



DEFT Program Briefing

Dr. Bonnie Dorr

The BAA posted at FBO takes precedence
over any information contained herein



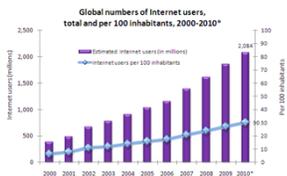
DEFT Goals

- Support the ability to see through language to meaning in text, make use of key information, prioritize documents containing new developments, and automate the initial stages of report writing
- HLT Research on Filtering, Anomalies, Relations (not exhaustive)
 - Deep semantic understanding of people, events, causal relationships, beliefs, anomalies, coreference, disfluencies, ambiguity, uncertainty, contradiction, unstated or implied information
- Development of robust fully integrated, configurable, trainable Alert System that supports human analysts with notifications of new developments and pre-report summaries
 - Notification of new developments (event-focused document cluster)
 - Pre-report summary of "points of interest"
- Transition of major capabilities to operational users



Why DEFT?

Exponential growth in Internet users and access, social media, and types of



2.3B Internet users (out of world population of 7B)

+



Fast, 24x7 access

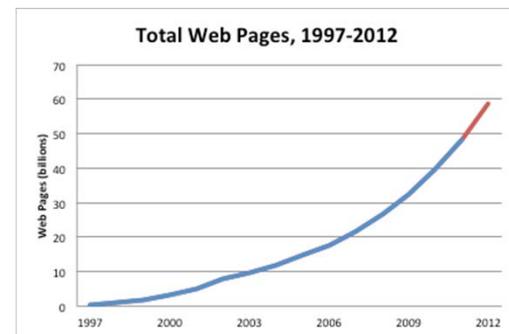
+



Text files, Web pages, articles, Craigslist ads, blog posts, Facebook updates, Tumblr posts, email messages, instant messages, Twitter tweets, etc.

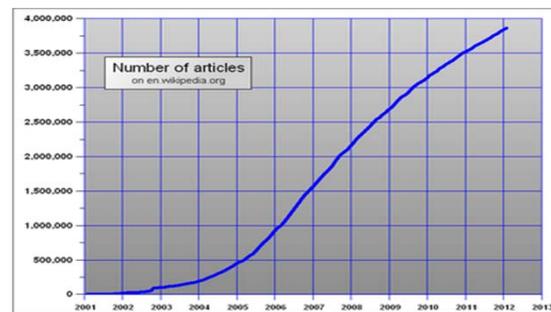
Leads to exponential growth in open-source text

- ~ 50B indexable Web pages today
- Google Books to contain 130M books (4B pages) by 2020
 - Now contains 25M books (750M pages)
- Library of Congress contains 32M books (200TB, scanned)
- English Wikipedia has 3.9M articles



Text growth similar at potential end-user sites

- Document bases exploding
 - One topic grew from 390K to 3.5M documents in a few years
 - One potential end-user's estimate: analysts would have to read 1 msg every 15 seconds!



English Wikipedia Articles



Deep Language Understanding from Reports

Example Report Text

On Friday February 17, terrorist organization X managed to sneak a car laden with explosives into the central intelligence building in capitol of country Y. According to authorities, two terrorist attack group members stole mobile phones as a ruse to get themselves arrested, and get their vehicle inside. The explosives were then detonated by remote control injuring several soldiers. Since I arrived here last week, several car bombs have been discovered, just in time, before they could wreak any major damage. Deputy Commander Z, is worried. He believes "members of terrorist organization X are hiding themselves among people returning" from a certain corridor just outside the capitol of country Y. It isn't difficult for them to do, thousands of people have been fleeing the region, concerned about an imminent attack.

Ambiguity, Vagueness

People & Groups

Facilities, Equipment

Geo-spatial-temporal(when, where)

Modality/Beliefs

Anomaly, Novelty, Emerging trends

Inter-relatedness

Events (who, what)

Dynamic, Changing Situation

Activities & processes (who, how, where)

Uncertainty, Inconsistency

Causal Explanation (why, how)

Unstated, Implied information



Deep Language Understanding from Informal Communication

Example Information Communication

A: Where were you? We waited all day for you and you never came.

B: I couldn't make it through, there was no way. They...they were everywhere. Not even a mouse could have gotten through.

A: You should have found a way. You know we need the stuff for the...the party tomorrow. We need a new place to meet...tonight. How about the...uh...uh...the house? You know, the one where we met last time.

B: You mean your uncle's house?

A: Yes, the same as last time. Don't forget anything. We need all of the stuff. I already paid you, so you had better deliver. You had better not \$%*! this up again.

Ambiguity, Vagueness

People & Groups

Facilities, Equipment

Geo-spatial-temporal(when, where)

Modality/Beliefs

Anomaly, Novelty, Emerging trends

Inter-relatedness

Events (who, what)

Dynamic, Changing Situation

Disfluency, Disjointedness

Uncertainty, Inconsistency

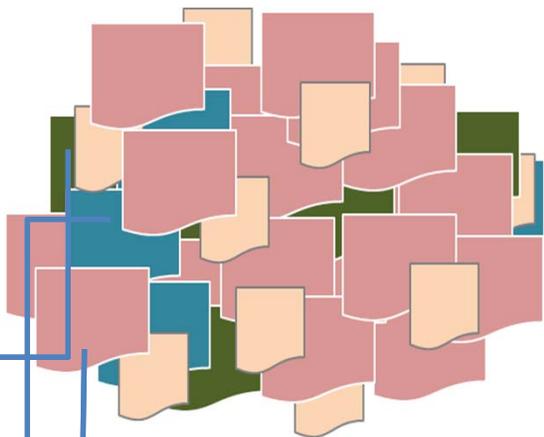
Causal Explanation (why, how)

Unstated, Implied information



Limitations in Current Process

Today: Continuous streams of massive language data



Irrelevant material
Potentially relevant material
Relevant implicit material

Keyword search will miss implicitly conveyed information.

← Analyst creates keyword search to find data.



→ Analysts gets back all documents with terms that match search keywords.



