

TECHNOLOGY TRANSFER PRESENTS

# MIKE FERGUSON

## MODERN DATA ARCHITECTURE

Building a Foundation for the Data Driven Enterprise

ONLINE LIVE STREAMING

**DECEMBER 10-11, 2024**



info@technologytransfer.it  
www.technologytransfer.it

## ABOUT THIS SEMINAR

In the last five years, digital transformation has reshaped data management presenting new challenges and opportunities including:

- Data complexity caused by many more data sources with data now stored in SaaS applications, on multiple clouds, on-premises and streaming in from the edge
- Business units are buying data catalogs to help understand, govern and provision data
- Multiple new siloed analytical systems like streaming analytics, lakehouses and graph databases have appeared beyond the data warehouse offering alternative new data architectures
- Data modelling seems to have disappeared
- Data engineering is now happening everywhere and new technologies like Data Fabric and Modern Data Stacks have emerged offering way more than ETL

Simultaneously CEO's now see data and AI as strategic and needed in every part of the business. They are demanding a way forward to speed of development.

So how do you make sense of all this? Is there a future for the data warehouse? Is data modelling dead? With so many competing data architectures, which one is best? How do you meet all requirements and prevent chaos?

This new 2-day class looks at this problem.

### WHO SHOULD ATTEND

- How you assess your existing environment
- How to define future requirements and design a new modern Data Architecture that modernises your Data Warehouse, and makes it possible to merge it with multiple Analytical workloads like Data Science, Streaming Analytics, and Graph Analysis
- How the Architecture allows you to use a Data Catalog, and Data Fabric and Data Observability to build resilient DataOps pipelines to create a Data Mesh of reusable data products published in a Data Marketplace that help shorten time to value by enabling new insights and AI to be delivered more rapidly

### AUDIENCE

CDOs, CIO's, IT Managers, CTOs, Business Analysts, Data Scientists, BI Managers, Data Warehousing Professionals, Enterprise Architects, Data Architects, Solution Architects, Business Intelligence Specialists, IT Strategists, Database Administrators, IT Consultants

# OUTLINE

## 1. What is Data Architecture?

This starting session looks at what data architect is, why companies need one and what current issues are with respect to Data Architecture today.

- What is Data Architecture?
- Why do you need a Data Architecture?
- What are the differences between Data Architecture, Solution Architecture and Enterprise Architecture?
- What are the main capabilities of a Data Architecture?
- Reference Data Architectures and their pros and cons
- What is the difference between Batch, Lambda, Kappa, Zero-copy integration?
- Popular Data Architectures and their pros and cons e.g. Data Warehouse, Data Lake, streaming Analytics, transaction system
- What data storage, data processing and data analytics technologies are used in Data Architecture
- What are the pros and cons of Cloud computing in a Data Architecture?
- How does it all fit together across all environments?

## 2. Assessing your existing Data Architecture

This session looks at how to assess your existing Data Architecture to understand how well it is serving your business needs from both an operational and analytical perspective. This includes documenting the current problems and identifying stand-alone projects that currently bypass your Architecture. The module will cover assessing your existing:

- Data sources you connect to
- Data ingestion techniques
- Data stores in use and their purpose
- Data stores on the Cloud Vs on-premises
- Operational applications, processes and Data Flows
- Analytical data capabilities e.g. Data Warehouses, Data Lakes, feature stores
- Data integration Data Flows for analytics
- Tools used e.g. data catalog, data quality, data governance, ETL, self-service data preparation, data science

- Identifying and documenting issues with your Data Architecture e.g.,
  - Data complexity
  - Siloed analytical systems
  - Too many copies of data
  - Data that is too big to move
  - Brittle data engineering pipelines
  - Unmanaged self-service data engineering
- Managing data in a distributed and multi-Cloud hybrid computing environment
  - Missing data & analytical capabilities
  - Integration complexity
  - Rapid growth in APIs
- How well does existing Data Architecture meet your business needs?
  - E.g. Inability to respond in a timely way
  - Are people bypassing your Data Architecture and creating their own?

