DTU

Kurt S. Hansen & Nikola Vasiljević, DTU Wind Energy

Rescuing legacy data from obsolescence:

The WindData.com user-case

Short CV - Kurt S. Hansen
Emeritus (Senior Scientist)

ORCID: 0000-0001-7109-3855

- MSc. Mech Engineering (1977) DTH

- DTU Wind Energy
  - Working area: research & education

- 44 years of experience within wind energy
  - Initial projects: (large) proto wind turbines, design & testing.
  - **Database on Wind Characteristics** & data analysis as part of research projects.
  - Flow analysis of onshore and offshore wind farms.

Introduction

- Research groups often possess a rich heritage of data spanning periods of several decades.
- As a consequence, the legacy data are typically unreadable and unaccessible due to obsolete formats and technologies in which they are provided.
- To prevent this from happening, DTU Wind Energy executed a pilot data rescue project and FAIRified the WindData.com data collection, moving it to a new technology stack.
- In this talk, we will present **process** and the **result** of this work.

Outline

- Database on Wind Characteristics
  - Introduction to the legacy dataset;
  - Vision, technology & contens;

- Requirements for the new database

- Implementation
  - Structure & transfer process
  - Data documentation
  - Examples of access to data and documentation

- Status for the transfer
- Wrap-up

Database on Wind Characteristics [objectives]

- Vast amounts of wind
  locations

Visions

The database should be useful as a supporting tool to identify:

- Rare meteorological events;
- Distributions of wind speed and turbulence (time & shear);
- Spatial structure of turbulence;
- **Interaction**: meteorology ⇔ wind turbine
- Spacing parameters in wind farms.
- ....

ublishing a small,

ese data was initiated in 1996.

Database on Wind Characteristics – FAIR before FAIR acronym was defined

The procedure when receiving data included:

- Data conversion to a **standard format**.
- Screening for errors (=>quality index).
- Re-calculation of the statistics.
- Documentation of setup and instrumentation based on a **template**.

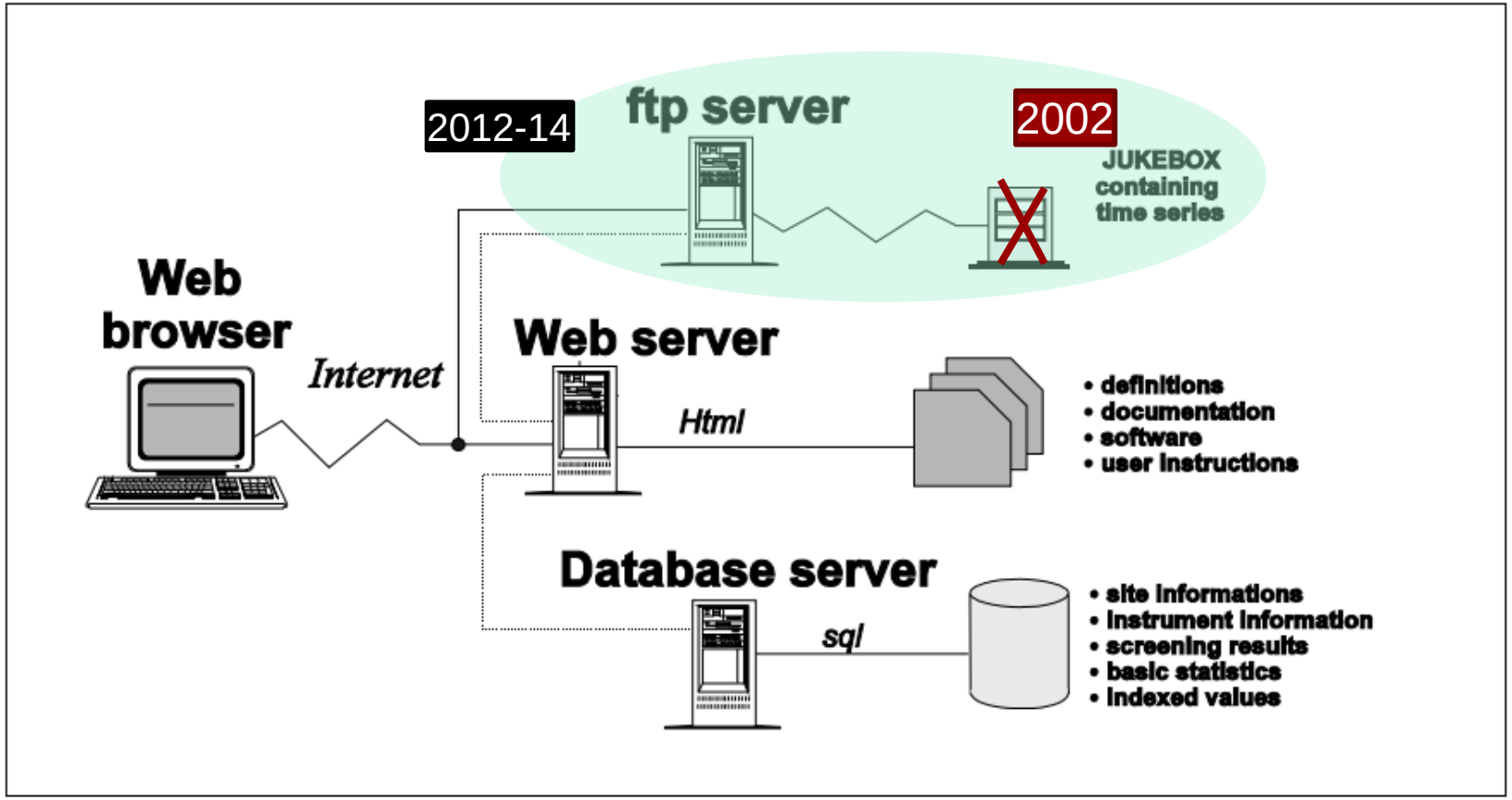
Technology – initial [1996]

