Project abbreviation: CEFRSERV

Project name: CEFR Labelling and Assessment Services

Project coordinator: Mark Breuker

Project consortium: EDIA

Funding: €137.560

Project duration: 12 months

Main key words: authoring, CEFR, readability, content curation

Goal of the project: Develop a set of data collection and annotation tools to facilitate the creation of data sets for CEFR readability assessment.

Project abstract: This project aims to develop a set of data collection and annotation tools to facilitate the creation of data sets which can be used to train text classification models. These models can automatically assess a text's reading difficulty against the Common European Framework of Reference (CEFR). The ability to accurately and consistently check the readability level of texts is crucial to authors and teachers. It will allow them to create and curate content that meets the needs of students with different backgrounds and skill levels.

EDIA already provides automated readability assessment technology (available as API and authoring tool) for the CEFR which is currently available for English. Through this project, additional languages will be supported (ie. Dutch, German, French and Spanish). As part of the project, we will also build an infrastructure that will pave the way for adding other languages in the future.

Publications:

https://www.edia.nl/edia-blog/2021/3/31/eu-funds-edias-efforts-towards-automated-readability-assessment-in-nl-de-es-and-fr