# Data Products

## From Design, to Build, to Publishing and Consumption

**Half day workshop
by Mike Ferguson**

///AdeptEvents

- Strengths and weaknesses of centralised data architectures used in analytics
- The problems caused in existing analytical systems by a hybrid, multi-cloud data landscape
- The benefits that data products offer and what are the implementation options?
- The critical importance of a data catalog in understanding what data is available
- What software is required to build, operate and govern a data mesh of data products?
- What is data fabric software, how does it integrate with data catalogs?
- An Implementation methodology to produce ready-made, trusted, reusable data products
- Managing data quality, privacy, access security, versioning, and the lifecycle of data products

**LANGUAGE**
English

**VENUE**
Utrecht / Hilversum

**TIME**
9:00 – 12:30 hours

**REGISTRATION**
www.adeptevents.nl

# Data Products

## From Design, to Build, to Publishing and Consumption

Most companies today are storing data and running applications in a hybrid multi-cloud environment. Analytical systems tend to be centralised and siloed like data warehouses and data marts for BI, cloud storage data lakes for data science and stand-alone streaming analytical systems for real-time analysis. These centralised systems rely on data engineers and data scientists working within each silo to ingest data from many different sources and engineer it for use in a specific analytical system or machine learning models. There are many issues with this centralised, siloed approach including multiple tools to prepare and integrate data, reinvention of data integration pipelines in each silo and centralised data engineering with poor understanding of source data unable to keep pace with business demands for new data.

To address these issues, a new approach called Data Mesh emerged in late 2019 attempting to accelerate creation of data for use in multiple analytical workloads. Data Mesh is a decentralised business domain-oriented approach to data ownership and data engineering to create a mesh of reusable data products that can be created once and shared across multiple analytical systems and workloads.
This half-day workshop looks at the development of data products in detail and also, how can you use a data marketplace to share and govern the sharing of data products across the enterprise to shorten time to value.

**Learning objectives:**
• Strengths and weaknesses of centralised data architectures used in analytics

- The problems caused in existing analytical systems by a hybrid, multi-cloud data landscape
- The emergence of data mesh and data products
- What exactly a data product is and the types of data products that you can create
- The benefits that data products offer and what are the implementation options?
- How to organise to create data products in a decentralised environment so you avoid chaos?
- How business glossaries can help ensure data products are formally defined, understood by business users and semantically linked
- The critical importance of a data catalog in understanding what data is available
- What software is required to build, operate and govern a data mesh of data products for use in a data lake, a data lakehouse or data warehouse?
- What is data fabric software, how does it integrate with data catalogs and connect to data in your data estate
- An Implementation methodology to produce ready-made, trusted, reusable data products
- Collaborative domain-oriented development of modular and distributed DataOps pipelines to create data products

- How a data catalog and automation software can be used to generate DataOps pipelines
- Managing data quality, privacy, access security, versioning, and the lifecycle of data products
- Publishing semantically linked data products in a data marketplace for others to consume and use
- Governing the sharing and use of data products in a data marketplace
- Consuming data products in an MDM system
- Consuming and assembling data products in multiple analytical systems like data warehouses, lakehouses and graph databases to shorten time to value.

### Who is it for?

This seminar is intended for business data analysts, data architects, chief data officers, master data management professionals, data scientists, IT ETL developers, and data governance professionals. It assumes you understand basic data management principles and data architecture plus a reasonable understanding of data cleansing, data integration, data catalogs, data lakes and data governance.

## MIKE FERGUSON

**MIKE FERGUSON** is Managing Director of Intelligent Business Strategies Limited. As an analyst and consultant he specialises in business intelligence / analytics, data management, big data and enterprise business integration. With over 40 years of IT experience, Mike has consulted for dozens of companies on business intelligence strategy, technology selection, enterprise architecture, and data management. He has spoken at events all over the world and written numerous articles.

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. Mike teaches in Data Warehouse Modernisation, Big Data Architecture & Technology, Centralised Data Governance of a Distributed Data Landscape, Practical Guidelines for Implementing a Data Mesh, Embedded Analytics, Intelligent Apps & AI Automation, Migrating Data Warehouse to the Cloud, Modern Data Architecture and Data Virtualisation.

Most companies today are storing data and running applications in a hybrid multi-cloud environment. Analytical systems tend to be centralised and siloed like data warehouses and data marts for BI, cloud storage data lakes or Hadoop for data science and stand-alone streaming analytical systems for real-time analysis. These centralised systems rely on data engineers and data scientists working within each silo to ingest data from many different sources, clean and integrate it for use in a specific analytical system or machine learning models. There are many issues with this centralised, siloed approach including multiple tools to prepare and integrate data, reinvention of data integration pipelines in each silo and centralised data engineering with poor understanding of source data unable to keep pace with business demands for new data. Also, master data is not well managed.

To address these issues, a new approach emerged in late 2019 attempting to accelerate creation of data for use in multiple analytical workloads. That approach is Data Mesh. Data Mesh is a decentralised business domain-oriented approach to data ownership and data engineering to create a mesh of reusable data products that can be created once and shared across multiple analytical systems and workloads. A Data Mesh can be implemented in a number of ways. These include using one or more cloud storage accounts on cloud storage, on an organised data lake, on a Lakehouse, on a data cloud, using Kafka or using data virtualisation. Data products can then be consumed in other pipelines for use in streaming analytics, Data Warehouses or Lakehouse Gold Tables, for use

in business intelligence, feature stores for use data science, graph databases for use in graph analysis and other analytical workloads.

