



EUROPEAN LANGUAGE DATA SPACE



Language Data Production, Management and Market Development: Overcoming Obstacles

Panel Session

Miroslav Zúmrík (Ľ. Štúr Institute of Linguistics of Slovak Academy of Sciences, Slovakia) – Moderator
miroslav.zumrik@korpus.juls.savba.sk

07-11-2024 LDS Slovakia Workshop, Bratislava
<https://language-data-space.ec.europa.eu>



**EUROPEAN
LANGUAGE
DATA SPACE**

Participants

Moderator: Miroslav Zumrík

Ľ. Štúr Institute of Linguistics of Slovak Academy of Sciences

Panelist: Filip Bednárík

Essential Data

Panelist: Andrej Greguš

Nettle.ai

Moderator

- We are aware of the need of language data and language models
- What are the main obstacles or challenges preventing the availability of data?
- Do policy level and legal frame have an impact on this? Any specific points?
- What can be done to improve the current situation and to facilitate the sharing in Slovakia?

Industrial Panelists

- There is a need to improve productivity and competitiveness at both local and European levels
- Difficulty to compete with existing lack of data, in particular for Slovak
- The European Language Data Space can play a crucial role setting the means to both obtain and provide data in a financially viable manner

Private Sector Data Holders

- Complexity of data monetising, in particular for data holders living on other sorts of income: time consuming, complex, legal risks and obstacles
- Needs and concerns are considerable: need of legal advice
- A single-entry point platform that helps establish links between the different stakeholders, that eases the transactions and that handles data protection and other legal issues is highly needed: the European Language Data Space

Panel Summary: Topics Discussed

- Need for more data for LT and AI-tools development: bottleneck with unavailable data
- What can be done to improve it?
 - We need to ease the full procedure of data sharing
 - We need to make it financially profitable: monetising language data
 - We need to simplify the technical procedure: the LDS platform can achieve this!
- Next steps: start collaborating with the European Language Data Space, be prepared to integrate its data ecosystem



Common European Language Data Space

Thank you!



A Common European Language Data Space – funded under contract LC-01936389 with the European Union.

Miroslav Zumrík (Ľ. Štúr Institute of Linguistics of Slovak Academy of Sciences, Slovakia) – Moderator
miroslav.zumrik@korpus.juls.savba.sk

07-11-2024 LDS Slovakia Workshop, Bratislava
<https://language-data-space.ec.europa.eu>