Semantic Scholar

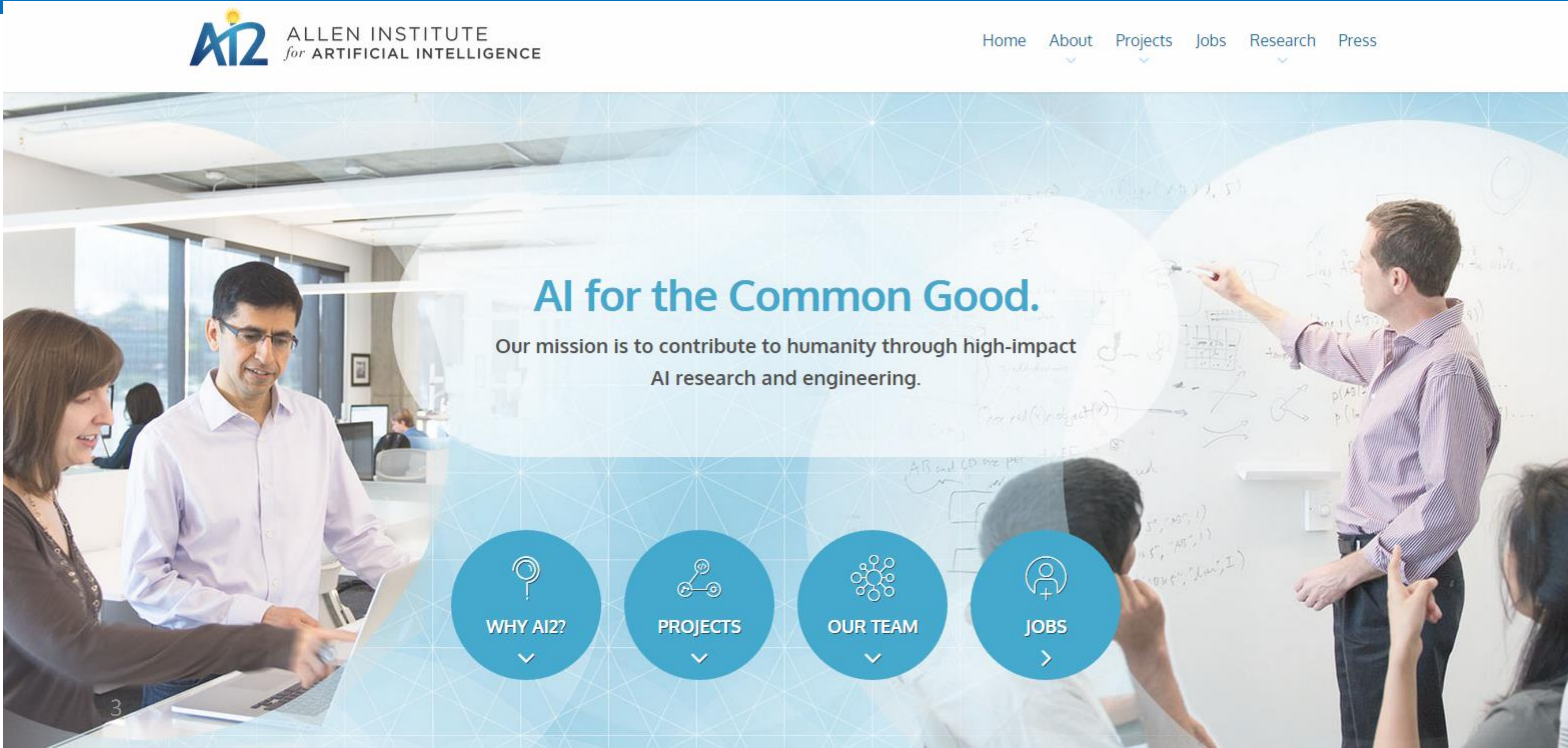
Oren Etzioni, CEO
Allen Institute for AI (AI2)

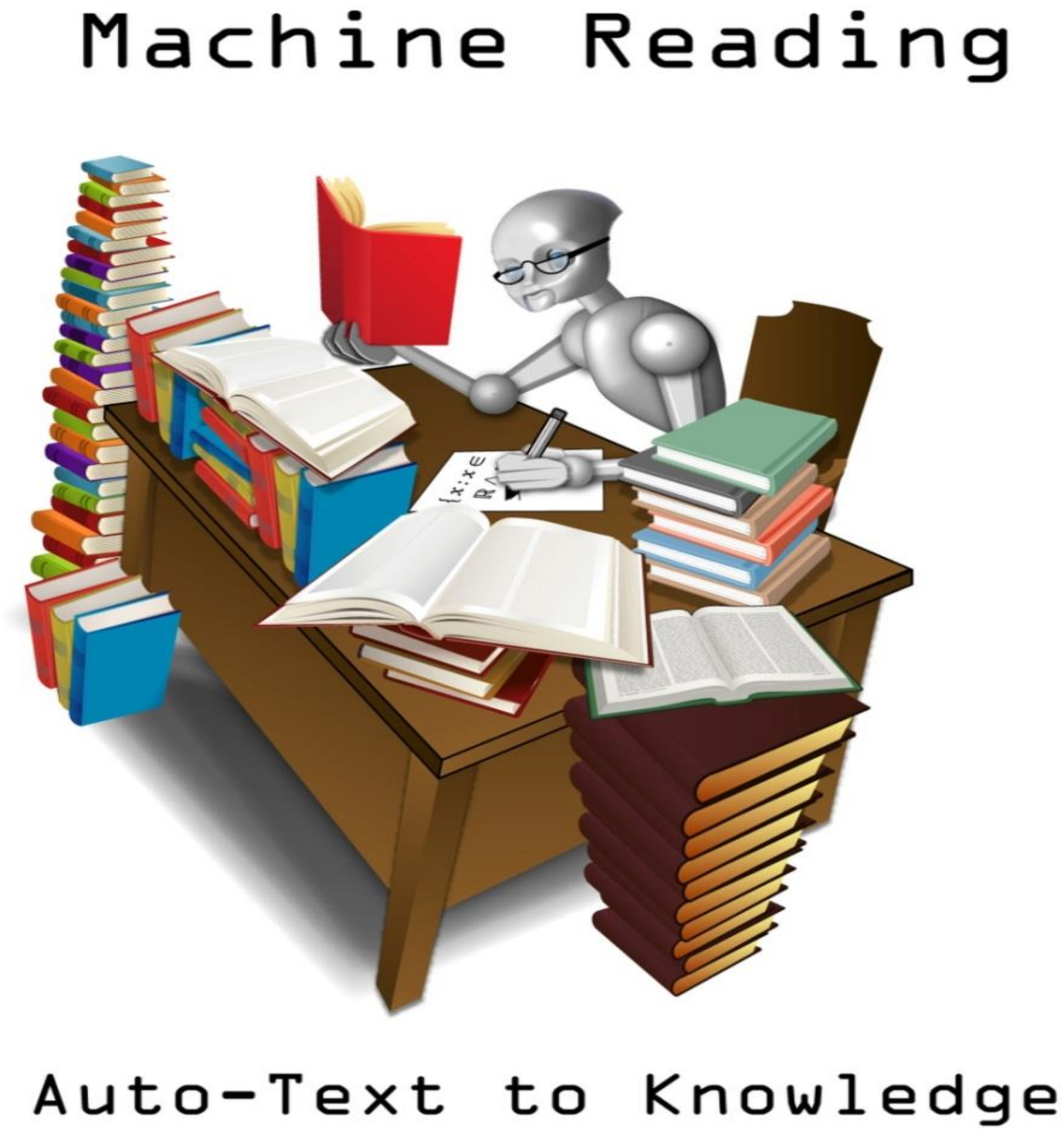
Microsoft Research
Faculty Summit
2015

Semantic Scholar

Oren Etzioni, CEO
Allen Institute for AI (AI2)

