# Text Analysis Conference
# TAC 2016

Hoa Trang Dang
*National Institute of Standards and Technology*

# TAC Goals

- To promote research in NLP based on large common test collections
- To improve evaluation methodologies and measures for NLP
- To build test collections that evolve to meet the evaluation needs of state-of-the-art NLP systems
- To increase communication among industry, academia, and government by creating an open forum for the exchange of research ideas
- To speed transfer of technology from research labs into commercial products

# Features of TAC

- Component evaluations situated within context of end-user tasks (e.g., summarization, knowledge base population)
  - opportunity to test components in end-user tasks
- Test common techniques across tracks
- "Small" number of tracks
  - critical mass of participants per track
  - sufficient resources per track (data, annotation/assessing, technical support)
- Leverage shared resources across tracks (organizational infrastructure, data, annotation/assessing, tools)

# Workshop

- Targeted audience is participants in the shared tasks and evaluations
- "Working workshop" – audience participation encouraged
- Presenting work in progress
- Objective is to improve system performance
  - Clarify task requirements, correct any false assumptions
  - Improve evaluation specifications and infrastructure
  - Learn from other teams

- 2016 evaluations largely in support of (and supported by!) DARPA DEFT program

# TAC 2016 Track Participants

- Track coordinators
  - EDL: Heng Ji; also Joel Nothman
  - Cold Start KB/SF/SFV: Hoa Dang, Shahzad Rajput
  - Event: Marjorie Freedman and BBN team (Event Arguments); Teruko Mitamura, Ed Hovy, and CMU team (Event Nuggets)
  - Belief and Sentiment: Owen Rambow
- Linguistic resource providers:
  - Linguistic Data Consortium (Joe Ellis, Jeremy Getman, Zhiyi Song, Stephanie M. Strassel, Ann Bies ….)
- 44 Teams: 10 countries (24 USA, 11 China, 2 Germany,….)

# TAC KBP 2016 Tracks

- Entity Discovery and Linking
- Cold Start KBP (CS)
  - KB Construction (CSKB)
  - Slot Filling (CSSF)
  - Slot Filler Validation (SFV)
- Event
  - Nugget Detection and Coreference (EN)
  - Argument Extraction and Linking (EAL)
- Belief and Sentiment (BeSt)

# TAC KBP 2016

| | Languages | Cross-Lingual | Docs Input | Docs evaluated, by gold standard annotation |
|---|---|---|---|---|
| EDL | ENG, CMN, SPA | Y | 90,000 / 3 | 500 / 3 |
| KB/SF/SFV | ENG, CMN, SPA | Y | 90,000 / 3 | (assessment) |
| Event Argument | ENG, CMN, SPA | Y | 90,000 / 3 | 500 / 3  (+assessment) |
| Event Nugget | ENG, CMN, SPA | N | 500 / 3 | 500 / 3 |
| Belief and Sentiment | ENG, CMN, SPA | N | 500 / 3 | 500 / 3 |

# 2016 Entity Discovery and Linking Track

- Task:
  - Entity Discovery and Linking (EDL): Given a set of documents, extract each entity mention, and link it to a node in the reference KB, or cluster it with other mentions of the same entity
- Entity types: PER, ORG, GPE, FAC, LOC
- Mention types: NAM, NOM
- 2015/2016 Reference KB:
  - Derived from Freebase snapshot
- Source documents: KBP 2016 Source Corpus
  - English, Chinese, Spanish
  - Newswire and discussion forum

# 2016 Cold Start KBP Track

- Goal: Build a KB from scratch, containing all attributes about all entities as found in a corpus
  - ED(L) system component identifies KB entities and all their NAM/NOM mentions
  - Slot Filling system component identifies entity attributes (fills in "slots" for the entity)
- Inventory of 41+ slots for PER, ORG, GPE
  - Filler must be an entity (PER, ORG, GPE), value/date, or (rarely) a string (per:cause_of_death)
  - Filler entity must be represented by a name or nominal mention
- Post-submission slot filling evaluation queries traverse KB starting from a single entity mention (entry point into the KB):
  - Hop-0: "Find all children of Michael Jordan"
  - Hop-1: "Find date of birth of each child of Michael Jordan"

# Cold Start KB/SF Task Variants and Evaluation

- Task Variants:
  - Full KB Construction (CS-KB): Ground **all** named or nominal entity mentions in docs to newly constructed KB nodes (ED, clustering); extract **all** attested attributes about **all** entities
  - SF (CS-SF): Given a query, extract specified attributes (fill in specified slots) for the query entities.
- (Primary) Slot filler evaluation:
  - Evaluation: P/R/F1 over slot fillers
  - Fillers grouped into equivalence classes (same entity, value, or string semantics); penalty if system returns multiple equivalent fillers.
  - Prefer named fillers over nominal fillers, if name exists in corpus
- (Diagnostic) Entity Discovery Evaluation for KBs:
  - Same as for EDL track, but ignore metrics for linking to a reference KB

# 2016 Event Track

- Given:
  - Source documents: KBP 2016 Source Corpus
    - EAL: all 90,000 docs
    - EN: 500 docs
  - Event Taxonomy: ~18 event types and their roles (Rich ERE, reduced set of types)
- Event Nugget:
  - Detection all mentions of events from the taxonomy, and corefer all mentions of the same event (within-doc)
- Event Argument:
  - Extract instances of arguments that play a role in some event from the taxonomy, and link arguments for the same event (within-doc)
  - Link coreferential event frames across the corpus
  - Don't have to identify all mentions (nuggets) of the event

# 2016 Belief and Sentiment

- Input:
  - Source Documents: ~500 docs from KBP 2016 Source Corpus
  - ERE (Entity, Relation, Event) annotations of documents
    - Gold
    - Predicted
- Task: Detect belief (Committed, Non-Committed, Reported) and sentiment (positive, negative), including source and target
  - Belief and Sentiment Source: Entity (PER, ORG, GPE)
  - Belief target: Relation ("John believed Mary was born in Kenya"), Event ("John thought there might have been demonstrations supporting his election")
  - Sentiment target: Entity, Relation, Event

# TAC KBP Evolution

- Goal: Populate a knowledge base (KB) with information about entities as found in a collection of source documents, following a specified schema for the KB
- KBP 2009-2011: Focus on augmenting an existing KB.
  - Decompose into 2 tasks: entity-linking (EL), slot-filling (SF)
- KBP 2012: Combine EL and SF to build KB -> Cold Start (CS).
- KBP 2013-2014:
  - + Conversational, informal data (discussion forum)
  - EL -> Entity Discovery (full-document NER) and Linking
  - + Event Argument Extraction
- KBP 2015: Fold SF track into Cold Start KB
  - + Event Nuggets and Argument linking
- KBP 2016: extend all tasks to 3 languages
  - + Belief and Sentiment
- KBP 2017: Fold Events, Belief, and Sentiment into Cold Start KB

# TAC 2017++ Session

- TAC 2017
  - Trilingual Cold Start++ KB
    - Entities (EDL), Relations (SF), Events (Arguments), Belief and Sentiment
  - Event Sequencing (tentative)
  - Adverse Reaction Extraction from Drug Labels

- Panel: What next, after 2017
  - KBP has been supporting DARPA DEFT program since 2013
  - DEFT ends in 2017
  - What next?