

Case Study



Microsoft Fabric Delivered in Six Weeks Powered by the Datahub Microsoft Fabric Accelerator

6 weeks

Discovery to production,
at a fixed price

Hourly

Refreshes, previously
a single overnight load

53m

Rows in the largest
fact table

1

Conformed customer
dimension

Industry

The organisation delivers manned guarding, building security, aviation security and facility services across hundreds of client sites, with offices and customers across the UK and Europe. The customer wanted reporting it could trust, close to near real-time as possible, and it wanted the platform live in weeks rather than months.

Objective

The organisation's operational and financial data sat in two systems: an ERP platform running across several separate legal entities, and a workforce management system holding scheduling, duty records and pay rates. Reporting came from an

ageing data warehouse that had grown organically over many years.

The objectives of this project were to:

- Replace the ageing warehouse with a modern, governed Microsoft Fabric platform built on dimensional modelling best practice.
- Give the internal data team a foundation they could extend themselves, rather than calling in consultants every time they needed a new table.
- Bring reporting data closer to real time using incremental loading instead of overnight full refreshes.
- Meet UK and EU data protection obligations and ISO 27001 control expectations, given the volume of employee data involved.

Challenge

The existing warehouse did not follow dimensional modelling best practice: no clear separation between raw, cleansed and reporting data, no conformed dimensions, no consistent handling of history. Business logic sat across views, stored procedures and reports, so the same measure could return different answers depending where it was run.

That left the business with four problems.

- **Trust.** When two reports disagreed, nobody could say which one was right. Analysts spent more time reconciling numbers than interpreting them.
- **Speed of change.** Adding a source table meant writing bespoke extract and load code. A request that should have taken a day took weeks, and the backlog grew.
- **Latency.** The warehouse reloaded every table overnight, so teams were working on data up to a day old, and the refresh window was getting longer every month.
- **Governance.** The platform held a lot of employee data, including scheduling, duty and pay information, with limited visibility of where it went, who could see it and how it was controlled once it left the source system.

All of this in six weeks, at a fixed price, without interrupting live reporting.

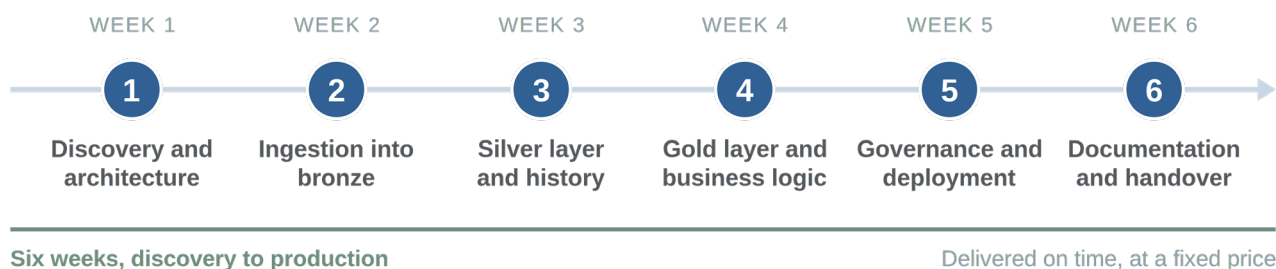
Solution

Datahub delivered the engagement using the Datahub Microsoft Fabric Accelerator. The accelerator is a framework, not a library of generic templates. It gives us proven engineering standards on day one for ingestion, logging, orchestration and deployment control, so no part of the six weeks goes on rebuilding plumbing. Everything above that is designed, built and coded for the organisation in front of us. The dimensional model, the transformation logic and the rules for reconciling records across both systems were written for this business and its data, because no two organisations have the same source systems, the same legal entities or the same questions to answer.

The engagement delivered the platform up to the gold layer. The semantic model and Power BI reports are being built on it by the client's own team.

Because the platform holds so much employee data, data protection and ISO 27001 controls were designed in during discovery rather than reviewed at the end, with our data protection lead auditor alongside the engineers.

