

CALL FOR PARTICIPATION

Text Analysis Conference (TAC 2012)

<http://www.nist.gov/tac/2012/KBP/>

Task Evaluations: February - October 2012

Workshop: November 5-6, 2012

Conducted by: **National Institute of Standards and Technology (NIST)**

With support from: **U.S. Department of Defense**

INTRODUCTION

The Text Analysis Conference (TAC) is a series of evaluations and workshops organized to promote research in Natural Language Processing and related applications, by providing a large test collection, common evaluation procedures, and a forum for organizations to share their results. NIST provides test data for each task, and participants run their NLP systems on the data and return their results to NIST for evaluation. The conference culminates in a November workshop at NIST in Gaithersburg, Maryland, USA.

You are invited to participate in TAC 2012. Organizations may choose to participate in any or all of the TAC 2012 tasks. All results submitted to NIST are archived on the TAC web site, and all evaluations of submitted results are included in the conference proceedings. Dissemination of TAC work and results other than in the conference proceedings is welcomed, but the conditions of participation specifically preclude any advertising claims based on TAC results.

TASKS

TAC 2012 will field shared tasks in three areas, all aimed at improving the ability to automatically populate knowledge bases (KBs) from text:

- **Entity linking task:** Determine which entity in an existing KB is meant when an entity such as a person, organization or location is mentioned in a text.
- **Slot filling task:** Find more information about an entity in a KB by mining a large text collection.
- **Cold start knowledge base population task:** Given a KB schema with an empty knowledge base, build the KB from scratch by mining a large text collection.

To promote research in populating probabilistic knowledge bases, systems may augment each assertion they return with a confidence score.

Languages:

- **English:** All tasks
- **Chinese:** Entity linking
- **Spanish:** Entity linking and slot filling

KBP Track coordinators:

- James Mayfield (james.mayfield@jhuapl.edu)
- Javier Artiles (javier.artiles@mail.rakuten.com)

REGISTRATION

Organizations wishing to participate in any of the TAC 2012 tasks are invited to [register online](#) by June 15, 2012. Participants are advised to register and submit all required agreement forms as soon as possible in order to receive timely access to track resources, including any sample and training data. Registration for a track does not commit you to participating in the track, but is helpful to know for planning. Late registration will be permitted only if resources allow. Any questions about conference participation may be sent to the TAC project manager: tac-manager@nist.gov.

- Track registration form: <http://www.nist.gov/tac/2012/KBP/track-app.html>

WORKSHOP

The TAC 2012 workshop will be held November 5-6, 2012, in Gaithersburg, Maryland, USA. The TAC workshop is a forum both for presentation of results (including failure analyses and system comparisons), and for more lengthy system presentations describing techniques used, experiments run on the data, and other issues of interest to researchers in Information Extraction and NLP. Track participants who wish to give a presentation during the workshop will submit a short abstract in September describing the experiments they performed. As there is a limited amount of time for oral presentations, the abstracts will be used to determine which participants are asked to speak and which will present in a poster session.

SCHEDULE

TAC 2012 Schedule (Preliminary)	
June 15	Deadline for registration for track participation
July - August	Deadlines for results submission (varies by task)
September 26	Deadline for workshop presentation proposals
By late September	Release of individual evaluated results to participants
October 3	Notification of acceptance of presentation proposals
mid October	Deadline for system reports (workshop notebook version)
November 5-6	TAC 2012 workshop in Gaithersburg, Maryland, USA
February 5, 2013	Deadline for system reports (final proceedings version)

ORGANIZERS

Javier Artiles (Coordinator, Rakuten Institute of Technology, javier.artiles@mail.rakuten.com)
James Mayfield (Coordinator, Johns Hopkins University, james.mayfield@jhuapl.edu)
Hoa Trang Dang (NIST, hoa.dang@nist.gov)
Joe Ellis (LDC, joellis@ldc.upenn.edu)
Kira Griffitt (LDC, kiragrif@ldc.upenn.edu)
Ralph Grishman (New York University, grishman@cs.nyu.edu)
Heng Ji (City University of New York, hengji@cs.qc.cuny.edu)
Xuansong Li (LDC, xuansong@ldc.upenn.edu)
Paul McNamee (Johns Hopkins University, paul.mcnamee@jhuapl.edu)
Boyan Onyshkevych (U.S. Department of Defense)
Stephanie M. Strassel (LDC, strassel@ldc.upenn.edu)

Cold Start: Putting the KB in KBP

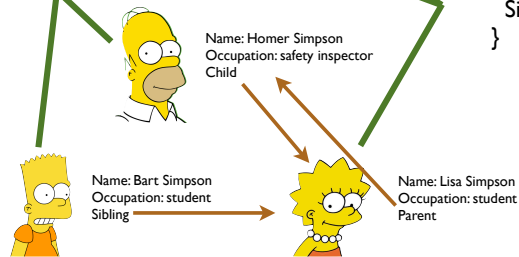
I. Given: Text with identified names

Meanwhile,
Lisa's father,
Homer, takes the
angel food, and
charges
admission to his
house for people
to view it.

Lisa and Bart
the typical
brother and sister
they're just like
me and my
brother .. I'd like
to make tracks on
me just like Bart.

3. Evaluate: SPARQL queries against KB using names in text as entry points

```
SELECT ?person
WHERE
{ Parent-of <mention("Homer")> ?person .
  Sibling-of <mention("Bart")> ?person
}
```



2. Produce: KB with links to names