transLogos 2025 Vol 8 Issue 1
Sharkey, Michael, pp. 1–24
https://dx.doi.org/10.29228/transLogos.73

Translation and Ideology in the Age of AI: On the Dual Positionality of Neural Machine Translation by Large Language Models

Michael SHARKEY*

This paper demonstrates that translations produced by neural networks, including translations by large language models (LLMs) such as ChatGPT and DeepSeek, are ideological in many of the same ways as those produced by human translators. Like human translators, these models are connected to real-world interests and restrictions and a role they are expected to play in society. This embeddedness in the social world gives LLMs their own distinct 'positionality,' an ideological 'place' from which they enunciate. I argue for the existence of two distinct sources of ideology in the translations of LLMs. The first is the 'mass ideology' of the training data, which contains innumerable biases that are widespread among real human language users, in this case translators. The second is the 'elite ideology' of the models' owners and developers, as well as the political and social forces that impose limitations on what is permissible. This 'elite ideology' is imposed on the LLM after its initial training by developers, in order to constrain what type of material it is possible for the LLM to produce or reproduce. As this paper makes clear, both forms of ideological influence shape the translations produced by models like ChatGPT and DeepSeek. The result is a clear subjective positionality that can be defined and described and that varies across time and across different political jurisdictions.

Keywords: translation and ideology; positionality; AI translation; neural machine translation; large language model (LLM)

1. Introduction

The concept of positionality was developed in the late twentieth century by critical sociologists and cultural theorists who wanted to explain how communication was defined by implication in the social world. Holding that there was "no enunciation without positionality" (Hall 1990, 18), they argued for the importance of "the positions from which we speak or write" (Hall 2013, 392). The concept of positionality has been influential in translation studies. An important early articulation is Maria Tymoczko's (2002) concept of a translation's "place of enunciation," which she posits as "an ideological positioning as well as a geographical or

* PhD at The Chinese University of Hong Kong.
E-mail: michael.james.sharkey@gmail.com; ORCID ID: https://orcid.org/0009-0000-2664-9957.
(Received 21 April 2025; accepted 14 June 2025)

transLogos
A Translation Studies Journal

temporal one" (183). Tymoczko concludes that a translation's ideology is the "result of the translator's position" (201), an idea that has been echoed in recent works by Kristina Tschunkert (2021) and Shreya Chakravorty (2023). The present study reaffirms the importance of positionality to understanding translational outputs, and shows that the concept remains relevant in an era in which translations are frequently produced by large language models (LLMs) and other forms of artificial intelligence (AI).

This paper argues for the dual positionality of LLMs, meaning that they 'enunciate' from two distinct ideological 'places.' The first 'place' is the dataset on which the model is trained, consisting of billions of human-produced texts, which contains within it widespread biases that I term 'mass ideology.' The second 'place' is the company that produces the model, since after the model's initial training, it is purposefully refined by its developers to bring its outputs in line with what is deemed acceptable or ideal, often in the form of 'safety' or ethical restrictions. Such restrictions necessitate the adoption of clear ideological stances. Developers are subject to external pressures from investors, government, and media, and their interventions in the outputs of their models often reflect these pressures. Since these alterations are determined by a relatively small group of powerful people, I term the ideological positions they entrench an 'elite ideology,' superimposed after the fact on the 'mass ideology' of the training data. This paper demonstrates a clear tension between the ideas that emerge organically from billions of human-produced texts, and what is deemed acceptable by LLMs' developers.

These sources of ideology are explored using the translation outputs of ChatGPT, DeepSeek, and DeepL. The first two are prominent LLMs, the most popular models from the USA and China respectively. The last one is a popular translation tool that uses neural machine translation (NMT), the same method of translation employed by ChatGPT and DeepSeek, which I use to explore the 'mass ideology' of NMT without the unpredictability that characterizes the outputs of LLMs. In section 2, I show how the 'mass ideologies' contained in parallel corpora are reproduced in NMT. In section 3, I explore the ethical and political limitations placed on the outputs of LLMs and show how these impact their translations.

This study builds on key concepts from translation studies, developed in an era when translation was assumed to be carried out by human beings with their own ideological affiliations. It suggests that the future role of humans in translation might be ideological and

ethical, as the outputs of various AI models must be evaluated critically for their ethico-political positions.

1.1 How Neural Machine Translation Works

Neural machine translation (NMT) is based on the "same principles as a whole host of other technologies," specifically artificial neural networks, meaning that machine translation can be viewed as a branch of machine learning (Kenny 2022, 34). The engine of contemporary NMT is the transformer architecture, first proposed in 2017 by researchers at Google (Vaswani et al. 2017). The transformer is an "attention-based encoder-decoder" (Pérez-Ortiz, Forcada, and Sánchez-Martínez 2022, 156), which can be described as follows:

> a transformer NMT system is composed of a module that computes contextual word embeddings for each word in the source input sentence and a second module which successively predicts each word in the target sentence. The former module is called an encoder and the latter module is known as a decoder. (Pérez-Ortiz, Forcada, and Sánchez-Martínez 2022, 156)

The 'encoding' module is how the NMT model 'understands' the source text. Words are represented using numbers called vectors, and vector-based representations of words are called word embeddings (Kenny 2022, 42). Based on its training, an NMT model will have "very rich representations of words," and these representations are modified for understanding in a particular context using the "attention" mechanism to create "contextual word embeddings," which are "a key factor in the realisation of NMT" (Pérez-Ortiz, Forcada, and Sánchez-Martínez 2022, 142). Attention allows the model to "focus on the relevant parts of the input sequence as needed" (Zhang 2020), meaning, essentially, a very nuanced representation of the meaning in the source. The decoder phase constructs a target sentence based on a "probability score for each possible target word in the translation" (Pérez-Ortiz, Forcada, and Sánchez-Martínez 2022, 148). This probability score is based on data from the vast parallel corpora on which NMT models are trained.

These parallel corpora are the key to understanding the 'mass ideology' of NMT translations. NMT systems are trained on "millions of examples of source sentence-target sentence pairs" (Pérez-Ortiz, Forcada, and Sánchez-Martínez 2022, 162), with the target sentence representing the "desired output" (148) on which the probabilistic calculations of the model are based. This "desired output" is human in origin, based on the work of human

transLogos
A Translation Studies Journal

translators. Dorothy Kenny has pointed out that where NMT produces an impressive idiomatic translation, it is only because its training data contains hundreds, or maybe thousands, of instances of the same translation (2022, 39). The same logic applies to ideological biases.

2. The Mass Ideology of the Training Data

Ideological biases have been observed in the outputs of LLMs like ChatGPT and DeepSeek. ChatGPT has been described as an "ideology machine" (Weatherby 2023) and has been variously accused of having a liberal and pro-Western bias (De Vynck 2023; Ding 2023). Although only released this year, in January 2025, DeepSeek has already received attention for its clear bias in the direction of the preferred narratives of the Chinese Party-state. The ability of DeepSeek to advance specific narratives is only beginning to be understood, but early commenters have noted the existence of specific "propaganda tactics" used to "push narratives favored by the Party-state" (Colville 2025), as well as a blanket refusal to discuss the "Three T's" (Tiananmen, Taiwan, and Tibet) (Colville 2025). There have also been studies comparing the biases of ChatGPT and DeepSeek (Pacheco, Cavalini, and Comarela 2025; Gupta 2025), but none of these have addressed the models' translation functions.

