Introduction

Yaşar Tonta
Editor and Lecture Series Director
Hacettepe University
Department of Information Management
06532 Beytepe, Ankara
TURKEY

tonta@hacettepe.edu.tr

Bibliographic, full-text and multimedia databases available through the intranets, extranets and the Internet are of paramount importance to all organizations large and small. Networked information services proved to be an indispensable part of every day lives of users working for both commercial and non-profit organizations as well as of more casual users with personal interests to pursue. Almost half a billion people try to get access to networked information sources and services every day. More often than not they are confronted with too much information. Although search engines, "knowbots," and "intelligent agents" are of some use in this area, trying to find information among billions of electronic sources is likened to trying to "drink water from a fire hydrant." Well-designed electronic information management systems and services can facilitate users' task and enable them to better cope with too much information in their private and professional lives.

Organized by the Information Management Committee (IMC) of the Research and Technology Organization (RTO) of NATO for the Partnership for Peace (PfP) Nations, Lecture Series on Electronic Information Management aims to review current developments on electronic information management. It explores a wide variety of operational and policy issues with regards to electronic information management ranging from available sources and services to the description, organization, management, preservation and archiving of electronic information collections, to infrastructure, economics and intellectual property rights of electronic information provision.

Available to participants prior to the Lecture Series, this book provides background information on the topic and can serve as an additional resource to support the lectures. It contains 10 papers of lecturers on various aspects of electronic information management. References to both printed and electronic information sources listed in each paper can be useful. Full-text of papers including links to cited sources will be made available through the Web site of IMC (http://www.rta.nato.int). What follows is a brief overview of each paper in the order of their appearance in the book.

In the first paper, "Internet and Electronic Information Management," Dr. **Yaşar Tonta** reviews the latest developments in the electronic information scene. He draws attention to the amount of information produced annually in the world (about five exabyte), increasing processing, storage and transmission capacities of computers as well as declining costs of computer hard drives and network bandwidths. He discusses issues of electronic information description and organization in detail along with development and management of electronic information collections. He reviews the developments in information technologies that gave way to customization and personalization of electronic information services. Dr. Tonta emphasizes the importance of preserving and archiving electronic information and speculates whether publishers, information centers and aggregators would assume this responsibility. Dr. Tonta then goes on to introduce Davenport's ecological model for information management. In his ecological model Davenport sees information, its collection, description, organization, management, and use in a broader context and thinks that information can be better managed if we take its three

Paper presented at the RTO IMC Lecture Series on "Electronic Information Management", held in Sofia, Bulgaria, 8-10 September 2004, and published in RTO-EN-IMC-002.

interrelated and interdependent environments into account, namely (1) the more immediate "information environment," which consists of the whole set of cross-relationships among information people, strategies and policies, processes, technology, information culture and behavior; (2) the "organizational environment," and (3) the "external environment." Dr. Tonta points out that Davenport's ecological model is also applicable to electronic information management.

In their paper, "New Initiatives for Electronic Scholarly Publishing: Academic Information Sources on The Internet," Professors **Ana Maria Ramalho Correia** and **José Carlos Teixeira** review the evolution of scientific communication and discuss in detail the current developments in electronic journals, electronic prints, electronic theses and dissertations (ETDs) and other digital collections of "grey literature" (i.e., technical reports). They provide several examples of such repositories containing electronic information sources that can be used for academic research. The Los Alamos Physics Archive, providing access to e-prints of some 200.000 articles on high-energy physics; and the Networked Digital Library of Theses and Dissertations (NDLTD), providing free access to graduate theses and dissertations in a distributed environment, are among them. Correia and Teixeira also describe initiatives to create a global network of archives of digital research materials (e.g., Open Archives Initiative, Budapest Open Access Initiative) and discuss such policy issues as prior publication, and the possible roles of library and information professionals in self-publishing schemes and in creating digital archives of ETDs.

In his paper, "Information Discovery and Retrieval Tools," **Michael T. Frame** reviews the principles of how search engines work by means of a model. He describes the ways in which search engines discover the existence of Web documents and provides a list of metatags that are used most frequently by search engines for discovery and indexing. He also touches upon the issue of "spam," which some Web site developers are inclined to use to falsify search engines so that their content will be indexed more favorably by search engines and retrieved before the other sites in the retrieval output. His paper ends with a list of features and capabilities of search engines along with some recommendations to content and software developers to improve the discovery and retrieval of their content.

