Chapter VI
A Conceptual Structure for Designing Personalized Information Seeking and Retrieval Systems in Data-Intensive Domains

Nong Chen
Delft University of Technology, The Netherlands

Ajantha Dahanayake
Georgia College & State University, USA

ABSTRACT

Personalized information seeking and retrieval is regarded as the solution to the problem of information overload in domains such as crisis response and medical networks. Personalization algorithms and techniques are maturing, but their centralized implementation solutions are becoming less efficient for dealing with ever-changing user information needs in data-intensive, dynamic, and distributed environments. In this chapter, we present a conceptual structure for designing personalized, multidisciplinary information seeking and retrieval systems. This conceptual structure is capable of serving as a bridge between information needs coming from an organizational process, and existing implementations of information access services, software, applications, and technical infrastructure; it is also capable of sufficiently describing and inferring users’ personalized information needs. We believe that it offers a new way of thinking about the retrieval of personalized information.

INTRODUCTION

The technological developments of the last 50 years have made more information more available to more people than at any other time in human history (Feather, 1998). The increase in widely available Internet communication tools, especially the World Wide Web, has provided a catalyst for a revolution in presenting, gathering, sharing, processing, and using information. Enabled by
several distributed infrastructure and technologies based on Microsoft’s Component Object Model (COM) and .NET (Microsoft), Object Management Group’s (OMG) Common Object Request Broker Architecture (CORBA) (Siegel, 2000), or Java-based tools (Sun Microsystems; Stojanovic, 2003), information is accessible globally today simply via the Internet, middleware, or Web service bus. Furthermore, the availability and popularity of small mobile devices have accelerated the growth of user mobility. Organizations and information seekers now have the privilege of anywhere/anytime information access via wired or wireless networks. Technology availability has significantly encouraged information sharing between business, scientific, or other organizational coordination processes distributed over various independent locations. Attempting to expand the sophistication and scope of data-intensive applications to share and retrieve information over disciplines, organizational, and geographic boundaries stimulates people in domains such as crisis response, medical and healthcare networks, national and international security networks, and so forth to develop complex, Web-enabled, multidisciplinary information seeking and retrieval applications and services. Our world is becoming increasingly interconnected.

This increase in information availability cannot guarantee that organizations and information seekers are able to retrieve and access the information they really need. One of the biggest problems organizations are facing today is the sheer amount of information received and created that has to be catalogued and securely shared. An overwhelming amount of information from many sources must be dealt with as part of their work. The volume of information causes problems with trying to search an immense collection of data for a small and specific set of knowledge, and with dealing with inconsistencies, errors, and useless and conflicting information (Nelson, 2001). Heterogeneous information resources exacerbate the problem of information access. New information types, such as image, animation, video, music, and so forth, and databases or information systems built from a variety of purposes, different technologies, and using different methodologies, make information seeking and retrieval even more complex.

In addition, organizational and information seekers’ information needs are changing over time under different situations, scenarios, and even personal preferences; and many of these cannot be predicted or are short lived. The traditional IT approaches that tried to address inter-organizational information access over boundaries are no longer applicable, as the initial assumption of the design paradigm was based on a centralized system. In other words, bringing diverse information into a central store with a predefined data structure to manage and control the solution space cannot efficiently support rapidly changing information needs or the organizational structures to be formed in dynamic and distributed environments. Changes in an organization’s or a person’s information needs may lead to the need to redesign a complete application.

Obviously, the huge amount of available information, the heterogeneous nature of the information resources, and the dynamically changing information needs of those seeking information make it increasingly difficult to find the “right information” in the “right format” at the “right time.” Dealing with the problems of information seeking and retrieval in data-intensive domains shows that it is no longer realistic to design the large information systems of the past. To build such a multiple disciplinary information seeking and retrieval system, the ever-increasing availability of component-based design methods, service-oriented architectures, distributed infrastructures, and other technological achievements provides us with a technical foundation we can use to address the requirement for flexibility and extendibility that we come up against when designing information systems today. Modularization of complex systems into components or services that interoperate primarily via exchang-
Related Content

Performance Evaluation of Hierarchical SOAP Based Web Service in Load Balancing Cluster-Based and Non-Cluster-Based Web Server

A New Approach Based on the Detection of Opinion by SentiWordNet for Automatic Text Summaries by Extraction
[www.igi-global.com/article/a-new-approach-based-on-the-detection-of-opinion-by-sentiwordnet-for-automatic-text-summaries-by-extraction/161659?camid=4v1a](www.igi-global.com/article/a-new-approach-based-on-the-detection-of-opinion-by-sentiwordnet-for-automatic-text-summaries-by-extraction/161659?camid=4v1a)

Design of a Least Cost (LC) Vertical Search Engine based on Domain Specific Hidden Web Crawler
[www.igi-global.com/article/design-of-a-least-cost-lc-vertical-search-engine-based-on-domain-specific-hidden-web-crawler/177280?camid=4v1a](www.igi-global.com/article/design-of-a-least-cost-lc-vertical-search-engine-based-on-domain-specific-hidden-web-crawler/177280?camid=4v1a)

A Roadmap to Integrate Document Clustering in Information Retrieval
[www.igi-global.com/article/roadmap-integrate-document-clustering-information/53125?camid=4v1a](www.igi-global.com/article/roadmap-integrate-document-clustering-information/53125?camid=4v1a)