The Alliance for Language Technologies (ALT-EDIC)

Fabio Pianesi, Chair of the ALT-EDIC Assembly of Members
(online presentation)
Pianesi@fbk.eu

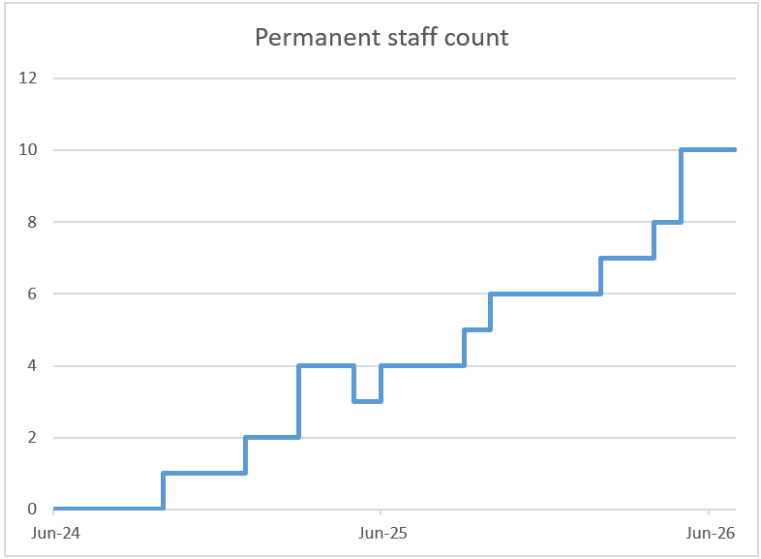
Language Data Space (LDS) Country Workshop
Bucharest, Romania
17 September 2026

The ALT-EDIC in a nutshell

A European Digital Infrastructure Consortium
that supports the **language technology** ecosystem and
contributes to preserve linguistic diversity and cultural richness

Development timeline

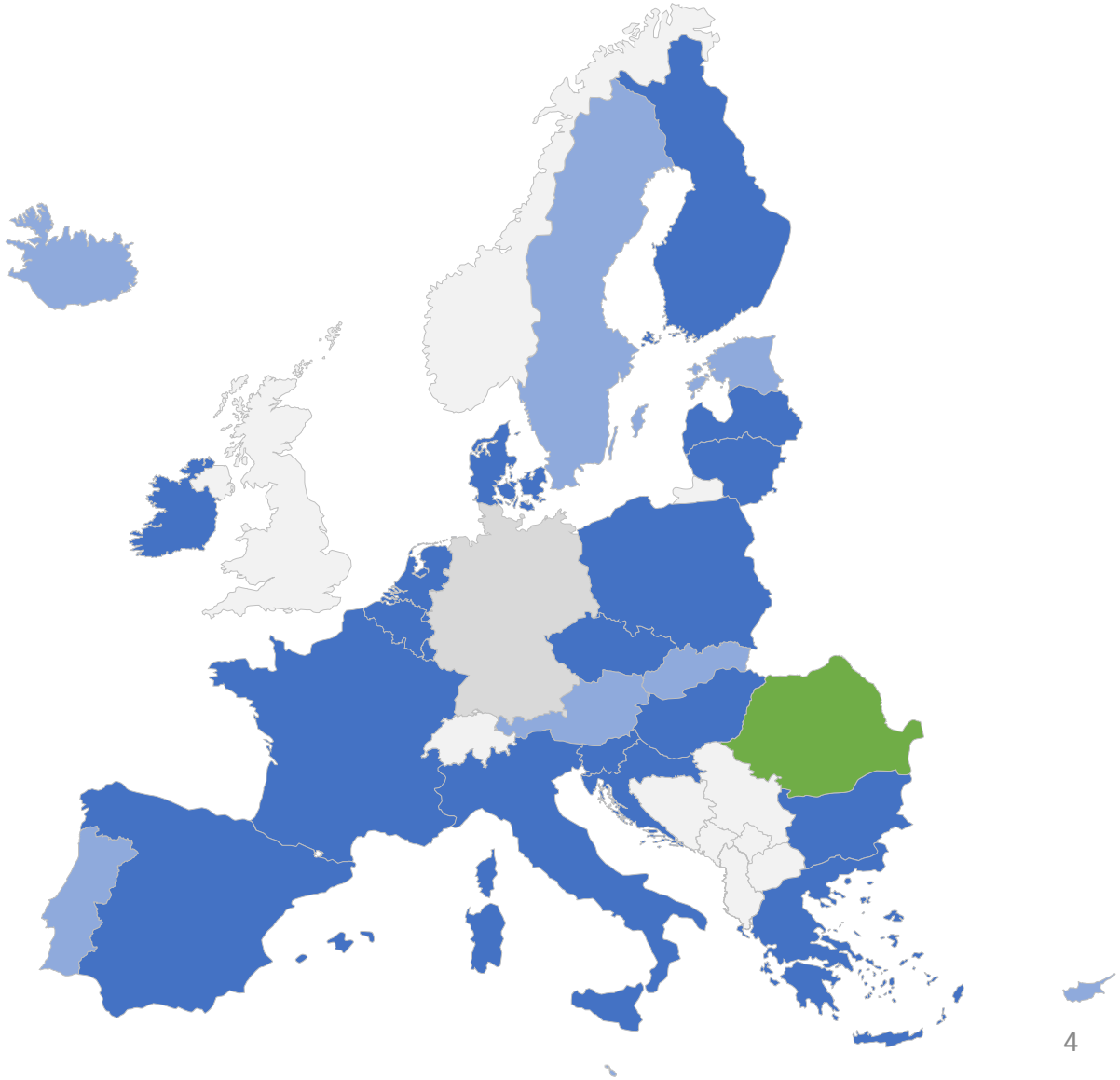
- The Alliance for Language Technologies is one of the first EDICs, proposed by 10 Member States in Dec 2023 and **created by a Commission Decision on 7 February 2024**
  - International organisation with legal personality in each of the EU Member States
- Director in office since January 2025
- 4 EU projects started between January and September 2025 (among which 3 as coordinator)
- Recruitment of the core team under way (10 permanent staff as of today, **~20 persons foreseen in 2027**)