In their paper, **Gladys Cotter**, **Bonnie Carroll**, **Gail Hodge** and **Andrea Japzon** provide a comprehensive overview of electronic collection management and electronic information services. After a brief discussion on the digital revolution that is currently taking place in library and information centers, they first tackle the issue of electronic collection management and review the major collection management strategies. They identify the key challenge in collection management as being that of "ownership vs. access" and stress that the move to electronic information management is transforming information centers to "access-based organizations." They discuss the issues of selection, acquisition, cataloging, and archiving of electronic information in detail. Next, authors review the electronic information services and concentrate on electronic reference, information delivery, and education of users and personnel. They conclude that electronic collection management and electronic information services are in a period of rapid transition, and the technology used to manage the information allows for extensive innovation in information selection, description, distribution, retrieval, and use.

In his paper, "Economics of Electronic Information Provision," **Graham Cornish** covers the economics of preparing and providing published information. He examines the role of different "players" in the publishing chain including authors, editors, publishers, distributors, and users. He also does this for the provision of electronic information and reviews the roles of libraries. He challenges the view that libraries are supermarkets and argues that libraries are not solely run on the basis of commercial motives and that their purchasing policy is not dictated by commercial needs. Libraries make strenuous efforts to collect materials for all types of users and they do not discourage certain types of users such as the children and the elderly. He considers the question of who will pay for those unable to afford access. Finally, he discusses alternative models of electronic information provision (e.g., SPARC, the Scholarly Publishing and Academic Resources Coalition) and reviews the roles of licensing consortia such as ICOLC (International Coalition of Library Consortia) and JSTOR (Journal Storage Online).

Gail Hodge reviews the issues with regards to metadata for electronic information resources. She points out that the rationale for creating metadata remains the same for both electronic and printed resources (to facilitate resource discovery and access), although the terminology has changed (from cataloging and indexing to "metadata"). She describes the purpose of metadata (to discover, locate and organize electronic information resources) and the methods by which metadata can be created (manual vs. through metadata editors and generators). She provides a basic metadata structure and summarizes the characteristics of major metadata schemes (Dublin Core, GILS, TEI, and EAD, to name a few). She also discusses the issues of "metadata interoperability" among different schemes and the importance of controlled vocabularies for subject indexing of electronic information sources.

In her second contribution, entitled "Preservation of and Permanent Access to Electronic Information Resources," **Gail Hodge** starts with the definitions of basic terms such as "digital archiving," "digital preservation" and "long-term access" and gives an outline of major projects including JSTOR, InterPARES (International Research on Permanent Authentic Records in Electronic Systems) and ERPANET (Electronic Resources Preservation and Access Network). She offers a framework for archiving and preservation of electronic information and addresses a number of issues comprehensively. Among them are the creation and acquisition of electronic information, collection development, metadata and archival storage formats for preservation, migration and emulation, access, rights management and security requirements. She also discusses the emerging stakeholder roles and identifies key issues in archiving and preservation of electronic information such as long-term preservation and intellectual property rights.

In his second contribution, **Graham Cornish** addresses the intellectual property rights in the context of electronic information management. He clarifies the use of such basic terms as "copyright," "copy," "author," "publisher," "user" and "fair use" in the digital environment. He explains access control devices including fingerprinting, watermarking, and stamping, and gives examples of their use in the European Union (EU) projects such as CITED (Copyright in Transmitted Electronic Documents) and COPY SMART. He discusses the impact of the latest EU directive on copyright and information society and the complexities of implementing this directive in different legal regimes and cultural environments.

Gregory D. Twitchell and **Michael T. Frame**'s paper addresses the infrastructure of electronic information management. Using a non-technical language as much as possible, they describe the following tools and technologies: network infrastructure, mass storage devices, JAVA, proxy servers, network address translation, firewalls, tunnelling, forwarding, encryption, and routing. They emphasize that the key to a robust, flexible, secure, and usable network systems is to establish a strong network infrastructure and point out that network hardware and applications are co-dependent. They conclude that a standard component of a good network management is the planning and review process, and organizations must take a proactive role in this process to make sure that they have a secure, reliable, usable and scalable network.

The book concludes with Dr. **Dincho Krastev**'s piece on the development of digital libraries in Bulgaria. Dr. Krastev provides a summary of some of the early digital library projects and gives a detailed description of the digitization of Slavic manuscripts that was carried out in the Central Library of the Bulgarian Academy of Sciences in cooperation with the national and international institutions. He ends his paper by emhasizing the importance of team work to succeed in such collaborative digitization projects involving specialists in Slavic manuscripts and medieval texts, computational medieval studies and computational humanities.