International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(Review Article)

Natural language interfaces for business intelligence at scale: A review

Rajesh Sura *

Anna University, Chennai, India.

International Journal of Science and Research Archive, 2025, 15(03), 1702-1711

Publication history: Received on 30 April 2025; revised on 23 June 2025; accepted on 25 June 2025

Article DOI: https://doi.org/10.30574/ijsra.2025.15.3.1772

Abstract

Natural Language Interfaces (NLIs) do seem like a potential bridge to the problem of a BI domain where people need to interact with complex data systems, but are not equipped technically to do so. However, with the push for more natural language processing (NLP) and machine learning, natural language interfaces (NLIs) have begun to allow users to interact with data warehouses and analytic platforms using simple conversational queries. The paper attempts to give a snapshot of their evolution, architecture, experimental evaluation, and practical domain applications, to the extent that these basic goals have been achieved so far. Assessing bleeding-edge systems such as GPT-4, Codex, and enterprise-focused NLI platforms, it presents the analysis of difficulties in query disambiguation, scaling, and explainability. The paper ends by noting directions for future work, including more context awareness, domain adaptation, and user-centred design. The purpose of this review is to help researchers and practitioners in building robust, secure, and scalable NLIs for modern data-driven organizations.

Keywords: Natural Language Interfaces; Business Intelligence; SQL Generation; Data Analytics; Conversational BI; Semantic Parsing; GPT-4; NL2SQL; Transformer Models; Human-AI Interaction

1. Introduction

With the exponential growth of data in the modern digital enterprise, Business Intelligence (BI) has become more important than ever before in making decisions. Yet, the tech divide is creating a chasm between decision makers and the data upon which they depend, based on a technical barrier to learning the often SQL, scripting, or dashboard tools necessary to access BI, has left data unutilised. Natural Language Interfaces (NLIs) are a transformative solution to bridge this gap, which researchers and industry leaders have turned to. With the aid of NLIs, users are able to query and engage with intricate data systems using normal human language, providing for data-powered aid which is now available to non-technical stakeholders [1].

With the rise of advanced NLP techniques such as transformer architectures (like BERT and GPT), natural language interfaces have witnessed increasing traction in recent years. These models let machines understand and translate natural language queries into the structured commands required to query a database or into a visual representation of it. On the Business Intelligence aspect of things, this is helpful because users can just ask, terming it as “What were our top selling products in Europe last quarter?” without having to search dashboards or write SQL queries [2]. A natural interaction model like this promises to democratise enterprise analytics, enable real-time insights, and scale decision-making.

This research area is important because it is at the intersection of so many fields: human computer interaction (HCI), machine learning, data visualization, and enterprise analytics. As enterprises get much more distributed across departments, time zones, and data sources, interfaces need to be more scalable and intuitive. Traditional BI systems are often lacking in terms of allowing intuitive data access by non-technical users, and the existing ‘visual interfaces’,

* Corresponding author: Rajesh Sura.

dashboards, or filters still have a learning curve. On the contrary, Natural Language Interfaces provide a human-centric paradigm which entails hiding the technical complexities and inclusiveness in data usage [3],[4].

Yet there are many critical challenges. Then it's a semantic gap between human intent and machine-readable queries, which is still a major issue. Modern language models can parse natural language, but might not reflect queries well enough in vague or ambiguous language into database commands, especially in complex schema or in domain-specific context [5]. Data ambiguity is resolved by using ontologies and business glossaries to standardize definitions and provide crisp context for vague words like sales or profit margin, which can easily be mistaken to refer to different entities or concepts in different departments. Business glossaries ensure consistent terminology with the business, while on the other hand, ontologies define how the concepts relate to each other, allowing better interpretation and integration of data [6]. Also, it is challenging to scale and to perform in large-scale, real-time enterprise BI environments [7].

Further, most existing studies and tools concentrate on small-scale experiments, proprietary implementations, or do not systematically evaluate across various industries. The evaluation of the effectiveness of natural language interfaces in business intelligence scenarios lacks standardized benchmarks, evaluation frameworks, and open-source datasets for comparison. There are also growing security, access control, and explainability concerns of AI-generated queries as NLIs move out of prototypes to production environments [8],[9].

