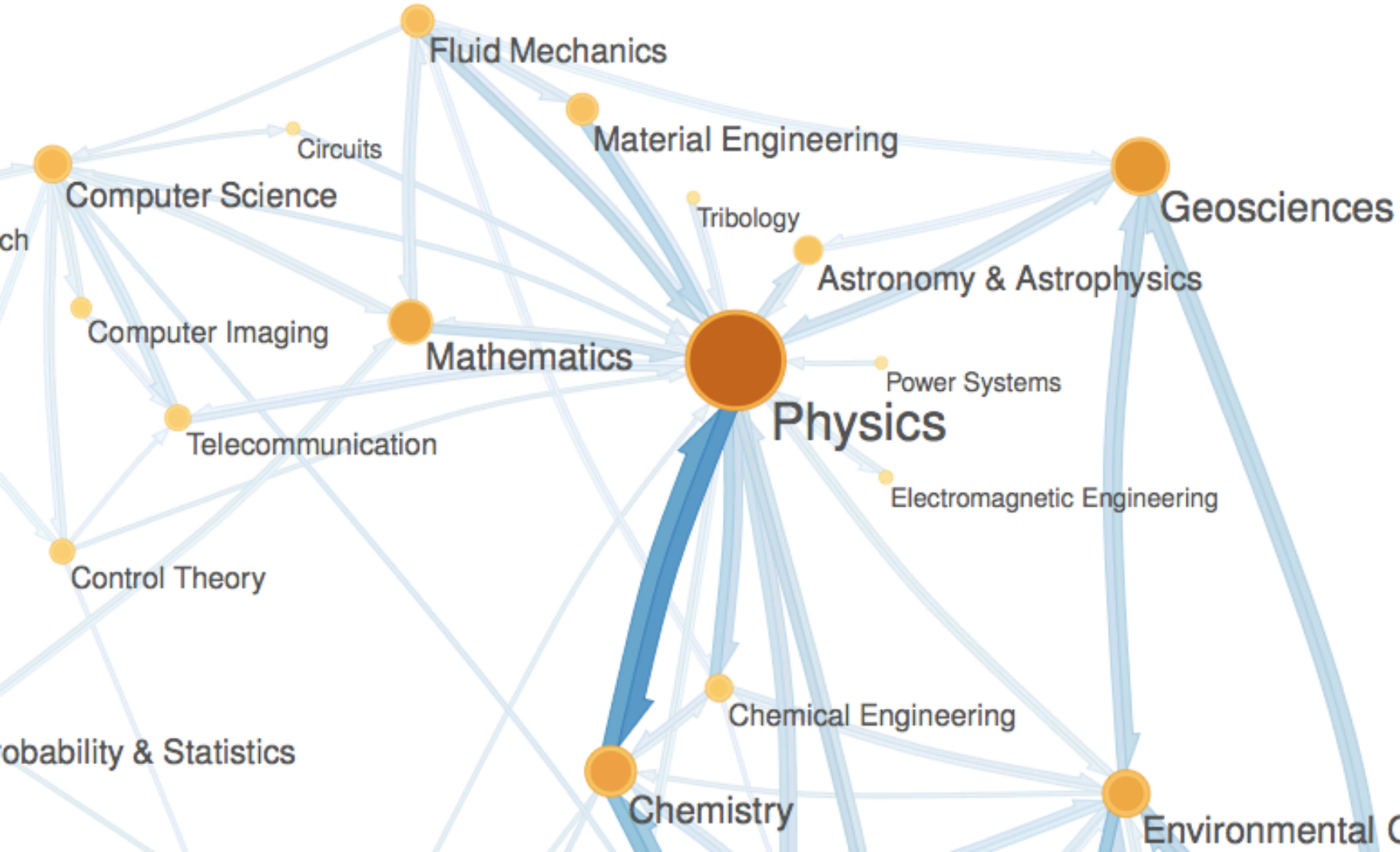
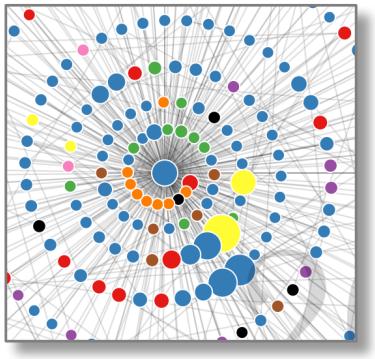
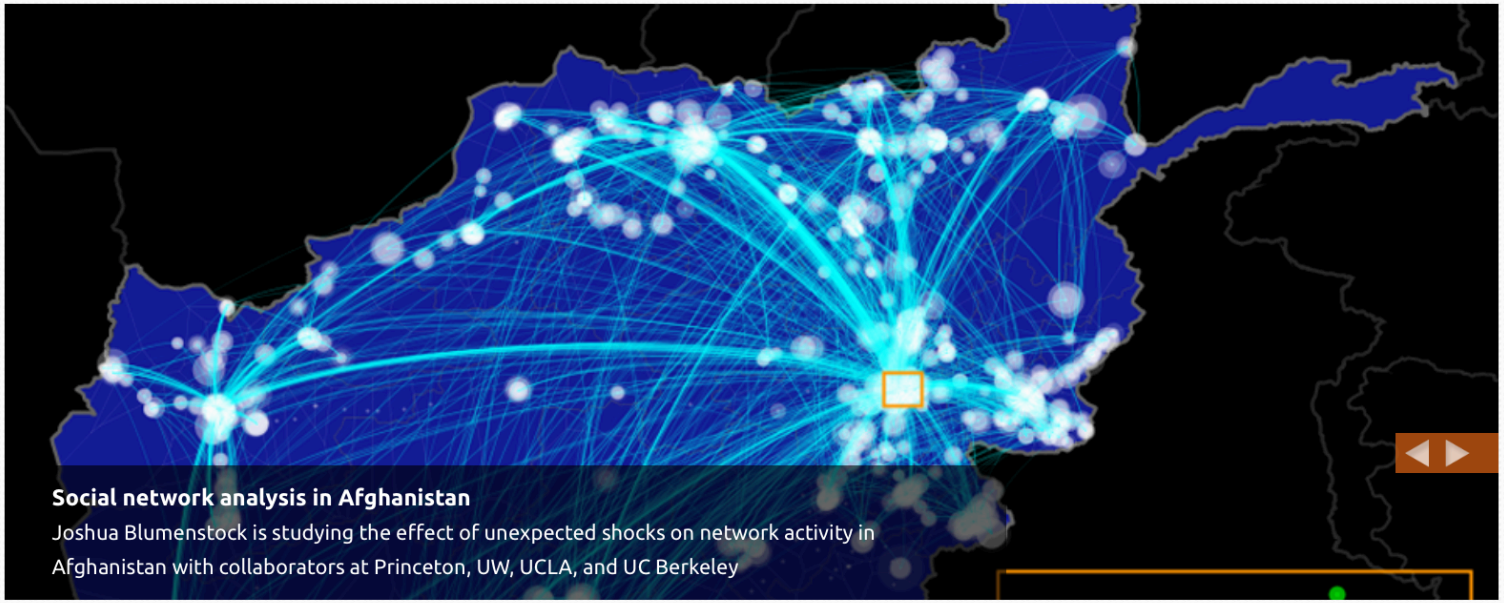
The Science of Science

Jevin West, Information School, University of Washington



Science of Science

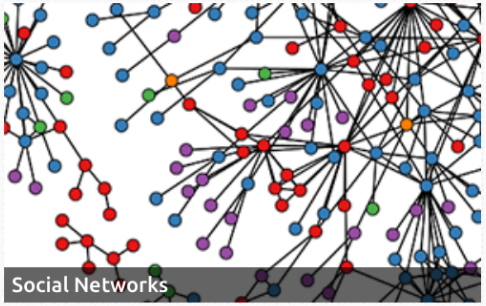


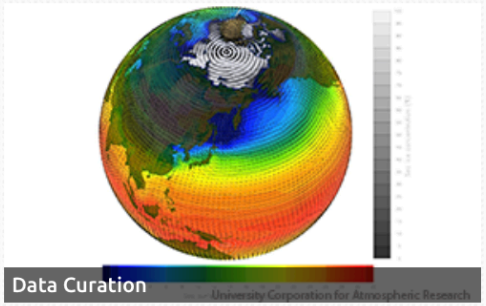


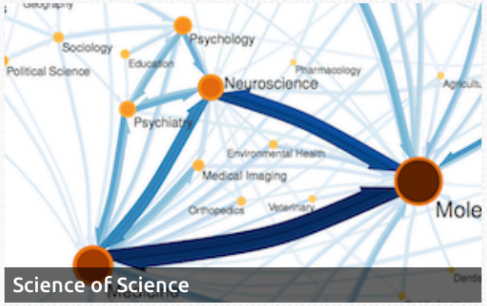
Social network analysis in Afghanistan

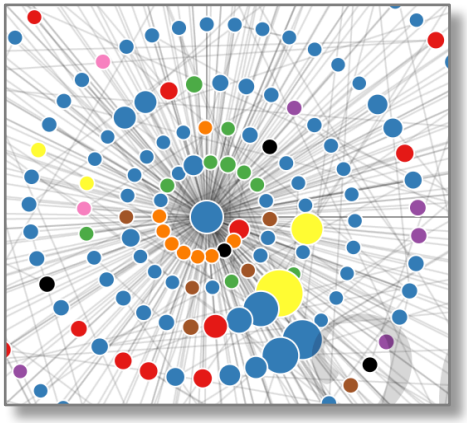
Joshua Blumenstock is studying the effect of unexpected shocks on network activity in Afghanistan with collaborators at Princeton, UW, UCLA, and UC Berkeley

Research Focus Areas

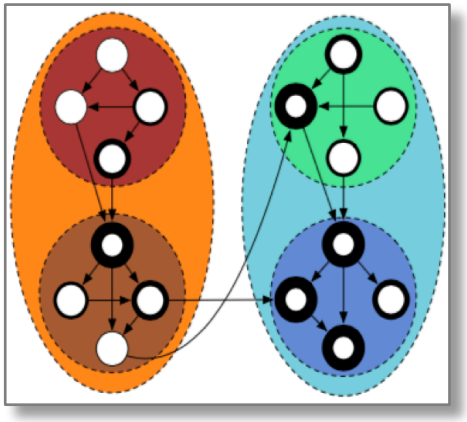






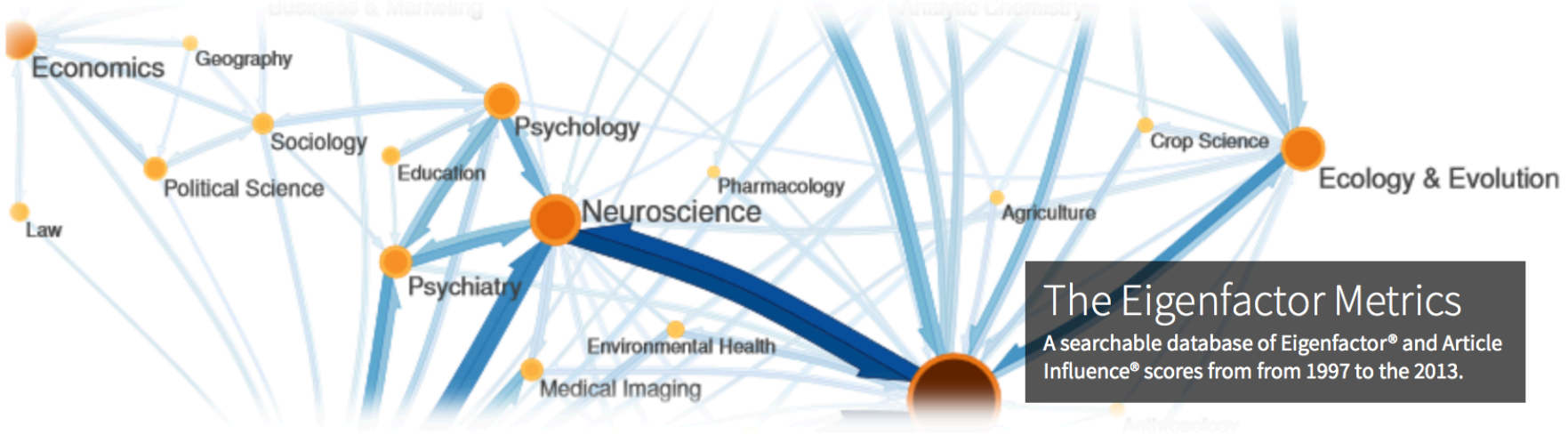


Knowledge Science



Knowledge Engineering

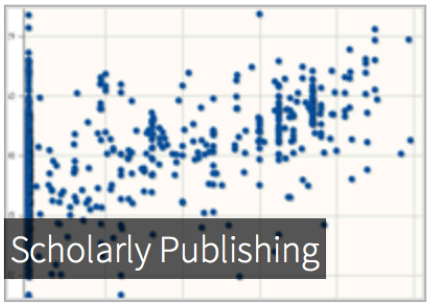
EIGENFACTOR.org



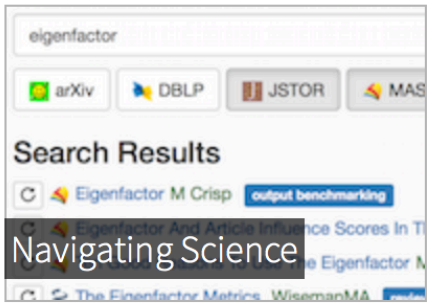
The Eigenfactor Metrics

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

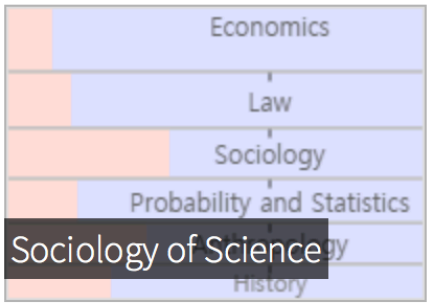
RESEARCH AREAS


Scholarly Publishing

Mapping Science



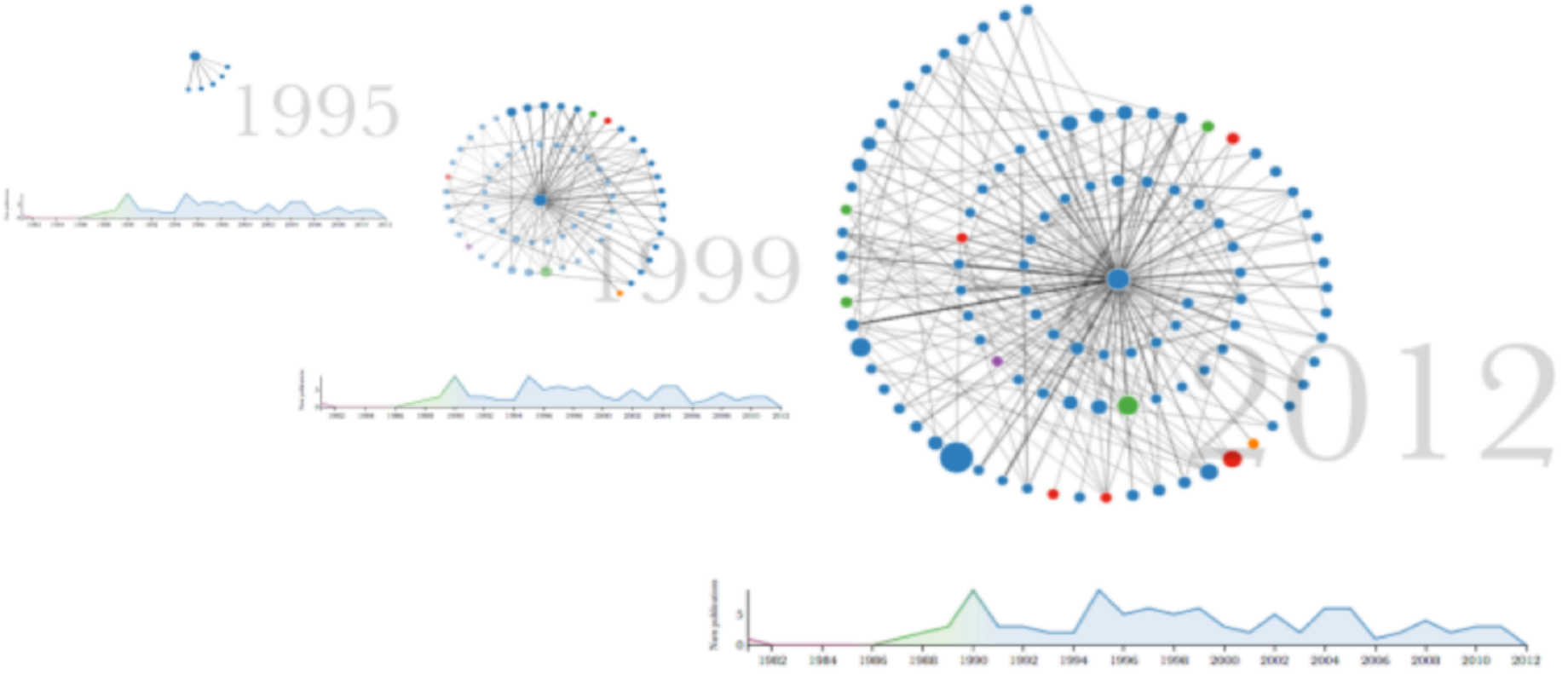
Navigating Science



Sociology of Science

Explore the data

scholar.eigenfactor.org



* Please use Chrome web browser for best results

Maladies of Science…

IS THERE A REPRODUCIBILITY CRISIS?

