Contents lists available at ScienceDirect

SoftwareX

journal homepage: www.elsevier.com/locate/softx

Digital ecosystem for FAIR time series data management in environmental system science

J. Bumberger[a,b,c,*], M. Abbrent[a,b], N. Brinckmann[d], J. Hemmen[a,b], R. Kunkel[e], C. Lorenz[f], P. Lünenschloss[a,b], B. Palm[a,b], T. Schnicke[a,g], C. Schulz[a,g], H. van der Schaaf[h], D. Schäfer[a,b]

[a] Helmholtz Centre for Environmental Research – UFZ, Research Data Management - RDM, Permoserstraße 15, Leipzig 04318, Germany
[b] Helmholtz Centre for Environmental Research – UFZ, Department Monitoring and Exploration Technologies, Permoserstraße 15, Leipzig 04318, Germany
[c] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstraße 4, Leipzig 04103, Germany
[d] Helmholtz Centre Potsdam – GFZ German Research Centre for Geoscience, Department Geoinformation, eScience Centre, Telegrafenberg, Potsdam 14473, Germany
[e] Forschungszentrum Jülich - FZJ, Institute of Bio- and Geosciences (IBG), Agrosphere (IBG-3), Wilhelm-Johnen-Straße, Jülich 52428, Germany
[f] Karlsruhe Institute of Technology – KIT, Institute of Meteorology and Climate Research Atmospheric Environmental Research (IMK-IFU), Kreuzeckbahnstraße 19, Garmisch-Partenkirchen 82467, Germany
[g] Helmholtz Centre for Environmental Research – UFZ, IT Department, Permoserstraße 15, Leipzig 04318, Germany
[h] Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Fraunhoferstraße 1, Karlsruhe 76131, Germany

ARTICLE INFO

Keywords:
Environmental data management
Time series data
Sensor management
Automated quality control
Data integration
Metadata
FAIR principles
Data infrastructure
Scalable data infrastructure
Cloud-ready infrastructure

ABSTRACT

Addressing the challenges posed by climate change, biodiversity loss, and environmental pollution requires comprehensive monitoring and effective data management strategies that support real-time analysis and applicable across various scales in environmental system science. This paper introduces a versatile and transferable digital ecosystem for managing time series data, designed to adhere to the FAIR principles (Findable, Accessible, Interoperable, and Reusable). The system is highly adaptable, cloud-ready, and suitable for deployment in a wide range of settings, from small-scale projects to large-scale monitoring initiatives. The ecosystem comprises three core components: the Sensor Management System (SMS) for detailed metadata registration and management; time.IO, a platform for efficient time series data storage, transfer, and real-time visualization; and the System for Automated Quality Control (SaQC), which ensures data integrity through real-time analysis and quality assurance. With its modular and scalable architecture, the ecosystem enables automated workflows, enhances data accessibility, and supports seamless integration into larger research infrastructures, including digital twins and advanced environmental models. The use of standardized protocols and interfaces ensures that the ecosystem can be easily transferred and deployed across different environments and institutions. This approach enhances data accessibility for a broad spectrum of stakeholders, including researchers, policymakers, and the public, while fostering collaboration and advancing scientific research in environmental monitoring.

Metadata

Nr	Code metadata description	*Please fill in this column*
C1	Current code version	***SMS:*** 1.17.1 **time.IO:** 0.1 **SaQC:** 2.6
C2	Permanent link to code/repository used for this code version	***SMS:*** https://github.com/sensor-management-system **time.IO:** https://github.com/time-IO **SaQC:** https://github.com/saqc
C3	Permanent link to reproducible capsule	***SMS:*** https://zenodo.org/doi/10.5281/zenodo.13329925 **time.IO:** https://zenodo.org/doi/10.5281/zenodo.8354839 **SaQC:**https://zenodo.org/doi/10.5281/zenodo.5888547

(*continued on next page*)

* Corresponding author.
E-mail address: jan.bumberger@ufz.de (J. Bumberger).

https://doi.org/10.1016/j.softx.2025.102038
Received 11 September 2024; Received in revised form 2 January 2025; Accepted 6 January 2025
Available online 15 January 2025

(*continued*)

Nr	Code metadata description	*Please fill in this column*
C4	Legal code license	**SMS:** EUPL-1.2 **time.IO:** EUPL 1.2 **SaQC:** GNU GPL 3.0
C5	Code versioning system used	GIT
C6	Software code languages, tools and services used	**SMS:** Docker, Elasticsearch, MinIO, nginx, PostGIS Python, TypeScript **time.IO:** Alpine, CAdvisor, Django, django-helmholtz-aai, Docker CE, Docker Compose, FastAPI, FROST, Grafana, MinIO, Mosquitto MQTT Broker, nginx, NumPy, Pandas, Python Click, TimescaleDB, Tomcat **SaQC:** Python
C7	Compilation requirements, operating environments and dependencies	**SMS:** Docker, Docker Compose **time.IO:** Docker, Docker Compose **SaQC:** Python 3.8+
C8	If available, link to developer documentation/manual	**SMS:**https://hdl.handle.net/20.500.14372/SMS-Readme https://hdl.handle.net/20.500.14372/SMS-Wiki **time.IO:** https://codebase.helmholtz.cloud/ufz-tsm **SaQC:** https://rdm-software.pages.ufz.de/saqc/index.html
C9	Support email for questions	SMS: sms-core-team@listserv.dfn.de time.IO: rdm-contact@ufz.de SaQC: saqc-support@ufz.de

1. Motivation and significance

Climate change, biodiversity loss, environmental pollution and related anthropogenic impacts are leading to significant pressures on the world's ecosystems and their functions, requiring a comprehensive quantification of these impacts [1,2]. To address this, large-scale and standardised monitoring observatories such as NEON [3], eLTER [4,5] and TERENO [6,7] have been established to provide essential data for the long-term monitoring with sensor systems and modelling of environmental systems. In addition, e.g. MOSES [8] investigates the evolution and impacts of highly dynamic, often extreme events (e.g., heat waves or hydrological extremes) using a systemic monitoring approach including sensor systems. These event-oriented, cross-compartment datasets are needed to understand the impacts of climate change, biodiversity loss and pollution, and to develop effective adaptation strategies and address the development of a federated Global Ecosystem Research Infrastructure [9]. These observation networks are continuously expanding in terms of sensor density and geographical coverage. However, the resulting increase in data volumes poses challenges in handling and processing these continuously growing data streams in real-time while adhering to FAIR principles (Findable, Accessible, Interoperable, Reusable) and leveraging standardised interfaces along with associated sensor- and datastream-related metadata [10,11]. The integration of sensor data into such data infrastructures ensures subsequent real-time availability for further analysis, application of advanced data science methods, quantification of earth observation-based indicators, and integration into environmental models as well as digital twins and information systems [12–15]. Therefore, sensor management, effective storage and automatic quality assurance of time series data require the development of scalable, high-performance, and transferable data infrastructures to support real-time availability and handling of the rapidly increasing volume of sensor data [11,16–20].

1.1. *Challenges in data management and existing solutions*

In the field of geoinformatics, data management infrastructures for in-situ sensing are often considered an integral part of Spatial Data Infrastructures (SDI) or Geospatial Information Systems (GIS). As such, they are rarely addressed independently in the literature [11,12]. In environmental system sciences, however, there are solutions specifically tailored to the handling of sensor data, some of which offer reusable frameworks [17,21–24]. These solutions prioritize the standardization of data and metadata flows to ensure the subsequent reuse of data. In contrast, data infrastructures for sensors within the Internet of Things (IoT) domain are often designed for short-term data storage and/or operational control and are predominantly used in industries such as manufacturing or automation. Moreover, these systems are typically metric-oriented, focusing on derived values rather than the original measurement data [25,26]. Historically, such data management infrastructures have often been developed independently within institutions and integrated into existing IT frameworks, which complicates their scalability and promotion. The rapid growth of sensors and observatories has surpassed the scalability of these legacy systems. As a result, interdependencies have emerged that make comprehensive system overhauls resource-intensive, often limiting progress to incremental improvements rather than holistic redevelopment [11,12]. Cloud-based solutions that can be implemented universally have begun to emerge, offering a promising alternative to these traditional approaches. However, these solutions are often domain-specific, such as those used in Critical Zone Observatories or hydrology [23,24]. Furthermore, robust time series data infrastructures facilitate the integration of data from diverse sources into distributed data infrastructures at institutional, regional, national and continental levels. These solutions also enable real-time applications in digital twins, integration into Spatial Data Infrastructures, and assimilation into environmental models. Such capabilities ensure the dissemination of data to a wide range of stakeholders, including scientists, policymakers, resource managers and the general public.

