



## **Non-Operational IT Capex Re-Opener: NGET Data Portal**

**File name: Data Re-Opener – V3.0**

**Date: 15 September 2023**

## 1 Summary table

<b>Name of scheme</b>	NGET Data (Portal and Product Investment)
<b>Primary drivers</b>	Investment to realise enhanced Data Best Practice (DBP) and the provision of DBP portal requirements and associated content (products) and achieve the data transformation plan to increase availability and transparency of data to Ofgem, Industry and the consumer.
<b>Cost</b>	£6,696,026
<b>Outputs included in previous RIIO Business plan</b>	£0

## 2 Executive Summary

- 2.1.1 This reopener submission is seeking funding of £6,696,026 for an investment in the underlying technology and content which will provide NGET's stakeholders with simple, direct, and easy access to NGET data and insights. It will ensure continued compliance with data best practices, provide a foundation to accelerate the provision of valuable insights and analytics to the market, rapidly increase efficiency, reduce waste in the provision of data and simplify the overall data landscape with whole system thinking to easier adaption to change.
- 2.1.2 This submission looks at two key elements, the Data Portal Infrastructure and Data Product Acceleration, these are key enablers for NGET to meet its strategic objectives and ensure that we remain at the forefront of data management in the energy sector.
- 2.1.3 For nearly 30 years we have been operating with big, complex data platforms which make connecting, enriching and consuming data difficult. Our objective is to modernise our approach through the creation of a single and connected platform (Data Fabric) unifying existing components and avoiding the future creation of disconnected, resource dependent and expensive technical components of a traditional data platform.
- 2.1.4 The outcome of this investment will be:
- Accessible data connections resulting in fast access to data for those that need it and improving efficiency in the delivery of key insights.
  - Reduction in data duplication and storage so that Data infrastructure shows the same focus on reduction in the emissions associated with the adding infrastructure.
  - Creation of, simple tools to enrich and add value, consistent data descriptors through automation, adding value to the industry by making easier to connect and understand the data available.
  - Ensure that data is easy to find and govern, improving quality and reducing manual exercises to answer key questions.
  - Make enriched data products instantly available to interested parties as part of one platform, providing value to the consumer by reducing waste in ad hoc provision of data via manual means.

- 2.1.5 Ahead of this submission, NGET participated in an Request for Information (RFI) on data best practices and held several sessions with the Ofgem Data team around our strategic direction and objectives, during which we highlighted the need for investment to achieve our collective goals. One of the key elements of these discussions was the portal and catalogue, and how they would be of much higher value as part of a complete platform which connects everyone directly (internal or external) to data. This has been further highlighted as enhanced DBP requirements around these topics were agreed as part of the recent consultation.
- 2.1.6 This is a critical investment for NGET to support our data strategy, it enables the continued development and value building on Data Best Practices, but more than that, it creates future-proofed connected platform, and content which provides continued consumer value with faster, open, and easy access to all data products across our estate to support decision making as we drive towards Net Zero and the Future of Electric.
- 2.1.7 This re-opener submission seeks support for the need and full allowances to deliver two key deliverables. The first is Data Portal Infrastructure, the platform and services which enables the creation, enrichment, and access to data ensuring the consumer has full visibility of available insights. The second is Data Product Acceleration ensuring we provide data governance, quality and accelerate the creation of the value-added data products which aid decision making, reporting, transparency, and insight.

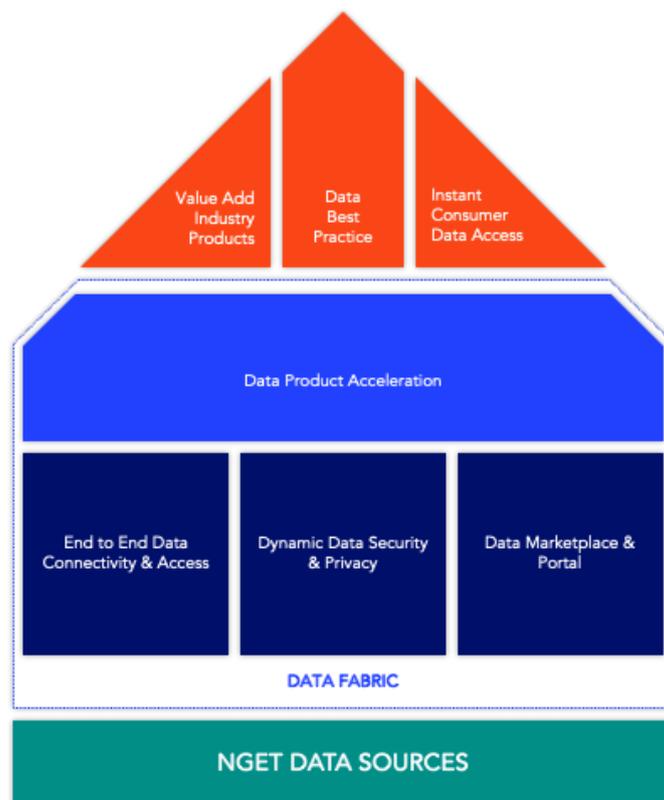


Fig 1 (NGET's Data Portal Infrastructure provides the foundation for the Data Product Acceleration which delivers Ofgem Data Best Practice, EDiT recommendations and NGET Network Strategic outcomes)

- 2.1.8 The Data Portal Infrastructure, which for brevity will be called Data Fabric throughout this document, will unify and further develop all components, including the catalogue and portal, of a modern data environment, automate the connectivity and enrichment of data, ensures all value-add data products are automatically defined and described and published to an open marketplace where the consumer can find and use them.
- 2.1.9 To further simplify the design, it will provide access control down to a single data element and dynamically change the view of these products based on the appropriate level of consumer access ensuring security and privacy is at the heart of the platform. This will enable NGET to remove the restrictions of traditional platforms so data can be collected faster, identify and resolve issues with data quality, accelerate insights from data, and better inform business decisions.
- 2.1.10 The benefits of the Data Fabric (Including the integrated portal and enhanced catalogue) include:
- Increases the efficiency in the implementation of common data models, data interoperability and product development.
  - Future-proofs the Platform, Capabilities, and Content which improves speed to insight and reduces wasted effort in the connection and consumption of data.
  - Provides Direct Data Customer access and data products and removes the need to replicate, copy or extract data saving time and wasted resource costs.
  - Simplifies the data landscape and will significantly reduce the volume of vendors and partners used, lowering long terms costs to the consumer.
  - Single front end for consuming data, for both internal and external data sources.
  - Creates the foundation for interoperable industry wide products.
  - Enables the compliance with enhanced Data Best Practices from the recent consultation in June, this included Enhanced Cataloguing, Dublin Core Adoption and Open Data Standards.
- 2.1.11 The Data Product Acceleration and delivery elements will enable NGET to easily adapt to changes and standards by developing industry standard data products and treating data as a product. This will provide the wider industry with easily accessible data content which can be directly accessed and consumed.

The benefits of treating Data as a Product include:

- All products and sources of data are catalogued and described by default, making them easy to find and use by customers.
- All products are contracted, which ensures that they have owners, quality metrics, and end-to-end lineage. This makes it easy to manage, maintain, support, and govern.
- In addition, there are the following Data Mesh and Fabric Benefits:
  - The Data Platform enables Content to provide an interoperable mesh (connected Data Products). This means that products can be shared internally and externally (subject to the portal) and reused across different teams, departments or externally to consumers and stakeholders.
  - By contracting data products through the fabric, it ensures they are interoperable, which makes it easy to combine them for more insight.
  - By enabling data connectivity and product development within the fabric. it enables us to easily adapt to changes and standards and provide the wider industry with easily accessible data content.

2.1.12 In preparation for the submission, we have reviewed and assessed a variety of options to ensure we achieve the outcomes and add the value we need; we considered three core approaches.

- Buy the Data Fabric components as Software as Service from the market, via our existing contracts.
- Hybrid use open source to provide portal component, without a fabric or advanced security controls.
- Build creates an in-house capability to develop and enhance the Data Fabric and content.

Following the assessment “buy” has been established as the recommended approach because it meets NGET’s strategic needs, it aligns with our architecture principles (Appendix 6) is the lowest risk and provides the best value to the consumer. [Section 9](#) Project Justification Optioneering goes into further detail.

2.1.13 As part of the optioneering, assessment and the overall design process we have broken the overall platform architecture in easily consumable groups which will point to several activities regarding Procurement of services and capabilities. We have assessed our existing contracts and frameworks to identify whether the capabilities we require are available, to simplify and accelerate the procurement process, delivery, and benefit realisation. We will use our Strategic Sourcing Process, which assures a robust procurement process is followed, and complies with Utilities Contract Regulations 2016 (UCR).

2.1.14 The Data Fabric will be implemented by NGET, with additional resources from partners. The Data Product Acceleration will be delivered by NGET, and these resources are part of the business units, to improve delivery of business value and increase domain expertise and organisational memory.

2.1.15 There are several risks to the implementation, development, and maintenance of the Data Fabric. They fall into Technical, Quality, Programmatic and Business Risks.

- Technical risks include integration of the Data Fabric.
- Quality risks include meeting security best practice requirements.
- Programmatic risks include scaling the team and finding appropriate skills.
- Business risks include cost over-run and failing to deliver business value.

2.1.16 We have identified a range of mitigations for these risks, they include:

- Inclusion of industry standard approaches to integration for software and SaaS.
- Inclusion of NGET’s security baseline requirements into the procurement phase to ensure that the solution meets our quality requirements.
- Early engagement with our suppliers and partners, with skills and capacity shared early to build a talent runway.
- Good governance and business value tracking and adjustment through our Data Product and utilisation of our SAFe Scaled Agile delivery framework.

2.1.17 The programme of work will commence in January 2024. There will be two streams of work, the Data Fabric, and the Data Product accelerator.

- The Data Fabric has three phases.



Fig 2.1 Data Fabric phases

- The Data Product Accelerator is a continuous phase of work starting in January 2024 and runs until January 2026.

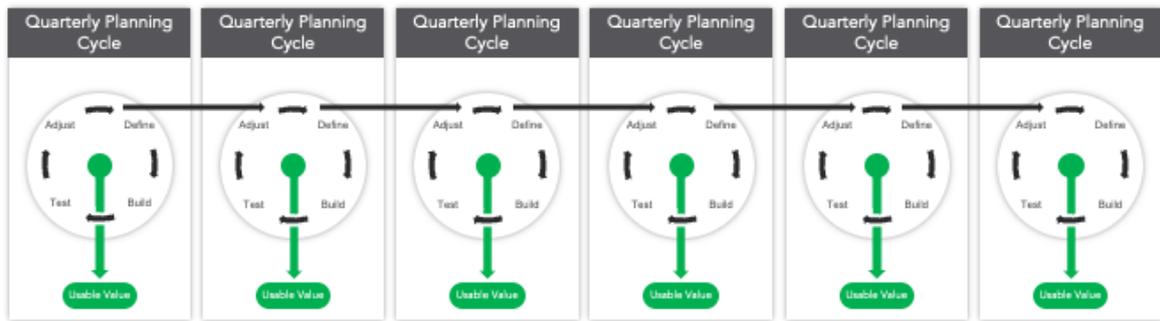


Fig 2.2 Data product accelerator continuous phase

2.1.18 Summary Cost Breakdown:

	FY22/23	FY23/24	FY24/25	FY25/26	Total
CAPEX	0	0	£ 2,496,670	£ 3,262,826	£ 5,759,496
OPEX	0	0	£ 160,349	£ 352,049	£ 512,398
RISK	0	0	£ 142,211	£ 281,921	£ 424,132
TOTEX	0	0	£ 2,799,230	£ 3,896,796	£ 6,696,026

Table 1: Summary of Programme Cost (18/19 prices)

## 1.1 Version Control

Version Number	Summary of Changes	Name	Date
1.0	Initial draft template	John Brookes	30 June 2023
2.0	Draft for sign off	Matt Clark	11 September 2023
3.0	Approved for submission	Sarah Milton-Hunt	15 September 2023

For more information, please contact:

John Brookes, Regulatory Manager

Email: [john.brookes1@nationalgrid.com](mailto:john.brookes1@nationalgrid.com)

Matt Clark, Data Director

Email: [matthew.clark1@nationalgrid.com](mailto:matthew.clark1@nationalgrid.com)

### 3 Table of Contents

2	Executive Summary.....	3
3	Table of Contents.....	9
4	Glossary of terms.....	10
5	Project Justification: Challenges and Opportunity.....	14
6	Project justification: The strategic context.....	16
7	Project Justification: Needs Case.....	18
8	Project Justification: Engaging our Stakeholders and whole system thinking.....	28
9	Project Justification: Optioneering.....	31
10	Cost Benefit Analysis.....	36
11	Project Definition: Delivery Strategy.....	39
12	Project Definition: Risks.....	44
13	Procurement and Sourcing Strategy.....	48
14	Detailed resource plan.....	51
15	Cost Assurance.....	53
16	Concluding Statement.....	58
17	Appendices.....	59

## 4 Glossary of terms

4.1.1 Table below provides a useful glossary of terms to support the assessment of this reopener.

REF	Term	Description
01	<b>Data Mesh</b>	Data Mesh is the term used to describe our Data Operating Model, the purpose is to push as much of the data interaction and value-added activities towards the business subject matter experts who understand the data best and can add the most value. It is a non-technical construct based around value generation, data creation and data ownership
02	<b>Data Fabric</b>	A Data Fabric is the term used to describe the technical platform which connects, catalogues, and provides access to Data for Product teams to work and develop on that data. It supports and enables it accelerating access to data and enabling automated governance & management.
03	<b>Data Catalogue</b>	Refers to the Catalogue Platform which holds all details of a Data Product and the corresponding Data Contract. Details include Structural Metadata, Tagging, Quality Metrics, Data Types, Privacy Controls, Categorisation, Lineage and ownership.
04	<b>Data Product</b>	A Data Product as a self-contained but interoperable data “container” that directly solves a business problem. They are built for internal or external users at various levels of maturity.
05	<b>Data Contract</b>	A data contract is the central component of a Data Product which ensure that every product is built to the required standards and against prescribed governance and best practices. The contract contains details include Data Quality Assessment criteria, Model Alignment, Data Ownership, Product Ownership, SLA’s, Subject Matter, Descriptions and Lineage details.
06	<b>Data Product Manager</b>	The Data Product Manager is the individual who is accountable for the creation of Data Products for a specific line of Business, they may also be the owner of Data Domain. The Data Product Manager has end to end accountability for data (including quality and security) from its originating source through to its use in Data Products.
07	<b>Data Domain</b>	A Data domain is the categorisation of data within a broader subject matter and the teams within who are accountable for the quality and use of data within that domain. An example of a data Domain would be Asset.
08	<b>Data Product Strategy</b>	This refers to NGET’s Data Strategy which is to treat Data as a Product and the drive to ensure end to end ownership of data from creation to use. It is through this strategy we are driving Data Products, Data Ownership, Quality Improvement and increased insights and Analytics as part of the broader Digital Strategy.
09	<b>Data Guild</b>	The Data Guild is the Decision Making and Policy Defining group within NGET and is made up of Core Data Enablement team and Data Product Managers, chaired by NGET’s Data Director. It runs to set the standards and policy which directly impact the improvement in data practices across NGET, manages risk and issues, sets governance and controls where needed.
10	<b>Communities of Practice</b>	The Communities of Practice (CoP) are groups of data professionals where best practices and opportunities are discussed and shared to increase talent and skills across the Data Community.
11	<b>Data Portal</b>	This refers to the externally facing website which will be developed to provide a user-friendly interface for access to NGET data as detailed in the Data Catalogue. It will include the ability to search and locate Open Data, log Quality Issues and Risks, contact a Data Owner and make requests for additional information or Data.

12	<b>Data Triage</b>	This refers to the categorisation of data where the Open Data Principle may not apply, it would include commercially sensitive data, PII data, critical infrastructure which may pose a security risk, internal data, and open data
13	<b>Data Product Lifecycle</b>	This refers to the end-to-end process for creating a Data Product and all the steps to ensure that the Data Contract and associated standards are adhered to in a products creation.
14	<b>System of Record</b>	The originating source of data, where it is created and drawn from, and considered the single source of the truth for that particular data.
15	<b>Data Customers</b>	As part of the NGET Data Strategy and Data Product implementation we refer to users or consumers of Data as Customers. We do this to focus Data Product providers as facilitators of good services to their customer base.
16	<b>Data Portal</b>	A component of the Data Fabric which enables seamless access to the underlying data products
17	<b>Marketplace</b>	Interchangeable term with data portal
18	<b>Dynamic Security and Privacy</b>	This term refers to the ability to flex the view of a data product dependent on the type of customer accessing it. It reduces the needs for multiple products with dynamic shifting and removal of sensitive or protected data on access.
19	<b>WSJF</b>	Weighted Shortest Job First, a method to prioritise work for maximum economic value.
20	<b>SAFe</b>	Scaled Agile Framework, aims to improve business, time to market, improve quality, increase productivity, and have happier engaged employees.

*Table 2: Glossary of terms.*

## 4.2 Key supplementary documents to support the assessment of this re-opener.

4.2.1 The re-opener references work undertaken by Energy Data Taskforce, Ofgem and NEG, this work influences and informs the re-opener. The following documents provide additional detail context to the consumer, government, and energy sector environment.

