



CELIA: Common European Language Indication and Analysis - phase 1

Co-funded by the Asylum, Migration and Integration Fund AMIF of the European Commission (HOME-Funds/2023/46). For more information contact CELIA@ind.nl

Authorities spend considerable time and effort investigating the identity and country of origin of asylum applicants, based on their testimonies. Analysing the language variety of asylum applicants helps verify their claimed country of origin. Developing modern language assessment tools can significantly raise the quality of the identification process and contribute to faster procedures.

Ultimate goal: establishing a common process for language assessments accessible to the migration authorities of EU+ countries. The process will be designed to combine the strengths of various analysis techniques (human, semi-automatic and fully automatic) to support asylum decision-making.



Scope for phase 1 (2024-2027)

- method development
- preparation for implementation
- Arabic dialects only

Main products

Validation of (semi-) automatic language analysis

Method development and documentation within the framework set by the EU Artificial Intelligence Act proposal.

Certification of human analysts

Development of peer reviewed training and selection procedure for human analysts.

Standardisation of methodologies

Development of peer reviewed standards and explanatory materials that ensure responsible and ethical use.

Artificial Intelligence Act

AI systems used in migration, asylum and border control management are to be classified as high-risk. This implies strict requirements regarding the quality of data sets used, technical documentation and record-keeping, transparency, the provision of information to users, human oversight, robustness, accuracy and cybersecurity.

Exhibit of expertise

Candidates for the pool of certified human language analysts must follow specified training and submit an exhibit of expertise, showcasing their linguistic knowledge of the relevant dialects. These exhibits of expertise will be peer reviewed by an external body, provisionally called the Linguistic Advisory Panel.

The European Union acquis does not provide a uniform procedure for language assessment. National rules and practices are diverse. Commonly accepted scientific standards are to be agreed for language indication and analysis, similarly to other language forensic methodology.

Written report on the legal and technical requirements for implementation across the EU+ countries

Preparations for a legal basis and a secure digital platform, in compliance with the GDPR.

Knowledge is derived from:

- Over 30 years of experience with traditional human language analyses
- Scientific research on automatic Arabic dialect recognition
- International Expert Group: network of scientists and professionals
- Standards developed for forensic speech evidence

Main frame of reference:

- EU Artificial Intelligence Act (AI Act)
- EU General Data Protection Regulation
- EU Migration Pact
- Best practices in provision of forensic evidence