EERA data

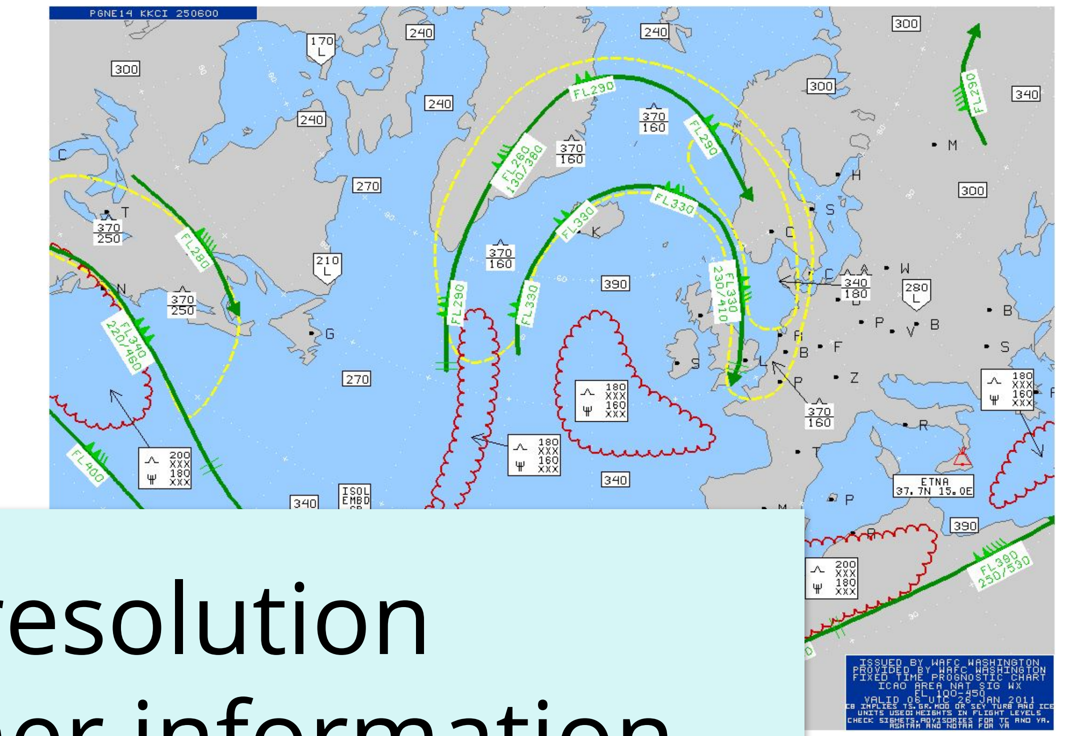
TOWARDS A FAIR AND OPEN DATA ECOSYSTEM IN THE LOW-CARBON ENERGY RESEARCH COMMUNITY

August Wierling
Western Norway University of Applied Sciences

Advancing metadata standards for low carbon energy research

1 Data for low carbon energy transition
2 Metadata and FAIR
3 How to come up with metadata?
4 What has to be done!

Data for the low carbon energy transition



High resolution weather information is needed. On-side conditions matter.

Digital twins are designed and used.

New actors invest in the energy market.

Consumers are metered.

