Mapping Science Introduction: Past, Present and Future
by Katy Börner, Todd N. Theriault and Kevin W. Boyack

EDITOR’S SUMMARY
From early cartography to modern science maps, visual presentations facilitate understanding of large amounts of data. A traveling exhibit entitled Places and Spaces: Mapping Science has presented outstanding maps illustrating different designs and applications since 2005. The 10th year of the exhibit focuses on the future of science mapping and features five maps described in this special section of the Bulletin. Topics include the history of physics and key contributors, the development of the Internet and the structure of fields and topics in science and technology. Each emerged from latent relationships among elements in large volumes of data, made clear through visualization in an easily understandable format. Given high quality data, processing tools, design and analysis expertise and research funding, science mapping can be expected to expand in application and usefulness. Key challenges include insufficient numbers of experts, lack of sophisticated tools, low literacy in data visualization and absence of design standards.

KEYWORDS
data maps scientometrics
scientific and technical information trends
electronic visualization

For centuries, visual depictions of data and of the world have been used to understand and communicate information in novel ways. A prime example is cartographic maps of earth and water that have guided human exploration. They have marked the border between the known and the unknown, firing the imagination and fueling the desire for new knowledge and new explorations. Over time, geographic maps have become more detailed, more accurate and more sophisticated. Despite these advances, the thirst for exploration and discovery, along with the need for maps to guide our travels, remains undiminished.

Today, our opportunities for discovery reside less in physical places than in abstract information spaces that are ever growing, ever changing. Indeed, big data has become a buzzword that crosses all fields. But while search engines can retrieve facts from oceans of data, they cannot answer larger questions about those data as a whole: How big is this ocean? How can we navigate to the useful islands of knowledge and wisdom? How are knowledge and expertise interlinked on a global scale? In which areas is it worth investing time, effort, resources and compassion? Advances in science and technology are increasingly dependent on effective navigation and management of information spaces. But how do we make sense of all the data, information, knowledge and expertise that is relevant for our daily decision-making?

The field of information science is well situated to make contributions to all fields that are awash in data and information. Coping with this flood is where science maps and, more generally, visualizations can help. They serve as visual interfaces to immense collections of data, depicting myriad objects in ways that allow us to effectively discern apparent outliers, clusters and trends. This issue of the Bulletin of the Association for Information Science and Technology focuses on science mapping. It presents five short papers that
provide detail on recent maps from the *Places & Spaces: Mapping Science* exhibit. While the exhibit aims to introduce science mapping techniques to the general public and to experts across diverse disciplines for educational, scientific and practical purposes, this issue of the *Bulletin* aims to make the purposes, methods and lessons learned from science mapping more transparent to the information science community. Although detailed recipes for creating visualizations are not presented here, our hope is that the suite of examples given in the following papers will provide a sense of the types and scope of analyses that can be done using science mapping and visualization techniques. Perhaps these authors and their works will inspire students and practitioners in information science to make greater use of these tools [1].

*Places & Spaces: Mapping Science Exhibit – 100 Maps from 236 Mapmakers in 10 Years*

The *Places and Spaces: Mapping Science* exhibit began in 2005 with an initial set of 10 maps from 12 mapmakers (see Figure 1, left) and a 10-year goal and plan to curate 10 exemplary maps each year (through competitive selection), ultimately resulting in an exhibit with 100 maps. This goal has been reached; the 10th annual iteration of the exhibit is now complete and the exhibit features 100 maps from 236 mapmakers that render science and technology data into actionable insights (see exhibit display at the University of Miami in Figure 1, left, and thumbnail versions of all maps in Figure 2). Each iteration of the exhibit has had a different focus, and the resulting maps provide a very rich set of exemplars of the design and uses of science mapping and visualization. Although only five maps are discussed in the following papers, detail on all 100 maps is available at http://scimaps.org. Background on the processes used in many of these maps and visualizations are also described in [2, 3].

How Are Science Maps Used?

