

# Semantic Search via XML Fragments: A High-Precision Approach to IR

Jennifer Chu-Carroll  
IBM T.J. Watson Research Center

April 2006

Joint work w/ John Prager, Krzysztof Czuba,  
David Ferrucci & Pablo Duboue

## Overview

- **Why Semantic Search?**
- **Corpus Analysis for Semantic Search**
- **Using XML Fragments for Semantic Search**
  - Three XML Fragments operators
  - Applied to address four query-time semantic needs
- **Evaluation**
- **Pseudo demo**
- **Conclusions**

## Why Semantic Search?

- **Plenty of people find what they need with keywords**
  - Web search
  - Desktop search
- **These applications work well with keywords because**
  - Web search: corpus is huge
  - Desktop search: looking for known item
- **Keyword search does not work quite so well for**
  - Enterprise search: limited-size corpus
  - Exploratory research: user is researching a topic of interest and does not know what they are looking for
- **Semantic search can help in the latter classes of applications**

# Semantic Search vs. Keyword Search

SEARCH: Going rate for leasing a billboard near Triborough Bridge



Top hits from popular search engines miss the mark...

*Keywords may match*  
*BUT **WRONG** content returned*  
*And right content **MISSED***

# SEARCH: Going rate for leasing a billboard near Triborough Bridge

## [Wired New York Forum - Long Island City Development](#)

"it's Right over The 59th Street **Bridge**." ... it's **going** to become our SoHo. The Powerhouse is most likely **going** to become a performing Arts facility for ...  
[www.wirednewyork.com/forum/printthread.php?t=5160&pp=50](http://www.wirednewyork.com/forum/printthread.php?t=5160&pp=50) - 150k - Supplemental Result - [Cached](#) - [Similar pages](#)

## [Long Island City Development \[Archive\] - Wired New York Forum](#)

... Creek to the **Triborough Bridge** - would become an attractive destination. ... Several Long Island City sites are now in play - Junger himself's **going** ...  
[www.wirednewyork.com/forum/archive/index.php/t-5160.html](http://www.wirednewyork.com/forum/archive/index.php/t-5160.html) - 255k - [Cached](#) - [Similar pages](#)  
 [ [More results from www.wirednewyork.com](http://www.wirednewyork.com) ]

## [\[PDF\] Austin's Wireless Future](#)

File Format: PDF/Adobe Acrobat - [View as HTML](#)  
 In Section 2 we explore the **near**-term future of wireless-- ubiquitous mobile ... applications such as the **Triborough Bridge** installation, wireless sensors ...  
[www.ic2.org/publications/AustinsWirelessFuture.pdf](http://www.ic2.org/publications/AustinsWirelessFuture.pdf) - [Similar pages](#)

## [\[PDF\] Funding of the High-Speed Rail System](#)

File Format: PDF/Adobe Acrobat - [View as HTML](#)  
**Triborough Bridge**. and Tunnel. Authority\*\*. Oklahoma. Turnpike Authority ... management fees, reserves / on **going** TI's, and **leasing** commissions. ...  
[www.cahighspeedrail.ca.gov/plan/pdf/Financial\\_Plan.pdf](http://www.cahighspeedrail.ca.gov/plan/pdf/Financial_Plan.pdf) - [Similar pages](#)

## [The rail magazine for The computer age ...](#)

Surprisingly, thanks to 9600 baud modems, lower **rates**, and automated access ... his previous positions included President, **Triborough Bridge** and tunnel ...  
[www.railfan.net/railpix/text/ro\\_1\\_6.txt](http://www.railfan.net/railpix/text/ro_1_6.txt) - Supplemental Result - [Similar pages](#)

## [Checklist 1989](#)

**Triborough Bridge** and Tunnel Authority, selected procurement and contract ... Distribution of high school graduates and college-**going rate**, New York State. ...  
[www.nysl.nysed.gov/edocs/education/check89.htm](http://www.nysl.nysed.gov/edocs/education/check89.htm) - 513k - [Cached](#) - [Similar pages](#)

## [\[PDF\] www.icc.utexas.edu/publications/AustinsWirelessFut...](#)

File Format: PDF/Adobe Acrobat - [View as HTML](#)  
 Supplemental Result - [Similar pages](#)

“Going”, “Bridge”, “Rate”,  
“Leasing”, etc.

But All Misses

**SEARCH:** Going rate for leasing a billboard near Triborough Bridge

\* Were you looking for [Going rate for leaving a billboard near Triborough Bridge](#)

## Web Results

Page 1 of 4 results containing **Going rate for leasing a billboard near Triborough Bridge** (0.70 seconds)

### [Wired New York Forum - Long Island City Development](#)

... on Vernon Boulevard **near** the Queensboro **Bridge**, the founders of Silvercup ... Newtown Creek ... the **Triborough Bridge** - would become an attractive ... Pearson St. - where he's **going** to co ... 20-story ...

[www.wirednewyork.com/forum/printthread.php?t=5160&pp=50](http://www.wirednewyork.com/forum/printthread.php?t=5160&pp=50) [Cached page](#)

### [FINANCIAL PLAN](#)

**Triborough Bridge** and Tunnel Authority\*\* Oklahoma Turnpike Authority ... Average Annual Growth **Rate**) 1979-1982 10.9% 1980-1990 5 ... Leveraged **Leasing** of Rail Rolling Stock When rail rolling ...

[www.cahighspeedrail.ca.gov/plan/pdf/Financial\\_Plan.pdf](http://www.cahighspeedrail.ca.gov/plan/pdf/Financial_Plan.pdf) [Cached page](#) PDF file

### [NEW YORK: Stadiums/Arenas \(NETS, JETS, YANKS, and others\) \[Archive ...](#)

... backward from the adjacent **Triborough Bridge** elevated roadway that leads to ... I was **going** to the name of one of my companies on it to get ... The area **near** the Continental Arena has become construction ...

[forum.skyscraperpage.com/archive/index.php/t-65433.html](http://forum.skyscraperpage.com/archive/index.php/t-65433.html) [Cached page](#)

### [http://odur.let.rug.nl/ftp/pub/prolog-app/DutchBrillTagging/TaggerSoftware/Bin\\_and\\_Data/LEXICON.WSJ](http://odur.let.rug.nl/ftp/pub/prolog-app/DutchBrillTagging/TaggerSoftware/Bin_and_Data/LEXICON.WSJ)

**Billboard** NNP Billboarding NNP Billed VBN Billerica NNP Billheimer NNP Billiards NNP NN Billie NNP Billing NN Billings NNS NNP Billion CD Billionaire NN Billions NNS Billmeyer NNP Billock NNP Bills NNS ...

[odur.let.rug.nl/ftp/pub/prolog-app/DutchBrillTagging/TaggerSoftware/Bin\\_and\\_Data/LEXICON.W...](http://odur.let.rug.nl/ftp/pub/prolog-app/DutchBrillTagging/TaggerSoftware/Bin_and_Data/LEXICON.W...) [Cached page](#)

Another Search Engine  
Some Different Hits  
but  
Still All Misses

Makes you wonder...  
Could it be out there?

Windows Live Local  
powered by Virtual Earth Beta

[Use current map view]

What: Business name or category Where: Address, city, or other place

Welcome Scratch Pad Locate Me Permalink Add Pushpin Directions

road aerial **bird's eye**

WELCOME

Welcome to Windows Live Local

What's New?  
We've changed our name, changed our look, and added lots of new features!

