

## Advances In Automatic Text Summarization

These two volumes, LNCS 7076 and LNCS 7077, constitute the refereed proceedings of the Second International Conference on Swarm, Evolutionary, and Memetic Computing, SEMCCO 2011, held in Visakhapatnam, India, in December 2011. The 124 revised full papers presented in both volumes were carefully reviewed and selected from 422 submissions. The papers explore new application areas, feature new bio-inspired algorithms for solving specific hard optimization problems, and review the latest progresses in the cutting-edge research with swarm, evolutionary, and memetic computing in both theoretical and practical aspects.

This volume is based on the research papers presented in the 4th Computer Science On-line Conference. The volume Artificial Intelligence Perspectives and Applications presents new approaches and methods to real-world problems, and in particular, exploratory research that describes novel approaches in the field of artificial intelligence. Particular emphasis is laid on modern trends in selected fields of interest. New algorithms or methods in a variety of fields are also presented. The Computer Science On-line Conference (CSOC 2015) is intended to provide an international forum for discussions on the latest high-quality research results in all areas related to Computer Science. The addressed topics are the theoretical aspects and applications of Computer Science, Artificial Intelligences, Cybernetics, Automation Control Theory and Software Engineering. This handbook plays a fundamental role in sustainable progress in speech research and development. With an accessible format and with accompanying DVD-Rom, it targets three categories of readers: graduate students, professors and active researchers in academia, and engineers in industry who need to understand or implement some specific algorithms for their speech-related products. It is a superb source of application-oriented, authoritative and comprehensive information about these technologies, this work combines the established knowledge derived from research in such fast evolving disciplines as Signal Processing and Communications, Acoustics, Computer Science and Linguistics.

This book constitutes the refereed proceedings of the 14th Artificial Intelligence Conference sponsored by the Canadian Society for Computational Studies of Intelligence, AI 2001, held in Ottawa, Canada, in June 2001. The 24 revised full papers presented together with 14 posters were carefully reviewed and selected from around 70 submissions. Among the topics addressed are learning, data mining, searching, multi-agent systems, automated deduction, computational linguistics, constraint programming, agent learning, planning, classifier systems, heuristics, logic programming, and case-based reasoning.

This book lays out a path leading from the linguistic and cognitive basics, to classical rule-based and machine learning algorithms, to today's state-of-the-art approaches, which use advanced empirically grounded techniques, automatic knowledge acquisition, and refined linguistic modeling to make a real difference in real-world applications. Anaphora and coreference resolution both refer to the process of linking textual phrases (and, consequently, the information attached to them) within as well as across sentence boundaries, and to the same discourse referent. The book offers an overview of recent research advances, focusing on practical, operational approaches and their applications. In part I (Background), it provides a general introduction, which succinctly summarizes the linguistic, cognitive, and computational foundations of anaphora processing and the key classical rule- and machine-learning-based anaphora resolution algorithms. Acknowledging the central importance of shared resources, part II (Resources) covers annotated corpora, formal evaluation, preprocessing technology, and off-the-shelf anaphora resolution systems. Part III (Algorithms) provides a thorough description of state-of-the-art anaphora resolution algorithms, covering enhanced machine learning methods as well as techniques for accomplishing important subtasks such as mention detection and acquisition of relevant knowledge. Part IV (Applications) deals with a selection of important anaphora and coreference resolution applications, discussing particular scenarios in diverse domains and distilling a best-practice model for systematically approaching new application cases. In the concluding part V (Outlook), based on a survey conducted among the contributing authors, the prospects of the research field of anaphora processing are discussed, and promising new areas of interdisciplinary cooperation and emerging application scenarios are identified. Given the book's design, it can be used both as an accompanying text for advanced lectures in computational linguistics, natural language engineering, and computer science, and as a reference work for research and independent study. It addresses an audience that includes academic researchers, university lecturers, postgraduate students, advanced undergraduate students, industrial researchers, and software engineers.

