Traduzione Automatica di Documenti via Large Language Models - LLMs

In vari ambiti, e in particolare nell'ambito Industria 4.0, vi è una crescente esigenza di adottare sistemi di traduzione automatica efficaci per manuali tecnici e documentazione aziendale. Ciò permette di ridurre i costi associati alla traduzione manuale, migliorare la velocità di distribuzione dei manuali in diverse lingue e assicurare una maggiore precisione e consistenza terminologica nei contesti altamente tecnici. Inoltre, facilita la standardizzazione della documentazione tecnica, elemento fondamentale per garantire la sicurezza operativa, conformità normativa e trasferimento di competenze all'interno di realtà produttive globalizzate.
A questo scopo, viene presentata una soluzione basata sui Large Language Models (LLM), sviluppata per le sfide della traduzione automatica di manuali tecnici. La soluzione integra tecniche di IA e permette di personalizzare la traduzione in base al dominio applicativo, mantenendo la precisione terminologica e l'aderenza ai contesti settoriali.

Laboratory	IN4
Specialization Area	Digitale
Contacts	Cesare Stefanelli, Mauro Tortonesi
Keyword	Traduzione automatica di manuali tecnici, Large Language Models (LMM), Machine Learning (ML), Industria 4.0

"Una Traduzione Automatica Agile, Efficiente e Precisa basata su LLM"

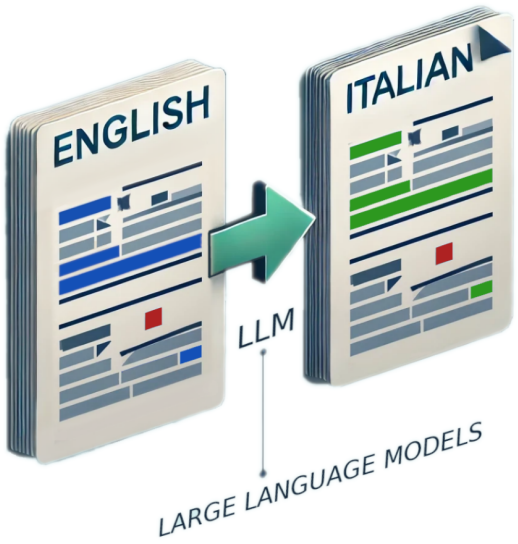


Fig. 1: Traduzione automatica di manuali e documenti tecnici



Fig. 2: Workflow del servizio di traduzione

Description

La soluzione tecnologica presentata consiste in un sistema software avanzato per la traduzione automatica di documentazione tecnica. Il sistema sfrutta approcci innovativi basati su modelli di Large Language Models (LLM) per garantire un'alta qualità delle traduzioni, integrando tecniche avanzate di Machine Learning (ML), in particolare Natural Language Processing (NLP) per l'elaborazione del linguaggio tramite meccanismi di adattamento contestuale. La soluzione è modulare e scalabile, adattabile a esigenze specifiche di aziende in diversi settori industriali.

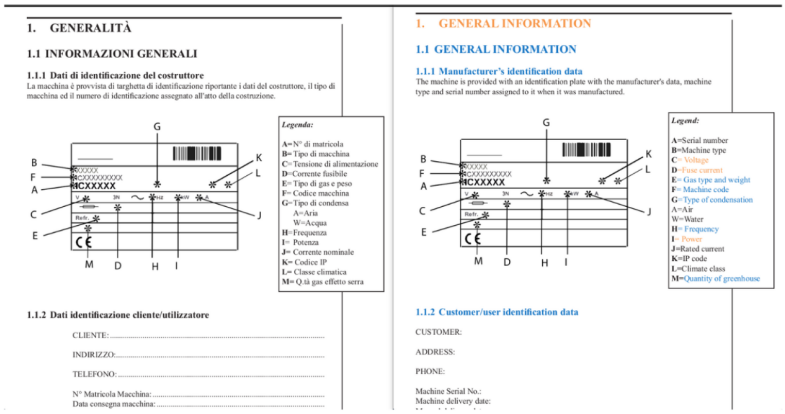
Il processo di implementazione si articola in cinque fasi principali. **(i) Preparazione dei dati**: raccolta e pre-elaborazione dei documenti tecnici, con operazioni di pulizia e normalizzazione. **(ii) Sviluppo di algoritmi personalizzati:** capaci di tradurre testi tecnici in maniera accurata e coerente con il dominio applicativo. **(iii) Utilizzo di molteplici metriche:** per valutare la bontà delle traduzioni ottenute ed eventualmente segnalare quelle da revisionare. **(iv) Addestramento dei modelli linguistici:** ottimizzazione degli LLM, come Llama-3.1, e fine-tuning sul dominio specifico dell'azienda di interesse. **(v) Sviluppo dell'interfaccia utente:** creazione di un'interfaccia che permette la gestione delle traduzioni con strumenti per il controllo qualità delle traduzione. Questo approccio consente di trasformare i complessi processi traduttivi in una soluzione tecnologica efficiente, affidabile e facilmente integrabile.

