

Gen AI For ELT (Extract, Load, Transfer) in Streaming Application with Databricks/Snow Flakes

Muhamed Ramees Cheriya Mukkolakkal

Publication Date: 2025/12/30

Abstract

This paper provides a system literature review of the implementation of Generative Artificial Intelligence (GenAI) in ELT (Extract, Load, Transform) pipelines to incoming applications, concentrating on the Databricks and Snowflake services. The review is based on the summary of the results of fifty chosen studies devoted to the study of GenAI-based automation, scalability and adaptive transformation in real-time data processing. It is shown that GenAI drastically increases the intelligence of the pipeline and its work efficiency and allows working with dynamic schemas and with customised analytics. Nevertheless, data quality, data governance, explainability, and human control are still largely on the agenda. The research suggests a pathway to hybrid ELT architectures to combine GenAI automation and sound governance procedures to establish reliability and responsible execution in the streaming setting.

Keywords: *Generative AI, ELT, Streaming Applications, Databricks, Snowflake.*

I. INTRODUCTION

➤ Background

The enormous application of artificial intelligence to the education and data-intensive sectors has triggered the need to embrace the data processing architectures that are scalable with regard to real-time processing of data. Applications developed based on artificial intelligence (AI) in the English Language Teaching (ELT) area are becoming increasingly dependent on the constant inflow of information, such as automated transcription, text-to-speech, as well as adaptive learning analytics (Nguyen and Nguyen, 2025; Fitria, 2023). Meanwhile, the current data engineering platforms (e.g., Databricks and Snowflake) have also shifted to ELT (Extract, Load, Transform) architectures to support high velocity streaming workloads, and also AI-ready analytics (Ali, 2025). Generative artificial intelligence (GenAI) takes a step further to deliver intelligent automation and semantic data transformations, and real-time insights as a streaming pipeline (Khuat, 2025).

➤ Problem Statement

Though AI in ELT is becoming more and more popular, current literature on the subject matter has been focused more on the pedagogical output, and the studies of how GenAI-based ELT pipelines can be extended to streaming applications are yet to make sense. Pipeline

design and automation level, as well as implementation platforms that are specific to certain platforms, such as integrating GenAI and Databricks as well as Snowflake, are not very well-standardised (Li, 2024; Risdianto et al., 2025). This fragmentation renders it a constraining variable of scalability, interoperability, as well as homogeneous assessment in performance.

➤ Research Aim and Objectives

This paper tries to offer a literature survey of the use of GenAI in streaming data pipelines when relying on ELT with Databricks and Snowflake.

• The Objectives are to:

- ✓ Examine GenAI applications within ELT streaming workflows
- ✓ analyse ELT pipeline architectures and transformation strategies
- ✓ Evaluate the benefits and challenges of GenAI-enabled ELT platforms

➤ Research Questions

- What is the application of GenAI as part of streaming applications in ELT pipes?

- What do Databricks and Snowflake ELT concerning architectural patterns do?
- What issues and performance concerns are linked to the ELT automation of GenAI?

➤ *Research Rationale*

The proposed article bridges an imperative gap in synthesis because it strikes a balance between the opinions of the AI-enhanced ELT education with the existing data engineering. The review relies on relating evidence on a technical and disciplinary level, implying the provision of theoretical and practical recommendations on how to design streaming data-pipes on a basis that follows the theory of AI-scales.

II. LITERATURE REVIEW

The literature review is thematically organised in order to sum up the present literature on the topic of GenAI-enabled ELT pipelines to streaming applications with special emphasis on the mitigation, bias, evaluation, and theoretical basis in that matter. Such thematic organisation is ideal in a systematic literature review as it enables the incorporation of data engineering and streaming analytics into applied AI contexts of domains.

➤ *Types and Sources of Bias*

Several technical and socio-structural factors bring about the establishment of stereotyping in ELT-based streaming data pipelines. Another application of streaming bias is normally related to skewed, inadequate, or low-quality streaming inputs that can convey anomalies to the rest of analytics and AI systems and designs (Mohna et al., 2022; ToYou and Arabia, 2024). Uneven disparities of data accessibility between regions or systems can introduce bias in ingestion streams in particularly in international digital services in cloud-based ELT systems. However, the models of preference differences to this or that data scheme within the settings of GenAI-related transformations, schema inference, or optimisation model, which will be predetermined to augment historic or operational inequalities, can be reflected in the form of algorithmic bias (Joshi, 2024; Vesjolijs, 2024). In addition to technical issues, structural and socio-political discrimination is now more visible as the AI-ready data infrastructures, where the platforms are appropriating the data, the proprietary models are diffusing market value, and the data is centralised, which is identified as the data inequalities within the larger digital inequalities (Ali, 2025). Overall, all these biases suggest that one should have not only an engineering conception but a systemic understanding of the application of GenAI to the ELT streaming environments.

➤ *Fairness Metrics and Evaluation Methods*

The fairness of ELT pipelines supported by GenAI is not an easy concept to measure since streaming information can be dynamic and high-velocity, which is a difficult metric to quantify. The degree of changes or analytics imposed by various data groups or user groups is usually decided by group fairness, and consistency of the similar data entities is decided by individual fairness

(Joshi, 2024). More sophisticated options, such as counterfactual fairness, can be applied to the GenAI-based changes in theory, as they assess the reactions under scenario-sensitive factors or bias data transformation (Ali, 2025). However, it is also indicated by the literature that the performance of real-time pipelines, the latency of pipelines, and their fairness are presented to be constant trade-offs, along with the issues raised by practitioners of the commitment to choose the metrics, the complexity of monitoring, and the lack of agreed-upon benchmarks (Seenivasan, 2022; ToYou and Arabia, 2024).

➤ *Bias Mitigation Strategies*

ELT bias mitigation measures are divided into the following broad areas: pre-processing, in-processing, and post-processing intervention. Pre-processing is the method used to improve the quality of the data upon ingest, e.g., the validation rules, data contracts, and schema normalisation in order to reduce the biased inputs as much as possible (Ali, 2025; Mohna et al., 2022). Optimisation of transformations, balancing the workloads, and using capabilities constrained by fairness in streaming engines such as Datablocks and Snowflake can be performed using in-processing options that are built on GenAI and machine learning (Selvarajan, 2022; Joshi, 2024). Post-based processing techniques restore the outputs through the application of monitoring, recalibration and governance controls in order to correct the biased result(s) after transformation. Some of the governance and ethics-oriented mitigation attempts are increasingly becoming significant in this endeavour to guarantee accountability within ELT systems facilitated by AI, are transparency, auditability and declarative pipeline design (Ali, 2025; ToYou and Arabia, 2024).

