IEA Wind Energy Task 41: Distributed winds

PNNL is operated by Battelle for the U.S. Department of Energy

WORK PACKAGE 2: DATA CATALOG

Danielle Preziuso
Pacific Northwest National Laboratory

FAIR DATA Principle: The culture of sharing.

Findable **A**ccessible **I**nteroperable and **R**e-usable data

Anna Maria Sempreviva
Technical University of Denmark
DTU Wind Energy

PURPOSE

- Disseminate the sharing culture in wind energy sector
- Inform on the opportunities from adopting the FAIR data principle

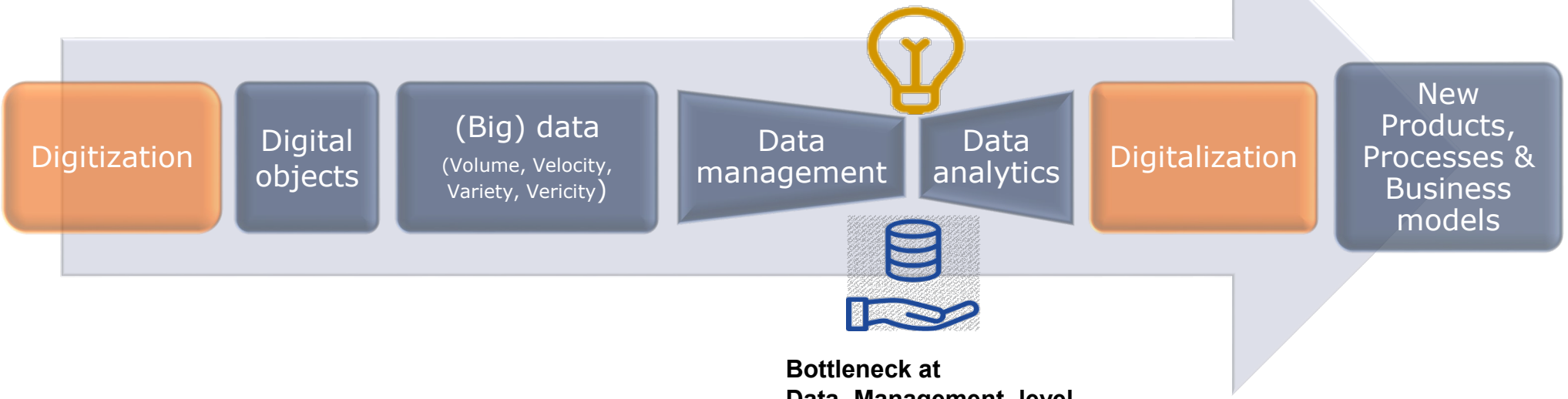
CONTENT

- Context:
  - The digital transformation: the pathway to innovation
  - FAIR, the culture of sharing: other's ideas meet your data
- Data Catalogue: collecting information on data availability on key topics
- Existing data platforms: metadata and taxonomies
- Conclusive remarks

Equivalences of terms in different environment

RESEARCH DATA	
(Academia)	Data, Codes, Workflows
(Research Data Alliance)	Digital Objects
(Industry)	Assets

From digitization to digitalization (innovation)



Bottleneck at Data Management level where, ideally, ideas should meet data.