- *Database*
  - Borland Interbase => **MySQL**
    - User accounts (guest, users & administration)
    - Simple queries & advanced queries

- *Webserver*: MS IIS => **Apache**
- *Online queries*: **PHP**
  - FTP => **sFTP**
  - Online graphic tool

- **Challenges**
  - Updates & ethical hackers

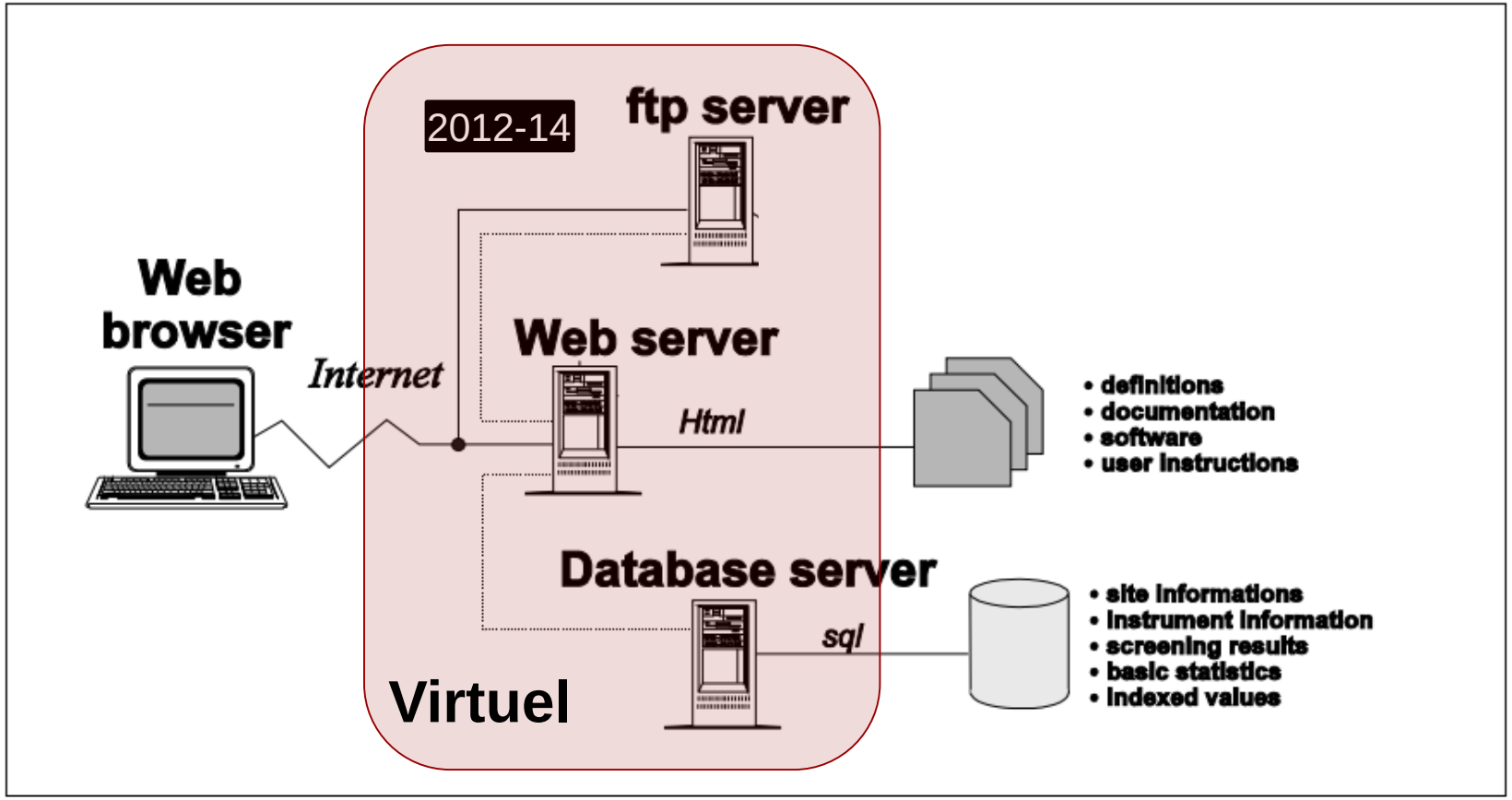
Structure of database server[s] in years 1996-2002



