

**Protocols for
Assembly Bill 1755, the
Open and Transparent Water Data Act**

Contents

Preliminary Protocols for Assembly Bill 1755, the Open and Transparent Water Data Act.....	1
Introduction.....	1
Foundational Concepts.....	1
Protocols	2
Data Standards	2
Interoperability.....	2
Initial Minimum Protocols	3
Identify a Data Steward.....	3
Publish and Document on an Open Data Platform	4
Metadata Requirements.....	4
Data Dictionary Requirements.....	4
Guideline for Optional Descriptive Text	5
Access Data.....	6
Continuing Protocol Development - Use Cases and the Open Water Information Architecture.....	6
Continuing Development of Protocols – Long Term.....	8
Interoperability Testbeds	9
References.....	9
Glossary	10
Acknowledgements.....	10
Appendix A: Open Data Metadata and Data Dictionary Requirement.....	A-1
Appendix B: Checklist for Publishing a Dataset in Compliance with Existing AB 1755 Protocols (April 2018).....	B-1

Tables

Table 1 Three Initial Protocols Developed to Support Early Implementation of AB 1755.....	3
Table 2 A Sample Technical Workflow for Publishing a Dataset on data.ca.gov or the California Natural Resources Agency Open Data Platform.....	4
Table 3 A Sample Workflow for Accessing Data on an Open Data Portal.....	6
Table 4 Summary of the Functional and Technical Requirement Categories Specified in the OWIA to Protocol Categories Identified by AB 1755	7
Table 5 Mapping of 41 Functional and Technical Requirement identified in OWIA to Protocol Categories Identified by AB 1755	7

Figure

Figure 1 Proposed Process for Protocol Development.....	8
--	----------

Acronyms and Abbreviations

AB 1755	Assembly Bill 1755, The Open and Transparent Water Data Act
CNRA	California Natural Resources Agency
DWR	California Department of Water Resources
CCST	California Council on Science and Technology
CVS	comma separated values
HTML	hypertext markup language
IETF	Internet Engineering Task Force
JSON	JavaScript Object Notation
OWIA	Open Water Information Architecture
PDF	portable document format
Progress Report	Progress Report for Implementing the Open and Transparent Water Data Act with Initial Draft Strategic Plan and Preliminary Protocols
USGS	U.S. Geological Survey
XML	extensible markup language

Protocols for Assembly Bill 1755, the Open and Transparent Water Data Act

Introduction

In the wake of the most recent drought, the Legislature passed Assembly Bill 1755, the Open and Transparent Water Data Act (AB 1755) with the goal of improving water resources management through development of an open source platform that integrates existing water and ecological data. AB 1755 highlights the value of accessible, discoverable, and usable water data for both water managers and users; and a desire for increased transparency and collaboration among State agencies. AB 1755's specific requirements for protocol development are as follows:

The (California Department of Water Resources [DWR]), in consultation with the California Water Quality Monitoring Council, the state board, and the Department of Fish and Wildlife, **shall develop protocols for data sharing, documentation, quality control, public access, and promotion of open source platforms and decision support tools related to water data** (emphasis added). The department shall develop and submit to the Legislature, in compliance with Section 9795 of the Government Code and by January 1, 2018, a report describing these protocols. The report shall be developed in collaboration with the California Water Quality Monitoring Council, the state board, the Department of Fish and Wildlife, relevant federal agencies, and interested stakeholders, including, but not limited to, technology and open data experts and water data users. (California Water Code Section 12406 [a])

As described further in the *Progress Report for Implementing the Open and Transparent Water Data Act with Initial Draft Strategic Plan and Preliminary Protocols* (Progress Report), DWR has worked in consultation with its partner agencies, the State Water Resources Control Board, California Department of Fish and Wildlife, California Water Quality Monitoring Council, California Council on Science and Technology (CCST), UC Water, and others on development of protocols in response to AB 1755.

This report maps out AB 1755 protocol development in stages. The first stage identifies minimum initial protocols to support early implementation of an open data portal (“Initial Minimum Protocols” section). The second stage develops additional protocols identified by the Open Water Information Architecture (OWIA) and use cases (“Continuing Protocol Development – Use Cases and the Open Water Information Architecture” section). The third stage implements a mechanism for allowing protocols and standards to be created and modified (“Continuing Development of Protocols – Long Term” section). At each stage of the protocol development, interoperability testbeds will be utilized to determine whether identified protocols and standards are effective in specific applications (“Interoperability Testbeds” section).

