

## LREC 2010 Workshop CFP

### **Resources and Evaluation for Entity Resolution and Entity Management**

This workshop, open to all LREC participants, builds on the success of the workshop held at LREC08 on Resources and Evaluation for Identity Matching, Entity Resolution and Entity Management. We hope to leverage results achieved since that workshop and advance the state of the art.

There is currently great interest in the extraction and management of structured information from unstructured text – a need that continues to grow with ever increasing production of unstructured information. Much of the information extracted focuses on people, and this information must be managed – and searchable – when stored in a structured repository. This workshop will address the intersection of the name/record matching, entity resolution, and entity management problems, with a focus on evaluation methodologies suited to such systems and the language resources necessary both to perform these tasks and to evaluate them.

Whereas entity extraction systems are usually evaluated using MUC methodology, for example, the case is less clear for these “follow-on” technologies. Even a seemingly simple task such as matching person names in a database context is problematic, and although **name** or **record** matching systems can be evaluated using IR evaluation methodologies (à la CLEF or TREC), there are still interesting methodological issues to resolve before we can refer to a “standard” evaluation methodology for name/record matching. Moreover, it is much less clear how to effectively evaluate **identity** matching, resolution, and management systems, or even what it means to perform an effective identity match.

We solicit papers discussing state-of-the-art approaches to name and record matching, identity resolution, and identity management, as well as novel proposals for methodologies to evaluate these technologies. Suggested topics include, but are not limited to:

- Research in name / record matching
- Research in entity disambiguation / resolution / deconfliction
- Descriptions of systems or best practices for entity management
- Descriptors of systems that incorporate entity resolution as an essential component
- Evaluation of record matching, entity disambiguation, or entity management in context
- Resources for identity matching, identity resolution, or identity management
- Resources for evaluation of these technologies

Position papers on the evaluation and integration of identity resolution technologies will also be considered for acceptance.

Although we are distributing a call for papers and there will be papers accepted for presentation at the workshop, this workshop will not be a “mini-conference”. Rather, this workshop will be an interactive and dynamic event. Presentations will serve to introduce topics for discussion and to shape the day. A significant portion of time will be devoted to interactive exercises, brainstorming, and other “work.” The workshop will be organized such that all attendees

(including the organizers) leave the workshop more informed—having made progress toward a contextualized and principled evaluation of identity matching, resolution, and management systems—as well as more connected, having begun interactive discussions that will continue after the conference.

Organizing Committee:

Keith J. Miller (The MITRE Corporation)  
Mark Arehart (The MITRE Corporation)  
Sherri Condon (The MITRE Corporation)  
Jason Duncan (M.I.T. Lincoln Labs)  
Louise Guthrie (University of Sheffield)  
Christopher Dozier (Thomson Reuters)  
Elizabeth Schroeder (The MITRE Corporation)  
Massimo Poesio (Universita' di Trento)  
Emanuele Pianta (Fondazione Bruno Kessler)

Important Dates:

Deadline for Paper Submission: Friday, 12 February 2010  
Notification to Authors: Friday, 26 February 2010  
Submission of Final Version: Friday, 12 March 2010

**Workshop at LREC: Saturday, 22 May 2010**

Submission Information:

Papers should be submitted to the START Conference Manager at <https://www.softconf.com/lrec2010/Entity2010/>. When submitting a paper from the START page, authors will be asked to provide essential information about resources (in a broad sense, i.e. also technologies, standards, evaluation kits, etc.) that have been used for the work described in the paper or are a new result of your research. For further information on this new initiative, please refer to <http://www.lrec-conf.org/lrec2010/?LREC2010-Map-of-Language-Resources>.