DTU

FIND THE DATA

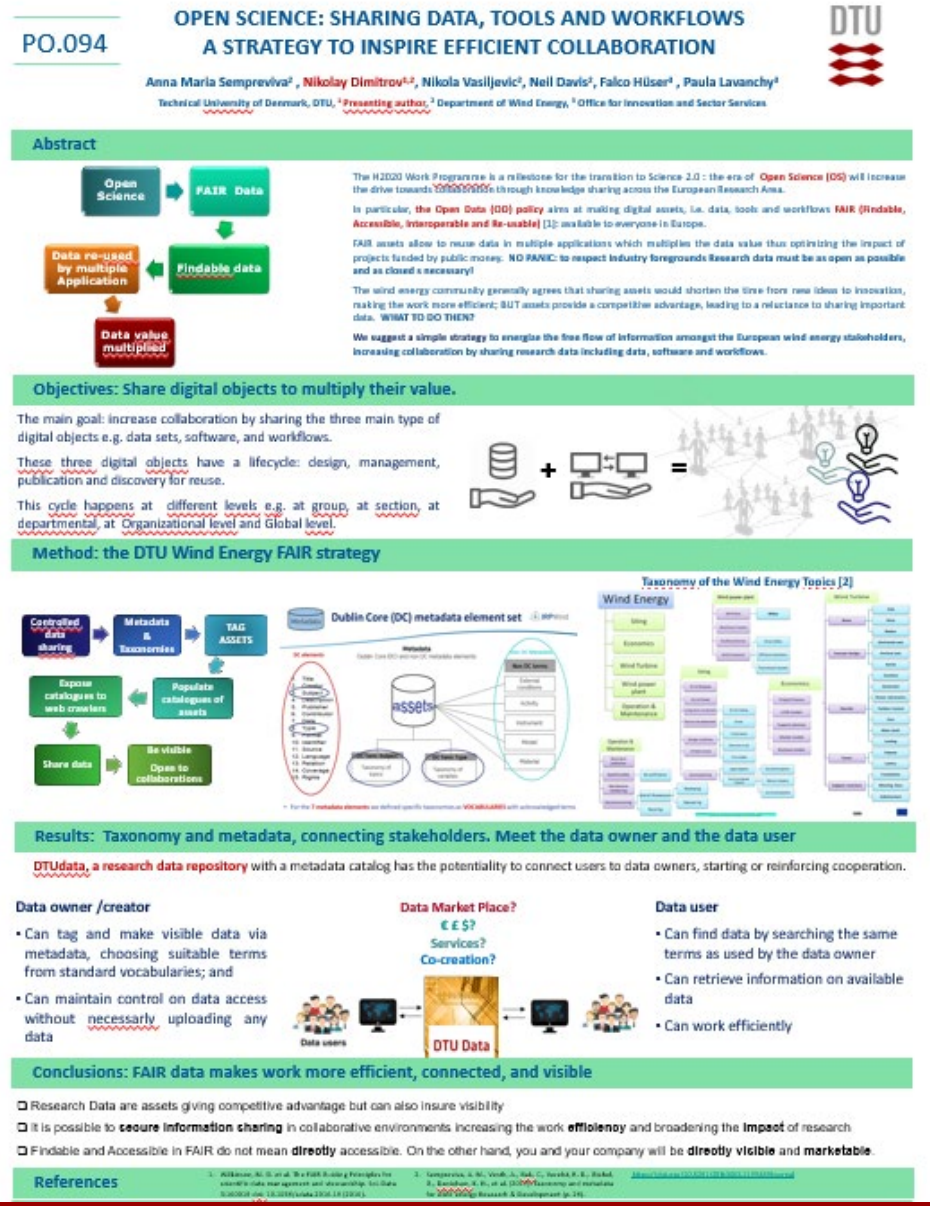
Find the data: Metadata and taxonomies for FAIR data shareuse in Wind Energy.

Anna Maria Sempreviva & Many More

Technical University of Denmark
DTU Wind Energy
Coordinator of metadata and taxonomy creation