Open access, freely available online

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance. Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias. In this essay, I discuss the implications of these problems for the conduct and interpretation of research.

factors that influence this problem and some corollaries thereof.

Modeling the Framework for False Positive Findings

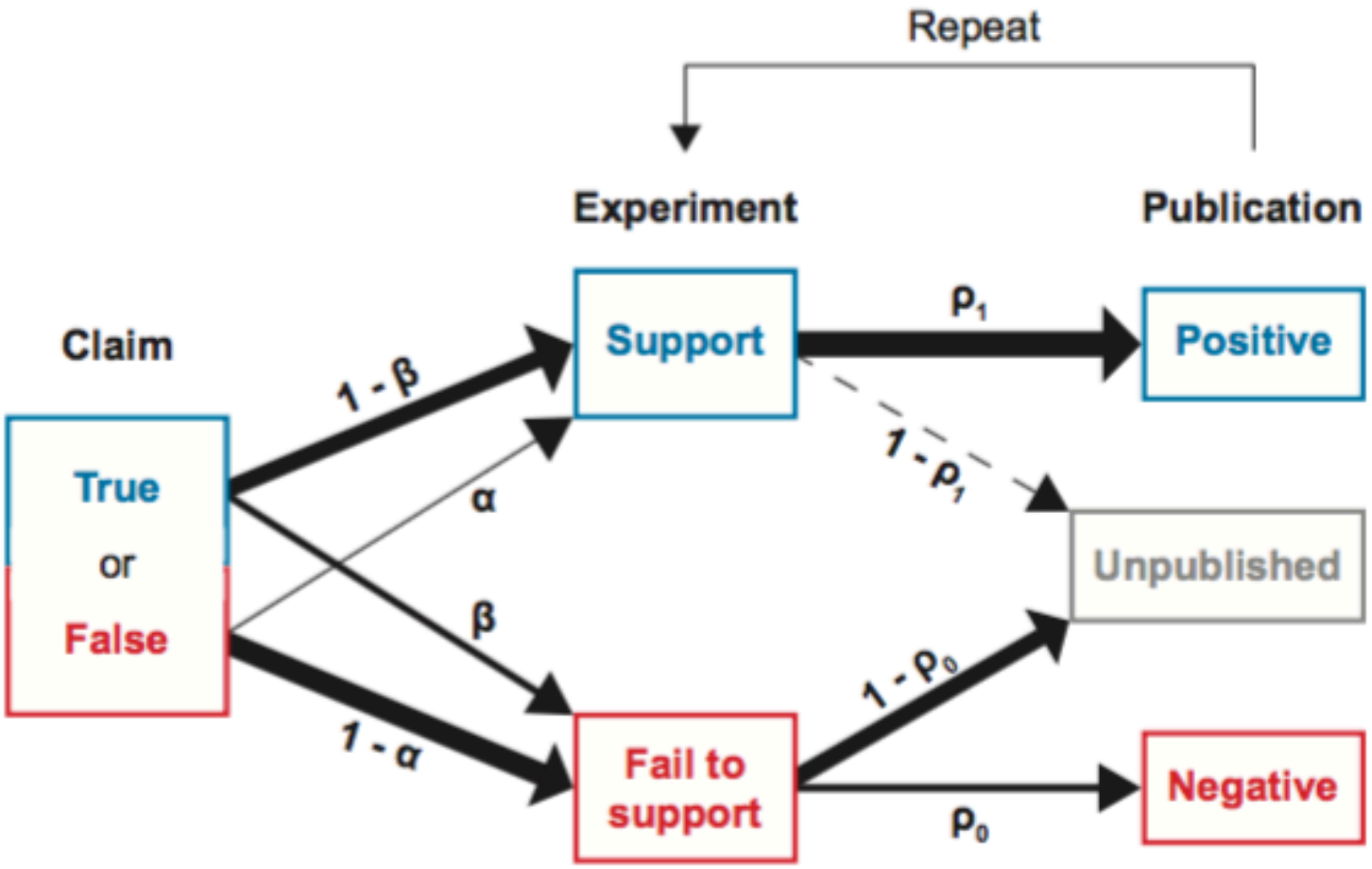
Several methodologists have pointed out [9–11] that the high rate of nonreplication (lack of confirmation) of research discoveries is a consequence of the convenient, yet ill-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a *p*-value less than 0.05. Research is not most appropriately represented and summarized by *p*-values, but, unfortunately, there is a widespread notion that medical research articles

It can be proven that most claimed research findings are false.

should be interpreted based only on *p*-values. Research findings are defined here as any relationship reaching formal statistical significance, e.g., effective interventions, informative predictors, risk factors, or associations. "Negative" research is also very useful.

is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is $R/(R+1)$. The probability of a study finding a true relationship reflects the power $1 - \beta$ (one minus the Type II error rate). The probability of claiming a relationship when none truly exists reflects the Type I error rate, α. Assuming that c relationships are being probed in the field, the expected values of the 2×2 table are given in Table 1. After a research finding has been claimed based on achieving formal statistical significance, the post-study probability that it is true is the positive predictive value, PPV. The PPV is also the complementary probability of what Wacholder et al. have called the false positive report probability [10]. According to the 2 $\times$ 2 table, one gets PPV = $(1 - \beta)R/(R$

Publication bias and the canonization of false facts



Nissen, Magidson, Gross & Bergstrom (2016) *arXiv:1609.00494*

Maladies of Science

Process

Incentives

The H-index impact on science

Jure Leskovec

Follow

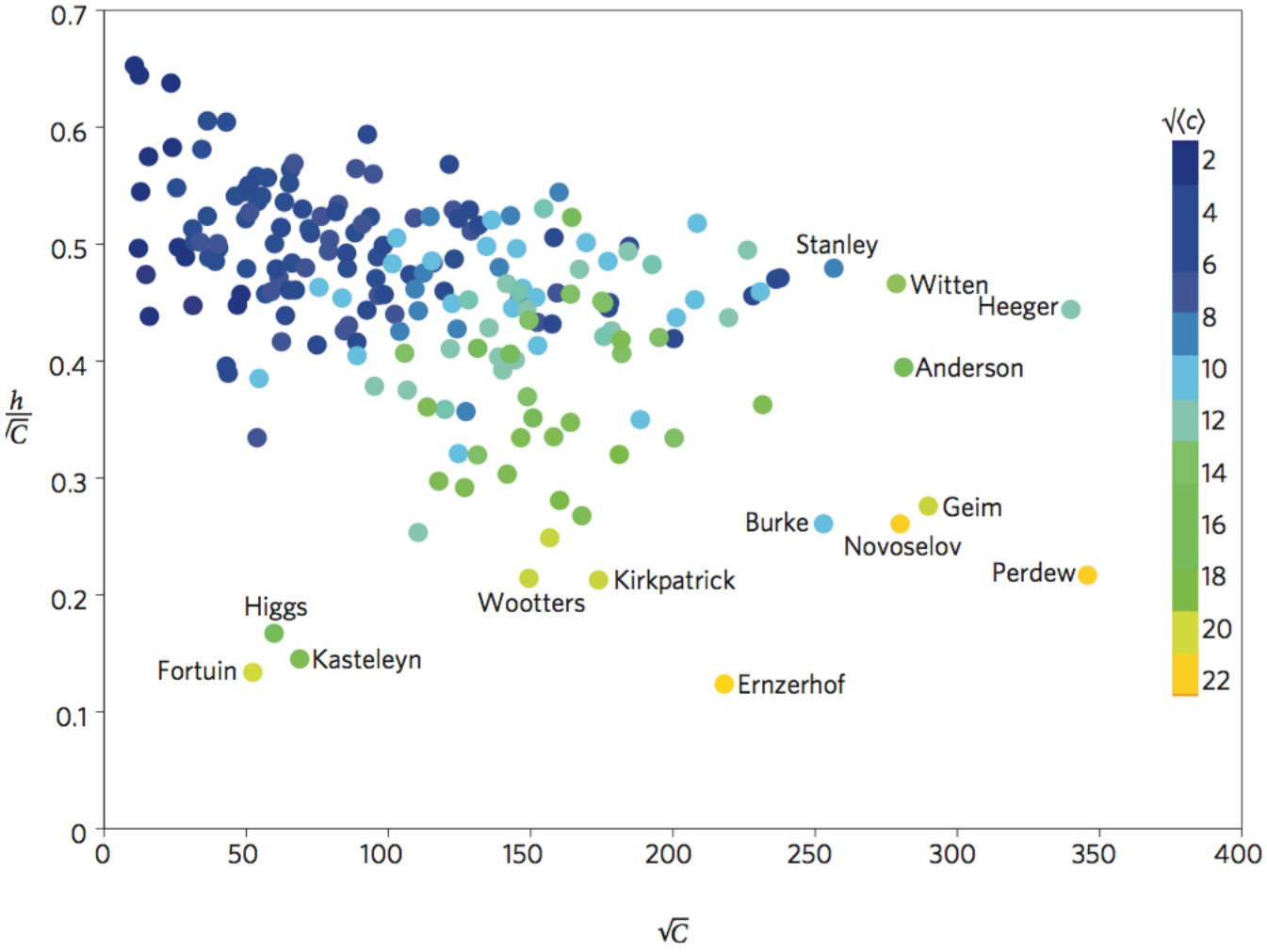
Professor of Computer Science, Stanford University
Data mining, Social Network Analysis, Information Networks
Verified email at cs.stanford.edu - Homepage

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		

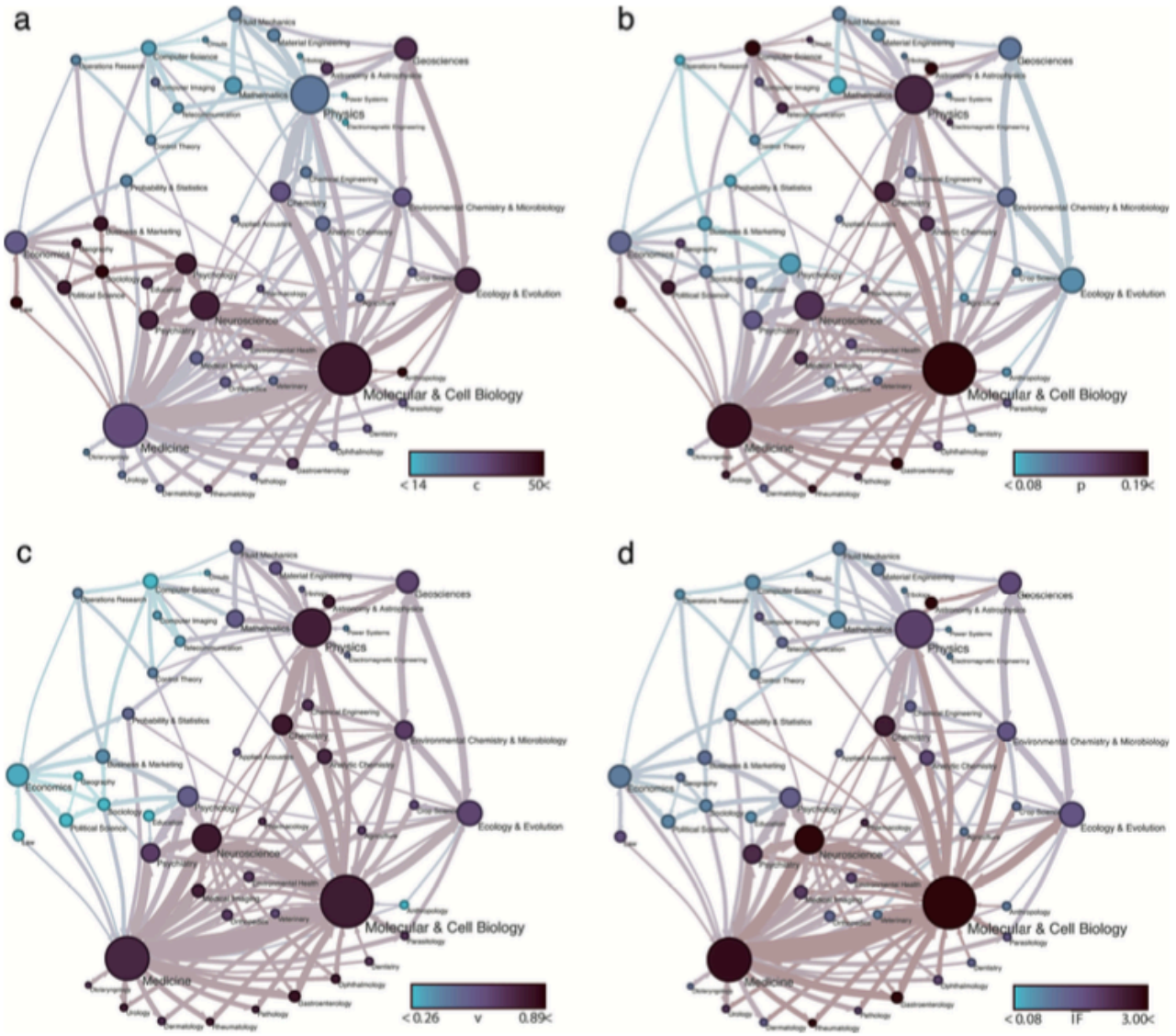
Citation indices	All	Since 2010
Citations	19409	17853
h-index	59	56
i10-index	103	101

Evisceration of the H-index



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

Impact Factor Inflation



Althouse et al (2009) JASIST

Impact factor drives…

The worse misuse of the H-index

You don't know me, but I am a fellow academic…...

I also noticed you have pretty good research prestige, as indicated by your **H-index** of 15. Even one paper of yours has been cited 170 times! This is great considering you are pretty new at academia.

My researcher prestige is not bad either, but I've been at it a bit longer: LINK

…

I know how hard it is for academics who are single to find a mate (I've been single for a long time). …if there is an off-chance that you are single, please pick me. I like your work. It is very interesting and I think you are quite attractive. I would like to go out on a date with you, what do you say? Let's start a wonderful romance.

San Francisco DORA
Declaration on Research Assessment

The San Francisco Declaration on Research Assessment (DORA), initiated by the American Society for Cell Biology (ASCB) together with a group of editors and publishers of scholarly journals, recognizes the need to improve the ways in which the outputs of scientific research are evaluated. The group met in December 2012 during the ASCB Annual Meeting in San Francisco and subsequently circulated a draft declaration among various stakeholders. DORA as it now stands has benefited from input by many of the original signers listed below. It is a worldwide initiative covering all scholarly disciplines. We encourage individuals and organizations who are concerned about the appropriate assessment of scientific research to sign DORA.

