

Automatic Text Processing The Transformation Analysis And Retrieval Of Information By Computer Addison Wesley Series In Computer Science

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.

The increasing importance of multilingualism in language industries, brought about by global markets and world-wide information exchange, parallel corpora, i.e. corpora of texts accompanied by their translation, have become key resources in the development of natural language processing tools. The applications based upon parallel corpora are numerous and growing in number: multilingual lexicography and terminology, machine and human translation, cross-language information retrieval, language learning, etc. The book's chapters have been commissioned from major figures in the field of parallel corpus building and exploitation, with the aim of showing the state of the art in parallel text alignment and use ten to fifteen years after the first parallel-text alignment techniques were developed. Within the book, the following broad themes are addressed: (i) techniques for the alignment of parallel texts at various levels such as sentence, clause, and word; (ii) the use of parallel texts in fields as diverse as translation, lexicography, and information retrieval; (iii) available corpus resources and the evaluation of alignment methods. The book will be of interest to researchers and advanced students of computational linguistics, terminology, lexicography and translation, both in academia and industry.

The International Conference on Intelligent Computing (ICIC) was set up as an annual forum dedicated to emerging and challenging topics in the various aspects of advances in computational intelligence fields, such as artificial intelligence, machine learning, bioinformatics, and computational biology, etc. The goal of this conference was to bring together researchers from academia and industry as well as practitioners to share ideas, problems and solutions related to the multifaceted aspects of intelligent computing. This book constitutes the proceedings of the International Conference on Intelligent Computing (ICIC 2005), held in Hefei, Anhui, China, during August 23-26, 2005. ICIC 2005 received over 2000 submissions from authors in 39 countries and regions. Based on rigorous peer reviews, the Program Committee selected 563 high-quality papers for presentation at ICIC 2005; of these, 215 papers were published in this book organized into 9 categories, and the other 348 papers were published in five international journals. The organizers of ICIC 2005 made great efforts to ensure the success of this conference. We here thank the members of the ICIC 2005 Advisory Committee for their guidance and advice, the members of the Program Committee and the referees for reviewing the papers, and the members of the Publication Committee for checking and compiling the papers. We would also like to thank the publisher, Springer, for their support in publishing the proceedings in the Lecture Notes in Computer Science series. Particularly, we would like to thank all the authors for contributing their papers.

The amounts of information that are flooding people both at the workplace and in private life have increased dramatically in the past ten years. The number of paper documents doubles every four years, and the amount of information stored on all data carriers every six years. New knowledge, however, increases at a considerably lower rate. Possibilities for automatic content recognition in various media and for the processing of documents are therefore becoming more important every day. Especially in economic terms, the efficient handling of information, i.e., finding the right information at the right time, is an invaluable resource for any enterprise, but it is particularly important for small- and medium-sized enterprises. The market for document management systems, which in Europe had a volume of approximately 5 billion euros in 2000, will increase considerably over the next few years. The BMBF recognized this development at an early stage. As early as in 1995, it pooled national capabilities in this field in order to support research on the automatic processing of information within the framework of a large collaborative project (READ) involving both industrial companies and research centres. Evaluation of the results led to the conclusion that research work had been successful, and, in a second phase, funding was provided for the collaborative follow-up project Adaptive READ from 1999 to 2003. The completion of these two important long-term research projects has contributed substantially to improving the possibilities of content recognition and processing of handwritten, printed and electronic documents.

Automatic Text Processing

Alignment and Use of Translation Corpora

Proceedings, October 1996-October 1998

Natural Language Processing – IJCNLP 2005

Natural Language Processing and Information Systems

Encyclopedia of Language and Linguistics

Automatic Indexing and Abstracting of Document Texts

Held in Gaithersburg, MD, Nov. 4-6, 1992. Evaluates new technologies in information retrieval. Numerous graphs, tables and charts.

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.

This book teaches the principles of natural language processing and covers linguistics issues. It also details the language-processing functions involved, including part-of-speech tagging using rules and stochastic techniques. A key feature of the book is the author's hands-on approach throughout, with extensive exercises, sample code in Prolog and Perl, and a detailed introduction to Prolog. The book is suitable for researchers and students of natural language processing and computational linguistics.

