Beyond Jeopardy!TM Adapting WatsonTM to new domains using Distributional Semantics

Alfio Massimiliano Gliozzo (gliozzo@us.ibm.com)
IBM T.J. Watson Research Center
Watson Technologies
Yorktown Heights, NY

Outline

- Watson™ and the JEOPARDY!™ challenge

- Distributional Semantics for Domain Adaptation

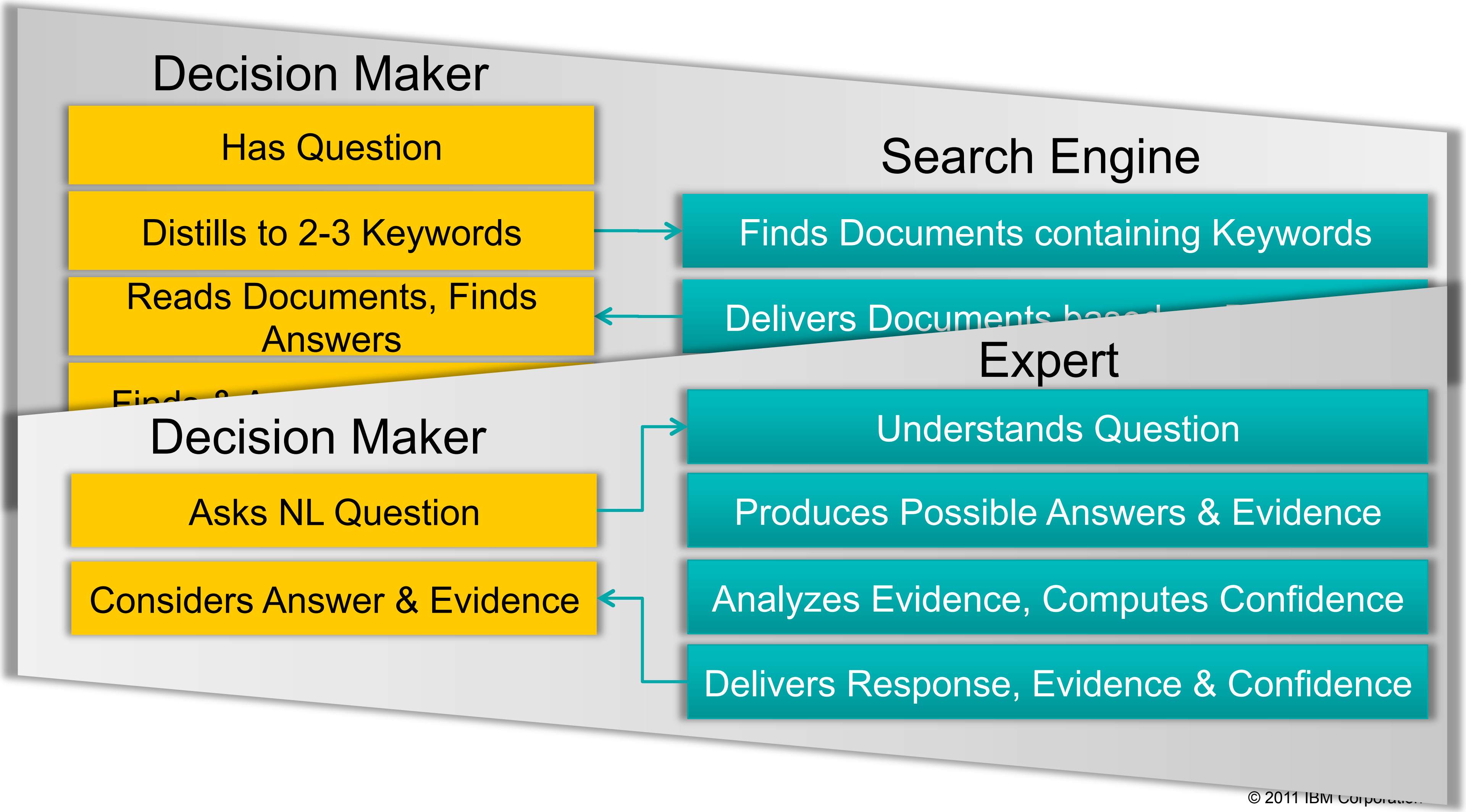
Automatic Open-Domain Question Answering

A Long-Standing Challenge in Artificial Intelligence to emulate human expertise

- Given
  - Rich **Natural Language Questions**
  - Over a **Broad Domain of Knowledge**

- Deliver
  - **Precise Answers:** Determine what is being asked & give precise response
  - **Accurate Confidences:** Determine likelihood answer is correct
  - **Consumable Justifications:** Explain why the answer is right
  - **Fast Response Time:** Precision & Confidence in <3 seconds

Informed Decision Making: Search vs. Expert Q&A



Decision Maker

Has Question

Distills to 2-3 Keywords

Reads Documents, Finds Answers

Finds & A

Search Engine

Finds Documents containing Keywords

Delivers Documents based

Decision Maker

Asks NL Question

Considers Answer & Evidence

Expert

Understands Question

Produces Possible Answers & Evidence

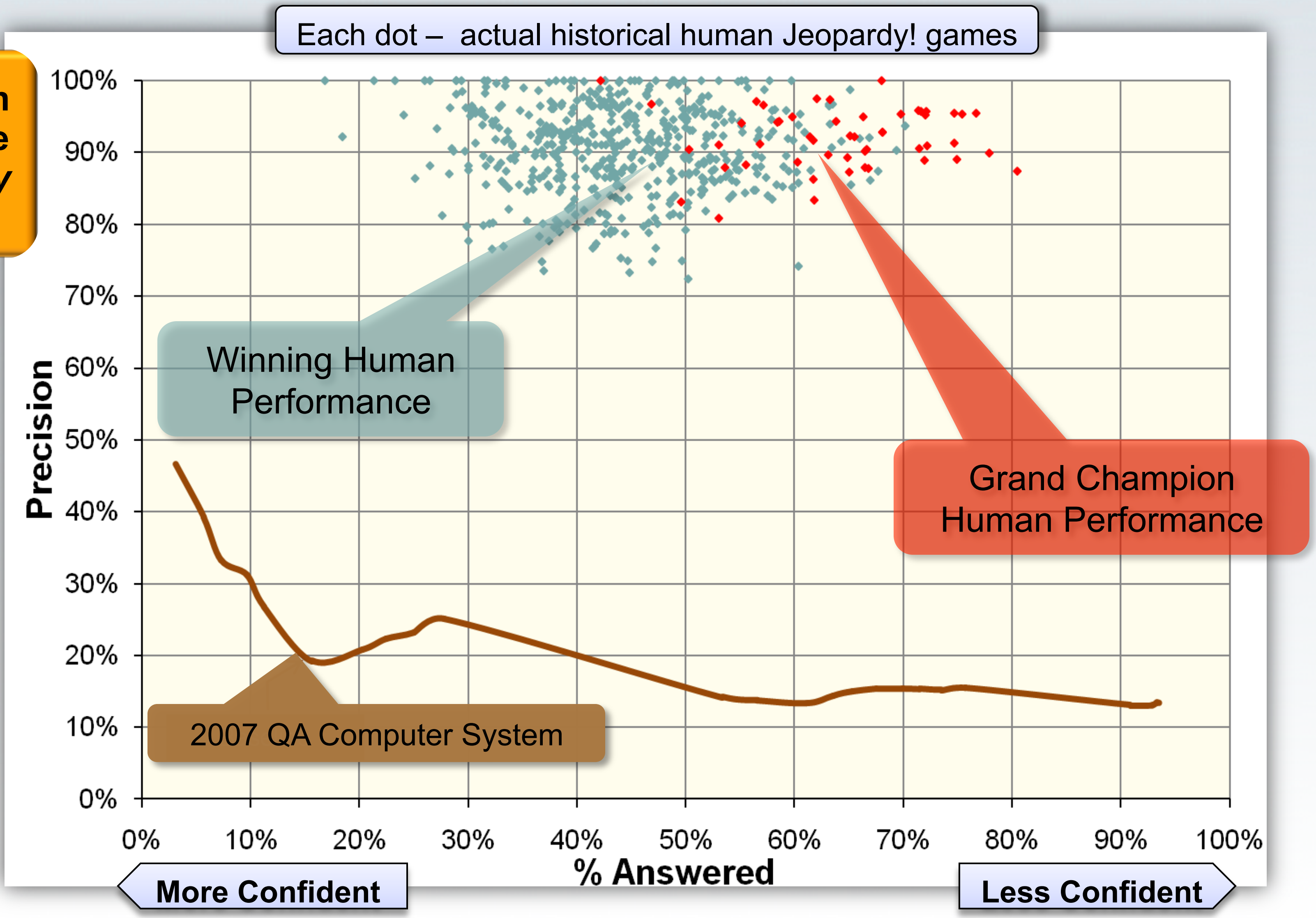
Analyzes Evidence, Computes Confidence

Delivers Response, Evidence & Confidence

What It Takes to compete against Top Human Jeopardy!^{TM} Players

Our Analysis Reveals the ***Winner's Cloud***

Top human players are *remarkably* good.



