Question Answering over Linked Data (QALD)

Horacio Rodríguez

TALP

Outline

- Introduction
- A bit of History of QA
- General QA systems
- Factoid QA
- Beyond Factoid QA
- DRQA
- CQA
- QALD
- Applications
- Conclusions

Introduction

- QA can be defined as the task of given a user information need, expressed as a NL question (basically in QA the user query consists of a question expressed in NL, sometimes, however, limited forms of NL, the so called Controlled NL, are used instead) provide to the user the correct answer to the question, not, as usual in Information Retrieval, IR, systems, a set of documents where likely the answer can be found.

Introduction

- Many NLP sub-tasks are involved in QA and many approaches can be followed for approaching these tasks.
- [Weston at al, 2015] have proposed a framework and a set of synthetic tasks for the goal of helping to develop learning algorithms for text understanding and reasoning and applying them to QA. The goal is to categorize different kinds of questions into skill sets, which become their tasks.
- The tasks are publicly available at http://fb.ai/babi
- Source code to generate the tasks is available at https://github.com/facebook/bAbI-tasks.

Introduction

- 1. Single Supporting Fact
  - Mary went to the bathroom, Where is Mary?
- 2. Two Supporting Facts
  - John is in the playground. John picked up the football. Where is the football?
- 3. Three Supporting Facts
  - John picked up the apple. John went to the office. John went to the kitchen. Where was the apple before the kitchen?
- 4. Two Argument Relations
  - The office is north of the bedroom. The bedroom is north of the bathroom. What is the bedroom north of?

Introduction

- 5. Three Argument Relations
  - Mary gave the cake to Fred. Fred gave the cake to Bill. Jeff was given the milk by Bill. Who did Fred give the cake to?
- 6. Yes/No Questions
  - John moved to the playground. Is John in the playground?
- 7. Counting
  - Daniel picked up the football. Daniel dropped the football. Daniel got the milk. Daniel took the apple. How many objects is Daniel holding?
- 8. Lists/Sets
  - Daniel picks up the football. Daniel drops the newspaper. Daniel picks up the milk. What is Daniel holding?

Introduction

- 9. Simple Negation
  - Fred is no longer in the office. Sandra is in the garden. Is Fred in the office?
- 10. Indefinite Knowledge
  - John is either in the classroom or the playground. Is John in the classroom?
- 11. Basic Coreference
  - Daniel was in the kitchen. Then he went to the studio. Where is Daniel?
- 12. Conjunction
  - Mary and Jeff went to the kitchen. Then Jeff went to the park. Where is Mary?

Introduction

- 13. Compound Coreference
  - Daniel and Sandra journeyed to the office. Then they went to the garden. Where is Daniel?
- 14. Time Reasoning
  - In the afternoon Julie went to the park. Yesterday Julie was at school. Julie went to the cinema this evening. Where did Julie go after the park?
- 15. Basic Deduction
  - Sheep are afraid of wolves. Cats are afraid of dogs. Mice are afraid of cats. Gertrude is a sheep. What is Gertrude afraid of?
- 16. Basic Induction
  - Lily is a swan. Lily is white. Greg is a swan. What color is Greg?

Introduction

- 17. Positional Reasoning
  - The triangle is to the right of the blue square. The red square is on top of the blue square. The red sphere is to the right of the blue square. Is the red square to the left of the triangle?
- 18. Size Reasoning
  - The football fits in the suitcase. The suitcase fits in the cupboard. The box is smaller than the football. Will the box fit in the suitcase?
- 19. Path Finding
  - The kitchen is north of the hallway. The bathroom is west of the bedroom. The den is east of the hallway. The office is south of the bedroom. How do you go from den to kitchen?
- 20. Agent's Motivations
  - John is hungry. John goes to the kitchen. John grabbed the apple there. Daniel is hungry. Where does Daniel go?

Introduction

- QA systems can be viewed as a natural extension of IR systems. In IR systems user information needs are expressed through a query, usually consisting on a set of keywords (sometimes more complex and expressive query languages can be used).
- The query is then used for retrieving from a dataset (a collection of documents, the whole Web, a domain restricted collection, a corporative textual database, …).
- The output of IR consists of a, sometimes ranked, set of documents retrieved from the dataset.

Introduction

- If the query is well formulated, a good IR system should retrieve in its best ranked documents those satisfying the user information needs. In QA systems, the query consists of a NL question (some systems use instead several forms of restricted NL) and the output of the system is not a set of likely relevant documents but the exact answer to the query.

Introduction

- Instead of allowing full NL expressivity, some NL applications, basically those heavily based on human-computer interaction, as QA, dialog-based systems, and so, prefer to work with limited forms of NL, the so called Controlled NL, CNL. Examples of these systems are SQUALL, [Ferré, 2013], AquaLog, and ORAKEL.
- Using CNL results on more robust interfacing, reduces the ambiguity inherent to NL and improves the parsing performance of the systems. There exist generic CNL but also domain-restricted CNL.

Introduction

- Currently, CNL systems are applied only to English, in fact the term Controlled English is frequently used for referring to these systems, although extending their capabilities to other languages does not seem to be, a priori, difficult.
- CNL implies several tasks:
  - Defining a grammar for a controlled language
  - Linguistic Engineering for building, testing, and maintaining the grammars
  - Tuning the grammars to new domains
  - Parsing

A bit of History of QA

- The origin of QA can be found in the eighties of last century, with the development of many NLI to computer applications, especially NLI to databases. The term QA was started to be used in the framework of QA tracks within TREC challenges (starting with TREC-8 in 1999).
- Conventional IR systems use basically statistical approaches, QA systems use, increasingly, as evolving towards more complex questions, NLP techniques, both for processing the question and for extracting the answer.

A bit of History of QA

- Some QA systems accessible through the Web are START (MIT) , AnswerBus , Webclopedia (ISI) , Ask (before AskJeeves) , LCC , PowerAnswer , IBM’s Watson , or Wolfram Alpha .
- Modern QA systems started with QA tracks in TREC contests (from 1999 to 2007), acquiring later a multilingual dimension in the framework of CLEF (from 2003) challenges, and currently included into the framework of TAC (from 2008).

A bit of History of QA

- Related disciplines are, obviously, IR and other closer disciplines as Information routing, filtering, harvesting, etc.) and also the Answer Finding, able to discover in a collection of question/answer pairs (as FAQ), the questions closest to the original one for retrieving the corresponding answer. Also related, although more distant are disciplines as Paraphrase detection, Textual Entailment, Information Integration, Knowledge Base Population, organized in TAC challenges into two tasks: Slot Filling and Entity Linking.

A bit of History of QA

- Initially QA was limited to Factoid Questions (Factoid QA), where the questions consisted on asking for a fact.
- Answering to these questions was not specially challenging because an assertive formulation of the answer is likely to be found in the collection.
  - For instance, for the question "Where was Barack Obama born", the assertion "Obama was born in Hawaii" probably occurs in the collection.
- One of the lines of research in QA was to increase the complexity and expressivity of the questions.

General QA systems

- Initially Conventional QA systems use to be structured into four modules (or steps):
  - Question Processing
  - IR of relevant documents
  - IR of relevant passages (fragments)
  - Answer Extraction.
- So, the language technologies involved have to cover these 4 tasks:
  - Linguistic analysis of questions using general or specific parsers (and grammars). This includes the definition of an appropriate tagset for classifying the questions (Question Type, QT)
  - IR engines (general or specific for the task)
  - IE techniques for extracting the answer

General QA systems

- For “simple” and “factoid” we mean that the question asks for a fact, the formulation of the question is simple (there are no additional constraints) and the answer can be retrieved from a single document with no additional processing. Most of the available QA systems satisfy this definition.

Factoid QA

Some examples of Factual Questions from TREC 8 Contest

How much folic acid should an expectant mother get daily?
Who invented the paper clip?
What university was Woodrow Wilson president of?
Where is Rider College located?
Name a film in which Jude Law acted.
Where do lobsters like to live?
Who was Picasso?

Beyond Factoid QA

- One of the lines of research in QA was to increase the complexity and expressivity of the questions. Some of these lines are the following:
  - Why QA, is the task of retrieving answers for a given Why Question, as “Why are tsunamis generated?”. [Oh et al, 2013], and [Oh et al, 2016] are excellent systems facing this task.
  - An in depth analysis of causality, causal relations, causal inference is needed.

Beyond Factoid QA

- …
  - Definitional QA, DQA, where the answer probably has to be synthesized from partial pieces of information extracted from several documents. Definitions can refer to people, “Who was Unamuno”, organizations, “what is the FAO”, or terms, “what is synonymy”. A good survey can be found in Rodrigo Alarcón’s thesis, [Alarcón , 2009].
  - As many DQA systems synthesize definitions from multiple sources NL Generation techniques, Slot Filling, Entity Linking and IE are widely used.

Beyond Factoid QA

- …
  - List QA, where the answer consists of a list. In some cases the whole list can be found in a document but frequently the members of the list have to be collected from different documents: “French president after world war II”
  - Linked questions: “Who was Picasso?”, “When and where He was born and dead?

Beyond Factoid QA

- …
  - Dialog based QA:
    - Time span of Georges Bush presidency
    - There are two USA presidents named Georges Bush, which do you refer to?
    - …
  - The form of dialog usually implies harder forms of anaphora.
    - Anaphora resolution
    - Dialog management
    - Dialog grammars
    - NL Generation

Beyond Factoid QA

- …
  - Time and space constrained questions: “Who was the second USA republican president after the Vietnam war?”. For answering this question a QA system should probably split the complex question into a set of related simpler questions:
    - “When did the Vietnam war ends?”, giving Date_1, “USA presidents after Date_1”, giving Person_1, Person_2, Person_3, …, “Party of Person_1”, …, “Party of Person_i”.
    - “Small cities, closed to Madrid having Romanic monuments”.
    - A good reference on time constrained QA is] Estela Saquete’s thesis, [Saquete, 2005].

