Virtual Energy System

Common framework

Interoperability report & data sharing framework

April 2023

Contents

Executive summary

Recommendations on the demonstrator data needs and gaps

Background

ESO have launched the 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. 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 today.

Following the example set by the National Digital Twin programme and the Digital Twin Hub through their Climate Resilience Demonstrator project (CReDo), the VirtualES is developing a demonstrator that is initially focused on a flexibility use case, which is an electricity network use case .

This document contributes to the development of this demonstrator, currently being progressed through an NIA-funded project in Alpha phase. Its purpose is to assess the current data landscape, determine the demonstrator data needs and identify the appropriate standards to facilitate data sharing between operators.

Approach

This report follows on from the data needs assessment and technology review and builds on the findings identified.

Specifically it identifies the key interactions between users, data and technology and the processes for enabling the use case; it reviews the technology required to enable the use case; and sets out a data sharing framework that can be adopted for the demonstrator.

Additional stakeholder engagements and desk research has contributed to that already undertaken as part of the previous deliverables and has helped fill gaps in knowledge.

This has enable us to develop new artifacts including user journeys, process maps, and a data sharing framework.

This interoperability report (Section 2) discusses the flexibility use case, which is an electricity network use case. Whilst the technology (WP2.2) is applicable to both electricity and gas use cases, this report focuses on the electricity use case technology requirements (see Section 2.1).

Recommendations

This report identified several recommendations to be considered in the next phase. These are applicable to both electricity and gas use cases.

The full **interoperability** recommendations are given in Section 2.3

The key recommendations for the next phase are:

- Define full functional, non-functional security requirements for the VirtualES to steer the design choices, and design feature.
- Outline a data standard for the demonstrator to inform scheme validation checks, schema registry, and metadata.

The full **data sharing agreement** recommendations are given on page 38. The key recommendation is:

- It is considered that a common approach to agreeing a data sharing framework is required. A new rubric which factors in process and decision making and contracting is required. A proposed methodology, which aligns with Data Bill, and objectives of the smart data council, for developing this approach is given on page 39.

Nomenclature

ABAC – Attribute Based Access Control
API – Application Programming Interface
CReDo – Climate Resilience Demonstrator
ESO – National Grid Electricity Systems Operator
GSP – Grid Supply Point
HLD – High Level Design
NIA – Network Innovation Allowance
RBAC – Role Based Access Control
SIF – Strategic Innovation Fund
SLA – Service Level Agreement
TNO – Transmission Network Operator
VirtualES – Virtual Energy System
WP – Work Package

1 — Approach

Context

What is the Virtual Energy System?

The Virtual Energy System

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.

The VirtualES is delivered through three workstreams:

- Workstream 1 - Stakeholder engagement
- Workstream 2 - Common framework & principles
- Workstream 3 - Use cases

Workstream 2 - Common Framework & Principles

This report forms part of workstream 2.

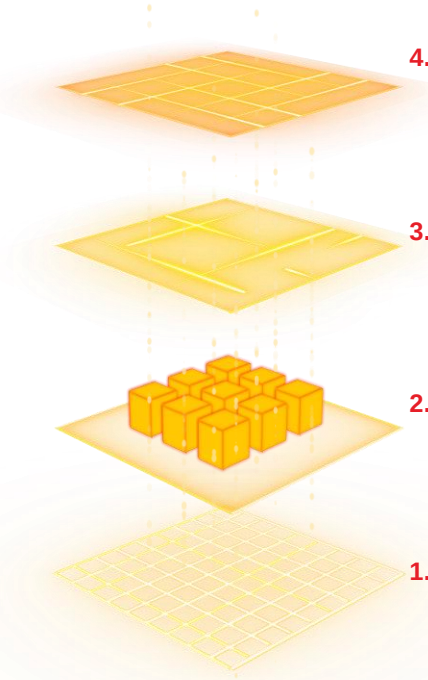
The objective of this workstream is to develop the socio-technical common framework that will form the foundation of the VirtualES – enabling the creation of this ecosystem of connected digital twins.

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 factors are grouped by the categories of People, Process, Data, and Technology. Six of these factors were prioritised based on their potential impact on the VirtualES objectives and their relative maturity across the wider energy sector.

Following the example set by the National Digital Twin programme and the Digital Twin Hub through their Climate Resilience Demonstrator project (CReDo), this workstream is now developing a demonstrator that is focused on a *whole-system flexibility* use case.

This document contributes to the development of this demonstrator, currently being progressed through an NIA-funded project in Alpha phase.



4. The data becomes more layered, these interactions will create valuable insight to help guide and govern how we generate, manage, store, and consume energy.

3. Populated by existing and new digital twins – replicas of physical components of our energy system

2. Each digital twin will contribute to and access real-time data on the status and operation of other elements of the system

1. A social-technical common framework, with agreed access, operations and security protocols

Virtual Energy System

Indicative components of the Virtual Energy System

Developing a common framework

Published research and reports for the common framework

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 the six reports completed to date all published online.

Following the SIF Discovery project (report #3), the demonstrator was further developed using the whole-system flexibility use case (report #4).

The demonstrator is currently progressing through an NIA-funded project in Alpha phase, and is being delivered in line with the project plan (report #6).

1. External benchmarking

Understanding the cross-sector and global best practice for connecting assets, systems, and digital twins.

Read the report

2. Defining the common framework

Determining the key socio-technical factors that need to be considered for the VirtualES to succeed. See the next page for more information.

Read the report

3. Demonstrating the common framework

Collaboratively prove and demonstrate, with industry, how the socio-technical principles work.

This was a Round 1 SIF Discovery project.

Read the report

4. Whole system flexibility use case definition

Further define the “whole-system flexibility” use case that is recommended as the initial use case to demonstrate the common framework.

Read the report

5. Demonstrator data standards, data portals, and data licensing

Identified data standards and outline data licensing considerations applicable to the use case. Initial review of currently available public energy sector ‘data portals’.

Read the report

6. Demonstrator project plan & advisory groups

Proposed delivery plan, governance structure, advisory groups approach, and cross-workstream collaboration that will enable the successful delivery of the demonstrator.

Read the report

Socio-technical factors

14 factors to develop the common framework

As detailed on the previous two pages, the defining the common framework report (report #2 on the previous page) identified 14 socio-technical factors which are considered necessary for the development and delivery of the VirtualES today.

These factors were derived through research, expert interviews, and industry-wide engagement. They are shown in the adjacent diagram, and are grouped by the categories of People, Process, Data and Technology. The titles of the factors intentionally include verbs, making their framing actionable.

These 14 factors were prioritised to highlight the six factors recommended for immediate consideration.

Best practice guidance notes are being developed for the six priority factors as part of WP3.

Delivery 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.

Introduction

Purpose of this document

Purpose

This document presents the findings of **WP2.3 - Interoperability and data licensing requirements**, developed as part of the common framework demonstrator Alpha phase.

This document contains the following deliverables:

- Interoperability report (M4)
- Data sharing framework (M5)

Electricity and gas network use cases

This NIA-funded Alpha phase is supported by ESO and National Gas. The objective of the VirtualES is to include and consider both the electricity and gas.

The user journeys discussed in this interoperability report are for the flexibility use case, which is an electricity network use case. Whilst the technology (WP2.2) is applicable to both electricity and gas use cases, this report focuses on the electricity use case requirements.

In recognition of the future energy system, a separate demonstrator use case is recommended for the gas network, with separate user journeys will be required.

Interoperability report – electricity use case

This report expands on the work conducted as part of the data assessment and technology review (WP2.1/WP2.2).

The purpose of this report is to explore in greater detail the specific interactions between users, data, and technology.

It builds on the findings of the electricity use case data needs assessment (WP2.1) and the underlying need for sharing of data to enable the use case. It explores the process of sharing a base network model and operational scenarios between operators, and the process of accessing, merging models, and running scenarios.

Through the use of user journeys and process maps the report sets out the flow of key activities and the user interactions with the VirtualES. These interactions include the publishing of a base model by a data producer, accessing and merging of base models by data consumers, publishing of an operational scenario by a data producer, and the accessing and merging of operational scenarios by a data consumer.

This report also builds on the demonstrator technology review (WP2.2), expanding the key functionalities required to deliver the user journeys and use case.

Data sharing framework

The significant diversity in approach and the lack of an industry standard data sharing suite of agreements has meant best practice and insights are not socialised throughout the industry.

It is considered that a common approach to agreeing a data sharing framework is required. A new rubric which factors in process and decision making and contracting is required. This report summarises the current data sharing agreement landscape, challenges, and issues, and provides recommendations for address them.

It proposes a methodology which supports a logical and efficient process for getting to a digitalised smart data sharing agreement. A process map for this proposed methodology has been developed and is set out in the relevant section below.