Each dot – actual historical human Jeopardy! games

Winning Human Performance

Grand Champion Human Performance

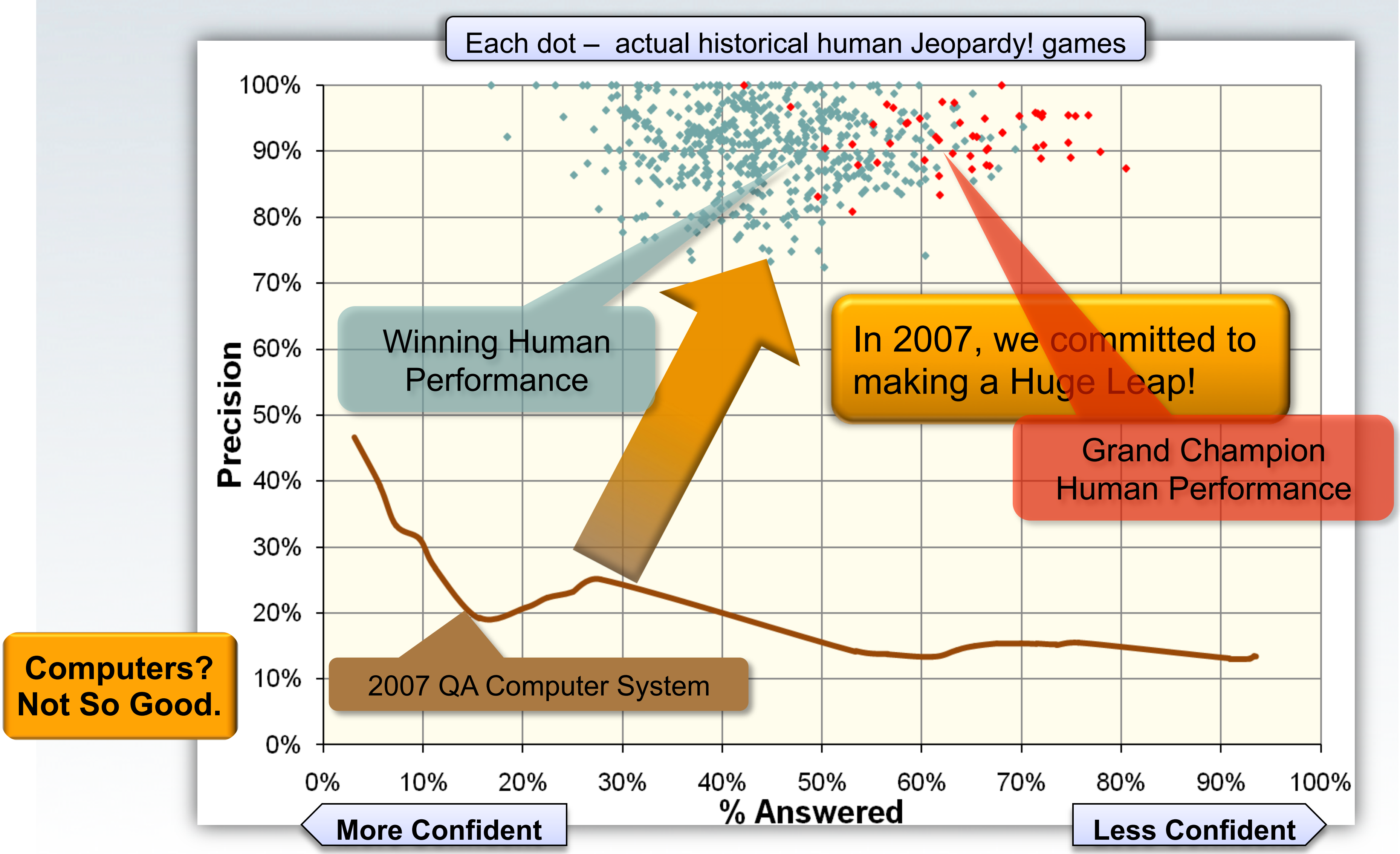
2007 QA Computer System

More Confident

Less Confident

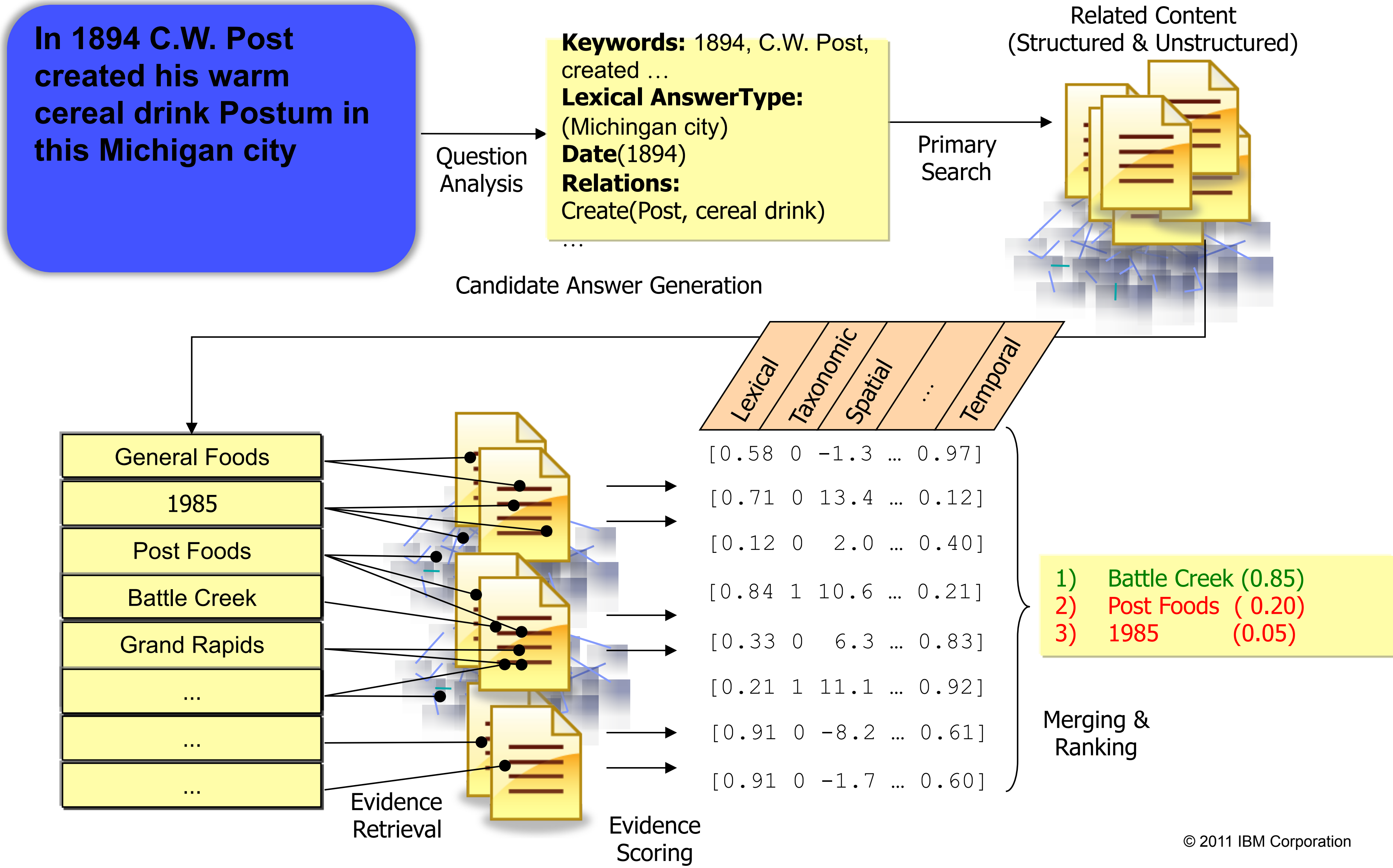
What It Takes to compete against Top Human Jeopardy!^{TM} Players

*Our Analysis Reveals the **Winner's Cloud***



Computers? Not So Good.

Example Question



In 1894 C.W. Post created his warm cereal drink Postum in this Michigan city

Question Analysis

Keywords: 1894, C.W. Post, created …
Lexical AnswerType: (Michigan city)
Date(1894)
Relations: Create(Post, cereal drink)
…

Primary Search

Related Content (Structured & Unstructured)

Candidate Answer Generation

General Foods
1985
Post Foods
Battle Creek
Grand Rapids
...
...
...

Evidence Retrieval

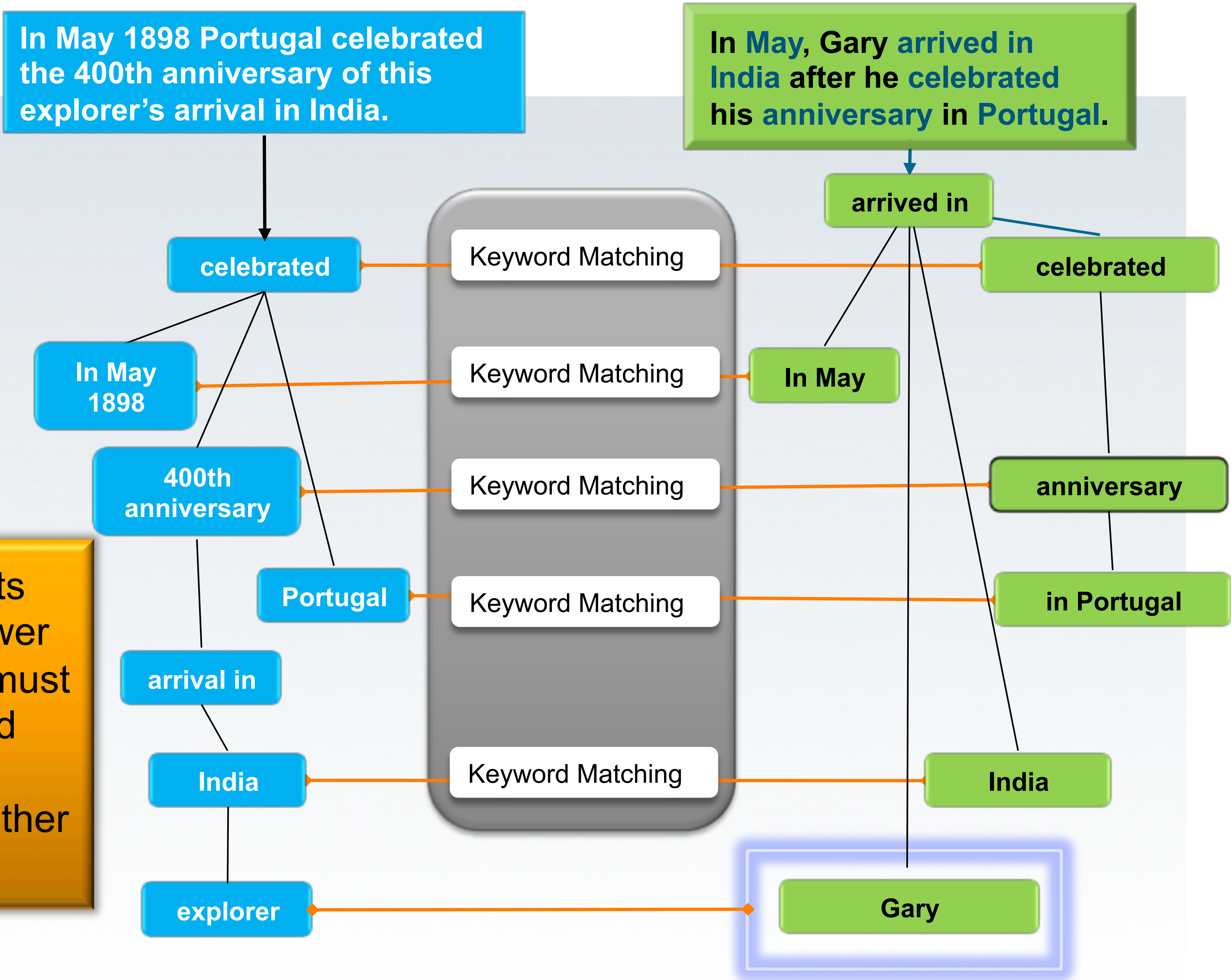
Evidence Scoring

```
[0.58 0 -1.3 … 0.97]
[0.71 0 13.4 … 0.12]
[0.12 0  2.0 … 0.40]
[0.84 1 10.6 … 0.21]
[0.33 0  6.3 … 0.83]
[0.21 1 11.1 … 0.92]
[0.91 0 -8.2 … 0.61]
[0.91 0 -1.7 … 0.60]
```

1) Battle Creek (0.85)
2) Post Foods (0.20)
3) 1985 (0.05)

