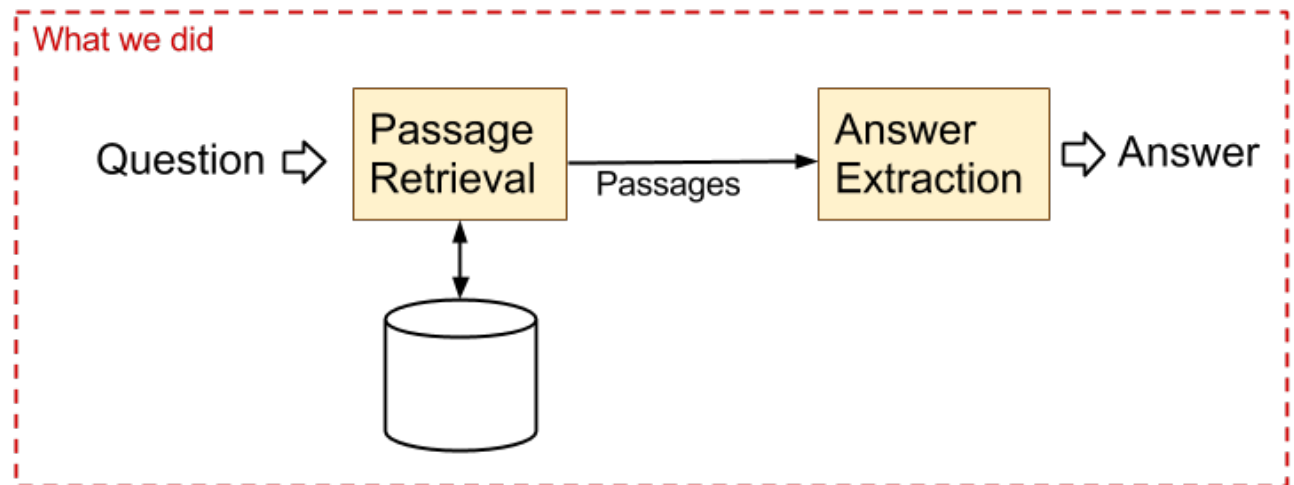
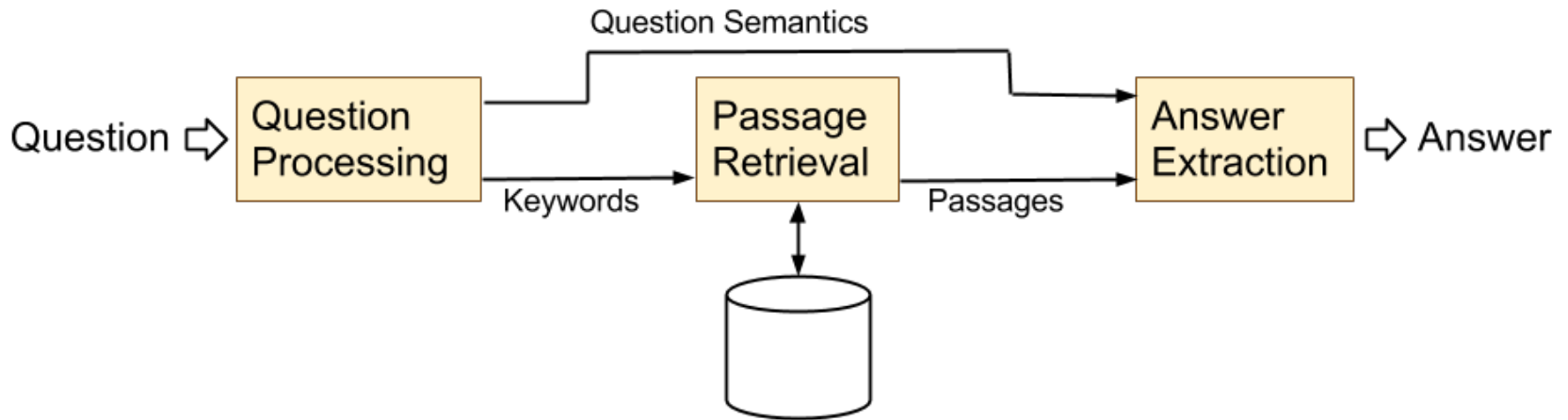


