Call for Papers: 2nd Workshop on Mining Unstructured Data (MUD'12)

Co-located with the 19th Working Conference on Reverse Engineering (WCRE'12)

Kingston, Ontario, Canada, October 15th – 18th, 2012.

SCOPE

Today's software developers are supported by a variety of tools, such as version control systems like GIT, Issue Tracking Systems like BugZilla, and mailing list services. As such we have a wide range of information recorded in the repositories in which these tools store their data. These repositories comprise two types of data: **Structured Data** and **Unstructured Data**. Structured data, such as source code or execution traces, has a well-established structure and grammar, and is straightforward to parse and use with computer machinery. Unstructured data, such as documentation, discussions, comments, customer support requests, consists of a mixture of natural language text, snippets of structured data, and noise. **Mining unstructured data** poses many hard challenges, since out-of-the box approaches adopted from related fields such as Natural Language Processing (NLP) and Information Retrieval (IR) cannot be directly applied in the Software Engineering domain. **The goal of this half-day workshop is to address these challenges and aims to make the knowledge contained in unstructured data repositories accessible to practitioners and researchers.**

TOPICS OF INTEREST

We seek contributions related to techniques and applications of mining unstructured data. Contributions would ideally, but not limited to, address one of the following topics:

- Tools and methods for extracting unstructured data from software development repositories.
- Adaptions of NLP and IR approaches to the Software Engineering domain.
- Lessons learned and pitfalls of working with unstructured data and potential solutions.
- Analysis of unstructured data embedded within structured data (e.g., comments in source code).
- Unstructured data for linking of software development artifacts and traceability.
- Usage of unstructured data for software evolution tasks, such as Feature and Concept Location, evolution of unstructured data over time, or extraction of semantics from software development repositories.
- Success and failure stories of mining unstructured data with a critical dissemination of reasons for the specific success or failure, and future steps.

SUBMISSIONS

We invite high-quality contributions to this workshop in the form of short (4 page) papers. Submissions should address specific tools, techniques, applications, stories of success or failure, adaptions of techniques from other domains, and lessons learned in the domain of Mining Unstructured Data. Papers must not have been accepted elsewhere for publication or be under review for another workshop, conference, or journal. Contributions are required to be formatted in the IEEE Computing Society conference format. For detailed formatting and submission guidelines, please visit the WCRE'12 website at http://www.distat.unimol.it/wcre2012/. Top-quality papers that are accepted for presentation at the workshop will be recommended to the Special Issue of Electronic Communications of the ECEASST journal (http://journal.ub.tu-berlin.de/eceasst).

IMPORTANT DATES

September 21st, 2012	Workshop paper submissions
September 25th, 2012	Notifications of acceptance
October 3rd, 2012	Camera-Ready version due
October 17th, 2012	Presentation of research at the workshop

WORKSHOP ORGANIZERS

Alberto Bacchelli	University of Lugano, Switzerland	http://www.inf.usi.ch/phd/bacchelli/
Nicolas Bettenburg	Queen's University, Canada	http://nicolas-bettenburg.com
Latifa Guerrouj	Ecole Polytechnique de Montreal	http://web.soccerlab.polymtl.ca

ADDITIONAL INFORMATION

For more information on the MUD'12 workshop, visit our website at http://sailhome.cs.queensu.ca/mud or contact the workshop organizers.

See you soon at MUD'12 in Kingston, Ontario! Alberto, Nicolas and Latifa.