WESC 2025
24 - 27 JUNE | NANTES, FR

Wind Energy Science Conference
24 - 27 June 2025
Nantes, France

Design characteristics and capabilities of the WiValdi Data Management System

Martin Jessen, Stephan Graeber, Bijan Barzgaran

Deutsches Zentrum für Luft- und Raumfahrt, Institut für Flugsystemtechnik

E-mail: martin.jessen@dlr.de

Keywords: data management, derived data, key-value store, archiving, data access, data security

1 Motivation for implementing a Data Management System

Together with the partners from the Research Alliance Wind Energy (ForWind, Fraunhofer IWES) the German Aerospace Center (DLR) has built the Krummendeich Research Wind Farm WiValdi.
An effective data management system is crucial for a research facility that generates a large volume of sensor data like WiValdi, enabling collaboration, reproducibility, and research advancement. The following abstract describes the implementation and the underlying requirements of this system.

2 Requirements for the WiValdi Data Management System

Based on the nature of WiValdi as a research facility that continuously generates data and addresses a wide research audience, the DMS must meet, among other things, the following requirements.

2.1 Acquisition and storage of data

The DMS must be able to gather data acquired by measurement systems of various manufacturers and store it shortly after capture. The resulting large amounts of data need to be permanently archived and it has to be ensured that raw data cannot be manipulated retrospectively to ensure data integrity.

2.2 Data processing and access to data

The DMS must generate 10- or 15-minute-average values and other similar aggregated data and make these aggregations accessible for researchers because data in this time-resolution is already widely used e.g., in the field of meteorology. This aggregated data is also important for use as meta data for use together with the higher frequency raw data. A sophisticated DMS must also be able to perform general calculations on the data, so that derived data can be generated and stored.
For researchers it is important to have an easy-to-use and efficient way to access the data. Thus, data has to be accessible through a browser-based web portal and APIs for direct interaction with common tools. It is also important to be able to access all data shortly after the measurement.

2.3 Data security

An important aspect for a real world DMS is data security. The system must be able to handle public as well as confidential data and make each available in an appropriate way. While for some research aspects, data shall be accessible as easily as possible, it is also crucial to be able to effectively restrict access to confidential data to protect the intellectual property of involved parties. To provide a secure access to authorized users, a finely configurable rights management is required.

2.4 Continuous unsupervised operation in a remote location

As wind turbines are designed for continuous unsupervised operation in remote locations, this is also a requirement for the DMS. I.e., it should run without interruptions with minimal or no maintenance for a long period of time. It is also important that any administrator interaction needed, e.g., during the initial operation phase, can be performed remotely.

3 Properties of the implemented system

The WiValdi DMS has been developed in collaboration with Infolytics AG to address the specific needs and requirements outlined in the previous chapter. As the first step in the development process, a comprehensive concept document was written to outline the functional specifications of the system. Building on this foundation, this chapter delves into some of the details of the system's architecture, including the database choices and data processing capabilities which enable efficient and reliable data management. Additionally, we will explore the rights management features that ensure secure access and control of the data.

3.1 General architecture

To ensure continuous and unsupervised operation in a remote place, the WiValdi DMS is divided into an onsite component located directly at the wind farm and an archive component installed in a DLR data center. This makes it possible to have local data collection and storage without reliance on internet connectivity, providing a layer of protection against data loss. The communication between the onsite component and the archive component is not time-critical, enabling efficient data transfer.

3.2 Data acquisition systems and interfaces to the Data Management System

The DMS is designed to handle various types of data, including equidistant and non-equidistant time series data, multidimensional data, and opaque data. The data is collected, processed, and stored in near real time, with a defined interface for connecting measurement systems to the DMS. The DMS is capable of simultaneous data acquisition from various measurement devices and the interface is designed to be open and manufacturer-independent, making it possible to connect a variety of different data acquisition systems and, e.g., having a direct link to the turbine controller.

