

EXTRA-TAF

Explainable LLM-based Translation for Tajik and Farsi Languages

Programm / Ausschreibung	KIRAS, Kooperative F&E-Projekte, KIRAS CS Kooperative F&E-Projekte (CS KFE_2025)	Status	laufend
Projektstart	01.10.2026	Projektende	30.09.2028
Zeitraum	2026 - 2028	Projektlaufzeit	24 Monate
Barwert der Projektförderung	€ 460.238		
Keywords	LLM-based translation; low resource language translation; explainable llms; trustworthy and secure llm-based modelling		

Projektbeschreibung

Die zunehmende internationale Vernetzung von Kriminalität, Migration und sicherheitsrelevanten Ereignissen stellt europäische Sicherheitsbehörden vor wachsende sprachliche Herausforderungen. Insbesondere bei ressourcenarmen Sprachen wie Tadschikisch und Farsi besteht ein erheblicher Mangel an qualifizierten Übersetzer:innen und Dolmetscher:innen. Dadurch entstehen Sprach- und Kommunikationsbarrieren, die die Analyse, Dokumentation und Bewertung sicherheitsrelevanter Inhalte erschweren. Diese Problematik wird durch abgeschottete soziale Strukturen und eine geringe Kooperationsbereitschaft in bestimmten Communitys zusätzlich verstärkt und entwickelt sich zu einem sicherheitskritischen Engpass.

Gleichzeitig haben sich große Sprachmodelle (LLMs) rasant weiterentwickelt und neue Potenziale für KI-gestützte Übersetzungssysteme eröffnet. Während für hochressourcige Sprachen bereits leistungsfähige Lösungen existieren, sind Übersetzungsqualität, Robustheit und Verlässlichkeit solcher Modelle für ressourcenarme Sprachen bislang unzureichend erforscht. Gerade im sicherheitsrelevanten Kontext ist dies problematisch, da fehlerhafte oder verzerrte Übersetzungen erhebliche Auswirkungen auf Analyse- und Entscheidungsprozesse haben können und häufig keine Transparenz über Modellentscheidungen besteht.

Vor diesem Hintergrund verfolgt das Projekt EXTRA-TAF (Explainable LLM-based Translation for Tajik and Farsi Languages) das Ziel, ein erklärbares und robustes KI-basiertes Übersetzungssystem für Tadschikisch und Farsi zu entwickeln.

Ein zentraler Innovations- und Forschungsschwerpunkt von EXTRA-TAF ist die systematische Evaluation der Übersetzungsqualität, die quantitative Metriken mit qualitativen Bewertungen durch professionelle Übersetzer:innen kombiniert. Für ressourcenarme Sprachen sind automatische Evaluationsmetriken nur eingeschränkt aussagekräftig, da Referenzdaten begrenzt sind und semantische, pragmatische sowie kulturelle Nuancen oft nicht adäquat erfasst werden. Die qualitative Analyse durch menschliche Expert:innen ermöglicht eine belastbare Bewertung von Bedeutungsgenauigkeit, terminologischer Konsistenz und kontextueller Angemessenheit und ist essenziell für praxisnahe Übersetzungsworkflows für Nischensprachen.

Die enge Verzahnung von Evaluation, Datenkuration und Modellfinetuning liefert gezielte Rückmeldungen für die iterative

Verbesserung der Modelle und schafft die Grundlage für hybride Mensch-KI-Ansätze, wie sie insbesondere im sicherheitskritischen Umfeld erforderlich sind. Ergänzend setzt EXTRA-TAF konsequent auf Explainable-AI-Methoden (XAI), um Übersetzungsentscheidungen nachvollziehbar zu machen, Unsicherheiten sichtbar zu machen und systematische Fehler oder Verzerrungen frühzeitig zu identifizieren. In Kombination mit qualitativen Evaluationsergebnissen entsteht so ein fundiertes Verständnis der Modellgrenzen und Einsatzbedingungen.

Zur Demonstration der praktischen Anwendbarkeit wird eine benutzerfreundliche Demonstrator-Anwendung entwickelt, die Übersetzungsfunktionen sowie erklärbarkeits- und evaluierungsrelevante Informationen wie Konfidenzen und Robustheitsindikatoren visualisiert. Das Projekt leistet damit einen direkten Beitrag zu den Ausschreibungszielen im Bereich KI-gestützter Übersetzungssysteme und adressiert grundlegende Forschungsfragen zu Datenqualität, Robustheit und Erklärbarkeit im Kontext ressourcenarmer Sprachen.

Abstract

The increasing international interconnectedness of crime, migration, and security-relevant events poses growing linguistic challenges for European security authorities. In particular, there is a significant shortage of qualified translators and interpreters for low-resource languages such as Tajik and Farsi. This leads to language and communication barriers that hinder the analysis, documentation, and assessment of security-relevant content. The situation is further exacerbated by socially closed community structures and limited willingness to cooperate in certain communities, resulting in a security-critical bottleneck.

At the same time, Large Language Models (LLMs) have advanced rapidly, opening new potential for AI-based translation systems. While high-resource languages already benefit from powerful solutions, translation quality, robustness, and reliability for low-resource languages remain insufficiently explored. This is particularly problematic in security-sensitive contexts, where inaccurate or biased translations can significantly affect analytical and decision-making processes, and where model decisions often lack transparency.

Against this backdrop, the project EXTRA-TAF (Explainable LLM-based Translation for Tajik and Farsi) aims to develop an explainable and robust AI-based translation system for Tajik and Farsi.

A central innovation and research focus of EXTRA-TAF is the systematic evaluation of translation quality by combining quantitative metrics with qualitative assessments conducted by professional translators. For low-resource languages, automatic evaluation metrics are only partially informative, as reference data are limited and semantic, pragmatic, and cultural nuances are often not adequately captured. Human expert evaluation enables a reliable assessment of semantic accuracy, terminological consistency, and contextual appropriateness, and is essential for practical translation workflows in niche languages.

The close integration of evaluation, data curation, and model fine-tuning provides targeted feedback for iterative model improvement and establishes the foundation for hybrid human-AI approaches, which are particularly required in security-critical environments. In addition, EXTRA-TAF consistently applies Explainable AI (XAI) methods to make translation decisions transparent, highlight uncertainties, and identify systematic errors or biases at an early stage. Combined with qualitative evaluation results, this leads to a comprehensive understanding of model limitations and operational boundaries.

To demonstrate practical applicability, a user-friendly demonstrator application will be developed, visualizing translation outputs alongside explainability and evaluation-related information such as confidence scores and robustness indicators. The project thereby directly contributes to the call objectives in the field of AI-based translation systems and addresses fundamental research questions related to data quality, robustness, and explainability in the context of low-resource languages.

Projektkoordinator

- AIT Austrian Institute of Technology GmbH

Projektpartner

- Bundesministerium für Inneres
- Universität Wien - Zentrum für Translationswissenschaft
- X-Net Services GmbH