

KB Representation of Text, Audio, Images, and Video

Boyan Onyshkevych

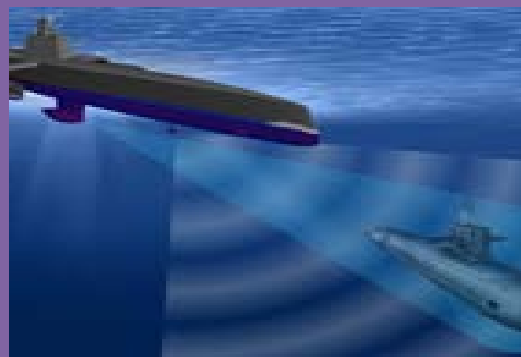
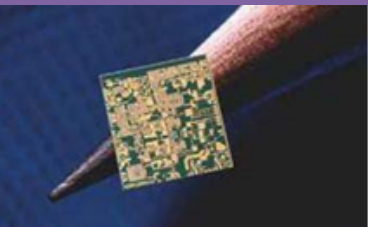
AKBC

December 2014





Defense Advanced Research Projects Agency (DARPA)



DARPA is the DoD's high-risk/high-reward research organization

DARPA's mission is to maintain the technological superiority of the U.S. military and prevent technological surprise from harming our national security.

We fund researchers in industry, universities, government laboratories, and elsewhere to conduct research and development projects that will benefit U.S. national security.



DARPA's Information Innovation Office (I2O)



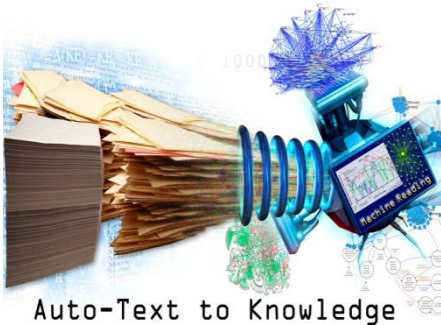
I2O aims to ensure U.S. technological superiority in all areas where information can provide a decisive military advantage.

Areas include conventional defense mission areas:

- ISR, C4I, networking, and decision-making.
- Emergent information technologies such as:
 - Artificial intelligence and/or big data
 - Crowd-based development paradigms
 - Natural language processing

The I2O capabilities enable the warfighter to:

- Understand the battlespace and the capabilities, intentions and activities of allies and adversaries;
- Discover insightful & effective strategies, tactics and plans;
- Securely connect the warfighter to the people and resources required for mission success.





Deep Exploration and Filtering of Text (DEFT)

Increasing the yield of actionable intelligence
from large volumes of data



Image sources: defenseimagery.mil and wikimedia.org

Goal: Identify explicit and implicit information from multiple unstructured text sources to support automated analytics and human analysts

Approach: Convert text to structured representation of alternatives

- Find and represent key information, including information on entities, relations, events, sentiment, beliefs, and intentions
- Combine information from multiple sources, detect inter-document relationships, and represent the information in a structured KB
- Identify emergent situations and anomalies through KB analysis



