**International Academy of Science,
Engineering and Technology**
Connecting Researchers; Nurturing Innovations
IASET

LEVERAGING AZURE DATA FACTORY FOR LARGE-SCALE ETL IN HEALTHCARE AND INSURANCE INDUSTRIES

Afroz Shaik[1], Shyamakrishna Siddharth Chamarthy[2], Krishna Kishor Tirupati[3], Prof. (Dr) Sandeep Kumar[4], Prof. (Dr) MSR Prasad[5] & Prof. (Dr) Sangeet Vashishtha[6]

[1]Cleveland State University, Cleveland OH, USA

[2]Scholar, Columbia University, Sakthinagar 2nd Ave, Nolambur, Chennai, India

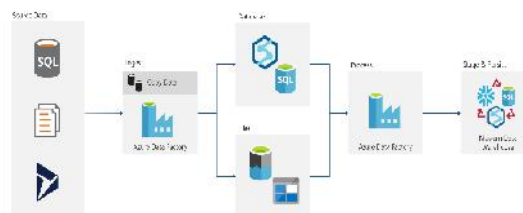*[3]International Institute of Information Technology Bangalore, India*

[4]Department of Computer Science and Engineering Koneru Lakshmaiah Education Foundation Vadeshawaram, A.P., India

[5]Department of Computer Science and Engineering Koneru Lakshmaiah Education Foundation Vadeshawaram, A.P., India

[6]IIMT University, Meerut, India

ABSTRACT

This paper explores the utilization of Azure Data Factory (ADF) for large-scale Extract, Transform, and Load (ETL) processes in healthcare and insurance industries. As data volume and variety increase in these sectors, there is a growing need for scalable, secure, and efficient data integration solutions. Azure Data Factory, a cloud-based data integration service, offers a robust platform for orchestrating complex ETL workflows while ensuring data accuracy and compliance with industry regulations. This study highlights how ADF streamlines data ingestion from disparate sources, enabling real-time data analytics, operational insights, and enhanced decision-making. Special focus is given to healthcare and insurance use cases such as claims processing, patient record integration, and fraud detection. The paper also discusses the advantages of using ADF, including scalability, seamless integration with other Azure services, and advanced monitoring capabilities. Moreover, it addresses the challenges of data privacy and compliance, outlining how ADF supports encryption and governance frameworks like HIPAA and GDPR. The findings suggest that ADF not only improves data management but also facilitates operational efficiency and predictive analytics, empowering organizations to meet the evolving demands of the healthcare and insurance industries.