

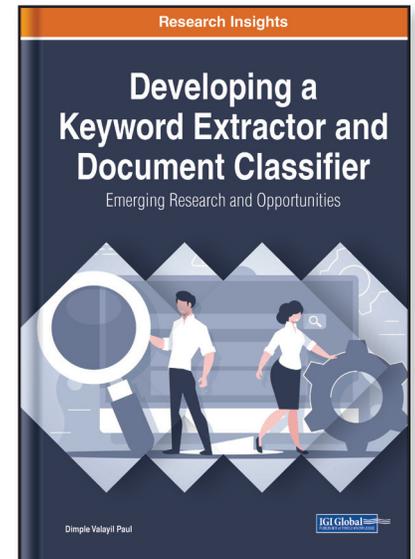
Developing a Keyword Extractor and Document Classifier: Emerging Research and Opportunities

Part of the Advances in Data Mining and Database Management Book Series

Dimple Valayil Paul (Goa University, India)

Description:

The main problems that prevent fast and high-quality document processing in electronic document management systems are insufficient and unstructured information, information redundancy, and the presence of large amounts of undesirable user information. The human factor has a significant impact on the efficiency of document search. An average user is not aware of the advanced option of a query language and uses typical queries. Development of a specialized software toolkit intended for information systems and electronic document management systems can be an effective solution of the tasks listed above. Such toolkits should be based on the means and methods of automatic keyword extraction and text classification. The categorization (or classification) of texts into predefined categories has witnessed a booming interest in the last 10 years due to the increased availability of documents in digital form and the ensuing need to organize them. Thus, research on keyword extraction, advancements in the field, and possible future solutions is of great importance in current times.



Developing a Keyword Extractor and Document Classifier: Emerging Research and Opportunities presents an information extraction mechanism that can process many kinds of inputs, realize the type of text, and understand the percentage of the keywords that has to be stored. This mechanism then supports information extraction and information categorization mechanisms. This module is used to support a text summarization mechanism, which leads—with the help of the keyword extraction module—to text categorization. It employs lexical and information retrieval techniques to extract phrases from the document text that are likely to characterize it and determines the category of the retrieved text to present a summary to the users. This book is ideal for practitioners, stakeholders, researchers, academicians, and students who are interested in the development of a new keyword extractor and document classifier method.

ISBN: 9781799837725

Pages: 300

Copyright: 2021

Release Date: January, 2021

Hardcover: \$195.00

Softcover: \$150.00

E-Book: \$195.00

Hardcover + E-Book: \$235.00

Topics Covered:

Data Mining
Document Classification
Input and Output Design
Keyword Extraction

Menu Design
Performance Measures
System Analysis and Design
System Testing

Text Categorization
Text Mining

Subject: Computer Science and Information Technology

Classification: Research Insights

Readership Level: Advanced-Academic Level (Research Recommended)

Research Suitable for: Advanced Undergraduate Students; Graduate Students; Researchers; Academicians; Professionals; Practitioners

Order Information

Phone: 717-533-8845 x100

Toll Free: 1-866-342-6657

Fax: 717-533-8661 or 717-533-7115

Online Bookstore: www.igi-global.com

Mailing Address: 701 East Chocolate Avenue, Hershey, PA 17033, USA