Foundational Concepts

The process of developing protocols for AB 1755 began with defining a few key concepts. Work on protocol development could not occur without a common understanding of the terms “protocols” and

“interoperability.” Likewise, much debate occurred over the need to include data standards. Ultimately, the following foundational concepts prevailed.

Protocols

AB 1755 provides clear direction for the development of protocols. But, AB 1755 does not define “protocols,” which can connote different meaning depending on context. The following working definition of protocols was used in the development of this document and will be revisited in later phases of protocols development:

Protocols are methods of implementing a set of objectives and requirements in a systematic way. In computing, protocols mean both specific implementations of methods such as HTTP and FTP and, more generally as described by the Internet Engineering Task Force, protocols are sequences of processing steps that are also referred to as procedures.

Data Standards

Data standards will be integral to many of the protocols developed to meet AB 1755 requirements. The following working definition is adapted from the U.S. Geological Survey (USGS):

Data standards are sets of rules by which data or processes are described and recorded.

Standards are critical to sharing, exchanging, and understanding data in a meaningful way. The importance of data standards is well-stated by the U.S. Bureau of Land Management: "Standards provide data integrity, accuracy and consistency, clarify ambiguous meanings, minimize redundant data, and document business rules." To evolve a system that allows for meaningful exchange of data between groups it is necessary to agree on common data standards.

The USGS highlights issues that may arise from not implementing data standards: “If different groups are using different data standards, combining data from multiple sources is difficult, if not impossible. Utilizing data standards allows the agency to move from ‘project-based’ data files to ‘enterprise’ data files - and vice versa. In other words, the data become usable to more than just the project or person that created the data, because you know the data will be in an expected format and you know what is represented by the data.”

When standards are executed correctly they become a cost-efficient way to assure interoperability between those who produce data and those who use data, across organizational boundaries. It is anticipated that future phases of protocol development will necessarily address data standards.

Interoperability

Interoperability is the ability of diverse computer systems, or software, to exchange and make use of common input data. Interoperability is critical to supporting decision-making, as it allows different relevant datasets to be analyzed together. Increasing the ability for data sets and decision support tools to interoperate will, over time, yield better-informed decisions for water management. Many protocols and standards developed for AB 1755 will directly relate to, or support, increased interoperability. Over time, consensus on protocols and standards will lead to a higher level of interoperability.

Initial Minimum Protocols

As discussed in the Progress Report, the approach to implementation of AB 1755 involves starting with accessible products and adapting in response to user feedback, changing program needs, and policy decisions. To support the initial implementation of AB 1755, DWR has consulted with the partner agencies and others to outline three initial minimum protocols, consistent with available open data platforms, to guide early implementation of the program (Table 1). The intent is to develop only what is necessary to facilitate early implementation to avoid creating barriers to sharing of data through an open data portal. These protocols will necessarily adapt over time in response to both changing software capabilities and the needs of the users of the open data portals to support a more efficient and transparent use of data. The section “Continuing Development of Protocols – Long Term,” highlights a tentative process by which these protocols, and others developed in the interim, might be changed.

Table 1 Three Initial Protocols Developed to Support Early Implementation of AB 1755

Protocol	Business Requirement
1. Identify a data steward	All datasets published by Partner Agencies on the open platform have Partner Agency “owners,” whom are responsible for maintaining and curating them for users.
2. Publish and document on an open data platform	All datasets published by Partner Agencies on the open platform have a place where they can be discovered.
3. Access data	All datasets published by Partner Agencies on the open platform are machine readable, well documents and accessible to users.

A detailed description of each of the three minimum initial protocols follows.

Identify a Data Steward

To facilitate dissemination of information and avoid orphaned datasets, each dataset on the open data platform must have a data steward assigned to it from the appropriate agency. The data steward is responsible for the data and for meeting any related data requests. This protocol allows for multiple levels of data stewardship, such as a data creator or author (originator of the data), data caretaker (inheritor, or external sponsor of the data), data sub-steward (person responsible for a subset of the data), and other roles beyond what is defined here. This protocol does not define specific roles for data stewards, it simply indicates the need to have at least one accessible person identified, and prescribes minimum required information for each data steward:

- Name of steward.
- Contact information.
- Organization.
- Roles.
- Dataset(s).

