

PARTIMOS EN BREVE

MUCHAS GRACIAS

Network Science

dataScience UDD



Cristian Candia-Castro Vallejos, Ph.D.

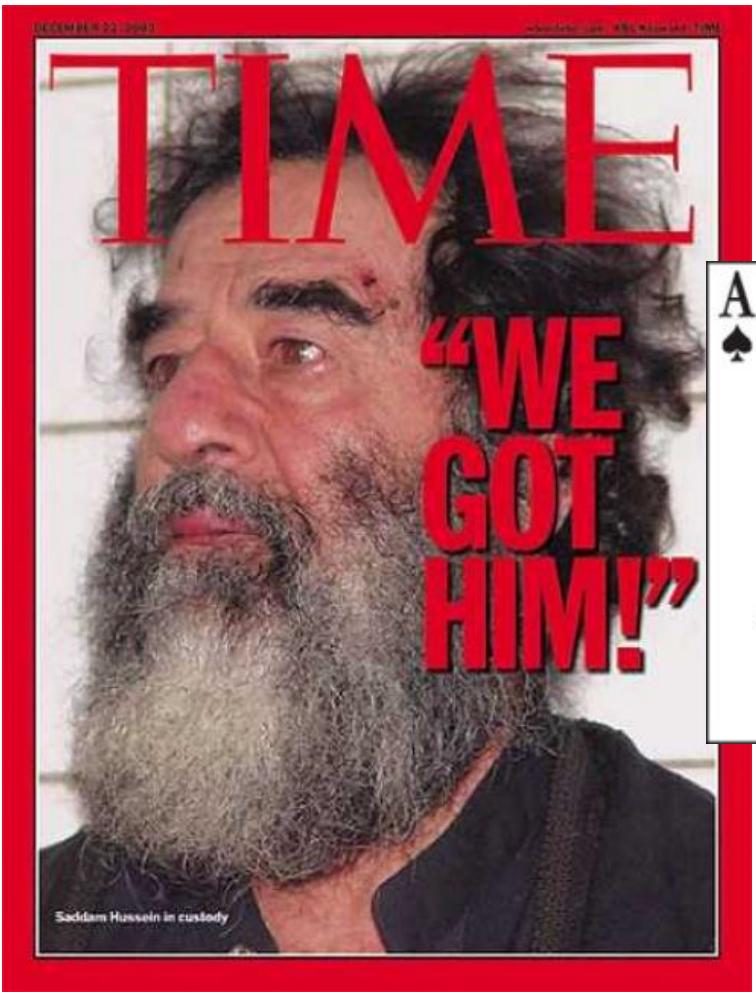
cristiancandia@udd.cl

Director Magister en Data Science UDD
Profesor Investigador, Facultad de Ingeniería, UDD
External Faculty Northwestern Institute on Complex Systems, Kellogg School of
Management, Northwestern University

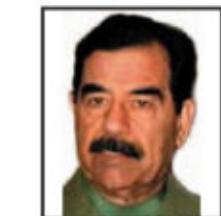
DESDE SADDAM HUSSEIN HASTA TEORÍA DE REDES

Una simple historia(1)

El destino de Saddam y network science



A
♠



SADDAM HUSSEY AL-TIKRITI
President

♦
A



Qusay Hussein



Abid Hamid



Uday Hussein
Majid



Saddam Hussein
Majid



Ibrahim Hasan
Khatab



Barzan Ibrahim
Khatab



Sabha Talfah
Musslit



Basim Latif



Kharallah Talfah
Musslit



Mohammad Ibrahim
Musslit



Sulwan Ibrahim
Musslit



Mohammad Khudayr



Rudman Ibrahim
Musslit



Nasir Yasin
Musslit



Faris Yasin
Musslit

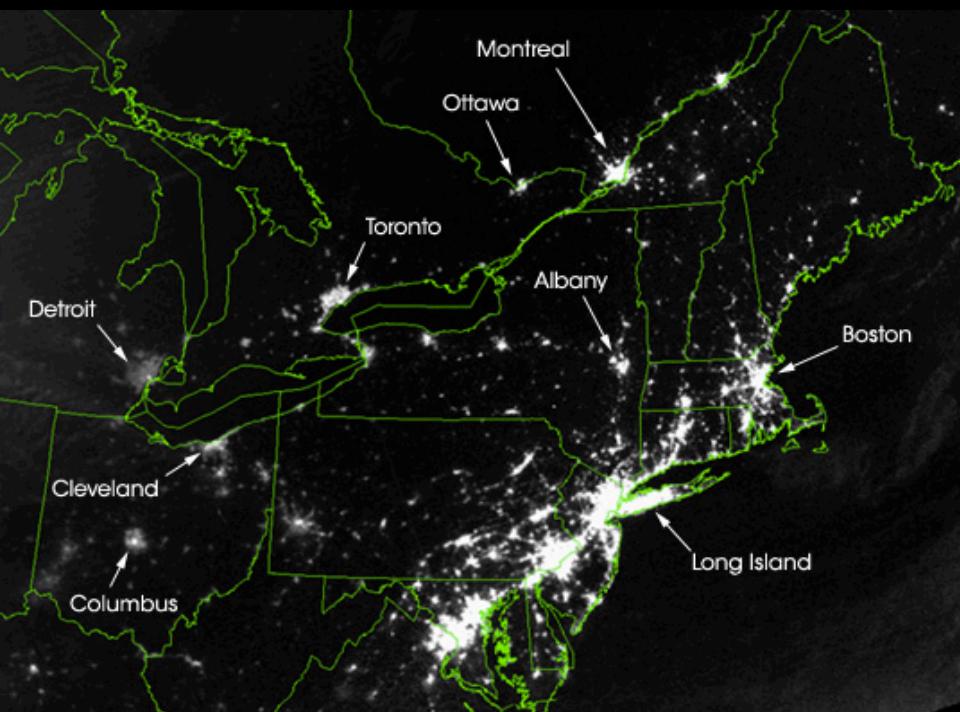
La captura de Saddam Hussein:

- Muestra el fuerte poder predictivo de las redes.
- subraya la necesidad de obtener mapas precisos de las redes que pretendemos estudiar; y las dificultades a menudo heroicas que encontramos durante el proceso de mapeo.
- demuestra la notable estabilidad de estas redes: la captura de Hussein no se basó en una nueva inteligencia, sino en sus vínculos sociales previos a la invasión, desenterrados de fotos antiguas apiladas en su álbum familiar.
- muestra que la elección de la red en la que nos enfocamos marca una gran diferencia: el árbol jerárquico, que capturó la organización oficial del gobierno iraquí, no sirvió de nada cuando se trató del paradero de Saddam Hussein.

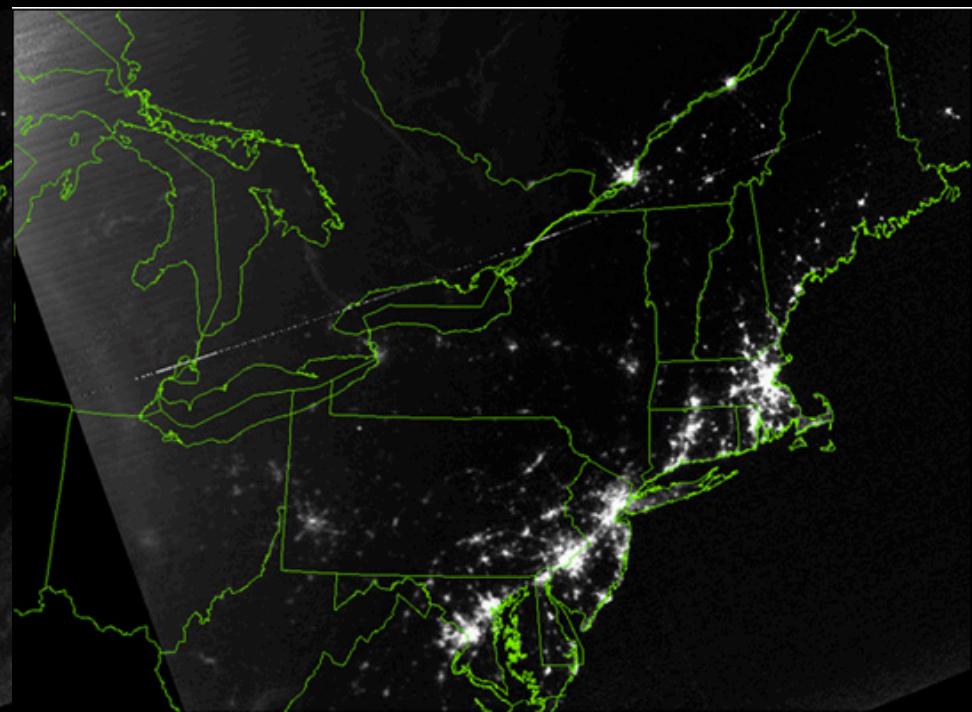
VULNERABILIDAD DEBIDA A LA INTERCONECTIVIDAD

UNA SIMPLE HISTORIA (2): Apagón del 15 de Agosto de 2003.

Robustez de un sistema complejo



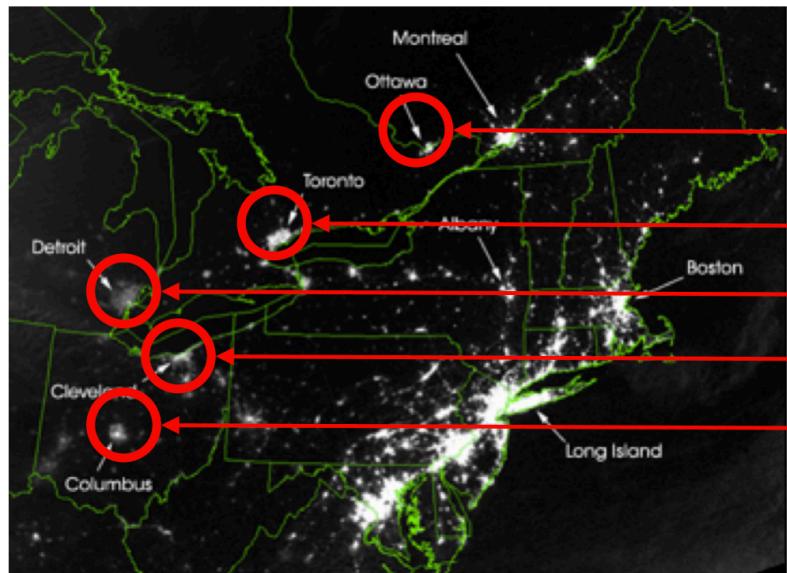
Agosto 14, 2003: 9:29pm EDT
20 horas antes



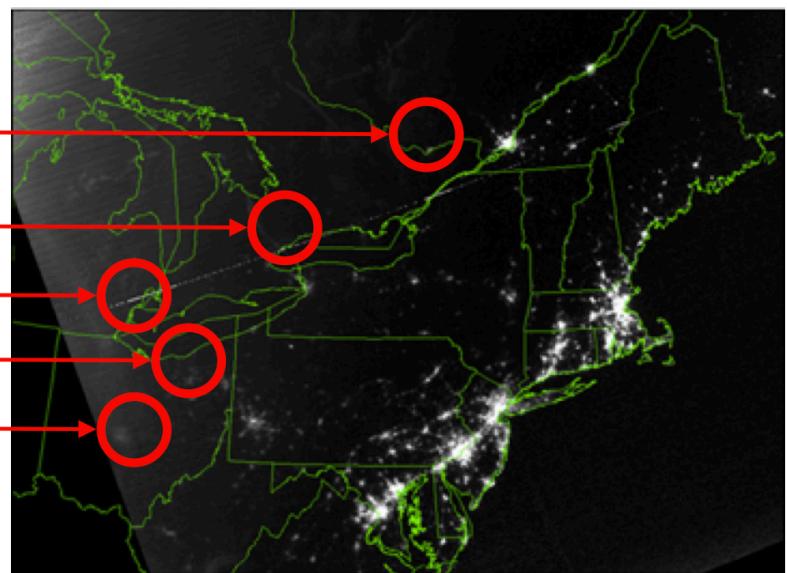
Agosto 15, 2003: 9:14pm EDT
7 horas después

August 15, 2003 Northeast Blackout

Cascade Failure



August 14, 2003: 9:29pm EDT
20 hours before



August 15, 2003: 9:14pm EDT
7 hours after

UNA SIMPLE HISTORIA (2): Apagón del 15 de Agosto 15 de 2003.

Un importante tópico en Network Science:

- Debemos entender como la estructura de las redes afecta la robustés de un Sistema complejo.
- desarrollar herramientas cuantitativas para evaluar la interacción entre la estructura de la red y los procesos dinámicos en las redes, y su impacto en las fallas.
- Aprenderemos que las fallas en la realidad siguen leyes reproducibles, que se pueden cuantificar e incluso predecir utilizando las herramientas de la ciencia de redes..

Redes en el Corazón de los Sistemas Complejos



*“I think the next century
will be the century
of complexity.”*

Stephen Hawking
23 de Enero del 2000

SISTEMAS COMPLEJOS

enclave
RAE

complejo, ja

Del lat. *complexus*, part. pas. de *complecti* 'enlazar'.

1. adj. Que se compone de elementos diversos.
2. adj. complicado (l) enmarañado, difícil).
3. m. Conjunto o unión de dos o más cosas que constituyen una unidad. *Complejo vitamínico*.
4. m. Conjunto de establecimientos industriales generalmente próximos unos a otros.
5. m. Conjunto de edificios o instalaciones agrupados para una actividad común.

Fuente: <http://dle.rae.es>

Complejidad, una **teoría científica** que afirma que algunos sistemas muestran fenómenos de comportamiento que son completamente inexplicables por cualquier análisis convencional de las partes constituyentes de los sistemas. Estos fenómenos, comúnmente conocidos como comportamiento emergente, parecen ocurrir en muchos sistemas complejos que involucran tanto organismos vivos, como un mercado de valores o el cerebro humano.

Fuente: John L. Casti, *Encyclopædia Britannica*

Complexity

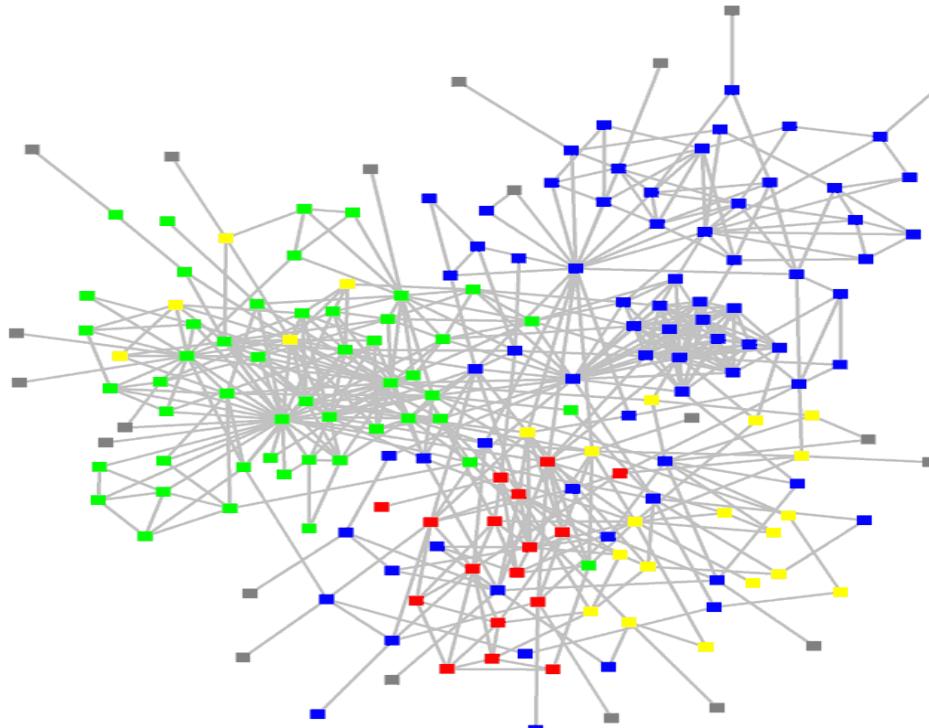
Detrás de cada sistema complejo hay una **red**, que define las interacciones entre los componentes.