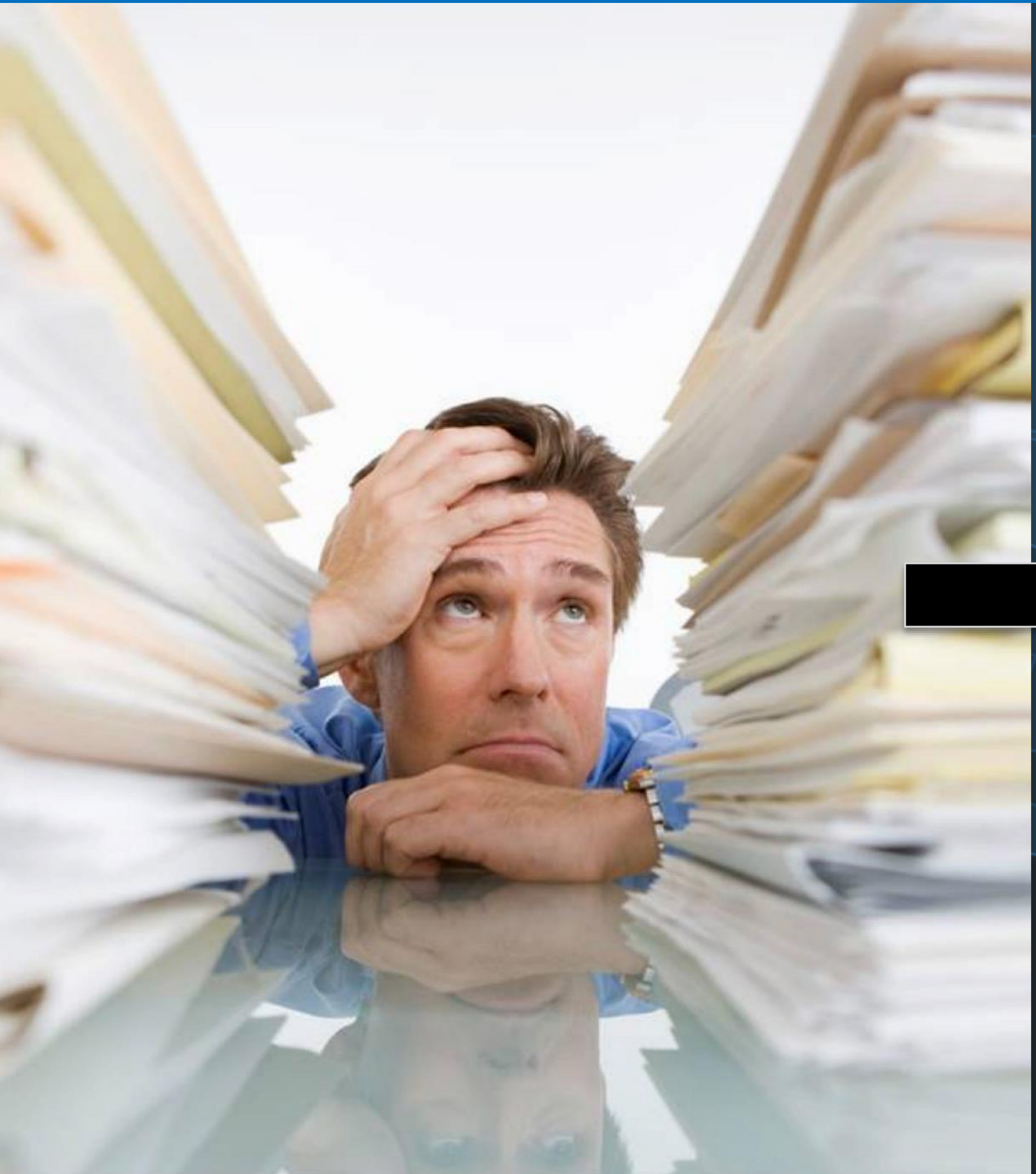
http://allenai.org

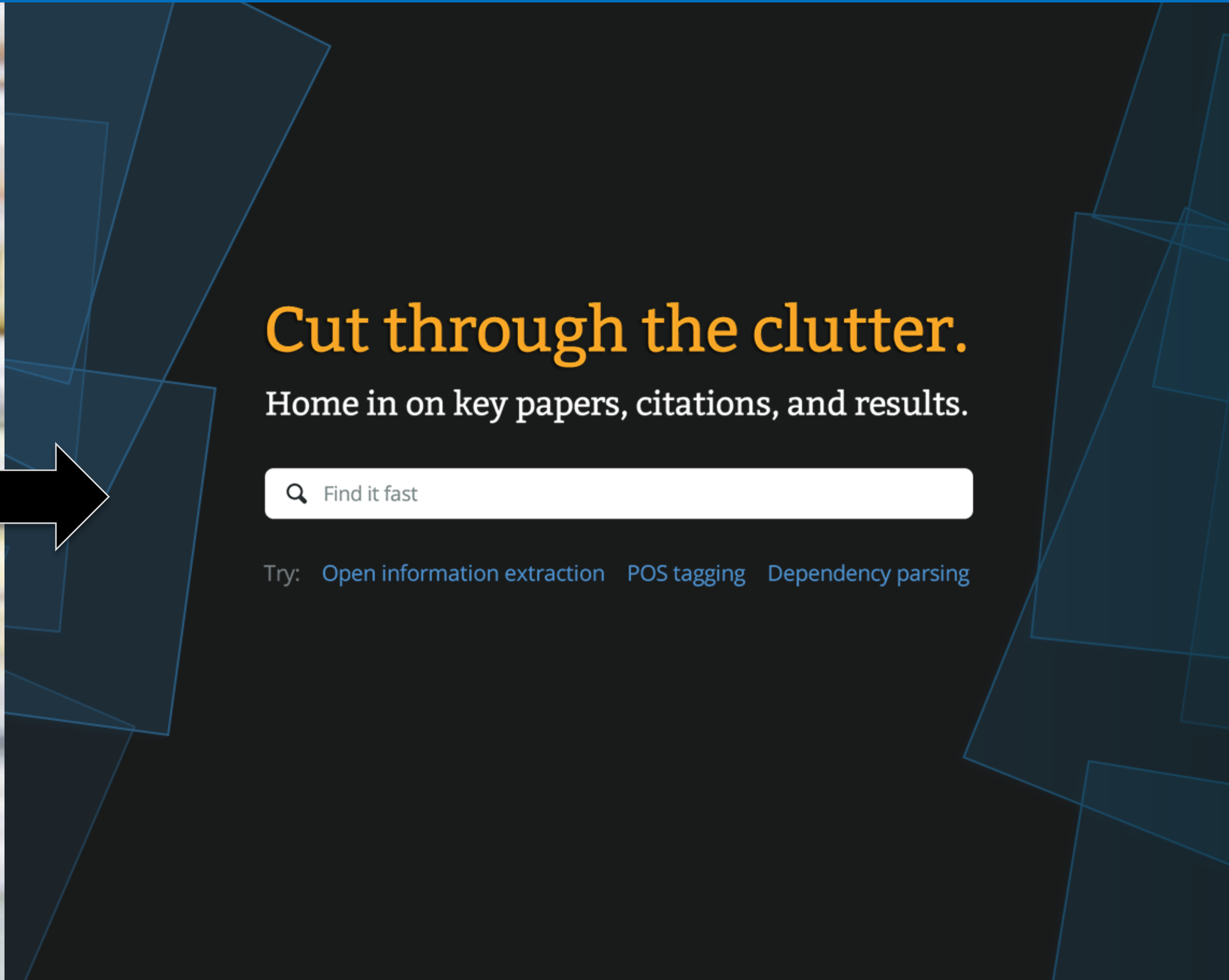




 Source: DARPA, Machine Reading initiative

Semantic Scholar (Scientific Search)

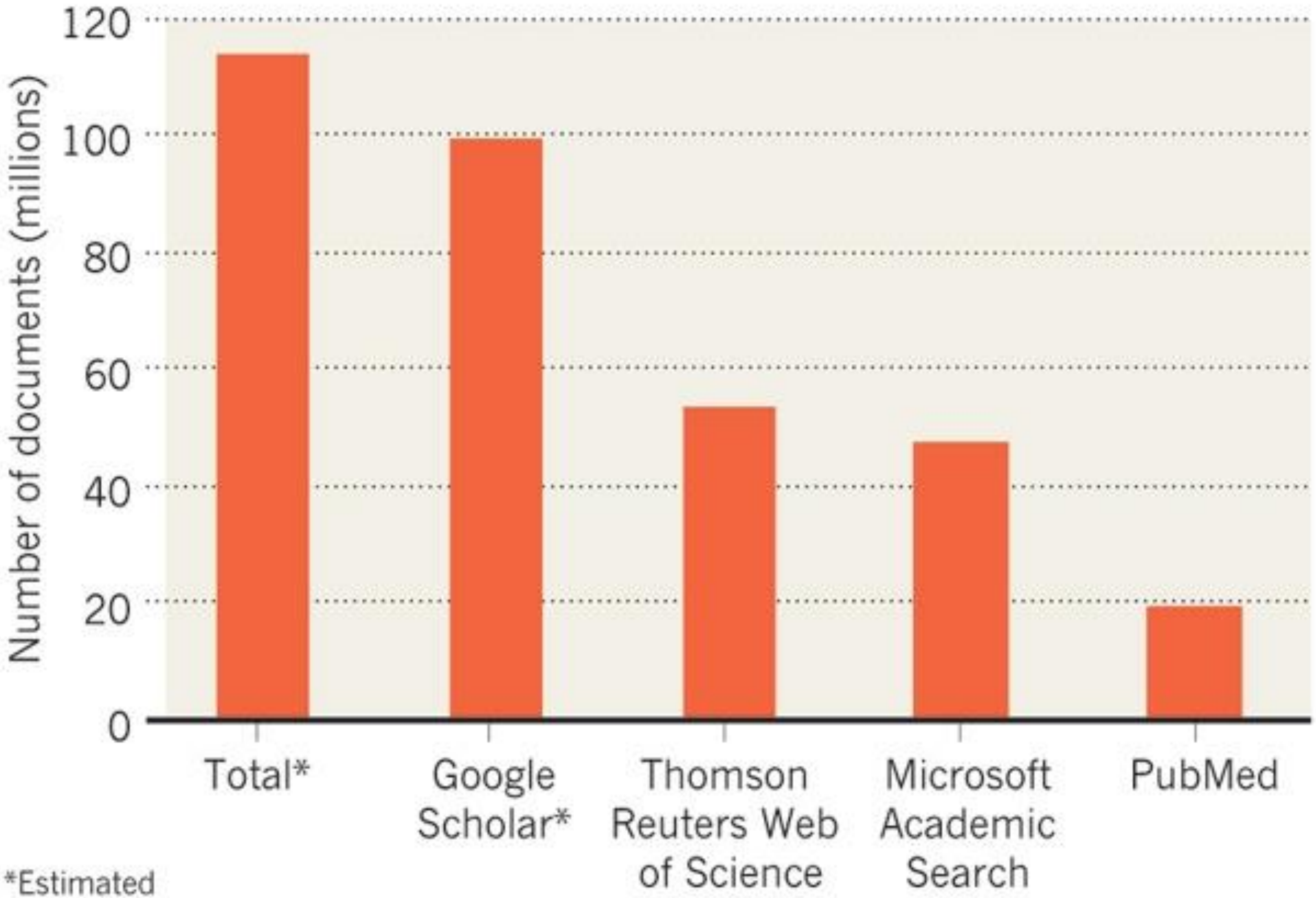




Motivation

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. [PDF] from berkeley.edu

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 [PDF] from aclweb.org Discover UIUC Full Text