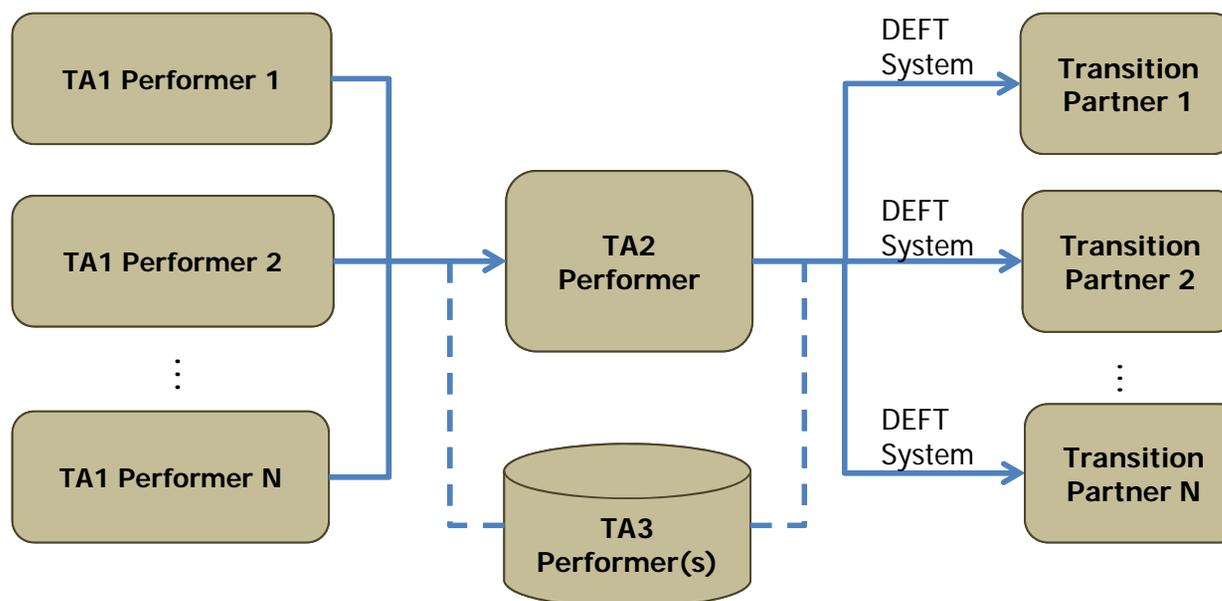
Why is now the time for DEFT?

- Urgent operational need to find implicit information in highly ambiguous or vague input
- Human language technology R&D now supports deep content analysis
 - Interpretation of words given context
 - Probabilistic linguistic models to enrich context models
- Prior DARPA programs did not focus on deep language understanding:
 - Infer causal relationships:
 - Thousands of people have been fleeing the region, [because the people are] concerned about an imminent attack.
 - A: We waited all day and you never came.
B: I couldn't make it through.
 - Detect novelty/anomaly:
 - Terrorist organization X managed to sneak a car laden with explosives into the central intelligence building in the capitol of country Y
 - You know we need the stuff for the...the party.
You had better not \$%*! this up again



DEFT Program Structure

- Proposers may bid on one or more technical areas:
 - TA1:** Implement Algorithms for Smart Filtering, Relational Analysis, Anomaly Analysis and System-Level Output
 - TA2:** Integrate TA1 implemented algorithms into a single DEFT system and Transition of DEFT to end-user sites
 - TA3:** Create and annotate data in support of DEFT evaluations (no overlap with TA1 or TA2 performers)
- Proposals not solicited for Evaluation:
 - Research Evaluation to be done by NIST language group in each phase:
 - Algorithm-level evaluations at NIST (3-4 sub-tasks each year)
 - DEFT system-level evaluations at NIST on unclassified proxy problems
 - Transition Evaluations (algorithm-level and system-level) at end-user sites





Technical Area 1: Implemented Algorithms*

- Smart Filtering Example Algorithm Topics:
 - Redundancy Detection
 - Semantic Equivalence Detection
 - Relational Analysis Example Algorithm Topics:
 - Event and Entity Coreference Resolution
 - Semantic Relation Detection
 - Anomaly Analysis Example Algorithm Topics:
 - Disfluency Detection
 - Novelty Detection
-
- Alert Generation:
 - Document Clustering
 - Summary Bullet Generation

HLT
Components

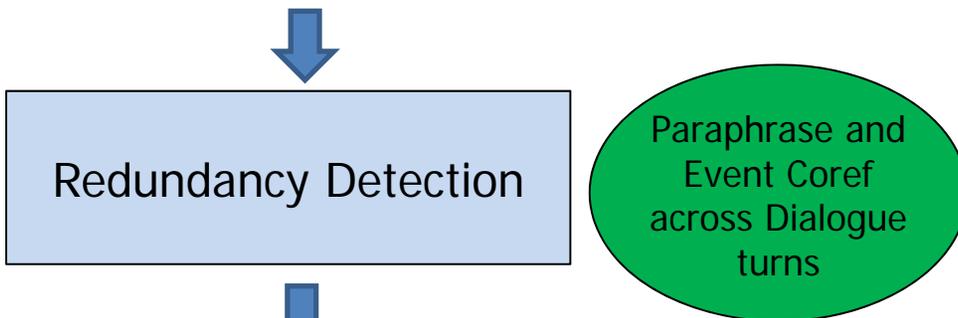
System-Level
Output

*Not exhaustive or exclusive



Smart Filtering (e.g., Redundancy Detection)

Input: Find information on Company X and their merger with Company Y



Output:

Retrieved Documents Determined to be Redundant

Company X and Company Y have merged in a 1.9 billion dollar deal.

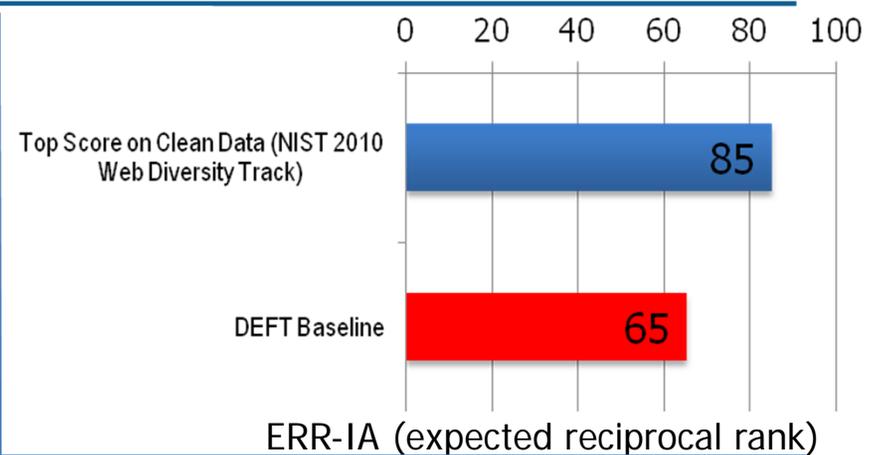
The merger between Company X and Company Y had a transaction value of \$1.9 billion.

Retrieved Documents Determined to be Non-Redundant

This merger is a cash plus stock reorganization.

The total merger consideration is \$29.394 per Company X share.

Company X and Company Y have completed their merger deal worth about \$1.9 billion.



