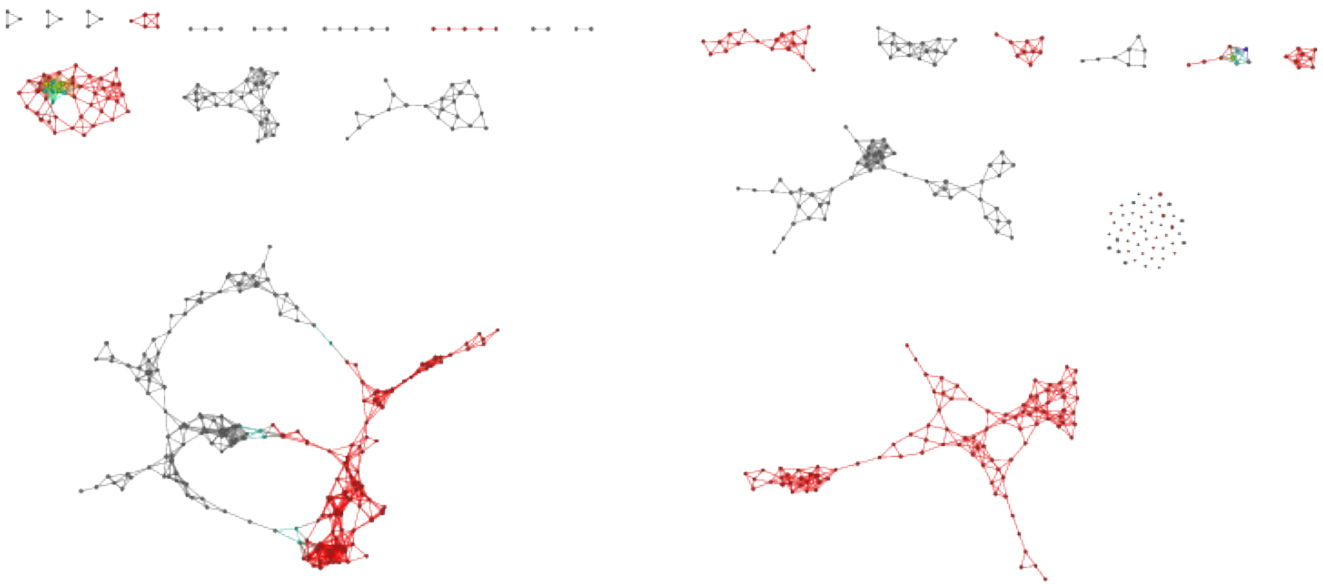
House of Representative.

1998



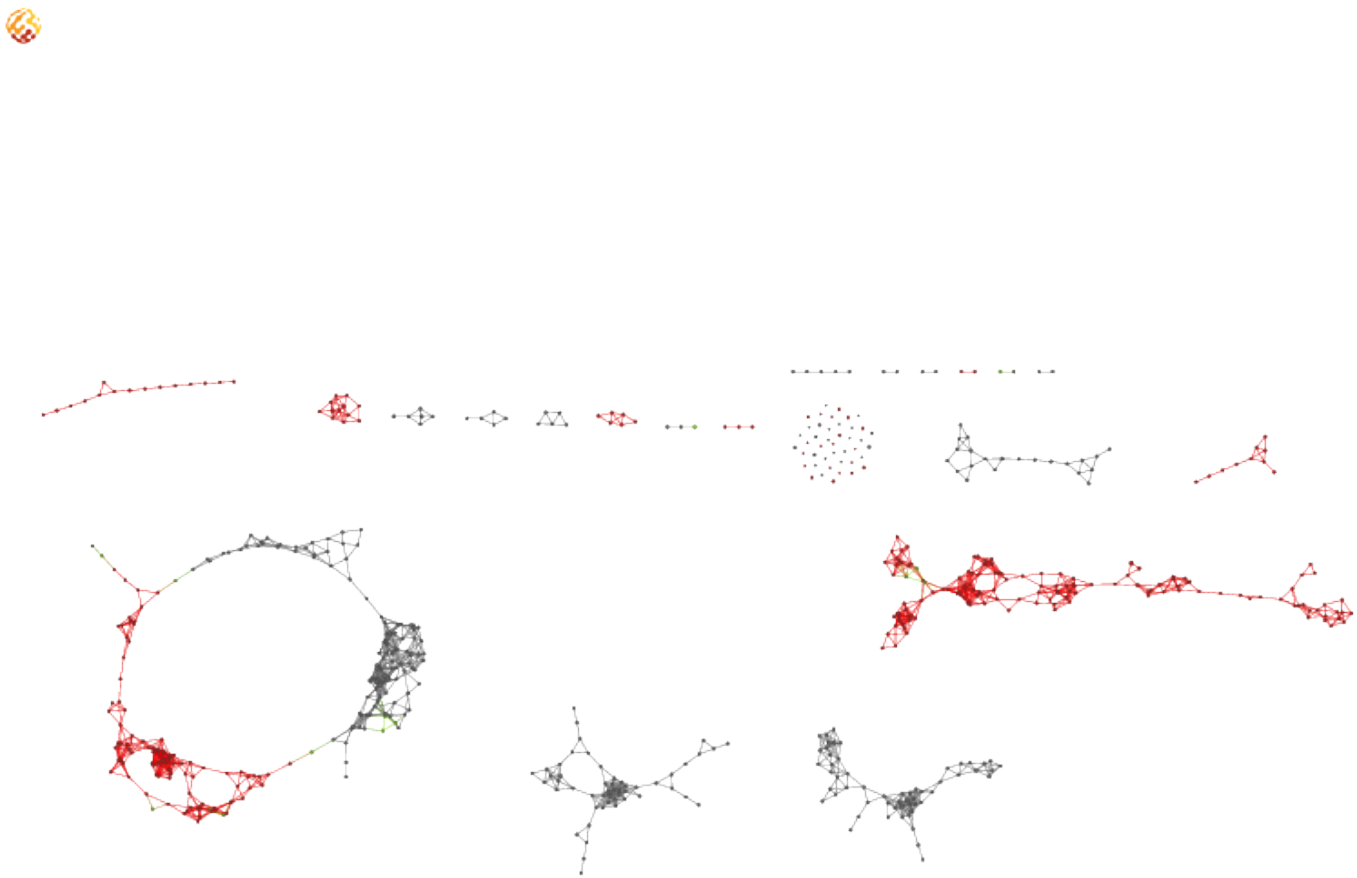
House of Representative.