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 [PDF] from aaai.org

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

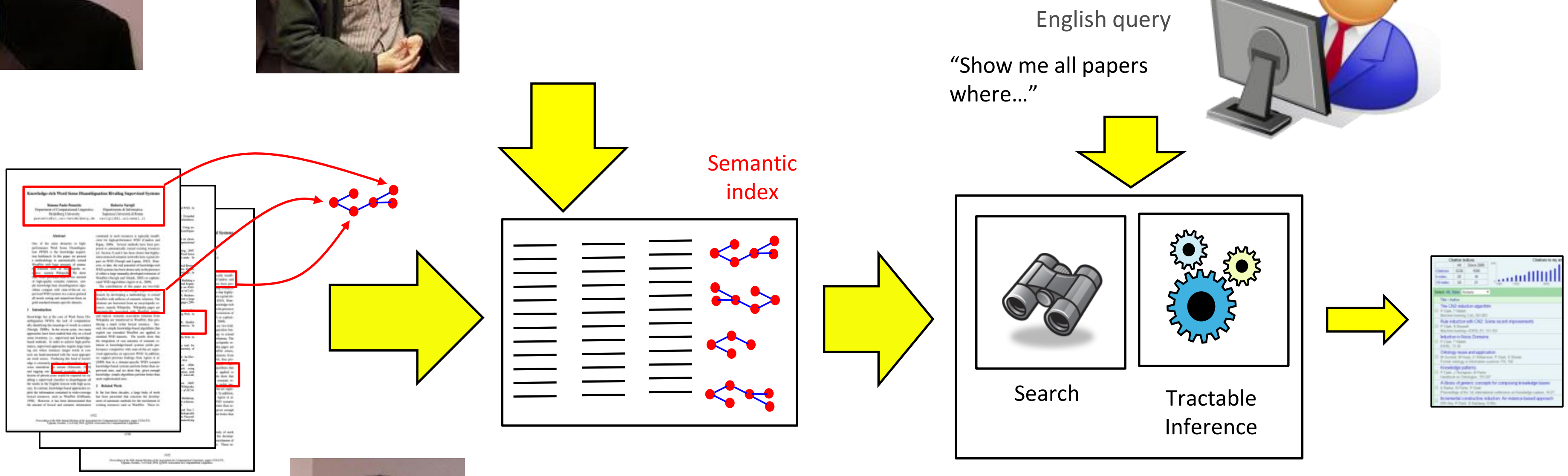
[PDF] Open **information extraction** for the web [PDF] from aaai.org

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



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.