Beyond Factoid QA

- …
  - In this use case the involved processors are more complex. Specially the question processing step is challenging (the process of decomposing the complex question into simpler ones, and the process of sequentializing the set of simple questions are usually challenging).
  - For temporal constraints a linguistic process including event and time tagging, a IE for recognizing temporal relations and possibly a temporal reasoner is needed.
  - For spacial constraints a spacial reasoner is needed. The identification of locations, GeoTagging, GeoDisambiguation, etc. are challenging tasks.

Beyond Factoid QA

- …
  - Opinion QA. Starting in TAC challenges and following the current trends in Opinion Mining and Sentiment Analysis, Opinion QA is currently subject of very active research.
    - Consider the question “Why do people enjoy Starbucks coffee? . Correctly answering this question implies not only locating a candidate answer but also check whether it is an opinion and whether its polarity is positive.
    - Other examples are “What Quevedo thought about Góngora?”, “Arguments pro and against arms control in USA”.
    - Alexandra Balahur thesis, [Balahur, 2012], is an excellent introduction to this topic.

Beyond Factoid QA

- …
  - Additional technologies are needed for dealing with this task:
    - The Question processing step is more complex because not only the QT and EAT have to be recognized but also the constraints related to opinion mining, sentiment analysis, polarity detection, …
    - In the answer extraction step some issues arise: i) classifying a detected assertion as informative, opinion, etc. ii) extracting its polarity, iii) detecting the subject of the opinion, etc.
    - Frequently the opinion information has to be extracted from social nets. Processing this kind of documents needs specialized linguistic processors.

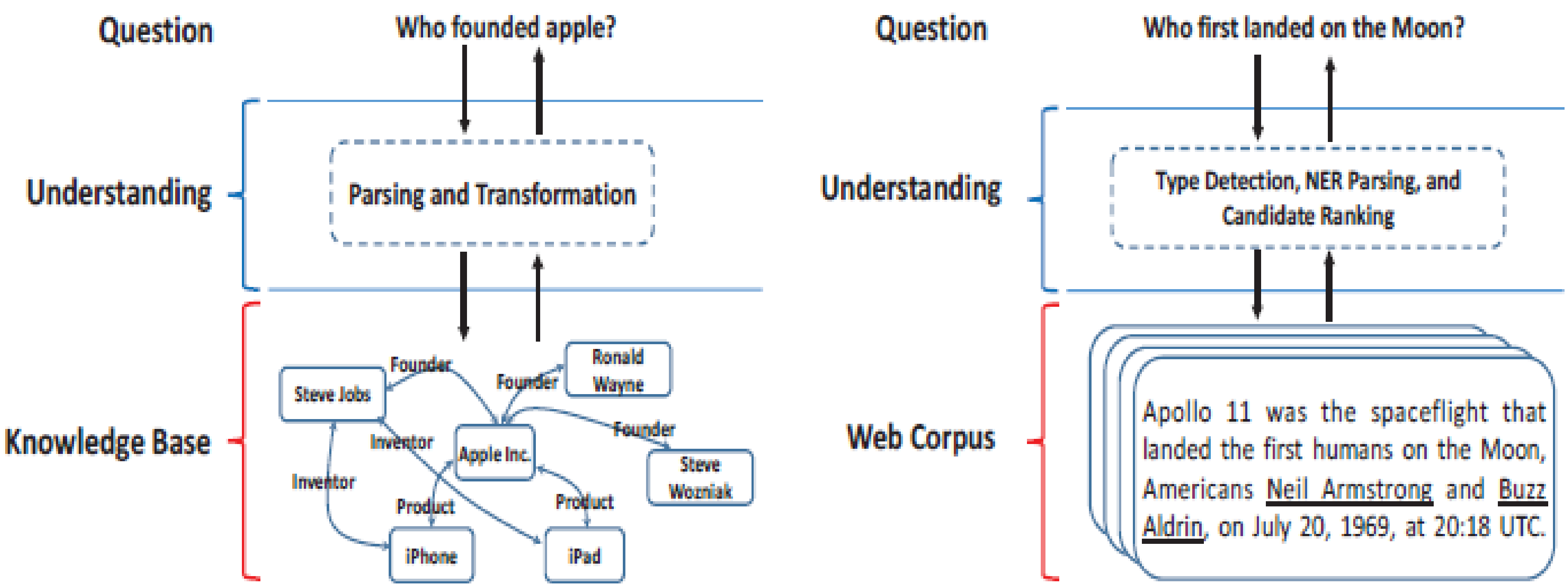
Beyond Factoid QA

- Another line of research in QA, somehow divergent from the above one, is to narrow the search space where the answer is expected to be found. We can found in this line:
  - QA for comprehension reading, where the questions are related with a document for checking the ability of the user to having understood the document content. Richardson et al. (2013) proposed the MCTest a set of 660 stories and associated questions intended for research on the machine comprehension of text. Each question requires the reader to understand different aspects of the story. QA4MRE challenges, organized in the framework of CLEF group the most interesting approaches to this type of QA.
  - QA for learning by reading, similar to the previous but in this case measuring the ability of the computer.

Beyond Factoid QA

- …
  - An important point is where the QA systems look for the answers of the questions. Basically two situation exist: Knowledge Base based QA (KB_QA) and Web based QA (WebQA).
  - A couple of downloadable state-of-the-art systems are: KB_QA: Sempre , ParaSempre, [Berant et al, 2013], [Berant, Liang, 2014] and WebQA: AskMSR+, [Tsai et al, 2015].

Beyond Factoid QA



(a) KB-based QA Systems (b) Web-based QA Systems

Beyond Factoid QA

- …
  - QA over domain ontologies, Ontology-based QA (ObQA). In this case the answers are looked up not in free text documents but in ontologies taking profit not only of the linguistic (terminological) data included into the ontology but also over their relations, properties, and inferential capabilities. An interesting example is Pythia, [Unger, Cimiano, 2011]. Pythia is based on an alignment between the question and a vocabulary aligned with the ontology. The process includes the semiautomatic generation of a grammar using LexInfo , a declarative model for lexicon-ontology interface.

Beyond Factoid QA

- …
  - In the framework of the Semantic Web there has been recently a huge growth of available open and closed domain resources. Many of these resources are included into the Linked Open Data (LOD) initiative. The most known and used resources in LOD are FreeBase and DBPedia as open domain LOD and BioPortal (medical and genomic) or LinkedGeoData (geographic) as closed domain LOD. QA systems using LOD as search space are referred as QALD and are the subject of this survey.

Beyond Factoid QA

- …
  - Somehow related to the preceding approach but in the opposite direction we can find the work of mapping SPARQL queries into their NL expressions (Verbalizing SPARQL). This task can provide us with useful training material for facing the mapping NL $\rightarrow$ SPARQL.
  - The work of Ngonga and colleagues, [Ngonga et al, 2013a], [Ngonga et al, 2013a], SPARQL2NL, is a nice example of this task.

DRQA

- …
  - Domain restricted QA, DRQA. Where both questions and search space are restricted to a given domain.
    - Many domains have been faced, geographic, tourism, economics, etc. Perhaps the domain object of the most applications is the medical. DRQA are applied to specific tasks and use domain specific lexicons, terminologies, KBs, ontologies, …. Search spaces are smaller and so approaches based on the redundancy of answers (as votation techniques) are useless. User’s requirements use to be high and system performance is more precision than recall oriented (no answer is better than a wrong answer). Questions and documents are challenging and frequently contain acronyms, non-textual content (tables, itemized lists, etc.), domain specific jargon, etc.

DRQA

- …
  - Domain restricted QA, DRQA. Where both questions and search space are restricted to a given domain.
    - thesis of Óscar Ferrández, [Ferrández, 2009], centered on the cinematographic domain, within the framework of the European project QALL-ME.
    - Within the Commercial Domain, but also related to LOD, we can find the Business to Client (B2C) scenario. Two interesting systems in this scenario are QALM, [Hallili et al, 2014], and SynchroBot, [Cabrio et al, 2015b].
    - More emphasis is posed on Semantic approaches due to the availability of domain specific semantic resources.
    - Some of the processors have to be tuned to the domain.
      - Domain Adaptation techniques

DRQA

- …
  - Clinical question answering (Clinical QA).
    - [Demner-Fushman et al, 2009]
    - Questions occurring in clinical situations could pertain to:
      - Information on particular patients
      - Data on health and sickness within the local population
      - Medical knowledge
      - Local information on doctors available for referral
      - Information on local social influences and expectation
      - Information on scientific, political, legal, social, management, and ethical changes affecting both how medicine is practiced and how doctors interact with individual patients

DRQA

- …
  - Clinical question answering (Clinical QA).
    - Some questions do not need NLP and can be answered directly by a known resource.
    - Questions about particular patients are currently answered by manually browsing or searching the EHR.
    - Answering these questions can be facilitated by summarization (which requires NLP if information is extracted from free-text fields) and visualization tools.
    - Facilitating access to medical knowledge by providing answers to clinical questions is an area of active NLP research.
    - The goal of clinical QA systems is to satisfy medical knowledge questions providing answers in the form of short action items supported by strong evidence.

CQA

- …
  - Community QA (CQA). In this scenario a member of the community formulates an initial query (a NL question) that triggers a thread of interventions of the community members that answer, refine, comment, the interventions of previous interventions. Members's interventions can be questions and answers related. CQA have been recently evaluated in the framework of SEMEVAL-2015, SEMEVAL-2016, SEMEVAL-2017. Both general purpose and topic-specific communities are growing in numbers for posting questions and obtaining direct answers in a short period of time.
    - Yahoo!Answers (Y!A), for example, provides a broad range of topics where as Stack-Overflow (SO), and Turbo Tax Live (TT) are quite focused and domain-specific.

