Measuring & Mapping Interdisciplinarity of a Research Program

Alan Porter
Search Technology, Inc.
&
Georgia Tech

alan.porter@isye.gatech.edu
404-384-6295

Agenda

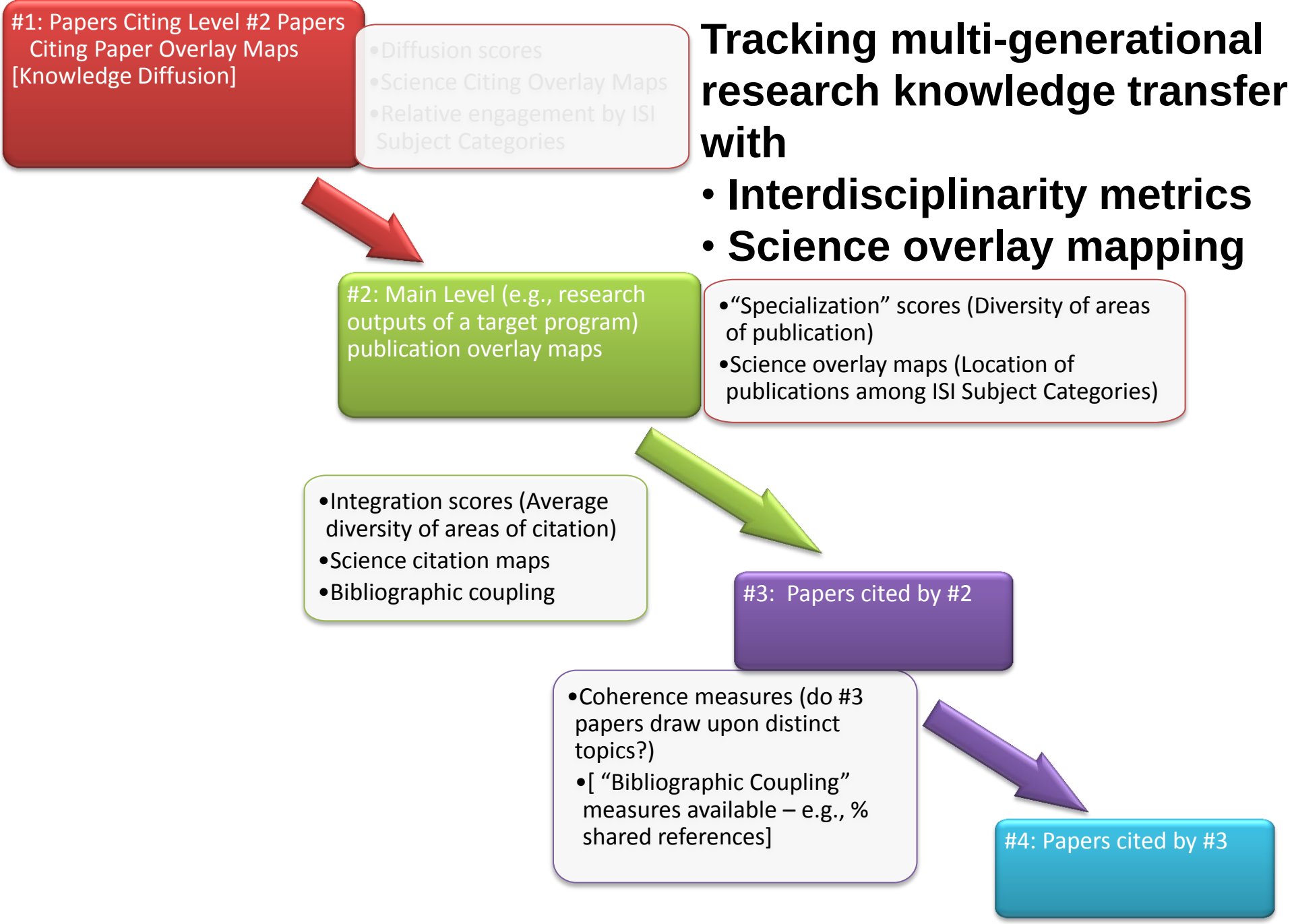
1. Measuring Interdisciplinarity: Integration, Specialization & Diffusion Scores

2. Visualization:
   - Science overlay maps (locating research activity)
   - Research networking maps

** Illustrations from NSF research assessments

Bases

1. Using multiple data resources for research assessment
   - Publications – via Web of Science
   - Proposal references – (using Web of Science)
   - Citations – via Web of Science
2. Data cleaning and analyses
   - Using VantagePoint software
3. Visualization
   - Using VantagePoint together with Aduna, Pajek, Excel, Gephi, etc.


#1: Papers Citing Level #2 Papers Citing Paper Overlay Maps [Knowledge Diffusion]
•Diffusion scores
•Science Citing Overlay Maps
•Relative engagement by ISI Subject Categories
Tracking multi-generational research knowledge transfer with
• Interdisciplinarity metrics
• Science overlay mapping
#2: Main Level (e.g., research outputs of a target program) publication overlay maps
•"Specialization" scores (Diversity of areas of publication)
•Science overlay maps (Location of publications among ISI Subject Categories)
•Integration scores (Average diversity of areas of citation)
•Science citation maps
•Bibliographic coupling
#3: Papers cited by #2
•Coherence measures (do #3 papers draw upon distinct topics?)
•["Bibliographic Coupling" measures available – e.g., % shared references]
#4: Papers cited by #3

Interdisciplinary Research Metrics

- National Academies Keck Futures Initiative (15-year program) to boost interdisciplinary research in the US
- Measure interdisciplinarity for program evaluation
- For a body of research
  - Extract papers' cited references
  - Associate cited journals to Web of Science (WOS) Subject Categories (SCs)
  - Matrix of SC by SC interrelationships
  - For given paper set, calculate
    - "Integration" – breadth of SCs drawn upon
    - "Specialization" – concentration of publication activity
    - "Diffusion" – diversity of SCs citing the research

Heuristics of diversity

(Stirling, 1998; 2007)
(Rafols and Meyer, 2009)

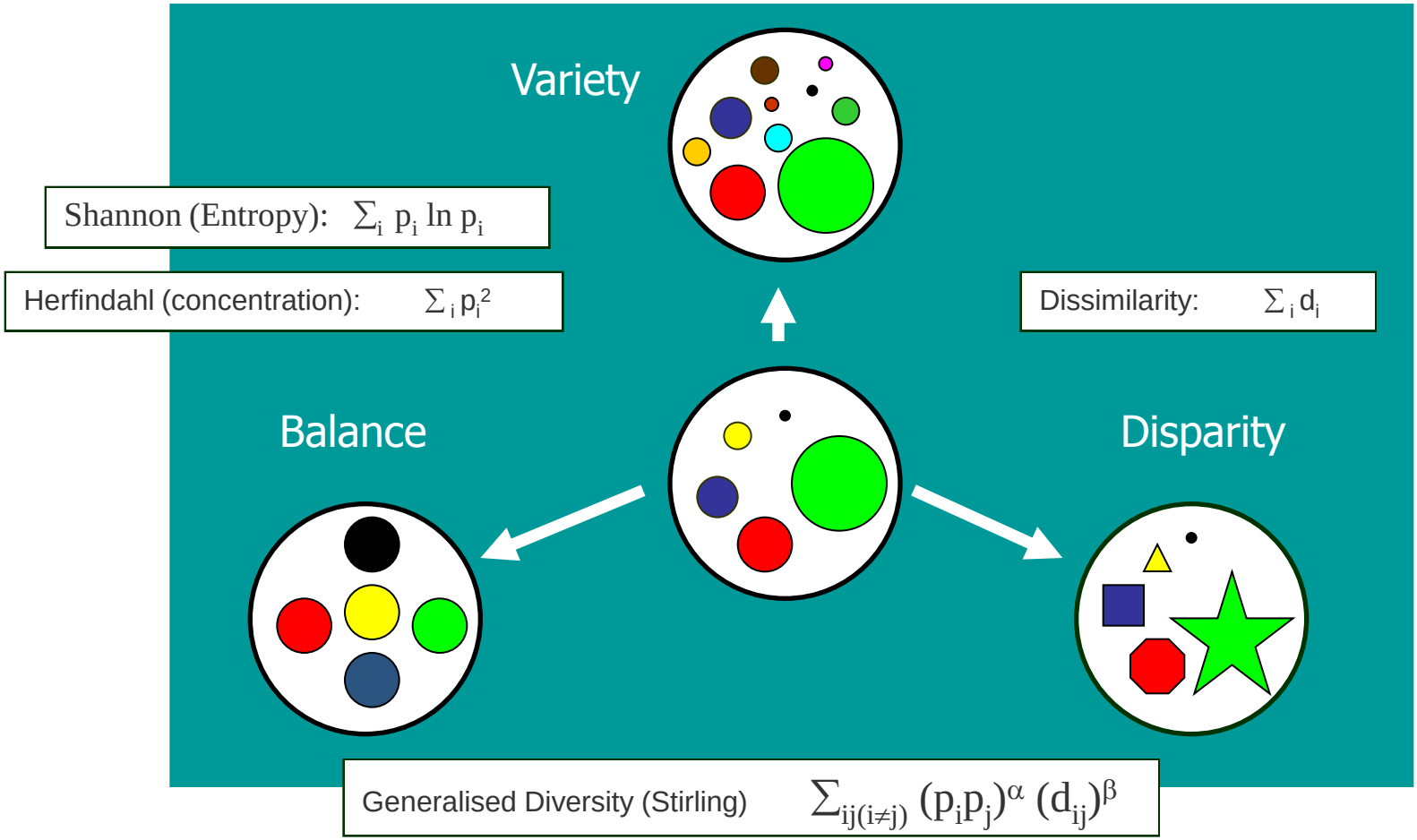
Diversity:

'attribute of a system whose elements may be apportioned into categories'

Characteristics:

Variety: Number of distinctive categories
Balance: Evenness of the distribution
Disparity: Degree to which the categories are different.