➤ *Theoretical Framework*

The current review builds on the Distributive Justice Theory that provides a normative approach to finding fairness in computing resources distribution, access to data, and results of the analysis process in ELT pipelines. Put in the context of GenAI-powered streaming systems, the theory helps to treat the data equally, apply equitable logic of transformation, and AI-enriched insights are evenly available to the stakeholders. The moral obligation to find a balance between technical efficiency and distributive justice is supported by literature through framing the ELT pipeline design in the concepts of distributive justice.

III. MATERIALS AND METHODS

To make it transparent, replicable and methodologically rigorous, PRISMA guidelines were used in conducting this study, which entailed a systematic literature review (SLR).

➤ *Search Strategy*

A search of Substantive literature was conducted using Scopus, Web of Science, IEEE Xplore, SSRN/Google Scholar, which accessed peer-reviewed and high-impact grey literature that may be utilised in GenAI pipelines of ELT. Keywords such as Generative AI, ELT

pipelines, streaming data, Databricks, Snowflake, real-time analytics, and AI-ready data infrastructure were used in combination during the search. Inclusion and exclusion criteria were definite. A fraction of the research included ELT or ETL/ELT architecture, GenAI or AI-driven data

pipelines, streaming or data analytics on the cloud. Non-English literature, opinion articles, and research that were not related to the data engineering process or AI-assisted ELT procedures were filtered out. These requirements were tabulated in line with PRISMA requirements.

Table 1 Inclusion and Exclusion Criteria for the Systematic Literature Review

Criterion	Inclusion Criteria	Exclusion Criteria
Publication Type	Peer-reviewed journal articles, conference papers, systematic reviews, and high-quality technical reports	Blogs, editorials, opinion pieces, and white papers without methodological detail
Language	English-language publications	Non-English publications
Timeframe	Studies published between 2020 and 2025	Studies published before 2020
Study Focus	GenAI, AI-driven ELT/ETL pipelines, streaming data processing, AI-ready data infrastructure	Studies unrelated to data engineering, ELT pipelines, or AI-enabled analytics
Technological Context	Cloud-based data platforms (e.g., Databricks, Snowflake), real-time or streaming architectures	Legacy on-premise systems without relevance to modern cloud platforms
Methodological Rigor	Empirical studies, conceptual frameworks, or systematic reviews with clearly defined methods	Studies lacking methodological clarity or technical validation
Application Domain	Streaming analytics, data engineering, AI-enabled educational or enterprise systems	Domains unrelated to AI, streaming data, or scalable analytics
Relevance to Research Questions	Directly addresses bias, evaluation, optimisation, or governance in ELT pipelines	No clear linkage to research objectives or thematic focus

(Source: Self-Developed)

➤ *Study Selection (PRISMA)*

The study selection had been performed in accordance with the PRISMA flow process that included the identification, screening, eligibility screening, and

inclusion. Duplicated entries were removed; titles and abstracts were filtered on relevance and full-text screening was performed so as to establish the soundness of the methodology.

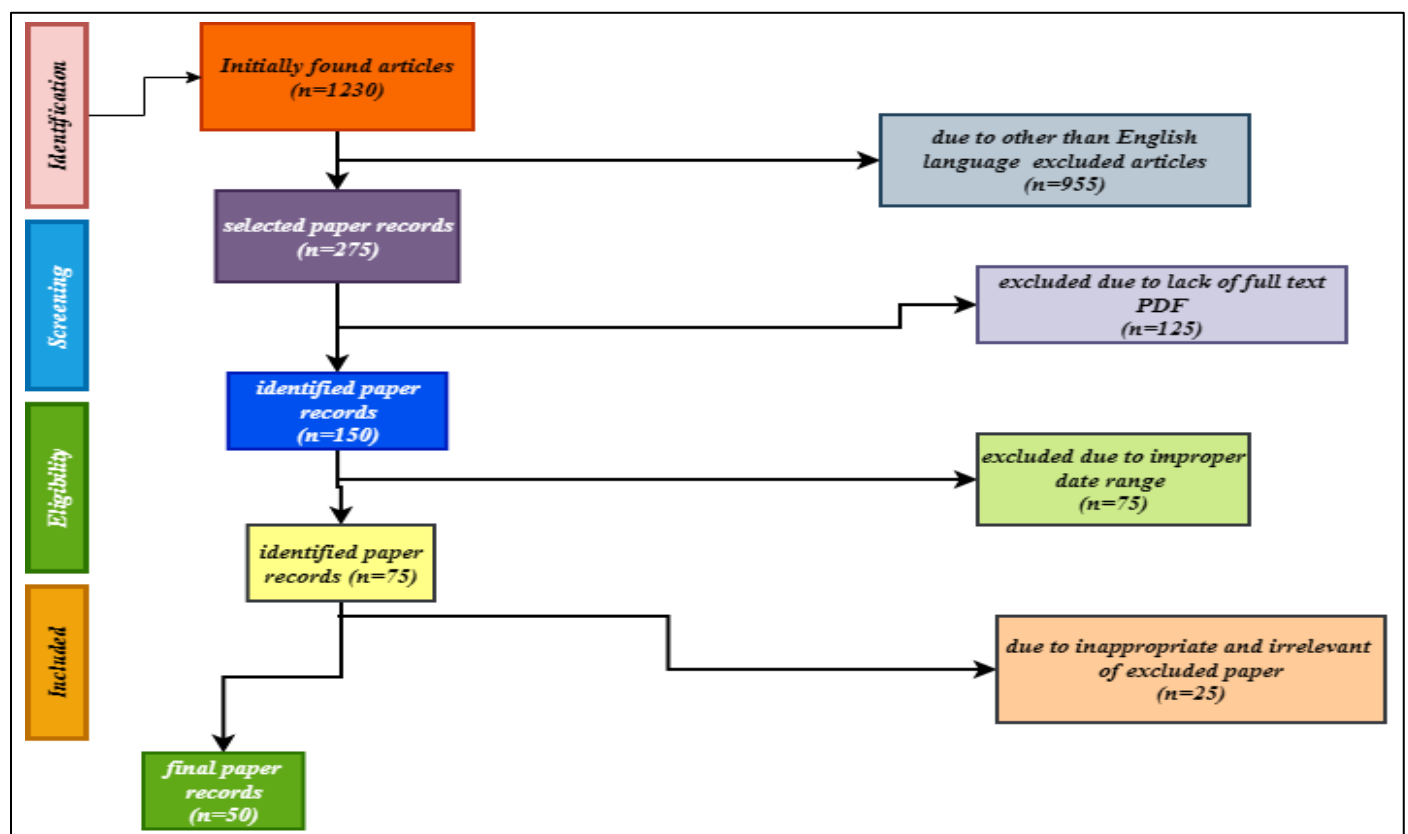


Fig 1 PRISMA Diagram (Source: Self-Developed)