Knowledge-Based Information Retrieval and Filtering from the Web contains fifteen chapters, contributed by leading international researchers, addressing the matter of information retrieval, filtering and management of the information on the Internet. The research presented deals with the need to find proper solutions for the description of the information found on the Internet, the description of the information consumers need, the algorithms for retrieving documents (and indirectly, the information embedded in them), and the presentation of the information found. The chapters include: -Ontological representation of knowledge on the WWW; -Information extraction; -Information retrieval and administration of distributed documents; -Hard and soft modeling based knowledge capture; -Summarization of texts found on the WWW; -User profiles and personalization for web-based information retrieval system; -Information retrieval under constricted bandwidth; -Multilingual WWW; -Generic hierarchical classification using the single-link clustering; -Clustering of documents on the basis of text fuzzy similarity; -Intelligent agents for document categorization and adaptive filtering; -Multimedia retrieval and data mining for E-commerce and E-business; -A Web-based approach to competitive intelligence; -Learning ontologies for domain-specific information retrieval; -An open, decentralized architecture for searching for, and publishing information in distributed systems.

ARIST, published annually since 1966, is a landmark publication within the information science community. It surveys the landscape of information science and technology, providing an analytical, authoritative, and accessible overview of recent trends and significant developments. The range of topics varies considerably, reflecting the dynamism of the discipline and the diversity of theoretical and applied perspectives. While ARIST continues to cover key topics associated with "classical" information science (e.g., bibliometrics, information retrieval), editor Blaise Cronin is selectively expanding its footprint in an effort to connect information science more tightly with cognate academic and professional communities. Contents of Volume 40 (2006): SECTION I: Information and Society Chapter 1: The Micro- and Macroeconomics of Information, Sandra Braman Chapter 2: The Geographies of the Internet, Matthew Zook Chapter 3: Open Access, M. Carl Drott SECTION II: Technologies and Systems Chapter 4: TREC: An Overview, Donna K. Harman and Ellen M. Voorhees Chapter 5: Semantic Relations in Information Science, Christopher S. G. Khoo and Jin-Cheon Na Chapter 6: Intelligence and Security Informatics, Hsinchun Chen and Jennifer Xu SECTION III: Information Needs and Use Chapter 7: Information Behavior, Donald O. Case Chapter 8: Collaborative Information Seeking and Retrieval, Jonathan Foster Chapter 9: Information Failures in Health Care, Anu MacIntosh-Murray and Chun Wei Choo Chapter 10: Workplace Studies and Technological Change, Angela Cora Garcia, Mark E. Dawes, Mary Lou Kohne, Felicia Miller, and Stephan F. Groschwitz SECTION IV: Theoretical Perspectives Chapter 11: Information History, Alistair Black Chapter 12: Social Epistemology and Information Science, Don Fallis Chapter 13: Formal Concept Analysis in Information Science, Uta Priss. Advances in Automatic Text Summarization MIT Press

In the last few years, there has been an increased advancement and evolution in semantic web and information systems in a variety of fields. The integration of these approaches to ontology engineering, sophisticated methods and algorithms for open linked data extraction, and advanced decision-making creates new opportunities for a bright future. Innovations, Developments, and Applications of Semantic Web and Information Systems is a critical scholarly resource that discusses integrated methods of research and analytics in information technology. Featuring coverage on a broad range of topics, such as cognitive computing, artificial intelligence, machine learning, data analysis, and algorithms, this book is geared towards researchers, academicians, and professionals seeking current information on semantic web and information systems.

As technology continues to become more sophisticated, a computer's ability to understand, interpret, and manipulate natural language is also accelerating. Persistent research in the field of natural language processing enables an understanding of the world around us, in addition to opportunities for manmade computing to mirror natural language processes that have existed for centuries. Natural Language Processing: Concepts, Methodologies, Tools, and Applications is a vital reference source on the latest concepts, processes, and techniques for communication between computers and humans. Highlighting a range of topics such as machine learning, computational linguistics, and semantic analysis, this multi-volume book is ideally designed for computer engineers, computer and software developers, IT professionals, academicians, researchers, and upper-level students seeking current research on the latest trends in the field of natural language processing.

This 4-Volume-Set, CCIS 0251 - CCIS 0254, constitutes the refereed proceedings of the International Conference on Informatics Engineering and Information Science, ICIEIS 2011, held in Kuala Lumpur, Malaysia, in November 2011. The 210 revised full papers presented together with invited papers in the 4 volumes were carefully reviewed and selected from numerous submissions. The papers are organized in topical sections on e-learning, information security, software engineering, image processing, algorithms, artificial intelligence and soft computing, e-commerce, data mining, neural networks, social networks, grid computing, biometric technologies, networks, distributed and parallel computing, wireless networks, information and data management, web applications and software systems, multimedia, ad hoc networks, mobile computing, as well as miscellaneous topics in digital information and communications.

