

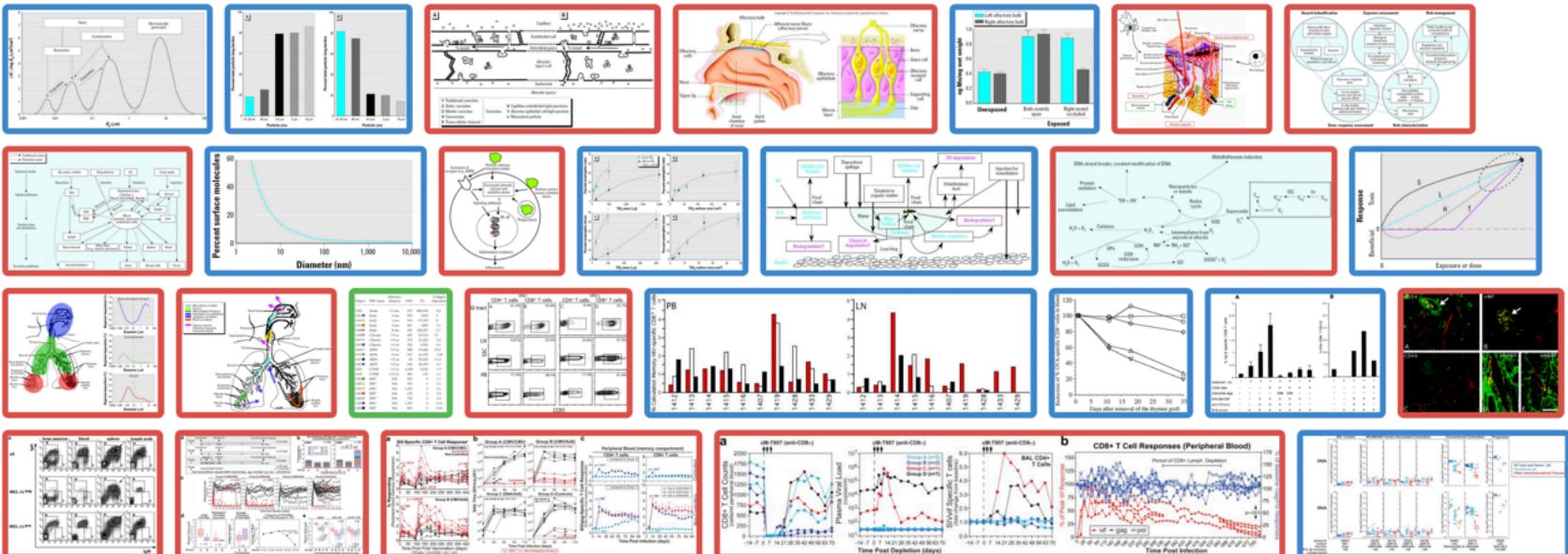
Figure-Centric Search Engine

 viziometrics.org

VizioMetrix About **Search** Crowdsourcing

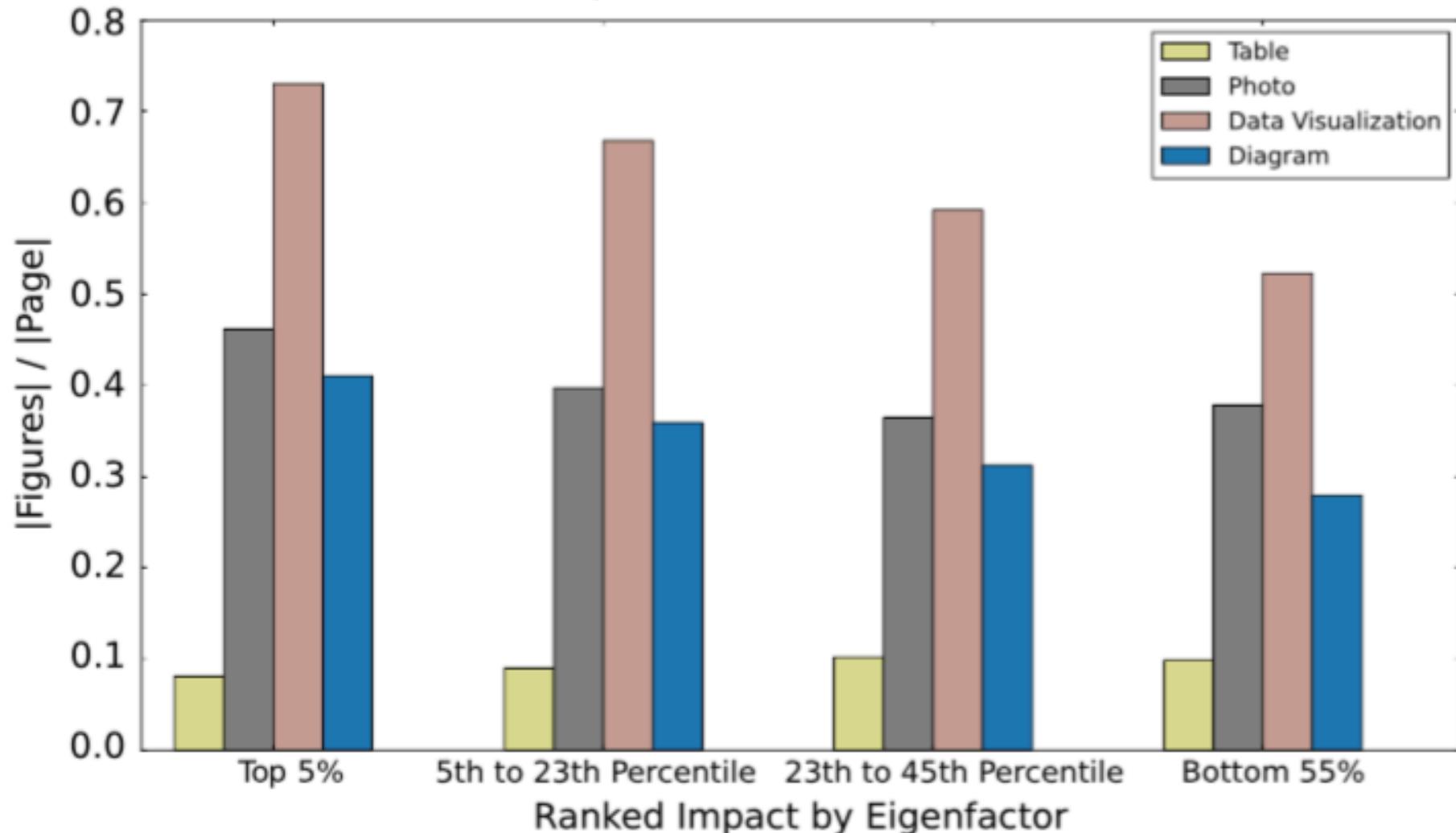
Impact blood lymph

Composite Equation Diagram Photo Plot Table

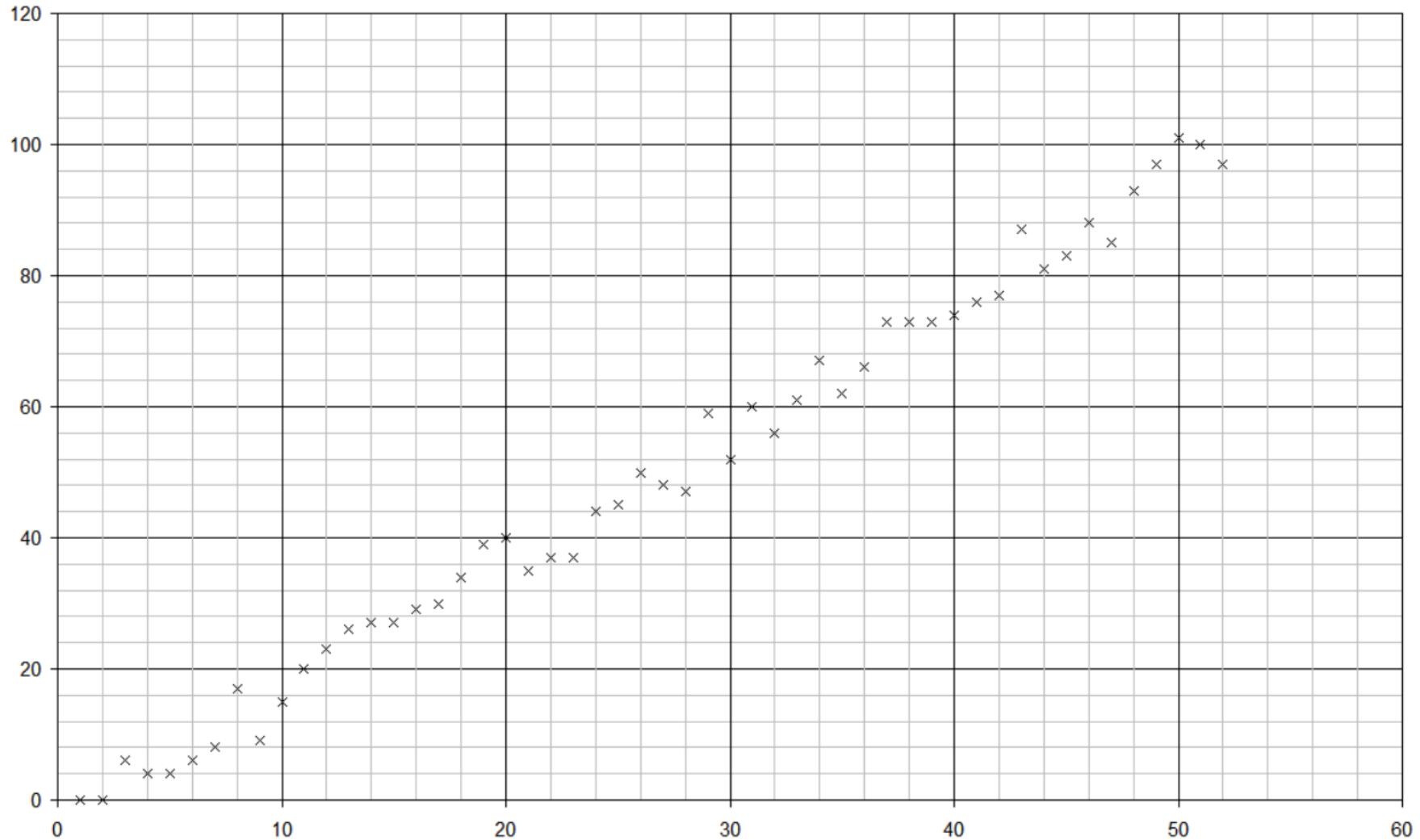


A project of the eScience Institute at the University of Washington

Impact versus Figure Density



Gridlines



New cases each week

Liberia

1

0

480

640

Sierra Leone

9

203

Guinea

2

9

Dec. 2013

May. 2015

Conservative

Labour

1923

1929

1974 (Feb.)

2010

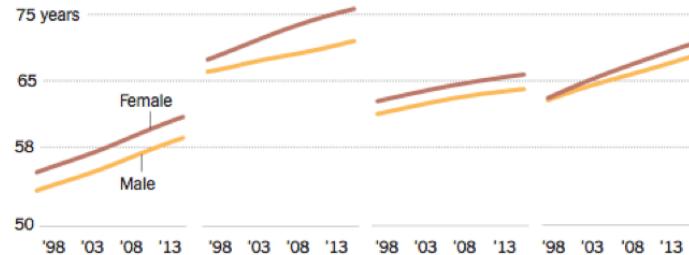
50
percent
of seats

1918 '22 '23 '24 '29 '31 '35 '45 '50 '51 '55 '59 '64 '66 '70 '74 '79 '83 '87 '92 '97 2001 '05 '10
Feb. Oct.

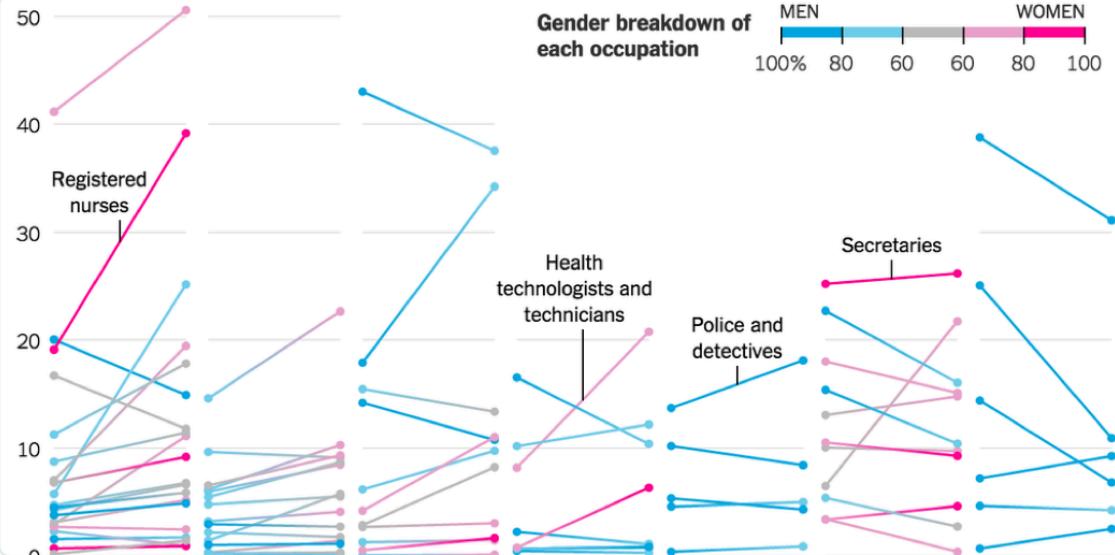
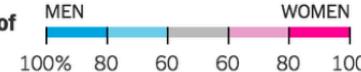
Acknowledgement: slides and discussions around gridlines with Carl Bergstrom

Life Expectancy

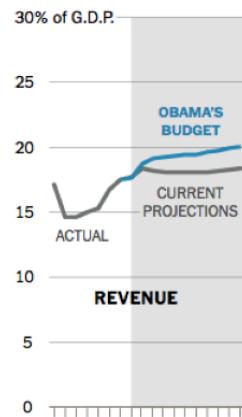
Worldwide, women are generally expected to live longer than men. This is true in Afghanistan as well. But life expectancy in Afghanistan is still well behind other countries in the region.



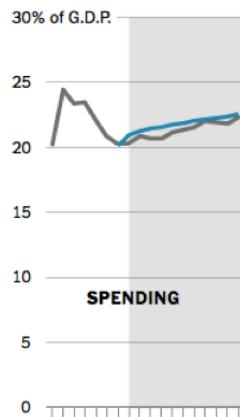
Gender breakdown of each occupation



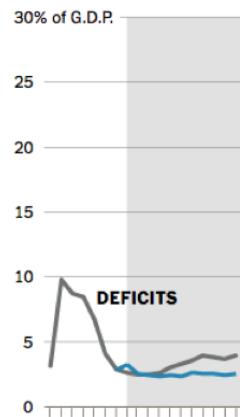
REVENUE



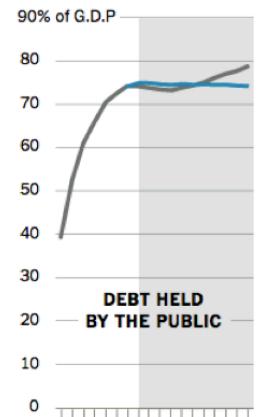
SPENDING



DEFICITS



DEBT HELD BY THE PUBLIC

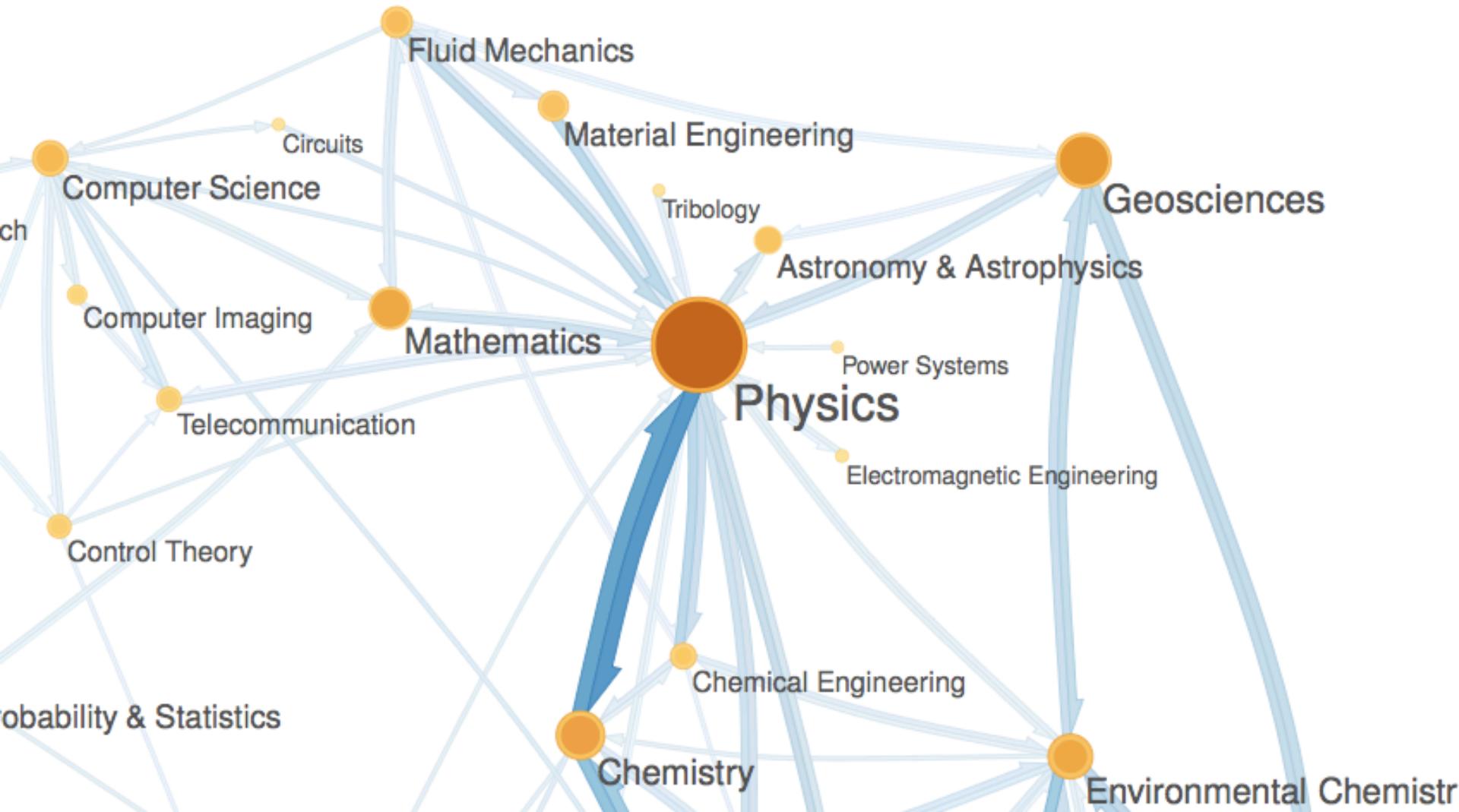


Gridline Rules

- 1) Gridlines should always be labelled.
- 2) Corollary: One should not use grid lines on any axis that lacks a scale.
- 3) If one does use grid lines, they should be ONLY be used to demarcate increments of the dependent variable.
- 4) Grid lines should be subtle elements of the background, and not draw the eye from the foreground elements.