3.3 Database choices for efficient storage and accessibility

As described in chapter 2, the system is required to handle large amounts of data and provide fast access for researchers. To reach this goal, a combination of a KVS (key-value store) and an SQL database for efficient storage and accessibility are employed. The KVS is used for storing raw data in 1-minute-packages, while the SQL database is used for storing aggregated data, such as minute-average values. A memory intensive aspect of the raw data processing is the way data is sent to the system and the way it needs to be stored. The incoming stream consists of a large amount of measured values arriving in parallel for every specific moment while the typical way to access data is to look at a specific value over a longer period of time. The necessary transpose operation is done in the so called “Staging Areas” on the Onsite DMS. The data is then transferred to the Archive DMS where the calculation of aggregations and permanent storage are handled.

3.4 Data processing capabilities and access to data

The DMS provides a set of APIs for data access and processing. These are used, e.g., for the creation of aggregated data (see above) and also enable the calculation of other derived data in so called “Signal Data Applications” running in the Archive DMS. The derived data can then be handled just as every other stored data.
Researchers can access the data through a web interface which includes an online visualization of the

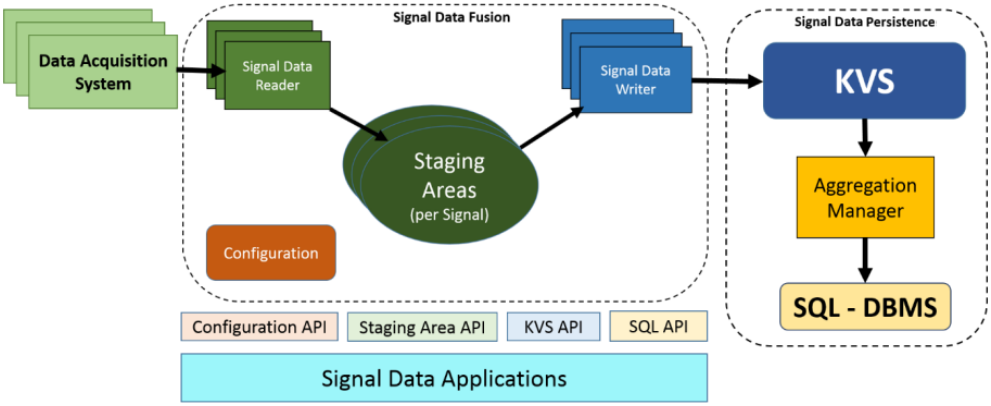

Figure 1: Data acquisition, processing and storage in the DMS developed by Infolytics AG

aggregated data. This can be used to identify interesting events and time intervals for raw data download. Alternatively, it is possible to directly access the raw data with common tools via an API. To ensure data quality, the DMS allows instrumentation operators to store corresponding insights as metadata.

3.5 Rights management and protection of intellectual property

The DMS implements a rights management system that allows for finely graduated control over access to data, down to the signal level. The system is designed with FAIR principles in mind. While it has a focus on simplicity of access to open data, an equally important aspect is the handling of confidential data and intellectual property. Thus, all data is protected through access controls, and the system allows for project-specific rights management. Additionally, the system enables the upload of additional data, but with controlled write access to ensure data integrity and relevance.

4 First results and outlook

WiValdi features an extensive range of equipment with all measurement data converging centrally in the DMS. This includes:

- Two Enercon E-115 EP3 E4 wind turbines with access to high frequency controller data
- Four highly instrumented measurement masts ranging in height from 100 to 150 meters
- Over 30 distributed measurement systems
- Over 1,000 electrical and fiber-optic sensors installed in each wind turbine, monitoring parameters such as vibration, acceleration, strain, foundation and pile movement, bearing condition, inverter behavior, turbine alignment, local air pressure, and more
- Multiple LiDAR systems
- Acoustic measurement devices for sound exposure and acoustic emission

The DMS described in the previous chapters has been gradually put into operation with a growing number of data-providing measurement systems from autumn 2022 and is now in regular operation, working as planned. So far, approximately 20 terabytes of measurement data have been recorded and archived and can be accessed by researchers.
Throughout the system's planned lifespan of 20 years, this amount of data will constantly grow and provide a valuable research resource. The APIs of the system open up possibilities for future expansions in the field of data processing, such as automated data analysis using machine learning algorithms.

Acknowledgements

This work was accomplished as part of the research project Deutsche Forschungsplattform für Windenergie. We greatly acknowledge the financial support of the German Federal Ministry for Economic Affairs and Climate Action, through FKZ 0325936, that enabled this work.