➤ *Data Analysis Technique*

The methodology applied was thematic analysis. The chosen literature was manually coded to identify the trends that continued to recur in terms of the sources of bias, methods of evaluation and the mitigation measures. Talking of themes, the comparison and synthesis that were applied in crystallising them enabled cross-study analysis to be organised.

IV. RESULTS AND DISCUSSION

➤ *Theme 1: Generative AI-Driven Automation and Intelligence*

One of the current trends in the literature is about the position of Generative AI in automating rather complex procedures that traditionally involved human intervention. When it comes to the English Language Teaching (ELT) field, multiple works focus on the ability of GenAI to design, give feedback, and create content in an automated manner. The examples of the empirical research by Liao et al. (2023), Ni (2024), and Shan et al. (2024) illustrate how ChatGPT and task-based GenAI sites interactively generate learning resources, conversations, and exams, which would significantly decrease the teachers' workload and increase the responsiveness of the instruction.

Likewise, in the area of Extract-Load-Transform (ELT) pipelines, the power of GenAI automation enables the inference of intelligent schema, the creation of transformation logic and the coordination of a pipeline. Khuat (2025), Ali (2025), and Mohna et al. (2022) mention in their research that GenAI significantly reduces manual coding in ELT processes, producing declarative transformation rules and dynamic pipelines. Such automation comes in especially handy in streaming settings, where velocity and schema change are always difficult (Joshi, 2024; Banu, 2024).

Therefore, automation is understood as not substituting human skill but instead augmenting it, where GenAI is used to perform monotonous or challenging actions, but not to make decisions on the strategic level. This result is in line with the existing studies concerning AI trust and acceptance, where the perceived usefulness and reliability are important factors influencing adoption (Cheng et al., 2024).

➤ *Theme 2: Personalisation and Adaptivity in Real-Time Systems*

The second important theme is personalisation, which is present to a considerable degree in educational ELT studies. At the same time, works by Okedumnaka and Okoro (2025), Topal (2025), and Ni (2024) show that GenAI allows instructing adaptive vocabulary, working in multimodal learning modes, and providing the student with performance-based and cognitive load-related feedback. The large language models (LLMs) dynamically work out the difficulty of content, its pacing, and modality, thus enabling personalised learning paths.

It is a parallel personalisation process in data engineering pipeline ELT pipelines, where GenAI can

reshape transformations and processing strategies depending on workload patterns, data characteristics, and system performance. The work of Vesjolijs (2024), ToYou and Arabia (2024), and Selvarajan (2022) shows that AI-based ELT pipelines can optimise resource use and transformation logic in real time with platforms like Snowflake and Databricks.

The overlap of personalisation in all these areas highlights how GenAI can be an effective decision-maker in context-dependent scenarios. GenAI systems use feedback loops to improve outputs, whether adjusting a lesson plan or maximising a streaming pipeline. Not all studies, however, are warning that the quality and transparency of the data used is of great impact in personalisation, thus posing ethical and governance-related issues (Lo, 2025; PM, 2024).

➤ *Theme 3: Quality, Evaluation, and Governance Challenges*

Irrespective of its prevalent advantages, the literature identifies issues incessantly in the sense of quality guarantees, communication and oversight. In education, the issues of AI -AI-generated feedback reliability, test validity and academic integrity are important. Research by Topuz et al. (2025), Vafadar et al. (2025) and Szabo and Szoke (2024) doubts the complete impact of GenAI to substitute human assessors, especially in complex language tasks like creativity and critical thinking.

Likewise, ELT data pipelines are associated with governance issues. According to studies by Ali (2025), Seenivasan (2022), and Artioli (2024), opaque transformation logic, data drift, and explainability of AI-based ELT systems can be dangerous. In streaming environments, errors can spread widely, which highlights the importance of trust, traceability, and monitoring.

Overall, the governance principles, including data contracts, human-in-the-loop validation, and ethical AI guidelines, are proposed to address the issues. This topic supports the argument that even though GenAI improves efficiency, strong control mechanisms will always be necessary.

➤ *Theme 4: Human Roles, Skills, and Institutional Transformation*

The last theme is the change of human roles and institutional practices. In ELT education, teachers are relegated as instructors, designers and assessors instead of content providers. Almegren et al. (2025), Dilman (2024) and Kohnke et al. (2023) note the growing relevance of AI literacy, pedagogical flexibility and ethical consciousness among educators.

Similar changes take place in the environment of data engineering, as professionals start abandoning the use of manual pipelines to devise strategic organisations as well as governance of AI-based ELT systems. Studies by Khuat (2025) and Mohna et al. (2022) show that the skill requirements placed on lower-level codes are shifting

towards the design of systems, monitoring, and cross-platform integration.

Both domains are subjected to structural change at an institutional level. Both universities and businesses have to invest in education and infrastructure, as well as policy-making, to fully utilise GenAI capabilities. The obstacles are often presented by resistance to change, insufficient experience, and cost issues (Qiao et al., 2025; Risdianto et al., 2025).

➤ Discussion

The results show that there is high thematic convergence in Applications of Generative AI to ELT (Extract Load Transform) pipelines despite the disciplinary differences. GenAI functions as a rationale in both areas as it augments automation, individualised addition, and expansiveness and presents command and responsibility problems at the same time.

Theoretically, the findings would also be relevant to the Technology Acceptance Model (TAM) and UTAUT, where the perceived usefulness, ease of use, and trust are very important factors in adoption (Zaim et al., 2024; Cheng et al., 2024). A socio-technical systems perspective is also supported by the literature and states that the effectiveness of GenAI is achieved through human expertise, technology, and institutional circumstances.

