Open-domain Cross-lingual Question Answering from Unstructured Documents

Günter Neumann

LT-lab, DFKI, Saarbrücken,

June, 2005

Motivation: From Search Engines to Answer Engines

Motivation: From Search Engines to Answer Engines

User Query:
KeyWrds, Wh-Clause, Q-Text

Search Engines

User still carries the major efforts in understanding

Motivation: From Search Engines to Answer Engines

User Query:
KeyWrds, Wh-Clause, Q-Text

Experienc
ed-based
QA cycles

Search Engines

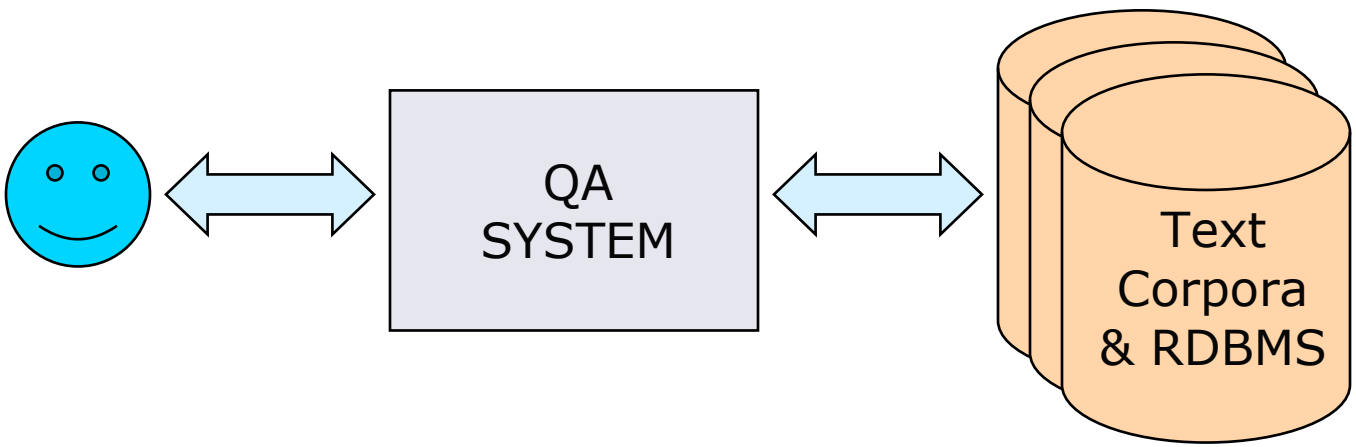
Answer Engines

Shift more „interpretation effort" to machines

User still carries the major efforts in understanding

Question Answering

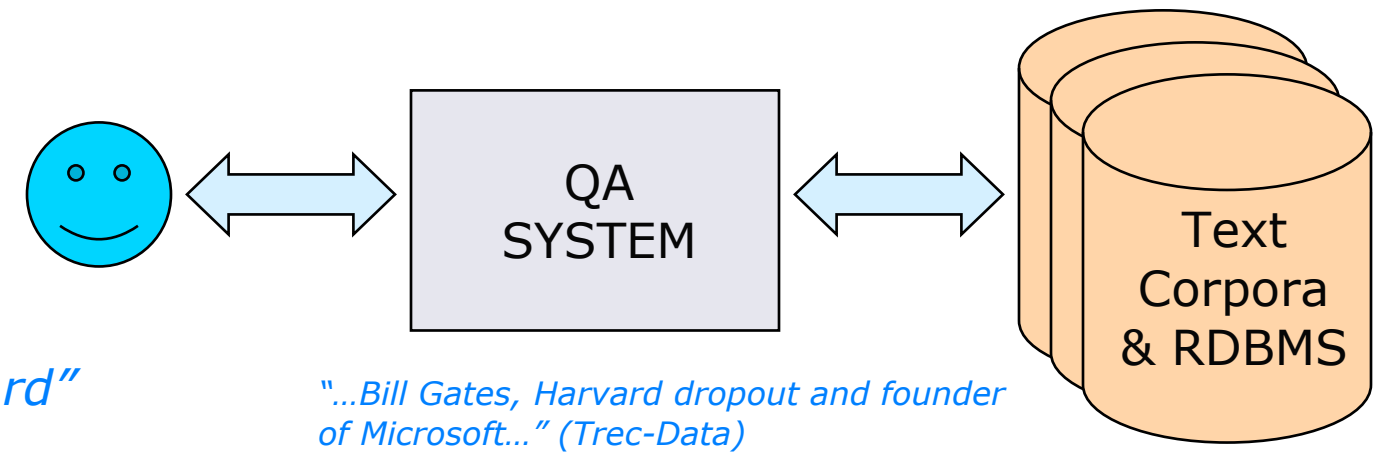
- Input: a question in NL; a set of text and database resources
- Output: a set of possible answers drawn from the resources



Question Answering

- Input: a question in NL; a set of text and database resources
- Output: a set of possible answers drawn from the resources

"Where did Bill Gates go to college?"



"Harvard"

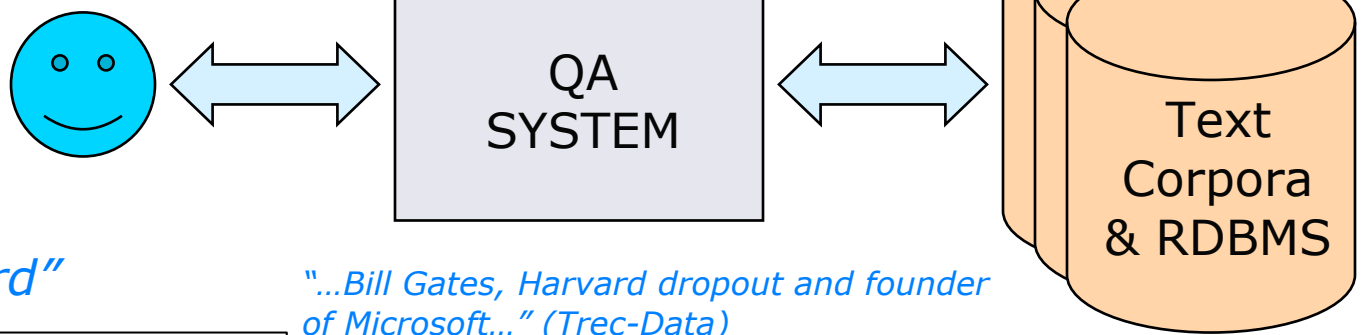
"…Bill Gates, Harvard dropout and founder of Microsoft…" (Trec-Data)

Question Answering

- Input: a question in NL; a set of text and database resources
- Output: a set of possible answers drawn from the resources

"Where did Bill Gates go to college?"

"What is the rainiest place on Earth?"



"Harvard"

"…Bill Gates, Harvard dropout and founder of Microsoft…" (Trec-Data)

"Mount Waialeale"

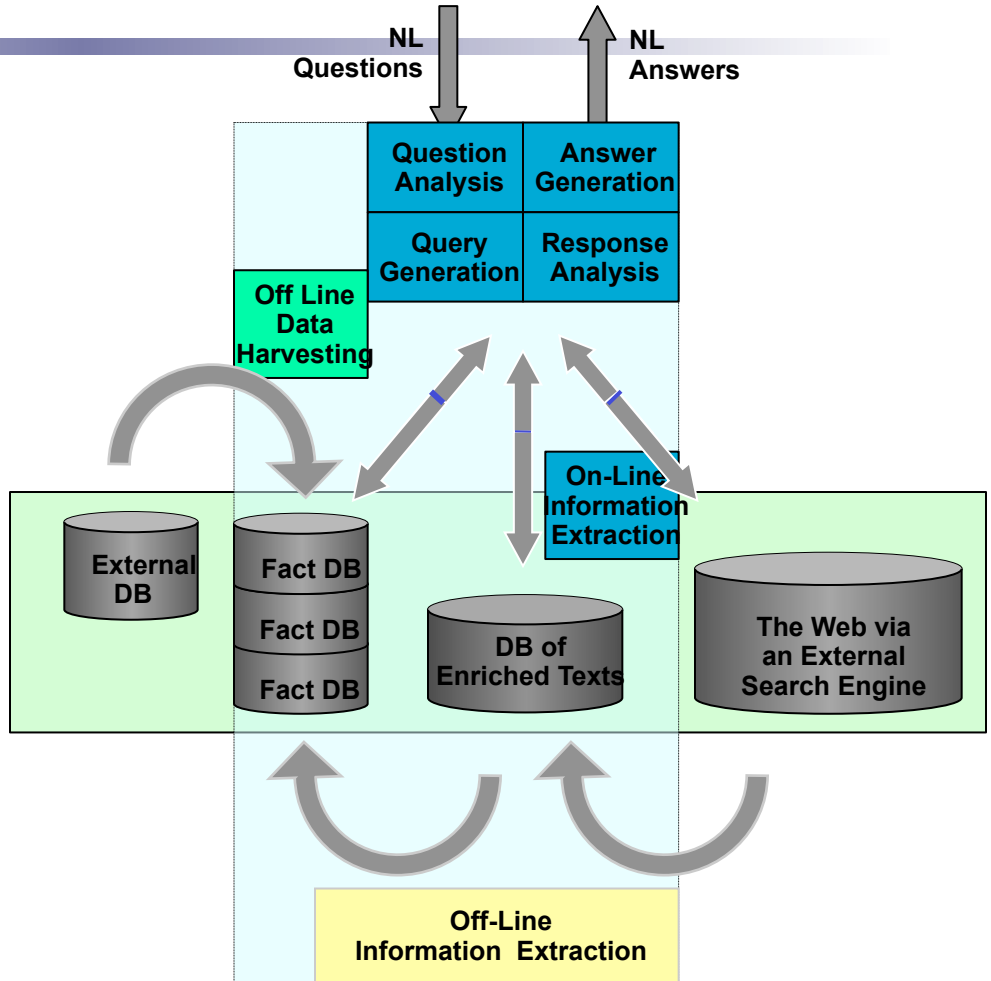
"… In misty Seattle, Wash., last year, 32 inches of rain fell.
Hong Kong gets about 80 inches a year, and even Pago Pago,
noted for its prodigious showers, gets only about 196 inches annually.
(The titleholder, according to the National Geographic Society,
is Mount Waialeale in Hawaii, where about 460 inches of rain falls each year.) …"
(Trec-Data; but see Google-retrieved Web page.)

Hybrid QA Architecture

Hypothesis

real-life QA systems will perform best if they can

- *combine* the virtues of domain-specialized QA with open-domain QA
- *utilize* general knowledge about frequent types and
- *access* semi-structured know-ledge bases

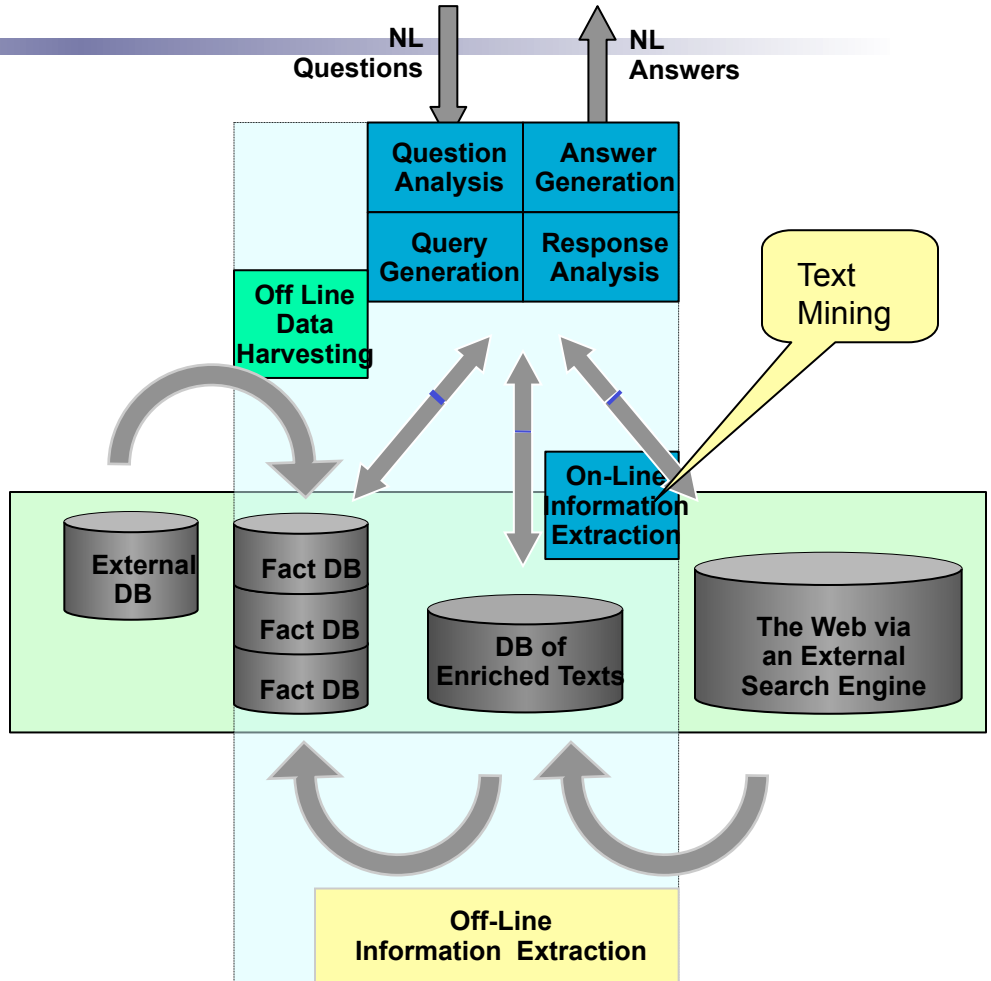


Hybrid QA Architecture

Hypothesis

real-life QA systems will perform best if they can

- *combine* the virtues of domain-specialized QA with open-domain QA
- *utilize* general knowledge about frequent types and
- *access* semi-structured know-ledge bases

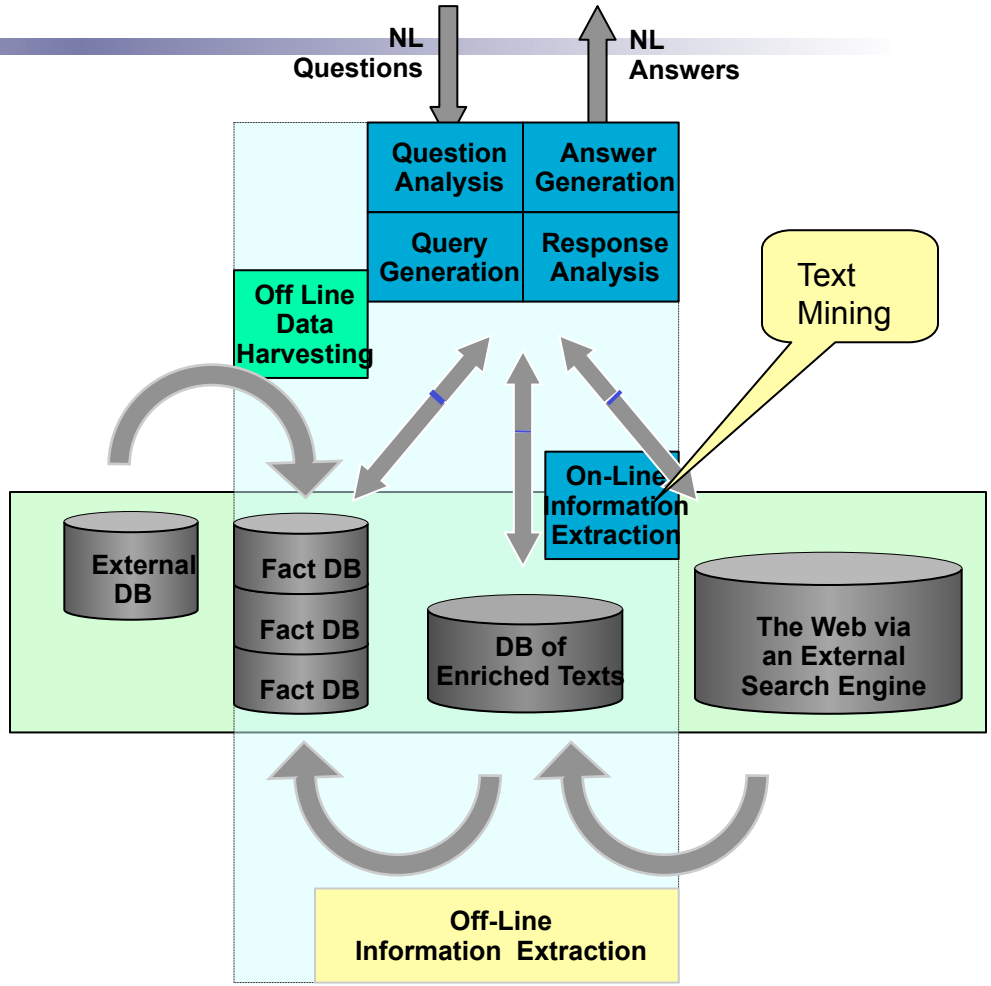


Hybrid QA Architecture

Hypothesis

real-life QA systems will perform best if they can

- *combine* the virtues of domain-specialized QA with open-domain QA
- *utilize* general knowledge about frequent types and
- *access* semi-structured know-ledge bases

