

Name of Dataset:

Date of Assessment:

Assessment conducted by:

Description of Dataset:

Please complete all questions in grey- Comment where Applicable

Data sharing restrictions

Question	Response	Comment (Complete where Applicable)
Are we legally or otherwise obligated to publish this information?	Yes	OFGEM requirements
Are we legally prohibited from providing external access to this data?	No	We do a data share with other DNOs yearly
Does this dataset relate to the affairs of any identifiable individual or to any particular business? (excluding SSEN) (Sections 105 of the Utilities Act 2000)	No	
Does this dataset contain any personally identifiable information?	No	
Does SSEN have the legal rights to publish this dataset? (eg are there any licencing restrictions?)	Yes	
Does publishing this dataset have any impact on sites of critical national infrastructure	No	

Risk Analysis

Category	Standard Enterprise Risk	Inherent Likelihood	Inherent Impact	Inherent Risk Score	Can this be mitigated?	Mitigation Approach	Mitigated Risk Likelihood	Mitigated Risk Impact	Final Risk Score	Comments
Regulatory Requirements	Published data conflicts with existing regulatory submissions resulting in reputational damage and regulatory action	N/A	N/A	0					0	
Quality	Published data is inaccurate or misleading, resulting in a series loss of reputation for SSEN.	Very Low	Very Low	1					1	We are audited
Security- Cyber	Published data enables someone with hostile intentions to compromise the cyber security of SSEN.	N/A	N/A	0					0	
Security- Physical	Published data enables someone with hostile intentions to compromise the physical security of SSEN.	N/A	N/A	0					0	
Privacy	Personally identifiable information is published with a legal basis resulting in legal action against SSEN.	N/A	N/A	0					0	
Legal	Published data breaches a licences or other intellectual property agreement resulting in legal action against SSEN.	N/A	N/A	0					0	
Commercial	Commercial stakeholders are able to gain a commercial advantage by abusing our published data to overcharge us.	N/A	N/A	0					0	
Ethics	Published data enables discrimination against individuals or a given community resulting in inequality	N/A	N/A	0					0	
Consumer	Published data has a negative impact on electricity markets resulting in a less favourable situation for consumers.	N/A	N/A	0					0	
Other	Are there any other risks you believe should be considered in deciding whether to publish this data?	N/A	N/A	0					0	

Are there any other risks you believe should be considered in deciding whether to publish this data?

Recommended Classification

Open

[Click Here for Further Information on Classifications](#)

Closed	Restricted	Shared	Public	Open
Total Score above 160	Total Score above 120 but less than 160	Total Score above 80 less than 120	Total Score above 40 less than 80	Total Score Below 40

Final Review

Can this dataset be published ?	Yes
Date of Completion:	19/07/2023
Making reference to documented risk ratings above and the opportunity of sharing this data, what is your justification for this conclusion?	Data is already widely shared.
What actions are required before this dataset is published? (Include documenting limitations, information for users, and ongoing monitoring required)	Pulling specific data together
Could we do anything to make this dataset more usable to the public? (eg providing definitions, adding in additional data, directing the user to an external party with complementary additional information)	Include descriptions/definitions
How often should this data be updated? (eg once a year, monthly, quarterly, once off?)	Yearly
This assessment has been approved by the following (Sub-Domain Owner or Higher)	Yes
The Risk Specialists consulted (required if any inherent risk score if over 10)	N/A

Guidance

This tab provides guidance for users, to support them in filling out the Data Triage Assessment tab, including definitions and more detailed description.

Key Terms

- "open data" is where data is freely available to any member of the public
- "published" refers to the sharing of this data
- "score" refers to the impact and likelihood of a specific risk occurring
- "trigger" is an event which causes an effect
- "risk specialist" is someone at SSE with specialist knowledge to advise on risk areas, but is not accountable for the risk being taken (e.g. a lawyer)
- "Data Sets" is a collection of related, discrete items of related data that may be accessed individually or in combination or managed as a whole entity.

Risk Likelihood and Impact

A risk is a trigger which leads to a specific event and has a defined consequence, and is quantified with a risk and likelihood (on a scale of 1 to 5) as set out below.

Rating	Score	Likelihood	Impact
Very High	5	Almost certain to occur (>90% likely)	Injury or loss of life
High	4	Is likely to occur (<75% likely)	Regulatory sanctions and reputational damage
Medium	3	Is not unlikely to occur (<50% likely)	Reputational damage and financial impact
Low	2	May occur in rare cases (<25% likely)	Financial impact only
Very Low	1	Possible but not expected to occur (<10% likely)	Minor effect on efficiency of operations
N/A	0	Not possible to occur (0% likely)	There would be no impact whatsoever

		Likelihood					
		0	1	2	3	4	5
Impact	0	0	0	0	0	0	0
	1	0	1	2	3	4	5
	2	0	2	4	6	8	10
	3	0	3	6	9	12	15
	4	0	4	8	12	16	20
5	0	5	10	15	20	25	

Score

The risk score is calculated by multiplying the impact and likelihood of a risk. Depending on how many risks are over a certain threshold, a "result" will be proposed on a sliding scale of how much risk would be posed by sharing the data. The scoring criteria is explained below. Note for any individual risk with a score greater than 10, please consider the mitigations below.