This half-day workshop looks at the development of data products in detail. It also looks at the strengths and weaknesses of data mesh implementation options for data product development. Which architecture is best to implement this? How do you co-ordinate multiple domain-oriented teams and use common data infrastructure software like Data Fabric to create high-quality, compliant, reusable, data products in a Data Mesh. Is there a methodology for creating data products? Also, how can you use a data marketplace to share and govern the sharing of data products? The objective is to shorten time to value while also ensuring that data is correctly governed and engineered in a decentralised environment. It also looks at the organisational implications of Data Mesh and how to create sharable data products for use as master data, in a data warehouse, in data science, in graph analysis and in real-time streaming analytics to drive business value? Technologies discussed includes data catalogs, data fabric for collaborative development of data integration pipelines to create data products, DataOps to speed up the process, data orchestration automation, data observability and data marketplaces.

- What are data products?
- What makes creating data products different from other approaches to creating data for use analytical workloads?
- A best practice methodology for creating data products
- How to design semantically linked data products to enable rapid consumption and use of data to produce new insights
- Quick start mechanisms to speed up data product design
- Defining common business data names for data products in a business glossary
- Data modelling techniques for data products
- Discovering data needed to build data products using a data catalog
- Developing DataOps pipelines to engineer the data needed using data fabric
- Publishing data products – the role of the data marketplace
- Governing access to and use of data products across the enterprise
- Consuming and assembling data products for use in multiple analytical workloads
- Technologies and skills needed.

## DATE AND TIME

The workshop will take place on March 28, 2024 from 9 am to 12:30. We may continue to run this course once or twice a year with the exact date and time available on our website.

In case of virtual half day sessions, please visit our website for the exact schedule and remember to log in well in advance to check your video and audio settings.

## VENUE

Adept Events works with several venues in and near Amersfoort and Utrecht. We strive to provide you with the location details as soon as possible. The exact venue will be on our website and in the confirmation e-mail that you will receive one week prior to the event. Always check our website prior to your departure to ensure you have the exact location and directions.

## HOW TO REGISTER

Please register online at **www.adeptevents.nl**. For registering by print, please scan the completed registration form and send this or your Purchase Order to **customerservice@adeptevents.nl**. We will confirm your registration and invoice your company by e-mail therefore please do not omit your e-mail address when registering.

## REGISTRATION FEE

Taking part in this half-day workshop will only cost € 351,- when registering 30 days beforehand and € 390,- per person afterwards. This also covers documentation and lunch.

If combined with attending our DW&BI Summit on March 27th, 2024, the fee will be € 333,- when registering 30 days beforehand and € 370,- per person afterwards.

In case of virtual delivery, the registration fee is € 297,- when registering 30 days beforehand and € 330,- per person afterwards (excl. 21% Dutch VAT).

In completing your registration form you declare that you agree with our **Terms and Conditions**.

## Team discounts

Discounts are available for group bookings of two or more delegates representing the same organization made at the same time. Ten percent off for the second and third delegate and fifteen percent off for all delegates when registering four or more delegates (all delegates must be listed on the same invoice).

This cannot be used in conjunction with other discounts.

All prices are VAT excluded.

## PAYMENT

Full payment is due prior to the workshop. An invoice will be sent to you containing our full bank details including BIC and IBAN. Your payment should always include the invoice number as well as the name of your company and the delegate name.

**Payment by credit card** is also available. Please mention this in the Comment-field upon registration and find further instructions for credit card payment on our **customer service page**.

## Cancellation Policy

Cancellations must be received in writing at least three weeks before the commencement of the workshop and will be subject to a € 75,- administration fee. It is regretted that cancellations received within three weeks of the workshop date will be liable for the full workshop fee. Substitutions can be made at any time and at no extra charge.

## Cancellation Liability

In the unlikely event of cancellation of the workshop for any reason, Adept Events' liability is limited to the return of the registration fee only. Adept Events will not reimburse delegates for any travel or hotel cancellation fees or penalties. It may be necessary, for reasons beyond the control of Adept Events, to change the content, timings, speakers, date and venue of the workshop.

## MORE INFORMATION

📞 +31(0)172 742680

🌐 **https://www.adeptevents.nl/mdp-en**

@ **seminars@adeptevents.nl**

🐦 @AdeptEventsNL / **https://twitter.com/AdeptEventsNL**

in **https://www.linkedin.com/company/adept-events**

f **https://www.facebook.com/AdeptEventsNL**

Visit our Business Intelligence and Data Warehousing website **www.biplatform.nl** and download the App

Visit our website on Software Engineering, **www.release.nl** and download the App

## IN-HOUSE TRAINING

Would you like to run this course in-company for a group of persons? We can provide a quote for running an in-house course, if you offer the following details. Estimated number of delegates, location (town, country), number of days required (if different from the public course) and the preferred date/period (month). Please find more info on the **In-house page on our website**.