1999



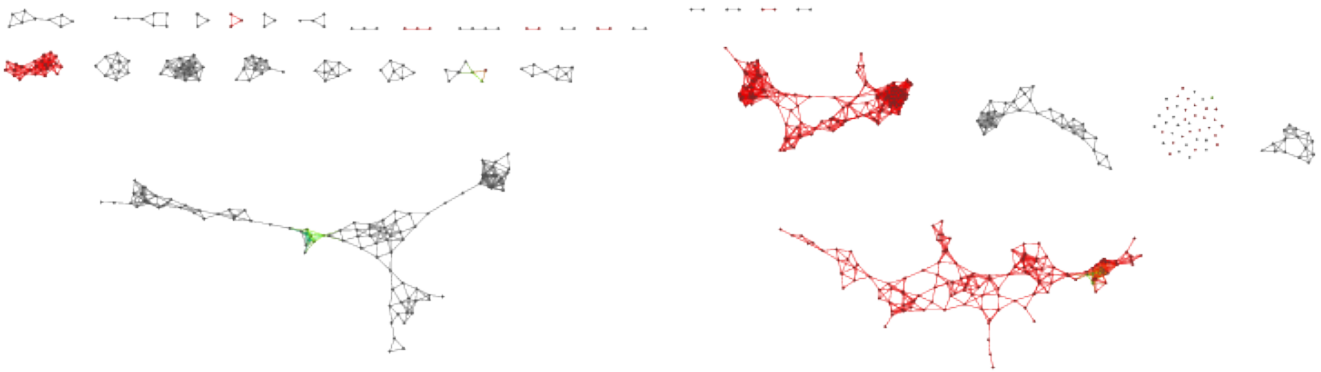
House of Representative.