El “grafo social” detrás de Facebook

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

ESTRUCTURA DE UNA ORGANIZACIÓN



■ ■ ■ : departments

■ : consultants

■ : external experts

www.orgnet.com



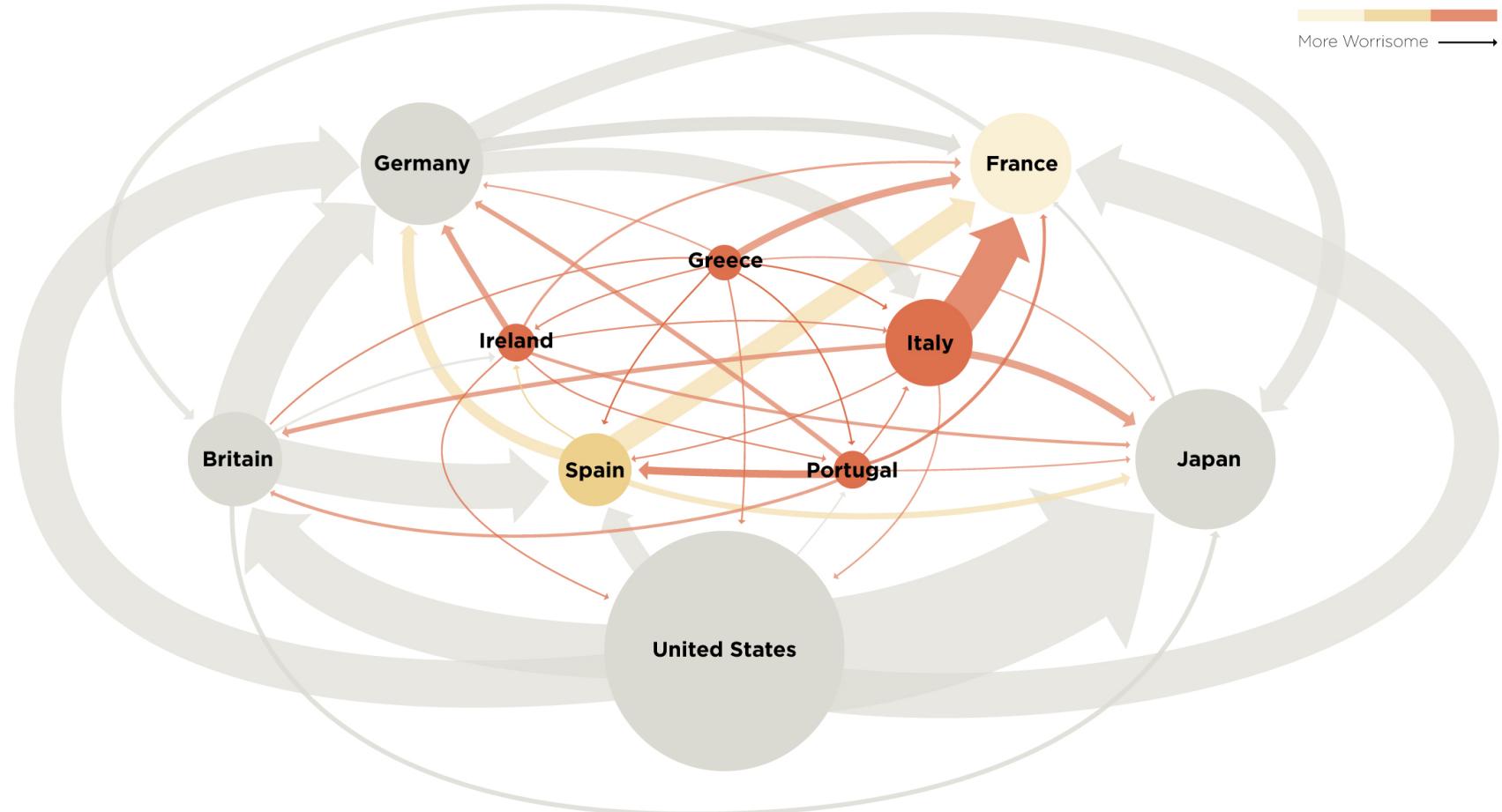
El cerebro humano tiene entre 10 y 100 billones de neuronas.

Las (sutiles) redes financieras.

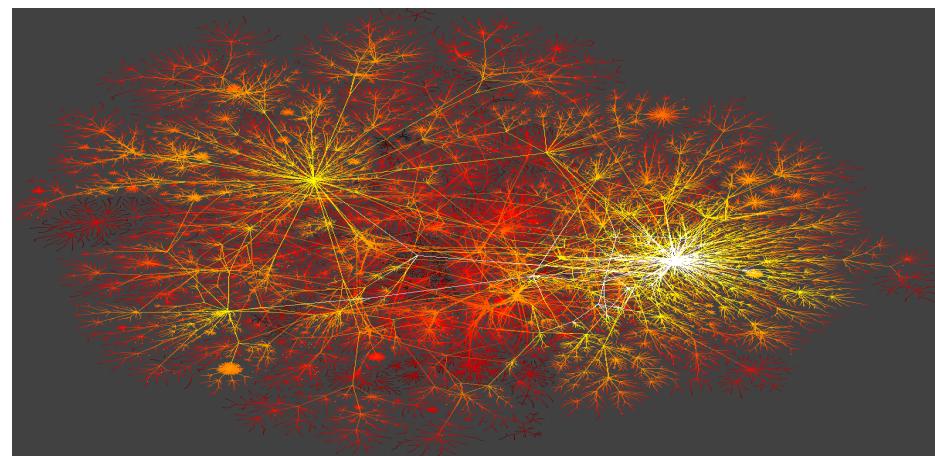
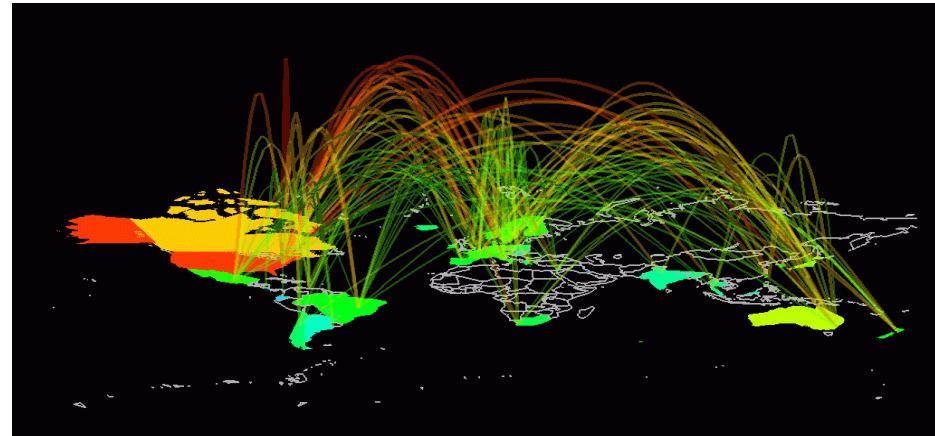
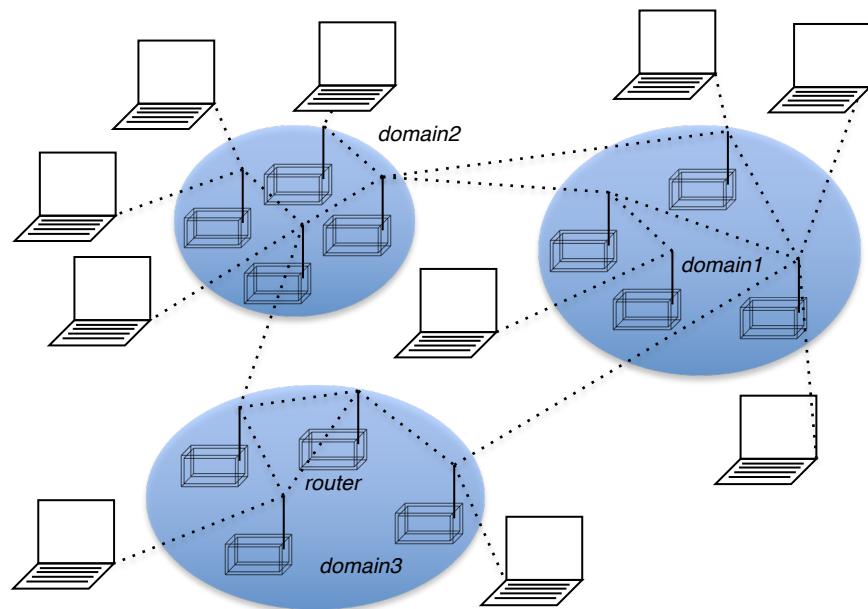


Emiratos Arabes
HSBC (Londres)
VISA (USA)
Regla transacciones Islámicas
También para países no-musulmanes

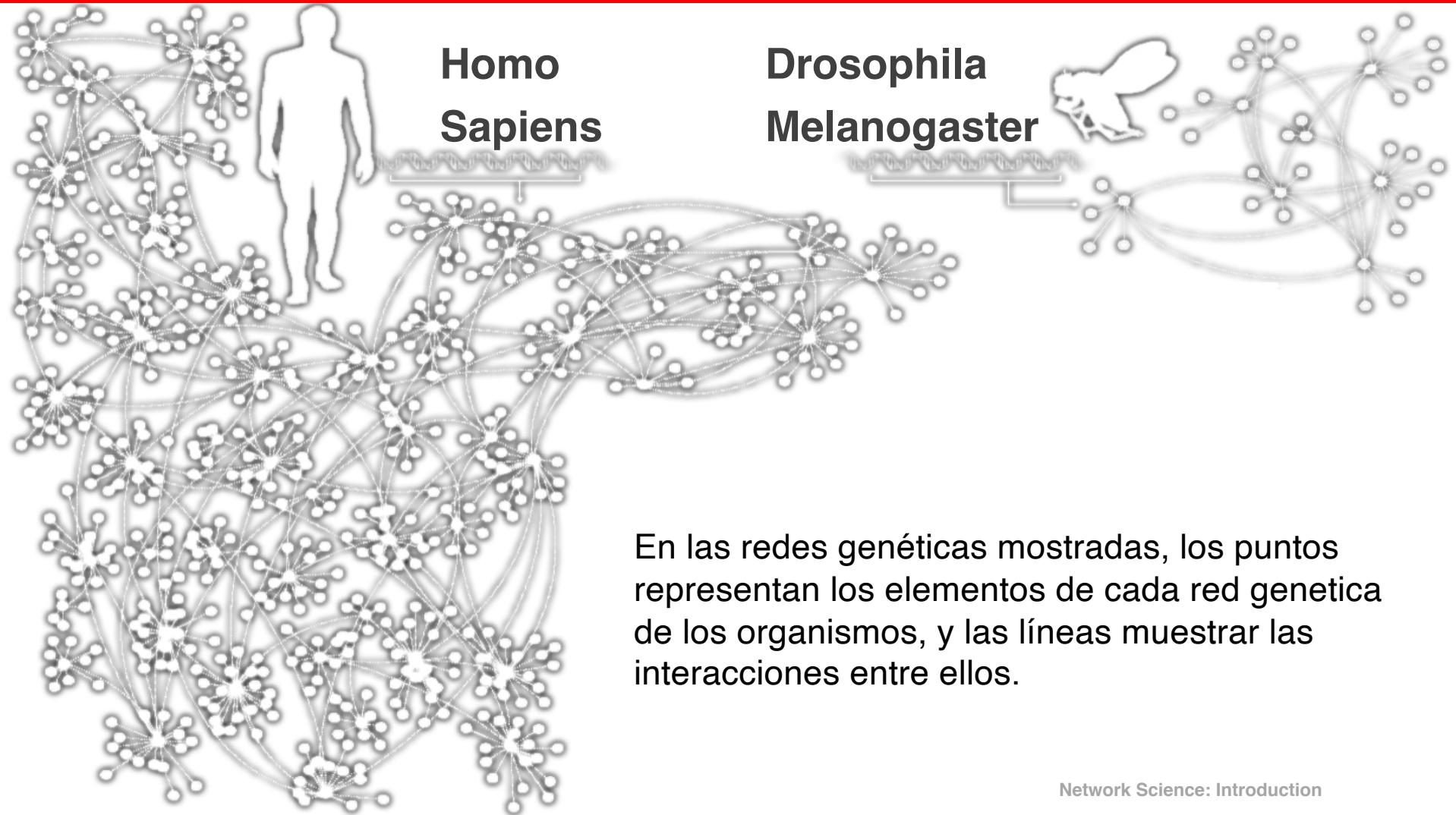
La no-muy-util red financiera: 2011



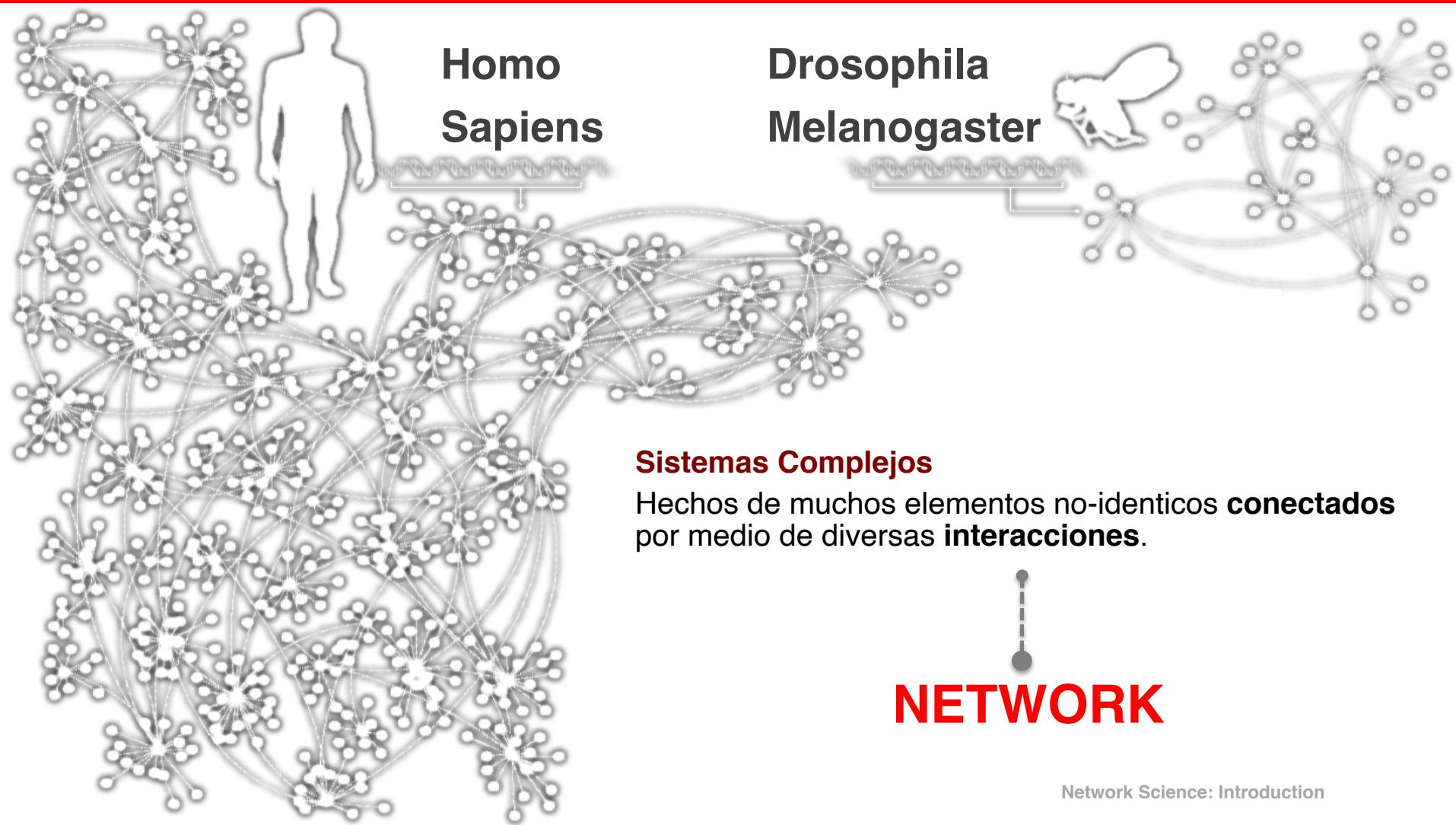
INTERNET



GENES HUMANOS



GENES HUMANOS



Sistemas Complejos

Hechos de muchos elementos no-identicos **conectados** por medio de diversas **interacciones**.



NETWORK

EL ROL DE LAS REDES

Detrás de cada sistema estudiado en complejidad, hay un intrincado diagrama de cableado, o una red, que define las interacciones entre los componentes.

Nunca entenderemos un sistema complejo a menos que mapeemos y entendamos las redes detrás de ellos.

DOS FUERZAS AYUDARON A LA EMERGENCIA DE NETWORK SCIENCE

LA HISTORIA DEL ANÁLISIS DE REDES

¿Por qué no surgió la ciencia de redes doscientos años antes?

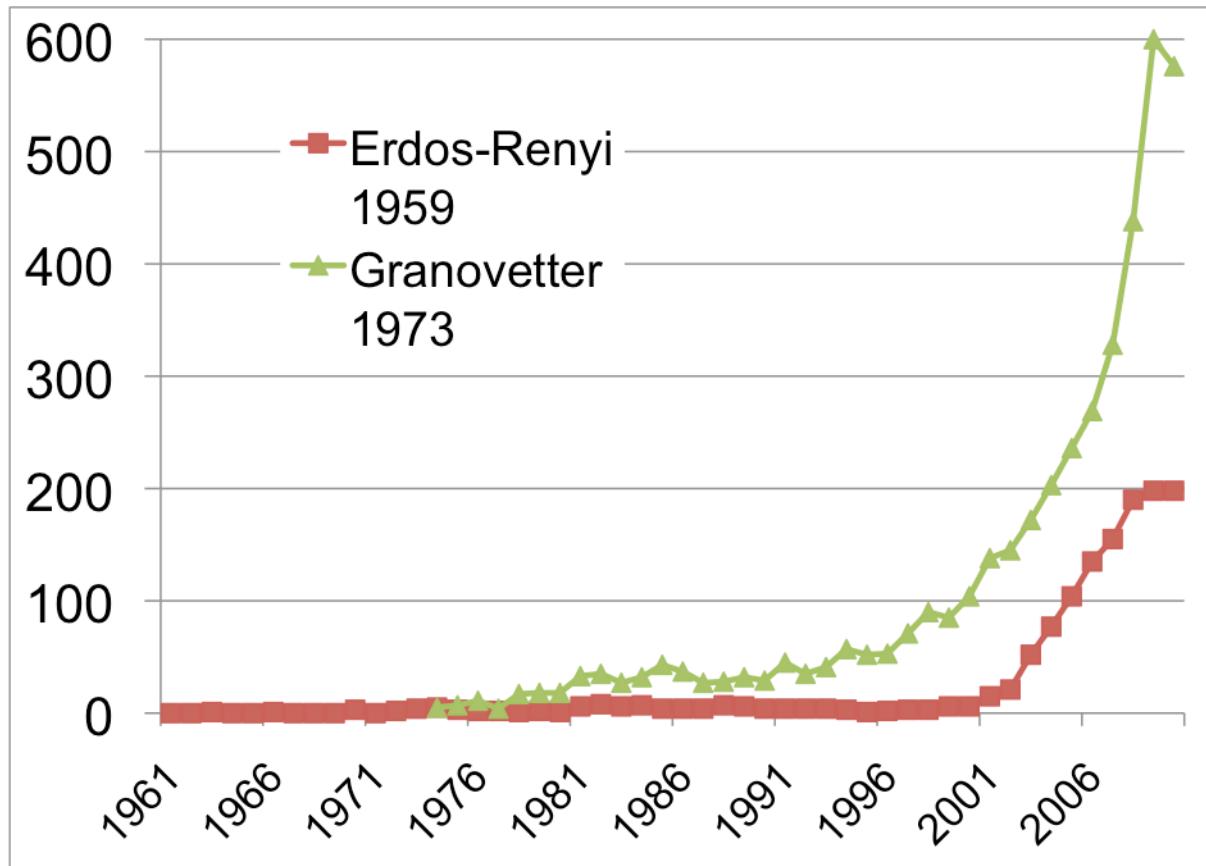
Teoría de grafos: 1735, Euler

Investigación en redes sociales: 1930s,
Moreno

Comunicación redes/internet: 1960s

Redes Ecológicas: May, 1979.

