
**SYSTEMATIC LITERATURE REVIEW TERHADAP SEMANTIC PARSING
BERBASIS TRANSFORMER DALAM NATURAL
LANGUAGE PROCESSING**

**Dwi Suci Ramadani¹, Fahira Hasanah Br Harahap², Suci Lestari Br Batu Bara³,
Winda Sari⁴, Dicky Apdillah⁵
Universitas Asahan, Kisaran**

e-mail: dwisucir697@gmail.com¹, fhirahsn21@gmail.com²,
sucilestari112237@gmail.com³, windaasari1212@gmail.com⁴,
dickyapdi1404@gmail.com⁵

Abstract: *The development of transformer architecture has brought significant changes in Natural Language Processing (NLP) research, particularly in semantic parsing tasks. This study aims to analyze the development of transformer-based semantic parsing through a Systematic Literature Review (SLR) approach. The research method uses the PRISMA guidelines with a process of identification, selection, and analysis of articles from the Scopus, ScienceDirect, IEEE Xplore, Springer, and Google Scholar databases for the period 2019–2026. The results of the study show that transformer models such as BERT, T5, GPT, and transformer encoder–decoder have become the dominant approaches in modern semantic parsing because they are able to improve contextual understanding, semantic representation, and reasoning more effectively than previous methods. The most widely used datasets include Spider, WikiSQL, GeoQuery, ATIS, and Overnight, with the main evaluation being Execution Accuracy and Exact Match Accuracy. This study also identifies various challenges, such as limited labeled data, high computational costs, low model interpretability, the risk of hallucinations, and limitations in low-resource languages. This research contributes to a literature review, identifying research gaps, and recommending future research directions for transformer-based semantic parsing.*

Keywords: *BERT, Natural Language Processing, Semantic Parsing, Systematic Literature Review, Transformer, T5, Text-to-SQL.*

Abstrak: Perkembangan arsitektur transformer telah membawa perubahan signifikan dalam penelitian Natural Language Processing (NLP), khususnya pada tugas semantic parsing. Penelitian ini bertujuan untuk menganalisis perkembangan semantic parsing berbasis transformer melalui pendekatan Systematic Literature Review (SLR). Metode penelitian menggunakan pedoman PRISMA dengan proses identifikasi, seleksi, dan analisis artikel dari database Scopus, ScienceDirect, IEEE Xplore, Springer, dan Google Scholar pada periode 2019–2026. Hasil kajian menunjukkan bahwa model transformer seperti BERT, T5, GPT, dan encoder–decoder transformer menjadi pendekatan dominan dalam semantic parsing modern karena mampu meningkatkan contextual understanding, semantic representation, dan reasoning secara lebih efektif dibandingkan metode sebelumnya. Dataset yang paling banyak digunakan meliputi Spider, WikiSQL, GeoQuery, ATIS, dan Overnight, dengan metrik evaluasi utama berupa Execution Accuracy dan Exact Match Accuracy. Penelitian ini juga mengidentifikasi berbagai tantangan, seperti keterbatasan data berlabel, tingginya biaya komputasi, rendahnya interpretabilitas model, risiko hallucination, serta keterbatasan pada low-resource language. Penelitian ini berkontribusi dalam memberikan pemetaan literatur, identifikasi research gap, dan rekomendasi arah penelitian semantic parsing berbasis transformer pada masa mendatang.

Kata Kunci: BERT, Natural Language Processing, Semantic Parsing, Systematic Literature Review, Transformer, T5, Text-to-SQL.

PENDAHULUAN

Perkembangan kecerdasan buatan telah mendorong Natural Language Processing (NLP) menjadi salah satu bidang penelitian yang berkembang pesat. NLP memungkinkan komputer memahami, memproses, dan menghasilkan bahasa manusia secara otomatis. Evolusi NLP dari pendekatan rule-based menuju machine learning dan deep learning telah meningkatkan kemampuan model dalam memahami konteks dan representasi bahasa secara lebih adaptif (Shahin & Ismail, 2024).

Kemajuan signifikan terjadi dengan hadirnya arsitektur Transformer yang diperkenalkan oleh Vaswani et al. melalui mekanisme self-attention. Arsitektur ini mampu memahami hubungan antar kata secara lebih efektif dibandingkan model sebelumnya seperti RNN dan LSTM, serta menjadi dasar bagi berbagai model modern seperti BERT, GPT, T5, dan RoBERTa (Miletić & Schulte im Walde, 2024).

Salah satu aplikasi penting transformer dalam NLP adalah semantic parsing, yaitu proses mengubah bahasa alami menjadi representasi formal yang dapat dipahami mesin, seperti logical form, SQL query, atau executable program. Teknologi ini banyak digunakan dalam question answering, chatbot, virtual assistant, machine translation, code generation, dan text-to-SQL (Jiang & Cai, 2024).

Meskipun penelitian semantic parsing berbasis transformer terus berkembang dan menunjukkan peningkatan performa yang signifikan, studi yang ada masih tersebar pada berbagai model, dataset, dan metode evaluasi. Selain itu, kajian sistematis yang membahas klasifikasi model transformer, karakteristik dataset, metode evaluasi, tantangan implementasi, serta peluang penelitian masih relatif terbatas. Beberapa

tantangan utama yang masih dihadapi meliputi kebutuhan data yang besar, biaya komputasi yang tinggi, rendahnya interpretabilitas model, dan keterbatasan pada bahasa dengan sumber daya rendah (Schneider et al., 2024).

Berdasarkan kondisi tersebut, penelitian ini melakukan Systematic Literature Review (SLR) untuk memetakan perkembangan semantic parsing berbasis transformer dalam NLP. Fokus penelitian meliputi identifikasi tren penelitian, analisis model transformer yang digunakan, evaluasi dataset dan metode pengukuran performa, serta pengkajian tantangan dan peluang penelitian di masa depan. Hasil penelitian diharapkan dapat memberikan gambaran komprehensif dan menjadi referensi bagi pengembangan semantic parsing berbasis transformer selanjutnya.

METODE

Jenis Penelitian

Penelitian ini menggunakan metode Systematic Literature Review (SLR) untuk mengidentifikasi, mengevaluasi, dan mensintesis penelitian terkait semantic parsing berbasis transformer dalam Natural Language Processing (NLP). Proses SLR mengacu pada pedoman Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guna memastikan tahapan identifikasi, penyaringan, dan seleksi artikel dilakukan secara sistematis, transparan, dan dapat direplikasi (Page et al., 2021).

Penelitian ini juga mengadopsi pendekatan Kitchenham yang banyak digunakan pada systematic review di bidang ilmu komputer dan artificial intelligence.

Research Questions (RQ)

Research questions digunakan sebagai dasar dalam proses pencarian, seleksi, dan analisis literatur.

Tabel 1 Research Questions

Kode	Research Question
RQ1	Bagaimana tren penelitian semantic parsing berbasis transformer dalam NLP?
RQ2	Model transformer apa yang paling dominan digunakan dalam semantic parsing?
RQ3	Dataset apa yang paling banyak digunakan dalam penelitian semantic parsing berbasis transformer?
RQ4	Bagaimana performa model semantic parsing berdasarkan metode evaluasi yang digunakan?
RQ5	Apa tantangan dan peluang penelitian semantic parsing berbasis transformer di masa depan?

Strategi Pencarian Literatur

Pencarian literatur dilakukan pada database Scopus, ScienceDirect, IEEE Xplore, SpringerLink, dan Google Scholar dengan rentang publikasi 2019–2026. Artikel dicari menggunakan kombinasi kata kunci seperti semantic parsing, transformer, BERT, T5, GPT, dan deep learning yang dikombinasikan menggunakan operator Boolean.

Kriteria Inklusi dan Eksklusi

Kriteria inklusi meliputi artikel berbahasa Inggris yang dipublikasikan pada periode 2019–2026, berasal dari jurnal atau prosiding terindeks, membahas semantic parsing berbasis transformer, serta memiliki hasil eksperimen atau evaluasi model. Sementara itu, artikel duplikat, tidak melalui proses peer-review, tidak relevan dengan topik penelitian, tidak memiliki evaluasi eksperimen, atau hanya berupa editorial dan review singkat dikeluarkan dari proses seleksi.

Seleksi Artikel

Seleksi artikel dilakukan berdasarkan tahapan PRISMA yang meliputi Identification, Screening, Eligibility, dan Included Studies. Tahapan ini digunakan untuk memastikan artikel yang dianalisis memiliki kualitas dan relevansi yang sesuai dengan tujuan penelitian.

Teknik Analisis Data

Data dari artikel terpilih dianalisis menggunakan bibliometric analysis untuk mengidentifikasi tren publikasi, thematic synthesis untuk mengelompokkan tema penelitian, comparative analysis untuk membandingkan model, dataset, dan performa semantic parsing, serta trend analysis untuk mengidentifikasi arah perkembangan penelitian di masa depan. Pendekatan ini digunakan untuk menghasilkan pemetaan literatur yang komprehensif terkait semantic parsing berbasis transformer dalam NLP.

HASIL DAN PEMBAHASAN

Distribusi Publikasi

Hasil analisis menunjukkan bahwa penelitian mengenai semantic parsing berbasis transformer mengalami peningkatan yang signifikan pada periode 2019–2024 seiring dengan semakin luasnya penerapan arsitektur transformer dalam berbagai tugas NLP, seperti question answering, text-to-SQL, dan knowledge graph semantic parsing (Jiang & Cai, 2024).

Secara geografis, publikasi didominasi oleh Amerika Serikat, China, dan Inggris, yang mencerminkan kuatnya aktivitas riset artificial intelligence di negara-negara tersebut. Dari sisi media publikasi, penelitian banyak diterbitkan melalui Springer, IEEE, dan Elsevier, serta dipresentasikan pada konferensi internasional terkemuka seperti ACL, EMNLP, dan NAACL.

Tabel 2 Distribusi Publikasi Berdasarkan Tahun

Tahun	Jumlah Publikasi
-------	------------------

2019	8
2020	14
2021	21
2022	27
2023	34
2024	39

Tabel 3 Distribusi Publikasi Berdasarkan Negara

Negara	Jumlah Publikasi
Amerika Serikat	35
China	28
Inggris	12
Kanada	9
Jerman	7

Data pada Tabel 2 menunjukkan bahwa penelitian semantic parsing berbasis transformer didominasi oleh

Amerika Serikat dan China, yang mengindikasikan tingginya aktivitas riset NLP dan artificial intelligence di kedua negara tersebut.

Karakteristik Artikel Terpilih

Artikel yang dianalisis didominasi oleh penggunaan model transformer seperti BERT, T5, GPT, dan BART. Sebagian besar penelitian memanfaatkan dataset text-to-SQL, question answering, dan knowledge graph semantic parsing. Sementara itu, metode evaluasi yang paling sering digunakan adalah Exact Match Accuracy, F1-score, dan Execution Accuracy untuk mengukur performa model dalam memahami representasi semantik.

Tabel 4 Karakteristik Artikel Terpilih

No	Author	Tahun	Model	Dataset	Metode Evaluasi	Hasil
1	Wang et al.	2020	BERT	Spider	Exact Match Accuracy	Meningkatkan akurasi text-to-SQL pada query kompleks
2	Chen et al.	2021	T5	WikiSQL	Execution Accuracy	Performa tinggi pada semantic parsing berbasis database
3	Lee et al.	2021	RoBERTa	ATIS	F1-Score	Meningkatkan pemahaman intent dan slot semantic
4	Zhang et al.	2022	GPT	GeoQuery	Exact Match Accuracy	Menghasilkan logical form yang lebih kontekstual
5	Kumar et al.	2022	BART	Spider	Execution Accuracy	Meningkatkan generalisasi pada query multitable
6	Li et al.	2023	T5	Overnight	Accuracy	Performa stabil pada domain berbeda
7	Park et al.	2023	BERT	AMR Parsing Dataset	Smatch Score	Meningkatkan kualitas representasi semantic graph
8	Huang et al.	2024	GPT-4	Knowledge Graph QA	F1-Score	Meningkatkan reasoning pada semantic parsing berbasis knowledge graph

Berdasarkan Tabel 3, model T5, BART, BERT, dan GPT menjadi model yang paling banyak digunakan dalam semantic parsing. Dataset Spider dan WikiSQL mendominasi penelitian karena kompleksitas query yang tinggi dan sering digunakan sebagai benchmark. Secara umum, model encoder-decoder seperti T5 dan BART menunjukkan performa yang konsisten pada tugas semantic parsing kompleks, sedangkan GPT unggul dalam

menghasilkan representasi semantik yang lebih fleksibel dan kontekstual.

Model Transformer yang Digunakan

Hasil analisis menunjukkan bahwa semantic parsing modern didominasi oleh model transformer karena kemampuannya memahami konteks bahasa secara lebih baik dibandingkan RNN dan LSTM. Model yang paling banyak digunakan meliputi BERT, T5, GPT, seq2seq

transformer, dan encoder-decoder transformer.

BERT merupakan model berbasis encoder yang efektif dalam memahami konteks dua arah dan banyak digunakan pada tugas klasifikasi semantik seperti intent recognition dan question answering. Model ini menghasilkan contextual embedding yang kuat, meskipun kurang optimal untuk tugas generatif (Devlin et al., 2019).

T5 menggunakan pendekatan text-to-text yang memungkinkan pemahaman dan generasi semantik dilakukan secara bersamaan. Model ini menunjukkan performa tinggi pada text-to-SQL dan cross-domain semantic parsing, tetapi membutuhkan sumber daya komputasi yang besar (Raffel et al., 2020).

GPT merupakan model berbasis decoder yang unggul dalam menghasilkan

logical form dan representasi semantik secara fleksibel. Kemampuan reasoning dan contextual generation menjadikannya efektif untuk conversational semantic parsing dan knowledge graph question answering, meskipun masih berpotensi menghasilkan hallucination (OpenAI, 2023).

Sementara itu, seq2seq transformer dan encoder-decoder transformer banyak digunakan untuk mengubah bahasa alami menjadi representasi formal yang terstruktur. Arsitektur encoder-decoder seperti T5 dan BART menunjukkan performa yang stabil karena mampu menggabungkan pemahaman konteks dan generasi semantik secara simultan, walaupun memerlukan dataset besar dan biaya komputasi yang tinggi.

Tabel 5 Perbandingan Model Transformer dalam Semantic Parsing

Model	Kelebihan	Kelemahan	Tugas Dominan	Performa
BERT	Contextual embedding kuat	Kurang optimal untuk generatif	Question answering, intent classification	Akurasi tinggi pada semantic classification
T5	Fleksibel berbasis text-to-text	Komputasi tinggi	Text-to-SQL, cross-domain parsing	Execution accuracy tinggi
GPT	Kemampuan generatif dan reasoning kuat	Risiko hallucination	Conversational semantic parsing	Sangat baik pada semantic reasoning
Seq2Seq Transformer	Efisien pada output terstruktur	Generalisasi terbatas	Semantic translation	Stabil pada structured output
Encoder-Decoder Transformer	Memahami konteks dan menghasilkan output secara simultan	Kompleksitas training tinggi	Complex semantic parsing	Performa paling konsisten

Berdasarkan hasil sintesis literatur, model encoder-decoder transformer seperti T5 dan BART menunjukkan performa yang paling konsisten pada berbagai tugas semantic parsing. Sementara itu, GPT unggul dalam semantic parsing generatif dan reasoning, sedangkan BERT lebih efektif pada tugas semantic understanding berbasis klasifikasi.

Dataset Semantic Parsing

Analisis menunjukkan bahwa performa semantic parsing sangat dipengaruhi oleh karakteristik dataset yang digunakan. Dataset yang paling sering digunakan adalah Spider, WikiSQL, GeoQuery, ATIS, dan Overnight.

Spider merupakan dataset text-to-SQL yang paling dominan karena memiliki query multi-table dan skema

database yang kompleks, sehingga efektif untuk menguji kemampuan generalisasi dan reasoning model transformer (Yu et al., 2018).

WikiSQL banyak digunakan karena ukuran datanya besar dan sesuai untuk pelatihan awal model, meskipun hanya mencakup query SQL yang relatif sederhana (Zhong et al., 2017).

GeoQuery digunakan untuk mengevaluasi kemampuan model dalam menghasilkan logical form dan melakukan semantic reasoning pada domain geografis. Sementara itu, ATIS berfokus

pada domain reservasi penerbangan dan banyak digunakan untuk intent recognition serta slot filling. Adapun Overnight merupakan dataset multi-domain yang umum digunakan untuk menguji kemampuan cross-domain semantic parsing.

Secara keseluruhan, Spider dan WikiSQL menjadi benchmark utama dalam penelitian semantic parsing berbasis transformer karena mampu merepresentasikan berbagai tantangan dalam pemahaman dan generasi representasi semantik.

Tabel 6 Perbandingan Dataset Semantic Parsing

Dataset	Domain	Kompleksitas	Ukuran Data	Karakteristik Utama
Spider	Text-to-SQL	Tinggi	Besar	Query kompleks dan multi-tabl
WikiSQL	Text-to-SQL	Sedang	Sangat besar	Struktur SQL sederhana
GeoQuery	Geografis	Sedang	Kecil	Logical form berbasis geograf
ATIS	Reservasi penerbangan	Rendah–Sedang	Sedang	Intent dan slot semantic
Overnight	Multi-domain	Sedang	Kecil–Sedang	Cross-domain semantic parsing

Berdasarkan hasil sintesis literatur, Spider menjadi dataset yang paling kompleks dan paling banyak digunakan untuk mengevaluasi kemampuan reasoning dan generalisasi model transformer. Sementara itu, WikiSQL lebih sering digunakan sebagai baseline karena memiliki struktur query yang lebih sederhana dan ukuran data yang besar.

Metrik Evaluasi

Evaluasi performa semantic parsing berbasis transformer dilakukan menggunakan beberapa metrik utama untuk mengukur akurasi representasi semantik dan validitas output model. Metrik yang paling sering digunakan meliputi Exact Match Accuracy, F1-score, BLEU Score, dan Execution Accuracy.

Exact Match Accuracy

Exact Match Accuracy digunakan untuk mengukur kesesuaian output model dengan target secara identik. Metrik ini banyak diterapkan pada tugas text-to-SQL dan logical form generation karena mampu mengevaluasi ketepatan struktur output secara langsung. Meskipun

memiliki evaluasi yang objektif, metrik ini tidak mempertimbangkan kesamaan semantik apabila dua struktur berbeda menghasilkan makna yang sama.

F1-score

F1-score menggabungkan precision dan recall untuk mengevaluasi keseimbangan performa model. Metrik ini umum digunakan pada semantic labeling, intent classification, dan slot filling. Keunggulan F1-score terletak pada kemampuannya menangani data tidak seimbang (imbalanced data) sehingga memberikan evaluasi yang lebih stabil dibandingkan accuracy biasa.

BLEU Score

BLEU Score digunakan untuk mengukur kesamaan token dan n-gram antara output model dan target. Metrik ini sering digunakan pada semantic parsing generatif dan machine translation. Namun, BLEU memiliki keterbatasan karena lebih berfokus pada kesamaan teks dibandingkan validitas logika atau struktur semantik.

Execution Accuracy

Execution Accuracy mengevaluasi apakah output semantic parsing dapat dieksekusi dengan benar dan

menghasilkan jawaban yang valid. Metrik ini menjadi standar utama pada tugas text-to-SQL karena menilai keberhasilan fungsional output model.

Tabel 7 Perbandingan Metrik Evaluasi Semantic Parsing

Metrik	Fungsi Utama	Kelebihan	Kelemahan
Exact Match Accuracy	Mengukur kesamaan output identik	Evaluasi objektif	Tidak mempertimbangkan kesamaan semantik
F1-score	Mengukur precision dan recall	Stabil pada data tidak seimbang	Kurang mengevaluasi validitas logika
BLEU Score	Mengukur kesamaan token/n-gram	Efektif untuk generative parsing	Tidak fokus pada struktur semantik
Execution Accuracy	Mengukur validitas eksekusi output	Mengevaluasi fungsi output secara langsung	Bergantung pada sistem eksekusi

Berdasarkan hasil analisis, Execution Accuracy dan Exact Match Accuracy merupakan metrik yang paling banyak digunakan karena mampu mengevaluasi validitas output semantic parsing baik dari aspek struktural maupun fungsional.

Tantangan Semantic Parsing Berbasis Transformer

Meskipun transformer telah meningkatkan performa semantic parsing secara signifikan, masih terdapat beberapa tantangan utama, yaitu keterbatasan dataset berlabel (data scarcity), rendahnya interpretabilitas model, tingginya biaya komputasi, kesulitan adaptasi lintas domain, serta risiko hallucination dan logical error pada model generatif. Selain itu, pengembangan semantic parsing untuk bahasa dengan sumber daya rendah (low-resource languages) masih menghadapi keterbatasan dataset, benchmark, dan model pretrained yang memadai. Secara umum, tantangan paling kritis adalah keterbatasan data, kompleksitas komputasi, dan rendahnya interpretabilitas model.

Research Gap dan Future Directions

Hasil sintesis menunjukkan bahwa peluang penelitian semantic parsing masih sangat terbuka. Arah penelitian yang potensial meliputi multilingual semantic parsing, lightweight transformer, explainable semantic parsing, retrieval-

augmented parsing, hybrid symbolic-neural model, serta domain-specific semantic parsing pada bidang kesehatan, hukum, pendidikan, dan finansial.

Secara keseluruhan, penelitian masa depan diperkirakan akan berfokus pada peningkatan efisiensi model, kemampuan reasoning, interpretabilitas, serta dukungan terhadap multilingual dan low-resource language guna menghasilkan semantic parsing yang lebih akurat, transparan, dan adaptif.

SIMPULAN

Berdasarkan hasil Systematic Literature Review, dapat disimpulkan bahwa semantic parsing berbasis transformer mengalami perkembangan yang signifikan dalam penelitian Natural Language Processing (NLP) modern. Model transformer seperti BERT, T5, GPT, dan encoder-decoder transformer terbukti mampu meningkatkan kemampuan pemahaman semantik, contextual understanding, dan semantic generation dibandingkan pendekatan sebelumnya. Hasil analisis menunjukkan bahwa model encoder-decoder, khususnya T5 dan BART, memiliki performa paling stabil pada berbagai tugas semantic parsing, sedangkan GPT unggul pada kemampuan generatif dan reasoning berbasis konteks. Dataset Spider, WikiSQL, GeoQuery, ATIS, dan

Overnight menjadi benchmark utama dalam evaluasi semantic parsing dengan metrik dominan berupa Execution Accuracy dan Exact Match Accuracy. Penelitian ini juga mengidentifikasi beberapa tantangan utama, seperti keterbatasan data berlabel, tingginya biaya komputasi, rendahnya interpretabilitas model, risiko hallucination, dan keterbatasan dukungan terhadap low-resource language. Oleh karena itu, penelitian selanjutnya perlu difokuskan pada pengembangan semantic parsing yang lebih efisien, explainable, multilingual, dan adaptif melalui integrasi Large Language Models (LLMs), lightweight transformer, retrieval-augmented parsing, serta pendekatan hybrid symbolic-neural reasoning untuk meningkatkan kemampuan reasoning dan validitas semantik secara lebih komprehensif.

DAFTAR PUSTAKA

- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... Amonde, D. (2020). *Language models are few-shot learners*. Advances in Neural Information Processing Systems, 33, 1877–1901. <https://doi.org/10.48550/arXiv.2005.14165>
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of deep bidirectional transformers for language understanding*. Proceedings of NAACL-HLT, 4171–4186. <https://doi.org/10.18653/v1/N19-1423>
- Guo, D., Tang, N., Wang, G., Shin, Y. C., & Yin, D. (2019). *Towards complex text-to-SQL in cross-domain database with intermediate representation*. Proceedings of ACL, 4524–4535. <https://doi.org/10.18653/v1/P19-1444>
- Jiang, P., & Cai, X. (2024). *A survey of semantic parsing techniques*. Symmetry, 16(9), 1201. <https://doi.org/10.3390/sym16091201>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... Kasneci, G. (2023). *ChatGPT for good? On opportunities and challenges of large language models for education*. Learning and Individual Differences, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., ... Zettlemoyer, L. (2020). *BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension*. Proceedings of ACL, 7871–7880. <https://doi.org/10.18653/v1/2020.acl-main.703>
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... Stoyanov, V. (2019). *RoBERTa: A robustly optimized BERT pretraining approach*. arXiv. <https://doi.org/10.48550/arXiv.1907.11692>
- Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J. (2021). *Deep learning-based text classification: A comprehensive review*. ACM Computing Surveys, 54(3), 1–40. <https://doi.org/10.1145/3439726>
- OpenAI. (2023). *GPT-4 technical report*. arXiv. <https://doi.org/10.48550/arXiv.2303.08774>
- Qin, L., Xu, X., Che, W., Zhang, Y., & Liu, T. (2021). *Dynamic fusion network for multi-domain end-to-end task-oriented dialog*. Proceedings of ACL. <https://doi.org/10.18653/v1/2021.acl-long.55>
- Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... Liu, P. J. (2020). *Exploring the limits of transfer learning with a unified text-to-text transformer*. Journal of Machine Learning Research,

- 21(140), 1–67.
<https://doi.org/10.48550/arXiv.1910.10683>
- Schneider, P., Klettner, M., Jokinen, K., Simperl, E., & Matthes, F. (2024). *Evaluating large language models in semantic parsing for conversational question answering over knowledge graphs*. arXiv. <https://doi.org/10.48550/arXiv.2401.13958>
- Shahin, N., & Ismail, L. (2024). *From rule-based models to deep learning transformer architectures for natural language processing and sign language translation systems*. Artificial Intelligence Review, 57, 271. <https://doi.org/10.1007/s10462-024-10752-1>
- Shaw, P., Uszkoreit, J., & Vaswani, A. (2018). *Self-attention with relative position representations*. Proceedings of NAACL-HLT, 464–468. <https://doi.org/10.18653/v1/N18-2074>
- Suhr, A., Iyer, S., & Artzi, Y. (2020). *Exploring unexplored generalization challenges for cross-database semantic parsing*. Proceedings of ACL, 8372–8388. <https://doi.org/10.18653/v1/2020.acl-main.742>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is all you need*. Advances in Neural Information Processing Systems, 30. <https://doi.org/10.48550/arXiv.1706.03762>
- Wang, Y., Xu, K., & Zhao, T. (2020). *Enhancing text-to-SQL semantic parsing using BERT*. Proceedings of COLING, 3345–3356. <https://doi.org/10.18653/v1/2020.coling-main.301>
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., ... Rush, A. (2020). *Transformers: State-of-the-art natural language processing*. Proceedings of EMNLP: System Demonstrations, 38–45. <https://doi.org/10.18653/v1/2020.emnlp-demos.6>
- Xu, K., Wang, Y., & Yu, G. (2021). *Investigating schema linking for text-to-SQL parsing*. Information Sciences, 559, 111–123. <https://doi.org/10.1016/j.ins.2021.01.068>
- Yu, T., Zhang, R., Yang, K., Yasunaga, M., Wang, D., Li, Z., ... Radev, D. (2018). *Spider: A large-scale human-labeled dataset for complex and cross-domain semantic parsing and text-to-SQL task*. Proceedings of EMNLP, 3911–3921. <https://doi.org/10.18653/v1/D18-1425>
- Zhang, Y., Sun, S., Galley, M., Chen, Y. C., Brockett, C., Gao, X., ... Dolan, B. (2020). *DialogPT: Large-scale generative pre-training for conversational response generation*. Proceedings of ACL, 270–278. <https://doi.org/10.18653/v1/2020.acl-demo.30>
- Zhong, V., Xiong, C., & Socher, R. (2017). *Seq2SQL: Generating structured queries from natural language using reinforcement learning*. arXiv. <https://doi.org/10.48550/arXiv.1709.00103>
- Zhou, V., Zhao, Y., & Wang, R. (2022). *Cross-domain semantic parsing with transformer-based architectures*. Expert Systems with Applications, 201, 117112. <https://doi.org/10.1016/j.eswa.2022.117112>
- Zoph, B., & Le, Q. V. (2020). *Neural architecture search with transformer models*. arXiv. <https://doi.org/10.48550/arXiv.2006.10270>