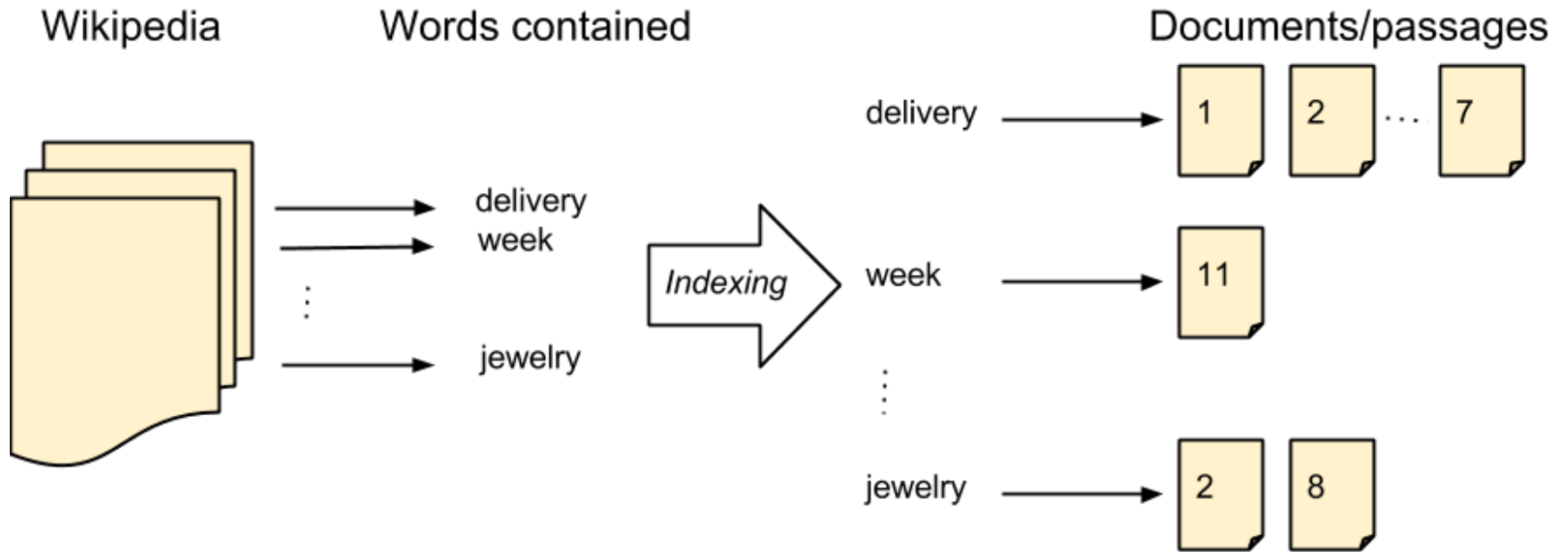
Passage Retrieval

Juri Pyykkö, Rebecka Weegar

Question Answering System



Indexing



Searching

[delivery, home, cheap, week, jewelry, in] (vocabulary)

$d = [1 , 1 , 2 , 0 , 1 , 0]$ (document)

$q = [0 , 0 , 1 , 0 , 1 , 0]$ (query)