Here are a few of the great new features:

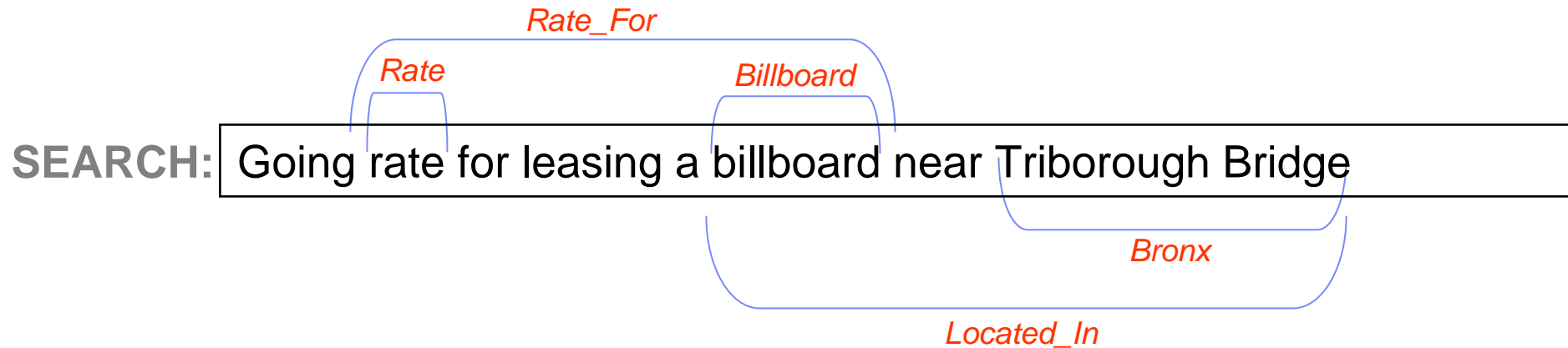
- See incredible [Bird's Eye images](#).
- Create [driving directions](#) and then customize how they [print](#).
- Add your own [custom pushpins](#) anywhere on the map (just [right-click](#) on the map).

Remarkably...With some **Location** semantics

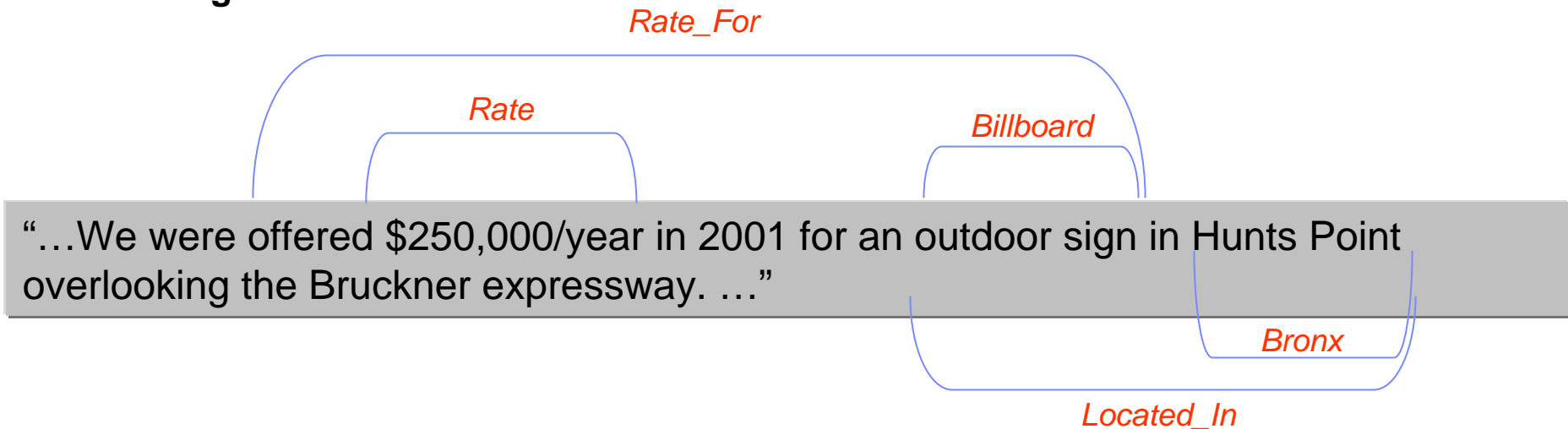
We can quickly find Hi-Res examples of area of interest

