Analisi delle Reti Sociali

http://didawiki.cli.di.unipi.it/doku.php/dm/sna.ingegneria2011

Introduzione & Motivazioni

Fosca Giannotti & Michele Berlingerio, KDD Lab. ISTI-CNR

kdd.isti.cnr.it/

fosca.giannotti@isti.cnr.it, michele.berlingerio@isti.cnr.it

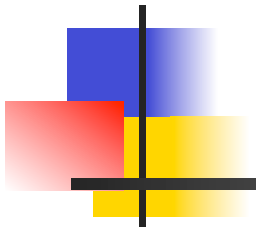
dal corso di Dino Pedreschi

Web Mining ed Analisi delle Reti Sociali

http://didawiki.cli.di.unipi.it/doku.php/wma/start

Dipartimento di Informatica, Università di Pisa

Organizzazione del corso



Sommario

- Social Network Introduction
- Graph theory and social networks: basic measures
- The social behaviours of networks: small-world, strong&weak ties:, homophilia
- Community detection
- Network Dynamics: information diffusion, generative models, link prediction

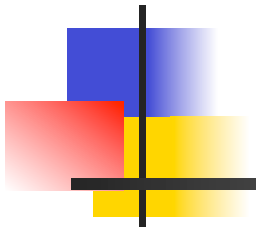
Materiale didattico

- Slides free composition by: Jiawei Han, Univ. of Illinois at Urbana-Champaign, Yure Leskovec: Stanford Univ., Laszlo Barabasi, Northeastern Univ., Cesar Hidalgo MIT
- M. E. J. Newman, ***The structure and function of complex networks***, SIAM Review, Vol. 45, p. 167-256, 2003.
- David Easley, Jon Kleinberg: Networks, Crowds, and Markets. http://www.cs.cornell.edu/home/kleinber/
- Chapter 9.2 of the book: Jiawei Han and Micheline Kamber. ***Data Mining: Concepts and Techniques***, 2nd ed. Morgan Kaufmann Publishers, 2006.
- Consultazione:
  - Duncan J. Watts. ***Six Degrees: The Science of a Connected Age***. (Norton, New York, 2003).
  - A.-L. Barabasi. **Linked**. PLUME, Penguin Group, 2002.

Valutazione

- Seminario + relazioncina (su 2-3 paperi a scelta)
  - 1 page: breve sommario: (quali contenuti tecnici
  - 1 page: critica: quale relazione con il corso e cosa gli autori non hanno considerato o cosa di non realistico
  - 1 page: Brainstorming ♣ What are promising further research questions in the direction of thepapers?♣ How could they be pursued?♣ An idea of a **better model for something? A better algorithm? A test of a model or algorithm on a dataset or simulated data?**
- Progettino
  - Costruire (crawling) ed analizzare una rete

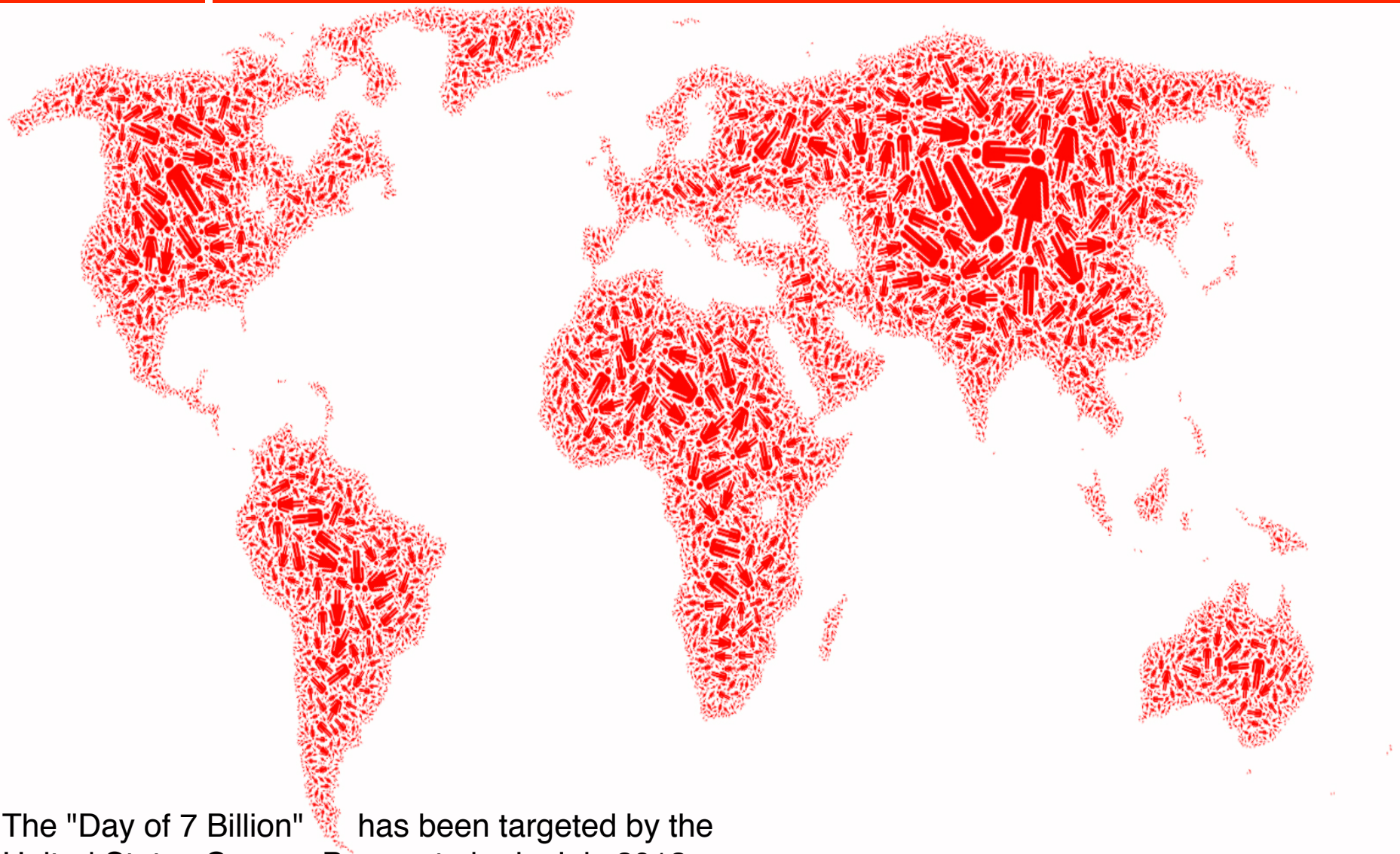
Perché ci interessiamo di Reti Sociali



COMPLESSITA' e Grandi Numeri

SOCIETY

Factoid:



The "Day of 7 Billion" has been targeted by the United States Census Bureau to be in July 2012.
http://en.wikipedia.org/wiki/World_population

COMPLEX SYSTEMS

Complex

[adj., v. kuh m-pleks, kom-pleks; n. kom-pleks]

–adjective

1.

composed of many interconnected parts; compound; composite: a complex highway system.

2.

characterized by a very complicated or involved arrangement of parts, units, etc.: complex machinery.

3.

so complicated or intricate as to be hard to understand or deal with: a complex problem.

Source: Dictionary.com

Complexity

Complexity, a **scientific theory** which asserts that some systems display behavioral phenomena that are completely inexplicable by any conventional analysis of the systems' constituent parts. These phenomena, commonly referred to as emergent behaviour, seem to occur in many complex systems involving living organisms, such as a stock market or the human brain.

Source: John L. Casti, Encyclopædia Britannica

THE ROLE OF NETWORKS

Behind each complex system there is a **network**, that defines the interactions between the component.

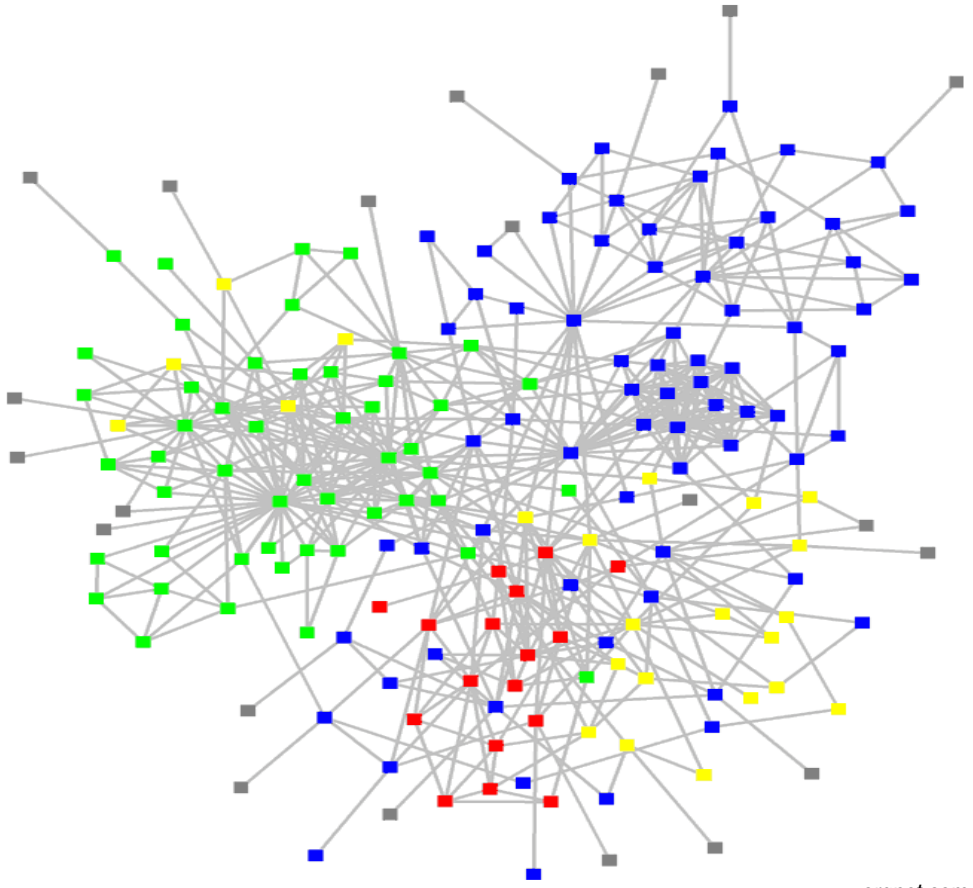
SOCIETY

Factoid:

