Studies of convergence using network analysis and natural language processing

Daniel McFarland

For presentation at Workshop on the Implications of Convergence for How the National Center for Science And Engineering Statistics Measures the Science and Engineering Workforce

What is convergence? (NRC 2014)

- "the coming together of *insights* and *approaches* from originally *distinct fields*" - will make fundamental contributions to providing *creative solutions* to the most difficult *problems* facing society (vii).
- "an approach to problem solving that *cuts across disciplinary boundaries*…to form a comprehensive *synthetic framework* for tackling" scientific challenges that exist *at the interfaces of multiple fields* (1).
- *cross-fertilization of ideas* with *diverse stakeholders* (funding, research, applied /basic, industry) – *expanded form of interdisciplinarity and translational research* (17).

My “charge”: Derived measures using network analysis and natural language processing

- Will discuss a series of papers that explore two vantages on convergence
  - *Social convergence*
    - How stakeholders relate to one another in the process of doing intellectual work (creating, consuming, applying knowledge)
  - *Intellectual convergence*
    - How conceptual relations and understandings come together from scientific discovery

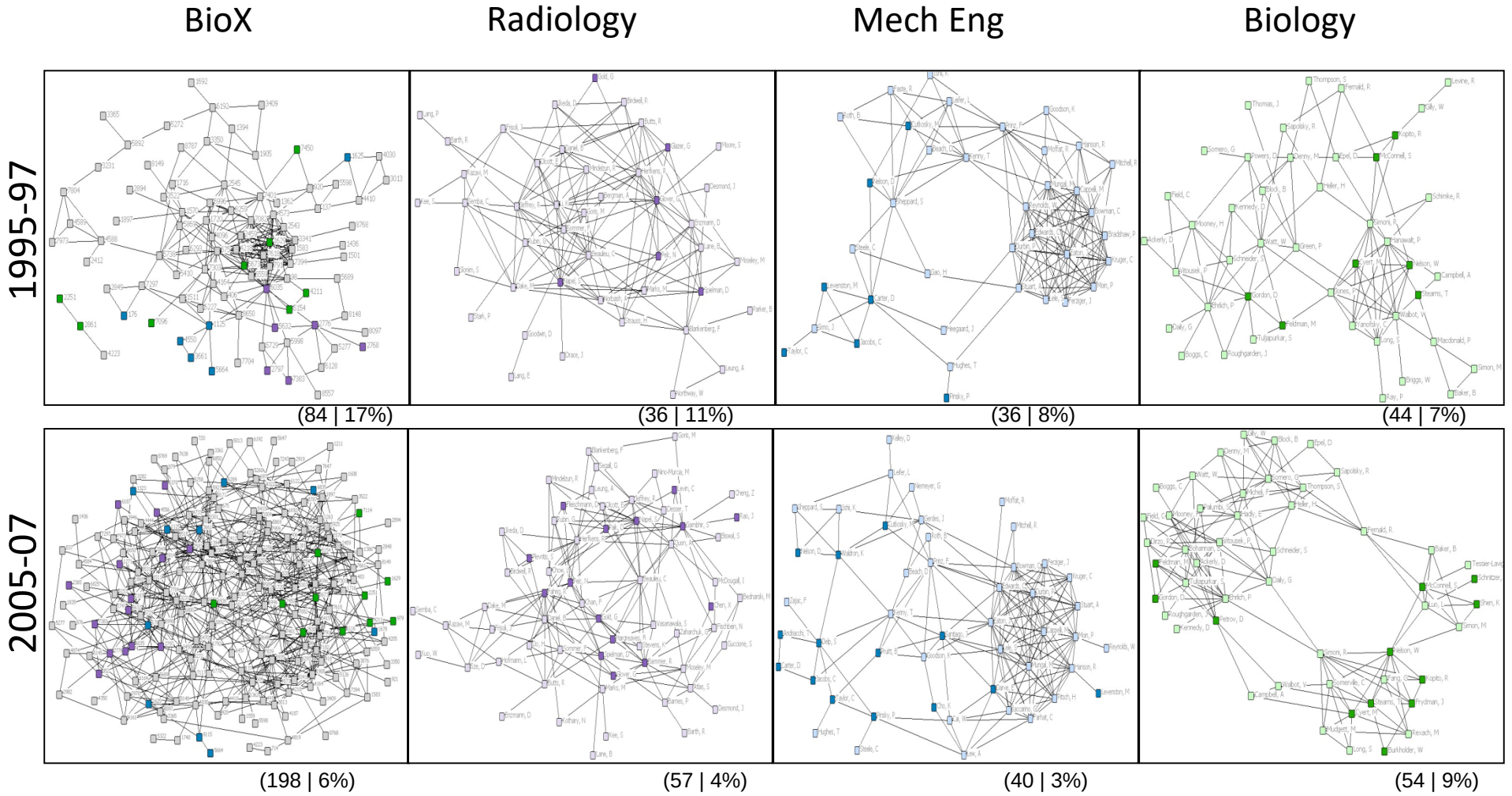
Considering social and intellectual convergence by unit of analysis