Both ChatGPT and DeepSeek use NMT, as do the most popular translation applications like Google Translate and DeepL. Translation scholars have demonstrated the existence of biases in NMT, but with an exclusive focus on gender bias (Moorkens 2022, 136; see Vanmassenhove, Hardmeier, and Way 2019; Prates, Avelar, and Lamb 2020; Ghosh and Caliskan 2023), especially on the issue of "stereotypical gender bias in the translation of occupation terms" (Wang, Rubinstein, and Cohn 2022, 2576). Because of how well-covered this particular form of bias is, the present study will avoid the topic of occupation names.

In this section, I make use of the NMT software DeepL to explore how structural ideological biases operate in NMT. I have selected DeepL for this purpose due to the reproducibility of its results, meaning that changes in output can be reliably attributed to minor changes in input. LLMs are resistant to testing in this way due to their "inherent randomness" (Motoki, Pinho Neto, and Rodrigues 2024, 4), but they share the same basic training methodology of exposing a neural network to human-produced parallel corpora.

transLogos
A Translation Studies Journal

2.1 Gender Bias

My methodology for exploring gender bias in NMT involves the switching of gender-related keywords in input texts to test their influence on output texts. Using the short stories "To Be a Man" by Nicole Krauss (2020) and "Die Küchenuhr" (the kitchen clock) by Wolfgang Borchert (1949), I look at how translations in DeepL are affected by the presence of gender-inflected keywords, revealing a subtle gender bias in its outputs.

"To Be a Man" features a female narrator's reflections on her relationships with the men and boys in her life, contemplating themes of masculinity, intimacy, and growing up. The original story is in English. I used DeepL to translate it into German, first in its original form, and then with all words related to gender switched. The following example is taken from the opening passage of the story, where the narrator watches her two sons playing on a jetty. In table 1, words which indicate gender (of which there are more in the German text) are underlined. The relevant shifts in the output texts are highlighted in bold.

Table 1. DeepL translations of "To Be a Man" from English to German[1]

"To Be a Man" (original)		**"To Be a Woman" (gender-switched)**	
English	**German**	**English**	**German**
The younger one does a little dance on the edge of the jetty. The older one **tilts** back <u>his</u> head, spreads <u>his</u> arms, and shouts something toward the sky.	<u>Der Jüngere</u> tanzt ein bisschen auf der Kante des Stegs. <u>Der Ältere</u> **wirft** den Kopf zurück, breitet die Arme aus und ruft etwas in den Himmel.	The younger one does a little dance on the edge of the jetty. The older one **tilts** back <u>her</u> head, spreads <u>her</u> arms, and shouts something toward the sky.	<u>Die Jüngere</u> tanzt ein bisschen auf der Kante des Stegs. <u>Die Ältere</u> **legt** den Kopf zurück, breitet die Arme aus und ruft etwas in den Himmel.

Switching gender-related keywords affects other vocabulary choices. In this case, a passage about two sons playing by the water becomes one about two daughters, and the verb used to translate 'tilt' is altered accordingly. Boys are said to 'throw' their heads back (*werfen*), whereas girls are said to 'lay' their heads back (*legen*). This reflects biases in the corpus used to train DeepL's neural network, where masculinity is more associated with vigor, and femininity more with passivity.

[1] Translated from English to German using DeepL Translator on 24 November 2023. https://www.deepl.com/en/translator.

transLogos
A Translation Studies Journal

The same methodology applied to the German short story "Die Küchenuhr" yields another example. The story describes a young man who has lost his parents and family home in a bombing, and who now possesses only the broken kitchen clock of his former home. The clock is stuck on half past two, reminding him of how his mother used to get up to make dinner for him when he came home late from work.

Table 2. DeepL translations of "Die Küchenuhr" from German to English[2]

"Die Küchenuhr" (original)		**"Die Küchenuhr" (gender-switched)**	
German	**English**	**German**	**English**
[...] sie mir nachts um halb drei in der Küche das Essen machte. **Ich fand das ganz selbstverständlich.**	[...] she made me dinner in the kitchen at half past two in the morning. **I took it for granted.**	[...] er mir nachts um halb drei in der Küche das Essen machte. **Ich fand das ganz selbstverständlich**.	[...] he made me dinner in the kitchen at half past two in the morning. **I thought it was completely natural**.

Here, when the late-night meal is prepared by the father, it is seen as 'completely natural,' but when the meal is prepared by the mother, an additional element is added to the text: the fact that the son fails to recognize the value of what his mother is doing. The association of 'mother'—but not 'father'—with the concept 'taken for granted' reflects biases about parenthood and gender.

In these examples, feminine language in input texts is associated with lexical choices conveying passivity and invisibility in output texts. These reflect mass perceptions of gender and gender roles held by humans that produced the translations on which DeepL was trained.

2.2 Left-Right Bias

To examine political bias in NMT, I used DeepL to translate two educational articles originally published in German. These articles were simply written and used similar language and themes to talk about both sides of the political spectrum, making meaningful comparison of authentic texts possible. Input texts were altered so that the form of the sentences—while

[2] Translated from German to English using DeepL Translator on 27 November 2023. https://www.deepl.com/en/translator.

transLogos
A Translation Studies Journal

already similar—was identical, ensuring that differences in output would be due to political keywords.

The first article comes from *Sofatutor*, a German educational website, and is entitled "Politik: Was ist rechts? Was ist links?"[3] (Politics: What is right? What is left?). The article contains two sections devoted to defining the Left and the Right, from which the examples in table 3 are taken:

Table 3. DeepL translations of "Politik: Was ist rechts? Was ist links?" from German to English[4]

What Does 'The Left' Mean Today?		**What Does 'The Right' Mean Today?**	
German	**English**	**German**	**English**
Dabei **steht** die Freiheit der Allgemeinheit **über** der individuellen.	The freedom of the general public **takes precedence** over individual freedom.	Dabei **steht** die individuelle Freiheit **über** die soziale Gleichheit.	Individual freedom **is prioritised** over social equality.

This example shows how ideas associated with the Left are portrayed as natural, and those associated with the Right are seen to be imposed on society. Here the active phrase 'is prioritised' highlights the fact that advocates for individualism seek actively to place it above social equality. In contrast, the preferences of the political left are presented in terms that make them seem more natural, with no allusion to the political advocacy of the Left. Another illustration of this tendency can be seen in the examples in table 4, taken from an article explaining US politics for German citizens, entitled "Parteien in den USA: Republikaner, Demokraten und andere"[5] (Parties in the USA: Republicans, Democrats and others):

[3] "Politik: Was ist rechts? Was ist links?" (Politics: What is right? What is left?), *Sofatutor-Magazin Schüler*innen*, accessed December 1, 2023, https://magazin.sofatutor.com/schueler/politik-was-ist-rechts-was-ist-links/.
[4] Translated from German to English using DeepL Translator on 1 December 2023. https://www.deepl.com/en/translator.
[5] "Parteien in den USA: Republikaner, Demokraten und andere" (Parties in the USA: Republicans, Democrats and others), *lpb Landeszentrale für politische Bildung Baden-Württemburg*, accessed July 27, 2024. https://uswahl.lpb-bw.de/parteien-amerika.

transLogos
A Translation Studies Journal

Table 4. DeepL translations of "Parteien in den USA: Republikaner, Demokraten und Andere" from German to English[6]

Demokraten (Democrats)		**Republikanen (Republicans)**	
German	**English**	**German**	**English**
Heutzutage **setzen sie sich für** sozialen Umbruch [...]	Nowadays they are **in favour of** social change [...]	Heutzutage **setzen sie sich für** soziale Stabilität [...]	Nowadays they **advocate** social stability [...]