CQA

- …
  - Community QA (CQA). In contrast to the traditional search engines such as Google, CQA services provide an alternative paradigm for seeking targeted information. [Zhang et al, 2014] is a good example of these kind of systems, see also [El Adlouni, et al, 2016] and [Nakov, et al, 2016, 2017]. The approach assumes that questions and answers share some common latent topics and are generated in a “question language" and “answer language" respectively following the topics. [Cong et al, 2008] presents an interesting system for facing the question detection and answer detection problems. [Xue et al, 2008] propose using retrieval models for detecting Q and A in Q&A archives (both FAQ archives and archives generated by CQA web services. The authors use as main source for learning the Wondir collection.

CQA

- …
- Community QA (CQA).
  - In the case of CQA the basic tasks are Question processing and Answer extraction, the IR steps are less important.
  - CQA implies semantic comparisons of elements (queries, answers, comments, etc.) of the query streams. Many similarity (and distance) computations have been used for the task.

Q&A over LOD

- Open-domain Question Answering
- answer question on any topic
  - query a KB with natural language
  - Semantic Representation = KB entities + relations

Q&A over LOD

- Question Answering with Subgraph Embeddings
  - A. Bordes, S. Chopra & J. Weston. EMNLP, 2014
- Paraphrase-Driven Learning for Open Question Answering
  - A. Fader, L. Zettlemoyer & O. Etzioni. ACL, 2013
- Open Question Answering Over Curated and Extracted Knowledge Bases
  - A. Fader, L. Zettlemoyer & O. Etzioni. KDD, 2014
- Large-scale Semantic Parsing without Question-Answer Pairs
  - S. Reddy, M. Lapata & M. Steedman. TACL, 2014.

Q&A over LOD

- Resources useful for QALD systems.
  - A list of available resources (with links) can be obtained from QALD site. I include next some of them not very known (other frequently used resources are Stanford CoreNLP, LingPipe, OpenNLP , Senna, MATE, WS4J , ClearNLP , and many others:
    - English lexicon for DBpedia 3.8 (in lemon7 format)
      - http://lemon-model.net/lexica/dbpedia_en/
    - PATTY (collection of semantically-typed relational patterns)
      - http://www.mpi-inf.mpg.de/yago-naga/patty/
    - DBpedia Spotlight
      - http://spotlight.dbpedia.org
    - FOX (Federated Knowledge Extraction Framework)
      - http://fox.aksw.org

Q&A over LOD

- Resources useful for QALD systems.
  - Wikipedia Miner
    - http://wikipedia-miner.cms.waikato.ac.nz/
  - WS4J (Java API for several semantic relatedness algorithms)
    - https://code.google.com/p/ws4j/
  - SecondString (string matching)
    - http://secondstring.sourceforge.net
  - PPDB (The Paraphrase Database)
    - http://www.cis.upenn.edu/~ccb/ppdb/
  - Wondir collection (about 1M Q&A pairs collected by Wondir)
    - http://wondir.com

Q&A over LOD

- Difficulties for dealing with RDF datasets
- Mapping questions into SPARQL is not an easy task. Several problems arise and have to be faced:
  - Different namespaces coexist in DBpedia, some of them belonging to DBpedia itself, and others corresponding to links from DBpedia to other ontologies, as Yago.
  - For instance, looking for the generic term 'Mountain' we find 217 categories in Yago namespace (e.g. http://dbpedia.org/class/yago/Mountain109359803), 10 DBpedia properties (e.g. http://dbpedia.org/ontology/highestMountain), and 6 DBpedia ontology categories (e.g. http://dbpedia.org/ontology/Mountain)

Q&A over LOD

- …
  - Lack of coherence in the nomenclature used for naming DBpedia entries (classes, properties and instances). Use of lower/upper case, singular/plural forms, abbreviations, order of simple components of the multi-word expressions, inclusion of parenthesis, underscores, and other orthographic marks is rather arbitrary or at least difficult to interpret. The following properties (among many others) were found in DBpedia when looking for number of members:
    - http://dbpedia.org/property/memberNo
    - http://dbpedia.org/property/members
    - http://dbpedia.org/property/member
    - http://dbpedia.org/property/numMembers
    - http://dbpedia.org/property/membersNumbers
    - http://dbpedia.org/property/noOfMembers

Q&A over LOD

- …
  - The habitual clash when mapping terms of the NL expression of the question into terms of the ontology is obviously present. The habitual problems of polysemy (a term of the question can be mapped to many terms of the ontology : classes, properties, and instances) and synonymy (an ontology term can be referred by different question terms) frequently occur.
  - The directionality of the relations in the ontology is not always clear. For instance, it is not clear whether http://dbpedia.org/property/mayor/ links a city to a person or a person to a city.
  - Depending on the Question Type (QT), the EAT and the complexity of the question, partially reflected in the constraints provided by the Question Processing module, resolving the mapping can be more or less difficult.

Q&A over LOD

- Question Answering with Subgraph Embeddings
  - Training data
    - Freebase is automatically converted into Q&A pairs closer to expected language structure than triples

QALD contests

- QALD5
  - http://greententacle.techfak.uni-bielefeld.de/~cunger/qald/index.php?x=challenge&q=5
  - Given a natural language question or keywords, retrieve the correct answer(s) from a repository containing both RDF data and free text.
  - QALD1, …, QALD6

QALD contests

- < question id ="272"
- answertype =" resource “
- aggregation =" true “
- onlydbo =" true “
- hybrid =" false " >
- < string lang =" en " > Which book has the most pages
- < string lang =" es " >¿Que libro tiene el mayor numero de paginas ?
- < keywords lang =" en " > book , the most pages
- < query >
  - PREFIX dbo : < http :// dbpedia . org / ontology / >
  - PREFIX rdf : < http :// www . w3 . org /1999/02/22 - rdf - syntax - ns # >
  - SELECT DISTINCT ?uri
  - WHERE {
    - ?uri rdf : type dbo : Book .
    - ?uri dbo : numberOfPages ? n }
  - ORDER BY DESC (?n) OFFSET 0 LIMIT 1

QALD contests

Questions solved by all systems	Questions solved by no systems
What is the capital of Canada?	Give me all members of Prodigy.
Who is the governor of Wyoming?	Does the new Battlestar Galactica series have more episodes than the old one?
What is the birth name of Angela Merkel?	Show me all songs from Bruce Springsteen released between 1980 and 1990.
How many employees does Google have?	Give me all B-sides of the Ramones.

QALD contests

CASIA

- QALD3, [He et al, 2013], [He et al, 2014a], QALD4, [He et al, 2014b])
- implements a pipeline consisting of question analysis, resource mapping and SPARQL generation.
- In QALD3, first transforms natural language questions into a set of query triples of the form <subject, predicate, object>, based on a shallow and deep linguistic analysis.
  - For instance, for the question “Who are the parents of the wife of Juan Carlos I?”, (id=67, test set), two query triples are produced:
  - <?who, are the parents of, ?wife>
  - <?wife, the wife of, Juan Carlos I >.

QALD contests

CASIA

- Second, it instantiates these query triples with corresponding resources from DBpedia, resulting in ontology triples. For every phrase in query triple, the corresponding resource in DBpedia has to be identified. For different types of resource, different techniques and resources are used. The output of this step is a list of ontology triples. One query triple will generate several possible ontology triples. Possible triples of the aforementioned example are:
  - <?Person, rdf:type, dbc:Person>
  - <?Person, dbp:parent, ?wife>
  - <?wife, dbo:spouse, dbr:Juan Carlos I Of Spain>

