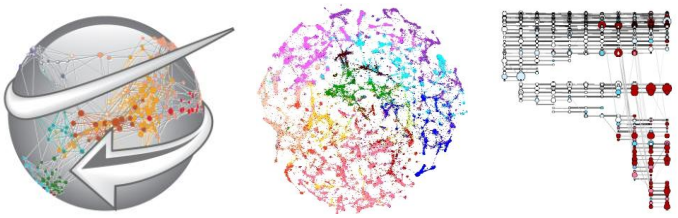
SciTech Strategies, Inc.

BETTER MAPS • BETTER DETECTION



Science Mapping of Gaps and Overlaps in a Research Portfolio

NIH Portfolio Analysis Symposium
Washington, DC
July 24, 2012

Kevin W. Boyack, SciTech Strategies

Agenda

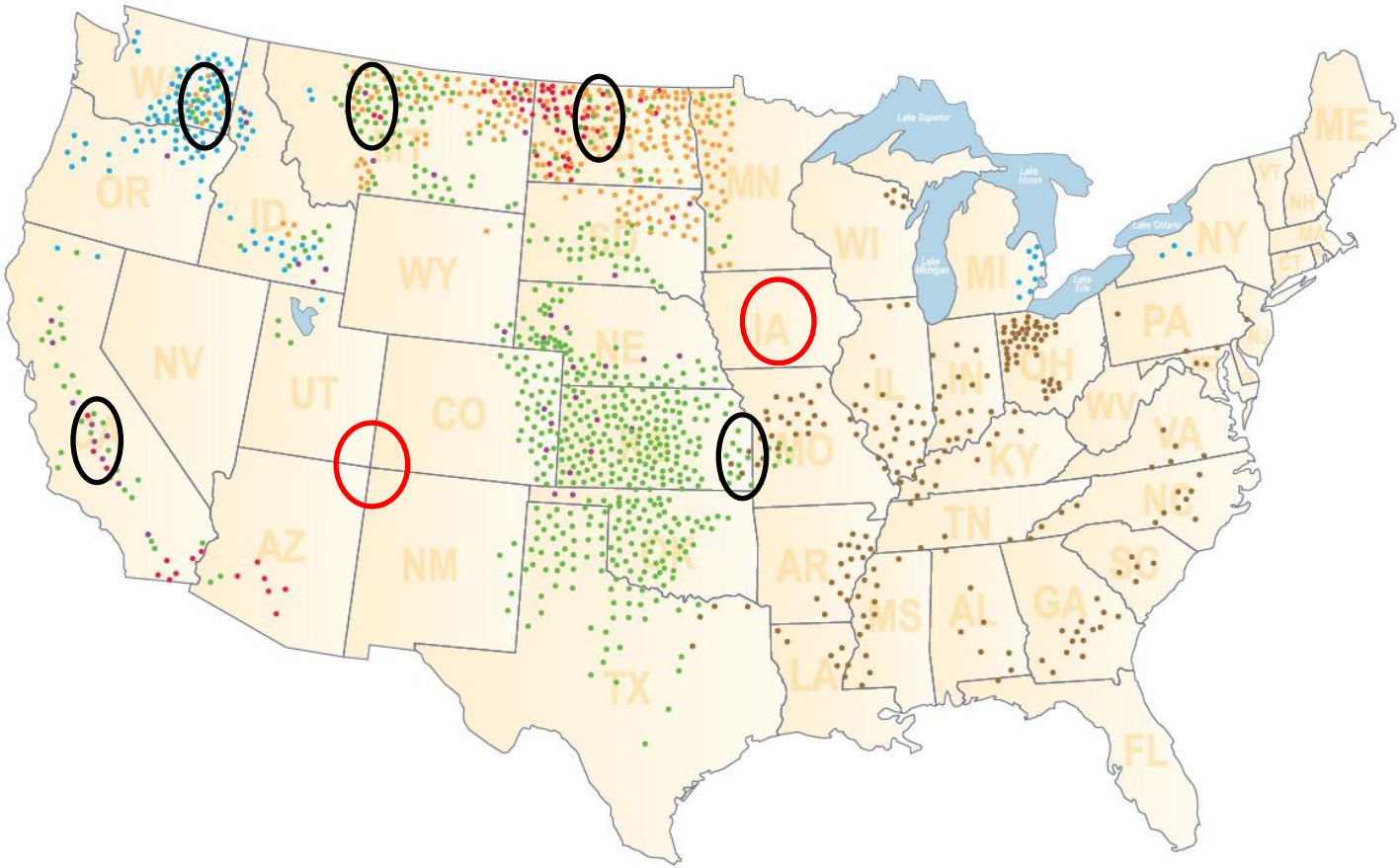
- A Geographic Analogy
- A Detailed Global Map of Science
- Overlaps
- Gaps
- Identifying Potentially Transformative Research
- Summary

Geographic Maps

- Maps show physical space
  - Existence
  - Location
  - Distance
- Maps provide a basis on which other features can be displayed and compared
  - Resources
  - Production
  - …
- Overlaps and gaps in features

Wheat Production in the U.S.

Overlaps in wheat types



Gaps in wheat production

HARD RED WINTER · HARD RED SPRING · SOFT RED WINTER · SOFT WHITE · HARD WHITE · DURUM

Science Maps Can Be Used the Same Way

- **GEOGRAPHIC MAPS**
- Maps show physical space
  - Existence
  - Location
  - Distance
- Maps provide a basis on which other features can be displayed and compared
  - Resources
  - Production
  - …
- Overlaps and gaps in features

- **SCIENCE MAPS**
- Maps show abstract space
  - Existence
  - Relative location
  - Relative distance
- Maps provide a basis on which other information can be displayed and compared
  - Resources (e.g., funding)
  - Production (e.g., articles)
  - …
- Overlaps and gaps in features

PORTFOLIO

Partition Size and Stability

SCALE

~ 10 Fields for all of science
(e.g., Clinical Medicine)

~ 500 Disciplines for all of science
(e.g., Endoscopic Surgery)

