Water Data Governance

In 2016 Governor Edmund G. Brown, Jr signed AB 1755, the Open and Transparent Water Data Act (Act). A Partner Agency Team (PAT) formed to provide the interagency collaboration necessary to implement the Act. The PAT is made up of the state entities named in the legislation and others with important roles related to data. The PAT recently finalized a Strategic Plan to define a vision and path forward. The legislation, the strategic plan, and a report by Redstone Strategy Group allude to or directly recommend the creation of a water data governing body external to the state. As a next step, stakeholders are considering the formal creation of a consortium to serve this purpose.

Goal of Consortium

Provide a shared governance space to support broad engagement, investment, adoption, use and persistence of open water data infrastructure. The Consortium will include representatives from the Partner Agency Team and nongovernmental representatives with diverse perspectives on water data and expertise in data collection, publication, management, and/or use.

Open Water Data Infrastructure Activities

	Consortium	Partner Agency Team
Tasks	Provide feedback and recommendations, in an iterative process, to improve data collection and publication	Collect and publish open data
	Provide feedback and recommendations, in an iterative process, to improve federation of existing and new portals	Maintain and expand federation of existing and future data portals
	Venue for developing agreements on standards and protocols	Adopt and implement standards and protocols
	Provide a pathway to solicit and leverage external (to the state) resources in coordination with state resources	Submit budget change proposals using existing state budget process and expend state funds as allocated in the state budget
	Promote innovation and lay the groundwork for pursuit of new opportunities	Coordinate with and anticipate new opportunities developed by the Consortium
	Build trust with data producers, providers, users, consumers	Be good partners in developing trust between state entities and others
	Assure that the open water-data infrastructure is accountable to shared, measurable benchmarks and performance measures	Adopt, as appropriate, benchmarks and performance measures and provide data and information to assess progress
	Develop use cases	Prioritize data transformation to serve use cases
	Develop recommendations to improve the process of data submission to the state	Consider and act upon recommendations, as appropriate
	Coordinate strategic direction for the data platform	Adopt and implement strategic direction for the data platform

Funding

The Consortium's formation is contingent on provision of base funding for 2 years of operation. Separate conversations are on-going among the philanthropic community, and other interested stakeholders, to develop the necessary resources. State agency participation in the Consortium will be funded by state funds unless, or until, the Consortium determines otherwise. During the first two years, the Consortium governing body will hire staff, develop, and implement an appropriate organizational structure, such as a not-for-profit or public benefit corporation, and develop and implement a sustainable financing strategy. The Consortium will also undertake efforts to articulate the value proposition of open and transparent water data and recommend opportunities to streamline the submission of data to the state.

Governing California's Open Water Data Infrastructure

California's climate is changing, its population is growing, and there is a backlog of infrastructure maintenance and environmental needs all putting pressure on the state's water resources. Responding to these challenges will require coordinated action at every level of government, along with tribes, local landowners, and individual water users. Data can provide a firm foundation for collaboration across all levels of government, between agencies, and with stakeholders.

Water data is collected and managed by a wide range of sources including local governments, water agencies, State agencies, federal agencies, tribes, land owners, and others. A person or entity wishing to use data to inform their decision-making will need to search out data sets from multiple sources. The data may need to be transformed so that various data sets can be aligned and combined for analysis. Each data set curation workflow will be different and will depend on the issues presented by how the data is stored. The data may need to be cleaned to eliminate incorrect data points. This activity is time consuming, it can be expensive, and it is done redundantly by data users wasting time and resources that could otherwise be put directly to analyzing the data and using it to inform decisions.

The need to transform water data and make it more accessible is already driving change across water-related agencies. See, for example, the California Water Quality Monitoring Council workgroup web interfaces, including the California Estuaries Portal and the California Harmful Algal Blooms Portal; the California Department of Water Resources' California Data Exchange Center (CDEC), and Sustainable Groundwater Management Act (SGMA) Data Viewer; and California Department of Fish and Wildlife's Biogeographic Information and Observation System (BIOS) Viewer. Assembly Bill (AB) 1755, the Open and Transparent Water Data Act, creates a pathway to build on the excellent work already being done and to coordinate across efforts. AB 1755 will streamline data access and alignment, saving time, money, and resources. Most importantly, it will improve the capacity to use data for informed decision-making.

Full implementation of AB 1755 will result in an open water-data system with useful, accessible, organized, documented, curated, and federated open water-datasets from federal, State, and local government data sources able to integrate with data from private and academic sources. This means that a data user will be able to visit one web page, use one search engine to find data sets, and download data sets that are in open data format. Open data are in compatible formats, including consistent metadata that provides important information about the data set, are machine readable, regularly updated, and publicly accessible.

Making Water Data Open

The process of developing open data infrastructure supports a shift in practice requiring consistent investment and coordinated management to develop and apply standards and practices to transform existing data, update collection standards for new data, and federate data access points. The process requires a combination of the following activities and skill sets:

- An agreed-upon set of standards and protocols, and a process to evolve them, to meet new needs and demands on the system.
- Organizational development projects (e.g., training, tools, performance measures, and management systems) to improve State partner agency data literacy so that everyone involved, from those who drive the collection of data, stewards responsible for publishing open data, and

managers charged with using and communicating data and information, have the needed capacity to succeed.
- Staff time to transform existing data sets.
- Staff time and management systems to address data quality and data structure issues associated with data coming from ongoing data collection efforts (e.g., public water supply reporting, etc.).
- New software and system architecture and engineering, requiring staff with skill sets not always available in every entity, to make existing data access points interoperable and searchable.
- Resources for training to ensure staff are able to learn new skills.