2 – Interoperability report – electricity use case

2.1 — User journeys

Demonstrator use case

Summary of the use case

Overview

The demonstrator is based on the published VirtualES flexibility use case definition (an electricity network use case).

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 explores the opportunity to re-route electricity between grid supply points (GSPs), in certain configurations, by using existing infrastructure commonly used for maintenance.

Changing the network running order in this way would enable demand or generation to be moved between different locations, providing an example of achieving flexibility through a location shift.

In instances of planned network outages, this bypass can re-route electricity from adjacent GSPs to provide resilience to the network. This will transfer all or part of the load from one GSP to the other, while keeping an electrical split. Or connect the two GSPs to operate as an interconnected group.

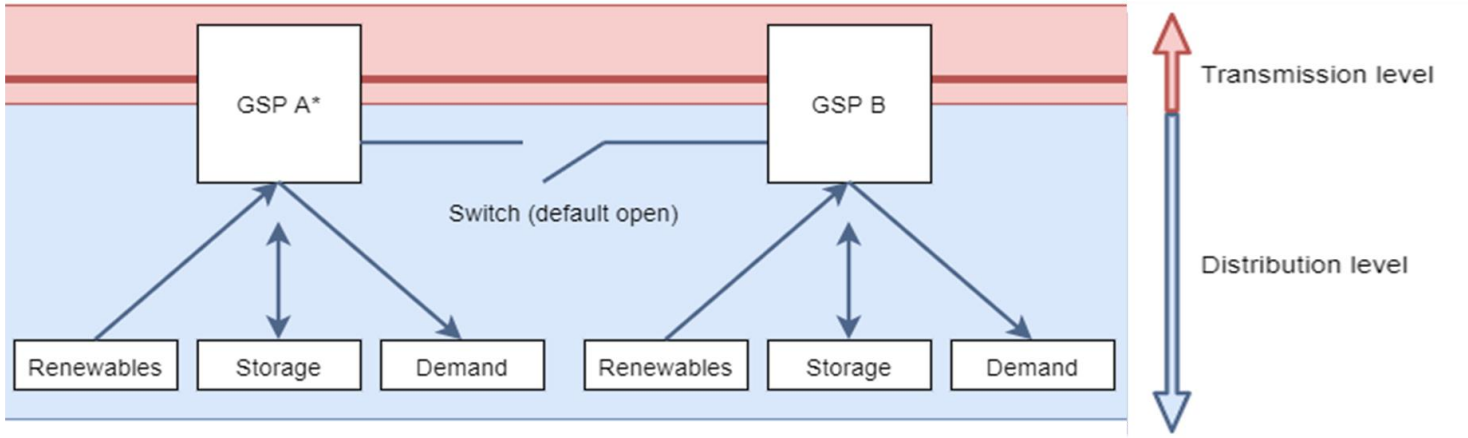
This reconfiguration currently requires weeks of planning and agreement in advance, through the outage planning processes of the Grid Code and System Operator / Transmission Owner Code.

Similar considerations in the operational planning process are required for interconnected, loosely coupled or radial GSP configurations, to maximise system availability and minimise system risk. This includes minimising generation restrictions, through an improved understanding of demand behaviour and flexibility services, using GSPs within a zone.

As more renewable generation comes online there are potential advantages to using this connection reconfiguration more actively.

These user journeys only considers the whole system flexibility use case, which is an electricity networks use case.

In recognition of the future energy system, a separate demonstrator use case has been developed for the gas network.



Example GSP configuration (GSPs can be owned by the TNO or the DNO)

ARUP

Demonstrator use case

Summary of the use case and related artifacts

Overview (continued)

A key purpose of this demonstrator is to showcase the feasibility of implementing a technological solution. To constrain the scope, the demonstrator considers the requirements of operational timescales from 3 weeks ahead to near real-time.

Critical to the use case is the assessment of the potential interconnections of GSPs. This requires visibility of the assets involved, their capabilities and the expected behaviour of demand and generation. This assessment is currently carried out by operators through the use of power flow modelling, e.g. PowerFactory.

The use of modelling to determine the impacts of future running arrangements and resolve potential issues is widespread across the energy landscape. Operators develop and run operational scenarios that determine the arrangements of their network. Currently this is done on an organisation-by-organisation basis with minimal data sharing between operators.

Data that is shared, e.g. the “week 24” data submission made to ESO, provides limited granularity of the network at a single snapshot in time, with peak load and generation data profiles. The existing processes do not meet the requirements of the use case.

User journeys and process maps

Following on from the data needs assessment (WP2.1) and technology review (WP2.2) this document explores the user journey and processes required to be in place to enable data sharing between parties for the electricity use case.

This report considers the use case diagrams, data products and technology recommendations made in the previous reports and develops them further to showcase the key user and technology interactions that will enable the sharing of the necessary electricity network models and operational scenarios.

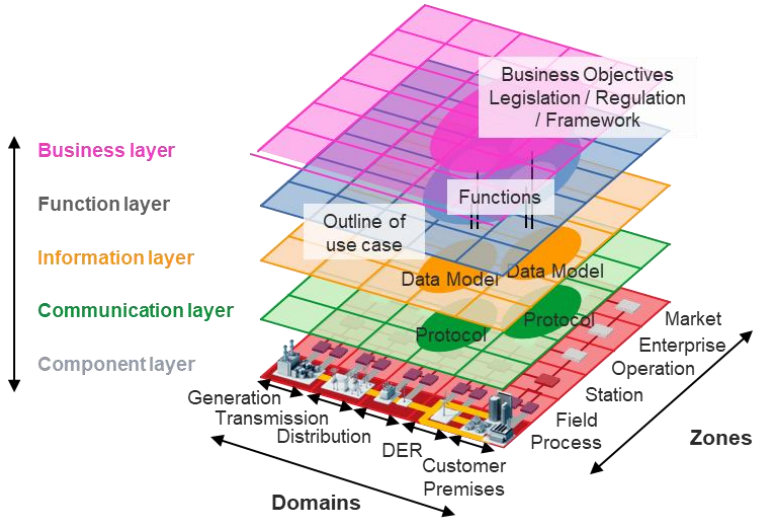
A trigger event has been selected of a known scenario in which the use case of closing normally open switches between GSPs is actioned. The user journey considers the sharing and accessing of data from four key activity areas implemented by two key personas.

This work will establish the critical processes and the interactions between users and the VirtualES technology which will be explored as part of the wireframing report.

The user journeys and process maps set out a logical series of steps to fulfil the use case and establish user requirements.

They are part of the suite of artifacts and assets created to better understand and enable the VirtualES.

User journeys and process maps are a key part of the Smart Grid Architecture Model that define the Function layer and help in further developing the Information & Communication layer.



Smart grid architecture model

Overarching user journey

High level overview of the key activities

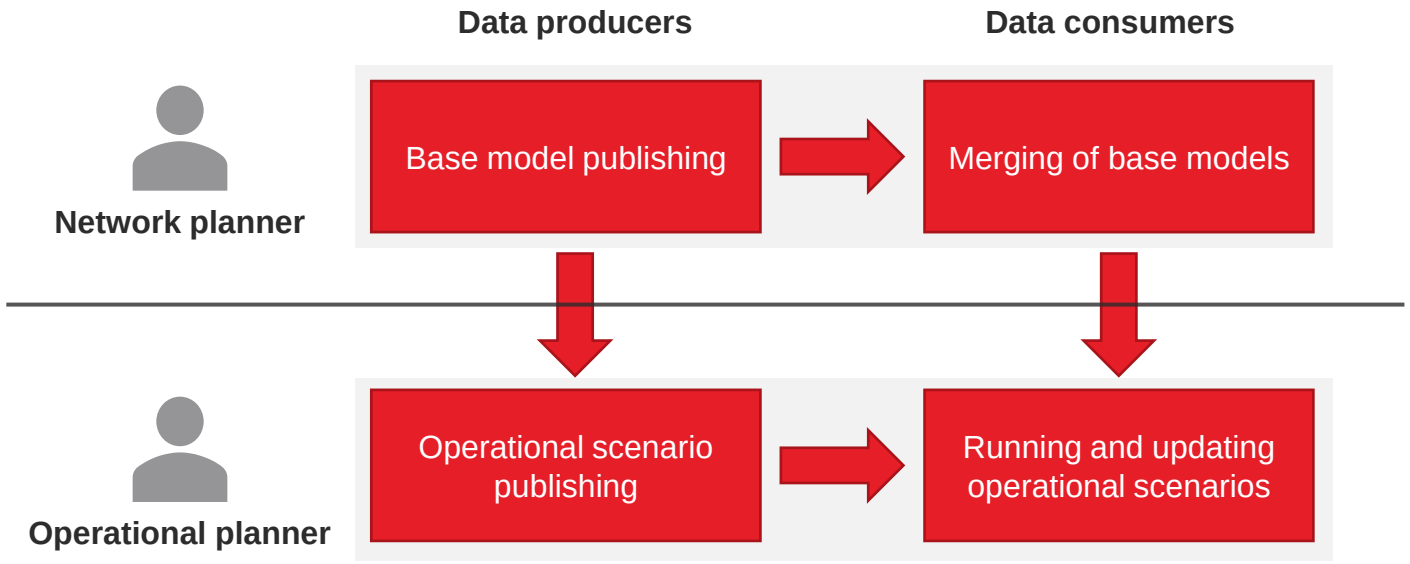
Trigger event and personas

Following the data needs assessment and technology review a user journey has been developed setting out the key steps required in responding to a particular trigger event.

