

A complex network visualization with numerous nodes (dots) in blue, yellow, and red, connected by thin lines, creating a dense web-like structure.

ARC²S Group

Applied Research on Computational Complex Systems

Dai Karate Club ai nostri gusti
musicali:
analizzare una rete e visualizzare
dati complessi



Tre mattine all'Università
Dipartimento di Informatica
28-02-2012

Prof. Giancarlo Ruffo
ruffo@di.unito.it
<http://arcs.di.unito.it>

Facciamo un gioco:
Cosa hanno in comune?

I Club di Karate



Il cinema



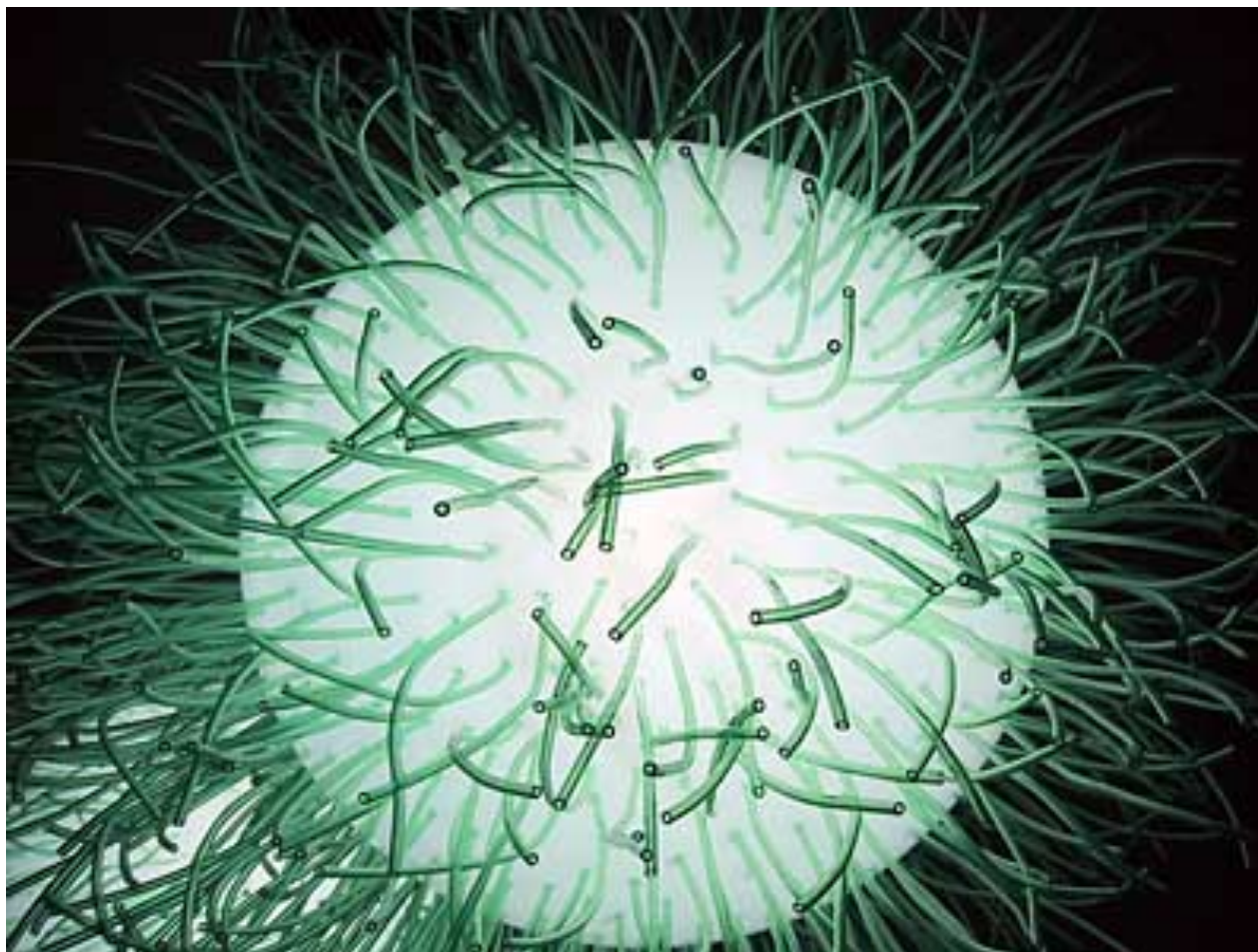
I voli di linea



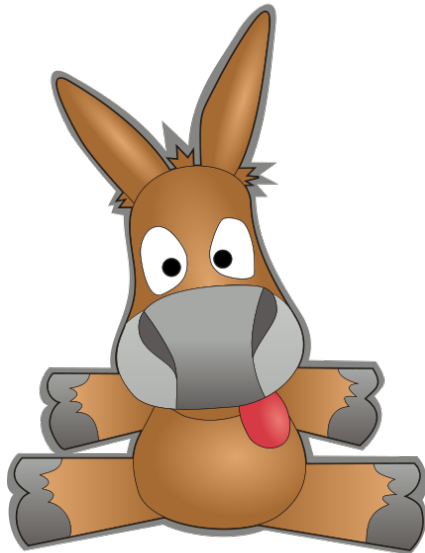
I consigli di amministrazione aziendale



Le epidemie



I Social Media ed il Web




I nostri gusti musicali



Music **Movies**

[See More >](#)

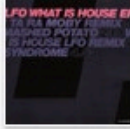
You have music by **Radiohead**





Ga Ga Ga Ga Ga (Bonus Track Ver...
Spoon
Released Jul 10, 2007
★★★★★ 726 Ratings
\$9.99 BUY




You have music by **The Dice Man**





What Is House - EP
LFO
Released Jan 20, 1992
\$3.96 BUY




You have music by **Peter Gabriel**





We Were Here
Joshua Radin
Released Mar 27, 2007
★★★★★ 277 Ratings
\$9.99 BUY




You have music by **Talking Heads**





Faithful
Todd Rundgren
Released Apr 1976
★★★★★ 11 Ratings
\$9.99 BUY




You have music by **Guns N' Roses**





Badmotorfinger
Soundgarden
Released Oct 08, 1991
★★★★★ 363 Ratings
\$9.99 BUY



You have music by **Black Sabbath**

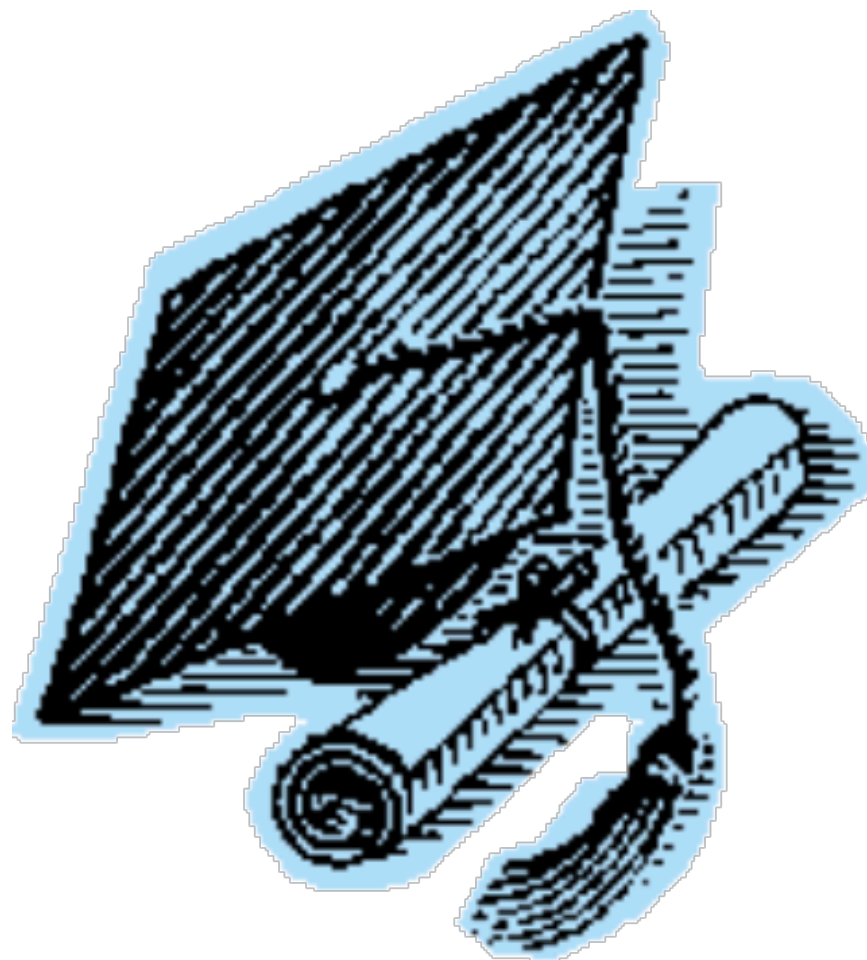


The Essential Ozzy Osbourne
Ozzy Osbourne
Released Jan 28, 2003
★★★★★ 231 Ratings
\$16.99 BUY



Ma, soprattutto,
cosa hanno a che fare
tutte queste cose
con ...

... la scelta di un corso di laurea?





DISTANZE, SEPARAZIONI E CONNESSIONI

Sembrano “mondi” lontanissimi ...



