WHITE PAPER

The Value of Using a Logical Approach to Data Management in Energy and Utilities

By Mike Ferguson
Intelligent Business Strategies
April 2023

Research sponsored by:

Table of Contents

INTRODUCTION – ASPIRING TO BECOME DATA-DRIVEN

Executives now see data, BI and AI as mission critical to their business

Executives in most companies today recognise that data and analytics have become central, if not mission critical to helping their business run smarter. Their expectation of what data, business intelligence, machine learning and artificial intelligence (AI) can do for them is huge. They see it as the way in which their companies can cut costs, improve efficiency, reduce risk, and seize opportunity by making timelier, more accurate decisions. They also see it as the way they will disrupt the markets they compete in and uncover new opportunities in new markets with new products and business models.

A high quality, secure and compliant data foundation of business-ready data needs to be created to shorten time to value

For this to happen, companies need to build a high quality, secure and compliant data foundation upon which such competitive insights can be created. Data needs to be 'business-ready' and available to share across the enterprise and beyond to help shorten time to value.

Data complexity is a major barrier to success

However, there are barriers in the way of success. One of the main challenges is data complexity caused by more and more data sources and more data stores that exist on-premises, in multiple clouds, in software-as-a-service (SaaS) application provider data centres and at the edge. Many companies now have data persisted across a truly hybrid, multi-cloud distributed data estate as shown in Figure 1. Data is stored in cloud and on-premises relational databases, NoSQL databases, cloud storage, file systems, and content management systems. It may also be streaming in from edge devices and ingested from multiple external data providers via many different mechanisms such as data marketplaces and data cleanrooms. So, we have a dilemma. On one hand, executive expectation is that data and AI means the company will be more agile and move at speed but on the other hand, the data needed to make it happen is becoming harder to find and integrate.

Many companies now have a hybrid, multi-cloud distributed data estate

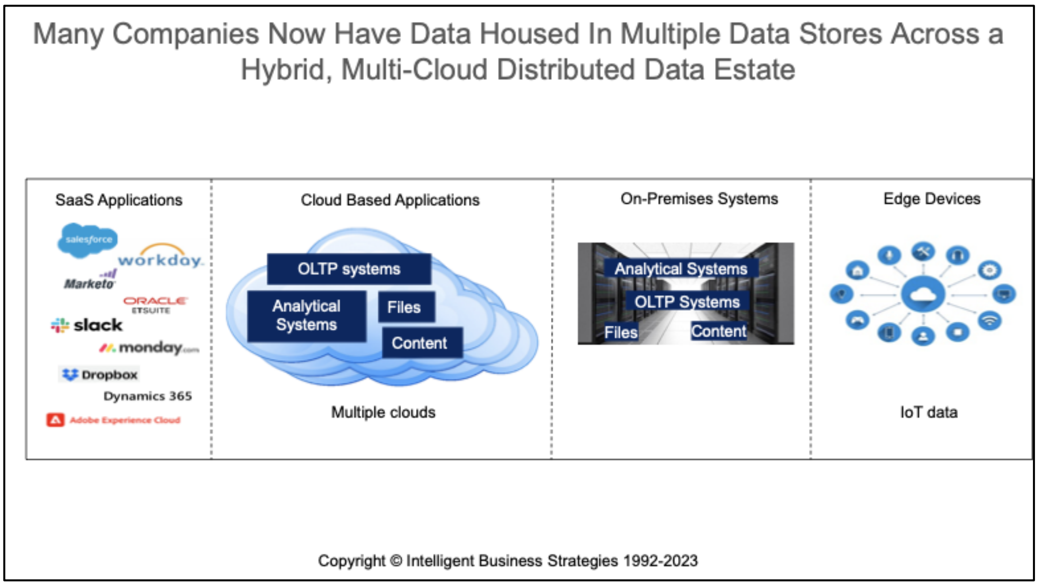

Figure 1

This paper looks at this problem and at how data virtualisation can provide a platform to help solve it so that Energy and Utilities companies can shorten time to value, meet executive expectations and achieve their goal of becoming data-driven enterprises.