Content

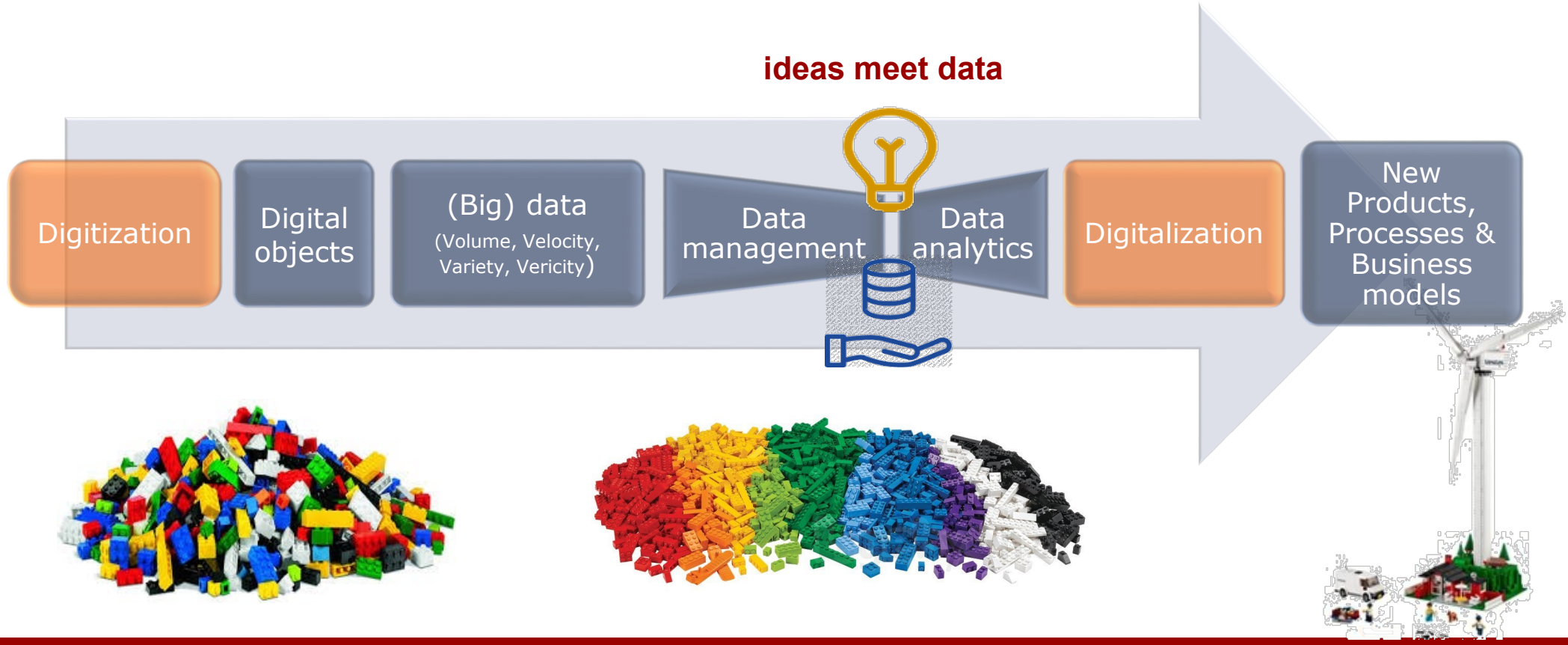
- Context:
  - The digital transformation of the Wind Energy sector,
  - The stakeholders: data owner, data user, the funding body
  - From Open to FAIR, the culture of sharing
- Metadata and Taxonomies for the Wind Energy what and why?
- Data Portal and data registry
- The FAIRification process
- Conclusive remarks



Context: Digital transformation of Wind Energy sector

Grand challenge: 2050 be the leading renewable

There is a bottleneck in whatever workflow if data is not properly <u>engineered and</u> managed



Context: the stakeholders

The data owner (Most Industry) demands:

- Control on data as assets
- Recognition
- Protection of competitive advantage

But willing to expose his data

The data user needs:

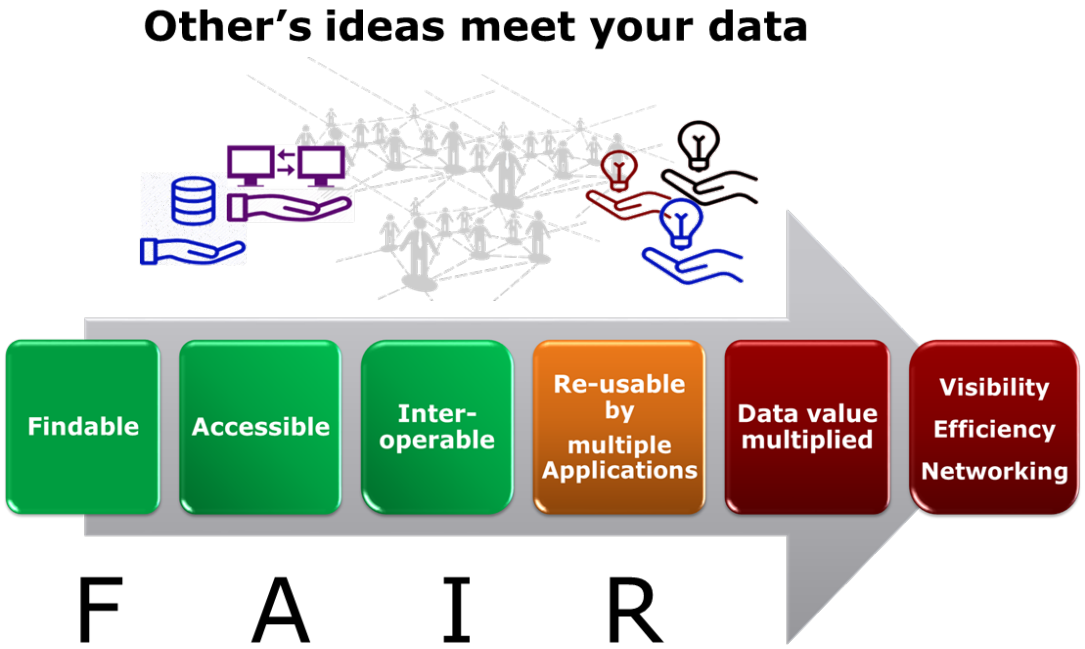
- Improved efficiency
- Money saving

Willing to sign NDAs & reward

Visibility – Efficiency – Networking - Control – Recognition – Reward

Context:The Funding body, H2020
from Open to FAIR

- 2014 H2020: Open Data …. **Panic**!
- 2016 H2020: FAIR Data Principle(s)*…… Milestone:
- **FAIR changes focus: From Available to Findable data**
- Data as open as possible as closed as necessary



ISSUE: How to make data findable but safe?

