

San Diego Integrated Regional Water Management Data Management System Basic Design Recommendations

SJ Steinberg and M Hankin

Southern California Coastal Water Research Project Authority, Costa Mesa, CA, US

ABSTRACT

This project, funded by Proposition 84, is derived from the 2013 SD-IRWM's planned goal to develop a regional web-based Data Management System (DMS). The demand for such a system stems from the need for data to be consistent and sharable to meet data requirements and desires of various agencies and organizations in the region. A comprehensive DMS for the region can help improve understanding of the monitoring resource allocations throughout a watershed by providing a complete picture of where data are available throughout the region, for what parameters and at what frequency. This information can facilitate more effective use of existing data while simultaneously supporting better decision-making regarding where, when and what data are most relevant to filling information gaps.

One of the underlying factors driving the interest in a web-based DMS for the SD-IRWM is the lack of a single, consistent source for water monitoring data for the region. Before proposing the development of a new system, the strengths and weaknesses were considered for several existing systems used in the region. Features found desirable in some of these systems include: Consistency of data and quality control, system scalability, use of open source software tools, and a federated DMS architecture.

Two stakeholder workshops were convened to provide input to this recommendation report. The overarching priority for a regional DMS was identified as a system that serves to support watershed health and sustainability. Stakeholders identified three priority functions for the DMS: communication to a range of audiences; provide access to monitoring information; and streamlining the permitting processes. Three criteria identified for prioritizing these tasks were: The system should meet the needs of stakeholders and provide benefits shared by multiple stakeholders in the region; promote interoperability of systems; and build on innovative technology to optimize data gathering, analysis and sharing, and is user-friendly.

The Advisory Workgroup recommends the SD-IRWM develop the desired DMS as a federated data system, using an open source software platform. A federated data system will allow most participating data generators to continue using whatever DMS platform they already have in place and minimize alterations to current their existing systems and workflows while providing a platform for sharing data with the regional DMS. Data incorporated into this system should follow open data standards evaluated by and overseen by a governing body. The initial focus should be on identifying a limited, common set of data relevant to assessment and support of watershed health and sustainability with emphasis on data collected and utilized by the majority of participating organizations. Where data are already available via an existing DMS, opportunities to harvest data from such systems should be the preferred option. Because historical data generally will require significant effort to prepare for inclusion in a DMS, we recommend against including historical data as an early priority. To build a strong architectural foundation, it is imperative that appropriate accommodations for well-developed, standards-based metadata be included in the DMS design and implementation. Additionally, the DMS should adopt open data standards to ensure interoperability among federated systems.

Development should be parsed into manageable subcomponents based on the time and cost associated with each aspect of the data management system. Hosting of the system should incorporate elastic cloud computing to allow the system to adapt to workload changes as needed. The web-based data system must have an intuitive and easy-to-use interface which make data discovery and download as easy as possible. Data outputs vary widely as their intended uses, some of the core system outputs to be considered include: simple aggregated datasets that consolidate data from multiple organizations into a consistent, comparable format for download or summarized data which is provided in calculated or visual formats representing the status, trends or other characteristics of the watershed. The web interface and data transfer services should be developed using available open-source software.

Finally, a governance structure should be established to define priorities and data requirements for data documentation, QA/QC availability, data sharing, including data privacy and sharing agreements. The governance group would also oversee development of system specifications for development, hosting and maintenance of the DMS. Because any DMS developed will require initial funding to establish and ongoing funding to maintain and support the system, it is essential to plan for these needs prior to initiation of system development. An investment of time and effort early in the planning process is essential to the long term success of the system.

Full text: [878_SanDiegoWaterDataManagementRecommendations.pdf](#)