Hiding the Complexity – The Need For Data Fabric

Given the data complexity that many companies are facing, there are some critical technology components that need to be in place to speed up the creation of a data foundation that enables companies to build insights, predictions, recommendations, and AI-automation to help create a data-driven enterprise. This includes the need for data fabric software.

Data fabric software connects to data across a distributed estate enabling companies to find, engineer and govern data

Data fabric is data management platform software that is capable of connecting to the wide range of data stores and sources across a hybrid, multi-cloud distributed data estate. It offers the full capabilities to support automated data and relationship discovery, data classification, governance, cleansing, transformation, and integration of data. It can be used by multiple IT professional and citizen data engineers to develop orchestrated pipelines that produce business-ready data products available for sharing securely across the enterprise in a compliant manner.

It often includes a data catalogue, support for collaborative development of DataOps pipelines and a data marketplace

Data fabric software includes support for secure connectivity to a broad range of data sources, a data catalogue, collaborative development, support for DataOps, performance and scalability, extensibility, unified data governance and a data marketplace.

It hides the complexity of a distributed data estate

Its job is to hide the complexity, simplify data access and accelerate the development of data engineering pipelines to produce reusable data products that can be consumed to create machine learning models and a data warehouse to support business intelligence analysis and reporting. This is shown in Figure 2.

Multiple teams of data producers can use common data fabric to build pipelines that integrate data to produce reusable data products

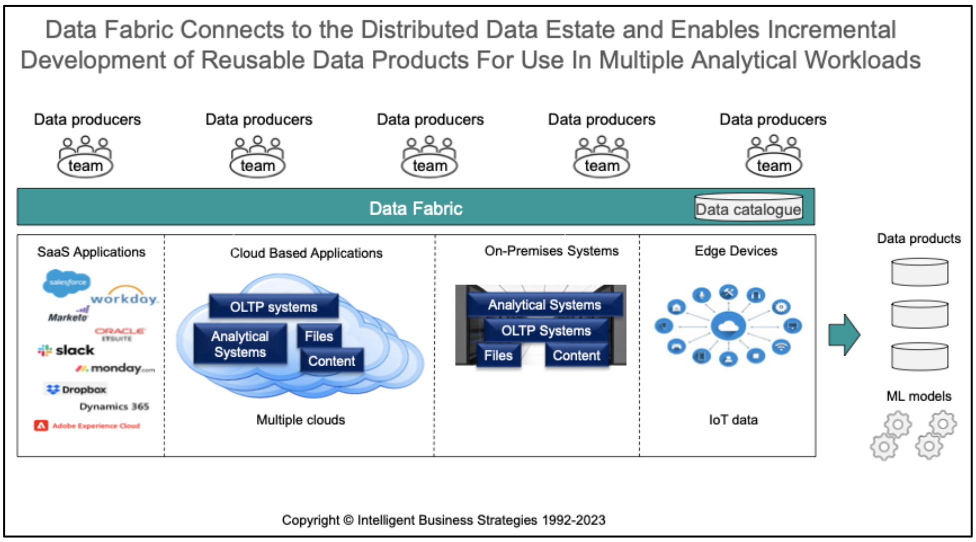

Figure 2

A Logical Data Fabric for the Data-Driven Enterprise

Logical data fabric technology is software that can be used as a platform to create virtual data products from data in multiple underlying data stores.

Data products can be created by defining virtual views of data entities using common business data names defined in a business glossary. The platform's data catalogue provides business glossary capabilities together with metadata on source data needed to create each data product. Identified data in multiple data sources can then be mapped to the schema of each virtual view describing a data product. This approach allows you to *incrementally* create a layer of semantically linked virtual data products that can be:

Logical data fabric enables virtual data products to be created and consumed for different analytical use cases

- Rapidly assembled into different virtual schema optimised for different analytical workloads e.g., data warehouse, data marts, data science.
- Queried to provide feature data for machine learning model development in data science.
- Used as input to create other virtual data products.
- Consumed via different interfaces e.g., SQL, REST, ODATA, GraphQL

This is shown in Figure 3.