2000



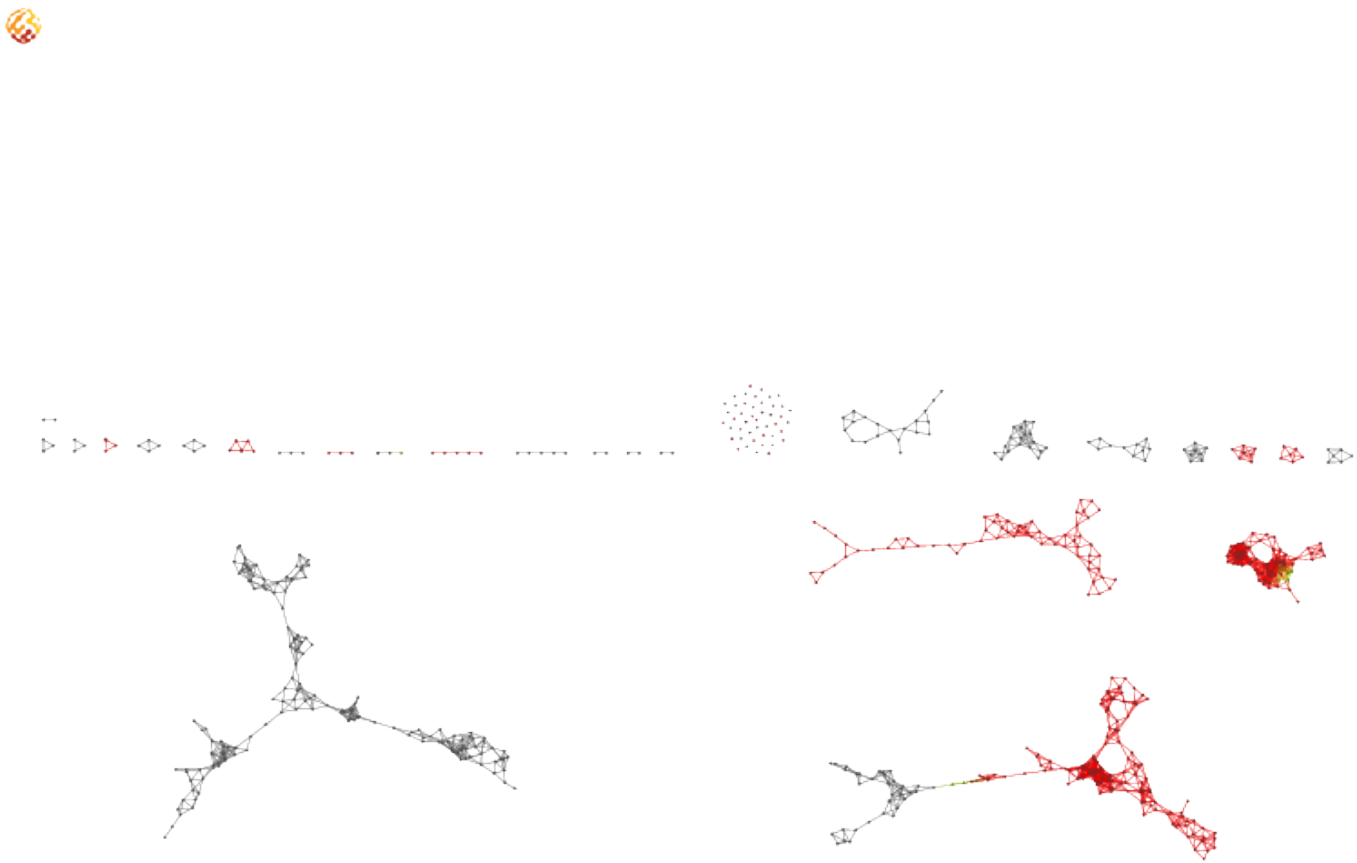
House of Representative.

2001



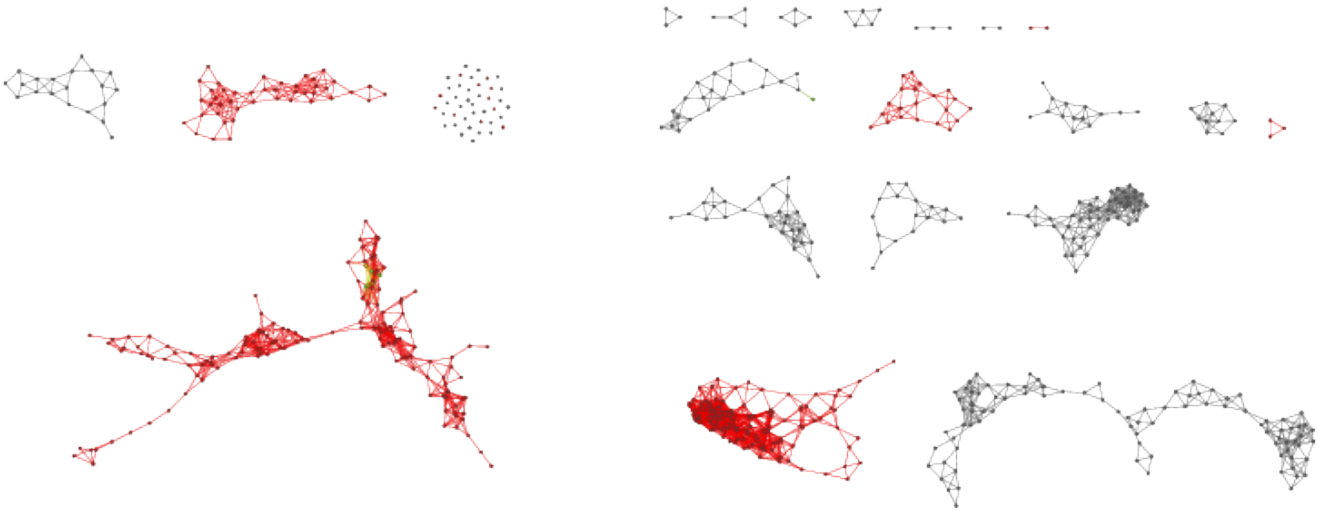
House of Representative.