How we planned the six weeks



Outcome

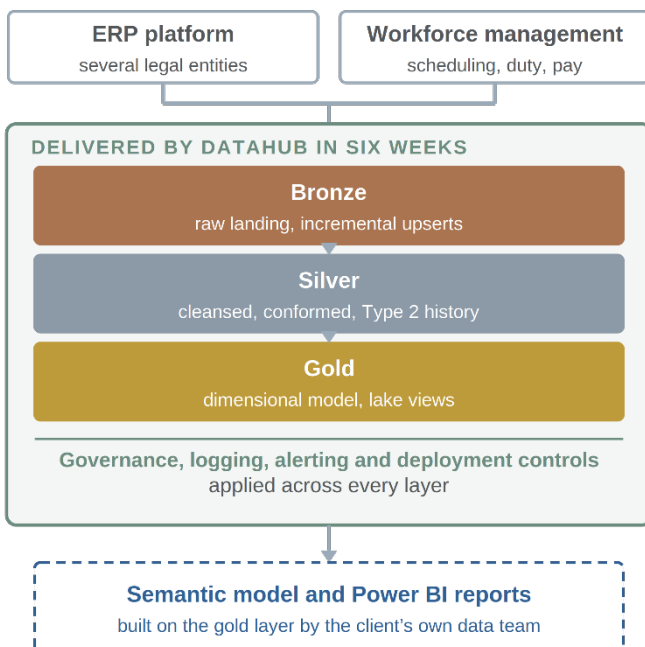
The engagement was delivered on time at a fixed price, and the organisation now has a governed medallion architecture in Microsoft Fabric:

- **Numbers the business trust.** Previously the system had been giving two versions of the same customer for years. There is now one customer dimension across both, with agreed rules for which source has the priority when records disagree. Also, Type 2 history so the business can report on what was true at the time as well as what is true today.
- **Reporting closer to the operational picture.** Incremental loading replaced the overnight reload, so data that refreshed once a day now refreshes every hour.
- **Failures visible in minutes.** Load status, durations and failure detail are logged centrally, with alerts to named staff.
- **Source data issues found and fixed.** Validation and quality checks surfaced several data issues, which we resolved inside the six weeks.

- **Regulatory alignment.** UK and EU data protection and ISO 27001 controls across the platform, with employee data in scope from the start.
- **A team able to extend the platform themselves.** New tables are added by configuration rather than a change request, with documentation and guides for their engineers.

Technology We Used

The platform in outline, from the two source systems through to the reporting layer the client is building on top of it.



OneLake
OneLake was the foundation of the solution, providing centralised storage using Delta Parquet tables across the bronze, silver and gold layers. That gave a consistent and optimised table structure for the reporting layer the business is now building.

Fabric Pipelines
Fabric Pipelines handled the copy activity into the bronze layer, using an upsert approach so loads could run incrementally. They also managed orchestration into the silver and gold layers, with error handling, monitoring and detailed logging of pipeline activity and load status.

PySpark Notebooks
PySpark Notebooks processed and transformed the data into the silver and gold layers. We also used them to profile data quality and flag anomalies, which gave the client far better visibility of the state of their source data.

Fabric Database
A Fabric Database held the control tables behind the metadata driven approach, along with load logs for bronze, silver and gold covering durations, load status and failure detail. New tables are added by configuration rather than code.

Materialized Lake Views
Materialized Lake Views replaced the complex stored procedures carried over from the old warehouse. The results are then materialized in Delta tables and can be loaded incrementally, so complex business logic is retrieved faster and at lower cost.

Microsoft Outlook
Any failure generates a notification to named staff with detail of the affected process, so problems are investigated within minutes rather than found the next morning.

Thinking about Microsoft Fabric Implementation?

The first conversation is not a pitch. Book a free 60-minute Fabric Readiness Review with the consultants who would do the work. Bring your current sources, the reporting you want and your timeline, and you will get an honest answer on whether six weeks is realistic for your estate, a sketch of the target architecture, and the governance questions worth settling before the build rather than after.

Book a Fabric Readiness Review: <https://www.datahubconsulting.co.uk/contact-us/>

Contact Information

Email: info@datahubconsulting.co.uk
Leicester Office: +44 (0)116 223 0689

London Office: +44 (0)20 3918 8799
European Office (Madrid): +34 911 11 58 11