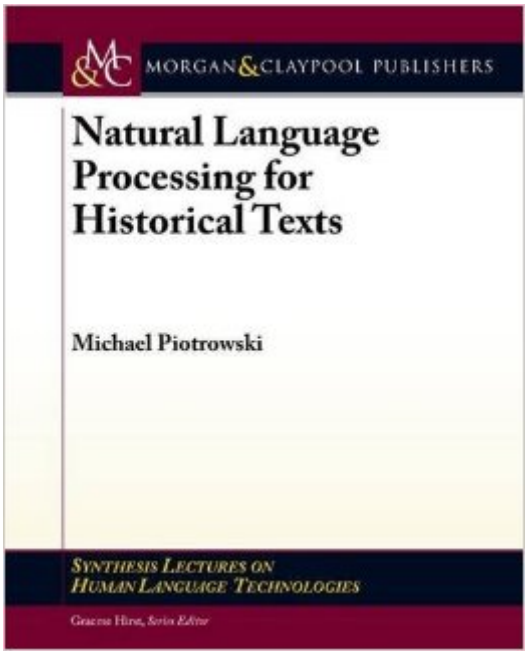
The book was found

Natural Language Processing For Historical Texts (Synthesis Lectures On Human Language Technologies)



PDF
Adobe
DOWNLOAD EBOOK

Synopsis

More and more historical texts are becoming available in digital form. Digitization of paper documents is motivated by the aim of preserving cultural heritage and making it more accessible, both to laypeople and scholars. As digital images cannot be searched for text, digitization projects increasingly strive to create digital text, which can be searched and otherwise automatically processed, in addition to facsimiles. Indeed, the emerging field of digital humanities heavily relies on the availability of digital text for its studies. Together with the increasing availability of historical texts in digital form, there is a growing interest in applying natural language processing (NLP) methods and tools to historical texts. However, the specific linguistic properties of historical texts -- the lack of standardized orthography, in particular -- pose special challenges for NLP. This book aims to give an introduction to NLP for historical texts and an overview of the state of the art in this field. The book starts with an overview of methods for the acquisition of historical texts (scanning and OCR), discusses text encoding and annotation schemes, and presents examples of corpora of historical texts in a variety of languages. The book then discusses specific methods, such as creating part-of-speech taggers for historical languages or handling spelling variation. A final chapter analyzes the relationship between NLP and the digital humanities. Certain recently emerging textual genres, such as SMS, social media, and chat messages, or newsgroup and forum postings share a number of properties with historical texts, for example, nonstandard orthography and grammar, and profuse use of abbreviations. The methods and techniques required for the effective processing of historical texts are thus also of interest for research in other domains. Table of Contents: Introduction / NLP and Digital Humanities / Spelling in Historical Texts / Acquiring Historical Texts / Text Encoding and Annotation Schemes / Handling Spelling Variation / NLP Tools for Historical Languages / Historical Corpora / Conclusion / Bibliography

Book Information

Series: Synthesis Lectures on Human Language Technologies
Paperback: 158 pages
Publisher: Morgan & Claypool Publishers (September 28, 2012)
Language: English
ISBN-10: 1608459462
ISBN-13: 978-1608459469
Product Dimensions: 7.5 x 0.4 x 9.2 inches
Shipping Weight: 13 ounces (View shipping rates and policies)

Average Customer Review: Be the first to review this item

Best Sellers Rank: #3,293,943 in Books (See Top 100 in Books) #79 inÂ Books > Computers & Technology > Software > Voice Recognition #190 inÂ Books > Computers & Technology > Digital Audio, Video & Photography > Speech & Audio Processing #241 inÂ Books > Computers & Technology > Computer Science > AI & Machine Learning > Natural Language Processing

Download to continue reading...

Natural Language Processing for Historical Texts (Synthesis Lectures on Human Language Technologies) Books of Breathing and Related Texts -Late Egyptian Religious Texts in the British Museum Vol.1 (Catalogue of the Books of the Dead and Other Religious Texts in the British Museum) Body Language: Body Language Training - Attract Women & Command Respect, by Mastering Your High Status Body Language (Body Language Attraction, Body Language ... Language Secrets, Nonverbal Communication) An Introduction to Language Processing with Perl and Prolog: 260 (Cognitive Technologies) Human Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, Third Edition (Human Factors and Ergonomics) Lectures on Antitrust Economics (Cairoli Lectures) The Birth of Biopolitics: Lectures at the College de France, 1978-1979 (Lectures at the CollÃƒÂ¨ge de France) The Government of Self and Others: Lectures at the College de France, 1982-1983 (Lectures at the CollÃƒÂ¨ge de France) Lectures on the Will to Know (Michel Foucault, Lectures at the CollÃƒÂ¨ge de France) Security, Territory, Population: Lectures at the College De France, 1977 - 78 (Michel Foucault, Lectures at the CollÃƒÂ¨ge de France) Feynman Lectures Simplified 4A: Math for Physicists (Everyone's Guide to the Feynman Lectures on Physics Book 12) Natural Remedies for Dogs : 101 Safe & Natural Essential Oils' Remedies for Your DOG: (Natural Remedies For Dogs, Essential Oils Remedies For Dogs, Natural Dog Care, Recipes For Dogs, Home Remedies) Introduction to Embedded Systems: Using ANSI C and the Arduino Development Environment (Synthesis Lectures on Digital Circuits and Systems) Narrowband Direction of Arrival Estimation for Antenna Arrays (Synthesis Lectures on Antennas) Circuit Analysis with Multisim (Synthesis Lectures on Digital Circuits and Systems) Mobile Robotics for Multidisciplinary Study (Synthesis Lectures on Control and Mechatronics) Optical Interconnects (Synthesis Lectures on Solid-State Materials and Devices) Essentials of Game Theory: A Concise, Multidisciplinary Introduction (Synthesis Lectures on Artificial Intelligence and Machine Learning) How To Analyze People: Body Language and Human Psychology Guide to Mastering Analyzing (Analyze People, Body Language, Human Psychology, How to Analyze People) The Organic Chemistry of Drug Synthesis, Volume 3 (Organic Chemistry Series of Drug Synthesis)

Dmca