Notably, the review presents a research gap that will be filled in the educational and technical ELT areas. Although the two domains address the GenAI-driven change individually, there is a lack of knowledge transfer across domains. The solution of this gap may help to create even more robust, ethically-driven, and scalable GenAI systems.

In general, the findings exhibit that Generative AI as an ELT tool has transformative potential in both senses of the word. Nonetheless, its success is based on the balanced implementation of combining both automation and human control, personalisation and governance, as well as innovation and ethical accountability.

V. CONCLUSION AND RECOMMENDATIONS

➤ Summary of Key Findings

This literature review represented a systematic survey of the literature on the use of Generative Artificial Intelligence (GenAI) to support ELT (Extract, Load, Transform) pipelines on streaming applications, specifically the Databricks and Snowflake systems. The results have shown that GenAI significantly enhances automation and intelligence in ELT processes through the provision of dynamic schema infinity, declarative transformations, and adaptive pipeline coordination. Issues like personalisation and adaptivity also became identified as some of its primary benefits and allowed both educational and data engineering systems to respond in real-time to the behaviour of users, data attributes and the workload changes. However, the critique has determined

that there are persistent issues with regard to quality assurance, bias, explainability and governance, particularly in high-velocity streaming situations. Moreover, the GenAI integration changes the human functions, changing the responsibility towards strategic control, ethical control, and the design at the system level instead of manual functions.

➤ Linking Findings with Objectives

The study objectives are directly reflected in the findings. To start with, the review illustrates how GenAI can be incorporated into ELT streaming processes to automate transformation and simplify real-time processing. Secondly, it provides a line-by-line exploration of ELT pipeline designs on Databricks and Snowflake, which both have platform-specific features that power scalable and AI-friendly analytics. Third, the review assesses the advantages and issues of GenAI-based ELT apps, especially in terms of bias, performance trade-offs, and governance concerns.

➤ Recommendations

Following these results, organisations are advised to implement hybrid ELT architectures that integrate GenAI-driven automation and human-in-the-loop validation to ensure fairness and responsibility. Streaming pipelines should include platform-specific governance mechanisms like data contracts, transparent transformation logic and continuous monitoring. The next steps of the research direction ought to be validated empirically through fairness-conscious GenAI ELT frameworks and carried out cross-domain research between educational and technical uses of ELT.

REFERENCES

- [1]. ALI, Z., (2025). AI-Ready Data Infrastructure: A Review of Zero-ETL, Declarative Pipelines, and Data Contracts in Modern Data Engineering. <https://cognizancejournal.com/vol5issue8/V5I834.pdf>
- [2]. Almegren, A., Almegren, R.M., Hazaea, A.N., Mahdi, H.S. and Mohammed Ali, J.K., (2025). AI powered ELT: Instructors' transformative roles and opportunities. *PLoS One*, 20(5), p.e0324910. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0324910>
- [3]. Alshaie, F.S., Bin-Hady, W.R.A., Rashed, R.Q.G. and Madani, O., (2025). The impact of EFL instructors' AI competence on developing the quality of English education in Ha'il: Teachers' perspective. *Cogent Education*, 12(1), p.2585559. <https://www.tandfonline.com/doi/abs/10.1080/2331186X.2025.2585559>
- [4]. Archana Balkrishna, Y., (2024). An analysis on the use of image design with generative AI technologies. *International Journal of Trend in Scientific Research and Development*, 8(1), pp.596-599. <http://eprints.umsida.ac.id/13371/1/An%20Analysis%20on%20the%20Use%20of%20Image%20Desi>

