



Università Ca'Foscari Venezia

PROJECT ACRONYM AND TITLE: IMSyPP – Innovative Monitoring Systems and Prevention Policies of Online Hate Speech

FUNDING PROGRAMME: Rights, Equality and Citizenship Programme (REC)

CALL: REC-AG-2019

SCIENTIFIC FIELDS: Rights, Equality

HOST DEPARTMENT: DAIS – Dipartimento di Scienze Ambientali Informatica e Statistica

SCIENTIFIC RESPONSIBLE: Fabiana Zollo

FINANCIAL DATA:

| Project total costs | Overall funding assigned to UNIVE |
|---------------------|-----------------------------------|
| € 520.210,03 | € 85.368,88 |

ABSTRACT:

Hate speech is spreading not only on social media, but also in the comments of online media news sites. Different ways of presenting news trigger different, sometimes inappropriate, responses from target audiences. The objectives of the project are to automatically detect hate speech in several languages, to identify triggers, to recommend effective counter narratives, and to propose policies to the European communication regulators.

The consortium consists of top research teams in the areas of Artificial Intelligence and machine learning (JSI), misinformation spreading (UNIVE), and hate speech counter narratives (UCY). Effective deployment and long-term sustainability will be ensured by an experienced SME (Textgain). The policies proposed will be disseminated to the European Regulators Group (ERGA), where the Italian communication regulator (AGCOM) is heading a task force.

The technical part of the project is data-driven. We closely collaborate with two prime national media and collect online content in Italian (Corriere della Sera) and Slovenian (RTV SLO public TV). We will also include English and Dutch social media contents from Twitter, Instagram, and Reddit. The data will be manually annotated for types of hate speech, and quality carefully controlled. We will develop hate speech models by applying advanced machine learning methods. The models will be deployed in online settings and used to analyze the large amounts of data already collected. This will allow to analyze mechanisms and trends in hate speech, and identify different types of triggers. We will conduct a pilot study on possible counter narratives and develop recommendations.

The dissemination activities of the project will target European regulators (ERGA), online media companies and journalists, and the scientific community. We will organize training activities for civil servants and journalists. There will be a communication campaign to raise awareness among the general public and youth.

| Planned Start date | Planned End date |
|--------------------|------------------|
| 2020, TBD | 24 MONTHS |

PARTNERSHIP:

| | | | |
|-----------|---|------------|--------------------|
| 1. | INSTITUT JOZEF STEFAN, the Coordinator | SLO | Coordinator |
| 2. | UNIVERSITA CA' FOSCARI VENEZIA | IT | Partner |
| 3. | UNIVERSITY OF CYPRUS | CY | Partner |
| 4. | Autorità per le garanzie nelle comunicazioni | IT | Partner |
| 5. | TEXTGAIN BVBA | BE | Partner |