Download the Declaration (PDF) | Download the DORA Logo (PDF) | Download the DORA Poster (PDF)

San Francisco Declaration on Research Assessment

Putting science into the assessment of research

There is a pressing need to improve the ways in which the output of scientific research is evaluated by funding agencies, academic institutions, and other parties. To address this issue, a group of editors and publishers of scholarly journals met during the Annual Meeting of The American Society for Cell Biology (ASCB) in San Francisco, CA, on December 16, 2012. The group developed a set of recommendations, referred to as the San Francisco Declaration on

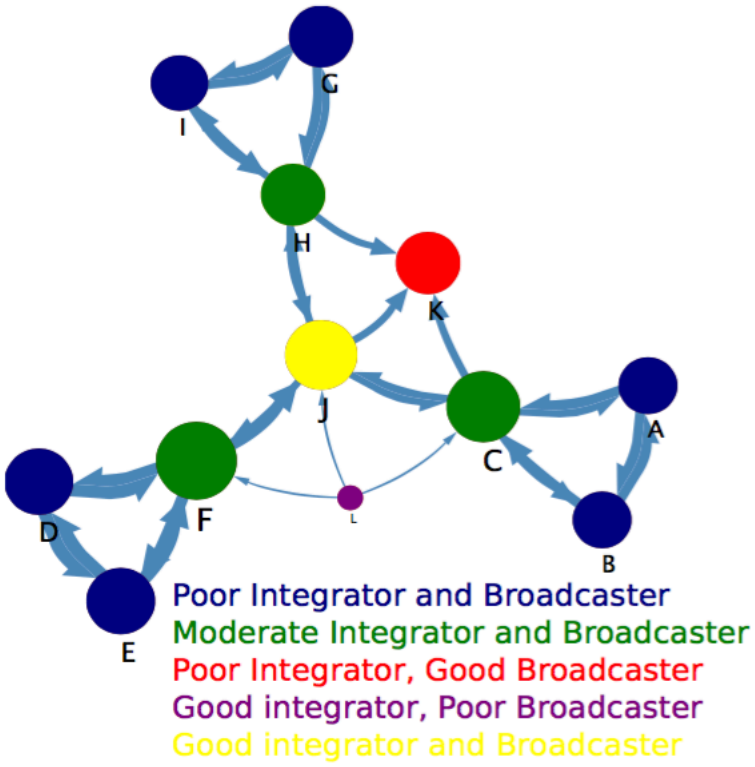
Stephen Curry on Why Universities Should Sign DORA

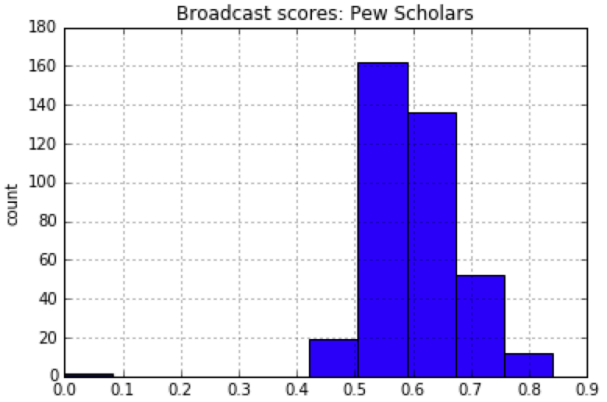
Read Now

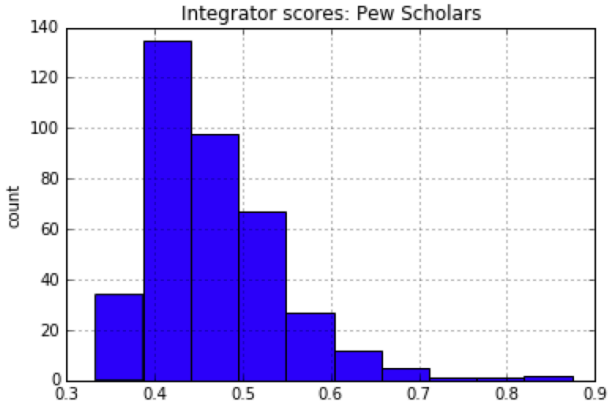
News about DORA (most recent news at top)

Bias Against Novelty in Science: A Cautionary Tale for Users of Bibliometric Indicators

Measuring Interdisciplinarity

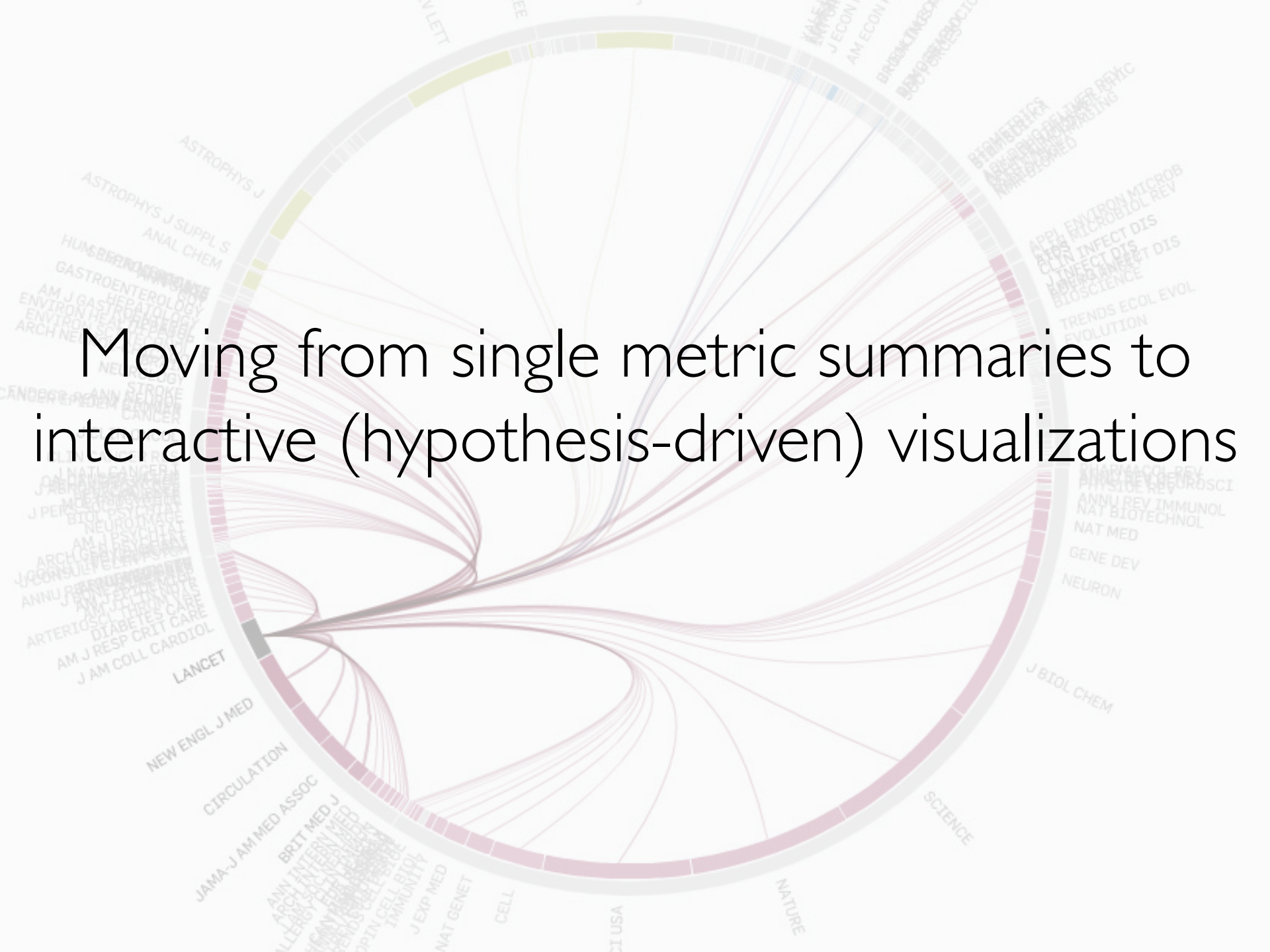






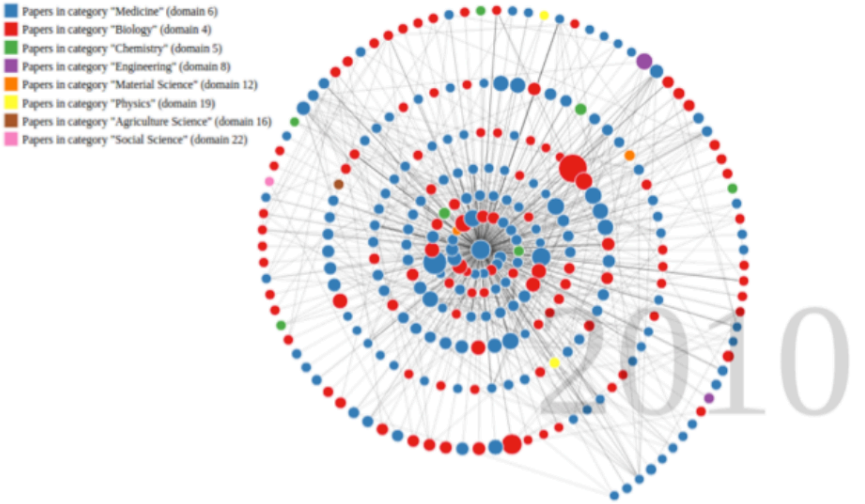
Bergstrom, CT, Foster, J, Portenoy, J, A. Misra,West, JD. (2016). Measuring interdisciplinarity without subject categories. (in prep)

Moving from single metric summaries to interactive (hypothesis-driven) visualizations



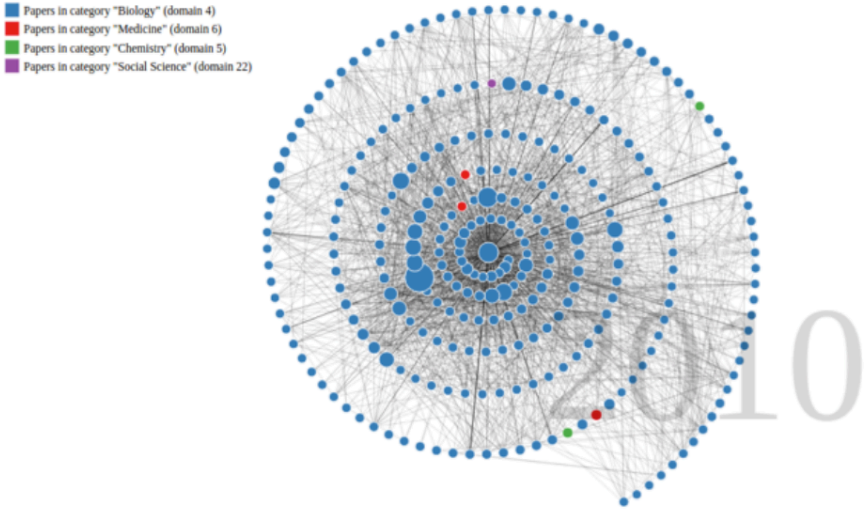
Visualizing Influence

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.

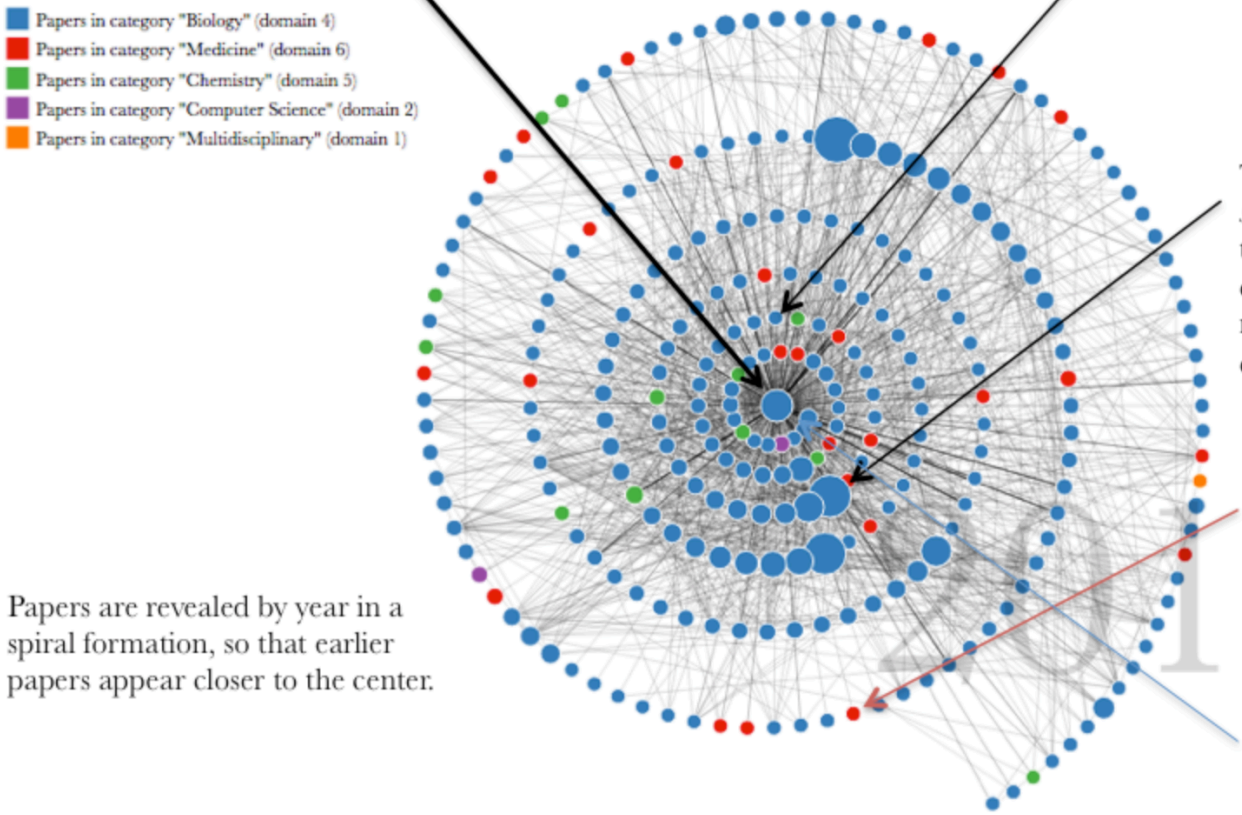


scholar.eigenfactor.org

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.

