Using Network Analyses to Identify Emerging Research Areas

Susan Makar
susan.makar@nist.gov

Amanda Malanowski
amanda.malanowski@nist.gov

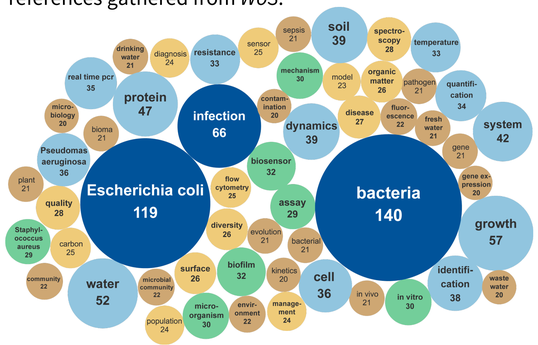
Amy Trost
amy.trost@nist.gov

Introduction

The Information Services Office (ISO) at the National Institute of Standards and Technology (NIST) is working to identify emerging research areas where NIST can have the greatest impact on industry and society. ISO staff have analyzed bibliometric data using a series of network analysis programs in order to reveal potential topics of interest in microbial metrology, a strategic priority area for NIST's Material Measurement Laboratory (MML).

Methodology

A *Web of Science* (*WoS*) database search identified microbial metrology papers for the years 2010-2016. The scope of this broad keyword search can be seen in the image below, which was created using CiteSpace and Tableau. ISO used these initial search results (1,459 papers) in Sci² and Gephi to create keyword and co-author network visualizations. To visualize in CiteSpace, the *WoS* search strategy was modified to analyze a smaller set of papers for the same time period that included the papers' cited references gathered from *WoS*.

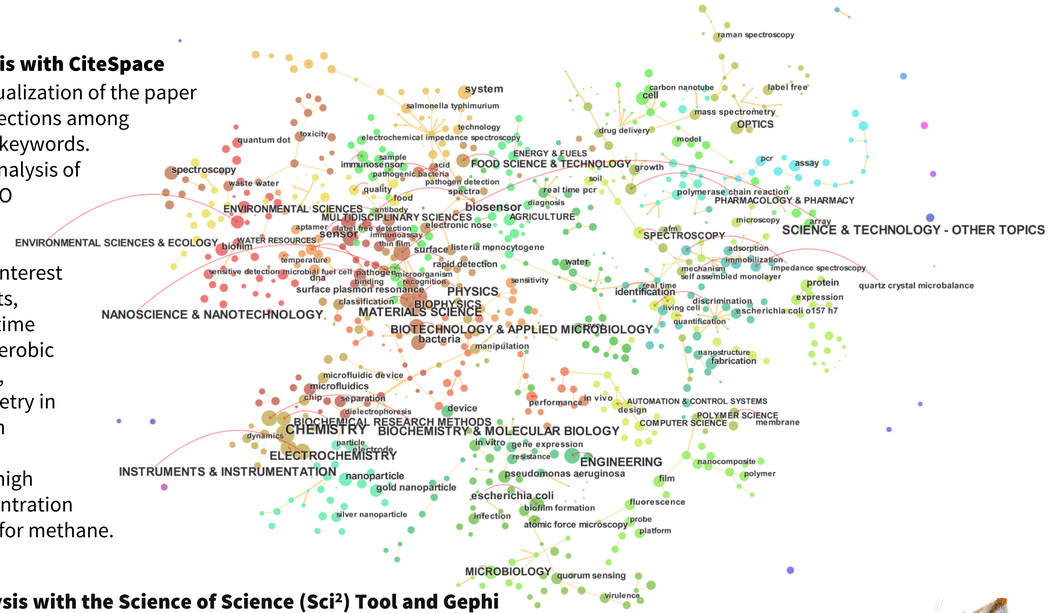


Most frequently occurring keywords in microbial metrology, 2010-2016

Findings

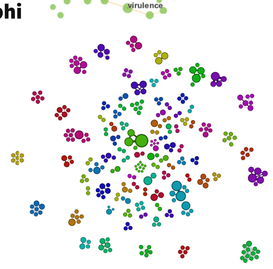
Topical Analysis with CiteSpace

A CiteSpace visualization of the paper set shows connections among disciplines and keywords. With a cluster analysis of this network, ISO was able to identify trends that may be of interest to MML scientists, including: real-time monitoring of aerobic biodegradation, spectrophotometry in food production facilities, and stable isotope/high precision concentration measurements for methane.

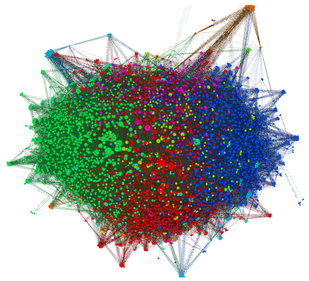


Network Analysis with the Science of Science (Sci²) Tool and Gephi

ISO extracted co-authorship and keyword networks from *WoS* (.isi) data using Sci² and visualized these networks in Gephi. Among microbial metrology papers, high levels of modularity (Q=0.97) in the co-authorship network indicate a lack of communication between sub-disciplines. Further investigation is needed to identify potential areas of collaboration among researchers. After normalizing the keywords in Sci², the resulting network mapped over 4,000 keywords and showed heavy connectivity. Further analysis of keywords while applying a time variable would be necessary to spot new/emerging research areas.



Co-Authorship Network



Keyword Network

Conclusions

Both Sci² and Gephi have well-developed user interfaces for evaluating co-authorship and co-citation networks. However, Sci² and Gephi do not have extensive keyword and semantic analysis capacity to adequately identify emerging research areas.

Multiple topical and cluster analyses could be performed with CiteSpace. However, its text mining algorithms may not yet be developed enough to provide detailed substantive guidance for NIST's scientists and research advisors. Additionally, several text analysis functions could not be consistently performed due to bugs in the software.

ISO learned that a successful bibliometric analysis has three challenging pieces:

- identifying a set of papers to examine,
- applying the right visualization tool(s) to answer the question, and
- interpreting the results in ways that inform the customer.

Interest in these tools and analyses will grow as text mining algorithms improve. In response, ISO is planning additional research and network analysis on a wide array of science mapping tools.

Related Work

Cobo, Manolo J., Antonio Gabriel López-Herrera, Enrique Herrera-Viedma, and Francisco Herrera. "Science mapping software tools: Review, analysis, and cooperative study among tools." *Journal of the American Society for Information Science and Technology* 62, no. 7 (2011): 1382-1402. doi: 10.1002/asi.21525.

Makar, Susan, Amanda Malanowski, and Katie Rapp. "Visualizing Forensic Publication Impacts and Collaborations: Presenting at a Scientific Venue Leads to Increased Collaborations between Researchers and Information Professionals." *Science & Technology Libraries* 35, no. 2 (2016): 109-118. doi: 10.1080/0194262X.2016.1184116.

NOTE: Identification of commercial entities in this poster is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology.