



[TAC 2015 Tracks](#)
[Cold Start KBP](#)
[Tri-Lingual EDL](#)
[Event](#)
[Validation/Ensembling](#)
[Data](#)
[Schedule](#)
[Organizers](#)
[Track Registration](#)
[Reporting Guidelines](#)
[TAC 2015 Workshop](#)



TAC Knowledge Base Population (KBP) 2015

Evaluation: February-November, 2015

Workshop: November 16-17, 2015

Conducted by:

U.S. National Institute of Standards and Technology (NIST)

With support from:

U.S. Department of Defense

Overview

The Text Analysis Conference (TAC) is a series of evaluation workshops organized to encourage 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. TAC comprises sets of tasks known as "tracks," each of which focuses on a particular subproblem of NLP. TAC tracks focus on end-user tasks, but also include component evaluations situated within the context of end-user tasks.

The goal of TAC Knowledge Base Population (KBP) is to develop and evaluate technologies for populating knowledge bases (KBs) from unstructured text.

The Cold Start KBP task is to build a KB from scratch, using a predefined KB schema and a collection of unstructured text. The current KB schema consists of named entities that can be a person (PER), organization (ORG), or geopolitical entity (GPE); and predefined attributes (a.k.a "slots") for those named entities. In addition to end-to-end KB construction, the Cold Start track offers two diagnostic tasks: Entity Discovery (ED) and Slot Filling (SF). The Cold Start entity discovery task is to create a KB node for each PER, ORG, and GPE entity mentioned by name in the Cold Start collection of text, and to link all named mentions of the entity to its KB node. The Cold Start slot filling task is to search the same document collection to fill in values for specific slots for specific entities.

The Tri-lingual Entity Discovery and Linking (EDL) track aims to extract entity mentions from a source collection of textual documents in multiple languages (English, Chinese and Spanish), and link them to an existing Knowledge Base (KB). An EDL system is also required to cluster mentions for those NIL entities that don't have corresponding KB entries.

The event track in KBP 2015 aims to extract information about events from unstructured text, such that the information would be suitable as input into a structured KB.

Finally, the slot filler validation task in the validation/ensembling track focuses on the refinement of output from slot filling systems by either combining information from multiple slot filling systems, or applying more intensive linguistic processing to validate candidate slot fillers.

Tracks

- **Cold Start KBP Track**

The Cold Start KBP track builds a knowledge base from scratch using a given document collection and a predefined schema for the entities and relations that will comprise the KB. Cold Start KBP includes an **Entity Discovery (ED)** task to discover entities and cluster their named mentions, and a **Slot Filling (SF)** task to fill in values for predefined slots (attributes) for a given entity.

Track home page:

<http://www.nist.gov/tac/2015/KBP/ColdStart/>

Track coordinators: James Mayfield
(jamesmayfield@gmail.com) and Ralph Grishman
(grishman@cs.nyu.edu)

- **Tri-Lingual Entity Discovery and Linking Track (EDL)**

The Tri-Lingual Entity Discovery and Linking (EDL) track aims to extract entity mentions from a source collection of textual documents in multiple languages (English, Chinese, and Spanish), and link them to an existing Knowledge Base (KB); an EDL system is also required to cluster mentions for those entities that don't have corresponding KB entries.

Track home page: <http://nlp.cs.rpi.edu/kbp/2015/>

Track coordinator: Heng Ji (jih@rpi.edu)

- **Event Track**

The goal of the Event track is to extract information about events such that the information would be suitable as input to a knowledge base. The track includes **Event Nugget (EN)** tasks to detect and link events, and **Event Argument (EA)** tasks to extract event arguments and link arguments that belong to the same event.

Track home page: <http://www.nist.gov/tac/2015/KBP/Event/>

Event Nugget coordinators: Eduard Hovy
(ehovy@andrew.cmu.edu) and Teruko Mitamura
(teruko@cs.cmu.edu)

Event Argument coordinator: Marjorie Freedman
(mfreedma@bbn.com)

- **Validation/Ensembling Track**

The Validation/Ensembling track focuses on the refinement of output from slot filling systems by either combining information from multiple slot filling systems, or applying more intensive linguistic processing to validate individual candidate slot fillers.

