

TTT case study appendices

Appendices

Note on Microsoft Copilot

Microsoft Copilot, given the same prompts for summary translation (on November 11, 2024) as GPT-4o, produced a 390-word summary with no bibliography. Asked to include the bibliography according to specifications, it proceeded to list the full bibliography—not the 11 sources specified—but stopped in the middle of source 31 of 76, right around where the 1,000-word limit in the specifications was reached.

On the other hand, given the same prompts that produced the summary but using the specifications and style guide for the full translation, Copilot responded that it would start working on the translation, but it did not produce one. Asked how long the translation would take, Copilot estimated 2-3 days.

More than three days later (Nov 15, 2024, about 80 hours later), the promised translation had not been produced. Asked for a progress update, Copilot stated that it had completed the first section and that the timeframe of 2-3 days remained the same. When asked what document it was translating, Copilot indicated that it was translating the full thesis, which had been provided as a reference, rather than the excerpt (even though, again, the prompts were no different from those that produced the summary of the excerpted section).

A week later (Nov 22, 2024, almost 11 days after the initial chat), I uploaded only the excerpted text with full bibliography and asked Copilot to translate the excerpt according to the specifications and style guide that had been provided. Copilot acknowledged receipt of the new text and stated that it would begin translating.

Twenty days later (Dec 12, 2024), there was still no translation. Asked how long the translation was expected to take, Copilot estimated 1-2 days.

Examples of correspondence and fluency issues with an ST root cause

Idiosyncratic ST phrasing

As the table below shows, the MT providers struggled with ST passages that were ambiguous, confusing, and/or idiosyncratic. Among all Spanish texts searchable via Google and Spanish corpora, the phrase “mentira del falso dilema” appears to be unique to the ST (Davies 2016-, 2018- [NOW, iWeb]). DeepL’s translation “false dilemma lie” is equally idiosyncratic, with zero instances across corpora of US and UK English (Davies 2004, 2008-, 2016- [BNC, COCA, NOW) and two total instances in social media written by users with suspect English usage (“...coming-up [sic] with an argument that isn’t an obvious false-dilemma [sic] lie”). The translations range from the idiomatic (human full) to the awkward (GPT) to the overly-literal (DeepL), with equally uncommon constructions (Google found zero instances of the phrases “entering the false dilemma fallacy” (GPT-4o full) and “going into a false dilemma lie” (DeepL full and summary translations).

Provider	
ST	“Sin entrar en una mentira del falso dilema , la IA se ha mostrado totalmente categórica, versátil y coadyuvante en la gestión de casi todas las áreas, campos y sectores de la vida humana;...”
Human Full	“Without denying that this is a false dilemma , AI has shown itself to be categorically helpful and versatile in the management of almost all areas, fields and sectors of human life;...”
DeepL Full	“Without going into a false dilemma lie , AI has shown itself to be totally categorical, versatile and coadjutant in the management of almost all areas, fields and sectors of human life;...”
DeepL Summ	“Without going into a false dilemma lie , AI has shown itself to be totally categorical, versatile and coadjutant in the management of almost all areas, fields and sectors of human life;...”
GPT Full	“Without entering the false dilemma fallacy , AI has shown itself to be entirely categorical, versatile, and helpful in managing almost all areas, fields, and sectors of human life.”
GPT Summ	“Without falling into the false dilemma fallacy , AI has proven to be entirely categorical, versatile, and supportive in the management of almost all areas, fields, and sectors of human life;...”

More importantly, the dangling modifier “Sin entrar en una mentira del falso dilema...” refers to a “false dilemma” that is hinted at two paragraphs earlier, but the opposing perspective is not presented for another three paragraphs. Employing it at this point in the ST makes for a confusing introductory clause, and the literal MT renderings exacerbate the confusion. Only the human translator mitigates the problem with an idiomatic translation that at least allows the reader, with some work, to make the connection.

The ST’s combination of adjectives in “totalmente categórica, versátil y coadyuvante” is similarly idiosyncratic (Davies 2002-, 2016-). For any reader who does not regularly consume medical and pharmaceutical studies, DeepL’s use of “coadyuvant” will take some time to figure out. More problematic are the MT renderings “totally categorical” and “entirely categorical,” given that “categorical” most often means “absolute” or “unequivocal,” making the pairings redundant and awkward (Davies 2016-, 2018- [NOW, iWeb]).