SOLUTION: Create a searchable data catalog for **distributed** data

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

How it started. A bottom up approach.

European Energy Research Alliance, Joint Programme on Wind Energy,

EERA JPWind Community

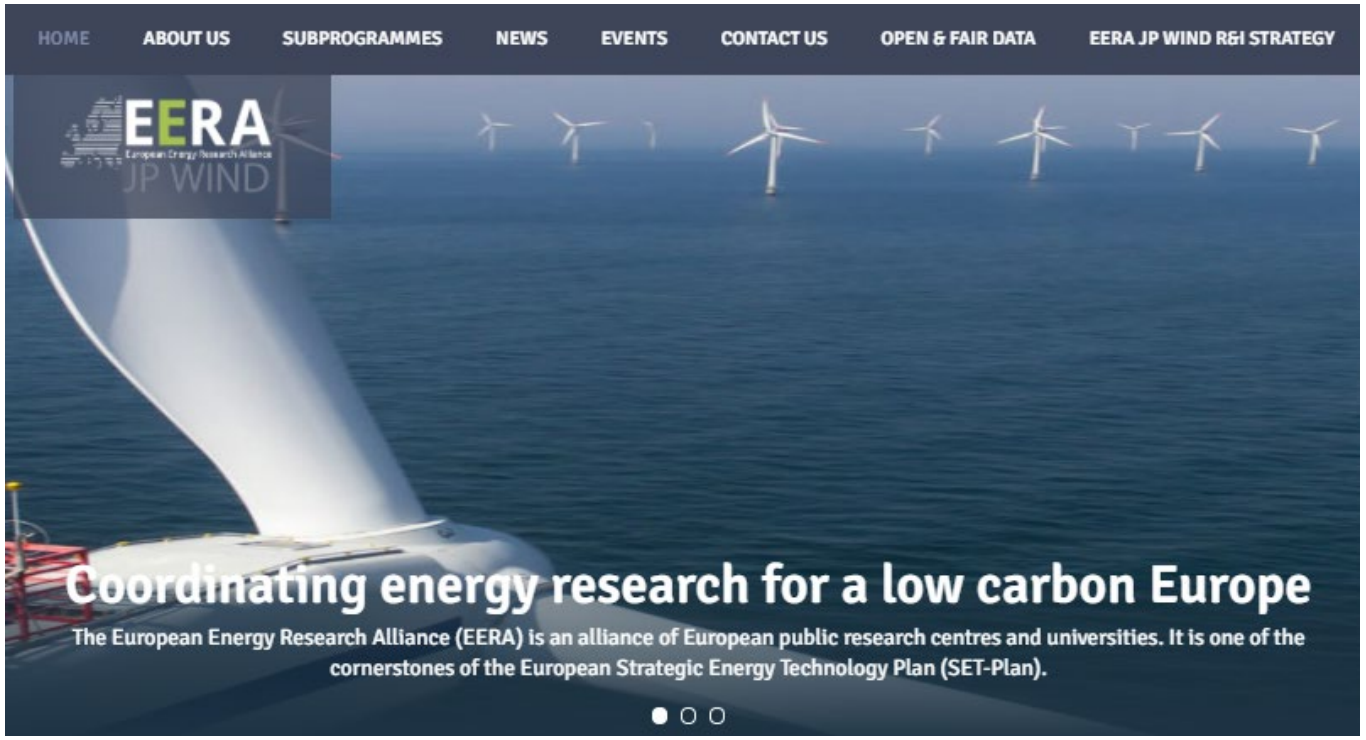
50+ European Research Organizations

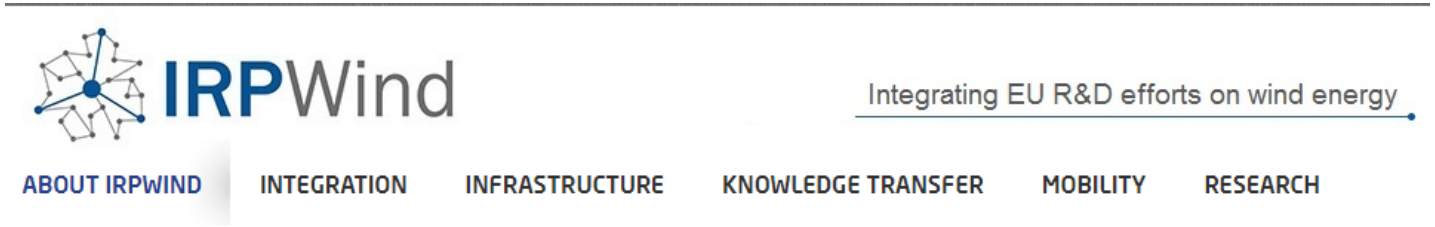
FP7 Integrated Research Programme in Wind Energy IRPWind project