DEFT Example English

Source Material

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

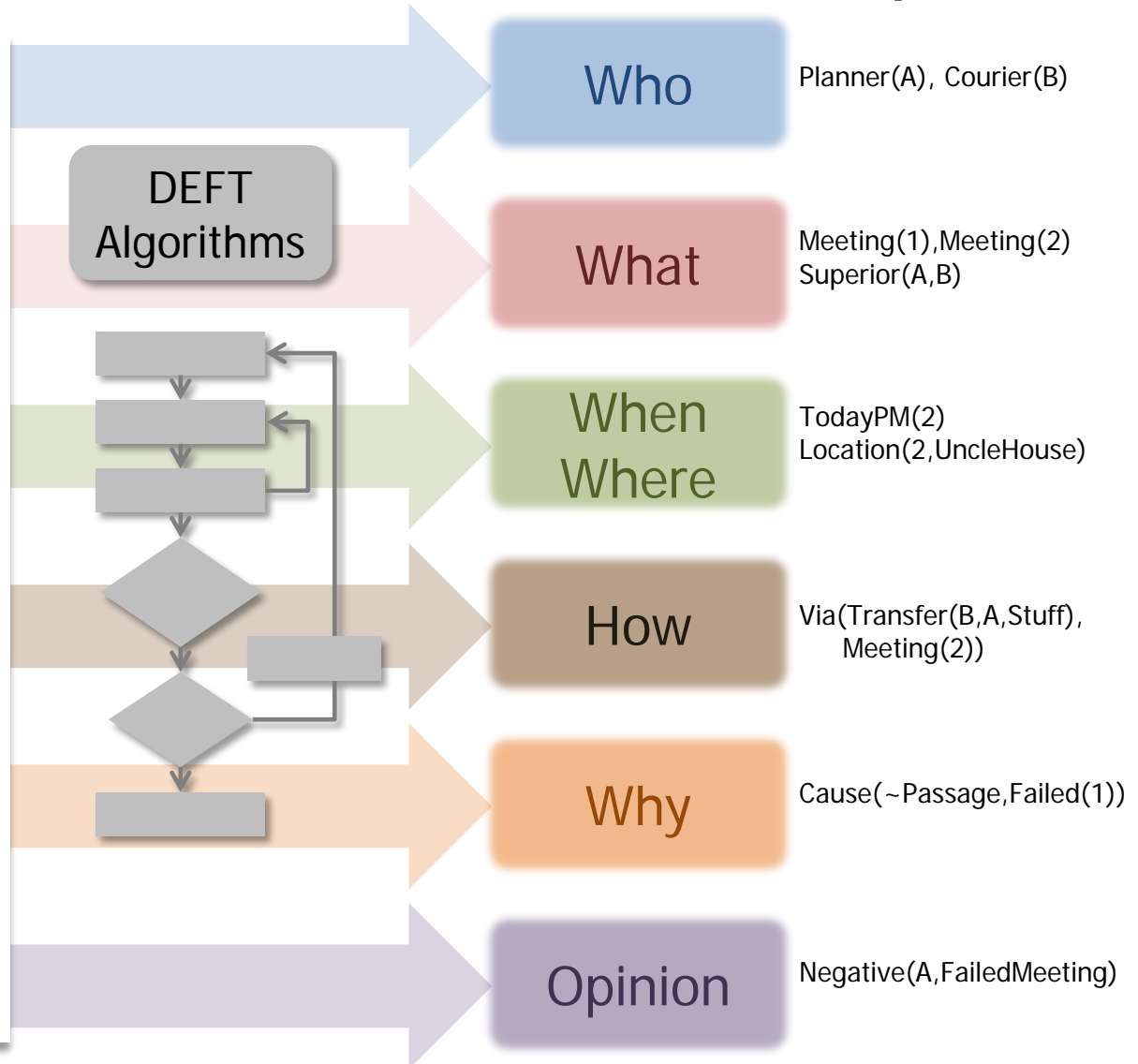
B: I couldn't make it through, there was no way. They...they were everywhere. Not even a mouse could have gotten through.

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 Output





DEFT Example Spanish

Source Material

A: ¿Dónde estabas? Te esperamos todo el día y nunca llegaste.

