GIS Shapefiles (SPEN_004)

Methodology

Method Statement Preparer:	Data Transformation and Analytics Lead GIS Analyst
Data Triage Representative:	Open Data Sharing Lead
Senior Manager:	Data Analytics Development Manager
Dataset Title:	GIS Shapefiles
Date of Method Statement:	March 2025
Refresh Date:	March 2026
Description of Dataset:	GIS shapefiles showing the key network assets across SP Distribution, SP Manweb and SP Transmission networks. These datasets allow users to visualise our key network assets. The data is published to give users an understanding of the location and characteristics of the key network assets. Users can either view the data on the Open Data Portal, or they can download the relevant datasets and visualise via GIS software. Our GIS Shapefiles are available on our Open Data Portal under our Shared Licence and are refreshed quarterly.

Production Timetable: *Provide info on: When does the process start; Key dates and milestones in the process.*
The GIS Shapefiles are refreshed on a quarterly basis, with the data available on our Portal under our shared data licence. The key milestones are: 1. The relevant asset information is extracted from the database (quarterly - third Monday of each quarter (March, June, September & December)). 2. The asset data files are processed and formatted (within 1 day of stage 1). 3. Revised files are quality assured (within 2 days of stage 2). 4. Approved files transferred to our Open Data Portal for final checks and sign off (within 2 days of stage 3). 5. Refreshed data published on our Open Data Portal and available for our stakeholders (by the end of the week in which the process was started). A risk assessment and data triage review are conducted every six months.
Process to collate data and Source Systems: *Explain the process undertaken to collate data and detail names of systems and type of data that is extracted from each.*
1. The source data is held in our corporate database, as part of a wider set of SPEN geospatial information. 2. The asset data is extracted from the corporate database into a set of standalone GIS files, processed to include only the relevant attribute information (asset identifier, status, voltage, and name), organised into the separate datasets. 3. Quality checks undertaken. 4. Published to Open Data Portal.
Assumptions: *Any interpretation of regulatory guidance; Any assumptions on the data source or its' application*
N/A
Additional Calculations: *Any calculations applied to the data to arrive at the final data table.*
N/A
Dependencies - Information sources: *Information, if any, that comes from other sources/departments.*
N/A
Control Points: *What checks are done during the process to confirm the accuracy of the content?*
The following QA checks are carried out as part of the process to ensure the content of the output files is accurate:

Methodology

- Date-stamps of the assets in the input datasets are checked to ensure they are current.
- Output files are loaded into a map to ensure they cover the expected extent.
- Attribute values of the output files are spot-checked to ensure:
    - Attribute fields are all fully populated.
    - Attribute values are accurate, based on the values of the input datasets.
    - Attribute values are in the agreed publication format, as specified in the metadata files.
  - Output files are checked to ensure they are the same size (and therefore contain roughly the same number of assets) as previous extracts.