While the availability of electronic documents increases exponentially with advancing technology, the time spent to process this wealth of resourceful information decreases. Content analysis and information extraction must be aided by summarization methods to quickly parcel pieces of interest and allow for succinct user familiarization in a simple, efficient manner. Trends and Applications of Text Summarization Techniques is a pivotal reference source that explores the latest approaches of document summarization including update, multi-lingual, and domain-oriented summarization tasks and examines their current real-world applications in multiple fields. Featuring coverage on a wide range of topics such as parallel construction, social network integration, and evaluation metrics, this book is ideally designed for information technology practitioners, computer scientists, bioinformatics analysts, business managers, healthcare professionals, academicians, researchers, and students.

The 7th IEEE/ACIS Conference and the 2nd IEEE/ACIS Workshop on e-Activity (IWEA 2008) featured researchers from around the world. The conference organizers selected 23 outstanding papers for this volume of Springer's Studies in Computational Intelligence.

The Meaning of Language illustrates the diversity of approaches in linguistics. The volume revolves around two main chapters authored by two internationally acknowledged Scandinavian scholars, Hans Basbøll and Stig Eliasson. Basbøll's contribution is the most detailed and coherent English-language presentation of the pioneering Danish 18th century linguist Jens Pedersen Høysgaard and his work, and Eliasson explores the intricacy of the issue of whether morphology can be borrowed between languages and the mechanisms of actual borrowings. The other contributions illustrate which topics may be taken up by language scholars today, from metaphor, regional phonology, morphology and syntax, language learning, discourse analysis, intensifier semantics, and Indo-European, to the interface between language and logic. The approaches invoke a wide spectrum of theoretical models and assumptions.

The refereed proceedings of the 6th International Workshop on Computational Processing of the Portuguese Language, PROPOR 2003, held in Faro, Portugal, in June 2003. The 24 revised full papers and 17 revised short papers presented were carefully reviewed and selected from 64 submissions. The papers are organized in topical sections on speech analysis and recognition; speech synthesis; pragmatics, discourse, semantics, syntax, and the lexicon; tools, resources, and applications; dialogue systems; summarization and information extraction; and evaluation. Textual information in the form of digital documents quickly accumulates to create huge amounts of data. The majority of these documents are unstructured: it is unrestricted text and has not

been organized into traditional databases. Processing documents is therefore a perfunctory task, mostly due to a lack of standards. It has thus become extremely difficult to implement automatic text analysis tasks. Automatic Text Summarization (ATS), by condensing the text while maintaining relevant information, can help to process this ever-increasing, difficult-to-handle, mass of information. This book examines the motivations and different algorithms for ATS. The author presents the recent state of the art before describing the main problems of ATS, as well as the difficulties and solutions provided by the community. The book provides recent advances in ATS, as well as current applications and trends. The approaches are statistical, linguistic and symbolic. Several examples are also included in order to clarify the theoretical concepts.

Information extraction (IE) and text summarization (TS) are powerful technologies for finding relevant pieces of information in text and presenting them to the user in condensed form. The ongoing information explosion makes IE and TS critical for successful functioning within the information society. These technologies face particular challenges due to the inherent multi-source nature of the information explosion. The technologies must now handle not isolated texts or individual narratives, but rather large-scale repositories and streams---in general, in multiple languages---containing a multiplicity of perspectives, opinions, or commentaries on particular topics, entities or events. There is thus a need to adapt existing techniques and develop new ones to deal with these challenges. This volume contains a selection of papers that present a variety of methodologies for content identification and extraction, as well as for content fusion and regeneration. The chapters cover various aspects of the challenges, depending on the nature of the information sought---names vs. events,--- and the nature of the sources---news streams vs. image captions vs. scientific research papers, etc. This volume aims to offer a broad and representative sample of studies from this very active research field.

