Access to energy information now and for the future

Reviewing the Energy Semantic Artefacts Landscape

Catherine Jones, Karen VanHaltren and Teagan Zoldoske

March 2025

https://doi.org/10.5286/UKERC.EDC.000987

Table of Contents

List of figures

1. Introduction

The Energy Data Centre (EDC) is a capability of the UK Energy Research Centre (UKERC) (www.ukerc.ac.uk), which is a multidisciplinary research centre performing whole systems energy research to support a just transition to net zero.

The EDC's repository (www.ukerc.rl.ac.uk) provides access to energy information by holding records on research data, grey literature and projects. It uses additional subject classification schemes to add context to both the digital objects held and metadata only records to assist in effective retrieval.

Being able to use common terminology helps in the discovery, use and reuse of objects in the energy sphere. The FAIR (Findable, Accessible, Interoperable and Reusable) principles [1] specify the use of domain specific common terminology which can be referenced and accessed by machine. The first interoperable principle states that "*(meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.*" In the EDC's current assessment of adherence to the FAIR principles [2] this is currently not met. There are two purposes for this work: firstly to establish if there is a single semantic artefact that satisfies this criteria for the energy community and secondly to identify possible candidates that would improve the current classification schemes in use.

The energy field for research, development and operation is very multi-disciplinary, it ranges over domains from psychology to heavy engineering, including materials development, economics, politics and agriculture. This spread of domains means that there is not a single community to develop agreed controlled vocabularies but many, each focussed on their needs. It should be noted that this report reflects activity in the semantic artefact domain, not necessarily how active the research areas are.

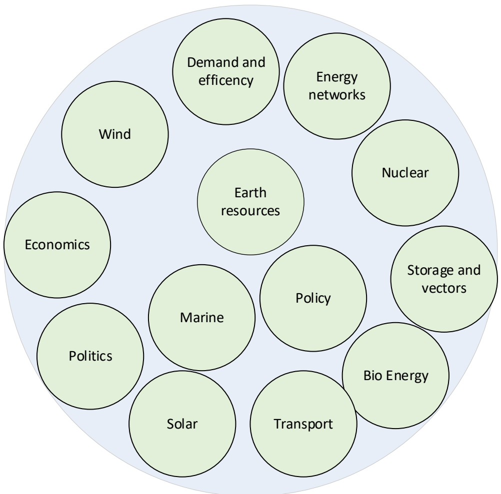

Figure 1 Visual representation of domains covered by the term Energy

In this report we use the term **semantic artefact** [3] to describe the variety of methods used to formally describe concepts in a specific domain. These methods range from glossaries which define the meaning of words (known as terms) to machine readable ontologies which define concepts and their relationships. All these methods and resulting semantic artefacts are designed to ensure a common understanding of the meaning of words and hence the context of the digital object, both for humans and for machines.

This report introduces the semantic artefact concept and subject classification in general, compares and contrasts selected schemes; discusses the practical application of subject terms to the EDC, demonstrates the issues around applying terms from different schemes to the same digital object and makes recommendations for next steps for both Data Infrastructure for National Infrastructures (DINI) and the EDC.

The table in Appendix I, shows the schemes that the authors have considered. We acknowledge this may not be a comprehensive list, but as we have considered different domains and sectors as part of the selection process, we propose it as a good starting point for our analysis. All schemes considered use English.

This work has been funded by the Department of Science, Innovation and Technology (DSIT) as part of the DINI pilot study. DINI aimed to explore the barriers to data sharing in the Infrastructure Systems Engineering domain with a focus on Energy, Water and Transport with a recommendation to design and build a federated infrastructure of data repositories. So, as the DINI project is considering the barriers to data sharing the complexity in this semantic artefact space has been identified as a barrier for the community but the EDC in particular to becoming FAIR-er.

1.1. Background

Using a standard agreed subject classification scheme to add meaning and context to the description of digital objects, or the metadata about an object is a well-established technique in the information management profession. Ranganathan [4] in the 1930s identified the fourth of his five laws of library science as "*Save time for the reader*" where he introduces the idea of subject classification to aid discovery.

By using a formal scheme, the terms used are within a conceptual view of the world and it reduces the use of synonyms which can impede effective retrieval. The different approaches bring a variety of levels of sophistication. However, the terminology describing these semantic artefacts can be confusing. For this report we use the following definitions for the variety of semantic artefacts that can be used to add additional meaning:

- **Controlled vocabularies:** Lists of terms and their definitions. Usually used to add additional meaning.
- **Thesauri:** List of terms with some relationships between terms, such as “narrower”, “broader”, “preferred term” or “deprecated term”. It is implementing a hierarchy of terms.
- **Taxonomies:** List of terms and their definitions. Usually hierarchical.
- **Ontologies:** A set of terms with a wide range of relationships possible. It is not necessarily hierarchical.
- **Glossaries**: Lists of terms and their definition. Usually generated from source material to ensure common understanding of meaning. Not usually used to add semantic meaning to the description of a digital object but to add clarity within it.

The FAIR principles for interoperability recommend the use of common vocabularies so that benefits of using semantic artefacts are gained and that the semantics can be understood by machines as well as humans.

1.2. Approaches to creating and applying energy semantic artefacts

In the terminologies we have looked at there are two main purposes for the creation and application of the scheme: to add subject information to metadata records to aid effective retrieval of information or to enable reporting of statistics for comparison with other countries/same country over a period of time.

Of the schemes identified, some are in active use in one or more system/process and some have been created for a specific project which may not persist after the end of the project funding. With the increase in computing capacity, there is more interest in either identifying terms through automated means or applying terms using machine learning beyond the traditional approach of humans creating and applying these vocabularies.

However, having a formal representation, or creating one, does not guarantee the consistent application of terms. As discussed by Cleverdon [5] if two or more people/groups each create a thesaurus, only 60% of the terms will be in common, if two or more experienced indexers use a controlled vocabulary for a given document only 30% of the terms will be in common, and finally that scientist/engineers (i.e. end users) review documents for a given subject there will only be 60% agreement. So, while the semantic artefacts themselves may be carefully constructed, the application and coverage need always to be considered and if consistency is key, quality assurance practices should be applied.

As noted in Chowdhury [6] “*the characteristics and specific needs of users determine the nature of the information to be collected by the system"* This concept is known as the

designated community [7] within the information and archival communities and all information retrieval systems should have a clear understanding of their target audience. By having a specific designated community in mind, the expectation is set of the level of knowledge of the users and hence the level of additional subject information. So, for a specialist service, then the subject terms assigned are likely to start at a more detailed level, and not necessarily use very generic subject terms than those assigned in a service that is aimed at a wider community. Applying high level terms, such as “energy” in an energy focused service where all the records would then have that term doesn’t add as much semantic value as adding “energy” in a service which covers more than the energy domain.

For all these semantic artefacts, when used in a specific context, it is likely that there would be instructions on how terms are applied in practice to ensure consistency across many potential people applying the scheme. Additionally, there need to be policy and procedure around how the service/ user addresses changes in the semantic artefact and how that is reflected in the use of the scheme. There needs to be active management to ensure the application of the ontology remains current.

2. Comparative analysis of semantic artefacts

The understanding of the purpose of applying a semantic artefact is key to choosing the appropriate scheme to use. However, the schemes must be discoverable with enough context to judge whether they are appropriate. This section discusses the landscape survey for ontologies and other semantic artefacts used in the energy space, see Appendix I for details, and the analysis of the overlaps between the different schemes and any possible gaps.

2.1. Methodology

Different types of energy resources or schemes such as ontologies, controlled vocabularies, taxonomies and glossaries were investigated. We performed a wildcard search for the term “Energy” within various search portals such as TIB Terminology Service. This search included results which had “Energy” as preferred terms, synonyms and within descriptions to include all possible terms related. In total, 28 resources composed of both energy specific and general schemes that had ‘Energy’ as a topic were identified as having potential. We selected schemes that reflected the various disciplines of energy research and after reviewing each scheme to determine if they involved the energy sector, e.g., terms relating to biological energy were removed, this number was refined down to 19.