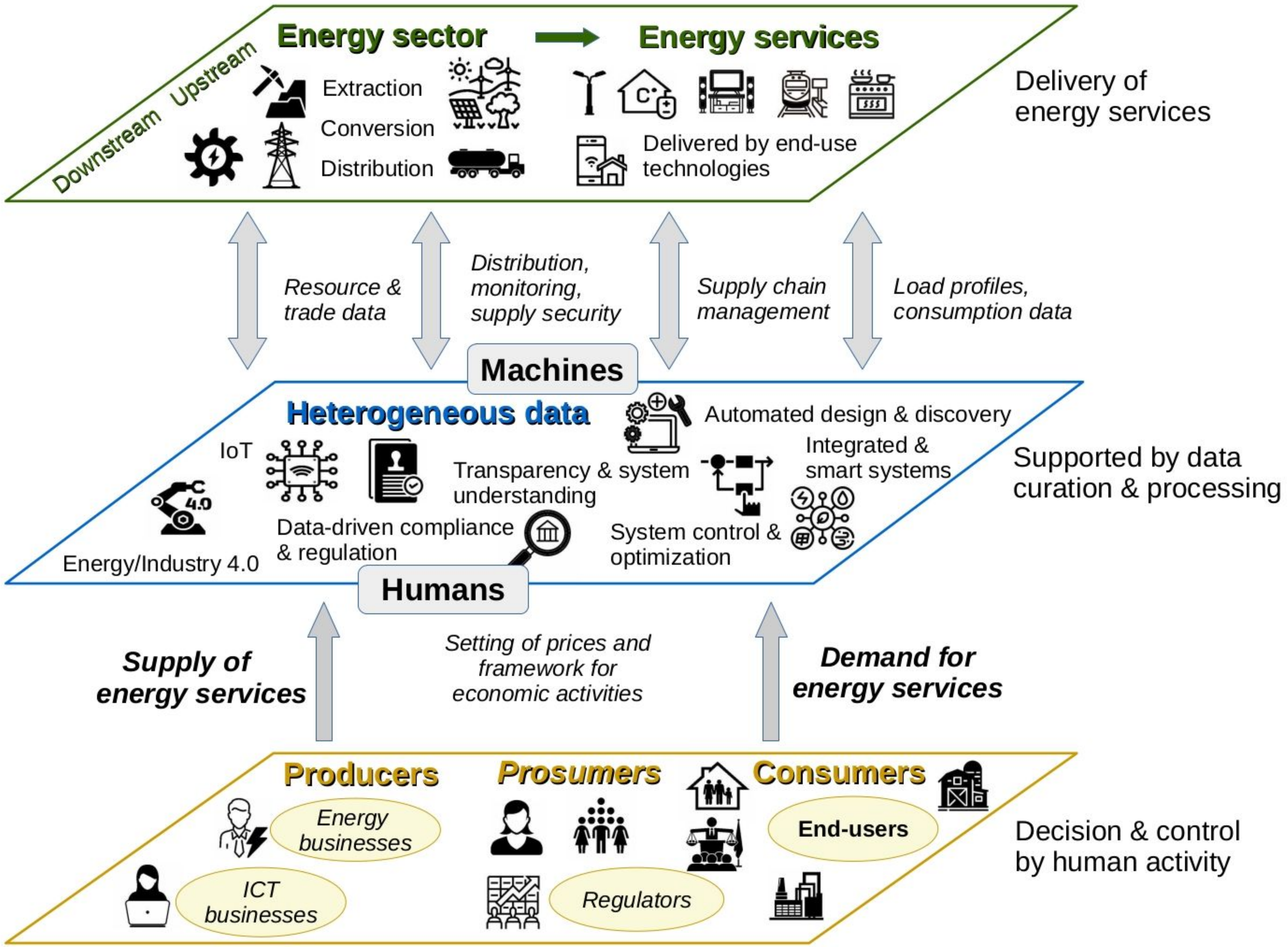
Buildings get energy passports.

Digital reporting & compliance is starting.

Equipment is monitored.

Source of photos: see Appendix

Data for the low carbon energy transition



- Energy systems is a highly complex system of machines and humans.
- Due to the many different components, very heterogeneous data are exchanged.
- Low carbon energy demands and liberalization increase number of actors and complexity of technologies and interactions.

Community workshops building on HLEG 2018

First Year *Define*	Second Year *Implement*	Third Year *Embed & Sustain*
Concepts	Ecosystem, culture, skills	Incentives
Workshop I *FAIR & Open Energy Data* June 2020. Community building.	**Workshop III** *FAIRification put into practice* Characterization of data and development of workflows.	**Workshop V** *Sustainable models* From licensing to business m.
Workshop II *Metadata concept* 30.11.-7.12. 2020.	**Workshop IV** *Supporting technologies* Fair & open energy data infrastructure.	**Workshop VI** *New trends in open science & steps beyond EERAdata*

HLEG 2018 recommendations. https://ec.europa.eu/info/sites/info/files/turning_fair_into_reality_1.pdf

What are metadata?

Metadata puts data into context ….

Classification of metadata:

- Administrative metadata: who collected, when collected, where etc.
- **Descriptive metadata: what is described by the data**
- Structural metadata: organization of data, file formats
- Reference metadata
- Statistical metadata
- Preservation/provenance metadata

Metadata serve different purposes (and are key) to make data FAIR !

Metadata and FAIR principles

F1. **(meta)data** are assigned a globally unique and persistent identifier;

F2. data are described with **rich metadata**;

F3. **metadata** clearly and explicitly include the identifier of the data it describes;

F4. **(meta)data** are registered or indexed in a searchable resource;

A1. **(meta)data** are retrievable by their identifier using a standardized communications protocol;

A1.1 the protocol is open, free, and universally implementable;

A1.2. the protocol allows for an authentication and authorization procedure, where necessary;

A2. **metadata** are accessible, even when the data are no longer available;

I1. **(meta)data** use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2. **(meta)data** use vocabularies that follow FAIR principles;

I3. **(meta)data** include qualified references to other (meta)data;

R1. **(meta)data** are richly described with a plurality of accurate and relevant attributes;

R1.1. **(meta)data** are released with a clear and accessible data usage license;

R1.2. **(meta)data** are associated with detailed provenance;

R1.3. **(meta)data** meet domain-relevant community standards;

