Virtual Energy System

Common framework – data & technology priority factors

Increasing visibility & enabling sharing
Creating an interoperable tech-stack
Aligning models & taxonomies
August 2023

Contents & Foreword

Foreword

Digitalisation is vital to reaching the UK's climate change targets. Continuing to innovate and adapt the energy system requires a range of new tools and a reimagining of how the entire industry can come together.

That is why ESO is today presenting this ambitious call to action for an Industry-wide mission to super charge digital twins across our energy system.

This Virtual Energy System common framework will create the common language, recommend infrastructure and processes to connect and federate individual digital twins from across the energy sector.

This open framework can proactively contribute to help meet today's energy mission to decarbonise the energy system and bridge the gap to net zero.

I am delighted to launch this shared framework and together as an industry, we can turn this vision into a reality.

Anna Carolina Tortora
Head of Digital Transformation & Innovation Strategy
ESO

Executive summary

A blueprint to create the VirtualES

Background

ESO have launched the Virtual Energy System (VirtualES) programme to enable the creation of an ecosystem of connected digital twins of the entire energy system of Great Britain, which will operate in synchronisation to the physical system to support the transition to net zero. It will include representations of electricity and gas assets and link up to other sectors.

Through research, expert interviews, and industry-wide engagement, 14 key socio-technical factors were identified which are considered necessary for the development and delivery of the VirtualES.

This report will explore the ***aligning models & taxonomies***, ***increasing visibility & enabling sharing***, and ***creating an interoperable tech-stack*** priority technical factors, three of six factors initially explored to support the VirtualES roadmap.

All three factors are intrinsically interdependent and must be considered together. They have been developed through a use case driven and iterative design approach, ensuring the features are delivered to provide benefits at the right time.

Purpose of this document

This report outlines the principles by which distributed data sharing within the VirtualES would be developed, following the principles of data mesh.

From these principles a high-level design of the VirtualES was developed. Each element of this design is described alongside details of its importance to the overall solution.

A set of six steps that describe how a data producer and data consumer will interact with the VirtualES are described through a user journey.

The key considerations to enable the establishment of VirtualES, such as security, governance, and aligned models and taxonomies, are detailed. With three initial case studies summarised that have been co-created with the sector.

Key findings

In addition to providing the high-level technical design for the VirtualES, this report provides tangible actions for future stakeholders to engage with the programme.

The key takeaways from this report are:

- The high-level design of the VirtualES incorporates several key components and functions that will ensure that distributed data sharing is possible in a secure and controlled environment.
- The need for decentralised ownership of data is fundamental to the success of the VirtualES. However, there must be the appropriate security, governance and user interfaces in place.
- Data provider and data consumer will follow key steps which can be supported by the VirtualES development ensuring compliance with standards, ability to search and find data, and for the distributed sharing of datasets via APIs.
- Use cases can benefit from the development of VirtualES and in doing so can support the development of the necessary components and functions of the VirtualES over time.

1

What is the Virtual Energy System?

ARUP

What is the Virtual Energy System?

A data sharing mechanism for the entire energy sector

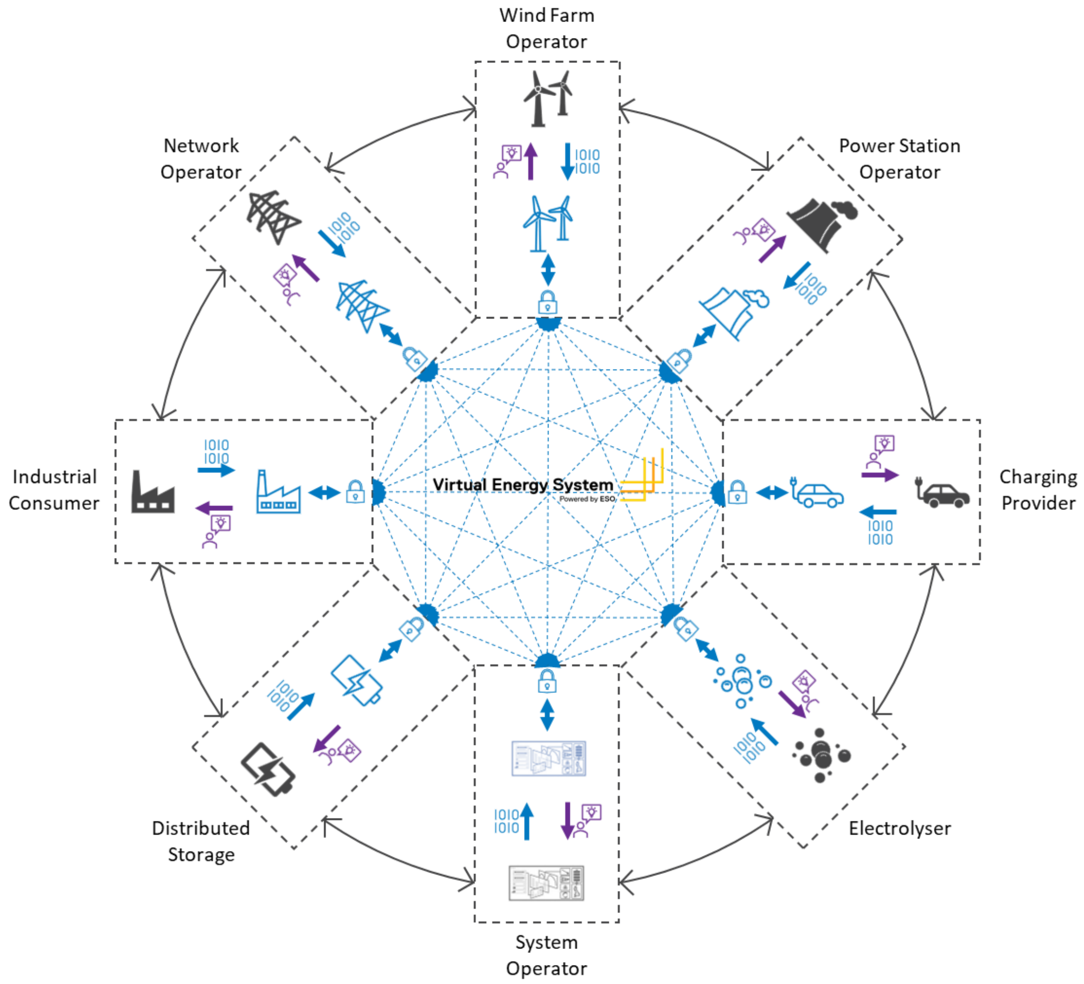
Overview

The ambition of the Virtual Energy System (VirtualES) programme is to enable the creation of an ecosystem of connected digital twins of the entire energy system of Great Britain, that will operate in synchronisation to the physical system. It will include representations of electricity and gas assets and link up to other sectors.

This ecosystem of connected digital twins will enable the secure and resilient sharing of energy data across organisational and sector boundaries, facilitating more complex scenario modelling to deliver optimal whole-system decision making.

These whole-system decisions will result in better outcomes for society, the economy, and environment by balancing the needs of users, electricity and gas systems and other sectors.

Creating the VirtualES is a socio-technical challenge that requires a collaborative and principled approach, aligned with the National Digital Twin Programme, and other energy sector digitalisation programmes.



How to enable a VirtualES?

A socio-technical challenge that requires a socio-technical solution

Overview

Enabling the VirtualES requires a common social-technical framework that outlines the agreed access, operation, and security protocols.

Through research, expert interviews, and industry-wide engagement, 14 key socio-technical factors were identified which are considered necessary for the development and delivery of the VirtualES today.

These 14 identified key factors are grouped by the categories of People, Process, Data, and Technology.

Six of these factors, the first among equals, were prioritised for immediate consideration

Priority socio factors:
- Raising awareness & fostering culture
- Engaging stakeholders
- Creating a governance framework

Priority technical factors :
- Aligning models & taxonomies
- Increasing visibility & enabling sharing
- Creating an interoperable tech-stack

People	**Defining roles & responsibilities** Formalise R&R for the VirtualES with the intentions of consumer benefits	**Raising awareness & fostering culture** Share vision, belief & behaviours. Enabling practices to support VirtualES objectives	**Building capabilities & skills** Understand skills & competency needs & develop capacity building strategies	PRIORITY FACTOR
Process	**Aligning around industry codes & standards** Identify standardised practices in industry & align around them	**Engaging Stakeholders** Nurture industrial, governmental and political support	**Creating a governance framework** Set strategy and operational governance of the VirtualES	**Determining operating environment** Business models, cross organisational legal, policy, & contractual framework
Data	**Aligning models & taxonomies** Harmonise existing data standards, taxonomies and ontologies.	**Establishing management & governance** Data management & governance requirements	**Increasing visibility & enabling sharing** Nurture effective data sharing to support interoperability	**Managing security** Set the core rules needed to address security, privacy and risk implications surrounding VirtualES data
Technology	**Connecting physical infrastructure** Physical infrastructure, devices and their connectivity required to operate the VirtualES	**Enhancing modelling and analysis** Modelling / simulation & analysis software used for current & future modelling	**Creating interoperable tech-stack** Communication, cooperation & sharing across VirtualES & other in/cross sector projects	

Descriptions of the six priority key factors

A socio-technical challenge that requires a socio-technical solution

Raising awareness & fostering culture

Raising awareness and fostering culture has been prioritised as it is critical the industry accepts and begins to advocate for the creation of data sharing infrastructure for the energy sector. Organisations must broaden their thinking beyond traditional business models and individualistic organisational objectives to understand the opportunities available through greater data-sharing across the energy sector.

This factor outlines the baseline requirements and recommended change management principles that will support an organisation to participate in the VirtualES.

Engaging stakeholders

This was prioritised as participating in the VirtualES will require input, collaboration and trust from numerous stakeholders across an organisation, including those that may not come from a technical background. guidance on how to approach this would therefore be valuable.

The engaging stakeholder's guidance note provides organisations with a methodology for identifying, engaging and evaluating the success of their stakeholder engagement plan.

Creating a governance framework

The design of the governance model was prioritised as it will define the core players that will establish the foundations of the VirtualES. An effective governance model is a key requirement to make informed decisions regarding national critical digital infrastructure and will also act as an enabler of the culture we want to foster within the VirtualES.

The proposed model seeks to legitimise a responsible orchestrator for coordination and conflict resolution with clear government backing.

Aligning models and taxonomies

There is broad range of terminologies for the energy sector, covering different aspects of both gas and electricity. Therefore, to achieve a successful, interoperable, standardised VirtualES, common ontologies need to be defined and specified that can enable alignment of models within the VirtualES.

This can be done through facilitating common vocabularies, outlining well-formed structures of definitions and relationships, and building on and using already existing vocabularies, ontologies, and domain knowledge.

Increasing data visibility & enabling sharing

Increasing data visibility starts with a distributed approach, allowing for data owners to retain control of their datasets, decide with whom to share their data, and have trust their data is secure from manipulation, or misuse.

For organisations to share their data or consume datasets from the VirtualES they will be required to meet a set of minimum characteristics. This standardised approach ensures an increase in visibility, searchability and accessibility of data for use.

Creating an interoperable technology stack

The goal of creating an interoperable technology stack is to enable smooth communication between organisations and systems, regardless of their underlying technologies or protocols. To enable interoperability in the sector, there needs to be a distributed network of digital spine nodes which can prepare and standardise data into a minimal operable standard. Through trusted protocols, the data can then be shared via the VirtualES by linking the digital spine nodes to enable the exchange of standardised data across the sector in a reliable, secure and governed way.

Six priority key factors

The relationship between the six priority key factors

These ‘first among equal’ socio-technical factors should be considered simultaneously together, along with the remaining eight key factors that will be explored next.

Raising awareness and fostering culture, creating a ***governance framework***, and ***engaging stakeholders*** were prioritised out of the socio factors as they act as the foundation of the common framework from a social perspective and will set the tone of the programme.

Culture is defined as a combination of the artefacts (e.g., organisational vision, assumptions, beliefs and values) that characterise the programme. Fostering an appropriate culture is crucial for success, particularly considering VirtualES is a first-of-a-kind programme that will involve sector-wide collaboration and involvement.

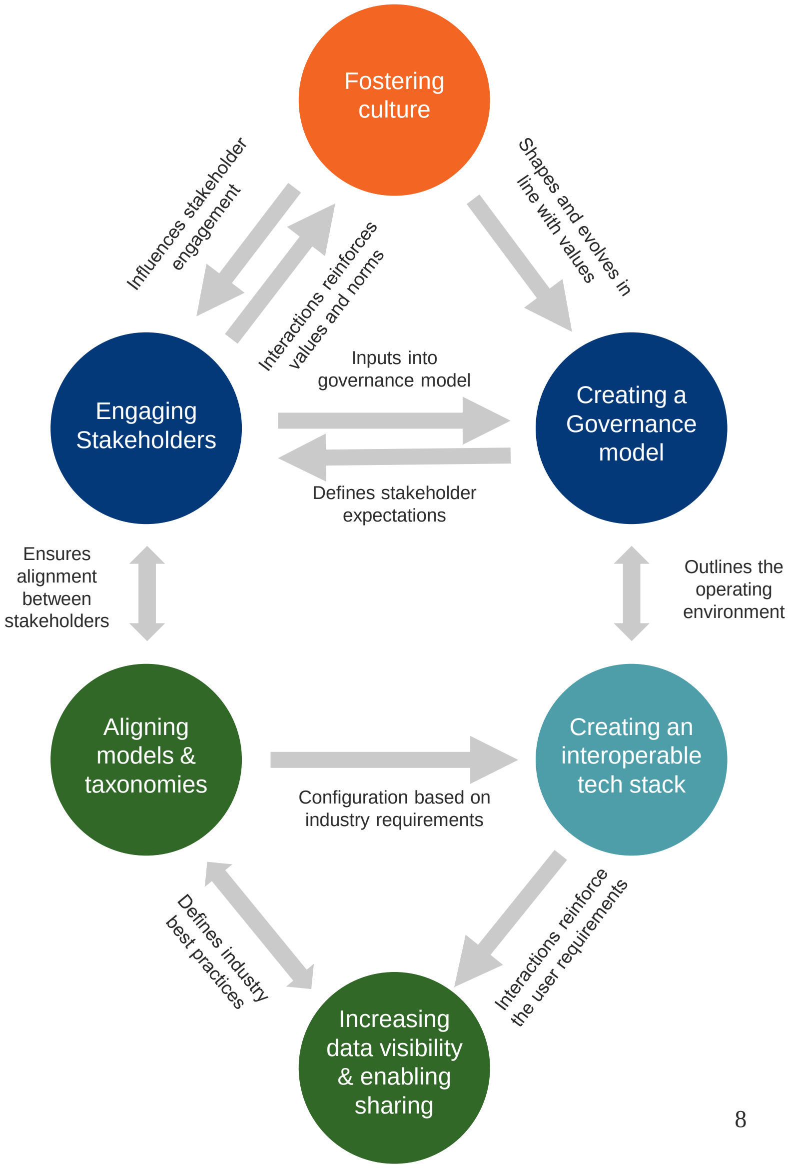
The governance model plays a significant role in reinforcing the cultural values, principles and norms that guide the behaviours and actions of individuals. When designing the governance model, it must be considered how it can reinforce and embed the cultural values, while meeting the needs of the sector.