Only data stewards can publish, update, maintain, or remove datasets published on the platform, and each dataset that is published must be assigned to an active data steward from the appropriate agency. The next protocol addresses publication in more detail.

Publish and Document on an Open Data Platform

For data to be discoverable, it must be published to, or made available to, an open data platform. To be published, all data must meet the minimum documentation standards outlined in this section, including the metadata standard, the data dictionary requirements, and the guidelines for optional descriptive text. Requiring minimum documentation helps ensure these items can be found by users of an open data portal, and once a user has found the dataset, that sufficient documentation on the dataset is available to answer most of the users' questions. A sample technical workflow is provided in Table 2 as guidance for a data steward trying to publish their dataset on one of the existing open data portals.

Table 2 A Sample Technical Workflow for Publishing a Dataset on data.ca.gov or the California Natural Resources Agency Open-Data Platform

Step	Activity	Actor	Required
1.	Log In	Data Steward	Always
2.	Ask system to create a new dataset	Data Steward	Always
3.	System creates new metadata template	System	Always
4.	Name dataset	Data Steward	Always
5.	Complete metadata for dataset <i>See Metadata Requirements, Machine Readability Requirements, and Guidelines for Optional Descriptive Text</i>	Data Steward	Always
6.	Identify category for dataset	Curator	Always
7.	Upload data as resource(s) for dataset	Data Steward	Always
8.	Identify keywords for resource(s)	Data Steward	Always
9.	Complete data dictionary <i>See Data Dictionary Requirements</i>	Data Steward	If applicable
10.	Identify keywords for dataset	Data Steward	Always
11.	Create API for dataset	System	If applicable
11.	Test API is functional	Data Steward	If applicable
12.	Notify curator that the dataset is published	Data Steward	Always
13.	Confirm dataset and resources appear and Test API	Curator	If applicable
14.	Check category	Curator	Always
15.	Check keywords	Curator	Always

Note: API = application programming interface

Metadata Requirements

To support initial implementation of AB 1755 on data.ca.gov and the California Natural Resources Agency (CNRA) Open Data Platform, the metadata requirements are to comply with the metadata standards identified by those portals. As additional necessary metadata elements are identified they will be added to the existing metadata requirements using a block structure format with the appropriate block elements, depending on the type of data. These requirements are included in the Appendix.

Data Dictionary Requirements

Similar to the metadata requirements, the data dictionary requirements are to follow those required by the respective open data portals, data.ca.gov and the CNRA Open Data Platform. These requirements are included in the Appendix.

Machine Readable Data Requirement

All tabular datasets published on an open platform must be machine readable. The Office of Management and Budget describes machine readable format in Circular A-11 Part 6 as: “a standard computer language (not English text) that can be read automatically by a web browser or computer system. (e.g.; xml).

Traditional word processing documents, hypertext markup language (HTML) and portable document format (PDF) files are easily read by humans but typically are difficult for machines to interpret. Other formats such as extensible markup language (XML), [JavaScript Object Notation] (JSON), or spreadsheets with header columns that can be exported as comma separated values (CSV) are machine readable formats. It is possible to make traditional word processing documents and other formats machine readable but the documents must include enhanced structural elements.” (Project Open Data)

Guideline for Optional Descriptive Text

To better meet user needs and reduce quantity of inquiries related to the data, data stewards are encouraged to provide optional descriptive text using the following guidelines:

For all datasets:

- Purpose
- Public license
- Distribution and Reuse Conditions
- Version
- Applicable temporal range
- Temporal accuracy
- When data was collected or produced
- Applicable spatial range
- Spatial accuracy
- Management procedures
- Data quality procedures
 - Records
 - Dataset
- Explanation of all controlled vocabulary used
- Explanation of all field domains used

For observations:

- Observation methods
- Instruments
 - Instrument calibration
 - Instrument accuracy

For derivative products and datasets:

- Lineage
 - Dataset
 - Reference
- Methods
 - Statistical formulae applied
 - Spatial aggregations
 - Temporal aggregations
 - Models
- Consistency
- Completeness

Access Data

Well-documented, published data with an appropriate data steward is not useful to the larger water community unless it is also accessible to the user of that data. To support this need, a sample workflow for user access to a dataset has been created in Table 3. While this workflow pertains to the user accessing the data set, it has significant implications on how state agencies should build platforms and organize data to support accessibility.

