



Cyprus
University of
Technology

Department of Electrical
Engineering and Computer
Engineering and Informatics

Bachelor Thesis

Comparative Evaluation of Storage and Retrieval Performance in Modern Big Data Architectures: Data Lake, Data Lakehouse, Knowledge Graph Lakehouse, Data Mesh, Data Fabric and Semantic Data Fabric

Christos Aristotelous

Limassol, May 2026

CYPRUS UNIVERSITY OF TECHNOLOGY

Faculty of Engineering and Technology

Department of Electrical Engineering, Computer Engineering, and Informatics

Bachelor Thesis

**Comparative Evaluation of Storage and Retrieval
Performance in Modern Big Data Architectures:
Data Lake, Data Lakehouse, Knowledge Graph
Lakehouse, Data Mesh, Data Fabric and Semantic
Data Fabric**

Christos Aristotelous

Advisor: Dr. Andreas S. Andreou

Limassol, May 2026

Copyrights

Copyright © 2026 Christos Aristotelous

All rights reserved.

The approval of the dissertation by the Department of Electrical Engineering, Computer Engineering, and Informatics does not necessarily imply the approval by the Department of the views of the writer.

Acknowledgements

I would like to express my sincere appreciation to everyone who supported and encouraged me throughout the course of this thesis. I am especially grateful to my professors, Dr. Andreas Andreou and Dr. Michalis Pingos, for their guidance, valuable feedback, and support during the development of this work. Their advice helped me improve the direction, structure, and quality of this thesis.

I would also like to thank them for giving me the opportunity to work on a topic that allowed me to expand my knowledge in Big Data architectures, fuzzy logic, and data-driven systems. Their support played an important role in the completion of this research.

ABSTRACT

This thesis presents a comparative evaluation of modern Big Data architectures, focusing on Data Lake, Data Lakehouse, Data Mesh, Data Fabric, Semantic Data Fabric, and Knowledge Graph Lakehouse. The aim is to examine how these architectures perform in important areas such as storage and retrieval, data organization, metadata management, governance, semantic support, processing patterns, performance, and scalability.

To achieve this, a fuzzy logic-based methodology is applied. Each architecture is evaluated using Low, Medium, and High satisfaction levels, supported by weighted criteria and trapezoidal membership functions. The final results are aggregated across all evaluation fields to produce an overall ranking.

The results show that Semantic Data Fabric achieved the strongest overall performance, followed by Data Mesh and Knowledge Graph Lakehouse. Data Lakehouse and Data Fabric also showed strong practical value, while Data Lake ranked lowest due to weaker support for governance, semantics, ownership, and systematic evolution.

Keywords: Big Data architectures, Data Lake, Data Lakehouse, Data Mesh, Data Fabric, Semantic Data Fabric, Knowledge Graph Lakehouse, fuzzy logic, storage performance, retrieval performance.