[** Shannon & Herfindahl do not include Disparity]



Web of Science ("WOS")

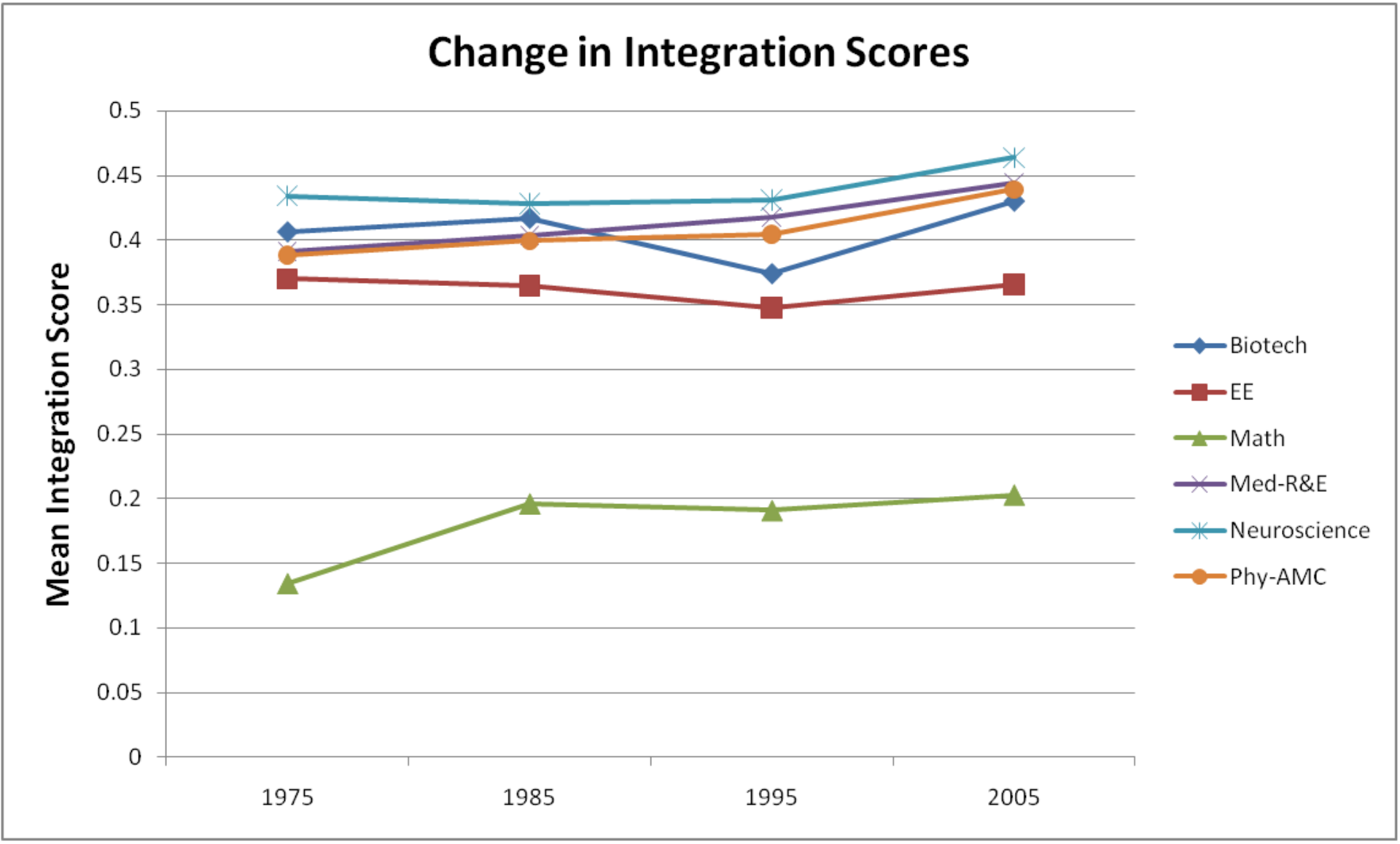
- Indexes publications from ~12,000 leading journals
- Recently >1.5 million papers per year
- Includes several databases
  - **Science Citation Index Expanded (SCI)**
  - **Social Sciences Citation Index (SSCI)**
  - Arts & Humanities Citation Index (A&HCI)
  - Conference Proceedings
- Provides field-structured abstract records
  - Classify journals into **Subject Categories ("SCs") –** presently, 224 for SCI + SSCI
  - Provide Cited References for each paper – we apply thesauri to associate to **Cited SCs**
  - Separately search for Citing records for each paper to discern **Citing SCs**

Sample WOS Abstract Record (excerpted)

AU Oliver-Hoyo, M
Gerber, RW
TI From the research bench to the teaching laboratory: Gold nanoparticle layering
SO JOURNAL OF CHEMICAL EDUCATION
DT Article
C1 N Carolina State Univ, Dept Chem, Raleigh, NC 27695 USA.
AB …
CR BENTLEY AK, 2005, J CHEM EDUC, V82, P765
BOLSTAD DB, 2002, J CHEM EDUC, V79, P1101
HALE PS, 2005, J CHEM EDUC, V82, P775, …
NR 16
TC 1
PY 2007
VL 84
IS 7
BP 1174
EP 1176
SC Chemistry, Multidisciplinary; Education, Scientific Disciplines

Use thesauri to associate "J Chem Educ" with its SCs

Benchmarking Integration Scores



Porter and Rafols (2009)

NSF Research Assessments

- RCN (Research Coordination Networks) Program
  - Can we see researcher network enrichment, Before to After?
- HSD (Human & Social Dynamics) and CMG (Collaborations in Math & Geosciences) Programs
  - How interdisciplinary (compared to ~similar projects)?
- REESE (Research & Evaluation on Education in Science & Engineering) Program
  - How is Cognitive Science engaging with STEM education, over time?

Research Design: Non-equivalent Control Group

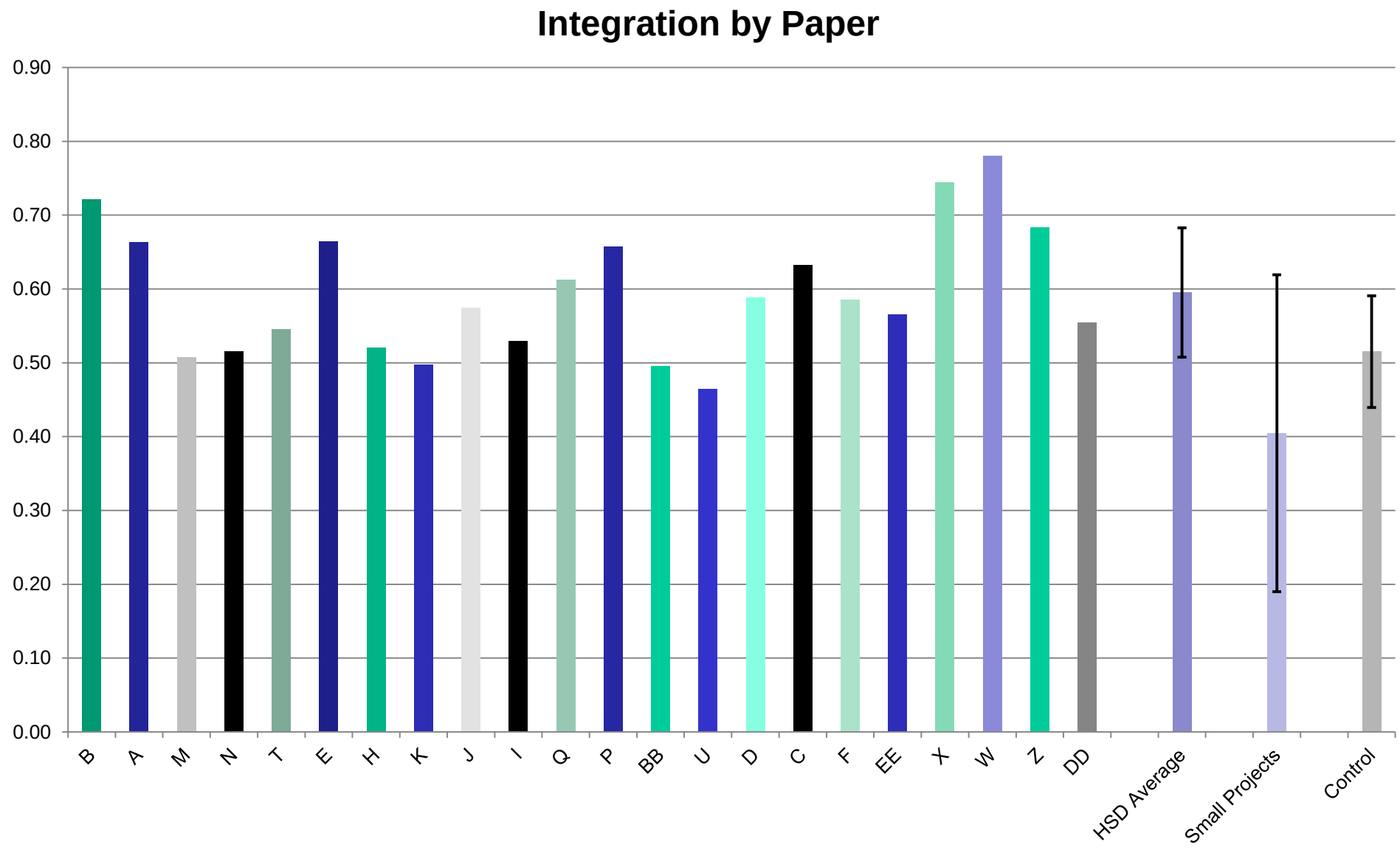
	RCN	Comparison Group
Before (1999-2001)	$O_{1\text{-RCN}}$	$O_{1\text{-Comp}}$
[Treatment]	X	
After (2006-2008)	$O_{2\text{-RCN}}$	$O_{2\text{-Comp}}$

