FOREWORD

Research groups worldwide are directing significant effort towards the creation of sophisticated Digital Libraries across a variety of disciplines. At the same time, a large number of companies, government agencies, and individuals are creating Digital Libraries that are accessible from corporate networks or the Internet. Today, a person must know where to search, how to query different media, and how to combine information from diverse resources. As Digital Libraries continue to proliferate, in a variety of media, and from a variety of sources, these problems of resource selection and data fusion become major obstacles. The emergence of multimedia, including text, recorded speech, and images, only exacerbates current problems and emphasises the need for new solutions. An answer to a query might be in a text document, an audio clip, or a newswire photograph. Effective, reliable information retrieval requires the ability to pose multimedia queries across many Digital Libraries.

The research and development area of *Distributed Information Retrieval* addresses issues that arise when people have routine access to thousands of multimedia Digital Libraries. When so many Libraries are available, the first information task users must undertake is to select the resources to search (resource selection). This is a complex manual task if users are unaware of the contents of each individual library in terms of their quantity, quality, information type, provenance and likely relevance. Accurate automatic selection tools to assist selection are required (resource acquisition and description). Once a set of libraries is selected and searched, a person must organise and interpret the retrieved information, possibly in different media (data fusion).

This workshop provides a venue for the presentation of recent results in Distributed Information Retrieval and, in particular, in the design and implementation of methods and tool for resource description, resource selection, data fusion, and user interaction. Thirteen papers were submitted to the workshop in response to the Call for Papers. Each paper was reviewed by the Program Committee listed below.

- Donatella Castelli, ISTI-CNR, Italy
- Jim French, University of Virginia, USA
- Norbert Fuhr, University of Duisburg, Germany
- Luis Gravano,Columbia University, USA
- Umberto Straccia, ISTI-CNR, Italy

Nine of the thirteen papers were selected for presentation at the workshop. The papers cover topics such as distributed crawling, architectures for distributed IR, resource selection, and fusion of information retrieved from different digital libraries.

We thank the Program Committee for their assistance in selecting a strong set of papers for presentation at the workshop. We also thank Jian-Yun Nie, SIGIR 2003 Workshops Chair, for his help. The workshop is partially sponsored by the European Commission Information Society and Technology (IST) project *MIND* on Resource Selection and Data Fusion in Multimedia Distributed Digital Libraries (IST-2000-26061). Information about MIND can be found at http://www.mind-project.net/.

The workshop schedule allows for plenty of discussion about the papers, recent successes in distributed information retrieval, and unmet challenges. A brief post-workshop report summarizing the discussion and major issues raised will appear in the Fall 2003 issue of SIGIR Forum. We encourage workshop participants to make their opinions known, and to make this a very interactive workshop.

Enjoy the workshop!

- Jamie Callan, Carnegie Mellon University, USA
- Fabio Crestani, University of Strathclyde, UK
- Mark Sanderson, University of Sheffield, UK