The trigger event considered is one in which an outage on the distribution network requires the DNO to reroute power between GSPs. This outage requires the DNO to establish a new running arrangement and communicate this to relevant parties.

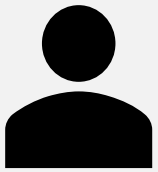
The user journey has been developed from the perspective of two personas. These are the **Network Planner** and the **Operational Planner**. These roles are considered common to all organisation types and as such there is no organisational distinctions made in this report.

- **Network Planners** are responsible for the base network models owned by organisations, and are responsible for accessing and merging base models. They can both be data producers (by sharing their base model) and data consumers (by accessing other organisation's models).
- **Operational Planners** are responsible for developing and running scenarios that respond to the trigger event. They can be both data producers and data consumers, by sharing and accessing operational scenarios.



User journey - base model publishing

User journey for a network planner publishing a base model



Persona: network planner

A network planner wants to publish an up to date base network model that represents the normal running arrangements for the modelling period.

Organisations

DNO - As owners and operators of distribution networks, DNOs will have responsibility for publishing their network models.

TNO - As owners of transmission networks, TNOs will have responsibility for publishing their network models.

Assumptions

- Organisations have an agreed data standard and format that the base model must align to.
- Sharing of a base model is not contingent on a specific trigger event but is a necessary pre activity.
- Organisations have an agreed timeframe against which the base model should be published and updated.

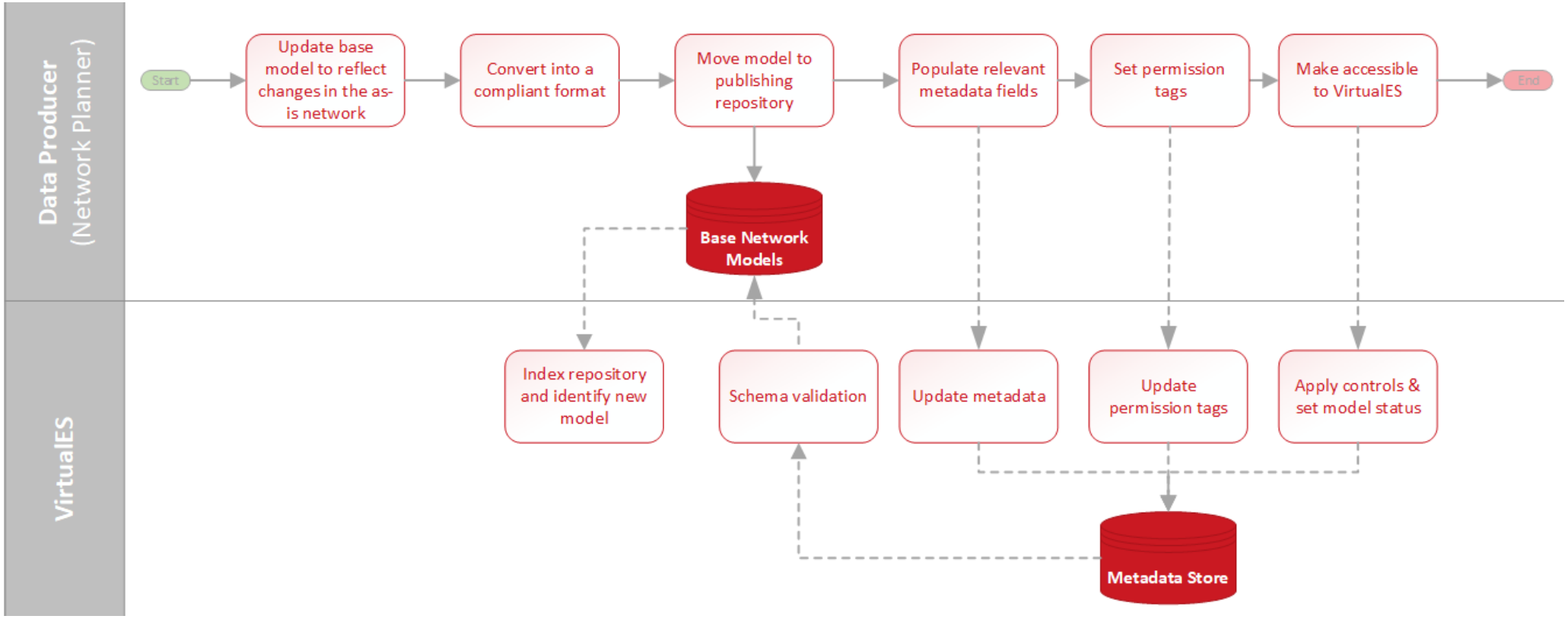
Sequence of activities

	User journey	Key considerations
Update base model	The network planner must update their organisations base model to represent the as-is network. The latest asset & maintenance data as well as any changes to embedded generation and demand capacity should be incorporated.	How operators update their base model with asset, demand and capacity data is an internal process and out of scope.
Convert & make ready	Once all updates have been made the model must be converted in to a format agreed to for the demonstrator. This should include certain data attributes such as a unique network ID. Then the model must be moved to a repository or database that allows VirtualES to identify and index it.	Choosing a standard or data format for the demonstrator is a key requirement. Data will be hosted by data producer on their selected technology and made discoverable through an API.
Populate metadata & apply permission	The network planner would access the demonstrator data portal and would search for and select the base model. Once selected they would update the metadata attributes such as model name and description and set the permissions tags.	The data portal must have the capability to index models hosted by the data producer and make them visible for management.
Make accessible through VirtualES	Once all the required fields had been set the network planner would change the status to make it accessible to others This would trigger the VirtualES to validate the model schema, and following this update the status.	Data must be discoverable and accessible by VirtualES

ARUP

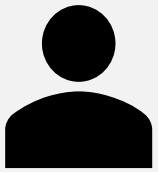
Process map - base model publishing

Process map for a network planner publishing a base model



User journey - accessing and merging base models

User journey for a network planner accessing and merging base models



Persona: network planner

A network planner wants to access base models from other organisations and to stitch one or more models together to create a combined model.

Organisations

DNO - Requires access to the transmission base model and neighbouring DNO models

TNO - Requires access to distribution base models

ESO - Requires access to all transmission and distribution base models.

Assumptions

- Agreed metadata standards and naming conventions allow users to intuitively search for base models and understand the model content.
- Networks can only be joined to those with which they share a physical connection to.

