The Future of AI:
The View from AI2

Oren Etzioni

www.allenai.org

ALLEN INSTITUTE *for* ARTIFICIAL INTELLIGENCE

Allen Institute for Artificial Intelligence (AI2)

AI2 ALLEN INSTITUTE
for ARTIFICIAL INTELLIGENCE

AI Present: Deep Learning Tidal Wave

Winograd Schema Challenge (Levesque, 2011)

The large ball crashed right through the table because **it** was ma... ***foam***

It = ...

Common-sense knowledge & tractable reasoning are necessary for NLP!

AI cannot be reduced to classification
(using realistic feature sets)

“You can’t play 20 questions with nature and Win.”
Allen Newell, 1973

What's Next?

Outline

I. AI2 Methodology
II. Euclid
III. Aristo
IV. Semantic Scholar Overview & Demo
V. Conclusions

AI is Increasingly Fragmented

SAT solver

MLN

SVM

Drones

Etc.

CNN

We need integrated AI Systems!

NLP

ALLEN INSTITUTE *for* ARTIFICIAL INTELLIGENCE

AI2 Research Methodology

Build AI Programs that take written tests

1. Externally-defined challenge tasks
2. Training data + unseen test data (“as is”)
3. Measurable progress, clear focus, ambitious goals

Key differences with Watson:

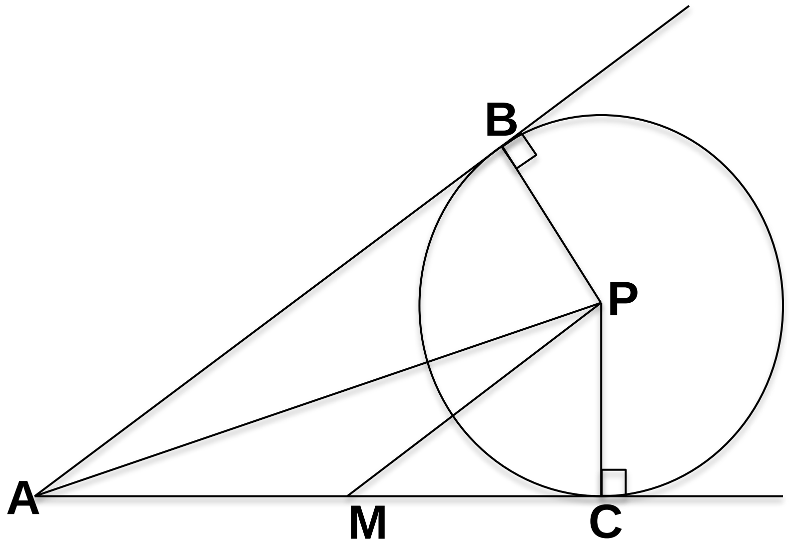
1. Deeper semantics & inference
2. Open model: publish, collaborate, open source

II. Euclid: Solving Geometry Questions

In the figure at the right, the circle has center P and radius r. Lines AB and AC are tangent to the circle at points B and C, respectively. If PM is the bisector of the tangent, and the measure of angle PMC equals the measure of angle MPC, what is the length of segment PA?

(a) $r\sqrt{3}$ **(b)** $r\sqrt{5}$ **(c)** $r\sqrt{2}$

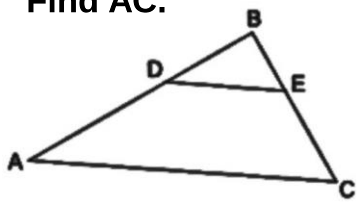
(d) $r+1$ **(e)** $2r$



Multi-Modal Parsing Approach

Geometry Question

In triangle ABC, line DE is parallel with line AC, DB equals 4, AD is 8, and DE is 5. Find AC.



Semantic Forest

(Noisy) Relations

Parallel(line(DE), line(AC))
Parallel(line(AC), line(DB))
Equal(length(line(AD)), 4)
Equal(length(line(AD)), 8)
...

G-Aligner

X^D, R^D, d

Grounded Relations

Parallel (l1,l2)
Parallel (l2,l3)
Equal(length(l3),4)
Equal(length(l4),8)
...

Selected Relations

Parallel (l1,l2)
Equal(length(l1),4)
Equal(length(l3),4)
Equal(length(l2),8)
Length(l2) = ?

Solver

Answer

Details in [EMNLP15]

ALLEN INSTITUTE *for* ARTIFICIAL INTELLIGENCE

Numerical Solver

- Translate a logical form to a non-linear equation

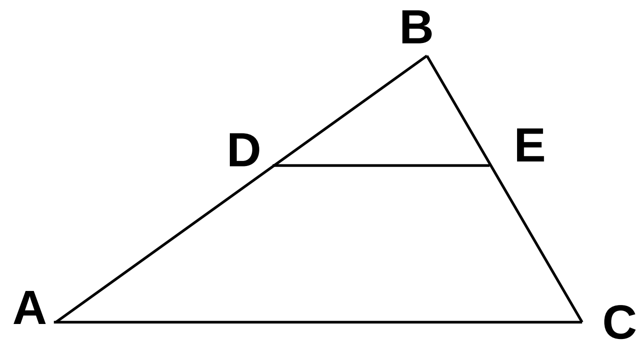
Formal Language	Equations
Equals(LengthOf(AB),d)	$(A_x-B_x)^2+(A_y-B_y)^2-d^2 = 0$
Parallel(AB, CD)	$(A_x-B_x)(C_y-D_y)-(A_y-B_y)(C_x-D_x) = 0$
LiesOn(B, AC)	$(A_x-B_x)(B_y-C_y)-(A_y-B_y)(B_x-C_x) = 0$
Perpendicular(AB,CD)	$(A_x-B_x)(C_x-D_x)+(A_y-B_y)(C_y-D_y) = 0$

Goal: Find an assignment to the variables that satisfies all the equations *simultaneously*

Optimization Result

"In triangle ABC, line DE is parallel with line AC, DB equals 4, AD is 8, and DE is 5. Find AC."

a) 2   b) 4   c) 6   d) 8   e) 10



IsTriangle(ABC)	0.96
Parallel(AC, DE)	0.91
Parallel(AC, DB)	0.74
Equals(LengthOf(DB), 4)	0.97
Equals(LengthOf(AD), 8)	0.94
Equals(LengthOf(DE), 5)	0.94
Equals(4, LengthOf(AD))	0.31
Find(LengthOf(AC))	0.90

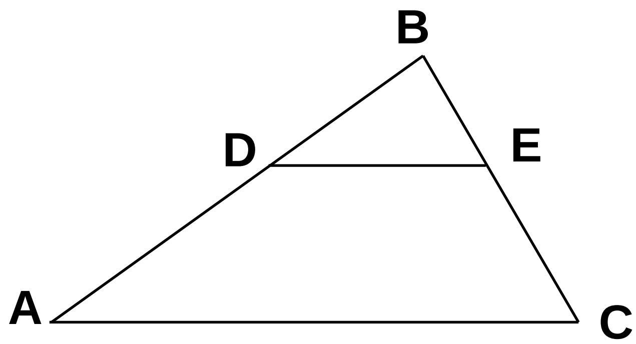
LiesOn(D, AB)	1.0
LiesOn(E, BC)	1.0
Parallel(AC, DE)	0.99
Parallel(AC, DB)	0.02

IsTriangle(ABC)	Parallel(AC, DE)
Equals(LengthOf(DB), 4)	Equals(LengthOf(AD), 8)
Equals(LengthOf(DE), 5)	Find(LengthOf(AC))

Optimization Result

"In triangle ABC, line DE is parallel with line AC, DB equals 4, AD is 8, and DE is 5. Find AC."