Appendix 1: Cost Benefit Analysis RIIO

“RIIO-ET\_CBA\_Data\_Portal.xlsb”

Appendix 2: Ofgem Data Best Practice

“Data\_Best\_Practice\_Guidance\_v1.pdf”

Appendix 3: Energy Data Taskforce (EDiT) Recommendations

“EDTF-Report-Appendix-1-Recommendation-Actions.pdf”

Appendix 4: NGET Data Review

“Becoming a Data Driven Business at National Grid.pdf”

Appendix 5: NGET Data Strategy

“NGET Data Strategy.pdf”

Appendix 6: NGET Architecture Principles

“NGET Digital Guiding Principles.pdf”

Appendix 7: EDiT Recommendations Mapping

“EDiT Recommendations Mapping.xlsx”

Appendix 8: Ofgem Data Best Practice Mapping

“Ofgem Data Best Practice Mapping.xlsx”

Appendix 9: Data Contract example

“Example data contract.json”

Appendix 10: Draft Data Fabric Capabilities

“Draft Data Fabric Capabilities.xlsx”

Appendix 11: Data Product Lifecycle

“Data Product Lifecycle.pdf”

Appendix 12: Assumptions

Table

Appendix 13: PARTNER Resource Costs

Table

Appendix 14: NGET Resource Costs

Table

### 4.3 Project justification: What is this re-opener requesting allowances for?

- 4.3.1 This reopener submission is seeking support from Ofgem on the need and the assessment of the allowances for an investment in new Data Fabric, to update NGET’s data architecture, and a Data Product Accelerator to deliver new Data Products insights from data within the portal to further develop the current data best practice compliance and service the enhanced data best practice outcomes.
- 4.3.2 This proposed investment will simplify the way data is accessed, transformed, and consumed, the outcome is that internal users, consumers, and wider industry, will have access to the open data licenced products needed to fulfil NGET’s strategy, Ofgem’s Data Best Practice Enhancements, and Energy Data Taskforce’s recommendations.
- 4.3.3 The Data Fabric will consist of three capabilities itself each provided by different SaaS (Software as a Service) products.
  - Data Product Catalogue – which will provide a single portal for data customers to discover, access and request Data Products.
  - Dynamic Security Access – which will provide Attribute Based Access Control covering Security and Privacy for Data Assets. Enabling the same Data Product to serve several different customers.
  - Data Fabric – which will allow Data Assets to be published to the Portal, without moving data from Systems of Record. This will allow us to provide near real time data products and develop and deploy them faster.

#### 4.3.4 NGET Supporting Data Platform Architecture

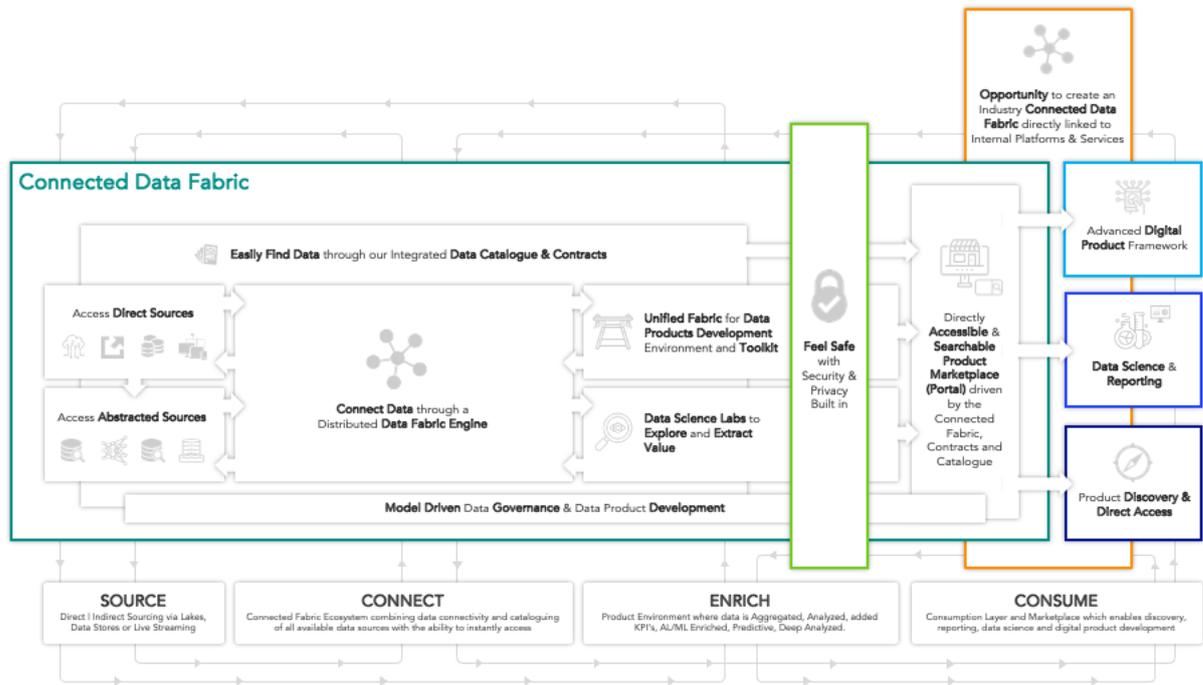


Fig 3 (Example National Grid Electricity Transmission Data Products Fabric Infrastructure)

The diagram starts on the left with the principle that a source is a source even if it is abstracted from the system of record, the fabric engine is a catalogue, a marketplace, a contracting platform, and a development environment with built in dynamic security and privacy. All of this enables the main objectives for data, delivery of digital products, data science, insight provision and data discovery and access.

4.3.5 The Data Product Accelerator will enable NGET to create new data insights and predictions via Data Products. We will as part of this work create and publish standards for federated governance (meaning automated and self-service governance managed through the data fabric to simplify and enable product creators) managed by Data Contracts.

- Data contracts (see appendix 9) will use standardised Metadata and logical data models, to allow inter-Data-Product and inter-Data Mesh interoperability. Including the Ofgem expected Dublin Core metadata standard tags and an extensible model which supports machine readable data contract.
- Federated governance to provide a consistent business information model and glossary will be embedded within the Data Products and their contracts enabling federated governance and data interoperability.
- Data Products will deliver a composable set of data, that delivers business value, and can be consumed on its own or with other Data Products.

## 5 Project Justification: Challenges and Opportunity

### 5.1 Challenges and programme drivers

- 5.1.1 Data is a key asset to any organisation, but to meet the challenges of decarbonising the transmission network we need to significantly accelerate our ability to expose and share data within NGET, the energy sector, our customers, and wider stakeholder communities. This demand means we need to adapt our capabilities to ensure that we fulfil our obligations, but also so we can provide a greater level of transparency, access and added value to relevant data consumers.
- 5.1.2 Ofgem has recently increased focus on Open Data, interoperability of data, increased speed of value-added analytics, increased reporting needs and the ability to easily find and explore available data via easy-to-use interfaces has prompted the case for change. It is understood that through these needs we improve the availability and quality of data, but this has highlighted opportunities to add real value through connecting industry wide knowledge and dramatically improving the diversity in the data we use to make decisions, of course a broader industry adoption would be needed to fully realise the opportunity.
- 5.1.3 There are three key drivers for the Data Fabric and Data Product Accelerator investment, which are:
- NGET using Data Best Practice and Energy Data Taskforce (EDiT) recommendations to drive a leading service in the provision of Open and Transparent Data Products to support the regulator and the industry.
  - Creating a data environment which will provide the capabilities and insights required to fulfil NGET's Network Strategy and enable the transition to a net zero power system.
  - Contributing to "technology whole system" thinking through democratization of data enabling staff, consumers, and stakeholders to discover high quality, accurate data.

This chapter describes these drivers in more detail and sets out why the heritage data practices are no longer suitable to support the demands we have on data today. These are common drivers to many of our IT investments as seen in section 6.2 where our Data Strategy is explained underpinning both this investment and the wider case for change if NGET is to deliver its strategy objectives.

## 5.2 Achieving Data Best Practice and EDiT recommendations.

- 5.2.1 NGET provided Ofgem our Data Best Practice Improvement Plan to ensure compliance with the Data Best Practice. This reopener paper sets out the Data Fabric and Data Product Accelerator investment that builds on the foundation of the NGET DBP Improvement Plan and in conjunction with the reopener will ensure full compliance with the enhanced Data Best Practice Guidance and enable a faster reaction to any future changes.
- 5.2.2 NGET has not submitted for Data Fabric and Data Product Accelerator funding to date. This means that there are no existing allowances within the current NGET RIIO T2 arrangements available to implement the Data Fabric and Data Product Accelerator outcomes required by Ofgem. The development and delivery of the Data Fabric and Data Product Accelerator will ensure NGET complies with its obligations under Part D of Special Licence Condition (“SpC”) 9.5 (Digitalisation) of the Electricity Transmission licence.
- 5.2.3 The Data Fabric can be used to adopt new platforms, capabilities, governance frameworks, and data products. This will help to achieve the wider objectives of the Energy Network, such as improving efficiency, reducing emissions, and increasing resilience.

## 5.3 Achieving Net Zero

- 5.3.1 NGET’s contribution to the government’s Net Zero targets and ambitions is predicated on NGET having good quality, trusted data to support sound decision making. The Data Fabric will provide a self-service platform enabling staff to find and request the data they need to gain insights to support high consumer value and efficient projects.
- 5.3.2 Through the Data Fabric we will be able connect to IoT sensor data (such as heat, performance, throughput from the ET infrastructure), allowing the us to build real-time data products and digital twins which can be used to run network scenarios and ultimately make predictive changes. This will allow the network to use predictive analysis techniques to prevent issues before they occur.
- 5.3.3 The Data Fabric simplifies integration to both internal and external systems. We will create the facility to combine external data sets with NGET customer information and using predictive machine learning models be able to build products that calculate customer predication to connect allowing NGET to better assign resources and operate efficiently.

## 5.4 Supporting “Technology Whole System” thinking

- 5.4.1 “Technology Whole System” thinking is a method to understand how things (elements and systems) are related, and how they influence one another. This term is often used to refer to the electricity network, however the data created from that network is an integral part of understanding it and why we use this terminology when talking about it here. The Data Fabric and Data Products enable the Data Mesh. Each Data Product on the Mesh includes a contract which describes the code, the data, and the component of the information model it describes. The Data Products interact and change in response to changes in the source data (SoR) and other Data Products. This collection of Data Products and their relationships is defined as the Data System for NGET.
- 5.4.2 This Data Fabric can determine the inter-relationships of systems, business process, and their data flow. This will allow NGET to abstract integration complexity by using Data Products as the integration layer.

- 5.4.3 The Data Fabric and Data Products will democratise data by supplying trustworthy data within NGET, stakeholders and consumers. The Data Portal enables our Data Product consumers to provide feedback and request new Data Products. This information creates new insights and knowledge as users of Data Products will bring new insights and value from NGET's data.

## 6 Project justification: The strategic context

### 6.1 Strategic Context

- 6.1.1 The Energy Data Taskforce (EDiT) was commissioned by BEIS, Ofgem and Innovate UK to focus on modernising the energy system to unlock flexibility and drive clean growth towards Britain's 2050 Net Zero target. EDiT developed a set of recommendations for how industry and the public sector can work together to facilitate greater competition, innovation, and markets in the energy sector through improving data availability and transparency. They determined that Data is fundamental to the future of our economy, which is why it is the focus of one of the Grand Challenges in our Modern Industrial Strategy. EDiT made the following recommendations, we will cover how these are fulfilled in the needs case.
- 6.1.2 In November 2021 Ofgem released its Data Best Practice and followed up by a consultation in May of 2023 which made Catalogue, Open Data and Metadata Standards adjustment these must be in place by the end of 2024. The purpose of the DBP is to ensure that data is treated as an asset, and used effectively for the benefit of consumers, stakeholders, and in the public interest. This submission focusses on ensuring the infrastructure and content provide a much higher and open level of transparency to the consumer through easily available data products.
- 6.1.3 NGET has a vital role in delivering the UK's decarbonisation commitments. To accomplish a green-backed electricity network by 2035 NGET is transitioning to a net zero power system. NGET needs to deliver the 'SuperGrid 2.0' to deliver the capacity, capabilities and intelligence needed of the network in the future. It needs to do this while also meeting our stakeholder expectations that the transmission network will maintain reliability and resilience, availability, affordability, safety, and sustainability.
- 6.1.4 All the questions that NGET needs to answer to meet the UK's energy transition, requires access to high quality data. We must move from manual reports, spreadsheets, with hard to validate data to new ways of working with predictive insights, based on machine learning, adaptive autonomous digital twins able to simulate decision making as they continuously learn and adapt to new environmental uncertainties.

### 6.2 NGET 2022 strategic data review

- 6.2.1 In July 2022 NGET undertook a review of the existing data architecture and operating model (NGET Data Review appendix 4)
- 6.2.2 The review formed the basis of a new NGET Data Strategy "data democratisation by default" (NGET Data Strategy appendix 5) which recommends a shift in mindset to how data is used, underpinned by an operating model that aligns:
- domain-oriented data ownership with the organisational knowledge,
  - applying product-based thinking to data,
  - providing automated self-service platforms to data consumers,
  - implementing modern data engineering practices; and
  - federated data governance & management practices.

- 6.2.3 Delivering and realising the data strategy will transform our approach to data and create a new architecture and organisational design, to deliver the data and insights at the pace NGET needs to enable the UK's energy transition.
- 6.2.4 Before the Data Strategy can be realised, two core changes are required to our existing operating model and architecture:
1. A Data Mesh, a social-technical architecture, where product thinking is applied to data by teams within the business who have domain expertise of the data they're working with. The Data Mesh needs a self-service architecture (the Data Fabric) to allow the teams within the business teams developing the Data Products to source the data and publish them to the consumers. The architecture must allow the consumers to self-serve, provide feedback, and create their own Data Products with appropriate boundaries and management.
  2. NGET's data needs to be democratised. This looks at how we use data within NGET. Distributing the teams with the skills and capabilities closest to the challenges and requirements in the business to build value aligned Data Products. It also includes development pathways for staff to improve their data skills and for the organisation to build its overall data capabilities. And in conjunction with the self-service architecture enabling staff, consumers, and stakeholders to discover, develop and realise insights from our data.

## 7 Project Justification: Needs Case

- 7.1.1 This section sets out how the data portal, products, and capability respond to the strategic context, programme drivers and NGET investment objectives.
- 7.1.2 All NGET's Strategy outcomes are enabled by good quality data, delivered in time to provide the insights the business needs.
- 7.1.3 The 2022 review of NGET's current Data Architecture identified that the Data Warehousing model would not support real-time data needs for integrated IoT, digital twins, Artificial Intelligence and Machine Learning needed to develop NGET data and insights. We need to implement a new architecture to support these outcomes.

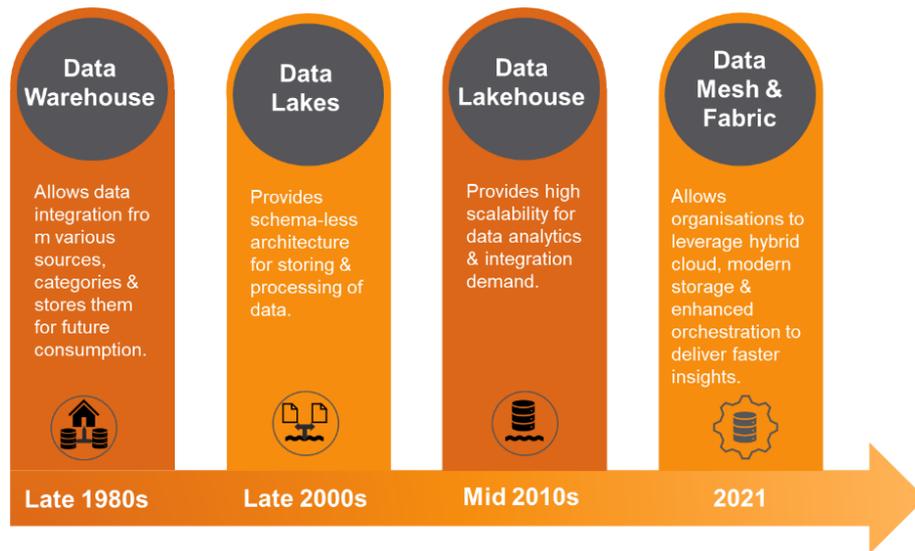


Fig 4 Evolution for Data Architecture over time.

- 7.1.4 Our current Data Warehouse architecture is limiting; slow to add new data sources, make changes to data models (relationships between different data), slowed by waste in handoff processes between business and technical teams, impeded by performance and scalability due to singular architecture, and making access to data is constrained rather than open by default.
- 7.1.5 The traditional data Warehouse, data storage model, means that there is a bottle neck in delivery of data. This model also requires lots of hand offs between teams and the movement of data from the source systems into the Warehouse, which slows processes and introduces errors.
- 7.1.6 The Data Mesh Operating and Data Product Model supported by a Modern Data Fabric removes our existing constraints and provides the freedom and acceleration in insights, that data warehouses, lakes and lake houses do not. Instead of moving and migrating data, the focus shifts to connecting and enriching data, and instead of central teams having business problems explained to them, those business teams can create the insights themselves. It is a positive step forward to breakdown data silos and connect data to those who need it. This is a key step in democratising data.

## 7.2 Need overview: Data Fabric

- 7.2.1 The Data Fabric provides the self-service platform that underpins the Data Mesh. It will enable NGET to make a marked strategic capability shift and provide the foundation to enable NGET to meet Data Best Practice and EDiT recommendations. (Appendix 7 and 8 covers how this investment meets fulfils the requirements).
- 7.2.2 Data Fabric will provide:
- 1 All customers will be able to discover all available Data Products
  - 2 Ability to review and search contextual supporting information about the Data Products
  - 3 Rapid increase in the speed, agility and efficiency in the connection and use of data internally externally.
- 7.2.3 The Data Portal will provide:
- All Data Products to be published to the Portal, without moving data.
  - Not just a Catalogue, but a direct window to published product.
  - Provide feedback to improve or request new data products.
- 7.2.4 The Dynamic Security Access will provide:
- Controls for internal and external consumers of Data Products.
  - Attribute Based Access Control covering security for Data Products.
  - Enables the same Data Product to serve several different customers.
  - Reduces product duplication and improves efficiency.