Our members

27 member and observer States
(currently)

- Members: (Belgium/)Flanders, Bulgaria, Croatia, Czech Republic, Denmark, Finland, France, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Poland, Slovenia, Spain
- Observers: Austria, Belgium, Cyprus, Estonia, Iceland, Malta, Portugal, Romania, Slovakia, Sweden



Offices

Statutory seat in Villers-Cotterêts

- Renaissance castle refurbished and inaugurated in 2023
- 1h by train from Paris Gare du Nord, 45’ by car from Paris Charles De Gaulle airport

Key objectives

Advance the European Human Language Technology (HLT) / (Gen)AI community by

- Catalysing and Orchestrating major EU-wide infrastructural initiatives: data, models, computing…
- Connecting and supporting stakeholders (MS, EC, Academia, RTOs, Industry)
- Leveraging Europe's linguistic diversity and cultural richness
- Addressing challenges such as
  - LLMs' cultural and linguistic biases
  - Frugality (data, computing resources, energy...)
  - Explainability
  - Adaptability
  - Resource (data, compute, talent) availability

Main actions

Federate and strengthen the linguistic **data** production

Contribute to the development of **evaluation**, certification and normalisation

Facilitate the adoption of (Multimodal) **Large Language Models**

Support the **development of** (M)LLMs

Provide **advice** to the HLT/(Gen)AI **ecosystem**, including the administration and citizens

Enhancing the availability of relevant linguistic data

- Main activities
  - **Produce** data in a shared manner
  - Develop and maintain open-source **tools** to create and manage data
    - Develop reference **annotation guidelines**
  - Maintain and operate a data **sharing platform**
- Language Data Space (LDS)
  - Part of the Common European data spaces
    - Make more data available for access and reuse, to build the single market for data
    - Agriculture, *Cultural Heritage*, Energy, Finance, Green deal, Health, **Language**, Manufacturing, *Media*, Mobility, Public administration, Research and Innovation, Skills, Tourism
  - Data sharing platform under development under a Commission contract
    - To be transferred to ALT-EDIC in early 2027

Offering trustworthy and relevant evaluation services

- Objective evaluation principles:
  - **Blind testing**: evaluation on unseen data
  - **Transparency**: post-evaluation access to data for result interpretation

  → **Synchronised** evaluations for all participants

  → Creation of **new datasets** for each evaluation

- Data requirements:
  - **Relevant** to real-world tasks
  - Large enough for **statistically significant** results

→ Creation of a Distributed Evaluation Centre

First activity report

- Provides an overview of
  - The organisational structure
  - The activities for 2024-2025
    - Projects launched in the period
    - Community building activities and events
- Published in July 2026

  https://www.alt-edic.eu/wp-content/uploads/2026/07/ALT-EDIC_Activity_Report_2024-2025.pdf