1.2. *Towards a holistic framework for sensor data management*

In response, we have pursued an innovative approach to collaboratively design and systematically develop a user-centric comprehensive data infrastructure specifically for time series data management of sensors in environmental science with the following characteristics and requirements. These are:

- **Interoperability:** Use of standardised interfaces and metadata standards as well as standard protocols for sensor integration, enabling compatibility with geospatial infrastructures for seamless spatial data handling
- **Modularity:** Application of Sensor management components with persistent identifiers, sensor data infrastructure component with the ability to connect different storage solutions, and quality assurance/ data processing component, each independently deployable, and extendable to geospatial systems
- **Transferability and cloud readiness:** Use open source solutions in a microservice architecture, and container-based deployment for easy scalability and integration with existing IT infrastructures
- **Authentication system:** Leverage cross-institutional identity management to enable collaboration and the ability to use own authentication systems
- **User-friendly:** With simple responsive web interfaces for operation and integrated data viewers for data dissemination
- **Generic application:** Applicable to all domains using sensor-based data, to go beyond environmental system science

2. Digital ecosystem for FAIR time series data management

The conception of the digital ecosystem for FAIR time series data management began in 2019, building on the experience gained in the development of an interoperable data infrastructure for the TERENO observatories since 2009 [6,7,22]. Particular emphasis was placed on ensuring that both the (meta)data and the data flows meet the FAIR criteria [10]. Furthermore, the digital ecosystem and its software components were designed to comply with the FAIR principles for research software [27], establishing a state-of-the-art solution. This approach was intended to develop a reference model for research infrastructures in Environmental System Science. The adherence to the FAIR criteria for (meta)data as outlined by Wilkinson et al. [10] is detailed in Appendix 1. A key highlight is the enrichment of metadata during data processing, which ensures comprehensive reproducibility. Further attention has been paid to the modularity of the overall system and to the ability of the components to operate independently. One of the key innovations of this integrated system is its transferability and usability by other research institutions, authorities, companies and organizations worldwide. It has been designed to be easily deployable and adaptable to diverse needs, significantly improving the accessibility and usability of time series data for end users, operators, system integrators, and maintainers. The digital ecosystem for FAIR time series data management has been designed under specific conditions and requirements and consists of three core components or independently usable systems:

(i) **Sensor Management System (SMS):** Facilitates the detailed registration of sensors with an internal and globally available persistent identifier (EUDAT B2INST[1]) and the management of sensor metadata using international standards (OGC SensorML[2]). Authentication is managed by the Helmholtz AAI[3], which allows all users from institutions connected to GÉANT EduGAIN[4] to log in. It also includes the option to use an alternative institutional or other identity provider system.

(ii) **Data Infrastructure (time.IO):** Provides the infrastructure for storing and managing time series data. Standard input protocols ((S-)FTP and MQTT[5]) are used for data ingestion, while the standardised OGC SensorThings API (OGC STA[6]) is used for data access. Authentication is handled by the Helmholtz AAI[3], enabling all users from institutions connected to GÉANT EduGAIN[4] to log in. Additionally, it offers the option to use an alternative institutional or other identity provider system.

(iii) **System for Automated Quality Control (SaQC):** Enables real-time analysis, annotation, and processing of data using predefined or custom quality schemes in real-time, including end-to-end metadata enrichment for each data point. This can be done using pre-configured standard procedures and user-configured methods for a variety of environmental variables.

The integration of the three primary components into a comprehensive digital system for sensor-based time series data (Fig. 1) is achieved through specific strategies. For example, SaQC is directly embedded within the time.IO orchestration framework as a dedicated subsystem operating in a separate container. In contrast, the connection between the Sensor Management System (SMS) and time.IO is established via a REST API. This modular design streamlines the integration process: time.IO orchestration inherently includes SaQC, while the linkage between an active SMS instance and time.IO is managed entirely through time.IO configuration settings.

The overall system includes all essential components, such as user-centric web-based front-ends for the sub-components, a versatile data integration layer, robust time-series database, efficient object storage, real-time quality control, and comprehensive real-time data visualization functions. It supports modern and legacy data transfer protocols ((S-)FTP and MQTT) and ensures compliance with OGC standards for data access and sensor metadata. In addition, the fully integrated containerized solution offers the convenience of rapid deployment and seamless integration with existing institutional services such as databases, identity providers, and object stores. The following sections will provide detailed descriptions of the three main components of the Digital Ecosystem for FAIR Time Series Data Management.

2.1. Sensor management system (SMS)

The Sensor Management System (SMS) is the foundation for sensor metadata handling. It allows detailed registration and management of sensors and specific measurement parameters, documenting changes over time. Additionally, the SMS supports the planning and management of complex measurement setups and campaigns over time. A controlled vocabulary ensures metadata consistency. Following the JSON:API[7] specification and adhering to international standards, such as OGC SensorML, SMS ensures that data sources are discoverable and accessible through standardized interfaces (Fig. 2). The SMS uses the EUDAT B2Inst identifier as its PID system to ensure metadata integrity throughout the data lifecycle. This is particularly important for integration into larger data infrastructures and for providing consistent metadata management. It uses a container-based deployment model for easy integration and scalability within existing IT infrastructures.

Its main features include:

- **Sensor Registration and management**: Allows for the detailed registration and management of sensors and specific measurement parameters, documenting changes over time. The permanent registration of the instruments is done with the globally available persistent identifier (EUDAT B2INST) to ensure metadata integrity throughout the data lifecycle.
- **Standard interfaces**: Uses JSON:API and follows international standards (OGC SensorML) for the provision of sensor metadata and measurement parameters.
- **Accessibility**: Ensures that all sensors with their metadata are findable and accessible through standardized interfaces.
- **Deployment**: Provides a container-based deployment model for easy integration and scalability within existing IT infrastructures.

For more details, see Brinckmann et al. [28] and/or https://hdl.handle.net/20.500.14372/SMS-Wiki

2.2. time.IO

Building on the foundation of sensor metadata using SMS, time.IO provides the infrastructure for storing and managing time series data. It supports the entire lifecycle of time series data, providing efficient data transfer and storage, real-time data visualisation using Grafana, and integrated data analysis and quality control with SaQC. The container-based deployment model facilitates easy integration and scalability within existing IT infrastructures, including seamless connection to geospatial infrastructures such as spatial.IO [29] for advanced spatial data analyses. time.IO also links to the SMS for consistent and standardised metadata management, ensuring a cohesive data management process (Fig. 3). For data access, the standardised OGC SensorThings API (OGC STA) is used and utilises the FROST-Server as a reference implementation for the OGC STA interface [30].

Its functionalities include:

- **Data transfer and storage**: Efficiently handles the transfer and storage of large volumes of time series data. time.IO provides transfer endpoints such as (S-)FTP servers or MQTT and supports the integration of existing data transfer infrastructures. For data access, the

[1] https://b2inst.gwdg.de/
[2] https://www.ogc.org/standard/sensorml/
[3] https://hifis.net/doc/helmholtz-aai/
[4] https://edugain.org/
[5] https://mqtt.org/
[6] https://www.ogc.org/standard/sensorthings/
[7] https://jsonapi.org/