Metadata for machines

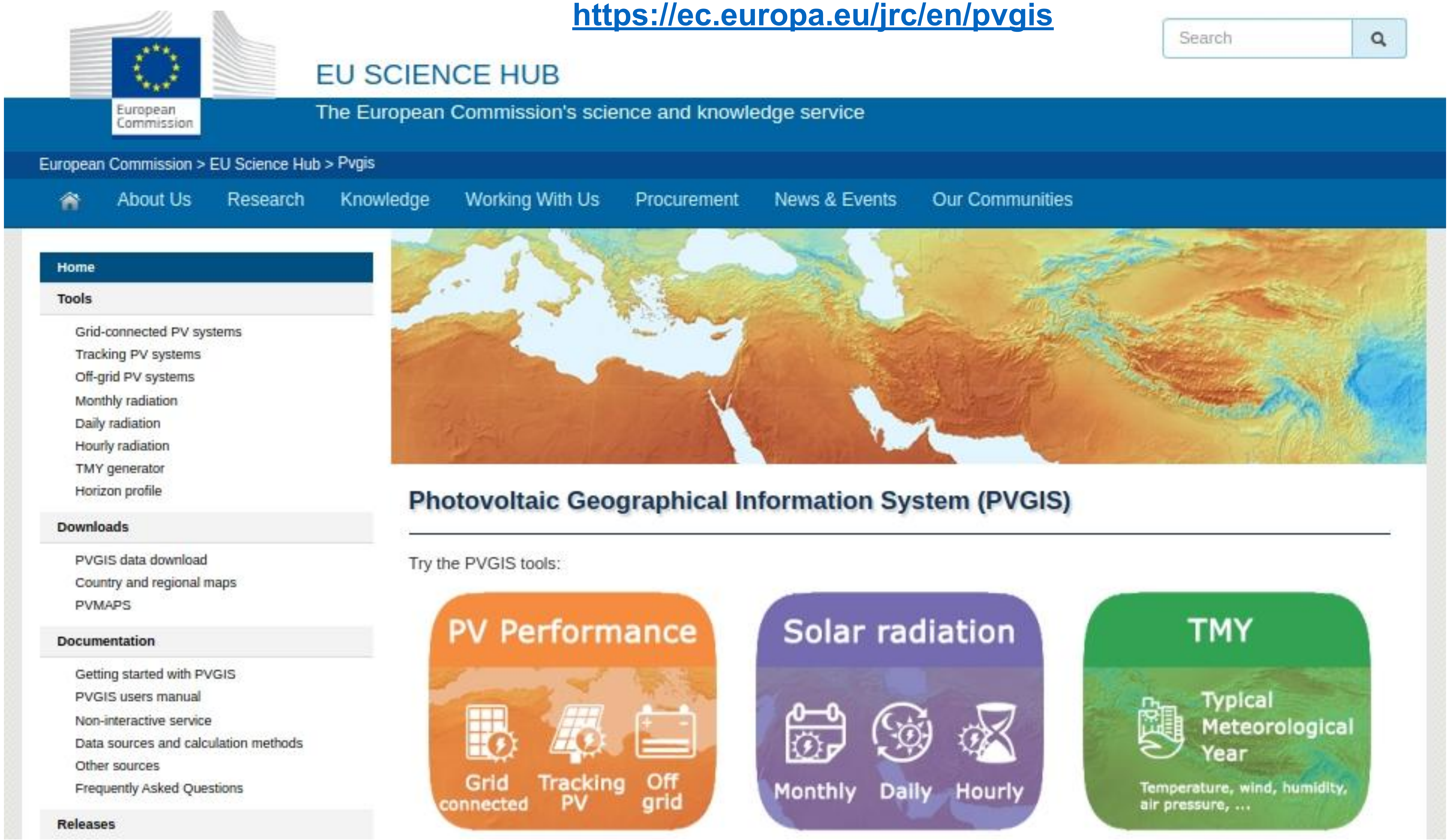
"FAIR must work for humans and for machines: unlocking the potential of analysis and data integration at scale and across a distributed, federated infrastructure is one of the key benefits of making FAIR a reality."
(Turning FAIR into reality)

"When the resource is FAIR, 'machines know what it means'."
(Barend Mons, president CODATA)

see also:
https://www.go-fair.org/resources/go-fair-workshop-series/metadata-for-machines-workshops/