The “Social Graph” behind Facebook

Keith Shepherd's "Sunday Best". http://baseballart.com/2010/07/shades-of-greatness-a-story-that-needed-to-be-told/

STRUCTURE OF AN ORGANIZATION



www.orgnet.com

: departments

: consultants

: external experts

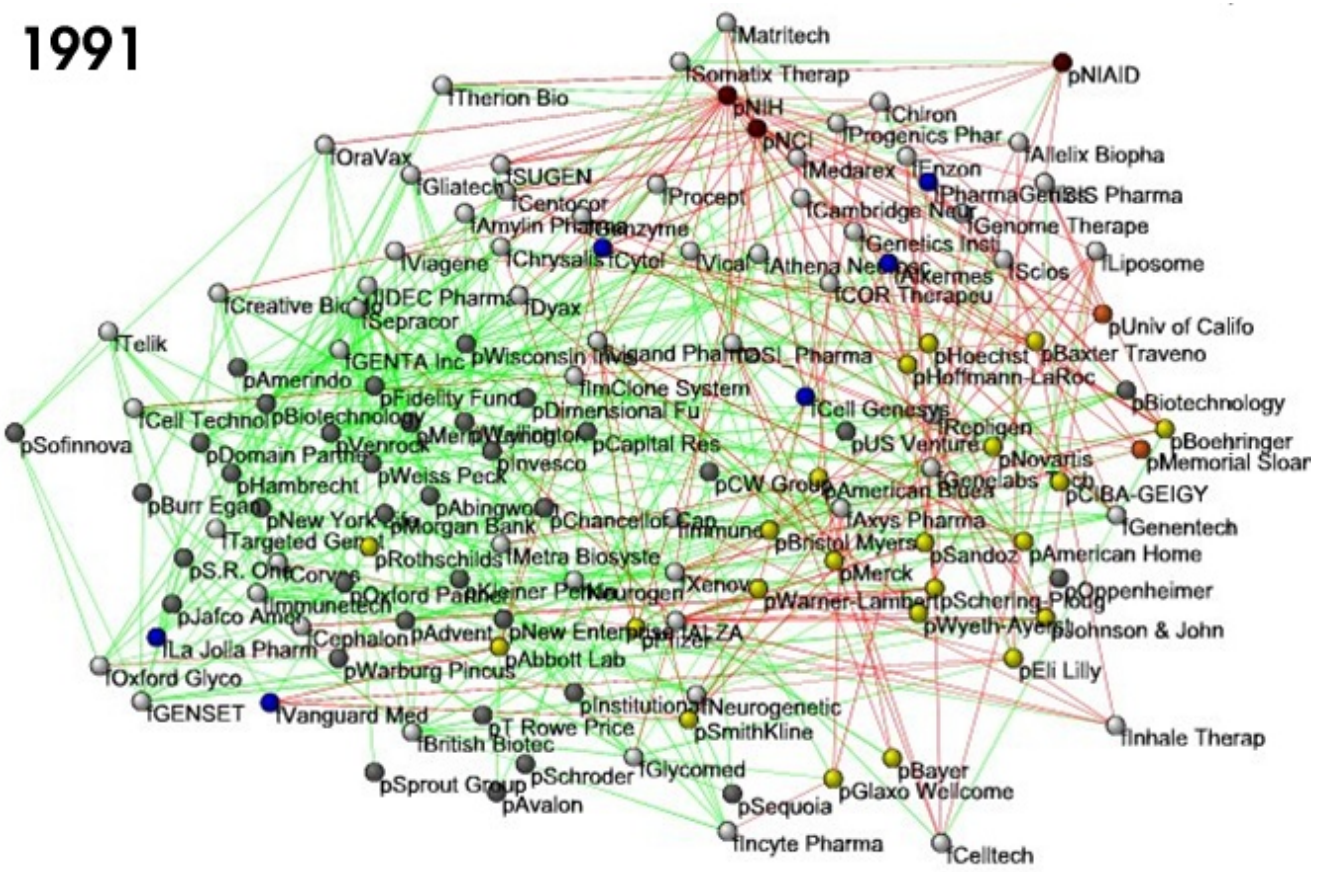
BUSINESS TIES IN US BIOTECH-INDUSTRY

Nodes:

- Companies
- Investment
- Pharma
- Research Labs
- Public
- Biotechnology

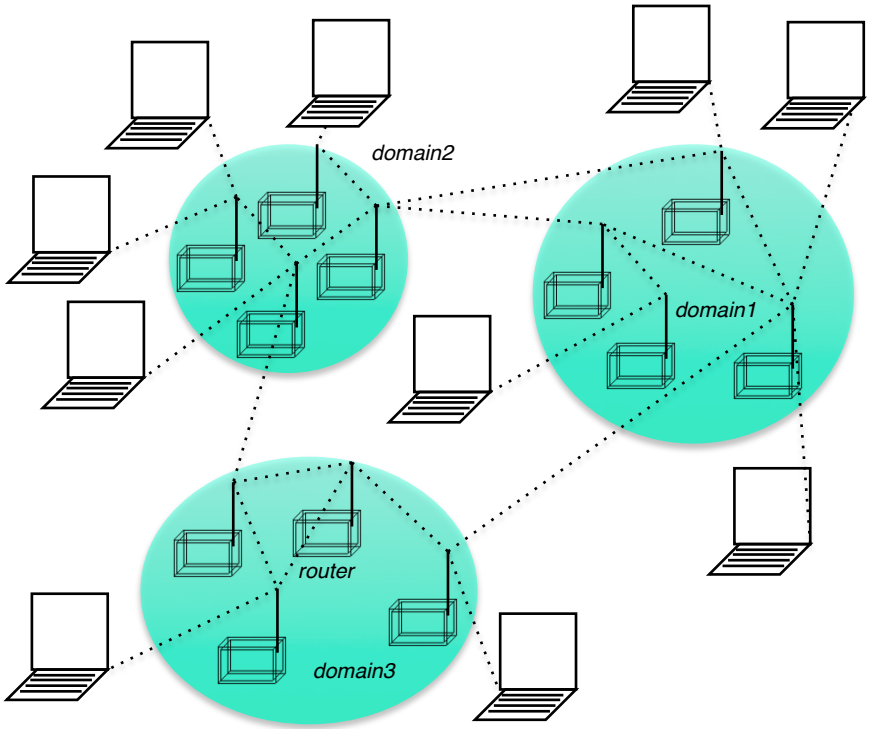
Links:

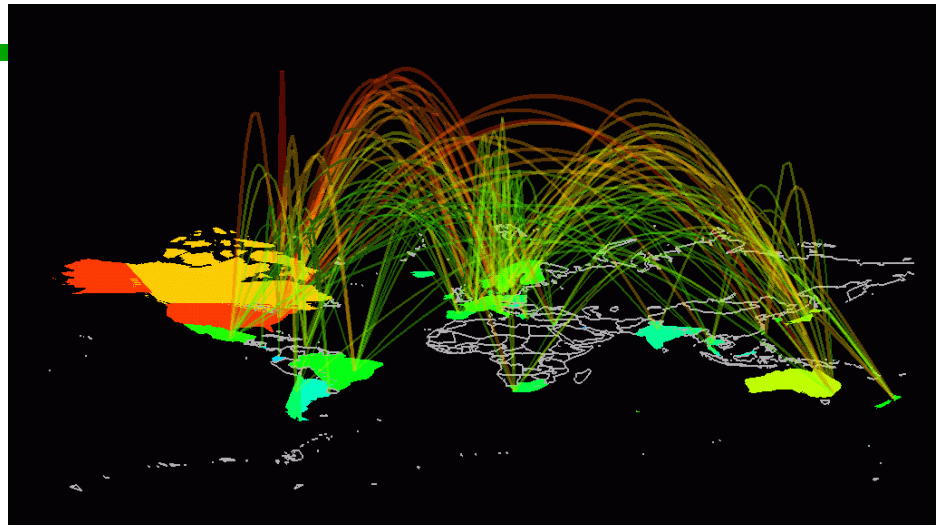
- Collaborations
- Financial
- R&D

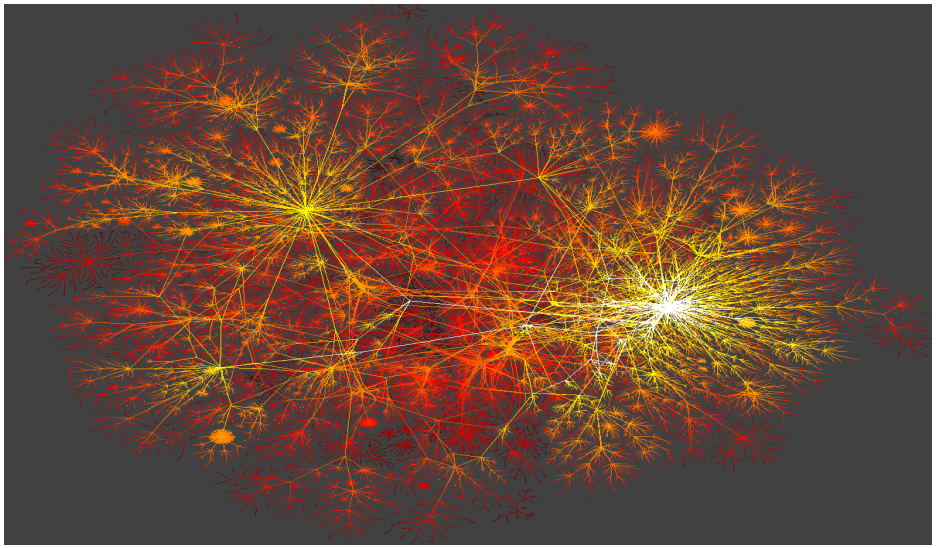


http://ecclectic.ss.uci.edu/~drwhite/Movie

INTERNET



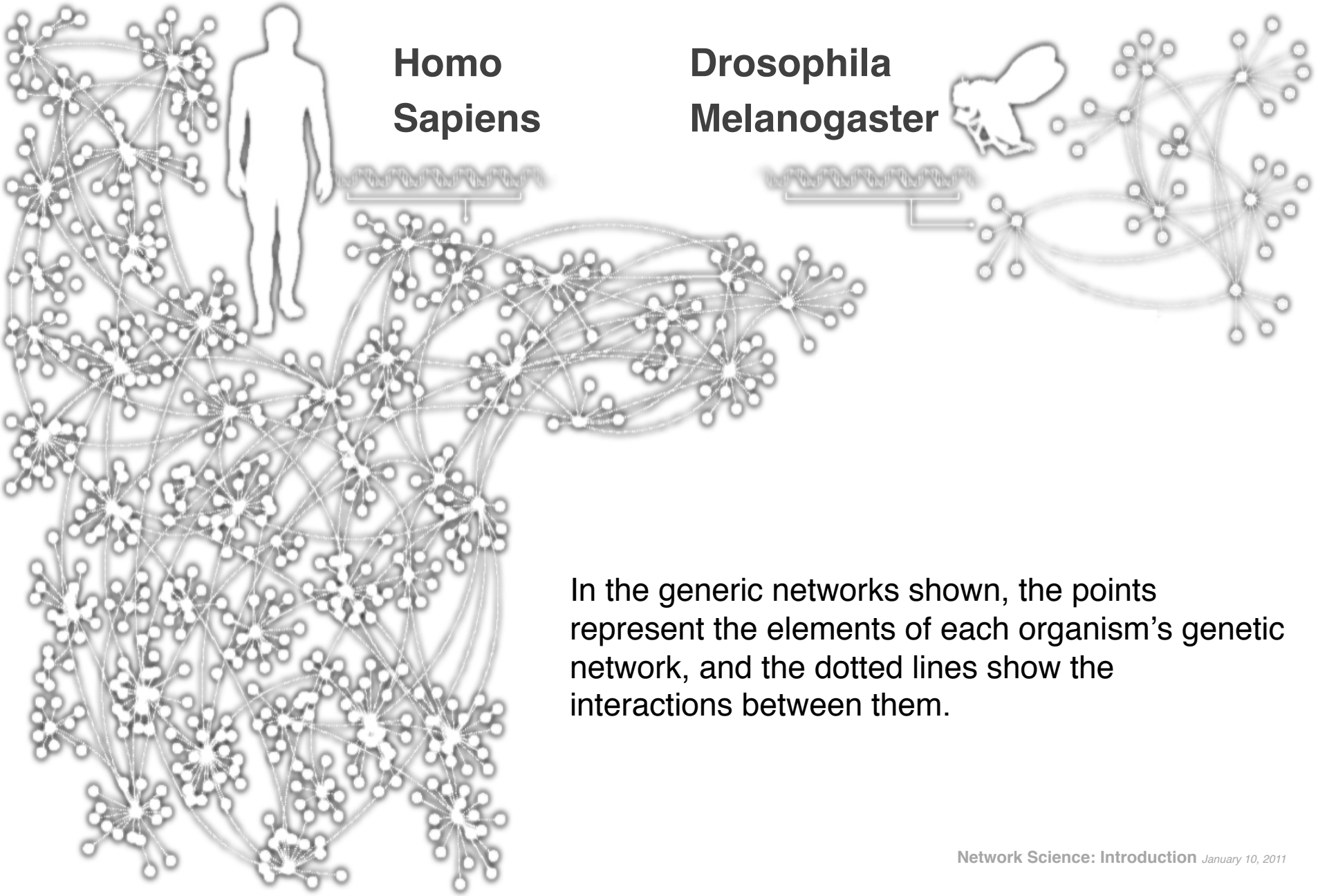




HUMANS GENES

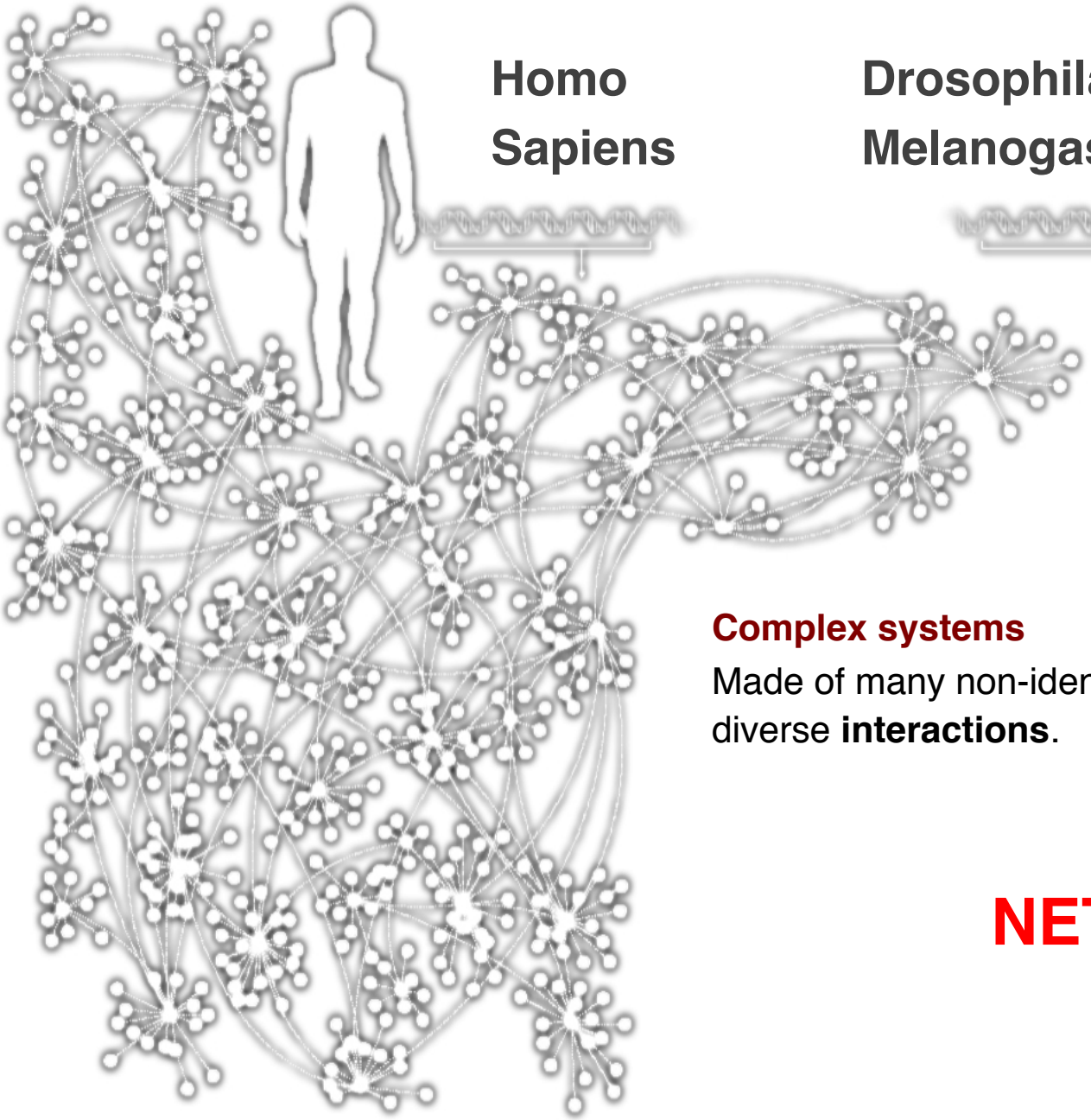
Humans have only about three times as many genes as the fly, so human complexity seems unlikely to come from a sheer quantity of genes. Rather, some scientists suggest, each human has a network with different parts like genes, proteins and groups.