LA HISTORIA DEL ANÁLISIS DE REDES



La emergencia de los mapas de redes

Red de actores (películas), 1998;

World Wide Web, 1999.

C elegans diagrama de cableado
neural 1990

Red de Citaciones, 1998

Red metabólica, 2000;

Red de interacción proteína-proteína
(PPI), 2001

La universalidad de las características de las redes:

La arquitectura de las redes que emergen en varios dominios de la ciencia, la naturaleza y la tecnología son más similares entre sí de lo que uno hubiera esperado.

Conjunto de herramientas matemáticas común

LAS CARACTERÍSTICAS DE LA CIENCIA DE REDES

THE CHARACTERISTICS OF NETWORK SCIENCE

Interdisciplinaria

Empírica

Cuantitativa y Matemática

Computacional

THE CHARACTERISTICS OF NETWORK SCIENCE

Interdisciplinaria

Empírica, data-driven

Cuantitativa y Matemática

Computacional

THE CHARACTERISTICS OF NETWORK SCIENCE

Interdisciplinaria

Empírica

Cuantitativa y Matemática

Computacional

THE CHARACTERISTICS OF NETWORK SCIENCE

Interdisciplinaria

Empírica

Cuantitativa y Matemática

Computacional

EL IMPACTO DE LA CIENCIA DE REDES

IMPACTO ECONÓMICO



Google

Market Cap(Jan 1, 2010):
\$189 billion

Cisco Systems

networking gear Market
cap (Jan 1, 2010):
\$112 billion

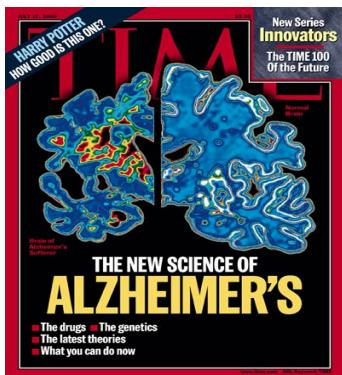
Facebook

market cap:
\$50 billion

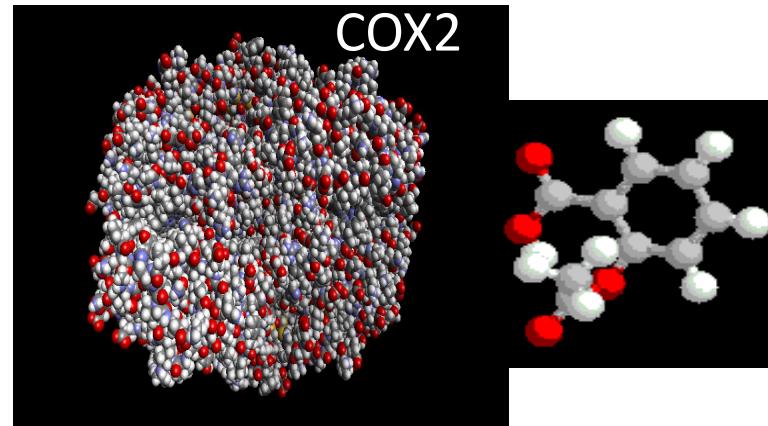
www.bizjournals.com/austin/news/2010/11/15/facebook... - Cached

DISEÑO DE DROGAS E INGENIERÍA METABÓLICA

Reduce
Inflamación
Fiebre
Dolor

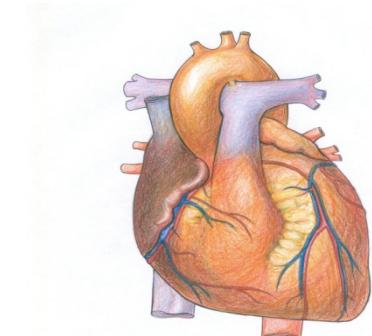


Reduce el riesgo de
Alzheimer



Reduce el riesgo de
Cancer de mama
Cancer de Ovario
Cancer colorectal

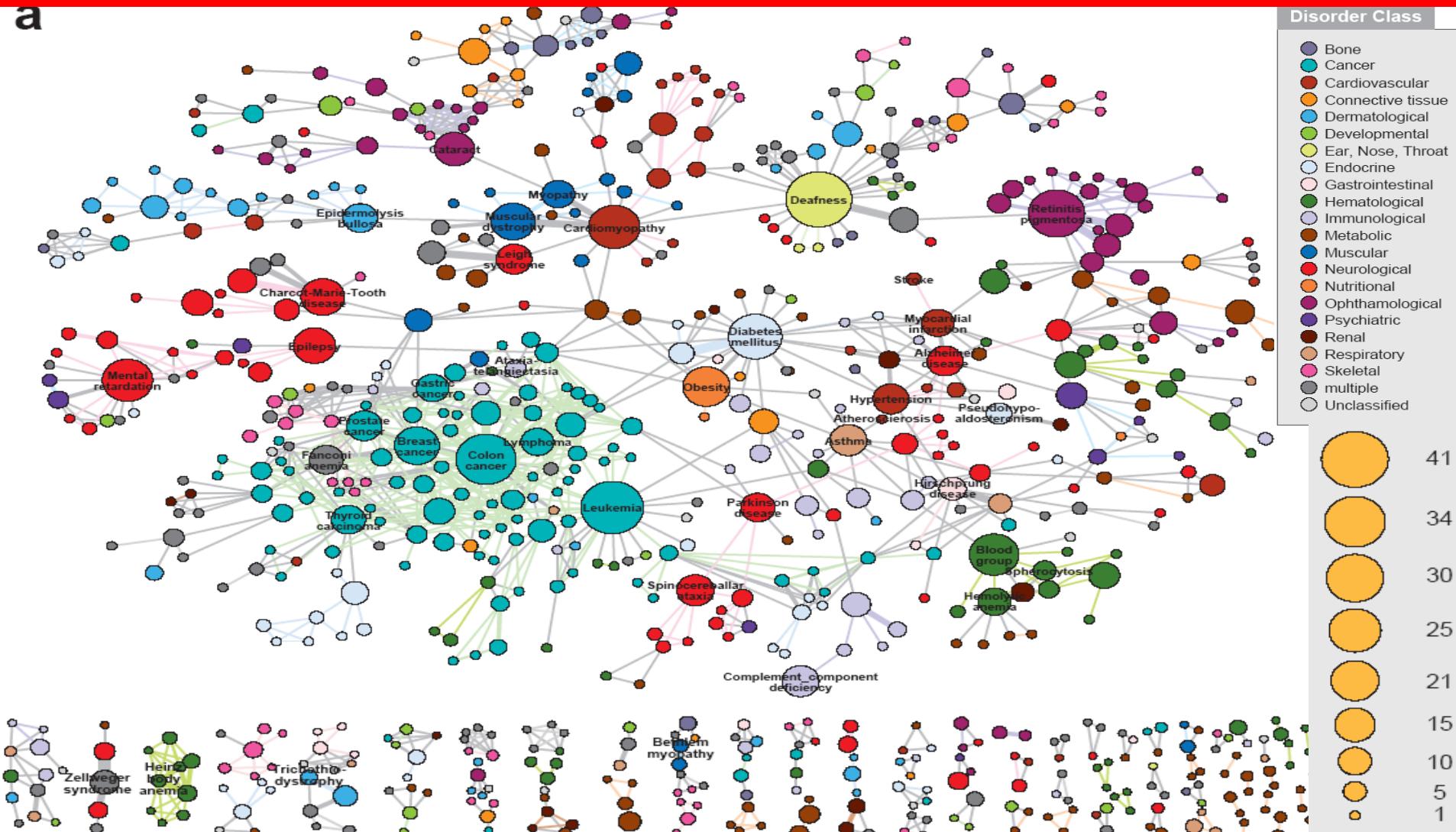
Previene
Ataques al
corazón



Causes
Ulcera

RED DE ENFERMEDADES HUMANAS

a



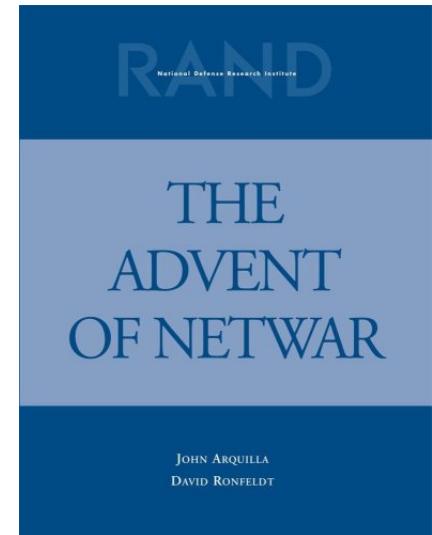
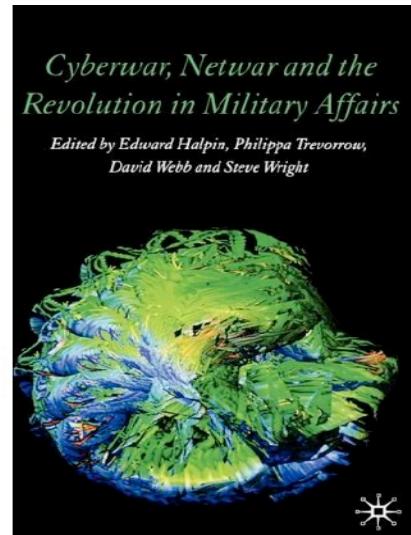
Network Biology/Network Medicine



LUCHA CONTRA EL TERRORISMO.

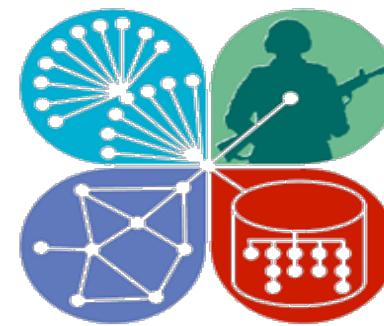


<http://www.slate.com/id/2245232>



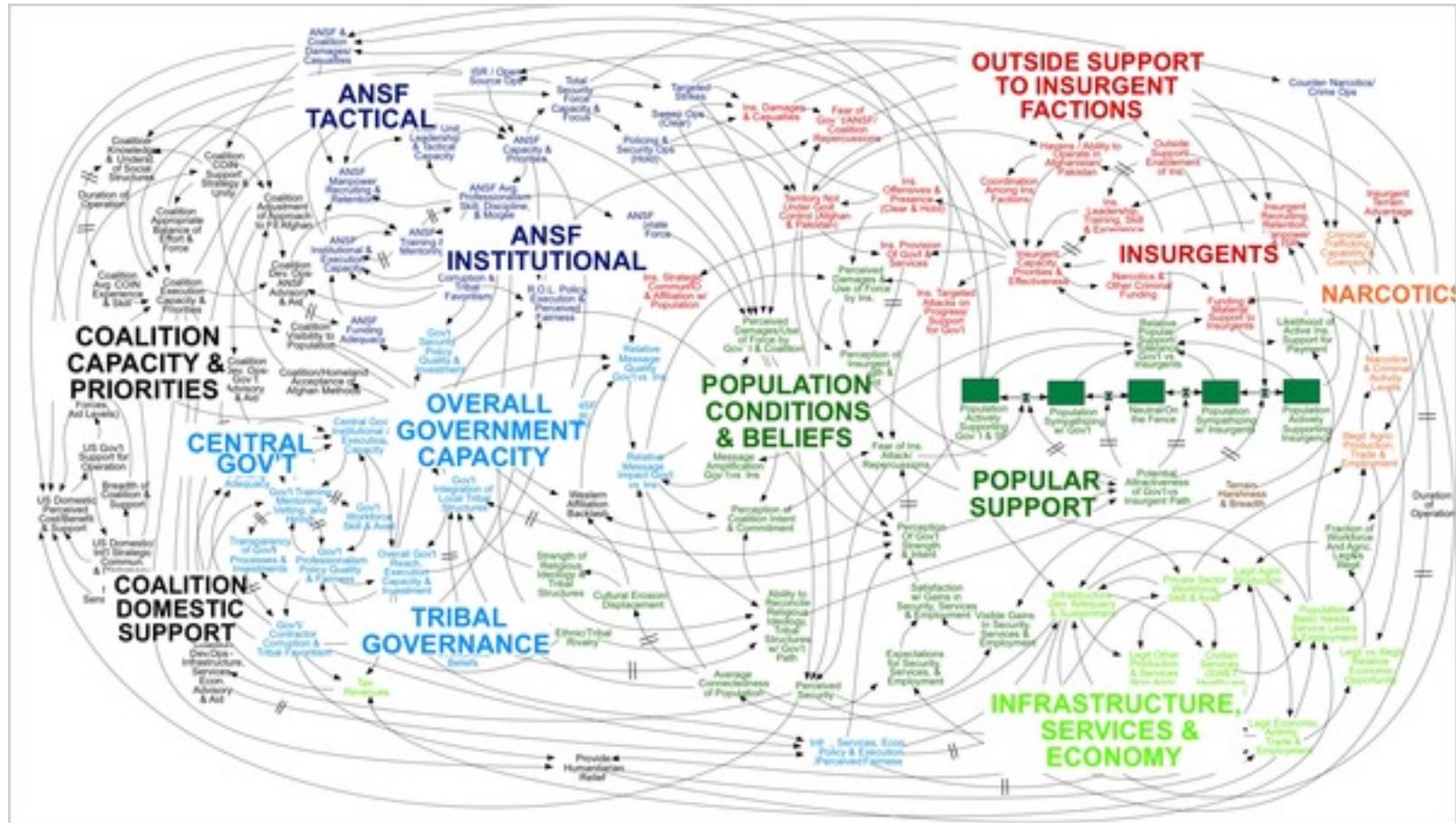
LUCHA CONTRA EL TERRORISMO.

Network Science Center
West Point 



<http://www.ns-cta.org/ns-cta-blog/>

La red detrás de un combate militar.



Predicting the H1N1 pandemic

Feb 18 2009



Chicago
New York
Los Angeles
Houston
Toronto
Vancouver
Calgary
Indianapolis

La Gloria
Sao Paulo
Mexico City
Rio De Janeiro
San Juan
Bogota

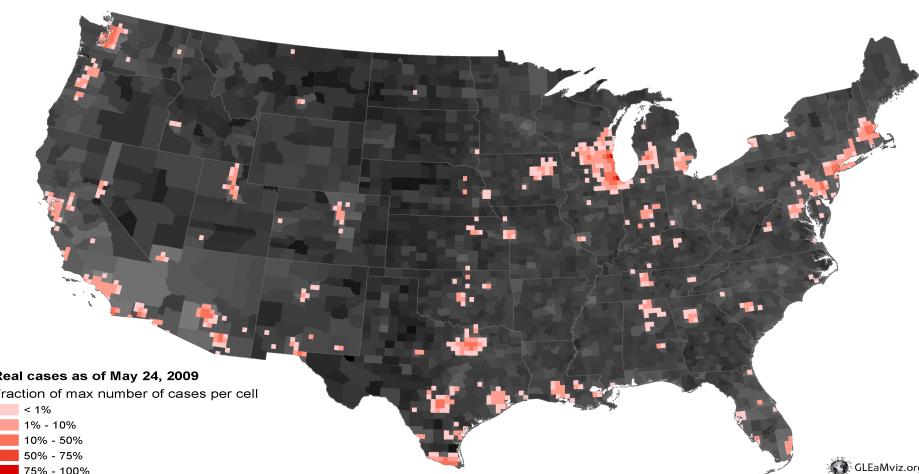
Johannesburg
Cairo
Cape Town
Nairobi

Paris
Frankfurt
Amsterdam
Rome
Milan
Moscow
Dublin

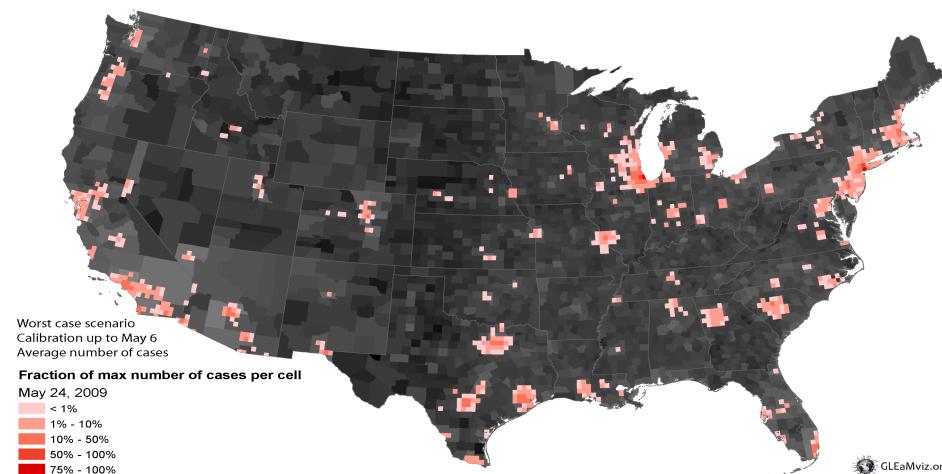
Hong Kong
Tokyo Narita
Bangkok
Singapore
Beijing
Manila

Sydney
Brisbane
Auckland
Perth

Real



Projected

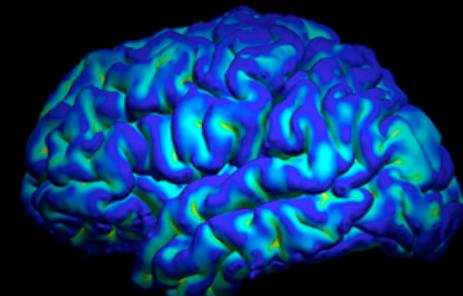
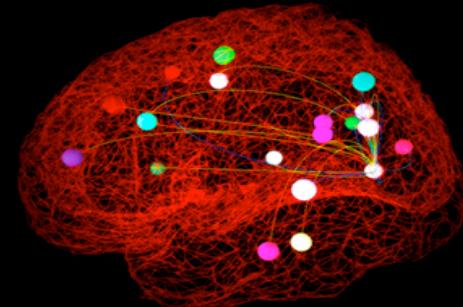