Table 3 A Sample Workflow for Accessing Data on an Open Data Portal

Step	Activity	Actor	Required
1.	Access open data portal via internet	Data Consumer	Always
2.	Search using keywords or tags	Data Consumer	Always
3.	Generate results list sorted by relevance to search terms	Open Data Platform	Always
4.	Select desired dataset from search results	Data Consumer	Always
5.	Take user directly to data or to data location	Open Data Platform	Always
6.	Query and visualize results using basic in-browser tools	Data Consumer	If applicable
7.	Download full or relevant queried portion of the dataset	Data Consumer	If applicable
8.	Connect to dataset directly via API	Data Consumer	If applicable

Note: API = application programming interface

Continuing Protocol Development–Use Cases and the Open Water Information Architecture

The next stage of protocol development will revolve around use cases and the OWIA. Use cases, as described in Data for Water Decision Making are “short examinations of how decision processes employ data – to inform a decision-driven water data system.” In February and May 2017, DWR, CCST, and UC Water co-hosted workshops to engage stakeholders in the development of use cases for AB 1755. As a result of the workshops, 20 draft use cases were developed (Cantor 2018).

The OWIA, in turn, applied the use cases to develop a list of functional and technical requirements. The OWIA document addresses the intended outcomes (functional requirements) and system details (technical requirements) to ensure that both executives and engineers remain aligned in common purpose. The OWIA outlines the protocols, procedures, resources, governance, and minimum standard of technology required to meet the needs of California’s water community, while also promoting greater levels of openness, transparency, and comparability for the information needed to manage water-related resources more effectively.

Each of the 41 functional and technical requirements identified in the OWIA is expected to result in a corresponding protocol. Protocols developed to meet the OWIA functional and technical requirements are mapped to AB 1755 protocol requirements in summarized form in Table 4. Table 5 includes the full list of functional and technical requirements as presented in the OWIA report. For more information about the functional and technical requirements, see the attached OWIA report.

Table 4 Summary of the Functional and Technical Requirement Categories Specified in the OWIA to Protocol Categories Identified by AB 1755

OWIA Functional/Technical Requirement Categories	Required Protocol Category Identified by AB 1755					
	Data Sharing	Documentation	Quality Control	Public Access	Open Source Platforms	Decision Support Tools
Data Acquisition	X					
Quality Control	X	X	X		X	X
Publication	X	X		X		
Data Traceability	X	X		X		
System Portability					X	X
External Interfaces	X				X	X

Notes: AB 1755 = Assembly Bill 1755, The Open and Transparent Water Data Act; OWIA = Open Water Information Architecture

Table 5 Mapping of 41 Functional and Technical Requirement Identified in OWIA to Protocol Categories Identified by AB 1755 (Helly 2017)

Name	Data Sharing	Documentation	Quality Control	Public Access	Open-source platforms and decision support tools
Data Acquisition	X				
*-Manual-	X				
*-Automated-	X				
Quality Control-*	X		X		
*-Verification-	X		X		
*-*Documentation	X	X	X		
*-*Reproducibility	X		X		
*-*Data Traceability	X		X		
*-Standardization-	X	X	X		X
*-*File-naming Conventions	X	X	X		X
*-Interoperable Transformation-	X		X		X
*-*Separation of Data and Computation	X		X		X
*-*Data Interoperability	X	X	X		X
*-*Products or Resources	X		X		X
Publication-*	X	X		X	
*-Cross-Referencing-Service-	X	X		X	
*-*Assignment of Digital Object Identifiers	X	X		X	
*-Packaging-	X			X	
*-*Compression Methods	X			X	
*-*Archive File Formatting	X			X	
*-Archival-	X			X	
*-*Open Access Distribution	X			X	
Data Traceability-*	X	X		X	
*-Metadata Production-	X	X		X	
*-Intellectual Property Rights Management-	X	X		X	
*-Public Law Compliance-	X	X		X	
*-Licensing-	X	X		X	
*-Liability-	X	X		X	
*-Searching-	X			X	
*-*Cross-referencing System Integration	X			X	
*-*Search Engine Optimization	X			X	
*-Version Control-	X	X			
*-*Binary Data	X	X			
*-*Non-Binary Data	X	X			
*-Anomaly Reporting-	X	X			
System Portability-*					X
*-Backup and Restore-					X
*-Platform Portability-					X
External Interfaces-*	X				X
*-Data and Metadata Acquisition-	X				X
*-Data and Metadata Distribution-	X				X

