Notes on Completion: Please refer to the appropriate NIA Governance Document to assist in the completion of this form. The full completed submission should not exceed 6 pages in total.

NIA Project Registration and PEA Document

Date of Submission

Nov 2019

Project Reference Number

NIA_WPD_048

Project Registration

Project Title

Presumed Open Data (POD)

Project Reference Number

NIA_WPD_048

Project Licensee(s)

National Grid Electricity Distribution

Project Start

January 2020

Project Duration

1 year and 6 months

Nominated Project Contact(s)

Sam Rossi Ashton

Project Budget

£580,366.00

Summary

There is a large amount of useful data which is published about the energy sector through mandatory reports, innovation trials and consumer tools. However, datasets are often published on standalone webpages with limited descriptions. This makes it very difficult for both incumbents and innovators to discover, search and understand datasets.

This project will review all data held by WPD to ascertain the extent that it can be shared with third-parties. Use cases will then be developed using the data that can be shared. The dataset with the highest value will be processed, standardised, and published so the identified use cases can be fully realised.

Third Party Collaborators

Energy Systems Catapult

Centre for Sustainable Energy

Problem Being Solved

There is a large amount of useful data which is published about DNOs through mandatory reports, innovation trials and consumer tools. However, datasets are often published on standalone webpages with limited descriptions. This makes it very difficult for both incumbents and innovators to discover, search and understand datasets.

Data should be easy to find and accompanied by the information needed to understand their content. Reducing barriers to access data will attract innovators who can create operational efficiencies, develop new business models and define new value propositions. Increasing the speed at which new markets can be developed may improve and attract investment in the energy system due to the improved ability to understand the risks and opportunities up front.

Issues can occur across privacy, security, consumer impact and commercial domains but these can be mitigated using anonymisation, aggregation, redaction or introduction of noise. If issues cannot be resolved through the above techniques, it may be appropriate to

limit rights (public data) or limit access (shared data), such that key parties can safely utilise data to create value whilst protecting consumers.

Method(s)

The project will use WPD data as a worked example of how the recommendations of the Energy Data Task Force Digitalised Energy System Report could be implemented by a DNO in the following stages:

1. Data Discovery and Classification

- Identify data that is currently being collected and held by us.
- Review available metadata standards and apply the most appropriate one to the data catalogue.

2. Use Case Development

An iterative process whereby our data users and the Energy Systems Catapult (ESC) will develop compelling use cases for the data identified in stage 1. These use cases will then be taken to two or three half-day consultation workshops that will be open to third-parties (e.g. The ENA Data Group, Councils, Community Groups, Academics, I/C generation/demand site developers, suppliers and energy service companies) and WPD data owners/users. The workshops aims are to:

- Provide an overview of the project;
- Further develop the established use cases so they optimally benefit third-parties; and
- Establish the extent that preliminary processing and context is required to make data accessible for parties with lesser resource for such activity.

The use cases will be assessed against a range of criteria (Net Zero impact, customer goals, innovation impact, etc.) to prioritise use cases. This will in turn be used to rank the importance of our data and drive the creation of a data value assessment.

3. Data Openness Assessment & Processing

Review the sensitivity issues related to each data set (such as consumer privacy, negative consumer impact, security, and commercials) and develop a generic methodology that can be used by all LNOs to classify the openness of each datasets in accordance with the Open Data Institute's Data Spectrum:

- Open: Data is made available for all to use, modify and distribute with no restrictions
- Public: Data is made publicly available but with some restrictions on usage
- Shared: Data is made available to a limited group of participants possibly with some restrictions on usage
- Closed: Data is only available within a single organisation.

The datasets required to facilitate the use cases identified in stage two go through:

- Data quality upscaling to rectify issues with incomplete datasets that could render them unfit for use by third-parties.
- Openness upscaling to see if minor changes (aggregation, redaction, or adding noise) can be made to datasets to make them less sensitive.
- Preliminary processing identified it stage 2 to promote third-party accessibility.

If data required to facilitate use cases is not currently collected by WPD the method and format in which it would be collected would go through the same review process.

4. Hub Development

Development of recommendations for a public-facing data hub to be hosted online where:

- All data is stored in one central location;
- Appropriate means of access (registration/verification of identity) are required for datasets than can be considered Public or Shared;

- Data can be easily downloaded upon necessary verification; and
- Stakeholders can register to be notified when new data sets are published.

The hub development will look to exploit techniques that are already available regarding the automation of data correction, meta-tagging, and the flagging of data issues when uploading new data.

