



# SELMA

## Shaping speech and text technologies for media monitoring & the newsroom

### Objectives

1. **Scalability:** *Massive* processing of audio/video/text data streams
2. **Research:** Unsupervised multilingual language models in a shared space for 30 languages
3. **Research II:** Knowledge transfer across tasks and languages
4. **Decision-Making:** Enable media monitoring analytics for decision-making
5. **Editorial Work:** Enable multilingual content production workflow
6. **Use feedback:** Fine-tune deep learning models from user feedback
7. **The platform:** exploitation of the SELMA platform
8. **Impact:** Dissemination and communication of the SELMA project outputs

### Partners

Deutsche Welle  
Avignon University (LIA)  
University of Latvia (IMCS)  
Priberam  
Fraunhofer-Gesellschaft (FhG)

### Aim

We aim to build a continuous deep learning multilingual media platform using extreme analytics.

“Our aim is to deliver news content & media data better, faster and more diverse.”

### Use Cases

1. **Media Monitoring** - it helps journalists and media monitors to make sense of huge content streams (Big Data Analysis).
2. **Multilingual Content Production** - it helps newsrooms to enrich audiovisual output through entity recognition, translation and voice-over and thus makes it more accessible.

### Impact

The project builds an open-source, multilingual platform that can process (very) high volumes of content.

### Contact

info@selma-project.eu  
twitter.com/selma\_project  
www.selma-project.eu