Notes: AB 1755 = Assembly Bill 1755, The Open and Transparent Water Data Act; OWIA = Open Water Information Architecture

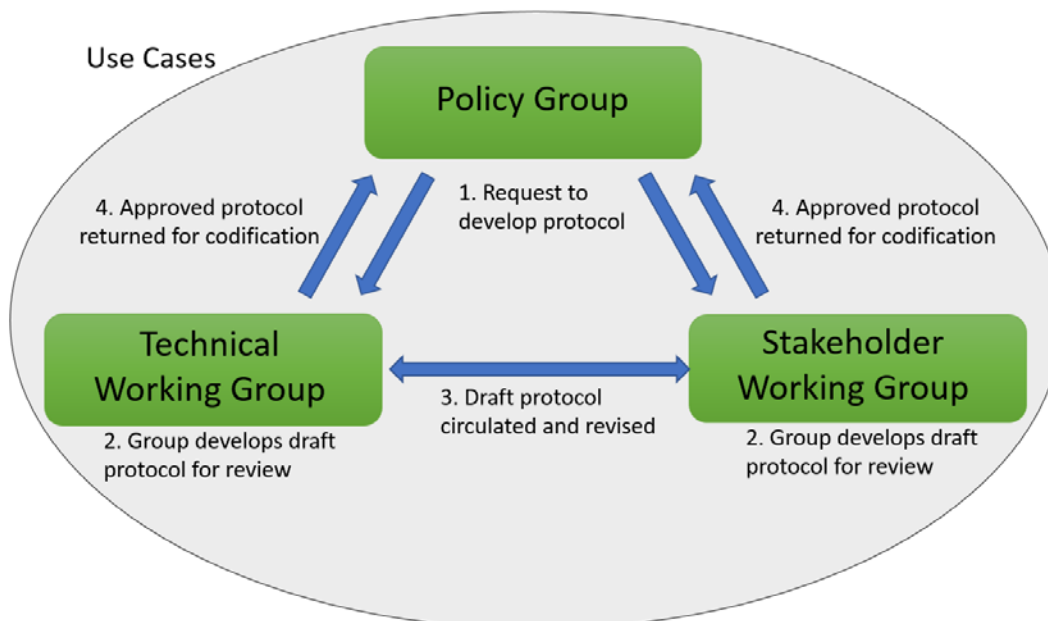
Continuing Development of Protocols – Long Term

The implementation of AB 1755 is based on the idea of iterative improvement as outlined in the *Initial Draft Strategic Plan*. To facilitate iterative improvement, a mechanism for adopting new protocols and standards is needed. To that end, the Partner Agencies are considering implementation of a modified form of the Internet Engineering Task Force (IETF) Internet Standards Process.

The IETF concept of working groups is particularly well-suited to AB 1755. Briefly, from the IETF website: “Working Groups (WGs) are the primary mechanism for development of IETF specifications and guidelines, many of which are intended to be standards or recommendations.” Leveraging the IETF process, the Partner Agencies are considering the creation of a governance structure that could support the creation of a fluid set of groups to develop and adopt protocols and standards to support interoperability and smooth functioning of the AB 1755 open data portal.

Conceptually, this might take the form of three interactive groups, a policy group consisting of an AB 1755 governance structure, a stakeholder working group consisting of people working with the data or utilizing the data for decision making, and a technical working group consisting of the people collecting and providing the data. These groups would interact through the lens of the use cases, which define “WHO needs WHAT data in WHAT form for WHAT decisions.” For example, the policy group could put forth a request to the working groups to develop a standard to support interoperability. If the technical working group and stakeholder working group agree that the requested standard is needed, the two working groups would decide which group will develop the standard. Because of the technical nature of this particular request, the technical working group would develop the interoperability standard and pass it to the stakeholder working group for review. If the stakeholder working group agreed the proposed standard was feasible, it would then be returned to the policy group. If approved by the policy group, they would then formally codify the standard. A simple diagram of what this process might look like is provided in Figure 1.