Databases for low carbon energy and metadata

https://ec.europa.eu/jrc/en/pvgis

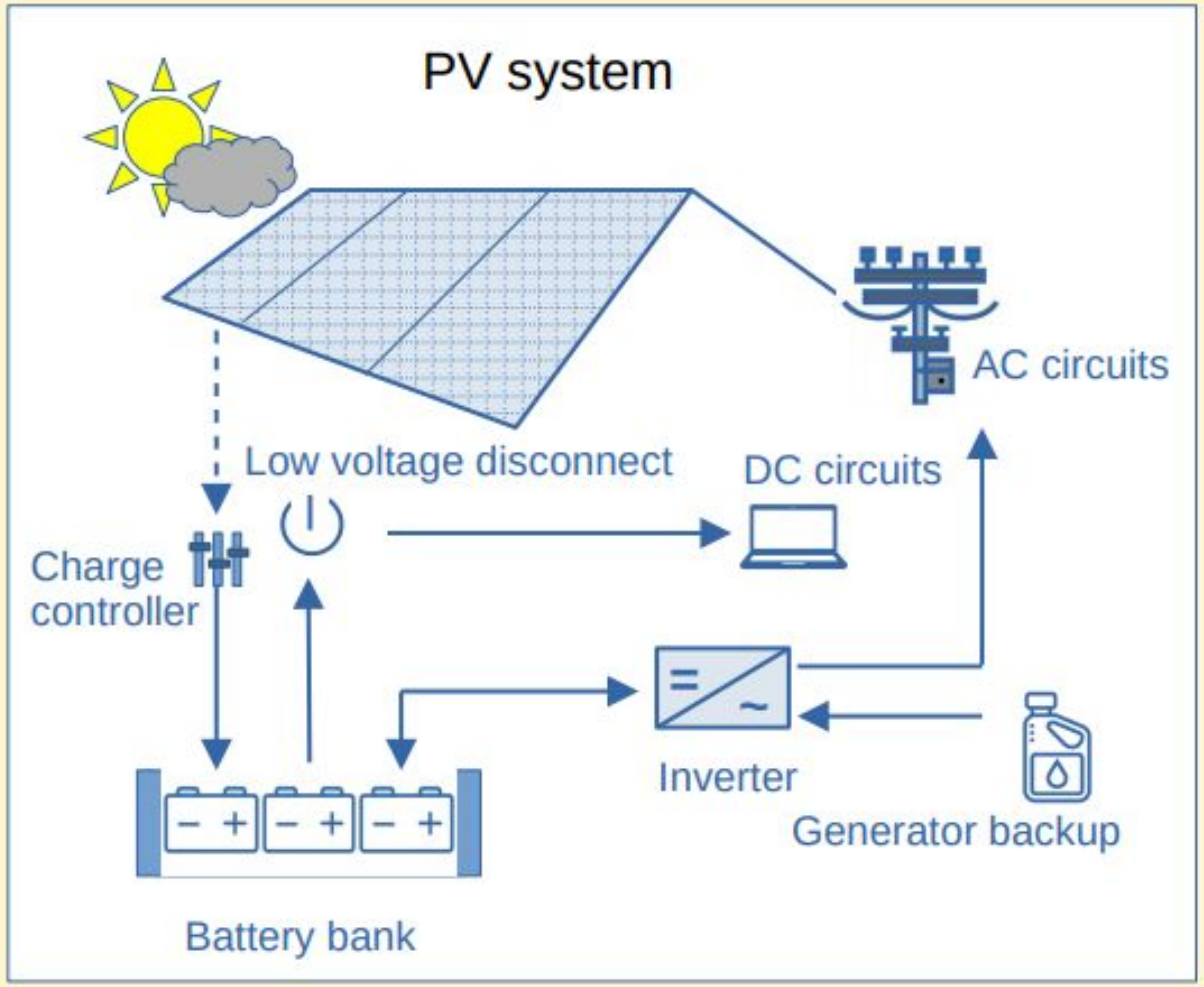


PV data system - an example of a research workflow

PV system monitoring

Input data

Weather
Location
Type of module
Orientation of roof
...



Output data

Electricity to grid
Solar irradiance
Efficiency
Use profiles
...

PVGIS data output: example for information in csv file

Input section

Latitude (decimal degrees):	51.95
Longitude (decimal degrees):	14.711
Radiation database:	PVGIS-SARAH
Nominal power of the PV system (c-Si) (kWp):	1
System losses(%):	14
Fixed slope of modules (deg.):	35
Orientation (azimuth) of modules (deg.):	0

Output section

Fixed angle

Month	E_d	E_m	H(i)_d	H(i)_m	SD_m
1	0.98	30.37	1.14	35.33	6.54
2	1.69	47.38	1.95	54.47	16.48
3	2.73	84.55	3.25	100.84	18.18
4	4.06	121.83	5.01	150.23	21.45
5	4.13	127.89	5.22	161.83	19.78
6	4.21	126.22	5.42	162.53	9.68
7	4.15	128.61	5.42	168.15	16.14
8	3.93	121.68	5.05	156.57	14.51
9	3.44	103.09	4.3	128.97	17
10	2.31	71.5	2.8	86.71	16.17
11	1.26	37.85	1.5	45.01	13.49
12	0.95	29.42	1.11	34.37	6.23
Year	2.82	85.87	3.52	107.08	5.52

	AOI loss (%)	Spectral effects (%)	Temperature and low irradiance loss (%)	Combined loss (%)
Fixed angle:	-3.05	1.82	-5.54	-19.82

