

## **English Version (Italian below)**

Translation of just-in-time culturally loaded contents: Machine translation, crowdsourced translation and professional translation for local promotion

### **State of the art**

In the whole of Italy, and in Emilia-Romagna in particular, food constitutes a whole subculture with high economic and social relevance. Restaurants often rely on traditional recipes and local ingredients, and pride themselves in offering a daily changing menu. This type of content – local ingredients, recipes, dish names – is rarely if at all translated, due to the short lifespan of the content (daily menus), the lack of budget for professional translation, and the still limited access to high quality machine translation, particularly by small businesses in this sector. Furthermore, the range of possible target languages continues to expand, due to both international travel and migration from all over the world, making the problem more difficult to solve.

While research on machine translation has produced solutions that nowadays can be leveraged by small and medium enterprises, it has not been applied nor made readily available to the food service sector. Showcase examples of crowdsourced translation have been applied to other domains like music lyrics (Desblache, 2021), and to the problem of low-resource languages (Zaidan and Callison-Burch 2011). Furthermore, preliminary research hints that simply translating dish names and ingredients may not lead to full comprehension and to an enjoyable experience (Gallo et al. 2020), but that additional visual and textual clues should be provided to customers.

### **Aims and expected results**

This project aims to evaluate possible application scenarios for low-cost, high-speed translation within the food service industries, particularly of the Emilia-Romagna region, focusing on one or more case studies and making suggestions for real-life applications of interest for regional actors operating in this sector.

In addition, it will explore the application of the same technologies to the overarching field of tourism, targeting low-resource destinations and attractions, leveraging content generated by local guides and newly established residents and making it available in many languages. As a result of this project, a platform for community-based translation aimed at food service providers (such as restaurants, cafeterias, take-out and delivery establishments) will be created and evaluated through realistic testing in the field. As part of the platform, a customized machine translation engine will be set up and its output evaluated with respect to translations produced in other ways (crowdsourced, professional, out-of-the-box Machine translation systems). Other case studies will be identified based on the needs and interests of regional stakeholders promoting *food tourism* and events.

### **Activity Plan**

The research assistant will carry out the following activities, under the supervision of the Tutor and the national and international experts who collaborate with him:

1. Survey of potential application areas for machine translation to food and tourism industries, particularly those based in Emilia-Romagna.
2. Design, creation and output evaluation of a machine translation engine for food-related content (menus, ingredients, recipes).
3. Creation of a mobile-oriented prototype application for restaurant owners and casual translators to interact and provide just-in-time translations for daily menus.
4. Design and delivery of a field experiment using the app created.
5. Analysis of the results in terms of quality of the translations produced and satisfaction of both business owners and customers.
6. Extension of the approach to machine translation and translation evaluation to other domain of the tourism industry.

### **Competencies and expertise required**

Given the limited time available and the tasks envisaged, the research assistant should have a strong background in natural language processing/computational linguistics, particularly as concerns machine learning and machine translation. Knowledge of machine learning models (deep learning) applied to natural language processing is indispensable, as are the ability to implement them through standard libraries (e.g., scikit, TF, Keras, spacy, huggingface), and the ability to create customized machine translation engines. In addition, prior experience in creating real-world mobile applications is required. (Near-)native knowledge of Italian and English are also indispensable.

### **Training/Supervision**

Training/Supervision will be provided by Prof. Alberto Barrón Cedeño, with support from the national and international experts who collaborate with the Department on research on machine translation and corpus/computational linguistics.

### **References**

Zaidan, Omar F. and Callison-Burch, Chris. 2011. "Crowdsourcing Translation: Professional Quality from Non-Professionals". In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 1220–1229). Association for Computational Linguistics.

Gallo, Danilo et al. 2020. "Restaurant Menu Understanding: Illustrating the Need for Culturally Augmented Translation". *International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '20)*, October 5–8, 2020, Oldenburg, Germany.

Lucile Desblache 2021. "Live music and translation: The case of performances involving singing". *Journal of Specialised Translation*, issue 35.

[Traduzione just-in-time di contenuti culturalmente rilevanti: traduzione automatica, traduzione crowdsourced e traduzione professionale per la promozione del territorio](#)

## **Stato dell'arte**

In tutta Italia, e in Emilia-Romagna in particolare, il cibo costituisce un'intera sottocultura ad alta rilevanza economica e sociale. I ristoranti spesso si basano su ricette tradizionali e ingredienti locali, e sono orgogliosi di offrire un menu che cambia ogni giorno. Questo tipo di contenuto – ingredienti locali, ricette, nomi di piatti – viene tradotto raramente, a causa dei cambiamenti frequenti del contenuto (menu giornalieri), della mancanza di budget per la traduzione professionale e dell'accesso ancora limitato alla traduzione automatica di alta qualità, in particolare da parte delle piccole imprese del settore. Inoltre, la gamma di possibili lingue di destinazione continua ad espandersi, sia a causa del turismo internazionale che delle migrazioni da tutto il mondo, rendendo il problema più difficile da risolvere.