~ 8,000 Specialties

~ 100,000 Problems

10^6

10^5

10^4

10^3

10^2

10^1

10^0

documents / partition / year

EVOLUTION

No morphological change

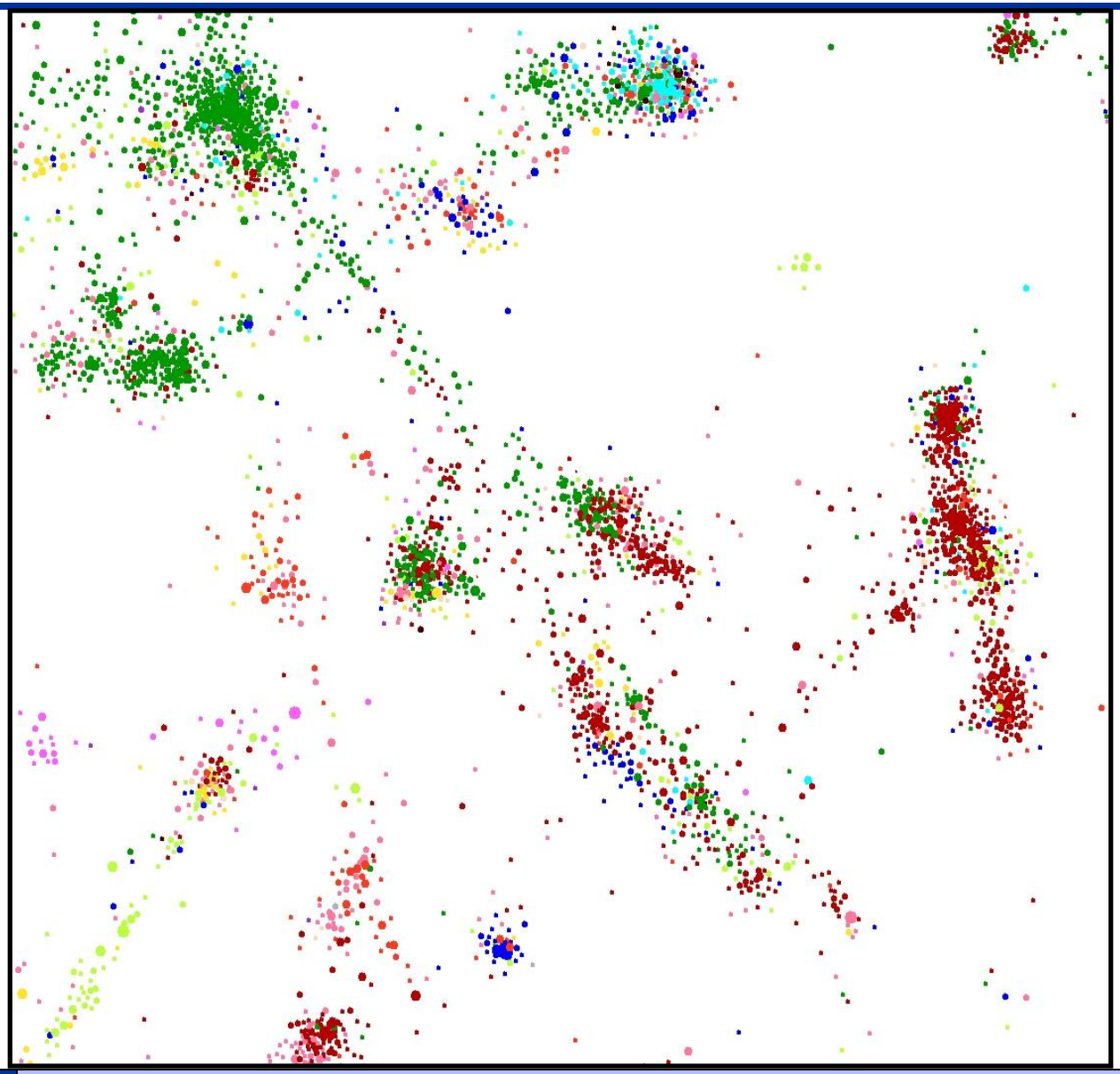
99+% Survival Rate/year

95% Survival Rate/year

56% Survival Rate/year

A Detailed Global Map of Science

Scopus data
2010
1.7M articles
116,000 topics



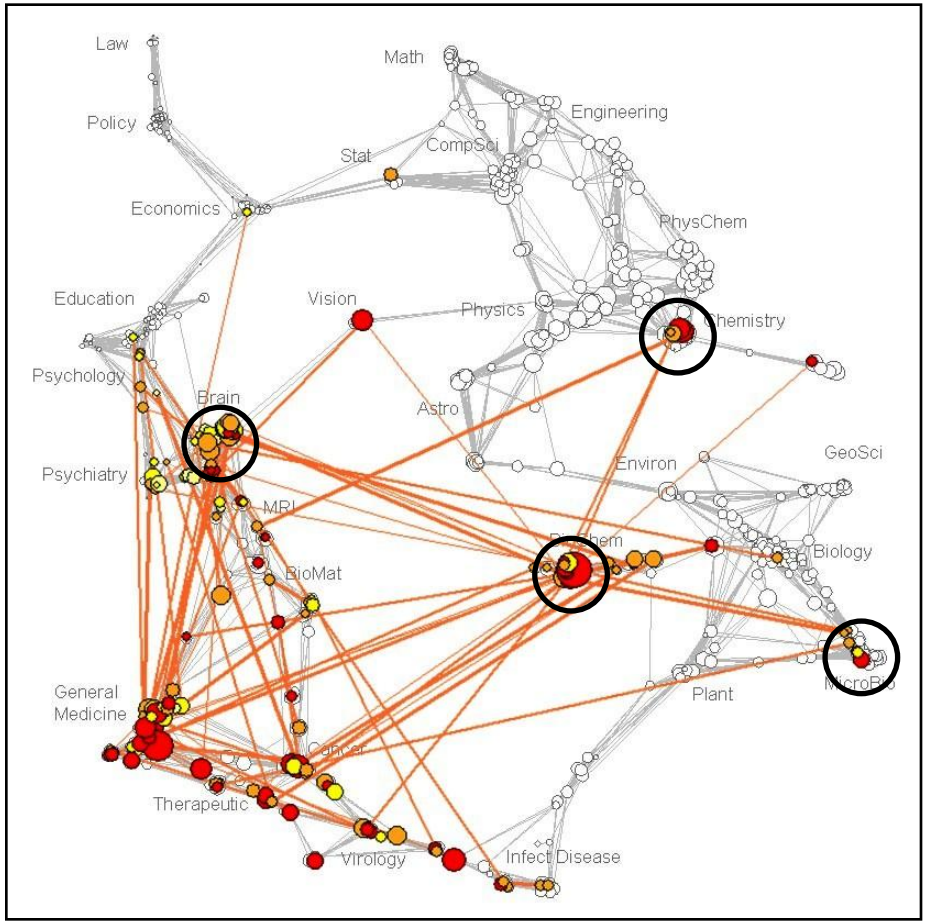
- Infectious Disease
- Medical Specialties
- Health Sciences
- Brain Research
- Humanities
- Social Sciences
- Computer Sciences
- Math / Physics
- Chemistry
- Engineering
- Earth Science
- Biology
- Biotechnology