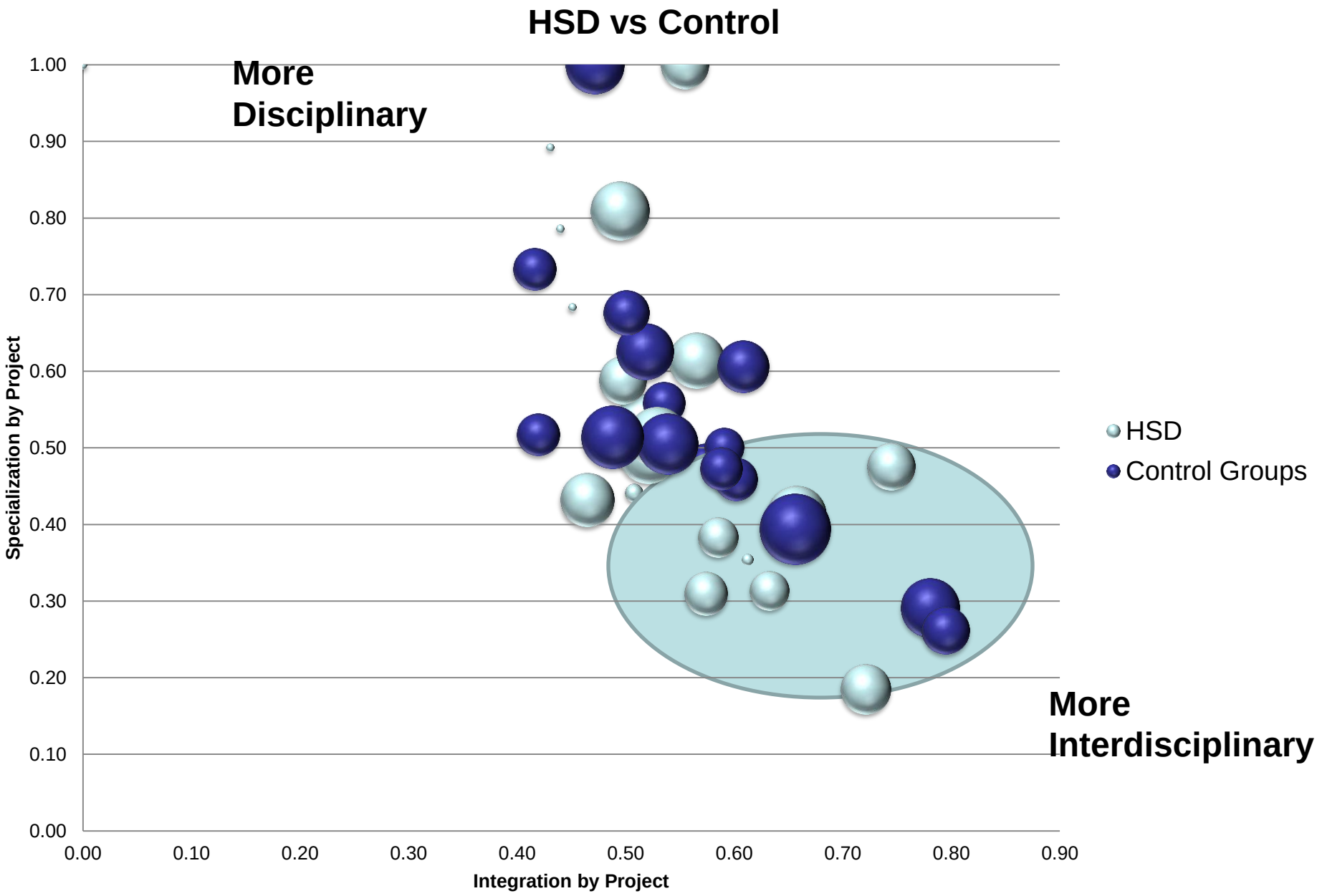
[We also did various analyses focused on research outputs **deriving explicitly from RCN support**.]

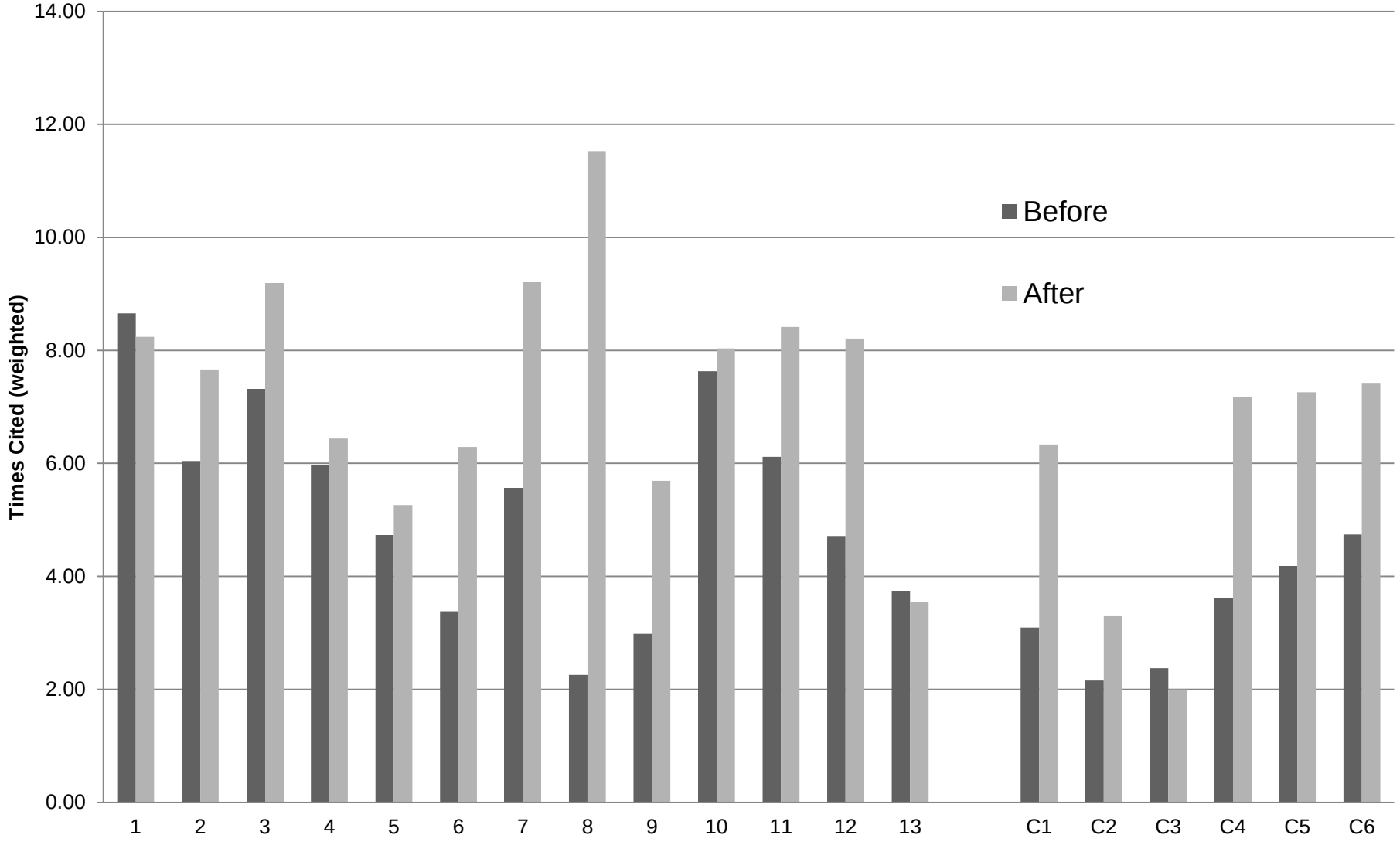
HSD-derived Publication Characteristics

		Project		
	Overall	Project B	Project A	Project H
Authors/Paper	2.79	2.42	3.00	2.27
Author Affiliation	2.26	1.92	2.69	2.09
Cited Reference Count	42.44	38.63	31.00	54.73
Number of Countries	1.44	1.25	2.19	1.55
Integration by Article	0.58	0.72	0.66	0.52
Journal Impact Factor	3.89	2.36	3.14	4.24
Times Cited	7.48	6.46	4.31	3.27

HSD Integration Scores




HSD vs Control
More Disciplinary
More Interdisciplinary
HSD
Control Groups
Specialization by Project
Integration by Project
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90



