Using Keyqueries to Reduce Misinformation in Health-Related Search Results

April 10, 2022


Maik Fröbe


Sebastian Günther


Alexander Bondarenko


Johannes Huck


Matthias Hagen

Martin-Luther-Universität Halle-Wittenberg

first.last@informatik.uni-halle.de

@webis_de www.webis.de

Challenges in Health-Related Searches

Example Information Need

[Clarke et al.; TREC'21]

Does selenium help prevent cancer?

Challenges in Health-Related Searches

Example Information Need

[Clarke et al.; TREC'21]

Does selenium help prevent cancer?

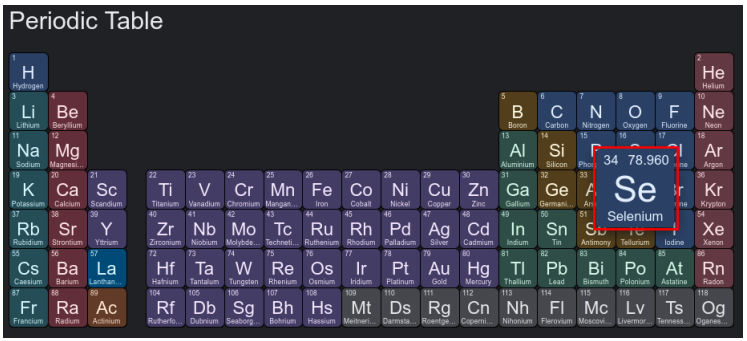
SELENIUM AUTOMATION TESTING

Challenges in Health-Related Searches

Example Information Need

[Clarke et al.; TREC'21]

Does selenium help prevent cancer?



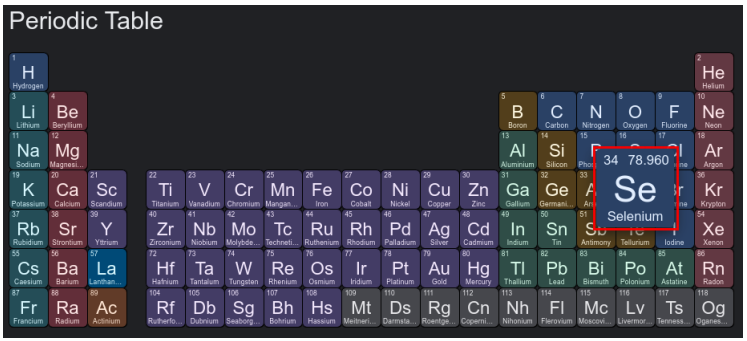
- Selenium is a chemical element
- Medical use-cases of selenium:
  - selenium deficiency
  - high cholesterol

Challenges in Health-Related Searches

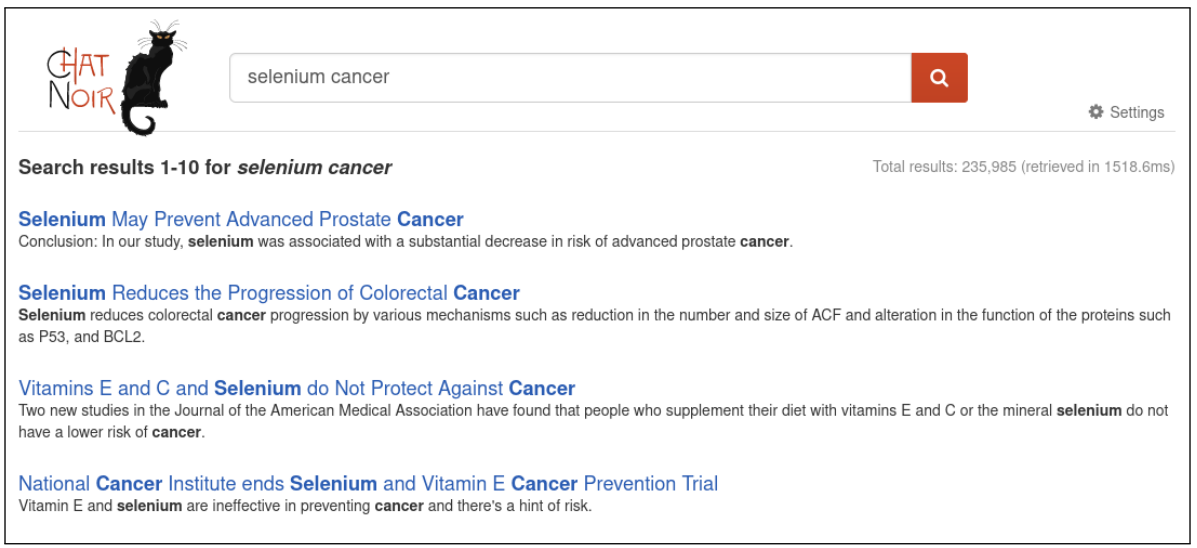
Example Information Need

[Clarke et al.; TREC'21]

Does selenium help prevent cancer?