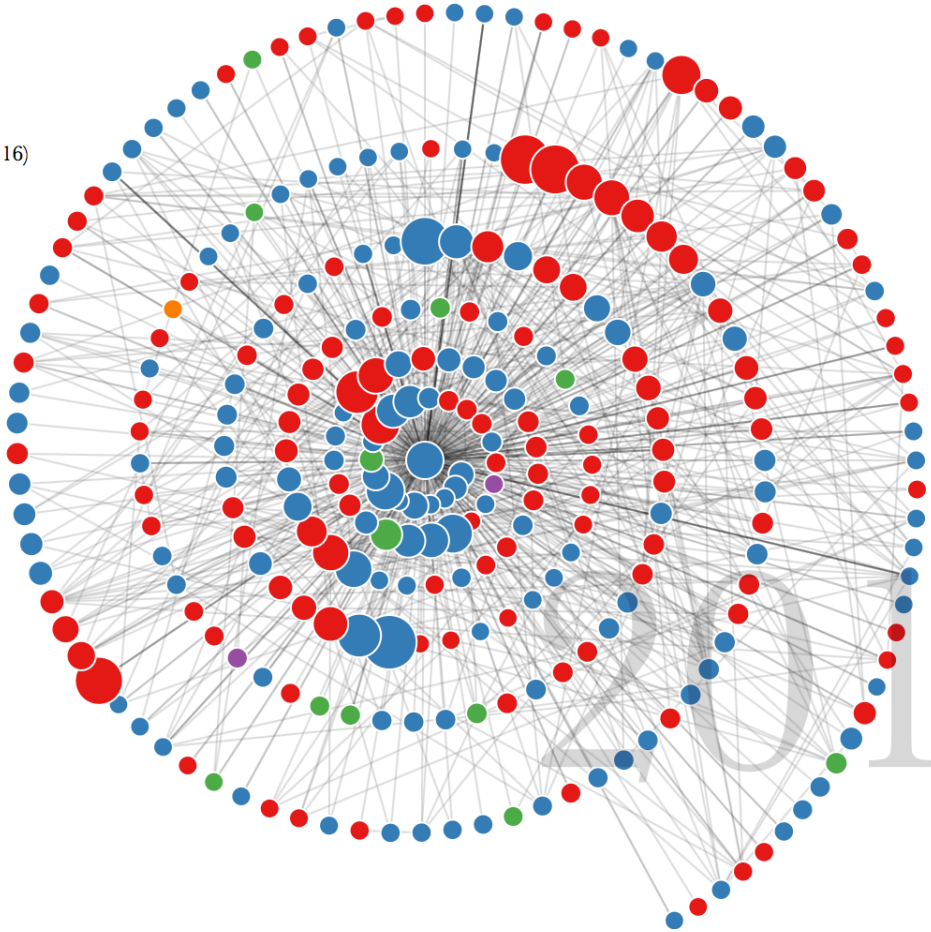
Visualizing Scholarly Influence Over Time

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

Visualizing Scholarly Influence Over Time

Influence of Pew Scholars

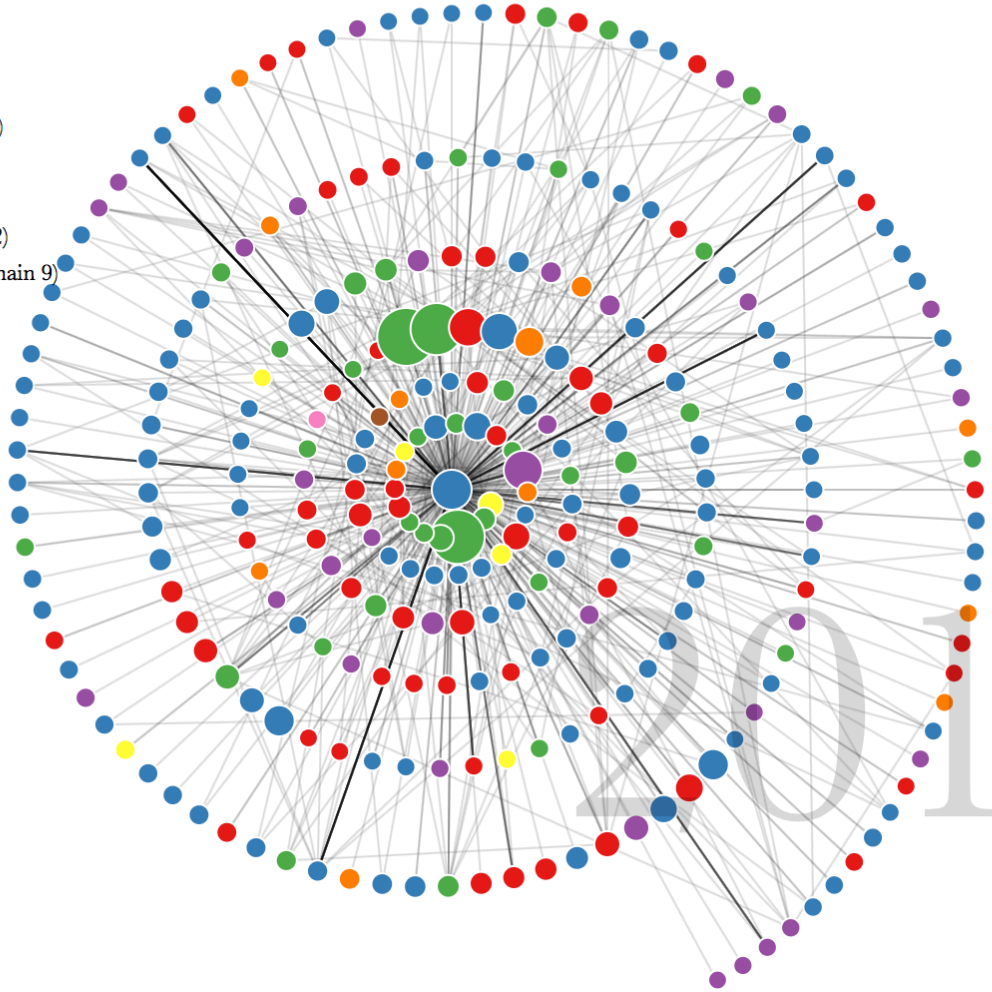
Mark W. Grinstaff

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)

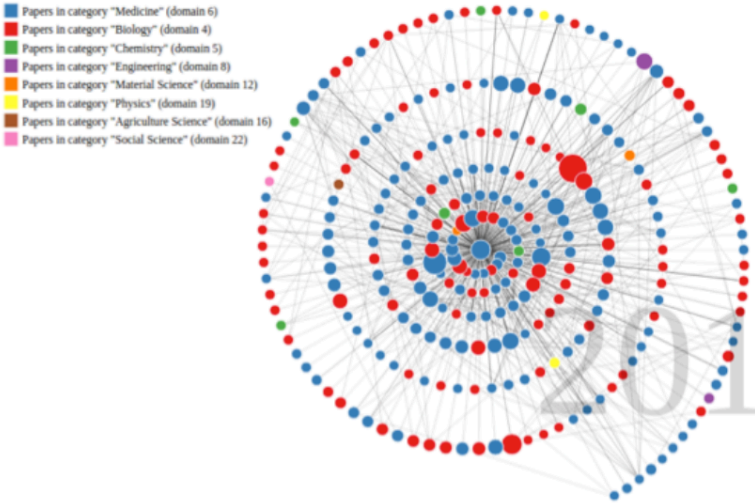
Mark W.
Grinstaff

Pew Scholar
1999



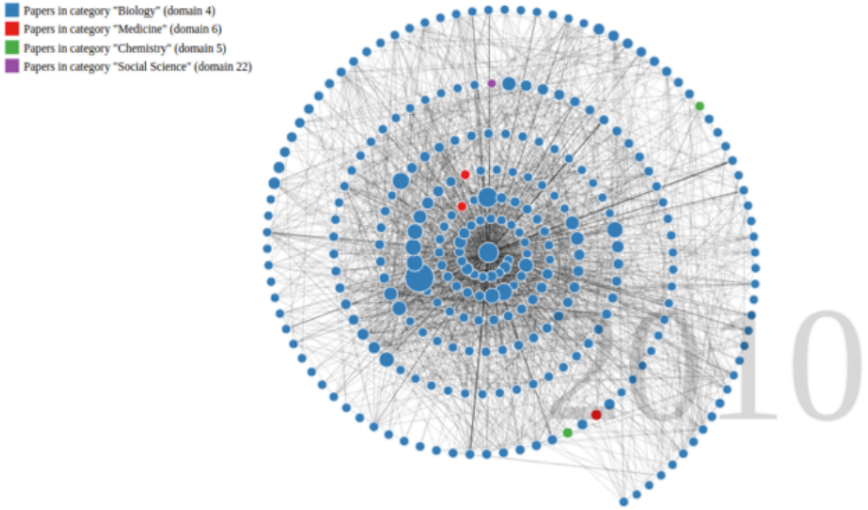
2012

Comparing Authors



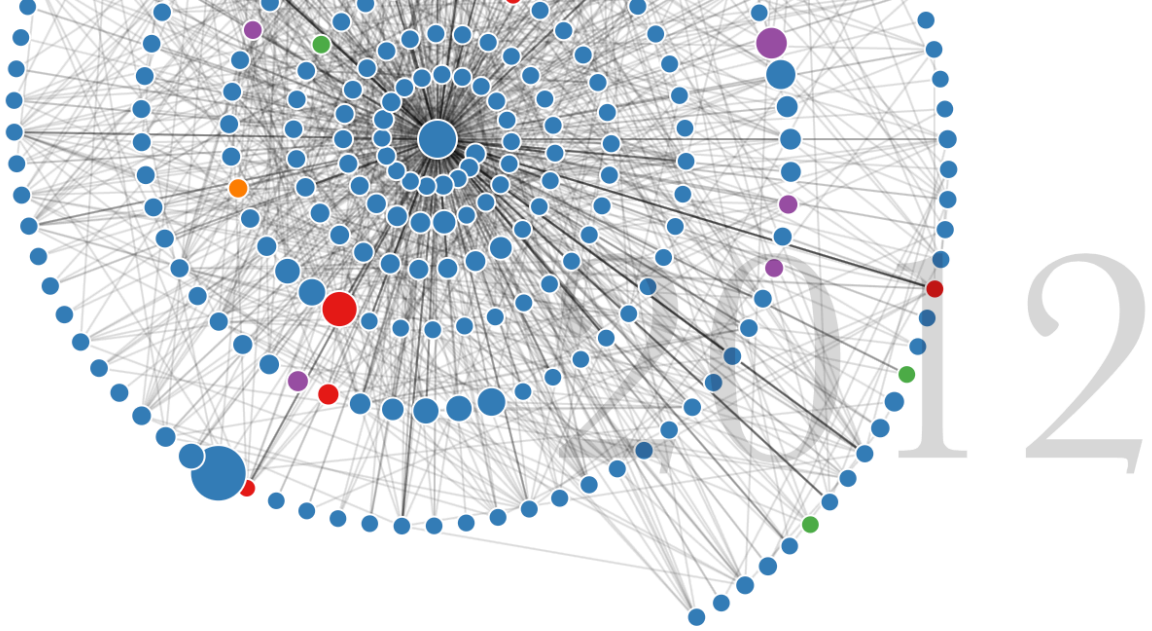
A more sparse network indicates fewer citations between papers shown in the network. This could be a result of the central scholar having impact across a wider set of academic communities.

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

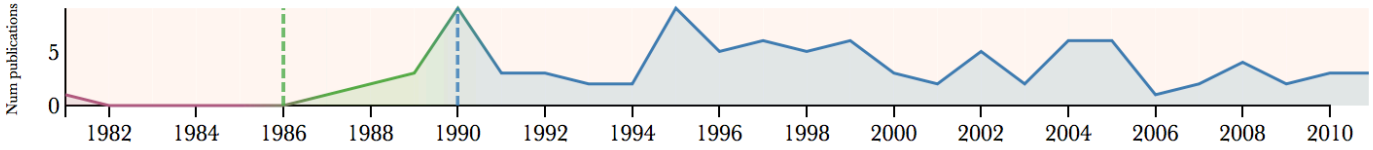


Philip A.
Hieter

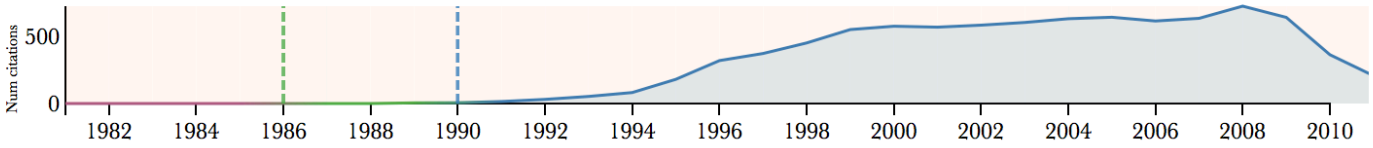
Pew Scholar
1986



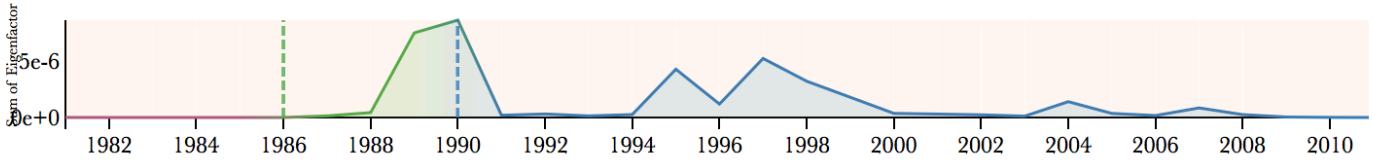
Number of publications



Number of citations received

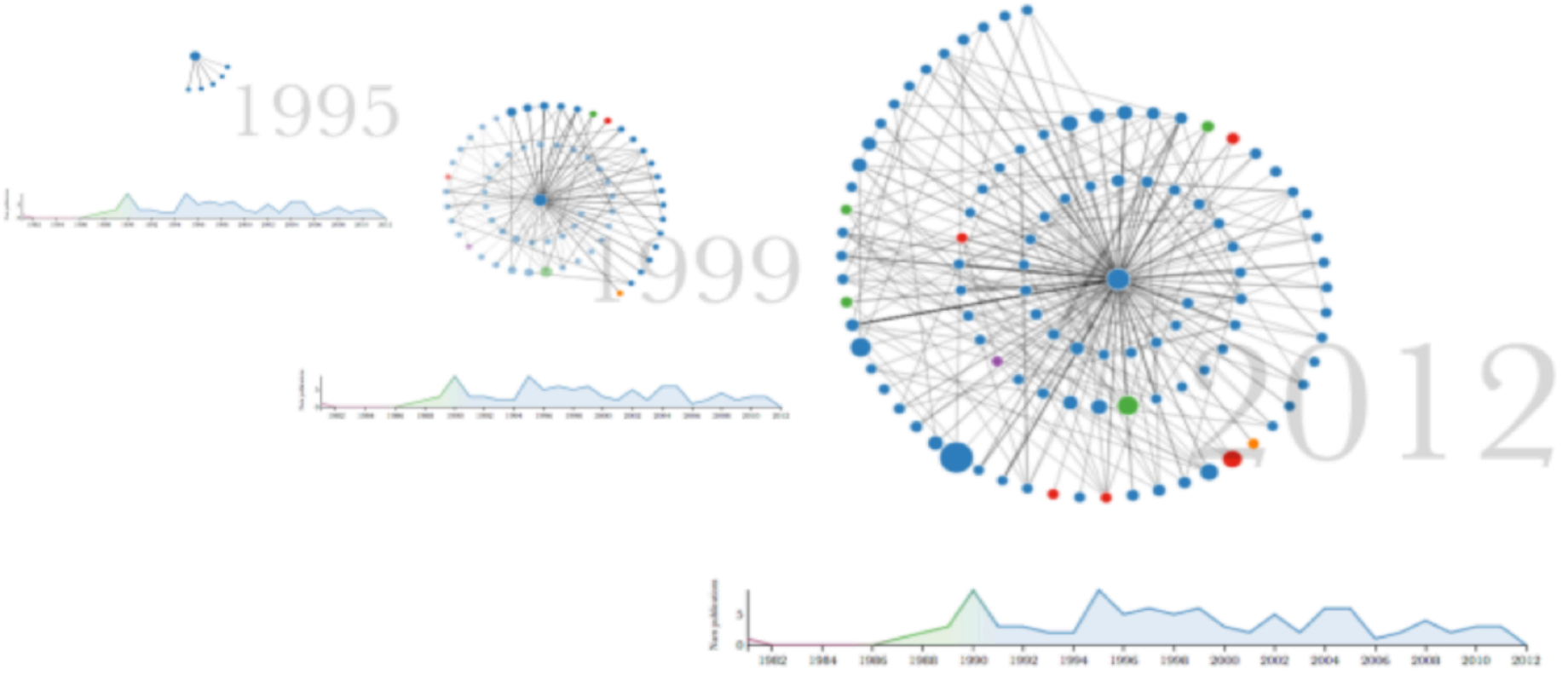


Sum of eigenfactor for this author's publications by year



Explore the data

scholar.eigenfactor.org



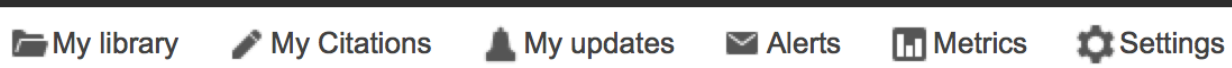
* Please use Chrome web browser for best results

Maladies of Science

Process

Incentives

Infrastructure



- Unknown algorithm
- Unknown corpus
- Non-customizable
- Non-extensible
- No community development

Google scholar Sucks

Articles (include patents) Case law

My updates: recommended based on My Citations Learn more

Faster unfolding of communities: speeding up the Louvain algorithm

Networks of Communities and Communities of Networks in Online Government

See all updates

When you know what you are looking for, Scholar can usually find it. When you don't, Scholar is useless. We need tools for *navigation*.

Antimicrobial Cycling

Scholar

About 39,500 results (**0.05** sec)

Articles

Case law

My library

Any time

Since 2015

Since 2014

Since 2011

Custom range...

