

# Network Analysis and Modeling

Aaron Clauset

 @aaronclauset

Assistant Professor of Computer Science

University of Colorado Boulder

External Faculty, Santa Fe Institute

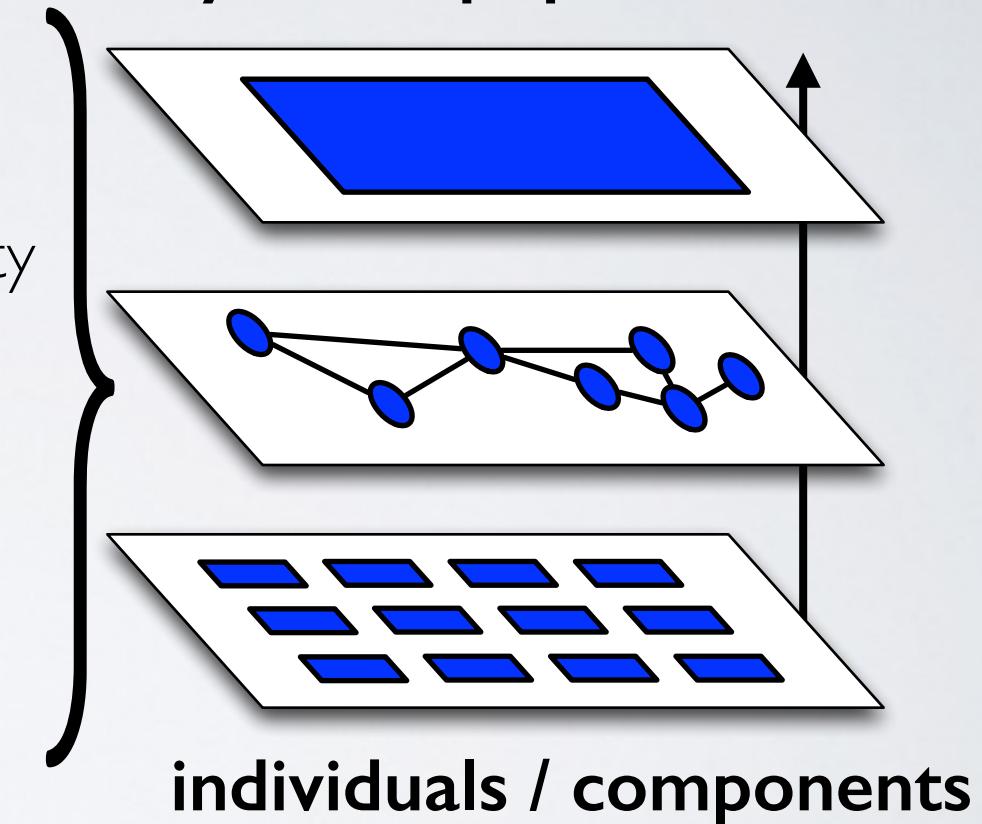
**lecture I: what are networks and how do we talk about them?**

**what are networks?**

## what are networks?

- an approach
- a mathematical representation
- provide structure to complexity
- *structure above*  
individuals / components
- *structure below*  
system / population

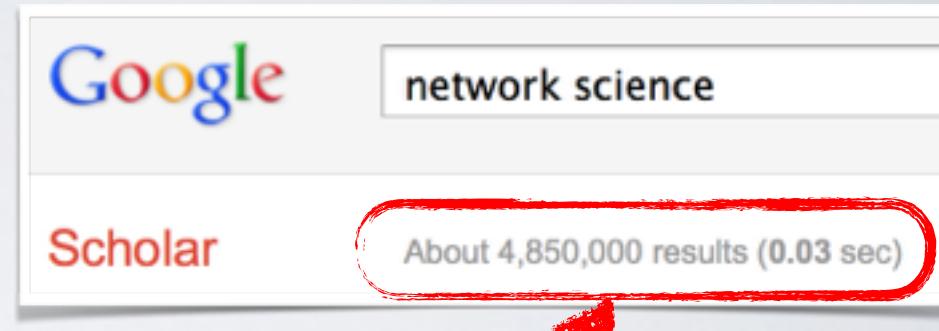
**system / population**



**individuals / components**

## this course

- build intuition
- teach technical methods
- expose key concepts
- highlight big questions
- network mathematics
- network computations
- lots of examples and real data sets



it's a big field now

course webpage: <http://santafe.edu/~aarong/courses/5352/>

The screenshot shows a web browser window with the URL [tuvalu.santafe.edu/~aarong/courses/5352/](http://tuvalu.santafe.edu/~aarong/courses/5352/) in the address bar. The page has a red header bar with the word "Teaching". Below the header, there is course information: "Network Analysis and Modeling" and "CSCI 5352, Fall 2016". It also lists the time as "Monday and Wednesday, 9:30am - 10:45am" and the room as "ECCS 1B12". Below this, there is information about the instructor: Aaron Clauset, ECOT 743, office hours from Monday 1:15-2:45pm, and an email address (zzilm.xozfhvg@xolizwl.vwf) which is a link to an Atbash cipher. There are also links for "Syllabus", "Description", "Course work and grading", "Schedule and lecture notes", "Problem sets", and "Supplemental readings". To the right of the text, there are two network visualization diagrams. One is a hierarchical tree-like structure with nodes branching downwards. The other is a complex, dense network graph with many nodes and edges, colored in various shades of red, orange, blue, green, and purple.

**tuvalu.santafe.edu/~aarong/courses/5352/**

# Teaching

Network Analysis and Modeling  
CSCI 5352, Fall 2016

Time: Monday and Wednesday, 9:30am - 10:45am  
Room: ECCS 1B12

Instructor: Aaron Clauset  
Office: [ECOT 743](#)  
Office hours: Monday, 1:15-2:45pm  
Email: [zzilm.xozfhvg@xolizwl.vwf](mailto:zzilm.xozfhvg@xolizwl.vwf) ([an Atbash cipher](#))

[Syllabus](#)

[Description](#)

[Course work and grading](#)

[Schedule and lecture notes](#)

[Problem sets](#)

[Supplemental readings](#)

### Description

Network science is a thriving and increasingly important cross-disciplinary domain that focuses on the representation, analysis and modeling of complex social, biological and technological systems as networks or graphs. Modern data sets often include some kind of network. Nodes can have locations, directions, memory, demographic characteristics, content, and preferences. Edges can have lengths, directions, capacities, costs, durations, and types. And,

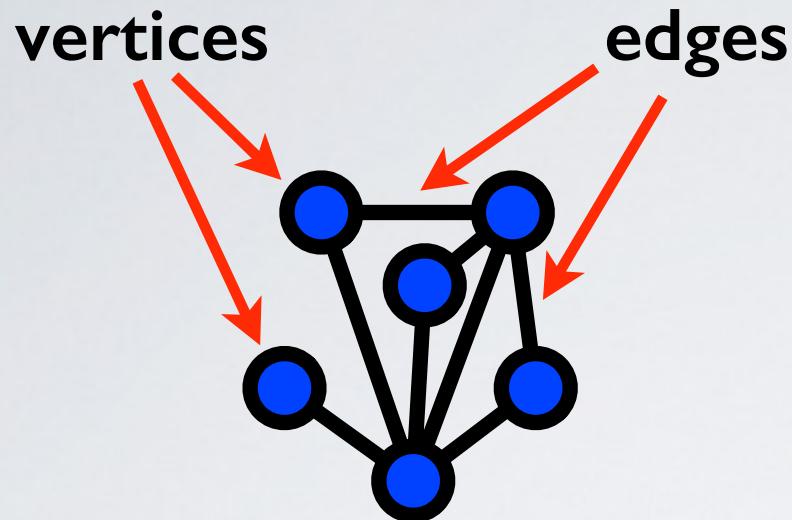
**PROFESSIONAL**

[Research](#)  
[Lab Group](#)  
[Data and Code](#)  
[Teaching](#)  
[Vitae](#)  
[Blog](#)  
[Twitter](#)

**PERSONAL**

[About Me](#)  
[Contact](#)

1. defining a network
2. describing a network



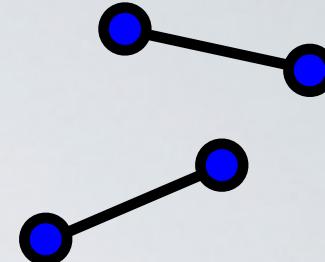
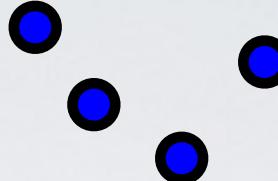
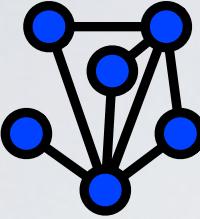
**what is a vertex?**

$V$  distinct objects (vertices / nodes / actors)

**when are two vertices connected?**

$$E \subseteq V \times V$$

pairwise relations (edges / links / ties)