Figure 1 Proposed Process for Protocol Development



Interoperability Testbeds

During the development and implementation of protocols and standards, there is a need for testing to see if the requirements imposed achieve the business requirements identified. This section discusses the concept of testing protocols against use cases, an activity termed interoperability testbeds.

An interoperability testbed allows innovative users to test proposed protocols and standards to address a specific use case. Conducting a series of testbeds is critical to the development and vetting of protocols in different applications of any given use case. Interoperability testbeds would include, at a minimum, State and federal agencies responsible for providing data under AB 1755. There are additional groups, not mentioned in the bill, who could be helpful in the formation of a long-term well-functioning water data system. These groups include other federal and State agencies, the research sector, local governments, non-governmental organizations, groundwater sustainability agencies, and the private sector.

Once participants in interoperability testbeds are selected, they would work together as a group to select data formats, quality assurance/quality control levels, update frequencies, and exchange protocol choices. Conducting these testbeds will help to ensure that proposed protocols and standards utilized by the groups are beneficial in achieving the goals of the Strategic Plan, that data are accessible, sufficient, useful, and used. It is also necessary to assure that relevant data is properly inventoried and available for the selected use cases. Each testbed would conclude with a documentation of lessons learned and analysis of beneficial, or non-beneficial, protocols and standards as well as recommendations for future actions.

References

- Cantor A., Kiparsky M., Kennedy R., Hubbard S., Bales R., Cano Pecharroman L., Guivetchi K., McCready C., Darling G. 2018. Data for Water Decision Making: Informing the Implementation of California's Open and Transparent Water Data Act through Research and Engagement. Center for Law, Energy & the Environment, UC Berkeley School of Law, Berkeley, CA. 54 pp. Available at: <https://doi.org/10.15779/J28H01> or law.berkeley.edu/datafordecisions
- Dodd, B. (2016). Assembly Bill No. 1755. The Open and Transparent Water Data Act. Available at https://leginfo.ca.gov/faces/billNavClient.xhtml?bill_id=201520160AB1755 Cal. Water Code §12400
- Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing, RFC 7230 ISSN: 2070-1721, IETF, June 2014.
- Internet Engineering Task Force (IETF), The File Transfer Protocol, Rfc 354, IETF, 1972.
- Open Water Information Architecture AB 1755 Protocol Specification DRAFT Version 0 OWIA Technical Working Group 5, OWIA Technical Working Group, July 23, 2017
- Open Water Information Architecture Technical Working Group, OWIA System Requirements Document, Version 2.0, OWIA Technical Working Group, January 7, 2018

Project Open Data. 2017. Applying the Open Data Policy to Federal Awards. Viewed online at: <https://project-open-data.cio.gov/federal-awards-faq/>. Accessed: Jan. 2018

The Internet Standards Process – Revision 3, Tech. Report RFC 2026, IETF, October 1996.

U.S. Geologic Survey. 2018. USGS Data Management. Viewed online at: <https://www2.usgs.gov/datamanagement/plan/datastandards.php>. Accessed: Jan. 2018

Glossary

AB 1755: The Open and Transparent Water Data Act, legislation passed in 2016 that requires the creation, operation, and maintenance of a statewide integrated water data platform.

API: An application programming interface is a set of programming library calls and supporting compile and run-time libraries. These exist on both the client-side and server-side of a computer application although they are usually asymmetrical in terms of what the server implements versus what the client implements. The purpose for providing APIs is to standardize and simplify the programming required to add functionality to a software application and to enhance the portability and interoperability of software across both platforms and data.

Data system: A software or hardware system that collects, organizes, archives, distributes, or integrates data.

Data: Quantitative or qualitative representations or measurements of basic properties of the world.

Data-driven decision making: The practice of making decisions based on analysis of data rather than experience or intuition.

Decision support system: A modelling or analytic tool used to help guide decisions by processing and synthesizing data into information.