Layers of virtual views can be defined that help to rapidly integrate data from multiple sources and make it available in business terms for consumption

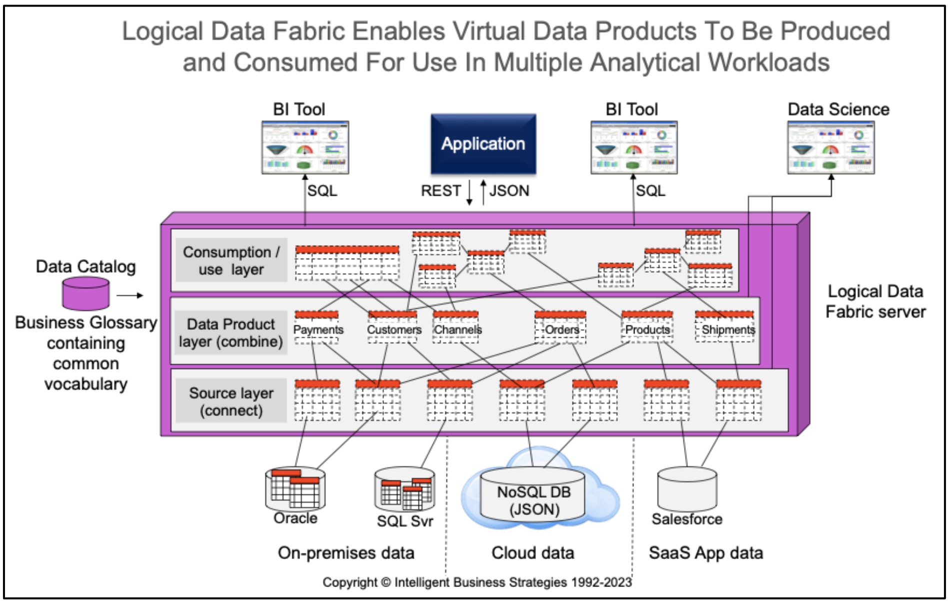

Figure 3

Data products can be quickly combined to create other virtual data structures

Data can also be persisted if required

Note that this ability to quickly 'assemble' semantically linked data products to create other virtual schema signals a major step change in productivity that increases agility and shortens time to value. It provides business-ready data to jumpstart multiple analytical projects as well as enabling component-based schema development. In addition, it enables a common semantic layer to be created so that all tools and applications are served data with the same common business data names.

CROSS INDUSTRY USE CASES

Logical data fabric can be used by companies in any vertical industry

Logical data fabric can be used in a number of ways that would apply to every industry. Popular use cases would include using platform data virtualisation capabilities to integrate data from multiple underlying data sources:

- To create a customer data platform for an omni-channel customer intelligent front office
- To produce the data needed for timely regulatory reporting associated with compliance obligations in one or more legal jurisdictions.
- Modernise data warehouses.

Let's look at these in more detail.

USING LOGICAL DATA FABRIC TO CREATE A CUSTOMER DATA PLATFORM

A very popular use case is to use it to integrate data to create a 360º view of a customer

To compete in a digital economy, companies need a lot more data on customers than is stored in their transaction processing application databases and data warehouses. Today, companies have to know not only who their customers are, and their transaction activity but also all their interactions with your company and what their opinions are, to provide an understanding of their online behaviour, customer location and product or service usage behaviour. This means that you need more data from both inside and outside the enterprise.

Creating a single view of the customer in a customer data platform using logical data fabric means data can be analysed using machine learning to drive consistent personalised recommendations across all front-office applications for each customer. A detailed understanding of how to create a customer data platform using a logical data fabric can be found here .

USING LOGICAL DATA FABRIC FOR REGULATORY REPORTING

Logical data fabric can also integrate data and make it available via a virtual view to simplify regulatory reporting

Logical data fabric can also be used to quickly connect to multiple data sources and integrate data for regulatory reporting. Also, virtual views can be created on the platform to integrate the data needed to provide reporting for each different regulation on one or more legal jurisdictions. This dramatically simplifies the regulatory reporting task. Also, data snapshots can be kept that represent data at specific points in time and data in virtual views will always be up to date. Also, by using common data names in virtual views, a common semantic layer can be created for regulatory reporting that is consistent across all reporting tools used.