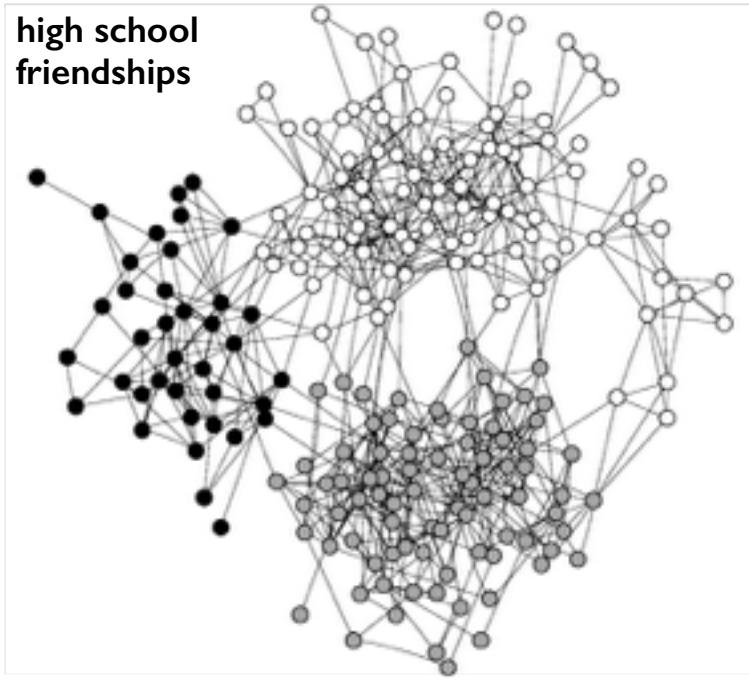


	<b>network</b>	<b>vertex</b>	<b>edge</b>
informational	Internet(1)	computer	IP network adjacency
	Internet(2)	autonomous system (ISP)	BGP connection
	software	function	function call
	World Wide Web	web page	hyperlink
	documents	article, patent, or legal case	citation
transportation	power grid transmission	generating or relay station	transmission line
	rail system	rail station	railroad tracks
	road network(1)	intersection	pavement
	road network(2)	named road	intersection
social	airport network	airport	non-stop flight
	friendship network	person	friendship
	sexual network	person	intercourse
biological	metabolic network	metabolite	metabolic reaction
	protein-interaction network	protein	bonding
	gene regulatory network	gene	regulatory effect
	neuronal network	neuron	synapse
	food web	species	predation or resource transfer

# social networks

**vertex:** a person

**edge:** friendship, collaborations, sexual contacts, communication, authority, exchange, etc.

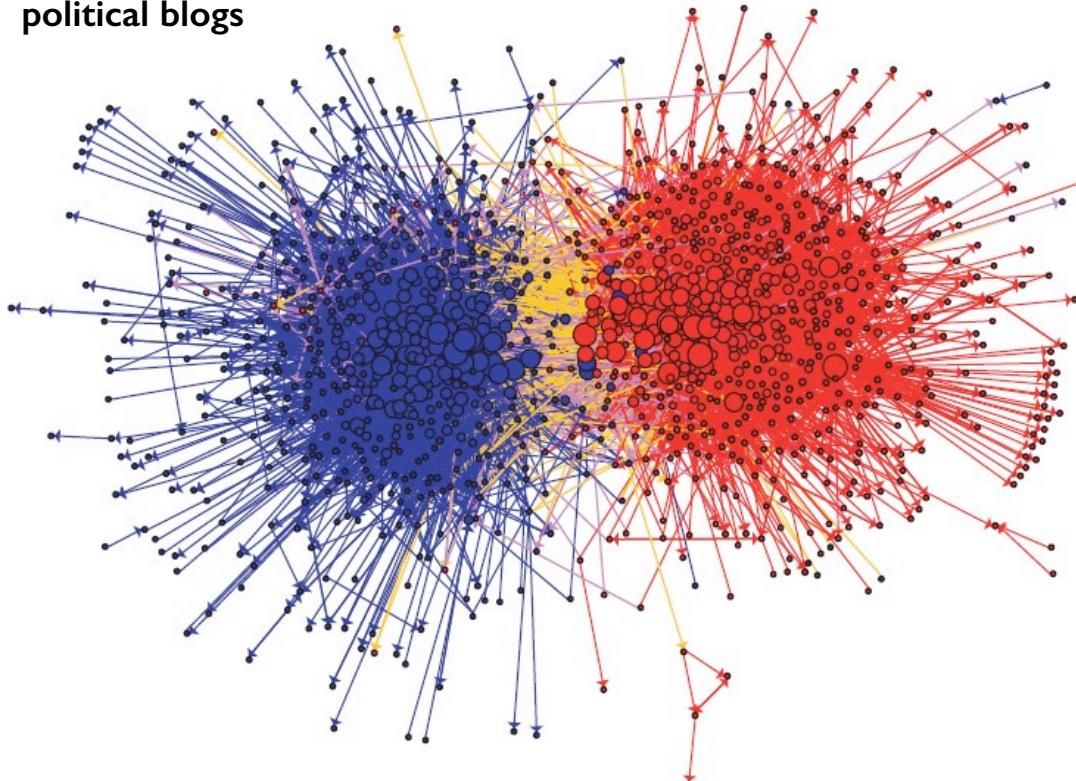


# information networks

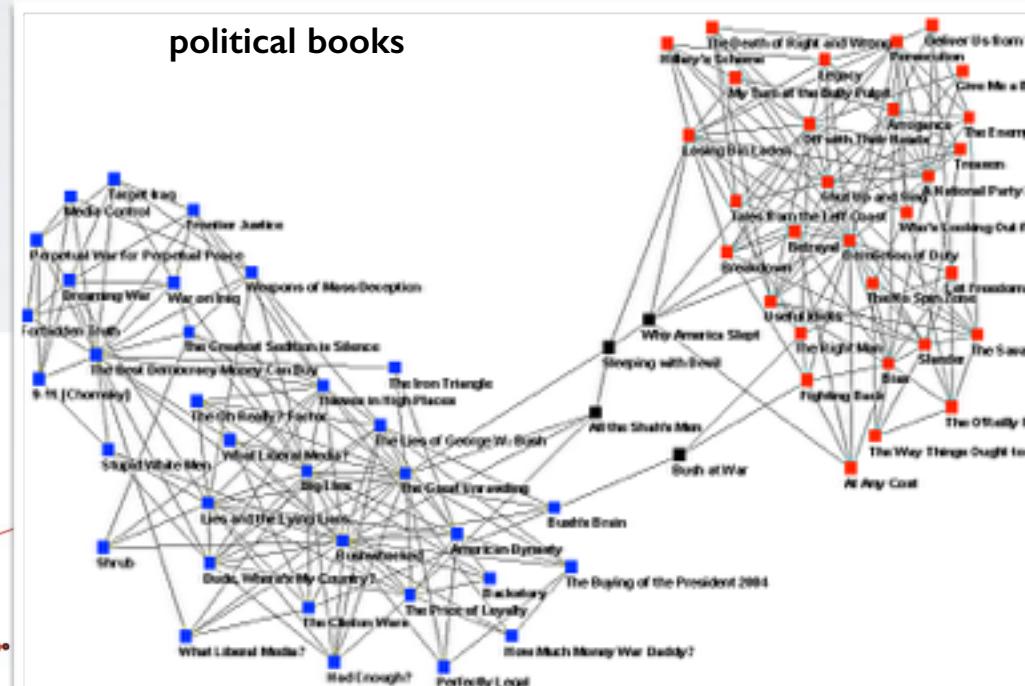
**vertex:** books, blogs, webpages, etc.

**edge:** citations, hyperlinks,  
recommendations, similarity, etc.

political blogs



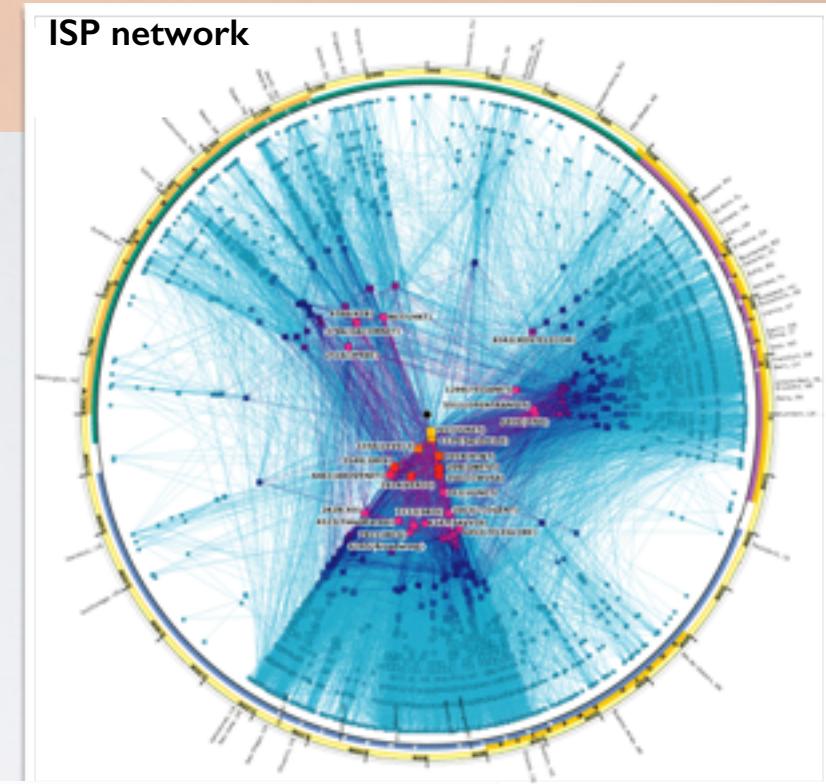
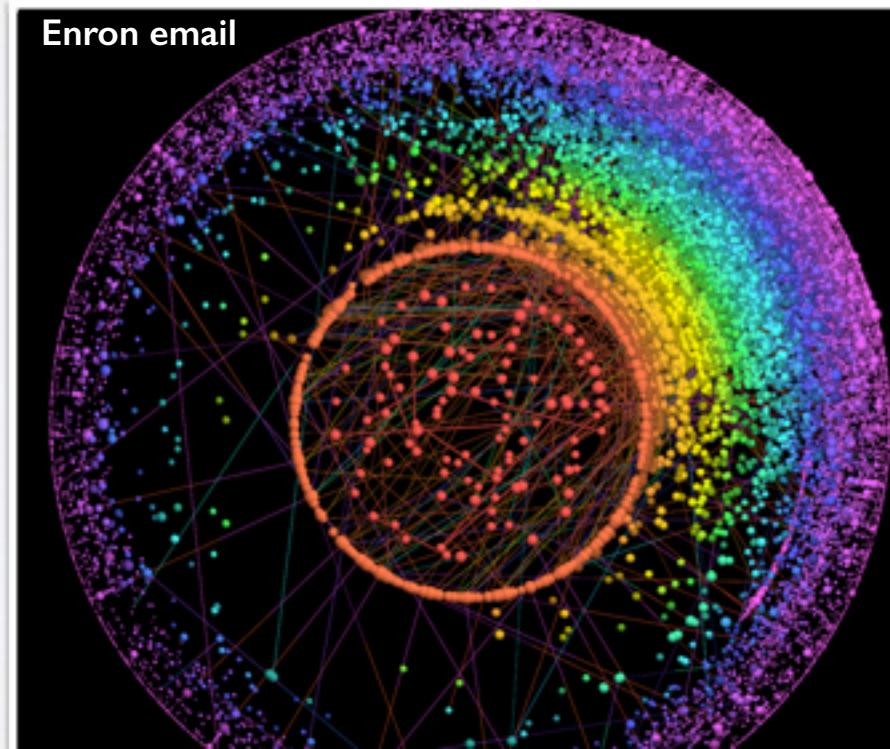
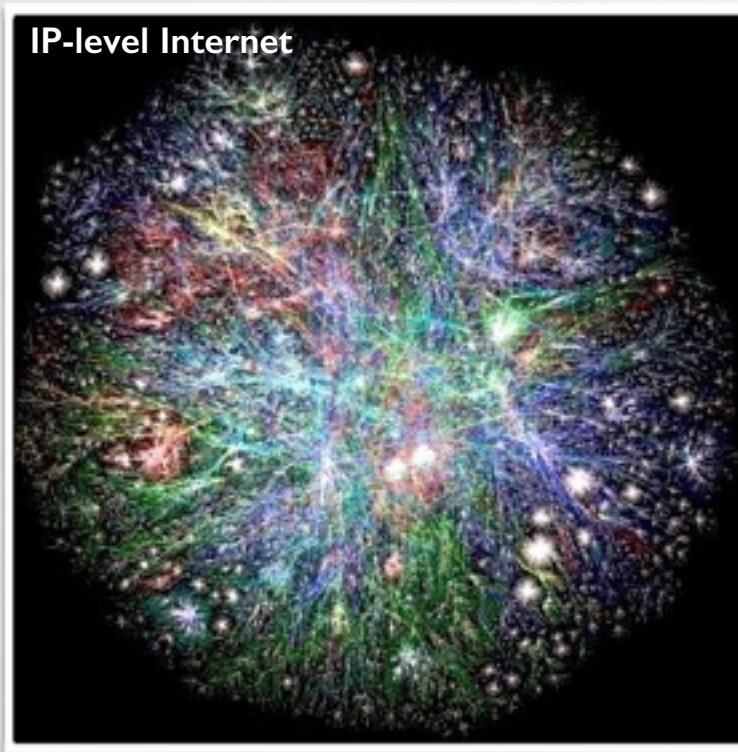
political books