## 3. New Data Architecture and Technologies Options

Having understood Data Architecture capabilities, technologies and assessed your existing architecture, this module looks at multiple new Data Architectures. It also looks at how the emergence of open table formats, advances in SQL, universal Data Lake APIs, federated query, data catalogs, and data automation are changing Data Architecture to provide a data foundation for an AI-driven enterprise.

- Data Architecture options for operational processing
  - Data integration hub
  - Message bus
- Data Architecture options for analytical processing and their pros and cons
  - Zoned Data lake
  - Cloud Data Warehouse
  - Lakehouse
  - Cloud Data Warehouse and Data Lake integration
  - Logical Data Warehouse
  - Data Fabric and Data Mesh
  - Customer data platforms
- Data Architecture options for both Operational and Analytical Processing

- Databricks zero copy integration for data collaboration
- Centralised versus Federated architectures
- Next generation analytical Data Architecture
  - The impact of Open table formats
  - The OneTable initiative
  - Merging of Data Warehouse, Lake house and streaming on multi-purpose open tables
  - Separation of storage from compute
  - Multiple query and Data Engineering engines on shared data
  - Universal execution engine for query and Data Engineering workloads e.g. Velox
  - Universal administration of multiple engines
  - Advances in SQL to support graph Analytics
  - Federated query across open tables on multiple Clouds and on-premises storage
- Technology options and their pros and cons
  - Single vendor data and AI stack
  - Data Fabric including:
    - Data catalog
    - Collaborative data engineering
    - Data Automation – using a data catalog and generative AI to auto generate pipelines, virtual views, APIs and Architecture
      - DataOps and Data Observability
      - Data virtualization for on-demand data integration
  - Data platform
  - AI platform
  - End-to-end data and AI governance
  - Modern data stack
  - Best of breed tools

#### 4. Considerstions for Designing, Managing and Operating a Modern Data Architecture

This module focusses on what you need to think about with respect to designing a data architecture and how you operate and manage it:

- Ensuring a common approach to integrating transaction processing applications
- Introducing continuous change and change data capture in transaction systems to lower data latency

- Ensuring tools are integrated to share metadata and enable reuse of business terms, data models, data transformations, data quality rules and data governance policies
- Data Warehouse modernisation
  - Modernising ETL processing
  - Data Warehouse automation
  - Migration to the Cloud
  - Virtual Data Marts
- Where does data modelling fit in?
- Data engineering
  - Supporting citizen data engineers, IT professional data engineers and code in data engineering pipelines
  - Establishing a universal approach to data ingestion
  - Incorporating DataOps, CI/CD, data orchestration and data observability into data engineering pipelines
  - Dealing with sensitive data to govern data privacy
  - Establishing a data marketplace to share data

#### 5. Defining your future Data Architecture Requirements

This session looks at how you can use a multi-purpose Data Lake, Lakehouse and Data Mesh to accelerate ETL processing and integration of data for your Data Warehouse

- Ability to process any type of data
- Support hybrid multi-Cloud as standard
- Documented sharable metadata rules on cleansing, transformation, matching and integration with full metadata lineage
- Governed data ingestion, staging and transformation
- Data automation to accelerate development
- Reusable high-quality data products available as a service
- Shared data products to reduce data redundancy
- Data contracts to govern data sharing
- Consuming data products via APIs
- Integrated data governance
- Integrated analytics e.g. BI and ML, ML and streaming, graph analysis and BI, graph and ML, streaming and graph analysis

- Scalability to handle
- Data volumes, data velocity and an increasing number of concurrent users using tools with natural language Generative AI user interfaces
- History of changes to transaction data and/or data products used in transaction systems

## 6. Designing an innovative Data Architecture

In this session we focus on how to design a modern Data Architecture and merge workloads to create a multi-purpose platform supporting multiple analytical workloads.