The last 10 years have seen a Cambrian explosion of science maps. While a timeline of major maps from the 1930s to 2007 has been compiled previously [2], it would be difficult if not impossible to map all the diverse developments across the sciences since 2008. The *Mapping Science* exhibit promotes map-making excellence, and the different themed iterations that introduce science maps for different stakeholders – from science policy makers, industry leaders and scholars to kids and the general public – have introduced the value and utility of topic maps to many.
The 10th iteration of the exhibit focuses on the future of science mapping. It showcases some of the best science maps in existence and reviews existing challenges and opportunities. The exhibit maps are large, high-resolution graphics – they are 24" x 30" (76 x 61 cm) and thus cannot be properly reproduced in a standard-size scholarly journal. All of the maps aim to help us understand and communicate complex S&T data. Many of the maps represent advances in data federation, mining and visualization approaches. Ultimately, each map tells a story. The best of these maps tell stories that can be understood by both experts and novices.

The 10 maps in the 10th iteration address a variety of topics and are used to tell stories about a variety of questions. One key feature of many maps is that they provide a broad context within which specific points can be addressed. Here are several examples:

- A map of physics (http://scimaps.org/mapdetail/being_a_map_of_physi_171) shows key points and people in its historical development from the 6th century to the present. The broad perspective not only enables a view of the entire field, but also allows key theoretical starting points, streams of thought and key figures to be seen within the context of the whole.

- Exploring a highly detailed map of the Internet can easily be consuming. Such a map (http://scimaps.org/mapdetail/map_of_the_internet_172) features old and new continents of companies, websites, services and social media. This map confirms relationships that we may already know about while also surprising us with unexpected connections. The best maps present familiar knowledge in a way that is consistent with our understanding and then build off that familiar knowledge to suggest things that we did not know. Indeed, hypothesis generation is one of the key uses of science maps.

- Maps of sets of documents, such as scientific papers or patents (see papers by Klavans & Boyack and map at http://scimaps.org/mapdetail/exploring_the_relati_180; Kay et al. and map at http://scimaps.org/mapdetail/mapping_graphene_sci_179) or of grants (http://scimaps.org/mapdetail/visual_funding_portf_178), are often used to show the structure of science and technology and the relationships between fields and topics. Large-scale maps of papers, patents or grants can then be used as basemaps or templates upon which other information can be overlaid. For example, the maps of Kay et al. (http://scimaps.org/mapdetail/mapping_graphene_sci_179) show that while the science of graphene is relatively focused in materials science, its application space as measured by patents is much broader.

Science maps often take advantage of geographic maps and geographic information systems to situate scientific data within the geographic contexts that are familiar to us all. For example, geographic maps can be used to show the potential spread of pandemic threats (http://scimaps.org/mapdetail/predict_healthmap_173) and demonstrate how different interventions can change the timing and breadth of that spread. They can also be used in concert with a flowmap metaphor to illustrate the impact of different variables on travel times in both current and historical contexts (see the paper by Meeks and map at http://scimaps.org/mapdetail/orbis_174).

Maps can also be used to uncover relationships between a multitude of different object types. For instance, maps can show the relationships between words and their usage (see the paper by Bertin et al. and map at http://scimaps.org/mapdetail/the_linguistic_conte_177) or between organizations of different types (see the paper by Oberg et al. and map at http://scimaps.org/mapdetail/interstitial_organz_181), thus revealing patterns that can help us to better understand our world. Maps may also work on multiple levels – from the micro (individual) level to the macro (population) level – revealing patterns and trends across and between multiple scales. An example is the money map by Randall Munroe (http://www.scimaps.org/mapdetail/money_176) that shows how economic resources are used and distributed at the one dollar to trillion dollar levels.

This issue includes maps by Meeks, Kay et al., Bertin et al., Klavans & Boyack and Korff. Ideally, a close examination of different maps will empower and inspire individuals to tell their very own stories about the enormous complexity, sheer beauty, practical utility and societal value of science and technology.
Opportunities and Challenges

The design of actionable science maps requires four critical ingredients: high-quality data; algorithms and tools to process this data; expertise to design valid workflows and to interpret results; and funding to pay for personnel, data storage, compute power and so forth. These four ingredients are discussed below.

In the last 10 years, the size of datasets, their coverage (for example, book data can now be used to map scholarly activities of arts and humanities scholars) and interlinkage (for example, http://linkeddata.org interconnects many previously isolated data silos, while the International Researcher Network at http://nrm.cns.iu.edu allows anyone to browse publication, teaching and funding data from institutions around the globe) have increased enormously. Thanks to the open access movement, many high quality datasets are becoming available online free of charge (such as MEDLINE publications at http://nlm.nih.gov). Increasingly, data covering other areas of scholarship and creative activity are becoming available, including datasets of genes, proteins, diseases, films and film characters, music, photos and social media data.