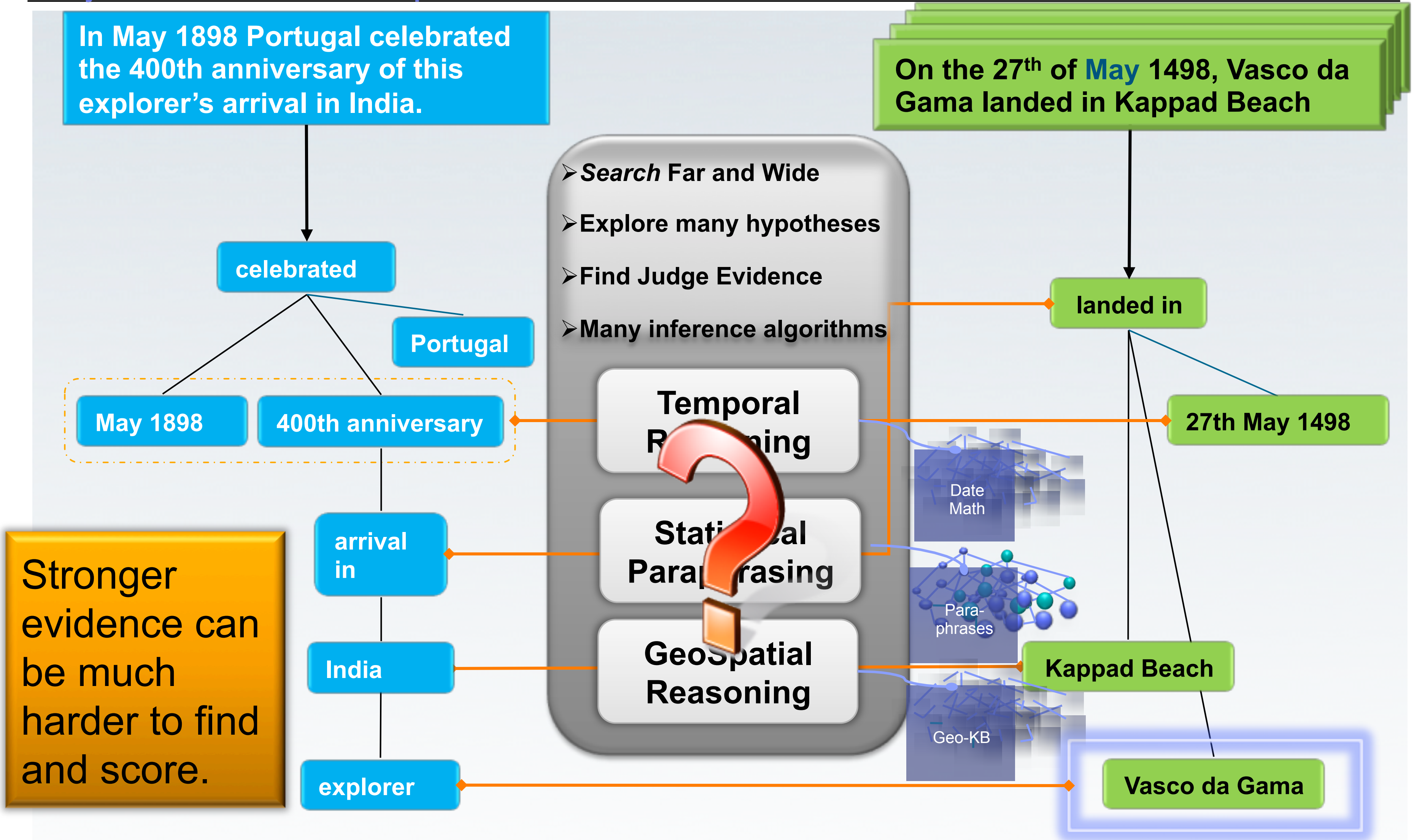
Merging & Ranking

Keyword Evidence



Evidence suggests "Gary" is the answer BUT the system must learn that keyword matching may be weak relative to other types of evidence

Why Semantics? Deeper Evidence



In May 1898 Portugal celebrated the 400th anniversary of this explorer's arrival in India.

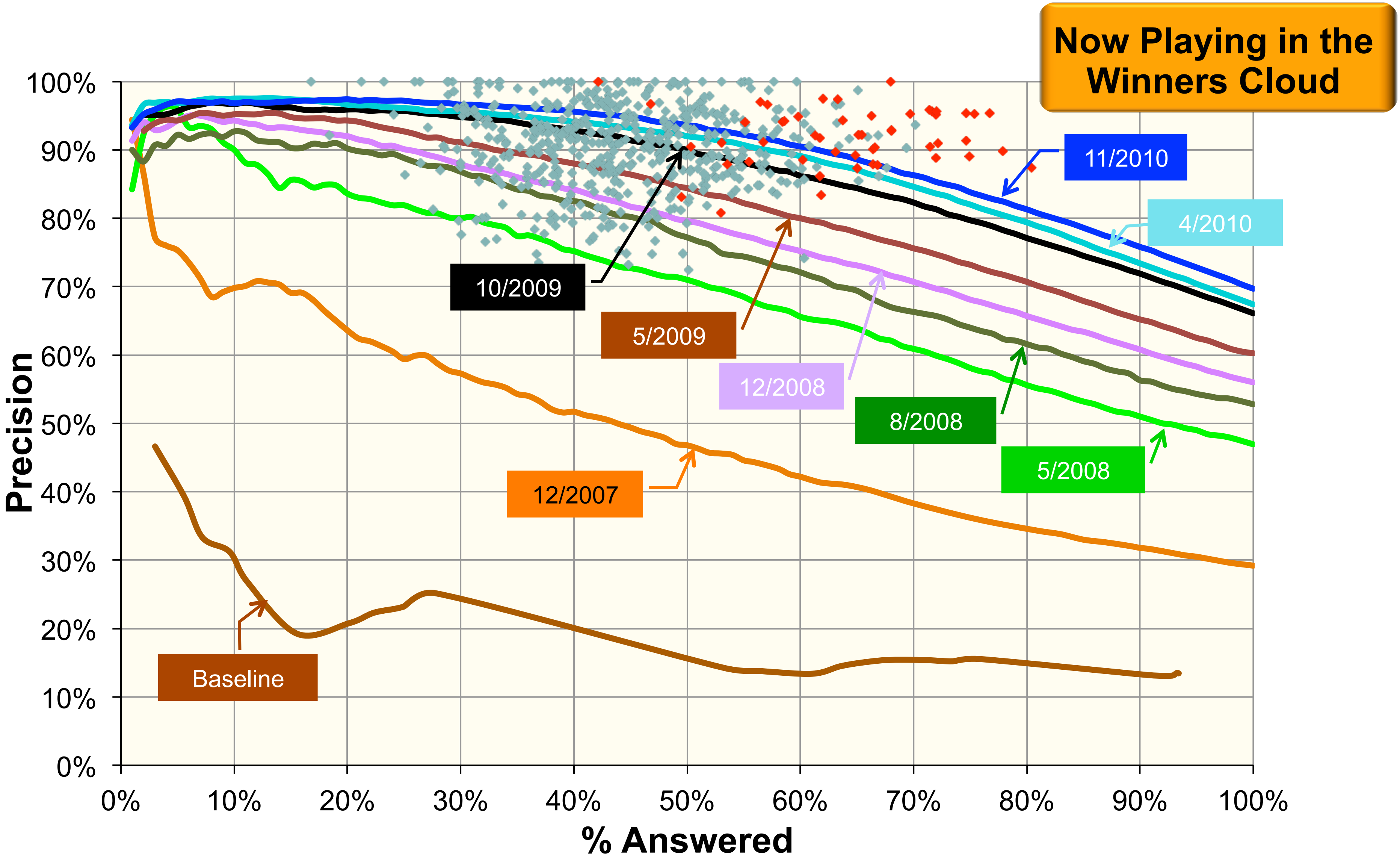
On the 27th of May 1498, Vasco da Gama landed in Kappad Beach

- *Search* Far and Wide
- Explore many hypotheses
- Find Judge Evidence
- Many inference algorithms

Stronger evidence can be much harder to find and score.

The evidence is still not 100% certain.

Compare Experiments



Outline

- Watson^{TM} and the JEOPARDY!^{TM} challenge
- Distributional Semantics for Domain Adaptation

Adaptation: What do we have in a new domain?

Content Adaptation

New Text Content

Structure and ingest text content

Training Adaptation

New "Questions"

Train the system on target scenarios

58-year-old woman presenting to her primary care
physician after several days of dizziness, anorexia,
dry mouth, increased
She had also had a f
would "get stuck" wh
reported no pain in h
no cough, shortness
Her family history inc
in her mother, Grave

What inflammation is characterized by nasal mucosal atrophy and foul-smelling crusts in the nasal passages?

Functional Adaptation

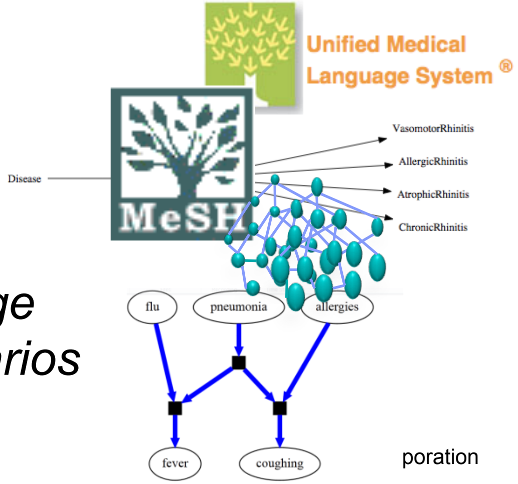
New Concepts / Reasoning / Discourse

Enhance the functional capabilities with domain-specific

- Concepts: entities, relations from domain modeling

- Reasoning: domain axioms and background knowledge

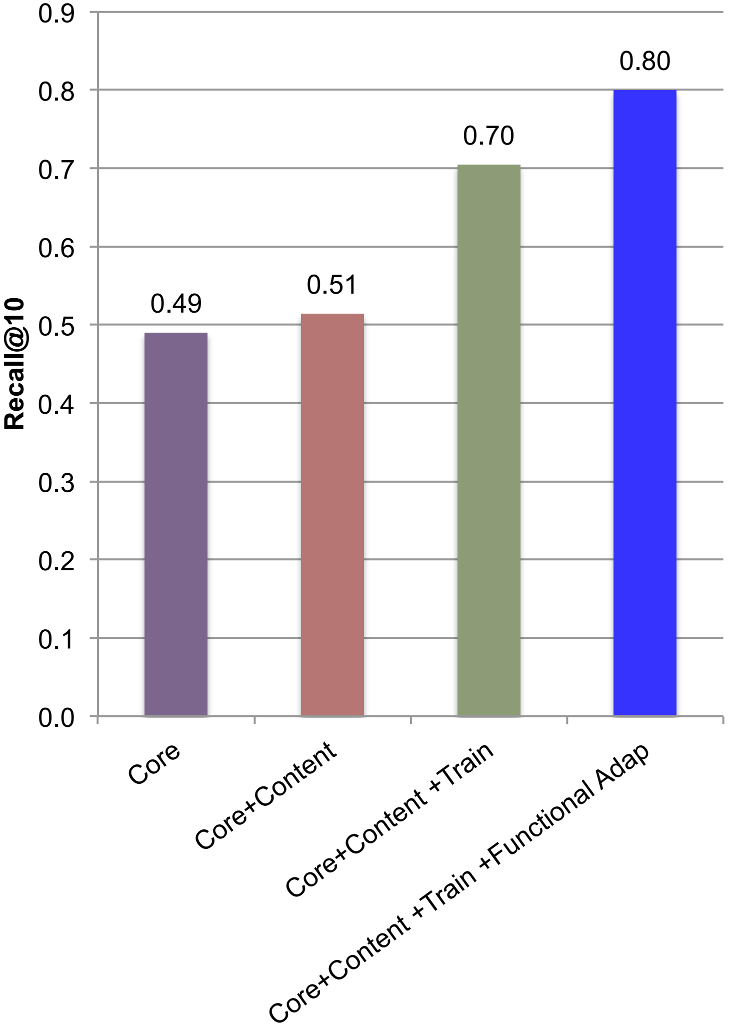
- Discourse: algorithms for domain text / problem scenarios