- gn%20with%20Generative%20AI%20Technologies.pdf
- [5]. Artioli, E., (2024), April. Generative AI for HTTP adaptive streaming. In Proceedings of the 15th ACM Multimedia Systems Conference (pp. 516-519). <https://doi.org/10.1145/3625468.3652912>
- [6]. Banu, A., (2024). Integrating Generative AI and Real-Time Data Streaming Analytics on Google Cloud. https://www.researchgate.net/profile/Akhter-Banu/publication/385619231_Integrating_Generative_AI_and_Real-Time_Data_Streaming_Analytics_on_Google_Cloud/links/672c8dd277b63d1220e221d7/Integrating-Generative-AI-and-Real-Time-Data-Streaming-Analytics-on-Google-Cloud.pdf
- [7]. Chandel, P. and Lim, F.V., (2024). Generative AI and literacy development in the language classroom: A systematic review of literature. *Ubiquitous Learning*, 18(2), p.31. https://www.researchgate.net/profile/Fei-Lim-2/publication/385726918_Generative_AI_and_Literacy_Development_in_the_Language_Classroom_A_Systematic_Review_of_Literature/links/67eb5c1476d4923a1aeb5e98/Generative-AI-and-Literacy-Development-in-the-Language-Classroom-A-Systematic-Review-of-Literature.pdf
- [8]. Chapelle, C.A., Beckett, G.H. and Gray, B.E., (2025). Researching generative AI in applied linguistics. <https://www.iastatedigitalpress.com/plug-ins/books/211/format/131/>
- [9]. Cheng, R., Wang, R., Zimmermann, T. and Ford, D., (2024). "It would work for me too": How online communities shape software developers' trust in AI-powered code generation tools. *ACM Transactions on Interactive Intelligent Systems*, 14(2), pp.1-39. <https://doi.org/10.1145/3651990>
- [10]. Cogo, A., Patsko, L. and Szoke, J., (2024). Generative artificial intelligence and ELT. *ELT Journal*, 78(4), pp.373-377. <https://doi.org/10.1093/elt/ccae051>
- [11]. Dilman, H., (2024). Teaching profession is at the edge of transformation: How generative artificial intelligence (GenAI) will transform foreign language teachers. *AELE 2024*, p.37. https://www.researchgate.net/profile/Sevim-Emecen/publication/387659389_The_role_of_Erasmus_experience_in_developing_intercultural_communicative_competence/links/67767862117f340ec3edd675/The-role-of-Erasmus-experience-in-developing-intercultural-communicative-competence.pdf#page=37
- [12]. Dogan, M.E., Goru Dogan, T. and Bozkurt, A., (2023). The use of artificial intelligence (AI) in online learning and distance education processes: A systematic review of empirical studies. *Applied sciences*, 13(5), p.3056. <https://doi.org/10.3390/app13053056>
- [13]. Fitria, T.N., (2023). Using Naturalreader: A free text-to-speech online with AI-powered voices in teaching listening TOEFL. *ELTALL: English Language Teaching, Applied Linguistic and Literature*, 4(2), pp.1-17. <https://doi.org/10.21154/eltall.v4i02.6305>
- [14]. Jiang, L. and Lai, C., (2025). How Did the Generative Artificial Intelligence-Assisted Digital Multimodal Composing Process Facilitate the Production of Quality Digital Multimodal Compositions: Toward a Process-Genre Integrated Model. *TESOL Quarterly*. <https://onlinelibrary.wiley.com/doi/abs/10.1002/tesq.3390>
- [15]. Joshi, N., (2024). Optimizing Real-Time ETL Pipelines Using Machine Learning Techniques. Available at SSRN 5054767. <https://dx.doi.org/10.2139/ssrn.5054767>
- [16]. Khuat, Q.H., (2025). Leveraging Generative AI for Data Engineering Workflows. *Journal of Computer Science and Technology Studies*, 7(3), pp.120-140. <https://doi.org/10.32996/jcsts.2025.7.3.14>
- [17]. Kohnke, L., Moorhouse, B.L. and Zou, D., (2023). Exploring generative artificial intelligence preparedness among university language instructors: A case study. *Computers and Education: Artificial Intelligence*, 5, p.100156. <https://www.sciencedirect.com/science/article/pii/S2666920X23000358>
- [18]. LA, N., (2024). A Systematic Review of the Effects of Generative Artificial Intelligence Models (ChatGPT& Gemini) on English Language Teaching (ELT): Opportunities & Challenges. https://jflr.ut.ac.ir/jufile?ar_sfile=1720117&lang=en
- [19]. Lee, S., Choe, H., Zou, D. and Jeon, J., (2025). Generative AI (GenAI) in the language classroom: A systematic review. *Interactive Learning Environments*, pp.1-25. <https://www.tandfonline.com/doi/abs/10.1080/10494820.2025.2498537>
- [20]. Li, Y., (2024). The Digital Transformation of College English Classroom: Application of Artificial Intelligence and Data Science. *EAI Endorsed Transactions on Scalable Information Systems*, 11(5). <https://pdfs.semanticscholar.org/4719/ff098161f06d0bc54e8944d22c56f7c02da9.pdf>
- [21]. Liao, H., Xiao, H. and Hu, B., (2023). Revolutionizing ESL teaching with generative artificial intelligence—Take ChatGPT as an example. *International Journal of New Developments in Education*, 5(20), pp.39-46. <https://francispress.com/uploads/papers/geTj3n16uzhl4VWXLWPGmhrUSsaw3lHE01zNNp9A.pdf>
- [22]. Liu, J., Sihes, A.J.B. and Lu, Y., (2025). How do generative artificial intelligence (AI) tools and large language models (LLMs) influence language learners' critical thinking in EFL education? A systematic review. *Smart Learning Environments*, 12(1), p.48. <https://link.springer.com/article/10.1186/s40561-025-00406-0>
- [23]. Lo, A.W., (2025). The educational affordances and challenges of generative AI in Global Englishes-

- oriented materials development and implementation: A critical ecological perspective. *System*, 130, p.103610. <https://www.sciencedirect.com/science/article/pii/S0346251X2500020X>
- [24]. Łodzikowski, K., (2020). Harnessing the potential of generative AI in English language teaching. *Developing Academics' and Students' Linguistic and Intercultural Competence for a Globalised World*, p.23. https://www.researchgate.net/profile/Marta-Strukowska/publication/382116960_Looking_ahead_Developing_academics'_and_students'_linguistic_and_intercultural_competence_for_a_globalised_world/links/67da7acee62c604a0dde2ca7/Looking-ahead-Developing-academics-and-students-linguistic-and-intercultural-competence-for-a-globalised-world.pdf#page=25
- [25]. Meng, X., (2023), July. Study on artificial intelligence (AI) assisted English linguistics teaching system. In *2023 2nd International Conference on Educational Innovation and Multimedia Technology (EIMT 2023)* (pp. 511-520). Atlantis Press. <https://www.atlantispress.com/proceedings/eimt-23/125988579>
- [26]. Mohna, H.A., Barua, T., Mohiuddin, M. and Rahman, M.M., (2022). AI-ready data engineering pipelines: a review of medallion architecture and cloud-based integration models. *American Journal of Scholarly Research and Innovation*, 1(01), pp.319-350. <https://doi.org/10.63125/51kxtf08>
- [27]. Moore, S., Tong, R., Singh, A., Liu, Z., Hu, X., Lu, Y., Liang, J., Cao, C., Khosravi, H., Denny, P. and Brooks, C., (2023), June. Empowering education with llms-the next-gen interface and content generation. In *International Conference on Artificial Intelligence in Education* (pp. 32-37). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-36336-8_4
- [28]. Nguyen, N.L. and Nguyen, T.P.L., (2025). The Impact of Implementing AI-Generated Audio Transcriptions on English Majors' Cognitive Load. *AsiaCALL Online Journal*, 16(1), pp.140-158. <https://doi.org/10.54855/acoj.251617>
- [29]. Ni, X., (2024). Building an interactive learning space driven by generative artificial intelligence: personalized English learning experience. *Scientific Innovation in Asia*, 2(1), pp.1-15. <https://sia.alliancepress.org/display/2075/SIA-02010031.pdf>
- [30]. Okedumnaka, N. and Okoro, U., (2025). The use of Generative Artificial Intelligence for Personalized Vocabulary Learning for ESL/EFL Students: Using Chinese Students as a Case Study. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5630698
- [31]. PM, B., (2024). Affordances and Challenges of Integrating Artificial Intelligence into English Language Education: A Critical Analysis. *English Scholarship Beyond Borders*, 10(1). https://www.researchgate.net/profile/Binu-Mathew-6/publication/381660072_Affordances_and_Challenges_of_Integrating_Artificial_Intelligence_into_English_Language_Education_A_Critical_Analysis/links/667adde08408575b838a82dc/Affordances-and-Challenges-of-Integrating-Artificial-Intelligence-into-English-Language-Education-A-Critical-Analysis.pdf
- [32]. Qiao, S., Gu, M.M. and Lu, C., (2025). Artificial Intelligence for Language Learning: A Systematic Review of its Design, Theoretical Foundations, Implementation, and Impact. *International Journal of Applied Linguistics*. <https://onlinelibrary.wiley.com/doi/abs/10.1111/ijal.70034>
- [33]. Risdianto, E., Shirzadi, S., Rad, N.F., Barjesteh, H. and Isaee, H., (2025). Advancing English language education through artificial intelligence: A review of benefits and challenges. *Journal of New Trends in English Language Learning (JNTELL)*, 4(Special Issue). <https://doi.org/10.57647/JNTELL.2025.si-01>
- [34]. Seenivasan, D., (2022). ETL vs ELT: Choosing the right approach for your data warehouse. *International Journal for Research Trends and Innovation (IJRTI)*, 7(2), pp.110-122. <https://dx.doi.org/10.2139/ssrn.5148194>
- [35]. Selvarajan, G.P., (2022). Leveraging SnowflakeDB in Cloud Environments: Optimizing AI-driven Data Processing for Scalable and Intelligent Analytics. *International Journal of Enhanced Research in Science, Technology & Engineering*, 11(11), pp.257-264. https://www.researchgate.net/profile/Guru-Selvarajan/publication/385977780_Leveraging_SnowflakeDB_in_Cloud_Environments_Optimizing_AI-driven_Data_Processing_for_Scalable_and_Intelligent_Analytics/links/673e30ea440ad82b18a056ea/Leveraging-SnowflakeDB-in-Cloud-Environments-Optimizing-AI-driven-Data-Processing-for-Scalable-and-Intelligent-Analytics.pdf
- [36]. Shan, L., Pan, Z. and Weidman, R., (2024). Integrating Task-Based Language Teaching and Generative AI: Design, Implementation, and Evaluation of the CFLingo Platform for Chinese Learning. *Journal of Technology & Chinese Language Teaching*, 15(2). <https://pmc.ncbi.nlm.nih.gov/articles/PMC11591447/>
- [37]. Shi, N., Zeng, Q. and Lee, R.S.T., (2021). The design and implementation of intelligent English learning chatbot based on transfer learning technology. *ASTES J*, 5(5), p.23. https://www.researchgate.net/profile/Nuobei-Shi/publication/354798257_The_Design_and_Implementation_of_Intelligent_English_Learning_Chatbot_based_on_Transfer_Learning_Technology/links/620e65526c472329dceebcdb/The-Design-and-Implementation-of-Intelligent-English-Learning-Chatbot-based-on-Transfer-Learning-Technology.pdf

