

for energy consumers

Style Definition: Text - bulleted: Tab stops: 0.63 cm, List tab

Style Definition: Text - Roman numbered: Indent: Left: 0 cm, Hanging: 0.63 cm, Tab stops: 0.63 cm, List tab + Not at 0.32 cm

Guidance

Data Best Practice Guidance					
Publication date:	25 May<u>15</u> November 2021	Contact:	Greg Johnston, Senior Manager – Market Data Policy		
		Team:	Modernising Energy Data		
		Tel:	0141 354 5415		
		Email:	ofgemdataservices@ofgem.gov.uk		

This document is Ofgem's Data Best Practice Guidance. Its requirements must be complied with by <u>licenseesLicensees</u> of the RIIO-2 price controls. It is part of Ofgem's standards for data and digitalisation, to be followed by relevant Licensees and by Ofgem itself.

© Crown copyright 2021

The text of this document may be reproduced (excluding logos) under and in accordance with the terms of the **Open Government Licence**.

Without prejudice to the generality of the terms of the Open Government Licence the material that is reproduced must be acknowledged as Crown copyright and the document title of this document must be specified in that acknowledgement.

Any enquiries related to the text of this publication should be sent to Ofgem at:

10 South Colonnade, Canary Wharf, London, E14 4PU. Alternatively, please call Ofgem on 0207 901 7000.

This publication is available at <u>www.ofgem.gov.uk</u>. Any enquiries regarding the use and re-use of this information resource should be sent to: <u>psi@nationalarchives.gsi.gov.uk</u>

Guidance - Data Best Practice

Document version	Description	Changes since previous
		document
Data Best Practice Guidance	Proposed version May 2021	N/A
v0.3		
Data Best Practice Guidance	Version as a result of	As set out in "Track Changes
<u>v1.00</u>	consultation May 2021	DBP Guidance v1.00"

Contents

1. Introduction
General feedback
2. Summary
Data Best Practice principles
Definitions
3. The Data Best Practice Principles
1. Identify the roles of stakeholders of Data Assets
Explanation
2. Use common terms within Data Assets, Metadata and supporting information
Explanation
3. Describe data accurately using industry standard Metadata
Explanation
4. Enable potential Data Users to understand the Data Assets by providing supporting
information
Explanation
5. Make Data Assets discoverable to potential Data Users
Explanation
6. Learn and deliver to the needs of current and prospective Data Users
Explanation
7. Ensure data guality maintenance and improvement is prioritised by Data User needs19
Ensure data quality maintenance and improvement is prioritised by bata over needs 19
8. Ensure Data Assets are interoperable with Data Assets from other data and digital
services20
Explanation
9. Protect Data Assets and systems in accordance with Security, Privacy and Resilience
best practice
Explanation
10. Store, archive and provide access to Data Assets in ways that ensures sustained
benefits
Explanation
11. Treat all Data Assets, their associated Metadata and software scripts used to process
Data Assets as Presumed Open
Explanation
1. Introduction
General feedback

Guidance - Data Best Practice

2. Summary
Data Best Practice principles
Definitions
3. The Data Best Practice Principles
1. Identify the roles of stakeholders of Data Assets13
2. Use common terms within Data Assets, Metadata and supporting information14
3. Describe data accurately using industry standard Metadata15
4. Enable potential Data Users to understand the Data Assets by providing supporting
information16
5. Make Data Assets discoverable to potential Data Users
6. Learn and deliver to the needs of current and prospective Data Users
7. Ensure data quality maintenance and improvement is prioritised by Data User needs19
8. Ensure Data Assets are interoperable with Data Assets from other data and digital
services20
9. Protect Data Assets and systems in accordance with Security, Privacy and Resilience
(SPaR) best practice
10. Store, archive and provide access to Data Assets in ways that ensures sustained
benefits22
11. Treat all Data Assets, their associated Metadata and Software Scripts used to process
Data Assets as Presumed Open23