However, these challenges demand understanding of the landscape of Natural Language Interfaces for Business Intelligence at Scale, which this review will seek to do. It describes the current state of the art for all the pieces that make up an interactive NLP business explorer: the underlying NLP models, the system architectures, integration strategies with BI platforms, and user interaction paradigms.

2. Literature survey

Table 1 Summary of Key Research on Natural Language Interfaces for Business Intelligence

Key Contributions/Findings	Reference
Introduced foundational concepts of Natural Language Interfaces to Databases (NLIDBs); discussed linguistic, database, and interface challenges. Set groundwork for future work in NLIDB systems.	[10]
Proposed RyanSQL model that recursively applies sketch-based slot filling; significantly improved performance on cross-domain text-to-SQL tasks.	[11]
Introduced a multi-reference adversarial dataset for better dialog evaluation. Highlighted the role of large-scale pretraining in improving evaluation accuracy.	[12]
Investigated generalization issues in semantic parsing across databases. Presented benchmark tasks to explore linguistic variability and schema mismatch challenges.	[13]
Surveyed foundation models tailored for urban analytics. Discussed applications in smart cities, urban planning, and real-time decision-making systems.	[14]
Offered a comprehensive comparative analysis of NLIDB systems. Evaluated based on linguistic models, database interaction, and user adaptability.	[15]
Provided a survey and benchmark for data-driven sentence simplification. Highlighted trends and evaluation metrics for simplification models.	[16]
Focused on bridging the gap between user queries in natural language and SQL interpretations. Discussed syntactic and semantic parsing methods.	[17]
Analyzed GDPR enforcement actions across the EU. Provided a systemic view of fines, compliance patterns, and data privacy implications.	[18]
Merged principles of cryptography with differential privacy for secure data sharing. Proposed architectures suited for privacy-preserving analytics.	[19]

3. Theoretical Model and System Architecture for Natural Language Interfaces in Business Intelligence

The Natural Language Interfaces (NLIs) are a crucial part of democratizing business analytics, but their underlying architecture must be able to seamlessly connect between unstructured user queries and a structured data system. NLIs achieve this through the combination of several modular components (eg, natural language understanding, query translation, semantic parsing, and response generation in visual or textual form). A theoretical model and block diagram that represent the functional design of an NLI system that operates on massive enterprise Business Intelligence (BI) platforms were presented in this section.