INVESTIGACIÓN CEREBRAL (NEUROCIENCIA)

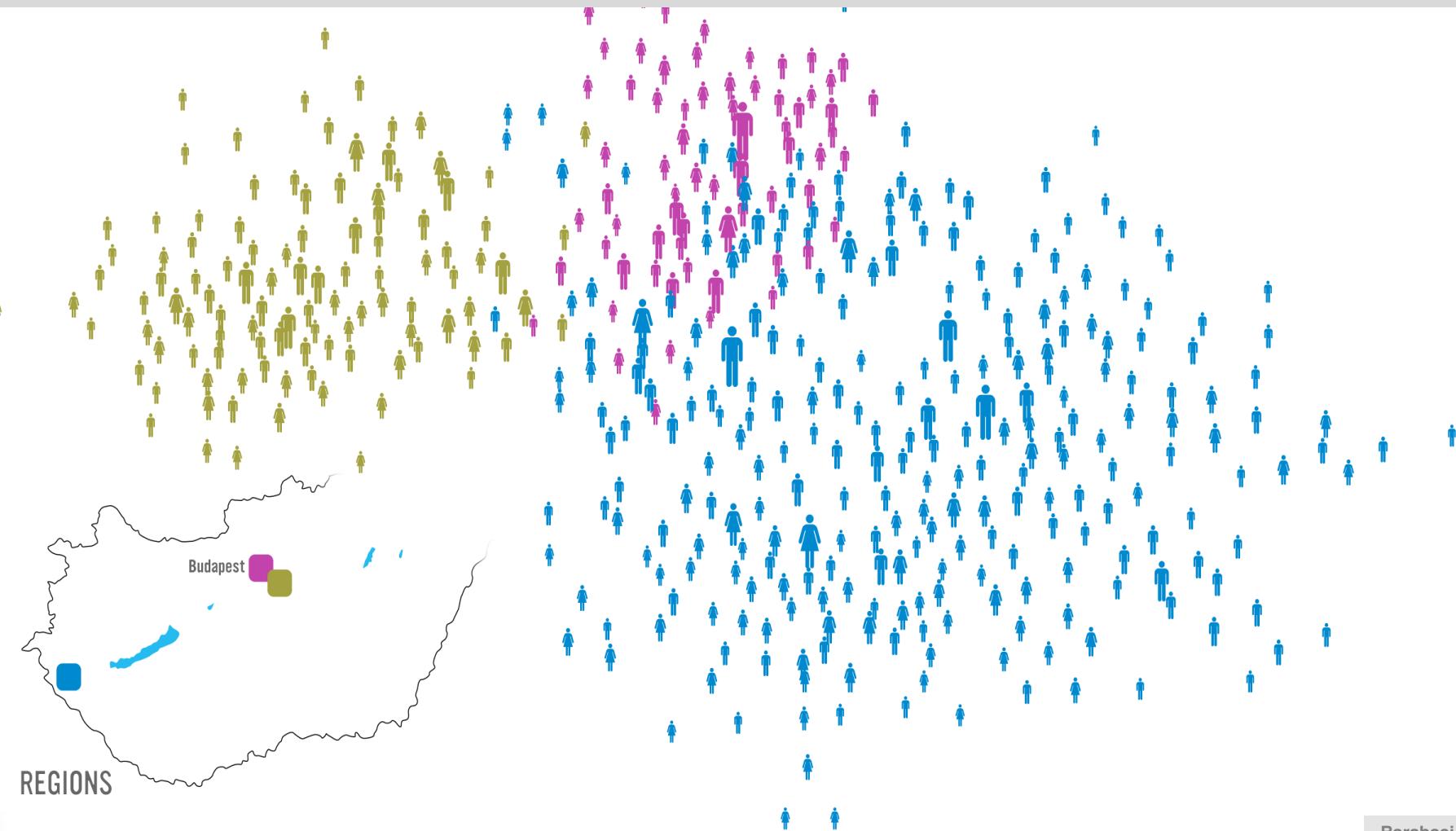
En septiembre de 2010, los Institutos Nacionales de Salud otorgaron \$ 40 millones a investigadores de Harvard, la Universidad de Washington en St. Louis, la Universidad de Minnesota y UCLA, para desarrollar las tecnologías que podrían mapear sistemáticamente los circuitos cerebrales.

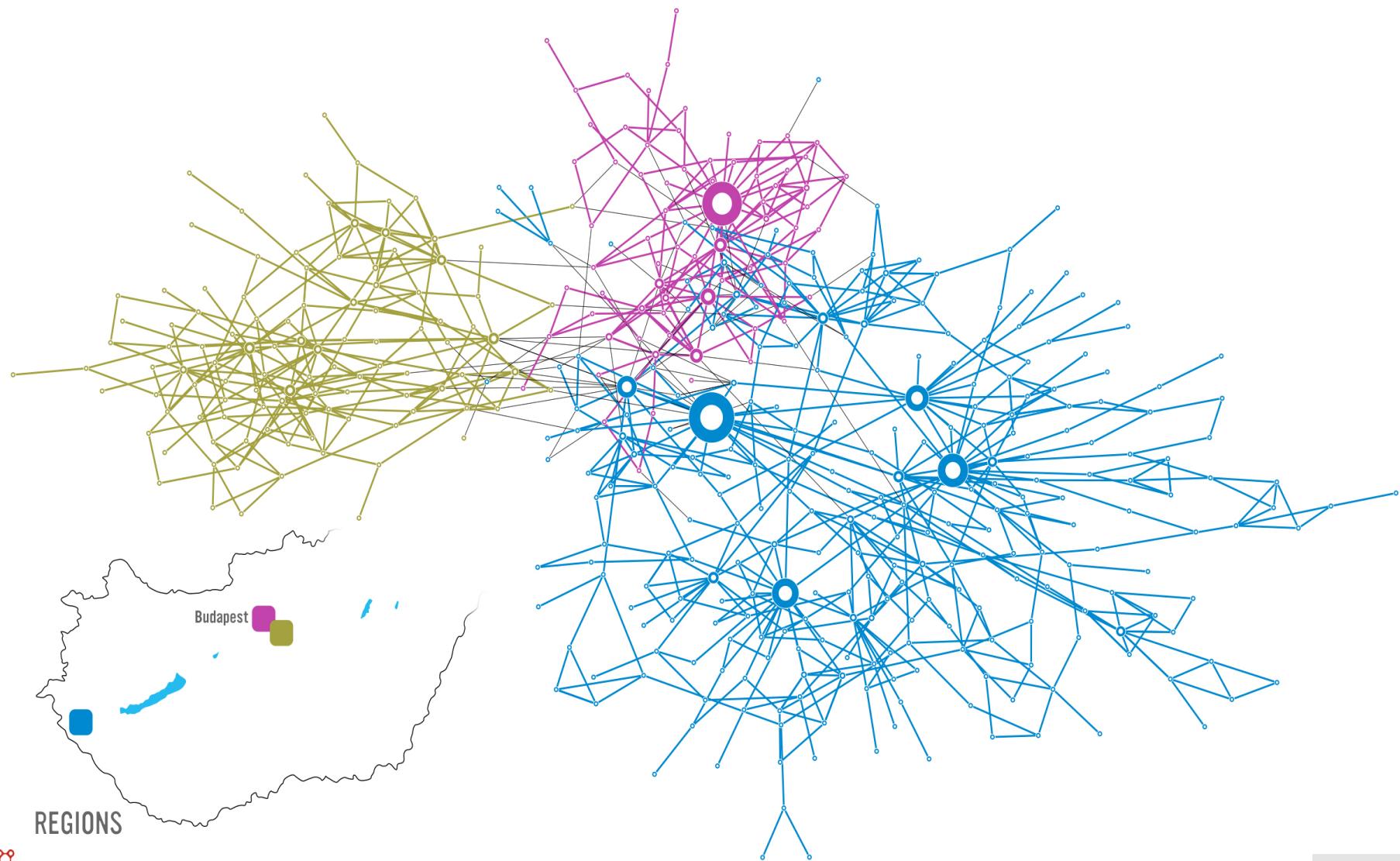
The Human Connectome Project (HCP) con el ambicioso objetivo de construir un mapa de las conexiones neuronales funcionales y estructurales completas *in vivo* dentro y entre los individuos.

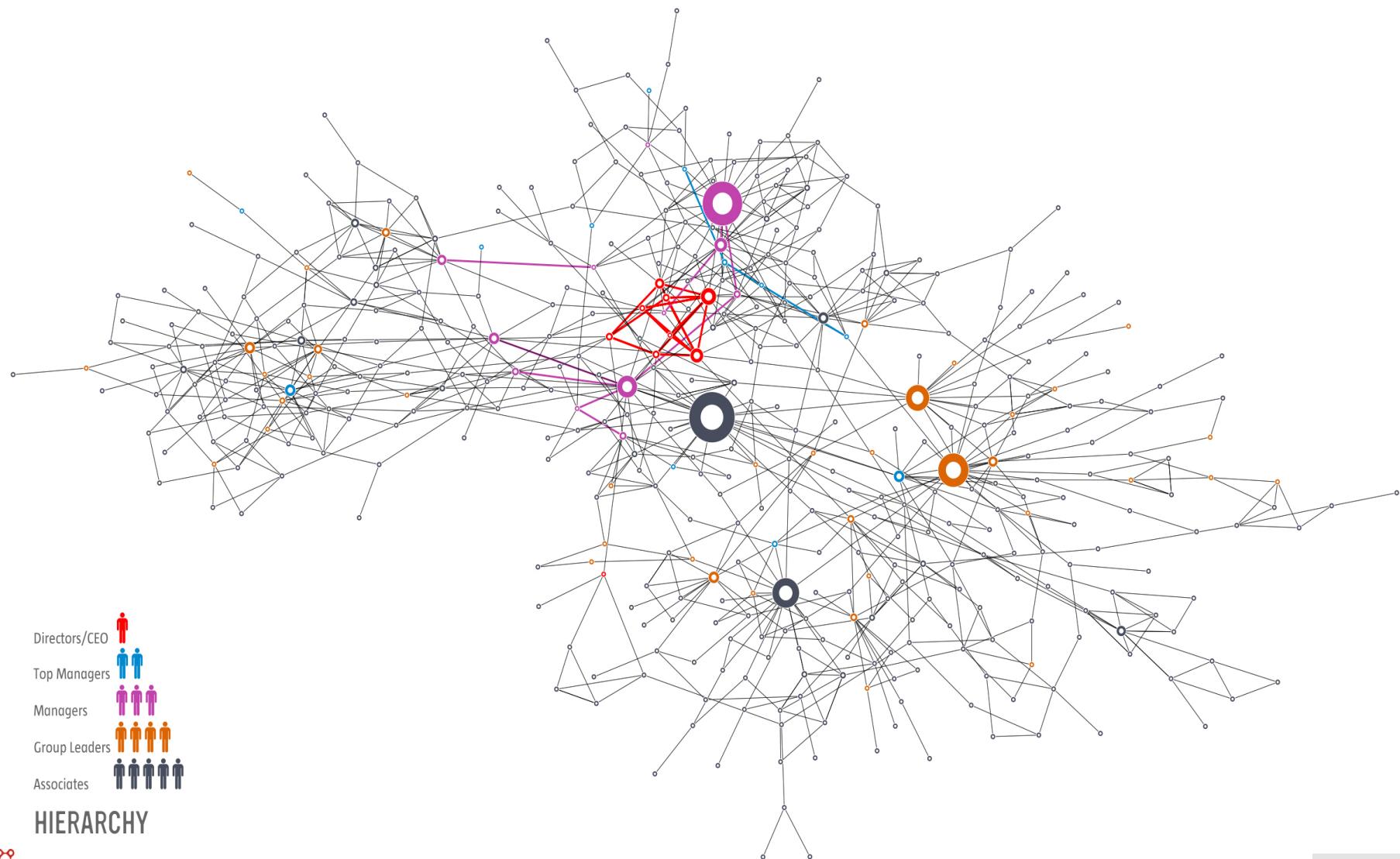
<http://www.humanconnectomeproject.org/overview/>

