

soasis&t

Workshop Series

**The Southern Ohio Chapter of the American Society for
Information Science & Technology &
The LexisNexis Technical Library Sponsor**

***Aboutness: Automated Indexing &
Categorization***

June 21, 2001, Miamisburg, Ohio

9:00 – 9:20 Registration & Continental Breakfast

Introductions by Patricia Carter

Patricia Carter, SOASIST Chair-Elect and 2000 Chapter Member-of-the-Year, has been the Technical Librarian at LexisNexis for almost nine years. Prior to LexisNexis, she worked at the IBM Austin, TX Technical Library for one year and the previous ten years at the University of Texas at Austin's Engineering Library, Chemistry Library, Life Sciences Library, Fine Arts Library, and managed the Search/Ordering section of the Acquisitions department in the Main Library. She holds both a Masters in Library Science and a Bachelors of Fine Arts in Art History from the University of Texas at Austin.

9:20 – 10:20 Rich Miller, LexisNexis, "Visual Thinking"

An introduction to a collection of concepts and methods for perceiving and communicating visually. These concepts and methods will be used to stimulate thought and communication about the main topics of the day: automated indexing and categorization.

Rich Miller is a Research Scientist in the Alliances and New Technology group at LexisNexis, specializing in areas such as streaming media, data visualization, and user interface design. Rich started at LexisNexis in 1990 and served as a User Interface Designer before joining Structural Dynamics Research Corporation in Cincinnati as Manager of a User Interaction group, returning to LexisNexis in 2000. Rich earned his Ph.D. in Experimental Psychology from Miami University.

**10:20 – 11:20 Jean Godby, OCLC Office of Research, "Terminology
Identification from Full Text: OCLC's WordSmith Experience"**

A discussion of philosophical and practical issues involved in extracting topical terminology from full-text documents.

Carol Jean Godby is a Senior Research Scientist in the Office of Research at OCLC and manager of the Automatic Classification project. Since starting at OCLC's Office of Research in 1990, she has contributed to a variety of projects, including an early large-scale prototype of a digital library, numerous information retrieval and automatic indexing experiments, the Dublin Core metadata standard, and electronic adaptations of the Dewey Decimal Classification. She managed the WordSmith project, whose goal was to incorporate tractable natural language processing solutions into OCLC's database products. She is also finishing a Ph.D. in computational linguistics at The Ohio State University. When she is avoiding work on her dissertation, she is volunteering for the Rails to Trails Conservancy or tinkering with her house in a historic neighborhood in Columbus.

11:20 – 11:30 Break

11:30 – 12:30 Mark Shewhart, LexisNexis, “Portable Classification Tools”

With the advent and popularity of Intranet portals, customers often want to simultaneously and consistently search internal proprietary materials, the Internet, and vendor-provided materials. One approach to this customer demand is to find ways to apply indexing technologies to the Internet and Customer data that will allow a single search from a topic hierarchy. In addition to issues of document classification, we will discuss the powerful tools that are enabled by a consistent controlled classification scheme across all content.

Mark Shewhart is a Consulting Research Scientist at LexisNexis, working on advanced data enhancement systems since 1993. Prior to LexisNexis, Mark worked in the Artificial Intelligence Program Management Office at Wright Patterson Air Force Base also in Dayton, Ohio. He has a B.S. in Computer Engineering and a M.S. in Applied Mathematics from the University of Illinois at Urbana-Champaign and a M.S. in Computer Science focused on artificial intelligence and neural networks from Wright State University in Dayton. Mark lives in Bellbrook, Ohio with his wife Pat and their five children, ages 4 to 11.

12:30 – 1:00 Buffet Lunch

1:00 – 2:00 Tim Humphrey, LexisNexis, “Machine Learning Classification”

An introduction with exercises that shows how computers can learn how to classify documents by analyzing example documents that have already been classified.

Tim Humphrey is a Consulting Research Scientist at LexisNexis where he has worked for 16 years. He currently is investigating machine learning technologies for document classification. His BS is in mathematics and did both his Masters and Ph.D. work is in Computer Science. In addition, last year he founded the organization, Elders Connected, for the purpose of teaching senior citizens how to use email and the Internet.

**2:00 – 3:30 Jill Sellers, Leslie Denton & Patricia Neff, LexisNexis,
“Aboutness: the Human Factor”**

A hands-on activity during which workshop participants will break into small groups, review recent news stories, determine the subjects covered, and assign controlled vocabulary terms from a list provided. Discussion following the exercise will explore the process of labeling concepts and what counts as evidence of “aboutness.”

Jill Sellers has been with LexisNexis for 4 years. She has an MA in Library & Information Science from the University of Chicago and a B.A. in Philosophy from Bryn Mawr College. Before coming to LexisNexis in 1997, Jill served as an index coordinator at Matthew Bender and a manuscript editor at the University of Chicago Press. She is currently a Program Manager, responsible for development and maintenance of news and business indexing products.

Leslie Denton has been with LexisNexis for 14 years. She holds a JD and an MLS in Library & Information Science from Indiana University, as well as a BA in Literature from Purdue University. Leslie served as a librarian and media specialist in the Indiana University library system before coming to LexisNexis in 1986. She is currently a Program Manager, responsible for development and maintenance of legal indexing and categorization products.

Patricia Neff is a Senior Standards Analyst, Legal Classification, and has been with LexisNexis for 4 years. She holds an MA from Antioch University McGregor, an MLIS in Library & Information Studies from the University of California, Berkeley, and an AB in Comparative Literature and Slavic Languages & Literatures from UC Berkeley. She has worked in a variety of academic, special, and public libraries, most recently as director of technical services at two law school libraries and as a library systems and training consultant at Innovative Interfaces, Inc.

3:30 – 3:45 Wrapup, Evaluation

4:00 Optional Data Center Tours