USING LOGICAL DATA FABRIC TO MODERNISE DATA WAREHOUSES

Another popular use case is data warehouse modernisation

Logical data fabric can be used to modernise data warehouses. For example, it can be used to break the dependency between business users using self-service BI tools and the physical schema of a data warehouse and / or data marts. This enables data warehouses and data marts to be migrated to the cloud while shielding business users from any structural changes in the data or changes in DBMS software that have occurred as a result of the migration.

In addition, data warehouse architecture can be modernised and simplified by using the data virtualisation capability of logical data fabric software to replace physical data marts with virtual data marts. This can be done in a non-disruptive way and makes it possible to reduce the number of data stores in the end-to-end data warehouse architecture.

THE VALUE OF A LOGICAL DATA MANAGEMENT APPROACH IN ENERGY AND UTILITIES

Energy and utilities companies can benefit from the use of data fabric software

To understand the value of a logical data management approach in energy and utilities it is first important to understand current problems in the industry, what the business priorities are, the insights needed to help solve these problems, the relevant data sources and how data virtualisation can be used to integrate data from these sources to enable required insights to be easily produced.

PROBLEMS IN ENERGY AND UTILITIES

Energy and utilities companies face a public backlash on rising energy costs

The impact of the Ukraine war has caused major turbulence in energy markets over the last year. Energy prices have soared which has resulted in oil and gas producers announcing record level profits. However, on the retail side of the market, many energy retailers have struggled to stay afloat. Several companies have gone out of business and price rises have been pushed on to consumers. This has resulted in a public backlash on pricing levels, with governments stepping in to subsidise citizens and businesses to attempt to minimise fuel poverty levels and keep energy payments affordable.

Customer churn and bad debt are on the increase

Despite these efforts, the industry has seen an erosion in consumer confidence, with customer loyalty disappearing rapidly as businesses and consumers seek to reduce energy bills in every way possible. Fuel poverty also increases the focus on revenue assurance to minimise build-up in bad debt.

Climate change is also impacting service levels and causing outages

In addition to this, issues associated with climate change continue. Weather and environmental catastrophes are increasing and can impact service levels. Grid health and resiliency is being tested to avoid outages caused by extreme weather. This also has implications on prioritising field service crews in areas in which extreme weather catastrophes are predicted to impact grid networks. It is also important to ensure that supply chains and inventory levels are sufficient to accommodate demand on materials and parts caused by weather incidents.

Resiliency, preventative maintenance, and field service optimisation are important to avoid and respond to incidents

The impact of climate change on energy and utilities companies is broader than just maintaining service levels in more extreme weather. Customers are increasingly seeking to sign up with renewable energy providers. Therefore, utilities companies need to upgrade infrastructure to boost adoption of renewable energy and keep customer loyalty in sustainable energy providers.

Customers want to buy from renewable energy providers and participate in distributed energy generation

Also, climate change is impacting customer buying patterns as consumers and businesses look to switch to renewable energy and purchase more energy-efficient products, services, and equipment. The increase in purchase of electric cars is a good example of this which is causing major growth in installation of home charging points. Also, growth in distributed energy generation is boosting the need for more gas and power storage points across the network rather than reliance on central power-generating stations. Such changes make it even more important to ensure grid networks have advanced metering and remain healthy.

HIGH PRIORITIES IN ENERGY AND UTILITIES

Given these problems, energy and utilities companies have currently prioritised the following areas:

Priority Area	Reasoning

Renewable energy, grid modernisation, increased resilience, improved customer experience, field service optimisation, revenue assurance and ESG reporting are all priorities

Decarbonisation and adoption of renewable energy	Reduce the emissions per unit of electricity generated (measured in grams of carbon dioxide per kilowatt-hour)
Grid modernisation and upgrading of infrastructure	Advanced metering and industrial IoT enables more efficient generation and provisioning of energy to meet demand. Introduce infrastructure supporting renewable energy generation
Resilience and risk management	Strengthen grid networks to maintain service levels, reduce outages and survive extreme weather incidents
Customer engagement and experience	Increase customer self-service, improve personalisation, cross-sell and upsell
Digital transformation to improve operational efficiency and optimise supply chains	Drive out cost, optimise the supply chain to improve field service responsiveness and grid maintenance, optimise energy generation and storage to align with demand and usage patterns across grid
Revenue assurance	Improve oversight on revenues, minimise bad debt, improve cashflow, streamline operations
Environmental, Social and Governance (ESG) reporting	Evidence of safe, reliable, low-cost provisioning of power and gas, while effectively managing social and environmental impacts, such as community safety, methane emissions etc. Promotes confidence in brand