HUMANS GENES



In the generic networks shown, the points represent the elements of each organism's genetic network, and the dotted lines show the interactions between them.

HUMANS GENES



Homo Sapiens

Drosophila Melanogaster

Complex systems

Made of many non-identical **elements** connected by diverse **interactions**.

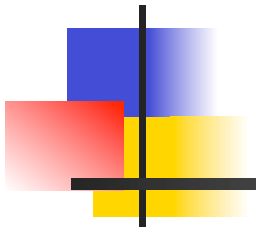
NETWORK

THE ROLE OF NETWORKS

Behind each system studied in complexity there is an intricate wiring diagram, or a **network**, that defines the interactions between the component.

We will never understand complex system unless we map out and understand the networks behind them.

Quali tipi di reti



Society

Nodes: individuals

Links: social relationship (family/work/friendship/etc.)

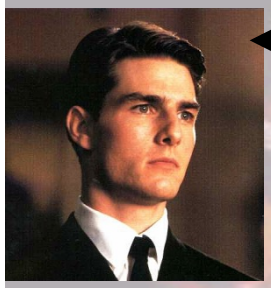
S. Milgram (1967)

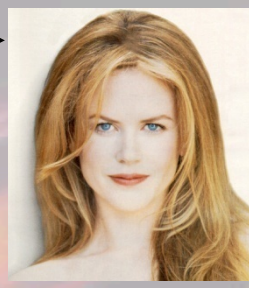
John Guare

Six Degrees of Separation

Social networks: Many individuals with diverse social interactions between them.

Social networks: Actor Connectivity





Days of Thunder (1990)
Far and Away (1992)
Eyes Wide Shut (1999)

N = 212,250 actors

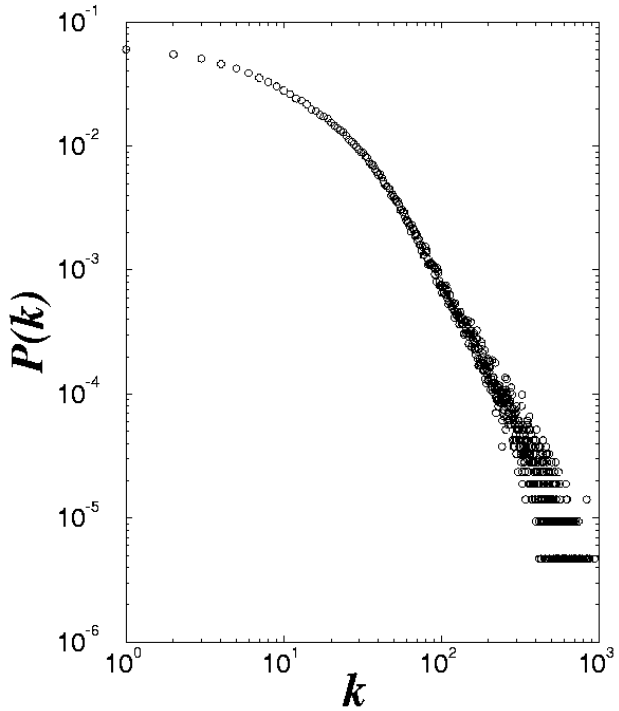
⟨k⟩ = 28.78

$P(k) \sim k^{-\gamma}$

$\gamma = 2.3$

Nodes: actors

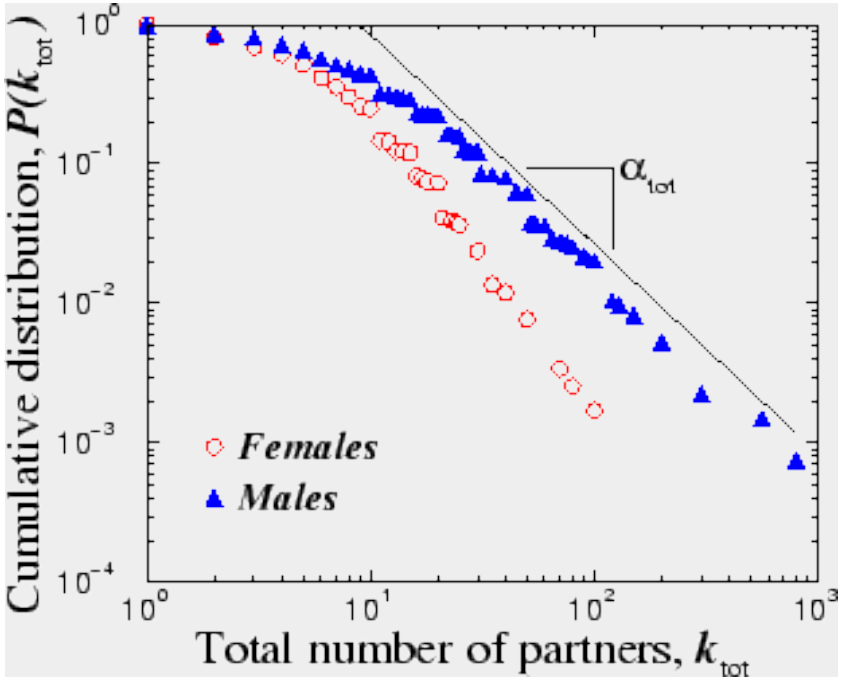
Links: cast jointly



Social networks: Sex-Web

Nodes: people (Females; Males)

