Translation in the 21st C.: Human Language Translators + Al

Bruce D. Marron January 9, 2025



Is Al a threat to human language

translators?

The answer is perhaps...but let's investigate!

What is Al afterall?

Artificial Intelligence (AI) and Machine Learning (ML)

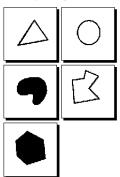
Artificial intelligence: the general ability of computers to emulate human thought and perform tasks in real-world environments.

Machine learning: the technologies and algorithms that enable systems to identify and recognize patterns, make decisions, and improve themselves through experience and data.

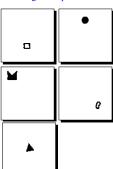
Machine learning is a pathway to artificial intelligence.

Machine Learning (ML)

Training Examples: Class 1



Training Examples: Class 2



Test example: Class = ?



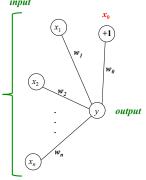
Machine Learning (ML)

Perceptrons as simplified "neurons"

 w_0 is called the "bias"

 $-w_0$ is called the "threshold"

input



Input is
$$(x_1, x_2, \dots x_n)$$

Weights are
$$(w_1, w_2, \dots w_n)$$

Output y is +1 ("the neuron fires") if the sum of the inputs times the weights is greater or equal to the threshold:

If
$$w_1x_1 + w_2x_2 + ... + w_nx_n > threshold$$

then $y = 1$, else $y = -1$

If
$$w_1x_1 + w_2x_2 + ... + w_nx_n > -w_0$$

then $y = 1$, else $y = -1$

If
$$w_0 + w_1 x_1 + w_2 x_2 + ... + w_n x_n > 0$$

then $y = 1$, else $y = -1$

If
$$w_0 x_0 + w_1 x_1 + w_2 x_2 + ... + w_n x_n > 0$$

then $y = 1$, else $y = -1$

A Few Definitions

Acronyms in Al Translation

LLM Large Language Model

NMT Neural Machine Translation [Google]

GPT Generative Pre-trained Transformer [OpenAl]

API Application Programming Interface

SDK Software Development Kit

IDE Integrated Development Environment

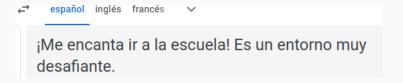
MIME Multipurpose Internet Mail Extension; type of file format

(free with limitations)

Public AI Translators

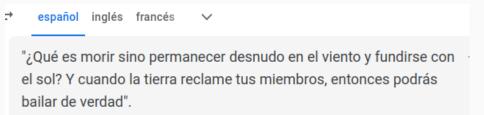
Google Translator - Simple Prose

I love going to school! It's such a challenging environment.



Google Translator - Poetry

"For what is it to die but to stand naked in the wind and to melt into the sun? And when the earth shall claim your limbs, then shall you truly dance." (Kahlil Gibran, artist and poet)



Google Translator - Science Text

The Anthropocene is a proposed new geological epoch based on the observation that human impacts on essential planetary processes have become so profound that they have driven the Earth out of the Holocene epoch in which agriculture, sedentary communities, and eventually, socially and technologically complex human societies developed. The formalization of the Anthropocene as a new geological epoch is being considered by the stratigraphic community, but regardless of the outcome of that process, it is becoming apparent that Anthropocene conditions transgress Holocene conditions in several respects. The knowledge that human activity now rivals geological forces in influencing the trajectory of the Earth System has important implications for both Earth System science and societal decision making. [Steffen et. al. (2018) Trajectories of the Earth System in the Anthropocene, PNAS

Google Translator - Science Text

francés

español inglés

El Antropoceno es una nueva época geológica propuesta que se basa en la observación de que los impactos humanos en los procesos planetarios esenciales se han vuelto tan profundos que han hecho que la Tierra salga de la época del Holoceno, en la que se desarrollaron la agricultura, las comunidades sedentarias y, finalmente, las sociedades humanas social y tecnológicamente complejas. La comunidad estratigráfica está considerando la formalización del Antropoceno como una nueva época geológica, pero independientemente del resultado de ese proceso, se está volviendo evidente que las condiciones del Antropoceno transgreden las condiciones del Holoceno en varios aspectos. El conocimiento de que la actividad humana ahora rivaliza con las fuerzas geológicas en la influencia de la trayectoria del Sistema Tierra tiene implicaciones importantes tanto para la ciencia del Sistema Tierra como para la toma de decisiones sociales.

Google Translator - Legal Text

La Secretaría del Trabajo y Previsión Social deberá pronunciarse respecto de la solicitud de registro dentro de los veinte días posteriores a la recepción de la misma, de no hacerlo, los solicitantes podrán requerirla para que dicte la resolución correspondiente, dentro de los tres días siguientes a la presentación del requerimiento. Transcurrido dicho plazo sin que se notifique la resolución, se tendrá por efectuado el registro para los efectos legales a que dé lugar. [Artículo 15. LEY FEDERAL DEL TRABAJO, publicada 30-09-2024]

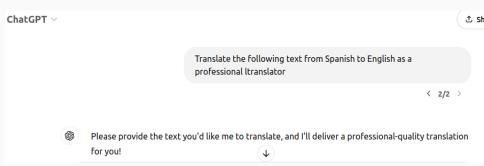
Google Translator - Legal Text



The Ministry of Labor and Social Welfare must decide on the registration request within twenty days of receiving it. If it fails to do so, the applicants may request that it issue the corresponding resolution within three days of submitting the request. After this period has elapsed without notification of the resolution, the registration will be deemed to have been made for the legal purposes to which it gives rise.