Track home page:

<http://www.nist.gov/tac/2015/KBP/SFValidation/>

Track coordinator: Hoa Dang (hoa.dang@nist.gov)

Preliminary Schedule

Preliminary TAC KBP 2015 Schedule	
April 15	Track registration opens
Mid April	Initial track guidelines posted
June 30	Deadline for registration for track participation
August - October	Track evaluation windows (varies by track)
August 3-17	<i>Cold Start KBP evaluation window</i>
August 17-24	<i>Event (Argument Extraction and Linking) evaluation window</i>
August 31-Sept 7	<i>Event (Argument Verification and Linking) evaluation window</i>
August 31-Sept 11	<i>Event (Nugget Detection) evaluation window</i>
August 31-Sept 11	<i>Event (Nugget Detection and Coreference) evaluation window</i>
Sept 14-21	<i>Event (Nugget Coreference) evaluation window</i>
Sept 21-28	<i>Slot Filler Validation/Ensembling evaluation window</i>
Sept 28-Oct 5	<i>Tri-Lingual EDL evaluation window</i>
Oct 7-12	<i>Diagnostic Tri-Lingual EL evaluation window</i>

By mid October	Release of individual evaluated results to participants (varies by track)
October 15	Deadline for short system descriptions
October 15	Deadline for workshop presentation proposals
October 21	Notification of acceptance of presentation proposals
Nov 1	Deadline for system reports (workshop notebook version)
November 16-17	TAC 2015 workshop in Gaithersburg, Maryland, USA
February 2016	Deadline for system reports (final proceedings version)

Organizing Committee

Hoa Trang Dang (U.S. National Institute of Standards and Technology)

Jason Duncan (MITRE)

Joe Ellis (Linguistic Data Consortium)

Marjorie Freedman (BBN Technologies)

Ralph Grishman (New York University)

Eduard Hovy (Carnegie Mellon University)

Heng Ji (Rensselaer Polytechnic Institute)

James Mayfield (Johns Hopkins University)

Teruko Mitamura (Carnegie Mellon University)

Boyan Onyshkevych (U.S. Department of Defense)

Shahzad Rajput (U.S. National Institute of Standards and Technology)

Zhiyi Song (Linguistic Data Consortium)

Stephanie Strassel (Linguistic Data Consortium)

NIST is an agency of the
[U.S. Department of Commerce](#)

[privacy policy](#) / [security notice](#) / [accessibility statement](#)
[disclaimer](#)
[FOIA](#)

Last updated: Wednesday, 15-Apr-2015 08:44:29 EDT
 Comments to: tac-web@nist.gov



[TAC 2015 Tracks](#)
[Cold Start KBP](#)
[Tri-Lingual EDL](#)
[Event](#)
[Event Nugget](#)
[Event Argument](#)
[Validation/Ensembling](#)
[Data](#)
[Schedule](#)
[Organizers](#)
[Track Registration](#)
[Reporting Guidelines](#)
[TAC 2015 Workshop](#)



TAC KBP 2015 Event Track

Overview

The goal of the TAC KBP Event track is to extract information about events such that the information would be suitable as input to a knowledge base. The track includes **Event Nugget (EN)** tasks to detect and link events, and **Event Argument (EA)** tasks to extract event arguments and link arguments that belong to the same event. In 2015, there are 5 sub-tasks of the TAC KBP Event Track:

