Preface

WWW has become currently an inseparable part of our life. Today we are living in digital society, where large amount of multimedia information is produced and stored daily. Retrieving relevant information from these large repositories is a challenging task. The present web generation, namely Web 2.0 focuses on embedding semantics with information. It is called as Semantic Web i.e. meaning oriented web, which means rather than returning information based on keyword matching, information would be returned on the basis of its context. Recently, the web includes text as well as images, audio, video and many other forms of unstructured data generated from messengers.

The semantic web term was coined in 2001, since then lot of research work had been done towards architecture of semantic web, towards developing methods and techniques for embedding semantics with textual information. However, since last one decade there had been tremendous growth in capturing, sharing and retrieving photographs, audio and video data on the web. With widespread popularity of hand-held smart mobile devices and availability of Internet anywhere, anytime has led to this growth of multimedia data. There is also increase in attempt of retrieving such information; however retrieving relevant multimedia information from web is still very tough task due to the lack of sufficient information about uploaded multimedia contents. Now, research community has started putting efforts towards incorporating semantics with multimedia contents so that relevant contents may be retrieved in future searches. For associating semantics with pictures, their low-level features need to be extracted and to be expressed in words. Further, automation of this process is very much desired, considering the frequency of contribution on the web.

Consequently, there is an urgent need for an edited collection elaborating techniques of embedding semantics with multimedia information. Specially, techniques for automatic semantic annotation of images must be presented.

OBJECTIVE OF THE BOOK

Objective of this book is to present existing methods and techniques of associating semantics with textual and visual information. Techniques used for associating semantics with information at low level are also elaborated. Methods used for extracting knowledge from web and techniques used for this purpose are explained. Further, extracting knowledge from Big-data available on web, significance of such knowledge discovery and challenge's prevailing needs to be focused. This book also aims to present various application areas of semantics oriented multimedia information.

ORGANIZATION OF THE BOOK

This book is comprised of 11 chapters arranged in four sections. Section 1 comprises of three chapters and focuses on methods and techniques of semantic information retrieval from textual contents as well as multimedia contents. Second section contains two chapters that provide low level details of associating semantics with information. The third section focuses on extracting knowledge from multimedia information through providing techniques for knowledge discovery from textual data as well as from big data present on the web. The fourth section of the book includes four chapters, which illustrating various dimensions of applicability of semantic based contents in present digital society. The summary for the book organization is as follows.

Section 1: Tools and Techniques of Information Retrieval (Chapters 1-3)

This section highlights the information retrieval techniques, which are currently an active research field with the evolution of World Wide Web.

Chapter 1: This chapter focuses on the elaborating techniques that are available for embedding meaning with textual and visual information. It includes several image retrieval process and techniques and also outlines intelligent software agents, which can be used for automatic context based image retrieval.

Chapter 2: This chapter included an insight into the information retrieval definitions, process and models. The authors addressed the traditional information retrieval procedures with adopted search engines. Furthermore, the CLIR process is introduced in brief. The role of external sources such as ontologies in information retrieval systems are also entails.

Chapter 3: This chapter highlighted the existing methods for information retrieval, such as the basic model, Boolean model, vector space model, latent semantic, and the probabilistic model. All these methods own pros and cons based on the kind of information that the user requires. Furthermore, the authors discussed the IR capabilities along with the text based and content based information retrieval methods that used for extracting the information. The authors established that the CBIR is more efficient and effective than the TBIR with accurate results. The evaluation of ranked and unranked retrieval result which retrieved after a query is also addressed.

Section 2: Associating Semantics With Information (Chapters 4-5)

The semantic information is the entry part that defines the different attributes for an entry. The main goal of this section is to establish the concept of semantic information firmly on using mapping techniques.

Chapter 4: This chapter discussed the importance of better mapping techniques for Relational Database (RDB) to Resource Description Framework (RDF). It represented the limitations and benefits of the data model to achieve superior results. Each transformed data form has its own significance in the field of data science. For example, the RDB can be employed in several applications especially the web, desktop, remote, embedded, and network-based applications. Although, the EXtensible Markup Language (XML) is involved for data for transferring among all computer related resources regardless of their type, shape, place, capability and capacity due to its form is in application understandable form. The authors established that the proposed RDF can be used to achieve the nearest form of maximized customization when transforming one data model to another.