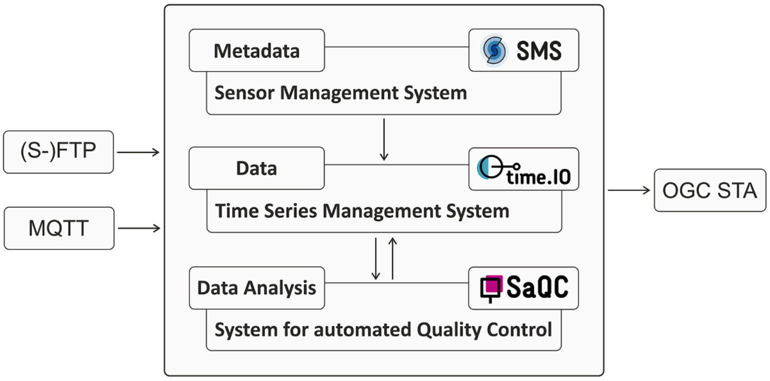
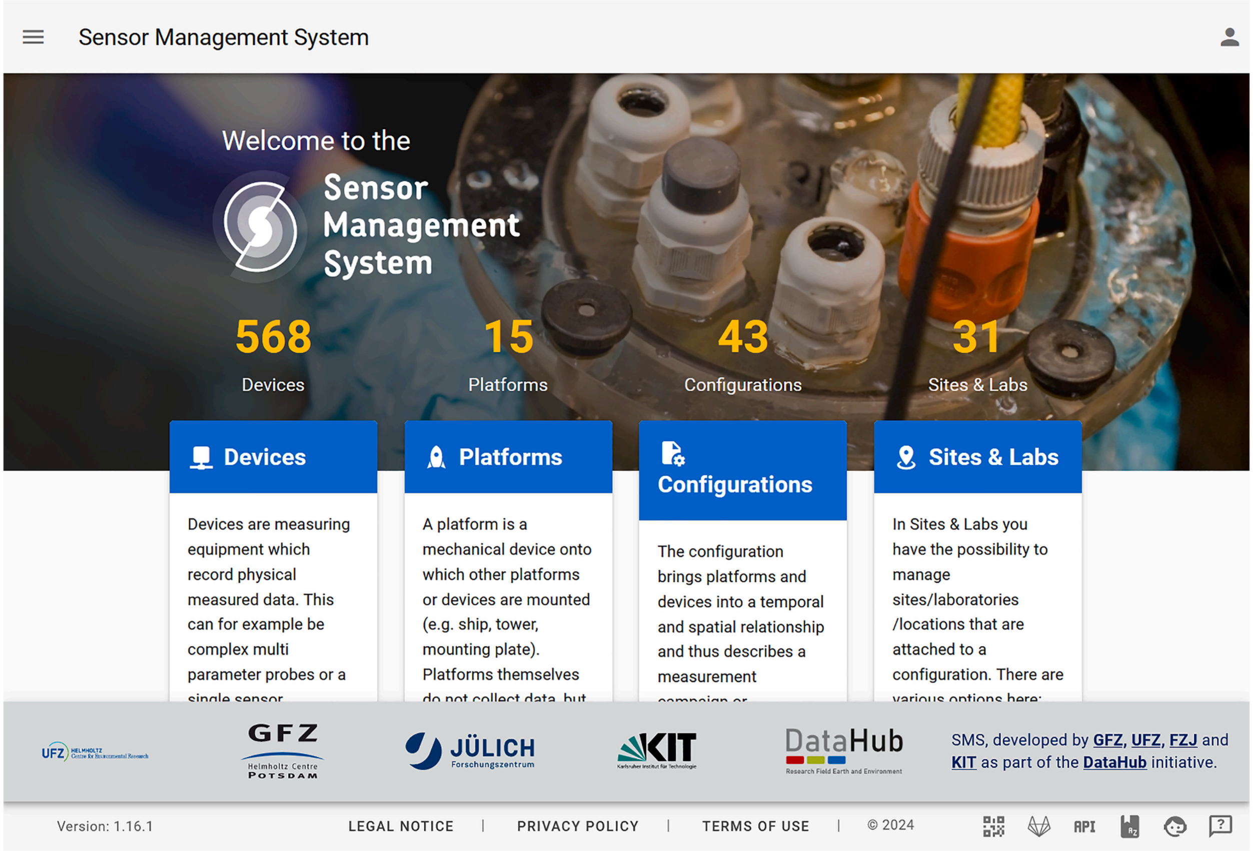


Fig. 1. Software architecture of the digital ecosystem for FAIR time series data management with the three main components for the management of sensor metadata (SMS), data (time.IO), and analysis or automated quality control of the data (SaQC).


Sensor Management System

Welcome to the Sensor Management System

568 Devices

15 Platforms

43 Configurations

31 Sites & Labs

Devices

Devices are measuring equipment which record physical measured data. This can for example be complex multi parameter probes or a single sensor

Platforms

A platform is a mechanical device onto which other platforms or devices are mounted (e.g. ship, tower, mounting plate). Platforms themselves do not collect data, but

Configurations

The configuration brings platforms and devices into a temporal and spatial relationship and thus describes a measurement campaign or

Sites & Labs

In Sites & Labs you have the possibility to manage sites/laboratories /locations that are attached to a configuration. There are various options here:

SMS, developed by GFZ, UFZ, FZJ and KIT as part of the DataHub initiative.

UFZ HELMHOLTZ Centre for Environmental Research
GFZ Helmholtz Centre Potsdam
JÜLICH Forschungszentrum
KIT Karlsruher Institut für Technologie
DataHub Research Field Earth and Environment
Version: 1.16.1
LEGAL NOTICE | PRIVACY POLICY | TERMS OF USE | © 2024
API

Fig. 2. Front-End SMS. The sensor management system (SMS) is designed to handle the acquisition and management of metadata from various sensors.

OGC SensorThings API is used, utilizing the FROST-Server as a reference implementation.

- **Data visualisation**: Uses Grafana[8] to provide real-time visualizations of time series data within automatically setup, preconfigured and shareable dashboards.
- **Quality control**: Integrates seamlessly with SaQC to ensure data quality and integrity.
- **Metadata management**: Uses the SMS for consistent and standardized metadata management.
- **Deployment**: Provides a container-based deployment model for easy integration and scalability within existing IT infrastructures.

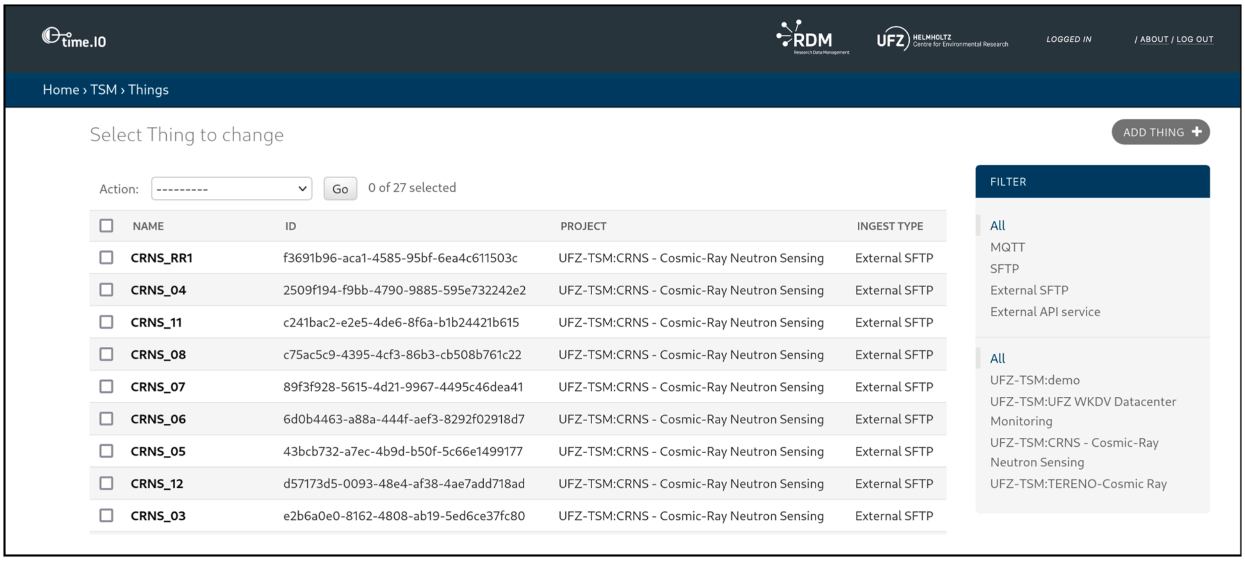
For more details, visit Schäfer et al. [31] and/or https://codebase.helmholtz.cloud/ufz-tsm

2.3. System for automated quality control (SaQC)

Completing the ecosystem, SaQC automates the quality control of time series data, improving traceability and reproducibility. It supports detailed data analysis based on a catalogue of state-of-the-art time series analysis, data processing, and annotation of data using predefined or custom quality schemes (Fig. 4). The meticulous collection of metadata throughout all operations ensures that resulting data meets quality standards such as traceability and reproducibility. This automated quality control is essential to maintain the integrity of data used in environmental research [32].