But NOT the information we need

# Semantic Search Can Improve Recall

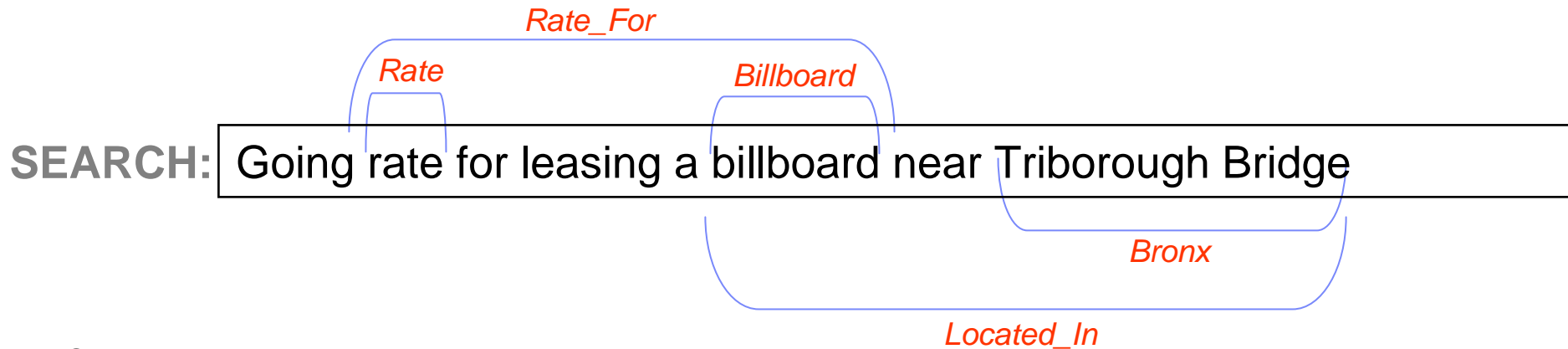


**No Keywords in Common  
But a good “hit”**

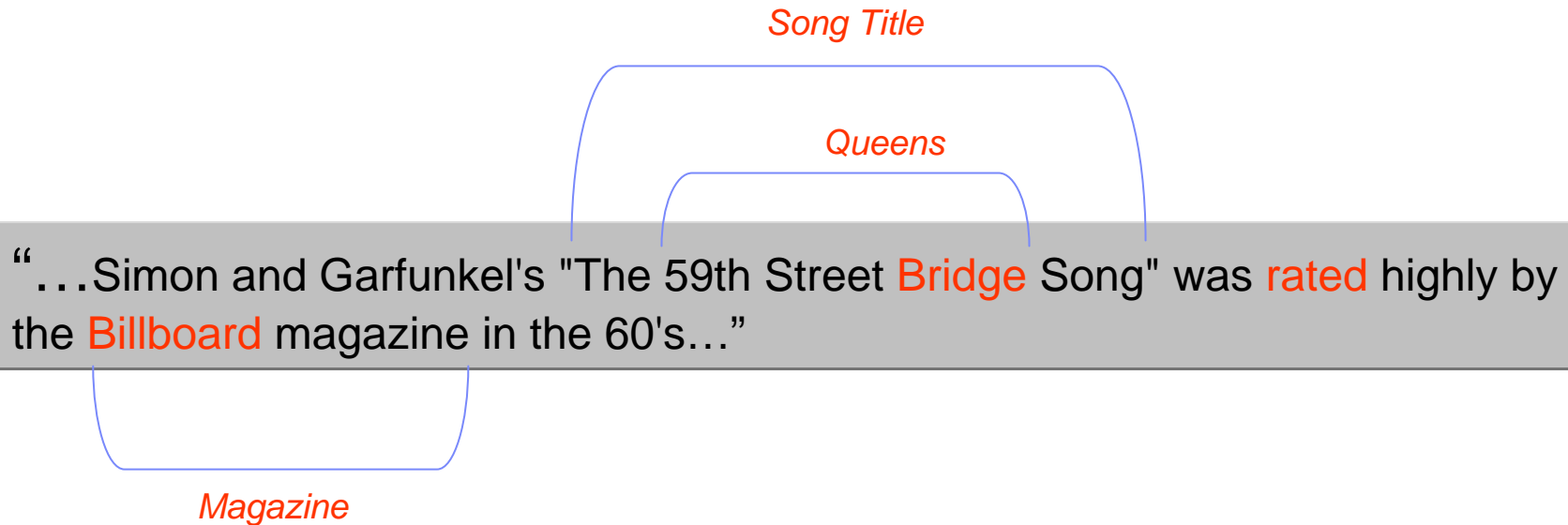




# Semantic Search Can Improve Precision



**Common Keywords**  
**Bad Semantic Match**



# Semantic Search on Semantically-Encoded Corpora

- Corpus: XML documents
  - <Book>
    - <Title>My Life</Title>
    - <Author>
      - <FirstName>Bill</FirstName>
      - <LastName>Clinton</LastName>
    - </Author>
    - <Publisher>Knopf</Publisher>
    - <PubDate>2004</PubDate>
  - </Book>
  - <Book>
    - <Title>The Survivor: Bill Clinton in the White House</Title>
    - <Author>
      - <FirstName>John</FirstName>
      - <LastName>Harris</LastName>
    - </Author>
    - <Publisher>Random House</Publisher>
    - <PubDate>2005</PubDate>
  - </Book>
- **Semantics enables more precise queries**
  - <Book>Bill Clinton</Book> retrieves both documents
  - <Book>
    - <Author>Bill Clinton</Author></Book> retrieves only the first document
- **XML document corpora**
  - Automatically generated from databases
  - Manually marked up
  - From the semantic web (vision)
- **But... most existing electronic documents are unannotated or sparsely annotated**

## Corpus Analysis for Semantic Search

- **Apply semantic search to previously unannotated documents**

- Use existing text analysis techniques for document processing
  - Named entity recognition: PERSON, ORGANIZATION, DATE,...
  - Relationship extraction: ALIAS, OWNER, EXPORTSTO,...
- Annotate and index documents with extracted semantic information
- Leverage automatic annotations for semantic search

- **Example**

<BirthPlaceOf> <BirthDateOf> <Alias> <Person> President Clinton  
</Person> was born <Person> William Jefferson Blythe IV </Person>  
</Alias> on <Date> August 19, 1946 </Date> </BirthDateOf>, in  
<City> Hope, Arkansas </City> </BirthPlaceOf>, three months after his  
father died in a traffic accident.

## XML Fragments for Semantic Search

### ■ XML Fragments query language

- Keyword queries augmented with XML tags
  - `<CeoOf> <Person> Center </Person> </CeoOf>`
- Supports classic query operators
  - Phrase: applies to content within XML tags
    - `<Person> “Thomas Ryan” </Person>`
  - “+” and “-”: apply to content within XML tags or XML fragments
    - `<Organization> +Clinton Institute </Organization>`
    - `+<Organization> +Clinton Institute </Organization> -Hillary`
- Query language specifies valid syntactic representations but are semantically

### ■ JuruXML indexer and search engine

- Supports indexing of keywords and annotations
- Supports XML Fragments query language
- Developed at IBM Haifa Lab

# Three XML Fragment Operators

## ■ Conceptualization

- Generalizes a lexical string to a concept represented by that string
- “animal” vs. “<Animal></Animal>”

## ■ Restriction

- Constrains the XML tags in which keywords in relevant documents can appear
- “bass” vs. “<Animal> bass </Animal>”

## ■ Relation

- Specifies relations between terms that appear in relevant documents
- Syntactic: “<SubjectVerb> Unabomber kill </SubjectVerb>”
- Semantic: “<Kill> Unabomber <Person></Person> </Kill>”
- Pragmatic: “<HasNegOpinion> Clinton war Iraq </HasNegOpinion>”

## Application of XML Fragment Operators

- **XML Fragments operators**
  - Enhance the expressiveness of search queries
  - Should lead to more relevant search results
- **Operators applied to address four query-time semantic needs**
  - Specify target information
  - Disambiguate keywords
  - Specify search term context
  - Specify relations between search terms

## Target Information Specification

- Application of the **conceptualization** operator
- Uses XML tags to represent target information as concepts
- E.g., user wants to know the zip code of the White House
  - Keyword query: +“white house” +zip +code
  - Semantic query: +“white house” +<Zipcode></Zipcode>
  - Matches: “*The White House, 1600 Pennsylvania Avenue NW, Washington, DC 20500.*”
- Useful when
  - Query consists primarily of low idf terms; and
  - Target information does not frequently co-occur with other search terms
- Adopted in PIQUANT factoid QA system to represent answer types

## Search Term Disambiguation

- Application of the **restriction** operator
- Uses XML tags for disambiguation of keywords
- E.g., User wants to know George Washington's birthday
  - Query w/o disambiguation: `+<Date></Date> bear +George +Washington`
  - Retrieves: *"A bear was spotted on the campus of George Washington University yesterday"*
  - Query w/ disambiguation: `+<Date></Date> bear +<Person> +George +Washington </Person>`
- Useful when
  - Query term has multiple word senses in the corpus; and
  - User is interested in minority word sense
- Adopted in
  - PIQUANT QA for automatic disambiguation
  - SAW (Semantic Analysis Workbench) for interactive query refinement



## Search Term Context Specification

- Application of the **restriction** operator
- Uses XML tags to specify the context in which query terms should appear
  - subject/object
  - agent/patient
  - request/suggest/opinion
- E.g., User wants to know what was the Pentagon panel's position was with respect to the dispute over the US Navy training range in the island of Vieques
  - Keyword query: *+Pentagon panel +US +Navy training range island +Vieques*
  - Query w/ context: *+Pentagon panel +<Quotation> +US +Navy training range island +Vieques </Quotation>*
- Adopted in AQUAINT 2005 opinion pilot

## Relation Specification

- Application of the **relation** operator
- Uses XML tags to specify relations between query terms
- E.g., User wants to know what biological weapons Iraq possesses
  - Query w/o relation: `+Iraq +<BiologicalWeapon></BiologicalWeapon>`  
*possess*
  - Matches: *“Iraq possesses the technology to produce smallpox.”*
  - Relation query: `<Owner> +Iraq`  
`+<BiologicalWeapon></BiologicalWeapon></Owner>`
  - Matches: *“Iraq’s anthrax stockpile is estimated to be at least 8500 liters.”*
- Adopted in SAW for automatic relation query generation and interactive query refinement

# Evaluation

- **Corpora and corpora analysis**
  - AQUAINT corpus (3GB; 1M+ documents)
    - Annotated with ~100 named entity types
  - CNS corpus (~38,000 documents)
    - Annotated with ~100 named entity types and ~15 relation types
- **Standard test set and judgments used when possible**
- **In all experimental setups**
  - Application of XML Fragments operator constrains baseline query
  - Hypothesis: our applications of semantic search yield significant improvement in the precision of retrieved results
- **Evaluation metrics**
  - R-Prec: Precision at R (R is # of known relevant documents)
  - MAP: Mean Average Precision
  - Exact Precision

# Target Information Specification Evaluation

## ■ Experimental setup:

- Test set: 50 factoid questions from TREC 2005 QA document task with relevant document judgments
- Baseline query: keywords extracted from question
  - +White +House zip code
- Query with target: keywords + answer type
  - +White +House zip code +<Zipcode></Zipcode>

## ■ Evaluation Results

	<b>R-Prec</b>	<b>MAP</b>	<b>Exact Precision</b>
<b>Baseline Query</b>	<b>0.4219</b>	<b>0.4329</b>	<b>0.0817</b>
<b>Query w/ Target</b>	<b>0.4342</b>	<b>0.4505</b>	<b>0.1124</b>

# Search Term Disambiguation

## ■ Experimental setup:

- Test set: 50 factoid questions from TREC 2005 QA document task with relevant document judgments
- Baseline query:
  - +White +House zip code +<Zipcode></Zipcode>
- Query with disambiguation: select keywords disambiguated
  - +<Facility>+White +House</Facility> zip code +<Zipcode></Zipcode>