27 EERA JPWind Organization

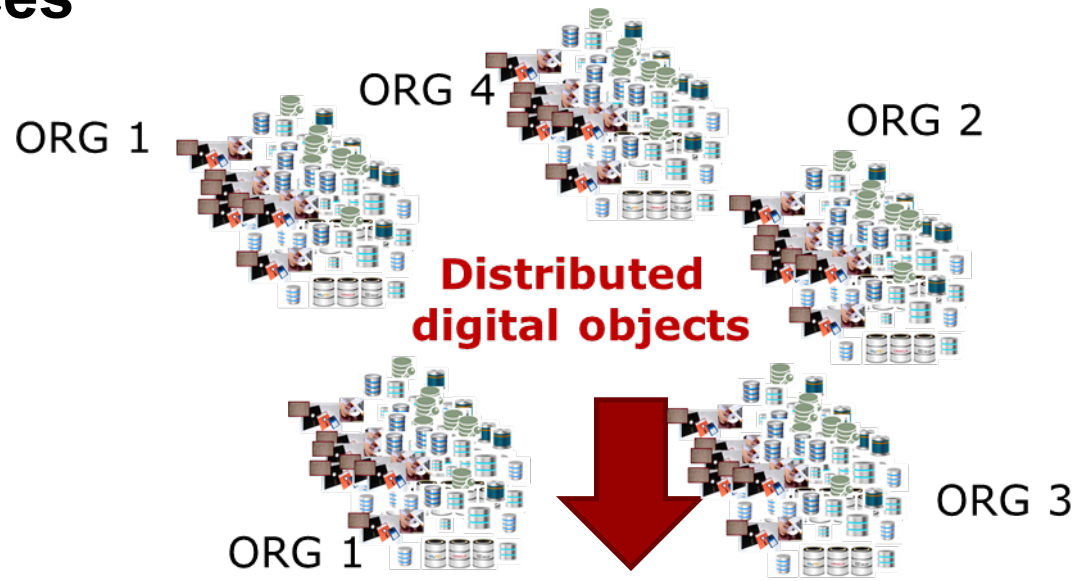
Goal : create a new WP for open data

EERA JP Wind digital object findability



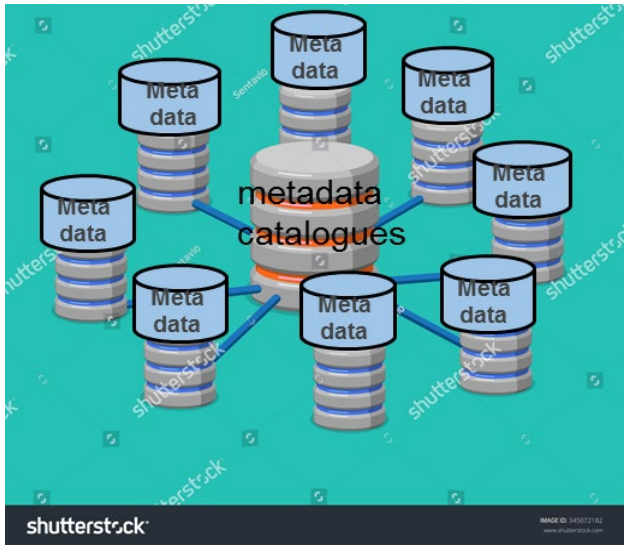


FIND EERA JPWind distributed resources



Issue EERA digital object **findability**

- Datasets are distributed in the "cloud", organized and stored in different ways.
- Datasets often miss **documentation (Metadata)**



Action: we needs 3 ingredients

- Create a list of **metadata**
- Assign **taxonomies** to metadata:controlled vocabularies to tag data
- Design **a data portal as a Virtual Library with a** metadata catalog

Courtsey: Pawel Garcansky, CENER

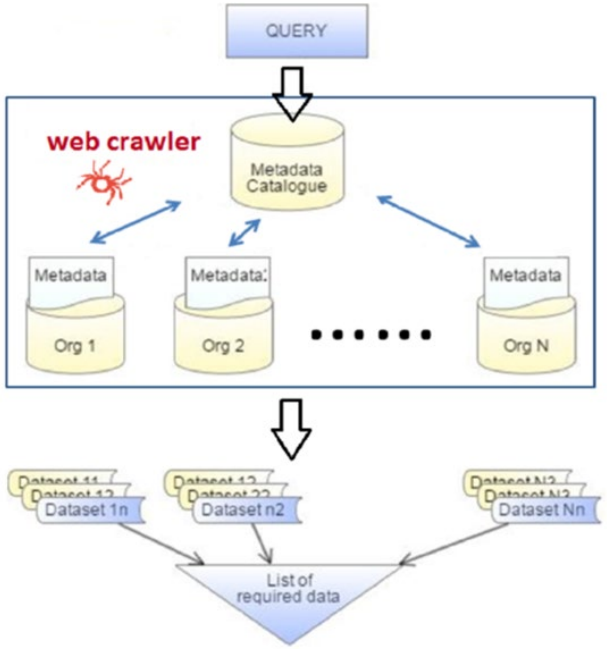
How virtual distributed data catalogs / libraries work?