# communication networks

**vertex:** network router, ISP, email address, mobile phone number, etc.

**edge:** exchange of information



# transportation networks

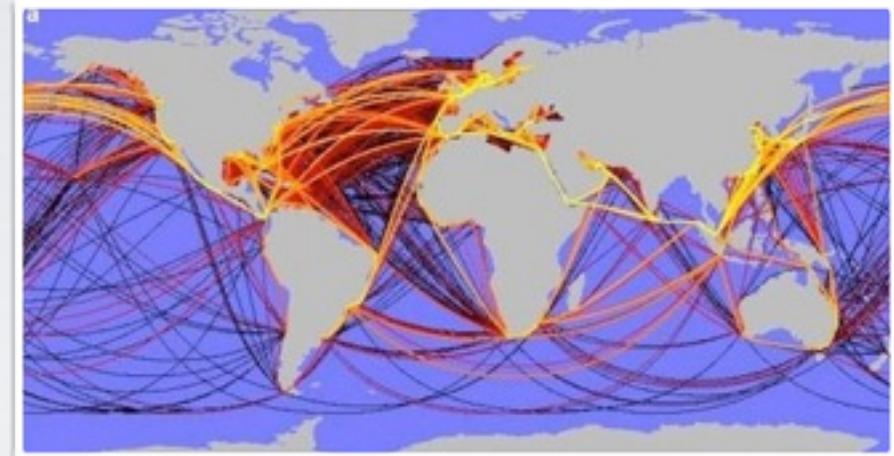
**vertex:** city, airport, junction, railway station, river confluence, etc.

**edge:** physical transportation of material



US Interstates

global shipping



global air traffic

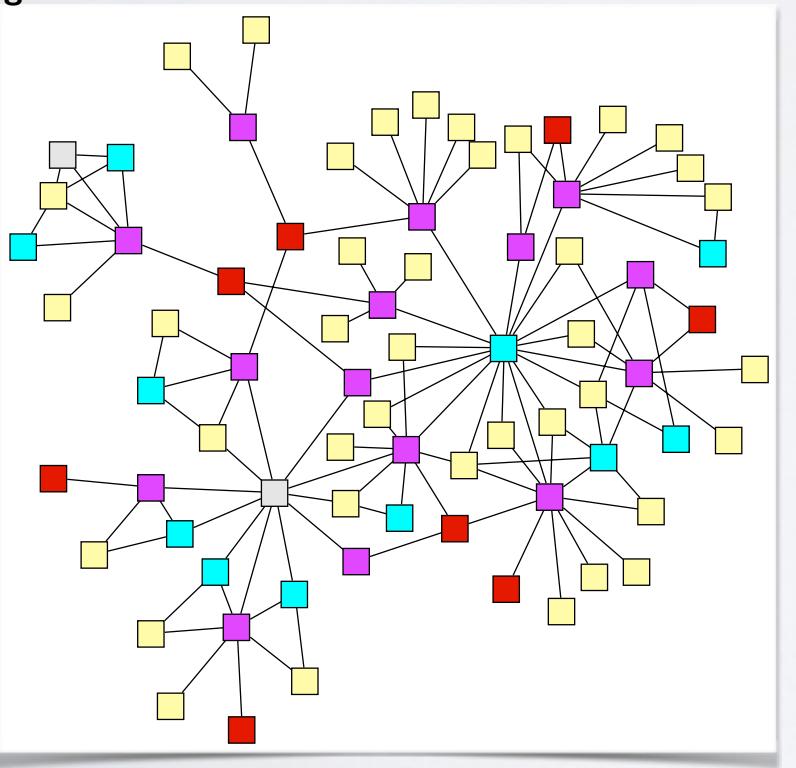


# biological networks

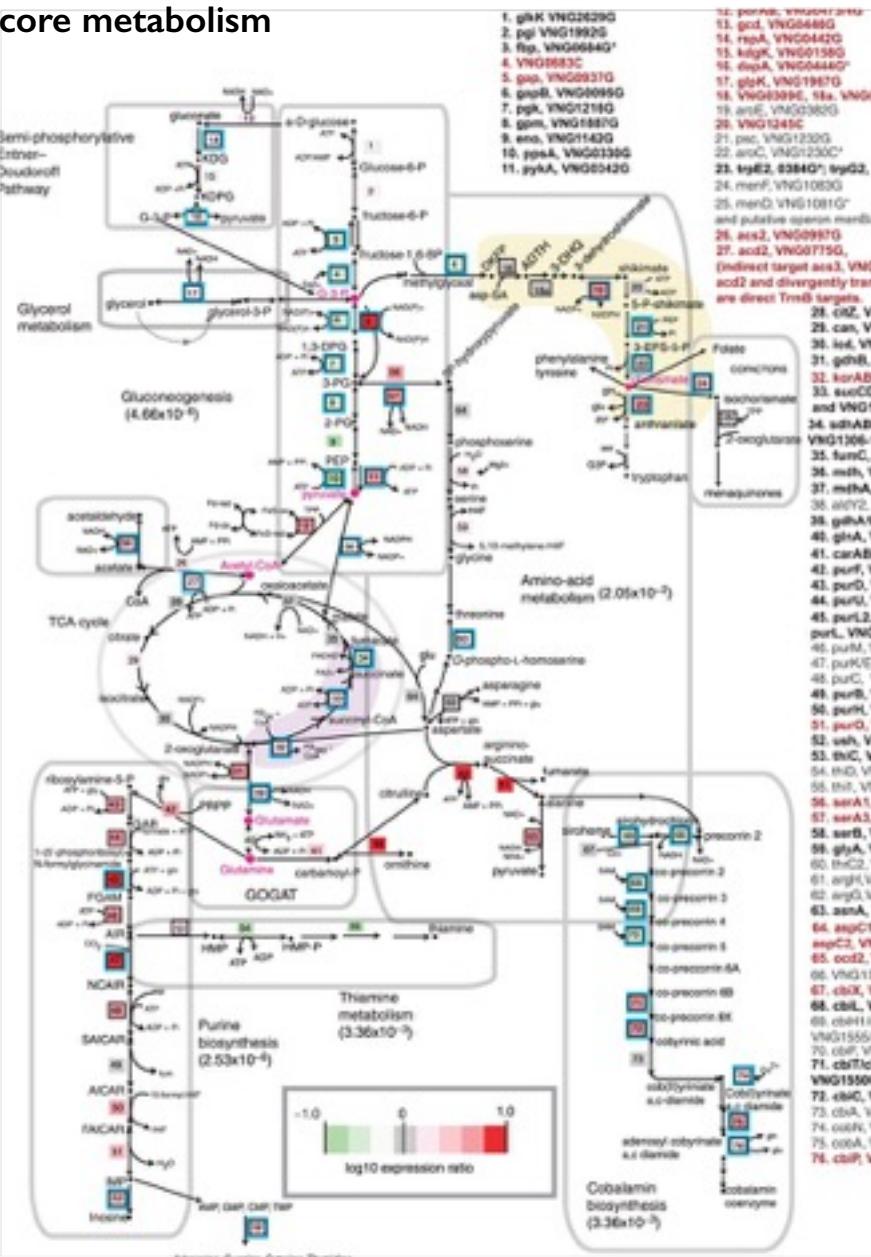
**vertex:** species, metabolic, protein, gene, neuron, etc.