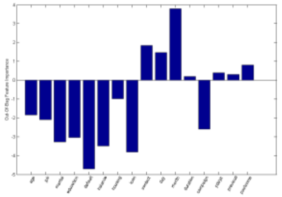
The need for pre-existing datasets can be avoided by using language itself as the source of the schema. This is the approach taken by OpenIE (Etzioni et al., 2008). Here surface patterns between mentions of concepts serve as relations. This approach requires no supervision and has tremendous flexibility, 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.

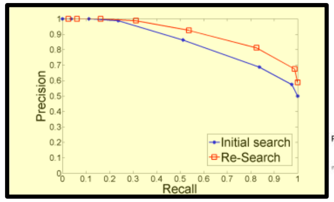
One way to gain generalization is to cluster textual surface forms that have similar meaning (Lin and Pantel, 2001; Pantel et al., 2007; Yates and Etzioni, 2009; Yao et al., 2011). While the clusters discovered by all these methods usually contain semantically related items, closer inspection invariably 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*

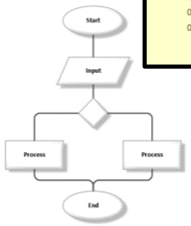
74

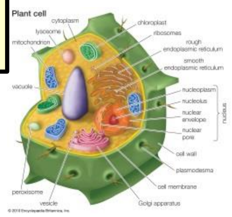
Proceedings of NAACL-HLT 2013, pages 74–84,
Atlanta, Georgia, 9–14 June 2013. ©2013 Association for Computational Linguistics

Extract Figures

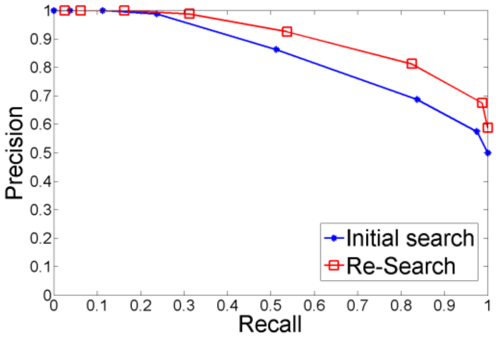








Classify Figures



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

SEMANTIC SCHOLAR

Google login

Cut through the clutter.

Home in on key papers, citations, and results.

discriminative model in machine translation

Try: Open information extraction POS tagging Dependency parsing

Made by

ALLEN INSTITUTE *for* ARTIFICIAL INTELLIGENCE

Powered by ACL Anthology Reference Corpus

SEMANTIC SCHOLAR

discriminative model machine translation

Google login

Filter results

Survey
Experimental
Theoretical
Software

Year

1992 2013

Authors

Klein, Dan (24)
Manning, Christopher D. (23)
Liu, Qun (16)
Ney, Hermann (16)
+ More

Venues

ACL (273)
EMNLP (147)
NAACL (111)
CL (59)
+ More

Topics

Statistical Machine Translation (29)
Penn Treebank (27)
Machine Translation (26)
GIZA++ (21)

967 results

Sort by: Relevance

A **Discriminative** Latent Variable **Model** for Statistical **Machine Translation**

Philip Blunsom, Trevor Cohn, Miles Osborne • ACL • 2008

Abstract: Large-scale **discriminative machine translation** promises to further the state-of-the-art, but has

Cited by 43 View PDF Add to reading list

A **Discriminative** Syntactic Word Order **Model** for **Machine Translation**

Pi-Chuan Chang, Kristina Toutanova • ACL • 2007

Abstract: We present a global **discriminative** statistical word order **model** for **machine translation**. Our **model** combines syntactic movement and surface movement information, and is discriminatively trained to choose among possible word orders. We show that combining **discriminative** training

Cited by 1 View PDF Add to reading list

Unsupervised **Discriminative** Language **Model** Training for **Machine Translation** using Simulated Confusion Sets

Zhifei Li, Ziyuan Wang, Sanjeev P. Khudanpur, Jason M. Eisner • COLING • 2010

Abstract: An unsupervised **discriminative** training procedure is proposed for estimating a language **model** (LM

Cited by 2 View PDF Add to reading list

Discriminative Feature-Tied Mixture Modeling for Statistical **Machine Translation**

Bing Xiang, Abraham Ittycheriah • ACL • 2011

Abstract: In this paper we present a novel **discriminative** mixture **model** for statistical **machine translation**

Cited by 1 View PDF Add to reading list

Inducing a **Discriminative** Parser to Optimize **Machine Translation** Reordering

Graham Neubig, Taro Watanabe, Shinsuke Mori • EMNLP • 2012

Abstract: This paper proposes a method for learning a **discriminative** parser for **machine translation**

Cited by 4 View PDF Add to reading list

Discriminative Training And Maximum Entropy Models For Statistical **Machine Translation**

Franz Josef Och, Hermann Ney • ACL • 2002

Abstract: We present a framework for statistical **machine translation** of natural languages based on direct

A Discriminative Latent Variable Model for Statistical Machine Translation

Philip Blunsom, Trevor Cohn, Miles Osborne • ACL • 2008 • View PDF • Add to reading list

Details | 43 Citing papers

80th	14.0%	37.8
SCHOLAR RANK PERCENTILE	6 OF 43 CITING PAPERS ARE KEY CITATIONS	CITATION RATE

Abstract

Large-scale discriminative machine translation promises to further the state-of-the-art, but has failed to deliver convincing gains over current heuristic frequency count systems. We argue that a principle reason for this failure is not dealing with multiple, equivalent translations. We present a translation model which models derivations as a latent variable, in both training and decoding, and is fully discriminative and globally optimised. Results show that accounting for multiple derivations does indeed improve performance. Additionally, we show that regularisation is essential for maximum conditional likelihood models in order to avoid degenerate solutions.

Selected Citation Contexts

Key Citation

Fast Generation of Translation Forest for Large-Scale SMT Discriminative Training
Xinyan Xiao, Yang Liu, Qun Liu, Shouxun Lin • 2011

1. Recent work have shown that SMT benefits a lot from exploiting large amount of features (Liang et al., 2006; Tillmann and Zhang, 2006; Watanabe et al., 2007; **Blunsom et al., 2008**; Chiang et al., 2009).
2. We use the forest to train a log-linear model with a latent variable as describe in **Blunsom et al. (2008)**.
3. Researchers have propose many learning algorithms to train many features: perceptron (Shen et al., 2004; Liang et al., 2006), minimum risk (Smith and Eisner, 2006; Li et al., 2009), MIRA (Watanabe et al., 2007; Chiang et al., 2009), gradient descent (**Blunsom et al., 2008**; Blunsom and Osborne, 2008).