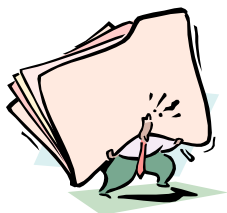


Advertisement:
DFKI project Quetal
2003-2005

Design Issues

- Foster bottom-up system development
  - Data-driven, robustness, scalability
  - From shallow to deep NLP

- Large-scale answer processing
  - Coarse-grained uniform representation of query/documents
  - Text zooming
    - » From paragraphs to sentences to phrases
  - Ranking scheme for answer selection

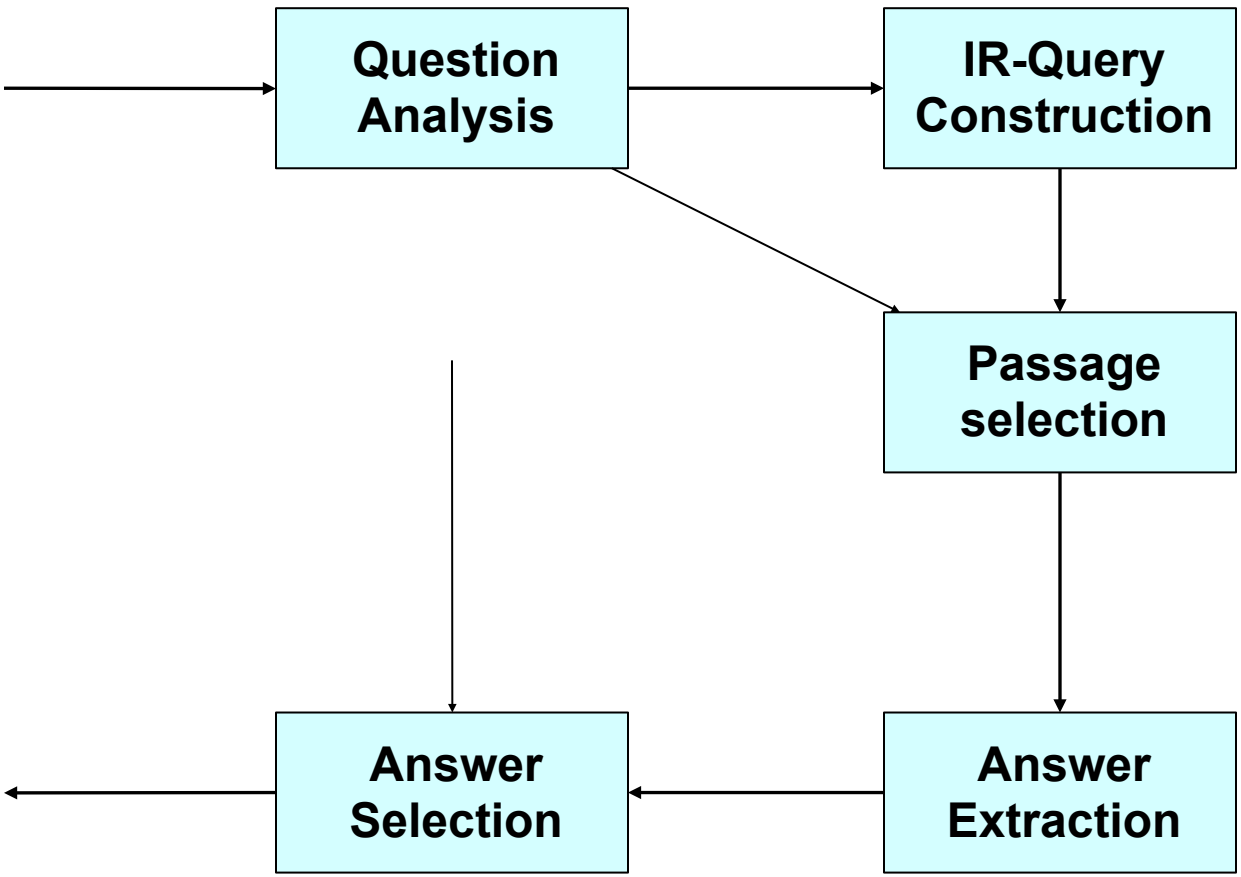


- Common basis for
  - Online Web pages
  - Large textual sources

Open-Domain Question Answering

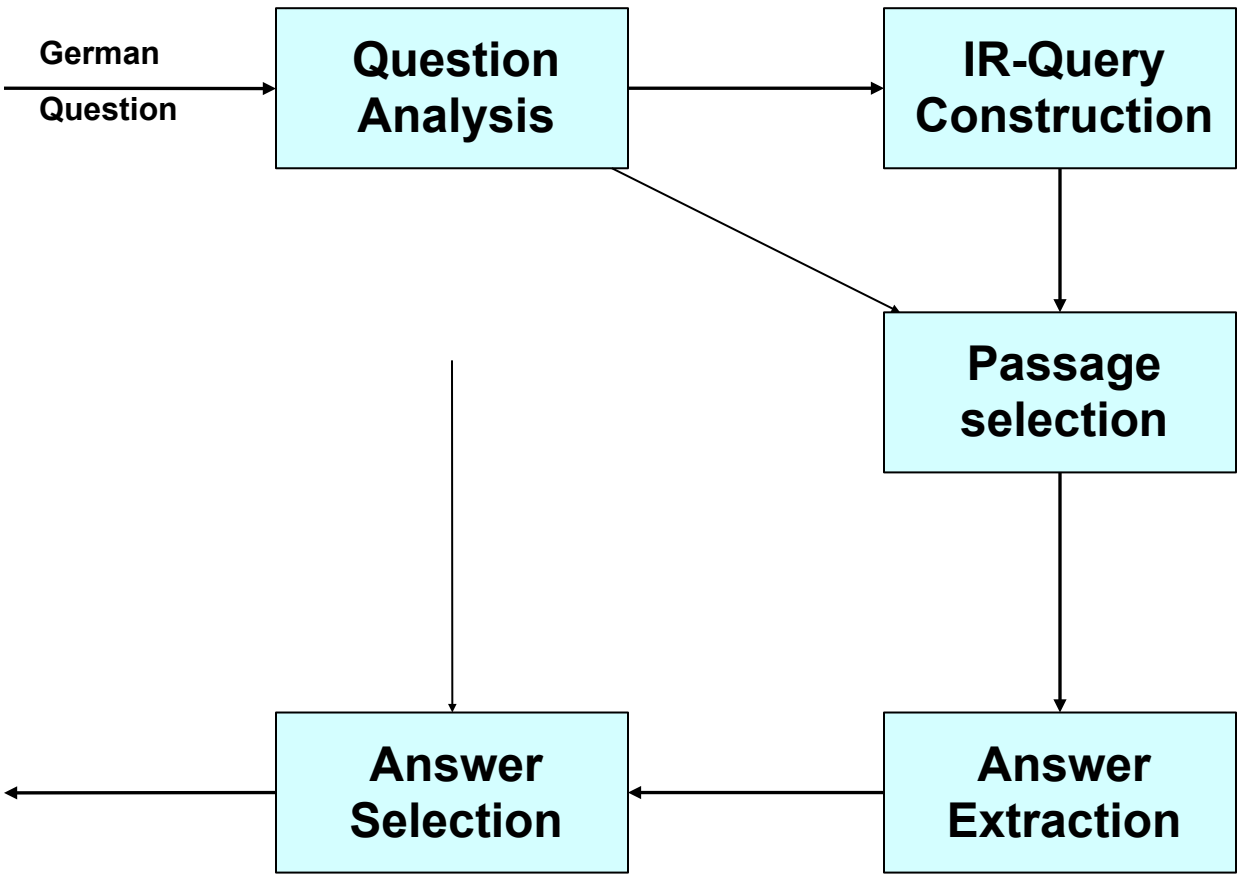
- Open domain
  - No restriction for the domain and type of question
  - No restriction on document source
- Combines
  - Information retrieval
  - Information extraction
  - Text mining
  - Computational Linguistics
- Cross-lingual ODQA
  - Express query in language X
  - Answer from documents in language Y

Open-Domain Question-Answering



Open-Domain Question-Answering

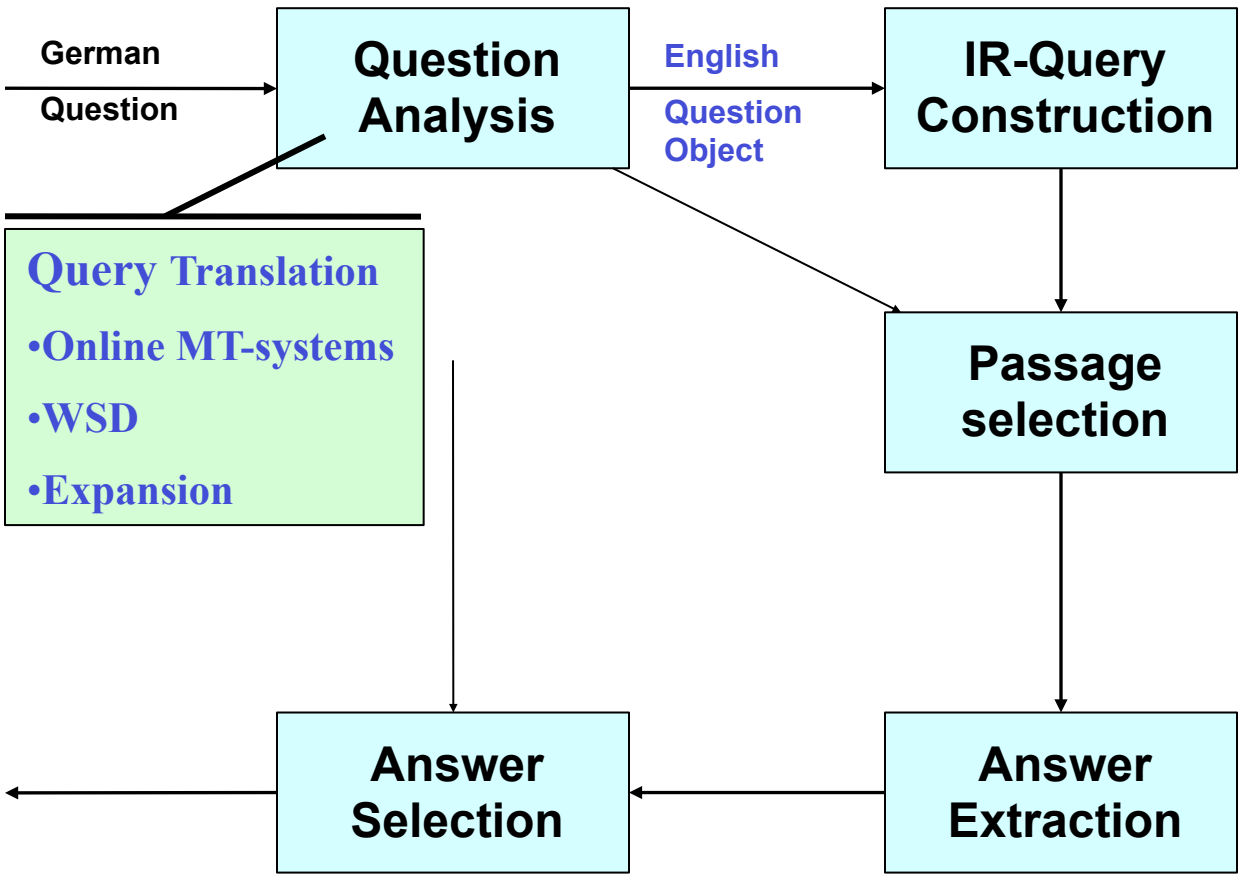
"Mit wem ist David Beckham verheiratet?"



Open-Domain Question-Answering

"Mit wem ist David Beckham verheiratet?"

{person:David Beckham, married, person:?}



Open-Domain Question-Answering

"Mit wem ist David Beckham verheiratet?"

{person:David Beckham, married, person:?}

German Question

Question Analysis

English Question Object

IR-Query Construction

IR-Google

IR-Lucene/XML

Annotated Corpus

Documents

Query Translation

- Online MT-systems
- WSD
- Expansion

Passage selection

Answer Extraction

Answer Selection

Open-Domain Question-Answering

"Mit wem ist David Beckham verheiratet?"

{person:David Beckham, married, person:?}

German Question

Question Analysis

English Question Object

IR-Query Construction

IR-Google

IR-Lucene/XML

Annotated Corpus

Documents

Query Translation

- Online MT-systems
- WSD
- Expansion

Question Object:

- Focus, Scope
- AnswerType

Passage selection

Passages

"David Beckham, the soccer star engaged to marry Posh Spice, is being blamed for England 's World Cup defeat."

Answer Selection

Answer Extraction

Open-Domain Question-Answering

"Mit wem ist David Beckham verheiratet?"

{person:David Beckham, married, person:?}

German Question

Question Analysis

English Question Object

IR-Query Construction

IR-Google

IR-Lucene/XML

Annotated Corpus

Query Translation

- Online MT-systems
- WSD
- Expansion

Question Object:

- Focus, Scope
- AnswerType

Documents

Passage selection

Passages

"David Beckham, the soccer star engaged to marry Posh Spice, is being blamed for England 's World Cup defeat."

Answer Selection

Candidates

Answer Extraction

{person:David Beckham, person:Posh Spice}

Open-Domain Question-Answering

"Mit wem ist David Beckham verheiratet?"

{person:David Beckham, married, person:?}

German Question

Question Analysis

English Question Object

IR-Query Construction

IR-Google

IR-Lucene/XML

Annotated Corpus

Documents

Query Translation

- Online MT-systems
- WSD
- Expansion

Question Object:

- Focus, Scope
- AnswerType

Passage selection

Passages

"David Beckham, the soccer star engaged to marry Posh Spice, is being blamed for England 's World Cup defeat."

Answer

Answer Selection

Candidates

Answer Extraction

Posh Spice

{person:David Beckham, person:Posh Spice}

Open-Domain Question Answering: Some Details

- Multi-dimensional annotation of unstructured text corpora
- Cross-lingual query processing
- Clef-2004 participation

Why multi-dimensional annotation of un-structured text?

- The assumption is that a structural analysis of un-structured texts towards the type of information that can be the focus of questions, will support the retrieval of relevant small textual information units through informative IR-queries.
  - From candidate document retrieval to candidate answer retrieval.

- However, since we cannot foresee all the different user's interests/questions, a challenging research question is:
  - How detailed can the structural analysis be made without putting over a "straitjacket" of a particular view on the un-structured source?