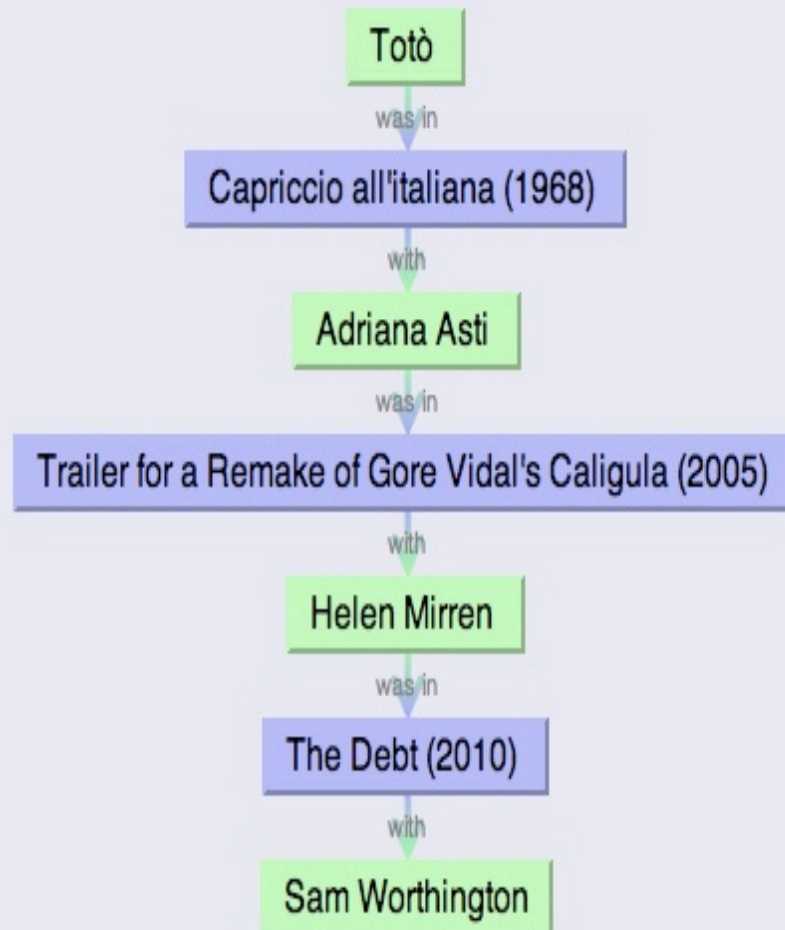
Eppure...

Che “*distanza*” c’è tra Totò e Sam
Worthington?

[The Oracle of Bacon](#)

Totò has a Sam Worthington
number of 3.

Find a different link



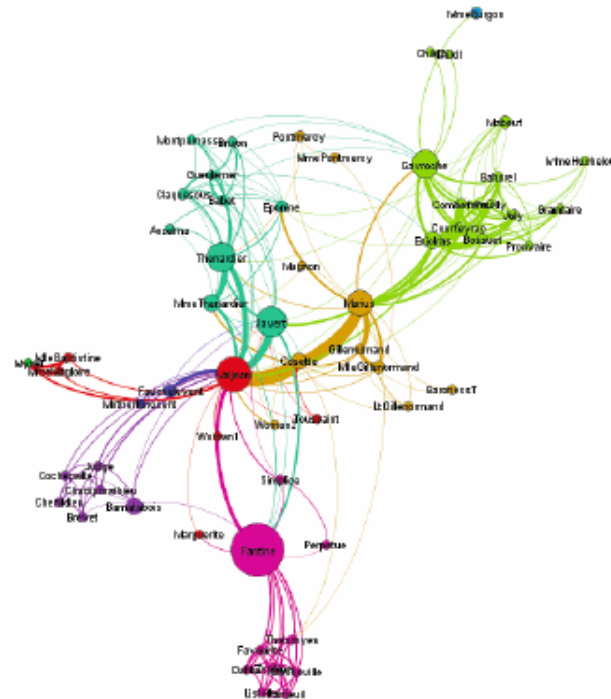
Quanto è piccolo il mondo!

La rete degli attori, così come altre reti sociali,
costituisce un “piccolo mondo”



Il paradigma delle reti (o grafi)

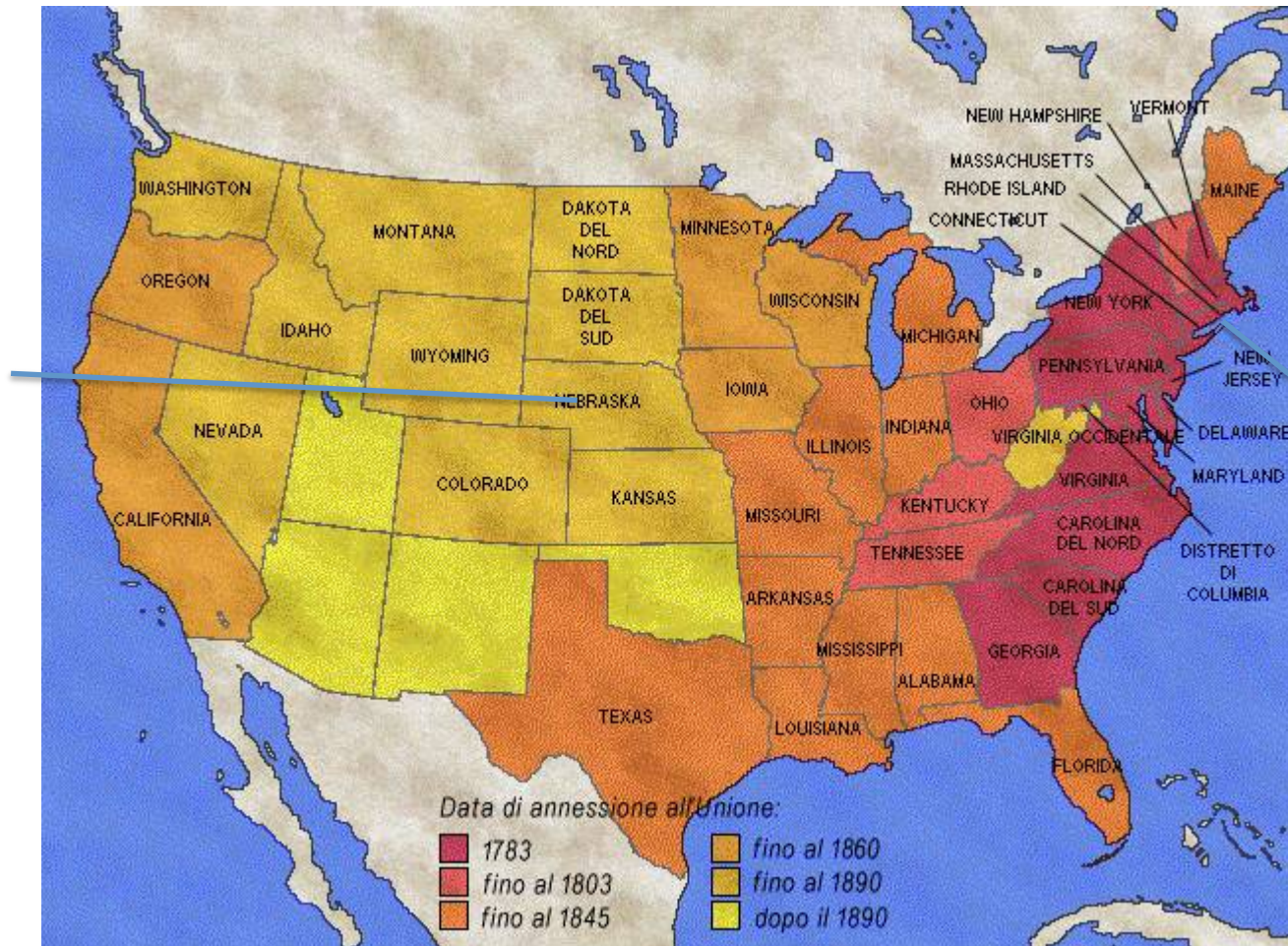
Un sistema sociale (così come altri sistemi **complessi**) si può rappresentare facilmente tramite un **grafo** o una **rete**



I 6 gradi di separazione

L'esperimento di Stanley Milgram (Harvard, anni '60)

Da 160
persone
prese a
caso
In Omaha,
Nebraska...



... ad un
agente di
borsa
di Boston!

Fenomeni emergenti

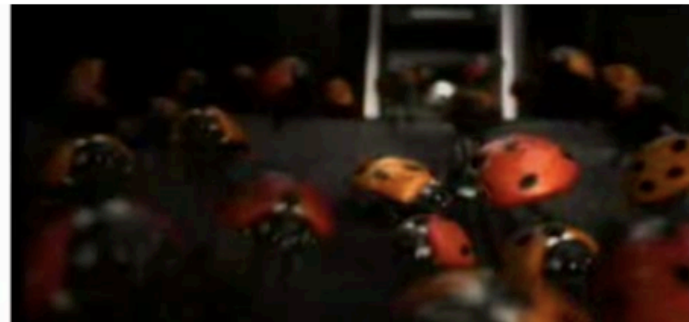
- Il “pattern” **small world** non è l’unico fenomeno che emerge frequentemente dalla complessità
- Altri parametri da osservare:
 - Misure di centralità;
 - Fattori di influenza;
 - Transitività;
 - Invarianza di scala;
 - ...

COS'È LA SCIENZA DELLA COMPLESSITÀ?

Da “elementi semplici” o puntuali



A comportamenti collettivi **complessi**



Riduzionismo e Costruttivismo

- I fenomeni di alto livello possono essere “*ridotti*” a quelli di basso livello;
- Scendendo di livello, “*puntiamo la freccia*” verso la **meccanica quantistica**;
- Può funzionare il programma “*costruttivista*” inverso?
- Se così fosse, potremmo **prevedere e spiegare *tutti*** i fenomeni a partire dalla fisica fondamentale

La complessità secondo P. W. Anderson (1972)

X	Y
Fisica dello stato solido o fisica dei molti corpi	Fisica delle particelle elementari
Chimica	Fisica dei molti corpi
Biologia molecolare e biologia cellulare	Chimica
*	*
*	*
*	*
Psicologia	Fisiologia
Scienze sociali	Psicologia

Questa gerarchia **non** implica che la scienza **X** sia la *semplice applicazione* di **Y**

Altro argomento di Anderson

- Pensiamo alla Macchina di A. Turing: universale
- Problema della fermata e tesi di A. Church
 - Nessun programma può decidere se un altro programma si fermerà
 - Macchina di Turing (o modello equivalente) *sufficiente* a calcolare un problema
- Qualsiasi sistema di entità capace di “trasportare” simboli 0 e 1 può essere usato per calcolare



Esempio “Computazionale” di Anderson

- *Il modello computazionale non è riducibile alla Fisica*
- Il modello computazionale è **multiplatforma**
 - potete implementare una *macchina di Turing* con dei secchi riempiti con livelli differenti di acqua per rappresentare “0” e “1” e poi trasportarli avanti ed indietro eseguendo certe regole (“esperimento” che il biologo Stuart Kauffman riporta di fare eseguire ai propri neolaureati...)