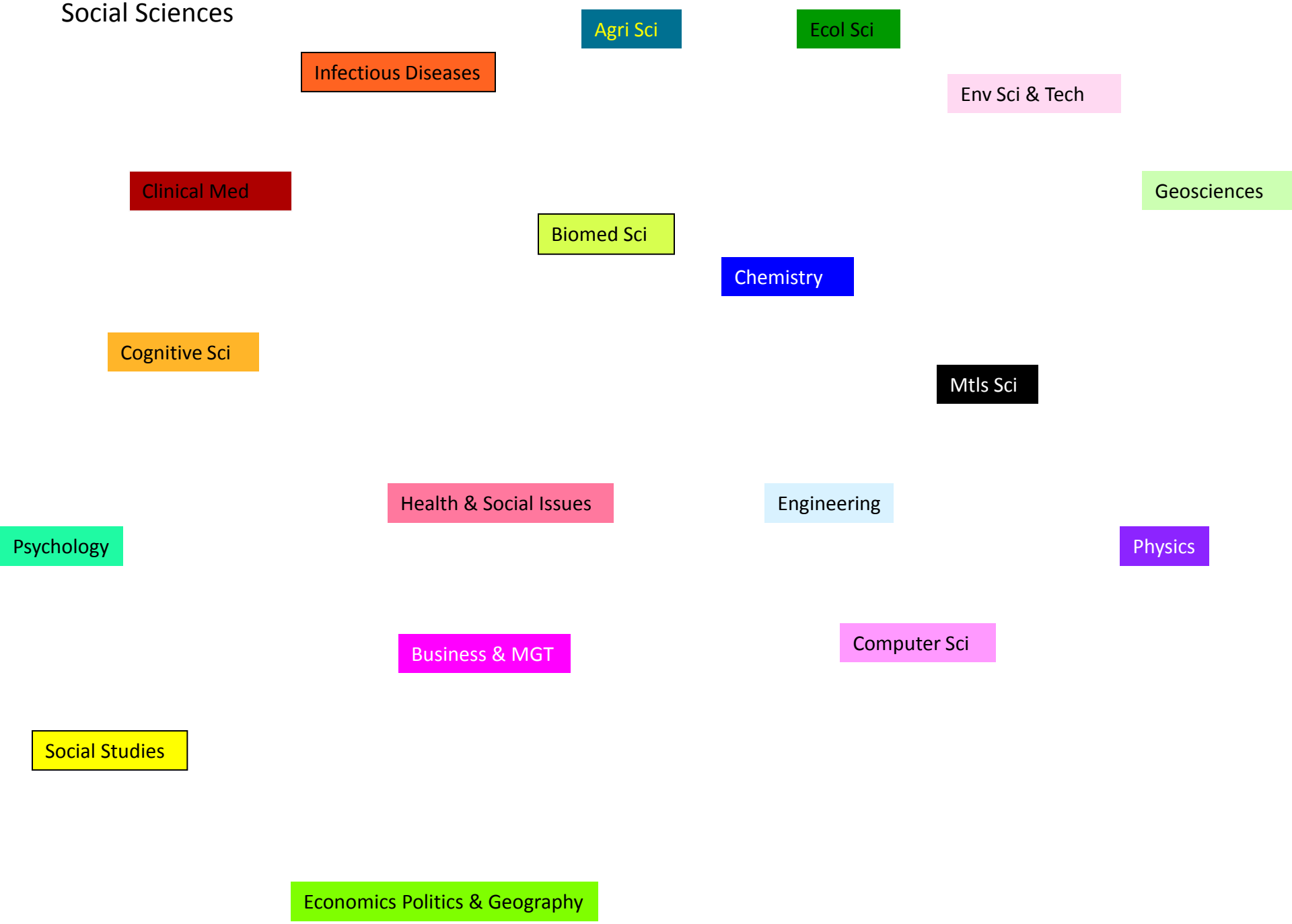
Comparable Increase in Times Cited (weighted):
RCN: paired t = 3.14, P = 0.009;
Control: paired t = 3.56, P = 0.016

Dual, Complementary Mapping

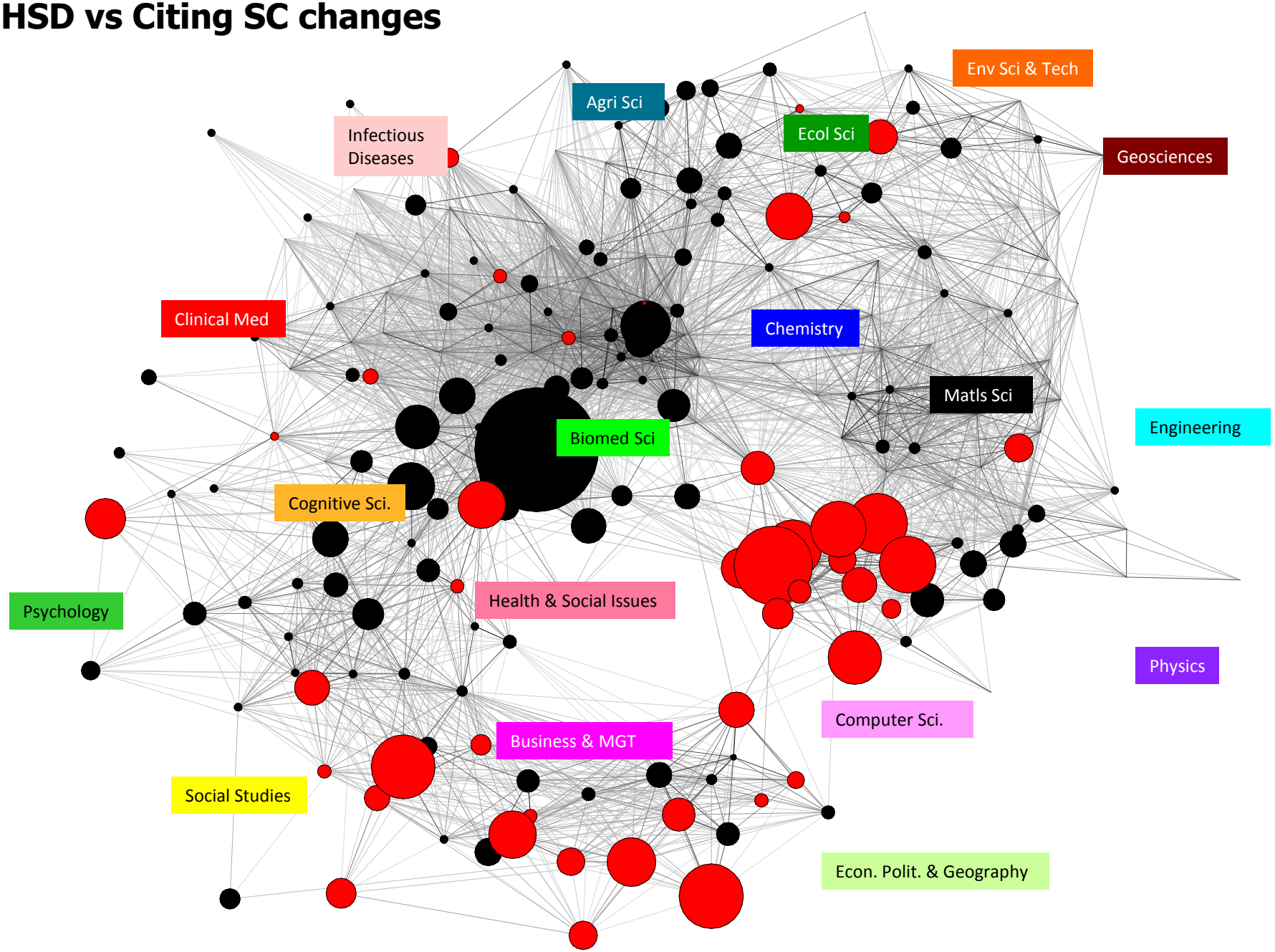
1) "Global" -- Science Overlay Maps: Show Diversity