- The assumptions here are:
  - Questions and answers are somewhat related ("questions influence the information geometry and hence, the information view and access", see also Rijsbergen, 2004)
  - There is a bias between off-line and on-line answer extraction.

Some initial experiments

We have performed some experiments focusing on the relationship between the size of information units and answer containment (using the QA-test set from Clef-2003).

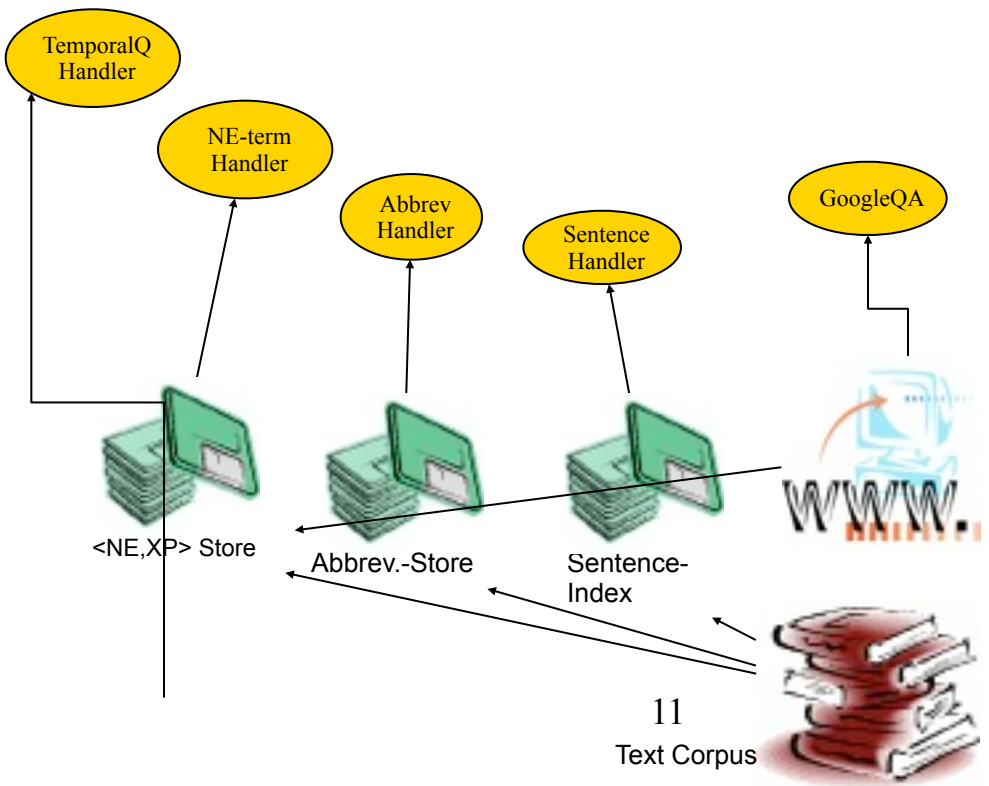
#N / Unit-Type	1	5	10	20	30	40	50	100
*Sentences**	37.9	58.2	65.8	69.6	70.8	72.1	74	75.9
Sentences	28.4	53.1	60.1	67	70.2	72.7	72.7	74.6
*Passages**	39.8	63.2	68.3	73.4	74	75.3	76.5	77.8
Passages	31.6	60.7	67.7	71.5	74.6	77.2	77.2	80.3
*Documents**	47.4	69.6	76.5	80.3	81.6	82.9	82.9	83.5
Documents	46.2	68.3	77.8	82.2	82.2	83.5	84.1	85.4

As a result we hyphothesized that it is reasonable to use NE-annotated sentences as major retrieval units for the IR-engine
⇒
Simplified answer extraction process & no need of special passage extraction methods

Precision of retrieval for different unit types and top N units retrieved, namely documents, passages, sentences – and their NE-annotated correspondents (marked by *).

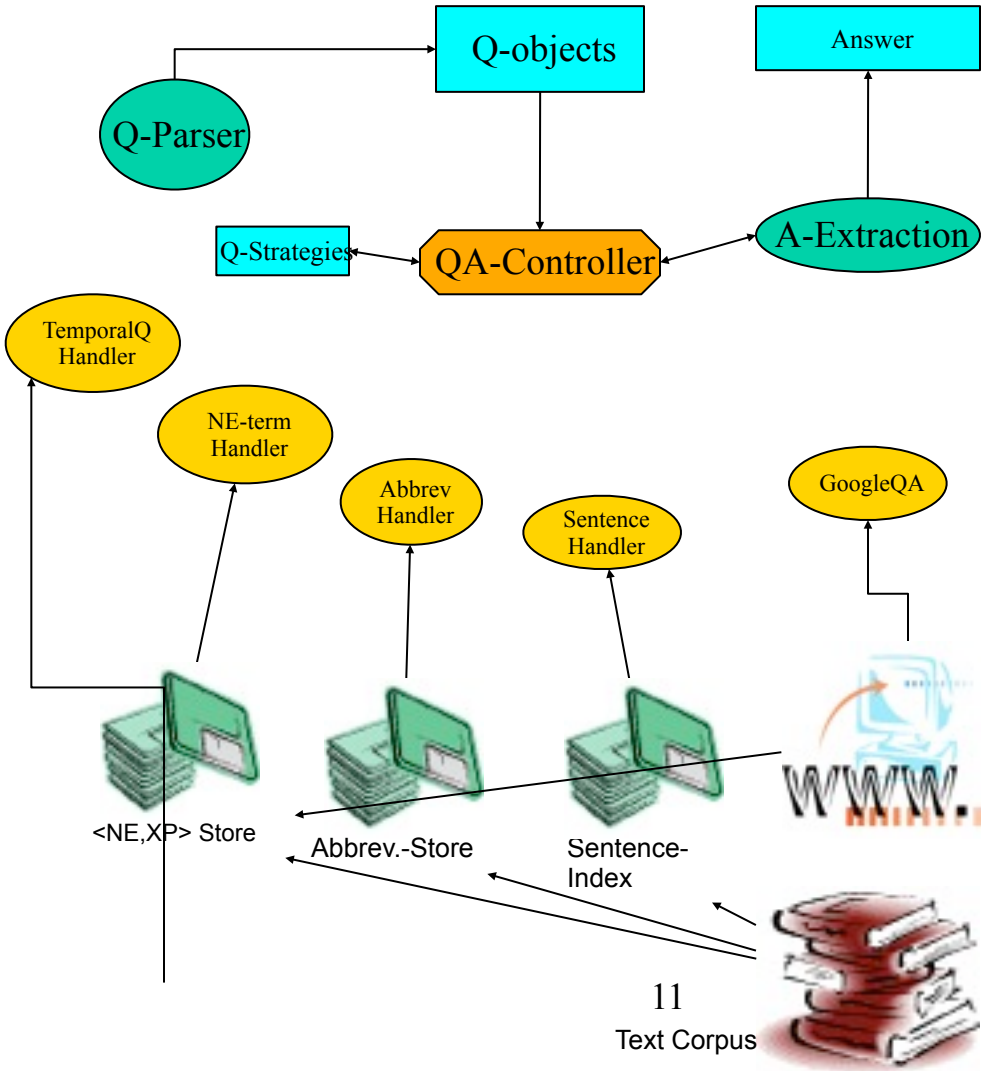
Open-domain Question Answering: Multi-dimensional annotation

- Idea: off-line annotation of the data collection, which support
  - Query-specific indexing (Q-strategies), and
  - Answer extraction
- Sentence-level pre-processing proved valuable
  - Sentences-boundary
  - Named Entity + Co-reference
  - Abbreviations
  - NE-anchored tuples



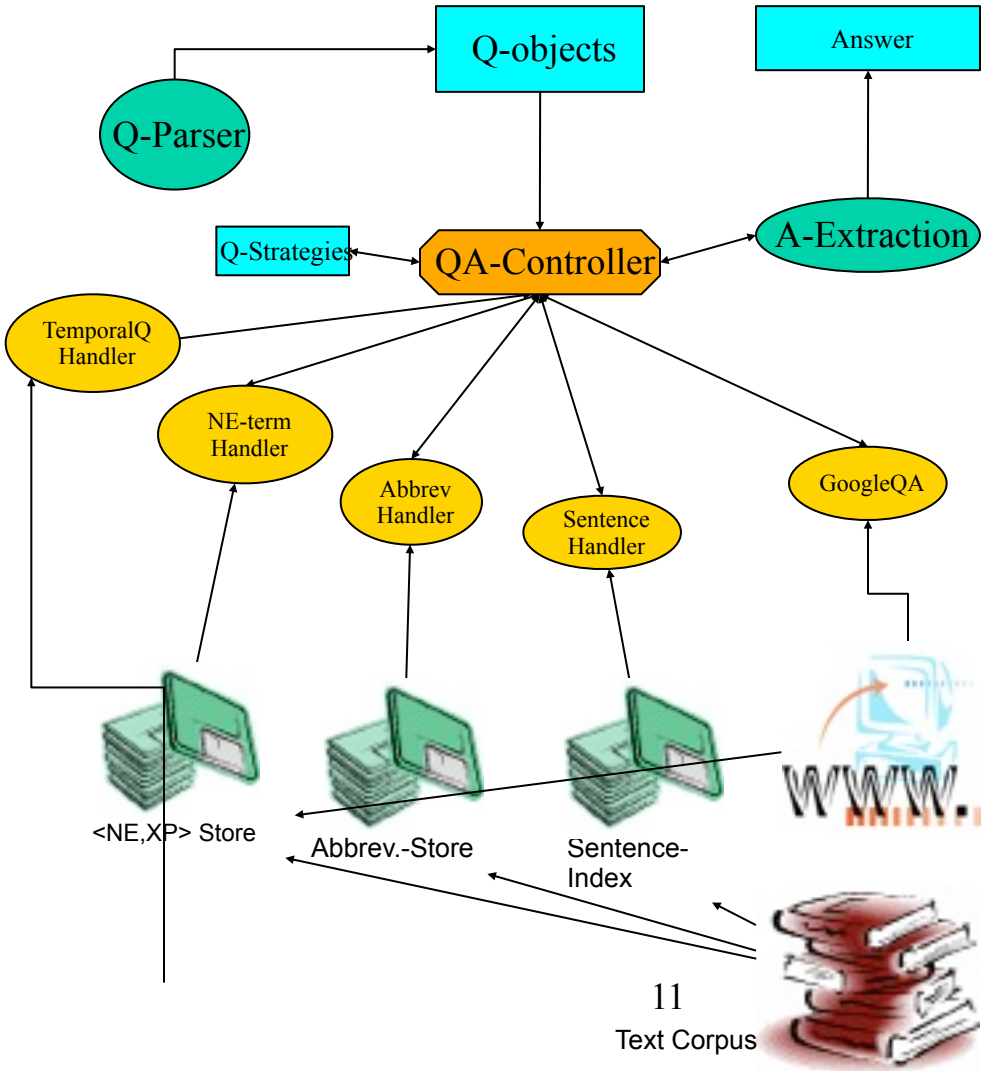
Open-domain Question Answering: Multi-dimensional annotation

- Idea: off-line annotation of the data collection, which support
  - Query-specific indexing (Q-strategies), and
  - Answer extraction
- Sentence-level pre-processing proved valuable
  - Sentences-boundary
  - Named Entity + Co-reference
  - Abbreviations
  - NE-anchored tuples



Open-domain Question Answering: Multi-dimensional annotation

- Idea: off-line annotation of the data collection, which support
  - Query-specific indexing (Q-strategies), and
  - Answer extraction
- Sentence-level pre-processing proved valuable
  - Sentences-boundary
  - Named Entity + Co-reference
  - Abbreviations
  - NE-anchored tuples



Why Query Analysis for Open-Domain?

- In principle possible:
  - No query analysis necessary
  - Send to IR (e.g, Lucene or Google) NL-Query as it is
- However
  - Would not allow to control IR through specific search queries
  - Answer extraction & selection not possible

Query Analysis

- Hence, we want to compute an internal query object, to support
  - Construction of “meaningful” IR-search queries, e.g.,
    - » “What is a battery?” -> “define:battery”
    - » “Wer hat das Schloss Charlottenburg erbaut?” -> “Schloss Charlottenburg”+ erbauen erbaute erbaut
    - » “Name a scientist who won the Nobel Price in physics.” -> neType:PERSON text:scientist text:won text:”Nobel Price”+ text:physics+
  - Control/guide answer selection
    - » Type of question (e.g., definition, completion)
    - » Expected answer type (e.g., PERSON, LOCATION, …)

Why is it Difficult?

- Open-domain:
  - Should not rely on sophisticated use of knowledge bases
  - Should also work, even if input is non-well-formed and in-complete
  - Nevertheless should help to direct search engine and answer processing

Robust Interpretation of NL Queries

- German syntax (SMES):
  - Topological parsing
  - Local Subgrammars for Wh-phrases
- Re-representation
  - Distributed Representation for Dependency Structure
- Query analysis
  - Major information
    - » Q-type (description/definition/…)
    - » A-type (Person/Location/…)
    - » Scope (further constraints for A-type)
  - Q-type determination using Wh-meta terms
    - » "What type of bridge is the Golden Gate Bridge?"
  - Corpus-driven approach for Wh-domain terms
    - » "What is the capital of Somalia?"
- Determines control-information for QA-strategy selection

Examples (and more)

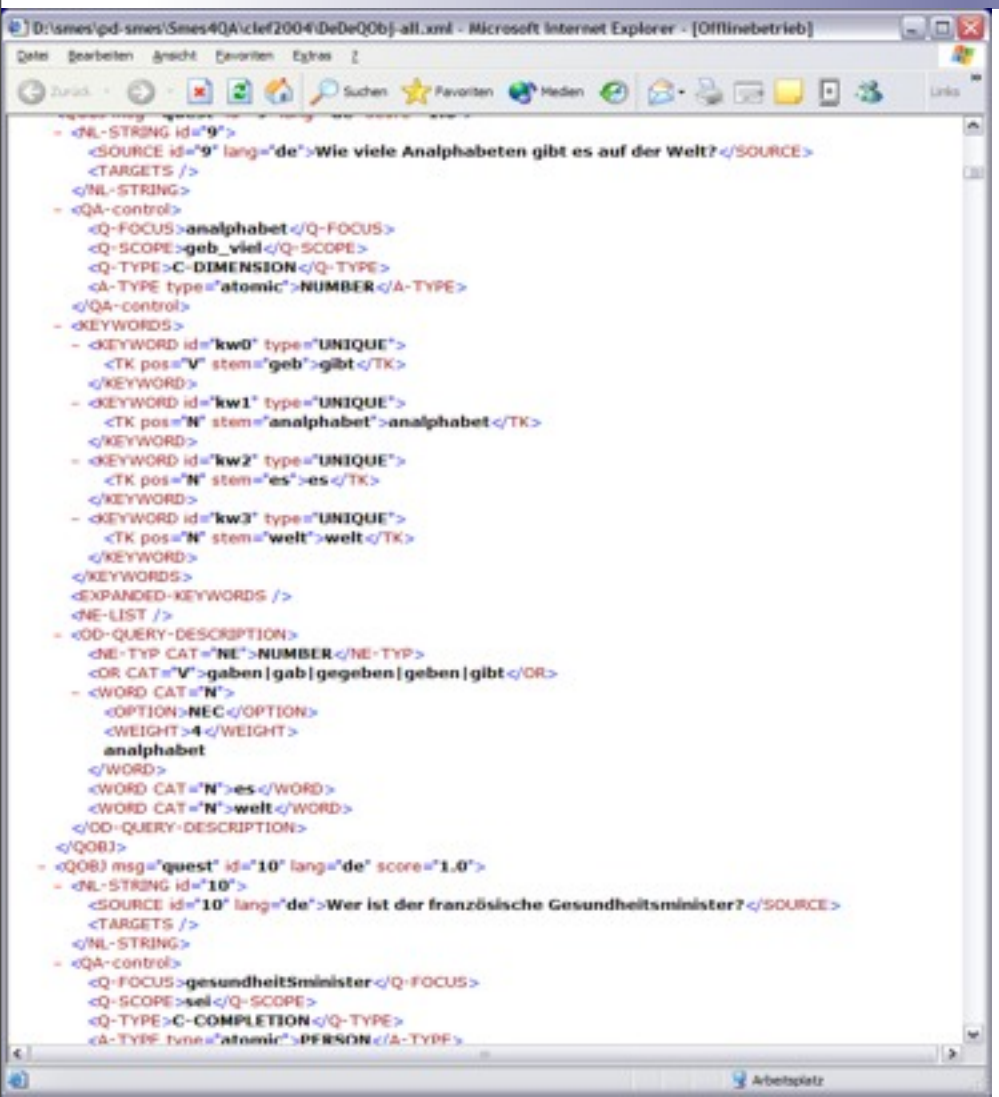
„Was für eine Art Tier ist der Hund? "

<IOOBJ msg='quest' q-type='C-HYPONYM' q-weight='1.0'>
<A-TYPE>ANIMAL</A-TYPE>
<SCOPE>hund</SCOPE>
…

“In welcher Stadt lebte Picasso?"