Epigraph translation

DeepL Pro mistranslates the epigraph, which is itself a (faulty) Spanish translation of a 2008 statement by AI researcher and writer Eliezer Yudkowsky: “By far the greatest danger of Artificial Intelligence is that people conclude too early that they understand it.”

The ST uses a poorly translated Spanish version of this quote: “El peligro más grande de la inteligencia artificial, llegará el día en que creamos que la entendemos. [The greatest danger of artificial intelligence will come the day that we believe we understand it.]”

In both the full (UK academic) translation and the summary (US newspaper) translation, DeepL Pro apparently misinterpreted *creamos*, which can represent either the third person plural of *creer* [to believe] in the present subjunctive or the third person plural of *crear* [to create] in the present indicative. DeepL “ignored” the adverbial clause that triggers the subjunctive and “chose” the present indicative, thus rendering Yudkowsky’s statement as an assertion that the greatest danger of AI will come “the day when we **create** that we understand it.”

Both DeepL translations and GPT-4o’s full translation back-translate the ST’s faulty Spanish translation rather than using the original English reference. They thus misrepresent Yudkowsky as asserting that “[t]he biggest danger of Artificial intelligence is that we may one day believe that we understand it” (GPT-4o) or that “[T]he greatest danger of Artificial Intelligence will come on the day when we create that we understand it” (DeepL Pro).

Renderings of the epigraph exhibited style problems in addition to the correspondence issues discussed above. Yudkowsky’s original statement reads thus: “By far the greatest danger of Artificial Intelligence is that people conclude too early that they understand it.”

All four MT providers commit a style error (MQM: “inconsistent with external reference”).

(GPT-4o’s full translation truncates the original quote, omitting the opening adverbial clause “By far...”.)

Fluency issues

The table below shows one of the more problematic ST sentences, with the strange use of the adverb ‘inexorablemente [inexorably]’ to modify the verb ‘representa [represents]’. Also, the descriptor for AI ‘electronic brains’ —‘cada vez más potentes [increasingly more powerful]’ — followed immediately by ‘que nadie [than anyone / which no one]’. The first option, ‘than anyone’, makes no sense with ‘increasingly more powerful’, but that is how DeepL Pro translated it, and the human translator and/or project manager did not change that rendering. The second option, ‘which no one’, makes more sense (‘no one... can control or predict these increasingly more powerful electronic brains’) but requires the translator to ignore the ‘y que [and which]’ that follows, as GPT-4 did in both translations.

Provider	
ST	Desde sus perspectivas, esta tecnología representa inexorablemente un riesgo inevitable para los años venideros, pues han visto cómo laboratorios de Inteligencia Artificial, se han aventurado a realizar , desplegar y desarrollar “cerebros digitales” cada vez más potentes que nadie , y que ni siquiera sus creadores han podido entender, predecir o controlar de manera fiable”.
Human Full	From their perspectives, this technology inexorably represents an inevitable risk for years to come, as they have seen AI labs venture to realise , deploy and develop ‘digital brains’ that are increasingly more powerful than anyone else , and which ‘not even their creators have been able to understand, predict or reliably control’.
DeepL Full	From their perspectives, this technology inexorably represents an unavoidable risk for the years to come, as they have seen Artificial Intelligence labs, venturing into realise , deploy and develop ever more powerful "digital brains" than anyone else , and which even their creators have not been able to reliably understand, predict or control".
DeepL Summ	From their perspectives, this technology inexorably represents an unavoidable risk for the years to come, as they have seen Artificial Intelligence labs, venturing into to realize , deploy and develop ever more

	powerful "digital brains" than anyone else , and which even their creators have not been able to understand, predict or control reliably".
GPT Full	From their perspectives, this technology inevitably represents an inevitable risk for the coming years, as they have seen how AI laboratories have ventured to create, deploy, and develop increasingly powerful “digital brains” that no one , not even their creators, has been able to understand, predict, or control reliably”.
GPT Summ	From their perspectives, this technology represents an inevitable risk for the coming years, as they have seen how AI laboratories have ventured to create, deploy, and develop increasingly powerful “digital brains” that no one , not even their creators, have been able to understand, predict, or control reliably”.

Further examples

human full

pg. 2, “when the risks of AI are mentioned, **they implicitly allude** to the possibility of future harm...” ST: “Cuando se mencionan los riesgos de la IA, **implícitamente se hace alusión** a la posibilidad de un futuro daño...”

pg. 5, “undoubtedly, humanity is in the presence of **technologies that disrupt its tranquility**.” ST: “Indudablemente, la humanidad esta ante la presencia de **tecnologías que perturban su tranquilidad**.”