a) 2   b) 4   c) 6   d) 8   e) 10



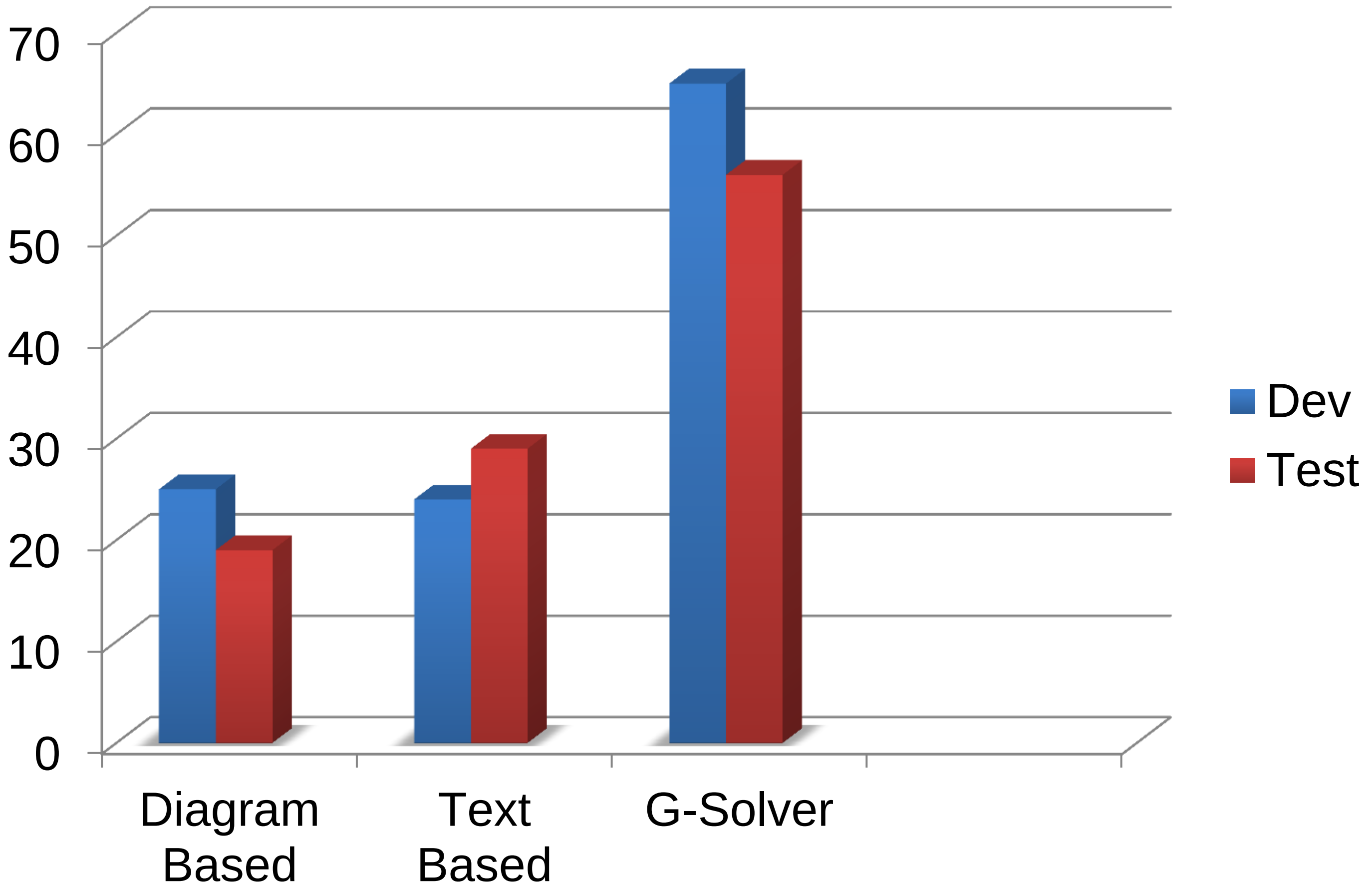
IsTriangle(ABC)	0.96
Parallel(AC, DE)	0.91
Parallel(AC, DB)	0.74
Equals(LengthOf(DB), 4)	0.97
Equals(LengthOf(AD), 8)	0.94
Equals(LengthOf(DE), 5)	0.94
Equals(4, LengthOf(AD))	0.31
Find(LengthOf(AC))	0.90

LiesOn(D, AB)	1.0
LiesOn(E, BC)	1.0
Parallel(AC, DE)	0.99
Parallel(AC, DB)	0.02

IsTriangle(ABC)	Parallel(AC, DE)
Equals(LengthOf(DB), 4)	Equals(LengthOf(AD), 8)
Equals(LengthOf(DE), 5)	**Equal(LengthOf(AC),10)**

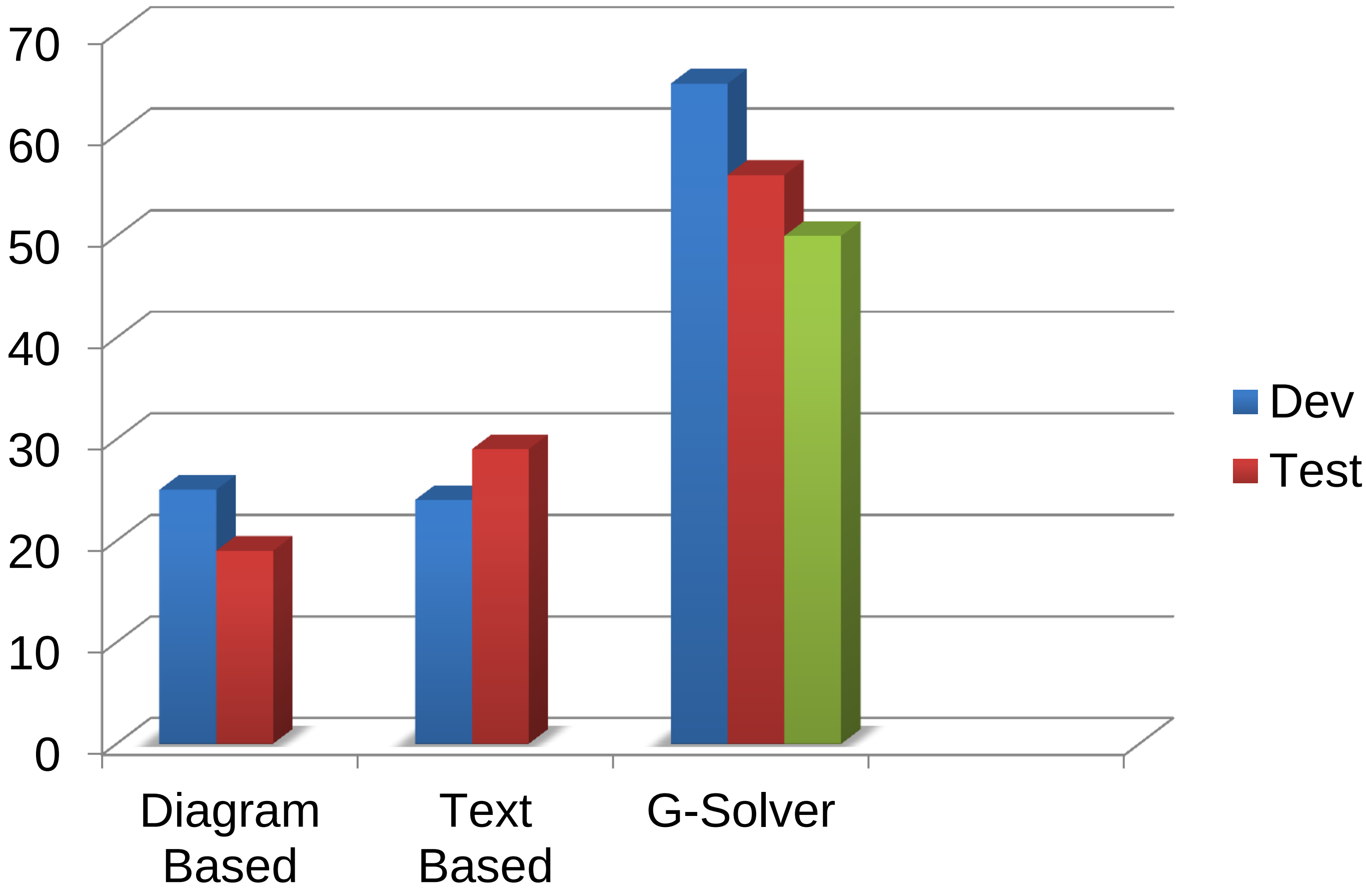
ALLEN INSTITUTE *for* ARTIFICIAL INTELLIGENCE

Evaluation



ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Evaluation



Just presented at [EMNLP2015]

III. Aristo: Elementary School Science Tests

THE UNIVERSITY OF THE STATE OF NEW YORK

GRADE 4

ELEMENTARY-LEVEL

SCIENCE TEST

WRITTEN TEST

JUNE 3, 2013

12 Which force causes a bicycle to slow down when the brakes are used?

A friction
B electricity
C gravity
D magnetism

Why Elementary Science is Challenging

- **Natural Language <u>Understanding</u>** is often needed
  - "Shallow" NLP maxes out <60% (guessing = 25%)

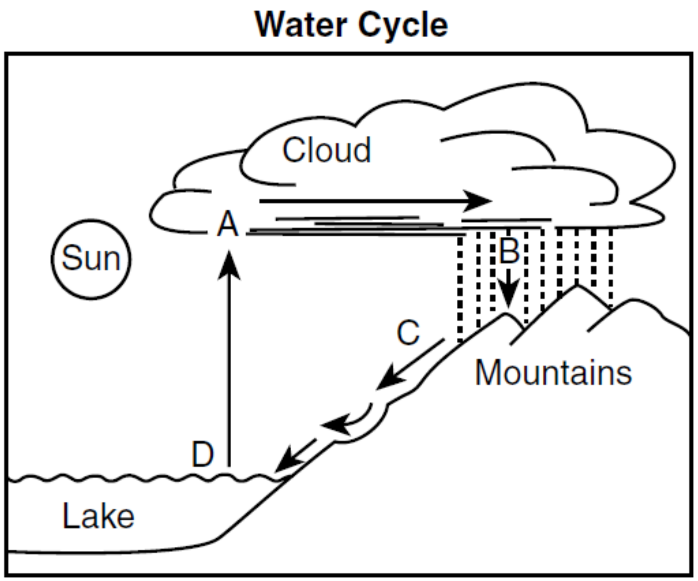
- **Scientific and World Knowledge** is needed

  A ball is tossed in the air and it comes back down due to (A) gravity (B) …

- **Linguistic Variability** is everywhere

  allow heat transfer ~ is a conductor
  get a better look at ~ view in more detail
  healthy lifestyle ~ maintain good health