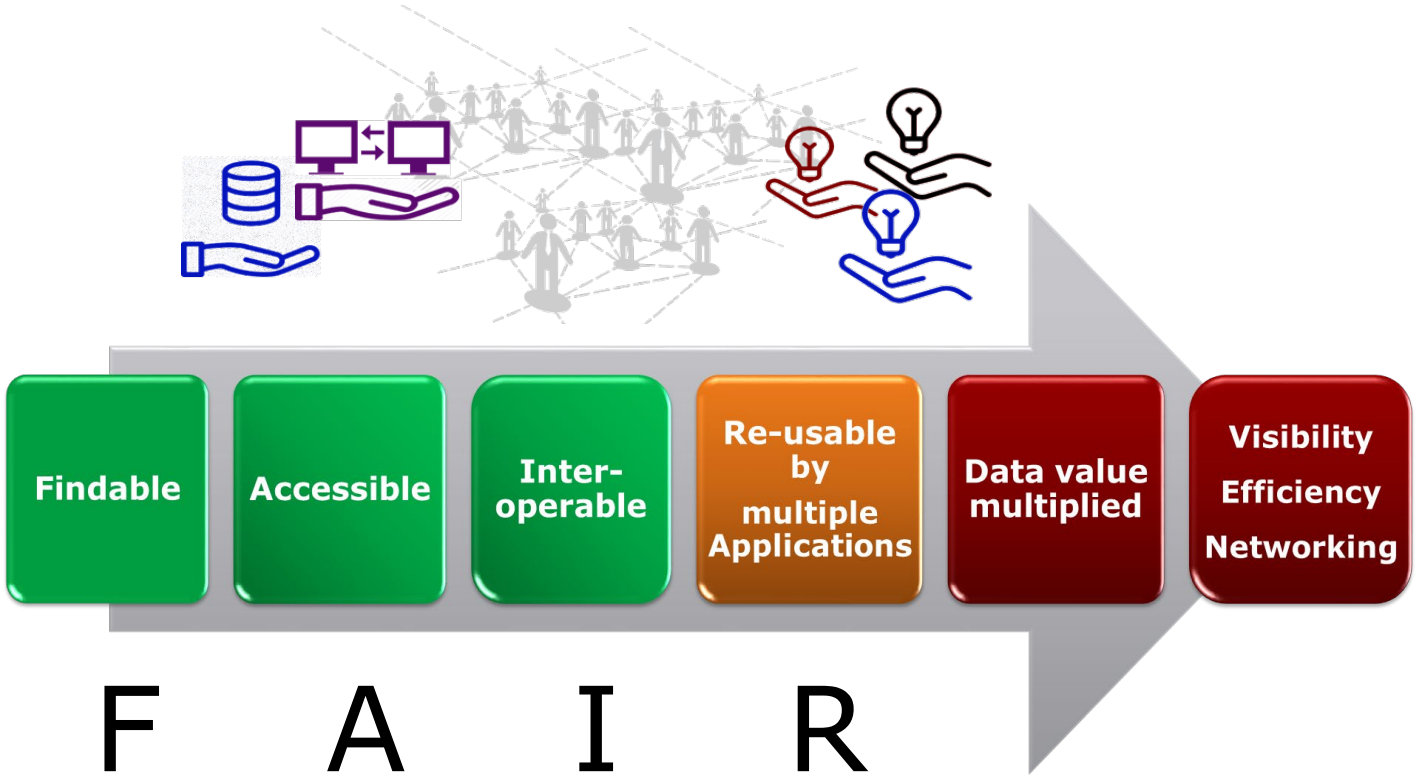
FAIR supports innovation: Find the data

- 2014 H2020 Open Data
- 2016 H2020 FAIR Data Principle changes the focus: **From Available to Findable data**

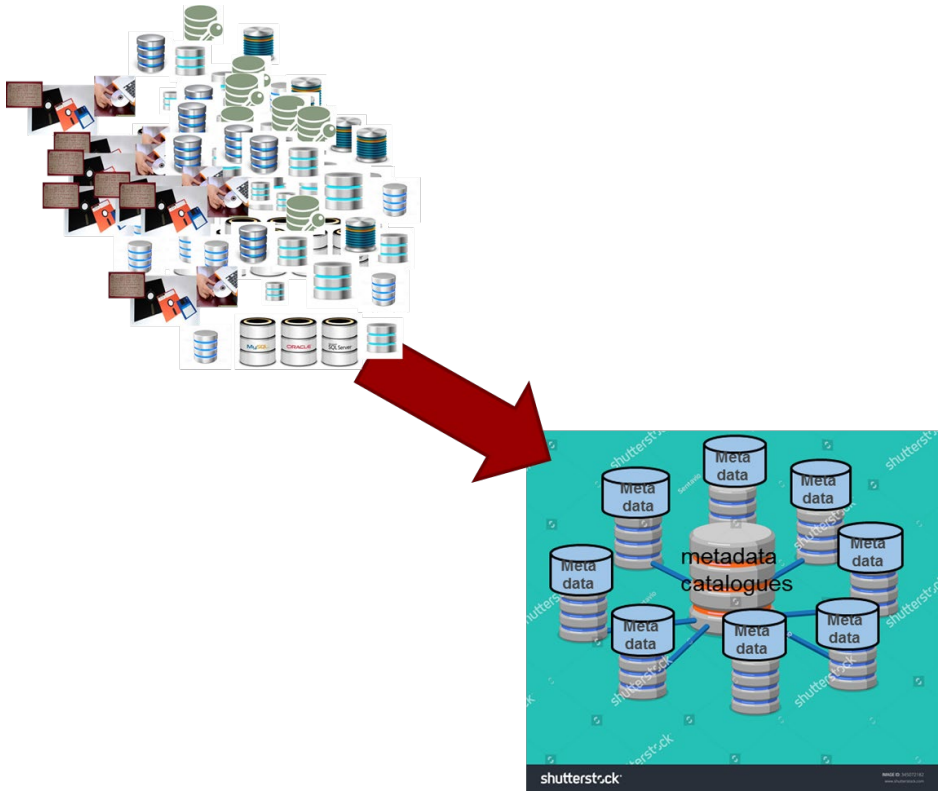
ISSUE: How to make data findable but safe?

