

## Supplementary Material to the manuscript “Towards FAIR and open data for low carbon energy - current state and call for action”

*Authors: Valeria Jana Schwanitz (Western Norway University of Applied Sciences), August Wierling (Western Norway University of Applied Sciences), Maria Bałazińska (Central Mining Institute), Mariusz Kruczek (Central Mining Institute), Demet Suna (AIT Austrian Institute of Technology), Christopher Burger-Scheidlin (AIT Austrian Institute of Technology), Manfred Paier (AIT Austrian Institute of Technology), Mehmet Efe Biresellioglu (IUE Izmir University of Economics), Muhittin Hakan Demir (IUE Izmir University of Economics), Massimo Celino (ENEA, IT)*

### **Content:**

1. Proof of representativeness
2. FAIR/O assessment methods
  - 2.1 Comparing methods for FAIR assessments
  - 2.2 FAIR assessment tool by ARDC
  - 2.3 Machine-based assessment of FAIR criteria
3. Results
  - 3.1 Aggregated results human vs. machine assessments
  - 3.2 Results from manual assessments
  - 3.3 Results from machine-based assessments
4. References

## 1. Proof of representativeness

Data is omnipresent in the energy system, flowing between its agents supported by IT infrastructure (Fig. 1, main text). We assess the compliance of energy databases with FAIR and openness criteria. To test whether the choice of databases is representative of data flows in the energy system, we reflect Fig. 1 with the help of the Global Energy Assessment Report, GEA 2012, and established classification schemes for energy data. We do so by mapping tested databases onto key tasks for enabling the energy transition. We utilize GEA 2012 as a reference framework, because it is a comprehensive review of the energy system drafted by numerous energy researchers from a broad range of disciplines. Altogether, 334 authors and 179 reviewers (including 5 anonymous reviewers) compiled a total of 1884 pages. This report on the energy system from 2012 is aiming to “... examine: the major global challenges and their linkages to energy, the technologies and resources available for providing energy services, future energy systems that address the major challenges, the policies and other measures to realize sustainable energy futures.” Sustainable energy is defined by the three ‘aspirational goals’: “(1) Stabilizing global climate change to 2 degrees above preindustrial levels to be achieved in the 21st century. (2) Enhance energy security by diversification and resilience of energy supply. (3) Eliminating household and ambient air pollution and universal access to modern energy services by 2030.” Box 1 below lists the chapters of GEA 2012. They illustrate the multi-disciplinary approach for examining the state of the art of the energy system and the sustainability challenges that lie ahead. Testing chapter by chapter (and sections), whether aspects covered in GEA 2012 are included in the energy system as presented in Fig. 1 (main paper), we find that this reference framework is well mirrored.

**Box 1:** Chapters of Global Energy Assessment Report, GEA 2012.

|  |  |
|--|--|
| <a href="#">Chapter 1: Energy Primer</a>                   | <a href="#">Chapter 12: Fossil Energy</a>  |
| <a href="#">Chapter 2: Energy, Poverty and Development</a> | <a href="#">Chapter 17: Energy Pathways for Sustainable Development</a>          |
| <a href="#">Chapter 3: Energy and Environment</a>          | <a href="#">Chapter 18: Urban Energy Systems</a>                                 |
| <a href="#">Chapter 4: Energy and Health</a>               | <a href="#">Chapter 19: Energy Access for Development</a>                        |
| <a href="#">Chapter 5: Energy and Security</a>             | <a href="#">Chapter 20: Land and Water: Linkages to Bioenergy</a>                |
| <a href="#">Chapter 6: Energy and Economy</a>              | <a href="#">Chapter 21: Lifestyles, Well-Being and Energy</a>                    |
| <a href="#">Chapter 7: Energy Resources and Potentials</a> | <a href="#">Chapter 22: Policies for Energy System Transformations</a>           |
| <a href="#">Chapter 8: Energy End-Use: Industry</a>        | <a href="#">Chapter 23: Policies for Energy Access</a>                           |
| <a href="#">Chapter 9: Energy End-Use: Transport</a>       | <a href="#">Chapter 24: Policies for the Energy Technology Innovation System</a> |
| <a href="#">Chapter 10: Energy End-Use: Buildings</a>      | <a href="#">Chapter 25: Policies for Capacity Building</a>                       |
| <a href="#">Chapter 11: Renewable Energy</a>               |  |

Next, we propose a structure for the energy system with the help of ontological concepts (Gruber, 1983). In this way, we expand the review of the energy system representation towards classification schemes used for energy and scientific data. We include the Standard International Energy Product Classification by UNSTATS and IRES ([SIEC](#))<sup>1</sup>, the Global Change Master Directory Keywords ([GCMD 2020](#)), JEL classification, and the European Science Vocabulary ([EuroSciVoc 2020](#)). Fig. 1 shows the result for the ontology with the energy system at the centre together with its agents (humans, machines), related activities (system monitoring, control, and integration) and available technologies (energy technologies, renewable energy generation technologies, smart technologies, environmental technologies, and others). A more detailed version of the ontology is available from the repository. Table 1 complements the classification and review by linking these classification schemes to a total of 80 energy and other databases which we evaluate in the section below. The databases may belong to one (or several) of

<sup>1</sup> IRES Compiler Manual: [https://unstats.un.org/unsd/energystats/methodology/documents/ESCM\\_Whitecover\\_170323.pdf](https://unstats.un.org/unsd/energystats/methodology/documents/ESCM_Whitecover_170323.pdf)

concepts of the energy system ontology (Fig. 1, Tab. 1). Keywords K1-K21 help to further characterize them. See also the details provided database by database in the attached spreadsheet documenting machine assessments.

**Fig. 1: Ontological concept of the energy system.** See also the repository to this publication for files.



**Table 1: Mapping ontological concepts and 80 databases selected for assessments.** Keywords K1-K21 further characterize the concepts. These are K1: demand & supply, K2: system integration, K3: objectives, K4: constraints, K5: data security & privacy, K6: policy process, K7: reporting of transition processes, K8: price signals, K9: resource endowments, K10: generation statistics, K11: energy types, K12: trade, K13: grid status, K14: consumption data, K15: load profiles, K16: control of energy services, K17: system flows incl. storage, K18: signals for plant operation, K19: alerts, K20: modeling, K21: technical parameters & physical constants.

| Classification concepts & related keywords  | Databases   |
|---|---|
| <p><b>Energy system:</b><br/>Resources; primary, secondary and final energy; energy services.</p> <p><i>Related keywords:</i><br/>K1, K9, K10, K12, K13, K14, K15, K17.</p>   | <p><a href="#">OECD Energy Data</a> , <a href="#">IEA Data &amp; Statistics</a>, <a href="#">National energy accounting (Example: National Statistics Poland)</a>, <a href="#">CommONEnergy Data Mapper</a>, <a href="#">Satellite Application Facility on Climate Monitoring (CMSAE)</a>; <a href="#">Simulations of hourly power output from wind and solar PV farms (Renewables.Ninja)</a>; <a href="#">ENTSO-E transparency platform</a>, <a href="#">German Federal Grid Regulator Open Data for Electricity Markets (SMARD)</a>; <a href="#">CommONEnergy Data Mapper</a>, <a href="#">Policies to Enforce the Transition to nearly zero energy buildings in EU-27 (Project: Entranze)</a>, <a href="#">Global Buildings Performance Network (GBPN)</a>, <a href="#">The open source mapping and planning tool for cooling and heating (project: Hotmaps)</a>, <a href="#">Mesures d'Utilisation Rationnelle de l'Energie (Mure database)</a>, <a href="#">Scenarios of market transition to nearly Zero Energy Buildings (nZEB)</a>, <a href="#">EU Building Stock Monitoring (Project: EPISCOPE)</a>; <a href="#">EUROSTAT energy balances</a>, <a href="#">IRENA Statistics</a>; <a href="#">European Energy Efficient Building District Database: from data to information to knowledge (Project: ExcFED)</a>; <a href="#">Global Energy Institute - US Chamber of Commerce</a>; <a href="#">FAOSTAT - energy use</a>; <a href="#">Joint Research Center (IRC) data catalogue</a>; <a href="#">Heat Roadmap Europe (PETA)</a>; <a href="#">Smart Integration of Energy Storages in Local Multi Energy Systems for maximising the Share of Renewables in Europe's Energy Mix (SmILES)</a>; <a href="#">Norwegian Power System Data by STATNET</a>; <a href="#">The WORLD BANK Data Catalog</a>; <a href="#">zenodo Repository</a>; <a href="#">Data Portal of the province Castilla and Leon</a>; <a href="#">EUROSTAT</a>; <a href="#">United Nations Open SDG Data Hub</a></p> |
| <p><b>Constraints &amp; objectives:</b><br/>Economic constraints, technological constraints, societal constraints, natural constraints.</p> <p><i>Related keywords:</i><br/>K3, K4, K5, K6, K8.</p>   | <p><a href="#">energydata.info (SDG 7)</a>, <a href="#">The WORLD BANK Data Catalog</a>, <a href="#">EU Energy Poverty Observatory</a>, <a href="#">Carbon Dioxide Information Analysis Center (CDIAC)</a>, <a href="#">United Nations Open SDG Data Hub</a>, <a href="#">European Environment Agency Data</a>, <a href="#">EU Critical Raw Materials - 4th list of CRMs</a>, <a href="#">EUROSTAT electricity prices</a>, <a href="#">NORDPOOL day-ahead electricity prices</a>, <a href="#">NORDPOOL historical market data</a>, <a href="#">International Monetary Fund - Fossil fuel subsidy data</a>; <a href="#">Access to European Union law (EUR-Lex)</a>; <a href="#">OECD Database on Policy Instruments for the Environment (PINE)</a>; <a href="#">UNEP Material Flow Database</a>; <a href="#">Ireland's Open Data Portal (case: Wicklow Wind Energy Strategy Wicklow County Development Plan 2016 - 2022)</a>; <a href="https://unstats-undesa.opendata.arcgis.com/">https://unstats-undesa.opendata.arcgis.com/</a>; <a href="#">Life Cycle Inventory Database (ecoinvent)</a></p>   |
| <p><b>Agents:</b><br/>Energy consumers (incl. households, public services and commerce with ICT businesses etc.), energy producers and distributors, energy system regulators</p> <p><i>Related keywords:</i><br/>K6, K7, K16, K18, K19</p> | <p><a href="#">IEA policies database</a>, <a href="#">EMODnet - THE EUROPEAN MARINE OBSERVATION AND DATA NETWORK</a>, <a href="#">Mesures d'Utilisation Rationnelle de l'Energie (Mure database)</a>, <a href="#">Access to European Union law (EUR-Lex)</a>, <a href="#">European Commission eCERTIS database</a>, <a href="#">Covenant of Mayors for Climate &amp; Energy Europe</a>, <a href="#">OECD Database on Policy Instruments for the Environment (PINE)</a>; <a href="#">UK Bus Open Data Service</a>, <a href="#">U.S. National Renewable Energy Laboratory (NREL)</a>; <a href="#">ENTSO-E transparency platform</a>; <a href="#">European Environment Agency Data</a>; <a href="#">EUROSTAT electricity prices</a>; <a href="#">FAOSTAT - energy use</a>; <a href="#">International Monetary Fund - Fossil fuel subsidy data</a>; <a href="#">Norwegian Felles datakatalog</a>; <a href="#">Database for support schemes, grid issues and policies about RE sources in EU-28, EFTA etc. (RES LEGAL)</a>; <a href="#">Norwegian Power System Data by STATNET</a>; <a href="#">TOMTOM Traffic Index</a>; <a href="#">German Business Registry</a>; <a href="#">Data Portal of the province Castilla and Leon</a>; <a href="#">NORDPOOL day-ahead electricity prices</a>; <a href="#">German Federal Grid Regulator Open Data for Electricity Markets (SMARD)</a></p>  |
| <p><b>Technologies:</b><br/>Energy t., environmental t., smart technologies &amp; ICT</p> <p><i>Related keywords:</i> K11</p>   | <p><a href="#">IIASA Publications Repository (IIASA PURE)</a>, <a href="#">Outdoor CPV monitoring (Project: Greco@zenodo)</a>, <a href="#">zenodo Repository</a>; <a href="#">Life Cycle Inventory Database (ecoinvent)</a>; <a href="#">German Open Government Data Portal (GovData)</a>, <a href="#">EU SCIENCE HUB (dataset: PV-GIS)</a>; <a href="#">Project Sharewind@zenodo</a>, <a href="#">EMODnet - THE EUROPEAN MARINE OBSERVATION AND DATA NETWORK</a>, <a href="#">EMODnet - THE EUROPEAN MARINE OBSERVATION AND DATA NETWORK</a>, <a href="#">Database for biomass and waste (Phyllis)</a>; <a href="#">European Energy Efficient Building District Database: from data to information to knowledge (Project: ExcFED)</a>; <a href="#">Global Buildings Performance Network (GBPN)</a>; <a href="#">The open source mapping and planning tool for cooling and heating (project: Hotmaps)</a>; <a href="#">IRENA Statistics</a>; <a href="#">Simulations of</a></p>   |

|   |  |
|---|--|
|   | <a href="#">hourly power output from wind and solar PV farms (Renewables.Ninja)</a> ; <a href="#">Heat Roadmap Europe (PETA)</a> ; <a href="#">opendata.swiss Windenergieanlagen</a> ; <a href="#">Ireland's Open Data Portal (Wave energy test site)</a> ;  |
| <b>Systemic change:</b><br>Technological c., societal c., environmental c., disruptive events<br><i>Related keywords: K20</i> | <a href="#">IIASA Data Repository (IIASA DARE)</a> ; <a href="#">ENERGYDATA.INFO (SDG 7)</a> ; <a href="#">The International Disaster Database (EM-DAT)</a> , <a href="#">Life Cycle Inventory Database (ecoinvent)</a> ; <a href="#">NORDPOOL historical market data</a> ; <a href="#">Statistics on power blackouts (Scientific publication)</a>   |
| <b>Inter-linkages with other systems</b><br>Economy, Environment, Society, Health<br><i>Related keywords: K2, K12</i>         | <a href="#">FAOSTAT (land-use)</a> , <a href="#">FAOSTAT (energy)</a> , <a href="#">FAOSTAT (water)</a> , <a href="#">WHO (e.g., air pollution)</a> , <a href="#">Satellite Application Facility on Climate Monitoring (CMSAF)</a> , <a href="#">FMODnet - THE EUROPEAN MARINE OBSERVATION AND DATA NETWORK</a> , <a href="#">The World Bank Data Catalog</a> , <a href="#">Copernicus Space Component Data Access system (CSCDA)</a> ; <a href="#">Carbon Dioxide Information Analysis Center (CDIAC)</a> ; <a href="#">International Monetary Fund (IMF) Data</a> ; <a href="#">IIASA Data Repository (IIASA DARE)</a> ; <a href="#">Nature Scientific Data (Scientific Publication)</a>   |
| <b>Data hubs</b> not associated with single keywords, <b>covering large ground in the energy domain</b>                       | <a href="#">DataCite - Research data repository</a> (>76687 data sets on energy), <a href="#">zenodo Repository</a> (>2226), <a href="#">OECD data</a> (>1028), <a href="#">Joint Research Center (JRC) data catalogue</a> (>1162), <a href="#">U.S. National Renewable Energy Laboratory (NREL)</a> (100), <a href="#">IEA Data &amp; Statistics</a> (unspecified but large number), <a href="#">The WORLD BANK Data Catalog</a> (>1993), <a href="#">Hub - energydata.infoD</a> (558), <a href="#">EUROSTAT</a> (4921); <a href="#">Kaggle repository</a> (>9000); <a href="#">Norwegian Felles datakatalog</a> (>28), <a href="#">Ireland's Open Data Portal</a> (>153); <a href="#">German Open Government Data Portal (GovData)</a> (>4269), <a href="#">United Nations Open SDG Data Hub</a> (tracking of SDG 7 indicators across countries) |

## 2. FAIR/O assessment methods

Different approaches to assess the compliance with the FAIR principles of Wilkinson et al. 2016 exist, but few of them are machine-actionable. Most of them are tools for self-assessment of databases by humans to support improving the FAIRification of data. These tools come in the form of questionnaires and checklists. A comparison of them is done in RDA (2020). To contrast human-performed and machine-actionable tests, we selected the ARDC DAIR data self-assessment tool (ARDC 2019), and compare it with the FAIR Maturity Evaluation Service (Wilkinson et al. 2019). Our choice is motivated by the following reasons. In regard to machine-actionable tests, Wilkinson et al. 2019 provides the only comprehensive test at this point. In regard to human-performed tests, we reviewed the different approaches listed in RDA (2020), weighing pro's and con's (Table 2). The reason to select the ARDC FAIR data assessment tool is that it aligns with the FAIR principles, has a good balance between technical and non-technical questions, and enables scoring. In this way, we can directly compare the two assessment perspectives.

