Information access through search engines and digital libraries

Maristella Agosti

Department of Information Engineering, University of Padua, Italy
maristella.agosti@unipd.it

Abstract. Information Retrieval (IR) identifies the activities that a person (user) has to conduct in order to choose from a collection of documents those that will satisfy a specific and contingent information need. The aim of IR therefore is to help and support the user in choosing those documents from among the available ones that are more likely to satisfy their information need. When the collection of documents reaches a size that makes manual inspection of the documents prohibitive, as is the case with web pages and web documents, the collection and the application of the retrieval function are automatically managed through an information retrieval system, also called a search engine. Digital Libraries (DL) have been steadily progressing over time and they now determine how citizens and organizations study, learn, access and interact with their document collections. Despite their name, DL are not only the digital counter-part of traditional libraries, but they also deal with other kinds of organizations, such as archives and museums, and also private companies. IR and DL have the common aim of satisfying user information needs across huge document collections. However, they deal with document collections that differ in the way they are built, in the type of documents that are managed and in the contents that are made available. The talk will show how in the early days both IR and DL were mostly concerned with retrieving and providing access to relevant information, which was then typically processed and used outside the system providing it. Today, however, there is an increasing need to engage users with the digital contents through personalized search or by allowing them to add their own content in different forms.

Keywords: search engines, digital libraries