

Icebreaker One response to Ofgem's consultation on updates to Data Best Practice Guidance and Digitalisation Strategy and Action Plan Guidance

FAO: Liam Bennett, Senior Manager - Data Policy and Regulation, Energy Systems Data Regulation Team

This is Icebreaker One's response to the consultation on updates to Data Best Practice Guidance and Digitalisation Strategy and Action Plan Guidance¹. It can be published openly.

Please note that throughout this consultation, Icebreaker One uses the terms Open, Shared and Closed data as defined [here](#)².

If you have any questions about our submission or require clarifications please do not hesitate to contact us via openenergy@icebreakerone.org. Thank you for considering our submission.

Call for input response:

Q1. Do you agree with our proposal to implement a structural change to DBP Guidance, introducing intended outcomes for each principle? If not, how do you suggest we could clarify the aim of each principle?

We support the use of clearly defined intended outcomes, and support the addition of stated outcomes for each principle.

Q2. What are your views on the proposed wording of our intended outcomes for each principle in DBP Guidance?

N/A

Q3. What are your views on our proposal to require the use of Dublin Core as the Metadata standard for companies obligated under DBP Guidance?

We support requiring the use of Dublin Core as the Core Standard. We agree with the recommendation for publishing metadata using the Dublin Core standard, but suggest further refining it to specify the Dublin Core-based Data Catalog Vocabulary (DCAT), which is supported by the main data catalogue platforms and is the metadata standard adopted for all EU public sector data publication. As part of the research work completed on the Energy Data Visualisation Project³ in partnership with Arup, Hippo Digital, our recommendations included:

¹<https://www.ofgem.gov.uk/publications/consultation-updates-data-best-practice-guidance-and-digitalisation-strategy-and-action-plan-guidance>

² <https://icebreakerone.org/open-shared-closed/>

³ <https://energy.icebreakerone.org/report-edvp/>

- Adopt Dublin Core Metadata Initiatives (DCMI) as the core, domain agnostic standards
- Apply DCMI in conjunction with multiple standards in the field to provide extensibility and fulfil specific data usage requirements.
- That Energy Data Best Practice guidance, by which networks must now comply, use Dublin Core metadata standards

As part of this research work, we concluded the following regarding open metadata standards:

- Adopt open technical standards for data by default.
- Metadata standards and glossaries should be presented in platform-neutral⁴ and open formats, allowing implementation in both the open source and proprietary technologies.
- We recommend metadata be licensed under an Open License by default and exceptions only applied if material risks apply (e.g. national security)
- The adoption of glossaries and controlled vocabularies to achieve common understanding, interpretation of datasets and interoperability.

Q4. If you do not agree with this proposal, are there alternative Metadata standards that should be utilised by licensees instead?

N/A

Q5. If you are a licensee required to comply with DBP Guidance, can you provide a timescale for the implementation of the proposal to adopt Dublin Core as your Metadata standard?

IB1 is not required to comply with DBP Guidance. IB1 can codify this requirement into Open Energy Membership agreements for those who are licensees.

Q6. What are your views on our proposal to require the use of the Creative Commons Attribution Licence or the Open Government Licence as the standard Open Data licence for companies obligated under DBP Guidance?

Icebreaker One has conducted extensive sector research supporting the creation of the Open Energy Trust Framework, including inquiry into whether this framework should mandate common standards for licensing, metadata, and data quality.

Adoption of *standardised licences* for Open Data should be prioritised. Research demonstrated that creation of bespoke 'Open' data licences presents *risks to data usability*, by increasing the *costs* (time, financial, legal) of interpreting non-standardised licences and/or by permitting data publishers to create licences which *do not actually produce Open Data* (either through error or evasion).

⁴ Platform-neutral means the metadata standards and glossaries should run/display properly on any device. This is done through a combination of a standard document encoding, such as UTF-8, and a standard document structure, such as JSON.

We would support Ofgem to standardise their approach to this issue within the DBP guidance. Rather than creating a new licence for this purpose, we strongly suggest that Ofgem mandates *existing well-known licences (see list below in Q7)*. In order to minimise unnecessary conflict which could surround the choice of a single Open Data licence, as well as costs of change for organisations that already adopt different but well-known standard Open Data licences, we suggest that Ofgem specifies a *limited set of well-established standard Open Data licences that must be used to publish Open Data under the DBP guidance*.

While we support mandating *existing well-known licences* such as Creative Commons Attribution Licence or Open Government Licence, requiring DBP compliant organisations to use only those two licences *may restrict the quantity and/or increase costs to organisations using other well-known open licences appropriate for their datasets*.