Innovative aspects

Soluzione software dedicata e completamente personalizzabile sulla base delle esigenze aziendali. La piattaforma presenta dei moduli di traduzione automatica ottimizzati per garantire un'accurata contestualizzazione terminologica nei manuali tecnici, adattandosi dinamicamente ai domini applicativi specifici.

Potential applications

La soluzione proposta trova applicazione in numerosi contesti in cui è necessario gestire e tradurre documentazione tecnica in modo efficiente e accurato. Può essere utilizzata per la traduzione automatica di manuali operativi, guide tecniche, e documenti specialistici, supportando settori quali l'industria manifatturiera, l'automazione, la medicina, l'ingegneria e molti altri settori. La soluzione risulta particolarmente utile per il supporto multilingue nella creazione e distribuzione di documentazione relativa al prodotto o processo a livello internazionale.



1. GENERALITÀ

1.1 INFORMAZIONI GENERALI

1.1.1 Dati di identificazione del costruttore

La macchina è provvista di targhetta di identificazione riportante i dati del costruttore, il tipo di macchina ed il numero di identificazione assegnato all'atto della costruzione.

Legenda:

A= N° di matricola
B= Tipo di macchina
C= Tensione di alimentazione
D=Corrente fusibile
E= Tipo di gas e peso
F= Codice macchina
G= Tipo di condensa
A=Aria
W=Acqua
H=Frequenza
I= Potenza
J= Corrente nominale
K= Codice IP
L= Classe climatica
M= Q.tà gas effetto serra

1.1.2 Dati identificazione cliente/utilizzatore

CLIENTE:

INDIRIZZO:

TELEFONO:

N° Matricola Macchina:

Data consegna macchina:

1. GENERAL INFORMATION

1.1 GENERAL INFORMATION

1.1.1 Manufacturer's identification data

The machine is provided with an identification plate with the manufacturer's data, machine type and serial number assigned to it when it was manufactured.

Legend:

A=Serial number
B=Machine type
C=Voltage
D=Fuse current
E= Gas type and weight
F= Machine code
G=Type of condensation
A=Air
W=Water
H=Frequency
I= Power
J=Rated current
K=IP code
L=Climate class
M=Quantity of greenhouse

1.1.2 Customer/user identification data

CUSTOMER:

ADDRESS:

PHONE:

Machine Serial No.:

Machine delivery date:

Fig. 3: Esempio di Manuale di Macchina di Gelato Carpigiani. Le frasi tradotte vengono inserite con una colorazione per indicare la bontà della traduzione ottenuta, l'arancione indica traduzioni incerte.

Involved partners	• IN4 Hub, Università di Ferrara • Carpigiani Group S.p.a
Implementation Time	4 mesi/persona
Technology Readiness Level	TRL 9 - sistema reale testato in ambiente operativo
Exploitation	IN4 Hub si propone alle aziende come partner nello sviluppo di soluzioni innovative per la traduzione automatica e semi-automatica di documentazione tecnica, grazie alla capacità di integrare modelli linguistici avanzati nel processo di traduzione affrontando le sfide specifiche della traduzione di testi complessi, garantendo elevata qualità, precisione terminologica e coerenza contestuale.