The inherent purpose for the VirtualES is to enable sharing and visibility of data through an interoperable technology stack that ensures models and taxonomies are aligned, whilst providing trust and interoperability between actors.

Interoperability plays a critical role in enabling data sharing, ensuring standardised data formats, and an overarching governance framework. Aligning models and taxonomies ensures consistency in data interpretation.

Underpinning the other factors is engaging stakeholders. As interactions between stakeholders increase during the programme, they will influence the culture of the VirtualES. Therefore, ensuring that stakeholders who interact with the VirtualES are aware of the values of the programme and act in line with those, is important to ensure a positive culture persists over time

These six factors support the development of an energy sector data sharing infrastructure. By addressing these factors together, actors can create a harmonised data architecture that will support the VirtualES.



Delivering the VirtualES

An overview of the development timeline of the VIrtualES

Project timelines

Throughout the development of the common framework, the approach has been industry-led, consultative, and collaborative.

This approach, coupled with explicit and proactive engagement within the energy sector and with cross-sector stakeholders, is necessary for the successful development of the common framework, delivery of the VirtualES, and ultimately in achieving sector-wide adoption.

All work has been conducted openly, with all the outputs to date published online.

Defined the vision *(Dec 21 – Feb 22)*

Benchmark & key factors

- Defined the cross-sector and global best practice for connecting assets, systems, and digital twins.
  Read the report
- Outlined the key socio-technical factors that need to be considered for the VirtualES to succeed.
  Read the report

Sketched the concept *(Mar 22 – Apr 22)*

Discovery phase (SIF)

- Collaboratively proved and demonstrates, with industry, how the socio-technical principles work to enable the VirtualES.
- Recommended a tangible demonstrator use case to put the framework in practice.
  Read the report

Developed enablers *(Jul 22 – Sep 22)*

Data standards, use cases & advisory groups

- Identified data standards and outlined data licensing considerations applicable to the use case.
  Read the report
- Defined a delivery plan, governance structure, and collaborative approach, for the successful delivery of the demonstrator.
  Read the report

Set design principles *(Nov 22 – Jul 23)*

Common framework demonstrator (Alpha)

- Developed industry-validated governance model and technical architecture.
- Created wireframes of VirtualES based on the demonstrator use case.
- Defined a cost-benefit analysis methodology to assess potential use cases.

Showcase benefit *(Expected Oct 23 – Sep 24)*

Development (Beta)

- Initiate the development and testing of the demonstrator.
- Further explore the security, operating model, and technical aspects.
- Continue on-going engagement with stakeholders across the sector.

Project team

Supporting the development of the social-technical common framework

The development of the common framework has been delivered by Arup and supported by the Energy Systems Catapult and Icebreaker One. It has been sponsored by the Electricity System Operator (ESO) and National Gas Transmission (NGT) through the Network Innovation Allowance (NIA).

The purpose of the RIIO-2 NIA is to provide funding to Gas Transporter and Electricity Transmission Licensees to allow them to carry out innovative projects, that focus on the energy system transition or addressing consumer vulnerability, which are outside of business-as-usual activities.

- **Electricity System Operator (ESO):** ESO is responsible to ensure a reliable, secure system operation to deliver electricity when customers need it. ESO balances the supply and demand on the system day to day, second by second, and coordinates with networks to transfer electricity from where it is generated to where it is needed.
- **National Gas (NGT):** National Gas own and operate the national gas network in addition to maintaining and managing the 7,000,000 domestic industrial and commercial combined gas assets around the UK.
- **Arup:** An employee owned, multinational organisation with more than 15,000 specialists, working across 90+ disciplines, with projects in over 140 countries and the mission to 'shape a better world'. Arup have extensive energy and cross-sector digital twin expertise, actively contributed to the National Digital Twin programme, and are members of the Digital Twin Hub.
- **Energy Systems Catapult (ESC):** An independent, not-for-profit centre of excellence that bridges the gap between industry, government, academia, and research. Set up to accelerate the transformation of the UK's energy system and ensure businesses and consumers capture the opportunities of clean growth. ESC are responsible for the Energy Data Task Force (EDTF) & Energy Digitalisation Task Force (EDiT).
- **Icebreaker One (IB1):** An independent, non-partisan, non-profit organisation with a mission to 'make data work harder to deliver Net Zero' by creating open standards for data sharing across agriculture, energy, transport, water, and the built world.

Together the five organisations assembled a delivery team to effectively collaborate and deliver the objectives of this workstream.

2

Data & technology priority factors

Increasing visibility & enabling sharing

Creating an interoperable tech-stack

Aligning models & taxonomies

Approach

Aligning models & taxonomies, increasing visibility & enabling sharing, and creating an interoperable tech-stack

Overview

The first three technical factors, aligning models & taxonomies, increasing visibility & enabling sharing, and creating an interoperable tech-stack are intrinsically interdependent factors that must be considered together.

The purpose of this priority key factor guidance note is to outline how industry best practices are used to create an interoperable technology stack. With this stack considering user experience and accessibility to increase visibility and enabling sharing of data. With this sharing underpinned by aligning the relevant models and taxonomies that underpin the sector.

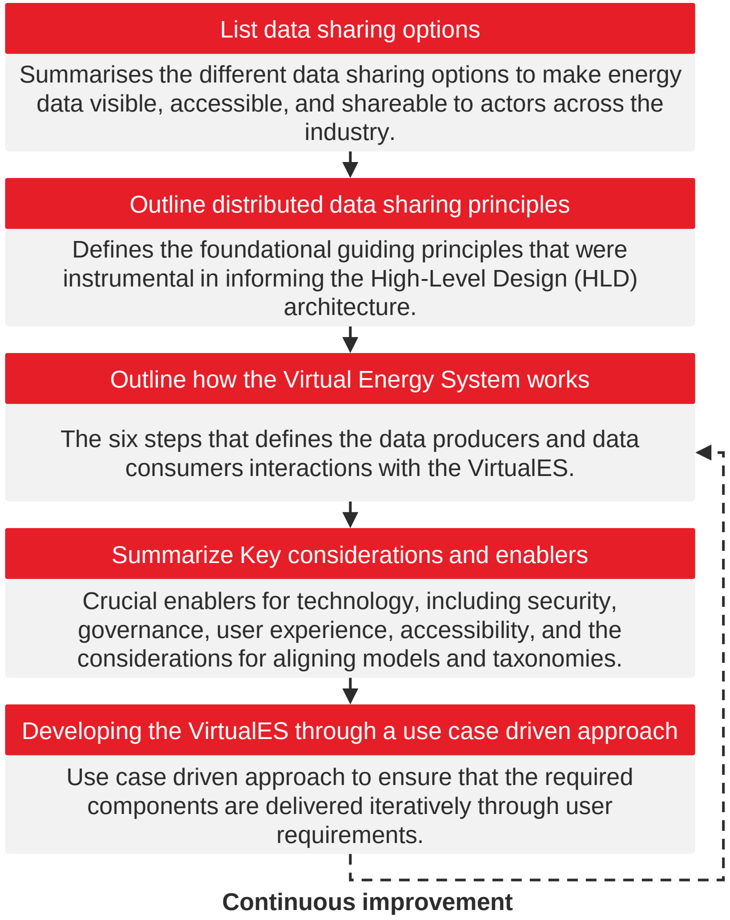
This is developed through a use case driven and iterative design approach, ensuring the features are delivered to provide benefits at the right time.

The labels at the bottom of each page indicate which of the key factors that the content is applicable to.

Aims of the priority technical factors

In the context of the VirtualES, the aim of the three technical factors are to:

- **Define and present an interoperable tech stack that underscores the VirtualES**: By using current academic research, insight from industry experts, and applying lessons learned from other related in-sector and cross-sector programmes, the interoperable tech stack is supported by latest thinking in the industry.
- **Outlines the typical journey for both data consumers and producers**: Providing a user journey that shows the common interactions and the features required to increase visibility of data between actors and enable sharing of data from one actor to another in the energy sector.
- **Summarises the interaction between an interoperable tech stack and the models and taxonomies that enables interoperability**: Shows the interactions between the core aspects of the tech stack such as security, governance, user experience and accessibility, and the standards and taxonomies that enable creation of models and digital twins.



2.1

Data sharing options

Data sharing options

Overview of the possible options for data sharing

Overview

There are various method and options for sharing data.

For the VirtualES to make energy data visible, accessible, and shareable to actors across the industry, a secure and scalable data sharing solution but be used.

Six data sharing options were identified, building on the work conducted by the data sharing architecture industry collaboration group.

Summary conclusions of data sharing options

- **Option 1 - point-to-point sharing:** Direct sharing (via emails) is not recommended because it does not meet the range of technology, data, governance and security requirements.
- **Option 2 - centralised datastore:** Suitable for use cases where the data has to be centralised and administered by a central owner, and where, for example, historical data is required for analysis.
- **Option 3 - data portals:** Suitable when acting as the user interface and data catalogue for users to search and discover their data.
- **Option 4 - distributed storage system:** Suitable to distribute data amongst different nodes and servers for instances where performance, security, and large volumes of data are required, perhaps for analytical purposes.
- **Option 5 - distributed streaming:** Suitable for sharing of real-time data, and orchestration of data sharing with users and data stores by implementing an event-driven architecture.
- **Option 6 - data virtualisation:** Is not considered as a preferred option, due to the its complexity and uncertainty on the applicability of its architecture for the future VirtualES.

Data sharing options VirtualES

The VirtualES is using a fully distributed architecture, which will involve a combination of technological solutions to accommodate multiple future use cases, and to meet a range of technology, data, and security, and governance requirements.

The design adopts data mesh principles to deliver an architecture suited for decentralised data sharing. For more information see Section 2.2.

The resulting high-level design of the VirtualES is summarized in Section 2.3.

Key factors applicable to:

Increasing data visibility & enabling sharing | Creating an interoperable tech stack

Data sharing options

Descriptions of the six data sharing options identified

Option 1: point-to-point sharing

A data producer will take a copy of the relevant data - which may be a CSV file or JSON/XML (CIM) and send the copy of the data to a data consumer.

For example, via email or conventional postal services.

Option 2: centralised datastore

A centralised and shared datastore is used to store relevant data and to connect the producers and consumers.

This will mean that the data producer will take a copy of the data, model it to store it in a database (relational, key-value, file format etc.) and makes it available for querying and consumption by a data consumer.

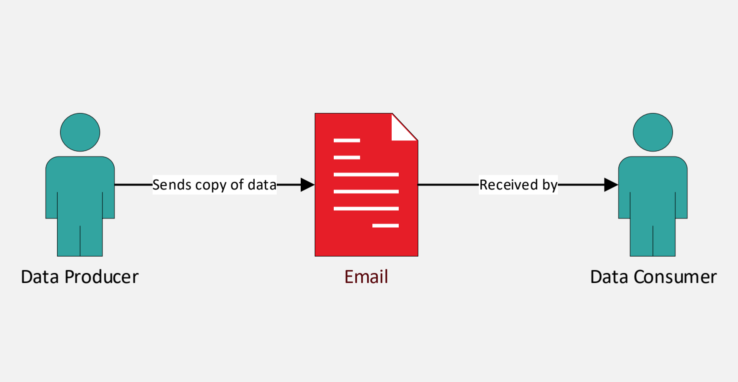
This could be used in instances where historical data is required for storage and consumption, and instances where the data does not necessarily have to be shared in real-time but, for example, through a batch process

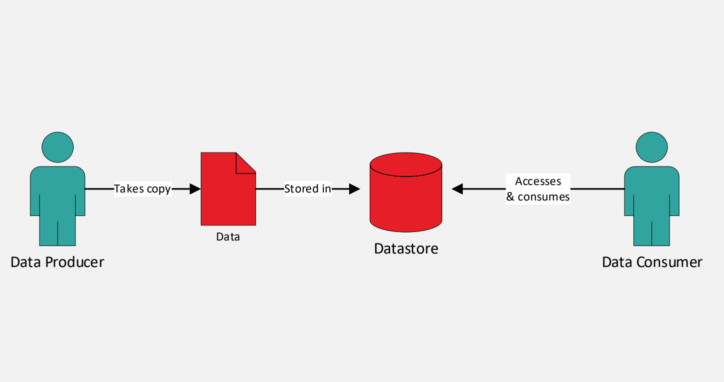
Option 3: data portal

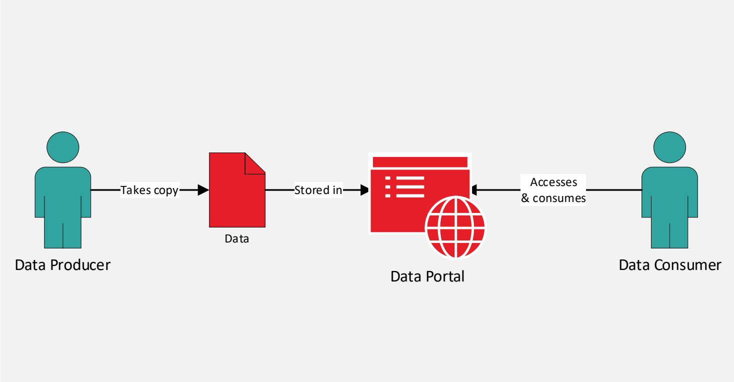
A data portal is a web-based location where data and datasets are aggregated and published in any variety of machine and non-machine readable formats.

The intention is for data consumers and producers to search, discover, access and securely share energy data. A variety of these data portals currently exist, with different formats, metadata and datasets.

Some data portals host the data by using a database, whilst others (like Open Energy) does not host the data, but provides a cataloguing capability and directs users to where the data is hosted with the central administrator.







Key factors applicable to:

Increasing data visibility & enabling sharing | Creating an interoperable tech stack

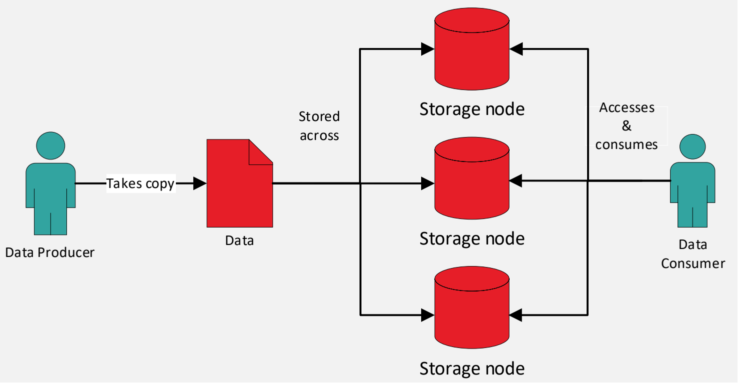
Data sharing options

Descriptions of the six data sharing options identified

Option 4: distributed storage system

A distributed storage system is infrastructure that can divide and split the data across multiple nodes/servers/data centres thereby allowing for massively scalable storage systems to be accessed by multiple consumers.