With the rising importance of multilingualism in language industries, brought about by global markets and world-wide information exchange, parallel corpora, i.e. corpora of texts accompanied by their translation, have become key resources in the development of natural language processing tools. The applications based upon parallel corpora are numerous and growing in number: multilingual lexicography and terminology, machine and human translation, cross-language information retrieval, language learning, etc. The book's chapters have been commissioned from major figures in the field of parallel corpus building and exploitation, with the aim of showing the state of the art in parallel text alignment and use ten to fifteen years after the first parallel-text alignment techniques were developed. Within the book, the following broad themes are addressed: (i) techniques for the alignment of parallel texts at various levels such as sentence, clause, and word; (ii) the use of parallel texts in fields as diverse as translation, lexicography, and information retrieval; (iii) available corpus resources and the evaluation of alignment methods. The book will be of interest to researchers and advanced students of computational linguistics, terminology, lexicography and translation, both in academia and industry.

The International Conference on Intelligent Computing (ICIC) was set up as an annual forum dedicated to emerging and challenging topics in the various aspects of advances in computational intelligence fields, such as artificial intelligence, machine learning, bioinformatics, and computational biology, etc. The goal of this conference was to bring together researchers from academia and industry as well as practitioners to share ideas, problems and solutions related to the multifaceted aspects of intelligent computing. This book constitutes the proceedings of the International Conference on Intelligent Computing (ICIC 2005), held in Hefei, Anhui, China, during August 23-26, 2005. ICIC 2005 received over 2000 submissions from authors in 39 countries and regions. Based on rigorous peer reviews, the Program Committee selected 563 high-quality papers for presentation at ICIC 2005; of these, 215 papers were published in this book organized into 9 categories, and the other 348 papers were published in five international journals. The organizers of ICIC 2005 made great efforts to ensure the success of this conference. We here thank the members of the ICIC 2005 Advisory Committee for their guidance and advice, the members of the Program Committee and the referees for reviewing the papers, and the members of the Publication Committee for checking and compiling the papers. We would also like to thank the publisher, Springer, for their support in publishing the proceedings in the Lecture Notes in Computer Science series. Particularly, we would like to thank all the authors for contributing their papers.

The amounts of information that are flooding people both at the workplace and in private life have increased dramatically in the past ten years. The number of paper documents doubles every four years, and the amount of information stored on all data carriers every six years. New knowledge, however, increases at a considerably lower rate. Possibilities for automatic content recognition in various media and for the processing of documents are therefore becoming more important every day. Especially in economic terms, the efficient handling of information, i.e., finding the right information at the right time, is an invaluable resource for any enterprise, but it is particularly important for small- and medium-sized enterprises. The market for document management systems, which in Europe had a volume of approximately 5 billion euros in 2000, will increase considerably over the next few years. The BMBF recognized this development at an early stage. As early as in 1995, it pooled national capabilities in this field in order to support research on the automatic processing of information within the framework of a large collaborative project (READ) involving both industrial companies and research centres. Evaluation of the results led to the conclusion that research work had been successful, and, in a second phase, funding was provided for the collaborative follow-up project Adaptive READ from 1999 to 2003. The completion of these two important long-term research projects has contributed substantially to improving the possibilities of content recognition and processing of handwritten, printed and electronic documents.

Automatic Text Processing

Alignment and Use of Translation Corpora

Proceedings, October 1996-October 1998

Natural Language Processing – IJCNLP 2005

Natural Language Processing and Information Systems

Encyclopedia of Language and Linguistics

Automatic Indexing and Abstracting of Document Texts

Held in Gaithersburg, MD, Nov. 4-6, 1992. Evaluates new technologies in information retrieval. Numerous graphs, tables and charts.

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.

This book teaches the principles of natural language processing and covers linguistics issues. It also details the language-processing functions involved, including part-of-speech tagging using rules and stochastic techniques. A key feature of the book is the author's hands-on approach throughout, with extensive exercises, sample code in Prolog and Perl, and a detailed introduction to Prolog. The book is suitable for researchers and students of natural language processing and computational linguistics.

