

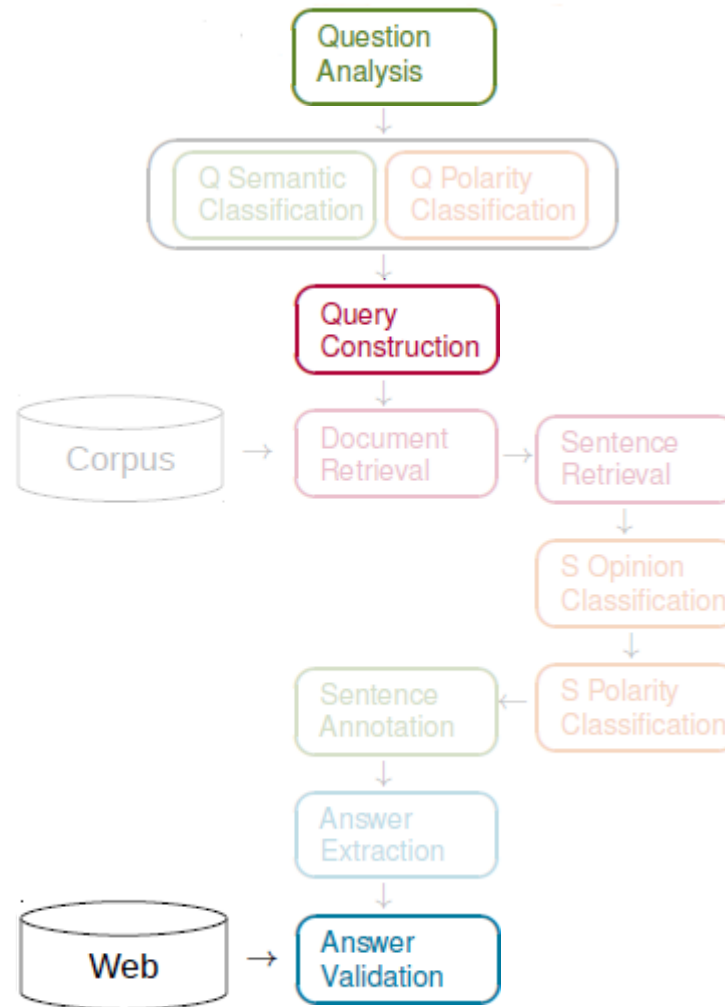
Answer Validation

The final step in Question Answering

Matthias Kohnen

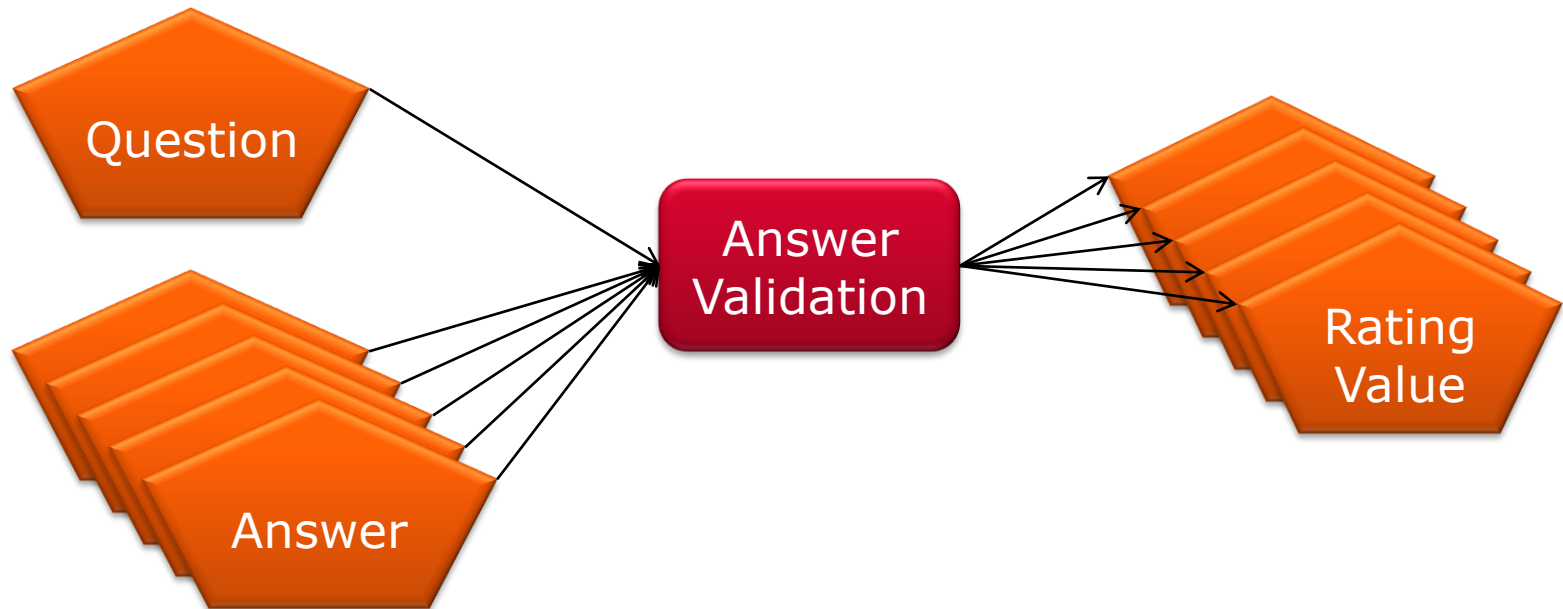
Today's Topic

2



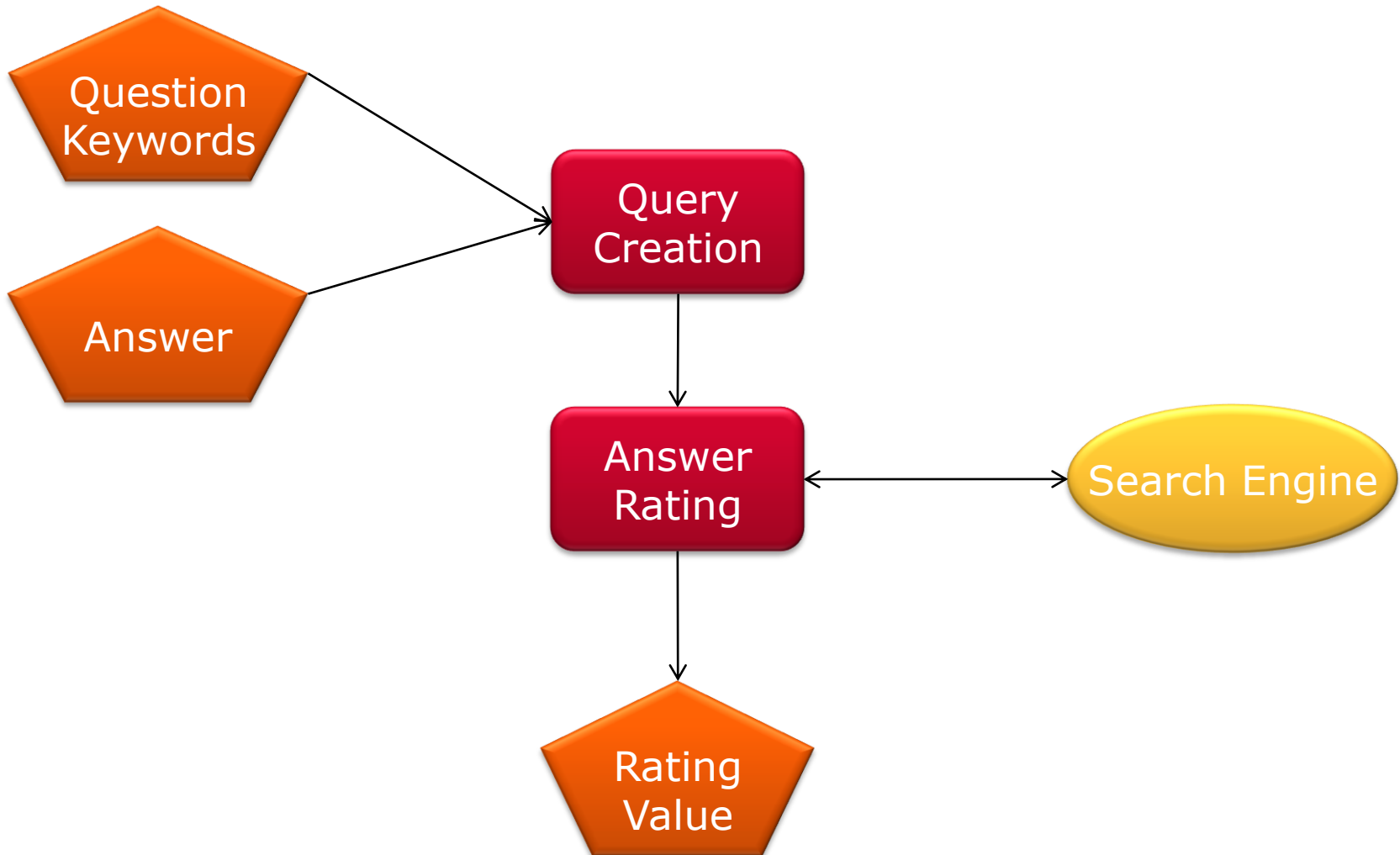
Today's Topic

3



The Basic Way

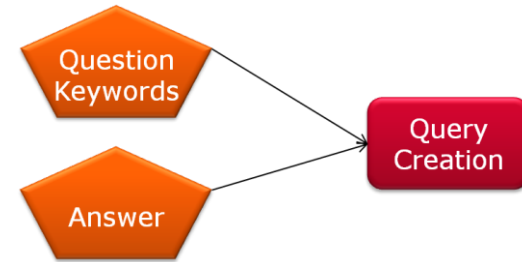
5



Query Creation

6

- Combining the answer with a subset of the question keywords
- If available sequences of keywords may be used
- Different attempts to choose the subset, various combinations of following
