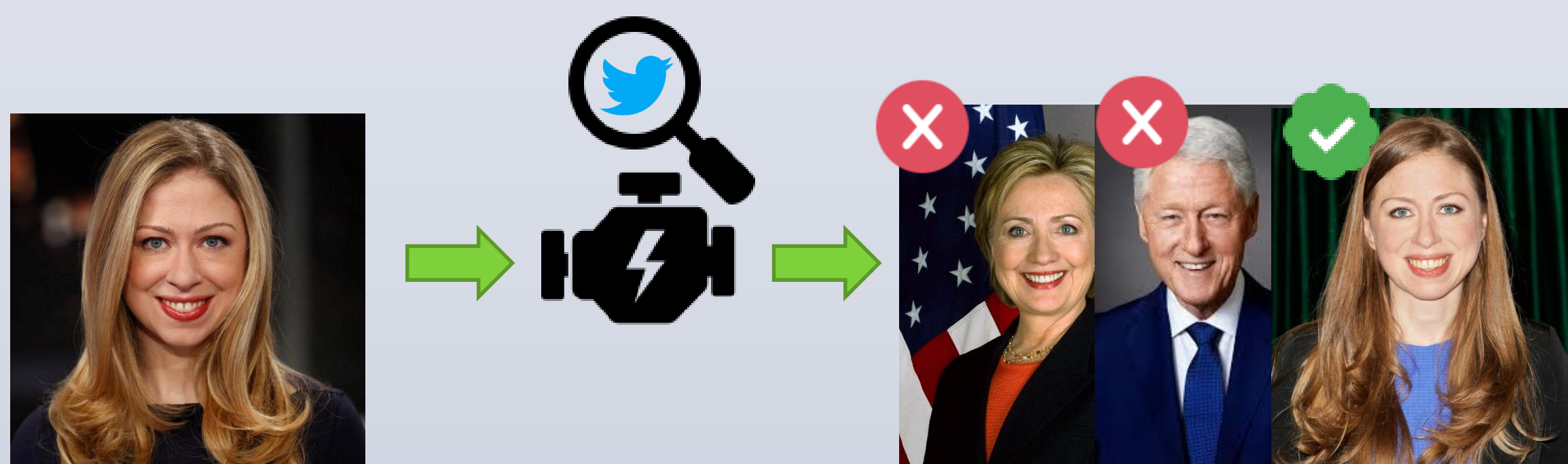


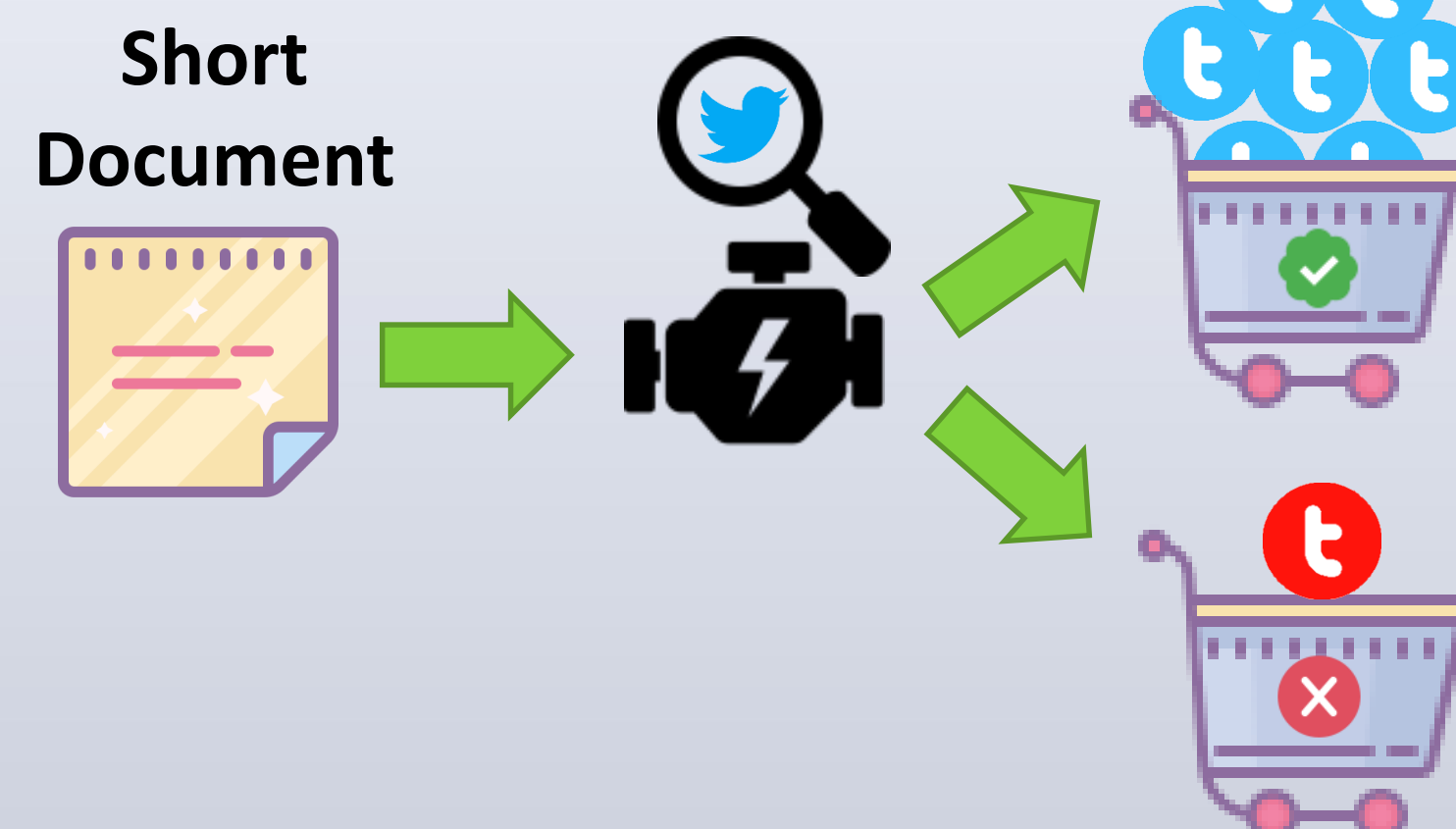
## MOTIVATION

- Short keyword queries are one of the main tools of any user seeking information through the ubiquitous search engines available on the Web.
- Unfortunately, in many cases, the ambiguity of short keyword queries causes poor performance.
- The problem of ambiguity is exacerbated when working with opaque search engines.