Report:

...**returning** from a certain corridor just outside the capitol ... **fleeing** the region

Informal Communication:

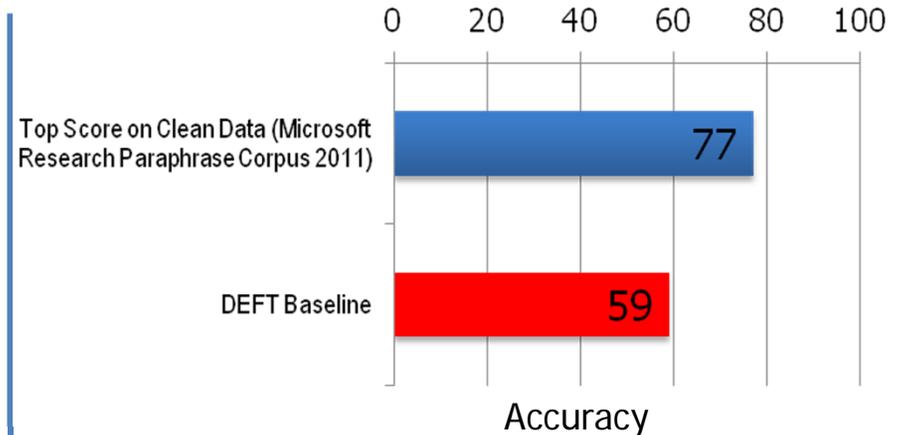
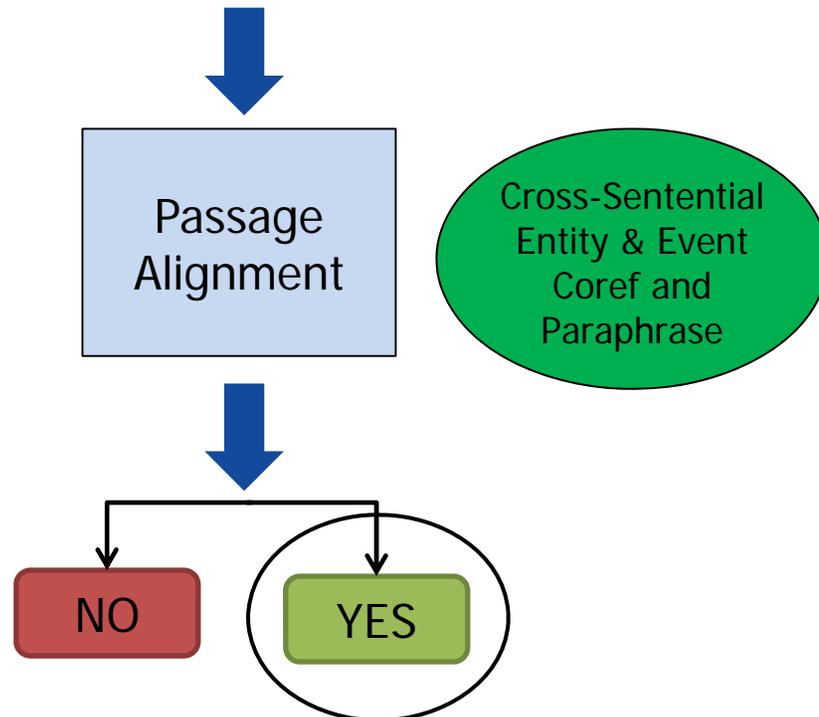
We need a **new place to meet** ...tonight.
You know, **the one where we met last time.**



Smart Filtering (e.g., Semantic Equivalence Detection)

About 120 potential jurors were being asked to complete a lengthy questionnaire.

The jurors were taken into the court room in groups of 40 and asked to fill out a questionnaire.



Report:

...**sneak** a **car** laden with explosives **into** ...
... **get** their **vehicle inside** ...

Informal Communication:

A: How about the...uh...uh...
house? You know, the **one** where we met last time.

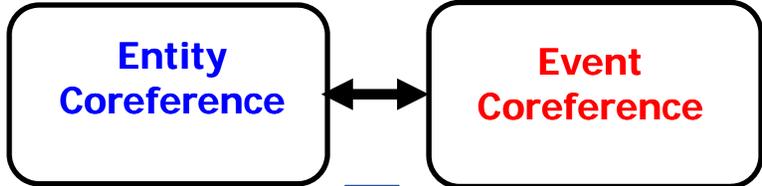
B: You mean **your uncle's**?



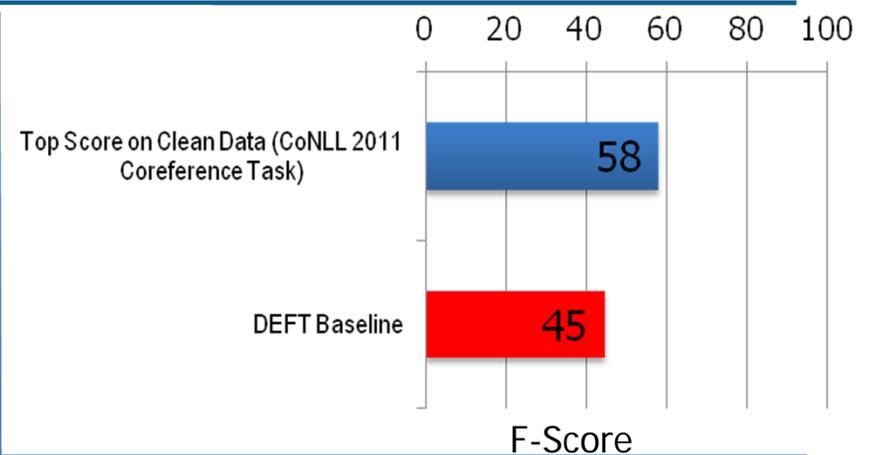
Relational Analysis (e.g., Coreference Resolution)

At least 40 people were hurt in the earthquakes... A series of powerful earthquakes injuring dozens and destroying...

Complex Paraphrase Detection for relating Redundant Events



At least 40 people were hurt in the earthquakes... A series of powerful earthquakes injuring dozens and destroying...



Report: Terrorist organization X managed to sneak a car laden with explosives into the central intelligence building... get their vehicle inside ... the explosives were then detonated...

Informal Communication:

I couldn't make it through, there was no way ...

Not even a mouse could have gotten through.

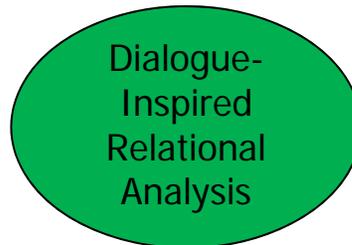


Relational Analysis (e.g., Causal Relation Detection)

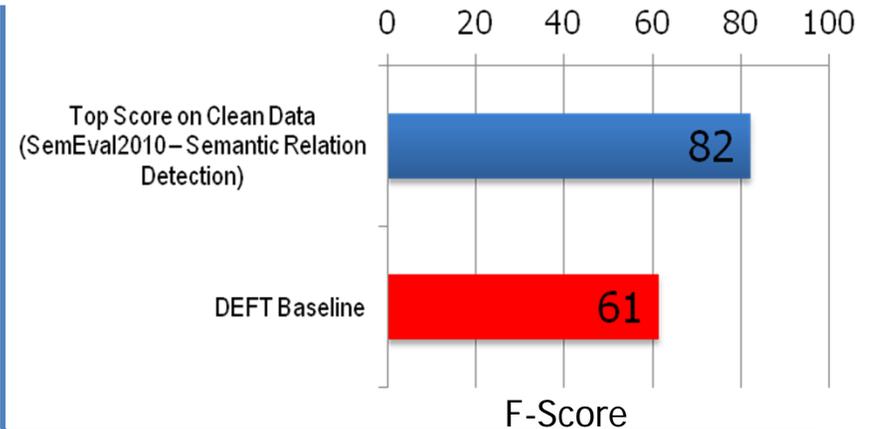
Forward **[motion]E1** of the vehicle through the air ended up in a **[suction]E2** on the road draft tube.