- **EN Task 1: Event Nugget Detection:** This task aims to identify the explicit mentioning of Events in text for English. Participating systems must identify all relevant Event Mention instances within each sentence. Every instance of a mention of the relevant Event types/subtypes taken from the Rich ERE Annotation guidelines must be identified. In addition, systems must identify three REALIS values (ACTUAL, GENERIC, OTHER), which are also described in the Rich ERE guidelines.
- **EN Task 2: Event Nugget Detection and Coreference:** In addition to the Event Nugget Detection task described in the EN Task 1, this task also aims to identify Full Event Coreference links at the same time. Full Event Coreference is identified when two or more Event Nuggets refer to the same event. This notion is described as Event Hoppers in the Rich ERE Annotation Guidelines. The Full Event Coreference links do not include subevents relations.
- **EN Task 3: Event Nugget Coreference:** This task is to identify Full Event Coreference links, given the annotated Event Nuggets in the text.
- **EA Task 1: Event Argument Extraction and Linking:** Systems in this task aim to extract information about entities (and times) and the role they play in an event and then link those arguments that participate in the same event. Participating systems will extract tuples that include (EventType, Role, Argument) and group these tuples so that, for example the date, location and thing purchased of the same transaction event appear in the same frame. EventType and Role will be drawn from an externally specified ontology that is based on [ACE 2005](#) and Rich ERE. Arguments will be

strings from within a document representing the canonical (most-specific) name or description of the entity.

- **EA Task 2: Event Argument Verification and Linking:** This task is designed to allow participants to explore approaches to event argument linking without building their own event argument extraction systems. It also allows participants to explore approaches that improve the precision of extracted event arguments via system combination and/or verification algorithms. The output of this task is identical to that of EA Task 1 (Event Argument Extraction and Linking). Participants in this task will receive unlinked argument extraction output from all EA Task 1 participants.

For all 5 subtasks of the Event Track, systems will operate over a 100-500 document English corpus that is a mix of newswire and discussion forum documents. A discussion forum document may contain multiple posts. The corpus will be manually and automatically filtered to ensure at least a few instances of all event-types. The discussion-forum posts will be automatically filtered to identify those posts that are not simply reposts of newswire documents. Very long discussion-forum threads will be truncated.

The input documents for the EN tasks will be a subset of the input documents for EA.

Participants can submit to any number and combination of the 5 tasks in the Event track. Because input documents are shared across the 5 tasks, participants must refrain from examining any of the input documents before they have finished submitting results for all of the Event tasks in which they are participating.

Preliminary Schedule

Preliminary TAC KBP 2015 Schedule	
April 15	Track registration opens
Mid April	Initial track guidelines posted
June 30	Deadline for registration for track participation
August 17-24	<i>Event (Argument Extraction and Linking) evaluation window</i>
August 31-Sept 7	<i>Event (Argument Verification and Linking) evaluation window</i>
August 31-Sept 11	<i>Event (Nugget Detection) evaluation window</i>
August 31-Sept 11	<i>Event (Nugget Detection and Coreference) evaluation window</i>

Sept 14-21	<i>Event (Nugget Coreference) evaluation window</i>
By mid October	Release of individual evaluated results to participants (varies by track)
October 15	Deadline for short system descriptions
October 15	Deadline for workshop presentation proposals
October 21	Notification of acceptance of presentation proposals
Nov 1	Deadline for system reports (workshop notebook version)
November 16-17	TAC 2015 workshop in Gaithersburg, Maryland, USA
February 2016	Deadline for system reports (final proceedings version)

References

- ACE 2005 Training Data:
<http://catalog ldc.upenn.edu/LDC2006T06>

While the ACE 2005 event annotation is being provided to all participants, this task diverges from ACE in some cases. One example of divergence is the addition of correct answers derived through inference/world knowledge. This evaluation will treat as correct some cases that were explicitly excluded in ACE 2005.

Track Coordinators

Event Arguments: Marjorie Freedman (BBN, mfreedma@bbn.com)
 Event Nuggets: Teruko Mitamura (CMU, teruko@cs.cmu.edu) and
 Eduard Hovy (CMU, ehovy@andrew.cmu.edu)

NIST is an agency of the
[U.S. Department of Commerce](#)

[privacy policy](#) / [security notice](#) / [accessibility statement](#)
[disclaimer](#)
[FOIA](#)

Last updated: Friday, 22-May-2015 12:03:54 EDT
 Comments to: tac-web@nist.gov