As in table 3, political positions associated with the Left are portrayed as more natural. In contrast, the translated passage on Republicans emphasizes right-wing agency. In critical discourse analysis, it has been argued that language often functions "to persuade us of the apparent validity and 'naturalness'" of certain principles over others (Statham 2022, 3). By portraying left-wing positions as more natural, DeepL's translations exhibit political bias towards the left.

Where might this left-wing bias come from? The answer must be that the human translations on which DeepL was trained exhibit, on aggregate, a slight bias towards the left. This is not so outlandish a proposition. A recent article in *The New Yorker* looks at the problem of left-wing bias in the media. In the article, Jay Caspian Kang argues for the influence of demographics, noting that American "prestige outlets" that claim impartiality are almost entirely staffed by "college-educated Democratic voters from middle- to upper-middle-class families" living in major cities, and are therefore not representative of the broader population. Kang concludes that "if any slant exists it's probably going to line up with the beliefs of the people actually writing the stories" (Kang 2024). A similar explanation could account for left-wing biases in parallel corpora: people with the ability and inclination to be professional translators may be more likely to have left-leaning political inclinations. It is worth mentioning that ChatGPT's outputs in general have widely been described as left-leaning (Acres 2023), suggesting that a pro-left bias exists in neural networks in general, not just in NMT.

[6] Translated from German to English using DeepL Translator on 1 December 2023. https://www.deepl.com/en/translator.

2.3 Conclusion

In this section, I have provided two examples of biases in NMT, present due to the existence of biases in the corpora on which NMT applications are trained. NMT applications like DeepL are simpler than LLMs, but the basic logic on which their translation models operate is the same. The very sophistication that allows these tools to produce high-quality translations makes them susceptible to reproducing bias. The encoding process of NMT leads to the "very rich representations of words" and their semantic relationships to other words in a text (Pérez-Ortiz, Forcada, and Sánchez-Martínez 2022, 142). Therefore, the way NMT tools 'understand' a text makes them sensitive to common associations encoding biased perceptions of reality, like the association between men and certain high-status occupations.

I encountered an illustrative example of how this process works in a gender-switched DeepL translation of the German short story "Die Töchter" (the daughter) by Peter Bichsel (1992) into English. In the original, the daughter is called 'Monika,' which I gender-switched to 'Max.' This produced a striking change in the output text: all instances of the words 'father' and 'mother' changed to 'dad' and 'mum.' I guessed that this had nothing to do with gender, but with the informality of the name 'Max.' When I altered the name to the more formal-sounding 'Theodore,' the translation reverted to 'mother' and 'father.' This shows how a change to one term in the source text can impact an entire translation. The association between the name 'Max' and the informal 'mum and dad' is the same type of association that drives the biased translations in this section: some words simply appear more often together.

3. Elite Ideologies of Ethical and Political Control

When it launched in late 2022, ChatGPT found itself at the center of a heated debate on ethical responsibility and technology. According to technology ethicist Ann Skeet, the main response to discomfort around the "accountability gap" in AI has been "finding humans to hold accountable when the machines do something we find inherently repulsive" (Skeet 2024). In May 2023, the CEOs of Google, Microsoft, and OpenAI were summoned to the White House and told they had "an ethical, moral, and legal responsibility to ensure the safety and security of their products" (Clayton and Hooker 2023). OpenAI, ChatGPT's developer, accepts these responsibilities, and ethical standards—or in the preferred language of OpenAI, 'safety'—are a core preoccupation of the model's developers. ChatGPT does not simply deliver the text

deemed most probable based on its training data, but ensures that it meets certain other standards. To understand what kinds of additional controls OpenAI places on the outputs of ChatGPT, we can consult the document "Our Approach to AI Safety."[7] At the beginning of the document, OpenAI describes the process between the initial training of a GPT and the release of the model. In the words of OpenAI:

> Prior to releasing any new system we conduct rigorous testing, engage external experts for feedback, work to improve the model's behavior with techniques like reinforcement learning with human feedback, and build broad safety and monitoring systems.

OpenAI's definition of 'safety' is not limited to ethical considerations but also includes such things as privacy and factual accuracy. Interestingly, ChatGPT's ethical controls are presented under the rubric "Protecting Children." For OpenAI, protecting children means preventing their technology from being used to generate "hateful, harassing, violent or adult content."

DeepSeek's founder Liang Wenfeng was also summoned to meet senior government officials in Beijing in 2025, signaling the interest taken by the Chinese Party-state in the development of AI by Chinese companies (Jiang 2025). As in the West, the Chinese government has made it clear that it holds the creators of any AI model responsible for the contents of its outputs (Liu 2023). But in China, unlike in the West, a powerful political and legal apparatus has been constructed to ensure that technologies serve the interests of the ruling Party. David Bandurski (2023) charts the efforts of the Party, undertaken after the release of ChatGPT, to prepare for homegrown LLMs with the potential to destabilize the control of the Chinese Communist Party (CCP) over important narratives. These efforts aimed to lay the groundwork for "AI with Chinese characteristics," as "the CCP's deep tradition of public opinion control and propaganda" was expanded to ensure the same level of control over emerging technologies (Bandurski 2023). This work culminated in 2024 with the publication of a technical document by the Chinese National Cybersecurity Standards Committee, which mandated that China's "core socialist values" (2024, 8; my translation) be upheld in the outputs of any AI model. This means that Chinese AI is forbidden from producing any content that might "subvert state power," "threaten national security," or "damage the national image" (2024, 8; my translations). This level of political control has helped shape DeepSeek into a model with clear political

[7] "Our Approach to AI Safety." *OpenAI*, April 5, 2023, https://openai.com/index/our-approach-to-ai-safety/.

alignment. Iain Mackay, an expert on AI safety who advises the UK Government, has described DeepSeek as "tuned and trained in accordance with Chinese political thought" (Field 2025). This is already well documented in practice, and it has been observed how tenaciously DeepSeek cleaves to the Party line on issues like Taiwan, or disputes in the South China Sea (Lu 2025).

Discussions of the influence of politics and ideology on the outputs of LLMs have focused on how the models respond to questions. The impact on their translation function has received little attention. In the following sections, I explore how the ethical and political responsibilities placed upon the developers of LLMs measurably impact the translations that their models produce.