INSIGHTS ARE NEEDED TO SOLVE ENERGY AND UTILITIES PROBLEMS

To solve these problems, energy and utilities companies need to utilise data to produce insights that support decision making in each of the aforementioned high priority areas. That means creating insights associated with:

Insights are needed in asset management to understand and efficiently operate grid networks

- Asset management
    - Power assets such as wire and cables, transformers, breakers, structures, shunt compensators, voltage, and current transducers
    - Gas assets such as pipes, pipe fittings (couplings, bends, tees, end caps, taps), regulators, valves, compressors, anodes etc.
    - Electric and gas meters

  Asset management insights help energy and utilities companies visualise and operate energy grid networks, understand asset reliability scores, asset propensity to fail and cost of operating assets, etc. This also helps manage inventory, optimise field service supply chains, and support field service.
- Asset health to:
    - Optimise grid performance, identify underperforming assets and assets at capacity, monitor energy storage levels etc.
    - Measure quality of service, performance and usage of power and gas grid networks vs capacity - including peaks and troughs
    - Track inspections, inspection scores and treatments, asset installation and removal / decommissioning
    - Assess climate related physical risks from hotter temperatures, severe weather, rising sea levels, extreme winds, and heavy rain

Insights are needed on asset health to measure quality of service, identify underperforming assets, assess the impact of climate change, implement predictive maintenance, and optimise field service

    - Detect events that identify faults (e.g., gas leaks, power failures), fault causes (e.g., weather), affected assets, grid impact areas and customers impacted by outages and assets at risk of failing, causing outage and/or harm to human life or the environment
    - Analyse faults, affected assets, risk to network, and mitigate risks by optimising field service to carry out preventative maintenance to minimise outage
    - Take timely actions to re-prioritise work and respond to outages
    - Increase resiliency and minimise cost by identifying outage patterns, outage regions, outage restoration times, and gas leaks

Insights are needed to help budget and plan infrastructure upgrades to increase the adoption of renewables

- Asset financial planning
    - To analyse the cost of corrective maintenance and preventative maintenance as a percentage of total maintenance cost
    - To support budgeting and planning of major grid modernisation infrastructure projects including investments in renewable energy infrastructure for lower cost generation and in decarbonisation

It is also important to measure demand and grid usage to optimise energy distribution, energy storage and exploit distributed energy generation

- Grid usage analysis to understand and forecast:
    - Consumption and renewable energy usage
    - Peak loads, energy storage, grid health, load planning
    - Demand for gas entering / leaving citygate measuring stations
- Energy distribution analysis
    - To analyse generating capacity vs demand
    - To understand power stations, substations, nodes, distributed gas and power storage as well as amounts and value of power generated by net metering customers (solar customers)
    - To optimise energy distribution using distributed generation together with distributed gas and power storage
    - To understand grid capacity degradation (e.g., current state of power / gas storage)

The creation of customer living profiles is needed to improve marketing, enable personalisation, and improve risk management

- Customers
    - To create 360º living profiles that understand customer agreements, locations, tariffs, usage, billing, transactions, payment methods, collections, and contact points. This helps to improve customer experience, personalisation, cross-sell/up-sell, and service
    - To enable customer segmentation for more effective marketing
    - To identify and predict customers at risk or likely to churn

Insights are also needed to manage and help avoid bad debt

- Preventing and managing bad debt by:
    - Understanding payment methods and payment history
    - Predicting accounts likely to default and take preventative action
    - Understanding customer interaction history on overdue accounts
    - Managing collection activity

Field service optimisation is needed to improve service levels and respond to extreme weather incidents