QALD contests

CASIA

- The most challenging part of this process is the mapping into properties. For doing so CASIA uses PATTY in order to get the relation patterns. Based on the ontology triples and question type, SPARQL queries are constructed. In the example a possible query is shown following.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX dbo: <http://dbpedia.org/ontology/>
PREFIX dbp: <http://dbpedia.org/resource/>
SELECT DISTINCT ?URL
WHERE {
    ?URL rdf:type dbo:Person.
    ?URL dbo:parent ?wife.
    ?wife dbo:spouse dbp:Juan_Carlos_I_of_Spain.
}
```

QALD contests

CASIA

- Finally, the candidate queries are validated and ranked, and the best query is selected. As a result, each step can be subject to global optimization. It makes use of the Stanford NER, the PATTY and ReVerb resources, as well as thebeast tool for weight learning and MAP inferencing.

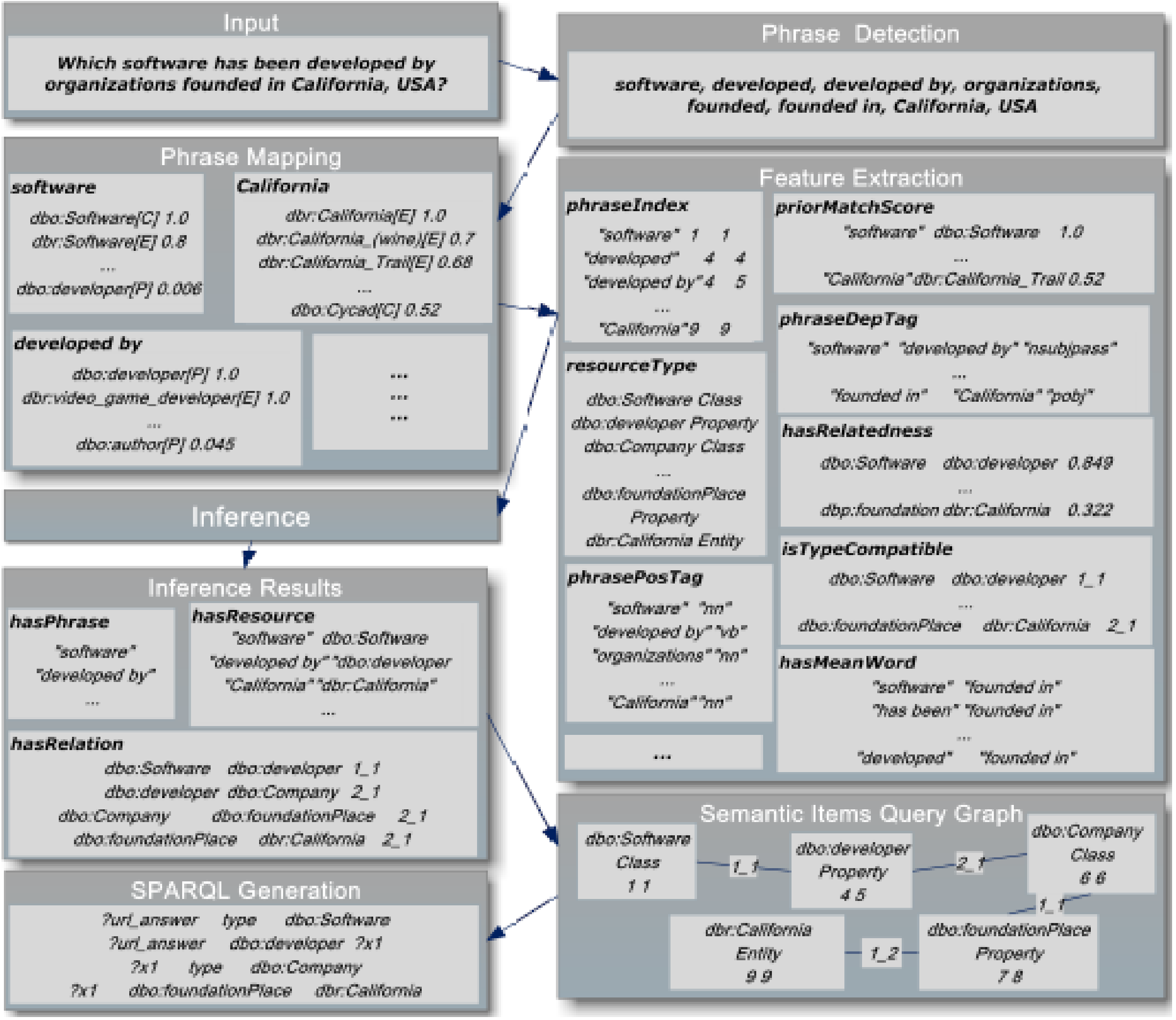
QALD contests

CASIA

- Several improvements have been introduced into QALD4.
- First the recognized phrases are mapped into embedded representations using word2vec in order to face the problem of lexical variation.
- Besides PATTY, for mapping phrases to properties, Anthony Fader's **ReVerb** is used.
- An additional step for disambiguating as much as possible ambiguities in phrase detection and mapping-phrase-to-semantic-item. This step consists in the resolution of these ambiguities and determine the relations among the mapped semantic items. **Markov Logic Networks, MLN**, are used for this purpose.

QALD contests

CASIA



QALD contests

CASIA relation types

Type	Example	Question
1_1	*dbo:height 1_1 dbr:Michael_Jordan*	*How tall is Michael Jordan?*
1_2	*dbo:River 1_2 dbo:crosses*	*Which river does the Brooklyn Bridge cross?*
2_1	*dbo:creator 2_1 dbr:Walt_Disney*	*Which spaceflights were launched from Baikonur?*
2_2	*dbo:birthPlace 2_2 dbo:capital*	*Which actors were born in the capital of American?*

QALD contests

CASIA Observed predicates

Describing the attributes of phrases and relation between two phrases	
phraseIndex(p, i, j)	The start and end position of phrase *p* in question.
phrasePosTag(p, pt)	The POS tag of head word in phrase *p*.
phraseDepTag(p, q, dt)	The dependency path tags between phrase *p* and *q*.
phraseDepOne (p, q)	If there is only one tag in the dependency path, the predicate is true.
hasMeanWord (p, q)	If there is any one meaning word in the dependency path of two phrases, the predicate is true.
Describing the attributes of semantic item and the mapping between phrase and semantic item	
resourceType(r, rt)	The type of semantic item r. Types of semantic items include *Entity*, *Class* and *Property*
priorMatchScore(p, r, s)	The prior score of phrase *p* mapping to semantic item *r*.
Describing the attributes of relation between two semantic items in knowledge base	
hasRelatedness(p, q, s)	The semantic coherence of semantic items.