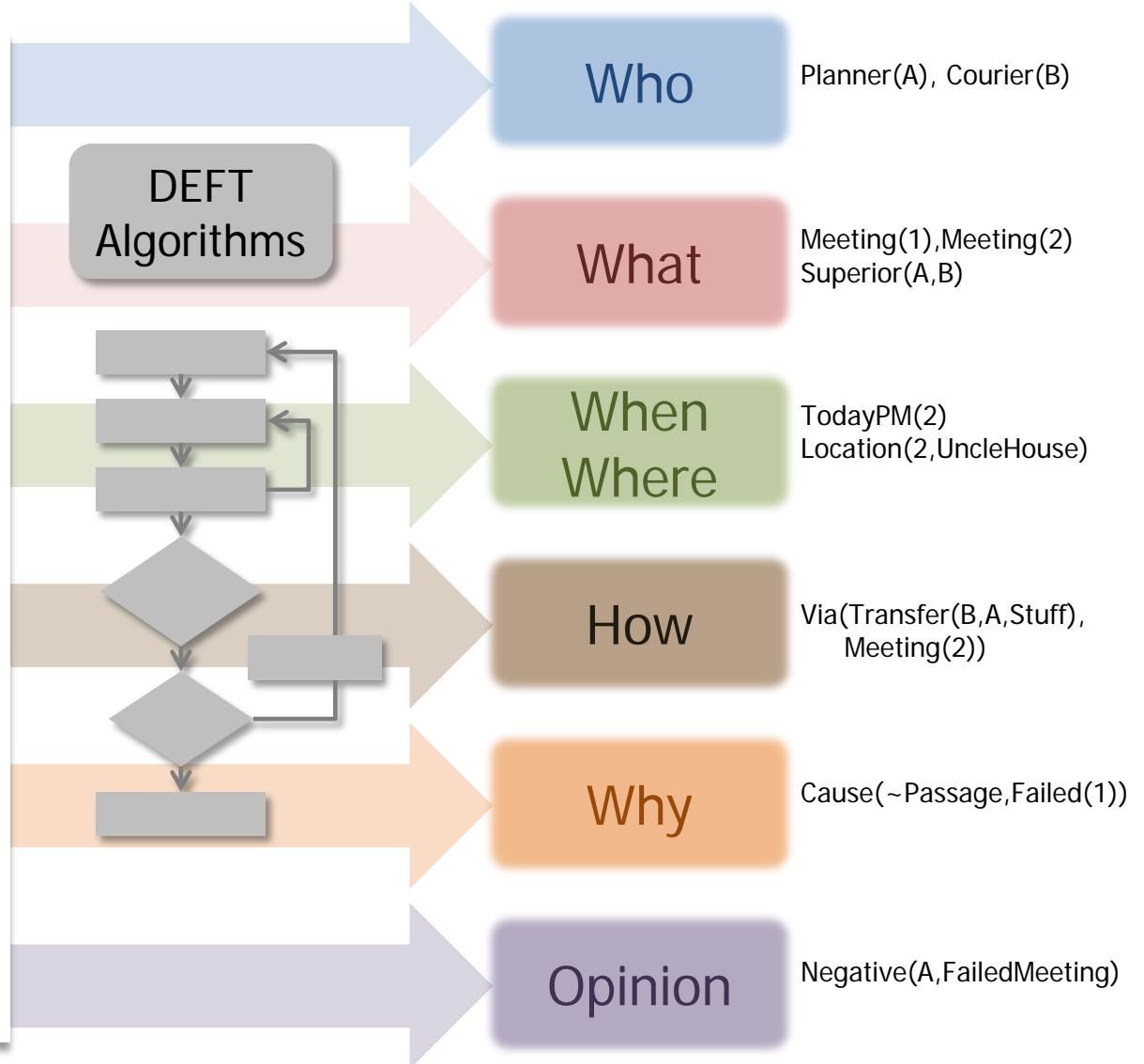
B: No pude venir, no había pasaje. Ellos ... ellos estaban por todas partes. Ni siquiera un ratón podría pasar.

A: Debieras haber encontrado pasaje. Sabes que tenemos las cosas para la ... la fiesta mañana por la noche. Necesitamos otro lugar reunir... esta noche. ¿Y este ... este ... la casa? Ya sabes, donde nos reunimos la última vez.

B: ¿Te refieres a la casa de tu tío?

A: Sí, la misma que la última vez. No te olvides nada. Necesitamos todas las cosas. Ya te pagué, así que debieras cumplir. Que no \$%*! esta vez.