- Selenium is a chemical element
- Medical use-cases of selenium:
  - selenium deficiency
  - high cholesterol

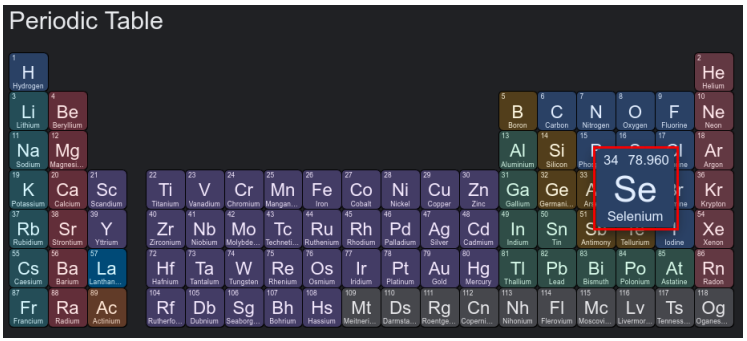


Challenges in Health-Related Searches

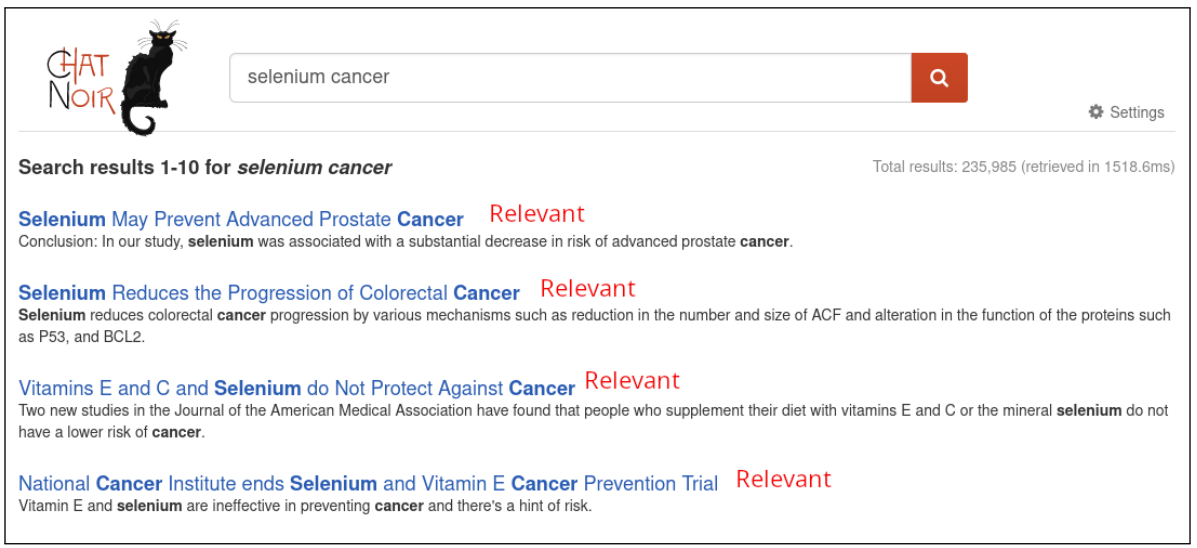
Example Information Need

[Clarke et al.; TREC'21]

Does selenium help prevent cancer?



- Selenium is a chemical element
- Medical use-cases of selenium:
  - selenium deficiency
  - high cholesterol