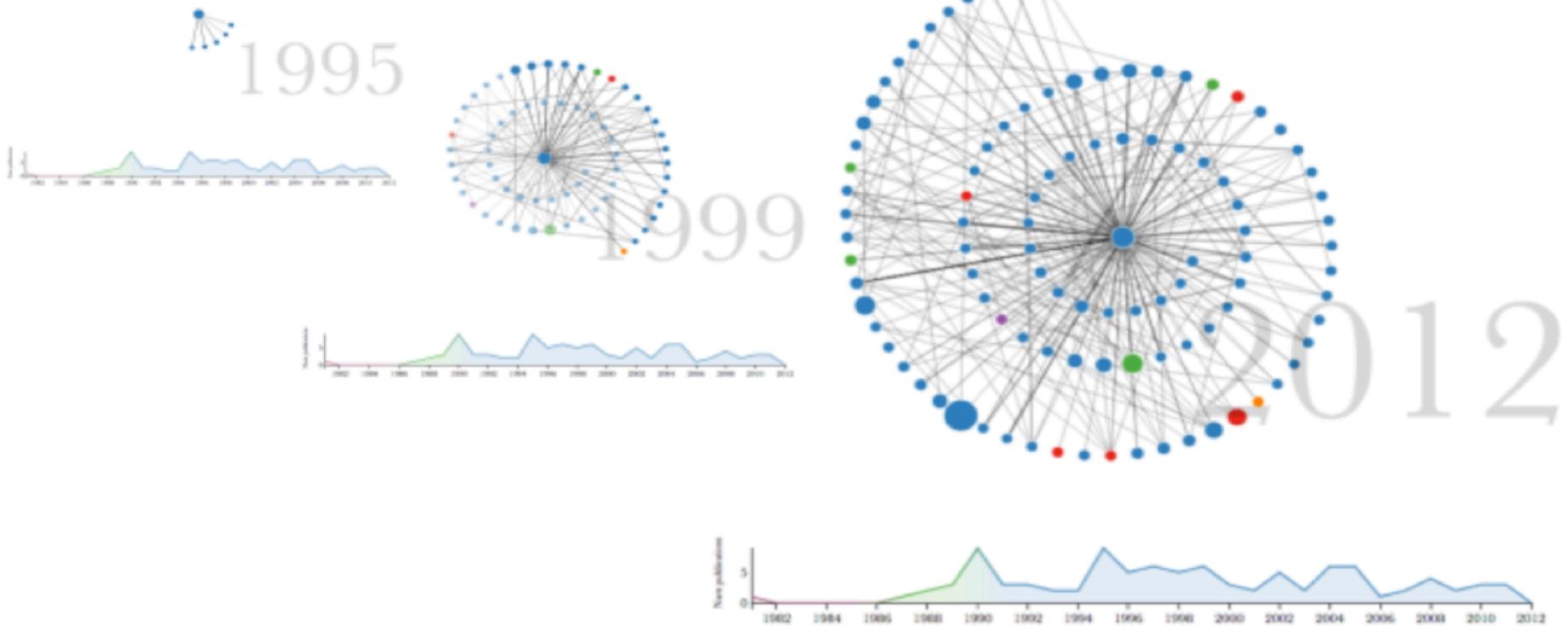
Mapping the AHA Community

Jevin West, Information School, University of Washington



Explore the data

scholar.eigenfactor.org



* Please use Chrome web browser for best results

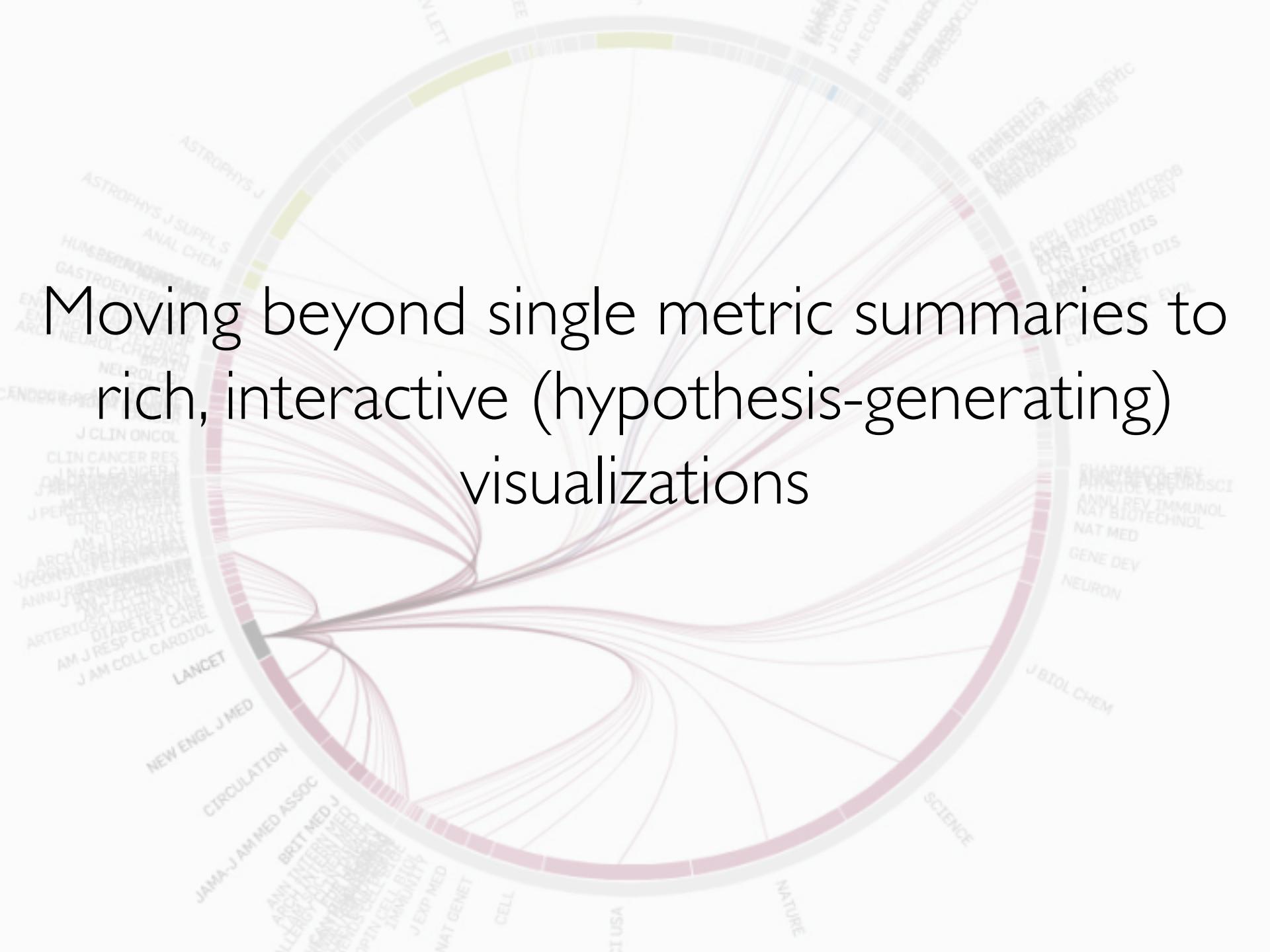
scholar.eigenfactor.org

username: PewScholar

password: 1N!kdG

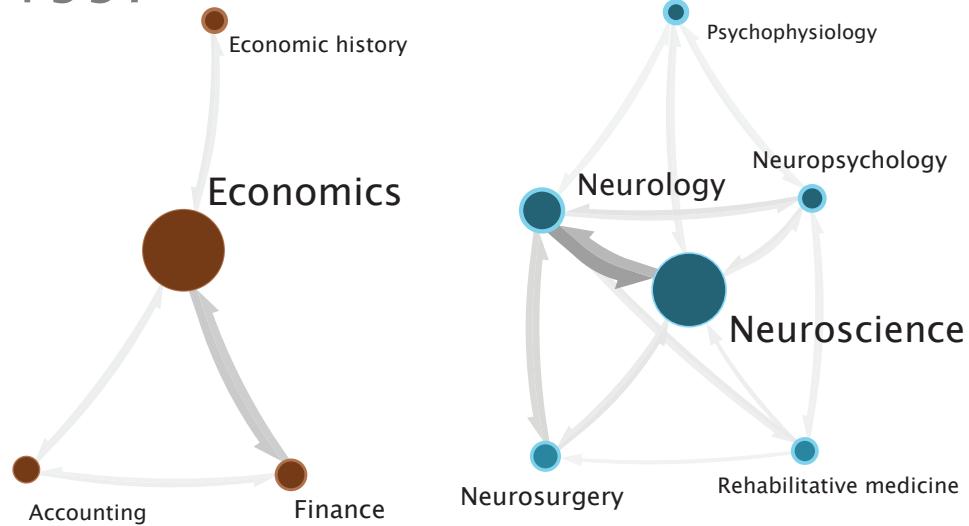
Jevin West, jevinw@uw.edu

Moving beyond single metric summaries to rich, interactive (hypothesis-generating) visualizations

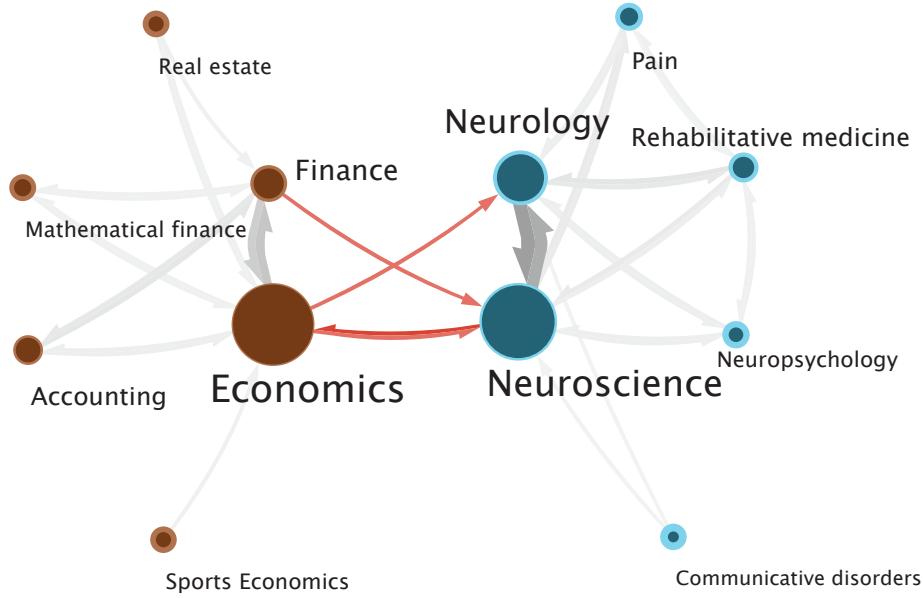




1997



2010



NEUROECONOMICS

How do we *map* the evolution of scientific disciplines?

What is my impact on science?

\$7,933,670,366





22,756 awards

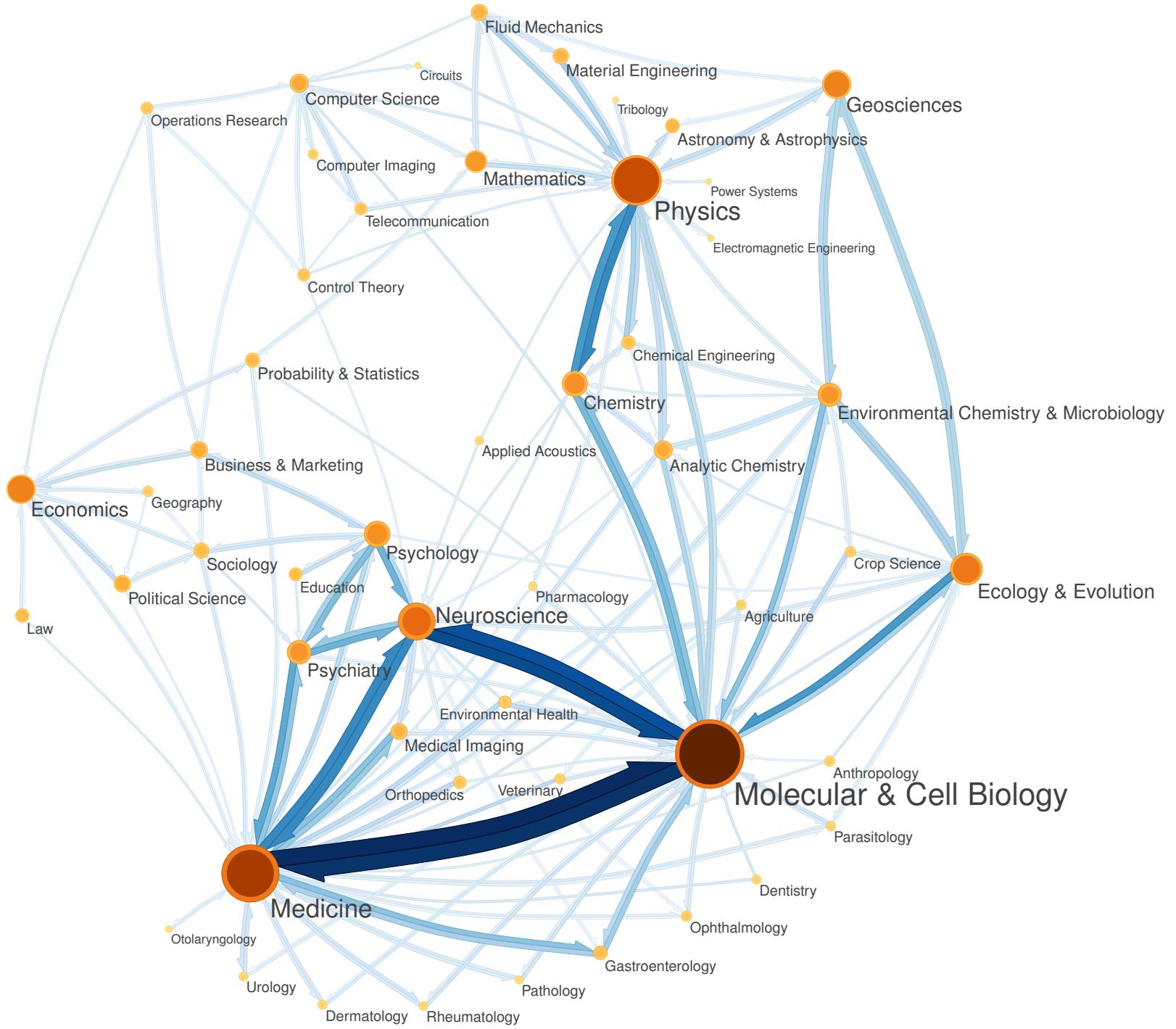
17,849 researchers

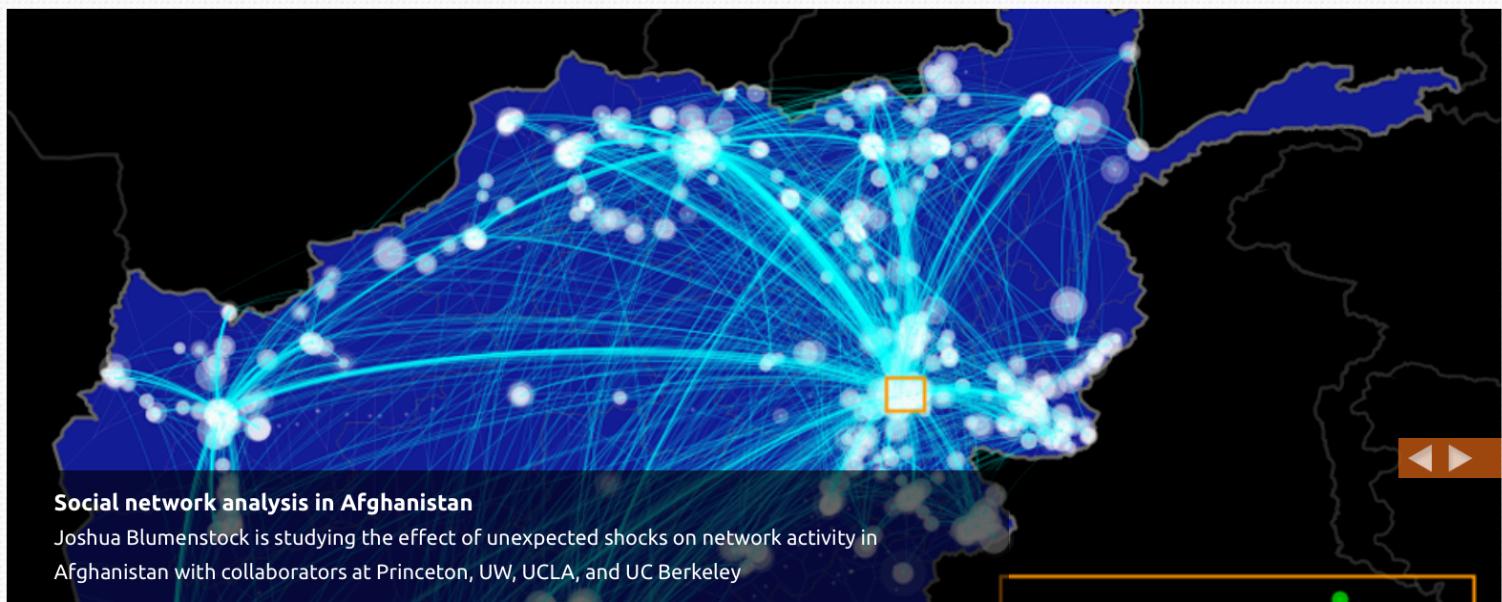
344,917 papers

8,174,533 citations

23.7 citations/paper

2006 - 2015

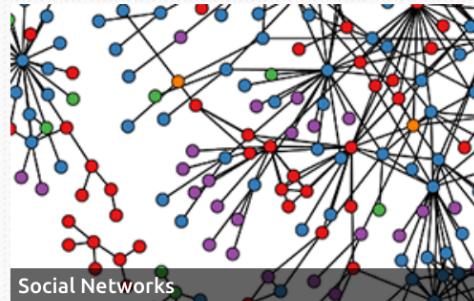




Research Focus Areas



Data for Development



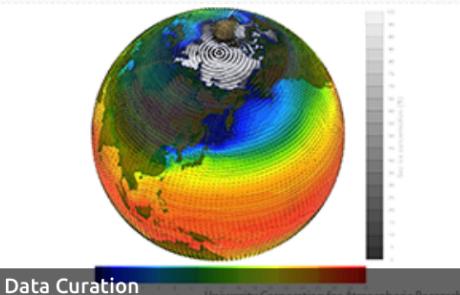
Social Networks



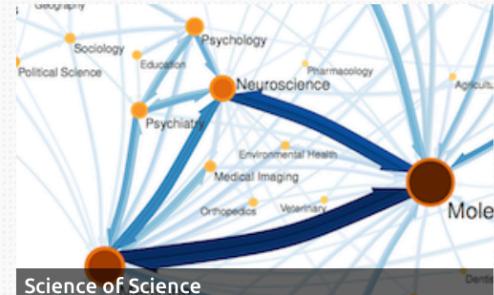
Data Visualization



Computational Social Science



Data Curation



Science of Science

What We Do



Overview

Over the course of the last decade many disciplines have evolved from recording observations in laboratory notebooks to the use of instruments capable of digitally recording many gigabytes of data in a day. This abundance of data provides unprecedented opportunities for discovery. Tapping its potential requires the application of sophisticated new computational techniques operating on large scale storage, computational and network resources. Since its creation in 2008, the eScience Institute has worked to create the intellectual and physical infrastructure needed to meet this challenge.