Overlaps

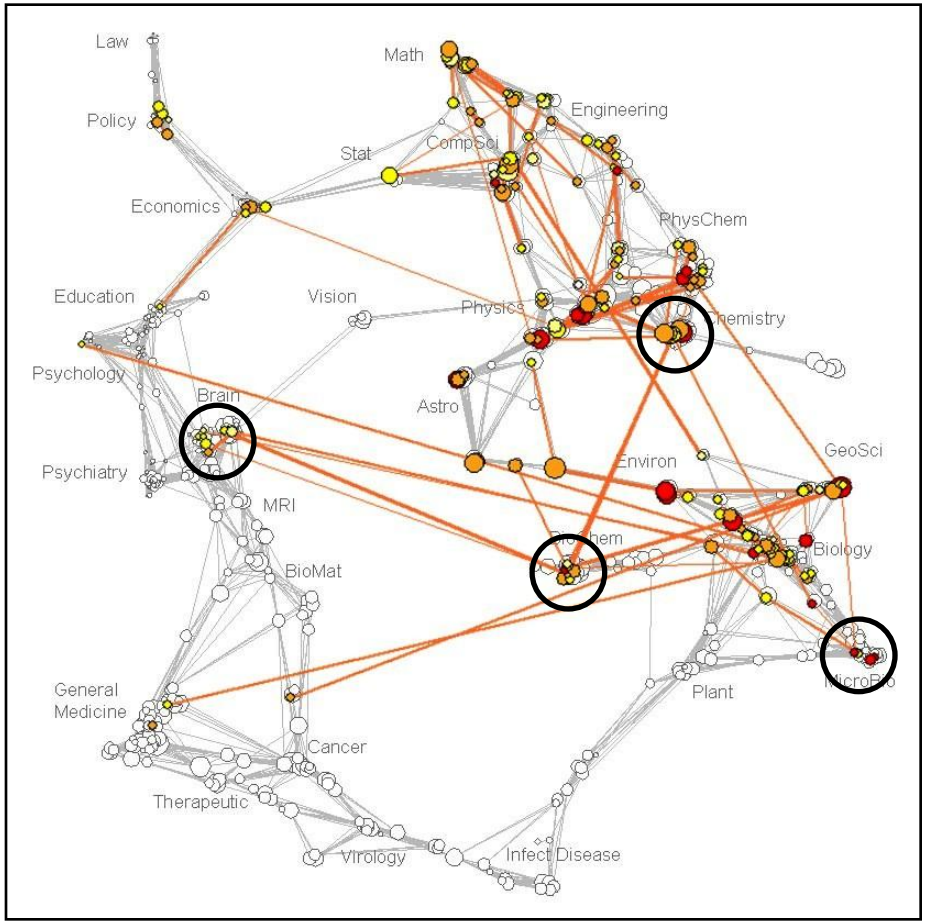
- Relatively simple to calculate or display
  - Can be shown on a single map or side-by-side
- Can occur on a variety of levels
  - Agency (e.g., NIH, NSF)
  - Institute (e.g., NCI, NLHBI)
  - Study sections
  - Programs