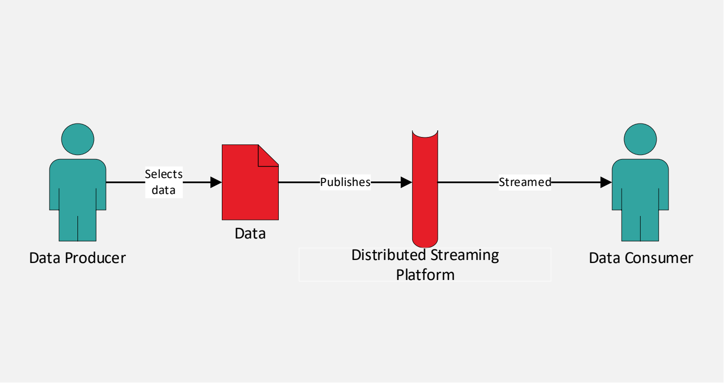
Popular examples of these are cloud storage technologies e.g. Amazon S3, Google Cloud Storage, Azure Blob etc. Another prominent example is Hadoop, an open-source framework which utilises clusters of servers to process and store data – typically used to manage big data.



Option 5: distributed streaming platform

A streaming platform is used to share event data between data producers (publishers) and consumers (subscribers). The streaming platform acts as a single platform to connect multiple users to relevant data using a publisher/subscriber model.

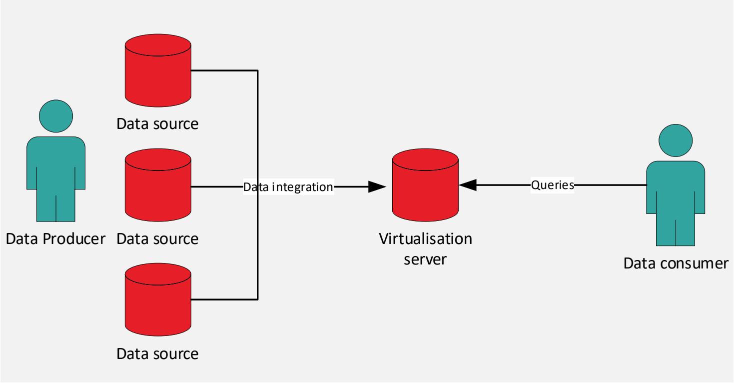
Data is streamed using a central log by publishers, and subscribers are notified and granted access via a real-time subscription to that data. The platform acts as a distributed log whereby data is stored as events in the order they were produced in a persistent and fault-tolerant way, thereby allowing for processing of event streams in real-time (similar to an enterprise messaging system).



Option 6: data virtualisation

Data virtualisation involves creating a virtual layer that sits on top of the underlying data sources and provides a unified, real-time view of the data. This virtual layer, known as a data virtualisation server, acts as an intermediary between the data sources and the consumers. This allows the data to be accessed and queried without physically copying or moving the data.

As this technique also allows data from multiple sources to be combined and presented in a single view, it eliminates the need for consumers to access each data source individually. This approach works well when organisations need a unified view of the data, but does not need to store all the data in a single location.



Key factors applicable to:

Increasing data visibility & enabling sharing | Creating an interoperable tech stack

2.2

Distributed data sharing principles

Distributed data sharing

What is distributed data sharing

What is it?

The technology architecture for the VirtualES requires various considerations for the security, data and technology requirements. These factors include the needs for high availability and scalability, including elements around asynchronous data sharing in an event-driven architecture. Overall, the goal is to design a data sharing infrastructure that can support the effective and secure operations of the VirtualES.

Based on the research conducted by the VirtualES, and the work conducted by the data sharing architecture industry collaboration group, a **distributed data sharing architecture** has been considered as the most appropriate for the VirtualES infrastructure.

The goal of distributed data sharing is to enable the seamless access to data for multiple users, regardless of its physical location. The data is decentralised i.e., it is not stored in a centralised repository with a central owner but rather the locality and ownership of the data lies with the organisations providing that data.

Why is it important?

A distributed approach provides several benefits over a centralised approach:

- **Ownership & governance:** the locality and ownership of the data remains with the data producers – rather than with a central entity.
- **Scalability:** enables increased data demands without overwhelming a centralised storage system.
- **Performance:** latency can be minimised leading to faster data retrieval and processing – especially in scenarios requiring near real-time data.
- **Availability:** the failure of one computing instance or datastore will not affect the entire data infrastructure – as it would with a centralised approach.

Whilst it is appreciated that a distributed approach offers the most promising solution to sharing large volumes of data with multiple users, there may be use cases where instances of centralised storage would be appropriate for the VirtualES but this will be subject to data governance and management agreements.

Examples

A distributed approach has been adopted by other initiatives and organisations which aim to tackle similar challenges as the VirtualES. Some of these include: Health Information Exchanges (HIEs), Interbank Payment Networks, and CReDo. Industries adopting this approach span healthcare, finance, tech, and research.

It is important to note that a distributed architecture contains many complex parts; it is not necessarily just distributed storage as considerations around distributed processing, communication, orchestration, governance, scalability etc., should also be made. Considering the scale, ambition, and complexity of future use cases that the VirtualES may need to address, an amalgamation of techniques and technologies may need consideration to achieve a performant and secure data sharing infrastructure.

To support the creation of an interoperable tech stack that enables sharing, the VirtualES will adopt some of best practices and principles of a ‘data mesh’ approach, which is an industry approach for implementing distributed architectures. These principles will not only support the technology decisions but also guide the behaviour of the participants using the VirtualES.

Key factors applicable to:

Increasing data visibility & enabling sharing | Creating an interoperable tech stack

Data mesh

Overview of distributed approaches to data sharing by using data mesh

Introduction

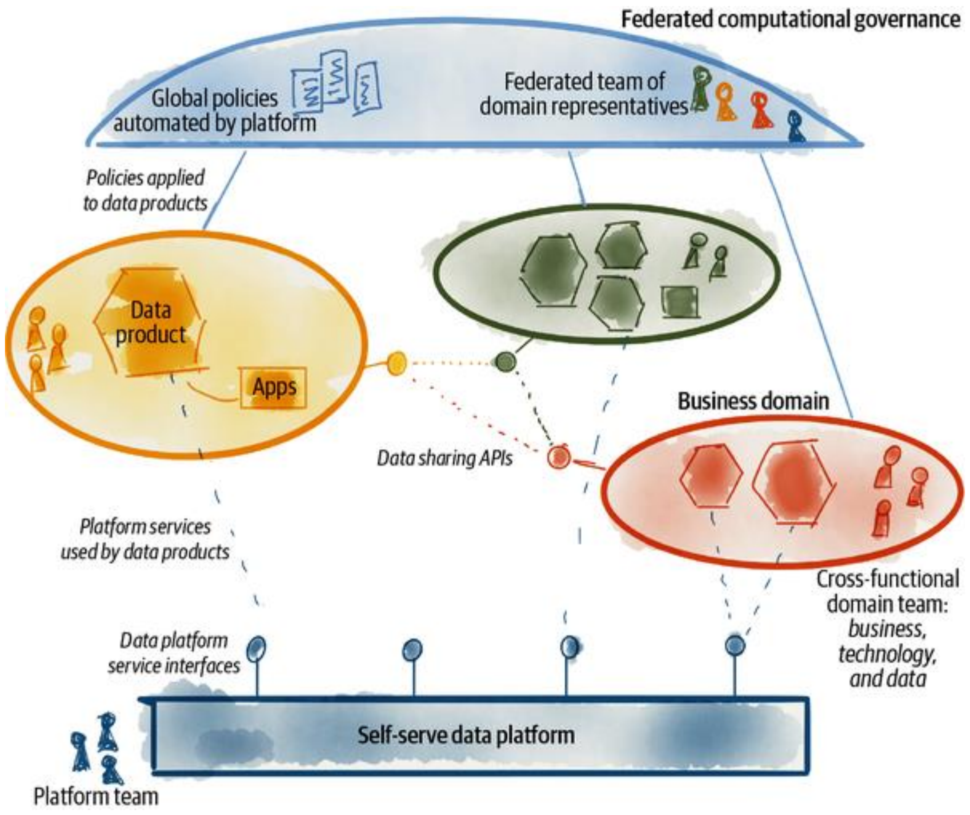
Data mesh is a decentralised socio-technical approach to the distributed sharing of data in large and complex environments – within or across sectors.

In essence, it is about organisations sharing 'data products' that are discoverable, addressable, trustworthy, interoperable, secure and governed by global standards and access control i.e., the data should be treated as a product that serves users and is reliable.

Data mesh sees key shifts in approaches to organisations and architectures and adopts a set of principles including self-serve data platform, data as a product, domain ownership of data, and federated computational governance. This approach sees a shift from traditional sharing of analytical data using monolithic data platforms and centralised ownership, to a distributed approach with decentralised ownership.

Whilst the approach is more commonly applicable to the sharing of data across domains within an enterprise, the core elements can also be applied to data sharing across organisations.

The aim when designing the VirtualES is to adopt some the guiding principles and approaches of data mesh.



High level diagram of Data Mesh applied to an enterprise
Source: Data Mesh, Zhamak Dehghani

Key factors applicable to:

Increasing data visibility & enabling sharing | Creating an interoperable tech stack

Data mesh in practise

Examples of how data mesh is being adopted in the wider industry to achieve what and how this can be used to inform the VirtualES

Data mesh in industry

Data mesh was first popularised in 2019 and is still a relatively new socio-technical approach for decentralised data sharing. It aims to create a more scalable, agile and collaborative data architecture that enables organisations to leverage their data more effectively whilst maintaining ownership and governance of the data.

Since then, data meshes have been adopted by a handful of companies including Netflix, PayPal, JPMorgan Chase and more.

Furthermore, the Isle of Wight Digital Twin pilot for the National Digital Twin Programme leverages a similar methodology through the adoption of Telicent's software platform to enable distributed data sharing.

Whilst it remains a relatively new concept, data mesh's adoption and popularity is gaining traction across several industries, as it aims to overcome some of the key challenges facing organisations when it comes to managing and leveraging their data sets, which pertain to: centralised data monoliths, data fragmentation, data governance & poor compliance, and slow time-to-value and lack of agility.

Comparison with other architecture approaches

Despite it being a relatively new architectural approach, data mesh shares similarities with other approaches and concepts. In addition to the data mesh principles, the VirtualES will also look to adopt relevant elements of other proven approaches:

- **Data lakes:** VirtualES will focus on organisations storing their data products within their own data stores and shared with others, as opposed to hosting all the data in centralised data lakes.
- **Microservices:** autonomous and independent components will be used in developing the data infrastructure for the VirtualES to enable scaling and agility.
- **Open banking:** similar to how Open Banking focuses on API-based sharing of financial data securely in a standardised way, the VirtualES will also enable this across the energy sector.
- **Data fabric:** whilst data fabric focuses on centralised ownership and governance of the data, the ability to consume data through a standardised interface where several technology and data sources are abstracted may enable a more seamless user experience through the adoption of a data portal for the VirtualES.

An evolving landscape

The technology landscape is constantly evolving, with new architectural approaches emerging and gaining traction over time. Therefore, the VirtualES must be open and adaptable to new evolving best practices and approaches that the VirtualES may need to adopt in the future.

Furthermore, like most technology concepts and approaches, it is not always feasible to implement a like for like replica of a data mesh and all its guiding principles for the VirtualES (this remains true for other industries and organisations adopting it).

The architecture of the VirtualES will align and adopt the best practices described in data mesh, however it may not always be possible to implement all of its recommendations, as compromises or different approaches may be necessary in some instances. This is true for any architecture approach.

However, by aligning with data mesh, the VirtualES can create a solid foundation that is aligned to industry best practices.

Key factors applicable to:

Increasing data visibility & enabling sharing | Creating an interoperable tech stack

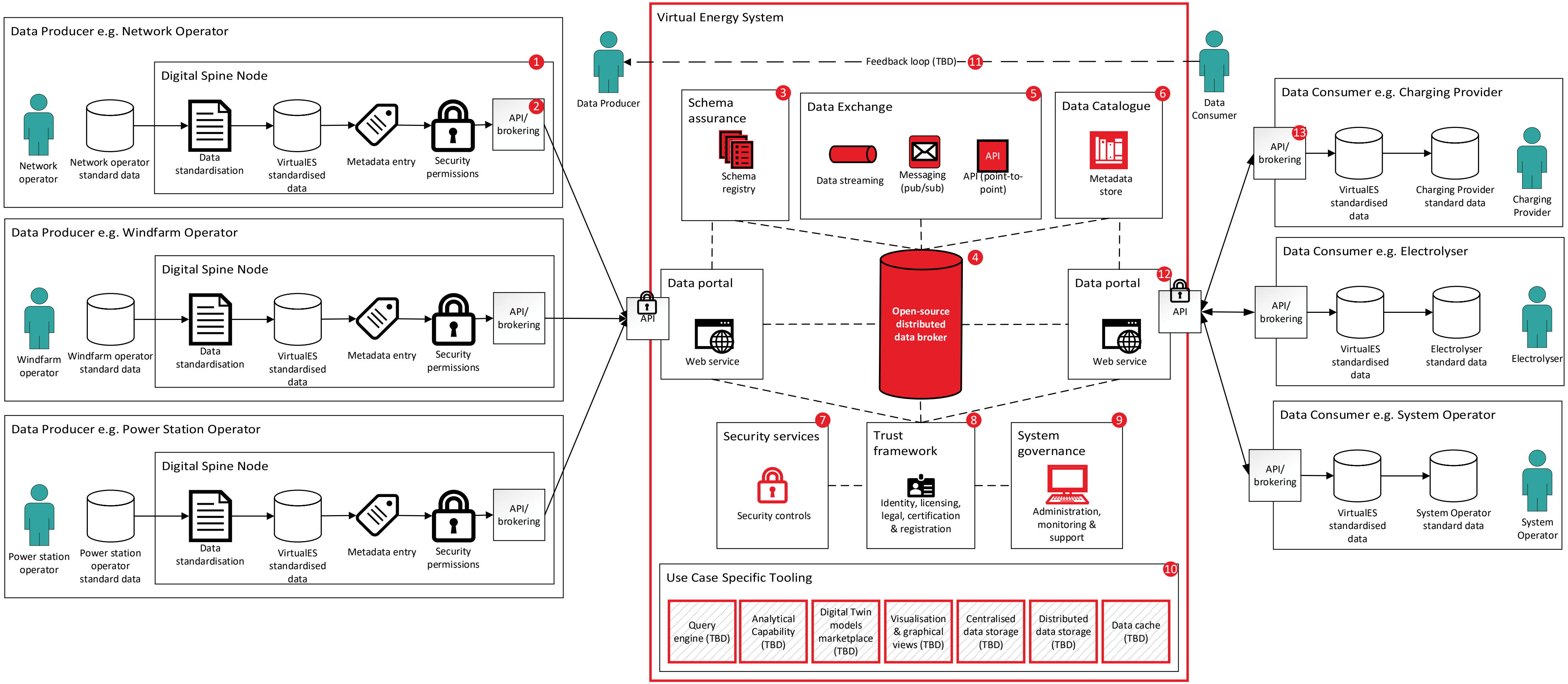
2.3

How the Virtual Energy System works

ARUP

How the Virtual Energy System works

The high-level design of the VirtualES and how data providers and data consumers will interact with it



Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

High-level design of the Virtual Energy System

Description of the components for VirtualES high-level design

Description of high-level design labels

The description numbers below align with the labels on the diagram on the previous page.