2002



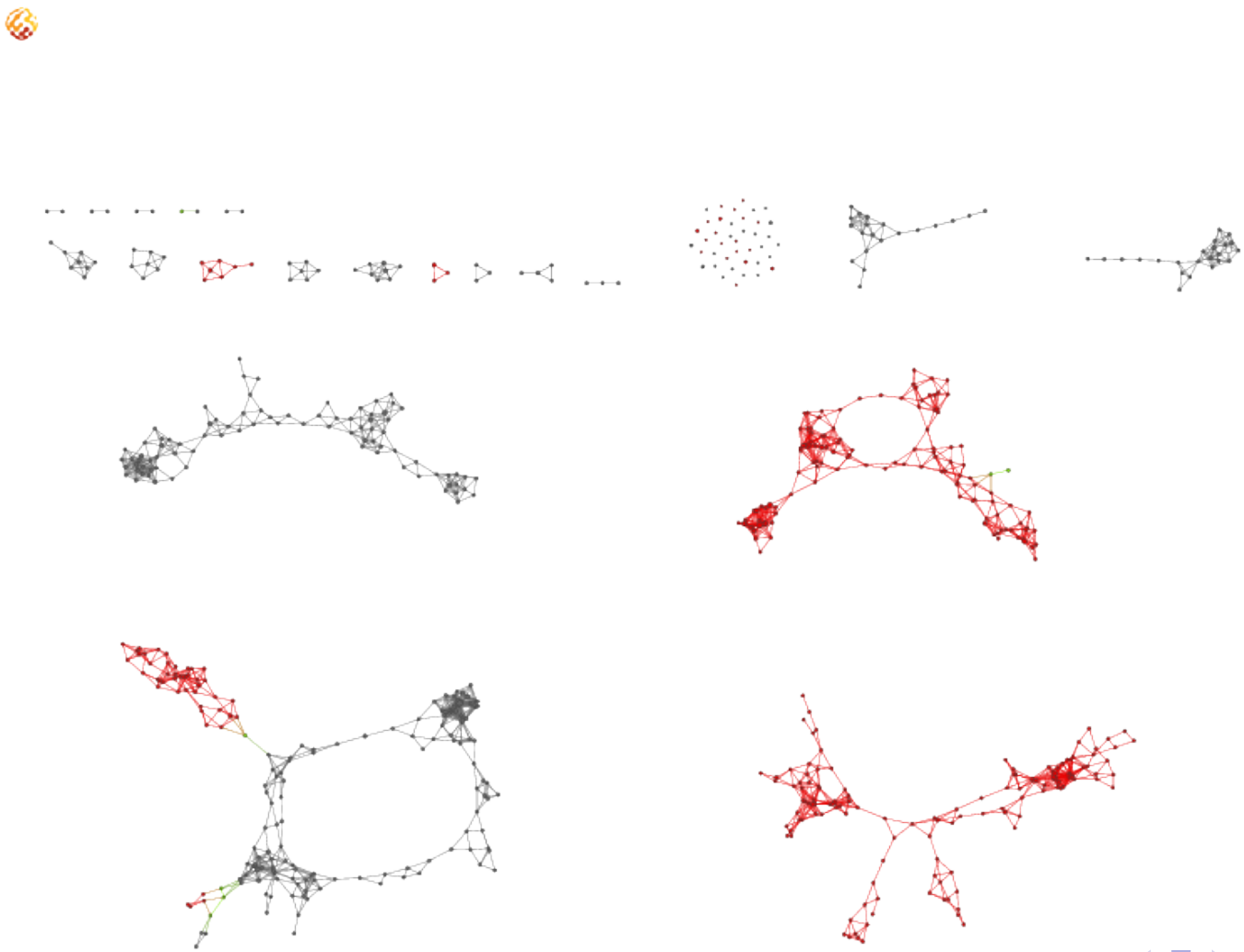
House of Representative.

