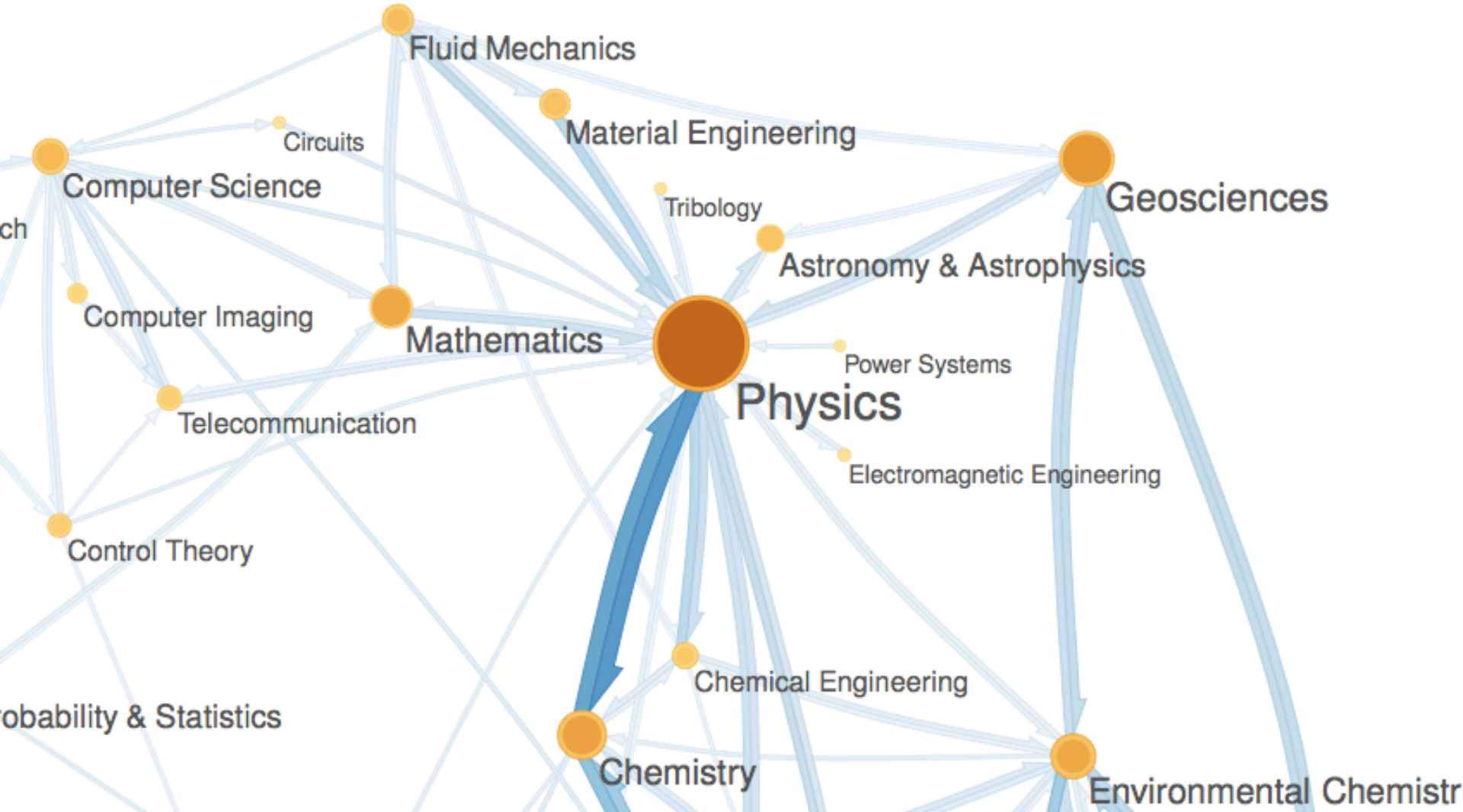
Mapping the HRA Community

Jevin West, Information School, University of Washington



JW

What is my impact on science?

$7,933,670,366

JW

22,756 awards

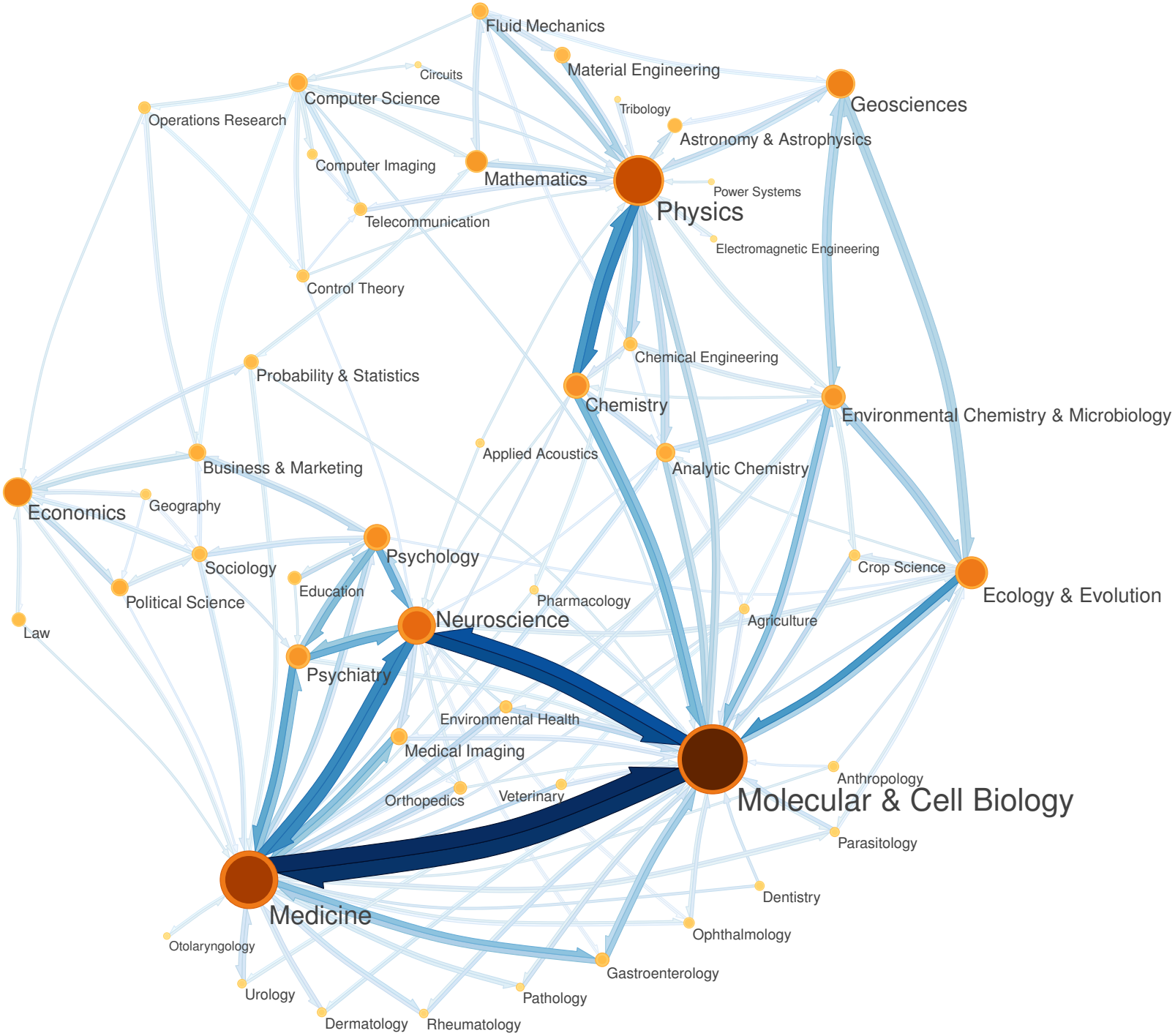
17,849 researchers

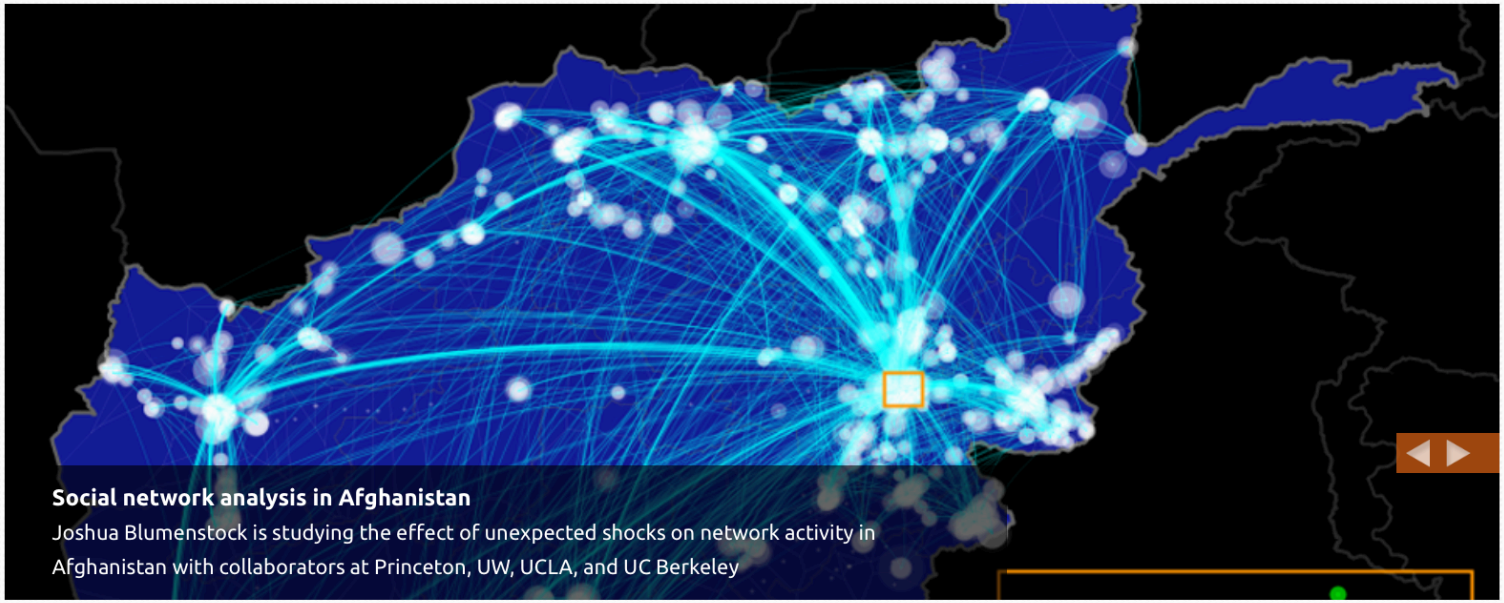
344,917 papers

8,174,533 citations

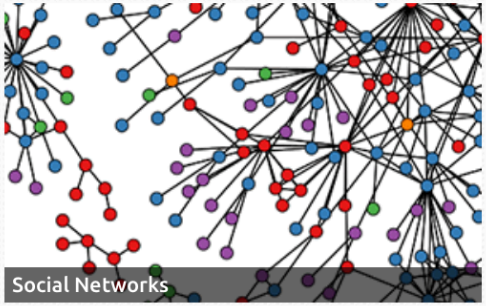
23. 7 citations/paper

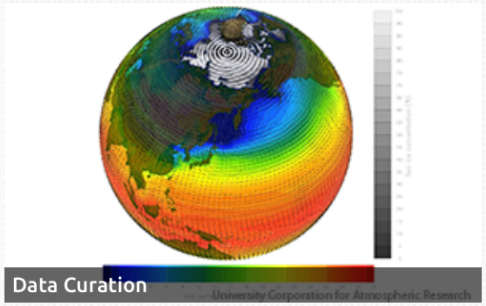
2006 - 2015

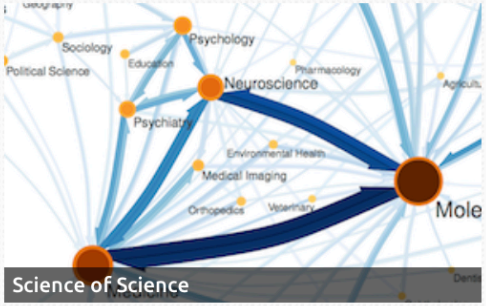

Fluid Mechanics
Material Engineering
Circuits
Computer Science
Geosciences
Operations Research
Tribology
Astronomy & Astrophysics
Computer Imaging
Mathematics
Power Systems
Physics
Telecommunication
Electromagnetic Engineering
Control Theory
Chemical Engineering
Probability & Statistics
Chemistry
Environmental Chemistry & Microbiology
Applied Acoustics
Business & Marketing
Analytic Chemistry
Economics
Geography
Psychology
Crop Science
Sociology
Education
Political Science
Pharmacology
Ecology & Evolution
Neuroscience
Law
Agriculture
Psychiatry
Environmental Health
Medical Imaging
Anthropology
Orthopedics
Veterinary
Molecular & Cell Biology
Parasitology
Dentistry
Medicine
Ophthalmology
Otolaryngology
Gastroenterology
Urology
Pathology
Dermatology
Rheumatology



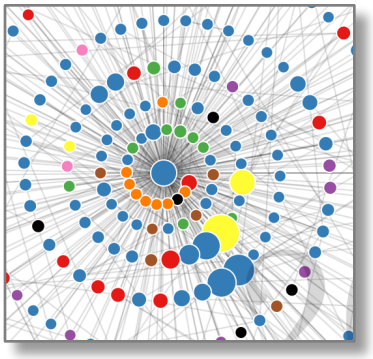
Research Focus Areas

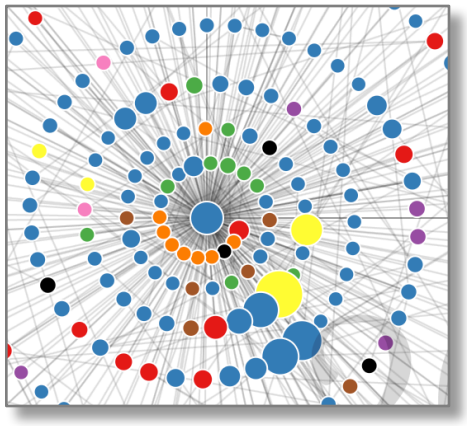




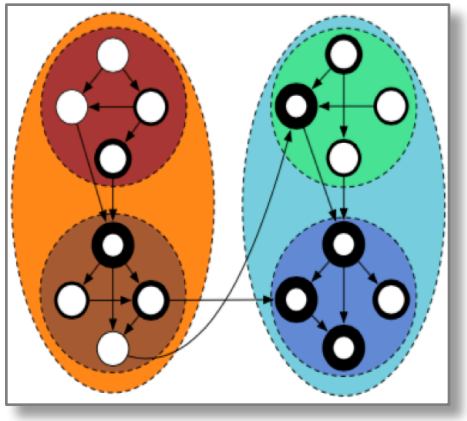


Science of Science



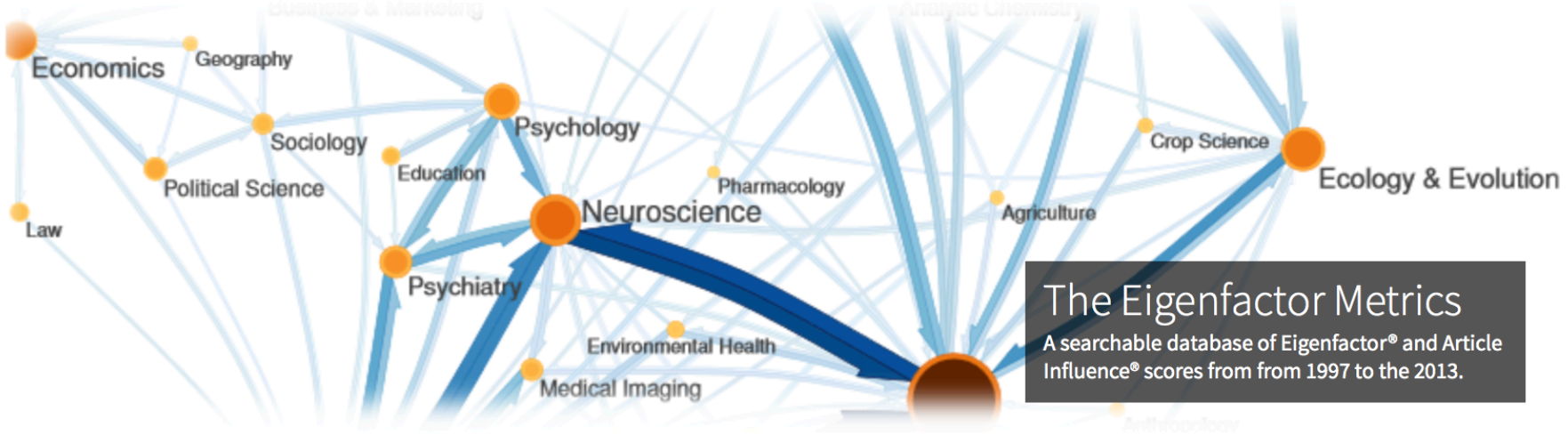


Knowledge Science



Knowledge Engineering

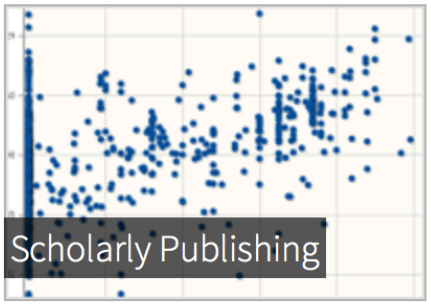
EIGENFACTOR.org



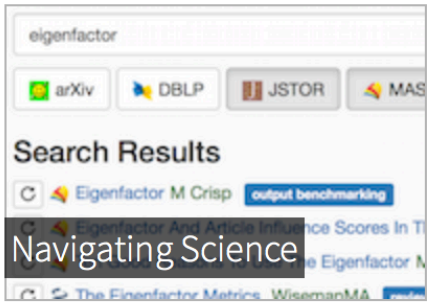
The Eigenfactor Metrics

A searchable database of Eigenfactor® and Article Influence® scores from from 1997 to the 2013.

