

# Introduction to the Digital Libraries Minitrack

Michael Shepherd and James W. Cooper

In the ever-burgeoning field of information overload and information retrieval, how we organize digital documents into libraries becomes increasingly important. Digital libraries can encompass electronic collections of essentially print documents or they can be libraries of documents which have never been printed. More likely digital libraries are an amalgam of formal and informal writing, as well as pictures, sound, video, and various types of animated representations of information.

In the Digital Libraries Minitrack, we present three widely differing papers on aspects of this technical area. Ihlstrom and Lundberg continue reports on their studies of how on-line newspapers behave as a genre, and studied the users of nine different on-line papers to discover how they used the papers and how effectively they presented information.

In order to retrieve the most relevant information from a digital library, it is important that the digital “librarian” be able to determine what you are most interested in retrieving. To this end, Poo, Chng and Goh developed a conceptual model consisting of a hybrid of user profiling and information filtering, and report on their preliminary experiments.

Finally, the ever-productive Watters, Basu and Shepherd examine text categorization techniques, to the end of discovering whether automatic classification of news can be improved by prefiltering the vocabulary. They compare neural networks and vector machine algorithms for accomplishing this classification.