This book constitutes the thoroughly refereed proceedings of the Second International Joint Conference on Natural Language Processing, IJCNLP 2005, held in Jeju Island, Korea in October 2005. The 88 revised full papers presented in this volume were carefully reviewed and selected from 289 submissions. The papers are organized in topical sections on information retrieval, corpus-based parsing, Web mining, rule-based parsing, disambiguation, text mining, document analysis, ontology and thesaurus, relation extraction, text classification, transliteration, machine translation, question answering, morphological analysis, text summarization, named entity recognition, linguistic resources and tools, discourse analysis, semantic analysis NLP applications, tagging, language models, spoken language, and terminology mining.

Foundations of Computational Intelligence Volume 6: Data Mining: Theoretical Foundations and Applications Finding information hidden in data is as theoretically difficult as it is practically important. With the objective of discovering unknown patterns from data, the methodologies of data mining were derived from statistics, machine learning, and artificial intelligence, and are being used successfully in application areas such as bioinformatics, business, health care, banking, retail, and many others. Advanced representation schemes and computational intelligence techniques such as rough sets, neural networks; decision trees; fuzzy logic; evolutionary algorithms; artificial immune systems; swarm intelligence; reinforcement learning, association rule mining, Web intelligence paradigms etc. have proved valuable when they are applied to Data Mining problems. Computational tools or solutions based on intelligent systems are being used with great success in Data Mining applications. It is also observed that strong scientific advances have been made when issues from different research areas are integrated. This Volume comprises of 15 chapters including an overview chapter providing an up-to-date and state-of-the research on the applications of Computational Intelligence techniques for Data Mining. The book is divided into 3 parts: Part-I: Data Click Streams and Temporal Data Mining Part-II: Text and Rule Mining Part-III: Applications Part I on Data Click Streams and Temporal Data Mining contains four chapters that describe several approaches in Data Click Streams and Temporal Data Mining.

This book deals with "Computer Aided Writing", CAW for short. The contents of that is a sector of Knowledge based technics and Knowledge Management. The role of Knowledge Management in social media, education and Industry 4.0 is out of question. More important is the expectation of combining Knowledge Management and Cognitive Technology, which needs more and more new innovations in this field to face recent problems in social and technological areas. The book is intended to provide an overview of the state of research in this field, show the extent to which computer assistance in writing is already being used and present current research contributions. After a brief introduction into the history of writing and the tools that were created, the current developments are examined on the basis of a formal writing model. Tools such as word processing and content management systems will be discussed in detail. The special form of writing, "journalism", is used to examine the effects of Computer Aided Writing. We dedicate a separate chapter to the topic of research, since it is of essential importance in the writing process. With Knowledge Discovery from Text (KDT) and recommendation systems we enter the field of Knowledge Management in the context of Computer Aided Writing. Finally, we will look at methods for automated text generation before giving a final outlook on future developments.

The marketing of a destination necessitates strategic planning, decision making, and organization. Effective positioning will result in a strong brand that develops an emotional and productive two-way relationship. Notwithstanding, destination managers should possess relevant knowledge and understanding on traditional and contemporary marketing channels to better engage with prospective visitors. Destination Management and Marketing: Breakthroughs in Research and Practice focuses on utilizing destination branding and content marketing for sustainable growth and competitive advantage within the tourism and hospitality industry, including tools and techniques for travel branding and best practices for better tourism management strategies. Highlighting a range of topics such as service quality, sustainable tourism, and competitiveness model, this publication is an ideal reference source for government officials, travel agencies, advertisers, marketers, tour directors, hotel managers, restaurateurs, industry professionals including those within the hotel, leisure, transportation, and theme park sectors, policymakers, practitioners, academicians, researchers, and students.