We suggest that Ofgem extends the guidance to support all organisations to produce *appropriately descriptive metadata* in order to *improve Open Data discoverability and searchability*.⁵ This could potentially take the form of specified minimum descriptive metadata - for example stating that Open Data published under the DBP Guidance must be associated with metadata containing basic items such as: publisher details, title, associated licence, brief description of contents, publication date, update frequency, etc. These fields have a high value in metadata discovery as it improves searchability in search engines when they can offer filters for these criteria to users. While this may seem a basic approach in contrast to more extensive technical standards, our research supports the view that the introduction of minimum descriptive requirements could have a significant impact in practically supporting the discovery and use of Open Data in the sector, without requiring difficult political consensus or significant additional cost.

We recommend metadata should be licensed under an Open License *by default* and exceptions only applied if material risks apply (e.g. national security). This should apply to both Open Data and Shared Data that needs some manner of restriction or access control to the underlying data.

Q7. If you do not agree with this proposal, can you suggest alternative Open Data licences to be utilised as a common Open Data licence?

Through our Open Energy research, we found supporting existing well-known licences for data publishing to Open Energy allowed for participants to easily publish and share datasets. Open Energy and Open Net Zero can support the following Open Data licences:

- Creative Commons Attribution
- Creative Commons Attribution Share-Alike
- Creative Commons CC0 ('CC Zero')
- Creative Commons Non-Commercial (Any)

⁵ As also echoed by academic research:

https://geography.exeter.ac.uk/media/universityofexeter/schoolofgeography/images/researchgroups/epg/Data_governance_2021.11.01_EJ_DIGITALISATION_WORKSHOP.pdf

- GNU Free Documentation License
- Open Data Commons Attribution Licence
- Open Data Commons Open Database License (ODbL)
- Open Data Commons Public Domain Dedication and License (PDDL)
- Other (Attribution)
- Other (Non-commercial)
- Other (Open)
- Other (Public Domain)
- UK Open Government License (OGL)

Icebreaker One has conducted extensive research around data sharing in the energy sector, our position reflects two key research findings from the Open Energy project:

1. General demand for data-sharing between regulated and non-regulated stakeholders stems primarily from *Shared Data*, not Open Data.
2. Open Data alone will not be sufficient to achieve net zero. Data exchange between regulated and non-regulated stakeholders for net zero purposes/projects is also primarily driven by *Shared Data*, not Open Data. This concerns data shared by and to regulated entities.
3. Currently, sharing Shared Data is *high friction* and *high cost*

Reflecting the above points, we recommend that future investment in the ecosystem of energy and related data should focus on specifying a *unified approach to sharing types of data that do not meet the criteria to be published as Open Data*. While we fully support all efforts to improve the energy sector Open Data ecosystem, our research and technology development as MEDA competition winners (Open Energy project) strongly suggests that the *development of Shared Data infrastructure is of high importance to the role of data and digitalisation in enabling net zero*. As such, we suggest that this be appropriately prioritised alongside new initiatives to expand DBP guidance.

We suggest that a future approach builds from industry-tested foundations developed through the Open Energy Trust Framework⁶ in order to implement a clear and consistent approach to Shared Data governance which can support an ecosystem of commercially, socially and environmentally valuable Shared Data. We recommend that this includes:

- The establishment of a system of *data sensitivity classes*, to improve consistency in sector understanding and categorisation of Shared Data.
- The implementation of a *standardised approach to Shared Data access*, based on a standardised range of access conditions, to enable flexible governance of data access within well-defined parameters.
- The introduction of a *standardised approach to Shared Data licensing*, based on a standardised range of licence conditions, to enable flexible creation of licences within well-defined parameters.

⁶ <http://icebreakerone.org/ib1-trust-framework-for-data-sharing/>

Q8. If you are a licensee required to comply with DBP Guidance, can you provide a timescale for the implementation of the proposal to adopt the Creative Commons Attribution Licence or the Open Government Licence as your Open Data licence?

IB1 is not required to comply with DBP Guidance.

Q9. What are your views on our proposal to require licensees to create and publish a Data Catalogue of their Data Assets?

We support this proposal and applaud efforts to increase the ease with which data can be published, and improve the discoverability of datasets. It should be noted that publishing a data catalogue does not mean licensees require a (meta)data management platform, although it is usually the best solution for anyone with a significant number of datasets and/or need for careful governance. For basic publication, meeting this requirement could be as simple as using a DCAT editor to generate a static file on the website e.g. <https://rdforms.com/editors/dcat/>.