### 7.3 Data Fabric

- 7.3.1 This reopener requests a total allowance of £ [REDACTED] to deliver the Data Portal. If this reopener is successful design and build will commence in January 2024 with the Portal being live from September 2024, subject to funding release timescales.
- 7.3.2 The Data Fabric will improve the visibility of Electricity Transmission Data, it will help customers identify if a Data product exists, which business domain owns it, and then gain access to it. It will fulfil recommendation three from the Energy Data Taskforce “Visibility of Data”.
- 7.3.3 The Data Fabric consists of three capabilities Data Connectivity, Dynamic Security Access, and Data Product Catalogue which together will allow visibility of near real-time Data Products.
- 7.3.4 The Data Fabric will use the same technology for both internal and external customers. By combining a persona aware data portal with dynamic security access, we can create a customised attribute-based access to the same Data Product, with different views depending on what, where and who is accessing the data. The Data Connectivity is the glue that spans the platform, it provides an augmented integration layer allowing Data Products to query data without moving it, creating a Knowledge Graph (Think Google or Wikipedia, connecting related data types together) to provide curated persona aware inventories of Data Products in the portal.
- 7.3.5 Specific investment components diagram:

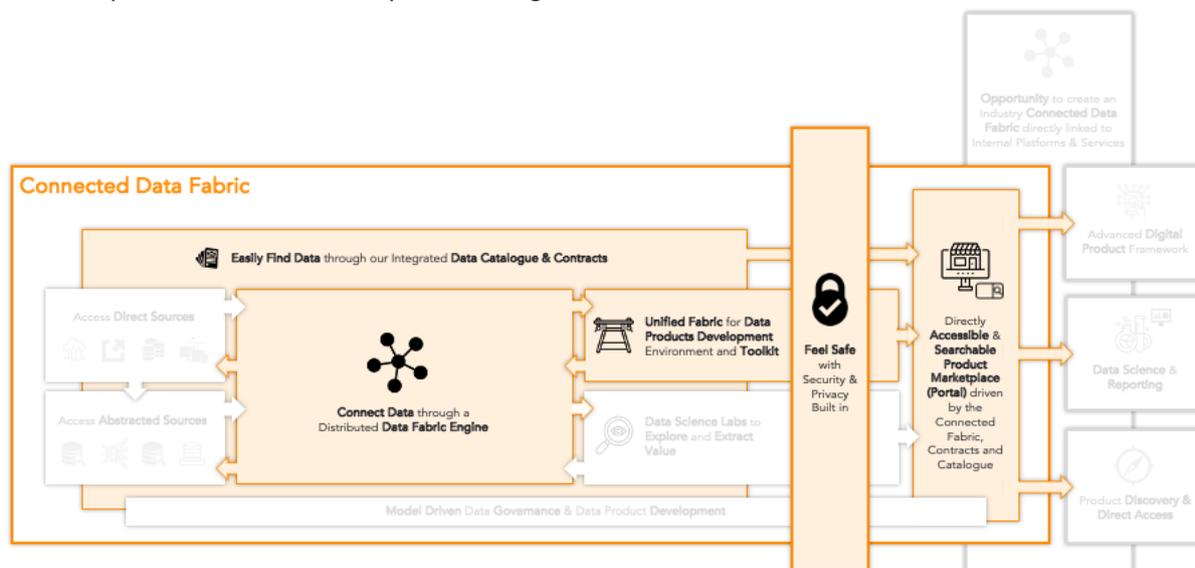


Fig 5 (High Level Architecture Framework for Data Fabric)

### 7.3.6 Specific investment components diagram (Lower Details):

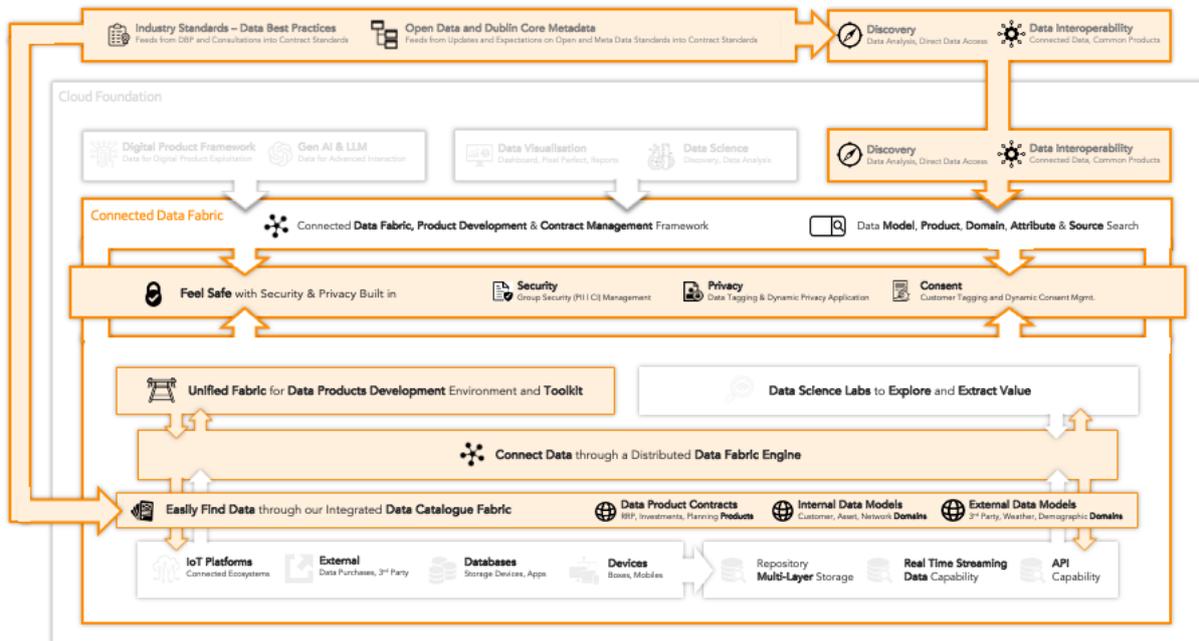


Fig 6 (Level 1 Architecture Framework for Data Fabric Aspects)

7.3.7 The diagrams above show the three capabilities in Orange that make up the Data Fabric.

7.3.8 The Data Product Catalogue investment of [REDACTED] will provide the following benefits:

- Greater transparency of data which is being collected and by which business unit.
- Identification of data quality and granularity issues.
- Increased adoption of metadata standards and information management best practice.
- Enable objective evaluation of increasing data openness by business units.
- Support demand driven prioritisation of data product build.

7.3.9 The Data Product Catalogue will include a Knowledge Graph technology, this allows for multiple relationships to be made between Data Products in the catalogue. Some of the benefits of this approach are listed below:

- Connected Data – A knowledge graph allows you to see relationships between different data assets and sources. This provides greater context and makes it easier to find relevant data.
- Discovery – Searching and querying a knowledge graph allows you to discover new datasets and tables that you may not have been aware of before. You can find related data more easily.
- Metadata Enrichment – A knowledge graph enhances the metadata about data sources. More semantic metadata makes data more understandable and usable.
- Data Governance – Linking data via a knowledge graph provides more oversight into the use and lineage of data. This improves data governance, compliance, and security.
- Agility – Comprehensive organisational knowledge of data assets allows companies to respond faster to business needs and opportunities. Adding new data is also faster.
- Automation – AI and ML algorithms can automate metadata tagging, relationship mapping and other knowledge graph capabilities to keep improving the catalogue over time.

- Collaboration – Knowledge graphs give a common language and understanding of data across teams. This improves communication and data democratisation.

7.3.10 The Data Product Catalogue will also allow customers to review Data Products in development, request new Data Products and updates to existing Data Products.

7.3.11 The Dynamic Security Access investment of [REDACTED] will work alongside the Data Catalogue to ensure that Data Products are protected in accordance with Security, Privacy and Resilience (SpaR).

7.3.12 Dynamic Security Access will use Attribute-Based Access Control (ABAC), rather than Role-Based Access Control (RBAC). This has several advantages, when applied to securely sharing data with different customers.

- More flexibility – With ABAC, access is granted based on attributes of the user, environment, and resource. This allows for more dynamic, context-aware access control compared to predefined roles.
- Better segmentation – ABAC can provide access segmentation down to the row or column level based on user attributes.
- Policy consistency – Policies follow the data attributes in ABAC. In RBAC, policies need to be manually aligned with related data.
- Granularity – Assigning access based on attributes allows for very granular control over access. Roles are broader and can overprovision access.
- Scalability – ABAC policies work the same regardless of company size. RBAC requires more roles to be defined as companies grow.
- Auditability – Determining why a user has access is easier in ABAC based on assigned attributes. RBAC requires matching access to roles.
- Automated Compliance – Attributes like location, time, etc. can automate compliance with regulations like GDPR (General Data Protection Regulation). Manual role setup is required in RBAC.
- Unified Policies – ABAC provides centralised attribute-based policies across all data platforms which make up the Data Portal, whereas RBAC is siloed to each capability/platform.
- No Access Sprawl – Adding a new access requirement simply needs a new policy in ABAC. In RBAC, this often leads to a new role being defined.
- Faster Auditing – ABAC's centralised logs make it easy to audit and trace the attribute-based policies that allowed access. RBAC logs require role correlation.

7.3.13 The Dynamic Security Access will allow one Data Product to fulfil all persona types from Data Scientists, Data Engineers, Business Users, Regulators, Non-Profits, and Industry Partners without having to re-engineer to protect access to PII (Personally Identifiable Information), GDPR and other sensitive data.

7.3.14 The advantage of this method is that it will reduce the time to create and deploy data products, whilst ensuring the SpaR principles are met.

- 7.3.15 The Data Connectivity investment of [REDACTED] will work with Dynamic Access and the Data Portal to provide access to Data Products in ways that are stored, archived, and provide sustained benefits, delivered in line with the fabric timeline.
- 7.3.16 The Data Connectivity will allow Data from Systems of Record to create a Data Product without having to move Data, this aligns to the data protection principles set out by the UK GDPR, which includes principles such as purpose limitation, data minimisation, and storage limitation that aim to ensure personal data is not copied or used beyond its original stated purpose.
- 7.3.17 The Data Connectivity will provide the following benefits:
- Unified access and management – The data fabric provides a single point of access and management across disparate data sources and environments. This simplifies working with distributed data.
  - Increased performance – Features like pre-caching, optimised querying and data virtualisation boost performance when working with large, distributed datasets.
  - Data governance – Consistent security, compliance and governance can be applied across all data sources in the fabric from a centralised interface.
  - Agility – Faster access to data with consistent meaning and context fosters business agility and innovation powered by deeper data insight.
  - Reduction in data duplication and consequential lowering of emissions related to data storage.

#### 7.4 Need overview: Data Product Accelerator.

- 7.4.1 Funding of £[REDACTED] for the Data Product Acceleration and delivery elements will ensure that NGET can easily adapt to changes and standards via the following content benefits:
- Data Content: The opportunity for the development of industry standard data products provides the wider industry with easily accessible data content which can be directly accessed and consumed.
  - The Data Products provide the basis for analytics, insights, and management within NGET but will also provide the foundation for Open Data principles and the easy publication of these products to the wider industry and interested parties.
- 7.4.2 By treating our Data as Product, we can ‘contract’ and ‘catalogue’ our data in the most effective manner ensuring the following:
- All Products are catalogued by default.
  - All sources connected to a Data Product are catalogued by default.
  - All products are contracted providing the following benefits:
    - Contracted Products all have Owners.
    - Contracted Products all have Quality Metrics
    - Contracted Products have end to end Lineage, Metadata and Context
    - Contracted Products have recognised frequency and Service Level Objectives (SLO)
    - Contracted Products are Validated Models for Interoperability

- 7.4.3 The Data Fabric and Data Products create a Mesh of federated Data Products. Teams with domain expertise develop their own Data Products aligned to business value. The Data Contracts all the Data Products to interoperate.
- 7.4.4 By adopting a product oriented and mesh approach to data we enable the ability to maintain compliance to ensure “business separation”, while making it easier to manage. The product ownership and connectivity capabilities ensure that products can be easily ported should things change, repointing data connections rather than migrating data which can be complicated and time consuming.
- 7.4.5 Within National Grid we will be creating a regulatory compliant Mesh of Connected Data Products.

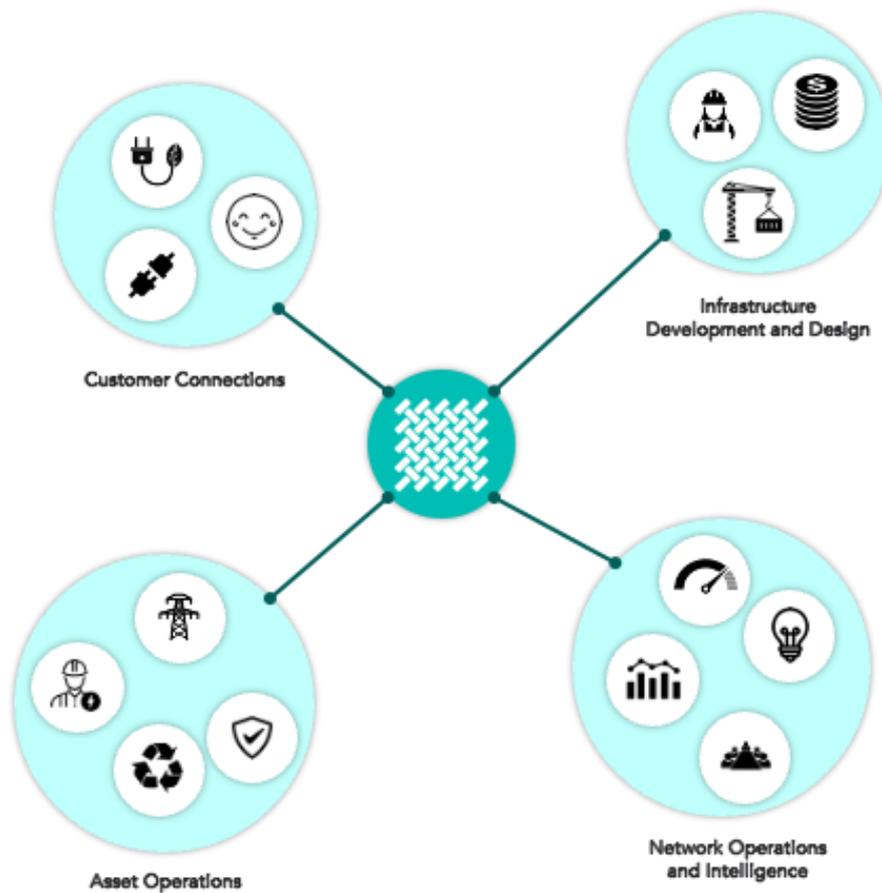


Fig 7 (Example National Grid Electricity Transmission Data Products Fabric)

- 7.4.6 As part of the Data Product Accelerator, funding lines will allow Data Product Teams within NGET to create Data Products from Systems of Record and from other existing Data Products, this allows the teams to develop new Data Products faster because the integrations are already in-place, and we get agglomeration effects from them as more products are built.
- 7.4.7 Through the Data Product Accelerator, Data Products will be developed by the teams within the business to support numerous existing programmes.
- 7.4.8 Data Products provide the reusable baselines for onward use in Digital Products and Platforms. This is a shift in our current delivery model which looks simply to provide data to a digital product or platform on case-by-case basis where data satisfies a single object, the mesh and fabric simplify this making existing products easier to find and evolve reducing effort over time.
- 7.4.9 This allows us to apply “technology whole system” thinking to Data, rather than data being specific to a project. This will bring value to consumers by high levels of re-usability where one Data Product may fulfil several business needs, as the Data Mesh scales it becomes faster to build Digital Products on existing Data Products.

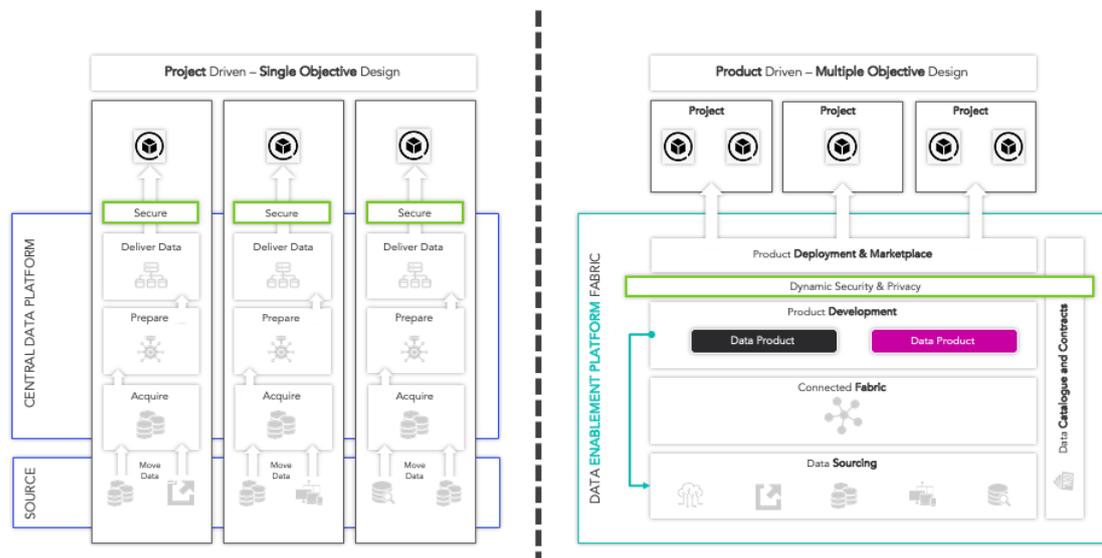


Fig 8 (Pivot to re-usable product model for single object data efforts)

Figure xx – project vs product thinking approach

From left to right, project driven data practices represent the as is process for most organisations today, moving right we see the shift being undertaken within NGET to move to reusable data products supporting multiple outcomes, connected via a data fabric.

- 7.4.10 The Data Products will be developed following National Grid’s Agile Model which uses SAFe (Scaled Agile Framework). The Data Product Managers in the business Product Line teams will be responsible for creating a backlog of Data Products, they will prioritise and sequence the work based upon assessment of value and reach in line with the SAFe Framework. This determines the economic value of delaying the work divided by the effort required to deliver the work. This approach ensures that the work with the most consumer value is undertaken, whilst providing the agility to meet the changing needs of our customers and stakeholders.