1. Introduction

- 1.1. This document provides principles and explanations that describe requirements for complying with Data Best Practice Guidance- ("DBP Guidance").
- 1.2. Energy network companies who are licensed under the RIIO-2 price controls (gas and electricity transmission, gas distribution network companies and the electricity system operator) are required to comply with this guidance when they are preparing and updating their Digitalisation Strategy and Digitalisation Action Plan.
- 1.3. To find out more about this licence obligation, please visit these following links:
 - RIIO-2 Final Determinations for Transmission, Gas Distribution and Electricity System Operator¹;
 - RIIO-ED2 Sector Specific Methodology Decision²; and
 - Decision on the proposed modifications to the RIIO-2 Transmission, Gas Distribution and Electricity System Operator licences³.
- 1.4. This guidance is part of Ofgem's standards for data and digitalisation, to be followed by relevant Licensees and by Ofgem itself⁴. Both the Data Best PracticeDigitalisation Strategy and Action Plan Guidance and this guidance are part of our standards for data and digitalisation.
- 1.5. The work of the other organisations, such as GO FAIR and Government Digital Service (GDS) have strongly informed the development of this guidance. The Government Digital Service (GDS) provides wide-ranging support for topics relating to data and digitalisation; it gives information and methods that span all the principles in the guidance. Of particular relevance are the The following are particularly relevant:
 - GO FAIR and its FAIR data principles⁵;

Formatted: Indent: Left: 0 cm, Hanging: 1.25 cm Formatted: Font color: Text 1

¹ <u>https://www.ofgem.gov.uk/publications-and-updates/riio-2-final-determinations-transmission-and-</u> gas-distribution-network-companies-and-electricity-system-operator

https://www.ofgem.gov.uk/publications-and-updates/riio-ed2-sector-specific-methodology-decision ³ https://www.ofgem.gov.uk/publications-and-updates/decision-proposed-modifications-riio-2-

transmission-gas-distribution-and-electricity-system-operator-licences

https://www.ofgem.gov.uk/publications-and-updates/forward-work-programme-202122 ⁵ <u>https://www.go-fair.org/fair-principles/</u>

- GDS Service Standard⁶;
- GDS Technology Code of Practice7; and the
- GDS Service Manual⁸.

General feedback

- We believe that feedback is at the heart of good policy development. We are keen to 1.6. receive your comments about this guidance. We'd also like to get your answers to these questions:
 - Do you have any comments about the overall quality of this guidance?
 - Do you have any comments about its tone and content?
 - WasIs it easy to read and understand? Or could it have been better written?
 - Any further comments?
- 1.7. Please send any general feedback comments to <u>ofgemdataservices@ofgem.gov.uk</u>.

⁶ <u>https://www.gov.uk/service-manual/service-standard</u>
⁷ <u>https://www.gov.uk/government/publications/technology-code-of-practice/technology-code-of-</u>

practice ⁸ https://www.gov.uk/service-manual

2. Summary

2.1. DBP <u>guidanceGuidance</u> is designed to ensure data is treated as an asset and used effectively for the benefit of consumers, <u>stakeholders</u> and the Public Interest. It is a principles-based <u>set of guidanceapproach</u> which provides guidance on the quality, accuracy and accessibility of data, <u>and its requirements are not specific to the energy sector</u>. It includes the principle that <u>"data assets"Data Assets</u> must be treated as <u>'presumed open'Presumed Open</u>⁹ which means that data must be made available for all people to use, unless the organisation responsible for handling the data provides <u>specific</u> evidence <u>of a specific reason for needing</u> to <u>reduceshow that the data should</u> <u>be withheld or</u> its availability <u>reduced (e.g. to protect individuals' rights to privacy)</u>. By complying with this guidance organisations will enable the full benefits of data to be unlocked for consumers.

Data Best Practice principles

- 1. Identify the roles of stakeholders of Data Assets.
- 2. Use common terms within Data Assets, Metadata and supporting information.
- **3.** Describe data accurately using industry standard Metadata.
- Enable potential Data Users to understand Data Assets by providing supporting information.
- 5. Make Data Assets discoverable for potential Data Users.
- 6. Learn and deliver to the needs of current and prospective Data Users.
- 7. Ensure data quality maintenance and improvement is prioritised by Data User needs.
- Ensure Data Assets are interoperable with Data Assets from other data and digital services.
- Protect Data Assets and systems in accordance with Security, Privacy and Resilience (SPaR) best practice.
- **10.** Store, archive and provide access to Data Assets in ways that ensure sustained benefits.
- Treat all Data Assets, their associated Metadata and software scripts used to process Data Assets as Presumed Open.