SaQC is designed to automate the quality control of time series data,

[8] https://grafana.com/

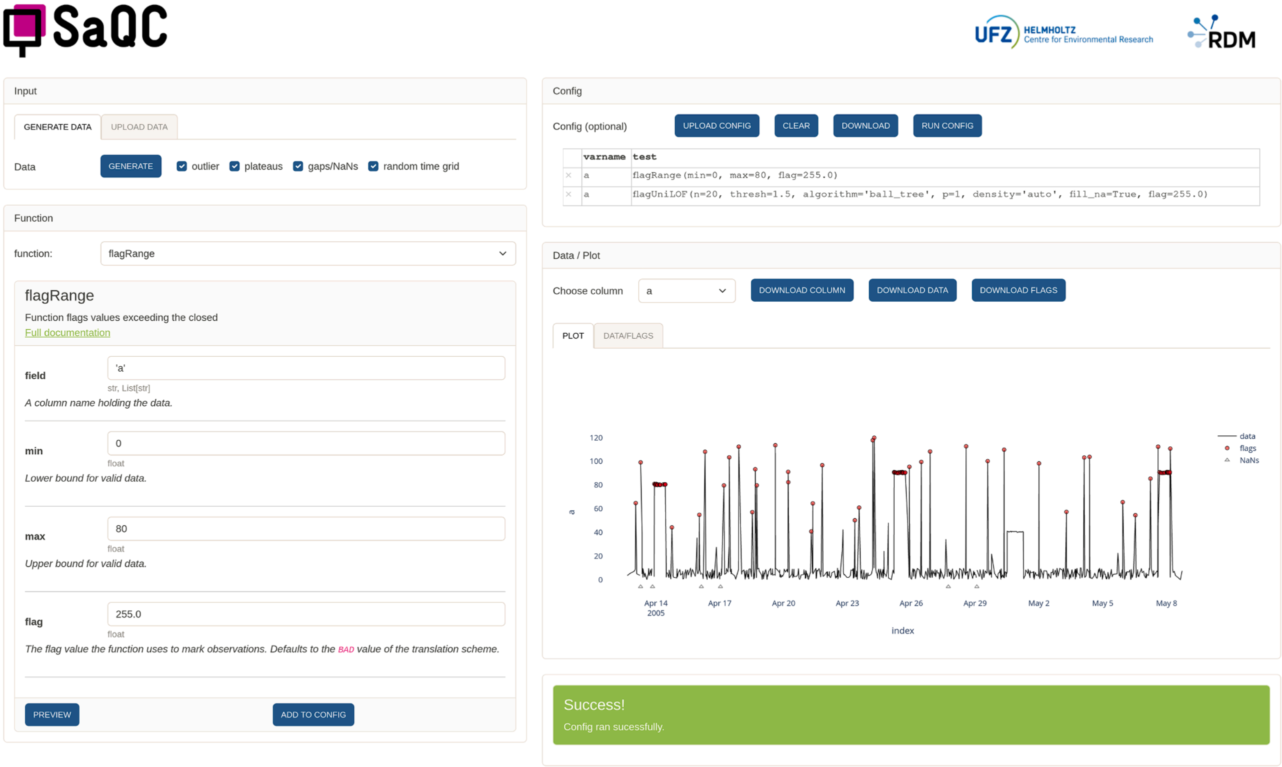

time.IO
RDM
UFZ HELMHOLTZ Centre for Environmental Research
LOGGED IN
/ ABOUT / LOG OUT
Home › TSM › Things
ADD THING
FILTER
All
MQTT
SFTP
External SFTP
External API service
All
UFZ-TSM:demo
UFZ-TSM:UFZ WKDV Datacenter Monitoring
UFZ-TSM:CRNS - Cosmic-Ray Neutron Sensing
UFZ-TSM:TERENO-Cosmic Ray

Select Thing to change

Action: --------- Go 0 of 27 selected

□	NAME	ID	PROJECT	INGEST TYPE
□	CRNS_RR1	f3691b96-aca1-4585-95bf-6ea4c611503c	UFZ-TSM:CRNS - Cosmic-Ray Neutron Sensing	External SFTP
□	CRNS_04	2509f194-f9bb-4790-9885-595e732242e2	UFZ-TSM:CRNS - Cosmic-Ray Neutron Sensing	External SFTP
□	CRNS_11	c241bac2-e2e5-4de6-8f6a-b1b24421b615	UFZ-TSM:CRNS - Cosmic-Ray Neutron Sensing	External SFTP
□	CRNS_08	c75ac5c9-4395-4cf3-86b3-cb508b761c22	UFZ-TSM:CRNS - Cosmic-Ray Neutron Sensing	External SFTP
□	CRNS_07	89f3f928-5615-4d21-9967-4495c46dea41	UFZ-TSM:CRNS - Cosmic-Ray Neutron Sensing	External SFTP
□	CRNS_06	6d0b4463-a88a-444f-aef3-8292f02918d7	UFZ-TSM:CRNS - Cosmic-Ray Neutron Sensing	External SFTP
□	CRNS_05	43bcb732-a7ec-4b9d-b50f-5c66e1499177	UFZ-TSM:CRNS - Cosmic-Ray Neutron Sensing	External SFTP
□	CRNS_12	d57173d5-0093-48e4-af38-4ae7add718ad	UFZ-TSM:CRNS - Cosmic-Ray Neutron Sensing	External SFTP
□	CRNS_03	e2b6a0e0-8162-4808-ab19-5ed6ce37fc80	UFZ-TSM:CRNS - Cosmic-Ray Neutron Sensing	External SFTP

Fig. 3. Front-End time.IO. The time.IO is developed to handle the entire lifecycle of sensor based data streams.



Data ☑ outlier ☑ plateaus ☑ gaps/NaNs ☑ random time grid

function: flagRange

flagRange

Function flags values exceeding the closed

Full documentation

field 'a'
str, List[str]
A column name holding the data.

min 0
float
Lower bound for valid data.

max 80
float
Upper bound for valid data.

flag 255.0
float
The flag value the function uses to mark observations. Defaults to the BAD value of the translation scheme.

Config (optional)

	varname	test
×	a	flagRange(min=0, max=80, flag=255.0)
×	a	flagUniLOF(n=20, thresh=1.5, algorithm='ball_tree', p=1, density='auto', fill_na=True, flag=255.0)

Choose column a

Success!
Config ran sucessfully.

Fig. 4. Front-end SAQC (GUI): Users can add tests for their data in the panel on the left and see the resulting flags on the right side. These configurations can then be used for automated testing, allowing for continuous monitoring and quality assurance of the data.

improving traceability and reproducibility. Key features include:

- **Data Analysis:** Provides state of the art algorithms for detailed analysis of time series data.
- **Data Processing:** Exposes a large set of data processing features
- **Data Annotation**: Supports the annotation of time series data using predefined or custom quality schemes.
- **End-to-end metadata enrichment**: Enrich metadata from initial data collection to final use, ensuring metadata integrity throughout the data lifecycle.
- **User interfaces**: Provides flexibility through a Python API, text-based configuration, and a web application.

For more details, visit Schmidt et al. [32], Schäfer et al. [33] and/or https://git.ufz.de/rdm-software/saqc.

3. Illustrative example

The presented toolchain encompasses every stage of the typical sensor data lifecycle through user-friendly frontends, which include (i) metadata registration during sensor deployment, (ii) the setup of data transmission, including the automatic generation of endpoints, accounts, and credentials, (iii) real-time monitoring and provisioning of incoming data streams, and (iv) the configuration and parameterization of quality control pipelines. This self-service approach is designed specifically for data producers, data managers, and technicians, abstracting the technical complexities of underlying technologies such as data transfer and storage (see Fig. 5).

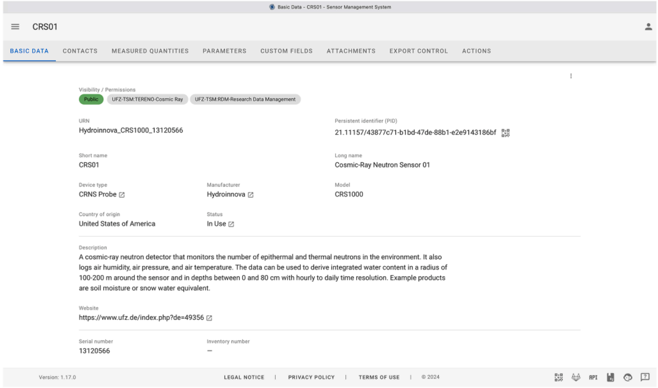
3.1. *Cosmic ray neutron sensing (CRNS)*

CRNS - Cosmic Ray Neutron Sensing is a method for determining average soil moisture (non-contact technology) for areas of about 5-15 hectares [34]. This novel approach to soil moisture measurement is one of our reference use cases for time.IO and well established in the TERENO community.