5. Data Science Challenge

This will include publicly launching the proactive data publication roadmap, launching the data hub implementation phase. In addition, a specific data use case will be selected and launched as a data science challenge to drive engagement with our data and deliver immediate value.

Scope

The scope of this project is to deliver on recommendations, made by the Energy Data Task Force in it's a Strategy for a Modern Digitalised Energy System Report.

Recommendation 2 - Maximising the value of data

Government and Ofgem should direct the sector to adopt the principle that Energy System Data should be Presumed Open, using their range of existing legislative and regulatory measures as appropriate, supported by requirements that data is 'Discoverable, Searchable, Understandable', with common 'Structures, Interfaces and Standards' and is 'Secure and Resilient'.

Recommendation 3 - Visibility of data

A Data Catalogue should be established to provide visibility through standardised metadata of Energy System Datasets across Government, the regulator and industry. Government and Ofgem should mandate industry participation though regulatory and policy frameworks.

Objective(s)

The project has two objectives:

1. **Maximise the visibility of data.** The data hub will make data discoverable, searchable, and provide visibility through standardised metadata.
2. **Maximise the value of data.** The data hub will make understandable by employing common structures and interfaces.

Consumer Vulnerability Impact Assessment (RIIO-2 Projects Only)

n/a

Success Criteria

The project will be deemed successful if an implementation plans for a functional energy data hub is published and an initial set of network data is used by customers to facilitate their own objectives.

Project Partners and External Funding

We will deliver this project with two partners:

- Energy Systems Catapult (ESC) – The ESC was fundamental in the drafting of the Modern Digitalised Energy System Strategy. As such, it is best placed to assist with data discovery, use case development, data prioritisation, and openness assessment.
- Centre for Sustainable Energy (CSE) – CSE have successfully delivered projects with us previously regarding external engagement and web developed. As such they are well placed to assist in these areas of the project.

Potential for New Learning

Maximising the visibility and value of data could:

- Faster decarbonise the energy system through easier identification of capacity for the connection of LCT assets.
- Provide a better 'whole of system' view promoting better system security.
- Enhance third-party interactions promoting better flexibility response behaviour and greater opportunities for flexibility revenue.
- Facilitate the realisation new commercial opportunities from the creation of new markets with new players.

- Reduce customer bills through a more strategic deployment of community-owned LCTs.
- Optimise procurement regarding asset location, size, or function.
- Better identify whether flexibility or build is a better solution where constraint exists.

Scale of Project

The project aims to utilise a comprehensive, but not exhaustive, catalogue of data held by us. Datasets will be scored after use cases have been fully developed and the datasets considered to have the highest value will be used to trial the hosted data hub.

Technology Readiness at Start

TRL6 Large Scale

Technology Readiness at End

TRL8 Active Commissioning

Geographical Area

This project is non-geographical but will likely utilise data from all of our operating areas.

Revenue Allowed for the RIIO Settlement

N/A

Indicative Total NIA Project Expenditure

The project budget is £580,366, of which £522,329 is NIA expenditure.

Project Eligibility Assessment Part 1

There are slightly differing requirements for RIIO-1 and RIIO-2 NIA projects. This is noted in each case, with the requirement numbers listed for both where they differ (shown as RIIO-2 / RIIO-1).

Requirement 1

Facilitate the energy system transition and/or benefit consumers in vulnerable situations (Please complete sections 3.1.1 and 3.1.2 for RIIO-2 projects only)

Please answer **at least one** of the following:

How the Project has the potential to facilitate the energy system transition:

n/a

How the Project has potential to benefit consumer in vulnerable situations:

n/a

Requirement 2 / 2b

Has the potential to deliver net benefits to consumers

Project must have the potential to deliver a Solution that delivers a net benefit to consumers of the Gas Transporter and/or Electricity Transmission or Electricity Distribution licensee, as the context requires. This could include delivering a Solution at a lower cost than the most efficient Method currently in use on the GB Gas Transportation System, the Gas Transporter's and/or Electricity Transmission or Electricity Distribution licensee's network, or wider benefits, such as social or environmental.

Please provide an estimate of the saving if the Problem is solved (RIIO-1 projects only)

The value of Open Data across the economy has been estimated to be 0.5% of GDP and the value of flexibility and increased system efficiency has been estimated to be £17-40 Billion between 2016 and 2050.

An example use case is that of fault detection and mitigation. It is estimated that, if Open Data could improve fault detection to reduce the number of WPD faults by 10%. In 18/19 WPD's customers experienced 3.6mn Customer Interruptions (CI) and 200mn Customer Minutes Lost (CML). Ofgem's Value of Lost Load (VoLL) assessment suggests the cost of society of these is £15.44 per CI and £0.38 per CML. Multiplying these gives the cost of CIs as £56mn and the cost of CMLs as £77mn, totaling £133mn. Reducing the faults causing these interruptions by 10% would represent a £13.3mn annual saving to society.