Variable description

E_d: Average daily energy production from the given system (kWh/d)
E_m: Average monthly energy production from the given system (kWh/mo)
H(i)_d: Average daily sum of global irradiation per square meter received by the modules of the given system (kWh/m2/d)
H(i)_m: Average monthly sum of global irradiation per square meter received by the modules of the given system (kWh/m2/mo)
SD_m: Standard deviation of the monthly energy production due to year-to-year variation (kWh)

PVGIS data output: example for information in json-ld file

Data


```
{"inputs": {"location": {"latitude": 51.95, "longitude": 14.71, "elevation":
44.0}, "meteo_data": {"radiation_db": "PVGIS-SARAH", "meteo_db": "ERA-Interim",
"year_min": 2005, "year_max": 2016, "use_horizon": true, "horizon_db": "DEM-
calculated"}, "mounting_system": {"fixed": {"slope": {"value": 35, "optimal":
false}, "azimuth": {"value": 0, "optimal": false}, "type": "free-standing"}},
"pv_module": {"technology": "c-Si", "peak_power": 1.0, "system_loss": 14.0},
"economic_data": {"system_cost": null, "interest": null, "lifetime": null}},
"outputs": {"monthly": {"fixed": [{"month": 1, "E_d": 0.98, "E_m": 30.37,
"H(i)_d": 1.14, "H(i)_m": 35.33, "SD_m": 6.54}, {"month": 2, "E_d": 1.69, "E_m":
---------------------------------------------------------
```


Metadata


```
"meta": {"inputs": {"location": {"description": "Selected location",
"variables": {"latitude": {"description": "Latitude", "units": "decimal
degree"}, "longitude": {"description": "Longitude", "units": "decimal degree"},
"elevation": {"description": "Elevation", "units": "m"}}}, "meteo_data":
---------------------------------------------------------
"outputs": {"monthly": {"type": "time series", "timestamp": "monthly averages",
"variables": {"E_d": {"description": "Average daily energy production from the
given system", "units": "kWh/d"}, "E_m": {"description": "Average monthly energy
production from the given system", "units": "kWh/mo"}, "H(i)_d": {"description":
"Average daily sum of global irradiation per square meter received by the
modules of the given system", "units": "kWh/m2/d"}, "H(i)_m": {"description":
---------------------------------------------------------
"%"}, "l_spec": {"description": "Spectral loss", "units": "%"}, "l_tg":
{"description": "Temperature and irradiance loss", "units": "%"}, "l_total":
{"description": "Total loss", "units": "%"}}}}}}
```

PVGIS metadata - pros and cons

Pros:

- contains all relevant scientific context
- uses json-ld as format
- contains description on variables and units
- contains human description on variables
- ...

Cons:

- Variable names are not linked to a unique standards
- Units are not linked to unique standards
- No links to input databases
- Some administrative metadata is missing and could be implemented using Dublin core
- ...

Asking the practitioners: Questionnaire on metadata etc.

GUIDELINES FOR THE MONITORING OF PV SYSTEMS

Summary of questionnaire

COST ACTION PEARL PV: CA16235

Working Group 1, Task 1.1

Basant Raj Paudyal, Anne Gerd Imenes (University of Agder)

Date: 9 May 2019

Quality Control: In terms of quality control, most respondents require that instrumentation/sensor accuracy should be specified, data quality control processes should be in place, and a minimum requirement for data accuracy should be set.