The number, sophistication and scalability of data processing, mining, modeling and visualization algorithms have improved substantially as well over the past decade. Many different open-source libraries, tools and services exist to perform temporal, geospatial, topical and network studies. Several of the tools come with extensive online documentation on how to run expert-validated workflows (that is, sets of algorithms that are executed in a well-defined sequence with specific parameter values) so that users can visualize their own data using advanced workflows. While many government agencies, researchers and practitioners are not able (or willing) to share their data, they can now apply the very same workflows and compare results across institutional and disciplinary boundaries.

The number of experts able to advance data mining and visualization research and development has grown enormously but not as fast as the need to render data into actionable insights. The McKinsey Global Institute (2011) forecasts a 50-60% gap between the supply and demand of people with deep analytical talent and projects. They project that by 2018 the United States may experience “a potential shortfall of 1.5 million data-savvy managers and analysts” [4]. Domain expertise is needed to validate and interpret S&T maps and to tell the stories that make their content relevant for many. The design of truly actionable science maps frequently requires a close collaboration between data analysts and domain experts.

Funding and attention for the science of science research has increased enormously over the last 10 years. The National Science Foundation’s Science of Science and Innovation Policy (SciSIP) program has co-funded 108 projects since 2009 alone, with a total award amount of about $58 million dollars. The recent 1 billion euro FutureICT flagship project proposal brought together hundreds of the best scientists from Europe and elsewhere to design a data-simulation-visualization platform that can help accelerate science, technology and innovation (http://www.futureict.eu). In a time of tightened budgets, federal and personal funds need to be spent more wisely, and accountability for the impact of investments becomes more important. We note, however, that visualizations can and should be a part of projects funded across all of science and not just by projects aligned with policymaking.

While access to the four critical ingredients is getting easier, there are a number of serious challenges that need to be addressed to make science maps truly useful for different stakeholders. First, insight needs and datasets are growing faster than both the number of experts and the scalability and sophistication of algorithms and tools. Varied algorithms exist for the study of heterogeneous networks (networks with multiple node types and link types such as authors and papers with author-author, paper-paper and author-paper links). However, fewer approaches exist to study problems that are based in multi-level networks (for example, the impact of a certain nationwide policy decision on the career trajectories of individual researchers).

Second, data visualization literacy, the ability to read and make visualizations, is rather low. A recent study of 900 youth and adult visitors across five science museums in the United States revealed that many people cannot interpret basic data visualizations, such as scatter plots or geospatial maps with data overlays. Asked to read more advanced visualizations such as network layouts, many had no understanding of or vocabulary to refer to...
key features such as high-degree nodes or network clusters. Additional studies showed that the composition tasks – asking visitors to compose visualizations from multiple data layers – led to a higher level of understanding than asking them to decompose visualizations. This finding is one of several reasons why Phase II of *Mapping Science* (coming soon) will focus on interactive visualizations and invite submission of tools that empower users to render data into insights. That is, future iterations of the exhibit will not only invite a general audience to view and read S&T maps but also to make S&T maps. It is our hope that the ability to see the impact of different data cleaning, analysis and visualization algorithms will help to increase the data visualization literacy of users. The maps and tools will be made available in formal education in schools and informal encounters in (science) museums, libraries and other public spaces.

Third, few guidelines and even fewer standards exist for the design of science maps (but also for science metrics or science classifications). Just like early maps of the world did not place north on top, used different color schemas and did not show latitude and longitude information, today’s maps of science come in all shapes and forms. However, many science maps seem to have the very same general structure – a comparison of 20 maps of science generated from different datasets manually or using diverse algorithms showed a very high level of correspondence [5]. The UCSD science map and classification system [6] is now used across 10 different tools and online services, and the color schema for the different scientific disciplines has been adopted for several other maps. Work on multi-level classification systems and a map of science that also covers book and proceedings data is in progress. Ultimately, it seems highly desirable to have free datasets, tools and workflows in place that anyone can plug-and-play to replicate and advance the generation of science base maps and data overlays. Different stakeholders and application domains might require different base maps. However, concordances (also called alignments or mappings) between classification system and maps will make it possible to visualize the very same dataset on different maps, enabling comparisons and human subject studies that can increase our understanding of what visualization to use when.

Additional opportunities and challenges as well as future works are discussed in the papers presented in this special section of the *Bulletin*.

---

**Resources Mentioned in the Article**