Key Citation

The CMU-ARK German-English Translation System
Chris Dyer, Kevin Gimpel, Jonathan H. Clark, Noah A. Smith • 2011

Problems

- Statistical Machine Translation

Techniques

- MERT
- Hiero
- Discriminative Model

Datasets

- Europarl

Topics

- Discriminative Latent Variable Model
- Model for Statistical Machine Translation

Semantic Scholar

4

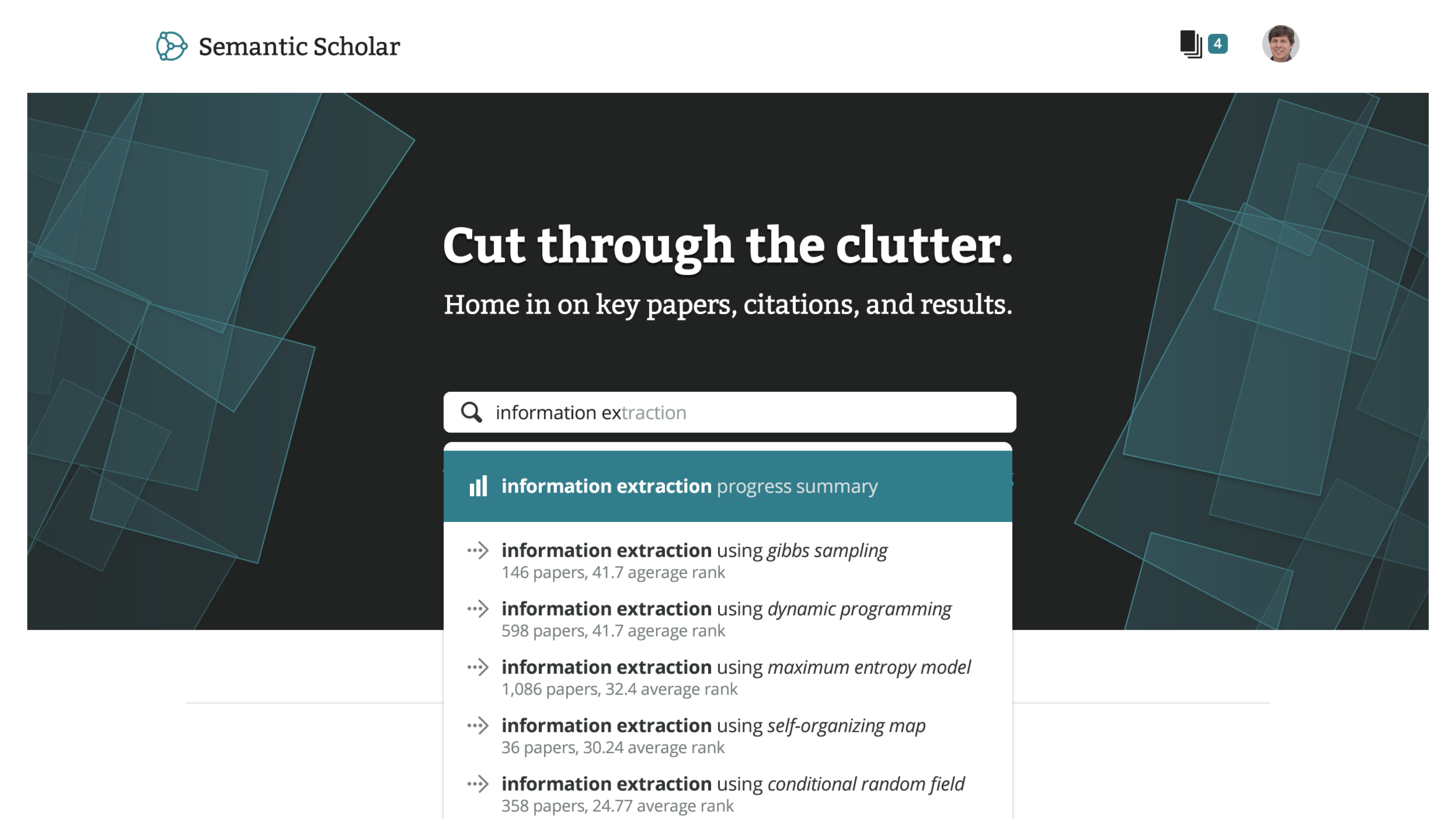
Cut through the clutter.

Home in on key papers, citations, and results.

Find it fast

Try: Open information extraction POS tagging Dependency parsing

Made By AI2 • Powered by ACL Anthology Reference Corpus


Semantic Scholar
4
Cut through the clutter.
Home in on key papers, citations, and results.
information extraction
information extraction progress summary
information extraction using gibbs sampling
146 papers, 41.7 agerage rank
information extraction using dynamic programming
598 papers, 41.7 agerage rank
information extraction using maximum entropy model
1,086 papers, 32.4 average rank
information extraction using self-organizing map
36 papers, 30.24 average rank
information extraction using conditional random field
358 papers, 24.77 average rank

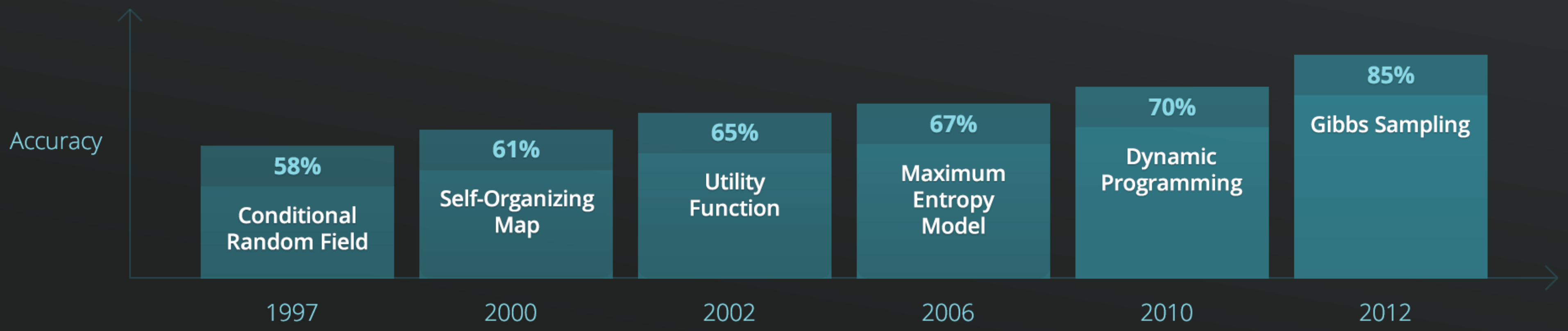
Semantic Scholar

information extraction

4

Information Extraction

Progress Summary



FILTER RESULTS

CLASSIFICATION

- Survey
- Experimental
- Theoretical
- { } Software

YEAR

6,962 results

Sort by: Relevance

Incorporating Non-Local **Information** Into **Information Extraction** Systems By Gibbs Sampling

