

Your Online Data Architecture

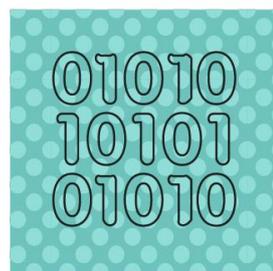
Data Licensing and Standardisation

Richard Dobson

Practice Manager Data Systems

9th July 2020

SIEMENS
Ingenuity for life



Contents

Foreword.....	2
1. Introduction to Data Licensing.....	3
1.1. License Options.....	3
1.2. Standard Licenses.....	4
1.3. Current Data License Usage.....	5
1.4. User Research.....	7
2. Data Standards.....	8
2.1. Current Landscape.....	8
2.2. Emerging Standards.....	9
3. Summary and Recommendations.....	11

Foreword

The path to a zero-carbon future is exciting, challenging and realising this ambition will call for innovative and disruptive ideas, alongside maximising proven sustainable and low carbon solutions. To ensure the UK achieves our carbon neutral commitments, it is critical that the UK energy sector maximises the value of existing (and future) national infrastructure, resources and the significant levels of existing and emerging data.

Realising this value however will call upon the energy sector to undertake a data-enabled cultural evolution. An evolution which actively enables open-data exchange, utilising data fed from multiple energy vectors, encompassing the breadth of the energy system and the plethora of stakeholders within. A true 'whole systems' approach which will give rise to a secure and managed service, a service affording access to the central data exchange for stakeholders wishing to illicit positive, innovative and lasting change within the energy sector.

Siemens, together with its partners the Energy Systems Catapult (ESC) and National Innovation Centre for Data (NICD) will deliver a 'digitally integrated energy system' which supports a Common Data Architecture concept. Underpinning the vision is an implementation of the open-data platform, constructed upon the requirements of the users and employing a sector specific metadata standard to drive commonality, enabling data-exchange. To do this Siemens will create 'Your Online Digital Architecture'.

The platform will be constructed upon the three relevant building blocks identified within the report 'Energy Data Taskforce: A Strategy for a Modern Digitalised Energy System' – incorporating asset registration strategy, data catalogue and digital system mapping.

Siemens promotes an inclusive approach to successful deployment, one which will be employed from beginning to end, facilitating user requirement capture workshops alongside show and tell events to provide insight toward the project outcomes and providing best practice guidelines. These events will support organisations who wish to utilise and embed the insight and outcomes, covering topics such as data transparency, data licensing and liability wavers, and data obfuscation / data protection techniques. Each of which are reflective of the metadata standard and Common Data Architecture underpinning 'Your Online Digital Architecture'.

The challenge is clear – the true value in data, in support of the transition to a low carbon economy, is in enabling visibility, access and insight throughout the energy value-chain – the industry must embrace this opportunity as a true catalyst for change, creating an open, yet secure, data marketplace which will create a modern, digitalized, energy system – one which drives system cost reduction, increases levels of asset visibility, improves system stability, provides capability for informed system management approaches and enables impactful innovation at scale. All critical factors in a decarbonized, digitalized and decentralized energy system.

Andrew Smyth - Head of Customer Success, Software

Siemens Smart Infrastructure

1. Introduction to Data Licensing

Data Licences are essential to ensure that both data providers and data consumers understand what data can be used for and are protected from legal ambiguity. Within this section we discuss the licencing and associated warranty statements / liability wavers which are currently in use, what standards exists and what has been identified as a core user need from stakeholder workshops.

Whilst there is a distinct lack of maturity in the Energy sector with regards to data licencing, this is not a new topic and many other industries have had to tackle this issue. For example, the creative arts sector has had to adapt from an industry where value is largely based on scarcity of supply (e.g. limited venue capacity) and difficulty to replicate (e.g. specialist equipment required to duplicate vinyl or photographs) to one largely based on digital media which is easily duplicated with minimal cost or technical skill. This change has provided many opportunities to increase the size of an audience and maximise impact but has also presented a challenge to ensure that artists are recognised and rewarded for their efforts.