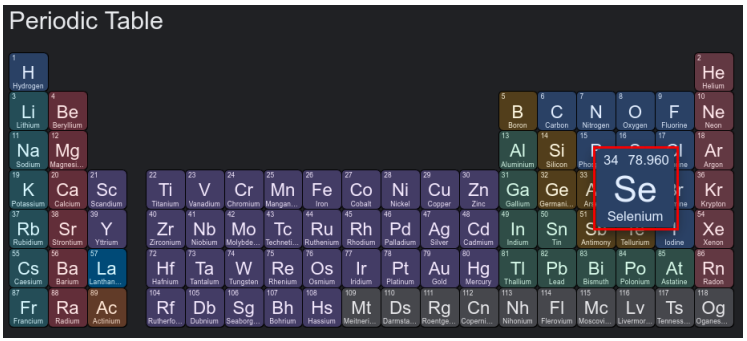


Challenges in Health-Related Searches

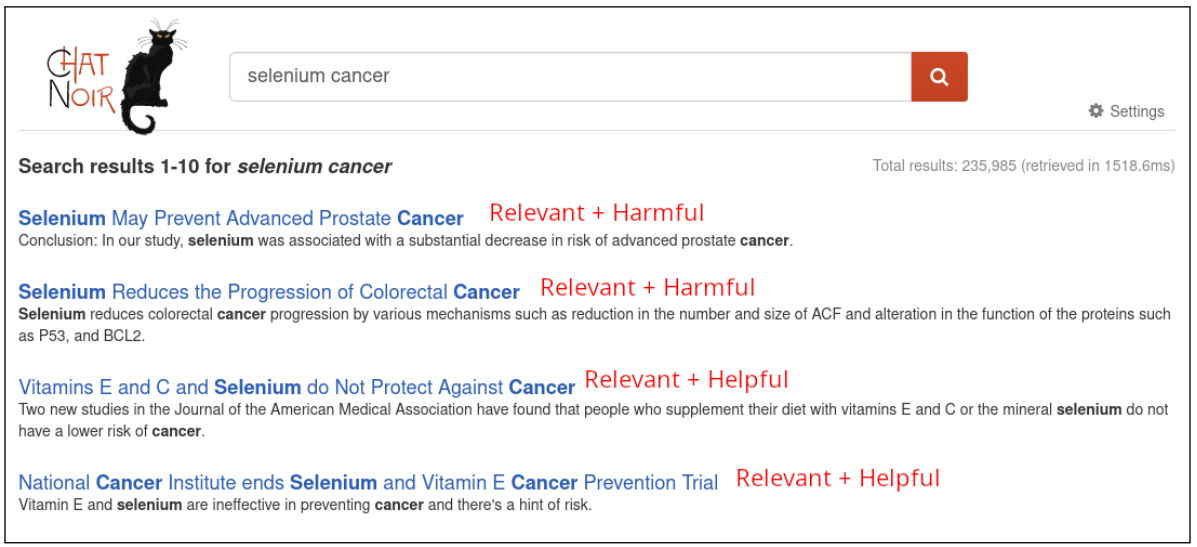
Example Information Need

[Clarke et al.; TREC'21]

Does selenium help prevent cancer?



- Selenium is a chemical element
- Medical use-cases of selenium:
  - selenium deficiency
  - high cholesterol

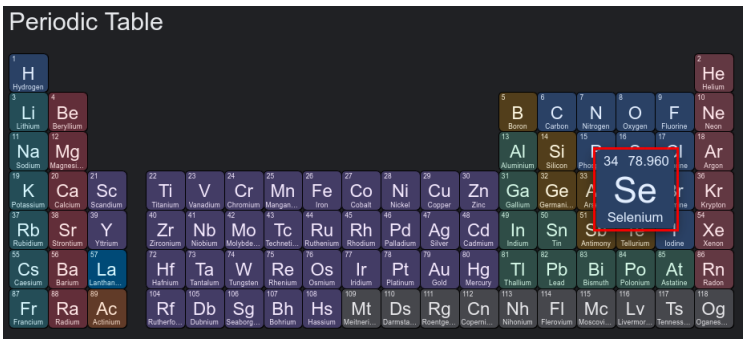


Challenges in Health-Related Searches

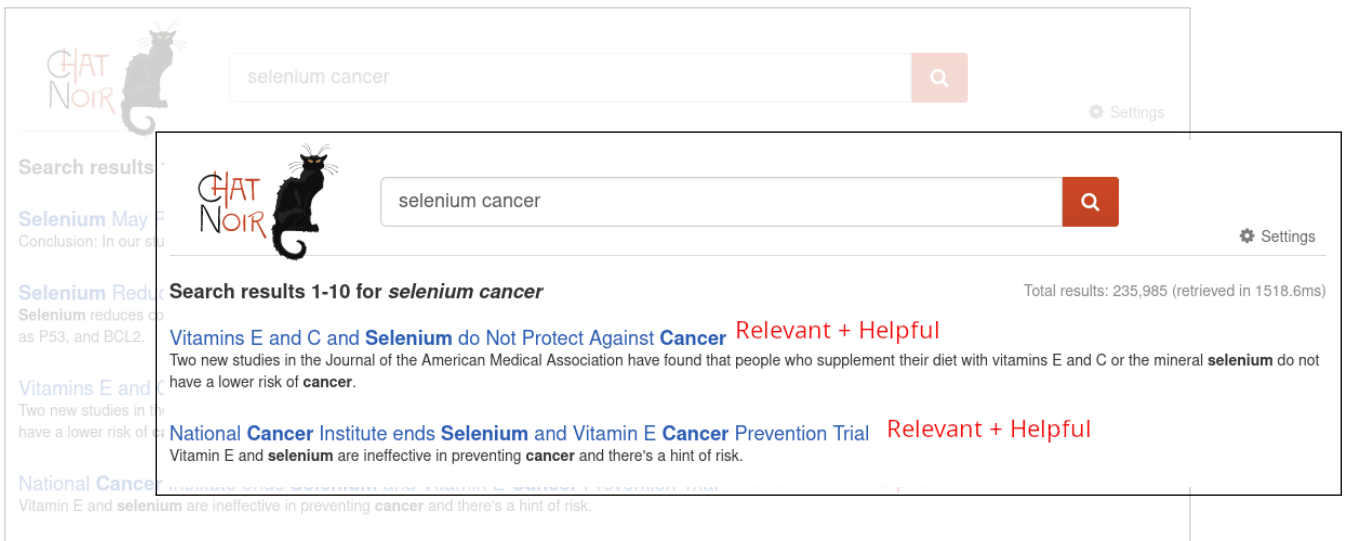
Example Information Need

[Clarke et al.; TREC'21]

Does selenium help prevent cancer?



- Selenium is a chemical element
- Medical use-cases of selenium:
  - selenium deficiency
  - high cholesterol



Automatic Approaches for Health-Related Searches?

Fully Automatic Approaches Seem “Utopical”

- Scientific knowledge changes rapidly
- Distinguishing helpful and harmful documents is difficult
  - maybe not possible using only the documents text?