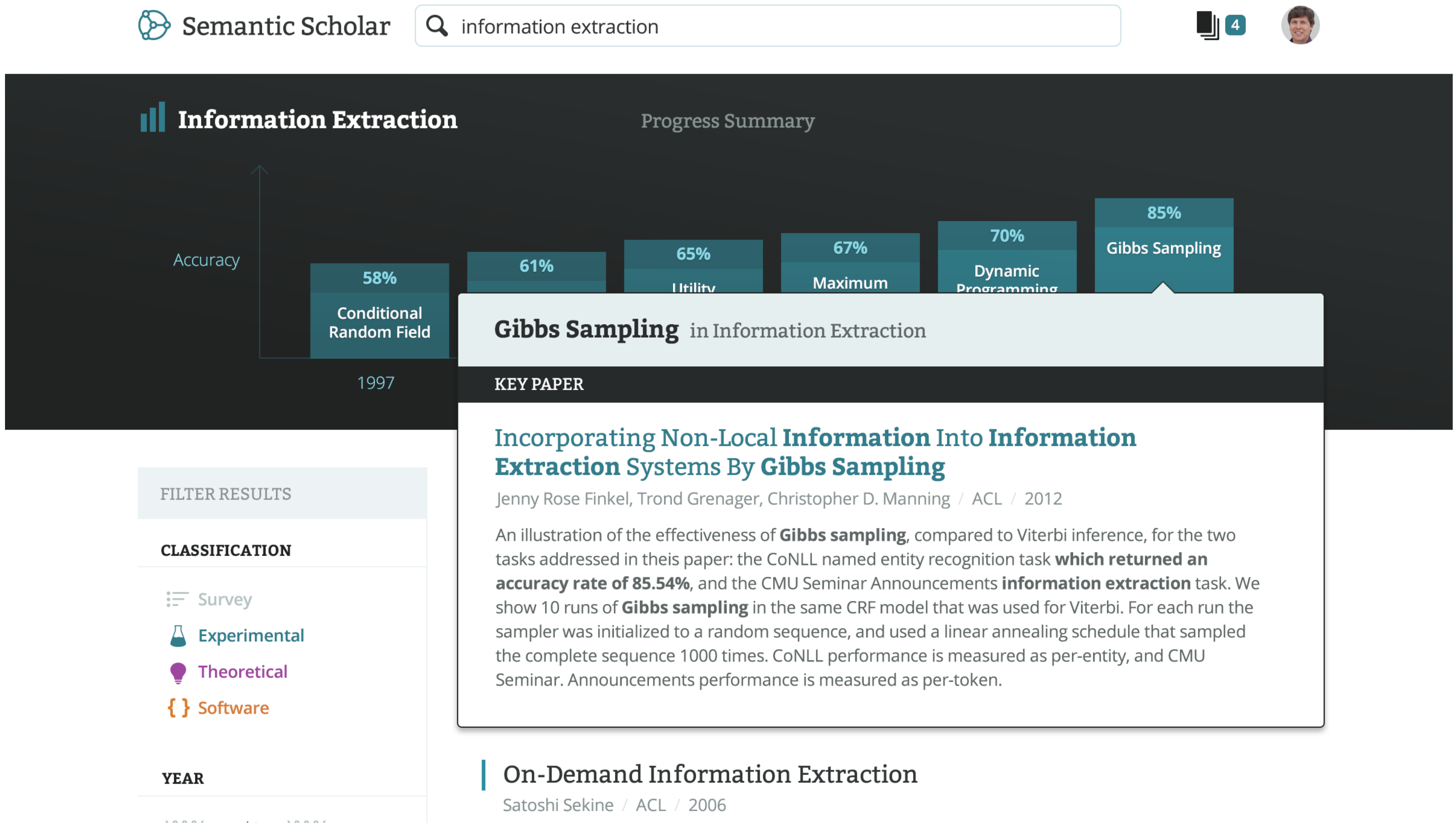
Jenny Rose Finkel, Trond Grenager, Christopher D. Manning / ACL / 2005

Cited by 166 / Abstract / View PDF / Add to reading list

structure that is prevalent in language use. We show how to solve this dilemma with Gibbs sampling **information extraction** task. We show 10 runs of Gibbs sampling in the same CRF...

On-Demand Information Extraction

Satoshi Sekine / ACL / 2006


Semantic Scholar
information extraction
4
Information Extraction
Progress Summary
Accuracy
58%
Conditional Random Field
1997
61%
65%
67%
Maximum
70%
Dynamic Programming
85%
Gibbs Sampling
Gibbs Sampling in Information Extraction
KEY PAPER
Incorporating Non-Local Information Into Information Extraction Systems By Gibbs Sampling
Jenny Rose Finkel, Trond Grenager, Christopher D. Manning / ACL / 2012
An illustration of the effectiveness of Gibbs sampling, compared to Viterbi inference, for the two tasks addressed in theis paper: the CoNLL named entity recognition task which returned an accuracy rate of 85.54%, and the CMU Seminar Announcements information extraction task. We show 10 runs of Gibbs sampling in the same CRF model that was used for Viterbi. For each run the sampler was initialized to a random sequence, and used a linear annealing schedule that sampled the complete sequence 1000 times. CoNLL performance is measured as per-entity, and CMU Seminar. Announcements performance is measured as per-token.
FILTER RESULTS
CLASSIFICATION
Survey
Experimental
Theoretical
Software
YEAR
On-Demand Information Extraction
Satoshi Sekine / ACL / 2006

Semantic Scholar

4

Cut through the clutter.

Home in on key papers, citations, and results.

gibbs sampling

gibbs sampling overview of applications

gibbs sampling in *information extraction*
162 papers, 41.7 agerage rank

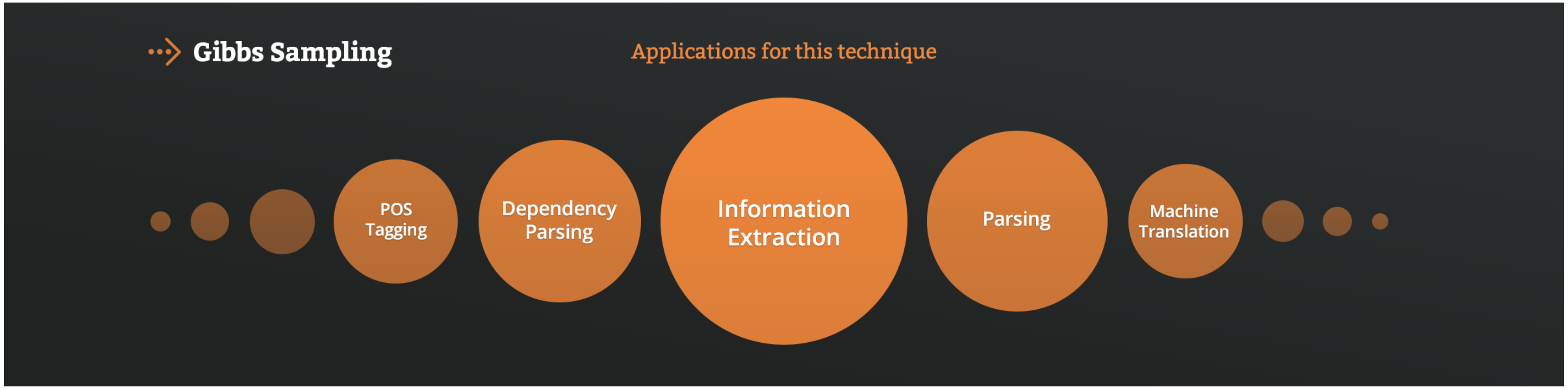
gibbs sampling in *dependency parsing*
105 papers, 18.71 agerage rank

gibbs sampling in *parsing*
159 papers, 32.4 average rank

gibbs sampling in *machine translation*
163 papers, 30.24 average rank

gibbs sampling in *POS tagging*
87 papers, 23.31 average rank

Semantic Scholar | gibbs sampling | 4



FILTER RESULTS

CLASSIFICATION

- Survey
- Experimental
- Theoretical
- { } Software

YEAR

yyyy to yyyy

VENUES (15)