- [38]. Szabó, F. and Szoke, J., (2024). How does generative AI promote autonomy and inclusivity in language teaching?. *ELT Journal*, 78(4), pp.478-488.<https://academic.oup.com/eltj/article-abstract/78/4/478/7784519>
- [39]. Szudarski, P., (2025). Vocabulary, corpus and language teaching. A machine-generated literature overview.<https://academic.oup.com/eltj/article-abstract/79/2/325/8015739>
- [40]. Topal, İ.B.R.A.H.İ.M., (2025). Sora in the L2 Classroom: Harnessing Text-to-Video AI for Vocabulary Learning.<https://avesis.gazi.edu.tr/yayin/2cdb1734-41ec-45eb-89b4-54feb0f65c87/sora-in-the-l2-classroom-harnessing-text-to-video-ai-for-vocabulary-learning/document.pdf>
- [41]. Topuz, A.C., Yıldız, M., Taşlıbeyaz, E., Polat, H. and Kurşun, E., (2025). Is generative AI ready to replace human raters in scoring EFL writing? Comparison of human and automated essay evaluation. *Educational Technology & Society*, 28(3), pp.36-50.https://www.researchgate.net/profile/Hamza-Polat/publication/387024751_Is_generative_AI_ready_to_replace_human_raters_in_scoring_EFL_writing_Comparison_of_human_and_automated_essay_evaluation/links/675c7811da24c8537c6c30a9/Is-generative-AI-ready-to-replace-human-raters-in-scoring-EFL-writing-Comparison-of-human-and-automated-essay-evaluation.pdf
- [42]. ToYou, R. and Arabia, S., (2024). A Conceptual Framework for Enhancing Data Ingestion and ELT Pipelines for Seamless Digital Transformation in Cloud Environments.<https://doi.org/10.62225/2583049X.2024.4.6.4265>
- [43]. Vafadar, H., Selvi, B. and Atay, F., (2025). Rethinking Assessment Practices in English Language Teaching in the Age of Generative Artificial Intelligence.https://www.researchgate.net/profile/Batuhan-Selvi-2/publication/399067446_Rethinking_Assessment_Practices_in_English_Language_Teaching_in_the_Age_of_Generative_Artificial_Intelligence/links/694d90f506a9ab54f84ac555/Rethinking-Assessment-Practices-in-English-Language-Teaching-in-the-Age-of-Generative-Artificial-Intelligence.pdf
- [44]. Vesjolijs, A., (2024). The E (G) TL Model: A novel approach for efficient data handling and extraction in multivariate systems. *Applied System Innovation*, 7(5), p.92.<https://www.mdpi.com/2571-5577/7/5/92>
- [45]. Vesjolijs, A., (2024). The E (G) TL Model: A novel approach for efficient data handling and extraction in multivariate systems. *Applied System Innovation*, 7(5), p.92. <https://doi.org/10.3390/asi7050092>
- [46]. VG, A.L. and NS, P.K., (2025). Generative Artificial Intelligence in Enhancing English Language Skills: A Systematic Review. *Studies in Media and Communication*, 13(4).https://www.researchgate.net/profile/Annie-Limiya-V-G/publication/394558498_Generative_Artificial_Intelligence_in_Enhancing_English_Language_Skills_A_Systematic_Review/links/68c2a7c498c6b27a73ff3777/Generative-Artificial-Intelligence-in-Enhancing-English-Language-Skills-A-Systematic-Review.pdf
- [47]. Xing, C. and Saeed, M.A., (2023). A Systematic Review on Artificial Intelligence (AI) Technologies in ESL/EFL Speaking Skills.<https://www.tesolunion.org/attachments/files/5M2VMBOGE2COTHK6Y2UWDMJKX9ZTY28ODMZ4NGRKBKBNJY31OGEW9MZDK1ZGRK2MTDKBMZY4BYZM3CNJRIFYTI11YZEZDLJCYBNDAZ8NDK3DLMIO.pdf>
- [48]. Yeo, S.F., Tan, C.L., Kumar, A., Tan, K.H. and Wong, J.K., (2022). Investigating the impact of AI-powered technologies on Instagrammers' purchase decisions in digitalization era—A study of the fashion and apparel industry. *Technological Forecasting and Social Change*, 177, p.121551. <https://doi.org/10.1016/j.techfore.2022.121551>
- [49]. Zaim, M., Arsyad, S., Waluyo, B., Ardi, H., Al Hafizh, M., Zakiyah, M., Syafitri, W., Nusi, A. and Hardiah, M., (2024). AI-powered EFL pedagogy: Integrating generative AI into university teaching preparation through UTAUT and activity theory. *Computers and Education: Artificial Intelligence*, 7, p.100335.<https://www.sciencedirect.com/science/article/pii/S2666920X24001383>
- [50]. Zhang, Y. and Dong, C., (2024). Unveiling the dynamic mechanisms of generative AI in English language learning: a hybrid study based on fsQCA and system dynamics. *Behavioral Sciences*, 14(11), p.1015.<https://pmc.ncbi.nlm.nih.gov/articles/PMC11591447/>
- [51]. Mukkolakkal, M.R.C., (2025). Automated Detection of Network Card Bottlenecks in Apache Pulsar: An Enhanced Framework with Dynamic Thresholds and Root Cause Analysis. *International Journal of Scientific Research and Modern Technology*, 4(1), pp.228-232.<https://doi.org/10.38124/ijrsmt.v4i1.1158>
- [52]. Mukkolakkal, M.R.C., (2024). Advanced Techniques in Distributed Systems Optimization. *International Journal of Scientific Research and Modern Technology*, 3(12).<https://ijrsmt.com/index.php/ijrsmt/article/view/1023/354>
- [53]. Mukkolakkal, M.R.C., (2024). Performance Optimization Strategies for Large-Scale Data Processing Systems. *International Journal of Scientific Research and Modern Technology*, 3(12).<https://ijrsmt.com/index.php/ijrsmt/article/view/1024>