1. **Data preparation:** data producers will need to prepare their data for sharing by transforming it to an agreed standard, and understanding their data licensing conditions, metadata, and security permissions. The Digital Spine may provide some of this functionality, subject to an ongoing feasibility study.

2. **Data publishing:** after the data is prepared, it will be published to the VirtualES via secure API endpoints and data brokers (with the data ownership remaining with the data producer). This also includes publishing the data handling characteristics i.e., data licenses, legal conditions and data contracts. In addition, the metadata is also indexed and published to a data catalogue via the data broker.

3. **Schema assurance:** the data schema undergoes schema assurance through schema validation checks against a schema registry, to ensure that approved and up-to-date standards and schemas are used e.g. CIM.

4. **Distributed data broker:** the data broker is used to orchestrate the data exchange and processing of data across the VirtualES, including sharing of large volumes of data using event-driven architecture patterns.

5. **Data exchange:** data is shared through different techniques depending on the requirements of the use-case and dataset e.g. API endpoints, messaging publish– subscribe (pub/sub) and streaming.

6. **Data catalogue:** data is indexed and catalogued so that consumers can search for and request access to producers' data.

7. **Security services**: range of security controls that is applied to the system, including entity security, communication security, and system security.

8. **Trust framework:** Provides the trust service to enable the exchange of data. This includes identity management, data handling policies, legal conditions, certifications and a registration portal to enable users to register with the trust framework to use the VirtualES. The trust framework will also provide the mechanism to store and enforce data licensing conditions.

9. **System governance:** provides the governance of the VirtualES system, including administration, monitoring of data and system use and system support. This ensures that data and system misuse can be acted upon.

10. **Use-case specific tooling:** range of analytical, storage and visualisation techniques for specific use-cases e.g. a digital twin marketplace to share twin models, ML analytical workloads and graphical views of data. These services are currently TBD.

11. **Feedback loops (TBD):** feedback mechanism between the consumers and producers to provide feedback on their data and digital twin models e.g. data quality, types of data etc.

12. **Data portal:** allows for data to be searchable and discoverable. Data consumers can request access to the required data using the portal. There may be instances where the portal can also offer web views of rendered data e.g. CIM.

13. **Data consumption:** data is consumed via secure APIs and protocols, after which the data can be transformed from the agreed VirtualES data standard to an organisation specific schema, if required.

Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

ARUP

How users will interact with the VirtualES

Defining the workflow and high-level steps for using the VirtualES

Overview

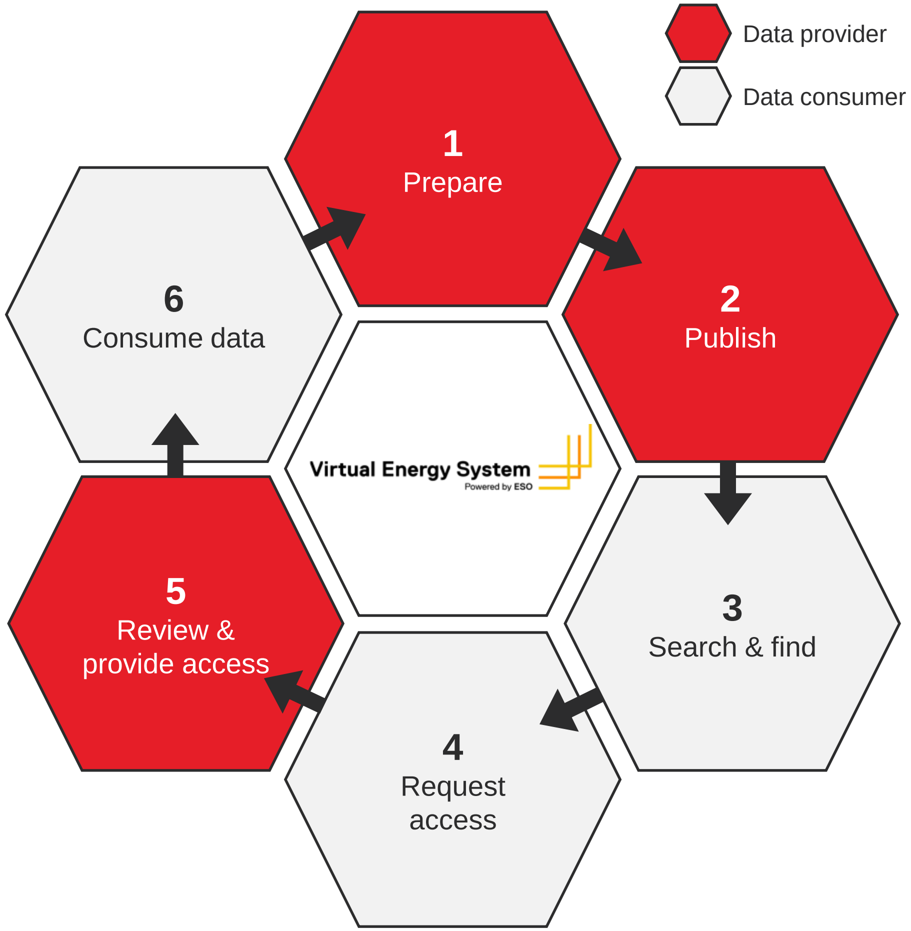
The workflow of how data producers and data consumers will interact with the VirtualES can be summarised in six steps.

1. **Prepare data**
2. **Publish data**
3. **Search and find data**
4. **Request access to data**
5. **Review and provide access**
6. **Consume data**

It is considered that these steps will broadly apply to all use cases.

Once the sixth step is complete, the process can potentially start again. For example, if a data consumer wishes to share the dataset via VirtualES derived from one they have just consumed, the process will initiate again from step one but this is subject to data license conditions of the original dataset.

Each of the steps have several associated activities and technology elements, which are explained in subsequent pages.



Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

ARUP

Step 1 - Prepare data

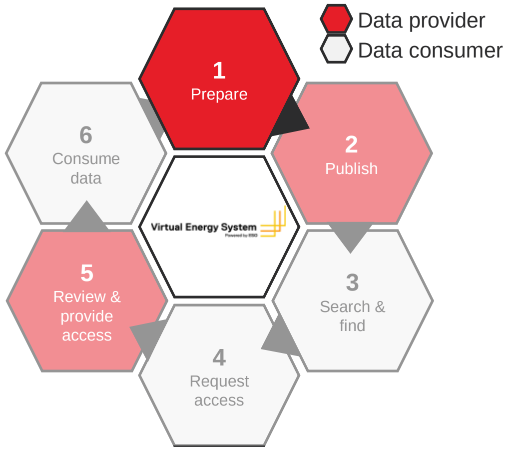
User: data producer

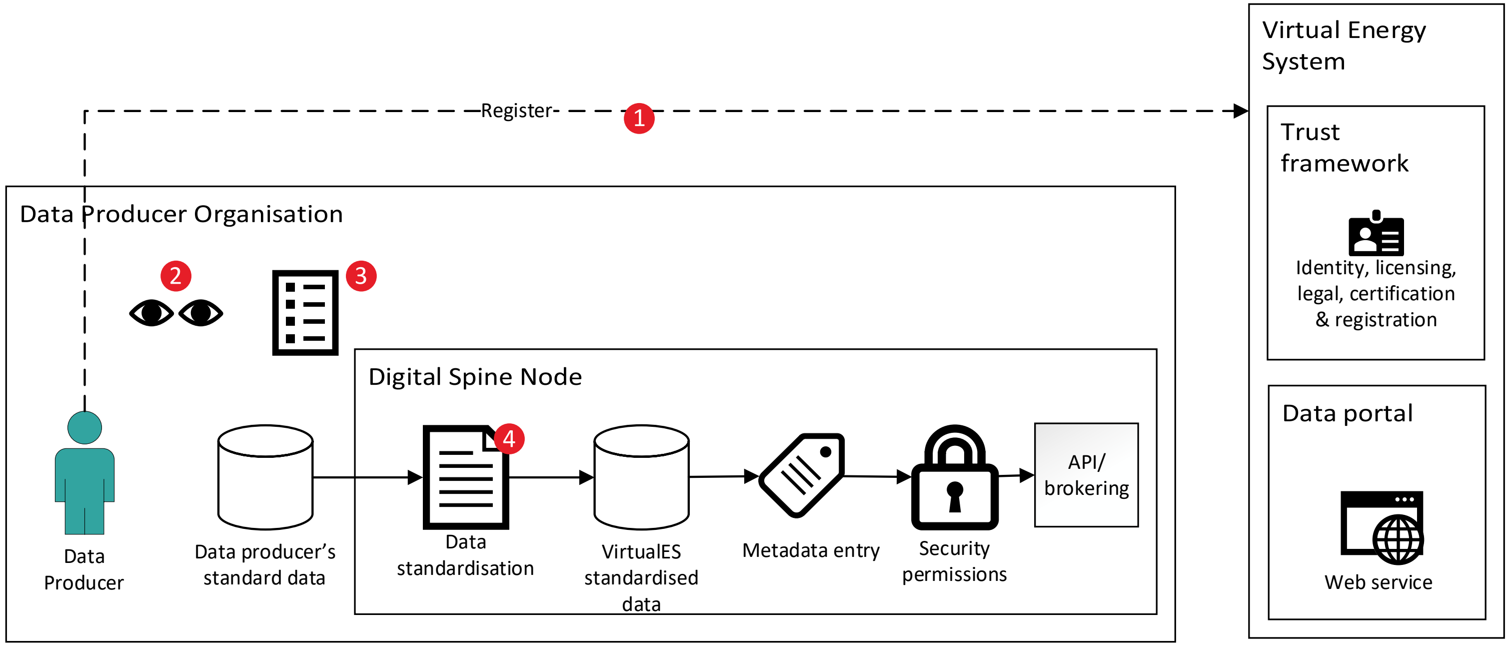
This first step involves the data producer transforming an organisational data resource into one suitable for sharing with the wider sector, and beyond, in an agreed open standard.

It must be fully documented in the form of metadata, access control permissions and other relevant characteristics e.g. licensing agreements etc.

This aligns with the outcomes of the separate digital spine feasibility study.

Descriptions of each of the numbered labels are given on the next page.





Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

Step 1 - Prepare data

User: data producer

#	Description	Importance
1	**Register with the VirtualES:** Before sharing data using the VirtualES, participants (including producers and consumers) will need to register with the trust framework in the VirtualES via a registration data portal. They will need to confirm their identity and additional details so that they can be registered in an identity management directory. This process includes providing registration information, validation and verification checks, and reviewing T&Cs. Once this is done, the data producer can be assigned certifications and appropriate access roles and permissions based on their identity.	Registering with the VirtualES through a trust framework will ensure that only verified and validated participants can exchange data.
2	**Identify data for sharing:** Data producers will need to identify the data they wish to share and where it is stored, and in what format, across their IT estate. This also includes understanding a host of security, compliance, regulatory, governance, data licensing agreements and legal implications for sharing that data, as the correct controls will need to implement as part of the Publish step.	By defining the appropriate security, governance and licensing controls will help ensure that the appropriate controls can be put in place for producers sharing their data. This includes with whom they want to share their data with, and how much of it.
3	**Identify data handling characteristics:** Once the data has been identified, producers will need to understand a host of security, compliance, regulatory, governance, data licensing agreements and legal implications associated with that data. This also includes additional metadata or characteristics to suitably describe it, including metadata, data quality and other useful information. The data may be a derived dataset, in which case there may be certain data licensing conditions that the producer may need to abide to. This will be subject to the use case, the dataset, data handling conditions, and intended use and target consumers for the data.	By defining the appropriate security, governance and licensing controls it will ensure that the appropriate measures can be put in place for producers sharing their data. This includes with whom they want to share their data with, and how much of it. Defining the appropriate characteristics of the data will make the data understandable, useful and self-describing. This, alongside the security & governance controls will ensure that the data is a treated as a product as contains all the required components and characteristics for sharing.
4	**Data standardisation:** Data producers will need to transform their datasets into an agreed common format and/or standard suitable for sharing with the wider sector. Once this is done, the data will sit in a standardised datastore where it will be published to the VirtualES through approved API endpoints and data brokering techniques (e.g., data streaming, and pub/sub) during the Publish step. This functionality could be provided by a digital spine node, which is referred to as a data standardisation mechanism node in its ongoing feasibility study. The deployment of a digital spine node provides a blueprint for preparing and publishing data into the VirtualES where participants could choose to either adopt the digital spine or use their own incumbent tooling to transform their data.	Standardising the data to an agreed format will enable interoperability of data throughout the energy sector and beyond. It will also enable key software / analysis packages to standardise on import / export functionality.

Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

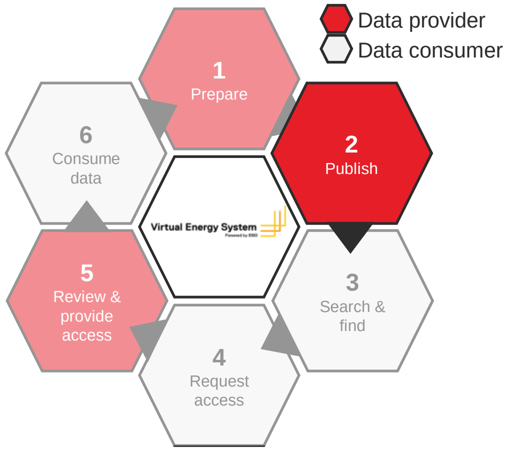
Step 2 - Publish data

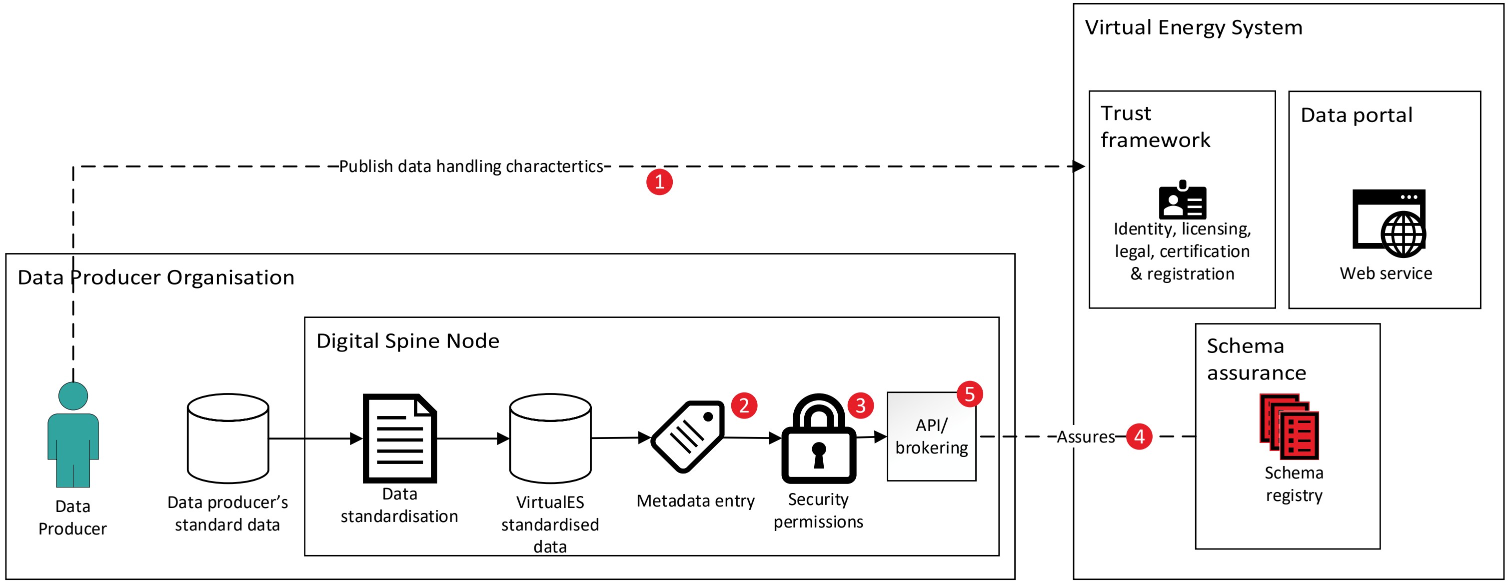
User: data producer

Once the data is prepared, transformed and documented it is ready to be published into the VirtualES by the data producer.

