Data Fragmentation
Challenges and the Call for Open Water Data

Cassidy White • March 2021
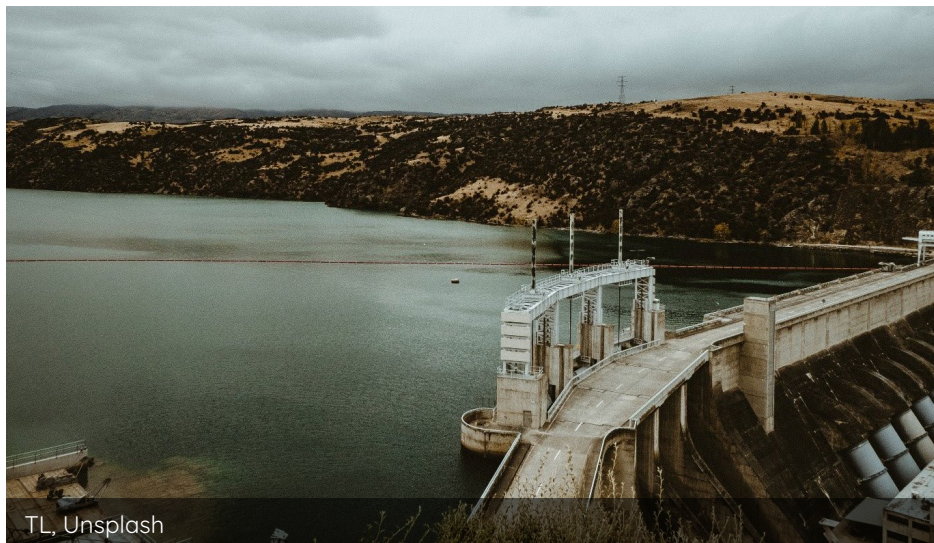

Jonathan-Kemper, Unplash

In the U.S., no one agency manages water resources. Rather, numerous agencies, each with unique missions and responsibilities, manage varying aspects of water. For example, the National Oceanic and Atmospheric Association manages weather and climate data; the US Geologic Survey monitors groundwater; the US Bureau of Reclamation manages federal dams, monitoring discharge and reservoir levels; the Army Corps of Engineers likewise manages dams for flood control and collects data on river discharge and precipitation; and the Environmental Protection Agency is charged with monitoring and regulating water quality.

State level governments are also home to myriad agencies, each unique to the state in which it was created. Many of these agencies collect water data to fulfill missions pertaining to their state's specific needs and demands. State Departments of Agriculture often gather data on water extraction, diversion, and returns as they pertain to irrigation and state environmental agencies collect their own data on water quality. Some states have up to 40 different agencies that collect and house water data! Each agency has its own techniques for recording, reporting, and storing data. As a result, water data is incredibly fragmented. This fragmentation across federal and state agencies and creates a major limitation to informed water management.

Seemingly everywhere and nowhere at once, water data may be widespread but is often difficult to find or is completely inaccessible due to the U.S. government's

Internet of Water

multi-tiered structure and disparate data collection across agencies. Likewise, the scattering of data across sources and host sites has exacerbated discrepancies in data standards that make manipulating or combining data sets unnecessarily complicated.

The governance structure of agencies in the U.S. is designed to divide responsibilities between central and local governments, and to give states the freedom to manage their resources and enforce regulations as they see fit. Nonetheless, this structure has inadvertently created a conundrum for water managers: although there is a plethora of data collected, the disparities in data collection and storage result in inadequate data for deriving meaningful insights and accurate analyses.

TL, Unsplash

Whereas one agency might measure discharge from a river, runoff, or hydroelectric dam in terms of cubic feet per second, another may measure in terms of acre-feet per year, leading to unit conversion challenges that make data difficult to work with.

Reliable water data is fundamental for enabling a resilient, sustainable, equitable water future. Data helps water managers detect patterns and trends to better define the uses, quantity, and quality of water and provides insights for regional planning and policy. While heavily utilized for assessing current practices and informing decision makers, data insights are only as reliable as the data upon which they are based.

If water leaders are misinformed, water use can exceed available supply or become misaligned with demands. Likewise, poor water quality data can lead to utility managers being under- (or over-) prepared to treat incoming water and can mislead the public about the safety of their drinking water or local recreational waterbodies (e.g., lakes and rivers). In addition, models and data evaluation techniques are limited by the availability and quality of input data. Although statistics can be used to deal with gaps in information, as water challenges become more and more significant, the errors associated with such techniques become a greater risk.