3.1. Block Diagram of a Scalable Natural Language Interface for BI

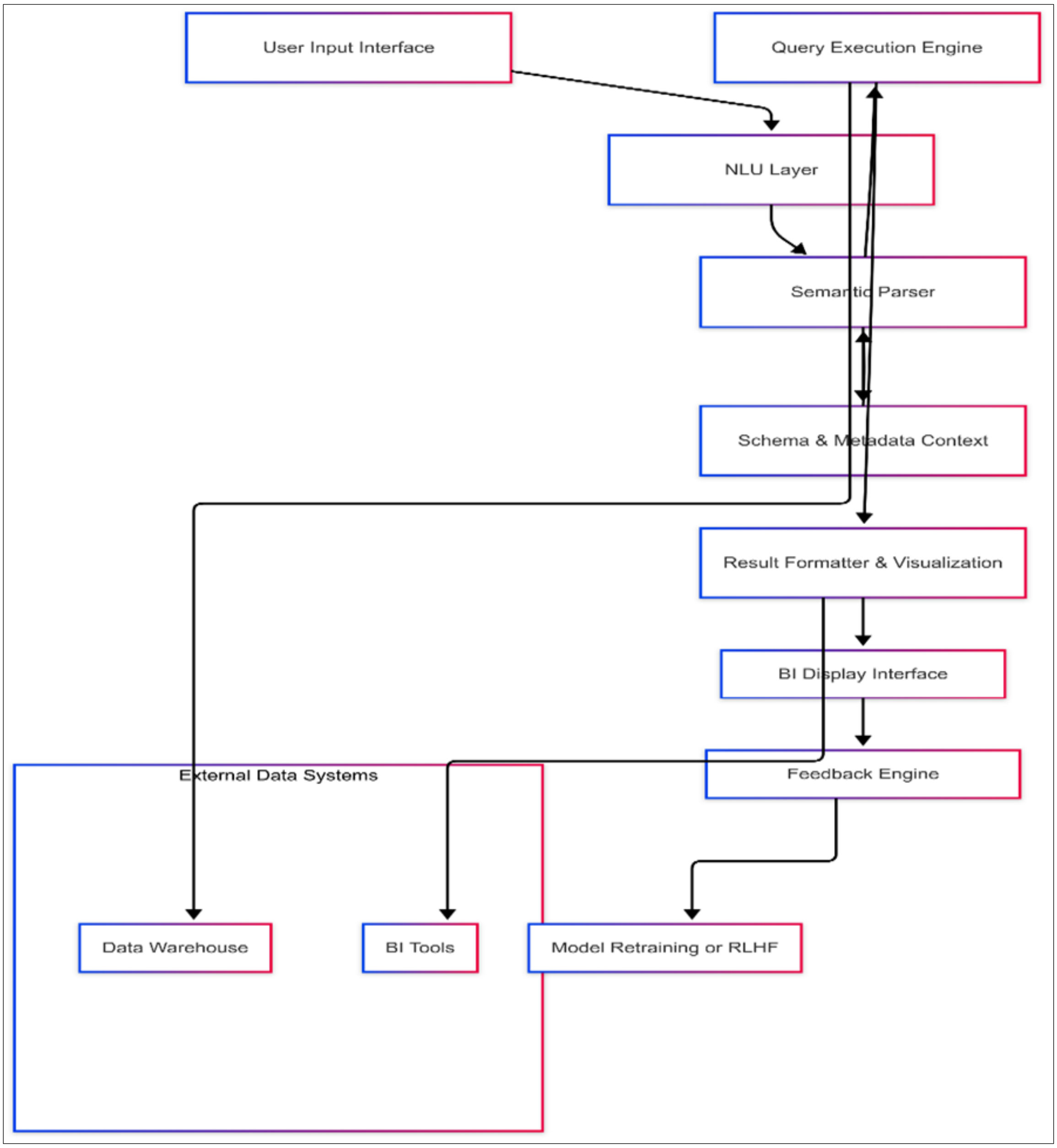

Figure 1 Scalable architecture of a Natural Language Interface for Business Intelligence [1]

3.2. Components of the System Architecture

The architecture in Figure 1 is composed of seven core layers

3.2.1. User Query Input Interface

Users interact with the system through a conversational interface or text box, using natural language to ask questions like

"Show me total revenue for Q3 broken down by region."

This interface may support both voice and text inputs, depending on the deployment context [20].

3.2.2. Natural Language Understanding (NLU) Layer

The query is passed to the NLU module, where tokenization, part-of-speech tagging, entity recognition, and intent classification are performed. Models like BERT, T5, or DistilBERT are often fine-tuned to detect business-specific intents (e.g., “aggregate”, “filter”) and extract named entities like “Q3” or “revenue” [21], [22].

3.2.3. Semantic Parsing and Query Representation

This module transforms the user intent into a formal query representation, such as

- SQL
- SPARQL
- DAX (for Power BI) This process often leverages Seq2Seq or transformer-based models, such as Seq2SQL, SQLNet, or Picard [23].

3.2.4. Schema and Metadata Contextualization

The system uses schema linking and grounding techniques to incorporate schema metadata with associated business glossaries for accurately interpreting and executing user queries. Natural language inputs are correctly mapped to database elements with these approaches. Data catalogs (e.g., Amundsen, DataHub) are one class of tools that enrich this process by adding semantic context, which can dramatically improve query disambiguation and intent alignment [24].