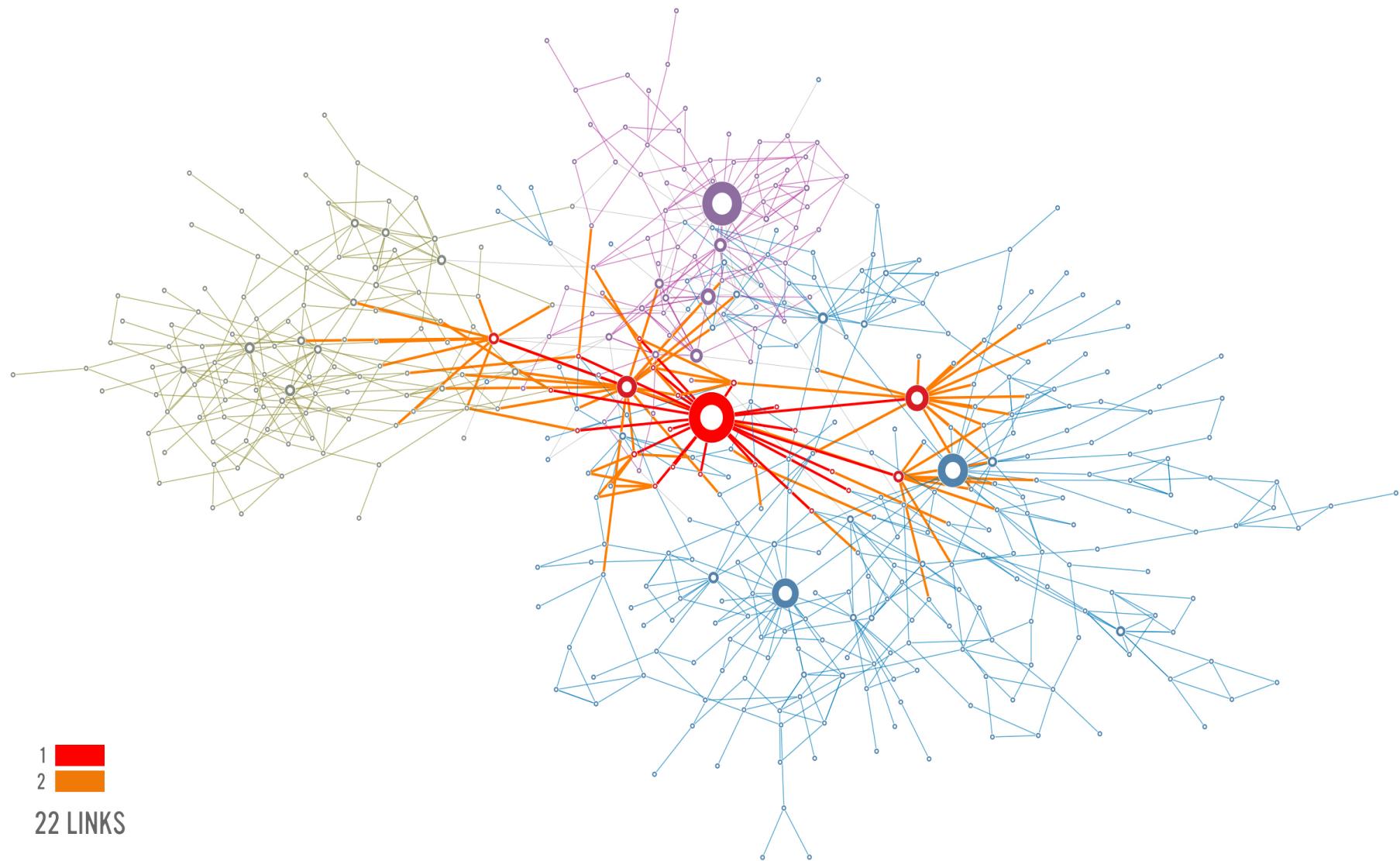


Management

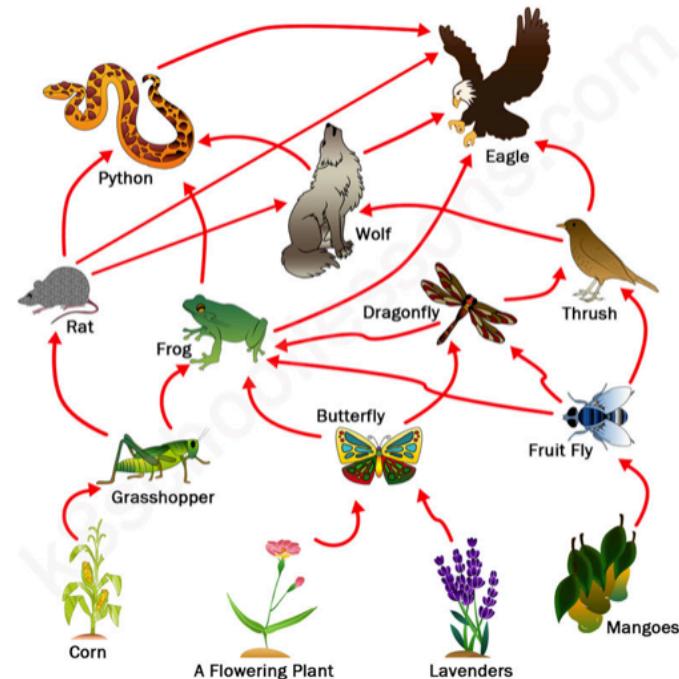






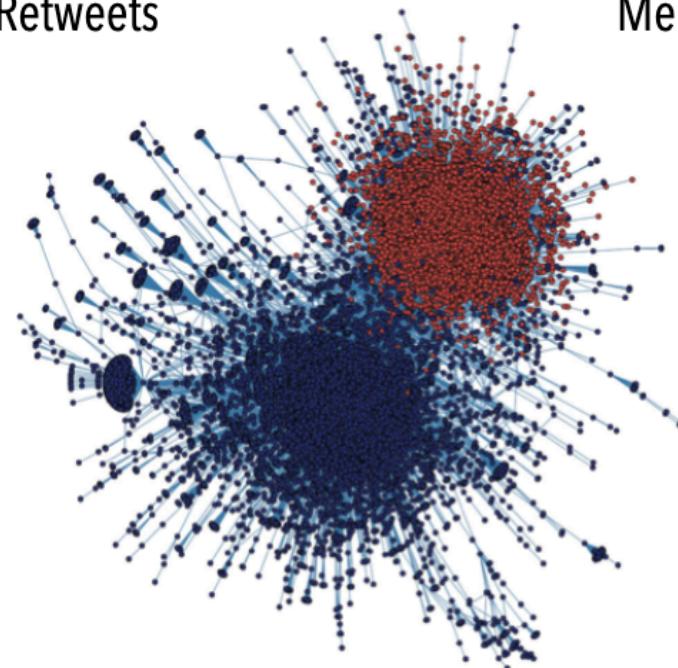


Ecological Networks

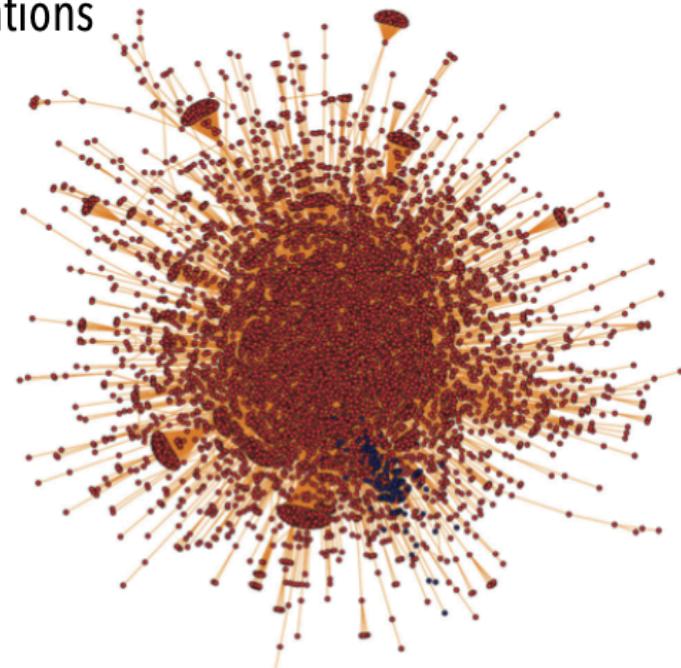


Social Networks - Twitter

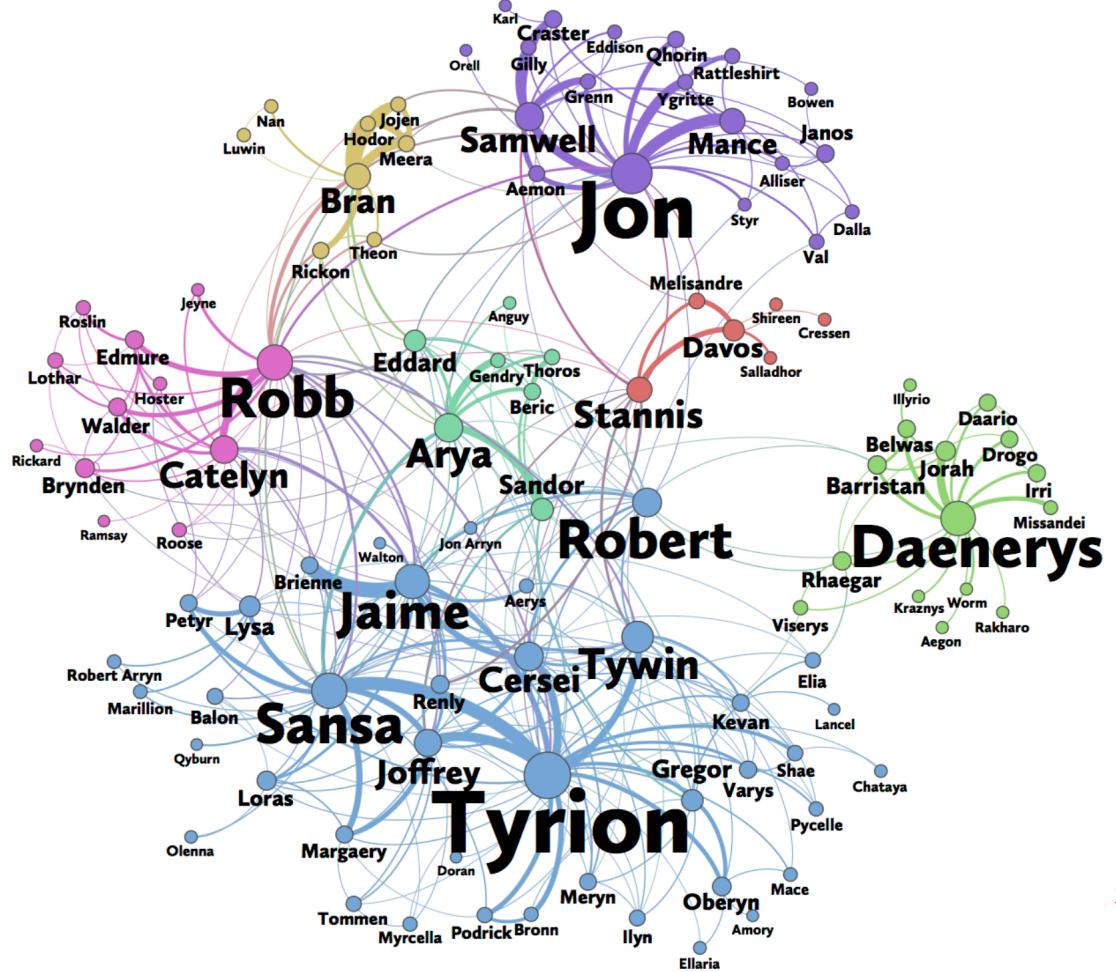
Retweets



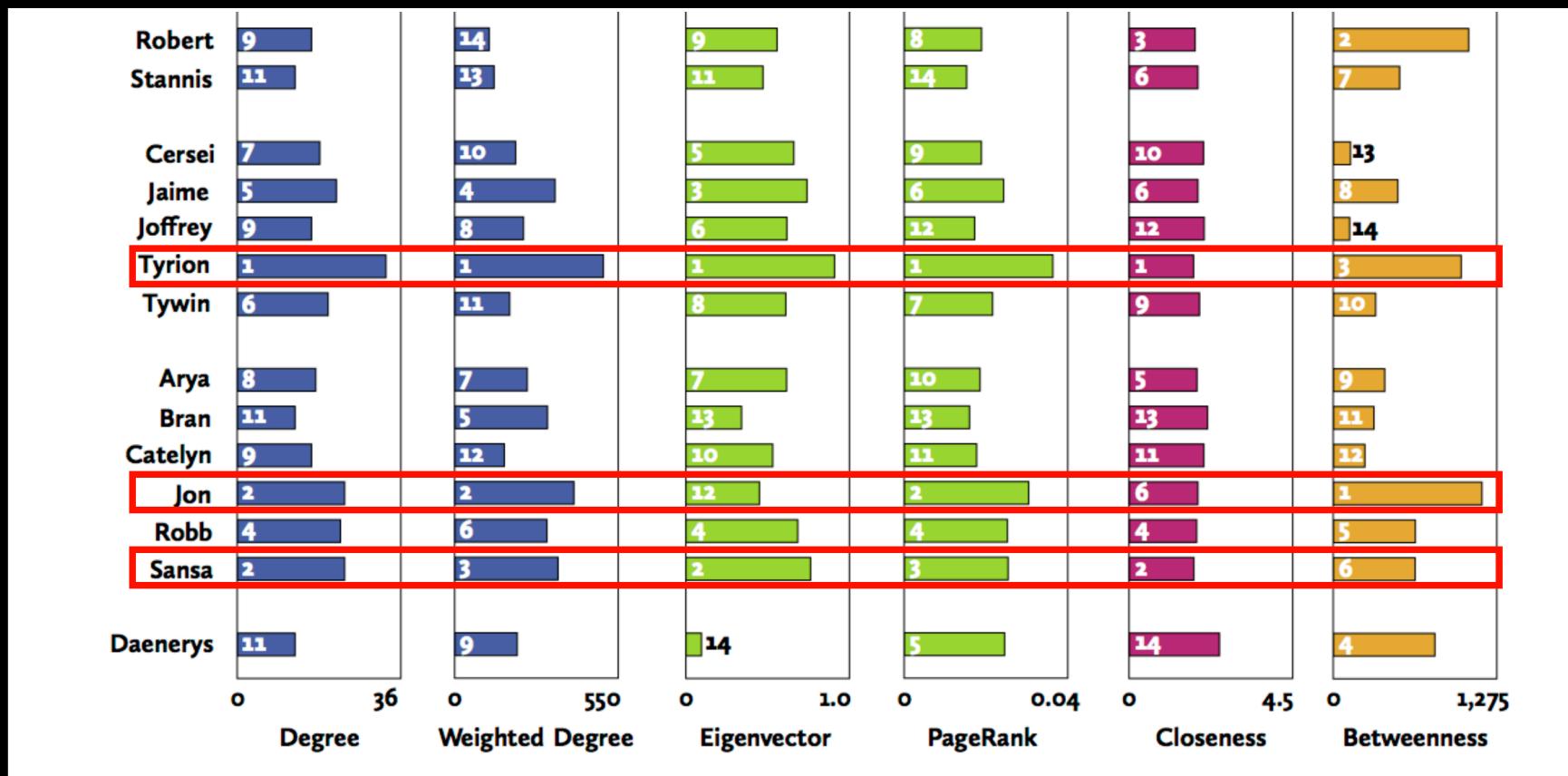
Mentions



● Republicans ● Democrats



Colours highlight the identified Network Communities



¿Quién es el personaje más importante?

Bipartite Networks

(a) RECIPES



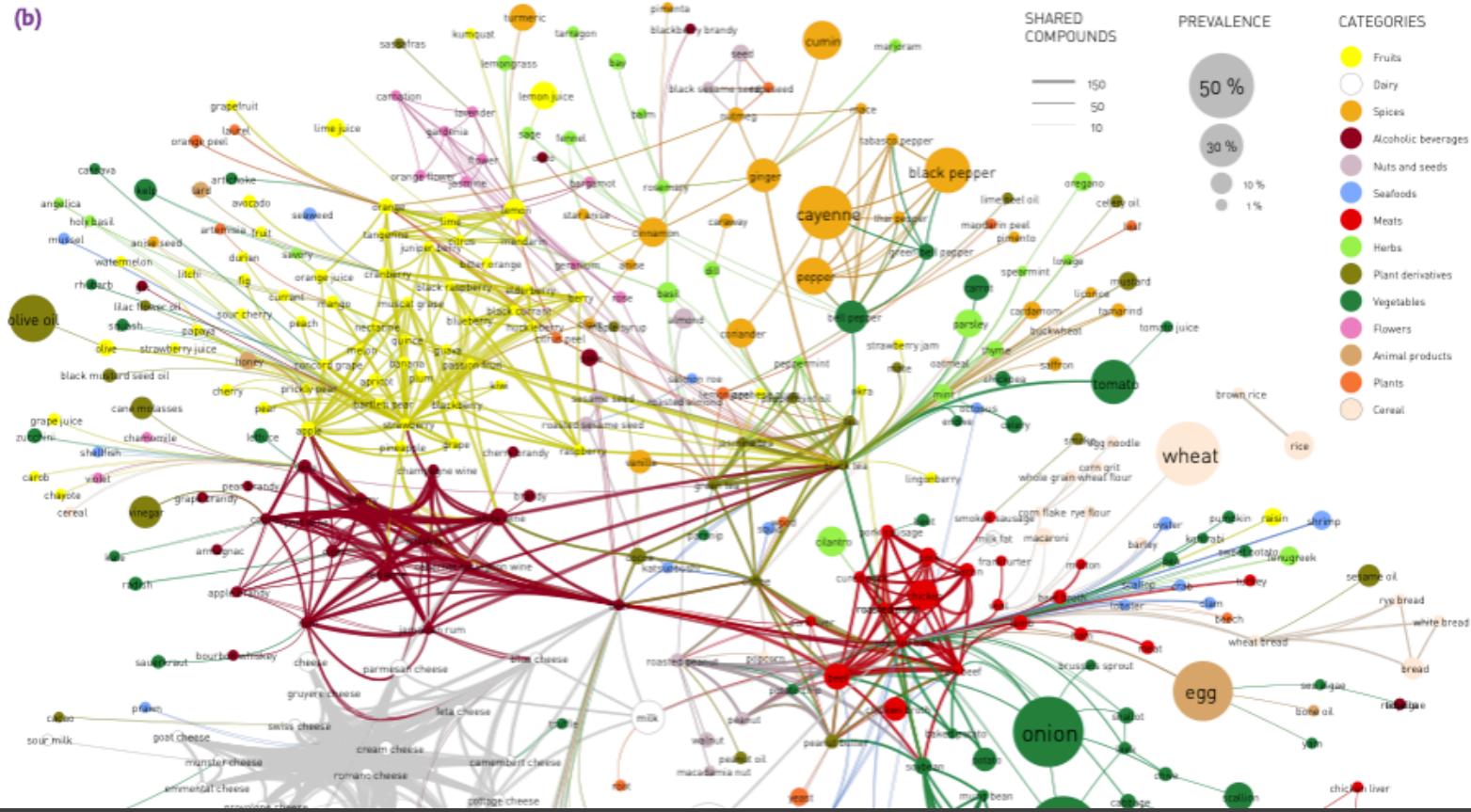
INGREDIENTS

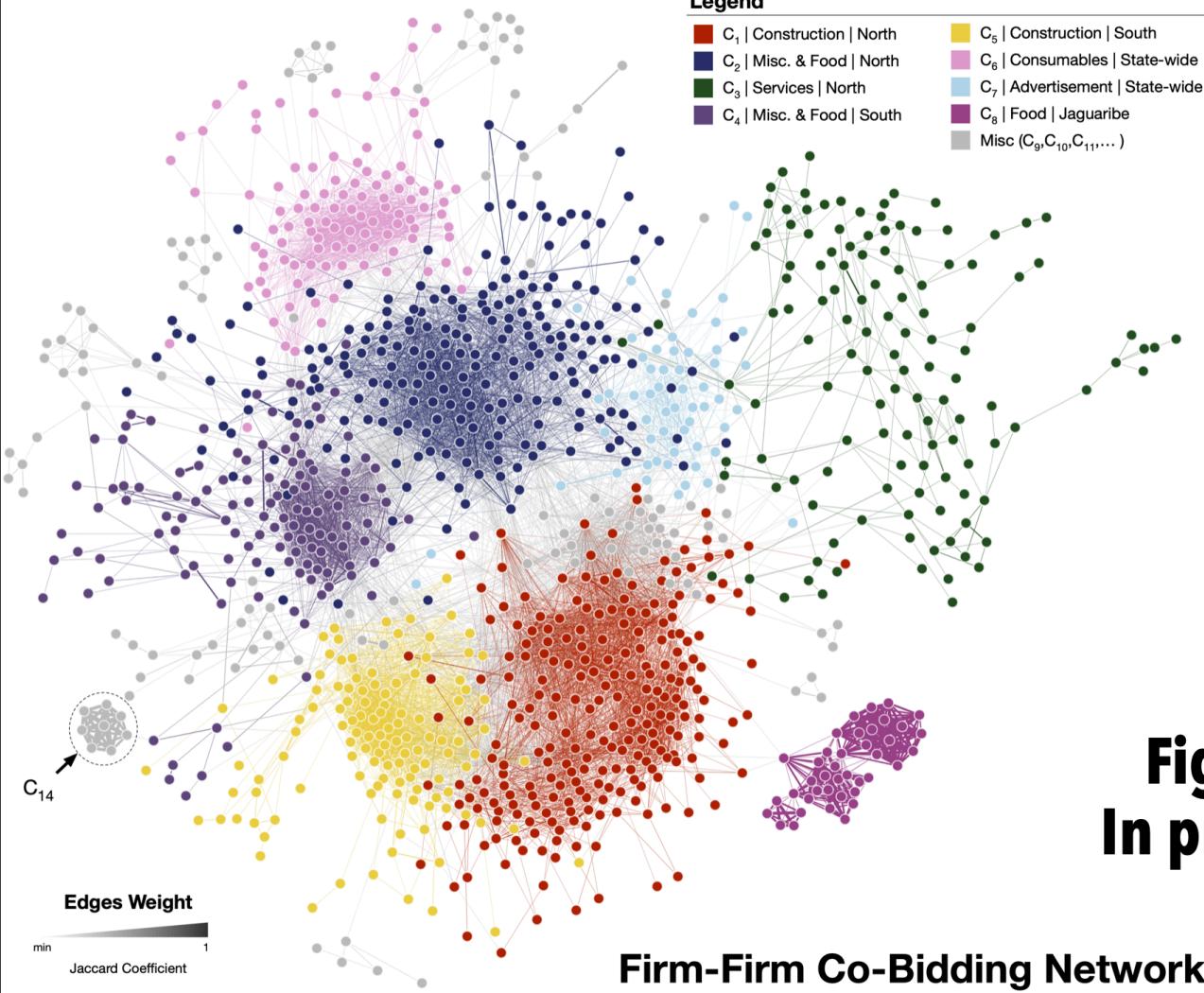
COMPOUNDS

PHENETHYL ALCOHOL
L-ASPARTIC ACID
BUTYRALDEHYDE
9-DECANOIC ACID
M-CRESOL
HYDROGEN SULFIDE
DELTA-TETRACALACTONE
ACETOIN
O-CRESOL
3-METHYL-2-BUTANOL
DECANOIC ACID
PYRROLIDINE
STYRENE
PROPENYL
PROPYL DISULFIDE
GERANIOL

Bipartite Networks

(b)