At the core of the eScience Institute are individuals who have proven track records in developing and applying advanced computational methods and tools to real world problems. Their task is to seek out and engage researchers across disciplines where eScience approaches are likely to have the greatest impact. To ensure that researchers have access to the necessary physical infrastructure, the Institute has undertaken coordinated planning and support for advanced local and remote computational platforms. This includes developing relationships with commercial and non-commercial service providers as well as the development of shared facilities on campus. This support extends to assistance in the preparation of select proposals where we are able to focus resources, improving their chances for success.

Also in... What We Do

[Appliance Gallery](#)

Find and use the eScience Institute's virtual machines equipped with software useful for specific applications.

[Campus Compute & Storage](#)

Learn about what UW is doing to support scalable scientific computing on campus

[Consulting & Services](#)

From algorithm development to database creation to cloud computing, we can help.

[Projects](#)

Explore some of our current collaborations with research scientists.

[Relevant Courses](#)

View a list of courses offered in eScience disciplines.

[SQLShare Success Stories](#)

[Tools](#)

Whether it's database management, visualization, or developer tools, learn about tools we can help you use.

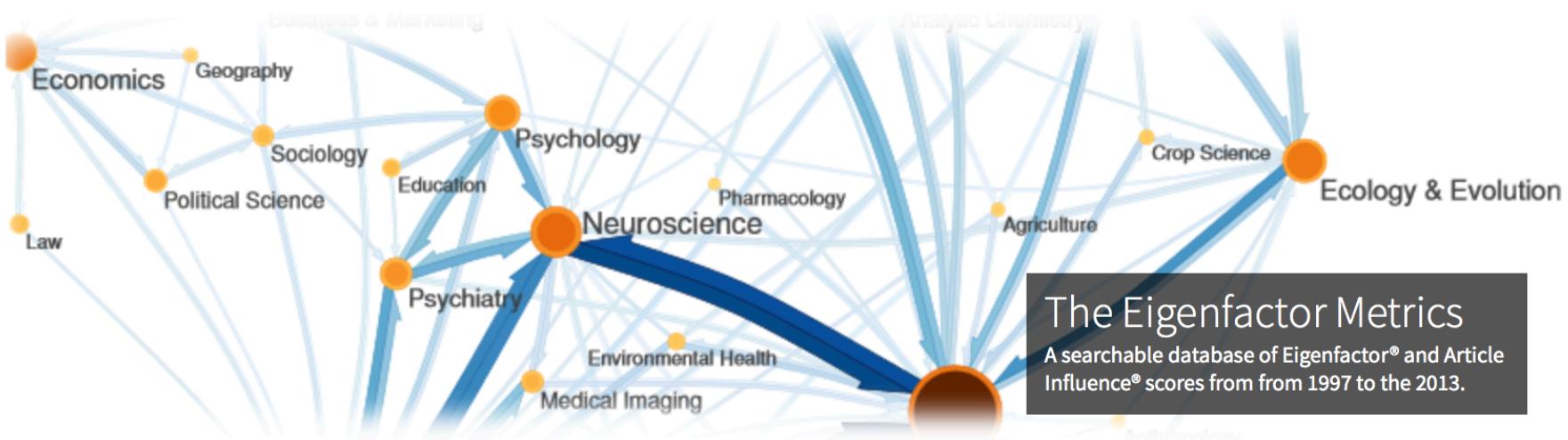
Latest eScience News

[Data Science Incubation Program - Winter 2016](#)

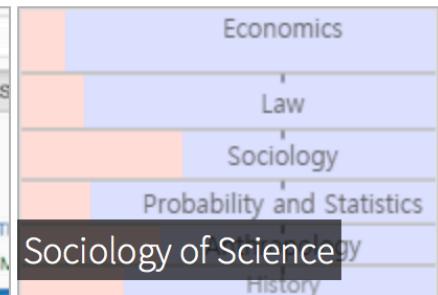
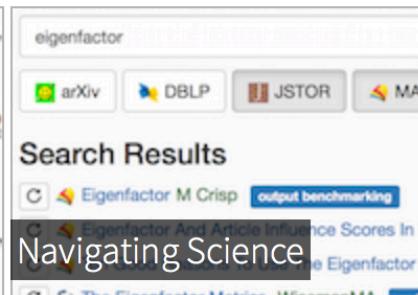
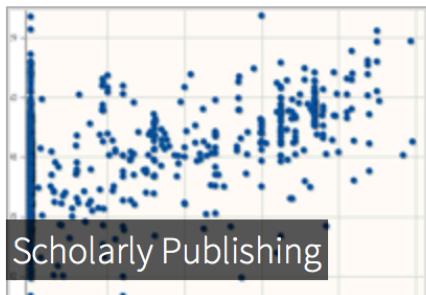
2 hours 4 min ago

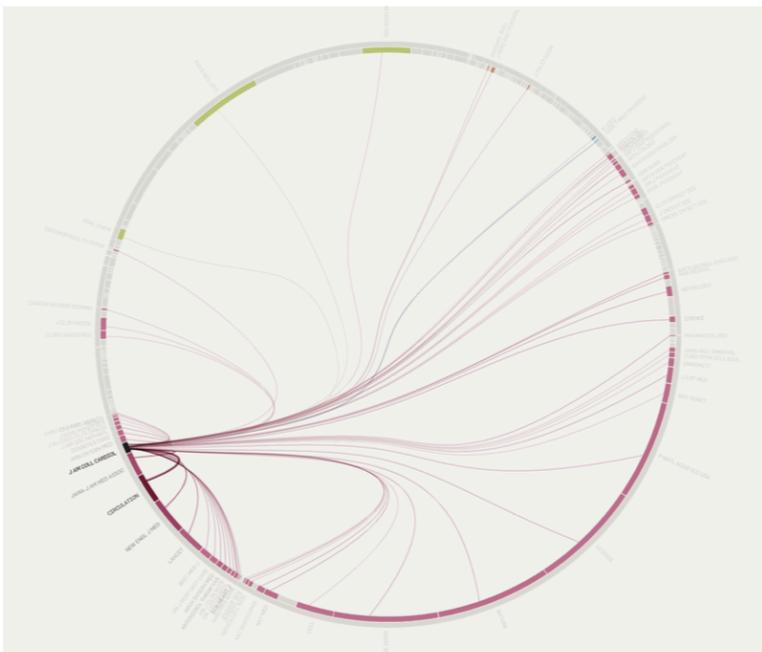
[Ben Marwick On How Computers Broke Science](#)

Search

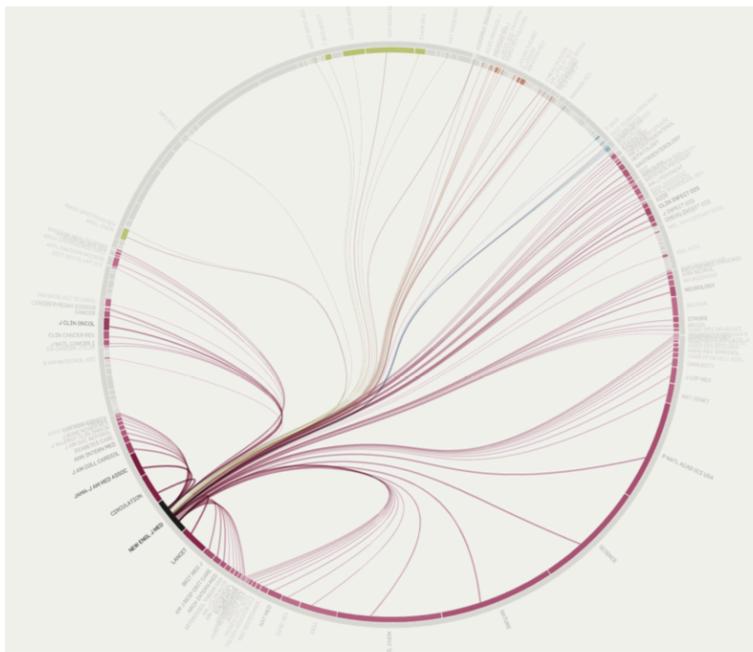


RESEARCH AREAS

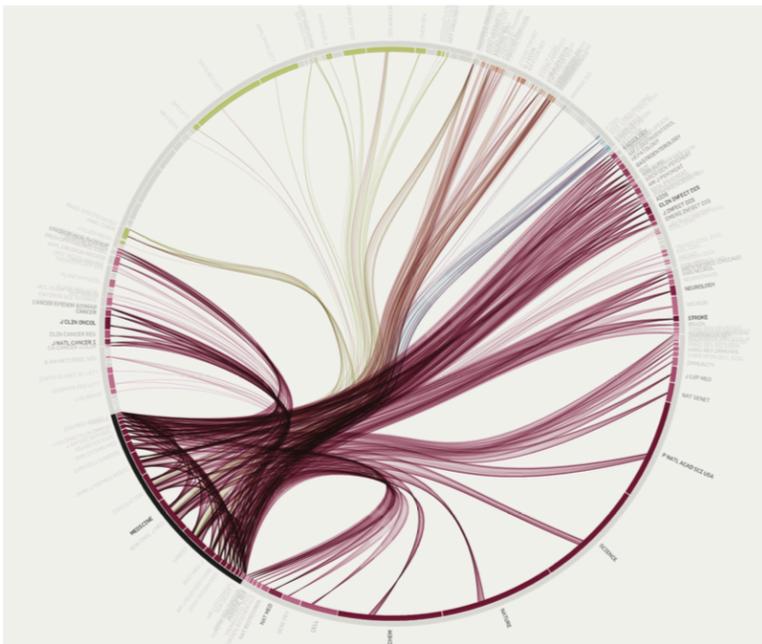




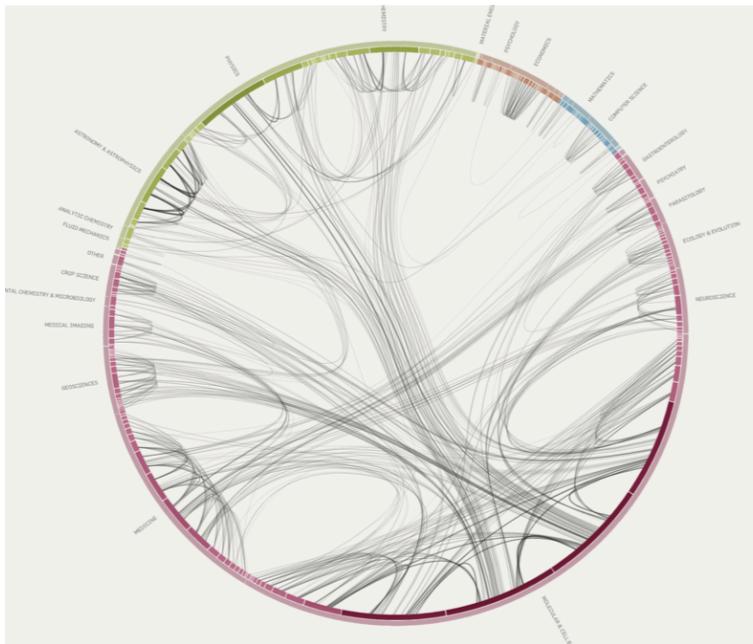
a) Journal of the ACC



b) New England Journal of Medicine



c) All of Medicine



d) All of Science

List 2: Medicine
Differences in Relative Ranking (2006 Data)

Impact Factor	Article Influence
CA-CANCER J CLIN	NEW ENGL J MED
NEW ENGL J MED	CA-CANCER J CLIN
LANCET	JAMA-J AM MED ASSOC
JAMA-J AM MED ASSOC	LANCET
J NATL CANCER I	8.635
ANN INTERN MED	5.884
PLOS MED	5.803
J CLIN ONCOL	5.772
ANNU REV MED	5.473
CLIN MICROBIOL REV	5.469
LANCET INFECT DIS	CIRCULATION
CIRCULATION	4.273
ANNU REV NUTR	4.004
CIRC RES	4.002
J AM COLL CARDIOL	3.919
BRIT MED J	3.665
AM J RESP CRIT CARE	3.475
J ALLERGY CLIN IMMUN	3.354
EPIDEMIOLOGY	3.287
CLIN PHARMACOL THER	3.271
DIABETES	2.895
ARCH INTERN MED	2.825
DIABETES CARE	2.673
MILBANK Q	2.631
ARTERIOSCL THROM VAS	2.426
CAN MED ASSOC J	2.426
ANNU REV PUBL HEALTH	2.307
AM J TRANSPLANT	2.224
MILBANK Q	2.211
HUM REPROD UPDATE	2.199
J BONE MINER RES	2.198
CRIT CARE MED	2.134
AM J CLIN NUTR	2.119
CLIN INFECT DIS	2.117
THORAX	2.097
HYPERTENSION	2.077
ATHEROSCLEROSIS SUPP	2.05
CARDIOVASC RES	2.049
J CLIN ENDOCR METAB	2.035
CURR OPIN LIPIDOL	2.033
AIDS	2.002
CLIN CHEM	1.980
STROKE	1.976
PHARMACOGENET GENOM	1.906
J INFECT DIS	1.89
ALLERGY	1.881
HEALTH TECHNOL ASSES	1.872
DIABETOLOGIA	1.829
AM J EPIDEMIOL	1.816
ONCOLOGIST	1.805
ANN ONCOL	1.792
MEDICINE	1.787
J THROMB HAEMOST	1.785
EMERG INFECT DIS	1.780
EUR RESPIR J	1.758
B WORLD HEALTH ORGAN	1.739
PEDIATRICS	1.733
ANTIVIR THER	1.712
J MOL CELL CARDIOL	1.695
CURR OPIN INFECT DIS	1.661
KIDNEY INT	1.637
ARCH DIS CHILD-FETAL	1.598
J INTERN MED	1.589
BREAST CANCER RES TR	1.582
ANN MED	1.572
CANCER	1.563
AM J MED	1.544
INT J EPIDEMIOL	1.51
DRUGS	1.505
INTENS CARE MED	1.488
CANCER TREAT REV	1.483
EPIDEMIOLOGY	1.482
BONE	1.478
AM J PHYSIOL-ENDOC M	1.478
CRIT CARE MED	1.455

List 3: Cardiac & Cardiovascular Systems
and Peripheral Vascular Disease
Differences in Relative Ranking (2006 Data)

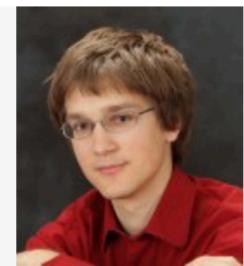
Impact Factor	Article Influence
CIRCULATION	CIRCULATION
CIRC RES	CIRC RES
J AM COLL CARDIOL	J AM COLL CARDIOL
ARTERIOSCL THROM VAS	ARTERIOSCL THROM VAS
EUR HEART J	EUR HEART J
HYPERTENSION	HYPERTENSION
ATHEROSCLEROSIS SUPP	ATHEROSCLEROSIS SUPP
CARDIOVASC RES	CARDIOVASC RES
CURR OPIN LIPIDOL	CURR OPIN LIPIDOL
STROKE	STROKE
J THROMB HAEMOST	J THROMB HAEMOST
J MOL CELL CARDIOL	J MOL CELL CARDIOL
TRENDS CARDIOVAS MED	TRENDS CARDIOVAS MED
J THORAC CARDIOV SUR	J THORAC CARDIOV SUR
HEART	HEART
AM HEART J	AM HEART J
HEART RHYTHM	HEART RHYTHM
HEART FAIL REV	HEART FAIL REV
ATHEROSCLEROSIS	ATHEROSCLEROSIS
AM J PHYSIOL-HEART C	AM J PHYSIOL-HEART C
J HYPERTENS	J HYPERTENS
CARDIOL CLIN	CARDIOL CLIN
AM J CARDIOL	AM J CARDIOL
J CARDIOVASC MAGN R	J CARDIOVASC MAGN R
J VASC SURG	J VASC SURG
THROMB HAEMOSTASIS	THROMB HAEMOSTASIS
J CARD FAIL	J CARD FAIL
ATHEROSCLEROSIS SUPP	ATHEROSCLEROSIS SUPP
PROG CARDIOVASC DIS	PROG CARDIOVASC DIS
CURR OPIN NEPHROL HY	CURR OPIN NEPHROL HY
MICROCIRCULATION	MICROCIRCULATION
J CARDIOVASC ELECTR	J CARDIOVASC ELECTR
AM J HYPERTENS	AM J HYPERTENS
J HEART LUNG TRANSPL	J HEART LUNG TRANSPL
HEART FAIL REV	HEART FAIL REV
THROMB HAEMOSTASIS	THROMB HAEMOSTASIS
CURR OPIN CARDIOL	CURR OPIN CARDIOL
J CARD FAIL	J CARD FAIL
SEMIN THROMB HEMOST	SEMIN THROMB HEMOST
NAT CLIN PRACT CARD	NAT CLIN PRACT CARD
J VASC RES	J VASC RES
PROG CARDIOVASC DIS	PROG CARDIOVASC DIS
MICROVASC RES	MICROVASC RES
J NUCL CARDIOL	J NUCL CARDIOL
MICROCIRCULATION	MICROCIRCULATION
J VASC INTERV RADIAL	J VASC INTERV RADIAL
ENDOTHELIUM-J ENDOTH	ENDOTHELIUM-J ENDOTH
ANN THORAC SURG	ANN THORAC SURG
INT J CARDIOL	INT J CARDIOL
REV ESP CARDIOL	REV ESP CARDIOL
EUR J VASC ENDOVASC	EUR J VASC ENDOVASC
J ENDOVASC THER	J ENDOVASC THER
CIRC J	CIRC J
EUR J CARDIO-THORAC	EUR J CARDIO-THORAC
RESP MED	RESP MED
THROMB RES	THROMB RES
CEREBROVASC DIS	CEREBROVASC DIS
CURR HYPERTENS REP	CURR HYPERTENS REP
EUR J CARDIOP REV R	EUR J CARDIOP REV R
CARDIOVASC PATHOL	CARDIOVASC PATHOL
KIDNEY BLOOD PRESS R	KIDNEY BLOOD PRESS R
NUTR METAB CARDIOVAS	NUTR METAB CARDIOVAS
CARDIOLOGY	CARDIOLOGY
J CARDIOVASC MAGN R	J CARDIOVASC MAGN R
CARDIOVASC DRUG REV	CARDIOVASC DRUG REV
J AM SOC ECHOCARDIOG	J AM SOC ECHOCARDIOG
J CARDIOVASC PHARM	J CARDIOVASC PHARM
CATHETER CARDIO INTE	CATHETER CARDIO INTE
CARDIOLOGY	CARDIOLOGY
VASC MED	VASC MED
ANN VASC SURG	ANN VASC SURG
J CARDIOVASC PHARM	J CARDIOVASC PHARM
ENDOTHELIUM-J ENDOTH	ENDOTHELIUM-J ENDOTH
J NUCL CARDIOL	J NUCL CARDIOL
HYPERTENS RES	HYPERTENS RES
EUROPACE	EUROPACE
CARDIOVASC DRUG THER	CARDIOVASC DRUG THER
HYPERTENS PREGNANCY	HYPERTENS PREGNANCY
CURR CONTR TRIALS C	CURR CONTR TRIALS C
CORONARY ARTERY DIS	CORONARY ARTERY DIS
CARDIOVASC INTER RAD	CARDIOVASC INTER RAD
CAN J CARDIOL	CAN J CARDIOL
NUTR METAB CARDIOVAS	NUTR METAB CARDIOVAS

What is my impact on science?