With the rising importance of multilingualism in language industries, brought about by global markets and world-wide information exchange, parallel corpora, i.e. corpora of texts accompanied by their translation, have become key resources in the development of natural language processing tools. The applications based upon parallel corpora are numerous and growing in number: multilingual lexicography and terminology, machine and human translation, cross-language information retrieval, language learning, etc. The book's chapters have been commissioned from major figures in the field of parallel corpus building and exploitation, with the aim of showing the state of the art in parallel text alignment and use ten to fifteen years after the first parallel-text alignment techniques were developed. Within the book, the following broad themes are addressed: (i) techniques for the alignment of parallel texts at various levels such as sentence, clause, and word; (ii) the use of parallel texts in fields as diverse as translation, lexicography, and information retrieval; (iii) available corpus resources and the evaluation of alignment methods. The book will be of interest to researchers and advanced students of computational linguistics, terminology, lexicography and translation, both in academia and industry.

The International Conference on Intelligent Computing (ICIC) was set up as an annual forum dedicated to emerging and challenging topics in the various aspects of advances in computational intelligence fields, such as artificial intelligence, machine learning, bioinformatics, and computational biology, etc. The goal of this conference was to bring together researchers from academia and industry as well as practitioners to share ideas, problems and solutions related to the multifaceted aspects of intelligent computing. This book constitutes the proceedings of the International Conference on Intelligent Computing (ICIC 2005), held in Hefei, Anhui, China, during August 23-26, 2005. ICIC 2005 received over 2000 submissions from authors in 39 countries and regions. Based on rigorous peer reviews, the Program Committee selected 563 high-quality papers for presentation at ICIC 2005; of these, 215 papers were published in this book organized into 9 categories, and the other 348 papers were published in five international journals. The organizers of ICIC 2005 made great efforts to ensure the success of this conference. We here thank the members of the ICIC 2005 Advisory Committee for their guidance and advice, the members of the Program Committee and the referees for reviewing the papers, and the members of the Publication Committee for checking and compiling the papers. We would also like to thank the publisher, Springer, for their support in publishing the proceedings in the Lecture Notes in Computer Science series. Particularly, we would like to thank all the authors for contributing their papers.

The amounts of information that are flooding people both at the workplace and in private life have increased dramatically in the past ten years. The number of paper documents doubles every four years, and the amount of information stored on all data carriers every six years. New knowledge, however, increases at a considerably lower rate. Possibilities for automatic content recognition in various media and for the processing of documents are therefore becoming more important every day. Especially in economic terms, the efficient handling of information, i.e., finding the right information at the right time, is an invaluable resource for any enterprise, but it is particularly important for small- and medium-sized enterprises. The market for document management systems, which in Europe had a volume of approximately 5 billion euros in 2000, will increase considerably over the next few years. The BMBF recognized this development at an early stage. As early as in 1995, it pooled national capabilities in this field in order to support research on the automatic processing of information within the framework of a large collaborative project (READ) involving both industrial companies and research centres. Evaluation of the results led to the conclusion that research work had been successful, and, in a second phase, funding was provided for the collaborative follow-up project Adaptive READ from 1999 to 2003. The completion of these two important long-term research projects has contributed substantially to improving the possibilities of content recognition and processing of handwritten, printed and electronic documents.

Automatic Text Processing

Alignment and Use of Translation Corpora

Proceedings, October 1996-October 1998

Natural Language Processing – IJCNLP 2005

Natural Language Processing and Information Systems

Encyclopedia of Language and Linguistics

Automatic Indexing and Abstracting of Document Texts

Held in Gaithersburg, MD, Nov. 4-6, 1992. Evaluates new technologies in information retrieval. Numerous graphs, tables and charts.

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.

This book teaches the principles of natural language processing and covers linguistics issues. It also details the language-processing functions involved, including part-of-speech tagging using rules and stochastic techniques. A key feature of the book is the author's hands-on approach throughout, with extensive exercises, sample code in Prolog and Perl, and a detailed introduction to Prolog. The book is suitable for researchers and students of natural language processing and computational linguistics.

