



Open data: A review of our progress so far

Pete Davies

Director of Data, Technology and Transformation
July 2023



Contents

Executive Summary	3
Wider industry context	5
Why we have commenced this journey	6
What we have achieved so far	7
Our open data foundations.....	7
Key Deliverables	7
Meta Data Catalogue.....	7
Roles & Responsibilities model.....	8
Data Item Classification model	9
Data Domain model	9
Data Quality Framework and log.....	9
Where are we going next?	10
Open Data Phase 2	10
Closing Remarks.....	11

Executive Summary

Globally and nationally, recognition is growing around the material benefits increased transparency and data publication can have on citizens, businesses, and public services. Where appropriate, it can increase the quality and transparency of public services like energy, build better foundations for future cost savings, bring greater efficiency in data/process management, reduce workloads, eliminate unnecessary duplication, guide policymakers through enhanced participation and collaboration between public and private sectors, and deliver new opportunities and initiatives which benefit consumers.

RECCo is accountable for delivering the REC Services as defined in the Retail Energy Code (REC). We deliver data services to support meter enquiry and switching. RECCo has obligations for these services, including, but not limited to, data controllership, data confidentiality and contract and service management. These require RECCo to conduct a periodic assessment and review to assure ourselves that the code and REC Services remain fit for purpose. For example, today, RECCo recognises that the access process put in place for the Enquiry Services requires amendment, and better alignment with Ofgem's Data Best Practices¹ (DBP) could be achieved at the same time.

In our 2023-2026 **Forward Work Plan**, we identified that a key priority area was developing and delivering our Data and Digitalisation Strategy. Part of this work was the creation of an

RECCo is a not-for-profit corporate vehicle ensuring the proper, effective, and efficient management of the multi-party Retail Energy Code (REC) arrangements. We operate this license-backed energy industry code and seek to promote trust, innovation and competition whilst delivering positive consumer outcomes.

'Open Data Project' to deliver a published asset via existing mechanisms (e.g., Energy Markets Data Specification) of the data governed under the REC. We have taken time to establish a clear picture of the REC data estate and documented this in the form of a metadata catalogue with roles and responsibilities, data classification definitions, a data domain model and a quality framework (this is further detailed in this report under “

¹ Ofgem's 2021 Data Best Practice (DBP) and Digitalisation Strategy & Action Plan (DSAP) Guidance regulatory requirements, expanded

What we have achieved so far”).

Within our **Forward Work Plan**, we committed to “Consult upon and implement the results of applying open data principles to REC controlled data”. This document summarises the activities and outcomes of Phase 1, which defines our open data principles and metadata catalogue, and outlines our approach to Phase 2, where we will ratify the outputs of Phase 1 and implement the agreed open data principles. Parties are invited to provide comments and feedback on our work to date via info@retailenergycode.co.uk.

What we have achieved so far

Wider industry context

Creating transparency around the national energy system and its data ecosystem is essential to assisting the GB journey towards a carbon-neutral, net zero society and a means to deliver the most effective, efficient, secure energy to consumers at the most cost-efficient prices. It can be used by energy participants, policymakers, product manufacturers, energy efficiency firms, start-ups, innovators, and researchers to better inform more efficient planning, building, operation, trading and the user experience of our national energy system.

The energy market's regulatory framework underpins the open, competitive, interoperable market. It ensures fair and transparent data practices through defined principles and requirements for data protection, confidentiality, and security management. This framework sits underneath the overarching general data protection regulations (GDPR), which the government is reviewing as part of its 'Data – a new direction' **reforms**. REC data is managed through this framework, and the REC Services we provide to code parties are further protected through our commercial contracts.

In recent years, the evolution of the data world has been reflected in various energy market regulatory policies and initiatives, working to make energy data more accessible, visible and

discoverable, e.g., Data Best Practice (DBP), Energy Data Taskforce (EDaT) – **final report**, Energy Digitalisation Taskforce (EDiT) – **final report**, or the Data Access & Privacy **framework**. Through the Ofgem-led DBP² and Digitalisation Strategy & Action Plan Guidance (DSAP) work, electricity networks have started on a journey to deliver a consistent, accessible, searchable view of all GB electricity network data assets.

RECCo fully supports the DBP/DSAP principles' aims and sees these as a key enabler to providing a fully interoperable and discoverable view of network Open Data Assets. Whilst DBAP/DSAP has focused on networks, we recognise that the retail and consumer-facing part of the industry needs to participate in the secure opening of data. This is, in part, why RECCo has commenced work to explore Open Data so the REC and parties can consider and evolve ahead of further DBP being mandated.