## ■ Evaluation Results

- Only 2 questions resulted in different queries

	<b>R-Prec</b>	<b>MAP</b>	<b>Exact Precision</b>
<b>Baseline Query</b>	<b>0.4464</b>	<b>0.4357</b>	<b>0.1443</b>
<b>Query w/ Disambiguation</b>	<b>0.4658</b>	<b>0.4409</b>	<b>0.1443</b>

# Search Term Context Specification

## ■ Experimental setup:

- Test set: 46 questions from the AQUAINT 2005 opinion pilot of the general form “*What does A think about B?*”
- Baseline query:
  - OpinionHolder + Opinion
- Query with context:
  - OpinionHolder + <Quotation> Opinion </Quotation>

## ■ Evaluation Results

	# Vital Nuggets	# Okay Nuggets
<b>Baseline Query</b>	<b>14</b>	<b>4</b>
<b>Query w/ Context</b>	<b>28</b>	<b>5</b>

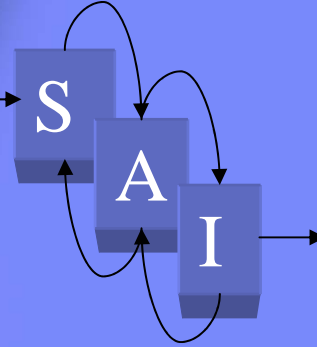
## Relation Specification Evaluation

### ■ Experimental setup:

- Test set: 25 semantic queries constructed from 10 relations in the national intelligence domain
- Relation query: constrained semantic search query with relation
  - `<ProducesWeapon>+Russia +<ChemicalWeapon></ChemicalWeapon></ProducesWeapon>`
- Baseline query: relation substituted by keywords
  - `+Russia +<ChemicalWeapon></ChemicalWeapon> produce`

### ■ Evaluation Results

	R-Prec	MAP	Exact Precision	# docs/Q
<b>Baseline Query</b>	<b>0.3895</b>	<b>0.4147</b>	<b>0.3530</b>	<b>9.72</b>
<b>Query w/ Relation</b>	<b>0.4139</b>	<b>0.4108</b>	<b>0.6108</b>	<b>6.32</b>



# SAW: Semantic Analysis Workbench



- ... MusicalInstrument
- + Organization
- + Person
- ... PhoneNumber
- ... Plant
- ... ProductionProcedure
- + Quantity
- ... Religion
- ... Sport
- ... SportsTeam
- + Timex
- ... Tool
- ... TravelReservation
- ... URL
- + Vehicle
- + Weapon
- + WmdAgent
- [-] TopRelation
  - + AgentArtifact
  - + ArgumentOfEvent
  - ... Casualties
  - ... Causes
  - ... DeliversWeapon
  - ... DestructionPotential
  - ... Discourse
  - + GpeAffiliation
  - ... HasIntention
  - ... HasProperty
  - ... HasQuality
  - ... HasReservation
  - ... HasWeaponCapability
  - ... HoldsDuring
  - ... MadePhoneCallTo
  - ... Motivation
  - + OrganizationalRelation
  - + PersonalAndSocial
  - + PersonalOrOrganization
  - ... Physical
  - ... ReservationFor
  - ... Residence
  - ... RouteEnd
  - + TemporalLink

New Query 1.ssq

**Descriptive comment**

Entity and Fact  Semantic, Natural-Language, and Keyword

What weapons does Iraq produce?

Interpret Query

Recall Precision

? ?produce

+ Weapon |

+ GPE | +Iraq

produce +<Weapon></Weapon> +<GPE>Iraq</GPE>

Show JURU from SSQE

Send this JURU to Piquant

• User can type keywords, phrases, or questions

• Query Analysis generates Semantic Search Query

Annotation View Hit List

500+ documents | 500+ passages

- ◆ [SADDAM CLOSE TO NUCLEAR WEAPON](#)  
The defector claimed that Iraq produced its 40kg of weapons-grade uranium using the electromagnetic isotope separation process.
- ◆ [HYPING THE IRAQI BOMB](#)  
On 22 November 1990, US President, George Bush, claimed that Iraq was close to producing nuclear weapons.
- ◆ [POLITICIANS IN THE LAB...AND SCUTTLING AN EASY WAY TO STOP NUCLEAR PROLIFER](#)  
Mark claimed that Iraq could produce two additional nuclear bombs with the 90 pounds of weapons-grade uranium that the Iraqis had.
- ◆ [EVIDENCE REVEALS IRAQI NUCLEAR POTENTIAL](#)  
He also confirmed that IAEA inspectors discovered 3 grams of plutonium that Iraq produced at its Tuwaiha reactor in violation of the NPT.
- ◆ [U.N. Envoy Says Arms Talks With Iraq Positive](#)  
Abstract: On 22 February 1995, UNSCOM head Rolf Ekeus stated that his inspectors, with IAEA assistance, had concluded "that Iraq has no capability to produce a nuclear bomb." Ekeus added, "Iraq has no intention of producing a nuclear bomb."
- ◆ [Iraq may still have the "talent" to produce nuclear weapons](#)  
According to the report, there are no indications that Iraq achieved its goal of producing nuclear weapons.
- ◆ [SAMs May Pose Most Direct Threat In Iraq's Arsenal](#)  
British intelligence said that Iraq could produce mustard gas immediately, nerve gas soon, and 350 liters of weapons-grade anthrax per week without monitoring by the United Nations.
- ◆ [IRAQI CHEMICAL ARMS: DIFFICULT TARGET](#)  
US analysts say Iraq could begin to produce nuclear weapons in three to five years, though practical devices mounted on missiles, for instance, would take a decade.
- ◆ [Not Available](#)  
There is no reliable information on whether Iraq is producing and stockpiling chemical weapons, or whether Iraq has - or will - establish its chemical warfare agent production facilities,' said the Defence Intelligence

• Intelligent passage selection  
• Hones in on relevant sections  
• Redefines the right grain



Web Images Groups News Froogle Local Desktop more »  
weapons Iraq produce Search Desktop Preferences Server Options Advanced Search

