

The Shift from Information Retrieval to Synthesis

Theme: The Grand Challenges

Catherine Blake

cablake@email.unc.edu

School of Information and Library Science
University of North Carolina, Chapel Hill

Caryn Anderson

caryn.anderson@simmons.edu

Graduate School of Library and Information Science
Simmons College, Boston

Grand challenges such as public health, security, genomics, environmental protection, education, and economics, are characterized by complexity, interdependence, globalization, and unpredictability. Although the unprecedented quantity of information surrounding these challenges can provide users with a new perspective on solutions, the data surrounding complex systems vary with respect to levels of structure and authority, and include vastly different contexts and vocabularies. To be successful in this domain we must extend our models of information science such that they operate successfully in environments where the quantity of relevant information far exceeds our human processing capacity. For example, the well-accepted precision and recall metrics break down when hundreds of thousands of documents are relevant. Solutions to grand challenges require that information scientists shift their focus from information retrieval towards information synthesis.

Systems that synthesize information support serendipity and experimentation, and enable the generation of patterns that no individual has previously considered. In keeping with the socio-technical core of information science, models of synthesis provide a mixed-initiative approach and “solve tasks cooperatively with a domain expert” [1]. Such systems build on theories and methodologies from the human-computer interaction, computer supported co-operative work, and information visualization communities. Information synthesis systems generate knowledge for human problem-solvers who may not yet recognize their information need.

Information scientists have developed several synthesis models. Swanson’s ABC model identifies connections between bibliographically disjointed literature [2] and has uncovered relationships resulting in medical treatments for Raynaud’s [3] and Alzheimer’s [5] diseases, migraines [4], and hepatitis C [6]. Small’s reflections on paradigms, citations, and maps of science capture the role that bibliometric analyses play in defining “normal science” [7] and in identifying new research fronts. Shneiderman combines information visualization and data mining technologies [8], and Blake extends meta-analytic techniques to provide both a quantitative and visual summary of information that was not the main point of a scientific article [9]. The lens provided by each of these information scientists, enables users to re-conceptualize existing information and to validate their new hypotheses.

Mao Tse-Tung said that *"We think too small. Like the frog at the bottom of the well. He thinks the sky is only as big as the top of the well. If he surfaced, he would have an entirely different view."* Schools of information science bring the interdisciplinary perspective that is required to engage in grand challenges. Information scientists revel in new ways to combine the power of people, technology and information. Their continued emphasis on synthesis will ensure the discovery of solutions to grand challenges.

1. Simon, H.A., R.E. Valdés-Pérez, and D.H. Sleeman, *Scientific Discovery and Simplicity of Method*. Artificial Intelligence, 1997. 91(2): p. 177-81.
2. Swanson, D.R. and N.R. Smalheiser, *An Interactive System for Finding Complementary Literatures: a Stimulus to Scientific Discovery*. Artificial Intelligence, 1997. 91: p. 183-203.
3. Swanson, D.R., *Fish-oil, Raynaud's Syndrome, and undiscovered public knowledge*. Perspectives in Biology and Medicine, 1986. 30(1): p. 7-18.
4. Swanson, D.R., *Migraine and magnesium: eleven neglected connections*. Perspectives in Biology and Medicine, 1988. 31: p. 526-557.
5. Smalheiser, N. and D. Swanson, *Linking Estrogen to Alzheimer's disease: an Informatics Approach*. Neurology, 1996. 47(3): p. 809-10.
6. Weeber, M., et al., *Generating Hypotheses by Discovering Implicit Associations in the Literature: A Case Report of a Search for New Potential Therapeutic Uses for Thalidomide*. Journal of the American Medical Association, 2003. 10(3): p. 252-9.
7. Small, H., *Paradigms, Citations, and Maps of Science: A Personal History*. Journal of the American Society for Information Science, 2003. 54(5): p. 394-9.
8. Shneiderman, B., *Inventing Discovery Tools: Combining Information Visualization with Data Mining*. Information Visualization, 2002. 1(1): p. 5 - 12.
9. Blake, C., *Information Synthesis: A Mixed-Initiative Meta-Analytic Approach to Facilitate Knowledge Discovery from Scientific Text*, in *School of Information and Computer Science*. 2003, University of California: Irvine.