The results of this search were collated into a spreadsheet. The spreadsheet was organised with each scheme receiving their own sheet and with two further sheets that was used for analysis. Within the first analysis sheet “raw”, individual key terms were identified from all the schemes and some basic mapping was completed manually. This mapping was primarily done to the Open Energy Ontology (OEO) due to the rich choice of terms available within it; however, some terms were mapped to other schemes at the author’s discretion in order to try and improve consistency within the verbiage used, e.g., all sources of energy would be listed as “X Energy”. In total, 2,607 unique terms were identified within all the schemes, and after mapping, 53 terms appeared to cover the same topics across at least two separate schemes and were placed in the other analysis sheet “combined”. Once all the key terms were identified and refined, they were listed alongside the schemes and counts were made for the number of times that the term appeared in each scheme with a total count shown next to each term in both analysis sheets. Using a filter within the spreadsheet, simple analysis could be done to see which terms were used most often and which schemes were using them.

Information on the schemes is available in Appendix I.

2.2. Findings

As noted in the introduction, the schemes examined are those which the authors have discovered and reflect not the academic research activity per se but areas where there is semantic artefact activity. In particular the dominance of Wind Energy related schemes can be explained by the active community in the Research Data Alliance and the International Energy Authority taskforce in this topic.

We also observed that the topic of energy does not belong to a single discipline or domain rather it is ‘a range of topics revolving around the intersection of energy technologies, fuels, and resources on one side; and social processes and influences—including communities of energy users, people affected by energy production, social institutions, customs, traditions, behaviours, and policies—on the other’ [8]. The diagram below shows the broad subject domains that the resources we identified fall into.

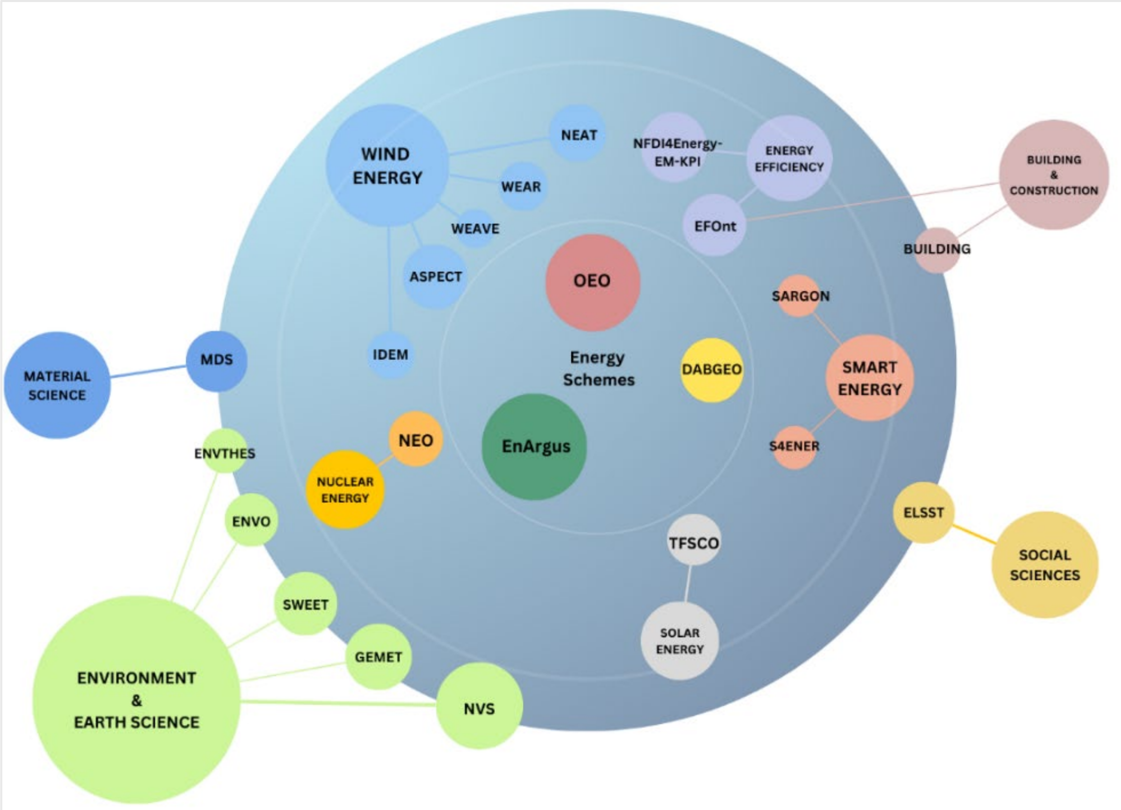

Figure 2 Diagram showing the resources identified and their broad domains

Looking at the different schemes, they largely fall into the following categories:

Controlled Vocabularies

Energy-domain specific:

- ASPECT: wind energy vAriableS ParametErs and ConsTants – *Wind energy*
- EnArgus – *Energy research and funding*

Other domain with Energy as a sub-category:

- NVS: The NERC Vocabulary Server - *Environment Science*

Thesauri

Other domain with Energy as a sub-category:

- GEMET: General Multilingual Environmental Thesaurus – *Environment Science*
- ELSST: European Language Social Science Thesaurus – *Social Science*
- ENVTHES: Environmental Thesaurus – *Environment Science*

Taxonomies

Energy-domain specific:

- IDEM: wInd energy moDEls Taxonomy – *Wind energy*
- NEAT: wiNd Energy tAxonomy of Topics – *Wind energy*
- WEAR: Wind Energy mAteRials Taxonomy – *Wind energy*
- IRENA: International Renewable Energy Agency – *Renewable energy*

Ontologies

Energy-domain specific:

- OEO: Open Energy Ontology – *Energy systems*
- WEAVE: Wind Energy ActiVitiEs – *Wind energy*
- NFDI4Energy-EM-KPI: NFDI4Energy-Energy Management-Key Performance Indicators Ontology – *Energy management*
- DABGEO: Domain Analysis-Based Global Energy Ontology – *Energy management*
- S4ENER – *Smart energy*
- SARGON: SmArt eneRGy dOmain oNtology – *Smart energy*
- NEO: Nuclear Energy Ontology – *Nuclear energy ontology*
- EFOnt: Energy flexibility Ontology – *Building/Energy efficiency*
- TFSCO: Thin-film solar cell ontology – *Solar energy*

Other domains with Energy as a sub-category:

- ENVO: The Environment Ontology – *Environment Science*
- SWEET: Semantic Web for Earth and Environmental Terminology – *Environment Science*
- MDS: Materials Data Science – *Material Science*
- BUILDING: Building Ontology – *Construction and building*

We further categorised them according to their size (total number of terms) and their focus on energy. We discovered that only a few large, comprehensive energy schemes exist, while many smaller energy schemes address specific energy domains. The general schemes with energy as a topic are also illustrated in the figure.

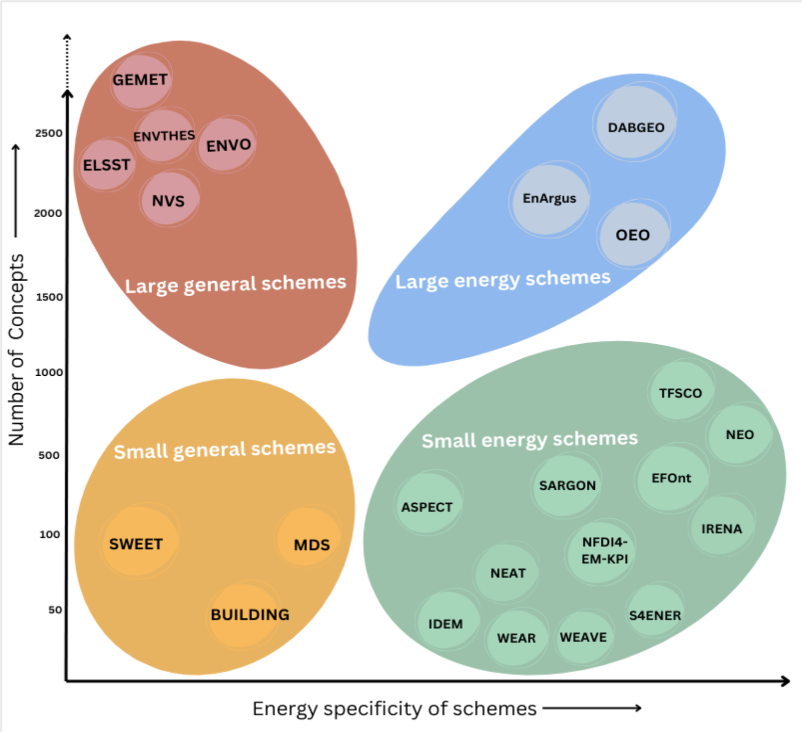

Figure 3 Diagram showing general size and specificity of the schemes reviewed

In addition to the aforementioned schemes, we also analysed glossaries and standard terminologies from various global organisations within the Energy sector. Our findings indicated that there is extensive coverage in specialised energy areas such as Renewables, Fuels and Transport, Energy storage etc. The organisations and energy sectors are shown below:

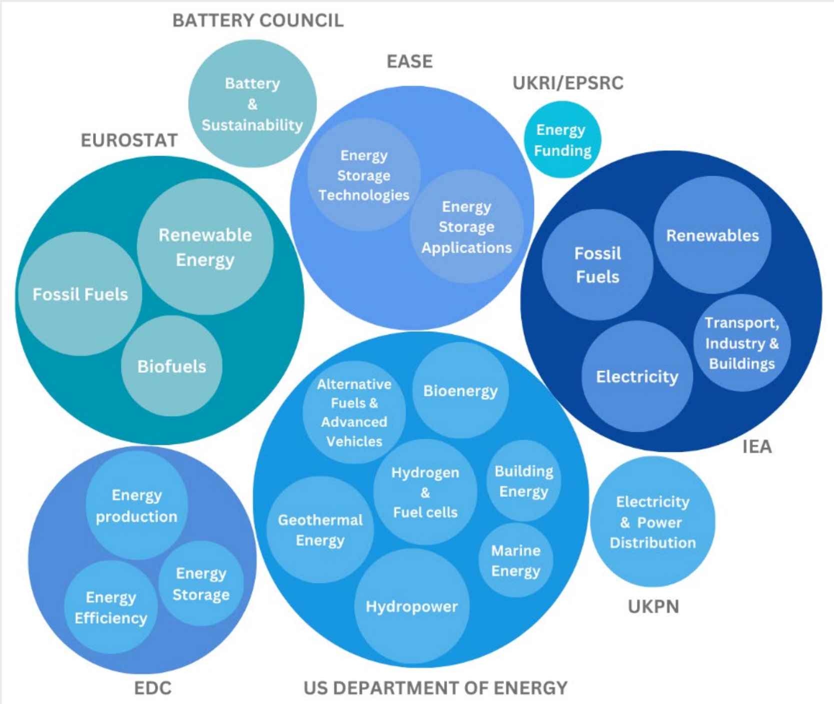

Figure 4 Diagram showing various global organisations and some of their sectors that include energy schemes

- UKPN: UK Power Networks – *Power Generation & Distribution sector*
- DUKES: Digest of United Kingdom Energy Statistics- *Business and economy*
- Eurostat – *Fossil fuels, Biofuels, Renewable energy sectors*
- Battery Council International – *Energy storage sector*
- EASE: European Association for Storage of Energy – *Energy storage sector*
- US Department of Energy – *Geothermal energy, Bioenergy, Marine energy, Hydropower, Hydrogen & Fuel cells, Building energy, Alternative Fuels & Advanced vehicles sectors*
- IEA: International Energy Agency – *Fossil fuels, Renewables, Electricity, Transport Industry & Buildings sectors*
- EPSRC: Engineering and Physical Sciences Research Council (UKRI) – Engineering *and physical sciences*
- EDC: Energy Data Centre (UKRI) – *Energy research*
- ERI-MAP: Energy Research Investment (UKRI)

Energy related terms from 19 resources were used in our further analysis. These included ontologies (OEO, Building, WEAVE, NFDI4Energy-EM-KPI, DABGEO, S4ENER, SARGON, SWEET), controlled vocabularies (ASPECT), thesauri (ELSST, ENVTHES), taxonomies (IDEM, NEAT, WEAR), together with EDC, ERI-MAP, DUKES, INSPEC and IEA's RD&D: Energy research, development and demonstration. The terms were then compared to the full list of terms on an overview page of the spreadsheet to identify what were the main concepts used within all the schemes. Through this comparison of the 19 schemes, the following terms were found most frequently:

- Solar Energy
- Wind Energy
- Geothermal Energy
- Hydro Energy
- Coal
- Renewable Energy Sources
- Biofuels
- Fuel Cells
- Battery
- Bioenergy
- Energy

We further reviewed the relationships between the most frequent terms and schemes and created a matrix of overlapping terms and schemes (Figure 4). We found OEO as the most comprehensive ontology having 10 out of the 11 frequently occurring terms and IEA and EDC as well as ERI-MAP as the most comprehensive controlled terminology schemes, with 9 and 8 out of the 11 terms, respectively.

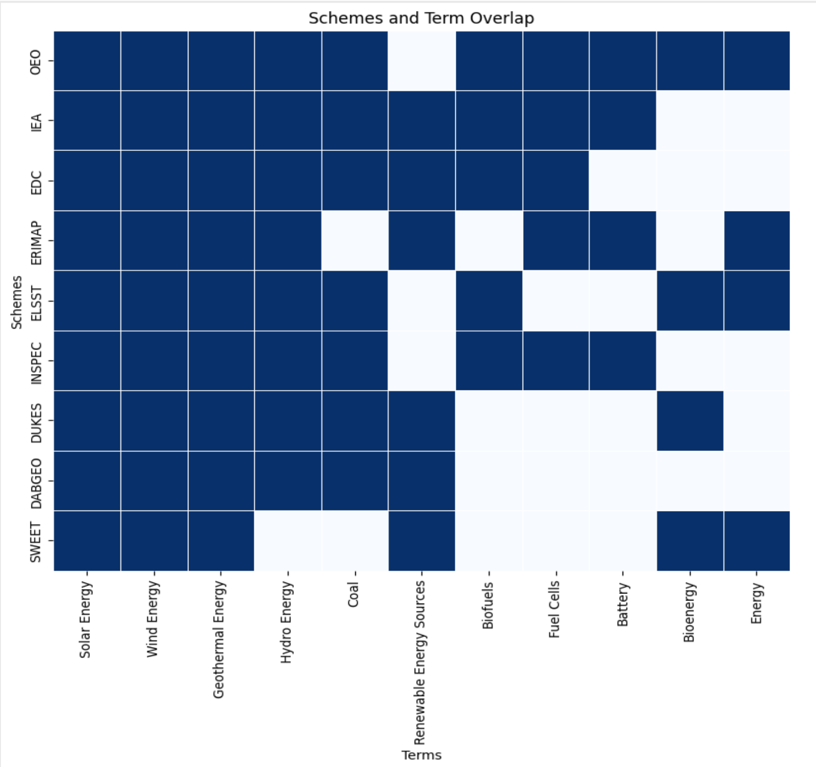

Figure 5 Plot showing the overlap in common terms in different schemes

There is the potential for further overlap, however, the individual terms of the various schemes were not standardised, so terms such as “Biofuel” and “Biofuels” would count as two separate terms as would “Biomass” and “Biomass Energy”. Some mapping was done for these terms to help determine similar types of concepts. We think that it's also important to note that there is minimal reference to non-renewable energy sources or transporting energy, just storage. That is not to say that these were left out of the schemes, only that they do not appear to have shared higher-level terms and will require more refinement of the shared schemes to group them into something that can be compared more easily.

2.3. Analysis of gaps

To ensure that gaps truly existed in the schemes analysed, a basic mapping of like terms was done. This helped minimise the potential for similar terms being missed, such as Hydro Energy and Hydropower. This mapping confirmed that most of the schemes identified provided terms for wind energy but the overlap about what related terms there were with wind energy was minimal, indicating that the schemes, while having a similar term, each focused on different elements relating to wind energy, e.g., materials, activities, parameters.

Further analysis of the schemes themselves showed minimal metadata outside of the terms which made determining the update and uptake of these schemes difficult. Many of the schemes appeared to be redundant or not maintained and over the course of the analysis. An example of this was the retirement of the Humanities and Social Science Electronic Thesaurus (HASSET) which was created and used by UK Data Archive (UKDA). It has been replaced by European Language Social Science Thesaurus (ELSST). This makes determining the reusability of the schemes difficult.

3. Case study: Application of semantic artefacts in the Energy Data Centre (EDC)

The EDC provides a portal to energy information and data management support to the UK Energy Research Centre. The portal (www.ukerc.rl.ac.uk) uses a variety of semantic artefacts to enhance discoverability and to ensure consistency. This case study demonstrates the benefits of using a formal scheme to add subject based context and discusses the policy and procedures that need to be in place to ensure that changes in the world, and formal scheme are reflected in a specific application.

