



# EUROPEAN LANGUAGE DATA SPACE



## Language Data and Language Technologies in Romania and for the Romanian language

### Panel Session

Vasile Păiș (ICIA, Romania) - Moderator  
vasile@racai.ro

22-10-2025 LDS Country Workshop, Cluj Napoca, Romania  
<https://language-data-space.ec.europa.eu>



**EUROPEAN  
LANGUAGE  
DATA SPACE**

**Participants  
(all experts in AI and Language  
technologies)**



**Moderator: Vasile Păiș**  
ICIA



**Panelist: Tiberiu Boroș**  
Adobe



**Panelist: Traian Rebedea**  
Nvidia



**Panelist: Corneliu Burileanu**  
Politehnica București



**Panelist: George Suci**  
Beia Consult



**Panelist: Vlad Popovici**  
UBB Cluj-Napoca

# Issues/challenges to be discussed

- Importance of language data
- Romanian language datasets
- Romanian language data vs other European language data
- What types of Romanian language data is needed ? How do we compare with other European countries ?
- Language data in the context of Large Language Models (LLMs)
- Romanian Large Language Models (LLMs)
- Problems in acquiring new (high-quality) language data
- Sharing platforms for language data (how to identify Romanian data?)
- Data generated with LLMs in Romanian language: quality, how to detect them ?
- Q&A with audience