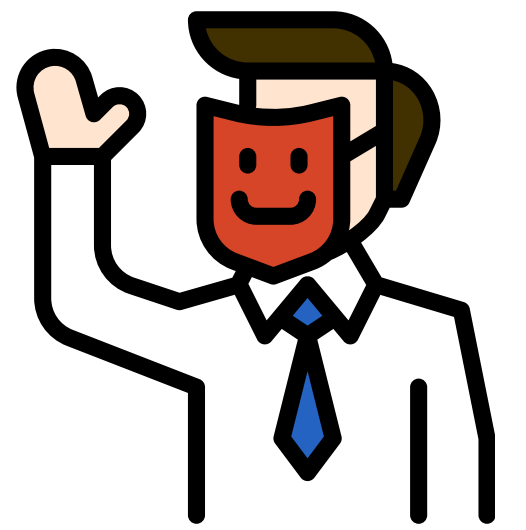




Politechnika
Wrocławska

FACT-CHECKING MUST GO ON!

Jakub Binkowski Denis Janiak Albert Sawczyn
albert.sawczyn@pwr.edu.pl



Introduction

The prevalence of fake news has made it unnecessary to explain it to anyone. Nevertheless, the problem of information verification has not been explicitly solved by artificial intelligence, even for the English language.

Our goal is to develop a system assisting in the process of fact-checking in Polish. As a result, we provide identification of key tasks and a proposal for their solution.

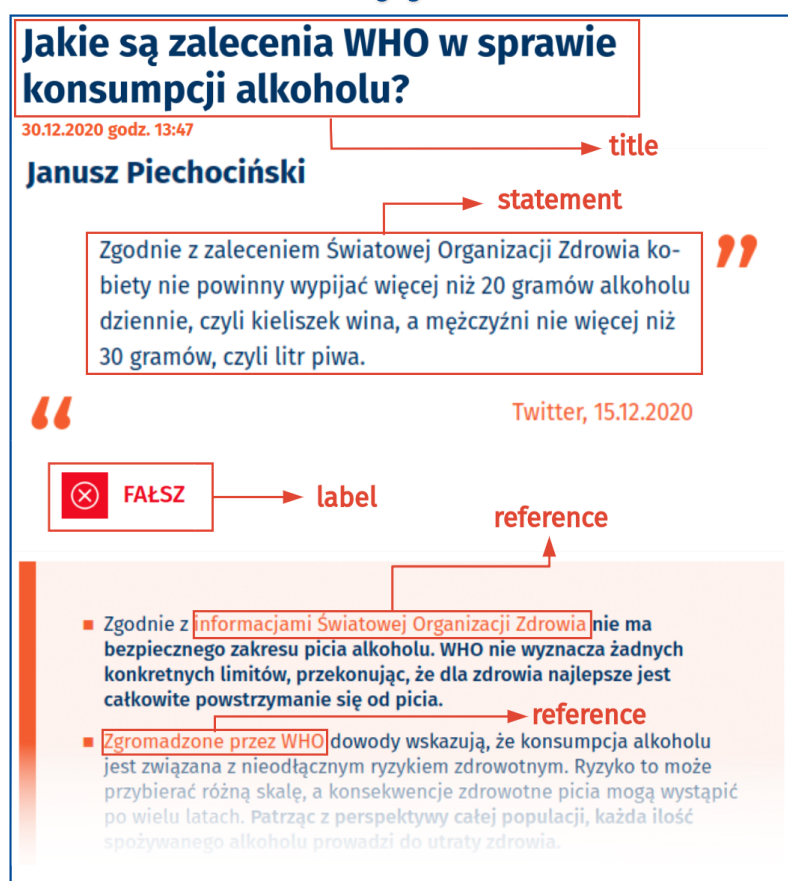
DATA

Data acquisition



- Our custom-designed scraper provided essential data from the website of Demagog fact-checking association, whose main goal is to improve the quality of public debate.
- The dataset was created by experts and contains a sophisticated explanation for each of the judgments.

