



CALL FOR PAPERS

"4th Edition of the Workshop on Interactions between Data Mining and Natural Language Processing" (DMNLP 2017)

ECML/PKDD 2017 Workshop

September, 2017, Skopje, Macedonia

<https://easychair.org/conferences/?conf=dmnlp17>

** IMPORTANT DATES

Paper submission deadline: **Monday, June 26, 2017**

Author notification: **Friday, July 21, 2017**

Final version: **Friday, July 28, 2017**

** SCOPE

On the one hand, in the field of Natural Language Processing (NLP), numerical Machine Learning methods (e.g., SVM, CRF) have been intensively explored and applied. Despite the good results obtained by the numerical methods, one major drawback is that they do not provide a human readable model. A promising direction is the integration of symbolic knowledge.

On the other hand, research in Data Mining has progressed significantly in the last decades, through the development of advanced algorithms and techniques to extract knowledge from data in different forms. In particular, for two decades Pattern Mining has been one of the most active field in Knowledge Discovery.

Recently, a new field has emerged taking benefit of both domains: Data Mining and NLP. The objective of DMNLP is thus to provide a forum to discuss how Data Mining can be interesting for NLP tasks, providing symbolic knowledge, but also how NLP can enhance data mining approaches by providing richer and/or more complex information to mine and by integrating linguistics knowledge directly in the mining process.

The workshop aims at bringing together researchers from both communities in order to stimulate discussions about the cross-fertilization of those two research fields. The idea of this workshop is to discuss future directions and new challenges emerging from the cross-fertilization of Data Mining and NLP and in the same time initiate collaborations between researchers of both communities.

** CALL FOR PAPERS

The workshop promotes works where the two following dimensions are combined in one as symbiosis. The first dimension is Data Mining, for instance Pattern Mining (itemsets, sequences, trees, graphs, association rules), classification (decision trees, FCA,...), inductive logic programming. The second dimension is NLP, for example question/answering systems, translation, information extraction, linguistic analysis (lexical analysis, terminology, syntax, semantics, discourse, stylistics), classification, knowledge extraction/ontology building from texts, information retrieval, corpus annotation, social/opinion mining.

A list of non-exhaustive topics that fit the scope of the workshop is thus:

- Pattern discovery for NLP
- Constraint-based pattern mining in text
- Data mining query languages for expressing NLP tasks
- Data representation (sequence, tree, graphs) for NLP
- Text modeling for Data Mining
- Mining complex data for / coming from NLP

- Representation learning in NLP (e.g., deep learning, compositional models...)
- Relationships between data mining and NLP
- Modeling and visualizing data mining results on text
- Integrating NLP characteristics in data mining
- Data mining approaches for linguistic knowledge building
- Knowledge discovery for linguistic analysis (e.g. stylistics, socio-linguistics,...)
- Knowledge-based data mining for NLP (e.g., ontologies, linked open data...)

Our main goal is to stimulate discussion, collaboration and the sharing of experiences. In that respect, we would have three submission types:

- unpublished works (max **16 pages**, double submissions allowed)
- extended abstracts and vision statements (max **6 pages**)
- recently published works (special oral-only track, **no page limits**)

For more details on submission, see the DMNLP webpage at <http://dmnlp.loria.fr>