Analytic dimension		Social			
	Unit of analysis	Domain	Communities	Teams/Labs	Individual(s)
Intellectual	Domain	<u>Interdisciplinary region</u> -*Trading zone*			
	Movements / Programs		<u>Shared project</u> -*Interdisciplinary center*		
	Arguments			<u>Shared problem</u> -paper, patent -*Boundary object*	
	Concept(s)				<u>Spanning relations</u> -*Bridging* collaboration or conceptual relations

Examples of Social Convergence

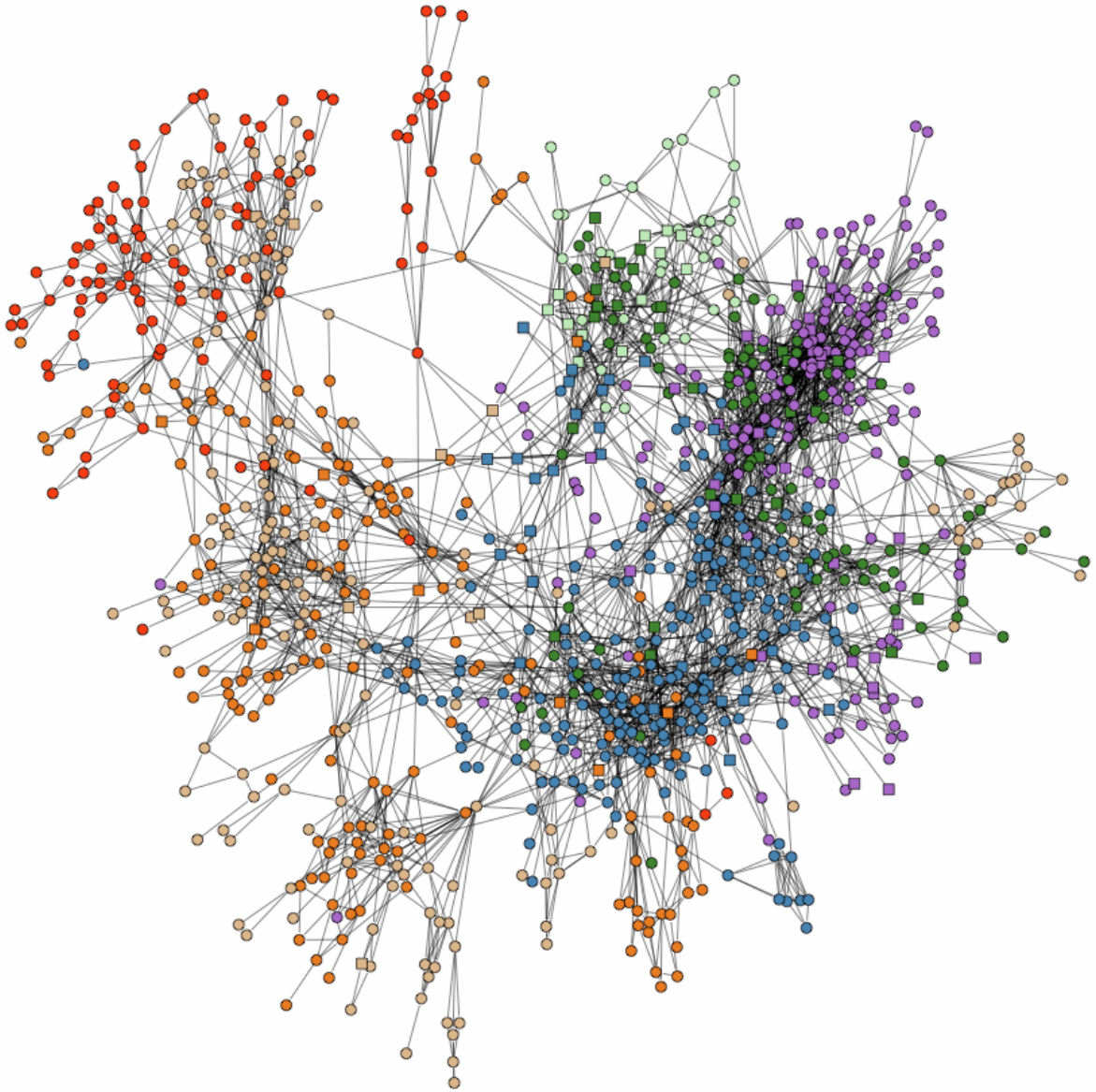
- Mostly focused on integrating, expanding, and accelerating collaboration networks
  - Q: How do interdisciplinary centers shape collaboration in universities?
    - Do centers generate productivity and new collaborations?
    - Are center collaborations “in addition” to disciplinary ones?
    - Do centers induce different forms of social convergence?
      - Focus is on Stanford and its interdisciplinary initiatives:
        - Thoroughfares and new communities (Biancani et al 2013)
        - More interdisciplinarity ➔ unequal recognition (Biancani et al 2018)

Comparison of BioX and Core Departments



Data: ties from coauthoring grants and publications or co-advising dissertations over 3 years.
Key: Dark nodes = center members. (X%) = percentage of active faculty who are isolates (disconnected).