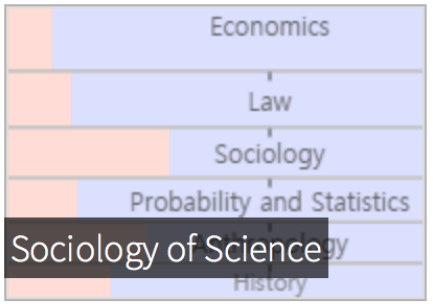
RESEARCH AREAS


Scholarly Publishing

Mapping Science



Navigating Science

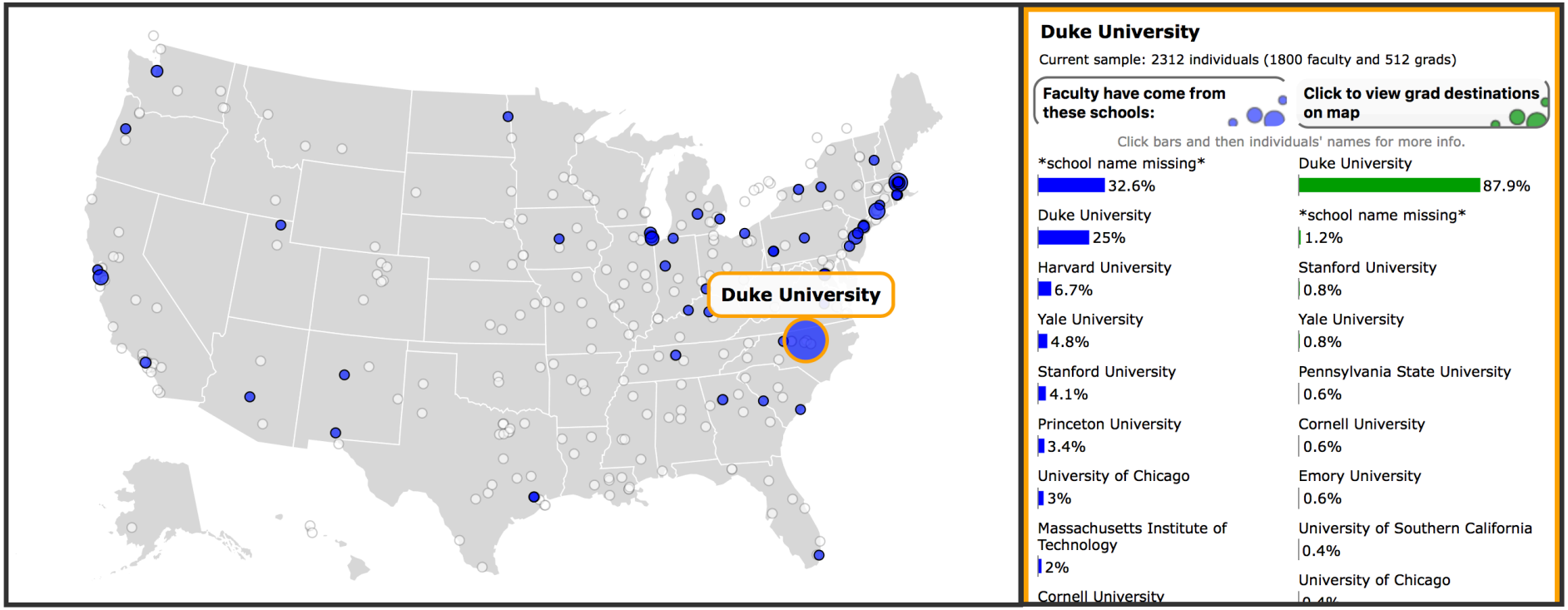


Sociology of Science

Faculty Migration

U.S. ACADEMIC MIGRATION MAP

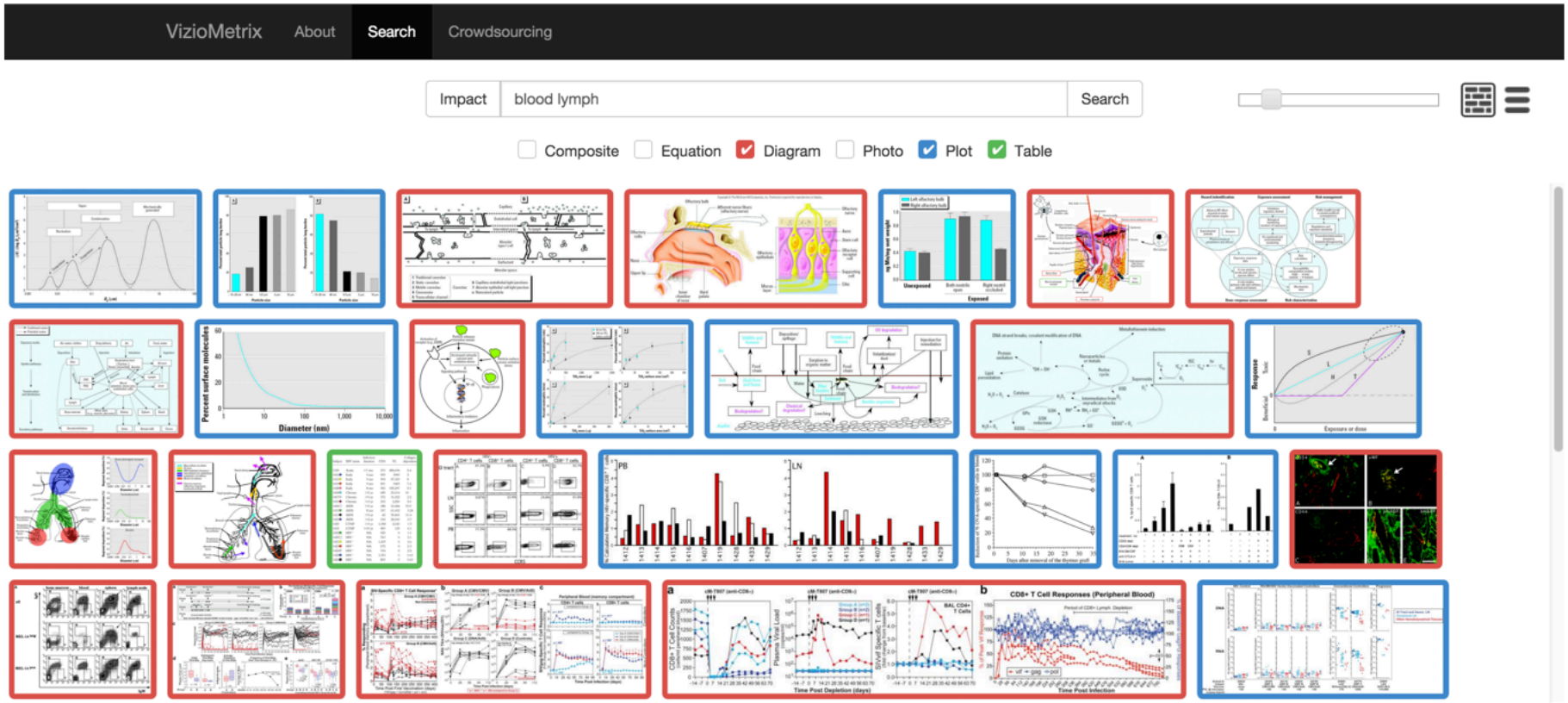
Where do people who pursue academic careers in the U.S. go on to land faculty positions after earning their advanced degrees?
Where do faculty come from? Click on a school to explore.



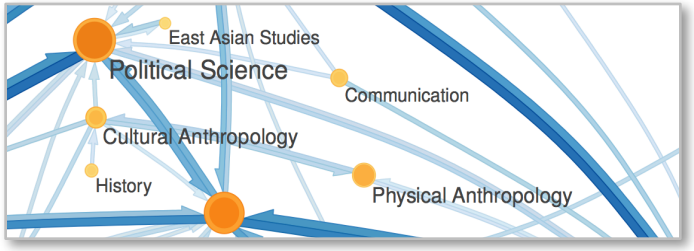
Eigenfactor.org/projects/migration/

Figure-Centric Search Engine

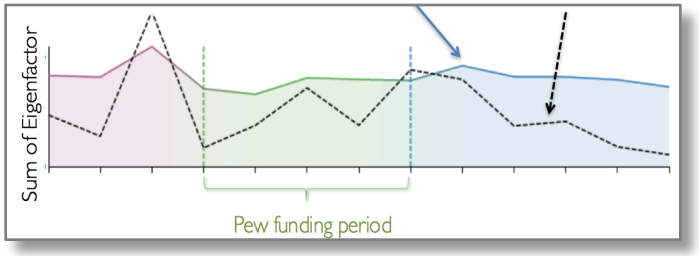
viziometrics.org



Talk Outline



Visualizing scholarly influence



Measuring funding impact

The H-index impact on science…

Jure Leskovec

Follow

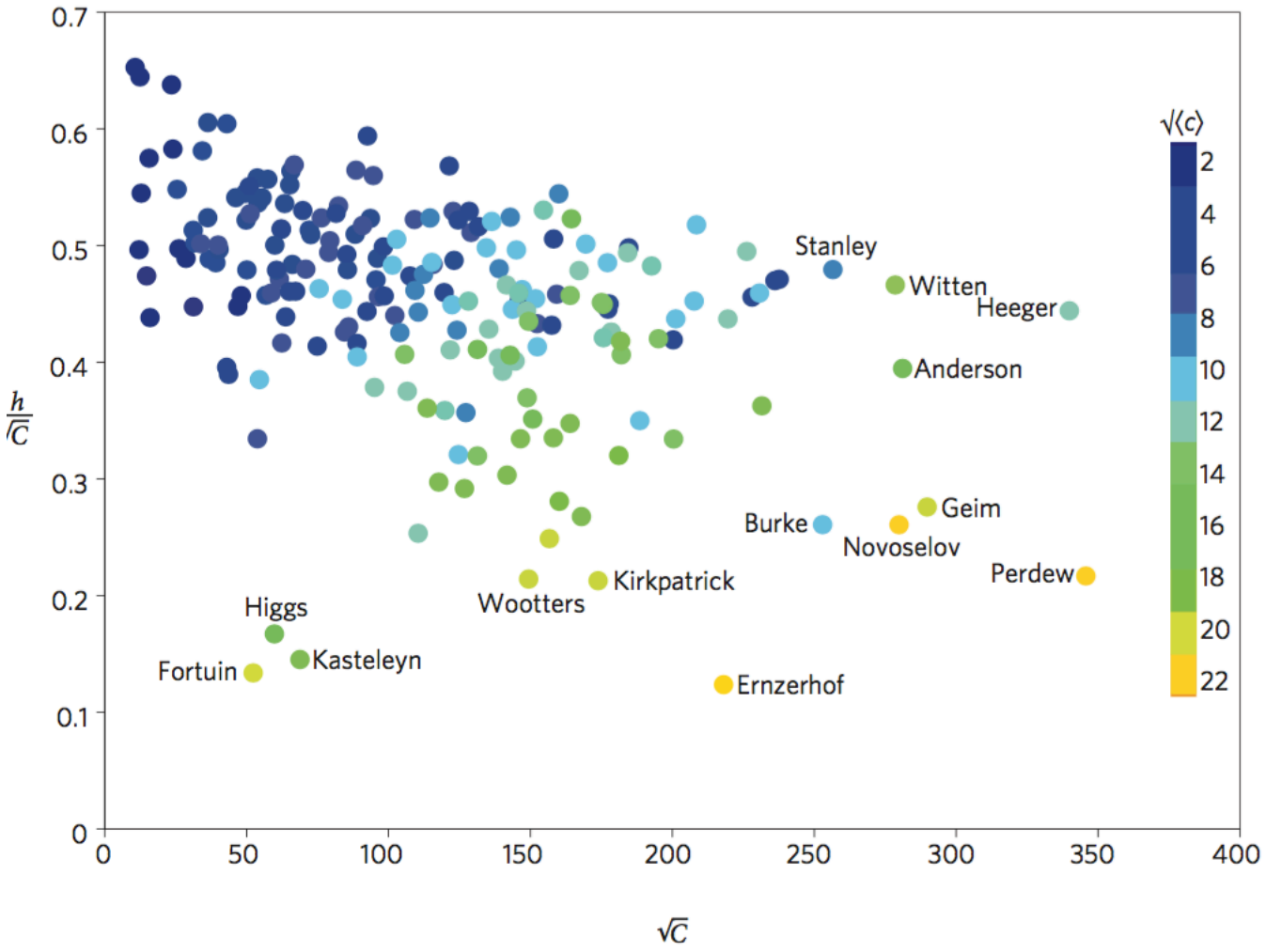
Professor of Computer Science, Stanford University
Data mining, Social Network Analysis, Information Networks
Verified email at cs.stanford.edu - Homepage

Google Scholar

Title 1–20	Cited by	Year
Graphs over time: densification laws, shrinking diameters and possible explanations J Leskovec, J Kleinberg, C Faloutsos Proceedings of the eleventh ACM SIGKDD international conference on Knowledge ...	1373	2005
The dynamics of viral marketing J Leskovec, LA Adamic, BA Huberman ACM Transactions on the Web (TWEB) 1 (1), 5	1338	2007
Cost-effective outbreak detection in networks J Leskovec, A Krause, C Guestrin, C Faloutsos, J VanBriesen, N Glance Proceedings of the 13th ACM SIGKDD international conference on Knowledge ...	887	2007
Meme-tracking and the dynamics of the news cycle J Leskovec, L Backstrom, J Kleinberg Proceedings of the 15th ACM SIGKDD international conference on Knowledge ...	885	2009
Graph evolution: Densification and shrinking diameters J Leskovec, J Kleinberg, C Faloutsos ACM Transactions on Knowledge Discovery from Data (TKDD) 1 (1), 2	853	2007
Friendship and mobility: user movement in location-based social networks E Cho, SA Myers, J Leskovec Proceedings of the 17th ACM SIGKDD international conference on Knowledge ...	728	2011
Community structure in large networks: Natural cluster sizes and the absence		

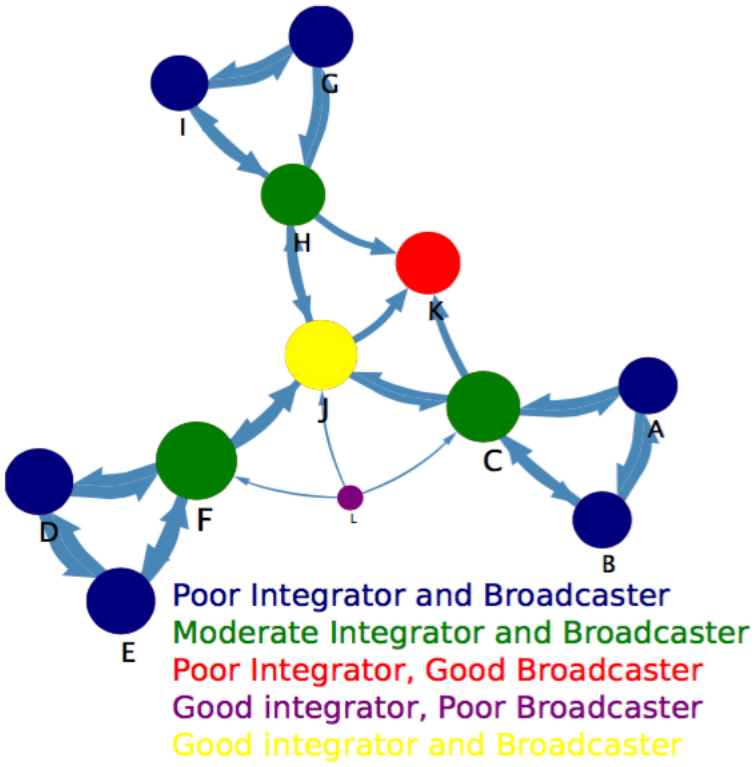
Citation indices	All	Since 2010
Citations	19409	17853
h-index	59	56
i10-index	103	101

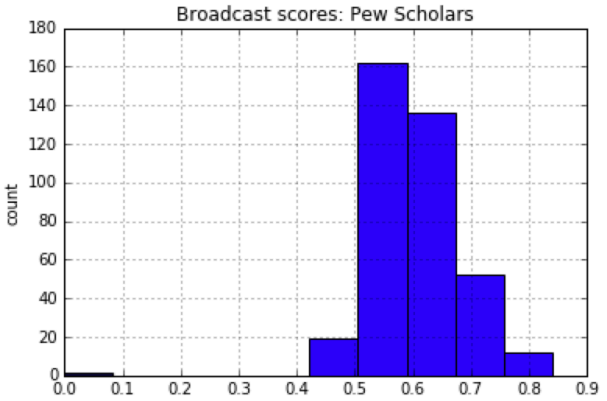
An evisceration of the H-index…

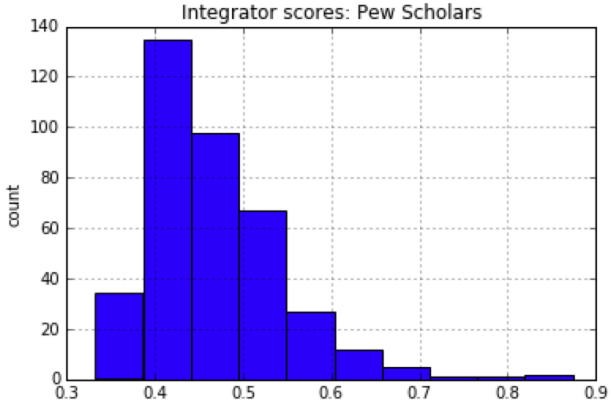


S. N. Dorogovtsev and J. F. F. Mendes (2015) Nature Physics

Measuring Interdisciplinarity



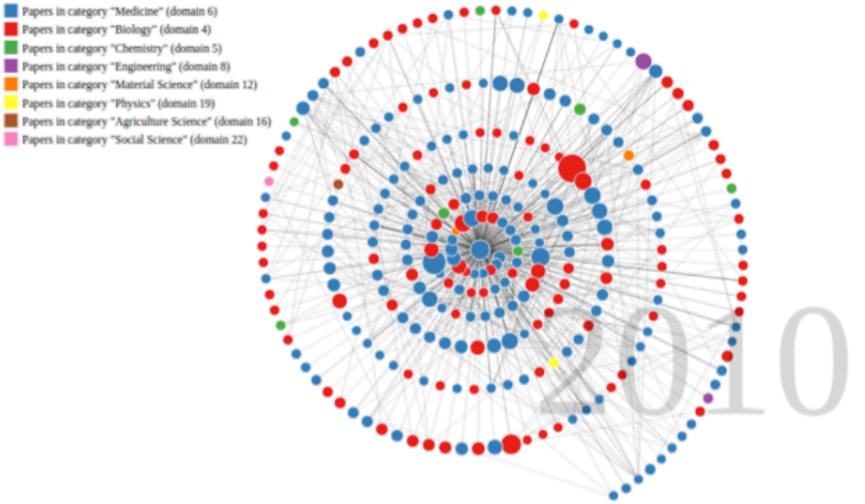




Bergstrom, CT, Foster, J, Portenoy, J, A. Misra, West, JD. (2016). Measuring interdisciplinarity without subject categories. (in prep)

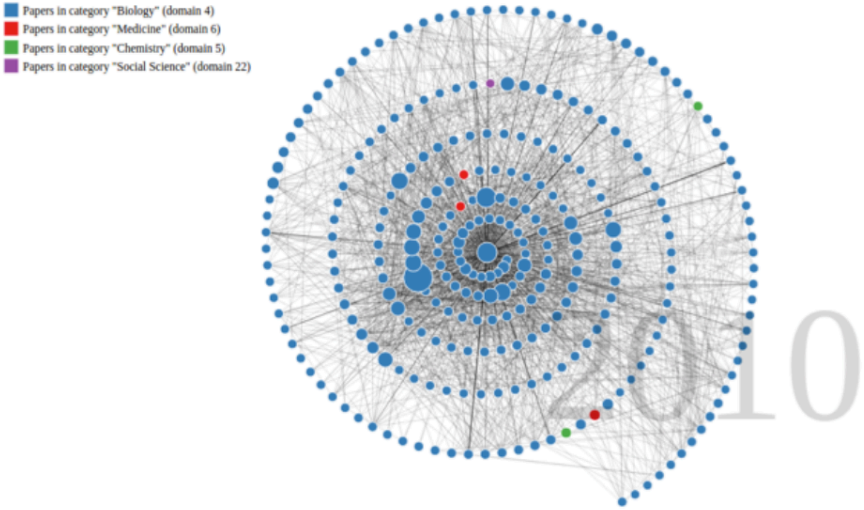
Visualizing Interdisciplinarity

Jason Portenoy



A more sparse network indicates fewer citations between papers shown in the network. This could be a result of the central scholar having impact across a wider set of academic communities.