Un altro esempio: le reti tecnologiche

- In particolare: Internet

Livello Applicazione (es. HTTP, SMTP, FTP)

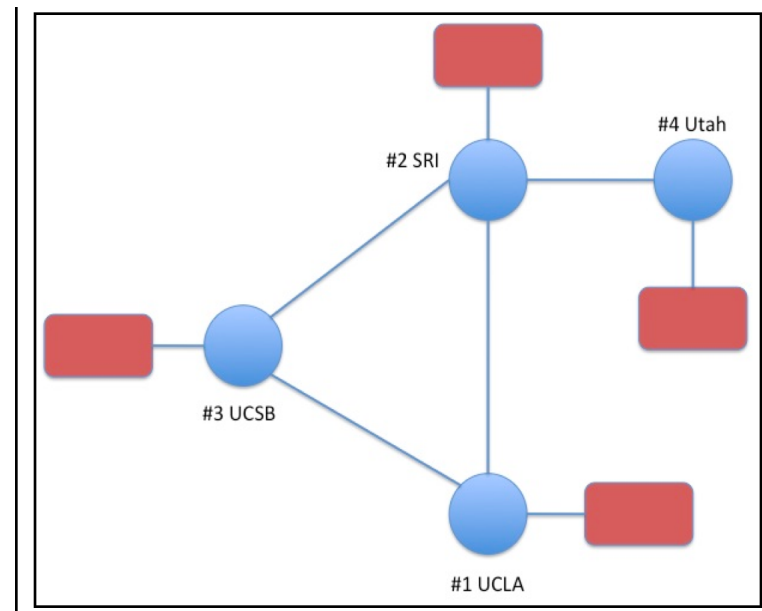
Livello Trasporto (es. TCP, UDP)

Livello Rete (es. IP, BGP)

Livello Collegamento (es. Ethernet, ARP, PPP, HDLC)

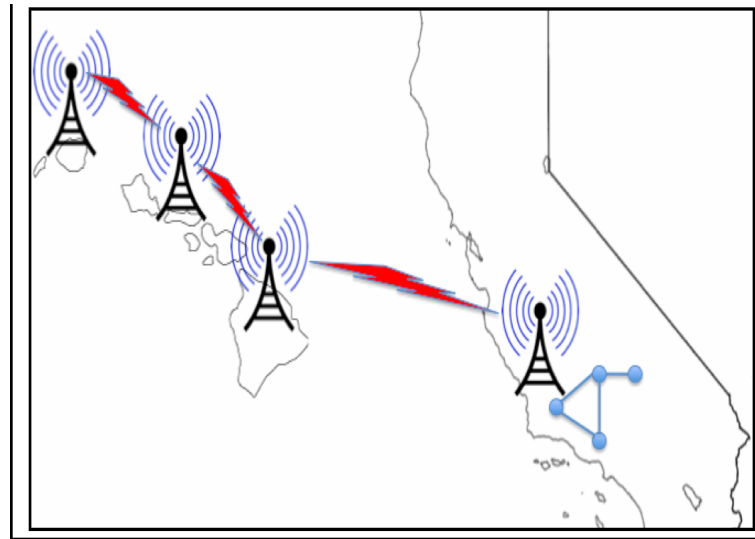
Livello Fisico (es. DSL, Wi-Fi, 10Base-T, OTN)

“Costruiamo” l’Internet



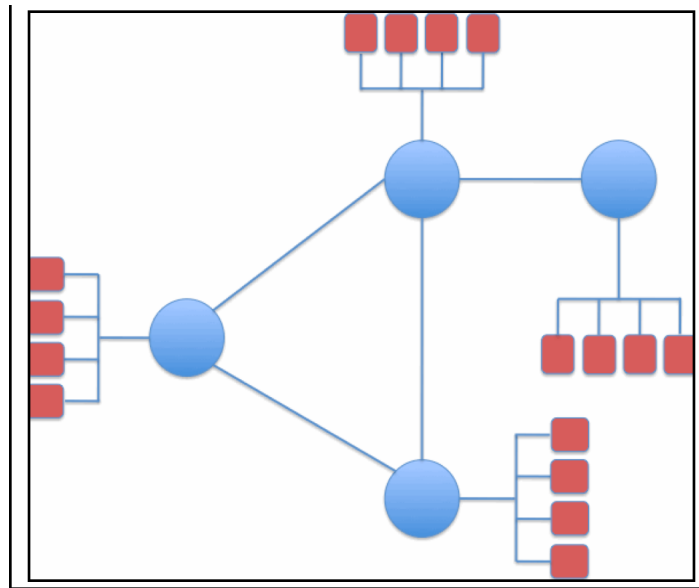
Arpanet nel 1969

“Costruiamo” l’Internet



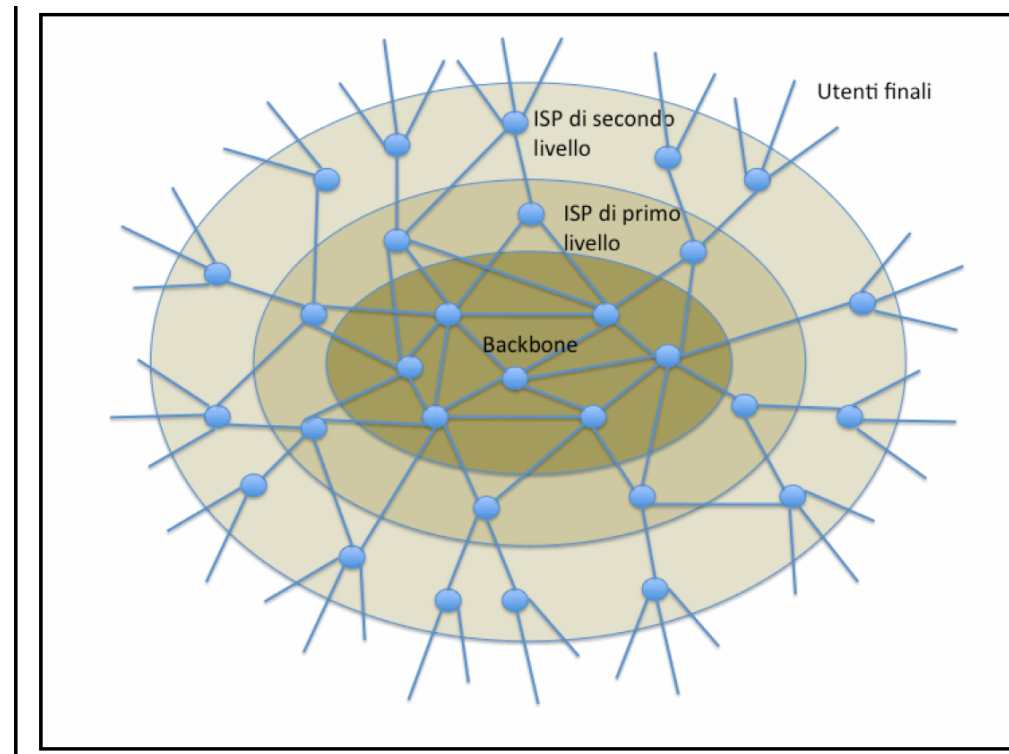
ALOHANet nel 1969

“Costruiamo” l’Internet



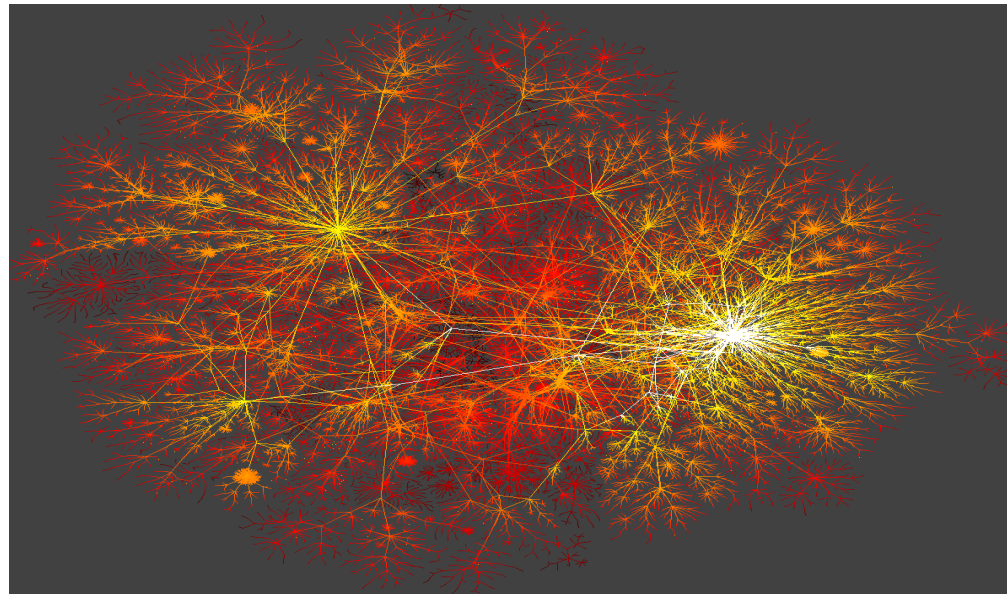
Ethernet, 1973

“Costruiamo” l’Internet



Internetworking e routing, TCP/IP, etc. 1978

“Costruiamo” l’Internet



AS network (“foto” inizio anni 2000)

... per arrivare al Web



ed i social media (in particolare)