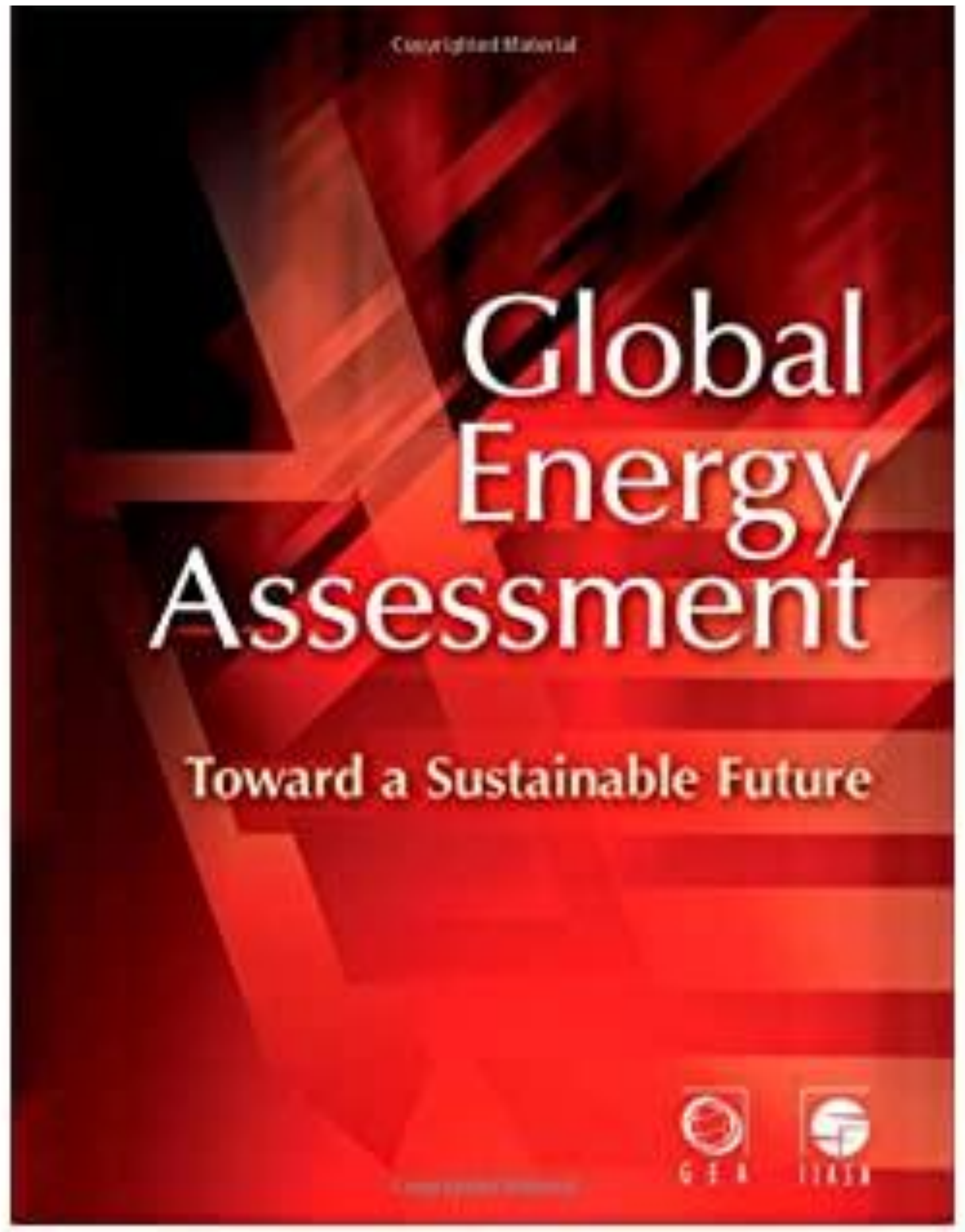
File Format: Most respondents prefer "csv" file format, with missing datapoints reported as "NaN" and using 1-second time stamp resolution (YYYY-MM-DDTHH:MM:SS). Two respondents require sub-second resolution data (SS.ssss), indicating special applications such as grid interaction analysis. Reporting of missing datapoints (e.g. number or time period of datapoints missing) is desired, not required.

Metadata - PV system and components: Generally, respondents agree that a relatively large amount of metadata should be made available to adequately document the system and technology used. Required information is: Site name and GPS coordinates, type of PV installation (fixed, tracking, BAPV/BIPV, roof/façade, etc), type of PV module technology, PV module characteristics (Impp, Vmpp, Isc, Voc, Pmpp, temperature coefficients), string design (number of modules, number of strings, connections to each inverter), shading (no shading), type of inverter technology (central, string, micro, transformerless, etc), inverter specifications (AC/DC power, frequency, rated efficiency, number of phases, reactive power, number of independent MPP's, total number of inverters, communication protocols). When applicable, the following BOS-components should be specified: Battery-system and own developed hardware/software. Information about other parts of the BOS is desired, not required.

taken from:
COST ACTION PEARL PV: CA16235, Working Group 1, Task 1.1, Authors: Basant Raj Paudyal, Anne Gerd Imenes (University of Agder)

Multidisciplinarity and interdisciplinarity in energy science

Global energy assessment, index on non-engineering solar energy issues:



- case study
- economics
- installed capacity
- market developments
- occupational health
- potential
- sustainability
- technology roadmap
- transitions management
- buildings
- forecasting
- health risks
- historical trends
- innovation policies

However:

Not every detail is needed in interdisciplinary studies.

Therefore, clean modularity of concepts & variables are needed.

See: https://iiasa.ac.at/web/home/research/Flagship-Projects/Global-Energy-Assessment/Home-GEA.en.html

Metadata - The babylonian confusion of tongues

Pieter Bruegel the Elder - The Tower of Babel. Source: Wikimedia Commons

Towards common understanding of concepts in energy science

Existing taxonomies/ontologies in energy science

- PV-TONS ontology (literature, Oxford solar house)
- IRPWind taxonomy (scientific literature, expert elicitation)
- SEMANCO - Semantic tools for Carbon Reduction in Urban Planning (standards, use cases, activity descriptions)
- energy accounting of statistical offices (e.g. Eurostat)
- Energistics: Energy Industry Profile of ISO 19115-1:2014
- Open Energy Ontology (community initiative)
- ...