**edge:** predation, chemical reaction, binding, regulation, activation, etc.

grassland foodweb



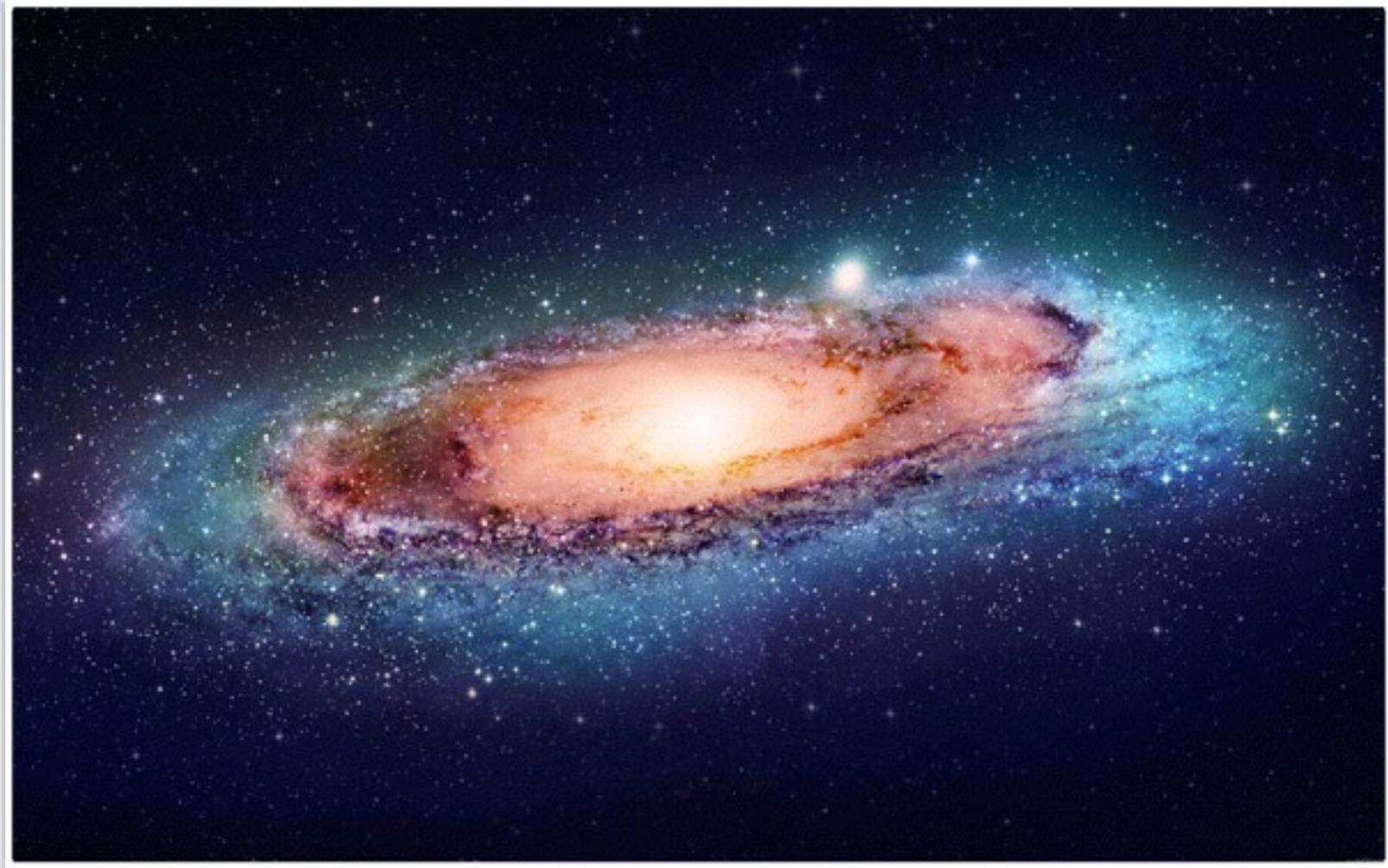
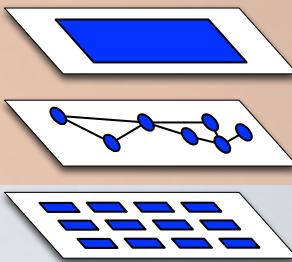
core metabolism



**what's a network?**

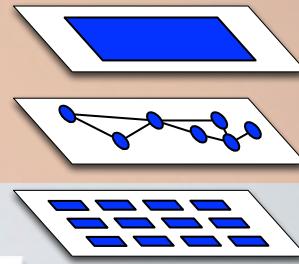
**pop quiz**

# what's a network?



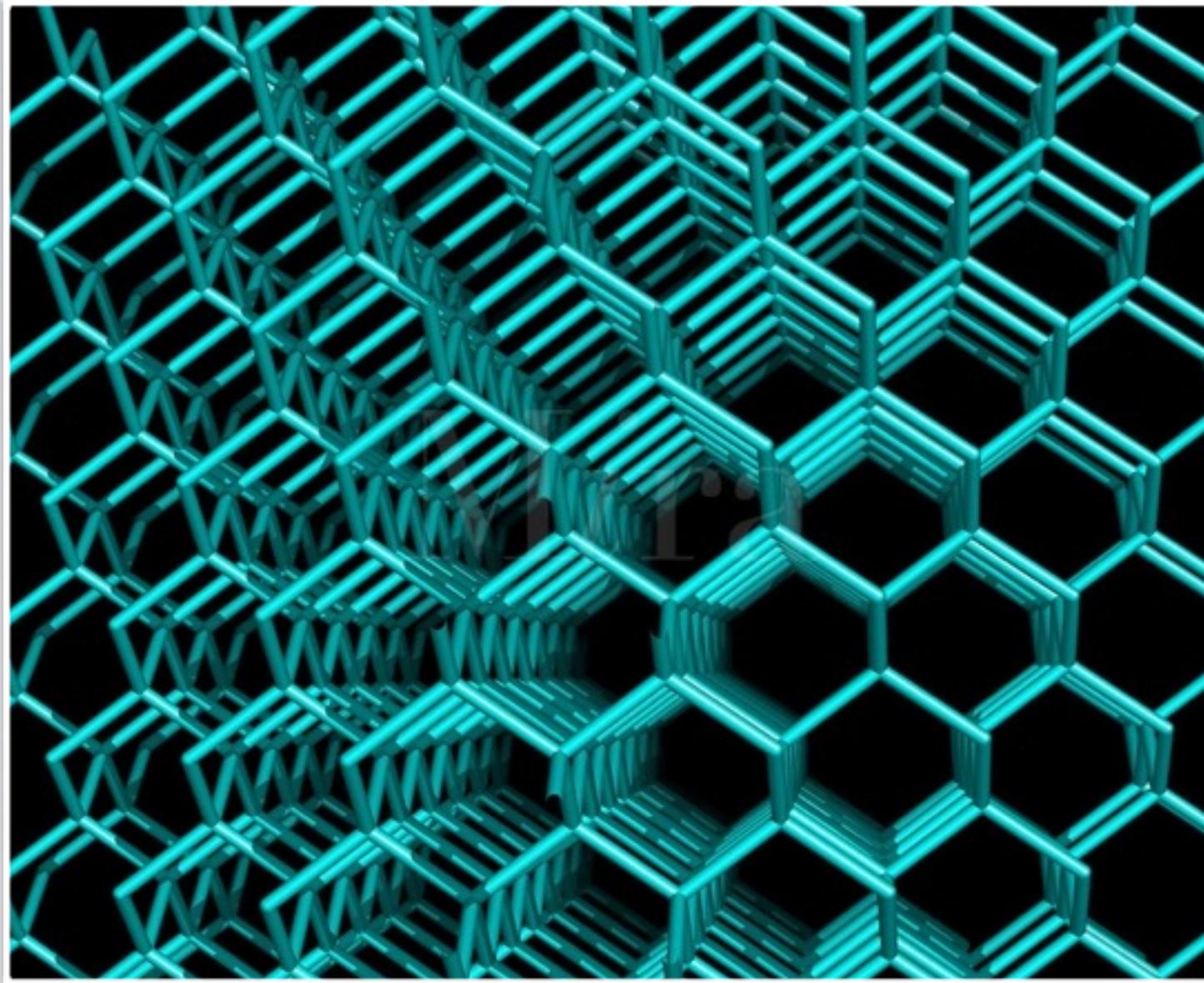
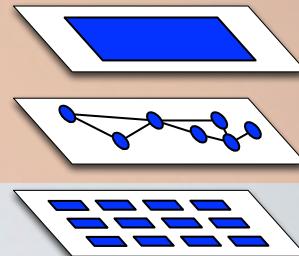
Andromeda galaxy

# what's a network?



cauliflower fractal

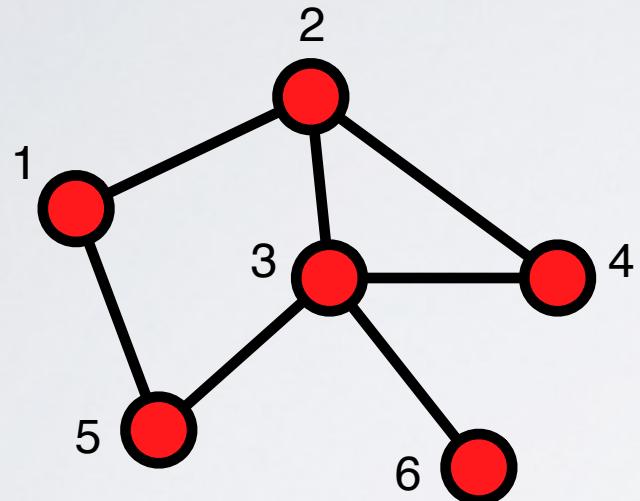
# what's a network?