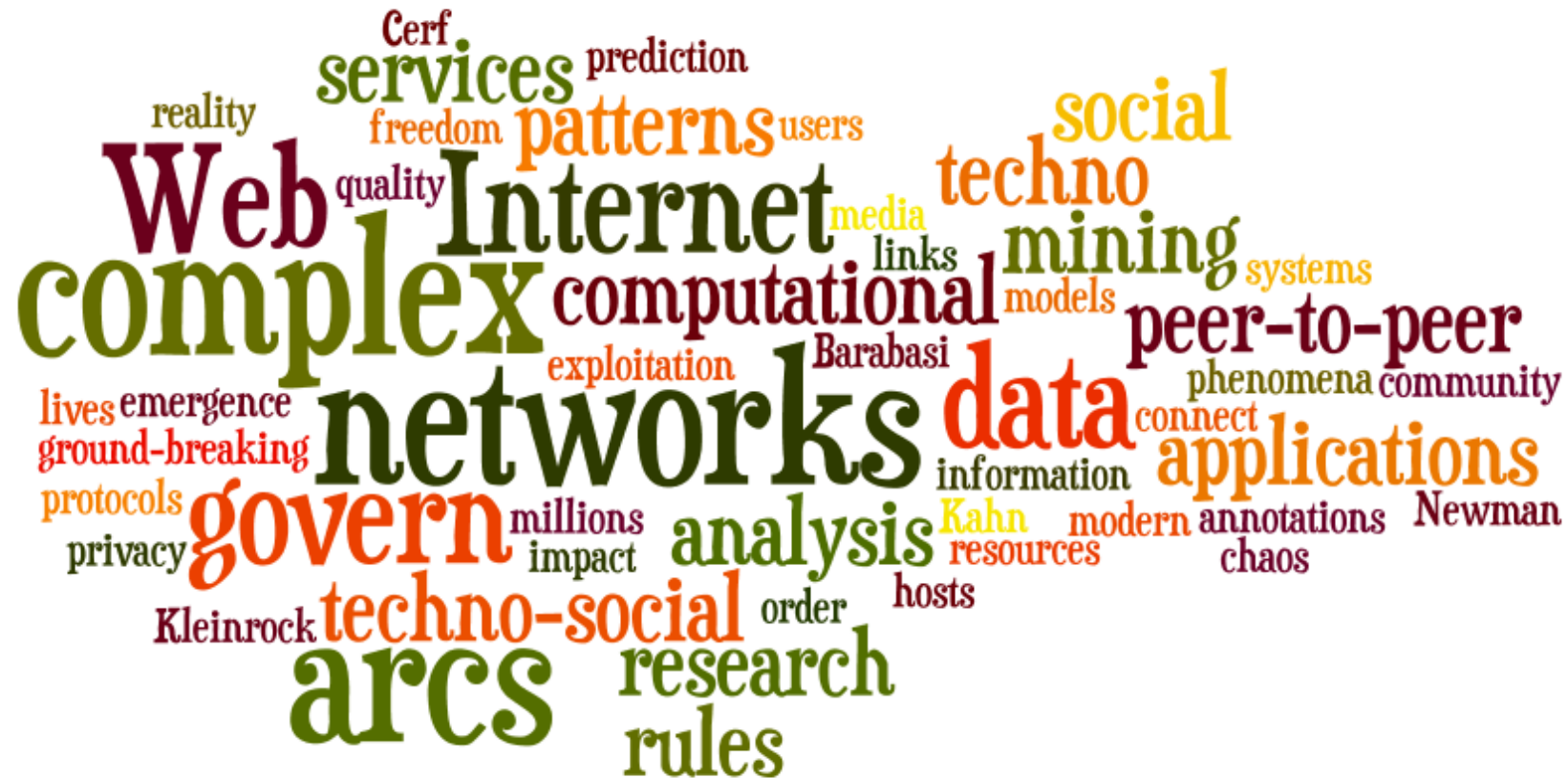


Avremmo potuto “prevedere” questo:

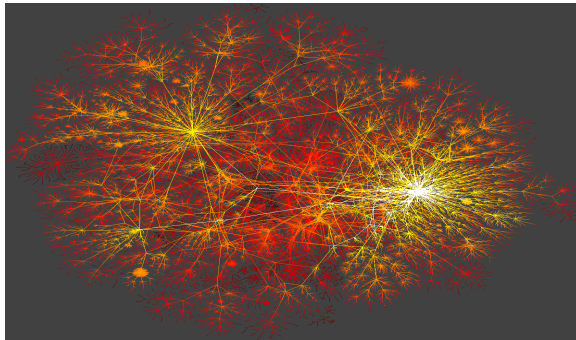


dalla fisica elementare?
o dai primi esperimenti di Kleinrock?

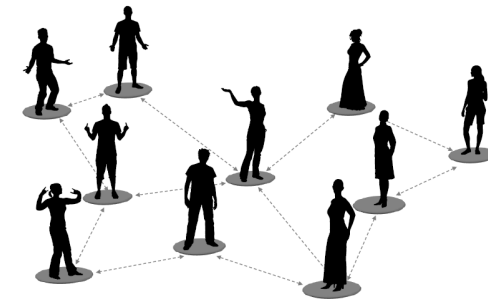
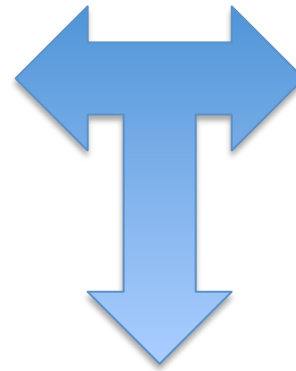
Techno-Social System



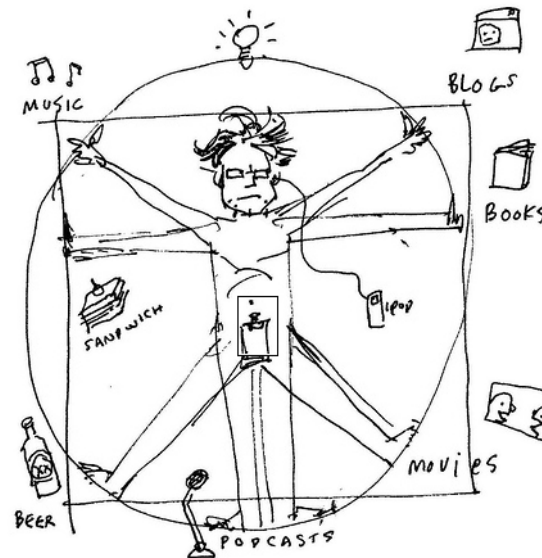
Livelli di complessità



reti tecnologiche



reti sociali



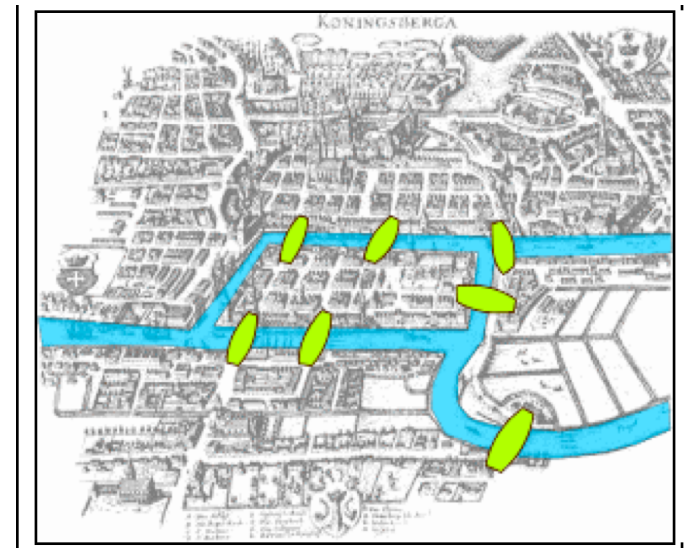
**Le reti
socio-tecnologiche**

Analizzare le reti

Un modello per analizzare reti

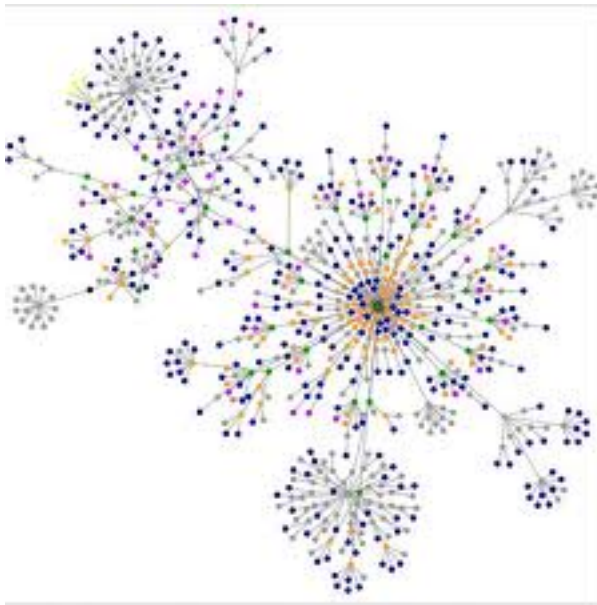


Il grafo!

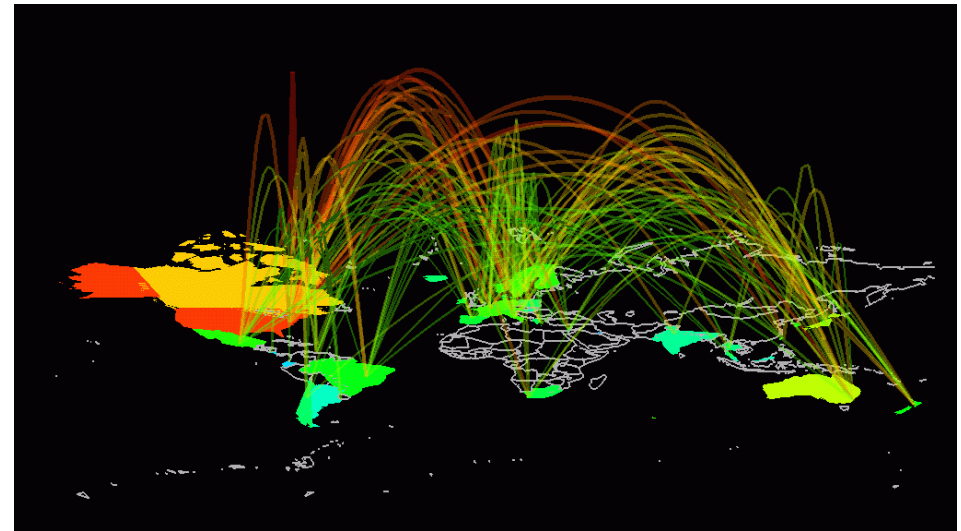


Ponti di Königsberg!