isTypeCompatible(p, q, rr)	If semantic items *p* is type-compatible with semantic items *q*, the predicate is true.
hasQueryResult(s, p, o, rr1, rr2)	If the triple pattern consisting of semantic items *s*, *p*, *o* and relation types *rr1*, *rr2* have query results, the predicate is true.

QALD contests

CASIA Boolean formulas

hf1	$hasPhrase(p) \Rightarrow hasResource(p,\_)$
hf2	$hasResource(p,\_) \Rightarrow hasPhrase(p)$
hf3	$\lvert hasResource(p,\_)\rvert \leq 1$
hf4	$!hasPhrase(p) \Rightarrow !hasResource(p,r)$
hf5	$hasResource(\_,r) \Rightarrow hasRelation(r,\_,\_) \vee hasRelation(\_,r,\_)$
hf6	$\lvert hasRelation(r1,r2,\_)\rvert \leq 1$
hf7	$hasRelation(r1,r2,\_) \Rightarrow hasResource(\_,r1) \wedge hasResource(\_,r2)$
hf8	$phraseIndex(p1,s1,e1) \wedge phraseIndex(p2,s2,e2) \wedge overlap(s1,e1,s2,e2) \wedge hasPhrase(p1) \Rightarrow !hasPhrase(p2)$
hf9	$resourceType(r,\text{“}Entity\text{”}) \Rightarrow !hasRelation(r,\_,\text{“}2\_1\text{”}) \wedge !hasRelation(r,\_,\text{“}2\_2\text{”})$
hf10	$resourceType(r,\text{“}Entity\text{”}) \Rightarrow !hasRelation(\_,r,\text{“}2\_1\text{”}) \wedge !hasRelation(r,\_,\text{“}2\_2\text{”})$
hf11	$resourceType(r,\text{“}Class\text{”}) \Rightarrow !hasRelation(r,\_,\text{“}2\_1\text{”}) \wedge !hasRelation(r,\_,\text{“}2\_2\text{”})$
hf12	$resourceType(r,\text{“}Class\text{”}) \Rightarrow !hasRelation(\_,r,\text{“}2\_1\text{”}) \wedge !hasRelation(r,\_,\text{“}2\_2\text{”})$
hf13	$!isTypeCompatible(r1,r2,rr) \Rightarrow !hasRelation(r1,r2,rr)$

QALD contests

CASIA Weighted formulas

sf1	$priorMatchScore(p, r, s) \Rightarrow hasPhrase(p)$
sf2	$priorMatchScore(p, r, s) \Rightarrow hasResource(p)$
sf3	$phrasePosTag(p, pt+) \wedge resourceType(r, rt+) \Rightarrow hasResource(p, r)$
sf4	$phraseDepTag(p1, p2, dp+) \wedge hasResource(p1, r1) \wedge hasResource(p2, r2) \Rightarrow hasRelation(r1, r2, rr+)$
sf5	$phraseDepTag(p1, p2, dp+) \wedge hasResource(p1, r1) \wedge hasResource(p2, r2) \wedge !hasMeanWord(p1, p2) \Rightarrow hasRelation(r1, r2, rr+)$
sf6	$phraseDepTag(p1, p2, dp+) \wedge hasResource(p1, r1) \wedge hasResource(p2, r2) \wedge phraseDepOne(p1, p2) \Rightarrow hasRelation(r1, r2, rr+)$
sf7	$hasRelatedness(r1, r2, s) \wedge hasResource(\_, r1) \wedge hasResource(\_, r2) \Rightarrow hasRelation(r1, r2, \_)$
sf8	$hasQueryResult(r1, r2, r3, rr1, rr2) \Rightarrow hasRelation(r1, r2, rr1) \wedge hasRelation(r2, r3, rr2)$

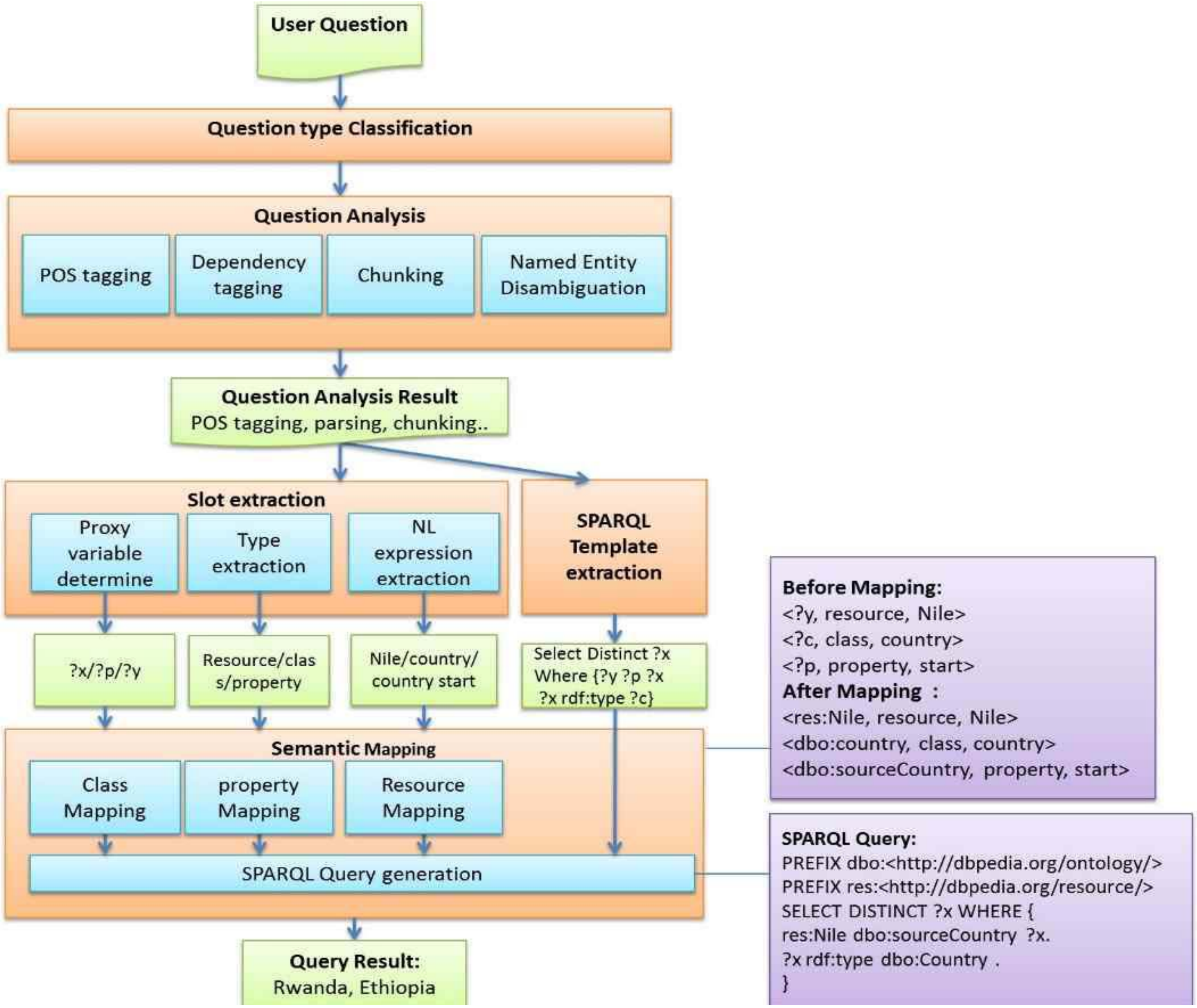
QALD contests

ISOFT

- QALD-4, [Park et al, 2014], QALD-5, [Park et al, 2015a). In its participation on QALD-4 ISOFT follows a template-based approach for transforming natural language questions into SPARQL queries. Based on a linguistic analysis of the input question, query templates and slots are determined, which are then filled by searching for appropriate concepts in the knowledge base, based on string similarity and Explicit Semantic Analysis (**ESA**) for mapping predicates in the user NL question to predicate uniform resource identifiers (URIs) in the KB. The analysis of the question is carried out using **ClearNLP.** For NE (i.e., resource) disambiguation, they used **AIDA.**