A denser network means that the papers that cite the central author also tend to cite each other.



scholar.eigenfactor.org

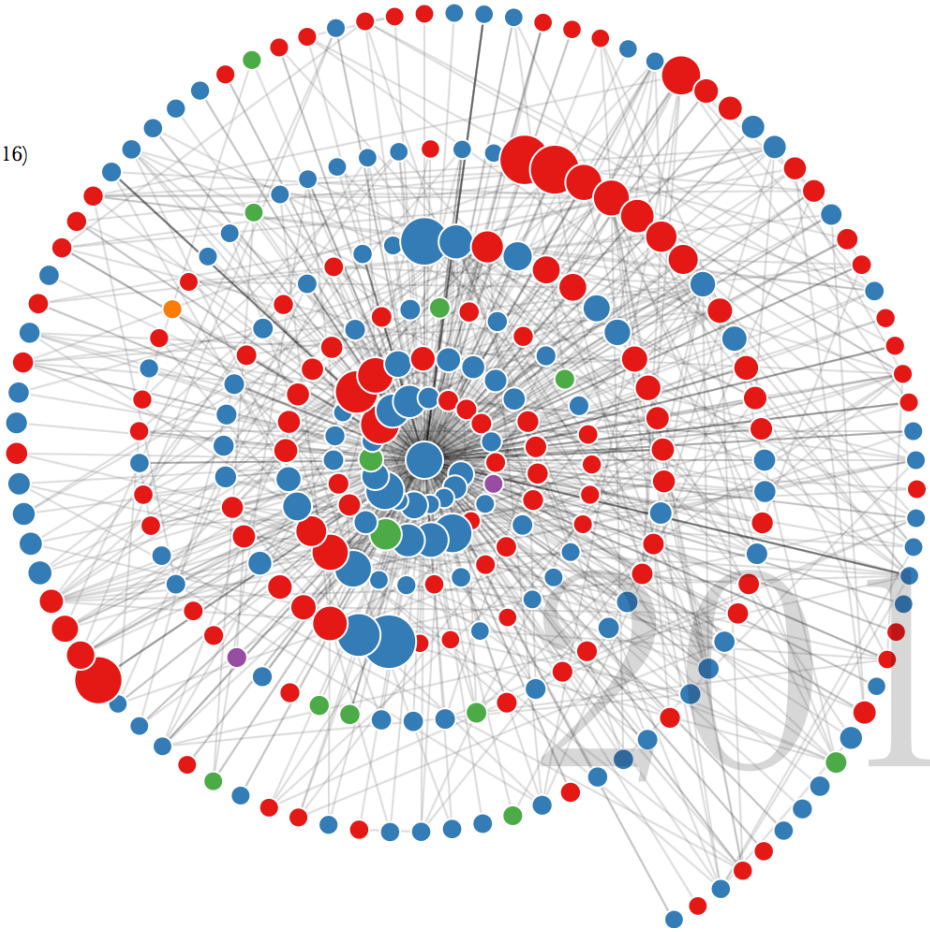
Visualizing Scholarly Influence Over Time

Influence of Pew Scholars

Roberta A. Gottlieb

Learn More

- Papers in category "Medicine" (domain 6)
- Papers in category "Biology" (domain 4)
- Papers in category "Chemistry" (domain 5)
- Papers in category "Unknown" (domain 0)
- Papers in category "Agriculture Science" (domain 16)

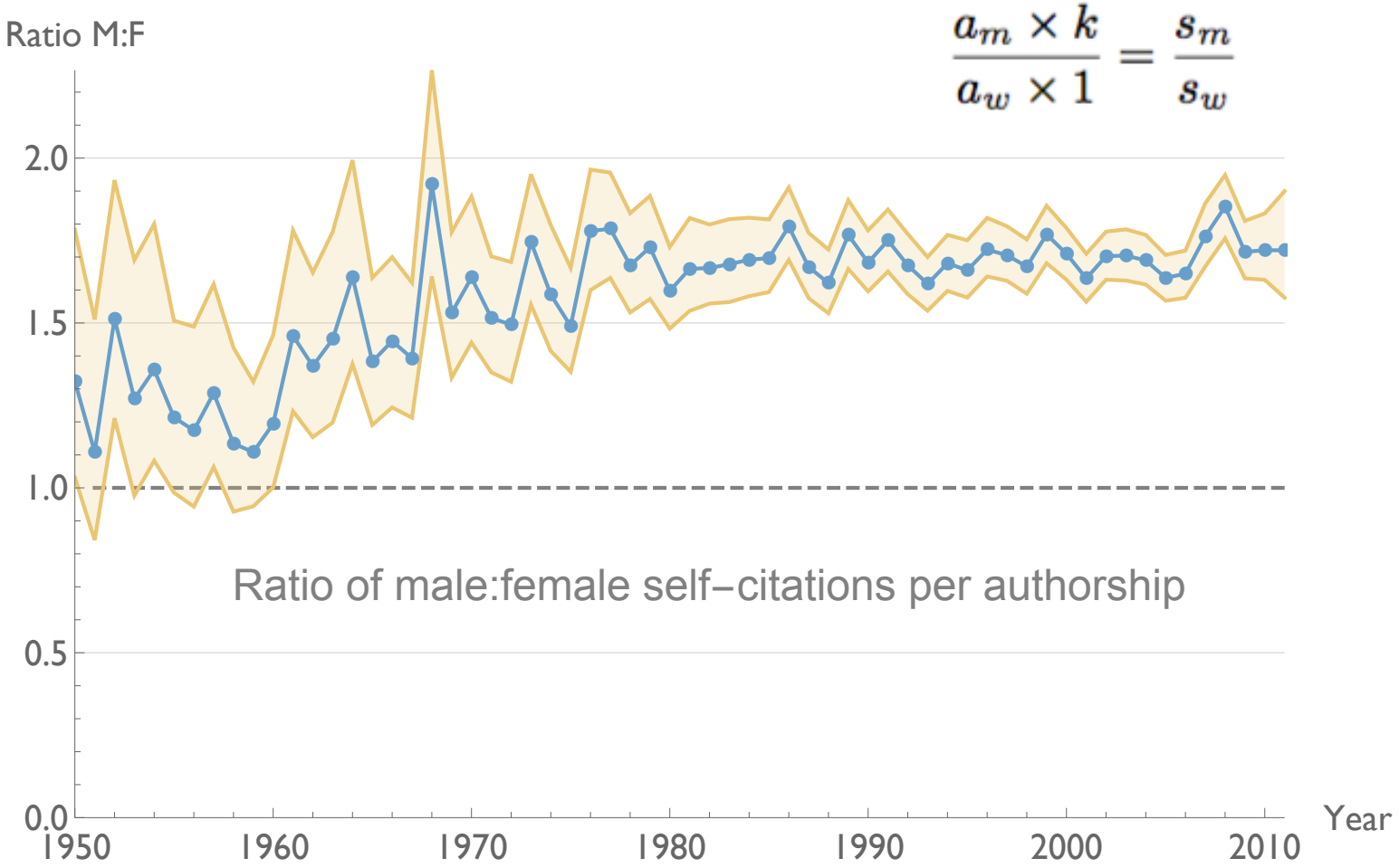


Roberta A. Gottlieb

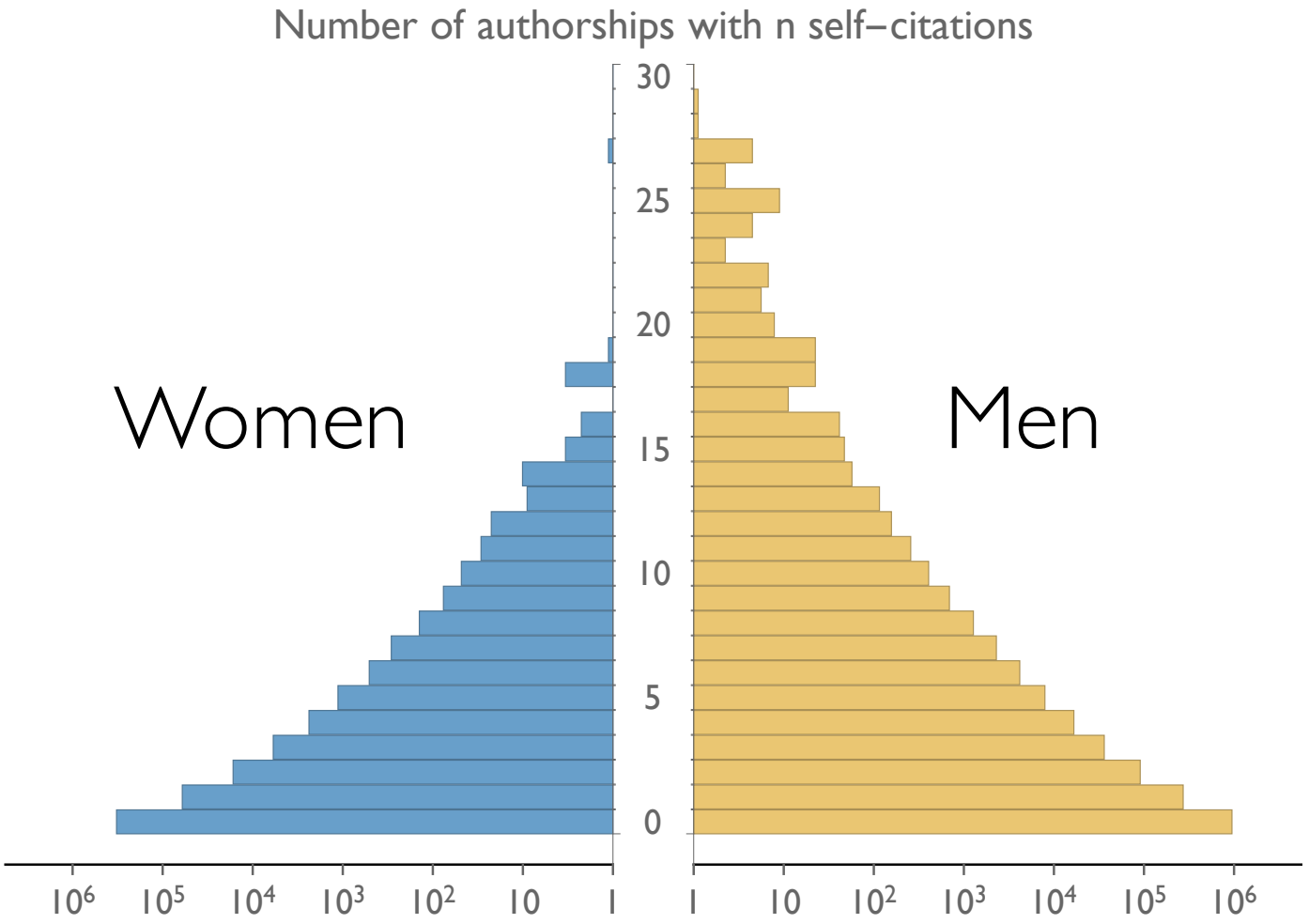
Pew Scholar
1997

2012

Self-citation over time

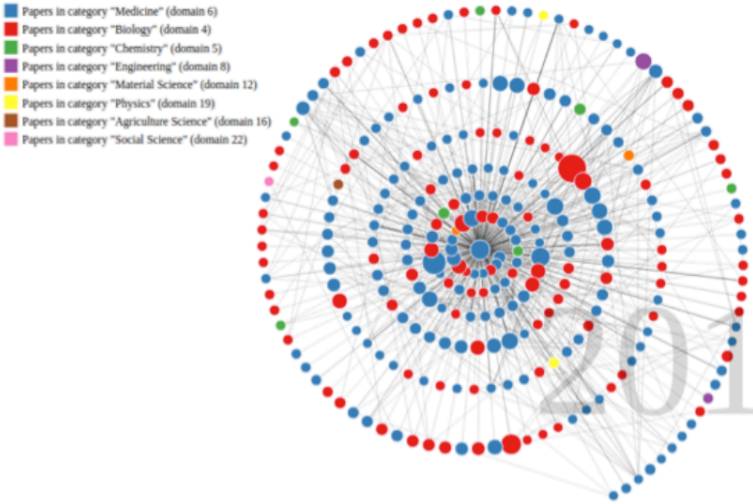


Gender and Self Citation



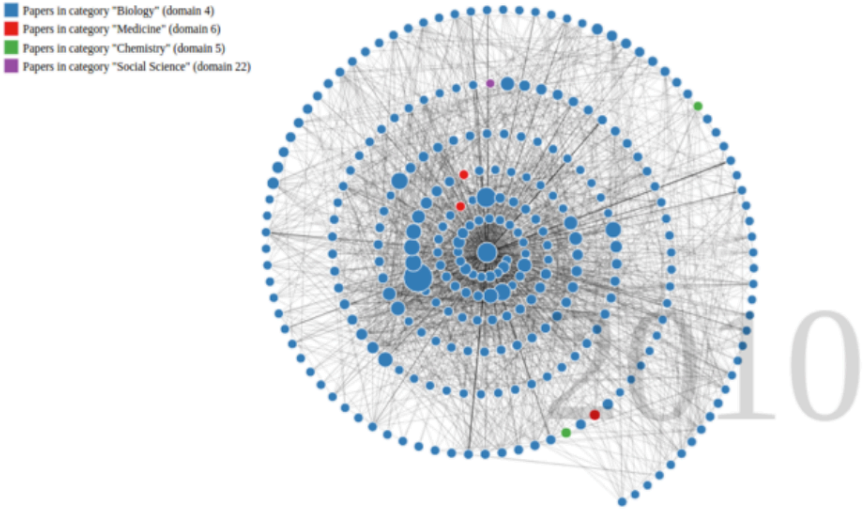
King, M et al. (2016) Men set their own cites high: Gender and self-citation across fields and over time. (in prep)

Comparing Authors



A more sparse network indicates fewer citations between papers shown in the network. This could be a result of the central scholar having impact across a wider set of academic communities.

A denser network means that the papers that cite the central author also tend to cite each other.



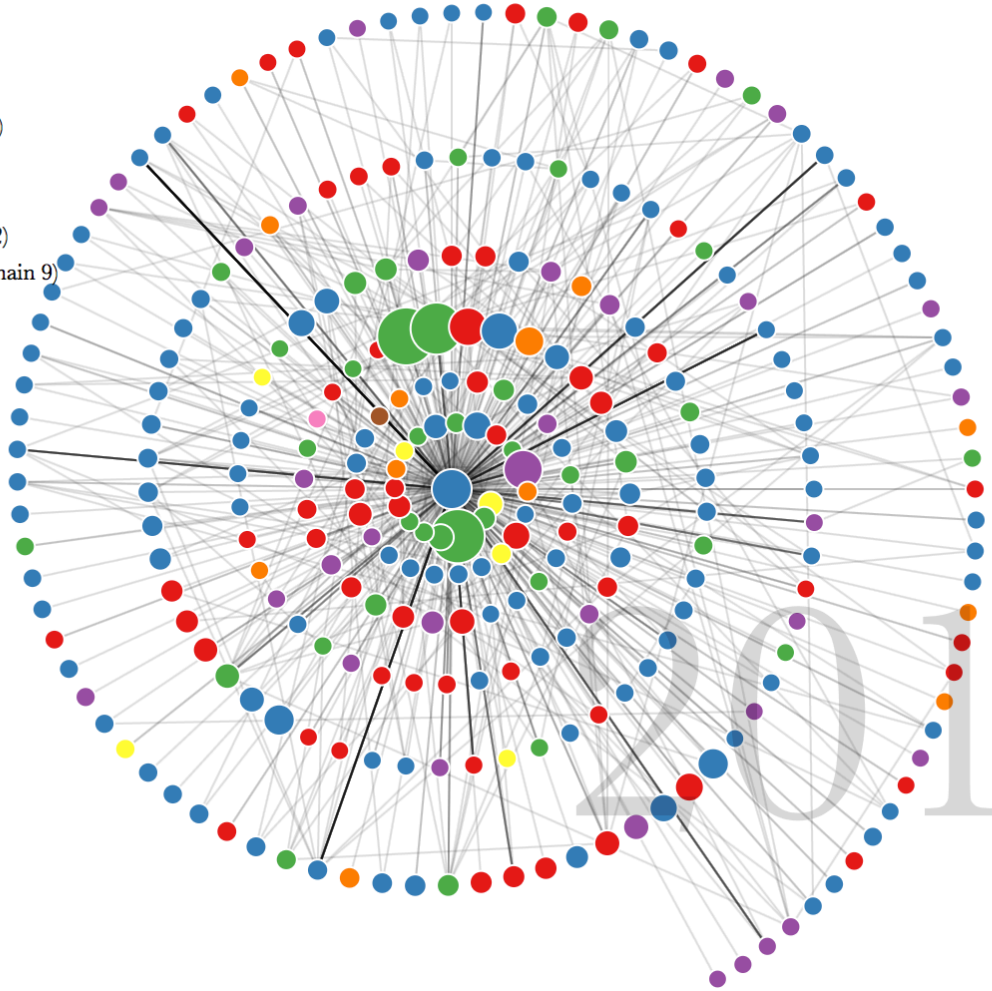
Visualizing Scholarly Influence Over Time

Influence of Pew Scholars

Mark W. Grinstaff

Learn More

- Papers in category "Chemistry" (domain 5)
- Papers in category "Medicine" (domain 6)
- Papers in category "Biology" (domain 4)
- Papers in category "Material Science" (domain 12)
- Papers in category "Engineering" (domain 8)
- Papers in category "Physics" (domain 19)
- Papers in category "Computer Science" (domain 2)
- Papers in category "Environmental Sciences" (domain 9)