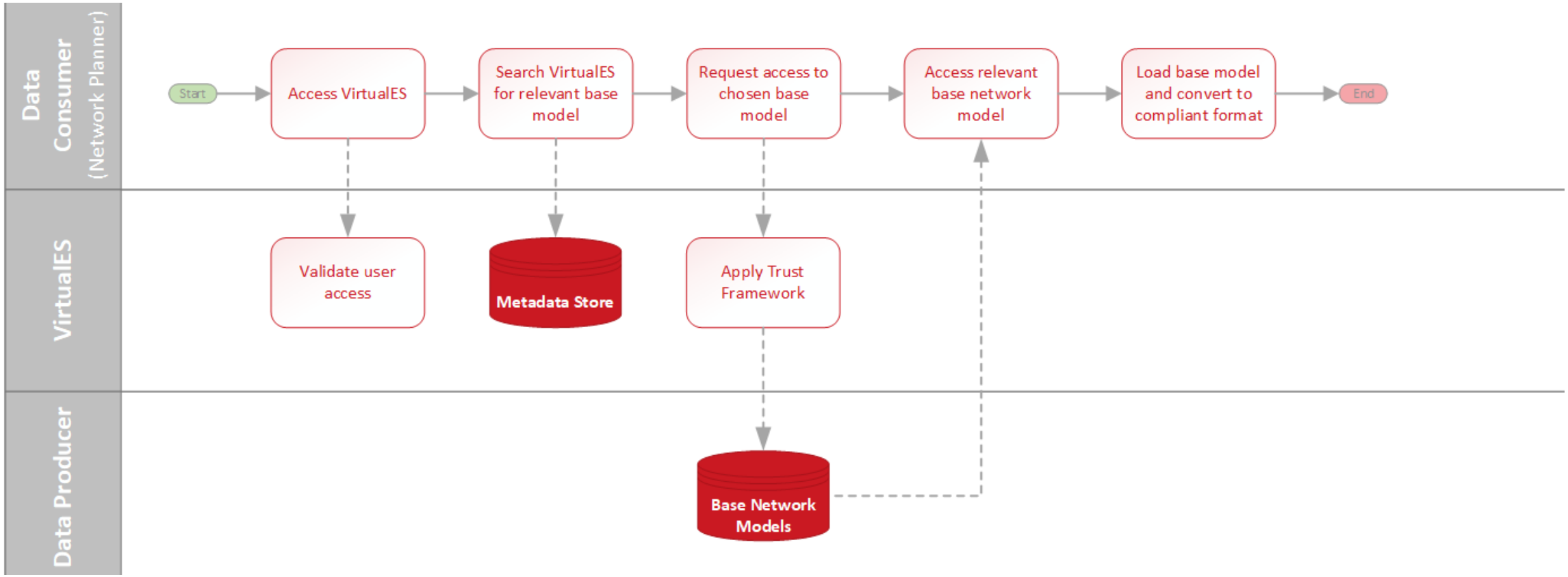
Sequence of activities

	User journey	Key considerations
Search & access relevant base models	The demonstrator will allow approved users, from different organisations, to search and access published base models. The metadata attributed to the model will be critical in allowing users to search and filter on various fields such as network name, network owner, published date, etc. VirtualES, through the use of the trust framework, will validate the user access, and allow them to subscribe to the selected base model.	Consideration should be given to the management of base network models and the archiving of models after a given time period. User access will need to be managed, likely at an organisational level. Consideration should be given to how this is done.
Load model & convert	The user will import the model in to their power flow modelling tool. The model may need to be converted from the agreed format or standard to the application's proprietary standard. Depending on the application's capability this function may be in-built or a separate conversion process may be required.	Power flow modelling tools may have capability to connect directly to APIs. Vendors should be engaged to understand if this capability exists and any constraints that may exist.
Identify boundaries & stitch networks	With a converted network the planner can select to join it with another neighbouring network, this could be their own network or a separate network loaded through the same process. The networks will then be stitched together using common nodes that sit at network boundaries.	Consideration should be given to the network drawing process to ensure no overlap when merged. The identification of network boundary nodes should be an automatic process

ARUP

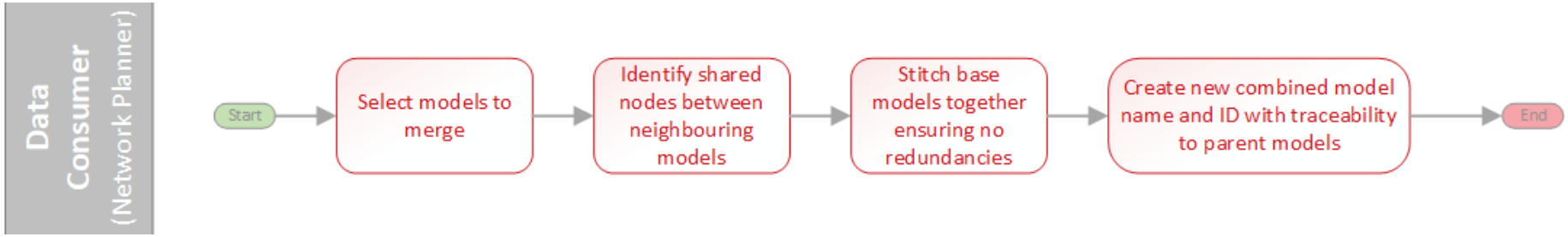
Process map - accessing base models

Process map for a network planner accessing base models



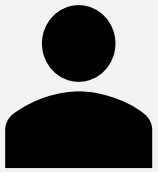
Process map - merging base models

Process map for a network planner accessing base models



User journey - scenario development and publishing

User journey for an operational planner developing and publishing scenarios



Persona: operational planner

An operational planner publishes a scenario with new network running arrangements, rerouting electricity between GSPs in response to an outage

Organisations

DNO - Will publish scenarios setting out proposed running arrangements for the distribution network.

ESO - After analysis of published DNO scenarios ESO may update and republish scenario to DNO & TO to resolve outstanding issues.

Assumptions

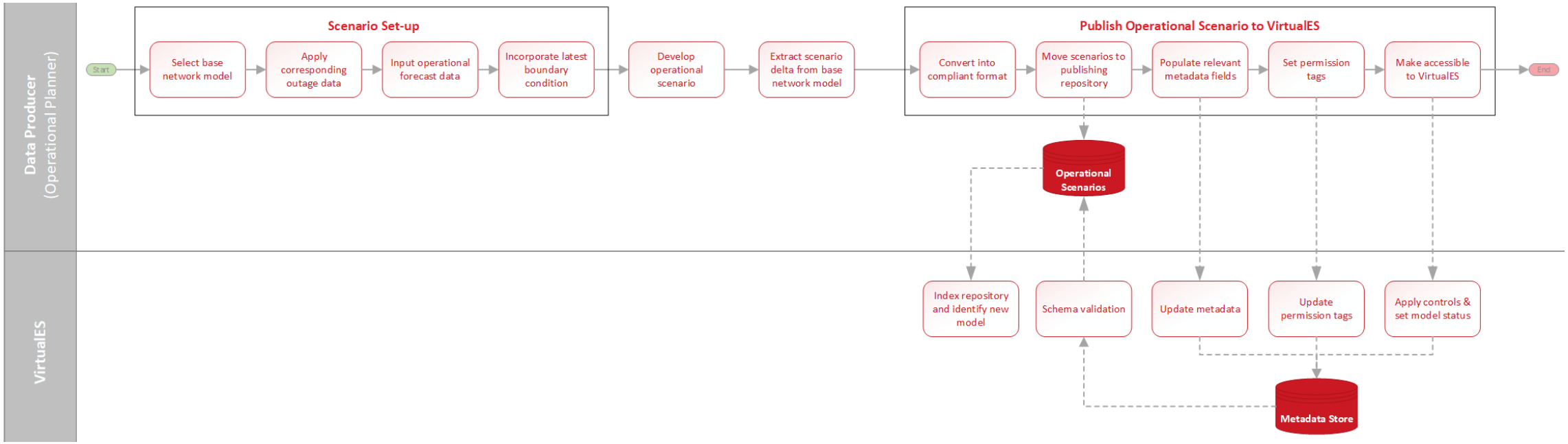
- Scenarios will be derived from a specific base model and this must be captured in the scenarios metadata.
- The forecasted load and generation data developed for the scenarios is considered an internal process and out of scope for this demonstrator.

Sequence of activities

	User journey	Key considerations
Scenario Set-up	For the demonstrator a representative trigger event will be used with a preselected running arrangement and historic load and generation data applied to the scenario. The operational modeller will set up and test the scenario on the agreed base model to ensure it converges.	Alignment on a process for including network outages within scenarios should be considered beyond the demonstrator.
Extract & convert	Having created and tested the scenario the user will then extract the scenario, critically defining the delta between the scenario and base model on which it was created. The scenario will then be converted in to the agreed format and moved to the operator's repository where VirtualES can identify and index it.	Scenario metadata must include base model ID that was used. Process of defining the difference between the scenario and base model is needed. This is ideally an automated process.
Populate metadata & apply permission	The operational planner would access the demonstrator data portal and would search for and select the scenario. Once selected they would update the metadata attributes such as the scenario name and description as well as set the permissions tags for access.	Scenario metadata will need to include information on when the scenario was developed and whether historic or forecast data has been used. If forecast data the scenario horizon should also be included.
Make accessible through VirtualES	Once all the required fields had been set the operational planner would request the portal to validate the scenario and following this update the data will be accessible to others.	Data must be discoverable and accessible by VirtualES

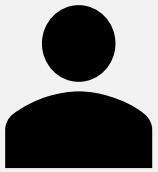
Process map - scenario development and publishing

Process map for an operational planner developing and publishing scenarios



User journey - scenario accessing and merging

User journey for an operational planner assessing and merging operational scenarios



Persona: operational planner

An operational planner wants to access a published scenario and run the scenario on a combined base network to determine impacts to their network.

Organisations

DNO - Will require access to updated scenarios made by ESO and neighbouring DNO scenarios.

