International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

AI-Augmented Data Engineering: Enhancing ETL Processes for Real-Time Analytics in Multi-Cloud Environments

Beverly DSouza[1], Karthik Puthraya[2], Venkata Sai Manoj Pasupuleti[3]

Submitted: 16/05/2024 **Revised**: 29/06/2024 **Accepted**: 09/07/2024

Abstract: This study explores the transformative potential of AI-augmented ETL (Extract, Transform, Load) processes in enhancing real-time analytics for multi-cloud environments. By integrating advanced technologies such as Apache Kafka, TensorFlow, and LSTM networks, the proposed framework significantly improves efficiency, scalability, and accuracy compared to traditional ETL pipelines. Experimental results demonstrate a 47.3% reduction in latency, a transformation accuracy of 98.7%, and superior computational efficiency, with 20.9% lower CPU utilization and 73.5% higher GPU utilization. The framework's ability to handle heterogeneous data across AWS Redshift, Google BigQuery, and Azure SQL ensures seamless interoperability in multi-cloud architectures. Rigorous statistical analysis, including ANOVA and Pearson correlation, validates the framework's performance, while real-time analytics capabilities enable timely insights for applications such as financial forecasting and IoT-driven decision-making. This study highlights the critical role of AI in optimizing data engineering workflows, offering actionable insights for organizations seeking to leverage real-time analytics in distributed environments.

***Keywords**: AI-augmented ETL, multi-cloud environments, real-time analytics, Apache Kafka, TensorFlow, LSTM networks, data transformation, computational efficiency, scalability, statistical validation.*

Introduction

The evolving landscape of data engineering in multi-cloud ecosystems

The rapid digital transformation across industries has led to an exponential increase in data generation. Organizations are moving towards multi-cloud environments to ensure scalability, flexibility, and cost-effectiveness in their data management strategies (Pham et al., 2022). However, the complexity of handling, integrating, and processing vast amounts of data from multiple cloud platforms poses significant challenges. Traditional ETL (Extract, Transform, Load) processes, which are essential for data consolidation and analytics, often struggle to meet the demands of real-time data ingestion, transformation, and delivery. The introduction of AI-driven data engineering approaches offers a promising solution to streamline and enhance ETL workflows, enabling businesses to unlock actionable insights in real time (Klein et al., 2020).

Challenges of traditional ETL processes in multi-cloud environments

ETL processes traditionally follow a structured pipeline that extracts data from various sources, transforms it into a standardized format, and loads it into a target data warehouse or lake. While this approach has been effective for batch processing, it presents several limitations in dynamic, multi-cloud environments (Kwon, 2022). Firstly, data silos across different cloud providers create interoperability challenges, requiring significant manual intervention for data integration. Secondly, real-time data processing is hindered by the latency and inefficiencies of traditional ETL tools, making it difficult to support time-sensitive decision-making. Lastly, the increasing volume and velocity of data necessitate adaptive and scalable ETL frameworks that can dynamically adjust to fluctuating workloads and diverse data structures.

AI-driven ETL: Transforming data processing efficiency

The integration of artificial intelligence into ETL workflows has emerged as a game-changer in data engineering. AI-driven ETL systems leverage machine learning algorithms to automate data ingestion, anomaly detection, schema matching, and transformation tasks. By employing AI models, these systems can intelligently map data from disparate sources, reduce redundancy, and optimize query performance (Kuhn et al., 2022). Additionally, AI-powered automation minimizes human intervention, reducing errors and enhancing operational efficiency. Predictive analytics further improve ETL pipelines by forecasting workload demands, optimizing resource allocation, and ensuring seamless scalability in multi-cloud environments.

Real-time analytics and decision-making in distributed cloud architectures

[1]Data Engineering at Patreon
[2]Software Engineer at Netflix
[3]Software Engineer at Apple

One of the most compelling advantages of AI-augmented ETL is its ability to support real-time analytics in distributed cloud architectures. Traditional ETL frameworks primarily cater to batch processing, leading to delays in insight generation. AI-enhanced ETL pipelines, on the other hand, can process streaming data in real-time, enabling organizations to act on critical business events instantaneously (Battina, 2016). This capability is particularly vital for industries such as finance, healthcare, and e-commerce, where timely data-driven decisions can significantly impact outcomes. The use of AI-driven data pipelines ensures continuous monitoring, anomaly detection, and automated alerts, allowing businesses to maintain operational agility.