Un modello per analizzare reti socio-tecnologiche



Il grafo!

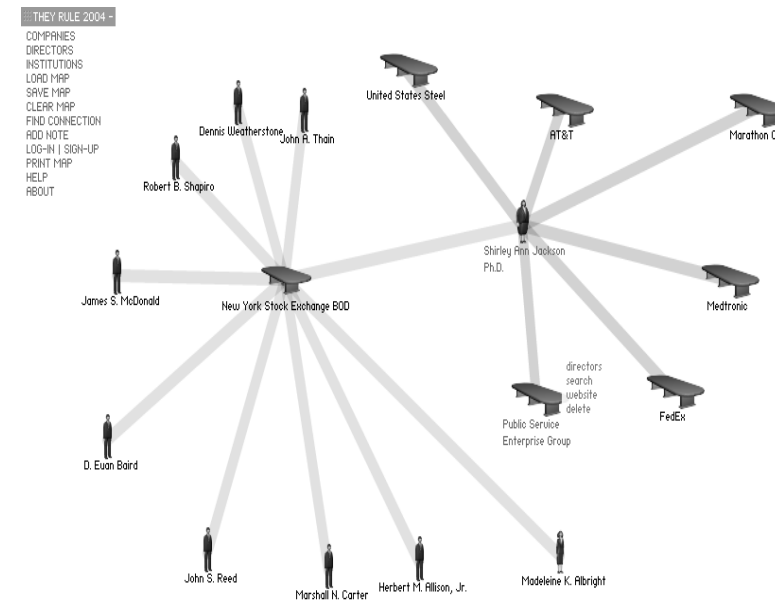


Reti di trasporto

Un modello per analizzare reti socio-tecnologiche



Il grafo!



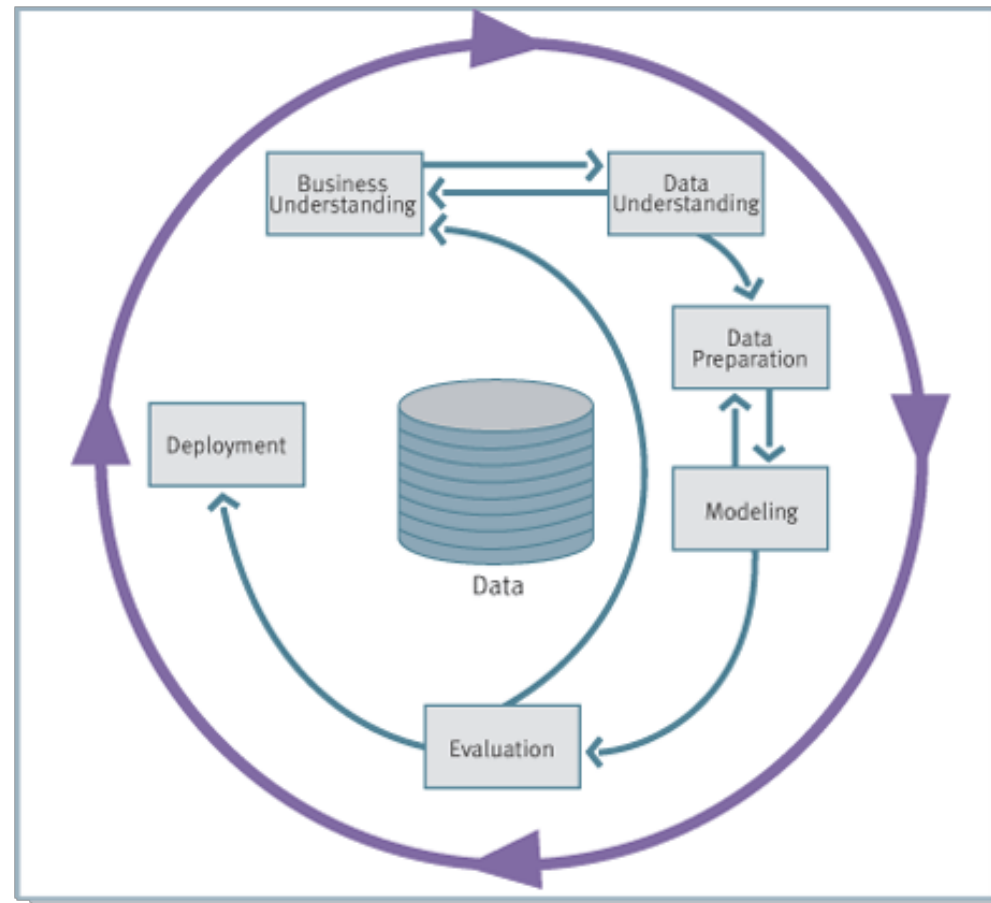
Consigli di amministrazione

Dati

- Rispetto al passato e alle classiche ricerche dei sociologi che attingevano alle informazioni in forma diretta (questionari), abbiamo risorse computazionali per raccogliere **enormi quantità di dati**

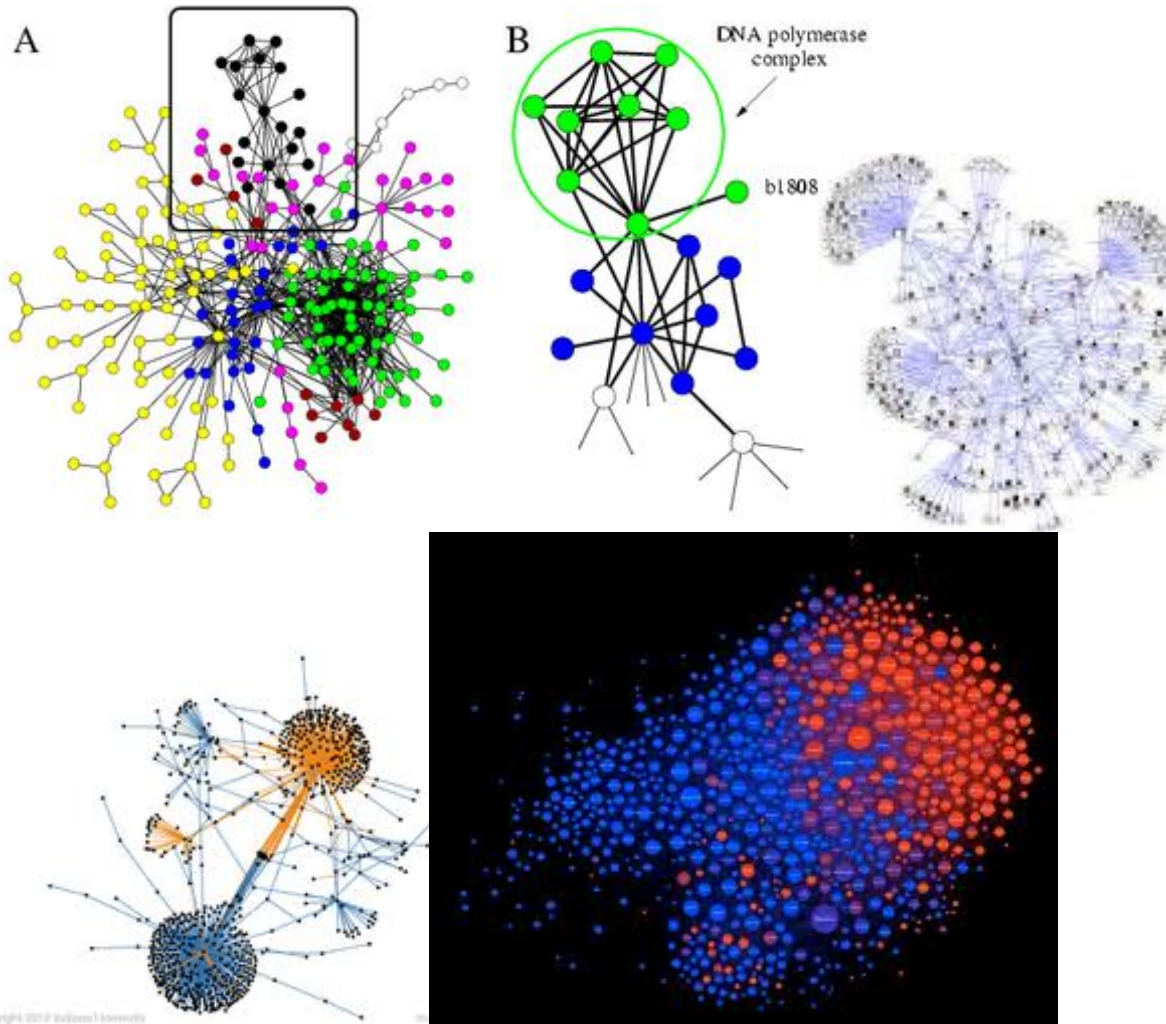


Analisi dei dati



Data Mining

Analisi di reti complesse



```
>>> g = networkx.DiGraph()
```

```
>>> g.add_edges_from([("A","B"), ("C","A")])
```

```
>>> print g.in_degree(with_labels=True)
{'A': 1, 'C': 0, 'B': 1}
```

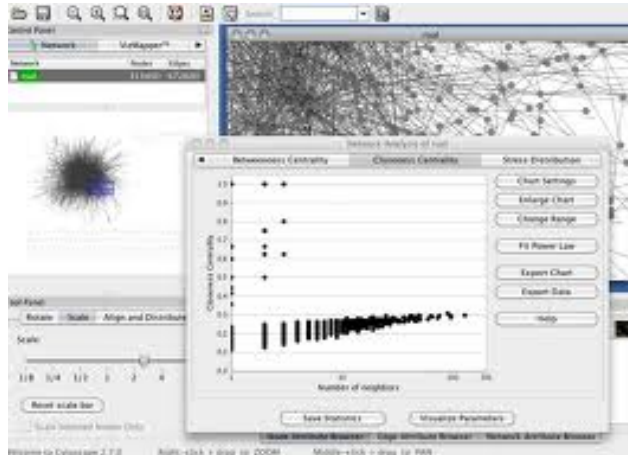
```
>>> print g.out_degree(with_labels=True)
{'A': 1, 'C': 1, 'B': 0}
```

```
>>> print g.neighbors("A")
['B']
```