To scrutinize our decision, we additionally tested the second-best candidate (SATISFYD) for two databases (Table 3). While both tools indicate ratings for FAIR categories, the questions and answers differ. Hence, the comparison of scores is limited. However, a notable difference in SATISFYD concerns the impact of single factors such as the existence of a license or whether metadata are available without the data itself. This information has an immense impact on the scoring. Furthermore, other factors are underrepresented in comparison to the ARDC tool. Overall, questions are posed in a more general way. At the same time, questions regarding controlled vocabulary, taxonomies and ontologies are very technical and difficult to answer for non-experts. The comparison of these two tools confirms that the ARDC assessment tool aligns better to the FAIR principles, meeting the intention of this study.

**Table 2: Brief comparison of FAIR assessment approaches and tools**

| Method/Tool   | Comment   |
|---|---|
| <b>ARDC:</b><br><a href="#">ANDS-NECTAR-RDS-FAIR data assessment tool</a>   | Online assessment tool, developed by the Australian Research Data Commons (ARDC). Designed for manual assessment by the user, <b>following all FAIR data principles of Wilkinson et al., 2016</b> . 12 technical and non-technical questions <sup>2</sup> . Users choose the answers from a drop-down menu. Contains short explanations of FAIR principles and terms. Results from the test are indicated with a progress bar as a new answer is given. Scores are not made explicit. However, one can infer them from the source code. The authors note that the test only serves as an orientation.   |
| <b>SATIFYD:</b> the <a href="#">DANS Self-Assessment Tool</a>   | Online assessment tool, developed by Data Archiving and Networked Services (DANS) and designed for manual assessment by the user. <b>Does not follow the FAIR data principles</b> , only touching upon them. Consists of different questions, e.g., regarding the use of standards to describe data (controlled vocabularies, taxonomies, ontologies). Consists of 12 non-technical questions, thus, easier to understand and answer. Contains short explanations of assessment principles, questions and terms. Indicates rating and scoring by percentages.   |
| <b>Fair enough?</b><br><a href="#">checklist</a>  | Online checklist developed by Data Archiving and Networked Services (DANS). Designed for manual assessment of FAIRness of data(sets) and the trustworthiness of the selected repository. <b>Does not follow the FAIR data principles</b> , only touching upon them. Consists of 11 non-technical questions, thus, easier to understand and answer. Contains explanations of terms/questions for corresponding paragraphs. Does not provide rating for FAIR and scoring of overall FAIRness.   |
| <b>OzNome: The CSIRO 5-star Data Rating</b> <a href="#">tool</a>  | Online assessment tool, developed by Commonwealth Scientific and Industrial Research Organisation, Australia. Designed for manual assessment of FAIRness of data(sets). <b>Follows the FAIR data principles</b> . Additionally includes questions on the openness of data (e.g., groups of users with full or limited access to the actual data). Consists of 14 technical and non-technical questions. Contains short explanations of terms/questions. Displays a chart using a 5-star data rating.  |
| <b>Stewardship Maturity Mix</b><br><a href="#">“Scientific Data Stewardship Maturity Assessment Model Template”</a><br><a href="#">(template)</a> | Online self assessment template developed by North Carolina Institute for Climate Studies (CICS-NC), the National Centers for Environmental Information (NCEI), and domain experts. Designed for measuring data stewardship practices and leveraging community best practices and standards. <b>Does not follow the FAIR data principles</b> . Key components are: preservability, accessibility, usability, production sustainability, data quality assurance, data quality control & monitoring, data quality assessment, transparency & traceability, and data integrity. Technical and non-technical checks of the components. Contains short explanations of terms/questions. Indicates a self assessment of the maturity level using a 5-level grading scheme from 1 (Ad hoc, not managed) to 5 (optimal/well managed). Does not include scoring relating to FAIR criteria. |
| <b>Data Stewardship Wizard</b> - <a href="#">questionnaire</a>  | Online questionnaire to create smart Data Management Plans (DMP) for FAIR Open Science tool, developed by Data Stewardship Wizard (DSW) in cooperation with six research organizations. Designed for DMP development for projects. <b>Does not follow all FAIR data principles</b> . Consists of 189 technical and non-technical questions, pertaining to experiment design, data design and planning, capture/ measurement, processing and curation, integration, interpretation, information and insights. Does not contain short explanations of terms/questions. Percentage scoring relating to FAIR criteria, “Good DMP Practice” and “Openness”.  |
| <a href="#">Checklist for Evaluation of Dataset Fitness for Use</a>   | Online checklist developed in cooperation with ICSU World Data System and Research Data Alliance. Designed for manual estimation of FAIRness of data(sets) and data curation. Dataset to be evaluated should be stewarded within a CoreTrustSeal-certified repository. <b>Does not follow strictly FAIR data principles</b> , only touching upon them. Consists of 20 non-technical questions, but it is not easy to follow the documentations and checklist. Some criteria are marked with asterisks denoting that assessing will require domain/discipline specific knowledge. Does not contain short explanations of terms/questions. Does not indicate ratings or scoring.  |
| <b>RDA-SHARC Evaluation,</b><br><a href="#">presentation</a>  | Poster presentation created by Research Data Alliance (RDA) and SHaring Rewards & Credits (SHARC). Designed for manual assessment by the user as a decision tree in each FAIR Principle. <b>Follows the FAIR data principles and Open Science Career Assessment Matrix</b> designed by EC Working Group on Reward under Open Science. Details on questions in the poster are limited. Unclear if the tool is ready for use. Details are limited.  |

<sup>2</sup> Technical questions and answers are defined as understandable for the researchers who have knowledge about data science and IT. Non-technical questions and answers are defined as understandable also for the researchers without data science and IT knowledge.

|  |   |
|--|---|
| <b>WMO-Wide Stewardship Maturity Matrix for Climate Data, <a href="#">tool</a></b> | Online assessment tool developed by WMO Stewardship Maturity Matrix for Climate Data (SMM-CD) Working Group. Designed for manual assessment by the user. <b>Follows “internationally-validated data stewardship best practices”</b> . Consists of 11 questions. Contains explanations/notes below each question. Indicates rating by maturity level 1 to 5 for the different questions, no overall score.   |
| <b>Data Use and Services Maturity Matrix, <a href="#">tool</a></b>                 | Online assessment tool developed by MM-Serv Working Group adopting the approach of the NCEI (National Centers for Environmental Information)/CICS-NC Data Stewardship Maturity Matrix (DSMM). Designed for the ESIP-DSC (Data Stewardship Committee)-wide review of the MM-Serv to ensure and improve its quality. <b>Follows MM-Serv Key Components and Maturity Level Criteria</b> . Consists of 9 Key Components, each being associated with maturity levels 1 to 5. Contains explanations of terms for corresponding maturity levels. Indicates rating based on maturity levels, no overall score |

**Table 3: FAIR data tests comparing the tools SATIFYD and ARDC Scorings**

| Database   | DANS  | ARDC   |
|--|---|--|
| Zenodo Repository (Data set 'Greco')                       | F: 89%, <b>A: 50%</b> , I: 100%, R: 87%<br>Total FAIR: 81%<br>O: 100% | F: 94%, <b>A: 90%</b> , I: 88%, R: 100%,<br>Total FAIR: 93%<br>O: 100% |
| EMODnet - THE EUROPEAN MARINE OBSERVATION AND DATA NETWORK | F: 44%, <b>A: 100%</b> , I: 75%, R: 74%<br>Total FAIR: 73%<br>O: 67%  | F: 53%, <b>A: 40%</b> , I: 50%, R: 71%,<br>Total FAIR: 54%<br>O: 67%   |

## 2.1 Comparing FAIR/O assessments

We apply and compare two assessment methods. One is the FAIR self-assessment tool developed by the Australian Research Data Commons ([Link](#)), which is carried out by humans. The other is the FAIR Metrics software developed by Wilkinson et al. 2019, ([Link](#)). The table below maps the criteria vis-a-vis with the original assessment criteria by Wilkinson et al. 2016. For comparing the methods quantitatively, we use the scoring developed by ARDC. We map this score onto the FAIR Metrics questions.

**Table 4: Comparison of scores assigned for manual assessments and machine assessments**

| FAIR criteria used for manual assessment based on ARDC <sup>3</sup> with associated points  | Score     | Maturity indicators used for machine assessment based on FAIR Metrics - associated scores <sup>4</sup>  | Score                           |
|---|-----------|---|---------------------------------|
| <b>F1: (Meta)data are assigned a globally unique persistent identifier</b>  |           |   |                                 |
| Does the dataset have any identifiers assigned?<br>A) Globally unique, citable and persistent (e.g. DOI, PURL, ARK or Handle) - 8 points<br>B) Web address (URL) - 3 points<br>C) Local identifier - 1 point<br>D) No identifier - 0 points<br><br><i>Note: No separation between data and metadata - likely that users only test the data and not the metadata</i> | up to 47% | <a href="#">FAIR Metrics Gen2- Unique Identifier</a> (F1): Metric to test if the metadata resource has a unique identifier. This is done by comparing the GUID to the patterns (by regexp) of known GUID schemas such as URLs and DOIs. Known schema are registered in FAIRSharing ( <a href="https://fairsharing.org/standards?q=&amp;selected_facets=type_exact:identifier%20schema">https://fairsharing.org/standards?q=&amp;selected_facets=type_exact:identifier%20schema</a> ). 25% of 8/17*100%=2/17*100%<br><a href="#">FAIR Metrics Gen2 - Identifier Persistence</a> (F1): Metric to test if the unique identifier of the metadata resource is likely to be persistent. Known schema are registered in FAIRSharing ( <a href="https://fairsharing.org/standards?q=&amp;selected_facets=type_exact:identifier%20schema">https://fairsharing.org/standards?q=&amp;selected_facets=type_exact:identifier%20schema</a> ). For URLs that don't follow a schema in FAIRSharing we test known URL persistence schemas (purl, oclc, fdlp, purlz, w3id, ark). 25% of 8/17*100%=2/17*100% | Scaled to meet up to 8/17 *100% |

<sup>3</sup> Note that in ARDC F2 and F3 are interchanged when comparing to the original order of the criteria.

<sup>4</sup> We assume that data and metadata have the same weight.



|   |                        |   |   |
|---|------------------------|---|---|
|   |                        | <a href="#">FAIR Metrics Gen2 - Data Identifier Persistence</a> (F1): Metric to test if the unique identifier of the data resource is likely to be persistent. Known schema are registered in FAIRSharing ( <a href="https://fairsharing.org/standards/?q=&amp;selected_facets=type_exact:identifier%20schema">https://fairsharing.org/standards/?q=&amp;selected_facets=type_exact:identifier%20schema</a> ). For URLs that don't follow a schema in FAIRSharing we test known URL persistence schemas (purl, oclc, fdlp, purlz, w3id, ark). 50% of $8/17*100\%=4/17*100\%$  |   |
| <b>F2: Data are described with rich metadata defined by R1.</b>   |                        |   |   |
| How is the data described with metadata?<br>A) Comprehensively (see suggestion) using a recognized formal machine-readable metadata schema? - 4 p.<br>B) Comprehensively, but in a text-based, non-standard format - 3 points<br>C) Brief title and description - 2 p.<br>D) The data is not described - 0 p.   | Not alone, may add 22% | <a href="#">FAIR Metrics Gen2 - Structured Metadata</a> (F2): Tests whether a machine is able to find structured metadata. This could be (for example) RDFa, embedded json, json-ld, or content-negotiated structured metadata such as RDF Turtle 50% of $4/17*100\%=2/17*100\%$<br><a href="#">FAIR Metrics Gen2 - Grounded Metadata</a> (F2): Tests whether a machine is able to find 'grounded' metadata. i.e. metadata terms that are in a resolvable namespace, where resolution leads to a definition of the meaning of the term. Examples include JSON-LD, embedded schema, or any form of RDF. This test currently excludes XML, even when terms are namespaced. Future versions of this test may be more flexible. 50% of $4/17*100\%=2/17*100\%$      | Scaled to meet $4/17*100\%$                   |
| <b>F3: Metadata are clearly and explicitly include the identifier of the data described.</b>  |                        |   |   |
| 2) Is the dataset included in all metadata records/files describing the data?<br>A) Yes - 1 point<br>B) No - 0 points<br><i>Note: Questions related to F2 and F3 are mixed in order when comparing Wilkinson and ARDC.</i><br><i>Note: If F1 is answered with 'no' then this question is by default zero.</i>   | May add up to 5%       | <a href="#">FAIR Metrics Gen2 - Data Identifier Explicitly In Metadata</a> (F3): Metric to test if the metadata contains the unique identifier to the data. This is done by searching for a variety of properties, including foaf:primaryTopic, schema:mainEntity, schema:distribution, sio:is-about, and iao:is-about. schema codeRepository is used for software releases. 50% of $1/17*100\%=1/34*100\%$<br><a href="#">FAIR Metrics Gen2- Metadata Identifier Explicitly In Metadata</a> (F3): Metric to test if the metadata contains the unique identifier to the metadata itself. This is done using a variety of 'scraping' tools, including DOI metadata resolution, the use of the 'extract' Python tool, and others...50% of $1/17*100\%=1/34*100\%$ | Scaled to meet potential plus of $1/17*100\%$ |
| <b>F4: (Meta)data are registered and indexed in a search-able resource</b>  |                        |   |   |
| What type of repository or registry is the metadata record in?<br>A) Data is in one place but discoverable through several registries - 4 points<br>B) Generalist public repository - 2 points<br>C) Domain-specific repository - 2 points<br>D) Local institutional repository - 2 points<br>E) The data is not described in any repository - 0 points | May add up to 23%      | <a href="#">FAIR Metrics Gen2 - Searchable in major search engine</a> (F4): Tests whether a machine is able to discover the resource by search, using Microsoft Bing  | Up to $4/17*100\%$                            |
| <b>A1: Metadata are retrievable by their identifier using a standardized communications protocol.</b>   |                        |   |   |
| <i>A1.1: The protocol is open, free, and universally implementable</i>  |                        |   |   |
| <i>A1.2: The protocol allows for an authentication and authorization procedure where necessary.</i>   |                        |   |   |
| <u>How accessible is the data?</u><br>A) Publicly accessible - 5 points<br>B) Fully accessible to persons who meet explicitly stated conditions,  | A1=<br>(Q1+Q2)/<br>90% | Tested through A1.1 and A1.2.   | $A1=(Q1+Q2)/90\%$                             |