The Ministry of Labor and Social Welfare (MLSW) must approve or reject the registration request within twenty days after its receipt in the offices of the MLSW. If the MLSW fails to do so, the applicants may formally require that the MLSW resolve the issue within three business days. If, after this second period has elapsed without notification of the status of the registration request, the registration request will be deemed approved for all applicable legal purposes. [Marron translation]

ChatGPT - Legal Text



The Ministry of Labor and Social Welfare must issue a decision regarding the registration request within twenty days of receiving it. If it fails to do so, the applicants may formally request that the corresponding resolution be issued within three days following the submission of such a request. If this period elapses without notification of the resolution, the registration will be considered completed for all applicable legal purposes.

Registered User AI Translators

(pay per limit)

OpenAl API: GPT-40

GPT-4o: A type of foundational LLM that can generate content in over 90 languages. Multiple types of rate limits apply.

"We're running GPT4 translations on english, german, french, spanish, italian, portugese, on a daily basis, and I can assure you it's the best available option hands down."

- "...the talk of the town with its very impressive language translation technology that's taking multilingual communication by storm. GPT-4 has the remarkable ability to analyze the context and meaning of words and sentences, which makes it able to make more accurate translations"
- "...it generates text 2x faster and is 50% cheaper. Additionally, GPT-4o has the best vision and performance across non-English languages of any of our models. GPT-4o is available in the OpenAl API to paying customers."

Access to GPT-40 API

Once you have applied for access, you must be reviewed and approved...

Once your application has been approved, you will receive an API key \dots

Once you have gained access to the API, you will need to install the necessary SDK (using an IDE) \dots

Once you have configured GPT-4 for your project, you need to initialize GPT-4 for translation....

And now you will be ready to start using the model for language translation.

Yikes!

Google Cloud Translation API

This API uses a pre-trained, custom model or a translation specialized large language model (LLMs). You can choose from different models based on your needs, such as:

Neural Machine Translation (NMT): For general text, like website content or news articles. Available in 194 languages!

Translation Large Language Model (LLM): For conversational text, like messages or social media posts

AutoML Translation: For highly specialized or technical content

NB. Google Translate uses sequence-to-sequence models. These models also are used in voice-activated gadgets, and online chatbots.

Al vs. Human Translators: The

Science

Peer-Reviewed Science

GPT-4 vs. Human Translators: A Comprehensive Evaluation of Translation Quality Across Languages, Domains, and Expertise Levels

Jianhao Yan^{1,2*} Pingchuan Yan^{3*} Yulong Chen^{4*}

Judy Li⁵ Xianchao Zhu⁵ Yue Zhang^{2,6,⊠}

¹ Zhejiang University

³ University College London

⁶ Institute of Advanced Technology, Westlake Institute for Advanced Study

To our knowledge, this study is the first to evaluate LLMs against human translators and analyze the systematic differences between their outputs, providing valuable insights into the current state of LLM-based translation and its potential limitations.

Peer-Reviewed Science

GPT-4 vs. Human Translators: A Comprehensive Evaluation of Translation Quality Across Languages, Domains, and Expertise Levels

ABSTRACT This study comprehensively evaluates the translation quality of Large Language Models (LLMs), specifically GPT-4, against human translators of varying expertise levels across multiple language pairs and domains... we find that GPT-4 performs comparably to junior translators in terms of total errors made but lags behind medium and senior translators....[we] find that GPT-4 translator suffers from literal translations, but human translators sometimes overthink the background information.

Peer-Reviewed Science

GPT-4 vs. Human Translators: A Comprehensive Evaluation of Translation Quality Across Languages, Domains, and Expertise Levels

CONCLUSION ... GPT-4 has made significant strides in approaching human-level translation quality...This suggests promising opportunities for collaboration and enhancement of translation workflows...we anticipate that LLMs will become increasingly valuable tools in the translation industry, working alongside human translators to improve productivity, efficiency, and overall translation quality.

Natural Language Translation

in the 21st C.

Is AI a Threat to Human Language Translators?

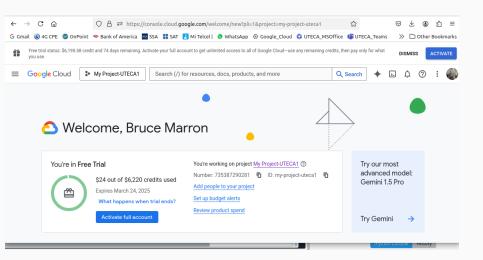
Perhaps, but unlikekly unless folks are unprepared.