Links: sexual relationships

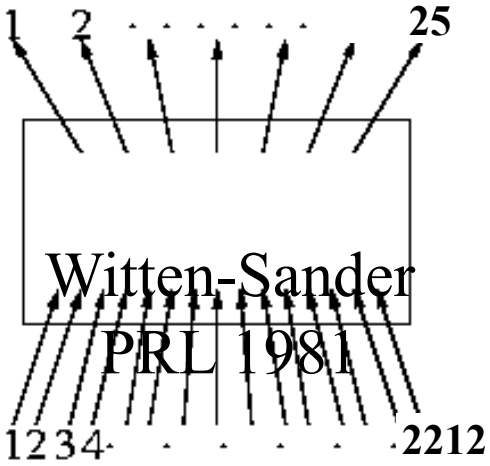


4781 Swedes; 18-74;
59% response rate.
Liljeros et al. Nature 2001

Information networks: Science Citation Index

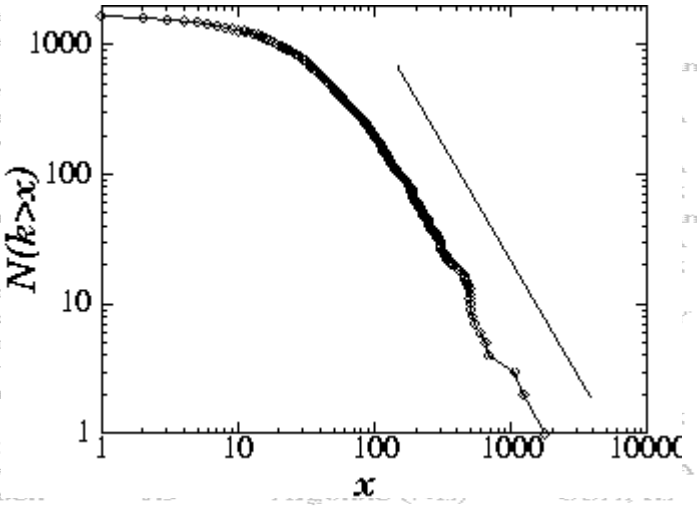
Nodes: papers

Links: citations



Witten-Sander PRL 1981

1736 PRL papers (1988)



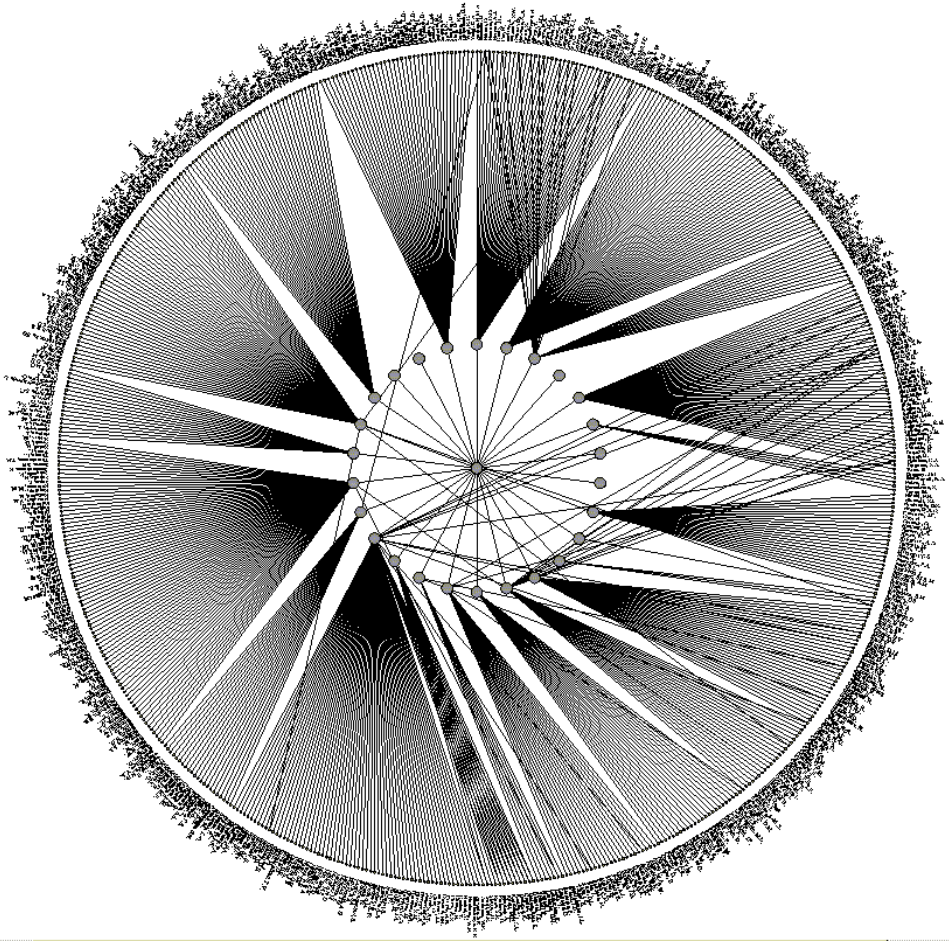
$P(k) \sim k^{-\gamma}$

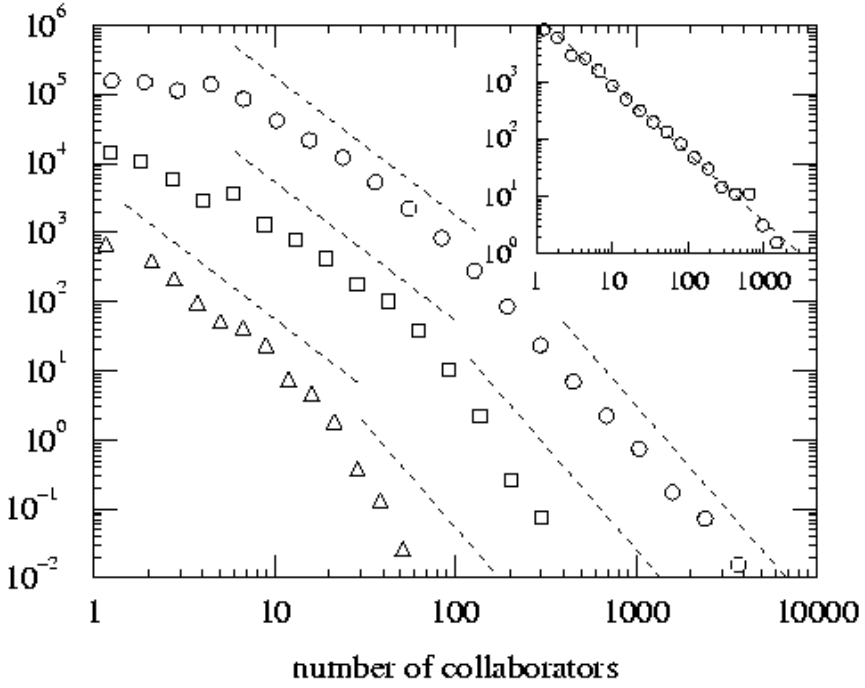
$(\gamma = 3)$

(S. Redner, 1998)

Information network: Science Coauthorship

Nodes: scientist (authors)
Links: write paper together



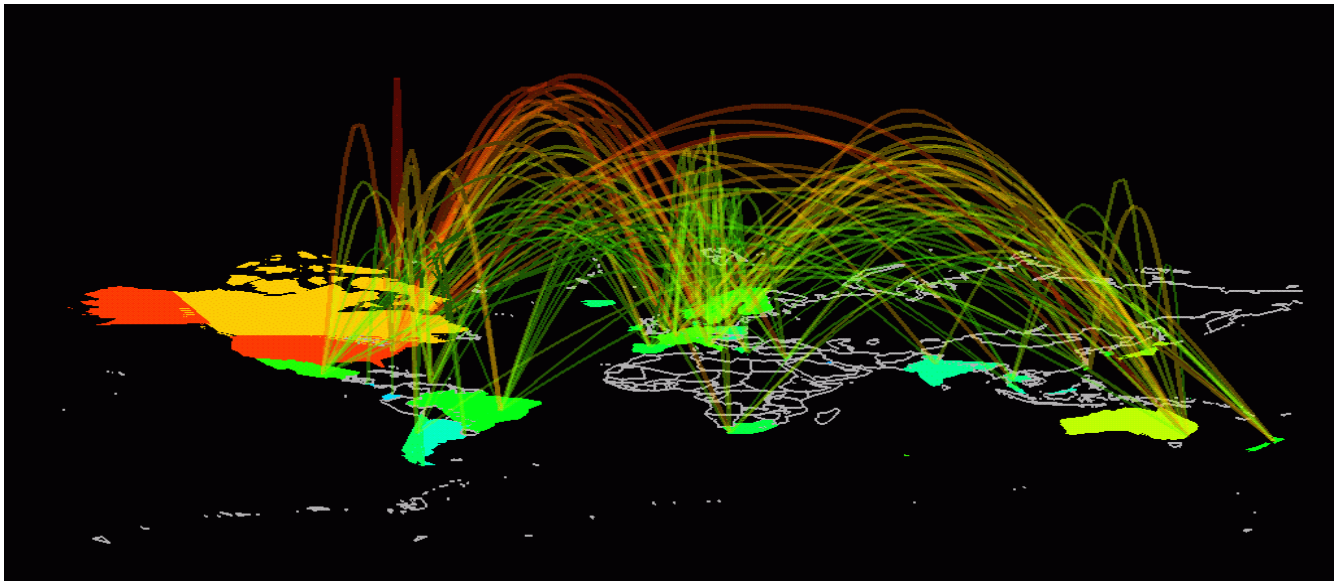


(Newman, 2000, H. Jeong et al 2001)

Communication networks

The Earth is developing an electronic nervous system, a network with diverse **nodes** and **links** are

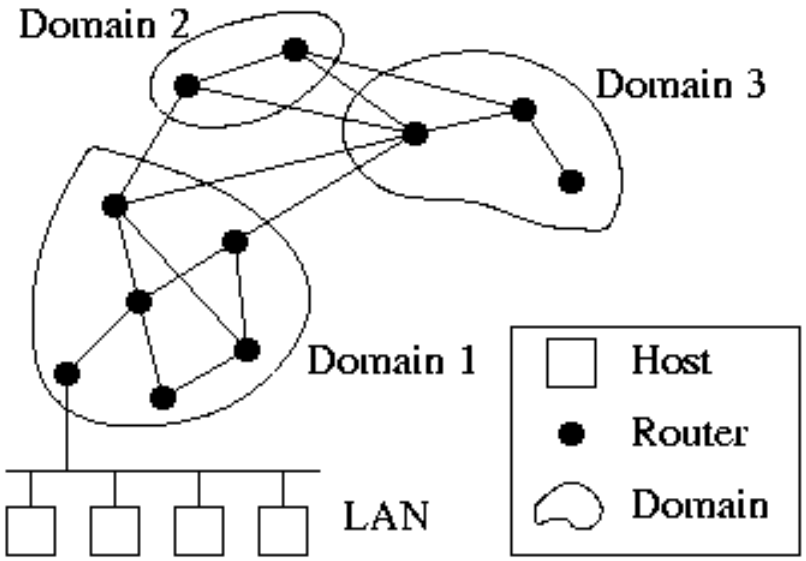
–computers	–phone lines
-routers	-TV cables
-satellites	-EM waves

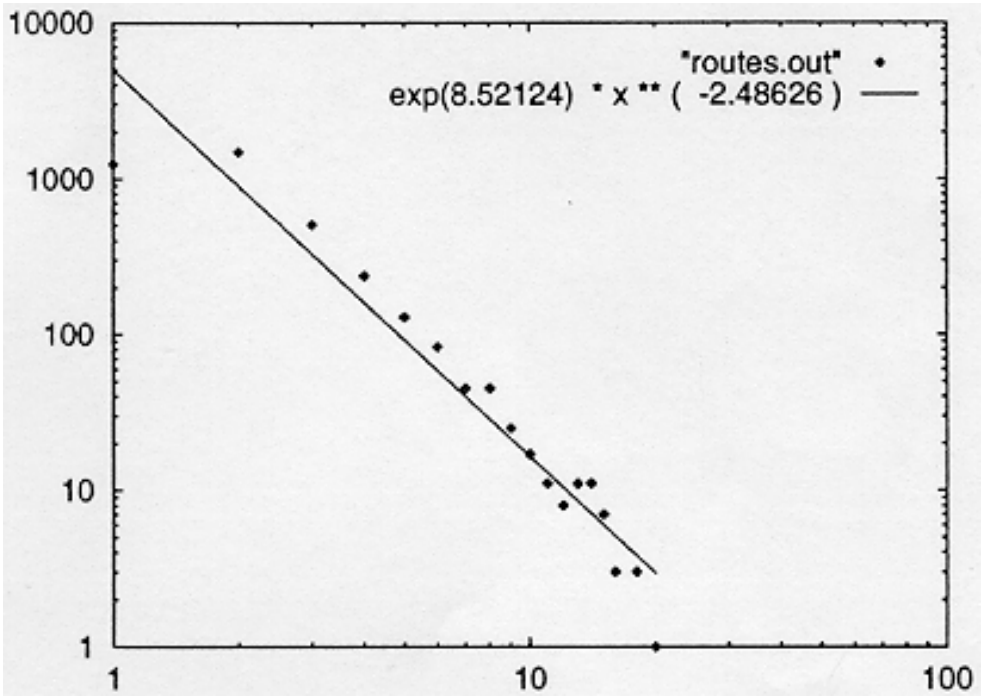


Communication networks: Many non-identical components with diverse connections between them.