Search engine

Registry

Storage

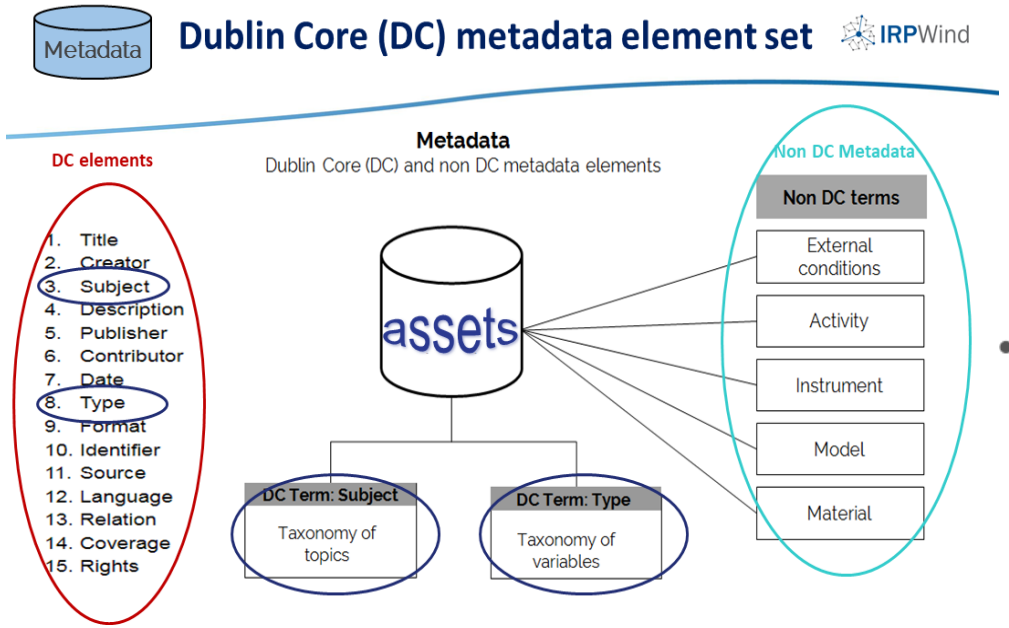
Search Engine



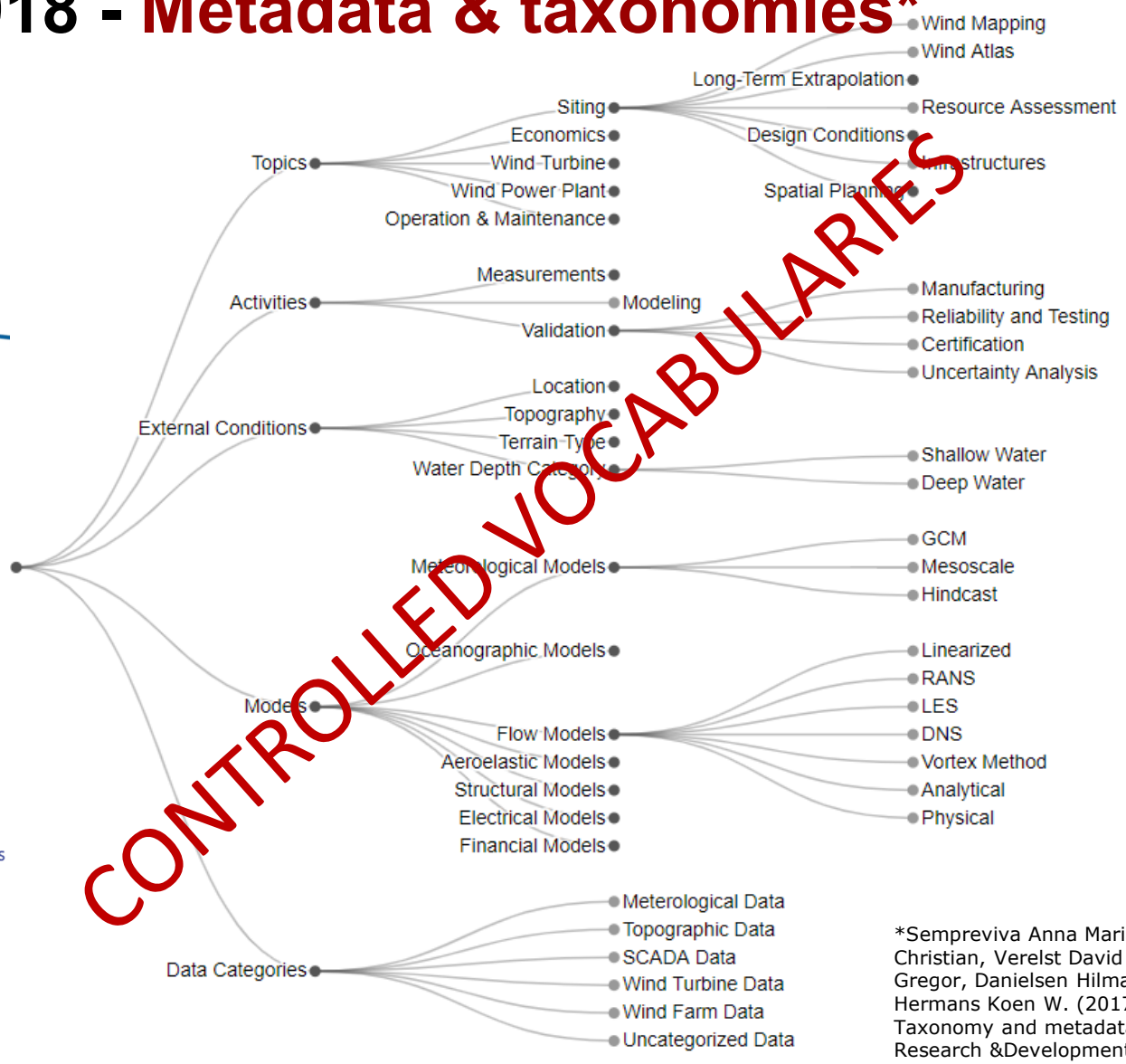
ADVANTAGES

- Data owner has control
- Data are visible without being accessible
- No uploading data and no need of storage and data preservation.
- By applying filters, users can accurately locate needed data

IRPWind Project 2014-2018 - Metadata & taxonomies*



