EUROPEAN LANGUAGE DATA SPACE

Common European Language Data Space
The LDS Workshop in Ireland

Jane Dunne (ADAPT Centre, DCU)
jane.dunne@adaptcentre.ie

10-10-2024 LDS Country Workshop, Dublin, Ireland
https://language-data-space.ec.europa.eu

EUROPEAN
LANGUAGE
DATA SPACE

Unleashing the potential of data – for EU businesses and citizens

Context: Language Technologies, Artificial Intelligence (Intleacht Shaorga) and (Large) Language Models

- Language Technologies (LT) and Artificial Intelligence (AI) need language data
- Large Language Models (LLMs) lead now the AI race, needing vast amounts of language data to train them
- Europe’s languages are largly under-resourced, except English
- Concerted effort needed to collect enormous amounts of language data for all European languages

EUROPEAN LANGUAGE DATA SPACE

LDS and the Digital Europe Programme

The Common European Language Data Space (LDS)

- Builds upon previous initiatives focusing on making language data available
- Data Spaces are an inherent part of the EU Data Strategy
- One of the approx. 20 official EU data space projects: strong focus on industry
- Looks to establish the European data market:
  - supporting Technology players and Language Data providers
  - focusing on data sharing and monetising

EUROPEAN LANGUAGE DATA SPACE

Workshop Programme

Time	Session
9:10 – 9:15	**Welcome** Jane Dunne, ADAPT Centre, DCU
9:15 – 9:25	**Introduction** Dr. Brian Davis, ADAPT Centre, DCU
9:15 – 10:45	Ireland in the Digital Age Tony Shannon, Department of Public Expenditure, NDP Delivery and Reform The Digital Europe Programme and the Common European Language Data Space Philippe Gelin, European Commission European Language Data Space: Developing a Market for Language Data and Services and Benefitting from a Joint Europe an Effort Stelios Piperdis, ILSP, LDS Consortium Language Data: The Human Story Alexandros Poulis, Transperfect
11:00 – 11:15	Coffee
11:15 – 11:45	**Session 1: Language Technology and AI in the Private Sector** Using Language Data & AI to drive transformation John Tallon, Storm Technology When Good Data Doesn't Grow on Trees: Challenges for Acquiring High Quality Language Data Dr. Carla Parra Escartín, RWS Language Weaver The Future of Language Technology Stephen Mackarel, Workair
11:45 – 12:15	**Session 2: Language Data and Language Technologies in Ireland and for Irish Public Sector** Towards a Greater Use of Irish in Public Life Dr. Eoin Droighneáin, Department of Tourism, Culture, Arts, Gaeltacht, Sport and Media Using Language Data to Provide New Services Máire Aoibhinn Ní Ógáin, TG4

Time	Session
12:15 – 12:45	**Panel 1: The Role NLP is Playing in Our Digital Society** Moderator: Dr. Abigail Walsh (ADAPT Centre, DCU) Panel: Daithí de Buitléir (DCC), Stephen Mackarel (Workair), Rena Maycock (Chirp), Dr. Joss Moorkens (DCU)
12:45 – 14:00	Lunch
14:00 – 14:50	**Session 3: Language Data Production, Management, and Market Development : Overcoming Obstacles** **Ongoing Data Collection Initiatives in Ireland** eSTÓR - ADAPT Centre, DCU - Teresa Clifford Gaois - Fiontar, DCU - Dr. Úna Bhreathnach Foras na Gaeilge - Edel Ní Chorráin Vicipéid - Amy O'Riordan Open Data Unit – Helena Campbell
14:50 – 15:20	**Panel 2: The Ups and Downs of the Language Data Lifecycle** Moderator: Dr. Ellen Rushe (DCU) Panel: Prof. Lorraine Leeson (TCD), Edel Ní Chorráin (Foras na Gaeilge), Máire Aoibhinn Ní Ógáin (TG4), Dr. Carla Parra Escartín (RWS Language Weaver)
15:20 – 15:30	**Summary and conclusions** Dr. Brian Davis, ADAPT Centre, DCU
15:30	**Coffee and Close**

EUROPEAN LANGUAGE DATA SPACE

Practical Information:

Please share your feedback with us: *Evaluation form*

Do you need a *Certificate of attendance*?

Common European Language Data Space

Thank you and enjoy the workshop!

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

Jane Dunne (ADAPT Centre, DCU)
jane.dunne@adaptcentre.ie

10-10-2024 LDS Country Workshop, Dublin, Ireland
https://language-data-space.ec.europa.eu