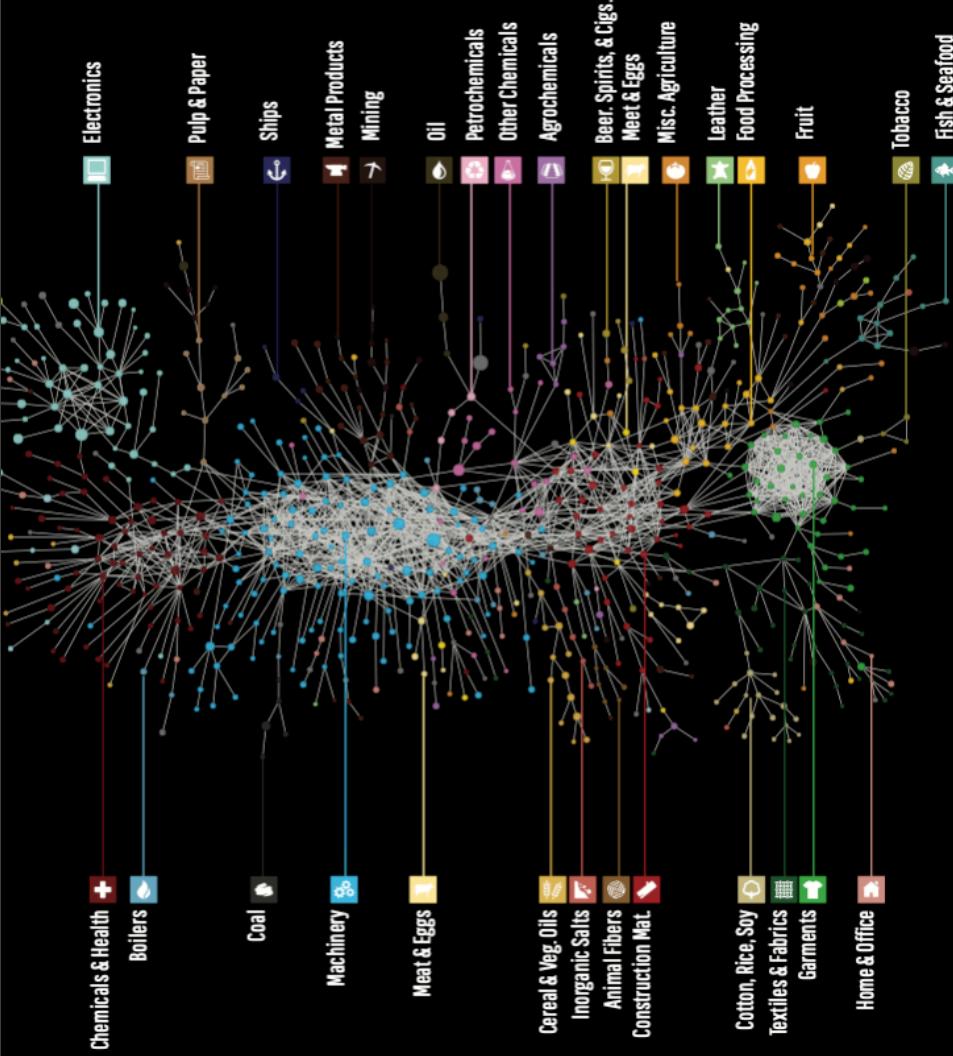




**Fighting Corruption
In procurement activity**

Product Space

Relates products
that countries export



Hidalgo, César A., et al.
"The product space conditions the development of nations."
Science 317.5837 (2007): 482-487.

The Higher Education Space

Lista de Preferencias

1. Medicina
2. Odontología
3. Tecnología Médica
4. Odontología
5. Cs. Físicas y Astronómicas

Consideramos todas las preferencias de cada postulantes.

Las preferencias repetidas indican una postulación a dos instituciones de educación superior distintas.



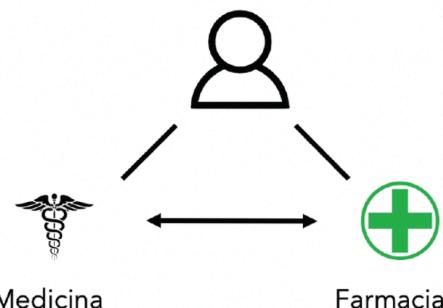
Pares de Carreras

Medicina	●—●	Odontología
Medicina	●—●	Tecnología Médica
Medicina	●—●	Odonotología
Medicina	●—●	Cs. Físicas y Astronómicas
Odontología	●—●	Tecnología Médica

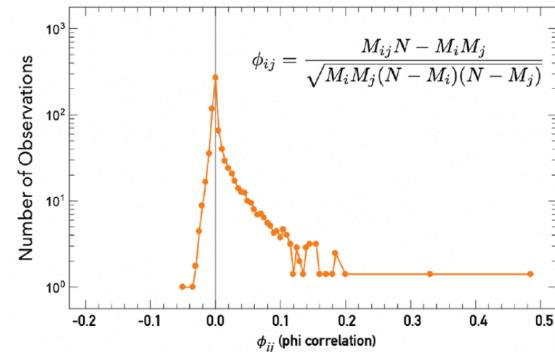
Odontología	●—●	Odontología
Odontología	●—●	Cs. Físicas y Astronómicas
Tecnología Médica	●—●	Odontología
Tecnología Médica	●—●	Cs. Físicas y Astronómicas
Odontología	●—●	Cs. Físicas y Astronómicas

Creamos todos los pares de carreras posibles. Luego, descartamos los que contienen la misma carrera en ambos extremos (color gris).

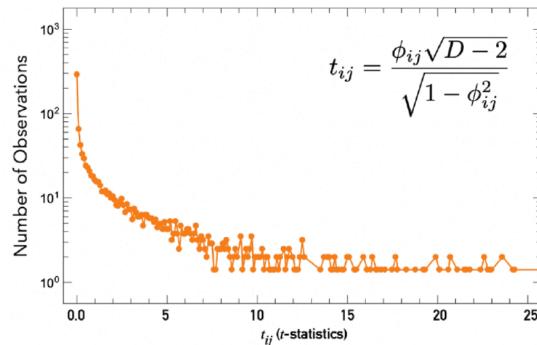
Conectamos carreras que se co-postulan



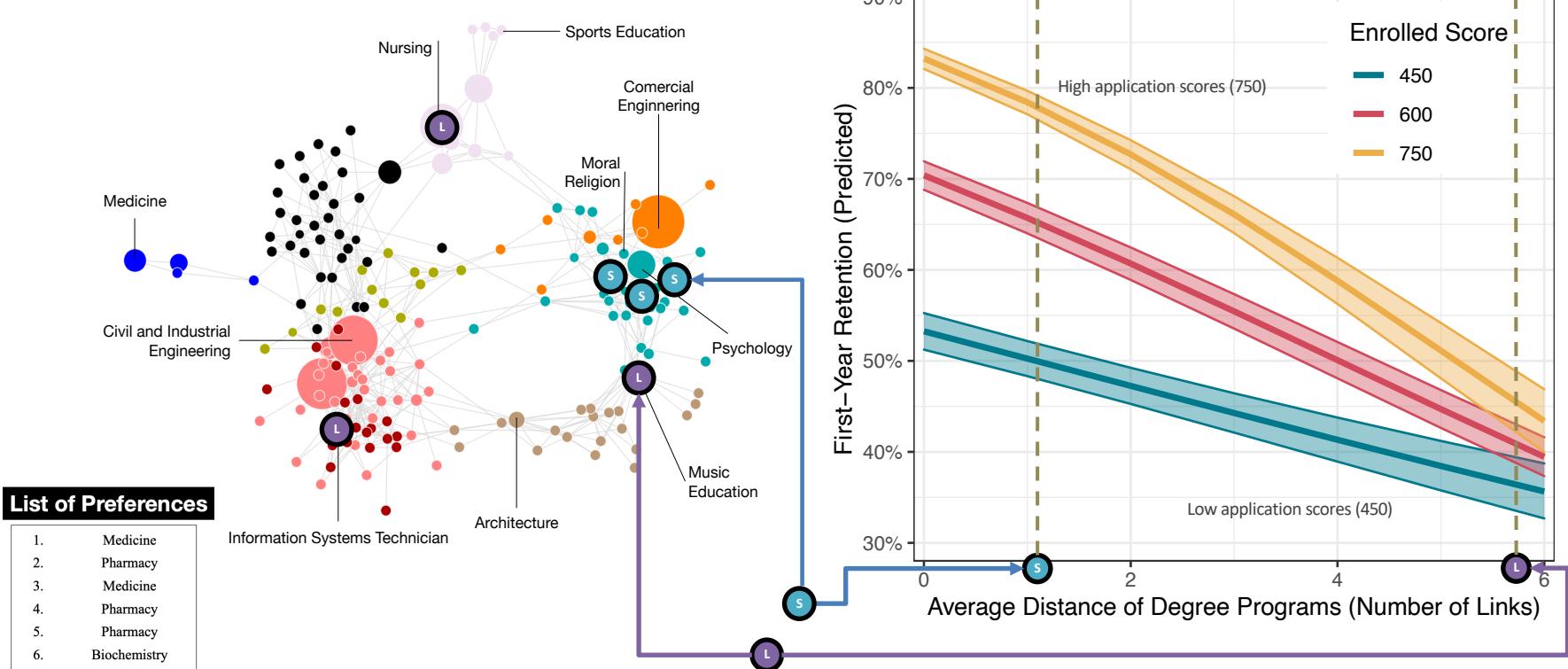
Calculamos la correlación entre carreras

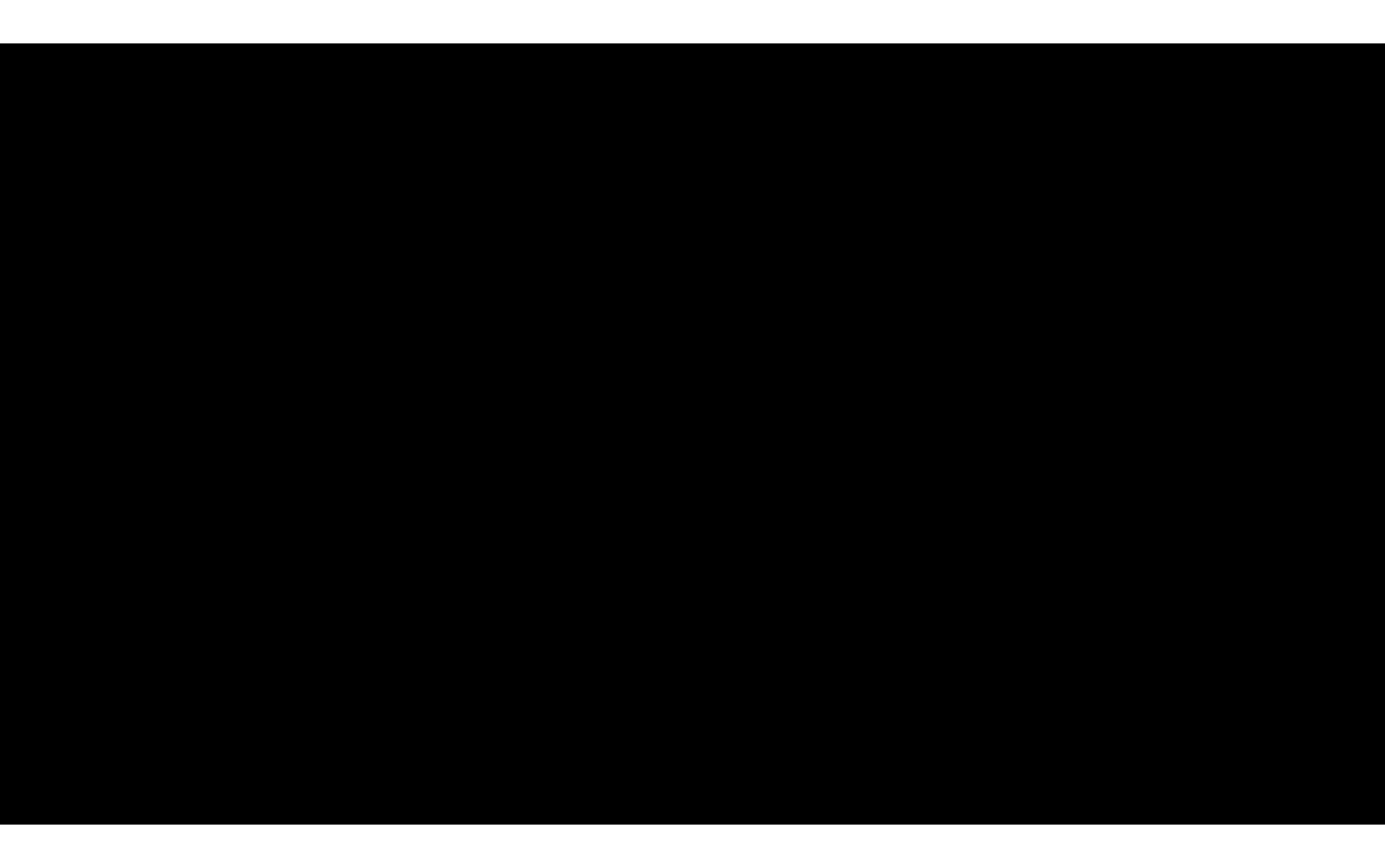


Calculamos la significancia de esas correlaciones



Applicant's inconsistent preferences for degree programs reduce their first-year retention in higher education.





Modified and Non-Anonymous Prisoner's Dilemma for Mapping Cooperative Networks in Elementary Education

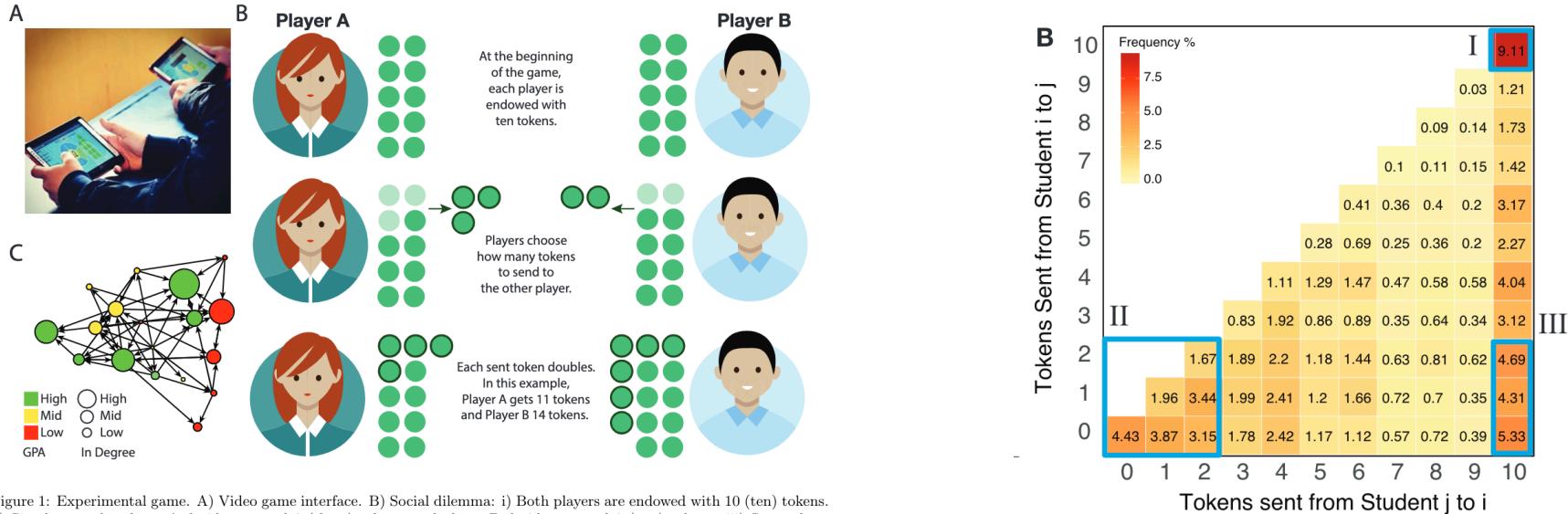
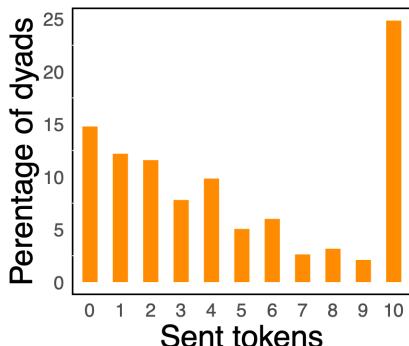


Figure 1: Experimental game. A) Video game interface. B) Social dilemma: i) Both players are endowed with 10 (ten) tokens. ii) Simultaneously, player A decides to send 3 (three) tokens, and player B decides to send 2 (two) tokens. iii) Sent tokens are doubled. iv) Player A receives 4 (four) tokens and player B receives 6 (six) tokens. C) Students' network, edges represent fully cooperative interactions.

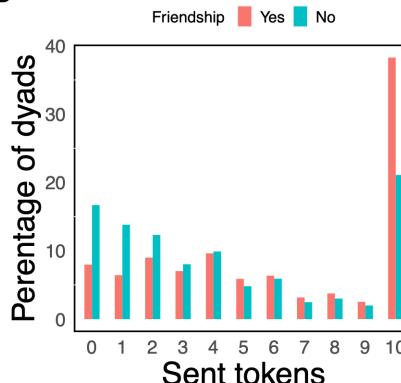
+1400 children between 8 and 10 years old

High GPA children tend to be less cooperative with their non-friends.