Federation: A federation is a group of data providers and users using jointly agreed-upon standards of operation in a collective fashion to ensure the interoperability of the resources they collectively hold and employ. The term may be used, for example, when describing the interoperation of distinct cyber infrastructure networks with different internal structures. The term may also be used when human groups agree to collectively manage cyberinfrastructure development and operation using commonly held, and managed, requirements, standards and conventions, and operating procedures to ensure the interoperability of distinct cyberinfrastructure resources.

Federated data system: A federated data system connects multiple independent data systems through common standards and conventions, while keeping those independent systems as autonomous entities.

Functional requirements: The translation of objectives into engineering terms and technical language describing how the objectives will be met.

Information system: A software or hardware system that supports the processing, analysis, or synthesis of data so they can be used to answer questions.

Information: Data that have been processed, analyzed, or synthesized so they can be used to answer questions.

Interoperability: The ability of diverse computer systems or software to exchange and make use of common input data.

Metadata: Data that describes and gives information about other data.

Objectives: The stakeholder-generated goals defined through use cases. The goals for the data system's intended uses and outputs.

Open Water Information Architecture (OWIA): An organizing structure for an open and transparent water data system created in response to the mandate of AB 1755.

Open: The provision of access to data using open-source and open-architecture protocols and methods.

Procedures: An established or official way of doing something.

Protocol: Protocols are methods of implementing a set of objectives and requirements in a systematic way. In computing, protocols mean both specific implementations of methods such as HTTP and FTP and, more generally as described by the Internet Engineering Task Force, protocols are sequences of processing steps that are also referred to as procedures.

Usability: Data that meets the needs of decision making processes in practice. Data that are readily available in formats that suit users' needs for making decisions.

Use case: For this report, defined as an example of a water decision making process and the data needs associated with that process. An answer to the set of questions of who needs what data in what form to make what decision.

Acknowledgements

This report was developed through a collaborative process led by the AB1755 Partner Agency Team reaching out to stakeholders who share interest in open and transparent water data for California. The effort began in July 2017 and continues today. Following is a list of agencies, organization and individuals who have provided invaluable input to this ongoing process.

AB 1755 Partner Agency Team

California Department of Fish and Wildlife

Steve Goldman
Tom Lupo
Karen Miner

California Department of Water Resources

Matt Correa
Gary Darling
Kamyar Guivetchi – AB1755 Program Sponsor
Julie Haas
Abdul Khan
Christina McCready – AB1755 Program Manager
Paul Shipman

California Natural Resources Agency

David Harris

California Water Quality Monitoring Council

Kristopher Jones
Nick Martorano

Delta Stewardship Council

George Isaac

Governor's Office of Planning and Research

Ken Alex
Debbie Franco

Government Operations

Stuart Drown
Angelica Quirarte

State Water Resources Control Board

Greg Gearheart
Rafael Maestu

AB1755 Collaborators

California Council on Science and Technology

Susan Hackwood
Brie Lindsey
Amber Mace
Shannon Muir

Center for Collaborative Policy

Ariel Ambruster
Alexandra Cole-Weiss

Delta Conservancy

Shakoora Azimi-Gaylon

Lawrence Berkeley National Laboratory

Debra Agarwal
Susan Hubbard
Peter Nico
Charu Varadharajan

Redstone Strategy Group

Nathan Huttner
Kathy King
John Whitney

San Diego Super Computer Center

John Helly
Michael Norman

S.D. Bechtel, Jr. Foundation

Joya Banerjee

University of California, Berkeley

Alida Cantor
Ronan Kennedy
Michael Kiparsky
Meredith Lee

University of California, Merced

Roger Bales
Leigh Bernacchi
Martha Conklin

University of California, Santa Barbara

Jeff Dozier

Water Foundation

Mike Myatt

OWIA Technical Working Group

Stephen Abrams, University of California
Deb Agarwal, Lawrence Berkeley Laboratory
Roger Bales, University of California, Merced
David Blodgett, U.S. Geological Survey
Martha Conklin, University of California, Merced
Matt Correa, California Department of Water Resources
Gary Darling, California Department of Water Resources
Greg Gearheart, California Water Boards
Kamyar Guivetchi, California Department of Water Resources
Tony Hale, San Francisco Estuary Institute
John Helly, University of California, San Diego
George Isaac, California Delta Stewardship Council
Sara Larsen, Western States Water Council
Christina McCready, California Department of Water Resources
Don Sullivan, NASA Ames
Dwane Young, U.S. Environmental Protection Agency