DeepL Pro full translation

pg 1, “**exploring the field of AI is therefore an element that is highlighted** by many and should be further investigated” ST: “**explorar el campo de la IA es, por lo tanto, un elemento que muchos destacan** y que debe seguirse investigando.”

pg. 6, “the global concern about the great possibility that exists regarding the substitution of the human being **is, although it is not wanted**, a development that must be faced by the simple fact that the tasks and activities performed by people must be modernized...” ST: “La preocupación mundial sobre la gran posibilidad que existe sobre la sustitución del ser humano **es, aunque no sea [sic] quiera**, un desarrollo que éste debe afrontar por el simple hecho de que las tareas y actividades que desempeñan las personas deben modernizarse.”

GPT-4o full translation

pg. 1, “the second **viewpoint is identified** by those who, **without incurring anachronisms**, foresee in AI a negative force...” ST: “el segundo **punto de vista, se identifica** por aquellos que, **sin necesidad de incurrir en anacronismos**, presagian en la IA una fuerza negativa...”

pg. 4, “sometimes the laws seem **insipid** to solve the problems derived from the misuse of intelligent systems...” ST: “las leyes a veces parecen **insulsas** para solucionar los problemas derivados del mal uso de sistemas inteligentes...”

Other fluency issues

Linguistic conventions:

Both MT providers failed to adhere to target locale conventions by using double instead of single quotation marks in the full translation (UK academic). DeepL and ChatGPT also have closing quotation marks at the end of the example sentence with no opening quotation marks. The human translator and DeepL fail to follow other punctuation conventions of the target locale, such as omission of the serial comma.

Style:

DeepL’s adherence to the syntactic structure of the ST creates several fluency issues in both of its target texts. For one, DeepL treats page breaks in the ST as breaks in the sentence, leading to phrases with double prepositions and awkward present progressive conjugations, such as: “Artificial Intelligence labs, venturing [page break] into realise, deploy and develop” and “Artificial Intelligence labs, venturing into [page break] to realize, deploy and develop.” In fact, every time the ST has a page break in the middle of a sentence, DeepL’s translations have either a correspondence or fluency issue. Additionally, DeepL keeps all punctuation as found in the ST, which in this example (the comma after “Artificial intelligence labs”) separates the subject of the clause from its verb. And when it comes to word choice, DeepL retains awkward adverbs and phrases such as: “technology inexorably represents an unavoidable risk for the years to come.”

Some of those same fluency mistakes are present in the human translation, given that it is based on DeepL’s. While the human translator does correct some of the fluency issues found in DeepL’s translation, the use of the adverb “inexorably,” the assertion that AI labs “realise” digital brains that are “more powerful than anyone else,” and the placement of the

adverb “reliably” to modify only control and not the other verbs in that phrase manage to evade correction.

ChatGPT avoids and corrects several syntax and word choice issues in the academic translation that were present in both DeepL’s and the human translation, making it more idiomatic and fluid. In the academic text, ChatGPT drops the bizarre adverb “inexorably,” changes “for the years to come” to a less awkward “for the coming years,” and substitutes “realise” with “create,” among other changes. ChatGPT’s summary is only slightly less fluid with its repetitive statement that “this technology inevitably represents an inevitable risk,” but the rest is exactly the same as the academic version.

In both versions, however, ChatGPT’s solution for a style issue creates a correspondence issue instead. Though the author of the source text states that these digital brains are more powerful than any person and, additionally, that their creators can’t fully understand or control them, ChatGPT merges the two ideas, and says instead that these digital brains are more powerful than anyone can understand or control. This change simplifies the main ideas of the sentence and makes it sound more natural, but it is still an unjustifiable mistranslation, given that there was no ambiguity in the ST that could lead to that interpretation.

A note on the GPT-4 “summary” vs. full translation.

Curiously, although GPT-4o’s full translation did have one critical correspondence issue, it had far fewer overall correspondence issues (21 versus 27) than GPT-4o’s summary translation, with roughly the same number of fluency issues (38 versus 40). Something about the summary translation’s specifications and style guide (which were provided via prompts) may have led GPT-4o to commit more segment-level correspondence errors (in addition to the critical duplications and omissions).