diamond lattice

# **representing networks**

# *a simple network*

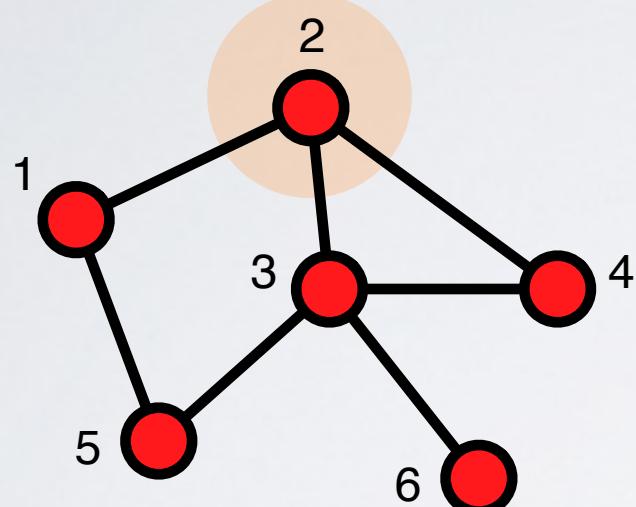


undirected

unweighted

no self-loops

# a simple network



undirected

unweighted

no self-loops

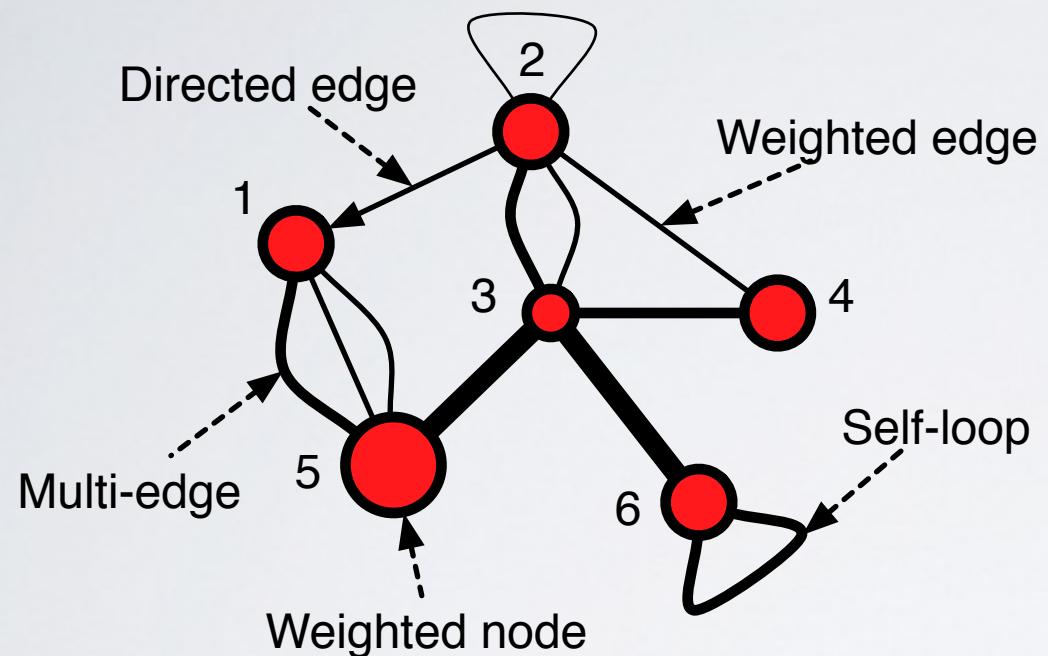
adjacency matrix

$A$	1	2	3	4	5	6
1	0	1	0	0	1	0
2	1	0	1	1	0	0
3	0	1	0	1	1	1
4	0	1	1	0	0	0
5	1	0	1	0	0	0
6	0	0	1	0	0	0

adjacency list

$A$
$1 \rightarrow \{2, 5\}$
$2 \rightarrow \{1, 3, 4\}$
$3 \rightarrow \{2, 4, 5, 6\}$
$4 \rightarrow \{2, 3\}$
$5 \rightarrow \{1, 3\}$
$6 \rightarrow \{3\}$

# a less simple network

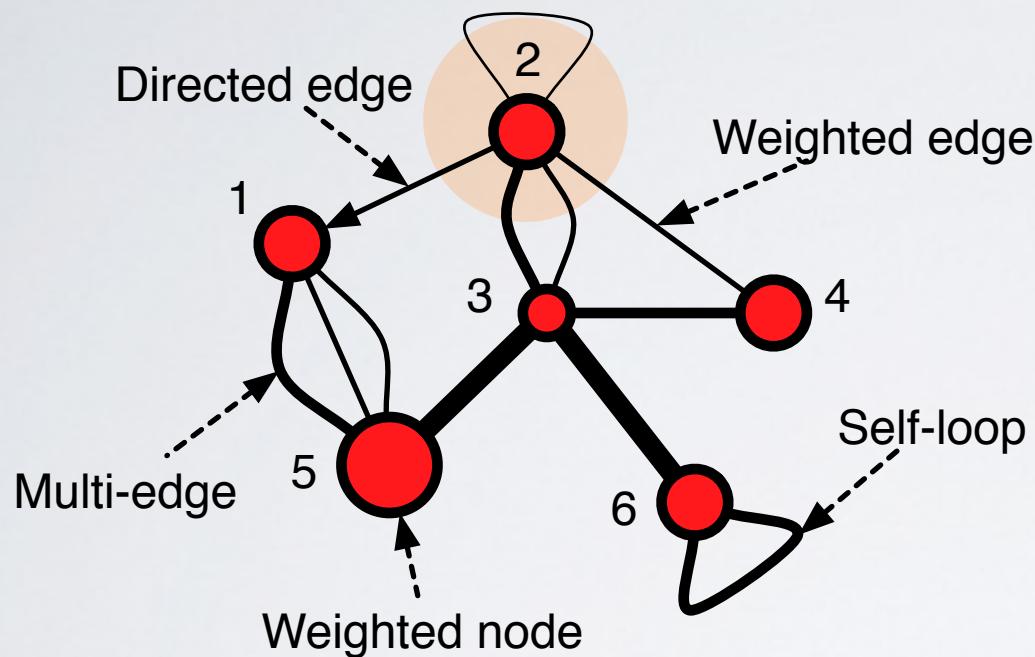


~~undirected~~

~~unweighted~~

~~no self loops~~

# a less simple network



adjacency matrix

$A$	1	2	3	4	5	6
1	0	0	0	0	{1, 1, 2}	0
2	1	$\frac{1}{2}$	{2, 1}	1	0	0
3	0	{2, 1}	0	2	4	4
4	0	1	2	0	0	0
5	{1, 1, 2}	0	4	0	0	0
6	0	0	4	0	0	2

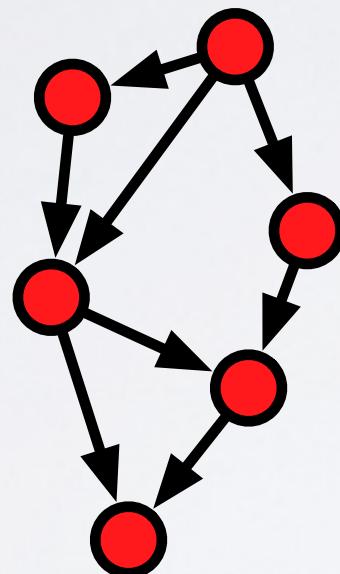
adjacency list

$A$
1 $\rightarrow \{(5, 1), (5, 1), (5, 2)\}$
2 $\rightarrow \{(1, 1), (2, \frac{1}{2}), (3, 2), (3, 1), (4, 1)\}$
3 $\rightarrow \{(2, 2), (2, 1), (4, 2), (5, 4), (6, 4)\}$
4 $\rightarrow \{(2, 1), (3, 2)\}$
5 $\rightarrow \{(1, 1), (1, 1), (1, 2), (3, 4)\}$
6 $\rightarrow \{(3, 4), (6, 2)\}$

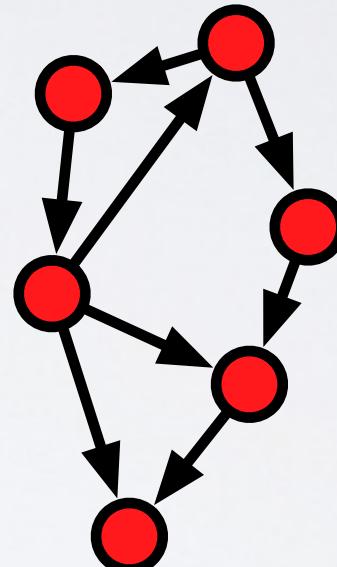
# directed networks

$$A_{ij} \neq A_{ji}$$

citation networks  
foodwebs\*  
epidemiological  
others?



directed acyclic graph

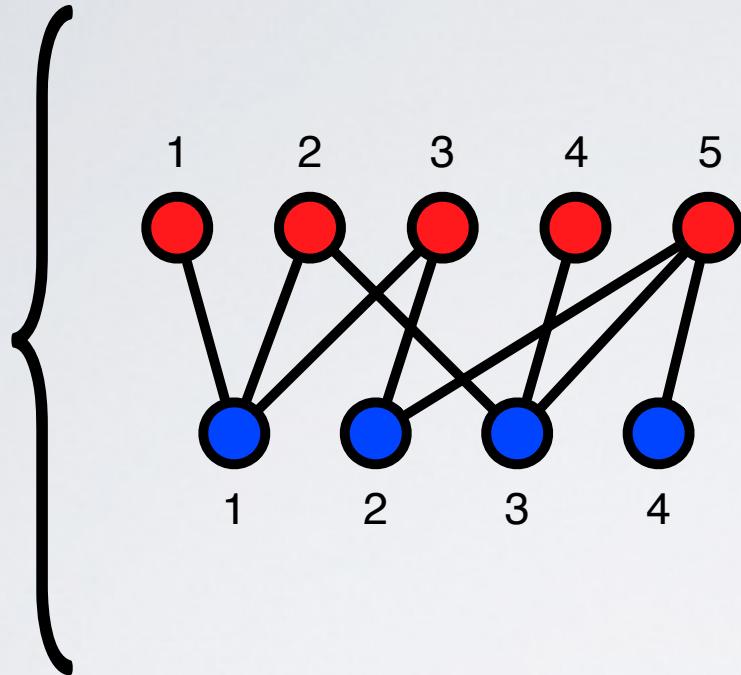


directed graph

WWW  
friendship?  
flows of goods,  
information  
economic exchange  
dominance  
neuronal  
transcription  
time travelers