 - Between 2 and 10 keywords
 - Rating of every keyword
 - Using of stop word lists
 - Using of conjugated forms of keywords

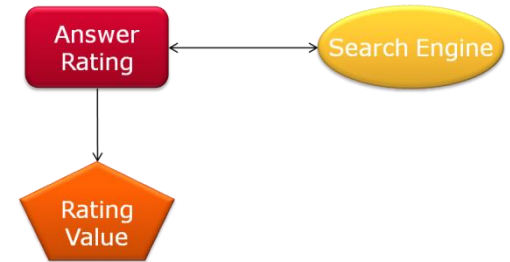


Answer Rating

7

- Passing the query to a search engine
- Analyzing the result of the search engine
 - Counting the results
 - Parsing the result snippets

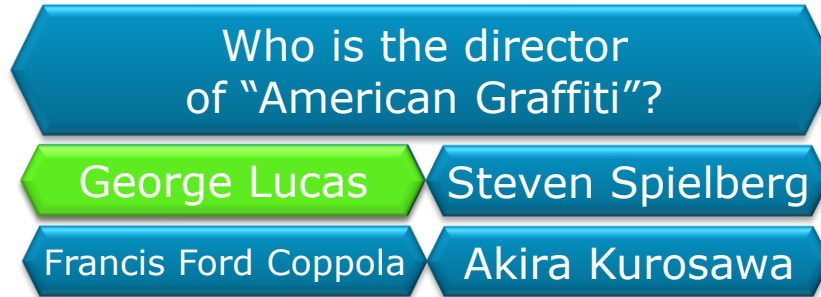
- Other possibilities: Using knowledge bases to find relations between question keywords and answer
 - More complex
 - Good for special question type
 - Quality depends on knowledge base quality



Keyword Association Counting

Who is the director of "American Graffiti"?

9



Google

Search

director "american graffiti" "george lucas"

About 2,460,000 results

Google

Search

director "american graffiti" "steven spielberg"

About 1,330,000 results (0.17 seconds)

Google

Search

director "american graffiti" "francis ford coppola"

About 244,000 results (0.17 seconds)