This entails of publishing the relevant metadata, applying security controls, validating the data schema and using approved protocols and APIs.

Descriptions of each of the numbered labels are given on the next page.





Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

Step 2 - Publish data

User: data producer

#	Description	Importance
1	**Publish data usage characteristics:** data producers will publish their data usage conditions to the trust framework. They will do this by assigning the correct data licensing conditions, data contracts, legal T&Cs to their data. This may also include additional metadata associated to the handling conditions to help enforce and communicate them across the VirtualES. These data handling conditions will be managed and retained by the trust framework using standardised formats and metadata.	By publishing the correct data handling characteristics, will mean that the trust framework can ensure that data consumers are aware and able to meet the data licensing conditions are part of the Request Access step 4.
2	**Publish metadata to the data catalogue:** the metadata associated with a dataset will be published and indexed to the data catalogue. This will require adoption of a suitable metadata standard that is aligned with sector data best practices and e.g. Ofgem's Data Best Practice Guidance, Dublin Core etc. When the producer publishes their data to their API endpoint, the distributed data broker will update and populate the data catalogue with relevant metadata automatically.	Publishing the metadata characteristics to the VirtualES will allow for the data to be indexed using a suitable standard and will enable data consumers to search and find the data they are interested in for consumption.
3	**Apply security controls:** data producers will then need to apply the appropriate access permissions and security controls to their data. This will include access control permissions e.g. fine-grained access control, so that the producer can describe what data they want to share and how much of it they want consumers to see. In addition, security controls pertaining to the data classification and specification for any PETs e.g. encryption, anonymisation techniques etc, may also be applied for sensitive datasets that require additional protection. These controls will be managed by the security controls in the VirtualES, which includes controls for message integrity protection, non-repudiation, message anonymisation etc.	Applying the correct security & governance controls will enable the data to be shared securely and compliantly in accordance with the handling arrangements for that data. In addition, data access requests can be approved or denied depending on the consumer's credentials and access control permissions for that data, as part of the Step 5 Review & Provide Access.
4	**Validate schema:** a schema validation check is conducted before the data is published to the VirtualES through the use of a schema registry which ensure that the data meets the correct schema validation check will be conducted on the producer's data against a schema registry in the VirtualES, to ensure that the data meets the required and up to date standard for inoperability. Conversely, a schema validation check can also be conducted when a consumer ingests data in Step 6 Consume Data.	A schema validation check is essential to ensure that only approved, complete and aligned schemas are used within the VirtualES. Datasets which haven't been validated may lead to issue around software import fails, and incorrect interpretation and use of the dataset.
5	**Publish via API & data brokers:** the data is published to the VirtualES through approved protocols such as API endpoints and data brokers. The data exchange will need to accommodate sharing of both static and dynamic data, between multiple producers and consumers - depending on the use-case. Depending on the use-case, the data can be published and shared via: 1. API endpoints: producers will expose their API endpoints for their data where they can be accessed by authorised consumers. Consumers can make requests to fetch that data when they require it through point-to-point sharing, typically for static datasets. 2. Data broker: the data is published to a distributed data broker whereby the data can be shared through event messages using a pub/sub method. Alternatively, the data can also be streamed in near real-time, depending on the use-case. Typically, this will be for sharing of dynamic data using asynchronous messaging, where high volumes of data is shared.	The adoption of approved and secure API endpoints/data brokers will ensure that data is shared securely, and effectively through a data exchange method that is most suited for the dataset and the use-case. Furthermore, publishing through trusted and secure APIs/data brokers will ensure that the producers are registered with the trust framework and are trusted participants. This will enable the application of security and governance controls to the data as it is exchanged using the VirtualES.

Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

ARUP

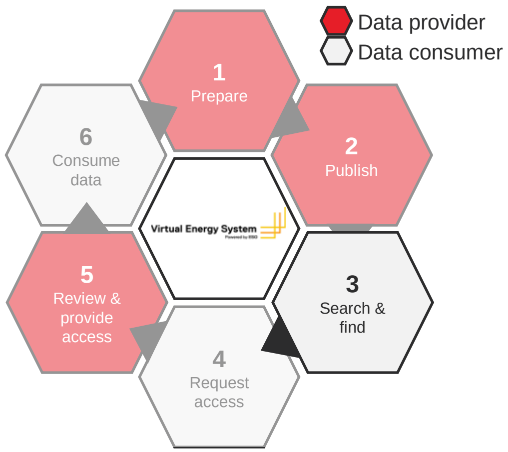
Step 3 - Search & find data

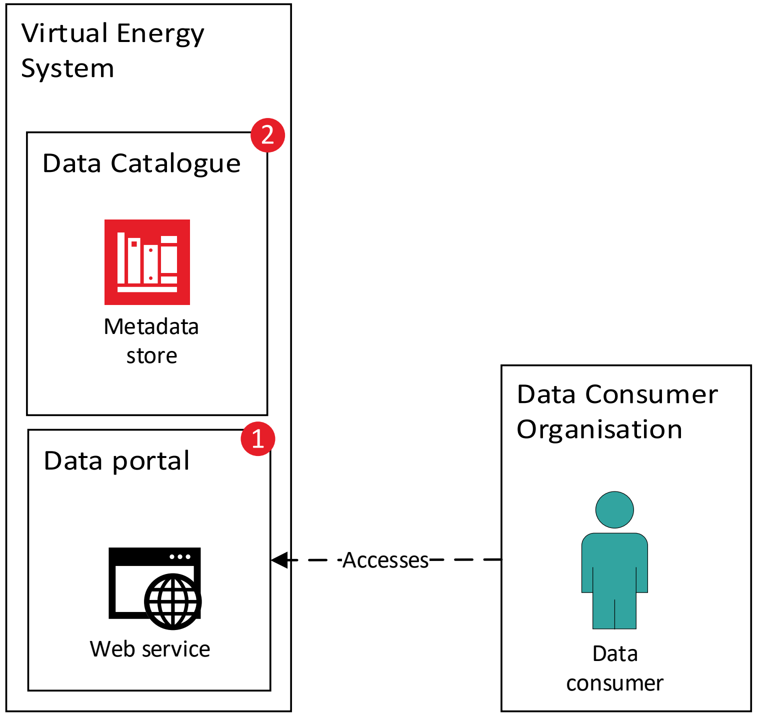
User: data consumer

This step is where the data consumer first enters the process.

They will be presented with the appropriate data available to them based on their registered details, and through a combination of search and filter process in the VirtualES data portal will find the datasets they are interested in using.

Descriptions of each of the numbered labels are given on the next page.





Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

Step 3 - Search & find data

User: data consumer

#	Description	Importance
1	**Access data portal**: consumers will access the VirtualES via a data portal as the front entrance/user interface. This will be via web link that they can access after they have registered with the trust framework as part of Step 1 Prepare.	It is critical that the VirtualES has a browser-based interface to enable anyone to access the datasets available. Registration to the portal also forms part of the trust framework to enable automatic data access depending on the permissions and access control policies. The data portal offers an intuitive and accessible user interface to navigate the VirtualES.
2	**Search using the data catalogue**: using the data portal's metadata store/data catalogue, data consumers will be able to search and discover the data they are interested in. There may be instances where the data portal can also offer thumbnail and/or web views of rendered data. The catalogued data will adhere to a common metadata standard (e.g. Dublin Core) as described in Step 2 Publish. If a specific dataset cannot be found, the data catalogue should also include links to other relevant external data portals, so that a more extensive search for that data can be conducted directly with possible data providers. This reflects the likely situation that not all data will exist within the VirtualES for a variety of reasons.	The data catalogue will allow users to search, find and understand the range of data available. The use of standard metadata fields enables the provision of filters and search criteria which will help users find the datasets they need, and understand its suitability for their requirements. Furthermore, the catalogue will also provide additional information pertaining to the characteristics of the data that was manually provided by the data producers in Step 2 Publish, for example, information around the quality of the data, and other fields describing its usefulness.

Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

ARUP

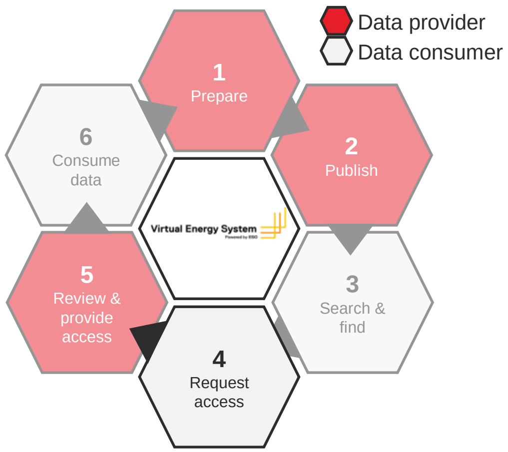
Step 4 - Request access to data

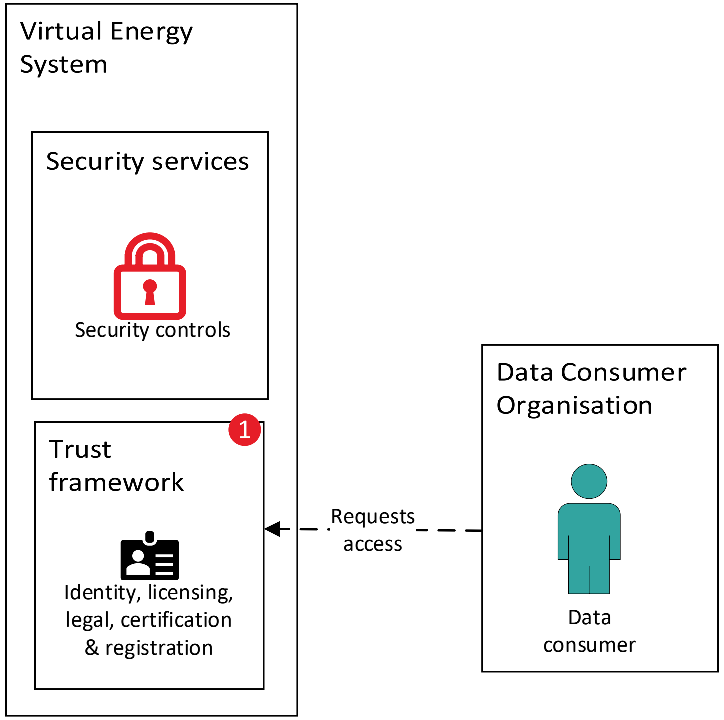
User: data consumer

The user has now found the dataset they require through review of the data catalogue information and by use of filter and search criteria.

The data consumer now wishes to use that dataset and they request access to it. Access will depend on the user's status and security group they belong to.

Descriptions of each of the numbered labels are given on the next page.





Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

ARUP

Step 4 - Request access to data

User: data consumer

#	Description	Importance
1	**Request access using the trust framework:** Once the data consumer has searched and found the data they are interested in, they can request access to it via the VirtualES. The trust framework will manage the approval process by ensuring the data handling characteristics and legal T&Cs are enforced and met through data contracts. The access control permissions and other security controls specified in Step 2 Publish will also come into effect through the security services. During this step, there may be two outcomes: a. The data consumer already has the correct credentials that meets the data handling & licensing conditions and security permission, and therefore they can subscribe and consume the data from the producer, without the need to request permission. b. The data consumer does not currently have access to the dataset and therefore they will need to request permission and justification for consuming the data through a webform in the trust framework. The dataset's access permissions will need to be updated to grant access to that consumer. They will also need to agree to the data licensing and handling conditions through the formalisation of a data contract in Step 5.	A trust framework will inform the user based on their status whether they have policies and permissions to consume a dataset immediately, or whether they have to request permission. If they are a member of a trusted user group – automatic access will speed up access and subscription to the data.

Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

ARUP

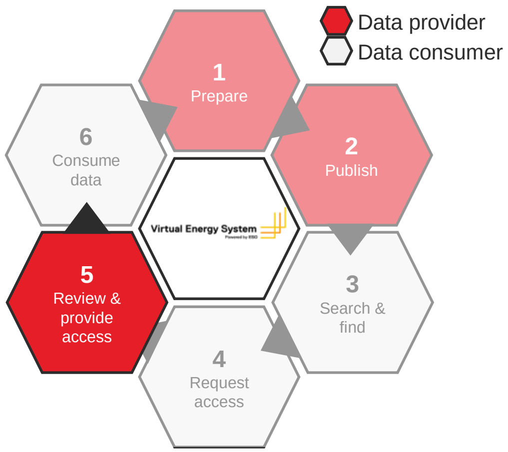
Step 5 - Review and provide access

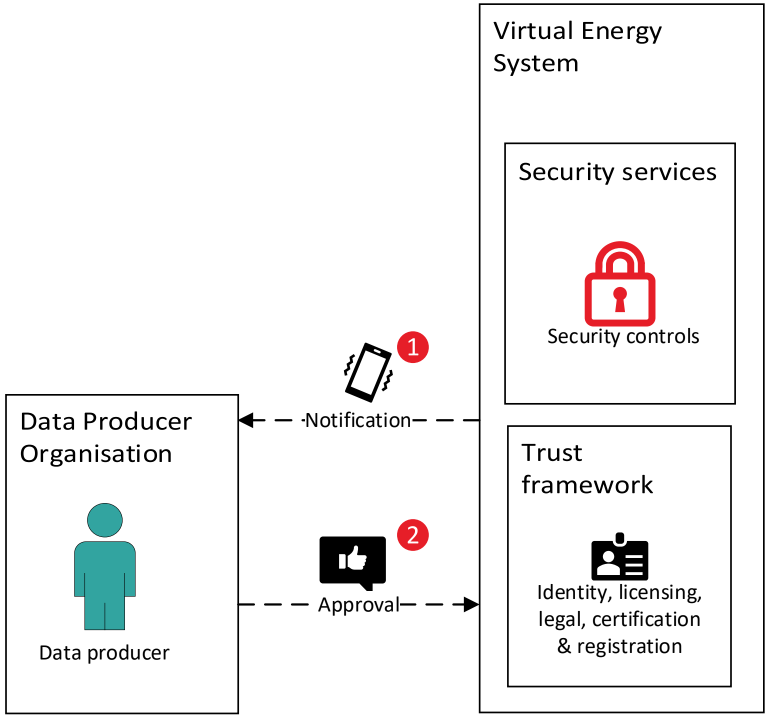
User: data producer

Back to the data producer, who will review the access request, including the justification for the data is being requested. This step is only relevant to data consumers who are not pre-approved for dataset use via the Trust Framework. They would go straight to Step 6.