Overlaps – Agency

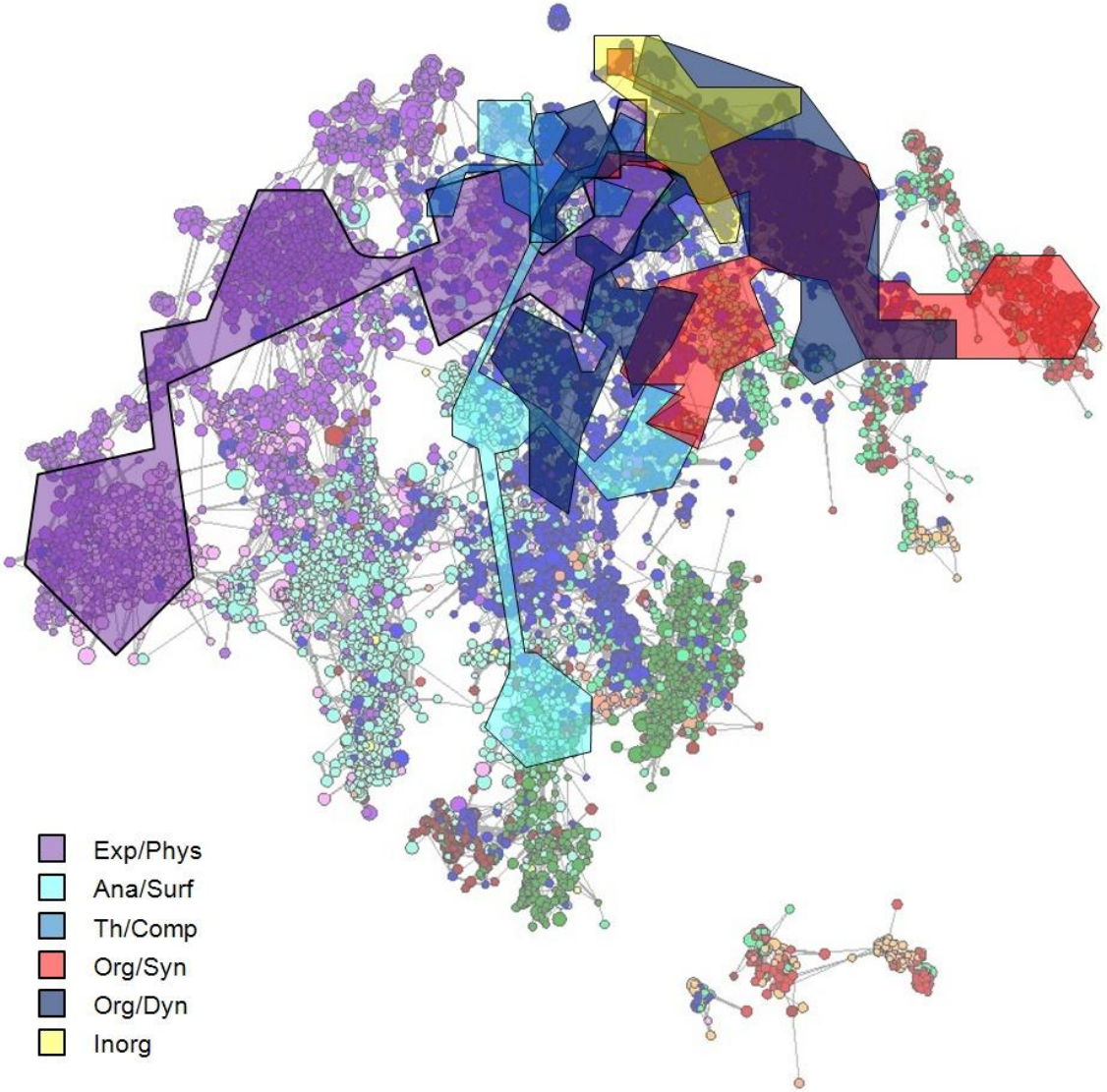
NIH



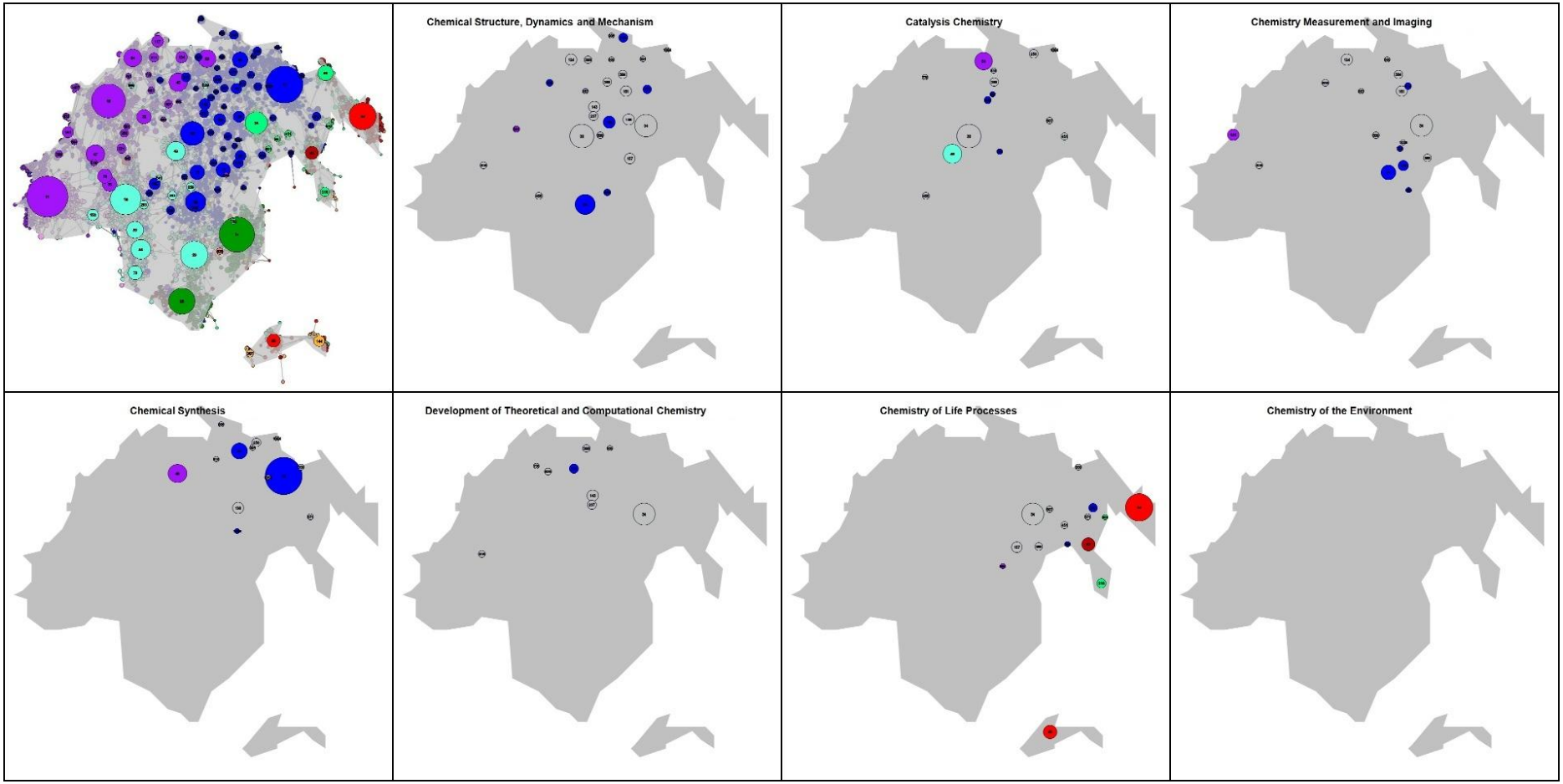
NSF



Overlaps – Program (Before)



Overlaps – Program (After)