Google

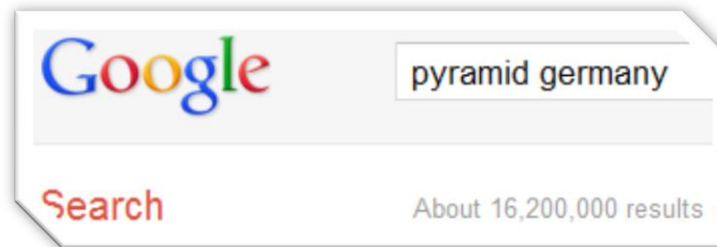
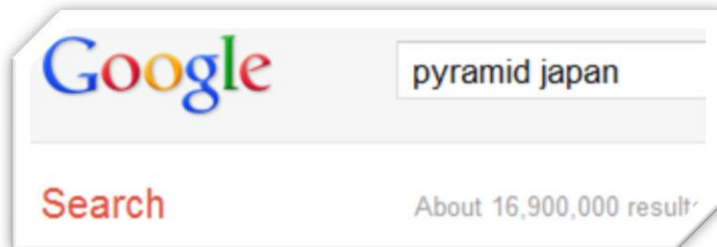
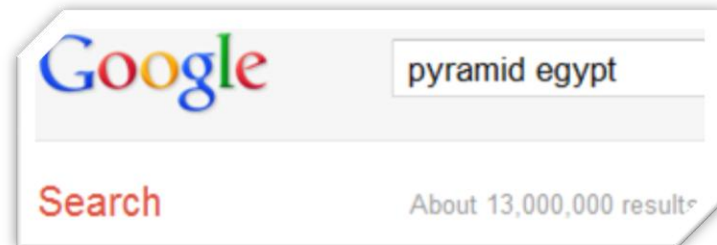
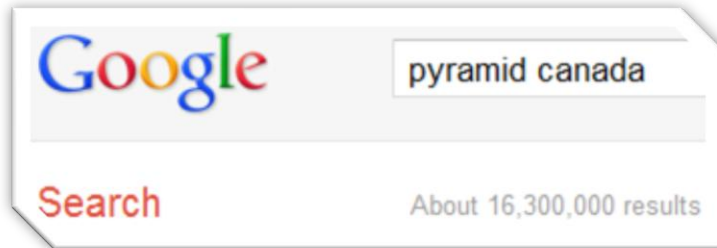
Search

director "american graffiti" "akira kurosawa"

About 137,000 results (0.53 seconds)

Where is Pyramid?

10



Where is Pyramid?

11

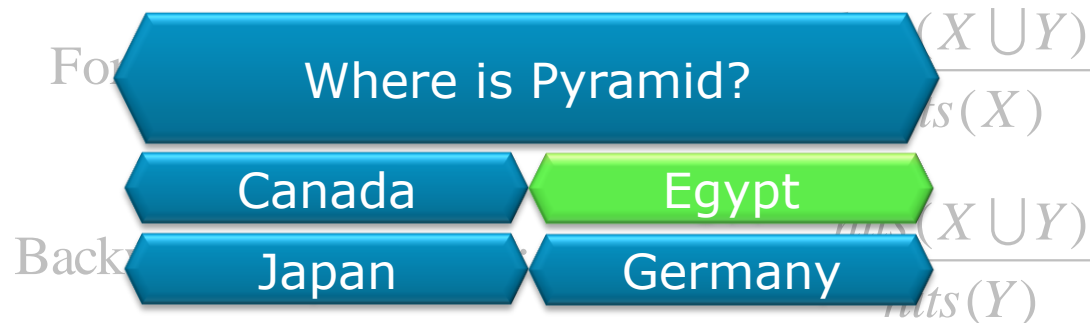
Query	Hits(Query)	Hits(Answer)	BA
pyramid canada	16.3 M	2650 M	0.006
pyramid egypt	13 M	633 M	0.021
pyramid japan	16.9 M	2430 M	0.007
pyramid germany	16.2 M	1820 M	0.009

Search

About 2,430,000,000 results

Search

About 1,820,000,000 results



Search Snippets Analysis

What is the Capital of Germany?

14

The image shows a collage of search results for the query "capital germany". Overlaid on the collage are four blue arrows pointing to the right, each containing a city name: "Cologne", "Hamburg", "Munich", and "Berlin". The "Berlin" arrow is highlighted in green, indicating it is the correct answer. The background search results include snippets from Wikipedia, GermanPlace, and other sources, all mentioning Berlin as the capital of Germany.