DEFT Output





DEFT Example Chinese

Source Material

A: 你在哪?我们等了您一整天,您都没来。

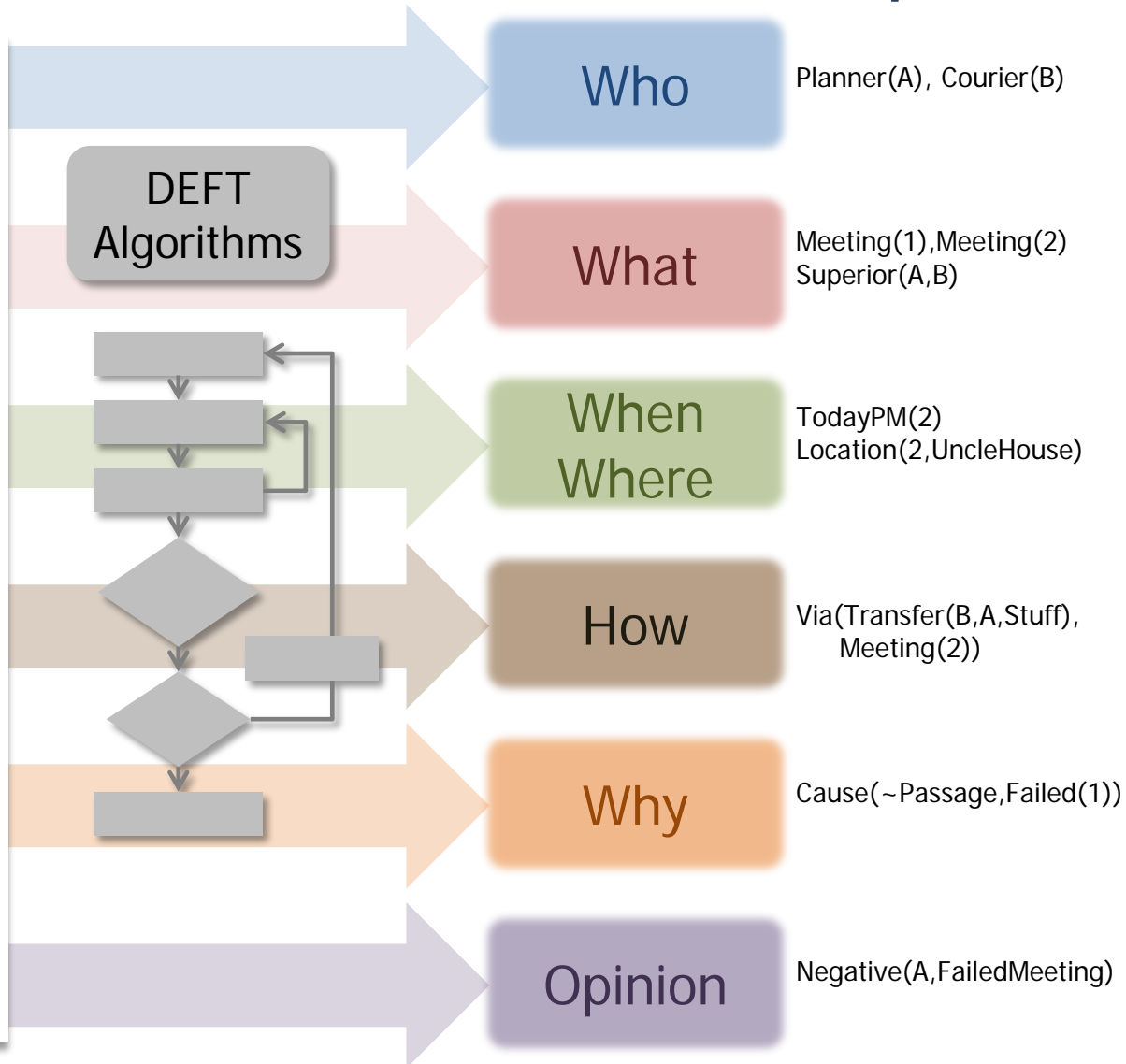
B: 我没法过去,没办法。他们...他们无所不在。连一只老鼠都没法穿过去。

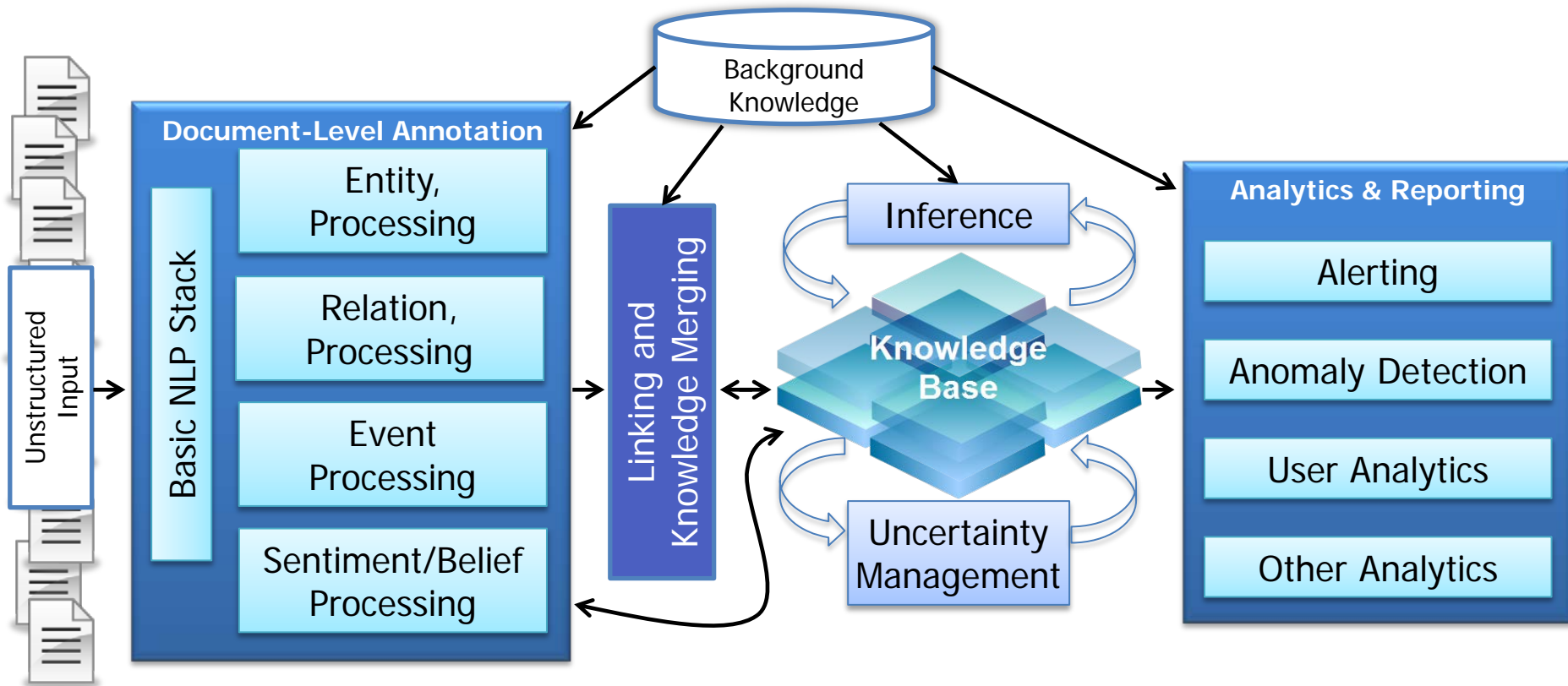
A: 你应该想个办法。你知道我们明天的...的派对需要这些东西。我们今晚需要一个新的地点见面。那间...嗯...嗯...房子怎么样?你知道的,我们上次见面的那间。

B: 你是说你叔叔家?

A: 对,跟上次一样。不要忘了任何东西。我们需要全部的东西。我已经付了你钱,所以你最好交差。你这次最好别再搞砸了。

DEFT Output





- IE / NLU / AKBC where the consumer is an automated analytic (in addition to a human) is very different than just producing output for humans
- Decision about reifying events vs. representing them through their arguments has major KB implications
- Streaming architecture has major algorithm implications



- **Coreference Resolution**
 - Inducing entity and event matches without direct explicit matches
 - Cross-document tracking of entities, relations, and events
- **Entities, Relations, and Events**
 - Explicit and implicit relations and learning of new expressions
 - Automatic representation of temporal, enabling, and attributive aspects
 - Recovery and representation of implicit arguments
- **Knowledge Bases**
 - Dynamic push and pull from active knowledge bases
 - How do DEFT algorithms add up to a knowledge base?
- **Opinions and Modality**
 - Infer opinion/modality, topic, attribution and change over time
 - Tackle ground-truth challenges (probabilistic, many annotators?)
 - Can belief/sentiment/etc. be represented as confidence in KB entries?
- **Novelty and Anomaly**
 - Prioritize data based on novelty and anomaly



New DEFT Focus on Events

- Event Issues
 - Granularity of event objects
 - Mention predicates for non-named entities
 - Representation of realis information, actual, generic, hypothetical, negated
 - All events, not just specific types
 - Temporal, causal, sub-event relations
 - Recovery and representation of implicit arguments
- Future Event Directions
 - Event hoppers/buckets? Instead of marking relations between events, put events in hoppers/buckets and relate the hoppers/buckets?
 - RED layered on top of AMR?