Chapter 5: Word Sense Disambiguation (WSD) is complex as it handles full language complexities to identify the meaning of words in context in a computational manner. This chapter proposed an exclusive approach to expand a machine translation system beginning higher language to lower language. Since Hindi is considered the regional language in India, a machine translation system has been proposed. An apertium platform has been used, where a lot of language pairs more specifically Indian language pairs. In this chapter, a machine translation system has been developed for strongly related language pair i.e Hindi.

Section 3: Extracting Knowledge From Information (Chapters 6-7)

The era of big data leads to the ability to process massive datasets from heterogeneous sources in real-time. But the conventional analytics can't be able to manage such a large amount of varied data. This section is concerned with the data discovery concept and tools.

Chapter 6: Knowledge Discovery is a significant process for finding new knowledge about an application domain from large data volumes for making decisions and strategies in related application areas. It has several applications in numerous scientific and business domains; however, the World Wide Web is one of the most fertile areas for data mining and knowledge discovery research as it contains huge amount of information. Web mining is the application of data mining techniques to discover and analyze potentially useful information from Web data. This chapter included a comparison of popular knowledge Discovery process models, where new types of data, new applications, and new analysis demands continue to emerge leading to novel data mining and knowledge discovery tasks. Finally, the authors depicted that the future seems to be dazzling for knowledge discovery. Increasing computational power and continuous creative solutions will be certainly revolutionizing the way for the data mining and information processing.

Chapter 7: This chapter addressed the different techniques that can be applied to design a high-performance computing platform for effective analytics on big data. In addition, the right mining schemes to get useful insights from voluminous big data are also included. The authors elaborated these challenges with an initial briefing about traditional data analytics followed by the mining algorithms that are suitable for emerging big data analytics. Subsequently, other issues and future scope are also presented to enhance the capabilities of big data.

Section 4: Suggested Reading – Applicability of Semantics-Driven Contents (Chapters 8-11)

Chapter 8: This chapter has extensively discussed the synergy between Semantic Web (SW) technologies and Web Personalization (WP) for demonstrating an intelligent interface for Personalized Information Retrieval (PIR) on web. Adding semantics to WP through ontologies and Software Agents (SA) realizes several benefits, which are also included. A brief introduction to the PIR process is provided followed by description of SW, ontologies and SA. An inclusive review of the existing web semantics for PIR has been included.

Chapter 9: The progress in computers leads to research benefits of scientific community for sharing data, service computing, building the frameworks and many more. E-Science becomes an active extending field by the increase of the data and tools. The main objective of this chapter is to discuss the use of semantic web applications to identify the components in the development of scientific workflows. It developed a framework which assists the scientific community to test and deploy the scientific experiments with the help of ontologies, service repositories, web services and scientific workflows. The overall objective of this chapter is to automate the use of semantic web services, to generate the workflows and to manage the search services and the ontologies by considering the web service composition.

Chapter 10: This chapter focuses on finding relevant information from the biological databases that contain information and knowledge collected from scientific experiments, published literature and statistical analysis of text, numerical, image and video data. Several tools have been developed to retrieve the information from these databases. This chapter is concerned with the present biological resources and their availability and integration at single platform. It provides an insight into existing biological resources with an aim to provide consolidated information at one place for ease of use and access by researchers, academicians and students.

Chapter 11: This chapter is briefly describes the issues and challenges of the social networks for teaching and working insertion. The authors assessed the social media tools, specifically Facebook, Blogs, Google groups, SkyDrive and Twitter that are used by the teachers and the students as well for academic purposes. Further, the relationship between usage of SNS and academic performance of students has been tested statistically by views of teachers in most of the universities and with affiliated colleges.

This book is expected to assist researchers, academicians, technology developers and engineers working in the field of Web Semantics. Furthermore, the book provides insights and support executives concerned with recent web semantics technologies that have magnetized much attention. It addressed innovative conceptual framework for various applications. The book is expected to serve as a reference for the post-graduate students as it offers the requisite knowledge for understanding the knowledge discovery techniques along with different applications. This book is based on a research studies carried out by experienced academicians and is expected to shed new insights for researchers; academicians, students and improves understanding of Web Semantics.

Aarti Singh
Guru Nanak Girls College, Yamuna Nagar, India

Nilanjan Dey
Techno India College of Technology, India

Amira S. Ashour
Tanta University, Egypt & Taif University, Saudi Arabia

V. Santhi
VIT University, India