Please provide a calculation of the expected benefits the Solution

It might be estimated that a project of this nature could deliver increased efficiency and system benefits of several million pounds over 3-year period. We expect it would be possible to develop a quantified estimate of the direct and indirect benefits of the project using a suitable methodology and approach.

Please provide an estimate of how replicable the Method is across GB

The methodology and data principles are replicable by all other DNOs.

Please provide an outline of the costs of rolling out the Method across GB.

The cost of each DNO to develop their own hub could vary depending on whether the project learning indicates that an off the shelf solution to be suitable for BaU rollout.

Requirement 3 / 1

Involve Research, Development or Demonstration

A RIIO-1 NIA Project must have the potential to have a Direct Impact on a Network Licensee's network or the operations of the System Operator and involve the Research, Development, or Demonstration of at least one of the following (please tick which applies):

☐ A specific piece of new (i.e. unproven in GB, or where a method has been trialled outside GB the Network Licensee must justify

repeating it as part of a project) equipment (including control and communications system software).

☑ A specific novel arrangement or application of existing licensee equipment (including control and/or communications systems and/or software)

☐ A specific novel operational practice directly related to the operation of the Network Licensees system

☐ A specific novel commercial arrangement

RIIO-2 Projects

☐ A specific piece of new equipment (including monitoring, control and communications systems and software)

☐ A specific piece of new technology (including analysis and modelling systems or software), in relation to which the Method is unproven

☐ A new methodology (including the identification of specific new procedures or techniques used to identify, select, process, and analyse information)

☐ A specific novel arrangement or application of existing gas transportation, electricity transmission or electricity distribution equipment, technology or methodology

☐ A specific novel operational practice directly related to the operation of the GB Gas Transportation System, electricity transmission or electricity distribution

☐ A specific novel commercial arrangement

Specific Requirements 4 / 2a

Please explain how the learning that will be generated could be used by the relevant Network Licensees

Whilst each DNO has their own IT systems, the methodology and data principles are replicable by all other DNOs.

The project (and partners as required) will provide a formal update at each of the ENA Data Working Group meetings to maximise the dissemination of learning to other parties and so that it can inform the work programme of the group. Additionally, the project team will host quarterly calls with other LNOs to inform them of the progress of the project and enable them to provide input.

Or, please describe what specific challenge identified in the Network Licensee's innovation strategy that is being addressed by the project (RIIO-1 only)

This meets the data challenge identified in our Innovation Strategy:

- Facilitating access to existing and future data sets – this will enable greater visibility of our network assets, current and planned operation as well as engaging with third parties to offer service and solutions based on data;
- Generating increased data to facilitate the future needs of the network and customers – utilising additional monitoring and analytics to increase network visibility and understanding.

☑ Has the Potential to Develop Learning That Can be Applied by all Relevant Network Licensees

Is the default IPR position being applied?

☑ Yes

Project Eligibility Assessment Part 2

Not lead to unnecessary duplication

A Project must not lead to unnecessary duplication of any other Project, including but not limited to IFI, LCNF, NIA, NIC or SIF projects already registered, being carried out or completed.

Please demonstrate below that no unnecessary duplication will occur as a result of the Project.

An Openness Assessment methodology has not been comprehensively applied to DNO operational data before.

If applicable, justify why you are undertaking a Project similar to those being carried out by any other Network Licensees.

n/a

Additional Governance And Document Upload

Please identify why the project is innovative and has not been tried before

An Openness Assessment methodology has not been comprehensively applied to DNO operational data before.

Relevant Foreground IPR

n/a

Data Access Details

n/a

Please identify why the Network Licensees will not fund the project as apart of it's business and usual activities

Previous project learning has indicated that simply publishing network data en masse does not suit third party needs. We need to investigate customers' requirements to develop the preliminary processing, format, and method of publication to provide maximum industry benefit. The benefits of open data are largely realized outside of WPD, as such it is currently uncertain how to best deliver them.

Please identify why the project can only be undertaken with the support of the NIA, including reference to the specific risks(e.g. commercial, technical, operational or regulatory) associated with the project

There is a risk that without the project's scope of engagement, the method of data sharing we decide upon would not be best suited to some types of network customers. This could increase false starts, stifle innovation, and be a barrier to market entry. Innovation funding is required to determine the correct audiences to ensure that the right levels of information is made available to the right people in the format expected.

This project has been approved by a senior member of staff

☑ Yes