- In contrast to transparent search engines, opaque search engines provide a very limited level of interactivity and hide all activities that the search engine performs, including the repository itself.



## GOALS

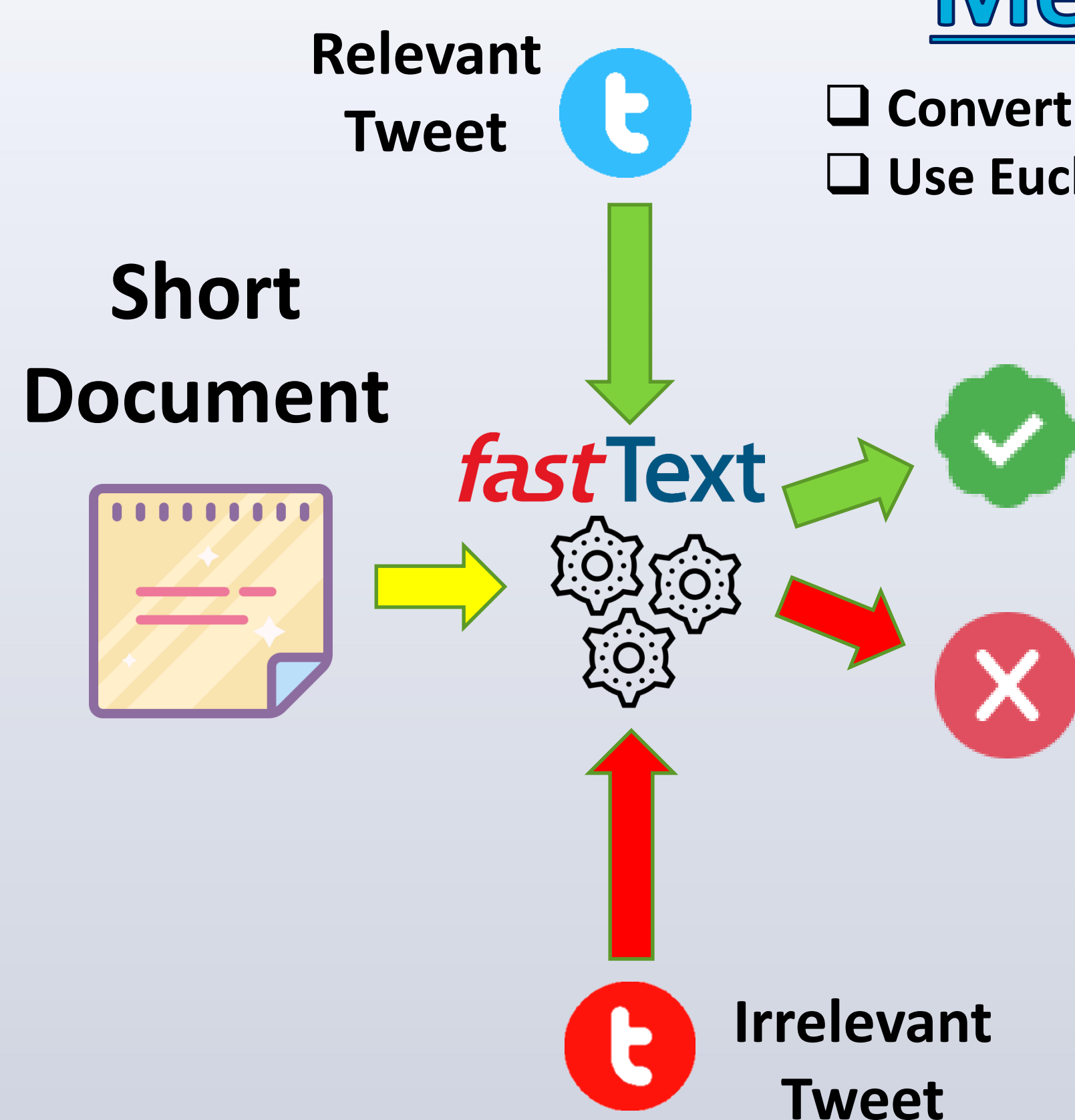
- In this paper, we suggest an automated iterative keyword optimization method for improving information retrieval from opaque search engines.



