Dmitry I. Ignatov, Sergei O. Kuznetsov, Jonas Poelmans (Eds.)

CDUD 2012 – Concept Discovery in Unstructured Data

Workshop co-located with the 10th International Conference on Formal Concept Analysis (ICFCA 2012)
May 2012, Leuven, Belgium

Volume Editors

Dmitry I. Ignatov
School of Applied Mathematics and Information Science
National Research University Higher School of Economics, Moscow, Russia

Sergei O. Kuznetsov
School of Applied Mathematics and Information Science
National Research University Higher School of Economics, Moscow, Russia

Jonas Poelmans
Faculty of Business and Economics
Katholieke Universiteit Leuven, Belgium

Printed in Belgium by the Katholieke Universiteit Leuven with ISBN 978-9-08-140991-9.

The proceedings are also published online on the CEUR-Workshop website in volume Vol-871 of a series with ISSN 1613-0073.

Preface

Concept discovery is a subarea of Knowledge Discovery in Databases (KDD) where concept models, such as Formal Concept Analysis (FCA), multimodal clustering, conceptual graphs and other, are used for gaining insight into the underlying conceptual structure of data. Traditional machine learning techniques are mainly focusing on structured data given by object-attribute tables, whereas most data available nowadays are given in unstructured, often textual, form. As compared to traditional data mining techniques, human-centered instruments of concept discovery actively engage domain experts in the discovery process.

This volume contains the papers presented at the 2nd International Workshop on Concept Discovery in Unstructured Data (CDUD 2012) held on May 10, 2012 at the Katholieke Universiteit Leuven, Belgium. This workshop welcomes papers describing innovative research on data discovery in complex data. Moreover, this workshop provides a forum for researchers and developers of data mining instruments, working on issues associated with analyzing unstructured data. This year the committee decided to accept 11 papers for publication in the proceedings. Each submission was reviewed by on average 3 program committee members.

A. Mestrovic presents an application of concept lattices to semantic matching in Croatian language. A. Chepovskiy et al. propose a method for automatic language identification for transliterated texts. X. Naidenova describes a novel neural network based data structure for inferring classification tests. A. Kravchenko et al. introduce an approach for expert search which is based on analyzing e-mail communication patterns. D. Ustalov et al. propose an ontology-based approach for text-to-picture synthesis. A. Skabin presents a computerized recognition system for hand-written historical manuscripts. A. Panchenko et al. extract semantic relations between concepts from Wikipedia using KNN algorithms. D. Fedyanin uses parameter identification methods for Markov models and applies them to influence analysis in social networks. S. Milyaev et al. discuss a new method for self-tuning semantic image segmentation. A. Vorobev proposes a probabilistic model for evaluating the quality level of projects, authors and experts in collaborative innovation platforms. D. Gnatyshak et al. present a novel pseudo-triclustering algorithm and applied it to online social network data. A. Bozhenyuk et al. discuss methods for maximum flow and minimum cost flow finding in fuzzy setting.

We would like to express our gratitude to all contributing authors and reviewers. We also want to thank our sponsors Amsterdam-Amstelland police, IBM Belgium, Research Foundation Flanders, Vlerick Management School, OpenConnect Systems and Higher School of Economics (Moscow, Russia). Finally, we should thank the authors of the EasyChair system which helped us to manage the reviewing process.

May 10, 2012
Leuven

Dmitry I. Ignatov
Sergei O. Kuznetsov
Jonas Poelmans

Organization

The 2nd International Workshop on Concept Discovery in Unstructured Data (CDUD 2012) was held on May 10, 2012 at the Katholieke Universiteit Leuven, Belgium. The workshop was co-located with the 10th International Conference on Formal Concept Analysis (ICFCA-2012). The inaugural edition of CDUD was held on June 25, 2011 at the Higher School of Economics in Moscow, Russia.

Program Chairs

Dmitry I. Ignatov	National Research University Higher School of Economics, Russia
Sergei O. Kuznetsov	National Research University Higher School of Economics, Russia
Jonas Poelmans	Katholieke Universiteit Leuven, Belgium

Program Committee

Simon Andrews	Sheffield Hallam University, United Kingdom
Guido Dedene	Katholieke Universteit Leuven, Belgium
Florent Domenach	University of Nicosia, Cyprus
Irina Efimenko	National Research University Higher School of Economics, Russia
Paul Elzinga	Amsterdam-Amstelland Police, The Netherlands
Boris Galitsky	University of Girona, Spain
Bernhard Ganter	Technische Universität Dresden, Germany
Yury Katkov	National Research University of Information Technologies, Mechanics and Optics, Russia
Natalia Loukachevitch	Moscow State University, Russia
Dmitry Mouromtsev	National Research University of Information Technologies, Mechanics and Optics, Russia
Xenia Naidenova	Military Medical Academy, Russia
Alexey A. Neznanov	National Research University Higher School of Economics, Russia
Sergei A. Obiedkov	National Research University Higher School of Economics, Russia
Simon Polovina	Sheffield Hallam University, United Kingdom
Uta Priss	Edinburgh Napier University, United Kingdom
Dominik Slezak	University of Warsaw and Infobright, Poland
Rustam Tagiew	Technische Universität Freiberg, Germany
Stijn Viaene	Katholieke Universiteit Leuven, Belgium

Johanna Voelker	University of Mannheim, Germany
Rostislav Yavorsky	Witology, Russia

Additional Reviewers

Ekaterina Cherniak, National Research University of Higher School of Economics, Russia
Alexandr Vorobev, Moscow State University and Witology, Russia

Sponsoring Institutions

Amsterdam-Amstelland police, The Netherlands
IBM, Belgium
OpenConnect Systems, USA
Research Foundation Flanders, Belgium
Vlerick Management School, Belgium
National Research University Higher of School Economics, Russia

Table of Contents