|  |  |  |                                     |
|--|--|--|-------------------------------------|
| <p>e.g., ethics approval for sensitive data - 5 points<br/> C) A de-identified/modified subset of the data is publicly accessible - 4 p.<br/> D) Embargoed access after a specified date - 3 points<br/> E) Unspecified conditional access, e.g., contact the data custodian for access - 2 points<br/> F) Access to metadata only - 1 point<br/> G) No access to data or metadata - 0 points</p> <p><u>Is the data available online without requiring specialized protocols or tools once access has been approved?</u></p> <p>A) Standard web service API - 4 points<br/> B) Non-standard web service - 3 points<br/> C) File download from online location - 2 points<br/> D) By individual arrangement - 1 points<br/> E) No access to data - 0 points</p> | <p>Q1: 51%<br/> Q2:<br/> Up to 39%</p> | <p><a href="#">FAIR Metrics Gen2 - Uses open free protocol for data retrieval</a> (A1.1): Data may be retrieved by an open and free protocol. Tests data GUID for its resolution protocol. Currently passes InChI Keys, DOIs, Handles, and URLs. Recognition of other identifiers will be added upon request by the community. 22.5%<br/> <a href="#">FAIR Metrics Gen2 - Uses open free protocol for metadata retrieval</a> (A1.1): Metadata may be retrieved by an open and free protocol. Tests metadata GUID for its resolution protocol. Currently passes InChI Keys, DOIs, Handles, and URLs. Recognition of other identifiers will be added upon request by the community. 22.5%<br/> <a href="#">FAIR Metrics Gen2 - Data authentication and authorization</a> (A1.2): Test a discovered data GUID for the ability to implement authentication and authorization in its resolution protocol. Currently passes InChI Keys, DOIs, Handles, and URLs. It also searches the metadata for the Dublin Core 'accessRights' property, which may point to a document describing the data access process. Recognition of other identifiers will be added upon request by the community. 22.5%<br/> <a href="#">FAIR Metrics Gen2 - Metadata authentication and authorization</a> (A1.2): Tests metadata GUID for the ability to implement authentication and authorization in its resolution protocol. Currently passes InChI Keys, DOIs, Handles, and URLs. Recognition of other identifiers will be added upon request by the community. 22.5%</p> | <p>45% for A1</p> <p>45% for A2</p> |
| <p><b>A2: Metadata are accessible, even when the data are no longer available.</b></p>   |  |  |                                     |
| <p>Will the metadata record be available even if the data is no longer available?<br/> A) Yes - 1 point<br/> B) No - 0 points<br/> C) Unsure - 0 points</p>  | <p>Up to 9%</p>                        | <p><a href="#">FAIR Metrics Gen2 - Metadata Persistence</a> (A2): Metric to test if the metadata contains a persistence policy, explicitly identified by a persistencePolicy key (in hashed data) or a <a href="http://www.w3.org/2000/10/swap/pim/doc#persistencePolicy">http://www.w3.org/2000/10/swap/pim/doc#persistencePolicy</a> predicate in Linked Data. 1/10*100%</p>   | <p>10%</p>                          |
| <p><b>I1: (Meta-)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.</b></p>  |  |  |                                     |
| <p>What (file) formats is the data available in?<br/> A) In a structured, open-standard, machine-readable format - 2 points<br/> B) In a structured, open-standard, non-machine-readable format - 1 point<br/> C) Mostly in a proprietary format - 0 points</p>  | <p>Up to 25%</p>                       | <p><a href="#">FAIR Metrics Gen2 - Metadata Knowledge Representation Language (weak)</a> (I1): Maturity Indicator to test if the metadata uses a formal language broadly applicable for knowledge representation. Broad view of what defines a 'knowledge representation language'; in this evaluation, anything that can be represented as structured data will be accepted. 1/16*100%=6.25%<br/> <a href="#">FAIR Metrics Gen2 - Metadata Knowledge Representation Language (strong)</a> (I1): Maturity Indicator to test if the metadata uses a formal language broadly applicable for knowledge representation. Broad view of what defines a 'knowledge representation language'; in this evaluation, a knowledge representation language is interpreted as one in which terms are semantically-grounded in ontologies. Any form of RDF will pass this test (including RDF that is automatically extracted by third-party parsers such as Apache Tika). 1/16*100%=6.25%<br/> <a href="#">FAIR Metrics Gen2 - Data Knowledge Representation Language (weak)</a> (I1): Maturity Indicator to test if the data uses a formal language broadly applicable for knowledge representation. Test takes a broad view of what defines a</p>  | <p>25%</p>                          |

|  |                |  |   |
|--|----------------|--|---|
|  |                | 'knowledge representation language'; in this evaluation, a knowledge representation language is interpreted as one in which terms are semantically-grounded in ontologies. Any form of structured data will pass this test. $1/16*100\%=6.25\%$<br><a href="#">FAIR Metrics Gen2 - Data Knowledge Representation Language (strong)</a> (I1): Maturity Indicator to test if the data uses a formal language broadly applicable for knowledge representation. This particular test takes a broad view of what defines a 'knowledge representation language'; in this evaluation, a knowledge representation language is interpreted as one in which terms are semantically-grounded in ontologies. Any form of ontologically-grounded linked data will pass this test. $1/16*100\%=6.25\%$ |   |
| <b>I2: (Meta)data use vocabularies that follow FAIR principles.</b>  |                |  |   |
| What best describes the types of vocabularies/ontologies/tagging schemas used to define the data elements?<br>A) Standardised, open and universal using resolvable global identifiers linking to explanations - 3 points<br>B) Standardised vocabularies/ ontologies/ schemas without global identifiers - 2 points<br>C) No standards have been applied in the description of data elements - 1 point<br>D) Data elements are not described - 0 points<br><i>Note: FAIR is not tested here!</i> | Up to 37%      | <a href="#">FAIR Metrics Gen2 - Metadata uses FAIR vocabularies (weak)</a> (I2): Maturity Indicator to test if the linked data metadata uses terms that resolve. This tests only if they resolve, not if they resolve to FAIR data, therefore is a somewhat weak test. $1/8*100\%=12.5\%$<br><a href="#">FAIR Metrics Gen2 - Metadata uses FAIR vocabularies (strong)</a> (I2): Maturity Indicator to test if the linked data metadata uses terms that resolve to linked (FAIR) data. Additionally $2/8*100\%=25\%$ .  | Either 0 or 12.5% or 37.5%              |
| <b>I3: (Meta)data include qualified references to other metadata.</b>  |                |  |   |
| How is the metadata linked to other data and metadata (to enhance context and clearly indicate relationships)?<br>A) Metadata is represented in a machine-readable format, e.g., in a linked data format such as resource-descriptive framework (RDF) - 3 points<br>B) The metadata record includes URI links to related metadata, data, and definitions - 2 points<br>C) There are no links to other metadata - 0 points  | Adds up to 38% | <a href="#">FAIR Metrics Gen2 - Metadata contains qualified outward references</a> (I3): Maturity Indicator to test if the metadata links outward to third-party resources. It only tests metadata that can be represented as Linked Data. $3/8*100\%=37.5\%$  | Either 0% or 37.5%                      |
| <b>R1: (Meta)data are richly described with a plurality of accurate and relevant attributes.</b><br><i>R1.1: (Meta)data are released with a clear and accessible data usage license.</i><br><i>R1.2: (Meta)data are associated with detailed provenance.</i><br><i>R1.3: (Meta)data meet domain-relevant community standards.</i>  |                |  |   |
| Which of the following best describes the license/usage rights attached to the data?<br>A) Standard machine-readable license, e.g., creative commons - 4 points<br>B) Standard text-based license - 3 points   | Up to 57%      | <a href="#">FAIR Metrics Gen2 - Metadata Includes License (strong)</a> (R1.1): Maturity Indicator to test if the linked data metadata contains an explicit pointer to the license. Tests: xhtml, dvia, dcterms, cc, data.gov.au, and Schema license predicates in linked data, and validates the value of those properties. $4/7*100\%=57\%$<br><a href="#">FAIR Metrics Gen2 - Metadata Includes License (weak)</a> (R1.1): Maturity Indicator to test if the metadata  | R1.1 + R1.2.<br>Either 0% or 43% or 57% |

|   |  |   |     |
|---|--|---|-----|
| C) Non-standard machine-readable license (clearly indicating under what conditions the data may be reused) - 3 points<br>D) Non-standard text-based license - 2 points<br>E) No license - 0 points  |  | contains an explicit pointer to the license. This 'weak' test will use a case-insensitive regular expression, and scan both key/value style metadata, as well as linked data metadata. Tests: xhtml, dvia, dcterms, cc, data.gov.au, and Schema license predicates in linked data, and validates the value of those properties. $3/7*100\%=43\%$<br><br><i>Note: Here it is tested whether the license is correct AND machine-readable as in A). Thus, it is much more than what is tested by ARDC (even in the weak case).</i> |     |
| How much provenance information has been captured to facilitate data re-use?<br>A) Fully recorded in a machine-readable format - 3 p.<br>B) Fully recorded in a text format - 2 points<br>C) Partially recorded - 1 points<br>D) No provenance information is recorded - 0 points   | Up to 43%  | -   | N/A |
| <b>O1: Openness - How open is the data?</b>   |  |   |     |
| A) Publicly accessible by humans and machines alike by standard protocol (open, no log in etc. required)<br>B) Publicly accessible but requires human intervention (e.g., may require login to download etc.)<br>C) Fully accessible to persons who meet explicitly stated conditions, (e.g., ethics approval for sensitive data)<br>D) Data are not open | A) 100%<br>B) $2/3*100\%$<br>C) $1/3*100\%$<br>D) 0% |   |     |

## 2.2 FAIR assessment tool by ARDC

The score of the ARDC assessment tool is based on points associated with the answers given. Twelve questions Q1-Q12 are posted. Answers are provided as options from a drop-down menu (between 2-7 options). Options are connected to points, and different answers may be similarly weighted. However, the highest points are connected with the first choice of an answer offered per question. The maximum scores that can be achieved for each FAIR criteria are F=17, A=10, I=8, R=7. The total score is calculated using the following formula:

- Formula:  $\sum_i \sum_j w_{ij} q_{ij} / \sum_i \max_j w_{ij}$  with  $i \in \{F,A,I,R\}$ , and  $w$  being the weights above and  $q$  being the answers to  $i$ =number of question and  $j$ =number of answer options.  $q$  takes either 0 (not being the answer) or 1 (being the answer).
- The following weights apply. They are taken from the points associated with answers:

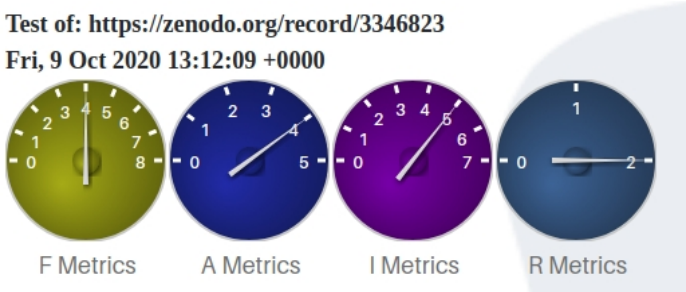
|                |                      |                |                   |
|----------------|----------------------|----------------|-------------------|
| Q1-F1: 0,1,3,8 | Q4-F4: 0,2,2,2,4     | Q7-A2: 0,0,1   | Q10-I3: 0,2,3     |
| Q2-F2: 0,2,3,4 | Q5-A1: 0,1,2,3,4,5,5 | Q8-I1: 0,1,2   | Q11-R1: 0,2,3,3,4 |
| Q3-F3: 0,1     | Q6-A1: 0,1,2,3,4     | Q9-I2: 0,1,2,3 | Q12-R1: 0,1,2,3   |

- Note an exception in Q1. If there is no identifier in the dataset, Q3 must also default to zero.
- The score is calculated from % scores of each FAIR component. All are weighted equally (25%).

### 2.3 Machine-based assessment of FAIR criteria

The machine-based assessments were developed by Wilkinson et al. 2019. They follow the FAIR principles in detail. Once installed, the software requires an input of an GUI ID to be tested along with the identifier of the person testing (e.g., ORCID). The tests take a few minutes and return two types of results, one being a summary picture (see Fig. 2) and the other being the returns from the software processing the GUI ID in form of a '0' (failure of test) or '1' (pass of test). A key for the software to return positive results is the existence of sufficient metadata details.

**Fig. 2** Screenshot from summary result of testing the instance from zenodo.



For comparing machine-based assessments of databases with manual assessments, we have to map the questions onto each other, using the same scoring and weighing of answers in each of the FAIR categories. The table below shows the assumptions in line with scores given in the ARDC-based manual assessment algorithm (Section 2.2) and their mapping onto the machine assessment (Table 5).

**Table 5:** Weights used to map results from machine assessments with manual assessments. The weights are used to calculate the scores in each of them.

| Test #          | Weight      | Answer    | Test #          | Weight | Answer    | Test #          | Weight | Answer    | Test #          | Weight | Answer    |
|-----------------|-------------|-----------|-----------------|--------|-----------|-----------------|--------|-----------|-----------------|--------|-----------|
|                 |             | (example) |                 |        | (example) |                 |        | (example) |                 |        | (example) |
| 1               | 0.12        | 0         | 1               | 0.225  | 0         | 1               | 0.0625 | 0         | 1               | 4/7    | 0         |
| 2               | 0.12        | 0         | 2               | 0.225  | 0         | 2               | 0.0625 | 0         | 2               | 3/7    | 0         |
| 3               | 0.24        | 1         | 3               | 0.225  | 1         | 3               | 0.0625 | 1         |                 |        |           |
| 4               | 0.12        | 1         | 4               | 0.225  | 0         | 4               | 0.0625 | 1         |                 |        |           |
| 5               | 0.12        | 1         | 5               | 0.1    | 1         | 5               | 0.125  | 1         |                 |        |           |
| 6               | 0.03        | 1         |                 |        |           | 6               | 0.25   | 1         |                 |        |           |
| 7               | 0.03        | 0         |                 |        |           | 7               | 0.375  | 1         |                 |        |           |
| 8               | 0.24        | 1         |                 |        |           |                 |        |           |                 |        |           |
| <b>F score:</b> | 0.74        |           | <b>A score:</b> | 0.33   |           | <b>I score:</b> | 0.88   |           | <b>R score:</b> | 0      |           |
| <b>Total:</b>   | <b>0.48</b> |           |                 |        |           |                 |        |           |                 |        |           |