There have been many solutions developed which range from open solutions to commercial licences. These allow the resource owners to decide what rights they are willing to give away and what they are seeking to get back in return e.g. attribution, fees, etc.

1.1. License Options

Data Licences provide the data user with a clear understanding of what they are able to do with data once it has been shared. There are a number of factors which need to be considered which have been outlined in the table below.

Factor	Description
Usage	<p>What is the data user able to do with the information?</p> <p>Options include:</p> <ul style="list-style-type: none"> • None <ul style="list-style-type: none"> ○ Data can be used for any purpose • Non-Commercial <ul style="list-style-type: none"> ○ Data can be used for any purpose providing it is not for direct commercial gain. • Research and Innovation only <ul style="list-style-type: none"> ○ Data can be used by research institutions or for innovation projects only • Non-competitive <ul style="list-style-type: none"> ○ Data can be used for any reason providing the use case does not directly impact the commercial interests of the publishing organisation • Bespoke restrictions <ul style="list-style-type: none"> ○ Limiting specific use cases for the data e.g. a photograph cannot be used to advertise certain products or services
Sharing and Sub-licencing	<p>Is the data user allowed to share this onwards?</p> <p>Options include:</p> <ul style="list-style-type: none"> • None <ul style="list-style-type: none"> ○ No onward sharing of data or derived products is allowed • Derived products only <ul style="list-style-type: none"> ○ Derived products can be shared providing the original data is not exposed • Matching Terms <ul style="list-style-type: none"> ○ Data can be shared providing matching licence terms are used, this may include pass through of revenue • No restrictions

	<ul style="list-style-type: none"> ○ Data can be shared in any format with any terms
Duration and Renewal	<p>How long can the data be stored and used for?</p> <p>Options include:</p> <ul style="list-style-type: none"> • Perpetual <ul style="list-style-type: none"> ○ Data can be stored and used indefinitely • Time limited <ul style="list-style-type: none"> ○ Data can be stored and used for a restricted period of time e.g. 1 year. The licence may 'auto renew' at the end of the term.
Updates and Latency	<p>Will the data be updated within the contract?</p> <p>Options include:</p> <ul style="list-style-type: none"> • No Updates <ul style="list-style-type: none"> ○ Data is static and any updates will need to be accessed separately • Scheduled Updates <ul style="list-style-type: none"> ○ Updates will happen at predefined points in time e.g. monthly • Continuous Updates <ul style="list-style-type: none"> ○ Data will be updated as soon as new information is available
Costs	<p>What cost (if any) is there for accessing the information?</p> <p>Options include:</p> <ul style="list-style-type: none"> • Free <ul style="list-style-type: none"> ○ No charge for access or use of the data • Flat fee <ul style="list-style-type: none"> ○ One off payment for use of the data within the terms of the licence (usage, sharing, duration) • Usage dependent <ul style="list-style-type: none"> ○ A fee that is linked to how the data is used e.g. price per API call

1.2. Standard Licenses

There are a number of standard licences which are already well understood which ensure that potential users know what they can and cannot do with data. The majority of these licences relate to open or permissive licences

Category	Description	Licence
Public Domain	Resource is dedicated to the public domain. Publisher is relinquishing all rights to the dataset	PDDL: This license is one of the open data commons licenses and is similar to public domain dedication. It allows you as a dataset owner to relinquish your rights in a dataset when you might otherwise not be able to dedicate your dataset to the public domain.
Creative Commons	A range of more permissive licences that make resources available to a wide audience with some limited restrictions	<p>CC-0: One of the most open creative commons licenses, similar to public domain and means the publisher has relinquished rights to the dataset, where they have not been able to dedicate to the public domain</p> <p>CC-BY: An open creative commons license, whereby the dataset user can share and adapt the dataset, but must give credit to publisher</p> <p>CC-BY-SA: Creative commons license where the user can share and adapt the data set, so long as they give credit to the publisher and distribute any additions, transformations or changes to the dataset under this license. Can be problematic, as user may decide not to use license as there is</p>