Jukebox with 150 CD’s

105 GB

Structure of database server after 2014



Database on Wind Characteristics [contents]

FTP server 109GB

- ***Time series of wind measurements***
  - Documentation
  - Statistics
  - Time series (1-40 Hz)
  - *Max. 62 sites*
    ~196187 hours of measurements

- ***Time series of wind turbine loads***
  - Documentation
  - Statistics
  - Time series (1-40 Hz)
  - 7 sites
    ~ 6533 hours of measurements

- ***Resource data*** *(statistics)*
  - Documentation
  - Statistics
  - *Max. 49 sites*
    ~ 1342132 hours of statistics

- ***Wind farm data*** *(SCADA; statistics)*
  - Documentation
  - Statistics
  - *2(3) sites*
    ~ 19141 hours of statistics

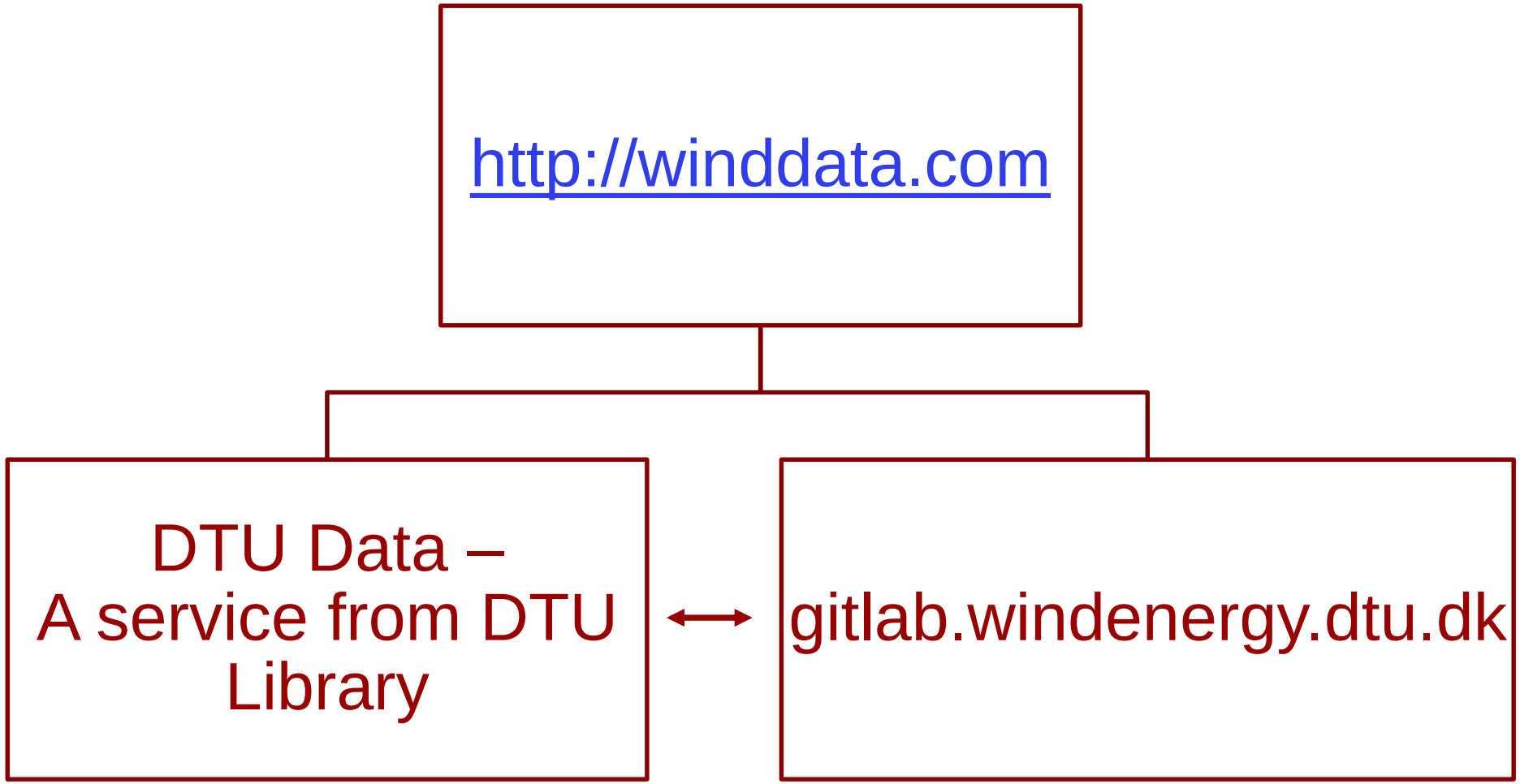
MYSQL: 16GB

> 100 sites

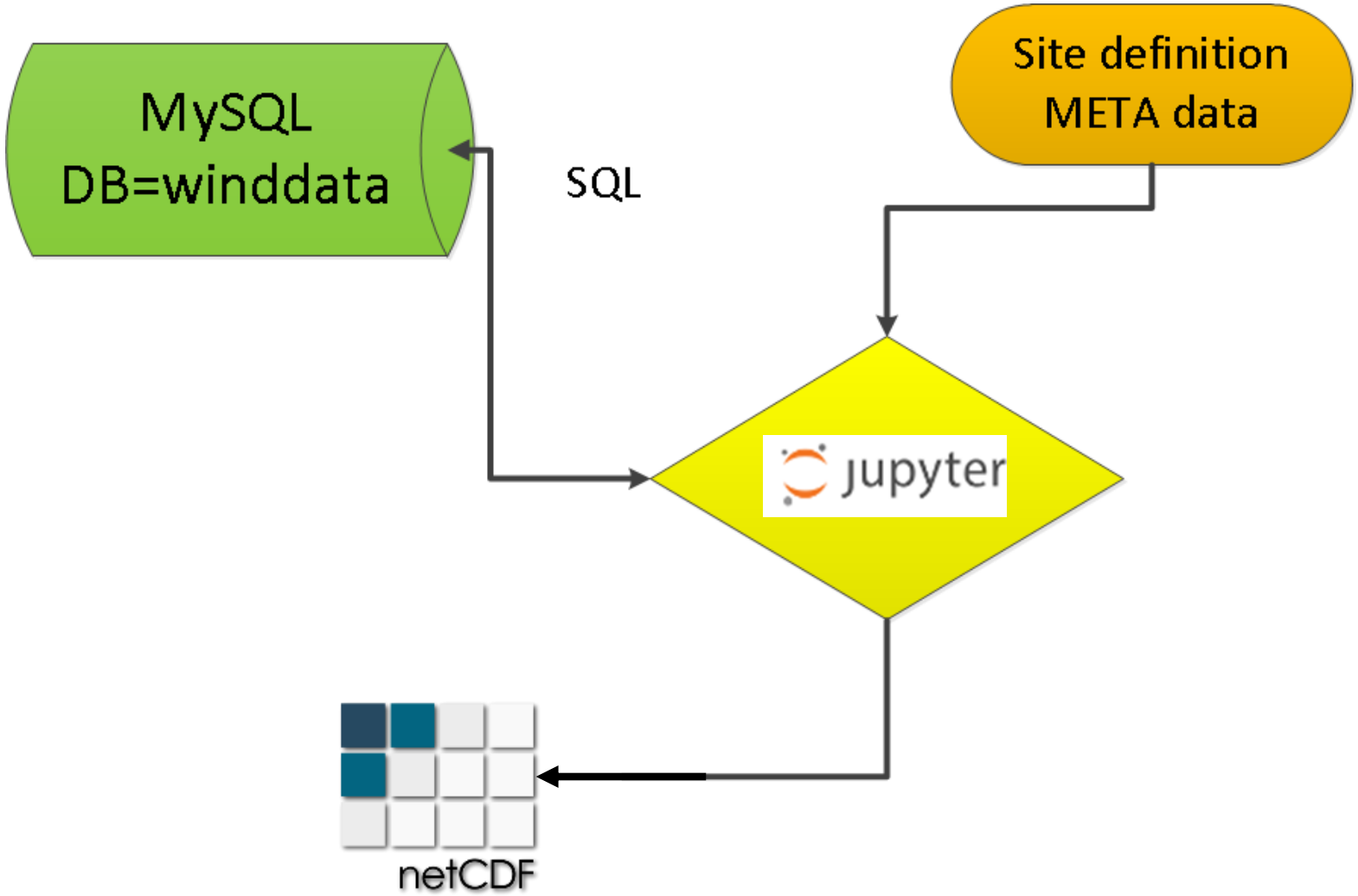
Requirements for new database

- Data should be FAIR;
- Global and persistent findability, but also discovery through **winddata.com**, a solid brand for quality wind data;
- All data should have a clear accessible criteria;
- Data should be reusable through detailed documentation and by clear copyright information;
- Interoperable file formats should be used for the data;