We agree with the recommendation for publishing metadata using the Dublin Core standard, but suggest further refining it to specify the Dublin Core-based Data Catalog Vocabulary (DCAT), which is supported by the main data catalogue platforms and is the metadata standard adopted for all EU public sector data publication.

Through our work in the Modernising Energy Data Access (MEDA), we developed Open Energy to make it easy to search, access, and use energy data. From our experience with Open Energy, we are able to easily index DNO's data catalogues which have DCAT endpoints, such as via OpenDataSoft data portals.

To support licensees to increase the data available to Data Users who can then cross-reference with other industry Data Assets, we recommend licensees to publish their data assets in a manner which is compatible with Open Energy and Open Net Zero *increase the ability to search, access, and use this data*.

DCAT is fully supported by Open Net Zero, and IB1 has added to the vocabulary to enable Shared Data descriptions. We may also extend DCAT with further useful sector-specific metadata, such as example indications of Ofgem regulatory compliance, or applicable data models.

Q10. Do you agree with our proposed position on treating aggregated smart meter consumption data as Energy System Data?

Through our research, we have found the potential need for half hourly data in two forms: 1. In aggregate at varying degrees of granularity (the majority of use cases), and 2. at the level of the individual smart meter.

Under the current energy codes and the GDPR, suppliers, rather than the DNOs, are the best placed parties to provide aggregated smart meter data, as they are the data controllers of individual smart meter data and have a direct relationship with their customers. Suppliers, [under the Statistics of Trade Act](#), already have an obligation to

provide energy consumption data to The Department of Energy Security and Net Zero (formerly BEIS) annually.

Suppliers also have additional data on the exact postal address (not just the MPAN location) of the smart meter. However, as suppliers aren't geographically focused, creating aggregated national, regional, city and street level datasets requires bringing the aggregated data from all the suppliers together. For the data to be useable and meaningful for the use cases identified, standards for publishing (technical, legal, operational) will be required. *We propose that the indexing of the aggregate data, and access controls for this data where it is not possible to be made Open, would be done via the Trust Framework applied to Open Energy.*

Whether the DNOs or Suppliers become responsible for aggregation, we recommend that the Office for National Statistics be involved in the design and standard setting of the data aggregation, to ensure it be both *meaningful statistically* and *aggregated in such a way to prevent de-anonymisation*.

To achieve the goal of aggregated smart meter consumption data, we recommend the Smart Energy Code be reassessed and revised to align with UK GDPR, and ensure clarity for consumers on the purposes for which smart meter data is used and the lawful basis for processing for each use.

As the Smart Energy Code was introduced in 2013, and GDPR was introduced in 2019, as a consequence the code had to address and provide rules for privacy and data protection concerns that the GDPR would later create a wider framework and set of principles for. Principles of transparency, purpose limitation, data minimisation and storage limitation would be used together with individual rights to design a smart meter ecosystem from scratch. We propose that the Smart Energy Code must be reassessed and revised to align with GDPR, with particular attention to the purposes of processing that consent is applied.

We are not able to offer definitive comment on this proposal without further detail regarding the exact methods of aggregation and de-personalisation that are proposed to be used.

We propose that the following questions be addressed as a next step:

- Will de-personalisation methods mirror DNOs' privacy plans in their current forms, or will any alternative methods be used?
- Will methods be standardised across different DNOs' geographic remits?
- Will data only be aggregated, or will other de-personalisation methods be used? If so, which methods and for what purposes?
- What size of 'n' will be considered appropriate and safe for smart meter data aggregation, particularly in sparsely populated areas where postcode level aggregation does not provide sufficient privacy?

In our own work, IB1 recommends that anonymised personal data should only be made Open if it complies with recognised robust methodologies, for example as set out by the

ICO⁷. We recommend that Ofgem formally adopts similar, standardised guidance for DNOs if this option is taken forwards.

We caution that no form of de-personalisation can occur without some level of risk, which must be appropriately managed at the sector level. We propose that *Ofgem takes responsibility for risk management* going forwards, incorporating monitoring and adaptation of guidance if required. While by no means exhaustive, the primary risk we identify in this space is the *risk of individual re-identification through dataset combination*. This issue is complex and can evolve in unforeseen ways as the available data landscape changes, particularly with increased opportunities to combine energy and non-energy data sources. If de-personalised smart meter data is to be made Open, we recommend that *Ofgem establishes a clear point of contact through which concerns regarding re-identification can be raised*. We also recommend that *re-identification risks are catalogued by Ofgem, made publicly available, and reviewed on a regular basis* as the data landscape evolves.