QALD contests

ISOFT



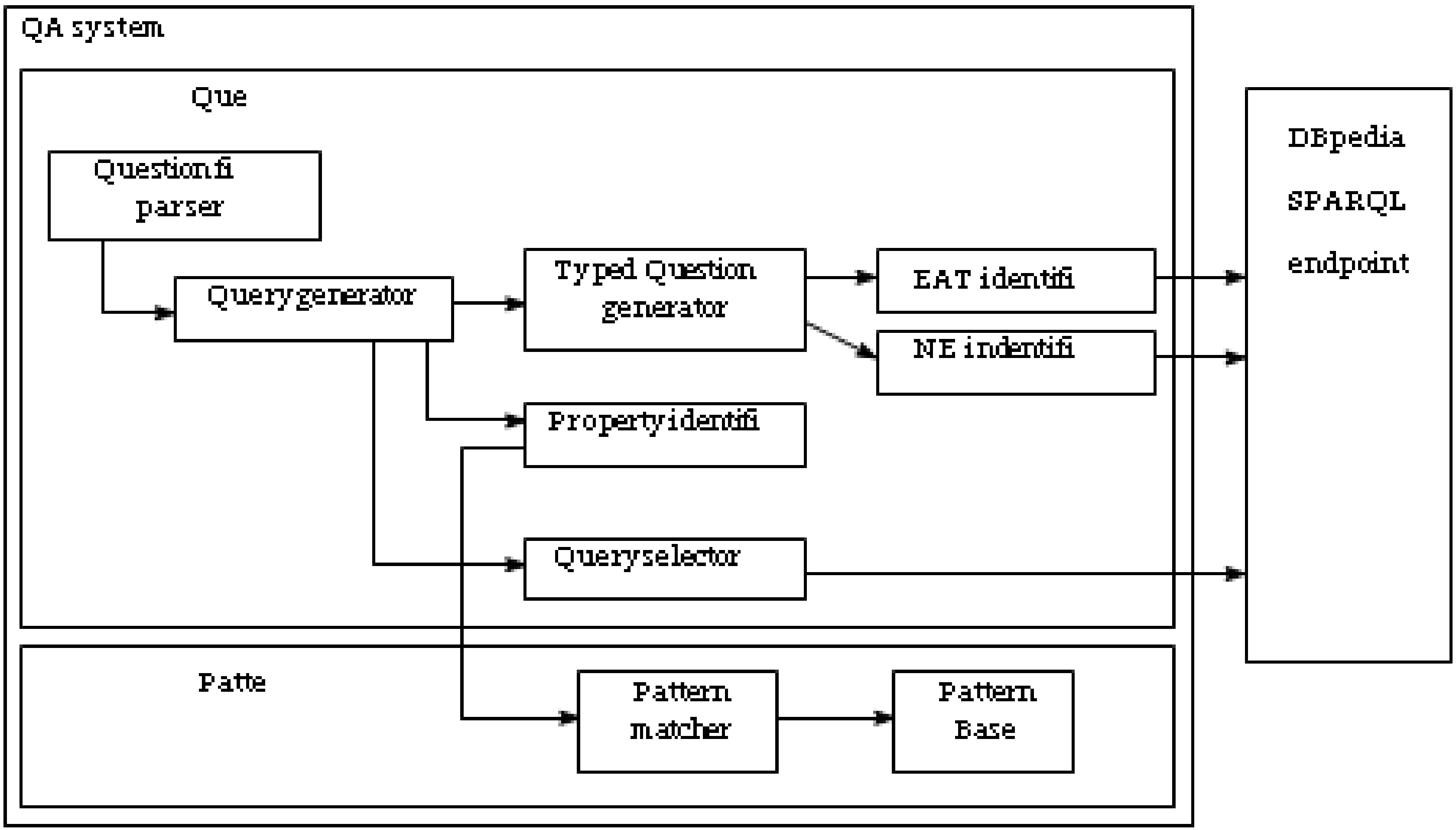
QALD contests

QAKIS

- [Cabrio at al, 2012]. The problem of question interpretation is addressed as the automatic identification of the set of relevant relations between entities in the natural language input question, matched against a repository of automatically collected relational patterns (i.e. the WikiFramework repository). Such patterns represent possible lexicalizations of ontological relations, and are associated to a SPARQL query derived from the linked data relational patterns. WP is used as the source of free text for the automatic extraction of the relational patterns, and DBpedia as the linked data resource to provide relational patterns and to be queried using a natural language interface. Goal of the WikiFramework is to establish a robust methodology to collect relational patterns in several languages, for the relations defined in DBpedia ontology. For example, an instance of the crosses relation is:
- <dbr:Golden_Gate_Bridge, dbo:crosses, dbr:Golden_Gate>

QALD contests

QAKIS



QALD contests

QAKIS *Examples of WikiFramework patterns*

Relation	Patterns
spouse	Person wife Person Person married Person Person husband Person
crosses	Bridge spanning River Bridge bridge River Bridge crossing River

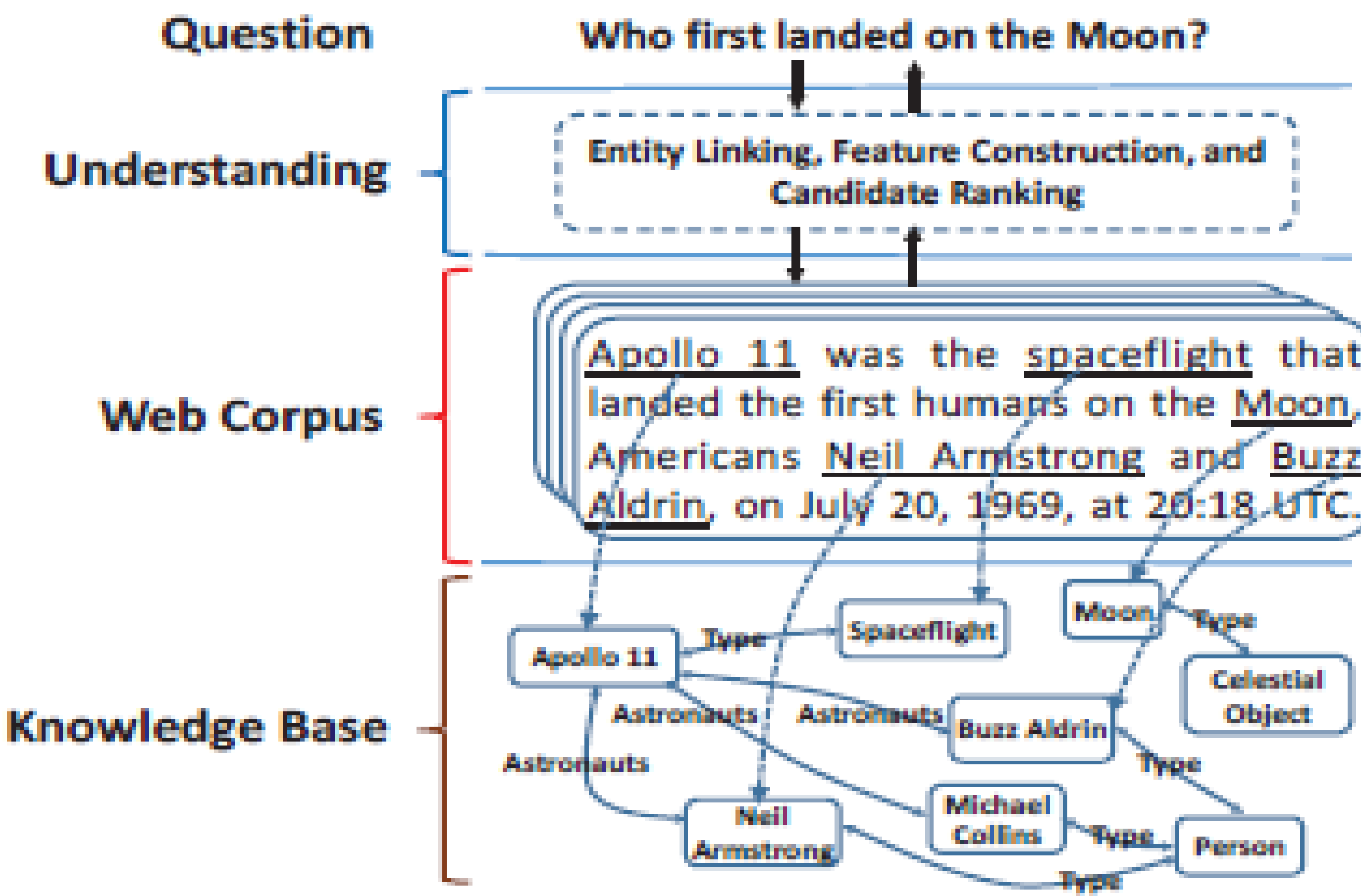
QALD contests

QuASE

- [Sun, 2015], i.e., question answering via semantic enrichment. The system extends the traditional Web-based QA system by linking answer candidates in the search texts to a knowledge base.
- Specifically, given a question, QuASE first selects a set of most prominent sentences from web resources. Then from those sentences, EL tools are used to detect answer candidates and link them to entities in Freebase. Once each answer candidate is mapped to the corresponding entity in Freebase, abundant information, such as their description texts and Freebase types, can be utilized for feature generation and modeling. A ranking algorithm is subsequently trained based on such features to rank correct answers as top choices.

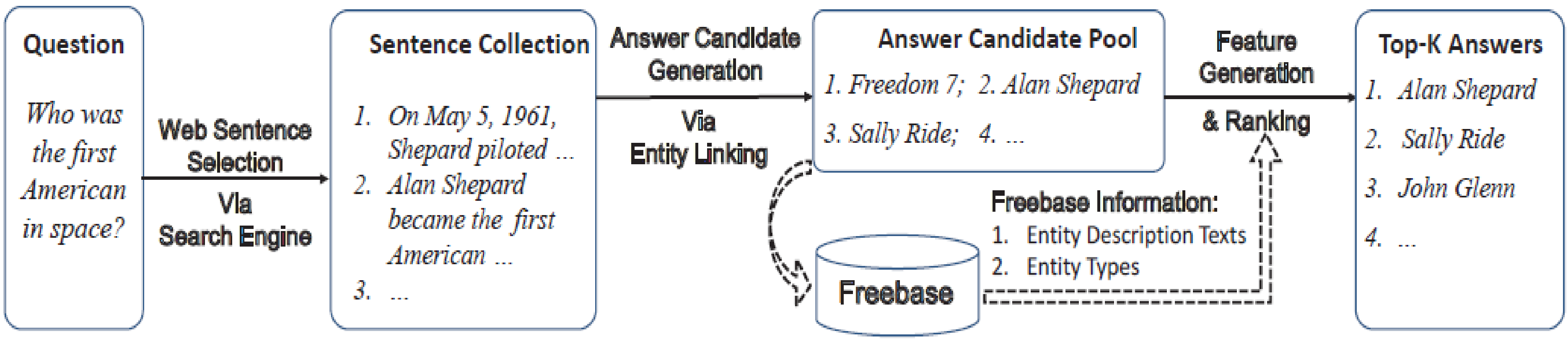
QALD contests

QuASE



QALD contests

QuASE Example of QuASE processing



QALD contests

Scalewelis

- [Guyonvarch, Ferré, 2013], is a faceted search system that guides the user through the search for an answer. Starting from an initial SPARQL query, facets are created for the first 1,000 results retrieved by that query, consisting of the classes the results belong to as well as properties that relate the results to other entities in the dataset. Scalewelis connects to SPARQL endpoints and uses partial result sets in order to scale to large datasets. The user's selection of a facet is then used to refine the query until the answer is found.

QALD contests

Scalewelis

Initial query