- For the **7 metadata elements** we defined specific taxonomies as **VOCABULARIES** with acknowledged terms



CONTROLLED VOCABULARIES

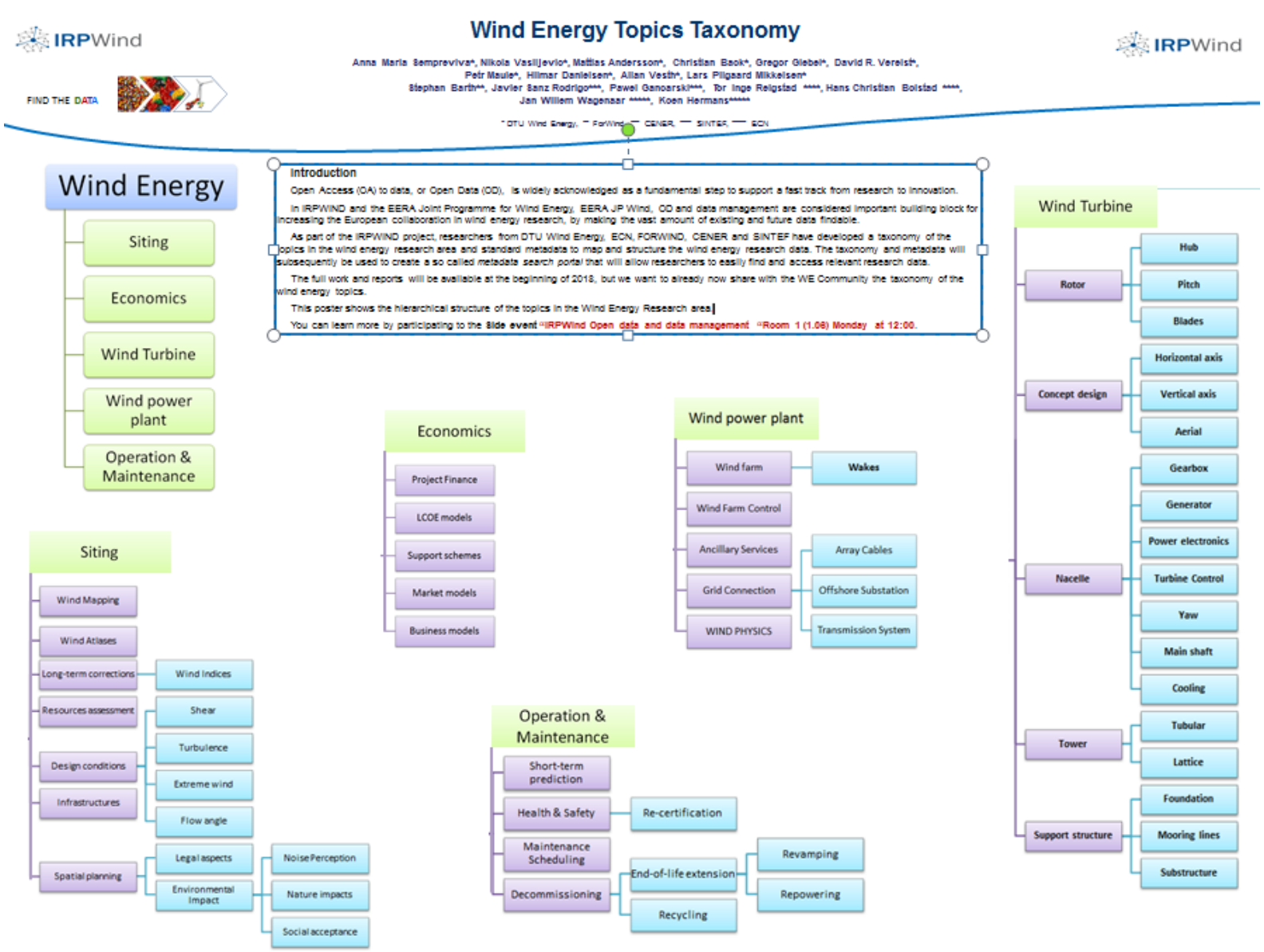
*Sempreviva Anna Maria, Vesth Allan, Bak Christian, Verelst David Robert, Giebel Gregor, Danielsen Hilmar Kjartansson, … Hermans Koen W. (2017, December 12). Taxonomy and metadata for wind energy Research &Development. Zenodo. http://doi.org/10.5281/zenodo.1199489