Comparison

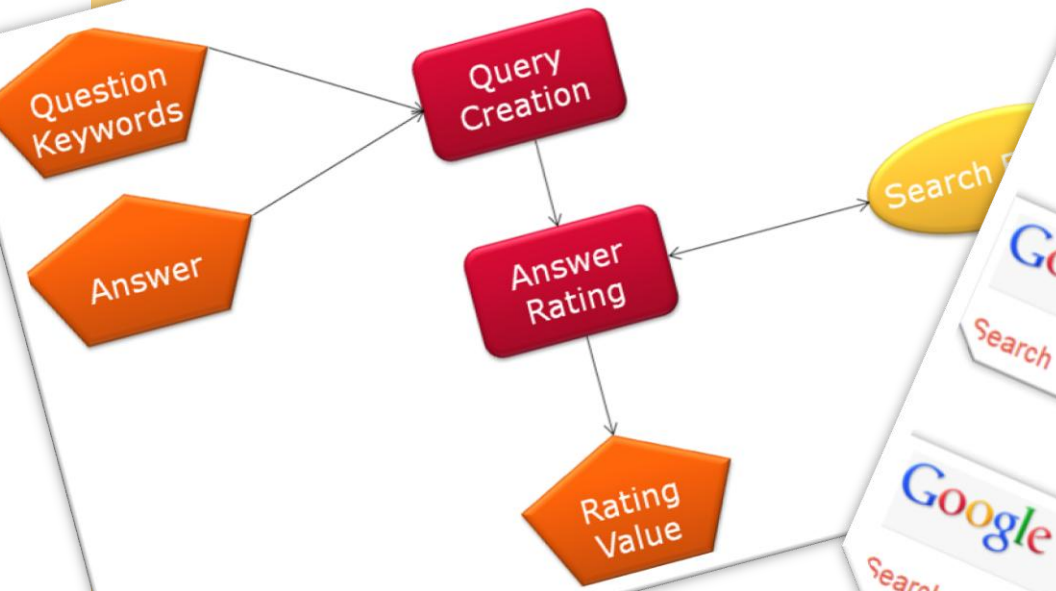
16

Keyword Association

- Japanese version of “Who wants to be a millionaire”
- Answers 76% of the questions
- Up to 3 queries for each answer option
- Very less calculation needed

Search Snippets Analysis

- QA track in TREC-2001
- Answers 86% of the questions
- Depending on question type from 78% to 93%
- Just 1 query for each answer option
- Much more text analysis needed



Google Search: director "american graffiti" "george lucas" About 2,460,000 results

Google Search: director "american graffiti" "akira kurosawa" About 1,330,000 results (0.17 seconds)

Google Search: director "american graffiti" "steven spielberg" About 137,000 results (0.53 seconds)

What is the Capital of Germany?

Cologne Hamburg

Munich Berlin

capital germ capital germ

[Cologne - Wikipedia](#)
Cologne was German: Ru Eau de Colog

[Hamburg - Wikipedia](#)
Hamburg is a m Award scheme, Hamburger SV - Po

[Munich German](#)
www.germanplace Since 1806 Munich media and high-pov

[Capital of Germany - Wikipedia, the free encyclopedia](#)
en.wikipedia.org/wiki/Capital_of_Germany
The Capital of Germany is the city state of Berlin. It is the seat of the ...
+ Show more results from wikipedia.org

[Berlin](#)
userpage.chemie.fu-berlin.de/adressen/berlin.html
14 Feb 1995 - Berlin is the capital and the biggest city of Germany. It has a population of about 3.5 million and extends over 889 square kilometers. ...

[German Capital Moves from Bonn to Berlin - Geography](#)
geography.about.com/library/misc/bibonnberlin.htm
17 Apr 1999 - Germany's capital city moves 370 miles to the northeast, from your About.com Guide.

[Welcome to Germany's capital city, Berlin](#)

[Cologne](#)
www.study-in Cologne is C underscores

[Hamburg \(Germany\)](#)
www.citypopulation Name, Abbr. Status 2005-12-31, E 2010

[Munich travel guide](#)
wikitravel.org/en/Mu Its outstanding posi Berlin becoming the

[Munich | German](#)

[Cologne](#)
www.timeanddate.c

Literature

18

Masatsugu Tonoike, Takehito Utsuro and Satoshi Sato
Answer Validation by Keyword Association
COLING Conference, 2004

Bernardo Magnini, Matteo Negri, Roberto Prevete and Hristo Tanev
Mining the web to validate answers to natural language questions
Data Mining Conference, 2002