2) "Local" – Research Network Maps: Show coherence

221 SC Base Map – Sciences + Social Sciences

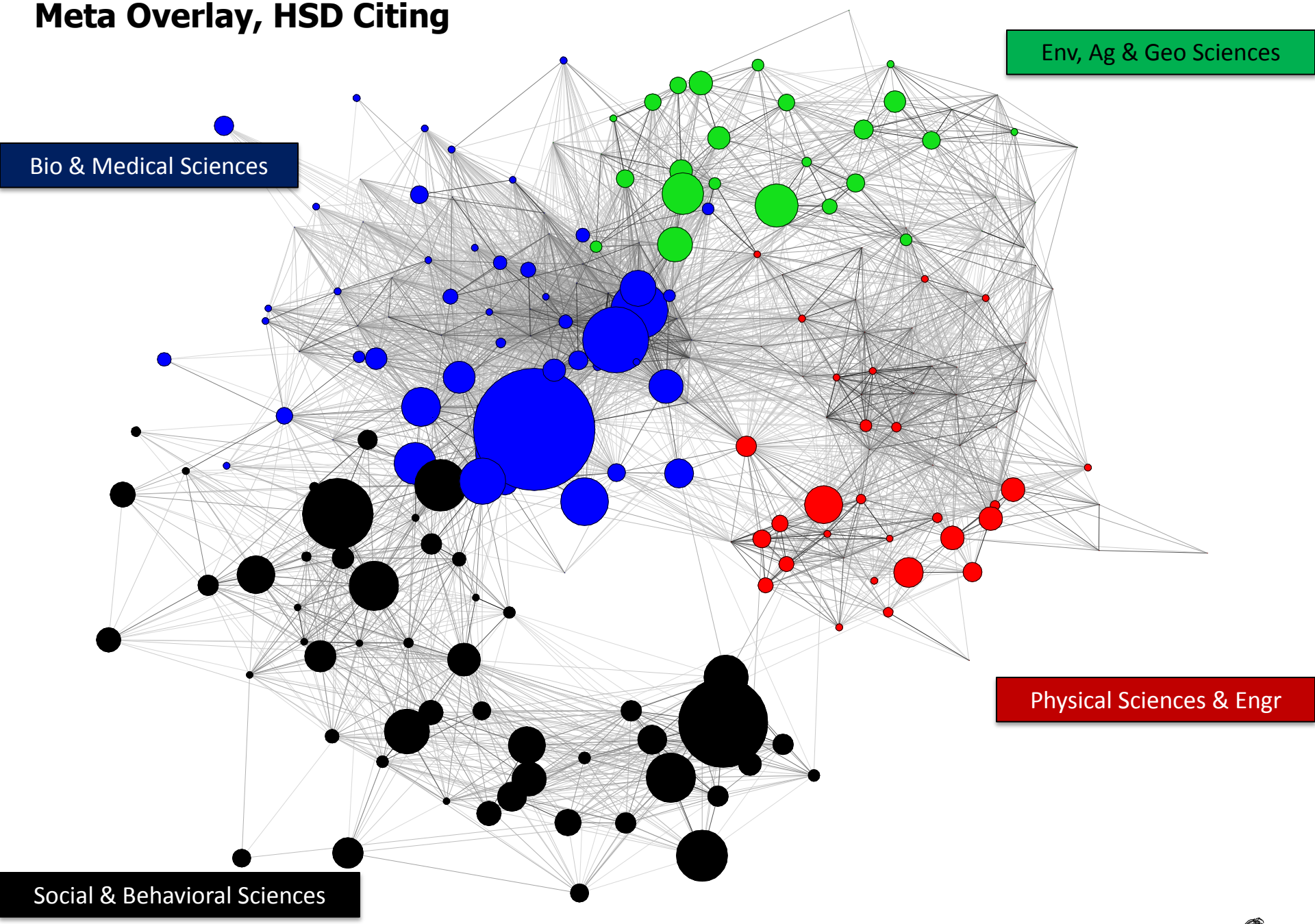


HSD vs Citing SC changes



Pajek

Meta Overlay, HSD Citing



[To identify research communities using a body of research knowledge]