3.1. Approach

The main subject classification scheme used, called *energy categories* within the portal is based on the IEA scheme [9] which is used to report national energy R&D statistics. It has been used in the EDC, since 2004, with only one period of amendment to address some areas not related to energy production and no modifications made by the IEA during the last 20 years have been adopted.

This scheme has seven headings covering a variety of energy domains, but with a focus on energy production. It implements a hierarchy of up to three levels but there are no formal explicit relationships (e.g. narrower term or broader terms) encapsulated in the EDC implementation of the scheme.

For the bulk of the first fifteen years the scheme was applied by a single person, and hence there was no need for classification guidance to ensure consistency. As the team has evolved these have been formalised into the EDC's Cataloguing and Classification Guidance.

1. → Introduction¶

This·document·formalises·the·cataloguing·rules·and·conventions·for·metadata·records·held·in·the·EDC·and·gives·guidance·on·classification/controlled·vocabulary.·Selection·for·cataloguing·follows·the·guidance·in·the·EDC·Collection·Management·Policy.··This·supersedes·previous·versions·of·the·cataloguing·policy¶

Figure 6 Introduction from the internal EDC Cataloguing and Classification guidance

The EDC uses this scheme to add additional context to the records in the portal. The records relate to funded projects, datasets and publications. All records, regardless of type, have at least one term and may have many assigned depending on the subject matter of the item being described. These energy categories are then used within the portal to discover, explore and filter content

RENEWABLE ENERGY SOURCES	Solar Energy	Solar heating and cooling (including daylighting)
		Photovoltaics
		Solar thermal power and high-temp. applications
	Wind Energy	
	Ocean Energy	
	Bio-Energy	Production of transport biofuels (incl. Production from wastes)
		Production of other biomass-derived fuels (incl. Production from wastes)
		Applications for heat and electricity
		Other bio-energy
	Geothermal Energy	
	Hydropower	Large hydropower (capacity of 10 MW and above)
		Small hydropower (less than 10 MW)
	Other Renewables	

Figure 7 Snippet of the energy category scheme used, showing the hierarchy

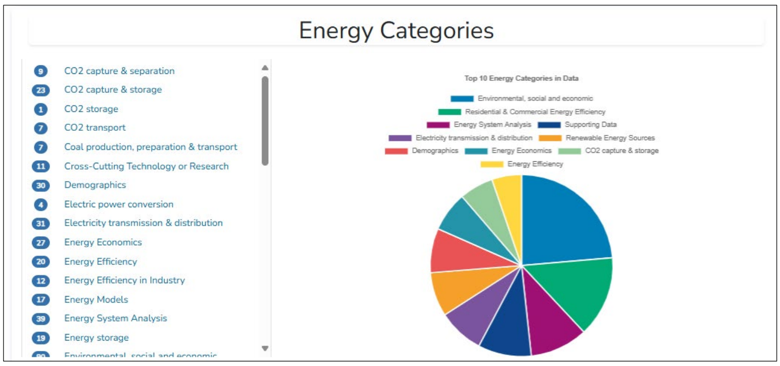

Figure 8 Energy categories as applied to datasets, used as a way of browsing content

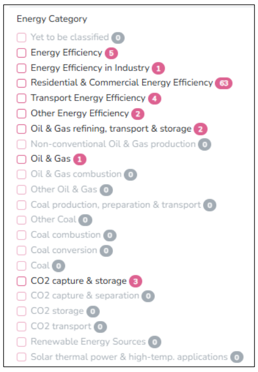

Figure 9 Filtering options after searching for "domestic" in publications

The different types of content that the EDC either hold or manage are at different levels of granularity as regarding assigning subjects. The two key differences are the scope of the item being described and where it is on its lifecycle. Funded projects (grants) cover a range of activities over a prolonged period with significant investment and are classified at the point of funding; whereas datasets and publications are usually focussed on a much narrower topic. Secondly for datasets and publications the artefact is available for inspection when classifying but for funded projects the information available to classify it is based on the researcher's grant application and intention. These differences can make the process of applying energy categories, which are at a fairly high level, challenging.

To address specificity, there are two further schemes that are applied to records. For projects one or more concepts from a "science and technology area" scheme is added to identify the type of research area that is contributing to the grant. This scheme was based on work done by UNESCO [10] and the Royal Academy of Engineering (scheme no longer discoverable) and maps at a high-level research domain, but with more detail in the science and engineering fields. It implements a basic hierarchy. As the whole systems approach increases, then more energy related grants are a mixture of technical, cultural and policy strands, this scheme doesn't reflect the nuances around cultural and policy aspects. For data and publications, a broad "subject" is added to the record which is used in a similar fashion as the science and technology area to focus in on the purpose of the material within a broad energy category. This approach is used in part to overcome issues with the energy categories not addressing energy demand. These subject terms have been identified by the data stewards within the EDC team and reflect the content that is held. Only one subject term is assigned to a record. This is a controlled vocabulary as there are no relationships recorded between the terms.

3.2. Reflections

The fact that a limited number of people have applied the terms to the content since 2004 has led to a consistent application of the concepts to a large corpus of records and as a result enabled effective retrieval reliability. Adapting an existing scheme has the advantage of being able to re-use the concepts and definitions already agreed and reduced effort in repeating work already done.

However, adopting a scheme and then not reflecting changes in the energy environment that the scheme models, has led to areas, such as heat, energy vectors or energy demand reduction not being described effectively, leading to gaps in discoverability. In addition, the purpose of the original scheme was for energy R&D statistics reporting which means that it is biased towards areas that statistics can be generated for. One of the aims of this work is to establish the best way of addressing the gaps in a systematic way that doesn't involve internal modifications to an existing scheme. The EDC team are currently identifying areas of the energy categories that do not reflect our requirements.

Any decision to amend the terms then leads to a philosophical and operational decision on how to implement any changes. Should the new terms be used from a set date and any old terms remain but are deprecated, or should the collection be reclassified using the updated scheme. Reclassifying an existing collection of thousands of records manually is an enormous task and it is unlikely that the EDC would have the resource to do so for the value added by doing so. It is more likely that new terms would be adopted, and old terms deprecated through our policies and procedures, adapting our cataloguing and classification policy to reflect new practices.

The EDC implementation of the schemes themselves are not machine readable and one of the next steps would be to enable this for the schemes we currently use. The EDC doesn't currently benefit from the navigational/browsing aspects that the application of the implicit hierarchy in the scheme uses provides or the wider benefits that using an ontology with richer relationships might provide.

Extending our existing schemes with terms from other schemes may address the known issues with the scheme and mapping between schemes is vital to enable wider data sharing between different repository services.

In conclusion, the EDC's portal benefits from the use of controlled vocabularies and subject classification schemes together with the consistency of application as it enhances the discoverability of the content. Using a formal scheme means that there are operational decisions to be made when there are changes to the concepts within the scheme that reflect changes in the domain the scheme models which have resourcing and discoverability consequences.

4. Case study: Comparison of application of terms

Choosing relevant subject terms from any type of semantic artefact is dependent on the purpose of adding them, the community that will be using them and the personal experience of the human, or machine that sets the terms. There is no one right answer in this space.

However, to make digital objects understandable in a federated world, some form of mapping between subject terms is needed and these examples show the potential issues involved in that process.

This section of the report takes examples from funded projects and research datasets classified within the EDC and also held elsewhere to demonstrate the similarities and difference of the terms assigned in the different systems. The authors acknowledge that it is a very small sample, but they highlight some of the issues.

For grants, the comparison services are: the EDC, Catalogue of Projects on Energy Data (CoPED) [11] currently run by the Energy Systems Catapult and developed by Coventry University and the UKRI Gateway to Research (GtR) [12]. CoPED was developed through the EnergyRev project [13] and focusses on energy projects and was aimed at academics and researchers from other sectors, but in the energy field. Gateway to Research shows information of projects funded by UKRI and so covers all domains, it is aimed at researchers and the general public to show what has been funded with public money.

For research data the comparison services are the EDC and the UK Data Archive (UKDA) provided by the UK Data Service (UKDS)[14]. The UKDA is a domain repository covering social, economic and population data, whose audience is researchers looking to re-use research data generated by others.