With the rising importance of multilingualism in language industries, brought about by global markets and world-wide information exchange, parallel corpora, i.e. corpora of texts accompanied by their translation, have become key resources in the development of natural language processing tools. The applications based upon parallel corpora are numerous and growing in number: multilingual lexicography and terminology, machine and human translation, cross-language information retrieval, language learning, etc. The book's chapters have been commissioned from major figures in the field of parallel corpus building and exploitation, with the aim of showing the state of the art in parallel text alignment and use ten to fifteen years after the first parallel-text alignment techniques were developed. Within the book, the following broad themes are addressed: (i) techniques for the alignment of parallel texts at various levels such as sentence, clause, and word; (ii) the use of parallel texts in fields as diverse as translation, lexicography, and information retrieval; (iii) available corpus resources and the evaluation of alignment methods. The book will be of interest to researchers and advanced students of computational linguistics, terminology, lexicography and translation, both in academia and industry.

The International Conference on Intelligent Computing (ICIC) was set up as an annual forum dedicated to emerging and challenging topics in the various aspects of advances in computational intelligence fields, such as artificial intelligence, machine learning, bioinformatics, and computational biology, etc. The goal of this conference was to bring together researchers from academia and industry as well as practitioners to share ideas, problems and solutions related to the multifaceted aspects of intelligent computing. This book constitutes the proceedings of the International Conference on Intelligent Computing (ICIC 2005), held in Hefei, Anhui, China, during August 23-26, 2005. ICIC 2005 received over 2000 submissions from authors in 39 countries and regions. Based on rigorous peer reviews, the Program Committee selected 563 high-quality papers for presentation at ICIC 2005; of these, 215 papers were published in this book organized into 9 categories, and the other 348 papers were published in five international journals. The organizers of ICIC 2005 made great efforts to ensure the success of this conference. We here thank the members of the ICIC 2005 Advisory Committee for their guidance and advice, the members of the Program Committee and the referees for reviewing the papers, and the members of the Publication Committee for checking and compiling the papers. We would also like to thank the publisher, Springer, for their support in publishing the proceedings in the Lecture Notes in Computer Science series. Particularly, we would like to thank all the authors for contributing their papers.

The amounts of information that are flooding people both at the workplace and in private life have increased dramatically in the past ten years. The number of paper documents doubles every four years, and the amount of information stored on all data carriers every six years. New knowledge, however, increases at a considerably lower rate. Possibilities for automatic content recognition in various media and for the processing of documents are therefore becoming more important every day. Especially in economic terms, the efficient handling of information, i.e., finding the right information at the right time, is an invaluable resource for any enterprise, but it is particularly important for small- and medium-sized enterprises. The market for document management systems, which in Europe had a volume of approximately 5 billion euros in 2000, will increase considerably over the next few years. The BMBF recognized this development at an early stage. As early as in 1995, it pooled national capabilities in this field in order to support research on the automatic processing of information within the framework of a large collaborative project (READ) involving both industrial companies and research centres. Evaluation of the results led to the conclusion that research work had been successful, and, in a second phase, funding was provided for the collaborative follow-up project Adaptive READ from 1999 to 2003. The completion of these two important long-term research projects has contributed substantially to improving the possibilities of content recognition and processing of handwritten, printed and electronic documents.

Automatic Text Processing

Alignment and Use of Translation Corpora

Proceedings, October 1996-October 1998

Natural Language Processing – IJCNLP 2005

Natural Language Processing and Information Systems

Encyclopedia of Language and Linguistics

Automatic Indexing and Abstracting of Document Texts

Held in Gaithersburg, MD, Nov. 4-6, 1992. Evaluates new technologies in information retrieval. Numerous graphs, tables and charts.

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.

This book teaches the principles of natural language processing and covers linguistics issues. It also details the language-processing functions involved, including part-of-speech tagging using rules and stochastic techniques. A key feature of the book is the author's hands-on approach throughout, with extensive exercises, sample code in Prolog and Perl, and a detailed introduction to Prolog. The book is suitable for researchers and students of natural language processing and computational linguistics.