² Ofgem's 2021 Data Best Practice (DBP) and Digitalisation Strategy & Action Plan (DSAP) Guidance regulatory requirements, expanded

Why we have commenced this journey

The REC arrangements were developed by Ofgem to be consistent with the DBP/DSAP principles, including the data access requirements set out in REC [Schedule 12](#). Whilst RECCo, and the REC, are not obliged to follow the DBP/DSAP principles at this time, we are committed to their aims and outcomes and recognise the need for the REC and its services to be reviewed and amended in a way that is consistent with these principles. Further, RECCo has committed to fully assess and adopt the principles more broadly in our future work.

RECCo is accountable for the delivery of key industry data services, including the:

- Electricity & Gas Enquiry services;
- Green Deal Central Charging database;
- Centralised Registration Services (CRS), including Central Switching Service (CSS);
- Secure Data Exchange Portal;
- Market Stabilisation Charge; and
- REC Performance Assurance.

RECCo has obligations for these services, including, but not limited to, data controllership, data confidentiality, and contract and service management. These require us to conduct periodic assessment and review to assure ourselves that the code and REC services continue to be fit for purpose and to consider future retail energy data needs, which we need to explore and facilitate. Within this, RECCo recognises the potential for the REC to drive and facilitate data and digitalisation change, which extends beyond simple incremental improvements to existing arrangements.

In our published [Forward Work Plan](#) and soon-to-be-published Data and Digitalisation strategy, we set out our strategy and roadmap for delivering improvements via an increase in agile delivery and included phased work to deliver secure open data. Our roadmap includes plans to:

- Commence our open data work; reviewing our data sources;
- Cataloguing, classifying, and considering what REC data can be accessed;
- Consider how RECCo, and the REC Services, would replicate the DBP and DSAP principles; and
- Applying approved open data principles for REC-controlled data.

What we have achieved so far

Our open data foundations

We have conducted data discovery analysis on REC data sources in collaboration with energy and utility experts to assess key data completeness, accuracy, and quality. This included exploring how a ‘Presumed Open’ basis, as defined under the Catapult Energy Data Taskforce Report, might be achieved under the REC.

Phase 1 of the project – the discovery phase – focused on metadata, which is information about the data, rather than the actual data. The analysis focused on the specified REC Services containing 911 items (see Table 1), which have been mapped back to specific items within the code.

Table 1: REC Data Services Analysed

Data Source	Meta Data Owner			
	REC	UNC	BSC	SEC
Electricity Enquiry Service (EES)	68		37	2
Gas Enquiry Service (GES)	43	154		
Green Deal Central Charging	89		10	
Secure Data Exchange Portal	28			
Central Switching Service (CSS)	131			
Market Stabilisation Charge (MSC)	85			
REC Performance Assurance Data (inc. Energy Theft)	264			

The key output of this phase was the development of a metadata catalogue and supporting assets – our foundation for future Open Data phases of work.


Key Deliverables

Meta Data Catalogue

This defines the data items through descriptive and business-oriented information, which provides value to the end users, i.e., the consumers of the data. It incorporates all the insights from the analysis of the data sets, along with the inputs from industry experts and subject matter experts, to provide the necessary context required to help understand data better.

This Meta Data Catalogue is currently being validated during **phase 2** (see below) by the respective Code Managers and system suppliers before requesting changes to the Energy Markets Data Specification. This activity will be managed through the standard REC Change Process to ensure visibility to all impacted parties.

Figure 1: Example of metadata vs actual data



“Field”	Meta Data	Actual Data Example 1	Actual Data Example 2
Name	The name of the shape	star	triangle
Number of Sides	The number of sides a shape has	10	3
Colour	The colour of the shape	purple	blue

Roles & Responsibilities model

A high-level model which demonstrates the different roles and responsibilities associated with the metadata attributes analysed. This model was developed from best practices and principles, metadata tools used across different industries, and common references from ICO and Ofgem.

Table 2: Open Data - Roles & Responsibilities Model

Role	Definition	Responsibility
Meta Data Owner	A metadata owner has the authority to make decisions about business term definitions, data quality, accessibility, retention, and classification requirements as they tie to the business needs.	<ol style="list-style-type: none"> 1. Create the rule book for maintenance of quality of the metadata item. 2. Ownership of items under the change process. 3. The body responsible for the governance of the data specification of that item published to the end user.
Data Master	A data master is responsible for the original source of data where its created and is responsible for physical data quality (E.g., Completeness and Accuracy)	<ol style="list-style-type: none"> 1. Creation of the actual data (the originator). 2. Maintenance of quality of the physical data item.
Authorised Providers	Responsible for the storage of the data item within a central service and provision of that data via Energy Market Messages to other Market Data Services. (e.g., The Electricity Enquiry Service provides this role for data	<ol style="list-style-type: none"> 1. Responsible for only the provision of data.

	items for which the CSS is the Data Master of some items).	
Data Controller / Custodian	A person, public authority, agency, or other body which, alone or jointly with others, determines the purposes and means of processing a specific Data Asset. Controller refers to personal data, and Data Custodian refers to non-personal data.	<ol style="list-style-type: none"> 1. Determines the purpose and means of the processing of a specific data set. 2. Responsibility of delegation provided by the Data Owner to determine the usage and purpose of the data.
Data Processor	A person, public authority, agency, or other body which processes Data Assets on behalf of the Data Controller.	<ol style="list-style-type: none"> 1. Responsible for processing the data to the guidelines set by the Data Controllers (E.g., Processing involves any operation (or set) performed on such as, but not limited to, collection, structuring, storage, use or disclosure).