Schema of Demagog article extraction



4.1k
statements

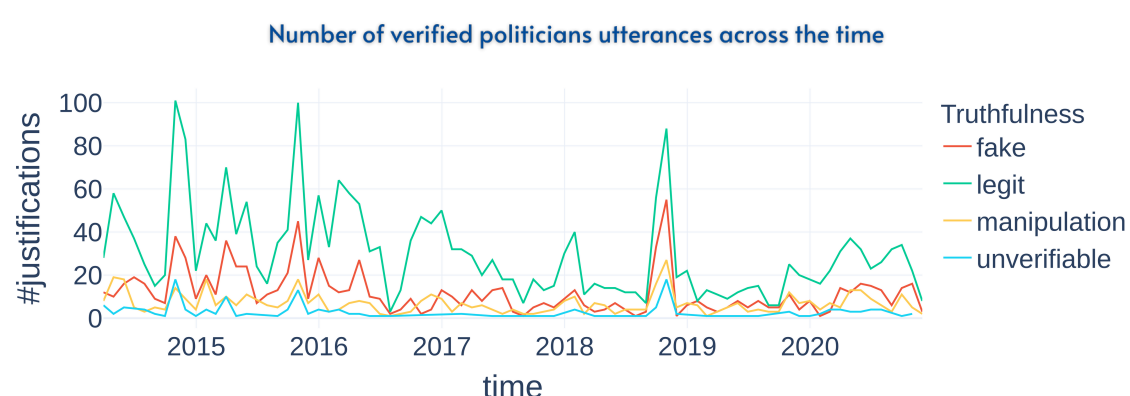
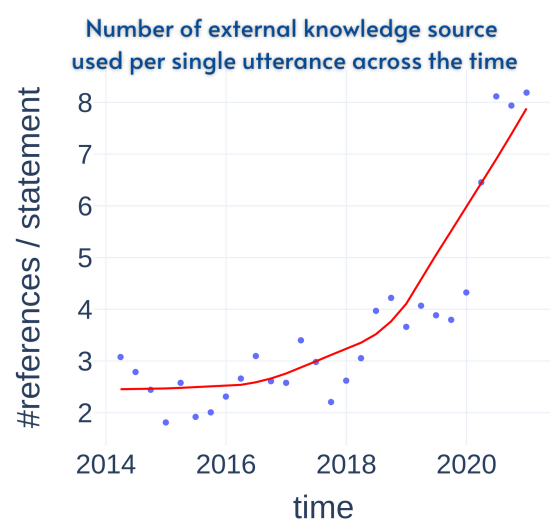
10k
explanations'
references

17.6k
Google Search
references

DO WE NEED AUTOMATION?

The conducted analysis and interviews have revealed the hidden problems of fact-checking. This expertise clarified the most important challenges to be overcome by the algorithms.

- During the elections in Poland, the need for fact-checking dramatically rises, since fake news strongly affects voters.
- Searching for trusted references is crucial and expensive.
- In recent years, number of references in a single verification has strongly increased.
- The automation of this process is the key for being up-to-date with politicians utterances.



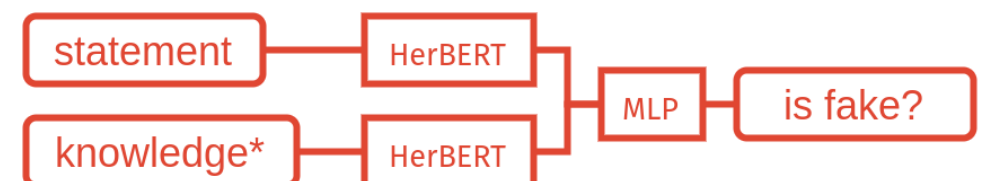
OUR MODELS

Knowledge Retriever



- Evaluating how relevant a document is with respect to the statement being reviewed.
- We use it to retrieve relevant references from an exhaustive knowledge base.
- It can be used as an assistant to a less experienced or overloaded fact-checker.

Fact-checker



*a bunch of documents from Knowledge Retriever or results from Google Search API

- Evaluating the veracity of a statement based on the necessary knowledge provided on the input in the form of a collection of documents.
- Two approaches to knowledge were developed during the research. The first was to provide html/pdf documents, which was the target of the previous model. The second one was to deliver documents obtained from Google Search API after searching by the Demagog article title.

RESULTS

Metrics



		ACC	PR	REC	F1
FACT CHECKING	RETRIEVER	0.76	0.76	0.76	0.76
	KNOWLEDGE BASE	0.44	0.31	0.79	0.44
	GOOGLE SEARCH	0.46	0.32	0.78	0.45

Conclusions & Further work



- At this stage, the presented Knowledge Retriever has achieved satisfactory results, which warrants its further use as an assistance to fact-checkers.
- Given the many assumptions we have used and the untested ideas that remain, we believe that the effectiveness of Fact-checker model can be significantly improved.

Are you interested? Check it out!
smart-wust.ml