"ended up" →
causal relation



[motion]E1 CAUSE [suction]E2



Report:

...people **have been fleeing** the region, concerned about an **imminent attack...**

Informal Communication:

A: We waited all day and **you never came.**

B: I **couldn't make it through.**



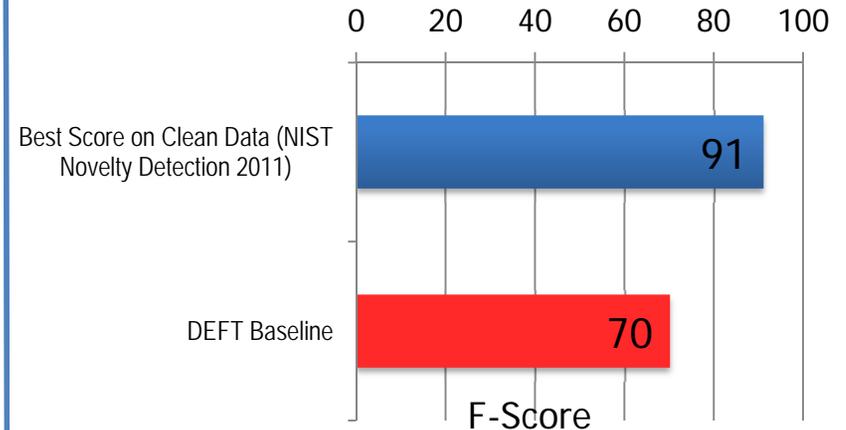
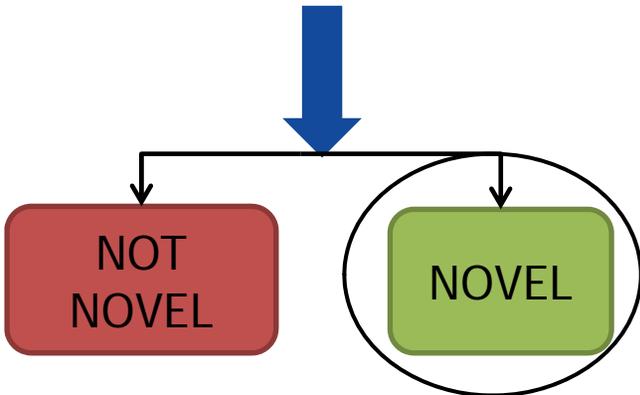
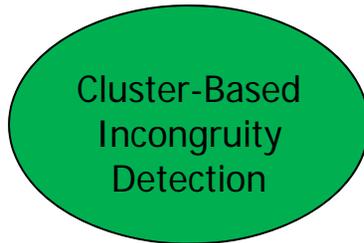
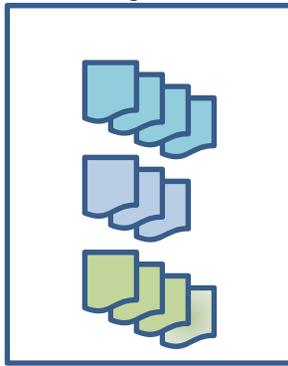
Anomaly Analysis (e.g., Novelty Detection)

News stories (input)



Hypothesis: Thousands of people were forced to leave their pets behind when they evacuated New Orleans

"Memory of events"



Report: Terrorist organization X managed to **sneak a car laden with explosives into** the central intelligence **building**...

Informal Communication:
You know we need the stuff for the...**the party**
...You had better not **\$%*! this up again.**



Anomaly Analysis (e.g., Disfluency Detection)

Input: speech with disfluencies

...a flight to Boston, uh, I mean to Houston on Friday...



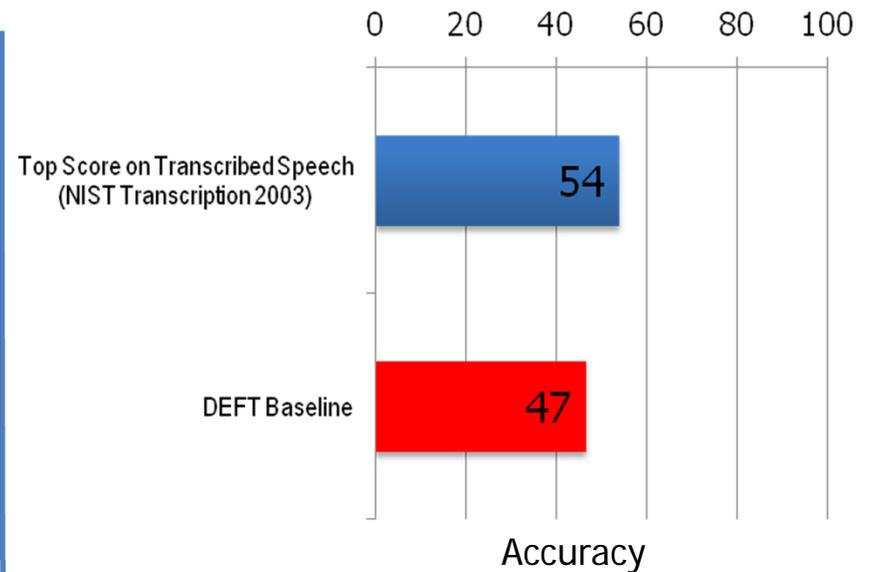
"uh" = "filled pause"
"I mean" = "repair"

