


Axel-Cyrille Ngonga Ngomo · Petr Křemen (Eds.)

Knowledge Engineering and Semantic Web

7th International Conference, KESW 2016
Prague, Czech Republic, September 21–23, 2016
Proceedings

 Springer

Editors
Axel-Cyrille Ngonga Ngomo
Leipzig University
Leipzig
Germany

Petr Křemen
Czech Technical University in Prague
Prague
Czech Republic

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-319-45879-3 ISBN 978-3-319-45880-9 (eBook)
DOI 10.1007/978-3-319-45880-9

Library of Congress Control Number: 2016949634

© Springer International Publishing Switzerland 2016

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Contents

Ontologies

- Multi-viewpoint Ontologies for Decision-Making Support 3
Sergey Gorshkov, Stanislav Kralin, and Maxim Miroshnichenko
- Ontological Anti-patterns in Aviation Safety Event Models 18
Jana Ahmad and Petr Křemen
- User-Driven Ontology Population from Linked Data Sources 31
*Panagiotis Mitziias, Marina Riga, Efstratios Kontopoulos,
Thanos G. Stavropoulos, Stelios Andreadis, Georgios Meditskos,
and Ioannis Kompatsiaris*
- Ontology for Performance Control in Service-Oriented System
with Composite Services 42
*Maksim Khagai, Dmitrii Zubok, Tatiana Kharchenko,
and Alexandr Maiatin*
- Privacy in Online Social Networks: An Ontological Model
for Self-Presentation 56
Javed Ahmed
- Design of an Ontologies for the Exchange of Software Engineering Data
in the Aerospace Industry. 71
Ricardo Eito-Brun

Information and Knowledge Extraction

- Family Matters: Company Relations Extraction from Wikipedia 81
Artem Kuznetsov, Pavel Braslavski, and Vladimir Ivanov
- A Bank Information Extraction System Based on Named Entity
Recognition with CRFs from Noisy Customer Order Texts in Turkish. 93
Erdem Emekligil, Secil Arslan, and Onur Agin
- A New Operationalization of Contrastive Term Extraction Approach Based
on Recognition of Both Representative and Specific Terms 103
*Aliya Nugumanova, Igor Bessmertny, Yerzhan Baiburin,
and Madina Mansurova*
- Ontology-Based Information Extraction for Populating the Intelligent
Scientific Internet Resources. 119
Irina R. Akhmadeeva, Yury A. Zagorulko, and Dmitry I. Mouromtsev

A New Operationalization of Contrastive Term Extraction Approach Based on Recognition of Both Representative and Specific Terms

Aliya Nugumanova^{1(✉)}, Igor Bessmertny², Yerzhan Baiburin⁴,
and Madina Mansurova³

¹ D. Serikbayev East Kazakhstan State Technical University,
Ust-Kamenogorsk, Kazakhstan
yalisha@yandex.kz

² Saint Petersburg National Research University of ITMO,
Saint Petersburg, Russia
igor_bessmertny@gmail.com

³ Al-Farabi Kazakh National University, Almaty, Kazakhstan
mansurova01@mail.ru

⁴ East Kazakhstan State University, Ust-Kamenogorsk, Kazakhstan
ebaiburin@gmail.com

Abstract. A contrastive approach to term extraction is an extensive class of methods based on the assumption that the words frequently occurring within a domain and rarely beyond it are most likely terms. The disadvantage of this approach is a great number of type II errors – false negatives. The cause of these errors is in the idea of contrastive selection when the most representative high frequent terms are extracted from the texts and rare terms are discarded. In this work, we propose a new operationalization of the contrastive approach, which supports the capture of both high frequent and low frequent domain terms. Proposed operationalization reduces the number of false negatives. The experiments performed on the texts of the subject domain “Geology” show promising of proposed approach.

Keywords: Contrastive term extraction · Termhood · Mutual information · LSA

1 Introduction

At present, in the field of information extraction there are numerous methods aimed at automated extraction of knowledge structures from natural language texts [1–4]. Among the knowledge structures being extracted, the simplest ones are lists of terms and the most complex ones are domain thesauri and ontologies. All these structure are designed for setting explicit specifications of subject domain to eliminate uncertainty and ambiguity in the knowledge exchange between humans and applications.

In this work, we focus on extraction of simple but valuable knowledge structures – lists of single word terms. Like the authors of [5], we consider domain terms to be words used by experts to describe conceptual apparatus of the domain. Lists of terms are used everywhere when it is necessary to convey in a structured compressed form