7.4.11 Some examples of programmes of work and already identified Data Products are below:

**Product:** Asset Foundation

**Description:** Universal attributes of assets like installation date, years in service, manufacturer, manufacturer's model name, commissioning status, identification numbers.

**Reuse Value:** Wide range of use cases all avoiding rework transforming source system data and using consistent data: RRP cost and volumes and NARM submission, EAM migration, SCADA migration. Additionally, we can minimise the impact of the EAM migration to make the consumption of asset data core system agnostic.

**Product:** SF6

**Description:** Complete list of NGET assets that contain SF6 insulating gas, the history of top-ups to replace gas leaked to the atmosphere and a forecast of expected leakage in future.

**Reuse Value:** Investment decision making and optioneering, long term strategic planning for SF6 inventory reduction, regulatory reporting (RRP table E1.4), internal monitoring of performance against SF6 incentives, potential to use externally for benchmarking with other transmission operators.

**Product:** Wound Plant Asset Health

**Description:** Collation of asset health data used to assess an asset's probability of failure including dissolved gas analysis, historical oil leaks, post-mortem data from other assets.

**Reuse Value:** Improving the annual asset health review process and availability of that data. Hubble for understanding flexibility of planned intervention, RRP/Price Control planning to support the development of engineering justification reports and technical appendices.

**Product:** Site Information

**Description:** A list of sites and key attributes like site code, location, voltage.

**Reuse Value:** Very widely reusable – immediately for Asset Foundational data to connect site workforce data to work, for SCADA to get site location information, for intelligent planning to support redistribution of staff to deliver work, for solution lab in customer connections.

**Product:** Investment Portfolio

**Description:** Detailed information on the Investment Portfolio including Investments and details of associated Projects at varying levels of detail.

**Reuse Value:** Wide range of use cases all avoiding rework transforming source system data and using consistent data: Regulation's RRP Data Quality dashboard, Composer tool for RRP, Capex data product, Project GIS Link data product, Hubble digital product, Finance's Cost data product, and more expected.

7.4.12 Further candidate Data Products are listed below, these are further back in the backlog for development.

- Performance
- Full set of Regulatory Reports
- System connectivity – Which will provide a map of connected Assets.
- Digital map (asset geography) – Asset by location and relationship.
- Asset Outage Impact Model – Model outages and outage windows across the transmission network.
- Outage Model – Upcoming outages.
- Nodal system forecast – Voltage forecasting across the transmission network.
- Climate risk model – A Geospatial model which looks at predicted climate change impacts to NGET assets.

- Investment Portfolio – Providing structured data model for investments, projects, and cost factors, enabling accurate and repeatable visibility of our network investments.
- Weather Data Products – Provides weather forecast data and actual recorded weather data. With other Data Products to potentially provide the ability to adjust the capacity of the grid based on asset class, age, and weather data. This may be an external data product.

## 7.5 Product Strategy Examples: RRP Cost and Volumes reporting and Mesh

- 7.5.1 NGET implemented a series of Data Products to reduce the time taken to create the Cost and Volumes component of the Regulatory Reporting Pack (RRP), this allowed us to test the Data Mesh and Product approach of our Data Strategy.
- 7.5.2 Before developing the Costs and Volumes Data Product, creating the Regulatory Reporting Pack (RRP) would rely on manual processes to achieve the results. The Data Product approach has allowed us to collect, consolidate and transform the data required for RRP much faster than before and automate the flow of data. We are better able to identify and rectify data quality issues in the process enabling us to improve processes to further improve the value of the data.
- 7.5.3 The Data Products and created to be interoperable, and through the Data Fabric they are available within NGET, National Grid, and externally to consumers. The example of Cost and Volumes RRP shows this, currently data is templated within manual methods, but if there was an industry product which adopted the same model, metadata, and content, it would provide the ability to connect whole industry data together and aggregate and interrogate it in a consolidated manner. The introduction of industry standard product models would greatly accelerate this capability and improve interoperability and industry wide aggregation.

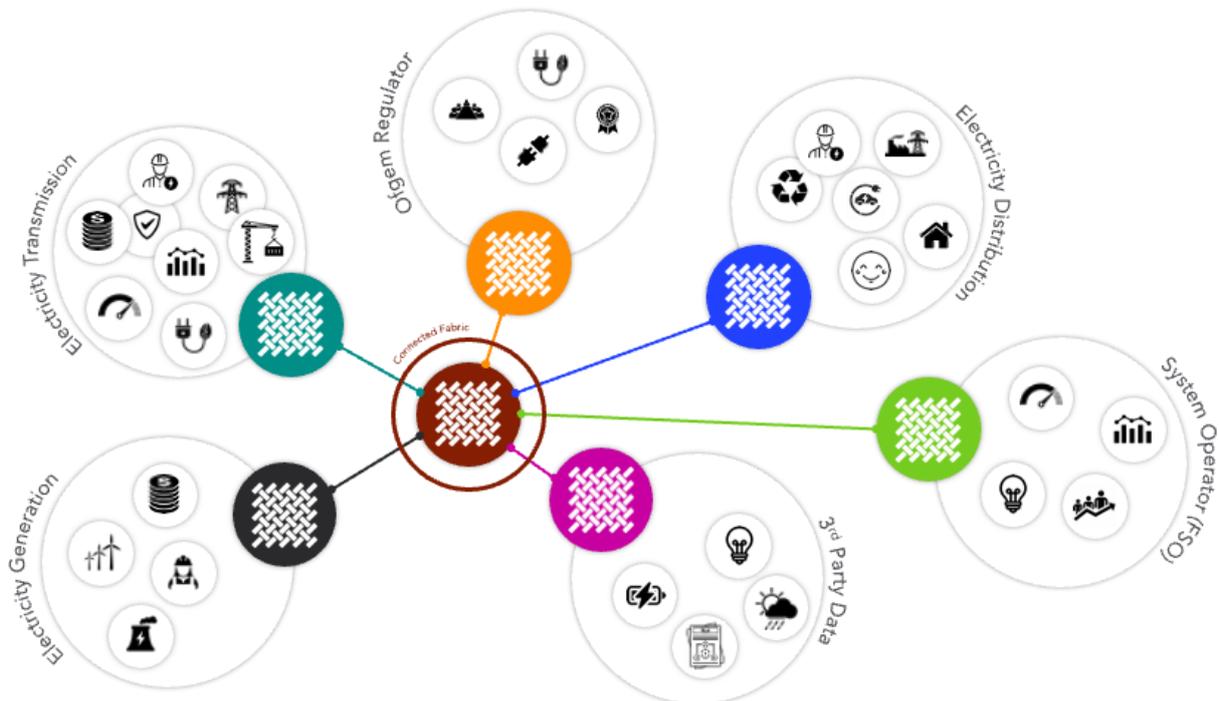


Fig 9 (Example Energy Industry Data Products)

- 7.5.4 This approach will allow each stakeholder to develop the Data Products that meet their business needs, whilst allowing these to develop the richness of the data products of the organisation and the product then being sharable and interoperable across the industry, without having to change the source data or data format.
- 7.5.5 This approach increases agility and speed of delivery of Data Products but could also create the opportunity for industry standard products accelerating regulatory reporting, planning and many other activities which currently require highly manual co-ordination activities.
- 7.5.6 The use of products and the associated catalogues, contracts and connectivity fabric enables NGET to react to regulatory standards and market changes, react faster to shifts in demands and information requests, provides easy adoption of Open Data principles, and will also allow us to easily feed future industry activities such as the digital spine or the strategic plan.

## 8 Project Justification: Engaging our Stakeholders and “Technology Whole System” thinking.

### 8.1 Stakeholder overview

- 8.1.1 In preparing this investment engagement has been undertaken with both internal and external stakeholders. We brought together different departments within NGET with the ENA (Energy Network Association Digital and Data Steering Committee) to share and review the NGET data strategy. Ofgem’s data team were also engaged in the development of the data strategy to ensure we fully adopted the data best practices. The preliminary Data Strategy implementation in 2022 ensured that we had the right level of engagement across NGET and the wider NG group, we also ensured that the NGET data strategy had alignment with the evolution IT strategy.
- 8.1.2 The RFI for the Data Best Practice improvement plan became a catalyst for reflection and with the change of Digital and Data leadership in the organisation setting out the most recent data strategy opportunity and intent has accelerated the development and work on what is needed to support continued compliance and accelerate the wider benefits of the platform and strategy. Internally we focussed on identifying, discussing, and developing the data capability and architecture framework needed to support product, organisational and capability development across NGET and meet the needs.
- 8.1.3 We have a regular governance and meeting cadence with our internal stakeholders, Product Line Directors (via our Digital Guild) and the Data Product Managers (via our Data Guild). In preparation for this submission, we also held sessions with Ofgem, firstly on the Improvement Plan and the April RFI response and latterly on the reopener submission to provide context to the ask and opportunities.

#### 8.1.4 Stakeholder discussions

Stakeholder	When	Engagement objectives and discussion points
Internal Executive	Ongoing	Achieve Executive buy in and sponsorship for the NGET data strategy, compliance to DBP, ongoing capabilities and organisational construction and roles and responsibilities in the development of the data fabric and product portfolio
Data Teams	Ongoing	Ongoing requirements and feedback loops for data product teams to inform and ensure the technical and capabilities roadmaps represent the needs. Ongoing governance, portfolio management and quality measurements in the creation of the data product portfolio
Digital Guild	Ongoing	Alignment of Platform and Product strategy in line with digital portfolio
Data Guild	Ongoing	Governance setting and design authority for platform and product strategy
Independent User Group	Ongoing	What the users, require from our data and their wider industry requirements.
Ofgem	Monthly	Over the course of the year, we have had several sessions to discuss the RFI on Data Best Practices, the RFI response and outcomes, Data Best Practice Improvement Plans, Data Fabric Infrastructure and Data Platform Reopener and to gain guidance and understanding of expectations and changes to best practices and new standards like Open data / Dublin core.
Energy Network Association Steerco	Monthly	Strategic Alignment, Data Best Practice Compliance, Architecture, Platforms, Interoperability, Data Products, and the wider alignment across the industry.

Table 3 (Stakeholder Engagement)

## 8.2 “Technology Whole System” thinking

8.2.1 Achieving a resilient secure energy transmission network at the lowest possible cost for consumers requires a “Technology Whole System” approach to be taken across industry. While data best practices are a unifying set of principles, deeper discussions on data as product, common products and open data are the mechanisms to bring this all together. Through forums such as the ENA data and digital steerco, it is our belief that this alignment will become stronger, and the use of common practices and open data expectations will bring the industry closer together. The NGET Data strategy and architecture can play a key role in leading this development.

## 8.3 Ofgem

8.3.1 NGET has held socialisation sessions with Ofgem colleagues on the investment since 2021. In March of 2023 an RFI on data best practice was submitted by Ofgem which NGET responded to achieving compliance, but with improvements identified and a marker for additional requirements previously unaccounted for, this resulted in the improvement plan shown below and the marker relating to this submission.

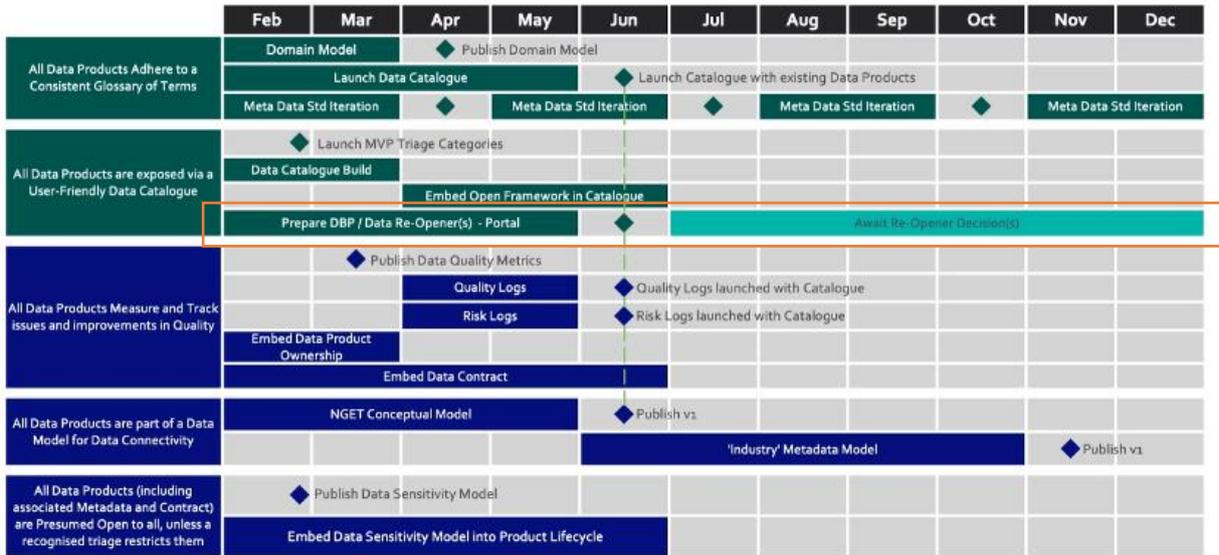


Fig 10 (Data Improvement Plan agreed with Ofgem including DBP/Data Portal Reopener)

## 9 Project Justification: Optioneering

### 9.1 Data Fabric Options considered.

9.1.1 The team undertook a review of Data Fabric options available and focused on three broad approaches were considered.

- 1) **Buy**, and integrate SaaS (Software as a service) Data Fabric
- 2) **Hybrid**, separate open-source Portal Data management system
- 3) **Build**, an internally developed solution

### 9.2 Buy an integrated SaaS Data Fabric

9.2.1 The proposed integrated technology stack provides an integrated, enterprise-ready solution to enable accelerated delivery of secured and governed data products. By leveraging the capabilities of all three platforms, NGET can achieve rapid agile development without compromising on security, governance, or the customer experience.

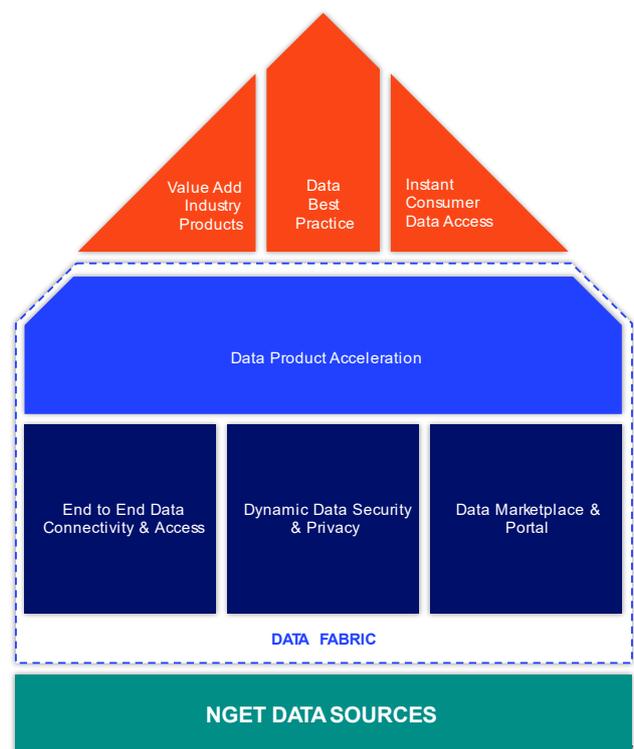
9.2.2 The Data Connectivity provides a high-performance SQL engine and distributed querying layer that allows us to query data across our entire data landscape regardless of source. This removes delays caused by having to move and consolidate data to build products. The Data Connectivity will enable our developers to query and combine data from across data stores and data lakes to quickly build new products.

9.2.3 The ABAC solution provides fine-grained access controls and security on top of the data sources. NGET can define policies to strictly control access to sensitive data, ensuring we comply with Ofgem data protection requirements. ABAC also enables us to provide different consumer groups with access to different views of the data based on their permissions and needs.

9.2.4 The Data Catalogue gives us persona aware access to Data Products, which are tailored to different internal teams and external stakeholder groups. Data producers, data consumers, and data stewards will each get personalised views of the catalogue that expose the Data Products relevant to them. This improves discoverability and the self-service capabilities of the platform.

9.2.5 By combining these three technologies, we get an agile, scalable, and unified data access platform that still allows us to apply the right level of security, governance, and customization. Tight integration between the products via pre-built connectors will reduce implementation overhead. While the licensing costs are higher than an open-source option, the benefits of faster development velocity, better security, and higher adoption by internal/external data consumers will provide significant value.

9.2.6 In summary, this recommended architecture meets all the stated requirements:



- Agile development: enabled by the data fabric distributed queries.
- Customized views: provided by persona aware, tailored data catalogues.
- Strong security: enforced by ABAC policies and access controls.
- Unified data access: delivered via the integrated platform.
- Follows Ofgem best practices.
- Complies with data governance needs.

### 9.3 Option two: Hybrid Portal Data management system

9.3.1 This solution proposes to use an open-source solution such as CKAN, which is widely used to provide portal services and provides a flexible and lower cost option to deliver a unified data access platform. As an open-source system, would allow NGET to build customised Data Portal. Open-source portals provide an extensible open-source framework for building open data portals and data management platforms. NGET can leverage the standard features within these tools like data catalogue, APIs (Application Programming Interfaces), visualisation, and access controls.

9.3.2 With an open-source portal, NGET would avoid the licensing costs of commercial solutions. This allows NGET to deliver a platform that meets Ofgem best practices for open data sharing in a cost-effective manner.

9.3.3 However, there are some downsides to the open-source approach:

- It does not have a Data Fabric, the lack of fabric and advanced security will mean that sourcing data, building Data Products will be much, much slower.
- The lack of the Fabric will impede federated governance because lineage and data quality will be hard to manage as data is moved from source systems to build Data Products.
- Agile development velocity may be slower compared to commercial solutions, since more custom development work is required for each new data product.
- While open-source portals have basic access control features, it lacks the robust, granular policy enforcement capabilities of a commercial solutions. Additional work would be needed to harden the platform from a security perspective.
- The data catalogue and self-service capabilities are also less sophisticated out-of-the-box compared to solutions that are purpose-built for ease of data discovery. Significant customisation would be required to achieve the desired consumer experiences.
- Performance and scalability may become issues for enterprise-scale deployments as open-source tooling is designed more for open data sharing than as a unified analytics data platform.
- Additional investments would still need to be made to provide a Data Catalogue for the System of Record because the open-source data portal does not provide this capability.