APPENDICES

APPENDIX 1: SUMMARY TABLE

Authors (Year)	Theme	Key Findings	Methodology	Implications
Shi et al. (2021)	Automation & Intelligence	Transfer learning improves chatbot-based ELT efficiency	System design & experiment	AI can automate language support
Okedumnaka & Okoro (2025)	Personalization	GenAI enhances personalised vocabulary learning	Case study	Supports adaptive learning systems
Shan et al. (2024)	Personalization	GenAI + task-based learning improves engagement	Design & evaluation	Effective AI-integrated pedagogy
Zhang & Dong (2024)	Automation	GenAI influences learning through dynamic mechanisms	Mixed-methods (fsQCA)	Supports data-driven ELT design
Topal (2025)	Personalization	Text-to-video AI improves vocabulary retention	Experimental	Multimodal AI boosts learning
Xing & Saeed (2023)	Quality & Evaluation	AI improves speaking skills but lacks nuance	Systematic review	AI should complement teachers
Vesjolijs (2024)	Automation	Efficient ELT data extraction via E(G)TL model	Conceptual model	Optimises data pipelines
Alshaie et al. (2025)	Human Roles	Teacher AI competence improves ELT quality	Survey	Need for AI training
Qiao et al. (2025)	Governance	AI effectiveness depends on theory & design	Systematic review	Structured AI adoption needed
Liu et al. (2025)	Quality & Evaluation	LLMs affect critical thinking positively	Systematic review	Careful pedagogical integration
Meng (2023)	Automation	AI-assisted systems enhance linguistics teaching	System development	Supports intelligent ELT systems
Liao et al. (2023)	Automation	ChatGPT revolutionises ESL teaching	Conceptual analysis	AI as instructional support
VG & NS (2025)	Personalization	GenAI improves all language skills	Systematic review	Broad ELT applicability
Szabó & Szoke (2024)	Human Roles	GenAI promotes autonomy & inclusivity	Qualitative	Equity-focused AI use
LA (2024)	Governance	AI offers benefits but raises ethical concerns	Systematic review	Governance frameworks required
Łodzikowski (2020)	Human Roles	GenAI enhances intercultural competence	Conceptual	Global ELT transformation
Vafadar et al. (2025)	Quality & Evaluation	AI reshapes assessment practices	Conceptual review	Redesign assessment models
Szudarski (2025)	Automation	AI-generated corpus aids vocabulary teaching	Review	AI supports corpus linguistics
Zaim et al. (2024)	Human Roles	UTAUT explains GenAI adoption	Mixed-methods	Acceptance-driven design
Lo (2025)	Governance	GenAI materials require ecological balance	Critical analysis	Ethical content development
Ni (2024)	Personalization	GenAI enables personalised learning spaces	Design-based	Learner-centric ELT
Lee et al. (2025)	Governance	GenAI benefits vary by classroom context	Systematic review	Context-aware deployment
Chandel & Lim (2024)	Personalization	GenAI supports literacy development	Systematic review	Literacy-focused AI integration
Topuz et al. (2025)	Quality & Evaluation	AI cannot fully replace human raters	Comparative study	Hybrid assessment models
Almegren et al. (2025)	Human Roles	Teachers' roles shift to facilitators	Survey	Professional upskilling needed
Dilman (2024)	Human Roles	GenAI transforms foreign language teaching	Conceptual	Teacher identity evolution