```
SELECT DISTINCT ?class WHERE {
    [] a ?class . }
```

Query 1: Computation of the partial results, limited to 1,000 entities

```
SELECT DISTINCT ?result WHERE {
    <Pattern>} LIMIT 1000
```

Query 2: Computation of class facets from the partial results

```
SELECT DISTINCT ?class WHERE {
    VALUES (?result) {res1 ... resN}
    ?result a ?class }
```

Query 3: Computation of property facets from the partial results

```
SELECT DISTINCT ?prop WHERE {
    VALUES (?result) {res1 ... resN}
    ?result ?prop [] }
```

Query 4: Computation of inverse property facets from the partial results

```
SELECT DISTINCT ?invProp WHERE {
    VALUES (?result) {res1 ... resN}
    [] ?invProp ?result }
```

QALD contests

Antoine Bordes

- [Bordes, 2014a], converts questions to embeddings which require no pre-defined grammars or lexicons and can query any KB independent of its schema. He focuses on answering simple factual questions on a broad range of topics, more specifically, those for which single KB triples stand for both the question and an answer (of which there may be many).
- For example, <parrotfish.e, live-in.r, southern-water.e> stands for “What is parrotfish’s habitat? For learning he uses Weak Supervision. The model, loosely based on [Fader et al, 2013] is able to take advantage of noisy and indirect supervision by:
  - automatically generating questions from KB triples and treating this as training data
  - Supplementing this with a data set of question collaboratively marked as paraphrases but with no associated answers.

QALD contests

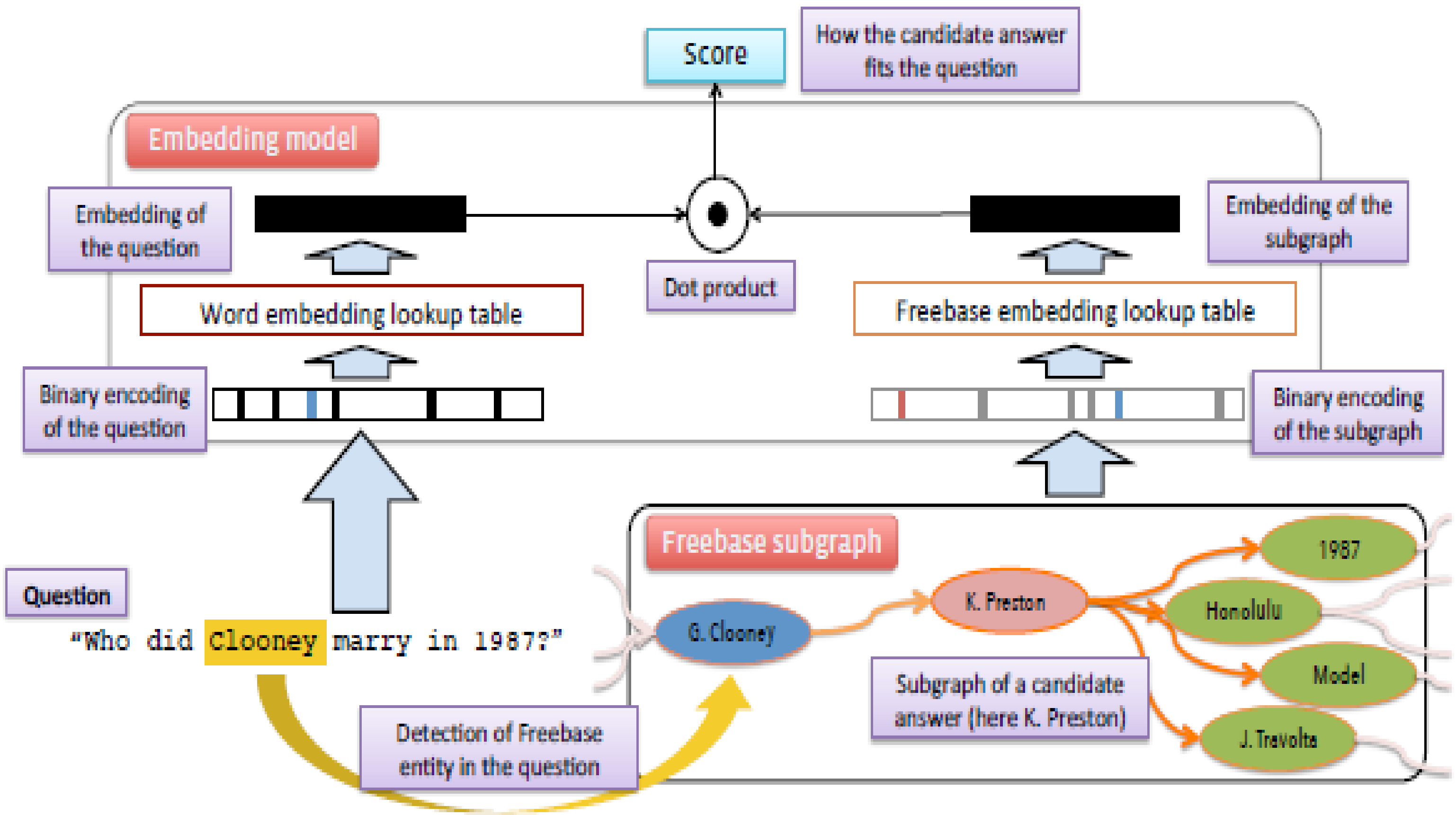
Antoine Bordes Patterns for generating questions from ReVerb triples

KB Triple	Question Pattern
(?, r, e)	*who r e ?*
(?, r, e)	*what r e ?*
(e, r, ?)	*who does e r ?*
(e, r, ?)	*what does e r ?*
(?, r, e)	*what is the r of e ?*
(?, r, e)	*who is the r of e ?*
(e, r, ?)	*what is r by e ?*
(?, r, e)	*who is e's r ?*

KB Triple	Question Pattern
(?, r, e)	*what is e's r ?*
(e, r, ?)	*who is r by e ?*
(e, r-in, ?)	*when did e r ?*
(e, r-on, ?)	*when did e r ?*
(e, r-in, ?)	*when was e r ?*
(e, r-on, ?)	*when was e r ?*
(e, r-in, ?)	*where was e r ?*
(e, r-in, ?)	*where did e r ?*

Q&A over LD

Bordes et al, 2014



Q&A over LD

Anthony Fader PHD 2014

- Identifying Relations for Open Information Extraction, which focuses on acquiring open-domain knowledge using a novel IE technique, [Fader et al, 2011]. For instance from the sentence “Windsor also does business in Cuban cigars, which are banned in the US." an Open IE system might extract two triples: (Windsor, does business in, Cuban cigars) and (Cuban cigars, banned in, US). The thesis proposes a system, **REVERB**, for doing this task.

Q&A over LD

Anthony Fader PHD 2014

- Paraphrase-Driven Learning for Open QA, which focuses on robust question interpretation using the paraphrase information available on WikiAnswers, [Fader et al, 2013]. The thesis presents a system, **PARALEX**, that learns a robust question-interpretation function from the paraphrase data available on WikiAnswers. Paralex uses ReVerb as a source of knowledge and is the first system to perform Open QA over an extracted knowledge base. Paralex uses a novel learning algorithm that generalizes from millions of paraphrase clusters.

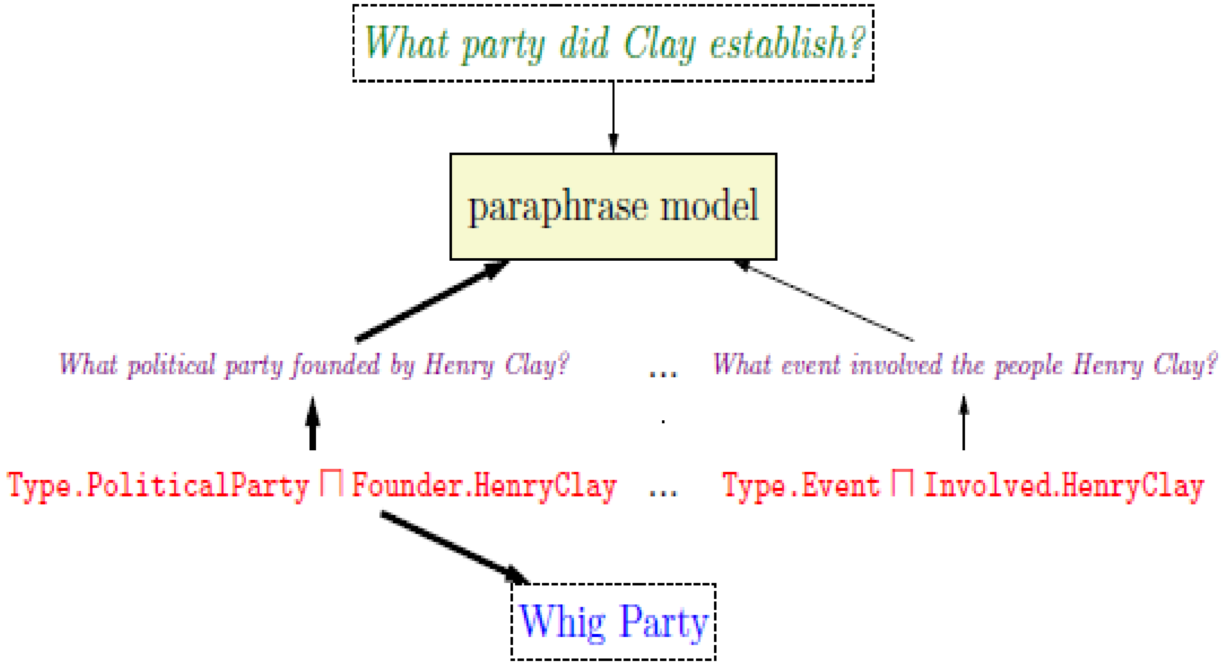
Q&A over LD

Anthony Fader PHD 2014

- Open QA Over Curated and Extracted Knowledge Bases, which focuses on combining knowledge from multiple sources and improving the accuracy of question interpretation, [Fader et al, 2014]. OQA overcomes the problems of Paralex decomposing the full QA problem into smaller problems that are easier to solve.
- PARALEX
  - large monolingual parallel corpora, containing 18 million pairs of question paraphrases from wikianswers.com, which were tagged as having the same meaning by users.

Q&A over LD

Fader thesis, 2013



Q&A over LD

Fader thesis, 2013
Sample of REVERB extractions containing the strings “legal”, “illegal”, or, “banned”

Argument 1	Relation Phrase	Argument 2
Gambling	is banned in	Islam
Foie gras	is not banned in	California
cocaine use	was legal in	the United States
Large breeds	were banned in	Beijing
Prostitution	is legal in	Amsterdam
Independent unions	are illegal in	China
Mah Jong	was completely banned in	China
Pyramid selling	is illegal in	Australia
GTA4	is banned in	the UAE.
humor	is illegal in	Poland
Hitch hiking	is illegal in	Oz
Dog and cat fur	should be banned in	Europe
same-sex marriages	were legalized in	California
common law marriages	are declared illegal in	England
Homosexuality	is illegal in	Mauritius
Cock fighting	was banned in	1849
the WOW	has been banned in	Manchester
Sabots	are illegal in	Colorado
gay marriage	became legal in	Massachusetts
poker	was recently legalized in	Catalonia
Slavery	is declared illegal in	the Oregon Country
Pornography	is legal in	Australia
Corporal punishment	is legal in	Wilkinson County

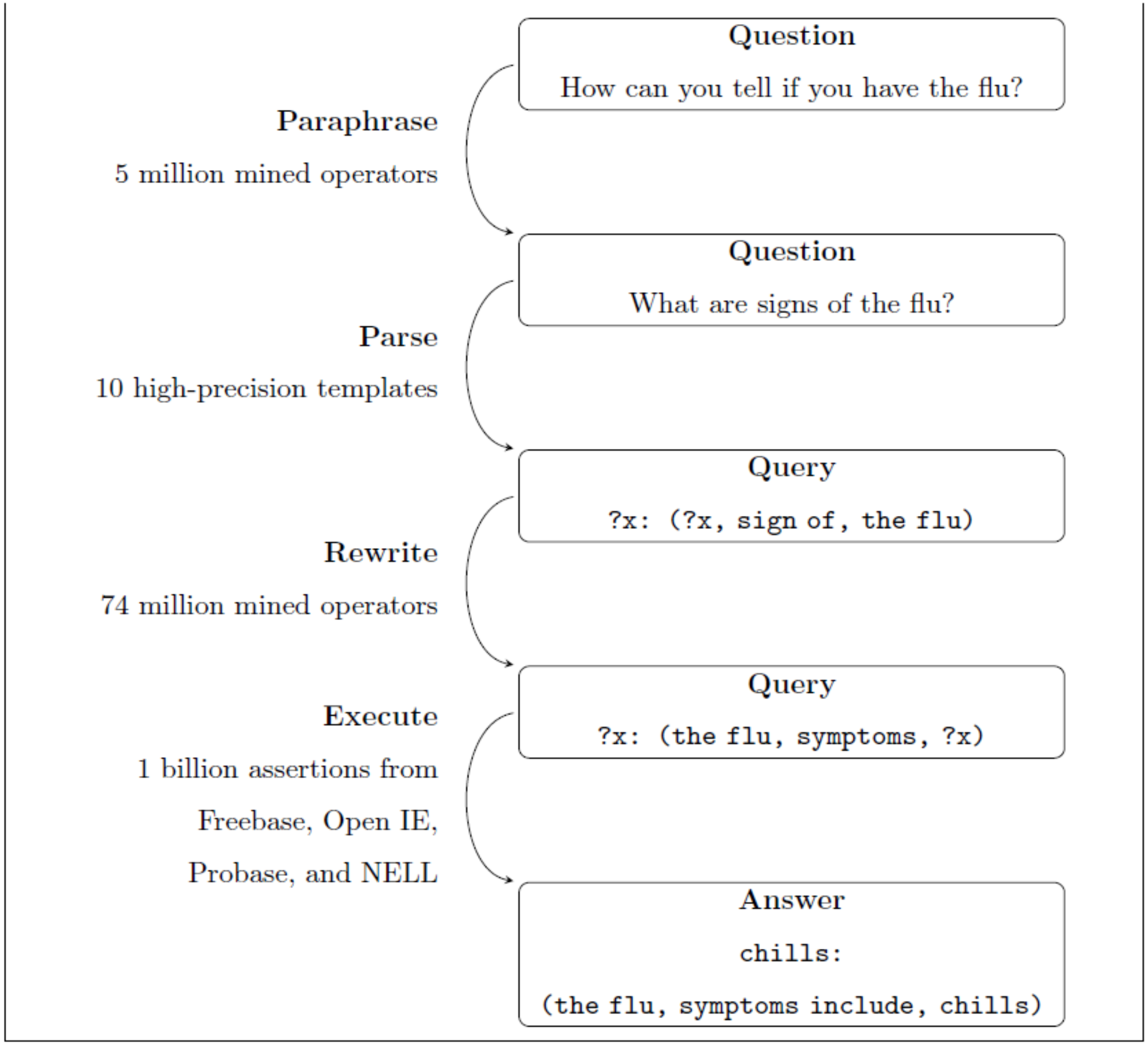
Q&A over LD

Fader thesis, 2013
An example cluster of questions that users on WikiAnswers have tagged as being paraphrases

Can us citizens gamble online?
Can you gamble online in america?
Internet gambling is legal in the US?
Is betting online legal in US?
Is it illegal to do online gamble?
