Richard Swinbank

richard@RichardSwinbank.net

Innovative, business-focused Data Platform Engineer with proven, in-depth skills and experience.
Microsoft Data Platform MVP, author, client recommendations at https://tinyurl.com/yc823f8s

Key skills

- Azure data platform (including IaC, Terraform) & Microsoft SQL Server
- Azure Data Factory (ADF), SSIS (including BIML, BimlScript)
- Azure Databricks (SQL & PySpark, Structured Streaming), Azure Synapse Analytics, Microsoft Fabric
- SSAS tabular, AAS, Power BI, semantic modelling
- DevOps toolchain components (git/svn/TFS, tSQLt/NUnit/pytest, Azure Pipelines, PowerShell)
- Expert T-SQL for queries, analysis, administration and process engineering
- Configuration/metadata-driven dynamic T-SQL
- C# (including SQL CLR assemblies, SSIS script tasks/components, BimlScript), Java
- Requirements-driven data modelling and warehouse/lakehouse development
- Effective communicator, able to convey complex technical issues to a range of audiences. Blog https://richardswinbank.net, X/Twitter https://twitter.com/richardswinbank.
- Data-driven, analytical problem solver & engineering thought leader
- Active in Microsoft Data Platform community: Organiser, Volunteer, Speaker (Data Relay, SQLBits)

Achievements

- Author of *Azure Data Factory by Example* (https://www.amazon.co.uk/dp/B0CRH4NLW4)
- Microsoft Data Platform MVP
- Microsoft Certified Data Engineer Associate DP-203
- Built greenfield cloud-native, infrastructure-as-code analytics platform, enabling proptech startup to interrogate and understand performance of product features for B2B clients.
- Automated resilient value (ETL) and innovation (CI/CD) pipelines using git (Azure DevOps Repos, Bitbucket, GitHub), automated testing (tSQLt, NUnit, pytest) and Azure DevOps Pipelines.
- Established high-performing, agile teams of Data Engineers & BI Analysts, operating with high degree of trust, collaboration & psychological safety.
- Created metadata-driven extract and staging framework, supporting rapid, repeatable implementation and maintenance of platform staging objects.
- Built ETL process controller for data platform in Microsoft Azure using Azure Durable Functions (C#)
- Designed multi-tenant data warehouse platform, allowing client to process multiple customers' data in a single instance, reducing administration overhead and customer cost.
- Enabled end-user master data management using Master Data Services (MDS) integrated into ETL.
- Built configurable master data management framework in T-SQL, allowing multiple source systems to be integrated into operational data store's global data dictionary (in the absence of MDS).
- Re-engineered existing data warehouse ETL in situ, replacing legacy corporate system sources with new ones while allowing valuable reporting assets to remain in use.
- Author of Sprockit (http://sprockit.info) – a free, open-source ETL process controller built in pure T-SQL with adaptive end-to-end ETL duration reduction and one-step resume after error.
- Contributor to and cited by Microsoft Learn Azure platform documentation.

Richard Swinbank

richard@RichardSwinbank.net

Experience

August 2024 – present (one renewal)

Senior Data Engineer, Tech Lead – CDP

Technical leadership & engineering to realise benefits of new data platform

- Lead team of engineers stabilising and operationalising new Azure Synapse Analytics platform.
- Migrate data ingestion from data flows to PySpark with an anticipated cost saving of 40%.
- Build metadata-driven frameworks for data ingestion, data pipeline quality metrics and on-demand data product production.
- Elevate knowledge level in team of junior/mid-level data engineers.
- Identify and advocate for platform improvements for higher quality, faster benefits realisation and reduced operating cost.

December 2023 – August 2024 (one renewal)

Senior Azure Pipelines Engineer – TPXimpact

Data engineering project for UK energy plc

- Built and tuned near real time data pipelines using Spark Structured Streaming.
- Implemented objects in Delta lake-based data vault model using Databricks & ADF.
- Designed and built function library for repeatable notebook development patterns.
- Created code generation utilities to construct or configure pipeline objects from metadata.
- Developed cost estimates and other activity analyses using KQL in Azure Log Analytics.
- Provided technical leadership, enhancing engineering practices and coaching junior engineers.

May 2022 – November 2023 (permanent role)

Senior Data Platform Engineer – Avanade

Technical leadership for a range of projects for national and multinational clients

- Lift & shift migration of on-premises analytics platform into Azure, rebuilding, reusing and reorchestrating components for enhanced resilience and improved performance (Azure SQL, Azure Data Factory including SSIS-IR, Key Vault, Terraform).
- Reverse-engineered undocumented metadata-driven legacy platform, enabling reimplementation.
- Refactored existing data transformation processes, reducing execution times by 50-70% (T-SQL).
- Implemented cloud-native integration of heterogenous mixed cloud/on-premises source systems (Azure Data Factory including managed VNet IRs).
- Operated highly automated data lakehouse using metadata-driven ingestion and transformation processes, including lifecycle management for personal data (Databricks, Azure Data Factory, Power BI).
- Empowered and led mixed (Avanade/client) engineering teams in agile, self-organised delivery of data analytics platforms.
- Transformed business requirements into deliverable work items, facilitated technical refinement.
- Mentored and advised junior engineers in processes, patterns, and technologies.

