

Uoc

Introducing ChatGPT

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer follow-up questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.

by OpenAI · Published 16 Nov 2022

La intel·ligència artificial (IA) i la traducció automàtica (TA)

Uoc

Universitat
Oberta
de Catalunya

eLearning
Innovation
Center



Amb l'arribada de **ChatGPT**, ha sorgit interès per la possibilitat d'utilitzar aquesta eina com a traductor automàtic. **ChatGPT té aquesta capacitat i, per tant, volem saber en quins casos seria possible treure'n profit** (principalment entre el català i el castellà, i entre aquests dos idiomes i l'anglès). Quines són les diferències entre les traduccions que surten de ChatGPT i les traduccions que surten de les eines de traducció automàtica que ja coneixem? En quins casos seria millor l'una o l'altra?

01

Quins són els trets diferencials de ChatGPT?



ChatGPT té més capacitat per aconseguir coherència al llarg del text que produeix que no pas les eines de TA. En general, les eines de TA tradueixen frase per frase, sense tenir en compte la traducció de la frase anterior. **ChatGPT, en canvi, analitza tot el text com a conjunt.**



ChatGPT pot oferir diverses traduccions del mateix text original. D'una banda, ens pot generar incertesa si no tenim clar quina versió és la millor, però de l'altra –potser més important–, ens permet refinar la traducció, adaptant-la a la finalitat i/o al públic.



És possible que ChatGPT prioritzi la llegibilitat del resultat per sobre de la fidelitat de la traducció. Per tant, pot fer traduccions que es llegeixen com a textos bons, però que tenen diferències respecte al sentit de l'original.



Hi ha diversos **dubtes legals** envers l'ús de ChatGPT en general: la propietat intel·lectual, els drets d'autor, el plagiat o la protecció de dades.

S'ha comprovat que, a vegades, ChatGPT "al·lucina"

(s'inventa coses). A causa d'aquests aspectes, la tasca de revisar la traducció serà intensiva, ja que les divergències respecte al sentit de l'original poden aparèixer en qualsevol lloc; no hi ha manera de preveure-les.

Quins són els trets diferencials de les eines de TA que ja coneixem?

La TA basada en regles

Com ara **Apertium**, el motor que fa servir traductor.uoc.edu en les traduccions CA<>ES



El fet que la **TA basada en regles** no utilitzi la intel·ligència artificial és bo. Per què? **Perquè això vol dir que no és capaç d'inventar-se coses, vol dir que el text de sortida ha de reflectir fidelment el text d'entrada.** També vol dir que és relativament fàcil introduir millores en el sistema (mitjançant regles i diccionaris); per exemple, sap que una PAC en català és una PEC en castellà.

El traductor de la UOC és capaç de traduir documents i mantenir el format (fins i tot el format especial d'XML de la UOC).



Amb la **TA basada en regles**, podem conèixer les tipologies de coses que tradueix bé i d'errors que fa. **Aquest coneixement de com funciona facilita molt la postedició de la traducció.**

La TA neuronal

Utilitzada en les combinacions CA/ES<>EN per traductor.uoc.edu i també DeepL o Google Translate



Per a combinacions de llengües més allunyades (com pot ser l'anglès i el català), **la traducció neuronal dona resultats més acceptables que els sistemes basats en regles.**



La traducció neuronal pot aprendre; hi ha la possibilitat de reentrenar el sistema amb textos ja traduïts per millorar-lo.

Aquesta millora no necessita un esforç humà, només el temps de computació.



Les tecnologies darrere la TA neuronal i ChatGPT s'assemblen força, i els resultats de traducció de cada eina també. Avui dia, possiblement **l'única raó per substituir la TA neuronal per la traducció amb ChatGPT seria per experimentar** (l'ús de ChatGPT per a la traducció és encara molt nou).



Un dels problemes de la TA neuronal **és la detecció d'errors difícils de veure a les traduccions** (que poden semblar correctes si no et fixes bé en el text de partida).

El mateix problema existeix amb ChatGPT, però la dificultat pot ser més alta, ja **que la fluïdesa és del text sencer, no només de cada frase individual.**



Només utilitzem la TA neuronal en determinats casos, per exemple, **per entendre el sentit general d'un text en un idioma desconegut**, o per tenir l'esborrany d'una traducció quan és acceptable que la traducció sigui força literal. Per a la comprensió general d'un text, **és possible que ChatGPT "al·lucini" més que la TA neuronal, i encara que no sempre sigui així, no hi ha raons evidents per utilitzar ChatGPT en comptes de la TA neuronal.**

Per a l'altre ús, si volem un resultat que es pugui publicar, cal la intervenció d'un traductor humà (una persona amb coneixements alts dels dos idiomes i del context) que postediti la traducció.

També caldria aquesta figura en el cas de la traducció per ChatGPT, però possiblement tindria més feina a causa de la dificultat d'identificar els errors (tal com hem comentat al punt anterior).

Conclusions

En el cas de les combinacions d'idiomes que fan un bon ús de la **TA basada en regles**, com ara català<>castellà, **no sembla que hi hagi arguments per utilitzar ChatGPT com a traductor**, cosa que implicaria una càrrega de feina molt superior en la postedició, i una pèrdua quant a la fiabilitat.

Pel que fa a la **TA neuronal**, també **hi hauria un increment en la feina de postedició si la canviéssim per la traducció amb ChatGPT**. Els possibles beneficis no estan comprovats, i no inclouen una solució per al problema més important: **que no hi ha garanties que les traduccions siguin lliures de falsos sentits**. Així doncs, no veiem clar, de moment, que puguem recomanar la traducció per mitjà de ChatGPT.

No obstant això, estarem atents a l'evolució de les eines d'intel·ligència artificial, que poden ser útils per a la traducció en el futur.



Aquest recurs es basa en un document elaborat pel **Servei Lingüístic de la UOC** en col·laboració amb Toni Oliver, professor dels Estudis d'Arts i Humanitats, el juliol de 2023.

El document original és un document viu i obert a canvis a mesura que vagi avançant aquesta tecnologia. Per tant, aquest recurs s'haurà d'anar actualitzant en paral·lel.

Barcelona
Bogotà
Ciutat de Mèxic
Madrid
Palma
Sevilla
València

Seu central
Av. del Tibidabo, 39-43
08035 Barcelona
(+34) 932 532 300

Tots els centres de la UOC a
seus.uoc.edu

uoc.edu/portal/ca/elearning-innovation-center

 @eLinC_UOC
 #eLinC_UOC
 @UOCuniversity
 @UOCuniversitat

 UOC
 UOC.universitat