With the rising importance of multilingualism in language industries, brought about by global markets and world-wide information exchange, parallel corpora, i.e. corpora of texts accompanied by their translation, have become key resources in the development of natural language processing tools. The applications based upon parallel corpora are numerous and growing in number: multilingual lexicography and terminology, machine and human translation, cross-language information retrieval, language learning, etc. The book's chapters have been commissioned from major figures in the field of parallel corpus building and exploitation, with the aim of showing the state of the art in parallel text alignment and use ten to fifteen years after the first parallel-text alignment techniques were developed. Within the book, the following broad themes are addressed: (i) techniques for the alignment of parallel texts at various levels such as sentence, clause, and word; (ii) the use of parallel texts in fields as diverse as translation, lexicography, and information retrieval; (iii) available corpus resources and the evaluation of alignment methods. The book will be of interest to researchers and advanced students of computational linguistics, terminology, lexicography and translation, both in academia and industry.

The International Conference on Intelligent Computing (ICIC) was set up as an annual forum dedicated to emerging and challenging topics in the various aspects of advances in computational intelligence fields, such as artificial intelligence, machine learning, bioinformatics, and computational biology, etc. The goal of this conference was to bring together researchers from academia and industry as well as practitioners to share ideas, problems and solutions related to the multifaceted aspects of intelligent computing. This book constitutes the proceedings of the International Conference on Intelligent Computing (ICIC 2005), held in Hefei, Anhui, China, during August 23-26, 2005. ICIC 2005 received over 2000 submissions from authors in 39 countries and regions. Based on rigorous peer reviews, the Program Committee selected 563 high-quality papers for presentation at ICIC 2005; of these, 215 papers were published in this book organized into 9 categories, and the other 348 papers were published in five international journals. The organizers of ICIC 2005 made great efforts to ensure the success of this conference. We here thank the members of the ICIC 2005 Advisory Committee for their guidance and advice, the members of the Program Committee and the referees for reviewing the papers, and the members of the Publication Committee for checking and compiling the papers. We would also like to thank the publisher, Springer, for their support in publishing the proceedings in the Lecture Notes in Computer Science series. Particularly, we would like to thank all the authors for contributing their papers.

The amounts of information that are flooding people both at the workplace and in private life have increased dramatically in the past ten years. The number of paper documents doubles every four years, and the amount of information stored on all data carriers every six years. New knowledge, however, increases at a considerably lower rate. Possibilities for automatic content recognition in various media and for the processing of documents are therefore becoming more important every day. Especially in economic terms, the efficient handling of information, i.e., finding the right information at the right time, is an invaluable resource for any enterprise, but it is particularly important for small- and medium-sized enterprises. The market for document management systems, which in Europe had a volume of approximately 5 billion euros in 2000, will increase considerably over the next few years. The BMBF recognized this development at an early stage. As early as in 1995, it pooled national capabilities in this field in order to support research on the automatic processing of information within the framework of a large collaborative project (READ) involving both industrial companies and research centres. Evaluation of the results led to the conclusion that research work had been successful, and, in a second phase, funding was provided for the collaborative follow-up project Adaptive READ from 1999 to 2003. The completion of these two important long-term research projects has contributed substantially to improving the possibilities of content recognition and processing of handwritten, printed and electronic documents.

Automatic Text Processing

Alignment and Use of Translation Corpora

Proceedings, October 1996-October 1998

Natural Language Processing – IJCNLP 2005

Natural Language Processing and Information Systems

Encyclopedia of Language and Linguistics

Automatic Indexing and Abstracting of Document Texts

Held in Gaithersburg, MD, Nov. 4-6, 1992. Evaluates new technologies in information retrieval. Numerous graphs, tables and charts.

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.

This book teaches the principles of natural language processing and covers linguistics issues. It also details the language-processing functions involved, including part-of-speech tagging using rules and stochastic techniques. A key feature of the book is the author's hands-on approach throughout, with extensive exercises, sample code in Prolog and Perl, and a detailed introduction to Prolog. The book is suitable for researchers and students of natural language processing and computational linguistics.

With the rising importance of multilingualism in language industries, brought about by global markets and world-wide information exchange, parallel corpora, i.e. corpora of texts accompanied by their translation, have become key resources in the development of natural language processing tools. The applications based upon parallel corpora are numerous and growing in number: multilingual lexicography and terminology, machine and human translation, cross-language information retrieval, language learning, etc. The book's chapters have been commissioned from major figures in the field of parallel corpus building and exploitation, with the aim of showing the state of the art in parallel text alignment and use ten to fifteen years after the first parallel-text alignment techniques were developed. Within the book, the following broad themes are addressed: (i) techniques for the alignment of parallel texts at various levels such as sentence, clause, and word; (ii) the use of parallel texts in fields as diverse as translation, lexicography, and information retrieval; (iii) available corpus resources and the evaluation of alignment methods. The book will be of interest to researchers and advanced students of computational linguistics, terminology, lexicography and translation, both in academia and industry.