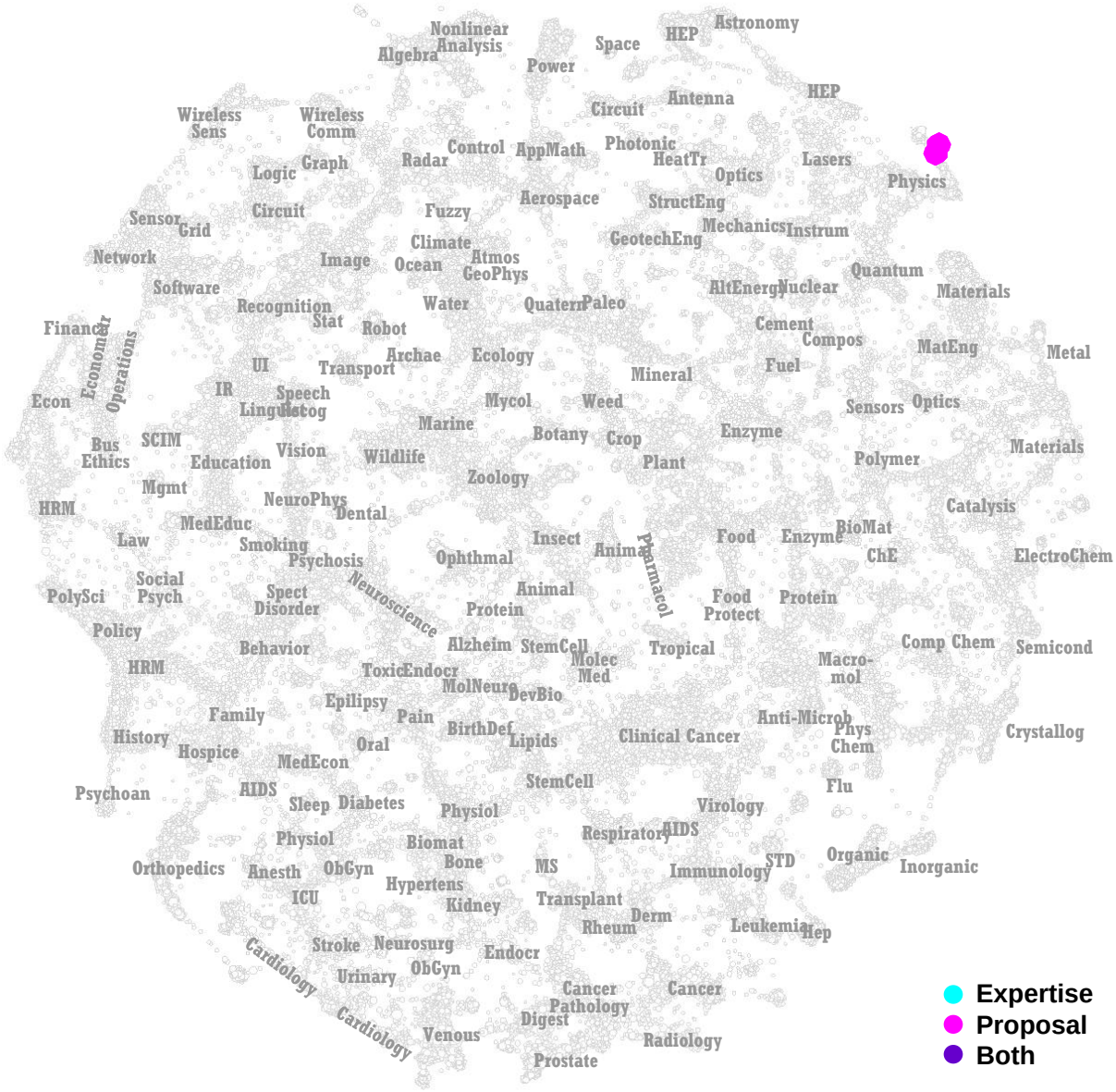


Gaps

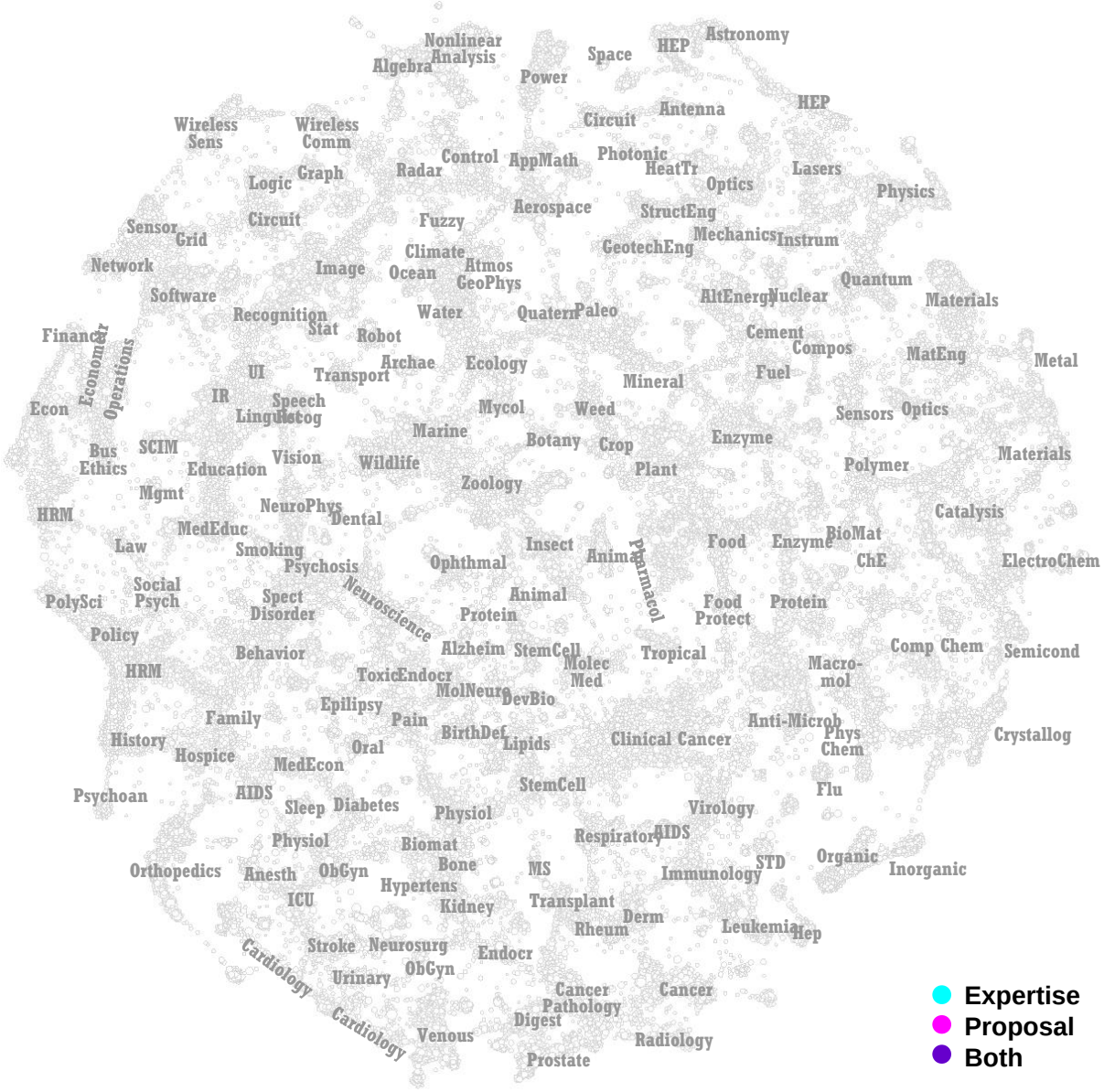
- Identifying gaps assumes a reference standard
  - The problem is that this standard often doesn't exist

- Two types of gaps
  - Counterpart to overlaps – these are "coverage gaps" or "mismatch gaps"
    - Relatively simple to calculate or display
    - Can be shown on a single map or side-by-side
  - Potentially transformative research
    - Much more difficult to identify than mismatch gaps