Sort by relevance

Sort by date

include patents

include citations

Create alert

Ecological theory suggests that **antimicrobial cycling** will not reduce **antimicrobial** resistance in hospitals

CT Bergstrom, M Lo, M Lipsitch - Proceedings of the …, 2004 - National Acad Sciences

Abstract Hospital-acquired infections caused by antibiotic-resistant bacteria pose a grave and growing threat to public health. **Antimicrobial cycling**, in which two or more antibiotic classes are alternated on a time scale of months to years, seems to be a leading **...**

Cited by 215 Related articles All 17 versions Cite Save

Cycling empirical **antimicrobial** agents to prevent emergence of **antimicrobial**-resistant Gram-negative bacteria among intensive care unit patients

DK Warren, HA Hill, LR Merz, MH Kollef… - Critical care …, 2004 - journals.lww.com

Wolters Kluwer Health may email you for journal alerts and information, but is committed to maintaining your privacy and will not share your personal information without your express consent. For more information, please refer to our Privacy Policy. **...** Skip Navigation Links Home > **...**

Cited by 98 Related articles All 5 versions Cite Save

Antimicrobial Cycling Lessons Learned From the Aminoglycoside Experience

DN Gerding - Infection Control, 2000 - Cambridge Univ Press

Abstract Several discrete strategies have been suggested to prevent or reduce microbial resistance to antimicrobials, including optimal use of the agents (also known as good stewardship); control, removal, or restriction of antimicrobials; use of antimicrobials in **...**

Cited by 71 Related articles All 5 versions Cite Save

Comparison of **antimicrobial cycling** and mixing strategies in two medical intensive care units*

JA Martínez, JM Nicolás, F Marco… - Critical care …, 2006 - journals.lww.com

Wolters Kluwer Health may email you for journal alerts and information, but is committed to maintaining your privacy and will not share your personal information without your express consent. For more information, please refer to our Privacy Policy. **...** Skip Navigation Links Home > February **...**

Cited by 79 Related articles All 6 versions Cite Save

Routine **cycling** of **antimicrobial** agents as an infection-control measure

RA Weinstein, SK Fridkin - Clinical infectious diseases, 2003 - cid.oxfordjournals.org

Abstract **Antimicrobial cycling** is the deliberate, scheduled removal and substitution of specific antimicrobials or classes of antimicrobials within an institutional environment (either

Recommendations

Results for:

Ecological Theory Suggests That Antimicrobial Cycling Will Not Reduce Antimicrobial Resistance In Hospitals - 2003

Expert

The Relationship Between The Volume Of Antimicrobial Consumption In Human Communities And The Frequency Of Re

Evaluating Treatment Protocols To Prevent Antibiotic Resistance - 1996

The Epidemiology Of Antibiotic Resistance In Hospitals: Paradoxes And Prescriptions - 1999

The Transmission Dynamics Of Antibiotic-Resistant Bacteria: The Relationship Between Resistance In Commensal Orga

Persistent Colonization And The Spread Of Antibiotic Resistance In Nosocomial Pathogens: Resistance Is A Regional Pr

Classic

The Crisis In Antibiotic Resistance - 1991

Epidemiology Of Drug Resistance: Implications For A Post-Antimicrobial Era - 1991

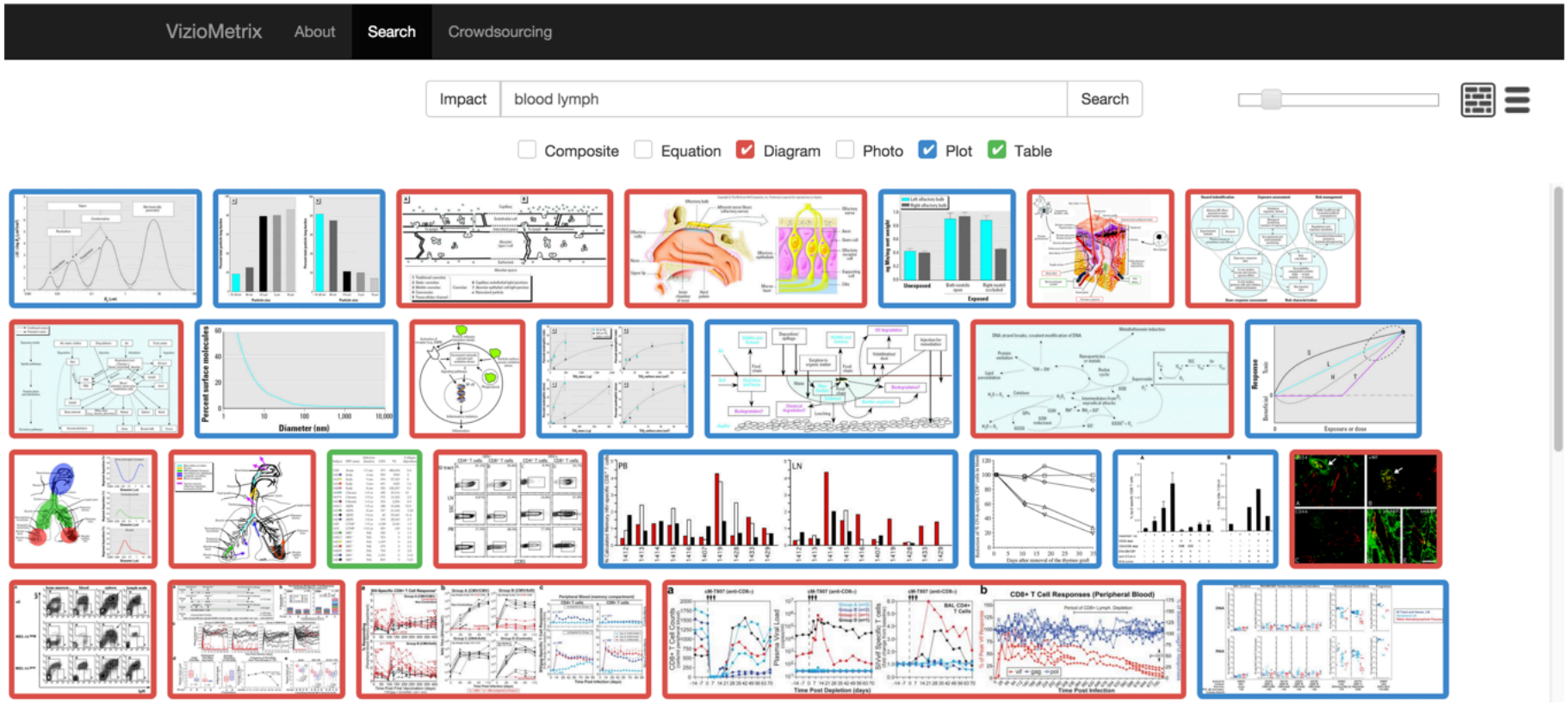
Drug-Resistant Salmonella In The United States: An Epidemiologic Perspective - 1985

The Relationship Between The Volume Of Antimicrobial Consumption In Human Communities And The Frequency Of Re