Univers
faculty

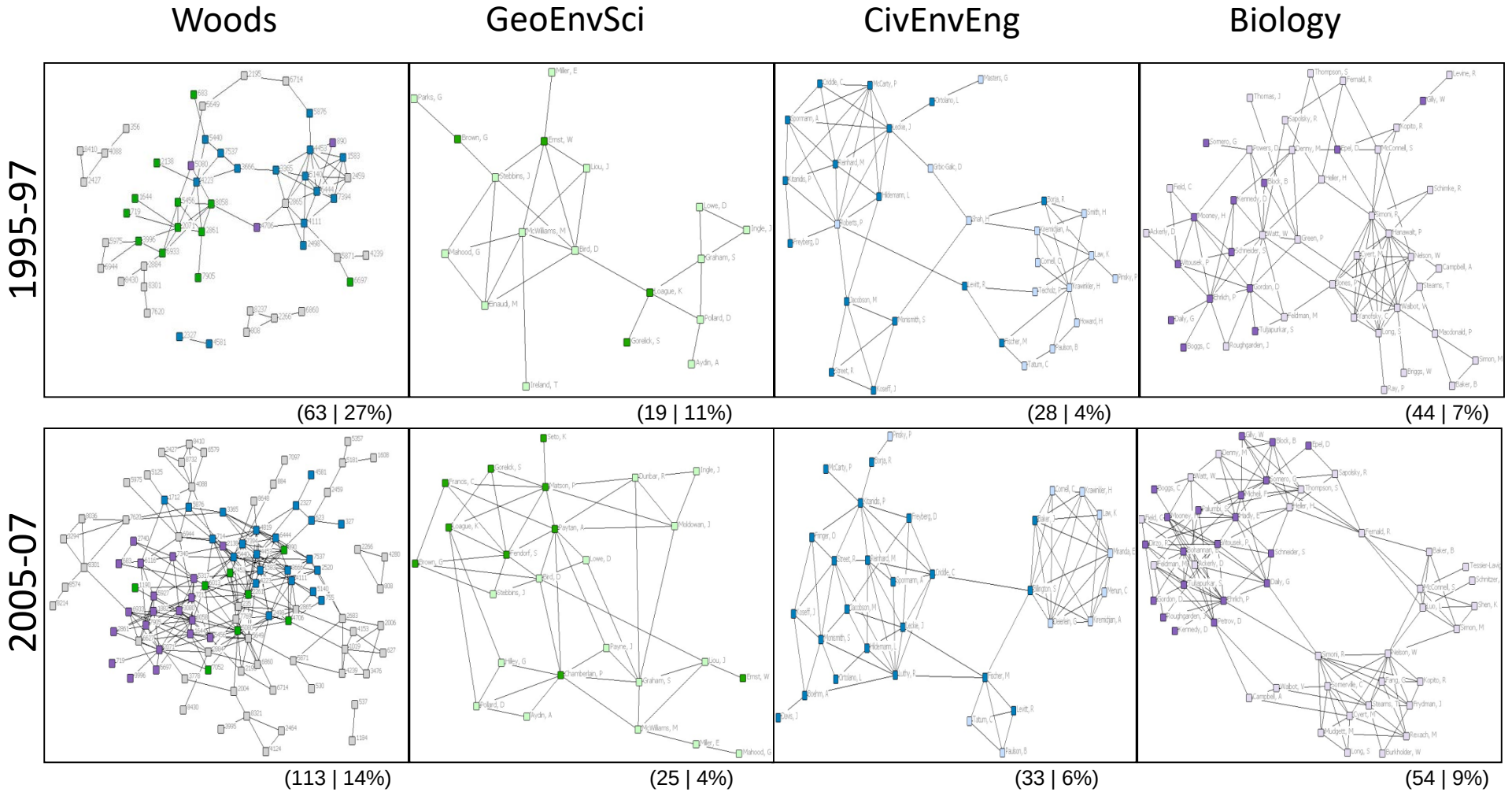


BioX
ittees

lty

Network

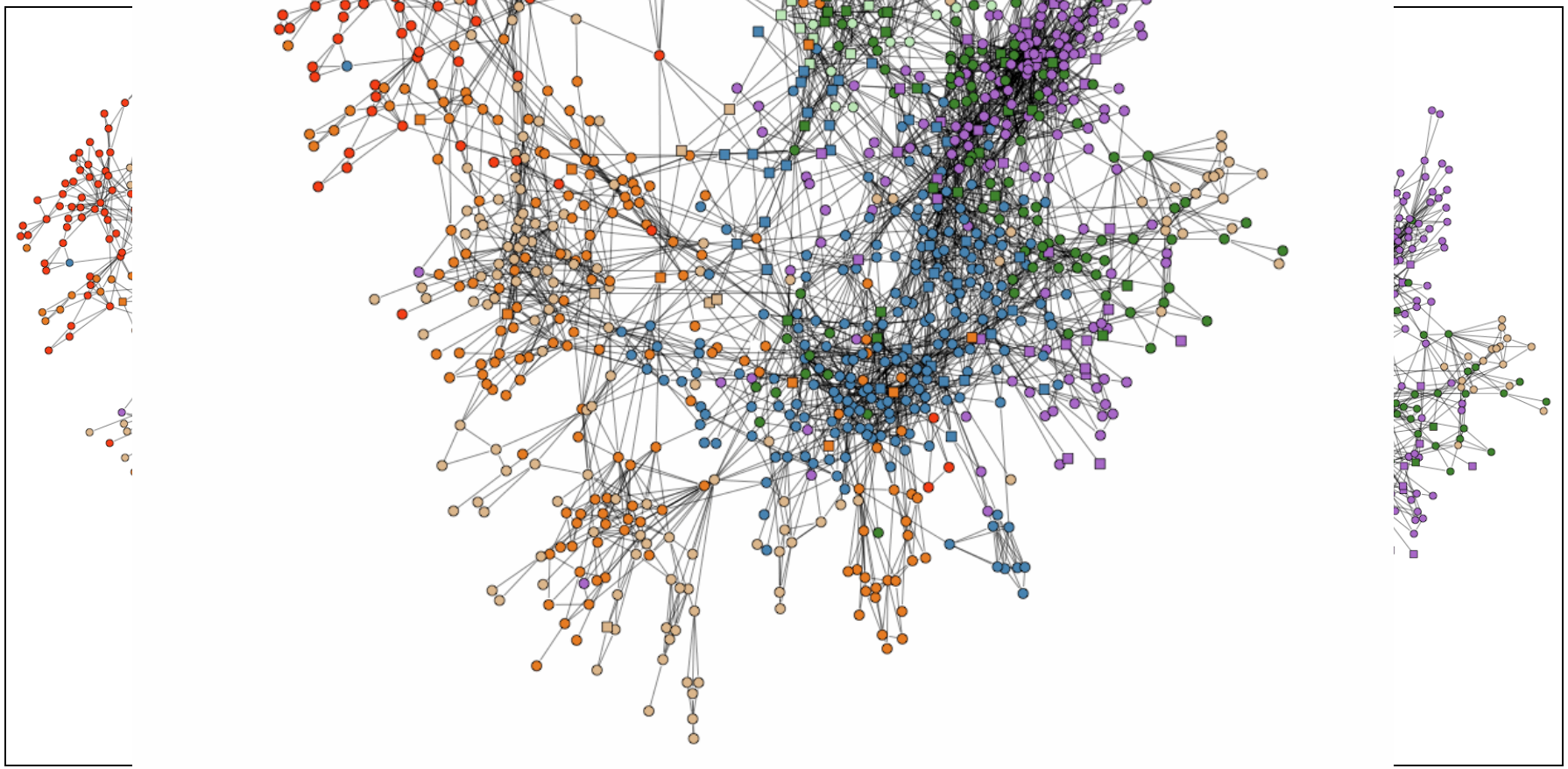
Comparison of Woods and Core Departments



Data: ties from coauthoring grants and publications or co-advising dissertations over 3 years.
Key: Dark nodes = center members. (X | X%) = # active faculty & % of active faculty who are isolates (disconnected).