Prosodic/
Intonation
Clue Detection



Output: annotation of identified disfluencies

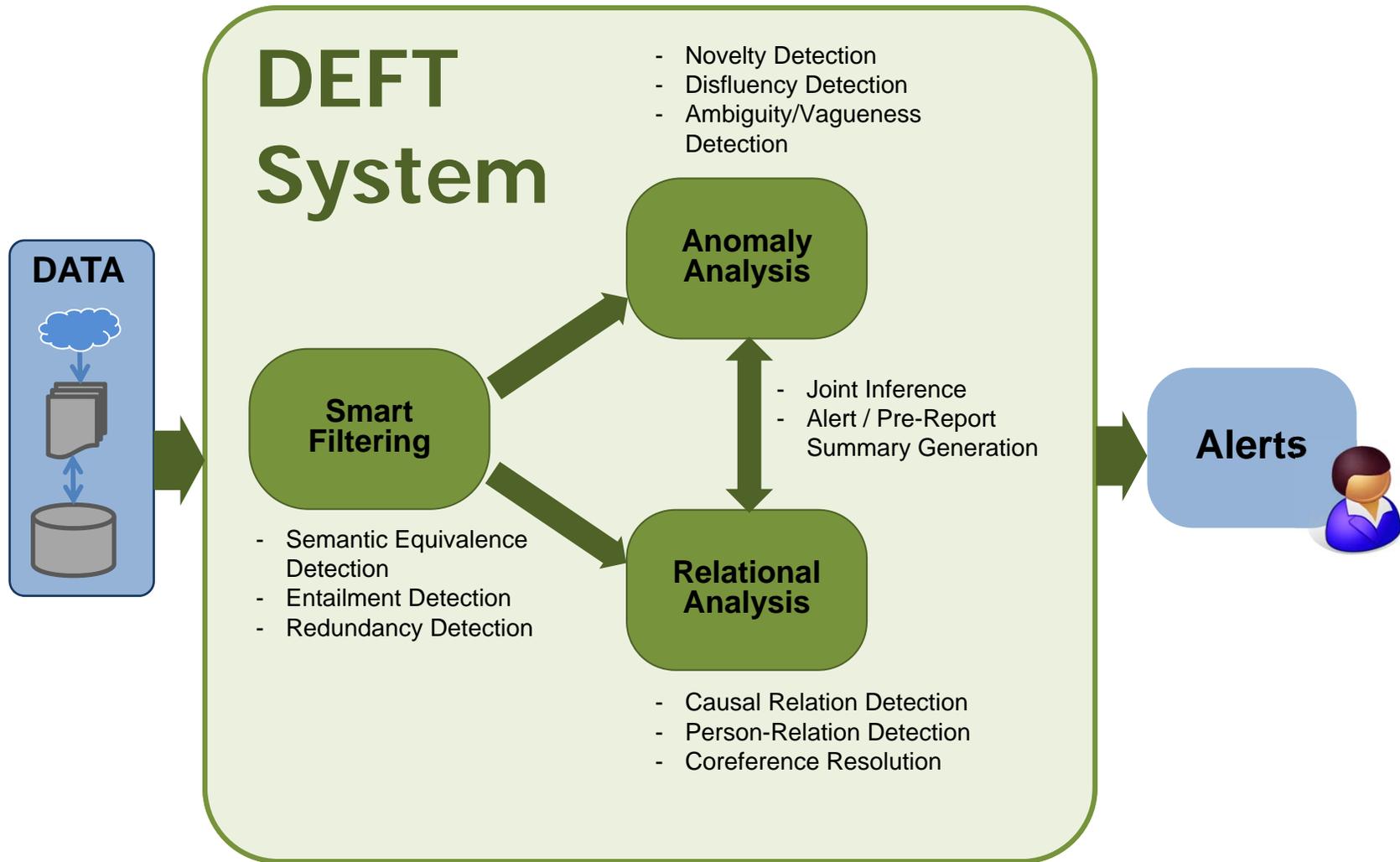
...a flight to Boston, **uh**, **I mean** to Houston on Friday...



Informal Communication:
we need **the stuff** for
the...the party tomorrow.



Technical Area 2: End-to-End DEFT System

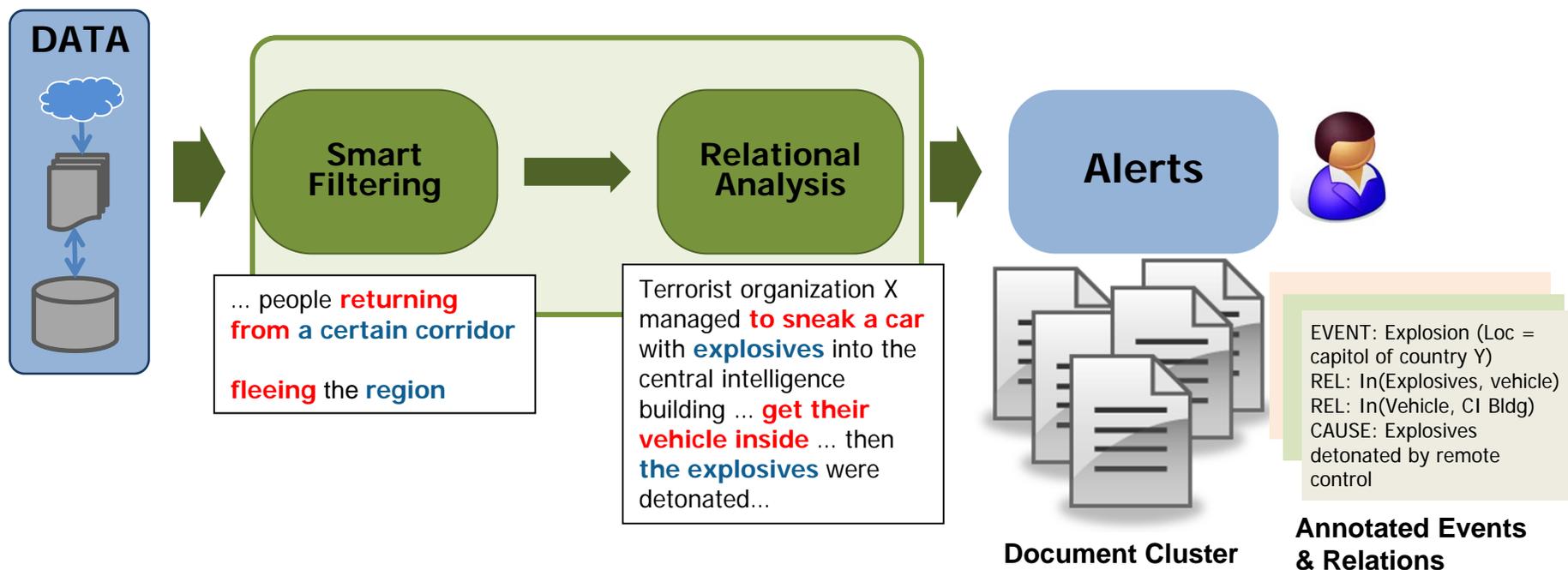




Phase 1 DEFT Challenge – Events and Relations

Task:

- Discover and annotate events and relations (both explicitly and implicitly expressed) that are of critical operational interest to transition partners
 - Candidate events include: Demonstrations, Bombings, Acquisitions, Border crossings, Elections, Campaigns, Government transitions, Equipment movements
 - Candidate relations include: interpersonal relationships, ownership, control, participation
- Generate alerts in the form of document clusters or passages of interest, with event and relation annotation
- TA2 performer transitions TA1 implemented algorithms directly to end-user sites