Is it legal to gamble online in the US?
Is it legal to gamble online in america?
Is online casinos legal in the us?
Is online gambling forbidden in the US?
Online gambling in US is legal?

Q&A over LD

Fader thesis, 2013
An example of how
OQA maps
the question
“How can you tell if you have the flu?"
to the answer “the chills."



Q&A over LD

Fader thesis, 2013
ReVerb

Weight	Feature
1.16	(x, r, y) covers all words in s
0.50	The last preposition in r is "for"
0.49	The last preposition in r is "on"
0.46	The last preposition in r is "of"
0.43	$len(s) \leq 10$ words
0.43	There is a WH-word to the left of r
0.42	r matches VW*P from Figure 3.1
0.39	The last preposition in r is "to"
0.25	The last preposition in r is "in"
0.23	10 words $< len(s) \leq 20$ words
0.21	s begins with x
0.16	y is a proper noun
0.01	x is a proper noun
-0.30	There is an NP to the left of x in s
-0.43	20 words $< len(s)$
-0.61	r matches V from Figure 3.1
-0.65	There is a preposition to the left of x in s
-0.81	There is an NP to the right of y in s
-0.93	Coord. conjunction to the left of r in s

Q&A over LD

Fader thesis, 2013
PARALEX

Who wrote the Winnie the Pooh books?
Who is the author of winnie the pooh?
What was the name of the authur of winnie the pooh?
Who wrote the series of books for Winnie the poo?
Who wrote the children's storybook 'Winnie the Pooh'?
Who is poohs creator?
What relieves a hangover?
What is the best cure for a hangover?
The best way to recover from a hangover?
Best remedy for a hangover?
What takes away a hangover?
How do you lose a hangover?
What helps hangover symptoms?
What are social networking sites used for?
Why do people use social networking sites worldwide?
Advantages of using social network sites?
Why do people use social networks a lot?
Why do people communicate on social networking sites?
What are the pros and cons of social networking sites?
How do you say Santa Claus in Sweden?
Say santa clause in sweden?
How do you say santa clause in swedish?
How do they say santa in Sweden?
In Sweden what is santa called?
Who is sweden santa?

Q&A over LD

Fader thesis, 2013
PARALEX

Question Pattern	Knowledge Base Query
Who r e?	(?, r, e)
What r e?	(?, r, e)
Who does e r?	(e, r, ?)
What does e r?	(e, r, ?)
What is the r of e?	(?, r, e)
Who is the r of e?	(?, r, e)
What is r by e?	(e, r, ?)
Who is e's r?	(?, r, e)
What is e's r?	(?, r, e)
Who is r by e?	(e, r, ?)
When did e r?	(e, r in, ?)
When did e r?	(e, r on, ?)
When was e r?	(e, r in, ?)
When was e r?	(e, r on, ?)
Where was e r?	(e, r in, ?)
Where did e r?	(e, r in, ?)

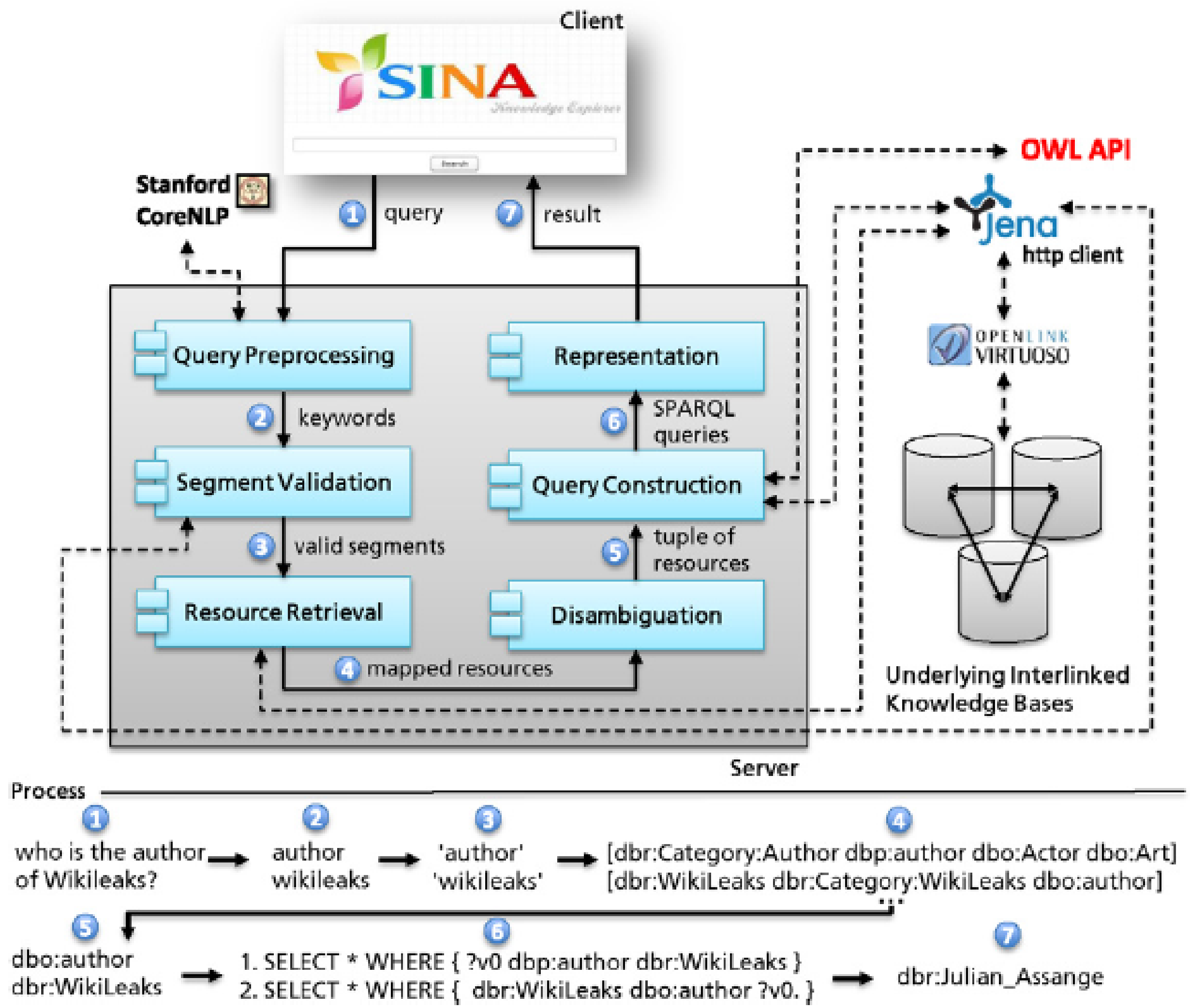
Q&A over LD

Ngonga Ngomo

- In [Shekarpour et al, 2014] the authors presents **SINA**, a scalable keyword search system that can answer user queries by transforming user-supplied keywords or NL queries into conjunctive SPARQL queries over a set of interlinked data sources. Sina uses a HMM to determine the most suitable resources for a user-supplied query from different datasets. Moreover, the framework is able to construct federated queries by using the disambiguated resources and leveraging the link structure underlying the datasets to query.
-

Q&A over LD

Ngonga Ngomo



Q&A over LD

Ngonga Ngomo

Category	Patterns	Pattern Schema
Instance-Property (IP)	IP.P1	$(s, p, ?o)$
	IP.P2	$(?s, p, o)$
	IP.P3	$(?s_1, ?p_1, o_1)(?s_1, p_2, ?o_2)$
	IP.P4	$(?s_1, ?p_1, o_1)(?o_2, p_2, ?s_1)$
	IP.P5	$(s_1, ?p_1, ?o_1)(?s_2, p_2, ?o_1)$
	IP.P6	$(s_1, ?p_1, ?o_1)(?o_1, p_2, ?o_2)$
Class-Instance (CI)	CI.P7	$(?s_1, a, c)(?s_1, ?p_1, o_1)$
	CI.P8	$(?s_1, a, c)(s_2, ?p_1, ?s_1)$
Instance-Instance (II)	II.P9	$(s, ?p, o)$
	II.P10	$(s, ?p_1, ?x)(?x, ?p_2, o)$
	II.P11	$(s_1, ?p_1, ?x)(s_2, ?p_2, ?x)$
	II.P12	$(?s, ?p_1, o_1)(?s, ?p_2, o_2)$
Class-Property (CP)	CP.P13	$(?s, a, c)(?s, p, ?o)$
	CP.P14	$(?s, a, c)(?x, p, ?s)$
Property-Property (PP)	PP.P15	$(?s, p_1, ?x)(?x, p_2, ?o)$
	PP.P16	$(?s_1, p_1, ?o)(?s_2, p_2, ?o)$
	PP.P17	$(?s, p_1, ?o_1)(?s, p_2, ?o_2)$

Q&A over LD

Ngonga Ngomo

Keywords	Answers
Instance characteristics.	
Kidman spouse	d:Kidman dp:spouse Keith Urban .
Iran language	d:Iran dp:Language d:Persian_language .
Titanic length	d:RMS_Titanic dp:Length 268.8336 .
Capital China	d:Republic_of_China dp:capital Beijing .
Michelangelo death	1. d:Michelangelo dp:deathDate "1564-02-18" . 2. d:Michelangelo dp:deathPlace "Rome, Italy" .
Associations between instances.	
Obama Clinton	d:Obama dp:predecessor d:Bush . d:Bush dp:predecessor d:Clinton .
Volkswagen Porsche	d:Volkswagen_Group dp:subsidiary d:Volkswagen .
Similar instances.	
Facebook Person	d:Facebook dp:keyperson d:Sheryl Sandberg . d:Sheryl Sandberg a d:Person .
Germany Island	1. d:Germany dp:Islands d:RÅijgen . d:RÅijgen a do:Island . 2. d:Germany dp:Islands d:FÅűhr . d:FÅűhr a do:Island . 3. d:Germany dp:Islands d:Sylt . d:Sylt a do:Island .
Lost Episode	1. d:Raised_by_Another dp:series dbp:Lost . d:Raised_by_Another a do:TVEpisode . 2. d:Homecoming dp:series dbp:Lost . d:Homecoming a do:TVEpisode . 3. d:Outlaws dp:series dbp:Lost . d:Outlaws a do:TVEpisode .
English Country	1. d:Ghana dp:officialLang d:English_language . d:Ghana a do:Country . 2. d:Cameroon dp:officialLang d:English_language . d:Cameroon a do:Country . 3. d:UK dp:officialLang d:English_language . d:UK a do:Country .