Creating a taxonomy

Expert elicitation

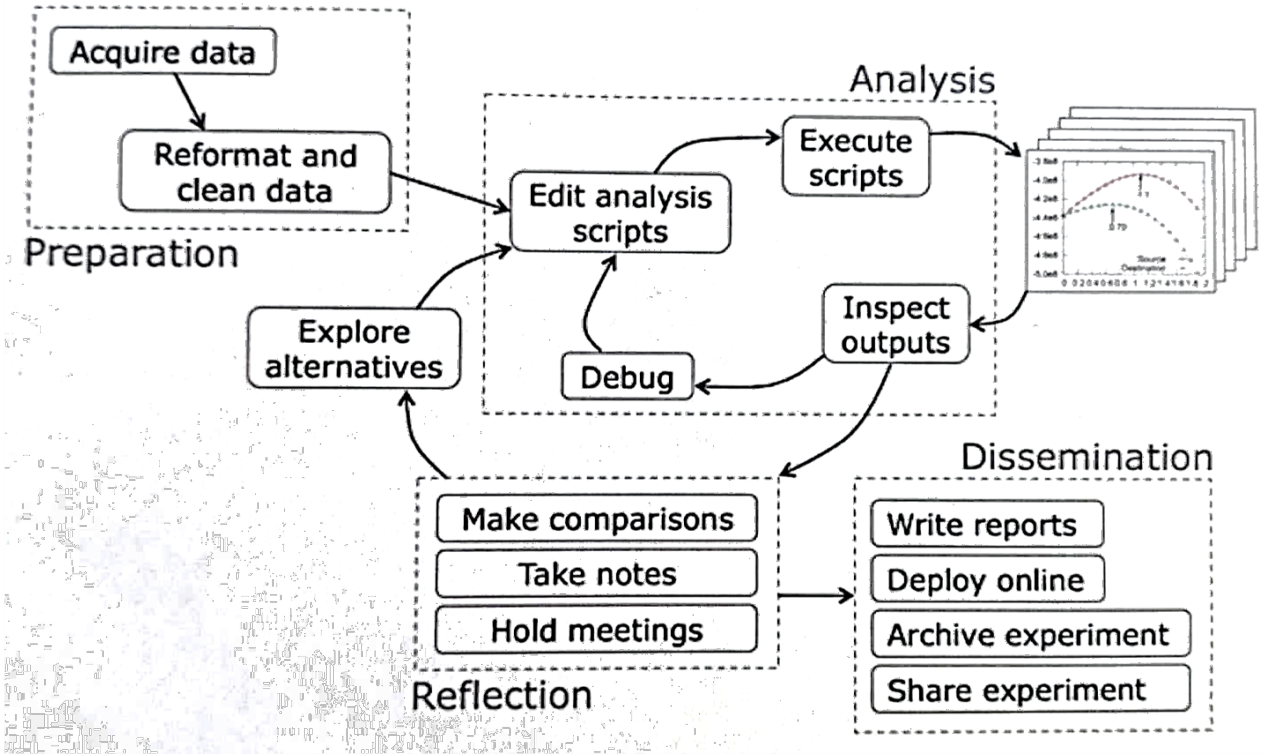
- 8 Experts from DTU were engaged in the task of creating taxonomy for WE topics

- Results were submitted and discussed to the IRPWind core group from Forwind, CENER, ECN, SINTEF