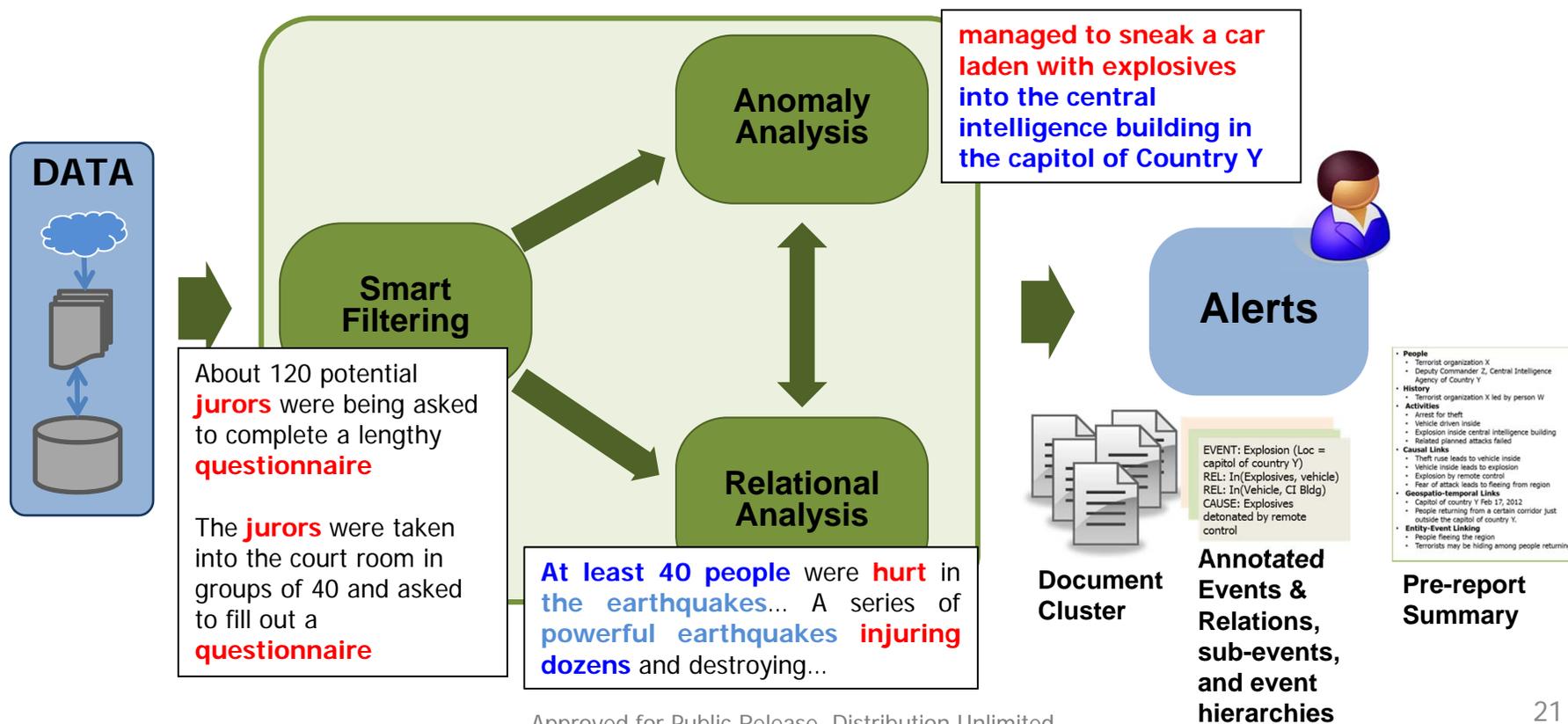




Phase 2 DEFT Challenge – End-to-End Capabilities

Task:

- Phase 2 will involve discovery and annotation document clusters, entity relationships (both explicitly and implicitly expressed), and anomalies on multiple data input genres
- Based on document analysis, generate alerts in the form of document clusters or passages of interest, with event and relation annotation, sub-events and event hierarchies, and a pre-report summary
- TA2 end-to-end prototype developed and integrated into end-user sites (Phase 2 & 3)

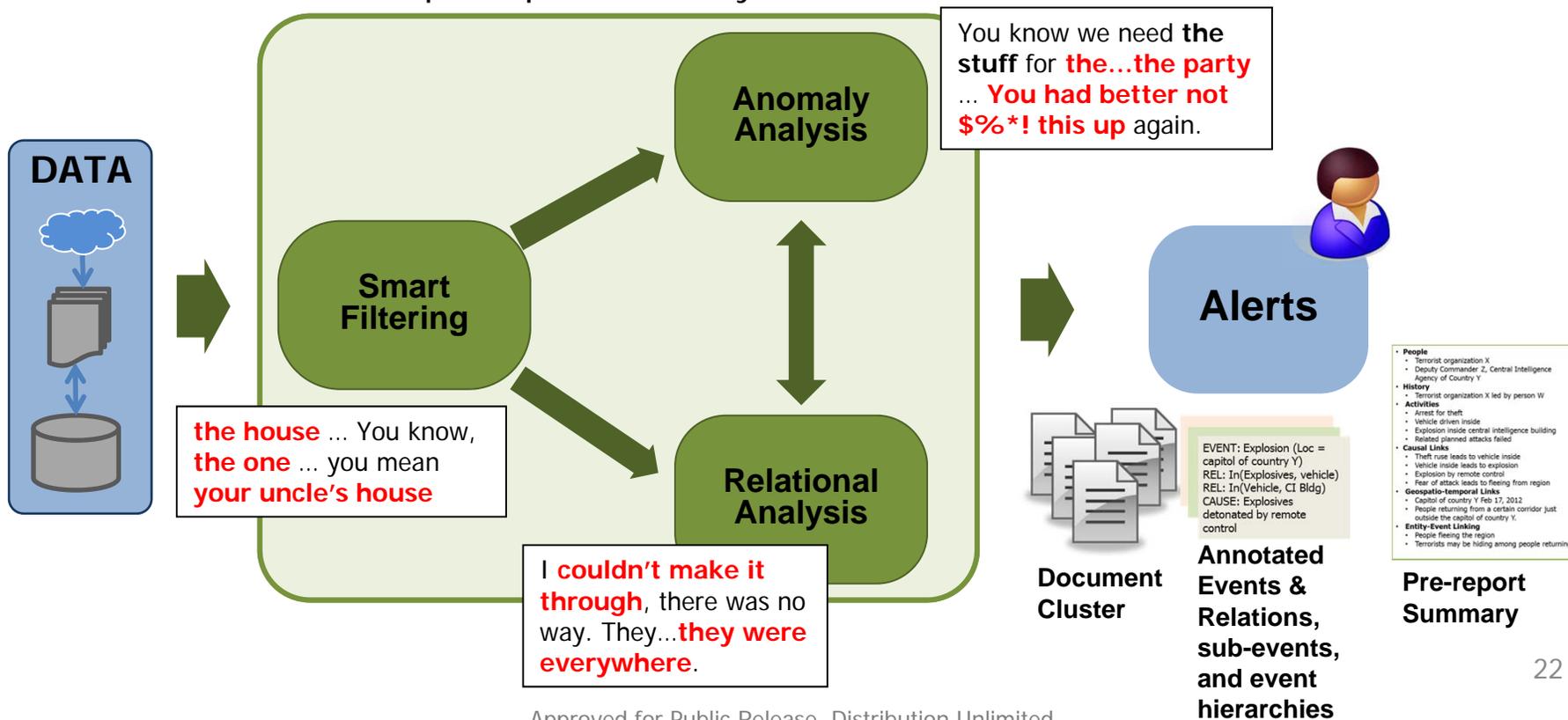




Phase 3 DEFT Challenge – Ambiguous language

Task:

- In addition to the Phase 1 and 2 challenges, the input data in Phase 3 will include ambiguous information about events and entities of interest.
 - Ambiguity may be achieved by using substitutions or vague language (such as “the stuff” and “the party”)
- Based on document analysis, generate alerts in the form of document clusters or passages of interest, with event and relation annotation, sub-events and event hierarchies, and a pre-report summary. Transition to end-user site.





What would DEFT produce? (Ex 1: Report Input)

Report Input

On Friday February 17, terrorist organization X managed to sneak a car laden with explosives into the central intelligence building in capitol of country Y. According to authorities, two terrorist attack group members stole mobile phones as a ruse to get themselves arrested, and get their vehicle inside. The explosives were then detonated by remote control injuring several soldiers. Since I arrived here last week, several car bombs have been discovered, just in time, before they could wreak any major damage. Deputy Commander Z, is worried. He believes "members of terrorist organization X are hiding themselves among people returning" from a certain corridor just outside capitol of country Y. It isn't difficult for them to do, thousands of people have been fleeing the region, concerned about an imminent attack.