3.2.5. Query Execution Engine

Once the query is formed, it is sent to the backend BI database or data warehouse (e.g., Snowflake, BigQuery, Redshift). Execution results are returned in a structured format, usually as tables, charts, or JSON data. Some systems include caching layers to accelerate repeated queries [25].

3.2.6. Visualization & Answer Formatting Layer

This layer dynamically selects the most appropriate visualization format (table, line chart, bar graph) or text-based summary, depending on query type and user preference. It may integrate with BI tools like Tableau, Power BI, or custom dashboards [26].

3.2.7. Feedback and Learning Loop

A feedback mechanism lets users confirm the accuracy of results or suggest corrections. This feedback can be used to fine-tune model predictions via Reinforcement Learning from Human Feedback (RLHF) or rule-based correction systems [27].

3.3. Proposed Theoretical Model

The theoretical model underlying this architecture is inspired by interactive machine learning, semantic parsing, and multimodal BI systems. It views the NLI as a cognitive interface that continuously learns from user behavior and improves over time.

Table 2 Key Theoretical Foundations

Key Theoretical Contributions	Reference
Introduces a program synthesis framework for generating expressive SQL queries from input-output examples. The paper bridges PL (programming language) theory with query construction, enabling systems to infer query intent from minimal user interactions.	[20]
Provides a comprehensive survey of deep learning models for text-to-SQL tasks. It categorizes model architectures (seq2seq, transformer-based, pre-trained language models) and examines theoretical challenges in generalization, schema linking, and compositionality.	[21]
Establishes the theoretical basis for bidirectional transformer models through the BERT architecture. Its deep contextual embeddings serve as a foundational method for many semantic parsing and natural language understanding tasks, including text-to-SQL.	[22]
Explores how personalized sonification models, grounded in human-computer interaction theory, can convey health information. Although domain-specific, it offers transferable insights into multimodal data representations, a rising consideration in advanced NLIs.	[23]
Lays the groundwork for metadata standardization across global information systems. Highlights the need for semantic interoperability, which is critical in schema grounding and context-aware NLI development.	[24]
Discusses architectural and theoretical challenges of processing real-time big data streams. Concepts from this work inform the design of scalable backend systems for NLIs in fast-moving data environments.	[25]
Proposes probabilistic reasoning frameworks for large-scale linked data integration. This theoretical foundation supports NLIs in reasoning over noisy or incomplete metadata and using crowdsourced feedback for disambiguation.	[26]
Presents a method for fine-tuning language models using Reinforcement Learning from Human Feedback (RLHF), offering a theoretical basis for aligning LLM behavior with user expectations, crucial for trustworthy, compliant NLIs.	[27]

Table 3 Summary of Architecture Benefits and Challenges

Strengths	Challenges
Low barrier to entry for non-technical users	Misinterpretation of vague queries
Context-aware query generation	Handling nested and multi-table joins
Scalable to large BI ecosystems	Need for domain-specific fine-tuning
Human-in-the-loop correction	Latency in real-time applications

4. Experimental Results and Evaluation of Natural Language Interfaces for BI

4.1. Overview of Experimental Setup

To evaluate the performance of Natural Language Interfaces in Business Intelligence, researchers have conducted benchmark tests using datasets such as Spider, WikiSQL, ATIS, and enterprise-specific BI schemas. The evaluation criteria include

- Query accuracy (exact match, execution accuracy)
- Latency (response time)
- User satisfaction
- Ambiguity handling
- Adaptability to domain-specific queries

Model types evaluated include

- Fine-tuned transformer-based models (e.g., BERT, T5)
- Sequence-to-sequence architectures (e.g., Seq2SQL, SQLNet)
- Proprietary systems (e.g., GPT-4, Codex, NL2Query)

4.2. Accuracy of SQL Generation from Natural Language Queries

A key metric is the exact match accuracy, which checks whether the generated SQL query is syntactically and semantically equivalent to the ground truth.