Additional Contributors

Commenters on January 2018 Progress Report Implementing the Open and Transparent Water Data Act with Initial Draft Strategic Plan and Preliminary Protocols

Patrick Atwater – California Data Collaborative
Derek Borba – Borba Farms
John Callaway – Delta Stewardship Council
Paul Cook – Irvine Ranch Water District
Richard Harasick – Los Angeles Department of Water and Power
Thomas Jabusch – Sacramento - San Joaquin Delta Conservancy
Jessica Law – Delta Stewardship Council
Elizabeth Lovsted – Eastern Municipal Water District
Jay Lund – UC Davis
Cathleen Pieroni – City of San Diego
Max Stevenson - Yolo County Flood Control and Water Conservation District
Devendra Upadhyay – Metropolitan Water District

Publications and Editorial Services by California Department of Water Resources

William O'Daly
Charlie Olivares
Scott Olling

Appendix A:

Open Data Metadata and Data Dictionary Requirement

Preliminary Open Data Metadata Requirement

Metadata Field	Requirement	Usage Description
Dataset Level		
Dataset Title	Required	Descriptive name for dataset
Dataset Description	Required	Summary explanation of dataset contents, purpose, origination, methods and usage guidance. Avoid jargon where possible.
Tags	Required	Enter descriptive keywords which describe the subject groups for the dataset and help it to be found in searches.
Organization	Required	Agency, department, board or commission publishing the dataset. Also known as Publisher.
Contact Name/Program	Required	Enter the name of the contact who maintains the dataset.
Contact Email	Required	The email for the dataset maintainer.
Public Access Level	Required	Whether this info could ever be made public. (Public, Restricted, Non-Public)
License	Required	List any restrictions on use of the data. Most often "Public Domain"
Program	Optional	Program or cross functional team name.
Spatial Coverage	Optional	Name of defined area or geometry of area data describes.
Temporal Coverage	Optional	Start and end time of events described in data.
Frequency	Optional	How often data needs to be updated.
Language	Optional	Most often English
Topic	Optional	Select a subject area from the defined list.
Homepage URL	Optional	URL for the page with useful information on the program creating the data.
Limitations	Optional	Appropriate usage notes, disclaimers and conditions of use.

Resource Distribution Fields		
File Title	Required	Descriptive name of the file.
File Description	Optional	Summary explanation of file contents, purpose, origination, methods and usage guidance. Avoid jargon where possible. Include for all in dataset or none.
Download URL	Optional	Optional if uploading
Format	Optional	File format such as CSV, PDF, XML, SHP or JSON
Data Standard	Optional	Established file structure defined for a particular use.

Preliminary Data Dictionary Requirement

Data Dictionary Element	Description
column	The name of the field from the data table
type	(text, numeric, timestamp)
label	Common English title for the data contained in this column. Please avoid abbreviations if possible.
description	Full description of the values included in the column. If the value is a date, document the time zone of recording, e.g. PDT (Pacific Daylight Time). If the column is a category, such as age group, then all categories or levels should be listed. If the values are calculated, the source of raw data and calculation method should be included.

Appendix B:

Checklist for Publishing a Dataset in Compliance with Existing AB 1755 Protocols (April 2018)

<input type="checkbox"/>	1. Identify a data steward with the following information:
<input type="checkbox"/>	a. Name of steward.
<input type="checkbox"/>	b. Contact information.
<input type="checkbox"/>	c. Organization.
<input type="checkbox"/>	d. Roles.
<input type="checkbox"/>	e. Dataset(s).
<input type="checkbox"/>	2. Publish and document on an open data platform:
<input type="checkbox"/>	a. Publish to one of the existing Federated Open Data Platforms (Open Gov or CNRA).
<input type="checkbox"/>	b. Complete the required metadata documentation.
<input type="checkbox"/>	c. Complete the Data Dictionary requirements.
<input type="checkbox"/>	d. Ensure the data is machine readable.
<input type="checkbox"/>	e. (Optional) Provide additional descriptive text.
<input type="checkbox"/>	3. Access data:
<input type="checkbox"/>	a. Ensure the dataset is discoverable via appropriate keywords or tags.
<input type="checkbox"/>	b. Test API and basic visualization tools to ensure the appropriate user experience.

