Call for papers

First International Workshop on Concept Discovery in Unstructured Data (CDUD)

June 25, 2011

Organized by *National Research University – Higher School of Economics, Moscow, Russia*

Homepage: http://rsfdgrc.hse.ru/cdud

Workshop chairs

Dmitry Ignatov, University - Higher School of Economics, Russia
Sergei Kuznetsov, Moscow Higher School of Economics, Russia
Jonas Poelmans, Katholieke Universiteit Leuven Belgium

Program committee

Bernhard Ganter, Technical University of Dresden, Germany
Guido Dedene, Katholieke Universiteit Leuven, Belgium
Stijn Viaene, Vlerick Leuven Management School, Katholieke Universiteit Leuven, Belgium
Henrik Scharfe, Aalborg University, Denmark
Simon Polovina, Sheffield Hallam University, United Kingdom
Richard Hill, University of Derby, United Kingdom
Sergei Obiedkov, University - Higher School of Economics, Russia
Alexey Neznanov, University - Higher School of Economics, Russia
Paul Elzinga, Katholieke Universiteit Leuven, Belgium, Amsterdam-Amstelland police Netherlands
Vladimir Selegey, ABBYY Software, Russia

Workshop theme:

Concept discovery is a Knowledge Discovery in Databases (KDD) research field that uses human-centered techniques such as Formal Concept Analysis (FCA), Biclustering, Triclustering, Conceptual Graphs etc. for gaining insight into the underlying conceptual structure of the data. Traditional machine learning techniques are mainly focusing on structured data whereas most data available resides in unstructured, often textual, form. Compared to traditional data mining techniques, human-centered instruments actively engage the domain expert in the discovery process.

This workshop welcomes papers describing innovative research on data discovery techniques. Moreover, this workshop intends to provide a forum for researchers and developers of data mining instruments, working on issues associated with analyzing unstructured data. First, we are interested in methods for transforming unstructured into semi-structured information. Unstructured information such as texts or images can be tagged, keywords can be extracted from texts by means of Natural Language Processing methods, etc. Second, in this workshop we also

particularly welcome research on using human-centered instruments such as FCA to analyse unstructured and semi-structured data. Applications in which we are interested include but are not limited to Text Mining, Web Mining including forums, blogs, social sharing systems like Twitter and Facebook, mining sociological interviews, etc. We are also interested in innovative instruments for dealing with knowledge incompleteness and asymmetry. Finally, innovations which increase the expressiveness of existing methods for unstructured data analysis such as fuzzy FCA, rough FCA, Latent Semantic Analysis, probabilistic concept discovery are also of interest to this workshop.

Subject coverage:
- Data Mining (Text Mining, Graph Mining, Web Mining, Association Rules, Frequent Closed Sets, etc.)
- Dealing with knowledge incompleteness and asymmetry
- Discovery techniques for conceptual models
- Efficient indexing and structuring algorithms
- Formal Concept Analysis (FCA), Fuzzy FCA and Rough FCA
- Knowledge discovery
- Innovative applications of FCA for discovery purposes
- Natural Language Processing, Latent Semantic Analysis
- Probabilistic concept discovery
- Structured methods for KDD
- Sociological applications of FCA and related methods
- Conceptual Clustering, Biclustering, Triclustering etc.

Important dates
Submission deadline: May 15, 2011
Revisions: May 30, 2011
Camera-ready copy: June 10, 2011

Proceedings: All accepted papers will be included in the workshop's proceedings to be published by U-HSE. Also, the proceedings will be published online on the CEUR-Workshop web site in a volume with ISSN.

Submission Procedure: Electronic version of full paper complete with authors' affiliations should be submitted through the conference electronic submission system.
        Use the submission link http://www.easychair.org/conferences/?conf=cdud2011 .
Manuscripts must be prepared with LaTeX or Microsoft Office and should follow the Springer format available at http://www.springer.de/comp/lncs/authors.html. The maximum number of accepted papers by an individual author that can be covered by the workshop's registration charge is 3. The papers over 12 pages are not allowed.

Location: Vorontsovo Pole Street, 5A, Moscow, Russia