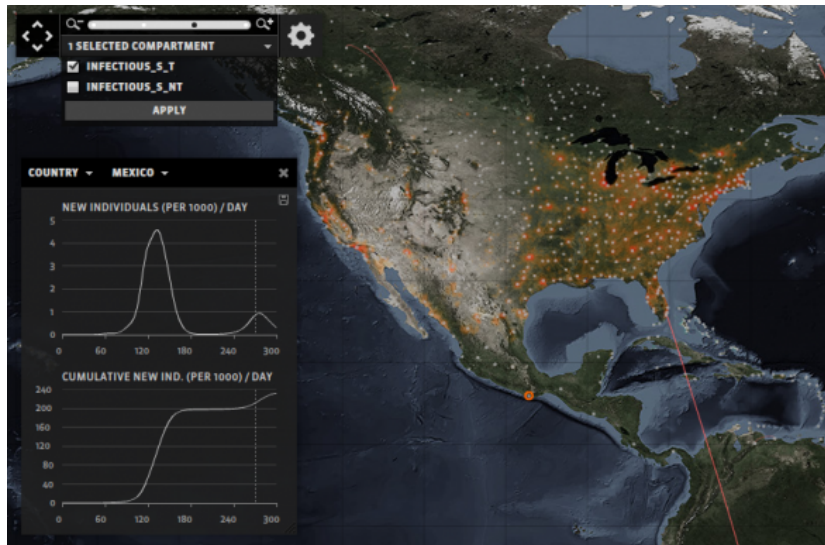
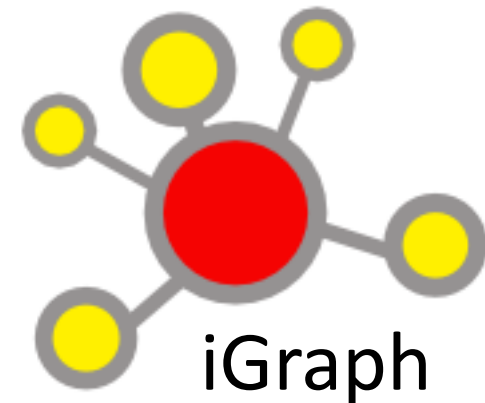
```
>>> print g.neighbors("B")
[]
```

```
>>> ug = g.to_undirected()
>>> print ug.neighbors("B")
['A']
```


Strumenti di analisi



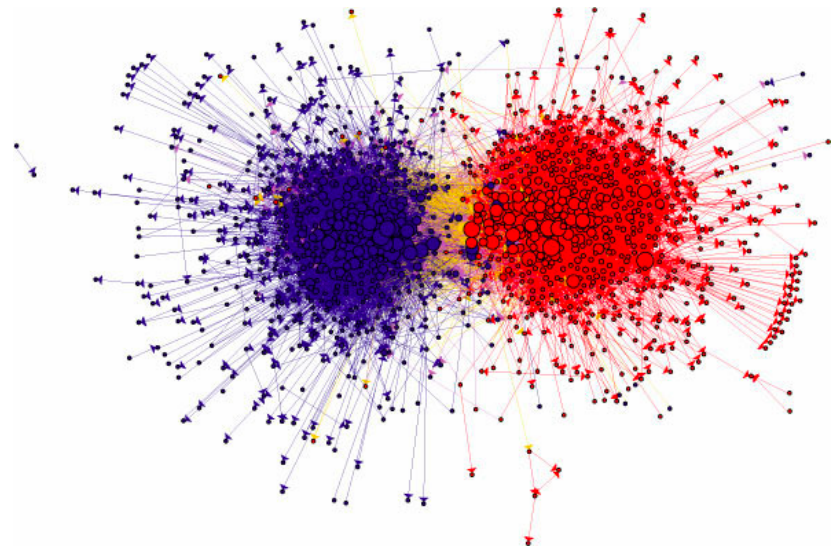
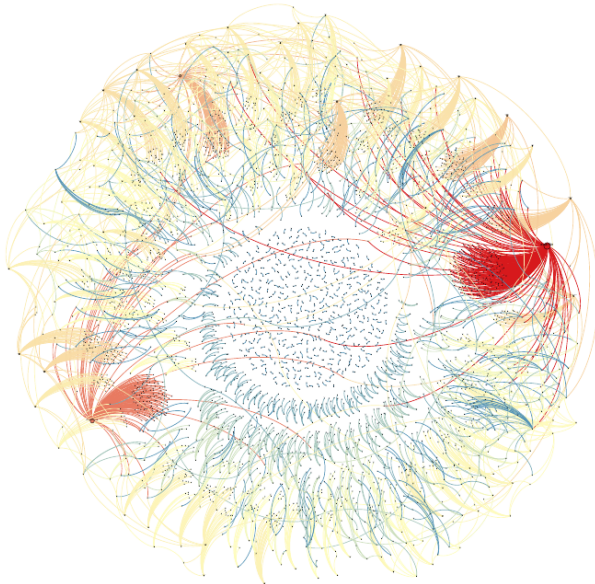
Gp Gephi



NetworkX

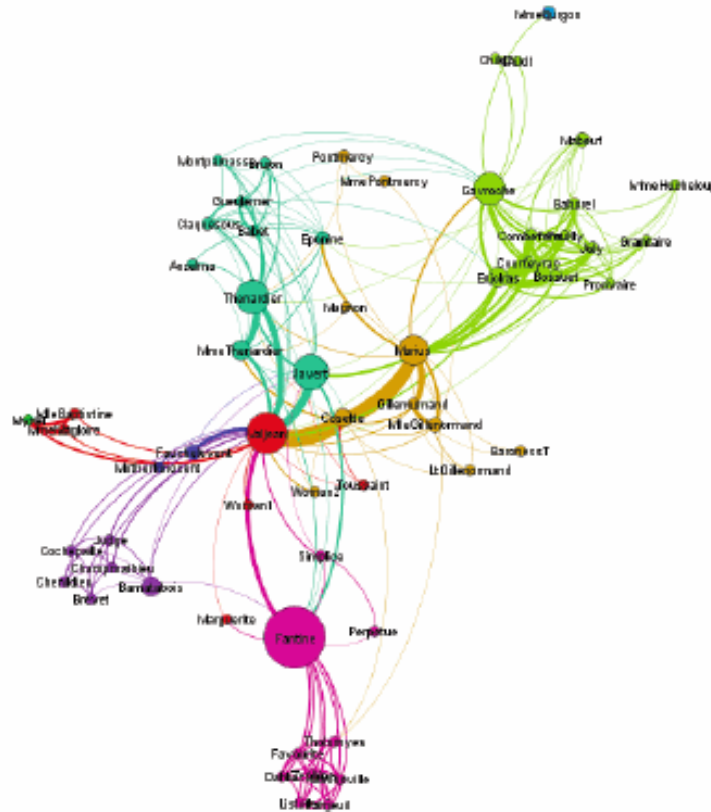
Aspetti di visualizzazione

- Domande possibili:
 - Come si propaga un'informazione (o un'infezione, una notizia falsa, etc.)
 - Quali “comunità” emergono?



Un esempio

- Usare  Gephi per analizzare una rete



Applicazioni e servizi

A cosa serve analizzare le reti?

Tecniche nascono in ambito matematico e si trasformano in strumenti attraverso
l'informatica

Due macro aree di applicazione:

Scienze sociali
ed umane

Scienze “dure”

L'esperto di Network Analysis assiste
spesso all'intero ciclo: si tratta di un ambito
molto **interdisciplinare**

Scienze sociali
ed umane

Stanley Milgram

- **Stanley Milgram (NewYork 1933 –NewYork 1984)**
sociologo americano noto per i suoi studi sull'autorità e l'obbedienza

- The Milgram experiment 18



- The lost-letter experiment



- The small-world experiment

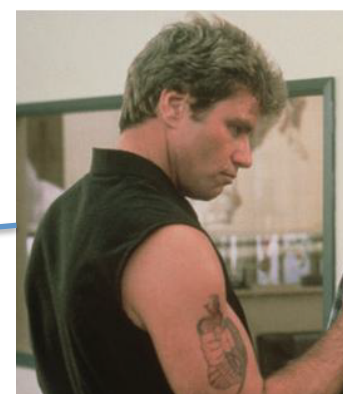
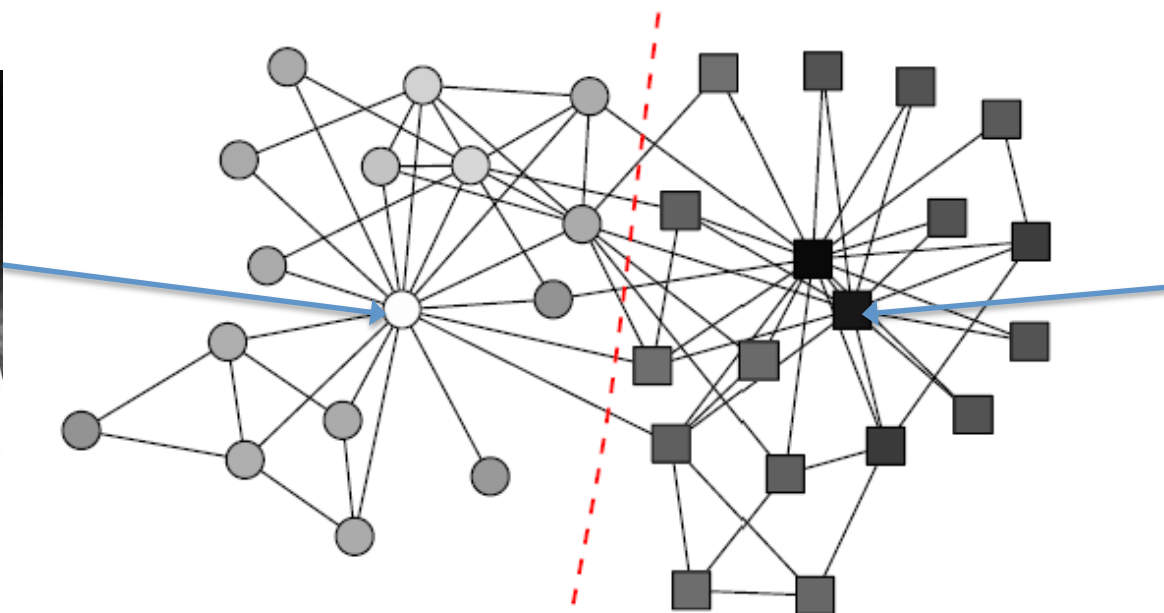
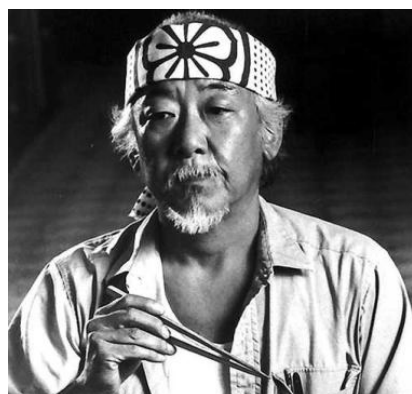