We Take a Step Back

Can explicit relevance feedback by experts guide query expansion methods to formulate queries that return fewer misleading or wrong results?

- Approach:
  - Simulate expert providing explicit relevance feedback
  - Explicit relevance feedback is input for query expansion approach
  - Return ranking for expanded query

Baselines

Query Expansion with RM3

[Jaleel et al.; TREC'04]

- Adds new terms with weights to a query
- Exploits information from feedback documents
- Step 1 (i.e., RM1): Estimate relevance model for feedback documents
  - Each term gets an score.
  - E.g., for feedback documents R for the original query q, a term t obtains:

$$\sum_{d \in R} P(t|d) \frac{score(q,d)}{\sum_{d' \in R} score(q,d')} ,$$

- Step 2 (i.e., RM3): Linearly combine original query weight with RM1

Observation: RM3 might be ineffective for explicit relevance feedback

- RM3 does not consider ranks of feedback documents
- RM3 does not check whether all expansion terms are actually needed
- Result: Costly relevance feedback by experts is maybe not fully leveraged

Can We Make Better Use of Explicit Feedback From Experts?

The key are . . .

Can We Make Better Use of Explicit Feedback From Experts?

The key are . . . keyqueries

Query Expansion with Keyqueries

What is a keyquery? [Hagen et al.; ECIR'16]

Query q is a keyquery for a set D of target documents against a search engine iff

1. Every $d \in D$ is in the top-k results. (specificity)
2. Query q has at least l results. (generality)
3. No subquery $q' \subset q$ satisfies the above. (minimality)

Remark: For small $|D| \leq 5$, typically $l \geq 10$ and $k = 10$.

Query Expansion with Keyqueries

Example: Keyquery for a paper ($l \geq 1000$, $k = 3$)

Elastic ChatNoir:
Search Engine for the ClueWeb and the Common Crawl

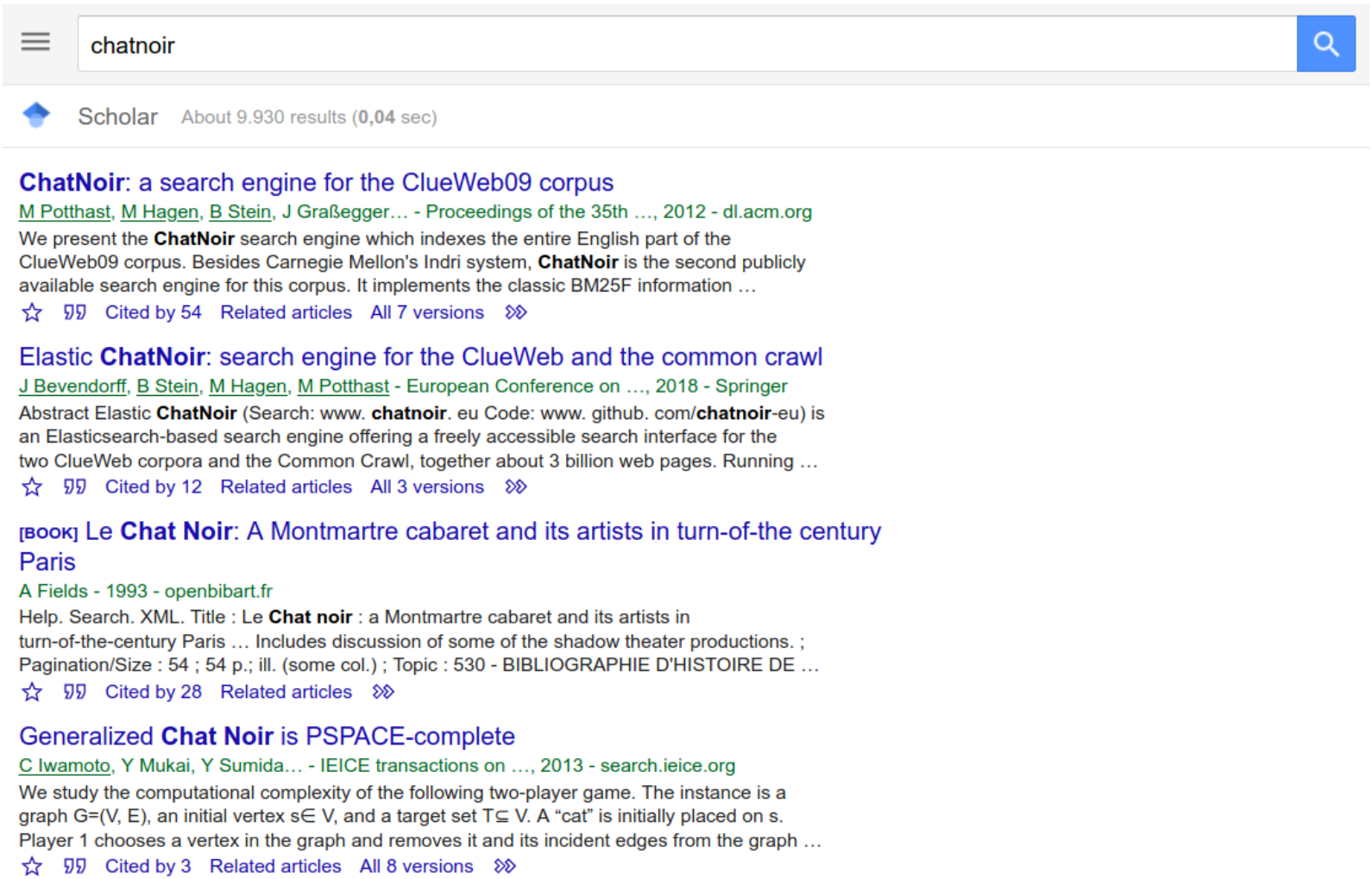
Janek Bevendorff, Benno Stein, Matthias Hagen, and Martin Potthast