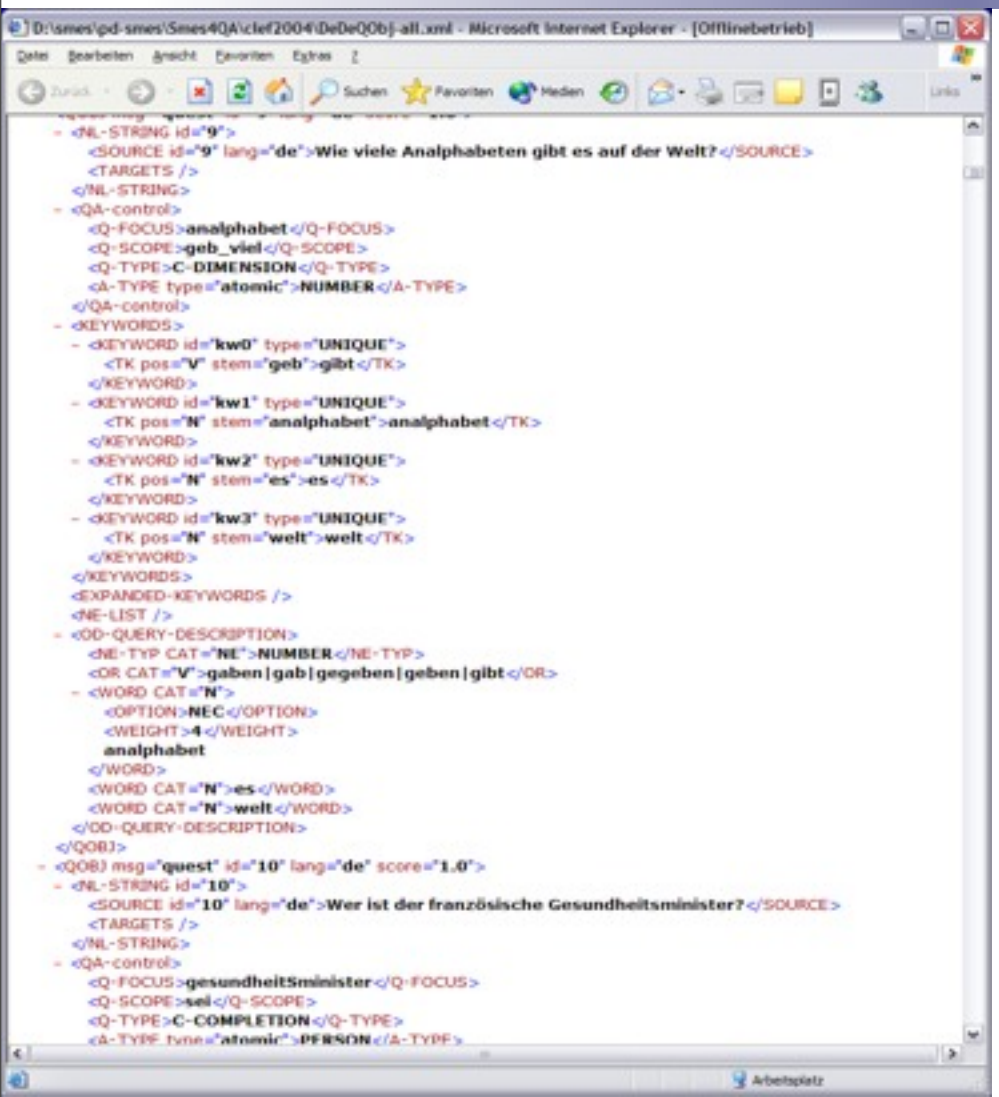
<IOOBJ msg='quest' q-type='C-COMPLETION' q-weight='1.0'>
<A-TYPE>LOCATION</A-TYPE>
<SCOPE>stadt</SCOPE>
…

Open-Domain NL-Query Analysis



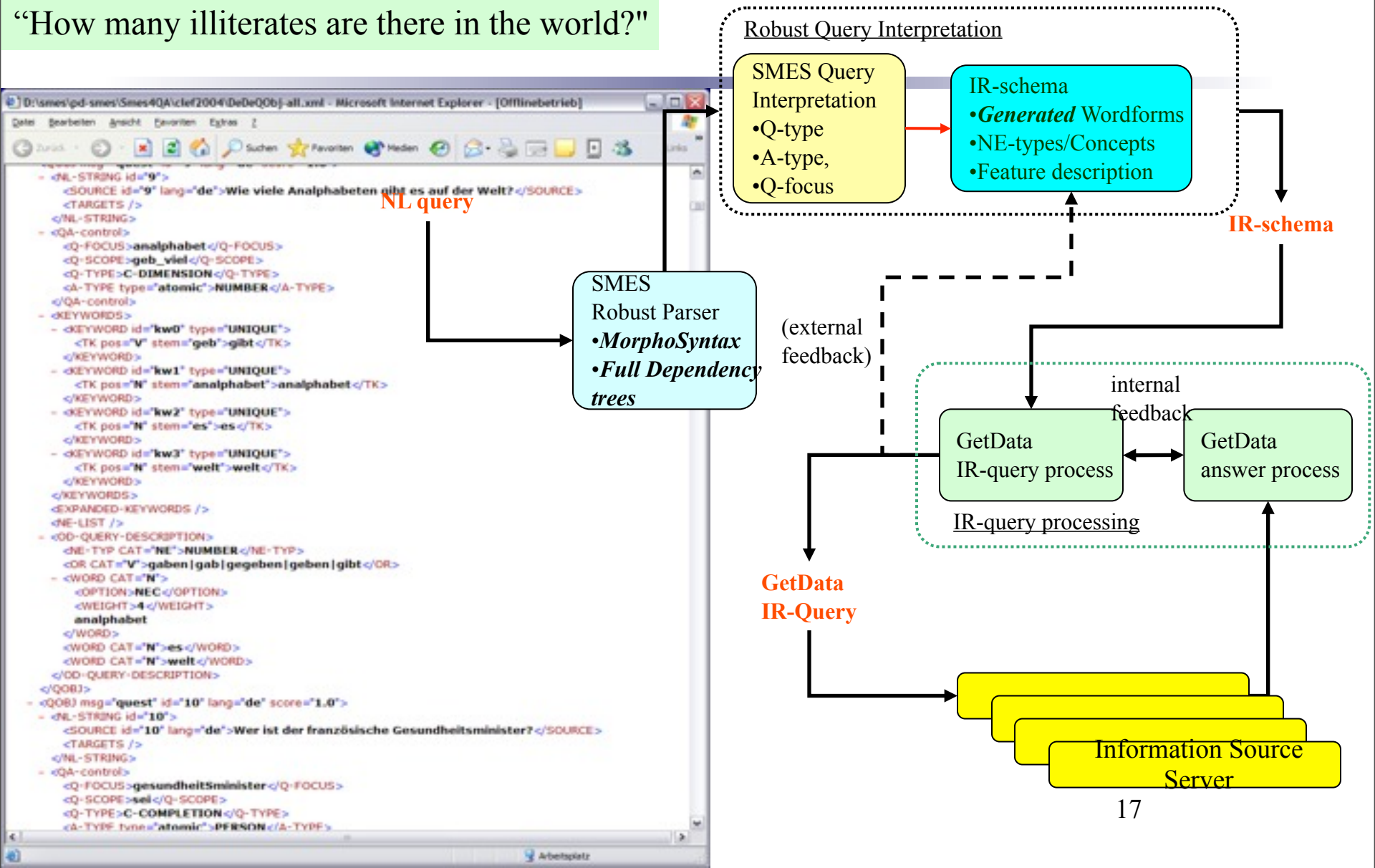
Open-Domain NL-Query Analysis

"How many illiterates are there in the world?"



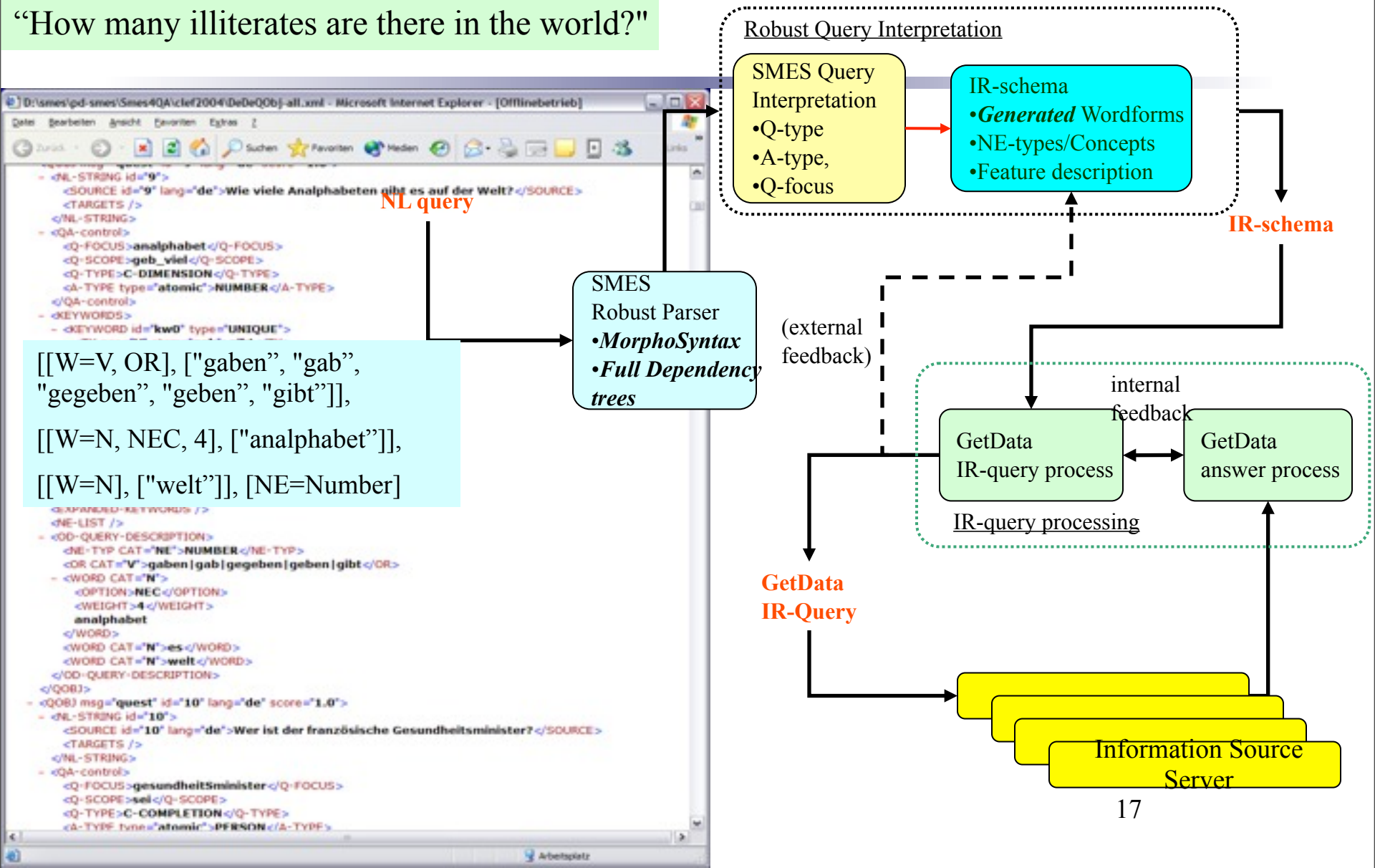
Open-Domain NL-Query Analysis

"How many illiterates are there in the world?"



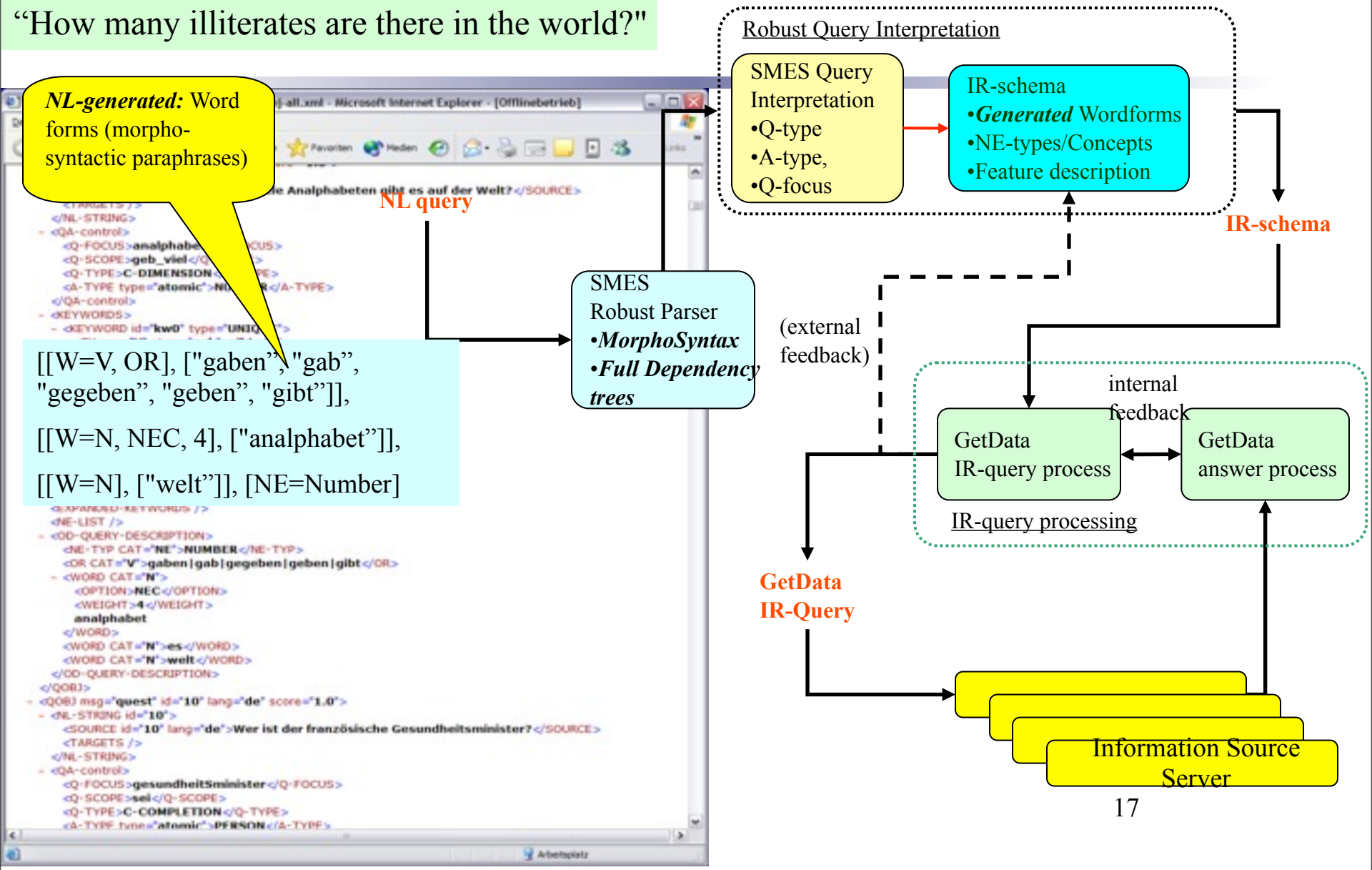
Open-Domain NL-Query Analysis

"How many illiterates are there in the world?"