Until now there has been no state-of-the-art collection of the most important writings in automatic text summarization. This book presents the key developments in the field in an integrated framework and suggests future research areas. With the rapid growth of the World Wide Web and electronic information services, information is becoming available on-line at an incredible rate. One result is the oft-decried information overload. No one has time to read everything, yet we often have to make critical decisions based on what we are able to assimilate. The technology of automatic text summarization is becoming indispensable for dealing with this problem. Text summarization is the process of distilling the most important information from a source to produce an abridged version for a particular user or task. Until now there has been no state-of-the-art collection of the most important writings in automatic text summarization. This book presents the key developments in the field in an integrated framework and suggests future research areas. The book is organized into six sections: Classical Approaches, Corpus-Based Approaches, Exploiting Discourse Structure, Knowledge-Rich Approaches, Evaluation Methods, and New Summarization Problem Areas. Contributors D. A. Adams, C. Aone, R. Barzilay, E. Bloedorn, B. Boguraev, R. Brandow, C. Buckley, F. Chen, M. J. Chrzanowski, H. P. Edmundson, M. Elhadad, T. Firmin, R. P. Futrelle, J. Gorlinsky, U. Hahn, E. Hovy, D. Jang, K. Sparck Jones, G. M. Kasper, C. Kennedy, K. Kukich, J. Kupiec, B. Larsen, W. G. Lehnert, C. Lin, H. P. Luhn, I. Mani, D. Marcu, M. Maybury, K. McKeown, A. Merlino, M. Mitra, K. Mitze, M. Moens, A. H. Morris, S. H. Myaeng, M. E. Okurowski, J. Pedersen, J. J. Pollock, D. R. Radev, G. J. Rath, L. F. Rau, U. Reimer, A. Resnick, J. Robin, G. Salton, T. R. Savage, A. Singhal, G. Stein, T. Strzalkowski, S. Teufel, J. Wang, B. Wise, A. Zamora

With the explosion in the quantity of on-line text and multimedia information in recent years, there has been a renewed interest in automatic summarization. This book provides a systematic introduction to the field, explaining basic definitions, the strategies used by human summarizers, and automatic methods that leverage linguistic and statistical knowledge to produce extracts and abstracts. Drawing from a

wealth of research in artificial intelligence, natural language processing, and information retrieval, the book also includes detailed assessments of evaluation methods and new topics such as multi-document and multimedia summarization. Previous automatic summarization books have been either collections of specialized papers, or else authored books with only a chapter or two devoted to the field as a whole. This is the first textbook on the subject, developed based on teaching materials used in two one-semester courses. To further help the student reader, the book includes detailed case studies, accompanied by end-of-chapter reviews and an extensive glossary. Audience: students and researchers, as well as information technology managers, librarians, and anyone else interested in the subject.

The prevalence of digital documentation presents some pressing concerns for efficient information retrieval in the modern age. Readers want to be able to access the information they desire without having to search through a mountain of unrelated data, so algorithms and methods for effectively seeking out pertinent information are of critical importance. Innovative Document Summarization Techniques:

Revolutionizing Knowledge Understanding evaluates some of the existing approaches to information retrieval and summarization of digital documents, as well as current research and future developments.

This book serves as a sounding board for students, educators, researchers, and practitioners of information technology, advancing the ongoing discussion of communication in the digital age.

This book constitutes the refereed proceedings of the 16th Conference of the Canadian Society for Computational Studies of Intelligence, AI 2003, held in Halifax, Canada in June 2003. The 30 revised full papers and 24 revised short papers presented were carefully reviewed and selected from 106 submissions. The papers are organized in topical sections on knowledge representation, search, constraint satisfaction, machine learning and data mining, AI and Web applications, reasoning under uncertainty, agents and multi-agent systems, AI and bioinformatics, and AI and e-commerce.

This book constitutes the refereed proceedings of the 17th Brazilian Symposium on Artificial Intelligence, SBIA 2004, held in Sao Luis, Maranhao, Brazil in September/October 2004. The 54 revised full papers presented were carefully reviewed and selected from 208 submissions from 21 countries. The papers are organized in topical sections on logics, planning, and theoretical methods; search, reasoning, and uncertainty; knowledge representation and ontologies; natural language processing; machine learning, knowledge discovery and data mining; evolutionary computing, artificial life, and hybrid systems; robotics and computer vision; and autonomous agents and multi-agent systems.