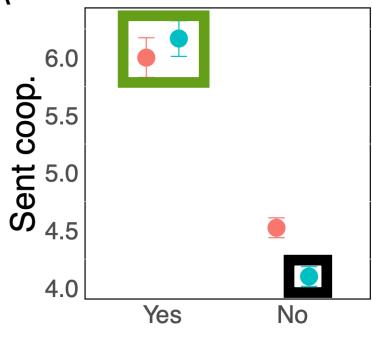
A



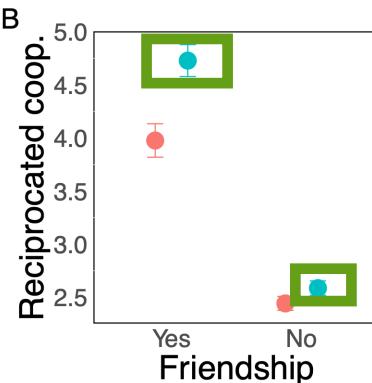
B



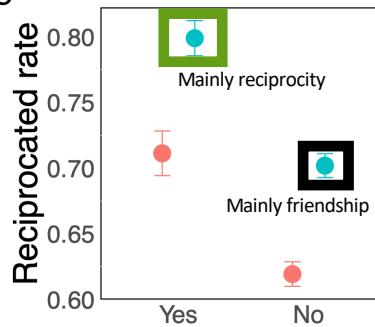
A



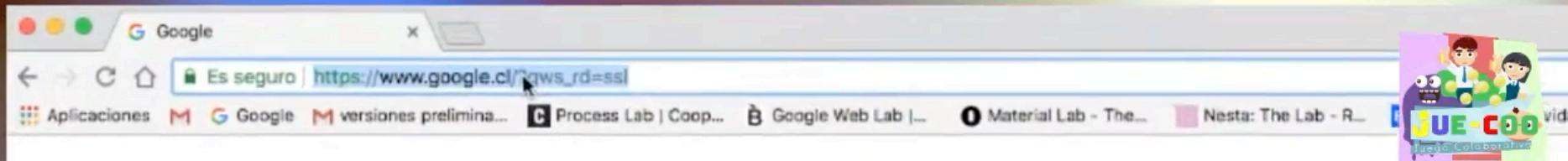
B



C



GPA ego ● Low ● High



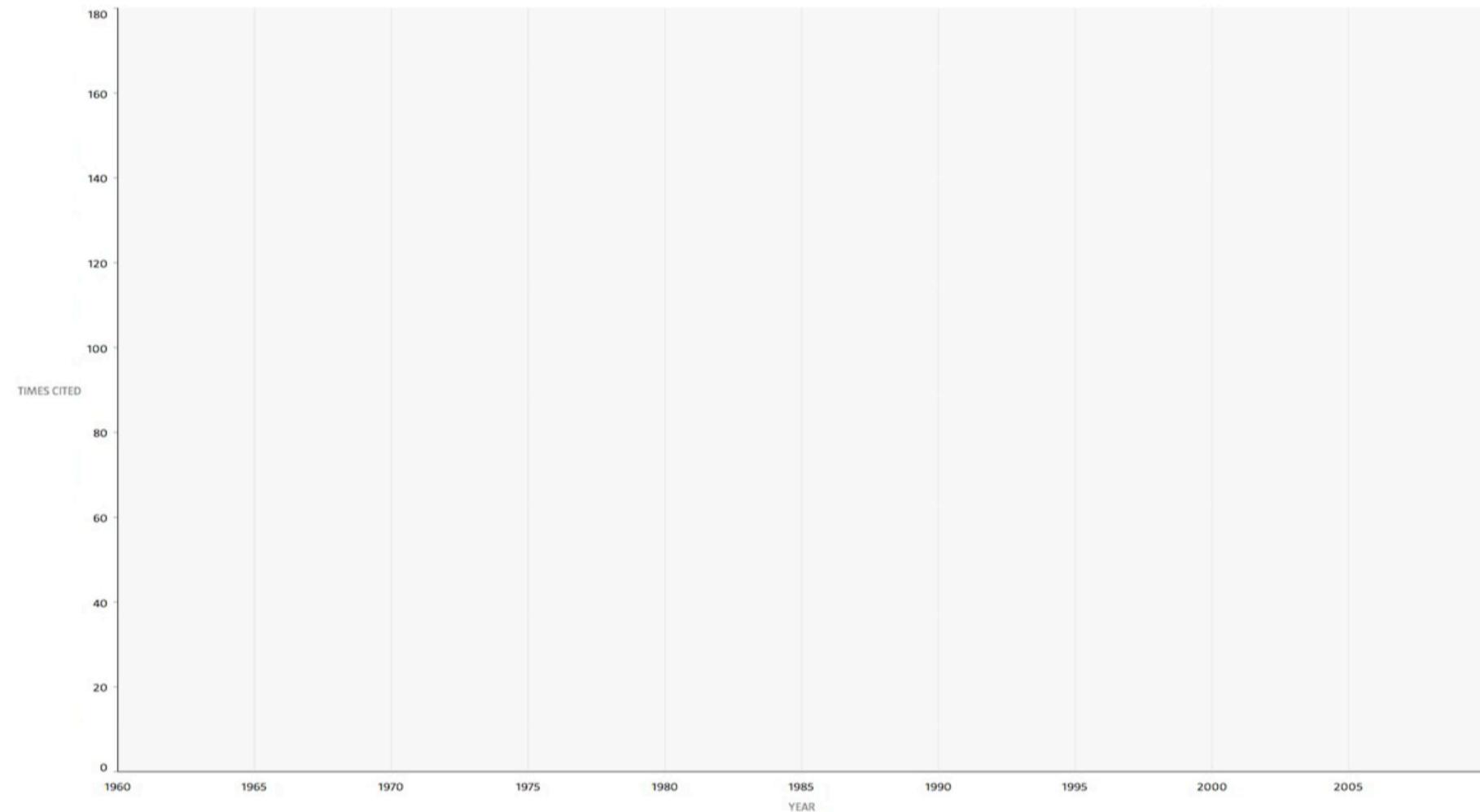
INGRESA A LA PÁGINA WEB
WWW.JUECOO.UDD.CL

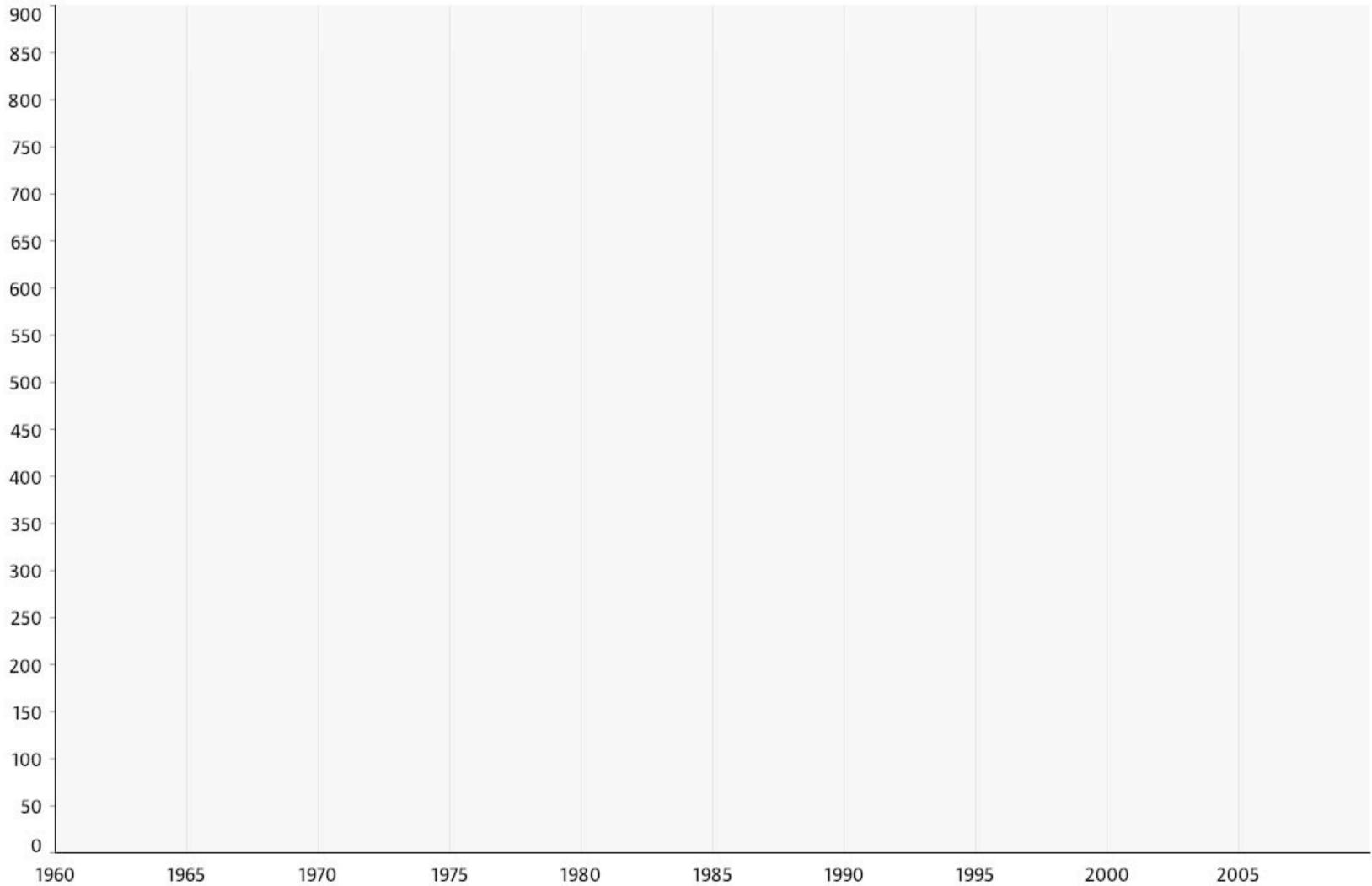
Buscar con Google

Me siento con s...

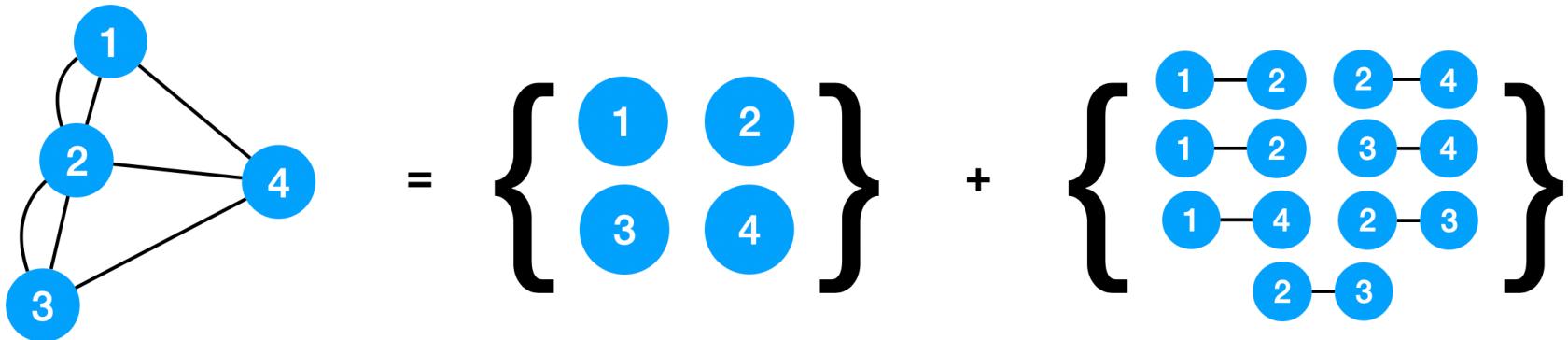
Ofrecido por Google en: Español (L...

IMPACTO CIENTÍFICO





Se ha desarrollado en distintas áreas del conocimiento



Complex Network
(Physics, Economics, Biology)

Nodes

Links

Graph
(Mathematics, Computer Science)

Vertices

Edges

Social Network
(Sociology, Political Sciences)

Actors

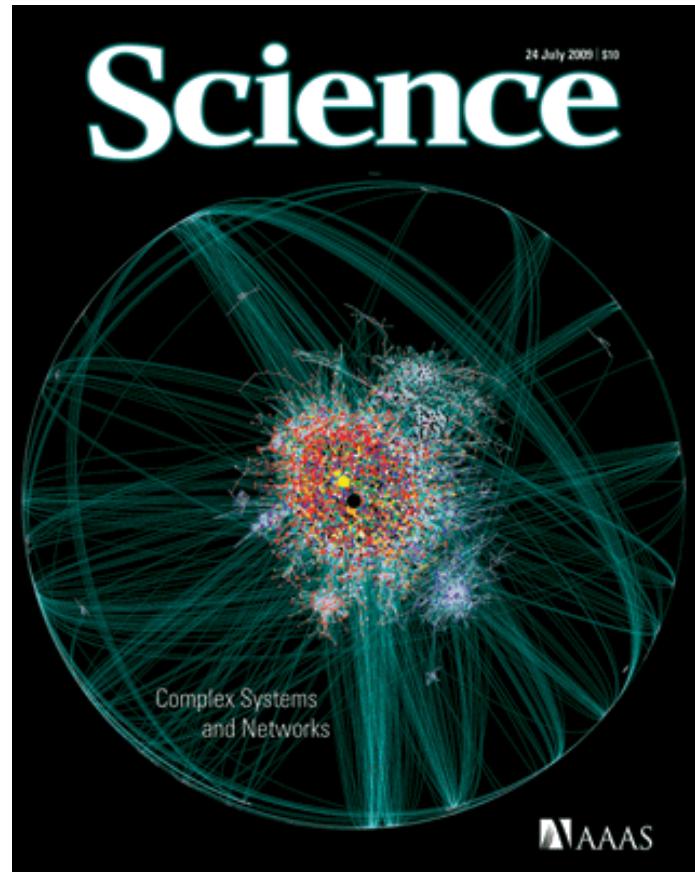
Relationships

¿Por qué Redes?

Complex systems and networks.

- Science:**

Número especial para el décimo aniversario del artículo de Barabasi & Albert 1999.



Original papers:

- 1998: artículo de Watts-Strogatz es la publicación **Nature** más citada de 1998; Destacado por ISI como uno de los diez artículos más citados en física en la década posterior a su publicación.
- 1999: El artículo de Barabasi y Albert es el artículo de **Science** más citado en 1999, destacado por ISI como uno de los diez artículos más citados en física en la década posterior a su publicación.
- 2001: Pastor -Satorras y Vespignani es uno de los dos artículos más citados entre los artículos publicados en 2001 por **Physical Review Letters**.
- 2002: Girvan-Newman es el artículo más citado en 2002 **Proceedings of the National Academy of Sciences**.

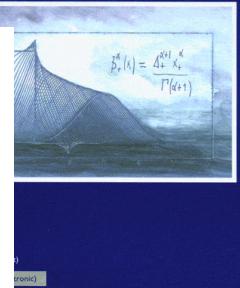
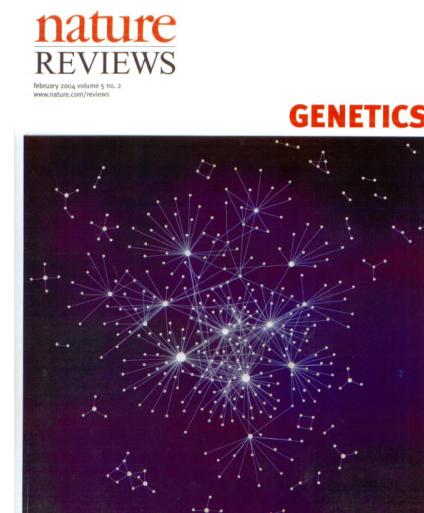
REVIEWS:

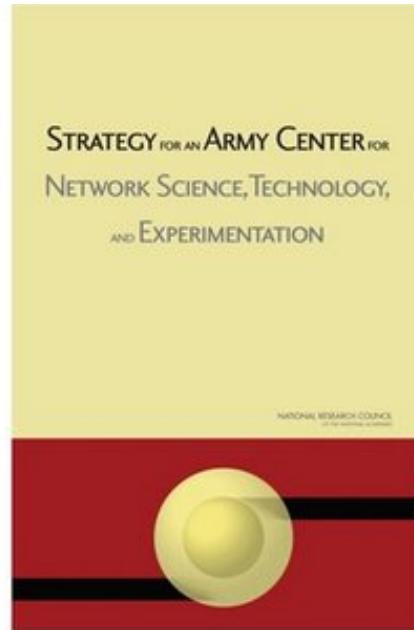
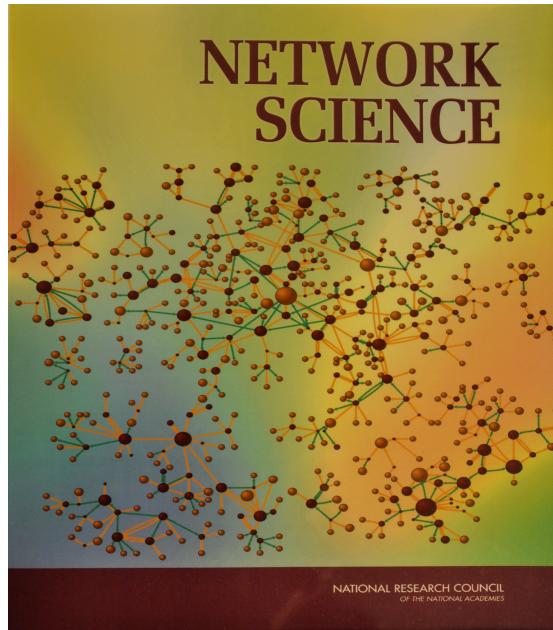
- La primera revisión de Network Science por Albert y Barabasi, 2001, es el segundo artículo más citado publicado en **Reviews of Modern Physics**, la revista de física de factor de mayor impacto, publicada desde 1929. La más citada es la revisión de Chandrasekhar de 1944 sobre procesos solares, pero será superado a finales de 2012 por Albert et al.

- Actualización de septiembre de 2016: Ahora superado, 17006 a 8545.

- La revisión SIAM de Newman sobre Network Science es el artículo más citado de todas las revistas de SIAM.

- BIOLOGÍA: “Network Biology”, de Barabasi y Oltvai (2004), es el segundo artículo más citado en la historia de Nature Reviews Genetics, la revista de reviews más importante en genética.



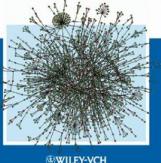


BOOKS

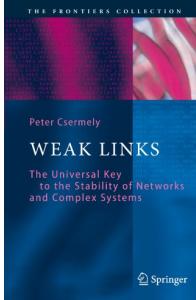
Stefan Bornholdt, Heinz Georg Schuster (Eds.)

Handbook of Graphs and Networks

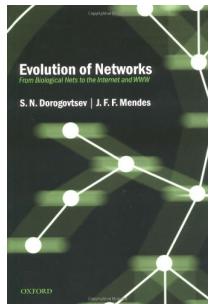
From the Genome to the Internet



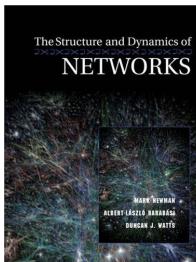
Handbook of Graphs and Networks: From the Genome to the Internet (Wiley-VCH, 2003).