Come si raccoglievano i dati?

- Questionari
- Svantaggi:
 - Limitato dal tipo di domanda (e dal formato della risposta)
 - Difficile lavorare sui grandi numeri
 - Come si ottengono i dati delle reti “clandestine”?

Yes...	No...	368
Yes.✓	No...	
Yes...	No...	
Yes...	No...	
Yes.✓	No...	
Yes...	No...	
Yes.✓	No...	3
Yes...	No...	
Yes...	No...	
1..2..3✓4..5..6..7..		6
8..9..10.. [See also sheet]		

Karate club di Zachary (1977)



Si possono studiare le relazioni all'interno di un gruppo sociale e come si formano le **sottocomunità**

Nel mondo digitale

- I Social Media hanno rivoluzionato lo studio dei comportamenti sociali
- Gli utenti mettono di loro iniziativa tantissimi dati a disposizione...
- ... di chi?
- In genere, chi ha i dati **se li tiene**



E se abbiamo accesso ai dati?

- Ad esempio, facendo **crawling** intensivo
- Non è così semplice elaborare tutti questi dati!
- **Massive data analysis**



“Raccontare storie” con la visualizzazione dei dati

Journalism in the Age of Data

A video report on data visualization as a storytelling medium
Produced during a 2009-2010 Knight Journalism Fellowship
Total Running Time: 54 Minutes; with related information and links



CHAPTERS

I. Introduction

II. Data Vis in Journalism

III. Telling "Data Stories"

IV. A New Era in Infographics

V. Life as a Data Stream

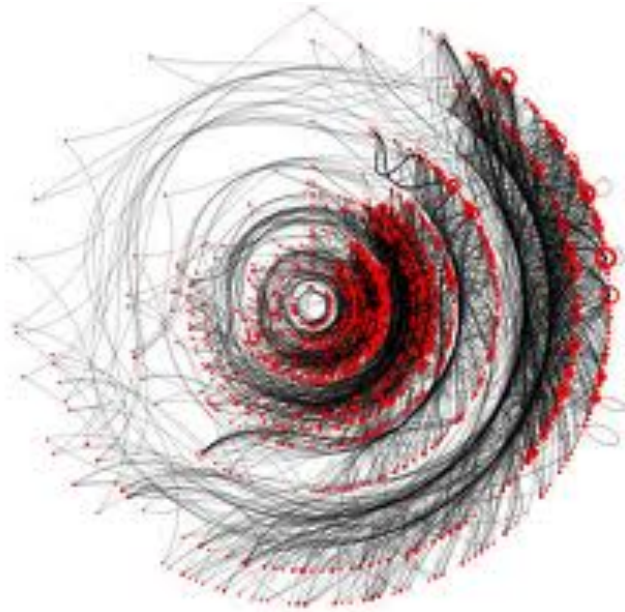
VI. Exploring Data

VII. Technologies and Tools

VIII. First Steps

RELATED INFORMATION FOR THIS CHAPTER

Visualizzare dati “collegati”



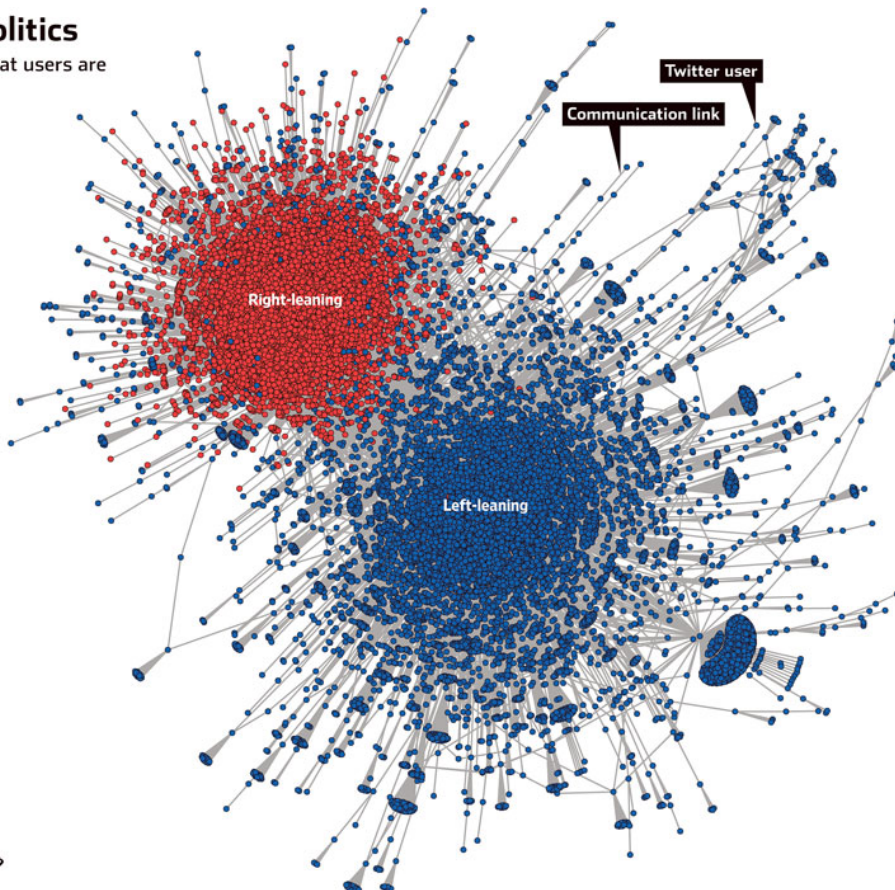
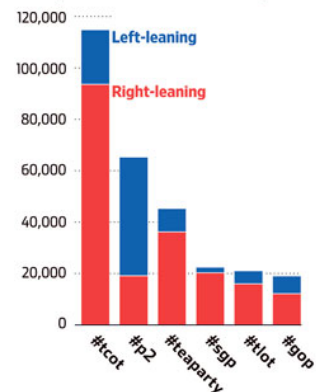
Scienze politiche

Twitter's Divided Politics

Political Twitter traffic reveals that users are polarized along party lines.*

Researchers at Indiana University analyzed 250,000 Twitter messages on political topics exchanged by 45,000 people during the 2010 mid-term congressional elections. This chart of 'retweets'—in which one user forwards another's message—shows that, though there were more left-leaning users, right-leaning users were more densely connected to one another. (Each dot is a Twitter user, and the lines show retweets.) Even so, as the chart illustrates, lines of communication do sometimes reach across the political divide.

The most popular hashtags (short codes signaling the message's content), shown by number of tweets. Researchers found that users on the left and right use each other's hashtags.



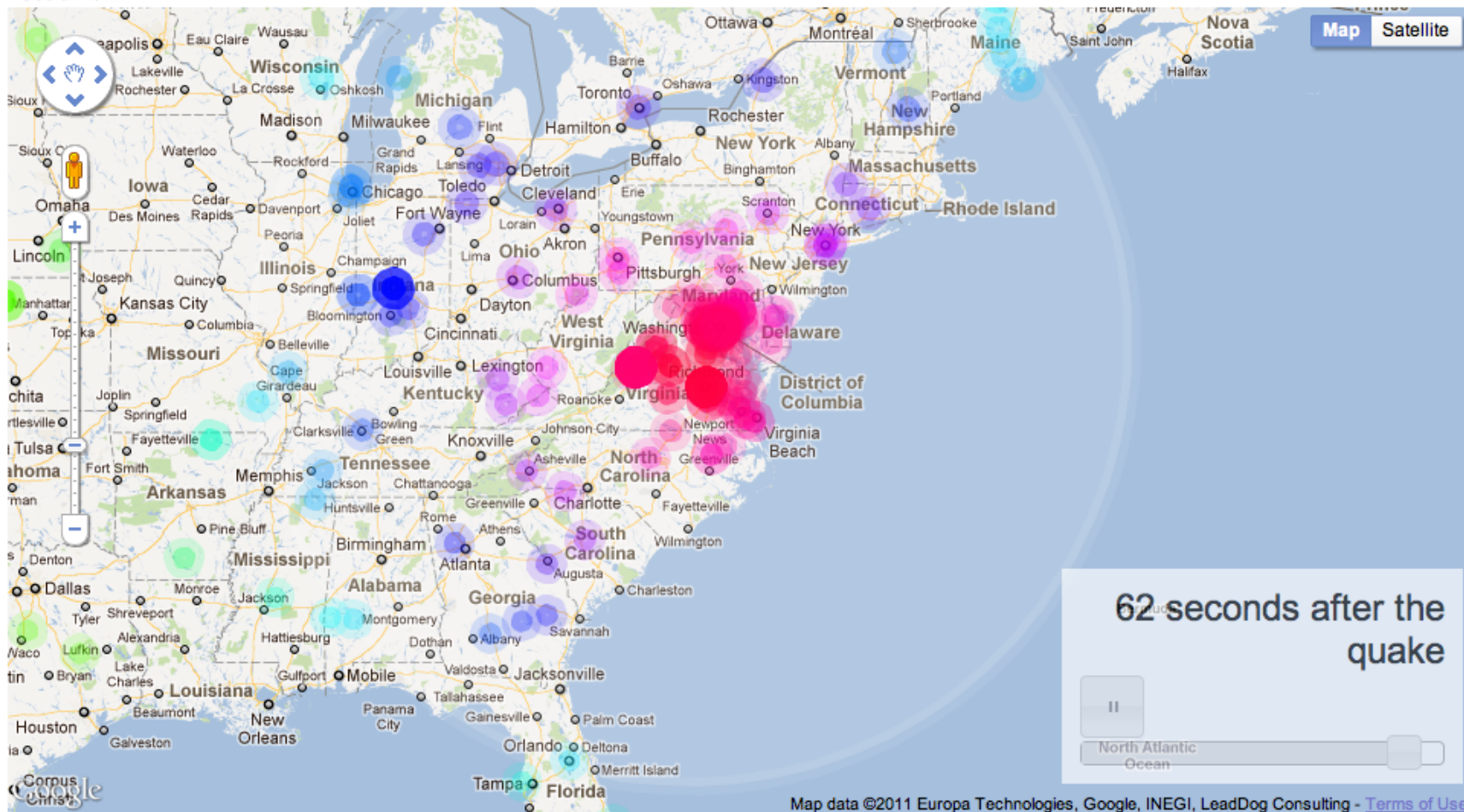
Hashtags: tcot, top conservatives on Twitter; p2, progressives 2.0; sgp, smart girl politics; tlol, top libertarians on Twitter.