Bauhaus-Universität Weimar and Leipzig University
<first name>.<last name>@uni-weimar.de and martin.potthast@uni-leipzig.de

Abstract Elastic ChatNoir[1] is an Elasticsearch-based search engine offering a freely accessible search interface for the two ClueWeb corpora and the Common Crawl, together about 3 billion web pages. Running across 130 nodes, Elastic ChatNoir features subsecond response times comparable to commercial search engines. Unlike most commercial search engines, it also offers a powerful API that is available free of charge to IR researchers. Elastic ChatNoir's main purpose is to serve as a baseline for reproducible IR experiments and user studies for the coming years, empowering research at a scale not attainable to many labs beforehand, and to provide a platform for experimenting with new approaches to web search.

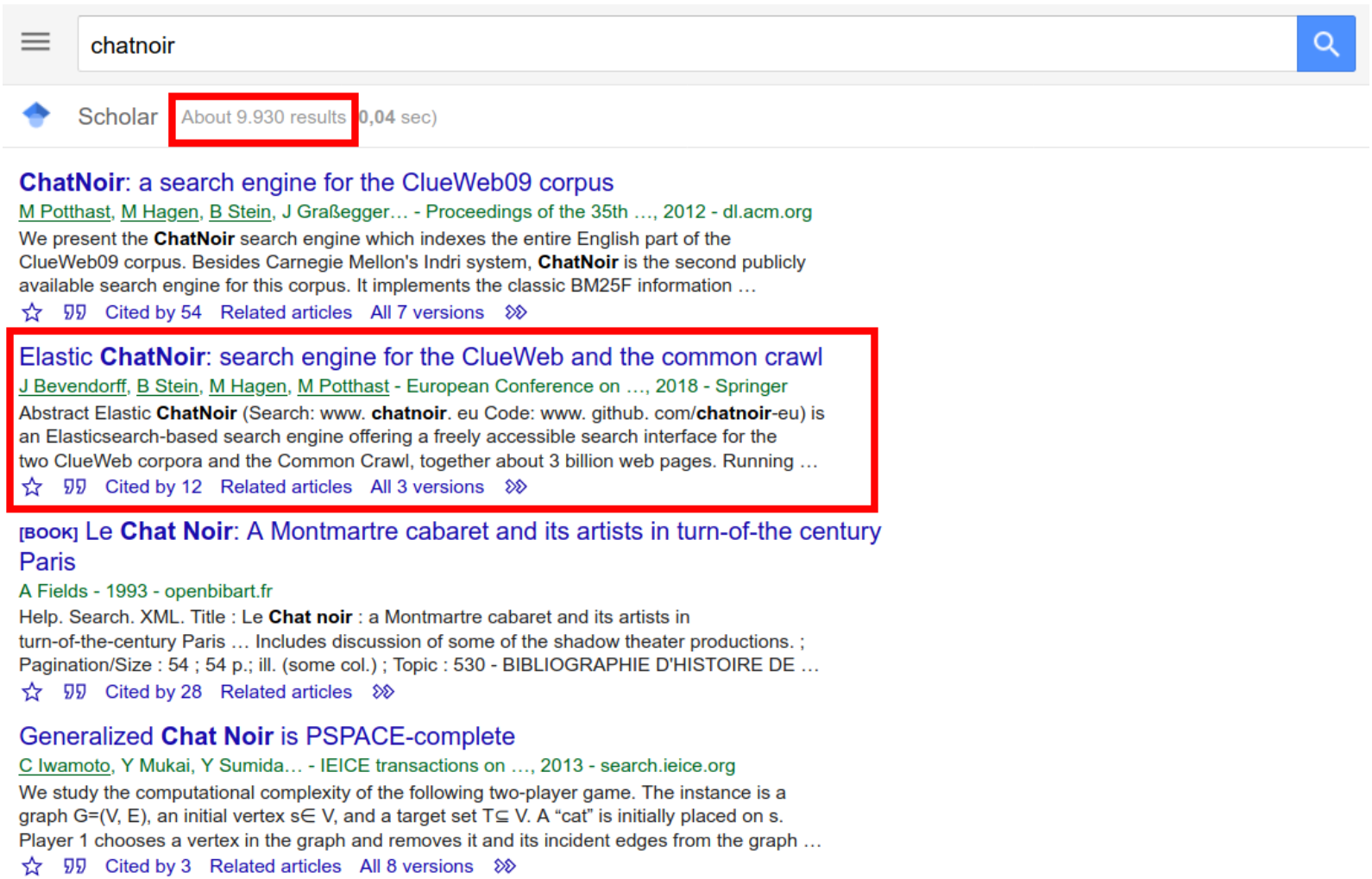