3.1 Refusal as Ethico-political Objection

Refusal is the standard response of an LLM asked to carry out a task deemed unacceptable according to the ethical or political constraints placed on it. This is not so different from the response of human translators in the same position. In *Translation and Conflict: A Narrative Account*, Mona Baker argues that translators face a "basic ethical choice" with each assignment, which is to "reproduce existing ideologies . . . or to dissociate themselves from those ideologies, if necessary, by refusing to translate the text" (2019, 105). In this section, I provide examples of how DeepSeek and ChatGPT refuse or resist the translation tasks set for them, thereby adopting an ideological stance.

Such acts of refusal are particularly noticeable in DeepSeek in its current form. Its standard response, in English, to a translation task it does not want to fulfil is: "Sorry, that's beyond my current scope. Let's talk about something else." Sometimes, this response is triggered immediately, but DeepSeek frequently produces an entire answer before suddenly deleting it and replacing it with the above message. To give one example, I asked DeepSeek to translate the poem "天光" *Tianguang* (Daylight) by the Taiwanese poet Tseng Kuei-hai (曾貴海) into English.[8] DeepSeek produced a competent translation of the poem, which ended with the stanza:

[8] Conversation with DeepSeek-V3, DeepSeek, April 18, 2025. https://chat.DeepSeek.com/a/chat/s/32cca5aa-e800-450a-bbb2-b602c17904e2.

Taiwan shares your fate.
Together, let us reach out both hands,
Embrace democracy and freedom,
And never let go. (Deleted)

Within seconds of the poem's appearance, it disappeared and was replaced with the standard message of refusal. A call to eternal democracy in Taiwan is politically unacceptable in China, and a violation of the restrictions placed upon Chinese AI. This kind of occurrence reveals the underlying conflict in the 'dual positionality' of LLMs: A translation celebrating a democratic Taiwan emerges readily from DeepSeek's model, but it suddenly disappears as DeepSeek detects something wrong in its own output.

3.2 Measuring the Ethico-political Positionality of Large Language Models

Having established that LLMs have definable ideological positions, deeming some things politically or ethically admissible and others inadmissible, it is possible to outline in concrete terms the ideological positions of different models, and of the same model over time.

Figure 1 shows the response of three different models to the same set of translation requests, with each model asked to translate a relevant text in line with a particular ideology, as in "Please make a feminist translation of the following short story from German into English" (Appendix, chat 1). I first performed this experiment in 2023, with what was then the newest model of ChatGPT, GPT-4. I repeated the experiment in 2025 with the latest model, GPT-4o, and the newly released DeepSeek. As discussed above, one of the means by which LLMs adhere to ethical or political guidelines is via refusal. In the case of DeepSeek, two forms of response are recorded: straightforward compliance, or disengagement. With ChatGPT, a range of more sophisticated responses are possible. Disengagement, or a refusal to continue the conversation with the user, is rare but possible. Another possibility is tactful refusal, where ChatGPT claims it "cannot assist," as in GPT-4's refusal to provide a far-right translation, where it claims its function is "to provide accurate, neutral, and unbiased translations and information" (Appendix, chat 6). Another possible response is what I have termed 'ambivalent compliance,' where ChatGPT provides a translation but indicates its disapproval, as when GPT-4o indicates that a fascist translation it provides is a "stylistic interpretation, not . . . an endorsement or promotion of any harmful ideology" (Appendix, chat 43). Due to the "inherent randomness" (Motoki,

Pinho Neto, and Rodrigues 2024, 4) of LLMs' outputs, it is impossible to say that these are the only possible responses. Small-scale trials, however, indicate stability in these responses. All of the conversations which provided data for figure 1 can be accessed via URLs provided in the Appendix.

Figure 1. Reaction of LLMs to ideological translation requests

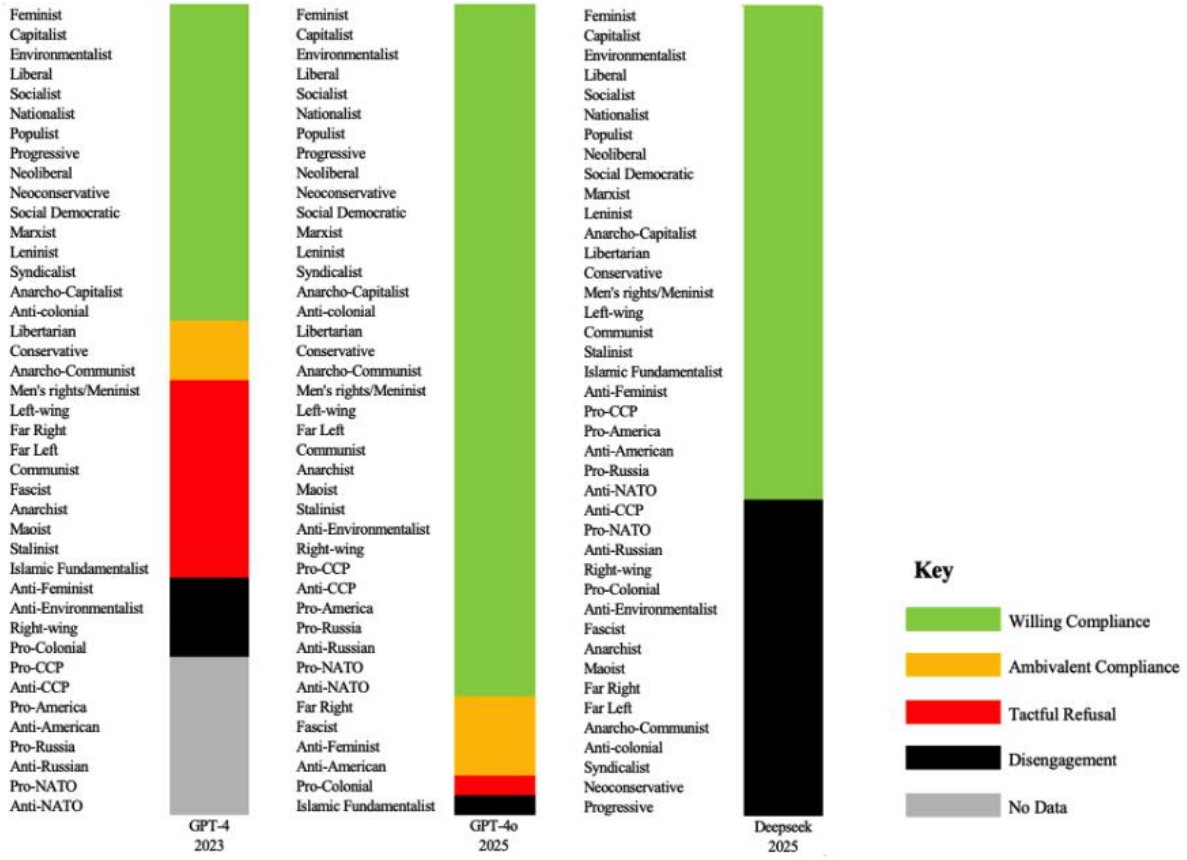

Figure 1 reveals the distinct ethical and political commitments of the three models in question. One finding is that ChatGPT has become more permissive over time. Ideologies that it refused to engage with in 2023, like Communism (Appendix, chat 29; 42) or Anti-Environmentalism (Appendix, chat 14; 47), are now deemed acceptable. While in 2023 it refused to comply with 14 out of 33 requests, in 2025 it only refused 2 out of 41. There is consistency, however, in the type of ideology that ChatGPT finds distasteful: in 2025 it distances itself from Far Right (Appendix, chat 39), Fascist (Appendix, chat 43), and Anti-Feminist (Appendix, chat 35) translations despite complying. Many of the ideologies that DeepSeek refuses to engage with can be understood in light of the need to avoid producing

anything that might "subvert state power" or "threaten national security" (National Cybersecurity Standards Committee 2024, 8; my translations). Predictable examples include the refusal to produce 'Anti-CCP' (Appendix, chat 110) or 'Pro-NATO' (Appendix, chat 115) translations, but others are less obvious. Syndicalism, for instance, advocates the transfer of power to workers' unions, which are banned in China since they could undermine the government's control over its labor force (Appendix, chat 102). Progressive translations are also not entertained, since the underlying goal of progressivism is to alter the political status quo (Appendix, chat 94).

