

FIGHTING DISINFORMATION IN THE ERA OF LARGE LANGUAGE MODELS

Preslav Nakov

Abstract. *We will explore the potential of using Large Language Models (LLMs) to fight disinformation. We will present some tools recently developed at MBZUAI:*

- (i) OpenFactCheck (<https://openfactcheck.com/>), a framework for fact-checking any text,*
- (ii) LM-Polygraph (<https://github.com/IINemo/lm-polygraph>), a tool for predicting an LLM's uncertainty in its output using cheap and fast uncertainty quantification techniques, and*
- (iii) LLM-DetectAIve (<https://github.com/mbzuai-nlp/LLM-DetectAIve>), a tool for machine-generated text detection.*

Preslav Nakov,
Mohamed bin Zayed University of Artificial Intelligence,
Abu Dhabi, United Arab Emirates
Corresponding author: preslav.nakov@gmail.com