DEFT Summary Bullets

- **People**
 - Terrorist organization X
 - Deputy Commander Z, Central Intelligence Agency of Country Y
- **History**
 - Terrorist organization X led by person W
- **Activities**
 - Arrest for theft
 - Vehicle driven inside
 - Explosion inside central intelligence building
 - Related planned attacks failed
- **Causal Links**
 - Theft ruse leads to vehicle inside
 - Vehicle inside leads to explosion
 - Explosion by remote control
 - Fear of attack leads to fleeing from region
- **Geospatio-temporal Links**
 - Capitol of country Y Feb 17, 2012
 - People returning from a certain corridor just outside the capitol of country Y.
- **Entity-Event Linking**
 - People fleeing the region
 - Terrorists may be hiding among people returning



What would DEFT produce? (Ex 2: Conversation Input)

Informal Communication Input

A: Where were you? We waited all day for you and you never came.

B: I couldn't make it through, there was no way. They...they were everywhere. Not even a mouse could have gotten through.

A: You should have found a way. You know we need the stuff for the...the party tomorrow. We need a new place to meet...tonight. How about the...uh...uh...the house? You know, the one where we met last time.

B: You mean your uncle's house?

A: Yes, the same as last time. Don't forget anything. We need all of the stuff. I already paid you, so you had better deliver. You had better not \$%*! this up again.

DEFT Summary Bullets

- **People**
 - Person A – Superior; planner.
 - Person B – Inferior; courier.
 - Unspecified group “we”
- **Associations**
 - B has previously received a call from a person of interest
- **Activities**
 - Meeting between A and B
 - One failed attempt today
 - Planning another attempt for tonight. Time unspecified.
 - Planned delivery of “stuff” by B to A
- **Causal Links**
 - Attempted meeting did not happen due to security in area of meeting.
- **Geospatio-Temporal Links**
 - The planned event is going to happen tomorrow
 - Location: “the house”, point of call origination
- **Entity-Event Linking**
 - Replanned meeting associated with “your uncle's house”
 - Previous meetings at “your uncle's house”²⁴
 - B paid an unspecified amount for services.

Information from other Sources



Technical Area 3: Data Creation and Annotation

Months from program kickoff	8	24	40
Narrative English text	400k words	400k words	400k words
Conversational English text	400k words	400k words	400k words
Conversational English speech, automatically transcribed	53 hours	53 hours	53 hours
Conversational foreign language text, automatically translated	400k words (in each proposed language)	400k words (in each proposed language)	400k words (in each proposed language)
Conversational foreign language text, not translated		400k words (in each proposed language)	400k words (in each proposed language)
Conversational foreign language speech, automatically transcribed and translated		53 hours (in each proposed language)	53 hours (in each proposed language)
Conversational foreign language speech, automatically transcribed but not translated			53 hours (in each proposed language)



Initial Data for DEFT Program

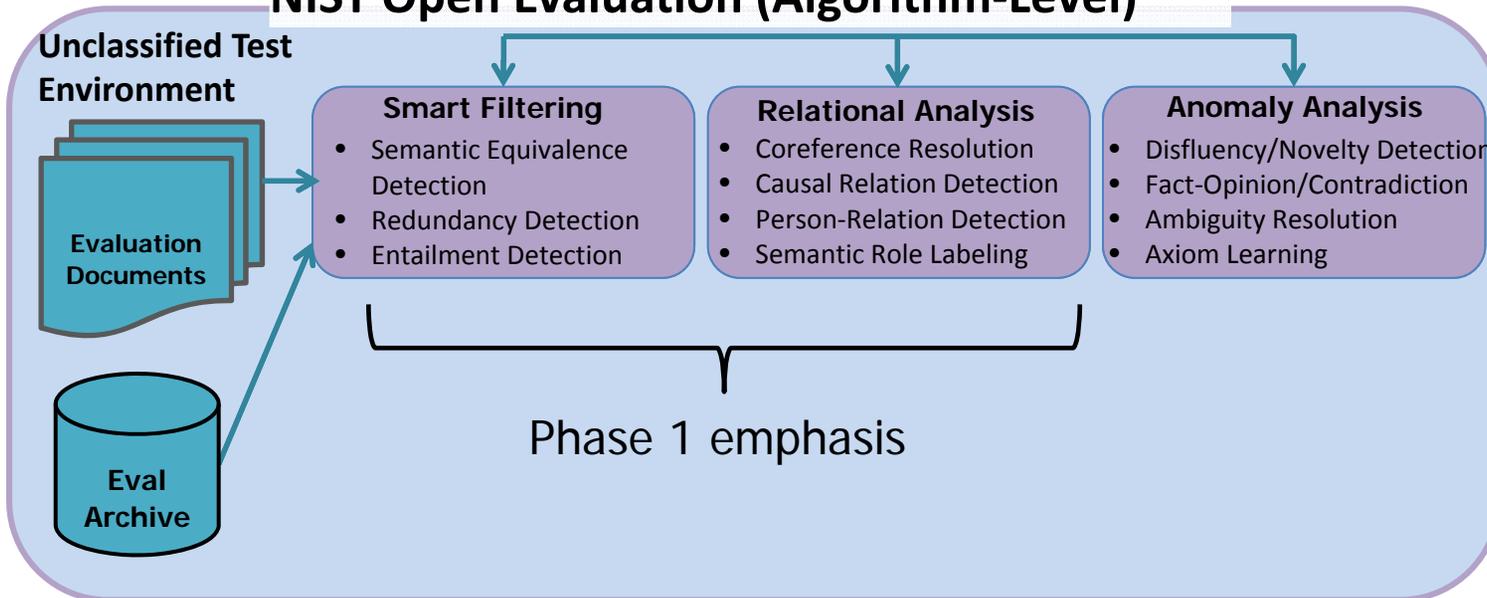
Advance Data Creation

- Anticipated that initial data will be available by DEFT kick-off
- Three types of conversations: no topic, explicit topic, and implicit topic
 - Speech – about 300k words
 - Constructed phone conversations
 - Data to be manually transcribed with disfluencies and words of interest (explicit words and substitution words) annotated
 - Text – about 300k words
 - Constructed email, text, and chat conversations
 - Words of interest (explicit words of interest and substitution words) annotated