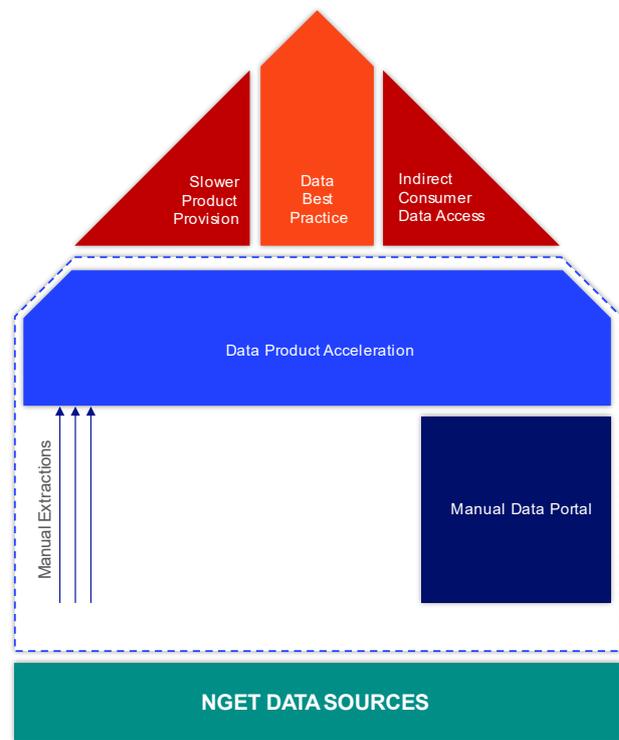


Fig 12 – Hybrid Data Portal

- While the flexibility and lower licensing costs of open-source portals are advantageous, the development overhead and lack of enterprise-grade features out-of-the-box make it higher risk compared to integrated commercial solutions. For our use case where agility, security, and consumer experience are critical, the commercial products are likely to deliver greater overall business value.

## 9.4 Option three: Build an internally developed solution.

9.4.1 Developing a bespoke solution for NGET would provide complete control and enable NGET to tailor the platform to our precise needs, the increased costs and delivery risks make this option less advisable compared to commercial or open-source alternatives.

9.4.2 An in-house build approach would allow NGET to design the solution to our specific use cases and user groups. NGET would not be constrained or limited by the functionality provided by vendor products. Our developers could build custom features, integrations, and experiences that fully support our business requirements.

9.4.3 This would also avoid the ongoing licensing fees associated with commercial solutions from vendors. An in-house build would provide full control over the product roadmap, allowing us to prioritise capabilities based on business value rather than being forced to upgrade on a vendor's timeline.

9.4.4 However, developing a production-ready platform with enterprise scale, security, reliability, and supportability comparable to leading vendor solutions would require a massive internal development effort. It could take years of sustained investment to achieve the robustness of commercial solutions that have decades of real-world hardening.

9.4.5 The opportunity cost of dedicating valuable developer resources to rebuilding existing capabilities also needs consideration. Focusing our teams on differentiating value-add business functionality, rather than core data management plumbing, is a better strategic use of resources.

9.4.6 An in-house build would also make it harder to take advantage of the latest innovations in the data management space, as commercial vendors invest heavily in Research and Development. We would constantly need to assess new technologies and standards to keep pace ourselves. Staying on top of industry best practices around governance, metadata management and security would be an ongoing challenge.

9.4.7 Whilst a custom in-house build offers ultimate flexibility, the substantial increase in costs, risks and delivery timeline make it difficult to justify over leveraging proven commercial and open-source technologies. We must weigh the benefits of complete control versus faster time-to-value, reduced development overhead and risk mitigation provided by robust vendor solutions and platforms.

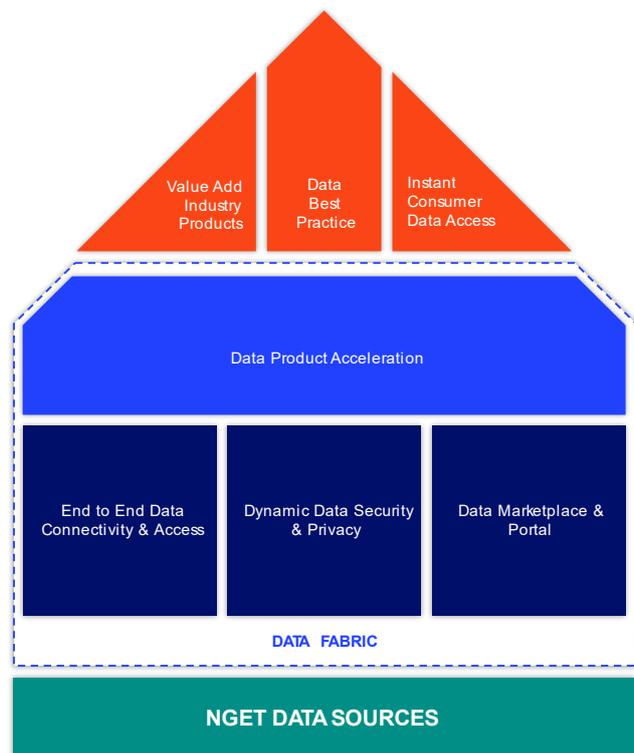


Figure 13 Build Data Portal

## 9.5 Options Review and Preferred Option

Solution Option	Development Cost	Time to Deploy	Scalability	Security	Customisability	Total Score
<b>Option 1 Buy</b>	5	5	3	5	3	<b>21</b>
<b>Option 2 Hybrid</b>	4	3	3	3	4	<b>17</b>
<b>Option 3 Build</b>	1	1	4	3	5	<b>14</b>

Table 4 (Optioneering Scores)

### Criteria Rating Scale:

Development Cost – The estimated cost of development resources and tools needed to deliver the solution. Includes internal development labour as well as external contractor costs. Lower is better.

Time to deploy – An estimate of how long it would take from project kick-off to having an initial production-ready version (MVP (Minimal Viable Products)) available. Shorter time to market is preferred.

Scalability – The ability of the solution to scale up in terms of data volume, users, transactions etc. to meet projected growth for the foreseeable future. Higher scalability is better.

Security – The degree to which industry best practices for data security, access controls, and governance can be met with the solution. Higher security is preferred.

Customisability – The level of effort required to customize and extend the platform to meet additional specialized requirements beyond core functionality. Higher customisability is better.

### Scoring Scale:

5 – Excellent

4 – Good

3 – Average

2 – Poor

1 – Very Poor

## 9.6 Summary and selected option

- 9.6.1 Option 1 scores very well on development cost, time to market, and enterprise scalability by leveraging mature commercial solutions. But has less flexibility than open source or custom options.
- 9.6.2 Option 2 reduces licensing costs and allows more control over customisation, but at the expense of enterprise-grade features and longer development cycles. This risks delivery of NGETs strategic outcomes and adherence to Data Best Practice because the infrastructure will not support the needs of the business and the manually processes, risk data quality and governance.
- 9.6.3 Option 3 gives full customisability but at an extremely high cost in terms of development resources, delivery timeline, and scalability.
- 9.6.4 **Option 1** is the recommended approach given its strengths in minimising development costs and time to market while still providing robust enterprise capabilities out-of-the-box.

## 10 Cost Benefit Analysis

### 10.1 Overview to the CBA (Cost Benefit Analysis)

10.1.1 The funding requested in the re-opener is for the years 2024 and 2025. The programme is intended to continue through 2026 and beyond and will form part of the investment plan for the next NGET price control.

10.1.2 The CBA considers the short-listed options derived from optioneering and these are listed here:

- Do nothing (base line scenario).
- Buy – which is the preferred option.
- Hybrid
- Build

### 10.2 Do Nothing (baseline scenario)

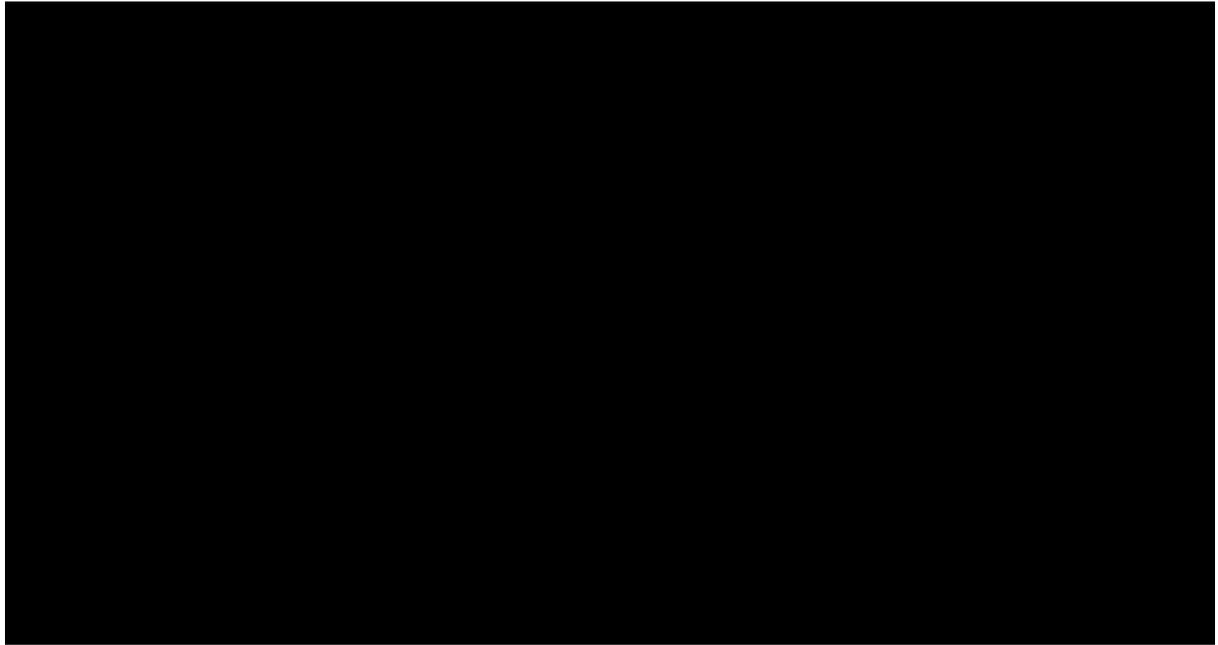
10.2.1 This option is the base case option and assumes that no work is undertaken and assumes there are no future developments to Data and there are no future benefits to the consumer. Given that the improvement plan and best practices called for these additional components, we have assumed that NGET would receive a fine and would need to remediate this and set aside a net cost of £1M. Practically this option is rejected because it does not provide any benefits and goes against the commitments to open data and transparency, as such places NGET at risk of not fulfilling its licence conditions.

### 10.3 Buy

10.3.1 This option proposes a fully integrated data fabric and portal, this option fulfils all the needs of the Data Best practice resolving the risk to NGET's.

10.3.2 The initial investment in the SaaS solution is higher than the Hybrid option, but because the integrated solution allows NGET to develop Data Products faster, with a lower unit cost for both the build and on-going support. *We have been conservative in our application benefits against supplier estimates of 90%+ speed to insight, and 50% lower costs (these are outlined in the CBA in appendix 1).*

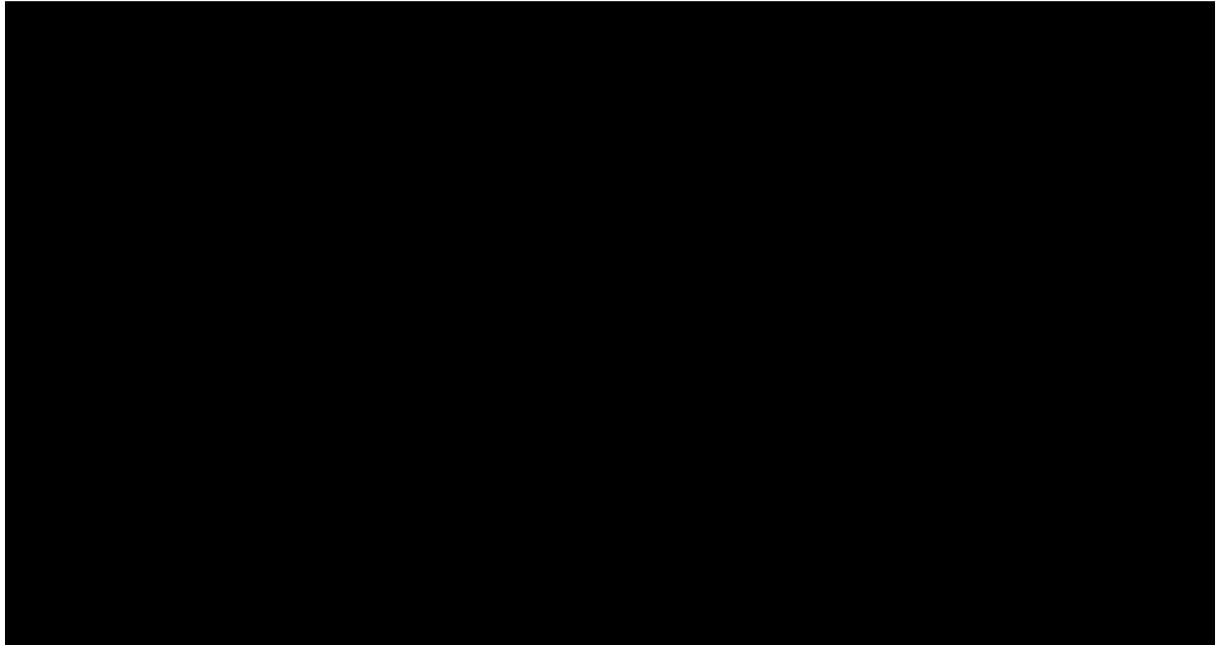
10.3.3 Using SaaS tooling, will allow NGET to build and deliver the solution faster than either the Hybrid or Build options, allowing NGET to deliver Data Products and their benefits earlier than the Hybrid and Invest options.



*Table 5 (Buy option)*

10.4 Hybrid

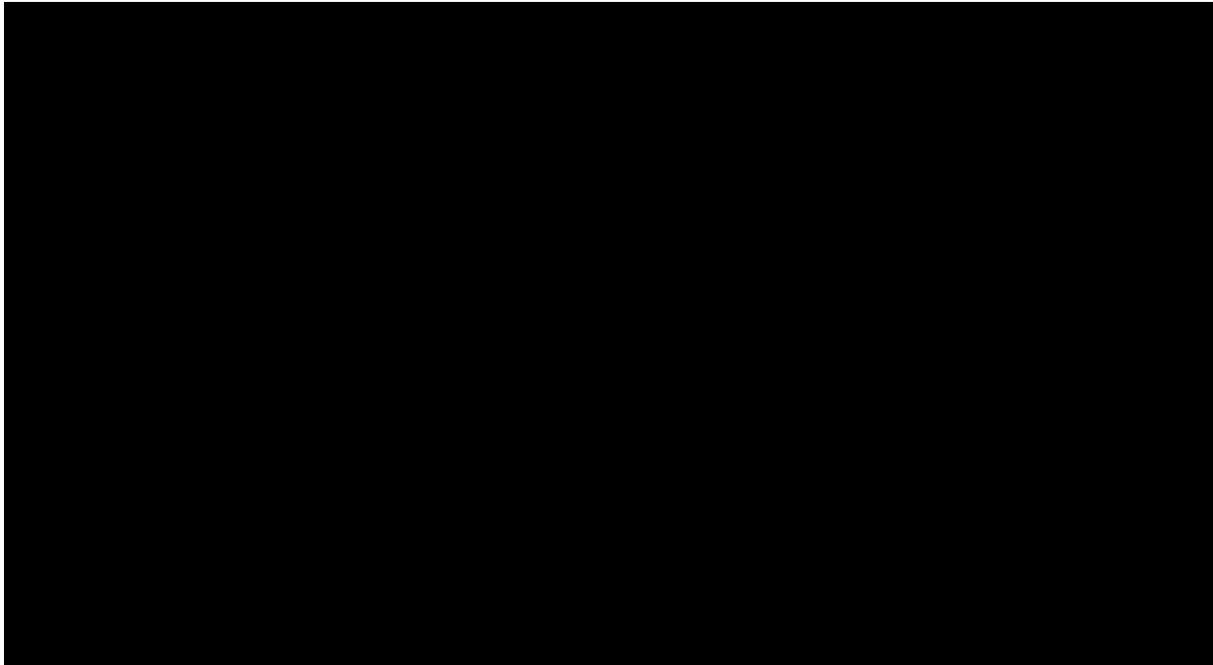
- 10.4.1 This option proposes to implement a separate opensource portal. The Data Portal allows us to share Data Products with external customers, but it lacks the automation of a fabric, and the controls of the ABAC solution. It does not remove the constraints of NGET’s existing Data Warehouse which makes the process of fulfilling Ofgem Data Best Practice a manual task with the inherent risks that this brings.
- 10.4.2 The initial investment in the portal is lower, but the lack of supplier support during the design and build phases we estimate that this will take longer to implement. Without the automation capabilities the time, cost and on-going support costs for each Data Product will remain the same. Overall, this option provides fewer benefits versus the costs than the Buy option.



*Table 6 (Hybrid option)*

## 10.5 Build

- 10.5.1 This option proposes that NGET develops the software to create an integrated portal, fabric, and an attribute-based security and access solution (or dynamic security). The option is provided for completeness, but it is not recommended because it is also contrary to our buy not build design preference.
- 10.5.2 This option doesn't represent high consumer value and carries significant risk because NGET is not a software development organisation, and these capabilities are complex to design and implement. The complexity to leads to a higher build cost and implementation time, which leads to the benefits in terms of data product development time, lower unit cost for both builds and on-going support arriving later than the Buy option. Once delivered we have assumed the same benefits as the Buy option once delivered.