- Easy citation of specific dataset;
- **Low maintenance**.

New platform structure



Data transfer process



data.dtu.dk

Download all

gitlab.windenergy.dtu.dk

Documentation

Resource data from the Godhavn met mast

Cite | Download all (19.13 MB) | Share | Embed | + Collect | •••

Dataset posted on 15.04.2021, 10:44 by Kurt Schaldemose Hansen, Nikola Vasiljevic, Steen Arne Sørensen

This resource dataset consists of 10 minute wind speed and wind direction measurements from a small met mast at Qeqertarsuaq (Godhavn),Greenland. The period includes 10 years of measurements, which starts in 1991.
Detailed documentation can be found here.

Data files:
1) Godhavn_all.nc includes all periods
2) Godhavn_concurrent.nc only includes concurrent periods for all channels.

FUNDING

ternal

DTU Data
Ex. site=Godhavn

HISTORY

- **14.04.2021** - Submission date
- **15.04.2021** - First online date, Posted date

.NC

godhavn_all.nc (9.56 MB)

REFERENCES

- http://doi.org/10.5281/zenodo.4580574
- http://doi.org/10.5281/zenodo.4580722
- https://gitlab.windenergy.dtu.dk/fair-data/winddata-revamp/winddata-documentation/-/blob/master/godhavn.md