\$7,933,670,366



The H-index impact on science...



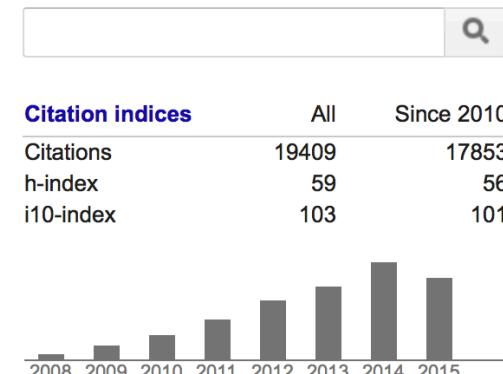
Jure Leskovec

[Follow](#)

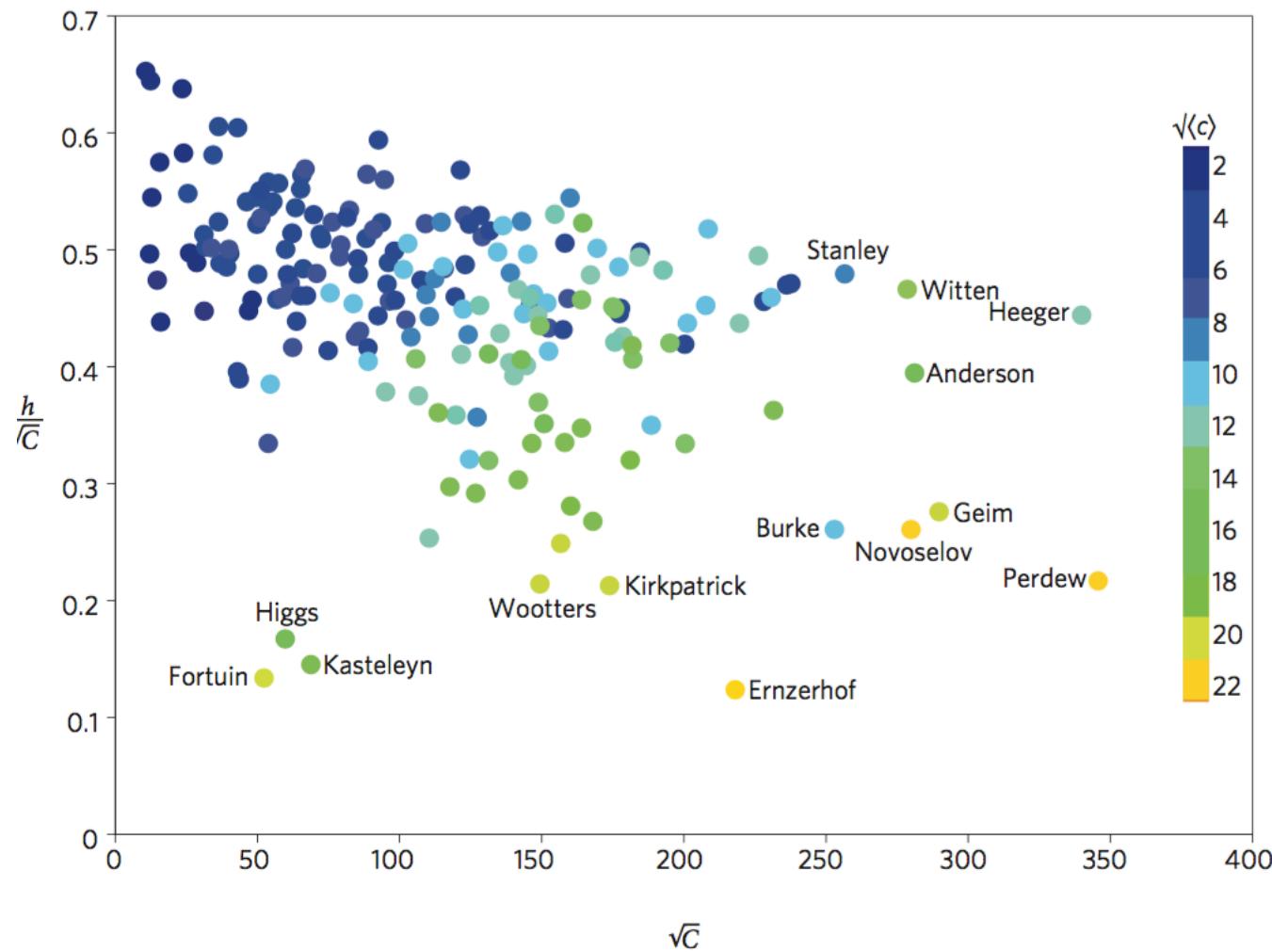
Professor of Computer Science, Stanford University
Data mining, Social Network Analysis, Information Networks
Verified email at cs.stanford.edu - [Homepage](#)

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

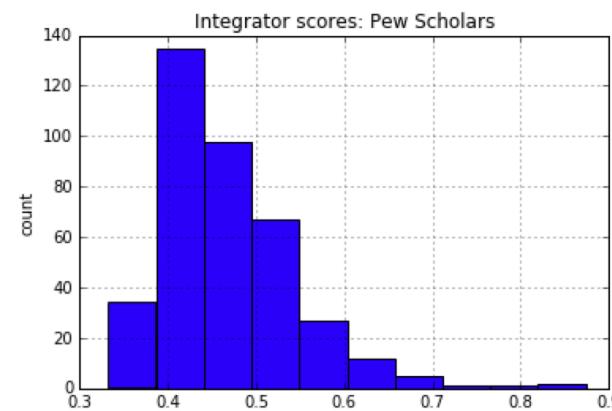
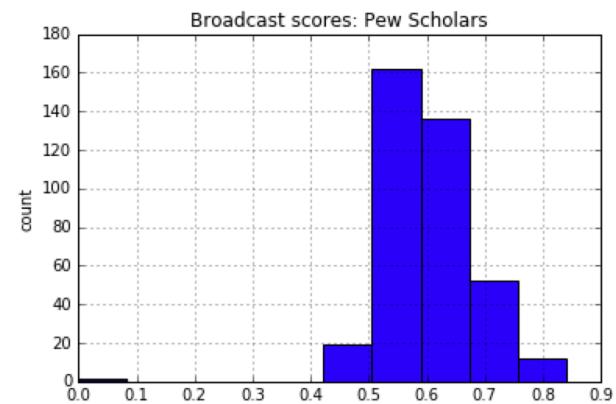
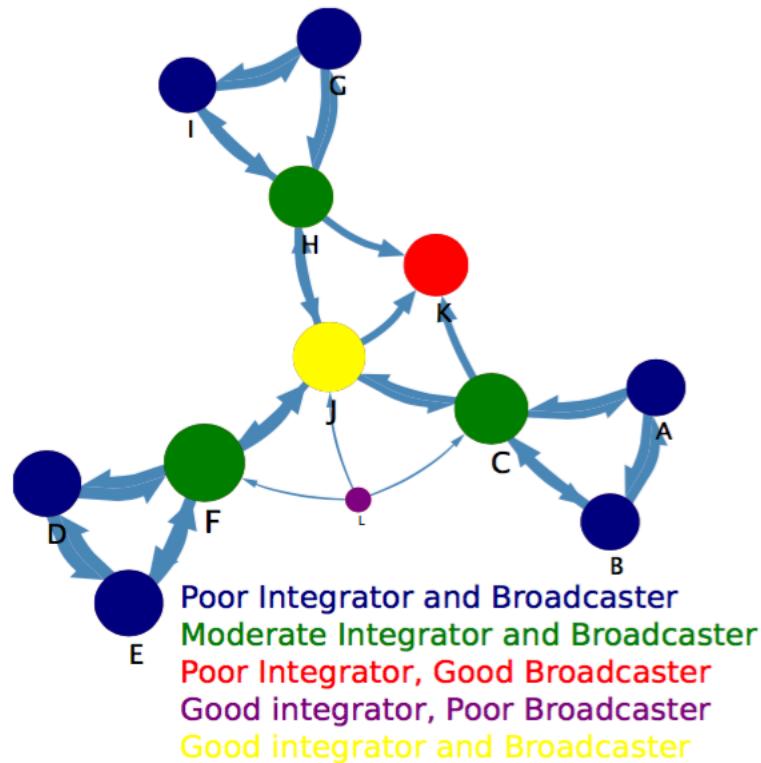
Google Scholar



An evisceration of the H-index...



Measuring Interdisciplinarity



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

Visualizing Scholarly Influence Over Time

Influence of Pew Scholars

Roberta A. Gottlieb

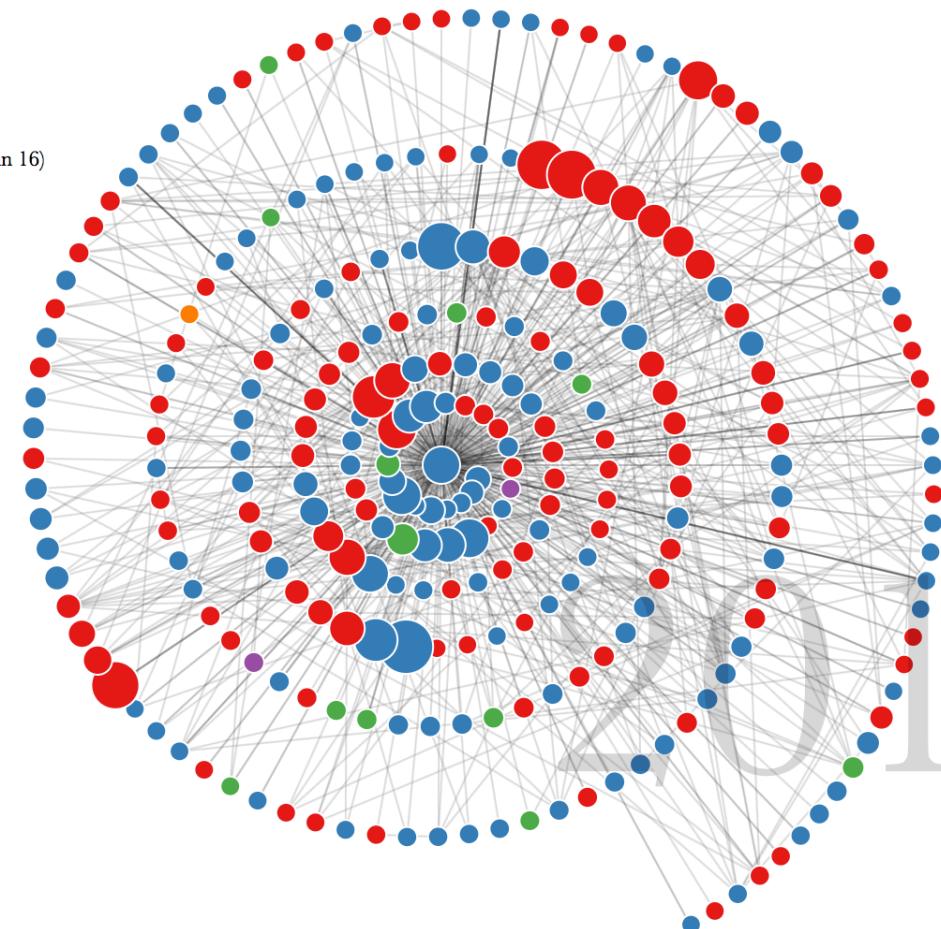
[Learn More](#)

- █ 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



1997

scholar.eigenfactor.org

username: PewScholar

password: 1N!kdG

Jevin West, jevinw@uw.edu

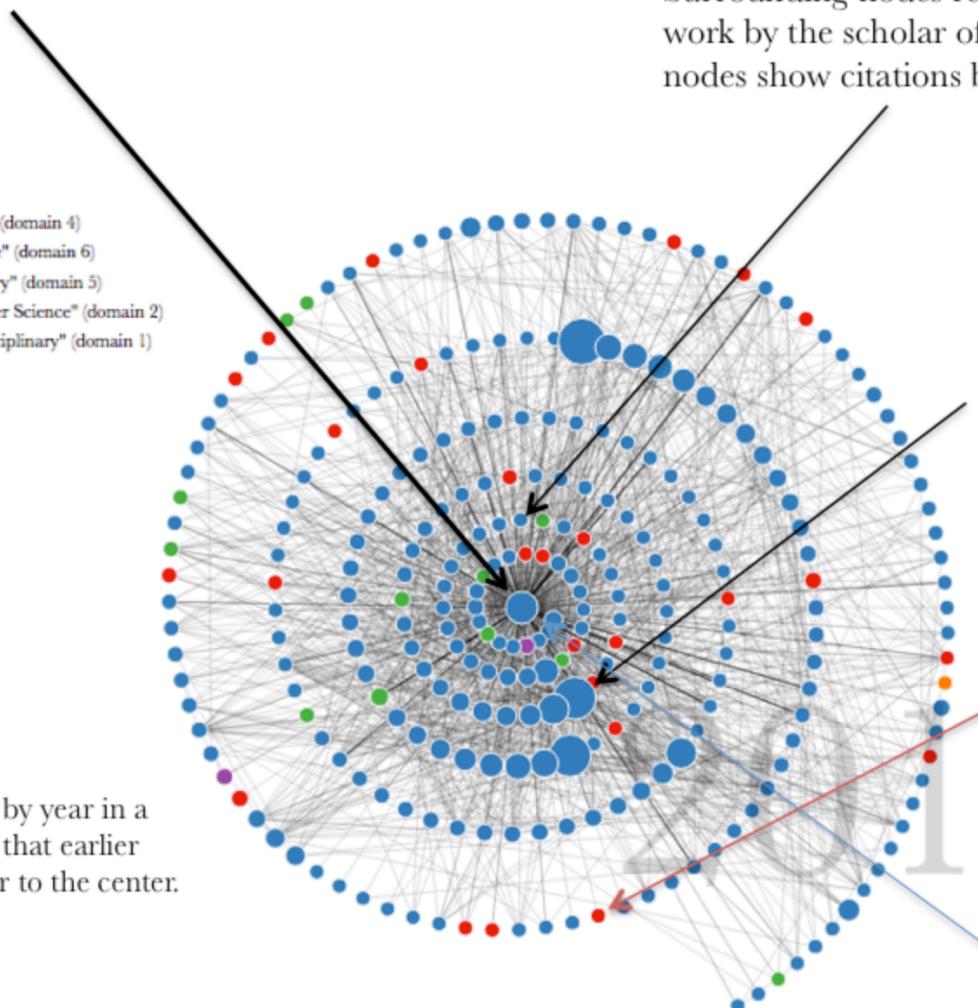
Pew Influence

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.

- Papers in category "Biology" (domain 4)
- Papers in category "Medicine" (domain 6)
- Papers in category "Chemistry" (domain 5)
- Papers in category "Computer Science" (domain 2)
- Papers in category "Multidisciplinary" (domain 1)

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



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.

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

Citation Data



49 million scholarly publications

260 million citations

354 Pew Scholars

22,000 publications

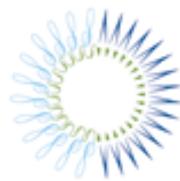
62 publications/scholar

Pew EF 3 times the average EF

field classification



Scholars Program
in the Biomedical Sciences



THE
PEW
CHARITABLE TRUSTS

Scholars Program
in the Biomedical Sciences

Science

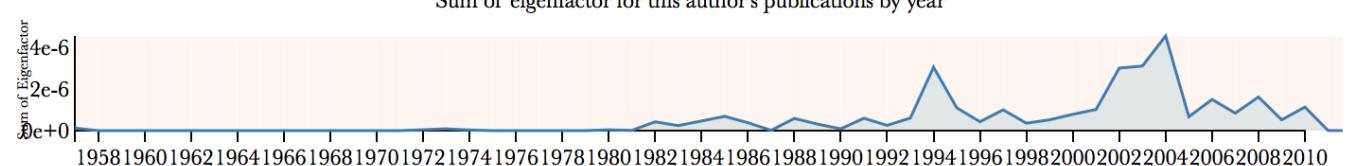
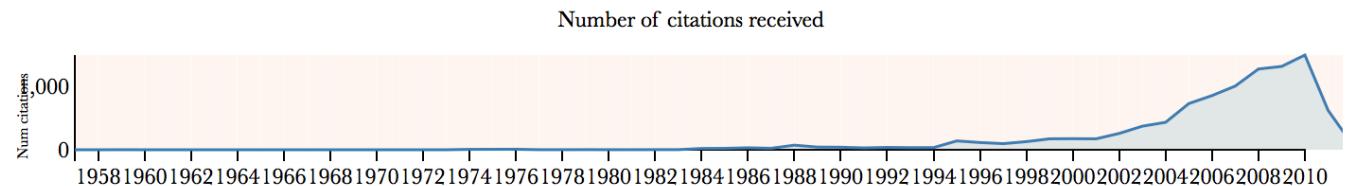
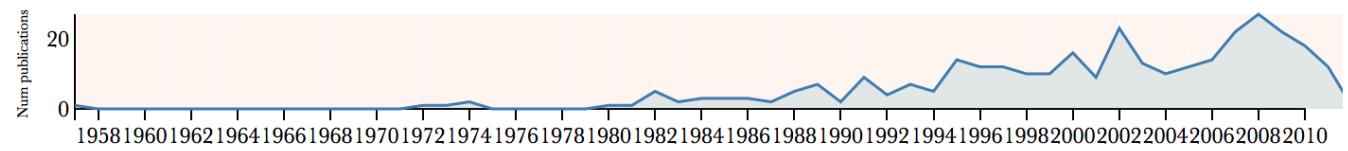
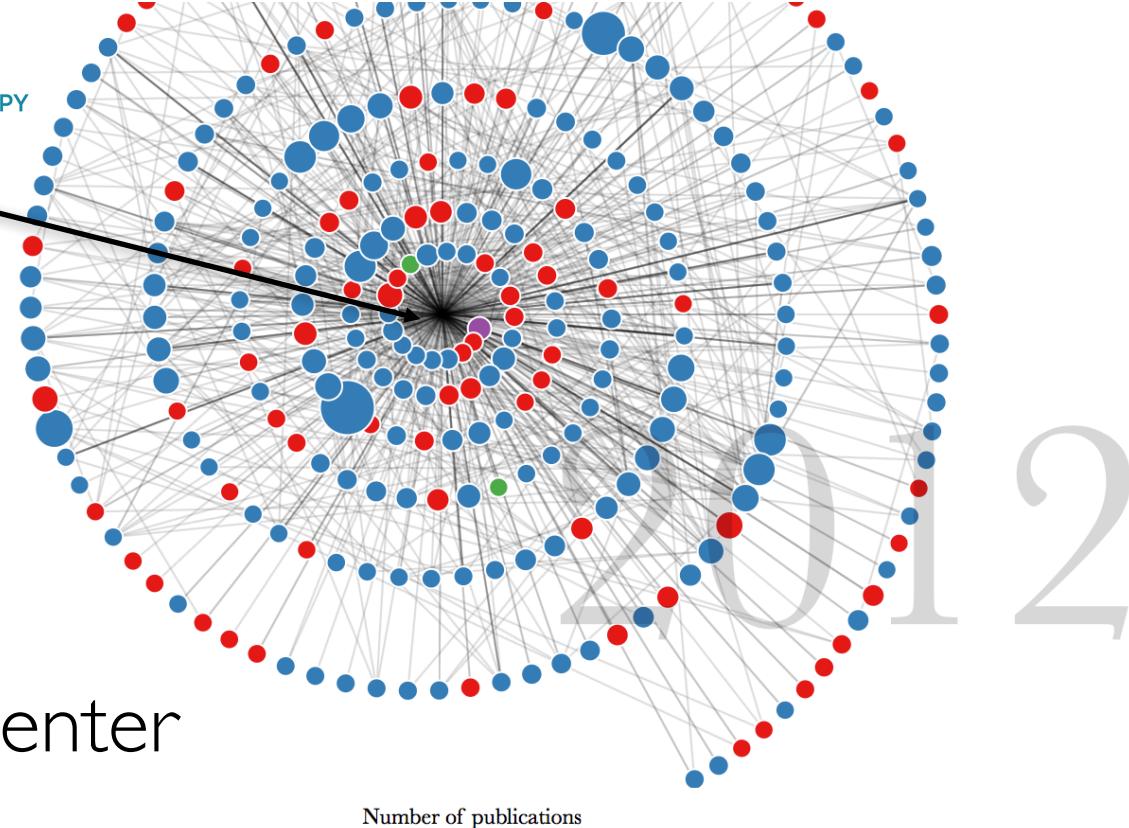
~ 37 citations/paper

