Dijlah Journal of Engineering Science (DJES) Vol. 3, No.1, March, 2026, pp. 72-78
ISSN: Printed: **3078-9656, Online: 3078-9664**, paper ID: 111

Artificial Intelligence in Arabic Natural Language Processing: A Review of Models, Datasets, and Applications

Rusul Hussein Hasan [1] Rasha Majid Hassoon [2], Inaam Salman Aboud [3]
[1, 2] University of Baghdad, Iraq, Baghdad
[3] College of Education, Al- Mustansiriya University, Iraq, Baghdad.
Email: [1] russl@colaw.uobaghdad.edu.iq, [2] rasha.m@uobaghdad.edu.iq [3] inaamsalman89@uomustansiriyah.edu.iq

Article Info
Article history:
Received Oct., 12, 2025
Revised Nov.,25, 2025
Accepted Jan., 5, 2026

Keywords:
Artificial Intelligence
Natural Language Processing
Arabic Language, Transformer Models
Deep Learning
Sentiment Analysis

ABSTRACT
Arabic language processing with artificial intelligence has evolved significantly in the past decades, from traditional rule- and dictionary-based techniques, through statistical models to modern deep and transformer models. This review intends to present an overview of the most well-known Arabic models as well as datasets used for its training, and the main practical applications such as sentiment analysis, machine translation, speech recognition, and smart assistant.AI-based Arabic NLP has had good progress in the previous decades, from rule and dictionary-based approaches to statistical methods and deep transformative learning models nowadays. In addition to it, the most popular state-of-the-art models that are fine-tuned for the Arabic language and their corpora of training data will be considered as well as major applications such as Sentiment Analysis, Machine Translation, Speech Recognition, and Virtual Assistant.This article review outlines the necessity for investment in language resources and advanced models to improve AI systems' ability to accurately understand Arabic natural language, a contribution that will support real-life applications and smart services associated with its present formalized variant of AI model capabilities.

Corresponding Author:
Rusul Hussein Hasan
University of Baghdad
Baghdad, Iraq
Email: russl@colaw.uobaghdad.edu.iq

1. INTRODUCTION

One area where this has been the biggest is natural language processing (NLP), which is devoted to teaching machines how to make sense of, interpret and express human language in a manner that is indiscernible from humans. That is why it has emerged as the standard technology for writers who want to play with developing applications in English from speech or text. [1 2] Employing classical methods for Arabic data , no way to do that effectively. In addition there is any attempt would have appeared on the horizon as it fails to satisfy reasonable requirements of performance. This strategy of translation is entirely revolutionary, unlike past attempts. [2] In addition to its noun-adjective word order, Arabic has a wide range of syntactic forms and special verb-particle combinations which are quite difficult to master properly for unsophisticated models. In the Arabic-speaking world Arabic proper consists of a great many different dialects (heavily influenced by both its complex grammar and vast number of different regional varieties) each one of which may be influenced further still by mother tongue influence over generations due to immigration history. All these factors make it difficult to use purely empirical methods when trying to produce fully precise linguistic models for every way Arabic is used. [3 4]Traditional and translational approaches do not deliver high performance; particularly in such areas as text generation or semantic comprehension where Arabic-language models have to be developed on specific data themselves. It is clear that we need to create an Arabic-generated model and process original information in order support this trend: [3 5 6]This review article aims to provide an overview of recent artificial intelligence advances in Arabic language processing by looking into the most widely-known models and training data sources used for them. We also delineate which are the most serious real-world applications for translation, text analysis and linguistic interaction in this particular field.

Thus the article has another purpose too: to consider current difficulties and potential breakthroughs in creating more precise and effective Arabic-processing systems.

2. DEVELOPMENT OF ARABIC LANGUAGE PROCESSING TECHNIQUES

Arabic language processing has made progress in last few decades due to global advancement of artificial intelligence and machine learning technologies. While you once upon a time had to turn to grammar and manual dictionaries for that, modern models instead learn from big data and deep neural networks to provide unparalleled linguistic accuracy and understanding. [7]

1.Traditional Methods based on Rules and Dictionaries [8]

Began by relying on grammar-based linguistic analysis (Parsing) that uses hand-crafted grammatical and morphological rules to analyze Arabic text. They were based on linguistic dictionaries that identified roots, weights, and pattern, the conjugation (Conjugation) algorithms of grammatical analysis to find relationships between words. Though of this approach is accurate in part fields, it has met with difficulty due to dialectal plurality, meaning of words have multiple shades and Arabic Police possess shape unusually. In consequence it doesn't have generalizability.

2.Statistical Models Stage [9]

With the beginning of the new millennium, statistical models emerged that replaced manual rules with probabilities extracted from data. The most prominent of these methods are:

- N-grams that rely on the probability of a word sequence to predict the next word.
- Hidden Markov Models (HMM) that have been used in sequence analysis such as speech recognition and grammatical tagging (POS Tagging.

This stage allowed clear progress in tasks such as machine translation, text analysis, and entity extraction, but remained limited by its reliance on statistical repetition without a deep understanding of the meaning or context of the language.

3.Deep Learning Stage [10]

With the development of computing and the emergence of deep neural networks (Deep Neural Networks), Arabic language processing has moved to a new stage. Models such as:

- Recurrent neural networks (RNNs) and their developed versions such as LSTM and GRU for processing word sequences and understanding temporal context.
- Bypass networks (CNNs) to extract distinct patterns in texts, especially in text classification and sentiment analysis tasks.

These models contributed to significantly improved performance, but their ability to understand distant relationships between words was limited, paving the way for the next stage.

4. Transformers [11]

When the Transformer model came out in 2017, a lot of NLP projects made great improvement on Chinese text and Arabic have been working fine so far. These models were constructed with a self-attention mechanism, which will allow one to learn and access interaction between every word at a time of sentence level to know what meaning will be made next. The most common applications of this technique are

- All other models, namely BERT and some of its Arabic offspring as well as our model Robert: MARBERT, Arabic BERT and Camel BERT.
- Language generation models (generative) such as GPT and Jaisaendeavor to maintain the Arabic language during both comprehension and generation.
- They jump-started translation, messages machinated in Hammond blears and feelings sounded out and questions posed. Powerful large-scale Arabic models from local data and multi-dialect data that enables the Arab world to be part of the language AI revolution cycle in action on the global scale.

3. MODELS

These models led the way to a qualitative shift in understanding of and representation for Arabic thanks to a new breed of transfer-based deep linguistic models that describe the syntactical structure of a language. They can broadly be divided into two types of models, one that only considers Arabic and multilingual ones with Arabic being just one of many languages.

1. Models for Arabic-only

They are created specifically for solving the problems of Arabic's morphological, grammatical and dialectal based which include but not limited to:

- AraBERT: As one of the first transformational models designed specially for Arabic, it has been trained by using data sources as diverse as news text, Wikipedia, and social media sites, and has shown very good performance in text classification and Sentiment Analysis tasks.
- MARBERT: Developed for the purpose of supporting Arabic dialects in addition to classical Arabic, it was trained on large-scale social media data making it more adept at understanding informal texts.
- Camel BERT: Developed by the Camel Lab, it has multiple versions suitable for different tasks such as morphological analysis, text classification, and entity recognition.
- Arabic BERT and Gigabit: They were trained with large quantities of Arabic and English texts, and thus were capable of having a strong linguistic quality which was good for tasks in which linguistic comprehension is needed
- Jays: One of the latest open-source Arabic models, it is designed to handle both classical Arabic and a variety of dialects, with advanced text generation capabilities, so it is an important step towards the development of a large Arabic language model similar to GPT.
- The models are distinguished from global models in that they take the linguistic specificity of Arabic into account and therefore deal with its morphological and dialectal problems more adeptly.

2. Multilingual Models [14 15]

This category is based on a single model being trained in many languages, with Arabic included among them, so that it has the advantage of being able to obtain knowledge from common relationships between languages. The most important ones are as follows:

- Met (Multilingual BERT) : One of the first models supporting the Arabic language, but with only a very feeble Arabic representation within the language group. As a result, its performance is relatively poor.
- XLM-Roberta: A heavyweight multilingual model, it showed higher performance than met on Arabic text-related tasks by virtue of more diverse and high quality training data.
- GPT-4: A huge generation model that is capable of understanding Arabic very well, even though its training was not mainly in Arabic. Due to its huge size, it is doing quite well on both understanding and generation tasks in Arabic.
- LLaMA-2: An open source model from Meta which supports Arabic in a number of languages and is used in language generation, summarization and translation research.
- Although multilingual models have great flexibility and powerful capabilities in general tasks, they often perform poorly when dealing with Arabic because they are not trained to the extent that other models are in languages during an unbiased distribution.

3. Performance Comparison between Models [16]

Comparative studies have shown that Arabic-oriented models such as Robert and MARBERT clearly excel at the tasks of text classification, sentiment analysis, and named entity extraction, due to their focus on the characteristics of the Arabic language. Multilingual models such as XLM-Roberta and GPT-4 excelled in machine translation and text generation tasks thanks to the diversity of linguistic data they were trained on. In general, monolingual models arguably excel at local understanding of language, while multilingual models excel at tasks that require cross-linguistic knowledge transfer (Cross-lingual Transfer)

4. DATABASES

Linguistic databases are the cornerstone upon which all AI models in natural language processing are built. The more diverse, quality, and appropriate the data is to the target language, the more accurate and efficient the models are to understand, analyze, and generate. Since the Arabic language is characterized by its morphological

complexity and multiple dialects, the availability of appropriate data has posed a major challenge to developing strong models that serve Arabic effectively.

1. **Public Databases [17 18 19]**

This category of data includes open linguistic sources containing general and diverse texts typically used in training basic models or for the purpose of academic research. The most prominent of which are:

- OSIAN (Open Source International Arabic News Corpus): A database consisting of Arabic news articles from multiple sources, used for tasks such as text classification and sentiment analysis.
- Arabic Wikipedia: It is one of the largest open text resources in the Arabic language, and is widely used in training transformational models such as Robert and Arabic BERT.
- Shekel: A huge database containing fully formed Arabic texts, making it an important resource in automated formation, text correction, and morphology analysis tasks.
- Arabia Corpus: This rule combines literary, journalistic, and educational texts, and is used in the tasks of analyzing meanings and classifying content.

These rules are comprehensive and accessible, but most lack the dialectal diversity and thematic focus required for specialized applications.

1. Task-Specific Datasets [20]

A large number of databases have been developed for specific tasks in Arabic language processing, the most prominent of which are:

- Text analysis and classification: such as HARD and OSAC, which are used to classify news or tweets according to theme or tone.
- Sentiment Analysis (Sentiment Analysis): Like Arendt-Lev, ASTD, and LABR, it is used to study positive and negative emotions in Arabic texts, especially on social media.
- Machine translation (Machine Translation): such as OPUS and QED Arabic-English Corpus, which contain sentences translated between Arabic and English to train neural translation systems.
- Arabic dialects (Dialectal Arabic): such as MADAR Corpus and NADI Dataset, which aim to identify different regional dialects (Gulf, Egyptian, Levantine, Maghreb.(
- Questions and Answers (Question Answering): Such as Arabic SQuAD, ARCD, and TyDi QA, they are used to train models capable of understanding texts and answering direct questions.

These rules contribute to enhancing the performance of models within specialized tasks, but most remain limited in size or geographical coverage of dialects. [21 22]

2. Challenges Related to Arabic Data

In recent years, progress has been fantastic. However, there are still some fundamental problems with Arabic data bases that users or developers should pay attention to. Topics include but are not limited to: [23 24]

. Size: Most of Arabic Databases are still much smaller than what the International languages have; such as English, Chinese or Russian – limiting model generalization capability at best.

. Variety of dialects: The wide range of data available also poses a challenge; the Arabic dialects differ so much in vocabulary and grammar that one version may be unintelligible to speakers of another.

. Quality of localization and purification: It is well known that there are errors in spelling and grammar for a number of text sources. Licensing & IP issues.

Applications

The current availability of AI artificial intelligence research and natural language processing have enabled a number of practical applications with respect to Arabic, a language that is significant to every aspect of our lives as well as being key in many industries. The most typical are: [25 26]

1. Mood Analysis and Commentates

The posts on social networks and comments on news sites are subject of the analyses based on artificial intelligence. Data is aggregated on general vibe of whether an audience may be in a good or bad mood regarding business strategies, product monitoring for public opinion sensitivity or policy implementation and service.

2. Machine Translation

Arabic to English, by using hybrid models, cans synthesis Arabic-English translation of high quality also 5 Information Channels with the international world. Systems developed using these methods allow for more efficient, robust translations of both formal and informal text by transcending any linguistic chasms.

3.Speech Recognition

Modern applications are now based on recognition models for spoken Arabic, a variety of dialects included. This is used for voice assistants, interpretation systems and automated lecture-notes interviews.

4.**Intelligent Assistants & Arabic Chabot's**

Artificial intelligence has enabled the development of smart assistants and chat systems that interact in Arabic, such as Arabic versions of Catgut, to provide support services, answer questions, make recommendations, and interact with users in a natural and seamless manner.

5. Text Correction and Morphological Analysis

Automatic spelling, grammar, and parsing correction tools are important practical applications, helping students, writers, and editors produce accurate texts free of linguistic errors, and contributing to supporting research tasks and large-scale text analysis.These applications demonstrate how artificial intelligence has become an essential part of the development of the digital Arabic language, whether in the fields of communication, education, or smart services, while enhancing access to Arabic texts and content more effectively and accurately.

5. EXISTING CHALLENGES

Challenges although significant progress has been made in artificial intelligence and Arabic language processing technologies, there exist a number of core challenges that hinder the development of accurate and dependable systems: [27 28 29]

1. **Scarcity of high-quality Arabic data:**

The majority of Arabic datasets are either small or narrow and do not allow for models to effectively learn from text in a representative manner, especially when it comes to complex or specialized domains.

2. **Poor Multiple Dialect Support**

Hence, most of the research and corpora is more centered on Classical Arabic. But they have to go through a gang text that does upchuck full of slang or the local dialect, it just doesn’t help.

3. **Observation and expression**

Another fog-like fabrication is its habitat in the Qur'an, where it twists characters in words and decomposes a sentence into meanings.

4. **Recognition of Arabic Investment**

In any large dwarf model, trained on big data sets, Arabic is represented weakly. Which means that one way or another, truncating such models even in their prime has to be more costly especially in fields like meaning comprehension and text generation?

6. Future Trends

The area of artificial intelligence is a fast evolving field and as such it spawns several future trends of research. This wave tries to enhance the performance of models for illumination the Arabic language computationally speaking in all shades:

1. Building a Big-Open Source Arabic Model:

Formulating a general model from the relaxed data is a great step forward. The SLUAE model achieves state-of-the-art performance for understanding and generation of Arabic with high adaptability when it is transferred to a number of tasks in various manners.

2.Assistants for Arabic dialects and their constipation in traditional training methods:

One of the biggest challenges is how to integrate Arabic dialects (alpha, alpha) and put them in training data? The approach will help models deal with more colloquial text and make it easier for them to understand all of you asking don’t want shoes and maybe even use smart assistants have a go.

3.Audio, Text and Image for Multimodal Language Processing (Multimodal LP):

If, with the inclusion of audio data and written text and images to further train models for intelligent systems, that's not a huge leap forward then I don't know what is. In addition to this speech recognition, understanding, text generation and understanding of the spatial context would enhance so much in a speaker's experience as far as his own Arabic language is concerned.

4. Arabia and interpretation of models are necessary:

With the growing use of deep models, we need to make these models explainable. Or where the decision is made and how the decision is taken also that through what Arab model even if it can be interpreted –this is a very important part- of trust, transparency and provides some mechanisms to bring these models into primary sensitive application areas in education, health and security. These have positioned, and are yet to position Arabic NLP for continued research in text mining, IR, HLT among others in a situation of minute presence where core linguistic resource development and advanced models sensitive to the language features themselves. Transaction With global development in the UI scenes have you learned a new language recently.

7. Conclusion

This review highlights notable developments in the field of Arabic language processing using artificial intelligence, as technologies have moved from traditional methods based on manual grammar, through statistical models, to modern deep and transformative models that have achieved a qualitative leap in understanding and representing language semantically.Despite significant achievements, the Arabic language still faces fundamental challenges related to scarcity of high-quality data, multiple dialects, difficulty of formation, and bias in multilingual models. These challenges underscore the need to develop greater linguistic resources, specialized models, and improved interpretive mechanisms to ensure reliable performance of Arabic models.Future trends point to the need to create large-scale, open source Arabic core models, support multiple dialects, develop multimedia processing, and improve the interpretability of models. Through these efforts, artificial intelligence can effectively contribute to enabling practical applications of the Arabic language, enhancing human-machine interaction, and ensuring the position of the Arabic language in the global revolution in natural language processing.

References

1. Khedimi, S., Bouziane, A., & Bouchiha, D. (2024). Advancements and challenges in Arabic question answering systems: a comprehensive survey. *Brazilian Journal of Technology*, *7*(4), e75604. https://doi.org/10.38152/bjtv7n4-028
2. *GPTAraEval: A Comprehensive Evaluation of ChatGPT on Arabic NLP*. (2023). https://doi.org/10.48550/arxiv.2305.14976
3. Khondaker, Md. T. I., Waheed, A., Nagoudi, E. M. B., & Abdul-Mageed, M. (2023). GPTAraEval: A Comprehensive Evaluation of ChatGPT on Arabic NLP. *arXiv.Org*, *abs/2305.14976*. https://doi.org/10.48550/arXiv.2305.14976
4. Alammary, A. (2022). BERT Models for Arabic Text Classification: A Systematic Review. *Applied Sciences*, *12*(11), 5720. https://doi.org/10.3390/app12115720
5. Jeetbala, Smt., Singh, D., & Goyal, K. D. (2025). *Natural language processing (nlp)*. 46–59. https://doi.org/10.58532/nbennurtech6
6. Bellaouar, S., Nehar, A., Souffi, S., & Bouameur, M. (2025). Dhati+: Fine-tuned Large Language Models for Arabic Subjectivity Evaluation. *arXiv.Org*, *abs/2508.19966*. https://doi.org/10.48550/arxiv.2508.19966
7. Jeetbala, Smt., Singh, D., & Goyal, K. D. (2025). *Natural language processing (nlp)*. 46–59. https://doi.org/10.58532/nbennurtech6
8. Alderazi, F., Algosaibi, A., Al-Abdullatif, M., Ahmad, H. F., Qamar, A. M., & Albarrak, A. (2024). Generative artificial intelligence in topic-sentiment classification for Arabic text: a comparative study with possible future directions. *PeerJ*, *10*, e2081. https://doi.org/10.7717/peerj-cs.2081
9. Al Moaiad, Y., Alobed, M., Alsakhnini, M., & Momani, A. M. (2024). Challenges in natural Arabic language processing. *Edelweiss Applied Science and Technology*, *8*(6), 4700–4705. https://doi.org/10.55214/25768484.v8i6.3018
10. Al-Ballaa, S. R., Al-Twairesh, N., Al-Salman, A. S., & Alfarhood, S. (2025). GATmath and GATLc: Comprehensive benchmarks for evaluating Arabic large language models. *PLOS ONE*, *20*(9), e0329129. https://doi.org/10.1371/journal.pone.0329129
11. Saiah, N. L. D., & Wafaà, D. N. (2025). The Role of Artificial Intelligence in Arabic Language Processing: Ethical Challenges and Development Prospects. *Al-Wazan*, *3*(1), 13–31. https://doi.org/10.58223/al-wazan.v3i1.345
12. Al Deen, M. M. S., Pielka, M., Hees, J., Abdou, B. S., & Sifa, R. (2023). Improving Natural Language Inference in Arabic using Transformer Models and Linguistically Informed Pre-Training. *arXiv.Org*, *abs/2307.14666*. https://doi.org/10.48550/arxiv.2307.14666
13. Haouhat, A., Bellaouar, S., Nehar, A., Cherroun, H., & Abdelali, A. (2025). Arabic Multimodal Machine Learning: Datasets, Applications, Approaches, and Challenges. *arXiv.Org*, *abs/2508.12227*. https://doi.org/10.48550/arxiv.2508.12227

14. Hanandeh, A., Ayasrah, S., Kofahi, I., & Qudah, S. (2024). Artificial Intelligence in Arabic Linguistic Landscape: Opportunities, Challenges, and Future Directions. *TEM Journal*, 3137–3145. https://doi.org/10.18421/tem134-48

15. Mohmiddin, N. A. M., & Abu Bakar, R. (2025). The Role Of Artificial Intelligence In Learning Arabic For Non-Native Speakers: A Systematic Literature Review. *Al-Qanatir: International Journal of Islamic Studies*, *34*(03), 198–216. https://doi.org/10.64757/alqanatir.2025.3403/1121

16. Mohamed, M., & Alosman, K. (2024). A Comparative Study of Deep Learning Approaches for Arabic Language Processing. *Jordan Journal of Electrical Engineering*, *11*(1), 1. https://doi.org/10.5455/jjee.204-1711016538

17. Deen, M., Pielka, M., van Hees, J. J., Abdou, B. S., & Sifa, R. (2023). *Improving Natural Language Inference in Arabic Using Transformer Models and Linguistically Informed Pre-Training*. https://doi.org/10.1109/ssci52147.2023.10371891

18. Darwish, K., Habash, N., Abbas, M., Al-Khalifa, H. S., Al-Natsheh, H. T., El-Beltagy, S. R., Bouamor, H., Bouzoubaa, K., Cavalli-Sforza, V., El-Hajj, W., Jarrar, M., & Mubarak, H. (2020). A Panoramic Survey of Natural Language Processing in the Arab World. *arXiv: Computation and Language*. https://arxiv.org/pdf/2011.12631.pdf

19. Darwish, K., Habash, N., Abbas, M., Al-Khalifa, H. S., Al-Natsheh, H. T., El-Beltagy, S. R., Bouamor, H., Bouzoubaa, K., Cavalli-Sforza, V., El-Hajj, W., Jarrar, M., & Mubarak, H. (2020). A Panoramic Survey of Natural Language Processing in the Arab World. *arXiv: Computation and Language*. https://arxiv.org/pdf/2011.12631.pdf

20. Alsaied, M. A. (2024). The Contribution of Artificial Intelligence Technologies in Refining the Arabic Language Skills of Its Native Speakers: Achieved and Possible. *Journal of Ecohumanism*, *3*(8). https://doi.org/10.62754/joe.v3i8.5126

21. Alsaied, M. A. (2024). The Contribution of Artificial Intelligence Technologies in Refining the Arabic Language Skills of Its Native Speakers: Achieved and Possible. *Journal of Ecohumanism*, *3*(8). https://doi.org/10.62754/joe.v3i8.6576

22. Khondaker, M. T. I., Waheed, A., Nagoudi, E. M. B., & Abdul-Mageed, M. (2023). *GPTAraEval: A Comprehensive Evaluation of ChatGPT on Arabic NLP*. https://doi.org/10.18653/v1/2023.emnlp-main.16

23. Alshammari, H., El-Sayed, A., & Elleithy, K. (2024). AI-Generated Text Detector for Arabic Language Using Encoder-Based Transformer Architecture. *Big Data and Cognitive Computing*. https://doi.org/10.3390/bdcc8030032

24. Ghaddar, A., Wu, Y., Bagga, S., Rashid, A., Bibi, K., Rezagholizadeh, M., Xing, C., Wang, Y., Duan, X., Wang, Z., Huai, B., Jiang, X., Liu, Q., & Langlais, P. (n.d.). Revisiting Pre-trained Language Models and their Evaluation for Arabic Natural Language Processing. *Conference on Empirical Methods in Natural Language Processing*, 3135–3151. https://www.aclanthology.org/2022.emnlp-main.205.pdf

25. Mashaabi, M., Al-Khalifa, S., & Al-Khalifa, H. S. (2024). *A Survey of Large Language Models for Arabic Language and its Dialects*. https://doi.org/10.48550/arxiv.2410.20238

26. Jarrar, M., Birim, A., Khalilia, M., Erden, M., & Ghanem, S. (2023). *ArBanking77: Intent Detection Neural Model and a New Dataset in Modern and Dialectical Arabic*. *abs/2310.19034*. https://doi.org/10.48550/arxiv.2310.19034

27. Al Omari, M., & Al-Hajj, M. (2018). *Classifiers for Arabic NLP: survey*. *1*(1), 231. https://doi.org/10.1504/IJCCIA.2018.10019805

28. Alghamdi, A., Duan, X., Jiang, W., Wang, Z., Wu, Y., Xia, Q., Wang, Z., Zheng, Y., Rezagholizadeh, M., Huai, B., & Ghaddar, A. (2023). AraMUS: Pushing the Limits of Data and Model Scale for Arabic Natural Language Processing. *arXiv.Org*, *abs/2306.06800*. https://doi.org/10.48550/arXiv.2306.06800

29. *AraMUS: Pushing the Limits of Data and Model Scale for Arabic Natural Language Processing*. (2023). https://doi.org/10.48550/arxiv.2306.06800