*Table 7 (Build option)*

## 10.6 Summary

- 10.6.1 The recommended option is to Buy a fully integrated Data Fabric. This option is the best choice because it fulfils the needs of the enhanced Data Best Practice, allows NGET to develop Data Products faster, has a lower unit cost, and can be implemented quicker than both the Hybrid or Build options.
- 10.6.2 Specifically, the Buy option will allow NGET to fulfil its licence obligations, improve trust through the transparency, deliver value to the business and wider energy industry faster, and provide leadership on using data to achieve our Net Zero goals.

## 11 Project Definition: Delivery Strategy

### 11.1 Scope of Programme

11.1.1 The programme of work is split into two parts, the Data Fabric, and Data Products Accelerator.

### 11.2 Delivery Methodology

11.2.1 The Data Fabric will use a hybrid approach of a traditional waterfall methodology to manage the procurement, solution design, and solution build phases. Then, the programme will move to Agile for the development and enhancement phase.

11.2.2 The Data Product Accelerator will use the Agile methodology throughout.

11.2.3 This is because the Data Fabric is cyber-physical asset, which needs to be integrated into the wider NGET IT estate, whereas the Data Products are built upon the Data Platform, and use the existing integrations provided by the data platform.

### 11.3 The Data Fabric

The Data Fabric has two broad phases, firstly the initial build phase where the infrastructure is designed and built. Then the enduring phase where the solution iteratively developed and maintained.

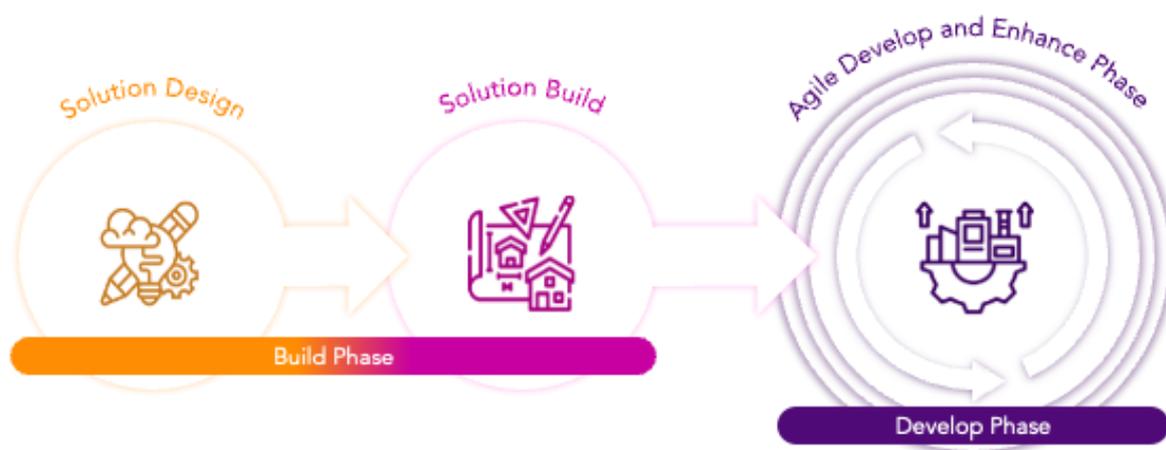


Fig 14 Describes the two phases of the Data Fabric.

11.3.1 The Build Phase of the Data Fabric are set out below.



Fig 15 Describes the build phases of the Data Fabric.

### 11.4 Key Milestones Data Fabric Build.

11.4.1 \*The dates in the build are predicated on the assumption that funds are released by December 2023.

11.4.2 The key programme milestones for this investment are:

#### 1) Sprint Zero: January 2024\*

- This is where we create a product vision and basic product backlog and allows us to create an estimated milestones for product release. We will also include a design sprint at the beginning, because creating a lightweight functional and architectural design as part of the product vision helps minimise risks and aligns stakeholders.

## **2) Design and Procurement:** January to March 2024\*

As per our optioneering, the preferred option is to buy, preferring SaaS, to PaaS to IaaS (Infrastructure as a Service). This aligns with the National Grid's IT Architecture Principles which lay out how we design solutions to maximise value, suitability, quality, and consistency. (Appendix 6 NGET Architecture Principles). The test strategy will be developed in parallel to the design, with the test plan to be created in time for the non-production build. During this phase the specific SaaS tooling integration will be agreed. We are already on the procurement strategy, a supplier qualification to shortlist the suppliers who could fulfil the requirements and will build the tender pack. Once the funding and sanction are approved, we will follow our Strategic Sourcing Process, to contract the SaaS services (section 13 Procurement and Sourcing Strategy)

## **3) Training:** March to April 2024\*

- Training materials will be created in this stage to roll out the Data Portal across NGET. The training strategy is a "train the trainer" approach upskilling local champions in each of the business delivery vehicles, so team members can maintain onward training within NGET departments in addition to disseminating training information and best practice guidance.

## **4) Non-Production Build:** April to July 2024\*

- During this phase the non-production environment for the Data Fabric will be built and tested. The security controls to ensure data Initial design assumptions will be tested, and the product design refined.

## **5) Design Update and Approval:** July 2024\*

- The design will be updated to reflect the feedback from the build of the non-production environment. This design will be approved at National Grid Electricity Transmission's Solution Design Authority.

## **6) Production Build:** Aug-Sept 2024\*

- The production environment is built during this phase. It will follow the approved design and use the code used to build the non-production environment. Testing during this phase will be extensive for availability, usability, and security.

## **7) Service Transition:** Sept-Oct 2024\*

- This is the period when the new Data Fabric shifts from build, to a Develop and Run. There is an enduring team, who will be part of the transition from the build team. During this phase the solution will be integrated into the wider service delivery framework (Helpdesk). The Develop and Run team, will continue to develop, and maintain the solution, during the products lifespan.

### 11.5 Data Fabric Agile Develop and Enhance phase.

- Agile Develop and Enhance: Oct 2024 – Dec 2025\* and onwards.
  - The SaaS technologies that underpin the Data Fabric are updated with minor and major updates on a regular basis. The updates will add new functionality and capabilities to the solution, this is providing additional business value. The develop and run team, will review updates, and productise them into the Data Fabric and wider Data Enablement Platform

- The Data Fabric must be integrated in to NGET’s IT environment, it therefore follows standard phases.

## 11.6 Data Product Accelerator

11.6.1 The Data Product Accelerator creates Data Products. Creating Data Products is like software development, where software engineers create software products, data engineers create Data Products.

11.6.2 We will apply the same development approaches to data products that are used in the software development.

11.6.3 The data products, follow the product lifecycle:

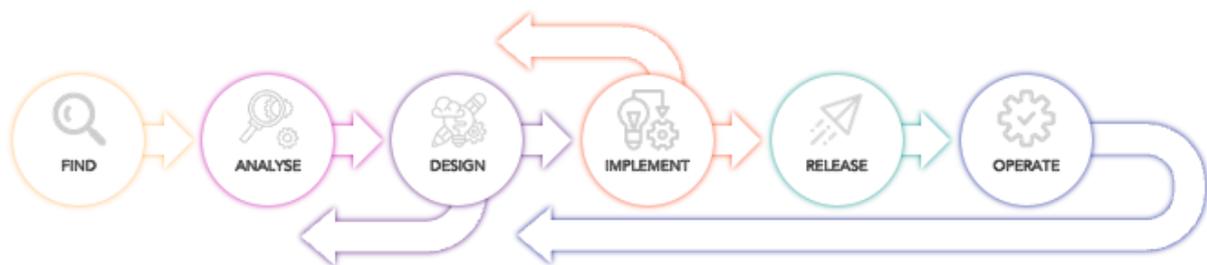


Fig 16 (Data Product Lifecycle High level)

11.6.4 The Data Product Accelerator enables Data Teams in each of the business Delivery Vehicles (DVs) to build the content (data products) needed to populate the fabric. This will be a managed and controlled process ensuring product builds provide the highest value to the consumer.

3. **Find** – each business Delivery Vehicles (NGET organisational groups) via the respective Product Line has a data team with a Product Manager who gathers, manages, and prioritises their customers’ needs. The Product Managers identify the business need and create candidate Data Products the Data Product backlog.
4. **Analyse** – The Product Managers works with the Product Owner and business SMEs in the Data Team to prioritise the items on the Data Product backlog. This is done by assessing which of the candidate Data Products delivers the highest potential value in the shortest delivery time. The outcome from this stage is to ensure we are delivering as maximum economic value. Once the highest value Data Product has been identified it is placed in the team’s backlog.
5. **Design** – The Data product is them worked on by the team, they will look to:
  - Define:
    - Data owners
    - Information Model (business logic and context of the product)
    - Functional requirements
    - Non-Functional Requirements
    - Data Quality Metrics
    - Metadata
    - Privacy requirements
  - Source Connection to the data
  - Create and validate the Data Product design.
6. **Implement** – The team will them implement the product.
  - Develop the Extract, Load and Transformation code to build the Data Product.

- Create and perform the functional testing, to assume the Data Product fulfils its Data Contract.
7. **Release** – We will follow continual development process, where the Data Product is released, and customers can use and provide feedback. The business will deploy the Data Product as time of their suiting, this will be via the standard change process.
  8. **Operate** – Once the Data Product has been deployed, it will be monitored to ensure that it fulfils its Data Contract, telemetry around usage, re-use (source of composite Data Products). Version updates to the Data Product may be incorporated but go through the Find and Analyse steps to determine whether this offers the best value.

11.6.5 Additional detail on the Data Product lifecycle can be found in Appendix 11.

11.6.6 This work of creating the Data Products is managed inside an Agile Release Train (ART), as per the SAFe framework. Each delivery vehicle operates an ART. To ensure engagement with stakeholders, and continued monitoring and learning, there are several key events that take place and are summarised below:

<b>EVENT</b>	<b>PURPOSE</b>	<b>PARTICIPANTS</b>
<b>QUARTERLY PLANNING</b>	Plan the work and outcomes aimed for in the quarter	ART team members Impacted stakeholders
<b>SPRINT PLANNING</b>	Planning the work for the sprint	Team Members
<b>SYSTEM DEMOS</b>	Review what has been delivered and invite feedback from stakeholders	ART team members Impacted stakeholders
<b>RETROSPECTIVES</b>	Reflection on the success of the previous quarter and decision on how to adapt to improve for next quarter	ART team members

*Table 8 (SAFe Stakeholder Engagement)*

Each of these events allow for a review of how well the Data Product meets business needs, and so allowing teams to adjust, if necessary. Through these frequent reviews and engagement with stakeholders we ensure the delivery of the technology stays in line with expectations and the outcomes required.

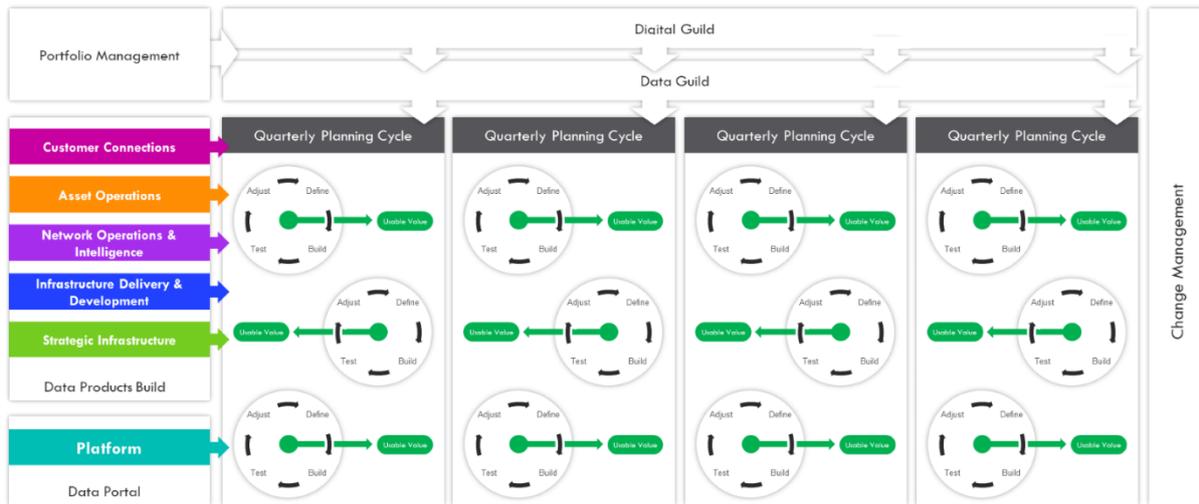


Fig 17 (SAFe Portfolio Planning and Governance)

## 11.7 Portfolio Management

- 11.7.1 The Digital Guild operates to provide Lean Portfolio Management, to ensure the work being undertaken across NGET aligns with the overarching NGET direction and strategy and the investments made contribute to and service that.
- 11.7.2 Such is the importance of Data; the Data Guild operates as an additional layer to review the Data portfolio across NGET in the same fashion with DV focus but a cross organisation context and overall outcome.

## 11.8 Dependencies

- 11.8.1 The design of the teams is to minimise dependencies by enabling each area to autonomously deliver. However, where dependencies exist, they are managed through the Portfolio Management function and through each ART. In each ART they ensure dependent areas are included in their quarterly planning and each team carries relevant tasks to support the other in their backlogs.

## 11.9 Business change and becoming a Data driven organisation.

- 11.9.1 The transformation to a Data driven organisation is integral to NGET realising the full potential of its Data Strategy. NGET will use our existing training processes and procedures to create new learning pathways for staff. We will work with our existing Data communities of practice (CoP) (Data Engineering, Data Science, Data Governance, and Visualisation) to identify the skills that our employees need, working with our training teams to create learning pathways that are tailored to their specific needs. This will involve a mix of formal training courses, on-the-job training, and self-paced learning resources, some of these will be provided by the suppliers of the SaaS tooling which the Data Portal is built upon.

11.9.2 We will make sure that the learning pathways are engaging and relevant to our employees. With a mixture of interactive content, real-world examples, and gamification of training. We will provide the opportunity through our CoP, and code café events we provide the opportunity for staff to practice their new skills. The pathways will allow staff to track their progress and find new opportunities to in data. Finally, we will celebrate their success through our regular company team updates and communities of practice.

## 12 Project Definition: Risks

### 12.1 Overview

12.1.1 The Data Fabric will be delivered via SAFe, which incorporates risk management into the core of the delivery methodology. There will be some risks which may be identified upfront, which will be outlined below. However new risks will occur during delivery and the SAFe approach to risk management will be used.

12.1.2 The Procurement will follow the Sustainable Sourcing Process (SSP) which has a Waterfall gated process. We will follow the procurement governance process. For clarity this the only Waterfall exception in the delivery approach, as a whole we will be following an Agile delivery methodology.

### 12.2 Approach to Risk Management

12.2.1 The SAFE (Scaled Agile Framework) approach to risk is that it is not a separate process, but continuously embedded throughout the agile delivery lifecycle. Early transparency, adaptation and iteration are the key to managing risk. The keyways this will be incorporated as part of the SAFe delivery methodology are:

- Continuous risk identification - Risks are identified continuously throughout projects, iteratively re-evaluating risks each sprint/iteration. This allows early risk detection.
- Embedded mitigation planning - Risk mitigation tactics are embedded into the actual development process through backlog management. They are not separate processes.
- Transparency via Kanban - Risks and mitigation status are visually tracked on Kanban boards to communicate progress and priority.
- Regular team-level assessment - Agile teams constantly assess risks in their iterations, supported by Scrum masters and product owners.
- Risk reviews - At higher levels, risks are regularly reviewed at the program and portfolio levels via governance meetings. Metrics are tracked.
- Risk cadence alignment - Risk activities are conducted at natural cadences aligned with team sprints, program increments, etc. rather than artificially separated processes.

12.2.2 Risks are categorised into four levels: high, significant, moderate, and low. These levels are determined based on the likelihood of the risk occurring and the potential impact it could have on the project or programme.

12.2.3 Risks are further categorised into types: Technical, Quality, Programmatic and Business. Examples of these are provided below.

- Technical Risks - These are risks related to technical execution within the project/software development.
  - Unproven or complex technologies
  - Integration issues with existing systems
  - Technical debt accumulation

- Infrastructure failures/deficiencies
- Performance issues
- Quality Risks - Risks impacting the overall quality of the solution.
  - Architecture shortcomings
  - Requirements gaps or defects.
  - Security vulnerabilities
  - Software defects and bugs
  - User experience flaws
- Programmatic Risks - Organizational and procedural risks.
  - Inadequate Skills and Staffing
  - Clarity on Roles & Responsibilities
  - Commitment of Executive Sponsors
  - Poor Management of Development Cycles
  - Insufficient Testing & QA Process
- Business Risks - Higher-level risks tied directly to business goals. Examples:
  - Lack of customer/user engagement
  - Market changes
  - Competitive threats
  - Legal/regulatory compliance
  - Insufficient ROI (Return on Investment)
  - Brand/reputational risks

12.2.4 Risk register is included in the Appendix.

### 12.3 Key Risks

KR	Category	Risk	Likelihood	Impact
KRT1	Technical	Delay due to complexity of integrating 3 new technologies into the existing Data Enablement Platform.	Yellow	Yellow
KRT2	Technical	Data Fabric layer is not performant due to integration with legacy systems of record.	Yellow	Yellow
KRT3	Technical	Meeting baseline security requirements adds complexity to the design, increases costs, and delay.	Yellow	Yellow
KRQ1	Quality	Initial Data Fabric and Data Product Catalogue functionality is low, due iterative delivery model.	Red	Green
KRQ2	Quality	Security risk in process, may lead to inadvertent publishing of controlled data.	Green	Red
KRQ3	Quality	Data Portal adds security risk as this provides access to NGET network.	Green	Red
KRQ4	Quality	Effort required to support and maintain is greater than assumed.	Green	Yellow
KRP1	Programmatic	Sourcing project team members with the requisite skills may be difficult.	Green	Yellow

<b>KRP2</b>	Programmatic	Time to stand-up the team, project tooling, etc. will take longer than anticipated.		
<b>KRP3</b>	Programmatic	Team will not have the requisite training required to fulfil install and operate the new Data Portal solution.		
<b>KRB1</b>	Business	Sourcing/procurement of technical solutions takes longer than anticipated.		
<b>KRB2</b>	Business	Can not source solution via existing routes to market.		
<b>KRB3</b>	Business	Incorrect publishing of data products leads to reputational risk.		
<b>KRB4</b>	Business	Assumption on Data Product T-shirt sizing may be wrong.		
<b>KRB5</b>	Business	Cost of SaaS software is greater than forecast.		
<b>KRB6</b>	Business	Cost of Hosting is greater than forecast.		