median citations = ||

~ 5 citations/paper

median citations = 0

Institution as Center



Visualizing Scholarly Influence Over Time

Influence of Pew Scholars

Mark W. Grinstaff

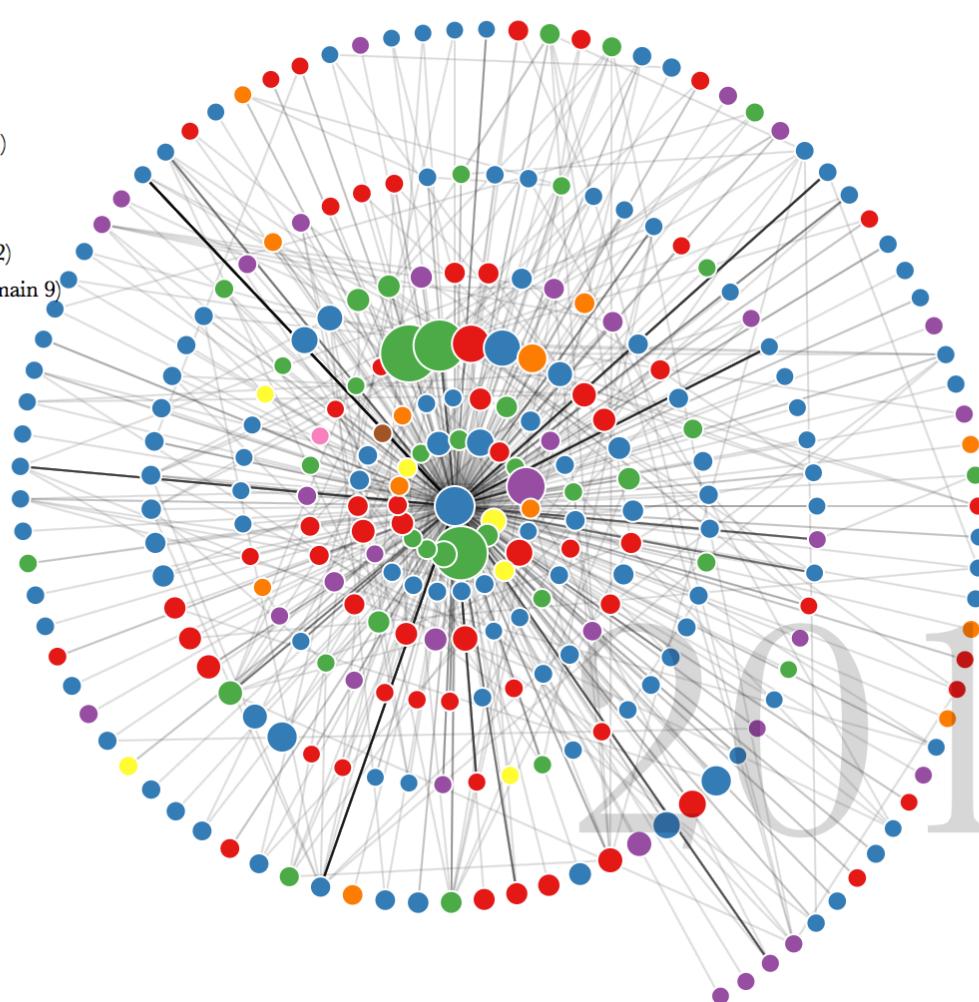
[Learn More](#)

- 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

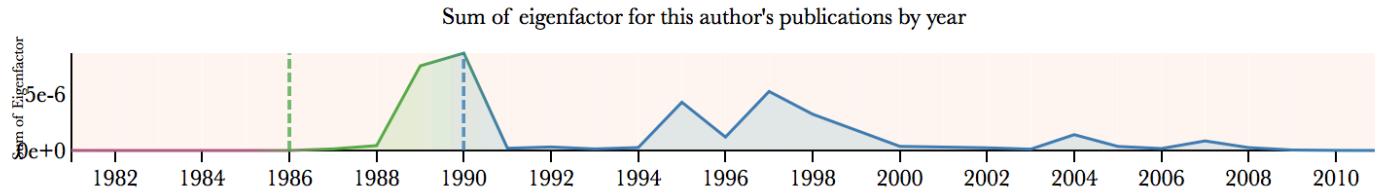
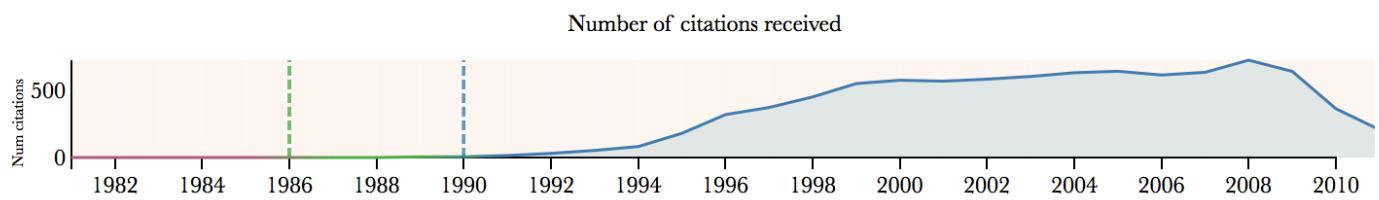
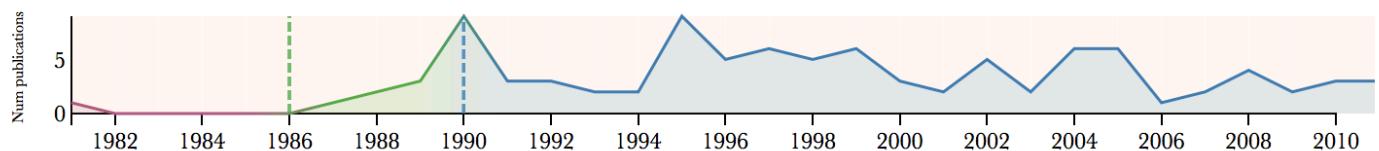
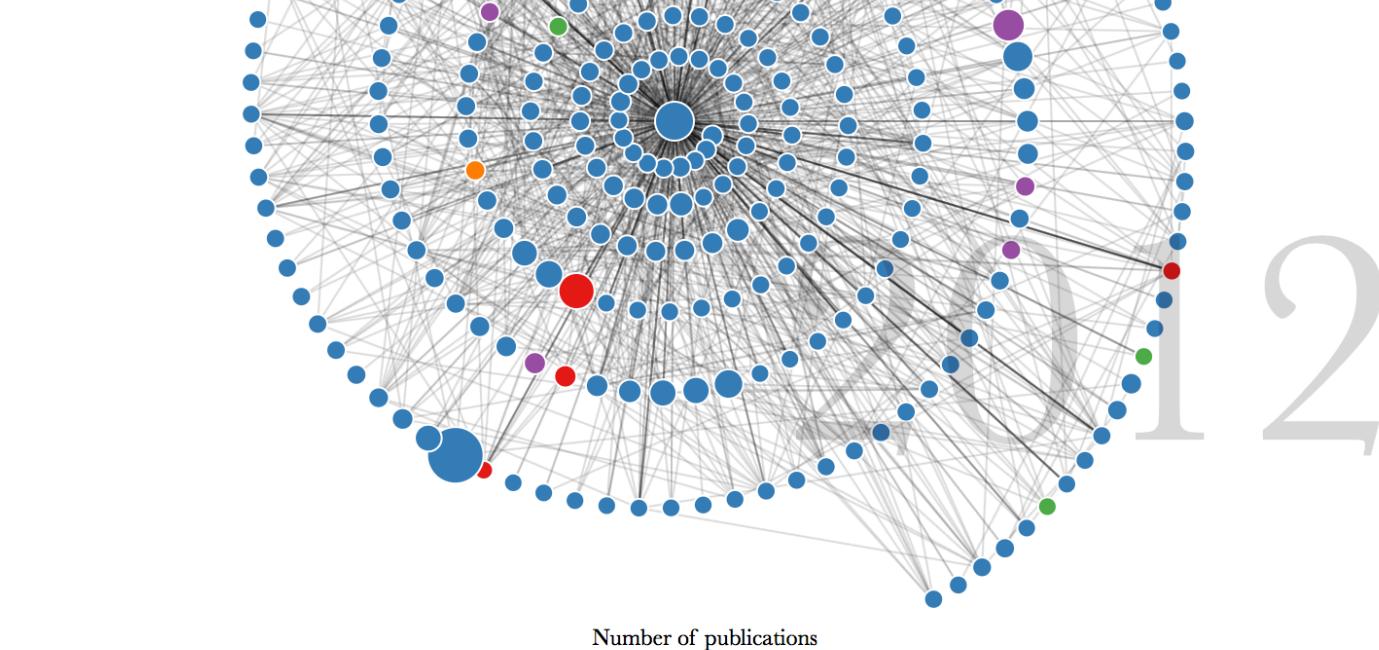


12

Philip A.
Hieter



Pew Scholar
1986



Funding

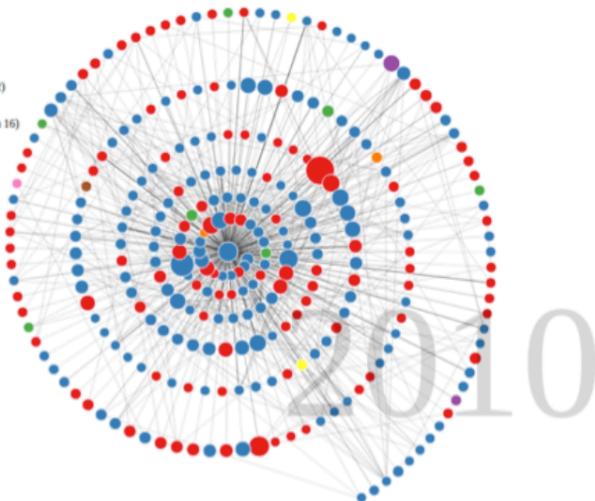
12

Visualizing Interdisciplinarity



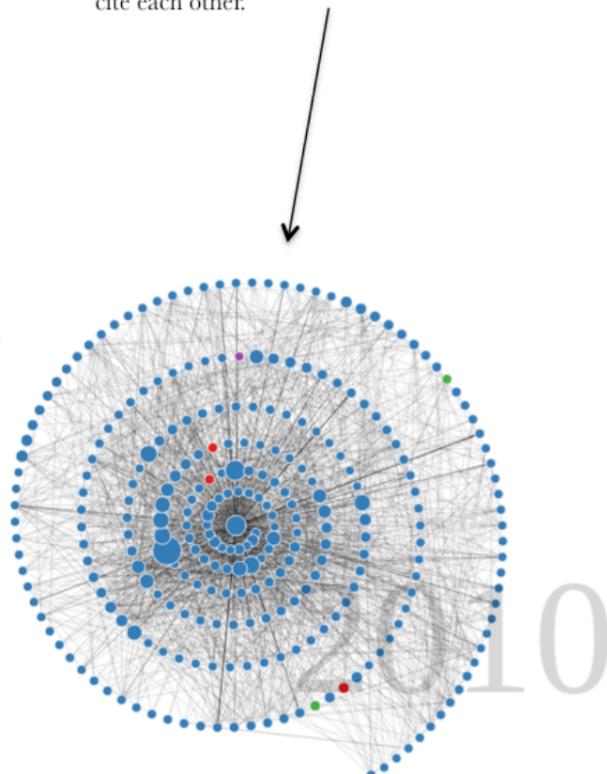
Jason Portenoy

- Papers in category "Medicine" (domain 6)
- Papers in category "Biology" (domain 4)
- Papers in category "Chemistry" (domain 5)
- Papers in category "Engineering" (domain 8)
- Papers in category "Material Science" (domain 12)
- Papers in category "Physics" (domain 19)
- Papers in category "Agriculture Science" (domain 16)
- Papers in category "Social Science" (domain 22)



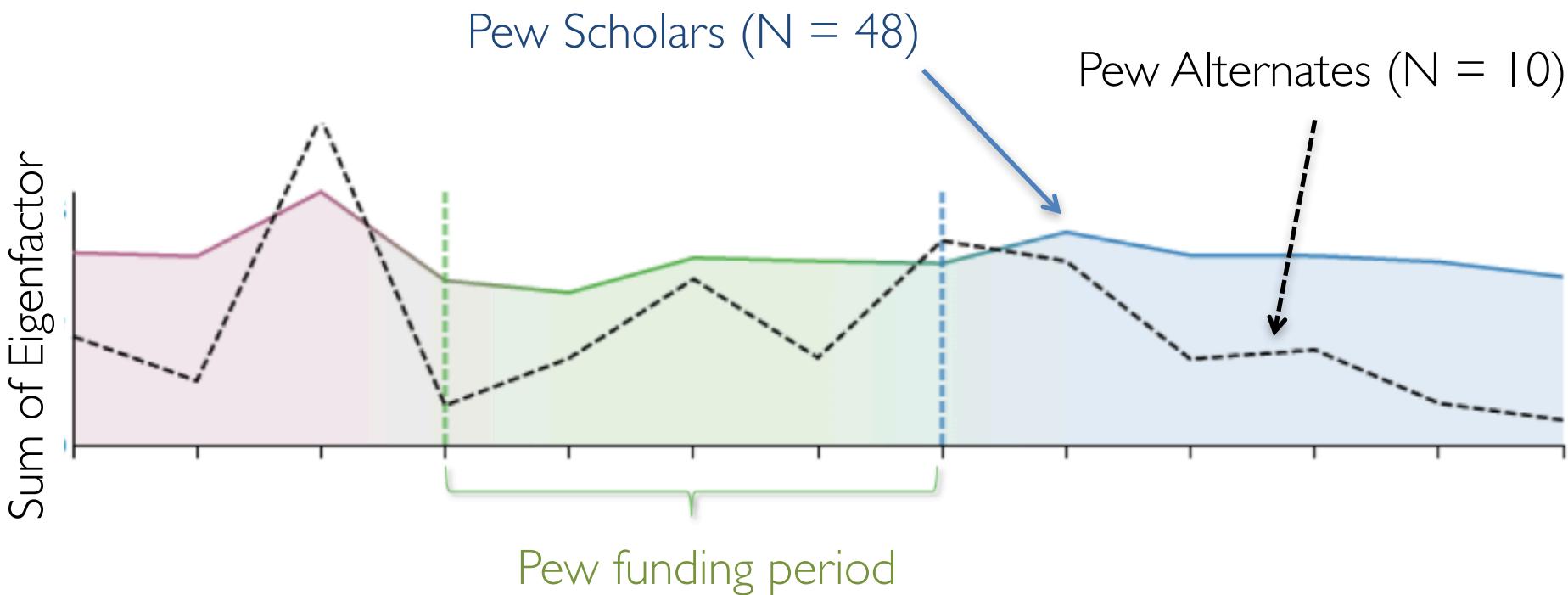
A more sparse network indicates fewer citations between papers shown in the network. This could be a result of the central scholar having impact across a wider set of academic communities.

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



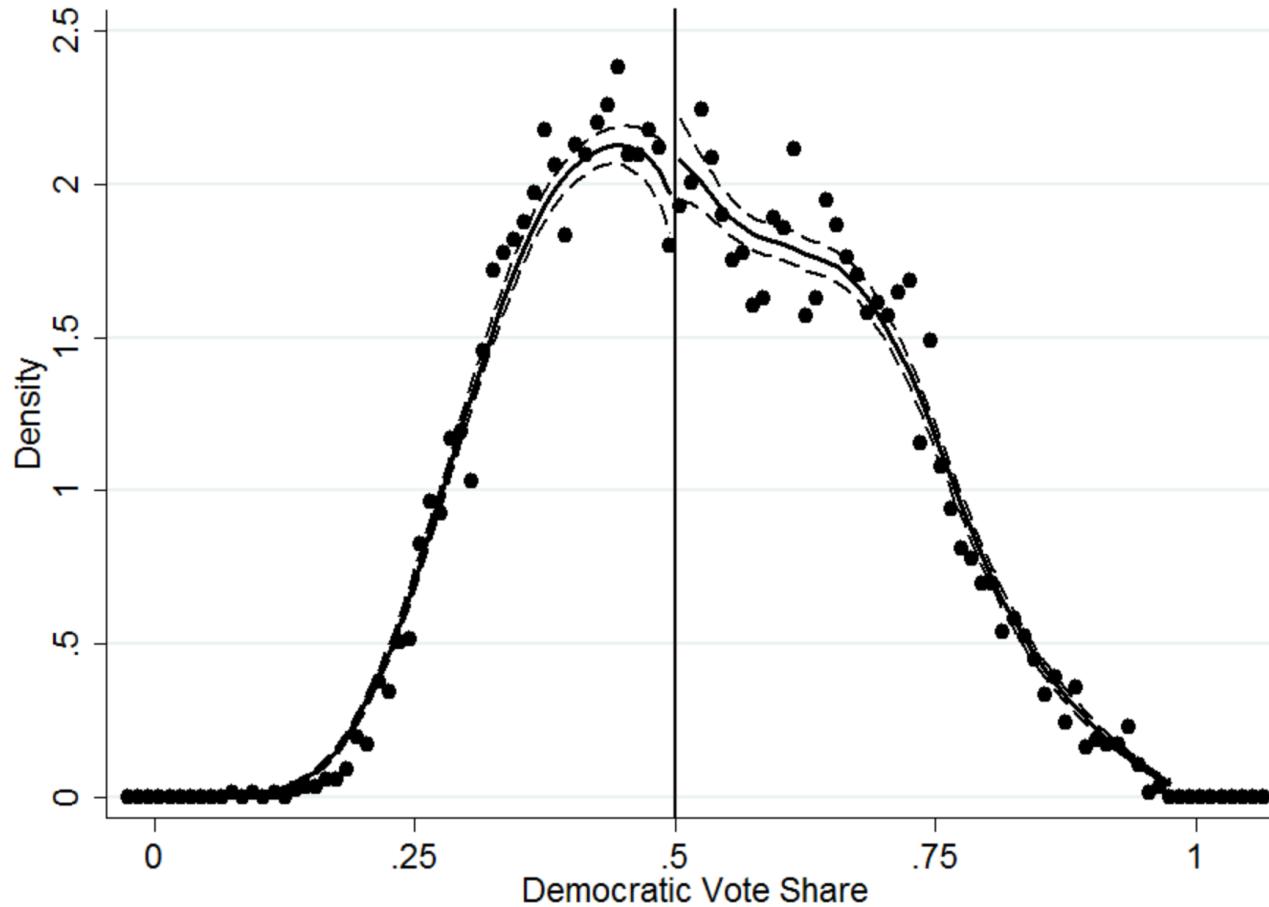
The Pew Impact...