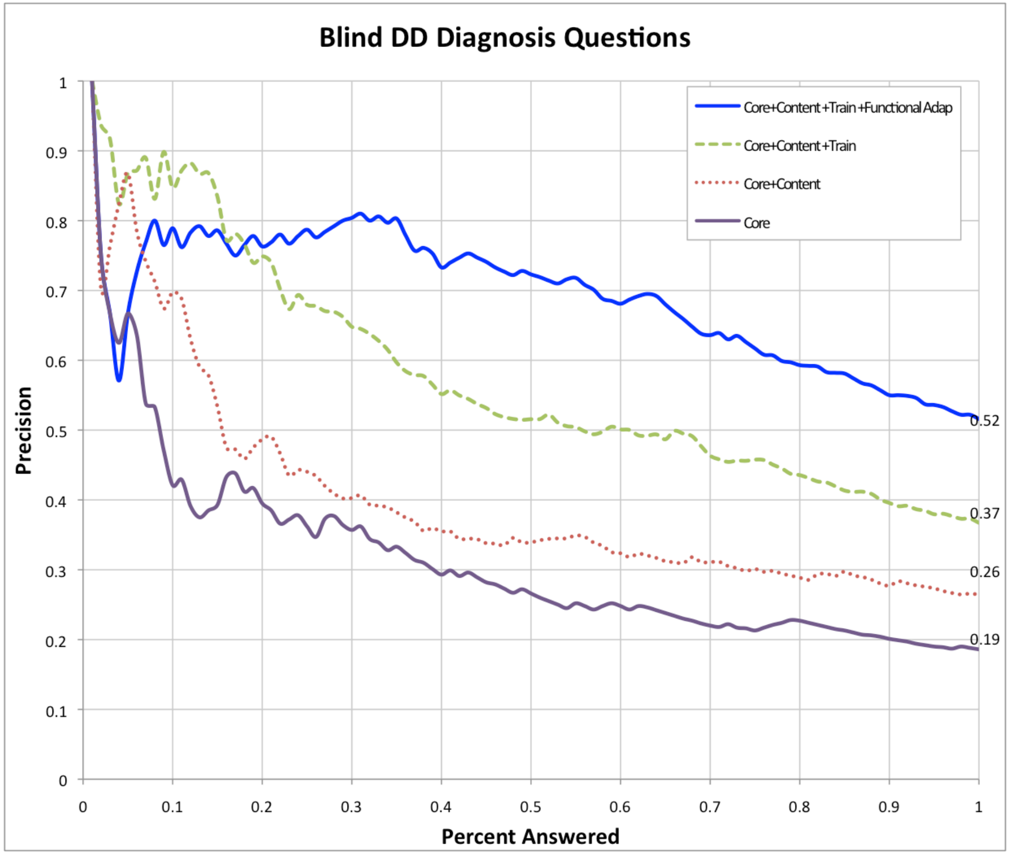


poration

Medical Adaptation - Results

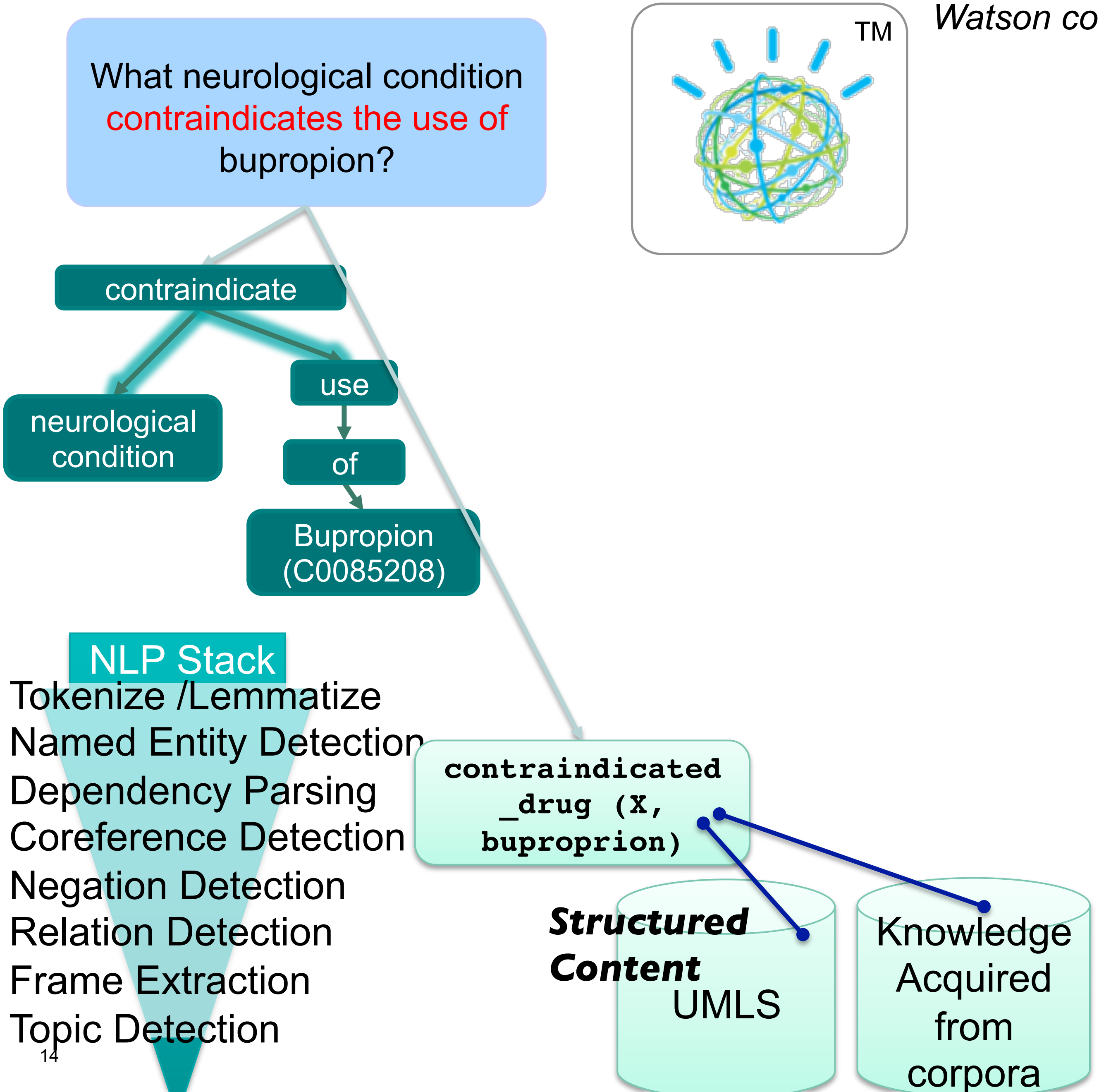
Accuracy: 52%





Watson uses an ensemble of NLP techniques justify an answer

Watson considers…



What neurological condition contraindicates the use of bupropion?

TM

contraindicate

neurological condition

use

of

Bupropion (C0085208)

NLP Stack

- Tokenize /Lemmatize
- Named Entity Detection
- Dependency Parsing
- Coreference Detection
- Negation Detection
- Relation Detection
- Frame Extraction
- Topic Detection

```
contraindicated
_drug (X,
buproprion)
```

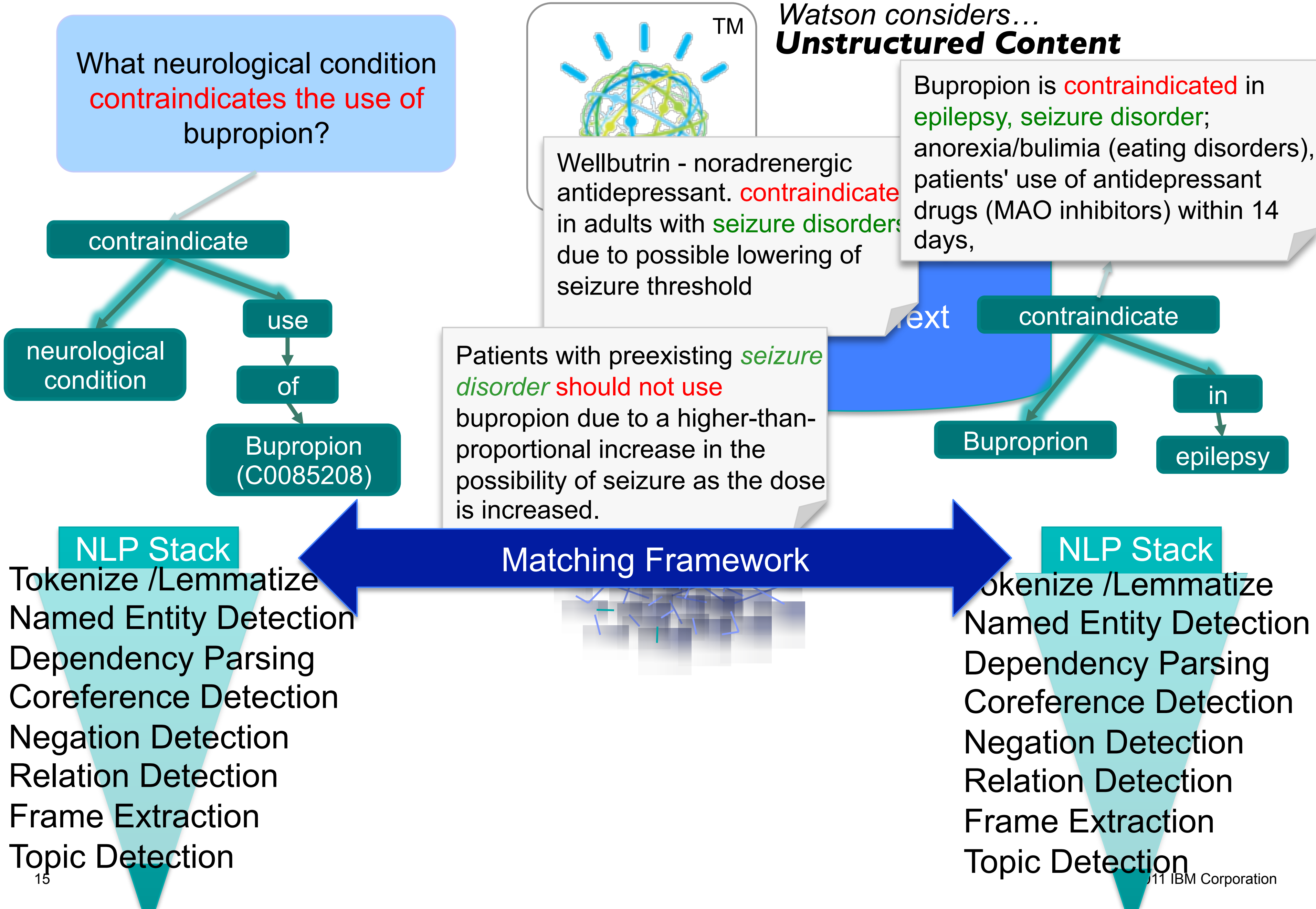
Structured Content

UMLS

Knowledge Acquired from corpora

Watson uses an ensemble of NLP techniques justify an answer

IBM WATSON



What neurological condition contraindicates the use of bupropion?