Universi faculty t

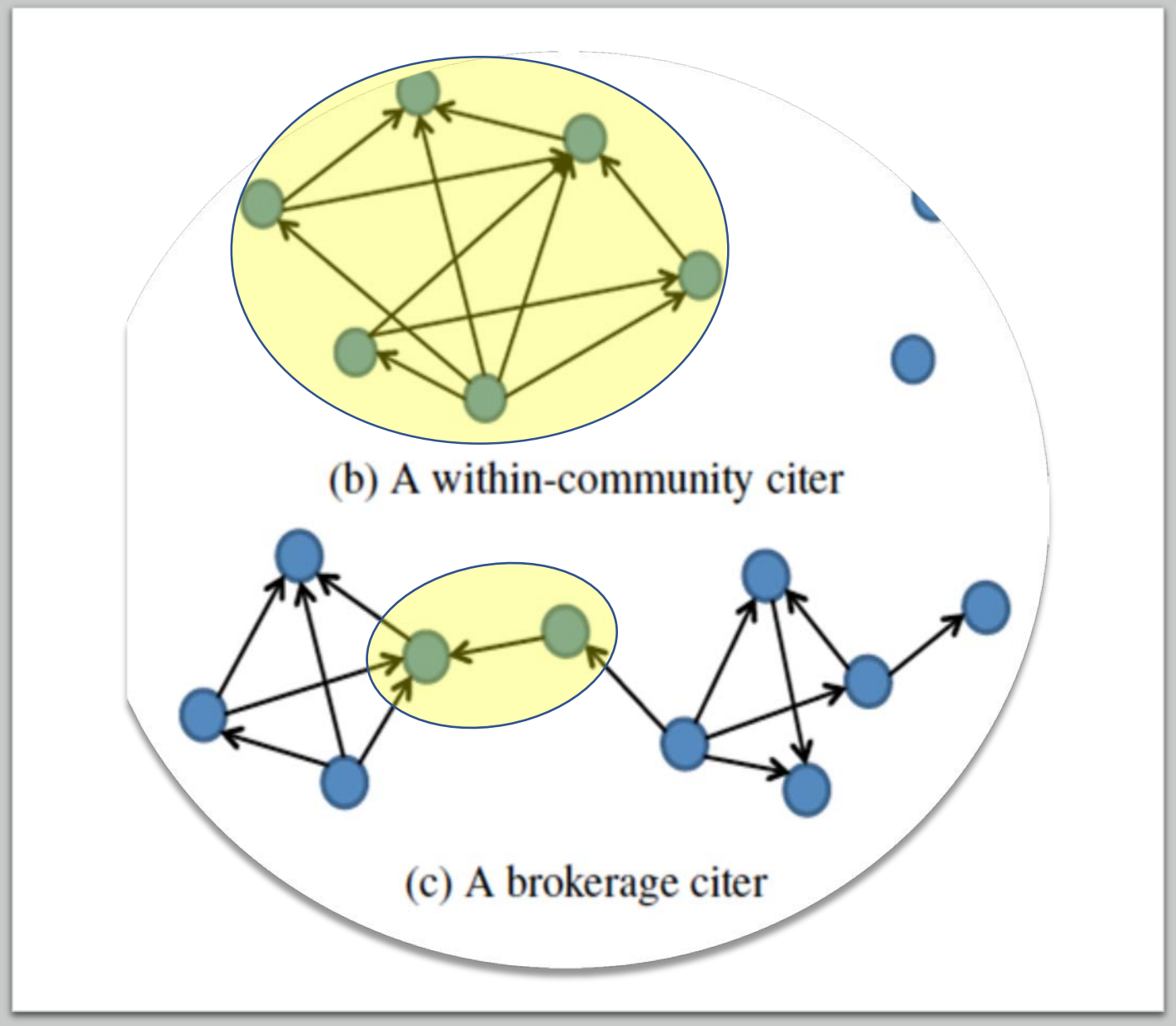
Woods ttees



Network c

Examples of Intellectual convergence

- Integration of intellectual efforts
  - Citation as proxy (i.e., pattern of intellectual consumption)
    - Papers with bridging citations interrelate distinct subfields and result in high risk high reward (Shi et al 2010)



Examples of Intellectual convergence

- Co-citation as proxy continued…
  - Study how annual review papers – attempting synthesis – shape citation networks (McMahan & McFarland)
  - Reviews poach citations.
  - Reviews perform *creative destruction* on citation networks.