Criteria	Rating Scale	Justification	Action
Two or more risks with a score above 10	Closed	If there are two risks that cannot be mitigated below a score of 10, then the dataset should not be shared publicly.	This dataset should not be shared publicly unless the risks can be mitigated.
One risk with a score above 10		Where there is a single risk that can not be mitigated below a 10, it means there is a risk that is somewhat likely to occur, and have a considerable impact. This does not necessarily mean the data must not be shared, but the risk should be carefully considered if deciding to proceed.	To determine whether the dataset can be shared given the risk profile, consult with the relevant specialists (see guidance) for the risks posed.
2+ risks have a score between 8 & 10 (inclusive)	Shared	When a dataset has more than 1 risks with a score higher than 8 (but lower than 10) it means there are valid risks to consider.	The data owner and the Open Data Team should discuss if this data can be openly published, and consider providing it to a limited audience.
1 risk has a score above an 8		Datasets with a single risk between 8-10 means there is perceived to be just one category where a valid risk needs to be considered before publishing the data. The likelihood and impact are not seen to be very high.	In most cases it should be possible to share this data openly, unless the data owner believes any mitigations or restrictions would be appropriate.
No scores above a 7	Open	Of all the risks presented, none that would have a significant impact are seen to be likely	These datasets should be shared openly without restrictions.

Mitigation Technique	Summary Description of approach	Effect on risk	Reference
Access Control	Data is released under access control, such as a user name and password, to manage readership for licensing or technical reasons. This allows us to share the data to some extent, but is	This lowers the probability of sensitive data being accessed by a wider audience, making it easier to monitor who is using the data.	https://odileeds.github.io/open-data-tips/technique/access-control
Aggregation	Combining data to reduce the level of detail in terms of time, physical space or individuals	The probability of deliberate or accidentally identification is reduced, but it may in turn become less useful	https://odileeds.github.io/open-data-tips/technique/aggregation
Anonymisation	Removing personal identifiers, both direct and indirect, that may lead to an individual being identified	Lowers or avoids risks associated with information being attributed to an individual	https://odileeds.github.io/open-data-tips/technique/anonymisation
Data Binning	Replaces a specific field such as age with a reference to a range (e.g. replacing a person's age with "18-25")	Lowers risks associated with information being attributed to an individual	https://odileeds.github.io/open-data-tips/technique/binning
Delayed Publication	Data is published after a pre-defined delay so that the user can not see "real time" data	Can reduce risks associated with data being used to follow an individual or organisations activity in real time	https://odileeds.github.io/open-data-tips/technique/delayed-publication
Obfuscation	Hiding original data with modified content	Reduces the accuracy of the dataset in the interest or risk reduction, but may make the dataset less useful	https://odileeds.github.io/open-data-tips/technique/obfuscation
Pseudonymisation	Separating the personal information from the dataset, and replacing it with a reference to the information held elsewhere (e.g. a staff ID number instead of the employee name)	Lowers risks associated with information being attributed to an individual, but allows it to be easily re-attributed by SSEN if needed	https://odileeds.github.io/open-data-tips/technique/pseudonymisation
Randomness	Altering the data to introduce noise which makes it less accurate	Can help reduce the risk of identifying individuals from a dataset, but may make it less useful	https://odileeds.github.io/open-data-tips/technique/randomness
Redaction	Removing certain data or replacing it with dummy data (e.g. "REMOVED"), including entire fields, or entire records	Avoid releasing sensitive data while being transparent about what has been removed, but may make it less useful	https://odileeds.github.io/open-data-tips/technique/redaction
Restrictive Licensing	A license is applied which sets restrictions on how the data can be used, and defines permissions for onward sharing	Reduces risks related to commercial sensitivity or security concerns related to specific groups, but is no longer open	https://odileeds.github.io/open-data-tips/technique/restrictive-licensing
Synthetic Data	Generates a dataset with the same properties as a real dataset, but using fake data. For example a fake list of employee names, that accurately reflects the real gender and race demographics of	Is arguably not open data, but allows us to avoid the risk of sharing personally identifiable information at the cost of accuracy	https://odileeds.github.io/open-data-tips/technique/synthetic-data

Referencing the UKPN templates with the CC by 4.0 Attribution licence