Watson considers…

Unstructured Content

Bupropion is contraindicated in epilepsy, seizure disorder; anorexia/bulimia (eating disorders), patients' use of antidepressant drugs (MAO inhibitors) within 14 days,

Wellbutrin - noradrenergic antidepressant. contraindicate in adults with seizure disorders due to possible lowering of seizure threshold

Patients with preexisting *seizure disorder* should not use bupropion due to a higher-than-proportional increase in the possibility of seizure as the dose is increased.

NLP Stack

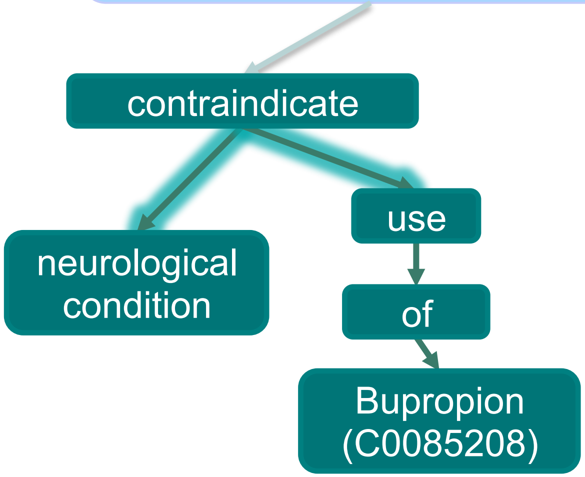
- Tokenize /Lemmatize
- Named Entity Detection
- Dependency Parsing
- Coreference Detection
- Negation Detection
- Relation Detection
- Frame Extraction
- Topic Detection

NLP Stack

- okenize /Lemmatize
- Named Entity Detection
- Dependency Parsing
- Coreference Detection
- Negation Detection
- Relation Detection
- Frame Extraction
- Topic Detection

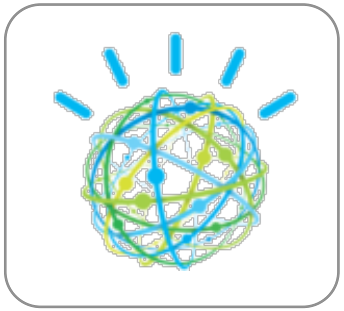
Watson uses an ensemble of NLP techniques justify an answer

What neurological condition contraindicates the use of bupropion?



NLP Stack

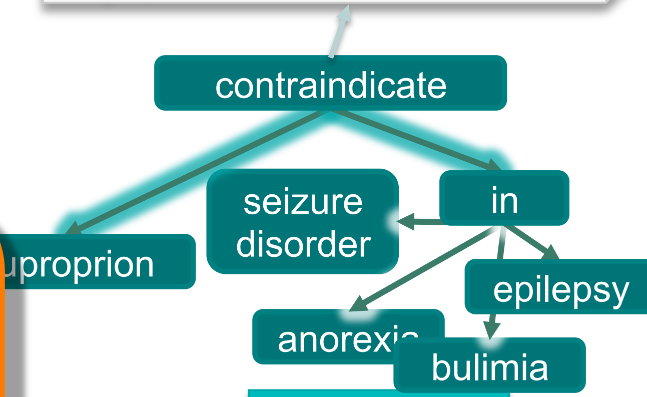
- Tokenize /Lemmatize
- Named Entity Detection
- Dependency Parsing
- Coreference Detection
- Negation Detection
- Relation Detection
- Frame Extraction
- Topic Detection



Watson considers…

Unstructured Content

Bupropion is contraindicated in epilepsy, seizure disorder; anorexia/bulimia (eating disorders), patients' use of antidepressant drugs (MAO inhibitors) within 14 days,



NLP Stack

- Tokenize /Lemmatize
- Named Entity Detection
- Dependency Parsing
- Coreference Detection
- Negation Detection
- Relation Detection
- Frame Extraction
- Topic Detection

Need to consider the type ("**neurological condition**") of the answer for possible candidates:

- Epilepsy
- Seizure disorder
- Anorexia
- Bulimia

Structured Content

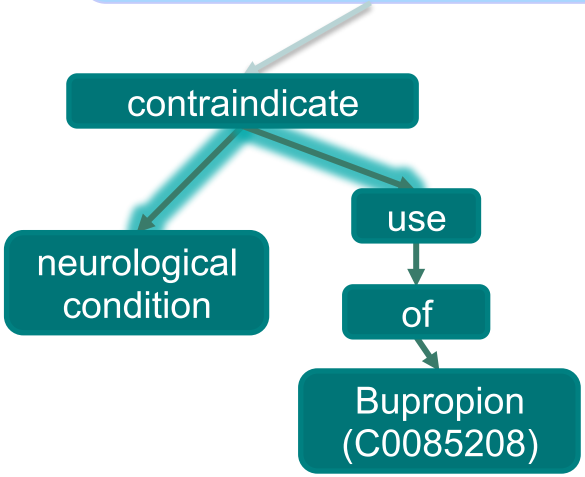
UMLS

Knowledge Acquired from corpora

Watson uses an ensemble of NLP techniques justify an answer

IBM WATSON

What neurological condition contraindicates the use of bupropion?



TM

Watson considers…
Unstructured Content

Wellbutrin - noradrenergic antidepressant. contraindicated in adults with seizure disorders due to possible lowering of seizure threshold

Use background medical knowledge (**Wellbutrin** is a brand name of **bupropion**)

NLP Stack

Tokenize /Lemmatize
Named Entity Detection
Dependency Parsing
Coreference Detection
Negation Detection
Relation Detection
Frame Extraction
Topic Detection

Structured Content

UMLS

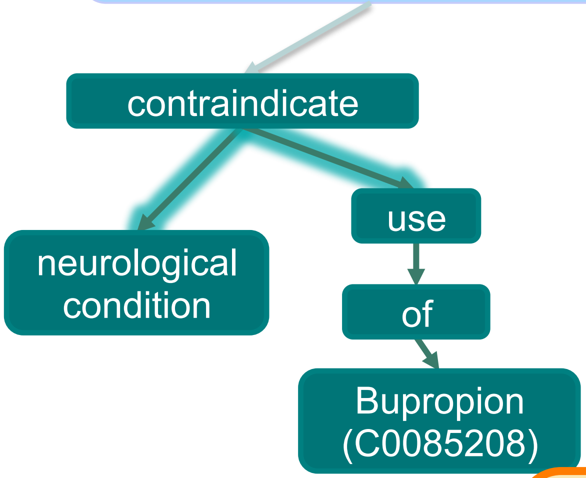
Knowledge Acquired from corpora

NLP Stack

Tokenize /Lemmatize
Named Entity Detection
Dependency Parsing
Coreference Detection
Negation Detection
Relation Detection
Frame Extraction
Topic Detection

Watson uses an ensemble of NLP techniques justify an answer

What neurological condition contraindicates the use of bupropion?



NLP Stack

Tokenize /Lemmatize
Named Entity Detection
Dependency Parsing
Coreference Detection
Negation Detection
Relation Detection
Frame Extraction
Topic Detection

Consider paraphrases in medical language:
(should not use = contraindicate)

Patients with preexisting *seizure disorder* should not use bupropion due to a higher-than-proportional increase in the possibility of seizure as the dose is increased.

NLP Stack

Tokenize /Lemmatize
Named Entity Detection
Dependency Parsing
Coreference Detection
Negation Detection
Relation Detection
Frame Extraction
Topic Detection

Structured Content

UMLS

Knowledge Acquired from corpora

Early Medical Adaptation Lessons

- Domain adaptation is difficult!

- Requires:
  - Deeply skilled research team across all the key disciplines (ML, NLP, IR, KR)
  - Domain Experts (Doctors) for annotation/vetting and design reasoning strategies
  - Collaboration between the two groups !
  - Background knowledge for new domains (e.g. UMLS) and analytics exploiting that
  - Rigorous methodological discipline (e.g., blind test!)

- Future Challenge: Scalable and cost effective functional adaptation process
  - Acquiring Domain Knowledge from Text
    - Taxonomy induction
    - Statistical Paraphrasing
    - Sense Induction/ Unsupervised WSD
  - Using the same analytics (e.g. matching, tycor) across domains

- We call it Distributional Semantics!

The Distributional Semantics Paradigm

- The challenge: Fully Unsupervised Computational Semantics
  - Input: few Gigabytes of raw text in a specific domain
  - Output: Semantic Analyzer having the following capabilities
    - Term/Text Similarity beyond Keyword Matching
    - WSD, Lexical Substitution
    - Matching: terms, relations
    - Linking text to knowledge bases
  - Radical Approach:
    - Mining (clustering) big data
    - No Rules, No labeled data
- Making it scalable (Hadoop)
  - More text = more hardware = same time
  - Fast semantic parsing
  - Web size Distributional Semantics to capture background knowledge

Distributional Hypothesis and Structuralism

The **Distributional Hypothesis** in linguistics is the theory that words that occur in similar contexts tend to have similar meanings (paradigmatic relations).

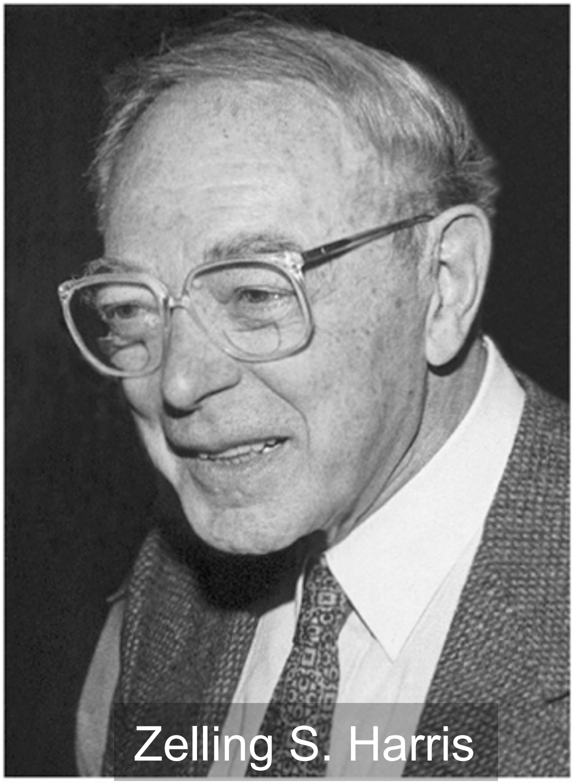
The Distributional Hypothesis is the basis for Distributional Semantics.

