

## BOKU-WaterRDM

BOKU Research Data Management Infrastructure for Water Research

|                                 |                                                                                                      |                        |            |
|---------------------------------|------------------------------------------------------------------------------------------------------|------------------------|------------|
| <b>Programm / Ausschreibung</b> | Implementing EU Missions, Implementing EU Missions, Forschungsdateninfrastruktur Ausschreibung 2024  | <b>Status</b>          | laufend    |
| <b>Projektstart</b>             | 01.12.2025                                                                                           | <b>Projektende</b>     | 30.11.2028 |
| <b>Zeitraum</b>                 | 2025 - 2028                                                                                          | <b>Projektlaufzeit</b> | 36 Monate  |
| <b>Keywords</b>                 | Water Research; Research Data Management (RDM); FAIR Principles; Interoperability; Big Data Analysis |                        |            |

### Projektbeschreibung

Wasserforschung erfordert groß angelegte, räumliche und zeitliche Analysen vielfältiger Datensätze, einschließlich hydrologischer, ökologischer und umweltbezogener Überwachungsdaten. Die BOKU Universität, eine führende Institution der österreichischen Wasserforschung, arbeitet mit nationalen und internationalen Partnern zusammen, um Wasserressourcen vor Verschmutzung durch Schadstoffe, Krankheitserreger und Umweltgefahren zu schützen. Zur Unterstützung dieser Bemühungen soll im Rahmen dieses Projekts BOKU-WaterRDM, eine skalierbare Forschungsdatenmanagement-Infrastruktur (RDM) für Sequenzierungsdaten, chemische Analysen und Multiskalen-Bildgebung, etabliert werden.

Inspiziert von führenden europäischen Institutionen wie der Universität Tübingen und der Vrije Universiteit Amsterdam wird BOKU-WaterRDM das Datenmanagement zentralisieren, um eine hohe Datenintegrität, Zugänglichkeit und Interoperabilität zu gewährleisten. Die BOKU Core Facility Bioinformatik wird das Projekt leiten und dabei ihre Expertise in Big-Data-Analyse und Metadatenmanagement nutzen. Die enge Zusammenarbeit mit den Core Facilities für Multiskalen-Bildgebung und Massenspektrometrie stellt eine umfassende Abdeckung aller relevanten Datentypen sicher.

- BOKU-WaterRDM wird auf iRODS (Integrated Rule-Oriented Data System) basieren und folgende zentrale Funktionen bieten:
- Standardisiertes Metadatenmanagement, um Daten auffindbar und wiederverwendbar zu machen.
- Datenzugriffsrichtlinien im Einklang mit den FAIR-Prinzipien.
- Verfolgung von Datenherkunft und Integrität, um langfristige Nutzbarkeit sicherzustellen.
- Interoperabilität mit österreichischen und europäischen Forschungsnetzwerken (z. B. eLTER RI, Danubius RI) für den Datenaustausch.
- Optimierte HPC-Workflows, mit Integration des Vienna Scientific Cluster 5 für groß angelegte Analysen wie Metagenomik und Klimasimulationen.
- Anbindung an den BOKU-Cluster für Biomolekulare Modellierung und Bioinformatik, um KI-gestützte Anwendungen zu ermöglichen.
- SQL-Datenbanken (PostgreSQL, MySQL) und Objektspeicher (MinIO, S3) für sicheres und skalierbares Datenmanagement.
- Effiziente Langzeitspeicherung, durch die Verknüpfung von iRODS mit der bestehenden BOKU-Infrastruktur zur sicheren Datenarchivierung.

Aktuell fehlt es der Wasserforschung an einer zentralisierten, automatisierten Datenmanagementlösung, was zu fragmentierter Speicherung, manueller Metadatenverwaltung und begrenzter Interoperabilität führt. BOKU-WaterRDM adressiert diese Herausforderungen, indem es eine skalierbare, FAIR-konforme Infrastruktur bereitstellt, die Datenintegrität erhöht, Workflows optimiert und eine nahtlose Datenübermittlung sowie Archivierung ermöglicht. Die Integration mit den BOKU Core Facilities gewährleistet langfristige Stabilität, erhöht die Sichtbarkeit des Projekts und fördert interdisziplinäre Zusammenarbeit.

BOKU-WaterRDM steht im Einklang mit nationalen und europäischen Strategien zur Wasserforschung, indem es interdisziplinäre Forschung erleichtert, Entscheidungsträger mit standardisierten Daten unterstützt und Transparenz sowie Open Science fördert.

Durch die Etablierung der ersten spezialisierten RDM-Infrastruktur für Wasserforschung in Österreich positioniert sich BOKU als Vorreiter in nachhaltiger Wasserforschung und Klimaanpassung. Diese Infrastruktur stellt sicher, dass zukünftige Forschung modernste Datenmanagementlösungen optimal nutzen kann.

## **Abstract**

BOKU University leads a unique combination of water research in Austria covering areas such as hydrology, flowing water research, sustainable hydropower, water supply and wastewater management, ecological water status and, water quality analysis. To support these efforts, this project aims to develop BOKU-WaterRDM, a scalable Research Data Management system designed for large-scale spatial and temporal datasets, including sequencing data, chemical analyses, and multiscale imaging.

Inspired by leading European institutions such as the University of Tübingen and Vrije Universiteit Amsterdam, BOKU-WaterRDM will centralize data management, ensuring high data integrity, accessibility, and interoperability. The BOKU Core Facility Bioinformatics will lead the initiative, leveraging expertise in big data analysis and metadata management.

Collaboration with Core Facilities for Multiscale Imaging and Mass Spectrometry will ensure comprehensive data coverage.

Powered by iRODS (Integrated Rule-Oriented Data System), BOKU-WaterRDM will offer:

- Standardized metadata management for findable, reusable data.
- Data access policies aligned with FAIR principles.
- Provenance and integrity tracking for long-term usability.
- Interoperability with Austrian and European research networks (eLTER RI, Danubius RI) for data sharing.
- Optimized HPC workflows, integrating with Vienna Scientific Cluster 5 for large-scale analyses like metagenomics and climate simulations.
- Integration with BOKU's Biomolecular Modeling and Bioinformatics cluster, enabling AI-driven applications for environmental and molecular research.
- SQL databases (PostgreSQL, MySQL) and object storage (MinIO, S3) for secure and scalable data management.
- Efficient long-term storage, linking iRODS with BOKU's existing infrastructure for data vaulting.

Current water research lacks centralized, automated data management, leading to fragmented storage, manual metadata handling, and limited interoperability. BOKU-WaterRDM will address these challenges by providing a scalable, automated, FAIR-compliant infrastructure that enhances data integrity, streamlines workflows, and supports seamless data submission and archiving. This infrastructure will be integrated and supported by the BOKU Core Facilities, enhancing its visibility, ensuring long-term sustainability, and fostering interdisciplinary collaboration.

BOKU-WaterRDM aligns with national and EU water research strategies, fostering interdisciplinary research, facilitating policymaker engagement with standardized data, and promoting transparency and open science.

By establishing Austria's first dedicated RDM for water research, BOKU-WaterRDM positions BOKU as a pioneer in sustainable water research and climate adaptation, ensuring future research fully leverages state-of-the-art data management solutions supported by the BOKU Core Facilities.

### **Projektpartner**

- Universität für Bodenkultur Wien