Overlaps and Mismatch Gaps



Overlaps and Mismatch Gaps



Gaps = Potentially Transformative Research

- **This is where the action is!**
  - Structural holes (Chaomei Chen)
  - Interdisciplinarity (Alan Porter)
  - Emergence (FUSE)
  - Transformative research (Agencies)
  - Hot science

Gap Analysis: Many Methodological Choices

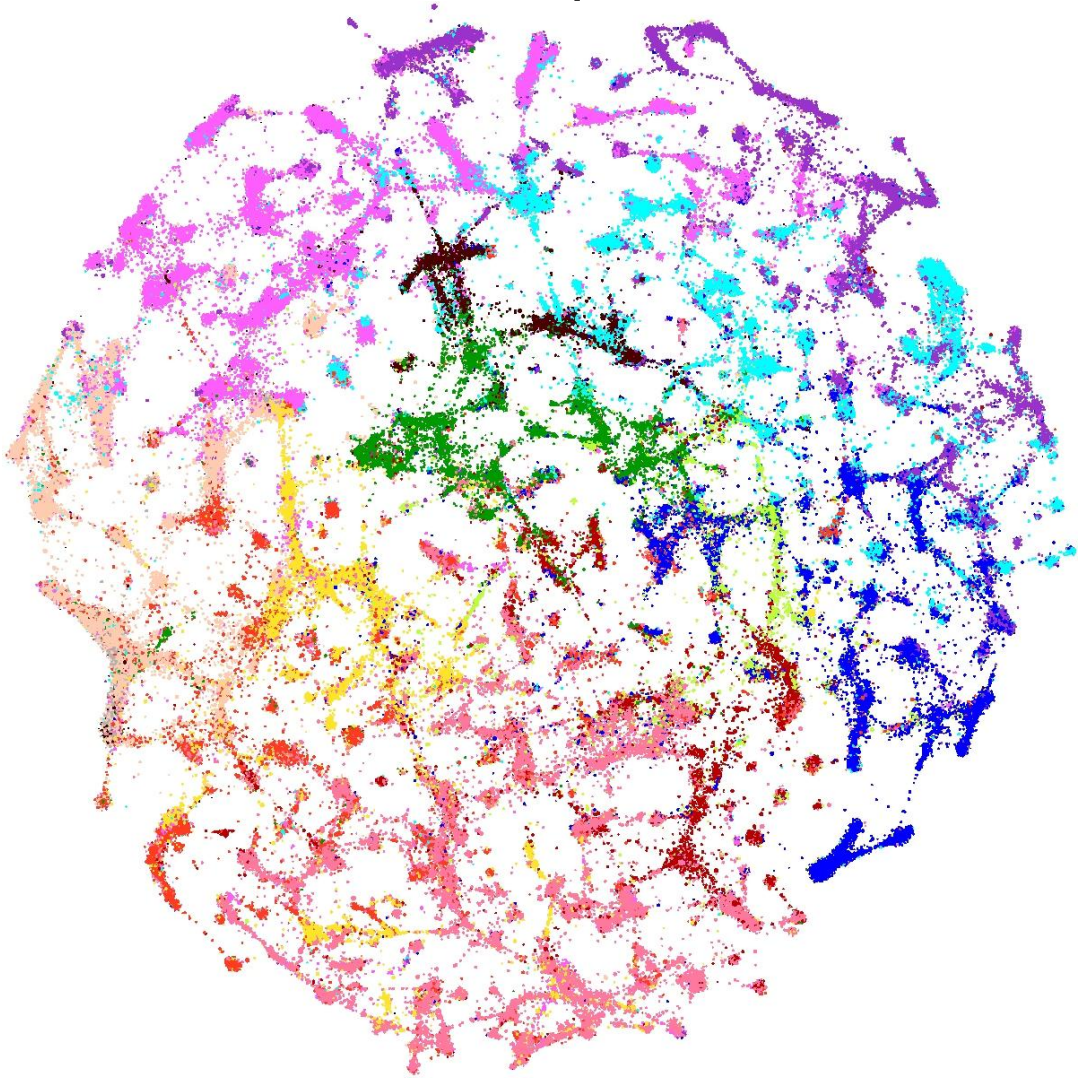
- We have much in common with the other presenters
  - » Common philosophies about science mapping and policy
    - Have co-authored with Katy & Chaomei and worked with Johan & Dewey
    - Correspond regularly with Alan & Loet
  - » Processes based on scientific method
  - » All interested in using more common terminology
- … and some differences
  - » Emphasis on accuracy
    - Global (all of science) rather than local (keyword defined sets)
    - Topic level rather than discipline level
    - Predictive modeling
    - We model a combination of stability and instability
- Details provided in the break-out session at 2:00 !

PTR – What Do We Look For?

- Should we evaluate white spaces?