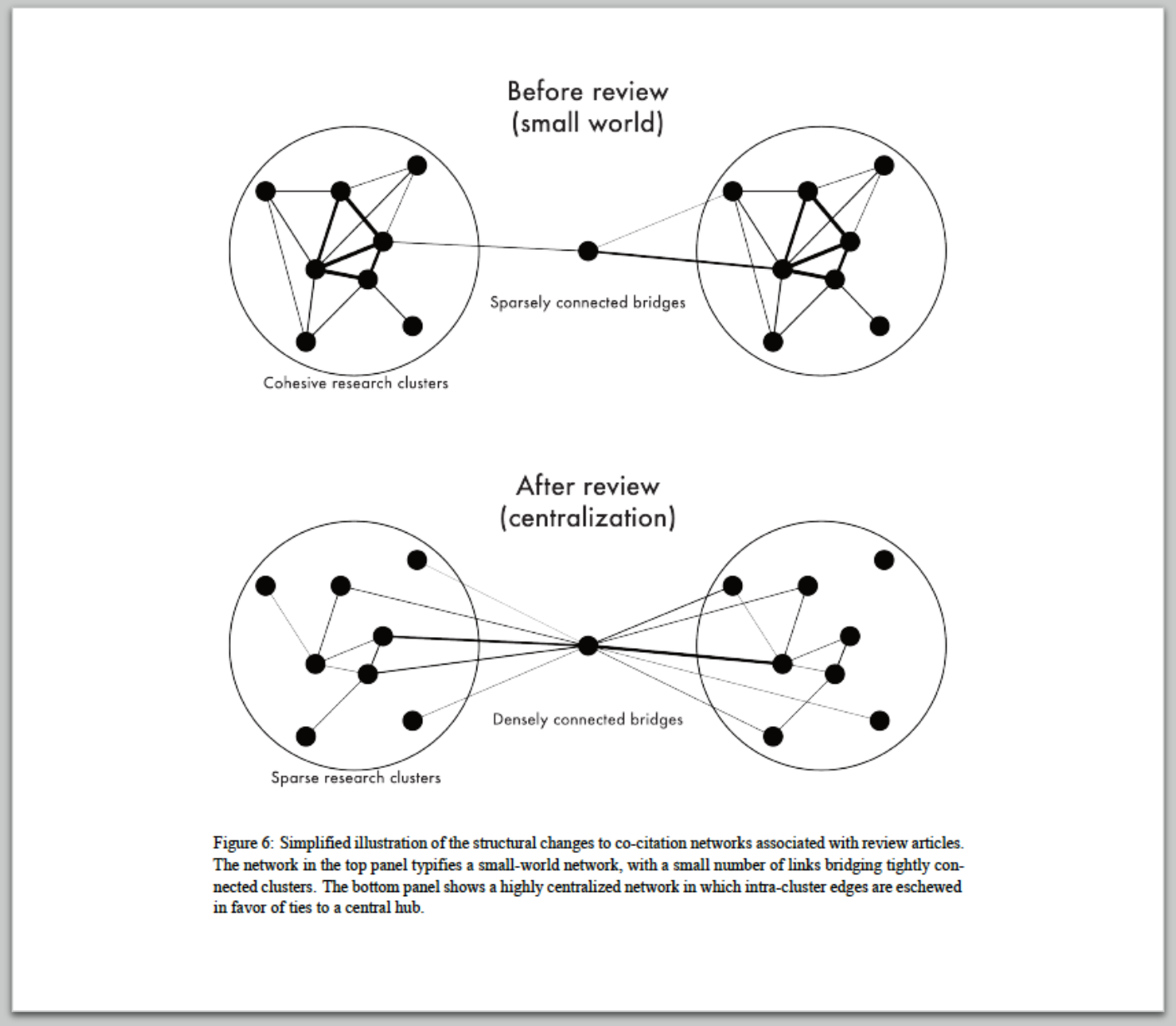


Figure 6: Simplified illustration of the structural changes to co-citation networks associated with review articles. The network in the top panel typifies a small-world network, with a small number of links bridging tightly connected clusters. The bottom panel shows a highly centralized network in which intra-cluster edges are eschewed in favor of ties to a central hub.

Examples of Intellectual convergence

- Q: How do we model social *and* intellectual convergence as co-evolving? (Stark et al 2020)
- Do faculty collaborate then converge intellectually or vice versa?
- What mechanisms facilitate social and intellectual connection?
  - Rank differences (tenure v. non)
  - Cultures differ (STEM, non-STEM)
  - Times change (pre/post-internet)