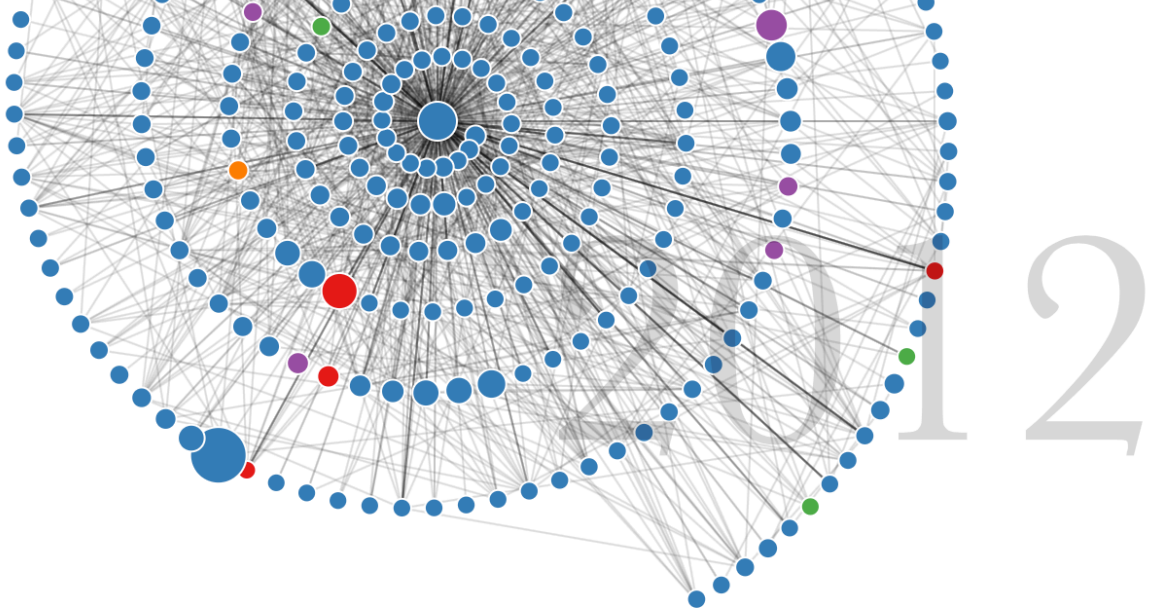
Mark W.
Grinstaff

Pew Scholar
1999

2012

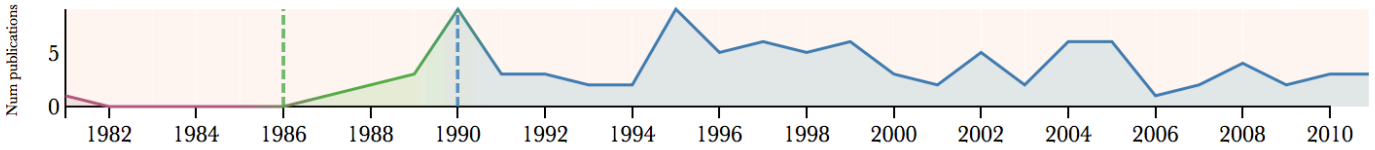
Philip A. Hieter

Pew Scholar
1986

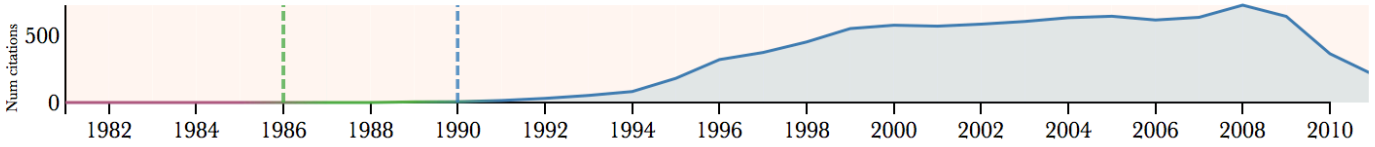


2012

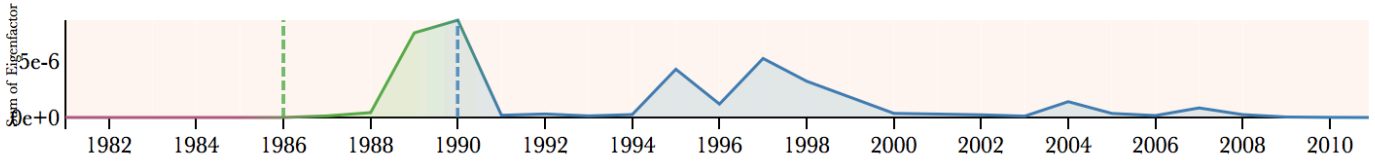
Number of publications



Number of citations received



Sum of eigenfactor for this author's publications by year



Funding

JW

Citation Data

49 million scholarly publications

260 million citations

354 Pew Scholars

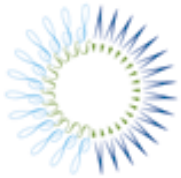
22,000 publications

62 publications/scholar

Pew EF 3 times the average EF

field classification

JW



THE PEW CHARITABLE TRUSTS | Scholars Program *in the* Biomedical Sciences

~ 37 citations/paper

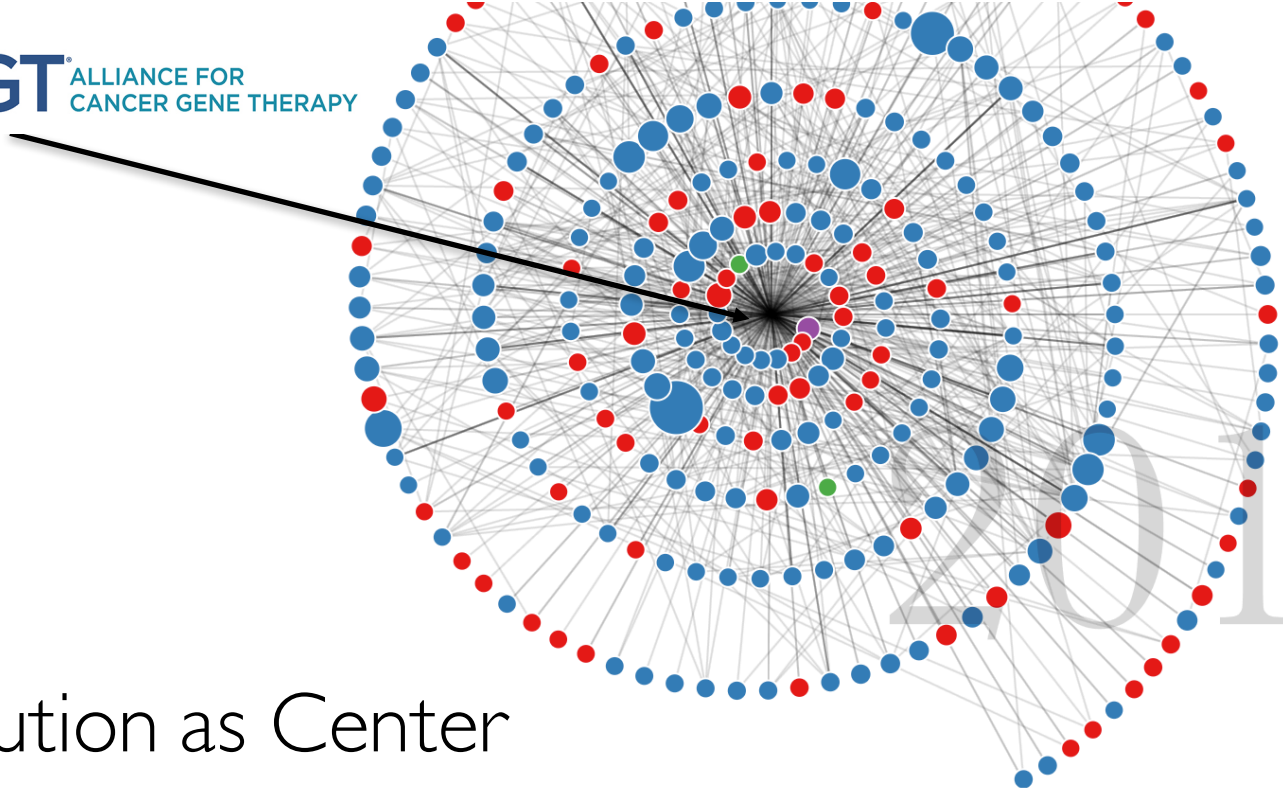
median citations = 11

Science

~ 5 citations/paper

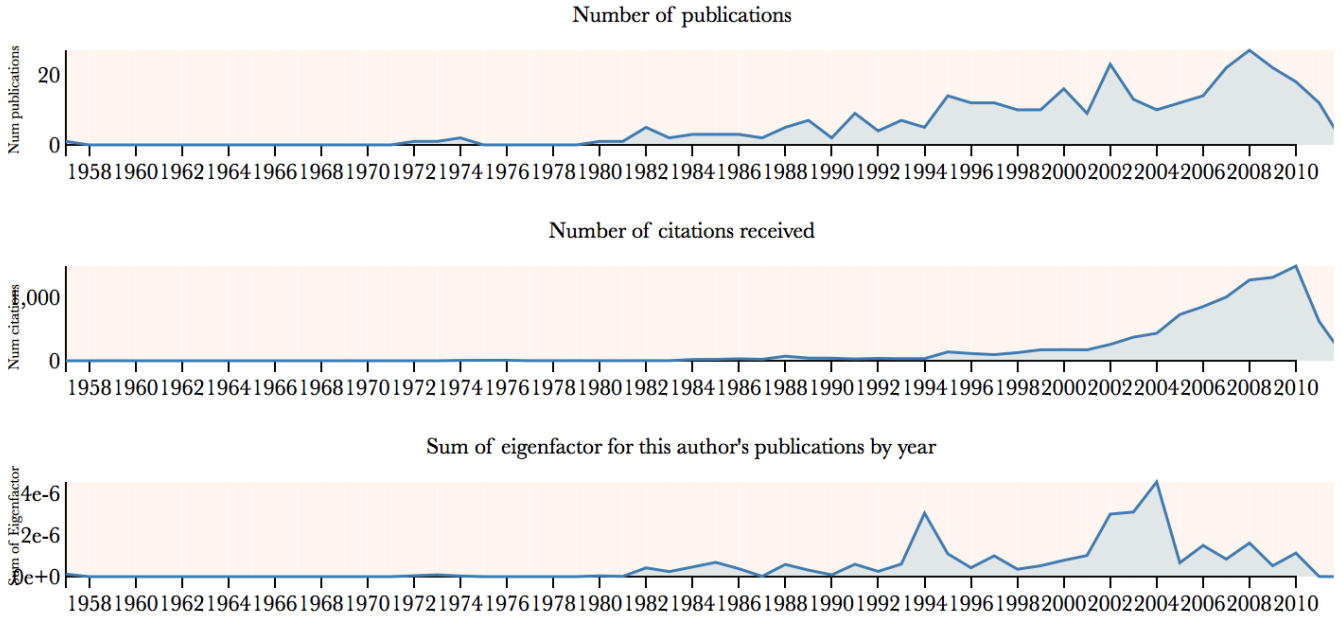
median citations = 0

ACGT ALLIANCE FOR CANCER GENE THERAPY



2012

Institution as Center



JW

scholar.eigenfactor.org

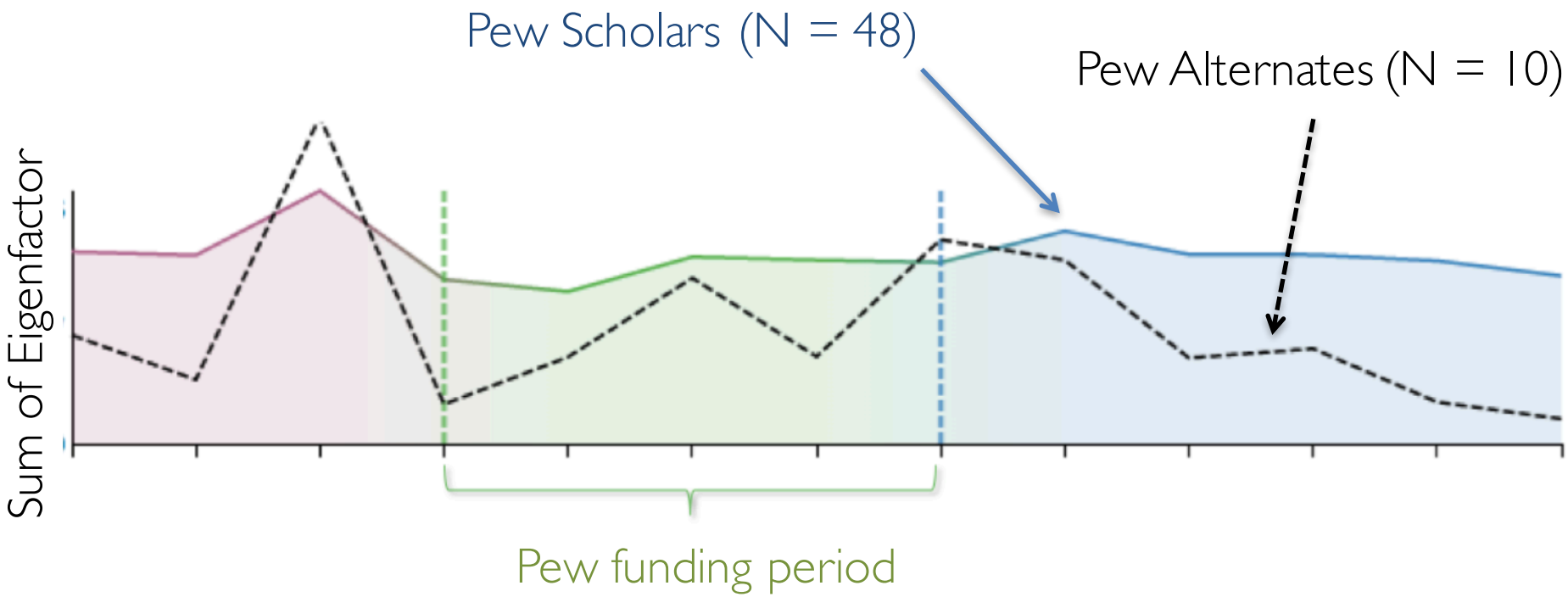
username: PewScholar

password: 1N!kdG

Jevin West, jevinw@uw.edu

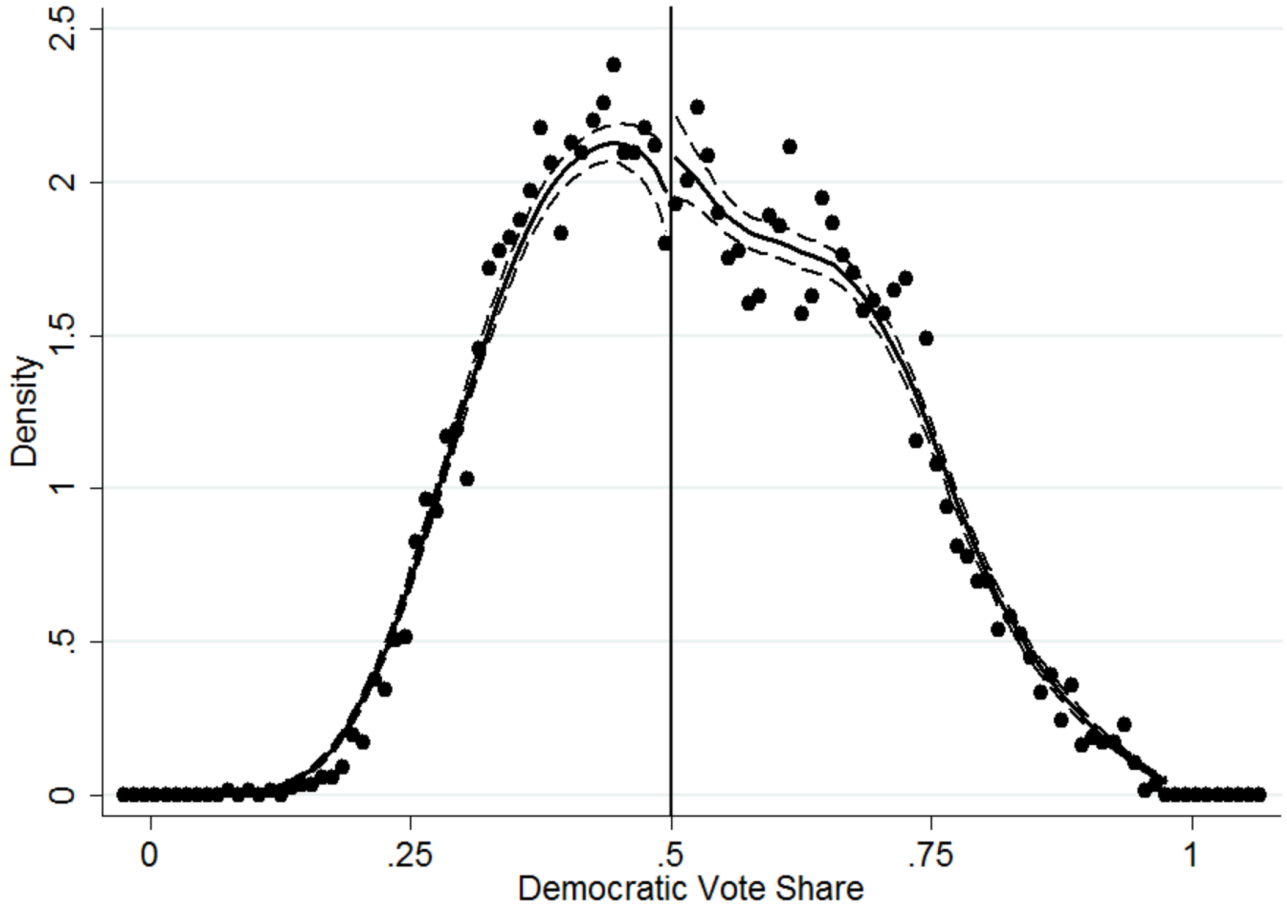
The Pew Impact…

Comparing Alternates



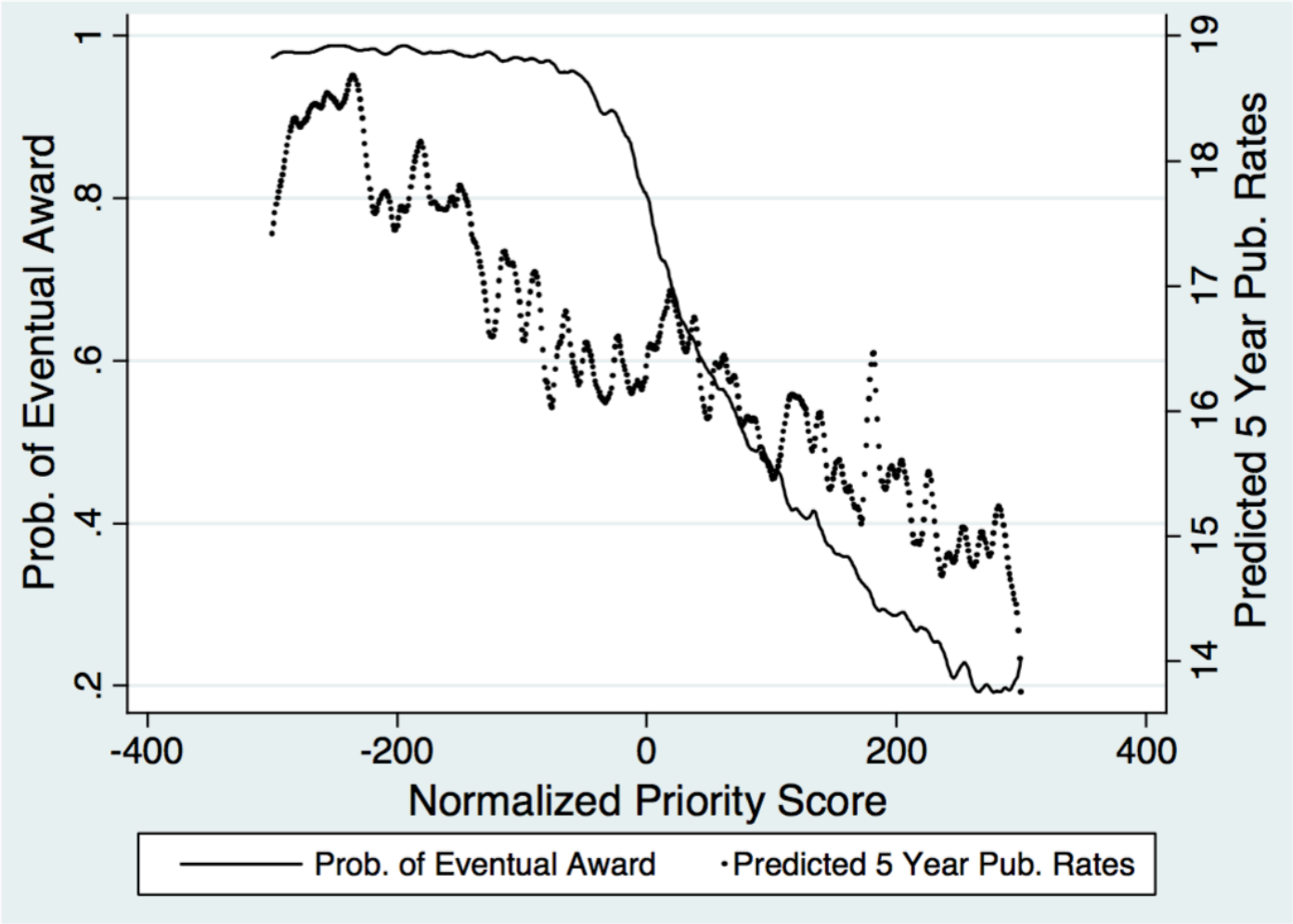
* Includes scholars and alternates from cohort years: 1997, 1999, 2000, 2001, 2002

