

Progetto: Usi innovativi della traduzione automatica e del trattamento automatico del linguaggio per riconoscere traduzione e non-traduzione

Progetto di ricerca

Questo progetto si inserisce in un'intensa attività di ricerca e terza missione che il Centro Dipartimentale su Corpora, Linguistica e Tecnologie del DIT e il Laboratorio di Terminologia e Traduzione Assistita stanno portando avanti con l'utilizzo di metodi computazionali e intelligenza artificiale per la comunicazione multilingue.

I due principali filoni di ricerca riguardano:

Analisi del testo eno-gastronomico

I progetti NoTrans e GastroWiki hanno gli obiettivi di sviluppare delle procedure, sia manuali che automatiche, per produrre spiegazioni in italiano e in lingua straniera di termini complessi e/o intraducibili, specifici del patrimonio intangibile italiano nell'ambito della gastronomia, allo scopo di migliorare la comprensione dei testi e valutarne l'usabilità in contesti di apprendimento (anche incidentale) dell'italiano.

Analisi del testo giornalistico

In ambito giornalistico la traduzione si affianca alla scrittura per produrre testi (ad esempio comunicati stampa) per la pubblicazione multilingue. Riuscire a distinguere i segmenti tradotti da quelli non-tradotti permette di analizzare i meccanismi di creazione del testo giornalistico (framing, discourse news values, evaluation) e le strategie comunicative in un'ottica monolingue, contrastiva e traduttologica. Questo progetto si propone di sviluppare e testare un metodo automatico per il riconoscimento dei segmenti tradotti che permetta la creazione di corpora comparabili e paralleli e lo studio di fenomeni linguistici in cui la traduzione gioca un ruolo rilevante ma non sempre evidente.

Sul filone enogastronomico sono state svolte due tesi di laurea magistrale (Martinelli 2021, Fericola 2021) e prodotto un articolo scientifico (Garcea et al 2023); sono state inoltre svolte presentazioni scientifiche e iniziative di divulgazione su questi temi, come la partecipazione all'Artusiana 2022 e 2023.

Il lavoro di ricerca sul testo giornalistico è oggetto di una tesi di dottorato attualmente in corso di svolgimento (Rodriguez Blanco 2024); è stato inoltre presentato al Convegno UCCTS7 (Bernardini et al 2023) con un intervento selezionato per la pubblicazione in rivista (Bernardini et al in preparazione).

Piano delle attività

L'assegnista di ricerca contribuirà ad entrambi i filoni di ricerca:

(a) Sull'enogastronomia del territorio attraverso l'estrazione semi-automatica, la classificazione e l'integrazione di contenuti provenienti dall'enciclopedia libera Wikipedia. Nello specifico, si occuperà di creare un motore di ricerca sia basato su piatti che su ingredienti per la ricerca di equivalenti gastronomici. Questa ricerca è conseguenza dei modelli a disposizione per l'estrazione e presentazione automatica di definizioni (Garcea et al 2023).

(b) Sul giornalismo sviluppando e testando modelli per il riconoscimento automatico dei segmenti tradotti in testi giornalistici.

Il peso relativo delle componenti potrà essere modificato in funzione del profilo dell'assegnista coinvolto.

La supervisione e formazione dell'assegnista di ricerca saranno affidate al team già operativo formato dalla responsabile scientifica, Prof.ssa Silvia Bernardini e dai Proff. A. Barrón-Cedeño, M. Milicevic Petrovic e A. Ferraresi.

Bibliografia

Bernardini, S. A. Ferraresi, F. Garcea, N. Rodriguez-Blanco. 2023. Corpus approaches to news translation. Can we do better than comparable? Paper presented at the UCCTS2023 Conference, Poznan 10-12 July 2023.

Fernicola, Francesco (2022) Return to the Source: Assessing Machine Translation Suitability based on the Source Text using XLM-RoBERTa. [Laurea magistrale], Università di Bologna, Corso di Studio in Specialized translation [LM-DM270] - Forlì <<http://amslaurea.unibo.it/view/cds/CDS9174/>>

Federico Garcea, Margherita Martinelli, Maja Miličević Petrović and Alberto Barrón-Cedeño. !Translate : When You Cannot Cook Up a Translation, Explain. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2023)*, pages 392–398, Varna, Bulgaria.

Martinelli, Margherita (2022) Methods of Definition Extraction and Linking for Food Recipes. [Laurea magistrale], Università di Bologna, Corso di Studio in Specialized translation [LM-DM270] - Forlì <<http://amslaurea.unibo.it/view/cds/CDS9174/>>