Category	Description	Licence
		<p>a risk that their work in the dataset will need to be shared also.</p> <p>CC-BY-NC : A more restrictive creative commons license, where the user can share and adapt the dataset, as long as they give credit to the publisher, but they mat not use the dataset for any commercial purposes.</p> <p>CC-BY-ND : Creative commons license where the user can share and adapt the dataset if they give credit to the publisher, but they cannot make any transformations, additions or changes to the dataset under this license.</p> <p>CC-BY-NC-SA : Creative commons license where user can share the dataset if they a) Give credit to the publisher, b) Do not use the data for commercial purposes and c) Distribute any additions and transformations or changes to the dataset under this license. Users will need to share their work under this license and any users of the adapted dataset and so on.</p> <p>CC-BY-NC-ND : A restrictive creative commons license, where users can share only the publishers unmodified dataset if they give credit to the publisher and do not share for any commercial purposes. Users can not make any transformations, additions or changes to the dataset under this license.</p>
Community Data Licence	Collaborative licenses to enable access, sharing and use of data openly among individuals and organizations. A Linux Foundation project.	<p>CDLA-Permissive 1.0 : One of the community data license agreements, similar to the permissive open source licenses. It allows the user to use, modify and adapt the dataset and the data within it, as long as they give credit to the publisher.</p> <p>CDLA-Sharing 1.0 : One of the community data license agreement licenses and was designed to embody the principles of ‘copyleft’ in a data license. It allows users to use, modify and adapt the dataset and the data within it, and to share the dataset and data with their changes, as long as they do so under the CDLA-Sharing and give credit to the publisher.</p>
Open Data Commons Licence	Open Data Commons exists to provide legal solutions for open data. An Open Knowledge Foundation project.	<p>ODC-BY : One of the open data licenses and allows users to share and adapt the dataset, as long as they give credit to the publisher</p> <p>ODC-ODbl : One of the open data commons licenses and allows users to share and adapt the dataset, so long as they give credit to you and publish any additions, transformations or changes under this license. Problematic as there is a risk that any work done on the dataset will too need to be shared under this license.</p>
Other		<p>MIT Licence</p> <p>Copyleft (e.g. GNU)</p> <p>Open Government Licence (OGL)</p>

1.3. Current Data License Usage

The Energy sector publishes a large amount of data via industry data portals. In general, these are used to publish open data as this is most widely publicised but there are many areas where data is shared with more restrictive terms or in return for payment. Note, many data sharing processes in the energy sector are legally underpinned by licence requirements or code agreements.

In the table below we have reviewed a sample of open and shared data portals and tried to identify the data licence in use and have noted if there was any warranty or liability waiver information included.

Company	Data Licence	Data Warranty Information
National Grid	Open <ul style="list-style-type: none"> Creative Commons License Open Data License Public Domain 	Information is licensed 'as is' and the information provider excludes all representation, warranties, obligations and liabilities in relation to the information
Western Power	Open <ul style="list-style-type: none"> Creative Commons License Public Domain Bespoke terms and conditions 	Included in terms and conditions for individual tools.
Scottish Power	Open <ul style="list-style-type: none"> CC-BY: Creative Commons By Attribution 	No warranty given or should be implied to the overall data quality
British Gas – Centrica	Licence could not be identified	Assurance note – more related to the validity of financial information
Renewables Ninja	Open <ul style="list-style-type: none"> CC-BY-NC: datasets available with creative commercial licensing. But cannot be used for commercial re-sale 	Data is available 'as-is' and without warranty. No guarantee of accuracy and accept no responsibility for any liability arising from use
EDF Energy	Licence could not be identified	-
The Energy Market Data Hub (Elektralink)	Open: <ul style="list-style-type: none"> Licence could not be identified Shared: <ul style="list-style-type: none"> DCUSA Code Agreement Bespoke legal document - 'AGREEMENT FOR THE PROVISION OF A DATA TRANSFER SERVICE' 	Discussed in code and bespoke agreement
Centre for Sustainable Energy	Open: <ul style="list-style-type: none"> Open Government Licence (OGL) 	-
UK Power networks	Open <ul style="list-style-type: none"> Open Government License Creative Commons License 	Included in terms and conditions for individual tools.
Elxon	Open <ul style="list-style-type: none"> BMRS Open Data Licence: You can copy, share and adapt the data, including for commercial use. You must acknowledge the source of data and including a link to the license where possible. Shared <ul style="list-style-type: none"> BSC Code Agreement P30, P114 and P315 Licence Agreements Bespoke Licences as required 	Data given as-is, no warranty given over the quality of the data
Energy Systems Catapult	Open and Shared <ul style="list-style-type: none"> Bespoke Licence – 'Energy Systems Catapult Data Sharing Licence' 	Within licence text
EPEX Spot	Shared <ul style="list-style-type: none"> Bespoke agreement 	Within agreement text