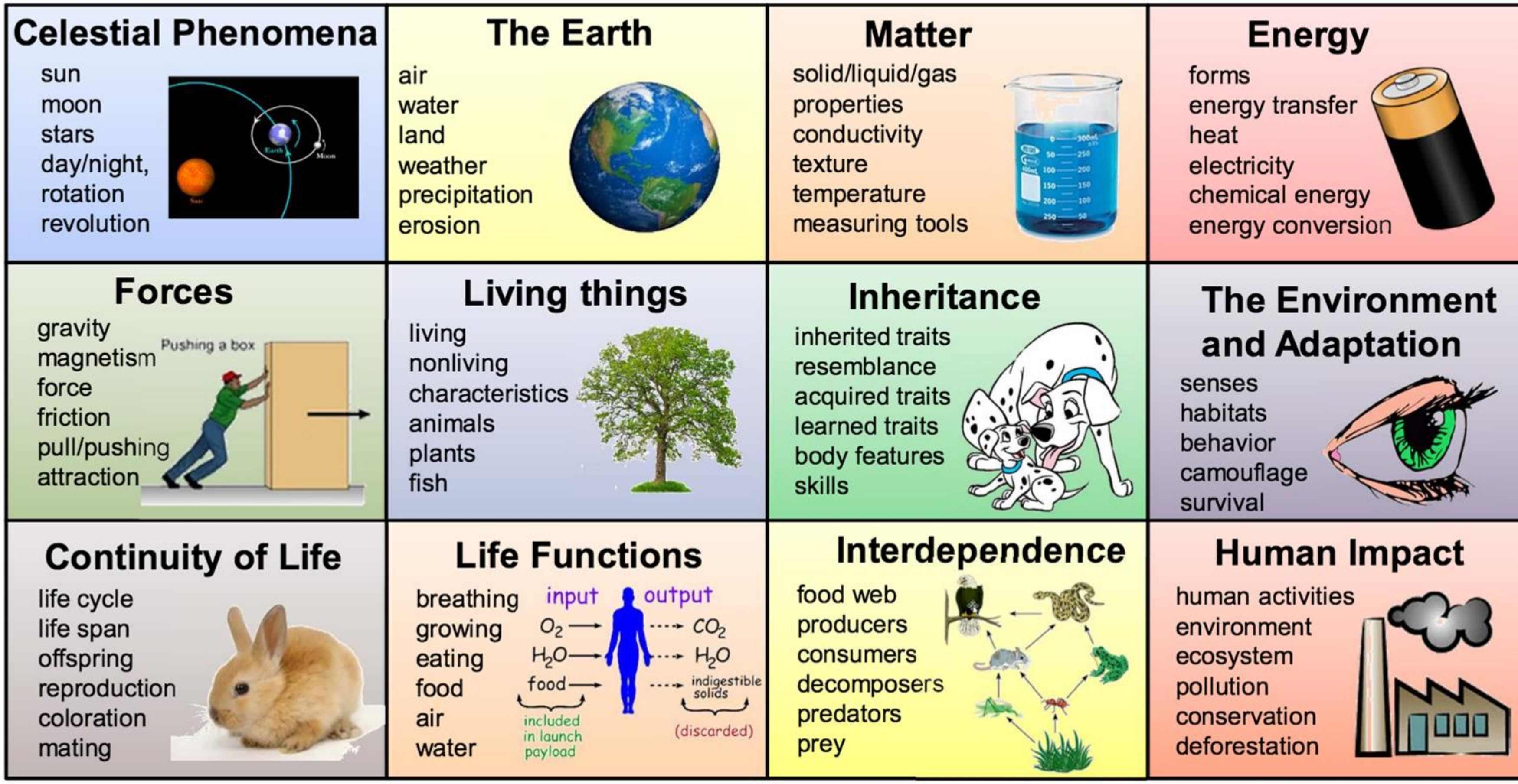
- 45% of questions involve **diagrams**



Which letter shows runoff?

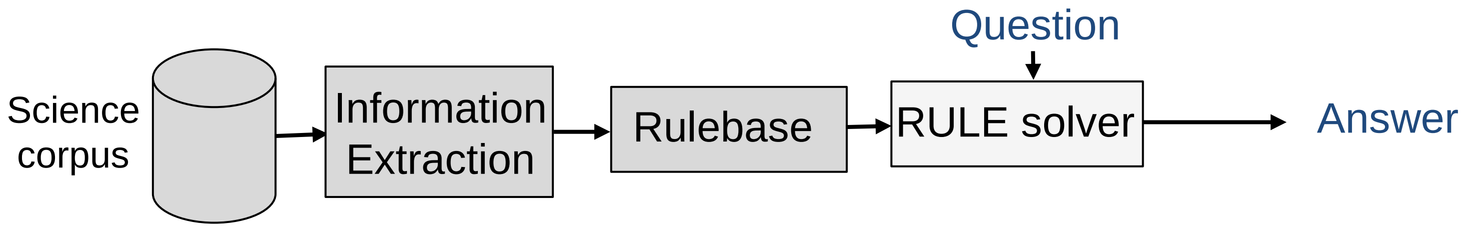
Project Aristo

What does Aristo need to know to pass the exam?