68 Highly **Citing** Authors, based on shared NSF ROLE/REESE PIs & co-PIs being cited

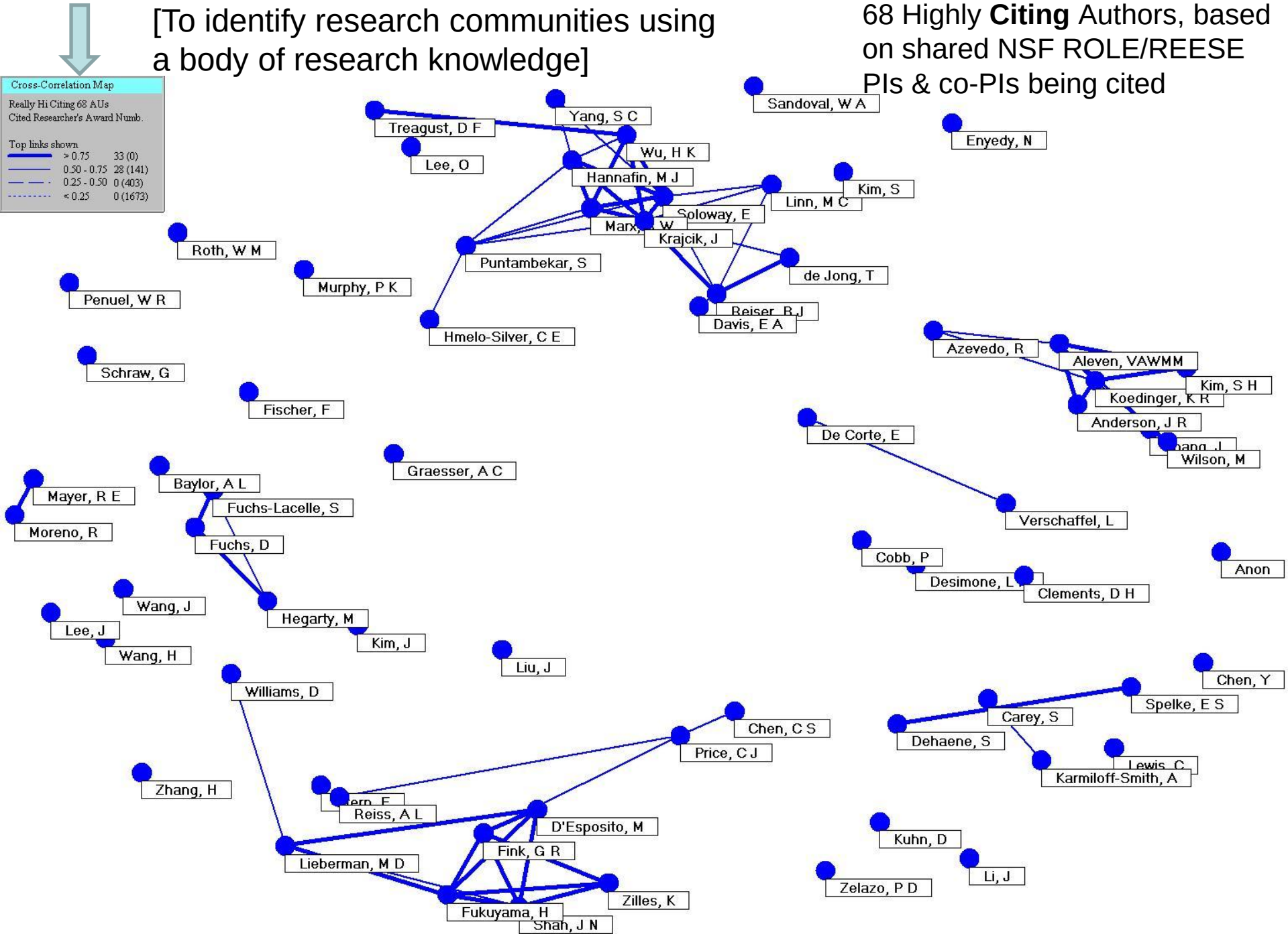
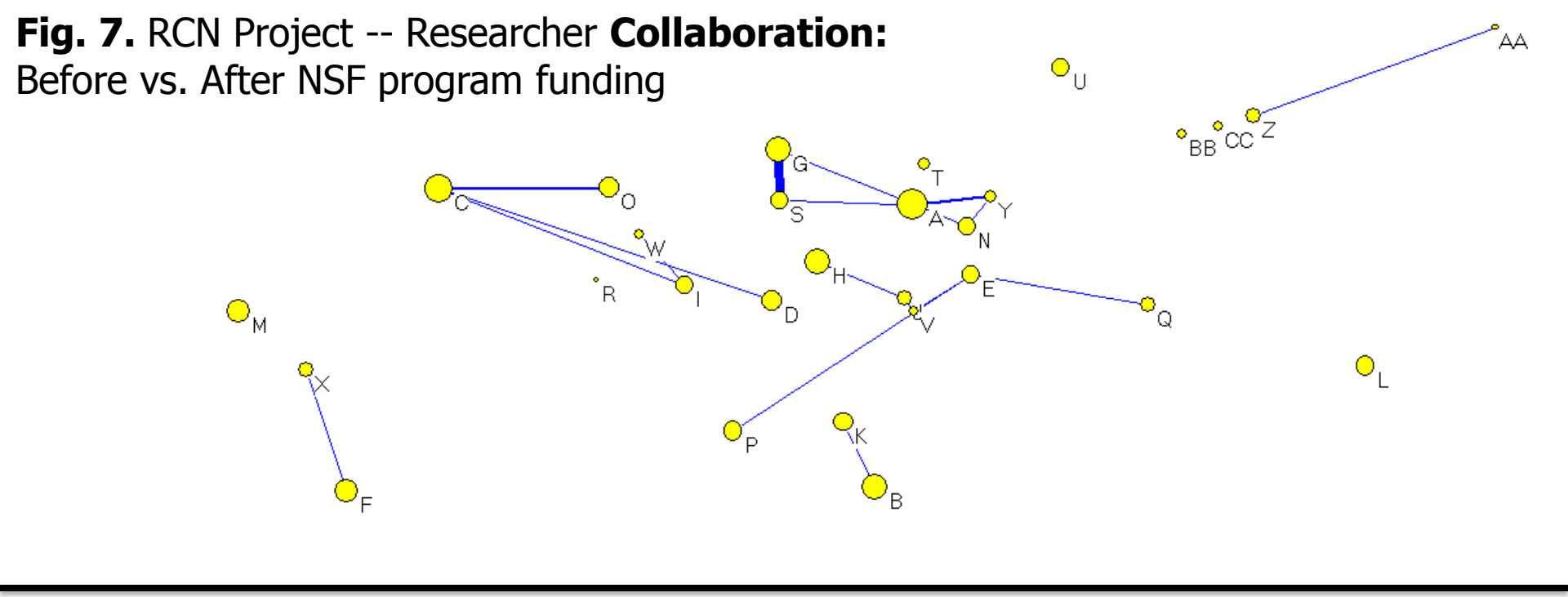
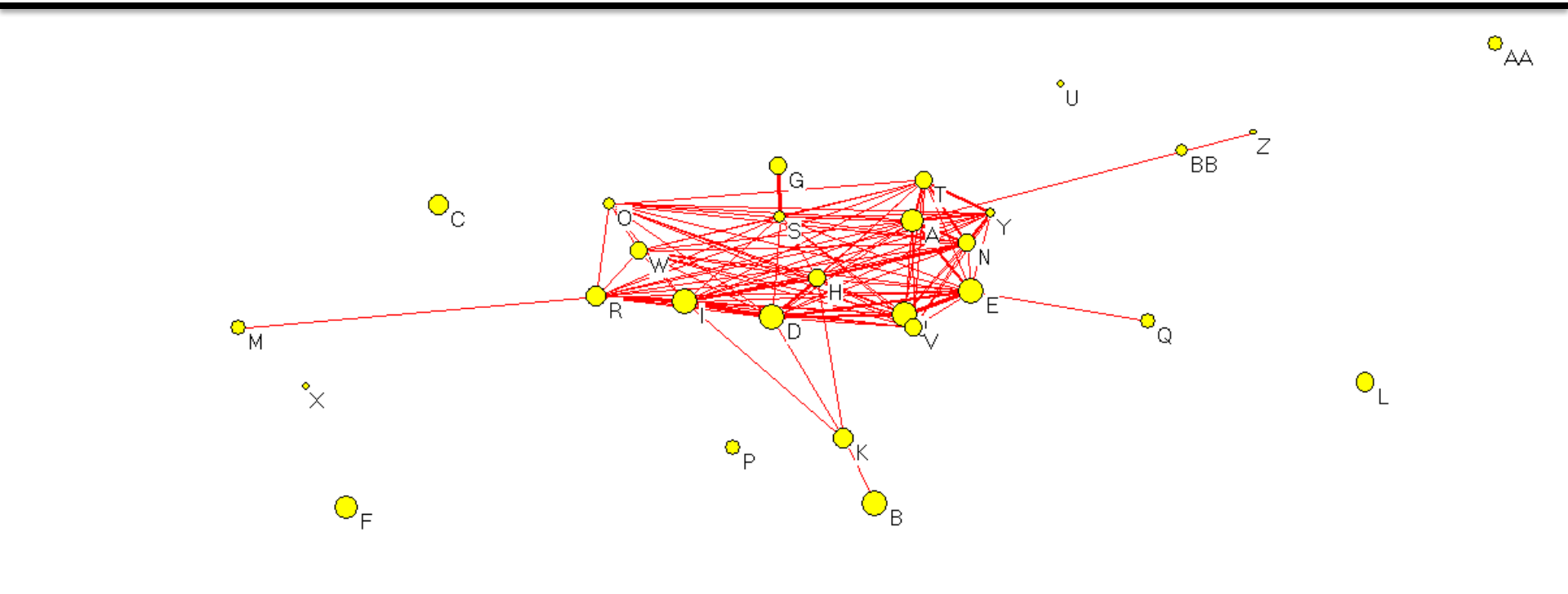
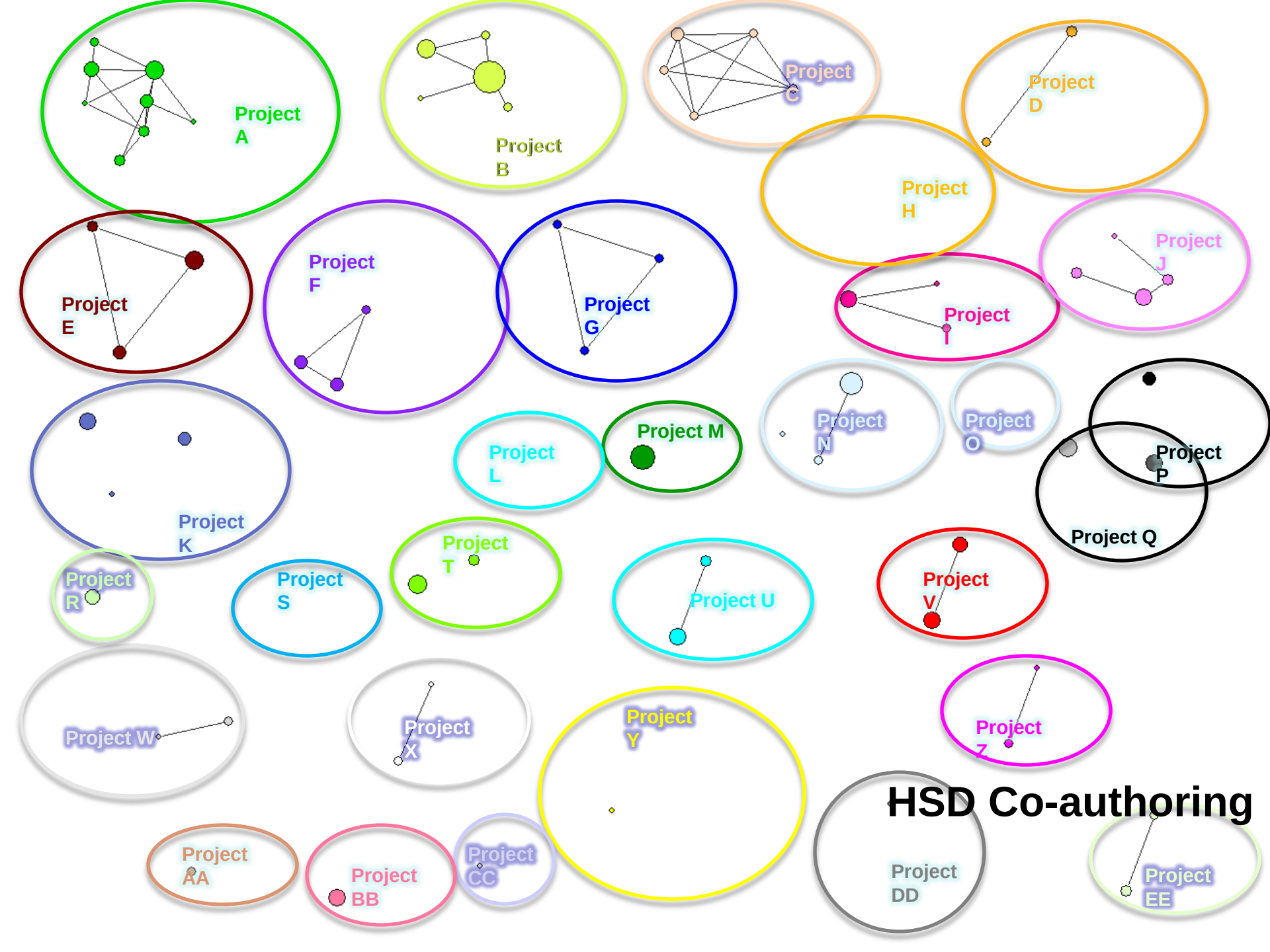
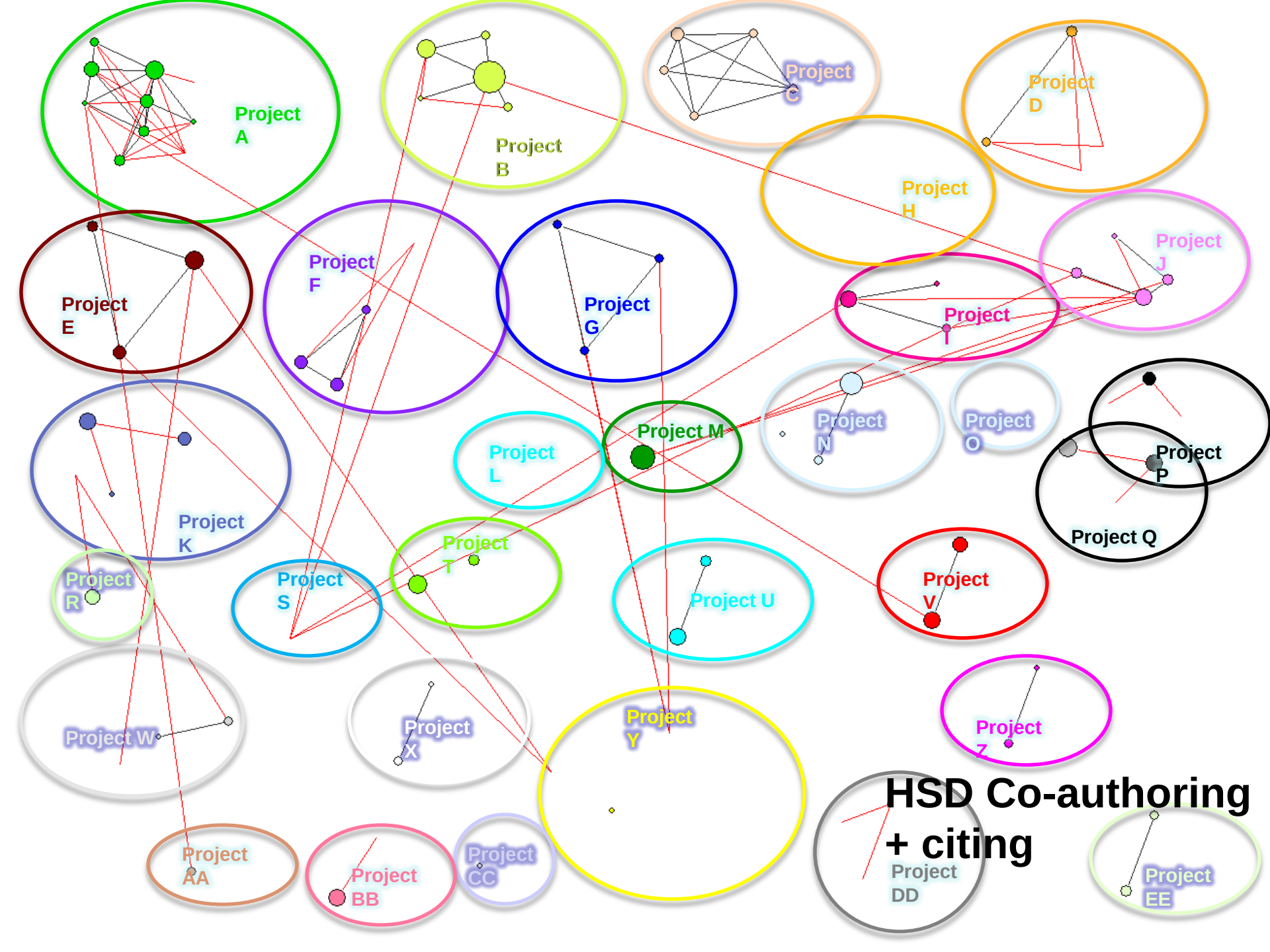