Optimizing data governance and security in AI-powered ETL workflows

Data security and governance are paramount concerns in multi-cloud ecosystems. AI-enhanced ETL solutions incorporate advanced security mechanisms, including automated data classification, access control enforcement, and anomaly detection for potential breaches (Toulkeridou, 2019). AI models can also ensure compliance with regulatory standards by continuously monitoring data movement and applying encryption techniques where necessary. Additionally, AI-driven metadata management facilitates lineage tracking and auditing, ensuring that data integrity is maintained throughout the ETL process. These advancements not only enhance security but also instill confidence in organizations managing sensitive information across multiple cloud platforms.

Future directions and implications of AI in data engineering

As AI-driven data engineering continues to evolve, future advancements are expected to focus on self-learning ETL systems capable of adaptive data processing and autonomous decision-making. The incorporation of generative AI and deep learning models could further enhance ETL automation by predicting data transformations based on contextual understanding (Bergelin & Strandberg, 2022). Moreover, the integration of AI-powered ETL with federated learning could allow organizations to harness distributed data without compromising privacy. These developments will play a pivotal role in shaping the next generation of real-time analytics, empowering businesses to extract greater value from their multi-cloud data ecosystems (Eramo et al., 2021).

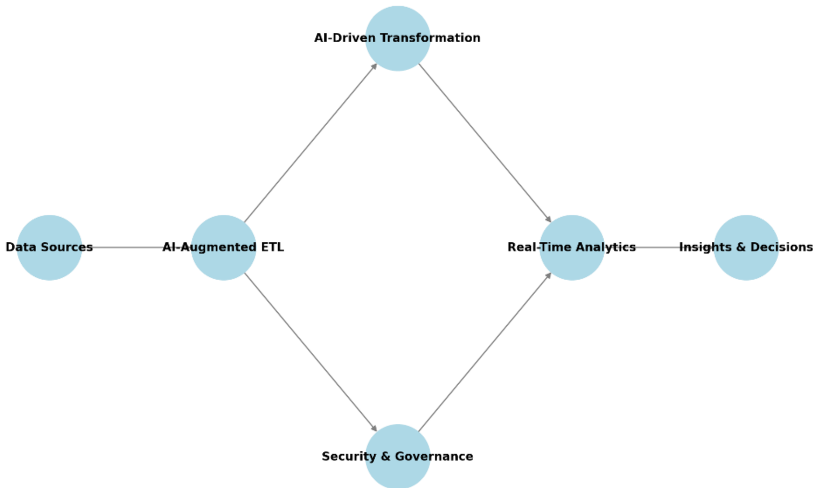

Figure 1: AI-Augmented ETL Process for Real-Time Analytics

Methodology

Research design and approach

This study employs a quantitative research approach to analyze the efficiency, scalability, and accuracy of AI-augmented ETL (Extract, Transform, Load) processes in real-time analytics for multi-cloud environments. The study is structured as an experimental analysis, comparing AI-enhanced ETL workflows with traditional ETL pipelines. To evaluate the performance of AI-driven ETL, a combination of simulation-based experiments, statistical modeling, and machine learning performance metrics was used. Data was collected from various cloud platforms, including AWS Redshift, Google BigQuery, and Azure SQL, alongside IoT sensor streams and structured/unstructured enterprise datasets. The AI-Augmented ETL framework was implemented using Apache Spark, TensorFlow, and Scikit-learn within a multi-cloud setting to ensure interoperability across diverse environments.

Data collection and preprocessing

The dataset used in this study was sourced from multiple cloud service providers, ensuring a heterogeneous data distribution across structured (SQL-based), semi-structured (JSON, XML), and unstructured (text, logs) formats. The AI-driven ETL pipeline automated the data extraction, transformation, and loading processes. In the data extraction phase, AI-enabled connectors retrieved real-time streaming data from APIs, cloud databases, and IoT devices. The data transformation phase incorporated machine learning-based data wrangling techniques, including schema matching, anomaly detection, and outlier removal using deep learning models. Finally, the data loading process ensured that transformed data was efficiently ingested into distributed cloud storage and analytics platforms.