Query Expansion with Keyqueries

Example: `chatnoir` is a keyquery for Google Scholar



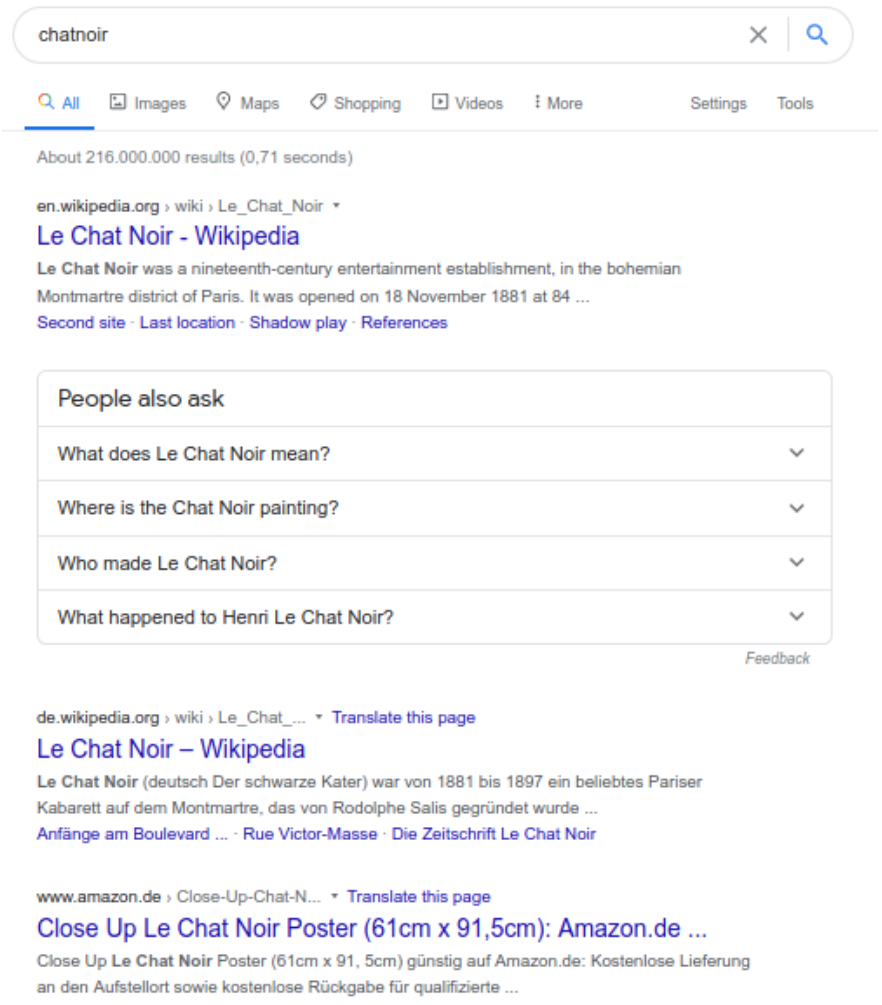
Query Expansion with Keyqueries

Example: `chatnoir` is a keyquery for Google Scholar



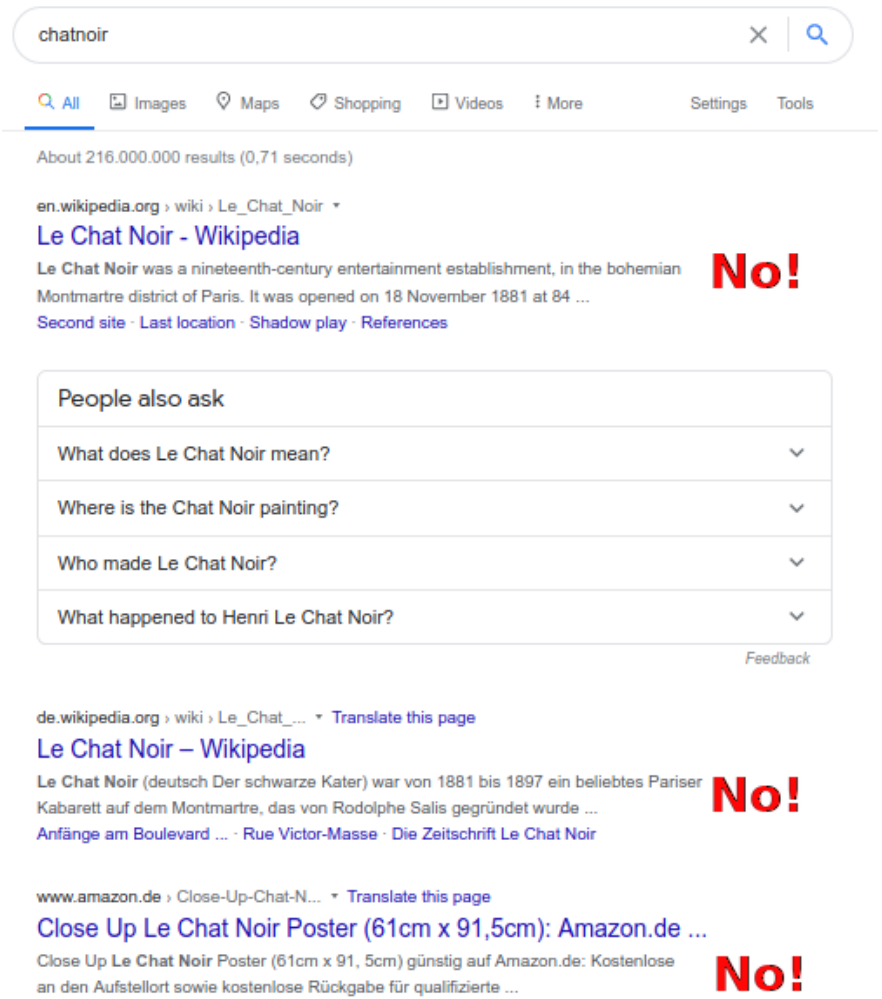
Query Expansion with Keyqueries

Example: . . . but not for Google



Query Expansion with Keyqueries

Example: . . . but not for Google



Query Expansion with Keyqueries

Simulation of explicit relevance feedback

- First k relevant and helpful documents that appear in the BM25 ranking

Implementation Details

- Generate vocabulary V of <term, weight> expansion terms with RM3
- $\mathcal{Q} = 2^V \setminus \{\emptyset\}$ are the meaningful queries over V
- Relaxation of the first keyquery criteria if necessary:
  - Retrieve $|R| - 1$, $|R| - 2$, ... target documents