2003



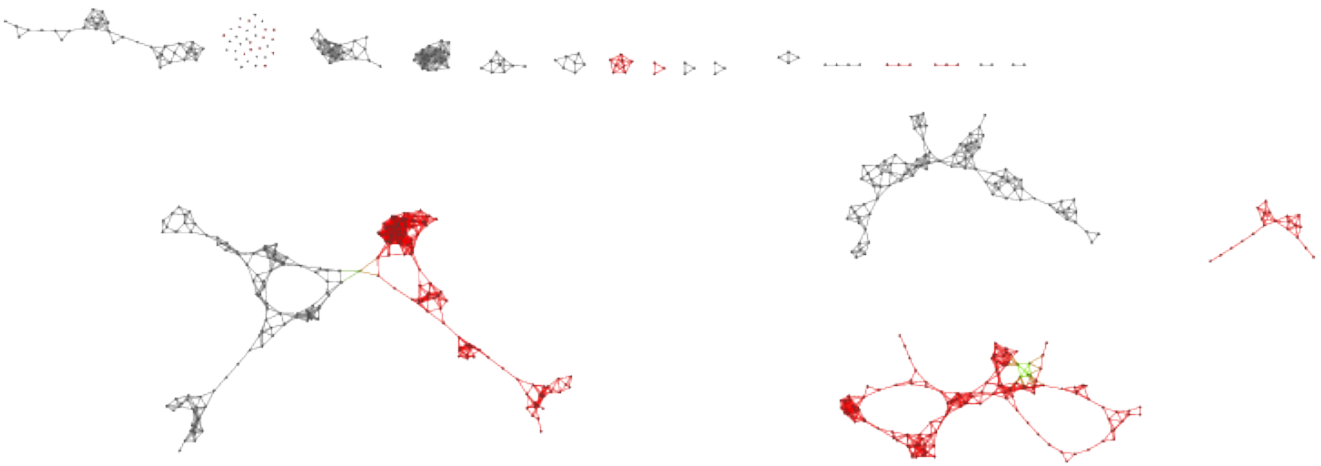
House of Representative.

2004



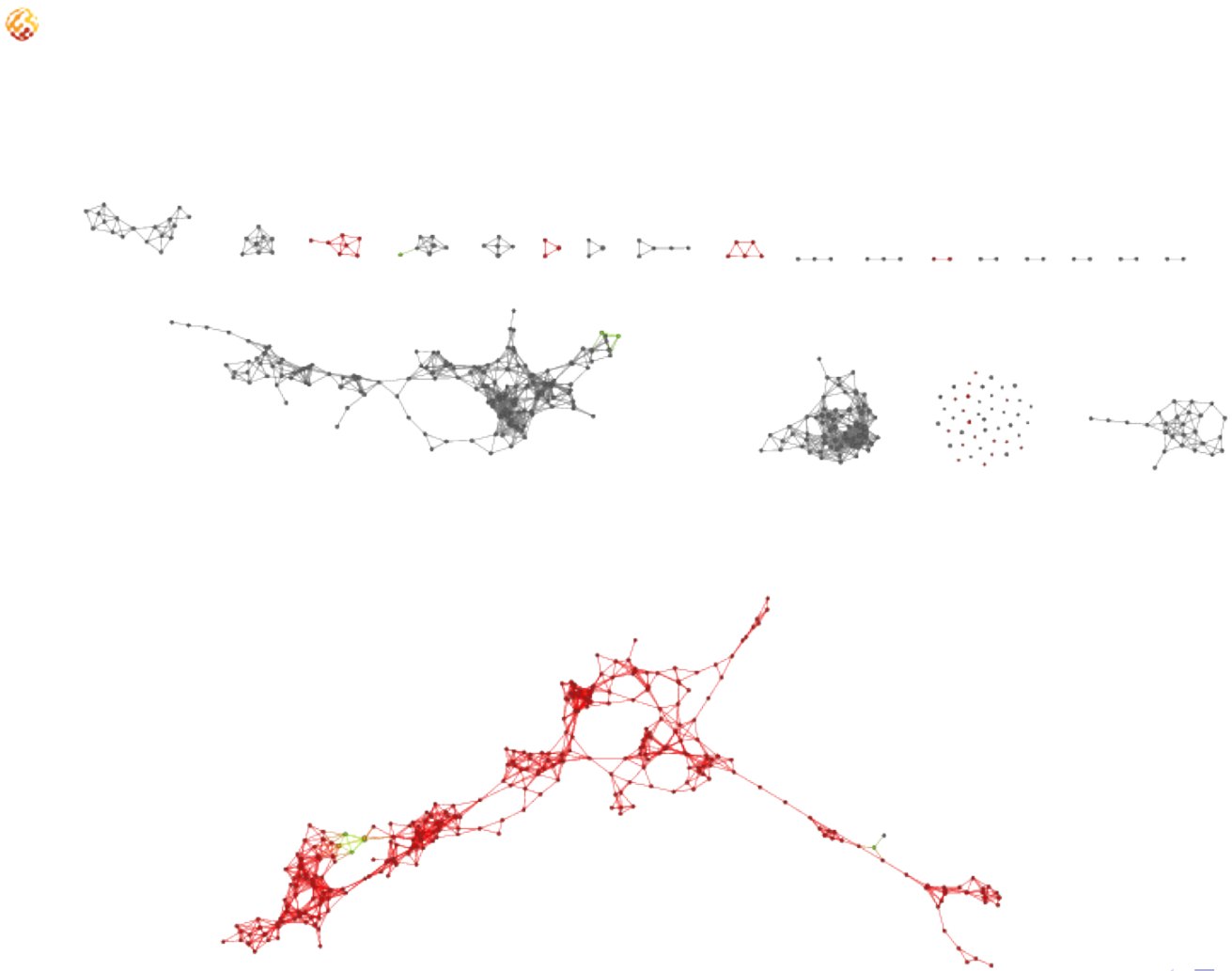
House of Representative.