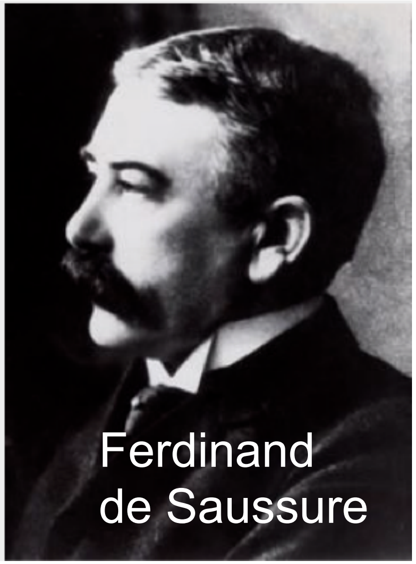
It states that the meaning of a word can be defined in terms of its context (properties).

- other words in the same sentence/document (bag of words)
- words in the immediate neighbors
- words along dependency paths
- Predicate Argument Structure
- Frame

➔ any process that builds a structure on sentences can be used as a source for properties

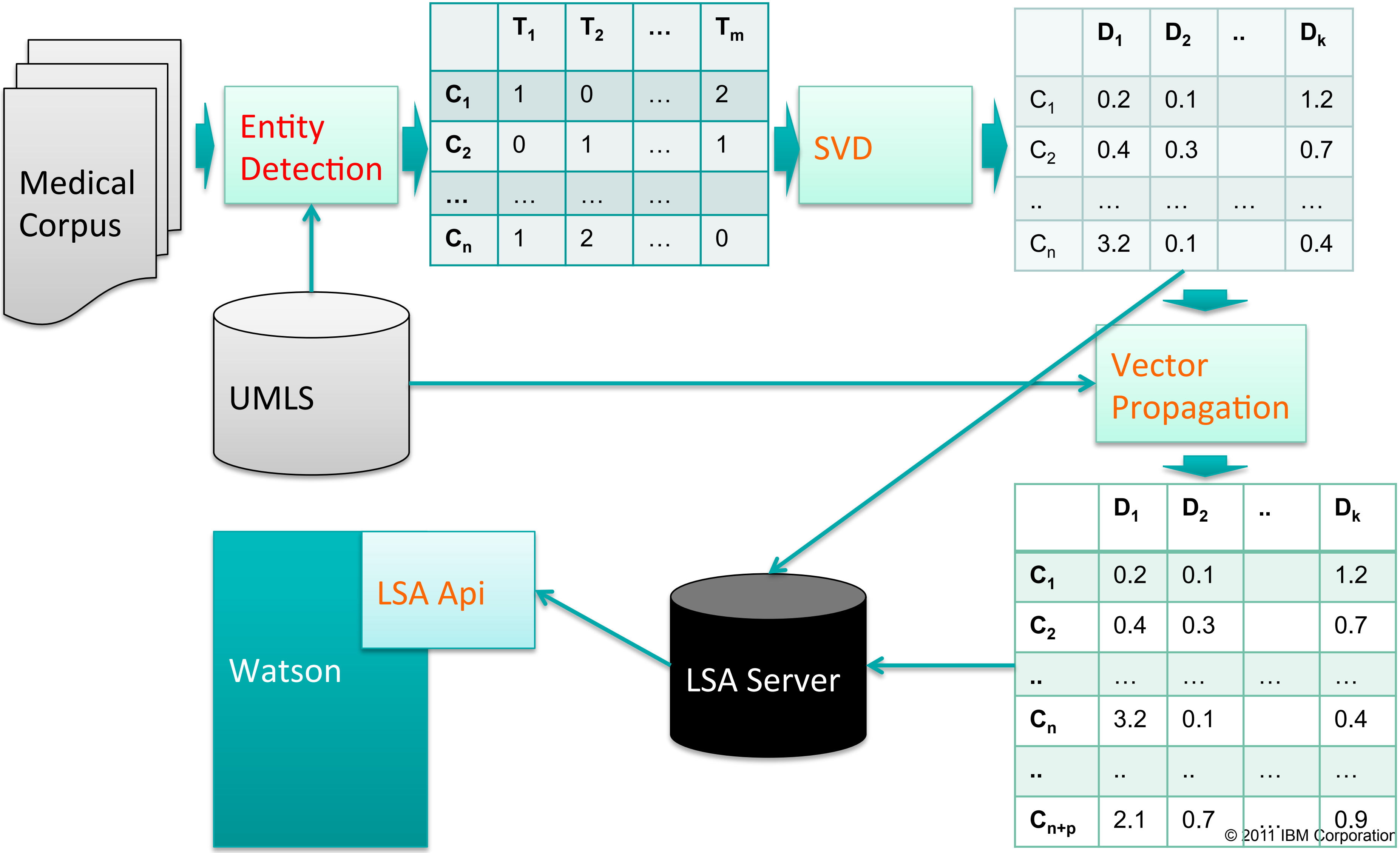
Z. Harris. (1954). Distributional Structure. Word 10 (2/3)


Zelling S. Harris


Ferdinand de Saussure

Distributional Semantics: Early Steps

Latent Semantic Analysis (2.0)



	T_1	T_2	…	T_m
C_1	1	0	…	2
C_2	0	1	…	1
…	…	…	…	
C_n	1	2	…	0

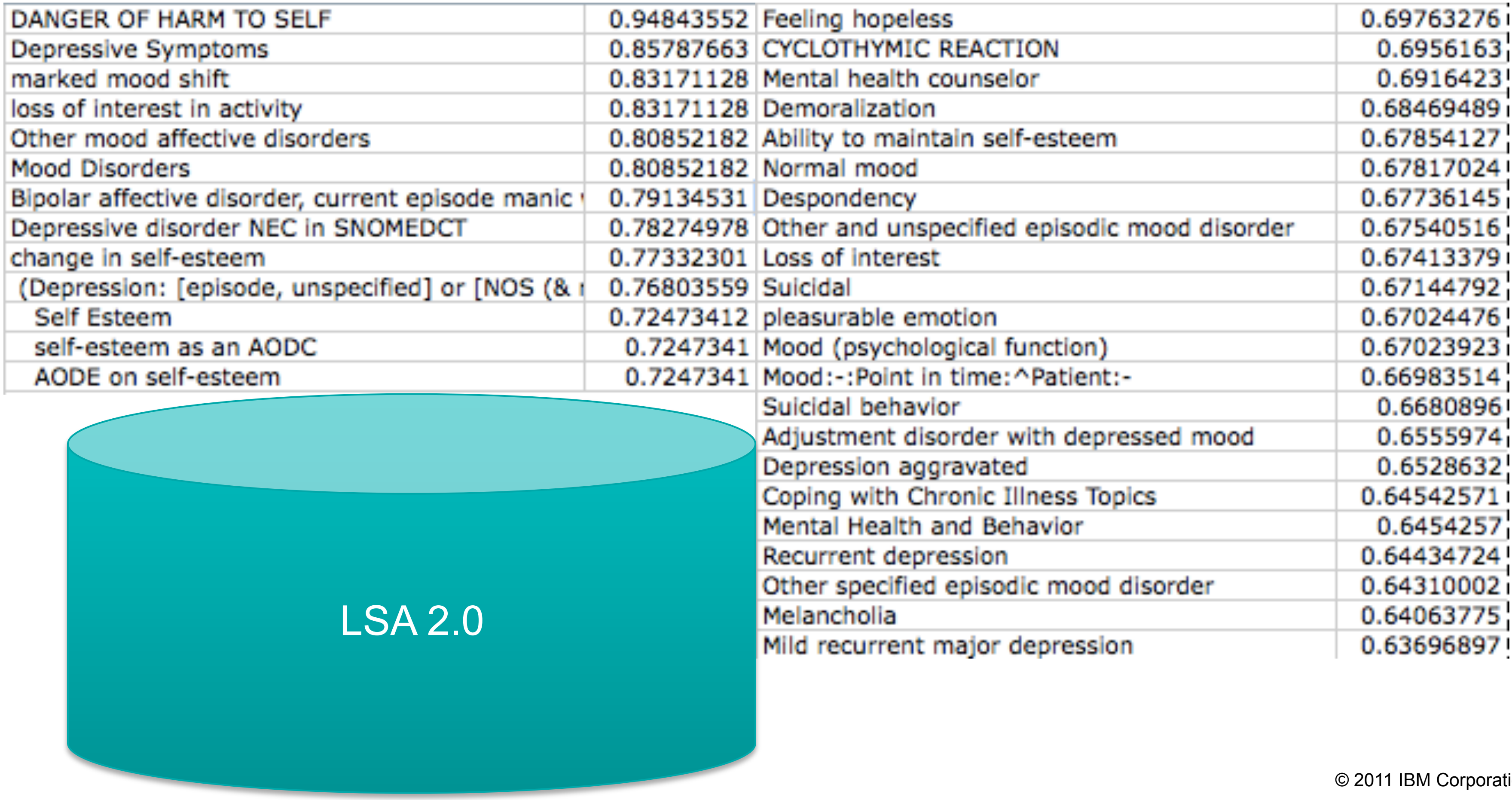
	D_1	D_2	..	D_k
C_1	0.2	0.1		1.2
C_2	0.4	0.3		0.7
..	…	…	…	…
C_n	3.2	0.1		0.4

	D_1	D_2	..	D_k
C_1	0.2	0.1		1.2
C_2	0.4	0.3		0.7
..	…	…	…	…
C_n	3.2	0.1		0.4
..	..	..	…	…
C_{n+p}	2.1	0.7	…	0.9

LSA 2.0: Factors related to suicide

	Accuracy	Precision@70
LSA	+0.66%	+0.47%
LSA 2.0	+1.13% (4.46%)	+1.58%(5.229%)

“Suicide”



DANGER OF HARM TO SELF	0.94843552	Feeling hopeless	0.69763276
Depressive Symptoms	0.85787663	CYCLOTHYMIC REACTION	0.6956163
marked mood shift	0.83171128	Mental health counselor	0.6916423
loss of interest in activity	0.83171128	Demoralization	0.68469489
Other mood affective disorders	0.80852182	Ability to maintain self-esteem	0.67854127
Mood Disorders	0.80852182	Normal mood	0.67817024
Bipolar affective disorder, current episode manic	0.79134531	Despondency	0.67736145
Depressive disorder NEC in SNOMEDCT	0.78274978	Other and unspecified episodic mood disorder	0.67540516
change in self-esteem	0.77332301	Loss of interest	0.67413379
(Depression: [episode, unspecified] or [NOS (&	0.76803559	Suicidal	0.67144792
Self Esteem	0.72473412	pleasurable emotion	0.67024476
self-esteem as an AODC	0.7247341	Mood (psychological function)	0.67023923
AODE on self-esteem	0.7247341	Mood:-:Point in time:^Patient:-	0.66983514
		Suicidal behavior	0.6680896
		Adjustment disorder with depressed mood	0.6555974
		Depression aggravated	0.6528632
		Coping with Chronic Illness Topics	0.64542571
		Mental Health and Behavior	0.6454257
		Recurrent depression	0.64434724
		Other specified episodic mood disorder	0.64310002
		Melancholia	0.64063775
		Mild recurrent major depression	0.63696897

LSA 2.0

The @@ operation

SENTENCE:

I suffered from a cold and took aspirin.