Evaluating Treatment Protocols To Prevent Antibiotic Resistance - 1996

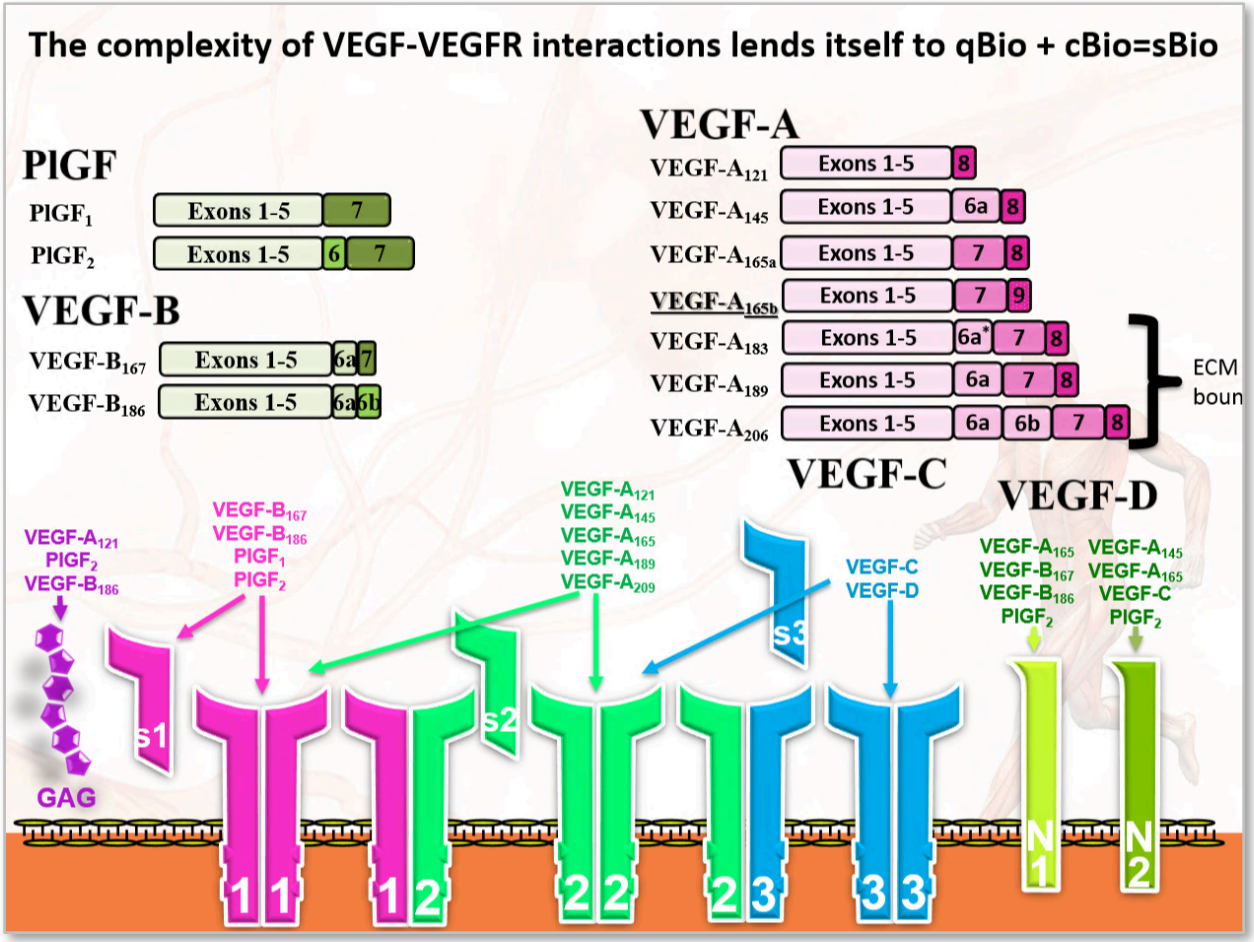
Figure-Centric Search Engine

viziometrics.org

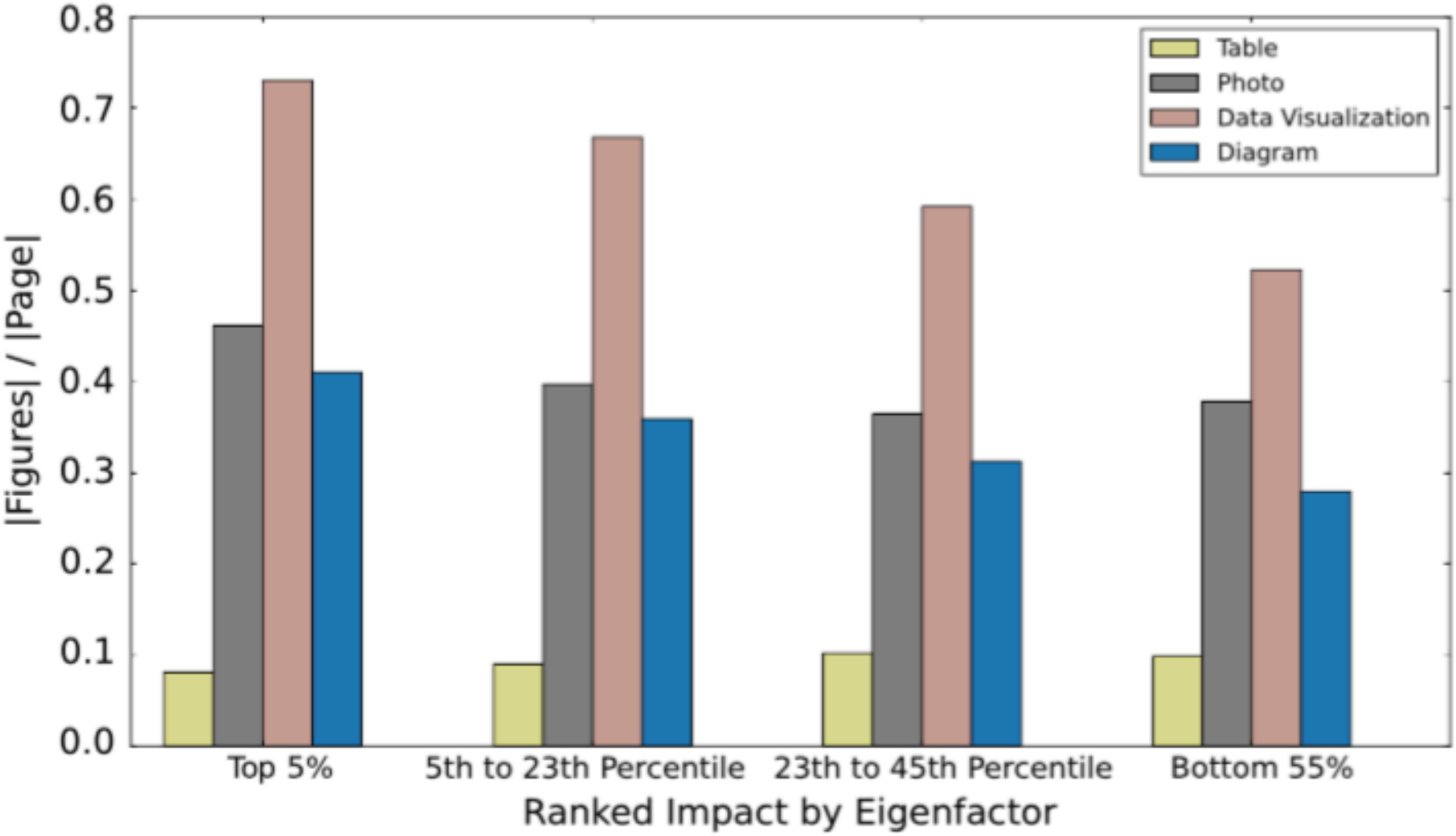


Mining for VEGF concentrations

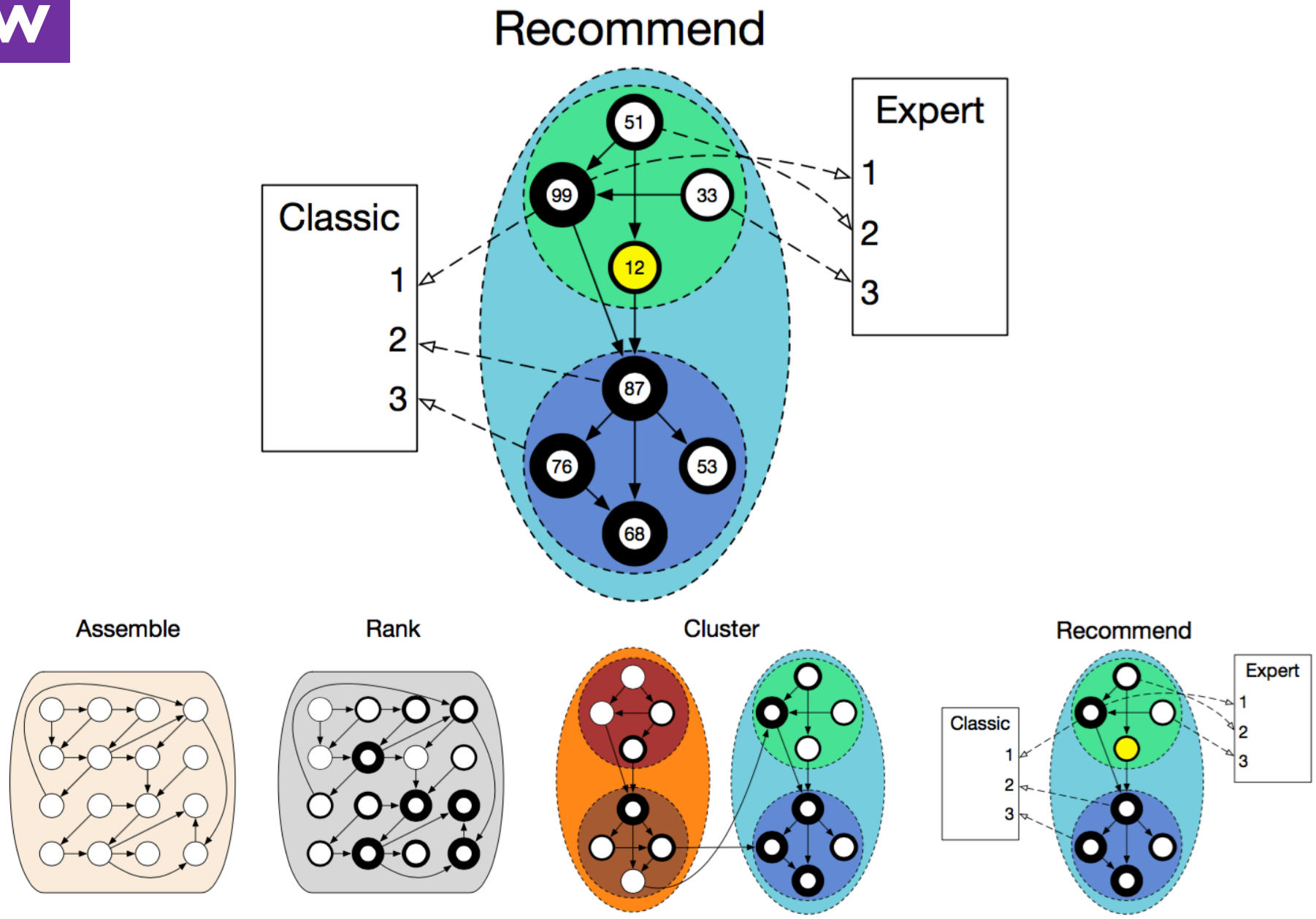
P.I. Imoukhuede
U. of Illinois



Impact versus Figure Density



Lee et al . (2016) Viziometrics: Analyzing Visual Information in the Scientific Literature. https://www.arxiv.org/pdf/1605.04951.pdf



West, Wesley-Smith, Bergstrom (2016) A recommendation system based on hierarchical clustering of an article-level citation network. *IEEE, Transactions on Big Data* (in press)

Babel Home About API

babel.eigenfactor.org

oren etzioni

Ian Wesley-Smith

DBLP JSTOR MAS PLOS PubMed

s

al Methods For Analyzing Speedup Learning Experiments O Etzioni satisfaction programs

-Face And Computer-Mediated Communities Amitai Etzioni, Oren Etzioni 1998 resources sustained

ument Clustering O Zamir document clustering

Communities: Virtual Vs. Real A Etzioni 1996 implications internet

al Methods For Analyzing Speedup Learning Experiments. O Etzioni 1993 scheduling problems

al Methods For Analyzing Speedup Learning Experiments O Etzioni 1993 generating abstractions

Get Related Web Document Clustering: A Feasibility Demonstration O Zamir 1997 document clustering

Get Related Web Document Clustering: A Feasibility Demonstration. O Zamir 1997 browsing large

Get Related Sound And Efficient Closed- World Reasoning O Etzioni proving problem

Get Related Appears In Comm. OfACM O Etzioni scalable comparison-shopping

« Previous 1 2 3 4 5 6 7 8 9 10 Next »

Papers related to Statistical Methods For Analyzing Speedup Learning Experiments O Etzioni satisfaction programs

Get Related Automatically Configuring Constraint Satisfaction Programs: A Case Study S Minton 1995 satisfaction programs

Get Related Abstraction Via Approximate Symmetry T Ellman 1992 satisfaction programs

Get Related Integrating Heuristics For Constraint Satisfaction Problems: A Case Study S Minton 1992 satisfaction programs

Get Related An Analytic Learning System For Specializing Heuristics S Minton 1992 satisfaction programs

Get Related Automated Synthesis Of Constrained Generators W Braudaway 1988 satisfaction programs

Maladies of Science

Process

Incentives

Infrastructure

Predatory Publishers

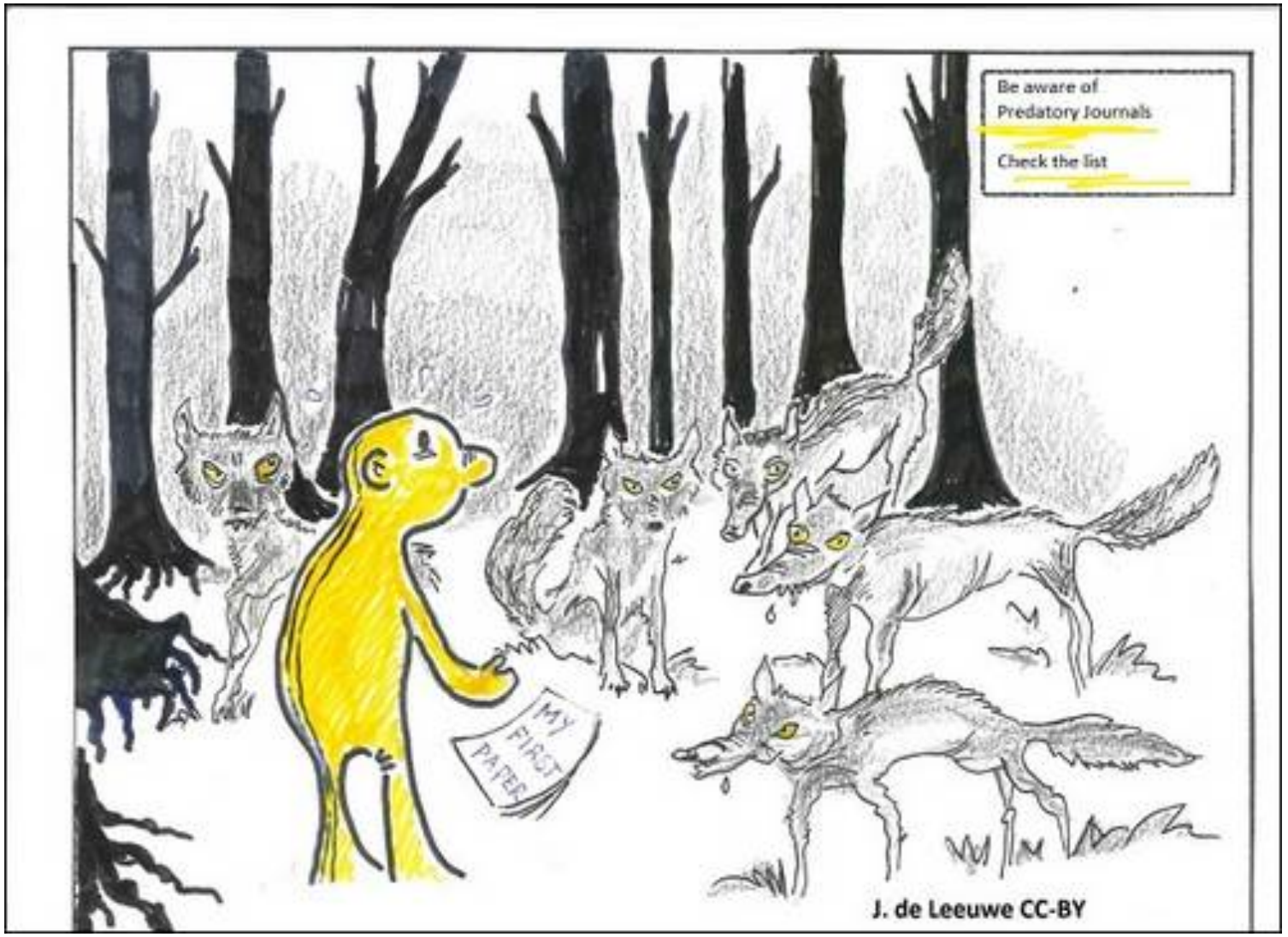
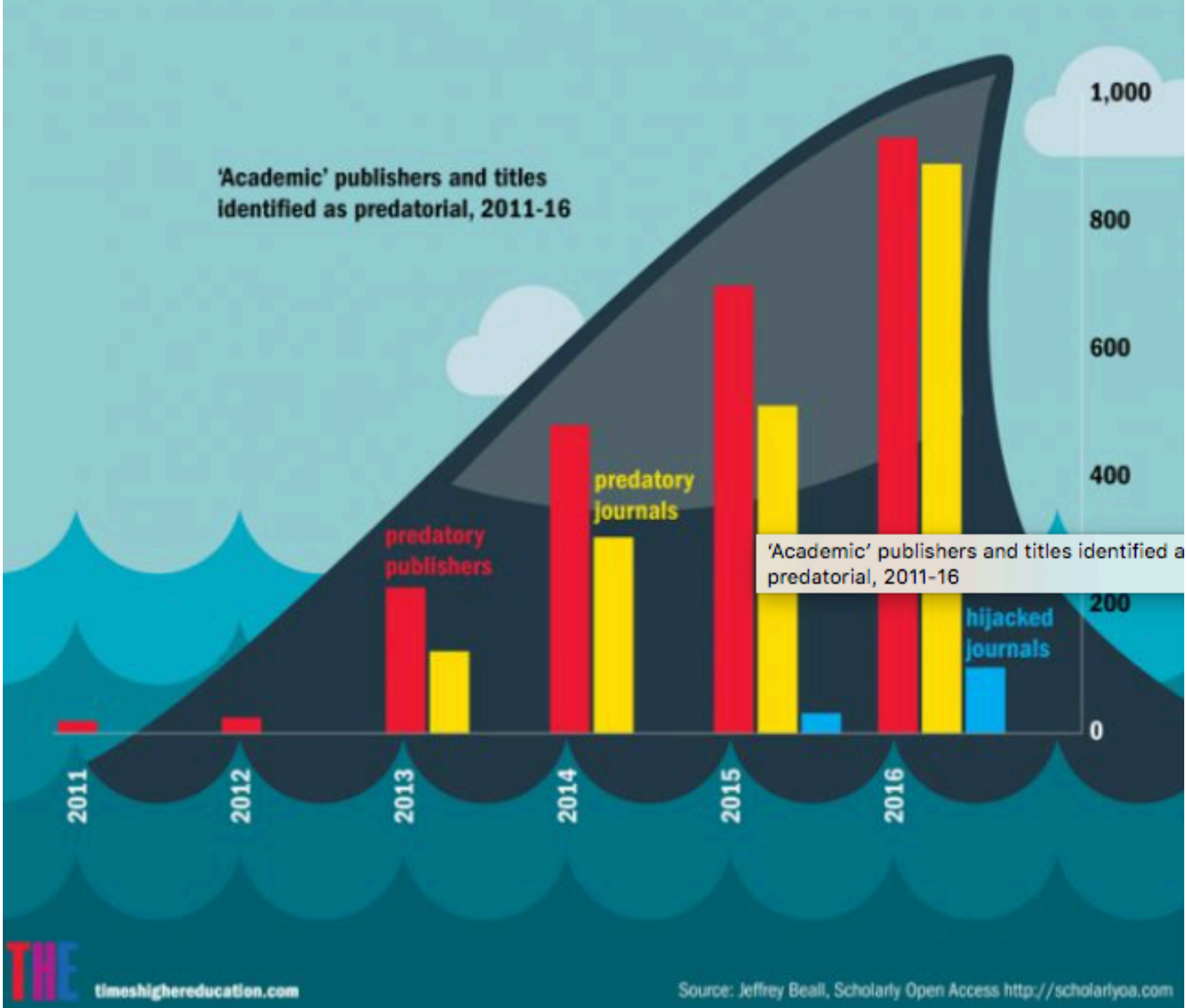
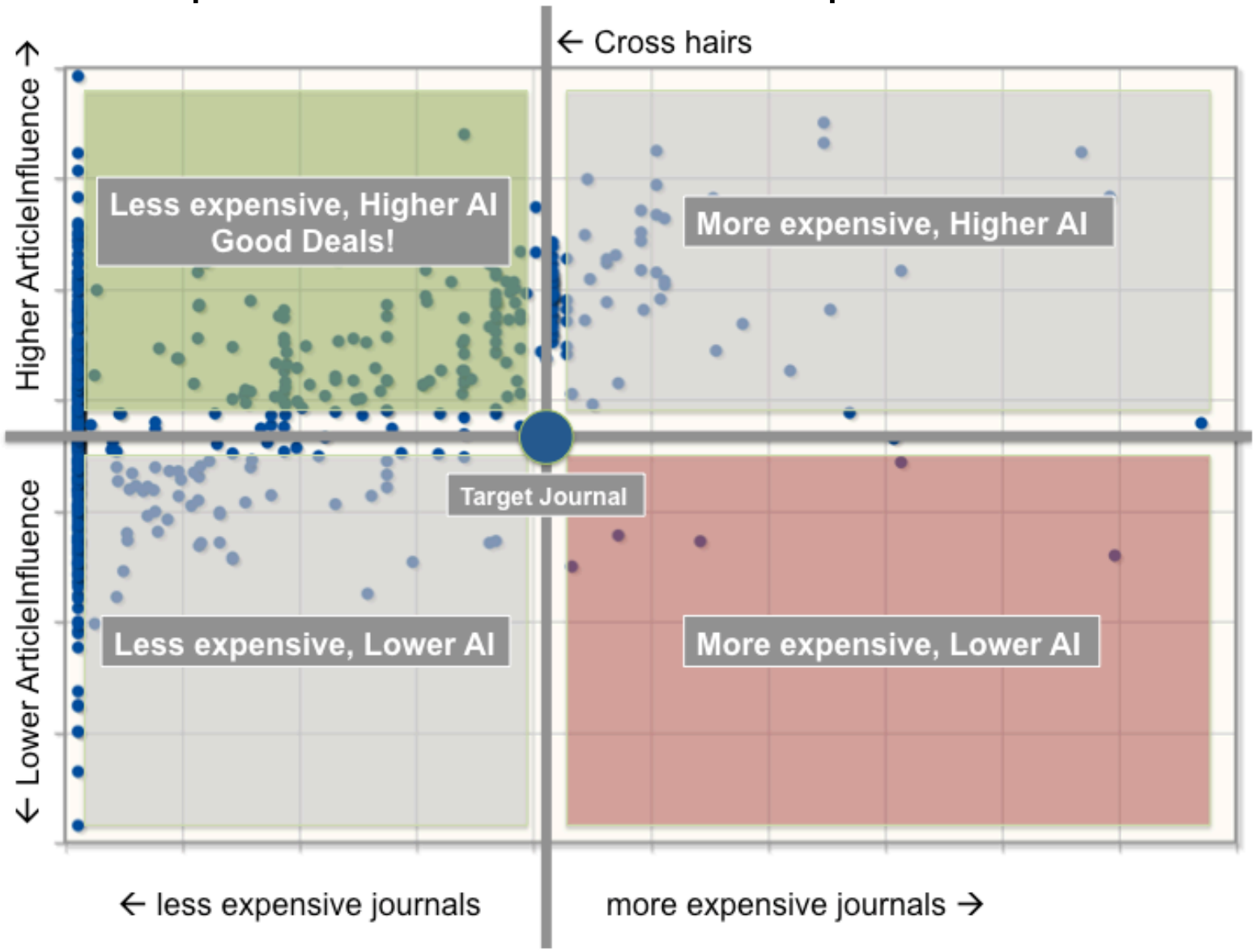


Image Source: http://www.library.tudelft.nl/en/support/publishing-advice/open-access/


'Academic' publishers and titles identified as predatorial, 2011-16
predatory publishers
predatory journals
hijacked journals
1,000
800
600
400
200
0
2011
2012
2013
2014
2015
2016
'Academic' publishers and titles identified a
predatorial, 2011-16
THE
timeshighereducation.com
Source: Jeffrey Beall, Scholarly Open Access http://scholarlyoa.com