A typical usage pattern begins with registering the CRNS sensor's metadata in the Sensor Management System, where detailed information about the device – such as type, manufacturer, model, and the quantities it measures, including air temperature, air pressure, and neutron counts – is provided (Fig. 6). Next, a 'Thing' is created in the time.IO frontend. The term 'Thing' is derived from the OGC SensorThings API data model and is best described as a data transmission unit. In practice, a Thing often corresponds directly to a data logger or sensor. The user is required to input information such as hostnames for (S-)FTP servers or MQTT brokers, as well as account settings and credentials (Fig. 7). Once the data transmission unit is operational, data monitoring can be conducted through Grafana. All incoming data becomes immediately visible in automatically set up and pre-configured dashboards. Subsequently, SaQC can be configured within the time.IO frontend to apply quality control and data processing directly within the data stream. Finally, data and metadata are provided both through the built-in Grafana dashboard (Fig. 8) for human users and via the OGC standard SensorThings API for machine-to-machine data exchange (Fig. 9).

4. Impact and discussion

This section provides a concise evaluation of the ecosystem's implications, divided into two primary dimensions: its limitations and constraints, and its broader impact on environmental research and practice. This approach highlights the system's strengths and areas for improvement, offering a balanced perspective for current and potential adopters.



Basic Data - CRS01 - Sensor Management System

CRS01

BASIC DATA | CONTACTS | MEASURED QUANTITIES | PARAMETERS | CUSTOM FIELDS | ATTACHMENTS | EXPORT CONTROL | ACTIONS

Visibility / Permissions
Public | UFZ-TSM:TERENO-Cosmic Ray | UFZ-TSM:RDM-Research Data Management

URN
Hydroinnova_CRS1000_13120566

Persistent identifier (PID)
21.11157/43877c71-b1bd-47de-88b1-e2e9143186bf

Short name
CRS01

Long name
Cosmic-Ray Neutron Sensor 01

Device type
CRNS Probe

Manufacturer
Hydroinnova

Model
CRS1000

Country of origin
United States of America

Status
In Use

Description
A cosmic-ray neutron detector that monitors the number of epithermal and thermal neutrons in the environment. It also logs air humidity, air pressure, and air temperature. The data can be used to derive integrated water content in a radius of 100-200 m around the sensor and in depths between 0 and 80 cm with hourly to daily time resolution. Example products are soil moisture or snow water equivalent.

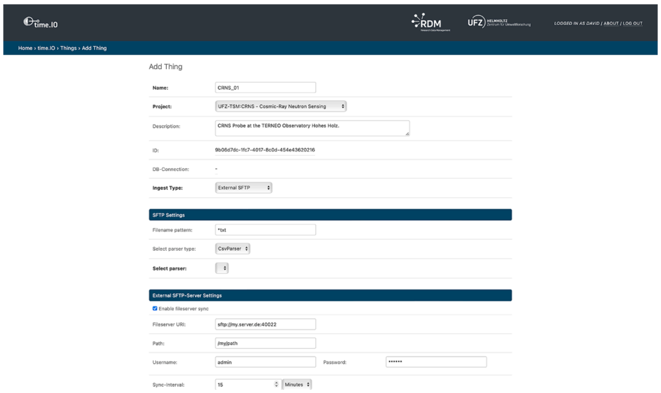
Website
https://www.ufz.de/index.php?de=49356

Serial number
13120566

Inventory number
—

Version: 1.17.0 LEGAL NOTICE | PRIVACY POLICY | TERMS OF USE | © 2024

Fig. 6. SMS – Manage sensor metadata.



LOGGED IN AS DAVID / ABOUT / LOG OUT

Home › time.IO › Things › Add Thing

Add Thing

Name: CRNS_01
Project: UFZ-TSM:CRNS - Cosmic-Ray Neutron Sensing
Description: CRNS Probe at the TERNEO Observatory Hohes Holz.
ID: 9b06d7dc-1fc7-4017-8c0d-454e43620216
DB-Connection: -
Ingest Type: External SFTP

SFTP Settings
Filename pattern: *txt
Select parser type: CsvParser
Select parser:

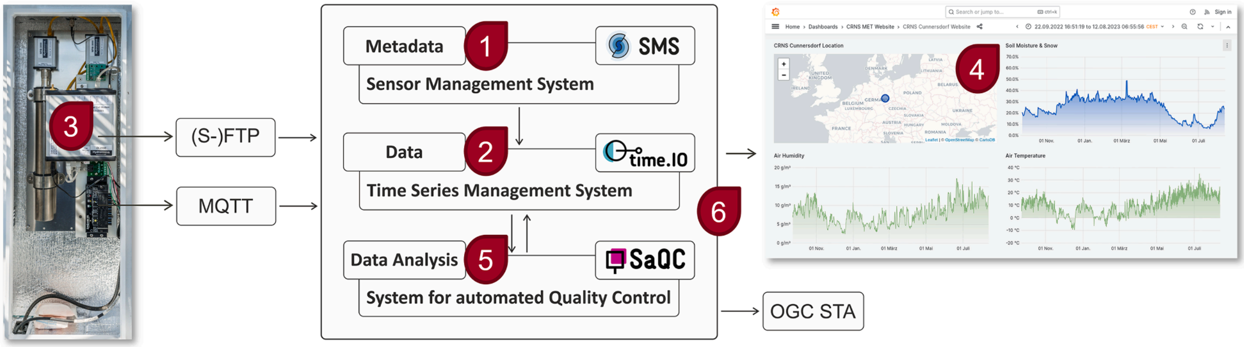
External SFTP-Server Settings
☑ Enable fileserver sync
Fileserver URI: sftp://my.server.de:40022
Path: /my/path
Username: admin Password: ******
Sync-Interval: 15 Minutes

Fig. 7. time.IO – Configure dataflows.

4.1. *Limitations and constraints*

The ecosystem emphasizes strict data integrity by employing a design that stores both raw and processed data, ensuring that raw data remains unaltered. This approach inevitably increases storage requirements, particularly when managing large datasets. Despite the added storage demands, the system prioritizes data integrity and persistence over minimizing storage usage, underscoring its commitment to data reliability.

The integration of advanced in-stream data processing



1. Register metadata with **SMS**
2. Setup Thing with **timeIO**
3. Configure logger and start data transfer
4. Monitor datastreams (**timeIO/Grafana**)
5. Setup quality control and attach custom processing steps (**SaQC**)
6. Share dashboards and data (**timeIO/Grafana/OGC STA**)

Fig. 5. The toolchain for implementing a CRNS sensor into the digital ecosystem and the necessary workflow steps.

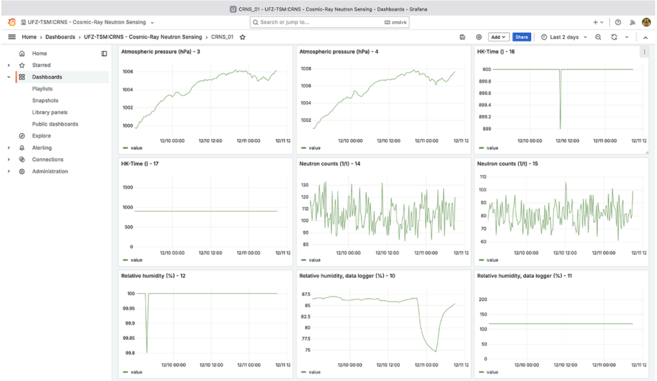
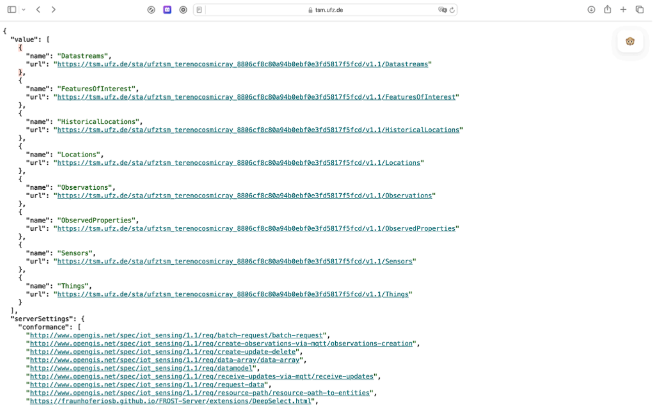