$$\text{sim}(q, d) = \frac{q \cdot d}{|q||d|}$$

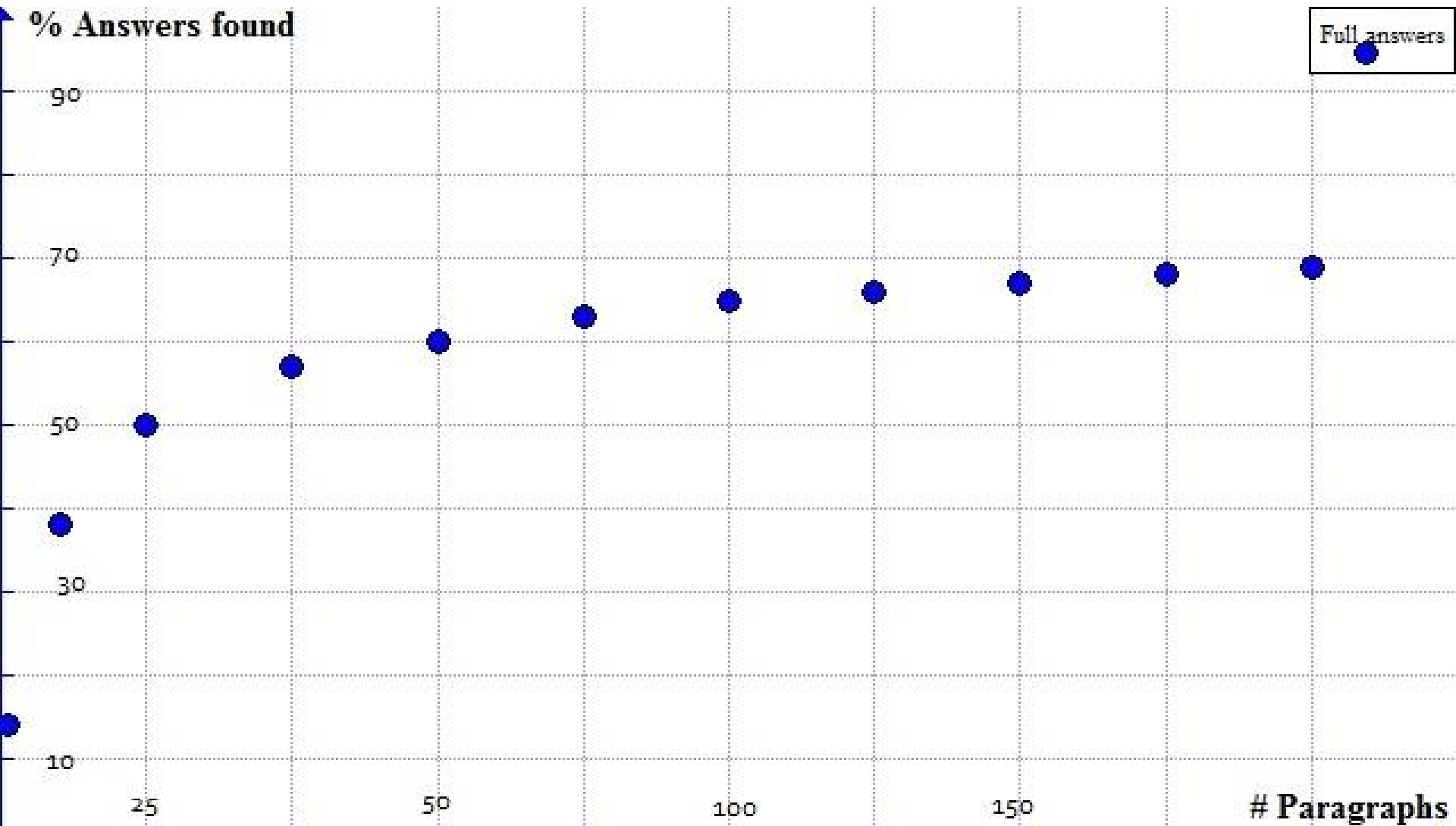
Answer Extraction



The class/type of the candidate answers should correspond to the expected class/type.

Results

Results



Matching answers

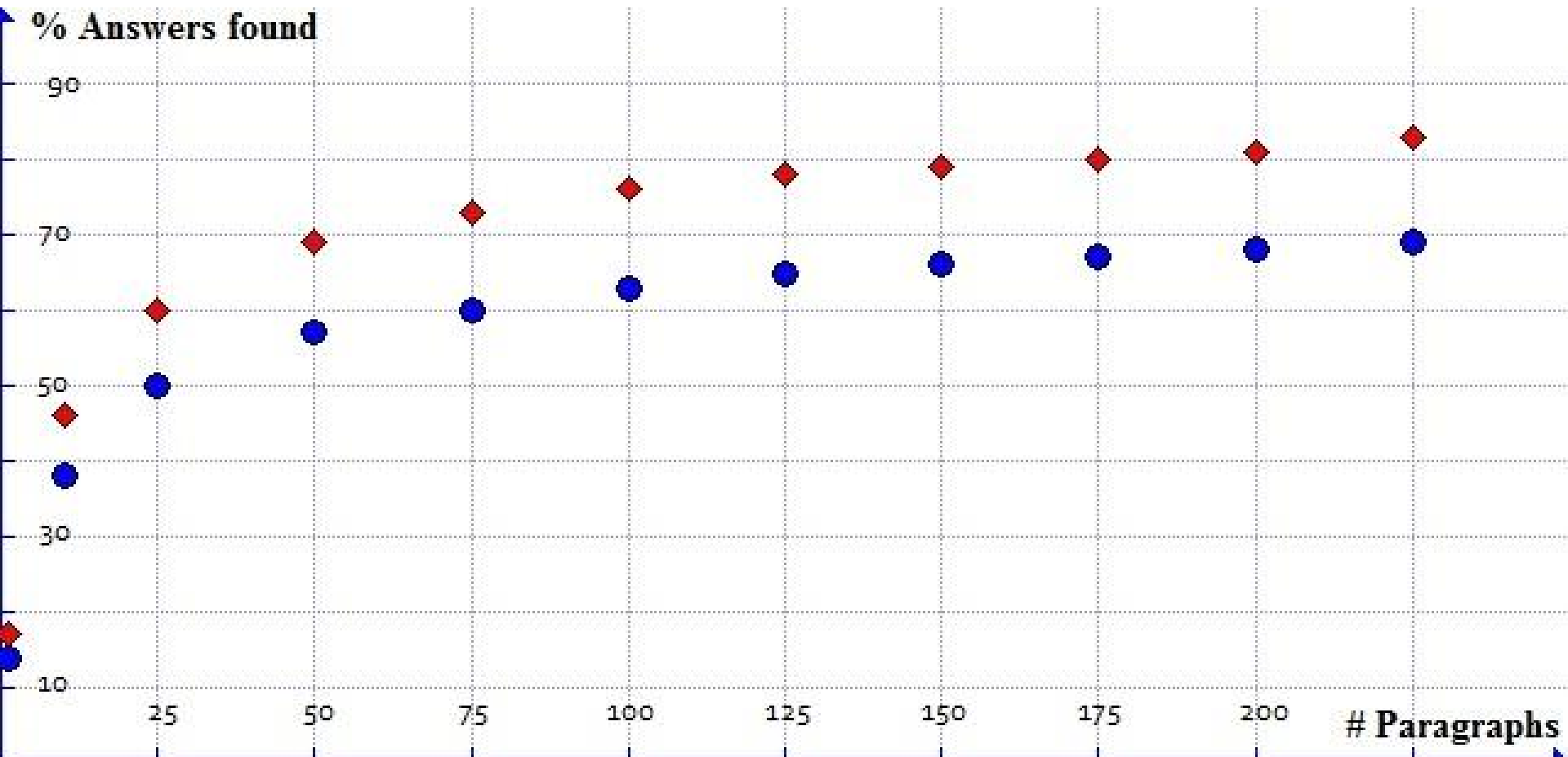
Q: Nämn ett av de två skolhusen på Hogwarts utom Slytherin och Gryffindor?

