



Rilevamento della lingua

Un esempio: il verbo “argumenter”. DeepL interpreta sistematicamente questo verbo come un sostantivo norvegese al plurale. Per me, invece, si tratta del verbo francese all’infinito. La traduzione varia di poco: “arguments” al posto di “argumenter”. Ma questo altera il senso della frase.

parti del discorso

Quando traduco gli articoli del nostro blog in altre lingue, le frasi vengono suddivise in “stringhe di caratteri”. Una frase può essere suddivisa in più stringhe, soprattutto se contiene un link ipertestuale. In questo caso, trattandosi di un modello linguistico, non è in grado di allineare le stringhe tra loro. Queste vengono elaborate come sequenze di segnali indipendenti l’una dall’altra. Ad esempio, recentemente ho tradotto un articolo sul metodo ABC Learning Design. A un certo punto compare una stringa che in francese termina con «je vous ai présenté la». La stringa successiva inizia con la parola «méthode». In italiano, spagnolo e portoghese, «méthode» è maschile (metodo). La stringa “je vous ai présenté la” diventa “vi ho presentato la” in italiano, ma è seguita da “metodo”, che è maschile. Ciò significa che devo correggere ogni stringa contenente gli articoli femminili e il sostantivo maschile. Non è un dramma, ma ridurrebbe la nostra credibilità agli occhi dei lettori italiani, spagnoli e portoghesi.

Modello predittivo

I grandi modelli linguistici sono sistemi probabilistici. Quando incontrano una stringa di caratteri, valutano la probabilità di ciò che seguirà. Ad esempio, in una frase che recita «la popolazione francese in», il modello calcolerà che c’è il 90% di probabilità che la parola successiva sia «generale». In alcune stringhe di caratteri, la probabilità è così elevata che il modello aggiunge la parola «generale» di propria iniziativa. Tuttavia, dato che il mio contesto è accademico e professionale, la parola successiva è probabilmente diversa. Ma se non rileggo questa sequenza, mi ritroverò con una traduzione in italiano o in portoghese che conterrà la parola “général”, nonostante questa non compaia nella frase originale e non abbia alcun senso né alcuna funzione nella traduzione prevista.

Non capisce il contesto

Una stessa parola può avere diversi significati, a seconda del contesto. Ad esempio, utilizzo la parola «qualification» in un contesto professionale. Significa quindi «idoneità, insieme di competenze legate a un mestiere o a una professione». Ma DeepL traduce sistematicamente «qualification» con «selezione», poiché questa parola significa «essere classificato, essere selezionato» nel mondo dello sport. Fortunatamente, una funzionalità permette di «fissare» il tipo di traduzione desiderato. Tuttavia, questo esempio, come altri, dimostra chiaramente che è necessario rileggere e verificare costantemente le traduzioni. Anche se si risparmia tempo, ciò non significa che non sia necessario farle revisionare da una persona. Ecco perché la legislazione

europea sulla trasparenza dell'IA esenta i contenuti revisionati da un essere umano dall'obbligo di etichettatura previsto per i contenuti gestiti dall'IA generativa.

Uso di uno glossario

Sia in DeepL che in Translate Press esiste una funzionalità che permette di imporre al software di tradurre una parola in un determinato modo. Si tratta del "glossario". È una sorta di dizionario o thesaurus che contiene un elenco di parole e la traduzione scelta. Ad esempio, quando abbiamo creato il MOOC Dys, nella prima parte del corso spiegavamo che i disturbi dell'apprendimento, come la dislessia, la disfasia, ecc., erano difficoltà derivanti dallo sviluppo cerebrale, ma non malattie. Tuttavia, in diverse occasioni, il software ha tradotto "disturbi" con "malattia". Cosa che i partecipanti al MOOC ci hanno subito segnalato. Grazie alla funzionalità "glossario", abbiamo potuto rivedere tutte le traduzioni assicurandoci che tale errore non si ripettesse.

Conclusione

Gli errori segnalati sopra non sono di per sé gravi. Tuttavia, evidenziano il fatto che qualsiasi contenuto generato dall'intelligenza artificiale generativa dovrebbe essere sottoposto a una revisione umana. Inoltre, quando si traduce un testo in un'altra lingua utilizzando l'IA, è meglio avere una conoscenza di base di quella lingua e dei suoi meccanismi. Altrimenti si rischia di perdere ogni credibilità agli occhi del proprio pubblico.

Traduzione automatica

YouTube ha recentemente integrato una funzione di traduzione automatica istantanea basata su un modello di linguaggio (LLM). I risultati sono spesso esilaranti. Alcune parole vengono tradotte in modo errato o addirittura inventate dal modello, al punto che il video originale diventa praticamente incomprensibile. Speriamo che le organizzazioni internazionali non adottino questo tipo di sistema per risolvere i conflitti tra Stati...

Marco Bertolini

Mappa mentale pubblicata sotto [licenza Creative Commons](#) International 4.0 BY-SA.

Le icone utilizzate in questa mappa provengono dal sito Icon8 [Promemoria](#) icona di [Icons8](#)