CRNS_01 - UFZ-TSM:CRNS - Cosmic-Ray Neutron Sensing - Dashboards - Grafana
UFZ-TSM:CRNS - Cosmic-Ray Neutron Sensing
Search or jump to...
cmd+k
Home › Dashboards › UFZ-TSM:CRNS - Cosmic-Ray Neutron Sensing › CRNS_01
Add
Share
Last 2 days
Home
Starred
Dashboards
Playlists
Snapshots
Library panels
Public dashboards
Explore
Alerting
Connections
Administration
Atmospheric pressure (hPa) - 3
1006
1004
1002
1000
Atmospheric pressure (hPa) - 4
1008
1006
1004
1002
HK-Time () - 16
900
899.8
899.6
899.4
899.2
899
HK-Time () - 17
1500
1000
500
0
Neutron counts (1/t) - 14
130
120
110
100
90
80
Neutron counts (1/t) - 15
110
100
90
80
70
60
Relative humidity (%) - 12
100
99.95
99.9
99.85
99.8
Relative humidity, data logger (%) - 10
87.5
85
82.5
80
77.5
75
Relative humidity, data logger (%) - 11
200
150
100
50
0
12/10 00:00
12/10 12:00
12/11 00:00
12/11 12
value

Fig. 8. Grafana – Visualize sensor data.



```
{
  "value": [
    {
      "name": "Datastreams",
      "url": "https://tsm.ufz.de/sta/ufztsm_terenocosmicray_8806cf8c80a94b0ebf0e3fd5817f5fcd/v1.1/Datastreams"
    },
    {
      "name": "FeaturesOfInterest",
      "url": "https://tsm.ufz.de/sta/ufztsm_terenocosmicray_8806cf8c80a94b0ebf0e3fd5817f5fcd/v1.1/FeaturesOfInterest"
    },
    {
      "name": "HistoricalLocations",
      "url": "https://tsm.ufz.de/sta/ufztsm_terenocosmicray_8806cf8c80a94b0ebf0e3fd5817f5fcd/v1.1/HistoricalLocations"
    },
    {
      "name": "Locations",
      "url": "https://tsm.ufz.de/sta/ufztsm_terenocosmicray_8806cf8c80a94b0ebf0e3fd5817f5fcd/v1.1/Locations"
    },
    {
      "name": "Observations",
      "url": "https://tsm.ufz.de/sta/ufztsm_terenocosmicray_8806cf8c80a94b0ebf0e3fd5817f5fcd/v1.1/Observations"
    },
    {
      "name": "ObservedProperties",
      "url": "https://tsm.ufz.de/sta/ufztsm_terenocosmicray_8806cf8c80a94b0ebf0e3fd5817f5fcd/v1.1/ObservedProperties"
    },
    {
      "name": "Sensors",
      "url": "https://tsm.ufz.de/sta/ufztsm_terenocosmicray_8806cf8c80a94b0ebf0e3fd5817f5fcd/v1.1/Sensors"
    },
    {
      "name": "Things",
      "url": "https://tsm.ufz.de/sta/ufztsm_terenocosmicray_8806cf8c80a94b0ebf0e3fd5817f5fcd/v1.1/Things"
    }
  ],
  "serverSettings": {
    "conformance": [
      "http://www.opengis.net/spec/iot_sensing/1.1/req/batch-request/batch-request",
      "http://www.opengis.net/spec/iot_sensing/1.1/req/create-observations-via-mqtt/observations-creation",
      "http://www.opengis.net/spec/iot_sensing/1.1/req/create-update-delete",
      "http://www.opengis.net/spec/iot_sensing/1.1/req/data-array/data-array",
      "http://www.opengis.net/spec/iot_sensing/1.1/req/datamodel",
      "http://www.opengis.net/spec/iot_sensing/1.1/req/receive-updates-via-mqtt/receive-updates",
      "http://www.opengis.net/spec/iot_sensing/1.1/req/request-data",
      "http://www.opengis.net/spec/iot_sensing/1.1/req/resource-path/resource-path-to-entities",
      "https://fraunhoferiosb.github.io/FROST-Server/extensions/DeepSelect.html",
```

Fig. 9. STA – Access sensor data and metadata.

functionalities introduces potential challenges, such as increased server loads, particularly in scenarios involving large datasets or complex processing workflows. To address these challenges, the ecosystem supports horizontal scaling of the time.IO component, allowing dynamic resource allocation to mitigate bottlenecks. The ability to represent complete data flows from source to end-user within a unified software system is considered sufficient justification for the associated hardware adjustments.

The ecosystem does not include integrated backup and recovery mechanisms. Instead, it relies on well-established and widely used software systems, namely MiniIO and PostgreSQL, for data storage. These systems provide robust backup and failure management functionalities. Consequently, users are encouraged to leverage these built-in mechanisms to fulfill their backup and recovery requirements.

4.2. *Impact*

The digital ecosystem for managing time series data has significantly advanced research and practical applications in environmental system science. The integration of SMS, time.IO, and SaQC within this digital ecosystem has opened up new research opportunities, particularly in real-time applications and data visualization. Automated, real-time quality control and immediate data viewers allow researchers to monitor and respond to environmental changes as they occur. This capability supports the study of dynamic events and enhances iterative research designs, enabling precise exploration of complex environmental interactions. The system's standardized framework also facilitates cross-disciplinary research by seamlessly integrating diverse sensor data, broadening the scope of environmental studies.

Daily practices have evolved with the system's deployment. Automated workflows reduce manual data management, allowing researchers to focus on analysis rather than logistics. User-friendly interfaces make advanced data management accessible, boosting productivity and enabling more efficient handling of large datasets.

The system's modular and cloud-ready architecture ensures its broad applicability and transferability across diverse research environments. By adopting standardized methods and protocols, the ecosystem enables consistent and reliable data management practices, making it possible to deploy identical infrastructures across different institutions. This standardization facilitates not only consistent workflows and data sharing but also seamless integration with geospatial infrastructures such as spatial.IO [29], thereby enhancing spatial analyses, interdisciplinary collaboration, and the comparability of research outcomes globally. The system's versatility ensures that it can be effectively utilized in various scales of research, from localized projects to large-scale monitoring networks, driving innovation and efficiency in environmental data management.

5. Conclusions

The integration of the Sensor Management System (SMS), time.IO for storage, transfer, and real-time visualization, and the System for Automated Quality Control (SaQC) represents a significant advancement in the field of environmental data management. This digital ecosystem not only ensures the comprehensive management of time series data but also enhances data integrity, accessibility, and usability. By combining data acquisition, temporal alignment, and real-time quality control, the system supports robust environmental research and informed policy-making. Additionally, its capabilities to facilitate real-time applications, dynamic event monitoring, and cross-disciplinary research significantly broaden its impact in advancing environmental system science.

The modular, cloud-ready architecture and use of standardized protocols make the ecosystem highly adaptable and transferable across various research environments. This flexibility allows for consistent data management practices across institutions, fostering collaboration and enabling the comparability of research outcomes. Its integration into larger infrastructures, such as digital twins and environmental models, highlights its potential to transform data-driven research and decision-making processes.

The digital ecosystem's ability to handle increasing data volumes and provide real-time insights supports emerging research needs, particularly in the context of climate change and biodiversity loss. It can be effectively deployed in both small-scale projects and large-scale monitoring networks, making it a versatile tool for advancing environmental science. By addressing current gaps in data management, it enables researchers to adopt innovative approaches and respond dynamically to environmental changes, fostering the development of effective adaptation strategies.

In summary, this digital ecosystem offers a powerful, standardized solution for managing environmental sensor data, supporting the pursuit of new research questions and improving the efficiency and reliability of existing studies. Its widespread adoption will likely drive further innovation in environmental monitoring and data management, benefiting a broad range of stakeholders in both scientific and practical contexts. The system's capacity to support iterative research designs, improve interdisciplinary collaboration, and provide actionable insights marks a substantial step forward in addressing global environmental challenges.

CRediT authorship contribution statement

J. Bumberger: Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Conceptualization. **M. Abbrent:** Writing – review & editing, Software, Methodology, Conceptualization. **N. Brinckmann:** Writing – review & editing, Software, Methodology, Conceptualization. **J. Hemmen:** Writing – review &

editing, Software, Methodology, Conceptualization. **R. Kunkel:** Writing – review & editing, Supervision, Methodology, Conceptualization. **C. Lorenz:** Writing – review & editing, Supervision, Software, Methodology, Conceptualization. **P. Lünenschloss:** Writing – review & editing, Software, Methodology, Conceptualization. **B. Palm:** Writing – review & editing, Software, Methodology, Conceptualization. **T. Schnicke:** Writing – review & editing, Supervision, Methodology, Conceptualization. **C. Schulz:** Writing – review & editing, Software, Methodology, Conceptualization. **H. van der Schaaf:** Writing – review & editing, Supervision, Methodology, Conceptualization. **D. Schäfer:** Writing – review & editing, Writing – original draft, Supervision, Software, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that there are no conflicts of interest relevant to this work.