Fig. 7. RCN Project -- Researcher **Collaboration:** Before vs. After NSF program funding






Project A
Project B
Project C
Project D
Project E
Project F
Project G
Project H
Project I
Project J
Project K
Project L
Project M
Project N
Project O
Project P
Project Q
Project R
Project S
Project T
Project U
Project V
Project W
Project X
Project Y
Project Z
Project AA
Project BB
Project CC
Project DD
Project EE
HSD Co-authoring


Project A
Project B
Project C
Project D
Project E
Project F
Project G
Project H
Project I
Project J
Project K
Project L
Project M
Project N
Project O
Project P
Project Q
Project R
Project S
Project T
Project U
Project V
Project W
Project X
Project Y
Project Z
Project AA
Project BB
Project CC
Project DD
Project EE
HSD Co-authoring
+ citing

Resources

- The text mining software used: www.theVantagePoint.com
- **Ongoing Research on Interdisciplinarity & to make your own science overlay maps:** //idr.gatech.edu/ or www.leydesdorff.net/overlaytoolkit
- Global Tech Mining Conference, in conjunction with S&T Indicators Conference, Sep., 2012, Montreal
- Global Tech Mining – forthcoming special issues of *Technological Forecasting & Social Change*, and of *Technology Analysis & Strategic Management*

Research Assessment References

- Porter, A.L., Newman, N.C., Myers, W., and Schoeneck, D., Projects and Publications: Interesting Patterns in U.S. Environmental Protection Agency Research, *Research Evaluation*, Vol. 12, No. 3, 171-182, 2003.
- Porter, A.L., Schoeneck, D.J., Roessner, D., and Garner, J. (2010). Practical research proposal and publication profiling, *Research Evaluation*, 19(1), 29-44.
- Carley, S., and Porter, A.L., A forward diversity index, *Scientometrics*, to appear -- DOI: 10.1007/s11192-011-0528-1.

Interdisciplinarity References

- National Academies Keck Futures Initiative: //www.keckfutures.org
- National Academies Committee on Facilitating Interdisciplinary Research, Committee on Science, Engineering and Public Policy (COSEPUP) (2005). *Facilitating interdisciplinary research*. (National Academies Press, Washington, DC).
- Klein, J. T. (1996), *Crossing boundaries: Knowledge, disciplinarities, and interdisciplinarities*. (University Press of Virginia, Charlottesville, VA.).
- Porter, A.L., Cohen, A.S., Roessner, J.D., and Perreault, M. Measuring Researcher Interdisciplinarity, *Scientometrics,* Vol. 72, No. 1, 2007, p. 117-147.
- Porter, A.L., Roessner, J.D., and Heberger, A.E., How Interdisciplinary is a Given Body of Research?, *Research Evaluation*, Vol. 17, No. 4, 273-282, 2008.
- Porter, A.L., and Rafols, I. (2009), Is Science Becoming more Interdisciplinary? Measuring and Mapping Six Research Fields over Time, *Scientometrics*, 81(3), 719-745.
- Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. Journal of The Royal Society Interface, 4(15), 707-719.
- Wagner, C.S., Roessner, J.D., Bobb, K., Klein, J.T., Boyack, K.W., Keyton, J., Rafols, I., and Borner, K. (2011), Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature, *Journal of Informetrics,* 5(1), 14-26.

Science Mapping References

Science Maps

- Chen, C. (2003) *Mapping Scientific Frontiers: The Quest for Knowledge Visualization*, Springer, London
- Boyack, K. W., Klavans, R. & Börner, K. (2005). Mapping the backbone of science. Scientometrics, 64(3), 351-374.
- Leydesdorff, L. and Rafols, I. (2009) A Global Map of Science Based on the ISI Subject Categories. Journal of the American Society for Information Science and Technology, 60(2), 348-362.
- Boyack, K. W., Börner, K. & Klavans, R. (2009). Mapping the structure and evolution of chemistry research. Scientometrics, 79(1), 45-60.
- Klavans, R. & Boyack, K. W. (2009). Toward a Consensus Map of Science. *Journal of the American Society for Information Science and Technology*, 60(3), 455-476.
- Places & Spaces: http://www.scimaps.org/

Science Overlay Maps

- Rafols, I. & Leydesdorff, L. (2009). Content-based and Algorithmic Classifications of Journals: Perspectives on the Dynamics of Scientific Communication and Indexer Effects. *Journal of the American Society for Information Science and Technology,* 60(9), 1823-1835.
- Rafols, I., Porter, A.L., and Leydesdorff, L., (2010) Science overlay maps: A new tool for research policy and library management, *Journal of the American Society for Information Science & Technology*, 61 (9), 1871-1887, 2010.
- Rafols, I. and Meyer, M. (2009) Diversity and Network Coherence as indicators of interdisciplinarity: case studies in bionanoscience. *Scientometrics*, 82(2), 263-287. DOI 10.1007/s11192-009-0041-y.
- Porter, A.L., and Youtie, J., Where Does Nanotechnology Belong in the Map of Science?, *Nature-Nanotechnology*, Vol. 4, 534-536, 2009.

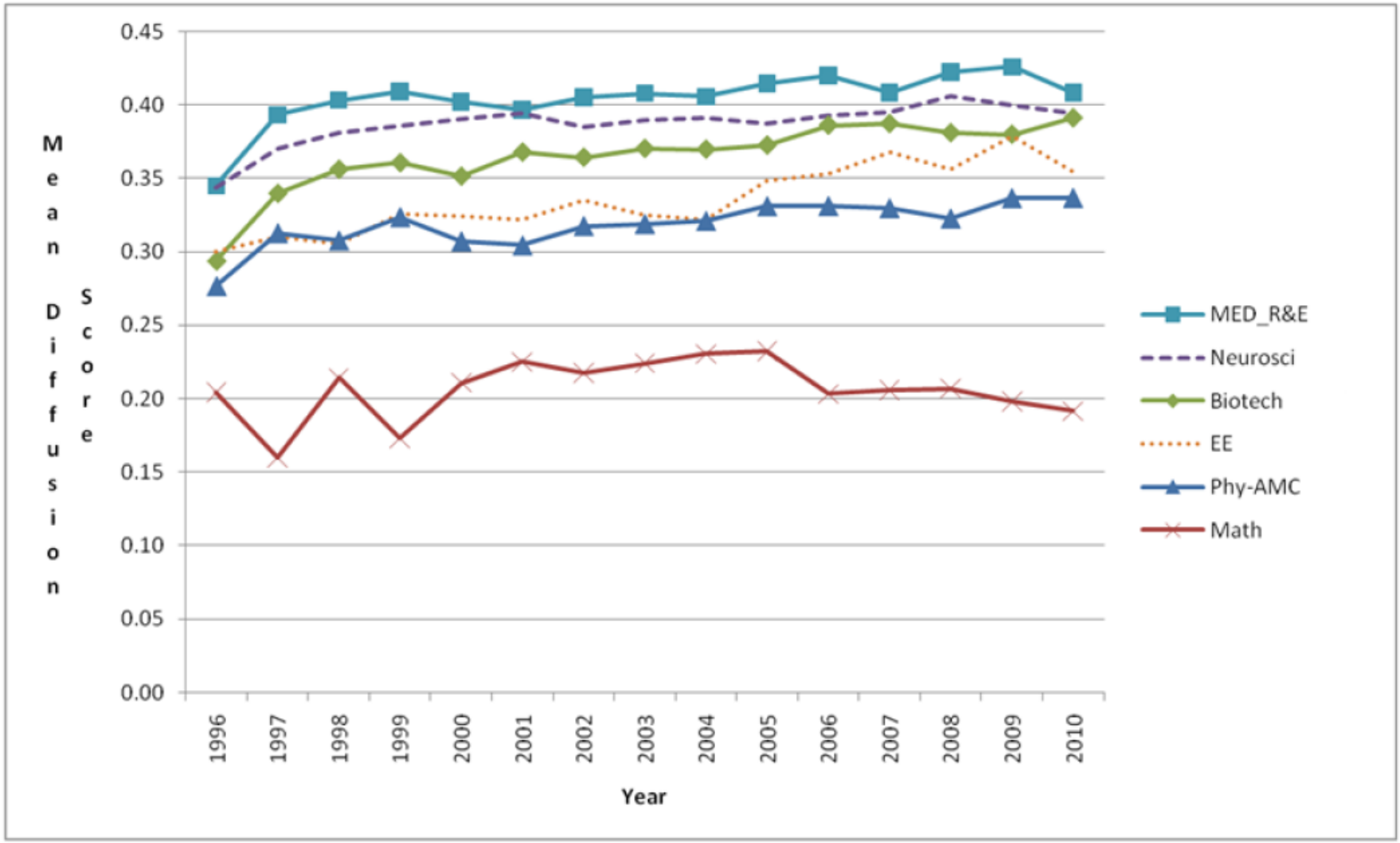
Summing Up

1. Framework for Tracking & Assessing research knowledge transfer
2. Search & retrieval from Web of Science
3. Measuring Interdisciplinarity
4. Visualization
   - Science overlay maps (locating research activity)
   - Research networking maps
5. Illustrations from NSF & other research assessments

Could you make use of these capabilities?