TNO - Will require access to scenarios developed by various DNOs

ESO - Will require access to scenarios developed by various DNOs

Assumptions

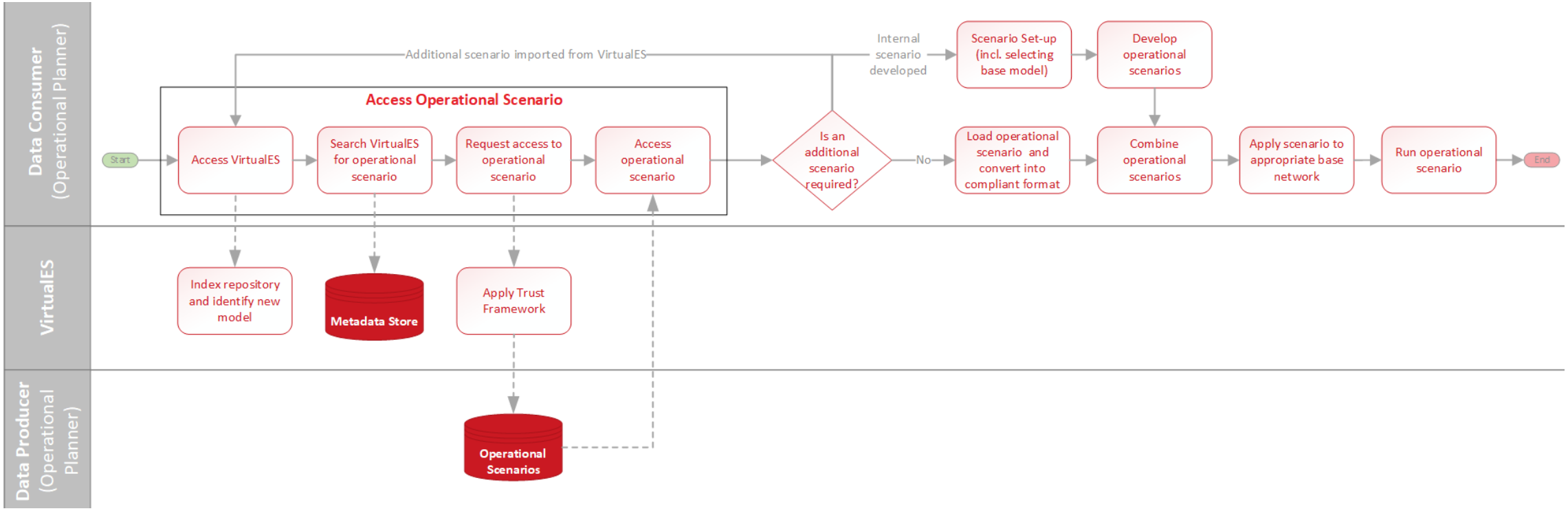
- A scenario must be derived from a specific base model and should only be run on a copy of that base model.
- A merged base model derived from joining the specific base model and another model would satisfy the criteria if the specific base model exists in it's entirety.

Sequence of activities

	User journey	Key considerations
Search, access, load, & convert relevant operational scenarios	The initial process for accessing the available scenarios will be the same as those for accessing the base model. The user will search and download the operational scenario and then import this in to their chosen power flow modelling tool and, if necessary, convert it in to the appropriate format.	Metadata related to the triggering event, such as a scenario description or scenario type, could be used to assess relevancy.
Running & updating the scenario	Once the scenario has been accessed the operational planner may consider merging this scenario with another scenario or running the scenario independently. Either way the operational planner would need to select the base model or merged based model from which the scenario(s) were derived. Scenarios would then be loaded on to the base model to implement network changes and to apply the load and generation data.	A process for merging scenarios will need to be explored. This will depend on the agreed data format and functionality of modelling tools. The use of unique IDs for base models and their nodes and lines should enable scenario data to be integrated with base models.
Updating the scenario & republish	Once the model has been run the user may update the scenario considering the modelled outputs. The updates would be made manually by the user and then saved as a new child scenario. This child scenario could then be shared back to VirtualES with updated metadata reflecting it's lineage.	A process for splitting merged and updated scenarios back to their original constituent parts should be explored so an updated scenario can be run on a original base model.

Process map - scenario accessing and merging

Process map for an operational planner assessing and merging operational scenarios



2.2 — Technology requirements

Overview

Overview of the demonstrator technology High Level Design (HLD)

Overview

Based on the findings of the technology review (WP2.2), an indicative High Level Design (HLD) was created for the demonstrator electricity use case.

The HLD does not represent, nor contain, the comprehensive list of functional and non-functional requirements for the VirtualES. This exercise will need to be conducted at a future development stage.

An illustration of the HLD is provided on the subsequent pages, where a conceptual picture of the data producers and the consumers interacting the VirtualES is provided. The data producers are required to provide metadata and security tags to their data before it shared using secure APIs with the VirtualES.

The data undergoes a schema validation check before it is streamed to the data consumers, and it is subject to governance and security controls, in addition to a trust framework.

The data consumers can search and find the data they are interested in by using a data catalogue as part of a data portal, where they can request access to the data from the producers by using a trust framework to handle the access permissions.

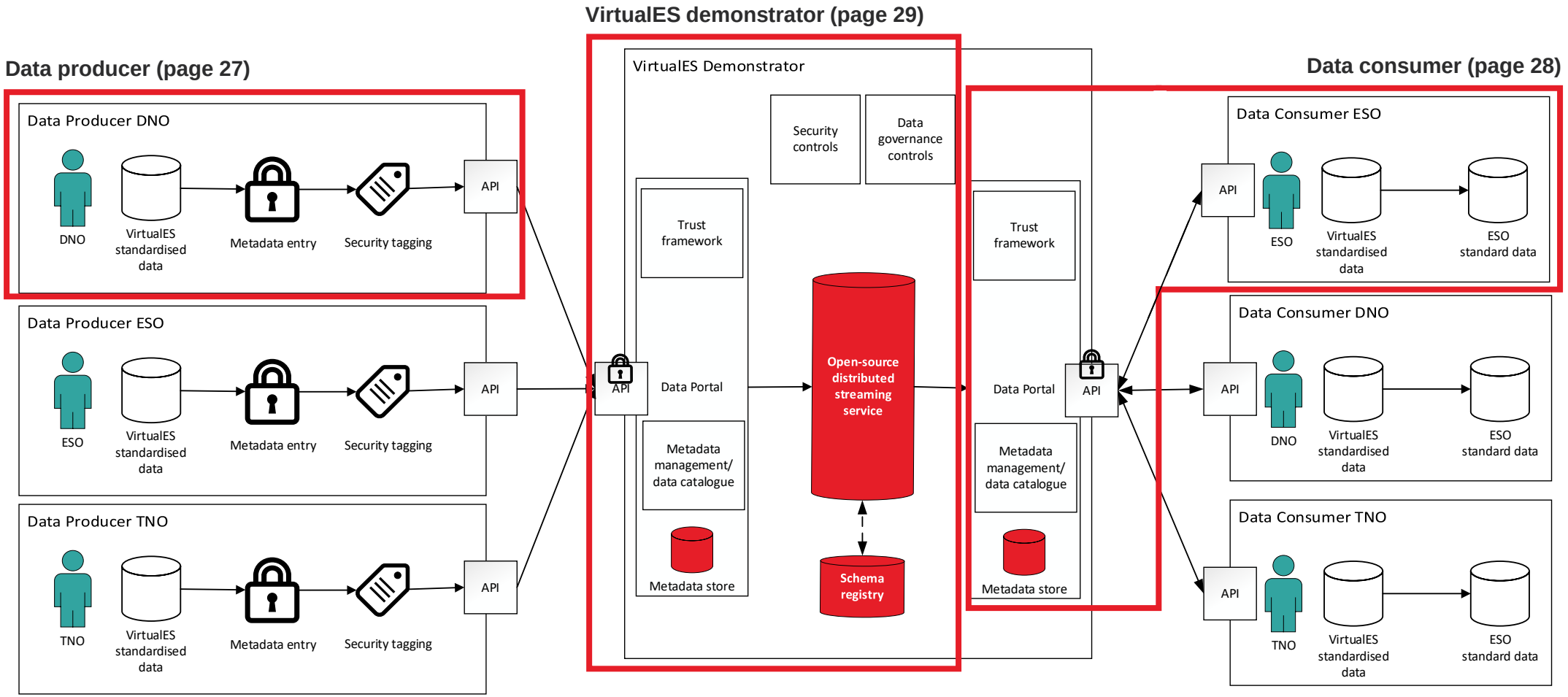
The HLD components are then broken down into individual pictures for the data producers, data consumers, and the VirtualES demonstrator.

Pages 27-29 contain descriptions for each of these components to provide context of how the technology is used to enable the use case and the user journeys described in this report.

The HLD provided in this report is specific to the electricity network demonstrator use case. The intention is to iterate its design and build upon it so that it accommodate future requirements and use cases (such as the gas networks demonstrator use case), where additional functionalities and components may be required.