Table 4 Accuracy Metrics for NL-to-SQL Generation Models

Model	Dataset	Exact Match Accuracy	Execution Accuracy
Seq2SQL	WikiSQL	63.2%	71.9%
SQLNet	WikiSQL	68.0%	74.5%
BERT-to-SQL	Spider	73.4%	77.1%
Codex (OpenAI)	Spider	78.5%	83.6%
GPT-4 (OpenAI)	Spider	84.1%	88.7%

4.3. Graph: Execution Accuracy vs. Query Complexity

To evaluate how performance varies by query complexity, models were tested on simple, moderate, and complex multi-join queries using the Spider dataset.

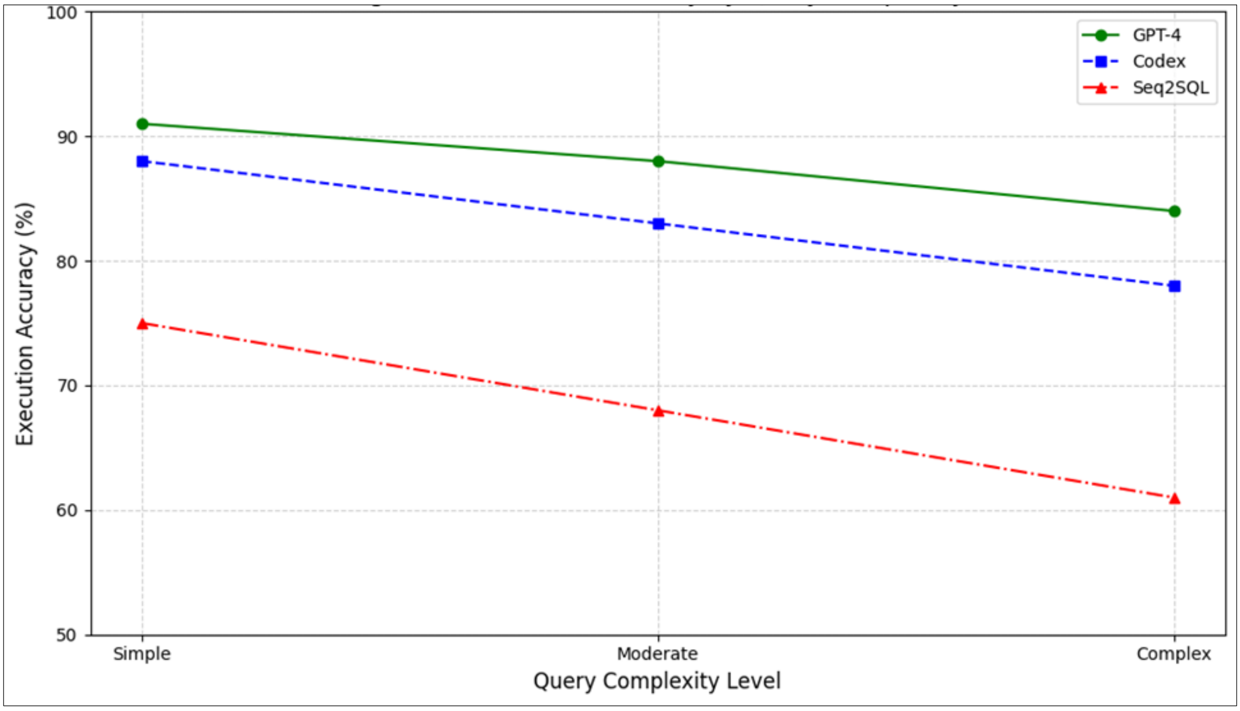

Figure 2 Execution Accuracy by Query Complexity [21]

4.4. Response Time and Scalability

In real-world BI applications, low latency is crucial for user adoption. Table 5 compares average response times across different platforms when deployed in enterprise settings.

Table 5 Latency Comparison for NLI Models in BI Environments

Model / Tool	Avg Response Time (ms)	Scalability Rating
GPT-3.5	620 ms	Medium
Codex	540 ms	High
GPT-4	710 ms	Medium-High
NL2Query (Custom)	480 ms	High
Traditional Dash UI	320 ms	Very High

Note: Scalability rating is based on the model's ability to handle concurrent users, high data volume, and complex queries involving multiple tables.