The International Conference on Intelligent Computing (ICIC) was set up as an annual forum dedicated to emerging and challenging topics in the various aspects of advances in computational intelligence fields, such as artificial intelligence, machine learning, bioinformatics, and computational biology, etc. The goal of this conference was to bring together researchers from academia and industry as well as practitioners to share ideas, problems and solutions related to the multifaceted aspects of intelligent computing. This book constitutes the proceedings of the International Conference on Intelligent Computing (ICIC 2005), held in Hefei, Anhui, China, during August 23-26, 2005. ICIC 2005 received over 2000 submissions from authors in 39 countries and regions. Based on rigorous peer reviews, the Program Committee selected 563 high-quality papers for presentation at ICIC 2005; of these, 215 papers were published in this book organized into 9 categories, and the other 348 papers were published in five international journals. The organizers of ICIC 2005 made great efforts to ensure the success of this conference. We here thank the members of the ICIC 2005 Advisory Committee for their guidance and advice, the members of the Program Committee and the referees for reviewing the papers, and the members of the Publication Committee for checking and compiling the papers. We would also like to thank the publisher, Springer, for their support in publishing the proceedings in the Lecture Notes in Computer Science series. Particularly, we would like to thank all the authors for contributing their papers.

The amounts of information that are flooding people both at the workplace and in private life have increased dramatically in the past ten years. The number of paper documents doubles every four years, and the amount of information stored on all data carriers every six years. New knowledge, however, increases at a considerably lower rate. Possibilities for automatic content recognition in various media and for the processing of documents are therefore becoming more important every day. Especially in economic terms, the efficient handling of information, i.e., finding the right information at the right time, is an invaluable resource for any enterprise, but it is particularly important for small- and medium-sized enterprises. The market for document management systems, which in Europe had a volume of approximately 5 billion euros in 2000, will increase considerably over the next few years. The BMBF recognized this development at an early stage. As early as in 1995, it pooled national capabilities in this field in order to support research on the automatic processing of information within the framework of a large collaborative project (READ) involving both industrial companies and research centres. Evaluation of the results led to the conclusion that research work had been successful, and, in a second phase, funding was provided for the collaborative follow-up project Adaptive READ from 1999 to 2003. The completion of these two important long-term research projects has contributed substantially to improving the possibilities of content recognition and processing of handwritten, printed and electronic documents.

Automatic Text Processing

Alignment and Use of Translation Corpora

Proceedings, October 1996-October 1998

Natural Language Processing – IJCNLP 2005

Natural Language Processing and Information Systems

Encyclopedia of Language and Linguistics

Automatic Indexing and Abstracting of Document Texts

Held in Gaithersburg, MD, Nov. 4-6, 1992. Evaluates new technologies in information retrieval. Numerous graphs, tables and charts.

This book constitutes the thoroughly refereed proceedings of the Second International Conference on Intelligent Text Processing and Computational Linguistics, CICLing 2001, held in Mexico City, Mexico in February 2001. The 38 revised full papers and 12 short papers presented together with three invited papers were carefully reviewed and selected from 72 submissions. The books is divided in topical sections on computational linguistic theories, semantics, anaphora and reference, disambiguation, translation, text generation, dictionaries and corpora, morphology, parsing techniques, text categorization, information retrieval, and structure identification and text mining.