Formatted: Text - Numbered, Indent: Left: 0 cm, Hanging: 0.63 cm, Line spacing: single

⁹ <u>https://es.catapult.org.uk/reports/energy-data-taskforce-report/</u>

Definitions

Data Asset: Any entity that is comprised of data. For example, a database is a data asset that is comprised of data records. A data asset may be a system or application output file, database, document, or web page. A data asset also includes a service that may be provided to access data from an application. For example, a service that returns individual records from a database would be a data asset. Similarly, a web site that returns data in response to specific queries (e.g., www.weather.com) would be a data asset.

This definition is taken from National instituteInstitute of Standards and Technology (NIST).¹⁰

Data Contact Point: An organisation or individual who is the primary point of contact about a Data Asset or Metadata associated with a Data Asset.

Data Controller: A person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of a specific Data Asset.

This is based on the <u>ICOInformation Commissioner's Office (ICO)</u> definition but has been modified by removing reference to personal data and replacing it noting the processing of a Data Asset.¹¹

Data Custodian: A person, public authority, agency or other body that has a legal right to process and publish a Data Asset as the Data Controller or otherwise.

Data Processor: A person, public authority, agency or other body which processes Data Assets on behalf of the Data Controller.

This is based on the ICO definition but has been modified by removing reference to personal data and replacing it noting Data Assets.¹²

Data Subject: The identified or identifiable living individual or entity to whom data relates.

Data User: An organisation or individual which utilises data held by a Data Custodian for any reason.

Formatted: Space Before: 6 pt, After: 6 pt

¹⁰ <u>https://csrc.nist.gov/glossary/term/data_asset</u>

¹¹ https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-

regulation-gdpr/controllers-and-processors/what-are-controllers-and-processors/ ¹² <u>https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/controllers-and-processors/what-are-controllers-and-processors/</u>

https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protectionregulation-gdpr/controllers-and-processors/what-are-controllers-and-processors/

Data Best Practice Guidance: means (1) the guidance document issued by the Authority¹³ in accordance with Part D of Special Condition 9.5 (Digitalisation) of the RIIO-2 price controls for Electricity Transmission, Gas Transmission and Gas Distribution and Special Condition 2.11 (Digitalisation) of the RIIO-2 price controls for the Electricity System Operator and (2) part of Ofgem's standards for data and digitalisation.

Digitalisation Action Plan: an organisation's plan to digitalise its Products and Services prepared and published in accordance with Part B of Special Condition 9.5 (Digitalisation) of the RIIO-2 price controls for Electricity Transmission, Gas Transmission and Gas Distribution and Part B of Special Condition 2.11 (Digitalisation) of the RIIO-2 price controls for the Electricity ServiceSystem Operator.

Digitalisation Strategy: the strategic approach taken by an organisation to digitalise its Products and Services and evidenced by the archive prepared and published by the <u>LicenseeLicensees</u> in accordance with Part A of Special Condition 9.5 (Digitalisation) of the RIIO-2 price controls for Electricity Transmission, Gas Transmission and Gas Distribution and Part A of Special Condition 2.11 (Digitalisation) of the RIIO-2 price controls for the Electricity System Operator.

Digitalisation Strategy and Action Plan Guidance: means (1) the guidance document issued by the Authority in accordance with Part C of Special Condition 9.5 (Digitalisation) of the RIIO-2 price controls for Electricity Transmission, Gas Transmission and Gas Distribution and Special Condition 2.11 (Digitalisation) of the RIIO-2 price controls for the Electricity System Operator and (2) part of Ofgem's standards for data and digitalisation.

DSAP: A combination of both Digitalisation Strategy and Digitalisation Action Plan.

Energy System Data: all Data Assets for which an entity is a Data Custodian as a consequence of it exercising its rights and obligations under a licence granted under section 6 (1) or (1A) of the Electricity Act 1989 or section 7, 7ZA, 7A or 7AB of the Gas Act 1986".