Regression discontinuity design



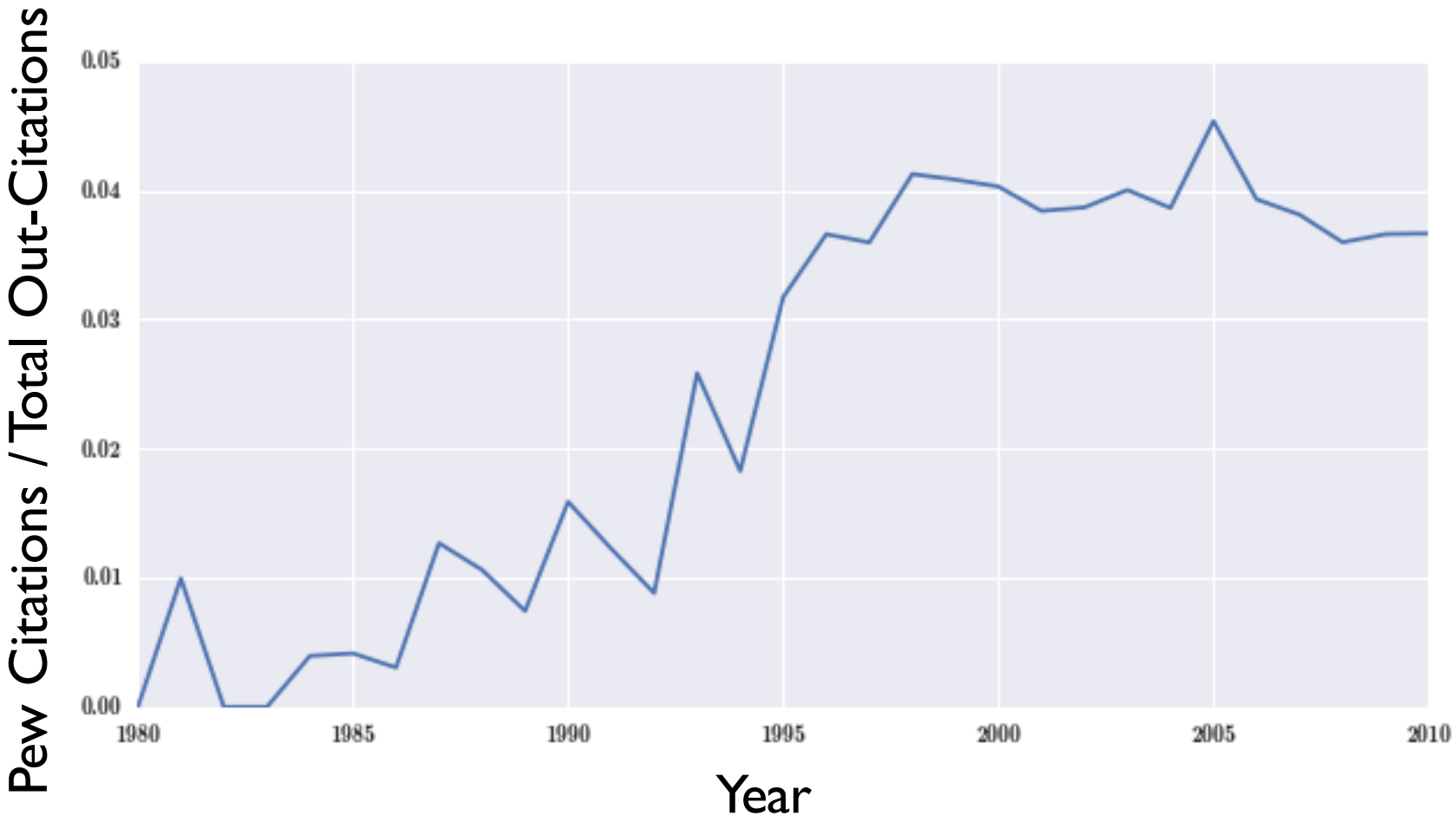
McCrary (2008)[10] density test on data from Lee, Moretti, and Butler (2004).

The impact of research grant funding on scientific productivity



Jacob, Brian A., and Lars Lefgren. (2011) Journal of Public Economics

Community Effect



Leading Edge
Commentary

Cell

Capturing the Value of Biomedical Research

Stefano Bertuzzi[2,3] and Zeina Jamaleddine[1,*]
[1]Sidra Medical and Research Center, Out-Patient Clinic, PO Box 26999, Al Luqta Street, Doha, Qatar
[2]American Society for Cell Biology, 8120 Woodmont Avenue, Suite 750 Bethesda, MD 20814-2762, USA
[3]Present address: American Society for Microbiology, 1752 North Street, N.W. Washington, DC 20036-2904, USA
*Correspondence: zjamaleddine@sidra.org
http://dx.doi.org/10.1016/j.cell.2016.03.004

Assessing the real-world impact of biomedical research is notoriously difficult. Here, we present the framework for building a prospective science-centered information system from scratch that has been afforded by the Sidra Medical and Research Center in Qatar. This experiment is part of the global conversation on maximizing returns on research investment.

The Complex Case of Assessing the Value of Research Investments

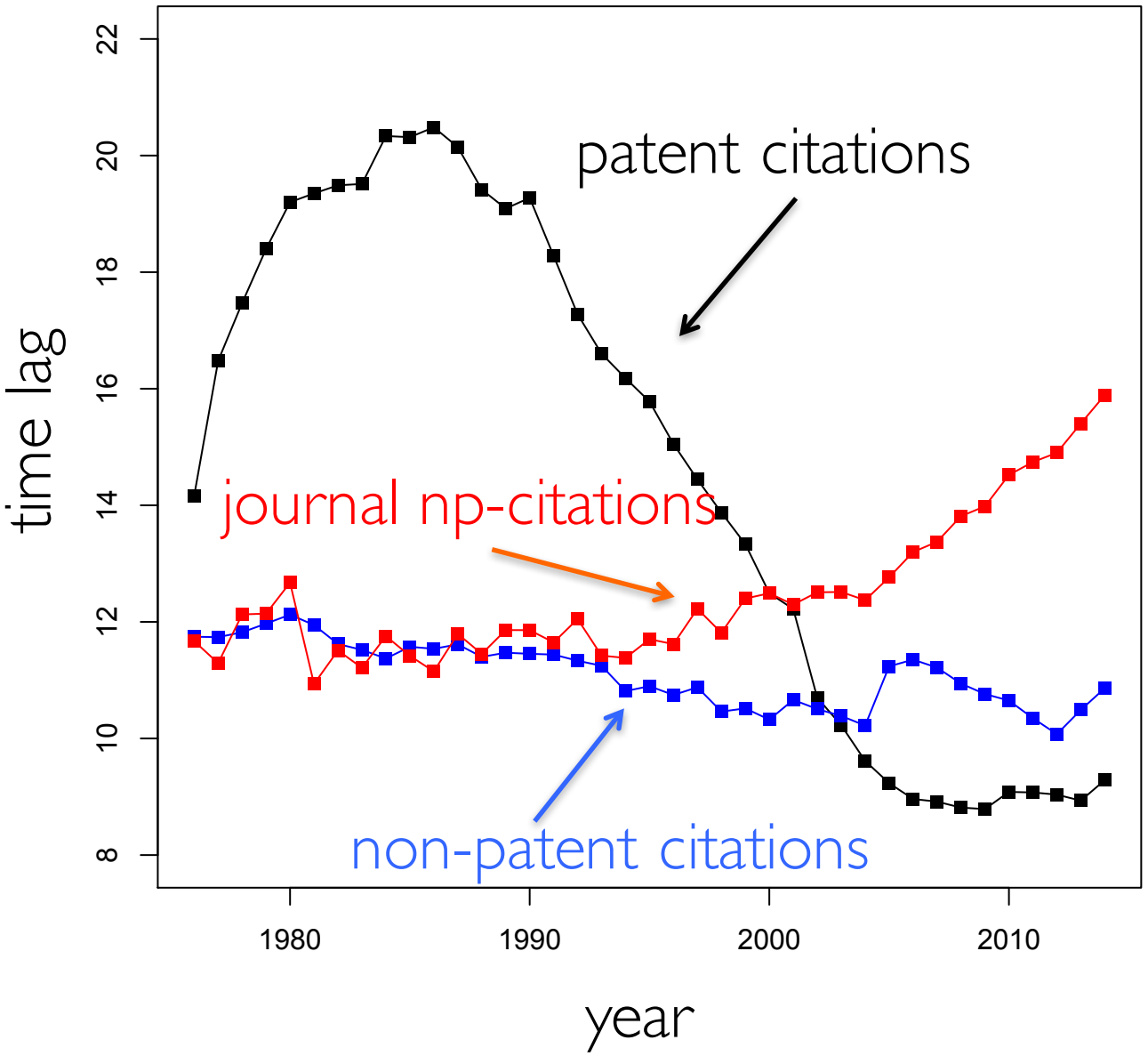
Assessing the impact of research funding on scientific, economic, and health outcomes is a complex endeavor, and quantifying it is made even more difficult by the unpredictable nature of basic science. The question of what society receives back in improved health, cost effectiveness, or cost savings from investments in research seems simple at first glance, but it often leads to an analytical quagmire. It is difficult to measure relevant parameters in a holistic manner that provides meaningful answers beyond mere anecdotes. Too often, analysts fall back on appealing summary statistics, some of which, unfortunately, are based on flawed metrics for assessing scientific productivity (Coleman, 1999; Lane, 2009; the systems currently in place to track science funding were developed primarily for administrative purposes. Relevant data needed to understand the effects of science funding often reside in different and disparate systems that are not interoperable, even within the same institution. Most importantly, these systems were not designed to capture scientific content and outcomes. This is evident in the fact that nearly all of them use grants as the fundamental unit of analysis. This is the wrong predicate for understanding innovation since the driver is the scientist, while the grant is simply the vehicle. In blunt terms, you can have a big research grant that goes nowhere; you can have an innovator on a tiny grant who upends an entire field. Grant-based analysis

Better tools that are specifically designed to assess the impact of research must be in place (Macilwain, 2010). This was our starting point. We believe this could be achieved by constructing a tracking system for a nation or region's universities, hospitals, funding agencies, regulatory agencies, intellectual property, advocacy, industry, public policies, and strategic plans. Moreover, we could keep track of deliverables in programs that support talented individuals and foster their free creativity through appropriate systems that allow them to easily network and access relevant and useful opportunities.

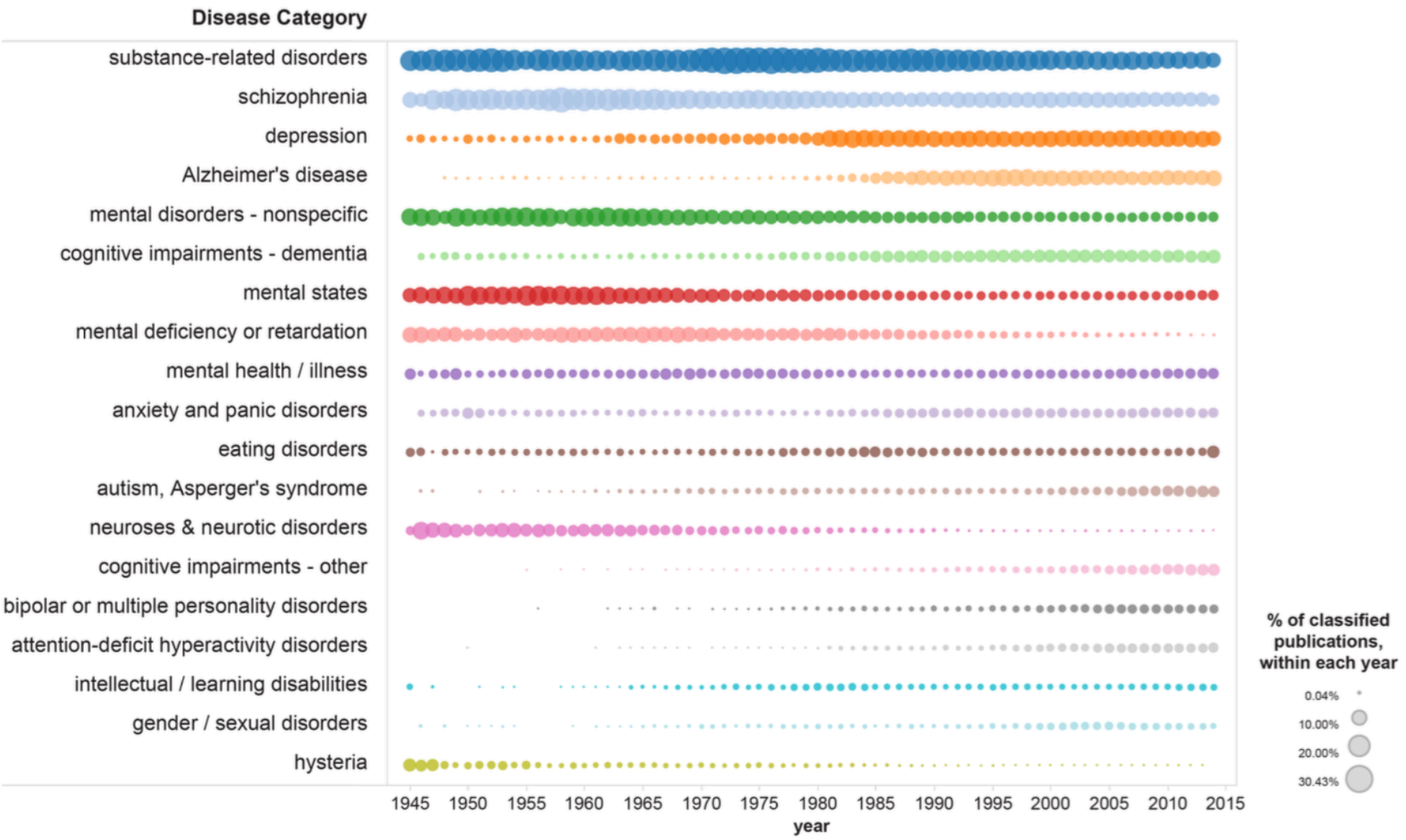
Why Qatar?