- Data Architecture design principles
- Defining the key data capabilities needed in your Data Architecture
- Alignment of data capabilities needed to meet business strategy objectives and priorities
- Designing a data concept model for your business
- Designing your Data Architecture and data flows to support:
  - Operational processing
  - Analytical processing
  - Knowledge Management
- Merging your Data Warehouse, Data Lake, Lakehouse, streaming and graph to create a new multi-purpose hybrid analytical data platform
  - Defining the core data concepts needed
  - Steps to migrating your Data Warehouse to the Cloud
  - Adoption of open tables formats
  - Using Data Fabric catalog metadata and generative AI to generate pipelines to build data products in open tables
  - Supporting streaming data in open tables
  - Consuming data products in your Cloud Data Warehouse
  - Building virtual Data Marts from data products
  - Building virtual feature stores for Data Science from data products
  - Supporting graph analysis in Data Warehouses

## 7. Getting Started

This final session looks at what you have to do to get started with a new Data Architecture initiative. In particular, it looks at:

- Creating an action plan to bring it to life
- Change Vs rebuild?
- What order do you do this in?
- How do you minimise impact on the business while you re-architect?
- How to you deal with a backlog of change when you are also trying to re-architect?
- Pros and cons of build Vs automating development
- What new skills are needed?
- Delivering new business value whilst re-architecting
- How do you involve business professionals in the re-architecting effort?

# INFORMATION

<p><b>PARTICIPATION FEE</b></p> <p>€ 1200</p> <p>The fee includes all seminar documentation.</p> <p><b>SEMINAR TIMETABLE</b></p> <p>9.30 am - 1.00 pm 2.00 pm - 5.00 pm</p>	<p><b>HOW TO REGISTER</b></p> <p>You must send the registration form with the receipt of the payment to: <b>TECHNOLOGY TRANSFER S.r.l.</b> Piazza Cavour, 3 - 00193 Rome (Italy) Fax +39-06-6871102</p> <p><b>PAYMENT</b></p> <p>Wire transfer to: Technology Transfer S.r.l. Banca: Credit Agricole Agenzia 1 di Roma IBAN Code: IT 03 W 06230 03202 000057031348 BIC/SWIFT: CRPPIT2P546</p>	<p><b>GENERAL CONDITIONS</b></p> <p><b>DISCOUNT</b></p> <p>The participants who will register 30 days before the seminar are entitled to a 5% discount.</p> <p>If a company registers 5 participants to the same seminar, it will pay only for 4. Those who benefit of this discount are not entitled to other discounts for the same seminar.</p> <p><b>CANCELLATION POLICY</b></p> <p>A full refund is given for any cancellation received more than 15 days before the seminar starts. Cancellations less than 15 days prior the event are liable for 50% of the fee. Cancellations less than one week prior to the event date will be liable for the full fee.</p> <p><b>CANCELLATION LIABILITY</b></p> <p>In the case of cancellation of an event for any reason, Technology Transfer's liability is limited to the return of the registration fee only.</p>
---	---	---

MIKE FERGUSON

**MODERN DATA ARCHITECTURE**

December 11-12, 2024

Registration fee: € 1200

first name .....

surname .....

job title .....

organisation .....

address .....

postcode .....

city .....

country .....

telephone .....

fax .....

e-mail .....



Stamp and signature

Send your registration form with the receipt of the payment to:  
**Technology Transfer S.r.l.**  
Piazza Cavour, 3 - 00193 Rome (Italy)  
Tel. +39-06-6832227 - Fax +39-06-6871102  
info@technologytransfer.it  
www.technologytransfer.it

If anyone registered is unable to attend, or in case of cancellation of the seminar, the general conditions mentioned before are applicable.

## ***SPEAKER***

**Mike Ferguson** is Managing Director of Intelligent Business Strategies Limited. As an analyst and consultant he specialises in business intelligence and enterprise business integration. With over 41 years of IT experience, Mike has consulted for dozens of companies. He has spoken at events all over the world and written numerous articles. Mike is Chairman of Big Data LDN – the fastest growing Big Data conference in Europe. Formerly he was a principal and co-founder of Codd and Date Europe Limited – the inventors of the Relational Model, a Chief Architect at Teradata on the Teradata DBMS and European Managing Director of Database Associates. He teaches popular master classes in Analytics, Big Data, Data Governance a Distributed Data Landscape, Data Warehouse Modernisation and Creating Data Products in a Data Mesh, Data Lake or Lakehouse for Use in Analytics.