### 3. Results

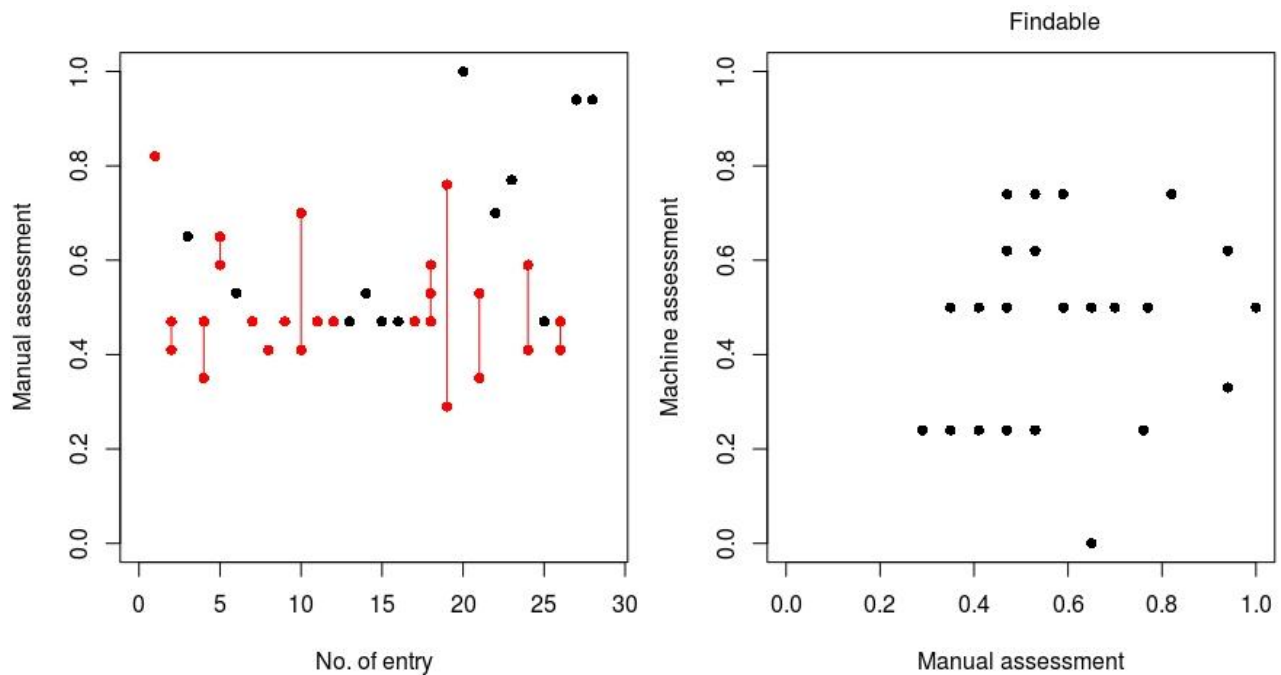
#### 3.1 Aggregated summary for manual vs. machine assessments

| Name of database  | Manual assessment   | Machine assessment                                  |
|---|---|---|
| Satellite Application Facility on Climate Monitoring (CMSAF)  | F: 82%, A: 80%, I: 63%, R: 29%<br>Total FAIR: 64%; O: 67%   | F: 74%, A: 33%, I: 88%, R: 0%<br>Total FAIR: 48%    |
| CommONEnergy Data Mapper  | F: 41%, A: 70%, I: 38%, R: 43%<br>Total FAIR: 48%; O: 67%   | F: 50%, A: 33%, I: 13%, R: 0%<br>Total FAIR: 24%    |
| Copernicus Space Component Data Access system (CSCDA)   | F: 65%, A: 90%, I: 88%, R: 86%<br>Total FAIR: 82%; O: 100%  | F: 0%, A: 0%, I: 0%, R: 0%<br>Total FAIR: 0%        |
| Covenant of Mayors  | F:47%, A: 70%, I: 50%, R: 29%<br>Total FAIR: 49%; O: 67%    | F: 50%, A: 33%, I: 88%, R: 0%<br>Total FAIR: 43%    |
| DataCite  | F: 59% , A: 90%, I: 75% , R: 57%, Total FAIR: 70%; O: 33%   | F: 50%, A: 33%, I: 13%, R: 0%<br>Total FAIR: 24%    |
| EMODnet - THE EUROPEAN MARINE OBSERVATION AND DATA NETWORK  | F: 53%, A: 40%, I: 50%, R: 71%,<br>Total FAIR: 54%; O: 67%  | F: 24%, A:33%, I: 0%, R: 0%,<br>Total FAIR: 14%     |
| Policies to Enforce the Transition to nearly zero energy buildings in EU-27 (Project: Entranze)         | F: 47%, A: 70%, I: 38%, R: 43%<br>Total FAIR: 49.5%; O: 67% | F: 24%, A:33%, I: 0%, R: 0%,<br>Total FAIR: 14%     |
| ENTSO-E transparency platform   | F: 41%, A: 70%, I: 0%, R: 0%<br>Total FAIR: 28%; O: 67%     | F: 50%, A:33%, I: 25%, R: 0%,<br>Total FAIR: 27%    |
| EUR-Lex (Document: Directive 95/46/EC )   | F: 47%, A: 80%, I: 63%, R: 57%<br>total FAIR: 62%; O: 67%   | F: 50%, A:33%, I: 25%, R: 0%,<br>Total FAIR: 27%    |
| Global Buildings Performance Network (GBPN)   | F: 70%, A: 70%, I: 50%, R: 14%<br>Total FAIR: 51%; O: 67%   | F: 50%, A:33%, I: 25%, R: 0%,<br>Total FAIR: 27%    |
| The open source mapping and planning tool for cooling and heating (project: Hotmaps)                    | F: 47%, A: 70%, I: 50%, R: 86%<br>Total FAIR: 63%; O: 67%   | F: 24%, A:33%, I: 0%, R: 0%,<br>Total FAIR: 14%     |
| IEA Policy database   | F: 47%, A: 50%, I: 38%, R: 57%<br>Total FAIR: 48%; O: 100%  | F:50%, A: 33%, I: 13%, R:0%,<br>Total FAIR: 24 %    |
| IEA World Energy Statistics and Balances  | F: 47%, A:50%, I: 38%, R: 57%,<br>Total FAIR: 48%; O: 0%    | F:74%, A: 33%, I: 88%, R:0%,<br>Total FAIR: 48 %    |
| JRC data catalogue  | F: 53%, A: 70%, I: 63%, R: 57%<br>Total FAIR: 61%; O: 100%  | F:62%, A: 78%, I: 88%, R:0%,<br>Total FAIR: 57 %    |
| JRC data catalogue - data collection  | F: 47%, A:100%, I: 38%,R: 43%,<br>Total FAIR: 57%, O: 100%  | F:50%, A: 33%, I: 13%, R:0%,<br>Total FAIR: 24 %    |
| JRC data catalogue - data set   | F: 47%, A:80%, I: 25%, R: 43%,<br>Total FAIR: 49%, O:67%    | F:62%, A: 78%, I: 88%, R:0%,<br>Total FAIR: 57%     |
| Mesures d'Utilisation Rationnelle de l'Energie (Mure database)  | F: 47%, A: 70%, I: 63%, R: 29%<br>Total FAIR: 52%, O: 67%   | F:50%, A: 33%, I: 13%, R:0%,<br>Total FAIR: 24%     |
| Kaggle Repository (data set)  | F:41%, A: 80%, I: 25%, R: 43%<br>Total FAIR: 47%; O: 67%    | F: 74%, A: 78%, I: 88%, R: 0%<br>Total FAIR: 60%    |
| Simulations of hourly power output from wind and solar PV farms (Renewables.Ninja)                      | F: 76%, A: 70%, I: 75%, R: 71%<br>Total FAIR: 73%; O: 67%   | F: 24%, A: 33%, I: 0%, R: 0%<br>Total FAIR: 14%     |
| U.S. National Renewable Energy Laboratory (NREL) data catalog   | F: 100%, A: 90%, I: 50%,R: 29%<br>Total FAIR: 67%; O: 100%  | F: 62%, A: 33%, I: 88%, R: 0%<br>Total FAIR: 45%    |
| Scenarios of market transition to nearly Zero Energy Buildings (nZEB)                                   | F: 53%, A: 60%, I: 38%, R: 0%<br>Total FAIR: 38%; O: 67%    | F: 24%, A: 33%, I: 0%, R: 0%<br>Total FAIR: 14%     |
| OECD hub (energy data)  | F: 70%, A: 90%, I: 75%,R: 100%<br>Total FAIR: 84%; O: 100%  | F: 50%, A: 33%, I: 13%, R:0%,<br>Total FAIR: 24%    |
| OECD hub (data set)   | F: 77% , A: 70%, I: 38% ,R: 29%<br>Total FAIR: 53%; O: 67%  | F: 50% , A: 33%, I: 13% , R: 0%;<br>Total FAIR: 24% |
| Heat Roadmap Europe (PETA)  | F: 59%, A: 60%, I: 38%, R: 0%<br>Total FAIR: 39%; O: 67%    | F: 50%, A: 33%, I: 50%, R:0%,<br>Total FAIR: 33%    |
| Database for support schemes, grid issues and policies about RE sources in EU-28, EFTA etc. (RES LEGAL) | F: 47%, A: 80%, I: 38%, R: 29%<br>Total FAIR: 49%; O: 100%  | F: 24% , A: 33%, I: 0% , R: 0%<br>Total FAIR: 14%   |

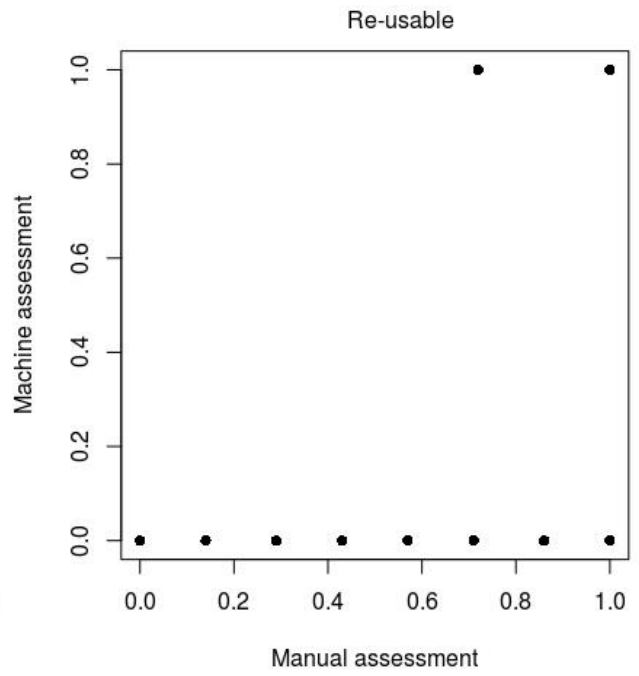
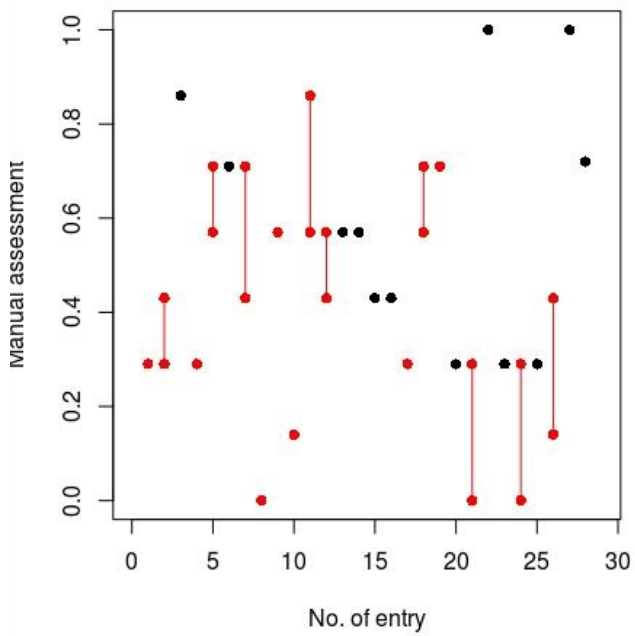
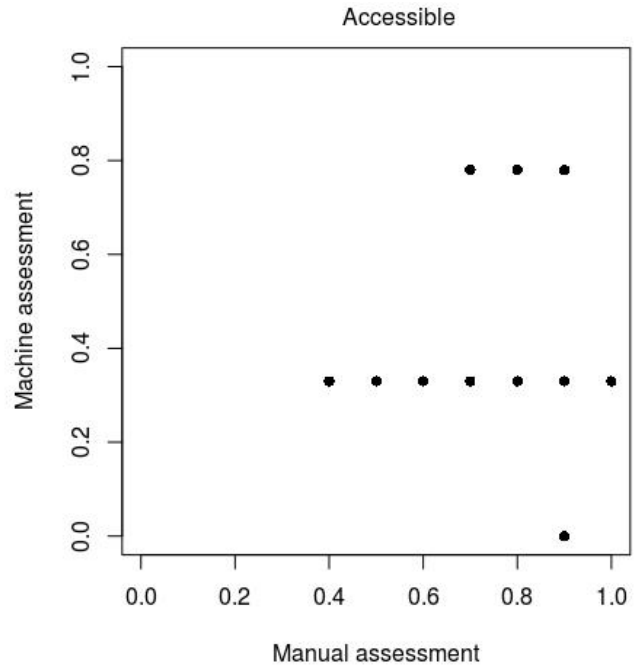
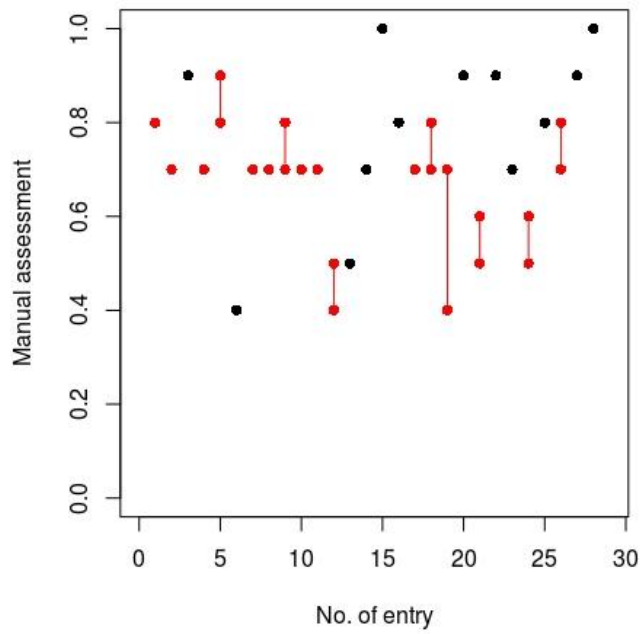
|  |  |   |
|--|--|---|
| EU Building Stock Monitoring (Project: EPISCOPE) | F: 47%, A: 70%, I: 25%, R: 43%<br>Total FAIR: 46%; O: 67%  | F: 24% , A: 33%, I: 0% , R: 0%<br>Total FAIR: 14% |
| Zenodo Repository (Data set 'Greco')             | F: 94%, A: 90%, I:88%,R: 100%,<br>Total FAIR: 93%; O: 100% | F:62%, A: 78%, I: 88%,<br>R:100%, Total FAIR: 82% |
| Zenodo Repository (Data set 'Sharewind')         | F: 100%, A: 80%, I: 63%, R:57%<br>Total FAIR: 75%; O: 100% | F: 38%,A: 33%,I: 88%, R:<br>100%; Total FAIR: 65% |

### 3.2 Results from individual manual assessments

The figures below contrast the assessment results before and after a joint briefing on how to interpret the FAIR assessment criteria. On the left hand side, black and red entries are shown, where black is the result for a single manual assessment and red shows the results for at least two assessments of the same database. If they are not connected by a line, assessments are on top of each other (agreement). The right hand side compares the results from manual assessments with those obtained through the machine-based assessment. Note that the results for interoperability are reported in the main document.