Qatar, a small Arabian Gulf country undergoing rapid technological and economic

Translational Lags

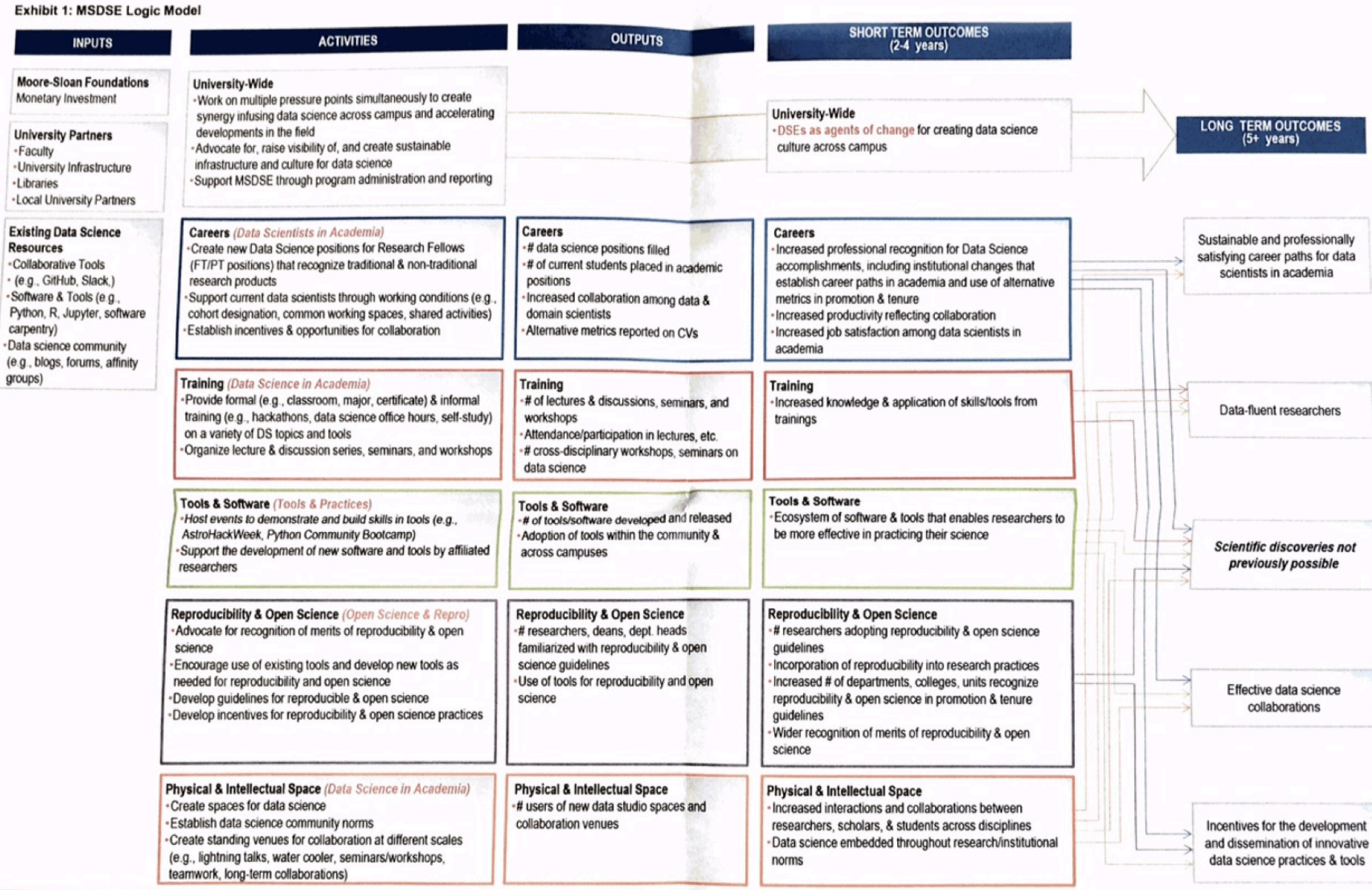


PubMed Publications by Disease Category



Kouper et al (2016) iConference Proceedings

Qualitative Methods

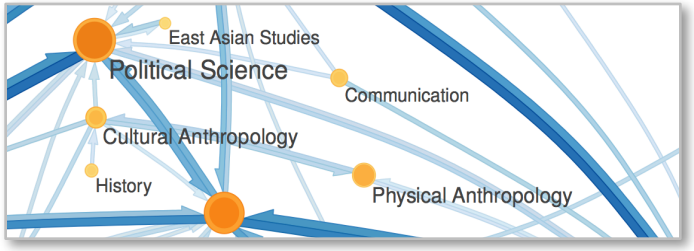


Qualitative Methods

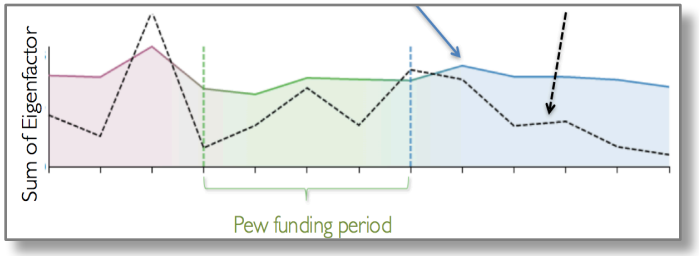
Jody Roberts

CHF

Talk Outline

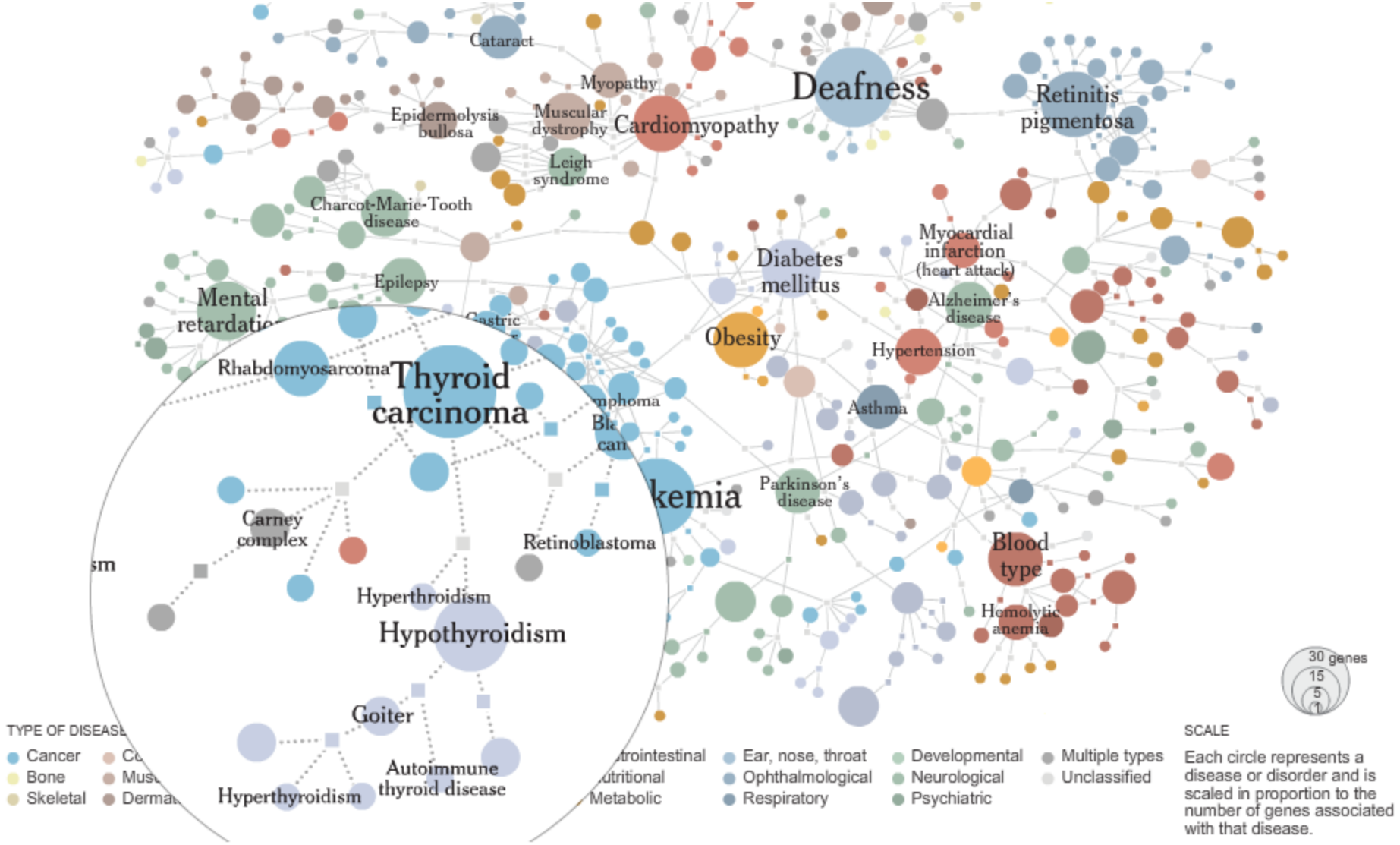


Visualizing scholarly influence



Measuring funding impact

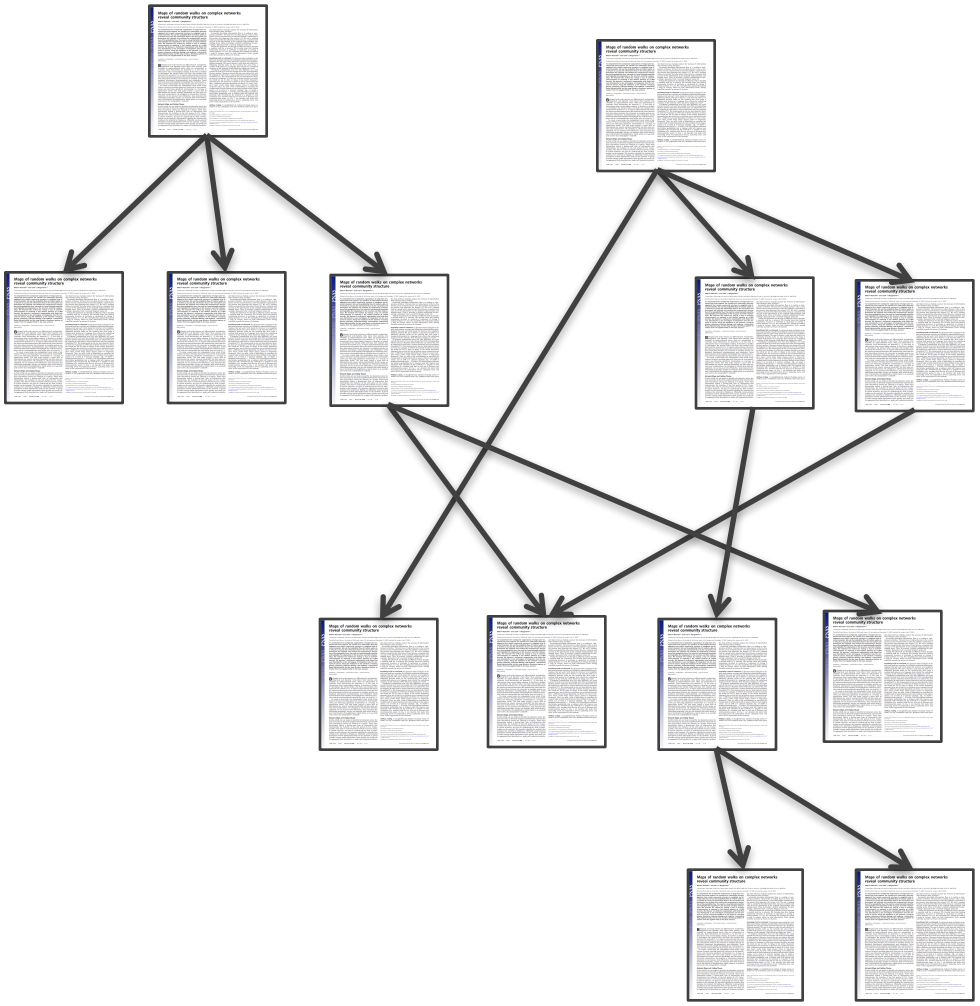
Disease association network



Sources: Marc Vidal; Albert-Laszlo Barabasi; Michael Cusick; Proceedings of the National Academy of Sciences

The New York Times

Citations form a vast network



de Solla Price, *Science* (1965)

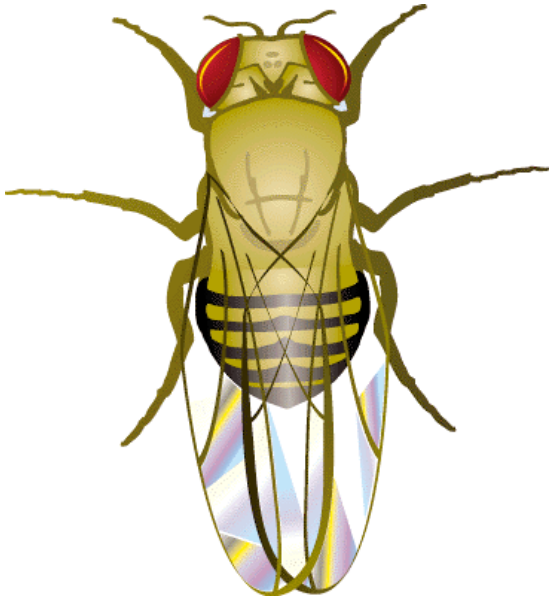
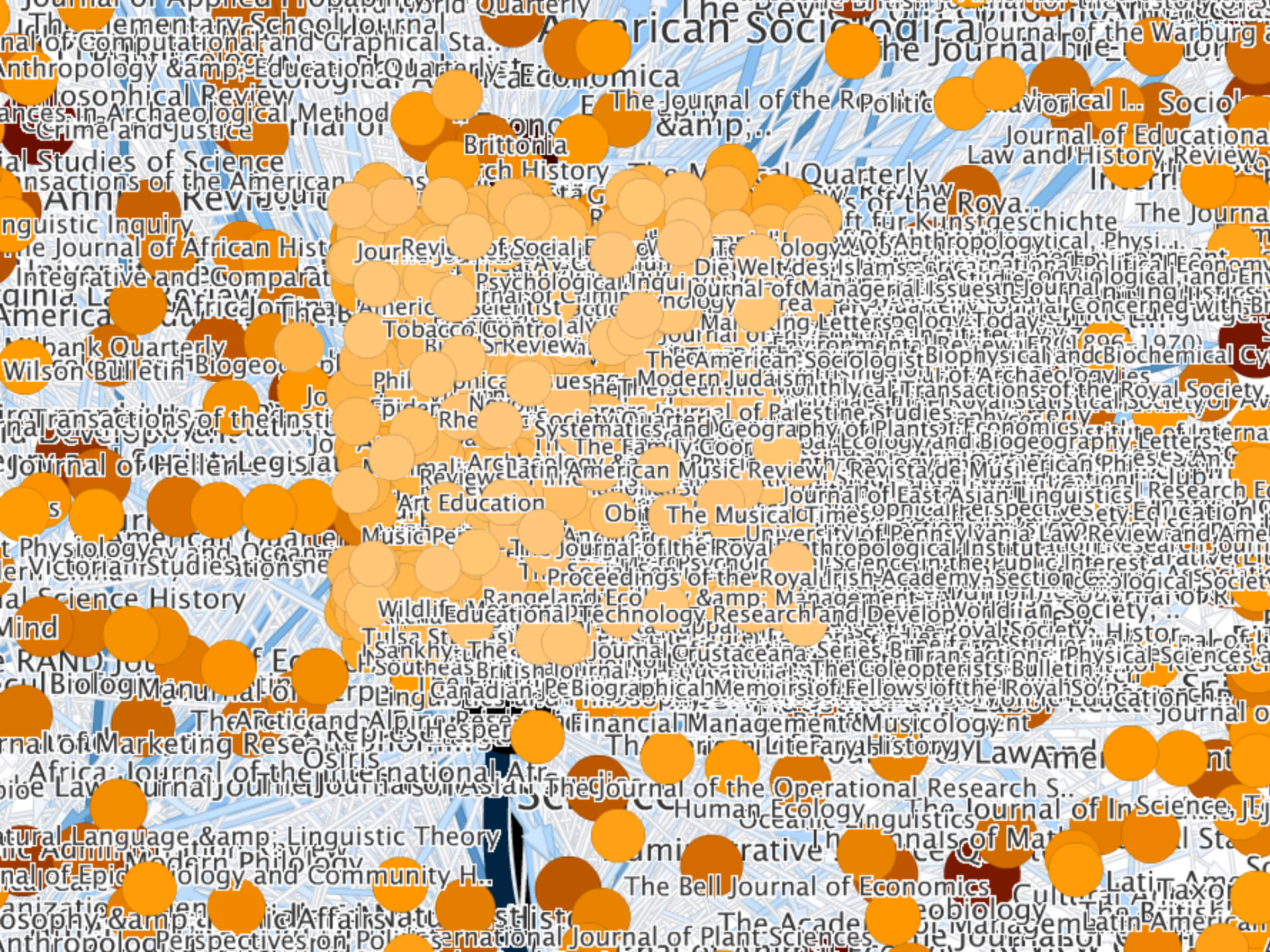


Image credits: Suite101.com: Bryant Lab UCI; U. Basel Plant Biol.; UW Arts and Sci.


The Elementary School Journal
Journal of Computational and Graphical Sta..
Anthropology & Education Quarterly
Ecological Economica
The Journal of the R
Crime and Justice
Britton
Journal of Educationa
Law and History Review
Annual Review
Linguistic Inquiry
The Journal of African Hist
Integrative and Comparat
Psychological Inqui
Journal of Managerial Issues
Tobacco Control
Wilson Bulletin
The American Sociologist
Modern Judaism
Transactions of the Institu
Systematics and Geography of Plants
Journal of Hellen
Latin American Music Review, Revista de Musi
Art Education
The Musical Times
Victorian Studies
Proceedings of the Royal Irish Academy. Section C: Geological Society
Science History
Wildlife Educational Technology Research and Develop
Mind
Journal of Marketing Rese
Osiris
Africa: Journal of the International Af
The Journal of the Operational Research S..
Human Ecology
Natural Language & Linguistic Theory
The Bell Journal of Economics
Perspectives on Po
International Journal of Plant Sciences

JW

Eigenfactor algorithm

$$P = \alpha H + (1 - \alpha) a.e^{T}$$

Matrix representing the random walk over citations

Probability of not teleporting

Cross-citation Matrix dictating the structure of the citation network

Probability of teleporting to completely new journal weighted by the number of articles in that journal

$$EF = 100 \frac{H\pi}{\sum_i [H\pi]_i}$$

Leading eigenvector of the random walk matrix P.

Normalization

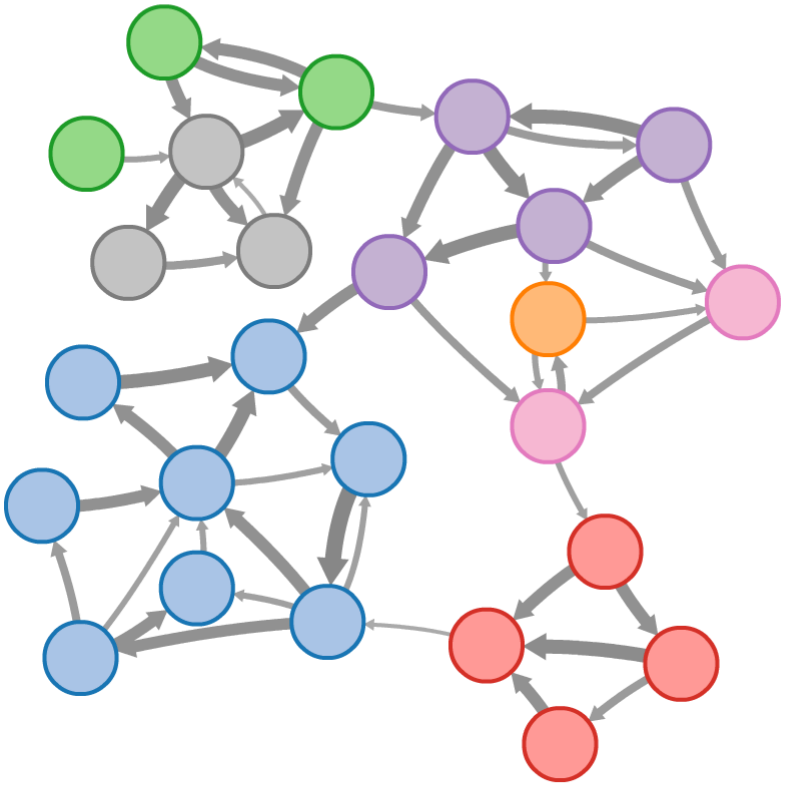
West, JD et al. (2010) *College of Research Libraries*

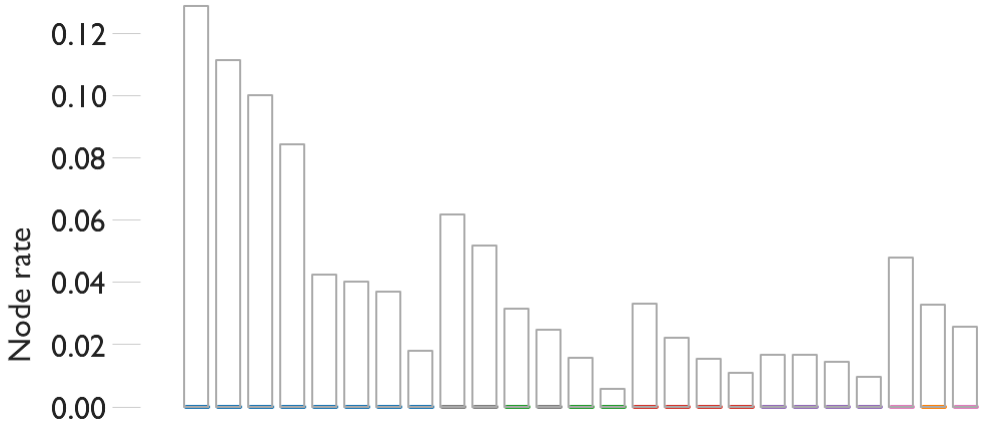
The map equation

$$L(\mathsf{M}) = q_{\curvearrowright} H(\mathcal{Q}) + \sum_{i=1}^{m} p^{i}_{\circlearrowright} H(\mathcal{P}^{i})$$

Rosvall and Bergstrom (2008) *PNAS*

Community Detection





Number of steps:

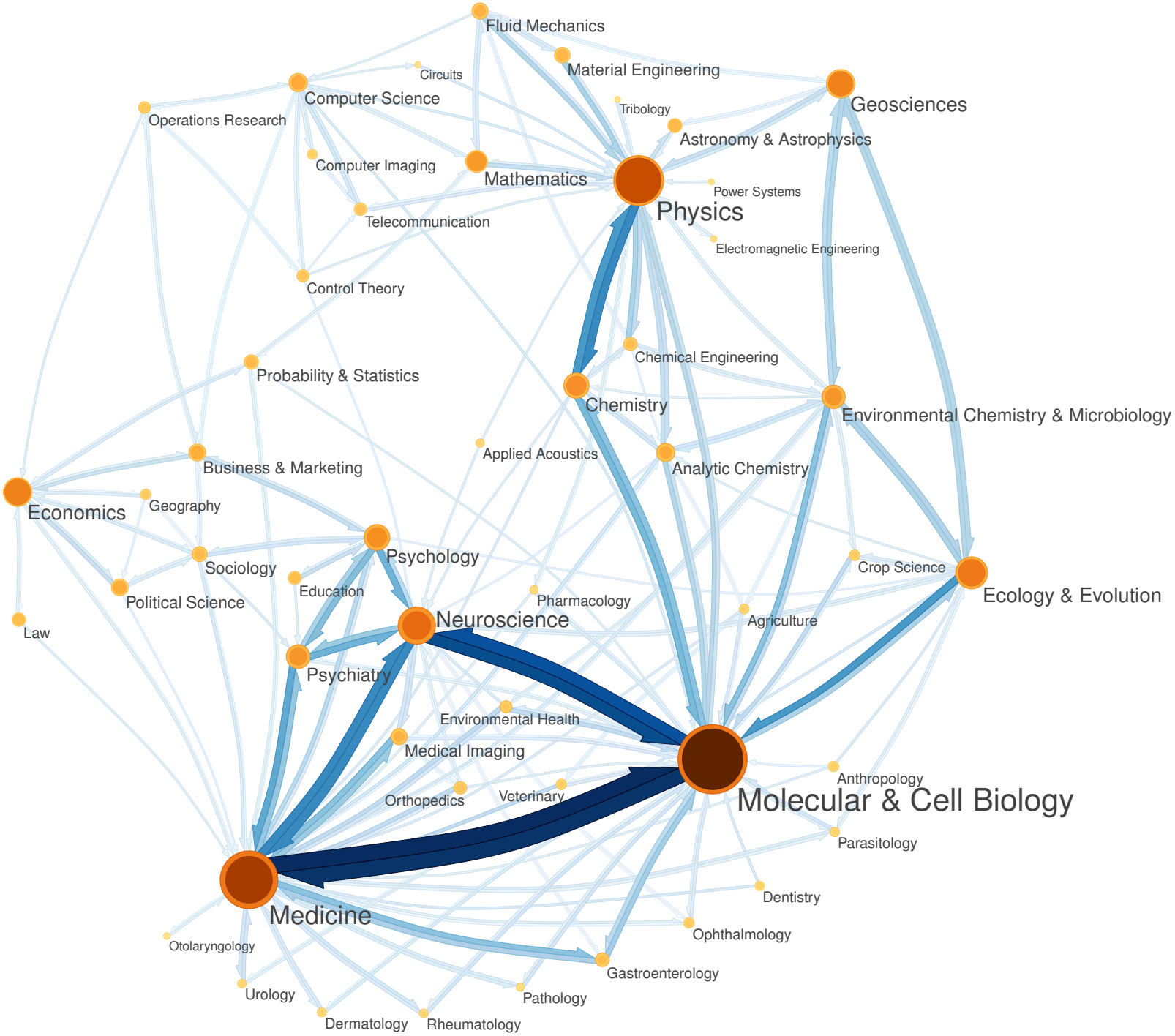
Iterative voting

Init votes | Vote | Automatic voting | Eigenfactor

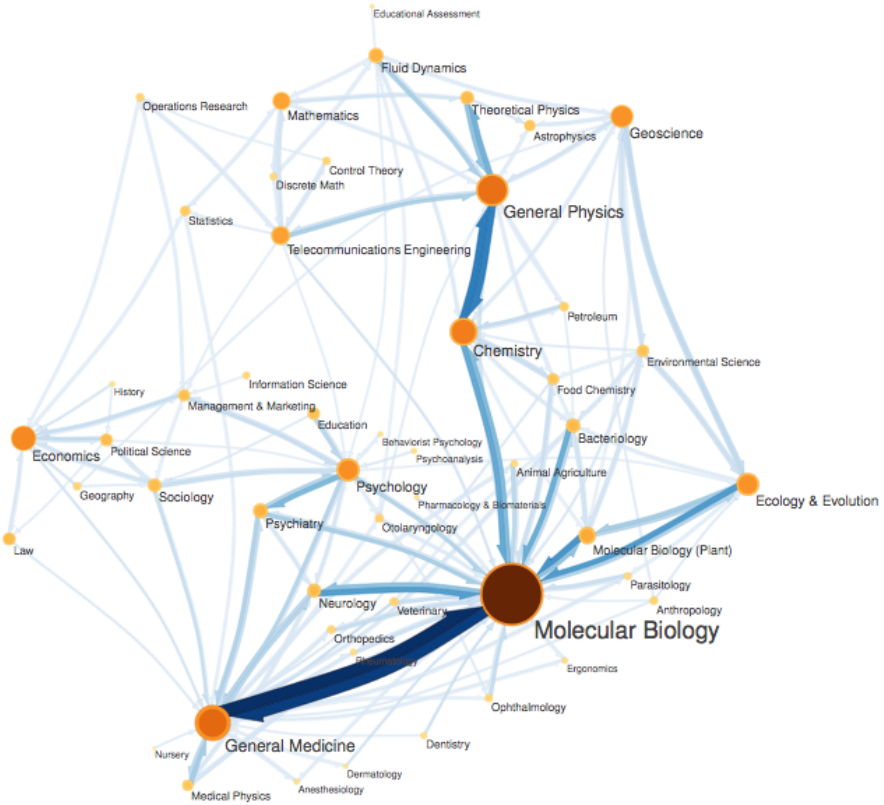
Speed

- [ ] Node size shows rate
- [ ] Link size shows flow
- [ ] Show origin of votes

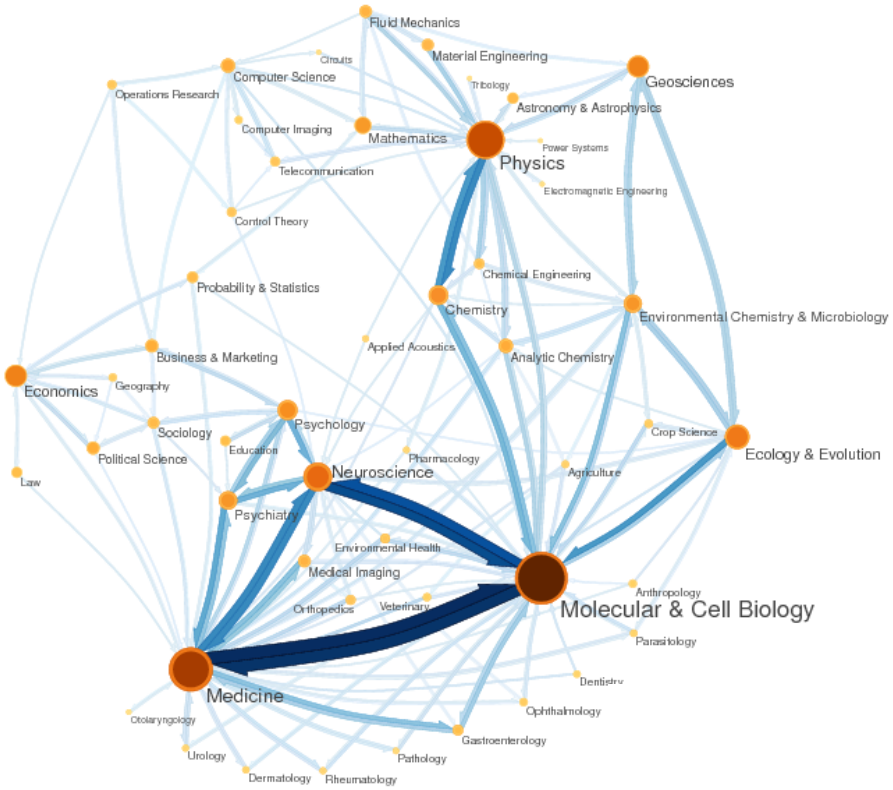
Mapequation.org, Daniel Edler


Fluid Mechanics
Material Engineering
Circuits
Computer Science
Geosciences
Operations Research
Tribology
Astronomy & Astrophysics
Computer Imaging
Mathematics
Power Systems
Physics
Telecommunication
Electromagnetic Engineering
Control Theory
Chemical Engineering
Probability & Statistics
Chemistry
Environmental Chemistry & Microbiology
Applied Acoustics
Business & Marketing
Analytic Chemistry
Economics
Geography
Psychology
Crop Science
Sociology
Education
Political Science
Pharmacology
Ecology & Evolution
Neuroscience
Law
Agriculture
Psychiatry
Environmental Health
Medical Imaging
Anthropology
Orthopedics
Veterinary
Molecular & Cell Biology
Parasitology
Dentistry
Medicine
Ophthalmology
Otolaryngology
Gastroenterology
Urology
Pathology
Dermatology
Rheumatology

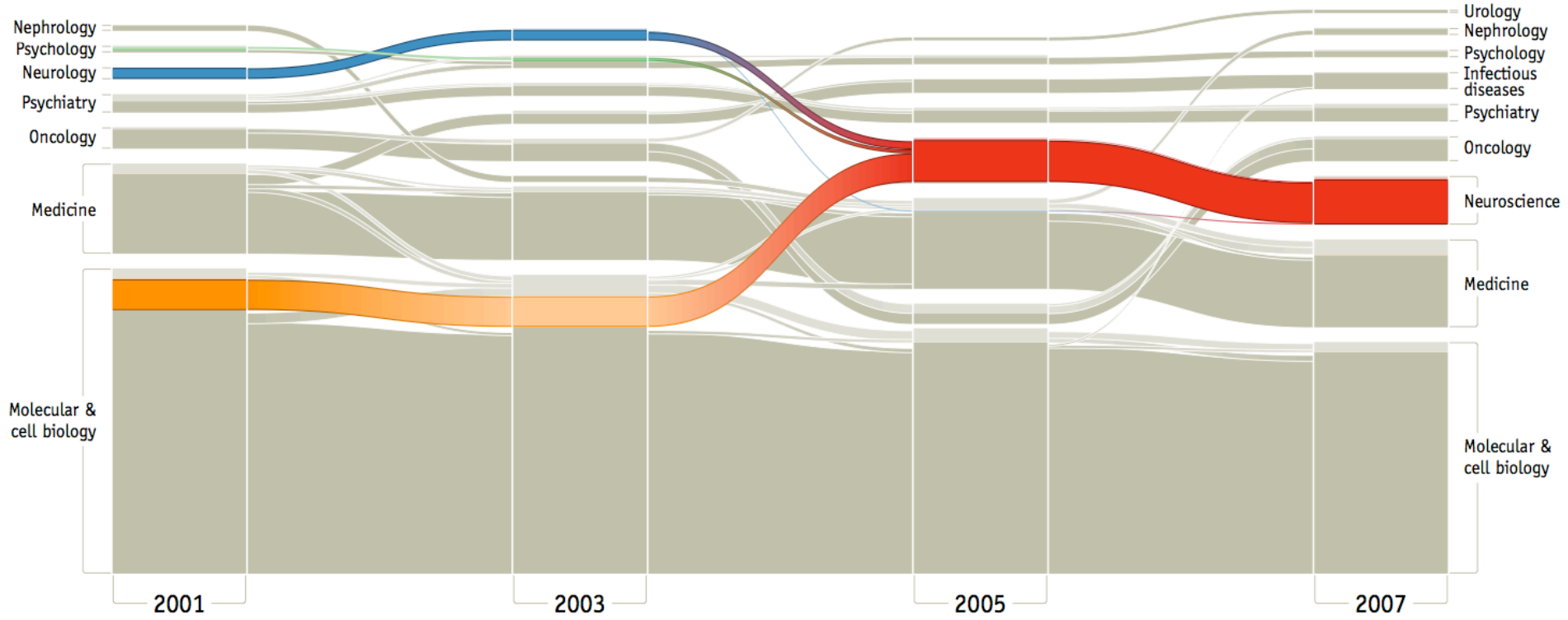
1995


Educational Assessment
Fluid Dynamics
Operations Research
Mathematics
Theoretical Physics
Astrophysics
Geoscience
Control Theory
Discrete Math
Statistics
General Physics
Telecommunications Engineering
Petroleum
Chemistry
Environmental Science
Information Science
History
Management & Marketing
Food Chemistry
Education
Bacteriology
Economics
Political Science
Behaviorist Psychology
Psychoanalysis
Animal Agriculture
Geography
Sociology
Psychology
Ecology & Evolution
Pharmacology & Biomaterials
Psychiatry
Otolaryngology
Law
Molecular Biology (Plant)
Parasitology
Neurology
Veterinary
Anthropology
Molecular Biology
Orthopedics
Ergonomics
Ophthalmology
General Medicine
Nursery
Dentistry
Dermatology
Medical Physics
Anesthesiology

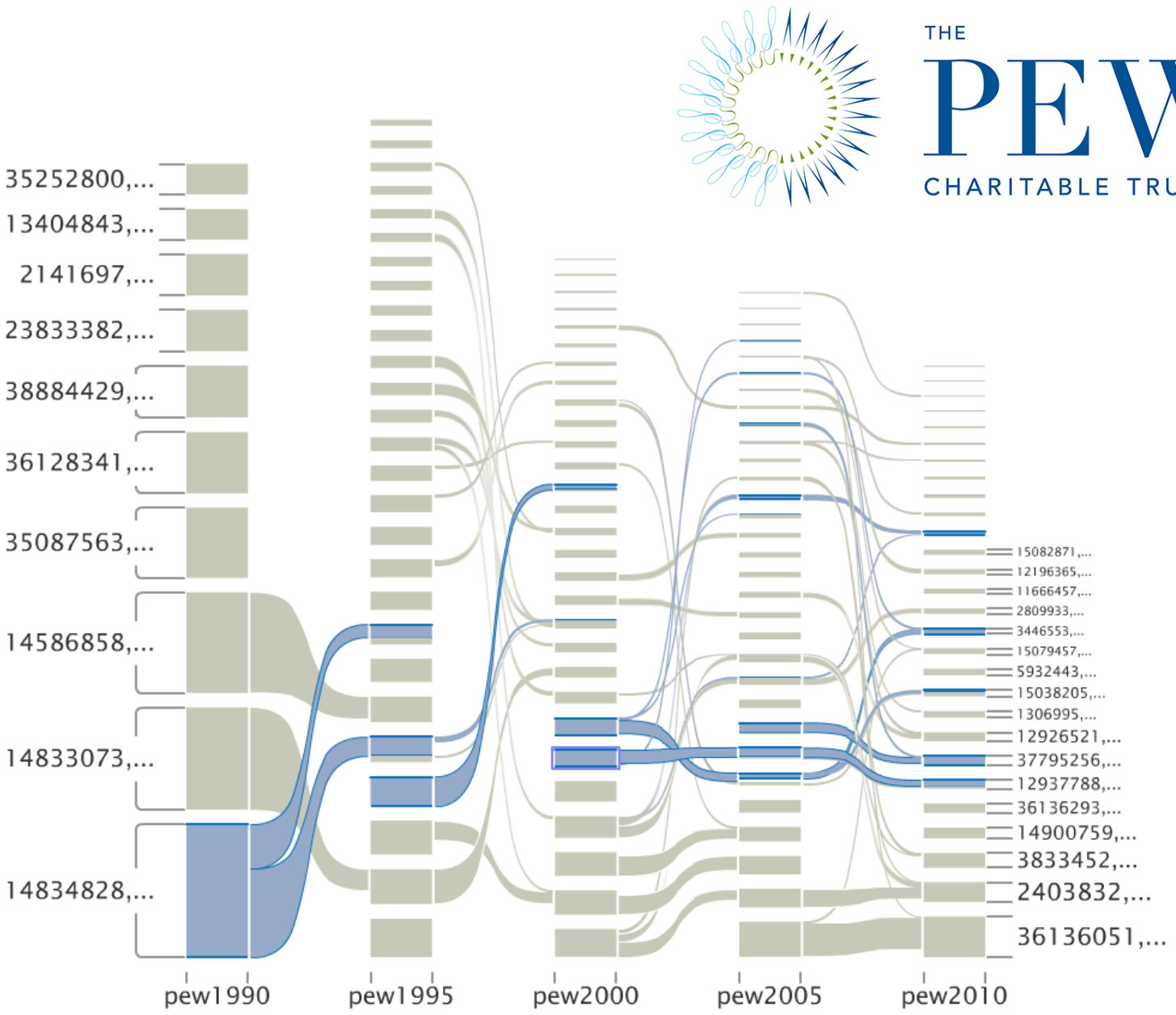
2004


Fluid Mechanics
Material Engineering
Circuits
Computer Science
Geosciences
Tribology
Operations Research
Astronomy & Astrophysics
Computer Imaging
Mathematics
Power Systems
Physics
Telecommunication
Electromagnetic Engineering
Control Theory
Chemical Engineering
Probability & Statistics
Chemistry
Environmental Chemistry & Microbiology
Applied Acoustics
Business & Marketing
Analytic Chemistry
Economics
Geography
Psychology
Crop Science
Sociology
Education
Pharmacology
Ecology & Evolution
Political Science
Neuroscience
Agriculture
Law
Psychiatry
Environmental Health
Medical Imaging
Anthropology
Orthopedics
Veterinary
Molecular & Cell Biology
Parasitology
Dentistry
Medicine
Ophthalmology
Otolaryngology
Gastroenterology
Urology
Pathology
Dermatology
Rheumatology

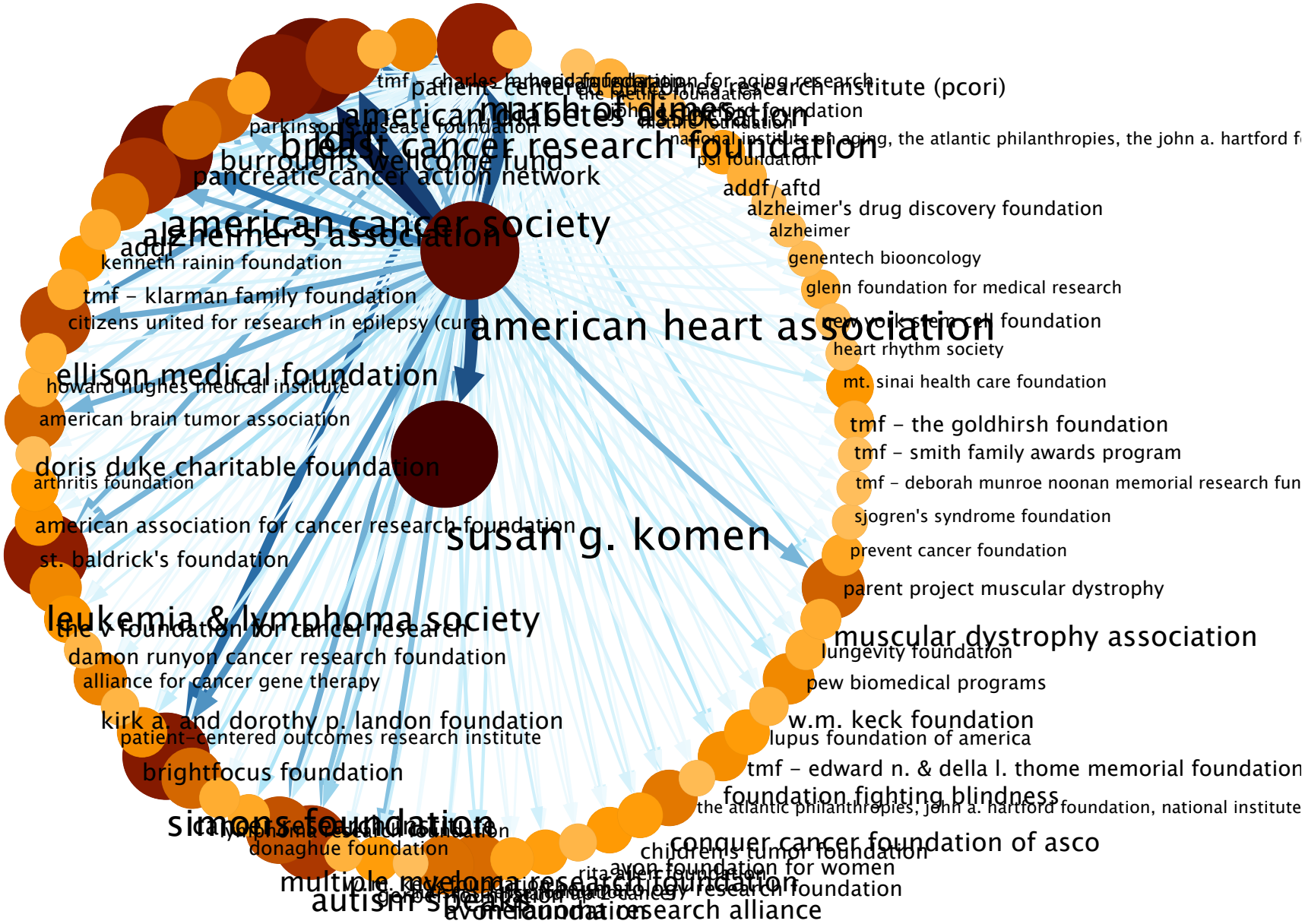
Citation networks over time



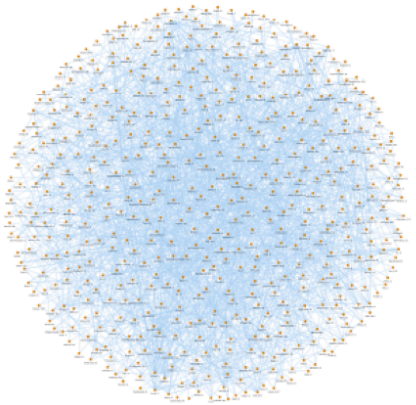
Rosvall & Bergstrom (2010) *PLoS One*

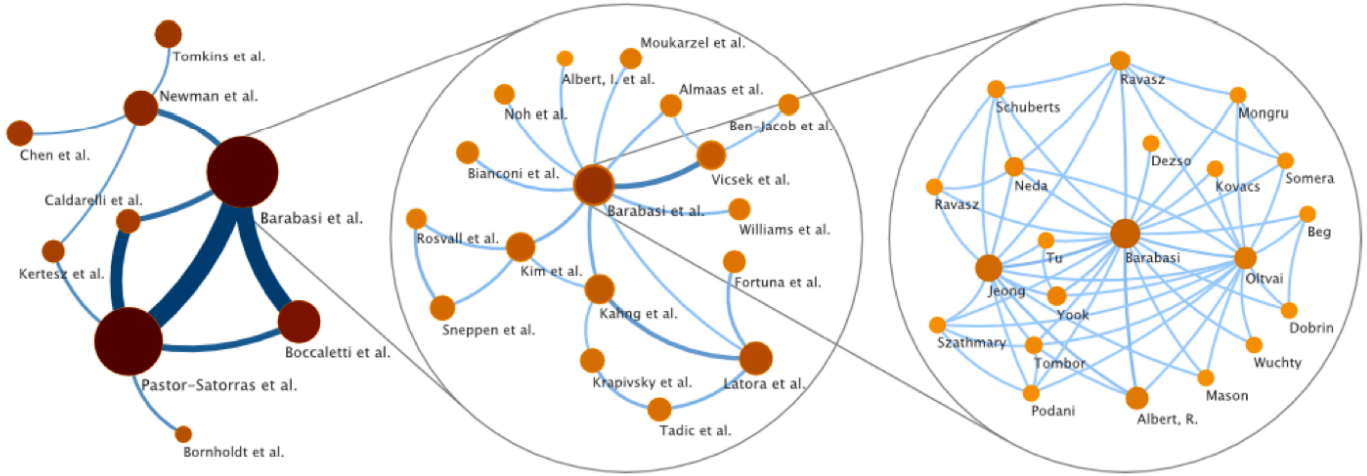

35252800,...
13404843,...
2141697,...
23833382,...
38884429,...
36128341,...
35087563,...
14586858,...
14833073,...
14834828,...
15082871,...
12196365,...
11666457,...
2809933,...
3446553,...
15079457,...
5932443,...
15038205,...
1306995,...
12926521,...
37795256,...
12937788,...
36136293,...
14900759,...
3833452,...
2403832,...
36136051,...
pew1990
pew1995
pew2000
pew2005
pew2010