A: Hufflepuff eller Ravenclaw

Q: Hur hittade prinsen Askungen?

A: Genom att prova hennes kvarglömda sko på alla flickor

Short answers



Headlines

Härskarringen

Q: Vem är den ondaste av dem alla? (Who is the most evil of them all?)

A: Sauron

Answer candidates

- One word answers only
- Questions with answers tagged as "Entity"
- Stagger, part-of-speech, nouns
- Ranking, frequency

Example

Headline: Hunden

Q: Vad heter hundens ungar?

A: Valpar

Example

Headline: Hunden

Q: Vad heter hundens ungar?

A: Valpar

Lucene query:

hund kall hund ung

Example, vad heter hundens ungar?

From wikipedia 700MB -> About 2000 words

...Tuaregerna lät unga hundar tidigt följa med på jakt tillsammans med vuxna hundar. På så sätt lärde sig unghundarna av de äldre hundarnas erfarenheter.

Den vetenskapliga läran om hundar kallas "kynologi." En hund av hankön kallas "hund" och en av honkön kallas för "tik" och dess avkomma kallas "valp". Den vetenskapliga läran om hundar kallas "kynologi."

Jaktformen kallas hetsjakt och jakten bedrivs normalt med två till sex hundar. Hetsjakt är inte längre tillåtet i västeuropeiska länder med ett fåtal undantag, så här betraktas rasen först och främst som familjehund. Hundutställningar, hundkapplöpning och lure coursing är områden där rasen framgångsrikt deltar...

Example - Stagger

11 en en RG RG
12 av av PP PP
13 honkön honkön NN NN
14 kallas kalla VB VB
15 för för PP PP
16 " " PAD PAD
17 tik tik NN NN
18 " " PAD PAD
19 och och KN KN
20 dess dess PS PS
21 avkomma avkomma NN NN
22 kallas kalla VB VB
23 " " PAD PAD
24 valp valp NN NN
25 " " PAD PAD
26 . . MAD MAD

2000 words -> 300 words

Ranked answer candidates

QUESTION: Hunden Vad kallas hundens ungar

NORMALIZED ANSWER: valp

CANDIDATE:FREQUENCY : hund:177, jakt:7, människa:7, varg:5, ras:5, räv:5, liv:5, roman:4, del:4, musikalbum:4, valp:4, skillnad:3, unge:3, katt:3...

Answer candidates

456 Questions, 325 answers found (71 %)

Ranking among found answers

4 % in top 5

13 % in top 10

29 % in top 20

71 % in top 100