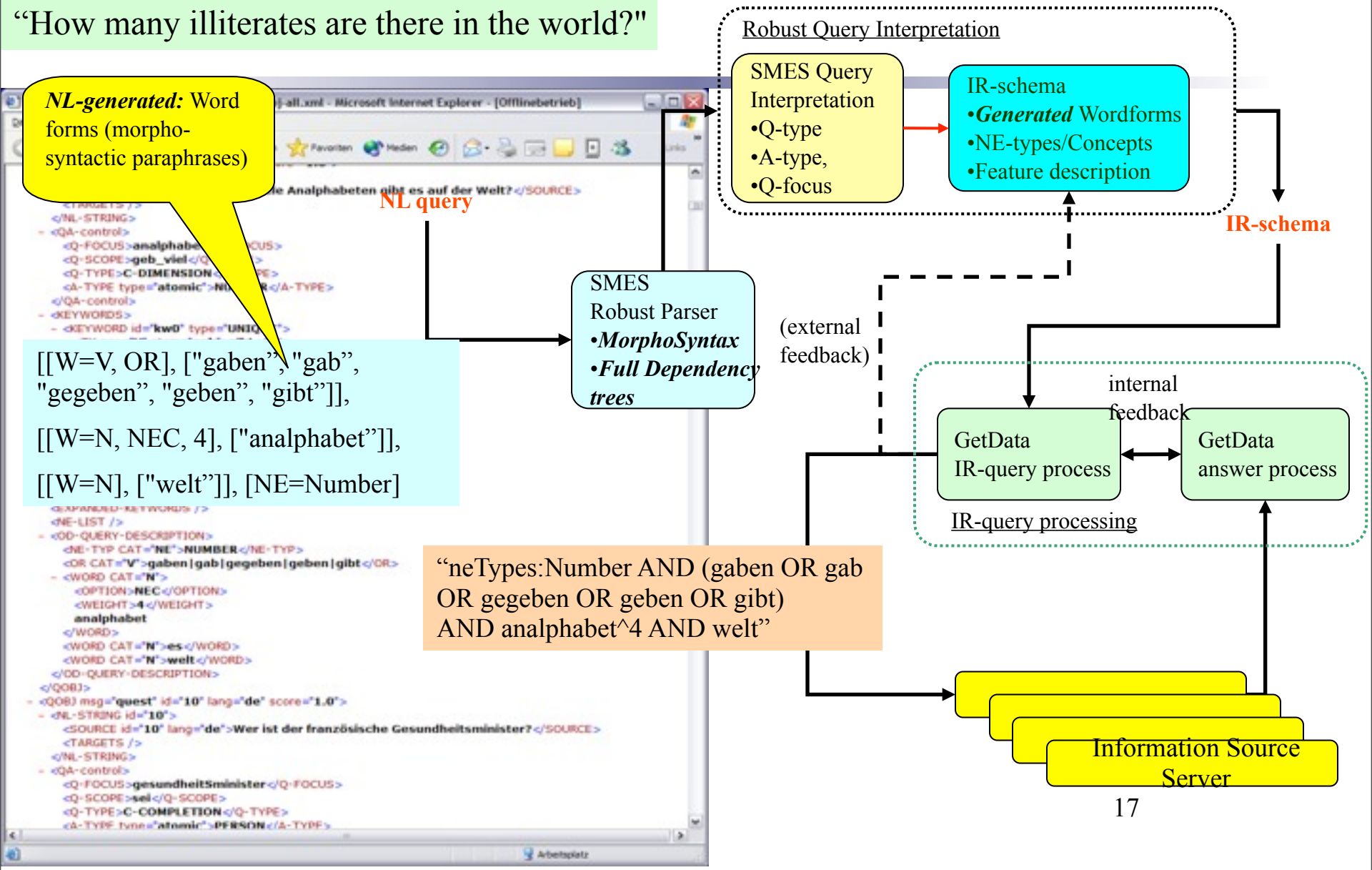
Open-Domain NL-Query Analysis

"How many illiterates are there in the world?"



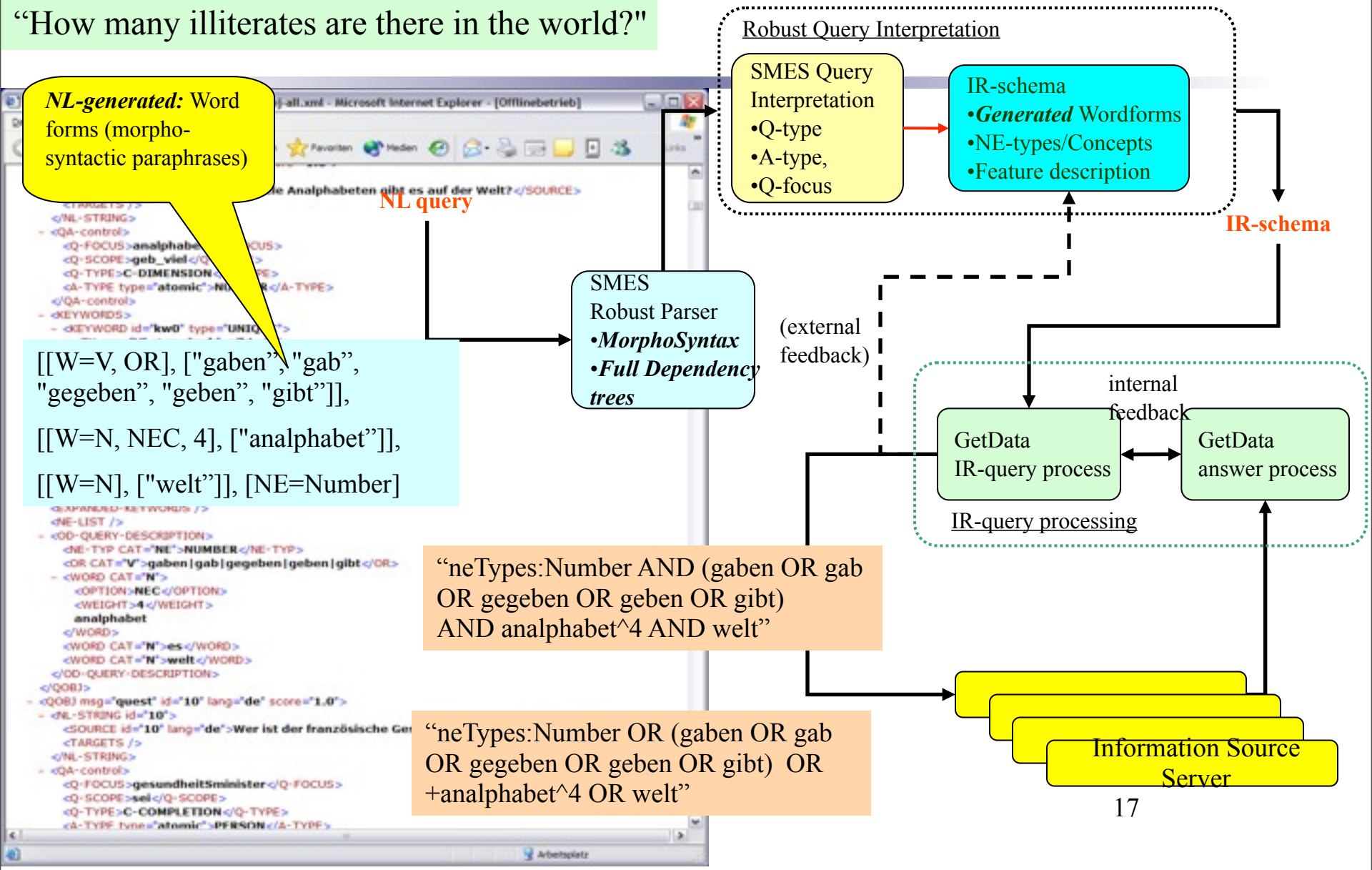
Open-Domain NL-Query Analysis

"How many illiterates are there in the world?"



Open-Domain NL-Query Analysis

"How many illiterates are there in the world?"



Approaches in CL QA

Two main different approaches used in Cross-Language QA systems:

1

translation of the question into the target language (i.e. in the language of the document collection)

question processing

answer extraction

2

question processing in the source language to retrieve information (such as keywords, question focus, expected answer type, etc.)

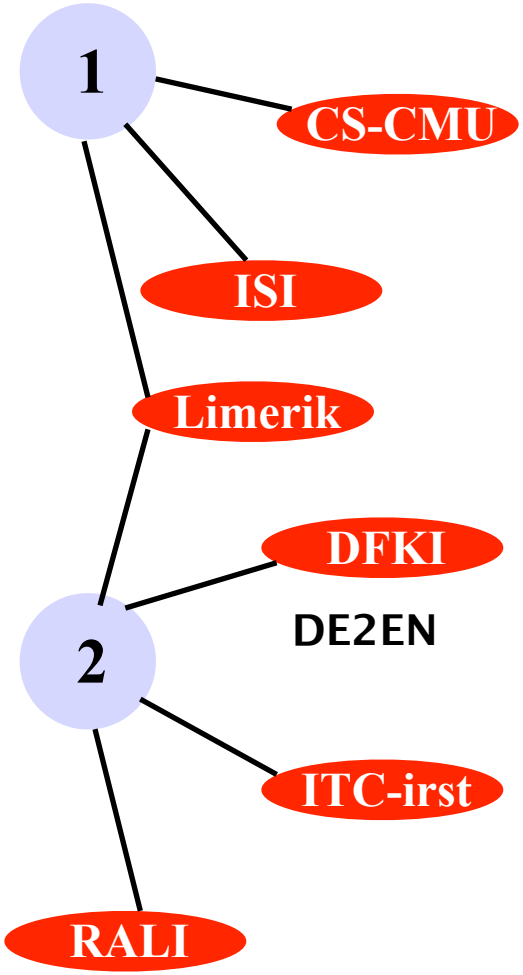
translation and expansion of the retrieved data

answer extraction

Slide from B. Magnini

Approaches in CL QA

Two main different approaches used in Cross-Language QA systems:



translation of the question into the target language (i.e. in the language of the document collection)

question processing

answer extraction

question processing in the source language to retrieve information (such as keywords, question focus, expected answer type, etc.)

translation and expansion of the retrieved data

answer extraction

Slide from B. Magnini

Query Translation & Expansion

- First idea:
  - Only use EuroWordNet
  - Defines a word-based translation via synset offsets
- Experience
  - EuroWordNet too sparse on German side
  - Neverless introduced too much ambiguity
  - NE-translation is crucial
- So far, not very much of help

Query Translation & Expansion

- First idea:
  - Only use EuroWordNet
  - Defines a word-based translation via synset offsets
- Experience
  - EuroWordNet too sparse on German side
  - Neverless introduced too much ambiguity
  - NE-translation is crucial
- So far, not very much of help

- Second idea:

Query Translation & Expansion

- First idea:
  - Only use EuroWordNet
  - Defines a word-based translation via synset offsets
- Experience
  - EuroWordNet too sparse on German side
  - Neverless introduced too much ambiguity
  - NE-translation is crucial
- So far, not very much of help

- Second idea:
  - Use EuroWordNet

Query Translation & Expansion

- First idea:
  - Only use EuroWordNet
  - Defines a word-based translation via synset offsets
- Experience
  - EuroWordNet too sparse on German side
  - Neverless introduced too much ambiguity
  - NE-translation is crucial
- So far, not very much of help

- Second idea:
  - Use EuroWordNet
  - Use external MT-services

Query Translation & Expansion

- First idea:
  - Only use EuroWordNet
  - Defines a word-based translation via synset offsets
- Experience
  - EuroWordNet too sparse on German side
  - Neverless introduced too much ambiguity
  - NE-translation is crucial
- So far, not very much of help

- Second idea:
  - Use EuroWordNet
  - Use external MT-services
  - Overlap-mechanism for query expansion

Query Translation & Expansion

- First idea:
  - Only use EuroWordNet
  - Defines a word-based translation via synset offsets
- Experience
  - EuroWordNet too sparse on German side
  - Neverless introduced too much ambiguity
  - NE-translation is crucial
- So far, not very much of help

- Second idea:
  - Use EuroWordNet
  - Use external MT-services
  - Overlap-mechanism for query expansion
- Crosslingual because

Query Translation & Expansion

- First idea:
  - Only use EuroWordNet
  - Defines a word-based translation via synset offsets
- Experience
  - EuroWordNet too sparse on German side
  - Neverless introduced too much ambiguity
  - NE-translation is crucial
- So far, not very much of help

- Second idea:
  - Use EuroWordNet
  - Use external MT-services
  - Overlap-mechanism for query expansion
- Crosslingual because
  - Q-type & A-type from DE-Question Analysis

Query Translation & Expansion

- First idea:
  - Only use EuroWordNet
  - Defines a word-based translation via synset offsets
- Experience
  - EuroWordNet too sparse on German side
  - Neverless introduced too much ambiguity
  - NE-translation is crucial
- So far, not very much of help

- Second idea:
  - Use EuroWordNet
  - Use external MT-services
  - Overlap-mechanism for query expansion
- Crosslingual because
  - Q-type & A-type from DE-Question Analysis
  - Synsets from EuroWN direct query expansion (online alignment)

Example

1. Translation services for Word Sense Disambiguation (WSD)

Wo wurde das Militärflugzeug Strike Eagles 1990 eingesetzt?

2. Query expansion using EuroWordNet

Example

1. Translation services for Word Sense Disambiguation (WSD)

Wo wurde das Militärflugzeug Strike Eagles 1990 eingesetzt?

2. Query expansion using EuroWordNet

Example

1. Translation services for Word Sense Disambiguation (WSD)

Wo wurde das Militärflugzeug Strike Eagles 1990 eingesetzt?

FreeTranslation: Where did the military airplane become would strike used Eagles 1990?

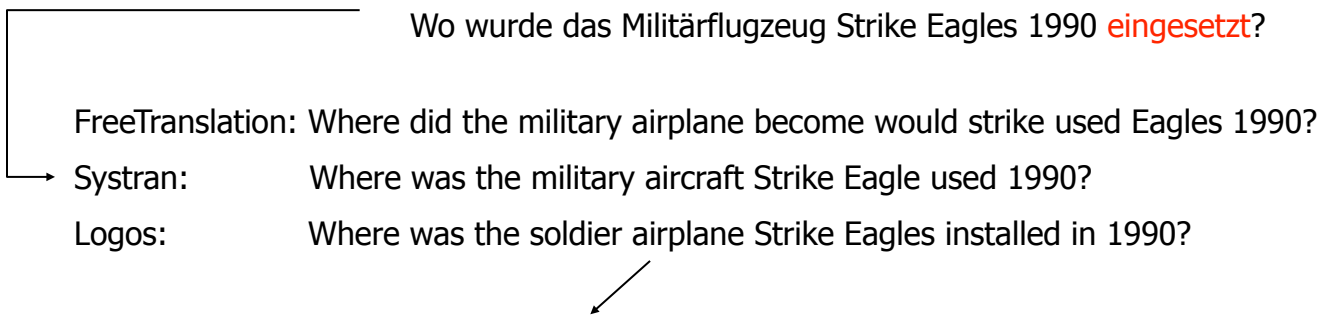
Systran: Where was the military aircraft Strike Eagle used 1990?

