

Analyzing Trends in Science & Technology Innovation

Cody Dunne¹, Pengyi Zhang², Chen Huang², Jia Sun²,
Ben Shneiderman¹, Ping Wang² & Yan Qu²

¹Department of Computer Science & Human-Computer Interaction Lab

²College of Information Studies

Contact: {cdunne, ben}@cs.umd.edu, {pengyi, chuang, jsun, pwang, yanqu}@umd.edu

Web site: <http://stick.ischool.umd.edu>

This project is part of the larger NSF-supported grant on Science and Technology Innovation Concept Knowledge-base (STICK). Previous research in trends of information technology innovations have found helpful patterns by studying the frequency of mention of key terms in trade and scholarly articles in bibliographic databases. The current efforts on this project extend this work to even more diverse sources of information and a richer model of technology innovation that includes academic papers, patents, trade press articles, news, blog mentions, product adoptions, organizations, people, and more careful topic analysis.

Our work focuses on several aspects:

- Data collection & processing
- The development of ontologies to guide database schema design
- Providing data & visualization tools for analyzing these topics over time
- Building a community around these datasets and their analyses

STICK Project Goals

The Science & Technology Innovation Concept Knowledge-base (STICK) project's goal is to overcome the bias in the Science of Science and Innovation Policy (SciSIP) towards popular or ultimately successful innovations by providing the much needed data and tools for analyzing innovations of all possible outcomes. This comprehensive endeavor enables SciSIP researchers to build and test theories that explain the differentiated trajectories of science and technology innovations and their associated communities. The project also spans disciplinary boundaries by bridging the artificial divide in SciSIP research between the production and the use of innovations, piecing together a holistic view of the dynamic supply and demand in the innovation ecosystem. Specifically, the project builds a large-scale, multi-source, longitudinal database, Science & Technology Innovation

Concept Knowledge-base (STICK), and develops a set of visual analytic tools for monitoring and understanding the emergence and revolution/evolution of innovations in three exemplar science and technology fields: information technology, biotechnology, and nanotechnology.

The knowledge-base captures innovations, the individual and organizational actors associated with the innovations, and the relationships among the innovations and the actors through a hybrid approach that combines computational

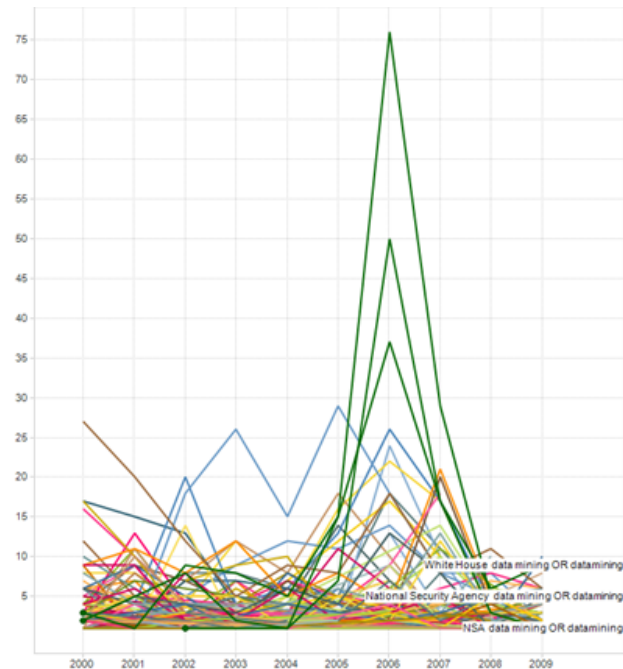


Figure 1: The frequency (vertical axis) of several terms relating to business intelligence (BI) co-occurring with each other. The chart shows 2000-2009 on the horizontal axis, with pairs of terms as lines. In 2006 there is a large spike for several pairs, highlighted in green. These correspond to data mining combined with the White House, NSA, and AT&T from the warrantless wiretapping scandal.

analysis of text (e.g., natural language processing) and social information processing (e.g., social tagging and collaborative writing by the users of the knowledge-base). State-of-the-art visualization tools are customized for SciSIP researchers and other innovation stakeholders to visualize innovation networks and analyze patterns and trends. The design of the knowledge-base and toolset is grounded in a demonstration study on the popularity of innovations. The study aims to address important questions concerning the complex relationships among innovations and the evolution of communities, with implications to the popularity and ultimate success of innovations.