STANFORD COLLAPSED DEPENDENCIES:

http://nlp.stanford.edu:8080/parser/

nsubj(suffered, I); nsubj(took, I); root(ROOT, suffered); det(cold, a); prep_from(suffered, cold); conj_and(suffered, took); dobj(took, aspirin)

WORD-PROPERTY PAIRS:

suffered	nsubj(@@, I)	1
took	nsubj(@@, I)	1
cold	det(@@, a)	1
suffered	prep_from(@@, cold)	1
suffered	conj_and(@@, took)	1
took	dobj(@@, aspirin)	1

Jo	Bim	
I	nsubj(suffered, @@)	1
I	nsubj(took, @@)	1
a	det(cold, @@)	1
cold	prep_from(suffered, @@)	1
took	conj_and(suffered, @@)	1
aspirin	dobj(took, @@)	1

JoBimText

Linking Language to Knowledge with Distributional Semantics

- www.jobimtext.org
- Open Source Software
  - Apache License
  - SourceForge
- Contributors
  - TU Darmstadt, Germany, FG Language Technology
    - **Chris Biemann** (**Bim**) , Martin Riedl
  - IBM T.J. Watson Research - Watson Technologies
    - **Alfio Gliozzo** (**Jo**) , Michael Glass, Bonaventura Coppola
- What's there
  - Scalable Distributional Similarity (Hadoop)
  - UIMA based text processing implementing @@ operation on different languages/NLP
  - Fast and Scalable Knowledge Management
  - Sense Clustering, WSD, lexical substitution, Thesauri induction, Paraphrasing, Entity Linking, …
  - Machine Learning: CRF, Chinese Whisper Clustering, …

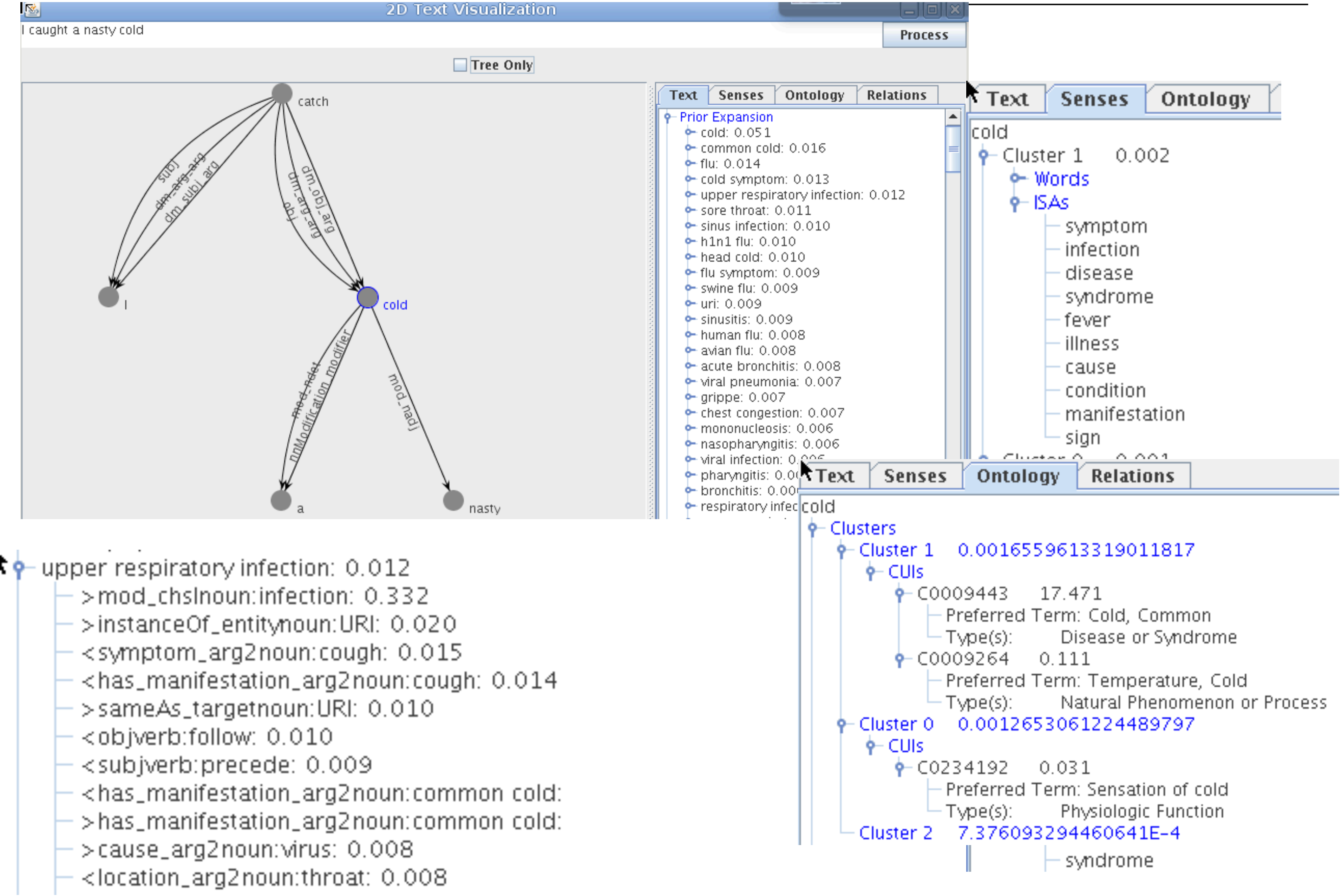
UIMA™ Unstructured Information Management Architecture
An Apache Project

hadoop

JoBimText for the Medical Domain

- Input:
  - Watson Medical Corpus
    - ~ 2 Gigabytes of text
    - UMLS
- Preprocessing:
  - Medical Extended Slot Grammar (ESG) Parser
    - Dependency Parser
    - Medical Adaptation of the Jeopardy Parser
  - TWREX
    - Relation Extraction system adapted to UMLS relations
- @@ system:
  - Terms are represented by
    - syntactic dependencies
    - TWREX relations
- Unsupervised learning on a Small Hadoop Cluster
- Watson Analytics for Answer Scoring, Matching, Passage Scoring
- Demo

JoBimText Demo: Ambiguity 1/2



I caught a nasty cold

Text | Senses | Ontology | Relations

- Prior Expansion
  - cold: 0.051
  - common cold: 0.016
  - flu: 0.014
  - cold symptom: 0.013
  - upper respiratory infection: 0.012
  - sore throat: 0.011
  - sinus infection: 0.010
  - h1n1 flu: 0.010
  - head cold: 0.010
  - flu symptom: 0.009
  - swine flu: 0.009
  - uri: 0.009
  - sinusitis: 0.009
  - human flu: 0.008
  - avian flu: 0.008
  - acute bronchitis: 0.008
  - viral pneumonia: 0.007
  - grippe: 0.007
  - chest congestion: 0.007
  - mononucleosis: 0.006
  - nasopharyngitis: 0.006
  - viral infection: 0.006
  - pharyngitis: 0.00
  - bronchitis: 0.00
  - respiratory infec

Text | Senses | Ontology

cold
- Cluster 1 0.002
  - Words
  - ISAs
    - symptom
    - infection
    - disease
    - syndrome
    - fever
    - illness
    - cause
    - condition
    - manifestation
    - sign
- Cluster 0 0.001

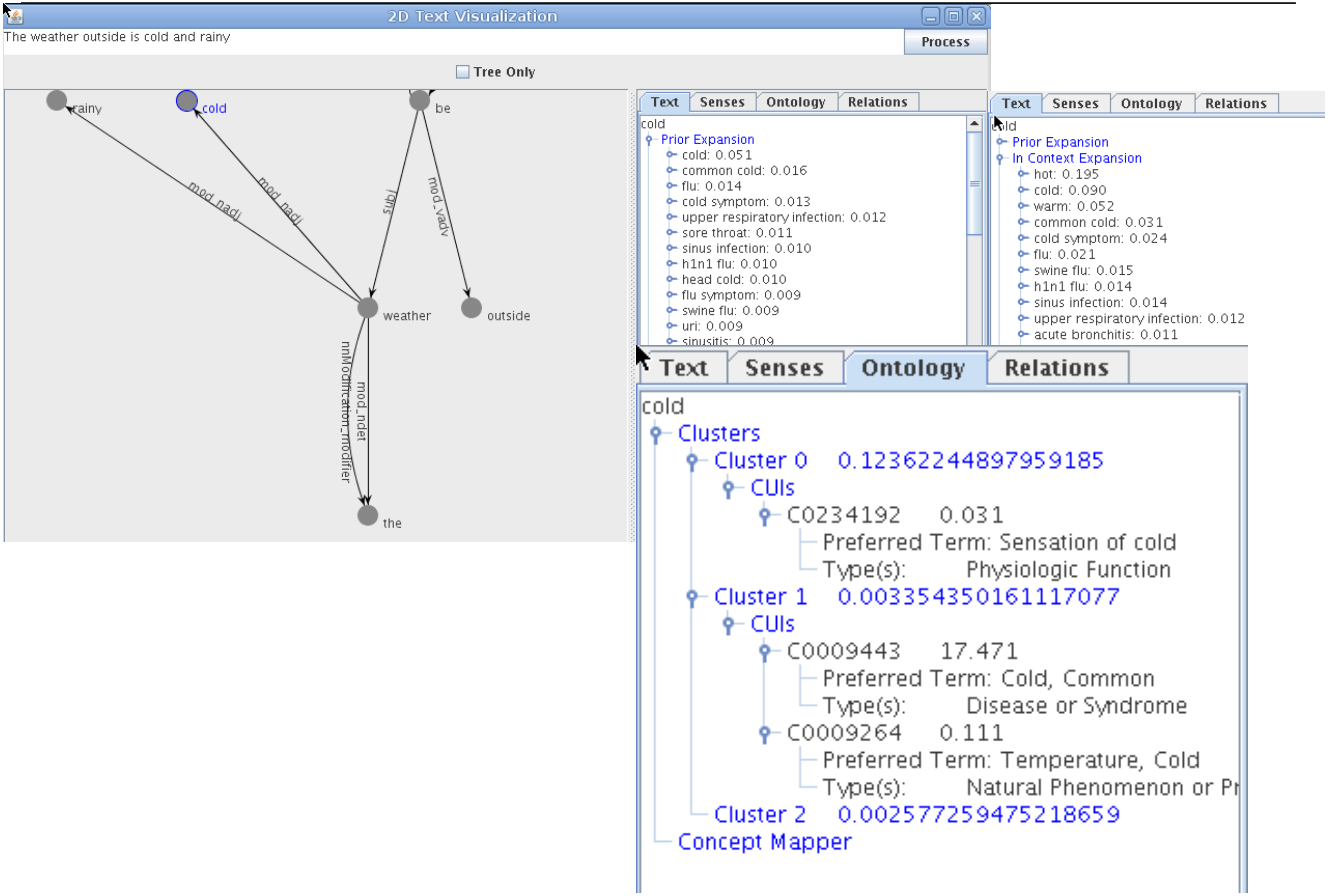
Text | Senses | Ontology | Relations

cold
- Clusters
  - Cluster 1 0.0016559613319011817
    - CUIs
      - C0009443 17.471
        - Preferred Term: Cold, Common
        - Type(s): Disease or Syndrome
      - C0009264 0.111
        - Preferred Term: Temperature, Cold
        - Type(s): Natural Phenomenon or Process
  - Cluster 0 0.0012653061224489797
    - CUIs
      - C0234192 0.031
        - Preferred Term: Sensation of cold
        - Type(s): Physiologic Function
  - Cluster 2 7.376093294460641E-4

syndrome

- upper respiratory infection: 0.012
  - >mod_chslnoun:infection: 0.332
  - >instanceOf_entitynoun:URI: 0.020
  - <symptom_arg2noun:cough: 0.015
  - <has_manifestation_arg2noun:cough: 0.014
  - >sameAs_targetnoun:URI: 0.010
  - <objverb:follow: 0.010
  - <subjverb:precede: 0.009
  - <has_manifestation_arg2noun:common cold:
  - >has_manifestation_arg2noun:common cold:
  - >cause_arg2noun:virus: 0.008
  - <location_arg2noun:throat: 0.008