Table 9 (High Level Risks)

## 12.4 Risk Mitigation

KR	Risk	Mitigation
<b>KRT1</b>	Delay due to complexity of integrating 3 new technologies into the existing Data Enablement Platform.	During the scope of the procurements the need for the solutions to interoperate will be laid out in the requirements. We will specify use of common integration patterns such as REST APIs or event streams to provide an interoperability between components.
<b>KRT2</b>	Data Fabric layer is not performing due to integration with legacy systems of record.	Use existing Extract and Load solutions to move the data of the legacy system to a shared repository, which can be used to surface the data to the Fabric. System performance will be specified in the requirements and measured in delivery. Non-compliance with SLAs for performance will be managed with the service provider including service credits and/or rebates as per standard service management and quality practices.
<b>KRT3</b>	Meeting baseline security requirements adds complexity to the design, increases costs, and delay.	Engage security early in the process. As per normal procurement, ensure baseline security requirements are included. Pen testing and security validation will be continuously conducted throughout the data platform and product lifecycles.
<b>KRQ1</b>	Initial Data Portal and Data Product Catalogue functionality is low, due to iterative delivery model.	This is expected as the catalogue will be added to. Roll out plan and timely communications will need to manage expectations. Early engagement with Communications Team. Clarity in the catalogue on current data product

		quality and data product ownership with feedback into the product backlog to inform and prioritise product development
<b>KRQ2</b>	Security risk in process, may lead to inadvertent publishing of controlled data.	Clear testing scenarios for the ABAC product and adherence to baseline security requirements.
<b>KRQ3</b>	Data Portal adds security risk as this provides access to NGET network.	Adherence to baseline security requirements. Use of standard architecture patterns and penetration testing.
<b>KRQ4</b>	Effort required to support and maintain is greater than assumed.	Assumed effort is based on 3 products to provide the capabilities, assign risk cost allocation to CBA model.
<b>KRP1</b>	Sourcing project team members with the requisite skills may be difficult.	Early work on detailed job role and skills, engaging our PARTNER partners. Early communication with partners to begin sourcing of prospective candidates.
<b>KRP2</b>	Time to stand-up the team, project tooling, etc. will take longer than anticipated.	Early work on detailed job role and skills. Early communication with partners to begin sourcing of prospective candidates. Engagement with delivery teams to stand-up environments.
<b>KRP3</b>	Team will not have the requisite training required to fulfil install and operate the new Data Portal solution.	Include training as part of the procurement with suppliers.
<b>KRB1</b>	Sourcing/procurement of technical solutions takes longer than anticipated.	Early engagement with procurement and potential suppliers.
<b>KRB2</b>	Can not source solution via existing routes to market.	Check existing routes to market before commencement (done).
<b>KRB3</b>	Incorrect publishing of data products leads to reputational risk.	Build in data quality and risks in the continuous deployment process.
<b>KRB4</b>	Assumption on Data Product T-shirt sizing may be wrong.	Reduce number of data products but maintain focus on those with the greatest value.
<b>KRB5</b>	Cost of SaaS software is greater than forecast.	Costs are currently based on unnegotiated list prices, working with procurement will hopefully lead to lower overall prices. We may be able to offset some increases in areas with decreases on others.
<b>KRB6</b>	Cost of Hosting is greater than forecast.	Costs are based on a forecast infrastructure requirement. Working with our suppliers to correctly right size the environment, will help control cost. Risk will set aside to cover any gap.

Table 10 (Risk Mitigations)

## 13 Procurement and Sourcing Strategy

### 13.1 Overview

13.1.1 The Data Fabric requires the sourcing of SaaS software and resources to build and implement the Data Fabric.

### 13.2 The Data Fabric SaaS software

13.2.1 The SaaS software will be procured via Sustainable Sourcing Process (SSP). This process has been developed to ensure a robust procurement process is undertaken. This process ensures that we deliver competitive value whilst also complying with the requirements of the Utilities Contract Regulations 2016 (UCR) which governs NGET's procurement activities. The UCR are designed to prevent contracts being awarded without having gone through a competitive process and to support the free movement of goods, services, people, and capital within the EU.

13.2.2 The SSP is modular, ensuring that we can modify our approach based on the spend, complexity, risk, marketplace and timescales of the requirement, to ensure that we drive the most effective outcome for the sourcing need. The eight-step process includes four governance gates that allow a steering group (including representatives from the businesses) to challenge outputs and to ensure that the event delivers best practice sourcing, fairness, transparency, value for money and robust contracts.

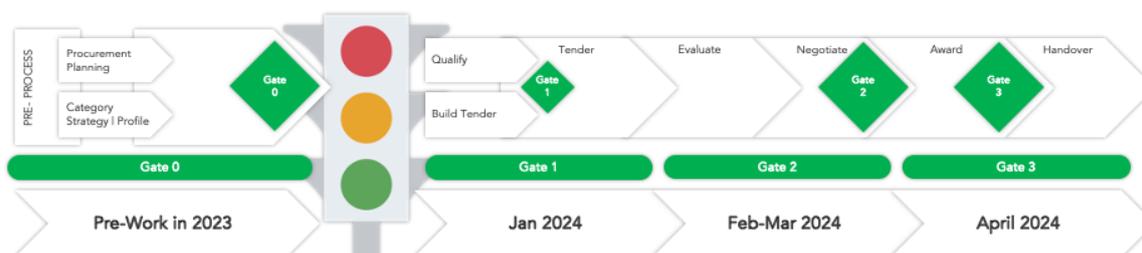


Figure 18 The high-level SSP process

13.2.3 The SSP process has eight stages :

1. Event Strategy – Outlines how the procurement event will be run, including the structure, e.g., set up a new Framework using appropriate procurement procedure available under the UCR or call off an existing Framework agreement, route to market, evaluation strategy, award criteria. Will utilise the relevant learning from the category strategies in place. Identifies what support will be needed from the business, i.e., individuals to be part of the tender team.
2. Qualification (Optional) – An option of shortlisting suppliers who meet the selection criteria.
3. Build Tender – Development of the tender pack and the evaluation process in line with the award criteria to differentiate tender returns.
4. Tender – Effective delivery of the tender in line with the agreed plan (taking account of the complexity of the event and necessary response time from Suppliers which may exceed the minimum UCR requirement), ensuring that business stakeholders and suppliers (including tender queries) are managed and communicated.
5. Evaluate – Completion of a thorough and consistent evaluation process, clearly documenting the outcomes to determine the shortlisting of tenderers.

6. Negotiate (Optional) – The assessment of tenderers against the contract award criteria, the conducting of negotiations covering both technical and commercial aspects and final evaluation.
7. Award – Notification of successful and unsuccessful tenderers and the management of the standstill period.
8. Handover – The facilitation of the formal handover from the SSP process to the Contract Management Process (CMP).

13.2.4 The SSP is a gate process with clear governance throughout.

- Gate 0 - Used as a formal gate to ensure the event strategy is agreed by all parties to ensure alignment before launch.
- GATE 1 (Optional) – Formal review milestone for qualification shortlisting and Tender documents.
- GATE 2 (Optional) – Used as a formal review point to shortlist tenderers for negotiation.
- GATE 3 – Formal recommendation on selection for contract award.

- 13.2.5 Determining the event strategy within the SSP process will vary based on factors such as related category strategies, the objectives of the tender, the associated risk, complexity, and scope. Based on this, we consider areas such as the agreement options, agreement structure, route to market, qualification strategy/options, native competition, and evaluation strategy to determine the appropriate strategy based on the risk associated with, and value of, the opportunity.
- 13.2.6 The use of SaaS products and the fact that there are a limited number of suppliers, reduces the scope, complexity, and risk for the SaaS components of the Data Fabric
- 13.2.7 Following the SSP process and subsequent alignment to the UCR ensures there is always an equal treatment of bidders and that a fair and competitive outcome is achieved and value for the consumer.

### 13.3 Resource Sourcing

- 13.3.1 Existing Data Enablement Platform Agile Release Trains will be used to provide the management and governance (section [Manage and Governing Delivery](#)). This approach ensures delivery interoperability to the wider portfolio of work being undertaken within NGET.
- 13.3.2 NGET staff will lead the solution build and we will use partner resources to augment the team during the build phase to flex the uplift in skills and volume needed. We will not use partner resources beyond the build and deployment. We plan to move to NGET staff for persistent ongoing development and run of the solution with no need to retain temporary resources beyond the build phase. We will look to have the NGET staff in place for transition as this ensures that NGET has the organisational memory for the solution.
- 13.3.3 The programme will source the build phase resources through our [REDACTED], contract as part of the partner framework. This is framework contracted by National Grid group and brings the groups size to bear, when achieving discounts.
- 13.3.4 There are several suppliers on the framework [REDACTED], [REDACTED], [REDACTED] and [REDACTED]. This reduces the risk of not being able to source suitable skills.
- 13.3.5 [REDACTED] is predicated on NGET owning the deliverables which marries with the delivery methodology.
- 13.3.6 The framework offers two models for sourcing resources.

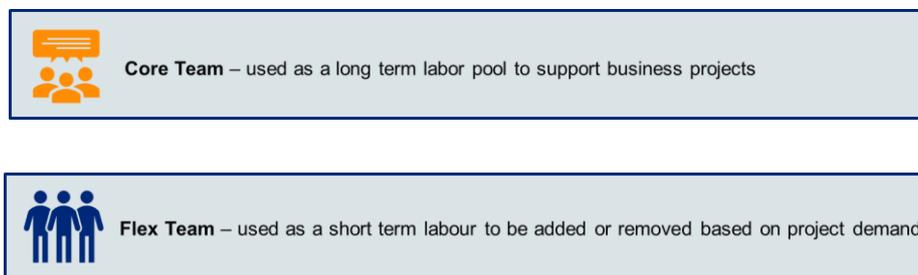


Fig 19 Core and Flex Team options

- 13.3.7 Flex Team is time and materials billed by the hour, whereas Core Team is billed per month.
- 13.3.8 We will use Core Team resources because we can predict our resource requirement. We will also lean towards offshore onshore mix of resource to further manage cost.
- 13.3.9 By using existing [REDACTED] contract through the partner framework, we can access the resources we need quickly, and with the confidence of an agreed framework sourced under competitive tender to assure customer value.

## 14 Detailed resource plan

14.1.1 Resource breakdown is broken to two sections one for the Data Fabric build and the other for the Data Products. The Data Fabric follows a more traditional design and build phases, whereas the Data Products have teams with a core skillset required to design and build the products, overtime additional team members are required to provide the on-going support and maintenance of the data products they create.

### 14.1.2 Data Fabric Resource Plan

Task	Role	Duration (Days)	Quantity
Discover and analyse	Business Analyst	25	2
Initial Design	Architect	25	2
Security Review	Security Architect	15	2
Architecture Review	Enterprise Architect	5	2
Procurement	Procurement Lead	45	2
Non-Production Build	IaC Engineer	30	4
CI/CD Integration	DevOps Engineer	30	4
Non-Production Source Integration	Data Engineer	45	4
Network Changes	Network Engineer	15	2
Non-Production Integration Test	Test Lead, Test Analyst	15	4
Design Update	Architect	15	2
Service Design	Service Architect	10	2
Architecture Final Approval	Enterprise Architect	5	2
Production Build	IaC Engineer	25	4
CI/CD Integration	DevOps Engineer	25	4
Production Integration	Data Engineer	40	4
Service Build	Service Designer	25	2
Production Integration Test	Test Lead	15	4
Production Security Test	Security Architect	5	2
Production Service Test	Test Lead, Test Analyst	8	4
Delivery Lead	Scrum Master	180	1
Change and Release	Service Manager	5	1

Table 11 (Resource Plan)

### 14.1.3 Data Product Resource Plan

NGET has already made the organisation changes moving staff from IT into the value aligned business units. Where possible the development of Data Products will be done by NGET staff, Data Products are owned by their Product Owner as per [Data Product Accelerator](#).

The table below shows the effort in person days to build a data product of a different size, the use of T shirt sizing for data product development is a standard practice and is impacted by many factors, for example establish connections resulting in lower set up time or evolution of an existing product

to meet the need over a new build. We have assumed that there will be a split in percentage of data products required as 70% small, 20% medium, 10% large.

Size	Number of Data Sources	Data Product Manager	Business Analyst	Data Engineering	Data Science	Data Visualisation	Total
Small	1	0.15	0.5	0.5	0.5	0.5	2.15
Medium	2 to 3	0.5	1	1	1	1	4.5
Large	3+	1	1	3	2	2	9

Table 12 (Data Product Build Resource Plan)

The table below shows the support and maintenance activity for each data product, the blended percentage for support and maintenance is 16.5%.

Size	Number of Data Sources	Data Product Manager	Business Analyst	Data Engineering	Data Science	Data Visualisation	Total
Small	1	0.1	0	0.1	0.1	0.1	0.4
Medium	2 to 3	0.2	0	0.2	0.2	0.2	0.8
Large	3+	0.3	0	0.3	0.3	0.3	1.2

Table 13 (Data Product Maintenance Resource Plan)

## 15 Cost Assurance

### 15.1 Overview

- 15.1.1 The Cost Assurance is for the preferred option however, the pricing has been used for both the Hybrid and Build scenarios. These are covered in the CBA template.
- 15.1.2 The cost analysis details the capital cost required to purchase the platforms and services needed for the implementation of the platform architecture, capability training and product acceleration. This takes the form of Capital Investment as it is the establishment costs which will deliver the platform and the ability to start delivering the value products on that foundation.
- 15.1.3 This section deals with the costs associated with the platform and content builds to meet the objectives of this investment.
- 15.1.4 The cost breakdown over 2 years for the Data Fabric and Data Products:

<b>CAPEX</b>	<b>OPEX</b>	<b>TOTEX</b>	<b>RISK</b>	<b>TOTAL</b>
£ 5,759,496	£ 512,398	£ 6,271,894	£ 424,132	£ 6,696,026

Table 13 (Total Investment Jan 2024-Jan 2026)

### 15.2 Cost Forecast

The cost will vary month to month because of the delivery phases. The chart below provides an estimated view of the spend profile over the period.

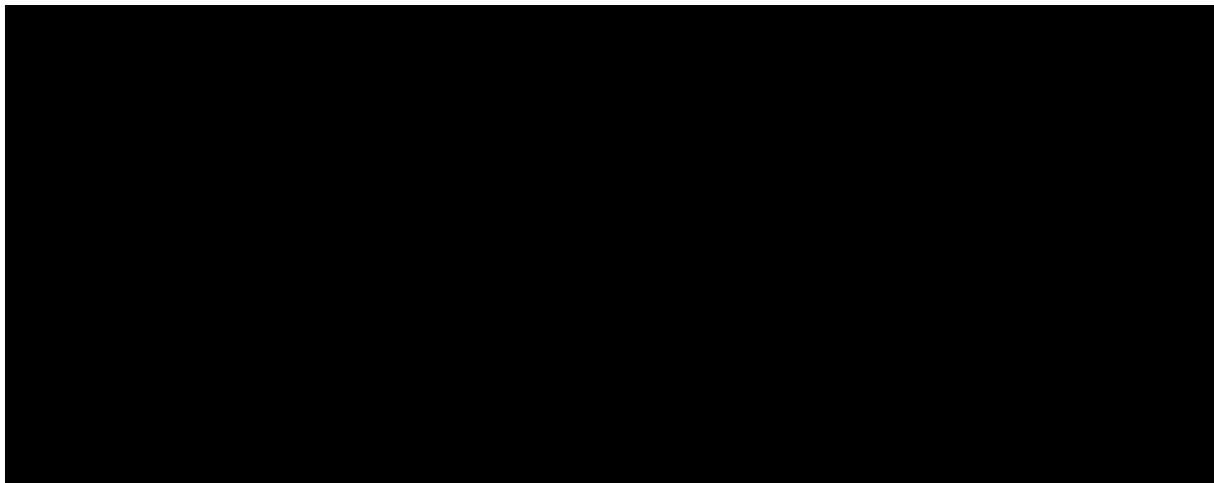


Table 14 (Monthly Spend)

This table is from the preferred option in the CBA, it shows the spend by month over the term of the investment.