Outtakes

Mean Annual Diffusion Scores for 6 Subject Categories



For most of the **1995 benchmarks,** Diffusion scores increase steadily with time. Mathematics is an outlier.

Quasi-Experimental Designs

- From publications
  - Mainly compare: Before vs. After
  - Special focus: Papers deriving from NSF support
- From citations
  - By researcher publications, or proposals
  - To researcher publications
- For Target & Comparison Group researchers
- Networks based on
  - Social links [e.g., co-authoring]
  - Intellectual links [e.g., cross-citing or bibliographic coupling on SCs, topics, or whatever]

National Academies Keck *Futures* Initiative
[*Facilitating Interdisciplinary Research*]

*www.**keckfutures**.org*

Interdisciplinary research (IDR) is a mode of research by teams or individuals that *integrates*

- **perspectives/concepts/theories and/or**
- **tools/techniques and/or**
- **information/data**

from two or more bodies of specialized knowledge or research practice. Its purpose is to advance fundamental understanding or to solve problems whose solutions are beyond the scope of a single field of research practice.

Examples of *bodies of specialized knowledge or research practice* include: low temperature physics, molecular biology, developmental psychology, toxicology, operations research, and fluid mechanics.

The 225 Web of Science Subject Categories [science & social science], used to categorize journals, reflect this granularity well

Integration Score

Porter et al. (2007)

"cos (SCi – SCj)" measures the association between two SCs, based on a national co-citation sample from Web of Science. It reflects the relative tendency of two particular SCs to be co-cited.

****equivalently,**

$$I = 1 - \sum_{i,j} p_i p_j s_{ij}$$

Rafols and Meyer (2009)

where p_i is the proportion of references citing the SC i in a given paper. The summation is taken over the cells of the SC x SC matrix. s_{ij} is the cosine measure of similarity between SCs i and j

[This measure is basically 1 – Stirling D.]

Multiple Mapping Approaches

- Science overlay mapping
  - Working on patent overlay maps
  - Working on biomedical overlay maps (MEDLINE)
- Geo-maps
- Research Network Mapping [Social Network Analyses]
  - Co-authoring; co-citation; co-term; etc.
  - Bibliographic coupling

Science Mapping

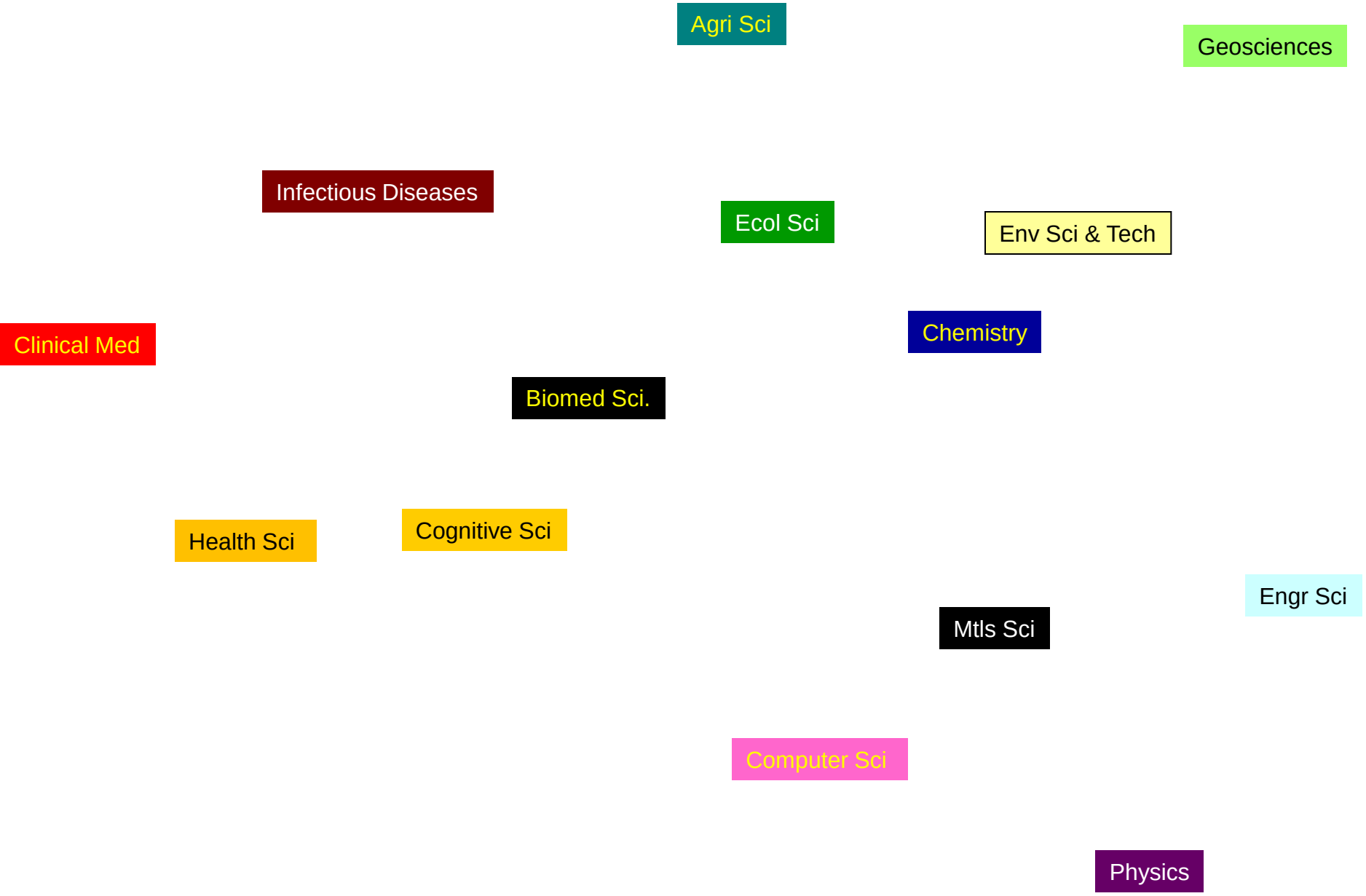
- Based on Bibliometrics
- Since the 1970's
- Chaomei Chen, *Mapping Scientific Frontiers*, 2003; CiteSpace site
- Usually **local** – research networking in a specific research arena
- Recently also -- **global** mapping -- "all" of science – Klavans, Boyack, Borner; Leydesdorff, Rafols, Meyer, Porter;
  - ~Robust to different data and representations

Science Overlay Mapping

- Rafols & Leydesdorff (with Meyer, Porter)
- Based on Web of Science (WoS)
  - Subject Categories (SCs; recast as Web of Science Categories – WCs – with WoS v. 5, late 2011)
  - Can do for Science (Science Citation Index) ~175 SCs, or
  - Science + Social Science (include Social Science Citation Index) ~224 SCs
- Base map
  - Nodes (SCs) and background links -- derive from an SC-by-SC cosine similarity matrix from a year of journal cross-citation data (recently, 2010)
  - Labels reflect groupings of SCs
- Overlays – a given body of research activity (e.g., a set of publications indexed in WoS)

Macro- and Meta-Disciplines

1. SC relatedness based on one year's data – WOS Journal X Journal **cross-citation** matrix
2. Loet Leydesdorff transforms to SC X SC matrix
   - Devise our interdisciplinarity metrics based on these
3. **Macro-Disciplines** come from Ismael Rafols' factor analyses:
   - 175 SC science base map (14 factors)
   - **224 SC science + social science base map (19 factors = Macro-Disciplines)**
4. **Meta-Disciplines** – we can further group to 4 or 6 overarching categories



"**Nanoscience & Nanotechnology" Subject Category** – 3863 articles,
partial 2008 – showing Top 40 Cited SCs Overlay over base 175 SC Science Map

Co-citation Map of the most cited authors by the 307 **nano social science** papers [Use Auto-corr on hi cited Authors]

Co-citation Map

- Mining the "CR" field – Cited References here [list]
- Imported the "Cited authors" field [really 1st authors]
- Cleaned them
- Selected those with 5 or more cites [list]
- #5) Map [note Seliktar & Stegemann]

Co-citing Map

- Different beast
- Need to search & retrieve the papers citing Nerem
- Work in a big, new VantagePoint file then
- Cross-correlation map – e.g., high citing authors with the measure of association being that they tend to cite the same Nerem papers
- Example (based on all 240 Nerem papers)