Concise Encyclopedia of Semantics is a comprehensive new reference work aiming to systematically describe all aspects of the study of meaning in language. It synthesizes in one volume the latest scholarly positions on the construction, interpretation, clarification, obscurity, illustration, amplification, simplification, negotiation, contradiction, contraction and paraphrasing of meaning, and the various concepts, analyses, methodologies and technologies that underpin their study. It examines not only semantics but the impact of semantic study on related fields such as morphology, syntax, and typologically oriented studies such as 'grammatical semantics', where semantics has made a considerable contribution to our understanding of verbal categories like tense or aspect, nominal categories like case or possession, clausal categories like causatives, comparatives, or conditionals, and discourse phenomena like reference and anaphora. COSE also examines lexical semantics and its relation to syntax, pragmatics, and cognitive linguistics; and the study of how 'logical semantics' develops and thrives, often in interaction with computational linguistics. As a derivative volume from Encyclopedia of Language and Linguistics, Second Edition, it comprises contributions from 150 of the foremost scholars of semantics in their various specializations and draws on 20+ years of development in the parent work in a compact and affordable format. Principally intended for tertiary level inquiry and research, this will be invaluable as a reference work for undergraduate and postgraduate students as well as academics inquiring into the study of meaning and meaning relations within languages. As semantics is a centrally important and inherently cross-cutting area within linguistics it will therefore be relevant not just for semantics specialists, but for most linguistic audiences. The first encyclopedia ever published in this fascinating and diverse field Combines the talents of the world's leading semantics specialists The latest trends in the field authoritatively reviewed and interpreted in context of related disciplines Drawn from the richest, most authoritative, comprehensive and internationally acclaimed reference resource in the linguistics area Compact and affordable single volume reference format

The Fourth Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2000) was held at the Keihanna-Plaza, Kyoto, Japan, April 18 - 20, 2000. PAKDD 2000 provided an international forum for researchers and application developers to share their original research results and practical development experiences. A wide range of current KDD topics were covered including machine learning, databases, statistics, knowledge acquisition, data visualization, knowledge-based systems, soft computing, and high performance computing. It followed the success of PAKDD 97 in Singapore, PAKDD 98 in Australia, and PAKDD 99 in China by bringing together participants from universities, industry, and government from all over the world to exchange problems and challenges and to disseminate the recently developed KDD techniques. This PAKDD 2000 proceedings volume addresses both current issues and novel approaches in regards to theory, methodology, and real world application. The technical sessions were organized according to subtopics such as Data Mining Theory, Feature Selection and Transformation, Clustering, Application of Data Mining, Association Rules, Induction, Text Mining, Web and Graph Mining. Of the 116 worldwide submissions, 33 regular papers and 16 short papers were accepted for presentation at the conference and included in this volume. Each submission was critically reviewed by two to four program committee members based on their relevance, originality, quality, and clarity.

The significance of big data can be observed in any decision-making process as it is often used for forecasting and predictive analytics. Additionally, big data can be used to build a holistic view of an enterprise through a collection and analysis of large data sets retrospectively. As the data deluge deepens, new methods for analyzing, comprehending, and making use of big data become necessary. Enterprise Big Data Engineering, Analytics, and Management presents novel methodologies and practical approaches to engineering, managing, and analyzing large-scale data sets with a focus on enterprise applications and implementation. Featuring essential big data concepts including data mining, artificial intelligence, and information extraction, this publication provides a platform for retargeting the current research available in the field. Data analysts, IT professionals, researchers, and graduate-level students will find the timely research presented in this publication essential to furthering their knowledge in the field.

The book presents high quality papers presented at 2nd International Conference on Intelligent Computing, Communication & Devices (ICCD 2016) organized by Interscience Institute of Management and Technology (IIMT), Bhubaneswar, Odisha, India, during 13 and 14 August, 2016. The book covers all dimensions of intelligent sciences in its three tracks, namely, intelligent computing, intelligent communication and intelligent devices. intelligent computing track covers areas such as intelligent and distributed computing, intelligent grid and cloud computing,

internet of things, soft computing and engineering applications, data mining and knowledge discovery, semantic and web technology, hybrid systems, agent computing, bioinformatics, and recommendation systems. Intelligent communication covers communication and network technologies, including mobile broadband and all optical networks that are the key to groundbreaking inventions of intelligent communication technologies. This covers communication hardware, software and networked intelligence, mobile technologies, machine-to-machine communication networks, speech and natural language processing, routing techniques and network analytics, wireless ad hoc and sensor networks, communications and information security, signal, image and video processing, network management, and traffic engineering. And finally, the third track intelligent device deals with any equipment, instrument, or machine that has its own computing capability. As computing technology becomes more advanced and less expensive, it can be built into an increasing number of devices of all kinds. The intelligent device covers areas such as embedded systems, RFID, RF MEMS, VLSI design and electronic devices, analog and mixed-signal IC design and testing, MEMS and microsystems, solar cells and photonics, nanodevices, single electron and spintronics devices, space electronics, and intelligent robotics.