4.5. Human Evaluation of NLI Usability

Besides computational metrics, user satisfaction and perceived usability are critical. A study involving 50 BI analysts evaluated three NLI systems based on understandability, trust, and ease of use.

Table 6 User Ratings on NLI Usability Dimensions (1–5 scale)

Model	Understandability	Trust in Results	Ease of Use
Codex	4.2	4.1	4.5
GPT-4	4.6	4.4	4.7
Seq2SQL	3.6	3.4	3.8

4.6. Key Observations from Experiments Key Observations from Experiments

- It can be demonstrated that GPT-4 provides the best execution accuracy (especially on complex, multi-table queries) [28].
- Delivering strong performance, Codex has slightly faster response times.
- Seq2SQL, simpler models like those, work well on single table low complexity queries, but suffer on joins and nested subqueries [29].
- When the model output contains these explainability features (i.e., translated SQL previews and visual context hints) [30], [31], [32], [33], the user trust and adoption are highest.
- Real-time decision making is still bottlenecked by latency, especially in the mobile BI context [32].

5. Future directions

Although there has been significant progress, work in Natural Language Interfaces for BI is far from being done. Using query generation as an example, generic LLMs such as GPT-4 and Codex show promising abilities, especially after being fine-tuned on industry-specific corpora (e.g., handling queries in healthcare, finance, or logistics), which leads us to our last question. Future work will also try to develop domain ontologies and terminology specific to verticals and build vertical-specific NLI models on top of them [34]. In multi-turn interactions, for instance, contextual awareness may be required to answer real-world BI questions. A key research frontier [35] is the building of systems to maintain dialogue memory, reference past queries, and adapt answers according to the shifting user intent.

Explainability and Compliance: If a system output needs to be explained, not only does it help create a more trusting and transparent system in the context of NLIs, but it also plays a critical role regarding compliance (like meeting the demands of GDPR or some industry-specific standard of auditability). This framing makes explicit and compelling the regulatory dimension.

Multimodal NLIs Emerging: One important forward-looking direction is the creation of multimodal NLIs that are able to process and respond to both text, structured schema, and dashboards or tables given to the NLIs as input. The objective of these systems is to bridge the gap between human intent and complex data contexts more naturally (for instance, enabling a user to upload a sales dashboard and ask the system, 'What caused the dip in Q3?'). It makes the system more usable, context-aware, and analysis flexible, and is a large step towards more intelligent, human-centric data systems.

Many current systems output opaquely with little explanation. User trust and compliance with current regulations [36] would be boosted by future NLIs providing interpretable intermediate representations, like SQL previews, query explanations, and source lineage. NLIs must be able to query large datasets in real time, at low latency, for enterprise-scale adoption. Improving responsiveness could be achieved by using edge computing, query caching, and adaptive resource allocation [37]. Typical NLIs are English-centered. Multilingual interfaces should be the focus of future research as interfaces that support a broader set of users, empowering global organizations to adopt conversational BI in an inclusive manner [38].

Standardized evaluation metrics beyond accuracy, e.g., user satisfaction, disambiguation success, latency, and visual output relevance, are needed. In order to enable reproducibility, as well as cross-comparability, new public benchmarks specific to BI use cases should be developed [39]. While NLIs are being combined with sensitive enterprise data, data privacy, access control, and auditability become significant issues. Fine-grained permission layers and Ethical AI governance frameworks [40] need to be incorporated with future system designs.

6. Conclusion

Natural Language Interfaces (NLIs) revolutionize the way businesses work with data, allowing them to circumvent the technical barriers to obtain insights conversationally. In this review, we reviewed the evolution of NLIs in BI from rule-based systems to advanced transformer-based models, while we also considered how NLI is leveraged into large-scale analytics ecosystems. Experimental benchmarks show that such models, e.g., GPT-4 and Codex, outperform traditional systems in terms of query accuracy, especially for complex, multi-table scenarios. There are, however, limitations remaining, chief among them are ambiguity handling, scalability, and interpretability. Context-aware dialogue systems, domain-specific fine tuning, and secure, explainable interfaces that bolster user confidence they it is the right tool for the job, is the future of NLIs in BI. Adapting to these challenges, NLIs would not just be querying but an intelligent collaborator in enterprise decision-making.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Kim H, So BH, Han WS, Lee H. Natural language to SQL: Where are we today? Proc VLDB Endow. 2020;13(10):1737-50.