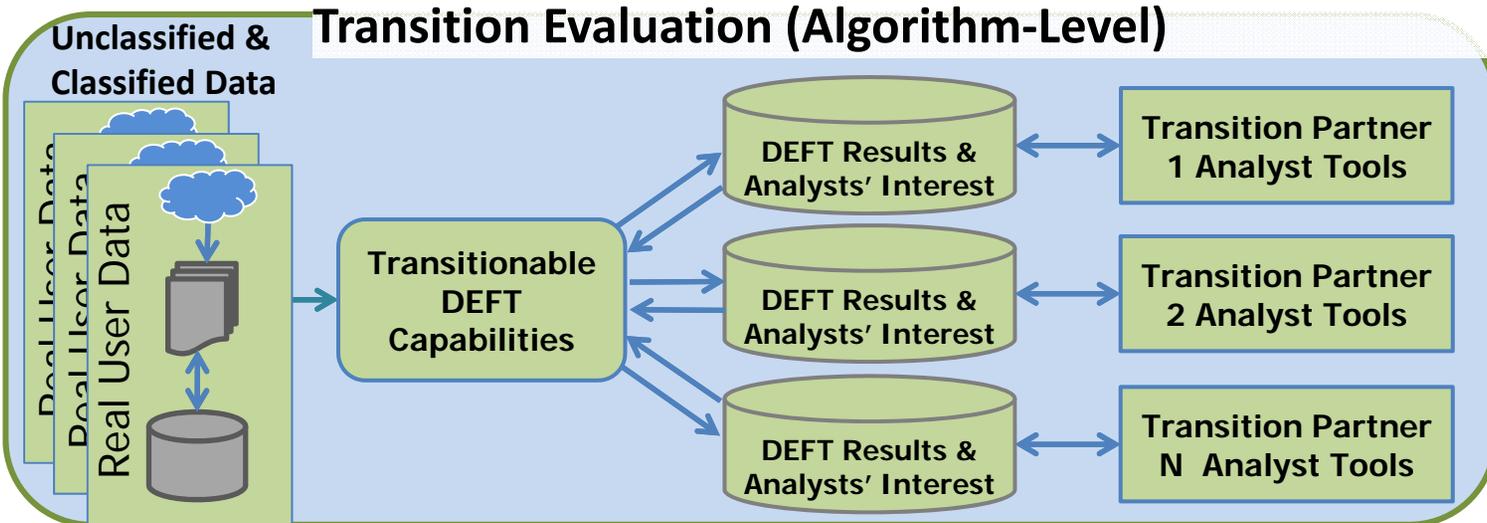


Anticipated DEFT Evaluation Plan: Phase 1

NIST Open Evaluation (Algorithm-Level)

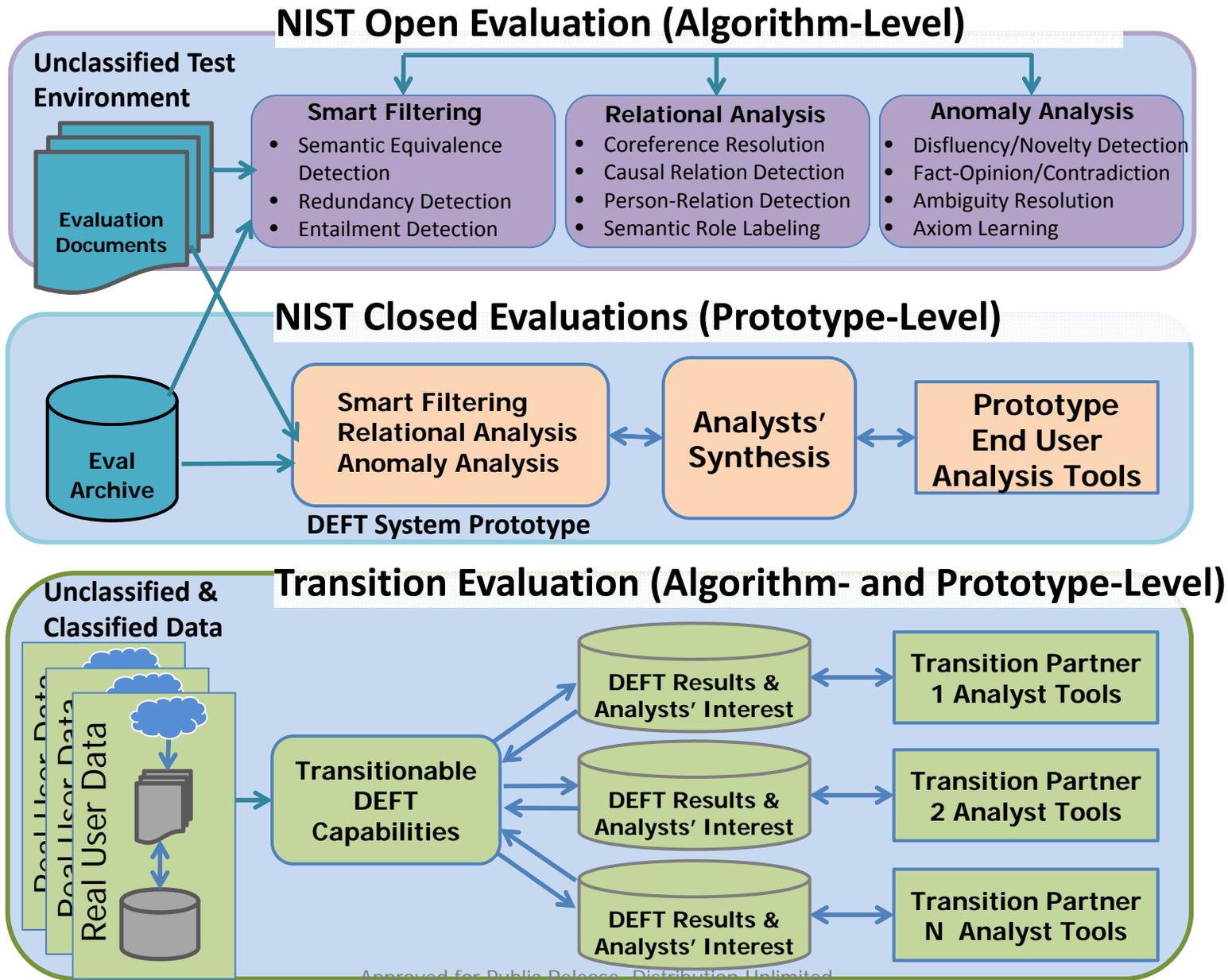


Transition Evaluation (Algorithm-Level)





Anticipated DEFT Evaluation Plan: Phases 2 and 3





Anticipated DEFT Evaluation Metrics

- NIST Open Evaluations (Algorithm-Level) (P1,P2,P3)
 - F-score of at least 90% by end of Phase 3.
 - TA1 proposers must describe an evaluation paradigm suitable for each implemented algorithm.
 - Smart Filtering Example: F-score computed from percentages of correct and incorrect assignments of passages to an appropriate group (cluster) of related passages.
 - Relational Analysis Example: F-score computed from percentages of correct and incorrect relation identifications.
 - Anomaly Analysis Example: F-score computed from percentages of correct and incorrect anomaly identifications.
- NIST Closed Evaluations (Prototype-Level) (P2, P3)
 - Probability of at least 80% of correct Alert detection by end of Phase 3.
- Transition Evaluations (Algorithms- and Prototype-Level) (P1, P2, P3)
 - F-score for event-focused annotations (P1, P2, P3), event hierarchies (P2, P3), and Alerts (P2, P3) at site-specific satisfactory performance
 - Site-specific utility of event-focused annotations, event hierarchies, and Alerts.

Submit Questions by 11:30 a.m.