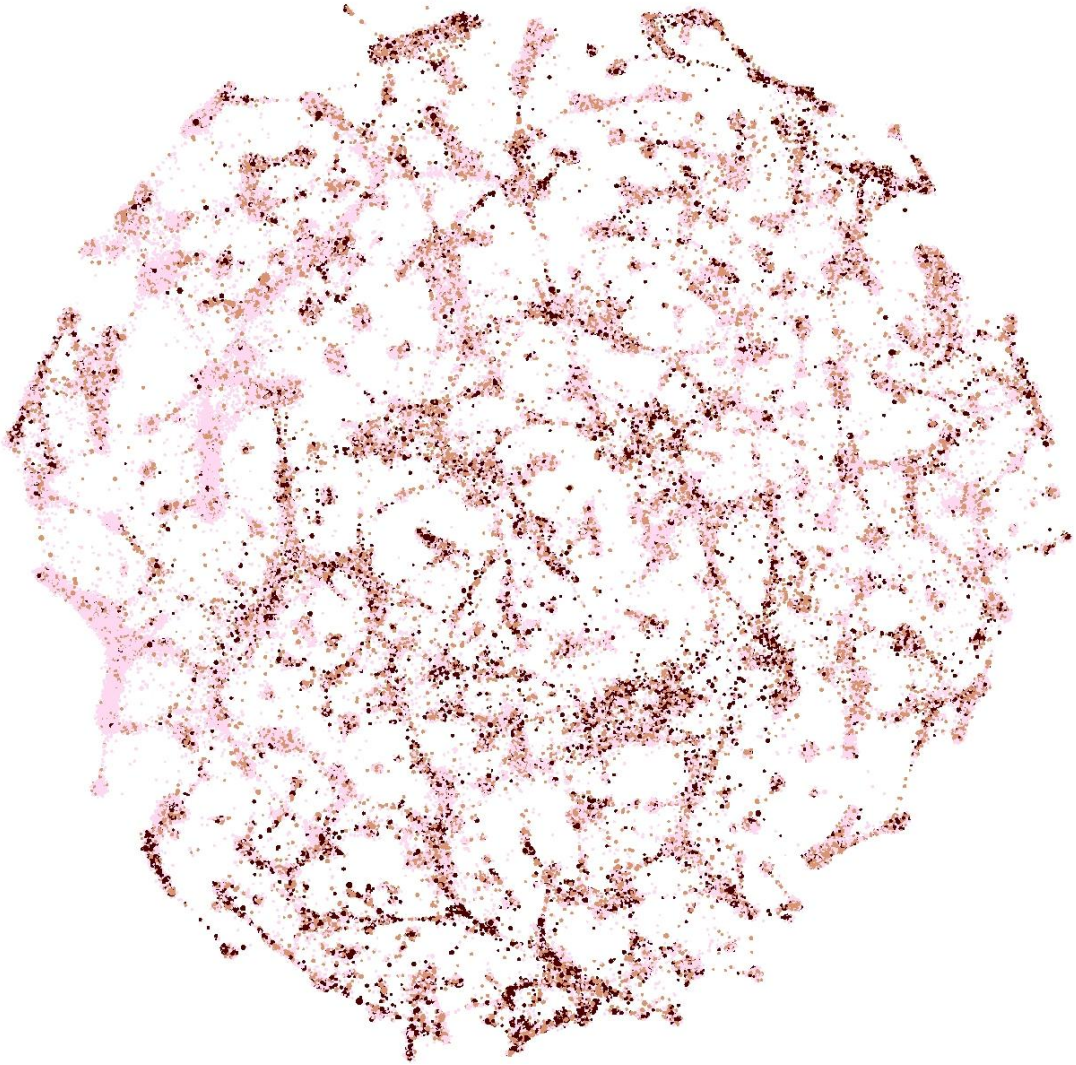
What Do We Look For?

- We add in ‘failures’ and ‘completed work’ from the past!



What Do We Look For?