# bipartite networks

bipartite  
network



**no within-type edges**

authors & papers

actors & movies/scenes

musicians & albums

people & online groups

people & corporate boards

people & locations (checkins)

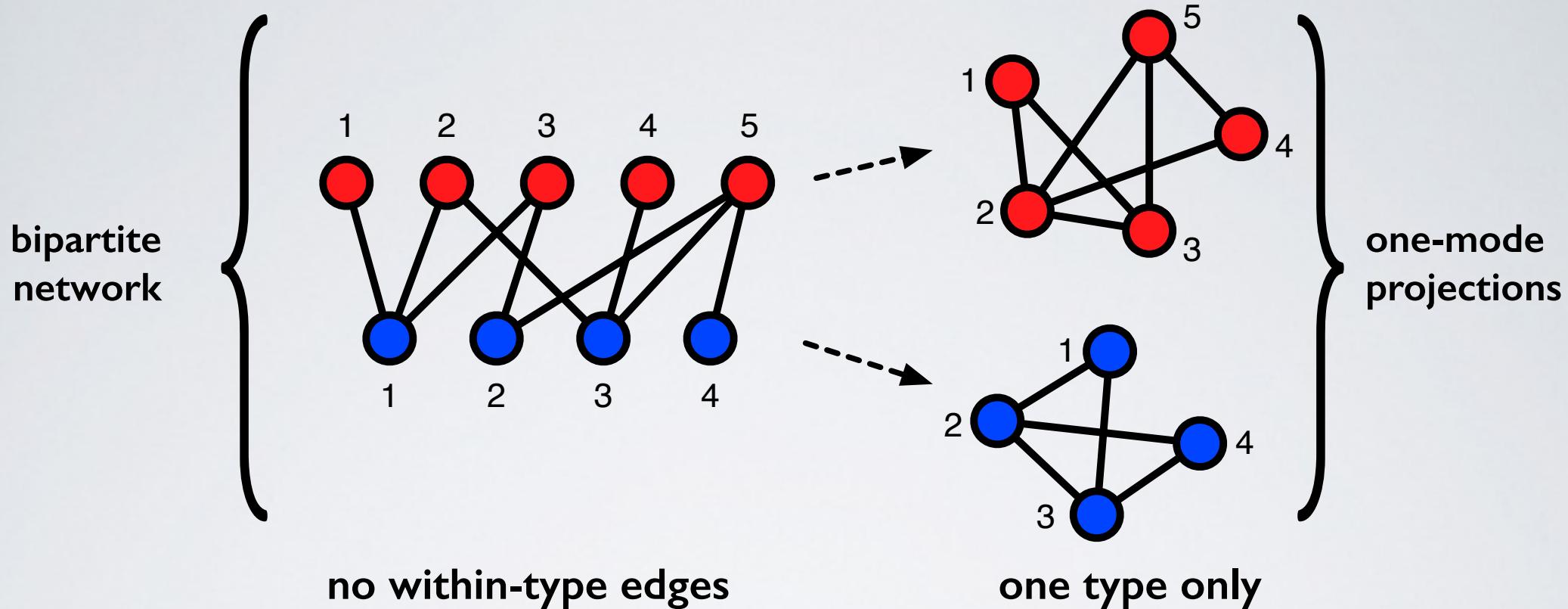
metabolites & reactions

genes & substrings

words & documents

plants & pollinators

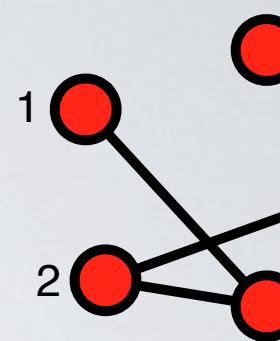
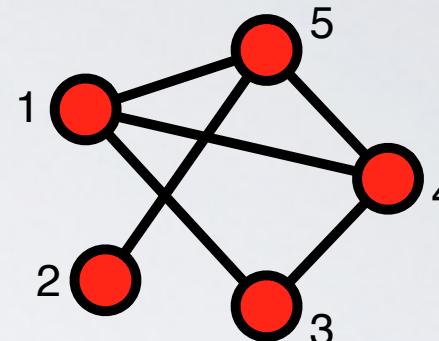
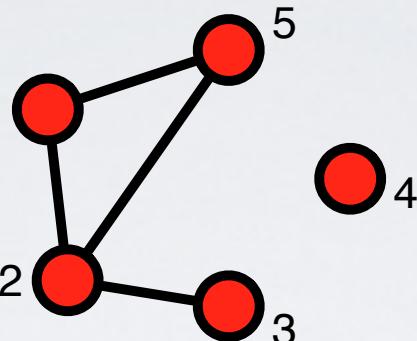
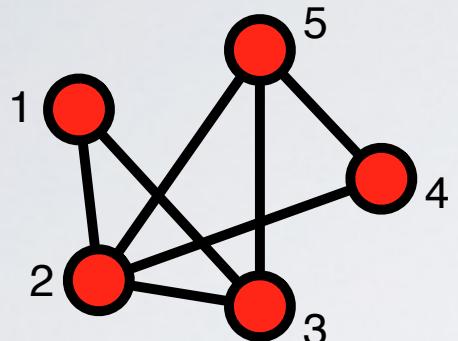
# bipartite networks



authors & papers  
actors & movies/scenes  
musicians & albums  
people & online groups  
people & corporate boards

people & locations (checkins)  
metabolites & reactions  
genes & substrings  
words & documents  
plants & pollinators