CoPED used machine learning techniques to identify and apply terms to their records, while UKDS, GtR and the EDC use humans, either expert staff or the depositor/grant Primary Investigator to identify and assign terms.

For all types of classification, whether human or machine, the available information such as the title and description/abstract are used to determine what terms should be applied, in accordance to the classification conventions. For text-based objects, then the content can also be inspected. Machine-assisted classification needs to ensure that the context of the word/phrase is adjusted for; e.g. "high energy physics" is not considered to be a branch of energy research. In Appendix II, this metadata on the chosen items discussed in the next sections are outlined.

4.1. Methodology

For grants, CoPED was used as a starting point to identify energy grants funded by EPSRC. EPSRC was chosen as a funder as it assigns research topics to the grants in GtR and the majority of energy research projects are funded by the EPSRC. Having chosen two records on different topics, using terms that mapped to areas of the IEA classification which has different levels of hierarchy, these were then looked up in the EDC and GtR. For all the records, the terms in fields labelled "subject", "topic" or "energy category" were recorded and compared. We did not consider other fields such as type of project, or other types of controlled vocabulary.

For research data, the starting point was metadata only records held in the EDC for data generated in UKERC Phase 4 where the data was lodged with the UKDA. Research data in this sector is usually held only in one service, however the EDC has a metadata only record for UKERC phase 4 data held in the EDC. For all the records, the subject(s) assigned to them were recorded and compared.

All the comparator services use terms from a controlled vocabulary; however, some are limited to energy concepts (EDC), some cover wider concept spaces (UKDA – social science and GtR – EPSRC portfolio) and the CoPED covers a mix of subject terms and other useful descriptive information. The record identifiers are in Appendix II.

4.2. Comparison of terms assigned: grants

The first grant is the "*MARLIN Modular Floating Platform for Offshore Wind: Concept Assessment*".

The table below shows the subject terms assigned in the three different systems:

CoPED terms	EDC terms	GtR terms
1. Renewable energy sources	Renewable Energy Sources (Wind Energy)	Energy - Marine and Hydropower
2. Projects		Energy
3. Energy policy		
4. Sustainable development		
5. Climate policy		
6. Development (active)		
7. Developing countries		

Figure 10 Table showing the terms assigned to the first project example

There is overlap in that all three systems assigned to renewable energy at a variety of levels. The EPSRC term is the narrowest for the domain and the CoPED term is the broadest. CoPED adds further subject terms beyond those associated with the energy research field, some of which may be covered by other metadata fields in the other systems or are not relevant for their community or are just not covered.

The EDC assigns terms from other controlled vocabularies to capture information on the impact/policy perspective.

The second example is "*EPSRC Centre for Doctoral Training in Fuel Cells and their Fuels - Clean Power for the 21st Century*"

CoPED terms	EDC terms	GtR terms
1. Fuel cells	Hydrogen and Fuel Cells (Fuel Cells, Stationary applications)	Unclassified
2. Hydrogen	Hydrogen and Fuel Cells (Fuel Cells, Mobile applications)	
3. Fuels	Hydrogen and Fuel Cells (Fuel Cells, Other applications)	
4. Biogas	Hydrogen and Fuel Cells (Hydrogen)	
5. Development (active)		
6. Renewable energy sources		
7. Training programmes		
8. Optimisation		

Figure 11 Table showing the terms assigned to the second project example

In this case the GtR record has not assigned a term, for the other two services there is an overlap in the main concept, but CoPED has some more detailed terms. , not all of which may be as relevant or accurate but have been added due to the description discussing using hydrogen as a fuel and so the mention of ".neutral feedstock, biological processes" has lead to the use of the term "biogas" which is not accurate in this context.

Again, some of the additional CoPED terms are covered in the EDC by other pieces of metadata – such as the fact it is a training grant.

4.3. Comparison of terms assigned: research data

It should be noted that the subjects covered by the UKDS are poorly served in the classification scheme that the EDC uses.

Considering the dataset "*Perception Spillover From Fracking, 2022" which* was generated from the UKERC Phase 4 theme on Energy Infrastructure Transitions.
This dataset considered the impact on other forms of renewable energy technology, in this case deep enhanced geothermal systems, and green hydrogen, of controversy about fracking technology.

UKDA	EDC
PERCEPTION	Fossil Fuels: Oil Gas and Coal
ATTITUDES	Oil and Gas
ENERGY	Non-conventional oil and gas production
GEOTHERMAL ENERGY	Hydrogen and Fuel Cells
	Hydrogen
	Other infrastructure and systems R&D
	Other Cross-Cutting Technologies or Research
	Environmental, social and economic impacts
	Renewable Energy Sources
	Geothermal Energy

Figure 12 Table showing the terms assigned to the second data example

It can be seen that the technologies considered in the study are more detailed in the EDC subject terms however it is clearer from the UKDA that the study was about human perceptions and not technical development.

Considering the dataset "Baseline Food Refrigeration Emissions in the UK, 2019-2020" This was generated by a UKERC Phase 4 Flexible Fund project. The following terms were used.

UKDA	EDC
CARBON DIOXIDE EMISSIONS	Other Cross-Cutting Technologies or Research
ENERGY	Environmental, social and economic impacts
ENERGY CONSUMPTION	Energy Models
GREENHOUSE GASES	

Figure 13 Table showing the terms assigned to the first data example

The UKDA has the high-level term ENERGY as it is in a wider social sciences disciplinary repository and so this additional context needs to be given, all the content within the EDC is about energy so this is implicit. The EDC can only represent energy consumption through using a broader term about impacts.

4.4. Reflections

While this case study only looks at a very small number of records, it demonstrates the point that the audience and purpose of assigning terms has an impact on the terms chosen. The EDC and CoPED remits are both focussed on the energy sector and hence there is no need to assign an overarching "Energy" term; whereas the UKDA serves the social science community and hence adding "Energy" as a term adds value to the discovery of the content.

The classification of records in the EDC is carried out by professional staff with a knowledge of energy research who inspect the record's title and metadata and underlying documentation. However, the EDC classification scheme is weighted towards energy production, rather than demand or impact.

While CoPED is a domain restricted service, the subjects assigned go beyond energy and in other services some of the information may be provided by other metadata.

These examples illustrate the challenges involved in comparing semantic meaning between services, using different schemes. Aligning or mapping is not straightforward activity but may bring benefit if resources are federated together. Careful consideration of the granularity or specificity of the underpinning resources is key to an effective federated resource.

5. Recommendations for DINI

This review has demonstrated the complexity of the semantic artefact space in the energy research domain, as DINI aims to support infrastructure systems engineers from other related domains such as water and transport, this will only increase the complexity. Being able to map terms between well-established and well used schemes will become more important while understanding what the required granularity and precision is needed for the community DINI is supporting.

As discussed in the introduction and the application case study, the consistent application of the scheme is as important as the scheme itself. The community building and outreach functions within DINI could act as a catalyst to bring together interested parties on the topic of interoperability of semantic meaning.

6. Conclusions

There is a need for consistent, agreed concepts identified through semantic artefacts to aid discovery and navigation of energy research data content. There is much activity about this in many fields within energy research and development, but there is no overall co-ordination at the top-level of the terminology.

The FAIR principles expect data creators to use common community vocabularies that are machine enabled. The landscape review demonstrates the range and scope of work in this area for the energy research community. The analysis undertaken defines the schemes by scope and shows the overlap and discusses the gaps within the schemes we have looked at.

The EDC has been applying subject classification terms to the content held within it for over twenty years. During this time the main energy classification scheme implemented has remained static and now there is an opportunity to address areas of the scheme which are not well conceptualised. It demonstrates the issues of applying semantic meaning consistently over the long term as the world changes and highlights the service decisions and resource requirements needed to implement these. Application of semantic artefacts needs proactive management to ensure that they remain a relevant tool.

This study has shown that a consistent application of a classification scheme which takes into account the audience of the service is very important. However, even the consistent application of a semantic artefact within a service may not make the discovery of data in a federated resource more effective - to do this requires a comprehensive mapping of different vocabularies, which although seemingly very complicated and time-consuming will be worthwhile in the longer term.

