

PATRIZIA COLLESI - Istat pcollesi@istat.it

Il poster presenta la proposta per il multilinguismo dei materiali di diffusione dell'Istat attraverso l'utilizzo di algoritmi proprietari di Intelligenza artificiale (IA)

PROGETTO

Per la comunicazione internazionale dei dati di diffusione, l'Istituto Nazionale di Statistica (Istat) richiede uno strumento affidabile ed innovativo: il **Traduttore multilingue di report statistici**, progettato per generare documentazione di alta qualità in inglese, francese, tedesco e sloveno partendo dall'italiano, lingua ufficiale Istat.

ESIGENZE

Istat deve tradurre e rendere pubblici documenti eterogenei tra loro, contenenti dati statistici, valori quantitativi e tabellari, spiegazioni testuali ed analisi dei dati, con terminologia statistica riguardanti tutti i settori di competenza: ad esempio ambientali, economici e sociali.

INNOVAZIONE

La MT garantisce la gestione della terminologia e del lessico di specialità, la coesione e la coerenza dello stile corporate, per assicurare precisione e coerenza nell'esito traduttivo dei materiali. L'obiettivo è di memorizzare frasi e termini di specialità già tradotte, consentendo di confrontare il testo da tradurre con i materiali memorizzati e fornendo suggerimenti basati su traduzioni precedenti.

QUALITÀ

Punto di forza dell'attività è la qualità dei dati utilizzati per l'addestramento dell'algoritmo: coppie bilingue Italiano-Inglese di materiali validati dal sito Istat.

TECNOLOGIA

Per queste tipologie documentali, il Traduttore combina l'esperienza umana con l'intelligenza artificiale trasformativa, utilizzando algoritmi di IA generativa quali i Modelli del Linguaggio di Grandi dimensioni (Large Language Models), in grado di produrre versioni multilingue e interpretazione dei documenti pubblici, sfruttando la funzionalità della Memoria di traduzione (MT), addestrata su coppie di traduzione.

SOLUZIONE

La soluzione prevede un'interfaccia utente (web app Python sviluppata con framework gradio dotata di logiche specifiche per la gestione ed elaborazione di "coppie di traduzione" a supporto della traduzione e integrata con servizi di Gen AI esterni in ambiente cloud) dove sarà possibile identificare la lingua di interesse tra quelle proposte dall'Istituto e caricare il documento da tradurre insieme alle relative coppie di traduzione, ottenendo il testo tradotto nella lingua di destinazione, tenuto conto delle specifiche indicate sopra.

VERSATILITÀ

Il Traduttore è versatile:

- migliora la qualità delle traduzioni e l'interpretazione dei dati statistici,
- può lavorare su documenti eterogenei
- può essere riutilizzato in diversi ambiti e use cases
- ottimizza le tempistiche di produzione dei report statistici, riducendo le revisioni, consentendo una distribuzione tempestiva delle informazioni e riducendo i costi per la redazione dei documenti.

Pertanto, il punto cruciale sarà la sua interazione con l'intervento umano:

- gli esperti di dominio
- i mediatori

che rivisitano i risultati ottenuti in tempi rapidi dal Traduttore, garantendo uniformità e coerenza nelle rese traduttive.

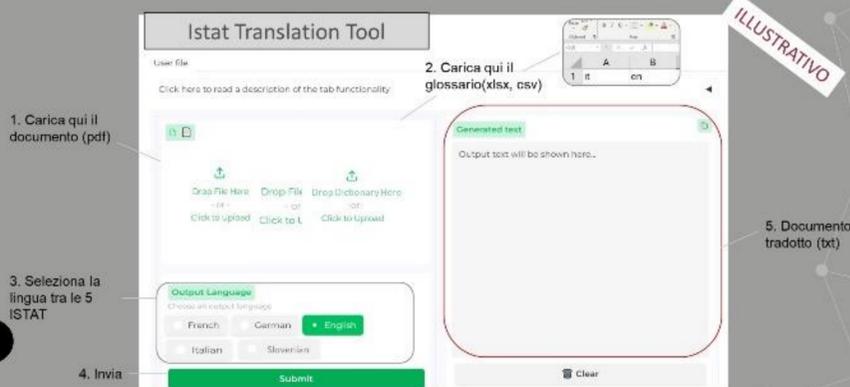
Si utilizza un framework tecnologico che favorisce l'adozione di soluzioni AI



Sistema di Traduzione automatica basato su soluzione KGRAIL + LLM (Large Language Model) esterno



Esempio UX



TEMPI DI REALIZZAZIONE

Dopo una prima fase di raccolta requisiti e materiali validati la sperimentazione partirà con la traduzione verso l'inglese dei primi prodotti, derivati dalle memorie di traduzione e si valuterà l'estensione alle altre lingue.

BIBLIOGRAFIA

Collesi, P., A. Serpente e M.T. Zanola, 2013 (a cura di), Terminologie e ontologie. Definizioni e comunicazione fra norma e uso.
 Collesi, P. et alii. (2021). Verso un glossario unico per la statistica ufficiale italiana, C. Grimaldi e M.T. Zanola (a cura di) Terminologie e vocabolari. Lessici specialistici e tesauri, glossari e dizionari. DOI: 10.36253/978-88-5518-364-2.05
 Ramos, F.P. (2024). Revisiting translator competence in the age of artificial intelligence: the case of legal and institutional translation, The Translator Trainer, 18:2, 148-173, DOI: 10.1080/1750399X.2024.2344942.
 Zanola, M.T. (2018), Che cos'è la terminologia.