- ACL
- Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data
- EMNLP

429 results

Sort by: Relevance

Incorporating Non-Local **Information** Into **Information Extraction** Systems By Gibbs Sampling

Jenny Rose Finkel, Trond Grenager, Christopher D. Manning / ACL / 2012

Cited by 166 / Abstract / View PDF / Add to reading list

structure that is prevalent in language use. We show how to solve this dilemma with Gibbs sampling information extraction task. We show 10 runs of Gibbs sampling in the same CRF...

Not-So-Latent Dirichlet Allocation: Collapsed **Gibbs Sampling** Using Human Judgments

Jonathan Chang / Proceedings of the NAACL HLT 2010 Workshop on Creating Speech ... / 2010

Cited by 1 / Abstract / View PDF / Add to reading list

Probabilistic topic models are a popular tool for the unsupervised analysis of text, providing both ... and cluster that annotation. This task simulates the **sampling** step of the collapsed **Gibbs** sampler

Sampling Alignment Structure under a Bayesian Translation Model

John DeNero, Alexandre Bouchard-Côté, Dan Klein / EMNMP / 2008

Cited by 31 / Abstract / View PDF / Add to reading list

We describe the first tractable **Gibbs sampling** procedure for estimating phrase pair frequencies

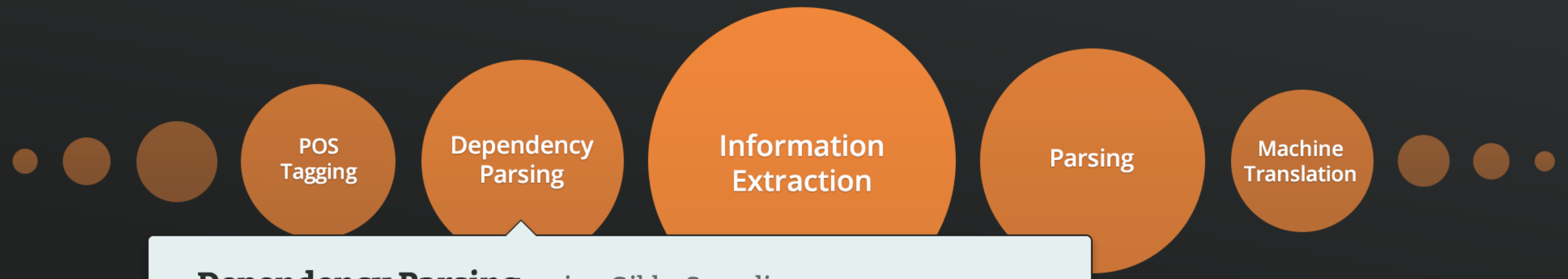
Semantic Scholar

gibbs sampling

4

Gibbs Sampling

Applications for this technique



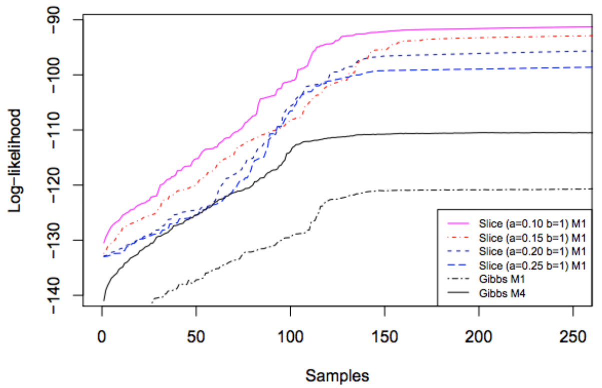
Dependency Parsing using Gibbs Sampling

KEY PAPER

Unsupervised Dependency Parsing using Reducibility and Fertility features

David Marecek, Zdeněk Zabokrtsky / NAACL / 2012

Inference	CoNLL	Seminars
Viterbi	85.51	91.85
Gibbs Sampling	85.54	91.85
	85.51	91.85
	85.49	91.85
	85.51	91.85
	85.51	91.85
	85.51	91.85
	85.51	91.85
	85.51	91.85
	85.51	91.86
Mean	85.51	91.85
Std. Dev.	0.01	0.004



FILTER RE

CLASSIFIC

Surve

Exper

Theor

{ } Softw

YEAR

Sort by: Relevance

Into **Information Extraction**

g / ACL / 2012

to solve this dilemma with Gibbs sampling
mpling in the same CRF...

lapsed **Gibbs Sampling** Using

Semantic Scholar

A Maximum-Entropy-Inspired Parser

Eugene Charniak / ANLP / 2000

Upvote 38 | Downvote | Add to reading list | Comment 3 | Share | View PDF

TOP 1% SCHOLAR RANK PERCENTILE

8.2% OF CITING PAPERS ARE IMPORTANT

52.3 CITATION RATE

ABSTRACT

We present a new parser for parsing down to Penn tree-bank style parse trees that **achieves 90.1% average precision/recall for sentences of length 40 and less, and 89.5% for sentences** of length 100 and less when trMned and tested on the previously established [5,9,10,15,17] "standard" sections of the Wall Street Journal treebank. This represents a 13% decrease in error rate over the best single-parser results on this corpus [9]. The major technical innovation is tire use of a "maximum-entropy-inspired" model for conditioning and smoothing that let us successfully to test and combine many different conditioning events. We also present some partial results showing the effects of different conditioning information, including a surprising 2% improvement due to guessing the lexical head's pre-terminal before guessing the lexical head.

CITATION CONTEXTS

"We train an English-to-Chinese translation system using the FBIS corpus, where 73,597 sentence pairs are selected as the training data, and 500 sentence pairs with no more than 25 words on the Chinese side are selected for both the development and test data.1 **Charniak (2000)**s parser, trained on the Penn Treebank, is used to generate the English syntax trees."

Semantic Role Features for Machine Translation
Ding Liu, Daniel Gildea / 2000

"A number of robust statistical parsers that oer solutions to these problems have now become available (**Charniak, 2000**; Collins, 1999; Henderson, 2003), but they typically produce CFG constituency data as output, trees that do not express long-distance dependencies."