7. References

[1] Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al*. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

[2] EDC's Adherence to FAIR principles https://ukerc.rl.ac.uk/cgi-bin/aboutPolicyFair.pl (Accessed Feb 2025)

[3] Le Franc, Y., Parland-von Essen, J., Bonino, L., Lehväslaiho, H., Coen, G., & Staiger, C. (2020). D2.2 FAIR Semantics: First recommendations (1.0). FAIRsFAIR. https://doi.org/10.5281/zenodo.5361930

[4] Ranganathan, S.R. (1931). The Five Laws of Library Science. New Delhi: Madras Library Association

[5] Cleverdon C. Optimizing convenient online access to bibliographic databases. Information Services and Use. 1984;4(1-2):37-47. doi:10.3233/ISU-1984-41-204

[6] Chowdhury, G.C. Introduction to modern information retrieval 3rd edition, Facet publishing, London, 2010

[7] Definition of Designated community - Wikipedia (accessed March 2025)

[8] Energy Research & Social Science (ERSS) retrieved from Homepage

[9] IEA Guide to Reporting Energy RD&D Budget/Expenditure Statistics (2011) https://iea.blob.core.windows.net/assets/3432ae79-1645-4cf1-a415-faa3588e6f29/RDDManual.pdf (Accessed March 2025)

[10] UNESCO Thesaurus, 1977 Available from https://vocabularies.unesco.org/browser/thesaurus/en/groups (Accessed March 2025)

[11] Energy Projects Catalogue (CoPED) available from https://coped.energy/projects/ (Accessed March 2025)

[12] UKRI's Gateway to research available from https://gtr.ukri.org/ (Accessed March 2025)

[13] EnergyREv project available from https://www.energyrev.org.uk/ (Accessed March 2025)

[14] UK Data Service available from https://ukdataservice.ac.uk/ (Accessed March 2025)

Appendix I: Schemes examined in the analysis

The table in this section outlines the schemes that were reviewed and some basic information about them.

List of Energy Ontologies and General ontologies with Energy as a subcategory

Ontology	Type	Topic	URL	Creation History	Scope	Features	License	Host	Creator /Funding
Open Energy Ontology (OEO)	Energy Ontology	Energy	https://openenergyplatform.org/viewer/oeo/	Created: 2021 Updated: 09/12/2024 Maintained	The Open Energy Ontology is an ontology for all aspects of the energy modelling domain. It is developed in four main modules: 'oeo-model', a module for all entities related to models and modelling; 'oeo-physical', a module for all entities related to the world of energy and energy generation; 'oeo-social', a module for all relevant social and economic aspects of the energy domain; 'oeo-shared', a module to cover entities and relations needed in all of the modules above to prevent them getting implemented multiple times. A supplementary module is the oeo-physical-axioms module, which contains general class axioms.	- Machine readable - Has description of concepts - Adheres to FAIR principles - Source code on Github - Has persistent identifiers	CC0-1.0	- Open Energy platform	Government funding (German)

NFDI4Energy-EM-KPI Ontology	Energy Ontology	Smart Grid	https://energy.linkeddata.es/em-kpi/ontology/index-en.html#	Created: 2019 Updated: N/A Not Maintained	The EM-KPI ontology describes the key performance indicators and master data domains (energy, building, utility, occupancy, observation, weather, location) in energy management at district and building levels. The ontology contains two parts: the KPI (key performance indicator) part and the EM (energy management) master data part; these respectively represent the multi-level performance information for energy performance tracking and the master data for data exploitation.	- No description of concepts - No information on FAIR adherence - No persistent identifier	CC BY-NC-SA 4.0	- Web interface	Government funding
DABGEO: Domain Analysis	Energy Ontology	Energy	https://innoweb.mondragon.edu/ontologies/dabgeo/index-en.html	Created: 2019 Updated: 22/02/2019 Not Maintained	DABGEO is a large-scale ontology that includes 97 modules. DABGEO provides a common representation of the energy domains represented heterogeneously by the available energy ontologies developed for specific applications.	- Ontology with smaller ontologies - Has description of concepts - No information on FAIR adherence - No persistent identifier	CC-BY 4.0	- Website	Javier Cuenca (Mondragon Unibertsitatea), Felix Larrinaga (Mondragon Unibertsitatea) Edward Curry/Government funding/Basque
S4ENER – Smart energy	Energy Ontology	Smart home	https://saref.etsi.org/saref4ener/Energy	Created: 2016 Updated: 29/03/2023 Not Maintained	SAREF4ENER focuses on demand response scenarios, in which customers can offer flexibility to the Smart Grid to manage their smart home devices by means of a Customer Energy Manager (CEM).	- Has description of concepts - No persistent identifiers - Many terms are deprecated - Number of Classes:146 - Number of Properties: 113 - Number of Individuals: 30	BSD-3-Clause	- TIB Terminology service - Website	Created by EEBus and Flexiblepower Alliance Network
SARGON: SmArt eneRGy dOmain oNtology	Energy Ontology	Smart assets	https://sargon-n5geh.netlify.app/ontology/1.0	Created: 2020 Updated: N/A	SARGON consists of an extensible dictionary of terms and concepts in and around building and smart grid, a set of relationships for linking and composing concepts together, and a flexible data model	- No description of concepts - No persistent identified - No information on FAIR adherence - Source code available	CC-BY 4.0	- Website - GitHub	Government funding/Germany

				Not Maintained	permitting seamless integration of SARGON with existing tools and databases	- Number of Classes: 170 - Number of Properties: 196 - Number of Individuals: 55			
EnArgus	Search engine / Energy Ontology	Energy	https://www.enargus.de/	Created: 2021 Updated: N/A Maintained	With EnArgus, the Federal Ministry of Economic Affairs and Climate Action provides an Internet portal that provides information on current and completed research projects related to energy research.	- WIKI terms with energy topics and description - No persistent identifiers	CC BY-SA 3.0 DE	- Website	Federal Ministry of Economic Affairs and Climate Action / Government funding
Nuclear Energy Ontology (NEO)	Energy Ontology	Nuclear Energy	https://github.com/emmo-repo/domain-neo	Created: N/A Updated: N/A Maintenance unknown	The Nuclear Energy Ontology (NEO) is a domain ontology developed under the EMMO framework for the representation of concepts in the field of nuclear energy		CC-BY- 4.0	- Github	N/A
Energy Flexibility Ontology (EFOnt)	Energy Otology	Building /Energy efficiency	https://github.com/LBNL-ETA/EnergyFlexibilityOntology	Created: N/A Updated: N/A Maintenance unknown	Energy Flexibility Ontology (EFOnt) is an open-source development effort to create a standardized schema for describing, characterizing, and quantifying building energy flexibility to support Grid-interactive Efficient Buildings.		N/A	- Github	USDOE Office of Energy Efficiency and Renewable Energy (EERE), Energy Efficiency Office. Building Technologies Office; USDOE Office of Energy Efficiency and Renewable Energy (EERE)
Thin-film solar cell ontology (TFSCO)	Energy Ontology	Solar Energy	https://matportal.org/ontologies/TFSCO	Created; 2022 Updated: 19/8/2024 Maintenance unknown	The Thin-film solar cell ontology (TFSCO) is a domain ontology that provides a model of the manufacturing and characterization of perovskite solar cells.	- Number of classes: 651 - Number of properties: 53	CC-BY- 4.0	- MatPortal	Dr. Jens Hauch (HIERN), Dr. José Márquez (Humboldt-Universität zu Berlin), Prof. Dr. Ulrich W. Paetzold (KIT), Prof. Dr. Eva Unger (HZB) and Dr. Thomas Unold (HZB)