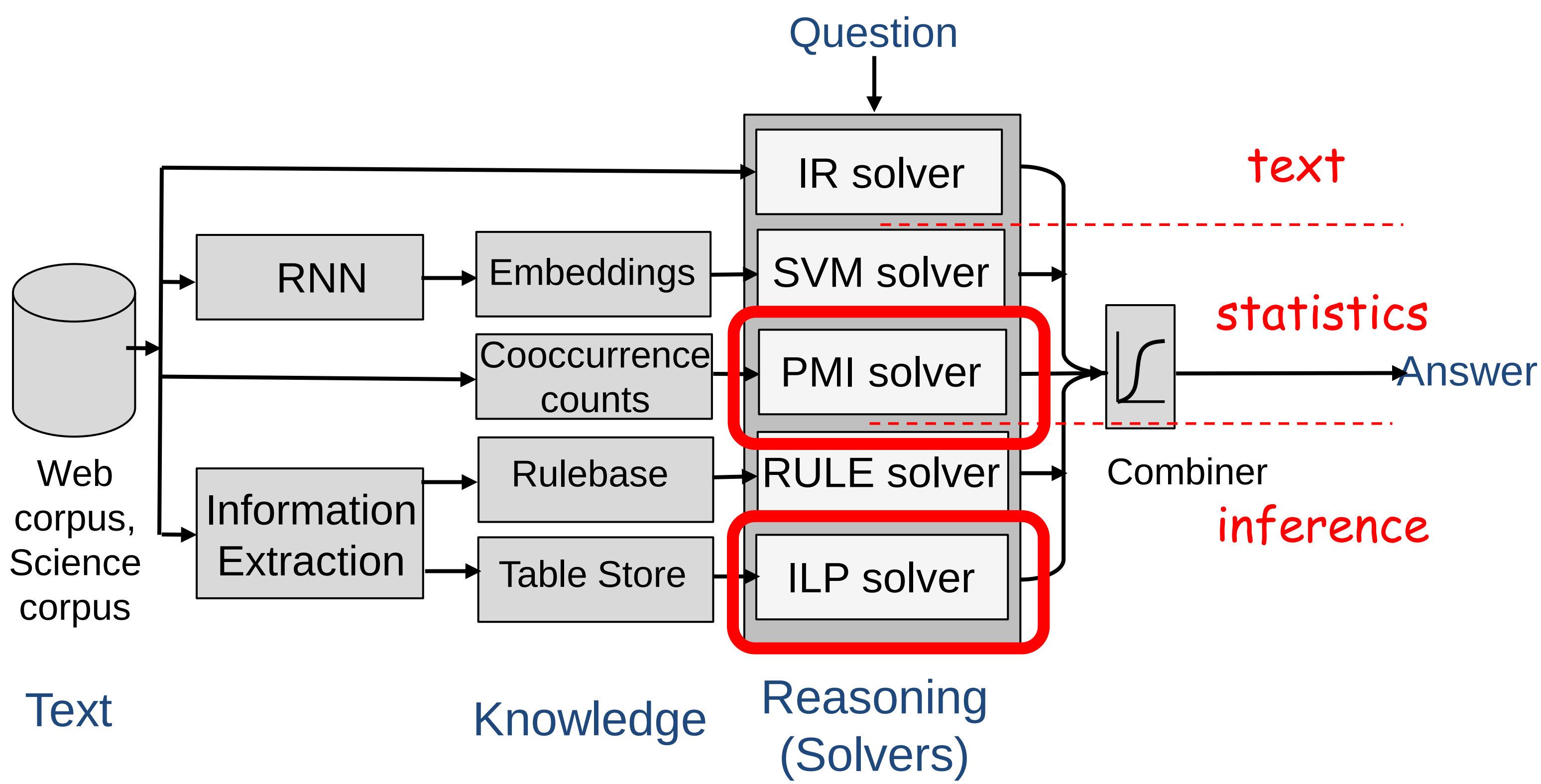


Evolving Design of Aristo

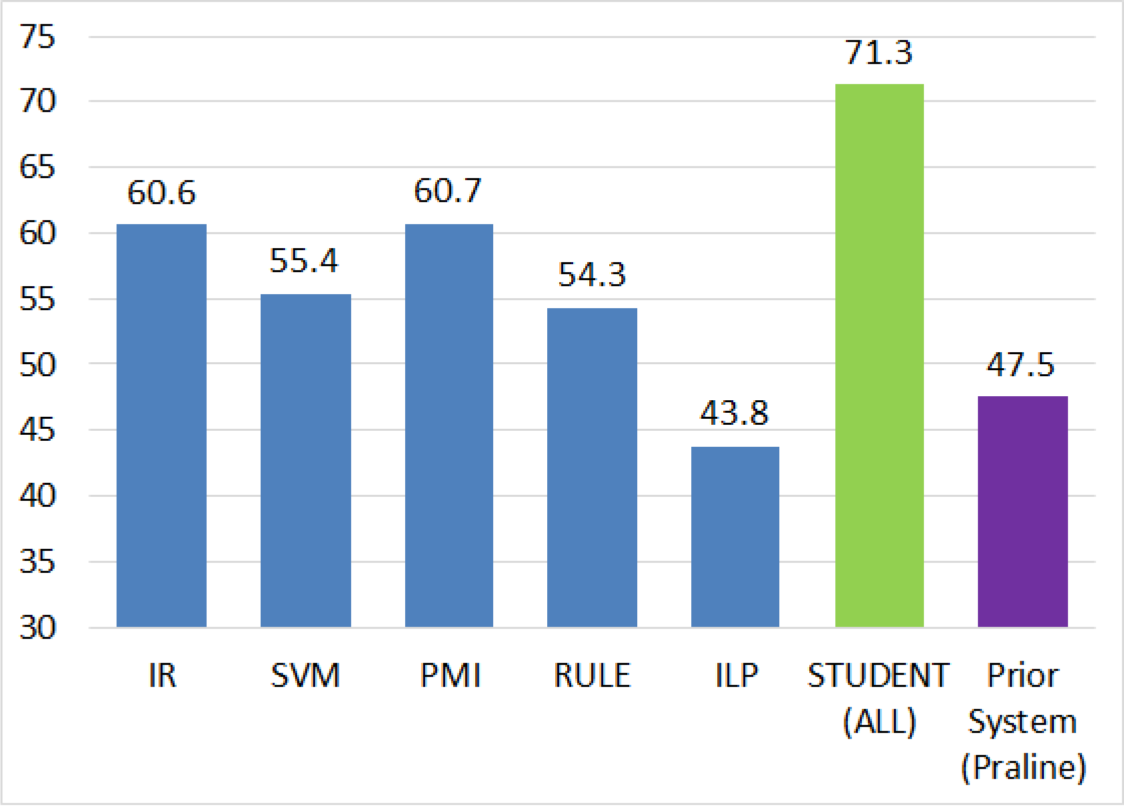
- **EMNLP 2015:** MLN "solver" operating on machine-extracted rules



- **2016:** Several levels of structure



Current Results (test score on ND Questions)

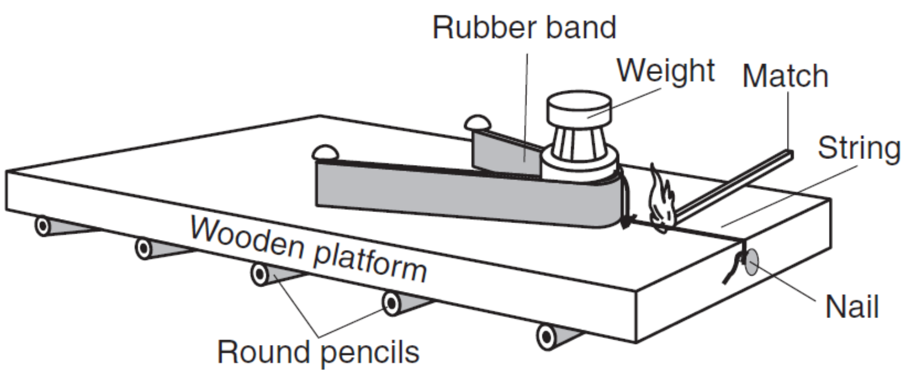


Aristo Demo

[Aristo]

8th Grade – A Significantly Increased Challenge

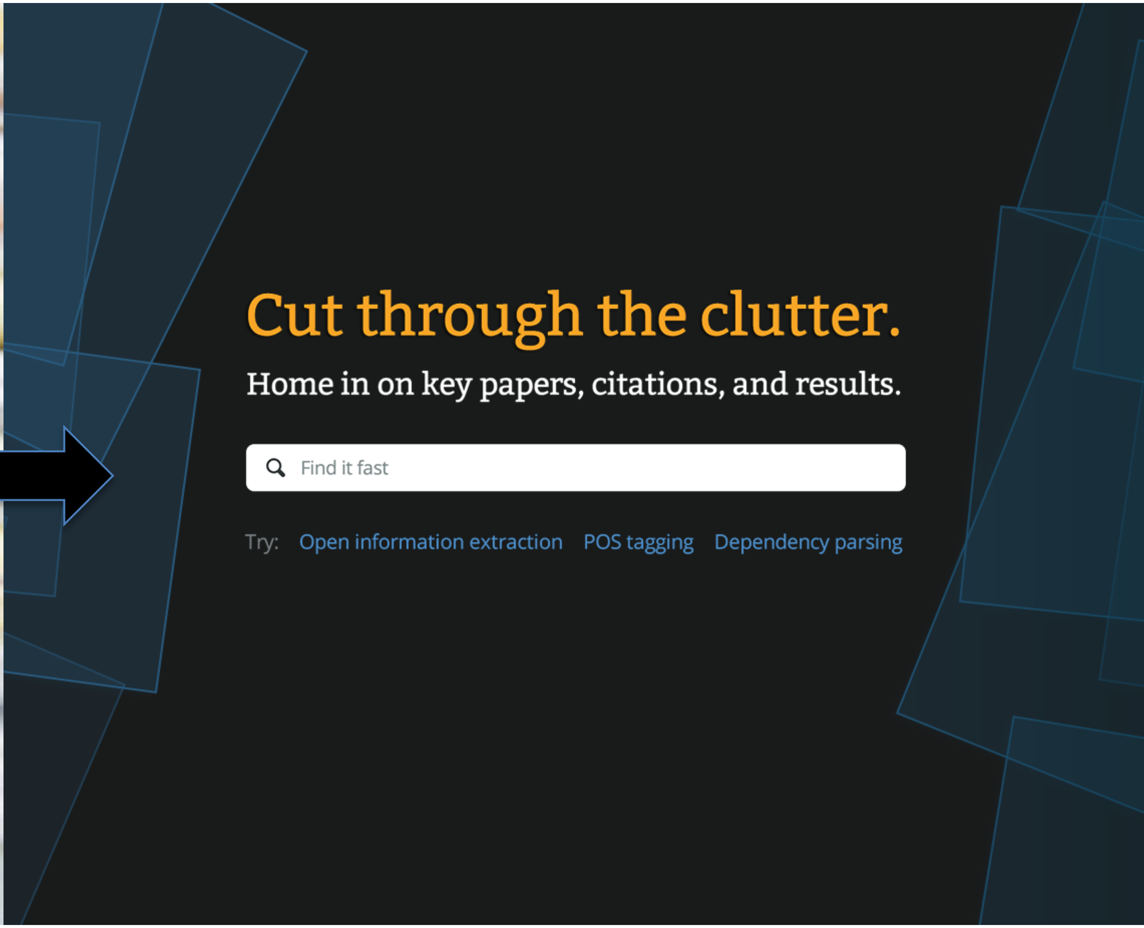
- 4th to 8th grade: major jump in challenges: 4 years of education and life experience for a student
- Majority of questions use diagrams
- Some are "AI hard", e.g. (in 2006 8th grade exam)



Source: Adapted from, Constantine Constant, *Earth Science Workbook*, AMSCO, 1972

57 On the diagram above, draw an arrow to represent the direction the **wooden platform** will move when the lit match burns through the string and the weight is propelled from the platform. [1]