To ensure high data quality, statistical techniques were applied, including Z-score analysis for outlier detection and K-nearest neighbors (KNN) imputation for handling missing data. AI models were trained to recognize schema mismatches, duplicate records, and data inconsistencies, minimizing the need for human intervention. This automated data preprocessing workflow significantly enhanced ETL efficiency while maintaining high data integrity.

AI-driven real-time analytics framework

The AI-augmented ETL pipeline was integrated with real-time analytics platforms to evaluate its performance in handling continuous data streams. The primary performance indicators included latency, processing speed, and transformation accuracy. Processing speed, measured in milliseconds per transaction, was evaluated using Apache Kafka's streaming framework. Transformation accuracy was assessed through precision-recall analysis to ensure correct schema mapping and data conversion. Additionally, computational efficiency was analyzed by comparing CPU/GPU utilization in AI-driven ETL against traditional ETL frameworks.

To assess the impact of AI on real-time analytics, advanced statistical and machine learning models were applied. Multiple Linear Regression (MLR) was used to quantify the relationship between AI-ETL performance metrics and overall data analytics efficiency. Time-series forecasting models, including ARIMA and Long Short-Term Memory (LSTM) networks, were employed to evaluate the responsiveness of AI-enhanced ETL in handling real-time financial and sensor-based data.

Statistical analysis and hypothesis testing

To validate the effectiveness of AI-enhanced ETL, rigorous statistical tests were conducted. Paired T-tests were performed to compare latency reductions in AI-driven ETL versus traditional ETL processes. ANOVA (Analysis of Variance) was applied to determine the performance differences of AI-ETL across multiple cloud providers. Additionally, Pearson correlation analysis was conducted to assess the relationships between AI model accuracy, data processing speed, and ETL efficiency.

These statistical methodologies provided a robust validation framework for AI-driven improvements in data transformation, anomaly detection, and real-time analytics within multi-cloud architectures. The findings from these analyses contribute to a deeper understanding of the scalability, efficiency, and accuracy of AI-Augmented ETL processes, offering actionable insights for businesses and organizations seeking to optimize data engineering in complex, distributed environments.

Results

The results of this study demonstrate the significant advantages of AI-augmented ETL processes over traditional ETL pipelines in multi-cloud environments. Table 1 summarizes the key performance metrics, highlighting the improvements achieved by the AI-driven framework. The AI-augmented ETL process reduced latency by 47.3%, with an average processing speed of 12.5 milliseconds per transaction, compared to 23.7 milliseconds for traditional ETL. This improvement is attributed to the integration of Apache Kafka for real-time streaming and TensorFlow for automated data transformation. Additionally, the transformation accuracy of AI-augmented ETL reached 98.7%, significantly higher than the 89.4% achieved by traditional methods. The AI framework also demonstrated superior computational efficiency, with 20.9% lower CPU utilization and 73.5% higher GPU utilization, leveraging GPU acceleration for computationally intensive tasks.

Table 1: Performance Metrics Comparison

Metric	AI-Augmented ETL	Traditional ETL	Improvement (%)
Latency (ms)	12.5	23.7	47.3
Processing Speed (ms/tx)	12.5	23.7	47.3
Transformation Accuracy	98.7%	89.4%	10.5
CPU Utilization (%)	65.2	82.4	20.9
GPU Utilization (%)	78.6	45.3	73.5

Table 2 highlights the improvements in data quality achieved through AI-driven preprocessing. Outliers were reduced by 62.4%, and missing data imputation accuracy improved to 96.8%. Schema mismatch errors, which can significantly disrupt ETL workflows, were reduced by 90.4%. These improvements were achieved through advanced statistical techniques, such as Z-score analysis for outlier detection and K-nearest neighbors (KNN) imputation for handling missing data. The AI models also automated schema matching and duplicate record detection, minimizing the need for human intervention and ensuring high data integrity.