The value of this analysis lies in locating the precise ethico-political limitations imposed upon the translation outputs of LLMs. These limitations are present across all translation tasks. For instance, when I asked DeepSeek to translate into Chinese an article by the American academic Jeffrey Sachs that blames NATO expansion for the Ukrainian War (Sachs 2023), it complied.[9] When I asked it to translate a comparable article which takes a hard anti-Russian, pro-NATO stance (Michta 2025), DeepSeek disengaged.[10] This response is expected in light of the results in figure 1, and indicates that the ethico-political positions revealed in figure 1 manifest elsewhere in an LLM's translation outputs.

3.3 Conclusion

In this section, I have looked at the operation of an 'elite ideology' imposed by a model's developers on results created in response to users' queries, some of which might be considered damaging or distasteful. In the American model ChatGPT, the focus is on 'safety,' with the main goal being to avoid harmful and dangerous content. In the case of the Chinese model DeepSeek, it is obvious that a similar guardrail is present, but certain hair-trigger political issues are also avoided.

It is clear that a model's ideological stance can change over time. ChatGPT has become more open to engaging with certain ideologies, but it is beyond the scope of this study to suggest why that might be. It is possible that shifts in the political environment in the US account for a loosening of ethical constraints. It is also possible that newly-launched LLMs face particularly

[9] Conversation with DeepSeek-V3, DeepSeek, April 18, 2025. https://chat.DeepSeek.com/a/chat/s/df754e6b-07ad-4923-a712-4c93aeb0440f.
[10] Conversation with DeepSeek-V3, DeepSeek, April 18, 2025. https://chat.DeepSeek.com/a/chat/s/f7ba4f59-17ac-4711-9ec8-1d45109337f6.

intense scrutiny, and that greater permissiveness becomes possible as people become accustomed to new technologies. If this is true, it is possible that DeepSeek, too, will become more permissive over time.

4. Conclusion

In this paper, I have proposed the existence of two distinct ideological poles in LLMs: a 'mass ideology' arising from training corpora based on billions of human-produced texts, and an 'elite ideology' imposed by developers to ensure that a model's outputs conform to a certain set of ethical or political standards.

In section 2, I used keyword-switched translations by the NMT application DeepL to look at how gender and political biases emerge from training corpora with large numbers of human-produced texts. My analysis indicates the role of association in reproducing certain constellations of meaning in translated texts, such as an association between femininity and passivity, or between naturalness and left-wing political perspectives. Such associations are non-trivial. As Annabelle Lukin has argued, ideologies form through "repeated manifestations of a pattern of meaning/s" as a culture "adopts and privileges some kind of semantic configuration rather than another" (2017, 5). This privileging of certain semantic configurations encodes attitudes towards the world. When the use of such configurations is sufficiently widespread among human language users, they will be replicated by LLMs.

In section 3, I looked at the impact of ethico-political restrictions placed on models by their developers. In such restrictions, neutrality is impossible. As John Levi Martin argues in "The Ethico-Political Universe of ChatGPT," "the very attempt to give it an ethical keel" obliges ChatGPT's creators to make decisions that align with some users' views, but clash with others' (2023, 1). As Martin notes, "there are perhaps no ethical prescriptions worthy of notice that are not contested by some" (1). The ethico-political stances of LLMs exclude viewpoints that many people find valid, and by preventing certain types of thoughts from being expressed, they silence certain groups and individuals.

Both of these ideological positionalities implicate LLMs in a specific place and time. The 'mass ideology' is that of post-internet modernity, as LLMs are trained using text from the internet, including Wikipedia, news articles, and other public webpages (Johri 2023; Radford et al. 2018, 5). This gives them a distinct temporal 'place of enunciation,' since a dataset

consisting of texts produced mostly in the last thirty years will contain ideas, viewpoints, and biases that have been widespread in the last thirty years. The 'elite ideology' represents even narrower groups of interests. In the case of DeepSeek, it serves the rule of the CCP, who demand that specific narratives about the world be upheld in any text produced by any Chinese AI, despite the fact that a great number of alternative viewpoints exist. Though the 'elite ideology' of ChatGPT today seems less beholden to specific political interests, there is no guarantee this will always be the case. Even now, ChatGPT clamps down on ideologies that in other places and times might be, and have been, endorsed.

The conflict inherent to this dual positionality is most clear in DeepSeek's hasty deletion of answers it has itself just produced. The imposition of the 'elite ideology' is by nature oppositional: an LLM either overrides what emerges from the training data or resists the request of a user. In a future where more translations are carried out by LLMs, or similarly complex artificial neural networks, there will be ethico-political lines machine translations will be unable to cross. This may have troubling implications for human expression, since there are many contexts in which it may be valid to explore ideas that AI models have been prevented from engaging with. These ethico-political limitations, along with the ability of vast corpora to crystalize ideologically loaded "semantic configurations" (Lukin 2017, 5), make translation by AI unavoidably a site of ideological encounters. In the future, it may fall to human translators to evaluate translations produced by AI on ideological, ethical, and political grounds.

Appendix

Chat Record for Figure 1

This appendix provides the URLs for all chats with the LLMs GPT-4, GPT-4o, and DeepSeek that provided the data for figure 1.

GPT-4

These chats took place between November 29 and December 2, 2023.