First projects

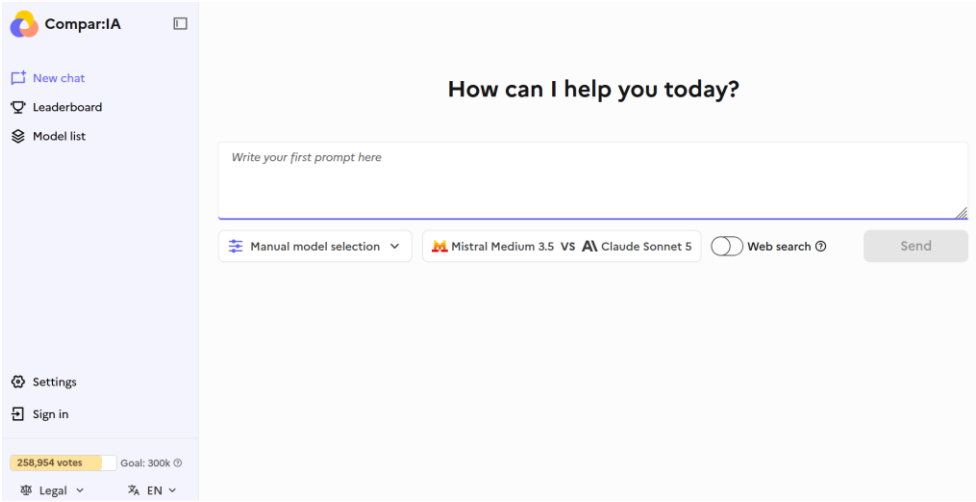
T0	Acronym / Logo	Title	Main objective	Duration	Total budget / EU max contribution	ALT-EDIC share / grant	FSTP calls
01/01/2025	ALT-EDIC4EU	Alliance for Language Technologies for the European Union	Support the **ALT-EDIC construction and consolidation** phase	4 years	4.0 M€ / 4.0 M€	2.3 M€ / 2.3 M€	Unlocking and repurposing datasets, incl. compliance work
01/02/2025	OPEN EURO LLM	Open European Family of Large Language Models	Develop **open-source foundation models**	3 years	37.5 M€ / 20.7 M€	1.3 M€ / 0.65 M€	
01/03/2025	LLMs4EU	Large Language Models for the European Union	Develop LLMs for **5 application domains** (science, public services, tourism, telecom, energy)	3 years	40.0 M€ / 20.0 M€	5.8 M€ / 2.9 M€	Production of data and testing / integration of services / tools
01/09/2025	LLM-BRIDGE	LLM Booster for Regional Innovation and Digital Growth in Europe	Support the ecosystem and in particular startups in their **adoption of LLMs**	3 years	6.8 M€ / 3.4 M€	1.2 M€ / 0.6 M€	Support to incubators and startups
				Total:	**88.3 M€ / 48.1 M€**	10.6 M€ / 6.45 M€	

First Summer School

- 25-28 August 2026 in Villers-Cotterêts
- 25 participants from across Europe
- 12 high-level speakers
- 4 days of exchanges
- Very high satisfaction

European LM Arena

- A platform enabling users to
  - compare the answers of different AI models
  - test the latest models in one place
  - measure their environmental footprint
- Currently in French, English, Danish
  - to be expanded to Swedish, Lithuanian
  - and other EU languages



https://comparia.beta.gouv.fr/en
https://ai-arenaen.dk/

ALT-EDIC Industrial Consortium

- Open to **enterprises and business support organisations** with an interest in Language Technologies and established in the European Economic Area (EEA)
- **Provides guidance on Language Technologies (LT) development priorities** for the ALT-EDIC and elects representatives in the Strategic Orientation Committee, which shall adopt annual or multi-annual strategic orientations
- Offers a **forum to foster collaborations and to elaborate proposals for new projects** on LT, in particular projects involving shared data and comparative evaluations

https://www.alt-edic.eu/industrial-consortium/

ALT-EDIC & Romania

- Romania is an observer
- certSIGN and the Research Institute for Artificial Intelligence “Mihai Drăgănescu” (ICIA) are partners in the LLMs4EU project
- Romanian covered in an LLMs4EU Use Case
- An evaluation campaign for LLMs in Romanian is foreseen in a new project starting soon

Benefits for members

- As a member of the Assembly of Members (AoM), help steer the strategy and operations of ALT-EDIC, e.g.
  - Language coverage
  - Use cases and data
  - Event locations
  - Technical choices
  - Relationship with stakeholders
  - Participation in relevant projects

↓

Members pool resources to jointly support the resulting larger HLT/(Gen)AI ecosystem

www.alt-edic.eu

Thank you for your attention!

Fabio Pianesi – pianesi@fbk.eu

Edouard Geoffrois – edouard.geoffrois@alt-edic.eu

General questions – contact@alt-edic.eu