The bomb exploded in a crowded marketplace. Five civilians were killed, including two children. Al Qaeda claimed responsibility.

Killed by whom? Responsibility for what?

Implicit arguments are important.

*Saucedo said that **guerrillas** in one car opened fire on police standing guard, while a second car carrying 88 pounds (40 kgs) of dynamite parked in front of the building, and a third car rushed the **attackers** away.*

Inferences can inform coreference.

Abstract Meaning Representation:

- guerilla – Arg0 of fire.01
- attacker – person – Arg0-of attack.01

Examples courtesy of Vivek Srikumar and Martha Palmer

Goal: Develop awareness about situations of concern to the US government based on partial information from multiple language, image, video, human, and structured sources – information that, in isolation, is incomplete, unreliable and/or cannot be fully analyzed with today's capabilities



CNS photo/Paul Haring

Microblog post

@news_va_en I am sorry that our [@Pontifex](#) do not put more emphasis on violence that Mexico suffers

News

Our lead video shows the Pope driving past throngs of people in León as he is transported in his signature 'Popemobile'.



Manuel Balce Ceneta/AP

Microblog post

Bin Laden's death doesn't mean we shouldn't still live in constant fear for our lives

Social media site

Do you believe Bin Laden was buried at sea? Something fishy about that..



Jewish News One

Audio from Video

Death to America. We want to keep our nuclear program

Message

Thousands of Iranians Shout Death to America, Burn Flags in Rallies ...