Comparing Alternates



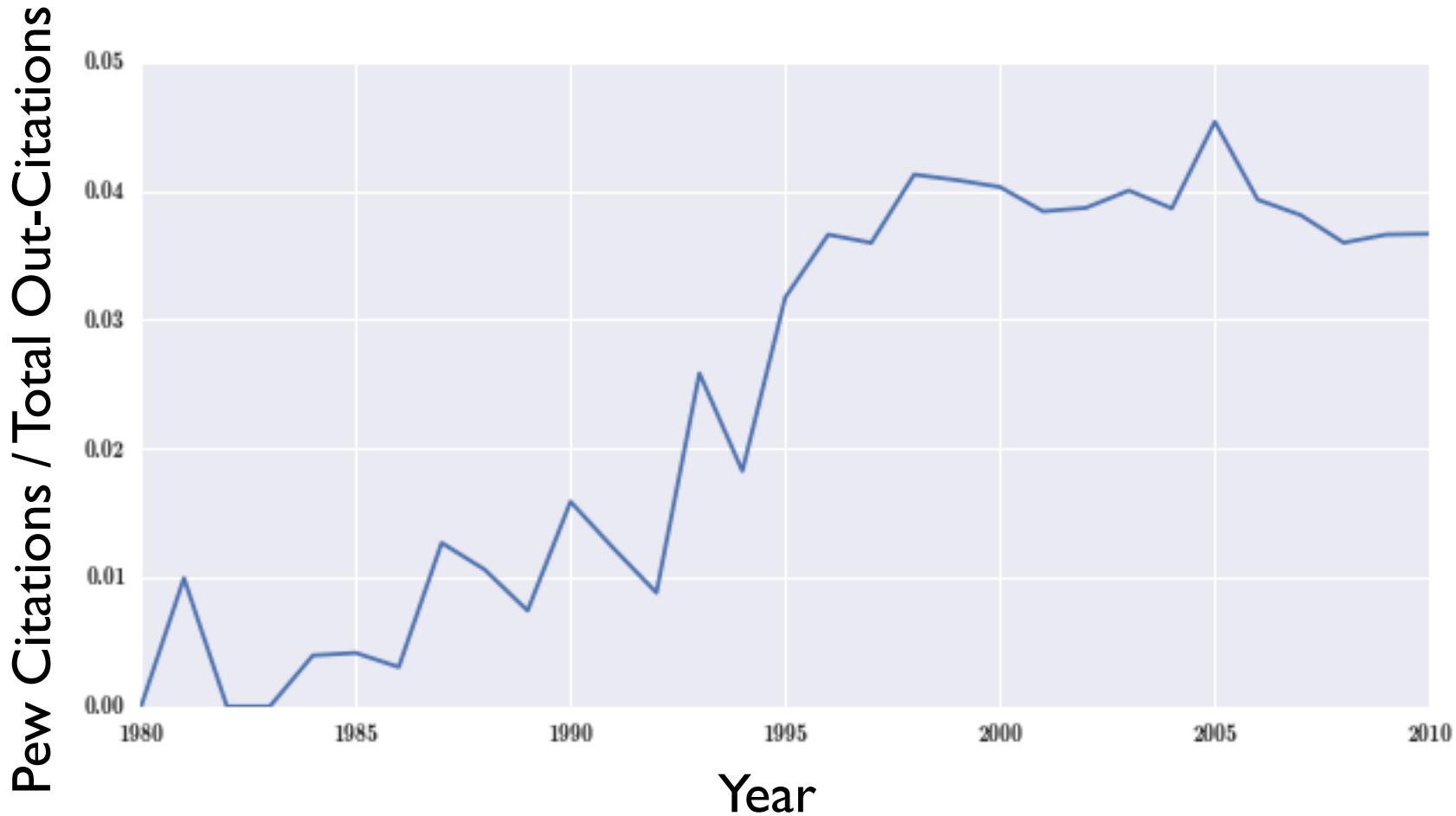
* Includes scholars and alternates from cohort years: 1997, 1999, 2000, 2001, 2002

Regression discontinuity design

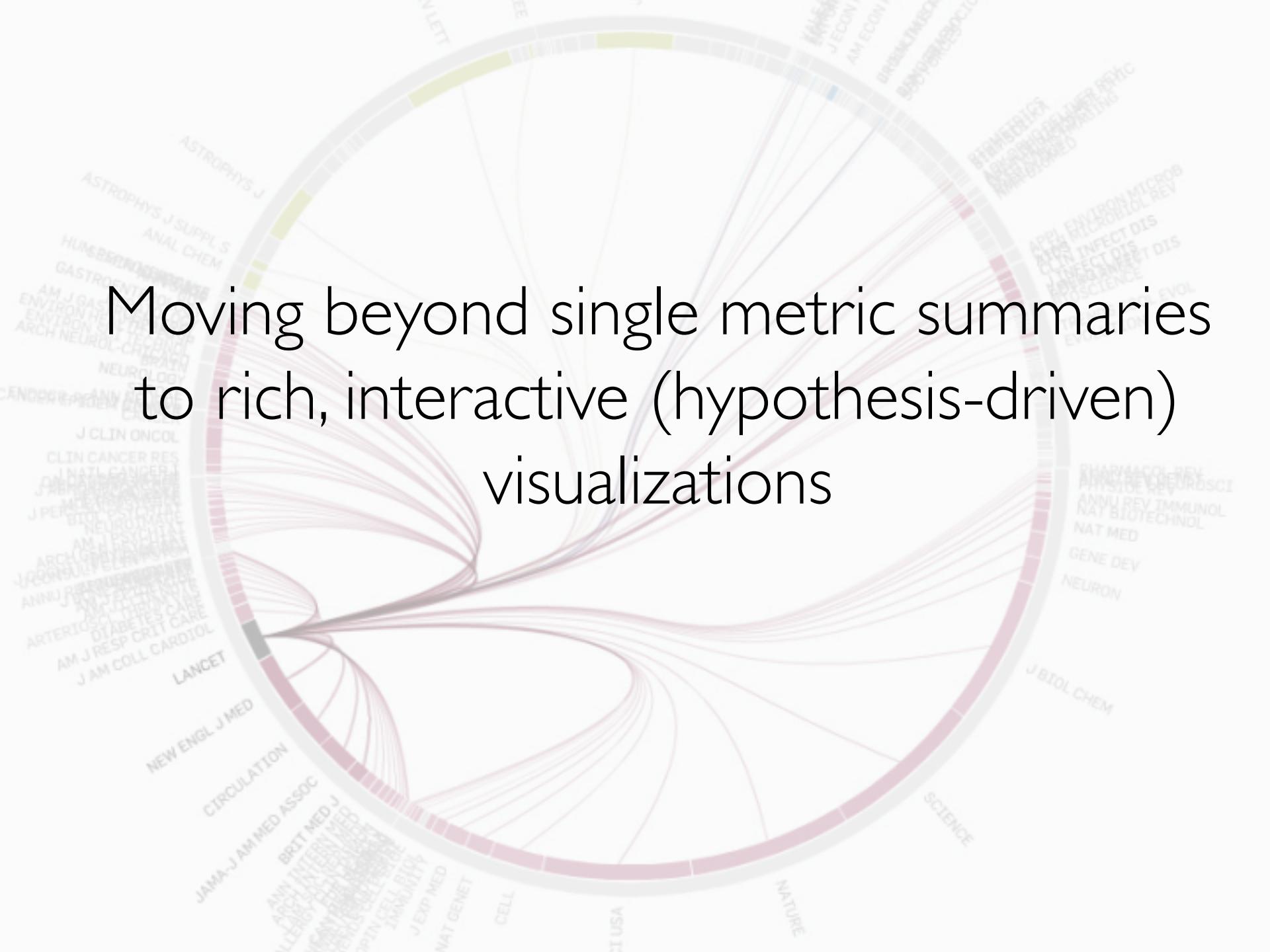


McCrary (2008)^[10] density test on data from Lee, Moretti, and Butler (2004).

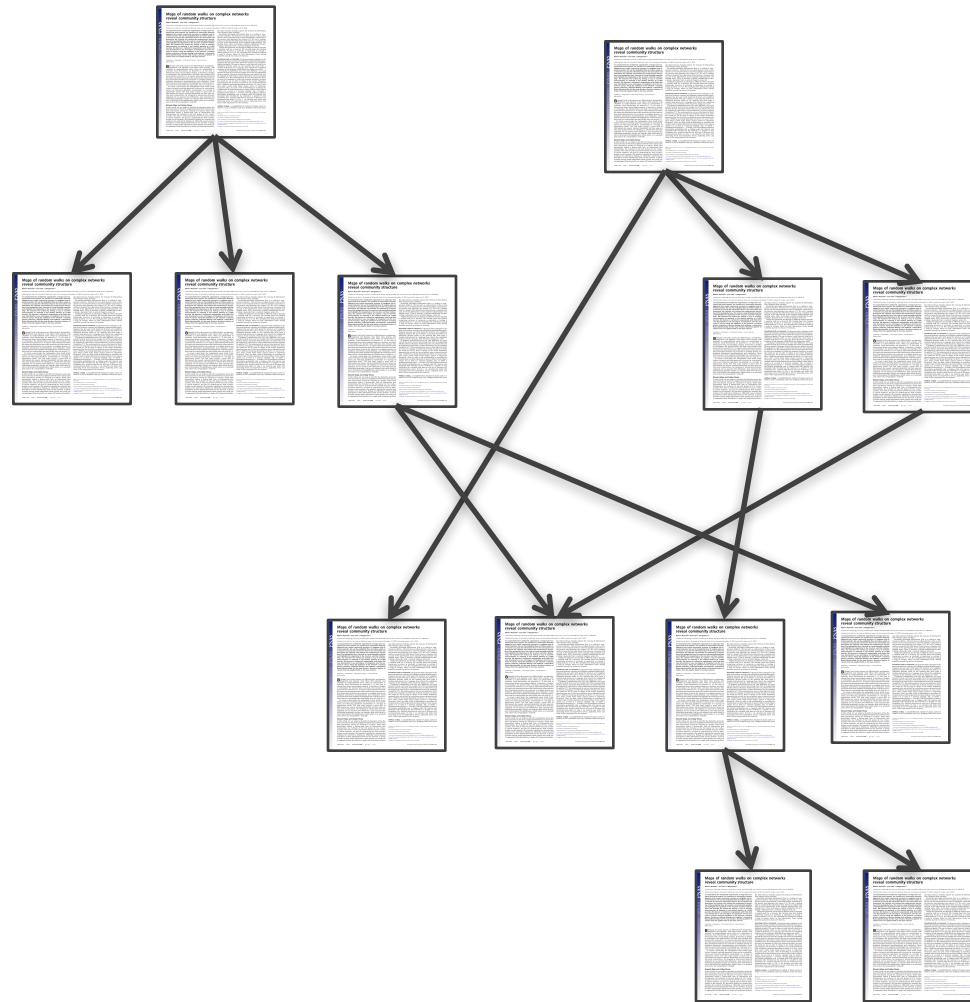
Community Effect



Moving beyond single metric summaries to rich, interactive (hypothesis-driven) visualizations

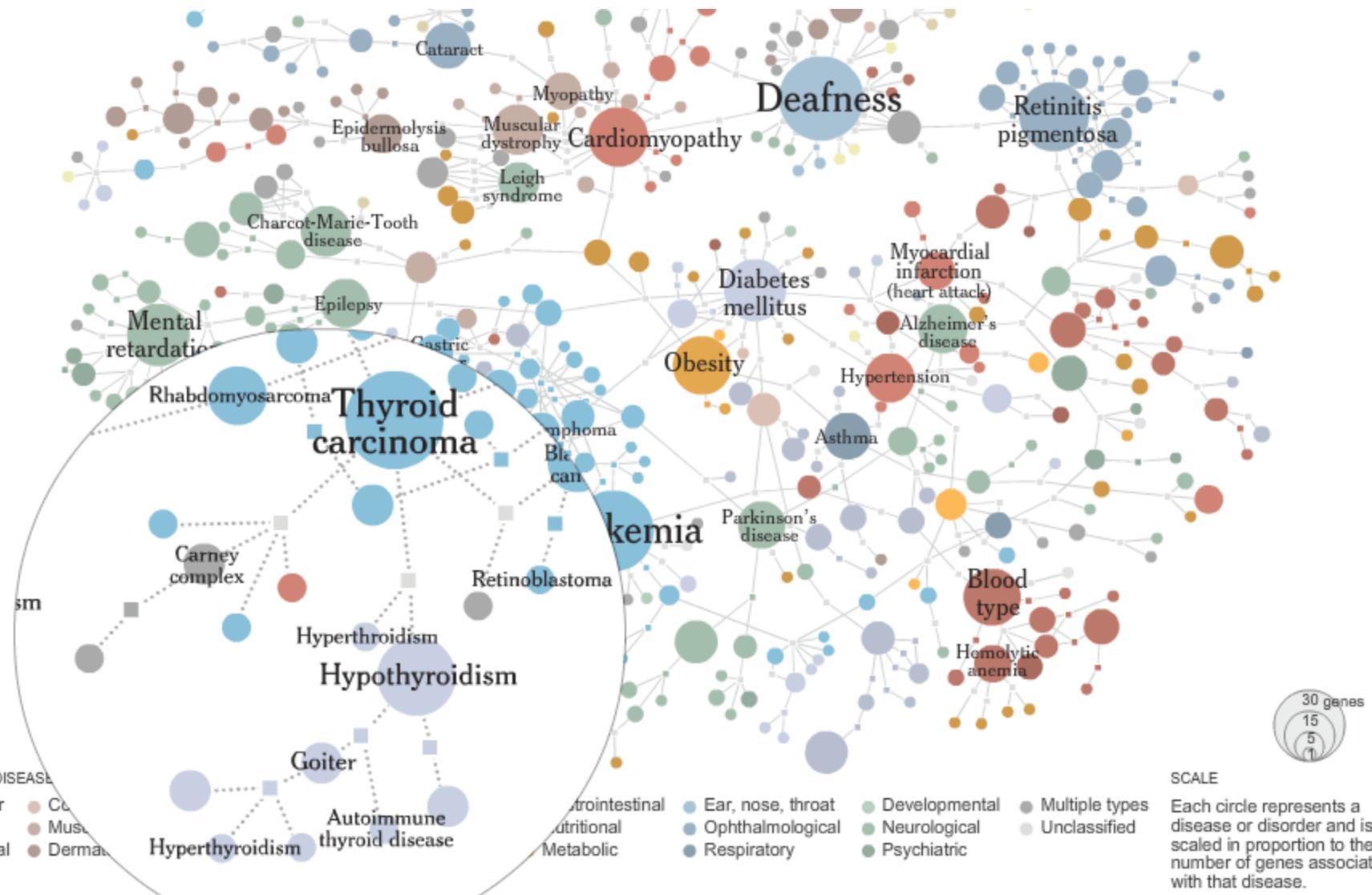


Citations form a vast network



de Solla Price, Science (1965)