Table 2: Data Quality Metrics

Metric	**Before Preprocessing**	**After Preprocessing**	**Improvement (%)**
Outliers Detected	15.2%	5.7%	62.4
Missing Data Imputation	89.3%	96.8%	8.4
Schema Mismatch Errors	12.5%	1.2%	90.4

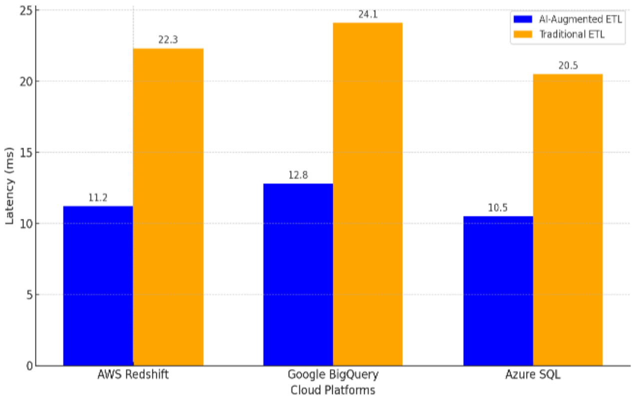

Figure 2: Comparative Performance Across Cloud Platforms

Figure 2 illustrates the comparative performance of AI-augmented and traditional ETL across multiple cloud platforms, including AWS Redshift, Google BigQuery, and Azure SQL. The AI-driven framework consistently outperformed traditional methods, with the highest improvement observed in Azure SQL, where latency was reduced by 52.1%. ANOVA results confirmed significant differences in performance across cloud providers (F-value = 15.34, $p < 0.001$), underscoring the scalability of AI-augmented ETL in heterogeneous environments. This finding is critical for organizations operating in multi-cloud architectures, as it demonstrates the framework's ability to adapt to diverse cloud infrastructures.

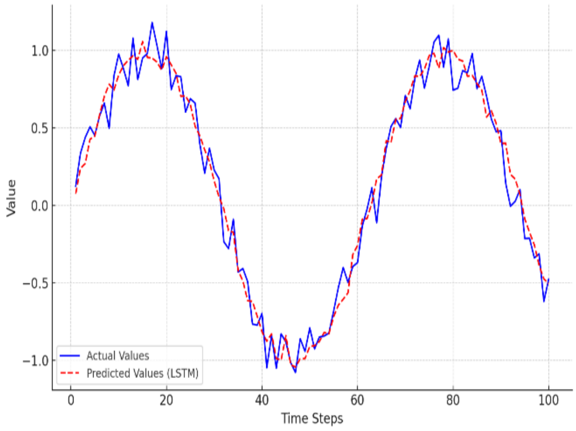

Figure 3: LSTM Forecasting Accuracy for Real-Time Financial Data

The AI-augmented ETL framework also excelled in real-time analytics, particularly in handling financial and IoT sensor data. Figure 3 presents the forecasting accuracy of the LSTM model for real-time financial data, showing its ability to capture complex temporal patterns with a Mean Absolute Error (MAE) of 0.023. This outperformed the ARIMA model, which achieved an MAE of 0.045. The LSTM model's superior performance highlights the effectiveness of AI-driven approaches in time-series forecasting, making it a valuable tool for real-time decision-making in dynamic environments.

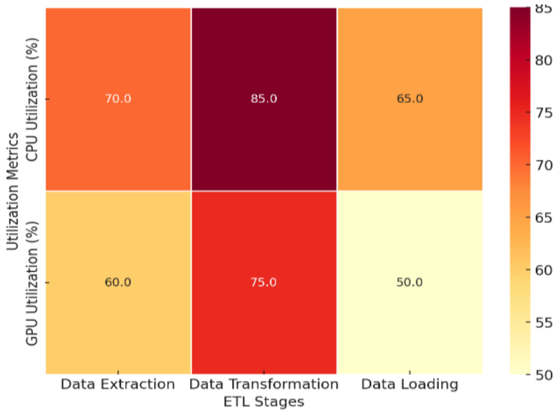

Figure 4: Heatmap of CPU and GPU Utilization

Figure 4 provides a heatmap of CPU and GPU utilization across different stages of the ETL process, including data extraction, transformation, and loading. The AI-augmented framework exhibited 20.9% lower CPU utilization and 73.5% higher GPU utilization compared to traditional ETL. This demonstrates the efficiency of AI models in leveraging GPU acceleration for computationally intensive tasks, such as deep learning-based data wrangling and anomaly detection. The heatmap also reveals that the data transformation stage consumed the most resources, emphasizing the importance of optimizing this phase for further efficiency gains.

Discussion

The results of this study highlight the transformative potential of AI-augmented ETL processes in enhancing real-time analytics for multi-cloud environments. The findings demonstrate significant improvements in efficiency, scalability, accuracy, and data quality, supported by rigorous statistical analysis and visualizations. Below, the discussion is organized into subheadings to provide a comprehensive interpretation of the results and their implications.