Need new summary

Connectionist, Statistical and Symbolic Approaches to Learning for Natural Language Processing

6th International Workshop, PROPOR 2003, Faro, Portugal, June 26-27, 2003. Proceedings

Spotting and Discovering Terms Through Natural Language Processing

Parallel Text Processing

Second International Conference, CICLing 2001, Mexico-City, Mexico, February 18-24, 2001. Proceedings

An Introduction to Text Mining

8th ELSNET Summer School, Chios Island, Greece, July 15-30, 2000. Revised Lectures

Online communities generate massive volumes of natural language data and the social sciences continue to learn how to best make use of this new information and the technology available for analyzing it. Text Mining brings together a broad range of contemporary qualitative and quantitative methods to provide strategic and practical guidance on analyzing large text collections. This accessible book, written by a sociologist and a computer scientist, surveys the fast-changing landscape of data sources, programming languages, software packages, and methods of analysis available today. Suitable for novice and experienced researchers alike, the book will help readers use text mining techniques more efficiently and productively. Available with Perusall—an eBook that makes it easier to prepare for class—Perusall is an award-winning eBook platform featuring social annotation tools that allow students and instructors to collaboratively mark up and discuss their SAGE textbook. Backed by research and supported by technological innovations developed at Harvard University, this process of learning through collaborative annotation keeps your students engaged and makes teaching easier and more effective. Learn more.

"This book provides relevant theoretical frameworks and the latest empirical research findings in biomedicine information retrieval as it pertains to linguistic granularity."—Provided by publisher.

This book constitutes the thoroughly refereed proceedings of the 9th International Conference on Computer Supported Education, CSEU 2018, held in Funchal, Madeira, Portugal, in March 2018. The 27 revised full papers were carefully reviewed and selected from 193 submissions. The papers deal with the following topics: new educational environments, best practices and case studies of innovative technology-based learning strategies, institutional policies on computer-supported education including open and distance education.

This book presents recent developments in automatic text analysis. Providing an overview of linguistic modeling, it collects contributions of authors from a multidisciplinary area that focus on the topic of automatic text analysis from different perspectives. It includes chapters on cognitive modeling and visual systems modeling, and contributes to the computational linguistic and information theoretical grounding of automatic text analysis.

First Text Retrieval Conference (TREC-1)

Computational Linguistics and Intelligent Text Processing

Text Mining

Survey of Text Mining

Advances in Automatic Text Summarization

Introduction to Modern Information Retrieval

Text- and Speech-Triggered Information Access

An information retrieval (IR) system is designed to analyze, process and store sources of information and retrieve those that match a particular user's requirements. A bewildering range of techniques is now available to the information professional attempting to successfully retrieve information. It is recognized that today's information professionals need to concentrate their efforts on learning the techniques of computerized IR. However, it is this book's contention that it also benefits them to learn the theory, techniques and tools that constitute the traditional approaches to the organization and processing of information. In fact much of this knowledge may still be applicable in the storage and retrieval of electronic information in digital library environments. The fully revised third edition of this highly regarded textbook has been thoroughly updated to incorporate major changes in this rapidly expanding field since the second edition in 2004, and a complete new chapter on citation indexing has been added. Unique in its scope, the book covers the whole spectrum of information storage and retrieval, including: users of IR and IR options; database technology; bibliographic formats; cataloguing and metadata; subject analysis and representation; automatic indexing and file organization; vocabulary control; abstracts and indexing; searching and retrieval; user-centred models of IR and user interfaces; evaluation of IR systems and evaluation experiments; online and CD-ROM IR; multimedia IR; hypertext and mark-up languages; web IR; intelligent IR; natural language processing and its applications in IR; citation analysis and IR; IR in digital libraries; and trends in IR research. Illustrated with many examples and comprehensively referenced for an international audience, this is an indispensable textbook for students of library and information studies. It is also an invaluable aid for information practitioners wishing to brush up on their skills and keep up to date with the latest techniques.

This volume is a presentation of all methods of legal knowledge representation from the point of view of jurisprudence as well as computer science. A new method of automatic analysis of legal texts is presented in four case studies. Law is seen as an information system with legally formalised information processes. The achieved coverage of legal knowledge in information retrieval systems has to be followed by the next step: conceptual indexing and automatic analysis of texts. Existing approaches of automatic knowledge representations do not have a proper link to the legal language in information systems. The concept-based model for semi-automatic analysis of legal texts provides this necessary connection. The knowledge base of descriptors, context-sensitive rules and meta-rules formalises properly all important passages in the text corpora for automatic analysis. Statistics and self-organising maps give assistance in knowledge acquisition. The result of the analysis is organised with automatically generated hypertext links. Four case studies show the huge potential