Acknowledgements

We thank the Helmholtz Association and the Federal Ministry of Education and Research (BMBF) for supporting the DataHub initiative of the Research Field Earth and Environment. The DataHub enables an overarching and comprehensive research data management, according to the FAIR principles, for all topics of the programme Changing Earth – Sustaining our Future.

The publication is a contribution in the context of the work of the NFDI4Earth – the consortium for the Earth System Sciences within the German National Research Data Infrastructure (NFDI) e. V.. The NFDI is financed by the Federal Republic of Germany and its 16 federal states, and the NFDI4Earth is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number: 460036893. The authors would like to express their gratitude for the funding and support.

This research has been supported by the European Commission, Horizon 2020 (H2020 Project eLTER PPP, grant no. 871126, eLTER PLUS, grant no. 871128 and eLTER EnRich, grant no. 101131751) in the context of the development of the eLTER cyberinfrastructure.

Appendix 1

Criterion	Description	Implementation in the Paper
To be Findable		
F1	(meta)data are assigned a globally unique and eternally persistent identifier.	Persistent identifiers (e.g., EUDAT B2INST) are used for sensor metadata, ensuring global uniqueness and persistence. SaQC workflows are version-controlled via Git, providing traceability for all processing steps.
F2	Data are described with rich metadata.	Detailed metadata is managed using international standards such as OGC SensorML in the Sensor Management System (SMS). SaQC enriches metadata during processing, adding details on quality flags, QC-tests performed, and versioning of workflows.
F3	(meta)data are registered or indexed in a searchable resource.	Metadata and data are indexed via searchable platforms enabled by SMS and time.IO components. SaQC outputs structured data (e.g., .csv or .parquet) with detailed metadata, facilitating integration into searchable systems.
F4	Metadata specify the data identifier.	Each dataset's metadata includes specific identifiers, ensuring clear referencing and traceability. SaQC metadata links data points to their processing history, including information on applied QC tests and configurations.
To be Accessible		
A1	(meta)data are retrievable by their identifier using a standardized communications protocol.	Standardized OGC SensorThings API, MQTT, and (S-)FTP protocols are utilized for data retrieval and transfer. SaQC enables consistent, accessible QC workflows, ensuring retrievability through standardized outputs.
A1.1	The protocol is open, free, and universally implementable.	Open protocols like MQTT and OGC SensorThings API are implemented, ensuring universal accessibility. SaQC is open-source, available via Python Package Index (PyPI) and Git repositories.
A1.2	The protocol allows for an authentication and authorization procedure, where necessary.	Authentication is managed using Helmholtz AAI, supporting institutional and cross-institutional logins via EduGAIN. SaQC ensures that processing steps are documented, enabling access control to data derived from authenticated workflows.
A2	Metadata are accessible, even when the data are no longer available.	Metadata management ensures independent accessibility, maintained via the Sensor Management System (SMS). SaQC preserves metadata for all processed datasets, even if raw data is no longer available.
To be Interoperable		
I1	(meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.	The system adheres to standards like OGC SensorML for knowledge representation. SaQC's configuration uses simple, standardized syntax (e.g., YAML-like text files) for QC tests and workflows, ensuring accessibility and compatibility.
I2	(meta)data use vocabularies that follow FAIR principles.	Vocabularies conforming to FAIR standards, integrated within the system, ensure interoperability and standardization. SaQC uses customizable and standardized flagging schemes to represent data quality, enabling integration with other FAIR systems.
I3	(meta)data include qualified references to other (meta)data.	Metadata links within the system connect data points, provenance, and related datasets using qualified references. SaQC records and references data provenance explicitly, including information about QC tests and their results.
To be Re-usable		
R1	(meta)data have a plurality of accurate and relevant attributes.	SMS and SaQC manage attributes accurately, enriching metadata during all stages of data handling. SaQC allows for detailed annotations of each data point, including quality flags and test results.
R1.1	(meta)data are released with a clear and accessible data usage license.	Licenses (e.g., EUPL-1.2 for SMS, GNU GPL 3.0 for SaQC) are clearly specified for software and data.
R1.2	(meta)data are associated with their provenance.	Provenance is documented throughout the data lifecycle, supported by SMS and SaQC for end- to-end metadata enrichment. SaQC maintains a complete log of all processing steps and metadata changes, ensuring full traceability.
R1.3	(meta)data meet domain-relevant community standards.	Domain standards, such as OGC SensorML, ensure compatibility with environmental and geoscience communities. SaQC's flexibility supports customization to meet community-specific QC needs, while adhering to widely accepted standards.

References

[1] Stocker TF, et al. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press; 2014. https://doi.org/10.1017/CBO9781107415324.

[2] Gupta J, Liverman D, Prodani K, Aldunce P, Bai X, Broadgate W, Ciobanu D, Gifford L, Gordon C, Hurlbert M, Inoue C, Jacobson L, Kanie N, Lade S, Lenton T, Obura D, Okereke C, Otto I, Pereira L, Verburg P. Earth system justice needed to identify and live within Earth system boundaries. Nat Sustain 2023;6(6):630–8. https://doi.org/10.1038/s41893-023-01064-1.

[3] Loescher HW, Kelly EF, Lea R. National ecological observatory network: Beginnings, programmatic and scientific challenges, and ecological forecasting. Terrestrial Ecosystem Research Infrastructures. CRC Press; 2017. p. 27–52. https://doi.org/10.1201/9781315368252.

[4] Mollenhauer H, Kasner M, Haase P, Peterseil J, Wohner C, Frenzel M, Mirtl M, Schima R, Bumberger J, Zacharias S. Long-term environmental monitoring infrastructures in Europe: Observations, measurements, scales, and socio-ecological representativeness. Sci. Total Environ 2018;624:968–78. https://doi.org/10.1016/j.scitotenv.2017.12.095.

[5] Ohnemus T, Zacharias S, Dirnböck T, Bäck J, Brack W, Forsius M, Mallast U, Nikolaidis NP, Peterseil J, Piscart C, Pando F, Terán CP, Mirtl M. The eLTER research infrastructure: Current design and coverage of environmental and socio-ecological gradients. Environ Sustain Indic 2024;23:100456. https://doi.org/10.1016/j.indic.2024.100456.

[6] Zacharias S, Bogena H, Samaniego L, Mauder M, Fuß R, Pütz T, Frenzel M, Schwank M, Baessler C, Butterbach-Bahl K, Bens O, Borg E, Brauer A, Dietrich P, Hajnsek I, Helle G, Kiese R, Kunstmann H, Klotz S, Munch JC, Papen H, Priesack E, Schmid HP, Steinbrecher R, Rosenbaum U, Teutsch G, Vereecken H. A network of terrestrial environmental observatories in Germany. Vadose Zone J 2011;10(4): 955–73. https://doi.org/10.2136/vzj2010.0139.

[7] Zacharias S, Loescher HW, Bogena H, Kiese R, Schrön M, Attinger S, Blume T, Borchardt D, Borg E, Bumberger J, Chwala C, Dietrich P, Fersch B, Frenzel M, Gaillardet J, Groh J, Hajnsek I, Itzerott S, Kunkel R, Kunstmann H, Kunz M, Liebner S, Mirtl M, Montzka C, Musolff A, Pütz T, Rebmann C, Rinke K, Rode M, Sachs T, Samaniego L, Schmid HP, Vogel H-J, Weber U, Wollschläger U, Vereecken H. Fifteen years of integrated terrestrial environmental observatories (TERENO) in Germany: Functions, services, and lessons learned. Earth's Fut 2024; 12(6). https://doi.org/10.1029/2024EF004510.

[8] Weber U, Attinger S, Baschek B, Boike J, Borchardt D, Brix H, Brüggemann N, Bussmann I, Dietrich P, Fischer P, Greinert J, Hajnsek I, Kamjunke N, Kerschke D, Kiendler-Scharr A, Körtzinger A, Kottmeier C, Merz B, Merz R, Riese M, Schloter M, Schmid H, Schnitzler J, Sachs T, Schütze C, Tillmann R, Vereecken H, Wieser A, Teutsch G. MOSES: A novel observation system to monitor dynamic events across Earth compartments. Bull Am Meteorol Soc 2022;103(2):E339–48. https://doi.org/10.1175/BAMS-D-20-0158.1.