*Fig 20 (Monthly Spend)*

### 15.3 Resource Costs

#### 15.3.1 Data Fabric Design and Build labour costs.

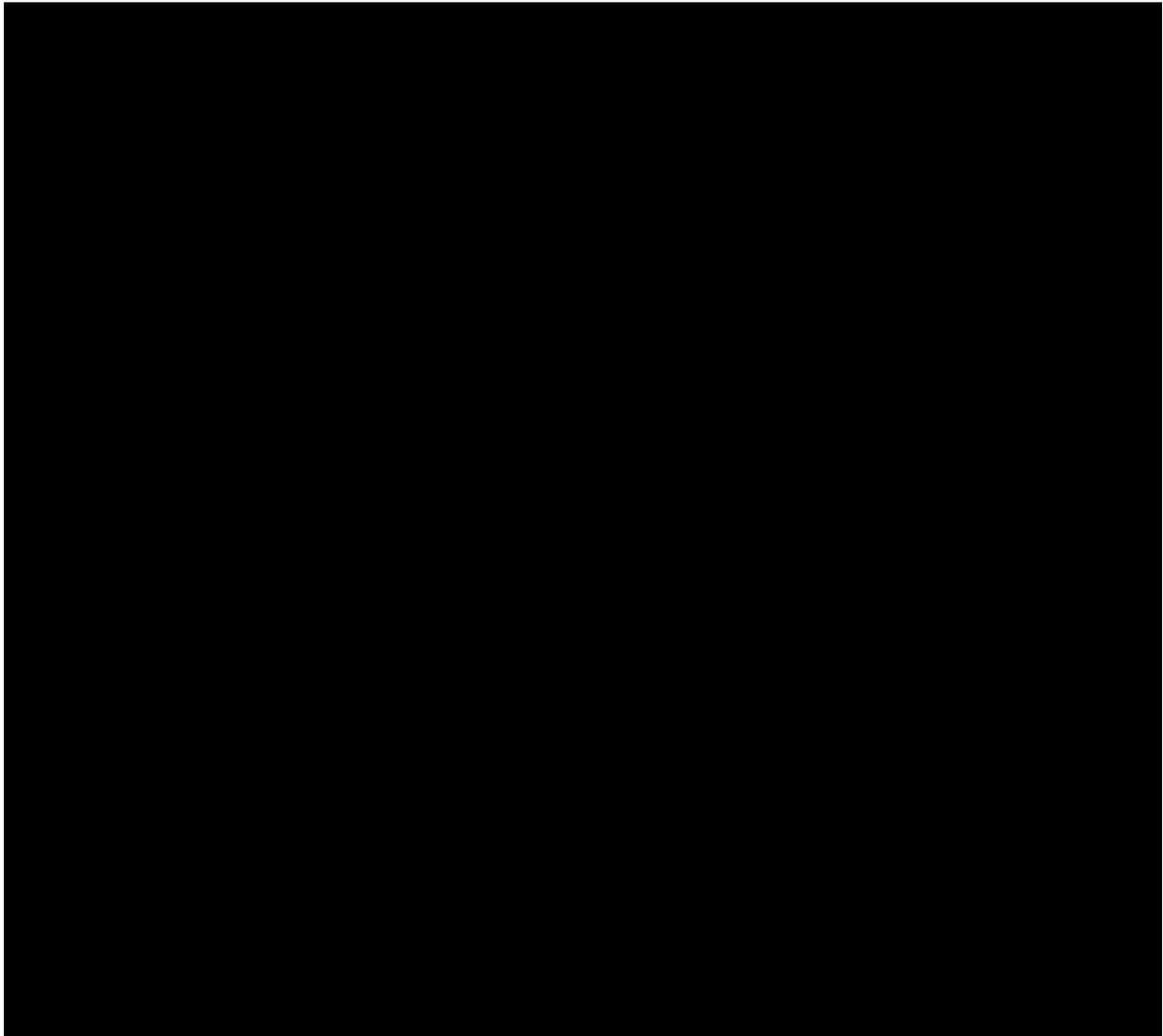
Table shows the cost break down per month for the design and build of the Data Fabric. Tables provides detailed breakdown of tasks, duration, and costs.

These resources are planned from PARTNER partners as they will in most cases be transitory and used during the build phase. For roles that persist we will endeavour for these to be NGET roles.

Data Portal Design resource costs.

*Table 15 (Resource Costs for Data Fabric Design)*

Data Fabric Build resource costs.

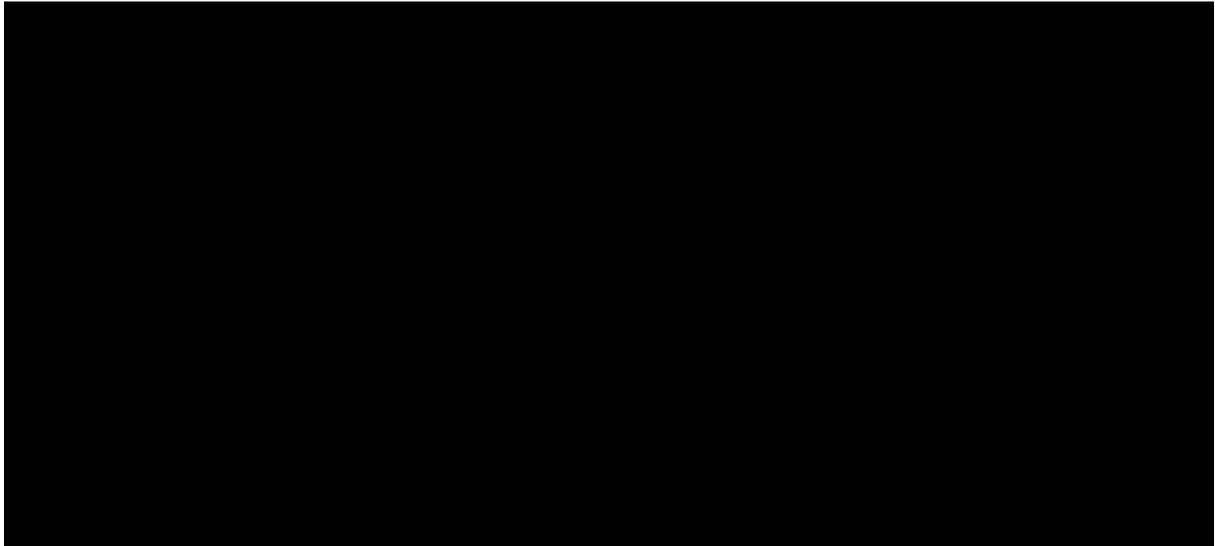


*Table 16 (Resource Costs for Data Fabric Design)*

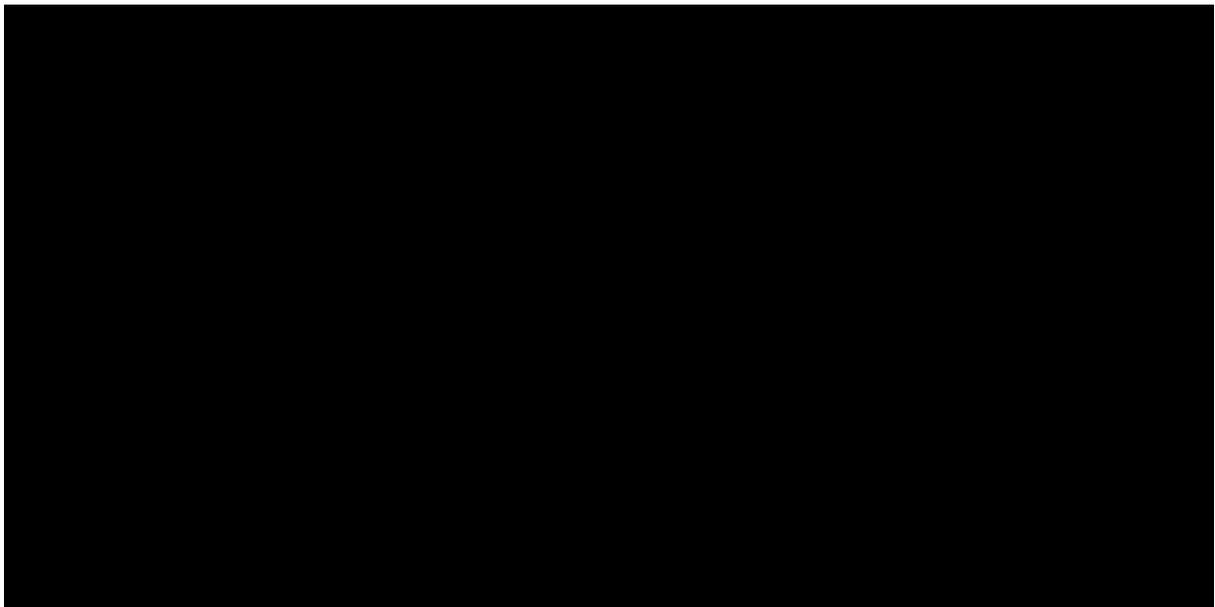
## 15.4 Overall Risk Allocation

Not all element of risk carries an associated cost. We have evaluated the identified high-level risks and mitigation in section [12.3](#)

The table and graph below provide the risk allocation over the programme period.



*Table 17 (Risk cost allocation by month)*



*Fig 21 (Risk cost allocation by month)*

**Risk cost allocation:**

The risks KRT1, KRT3, KRB1, and KRB2 identify risks which may cause delay to the design and build of the Data Portal. A risk allocation is added to the overall cost. The allocation is calculated by a percentage risk of slippage, with the risk reducing over each additional month. It is based on the cost of running the team, against the chance of slippage.

**Data Fabric Design Risk Cost allocation.**

Month	Risk of over run (assume 10% per month)	Cost of Design Team Month Cost	Cost of over run per month	Total Risk Allocation
1	95%			
2	85%			
3	75%			
4	65%			
5	55%			
6	45%			

Table 18 (Risk Model for Design Overrun)

**Data Fabric Build Risk Cost allocation.**

Month	Risk of over run (assume 10% per month)	Cost of Build Team Month Cost	Cost of over run per month	Total Risk Allocation
1	95%			
2	85%			
3	75%			
4	65%			
5	55%			
6	45%			

Table 19 (Risk Model for Build Overrun)

15.4.1 Data Fabric Run Risk Assumptions costing.

The risk KRQ4 calls out the assumption for the number of staff to support the Data Fabric is incorrect and is too low. The mitigation is to work with the suppliers to right size, but this has inherent risk as we have not yet engaged fully or procurement underway. The mitigation is to add additional to the risk pot, the cost of risk is an additional FTE (Full Time Equivalents) for the planned period of run.

15 months of a Platform Engineer [REDACTED]

Data Product Risk costs.

The risk KRB4 is that Data Product T-shirt sizing may be wrong, and we are not able to build as many Data Products as assumed. There is not risk cost allocated for this, the assumption is that fewer Data Products will be built and therefore the business benefit derived will be less.

15.5 SaaS and PaaS Costs

15.5.1 Licence costs

The Data Fabric is built using SaaS products, but their licencing models differ. The differing models, solution usage and growth are difficult to model. Example licence models are shown below.

Feature	Data Connection	Data Catalogue	ABAC
Number of licenses	vCPUs	Named Users or Capacity Licence	Seats or Capacity Licence
Cost per license	Varies by vCPU count	Varies by edition	Varies by edition

Pay-as-you-go	Yes	Yes	Yes
Flexibility	High	High	High
Scalability	High	High	High
Cost-effectiveness	High	High	High

Table 20 (SaaS Licence Models)

- 15.5.2 Each of the tools has a slightly different approach to licence and it may be financially beneficial to move from one model capacity to per named user or seat depending on the growth.
- 15.5.3 Some products are licensed per virtual core, and this can start small and grow with usage. There is however a downside in there is benefit in negotiating bulk discounts up front. We will need to work closely with procurement to get the contracting approach.
- 15.5.4 We conducted a market scanning exercise to determine the costs of the products, we called the top two suppliers in the area to get list price quotes. These are used in the pricing model.
- 15.5.5 Risk KRB5 identifies this risk and there is a suggest a mitigation. A risk pot of 10% of the total forecast SaaS licence costs, which is £[REDACTED] over the term.
- 15.5.6 Risk KRB6 covers hosting charges. There is a risk that Hosting requirements exceed what is forecast. This is because detailed technical and commercial conversations have not begun so sizing is an estimate. A 20% risk pot allocation to the Azure hosting costs over the period. This is £[REDACTED] over the term.

## 15.6 Training Costs

- 15.6.1 The Data Fabric adds three new technologies, the teams will need training. We have a made a forecast for what this will be, as part of the procurement from the suppliers the plan will be include training. We have assigned a forecast cost, there is an additional risk pot of £[REDACTED] which is 10% of the total forecast £[REDACTED] training costs.

## 16 Concluding Statement

### 16.1 Wrap up.

- 16.1.1 The objective of this submission is to reshape how we think about data, data access, the platforms that support us, the value we get from data and the improvement in transparency of what is available in line with improvement planning for data. We focus on continued compliance with best practices and their evolution, taking consultations and additions into account, and the application of those to achieve greater flexibility and value to consumer and use quality data securely, safely, and consistently.
- 16.1.2 We have considered the following during this submission: The need for change and investment, the requirements and value propositions, the direction and strategy which lead to the proposal, the opportunities, risks, and threats surrounding the submission, the delivery and outcomes from the investments and the range of benefits to the industry and consumer from this submission as well as the NGET organisation.
- 16.1.3 We are excited by the opportunity to lead the industry and support the evolving Data Best Practices, to deliver against this plan and welcome questions and further discussion on the submission and content.

## 17 Appendices

### 17.1.1 Appendix 12 - Assumptions

ID	Description
A1	Current Cost to create a data product may be extrapolated from existing costs
A2	Assume Data Fabric and ABAC get reduced Data Product support and maintenance costs with option 1 and 3 as they deliver the wider automation benefits. (Separate Portal, doesn't bring any automation benefits, there existing support costs continue).
A3	Assume current cost to produce a Data Product is £ [REDACTED] (Note EVN1)
A4	Assume cost to produce Data Product in future is £ [REDACTED] (based on Data Product Manager Feedback)
A5	Assume time to build Data Product currently is L = 1.5 Months, M = 1 Month, S = .5 Months
A6	Assume 50/50 split saving versus new value of Data Products
A7	Assume Agglomeration effects of 5% year on year for Data Products delivered via Fabric (Options 1 and 3)
A8	Assume 30 data products per year, based on feedback from Data Product Managers.
A9	Assume new Fabric (either Option 1 or 3) provide velocity acceleration of 10x from potential suppliers of products - Assume actual of 15%
A10	Assume 50/50 split between Staff cost savings and Business Value
A11	Products may be source via existing contracted mechanisms ([REDACTED])
A12	We can source resources without any delay
A13	Assumes licence costs incurred from first use at flat rate
A14	All costs are list/non negotiated for option 1 and 2
A15	Assumes SaaS licences can be capitalised because the services can run on-premises
A16	Assumes SaaS costs based on list price from top 2 providers in each category examples ([REDACTED], [REDACTED], [REDACTED], [REDACTED], [REDACTED], and [REDACTED]) - EVN5
A17	Assume Dev and Run is Pareto Principle split 20 maintain and 80 new value (Opex/Capex)
A18	CKAN Setup costs Support £ [REDACTED]/Mo Source: [REDACTED]
A19	Assumed Hybrid 20% Capex YoY for continual integrations as Data Products are onboarded
A20	Azure Hosting Costs. Largest cost is building ourselves (Opt3) (30% more), Buy (Opt1) is the mid option as we need 3 capabilities, and (Opt2) is the cheapest (20% less) as we are only adding external portal.
A21	20 working days in a month
A22	Work starts Jan 2024 and lasts 24 months to 2026 (Opt 1 and 2)
A23	Assumes 1 person can work on design and build of more than one solution (as there are 3 data portal, dynamic security access, data fabric)
A24	Data Product build assumes full automation with DEP platform
A25	Data products do not include digital frontend, this will be developed by the digital teams
A26	Data product build assumes the connectivity to the data source is already connected, in production & published. - note costs for connectivity is captured in model
A27	Data Products built on Data Products assumes existing data product exists and fulfils new data product requirements
A28	Data Product t-shirt size blend will be 70% small, 20% medium and 10% large
A29	Assume DEP have connected source required to build data products
A30	Assume DEP provide detail on SoR required to build data products
A31	Assumes Data Products built by NGET permanent staff
A32	Assumes build of Data Portal is from PARTNER partner resources
A33	All costs are list/non negotiated
A34	Assumes connection additional Data Product Source of 3 per month total of 18
A35	Assume circa 2.5 Data Products built a month
A36	PARTNER partners build the data portal
A37	NGET Employees create data products
A38	Assumes SaaS licences can be capitalised because the services can run on-premises
A39	Assume Dev and Run is Pareto Principle split 20 maintain and 80 new value (Opex/Capex)
A40	Assume no inflation in costs the models

Table 21 (Assumptions)

Assumption Notes

Note	Description
EVN1	Existing Data Product SF6 and RRP created 8 data products and cost £ [REDACTED] total, therefore £ [REDACTED] each
EVN2	It took 4 months to build SF6 and 6 months RRP data products (T-Shirt Medium and Large respectively). Large = 1.5 DP per month and Medium = 1 DP per month
EVN3	RRP created a saving in direct staffing costs of £ [REDACTED] PA, which is £ [REDACTED] per Data Product.
EVN4	Customer Connections queue optimiser data product predicted to make £ [REDACTED] year on year, for business value of a Data Product.
EVN5	Called Representative sample of suppliers to request pricing for their products, this is not to prejudice a future procurement.

Table 22 (Assumption Notes, background)

17.1.2 Appendix 13 – PARTNER Rate Card

Application Development via PARTNER partners

Role	Blended Rate (GBP)
Business Analyst	£ [REDACTED]
Architect	£ [REDACTED]
Security Architect	£ [REDACTED]
Enterprise Architect	£ [REDACTED]
Procurement Lead	£ [REDACTED]
IaC Engineer	£ [REDACTED]
DevOps Engineer	£ [REDACTED]
Data Engineer	£ [REDACTED]
Network Engineer	£ [REDACTED]
Test Lead/Test Analyst	£ [REDACTED]
Service Designer	£ [REDACTED]
Test Lead	£ [REDACTED]
Scrum Master	£ [REDACTED]
Service Manager	£ [REDACTED]

Table 23 (PARTNER Rate Card)

17.1.3 Appendix 14 – NGET Rate Card

22-May	Hourly	Daily
Band B	£ [REDACTED]	£ [REDACTED]
Band C	£ [REDACTED]	£ [REDACTED]
Band D	£ [REDACTED]	£ [REDACTED]
Level 8	£ [REDACTED]	£ [REDACTED]
Level 7	£ [REDACTED]	£ [REDACTED]
Level 6	£ [REDACTED]	£ [REDACTED]
Level 5	£ [REDACTED]	£ [REDACTED]
Level 4	£ [REDACTED]	£ [REDACTED]
Level 3	£ [REDACTED]	£ [REDACTED]

Table 24 (NGET Rate Card)