# temporal networks



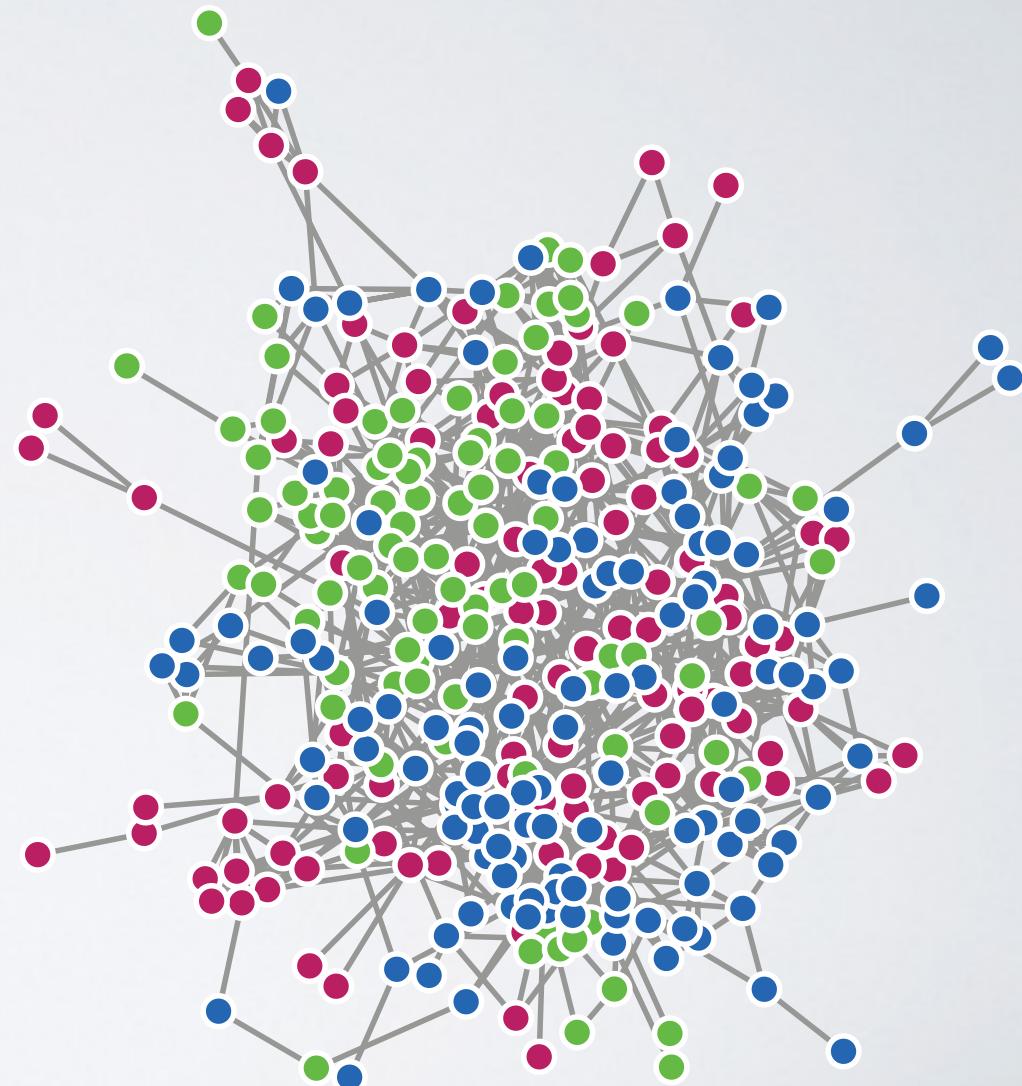
**any network over time**

discrete time (snapshots), edges  $(i, j, t)$

continuous time, edges  $(i, j, t_s, \Delta t)$

# describing networks

what networks look like



# describing networks

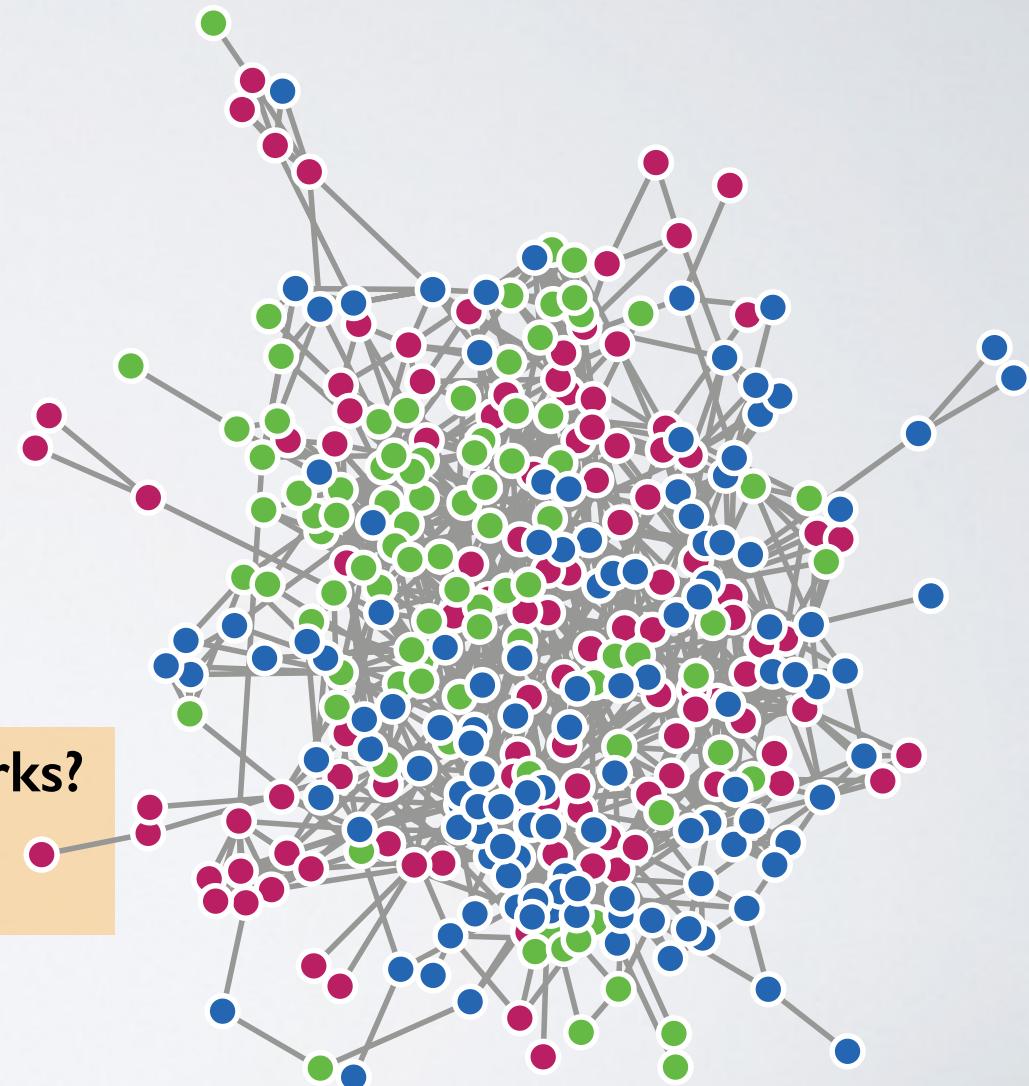
what networks look like

questions:

- **how are the edges organized?**
- **how do vertices differ?**
- **does network location matter?**
- **are there underlying patterns?**