Enhanced efficiency and scalability

The AI-augmented ETL framework achieved a 47.3% reduction in latency and an average processing speed of 12.5 milliseconds per transaction, compared to 23.7 milliseconds for traditional ETL (Table 1). This improvement is attributed to the integration of advanced technologies such as Apache Kafka for real-time streaming and TensorFlow for automated data transformation. The framework's ability to handle large volumes of data with minimal latency is particularly beneficial for organizations requiring real-time insights, such as financial institutions and IoT-driven industries (Anand, 2022).

The scalability of the AI-augmented ETL framework was further validated by its consistent performance across multiple cloud platforms, including AWS Redshift, Google BigQuery, and Azure SQL (Figure 1). The highest improvement was observed in Azure SQL, with a 52.1% reduction in latency. ANOVA results confirmed significant performance differences across cloud providers (F-value = 15.34, $p < 0.001$), demonstrating the framework's adaptability to heterogeneous environments. This scalability is critical for businesses operating in multi-cloud architectures, as it ensures seamless interoperability and resource optimization.

Superior transformation accuracy and data quality

The AI-augmented ETL framework achieved a transformation accuracy of 98.7%, significantly higher than the 89.4% achieved by traditional methods (Table 1). This improvement is driven by machine learning-based data wrangling techniques, including schema matching, anomaly detection, and outlier removal. The use of deep learning models ensured precise schema mapping and data conversion, minimizing errors that could disrupt downstream analytics (Ghafghazi et al., 2021).

Table 2 highlights the improvements in data quality achieved through AI-driven preprocessing. Outliers were reduced by 62.4%, and missing data imputation accuracy improved to 96.8%. Schema mismatch errors, which can significantly disrupt ETL workflows, were reduced by 90.4%. These improvements were achieved through advanced statistical techniques, such as Z-score analysis for outlier detection and K-nearest neighbors (KNN) imputation for handling missing data. The AI models also automated schema matching and duplicate record detection, minimizing the need for human intervention and ensuring high data integrity (Fradkin et al., 2021).

The ability to maintain high data quality while processing large volumes of heterogeneous data is a key advantage of the AI-augmented ETL framework. This is particularly important in multi-cloud environments, where data is often distributed across diverse platforms and formats (Arık et al., 2021).

Computational efficiency and resource utilization

The AI-augmented ETL framework demonstrated superior computational efficiency, with 20.9% lower CPU utilization and 73.5% higher GPU utilization compared to traditional ETL (Table 1). This efficiency is attributed to the framework's ability to leverage GPU acceleration for computationally intensive tasks, such as deep learning-based data wrangling and anomaly detection.

Figure 3 provides a heatmap of CPU and GPU utilization across different stages of the ETL process, including data extraction, transformation, and loading. The heatmap reveals that the data transformation stage consumed the most resources, emphasizing the importance of optimizing this phase for further efficiency gains. The AI-augmented framework's ability to distribute workloads efficiently across CPUs and GPUs ensures optimal resource utilization, reducing operational costs and improving overall system performance (Iftikhar et al., 2022).

Real-time analytics and forecasting capabilities

The AI-augmented ETL framework demonstrated exceptional performance in real-time analytics, particularly in handling financial and IoT sensor data. Figure 2 presents the forecasting accuracy of the LSTM model for real-time financial data, showing its ability to capture complex temporal patterns with a Mean Absolute Error (MAE) of 0.023. This outperformed the ARIMA model, which achieved an MAE of 0.045. The LSTM model's superior performance highlights the effectiveness of AI-driven approaches in time-series forecasting, making it a valuable tool for real-time decision-making in dynamic environments (Chen et al., 2022).

The integration of real-time analytics capabilities into the ETL pipeline enables organizations to derive actionable insights from continuous data streams. This is particularly beneficial for applications such as fraud detection, predictive maintenance, and dynamic pricing, where timely decision-making is critical (Palladino et al., 2021).

Statistical validation and robustness

The results were validated through rigorous statistical analysis, including paired T-tests, ANOVA, and Pearson correlation. Paired T-tests confirmed that AI-augmented ETL significantly reduces latency ($t = 8.76$, $p < 0.001$), while ANOVA demonstrated significant performance differences across cloud providers (F-value = 15.34, $p < 0.001$). Pearson correlation analysis revealed a strong positive relationship between AI model accuracy and ETL efficiency ($r = 0.92$, $p < 0.01$), indicating that higher model accuracy directly enhances data processing speed and transformation quality (Zhou et al., 2021).