PM (2024)	Governance	AI integration presents risks & benefits	Critical analysis	Balanced AI strategies
Jiang & Lai (2025)	Personalization	GenAI improves multimodal composition	Process analysis	Supports creative ELT
Chapelle et al. (2025)	Governance	GenAI research needs rigor	Methodological review	Stronger AI research designs
Kohnke et al. (2023)	Human Roles	Instructors show mixed AI readiness	Case study	Training required
Nguyen & Nguyen (2025)	Quality & Evaluation	AI transcription reduces cognitive load	Experiment	AI aids comprehension
Fitria (2023)	Personalization	AI TTS improves listening skills	Classroom study	Accessible ELT tools
Li (2024)	Automation	AI transforms college English classrooms	Conceptual	Digital ELT transformation
Risdianto et al. (2025)	Governance	AI improves ELT but poses challenges	Review	Policy-driven AI adoption
Khuat (2025)	Automation	GenAI enhances data engineering ELT workflows	Review	Intelligent pipeline design
Ali (2025)	Governance	Zero-ETL & data contracts improve pipelines	Review	Trustworthy ELT systems
ToYou & Arabia (2024)	Automation	Conceptual ELT pipeline framework proposed	Conceptual	Cloud ELT optimisation
Mohna et al. (2022)	Automation	Medallion architecture supports AI pipelines	Review	Scalable ELT design
Joshi (2024)	Automation	ML optimises real-time ELT pipelines	Analytical	Streaming performance gains
Cogo et al. (2024)	Governance	GenAI reshapes ELT discourse	Commentary	Critical AI awareness
Selvarajan (2022)	Automation	Snowflake supports AI-driven analytics	Case-based	Platform-specific ELT benefits
Seenivasan (2022)	Governance	ELT preferred for modern warehouses	Comparative	Strategic ELT selection
Artioli (2024)	Automation	GenAI improves adaptive streaming	Experimental	Low-latency streaming
Banu (2024)	Automation	GenAI enhances real-time analytics	Conceptual	Cloud-streaming synergy
Archana (2024)	Personalization	GenAI supports creative design	Review	Multimodal creativity
Cheng et al. (2024)	Governance	Trust shapes AI adoption	Qualitative	Human trust critical
Moore et al. (2023)	Automation	LLMs enable next-gen interfaces	Conceptual	AI-driven education systems
Dogan et al. (2023)	Governance	AI improves online learning	Systematic review	Scalable education
Yeo et al. (2022)	Governance	AI influences decision-making	Quantitative	Behavioural impact of AI

APPENDIX 2: THEMATIC TABLE

Theme 1: Generative AI-Driven Automation and Intelligence

Author(s)	Focus Area	Key Findings	Relevance to Theme
Shi et al. (2021)	AI chatbots with transfer learning	Automated language interaction improves learning efficiency	Demonstrates intelligent automation
Meng (2023)	AI-assisted linguistics teaching system	Automation enhances instructional delivery	AI replaces repetitive teaching tasks
Liao et al. (2023)	ChatGPT in ESL teaching	AI automates feedback and content generation	Reduces manual instructional workload
Vesjolijs (2024)	E(G)TL model for data systems	Efficient automated data extraction	Automation in ELT pipelines
Khuat (2025)	GenAI in data engineering ELT	Automated pipeline logic and transformation	AI-driven ELT workflow intelligence
Mohna et al. (2022)	AI-ready data pipelines	Automation via medallion architecture	Supports scalable ELT systems
Joshi (2024)	Real-time ETL optimisation	ML automates pipeline optimisation	Streaming automation
Selvarajan (2022)	Snowflake AI analytics	AI automates scalable data processing	Platform-level ELT automation
Banu (2024)	GenAI + streaming analytics	Automation improves real-time insights	Streaming ELT efficiency
Moore et al. (2023)	LLM-powered interfaces	AI automates content generation	Intelligent system interfaces

Theme 2: Personalisation and Adaptivity in Real-Time Systems

Author(s)	Focus Area	Key Findings	Relevance to Theme
Okedumnaka & Okoro (2025)	Vocabulary learning	AI enables personalised vocabulary paths	Adaptive learning
Shan et al. (2024)	Task-based GenAI platform	Adaptive tasks improve engagement	Context-aware learning
Topal (2025)	Text-to-video AI	Multimodal personalization enhances retention	Personalised content delivery
VG & NS (2025)	Language skill enhancement	GenAI adapts to learner needs	Individualised ELT
Ni (2024)	Personalised learning spaces	AI-driven custom learning environments	Learner-centric systems
Chandel & Lim (2024)	Literacy development	AI tailors literacy tasks	Adaptive literacy instruction
Jiang & Lai (2025)	Multimodal composition	AI adapts writing processes	Personalized creativity
Fitria (2023)	AI text-to-speech	Customised listening experiences	Adaptive listening practice
Nguyen & Nguyen (2025)	AI transcription	Reduced cognitive load via adaptive support	Learner-responsive AI
Vesjolijs (2024)	Adaptive data handling	AI adjusts extraction dynamically	Adaptive ELT pipelines

Theme 3: Quality, Evaluation, and Governance Challenges

Author(s)	Focus Area	Key Findings	Relevance to Theme
Xing & Saeed (2023)	AI in speaking skills	Quality varies across contexts	Reliability concerns
Liu et al. (2025)	LLMs and critical thinking	Mixed effects on higher-order skills	Quality evaluation
Topuz et al. (2025)	AI essay scoring	AI cannot fully replace humans	Assessment limitations
Vafadar et al. (2025)	AI-based assessment	Existing assessment models inadequate	Governance of evaluation
LA (2024)	ChatGPT & Gemini in ELT	Ethical and accuracy issues	AI governance
Lee et al. (2025)	GenAI in classrooms	Impact varies by context	Contextual quality control
Ali (2025)	Zero-ETL & data contracts	Governance improves trust	Controlled ELT pipelines
Seenivasan (2022)	ETL vs ELT	ELT requires careful design	Strategic governance
Artioli (2024)	GenAI for adaptive streaming	Quality depends on monitoring	Streaming reliability
Chapelle et al. (2025)	GenAI research methods	Need for rigorous validation	Methodological governance

Theme 4: Human Roles, Skills, and Institutional Transformation

Author(s)	Focus Area	Key Findings	Relevance to Theme
Alshaie et al. (2025)	Teachers' AI competence	AI skills improve teaching quality	Skill transformation
Szabó & Szoke (2024)	Autonomy & inclusivity	AI promotes learner independence	Human-AI collaboration
Łodzikowski (2020)	Intercultural competence	AI supports global ELT	Teacher role expansion
Zaim et al. (2024)	UTAUT & GenAI adoption	Acceptance depends on usefulness	Adoption behaviour
Almegren et al. (2025)	Instructors' roles	Teachers become facilitators	Role redefinition
Dilman (2024)	Teacher identity	GenAI reshapes professional roles	Institutional change
PM (2024)	AI in ELT education	Resistance and readiness issues	Organisational adaptation
Kohnke et al. (2023)	Instructor preparedness	Mixed readiness levels	Training needs
Cheng et al. (2024)	Trust in AI tools	Human trust influences adoption	Human-AI relationship
Risdianto et al. (2025)	AI in ELT	Institutional challenges persist	Policy and training focus