Tech. networks: Internet Backbone

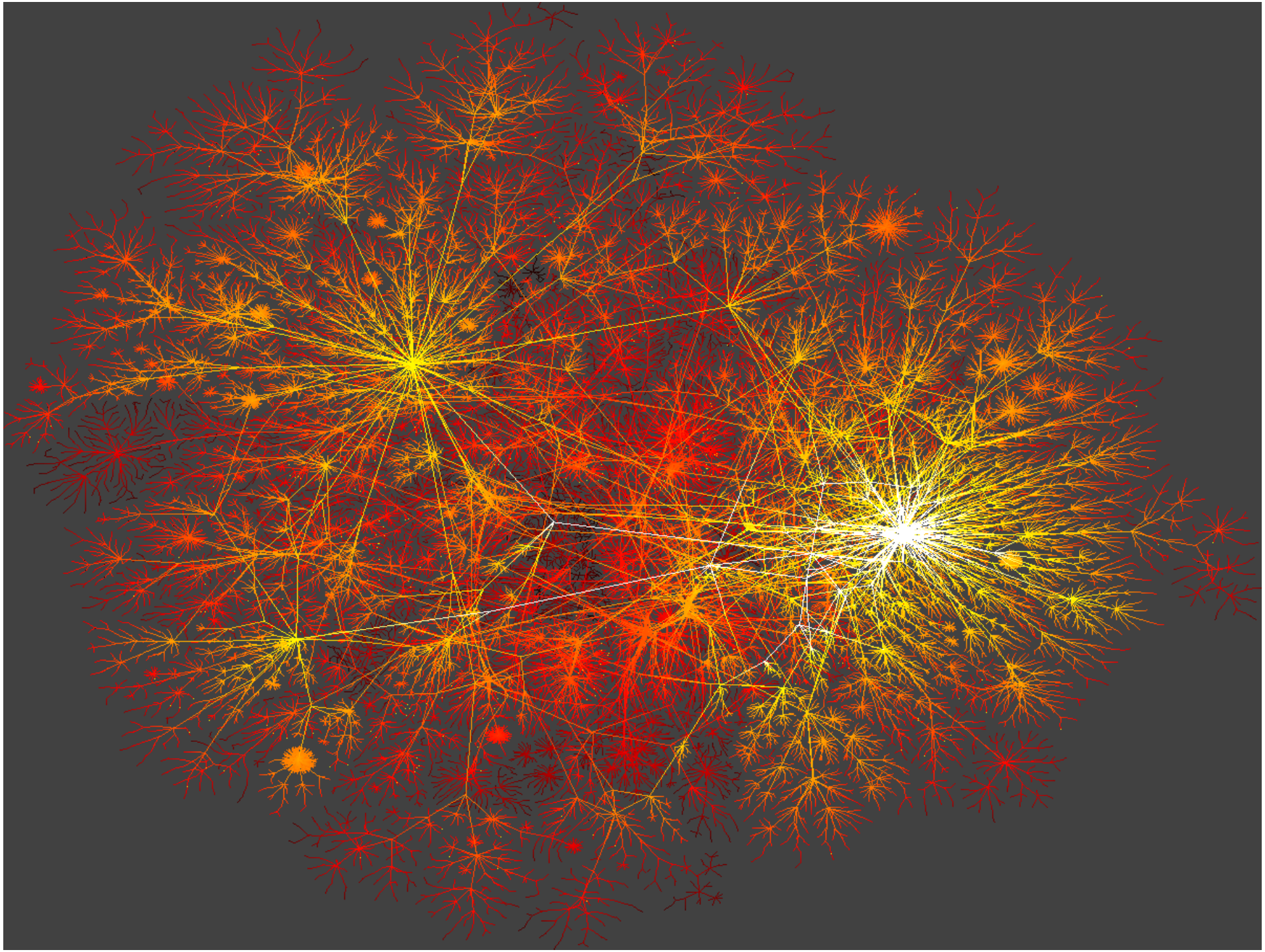
Nodes: computers, routers
Links: physical lines



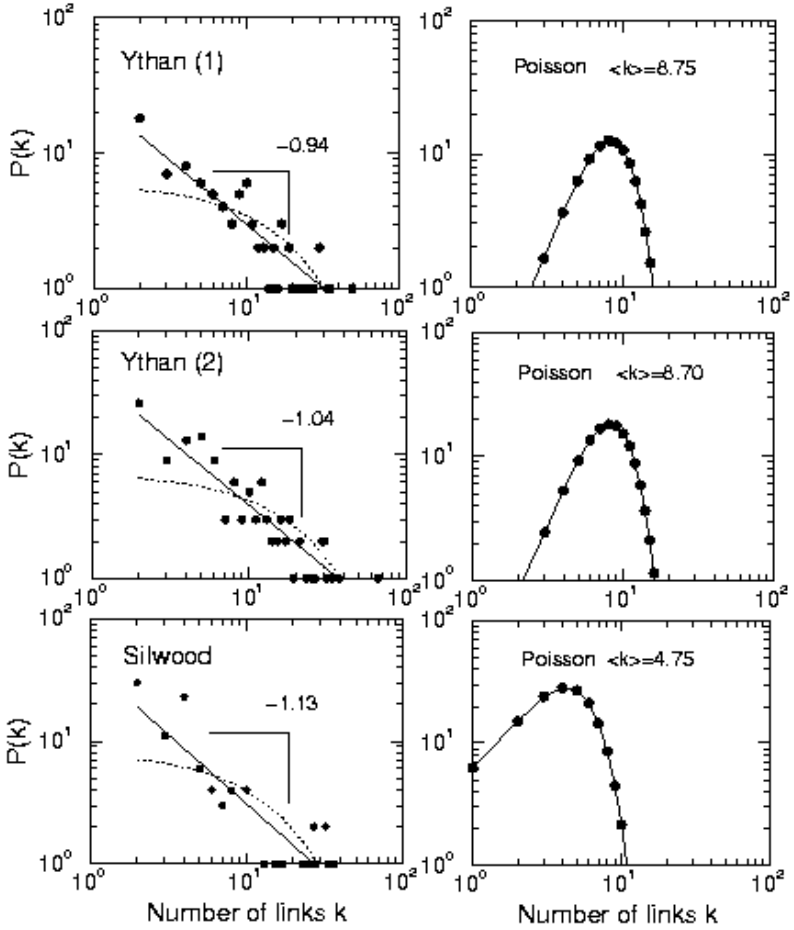


(Faloutsos, Faloutsos and Faloutsos, 1999)