- Field Service optimisation
    - To monitor work orders, projects, optimise dispatching, monitor status, costs, field service teams, and field service employee qualifications
    - To identify work delays and causes versus budget and forecasts

  - To optimise supporting field service supply chains

ESG reporting for power and gas is a priority to comply with regulations, protect brand and maintain customer confidence

- ESG power-and-gas-related metric reporting including compliance with Sustainable Accounting Standards Board (SASB) and reporting climate related risks in alignment with Task Force on Climate Related Financial Disclosures (TCFD)

REQUIRED DATA SOURCES

The number of data sources needed to provide data to support these analyses is increasing and includes:

The number of data sources now needed by energy and utilities companies to provide data for analysis is increasing rapidly

- Customer, product, and asset (e.g., gas and power assets) master data
- Agreements and customer account data
- Customer interaction data (outbound and inbound emails, chat)
- Customer usage data - smart metering data (readings, measurements)
- Financial ERP system data (invoices, payments, costs)
- Asset management systems data
- Grid/Network data (e.g., terminals, city gate stations, gas storage, power substations, power storage, net metering)
- Grid IoT data
- Location data
- Event history and outage data
- Supply chain data e.g., Suppliers, materials, purchase orders
- Catalogue item inventory data (warehouses, bins, items)
- Human resource data (employees, qualifications)
- Field service work allocation data
- Sustainability data (generation by renewables, net metering customers…)
- External weather, social network (e.g., outage reports and location on Twitter, Facebook…), social events (e.g., concerts, festivals that cause high demand etc.) and markets data

The ability to integrate this data in a timely manner is now a mission critical requirement

The data required to support analytical workloads is scattered across a hybrid, multi-cloud distributed data estate

Note also, that in many cases, more SaaS transaction applications are in use. Therefore, data needed can be in multiple SaaS applications, on-premises data stores, different data stores in multiple different clouds as well as huge amounts of live streaming IoT data coming off grid networks and smart meters at the edge. It is exactly like the distributed data estate shown in Figure 1.

USING LOGICAL DATA FABRIC TO INTEGRATE DATA FOR ANALYSIS IN ENERGY AND UTILITIES

The major challenge to overcome is the ability to engineer and integrate data quickly

Looking at the insights needed in energy and utilities and the growing number of data sources that energy and utilities companies need data from, there is no doubt that the major stand-out challenge that has to be overcome is the ability to engineer and integrate data quickly.

On-demand data integration is needed in every area, whether it be asset management, asset health, customer 360° live profiles, invoicing, marketing, bad debt, grid usage analysis or field service optimisation.

Companies need data fabric software to connect to the broad range of data sources needed and integrate data for specific kinds of analyses (see Figure 2).

Data fabric with data virtualisation capability enables on-demand data integration

This is where the data virtualisation capabilities of logical data fabric play a vital role. It helps to integrate the data needed to provide insights in all of the Energy and Utilities use cases mentioned. Let's look at a few examples.

Field Service Optimisation

Logical data fabric can help optimise field service

Figure 4 shows how data from disparate data sources can be integrated using the data virtualisation capabilities of logical data fabric to help optimise field service in Energy and Utilities companies. It can be done using a layered approach in Figure 3 that includes the creation of data products or as a logical data warehouse integrating multiple analytical systems or a combination of both.