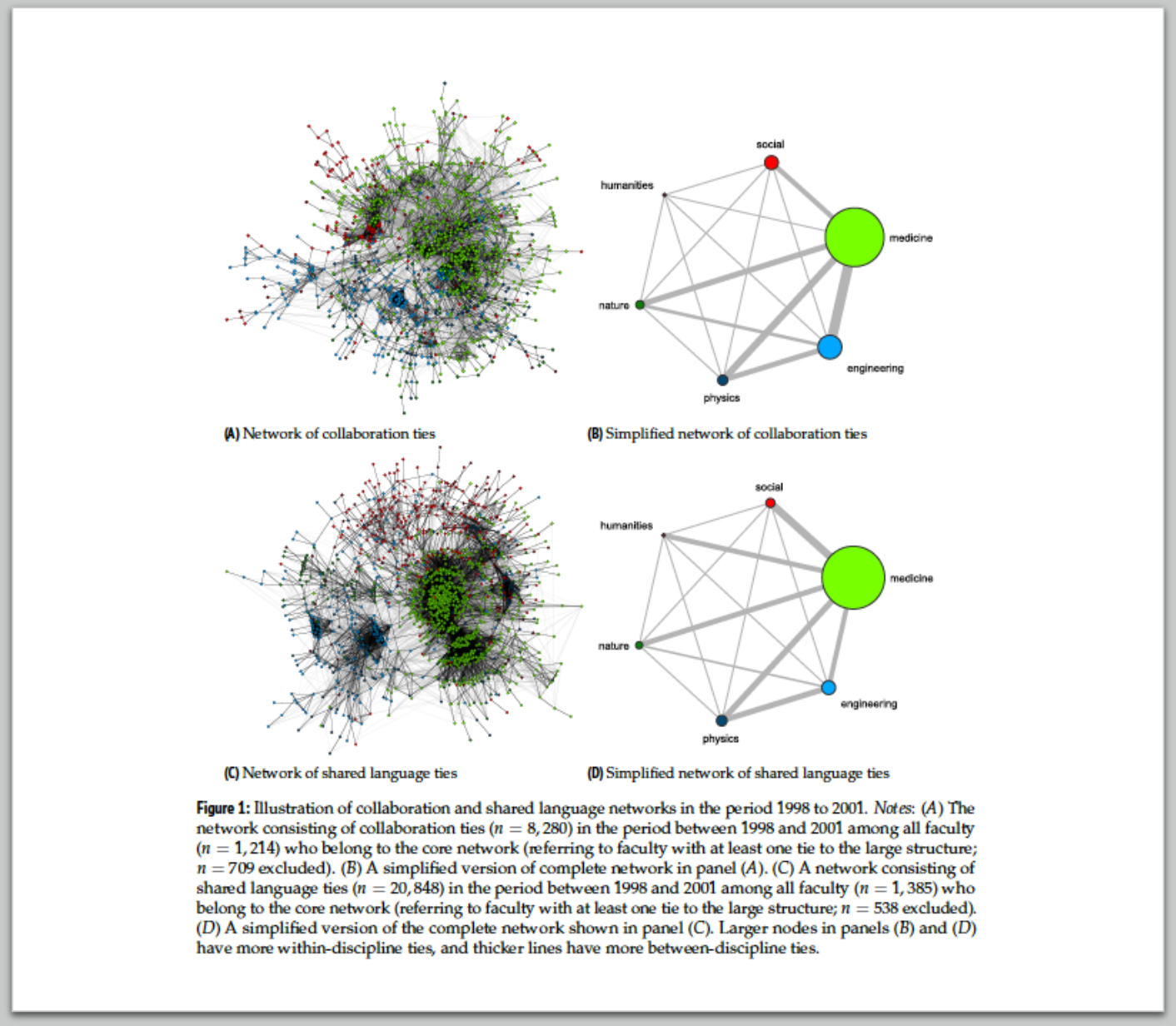


Figure 1: Illustration of collaboration and shared language networks in the period 1998 to 2001. *Notes*: (*A*) The network consisting of collaboration ties ($n = 8,280$) in the period between 1998 and 2001 among all faculty ($n = 1,214$) who belong to the core network (referring to faculty with at least one tie to the large structure; $n = 709$ excluded). (*B*) A simplified version of complete network in panel (*A*). (*C*) A network consisting of shared language ties ($n = 20,848$) in the period between 1998 and 2001 among all faculty ($n = 1,385$) who belong to the core network (referring to faculty with at least one tie to the large structure; $n = 538$ excluded). (*D*) A simplified version of the complete network shown in panel (*C*). Larger nodes in panels (*B*) and (*D*) have more within-discipline ties, and thicker lines have more between-discipline ties.

Examples of intellectual convergence

- Concept maps and their change (Kulkarni et al)
  - What are mechanisms of conceptual change and stability in the field of NLP?

Different concept effects:
-Phenomena anchor
-Methods turnover
-Cohort clumping

Different network mechanisms:
-Clusters stabilize
-Bridges turnover

Further efforts at intellectual convergence

- Ongoing questions:
  - How do we identify when scientific innovation occurs?
  - How do we trace the transfer and diffusion of innovations, and translational research?
  - How do we identify intellectual problems and their solution?