1.4. User Research

The user research which has been carried out as part of the YODA project found broad support for open data but it is clear that stakeholders from across the industry acknowledge that there will be a need for a more secure, controlled data sharing infrastructure if we are to unlock the full value of data.

Secure Data Sharing

Many stakeholders were interested in a secure data sharing solution which enables their organisation to confidently share data with the confidence that they know who the recipient is, that the data will be secure in transit and that the data is stored, processed and exploited appropriately.

This requirement is important for a number of reasons:

1. Data custodians and users want to ensure their systems are secure
2. Data custodians and users want to limit the risk of data sharing (regulatory, legal, data protection, etc.)
3. Data custodians want to ensure that the wishes of the data subject are respected (e.g. privacy or commercial sensitivity of data)

Data Monetisation

Whilst many stakeholders were interested in open data a number of stakeholders suggested that the platform should be capable of enabling the sharing of data with a range of licences including paid data.

This requirement is key to ensuring that the platform is able to be used by a whole range of users including those outside of the core regulated sector and helping to support energy data innovators.

Licence Clarity

Data users expressed a need to have a very clear set of licence conditions so that they are able to understand what they are able to do with data before they start working with it. In addition, there was an interest in datasets being made available with different licences. For example, a dataset might be made available with a free but restrictive licence which enables data exploration but no commercial use with the option to upgrade to another less restrictive paid licence when a business model has been developed.

2. Data Standards

The evolution of the energy sector, and the digital systems which support it, has resulted in a diversity of data structures and interfaces which have been designed to meet a specific organisational or sector need. Until recently, this approach has broadly met the needs of the sector and there has been little incentive to change. However, we are now at the starting to experience a seismic shift in the energy sector with the increasing prominence of renewable energy, deployment of high demand appliances (EVs, Heat Pumps, etc.), rapidly changing gas demand profiles and other factors all creating the need for better visibility and control of the whole system. This shift necessitates much more data sharing than ever before and exposes the lack of interoperability of data structure and interfaces.

Within this section we discuss energy data standardisation (structures and interfaces), focusing on both existing and developing solutions which have either been formally defined or have organically emerged. In some cases, data standardisation will be a by-product of other interventions such as engineering standards, codes or licences but these are not the main focus.