Data Item Classification model

Defined as part of the metadata catalogue, one of the most important aspects in creating the catalogue was assigning a classification to each data item analysed. The classification model considers how a 'Presumed Open' culture could be achieved in a simple, accessible, discoverable, and secure way.

Table 3: Open Data - Data item classification Model.

Category	Category Classification
Always Open	Data item does not contain personal, commercially-sensitive or security-sensitive data and can be made open.
Conditionally Open	Data item may become personal, commercially-sensitive or security-sensitive data if combined with other data items or sets.
Personal	Data item contains personally identifiable information (PII).
Closed	Data item contains personally-sensitive, commercially-sensitive or security-sensitive data that cannot be shared.

Data Domain model

A logical subset of data which has many related data items and describes the structure of the data classification. This resulted in 8 domain models for each system: the Electricity Enquiry Service (EES); Gas Enquiry Service (GES); Green Deal Central Charging database; the Secure Data Exchange Portal; Central Switching

Service (CSS); the Market Stabilisation Charge (MSC) and REC Performance Assurance Data (Core REC Performance Data, Energy Theft).

Data Quality Framework and log

The framework draws on industry best practices and frameworks from CRISP-DM and ISO Quality Management principles to measure data quality against six key criteria.

1. **Completeness** - not about ensuring 100% of data fields are complete but about determining what data is critical and what is optional.
2. **Uniqueness** - data is unique if it appears only once in a data set; this also builds trust in the dataset.
3. **Consistency** - data values don't conflict with other values within a record and across different data sets. E.g., a person's birth date should be the same across different datasets.
4. **Timeliness** - indicates whether the data is available when expected and needed. E.g., customer addresses may change over time, which may make those data items outdated.
5. **Validity** - the extent to which the data conforms to the expected format, type and range. E.g., MPAN number should have 21 digits and not 14. Note that validity doesn't always ensure accuracy. In a system storing eye colour, blue would be a valid value but inaccurate if the person's eye colour is brown.
6. **Accuracy** - when data reflects reality, for example, correct names and addresses that represent factual and up-to-date data. E.g., a 1cm tall person would be wrong and factually incorrect.

Where are we going next?

Through multiple agile phases, we are taking stock of the data we govern, developing our foundation knowledge base and confirming our future goals and aspirations. This evolving, iterative work will be maintained each year as new data items, sources, and associated uses increase, e.g., through REC changes and delivering the needs of programmes such as Market-wide Half Hourly Settlements (MHHS) and our Energy Theft Reduction work.

Feedback from this paper and other industry engagements will help us shape these outputs, considering the potential impacts, risks, and costs. Some of these will be recommendations for further work to policymakers; others will be changes proposed to existing REC-defined data, processes, services, systems, training, and materials. We will consider how best to build on the existing foundations delivered in the REC today, so we will initially seek to adapt processes, data, systems, services etc., before considering new systems or approaches. This approach will minimise the impact and cost to parties (ultimately paid by consumers).

Open Data Phase 2

In Phase 2, we aim to:

- Update the metadata catalogue based on stakeholder feedback;
- Gain approval of metadata item classification from metadata owners as per the Energy Markets Data Specification (EMDS);

- Raise a REC Change Proposal to improve and expand data access simply and securely; and
- Develop technical access to reflect any changes implemented.

Phase 2 will also use the outputs from Phase 1 to prepare updates to the Data Specification by adding additional information captured in the metadata catalogue and addressing some of the identified issues (noted in the quality assessment of each metadata item). This will ensure that users of the data items have access to information that enables them to make the best use of the data available.

We will utilise the REC change process to bring these updates forward.

Closing Remarks

The secure opening of data is critical to delivering net zero targets, increasing transparency and therefore trust, removing duplication and supporting efficiencies, and promoting and enabling innovation. To reduce the risk of unintended consequences, this needs to be delivered in a measured way and built upon existing governance mechanisms and legislation.

The number of organisations under Ofgem's Data Best Practices will continue to grow, and it is essential that RECCo acts now to ensure we can best support REC parties in delivering these obligations when they are implemented. As a result, RECCo has initiated this project. We have started reviewing data sources we are the Data Controller or Custodian of and have started to ratify early analysis with other code bodies and system providers before bringing through the REC change process.

We of course invite engagement through the formal change process at the relevant phase but welcome now comments and feedback on our work to date via info@retailenergycode.co.uk.