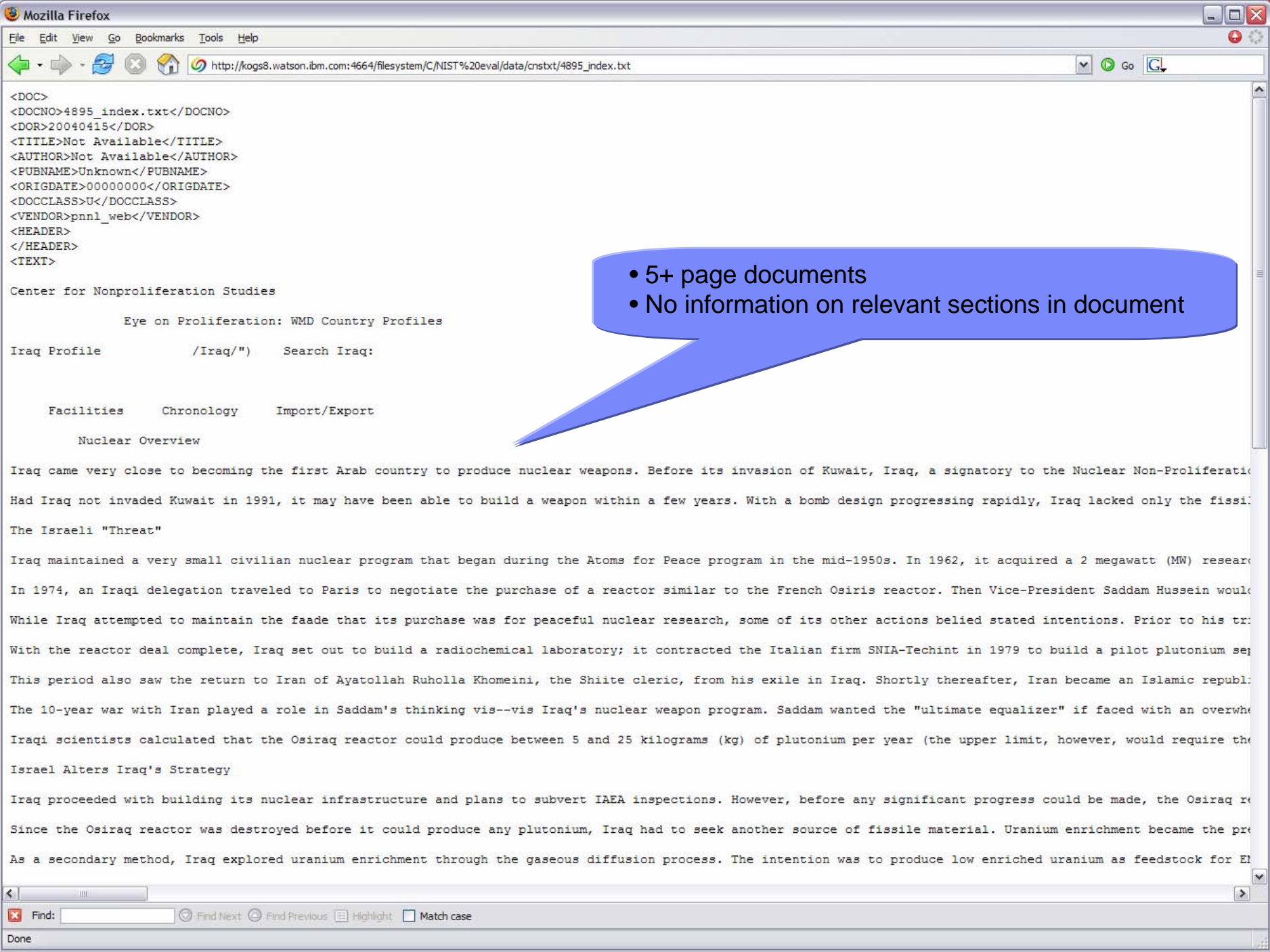
Host: kogs8.watson.ibm.com

Desktop: All - 0 emails - [237 files](#) - 0 web history - 0 chats - 0 other 1-10 of about 237 (0.02s)

[Remove from Index](#) | [Sorted by relevance](#) [Sort by date](#)

- [4895 index.txt](#)  
Iraq came very close to becoming the first Arab country to **produce** nuclear **weapons**. Before its ..  
Proliferation: WMD Country Profiles **Iraq** Profile /Iraq/ Search **Iraq**  
C:\NIST eval\data\cnstxt\4895\_index.txt - [Open folder](#) - [1 cached](#) - [Oct 26](#)
- [d1394\\_cbw053003.txt](#)  
VENDOR> HEADER> HEADER> TEXT> ChemBio **Weapons** and WMD Terrorism News -May 30 ..  
new report by the CIA estimates **Iraq** continued to **produce** mobile biological **weapons**  
laboratories  
C:\NIST eval\data\cnstxt\d1394\_cbw053003.txt - [Open folder](#) - [1 cached](#) - [Oct 26](#)
- [Msl 1652 m9807975.txt](#)  
**weapons** and 400 Scud missiles to Yemen and Sudan in the summer of 1991. In 1992, **Iraq**  
shipped nuclear .. s Task Force, added that **Iraq** maintains the capability to **produce weapons**  
C:\NIST eval\data\cnstxt\Msl\_1652\_m9807975.txt - [Open folder](#) - [1 cached](#) - [Oct 26](#)
- [Msl 2699 m9809103.txt](#)  
that provide parts for **weapons** systems. Hamza said that **Iraq** planned to **produce .. weapons**  
program Khidhir Hamza, who defected in 1995, stated that **Iraq** has  
C:\NIST eval\data\cnstxt\Msl\_2699\_m9809103.txt - [Open folder](#) - [1 cached](#) - [Oct 26](#)
- [4884 2117 3374.txt](#)  
Joseph Cirincione. 4] IAEA **Iraq** Action Team, Fact Sheet: **Iraq's Nuclear Weapons .. Proliferation:**  
WMD Country Profiles **Iraq** Profile /Iraq/ Search **Iraq**  
C:\NIST eval\data\cnstxt\4884\_2117\_3374.txt - [Open folder](#) - [1 cached](#) - [Oct 26](#)
- [5477 3490.txt](#)  
have assisted **Iraq** in establishing a .. plasmid used to **produce** interferon.along with .. genetic  
engineering in order to **produce** vaccines against diseases such as dengue  
C:\NIST eval\data\cnstxt\5477\_3490.txt - [Open folder](#) - [1 cached](#) - [Oct 26](#)
- [5478 3490 3491.txt](#)  
have assisted **Iraq** in establishing a .. plasmid used to **produce** interferon.along with .. genetic

• Keyword matches only  
• Patch-work passages



```

<DOC>
<DOCNO>4895_index.txt</DOCNO>
<DOR>20040415</DOR>
<TITLE>Not Available</TITLE>
<AUTHOR>Not Available</AUTHOR>
<PUBNAME>Unknown</PUBNAME>
<ORIGDATE>00000000</ORIGDATE>
<DOCCLASS>U</DOCCLASS>
<VENDOR>pnnl_web</VENDOR>
<HEADER>
</HEADER>
<TEXT>

```

Center for Nonproliferation Studies

Eye on Proliferation: WMD Country Profiles

Iraq Profile /Iraq/" Search Iraq:

Facilities Chronology Import/Export

Nuclear Overview

- 5+ page documents
- No information on relevant sections in document

Iraq came very close to becoming the first Arab country to produce nuclear weapons. Before its invasion of Kuwait, Iraq, a signatory to the Nuclear Non-Proliferation Treaty, had not yet developed a nuclear weapon. Had Iraq not invaded Kuwait in 1991, it may have been able to build a weapon within a few years. With a bomb design progressing rapidly, Iraq lacked only the fissile material to complete the weapon. The Israeli "Threat"

Iraq maintained a very small civilian nuclear program that began during the Atoms for Peace program in the mid-1950s. In 1962, it acquired a 2 megawatt (MW) research reactor. In 1974, an Iraqi delegation traveled to Paris to negotiate the purchase of a reactor similar to the French Osiris reactor. Then Vice-President Saddam Hussein would have signed the deal. While Iraq attempted to maintain the facade that its purchase was for peaceful nuclear research, some of its other actions belied stated intentions. Prior to his trip to Paris, Iraq had already contracted the Italian firm SNIA-Techint to build a pilot plutonium separation plant. With the reactor deal complete, Iraq set out to build a radiochemical laboratory; it contracted the Italian firm SNIA-Techint in 1979 to build a pilot plutonium separation plant. This period also saw the return to Iran of Ayatollah Ruhollah Khomeini, the Shiite cleric, from his exile in Iraq. Shortly thereafter, Iran became an Islamic republic. The 10-year war with Iran played a role in Saddam's thinking vis--vis Iraq's nuclear weapon program. Saddam wanted the "ultimate equalizer" if faced with an overwhelming military power. Iraqi scientists calculated that the Osiraq reactor could produce between 5 and 25 kilograms (kg) of plutonium per year (the upper limit, however, would require the reactor to operate at full capacity). Israel Alters Iraq's Strategy

Iraq proceeded with building its nuclear infrastructure and plans to subvert IAEA inspections. However, before any significant progress could be made, the Osiraq reactor was destroyed by an Israeli air strike in 1980. Since the Osiraq reactor was destroyed before it could produce any plutonium, Iraq had to seek another source of fissile material. Uranium enrichment became the preferred method. As a secondary method, Iraq explored uranium enrichment through the gaseous diffusion process. The intention was to produce low enriched uranium as feedstock for electricity generation.