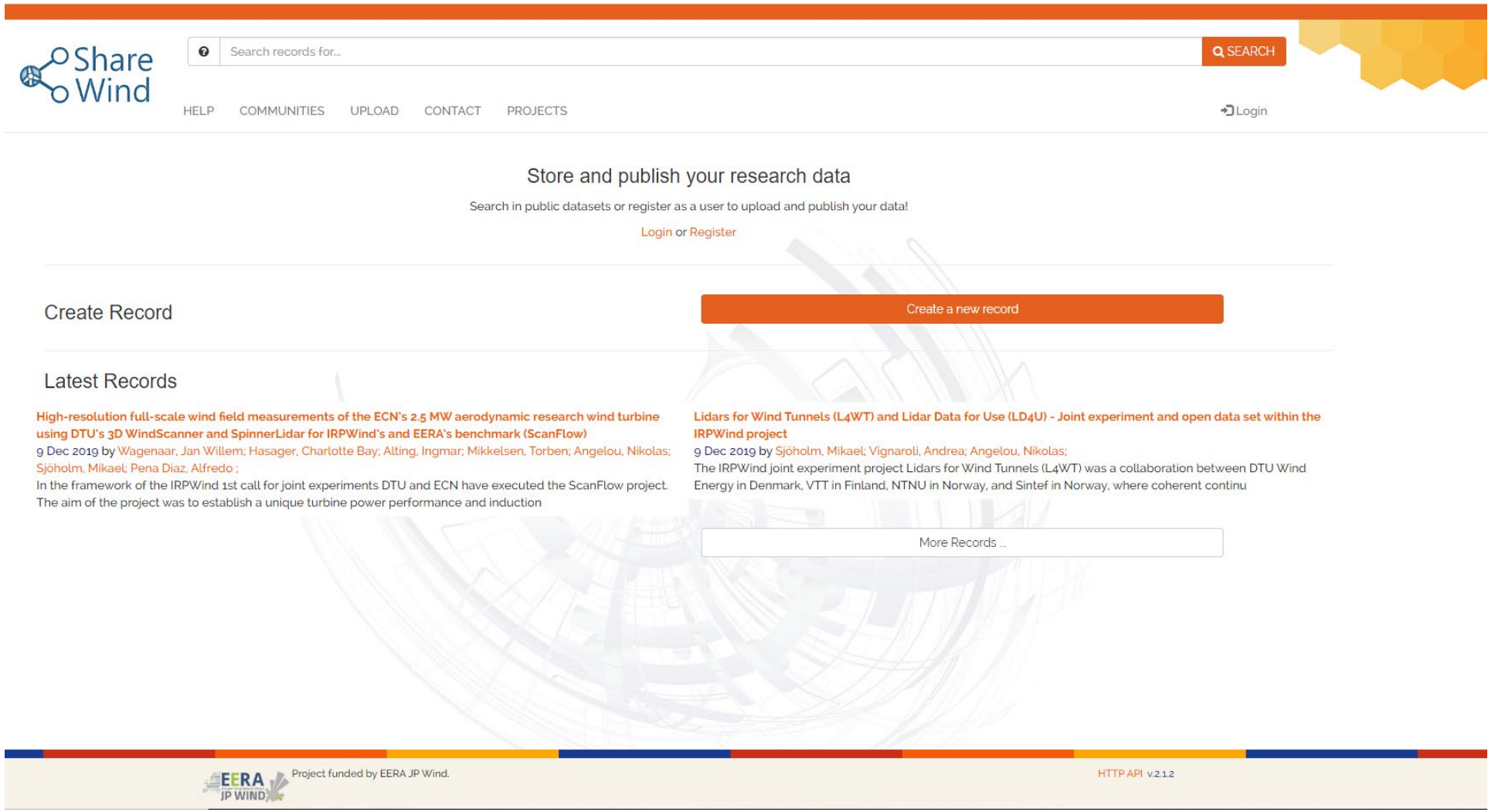
Metadata &Taxonomy: cataloguing workflow components

- Generic work flow can be specialized and automatized using element taxonomies



Courtesy: Pawel Garcansky

IRPWind – EERA JPWind Energy. www.ShareWind.eu
Data registry and metadata catalogue



Results – taxonomies

Conclusive remarks

A web data portal with a data catalog has a two-fold purpose

- **To connect safely data owners to users**
- **Inform on the availability of shared resources NOT necessary direct access**

Data owner /creator

- Can make visible data via metadata
  - without uploading any data, and
  - maintain control on data access
  - Multiply value of data

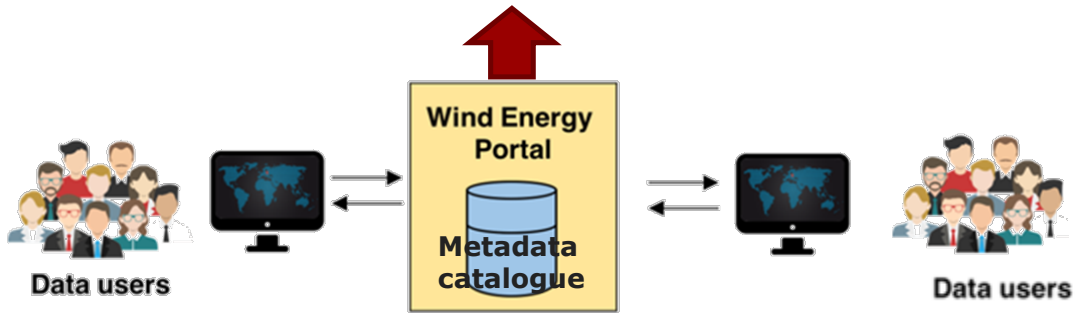
Data= asset

Market Place?

€ £ $?

Services?

Co-creation?



Data user

- Can find data accurately by searching the same terms used by the data owner
- Can retrieve information on available data
- Efficient path to results

Current work: taping into machine-actionability

Nikola Vasiljević

- Made taxonomies machine-actionable using FAIR Data Collective tools and workflows e.g., http://purl.org/neat
- Creating and exposing machine-actionable metadata templates and theirs instances https://www.youtube.com/watch?v=60ZRKeUe9D4
- Harvesting machine-actionable metadata for next generation share-wind portal
- Developing share-wind v2.0, https://www.share-wind.com *(in production from Oct 28th 2021)*
- Created rapid M4M workshops, which have been successful beyond wind energy domain

Thank you for your attention!