JoBimText Demo: Ambiguity 2/2



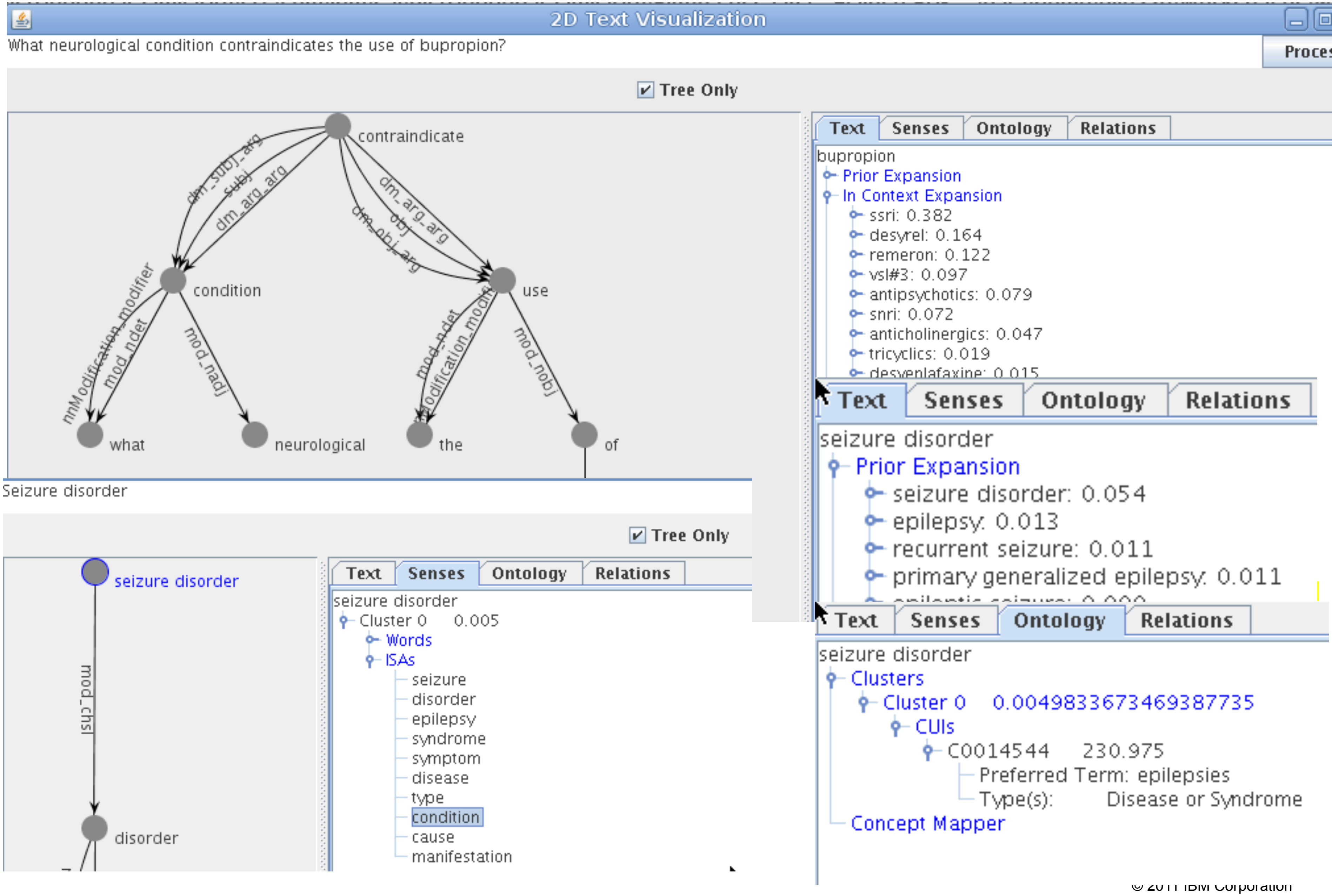
The weather outside is cold and rainy

```
cold
Prior Expansion
   cold: 0.051
   common cold: 0.016
   flu: 0.014
   cold symptom: 0.013
   upper respiratory infection: 0.012
   sore throat: 0.011
   sinus infection: 0.010
   h1n1 flu: 0.010
   head cold: 0.010
   flu symptom: 0.009
   swine flu: 0.009
   uri: 0.009
   sinusitis: 0.009
```

```
cold
Prior Expansion
In Context Expansion
   hot: 0.195
   cold: 0.090
   warm: 0.052
   common cold: 0.031
   cold symptom: 0.024
   flu: 0.021
   swine flu: 0.015
   h1n1 flu: 0.014
   sinus infection: 0.014
   upper respiratory infection: 0.012
   acute bronchitis: 0.011
```

```
cold
Clusters
   Cluster 0    0.12362244897959185
      CUIs
         C0234192    0.031
            Preferred Term: Sensation of cold
            Type(s):        Physiologic Function
   Cluster 1    0.0033543501611707 7
      CUIs
         C0009443    17.471
            Preferred Term: Cold, Common
            Type(s):        Disease or Syndrome
         C0009264    0.111
            Preferred Term: Temperature, Cold
            Type(s):        Natural Phenomenon or Pr
   Cluster 2    0.0025772594752186 59
Concept Mapper
```

JoBimText Demo: Question Processing



What neurological condition contraindicates the use of bupropion?

☑ Tree Only

Text | Senses | Ontology | Relations

bupropion
- Prior Expansion
- In Context Expansion
  - ssri: 0.382
  - desyrel: 0.164
  - remeron: 0.122
  - vsl#3: 0.097
  - antipsychotics: 0.079
  - snri: 0.072
  - anticholinergics: 0.047
  - tricyclics: 0.019
  - desvenlafaxine: 0.015

Seizure disorder

☑ Tree Only

Text | Senses | Ontology | Relations

seizure disorder
- Cluster 0 0.005
  - Words
  - ISAs
    - seizure
    - disorder
    - epilepsy
    - syndrome
    - symptom
    - disease
    - type
    - condition
    - cause
    - manifestation

Text | Senses | Ontology | Relations

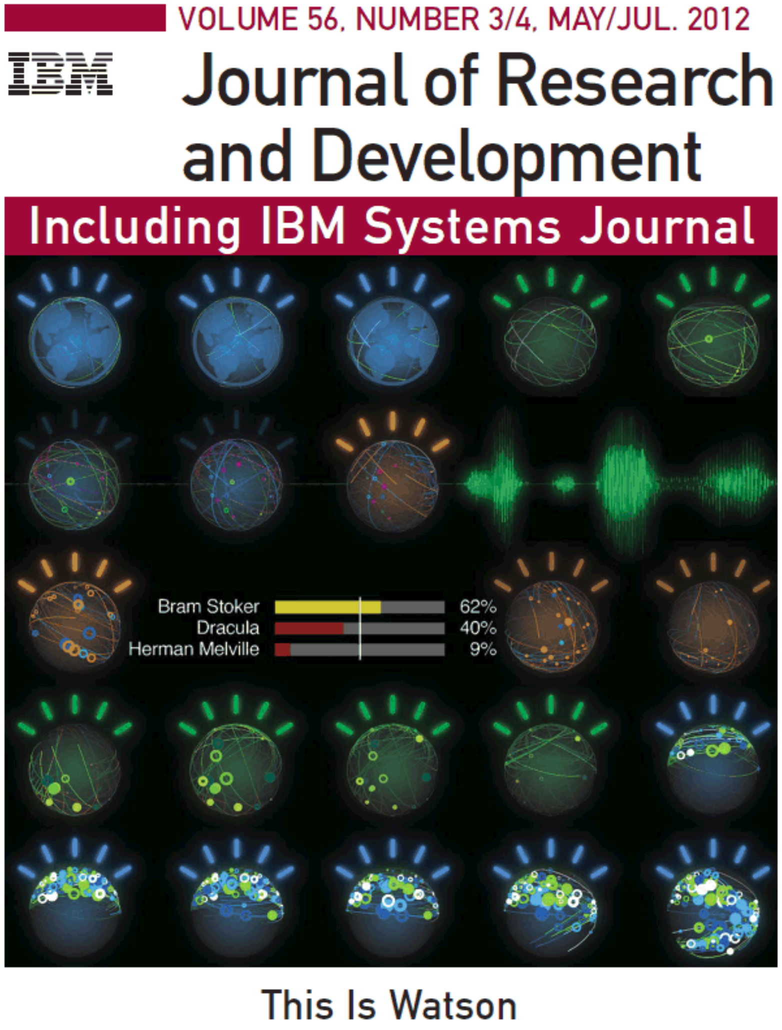
seizure disorder
- Prior Expansion
  - seizure disorder: 0.054
  - epilepsy: 0.013
  - recurrent seizure: 0.011
  - primary generalized epilepsy: 0.011
  - epileptic seizure: 0.009

Text | Senses | Ontology | Relations

seizure disorder
- Clusters
  - Cluster 0 0.004983367346938735
    - CUIs
      - C0014544 230.975
        - Preferred Term: epilepsies
        - Type(s): Disease or Syndrome
- Concept Mapper

References

- Ferrucci et al., **Building Watson: An Overview of the DeepQA Project,** AI Magazine, 2010
- *Ferrucci et al.,* **Watson: Beyond Jeopardy!,** 2011 RC25270, to appear in Artificial Intelligence Journal.
- Deep QA publications website
  - http://researcher.ibm.com/view_grouppubs.php?grp=2099
- Videos on Watson
  - http://www-03.ibm.com/innovation/us/watson/index.html



VOLUME 56, NUMBER 3/4, MAY/JUL. 2012

IBM Journal of Research and Development

Including IBM Systems Journal

This Is Watson

- http://ieeexplore.ieee.org/xpl/tocresult.jsp?reload=true&isnumber=6177717