- Type System...
- Display Types as Tree (or List)
- type-name...
- MusicalInstrument
- Organization
- Person
- PhoneNumber
- Plant
- ProductionProcedure
- Quantity
- Religion
- Sport
- SportsTeam
- Timex
- Tool
- TravelReservation
- URL
- Vehicle
- Weapon
- WmdAgent
- TopRelation
  - AgentArtifact
  - ArgumentOfEvent
  - Casualties
  - Causes
  - DeliversWeapon
  - DestructionPotential
  - Discourse
  - GpeAffiliation
  - HasIntention
  - HasProperty
  - HasQuality
  - HasReservation
  - HasWeaponCapability
  - HoldsDuring
  - MadePhoneCallTo
  - Motivation
  - OrganizationalRelation
  - PersonalAndSocial
  - PersonalOrOrganization
  - Physical
  - ReservationFor
  - Residence
  - RouteEnd
  - TemporalLink

New Query 1.ssq

**Descriptive comment**

Entity and Fact Semantic, Natural-Language, and Keyword

What weapons does Iraq produce?

Interpret Query

Recall Precision

- ProducesWeapon
  - Weapon
    - GPE | +Iraq

<ProducesWeapon> <Weapon> </Weapon> +<GPE>Iraq</GPE> </ProducesWeapon>

Show JURU from SSQE

Send this JURU to Piquant

• Precision query exploits relations  
 • 49 results (rather than 500+) using highly relevant relation

Annotation View Hit List

49 documents | 88 passages

- [US BELIEVES IRAQ HAS BUILT MAGNETIC ISOTOPE ENRICHMENT PLANT](#)  
 The Iraqi defector claims that Iraq produced 40kg of weapons-grade uranium from its calutrons, but US officials believe that Iraq only produced 1 to 5kg.
- [BUSH ORDERED AID TO IRAQI MILITARY](#)  
 Cyber 910B-400 workstation - to Huteen General Establishment, which was listed by the CIA as a bad end-user and was shown by the IAEA to be a main nuclear weapons design plant in Iraq.
- [POLITICIANS IN THE LAB...AND SCUTTLING AN EASY WAY TO STOP NUCLEAR PROLIFER](#)  
 An Iraqi defector told US officials that Iraq produced almost 90 pounds of weapons-grade uranium in a secret uranium enrichment plant.
- [Does Iraq Have The Bomb?](#)  
 The IAEA identifies these and other tools as "relevant to the production of gas ultra centrifuges, which Iraq intended to use to enrich uranium to weapons grade." Al-Rabiya is an Iraqi weapons plant, which o
- [UN Inspectors Head For Baghdad](#)  
 Abstract: On 27 September 1993, a UN weapons inspection team of more than 50 inspectors, the largest such team thus far, left for Iraq to conduct inspections of Iraqi weapons facilities and to provide infor
- [UN AIDES SAY IRAQ MAY BE CONCEALING NUCLEAR MATERIAL](#)  
 Paul L. Leventhal, president of the Nuclear Control Institute, said that US Administration officials had told him that he Iraqi defector claimed that Iraq produced 40 kilograms of weapons-grade uranium using m
- [Germany Says It Seeks To Curb Its Arms Exporters](#)  
 It is also alleged that German companies built poison gas plants in Iraq as well as sold sophisticated equipment such as precision timers to the Iraqi military.
- [Not Available](#)  
 The Iraqis had over two decades to develop these weapons.
- [Not Available](#)  
 Working with an Iraqi team, the U.N. specialists will take four or five days to eliminate the 155mm munitions at the al-Muthanna State Establishment, 40 miles northwest of Baghdad. Iraq's most important che

SAW-II

File Edit Window Help

Type System... New Query 1.ssq

Display Types as Tree (or List)

ype-name...

- GpeAdjective
- Language
- Machine
- Measurement
- MedicalCondition
- Medicine
- MentalState
- MiscellaneousEntity
- MusicalInstrument
- Organization
- Person
- PhoneNumber
- Plant
- ProductionProcedure
- Quantity
- Religion
- Sport
- SportsTeam
- Timex
- Tool
- TravelReservation
- URL
- Vehicle
- Weapon
- WmdAgent
- TopRelation
- AgentArtifact
- ArgumentOfEvent
- Casualties
- Causes
- DeliversWeapon
- DestructionPotential
- Discourse
- GpeAffiliation
- HasIntention
- HasProperty
- HasQuality
- HasReservation
- HasWeaponCapability
- HoldsDuring
- MadePhoneCallTo
- Motivation
- OrganizationalRelation
- PersonalAndSocial
- PersonalOrOrganization
- Physical
- ReservationFor
- Residence

Descriptive comment

Entity and Fact Semantic, Natural-Language, and Keyword

What weapons does Iraq produce?

Interpret Query

Recall Precision

- ProducesWeapon |
  - Weapon |
  - GPE | +Iraq

<ProducesWeapon> +<Weapon></Weapon> +<GPE>Iraq</GPE> </ProducesWeapon>

Show JURU from SSQE

Send this JURU to Piquant

- Annotated document view
- Reveals task-relevant concepts and relations
- Tracks source of annotation

Annotation View / Nuc\_13094\_n9105855.txt Hit List

49 documents | 88 passages

US BELIEVES IRAQ HAS BUILT  
MAGNETIC ISOTOPE ENRICHMENT PLANT.

Abstract: In June 1990, US experts briefed an Iraqi nuclear official who defected from Iraq. The US experts now believe that Iraq has at least one secret calutron uranium enrichment facility. US satellite photos taken several months ago may reveal calutron equipment. Since 1988, Iraq has been importing gas centrifuge uranium enrichment equipment from several western countries. In 1990, the US Department of Commerce blocked the sale to Iraq of vacuum diffusion pumps from CVC, a US company in NY, after US Customs determined that the pumps could be used in Iraq's centrifuge program. US officials now believe the pumps were intended for Iraq's calutron program. In the late 1980's, another US firm exported to Iraq \$1.2 million worth of 5-ampere, 40 to 45 kilovolt electrical generators, claimed to be for induction welding. US officials now believe that Iraq imported the generators for use as calutron power supplies. The generating equipment does not fall under dual-use technology US export control laws. Export and import of centrifuge-related dual use technology is strictly controlled by US export laws, while calutron-related equipment is not controlled. In 1949, the US declassified calutron equipment. Calutron equipment is not on the Zangger "trigger list," or the Nuclear Suppliers Group commodity control list. The Iraqi defector claims that Iraq produced 40kg of weapons-grade uranium from its calutrons, but US officials believe that Iraq only produced 1 to 5kg. In late 1990, press reports in the US claimed that Saddam Hussein had constructed centrifuge plants, uranium hexafluoride conversion facilities, a uranium mine, and an underground plutonium production reactor, all located in northern Iraq. US officials have not confirmed these press reports. A calutron plant may be located near a hydroelectric power station in