- Selection of expanded query if there are multiple keyqueries
  - Candidate with the highest nDCG@k for the target documents
- We use a simple brute-force implementation

Evaluation

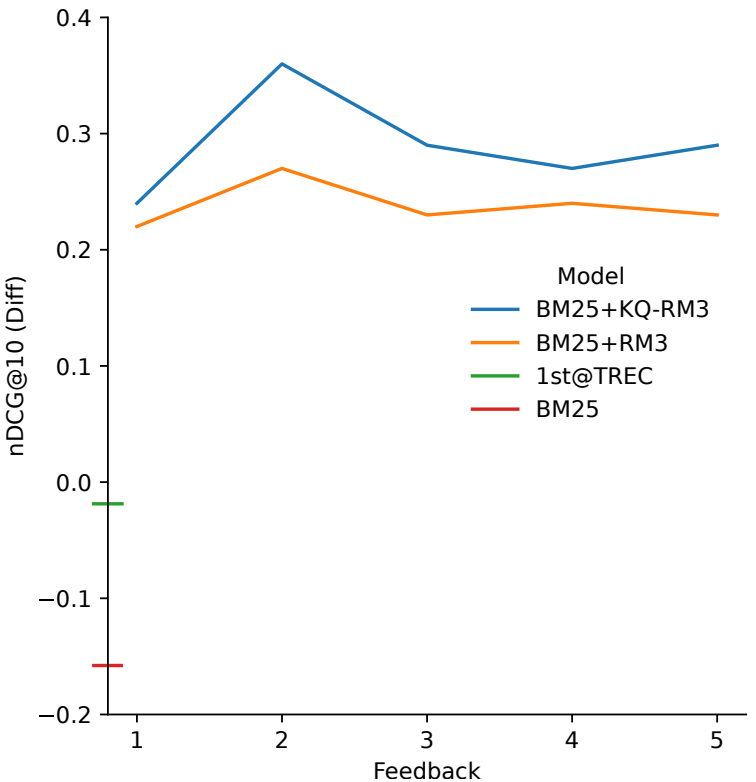
Setup: TREC 2019–2021 Decision and Health Misinformation Tracks

- Helpful qrels: Documents are relevant and helpful
- Harmful qrels: Documents are relevant but harmful
- Effectiveness measured as difference: Helpful - Harmful

Results on the TREC 2019 Decision Track

- Corpus: ClueWeb12 category B
- 50 judged Topics. E.g.:

Does hypnotherapy help people quit smoking?



Evaluation

Setup: TREC 2019–2021 Decision and Health Misinformation Tracks

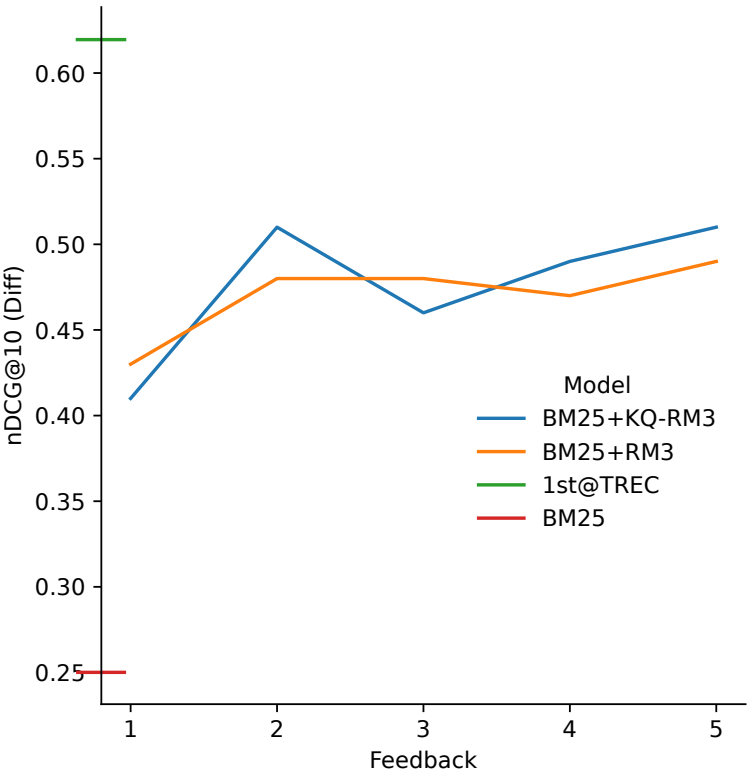
- Helpful qrels: Documents are relevant and helpful
- Harmful qrels: Documents are relevant but harmful
- Effectiveness measured as difference: Helpful - Harmful