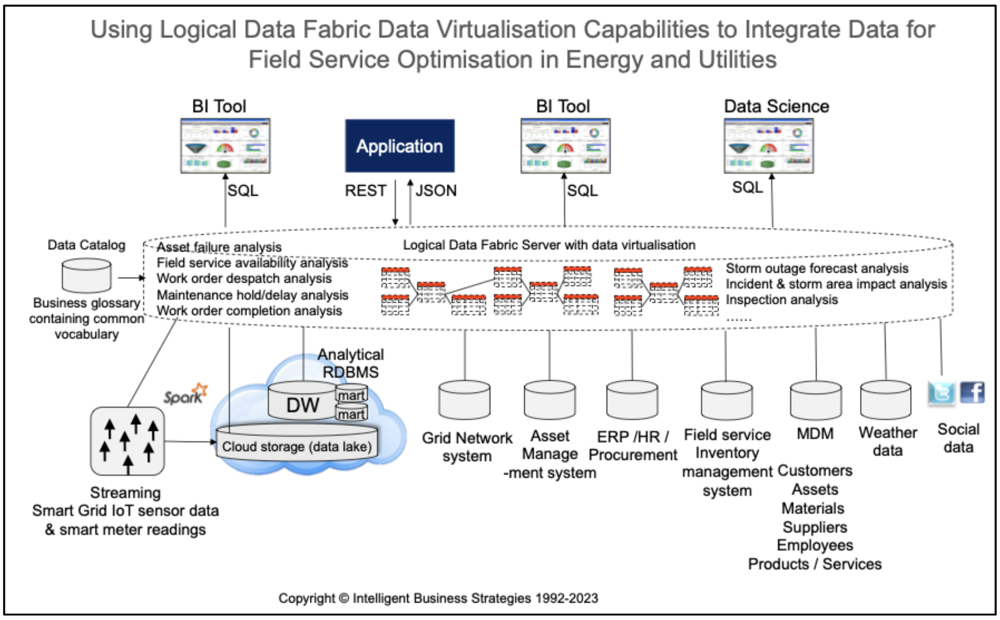

Figure 4

Integrating historical costed work order data in a data warehouse with real-time data on grid health, inventory, weather, and social network incident reports provides the holistic view needed to optimise field service

Here you can see how data can be accessed on:

- Real-time and historical grid IoT sensor data (grid / asset health)
- Real-time and historical smart meter reading (demand)
- The grid network (cables, towers, poles, pipelines, storage, substations…)
- Gas and power assets and asset maintenance work
- Field service catalogue inventory
- Field service inventory suppliers
- Field service contractor suppliers
- Field service employee data
- Weather data
- Social network data (social events, outage incidents, gas leaks etc.)

Existing field service-related analytical systems may already exist such as a data warehouse and/or a graph database representing the grid network. However, data from operational and analytical data sources can be integrated in virtual views to support different types of field service related analyses such as:

Multiple different analyses can be carried out on a rich set of integrated data available on-demand

- Asset performance, inspection, and reliability analysis
- Network risk and asset failure analysis
- Storm outage forecast analysis
- Incident & storm area impact analysis
- Field service availability analysis (crews, qualifications, inventory)
- Maintenance hold/delay analysis
- Work order despatch and completion analysis
- Work order cost Vs budget analysis

Credit Risk / Bad Debt Management

Figure 5 shows the use of data to manage credit and minimise bad debt.

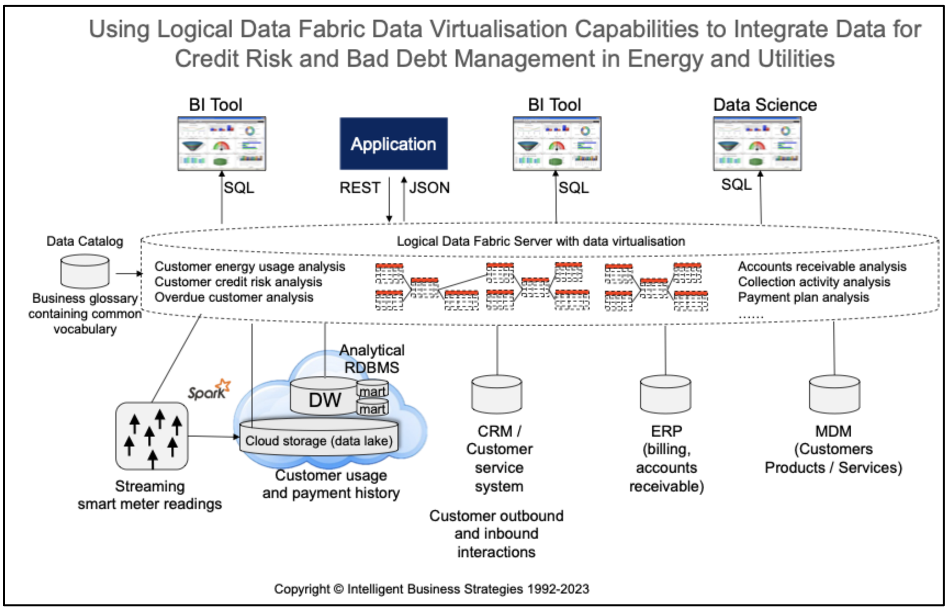

Figure 5

Logical data fabric can also be used to provide holistic virtual data views for credit risk management and to help minimise bad debt