| Database                   | Assessor | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | F | A | I | R | I | T<br>o<br>t<br>a<br>l<br>C<br>o<br>m<br>m<br>e<br>n<br>t | Differences<br>to<br>assessment<br>before<br>briefing | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | F | A | I | R | I | T<br>o<br>t<br>a<br>l |   |   |   |   |
|----------------------------|----------|---|---|---|---|---|---|---|---|---|----|----|----|---|---|---|---|---|--|---|---|---|---|---|---|---|---|---|---|----|----|----|---|---|---|---|---|-----------------------|---|---|---|---|
| nZEB                       | CBS      | 3 | 3 | 0 | 0 | 5 | 0 | 0 | 0 | 2 | 2  | 0  | 0  | 3 | 5 | 5 | 2 | 4 | 3  | 5   | 5 | 2 | 4 | 5 | 2 | 0 | 2 | 5 | 1 | 0  | 1  | 2  | 0 | 0 | 0 | 0 | 3 | 0                     | 8 | 0 | 8 | 3 |
| OECD data                  | CBS      | 3 | 2 | 0 | 2 | 5 | 4 | 0 | 2 | 1 | 2  | 0  | 1  | 4 | 9 | 6 | 1 | 5 | 1  | 1   | 0 | 3 | 4 | 2 | 3 | 4 | 1 | 4 | 5 | 4  | 0  | 2  | 2 | 2 | 4 | 3 | 0 | 0                     | 5 | 0 | 4 | 8 |
| OECD energy data           | CBS      | 5 | 2 | 0 | 2 | 5 | 4 | 0 | 2 | 1 | 2  | 0  | 2  | 5 | 9 | 6 | 2 | 5 | 3  | 0   | 3 | 9 | 9 | 8 | 3 | 0 | 2 | 5 | 2 | 0  | 2  | 1  | 0 | 2 | 0 | 7 | 0 | 8                     | 9 | 3 | 5 |   |
| PETA                       | CBS      | 3 | 2 | 0 | 2 | 5 | 0 | 0 | 0 | 1 | 0  | 0  | 2  | 4 | 5 | 1 | 2 | 3 | 1  | 0   | 3 | 9 | 3 | 5 | 3 | 0 | 2 | 5 | 1 | 0  | 1  | 2  | 0 | 0 | 0 | 9 | 0 | 8                     | 0 | 9 | 3 |   |
| MURE                       | MB       | 3 | 4 | 1 | 2 | 1 | 3 | 1 | 1 | 2 | 0  | 2  | 0  | 5 | 5 | 3 | 2 | 4 | 9  | 0   | 8 | 9 | 4 | 3 | 3 | 0 | 2 | 5 | 2 | 0  | 1  | 2  | 2 | 0 | 2 | 7 | 0 | 3                     | 9 | 2 | 5 |   |
| Nat. energy stats.         | MB       | 3 | 2 | 1 | 2 | 5 | 2 | 1 | 1 | 0 | 2  | 0  | 1  | 4 | 8 | 3 | 1 | 4 | 7  | 0   | 8 | 4 | 5 | - | - | - | - | - | - | -  | -  | -  | - | - | - | - | - | -                     | - | - | - | - |
| Kaggle - repository        | MB       | 3 | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 1 | 0  | 2  | 1  | 4 | 5 | 2 | 4 | 4 | 7  | 0   | 5 | 3 | 1 | - | - | - | - | - | - | -  | -  | -  | - | - | - | 7 | 0 | 5                     | 1 | 8 | 6 |   |
| Kaggle - repository        | MEB      | 3 | 3 | 1 | 2 | 5 | 2 | 0 | 2 | 2 | 2  | 3  | 2  | 5 | 7 | 7 | 7 | 6 | 3  | 0   | 5 | 1 | 7 | - | - | - | - | - | - | -  | -  | -  | - | - | - | - | - | -                     | - | - | - | - |
| Kaggle - data set          | MB       | 3 | 4 | 1 | 2 | 5 | 3 | 0 | 1 | 2 | 2  | 4  | 1  | 5 | 8 | 6 | 7 | 6 | 9  | 0   | 3 | 1 | 8 | - | - | - | - | - | - | -  | -  | -  | - | - | - | - | - | -                     | - | - | - | - |
| Kaggle - data set          | MEB      | 5 | 2 | 0 | 2 | 5 | 2 | - | 2 | 2 | 2  | 2  | 2  | 5 | 7 | 7 | 5 | 6 | 3  | 0   | 5 | 7 | 4 | - | - | - | - | - | - | -  | -  | -  | - | - | - | - | - | -                     | - | - | - | - |
| Ninja                      | MB       | 3 | 2 | 0 | 0 | 2 | 2 | 0 | 1 | 1 | 0  | 4  | 1  | 2 | 4 | 2 | 7 | 4 | 9  | 0   | 5 | 1 | 1 | 8 | 3 | 0 | 2 | 5 | 2 | 0  | 2  | 2  | 2 | 4 | 1 | 6 | 0 | 5                     | 1 | 3 | 7 |   |
| Ninja                      | MEB      | 8 | 2 | 1 | 2 | 5 | 2 | 0 | 2 | 2 | 2  | 3  | 2  | 7 | 7 | 7 | 7 | 7 | 6  | 0   | 5 | 1 | 3 | - | - | - | - | - | - | -  | -  | -  | - | - | - | - | - | -                     | - | - | - |   |
| SMILES                     | MC       | 1 | 0 | 0 | 2 | 5 | 2 | 1 | 0 | 1 | 0  | 0  | 0  | - | - | - | - | - | -  | -   | - | - | - | 3 | 4 | 0 | 2 | 5 | 2 | 0  | 2  | 2  | 0 | 0 | 1 | 3 | 0 | 0                     | 4 | 7 | 4 |   |
| Tabula                     | MC       | 1 | 3 | 1 | 2 | 5 | 2 | 1 | 0 | 1 | 0  | 0  | 1  | - | - | - | - | - | -  | -   | - | - | - | 3 | 3 | 0 | 2 | 5 | 2 | 0  | 0  | 2  | - | 2 | 1 | 7 | 0 | 5                     | 3 | 6 | 4 |   |
| Zenodo                     | MC       | 8 | 4 | 1 | 4 | 5 | 4 | 1 | 2 | 3 | 3  | 4  | 3  | - | - | - | - | - | -  | -   | - | - | - | - | - | - | - | - | - | -  | -  | -  | - | - | - | - | - | -                     | - | - |   |   |
| Zenodo - Dataset GRECO     | MC       | 8 | 4 | 1 | 4 | 5 | 4 | 1 | 2 | 2 | 3  | 4  | 3  | - | - | - | - | - | -  | -   | - | - | - | 8 | 4 | 0 | 4 | 5 | 4 | 0  | 2  | 2  | 3 | 4 | 3 | 4 | 0 | 8                     | 0 | 3 | 9 |   |
| Zenodo - Dataset SHAREWIND | MC       | 8 | 4 | 1 | 4 | 5 | 4 | 1 | 2 | 2 | 3  | 4  | 3  | - | - | - | - | - | -  | -   | - | - | - | - | - | - | - | - | - | -  | -  | -  | - | - | - | - | - | -                     | - | - |   |   |
| CMSAF                      | MP       | 8 | 3 | 1 | 2 | 5 | 2 | 1 | 1 | 2 | 2  | 0  | 2  | 8 | 8 | 6 | 2 | 6 | 2  | 0   | 3 | 9 | 4 | 8 | 3 | 1 | 2 | 5 | 2 | 1  | 1  | 2  | 2 | 0 | 2 | 2 | 0 | 3                     | 9 | 4 | 4 |   |

|                                   |     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|-----------------------------------|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|--|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| COMONEnergy                       | MP  | 3 | 3 | 0 | 2 | 5 | 2 | 0 | 1 | 2 | 2 | 0 | 2 | 4 | 7 | 6 | 2 | 5 |   |  |   |   |   |   |   |   |   |   |   |   | 4 | 7 | 3 | 4 | 4 |   |
| Copernicus – CAMS Climate Forcing | MP  | 3 | 4 | 1 | 4 | 5 | 3 | 0 | 2 | 3 | 3 | 3 | 3 | 1 | 7 | 8 | 0 | 8 | 8 |  |   |   |   |   |   |   |   |   |   |   |   | 6 | 9 | 8 | 8 | 8 |
| Covenant of Mayors                | MP  | 3 | 0 | 1 | 2 | 5 | 2 | 0 | 1 | 2 | 2 | 0 | 2 | 3 | 7 | 6 | 2 | 4 |   |  |   |   |   |   |   |   |   |   |   |   | 4 | 7 | 5 | 2 | 4 |   |
| Datacite – Repository             | MP  | 3 | 3 | 1 | 4 | 5 | 3 | 0 | 2 | 3 | 3 | 3 | 2 | 1 | 6 | 8 | 0 | 7 | 7 |  |   |   |   |   |   |   |   |   |   |   |   | 5 | 9 | 7 | 5 | 7 |
| MURE                              | MEB | 3 | 3 | 0 | 2 | 5 | 2 | 0 | 2 | 2 | 0 | 2 | 0 | 4 | 7 | 5 | 2 | 4 |   |  |   |   |   |   |   |   |   |   |   |   | 4 | 7 | 6 | 2 | 5 |   |
| National Energy Statistics        | MEB | 3 | 3 | 0 | 2 | 5 | 2 | 0 | 1 | 2 | 2 | 0 | 1 | 4 | 7 | 6 | 1 | 4 |   |  | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |   |
| IUR LEX                           | MHD | 3 | 3 | 0 | 2 | 5 | 2 | 0 | 1 | 2 | 2 | 4 | 0 | 4 | 7 | 6 | 5 | 5 |   |  |   |   |   |   |   |   |   |   |   |   | 4 | 8 | 6 | 5 | 6 |   |
| ENTSO-E                           | MHD | 3 | 2 | 0 | 2 | 5 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 7 |   |   | 2 |   |  |   |   |   |   |   |   |   |   |   |   | 4 | 7 |   |   | 2 |   |
| ENTRANZE                          | MHD | 3 | 3 | 0 | 2 | 5 | 2 | 0 | 1 | 2 | 0 | 2 | 1 | 4 | 7 | 3 | 7 | 5 |   |  |   |   |   |   |   |   |   |   |   |   | 4 | 7 | 3 | 4 | 5 |   |
| EMODnet - dataset                 | MHD | 3 | 4 | 1 | 2 | 5 | 2 | 0 | 2 | 2 | 2 | 3 | 2 | 5 | 7 | 7 | 7 | 6 |   |  | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |   |
| EMODnet - hub                     | MHD | 3 | 3 | 1 | 2 |   |   |   |   |   |   |   |   | 5 |   |   |   |   |   |  |   |   |   |   |   |   |   |   |   |   | 5 | 4 | 5 | 7 | 5 |   |
| EXCEED                            | MK  | 1 | 2 | 0 | 0 | 2 | 1 | 0 | 1 | 2 | 0 | 0 | 0 | 1 | 3 | 3 |   | 2 |   |  |   |   |   |   |   |   |   |   |   |   | 2 | 3 | 3 |   | 2 |   |
| GBPN                              | MK  | 3 | 2 | 0 | 2 | 5 | 2 | 0 | 1 | 1 | 2 | 0 | 1 | 4 | 7 | 5 | 1 | 4 |   |  |   |   |   |   |   |   |   |   |   |   | 7 | 7 | 5 | 1 | 5 |   |
| Global Energy Institute           | MK  | 3 | 3 | 1 | 2 | 4 | 2 | 1 | 1 | 1 | 2 | 0 | 1 | 5 | 7 | 5 | 1 | 5 |   |  | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |   |
| HOTMAPS                           | MK  | 3 | 3 | 0 | 2 | 5 | 2 | 0 | 2 | 2 | 0 | 3 | 1 | 4 | 7 | 5 | 5 | 6 |   |  |   |   |   |   |   |   |   |   |   |   | 4 | 7 | 5 | 8 | 6 |   |
| IEA - Policy database             | MK  | 3 | 3 | 0 | 2 | 2 | 2 | 0 | 1 | 1 | 0 | 2 | 1 | 4 | 4 | 2 | 4 | 3 |   |  |   |   |   |   |   |   |   |   |   |   | 4 | 5 | 3 | 5 | 4 |   |

The following tables give the results from manual assessments, including comments to motivate choices.

| Name of database                          | Answers to Q1-Q12 (ARDC)  | Score | Criteria |
|---|---|-------|----------|
| Satellite Application Facility on Climate | Q1: Globally unique, citable and persistent (e.g. DOI, PURL, ARK or Handle) | 8     |          |
|   | Q2: Comprehensively, but in a text-based, non-standard format               | 3     |          |

|  |   |                         |                |
|--|---|-------------------------|----------------|
| Monitoring (CMSAF)   | Q3: Yes   | 1                       |                |
|  | Q4: Domain-specific repository  | 2                       |                |
|  |   | 14/17= <b>82%</b>       | <b>F score</b> |
|  | Q5: Fully accessible to person who meet explicitly stated conditions, e.g. ethics approval for sensitive data | 5                       |                |
|  | Q6: File download from online location  | 2                       |                |
|  | Q7: Yes   | 1                       |                |
|  |   | 8/10= <b>80%</b>        | <b>A score</b> |
|  | Q8: In a structured, open standard, non-machine-readable format   | 1                       |                |
|  | Q9: Standardised vocabularies/ontologies/schema without global identifiers                                    | 2                       |                |
|  | Q10: The metadata record includes URI links to related metadata, data and definitions                         | 2                       |                |
|  |   | 5/8= <b>63%</b>         | <b>I score</b> |
|  | Q11: No license   | 0                       |                |
|  | Q12: Fully recorded in a text format  | 2                       |                |
|  |   | 2/7= <b>29%</b>         | <b>R score</b> |
|  | 82% x 0.25 + 80% x 0.25 + 63% x 0.25 + 29% x 0.25 = <b>64%</b>  | <b>FAIR total score</b> |                |
| O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | <b>67%</b>  | <b>Openness score</b>   |                |

| Name of database  | Answers to Q1-Q12 (ARDC)   | Score  | Criteria                |
|---|--|--|-------------------------|
| CommONErgy data Mapper  | Q1: Web address (URL)  | 3  |                         |
|   | Q2: Brief title and description  | 2  |                         |
|   | Q3: No   | 0  |                         |
|   | Q4: Generalist public repository   | 2  |                         |
|   |  | 7/17= <b>41%</b>   | <b>F score</b>          |
|   | Q5: Publicly accessible  | 5  |                         |
|   | Q6: File download from online location                                       | 2  |                         |
|   | Q7: Unsure   | 0  |                         |
|   |  | 7/10= <b>70%</b>   | <b>A score</b>          |
|   | Q8: In a structured, open standard, non-machine-readable format              | 1  |                         |
|   | Q9: Standardised vocabularies/ontologies/schema without global identifiers - | 2  |                         |
|   | Q10: There are no links to other metadata                                    |  |                         |
|   |  | 0  |                         |
|   |  | 3/8= <b>38%</b>  | <b>I score</b>          |
|   | Q11: Non-standard text-based license   | 2  |                         |
|   | Q12: Partially recorded  | 1  |                         |
|   |  | 3/7= <b>43%</b>  | <b>R score</b>          |
|   |  | 41% x 0.25 + 70% x 0.25 + 38% x 0.25 + 43% x 0.25 = <b>48%</b> | <b>FAIR total score</b> |
| B) Publicly accessible but requires human intervention (e.g., may require login to download etc.) | 2/3= <b>67%</b>  | <b>Openness Score</b>  |                         |

| Name of database  | Answers to Q1-Q12 (ARDC)   | Score  | Criteria                |
|---|--|--|-------------------------|
| Copernicus Space Component Data Access System (CSCDA)   | Q1: Web address (URL)  | 3  |                         |
|   | Q2: Comprehensively (see suggestion) using a recognized, formal machine-readable metadata schema                                     | 4  |                         |
|   | Q3: No   | 0  |                         |
|   | Q4: Data is in one place but discoverable through several registries   | 4  |                         |
|   |  | 11/17=65%  | <b>F score</b>          |
|   | Q5: Fully accessible to person who meet explicitly stated conditions, e.g. ethics approval for sensitive data                        | 5  |                         |
|   | Q6: Standard web service API (e.g. OGC   | 4  |                         |
|   | Q7: Unsure   | 0  |                         |
|   |  | 9/10=90%   | <b>A score</b>          |
|   | Q8: In a structured, open standard, machine-readable format  | 2  |                         |
|   | Q9: Standardized vocabularies/ontologies/schema without global identifiers   | 2  |                         |
|   | Q10: Metadata is represented in a machine readable format, e.g. in a linked data format such as Resource Description Framework (RDF) | 3  |                         |
|   |  | 7/8=88%  | <b>I score</b>          |
|   | Q11: Standard machine-readable license (e.g. Creative Commons)   | 4  |                         |
|   | Q12: Fully recorded in a text format   | 2  |                         |
|   |  | 6/7=86%  | <b>R score</b>          |
|   |  | $65\% \times 0.25 + 90\% \times 0.25 + 88\% \times 0.25 + 86\% \times 0.25 = 83\%$ | <b>FAIR total score</b> |
| O1: Publicly accessible by humans and machines alike by standard protocol (open, no log in etc. required) | <b>100%</b>  | <b>Openness score</b>  |                         |

