

Contenuto archiviato il 2024-06-18



Integration of Machine Translation Paradigms

Risultati in breve

Un sistema linguistico avanzato open-source

Uno strumento avanzato in grado di tradurre la lingua in modo efficiente e in tempo reale offrirebbe un enorme vantaggio alla società. Scienziati finanziati dall'UE hanno proposto un paradigma di traduzione automatica (machine translation, MT) avanzato per migliorare ulteriormente la qualità dei testi tradotti.



SOCIETÀ



© scyther5, Thinkstock

La MT è un settore altamente interdisciplinare e multidisciplinare, che richiede il contributo di professionisti quali traduttori, ingegneri, informatici, matematici e linguisti. Il progetto [IMTRAP](#) (Integration of machine translation paradigms), finanziato dall'UE, si è impegnato a sviluppare e convalidare un sistema MT ibrido open-source.

I ricercatori si sono concentrati sui molteplici aspetti della linguistica come la morfologia, la sintassi e la semantica. Il prototipo di sistema ibrido all'avanguardia creato combina diversi paradigmi MT – come la MT statistica e la MT basata su regole (rule-based MT, RBMT) – e può essere applicabile per qualsiasi coppia di lingue.

I ricercatori hanno introdotto con successo sistemi di base della MT statistica (statistical MT, SMT) dal cinese allo spagnolo e dall'inglese allo spagnolo, attraverso una raccolta di corpora di queste coppie di lingue.

Un altro importante risultato di IMTRAP è stato lo sviluppo del primo sistema ibrido open-source dal cinese allo spagnolo. Il contenuto in entrata di questo sistema è stato pretrattato con un sistema RBMT e quello in uscita da un sistema di SMT. La SMT usa modelli i cui parametri derivano dall'analisi di corpora monolingue e bilingue. È stata usata la RBMT per definire le regole di trasferimento strutturale per le frasi, mentre per il trasferimento lessicale delle parole è stata usata soltanto la SMT. Utilizzando le tecniche della SMT sono stati osservati notevoli miglioramenti nel risultato finale della traduzione.

Inoltre, il risultato di questo nuovo sistema ibrido è stato confrontato con un sistema d'avanguardia di SMT nel test di verifica fuori dominio. I risultati hanno mostrato che il nuovo sistema RBMT supera il sistema SMT in tutti i livelli linguistici, tranne che nel livello sintattico. In particolare, il nuovo sistema ibrido superava di gran lunga gli altri sistemi in termini di copertura lessicale.

Inoltre, IMTRAP ha raggiunto un livello più elevato di ibridazione nella SMT e RBMT. Il lavoro si è anche concentrato sull'estrazione di regole di trasferimento, assegnando una probabilità a una sequenza di n parole, nonché sull'introduzione di un modello di linguaggio per la fase di generazione.

I risultati della ricerca, che è riuscita a raggiungere i suoi obiettivi, sono stati pubblicati in articoli di riviste e libri, nonché attraverso conferenze internazionali. La commercializzazione di un sistema MT ibrido economico avrà una varietà di applicazioni nei sistemi di accesso all'informazione e di traduzione di documenti. La società in generale ne trarrà un enorme vantaggio, come anche i servizi civili e le relazioni internazionali europei, soprattutto con le parti asiatiche, in quanto il progetto si è inizialmente concentrato sulla lingua cinese.

Parole chiave

[Linguistico](#)

[traduzione automatica](#)

[IMTRAP](#)

[lingue](#)

[MT statistica](#)

Scopri altri articoli nello stesso settore di applicazione



Il monitoraggio in tempo reale basato su cloud a casa farà cantare il cuore

31 Ottobre 2019



Equità nell'elaborazione del linguaggio naturale

3 Agosto 2020



Una piattaforma di monitoraggio delle notizie rende il lavoro dei professionisti dei media più facile ed efficiente che mai

21 Giugno 2019



Utilizzare i dati per comprendere e prevedere le tendenze della moda

2 Giugno 2020



Informazioni relative al progetto

IMTRAP

Finanziato da

ID dell'accordo di sovvenzione: 299251

[Sito web del progetto](#) 

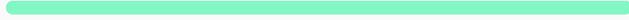
Progetto chiuso

Data di avvio

7 Dicembre 2012

**Data di
completamento**

28 Ottobre 2016



Specific programme "People" implementing the Seventh Framework Programme of the European Community for research, technological development and demonstration activities (2007 to 2013)

Costo totale

€ 173 212,80

Contributo UE

€ 173 212,80

Coordinato da

UNIVERSITAT POLITECNICA DE CATALUNYA



Spain

Ultimo aggiornamento: 8 Maggio 2017

Permalink: <https://cordis.europa.eu/article/id/147448-opensource-advanced-linguistic-system/it>

European Union, 2025