After the review, they can choose to provide the consumer with access to the data, thereby updating the access control permissions for that data.

Descriptions of each of the numbered labels are given on the next page.





Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

Step 5 - Review and provide access

User: data producer

#	Description	Importance
1	**Review access request:** the data producer is notified of the access request to their data from the data consumer, including their justification for access. The data producer can choose to accept the request based on who is requesting the access and why.	Data producers can see who is requesting access to their data and why, thereby ensuring the data is shared only with appropriate organisations if the data is potentially sensitive/closed.
2	**Provide access**: the consumer's permissions are updated to reflect their access rights in the security services, and their agreement to the data handling & licensing conditions for the data is formalised through a data contract in trust framework. The data consumer will now have the correct permissions to the data and can begin consuming it as part of Step 6.	Upon reviewing the request, the access control permissions and the policies of the data can be updated to reflect the correct permissions for the data consumer. The trust framework will enforce these permission and ensure that the data is shared securely through a combination of the VirtualES trust framework, security services and system governance controls.

Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

ARUP

Step 6 - Consume data

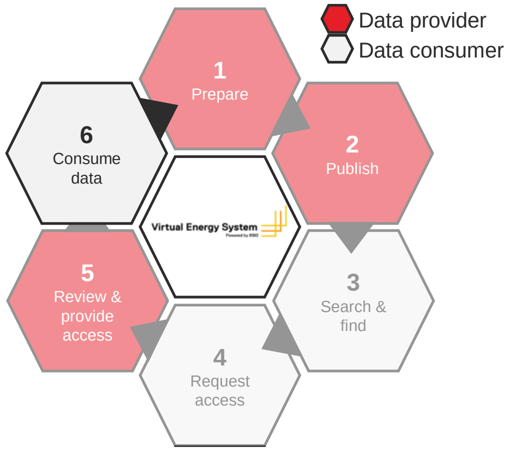
User: data consumer

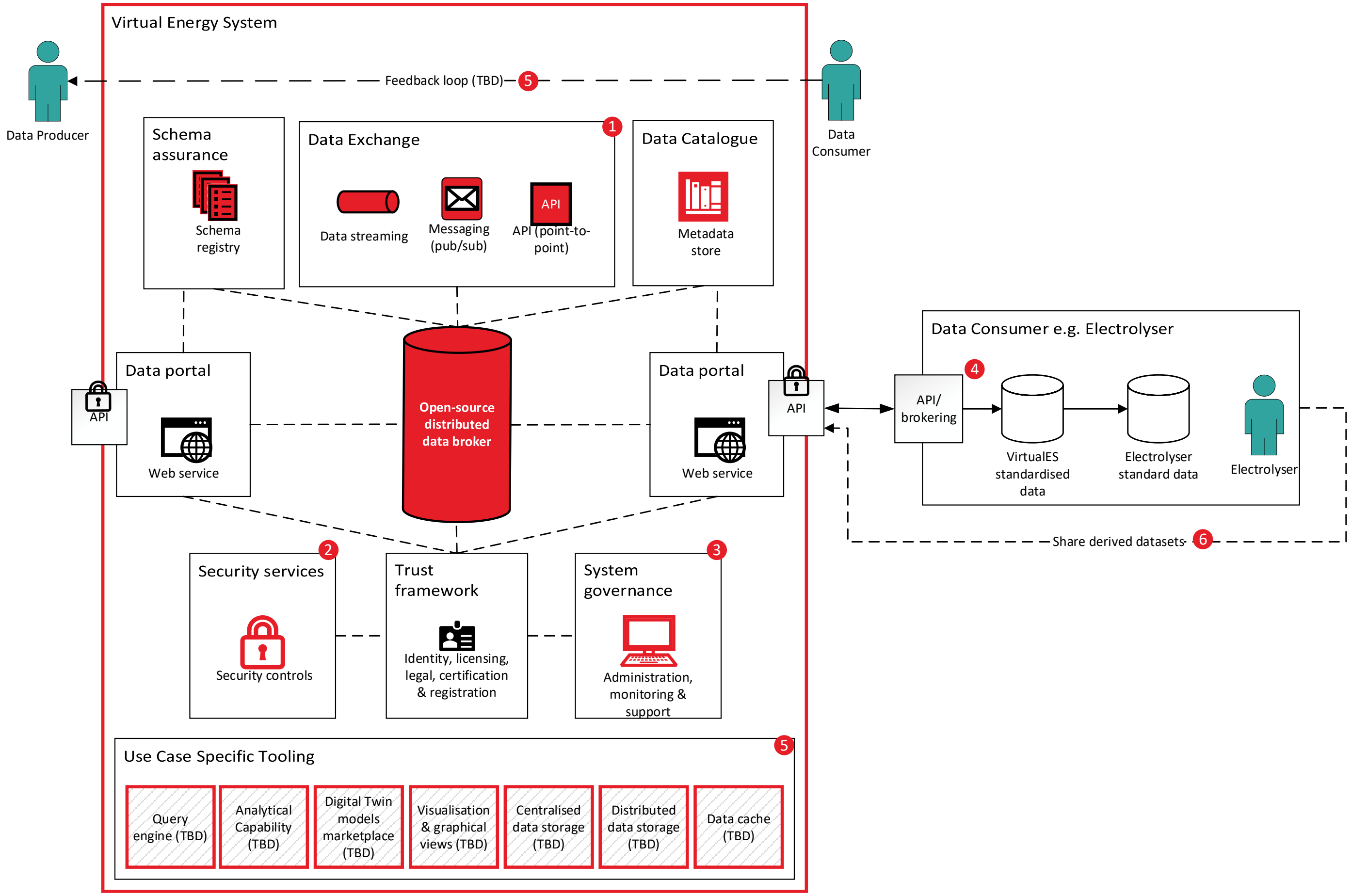
The final step is for the data consumer to consume the data offered from the data producer.

This is achieved through a data broker (pub/sub and data streaming) and directly through API endpoints.

Depending on the use case, consumers may also have access to a range of optional analytical and visualisation techniques to help them understand, access and make use of the data more effectively.

Descriptions of each of the numbered labels are given on the next page.





Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

Step 6 - Consume data

User: data consumer

#	Description	Importance
1	**Data exchange:** the producer can transmit their data to the consumer in different ways depending on the use-case offered to the consumer in 'right time, i.e. the timeframe most useful for its receipt. These include **data streaming** for sharing near real time via data broker, **messaging (pub/sub)** for a periodically updated dataset or an **API endpoint** for data to be consumed directly from the producers via an HTTPS request to fetch the data and ingest via a point-to-point sharing.	By offering different exchange patterns through streaming, messaging (pub/sub), API endpoints, means that producers and consumers can exchange data through the correct method for their datasets and use-cases. A distributed data broker will orchestrate the sharing of large volumes of data across the VirtualES using event-driven architecture patterns.
2	**Security services:** the data that is shared is subject to a number of security controls pertaining to access control, confidentiality, legal, and compliance to meet the security requirements for that dataset and use-case. These controls will be implemented by the security services architecture including entity security, communication security services and system security services. The security controls will be subject to the specific use-case and requirements for that data i.e., sensitive datasets may require the application of more restricted controls. Audit trails and security information is collected centrally by the system governance function for monitoring and remediation purposes.	Implementing a range security controls will provide the participants with the correct tools to protect and secure their data.
3	**System governance:** data exchange and activity across the VirtualES is monitored by a governance function. This function provides system monitoring through the collection of audit trails and log to help resolve system and data misuse. This also includes administration and system support to ensure that the VirtualES is running and performing accordingly and so that participants are using it appropriately.	System governance provides the administration, monitoring, and support to ensure that the VirtualES is operating as intended and to resolve any system and data misuse.
4	**Data ingestion:** The data consumer ingests the data requested through approved APIs via their specific data exchange technique. Once the data is ingested, consumers can choose to use or convert the data from the VirtualES standard into a propriety or system specific standard whereby it can be used for analytical and modelling purposes. This must be done in accordance with the data handling characteristics, including the data licensing conditions for that dataset.	The ingestion of standardised and well documented datasets via APIs will lead to efficiencies, additional modelling opportunities in order to derive new insights and planning outcomes in a timely manner. Sharing will be conducted within the confines and control of the virtual energy system streaming service enforcing the data provider conditions where feasible.
5	**Access to use-case specific tooling:** a range of analytical and visualisation services are provided to the participants based on the use-case. This may include graphical views of data, query engines, a digital twin marketplace to entire digital twin models, cached data and storage techniques for analytical purposes. In addition, there may also be a feedback mechanism between the consumers and producer for their data and digital twin models. These services are 'TBD' as future use cases and requirements will drive their implementation.	Offering a range of potential future components and applications will enable participants to use the VirtualES for a number of different use cases and requirements. The data infrastructure will grow to encompass a number of services to enable the exchange of digital twin data across the sector. It will lead to efficiencies, additional modelling opportunities in order to derive new insights and planning outcomes in a timely manner, which are presently difficult to achieve.
6	**Share derived datasets:** After the data is consumed, the data consumers may use that data to build derived datasets that they may offer to other participants through the VirtualES if the license conditions of the original data permit. In which case the consumer becomes the producer and cycle restarts from Step 1.	The data that is consumed can be used to inform digital twin models and other analytical workloads. The datasets produced as an outcome of this can be usefully shared back via the VirtualES to inform other models across the sector.

Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

2.4

Technology stack considerations

ARUP

Key considerations & enablers

Considerations and enablers for the VirtualES and organisation participating in it

Overview

There are several consideration and enablers that are necessary within and organisations and in the wider sector for the VirtualES to be successful.

Resources, roles, and skills

Data providers need to have resources in place to manage the preparation and publishing of data into the VirtualES, ensuring the correct metadata, licensing, data usage conditions and applied, and have reviewing the access requests to datasets.

Resources, roles, and skills will be further explored by separate key factors, *roles and responsibilities* and *building of capabilities and skills.*

Security

Security of data and users must be ensured within the VirtualES. Some datasets will be open, but others will only be accessible by trusted users. Therefore, security of the data infrastructure and the data must be embedded within the system through use of appropriate cyber security controls, (Privacy-Enhancing Technologies) and access control permissions.

Security will be further explored by a separate key factor, *managing security*.

Aligned models & taxonomies

Central to the VirtualES is the need for aligned models and taxonomies.

This enables data to be shared in ways that can be machine read, agnostic to specific proprietary software requirements.

This enables data exchange in a consistent and efficient manner. This key factor is detailed further in Section 2.5.

Governance (system and data)

The governance and administration of the VirtualES is crucial to the success of the system.

There must be appropriate processes and ownership of the system in place both for the management of the system and also custodianship of any limited data held within the system.

Governance for system and data will be addressed by separate key factor, *establishing management and governance.*

User experience & accessibility

The VirtualES design and user experience should enable users to easily engage with it.

It must take into account all user profiles' needs, including accessibility.

Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

2.5

Model alignment considerations

Aligning models and taxonomies

Supporting alignment of models and taxonomies is key to enabling robust data sharing and developing an interoperable tech stack

Introduction

The purpose of models and taxonomies is to provide data users with a structured and standardised framework for organising, categorising, and exchanging data in a consistent and efficient manner. This framework ensures that data is accurately and uniformly represented, allowing users to easily interpret, analyse, and compare information from diverse sources.

Models and taxonomies might be specific to a sector such as energy, aviation or healthcare or they may be used to describe components or concepts that are sector non-specific such as geospatial information or climate change.

Through alignment the focus is to both agree across the sector what models and taxonomies to use and for what purposes, and where there are identifiable gaps how they should be filled. Additionally, alignment ensures that different standards in use can be effectively integrated and interconnected. Establishing connections and mappings between different models and taxonomies enables seamless data exchange and interoperability across domains and disciplines. This will include agreeing which metadata standards, such as Dublin Core, should be adopted.

How this relates to the sector?

The Energy Data Task Force (EDTF) in their report A strategy for a Modern Digitalised Energy System carried out a review of the current energy data landscape and identified two key issues of data gaps and extracting value. These issues were seen to impeded innovation and prevent new ways of working that would deliver a more efficient system and ultimately benefit the end consumer.

The report made a series of recommendations, the second of which was to *Maximise the value of data*. This is a call to the sector to adopt the principle that Energy System Data should be *Presumed Open*. This principle requires data to be 'Discoverable, Searchable, Understandable', with common 'Structures, Interfaces and Standards' and is 'Secure and Resilient'.

The report set out three approaches to standardisation which address the variation across the energy sector. These are *Standardisation driven by Value*, *Government or Regulator Led Approach* and *Government or Regulator Led Development*. The report makes it clear that most standardisation can be developed by the market to respond to shared goals but that the Government or regulator will intervene where appropriate.

What alignment means for VirtualES and the sector

No one model or taxonomy describes all aspects of the energy sector and with the VirtualES looking to support a wide range of use cases alignment is key to ensuring an outcome that enables better integration of datasets.

Work by the sector to develop this alignment is underway through the establishment of panels and working groups, like that chaired by Ofgem for the Long Term Development Statement. The LTDS working group has adopted the Common Information Model (CIM) as a standard and are working, with input from the sector, towards the development of a GB CIM standard. Other working groups include those set up by ESO facilitating modifications to codes such as the Grid Code, impacting on critical data sharing between operators. These working groups and panels are driving the alignment through cross sector engagement and stakeholder buy in.

In addition to what exists in the sector and the current use cases being considered, other parallel initiatives such as the National Digital Twin Programme, are being actively engaged with to ensure future compatibility. Without this there is a risk of misalignment and an inability to integrate data and programmes in the future.

Key factors applicable to:

Aligning models & taxonomies

Methodology

Repeatable approach for aligning models and taxonomies to support use case implementation and ensure the success of the VirtualES

Approach overview

The methodology consists of a five-step process that will be used to identify and evaluate the relevant models and taxonomies required to support use case implementation. The approach, as shown in the diagram opposite, is repeatable and can be applied to any future VirtualES use cases. The five-steps consider:

1. **Review of existing data models, standards and taxonomies:** Given the data requirements, establish the existing relevant models, taxonomies and standards relevant to the use case.
2. **Stakeholder engagement:** Establish, through user research and interview, the pain points experienced and the existing "as-is" landscape for the use case.
3. **Requirements refinement:** Refine the understanding of relevant models and taxonomies and requirements needed to support the use case.
4. **Alignment:** Summarise the needs for the use case and outline suggested models, taxonomies and standards required to successfully implement it.
5. **Gap analysis:** Conduct a gap analysis between the, data requirements, and existing relevant models and taxonomies to inform any changes or adaptations that may be required.