III. Semantic Scholar

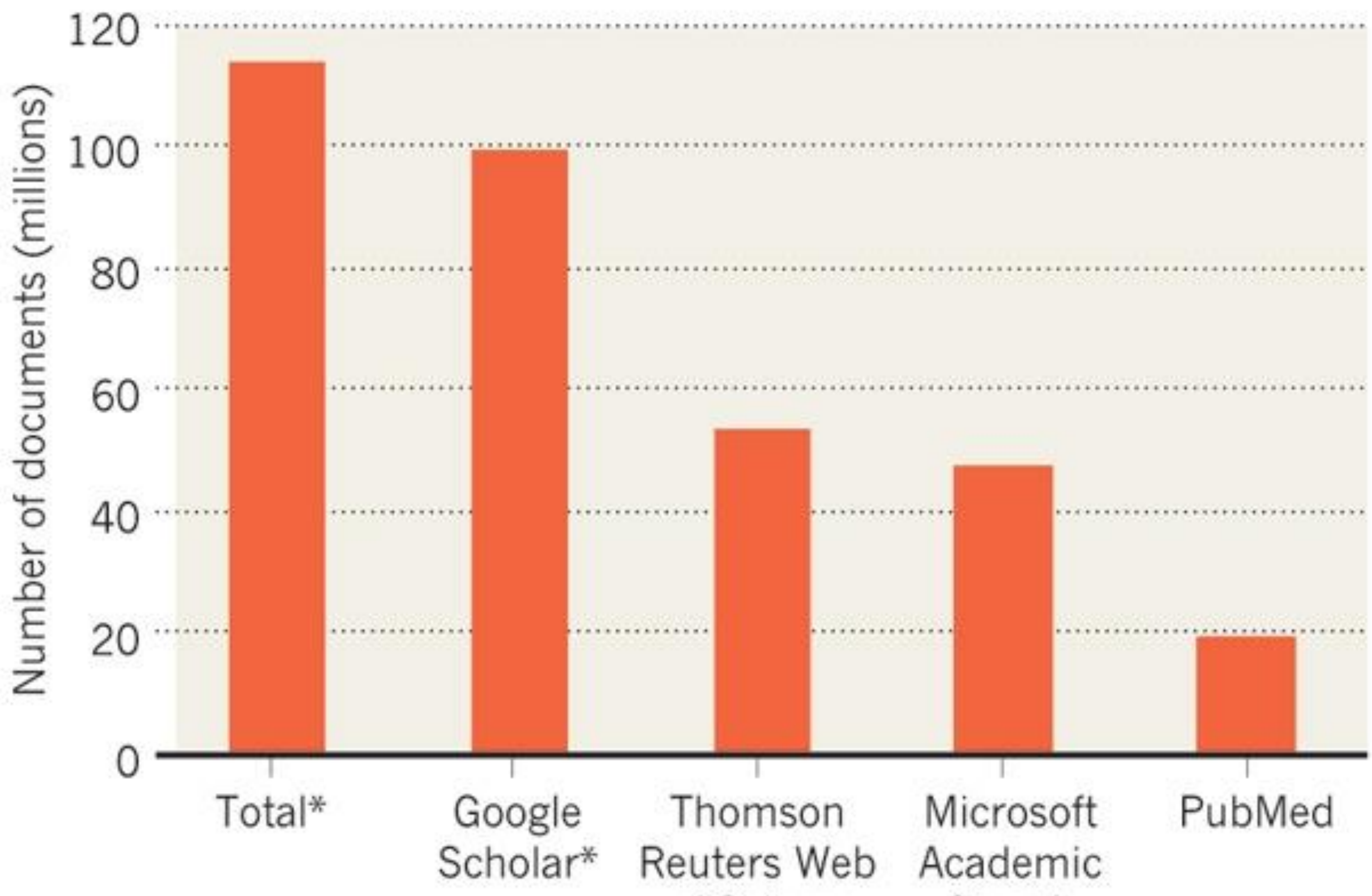


ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Motivation

THE WEB OF SCHOLARSHIP

Around 114 million English-language scholarly documents (including papers, books and technical reports) can be found on the web.



*Estimated

Web Images More...

Google "information extraction"

Scholar About 139,000 results (**0.12** sec)

Articles

Case law

My library

Any time

Since 2014

Since 2013

Since 2010

Custom range...

Sort by relevance

Sort by date

include patents

include citations

Create alert

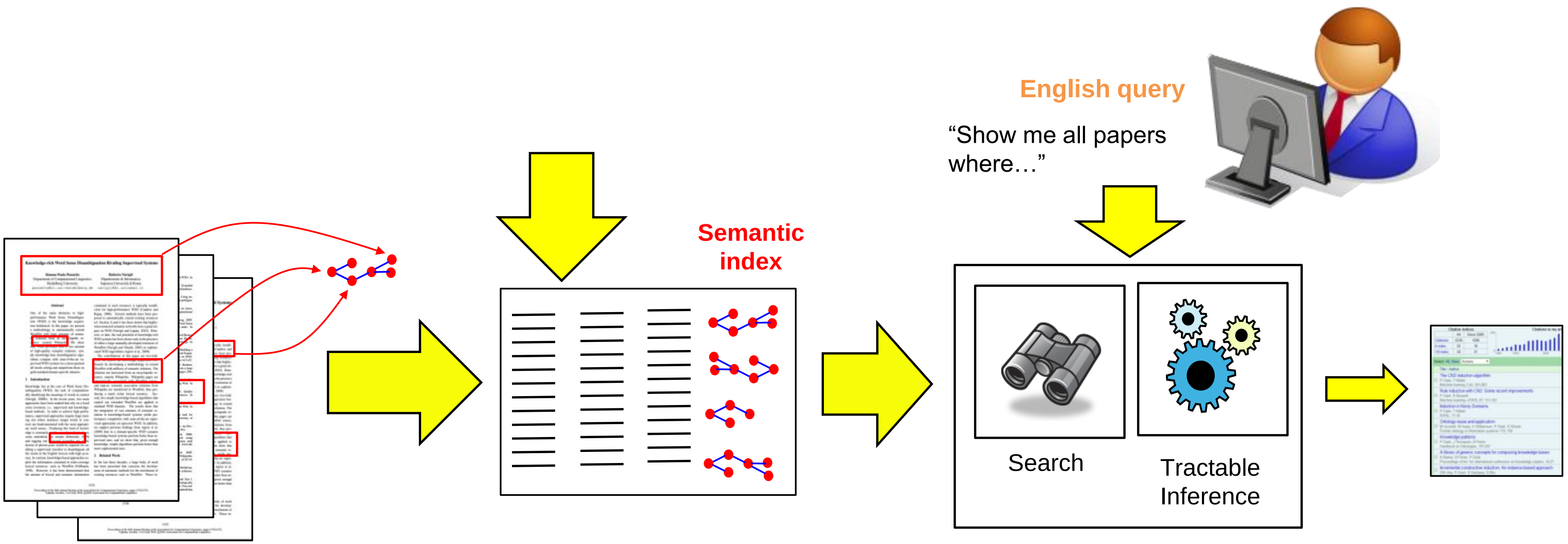
[PDF] Maximum Entropy Markov Models for **Information Extraction** and Segmentation.
A McCallum, D Freitag, FCN Pereira - ICML, 2000 - courses.ischool.berkeley.edu
Page 1. 1 Maximum Entropy Markov Models for **Information Extraction** and Segmentation Andrew McCallum, Dayne Freitag, and Fernando Pereira **...** Named entity recognition: <ORG>Mips</ORG> Vice President <PRS>John Hime</PRS> – **Information extraction: ...**
Cited by 1126 Related articles All 50 versions Cite Save More

Incorporating non-local information into **information extraction** systems by gibbs sampling
JR Finkel, T Grenager, C Manning - … of the 43rd Annual Meeting on …, 2005 - dl.acm.org
Abstract Most current statistical natural language processing models use only local features so as to permit dynamic programming in inference, but this makes them unable to fully account for the long distance structure that is prevalent in language use. We show how to **...**
Cited by 1129 Related articles All 23 versions Cite Save

[PDF] Learning dictionaries for **information extraction** by multi-level bootstrapping
E Riloff, R Jones - AAAI/IAAI, 1999 - aaai.org
Abstract **Information extraction** systems usually require two dictionaries: a semantic lexicon and a dictionary of extraction patterns for the domain. We present a multilevel bootstrapping algorithm that generates both the semantic lexicon and extraction patterns simultaneously. **...**
Cited by 753 Related articles All 18 versions Cite Save More

[PDF] Open **information extraction** for the web
M Banko, MJ Cafarella, S Soderland, M Broadhead... - IJCAI, 2007 - aaai.org
Abstract Traditionally, **Information Extraction** (IE) has focused on satisfying precise, narrow, pre-specified requests from small homogeneous corpora (eg, extract the location and time of seminars from a set of announcements). Shifting to a new domain requires the user to **...**
Cited by 806 Related articles All 31 versions Cite Save More

Leverage AI to Combat Information Overload



ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

DEMO

Our Approach to Figure Understanding

Relation Extraction with Matrix Factorization and Universal Schemas

Sebastian Riedel
Department of Computer Science
University College London
s.riedel@ucl.ac.uk

Limin Yao, Andrew McCallum, Benjamin M. Marlin
Department of Computer Science
University of Massachusetts at Amherst
{lmyao,mccallum,marlin}@cs.umass.edu

Abstract

Traditional relation extraction predicts relations within some fixed and finite target schema. Machine learning approaches to this task require either manual annotation or, in the case of distant supervision, existing structured sources of the same schema. The need for existing datasets can be avoided by using a *universal schema*: the union of all involved schemas (surface form predicates as in OpenIE, and relations in the schemas of pre-existing databases). This schema has an almost unlimited set of relations (due to surface forms), and supports integration with existing structured data (through the relation types of existing databases). To populate a database of such schema we present matrix factorization models that learn latent feature vectors for entity tuples and relations. We show that such latent models achieve substantially higher accuracy than a traditional classification approach. More importantly, by operating simultaneously on relations observed in text and in pre-existing structured DBs such as Freebase, we are able to reason about unstructured and structured data in mutually-supporting ways. By doing so our approach outperforms state-of-the-art distant supervision.

1 Introduction

Most previous work in relation extraction uses a predefined, finite and fixed schema of relation types (such as *born-in* or *employed-by*). Usually some textual data is labeled according to this schema, and this labeling is then used in supervised training of an automated relation extractor, e.g. Culotta and Sorensen (2004). However, labeling textual relations is time-consuming and difficult, leading to significant recent interest in distantly-supervised learning. Here one aligns existing database records with the sentences in which these records have been "rendered"—effectively labeling the text—and from this labeling we can train a machine learning system as before (Craven and Kumlien, 1999; Mintz et al., 2009; Bunescu and Mooney, 2007; Riedel et al., 2010). However, this method relies on the availability of a large database that has the desired schema.