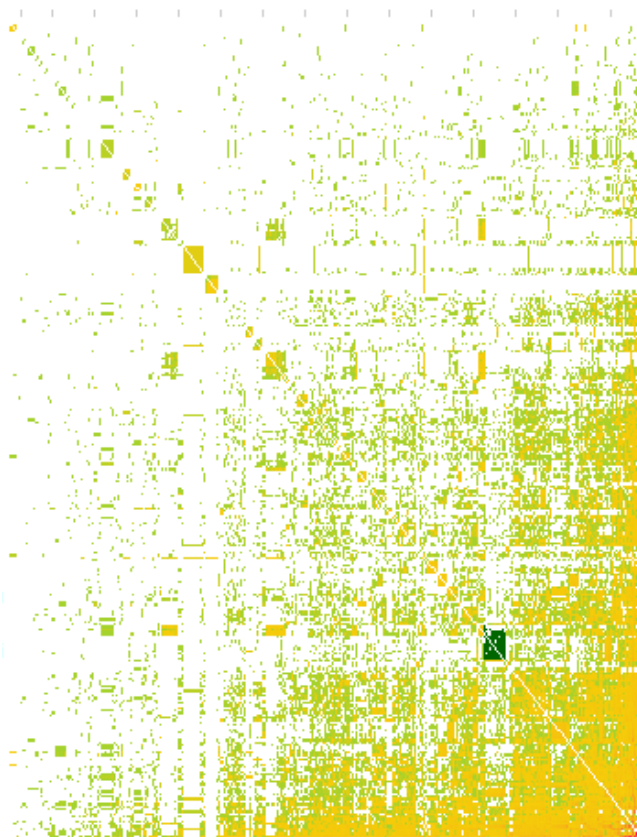


Figure 2: Clustered relationships between customer relationship management (CRM) terms. The highlighted cluster is the auto industry and armed forces collaborations. Following the diagonal to the top-left, we see three more big clusters. The first (going towards the top-left) is the U.S. military, Coast Guard, NASA, EPA, and DHHS. Next we have Native American associations from near Wyoming. Finally, several financial and pharmaceutical companies group together.

Acknowledgements: We appreciate the support of the National Science Foundation, Award 0915645 under the Science of Science and Innovation Policy program.

PAPERS

1. Zhang, P.Y., Qu, Y., and Huang, C. "Designing a Multi-Layered Ontology for the Science and Technology Innovation Concept Knowledge-Base," *Proc. 44th Hawai'i International Conference on System Sciences*, Kauai, HI, 2011. [DOI:10.1109/HICSS.2011.147](https://doi.org/10.1109/HICSS.2011.147)
2. Wang, P. "Chasing the Hottest IT: Effects of Information Technology Fashion on Organizations," *MIS Quarterly* (34:1), 2010, pp. 63-85
3. Tsui, C.-J., Wang, P., Fleischmann, K.R., Oard, D.W., and Sayeed, A.B. "Exploring the Relationships among ICTs: A Scalable Computational Approach Using K1 Divergence and Hierarchical Clustering," *Proc. 43rd Hawai'i Int'l Conference on System Sciences*, Kauai, HI, 2010. [DOI:10.1109/HICSS.2010.203](https://doi.org/10.1109/HICSS.2010.203)
4. Tsui, C.-J., Wang, P., Fleischmann, K. R., Sayeed, A. B., and Weinberg, A. "Building an IT Taxonomy with Co-Occurrence Analysis, Hierarchical Clustering, and Multidimensional Scaling," *iConference*, Champaign, IL., 2010. <http://www.ideals.illinois.edu/handle/2142/14918>
5. Shneiderman, B., Dunne, C., Sharma, P., and Wang, P., "Innovation Trajectories for Information Visualizations: Comparing Treemaps, Cone Trees, and Hyperbolic Trees", University of Maryland Human-Computer Interaction Lab Technical Report, 2010. <http://www.cs.umd.edu/localphp/hcil/tech-reports-search.php?number=2010-16>