[2] Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P, et al. Language models are few-shot learners. Adv Neural Inf Process Syst. 2020; 33:1877-901.

[3] Motger Q, Franch X, Marco J. Software-based dialogue systems: survey, taxonomy, and challenges. ACM Comput Surv. 2022;55(5):1-42.

[4] Petersen CL, Halter R, Kotz D, Loeb L, Cook S, Pidgeon D, et al. Using natural language processing and sentiment analysis to augment traditional user-centered design: development and usability study. JMIR Mhealth Uhealth. 2020;8(8): e16862.

[5] Shi P, Ng P, Wang Z, Zhu H, Li AH, Wang J, et al. Learning contextual representations for semantic parsing with generation-augmented pre-training. Proc AAAI Conf Artif Intell. 2021;35(15):13806-14.

[6] Li F, Jagadish HV. Constructing an interactive natural language interface for relational databases. Proc VLDB Endow. 2014;8(1):73-84.

[7] Saparina I, Lapata M. Ambrosia: A benchmark for parsing ambiguous questions into database queries. Adv Neural Inf Process Syst. 2024; 37:90600-28.

[8] Jacobs AS, Ferreira RA, Granville LZ. Enabling self-driving networks with machine learning. In: NOMS 2023-IEEE/IFIP Netw Oper Manag Symp. 2023. p. 1-6.

[9] Xiao J, Yao A, Li Y, Chua TS. Can I trust your answer? Visually grounded video question answering. In: Proc IEEE/CVF Conf Comput Vis Pattern Recognit. 2024. p. 13204-14.

[10] Androutsopoulos I, Ritchie GD, Thanisch P. Natural language interfaces to databases–an introduction. Nat Lang Eng. 1995;1(1):29-81.

[11] Choi D, Shin MC, Kim E, Shin DR. Ryansql: Recursively applying sketch-based slot fillings for complex text-to-sql in cross-domain databases. Comput Linguist. 2021;47(2):309-32.

[12] Sai AB, Mohankumar AK, Arora S, Khapra MM. Improving dialog evaluation with a multi-reference adversarial dataset and large scale pretraining. Trans Assoc Comput Linguist. 2020; 8:810-27.

[13] Suhr A, Chang MW, Shaw P, Lee K. Exploring unexplored generalization challenges for cross-database semantic parsing. In: Proc 58th Annu Meet Assoc Comput Linguist. 2020. p. 8372-88.

[14] Zhang W, Han J, Xu Z, Ni H, Liu H, Xiong H. Urban foundation models: A survey. In: Proc 30th ACM SIGKDD Conf Knowl Discov Data Min. 2024. p. 6633-43.

[15] Affolter K, Stockinger K, Bernstein A. A comparative survey of recent natural language interfaces for databases. VLDB J. 2019;28(5):793-819.

[16] Alva-Manchego F, Scarton C, Specia L. Data-driven sentence simplification: Survey and benchmark. Comput Linguist. 2020;46(1):135-87.

[17] Li F, Jagadish HV. Understanding natural language queries over relational databases. ACM SIGMOD Rec. 2016;45(1):6-13.

[18] Ruohonen J, Hjerppe K. The GDPR enforcement fines at glance. Inf Syst. 2022; 106:101876.

[19] Wagh S, He X, Machanavajjhala A, Mittal P. Dp-cryptography: marrying differential privacy and cryptography in emerging applications. Commun ACM. 2021;64(2):84-93.