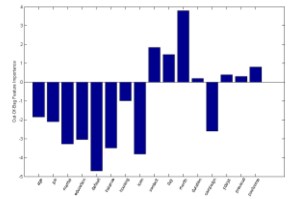
The need for pre-existing datasets can be avoided by using language itself as the source of the schema. This is the approach taken by OpenIE (Etzioni et al., 2008). Here surface patterns between mentions of concepts serve as relations. This approach requires no supervision and has tremendous flexibility, but lacks the ability to generalize. For example, OpenIE may find FERGUSON–*historian-at*–HARVARD but does not know FERGUSON–*is-a-professor-at*–HARVARD. OpenIE has traditionally relied on a large diversity of textual expressions to provide good coverage. But this diversity is not always available, and, in any case, the lack of generalization greatly inhibits the ability to support reasoning.

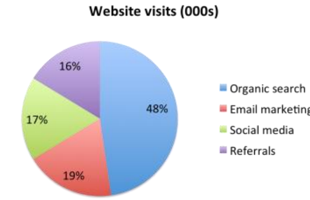
One way to gain generalization is to cluster textual surface forms that have similar meaning (Lin and Pantel, 2001; Pantel et al., 2007; Yates and Etzioni, 2009; Yao et al., 2011). While the clusters discovered by all these methods usually contain semantically related items, closer inspection invariably shows that they do not provide reliable implicature. For example, a typical representative cluster may include *historian-at*, *professor-at*, *scientist-at*, *worked-at*. Although these relation types are indeed semantically related, note that *scientist-at* does not necessarily imply *professor-at*, and *worked-at*

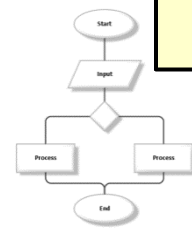
74

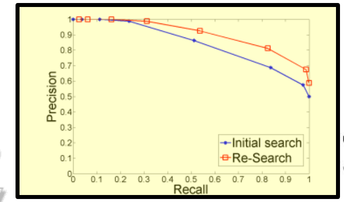
Proceedings of NAACL-HLT 2013, pages 74–84,
Atlanta, Georgia, 9–14 June 2013. ©2013 Association for Computational Linguistics

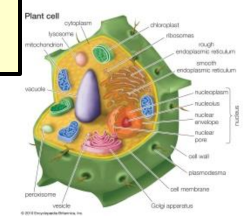
Extract Figures



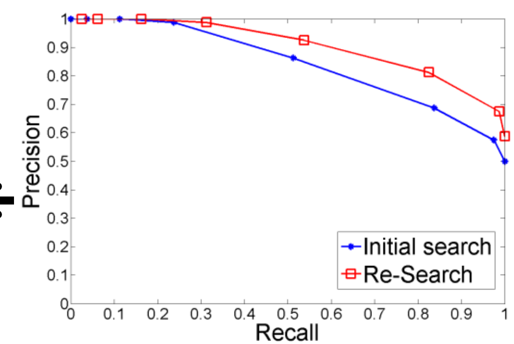








Classify Figures



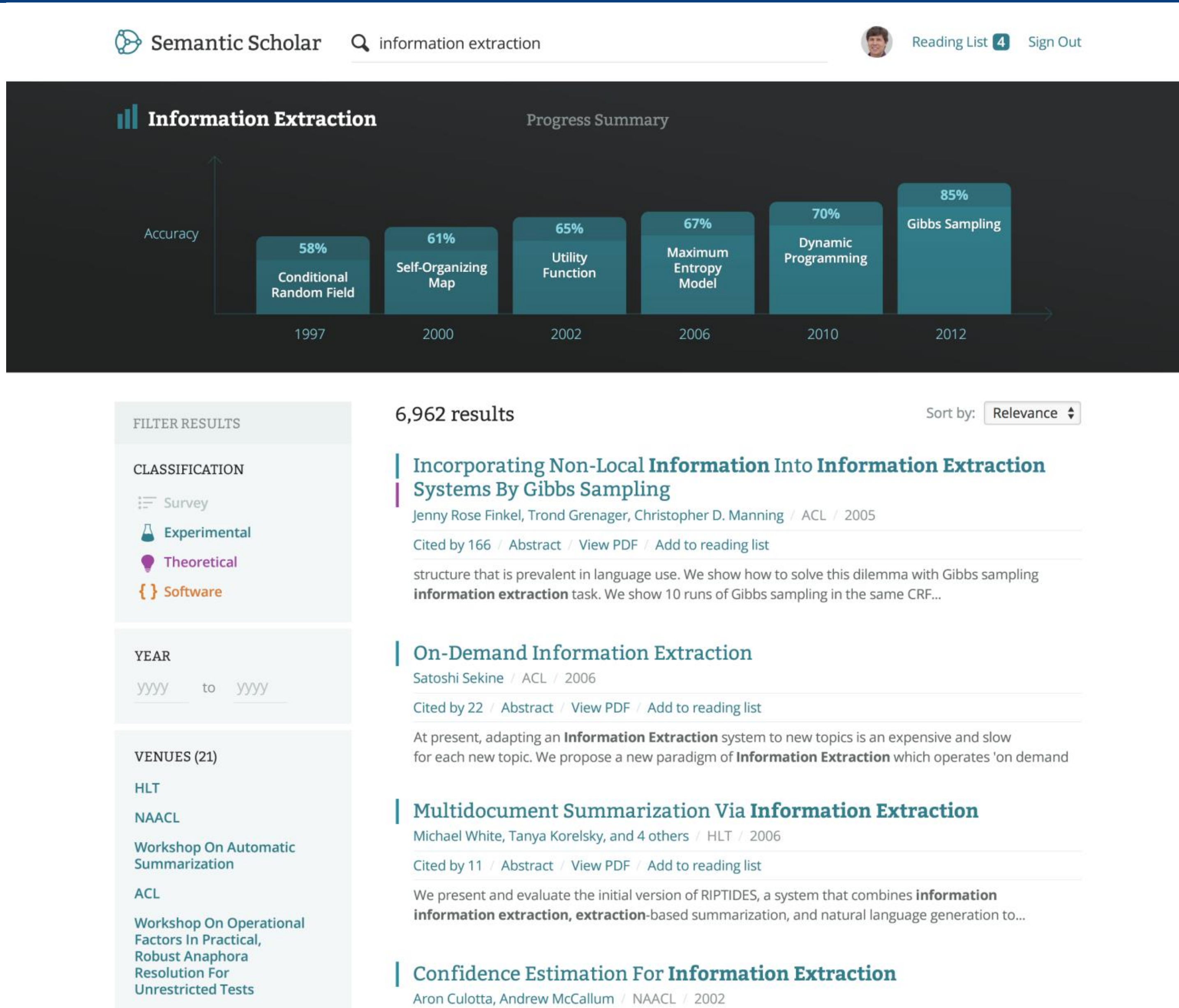
Curve Understanding

```
{"Type":"Figure",
"Number": 4,
"Page": 6,
"DPI": 100,
"Width": 827,
"Height": 1169,
"CaptionBB": [464,400,691,416],
"Caption": "Figure 4: predicted label probabilites",
"ImageBB": [427,85,729,386],
"ImageText" : [
{"Rotation": 0,"TextBB": [622.016,88.298,645.698,97.0056], "Text": "Label"},
        {"Rotation": 0,"TextBB": [648.12,88.298,700.152,97.0056], "Text":
"probabilities"},
        {"Rotation": 0,"TextBB": [702.573,88.298,718.71,97.0056], "Text": "(%)"},
        {"Rotation": 0,"TextBB": [445.871,99.3676,465.601,108.075], "Text": "True"},
        {"Rotation": 0,"TextBB": [468.022,99.3676,491.704,108.075], "Text": "Label"},
        {"Rotation": 0,"TextBB": [583.826,99.3676,623.46,108.075], "Text": "Predicted"},
        {"Rotation": 0,"TextBB": [642.26,99.4838,649.524,108.143], "Text": "..."},
        {"Rotation": 0,"TextBB": [651.946,99.3676,661.632,108.075], "Text": "10"},
        {"Rotation": 0,"TextBB": [664.053,99.3676,673.739,108.075], "Text": "11"},
        {"Rotation": 0,"TextBB": [676.16,99.3676,685.846,108.075], "Text": "12"},
        {"Rotation": 0,"TextBB": [690.689,99.4838,697.954,108.143], "Text": "..."},
        {"Rotation": 0,"TextBB": [702.797,99.3676,712.483,108.075], "Text": "17"},
        {"Rotation": 0,"TextBB": [714.904,99.3676,724.59,108.075], "Text": "18"},
        {"Rotation": 0,"TextBB": [445.871,111.107,456.041,119.766], "Text": "..."},
        {"Rotation": 0,"TextBB": [437.987,122.176,443.449,130.835], "Text": "="},
        {"Rotation": 0,"TextBB": [445.871,122.176,490.929,130.835], "Text":
"PastHistory"},
        {"Rotation": 0,"TextBB": [493.351,122.176,509.487,130.835], "Text": "[11]"},
        {"Rotation": 0,"TextBB": [520.442,122.176,533.886,130.835], "Text": "age"},
        {"Rotation": 0,"TextBB": [583.826,122.176,628.885,130.835], "Text":
"PastHistory"},
        {"Rotation": 0,"TextBB": [656.79,122.176,661.632,130.835], "Text": "0"},
        {"Rotation": 0,"TextBB": [664.054,122.176,673.74,130.835], "Text": "95"},
        {"Rotation": 0,"TextBB": [681.004,122.176,685.847,130.835], "Text": "0"},
        {"Rotation": 0,"TextBB": [707.64,122.176,712.483,130.835], "Text": "0"},
        {"Rotation": 0,"TextBB": [719.747,122.176,724.59,130.835], "Text": "0"},
        {"Rotation": 0,"TextBB": [437.987,133.246,443.449,141.905], "Text": "="},
        {"Rotation": 0,"TextBB": [445.871,133.246,490.929,141.905], "Text":
"PastHistory"},
        {"Rotation": 0,"TextBB": [493.351,133.246,509.487,141.905], "Text": "[11]"},
        {"Rotation": 0,"TextBB": [520.442,133.246,530.128,141.905], "Text": "63"},
        {"Rotation": 0,"TextBB": [583.826,133.246,628.885,141.905], "Text":
"PastHistory"}
```