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

Other's ideas meet your data



FIND THE DATA



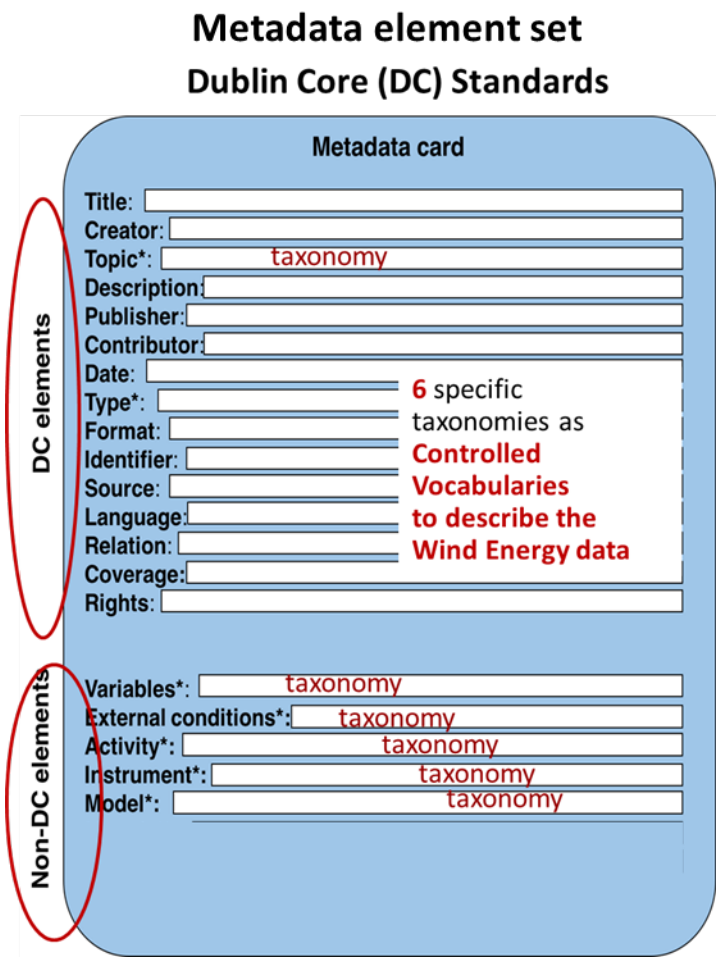
Issue: data findability

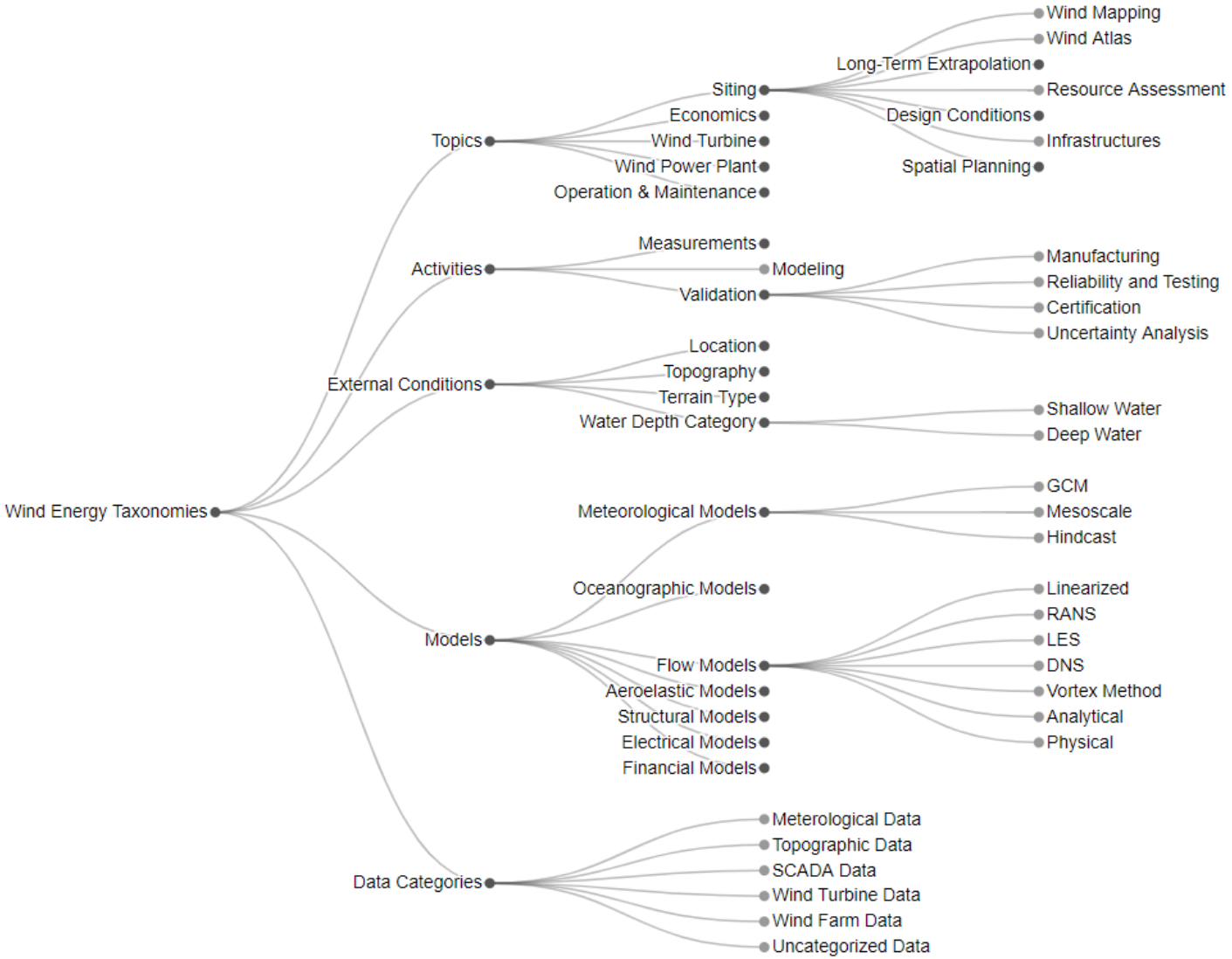
- Datasets are distributed in the "cloud", saved in and organized in different ways
- Datasets often miss **documentation (Metadata)**

Action: 3 ingredients

- Create **metadata** and
- Assign to metadata relevant controlled vocabularies (**Taxonomies**) to tag data
- Design **a data portal** for metadata catalogues

IRPWind Project 2014-2018 - Metadata & taxonomies





IEA Task 41 WP2 data catalog: Find the data

Deliverable D11:

- Fall 2020: Development of **data sharing**, **storage** and if needed **security protocols** for metadata to be stored on the platform. Specification of a potential **data sharing portal** that expands on the catalog.

Goals

Identify:

- Data contributors and users
- Needed shared resources
- Data availability on key topics
- Recommended practices for data collection, reporting, accessing, and storage

Catalog and Make Available:

- Metadata for distributed wind data sets

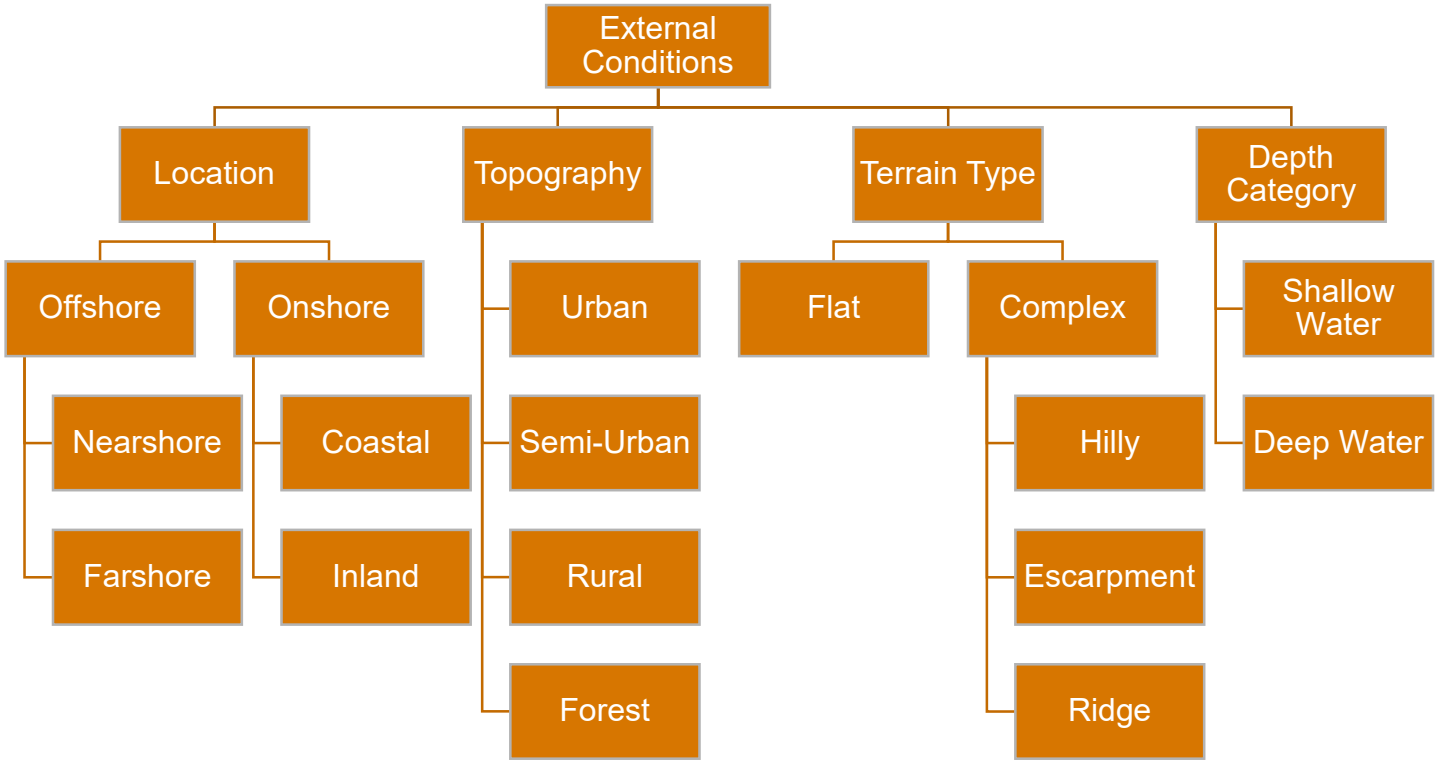
Consider:

- Including a catalog of data processing and decision support tools

PNNL-SA-152844

Taxonomy

- For some of the metadata elements, PNNL expanded the wind energy taxonomy developed by IRPWind to include some terms specifically relevant to the distributed wind energy community.



This is an example from the wind energy taxonomy. The full taxonomy is in Wind Energy Taxonomy Excel file.