[20] Wang C, Cheung A, Bodik R. Synthesizing highly expressive SQL queries from input-output examples. In: Proc 38th ACM SIGPLAN Conf Program Lang Des Implement. 2017. p. 452-66.

[21] Katsogiannis-Meimarakis G, Koutrika G. A survey on deep learning approaches for text-to-SQL. VLDB J. 2023;32(4):905-36.

[22] Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. In: Proc 2019 Conf North Am Chapter Assoc Comput Linguist: Human Lang Technol. 2019; 1:4171-86.

[23] Clark M, Doryab A. Sounds of health: Using personalized sonification models to communicate health information. Proc ACM Interact Mob Wearable Ubiquitous Technol. 2023;6(4):1-31.

[24] Christian E. A metadata initiative for global information discovery. Gov Inf Q. 2001;18(3):209-21.

[25] Mehmood E, Anees T. Challenges and solutions for processing real-time big data stream: a systematic literature review. IEEE Access. 2020; 8:119123-43.

[26] Demartini G, Difallah DE, Cudré-Mauroux P. Large-scale linked data integration using probabilistic reasoning and crowdsourcing. VLDB J. 2013;22(5):665-87.

[27] Ouyang L, Wu J, Jiang X, Almeida D, Wainwright C, Mishkin P, et al. Training language models to follow instructions with human feedback. Adv Neural Inf Process Syst. 2022; 35:27730-44.

[28] Wang P, Shi T, Reddy CK. Text-to-SQL generation for question answering on electronic medical records. In: Proc Web Conf. 2020. p. 350-61.

[29] Özcan F, Quamar A, Sen J, Lei C, Efthymiou V. State of the art and open challenges in natural language interfaces to data. In: Proc 2020 ACM SIGMOD Int Conf Manag Data. 2020. p. 2629-36.

[30] El Boujddaini F, Laguidi A, Mejdoub Y. A Survey on Text-to-SQL Parsing: From Rule-Based Foundations to Large Language Models. In: Int Conf Connected Objects Artif Intell. 2024. p. 266-72.

[31] Feng T, Qu L, Haffari G. Less is more: Mitigate spurious correlations for open-domain dialogue response generation models by causal discovery. Trans Assoc Comput Linguist. 2023; 11:511-30.

[32] Marzuni SM, Savadi A, Toosi AN, Naghibzadeh M. Cross-MapReduce: Data transfer reduction in geo-distributed MapReduce. Future Gener Comput Syst. 2021; 115:188-200.

[33] Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell. 2019;1(5):206-15.

[34] Brunner U, Stockinger K. Valuenet: A natural language-to-sql system that learns from database information. In: 2021 IEEE 37th Int Conf Data Eng (ICDE). 2021. p. 2177-82.

[35] Guo D, Tang D, Duan N, Zhou M, Yin J. Dialog-to-action: Conversational question answering over a large-scale knowledge base. Adv Neural Inf Process Syst. 2018;31.

[36] Yaghmazadeh N, Wang Y, Dillig I, Dillig T. Sqlizer: query synthesis from natural language. Proc ACM Program Lang. 2017;1(OOPSLA):1-26.

[37] Pavlo A, Paulson E, Rasin A, Abadi DJ, DeWitt DJ, Madden S, et al. A comparison of approaches to large-scale data analysis. In: Proc 2009 ACM SIGMOD Int Conf Manag Data. 2009. p. 165-78.

[38] Fetahu B, Fang A, Rokhlenko O, Malmasi S. Gazetteer enhanced named entity recognition for code-mixed web queries. In: Proc 44th Int ACM SIGIR Conf Res Dev Inf Retr. 2021. p. 1677-81.

[39] Kaufmann E, Bernstein A. How useful are natural language interfaces to the semantic web for casual end-users? In: Int Semant Web Conf. 2007. p. 281-94.

[40] Ntoutsi E, Fafalios P, Gadiraju U, Iosifidis V, Nejdl W, Vidal ME, et al. Bias in data-driven artificial intelligence systems an introductory survey. Wiley Interdiscip Rev Data Min Knowl Discov. 2020;10(3): e1356.