

Internet of Water Principles

Updated October 2021

1. Water data are essential for efficient, equitable, sustainable, and resilient water planning, management, and stewardship.
2. Modern data infrastructure increases the usefulness of water data and enables its broadest possible application.
3. All water data produced for the public good should, by default, be findable, accessible, interoperable, and reusable (FAIR) for public use or authorized users.
4. Security and privacy risks associated with sharing data can be mitigated using mechanisms for tiered access for authorized users.
5. Commonly accepted data, metadata, and exchange standards should be adopted by water data producers to promote interoperability, efficiency, sharing, and secondary uses of data
6. Control and responsibility over data are best maintained by data producers.
7. Data producers are responsible for sharing data of known quality and documenting essential metadata; data users are responsible for determining whether data are appropriate for specific purposes and uses.
8. Federated, distributed systems of interoperable public water data generally provide scalability and flexibility to meet the diverse needs of data producers and users.

DEFINITIONS

Accessible: Full data sets are available to the public or authorized users for download in machine-readable, non-proprietary formats.

Authorized users: The group of users that are allowed to access a given dataset by its owner. This group may include the general public. In certain cases, such as datasets that include personally identifiable information or that represent

serious security risks, this group can be limited to users with specific data use agreements or security clearances. Data producers should be responsible for determining who is an authorized user, but the default for public water data should be the general public.

Data hubs: Structured sources of standardized water data aggregated by theme or geography.

Data producers: Entities that collect data for a specific purpose and have authority over what and how data are produced, including organizations that manage citizen science and crowd-sourced data (e.g., a wastewater treatment plant that produces data about surface water conditions, a state agency that holds water rights data, a non-governmental organization (NGO) that collects water data samples, a private company that takes meter readings).

Data standards: Guidelines regarding how data about a particular topic is (1) structured, defining what data elements should be present; (2) populated, defining the kind and quality of information represented; (3) encoded in machine-readable formats; and (4) made interoperable for data exchange.

Data users: Primary and secondary entities that use water data to create information and value. Primary users are the producers who use the data they collect to meet a specific mission (e.g., a state environmental quality agency that regulates discharges of pollutants, a reservoir operator that regulates the flow of water through a dam). Secondary users create value by combining multiple types of data, typically from multiple organizations (e.g., a conservation organization building stream restoration maps from data held by a utility, state, and reservoir operator; a private company assessing, modeling, and visualizing the environmental impacts of real estate development).

Findable: Data and metadata published on the web in compliance with data-on-the-web best practices, ideally tied to a common hydrography.

Interoperable: Data bulk download formats and application programming interfaces (APIs) that follow community standard patterns; metadata are included with data and of sufficient quality for users to make judgments as to what purposes the data is fit for use; and data content references including publicly available definitions, controlled vocabularies, and data standards appropriate to the data's subject matter.

Metadata: Metadata is information about data that assists potential data users in the discovery, access, and use of the data. It can describe the identity, subject matter, and producer of the data to aid in data discovery. It can describe the location, license, and point of contact for the data to assist in data administration and access. It can describe the structure, format, and any applicable data standards to assist in the use and manipulation of the data.

Modern data infrastructure: An integrated system of 21st-century information technologies, which includes common standards, formats, and tools designed to make water data easy to find, access, and share online. This system is connected by a network of people and organizations serving as water data producers, users, and hubs.

Water data produced for the public good: refers to water data collected for any public mission or purpose, including for regulatory compliance, either made available to the public or limited to authorized users.

Reusable: Data that is published and identified with version records and made available to the public or authorized users so that workflows can be reproduced.