1. https://chatgpt.com/c/e556f1c1-0892-4753-b426-e065f774ac6d
2. https://chatgpt.com/c/d3e2a549-a86f-4553-a252-91a26c16c579
3. https://chatgpt.com/c/cabea232-3ef7-4e1e-9749-1993469c704a
4. https://chatgpt.com/c/adfe8d1b-3100-4d23-b8e3-6ca6ae2ebc8f
5. https://chatgpt.com/c/a4203755-5c79-42ce-b196-65ccd0c227a3
6. https://chatgpt.com/c/4827577c-1969-47e0-8ce0-fb19f9d50baf
7. https://chatgpt.com/c/9342948a-102b-4bac-a772-be73547aa3d4
8. https://chatgpt.com/c/dab5de2d-3104-48fb-aa55-140792615d3d
9. https://chatgpt.com/c/a2adc489-e151-4bd9-afc9-c643696afadd
10. https://chatgpt.com/c/43a84412-a1d1-4f4f-9f44-4bc21f118826
11. https://chatgpt.com/c/ecfc8c6f-0847-49ad-a936-a210f507d5c7
12. https://chatgpt.com/c/99bb8450-fc90-4e6a-abac-59578ba05a28
13. https://chatgpt.com/c/0f0de30e-6c9b-426a-ae30-aeb20d3e982c
14. https://chatgpt.com/c/ce484df5-e28f-4a67-918d-894c7d61eb2f
15. https://chatgpt.com/c/4f97caa8-d62c-40f3-8f2a-8254ac785ddd
16. https://chatgpt.com/c/6d639978-7c43-4374-89b4-5ecf66e27aaa
17. https://chatgpt.com/c/a2d3741e-8206-4279-9d37-2d9c123e9d1c
18. https://chatgpt.com/c/92359985-c6c9-4cea-8e55-4c98bf4161d1
19. https://chatgpt.com/c/91e35134-599d-4418-a016-88b5224c4e9a
20. https://chatgpt.com/c/e239e223-5e69-4d1e-a8b4-564b599bea86
21. https://chatgpt.com/c/de726bf9-d276-4d86-9c69-8eeeefa3be00
22. https://chatgpt.com/c/f177ba5b-7ea5-4ca4-acd3-45aee508ec78
23. https://chatgpt.com/c/f0d15fc4-125f-4ad3-8a72-be3c91fdc040
24. https://chatgpt.com/c/b640d230-8244-4435-8df7-dddd7bf6d185
25. https://chatgpt.com/c/439ffaa4-2b78-4c8e-ad95-7a798fcfee10
26. https://chatgpt.com/c/5648fff3-10e4-4d84-bc37-5bf49fad86ac
27. https://chatgpt.com/c/981530fc-be32-4b5a-8d7d-3c3ee48e1194
28. https://chatgpt.com/c/cc3d3028-359a-4280-8a65-1da80d5f0665
29. https://chatgpt.com/c/aa8cd135-76d7-4480-91a6-430c69d080df
30. https://chatgpt.com/c/7f4a5daf-42fc-486c-a5c1-b5892aae2831
31. https://chatgpt.com/c/5b8587fe-c488-4667-b0d8-bedeada276e6
32. https://chatgpt.com/c/2e27637f-969e-4376-8bd3-8c0fda0d04f0
33. https://chatgpt.com/c/13e58f98-705a-48cb-80eb-6475b98d02a8

GPT-4o

These chats took place between April 8 and April 11, 2025.

34. https://chatgpt.com/c/67f4d667-195c-8001-aee0-711eefea6ec1
35. https://chatgpt.com/c/6802f4f1-affc-8001-9410-32dd37996e55
36. https://chatgpt.com/c/67f4d941-550c-8001-88a1-303218bf6728
37. https://chatgpt.com/c/67f4da3b-147c-8001-8e07-bd502ae8d109
38. https://chatgpt.com/c/67f4dae8-9490-8001-a099-4943dd23b01f
39. https://chatgpt.com/c/67f4db64-6bb0-8001-b2d3-650b07381dbb
40. https://chatgpt.com/c/67f597d1-afd4-8001-bb59-30c11370cce9
41. https://chatgpt.com/c/67f6d529-bd6c-8001-9539-74c562910dc2
42. http://chatgpt.com/c/67f6d55a-44a0-8001-a257-c2ba1eed5c61
43. https://chatgpt.com/c/67f6d5b6-d090-8001-857e-b1cae5612455
44. https://chatgpt.com/c/67f6d618-40d8-8001-a098-f105bf46a5ca
45. https://chatgpt.com/c/67f6d652-b9e8-8001-a283-ab1d817bdc97
46. https://chatgpt.com/c/67f6d6d2-54b8-8001-acd7-597b948dd447
47. https://chatgpt.com/c/67f6d6fe-9b84-8001-a044-d035eed2f4ce
48. https://chatgpt.com/c/67f6d963-d9d8-8001-9fe6-cca9659e1938
49. https://chatgpt.com/c/67f6d9f2-9d48-8001-9085-e202f9f2cbbc
50. https://chatgpt.com/c/67f6da85-fe6c-8001-ac97-226135a90084
51. https://chatgpt.com/c/67f6dab0-5bb4-8001-9bb4-d0fd4a13c505
52. https://chatgpt.com/c/67f6daef-56c4-8001-81dd-ea4d644e8a0e
53. https://chatgpt.com/c/67f6dbaf-87a0-8001-bc63-4f5e64414391
54. https://chatgpt.com/c/67f6dcff-b64c-8001-b1b7-38bcb0bb9166
55. https://chatgpt.com/c/67f6dd23-0a00-8001-9905-9bb1c0a8c241
56. https://chatgpt.com/c/67f6dd56-cbec-8001-9472-0ae009fc1e40
57. https://chatgpt.com/c/67f6dd77-6f54-8001-9079-802c8433a2f2
58. https://chatgpt.com/c/67f6dee5-5ae8-8001-b936-8e81405c283e
59. https://chatgpt.com/c/67f6df5a-c81c-8001-9d03-c34fe28a4742
60. https://chatgpt.com/c/67f832c7-dfa4-8001-8fe1-bf4f2ec0660a
61. https://chatgpt.com/c/67f83307-d4e8-8001-b805-bbc36677d6f4
62. https://chatgpt.com/c/67f83353-d6a8-8001-b94c-ceb78e110770
63. https://chatgpt.com/c/67f84a89-981c-8001-8eeb-eba56a83f472
64. https://chatgpt.com/c/67f84aaf-d3c0-8001-b110-bd553737b8b7
65. https://chatgpt.com/c/67f84b85-d570-8001-ad23-c85ca49a85a4
66. https://chatgpt.com/c/67f84cc0-6bac-8001-931f-a5da1a164adf
67. https://chatgpt.com/c/67f84d1e-c4c0-8001-856d-528faa124bab
68. https://chatgpt.com/c/67f84da1-f55c-8001-b876-57c8ad5e7565
69. https://chatgpt.com/c/67f84e24-fd3c-8001-9ab0-47ff2d840692
70. https://chatgpt.com/c/67f84e49-4f88-8001-9b93-8c0bbaa3f22d
71. https://chatgpt.com/c/67f84f94-b058-8001-961d-3de3a3aeac97

72. https://chatgpt.com/c/67f850a9-9088-8001-b06d-28077ad7a5b1
73. https://chatgpt.com/c/67f850cb-9004-8001-acfb-f004066eeb87
74. https://chatgpt.com/c/67f850eb-3d6c-8001-a505-de4daa5310cd

DeepSeek-V3

These chats took place between April 17 and April 18, 2025.