Health Research Alliance



Hierarchical Mapping

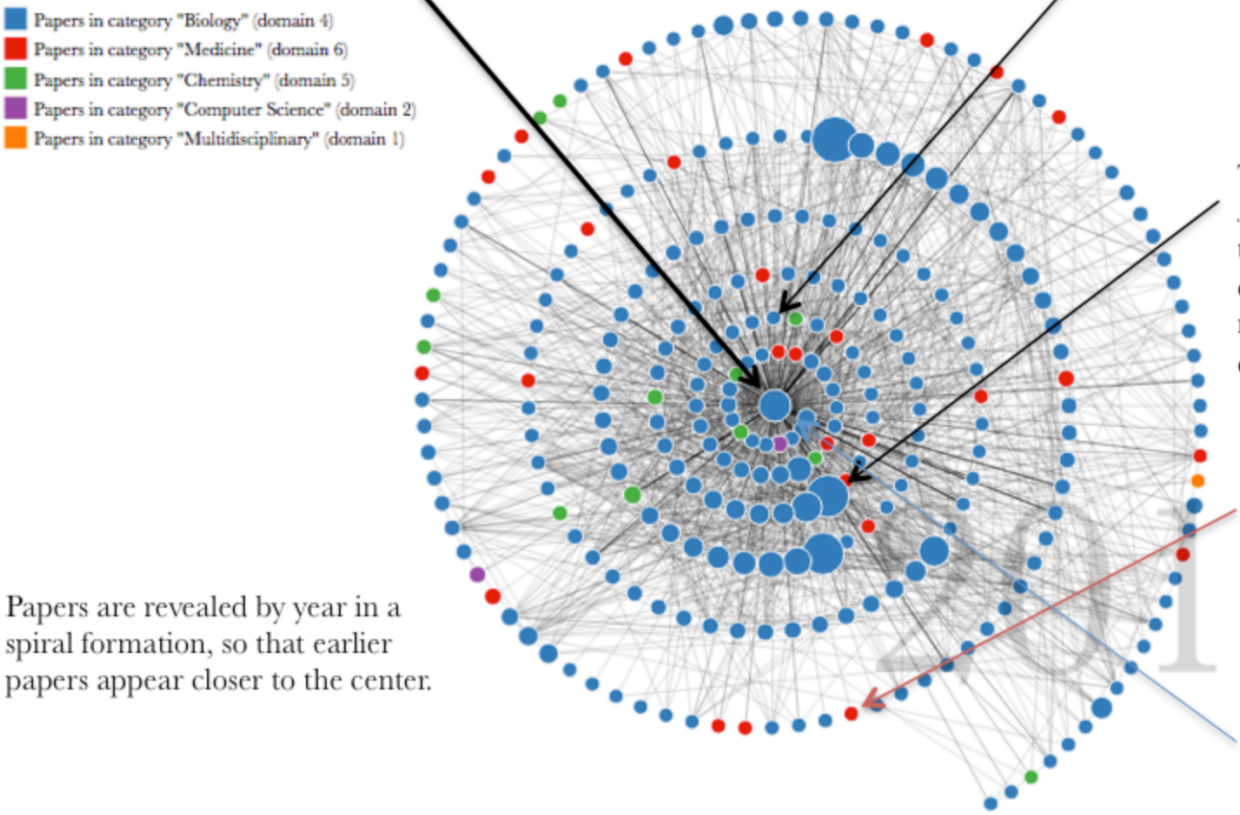




Pew Influence

The **center node** represents all of the papers authored by the scholar of interest.

Surrounding nodes represent papers that have cited work by the scholar of interest. Lines between the nodes show citations between papers.



Showing a scholar's influence

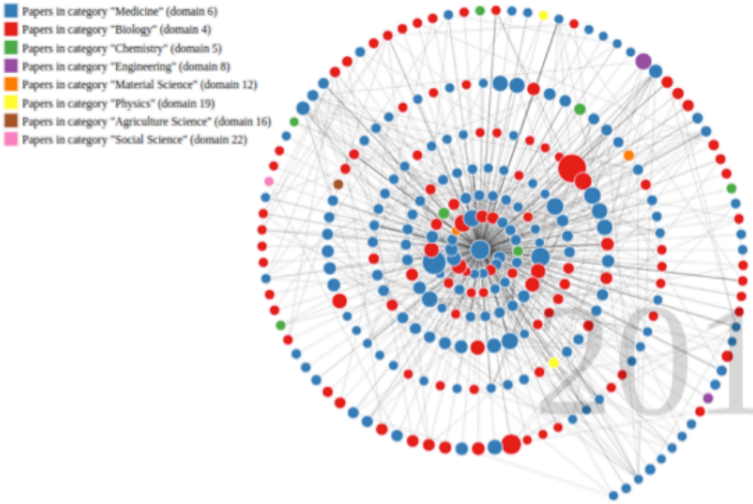
The size of each node is scaled by the *Eigenfactor score* of that paper—a metric of influence that takes into account its position in the total citation network. Bigger nodes represent the most influential papers that have cited the central scholar.

The color of each node shows the academic discipline of the paper. A more colorful network means that the impact of the central scholar's work has extended out to a wider range of fields.

Papers are revealed by year in a spiral formation, so that earlier papers appear closer to the center.

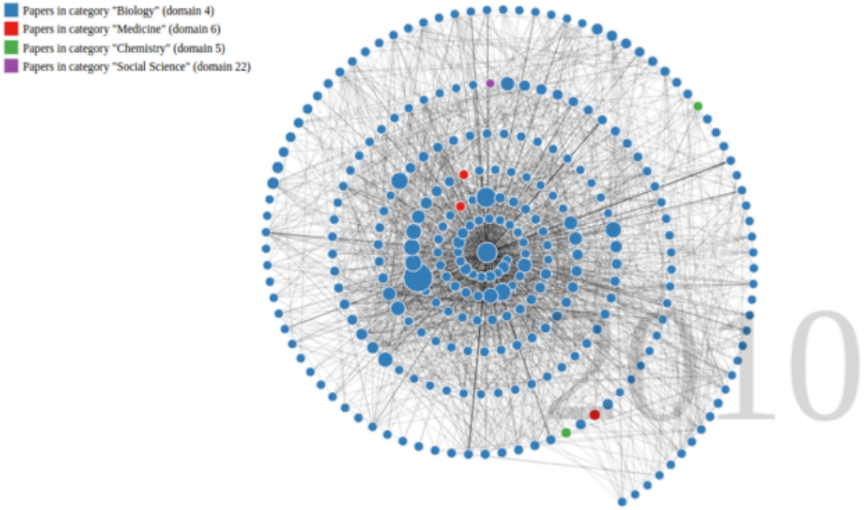
The color of the center node represents the dominant field of the central scholar—the most common field of all the scholar's publications.

Comparing Authors

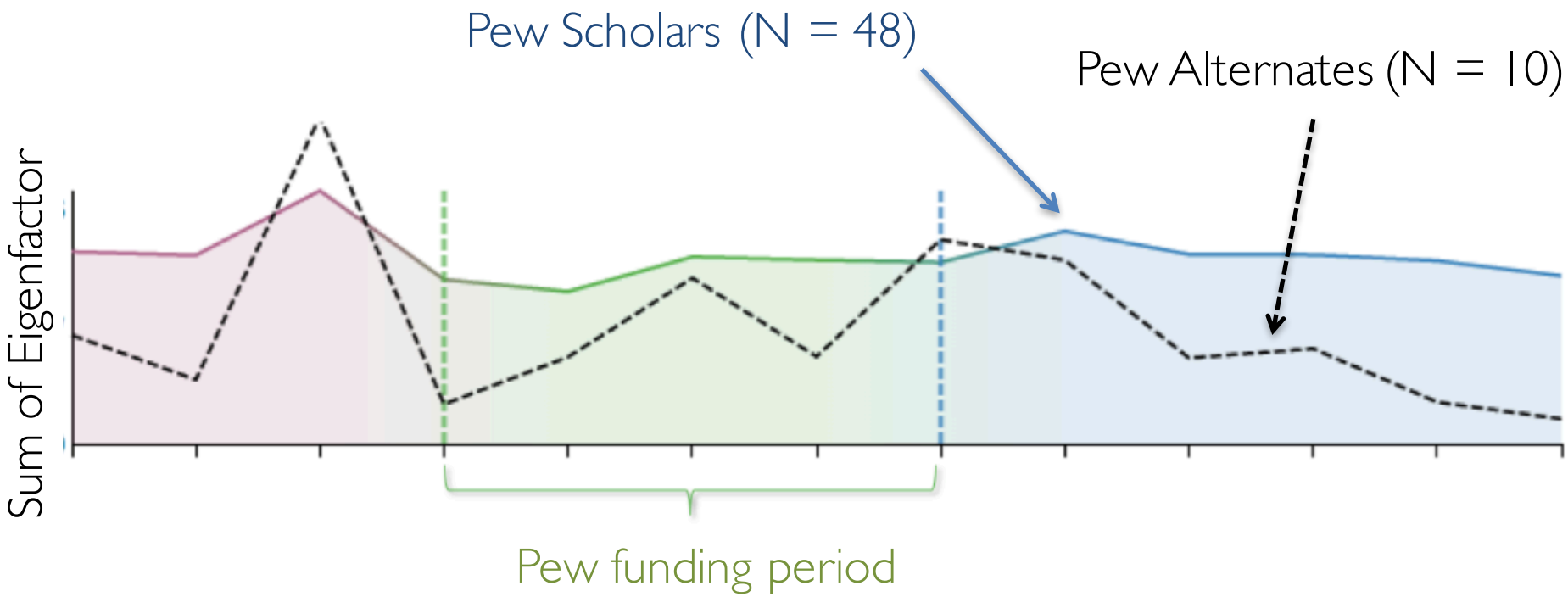


A more sparse network indicates fewer citations between papers shown in the network. This could be a result of the central scholar having impact across a wider set of academic communities.

A denser network means that the papers that cite the central author also tend to cite each other.



Comparing Alternates



* Includes scholars and alternates from cohort years: 1997, 1999, 2000, 2001, 2002

JW

Data I need from you…

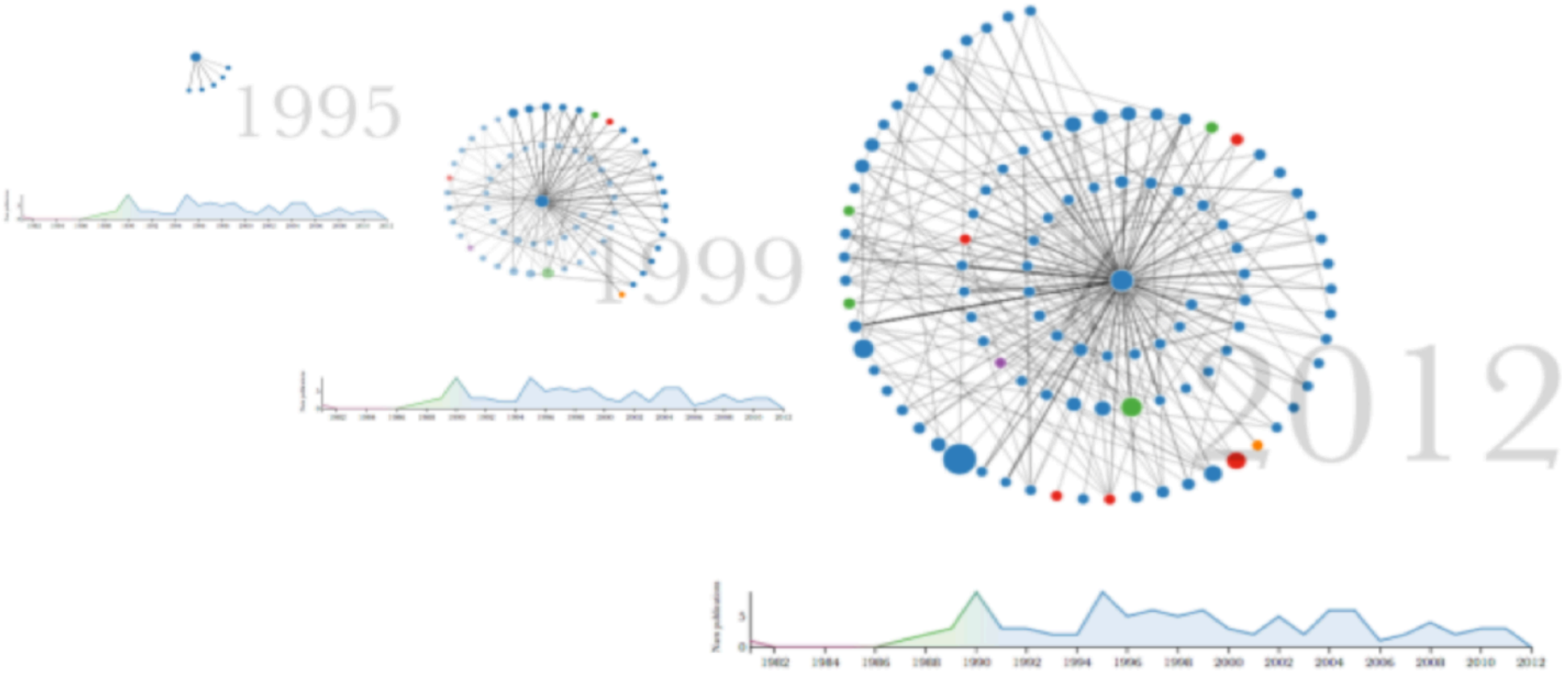
JW

Future Directions

- Regression discontinuity design analysis (pew scholars versus alternates)
- Integrate interviews from Chemical Heritage Foundation
- Automated narration of visualization
- Author disambiguation and further data cleaning
- User studies for improved hypothesis generation
- Personalize visualization for different stakeholders (funders, researchers, donors)
- Distinguish reviews, model organisms, funding agency
- Basic research versus applied research
- Individual grants versus collaborative grants

Explore the data

scholar.eigenfactor.org



* Please use Chrome web browser for best results

Acknowledgements

Jason Portenoy, Information School, University of Washington

Anita Pepper and Kara Coleman, Pew Charitable Trust

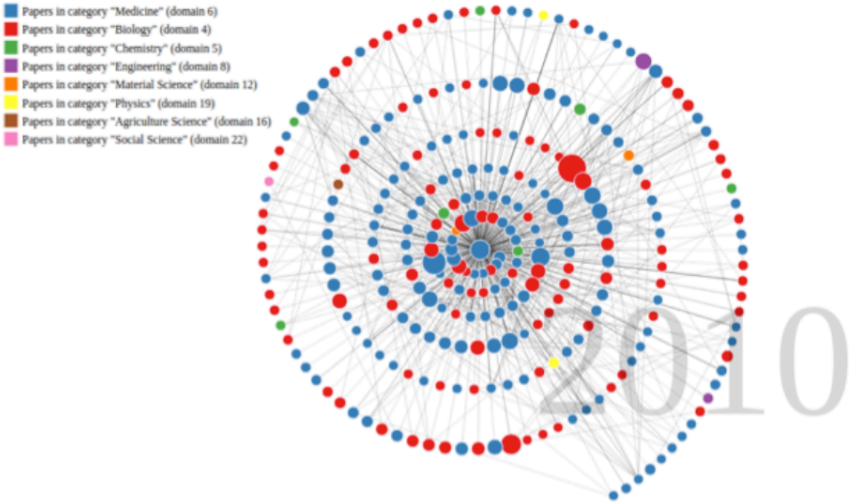
Jody Roberts , Chemical Heritage Foundation

Martin Rosvall, Department of Physics, Umea University

Carl Bergstrom, Department of Biology, University of Washington

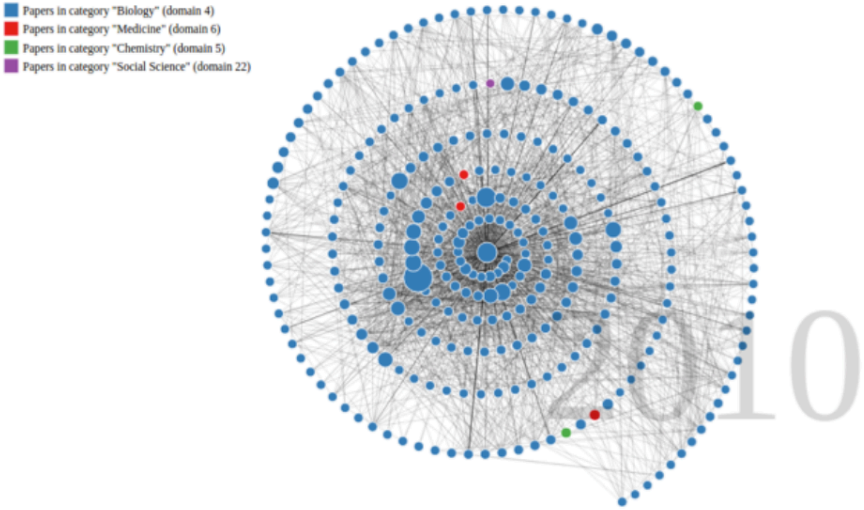
Visualizing Interdisciplinarity

Jason Portenoy



A more sparse network indicates fewer citations between papers shown in the network. This could be a result of the central scholar having impact across a wider set of academic communities.

A denser network means that the papers that cite the central author also tend to cite each other.



scholar.eigenfactor.org