Beyond the proposals outlined in this consultation, we agree with Ofgem's assessment that publishing Open, de-personalised smart meter data can be useful but has limitations. We emphasise the importance of developing a *parallel approach to Shared Data classification, access and licensing*. This combined approach to development of the sector's Open and Shared Data ecosystems will most effectively maximise the potential benefits of energy data to net zero, innovation, and wider society.

We emphasise that data that might be identifiable or materially de-anonymised should not be made available as Open Data. *Instead, a Trust Framework approach should be applied to enable access to it as Shared Data*. This will require risk assessments on a case-by-case basis.

Q11. What are your views on our position that this Data Asset should be published in a non-interoperable fashion by 14 October 2023, if the appropriate security controls are in place?

We strongly recommend that interoperability be made mandatory as a foundational requirement. Introducing any non-interoperable system into a digitalised market will increase friction, reduce success, increase costs, and be unfit for purpose.

Interoperability is core to the work of Icebreaker One, therefore, we recommend interoperability from the beginning. Publishing aggregated smart meter data in a standardised and interoperable manner will allow for greater access and use of this data to Data Users, and avoid possible duplication of effort, time, and money.

⁷ <https://ico.org.uk/media/1061/anonymisation-code.pdf>

Q12. What are your views on our proposal that DNOs collectively determine an interoperable methodology by 28 February 2024, for publishing aggregated smart meter consumption data?

Collective agreement can happen in parallel to interoperable design processes. We strongly recommend that interoperability be made a mandatory requirement for all systems from the outset of any work. This is a foundational architectural design attribute for a digitalised energy system.

“Interoperable by default” should mirror other DBP rules with a case to be made and justified (commercially, legally, technically, policy) as to use cases where it should not apply.

We propose that the following questions be addressed as a next step:

- What barrier is preventing licensees from building an interoperable methodology from the beginning? In our experience, it is slower and costlier to do this in a non-interoperable manner first.
- Under the current energy codes and the GDPR, Suppliers are the best placed parties to provide aggregated smart meter data, as they are the data controllers of individual smart meter data. Why would DNOs publish aggregated smart meter consumption data and not suppliers?
- What problem is trying to be solved here? From our research, we have found that for use cases prioritised by DNOs, they require more granular detail.
- Will there be a standardised methodology as to how DNOs aggregate the smart meter consumption data? From our research, each DNO aggregates the data differently (that we’re aware of).

We propose Ofgem reviews what DNOs can do with this data. Our industry and government-funded research shows that DNOs have a greater need for granularity. As mentioned above, we suggest a review and alignment of the Smart Energy Code and GDPR is at the heart of a lot of these challenges.

We also recommend that the collective determination of an interoperable methodology be convened and supported by a neutral party, who is able to include and engage with all relevant stakeholders (not just the DNOs) to ensure it meets user needs.

Q13. What are your views on our proposal that licensees treat Data Assets associated with flexibility market operation as Presumed Open?

‘Presumed open’ is not a legal basis under UK IP law. For data consumers to use any data (closed, shared or open) with confidence, at minimum, the data should be published under a licence. For data that cannot be published as Open Data - see the OE data sensitivity classes - we recommend *standardising shared licensing* to reduce legal friction. We have drafted Shared Data licensing proposals here.⁸

We *do not* recommend that asset-level data be made available as Open Data.

⁸ <https://docs.openenergy.org.uk/1.0.0/index.html>

We propose that asset-level data be available for discovery (via Open Metadata) and the asset-level data itself be made available as Shared Data to enable risks and controls to be managed, via a Trust Framework such as Open Energy. Open Energy was co-designed by industry, and funded by IUK and BEIS, to address this category of use case.

Q14. Do you foresee any specific barriers to treating Data Assets associated with flexibility market operation as Open Data?

We strongly recommend that Data Assets associated with flexible market operation not be published as Open Data.

There may be issues around areas including but not limited to national security (e.g. new attack vectors for bad actors), competitive interests, IP, privacy, and legal challenges.

Utilisation of a Trust Framework for permission / controlled access to such data will enable risks and controls to be applied, technically and legally. As a roadmap, there may be categories of data that can be 'moved to open' using the Trust Framework approach (i.e. taking a data set, sharing it with a small group of registered actors and, based on continued risk assessment, increasing the size of the group). Over time it may be deemed valid to make certain data open. However, any restriction (e.g. "only for use in the UK") will require the implementation of a Trust Framework.