However:

Many unrelated ontologies and islands of codified knowledge.

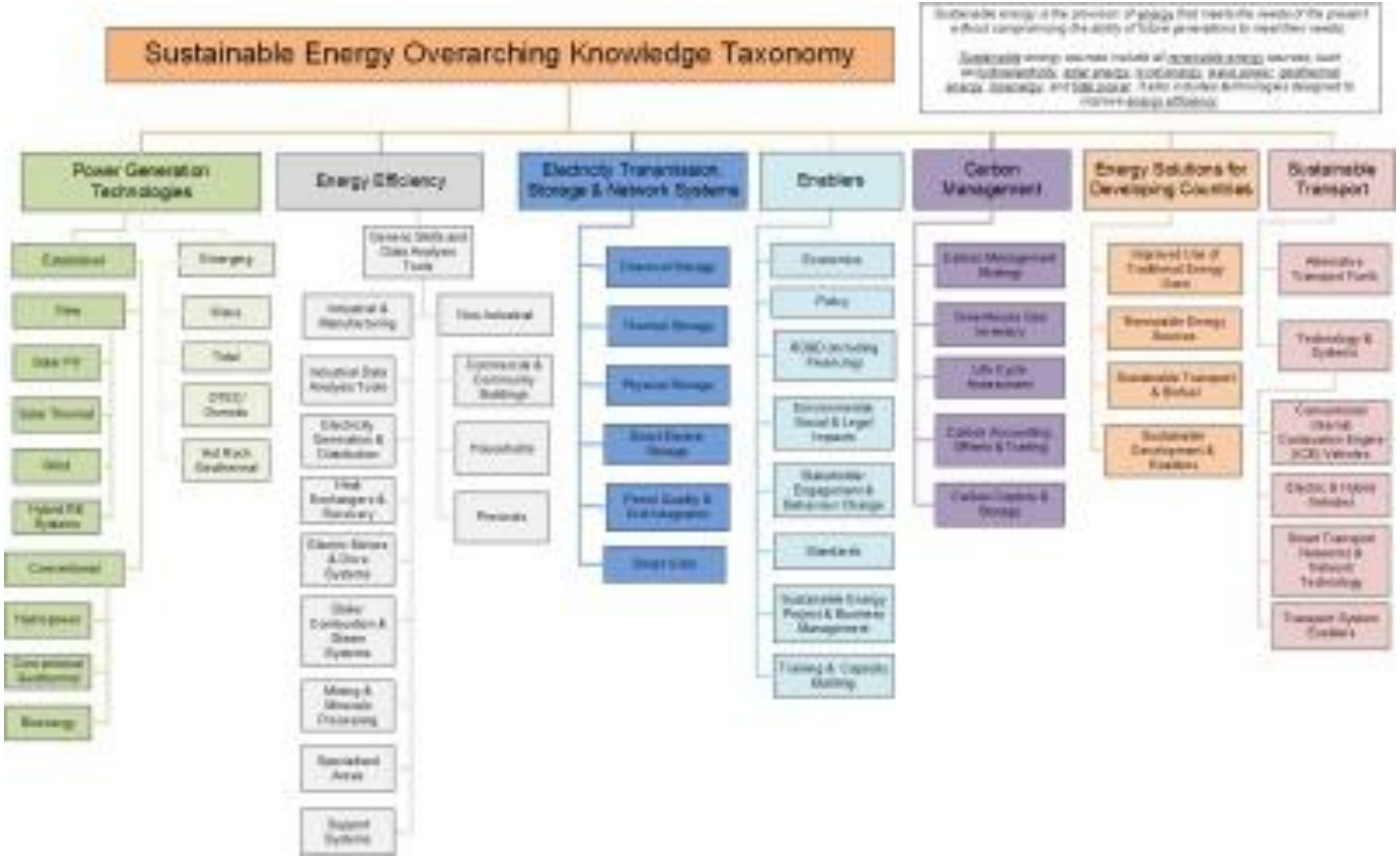
Alignment and linking is needed.

Support for navigation & visualization required.

Other options for (semi-)automatized sourcing of ontologies

- Textbooks and exam questions
- Wikipedia articles
- Models and simulation activities on aspects of the energy system

Example for a top-level taxonomy



https://www.murdoch.edu.au/School-of-Engineering-and-Information-Technology/Research/Renewing-the-Sustainable-Energy-Curriculum/Project-Reports-and-Papers/Knowledge-and-Generic-Attributes-Taxonomies/