.json

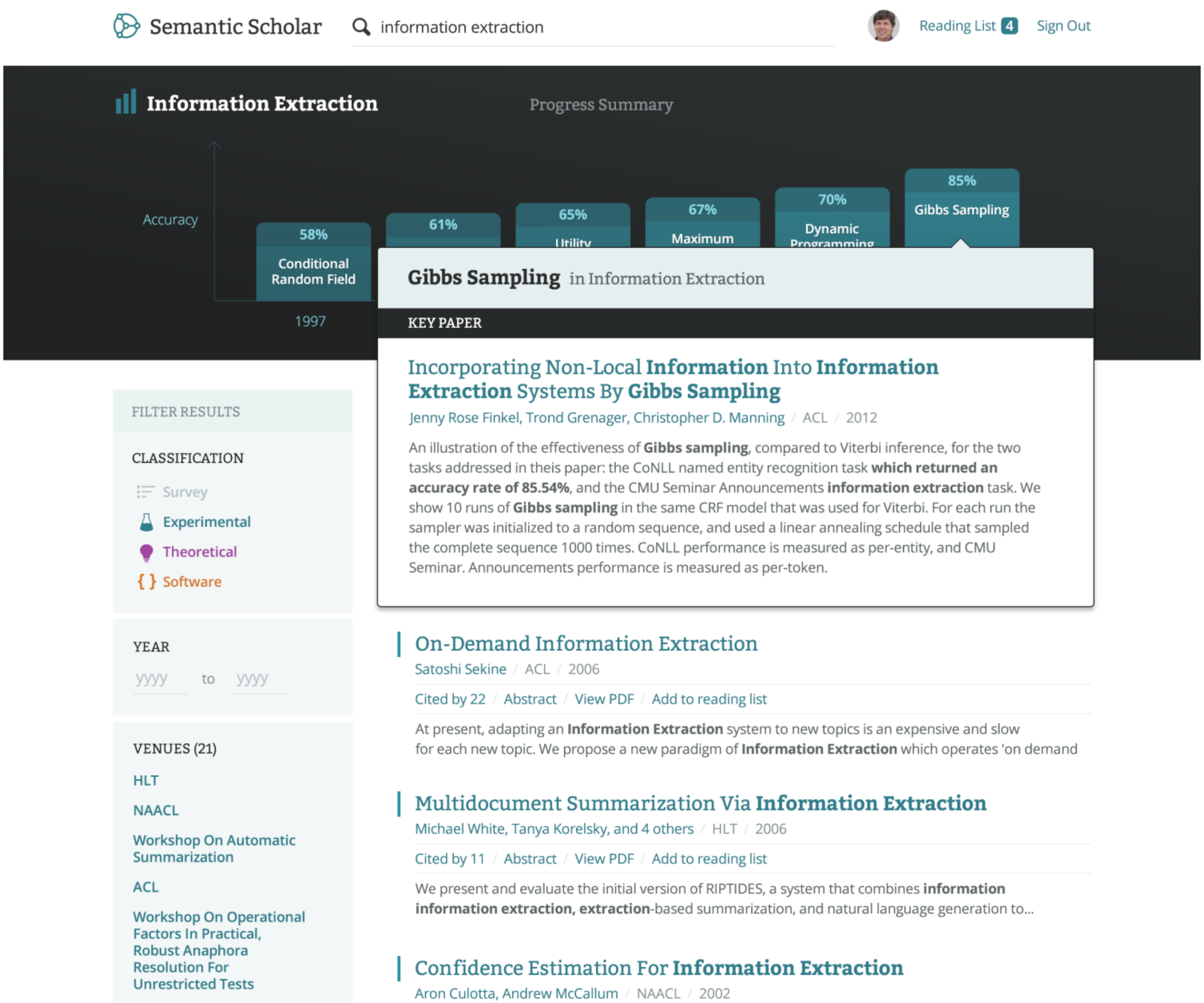
ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Progress Summary in an Topic



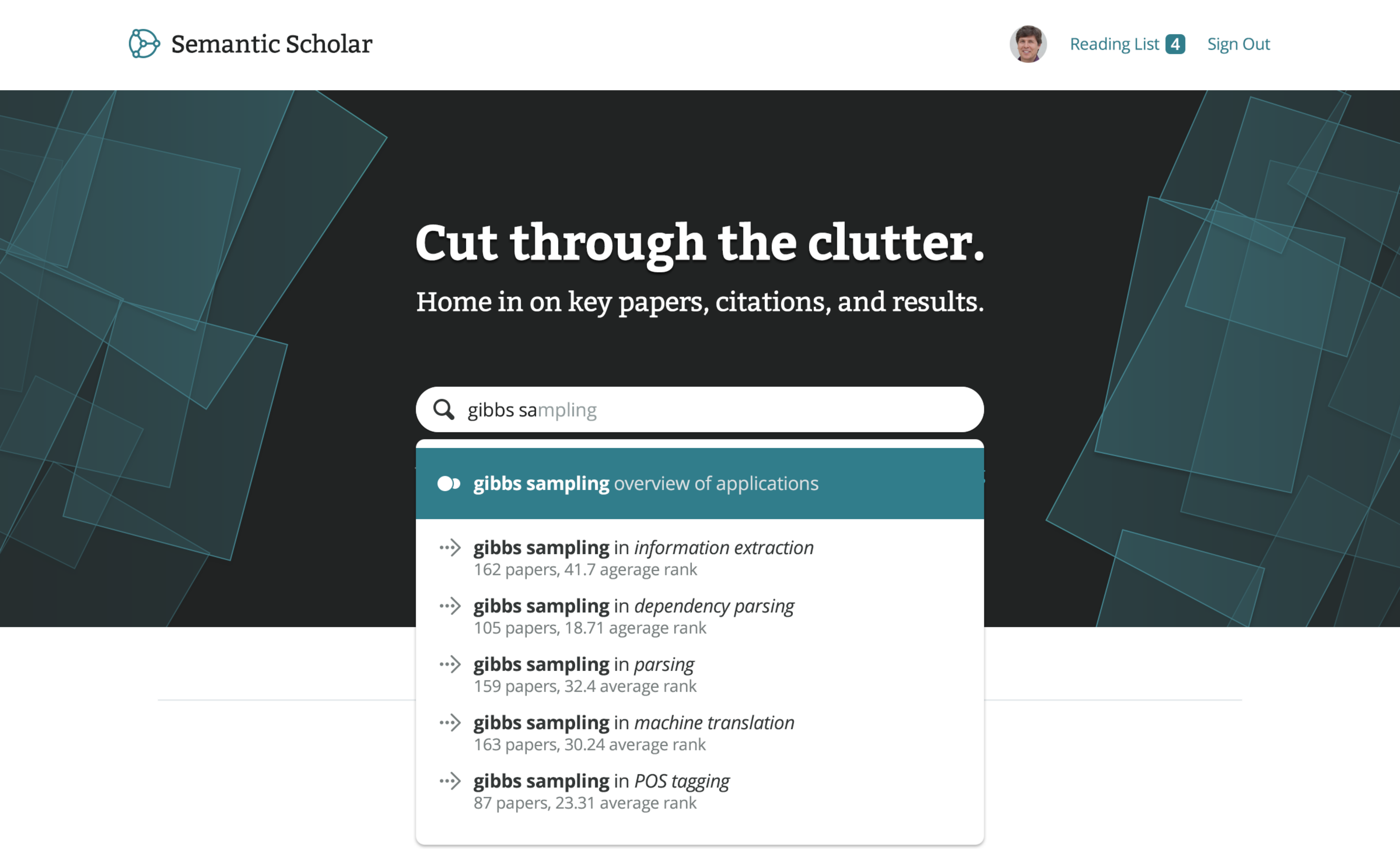
ALLEN INSTITUTE *for* ARTIFICIAL INTELLIGENCE