Metadata: a set of data that describes and gives information about other data.

 $^{^{\}rm 13}$ The terms "the Authority", "we", and "us" are used interchangeably in this document

Open Data: Data Assets, their associated Metadata and software scripts<u>Software Scripts</u> used to process Data Assets that are made available for anyone to use, modify and distribute with nowithout restrictions.

Open Data Triage: The process carried out by a Data Custodian to determine if there is any
evidence of sensitivities associated with Data Assets, their associated Metadata and
software
scriptsSoftware Scripts
used to process Data Assets if they are used as Open Data. These
sensitivities are limited to those that:

- (i). relate to peoples' rights to personal privacy;
- (ii). security needs;
- (iii). obligations from legislation and/or regulation;
- (iv). commercial requirements that, if not protected, will have a negative impact on Products and Services for end-consumers; and
- (v). would have a negative impact on the Public Interest.

Where any of the above sensitivities are identified, Open Data Triage will also include the determination of how the Data Custodian can mitigate any risk associated with them, while also making the Data Assets, their associated Metadata and software scriptsSoftware Scripts used to process Data Assets as open to stakeholders as possible. The Data Custodian should consider both processing of and/or whether providing different levels of access by different types of stakeholders to the Data Assets, their associated Metadata and software scriptsSoftware Scripts used to process Data Assets would help to mitigate any identified risk.

Presumed Open: The treatment of Data Assets, their associated Metadata and software scriptsSoftware Scripts used to process Data Assets as Open Data, subject to Open Data Triage.

Products and Services: Anything that a party can offer to a market for attention, acquisition, use or consumption that could satisfy a need or want.

Public Interest: The welfare or well-being of the general public and society.

Single Provider Product or Service: A product or service among the Products and Services provided by a Data Custodian where no alternative option or provider is available to parties seeking to access that product or service.

Software Scripts: A code and its programming documentation; including information on how to execute that code, that enables users to read, capture, process, store or transmit a Data Asset or Metadata.

Formatted: Space Before: 18 pt, After: 18 pt Formatted: Pattern: Clear

Formatted: Space Before: 6 pt, After: 6 pt

Guidance - Data Best Practice

the Authority: means the Gas and Electricity Markets Authority that is established under section 1 of the Utilities Act 2000

3. The Data Best Practice Principles

1. Identify the roles of stakeholders of Data Assets

Explanation

3.1. The <u>LicenseeLicensees</u> must identify the Data Assets that it is the Data Custodian of; for these, the <u>LicenseeLicensees</u> must also identify any relevant Data Subjects, Data Controllers and Data Processors. The <u>LicenseeLicensees</u> must keep this information in logs.

Formatted: Normal, Line spacing: single

2. Use common terms within Data Assets, Metadata and supporting information

Explanation

3.2. Licensees must enable Data Users to be able to search for and easily joinlink Data Assets and associated Metadata to Data Assets and Metadata provided by other organisations. Licensees must label and describe Data Assets and Metadata using a taxonomy that is commonly recognised by practitioners who use the Metadata across the relevant subject matter domain. Formatted: Font: Bold
Formatted: Normal, Line spacing: single

3. Describe data accurately using industry standard

Metadata

Explanation

- 3.3. The <u>LicenseeLicensees</u> must make it easy for Data Users to be able to <u>work withuse</u> and understand information that describes each Data Asset. The <u>LicenseeLicensees</u> must therefore provide Metadata associated with Data Assets and this Metadata must be made available to Data Users <u>independentlyindependent</u> of the Data Asset.
- 3.4. The <u>LicenseeLicensees</u> must treat the Metadata as a Data Asset-<u>itself</u>. When providing Metadata, the <u>LicenseeLicensees</u> must format and structure this in a widely recognised and accepted format-<u>that is machine readable.</u>
- 3.5. There is no requirement for the <u>LicenseeLicensees</u> to create Metadata about its Metadata associated with Data Assets.
- 3.6. When it updates or extends a Data Asset, the <u>LicenseeLicensees</u> must ensure that the Metadata reflects any such changes so that Data Users can<u>easily</u> identify additions or changes.