Legend

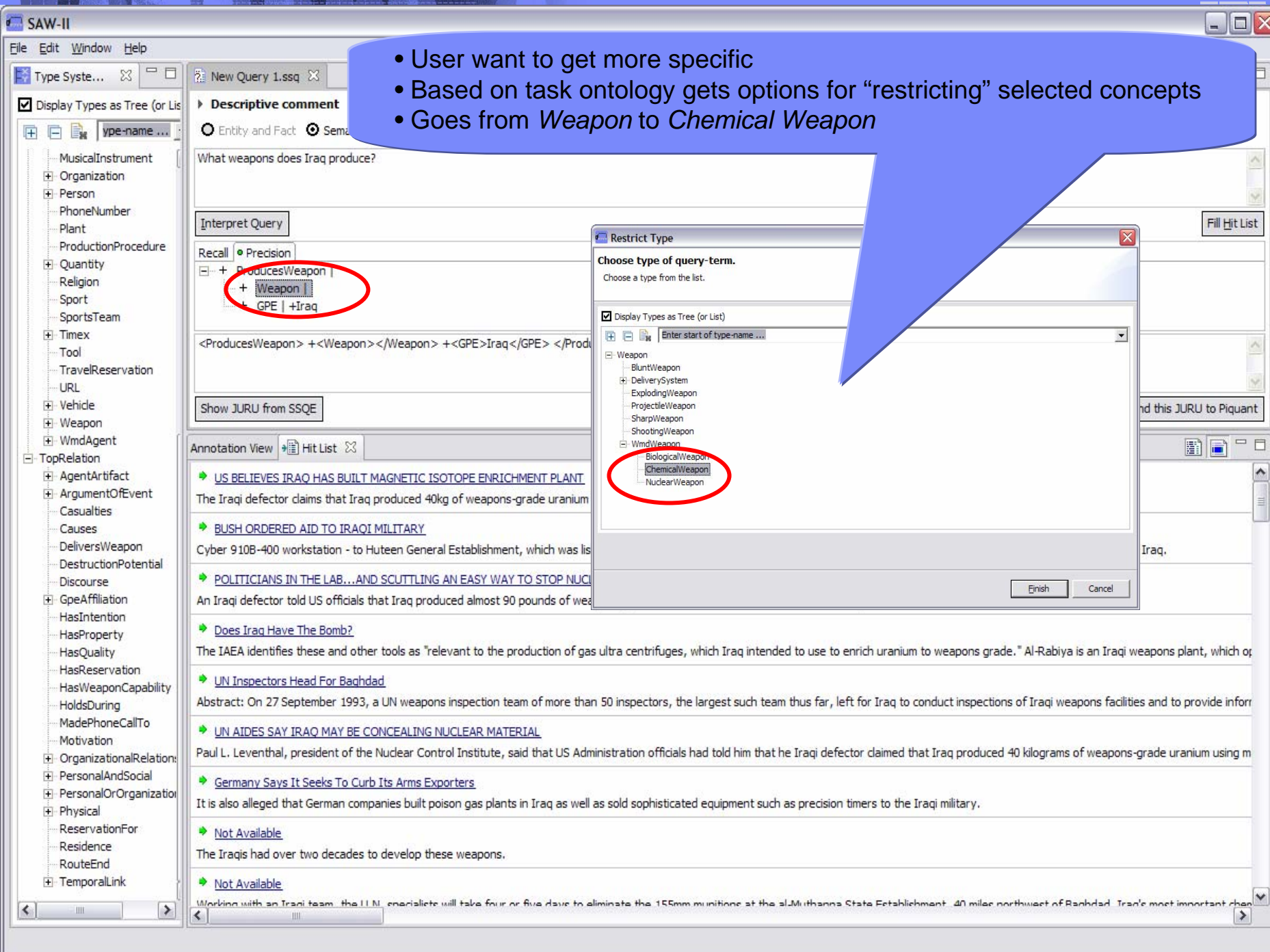
<input type="checkbox"/> HoldsDuring	<input type="checkbox"/> At	<input type="checkbox"/> ManagerOf	<input type="checkbox"/> BasedIn	<input type="checkbox"/> CitizenOf
<input type="checkbox"/> GeneralStaff	<input type="checkbox"/> NatoNation	<input type="checkbox"/> AxisOfEvilNati...	<input type="checkbox"/> MuslimNation	<input type="checkbox"/> Capital
<input type="checkbox"/> City	<input type="checkbox"/> UsState	<input type="checkbox"/> Nation	<input checked="" type="checkbox"/> GPE	<input type="checkbox"/> Location

Select All Deselect All Hide Unselected

Click In Text to See Annotation Detail

Annotations

- ProducesWeapon
  - ProducesWeapon ("Iraq produced 40kg of weapons-grade uranium")
    - begin = 1876
    - end = 1919
    - confidence = 0.0
    - componentId = KSRelationDetector
    - relationArgs = BinaryRelationArgs
      - domainValue = GPE ("Iraq")
      - rangeValue = NuclearWeaponAgent ("weapons-grade uranium")



- User want to get more specific
- Based on task ontology gets options for “restricting” selected concepts
- Goes from *Weapon* to *Chemical Weapon*

SAW-II

File Edit Window Help

Type System

Display Types as Tree (or List)

type-name ...

- MusicalInstrument
- Organization
- Person
- PhoneNumber
- Plant
- ProductionProcedure
- Quantity
- Religion
- Sport
- SportsTeam
- Timex
- Tool
- TravelReservation
- URL
- Vehicle
- Weapon
- WmdAgent
- TopRelation
- AgentArtifact
- ArgumentOfEvent
- Casualties
- Causes
- DeliversWeapon
- DestructionPotential
- Discourse
- GpeAffiliation
- HasIntention
- HasProperty
- HasQuality
- HasReservation
- HasWeaponCapability
- HoldsDuring
- MadePhoneCallTo
- Motivation
- OrganizationalRelation
- PersonalAndSocial
- PersonalOrOrganization
- Physical
- ReservationFor
- Residence
- RouteEnd
- TemporalLink

New Query 1.ssq

Descriptive comment

Entity and Fact Semantic

What weapons does Iraq produce?

Interpret Query

Recall Precision

- ProducesWeapon |
- Weapon**
- GPE | +Iraq

<ProducesWeapon> +<Weapon>/Weapon> +<GPE>Iraq</GPE> </ProducesWeapon>

Show JURU from SSQE

Restrict Type

Choose type of query-term.

Choose a type from the list.

Display Types as Tree (or List)

Enter start of type-name ...

- Weapon
  - BluntWeapon
  - DeliverySystem
  - ExplodingWeapon
  - ProjectileWeapon
  - SharpWeapon
  - ShootingWeapon
  - WmdWeapon
  - BiologicalWeapon**
  - ChemicalWeapon**
  - NuclearWeapon

Finish Cancel

Annotation View Hit List

- [US BELIEVES IRAQ HAS BUILT MAGNETIC ISOTOPE ENRICHMENT PLANT](#)  
The Iraqi defector claims that Iraq produced 40kg of weapons-grade uranium
- [BUSH ORDERED AID TO IRAQI MILITARY](#)  
Cyber 910B-400 workstation - to Huteen General Establishment, which was li
- [POLITICIANS IN THE LAB...AND SCUTTLING AN EASY WAY TO STOP NUC](#)  
An Iraqi defector told US officials that Iraq produced almost 90 pounds of wea
- [Does Iraq Have The Bomb?](#)  
The IAEA identifies these and other tools as "relevant to the production of gas ultra centrifuges, which Iraq intended to use to enrich uranium to weapons grade." Al-Rabiya is an Iraqi weapons plant, which op
- [UN Inspectors Head For Baghdad](#)  
Abstract: On 27 September 1993, a UN weapons inspection team of more than 50 inspectors, the largest such team thus far, left for Iraq to conduct inspections of Iraqi weapons facilities and to provide inform
- [UN AIDES SAY IRAQ MAY BE CONCEALING NUCLEAR MATERIAL](#)  
Paul L. Leventhal, president of the Nuclear Control Institute, said that US Administration officials had told him that he Iraqi defector claimed that Iraq produced 40 kilograms of weapons-grade uranium using m
- [Germany Says It Seeks To Curb Its Arms Exporters](#)  
It is also alleged that German companies built poison gas plants in Iraq as well as sold sophisticated equipment such as precision timers to the Iraqi military.
- [Not Available](#)  
The Iraqis had over two decades to develop these weapons.
- [Not Available](#)  
Working with an Iraqi team, the UN specialists will take four or five days to eliminate the 155mm munitions at the al-Muthanna State Establishment, 40 miles northwest of Baghdad, Iraq's most important chan

- Display Types as Tree (or Lis
- type-name ...
- GpeAdjective
- Language
- Machine
- Measurement
- MedicalCondition
- Medicine
- MentalState
- MiscellaneousEntity
- MusicalInstrument
- Organization
- Person
- PhoneNumber
- Plant
- ProductionProcedure
- Quantity
- Religion
- Sport
- SportsTeam
- Timex
- Tool
- TravelReservation
- URL
- Vehicle
- Weapon
- WmdAgent
- TopRelation
  - AgentArtifact
  - ArgumentOfEvent
  - Casualties
  - Causes
  - DeliversWeapon
  - DestructionPotential
  - Discourse
  - GpeAffiliation
  - HasIntention
  - HasProperty
  - HasQuality
  - HasReservation
  - HasWeaponCapability
  - HoldsDuring
  - MadePhoneCallTo
  - Motivation
  - OrganizationalRelation
  - PersonalAndSocial
  - PersonalOrOrganization
  - Physical
  - ReservationFor
  - Residence

New Query 1.ssq [x] [ ]

**Descriptive comment**

Entity and Fact  Semantic, Natural-Language, and Keyword

What weapons does Iraq produce?