| Name of database                   | Answers to Q1-Q12 (ARDC)                                      | Score    | Criteria  |
|------------------------------------|---|----------|---|
| Covenant of Mayors                 | Q1: Web address (URL)   | 3        |   |
|                                    | Q2: Comprehensively, but in a text-based, non-standard format | 3        |   |
|                                    | Q3: No  | 0        |   |
|                                    | Q4: Local institutional repository                            | 2        |   |
|                                    |   | 8/17=47% | <b>F score</b>  |
|                                    | Q5: Publicly accessible                                       | 5        |   |
|                                    | Q6: File download from online location                        | 2        | Open for discussion, data not explicitly downloadable, but copyable |
|                                    | Q7: Unsure  | 0        |   |
|                                    |   | 7/10=70% | <b>A score</b>  |
| Q8: Mostly in a proprietary format | 0   |          |   |

|  |  |  |                         |
|--|--|--|-------------------------|
|  | Q9: Standardised vocabularies/ontologies/schema without global identifiers                         | 2  |                         |
|  | Q10: The metadata record includes URI links to related metadata, data and definitions              | 2  |                         |
|  |  | 4/8= <b>50%</b>  | <b>I score</b>          |
|  | Q11: No license  | 0  |                         |
|  | Q12: Fully recorded in a text format   | 2  |                         |
|  |  | 2/7= <b>29%</b>  | <b>R score</b>          |
|  |  | 47% x 0.25 + 70% x 0.25 + 50% x 0.25 + 29% x 0.25 = <b>49%</b> | <b>FAIR total score</b> |
|  | O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | <b>67%</b>   | <b>Openness score</b>   |

| Name of database   | Answers to Q1-Q12 (ARDC)  | Score  | Criteria                |
|--|---|--|-------------------------|
| DataCite   | Q1: Web address (URL)   | 3  |                         |
|  | Q2: Comprehensively (see suggestion) using a recognized formal machine-readable metadata schema                 | 4  |                         |
|  | Q3: Yes   | 1  |                         |
|  | Q4: Generalist public repository  | 2  |                         |
|  |   | 10/17= <b>59%</b>  | <b>F score</b>          |
|  | Q5: Fully accessible to persons who meet explicitly stated conditions, e.g., ethics approval for sensitive data | 5  |                         |
|  | Q6: Standard web service API (e.g. OGC)   | 4  |                         |
|  | Q7: Unsure  | 0  |                         |
|  |   | 9/10= <b>90%</b>   | <b>A score</b>          |
|  | Q8: In a structured, open-standard, machine-readable format   | 2  |                         |
|  | Q9: Standardised vocabularies/ontologies/schema without global identifiers                                      | 2  |                         |
|  | Q10: The metadata record includes URI links to related metadata, data and definitions                           | 2  |                         |
|  |   | 6/8= <b>75%</b>  | <b>I score</b>          |
|  | Q11: Standard text-based license  | 3  |                         |
|  | Q12: Partially recorded   | 1  |                         |
|  |   | 4/7= <b>57%</b>  | <b>R score</b>          |
|  |   | 59% x 0.25 + 90% x 0.25 + 75% x 0.25 + 57% x 0.25 = <b>70%</b> | <b>FAIR total score</b> |
| O1: Fully accessible to persons who meet explicitly stated conditions, (e.g., ethics approval for sensitive data, membership fees) | <b>33%</b>  | <b>Openness score</b>  |                         |

| Name of database | Answers to Q1-Q12 (ARDC) | Score | Criteria |
|------------------|--------------------------|-------|----------|
|------------------|--------------------------|-------|----------|

|  |  |                         |                |
|--|--|-------------------------|----------------|
| EMODnet - THE EUROPEAN MARINE OBSERVATION AND DATA NETWORK   | Q1: Web address (URL)  | 3                       |                |
|  | Q2: Comprehensively (see suggestion) using a recognised, formal machine-readable metadata schema | 4                       |                |
|  | Q3: No   | 0                       |                |
|  | Q4: Local institutional repository   | 2                       |                |
|  |  | 9/17= <b>53%</b>        | <b>F score</b> |
|  | Q5: Unspecified conditional access e.g. contact the data custodian for access                    | 2                       |                |
|  | Q6: By individual arrangement  | 1                       |                |
|  | Q7: Yes (the metadata is present but for the data itself contact custodian)                      | 1                       |                |
|  |  | 4/10= <b>40%</b>        | <b>A score</b> |
|  | Q8: In a structured, open standard, non-machine-readable format                                  | 2                       |                |
|  | Q9: Standardised vocabularies/ontologies/schema without global identifiers                       | 2                       |                |
|  | Q10: There are no links to other metadata  | 0                       |                |
|  |  | 4/8= <b>50%</b>         | <b>I score</b> |
|  | Q11: Standard machine-readable license (e.g. Creative Commons)                                   | 4                       |                |
|  | Q12: Partially recorded  | 1                       |                |
|  |  | 5/7= <b>71%</b>         | <b>R score</b> |
|  | 53% x 0.25 + 40% x 0.25 + 50% x 0.25 + 71% x 0.25 = <b>54%</b>                                   | <b>FAIR total score</b> |                |
| O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | <b>67%</b>   | <b>Openness score</b>   |                |

| Name of database  | Answers to Q1-Q12 (ARDC)   | Score            | Criteria       |
|---|--|------------------|----------------|
| Policies to Enforce the Transition to Nearly Zero Energy Buildings in EU-27 (Project: ENTRANZE) | Q1: Web address (URL)  | 3                |                |
|   | Q2: Comprehensively, but in a text-based, non-standard format                | 3                |                |
|   | Q3: No   | 0                |                |
|   | Q4: Generalist public repository   | 2                |                |
|   |  | 8/17= <b>47%</b> | <b>F score</b> |
|   | Q5: Publicly accessible  | 5                |                |
|   | Q6: File download from online location                                       | 2                |                |
|   | Q7: Unsure   | 0                |                |
|   |  | 7/10= <b>70%</b> | <b>A score</b> |
|   | Q8: In a structured, open standard, non-machine-readable format              | 1                |                |
|   | Q9: Standardized vocabularies/ontologies/schema without global identifiers - | 2                |                |
|   | Q10: There are no links to other metadata                                    |                  |                |
|   |  | 3/8= <b>38%</b>  | <b>I score</b> |
|   | Q11: Non-standard text-based license   | 2                |                |
|   | Q12: Partially recorded  | 1                |                |
|   |  | 3/7= <b>43%</b>  | <b>R score</b> |

|  |  |   |                         |
|--|--|---|-------------------------|
|  |  | $47\% \times 0.25 + 70\% \times 0.25 + 38\% \times 0.25 + 43\% \times 0.25 = \mathbf{49.5\%}$ | <b>FAIR total score</b> |
|  | O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | 2/3=67%   | <b>Openness Score</b>   |

| <b>Name of database</b>             | <b>Answers to Q1-Q12 (ARDC)</b>  | <b>Score</b>            | <b>Criteria</b>       |
|-------------------------------------|--|-------------------------|-----------------------|
| ENTSO-E<br>Transparency<br>Platform | Q1: Web address (URL)  | 3                       |                       |
|                                     | Q2: Brief title and description  | 2                       |                       |
|                                     | Q3: No   | 0                       |                       |
|                                     | Q4: Local institutional repository   | 2                       |                       |
|                                     |  | 7/17= <b>41%</b>        | <b>F score</b>        |
|                                     | Q5: Publicly accessible  | 5                       |                       |
|                                     | Q6: File download from online location   | 2                       |                       |
|                                     | Q7: No   | 0                       |                       |
|                                     |  | 7/10= <b>70%</b>        | <b>A score</b>        |
|                                     | Q8: Mostly in a proprietary format   | 0                       |                       |
|                                     | Q9: Data elements not described  | 0                       |                       |
|                                     | Q10: There are no links to other metadata  | 0                       |                       |
|                                     |  | 0/8= <b>0%</b>          | <b>I score</b>        |
|                                     | Q11: No license  | 0                       |                       |
|                                     | Q12: No provenance information is recorded   | 0                       |                       |
|                                     |  | 0/7= <b>0%</b>          | <b>R score</b>        |
|                                     | $41\% \times 0.25 + 70\% \times 0.25 + 0\% \times 0.25 + 0\% \times 0.25 = \mathbf{28\%}$          | <b>FAIR total score</b> |                       |
|                                     | O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | <b>67%</b>              | <b>Openness score</b> |

| <b>Name of database</b>                         | <b>Answers to Q1-Q12 (ARDC)</b>   | <b>Score</b>     | <b>Criteria</b> |
|---|---|------------------|-----------------|
| EUR-Lex<br>(Document:<br>Directive<br>95/46/EC) | Q1: Web address (URL)   | 3                |                 |
|   | Q2: Comprehensively, but in a text-based, non-standard format   | 3                |                 |
|   | Q3: no  | 0                |                 |
|   | Q4: Domain-specific repository  | 2                |                 |
|   |   | 8/17= <b>47%</b> | <b>F score</b>  |
|   | Q5: Publicly accessible   | 5                |                 |
|   | Q6: Non-standard web service  | 3                |                 |
|   | Q7: Unsure (data a. metadata are available but not clear if metadata will still be available when the data itself is removed) | 0                |                 |
|   |   | 8/10= <b>80%</b> | <b>A score</b>  |
|   | Q8: In a structured, open-standard, non-machine-readable format   | 1                |                 |
|   | Q9: Standardised vocabularies/ ontologies/ schema without global identifiers  | 2                |                 |
|   | Q10: The metadata record includes URI links to related metadata, data, and definitions  | 2                |                 |
|   | 5/8= <b>63%</b>   | <b>I score</b>   |                 |
| Q11: Non-standard text-based license            | 2   |                  |                 |

|  |  |  |                         |
|--|--|--|-------------------------|
|  | Q12: Fully recorded in a text format   | 2  |                         |
|  |  | 4/7= <b>57%</b>  | <b>R score</b>          |
|  |  | 47% x 0.25 + 80% x 0.25 + 63% x 0.25 + 57% x 0.25 = <b>62%</b> | <b>FAIR total score</b> |
|  | O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | <b>67 %</b>  | <b>Openness score</b>   |

| Name of database  | Answers to Q1-Q12 (ARDC)  | Score                   | Criteria       |
|---|---|-------------------------|----------------|
| European Energy Efficiency Building District Database (Project: ExcEED) | Q1: Local identifier  | 1                       |                |
|   | Q2: Comprehensively, but in a text-based, non-standard format                 | 3                       |                |
|   | Q3: No  | 0                       |                |
|   | Q4: The data is not described in any repository                               | 0                       |                |
|   |   | 4/17= <b>23%</b>        | <b>F score</b> |
|   | Q5: Unspecified conditional access e.g. contact the data custodian for access | 2                       |                |
|   | Q6: By individual arrangement   | 1                       |                |
|   | Q7: No  | 0                       |                |
|   |   | 3/10= <b>30%</b>        | <b>A score</b> |
|   | Q8: In a structured, open standard, non-machine-readable format               | 1                       |                |
|   | Q9: Standardized vocabularies/ontologies/schema without global identifiers -  | 2                       |                |
|   | Q10: There are no links to other metadata                                     |                         |                |
|   |   | 3/8= <b>38%</b>         | <b>I score</b> |
|   | Q11: No license   | 0                       |                |
|   | Q12: No provenance information is recorded                                    | 0                       |                |
|   |   | 0/7= <b>0%</b>          | <b>R score</b> |
|   | 23% x 0.25 + 30% x 0.25 + 38% x 0.25 + 0% x 0.25 = <b>23%</b>                 | <b>FAIR total score</b> |                |
| O1: Data are not open   | 0%  | <b>Openness Score</b>   |                |

| Name of database                            | Answers to Q1-Q12 (ARDC)  | Score             | Criteria       |
|---|---|-------------------|----------------|
| Global Buildings Performance Network (GBPN) | Q1: Globally unique, citable and persistent (e.g. DOI, PURL, ARK or Handle) | 8                 |                |
|   | Q2: Brief title and description   | 2                 |                |
|   | Q3: No  | 0                 |                |
|   | Q4: Generalist public repository  | 2                 |                |
|   |   | 12/17= <b>70%</b> | <b>F score</b> |
|   | Q5: Publicly accessible   | 5                 |                |
|   | Q6: File download from online location                                      | 2                 |                |
|   | Q7: Unsure  | 0                 |                |
|   | 7/10= <b>70%</b>  | <b>A score</b>    |                |



|  |   |                         |
|--|---|-------------------------|
| Q8: In a structured, open standard, non-machine-readable format                                    | 1   |                         |
| Q9: No standards have been applied in the description of data elements                             | 1   |                         |
| Q10: The metadata record includes URI links to related metadata, data and definitions              | 2   |                         |
|  | 4/8= <b>50%</b>   | <b>I score</b>          |
| Q11: No license  | 0   |                         |
| Q12: Partially recorded  | 1   |                         |
|  | 1/7= <b>14%</b>   | <b>R score</b>          |
|  | $70\% \times 0.25 + 70\% \times 0.25 + 50\% \times 0.25 + 14\% \times 0.25 = \mathbf{51\%}$ | <b>FAIR total score</b> |
| O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | 2/3 = <b>67%</b>  | <b>Openness Score</b>   |

| Name of database   | Answers to Q1-Q12 (ARDC)   | Score   | Criteria                |
|--|--|---|-------------------------|
| The Open Source Mapping and Planning Tool for Cooling and Heating (Project: Hotmaps)               | Q1: Web address (URL)  | 3   |                         |
|  | Q2: Comprehensively, but in a text-based, non-standard format              | 3   |                         |
|  | Q3: No   | 0   |                         |
|  | Q4: Local institutional repository   | 2   |                         |
|  |  | 8/17= <b>47%</b>  | <b>F score</b>          |
|  | Q5: Publicly accessible  | 5   |                         |
|  | Q6: File download from online location                                     | 2   |                         |
|  | Q7: Unsure   | 0   |                         |
|  |  | 7/10= <b>70%</b>  | <b>A score</b>          |
|  | Q8: In a structured, open standard, non-machine-readable format            | 2   |                         |
|  | Q9: Standardised vocabularies/ontologies/schema without global identifiers | 2   |                         |
|  | Q10: There are no links to other metadata                                  | 0   |                         |
|  |  | 4/8= <b>50%</b>   | <b>I score</b>          |
|  | Q11: Standard machine-readable license (e.g. Creative Commons)             | 4   |                         |
|  | Q12: Fully recorded in a text format                                       | 2   |                         |
|  |  | 6/7= <b>86%</b>   | <b>R score</b>          |
|  |  | $47\% \times 0.25 + 70\% \times 0.25 + 50\% \times 0.25 + 86\% \times 0.25 = \mathbf{63\%}$ | <b>FAIR total score</b> |
| O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | <b>67%</b>   | <b>Openness score</b>   |                         |

| Name of database    | Answers to Q1-Q12 (ARDC)                                      | Score            | Criteria       |
|---------------------|---|------------------|----------------|
| IEA Policy database | Q1: Web address (URL)   | 3                |                |
|                     | Q2: Comprehensively, but in a text-based, non-standard format | 3                |                |
|                     | Q3: No  | 0                |                |
|                     | Q4: Local institutional repository                            | 2                |                |
|                     |   | 8/17= <b>47%</b> | <b>F score</b> |