[9] Loescher HW, Vargas R, Mirtl M, Morris B, Pauw J, Yu X, et al. Building a global ecosystem research infrastructure to address global grand challenges for macrosystem ecology. Earth's Fut. 2022;10:e2020EF001696. https://doi.org/10.1029/2020EF001696.

[10] Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJ, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PA, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone SA, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B. The FAIR guiding principles for scientific data management and stewardship. Sci Data 2016;3(1): 160018. https://doi.org/10.1038/sdata.2016.18.

[11] Zhang Y, Li J, Duan M, Chen W, Rio J, Xiang Z, Wang K, Liang S, Chen Z, Chen N, Di L, Hu C. Multi-sensor integration management in the earth observation sensor web: State-of-the-art and research challenges. Int J Appl Earth Observ Geoinform 2023;125:103601. https://doi.org/10.1016/j.jag.2023.103601.

[12] Zhang X, Chen N, Chen Z, Wu L, Li X, Zhang L, Di L, Gong J, Li D. Geospatial sensor web: A cyber-physical infrastructure for geoscience research and application. Earth Sci Rev 2018;185. https://doi.org/10.1016/j.earscirev.2018.07.00611.

[13] Li X, Feng M, Ran Y, Su Y, Liu F, Huang C, Shen H, Xiao Q, Su J, Yuan S, Guo H. Big data in Earth system science and progress towards a digital twin. Nature Rev Earth Environ 2023;4:319–32. https://doi.org/10.1038/s43017-023-00409-w.

[14] Selsam P, Bumberger J, Wellmann T, Pause M, Gey R, Borg E, Lausch A. Ecosystem integrity remote sensing—Modelling and service tool—ESIS/Imalys. Remote Sens (Basel) 2024;16(7):1139. https://doi.org/10.3390/rs16071139.

[15] Hazeleger W, Aerts J, Bauer P, Bierkens M, Camps-Valls G, Dekker M, Doblas F, Eyring V, Finkenauer C, Grundner A, Hachinger S, Hall D, Hartmann T, Iglesias-Suarez F, Janssens M, Jones E, Kölling T, Lees M, Lhermitte S, Vossepoel F. Digital twins of the Earth with and for humans. Commun Earth Environ 2024;5. https://doi.org/10.1038/s43247-024-01626-x.

[16] Hart J, Martinez K. Environmental sensor networks: A revolution in the Earth system science? Earth Sci Rev 2006;78:177–91. https://doi.org/10.1016/j.earscirev.2006.05.001.

[17] Horsburgh J, Tarboton D, Maidment D. Components of an environmental observatory information system. Comput Geosci 2011;37:207–18. https://doi.org/10.1016/j.cageo.2010.07.003.

[18] Dow AK, Dow EM, Fitzsimmons TD, Materise MM. Harnessing the environmental data flood: A comparative analysis of hydrologic, oceanographic, and meteorological informatics platforms. Bull Am Meteorol Soc 2015;96:725–36. https://doi.org/10.1175/BAMS-D-13-00178.1.

[19] Mons B, et al. Cloudy, increasingly FAIR; revisiting the FAIR data guiding principles for the European open science cloud. Inf Serv Use 2017;37(1):49–56. https://doi.org/10.3233/ISU-170824.

[20] Koedel U, Schuetze C, Fischer P, Bussmann I, Sauer PK, Nixdorf E, Kalbacher T, Wichert V, Rechid D, Bouwer LM, Dietrich P. Challenges in the evaluation of observational data trustworthiness from a data producer's viewpoint (FAIR+). Front Environ Sci 2022;9:772666. https://doi.org/10.3389/fenvs.2021.772666.

[21] Horsburgh J, Tarboton D, Piasecki M, Maidment D, Zaslavsky I, Valentine D, Whitenack T. An integrated system for publishing environmental observations data. Environmental Modelling & Software 2009;24:879–88. https://doi.org/10.1016/j.envsoft.2009.01.002.

[22] Kunkel R, Sorg J, Eckardt R, Kolditz O, Rink K, Vereecken H. TEODOOR: A distributed geodata infrastructure for terrestrial observation data. Environ Earth Sci 2013;69. https://doi.org/10.1007/s12665-013-2370-7.

[23] Braud I, Chaffard V, Coussot C, Galle S, Juen P, Alexandre H, Baillion P, Battais A, Boudevillain B, Branger F, Brissebrat G, Cailletaud R, Cochonneau G, Decoupes R, Desconnets J-C, Dubreuil A, Fabre J, Gabillard S, Gérard M-F, Squividant H. Building the information system of the French critical zone observatories network: Theia/OZCAR-IS. Hydrolog Sci J 2020;67. https://doi.org/10.1080/02626667.2020.1764568.

[24] Bastidas Pacheco CJ, Brewer J, Horsburgh J, Caraballo J. An open source cyberinfrastructure for collecting, processing, storing and accessing high temporal resolution residential water use data. Environ Mod Softw 2021;144:105137. https://doi.org/10.1016/j.envsoft.2021.105137.

[25] Abu-Elkheir M, Hayajneh M, Ali NA. Data management for the Internet of Things: Design primitives and solution. Sensors 2013;13(11):15582–612. https://doi.org/10.3390/s131115582.

[26] Diene B, Rodrigues J, Diallo O, Ndoye M, Korotaev V. Data management techniques for Internet of Things. Mech Syst Signal Process 2019;138. https://doi.org/10.1016/j.ymssp.2019.106564.

[27] Barker M, Chue Hong NP, Katz DS, Lamprecht AL, Martinez-Ortiz C, Psomopoulos F, Harrow J, Castro LJ, Gruenpeter M, Martinez PA, Honeyman T. Introducing the FAIR principles for research software. Sci Data 2022;9(1):622. https://doi.org/10.1038/s41597-022-01710-x.

[28] Brinckmann N, Alhaj Taha K, Kuhnert T, Abbrent M, Becker W, Bohring H, Breier J, Bumberger J, Ecker D, Eder T, Gransee F, Hanisch M, Lorenz C, Moorthy R, Nendel LJ, Pongratz E, Remmler P, Rosin V, Schaeffer M, Schaldach M, Schäfer D, Sielaff D, Ziegner N. Sensor management system - SMS (1.17.1). Zenodo 2024. https://doi.org/10.5281/zenodo.13329925.

[29] Schulz C, Lange R, Schnicke T, Bumberger J. spatial.IO - An integrated cloud-ready geospatial data management system (0.8.0). Zenodo 2024. https://doi.org/10.5281/zenodo.10391523.

[30] van der Schaaf, H., Moßgraber, J., Grellet, S., Beaufils, M., Schleidt, K., & Usländer, T. (2020). An environmental sensor data suite using the OGC SensorThings API. In I. Athanasiadis, S. Frysinger, G. Schimak, & W. Knibbe (Eds.), *Environmental Software Systems. Data Science in Action* (Vol. 554). Springer. https://doi.org/10.1007/978-3-030-39815-6_22.

[31] Schäfer D, Abbrent M, Gransee F, Kuhnert T, Hemmen J, Nendel L, Palm B, Schaldach M, Schulz C, Schnicke T, Bumberger J. timeIO - A fully integrated and comprehensive timeseries management system (0.1). Zenodo 2023. https://doi.org/10.5281/zenodo.8354839.

[32] Schmidt L, Schäfer D, Geller J, Lünenschloss P, Palm B, Rinke K, Rebmann C, Rode M, Bumberger J. System for automated Quality Control (SaQC) to enable traceable and reproducible data streams in environmental science. Environ Modell Softw 2023:105809. https://doi.org/10.1016/j.envsoft.2023.105809.

[33] Schäfer D, Palm B, Lünenschloß P, Schmidt L, Schnicke T, Bumberger J. System for automated Quality Control - SaQC (v2.6.0). Zenodo 2024. https://doi.org/10.5281/zenodo.5888547.

[34] Köhli M, Schrön M, Zreda M, Schmidt U, Dietrich P, Zacharias S. Footprint characteristics revised for field-scale soil moisture monitoring with cosmic-ray neutrons. Water Resour Res 2015;51(7):5772–90. https://doi.org/10.1002/2015WR017169.