Disease association network





The Scholarly Graph



PatentVector™



WIKIPEDIA
The Free Encyclopedia



PNAS





The Scholarly Graph



Tens of millions articles, patents, books



Billions of citation links

PatentVector™

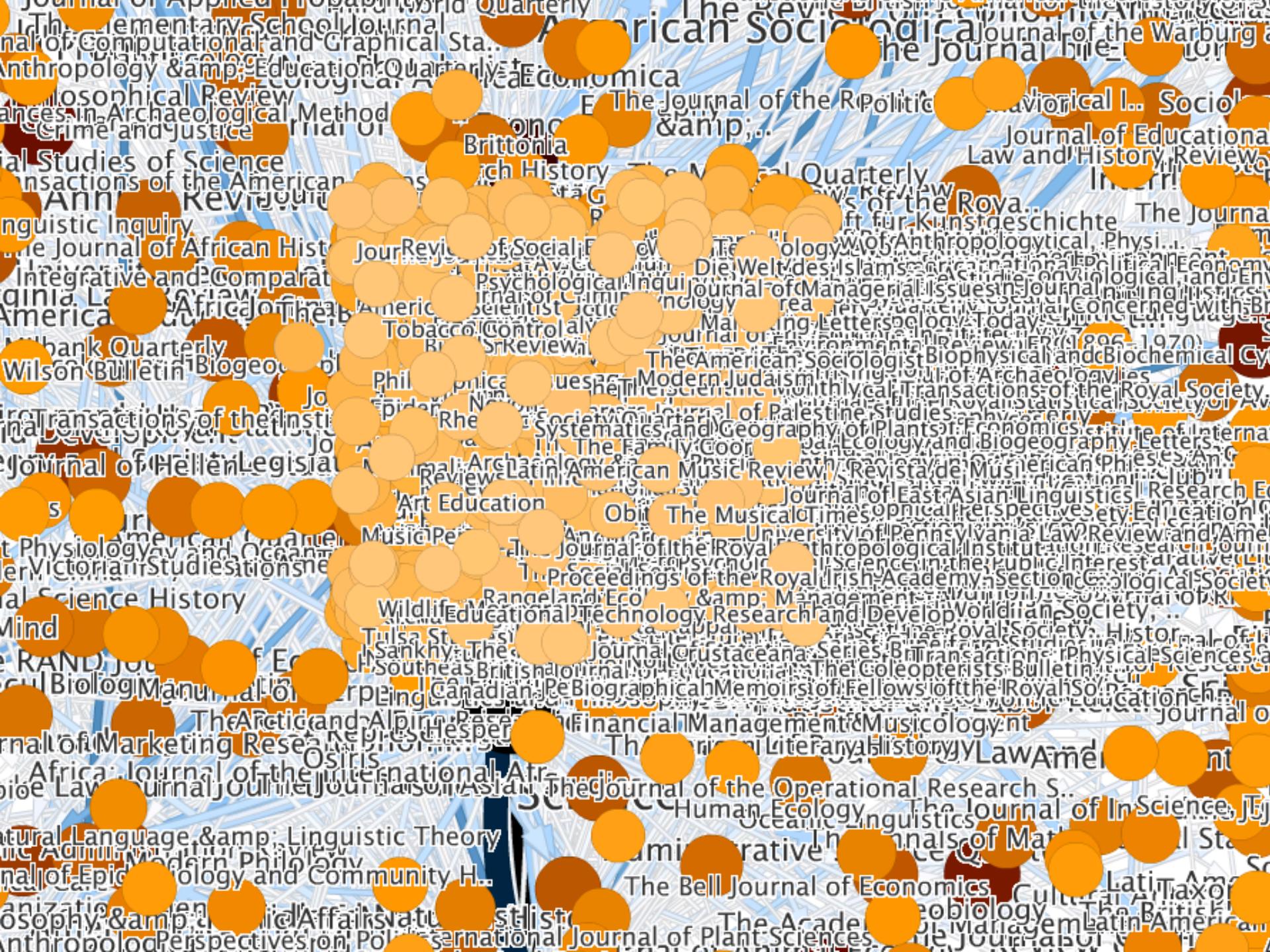


Years: 1600 - 2016



WIKIPEDIA
The Free Encyclopedia

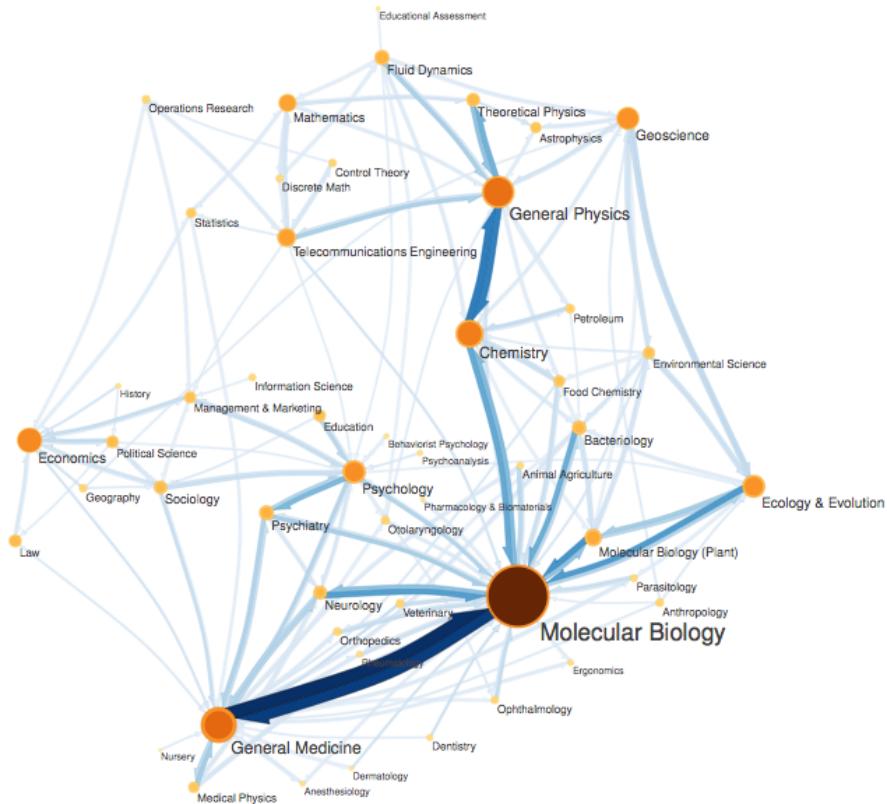




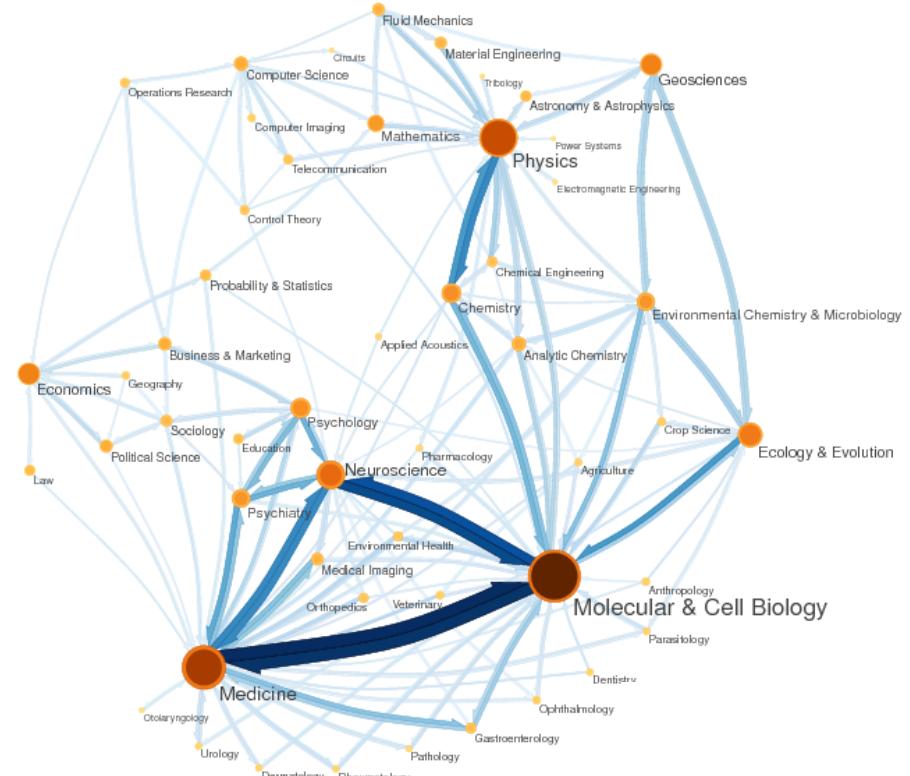
The map equation

$$L(M) = q_{\curvearrowright} H(Q) + \sum_{i=1}^m p_{\circlearrowleft}^i H(\mathcal{P}^i)$$

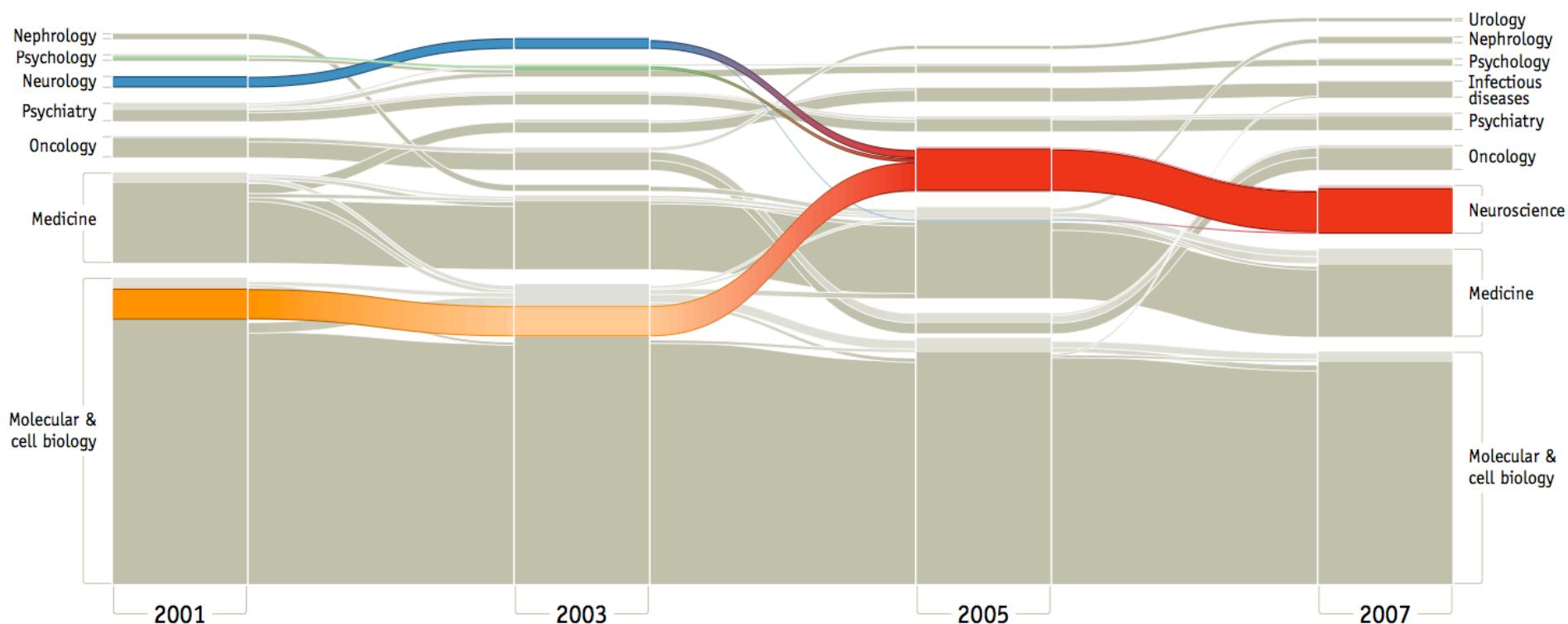
1995

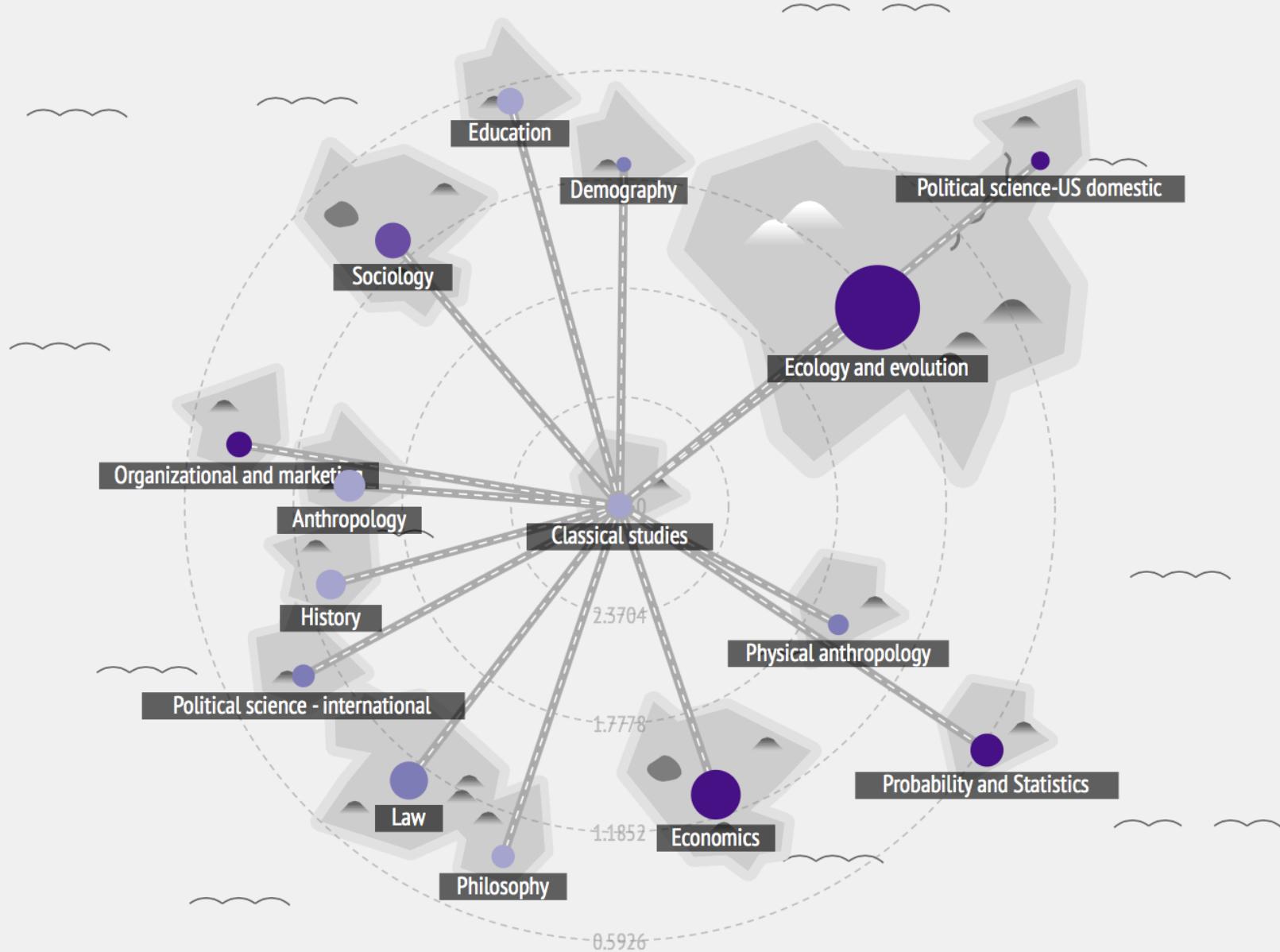


2004

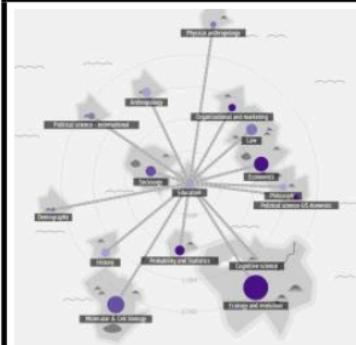
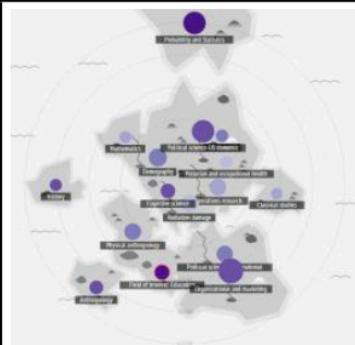
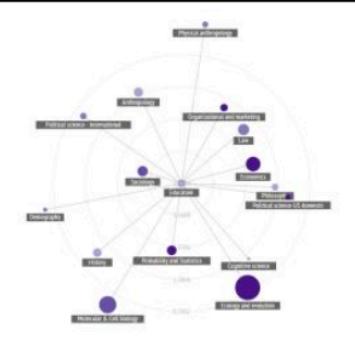
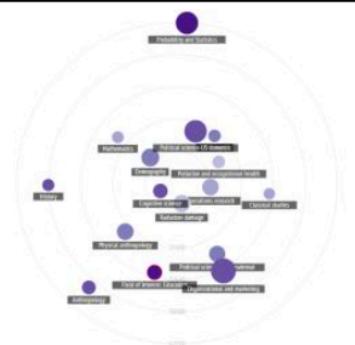
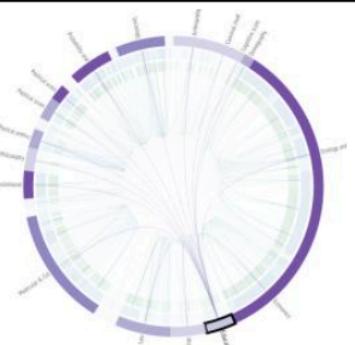


Citation networks over time

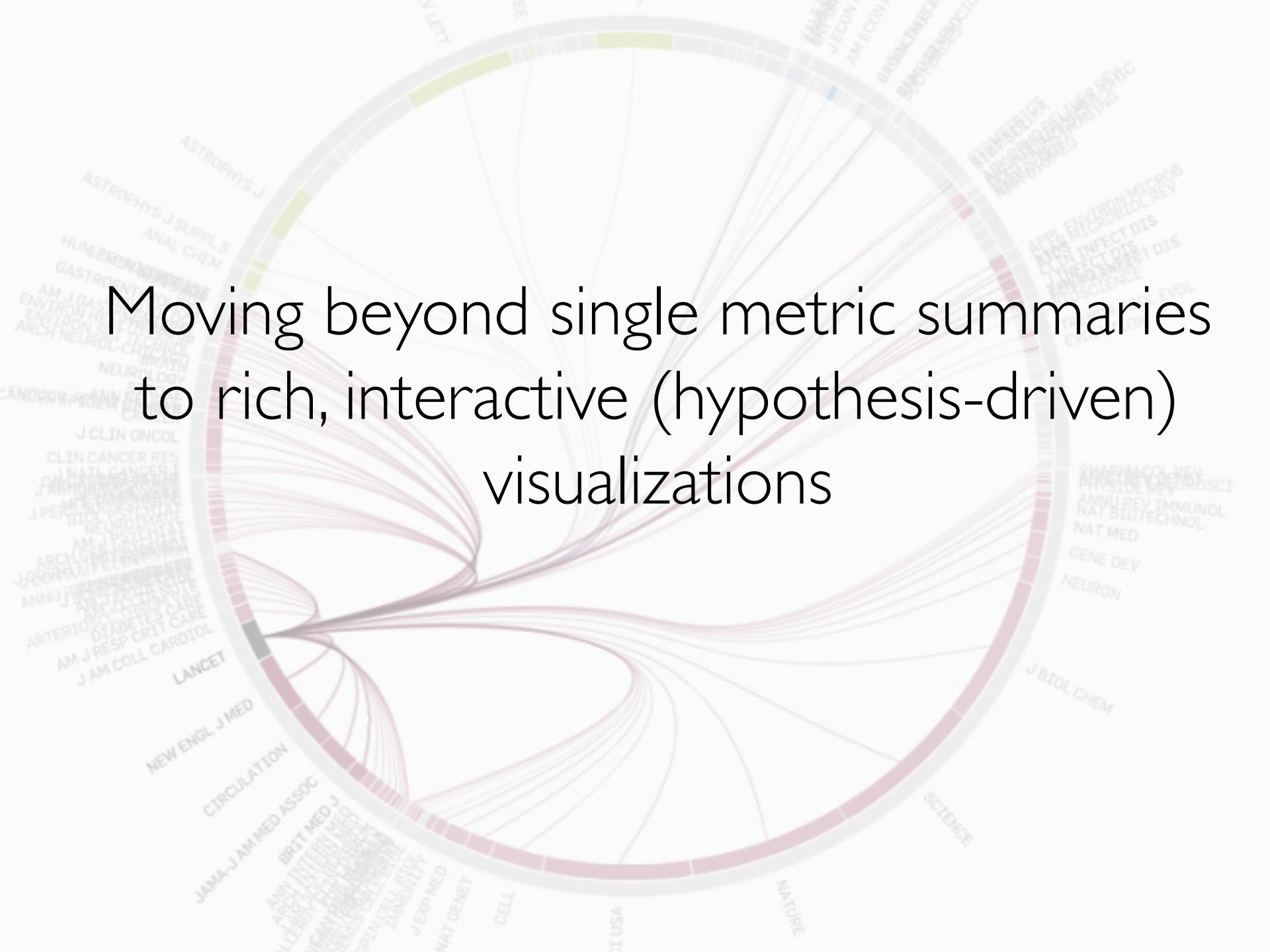




Navigating Hierarchical Knowledge Networks

1. Congruent Landscape	2. Incongruent Landscape	3. Congruent Abstract	4. Incongruent Abstract	5. Designer Baseline
 <p>Landscape visualization with data properties mapped to visual elements according to applicable image schemata</p>	 <p>Landscape visualization with data properties mapped to visual elements deliberately breaking with image schemata</p>	 <p>Identical to the Congruent Landscape tool but with all realistic details and overt “landscape” visuals removed</p>	 <p>Identical to the Incongruent Landscape tool but with all realistic details and overt “landscape” visuals removed</p>	 <p>Visualization designed by a hypothesis-blind designer attempting to make an effective visualization but without special emphasis on metaphor</p>

