

# QA Search

Yini Yang  
Nicolas Jonason

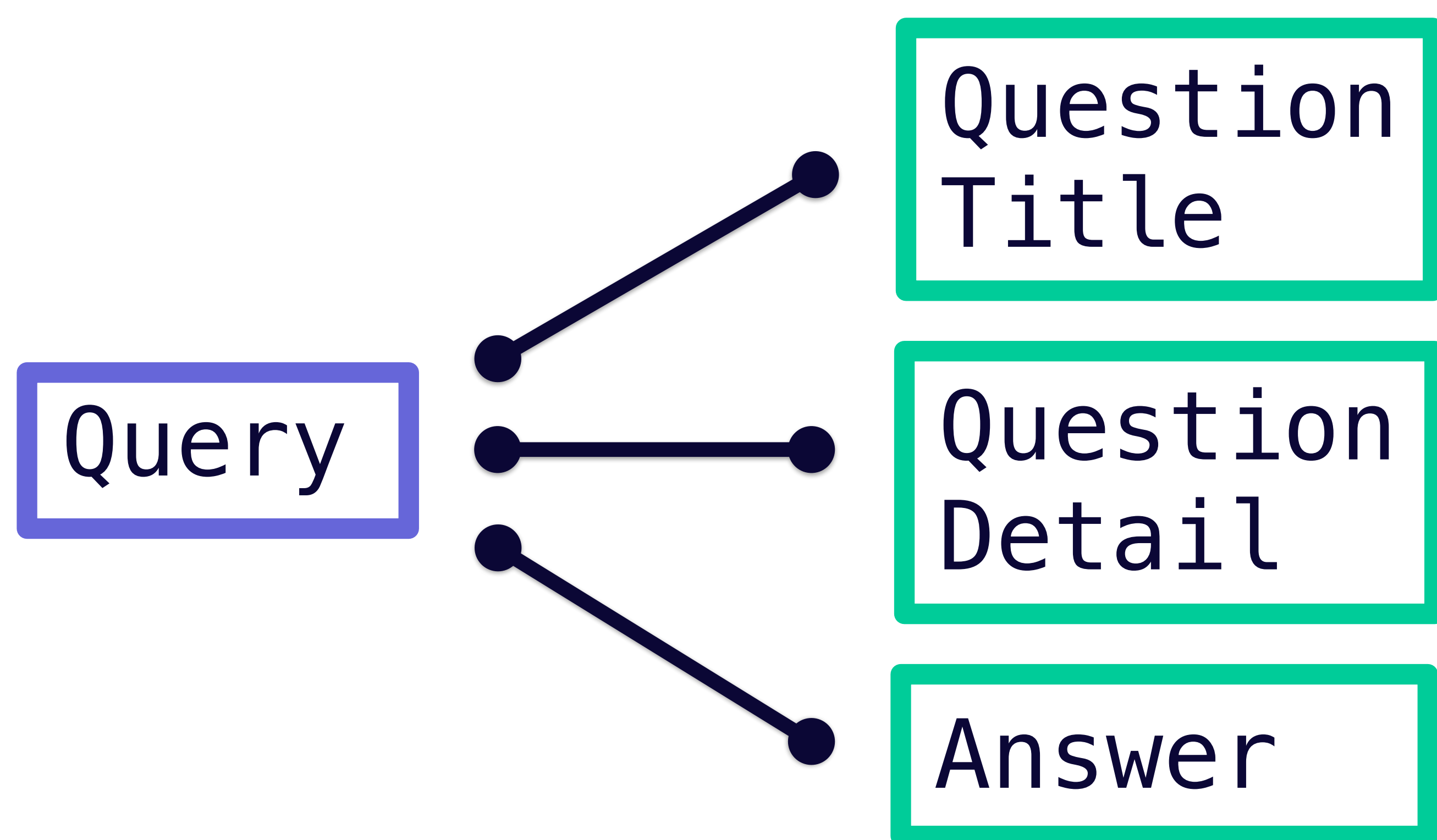
**Ruoyan Meng**  
**Lisa Vällfors**

## Our goal

- >Allow users to search a set of question answer pairs.
- >See what fields and retrieval techniques methods improve search in a QA context.