Logos: Where was the soldier airplane Strike Eagles installed in 1990?

2. Query expansion using EuroWordNet

Example

1. Translation services for Word Sense Disambiguation (WSD)



2. Query expansion using EuroWordNet

Example

1. Translation services for Word Sense Disambiguation (WSD)

Wo wurde das Militärflugzeug Strike Eagles 1990 eingesetzt?

FreeTranslation: Where did the military airplane become would strike used Eagles 1990?

Systran: Where was the military aircraft Strike Eagle used 1990?

Logos: Where was the soldier airplane Strike Eagles installed in 1990?

BoO_{EN} := {soldier, airplane, strike, eagle, install, 1990, military, become, strike, use, aircraft}

2. Query expansion using EuroWordNet

Example

1. Translation services for Word Sense Disambiguation (WSD)

Wo wurde das Militärflugzeug Strike Eagles 1990 eingesetzt?

FreeTranslation: Where did the military airplane become would strike used Eagles 1990?

Systran: Where was the military aircraft Strike Eagle used 1990?

Logos: Where was the soldier airplane Strike Eagles installed in 1990?

BoO_{EN} := {soldier, airplane, strike, eagle, install, 1990, military, become, strike, use, aircraft}

2. Query expansion using EuroWordNet

$\forall x \in BoO_{EN}$: lookup(EuroWN);
If x is unambiguous: extend BoO_{EN}
Else ∀readings(x):
get its aligned German readings &
Look them up in BoO_{GN}
If successfully then add English terms to BoO_{EN}

Example

1. Translation services for Word Sense Disambiguation (WSD)

Wo wurde das Militärflugzeug Strike Eagles 1990 eingesetzt?

FreeTranslation: Where did the military airplane become would strike used Eagles 1990?

Systran: Where was the military aircraft Strike Eagle used 1990?

Logos: Where was the soldier airplane Strike Eagles installed in 1990?

BoO_{EN} := {soldier, airplane, strike, eagle, install, 1990, military, become, strike, use, aircraft}

2. Query expansion using EuroWordNet

$\forall x \in BoO_{EN}$: lookup(EuroWN);
If x is unambiguous: extend BoO_{EN}
Else $\forall$readings(x):
get its aligned German readings &
Look them up in BoO_{GN}
If successfully then add English terms to BoO_{EN}

Reading-697925

EN: {handle, use, wield}

DE: {handhaben, hantieren}

Example

1. Translation services for Word Sense Disambiguation (WSD)

Wo wurde das Militärflugzeug Strike Eagles 1990 eingesetzt?

FreeTranslation: Where did the military airplane become would strike used Eagles 1990?

Systran: Where was the military aircraft Strike Eagle used 1990?

Logos: Where was the soldier airplane Strike Eagles installed in 1990?

BoO_{EN} := {soldier, airplane, strike, eagle, install, 1990, military, become, strike, use, aircraft}

2. Query expansion using EuroWordNet

$\forall x \in BoO_{EN}$: lookup(EuroWN);
If x is unambiguous: extend BoO_{EN}
Else ∀readings(x):
get its aligned German readings &
Look them up in BoO_{GN}
If successfully then add English terms to BoO_{EN}

Reading-697925
EN: {handle, use, wield}
DE: {handhaben, hantieren}

Example

1. Translation services for Word Sense Disambiguation (WSD)

Wo wurde das Militärflugzeug Strike Eagles 1990 eingesetzt?

FreeTranslation: Where did the military airplane become would strike used Eagles 1990?

Systran: Where was the military aircraft Strike Eagle used 1990?

Logos: Where was the soldier airplane Strike Eagles installed in 1990?

BoO_{EN} := {soldier, airplane, strike, eagle, install, 1990, military, become, strike, use, aircraft}

2. Query expansion using EuroWordNet

$\forall x \in BoO_{EN}$: lookup(EuroWN);
If x is unambiguous: extend BoO_{EN}
Else ∀readings(x):
get its aligned German readings &
Look them up in BoO_{GN}
If successfully then add English terms to BoO_{EN}

Reading-697925
EN: {handle, use, wield}
DE: {handhaben, hantieren}

Reading-1453934:
EN: {behave toward, use}
DE: not aligned

Example

1. Translation services for Word Sense Disambiguation (WSD)

Wo wurde das Militärflugzeug Strike Eagles 1990 eingesetzt?

FreeTranslation: Where did the military airplane become would strike used Eagles 1990?

Systran: Where was the military aircraft Strike Eagle used 1990?

Logos: Where was the soldier airplane Strike Eagles installed in 1990?

BoO_{EN} := {soldier, airplane, strike, eagle, install, 1990, military, become, strike, use, aircraft}

2. Query expansion using EuroWordNet

$\forall x \in BoO_{EN}$: lookup(EuroWN);
If x is unambiguous: extend BoO_{EN}
Else ∀readings(x):
get its aligned German readings &
Look them up in BoO_{GN}
If successfully then add English terms to BoO_{EN}

Reading-697925
EN: {handle, use, wield}
DE: {handhaben, hantieren}

Reading-1453934:
EN: {behave toward, use}
DE: not aligned

Example

1. Translation services for Word Sense Disambiguation (WSD)

Wo wurde das Militärflugzeug Strike Eagles 1990 eingesetzt?

FreeTranslation: Where did the military airplane become would strike used Eagles 1990?

Systran: Where was the military aircraft Strike Eagle used 1990?

Logos: Where was the soldier airplane Strike Eagles installed in 1990?

BoO_{EN} := {soldier, airplane, strike, eagle, install, 1990, military, become, strike, use, aircraft}

2. Query expansion using EuroWordNet

$\forall x \in BoO_{EN}$: lookup(EuroWN);
If x is unambiguous: extend BoO_{EN}
Else ∀readings(x):
get its aligned German readings &
Look them up in BoO_{GN}
If successfully then add English terms to BoO_{EN}

Reading-697925
EN: {handle, use, wield}
DE: {handhaben, hantieren}

Reading-1453934:
EN: {behave toward, use}
DE: not aligned

Reading-658243:
EN: {apply, employ, make use of, put to use, use, utilise, utilize}
DE: {anbringen, anwenden, bedienen, benutzen, einsetzen, ...}

Example

1. Translation services for Word Sense Disambiguation (WSD)

Wo wurde das Militärflugzeug Strike Eagles 1990 eingesetzt?

FreeTranslation: Where did the military airplane become would strike used Eagles 1990?

Systran: Where was the military aircraft Strike Eagle used 1990?

Logos: Where was the soldier airplane Strike Eagles installed in 1990?

BoO_{EN} := {soldier, airplane, strike, eagle, install, 1990, military, become, strike, use, aircraft}

2. Query expansion using EuroWordNet

$\forall x \in BoO_{EN}$: lookup(EuroWN);
If x is unambiguous: extend BoO_{EN}
Else ∀readings(x):
get its aligned German readings &
Look them up in BoO_{GN}
If successfully then add English terms to BoO_{EN}

Reading-697925
EN: {handle, use, wield}
DE: {handhaben, hantieren}

Reading-1453934:
EN: {behave toward, use}
DE: not aligned

Reading-658243:
EN: {apply, employ, make use of, put to use, use, utilise, utilize}
DE: {anbringen, anwenden, bedienen, benutzen, einsetzen, ...}

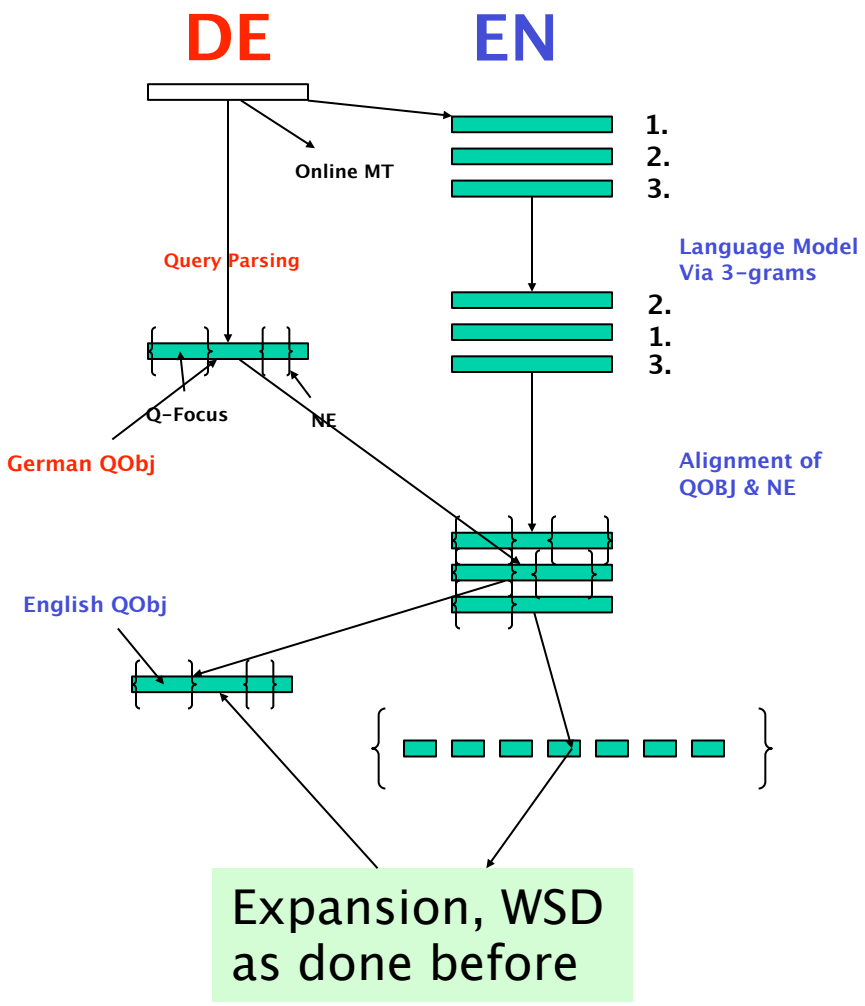
Cross-lingual Open-Domain Question Answering

- How we have done it so far:
  - Use external MT-services
  - Overlap-mechanism &
  - EuroWordNet for query expansion
- Experience
  - External MT services
    - » used ako WSD
    - » as well as query expansion
  - Reduced degree of ambiguity
- Problems/Restrictions
  - Important information from German QueryObject (e.g., Q-focus, Q-scope) cannot be carried over
    ⇒ Need of query parsing on English side
    ⇒ restricted view on cross-linguality
  - Translations of NE
  - Translated strings are neutral wrt. corpus

Hybrid NL-Query Translation for Cross-lingual ODQA

Improvements

- Language Model
  - translations from the on-line MT systems are ranked according to a language model
  - 3-gram model of document corpus
    ⇒ corpus-sensible ranking of translations
- Allignment of Query-Information
  - based on several filters (dictionary, PoS & string similarity)
    ⇒ "transformation" of DE-QueryObject (Q-Focus) onto to EN-translation
    ⇒ no need of parsing on English side
- NE-specific alignment
  - Not person names
  - but organizations, locations



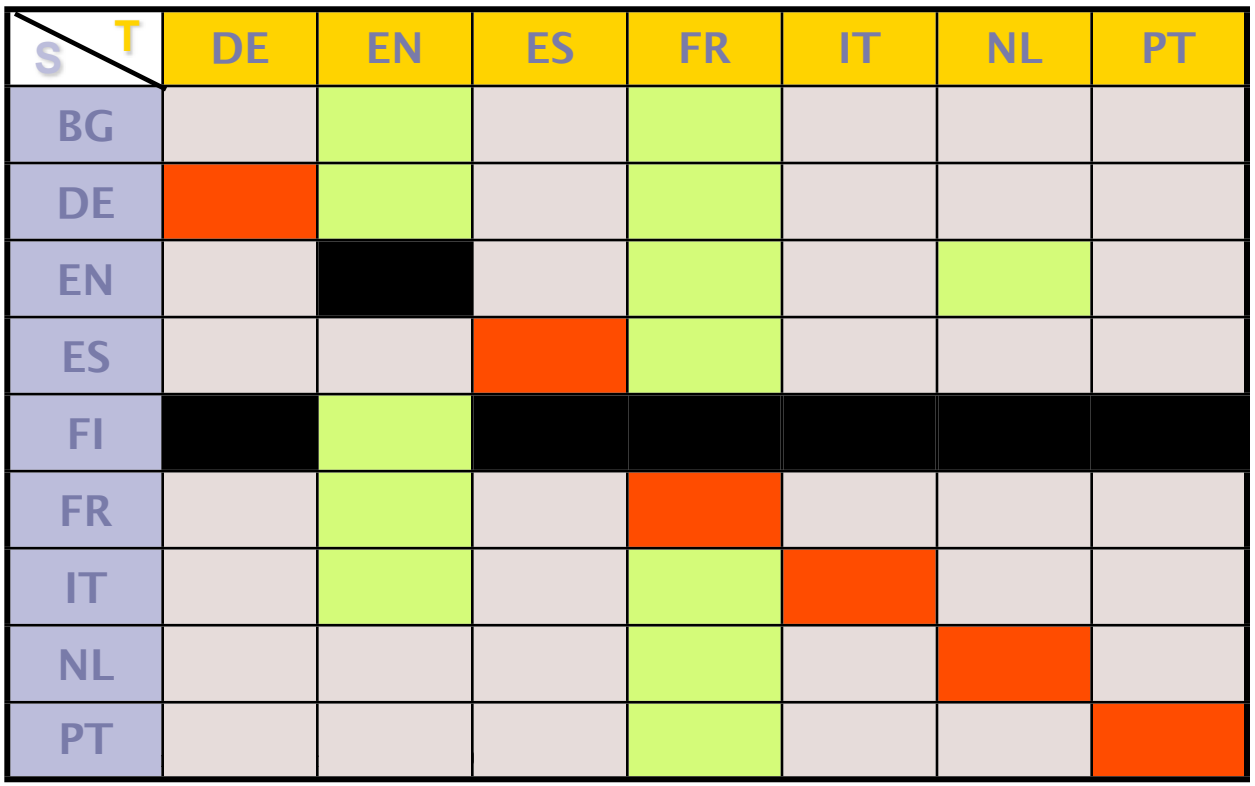
We wanted to use the Clef Forum for External Evaluation of Our ODQA approach: So, what is Clef?

- The Cross-Language Evaluation Forum (CLEF) supports global digital library applications by
  - developing an infrastructure for the testing, tuning and evaluation of information retrieval systems operating on European languages and
  - creating test-suites of reusable data which can be employed by system developers for benchmarking purposes.

- QA@Clef 2003:
  - Initial pilot evaluation for Cross-Lingual Open-Domain Question Answering (6 groups, among 3 from overseas)

QA@Clef 2004 Track Setup – Task Definition

Given **200 questions** in a source language, find ***one exact answer*** per question in a collection of documents written in a target language, and provide a justification for each retrieved answer (i.e. the `docid` of the unique document that supports the answer).



6 monolingual and 50 bilingual tasks.

18 Teams participated in 19 tasks, submitting 48 runs.

Slide from A. Vallin

Evaluation Exercise – Results (EN)

Results of the runs with English as target language.