Materials Data Science (MDS)	General Ontology	Material Science	https://matportal.org/ontologies/MDS	Created: 24/03/2024 Updated: 04/08/024 Maintenance unknown	Materials Data Science (MDS) is an ontology encompassing multiple domains relevant to materials science, chemical synthesis and characterizations, photovoltaics and geospatial datasets. The terms used for classes, subclasses and instances are mapped to PMDCo and BFO Ontologies.	- Number of classes: 256 - Number of properties: 12	N/A	- MatPortal - Industry portal	Alexander H. Bradley, Jonathan E. Gordon, Van Tran, Priyan Rajamohan, Quynh Tran, Gabriel Ponon, Yingjui Wu, Laura S. Bruckman, Erika I. Barcelos, Roger H. French
SWEET: Semantic Web for Earth and Environmental Terminology	General Ontology	Environment	https://bioportal.bioontology.org/ontologies/SWEET	Created: 2019 Updated: 14/07/2022 Maintained	Semantic Web for Earth and Environmental Terminology (SWEET) Ontologies is a highly modular ontology suite with ~6000 concepts in ~200 separate ontologies covering Earth system science.	- Has description of concepts - Adheres to FAIR (47%) - No persistent identifier - Concepts: 6000 - Number of Classes: 10238 - Number of Properties: 383 - Number of Individuals: 2148	CC0-1.0	- BioPortal - EarthPortal - GitHub	ESIP foundation
BUILDING: Building Ontology	General Ontology	Smart Home	https://terminology.tib.eu/ts/ontologies/building	Created: 2019 Updated: 20/10/2020 Maintenance unknown	The model is constructed as an extension of the BOT ontology that provides the vocabulary to describe the topology of a building as well as the relationships between their main components such as zones, spaces, and building elements.	- Number of Classes: 45 - Number of Properties: 44	CC-BY-4.0	- TIB Terminology service - BIMERR	María Poveda-Villalón, Serge Chávez-Feria/ Ontology Engineering Group, Universidad Politécnica de Madrid
ENVO: The Environment Ontology	General Ontology	Environment	https://sites.google.com/site/environmentontology/	Created: N/A Updated: 07/01/2024 Maintained	ENVO is an ontology which represents knowledge about environments, environmental processes, ecosystems, habitats, and related entities	- Has description of concepts - Has persistent identifiers - Number of Classes: 7040 - Number of Properties: 317 - Number of Individuals: 44	CC0-1.0	- Ontobee - OLS - Website - GitHub	Suzanna Lewis, Norman Morrison Christopher Mungall Pier Luigi Buttigieg / European Commission, the National Human Genome Research Institute, and the U.S. Department of Energy

List of Energy Taxonomies

Ontology	Type	Topic	URL	Creation History	Scope	Features	License	Host	Creator /Funding
IDEM WInd energy moDEls Taxonomy	Energy Taxonomy	Wind Energy	https://bioportal.bioontology.org/ontologies/IDEM	Created: 2017 Updated: 05/10/2021 Not Maintained	A classification of models used in wind energy	- No description of concepts - No persistent identifiers - No information on FAIR adherence	CC-BY 4.0	-NCBO Bioportal - Ontology viewer	Anna Maria Sempreviva
ASPECT: wind energy vAriableS ParametErs and ConsTant	Energy Taxonomy	Wind Energy	https://data.windenergy.dtu.dk/ontologies/view/aspect/en/	Created: N/A Updated: 14/10/2021 Not Maintained	Controlled vocabulary of variables, parameters and constants used in wind energy community.	- Preferred terms: 143 - Alternate terms: 35	CC0-1.0	- NCBO Bioportal - Ontology viewer	Technical University of Denmark, DTU Wind Energy
NEAT: wiNd Energy tAxonomy of Topics	Energy Taxonomy	Wind Energy	https://bioportal.bioontology.org/ontologies/WETAXTOPICS	Created: 2017 Updated: 14/10/2021 Not Maintained	A taxonomical organization of research topics in wind energy which follows a typical lifecycle of wind farm development.	- No description of concepts - No persistent identifiers - Adheres to FAIR principles - Machine readable - Preferred terms: 69 - Alternate terms: 8	CC-BY 4.0	- NCBO Bioportal - Ontology viewer	Nikola Vasiljevic, Danielle Preziuso, Anna Maria Sempreviva
Wind Energy mAteRials (WEAR) Taxonomy	Energy Taxonomy	Wind Energy	https://bioportal.bioontology.org/ontologies/WEAR	Created: 2017 Updated: 05/10/2021 Not Maintained	A classification of wind turbine materials	- No description of concepts - No persistent identifiers - Adheres to FAIR principles - Machine readable - Preferred Terms: 25	CC-BY 4.0	- NCBO Bioportal - Ontology viewer	Anna Maria Sempreviva

Wind Energy ActiVitiEs (WEAVE)	Energy Taxonomy	Wind Energy	https://bioportal.bioontology.org/ontologies/WEAVE	Created: 2017 Updated: 23/05/2023 Not Maintained	A classification of activities in which data are produced	- No description of concepts - No persistent identifiers - Adheres to FAIR principles - Machine readable - Preferred Terms: 19	CC-BY 4.0	- NCBO Bioportal - Ontology viewer	Nikola Vasiljevic, Anna Maria Sempreviva

List of General Thesauri with energy as a subcategory

Ontology	Type	Topic	URL	Creation History	Scope	Features	License	Host	Creator /Funding
European Language Social Science Thesaurus (ELSST)	Thesaurus	Social Science	https://thesauri.cessda.eu/elsst-5/en/	Created: 23-09-2023 Updated: 2024 Maintained	The European Language Social Science Thesaurus (ELSST) is a broad-based, multilingual thesaurus for the social sciences. The thesaurus consists of over 3,400 concepts and covers the core social science disciplines: politics, sociology, economics, education, law, crime, demography, health, employment, information and communication technology, and environmental science.	- 3300 concepts - Has Uniform Resource Identifier - Adheres to FAIR	CC-BY-SA 4.0	- Website	Consortium of European Social Science Data Archives (CESSDA) and national service providers
ENVTHES: Environmental Thesaurus	Thesaurus	Environment	https://ecoportal.lifewatch.eu/ontologies/ENVTHES?p=summary	Created: 2013 Updated: 04/01/2025 Maintained	EnvThes compiles a set of terms in order to describe in a harmonised way data resulting from observations and measurements of ecosystem processes across different domain specific sciences. It is used by DEIMS-SDR for common keywords for annotation and querying metadata purposes.	- Has description of concepts - Adheres to FAIR (61%) - No persistent identifier - Number of Classes: 1191 - Number of Properties: 153 - Number of Individuals: 2	N/A	- EcoPortal - eLTER	N/A

GEMET: General Multilingual Environmental Thesaurus	Thesaurus	Earth Science	https://www.eionet.europa.eu/gemet/en/about/	Created: 2004 Updated: 17/1/2024 Maintained	GEMET was conceived as a "general" thesaurus, aimed to define a common general language, a core of general terminology for the environment.	- Has description of concepts - Adheres to FAIR principles - Number of Classes: 7 - Number of Properties: 33 - Number of Individuals: 5739	CC-BY 4.0	- Website - AgroPortal	European Environment Agency (EEA), European Environment Information and Observation NETwork (EIONET)

List of Controlled vocabulary and terminology schemes

Ontology	Type	Topic	URL	Creation History	Scope	Features	License	Host	Creator /Funding
Energy Research Investment (ERI-MAP)	Controlled Vocabulary	Energy	Not publicly available	Not released	Designed to classify energy projects. To be able to add subject discovery to a portal for exploring the relationship between UK funded energy projects.		N/A	Not recorded	Rachel Freeman, EPSRC funded project
NERC Vocabulary Server (NVS)	Controlled Vocabulary	Environment	https://vocab.nerc.ac.uk/	Created: 2006 Updated: N/A Not Maintained	SKOS concept collections held in the NERC Vocabulary Server. In the NVS, concept collections are synonymous with controlled vocabularies or code lists. Each collection is associated with its governance body. An external website link is displayed when applicable.	- Has persistent identifier - No description of concepts	CC-BY-4.0	- NERC website - GitHub	British Oceanographic Data Centre at the National Oceanography Centre (NOC)
Energy Data Centre (EDC)	Controlled Vocabulary	Energy	UKERC EDC: About	Created: 2004 Updated: N/A Not Maintained	Based on the IEA reporting subjects from 2004. It is focussed on energy production and efficiency	- Has 70 terms in a 3-level hierarchy	CC-BY-4.0	Not recorded	UKERC EDC

Appendix II: Details of the records used for comparison

This appendix shows the unique identifiers for the record used for comparison of subject terms assigned and the description used for classification.