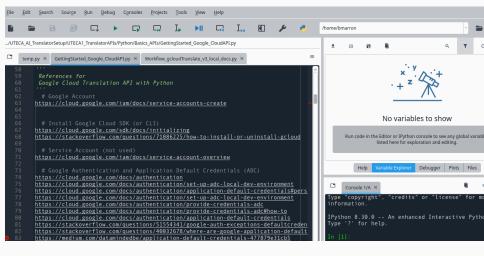
-

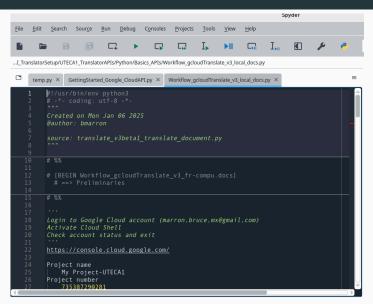
What is likely is that professional translators will increasingly be required to interface with LLMs through subscription-based APIs. This means adding machine language coding to your tool kit.

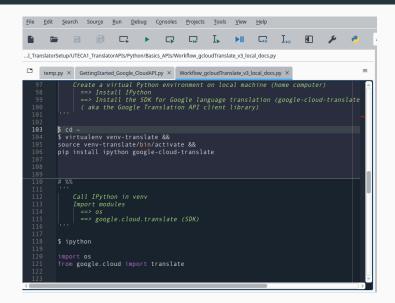
Using Google Cloud Translate API

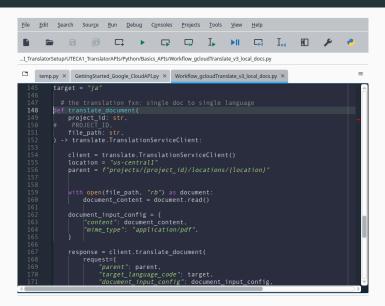
Not trivial!!

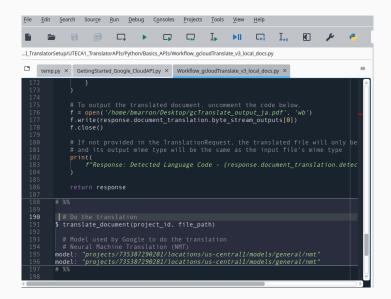














Marron Spring 2025

Course Title: Program and Track: Course ID:

Taller de Traducción Ingeniería y Adelantos Tecnológicos Interpretación y Traducción: 6° Cuatrimestre

Description

This course provides the modern foundations for linguistic translation in the engineering sciences and in the brander field of advanced technologies. The course will force our translation in selected engineering disciplines as well as translation related to advanced and emerging technologies that come primarily from the engineering sciences. The professional translational anapping nods developed in this course will be specific to Spanish and English translations, although such tools and methodologies can readily be extended to the translation of other languages.

Major elements of the course will include: 1) the conceptual framework for the science of engineering, in the structure, function, and dominant partiagnd on selected engineering disciplines; in the state of advanced and emerging technologies in selected engineering disciplines; vi) the elements of style and the use of standard conventions in English v) a review of research English gammatistic structure; vi) Version Control Systems (VCS) and elements of quality assurance and quality custom (QACC).

Scope

1. Conceptual Framework for the Science of Engineering

IT0629

- 1.1 Historical Context
- 1.2 Science vs. Engineering
- 1.3 Major Engineering Disciplines
- Structure, Function, and Dominant Paradigms
 2.1 Thomas Kuhn's "The Structure of Scientific Revolutions"
 2.2 Analysis of Selected Engineering Disciplines
- 3. The Elements of Style in Professional Writing
 - 3.1 Strunk and White's. "Elements of Style"
 3.2 Developing a Professional Voice
- Review of Essential English Grammar
- 4.1 Murphy's, "English Grammar in Use" 4.2 The Yale Graduate Writing Center
- 5. Document QAQC and Version Control Systems
 - 5.1 Introduction to Document QAQC

 5.1 Fundamental Concepts of QAQC
 - 5.1.1 Fundamental Concepts of QAQC 5.1.2 Frameworks for Document QAQC
 - Introduction to Version Control Systems (VCSs)
 Using Git and Github

Course Description in .pdf format Original language: English



Course Description in .pdf format Translated language: French



Course Description in .pdf format Translated language: German



Course Description in .pdf format Translated language: Japanese

Translation Stats

MIME type application/pdf

No. Pages 3

Processing Time 7 sec

Cost per translated doc 8 MXN

Recommendations

The 21st C. Professional Translator





Learn Python The world's Python resources are huge!

Anaconda is an open ecosystem for sourcing, building, and deploying data science and AI initiatives. Spyder is a solid IDE.

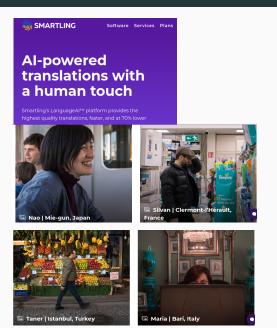
Apply for a Google Cloud Account Trial account gives you 6000 MXN and 90 days

Learn Google Cloud Client Libraries google-cloud-sdk and google-cloud-translate to start

Begin translating with Google's API for Python!

The Future

Smartling



Thank you!