LOCATION

Latitude: 69.250833, Longitude: -53.513611

DATE

Start date: 1994-06-01 Stop date: 2002-07-14

TOPIC

- Siting;>Wind Mapping
- Siting;>Design conditions;>Extreme wind

MODELS

- Not applicable

ACTIVITIES

- Measurements;>Field experiment

EXTERNAL CONDITIONS

- Location;>Offshore;>Nearshore
- Terrain type;>Flat

DATA CATEGORY

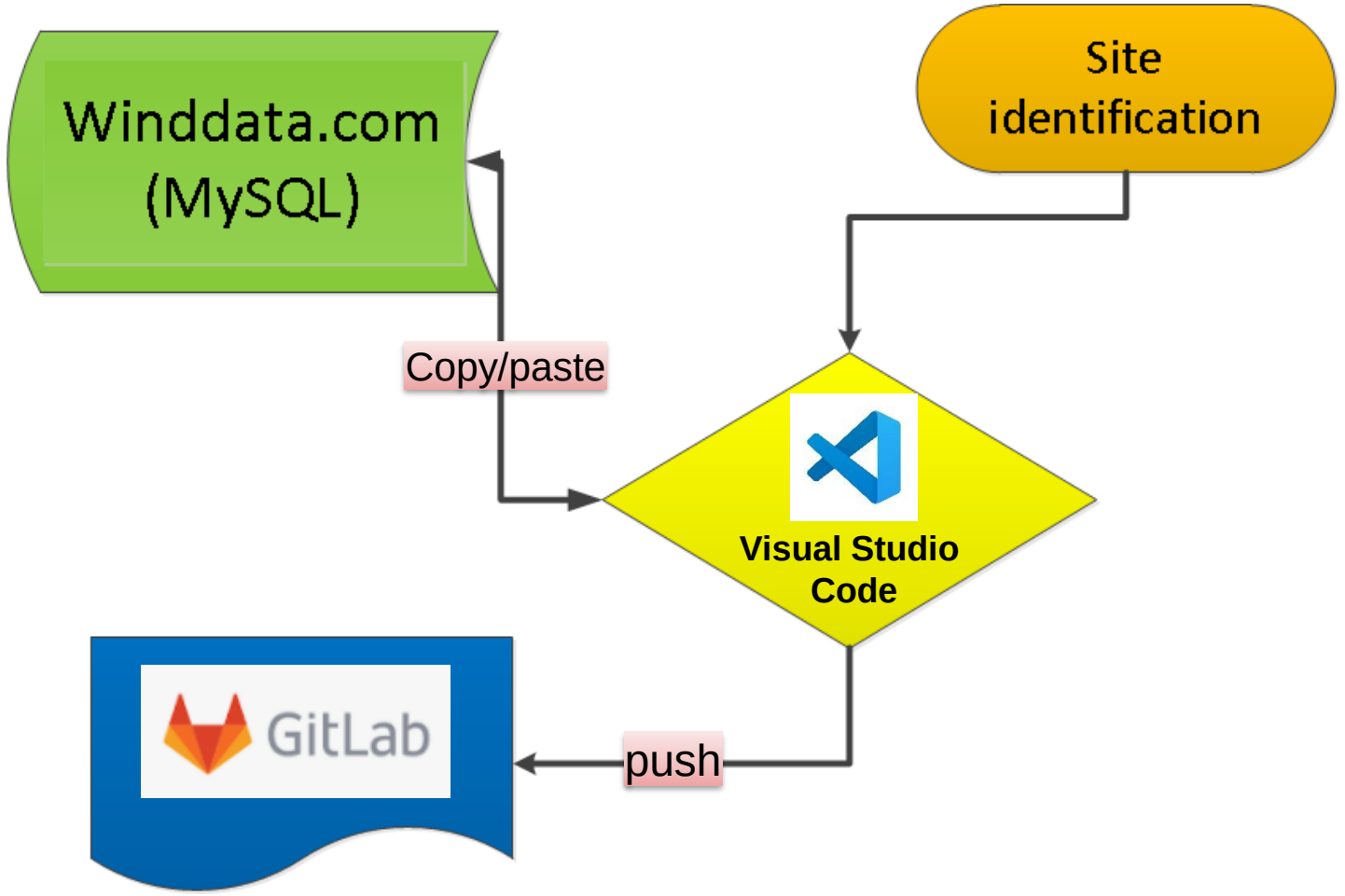
- Meteorological data

Wind Energy Topics
http://purl.org/neat

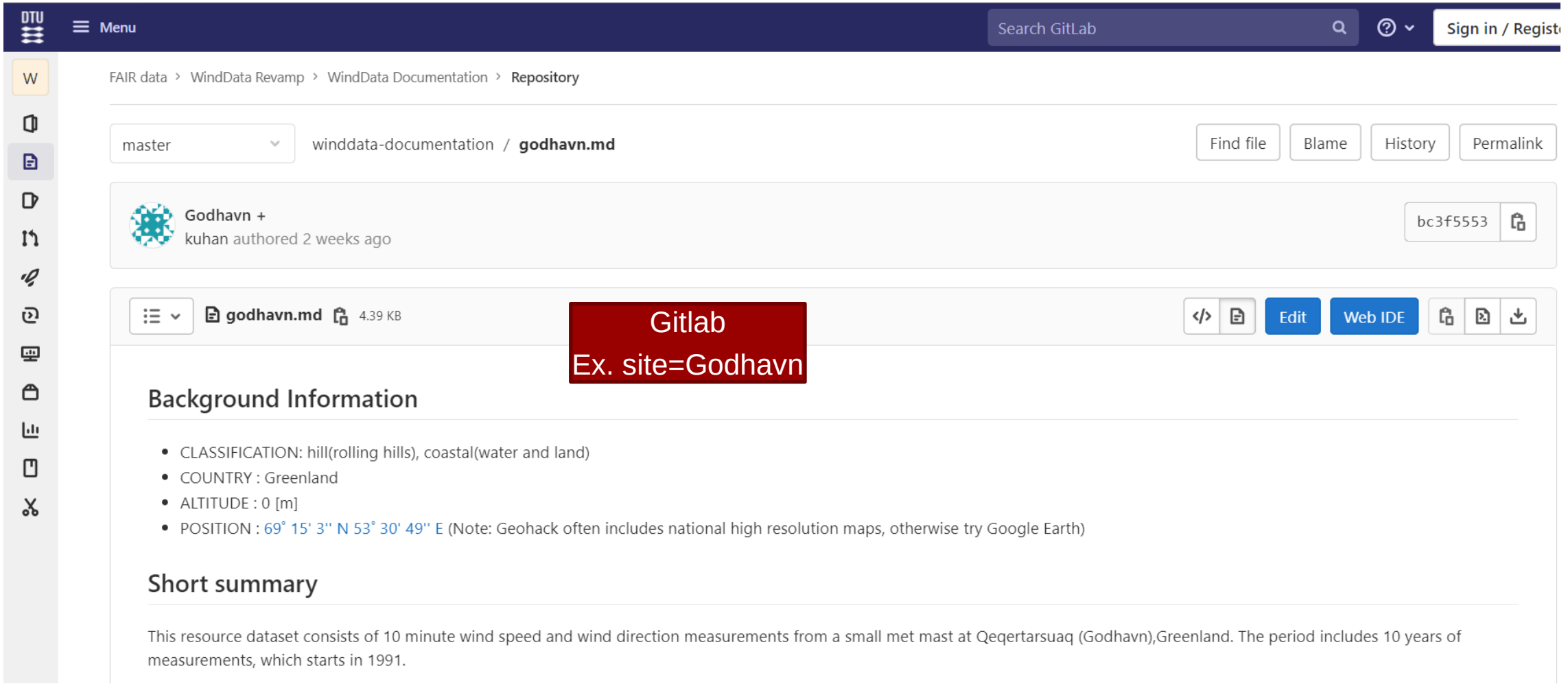
Wind Energy Acticities
http://purl.org/weave

