SCALABLE DATA INTEGRATION TECHNIQUES FOR MULTI-RETAILER E-COMMERCE PLATFORMS

Varun Garg[1] & Aayush Jain[2]

[1]*University of California[1], 9500 Gilman Dr, La Jolla, CA 92093, United States*

[2]*Vivekananda Institute of Professional Studies[2] -Pitampura, Delhi, India*

ABSTRACT

In the rapidly evolving landscape of e-commerce, multi-retailer platforms face the challenge of integrating diverse data sources to provide seamless and consistent shopping experiences. Scalable data integration techniques are essential for harmonizing product catalogs, standardizing metadata, and ensuring real-time updates across heterogeneous retailer systems. This study explores state-of-the-art methodologies and frameworks that address the complexities of integrating large volumes of structured and unstructured data from varied sources. Key techniques such as schema matching, ontology mapping, and machine learning-driven entity resolution are examined for their efficacy in creating unified data models. Additionally, the paper highlights the role of cloud-based infrastructure and distributed computing paradigms, such as Apache Spark and Kubernetes, in achieving scalability and fault tolerance. By leveraging advanced data integration strategies, multi-retailer e-commerce platforms can enhance catalog accuracy, improve operational efficiency, and deliver personalized customer experiences. The findings of this study provide actionable insights for developers and decision-makers aiming to implement scalable, efficient, and resilient data integration solutions.

KEYWORD: *Scalable Data Integration, Multi-Retailer E-Commerce, Product Catalog Harmonization, Schema Matching, Ontology Mapping, Entity Resolution, Cloud-Based Infrastructure, Distributed Computing, Real-Time Updates, Operational Efficiency*