Internet-Map

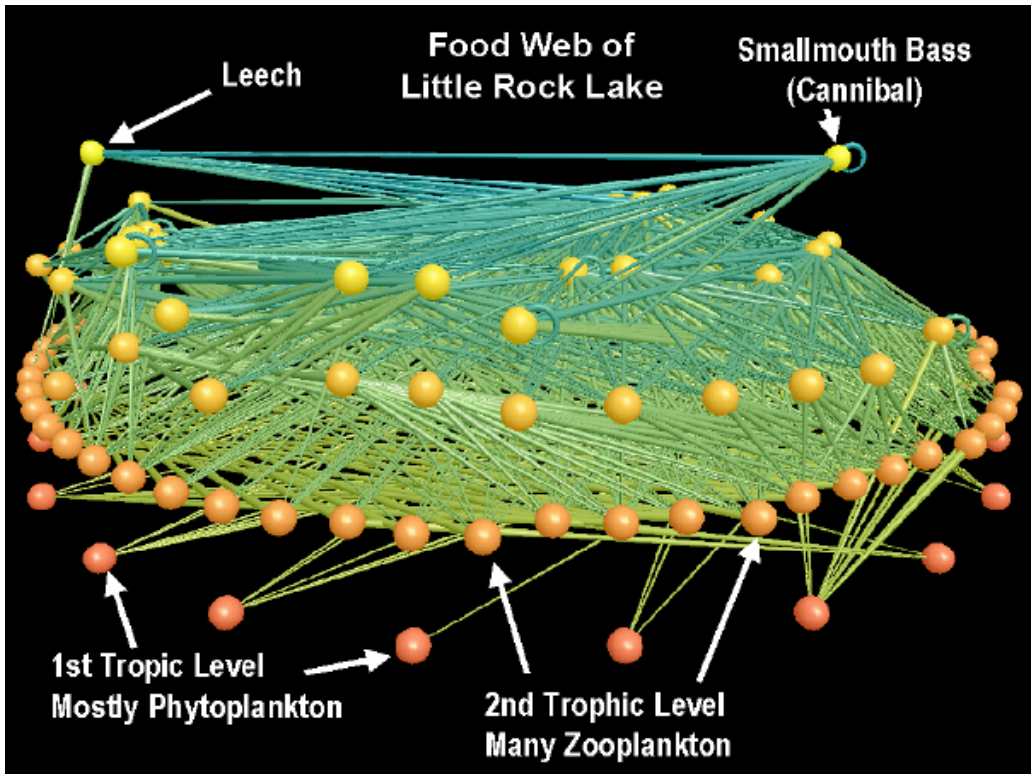


Biological networks: Food Web



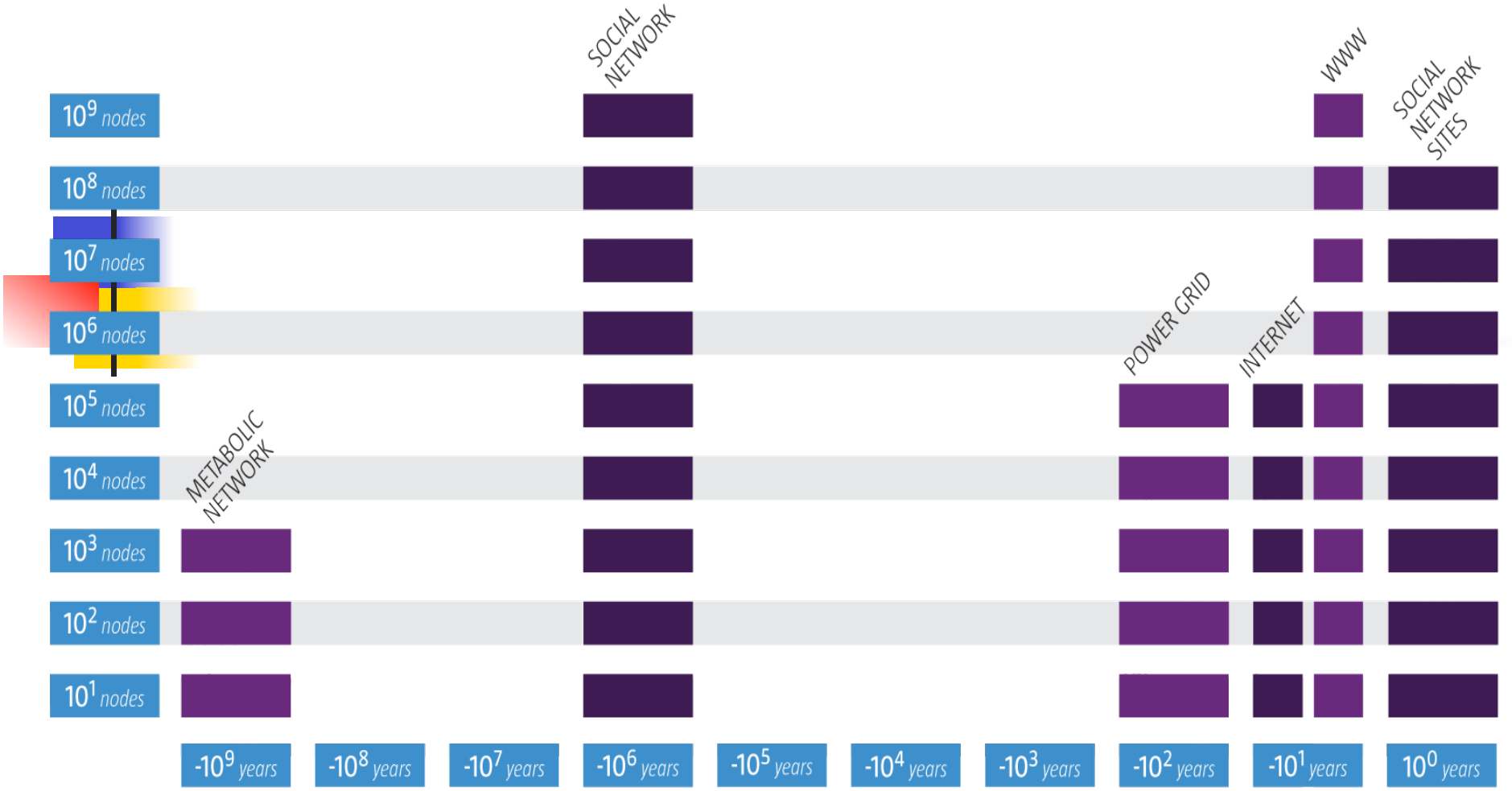
R. Sole (cond-mat/0011195)

Nodes: trophic species
Links: trophic interactions



R.J. Williams, N.D. Martinez *Nature* (2000)

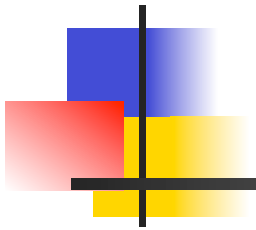
THE LIFE OF NETWORKS



THE EMERGENCE OF NETWORK SCIENCE

Data Availability:

Movie Actor Network, 1998;
World Wide Web, 1999.
C elegans neural wiring diagram 1990
Citation Network, 1998
Metabolic Network, 2000;
PPI network, 2001



Universality:

The architecture of networks emerging in various domains of science, nature, and technology are more similar to each other than one would have expected.

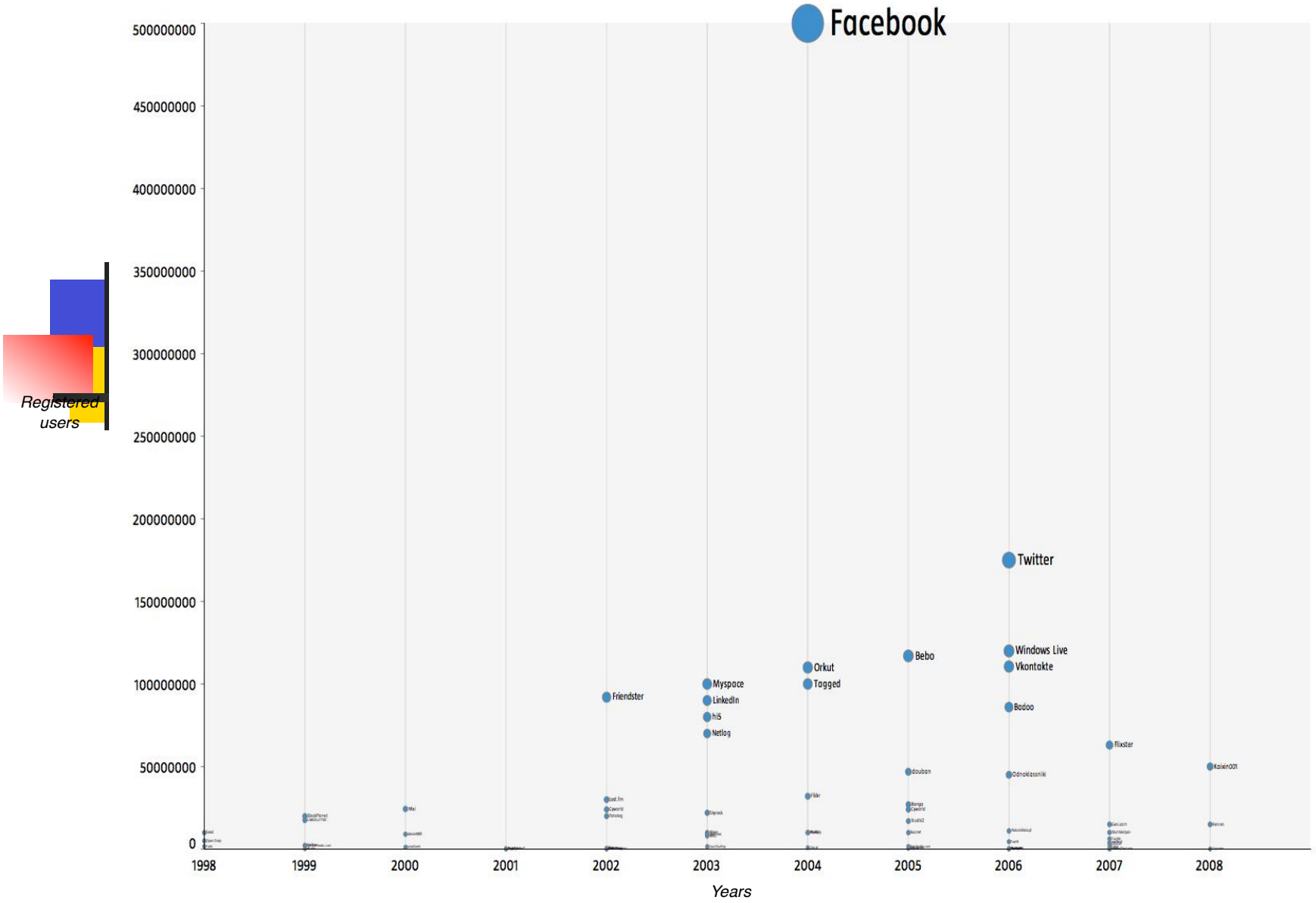
The (urgent) need to understand complexity:

Despite the challenges complex systems offer us, we cannot afford to not address their behavior, a view increasingly shared both by scientists and policy makers. Networks are not only essential for this journey, but during the past decade some of the most important advances towards understanding complexity were provided in context of network theory.