External Conditions Taxonomy
http://purl.org/extract

Documentation process



https://gitlab.windenergy.dtu.dk godhavn.md



https://gitlab.windenergy.dtu.dk/ => godhavn.md

1. List of mast signals

SCROLLED TO BOTTOM

ACKNOWLEDMENTS

- Acknowledgement : Niels Nielsen, Institute of Geography, University of Copenhagen.
- E-mail: nn@geogr.ku.dk
- INSTITUTION : Institute of Geography, University of Copenhagen
- ADDRESS : Øster Voldgade 10, DK-1350 Copenhagen K
- TEL/FAX : Ole Humlum / N.Nielsen
- CONTACTS : Niels Nielsen
- LINKS : http://www.geogr.ku.dk/
- COLLABS : Copenhagen University
- FUND AGENTS : Internal
- PERIOD : 1991-2001
- Naming_authority : 'DTU Data'
- DOI : 'https://doi.org/10.11583/DTU.14153210'

Gitlab
Ex. Godhavn

DOI to
DTU Data:
Godhavn

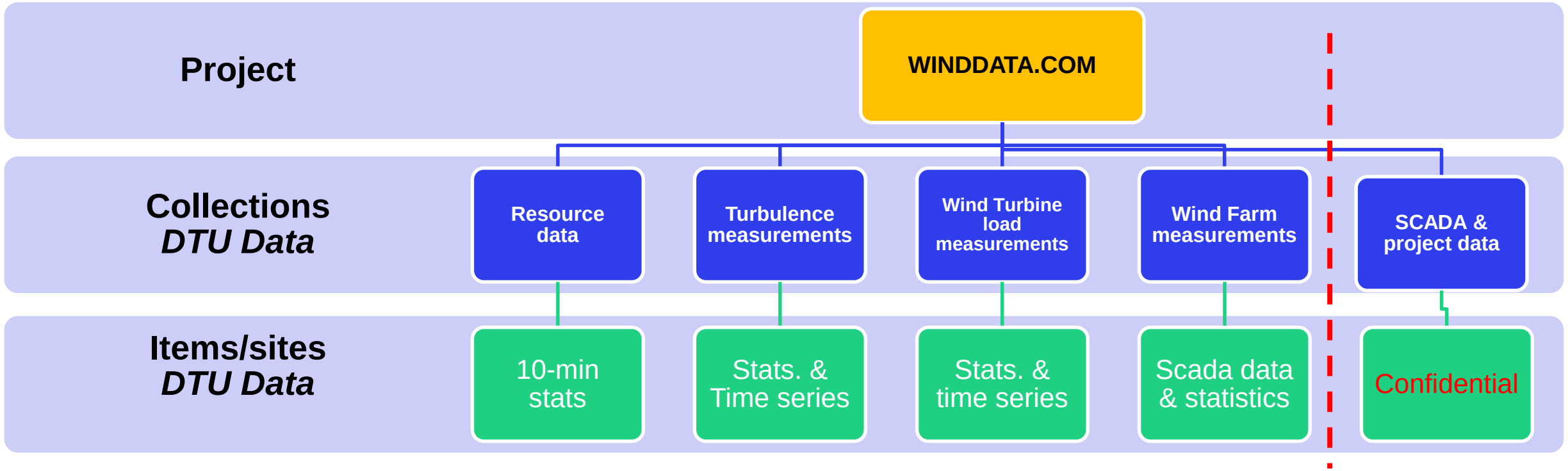
Public data

1. Resource data (NetCDF)

DTU Data structure

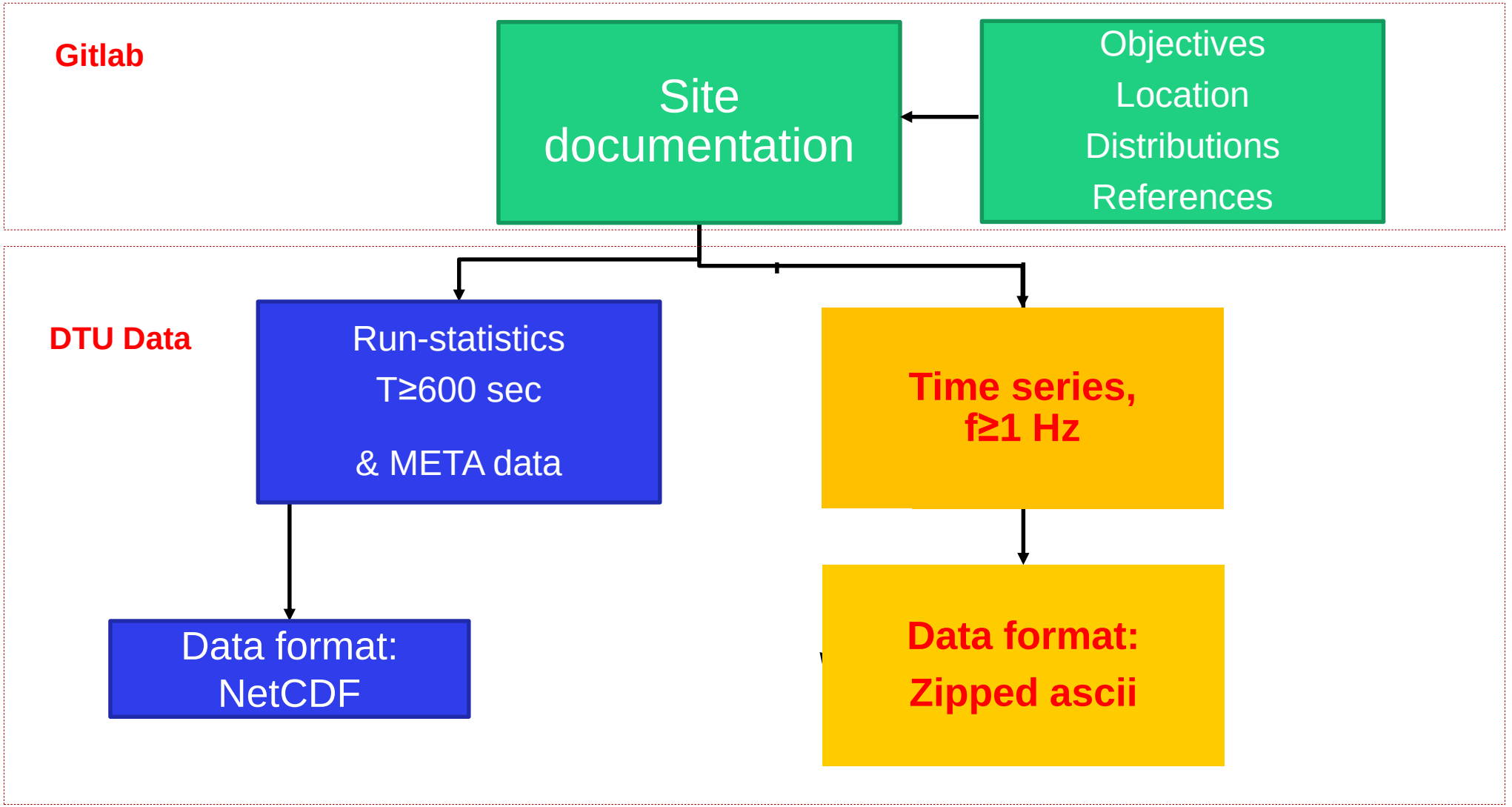
Terminology:
1) Item = site
2) Collection = group of items

DTU Data: Database on Wind Characteristics



Documentation is located on: *Gitlab/Fair data/Winddata Rewamp/*

Documentation of *Time series of measurements*



Status for winddata.com on DTU Data

- Collection of wind resource data finalized ☺
- Collection of time series of wind measurements (>85%)
- Time series of structural loads (66%)
- Wind farm data (33%)
- Internal SCADA & Project data - overview ☺

Database on Wind Characteristics
Wrap-up

- Five data collections have been defined
  – Wind resource; time series and SCADA data.

- Major part of winddata.com including documentation has been transferred.

- Metadata of both statistical and raw time series are publicly accessible.

- A platform for future publication of wind energy related dataset has been established.

Thanks for your attention

Questions?