Application example

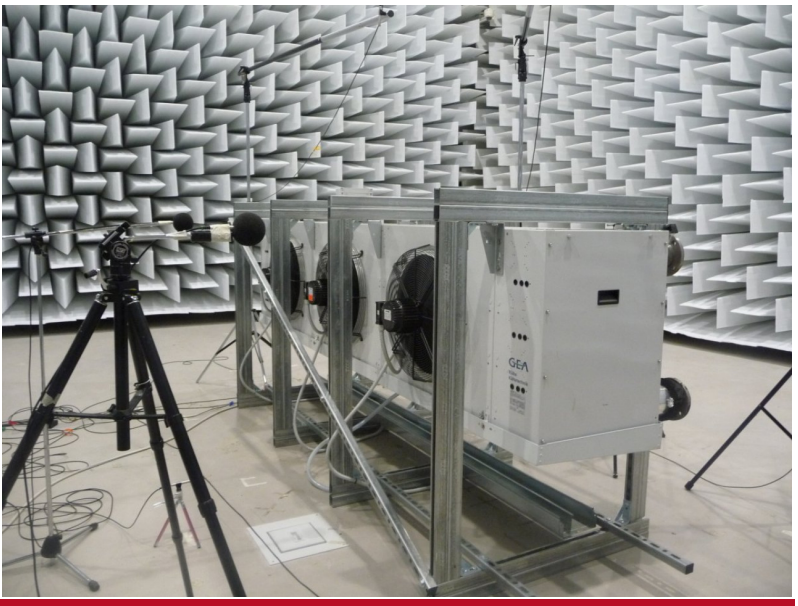
Il progetto ha trovato un'applicazione significativa presso Carpigiani Group, azienda manifatturiera con base in Emilia-Romagna e leader mondiale nella produzione di macchine per gelato. L'esigenza primaria dell'azienda era quella di ottimizzare la traduzione automatica dei manuali di istruzione delle proprie macchine in molteplici lingue, destinati a mercati globali con esigenze linguistiche diversificate.

Grazie al sistema sviluppato, basato su Large Language Models (LLM), Carpigiani è riuscita a integrare soluzioni di traduzione automatica che combinano alta precisione e velocità. Il sistema consente di tradurre documentazione tecnica in più lingue con una coerenza terminologica elevata, rispondendo alle esigenze specifiche dei clienti nei diversi mercati internazionali. La soluzione implementata, ospitata su una piattaforma cloud aziendale, supporta la gestione centralizzata e il controllo qualità delle traduzioni, riducendo significativamente i tempi di produzione dei manuali e migliorandone la fruibilità. Questo approccio ha anche permesso di ridurre i costi associati ai traduttori umani e di automatizzare gran parte del processo, garantendo al contempo una traduzione affidabile e conforme agli standard tecnici richiesti.
Attualmente, il sistema è in fase di estensione per includere altre tipologie di documentazione tecnica, ampliando ulteriormente il suo impatto positivo sull'efficienza operativa di Carpigiani.

Università degli Studi di Ferrara | IN4 innovazione ingegneria integrazione industria

IN4

IN4 Hub per l'Innovazione nell'Ingegneria e l'Integrazione nell'Industria



IN4 (ex MechLav) è il Laboratorio di ricerca industriale del Tecnopolo di Ferrara, che offre un ampio spettro di competenze ingegneristiche, soluzioni e servizi tecnologici principalmente nelle aree della meccanica e dell'ICT.
Il punto di forza di IN4 è dato dal fatto che i suoi ricercatori afferiscono ai Dipartimenti di Ingegneria e di Matematica e Informatica dell'Università di Ferrara, elemento che facilita la collaborazione tra le discipline e la creazione di team work sinergici con competenze complementari. Questo aspetto ha consentito lo sviluppo di numerosi progetti di ricerca industriale altamente innovativi, divenuti accordi di partnership, sia con aziende del territorio, sia con realtà di livello nazionale e internazionale.
Lavorano in IN4 ricercatori specializzati in: meccanica, manutenzione avanzata, acustica e vibrazioni, materiali, macchine a fluido, sistemi energetici, infrastrutture civili, ICT e industria 4.0, intelligenza artificiale, automazione e robotica, elettronica, compatibilità elettromagnetica, reti di comunicazione.
IN4 dispone di importanti attrezzature utilizzate per attività di ricerca e per i servizi alle imprese. A queste si aggiunge una Infrastruttura di Ricerca ed Innovazione comprensiva di una grande camera anecoica acustica e una tavola vibrante multi-assiale.
I servizi offerti spaziano dalla consulenza e formazione, alla caratterizzazione sperimentale, collaudo, progettazione e prototipazione rapida, simulazioni previsionali.

Website	https://mechlav.tecnopolo.fe.it/
Director	Giorgio Dalpiaz
Published on	20/12/2024