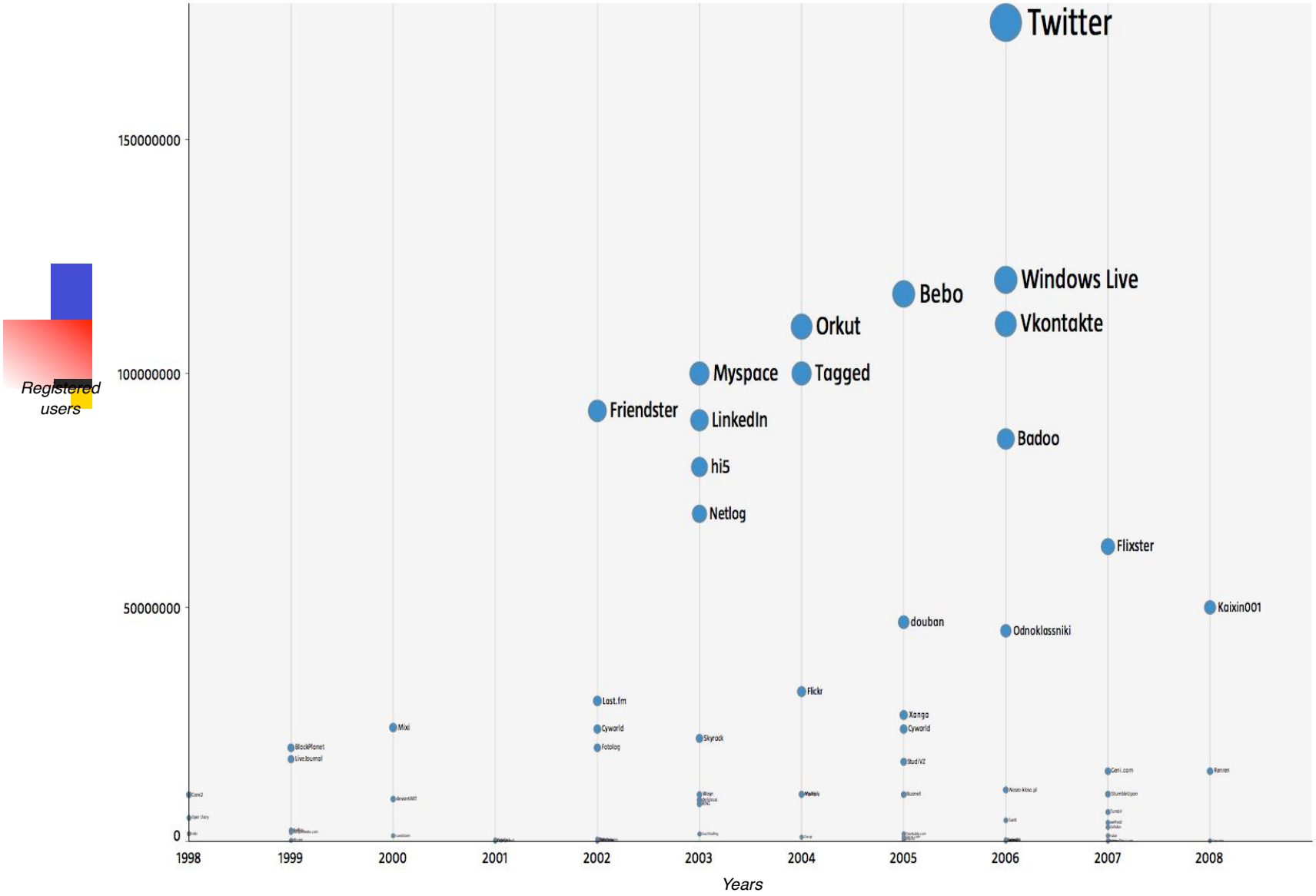
Network data is increasingly available

- ♣ On-line communities: Facebook (500 million users)
- ♣ Communication: Instant Messenger (~1 billion users)
- ♣ News and Social media: Blogging (250 million users)

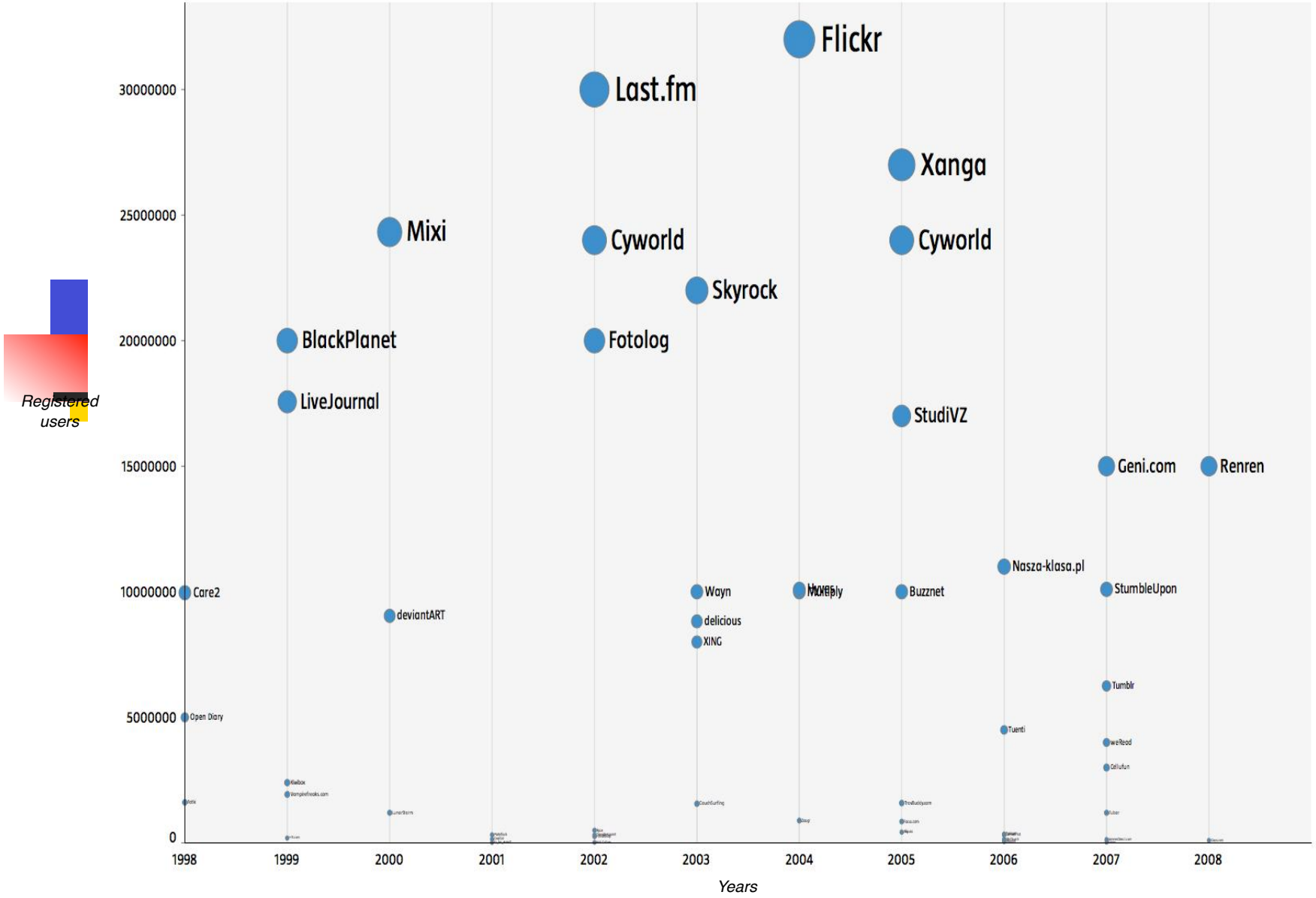
THE LIFE OF NETWORKS



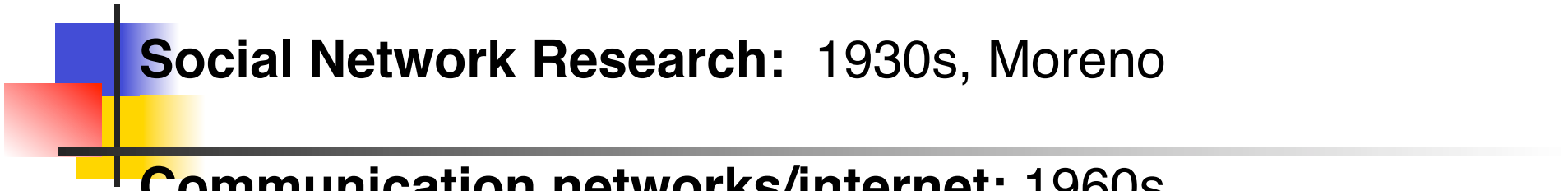
THE LIFE OF NETWORKS



THE LIFE OF NETWORKS



THE HISTORY OF NETWORK ANALYSIS



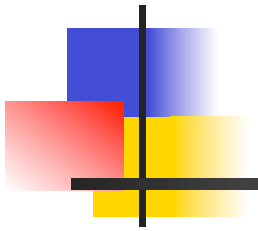
Graph theory: 1735, Euler

Social Network Research: 1930s, Moreno

Communication networks/internet: 1960s

Ecological Networks: May, 1979.

THE TOOLS OF MODERN NETWORK THEORY



- Graph theory
- Social network theory
- Statistical physics
- Computer science
- Biology
- Statistics

Reasoning on Networks

- How do we reason about networks
  - ♣ Empirical: look at large networks and see what you find
  - ♣ Mathematical models: probabilistic, graph theory
  - ♣ Algorithms for analyzing graphsϒ
- What do we hope to achieve from models of networks?
  - ♣ Patterns and statistical properties of network data
  - ♣ Design principles and models
  - ♣ Understand why networks are organized the way they are (predict behavior of networked systems)