Run Name	R	W	X	U	Overall Accuracy (%)	Accuracy over F (%)	Accuracy over D (%)	NIL Accuracy		CWS
								Precision	Recall	
bgas041bgen	26	168	5	1	13.00	11.67	25.00	0.13	0.40	0.056
dfki041deen	47	151	0	2	**23.50**	23.89	20.00	0.10	0.75	0.177
dltg041fren	38	155	7	0	19.00	17.78	30.00	0.17	0.55	-
dltg042fren	29	164	7	0	14.50	12.78	30.00	0.14	0.45	-
edin041deen	28	166	5	1	14.00	13.33	20.00	0.14	0.35	0.049
edin041fren	33	161	6	0	16.50	17.78	5.00	0.15	0.55	0.056
edin042deen	34	159	7	0	17.00	16.11	25.00	0.14	0.35	0.052
edin042fren	40	153	7	0	**20.00**	20.56	15.00	0.15	0.55	0.058
hels041fien	21	171	1	0	10.88	11.56	5.00	0.10	0.85	0.046
irst041iten	45	146	6	3	**22.50**	22.22	25.00	0.24	0.30	0.121
irst042iten	35	158	5	2	17.50	16.67	25.00	0.24	0.30	0.075
lire041fren	22	172	6	0	11.00	10.00	20.00	0.05	0.05	0.032
lire042fren	39	155	6	0	19.50	20.00	15.00	0.00	0.00	0.075

Slide from A. Vallin

QA@Clef 2005

- Temporal questions
- 23 groups
- DFKI
  - DE2DE, 2 runs
  - DE2EN
  - EN2DE, 2 runs

QA@CLEF-2005

The 67 runs that have been submitted are divided into:

- 2 in the tasks with *BG* as target (both monolingual)
- 6 in the tasks with *DE* as target (3 monolingual and 3 cross-language)
- 12 in the tasks with *EN* as target (all cross-language)
- 18 in the tasks with *ES* as target (13 monolingual and 5 cross-language)
- 2 in the tasks with *FI* as target (both monolingual)
- 13 in the tasks with *FR* as target (10 monolingual and 3 cross-language)
- 6 in the tasks with *IT* as target (all monolingual)
- 3 in the tasks with *NL* as target (all monolingual)
- 5 in the tasks with *PT* as target (4 monolingual and 1 cross-language)

Temporal Questions

- Examples
  - Who was world champignon in soccer 1966?
  - How often did German scientists won a Nobel Price between 1950 and 1980?
  - Who was US president during German's re-unification phase?
- Some challenges
  - Implicit temporal expressions
  - Answer candidates contain explicit mentioning of dates, but query refers to interval
- Our strategy
  - Construct date-related underspecified IR-query
  - Check consistency between date instances from answer candidates and query

Web-based Query Analysis

- In principle the same approach as for Lucene-IR-query; Major differences
  - Google-specific IR-query construction
  - IR-query construction: Return N-best snippets
  - Snippet post-processing
    - » A-type + snippet-structure,
      - e.g, use ... for distance measure
    - » Construction of possible well-formed string
      - "Snippet grammar"
      - application of NL parsing tool
    - » Latent semantic kernels for answer extraction
- WebQA-based answer validation
  - Q-type/Q-focus + Query string (Q) + Answer candidate (AC) = direct_answer_string
  - Google's Total Estimated Counts (TEC) for re-ranking
  - Ongoing: trustworthiness for answer candidates (utility theory; authority

Web-based Query Analysis

- In principle the same approach as for Lucene-IR-query; Major differences
  - Google-specific IR-query construction
  - IR-query construction: Return N-best snippets
  - Snippet post-processing
    - » A-type + snippet-structure,
      - e.g, use ... for distance measure
    - » Construction of possible well-formed string
      - "Snippet grammar"
      - application of NL parsing tool
    - » Latent semantic kernels for answer extraction
- WebQA-based answer validation
  - Q-type/Q-focus + Query string (Q) + Answer candidate (AC) = direct_answer_string
  - Google's Total Estimated Counts (TEC) for re-ranking
  - Ongoing: trustworthiness for answer candidates (utility theory; authority

Q: What is the capital of Germany?
AC: Berlin, Ney York

Web-based Query Analysis

- In principle the same approach as for Lucene-IR-query; Major differences
  - Google-specific IR-query construction
  - IR-query construction: Return N-best snippets
  - Snippet post-processing
    - » A-type + snippet-structure,
      - e.g, use ... for distance measure
    - » Construction of possible well-formed string
      - "Snippet grammar"
      - application of NL parsing tool
    - » Latent semantic kernels for answer extraction
- WebQA-based answer validation
  - Q-type/Q-focus + Query string (Q) + Answer candidate (AC) = direct_answer_string
  - Google's Total Estimated Counts (TEC) for re-ranking
  - Ongoing: trustworthiness for answer candidates (utility theory; authority

Q: What is the capital of Germany?
AC: Berlin, Ney York

Web-based Query Analysis

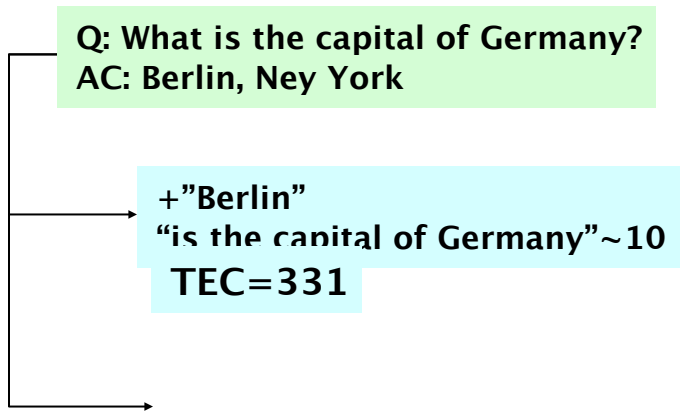
- In principle the same approach as for Lucene-IR-query; Major differences
  - Google-specific IR-query construction
  - IR-query construction: Return N-best snippets
  - Snippet post-processing
    - » A-type + snippet-structure,
      - e.g, use ... for distance measure
    - » Construction of possible well-formed string
      - "Snippet grammar"
      - application of NL parsing tool
    - » Latent semantic kernels for answer extraction
- WebQA-based answer validation
  - Q-type/Q-focus + Query string (Q) + Answer candidate (AC) = direct_answer_string
  - Google's Total Estimated Counts (TEC) for re-ranking
  - Ongoing: trustworthiness for answer candidates (utility theory; authority

Q: What is the capital of Germany?
AC: Berlin, Ney York

+"Berlin"
"is the capital of Germany"~10
TEC=331

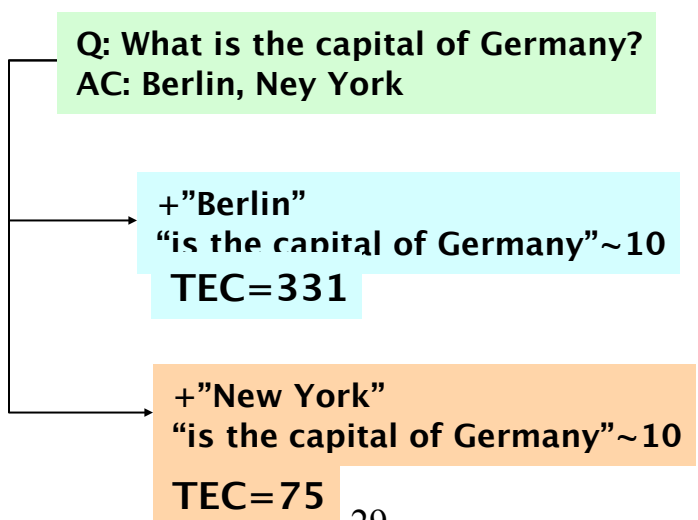
Web-based Query Analysis

- In principle the same approach as for Lucene-IR-query; Major differences
  - Google-specific IR-query construction
  - IR-query construction: Return N-best snippets
  - Snippet post-processing
    - » A-type + snippet-structure,
      - e.g, use ... for distance measure
    - » Construction of possible well-formed string
      - "Snippet grammar"
      - application of NL parsing tool
    - » Latent semantic kernels for answer extraction
- WebQA-based answer validation
  - Q-type/Q-focus + Query string (Q) + Answer candidate (AC) = direct_answer_string
  - Google's Total Estimated Counts (TEC) for re-ranking
  - Ongoing: trustworthiness for answer candidates (utility theory; authority



Web-based Query Analysis

- In principle the same approach as for Lucene-IR-query; Major differences
  - Google-specific IR-query construction
  - IR-query construction: Return N-best snippets
  - Snippet post-processing
    - » A-type + snippet-structure,
      - e.g, use ... for distance measure
    - » Construction of possible well-formed string
      - "Snippet grammar"
      - application of NL parsing tool
    - » Latent semantic kernels for answer extraction
- WebQA-based answer validation
  - Q-type/Q-focus + Query string (Q) + Answer candidate (AC) = direct_answer_string
  - Google's Total Estimated Counts (TEC) for re-ranking
  - Ongoing: trustworthiness for answer candidates (utility theory; authority



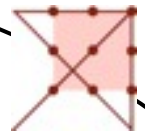
DFKI Related Future Work

- SmartWeb
  - Mobile access to the Semantic Web
  - ODQA for search in syntactic Web pages
  - BMBF Verbund project (partners from research and industries, e.g., BMW, Siemens, T-Systems)
- HyLab
  - QA for Personal Digital Memory
  - BMBF funded DFKI project (start 2006)

BMBF

Softwaresysteme

DaimlerChrysler



European Media Lab

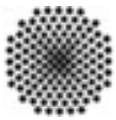
Fraunhofer Institut Rechnerarchitektur und Softwaretechnik

IT-2006 und Futur-Programm

BMBF: 13,7 Mio Euro (Dr. Reuse)

Leiter: W. Wahlster

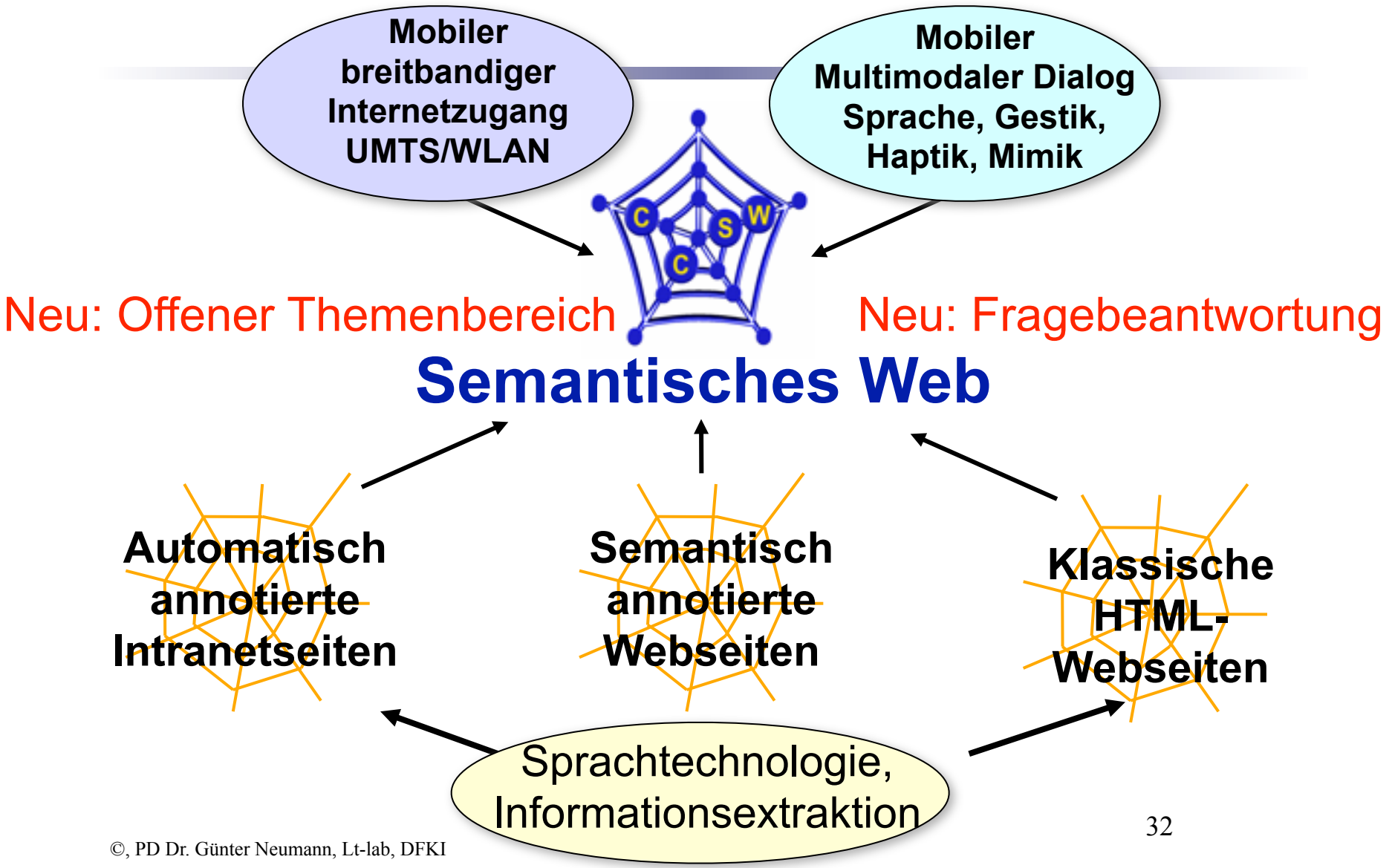
Laufzeit: 2004-2007



IMS Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart

Ludwig-Maximilians-Universität München

SmartWeb: Mobiler Breitbandzugang



DONE!

Thank you for your attention!

HyLaP

Hybrid Language Processing Technologies for a personal associative information access and management application