2005



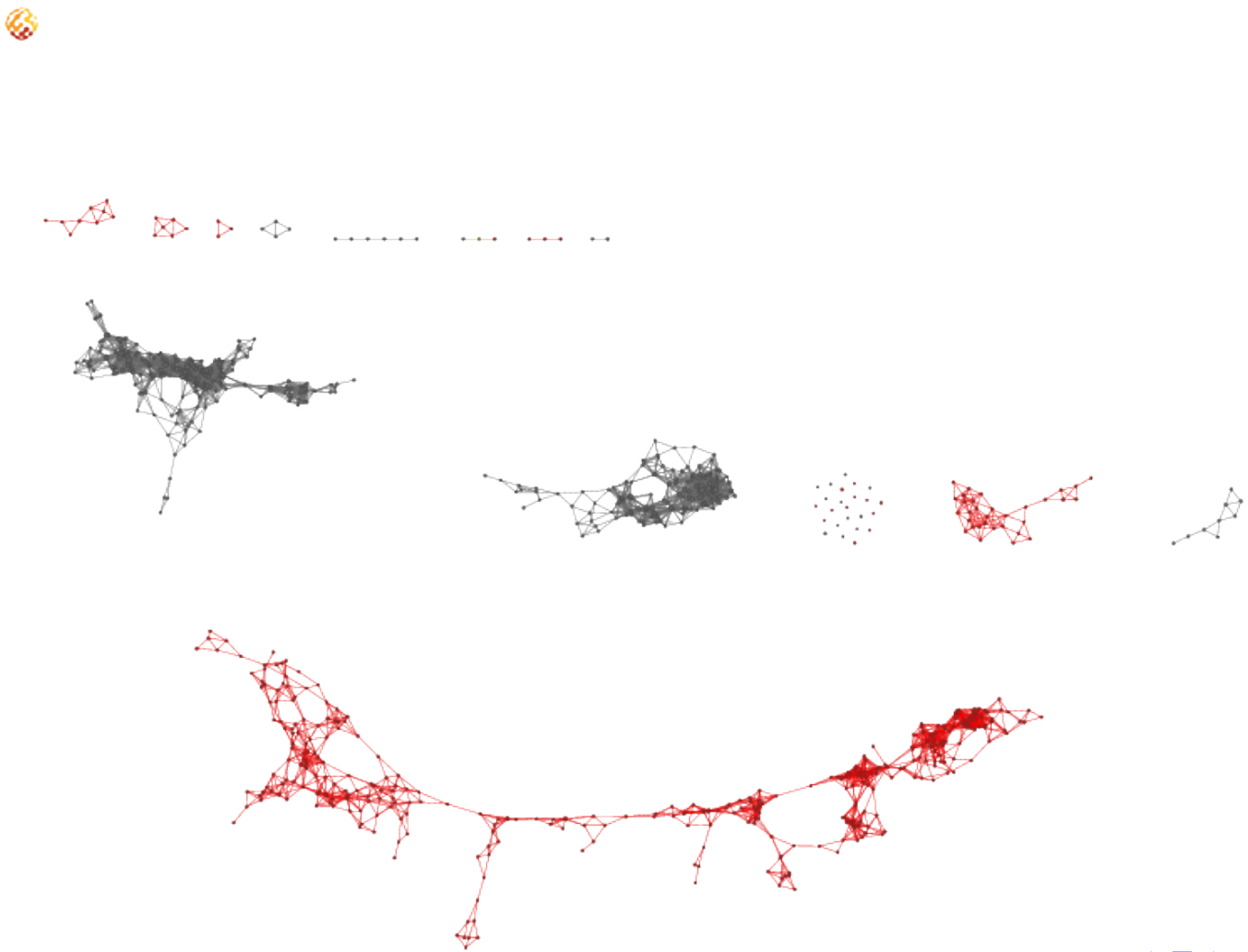
House of Representative.