Here you can see how historical and real-time metering data on energy usage can be brought together with customer master data, customer accounts, billing activity, accounts receivable and payments history and customer interaction. Historical usage data and payment history may reside in a data warehouse. Also live streaming usage data from smart meters can be landed in a data lake.

Holistic virtual data views can be created to enable trained ML models to predict credit risk before it happens so that payment assistance plans can be put in place to avoid bad debt

The ability to integrate historical and real-time data provides a holistic view of the situation. Data science can create machine learning models to predict customers likely to default on payments and/or who may become credit risks. Different virtual views can be created to support different types of analysis like:

Virtual data views can be created to support different types of analysis across the end-to-end process
Virtual data views can be created to support customer 360º, optimisation of energy distribution and to maximise the value of distributed energy generation

- Customer energy usage analysis
- Customer credit risk analysis
- Accounts receivable analysis
- Overdue customer analysis
- Collection activity analysis
- Payment plan analysis (payment assistance agreement plans)

In addition to these examples, many other kinds of analyses can be undertaken by discovering and integrating data using logical data fabric. This includes marketing, sales, and customer service oriented analyses using data in a customer 360º data platform, and grid optimisation of central versus distributed energy supply/generation and storage to meet peaks and troughs in demand. Also, financial planning and supply chain optimisation.

CONCLUSIONS

Energy and Utilities companies must be able to integrate data from multiple data sources in a timely manner if they are to become data and AI driven enterprises

If Energy and Utilities companies are to become data- and AI-driven, their future is heavily dependent on their ability to integrate data from an increasing number of data sources in a timely manner. This requirement is now becoming mission critical to building a high quality, secure, compliant, and rich data foundation to help produce new insights, recommendations, and predictions. It is this foundation that will enable them to create the insights they need to modernise their grid networks, switch to lower cost renewables, increase resiliency and improve grid health to offer better service that can handle the impact of climate change and deliver lower-cost energy to customers.

Logical data fabric hides the complexity of a distributed data estate, improves agility, and can expedite data integration

Multiple teams of data producers can create reusable virtual data views that shorten time to value

The main challenge is overcoming the complexity caused by an ever-increasing number of data sources and data stores scattered across a truly hybrid, multi-cloud distributed data estate. Implementing data fabric that can connect to these data sources and support the building of virtual data products while enabling source data to stay where it is, will help companies improve agility, overcome the bottleneck in data engineering and incrementally deliver the insights needed to improve business outcomes. Logical data fabric offers these capabilities and is therefore a candidate technology that energy and utilities companies should shortlist when selecting data management technology for a data-driven enterprise.

About Intelligent Business Strategies

Intelligent Business Strategies is an independent research, education, and consulting company whose goal is to help companies understand and exploit new developments in business intelligence, machine learning, advanced analytics, data management, big data, and enterprise business integration. Together, these technologies help an organisation become an *intelligent business*.

Author

Mike Ferguson is Managing Director of Intelligent Business Strategies Limited. As an independent IT industry analyst and consultant, he specialises in BI / analytics and data management. With over 40 years of IT experience, Mike has consulted for dozens of companies on BI/Analytics, data strategy, technology selection, enterprise architecture, and data management. Mike is also conference chairman of Big Data LDN, the largest data and analytics conference in Europe and a member of the EDM Council CDMC Executive Advisory Board. 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. He teaches popular master classes in Data Warehouse Modernisation, Big Data, Centralised Data Governance of a Distributed Data Landscape, Practical Guidelines for Implementing a Data Mesh, Machine Learning and Advanced Analytics, and Embedded Analytics, Intelligent Apps and AI Automation.

Telephone: (+44)1625 520700
Internet URL: www.intelligentbusiness.biz
E-Mail: info@intelligentbusiness.biz