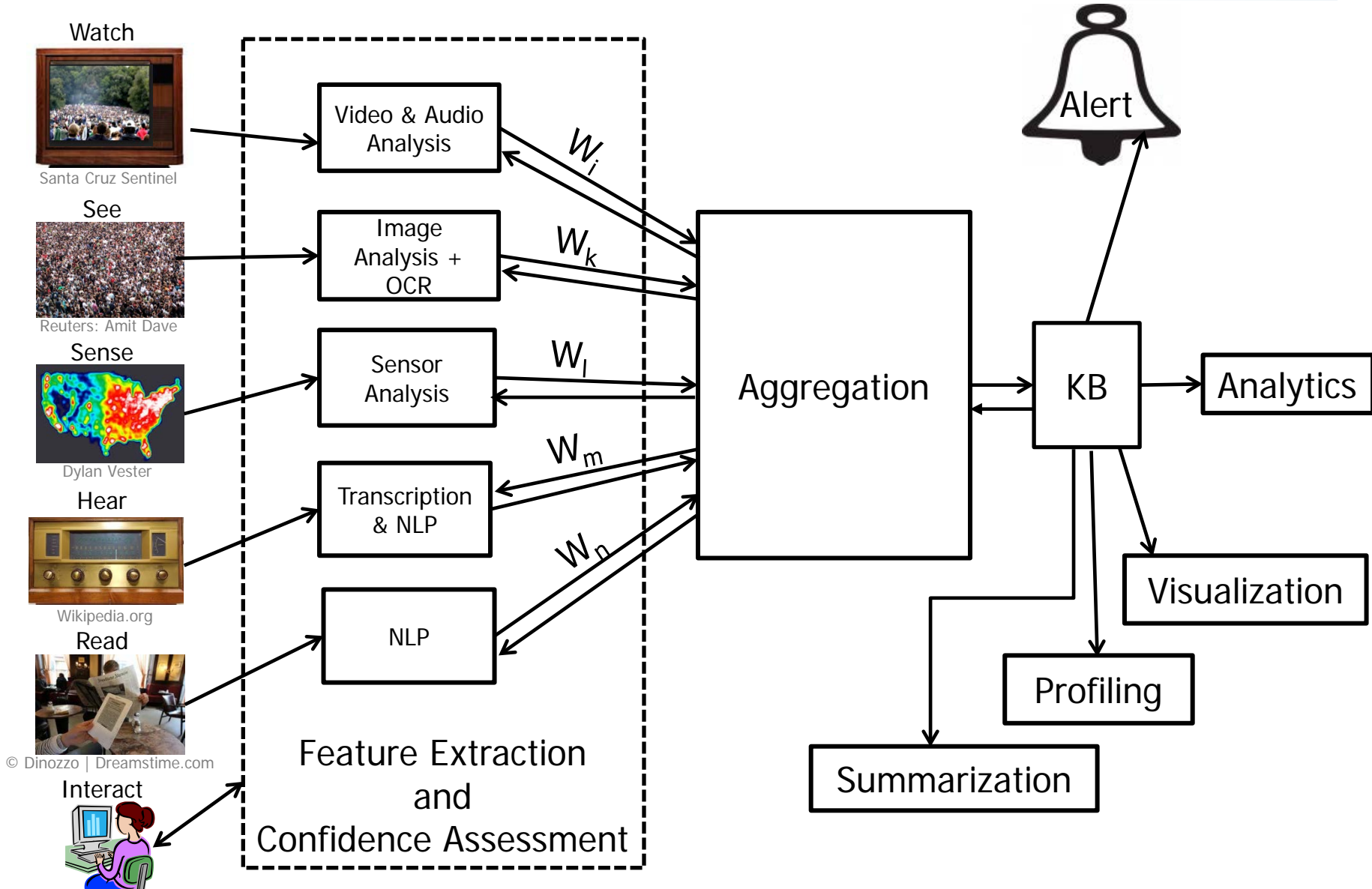


AKBC Challenges for the Future (outside of DEFT)

- Increasing number of media types, genres
- Stovepiped data flows
- Inadequate infrastructure for knowledge sharing across media
- Absence of a common language for knowledge sharing across media
- None of the media-specific technologies alone give high-fidelity situation awareness (because of accuracy and because of data completeness)
 - Image/Video (VIRAT, Mind's Eye, VMR, MADCAT, VACE, ALADDIN, JANUS)
 - Clustering
 - Scene Analysis
 - Event and Entity detection
 - In-scene text localization and OCR
 - Language (GALE, BOLT, DEFT, MEMEX, BABEL)
 - Entity, event, and relation discovery
 - Sentiment analysis
 - Metadata (DCAPS, SMISC)
 - Geo-location
 - Sensor data
 - Fusion (Insight)



Multimedia KB Architecture Concept



© Dinozzo | Dreamstime.com

- Joint inference:
 - Compute the most likely hypothesis from all observations
- Improve recognition based on dynamically-adjusted priors:
 - (Partial or complete) information from one medium adjusts the priors for other media
 - Either in a streaming model, or iterative "easy first" model
- Multiple flows:
 - Feed-back data path (adjustment of priors)
 - Feed-forward path (combination of evidence)



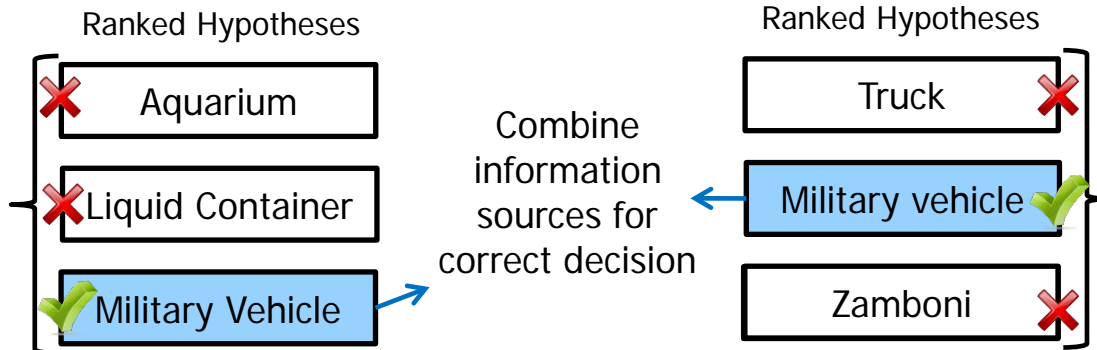
©Cision

PERSON: "Fred Brown"

PERSON: ??

Narrow the search space and reduce false alarms
(seed gallery with 1- or 2- hop social graph neighborhood of known individual)

There's a tank in front of my house!



Reuters



www.darpa.mil