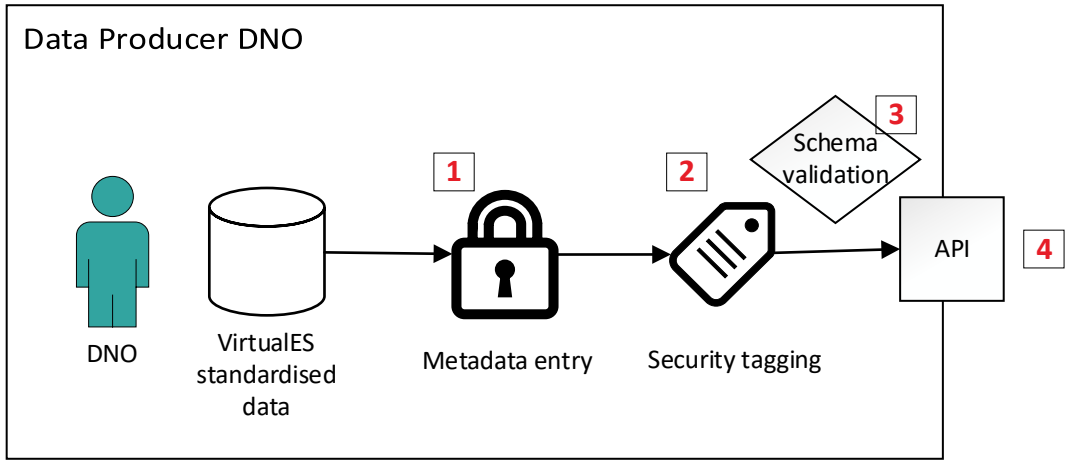
Electricity use case HLD

Proposed HLD for the electricity use case



HLD: data producer

Description of how the data producers interact with the VirtualES demonstrator

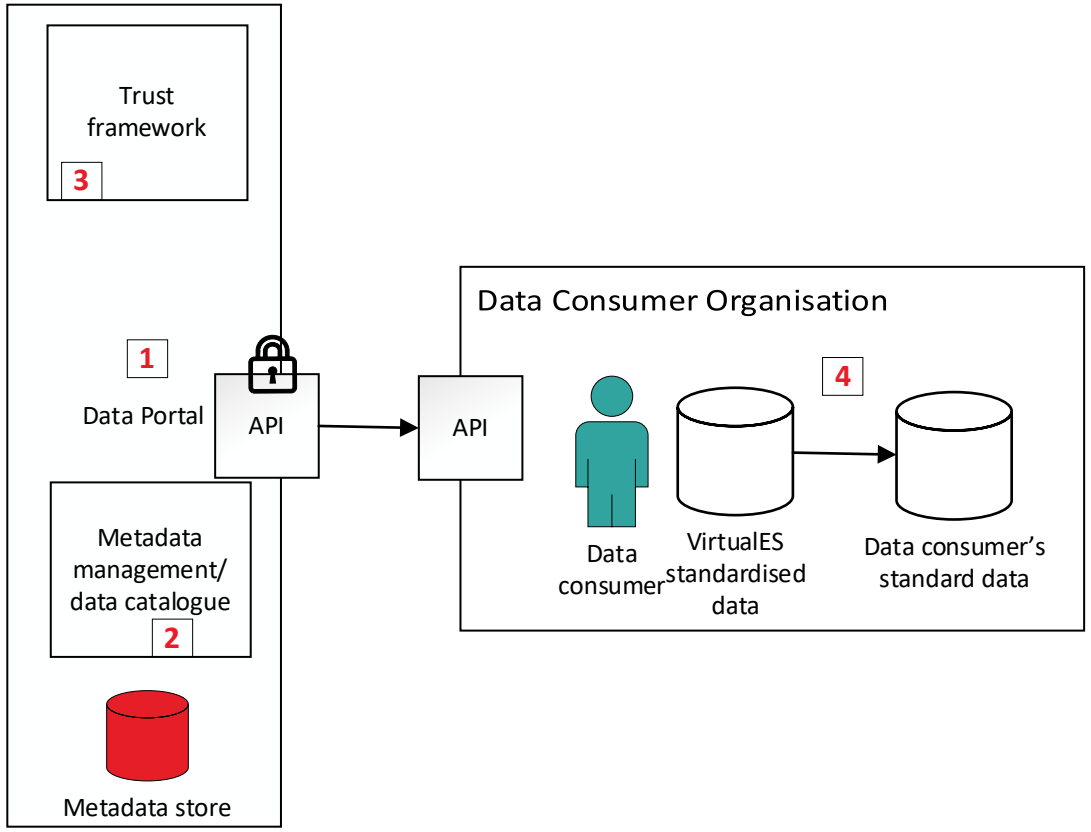


Description of diagram

1. **Metadata entry:** as part of publishing their data to the VirtualES, producers will need to provide a metadata entry to sufficiently describe the data they are sharing, so that it can be registered with a metadata store – as part of the data portal.
2. **Security tagging:** producers will also need to provide security and access control policies to the data so that it can be shared securely with the correct consumers. These security policies will be used by the Trust Framework to handle the permission controls.
3. **Schema validation:** once the data has the required characteristics for sharing with the VirtualES, it will undergo a schema validation check to ensure that the data standard conforms to an agreed standard for sharing. This validation check will be conducted by a streaming technology, where it will use a schema registry to check that the data conforms to the agreed format and is up to date.
4. **API:** the data is then shared securely with the VirtualES using approved and secure protocols.

HLD: data consumer

Description of how the data consumers interact with the VirtualES demonstrator



Description of diagram

1. **Data portal:** consumers will use a data portal as the front entrance/user interface for the VirtualES. This will be via web link that they can access.

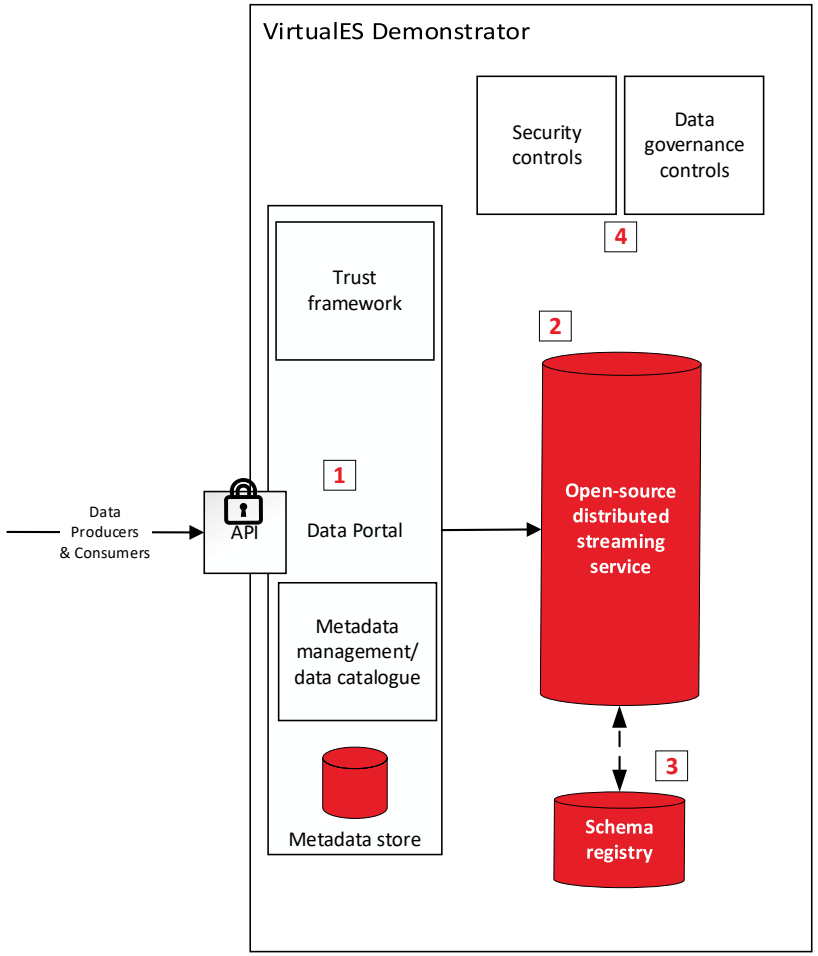
2. **Data catalogue:** using the data portal's metadata store/data catalogue, data consumers will be able to search and discover the data they are interested in. Furthermore, they will be able to use this to request access to the discovered data, where the data producers will be alerted of a request to the data thereby allowing them to approve subscriptions to their data feed. The catalogued data will adhere to a common metadata standard, and users publishing their data will be required to populate metadata fields to make sure that all data is sufficiently understandable and discoverable.