Semantic Role Features for Machine Translation
Ding Liu, Daniel Gildea / 2000

▸ View more citation contexts

Details
- Citing papers 371
- Important citing papers 30
- Paper clusters 16

PROBLEMS
- Dependency Parsing
- Information Extraction

TECHNIQUES
- Parsing
- Markov Grammar

DATA SETS
- Penn Tree-bank

TOPICS
- Wall Street Journal Penn Treebank
- Linguistic Information Processing

Semantic Scholar

A Maximum-Entropy-Inspired Parser

Eugene Charniak / ANLP / 2000

Upvote 38 Downvote Add to reading list Comment 3 Share View PDF

TOP 1% SCHOLAR RANK PERCENTILE

8.2% OF CITING PAPERS ARE IMPORTANT

52.3 CITATION RATE

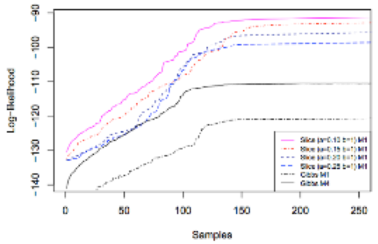
CITATIONS BY YEAR

ABSTRACT

We present a new parser for parsing down to Penn tree-bank style parse trees that **achieves 90.1% average precision/recall for sentences of length 40 and less, and 89.5% for sentences** of length 100

Parser	LR	LP	CB	0CB	2CB
	≤ 40 words (2245 sentences)				
Char97	87.5	87.4	1.00	62.1	86.1
Coll99	88.5	88.7	0.92	66.7	87.1
Char00	90.1	90.1	0.74	70.1	89.6
	≤ 100 words (2416 sentences)				
Char97	86.7	86.6	1.20	59.9	83.2
Coll99	88.1	88.3	1.06	64.0	85.1
Ratna99	86.3	87.5			
Char00	89.6	89.5	0.88	67.6	87.7

Figure 1: Parsing results compared with previous work



Treebank, is used to generate the English syntax trees."

Semantic Role Features for Machine Translation
Ding Liu, Daniel Gildea / 2000

"A number of robust statistical parsers that oer solutions to these problems have now become available (**Charniak, 2000**; Collins, 1999; Henderson, 2003), but they typically produce CFG constituency data as output, trees that do not express long-distance dependencies."

Semantic Role Features for Machine Translation
Ding Liu, Daniel Gildea / 2000

View more citation contexts

Details

- Citing papers 371
- Important citing papers 30
- Paper clusters 16

PROBLEMS

- Dependency Parsing
- Information Extraction

TECHNIQUES

- Parsing
- Markov Grammar

DATA SETS

- Penn Tree-bank

TOPICS

- Wall Street Journal Penn Treebank
- Linguistic Information Processing

Semantic Scholar

4

A Maximum-Entropy-Inspired Parser

Eugene Charniak / ANLP / 2000

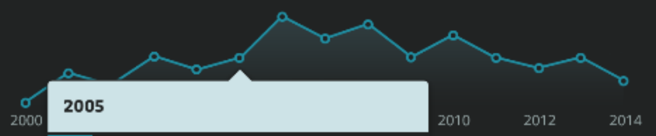
Upvote 38 Downvote Add to reading list Comment 3 Share View PDF

TOP 1%
SCHOLAR RANK PERCENTILE

8.2%
OF CITING PAPERS ARE IMPORTANT

52.3
CITATION RATE

CITATIONS BY YEAR



27 Citing papers / 2 are important (7.4%)

ABSTRACT

We present a new parser for parsing down to Penn tree-bank style parse trees that **achieves 90.1% average precision/recall for sentences of length 40 and less, and 89.5% for sentences** of length 100 and less when trMned and tested on the previously established [5,9,10,15,17] "standard" sections of the Wall Street Journal treebank. This represents a 13% decrease in error rate over the best single-parser results on this corpus [9]. The major technical innovation is tire use of a "maximum-entropy-inspired" model for conditioning and smoothing that let us successfully to test and combine many different conditioning events. We also present some partial results showing the effects of different conditioning information, including a surprising 2% improvement due to guessing the lexical head's pre-terminal before guessing the lexical head.

CITATION CONTEXTS

"We train an English-to-Chinese translation system using the FBIS corpus, where 73,597 sentence pairs are selected as the training data, and 500 sentence pairs with no more than 25 words on the Chinese side are selected for both the development and test data.1 **Charniak (2000)**s parser, trained on the Penn Treebank, is used to generate the English syntax trees."

Semantic Role Features for Machine Translation
Ding Liu, Daniel Gildea / 2000

"A number of robust statistical parsers that oer solutions to these problems have now become available (**Charniak, 2000**; Collins, 1999; Henderson, 2003), but they typically produce CFG constituency data as output, trees that do not express long-distance dependencies."

Semantic Role Features for Machine Translation
Ding Liu, Daniel Gildea / 2000

View more citation contexts

Details

Citing papers 371

Important citing papers 30

Paper clusters 16

PROBLEMS

Dependency Parsing

Information Extraction

TECHNIQUES

Parsing

Markov Grammar

DATA SETS

Penn Tree-bank

TOPICS

Wall Street Journal Penn Treebank

Linguistic Information Processing

Semantic Scholar

A Maximum-Entropy-Inspired Parser

Eugene Charniak / ANLP / 2000

▲ Upvote 38 · Downvote · Add to reading list · Comment 3 · Share · View PDF

TOP 1% SCHOLAR RANK PERCENTILE

8.2% OF CITING PAPERS ARE IMPORTANT

52.3 CITATION RATE

371 Citing papers **30 ARE IMPORTANT (8.3%)**

Sort by: Relevance

A Robust And Hybrid Deep-Linguistic Theory Applied To Large-Scale Parsing

Jenny Rose Finkel, Trond Grenager, Christopher D. Manning / ACL / 2012

IMPORTANT Cited by 166 / Abstract / View PDF / Add to reading list

Citation contexts:

"We train an English-to-Chinese translation system using the FBIS corpus, where 73,597 sentence pairs are selected as the training data, and 500 sentence pairs with no more than 25 words on the Chinese side are selected for both the development and test data.1 **Charniak (2000)**s parser, trained on the Penn Treebank, is used to generate the English syntax trees."

"The model The total 74,597 sentence pairs used in experiments are those in the FBIS corpus whose English part can be parsed using **Charniak (2000)**s parser."

TAG, Dynamic Programming, and the Perceptron for Efficient, Feature-Rich Parsing

Jonathan Chang / Proceedings of the NAACL HLT 2010 Workshop on Creating Speech ... / 2010

IMPORTANT Cited by 1 / Abstract / View PDF / Add to reading list

Citation contexts:

"A number of robust statistical parsers that oer solutions to these problems have now become available (**Charniak, 2000**; Collins, 1999; Henderson, 2003), but they typically produce CFG constituency data as output, trees that do not express long-distance dependencies."

"Statistical disambiguation such as (Collins and Brooks, 1995) for PP-attachment or (Collins, 1997; **Charniak, 2000**) for generative parsing greatly improve disambiguation, but as they model by imitation instead of by understanding, complete soundness has to remain elusive."

Comparing And Combining Finite-State And Context-Free Parsers

Details

Citing papers 371

Important citing papers 30

Paper clusters 16

PROBLEMS

Dependency Parsing

Information Extraction

TECHNIQUES

Parsing

Markov Grammar

DATA SETS

Penn Tree-bank

TOPICS

Wall Street Journal Penn Treebank

Linguistic Information Processing

Launch in Q4 2015

Sign up for notification here:

allenai.org/semantic-scholar.html