Richard Swinbank

richard@RichardSwinbank.net

July 2020 – May 2022 (permanent role)

Senior Data Engineer, Head of Data – Boomin

Construction of greenfield cloud-native enterprise analytics platform in Microsoft Azure. Recruitment and leadership of data platform team.

- Built enterprise analytics platform comprising Azure Data Lake Gen2, Azure SQL DB, Azure Analysis Services & Power BI
- Automated platform construction and development using Terraform, Azure DevOps YAML pipelines and a range of deployment tools (SSDT, YAML templates, PowerShell scripts)
- Implemented metadata-driven process orchestration framework using ADF
- Built metadata-driven extract and staging framework for rapid staging process development
- Authored Azure Function Apps to encapsulate non-standard platform processes (e.g. Google OAuth token acquisition, Firebase user extracts)
- Automated and version-controlled platform configuration (including user access) using JSON metadata serialisation and custom-built Azure pipelines tasks
- Improved developer experience using PowerShell scripts and git hooks
- Recruited and led high-performing team of Data Engineers & BI Analysts, nurtured a culture of collaboration, trust & psychological safety
- Managed complex backlog of business requirements, writing user stories & guiding task refinement

September 2019 – February 2020 (one renewal)

Data Engineer – Oakbrook Finance

Implementation of next-generation enterprise data platform in Microsoft Azure.

- Designed and built ETL process controller using Azure Durable Functions (C#), to manage execution of Azure Data Factory (ADF) pipelines and other processes
- Designed and built metadata transformation library, enabling ETL metadata to be both stored under version control as JSON, and imported to (and re-exported from) Azure SQL DB
- Built proof-of-concept Azure DevOps YAML pipelines to create disposable development environments in Azure using Terraform, SSDT deployments and metadata deserialisation
- Automated inference of normalised staging schemas from JSON source data in Azure blob storage
- Implemented extensive test suites for SQL objects (TSQL, tSQLt) and ETL process controller (C#, NUnit, Moq, FluentAssertions)
- Automated functional tests for ADF pipelines using NUnit and Azure DevOps YAML pipelines

January 2018 – September 2019 (four renewals)

Data Warehouse Developer – Dyson

Stabilisation and productionisation of SQL Server 2016 data warehouse platform.

- Instituted robust, well-documented deployment pipeline and change control processes
- Replaced centralised TFS-based version control with git feature branch and one-step deploy
- Refactored SSIS & TSQL ETL components, reducing nightly processing time by 60%
- Extended Kimball data model to keep pace with fast-evolving reporting requirements
- Delegated management of integration controls to end users using Master Data Services
- Integrated data from heterogeneous source systems including Teamcenter, MS Project Server, Active Directory (LDAP queries), SharePoint (CSOM API), JIRA (REST API) & Exchange (EWS API)
- Built API extract interfaces using SSIS script source components and JSON shredding
- Streamlined DW integration by calling Tableau extracts directly from ETL via REST API

Richard Swinbank

richard@RichardSwinbank.net

November 2017 – January 2018

DW/BI Consultant – Counter Intelligence Retail

Reimplementation and automation of SQL Server 2016 sales reporting platform to release analyst and developer time, permit integration with corporate master data and facilitate report suite redevelopment.

- Abstracted complex legacy reporting model into simplified star schema
- Prepared model for replacement of legacy dimensions with corporately-managed master data
- Automated data integration from a variety of source system Excel and flat file outputs using SSIS, saving 5+ analyst days per month
- Populated new model from historical data (T-SQL) and new sources (SSIS), demonstrating consistency with existing report outputs and ensuring continued availability of legacy BI assets

August 2016 – August 2017 (one renewal)

Data Warehouse Consultant – North Staffordshire Combined Healthcare NHS Trust

Integration of data from new Lorenzo patient system into existing SQL Server warehouse platform, enabling seamless reporting across source systems and retention of value in existing reporting assets.

- Built SSIS packages for flat file ETL
- Designed and built ETL metadata repository using SSIS to sync supplier metadata
- Used BimlScript to generate 850 SSIS packages to standard ETL pattern from metadata
- Automated T-SQL build of staging objects from metadata, supporting schema upgrades in-situ
- Implemented metadata-driven ETL application in C#
- Built searchable end-user view of historical system data using SSRS and legacy Oracle back end
- Developed master patient record in operational data store, automatically combining records from multiple sources

March 2015 – July 2016 (three renewals)

ETL Developer – Staffordshire & Shropshire Health Informatics Service

Incorporation of new acute hospital tenant into multi-tenant SQL Server warehouse platform.

October 2014 – March 2015 (one renewal)

BI Developer – Birmingham & Solihull Mental Health NHS Foundation Trust

Implementation and extension of new SQL Server reporting data model.

May 2013 – September 2014 (six renewals)

Data Warehouse Developer – Staffordshire & Shropshire Health Informatics Service

Architecture and development of shared, multi-tenant SQL Server data warehouse platform, hosted by a shared service for the use of multiple customer organisations.

Education

PhD, Computer Science

Virtual forced splitting in multidimensional access methods (http://etheses.bham.ac.uk/213/)

University of Birmingham, 2008

MSc, Computer Science

Distinction

University of Birmingham, 2001