



TRACES

AuTomatic
Recognition of
hum**A**n-written and
deepfake-generated
text disinformation
in so**C**ial m**E**dia
for a low-re**S**ourced
language

Project Concept & Objectives



- Addressing disinformation for its **intent** (not just untruth/fakeness).
- New Natural Language Processing methods for detecting (**traces**) of intentionally-spread **human and textual deepfake disinformation** for low-resourced languages
- Use-case: **Bulgarian** (a country with critical needs)
- **Interdisciplinary methods:** AI, Machine Learning, Journalism, Linguistics, Psycholinguistics, Natural Language Processing
- **In line with several European Union regulations** (GDPR, AI Act, EU Code of practice on disinformation, EU Charter of fundamental rights, European Convention of Human Rights).
- **Raise public awareness** through an educational event (winter school)

Expected Project Results



- Released ML models, scripts for Bulgarian
- New **large-scale datasets** for Bulgarian on health (Covid) and politics (elections)
- First **TRL4 tool for highlighting intentional disinformation** in Bulgarian to assist journalists and fact-checkers



TRACES

- **Guidelines for other low-resourced languages**
- Discovery of whether there is any deepfake disinformation in Bulgarian social media
- New psycholinguistic and linguistic resources for Bulgarian
- Raised society's skills to recognize disinformation (website, video, Twitter, winter school, media articles, research publications)

SOFIA UNIVERSITY
ST. KLIMENT OHRIDSKI



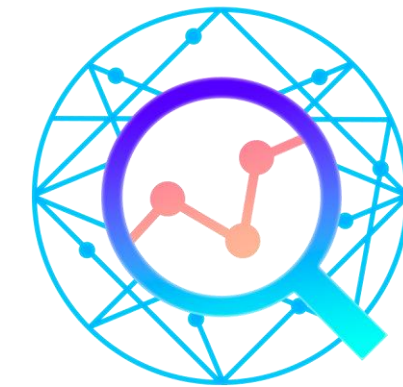
AI4media

GATE
big data for smart society



Our expectations from AI4Media

- Helping to identify potential expansion to other countries with similar needs
- Assistance with (indication of requirements) for integrating the resulting tool/models into AI4EU, AI4Media, Truly Media and TruthNest



TRACES



Team

- ★ **Institution:** Big Data for Smart Society Institute (GATE), part of Sofia University “St. Kliment Ohridski” (the oldest University in Bulgaria).



- ★ **Intelligent Government Department**
- ★ **Team:** complementary areas, track-record of working together



Dr. Irina Temnikova (Principal Investigator)
NLP, media bias, Psycholinguistics



Prof. Milena Dobрева
Media and information literacy, Big data



TRACES

Ruslana Margova

Journalist/news editor
Ph.D. student on the language of Bulgarian fake news



Prof. Ivan Koychev
Deep learning methods for fake content detection,
Machine Learning, Artificial Intelligence



Hristiana Nikolaeva
Bulgarian linguistics, disinformation



Silvia Gargova
Deepfake detection, NLP



Ivo Dzhumerov
Linguistics, Disinformation detection in social media

SOFIA UNIVERSITY
ST. KLIMENT OHRIDSKI



AI4media



GATE
big data for smart society