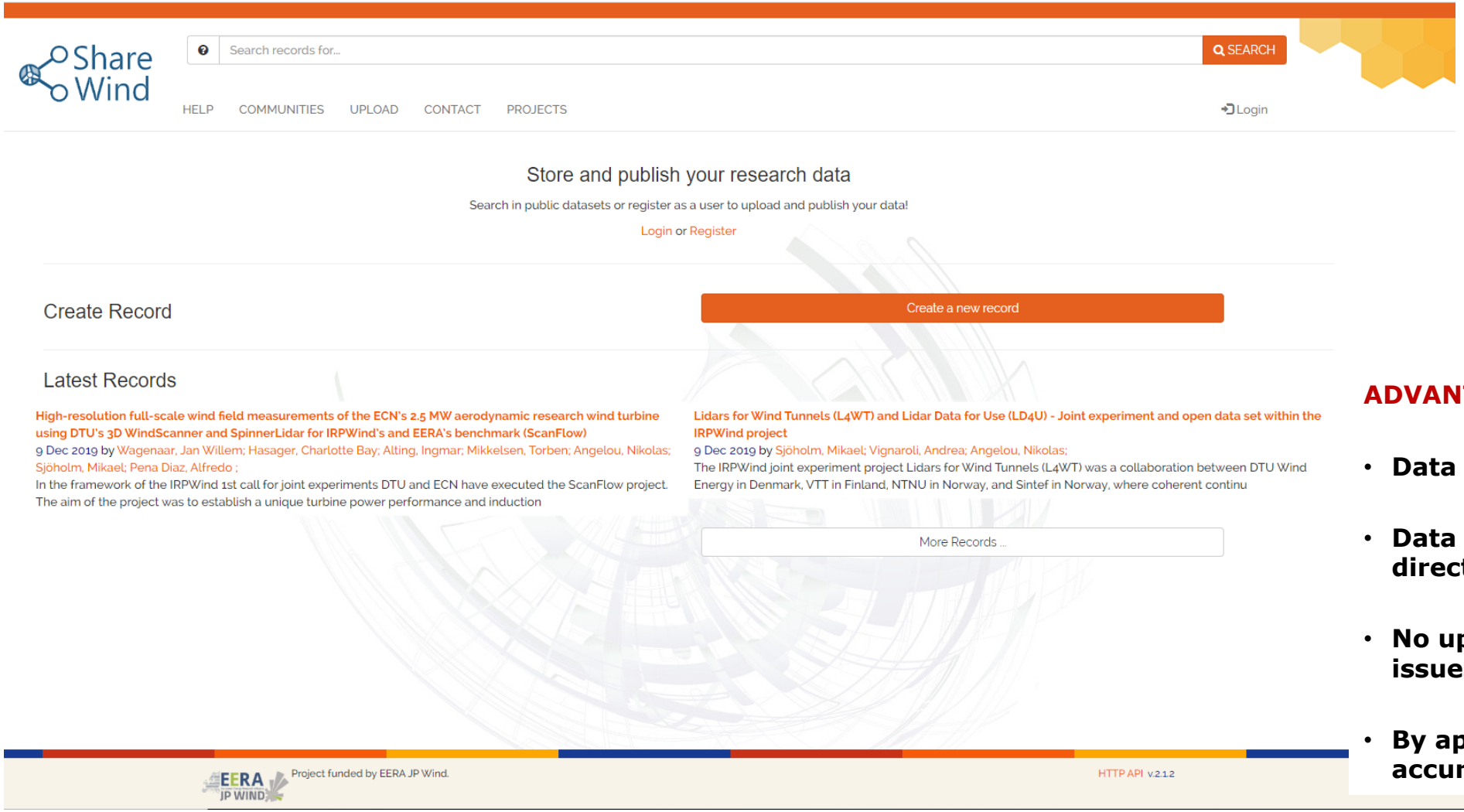
Existing Databases, Portals, and Catalogs

- Work to Date:
  - Establish and evaluate wind-related databases and catalogs that already exist
  - Identify opportunities for collaboration or to build upon existing work
  - Lessons learned from previous work
  - Outline a process for metadata collection and options for hosting Task 41 catalog

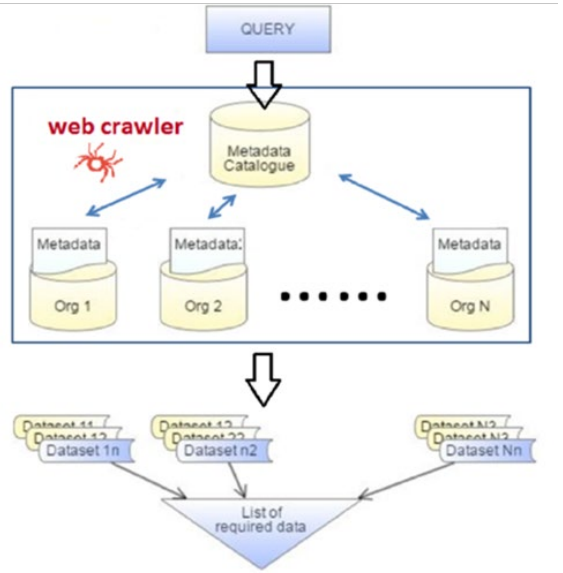
Tethys	Data Archive and Portal (DAP)	OpenEI	ShareWind