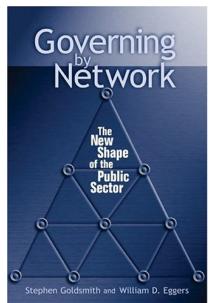
P. Csermely, Weak Links: The Universal Key to the Stability of Networks and Complex Systems (The Frontiers Collection) (Springer, 2006), 1st edn.



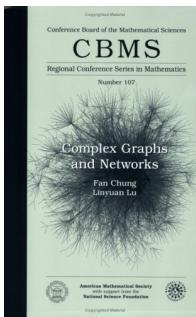
S. N. Dorogovtsev and J. F. F. Mendes, Evolution of Networks: From Biological Nets to the Internet and WWW (Oxford University Press, 2003).



M. Newman, A.-L. Barabasi, D. J. Watts, The Structure and Dynamics of Networks: (Princeton Studies in Complexity) (Princeton University Press, 2006), 1st edn.

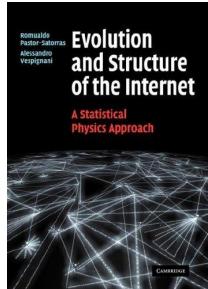


S. Goldsmith, W. D. Eggers, Governing by Network: The New Shape of the Public Sector (Brookings Institution Press, 2004).

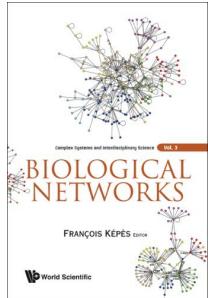


L. L. F. Chung, Complex Graphs and Networks (CBMS Regional Conference Series in Mathematics) (American Mathematical Society, 2006).

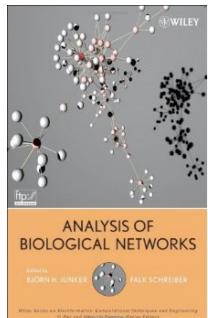
BOOKS



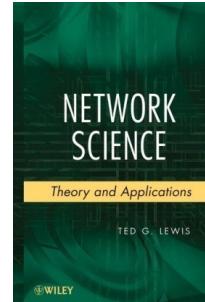
R. Pastor-Satorras, A. Vespignani, Evolution and Structure of the Internet: A Statistical Physics Approach (Cambridge University Press, 2007), rst edn.



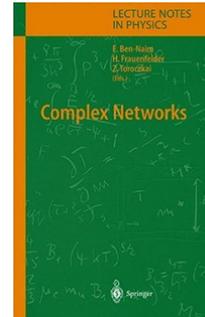
F. Képes, Biological Networks (Complex Systems and Interdisciplinary Science) (World Scientific Publishing Company, 2007), rst edn.



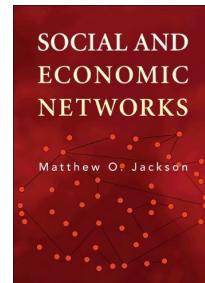
B. H. Junker, F. Schreiber, Analysis of Biological Networks (Wiley Series in Bioinformatics) (Wiley-Interscience, 2008).



T. G. Lewis, Network Science: Theory and Applications (Wiley, 2009).



E. Ben Naim, H. Frauenfelder, Z. Toroczkai, Complex Networks (Lecture Notes in Physics) (Springer, 2010), rst edn.

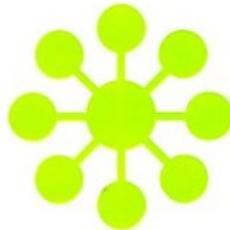


M. O. Jackson, Social and Economic Networks (Princeton University Press, 2010).

Público General

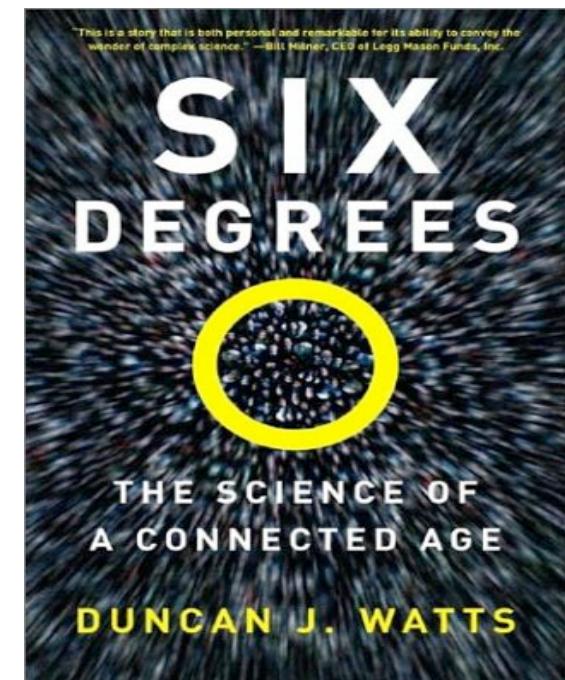
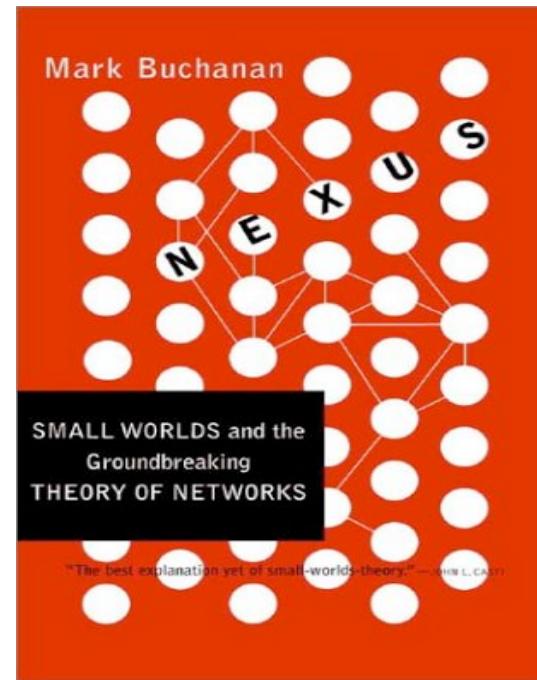
How Everything Is Connected to
Everything Else and What It Means for
Business, Science, and Everyday Life

Linked

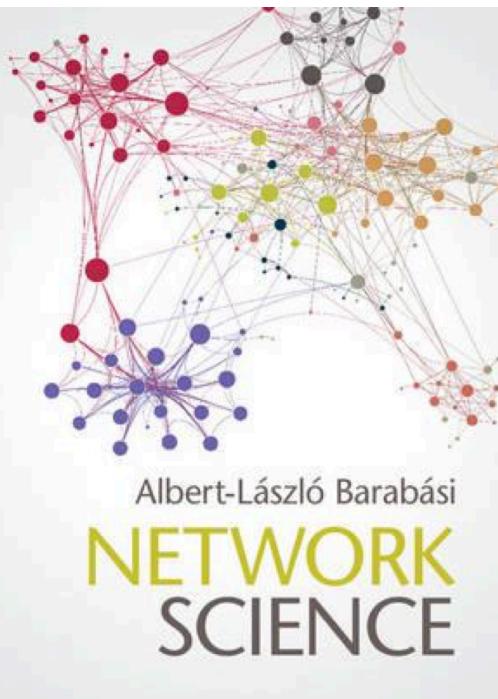


"*Linked* could alter the way we think about all of the networks that affect our lives." —The New York Times

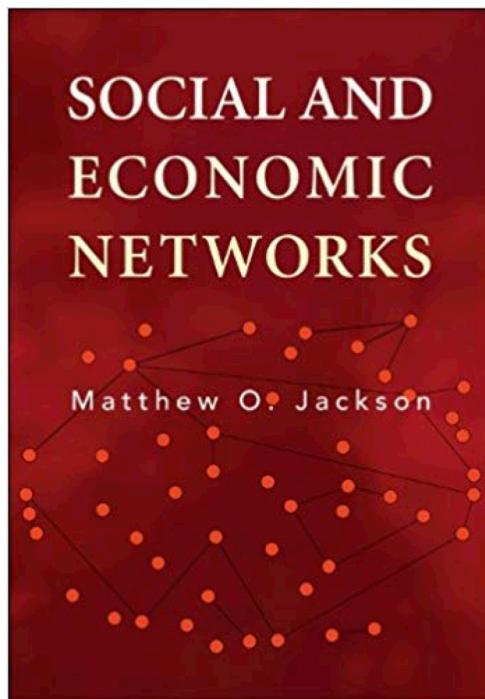
Albert-László Barabási
With a New Afterword



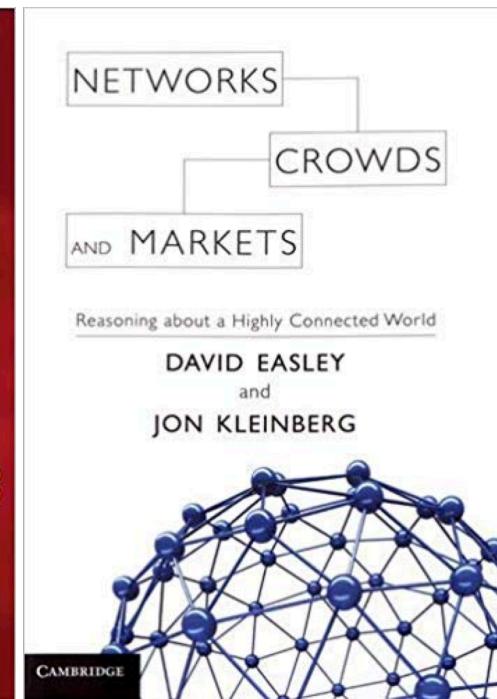
Bibliografía



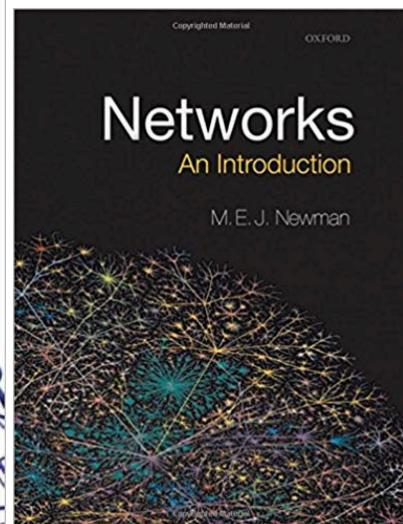
Obligatorio



Sugerido



Sugerido



Sugerido

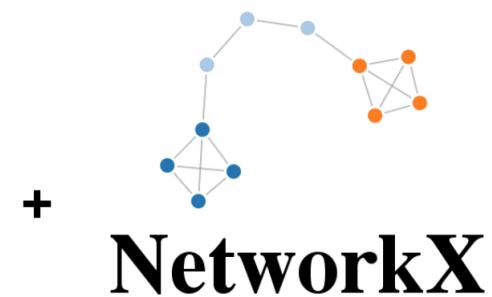
Beginner



Advanced

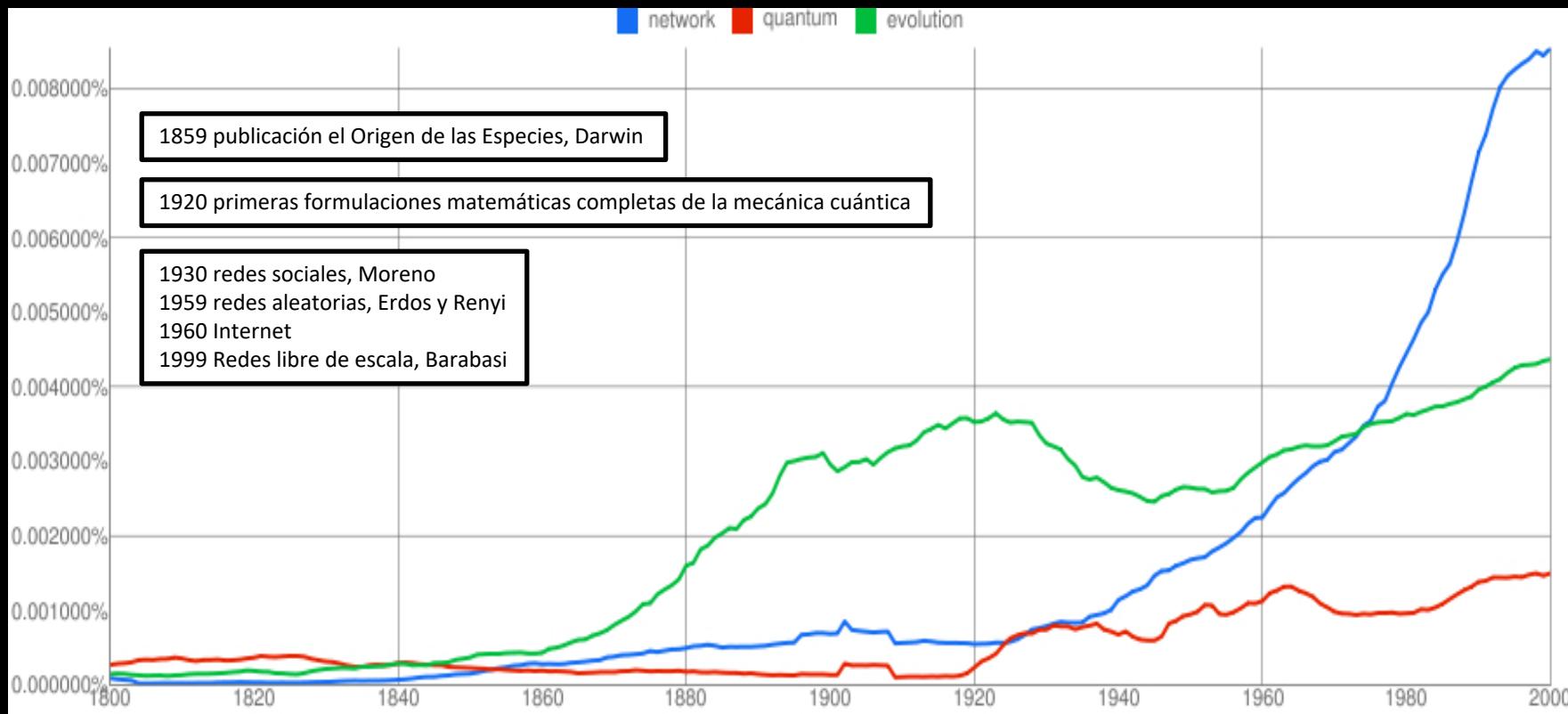


+ Pandas



DOCUMENTARY

RESUMEN



Si tuvieras que entender la propagación de enfermedades, **¿puedes hacerlo sin redes?**

Si tuviera que entender la estructura de WWW, la capacidad de búsqueda, etc., **sería inútil sin recurrir la topología de la Web.**

Si desea comprender las enfermedades humanas, **es inútil sin considerar el diagrama de cableado de la célula.**

Datasets

1) Actor:

Based on 2004 imdb data. Nodes represent actors, two nodes are connected if the corresponding actors made at least one movie together. Made for TV, direct-to-video and video game entries were removed. Preprocessed data courtesy of Hawoong Jeong.

2) Collaboration:

Scientific collaboration network based on the arXiv preprint archive's Condense Matter Physics category covering the period from January 1993 to April 2003. Each node represents an author, and two nodes are connected if they co-authored at least one paper in the dataset. Ref: Leskovec, J., Kleinberg, J., & Faloutsos, C. (2007). Graph evolution: Densification and shrinking diameters. ACM Transactions on Knowledge Discovery from Data (TKDD), 1(1), 2.

3) Internet

Network of routers connected to each other compiled by the Center for Applied Internet Data Analysis (CAIDA). Nodes are routers, and they are connected if they are directly connected via cables. For details of the non-trivial task of collecting the data see the reference. Ref:

http://www.caida.org/tools/measurement/skitter/router_topology/

4) Power Grid

Network representing the Western States Power Grid of the United States. Each node is a power plant, transformer or consumer, and two nodes are connected if they are physically connected via a cable. Ref: Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *nature*, 393(6684), 440-442.

5) Protein

Network representing the protein-protein interactions in yeast. Each node represents a protein, and they are connected if they physically interact within the cell. Original data: http://interactome.dfci.harvard.edu/S_cerevisiae/index.php?page=download Ref: Yu, H., Braun, P., Yildirim, M. A., Lemmens, I., Venkatesan, K., Sahalie, J., ... & Vidal, M. (2008). High-quality binary protein interaction map of the yeast interactome network. *Science*, 322(5898), 104-110.

6) Phone Calls:

Nodes represent a sample of cell phone users, they are connected if they have called each other at least once during the observed period. Data set contains a subset of most active users from (Song et al, 2010). Data provided by Chaoming Song. Ref: Song, C., Qu, Z., Blumm, N., & BarabÁsi, A. L. (2010). Limits of predictability in human mobility. *Science*, 327(5968), 1018-1021.

7) Citation:

Citation network of the APS journals (Physical Review Letters, Physical Review, and Reviews of Modern Physics). Each node represents a paper, and there is a directed link connecting node A to node B, if A cites B. Original data: <http://journals.aps.org/datasets> Details: Redner, S. (2004). Citation statistics from more than a century of physical review. arXiv preprint physics/0407137.

8) Metabolic

Network representing the metabolic reactions of the *E. coli* bacteria. Each node is a metabolite, and each directed link A->B means that there is a reaction where A is an input and B is a product (e.g. A+C -> B+D). Original data: <http://bigg.ucsd.edu/> Ref: Schellenberger, J., Park, J. O., Conrad, T. M., & Palsson, B. Ø. (2010). BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions. *BMC bioinformatics*, 11(1), 213.

9) Email:

Email network based on traffic data collected for 112 days at University of Kiel, Germany. Each node is an email address, and there is a directed link from node A to node B, if A sent at least one email to B. Ref: Ebel, H., Mielsch, L. I., & Bornholdt, S. (2002). Scale-free topology of e-mail networks. *Physical review E*, 66(3), 035103.

10) WWW