Formatted: Space Before: 6 pt, After: 6 pt

Formatted: Font: Bold Formatted: Normal, Line spacing: single Formatted: Space Before: 6 pt, After: 6 pt

4. Enable potential Data Users to understand the Data Assets by providing supporting information

Explanation

3.7. -Throughout the lifecycle of a Data Asset the <u>LicenseeLicensees</u> must make available supporting information that Data Users require for the <u>maximum</u> benefits to be gained by consumers, <u>stakeholders</u> and the Public Interest. The <u>LicenseeLicensees</u> must ensure a point of contact is provided for Data Users to raise and resolve enquiries about the Data Asset and its supporting information.

Formatted: Font: Bold
Formatted: Normal, Line spacing: single

5. Make Data Assets discoverable to potential Data Users

Explanation

3.8. Licensees must ensure that allany potential Data Users can identify the Data Assets that the LicenseeLicensees is the Data Custodian of, and how Data Users can pursue access to these Data Assets. Licensees must ensure that the Metadata associated to Data Assets is discoverable to Data Users, subject to the outcome of an Open Data Triage process.

Formatted: Font: Bold
Formatted: Normal, Line spacing: single

Formatted: Paragraph,numbered, Indent: Left: 1.25 cm, Line spacing: single

6. Learn and deliver to the needs of current and prospective Data Users

Explanation

3.9. The LicenseeLicensees must identify the Product and Service requirements of Data Users who use, or who wish to use, the Data Assets that it is theprovided by them as Data Custodian-of. The LicenseeLicensees must then develop and deliver Products and Services to meet Data Users' requirements, where doing so is of reasonable cost and would benefit at least one or both of the following: end-consumers-and, stakeholders or the Public Interest. Where the LicenseeLicensees is obligated to publish and update a DSAP, these Products and Services must be integrated into those publications.

Formatted: Font: Bold
Formatted: Normal, Line spacing: single

7. Ensure data quality maintenance and improvement is prioritised by Data User needs

Explanation

- 3.10. Where the LicenseeLicensees reasonably expects the Data Users' application of Data Assets for which it is the Data Custodian of to deliver a net benefit for end-consumers, stakeholders and/or the Public Interest, the LicenseeLicensees must ensure that Data Assets are of a quality that is sufficient to meet thereasonable requirements of its Data Users. Data Users must have an option for contesting decisions regarding the definition of sufficient data quality of a Data Asset.
- 3.11. Where data quality issues are identified the <u>LicenseeLicensees</u> must ensure that these issues are <u>logged</u>, <u>considered and</u> rectified as soon as practicable.

Formatted: Font: Bold
Formatted: Normal, Line spacing: single

8. Ensure Data Assets are interoperable with Data Assets from other data and digital services

Explanation

- 3.12. Licensees must enable interoperability, between the Data Assets thatfor which they are the-Data Custodian of withand Data Assets, of other Licensees as a minimum standard.
- 3.13. When the <u>LicenseeLicensees</u> makes Data Assets available, it must do so in ways that make it <u>reasonably</u> easy for Data Users to gain information and/or insight from those Data Assets in conjunction with other-Data Assets. <u>from other Licensees</u>. <u>There must also be sufficient information to align to Data Assets from other industries</u>.
- 3.14. The <u>LicenseeLicensees</u> must make data available in such a way that it is reasonably easy for Data Users to:
 - (i). exchange Data Assets between systems;
 - (ii). interface with Data Assets held in the Licensee'sLicensees systems; and
 - (iii). join Data Assets with other Data Assets, such as by using standard interfaces, standard data structures and/or common reference data.