|   |  |                         |
|---|--|-------------------------|
| Q5: Unspecified conditional access e.g. contact the data custodian for access                             | 2  |                         |
| Q6: File download from online location  | 2  |                         |
| Q7: Yes   | 1  |                         |
|   | 5/10=50%   | <b>A score</b>          |
| Q8: In a structured, open-standard, non-machine-readable format   | 1  |                         |
| Q9: Standardized vocabularies/ontologies/schema without global identifiers                                | 2  |                         |
| Q10: There are no links to other metadata   | 0  |                         |
|   | 3/8=38%  | <b>I score</b>          |
| Q11: Standard text-based license  | 3  |                         |
| Q12: Partially recorded   | 1  |                         |
|   | 4/7=57%  | <b>R score</b>          |
|   | $47\% \times 0.25 + 50\% \times 0.25 + 38\% \times 0.25 + 57\% \times 0.25 = 48\%$ | <b>FAIR total score</b> |
| O1: Publicly accessible by humans and machines alike by standard protocol (open, no log in etc. required) | <b>100%</b>  | <b>Openness score</b>   |

| Name of database                         | Answers to Q1-Q12 (ARDC)  | Score                   | Criteria       |
|--|---|-------------------------|----------------|
| IEA World Energy Statistics and Balances | Q1: Web address (URL)   | 3                       |                |
|  | Q2: Comprehensively, but in a text-based, non-standard format. Information spread across many files/websites. | 3                       |                |
|  | Q3: No. Note, data needs to be purchased.   | 0                       |                |
|  | Q4: Local institutional repository.   | 2                       |                |
|  |   | 8/17=47%                | <b>F score</b> |
|  | Q5: Unspecified conditional access e.g. contact the data custodian for access                                 | 2                       |                |
|  | Q6: File download from online location  | 2                       |                |
|  | Q7: Yes   | 1                       |                |
|  |   | 5/10=50%                | <b>A score</b> |
|  | Q8: In a structured, open-standard, non-machine-readable format   | 1                       |                |
|  | Q9: Standardized vocabularies/ontologies/schema without global identifiers                                    | 2                       |                |
|  | Q10: To the extent open, there seems no links to other metadata   | 0                       |                |
|  |   | 3/8=38%                 | <b>I score</b> |
|  | Q11: Standard text-based license for purchasing the data.   | 3                       |                |
|  | Q12: Information not accessible (partially recorded is likely).   | 1                       |                |
|  | 4/7=57%   | <b>R score</b>          |                |
|  | $47\% \times 0.25 + 50\% \times 0.25 + 38\% \times 0.25 + 57\% \times 0.25 = 48\%$                            | <b>FAIR total score</b> |                |
| O1: Data for purchase.                   | <b>0%</b>   | <b>Openness score</b>   |                |

| Name of database | Answers to Q1-Q12 (ARDC) | Score | Criteria |
|------------------|--------------------------|-------|----------|
|------------------|--------------------------|-------|----------|

|   |   |  |                         |
|---|---|--|-------------------------|
| JRC data catalog  | Q1: Web address (URL)   | 3  |                         |
|   | Q2: Brief title and description   | 2  |                         |
|   | Q3: Yes   | 1  |                         |
|   | Q4: Generalist public repository  | 2  |                         |
|   |   | 8/17=47%   | <b>F score</b>          |
|   | Q5: A de-identified / modified subset of the data is publicly accessible              | 4  |                         |
|   | Q6: File download from online location  | 2  |                         |
|   | Q7: Yes   | 1  |                         |
|   |   | 7/10=70%   | <b>A score</b>          |
|   | Q8: In a structured, open-standard, non-machine-readable format                       | 1  |                         |
|   | Q9: Standardized vocabularies/ontologies/schema without global identifiers            | 2  |                         |
|   | Q10: The metadata record includes URI links to related metadata, data and definitions | 2  |                         |
|   |   | 5/8=63%  | <b>I score</b>          |
|   | Q11: Standard text-based license  | 3  |                         |
|   | Q12: Partially recorded   | 1  |                         |
|   |   | 5/7=57%  | <b>R score</b>          |
|   |   | 53% x 0.25 + 70% x 0.25 + 63% x 0.25 + 57% x 0.25 = <b>61%</b> | <b>FAIR total score</b> |
| Q1: Publicly accessible by humans and machines alike by standard protocol (open, no log in etc. required) | <b>100%</b>   | <b>Openness score</b>  |                         |

| Name of database                   | Answers to Q1-Q12 (ARDC)   | Score   | Criteria                |
|------------------------------------|--|---|-------------------------|
| JRC data catalog (data collection) | Q1: Web address (URL)  | 3   |                         |
|                                    | Q2: Brief title and description  | 2   |                         |
|                                    | Q3: Yes  | 1   |                         |
|                                    | Q4: Generalist public repository   | 2   |                         |
|                                    |  | 8/17=47%  | <b>F score</b>          |
|                                    | Q5: Publicly accessible - the data set organization is visible to the users of the website showing the collection. | 5   |                         |
|                                    | Q6: Standard web services (http) is used.  | 4   |                         |
|                                    | Q7: Yes  | 1   |                         |
|                                    |  | 10/10=100%  | <b>A score</b>          |
|                                    | Q8: Data (=names of the datasets available) are presented in a structured, open-standard, machine-readable format  | 2   |                         |
|                                    | Q9: No specific standards have been applied.   | 1   |                         |
|                                    | Q10: There are no links to other metadata  | 0   |                         |
|                                    |  | 3/8=38%   | <b>I score</b>          |
|                                    | Q11: Standard text-based license   | 3   |                         |
|                                    | Q12: Not systematically recorded, only mentions 'updated irregularly'.   | 0   |                         |
|                                    |  | 3/7=43%   | <b>R score</b>          |
|                                    |  | 47% x 0.25 + 100% x 0.25 + 38% x 0.25 + 43% x 0.25 = <b>57%</b> | <b>FAIR total score</b> |

|  |   |             |                       |
|--|---|-------------|-----------------------|
|  | O1: Publicly accessible by humans and machines alike by standard protocol (open, no log in etc. required) | <b>100%</b> | <b>Openness score</b> |
|--|---|-------------|-----------------------|

| <b>Name of database</b>          | <b>Answers to Q1-Q12 (ARDC)</b>   | <b>Score</b>   | <b>Criteria</b>         |
|----------------------------------|---|--|-------------------------|
| JRC data catalogue (dataset)     | Q1: Web address (URL)   | 3  |                         |
|                                  | Q2: Comprehensively, but in a text-based, non-standard format   | 3  |                         |
|                                  | Q3: No  | 0  |                         |
|                                  | Q4: Generalist institutional repository   | 2  |                         |
|                                  |   | <b>8/17=47%</b>  | <b>F score</b>          |
|                                  | Q5: Public access when accepting the license.   | 5  |                         |
|                                  | Q6: File download from online location (with several redirections)  | 2  |                         |
|                                  | Q7: Yes   | 1  |                         |
|                                  |   | <b>8/10=80%</b>  | <b>A score</b>          |
|                                  | Q8: In a structured, open-standard, non-machine-readable format   | 1  |                         |
|                                  | Q9: No standards have been applied.   | 1  |                         |
|                                  | Q10: There are no links to other metadata   | 0  |                         |
|                                  |   | <b>2/8=25%</b>   | <b>I score</b>          |
|                                  | Q11: Standard text-based license  | 3  |                         |
|                                  | Q12: Not included in the metadata location. The zip-file from data download may contain such information, but is not accessible to humans without special visualization tools (e.g., using file-formats such as INF). | 0  |                         |
|                                  |   | <b>3/7=43%</b>   | <b>R score</b>          |
|                                  |   | <b>47% x 0.25 + 80% x 0.25 + 25% x 0.25 + 43% x 0.25 = 49%</b> | <b>FAIR total score</b> |
| O1: Human intervention required. | <b>66%</b>  | <b>Openness score</b>  |                         |

| <b>Name of database</b>   | <b>Answers to Q1-Q12 (ARDC)</b>  | <b>Score</b>    | <b>Criteria</b> |
|---|--|-----------------|-----------------|
| Mesures d'utilisation Rationelle d'Energi (MURE database)                             | Q1: Web address (URL)  | 3               |                 |
|   | Q2: Comprehensively, but in a text-based, non-standard format              | 3               |                 |
|   | Q3: No   | 0               |                 |
|   | Q4: Domain-specific repository   | 2               |                 |
|   |  | <b>8/17=47%</b> | <b>F score</b>  |
|   | Q5: Publicly accessible  | 5               |                 |
|   | Q6: File download from online location                                     | 2               |                 |
|   | Q7: Unsure   | 0               |                 |
|   |  | <b>7/10=70%</b> | <b>A score</b>  |
|   | Q8: In a structured, open-standard, non-machine-readable format            | 1               |                 |
|   | Q9: Standardized vocabularies/ontologies/schema without global identifiers | 2               |                 |
| Q10: The metadata record includes URI links to related metadata, data and definitions | 2  |                 |                 |

|  |  |  |                         |
|--|--|--|-------------------------|
|  |  | 5/8=63%  | <b>I score</b>          |
|  | Q11: No license  | 0  |                         |
|  | Q12: Fully recorded in a text format   | 2  |                         |
|  |  | 2/7=29%  | <b>R score</b>          |
|  |  | 47% x 0.25 + 70% x 0.25 + 63% x 0.25 + 29% x 0.25 = <b>52%</b> | <b>FAIR total score</b> |
|  | O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | 2/3 = <b>67%</b>   | <b>Openness score</b>   |

| Name of database                      | Answers to Q1-Q12 (ARDC)   | Score                   | Criteria       |
|---------------------------------------|--|-------------------------|----------------|
| Kaggle Repository                     | Q1: Yes, URL is assigned to metadata.  | 3                       |                |
|                                       | Q2: Very brief metadata description with title and description, only.                              | 2                       |                |
|                                       | Q3: The dataset is not included in the files describing the data.                                  | 0                       |                |
|                                       | Q4: The data is published in a generalist public repository.                                       | 2                       |                |
|                                       |  | 7/17=41%                | <b>F score</b> |
|                                       | Q5: Data is publicly accessible.   | 5                       |                |
|                                       | Q6: Data is available through download from an online location.                                    | 2                       |                |
|                                       | Q7: Yes, likely.   | 1                       |                |
|                                       |  | 8/10=80%                | <b>A score</b> |
|                                       | Q8: The data are available as csv, open but not machine-readable as the headers are not explained. | 1                       |                |
|                                       | Q9: No standards for vocabularies etc. have been used.   | 1                       |                |
|                                       | Q10: There are no links to other metadata.   | 0                       |                |
|                                       |  | 2/8=25%                 | <b>I score</b> |
|                                       | Q11: A standard text-based license is assigned.  | 3                       |                |
|                                       | Q12: Information about the updating frequency has not been provided.                               | 0                       |                |
|                                       | 3/7=43%  | <b>R score</b>          |                |
|                                       | 41% x 0.25 + 80% x 0.25 + 25% x 0.25 + 43% x 0.25 = <b>47%</b>                                     | <b>FAIR total score</b> |                |
| O1: Open for humans but not machines. | <b>66%</b>   | <b>Openness score</b>   |                |

| Name of database   | Answers to Q1-Q12 (ARDC)  | Score          | Criteria       |
|--|---|----------------|----------------|
| Simulations of hourly power output from wind and solar PV farms (renewables.ninja) | Q1: Globally unique, citable and persistent (e.g. DOI, PURL, ARK or Handle) | 8              |                |
|  | Q2: Comprehensively, but in a text-based, non-standard format               | 3              |                |
|  | Q3: No  | 0              |                |
|  | Q4: Domain-specific repository  | 2              |                |
|  |   | 13/17=76%      | <b>F score</b> |
|  | Q5: Publicly accessible   | 5              |                |
|  | Q6: File download from online location                                      | 2              |                |
|  | Q7: Unsure  | 0              |                |
|  | 7/10=70%  | <b>A score</b> |                |

|  |  |  |                         |
|--|--|--|-------------------------|
|  | Q8: In a structured, open standard, machine-readable format  | 2  |                         |
|  | Q9: Standardised vocabularies/ontologies/schema without global identifiers                         | 2  |                         |
|  | Q10: The metadata record includes URI links to related metadata, data and definitions              | 2  |                         |
|  |  | 6/8= <b>75%</b>  | <b>I score</b>          |
|  | Q11: Standard machine-readable license (e.g. Creative Commons)                                     | 4  |                         |
|  | Q12: Partially recorded  | 1  |                         |
|  |  | 5/7= <b>71%</b>  | <b>R score</b>          |
|  |  | 76% x 0.25 + 70% x 0.25 + 75% x 0.25 + 71% x 0.25 = <b>73%</b> | <b>FAIR total score</b> |
|  | O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | <b>67%</b>   | <b>Openness score</b>   |

| Name of database  | Answers to Q1-Q12 (ARDC)  | Score                   | Criteria       |
|---|---|-------------------------|----------------|
| U.S. National Renewable Energy Laboratory (NREL)  | Q1: Globally unique, citable and persistent (e.g. DOI, PURL, ARK or Handle)                     | 8                       |                |
|   | Q2: Comprehensively (see suggestion) using a recognized formal machine-readable metadata schema | 4                       |                |
|   | Q3: Yes   | 1                       |                |
|   | Q4: Data is in one place but discoverable through several registries                            | 4                       |                |
|   |   | 17/17= <b>100%</b>      | <b>F score</b> |
|   | Q5: Publicly accessible   | 5                       |                |
|   | Q6: Standard web service API  | 4                       |                |
|   | Q7: Unsure  | 0                       |                |
|   |   | 9/10= <b>90%</b>        | <b>A score</b> |
|   | Q8: In a structured, open-standard, machine-readable format                                     | 2                       |                |
|   | Q9: Standardized vocabularies/ontologies/schema without global identifiers                      | 2                       |                |
|   | Q10: There are no links to other metadata   | 0                       |                |
|   |   | 4/8= <b>50%</b>         | <b>I score</b> |
|   | Q11: No license   | 0                       |                |
|   | Q12: Fully recorded in a text format  | 2                       |                |
|   | 2/7= <b>29%</b>   | <b>R score</b>          |                |
|   | 100% x 0.25 + 90% x 0.25 + 50% x 0.25 + 29% x 0.25 = <b>67%</b>                                 | <b>FAIR total score</b> |                |
| O1: Publicly accessible by humans and machines alike by standard protocol (open, no log in etc. required) | <b>100%</b>   | <b>Openness score</b>   |                |

| Name of database                              | Answers to Q1-Q12 (ARDC)        | Score | Criteria |
|---|---------------------------------|-------|----------|
| Scenarios of market transition to nearly zero | Q1: Web address (URL)           | 5     |          |
|   | Q2: Brief title and description | 2     |          |
|   | Q3: No                          | 0     |          |

|  |  |   |                         |
|--|--|---|-------------------------|
| energy buildings (NZEB)  | Q4: Generalist public repository   | 2   |                         |
|  |  | 9/17= <b>53%</b>  | <b>F score</b>          |
|  | Q5: Publicly accessible  | 5   |                         |
|  | Q6: By individual arrangement  | 1   |                         |
|  | Q7: Unsure   | 0   |                         |
|  |  | 6/10= <b>60%</b>  | <b>A score</b>          |
|  | Q8: In a structured, open standard, non-machine-readable format              | 1   |                         |
|  | Q9: Standardized vocabularies/ontologies/schema without global identifiers - | 2   |                         |
|  | Q10: There are no links to other metadata                                    | 0   |                         |
|  |  | 3/8= <b>38%</b>   | <b>I score</b>          |
|  | Q11: No license  | 0   |                         |
|  | Q12: No provenance information is recorded                                   | 0   |                         |
|  |  | 0/7= <b>0%</b>  | <b>R score</b>          |
|  |  | 53% x 0.25 + 60% x 0.25 + 38% x 0.25 + 0% x 0.25 = <b>38%</b> | <b>FAIR total score</b> |
| O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | 2/3 = 67%  | <b>Openness Score</b>   |                         |