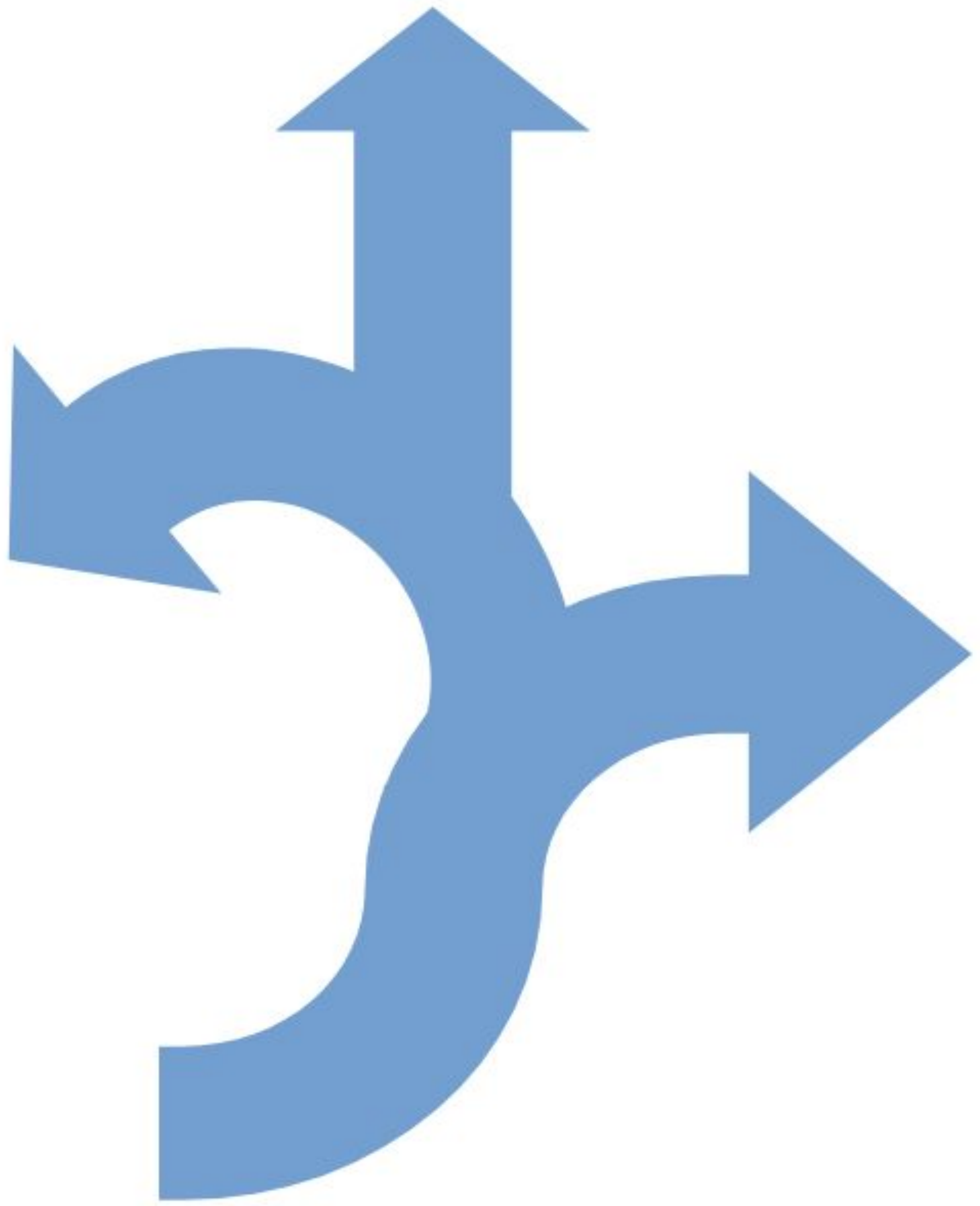
Bringing things together?

Convergence to a common metadata model

vs.

Interoperation among many metadata models

taken from: RDA Metadata Principles and their Use



What has to be done (@WS2 and beyond)!

- Compile existing metadata frameworks in the broader research community and within low carbon energy research.
- Classify, align, standardize existing metadata. Keep it flexible.
- Fill in the gaps.
- Implement a framework to publish/make accesible the standard.
- Issue recommendations: metadata suggestions and check-lists.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 883823.

Credits & sources

https://commons.wikimedia.org/wiki/File:Aviation_weather_forecast_chart.png
https://commons.wikimedia.org/wiki/File:Halberton,_electricity_transformer_station_-_geograph.org.uk_-_151077.jpg
https://en.wikipedia.org/wiki/File:Windmills_D1-D4_(Thornton_Bank).jpg
https://commons.wikimedia.org/wiki/File:Libro-sistemas-scada-3ra-ed-mecatronica-automatizacion-22680-MLM20234441553_012015-F.jpg
https://commons.wikimedia.org/wiki/File:Intelligenter_zaehler-_Smart_meter.jpg
https://www.nps.gov/goga/learn/management/energy-efficiency.htm

EERA data

TOWARDS A FAIR AND OPEN DATA ECOSYSTEM IN THE LOW-CARBON ENERGY RESEARCH COMMUNITY