2006



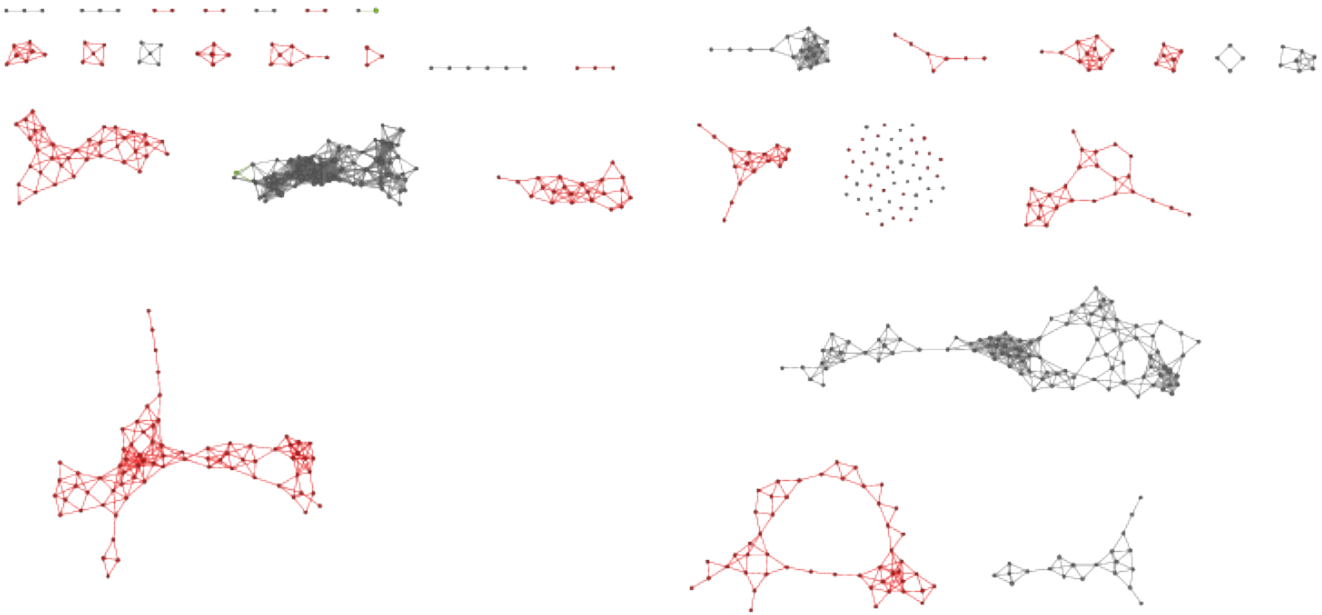
House of Representative.

2007



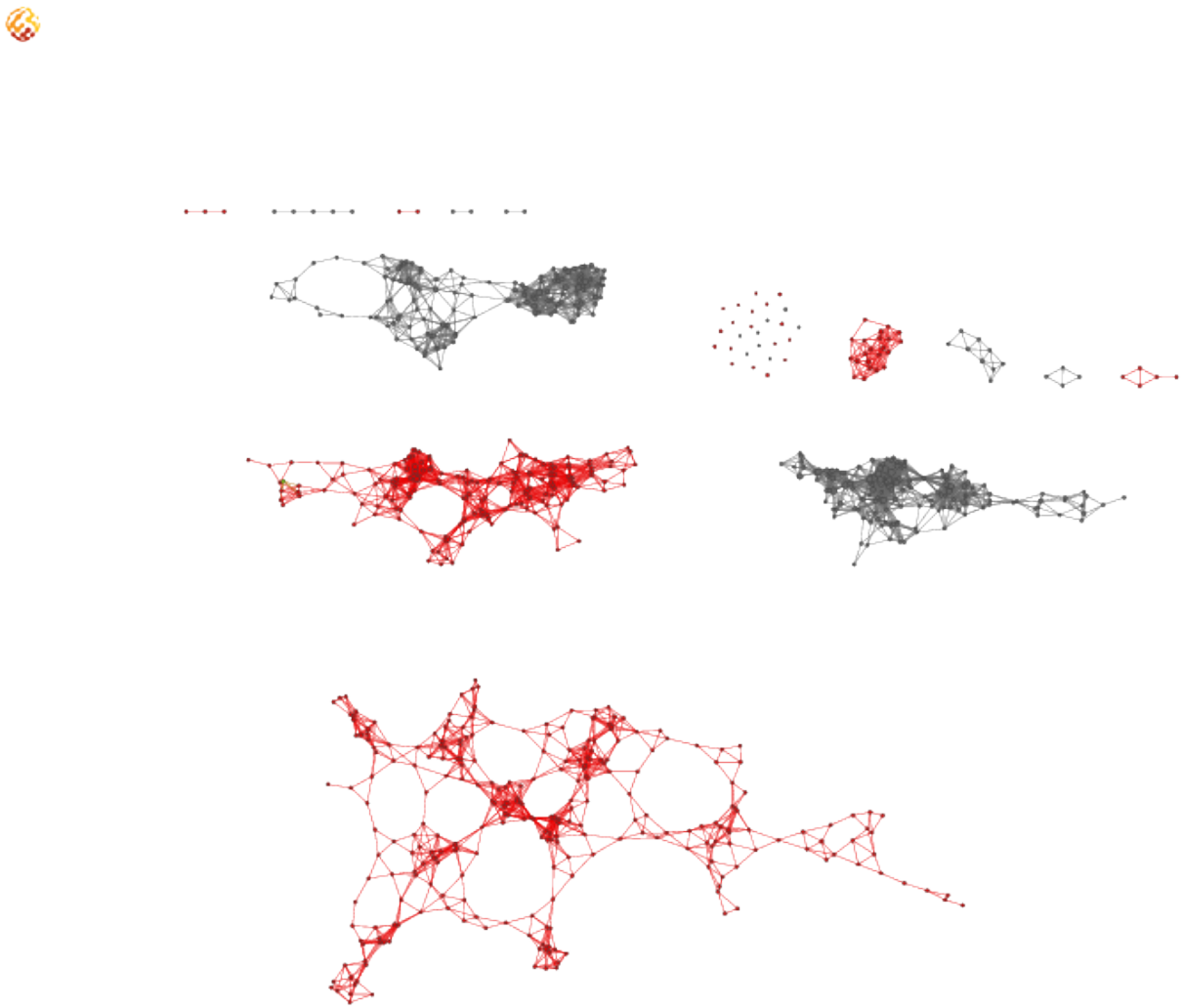
House of Representative.

