



EUROPEAN LANGUAGE DATA SPACE



Common European Language Data Space: The User perspective

Manuel Herranz (Pangeanic, Spain)

m.herranz@pangeanic.com

19-03-2025 LDS , Paris, France

<https://language-data-space.ec.europa.eu>

Pangeanic



- *In NLP / Machine Translation since 2009 (“DIY” automated training > neural > fine-tuning)*
- *Involved in Barcelona Supercomputer Center low-resource LLM (Catalan/ES/EN)*
- *Data Annotation & supply, testing to SV, EU projects,*

Why Pangeanic is a Member of the LDS User Group

- **Driving Innovation:** *LDS is a crucial platform for driving innovation in multilingual technologies and AI. By participating, we gain access to cutting-edge resources and help shape EU industry trends.*
- **Collaboration and Networking:** *LDS provides a valuable ecosystem for connecting with other stakeholders, including industry peers, research institutions, and public administrations.*
- **Shaping the Future:** *Actively contribute to the development of LDS, ensuring that it meets the needs of the language industry and its users.*
- **Benefits to Pangeanic:** *Access to a wider range of language data, models, and tools. Opportunities to develop new and improved services.*



pangeanic



pangeanic



Pangeanic

Role: Data Provider & Consumer



1. Produce valuable language data for fine-tuning, base models, applications (speech, text, annotation, etc.): Clear licensing, Data quality, GDPR compliant.
2. Consume language data (corpora, lexicons, etc.) to train MT engines, improve NLP models, and enhance language services, fine-tuning, etc.

Issues LDS addresses or solves

1. Data Accessibility: Breaking down data silos
2. Data Interoperability
3. Monetisation
4. Standard Compliance

Hopes, Expectations, and Demands

1. Robust Infrastructure
2. User-Friendly Interface
3. Clear Governance
4. Strong Security



pangeanic





Common European Language Data Space

Thank you!



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

Manuel Herranz (Pangeanic, Spain)
m.herranz@pangeanic.com

19-03-2025 LDS , Paris, France
<https://language-data-space.ec.europa.eu>