Eventually, all water and ecosystem data systems managed by the State partner agencies will be able to collect data, store it, and transform it in the most effective and efficient manner to ensure it can be easily published in open data formats. The work of maintaining the open data publication infrastructure will become part of day-to-day data management.

The transition to and maintenance of open data infrastructure, by nature of the need to align practices across a broad array of entities, requires ongoing coordination and collaboration. The level of coordination and collaboration is greater than existing practices and will be a perpetual need to maintain the integrity of the infrastructure. The ongoing coordination will assure that federation is maintained; standards and protocols evolve to meet new demands, interests, and interoperability; shared priorities are coordinated across participating entities, and other activities that will assure the infrastructure is responsive to the needs of data providers and data users.

Open Data Governance

The highest benefits of open data infrastructure can only be achieved if there is broad support and investment in the infrastructure. As the report from the Redstone Strategy Group describes, open data infrastructure is something akin to transportation infrastructure. A high level of coordination across practices and investments, and clear standards and protocols, allow various levels of government to align to produce the federal highway systems that connect with the state highway system and in turn, with local roads. In a similar way, successful implementation of AB 1755 should give data users a smooth connection across data providers and datasets.

A shared governance space will provide the organizational infrastructure to support broad engagement, investment, and adoption and persistence of open data infrastructure.
Shared governance provides the following benefits and services:

- A collaborative environment to align and set data collection, and publication standards and priorities, responsive to the needs of water management decision-makers.
- Opportunity to leverage State, federal, local, tribal, and private investment.
- A flexible and nimble organizational space able to adapt to new opportunities and demands on the open water data infrastructure.
- A means to build trust with data producers, providers, users, and consumers.
- Assure that the open water-data infrastructure is accountable to shared, measurable benchmarks, and performance measures.

In parallel with all that a shared governance space can provide, there are certain functions that are best maintained by data collectors and data stewards. All participating data providers will retain full responsibility, control, and authority over their data. Data providers will be expected to autonomously

transform, manage, and maintain their data according to the shared protocols, prioritize data management tasks internal to their organization, assure that their data collection and publication activities support their organizational mission and responsibilities, invest their own resources as they choose, and establish appropriate privacy protections. Open water-data infrastructure should not interfere with these ongoing activities but will help to streamline and align these activities.

Proposed Governance Structure

To achieve the scale of alignment and collaboration necessary to successfully implement AB 1755, the Redstone proposal offers two related governing bodies. The first is an internal State agency governing body, the second is an external consortium made up of State and non-state representatives. The Partner Agency Team recommends the following governance structure informed by the work of the Redstone Strategy Group.

State Governing Body

The purpose of the State governing body is to coordinate and align State activities around AB 1755 implementation. Coordination and alignment will be necessary for the following activities:

- Develop and staff an integrated AB 1755 budget change proposal.
- Coordinate prioritization of data transformation to support data user priorities and other State priorities.
- Maintain and expand federation of existing and future data portals.
- Adopt standards and protocols that are consistent internally, and consistent with the consortium.
- Assure prioritization of AB 1755 implementation across agencies.

The State governing body should include individuals with authority to allocate funding and staff resources, such as directors and chairs of respective organizations, to assure that commitments across agencies, and with the consortium, are implemented. At a minimum, the California Department of Water Resources, State Water Resources Control Board, California Department of Fish and Wildlife, and the Water Quality Monitoring Council should have representation on the State governing body. It may also be beneficial to have participation from California Natural Resources Agency, California Environmental Protection Agency, Governor's Office of Planning and Research, Government Operations Agency, Department of Technology, and Delta Stewardship Council. The State governing body should have capacity to expand membership as needed to achieve the outcomes. It is likely that two levels of organization will be necessary with connectivity between the two. One level, which may be a water principals group or something similar, should have sufficient authority to implement policy. The second level should include the staff who lead the implementation.

Water Data Consortium

The consortium has four components:

- Steering committee.
- Use case working groups.
- Technical working groups.
- Professional staff.

The steering committee will be made up of State and non-state seats that could include federal and tribal agencies, non-government organizations, private industry, technical experts, academics, etc. The steering committee will govern the consortium. Working groups will focus on development of user-specified use cases and the technical aspects of standards and protocols A small water-data science team will support State agencies and external data providers, particularly during the startup period. The consortium work would be organized and supported by a professional staff.

The steering committee would be responsible for:

- Coordinating the strategic direction for the data platform.
- Coordinating and managing resources for data platform activities.
- Setting standards, conventions, and protocols.
- Overseeing the administrative staff.

The use cases working group(s) should leverage other existing groups, such as the California Water Quality Monitoring Council's theme-specific workgroups, along with interested individuals. The use case working groups would be responsible for:

- Consulting and engaging around user priorities.
- Consulting on technical and functional requirements for use-case implementation with technical working group.
- Supporting awareness and adoption of standards, protocols, and conventions by user community.

The technical working group(s) would be responsible for:

- Identifying, developing, and recommending functional and technical requirements for the data platform, building on related efforts.
- Aligning with platforms and technical efforts external to the State.
- Proactively engaging with use-case working groups to support implementation.
- Ensuring interoperability and usability across the data platform.
- Supporting awareness and adoption of standards, protocols, and conventions by the data provider community.

The professional staff would be responsible for:

- Arranging consortium meetings.
- Supporting activities of the steering committee, use-case working groups, and technical working groups.
- Communicating steering committee reports and findings.
- Outreach.