Open Market of Open Access



http://www.eigenfactor.org/projects/openAccess/

Maladies of Science

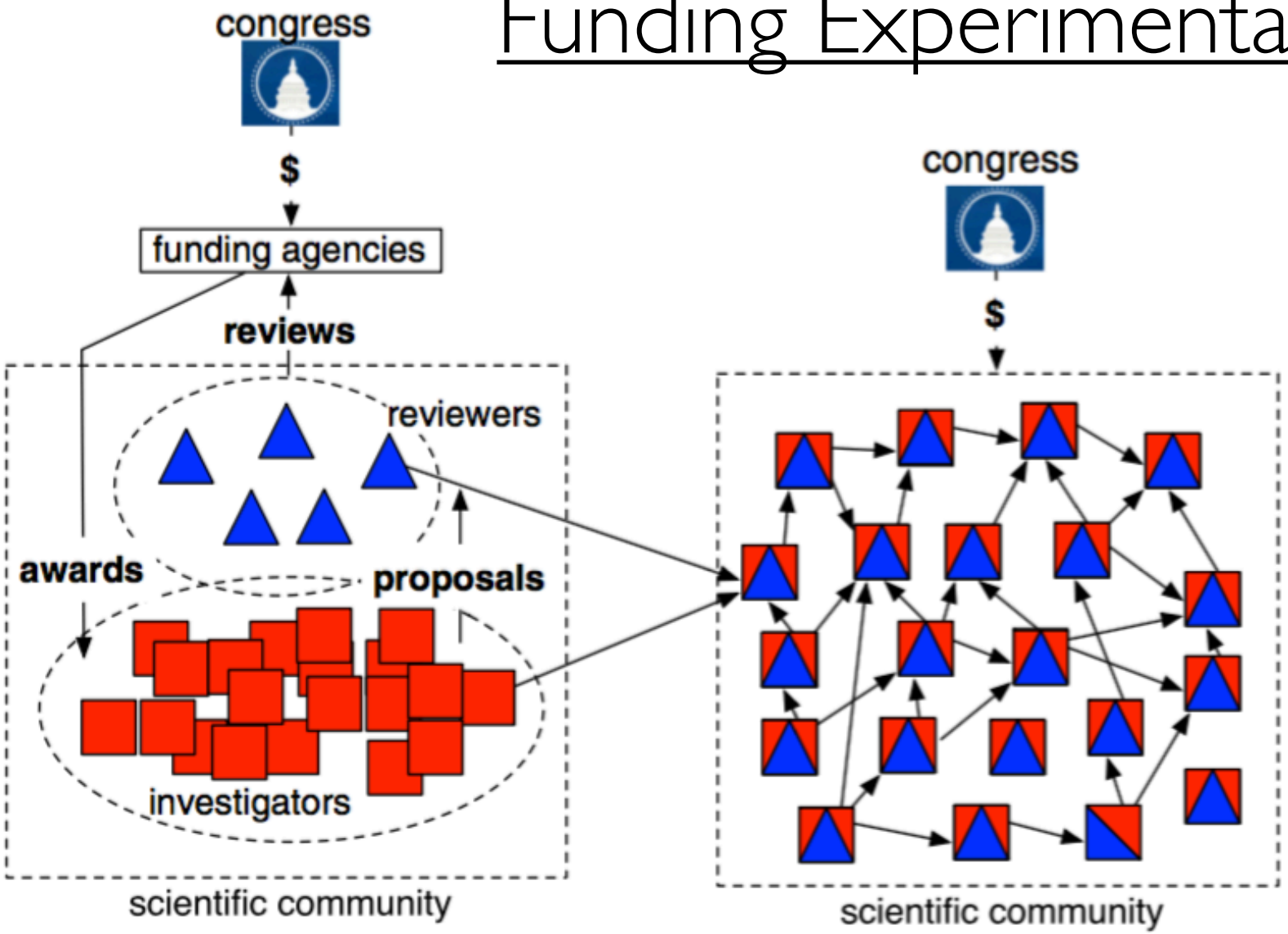
Process

Incentives

Infrastructure

Funding

Funding Experimentation



Bollen et al (2016) Scientometrics

JW

What is my impact on science?

$7,933,670,366

JW

22,756 awards

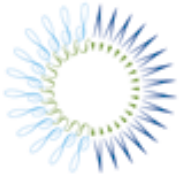
17,849 researchers

344,917 papers

8,174,533 citations

23. 7 citations/paper

2006 - 2015



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

~ 37 citations/paper

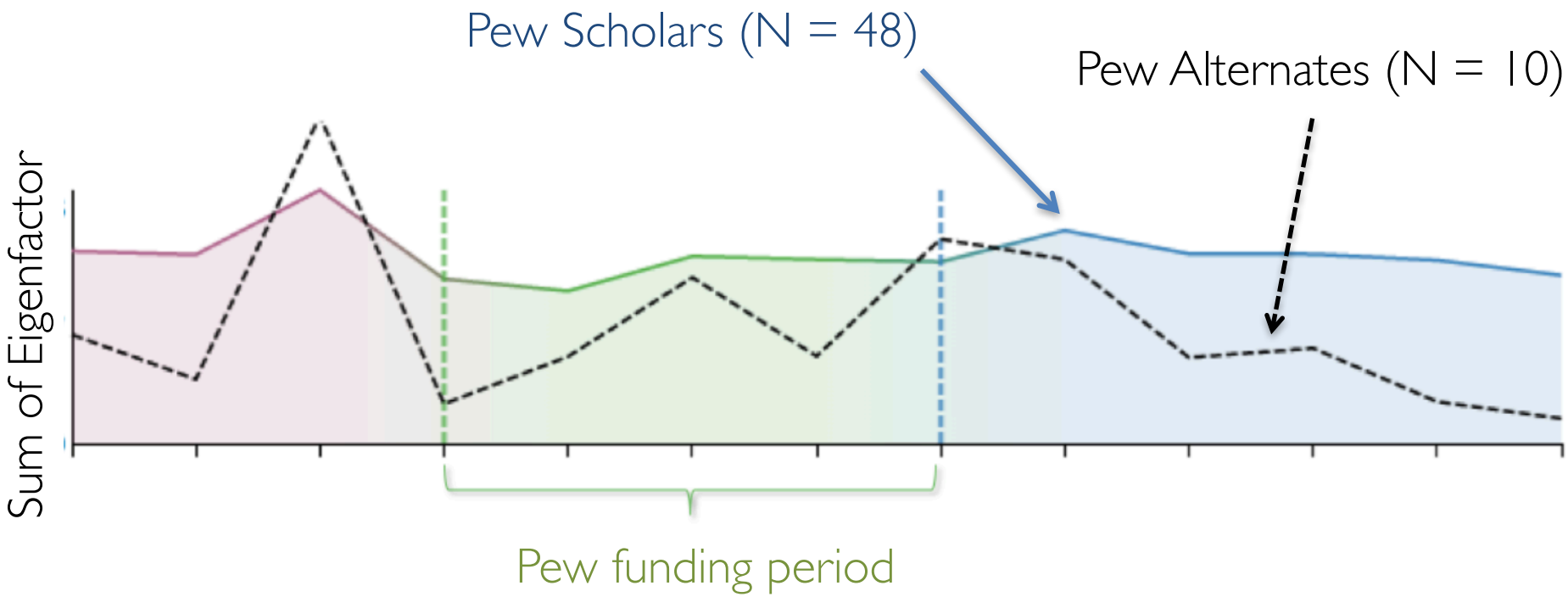
median citations = 11

Science

~ 5 citations/paper

median citations = 0

Comparing Alternates



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

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

The map equation

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

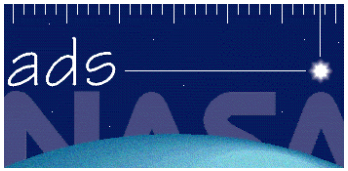
Rosvall and Bergstrom (2008) *PNAS*

Data

Compressing ⟷ Finding patterns

Minimum description length (MDL) statistics.

The Scholarly Graph

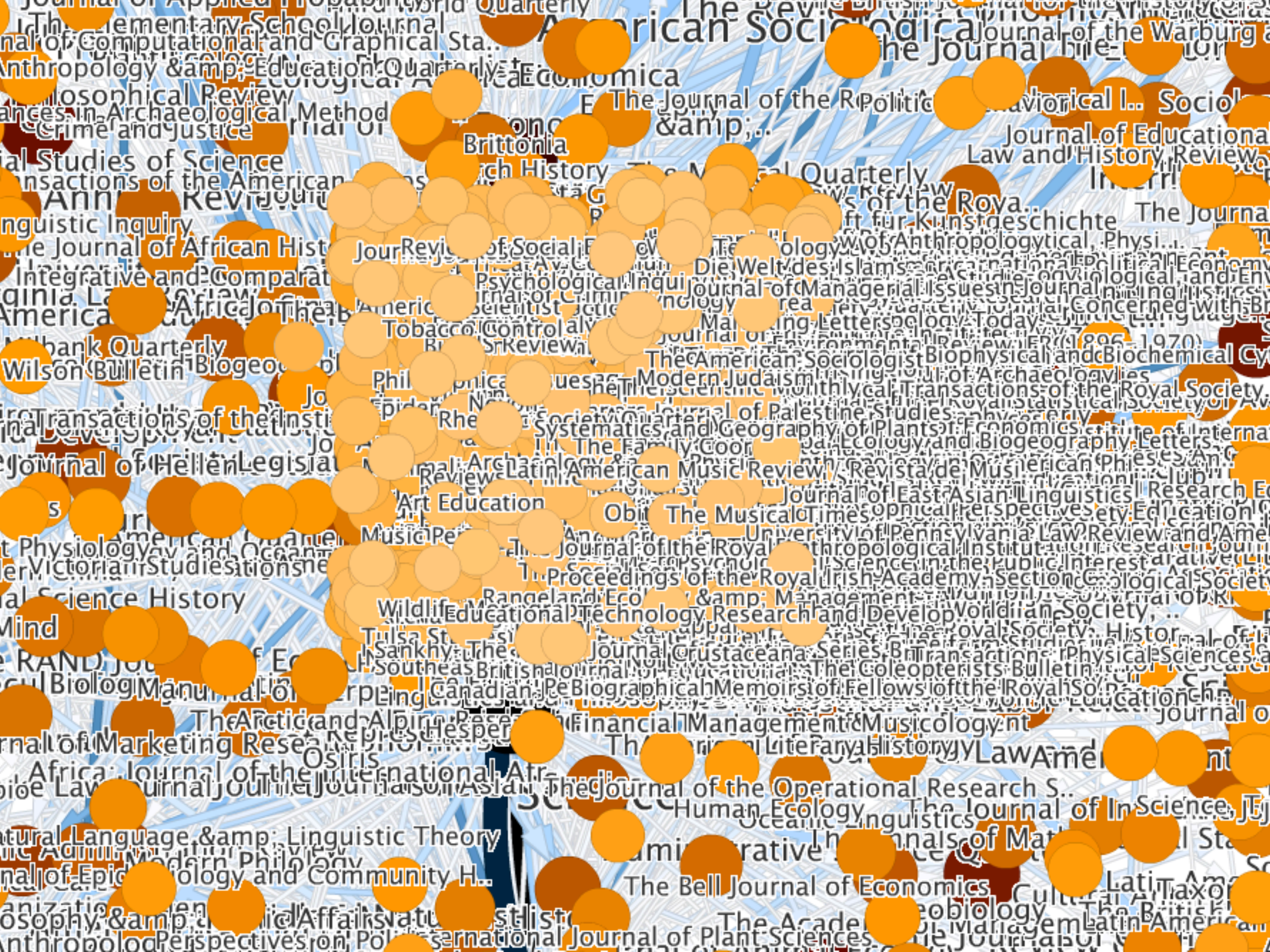


The Scholarly Graph

Tens of millions articles, patents, books

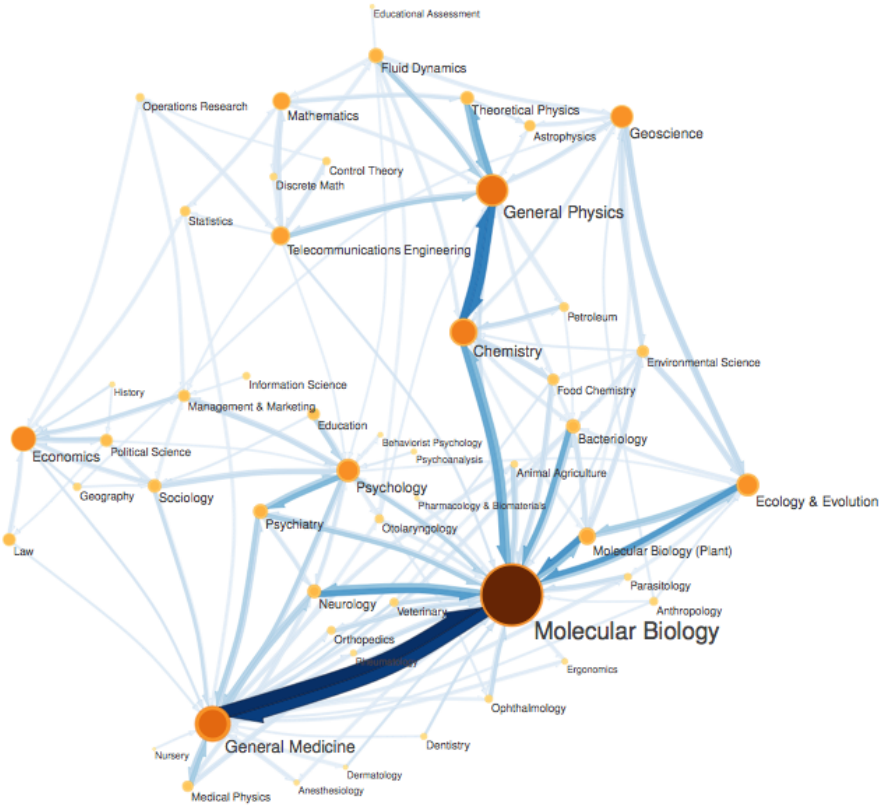
Billions of citation links

Years: 1600 - 2016

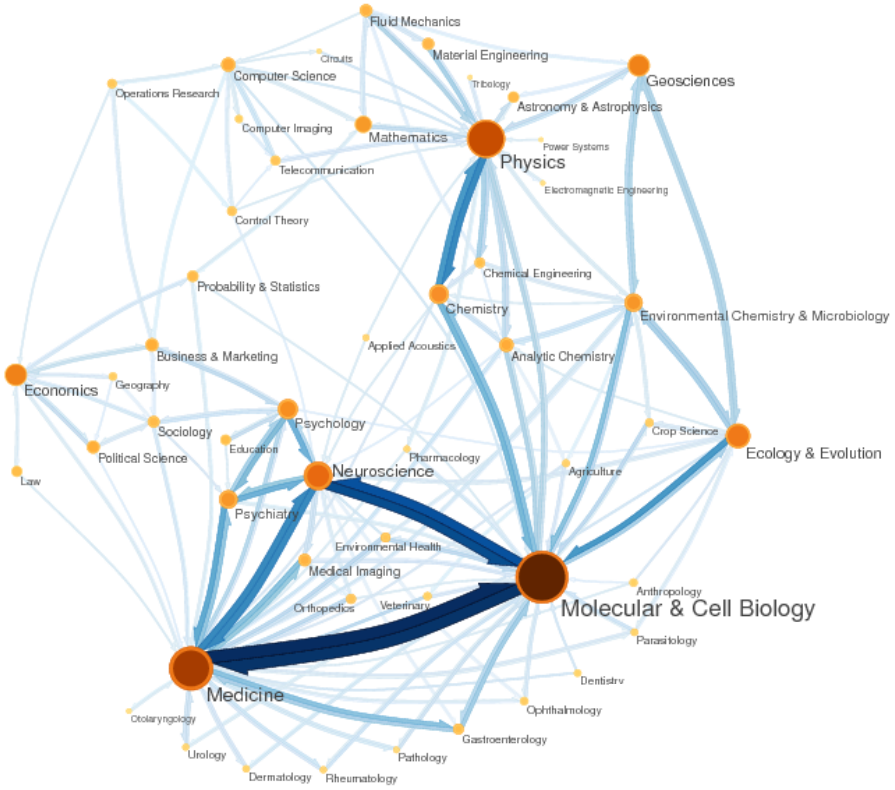


American Sociological
The Elementary School Journal
Journal of the Warburg
Anthropology & Education Quarterly
Economica
Philosophical Review
Crime and Justice
Brittonia
Law and History Review
Linguistic Inquiry
Wilson Bulletin
Tobacco Control
Psychological Inquiry
Die Welt des Islams
Journal of Managerial Issues
The American Sociologist
Modern Judaism
Journal of Palestine Studies
Systematics and Geography of Plants
Latin American Music Review
Journal of East Asian Linguistics
Art Education
The Musical Times
University of Pennsylvania Law Review
Victorian Studies
Proceedings of the Royal Irish Academy
Educational Technology Research and Develop
Mind
The Coleopterists Bulletin
Biographical Memoirs of Fellows of the Royal So..
Financial Management
Osiris
The Journal of the Operational Research S..
Human Ecology
Natural Language & Linguistic Theory
Modern Philology
The Bell Journal of Economics
International Journal of Plant Sciences