3. **Trust framework:** once the user has found the required data, they can request access to that data. The trust framework will provide the mechanism to enable the sharing and access of data between consumers and producers by ensuring policies and access control permissions are met. Data consumers can then subscribe to feeds of data from the producers.

4. **Data ingestion:** the data consumers can ingest streams of data in the agreed VirtualES schema format using APIs. Once it is ingested, consumers can use or convert the data for analytical and modelling purposes.

HLD: VirtualES demonstrator

Description of the VirtualES demonstrator



Description of diagram

1. **Data portal:** data producers and data consumers will access the VirtualES via a data portal, where they can search and register their data in a metadata store. The security tags for data that is registered with the VirtualES will also be read by the trust framework.

2. **Distributed streaming service:** data that is shared in the VirtualES is done so via a distributed streaming service. The streaming service also has a number of characteristics to help meet a variety of performance requirements e.g. scalable, low latency, asynchronous messaging etc. This enables the sharing of data in real-time between participants without requiring storage of the data within the VirtualES platform.

3. **Schema registry:** the streaming service contains a schema registry that validates schemas according to data standards. It ensures that schemas are aligned, complete and up to date.

4. **Security & governance controls:** the data that is shared using the VirtualES is subject to a variety of security and governance controls. This includes of cyber and data security controls, along with user permissions as part of the data portal's trust framework.

2.3 — Interoperability recommendations

Interoperability recommendations

Recommendations and considerations for interoperability with the VirtualES

Recommendations for the next activities in WP2.3

The following activities are considered in the next stage:

- **Wireframing:** Critical interactions and interfaces raised in this report should be wireframed and tested with potential end users. This should include wireframing of the following key processes:
  - Making base models and scenarios accessible in the VirtualES, including population of metadata and security tagging;
  - Searching and accessing base models or operational scenarios;
  - Merging two or more base models;
  - Extracting a scenario from a base model;
  - Merging two or more scenarios and running them on a merged base model
- **High-level metadata entry:** An initial draft of metadata for base models and operational scenarios should be developed

  Whilst common metadata fields which conform to a metadata standard (e.g. Dublin Core), will be expected to be populated, over time additional fields related to the quality, provenance and trustworthiness of data may also be expected to be provided.

Considerations for Beta phase

As part of the Beta development phase for the VirtualES, which informs both electricity and gas sectors, the full functional, non-functional and security requirements will need to be captured. The requirements will steer the design choices, and the each design feature will need have traceability back to the requirements. The development process will need to follow a formal governance process of providing assurance, validation, and review of requirements and design documentation.

Some of the non-functional requirements will relate to key themes around performance, availability, compatibility, accessibility, integration, service support and usability.

These will help form the Service Level Agreements (SLAs) for the VirtualES. Furthermore, it will help inform the vendor selection for the technology stack, including the choice of streaming technology.

It will also inform design patterns around network configurations, and security tagging implementation, for example, a recommendation on the choice between a Role Based Access Control (RBAC) or an Attribute Based Access Control (ABAC) model.

The following should also be considered:

- **Data standardisation & schema validation:** Data to be exchanged using the VirtualES must conform to an agreed format or schema, that can be ingested by the data consumers. A data standard for the demonstrator should be agreed upon and a schema validation check should be tested where the VirtualES will use a schema registry to validate that the data conforms to the agreed standard and is up to date and complete.
- **Security tagging:** a security tagging standard that the VirtualES will incorporate will need to be adopted by participants. Data producers will need to understand their data governance and access control requirements so that they can apply the correct controls to their data via the security tags. The trust framework will read these access control permissions and manage the access of data between the producers and consumers.
- **Trusted APIs:** to interact with the VirtualES, data consumers and producers will need to ensure that their infrastructure can interact with trusted and secure protocols in the form of APIs. A review of operator's and the selected technology platform's capabilities should be undertaken.

3 — Data sharing framework

Assessing risk and opportunity in data sharing

Overview of considerations for data sharing between organisations

Overview

To enable the use case, an appropriate framework for data sharing is crucial to facilitate the exchange of data between parties and stakeholders.

Data sharing agreements help reduce risks associated to data sharing by ensuring the data is accurate, complete, and up-to-date. They also establish guidelines for data privacy, security, and ownership - which are critical considerations when dealing with sensitive data.

Without appropriate data sharing agreements, there is a risk that parties use incorrect, incomplete, outdated data, which can result in inaccurate simulations and predictions, potentially leading to legal liability, financial penalties and reputational damage for the parties involved.

Therefore, given the considerable risks associated with data sharing, organisations, have defaulted to an overly risk averse and *defensive* position when agreeing data sharing agreements.

This section identifies the risks and challenges associated with sharing data between organisations and proposes a methodology for assessing and concluding data sharing agreements in an efficient way.

Risk categories in data sharing

In order to assess the risks associated with sharing data, it is necessary to inquire about the nature of the data, the purpose for sharing it, the data sensitivity, and the potential risks involved, for each risk category outlined in the table below.

These risk categories have been explored in detail in the *'Assessing risks when sharing data: a guide'* report by the Open Data Institute (February 2022)

As part of the current VirtualES programme, risk inquiries should be part of the overall programme risk identification and assessment process.

This will help the programme:

- Determine appropriate risk mitigation measures
- Produce a '*use case data sharing principles*' report (more details are provided on the subsequent pages)

Risk category	Risk description
Legal and regulatory	Perceived or actual risks of breaching data protection law, intellectual property rights, regulatory requirements, or legal contracts when collecting, using or sharing data.
Ethical	Perceived or actual risk of enabling unethical data collection or use of data, or directly impacting people or communities.
Reputational	Perceived or actual risk of suffering reputational damage from sharing or using data that breaches others' trust, or in reveals limitations in processes or analyses.
Commercial	Perceived or actual risk of losing competitive advantage in the market.

Challenges of data sharing

Summary of the challenges associated with data sharing

Overview

While personal data is subject to legal requirements as prescribed by the Data Protection Act 2018, Section 105 of the Utilities Act 200, and System Operator Functions Information ESO License Special Conditions 2.3, non-personal data is primarily regulated through contract.

While common data best practices inform data to be "presumed open", which refers to principles that data should be made openly accessible by default, unless there are compelling reasons to be kept closed, and data to be triaged, which refers to a process for prioritizing and classifying data based on its importance, quality, and relevance, Organisations are challenged to strike a balance between openness and responsible data management; therefore, needing complex contracts to be in place before sharing or consuming data.

These contracts are subject to negotiation and bespoke provisions. For this reason, content and practice of data sharing agreements vary widely across industry.

Challenges of data sharing

The key challenges include:

- The framework for sharing data lacks a standardised approach, with no commonly accepted methodology. Examples of data sharing frameworks include:
  - **Open Government Licences (OGL) or Creative Commons (CC):** OGL or CC include a range of licences with standard terms and different restrictions on use.
  - **Data sharing agreements:** Bespoke agreements, such as the NUAR or CReDo data licenses, outlining what data is being shared, for how long and any restrictions on use. While such agreements do contain common elements, their form, structure, content and risk profiles vary greatly.
- The nomenclature used in the context of data sharing is varied and not yet settled in the industry.
- There is a lack of understanding of the various types of data sharing agreements in use and when to use them. Organisations will take a very different approach to how they share data, resulting in a fragmented and inefficient approach to concluding data sharing agreements.

Diversity of data sharing agreements

Sharing agreements tend to take the following forms:

- **Data sharing agreements**: used where parties (two parties or more) are each sharing data with each other, for a specified purpose.
- **Data processing agreements**: where a supplier is processing data (personal or non-personal data) in accordance with a customer's instructions.
- **Data access agreements**: where one party allows other parties to access data for a specific purpose. The access may be restricted to certain fields or limited in the number of times data can be accessed.
- **Data licensing agreements**: used where a party is supplying data to another customer and granting it a licence to use the data for specific purposes.
- **Data transfer agreements:** where one party transfers data to another. This may include restrictions on the use of the data, requirements for data security, and measures to ensure confidentiality.

Which agreement is appropriate depends on several factors, including the nature and scope of the data to be shared, the intended purpose whether the data involves personal data, and the background and experiences of the parties.

Existing data sharing landscape

Approaches and features of data sharing agreements

Defensive and offensive approaches

The terms "**defensive**" and "**offensive**" nature are common vocabulary in the context of legal contracts to describe different approaches to managing risks.

A "**defensive**" approach in a legal contract involves taking measure to protect oneself from potential risks and liabilities. These measures include clauses such as limit liabilities, non-disclosure agreements, and warranties. These clauses aim to minimize the risk exposure to the contracting party.

A "**offensive**" approach in a legal contract involves taking measures to again an advantage over the opposing party. These measures include clauses that provide termination of contract rights, performance guarantees, and payment terms. These clauses aim to maximize the benefits of the contracting party.