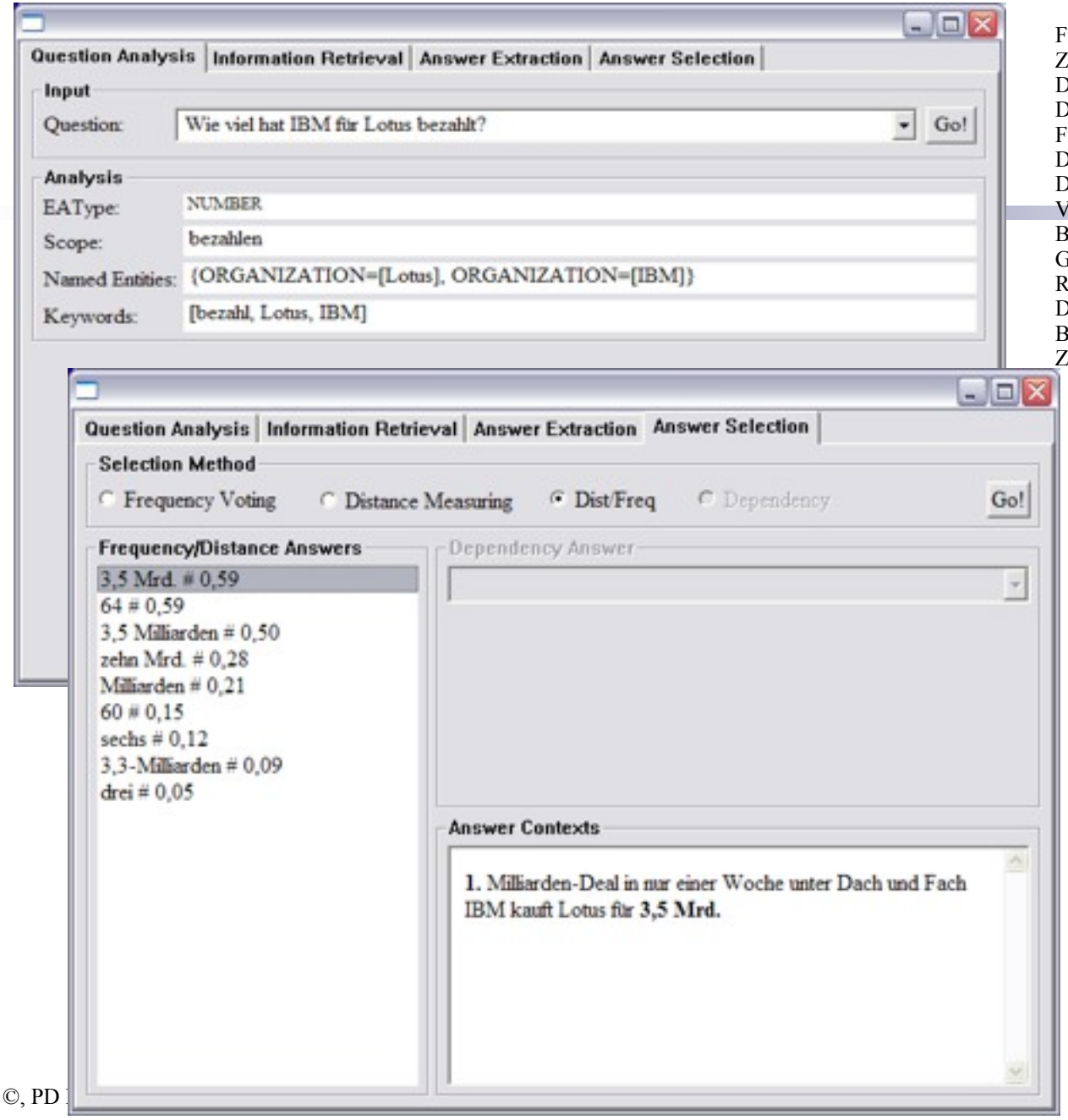
Günter Neumann & Hans Uszkoreit

BMBF funded DFKI project, ~1.5 Euro

Start: 2006

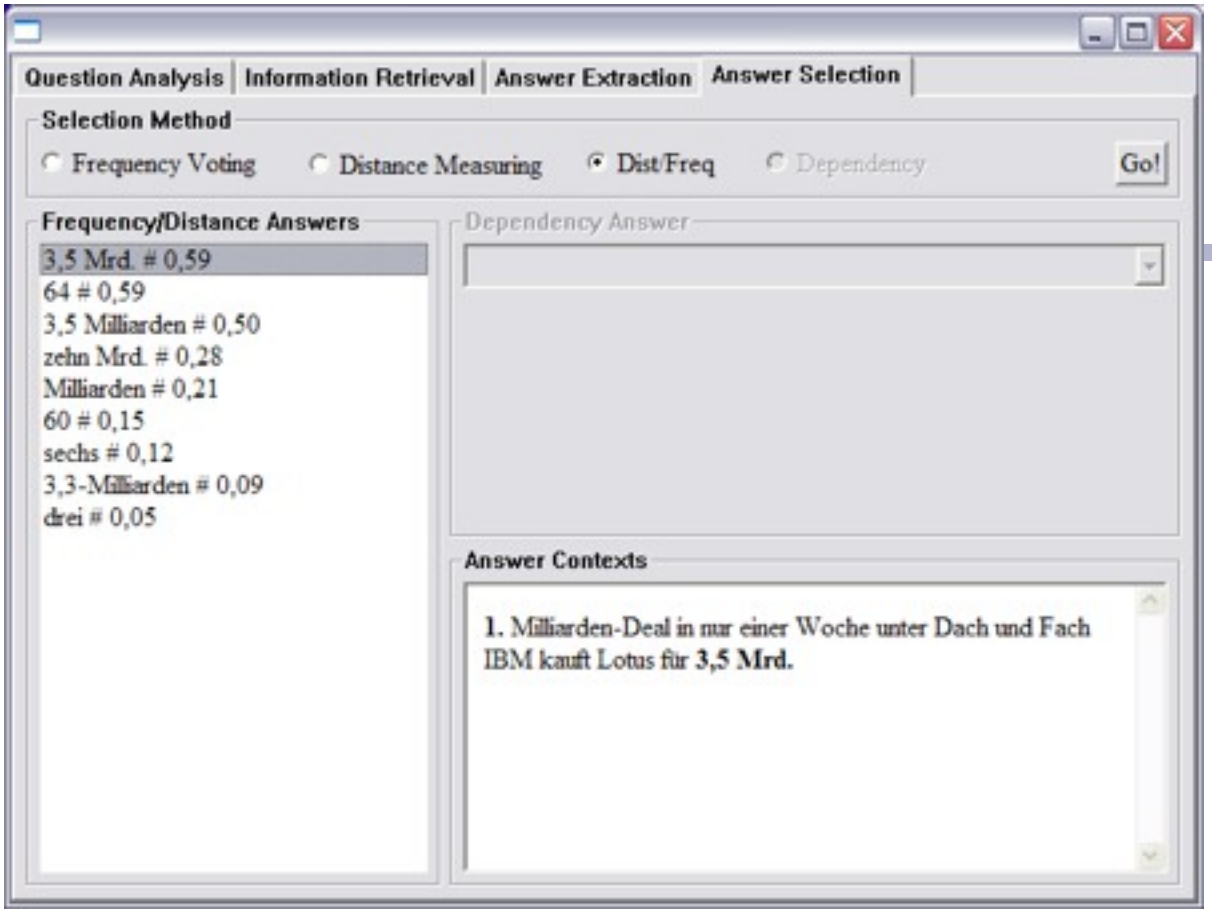
Assumptions

- Question answering is the most natural way of requesting information
- The management of personal digital content becomes a true challenge
- The exploitation of personal digital memory will change our lives
- Applications for authoring, browsing and commenting will converge
- Every document can become an interface to memory



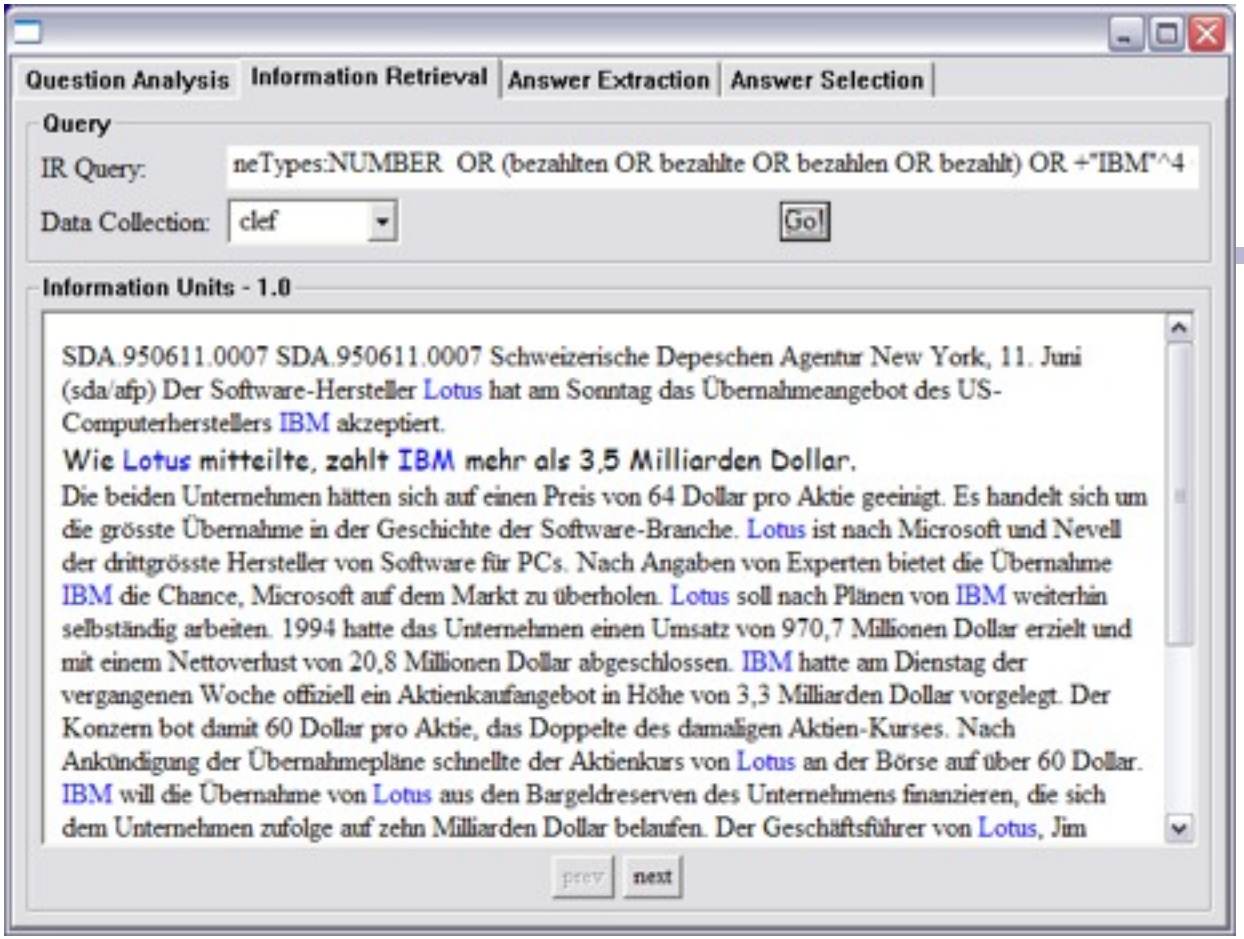
Frageanalyse:
Zentral ist die Berechung
Des Fragetyps (Faktenfrage,
Definitionsfrage) und des
Fragegegenstandes (Person,
Datum, Ort, …).
Da hierbei linguistischbasierte
Verfahren eingesetzt müssen,
Besteht die Kunst darin, dies
Gleichermassen Präzise und
Robust machen zu können ohne
Dem Benutzer eine Einschränkung
Bei seinen Formulierungsbemühungen
Zu machen.

Ergebnisauswahl:
Wenn das System eine Menge von Antwortkandidaten bestimmt hat, müssen gemäß ihrer Güte sortiert werden. Dadurch kann ´dem Benutzer zB nur die beste Antwort geliefert werden oder die N-besten, sodass der Benutzer zB selber noch entscheiden kann, welche er als die Beste Antwort auffast. Dies wird am besten dadurch realisiert, dass zu jedem Antwortkandidat auch der Satz, in dem Antwort vorkommt angezeigt wird. So kann der Benutzer stets auch den Kontext und auch die Güte des Systems selbst noch einmal begutachten.

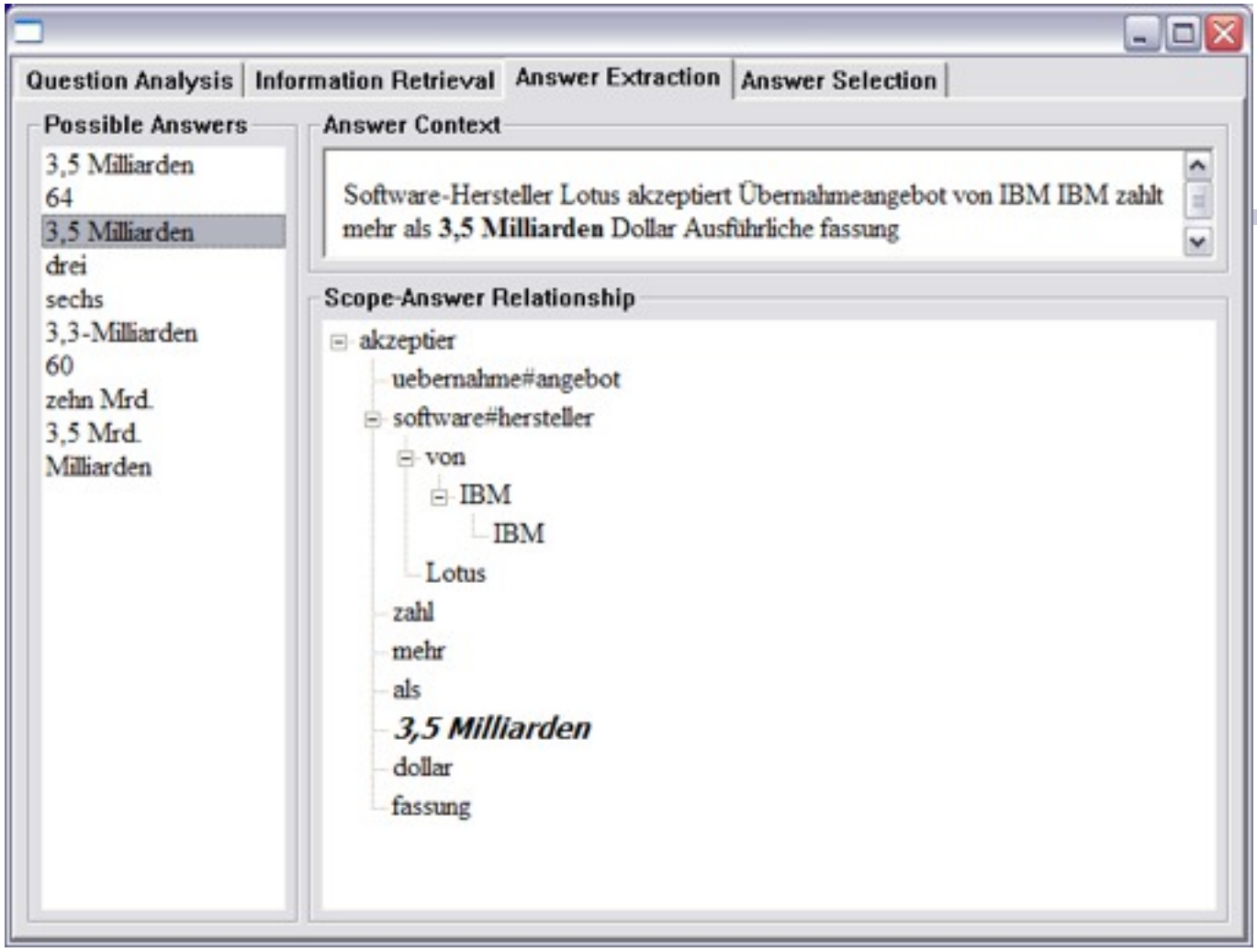


Ergebnisauswahl:
Wenn das System eine Menge von Antwortkandidaten bestimmt hat, müssen gemäß ihrer Güte sortiert werden. Dadurch kann ´dem Benutzer zB nur die beste Antwort geliefert werden oder die N-besten, sodass der Benutzer zB selber noch entscheiden kann, welche er als die Beste Antwort auffast. Dies wird am besten dadurch realisiert, dass zu jedem Antwortkandidat auch der Satz, in dem Antwort vorkommt angezeigt wird. So kann der Benutzer stets auch den Kontext und auch die Güte des Systems selbst noch einmal begutachten.
Bei der Anwser selection koennen verschiedene Verfahren eingesetzt werden, auch kombiniert:
-Frequenz
-Distanz zwischen Antwortkandidat und Fragewörtern
-Beides gemischt

Die Kunst hierbei ist es, solche Kriterien zu finden, die möglichst korrekte Antwortkandidaten von falschen Antworten unterscheiden. Jenachdem, ob das QA open-domain oder domain-spezifisch eingesetzt wird, können hjerzu auch Ontologien und spezielle Inferenzmechansismen eingesetzt werden. Desweiteren bietet sich hier die Möglichkeit, erfolgreiche Berechnungen zu speichern (episodic memory)



Dokumentenauswahl:
Das Ergenbis der Frageverarbeitung muss so umgeformt werden, das sie optimal zur Information retrieval eingesetzt werden kann. Hier benutzen wir sogar Verfahren zur linguistischen Generierung von Wörtern/Phrasen, um einen hohen recall zu erreichen.
Die Kunst hier ist es, die Frageanalyse in eine optimale IR-Suchanfrage zu überführen, sodass bereits beim Retrieval nur sehr wenige, aber relevante Dokumente bestimmt werden.



Antwortextraktion:
Die durch die IR-Engine bestimmten Paragrafen, müssen für die Identifikation und Extraktion der Antwortkandidaten weiter analysiert werden, wobei neben IE-Techniken auch eien robuste NL-Verarbeitung eingesetzt wird. Zum Beispiel ist es hilfreich, die Beziehungen zwischen den einzelnen Wörtern genau zu bestimmen, damit zB genauer Abstandsmasse eignesetzt werden können. Auch ist es gerade fürs Deutsche zentral, online zusammengesetzte Wörter zu identifizieren.
Die Kunst ist zum einen der Einsatz sehr robuster und sehr schneller Sprachtechnologie, aber vorallem der Einsatz automatischer Methoden zur Skalierung relevanter Extraktionsparameter.