Interpret Query [Fill Hit List]

Recall  Precision

- + ProducesWeapon |
- + ChemicalWeapon |
- + GPE | +Iraq

+<ProducesWeapon>+<ChemicalWeapon></ChemicalWeapon>+<GPE>Iraq </GPE></ProducesWeapon>

Show JURU from SSQE [Send this JURU to Piquant]

Germany Says It Seeks To Curb Its Arms Exports

It is also alleged that German companies built poison gas plants in Iraq as well as sold sophisticated equipment such as precision timers to the Iraqi military.

• Zoomed in on single highly relevant document

- Type System...
- Display Types as Tree (or Lis
- type-name ...
- GpeAdjective
- Language
- Machine
- Measurement
- MedicalCondition
- Medicine
- MentalState
- MiscellaneousEntity
- MusicalInstrument
- Organization
- Person
- PhoneNumber
- Plant
- ProductionProcedure
- Quantity
- Religion
- Sport
- SportsTeam
- Timex
- Tool
- TravelReservation
- URL
- Vehicle
- Weapon
- WmdAgent
- TopRelation
- AgentArtifact
- ArgumentOfEvent
- Casualties
- Causes
- DeliversWeapon
- DestructionPotential
- Discourse
- GpeAffiliation
- HasIntention
- HasProperty
- HasQuality
- HasReservation
- HasWeaponCapability
- HoldsDuring
- MadePhoneCallTo
- Motivation
- OrganizationalRelation
- PersonalAndSocial
- PersonalOrOrganization
- Physical
- ReservationFor
- Residence

New Query 1.ssq

**Descriptive comment**

Entity and Fact  Semantic, Natural-Language, and Keyword

What weapons does Iraq produce?

Interpret Query Fill Hit List

Recall  Precision

- + ProducesWeapon |
  - + ChemicalWeapon |
    - + GPE | +Iraq

+<ProducesWeapon>+<ChemicalWeapon></ChemicalWeapon>+<GPE>Iraq </GPE></ProducesWeapon>

Show JURU from SSQE Send this JURU to Piquant

Annotation View / Msl\_9146\_m9100142.txt Hit List

Germany Says It Seeks To Curb Its Arms Exporters.

Abstract: Germany has decided to seek new sanctions against illegal arms exporters and plans to create in the Office of Trade and Industry an export control office. Foreign Minister Hans-Dietrich Genscher called exporters "merchants of death," while Economics Minister J?ergen Nolleman says he is determined to close "the last loopholes" in German export laws. As the fifth largest arms exporter, Germany earns about 5% of its export earning from the sale of weapons abroad. Israel in fact accused Germany of "almost criminal laxity" in its export policies after being struck with Iraqi Scud missiles which, it is alleged, had been improved and their range extended by German companies. **It is also alleged that German companies built poison gas plants in Iraq as well as sold sophisticated equipment such as precision timers to the Iraqi military.** Under the proposed sanctions, violators of UN sanctions on trade with Iraq will be imprisoned for a minimum of one year, an increase from six months, with 10-year maximum sentences. At least nine German companies are under investigation on suspicion of violating the export laws.

**Legend**

<input type="checkbox"/> HoldsDuring	<input type="checkbox"/> At	<input type="checkbox"/> BasedIn	<input type="checkbox"/> NatoNation	<input type="checkbox"/> EuNation
<input type="checkbox"/> AxisOfEvilNati...	<input type="checkbox"/> MuslimNation	<input type="checkbox"/> Language	<input type="checkbox"/> Nation	<input checked="" type="checkbox"/> GPE
<input type="checkbox"/> Location	<input type="checkbox"/> GeographicEnt...	<input type="checkbox"/> Facility	<input type="checkbox"/> MilitaryOrgani...	<input type="checkbox"/> Organization

Select All Deselect All Hide Unselected

Click In Text to See Annotation Detail

Annotations

- ChemicalWeapon
- ProducesWeapon
- ProducesWeapon ("poison gas plants in Iraq")
  - begin = 1186
  - end = 1211
  - confidence = 0.0
  - componentId = KSRelationDetector
  - relationArgs = BinaryRelationArgs
    - domainValue = GPE ("Iraq")
    - rangeValue = ChemicalWeapon ("poison gas")



Google Desktop BETA  
iraq chemical weapons produce  
Search Desktop Preferences | Server Options  
Advanced Search

Host: kogs8.watson.ibm.com  
Desktop: All - 0 emails - 205 files - 0 web history - 0 chats - 0 other

1-10 of about 205 (0.03s)  
Remove from Index | Sorted by relevance | Sort by date

- [d1340\\_cbw011703.txt](#)  
4300950.htm 2) WMD -IRAQ http://www.nytimes.com .. the two. WMD -IRAQ 2A) Inspectors Find ..  
11 empty **chemical** warheads today at an ammunition storage depot in southern **Iraq**  
C:\NIST eval\data\cnstxt\d1340\_cbw011703.txt - [Open folder](#) - 1 cached - Oct 26
- [d942\\_cbw041400.txt](#)  
CW SITE 5) IRAQ MONITORING 6) WMD DEFENSE .. pounds of deadly **chemical** weap  
stored at .. fulfill its commitments to **produce** vaccine 'without further relief The  
C:\NIST eval\data\cnstxt\d942\_cbw041400.txt - [Open folder](#) - 1 cached - Oct 26
- [5078\\_3367\\_3393.txt](#)  
with an increasing fear that **Iraq** might resort to using **chemical weapons** .. mass-producing  
**chemical weapons**, but Iran .. capability to **produce chemical weapons**  
C:\NIST eval\data\cnstxt\5078\_3367\_3393.txt - [Open folder](#) - 1 cached - Oct 26
- [5400\\_3370.txt](#)  
relationships to help **Iraq** with its **chemical** .. prominently in **chemical weapons** discussions ..  
intentions to **produce chemical weapons**  
C:\NIST eval\data\cnstxt\5400\_3370.txt - [Open folder](#) - 1 cached - Oct 26
- [591\\_cia0200.txt](#)  
including nuclear **weapons**, **chemical weapons** .. attempting to develop an indigenous capability  
to **produce** various types of **weapons** nuclear, **chemical**, and biological  
C:\NIST eval\data\cnstxt\591\_cia0200.txt - [Open folder](#) - 1 cached - Oct 26
- [d1195\\_cbw012502.txt](#)  
**chemical** and biological **weapons**. The sanctions .. which could be used for **chemical** and  
biological **weapons**. Iran remains on .. nuclear **weapons**. Naming **Iraq** and North  
C:\NIST eval\data\cnstxt\d1195\_cbw012502.txt - [Open folder](#) - 1 cached - Oct 26
- [d1398\\_cbw060903.txt](#)  
**chemical weapons**. But a Defense Intelligence Agency report on **chemical weapons** ..  
**chemical weapons** or whether **Iraq** has or will establish its **chemical** agent  
C:\NIST eval\data\cnstxt\d1398\_cbw060903.txt - [Open folder](#) - 1 cached - Oct 26
- [4895\\_index.txt](#)

• Attempts "restriction" with Google  
• No reduction/precision gain  
• Finds documents with "Chemical Weapon"

## Conclusions

- **Semantic Search enables high precision IR**
  - Identified three XML Fragments operators
  - Applied operators to address four query time semantic needs
  - Shown significant improvement in precision in most applications
- **Semantic Search not intended as replacement for keyword search**
  - Keyword search is successful in web and desktop search
  - Semantic search intended to improve performance for
    - Enterprise search
    - Information gathering for exploratory research