Mentre la ricerca sulla traduzione automatica ha prodotto soluzioni che oggi possono essere sfruttate dalle piccole e medie imprese, non è stata applicata né resa prontamente disponibile al settore dei servizi alimentari. Esempi di traduzione crowdsourced sono stati applicati ad altri domini come i testi musicali (Desblache, 2021), e al problema delle lingue a basso contenuto di risorse (Zaidan e Callison-Burch 2011). Inoltre, la ricerca preliminare suggerisce che la semplice traduzione dei nomi dei piatti e degli ingredienti potrebbe non portare alla piena comprensione e a un'esperienza piacevole (Gallo et al. 2020), ma che ulteriori indizi visivi e testuali dovrebbero essere forniti ai clienti.

## **Obiettivi e risultati attesi:**

Il progetto ha lo scopo di valutare possibili scenari applicativi per la traduzione a basso costo e ad alta velocità nell'ambito della ristorazione, in particolare della regione Emilia-Romagna, focalizzando l'attenzione su uno o più casi di studio e fornendo suggerimenti per applicazioni di interesse reale per le aziende regionali che operano in questo settore.

Il progetto esaminerà inoltre l'applicazione delle stesse tecnologie al settore generale del turismo, in particolare a destinazioni e attrazioni meno conosciute, sfruttando i contenuti generati dalle guide locali e dai nuovi residenti, e rendendoli disponibili in molte lingue. Come risultato di questo progetto, verrà realizzata e valutata una piattaforma di comunità per la traduzione rivolta ai fornitori di servizi alimentari (ristoranti, caffetterie, centri di asporto e consegna) attraverso realistici test condotti sul campo. Come parte della piattaforma, verrà addestrato un motore di traduzione automatica personalizzato e verrà valutato il suo output rispetto alle traduzioni prodotte in altri modi (crowdsourced, professionali, sistemi di traduzione automatica pronti all'uso). Altri casi-studio saranno individuati sulla base delle esigenze e degli interessi degli enti regionali che promuovono il *turismo alimentare* e gli eventi ad esso connessi.

## **Piano dell'attività**

L'assegnista di ricerca svolgerà le seguenti attività, sotto la supervisione del Tutor e l'eventuale collaborazione di esperti nazionali ed internazionali:

1. Indagine sulle potenziali aree di applicazione della traduzione automatica alle aziende di ristorazione e del turismo, in particolare quelle con sede in Emilia-Romagna. industrie, in particolare quelle con sede in Emilia-Romagna.

2. Progettazione, creazione e valutazione dell'output di un motore di traduzione automatica per contenuti legati all'alimentazione (menu, ingredienti, ricette).
3. Creazione di un prototipo di applicazione mobile per gestori di ristoranti e traduttori occasionali per interagire e fornire traduzioni just-in-time dei menu quotidiani.
4. Progettazione ed esecuzione di un esperimento sul campo utilizzando l'app creata.
5. Analisi dei risultati in termini di qualità delle traduzioni prodotte e della soddisfazione sia degli imprenditori che dei clienti.
6. Estensione dell'approccio alla traduzione automatica e alla valutazione della traduzione ad altri settori dell'industria turistica.

### **Profilo in entrata**

Tenuto conto del tempo limitato disponibile e dei compiti previsti, l'assegnista dovrà possedere competenze avanzate negli ambiti del natural language processing / linguistica computazionale, in particolare per quanto riguarda i sistemi di traduzione automatica e i metodi di valutazione della sua qualità. La conoscenza dei modelli di apprendimento automatico (deep learning) applicati all'elaborazione del linguaggio naturale è indispensabile, così come la capacità di implementarli attraverso librerie standard (ad es. scikit, TF, Keras, spacy, huggingface) e la capacità di creare motori di traduzione automatica personalizzati. Inoltre, è necessaria avere esperienza nella realizzazione e distribuzione di applicazioni mobili. Il progetto richiede inoltre la conoscenza (quasi) nativa delle lingue italiana e inglese.

### **Formazione/Supervisione**

La formazione/supervisione sarà affidata al Tutor, Prof. Alberto Barrón Cedeño, coadiuvato dal gruppo di esperti italiani e internazionali che da diversi anni collaborano con il DIT in progetti di ricerca negli ambiti della traduzione automatica e della linguistica dei corpora e computazionale.

### **Bibliografia**

Zaidan, Omar F. and Callison-Burch, Chris. 2011. "Crowdsourcing Translation: Professional Quality from Non-Professionals". In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 1220–1229). Association for Computational Linguistics.

Gallo, Danilo et al. 2020. "Restaurant Menu Understanding: Illustrating the Need for Culturally Augmented Translation". *International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '20)*, October 5–8, 2020, Oldenburg, Germany.

Lucile Desblache 2021. "Live music and translation: The case of performances involving singing". *Journal of Specialised Translation*, issue 35.