With water quality and scarcity threats continuing to build and gather public attention, the demand for publicly available water data is rising. As a result, the limitations of data fragmentation have been highlighted in recent years, and the

Internet of Water

ensuing call for centralized water data is gaining traction. Water data, however, needs to be more than centralized—it needs to be open and accessible. Efforts need to go beyond data collection to ensure data are usable and transparent, and not just for water managers. Water data should be organized in a way that reduces the expertise required to manipulate and interpret it, allowing for its use without an extensive need to filter, transform, clean, standardize, or search for it.

Open data hubs (or portals) provide a solution. An open data hub combines the concept of free, easily accessible and interoperable data with a centralized repository. More so, open data hubs have standardized datasets and metadata, user-friendly interfaces, are categorized and tagged by topic, and can be easily searched and downloaded. Whereas open data laws can mandate the development of open data hubs, in the absence of legislation, local decision-makers can still develop policy at the agency or departmental level to modernize their data management practices and create open data portals. The benefits of such practices provide an incentive for agency leaders to pursue data sharing initiatives even without the support of legislation. Although there are countless advantages to integrating data through open data portals, added value can be generalized as seven key benefits:

Whereas open data laws can mandate the development of open data hubs, in the absence of legislation, local decision-makers can still develop policy at the agency or departmental level to modernize their data management practices and create open data portals.

1. **Operational Efficiency:** Large organizations (e.g. government agencies, utilities, and businesses) often operate within silos. Opening and integrating water data across an organization's operations (e.g., financial, customer, and consumption data) often results in increases in efficiency and greater capacity to understand and address issues regarding water use.

2. **Comparative Performance:** Open data enables water users (e.g., utilities, industry, agricultural users, etc.) to understand how they are performing relative to other users in their sector. This provides opportunities for them to seek advice and improve performance, lowering costs and reducing water consumption.

3. **"One Water":** One location or region may have multiple entities managing

Internet of Water

different aspects of water (e.g., drinking water, wastewater, and stormwater managers; conservation and environmental organizations; regional planners and developers; industry and agriculture; etc.). Open, integrated water data can provide each water user/manager with a more holistic understanding of water resources within their community or jurisdiction, promoting collaboration and informed decision-making.

4. **True Cost of Water:** The full cost of water is often not reflected in market prices, leading to an undervaluation and overuse of water as a resource. Open water data can provide an improved understanding of the social, environmental, and community costs of water use. Understanding the true cost of water can lead to a more efficient allocation of water and more informed planning and decision making by water users.

5. **Planning and Demand Forecasting:** Open water data can reveal a broader and more detailed operational picture of statewide water supply and demand, providing a more accurate understanding of current and future water availability. More accurate forecasts empower political leaders, regional planners, and water managers alike to prepare for future scenarios and ensure water resources are adequately managed to facilitate sustainable growth and development.

6. **Water Efficiency:** A more robust understanding of water resources and water use can 1) identify stressed water resources, informing the need for water efficiency programs, 2) identify anomalies and pain points (e.g., nonrevenue water), enabling targeted efficiency efforts, and 3) provide insights as to the success (or failure) of water efficiency efforts.

7. **Streamlined Reporting and Accountability:** Another value proposition for open water data is the creation of a streamlined process for reporting water data from agricultural communities, industry, and utilities to regulatory agencies. Streamlining can create transparency, build trust, reduce the cost burden (especially for small utilities), and enable improved identification of violations and increased accountability (e.g., holding utilities and industry accountable to water quality requirements and agricultural users to their water use permits).

In recent years, the federal government has taken steps to initiate the development of open data hubs for agencies, with a 2013 executive order and the 2016 Federal Open Data Act requiring federal agencies make their data available and machine readable. Such standards and requirements don't yet exist at the state level, with some exceptions, and states thus have the potential to realize numerous benefits from the consolidation and standardization of water data collection and organization. Whereas many states do have preexisting open data portals, only

two (California and Oregon) contain sections for water data, leaving many others with a missed opportunity for improved water management.

Ultimately, states will pursue open water data as their individual circumstances demand, yet those with open data will be better equipped to tackle the rising challenges of water quality and water scarcity. With open data hubs, we can overcome fragmentation, and equip state leaders and water managers with the data they need to make informed decisions, ensuring sustainable water management for current and future generations.

REFERENCES:

Internet of Water. (2020). National Assessment of Water Data Legislation. Duke Nicholas School of the Environment, 1-32.

Josset, L., Allaire, M., Hayek, C., Rising, J., Thomas, C., & Lall, U. (2019). The U.S. water data gap —A survey of state-level water data platforms to inform the development of a national water portal. Earth's Future, 7, 433–449.

Christofer Jeschke, Unsplash