- AND … identify existing topics with high novelty

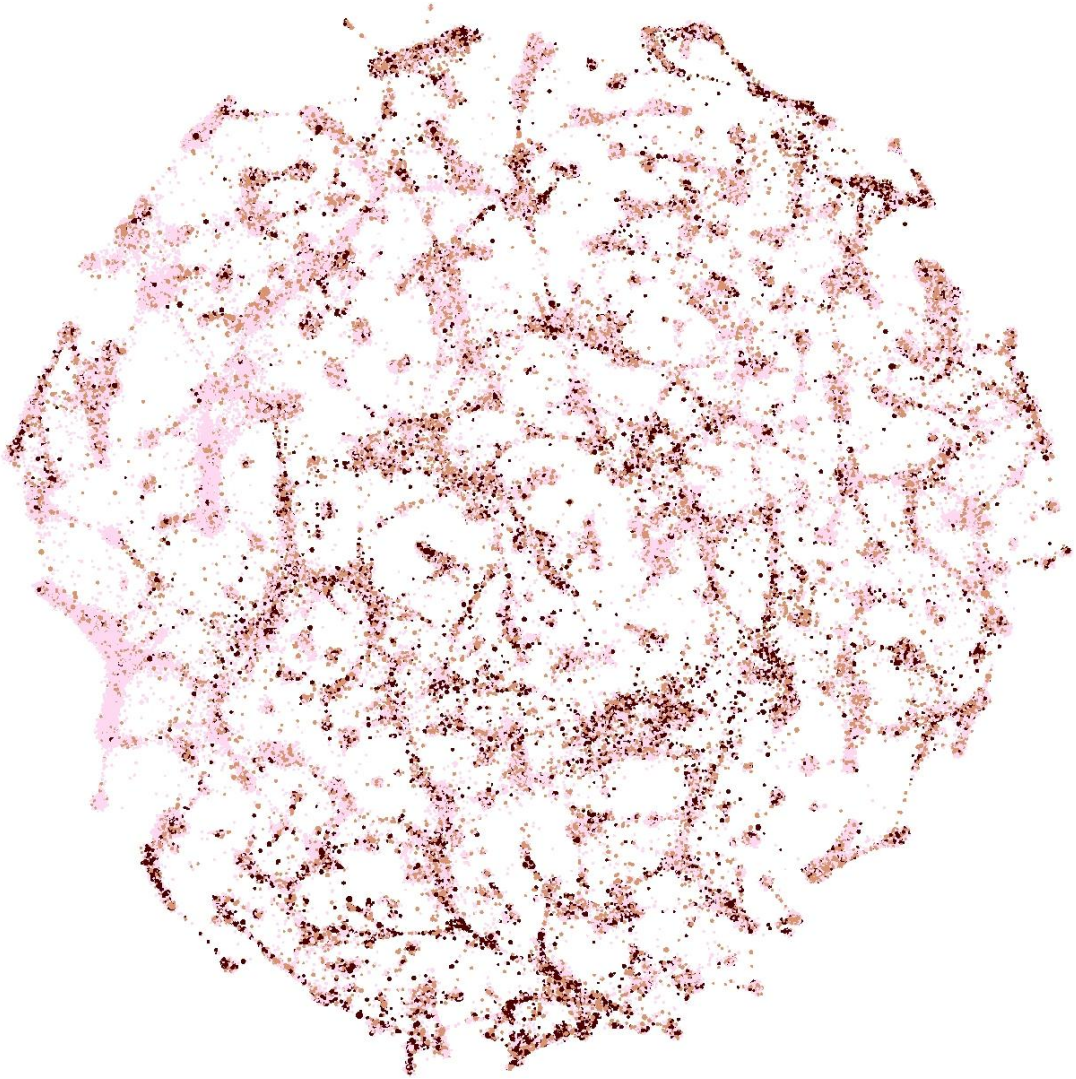


Not novel
Proven

Novel
Proven

Novel
Unproven

Top 20 PTR Opportunities

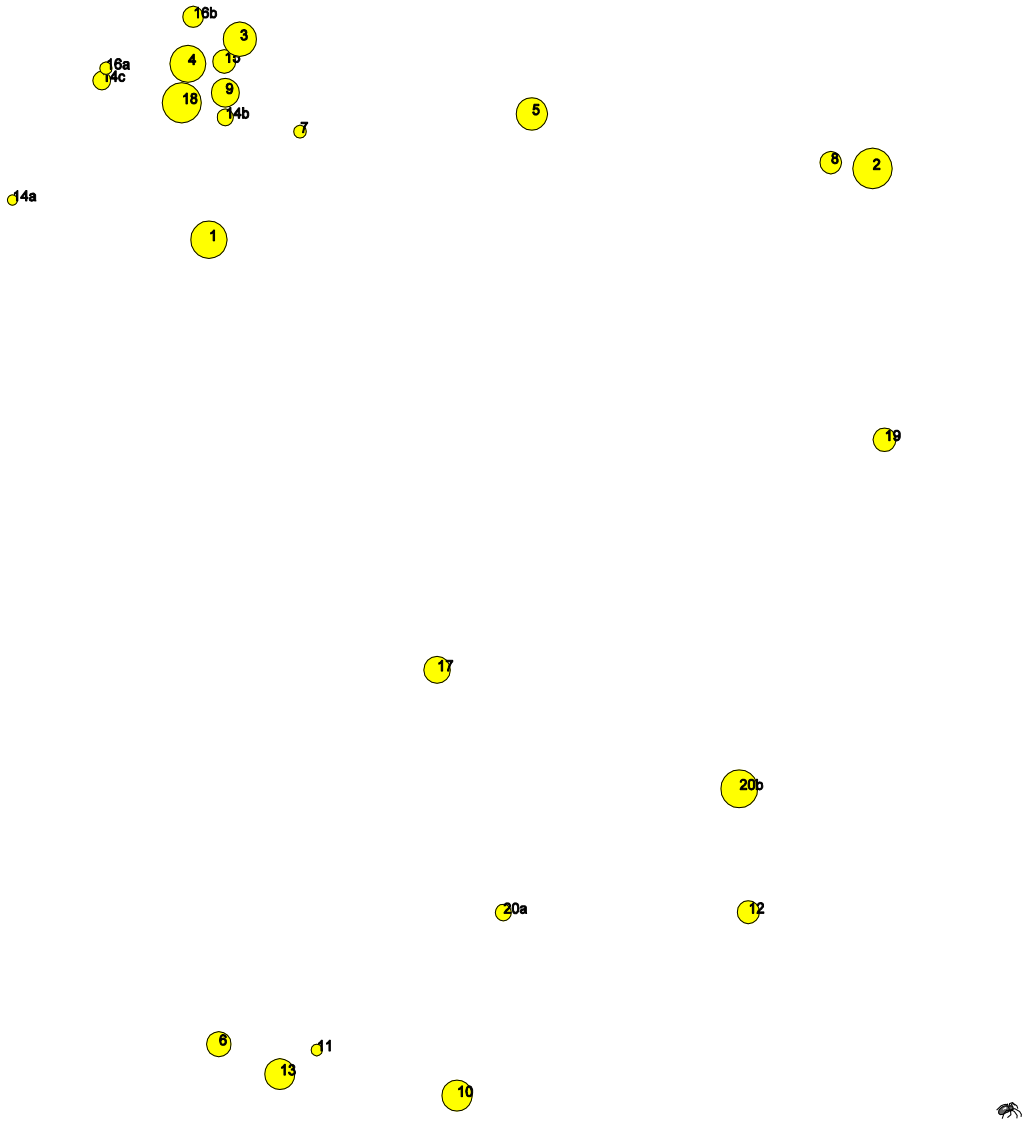


Not novel Proven

Novel Proven

Novel Unproven

Top 20 PTR Opportunities



Summary

- Science mapping can be used very effectively for portfolio analysis
  - Accuracy and detail are critical
  - Traditional overlaps and gaps (mismatches) are relatively simple to map
- Science mapping can also be used very effectively to identify potentially transformative research
  - The concept of novelty is key

Relevant Literature

- Börner, K., Klavans, R., Patek, M., Zoss, A. M., Biberstine, J. R., Light, R. P., Larivière, V., & Boyack, K. W. (2012). Design and update of a classification system: The UCSD map of science. ***PLoS One***, 7: e39464.
- Boyack, K. W., Newman, D., Duhon, R. J., Klavans, R., Patek, M., Biberstine, J. R., Schijvenaars, B., Skupin, A., Ma, N., & Börner, K. (2011). Clustering more than two million biomedical publications: Comparing the accuracies of nine text-based similarity approaches. ***PLoS One***, 6(3): e18029.
- Boyack, K. W., & Jordan, P. (2011). Metrics associated with NIH funding: A high-level view. ***Journal of the American Medical Informatics Association***, 18, 423-431.
- Klavans, R., & Boyack, K. W. (2011). Using global mapping to create more accurate document-level maps of research fields. ***Journal of the American Society for Information Science and Technology***, 62(1), 1-18.
- Boyack, K. W., & Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? ***Journal of the American Society for Information Science and Technology***, 61(12), 2389-2404.
- Klavans, R., & Boyack, K. W. (2010). Toward an objective, reliable, and accurate method for measuring research leadership. ***Scientometrics***, 82(3), 539-553.
- 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.
- Boyack, K. W. (2009). Using detailed maps of science to identify potential collaborations. ***Scientometrics***, 79(1), 27-44.
- Klavans, R., & Boyack, K. W. (2006). Quantitative evaluation of large maps of science. ***Scientometrics***, 68(3), 475-499.
- Klavans, R., & Boyack, K. W. (2006). Identifying a better measure of relatedness for mapping science. ***Journal of the American Society for Information Science and Technology***, 57(2), 251-263.
- Boyack, K. W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. ***Scientometrics***, 64(3), 351-374.
- Boyack, K. W. (2004). Mapping knowledge domains: Characterizing PNAS. ***Proceedings of the National Academy of Sciences***, 101, 5192-5199.
- Börner, K., Chen, C., & Boyack, K. W. (2003). Visualizing knowledge domains. ***Annual Review of Information Science and Technology***, 37, 179-255.

Acknowledgements

- Current funding
  - NSF EAGER IIS-1142795 †
  - IARPA FUSE ††
- Collaborators
  - Henry Small (SciTech)
  - Lyle Ungar (U. Penn)
  - Chaomei Chen (Drexel U.)
  - Katy Börner (Indiana U.)
  - Andrew McCallum (U. Mass)
  - Lance Ramshaw (BBN)
  - Nees Jan van Eck & Ludo Waltman (Leiden U.)

† This material is based upon work supported by the National Science Foundation under Grant No. IIS-1142795. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

†† This research is supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center (DoI/NBC) contract number D11PC20152. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/NBC, or the U.S. Government.

Thank you!