Key factors applicable to:

Aligning models & taxonomies

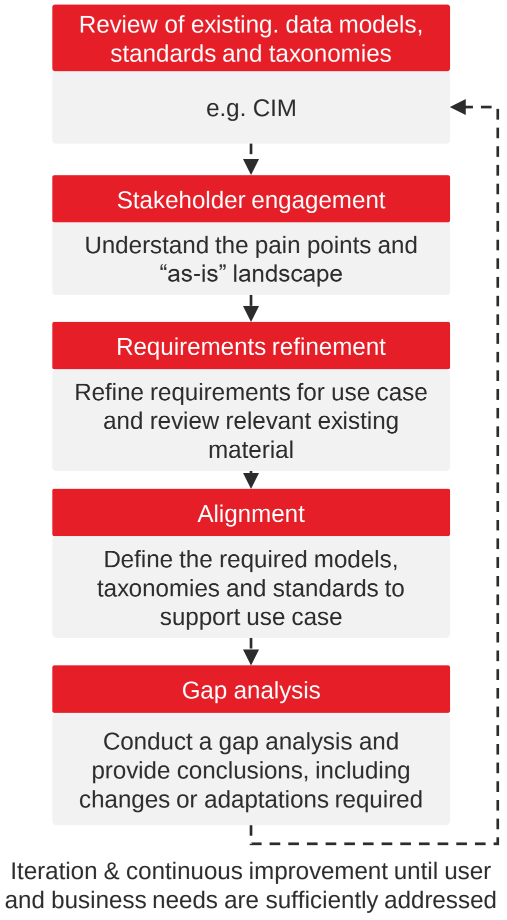
Where this methodology has worked before

The aviation sector, due to the global cooperation and interoperability required to run scheduled international flights, have instigated data sharing initiatives in several areas of its operation.

The Airline Industry Data Model (AIDM) and a suite of standards that respond to different functional activities within aviation, including AIDX (Aviation Information Data Exchange), have been developed to respond to challenges of inconsistent data definitions.

The AIDX project saw the upgrading of the messaging standards development capability between aviation operators. The development approach aligned to a new methodology which was owned by a board and delivered through working groups with industry wide representation. The methodology followed the same steps set out here and helped the industry to increase data consistency and a faster deployment of standards.

Across aviation there is no one single standard to align to but multiple standards for varying use cases and the AIDM has become the single point of access for industry agreed vocabulary and data definitions. This approach could be mirrored by VirtualES with the schema registry and validation being a centralised point against which multiple standards can be assessed.



2.6

Developing the VirtualES through a use case driven approach

Developing the VirtualES through a use case driven approach

Use cases enable a value driven, outcomes focussed approach to developing VirtualES

Defining

To develop the required technology stack for the VirtualES, a use case driven approach is being adopted to ensure that the required components are delivered iteratively through user requirements. This will provide continuous, incremental development in functionality, whilst demonstrating value from the outset.

In the first instance three related, high-value system-level use cases have been selected to start development of the VirtualES. Over time, additional use cases will be implemented. To deliver these a design thinking approach has been recommended. Design thinking is a user-centric design methodology focussing on the needs of the user based on the following principles:

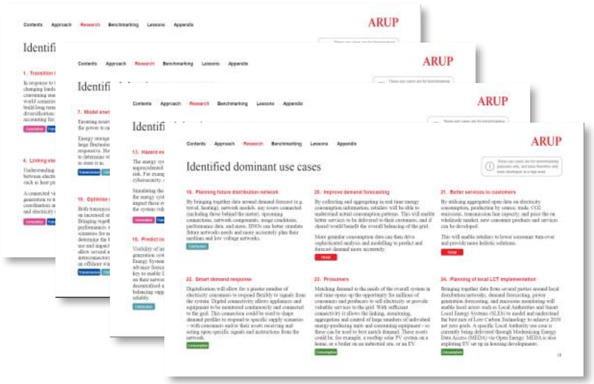
- **Empathise**: Understand the needs and perspectives of potential users of the VirtualES.
- **Define**: Define potential use cases, the scope, and the goals that need to be achieved.
- **Ideate**: Generate wide range of ideas and solutions.
- **Prototype**: Create tangible and testable solutions that can be quickly evaluated and refined.
- **Test**: Test the prototypes with users or stakeholders, gather feedback, and iterate based on the feedback.

Approach taken to identify use cases

The VirtualES is taking a use case driven approach.

Initial use cases were identified through research, expert interviews, industry-wide engagement, and validation through the Advisory Groups. These are summarised in the VirtualES common framework benchmarking report.

Further use cases have been identified by the sector and are being developed. It is considered that, as the VirtualES develops, users will be able to implement their own use cases building on lessons learned and outcomes from its implementation and increasing the breadth and coverage of its capability



Prioritising

By applying design thinking, the VirtualES can ensure that the solution is user-centred, cost-effective, and innovative.

To deliver value early, the use cases need to be prioritised by the sector based on value and development complexity.

- **Value** – how important is this use case to the users of the VirtualES, and to what extent does it contribute towards achieving net zero.
- **Complexity** – how difficult is this component to deliver, and are there any other barriers, such as legal, regulatory or technical limitations, that might make this use case expensive, slow, or costly to deliver.

The first three use cases have been selected to illustrate the complexity and the future needs of these initial ESO use cases.

Delivery of the use cases has been sequenced to iteratively build additional functionality for the VirtualES.

Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

Initial use cases co-created with the sector

Summary of three of the use cases being developed that were co-created with the sector

Overview

There are several use cases that are already under development which have been co-created with the sector.

Three of these use cases have been summarised in this section to illustrate the breadth and coverage the VirtualES is currently considering.

Use case	Description	Value
Common Framework Demonstrator	The use case considers the changing patterns of energy generation and demand and the need for a flexible grid that can be optimised to, for example, reduce the curtailment of renewable energy sources and facilitate bi-directional power from increased use of PVs and EVs.	The use case demonstrator ensures quick wins through improvement of existing operations. It requires minimal investment, and no changes to the existing regulatory landscape.
CrowdFlex	In addition to the changing patterns of energy generation, the CrowdFlex use case introduces the need for capabilities to enable domestic flexibility. This includes having the data licencing and standards to be able to interact with domestic consumers	The CrowdFlex use case will give the domestic consumer access to flexibility market delivering greater value from low-carbon technology, while supporting a more stable, secure system.
Advanced Dispatch Optimiser	Advanced Dispatch Optimiser (ADO) use case further builds on the capabilities to support an increase in the number of participants in the balancing mechanism to include domestic households.	ADO takes the flexibility scenario further in support a greater range of participation in flexibility markets.

Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

Use case - Common framework demonstrator

An overview of the use cases being developed that were co-created with the sector

Description

This use case will practically demonstrate the value of the fundamental components of the VirtualES and document processes and learnings to support effective delivery of future use cases. The use case* explores the opportunity to re-route electricity between grid supply points (GSPs), in certain configurations, by using the existing infrastructure commonly used for maintenance. This will enable greater flexibility and resilience in the network.

Currently this reconfiguration requires weeks of planning and agreements in advance, through the outage planning processes of the Grid Code and System Operator / Transmission Owner Code. Data is shared as part of the 'week 24' planning process and consists of peak load and generation data profiles. However, the data is only a snapshot in time, shared annually, providing limited granularity and doesn't support flexible operational use of this capability.

The demonstrator will explore improvements to the existing processes by developing the tools to enable more granular data sharing between the TNO, DNO and ESO facilitating the more regular re-routing of electricity between GSPs.

Value

The demonstrator will deliver greater flexibility in the network supporting a more stable and resilient system. It will ensure quick wins through improvement of existing operations, requires minimal investment, and no changes to the existing regulatory landscape.

Features

The demonstration will deliver core, high value features of the VirtualES providing the initial building blocks to test the concept and enable further development.

See the table for more information of the features.

Priority	Component	Justification
Must have	Schema assurance	Ensures that the correct, and up-to-date data schemas pertaining to base network models are shared.
	Data portal	Provides a user interface and link to access VirtualES, where search and discovery of the data is enabled by the data catalogue (which is part of the data portal).
	Security permissions	Provides a mechanism for participants to explicitly set permissions for whom to share their data with.
	Trust framework	Provides the mechanism to enable the sharing and access of operational and network data by ensuring policies and access control permissions are met.
Should/ could have	Data broker	A distributed data broker to enable sharing of the data through pub/sub messaging and streaming using event-driven architecture.
	Data licencing	Affords licensing and legal protection to data consumers and producers.
	Data catalogue	Enables search and discovery of the data by using a metadata store by authorised users.

* In current phase a gas use case has also been considered for future development

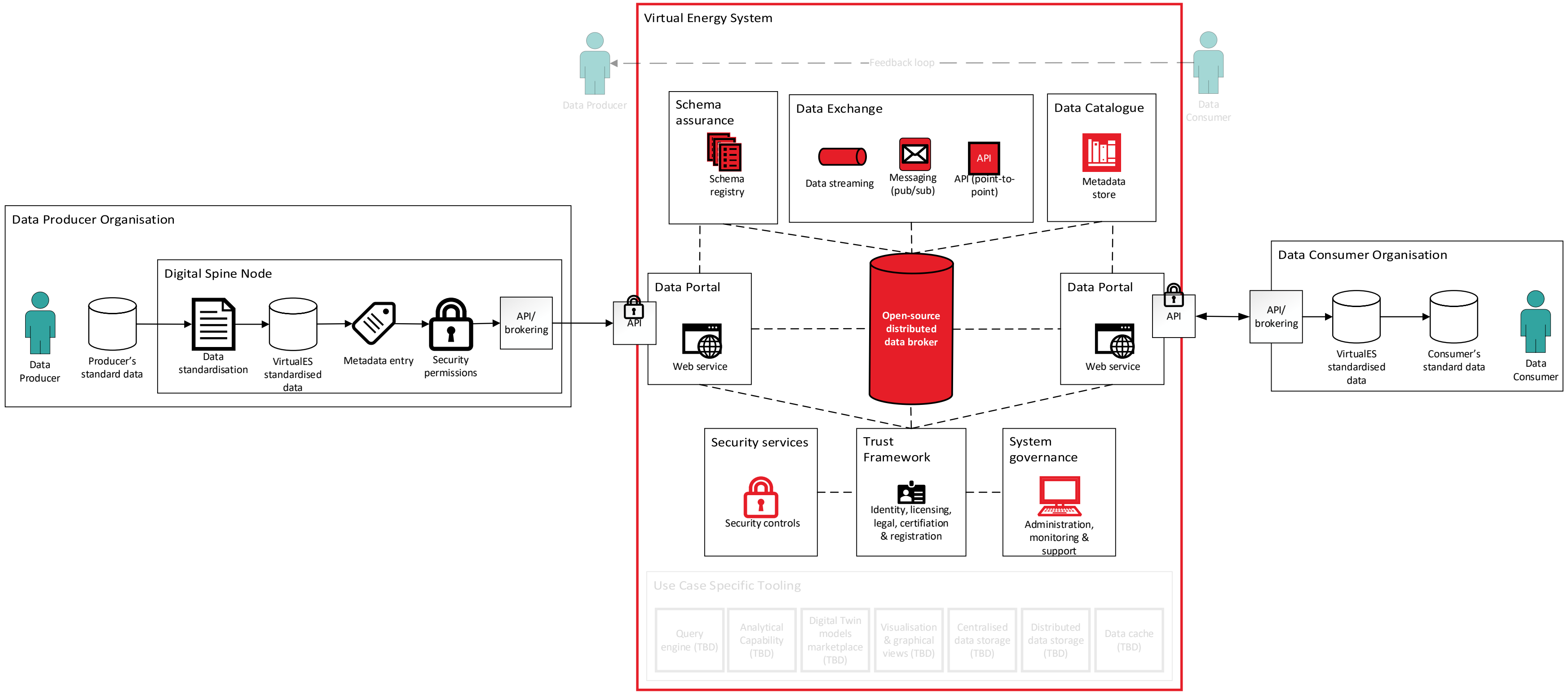
Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

ARUP

Use case - Common framework demonstrator

Components of the high-level design being developed through this use case



MoSCoW* for components

Must Have
- Schema Assurance
- System Governance
- Data Portal
- Security Permissions
- Trust Framework

Should/Could Have
- Data Broker
- Data Licensing
- Data Catalogue

Won't Have
- Analytical Capability
- Visualisation & Graphical Views
- Feedback Mechanism
- Digital Twin Model Marketplace
- Data Standardisation
- Data Cache
- Query Engine
- Distributed Storage
- Centralised Storage

** MoSCoW = Must-have, Should-have, Could-have, Won't-have prioritisation*

Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

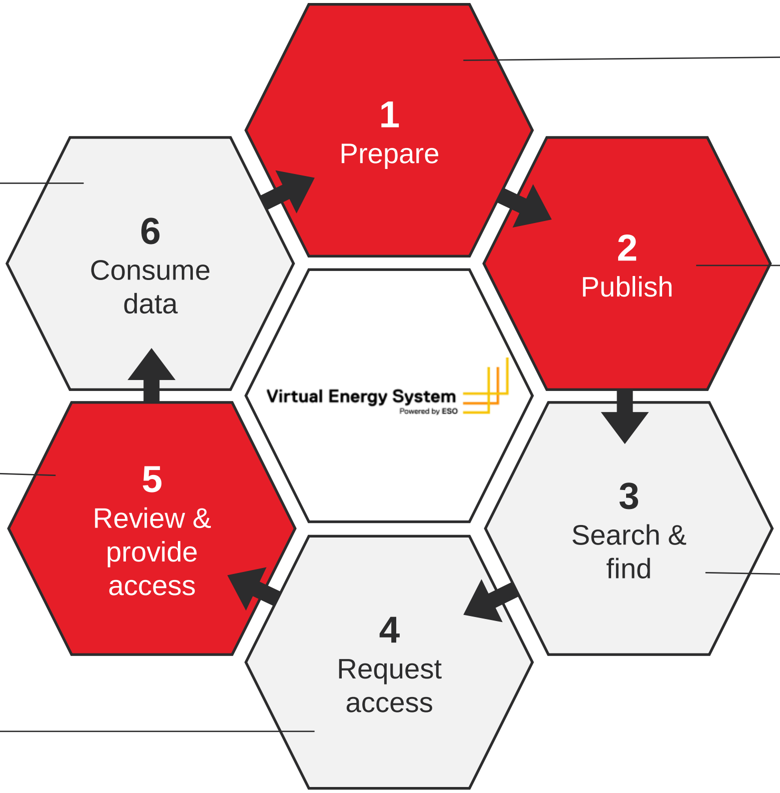
Use case - Common framework demonstrator

Indicative user journey for the use case

6. Consume Data
The base model is now available via the data brokers and API endpoints to the requesting data consumer. The security controls and trust framework ensures the secure and compliant sharing of the data. The data will enable the data consumer to assess the feasibility of rerouting between GSPs (in their own modelling software).