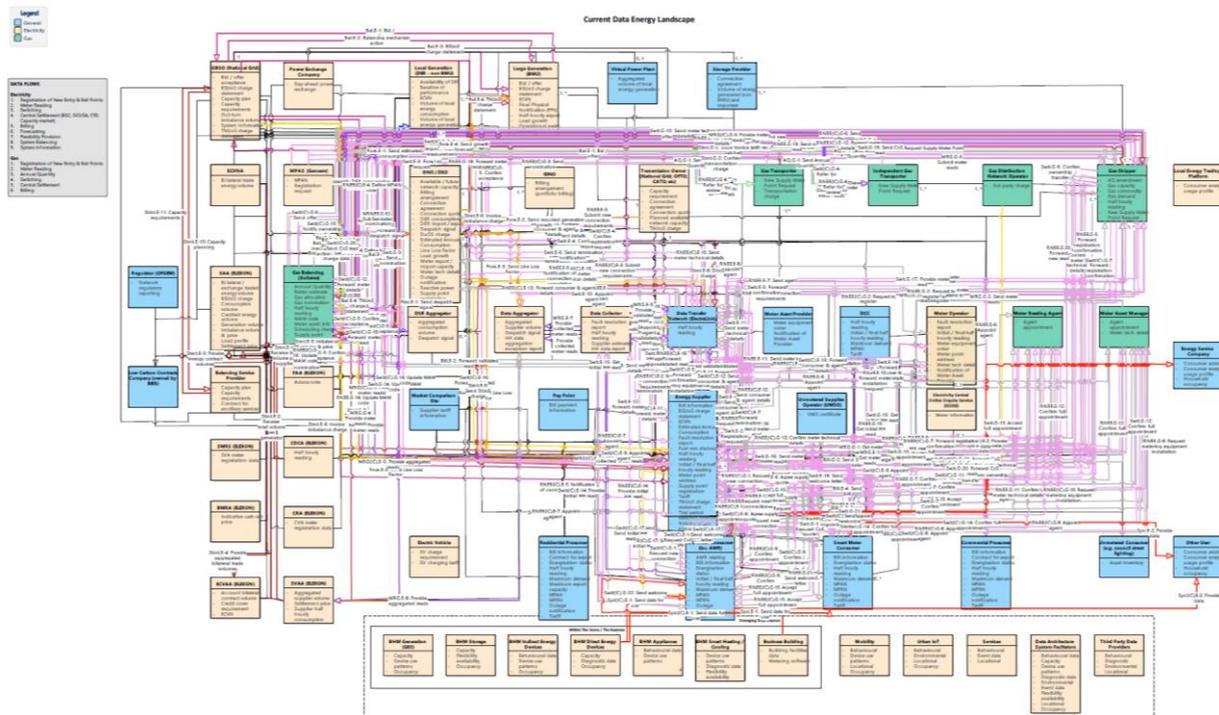
2.1. Current Landscape

At present, the vast majority of data across the energy sector is in a bespoke format which has been designed to fit the need of individual organisations but influenced by the IT solutions have been deployed, previous ownership structures and regulatory obligations. Whilst entirely understandable this has created a patchwork of data structures and interfaces which do not provide interoperability out of the box.

There are some areas where data sharing is integral to the sector e.g. meter readings, switching suppliers, balancing and settlement, system operation, etc. In these cases, data standardisation can occur through the appointment of a single entity that is responsible for delivering a service e.g. Data Transfer Service (Electralink), Balancing and Settlement (Elexon), Gas Data Services (Xoserve). In other cases, standardisation is more distributed and based on industry agreement of standardised requests and responses e.g. Week 24/42 submissions, connection requests, etc. In these cases, standards emerge organically based on a specific need (or business efficiency), the standards themselves may or may not be formally defined, governed and enforced.

- Data Services
 - Electralink Data Transfer Service – <https://www.electralink.co.uk/dts/>
 - Elexon BSC - <https://www.elexon.co.uk/guidance-note/bmrs-api-data-push-user-guide/>
 - XOServe Gas APIs - <https://www.xoserve.com/services/gas-api-services/>
 - Electricity Central Online Enquiry Service
- Industry Agreements
 - Week 24 / 42 submissions (National Grid ESO and DNOs)
 - Regulatory Reporting Requirements (Ofgem)
 - Flexibility definitions (DNOs and Flexibility platforms)

Energy Systems Catapult contracted with Baringa and Delta EE to develop a view of the data flows across the energy sector in 2018. The work focused on energy supply and found a complex landscape made up of many small but highly structured data transfers that underpin key operations. These flows are often defined by Codes but some are more flexible.



In contrast, there are a small number of areas where standards have been defined up front which have subsequently been adopted. The most prominent case of this is the SMETS 1/2 smart meter standards which, alongside describing the technical requirements for elements of the system, dictate what data is collected, how it is structured and how it is accessed. The SMETS standards are enforced through legal and technical and enforcement. The Common Information Model is an electricity network standard which is owned and maintained by the IEC, this has been in existence for some time but has recently started to gain traction in the UK thanks to reporting requirements and innovation projects.