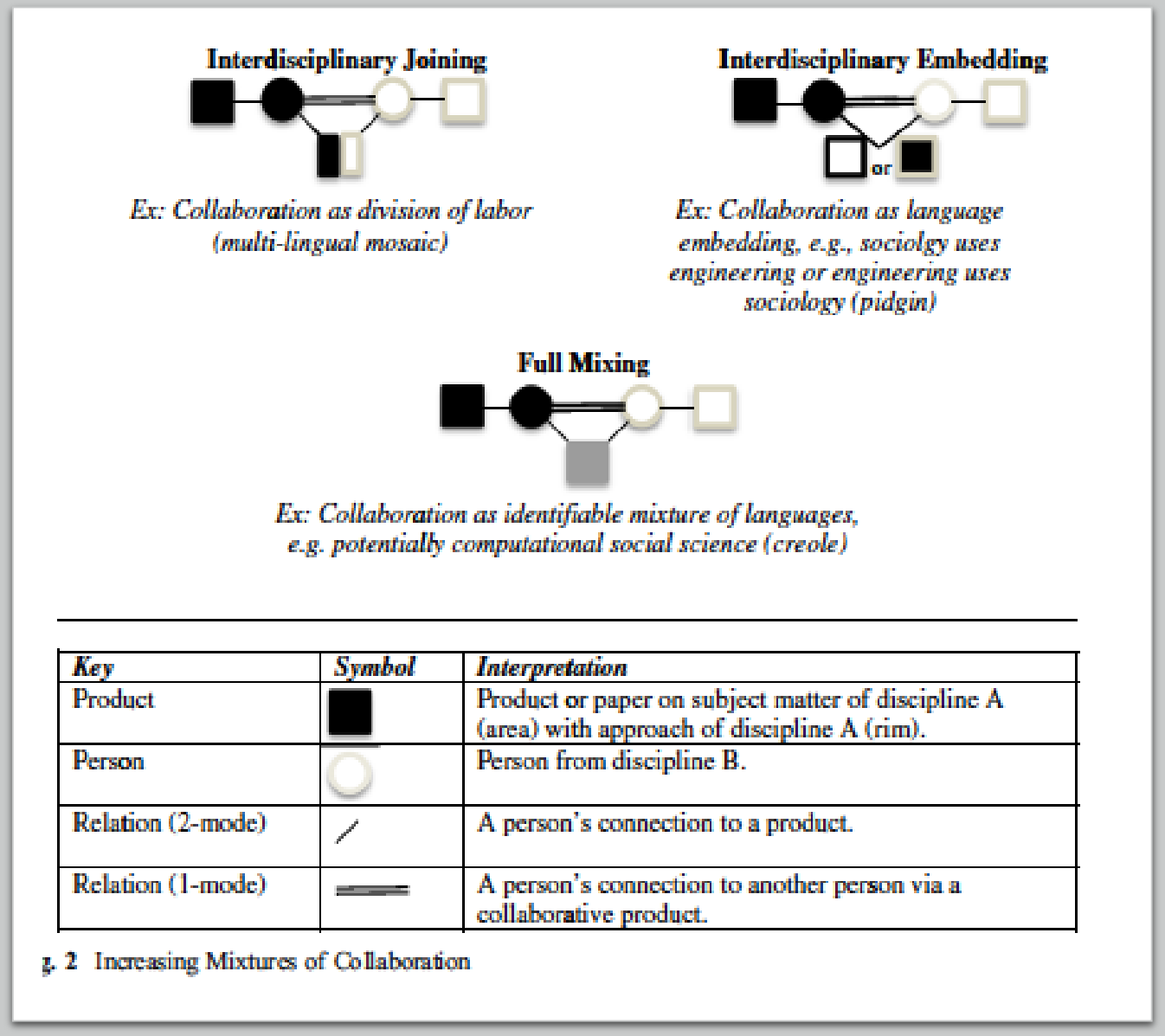
- Ongoing effort:
  - Shifts to document and concept levels (phrase extraction techniques)
  - Identify novel concepts and concept relations
  - Follow rates of adoption as “innovation” (Chen et al: working paper)
    - Ideas that bridge conceptual gaps **resonate** (a core aspect of convergence?) and get adopted
    - Proximal > distal (ease of use / familiarity issue; BUT favors short term > long term returns)
  - Follow adoption and changing sense across domains as “translation” (Cao et al 2020, EMNLP)

Cited works

- Biancani, Susan, Linus Dahlander, and Daniel A. McFarland. 2014. “The Semi-Formal Organization.” Organization Science 25(5): 1306-1324.
- Biancani, Susan, Linus Dahlander, Daniel A. McFarland, and Sanne Smith. 2018. “Superstars in the Making? The Broad Effects of Interdisciplinary Centers.” Research Policy 47: 543–557.
- Cao, Hancheng, Mengjie Cheng, Zhepeng Cen, Daniel A. McFarland, Xiang Ren. 2020. “Will This Idea Spread Beyond Academia? Understanding Knowledge Transfer of Scientific Concepts across Text Corpora.” *Empirical Methods in Natural Language Processing* (EMNLP 2020).
- Mäkinen, Elina I., Eliza D. Evans, & Daniel A. McFarland. 2020. “The Patterning of Collaborative Behavior and Knowledge Culminations in Interdisciplinary Research Centers.” Minerva 58: 71–95. https://doi.org/10.1007/s11024-019-09381-6
- McFarland, Daniel A., Kevin Lewis and Amir Goldberg. 2015. "Sociology in the Era of Big Data: The Ascent of Forensic Social Science.” The American Sociologist, 47(1), 12-35.
- McMahan, Peter and Daniel A. McFarland. 2021. “Creative Destruction: The structural consequences of scientific curation.” (Conditional accept) American Sociological Review.
- Shi, Xiaolin, Jure Leskovec, and Daniel A. McFarland. 2010. “Citing for High Impact.” *Joint Conference on Digital Libraries, (JCDL 2010*).
- Stark, Tobias H., J. Ashwin Rambaran and Daniel A. McFarland. 2020. “The Meeting of the Minds: Forging Social and Intellectual Networks within Universities.” Sociological Science.cd DOI: 10.15195/v7.a18

EXTRA:

Products and the sort of convergence that occurs in them (McFarland et al 2016)



Key	*Symbol*	*Interpretation*
Product		Product or paper on subject matter of discipline A (area) with approach of discipline A (rim).
Person		Person from discipline B.
Relation (2-mode)		A person's connection to a product.
Relation (1-mode)		A person's connection to another person via a collaborative product.

g. 2 Increasing Mixtures of Collaboration