Formatted: Normal, Line spacing: single	
Formatted: Font color: Text 1	_

Formatted: Font color: Text 1

5

9. Protect Data Assets and systems in accordance with Security, Privacy and Resilience (SPaR) best practice

Explanation

Formatted: Font: Bold

3.15. The <u>LicenseeLicensees</u> must ensure that compliance with this guidance does not negatively impact its compliance with all-relevant regulations, legislation, and Security, Privacy and Resilience (SPaR) requirements. Formatted: Normal, Line spacing: single

10. Store, archive and provide access to Data Assets in ways that ensures sustained benefits

Explanation

- 3.16. When Data Assets are not required by the <u>LicenseeLicensees</u>, the <u>LicenseeLicensees</u> must ask stakeholders whether they consider <u>ifthat</u> the Data Assets could create a future benefit if archived. The <u>LicenseeLicensees</u> must archive Data Assets when, taking account of <u>stakeholders views', stakeholders' views</u> it determines <u>that</u> the storage will be a net benefit to consumers-<u>, stakeholders</u> and/or the Public Interest.
- 3.17. When archiving, the <u>licenseeLicensees</u> must also ask stakeholders for views on the storage method and formats to use. In determining what to archive, it must consider:
 - Data Assets;
 - Metadata;
 - software scriptsSoftware Scripts used to process Data Assets;
 - data derived resulting from this processing of the original Data Asset; and
 - human-readable representations of the data and any other relevant supporting information.
- 3.18. The <u>LicenseeLicensees</u> must ensure <u>that</u> the risk of <u>unintentional or malicious deletion</u> of Data Assets, Metadata and software scriptsSoftware Scripts used to process Data Assets, <u>being unintentionally or maliciously deleted such that they are unrecoverable is</u> minimal.<u>effectively managed and monitored to ensure possible recovery.</u>

Formatted: Font: Bold

Formatted: Normal, Line spacing: single

Formatted: Font color: Custom Color(RGB(31,72,124))
Formatted: Font color: Auto

11. Treat all Data Assets, their associated Metadata and software scriptsSoftware Scripts used to process Data Assets as Presumed Open

Explanation

- 3.19. The <u>LicenseeLicensees</u> must treat all Data Assets, their associated Metadata and <u>software scriptsSoftware Scripts</u> used to process Data Assets <u>thatwhere</u> it is the Data Custodian-of, as Presumed Open and these must be subjected to Open Data Triage.
- 3.20. The <u>LicenseeLicensees</u> must treat information created during Open Data Triage as Open Data, except where this will result in a sensitivity listed in the Open Data Triage definition.
- 3.21. Where a sensitivity associated is identified with the Data Assets, their associated Metadata and software scriptsSoftware Scripts used to process Data Assets is identified, the LicenseeLicensees must take all reasonable steps to provide suitable options that describe how these can be madeto make them available in a format or version that mitigates the risk associated with any identified sensitivity. When identifying those options the LicenseeLicensees should additionally consider whether providing different stakeholders with different levels of access would mitigate any identified risk while also-minimising any reduction in the utility of the Data Asset.
- 3.22. The <u>LicenseeLicensees</u> must make available the Data Assets, their associated Metadata and <u>software scriptsSoftware Scripts</u> used to process Data Assets in the changed formats, versions or with the different levels of access to stakeholders, where it is beneficial to end consumers, <u>stakeholders</u> and the Public Interest to do so.
- 3.23. The <u>LicenseeLicensees</u> must record at least the following information about Open Data Triage processes:
 - what has been triaged;
 - when the process took place;
 - a description of the sensitivities <u>and risks</u>, if any, that have been identified including the type of sensitivity as defined by Open Data Triage;
 - the options considered for how to mitigate any sensitivities that wereor risks identified and the impact these have on the utility of the Data Assets, their associated Metadata and/or software scriptsSoftware Scripts used to process Data Assets; and
 - any decisions made.

Formatted: Font: Bold

Formatted: Normal, Line spacing: single Formatted: Space Before: 6 pt, After: 6 pt

- 3.24. The <u>LicenseeLicensees</u> must ensure there is a point of contact available to stakeholders forto allow them to seek information about Open Data Triage processes as well as to provide them with the opportunity to challenge decisions and escalate issues.
- 3.25. The <u>LicenseeLicensees</u> must keep under review its collection of available Data Assets, their associated Metadata and/or <u>software scriptsSoftware Scripts</u> used to process Data Assets for risks <u>ofor</u> sensitivities and must mitigate these as they arise.