This text covers the technologies of document retrieval, information extraction, and text categorization in a way which highlights commonalities in terms of both general principles and practical concerns. It assumes some mathematical background on the part of the reader, but the chapters typically begin with a non-mathematical account of the key issues. Current research topics are covered only to the extent that they are informing current applications; detailed coverage of longer term research and more theoretical treatments should be sought elsewhere. There are many pointers at the ends of the chapters that the reader can follow to explore the literature. However, the book does maintain a strong emphasis on evaluation in every chapter both in terms of methodology and the results of controlled experimentation.

This book constitutes the refereed proceedings of the 16th International Conference on Applications of Natural Language to Information Systems, held in Alicante, Spain, in June 2011. The 11 revised full papers and 11 revised short papers presented together with 23 poster papers, 1 invited talk and 6 papers of the NLDB 2011 doctoral symposium were carefully reviewed and selected from 74 submissions. The papers address all aspects of Natural Language Processing related areas and present current research on topics such as natural language in conceptual modeling, NL interfaces for data base querying/retrieval, NL-based integration of systems, large-scale online linguistic resources, applications of computational linguistics in information systems, management of textual databases NL on data warehouses and data mining, NLP applications, as well as NL and ubiquitous computing.

This 41st volume covers Application of Bayesian Belief Networks to Highway Construction to Virtual Reality Software and Technology.

Organizations generate and collect large volumes of textual data. Unfortunately, many companies are unable to capitalize fully on the value of this data because information implicit within it is not easy to discern. Primarily intended for business analysts and statisticians across multiple industries, this book provides an introduction to the types of problems encountered and current available text mining solutions.

2.1 Text Summarization “Text summarization is the process of distilling the most important information from a source (or sources) to produce an abridged version for a particular user (or users) and task (or tasks)” [3]. Basic and classical articles in text summarization appear in “Advances in automatic text summarization” [3]. A literature survey on information extraction and text summarization is given by Zechner [7]. In general, the process of automatic text summarization is divided into three stages: (1) analysis of the given text, (2) summarization of the text, (3) presentation of the summary in a suitable output form. Titles, abstracts and keywords are the most common summaries in Academic papers. Usually, the title, the abstract and the keywords are the first, second, and third parts of an Academic paper, respectively. The title usually describes the main issue discussed in the study and the abstract presents the reader a short description of the background, the study and its results. A keyword is either a single word (unigram), e.g.: ‘learning’, or a collocation, which means a group of two or more words, representing an important concept, e.g.: ‘machine learning’, ‘natural language processing’. Retrieving collocations from text was examined by Smadja [5] and automatic extraction of collocations was examined by Kita et al. [1].

In a globalized society, effective communication is critical, and study of language from a mathematical perspective can shed light on new ways in which to express meaning across cultures and nations. Computational Linguistics: Concepts, Methodologies, Tools, and Applications explores language by dissecting the phonemic aspects of various communication systems in order to identify similarities and pitfalls in the expression of meaning. With applications in a variety of areas, from psycholinguistics and cognitive science to computer science and artificial intelligence, this multivolume reference work will be of use to researchers, professionals, and educators on the cutting edge of language acquisition and communication science.

Karen Spärck Jones is one of the major figures of 20th century and early 21st Century computing and information processing. Her ideas have had an important influence on the development of Internet Search Engines. Her contribution has been recognized by awards from the natural language processing, information retrieval and artificial intelligence communities, including being asked to present the prestigious Grace Hopper lecture. She continues to be an active and influential researcher. Her contribution to the scientific evaluation of the effectiveness of such computer systems has been quite outstanding. This book celebrates the life and work of Karen Spärck Jones in her seventieth year. It consists of fifteen new and original chapters written by leading international authorities reviewing the state of the art and her influence in the areas in which Karen Spärck Jones has been active. Although she has a publication record which goes back over forty years, it is clear even the very early work reviewed in the book can be read with profit by those working on recent developments in information processing like bioinformatics and the semantic web.

[Copyright: 6d3af5bf09f352be7cc1ad3fee67ca7b](#)