Provenance with Text Understanding

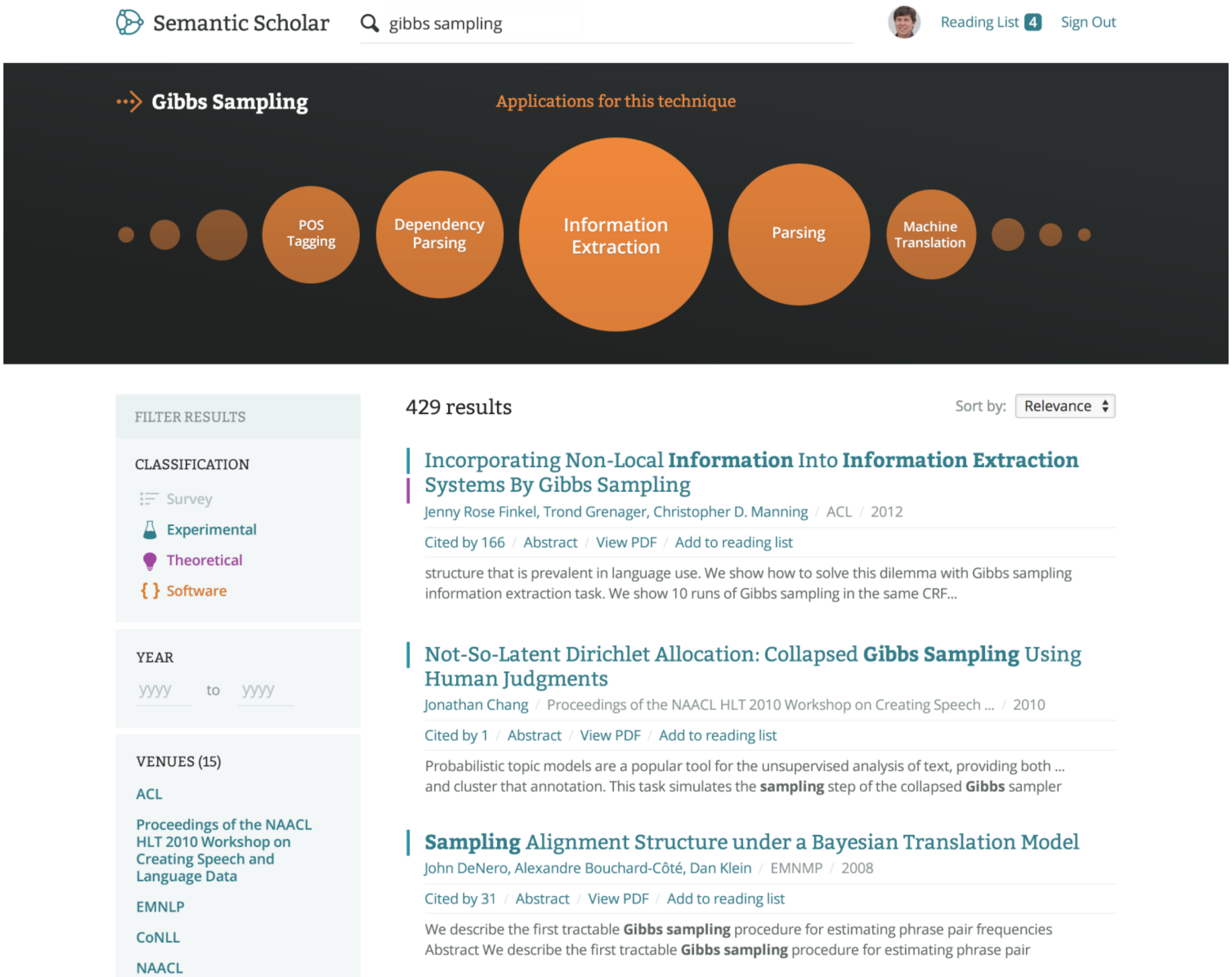


ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

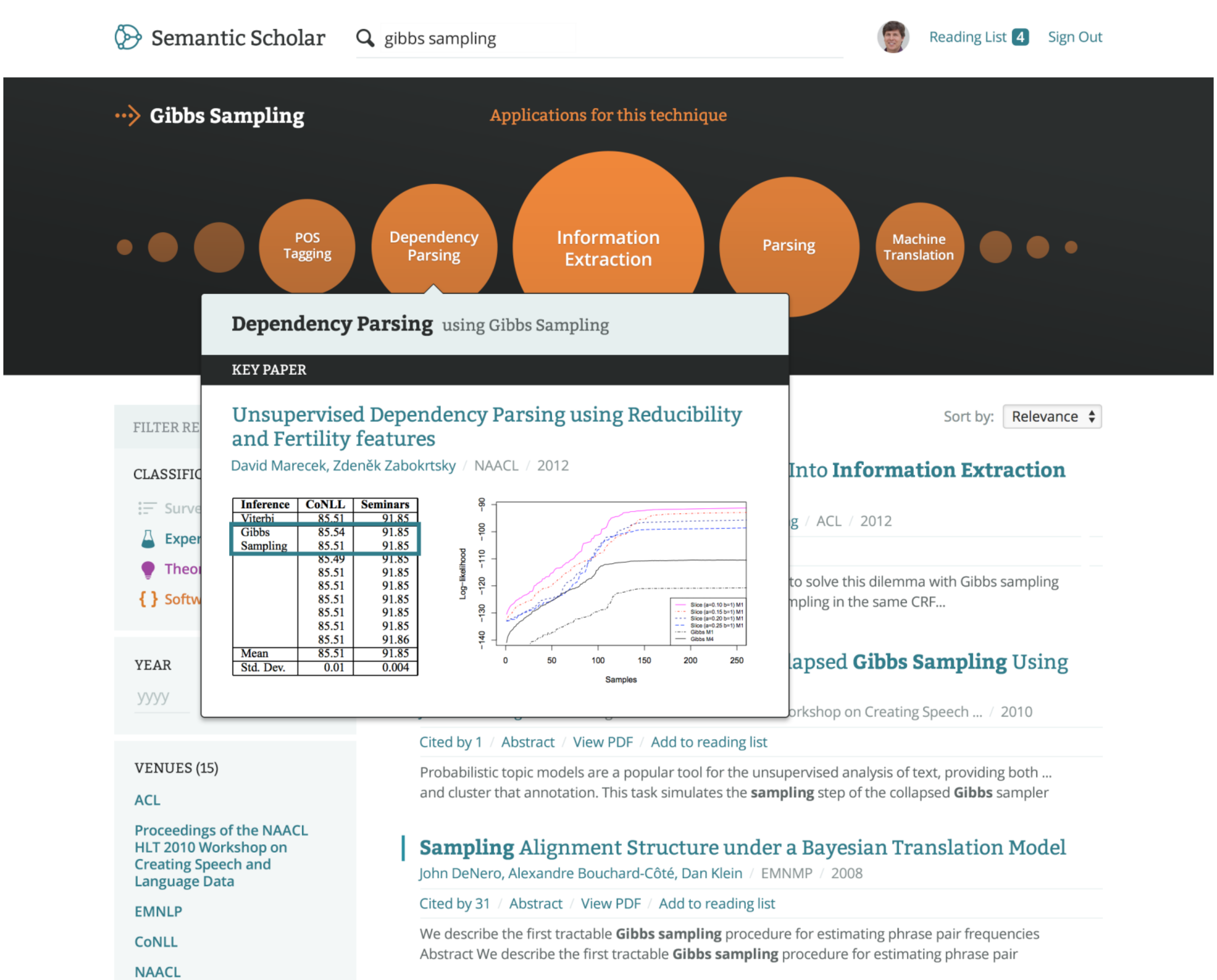
Applications of a Technique



Applications of a Technique



Provenance with Diagram Understanding



ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

Semantic Scholar to Launch in 2015

Sign up for notification here:

allenai.org/semantic-scholar.html

ALLEN INSTITUTE *for* ARTIFICIAL INTELLIGENCE

Eric Horvitz

“It's the **absence** of AI technologies that is **already** killing people.”

ALLEN INSTITUTE *for* ARTIFICIAL INTELLIGENCE

The Semantic Scholar Vision

“What if a cure for an intractable cancer is hidden within the tedious reports on thousands of clinical studies? In 20 years' time, AI will be able to read — and more importantly, understand — scientific text. These AI readers will be able to connect the dots between disparate studies to identify novel hypotheses and suggest experiments which would otherwise be missed. AI-based discovery engines will help to find the answers to science's thorniest problems and ultimately revolutionize science.”

Allen Institute for Artificial Intelligence

[Wired Magazine, September, 2015]

AI2 Core Projects:

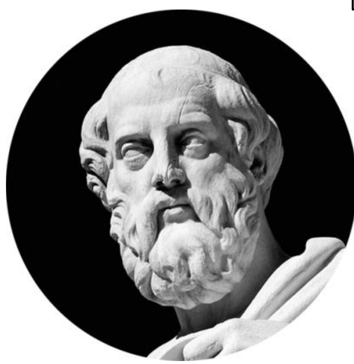
Science QA

Aristo

Semantic Search over papers

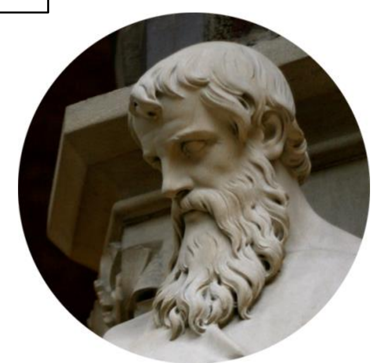
Semantic Scholar

Knowledge from images & diagrams



Plato

SAT Math QA



Euclid

AI2 Methodology

1. Grand Challenge Problems
2. Data-driven rather than single mechanism
3. Ambitious goals, but measurable progress
4. **Idea:** augment Turing Test with battery of standardized tests to measure AI progress
5. Semantic Scholar = AI to help Scientists

The Future of AI:
The View from AI2

Oren Etzioni

www.allenai.org

ALLEN INSTITUTE *for* ARTIFICIAL INTELLIGENCE

The View from AI2 (allenai.org)