- Formal Standards
 - CIM - <https://www.iec.ch/smartgrid/standards/>
 - Smart Meter Equipment Technical Specifications (SMETS) 2 - <https://smartenergycodecompany.co.uk/the-smart-energy-code-2/>

2.2. Emerging Standards

In addition to the informal and formal standards discussed above there are a number of emerging standards which are either exist or are in development. Many of these may not be deployed widely but others are likely to gain traction.

There are a number of industry initiatives underway that seek to develop and deploy data standards, some are driven by government and regulator, but others have more industry involvement. The government or regulator is leading on a range of projects that are seeking to improve the quality and consistency of data in order to improve consumer outcomes (e.g. Long Term Development Statement update, midata, Smart Tariff Comparison, etc.). Industry led initiatives include formal standards bodies projects such as the BSI Smart Appliances standards, initiatives led by companies such as Northern Power Grid’s DFES work with the ODI Leeds and academic / industry collaborations such as the Digital Framework Task Group work on digital twin infrastructure.

- Government and Regulator
 - Long Term Development Statement - https://www.ofgem.gov.uk/system/files/docs/2019/12/consultation_-_key_enablers_for_dso_programme_of_work_and_the_long_term_development_statement_0.pdf
 - Midata (supplier switching) – <https://www.ofgem.gov.uk/gas/retail-market/market-review-and-reform/midata-energy-project>
 - Smart Tariff Comparison - <https://www.gov.uk/government/publications/smart-meter-enabled-tariffs-comparison-project-smarter-tariffs-smarter-comparisons>
 - Metadata – Energy Data Visibility Project and Modernising Energy Data Access
- Industry
 - DFES Standardisation (Northern Power Grid) - <https://odileeds.org/blog/2019-09-26-building-data-standards-for-future-energy-systems>
 - BSI Smart Appliances - <https://www.bsigroup.com/en-GB/smart-appliances-flexible-energy/>
 - Digital Framework Task Group - https://www.cdbb.cam.ac.uk/files/approach_summaryreport_final.pdf

There are a number of open standards which have been developed over the years which present options for digital and data standardisation. These range from standards which focus on specific technologies such as EV charge points through to full representations of the energy system. LF Energy is a particularly interesting project in that it is backed by the Linux Foundation and aims to provide an internationally relevant set of open source standards from control system software through to energy data. Some of the Open standards included below have traditionally been developed by the academic sector and open source community but have been receiving more interest from industry in recent years.

- Open Charge Point Protocol (OCPP) - <https://www.openchargealliance.org/>
- LF Energy - <https://www.lfenergy.org/>
 - Open Energy Data Initiative - <https://www.lfenergy.org/projects/oedi/>
- Open Energy Ontology - <https://openenergy-platform.org/ontology/>

Open Street Map: Power networks - https://wiki.openstreetmap.org/wiki/Power_networks

3. Summary and Recommendations

From the research carried out on data licensing, it is apparent that though there are many different licenses available for publicly open data, there needs to be a clearer understanding about the current licensing landscape and the availability of data within the energy sector. From the user research as part of the YODA project, it is clear that work still needs to be carried out to encourage both data providers and data users that there are safeguards in place to keep data secure and the licenses will protect custodians from legal ambiguity.

There are a number of robust licences that exist in the open data space. However, there are many cases where an open licence is not possible or desirable but there is still significant value in sharing data. At the moment, these cases are dealt with through bespoke agreements – either regulated licences, code agreements or bilateral contracts. This creates significant overhead and does not scale well and restricts innovation.

There is a clear need for additional standard licences to be developed which provide data custodians with the ability to implement some common terms. It is advised that this be taken forward as an Alpha deliverable to further understand the range of licences which are required and implement a licence framework which enables these.

Energy Systems Catapult supports innovators in unleashing opportunities from the transition to a clean, intelligent energy system.

Energy Systems Catapult

7th Floor, Cannon House
18 Priory Queensway
Birmingham
B4 6BS

es.catapult.org.uk
info@es.catapult.org.uk
+44 (0)121 203 3700