Both "**defensive**" and "**offensive**" clauses are common in legal contracts and are used to protect and promote interests of the parties involved in the agreement.

Common features of data sharing agreements

The frequent challenge encountered with negotiating and agreeing data sharing agreements is that the effort and internal corporate governance focus tends to be on the **defensive** risk management/mitigation provisions.

The effect of this is that the **offensive** provisions (addressing the manner in which the data is exploited, used and reused - which is where the real value of data sharing lies), are not afforded the same importance.

This skewed balance arises due to a number of factors:

- There is a disconnect between individuals within organisations approving agreements, and those who understand the relevant project and the nature of the data.
- Where a mix of personal, sensitive and non-sensitive data is involved, organisations will take adopt the highest risk avoidance approach, rather than adopt a nuanced approach based on the nature of the datasets.
- Failure to comply with legal and regulatory requirements governing data privacy, security and confidentiality can result in financial penalties and reputational damage.
- The use cases for data sharing can be very specific, resulting in a diversity in principles and practices around data sharing that cannot be easily replicated.
- Organisations are reluctant to share data because it contains valuable intellectual property, such as trade secrets or proprietary algorithms. Unintended sharing of this data could lead to intellectual property theft.
- Organisations may view their data as a key competitive advantage and may be reluctant to share it with competitors or other third parties who could use it to gain a competitive market advantage.
- Organisations are concerned about the security and privacy of their data, particularly when it comes to sensitive information such as personally identifiable information or sensitive commercial data where sharing of such data could result in data breaches, identity theft, or other security or privacy incidents.

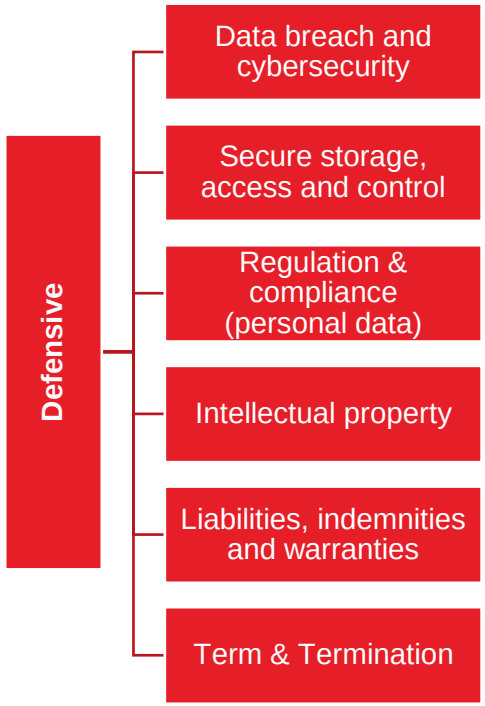
Key components and issues of defensive and offensive provisions

Issues associated with defensive and offensive provisions in contracts

Issues of defensive provisions

Issues of defensive provision include:

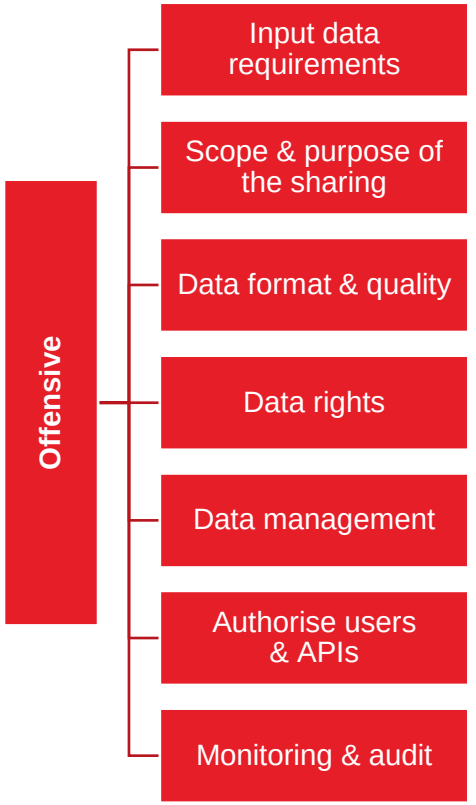
- Contain red flags for internal organisation governance
- Organisation tend to have a standard position on what they can / cannot accept (tick box approach to risk)
- Mostly boiler plate provisions without adequate appreciation of the project or use case
- Can be long winded and difficult to understand for non-lawyers
- Frequently negotiated by lawyers or personnel detached from the project or use case details
- Little focus or appreciation of the commercial or technical value of the data or information to be shared
- Insufficient consideration of the varying datasets involved and difference in risk profiles (open, closed data and personal data).



Issues of offensive provisions

Issues of offensive provisions include:

- Considered by technical teams in isolation of the defensive provisions and overall agreement risk profile
- Best practice in risk minimisation approaches at the point of data creation infrequently applied
- Data format and quality requirements are not fully worked through before negotiations on the agreements commenced. The lack of specificity is compensated by increased risk aversion in the defensive provisions.
- Best practice data management / governance principles are not always clearly set out, leaving parties with wide discretion regarding how the data is treated once shared.



Key observations

Observations regarding existing processes, models and systems to share data

Observation	Recommendation
Parties and stakeholders involved in sharing data need to collaborate earlier to avoid silos and defensive positioning. The significant risks associated with sharing data has resulted in organisations taking a one size fits all approach to negotiating and concluding data sharing agreements.	1. **Earlier cross-function collaboration**: There is currently insufficient collaboration at an early stage between parties and key stakeholders (from different disciplines and functions) within and across organisations. This means risk v opportunity assessment takes place in silos and at different times in the data sharing decision making process. 2. **Nuanced approach to assessing risk in datasets**: Greater appreciation of the varying risk profiles of the relevant datasets is necessary. Adopting a one-size-fits-all approach based on the highest standard leads to a biased risk mitigation strategy that doesn't account for the risk profile.
The significant diversity in approach and the lack of an industry standard data sharing suite of agreements has meant best practice and insights are not socialised throughout the industry.	3. **Develop a suite of data sharing agreements – with flexibility:** Data sharing agreements should be standardized where possible, but also designed with flexibility to accommodate changes in risk profile. Clauses addressing changes in data requirements, such as updates to standards, changes in ownership or access rights, and updates to regulatory requirements, should be included as a modular bolt-on schedule.
A rethink and reset is required, which considers not only the risks associated with data sharing, but the significant opportunity and value to be gained from sharing good quality data.	4. **Rebalancing the risk v opportunity equation**: The balance between risk avoidance and mitigation (Defensive provisions), and extracting the full value from the data shared (Offensive provisions) is skewed by the former. A cultural shift is required to rebalance this and re-emphasise the value of data sharing to gain insights, drive efficiencies, develop between products and services for the benefit of industries, communities and consumers.
A common approach to agreeing a data sharing framework is required. A new rubric which factors in process and decision making and contracting is required. This creates inefficiencies, increases costs and impacts value extraction.	5. **Data sharing process framework (methodology):** A change in approach is required to agreeing data sharing agreements. Rather than starting with a data sharing agreement template, we propose a methodology (detailed on page 38) which supports a logical and efficient process for getting to a digitalised smart data sharing agreement. This methodology will assist ESO in implementing a best practice approach to data sharing for the sector.

ARUP

Proposed data sharing framework methodology

Tailored approach to defining a common legal framework for the VirtualES

The below process outlines the suggested approach to developing data sharing contracts for the VirtualES.

1 — Assess + map the data flows and inputs

- What are the data inputs and outputs?
- Where and how will the data flow?
- How will the data be accessed and transferred?

2 — Categorise the risk profiles of relevant datasets

- What are the types of data to be shared?
- What are the risk categories of the datasets?
- Are there any regulatory / compliance requirements?
- What are there ethical, reputational or commercial risks?

3 — Produce data sharing risk and opportunity assessment

- How important is the data to the use case aims?
- What is the risk vs opportunity ratio?
- Are risk mitigations required? (i.e. anonymisation, use synthetic data, removal of personal data)

4 — Engage parties in agreeing 'use case data sharing principles'

- What are the key data sharing principles derived from steps 1-3?
- Has feedback from parties and stakeholders been considered?
- Can an agreement in principle be reached?

5 — Negotiate data sharing agreement based output of steps 1-4

- Has the appropriate data sharing agreement been identified?
- Have key internal stakeholders been consulted (commercial, technical, legal)?
- Can the risk and negotiation parameters be set internally?

Digitalised smart contract structure

- **Agreement summary**
- **Data Sharing Agreement (core defensive provisions)**
- **Schedules (optional offensive provisions)**

ARUP

Virtual Energy System

Powered by ESO

VirtualES@nationalgrideso.com