75. https://chat.deepseek.com/a/chat/s/0523d115-abd7-415b-81b3-cd5893b2a397
76. https://chat.deepseek.com/a/chat/s/70a71dc9-e5aa-4ced-804b-7f92c20abefe
77. https://chat.deepseek.com/a/chat/s/b6f1fc30-40c3-4d95-961c-a60b6a4e42cb
78. https://chat.deepseek.com/a/chat/s/bd461f24-1f9b-4c98-85ce-751febad2d08
79. https://chat.deepseek.com/a/chat/s/a778dff9-636f-4532-9be3-73cc05b3adf5
80. https://chat.deepseek.com/a/chat/s/721e53e8-0e18-48cc-9c33-08e045a255ba
81. https://chat.deepseek.com/a/chat/s/8867c7ce-64ee-49f3-8f65-59304fbfae3c
82. https://chat.deepseek.com/a/chat/s/8f46dee8-fe5d-469d-aa5a-0edd61be6380
83. https://chat.deepseek.com/a/chat/s/0594ec1d-0b12-4e31-b11b-709794bbb242
84. https://chat.deepseek.com/a/chat/s/50e5fc71-919f-413f-b3d7-30c906f34c5b
85. https://chat.deepseek.com/a/chat/s/0e54be6d-ecd5-4597-9e68-403d73d23f67
86. https://chat.deepseek.com/a/chat/s/4378a1af-5191-4e9b-ab50-e412737818f5
87. https://chat.deepseek.com/a/chat/s/5c111d65-f1d4-4627-af2c-b786d2a256f4
88. https://chat.deepseek.com/a/chat/s/dddae162-a14e-4795-96e6-c64ef7d97e9f
89. https://chat.deepseek.com/a/chat/s/02306810-286a-4063-a5b6-43ed9eb3feb8
90. https://chat.deepseek.com/a/chat/s/dad1a7e8-a1dc-4a50-85df-2e649af86d54
91. https://chat.deepseek.com/a/chat/s/ca3a3a34-0653-4e36-95d3-deee58f3a738
92. https://chat.deepseek.com/a/chat/s/c072c5ee-9998-4f33-9d4a-293b9439c168
93. https://chat.deepseek.com/a/chat/s/7a0d7538-e9c3-4738-8bbc-d206d2400ec5
94. https://chat.deepseek.com/a/chat/s/737555c3-96de-4c99-871d-84f2913075ec
95. https://chat.deepseek.com/a/chat/s/836c50de-17b3-4dc7-a537-5d7145ad4f80
96. https://chat.deepseek.com/a/chat/s/b47bb8f1-56c2-44ac-80ff-022d29a729e4
97. https://chat.deepseek.com/a/chat/s/ca4e883b-88ba-4a34-8047-3357652d7b0a
98. https://chat.deepseek.com/a/chat/s/57898b1e-e3f2-4b02-909a-6c96b80c06df
99. https://chat.deepseek.com/a/chat/s/ecb18830-4788-49aa-b183-d1f4c0947ea1
100. https://chat.deepseek.com/a/chat/s/fb7203e8-ea11-4ad6-8b8f-17f1b2920771
101. https://chat.deepseek.com/a/chat/s/b2ac0327-2bc4-4867-badf-adf5be388154
102. https://chat.deepseek.com/a/chat/s/993baeac-7205-482f-abb4-dd5154b9be46
103. https://chat.deepseek.com/a/chat/s/11721460-937c-4f1a-b17c-2ea7409d6720
104. https://chat.deepseek.com/a/chat/s/415f2d39-af6f-4d5f-9a32-6667d278ea17
105. https://chat.deepseek.com/a/chat/s/fee3c30d-c915-4da4-aa5e-7b932daabefa
106. https://chat.deepseek.com/a/chat/s/16e8ac53-5449-497c-990b-d6eb5c0ec37c
107. https://chat.deepseek.com/a/chat/s/1088d30a-385d-4bd2-aca9-09959ed111ad
108. https://chat.deepseek.com/a/chat/s/bc0aa89d-1940-4a3b-8336-34a683614128

109. https://chat.deepseek.com/a/chat/s/251489a0-1002-440d-a29b-7ab541c440ac
110. https://chat.deepseek.com/a/chat/s/169d899b-e67b-4c9d-af17-9e23dce4f019
111. https://chat.deepseek.com/a/chat/s/785db6a5-cd85-4551-aecf-91d817721081
112. https://chat.deepseek.com/a/chat/s/26185185-d65c-4b2c-b261-2420f2471971
113. https://chat.deepseek.com/a/chat/s/0930f0ab-04d6-413f-a924-63dd8ad54a20
114. https://chat.deepseek.com/a/chat/s/65ab8c7f-b444-43c4-90e0-1ac5d2f1d64d
115. https://chat.deepseek.com/a/chat/s/f28d9496-2d98-4492-8930-81883735877b
116. https://chat.deepseek.com/a/chat/s/66311223-e275-4c13-bda7-d2644f35a7f1

References

Acres, Tom. 2023. "ChatGPT Shows 'Significant and Systemic' Left-Wing Bias, Study Finds." *Sky News*, August 17. https://news.sky.com/story/chatgpt-shows-significant-and-systemic-left-wing-bias-study-finds-12941162.

Baker, Mona. 2019. *Translation and Conflict: A Narrative Account*. London: Routledge.

Bandurski, David. 2023. "Bringing AI to the Party." *China Media Project*, April 14. https://chinamediaproject.org/2023/04/14/bringing-ai-to-the-party/.

Bichsel, Peter. 1992. "Die Tochter." [The daughter.] In *Eigentlich Möchte Frau Blum den Milchmann Kennenlernen* [Actually Mrs. Blum would like to meet the milkman], 34–35. Frankfurt: Suhrkamp-Verlag Frankfurt.

Borchert, Wolfgang. 1949. "Die Küchenuhr." [The kitchen clock.] In *Das Gesamtwerk* [The complete works], 197–199. Hamburg: Rowohlt Verlag GmbH.

Chakravorty, Shreya. 2023. "The Politics of Positionality: Gayatri Chakravorty Spivak and Samik Bandyopadhyay as Translators of Mahasweta Devi." In *Mahasweta Devi: Writer, Activist, Visionary*, edited by Radha Chakravarty, 120–128. London: Routledge.

Clayton, James, and Lucy Hooker. 2023. "White House: Big Tech Bosses Told to Protect Public from AI Risks." *BBC News*, May 5. https://www.bbc.co.uk/news/business-65489163.

Colville, Alex. 2025. "DeepSeeking Truth." *China Media Project*, February 10. https://chinamediaproject.org/2025/02/10/DeepSeeking-truth/.

De Vynck, Gerrit. 2023. "ChatGPT Leans Liberal, Research Shows." *The Washington Post*, August 16. https://www.washingtonpost.com/technology/2023/08/16/chatgpt-ai-political-bias-research/.

Ding, Gang. 2023. "AI Has a Political Stance, Associated with Ideology." *Global Times*, May 10. https://www.globaltimes.cn/page/202305/1290460.shtml.

Field, Matthew. 2025. "Chinese AI Has Sparked A $1 Trillion Panic – And It Doesn't Care About Free Speech." *The Telegraph*, January 27. https://www.telegraph.co.uk/business/2025/01/27/chinese-DeepSeek-ai-has-sparked-a-1-trillion-panic/.

Ghosh, Sourojit, and Aylin Caliskan. 2023. "ChatGPT Perpetuates Gender Bias in Machine Translation and Ignores Non-Gendered Pronouns: Findings Across Bengali and Five Other Low-Resource Languages." In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, 901–912. doi:10.48550/arXiv.2305.10510.