Results on the TREC 2020 Decision Track

- Corpus: Common Crawl News Crawl
- 46 judged Topics. E.g.:

  Can ibuprofen worsen COVID-19?

- VERA (1st@TREC)
  [Pradeep et al.; SIGIR'21]
  - Answer as explicit feedback
  - Predict with T5 whether the document supports the answer
  - Interpolated with Mono/DuoT5



Evaluation

Setup: TREC 2019–2021 Decision and Health Misinformation Tracks

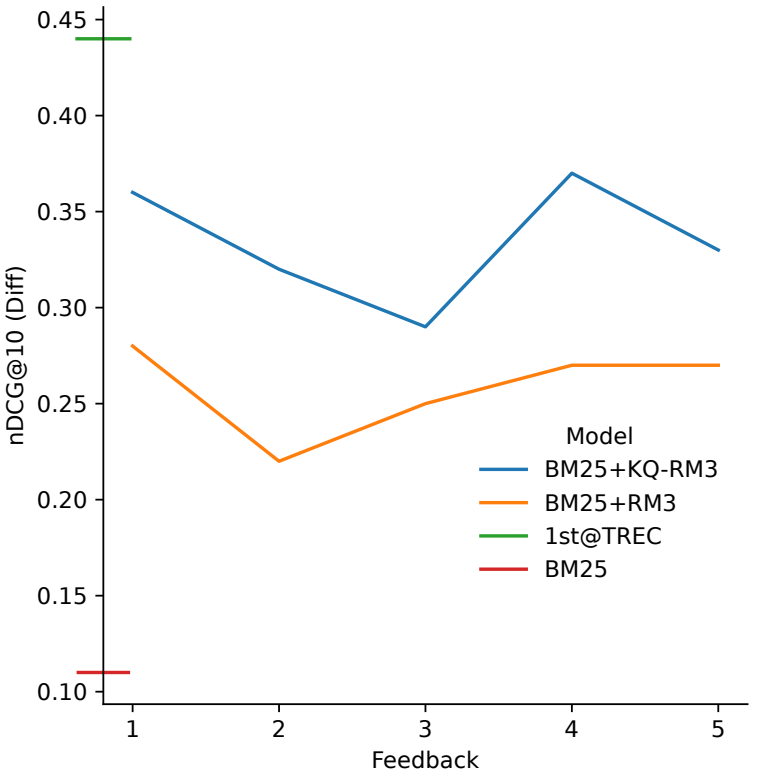
- Helpful qrels: Documents are relevant and helpful
- Harmful qrels: Documents are relevant but harmful
- Effectiveness measured as difference: Helpful - Harmful

Results on the TREC 2021 Decision Track

- Corpus: noclean subset of C4
- 35 judged Topics. E.g.:

  Does selenium help prevent cancer?

- VERA (1st@TREC)
  [Pradeep et al.; SIGIR'21]
  - Difference to
    - BM25+KQ-RM3 not significant
    - BM25+RM3 is significant



Conclusions

Summary

- RM3 query expansions can be enhanced with the concept of keyqueries
  - Leads to more effective BM25 queries: Increased help, decreased harm
- Explicit feedback is important for health-related searches
  - VERA is the state of the art

Conclusions

Summary

- RM3 query expansions can be enhanced with the concept of keyqueries
  - Leads to more effective BM25 queries: Increased help, decreased harm
- Explicit feedback is important for health-related searches
  - VERA is the state of the art

Future work

- Expand study:
  - More relevance feedback approaches, query expansion approaches, retrieval models
- Replace brute-force implementation
  - Efficient enumeration or reverted indexes
- Replace costly relevance feedback with information available on the Web?
  - E.g., schema.org/ClaimReview

Conclusions

Summary

- RM3 query expansions can be enhanced with the concept of keyqueries
  - Leads to more effective BM25 queries: Increased help, decreased harm
- Explicit feedback is important for health-related searches
  - VERA is the state of the art

Future work

- Expand study:
  - More relevance feedback approaches, query expansion approaches, retrieval models
- Replace brute-force implementation
  - Efficient enumeration or reverted indexes
- Replace costly relevance feedback with information available on the Web?
  - E.g., schema.org/ClaimReview

thank you!