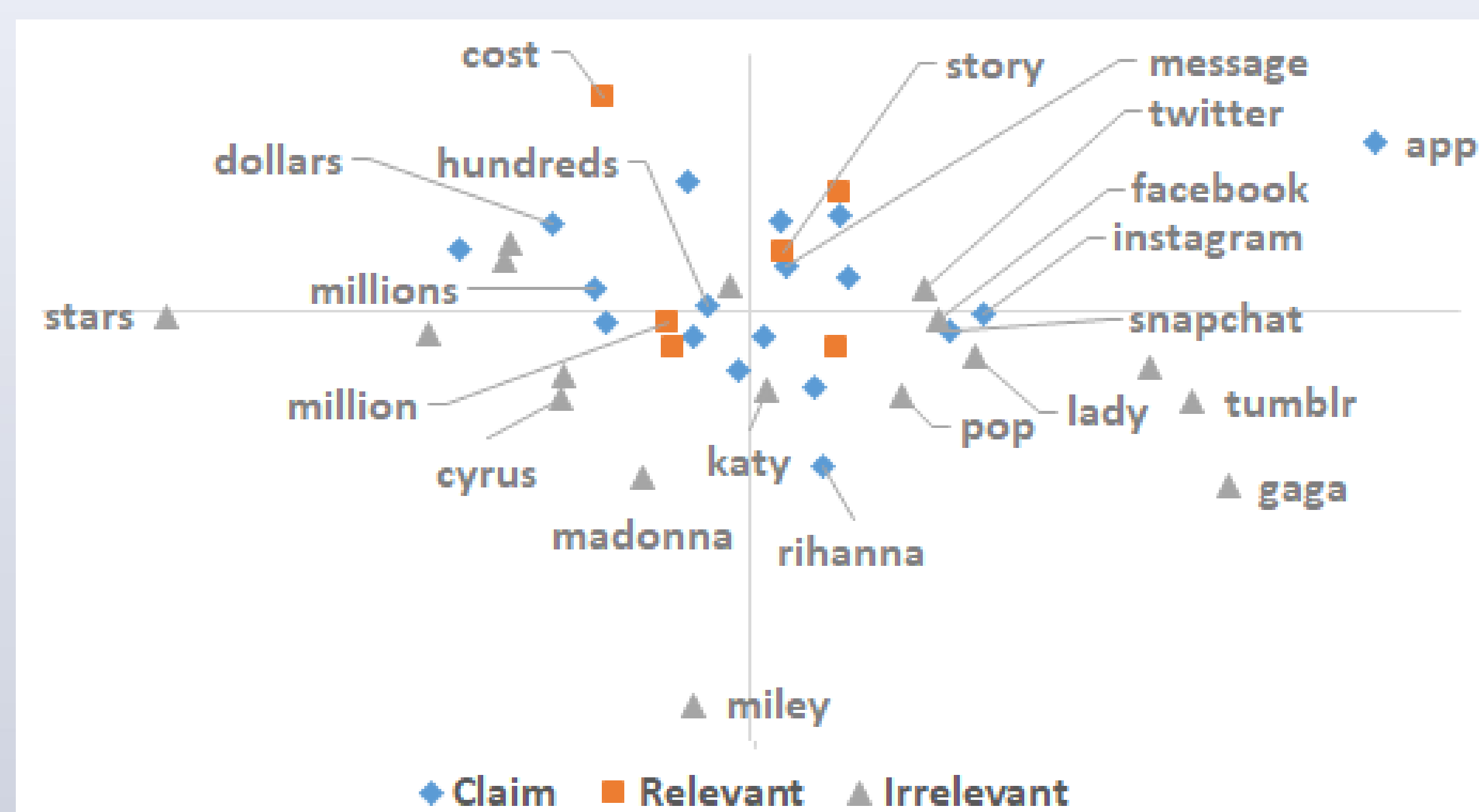
## CONTRIBUTIONS

- Estimating the relevance of posts to the claim by comparing their vector representations (AUC of 0.9 based on 1,078 tweets related to 20 claims).
- A greedy search algorithm that iteratively querying the OSM while optimizing the result relevance estimation.
- A dataset of 398 labeled claims and their corresponding tweets (~1.2M) and authors (~773K).
- An automatic application that retrieves relevant data for a given claim.

## Mean Relevance Error



- Convert words to vectors using FastText English model
- Use Euclidean distance between document and tweet



## Iterative Keyword Optimization