2008



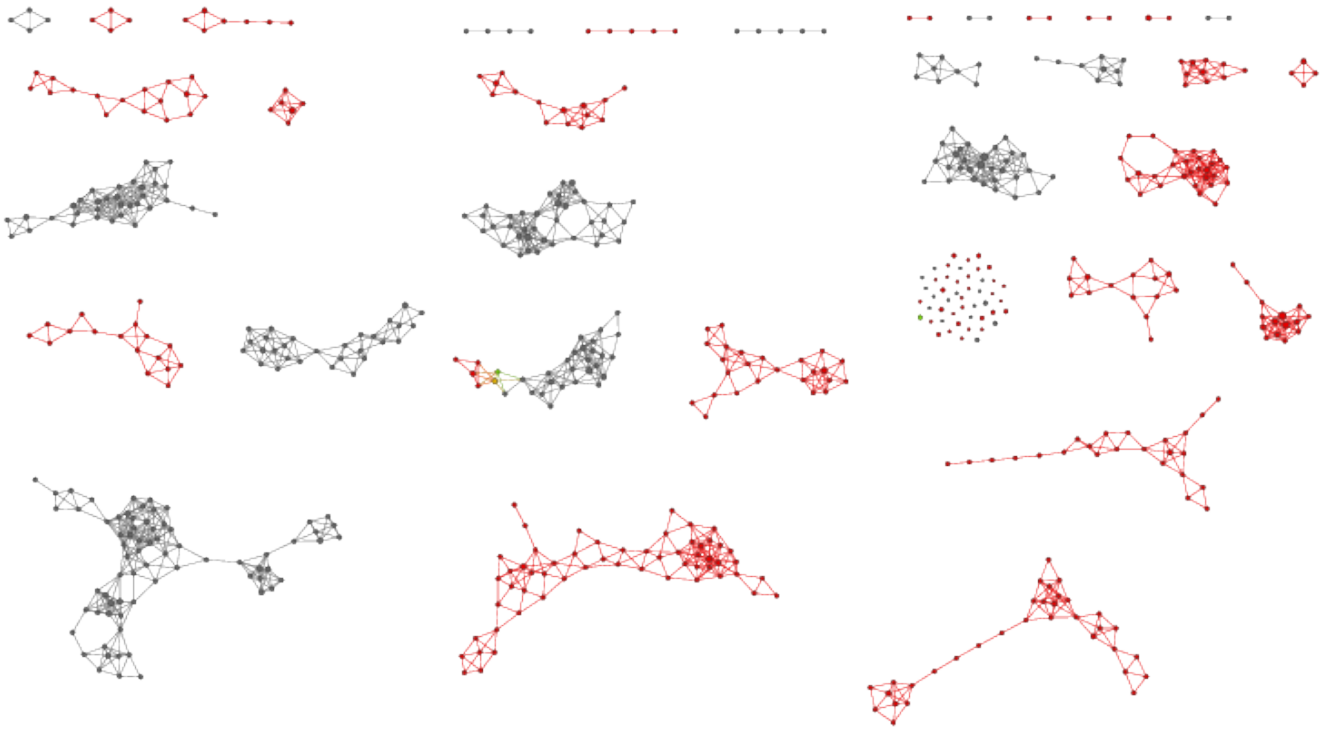
House of Representative.

2009



House of Representative.

2010



Questions?

Future directions

- More parliaments. We have (mostly un-analyzed) data for:
  - Sweden
  - European Union
  - Canada
  - US Senate
- Network analysis: who co-sponsors with whom, what are the parliamentary cliques?
- Meme spread: who takes over turns of phrase from whom?