## 1) Baseline

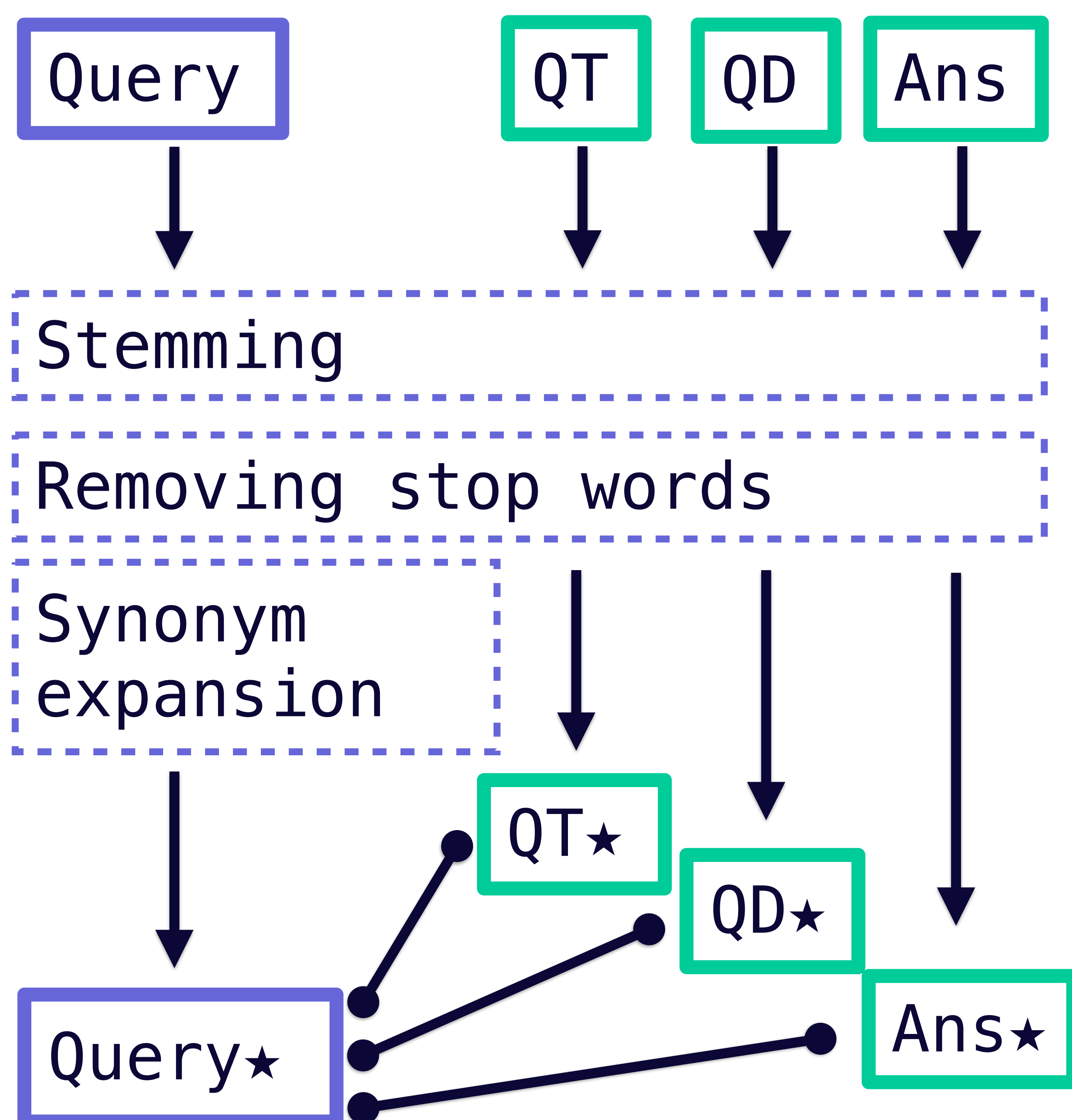
- >Elastic search on any subset of three fields.
- >Lucene practical scoring (tf\_idf & vector space model).



## 2) Advanced search

**All of the above as well as:**

- >Reduce words to their stems.
- >Remove stop words.
- >Expand query with synonyms (wordnet)



# Evaluation

## 'Jeopardy evaluation'

- >Take 50 random answers from the corpus.
- >Formulate new questions for each answer.
- >Query the engine and rate the returned answer w.r.t our question.

0. Not relevant at all
1. Same topic/somewhat relevant
2. Answers the question

matching question title

2	0	0	0	1	1	0	0	2	0	2	2	1	0	0	2	2	0	0	0	2	0	2	2	0
2	0	1	0	0	0	1	0	2	0	2	2	0	0	0	0	0	2	0	1	2	0	2	1	0

## matching question title & detail

2	0	0	0	1	0	1	0	2	0	2	2	0	0	0	2	2	0	0	0	2	0	2	2	0
2	0	1	0	0	0	1	0	2	0	0	2	0	0	0	0	0	2	0	1	2	0	2	1	0

### matching question title & detail & answer

2	2	1	2	2	1	0	2	2	2	2	2	1	2	0	2	2	0	1	0	2	2	2	2	0
2	0	1	0	0	0	1	0	2	0	0	2	0	0	0	0	0	2	0	1	2	0	2	1	0

*\*baseline above, advanced below*