

This is the table of contents of
Katy Börner, Chaomei Chen, & Kevin Boyack: Visualizing Knowledge Domains. Annual Review of Information Science & Technology, Volume 37, 2003. (in press)

Visualizing Knowledge Domains¹

Katy Börner

School of Library and Information Science, Indiana University, Bloomington, IN 47405, USA

katy@indiana.edu

Chaomei Chen

College of Information Science and Technology, Drexel University, Philadelphia, PA 19104, USA

chaomei.chen@cis.drexel.edu

Kevin W. Boyack

Sandia National Laboratories, Albuquerque, NM 87185, USA

kboyack@sandia.gov

"The purpose of computing is insight – not numbers."

R. W. Hamming (1962)

1	INTRODUCTION.....	2
2	HISTORY.....	4
2.1	SCIENTOMETRICS, BIBLIOMETRICS, AND CITATION ANALYSIS.....	4
2.2	MAP GENERATION AND VISUALIZATION.....	6
3	PROCESS FLOW OF VISUALIZING KNOWLEDGE DOMAINS	8
3.1	UNITS OF ANALYSIS.....	9
3.2	MEASURES AND SIMILARITY CALCULATION.....	11
3.2.1	<i>Measures</i>	11
3.2.2	<i>Simple Similarities</i>	11
3.2.3	<i>Vector Space Model</i>	12
4	ENABLING TECHNOLOGIES	13
4.1	DIMENSIONALITY REDUCTION TECHNIQUES.....	14
4.1.1	<i>Eigenvalue/Eigenvector Decomposition</i>	14
4.1.2	<i>Factor Analysis and Principal Components Analysis</i>	15
4.1.3	<i>Multidimensional Scaling</i>	15
4.1.4	<i>Latent Semantic Analysis</i>	16
4.1.5	<i>Pathfinder Network Scaling</i>	17
4.1.6	<i>Self-Organizing Maps</i>	19
4.2	CLUSTER ANALYSIS.....	20
4.3	SPATIAL CONFIGURATION.....	21
4.3.1	<i>Triangulation</i>	21
4.3.2	<i>Force Directed Placement</i>	22
4.4	VISUALIZATION AND INTERACTION DESIGN	22
4.4.1	<i>Visualization</i>	23
4.4.2	<i>Interaction Design</i>	24
4.4.3	<i>Focus+Context</i>	24
4.5	DISCUSSION.....	26

¹ All figures in this chapter will be available in color at <http://ella.slis.indiana.edu/~katy/arist02>.

5	THE ARIST DATA SET	26
5.1	DATA RETRIEVAL.....	26
5.2	COVERAGE	27
6	THE STRUCTURE OF THE SUBJECT DOMAIN	30
6.1	MULTIPLE MAPS OF THE DOMAIN.....	30
6.1.1	<i>ARIST-GSA/StarWalker</i>	31
6.1.2	<i>ARIST-ET-Map</i>	36
6.1.3	<i>ARIST-Cartographic-SOM Maps</i>	37
6.1.4	<i>ARIST-VxInsight</i>	39
6.2	COMPARISON OF MAPS.....	42
7	PROMISING AVENUES OF RESEARCH.....	47
8	CONCLUSIONS	50
9	ACKNOWLEDGEMENTS	50
10	BIBLIOGRAPHY	51

ABSTRACT

This chapter reviews visualization techniques that can not only be utilized to map the ever-growing domain structure of scientific disciplines but that also support information retrieval and classification. In contrast to the comprehensive surveys done in a traditional way by Howard White and Katherine McCain (1997; 1998), the current survey not only reviews emerging techniques in interactive data analysis and information visualization, but also visualizes bibliographical structures of the field as an integral part of our methodology. The chapter starts with a review of the history of knowledge domain visualizations. We then introduce a general process flow for the visualization of knowledge domains and explain commonly used techniques. In the interest of visualizing the domain this article reviews, we introduce a bibliographic data set of considerable size, which includes articles from the citation analysis, bibliometrics, semantics, and visualization literatures. Using a tutorial style, we then apply various algorithms to demonstrate the visualization effects produced by different approaches and compare the different visualization results. At the same time, the domain visualizations reveal the relationships within and between the four fields that together form the topic of this chapter, domain visualization. We conclude with a discussion of promising new avenues of research and a general discussion.

1 INTRODUCTION

Painting a big picture of scientific knowledge has always been desirable for various reasons. Traditional approaches are brute-force in nature – scholars have to sort through through the mountains of literature to conduct their surveys. Obviously, this is time-consuming, difficult to repeat, and subjective. The task is enormous in its complexity. Sifting through recently published documents to find ones that will later be recognized as important is labor-intensive. Traditional approaches are increasingly hard to keep up with the pace of information growth. When it comes to a multidisciplinary field of study, it is rather difficult to maintain an overview of what is going on. Painting the “big picture” of an ever-evolving scientific discipline has been akin to the situation described in a widely known Indian legend about blind men and an elephant. As the