Item	EDC Identifier	CoPED identifier	GtR identifier	UKDS identifier
MARLIN Modular Floating Platform for Offshore Wind : Concept Assessment	EP/P030645/1	5e8ee774-bbcb-44c4-b593-ea187925783b	EP/P030645/1	
EPSRC Centre for Doctoral Training in Fuel Cells and their Fuels - Clean Power for the 21st Century	EP/L015749/1	1f4cad26-dc99-4d9d-a182-41ade3e0832a	EP/L015749/1	
Baseline Food Refrigeration Emissions in the UK, 2019-2020	761ba953-0332-4807-81f6-1fdc9cb6513e			10.5255/UKDA-SN-855845
Perception Spillover From Fracking, 2022	8167dd41-1040-4f77-8c86-020d125fd9e6			10.5255/UKDA-SN-856047

Table 1 Unique identifiers

This table shows the abstract/description held in the EDC. Original text will have come from GtR or UKDA and will have been provided by the primary investigator/depositor.

Title	Abstract/description
MARLIN Modular Floating Platform for Offshore Wind : Concept Assessment	Project MARLIN will assess and develop a new concept for a modular floating platform system for offshore wind. The project will confirm technical and commercial feasibility of the novel method of construction and deployment of floating structures capable of supporting commercially relevant size wind turbines from ISO standard freight container-sized modules. Current demonstrator concepts in floating offshore wind require infrastructure of the scale unavailable or inaccessible in most of the world. Cost reductions

	needed to remove barriers to floating offshore adoption will come from development of methods not requiring large infrastructure and use of cost-effective mass manufacturing methods for making the construction modules. The proposed modular approach, with specially designed smaller and lighter building modules that could be towed out to sea for assembly, is significantly technically different from the current concepts and demonstrators. The concept will resolve the issue of prohibitively high cost of construction, logistics, and deployment in floating offshore wind. The main overarching research objective is to design the modules and the full structure, test those out as mathematical and physical models, carry out wave tank and sea conditions testing, and development of the manufacturing method. The project will deliver: design of a low-cost single module building block structure, design of a full modular configurable structure, creating physical and mathematical models, tank tests and sea test of physical models, analysis of manufacturing feasibility including a materials selection study and identification of coastal sites and new markets for adoption of the technology. Two of the University of Strathclyde engineering departments, AFRC and NAOME, will work together with the other members of the consortium. NAOME's role within the consortium is to develop a detailed hydrodynamic simulation model of the semi-submersible concept for two different types of floating modules - a passive one and a dynamic one which can have its buoyancy and orientation altered. Scaled models of the two module concepts under a range of different sea states representative of where the wind turbines will be deployed will be conducted. The results will be measured and analysed and a report provided to the lead partner on the findings from both tests and simulations. AFRC's role is to develop a finite element (FE) model for the initial and refined modules, to determine their suitability in terms of structural strength performance under different load cases. Once the best configuration for the module has been determined, the AFRC will develop a FE model for two different configurations of the final structural assembly made with the selected module and simulate the performance of the overall structures. A report will be provided, summarising the findings. Due to the complexity of the project, the geographical spread of the partners and the close collaborative nature of the project, AFRC will also support Frontier Technical in the management of the project
EPSRC Centre for Doctoral Training in Fuel Cells and	The CDT proposal 'Fuel Cells and their Fuels - Clean Power for the 21st Century' is a focused and structured programme to train >52 students within 9 years in basic principles of the subject and guide

their Fuels - Clean Power for the 21st Century	them in conducting their PhD theses. This initiative answers the need for developing the human resources well before the demand for trained and experienced engineering and scientific staff begins to strongly increase towards the end of this decade. Market introduction of fuel cell products is expected from 2015 and the requirement for effort in developing robust and cost effective products will grow in parallel with market entry. The consortium consists of the Universities of Birmingham (lead), Nottingham, Loughborough, Imperial College and University College of London. Ulster University is added as a partner in developing teaching modules. The six Centre directors and the 60+ supervisor group have an excellent background of scientific and teaching expertise and are well established in national and international projects and Fuel Cell, Hydrogen and other fuel processing research and development. The Centre programme consists of seven compulsory taught modules worth 70 credit points, covering the four basic introduction modules to Fuel Cell and Hydrogen technologies and one on Safety issues, plus two business-oriented modules which were designed according to suggestions from industry partners. Further - optional - modules worth 50 credits cover the more specialised aspects of Fuel Cell and fuel processing technologies, but also include socio-economic topics and further modules on business skills that are invaluable in preparing students for their careers in industry. The programme covers the following topics out of which the individual students will select their area of specialisation:- electrochemistry, modelling, catalysis;- materials and components for low temperature fuel cells (PEFC, 80 and 120 -130 degC), and for high temperature fuel cells (SOFC) operating at 500 to 800 degC;- design, components, optimisation and control for low and high temperature fuel cell systems; including direct use of hydrocarbons in fuel cells, fuel processing and handling of fuel impurities; integration of hydrogen systems including hybrid fuel-cell-battery and gas turbine systems; optimisation, control design and modelling; integration of renewable energies into energy systems using hydrogen as a stabilising vector;- hydrogen production from fossil fuels and carbon-neutral feedstock, biological processes, and by photochemistry; hydrogen storage, and purification; development of low and high temperature electrolysers;- analysis of degradation phenomena at various scales (nano-scale in functional layers up to systems level), including the development of accelerated testing procedures;- socio-economic and cross-cutting issues: public health, public acceptance, economics, market introduction; system studies on the benefits of FCH technologies to national and international energy supply.

	The training programme can build on the vast investments made by the participating universities in the past and facilitated by EPSRC, EU, industry and private funds. The laboratory infrastructure is up to date and fully enables the work of the student cohort. Industry funding is used to complement the EPSRC funding and add studentships on top of the envisaged 52 placements. The Centre will emphasise the importance of networking and exchange of information across the scientific and engineering field and thus interacts strongly with the EPSRC-SUPERGEN Hub in Fuel Cells and Hydrogen, thus integrating the other UK universities active in this research area, and also encourage exchanges with other European and international training initiatives. The modules will be accessible to professionals from the interacting industry in order to foster exchange of students with their peers in industry
Baseline Food Refrigeration Emissions in the UK, 2019-2020	This work benchmarks the existing UK cold chain and provides robust evidence-based data on emissions in 2020. Only emissions from refrigeration within UK borders was considered, both from refrigerant leakage and from electrical power usage. Energy consumption for energy consumption the Digest of UK Energy Statistics (DUKES) was widely used. This data is compiled by the Department for Business, Energy & Industrial Strategy (BEIS) and contains data for many years up until the current year. The United Kingdom Statistics Authority has designated these statistics as National Statistics, in accordance with the Statistics and Registration Service Act 2007 and therefore they were considered as the most accurate data available. DUKES data does not always differentiate the energy consumed by refrigeration systems in each of the cold chain sectors and therefore further analysis and assumptions were often required. Energy consumption values shown were collated per year for the years 2019/2020, unless otherwise stated. Emissions from refrigerant leakage. The main GHG refrigerants are the fluorinated gases (f-gases); hydrofluorocarbons (HFCs) and hydrochlorofluorocarbons (HCFCs). According to Brown et al (Brown et al., 2021) emissions of f-gases can occur at various stages of the refrigeration equipment life-cycle: During manufacturing, During installation, Over the operational lifetime, At disposal. The most comprehensive source of information for direct emissions is the UK Greenhouse Gas Inventory. This contains national greenhouse gas emission estimates for the period 1990-2019 and is the United Kingdoms National Inventory Report (NIR) submitted to the United Nations Framework Convention on Climate Change (UNFCCC). It includes losses during manufacture/initial charging and at decommissioning as well as losses in use.
Perception Spillover From Fracking, 2022	Public opposition to new energy technology can harm the chances of successful deployment. Less is known about knock-on effects on the wider energy system, including whether such opposition impacts

	public perceptions of other technologies. Here we present a mixed-methods study into perception spillover, examining whether the controversy over fracking for oil and gas affects public attitudes to two novel low-carbon energy technologies: deep enhanced geothermal systems, and green hydrogen. We argue that perception spillover is multi-faceted, and we conceptualise and test spontaneous, prompted and primed forms, examining how and why particular types occur. Using a nationally-representative UK survey and two focus groups, we show that perception spillover from fracking could lead to widespread negative perceptions of deep geothermal energy, influencing the conditions which deep geothermal would be expected to meet. Conversely, a minority of participants expressed more positive perceptions of green hydrogen because they deemed it dissimilar to fracking. Flexible fund project under the UK Unconventional Hydrocarbons project. Aims to understand the impact of fracking on public perceptions of other energy technologies. This is a linked dataset held in the UKDS.

Table 2 Abstracts/descriptions used for classification purposes