| Name of database                                 | Answers to Q1-Q12 (ARDC)   | Score             | Criteria       |
|--|--|-------------------|----------------|
| OECD hub (energy data)                           | Q1: Web address (URL)  | 3                 |                |
|  | Q2: Comprehensively (see suggestion) using a recognised, formal machine-readable metadata schema                               | 4                 |                |
|  | Q3: yes  | 1                 |                |
|  | Q4: A)Data is in one place but discoverable through several registries   | 4                 |                |
|  |  | 12/17= <b>70%</b> | <b>F score</b> |
|  | Q5: Publicly accessible  | 5                 |                |
|  | Q6: : Standard web service API (e.g. OGC)  | 4                 |                |
|  | Q7: Unsure (data and metadata are available but not clear if metadata will still be available when the data itself is removed) | 0                 |                |
|  |  | 9/10= <b>90%</b>  | <b>A score</b> |
|  | Q8: I In a structured, open standard, machine-readable format  | 2                 |                |
|  | Q9: Standardised vocabularies/ontologies/schema without global identifiers   | 2                 |                |
|  | Q10: The metadata record includes URI links to related metadata, data, and definitions   | 2                 |                |
|  |  | 6/8= <b>75%</b>   | <b>I score</b> |
|  | Q11: Standard machine-readable license (e.g. Creative Commons)   | 4                 |                |
| Q12: Fully recorded in a machine-readable format | 3  |                   |                |
|  | 7/7= <b>100%</b>   | <b>R score</b>    |                |

|  |   |   |                         |
|--|---|---|-------------------------|
|  |   | $70\% \times 0.25 + 90\% \times 0.25 + 75\% \times 0.25 + 100\% \times 0.25 = 84\%$ | <b>FAIR total score</b> |
|  | O1: Publicly accessible by humans and machines alike by standard protocol (open, no log in etc. required) | <b>100%</b>   | <b>Openness score</b>   |

| Name of database  | Answers to Q1-Q12 (ARDC)   | Score                   | Criteria              |
|-------------------|--|-------------------------|-----------------------|
| OECD - (data set) | Q1: Globally unique, citable, and persist  | 8                       |                       |
|                   | Q2: Comprehensive, but in a text-based, non-standard format  | 3                       |                       |
|                   | Q3: No   | 0                       |                       |
|                   | Q4: Local institutional repository   | 2                       |                       |
|                   |  | 13/17= <b>77%</b>       | <b>F score</b>        |
|                   | Q5: Publicly accessible  | 5                       |                       |
|                   | Q6: File download from online location   | 2                       |                       |
|                   | Q7: Unsure   | 0                       |                       |
|                   |  | 7/10= <b>70%</b>        | <b>A score</b>        |
|                   | Q8: In a structure, open standard, machine-readable format   | 2                       |                       |
|                   | Q9: No standards have been applied in the description of data elements.                            | 1                       |                       |
|                   | Q10: There are no links to other metadata.   | 0                       |                       |
|                   |  | 3/8= <b>38%</b>         | <b>I score</b>        |
|                   | Q11: Non-standard text-based license.  | 2                       |                       |
|                   | Q12: No provenance information is recorded.  | 0                       |                       |
|                   | 2/7= <b>29%</b>  | <b>R score</b>          |                       |
|                   | $77\% \times 0.25 + 70\% \times 0.25 + 38\% \times 0.25 + 29\% \times 0.25 = 53\%$                 | <b>FAIR total score</b> |                       |
|                   | O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | <b>67%</b>              | <b>Openness score</b> |

| Name of database           | Answers to Q1-Q12 (ARDC)   | Score             | Criteria       |
|----------------------------|--|-------------------|----------------|
| Heat Roadmap Europe (PETA) | Q1: Web address (URL)  | 5                 |                |
|                            | Q2: Comprehensively, but in a text-based, non-standard format                | 3                 |                |
|                            | Q3: No   | 0                 |                |
|                            | Q4: Generalist public repository   | 2                 |                |
|                            |  | 10/17= <b>59%</b> | <b>F score</b> |
|                            | Q5: Publicly accessible  | 5                 |                |
|                            | Q6: By individual arrangement  | 1                 |                |
|                            | Q7: Unsure   | 0                 |                |
|                            |  | 6/10= <b>60%</b>  | <b>A score</b> |
|                            | Q8: In a structured, open standard, non-machine-readable format              | 1                 |                |
|                            | Q9: Standardized vocabularies/ontologies/schema without global identifiers - | 2                 |                |
|                            | Q10: There are no links to other metadata                                    | 0                 |                |
|                            | 3/8= <b>38%</b>  | <b>I score</b>    |                |



|  |  |   |                         |
|--|--|---|-------------------------|
|  | Q11: No license  | 0   |                         |
|  | Q12: No provenance information is recorded   | 0   |                         |
|  |  | 0/7= <b>0%</b>  | <b>R score</b>          |
|  |  | 59% x 0.25 + 60% x 0.25 + 38% x 0.25 + 0% x 0.25 = <b>39%</b> | <b>FAIR total score</b> |
|  | O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | 2/3 = <b>67%</b>  | <b>Openness Score</b>   |

| Name of database  | Answers to Q1-Q12 (ARDC)   | Score                   | Criteria       |
|---|--|-------------------------|----------------|
| Database for support schemes, grid issues and policies about RE sources in EU-28, EFTA, etc. (RES LEGAL)  | Q1: Web address (URL)  | 3                       |                |
|   | Q2: Comprehensively, but in a text-based, non-standard format              | 3                       |                |
|   | Q3: No   | 0                       |                |
|   | Q4: Local institutional repository   | 2                       |                |
|   |  | 8/17= <b>47%</b>        | <b>F score</b> |
|   | Q5: Publicly accessible  | 5                       |                |
|   | Q6: File download from online location                                     | 2                       |                |
|   | Q7: Yes  | 1                       |                |
|   |  | 8/10= <b>80%</b>        | <b>A score</b> |
|   | Q8: In a structured, open-standard, non-machine-readable format            | 1                       |                |
|   | Q9: Standardized vocabularies/ontologies/schema without global identifiers | 2                       |                |
|   | Q10: There are no links to other metadata                                  | 0                       |                |
|   |  | 3/8= <b>38%</b>         | <b>I score</b> |
|   | Q11: No license  | 0                       |                |
|   | Q12: Fully recorded in a text format                                       | 2                       |                |
|   | 2/7= <b>29%</b>  | <b>R score</b>          |                |
|   | 47% x 0.25 + 80% x 0.25 + 38% x 0.25 + 29% x 0.25 = <b>49%</b>             | <b>FAIR total score</b> |                |
| O1: Publicly accessible by humans and machines alike by standard protocol (open, no log in etc. required) | <b>100%</b>  | <b>Openness score</b>   |                |

| Name of database  | Answers to Q1-Q12 (ARDC)   | Score            | Criteria       |
|---|--|------------------|----------------|
| Smart integration of energy storages in local multi-energy systems for maximizing the share of renewables in Europe's energy mix (Smiles) | Q1: Web address (URL)  | 3                |                |
|   | Q2: Comprehensively (see suggestion) using a recognised, formal machine-readable metadata schema | 4                |                |
|   | Q3: No   | 0                |                |
|   | Q4: Domain-specific repository   | 2                |                |
|   |  | 9/17= <b>53%</b> | <b>F score</b> |
|   | Q5: Publicly accessible  | 5                |                |
|   | Q6: File download from online location   | 2                |                |
|   | Q7: Unsure   | 0                |                |
|   | 7/10= <b>70%</b>   | <b>A score</b>   |                |
| Q8: In a structured, open standard, machine-readable format   | 2  |                  |                |

|  |  |  |                         |
|--|--|--|-------------------------|
|  | Q9: Standardised vocabularies/ontologies/schema without global identifiers                         | 2  |                         |
|  | Q10: There are no links to other metadata  | 0  |                         |
|  |  | 4/8= <b>50%</b>  | <b>I score</b>          |
|  | Q11: No license  | 0  |                         |
|  | Q12: Partially recorded  | 1  |                         |
|  |  | 1/7= <b>14%</b>  | <b>R score</b>          |
|  |  | 53% x 0.25 + 70% x 0.25 + 50% x 0.25 + 14% x 0.25 = <b>47%</b> | <b>FAIR total score</b> |
|  | O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | <b>67%</b>   | <b>Openness score</b>   |

| Name of database   | Answers to Q1-Q12 (ARDC)   | Score                   | Criteria       |
|--|--|-------------------------|----------------|
| EU Buildings Stock Monitoring (Project: EPISCOPE)  | Q1: Web address (URL)  | 3                       |                |
|  | Q2: Comprehensively, but in a text-based, non-standard format                | 3                       |                |
|  | Q3: No   | 0                       |                |
|  | Q4: Domain-specific repository   | 2                       |                |
|  |  | 8/17= <b>47%</b>        | <b>F score</b> |
|  | Q5: Publicly accessible  | 5                       |                |
|  | Q6: File download from online location                                       | 2                       |                |
|  | Q7: Unsure   | 0                       |                |
|  |  | 7/10= <b>70%</b>        | <b>A score</b> |
|  | Q8: Mostly in a proprietary format   | 0                       |                |
|  | Q9: Standardized vocabularies/ontologies/schema without global identifiers - | 2                       |                |
|  | Q10: There are no links to other metadata                                    |                         |                |
|  |  | 2/8= <b>25%</b>         | <b>I score</b> |
|  | Q11: Non-standard text-based license   | 2                       |                |
|  | Q12: Partially recorded  | 1                       |                |
|  |  | 3/7= <b>43%</b>         | <b>R score</b> |
|  | 47% x 0.25 + 70% x 0.25 + 25% x 0.25 + 43% x 0.25 = <b>46%</b>               | <b>FAIR total score</b> |                |
| O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.) | 2/3=67%  | <b>Openness Score</b>   |                |

| Name of database                     | Answers to Q1-Q12 (ARDC)   | Score | Criteria          |
|--------------------------------------|--|-------|-------------------|
| Zenodo Repository (Data set 'Greco') | Q1: Globally unique, citable and persistent (e.g. DOI, PURL, ARK or Handle)                      | 8     | Web address (URL) |
|                                      | Q2: Comprehensively (see suggestion) using a recognised, formal machine-readable metadata schema | 4     | Yes               |
|                                      | Q3: No   | 0     |                   |

|  |  |                         |
|--|--|-------------------------|
| Q4: Data is in one place but discoverable through several registries   | 4  |                         |
|  | 16/17=94%  | <b>F score</b>          |
| Q5: Publicly accessible  | 5  | Publicly accessible     |
| Q6: Standard web service API (e.g. OGC)  | 4  |                         |
| Q7: Unsure (data and metadata are available but not clear if metadata will still be available when the data itself is removed)       | 0  | No                      |
|  | 9/10=90%   | <b>A score</b>          |
| Q8: In a structured, open standard, machine-readable format  | 2  |                         |
| Q9: Standardised vocabularies/ontologies/schema without global identifiers   | 2  |                         |
| Q10: Metadata is represented in a machine readable format, e.g. in a linked data format such as Resource Description Framework (RDF) | 3  |                         |
|  | 7/8=88%  | <b>I score</b>          |
| Q11: Standard machine-readable license (e.g. Creative Commons)   | 4  |                         |
| Q12: Fully recorded in a machine-readable format   | 3  |                         |
|  | 7/7=100%   | <b>R score</b>          |
|  | 94% x 0.25 + 90% x 0.25 + 88% x 0.25 + 100% x 0.25 = 93% | <b>FAIR total score</b> |
| O1: Publicly accessible but requires human intervention (e.g., may require login to download etc.)                                   | 100 %  | <b>Openness score</b>   |

| Name of database  | Answers to Q1-Q12 (ARDC)   | Score          | Criteria       |
|---|--|----------------|----------------|
| Zenodo Repository (data set 'Sharewind')                | Q1: Using a globally unique, persistent identifier.  | 8              |                |
|   | Q2: Comprehensively, using a recognized formal machine-readable metadata schema (Open graph protocol).               | 4              |                |
|   | Q3: The dataset is part of the metadata description.   | 1              |                |
|   | Q4: Data is in one place (zenodo repository) but discoverable through several registries.                            | 4              |                |
|   |  | 17/17=100%     | <b>F score</b> |
|   | Q5: Embargoed to ensure first author publication.  | 3              |                |
|   | Q6: Standard web service API is used.  | 4              |                |
|   | Q7: Yes, likely.   | 1              |                |
|   |  | 8/10=80%       | <b>A score</b> |
|   | Q8: Files are given as bmp-files, with limited access for algorithms.  | 1              |                |
|   | Q9: Mixed - standardized but not systematically with global identifiers. Content itself not a controlled vocabulary. | 2              |                |
|   | Q10: The metadata record includes URI links to related metadata, data, and definitions.                              | 2              |                |
|   |  | 5/8=63%        | <b>I score</b> |
|   | Q11: Linked to a text-based license.   | 3              |                |
| Q12: Partially recorded through versioning information. | 1  |                |                |
|   | 4/7=57%  | <b>R score</b> |                |

|  |   |   |                         |
|--|---|---|-------------------------|
|  |   | $100\% \times 0.25 + 80\% \times 0.25 + 63\% \times 0.25 + 57\% \times 0.25 = 75\%$ | <b>FAIR total score</b> |
|  | O1: Publicly available for machines and humans alike. | <b>100%</b>   | <b>Openness score</b>   |

### 3.3 List of machine-based assessments of databases

Detailed results can be found in the Supplementary Material (spreadsheet): <https://doi.org/10.5281/zenodo.5577964>. The figure below illustrates a summary.

**Fig. 3 Hotmap summarizing findings from machine-based FAIR assessments.** F-1 to R-2 refer to FAIR maturity indicators in Wilkinson 2019. 80 databases have been assessed and sorted according to their compliance with the indicators. 17 points were the best performing databases. The test failed for 6 databases. Two databases were discontinued during the time when preparing this report.



*The European Science Vocabulary (EuroSciVoc)* <https://op.europa.eu/en/web/eu-vocabularies/euroscivoc>  
(2020) *Global Change Master Directory Keywords*. NASA.  
<https://earthdata.nasa.gov/earth-observation-data/find-data/gcmd/gcmd-keywords> (2020)

Global Energy Assessment Writing Team. *Global Energy Assessment - Toward a Sustainable Future*. International Institute for Applied Systems Analysis, Vienna, Austria and Cambridge University Press, Cambridge, UK and New York, NY, USA. (2012) ISBN: 9780521182935

Bahim, C., Dekkers, M., & Wyns B. Results of an Analysis of Existing FAIR assessment tools. Research Data Alliance. (2019) DOI: <https://doi.org/10.15497/RDA00035>

Wilkinson, M.D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). DOI: <https://doi.org/10.1038/sdata.2016.18>

Wilkinson, M.D. *et al.* Evaluating FAIR maturity through a scalable, automated, community-governed framework. *Sci Data* **6**, 174 (2019). DOI: <https://doi.org/10.1038/s41597-019-0>

*Standard International Energy Product Classification (SIEC)*  
<https://unstats.un.org/unsd/classifications/Family/Detail/2007> (2011)