*Data show 'retweets' of other users' messages. Political leaning designations are based on algorithmically-determined communities of users which correlate with political affiliation. Source: Center for Complex Networks and Systems Research, Indiana University

Gestione unità di crisi

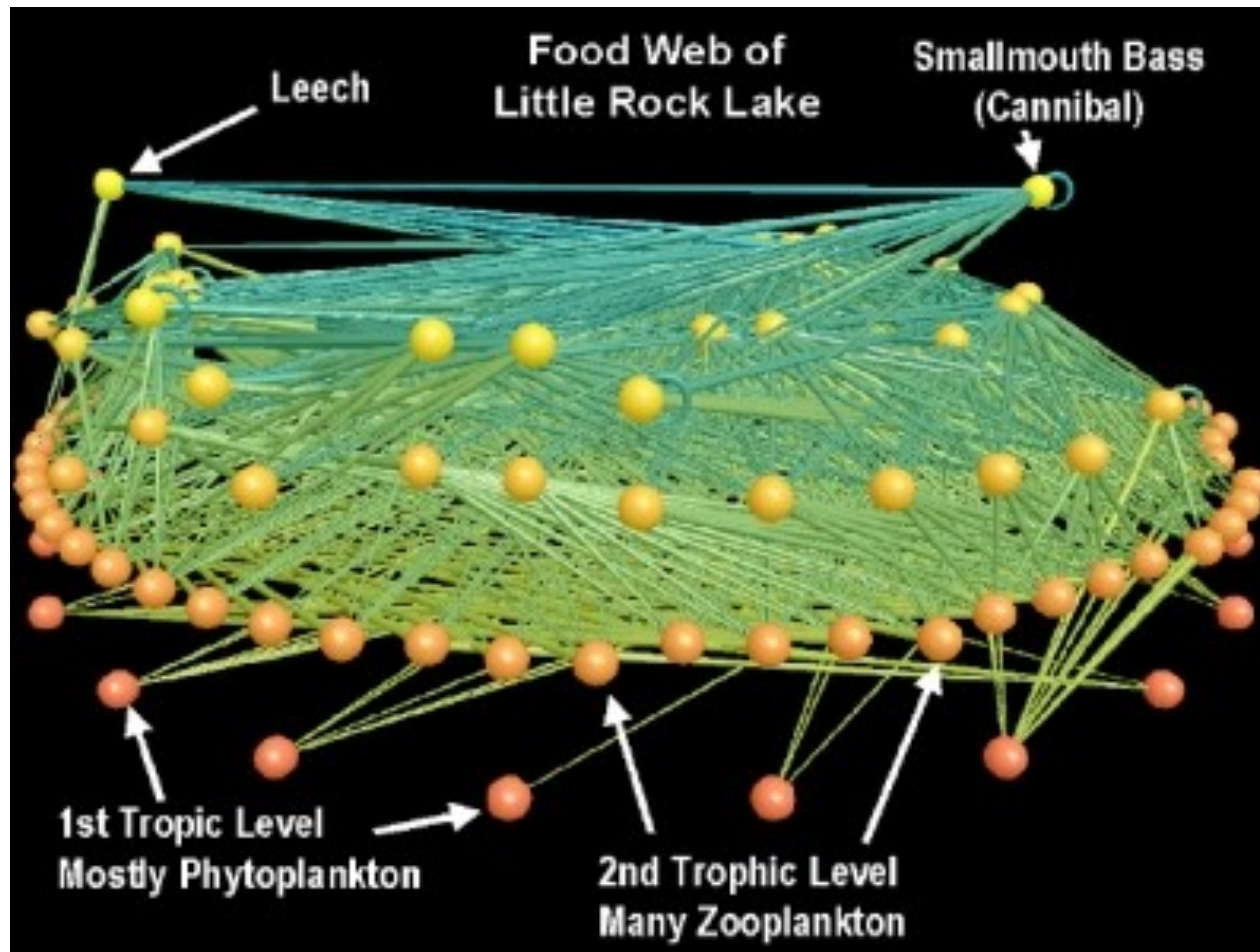
Twitter Responses to the Virginia Earthquake

The visualization below (created by SocialFlow) replays a 90-second spread of earthquake-related Tweets across North America, from the moment the earthquake hit Mineral Virginia (1:51PM). The colors represent distance from the epicenter and the circle size and opacity represent the number of tweets. *Sources: Twitter, SocialFlow.*



Scienze “dure”

Reti biologiche ed ecologiche

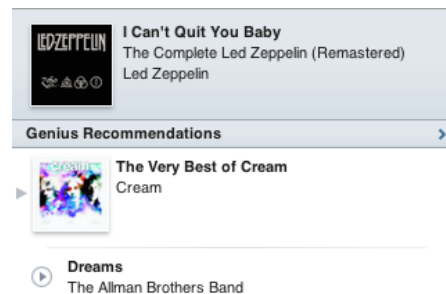
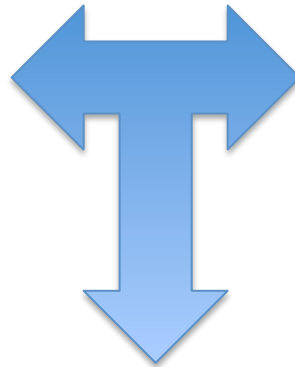
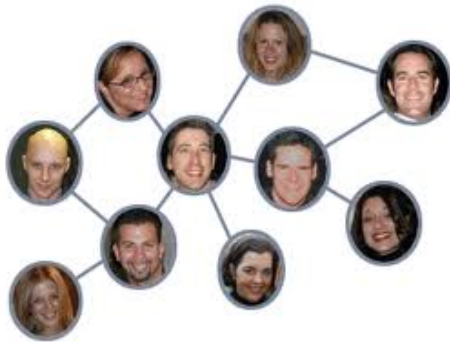


Predire la diffusione delle epidemie



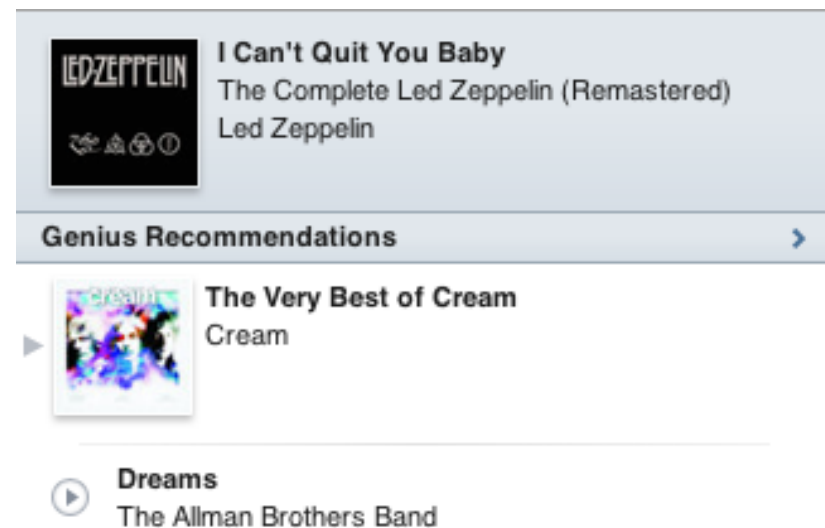
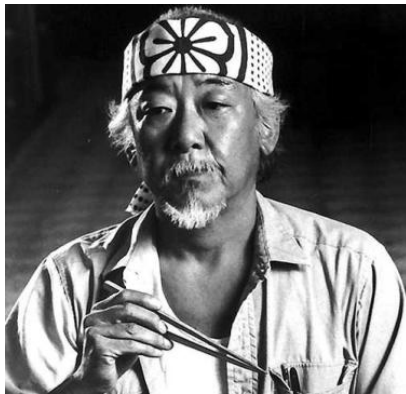
Reti di similarità e “raccomandazioni”

- *Costruiamo la rete da “somiglianze” esistenti (es., gusti musicali, generi)*



Conclusione

- Probabilmente hanno molte altre cose avranno in comune ...



- Queste cose aspettano solo di essere trovate!



Thank
you!

ARC²S People



G. Ruffo



M. Deplano



S. Spoto



L. Aiello



R. Schifanella



E. Sulis



A. Panisson