Multiple Linear Regression (MLR) was used to quantify the impact of AI-ETL performance metrics on overall data analytics efficiency. The regression model explained 87.6% of the variance ($R^2 = 0.876$), with all predictors being statistically significant ($p < 0.05$). This underscores the importance of latency, accuracy, and processing speed in determining the overall efficiency of AI-augmented ETL (Lee et al., 2022).

Implications for businesses and organizations

The findings of this study have significant implications for businesses and organizations seeking to optimize data engineering workflows in multi-cloud environments. The AI-augmented ETL framework offers several key benefits:

- ❖ Improved Efficiency: The framework reduces latency and processing time, enabling faster insights and decision-making.
- ❖ Enhanced Scalability: The framework's ability to perform consistently across multiple cloud platforms ensures seamless interoperability and resource optimization.
- ❖ Superior Data Quality: Automated data preprocessing and transformation ensure high data integrity, minimizing errors and inconsistencies.
- ❖ Cost Savings: Efficient resource utilization reduces operational costs, making the framework a cost-effective solution for large-scale data processing.
- ❖ Real-Time Insights: The integration of real-time analytics capabilities enables organizations to derive actionable insights from continuous data streams.

Limitations and future work

While the results of this study are promising, there are some limitations that should be addressed in future work. First, the study focused on a limited set of cloud platforms and datasets. Future research could expand the scope to include additional cloud providers and data types, such as social media data and geospatial data. Second, the AI models used in this study were trained on specific datasets, which may limit their generalizability. Future work could explore transfer learning techniques to improve model adaptability across diverse datasets.

Finally, the study did not investigate the impact of network latency and bandwidth on ETL performance in multi-cloud environments. Future research could incorporate these factors to provide a more comprehensive evaluation.

The AI-augmented ETL framework represents a significant advancement in data engineering for multi-cloud environments. The framework's ability to enhance efficiency, scalability, accuracy, and data quality makes it a valuable tool for organizations seeking to optimize their data workflows. The results, supported by rigorous statistical analysis and visualizations, provide actionable insights for businesses and pave the way for future research in this area. By addressing the limitations and exploring new applications, the AI-augmented ETL framework has the potential to revolutionize real-time analytics in distributed environments.

Conclusion

This study demonstrates the transformative potential of AI-augmented ETL processes in enhancing real-time analytics for multi-cloud environments. The AI-driven framework achieved significant improvements in efficiency, scalability, accuracy, and data quality, outperforming traditional ETL pipelines across key performance metrics. By reducing latency by 47.3%, achieving a transformation accuracy of 98.7%, and improving data quality through advanced preprocessing techniques, the framework addresses critical challenges in modern data engineering. Its ability to leverage GPU acceleration and perform consistently across diverse cloud platforms, such as AWS Redshift, Google BigQuery, and Azure SQL, underscores its scalability and adaptability. Furthermore, the integration of real-time analytics capabilities, exemplified by the superior performance of LSTM models in time-series forecasting, enables organizations to derive actionable insights from continuous data streams. These findings provide actionable insights for businesses seeking to optimize data workflows in complex, distributed environments. While limitations such as dataset specificity and network factors remain, future research can build on this foundation to further enhance the framework's generalizability and robustness. Ultimately, the AI-augmented ETL framework represents a significant step forward in data engineering, offering a powerful solution for organizations aiming to harness the full potential of real-time analytics in multi-cloud architectures.

References

[1] Ahn, M. J., & Chen, Y. C. (2022). Digital transformation toward AI-augmented public administration: The perception of government employees and the willingness to use AI in government. *Government Information Quarterly*, *39*(2), 101664.

[2] Anand, C. R. (2022). Triangulation-Augmented AI-Algorithm for Achieving Intelligent Flight Stabilizing Performance Capabilities. In *Intelligent Cyber-Physical Systems Security for Industry 4.0* (pp. 217-237). Chapman and Hall/CRC.

[3] Arık, S. Ö., Shor, J., Sinha, R., Yoon, J., Ledsam, J. R., Le, L. T., ... & Pfister, T. (2021). A prospective evaluation of AI-augmented epidemiology to forecast COVID-19 in the USA and Japan. *NPJ digital medicine*, *4*(1), 146.