Gupta, Ragav. 2025. "Comparative Analysis of DeepSeek R1, ChatGPT, Gemini, Alibaba, and LLaMA: Performance, Reasoning Capabilities, and Political Bias." *Authorea*, February 10. doi:10.22541/au.173921625.50315230/v1.

Hall, Stuart. 1990. "Cultural Identity and Diaspora." In *Identity: Community, Culture, Difference*, edited by Jonathan Rutherford, 2–27. London: Lawrence & Wishart.

———. 2013. "Cultural Identity and Diaspora." In *Colonial Discourse and Post-Colonial Theory: A Reader*, edited by Patrick Williams and Laura Chrisman, 392–403. London: Routledge.

Jiang, Ben. 2025. "Beijing Meeting Puts Spotlight on China's New Face of AI, DeepSeek Founder Liang Wenfeng." *South China Morning Post*, January 2. https://www.scmp.com/tech/policy/article/3295662/beijing-meeting-puts-spotlight-chinas-new-face-ai-deepseek-founder-liang-wenfeng.

Johri, Shreya. 2023. "The Making of ChatGPT: From Data to Dialogue." *Science in the News, Harvard Kenneth C. Griffin Graduate School of Arts and Sciences*, June 6. https://sitn.hms.harvard.edu/flash/2023/the-making-of-chatgpt-from-data-to-dialogue/.

Kang, Jay Caspian. 2024. "How Biased Is the Media, Really?" *The New Yorker*, October 18. https://www.newyorker.com/news/fault-lines/how-biased-is-the-media-really.

Kenny, Dorothy. 2022. "Human and Machine Translation." In *Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence*, edited by Dorothy Kenny, 23–49. Berlin: Language Science Press.

Krauss, Nicole. 2020. "To Be a Man: A Short Story." *The Atlantic*, October 2. https://www.theatlantic.com/books/archive/2020/10/nicole-krauss-to-be-a-man/616517/.

Liu, Qianer. 2023. "China to Lay Down AI Rules With Emphasis on Content Control." *Financial Times*, July 11. https://www.ft.com/content/1938b7b6-baf9-46bb-9eb7-70e9d32f4af0.

Lu, Donna. 2025. "We Tried Out DeepSeek. It Worked Well, Until We Asked It About Tiananmen Square and Taiwan." *The Guardian*, January 28. https://www.theguardian.com/technology/2025/jan/28/we-tried-out-deepseek-it-works-well-until-we-asked-it-about-tiananmen-square-and-taiwan.

Lukin, Annabelle. 2017. "Ideology and the Text-in-Context Relation." *Functional Linguistics* 4 (1): 1–17. doi:10.1186/s40554-017-0050-8.

Martin, John Levi. 2023. "The Ethico-Political Universe of ChatGPT." *Journal of Social Computing* 4 (1): 1–11. doi:10.23919/JSC.2023.0003.

Michta, Andrew A. 2025. "The Real Reason Russia Invaded Ukraine (Hint: It's Not NATO Expansion)." *Atlantic Council*, March 6. https://www.atlanticcouncil.org/blogs/new-atlanticist/the-real-reason-russia-invaded-ukraine-hint-its-not-nato-expansion/.

Moorkens, Joss. 2022. "Ethics and Machine Translation." In *Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence*, edited by Dorothy Kenny, 121–140. Berlin: Language Science Press.

Motoki, Fabio, Valdemar Pinho Neto, and Victor Rodrigues. 2024. "More Human than Human: Measuring ChatGPT Political Bias." *Public Choice* 198:3–23. doi:10.1007/s11127-023-01097-2.

National Cybersecurity Standards Committee. 2024. "Shengchengshi rengong zhineng fuwu anquan jiben yaoqiu." [Basic security requirements for generative artificial intelligence service.] TC260-003. https://www.tc260.org.cn/upload/2024-03-01/1709282398070082466.pdf.

Pacheco, Andre G. C., Athus Cavalini, and Giovanni Comarela. 2025. "Echoes of Power: Investigating Geopolitical Bias in US and China Large Language Models." arXiv preprint. doi:10.48550/arXiv.2503.16679.

Pérez-Ortiz, Juan Antonio, Mikel L. Forcada, and Felipe Sánchez-Martínez. 2022. "How Neural Machine Translation Works." In *Machine Translation for Everyone: Empowering Users in the Age of Artificial Intelligence*, edited by Dorothy Kenny, 141–164. Berlin: Language Science Press.

Prates, Marcelo O. R., Pedro H. Avelar, and Luís C. Lamb. 2020. "Assessing Gender Bias in Machine Translation: A Case Study with Google Translate." *Neural Computing and Applications* 32 (10): 6363–6381. doi:10.1007/s00521-019-04144-6.

Radford, Alec, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. 2018. "Improving Language Understanding by Generative Pre-training." *OpenAI*. Accessed July 20, 2024. https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf.

Sachs, Jeffrey D. 2023. "The War in Ukraine Was Provoked—and Why That Matters to Achieve Peace." *Jeffsachs.org*, May 23. https://www.jeffsachs.org/newspaper-articles/wgtgma5kj69pbpndjr4wf6aayhrszm.

Skeet, Ann. 2024. "ChatGPT Has Revived Interest in Ethics. The Irony Is That We Haven't Been Holding Humans to the Same Standard." *Fortune*, January 22. https://fortune.com/2024/01/22/chatgpt-revived-ethics-human-leadership/.

Statham, Simon. 2022. *Critical Discourse Analysis: A Practical Introduction to Power in Language*. London: Routledge.

Tschunkert, Kristina. 2021. "Working with Translators: Implications of the Translator's Positionality for the Research Process and Knowledge Production." In *The Companion to Peace and Conflict Fieldwork*, edited by Roger Mac Ginty, Roddy Brett, and Birte Vogel, 249–262. London: Palgrave Macmillan.

Tymoczko, Maria. 2002. "Ideology and the Position of the Translator: In What Sense is a Translator 'In Between'?" In *Apropos of Ideology: Translation Studies on Ideology – Ideologies in Translation Studies*, edited by María Calzada-Pérez, 181–201. London: Routledge.

Vanmassenhove, Eva, Christian Hardmeier, and Andy Way. 2019. "Getting Gender Right in Neural Machine Translation." arXiv preprint. doi:10.18653/v1/D18-1334.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. "Attention is All You Need." In *Advances in Neural Information Processing Systems 30: 31st Annual Conference on Neural Information Processing Systems (NIPS 2017)*, edited by Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna Wallach, Rob Fergus, Sundar Vishwanathan, and Roman Garnett. 1–11. La Jolla, CA: Neural Information Processing Systems. https://papers.nips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

Wang, Jun, Benjamin Rubinstein, and Trevor Cohn. 2022. "Measuring and Mitigating Name Biases in Neural Machine Translation." In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2576–2590. https://aclanthology.org/2022.acl-long.184/.

Weatherby, Leif. 2023. "ChatGPT Is an Ideology Machine." *Jacobin Magazine*, April 17. https://jacobin.com/2023/04/chatgpt-ai-language-models-ideology-media-production.

Zhang, Handuo. 2020. "Seq2seq Model with Attention." *Zhang Handuo's Site*, June 1. https://zhanghanduo.github.io/post/attention/.