PNNL-SA-152844

IRPWind Project 2014-2018. ShareWind.EU
Data portal and metadata catalogue



ShareWind



ADVANTAGES

- **Data are kept at the owner premises**
- **Data are visible without being directly accessible**
- **No uploading data and storage issues**
- **By applying filters users can accurately locate needed data**

Pacific Northwest
NATIONAL LABORATORY

Open Energy Information (OpenEI) - Overview

- A wiki platform for the energy community, including policymakers, developers, and researchers
- Renewable energy and energy efficiency focus
- Sponsored by US DOE, NREL, and a third party renewable energy search engine, reegle

PNNL-SA-152844

Conclusive remarks

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

- **To connect safely users to data owners**
- **Give information on the availability of shared resources and of Data on key topics**

Data owner /creator

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

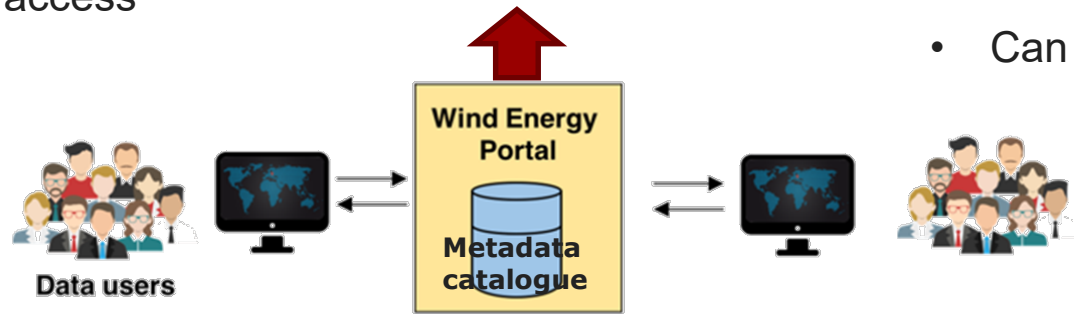
Data

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
- Can save time dedicated to the task

Thank you for your attention!

NEXT

Would you share your data?

BREAKOUT SESSION

- Have you a data catalogue Visible/Findable from outside your company?
- If no, would you use an established taxonomy to tag your data?
- If yes, how did you tag your data?

Could you please describe using three keywords the FAIR data principle

(e.g. ambitious, innovative, interesting, impossible, appealing, not-applicable, …..

Under which conditions would you share your data?

- Against a fix/variable amount of money;
- Against services;
- Against involvement in projects; for free if data is not used in competitive goals.

PESTEL Analysis

Political, Economic, Social, Technological, Environmental, Legal

- **Barriers/obstacles and framework conditions affecting FAIR impact**
- Sharing data, tools and workflows: a strategy to inspire efficient collaboration - Metadata catalogue: **Distributed data bases.**

	Barriers/Obstacles/Risks	Methods to resolve issues
Political	*Governmental funding agencies demand open data but at the same time Governments cut funding to universities demanding universities providing business models to support research.* *IPR and Patents are success criteria for universities*	Take actions to communicate that FAIR data is a good balance between Open data and IP protected data
Economic	*Data as competitive advantage*	Communicate the benefits of open data as a way of lowering project costs, enabling a faster project progress and enhancing replication in other markets
Social	*Managerial practices and skills, culture of open data*	Implement training programs for both early stage researchers and senior researchers.
Technological	*Lack of interoperability: access to data, data and software compatibility, lack of metadata*	Establish agreed standards to support interoperability and secure a better quality of data
Environmental	*Critical mass of data available*	System for recognition/ rewards to for their work. Ensures awareness and thereby generates interest in protecting the environment by being able to conduct research with open access data and develop innovative solutions.
Legal	*Constraints to the access to nationally funded research infrastructures by international consortia,* *Copyright and ownership* *Variety of EU directives, regulations and national laws and policies, as well as multinational initiatives, not fully coordinated such as the Research Data Alliance*	Greater coherence to the incentive, legal and regulatory frameworks governing research data and tools. Establish an information base of guidelines and instructional materials to secure legal reuse of data Delegate a body e.g. the EOSC, to play a coordinating role, of active initiatives