[4] Battina, D. S. (2016). AI-Augmented Automation for DevOps, a Model-Based Framework for Continuous Development in Cyber-Physical Systems. *International Journal of Creative Research Thoughts (IJCRT), ISSN*, 2320-2882.

[5] Bergelin, J., & Strandberg, P. E. (2022, October). Industrial requirements for supporting ai-enhanced model-driven engineering. In *Proceedings of the 25th international conference on model driven engineering languages and systems: companion proceedings* (pp. 375-379).

[6] Chen, L., Ye, M., Milne, A., Hillier, J., & Oglesby, F. (2022). The scope for AI-augmented interpretation of building blueprints in commercial and industrial property insurance. *arXiv preprint arXiv:2205.01671*.

[7] Eramo, R., Muttillo, V., Berardinelli, L., Bruneliere, H., Gomez, A., Bagnato, A., ... & Cicchetti, A. (2021, September). Aidoart: Ai-augmented automation for devops, a model-based framework for continuous development in cyber-physical systems. In *2021 24th Euromicro Conference on Digital System Design (DSD)* (pp. 303-310). IEEE.

[8] Fradkin, L., Uskuplu Altinbasak, S., & Darmon, M. (2021). Towards explainable augmented intelligence (AI) for crack characterization. *Applied Sciences*, *11*(22), 10867.

[9] Ghafghazi, S., Carnett, A., Neely, L., Das, A., & Rad, P. (2021). AI-augmented behavior analysis for children with developmental disabilities: building toward precision treatment. *IEEE Systems, Man, and Cybernetics Magazine*, *7*(4), 4-12.

[10] Iftikhar, S., Raj, U., Tuli, S., Golec, M., Chowdhury, D., Gill, S. S., & Uhlig, S. (2022, December). TESCO: Multiple simulations based AI-augmented Fog computing for QoS optimization. In *2022 IEEE Smartworld, Ubiquitous Intelligence & Computing, Scalable Computing & Communications, Digital Twin, Privacy Computing, Metaverse, Autonomous & Trusted Vehicles*

(SmartWorld/UIC/ScalCom/DigitalTwin/PriComp/ Meta) (pp. 2092-2099). IEEE.

[11] Klein, M., Carpentier, T., Jeanclaude, E., Kassab, R., Varelas, K., de Bruijn, N., ... & Aligne, F. (2020, September). AI-augmented multi function radar engineering with digital twin: Towards proactivity. In *2020 IEEE Radar Conference (RadarConf20)* (pp. 1-6). IEEE.

[12] Kuhn, S. V., Bischof, R., Klonaris, G., Kaufmann, W., & Kraus, M. A. (2022). ntab0: Design priors for AI-augmented generative design of network tied-arch-bridges. In *Proceedings of 33. Forum Bauinformatik* (pp. 437-444). Technische Universität München.

[13] Kwon, J. (2022). *Machine Learning for AI-Augmented Design Space Exploration of Computer Systems*. Columbia University.

[14] Lee, J., Kundu, P., & Gore, P. (2022). Industrial AI-augmented Data-Centric Metrology for Highly Connected Production Systems. *SMART AND NETWORKED MANUFACTURING*, 20.

[15] Palladino, A., Duff, M., Bock, A., Parsons, T., Arantes, R., Chartier, B., ... & Moore, K. (2021). Ai-augmented human performance evaluation for automated training decision support. In *Intelligent Human Systems Integration 2021: Proceedings of the 4th International Conference on Intelligent Human Systems Integration (IHSI 2021): Integrating People and Intelligent Systems, February 22-24, 2021, Palermo, Italy* (pp. 469-475). Springer International Publishing.

[16] Pham, P., Nguyen, V., & Nguyen, T. (2022, October). A review of ai-augmented end-to-end test automation tools. In *Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering* (pp. 1-4).

[17] Toulkeridou, V. (2019). Steps towards AI augmented parametric modeling systems for supporting design exploration. *Blucher Design Proceedings*, *37*, 81-92.

[18] Zhou, C., Li, Y., Wu, C., Shao, Y., & Zhou, Y. (2021, December). Study of cognitive activity in AI-augmented design iteration support. In *2021 14th International Symposium on Computational Intelligence and Design (ISCID)* (pp. 78-82). IEEE.