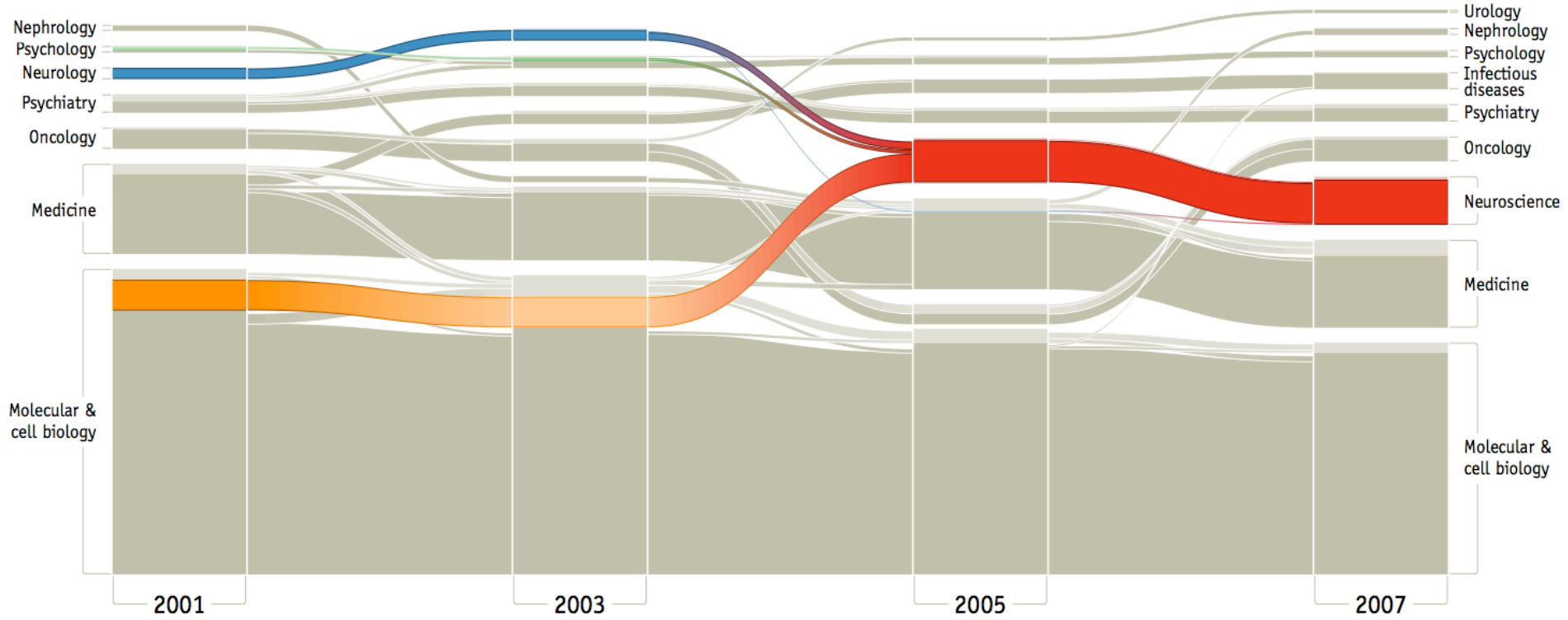
1995


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
Psychology
Animal Agriculture
Geography
Sociology
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

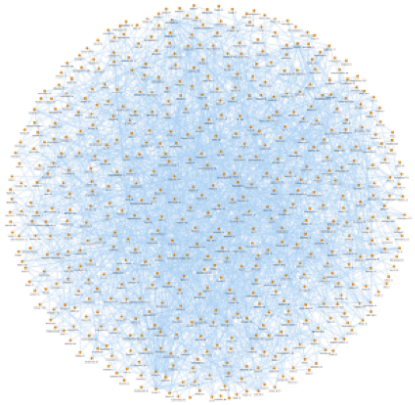

Fluid Mechanics
Material Engineering
Circuits
Computer Science
Geosciences
Tribology
Operations Research
Astronomy & Astrophysics
Computer Imaging
Mathematics
Power Systems
Physics
Telecommunication
Electromagnetic Engineering
Control Theory
Chemical Engineering
Probability & Statistics
Chemistry
Environmental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

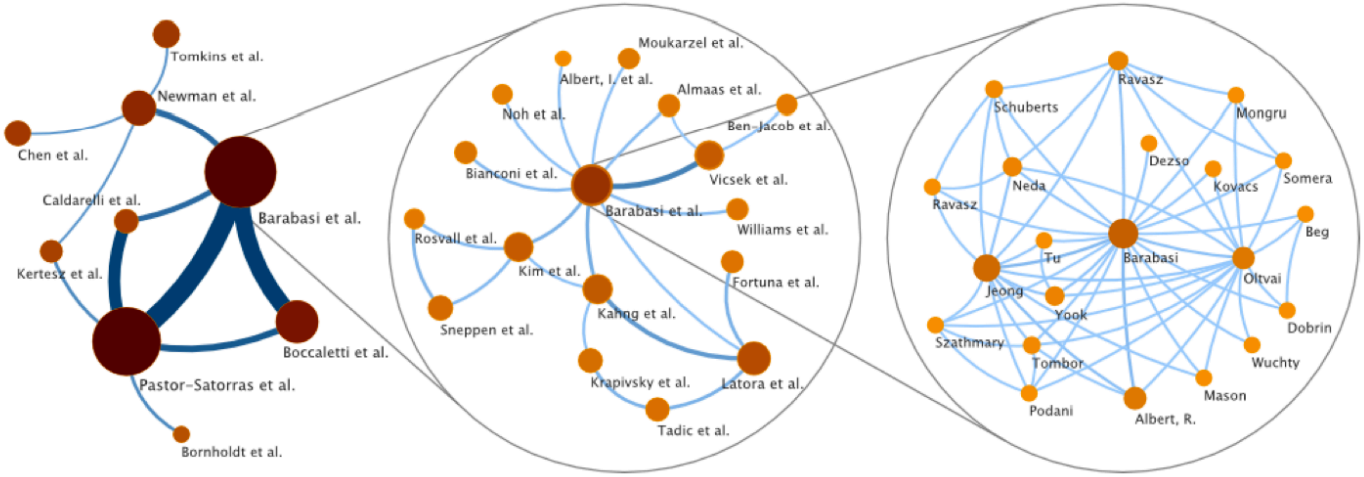
The Emergence of Neuroscience



Rosvall & Bergstrom (2010) *PLoS One*

Zoomability





Current data technologies give you one level.

I want to give you the ability to zoom.

Maladies of Science

Process

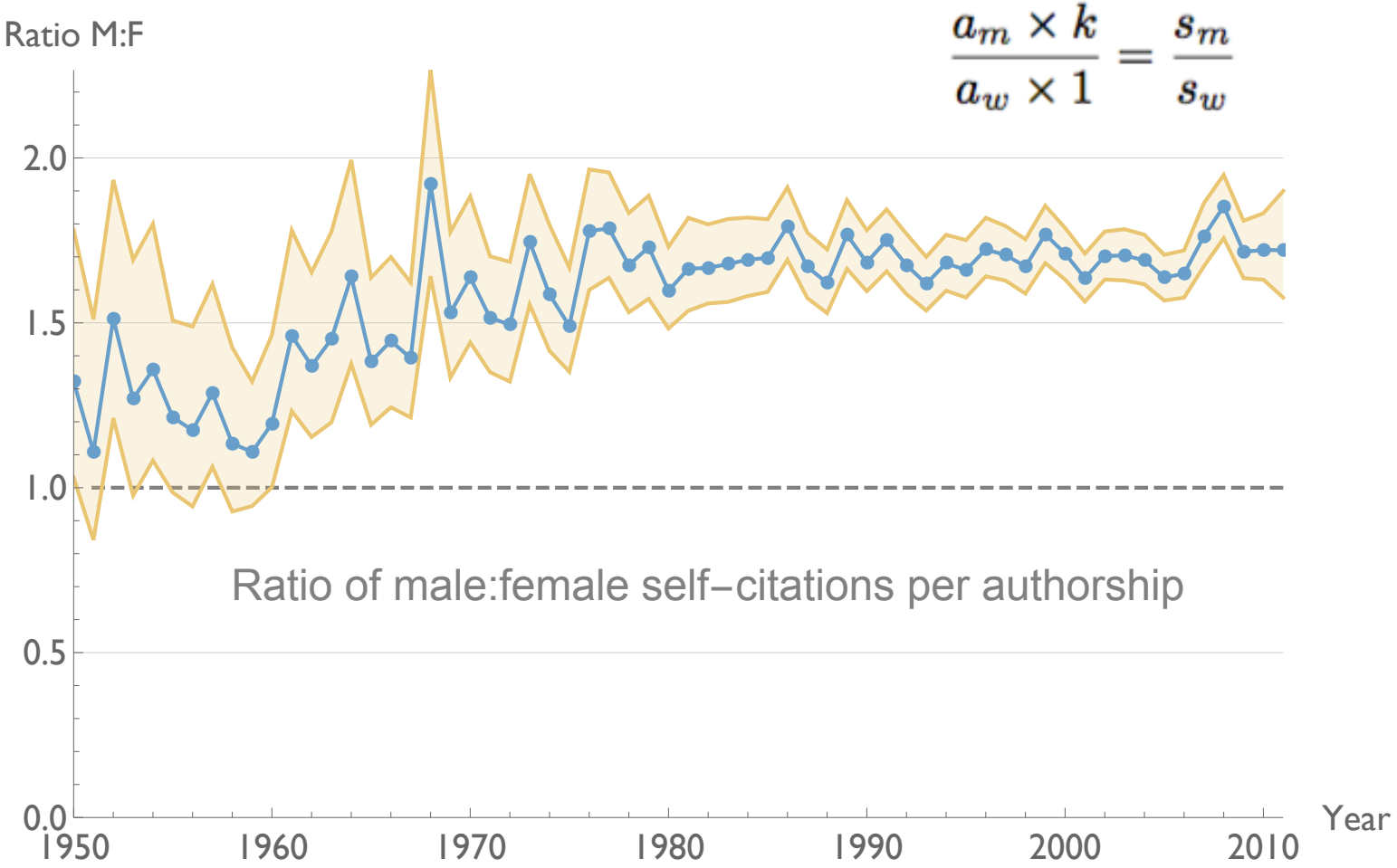
Incentives

Infrastructure

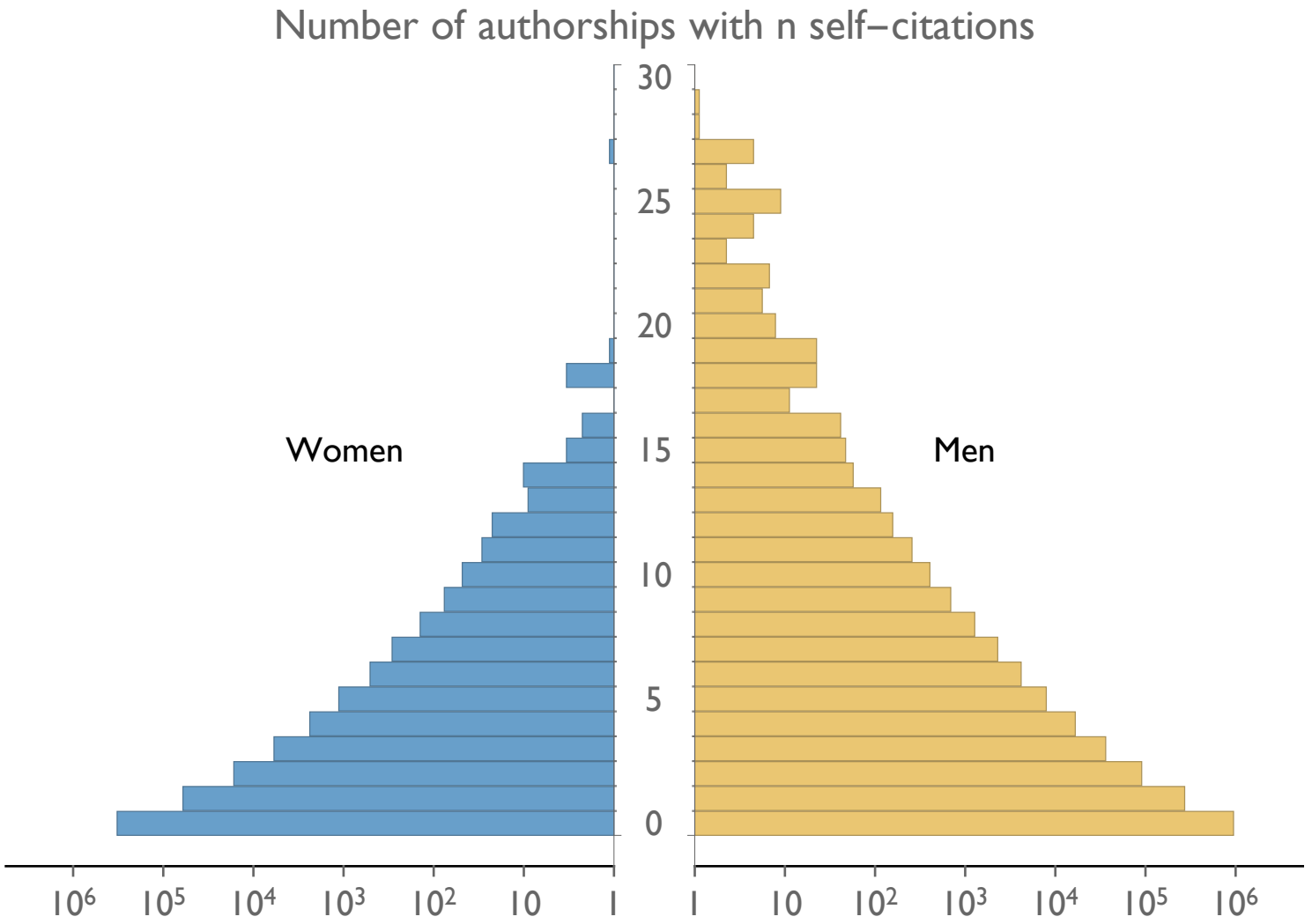
Funding

Bias

Self-citation over time



Excessive Self Citation



Summarizing the *Maladies* of Science

Process

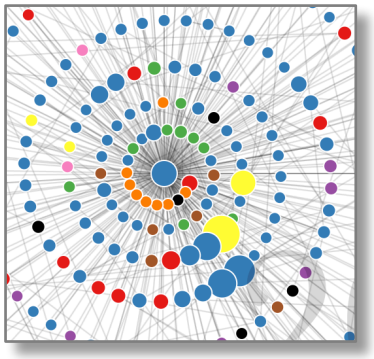
Incentives

Infrastructure

Funding

Bias

Science of Science



Acknowledgements

Jason Portenoy, Information School, University of Washington

Bill Howe, Information School, University of Washington

Poshen Lee, Electrical Engineering, University of Washington

Martin Rosvall, Department of Physics, Umea University

Carl Bergstrom, Department of Biology, University of Washington

Sloan Foundation, JSTOR, Metaknowledge Network, NSF