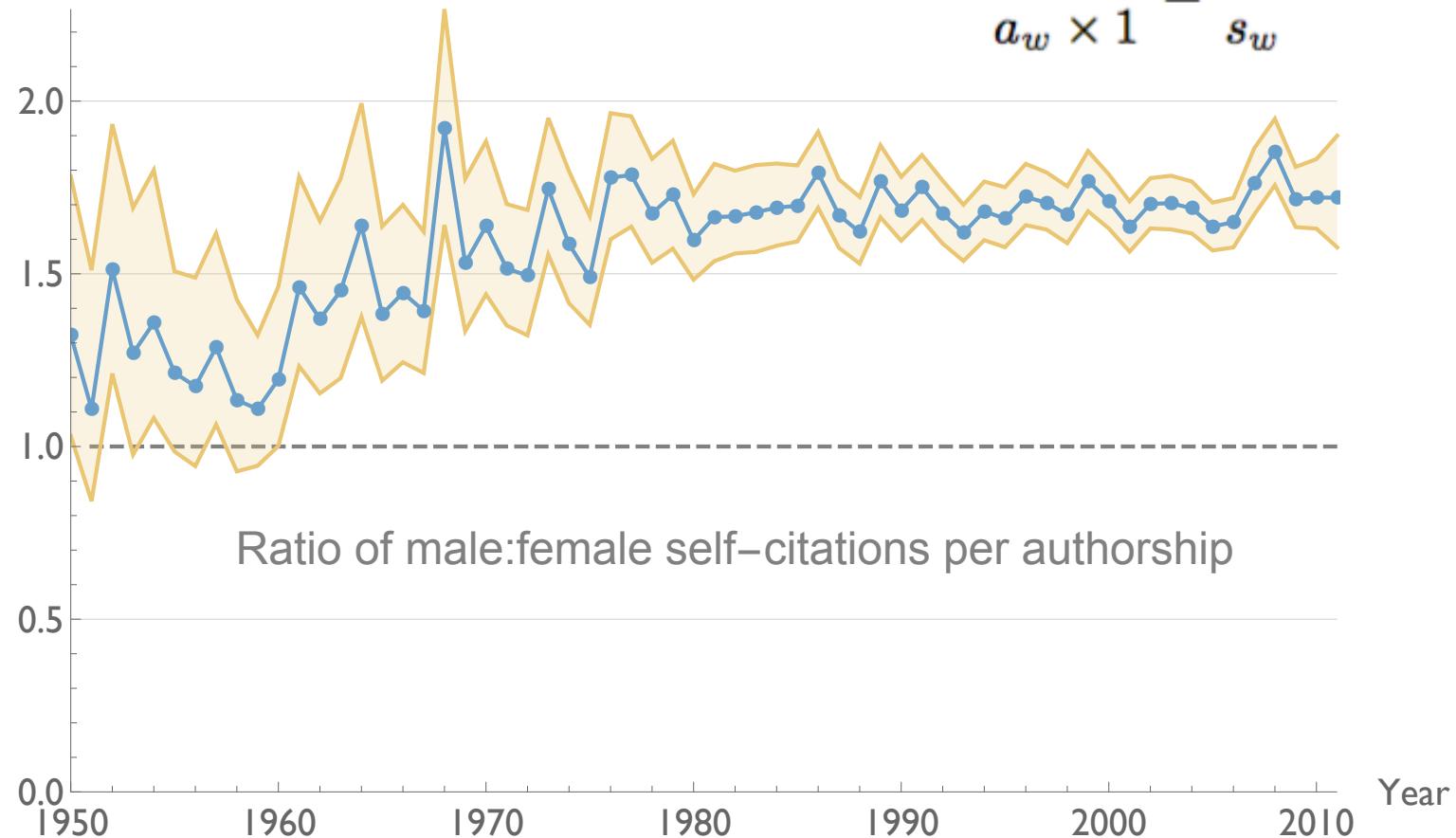
Moving beyond single metric summaries to rich, interactive (hypothesis-driven) visualizations



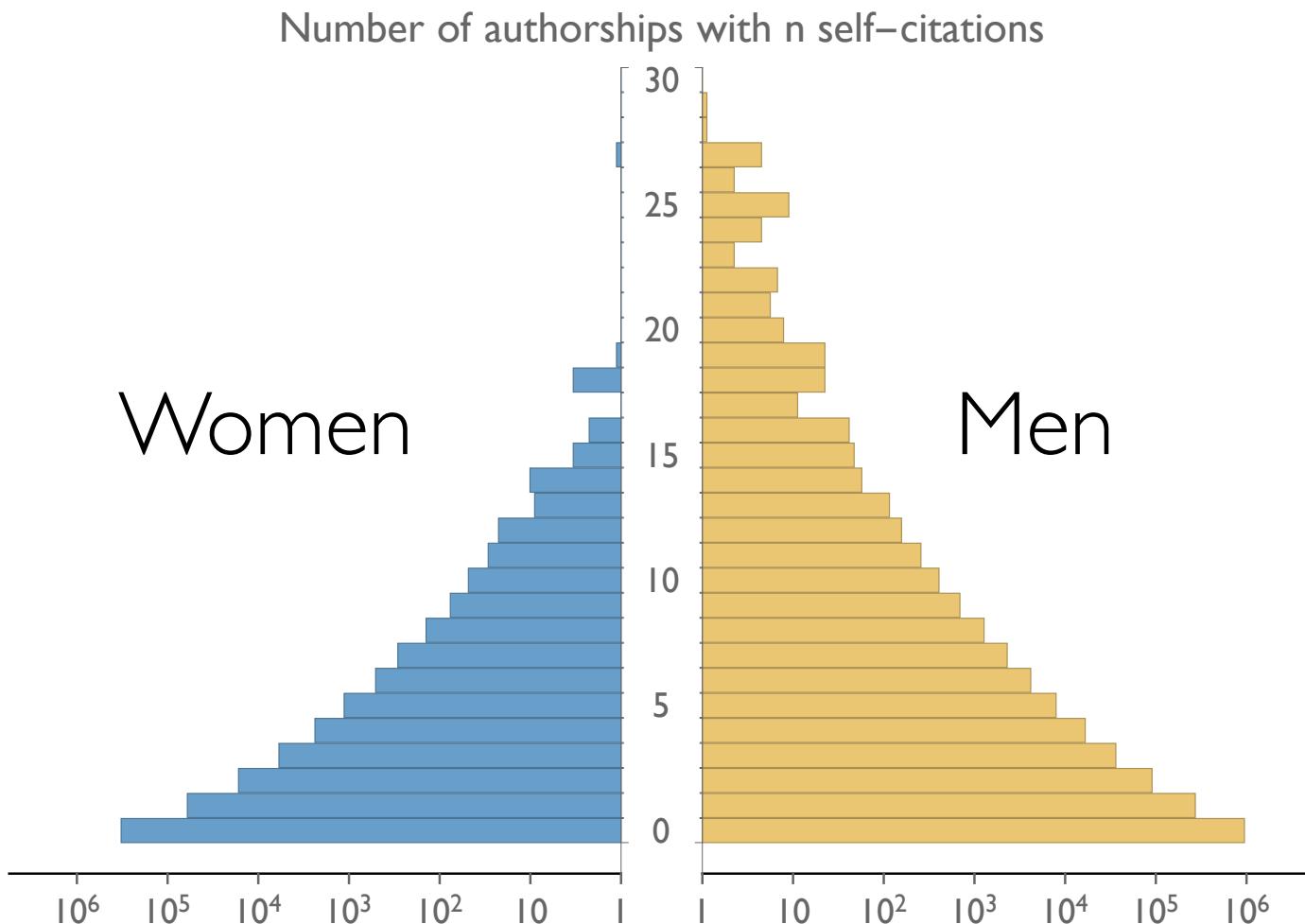
Self-citation over time

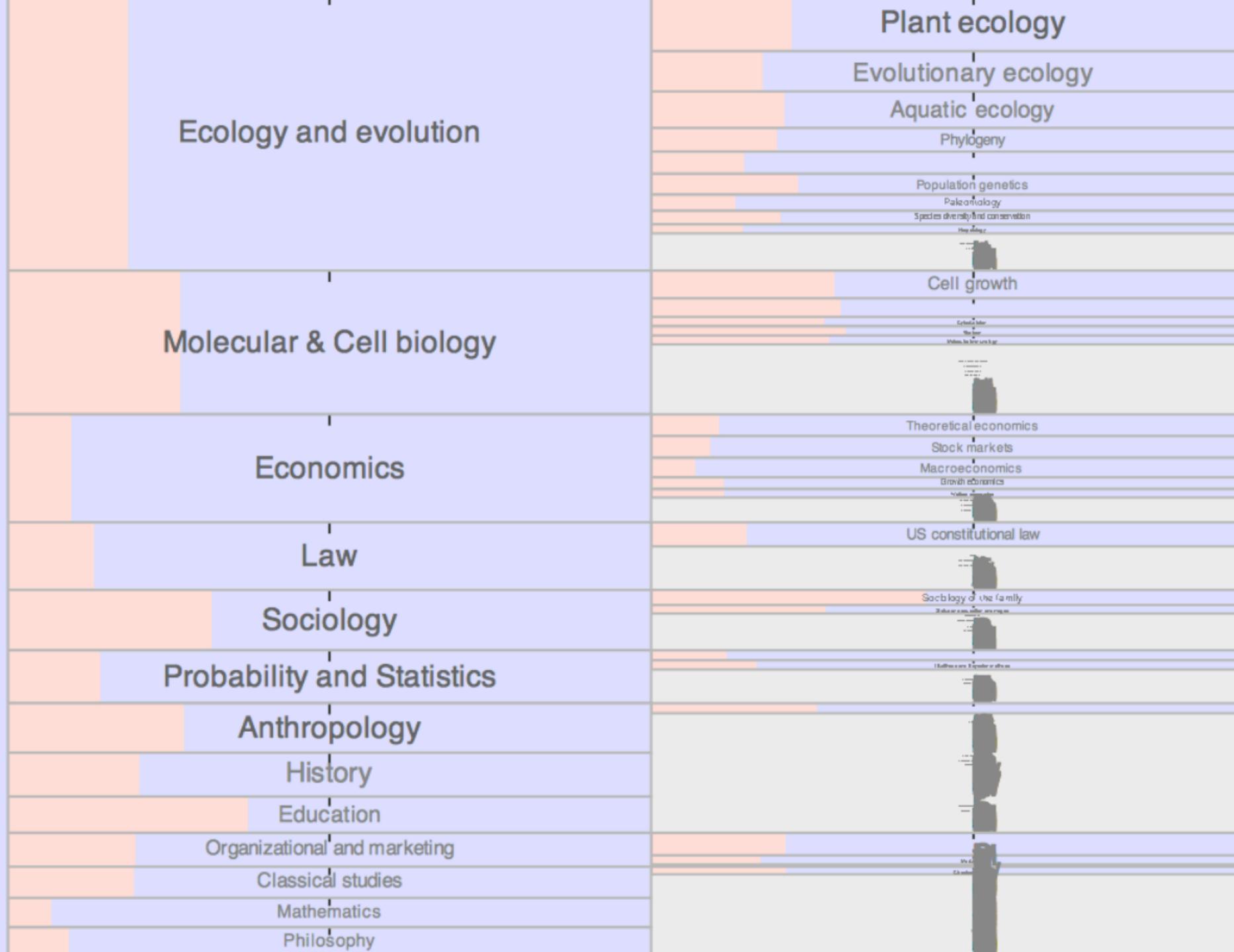
Ratio M:F

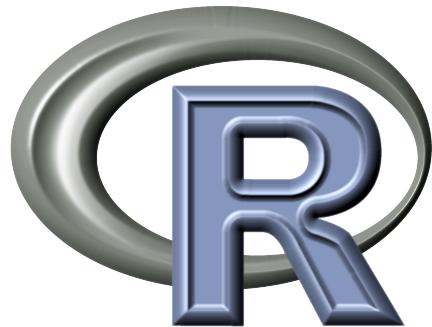
$$\frac{a_m \times k}{a_w \times 1} = \frac{s_m}{s_w}$$



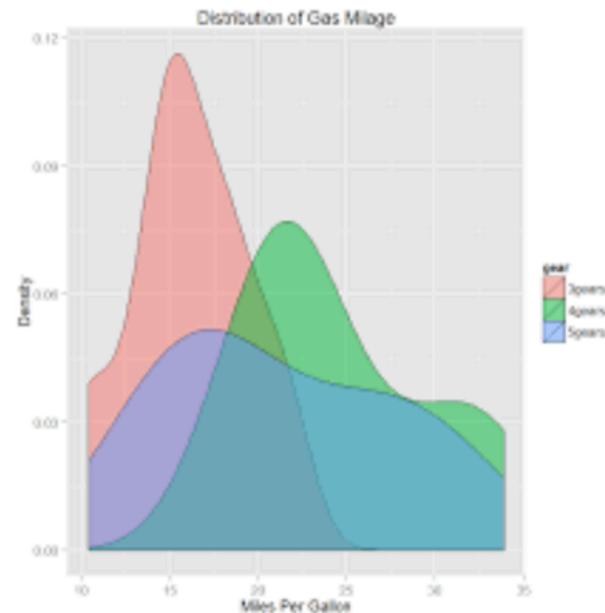
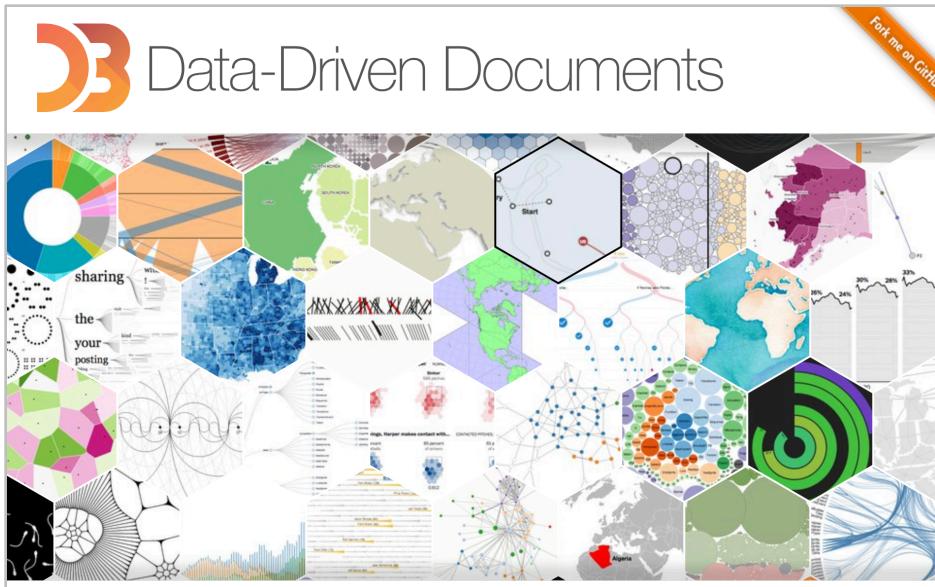
Gender and Self Citation







ggplot in R

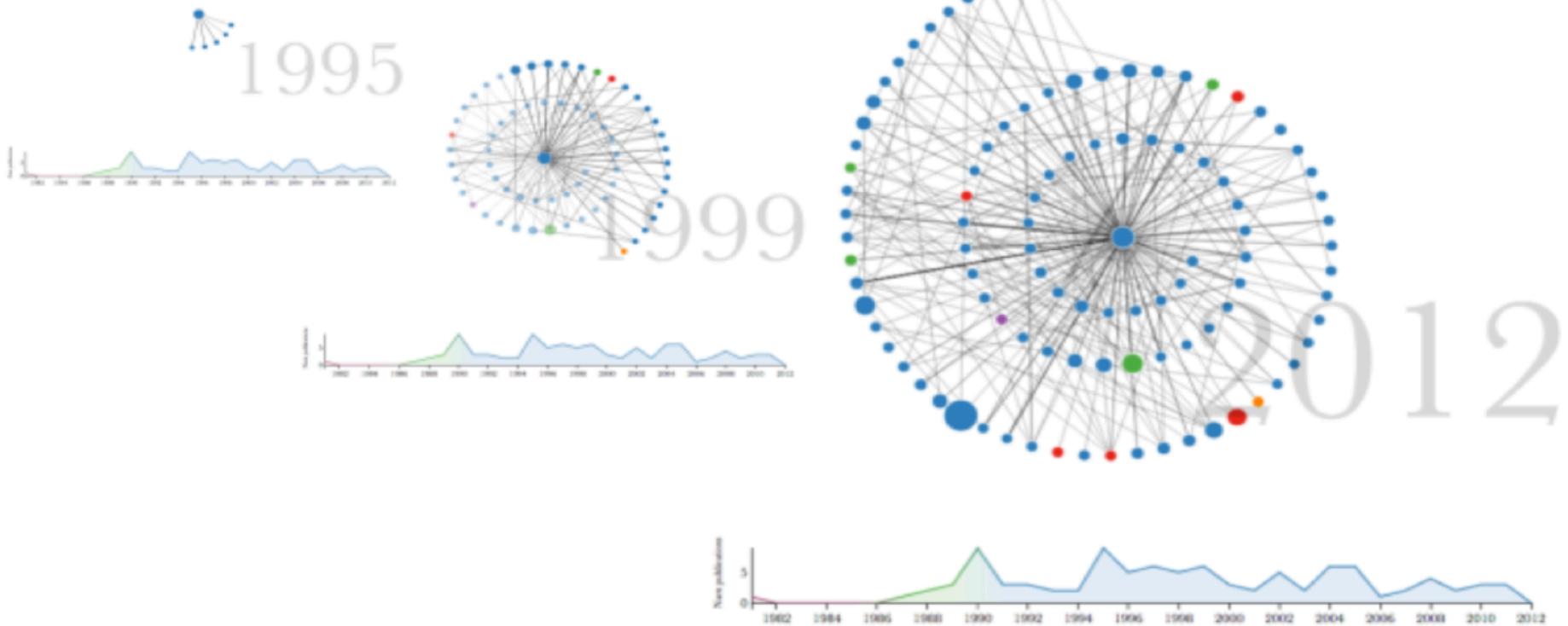


Summary

- Study the *Science of Science* and use visualizations for exploring questions
- Ask questions about the origin and evolution of ideas and fields, interdisciplinarity, impact assessment and sociology of science
- Single metrics to interactive visualization
- Building statistical and visualization tools that improve navigation, make relevant connections and facilitate knowledge discovery
- Tools: R, python, D3
- Challenges: scaling, mechanism
- Eigenfactor.org, Viziometrics.org, Scholar.eigenfactor.org

Explore the data

scholar.eigenfactor.org



* Please use Chrome web browser for best results

scholar.eigenfactor.org

username: PewScholar

password: 1N!kdG

Jevin West, jevinw@uw.edu

Acknowledgements

Jason Portenoy, Information School, University of Washington

Bill Howe, Information School, University of Washington

Jody Roberts , Chemical Heritage Foundation

Martin Rosvall, Department of Physics, Umea University

Carl Bergstrom, Department of Biology, University of Washington

jevinw@uw.edu

jevinwest.org

Eigenfactor.org

@jevinwest