

1. Record Nr.	UNINA9910484667103321
Titolo	String processing and information retrieval : 13th international symposium, SPIRE 2006, Glasgow, UK, October 11-13, 2006 : proceedings / / Fabio Crestani, Paolo Ferragina, Mark Sanderson (eds.)
Pubbl/distr/stampa	Berlin ; ; New York, : Springer, c2006
ISBN	3-540-45775-5
Edizione	[1st ed. 2006.]
Descrizione fisica	1 online resource (XIV, 370 p.)
Collana	Lecture notes in computer science, , 0302-9743 ; ; 4209 LNCS sublibrary. SL 1, Theoretical computer science and general issues
Altri autori (Persone)	CrestaniFabio FerraginaPaolo <1969-> SandersonMark
Disciplina	025.04
Soggetti	Text processing (Computer science) Information retrieval
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Bibliographic Level Mode of Issuance: Monograph
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Web Clustering and Text Categorization -- MP-Boost: A Multiple-Pivot Boosting Algorithm and Its Application to Text Categorization -- TreeBoost.MH: A Boosting Algorithm for Multi-label Hierarchical Text Categorization -- Cluster Generation and Cluster Labelling for Web Snippets: A Fast and Accurate Hierarchical Solution -- Principal Components for Automatic Term Hierarchy Building -- Strings -- Computing the Minimum Approximate ?-Cover of a String -- Sparse Directed Acyclic Word Graphs -- On-Line Repetition Detection -- User Behavior -- Analyzing User Behavior to Rank Desktop Items -- The Intention Behind Web Queries -- Web Search Algorithms -- Compact Features for Detection of Near-Duplicates in Distributed Retrieval -- Inverted Files Versus Suffix Arrays for Locating Patterns in Primary Memory -- Efficient Lazy Algorithms for Minimal-Interval Semantics -- Output-Sensitive Autocompletion Search -- Compression -- A Compressed Self-index Using a Ziv-Lempel Dictionary -- Mapping Words into Codewords on PPM -- Correction -- Improving Usability Through Password-Corrective Hashing -- Word-Based Correction for Retrieval of Arabic OCR Degraded Documents -- Information Retrieval

Applications -- A Statistical Model of Query Log Generation -- Using String Comparison in Context for Improved Relevance Feedback in Different Text Media -- A Multiple Criteria Approach for Information Retrieval -- English to Persian Transliteration -- Bio Informatics -- Efficient Algorithms for Pattern Matching with General Gaps and Character Classes -- Matrix Tightness: A Linear-Algebraic Framework for Sorting by Transpositions -- How to Compare Arc-Annotated Sequences: The Alignment Hierarchy -- Web Search Engines -- Structured Index Organizations for High-Throughput Text Querying -- Adaptive Query-Based Sampling of Distributed Collections -- Short Papers -- Dotted Suffix Trees A Structure for Approximate Text Indexing -- Phrase-Based Pattern Matching in Compressed Text -- Discovering Context-Topic Rules in Search Engine Logs -- Incremental Aggregation of Latent Semantics Using a Graph-Based Energy Model -- A New Algorithm for Fast All-Against-All Substring Matching.

---

## Sommario/riassunto

This volume contains the papers presented at the 13th International Symposium on String Processing and Information Retrieval (SPIRE), held October 11-13, 2006, in Glasgow, Scotland. The SPIRE annual symposium provides an opportunity for both new and established researchers to present original contributions to areas such as string processing (dictionary algorithms, text searching, pattern matching, text compression, text mining, natural language processing, and automata-based string processing); information retrieval languages, applications, and evaluation (IR modelling, indexing, ranking and filtering, interface design, visualization, cross-lingual IR systems, multimedia IR, digital libraries, collaborative retrieval, Web-related applications, XML, information retrieval from semi-structured data, text mining, and generation of structured data from text); and interaction of biology and computation (sequencing and applications in molecular biology, evolution and phylogenetics, recognition of genes and regulatory elements, and sequence-driven protein structure prediction). The papers in this volume were selected from 102 papers submitted from over 20 different countries in response to the Call for Papers. A total of 26 submissions were accepted as full papers, yielding an acceptance rate of about 25%. In view of the large number of good-quality submissions the Program Committee decided to accept 5 short papers, that have also been included in the proceedings. SPIRE 2006 also featured two talks by invited speakers: Jamie Callan (Carnegie Mellon University, USA) and Martin Farach-Colton (Rutgers University, USA).

---