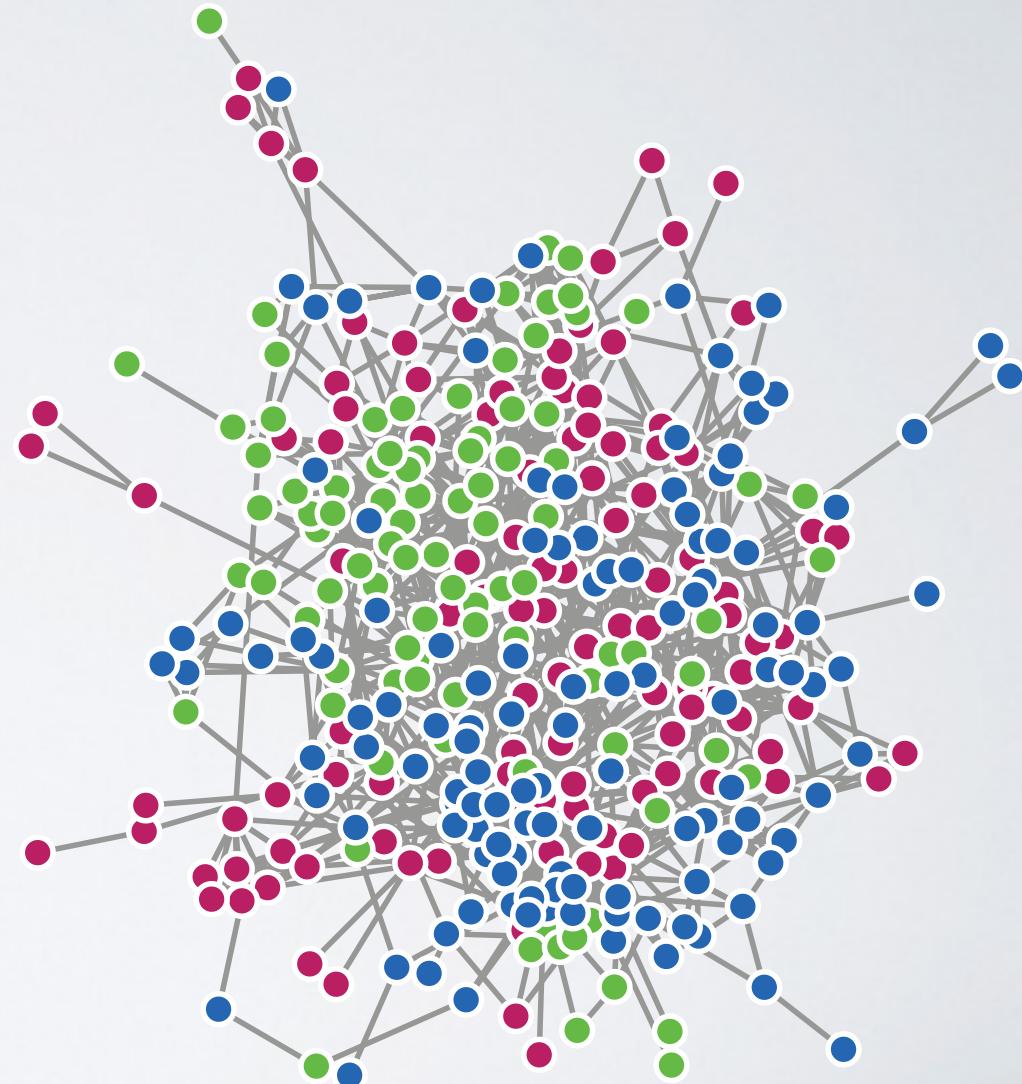
what we want to know

- **what processes shape these networks?**
- **how can we tell?**



# describing networks

a first step : **describe its features**

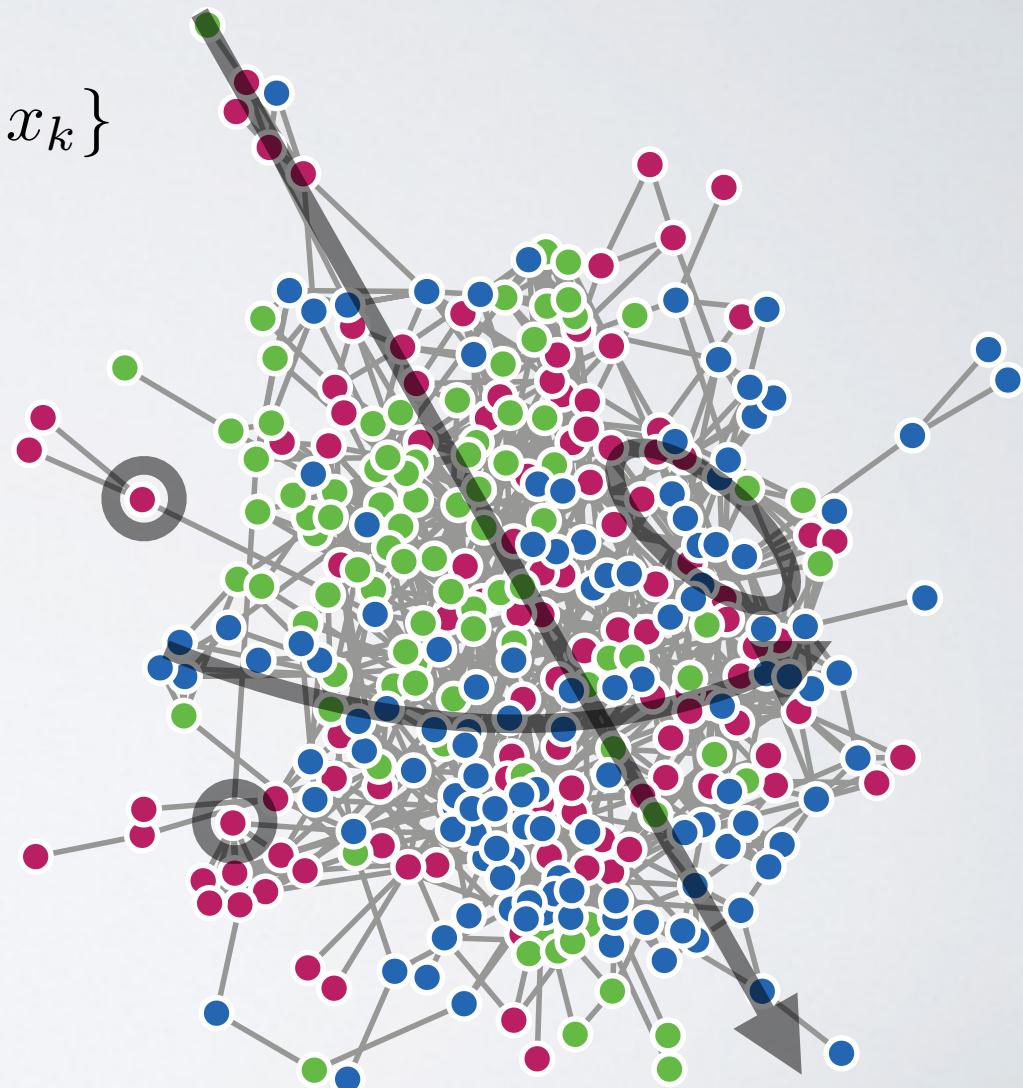


# describing networks

a first step : describe its features

$$f : G \rightarrow \{x_1, \dots, x_k\}$$

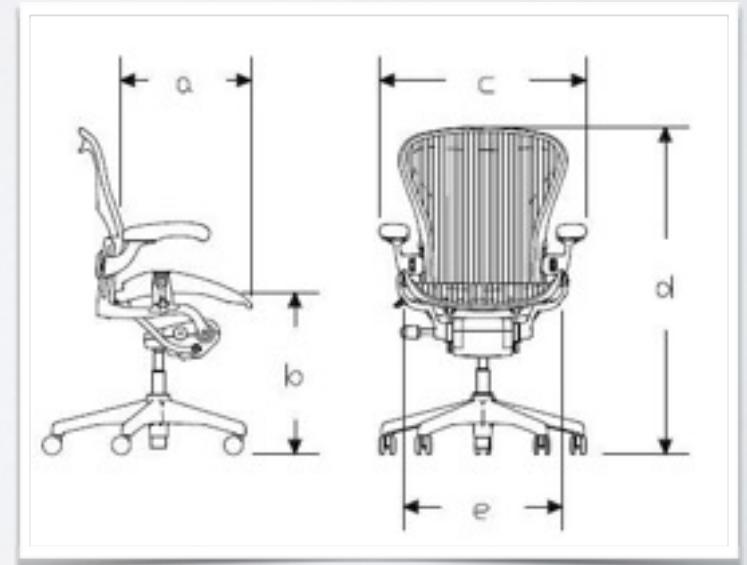
- **degree distributions**
- **short-loop density (triangles, etc.)**
- **shortest paths (diameter, etc.)**
- **vertex positions**
- **correlations between these**



# describing networks

a first step : **describe its features**

$$f : \text{object} \rightarrow \{x_1, \dots, x_k\}$$



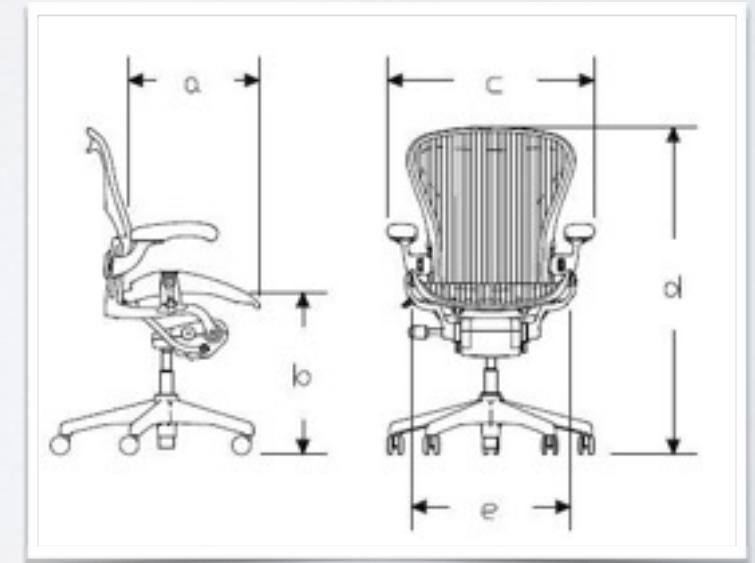
# describing networks

a first step : **describe its features**

$$f : \text{object} \rightarrow \{x_1, \dots, x_k\}$$

- **physical dimensions**
- **material density, composition**
- **radius of gyration**
- **correlations between these**

helpful for exploration, but not what we want...



# describing networks

what we want : understand its structure

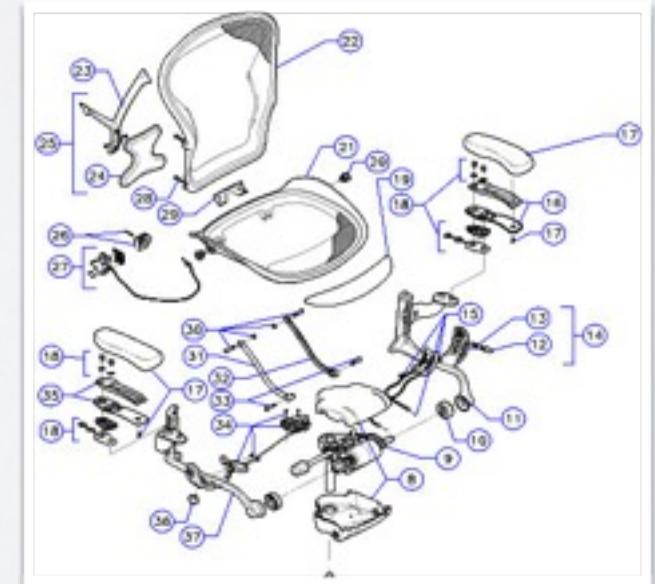
$$f : \text{object} \rightarrow \{\theta_1, \dots, \theta_k\}$$

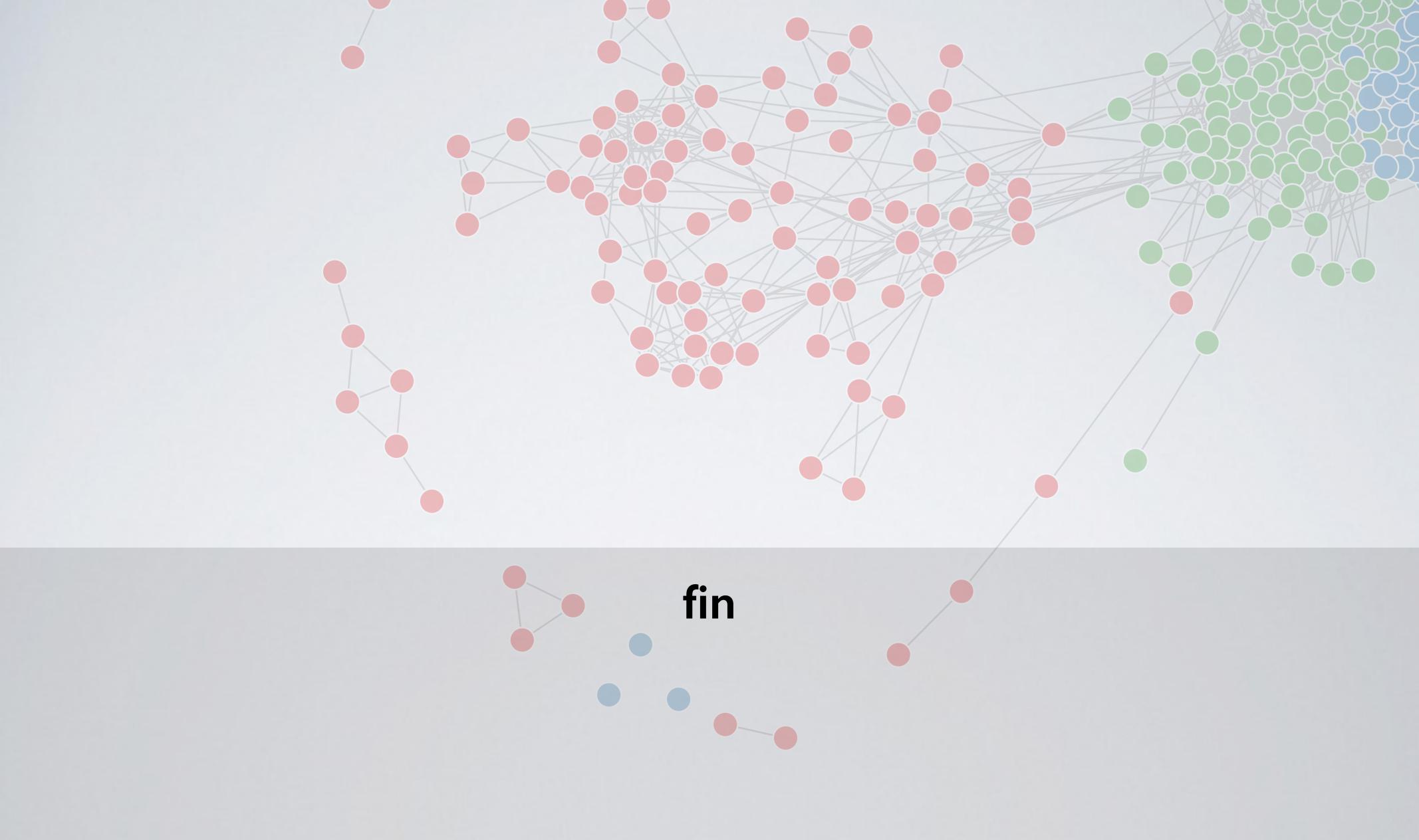
- what are the fundamental parts?
- how are these parts organized?
- where are the degrees of freedom  $\vec{\theta}$ ?
- how can we define an abstract class?
- structure — dynamics — function?

what does **local-level structure** look like?

what does **large-scale structure** look like?

how does **structure constrain** function?





**fin**