

Data Sharing Assessment

	<i>Insert information in boxes below</i>
Dataset name:	West Midlands BSP Transformer Flows
Dataset representation:	csv
Openness Result:	Open
Assessment completed by:	Ellie Mueller / Jay Chen
Date DSA completed:	22/01/2024
Approved for sharing by:	Phil Lawson
Date of approval:	07/03/2024

Form Completeness	Complete
--------------------------	-----------------

	Initial Risk	Mitigation	Residual Risk
Privacy	Green		Green
Security	Green		Green
Commercial	Green		Green
Negative Consumer Impact	Green		Green
Other	Green		Green

Summary	Green	100%
	Amber	0%
	Red	0%

Openness Recommendation	Open
Recommendation Notes	

#	Question	Guidance	Response	Initial Risk	Decision Tool Response	Guidance	Supporting Information and Notes	Mitigation Category	Residual Risk
0	Does NGED own or have a legal right to publish the data?	Has NGED (or a contracted third party) generated or collected the data? In some cases where data has been shared with NGED by the data owner, there may be rights (via data licensing) or obligations (e.g. Environmental Information Regulations or DCUSA Embedded Capacity Registers) which provide legal grounds for publication of the dataset or derived dataset. Where a dataset combines owned and shared data this should be recorded as shared data.	Yes: Data Owner	Green	Complete initial and residual risk assessment and proceed to next check		All assessed historical flows datasets have been created by NGED using NGED owned logger data.		Green
Privacy									
1	Does the dataset contain personal data?	Consider if sharing this dataset infringes the General Data Protection Regulation (the GDPR) or the Data Protection Act 2018. [1] Note: Additional guidance is given in LE5 - Data Protection Policy	No	Green	Complete initial and residual risk assessment and proceed to next check		Datasets contain only timestamps, substation details (names and numbers), transformer number, reading type, and a series of power / voltage readings. So does not contain any personal information		Green
Security									
2	Does this dataset contain information that creates an incremental security risk or exasperate an existing risk?	Consider if the dataset contains information that is not already public by another source, e.g. existing NGED reports, third party datasets (e.g. satellite imagery), etc. If an alternative source exists then consider if the NGED dataset contains more detailed data (granularity, latency, accuracy, etc.). Where the dataset contains novel information, it should be assessed for new security issues such as: security of supply (Distribution Code - Engineering Recommendation P2), security of assets, security of people, etc. Where the dataset replicates or extends existing public datasets, consideration should be given to the incremental risk publication would create. For example, does the increased granularity of data create more risk?	No	Green	Complete initial and residual risk assessment and proceed to next check		One supplementary data for these sites is already published every day for the previous day (midnight-midnight). This could be captured overtime and built into a historical flows dataset. The time granularity of this dataset will be lower, as it is half-hourly. The granularity of the readings is higher, split by transformer (not site) and providing MVA, MVA, MW, and V readings (including for measured flow, unmasked demand, and embedded generation)		Green
Commercial									
3	Does the dataset contain information that could damage the commercial interests of the company?	Data Best Practice Guidance described commercial information as "Data that relates to the private administration of a business or data which was not collected as part of an obligation / by a regulated monopoly and would not have been originated or captured without the activity of the organisation". Commercial issues could also arise due to prior data licensing which would preclude sharing the same data under different terms (e.g. exclusivity agreement, confidentiality agreement), data which is (or could be interpreted as) inconsistent with regulatory reporting or could be seen as exploiting NGED's licenced role (e.g. could undermine the Competition Act 1998 [3], Enterprise Act 2002 [4] or Enterprise and Regulatory Reform Act 2013 [5])	No	Green	Complete initial and residual risk assessment and proceed to next check				Green
Negative Consumer Impact									
4	Would sharing this dataset result in a negative impact on consumers?	Consider if access to this data is likely to drive actions, intentional or otherwise, which will negatively impact the consumer. For example, data about procurement assessment processes could drive up prices for products or services which increase costs for consumers.	No	Green	Complete initial and residual risk assessment and proceed to next check				Green
Other									
5	Are there any other concerns associated with sharing this dataset not covered above?	If there are any other issues that require mitigation before publishing your dataset that haven't been addressed by the questions above, please clearly state the reason(s) in the 'Supporting information' cell. For example, specific legislative or regulatory barriers.	No	Green	Complete initial and residual risk assessment and proceed to next check				Green
6	Is a dataset quality caveat required for end users?	Data quality is subjective. A dataset may be perfectly acceptable for one use case but entirely inadequate for another. Data accuracy can be more objective but there remain many instances where the required precision differs across use cases. Data quality should not be seen as a reason for not sharing as potential users may find the quality acceptable for their use, find ways to handle the quality issues or develop ways to solve issues which can improve the quality of the underlying data. However, known limitations (quality or accuracy) should be clearly documented and where there are uncertainties a robust quality disclaimer should be used.	Yes but mitigation in place	Amber	Capture evidence and reason for mitigation in the 'Supporting Information' column and proceed to next check.		There is a risk of data quality / accuracy issues, as the data comes from loggers on the network that can be subject to communications problems and inaccuracies etc. The volume of data is such that it would be difficult to identify and address all of these issues without committing considerable resource. A caveat to explain possible data issues would be suffice.	Disclaimer	Green