5. Review & Provide Access:
Base model provider is notified of the request for access, and after review approves the requestor. That access control policy is then updated to enable the exchange.

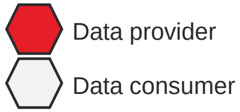
4. Request Access:
If the data consumer is already validated this step is not necessary to request access to the base model. For demonstration purposes this will be assumed to be necessary in this scenario.



1. **Prepare:**
The data owner e.g. TNO/DNO network planner converts base network model into suitable agreed format (e.g. CIM).

2. Publish:
The data owner can now publish this dataset into the VirtualES, publishing the dataset with appropriate security configurations and data handling characteristics. This is validated to ensure standard compliance and is published and made visible within the data catalogue for DNOs to see.

3. Search & Find:
A data consumer e.g. a network planner from a different DNO, searches the data portal for the base model presented in the data catalogue using key words, filters etc.



Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

Use case - CrowdFlex

An overview of the use cases being developed that were co-created with the sector

Description

Building on two pioneering projects (CrowdFlex: NIA and the Domestic Reserve Scarcity Trail); CrowdFlex will clarify the role domestic flexibility can play in addressing the system challenge of decarbonisation.

As more renewable generation is added to the network, it will become increasingly difficult to balance supply and demand. Domestic flexibility provides a huge opportunity during this energy transition to build a smart flexible energy system by enabling consumers to act as a new source of flexibility on the network.

The VirtualES will incrementally build on the demonstrator use case to include additional capabilities improving coordination across the network, while empowering consumers to be active players in reducing their energy bills via new tariffs and incentives.

CrowdFlex will support domestic flexibility by targeting four system needs:

- Frequency
- Thermal
- Adequacy
- Within-day flexibility

Value

The CrowdFlex use case will give domestic consumers access to domestic flexibility market, delivering greater value from low carbon technology, while supporting a more stable, secure system.

Features

The evolution from use case demonstrator to enabling crowd flex requires the VirtualES to scale, increase availability, and performance. This will require additional feature (see table below).

The use of the VirtualES will be for data exchange only with the modelling and analysis of datasets completed in the existing tools and systems following similar steps as outlined for the demonstrator use case.

Priority`	Component	Justification
Must have	CrowdFlex "Must have" components are the same as the common framework demonstrator use case	
	Data licensing	Data licencing affords licensing and legal protection to data consumers and producers, this will be important as consumer interact and share data.
	Data standardisation	Additional actors to the use case, introduces the need for a mechanism that ensures data is standardised to a common standard, such as CIM. This functionality could be provided by the digital spine (subject to an ongoing feasibility study).
Should / could have	**Data cache**	Caching will be required to store information that is readily required, such as: current grid conditions, energy prices, and energy usage patterns in a performant way.
	Query engine	The query engine provide system operators the ability to query and extract specific information on consumer behaviour, or system supply/demand. This will help system operators to make faster decisions based on consumer responses.

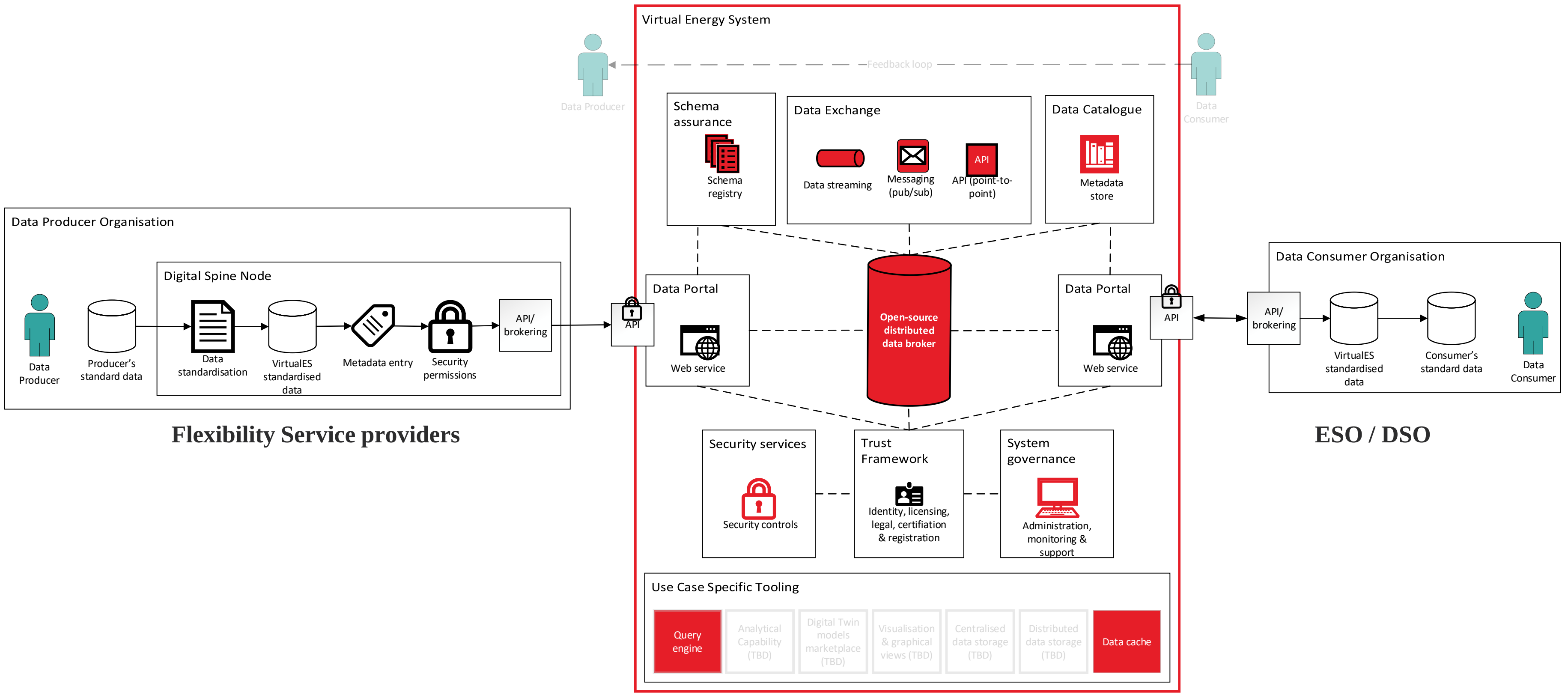
Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

ARUP

Use case - CrowdFlex

Components of the high-level design being developed through this use case



** MoSCoW = Must-have, Should-have, Could-have, Won't-have prioritisation*

MoSCoW* for components

Must Have

- Data Broker
- Schema Validation
- Trust Framework
- System Governance
- Data Portal
- Security Permissions

Should/Could Have

- Data Licensing
- Data Cache
- Data Standardisation
- Query Engine
- Data Catalogue
- Analytical Capability
- Visualisation & Graphical Views

Won't Have

- Feedback Mechanism
- Distributed Storage
- Centralised Storage
- Digital Twin Models Marketplace

Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

Use case - ADO (Advanced Dispatch Optimiser)

An overview of the use cases being developed that were co-created with the sector

Description

This project will research best practices globally, and the advanced technologies available (or being developed), to assess the feasibility of developing an advanced dispatch optimisation tool for the Balancing Mechanism (BM) and will build on CrowdFlex to allow for faster responses to flexibility.

The Advanced Dispatch Optimiser (ADO) will enable the BM to analyse and react to large quantities of data and use this to make complex decisions that are optimised for system security, reliability, cost and decarbonisation. BM is a "pay as bid" short term energy market, run by ESO in the last hour before real time. BM is ESO's primary means of balancing the system. The ADO will aim to increase the number of BM participants to include domestic households, giving the operator additional areas to procure flexibility.

The ADO should evaluate extraneous variables and be 'future proof' for a rapidly changing energy landscape, which includes more diverse providers, services and dynamic parameters across the system.

Value

It enables an increased number of balancing mechanism participants to improve flexibility, deliver better value for money to users, meet the increasing flexibility needs of the grid as it transitions to net zero, and support future proofing of the tools available to users.

Features

Enabling the ADO and future use cases will require further improvement to "future proof" the VirtualES through robust governance processes, and feature improvements.

These features are a snapshot in time and can change during design and requirements gathering phases of product development.

Priority	Component	Justification
Could Have	**Visualisation & Graphical Views**	Visualisation capabilities to enable participants to understand and consume data effectively. This may be provided through nodes & relationships of datasets
	Distributed Storage	Enables storage of data in a performant, secure and scalable way for analytical and simulation use cases. This may also be used to store large volumes of datasets or metadata which may need historical and aggregated views. A data broker can be used in conjunction to publish changes to the databases, where the changes are consumed dynamically.
	Centralised Storage	Enables storage of data for analytical purposes e.g., data warehousing. This also be used to store historical datasets that are accessed by multiple consumers.
	Feedback Loop	Provides a mechanism to enable data consumers and producers to provide feedback on datasets, digital twin models and analytical workloads.
	Analytical Capability	Analytical capabilities and workloads to derive insight from the participant's data.
	Digital Twin Models Marketplace	Marketplace service enabling users to consume complete digital twin models.

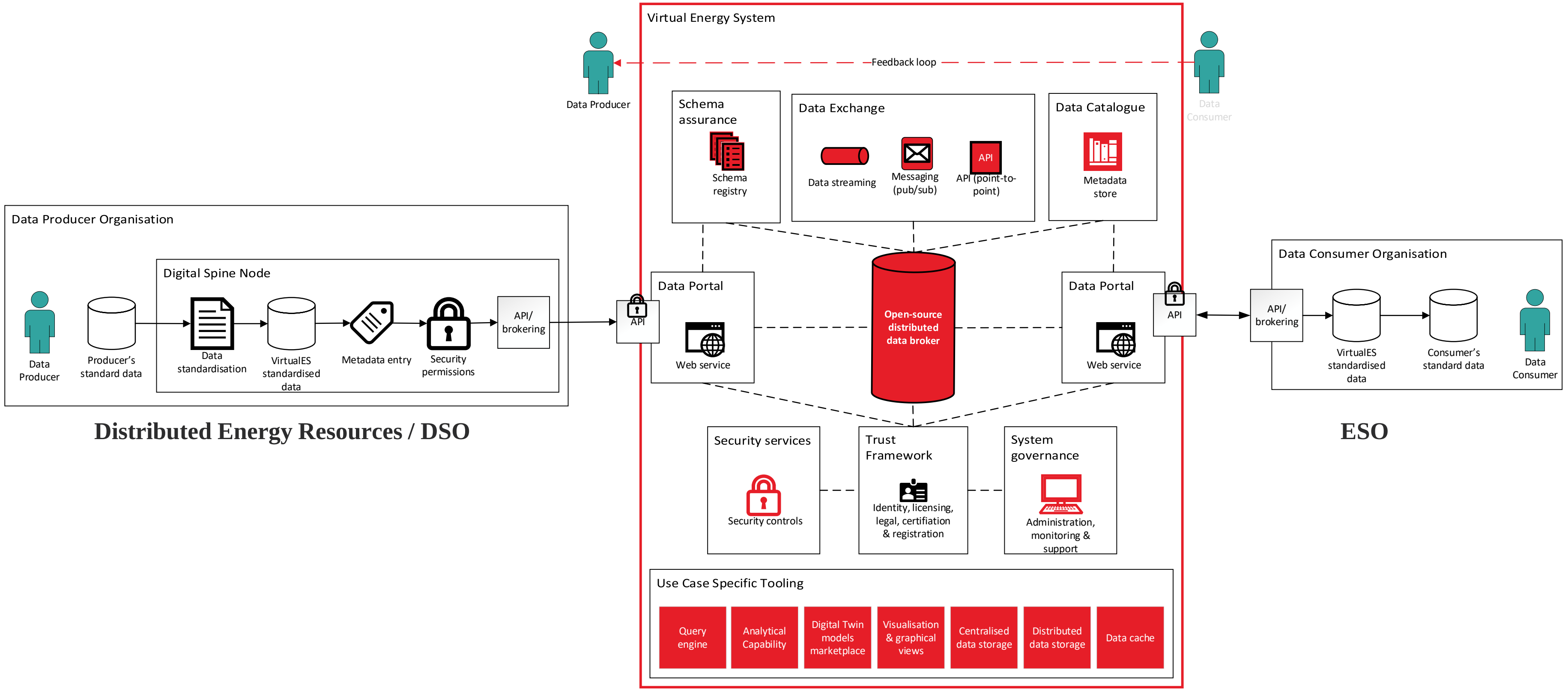
Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

ARUP

Use case - ADO (Advanced Dispatch Optimiser)

Components of the high-level design being developed through this use case



Distributed Energy Resources / DSO

ESO

** MoSCoW = Must-have, Should-have, Could-have, Won't-have prioritisation*

MoSCoW* for components

Must Have

- Data Broker
- Schema Validation
- Trust Framework
- System Governance
- Data Portal
- Security Permissions

Should/Could Have

- Data Catalogue
- Data Licensing
- Data Cache
- Data Standardisation
- Query Engine
- Visualisation & Graphical Views
- Analytical Engine
- Feedback Mechanism
- Centralised Storage
- Distributed Storage
- Digital Twins Model Marketplace

Won't Have

- N/A

Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

3

Call to action

Your input to support technology developments

A collaborative approach to sector buy-in

Overview

To develop the VirtualES further, use cases are needed to ensure alignment with the sector needs.

A mature VirtualES will support dual-fuel use cases and prove the blueprint for a wider national infrastructure integration. Therefore, the use cases that are prioritised will pave the roadmap for a wider sector implementation, and adoption.

Some potential use cases that can be enabled through the integration of VirtualES are city decarbonisation, local net zero planning, identifying consumers in vulnerable situations, testing potential electricity market structures and forms, and coupling across different power sectors.

This is a call to action to provide your input into potential use cases, and feedback on the proposed technology functional architecture.

Co-development

Co-development starts with identifying use cases, and on-boarding potential collaborators, sector champions, and priority stakeholders.

We are eager to listen and learn from your insights. A starting point is the review and feedback on the below proposals.

- Input into the demonstrator wireframes.
- Suggestions into assessment criteria for selecting possible service providers and/or tools which could support the delivery of the demonstrator.
- Apply to join our advisory groups.
- Discussions to align on other on-going initiatives.
- Understanding of on-going initiatives that are already under-way which can become dependencies for the VirtualES development and steady state.
- Engagement sessions in identifying key stakeholders developing ontologies and data standards relevant to the demonstrator use case, and the common framework.

Input into considerations

The technology considerations outlined in Section 2.5 are influenced and supported by input from in-sector and cross-sector subject matter experts.

We would value your specific input into the technology considerations and into the wider co-development opportunities.

Please share your feedback:
VirtualES@nationalgrideso.com

Visit our website to read the published reports.

Visit our website

Join the mailing list to be invited to future events or receive further communications.

Join the mailing list

Key factors applicable to:

Aligning models & taxonomies | Increasing data visibility & enabling sharing | Creating an interoperable tech stack

ARUP

Virtual Energy System

Powered by ESO

VirtualES@nationalgrideso.com