

CA Water Board Data Management Strategy and Open Data Initiative

Greg Gearheart, Deputy Director
Office of Information Management and Analysis

Western States Water Council - #westernwaternerdfest18
January 16-18, 2018

Water Governance in CA

- Federal agencies (usual suspects)
- State agencies (CA Water Boards and CA Dept. of Water Resources are primary)
- Local agencies (county and city governments in some cases)
- Special districts (lots)

Open Water Data in CA

- Pre September 2016 - voluntary, decentralized, topic driven efforts
- Post September 2016 - the Open and Transparent Water Data Act (AB 1755) has organized CA water agencies
 - State partners and a Data Management Strategic Plan (DWR leading)
 - Stakeholders curating use cases (UC Water)
 - Technical requirements (San Diego Supercomputing Center)

WB Data Management Strategy (2017 update)

Annual Civic Engagement Events

- Data Fair (open house)
- Data Innovation Challenge (hackathons)
- Water / Data Science Synthesis
- Brown Bag Series of Speakers
- Other partnerships

Our Data Management Strategy Framework

- Based on Principles
- Will guide:
 - Data driven management
 - “Water” decisions
 - “Technology” decisions
 - Quality program
- Lists our data management values
- Encourages “data literacy”

Databases and Datasets at the CA Water Boards

- Over 20 enterprise database applications
- Water quality, water rights, drinking water, etc.
- Program data (e.g., facilities, activities) and environmental / ambient data (e.g., surface water and groundwater quality, water use, water conservation, etc.)
- 18 data resources on data.ca.gov → more all the time

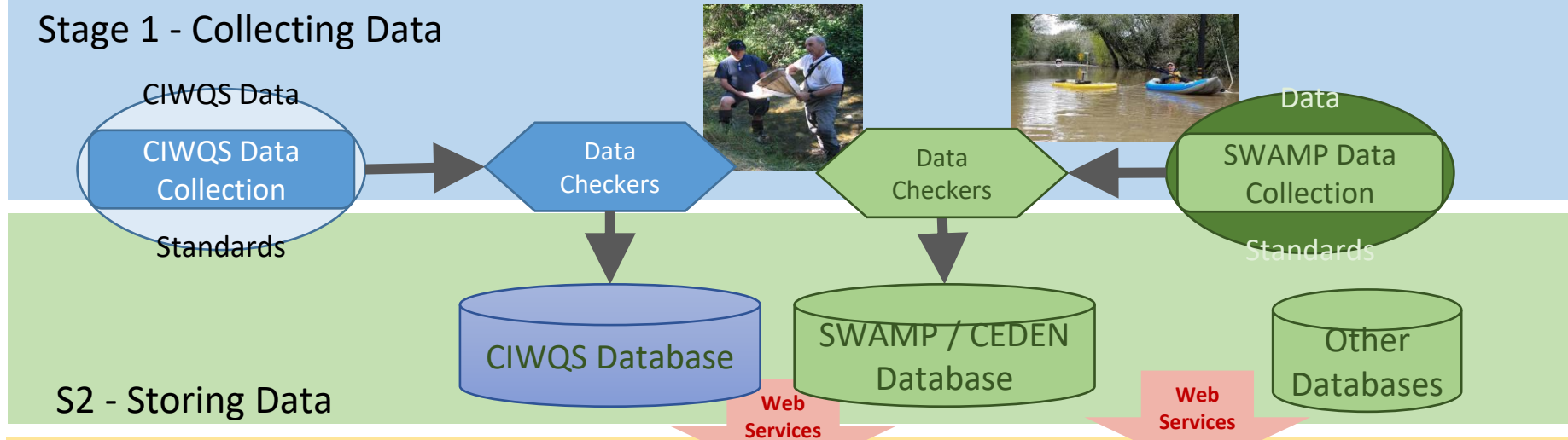
Why do we collect, use and produce data?

- To inform our **data-driven management** and planning activities – performance report cards, workplans, resource assignment/augmentation, evaluating program effectiveness, and many others examples;
- To inform our **critical decisions** regarding our mission(s) and water management responsibilities – water allocation and use, water quality planning and “policies,” permitting, program prioritization, and many other examples; and
- To provide **transparency** to our many partners and stakeholders for their use, interests and purposes.

Our Data Management Principles

- **Accessible (“Open first”)**: we deliver and maintain all critical data in machine readable datasets with metadata and data dictionaries
- **Quality and integrity**: our data is of known and acceptable quality and we deploy practices to protect its integrity with standards and protocols
- **Data literacy**: our organization understands its data needs and responsibilities and the concepts and terminology needed to achieve good data management practices
- **Govern with data**: we use our information technology to make data flow towards where it is needed or most valuable
- **Interoperability**: we invest in technological and semantic interoperability as the key to enabling and promoting access to and use of data.

Stage 1 - Collecting Data



S2 - Storing Data

S3 - Publishing Open Data

A	C	D	E	F	G	H	I
1	Latitude	Longitude	SampleDate	ProjectCode	LocationCode	CollectionTime	MatrixName
2	34.14699	-118.1633	8-Mar-16	BASMAA_TCT	Midchannel	8:57	samplewater
3	34.14699	-118.1633	8-Mar-16	BASMAA_TCT	Midchannel	10:08	samplewater
4	34.14699	-118.1633	8-Mar-16	BASMAA_TCT	Midchannel	11:53	samplewater
5	34.14699	-118.1633	8-Mar-16	BASMAA_TCT	Midchannel	12:34	samplewater
6	34.14699	-118.1633	8-Mar-16	BASMAA_TCT	Midchannel	12:34	samplewater
7	34.14699	-118.1633	8-Mar-16	BASMAA_TCT	Midchannel	11:18	samplewater
8	34.14699	-118.1633	8-Mar-16	BASMAA_TCT	Midchannel	13:14	samplewater
9	34.14699	-118.1633	8-Mar-16	BASMAA_TCT	Midchannel	13:46	samplewater
10	34.14699	-118.1633	7-Mar-16	BASMAA_TCT	Midchannel	12:47	samplewater
11	34.14699	-118.1633	7-Mar-16	BASMAA_TCT	Midchannel	13:31	samplewater
12	34.14699	-118.1633	7-Mar-16	BASMAA_TCT	Midchannel	13:31	samplewater
13	34.14699	-118.1633	7-Mar-16	BASMAA_TCT	Midchannel	14:07	samplewater
14	34.14699	-118.1633	7-Mar-16	BASMAA_TCT	Midchannel	14:40	samplewater
15	34.14699	-118.1633	7-Mar-16	BASMAA_TCT	Midchannel	15:24	samplewater
16	34.14699	-118.1633	7-Mar-16	BASMAA_TCT	Midchannel	16:03	samplewater

CalData data.ca.gov Portal

CIWQS

SWAMP
/ CEDEN

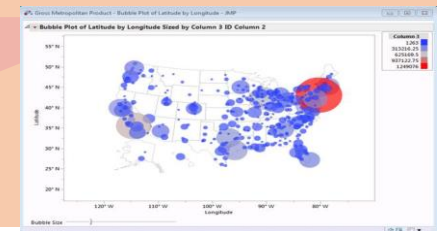
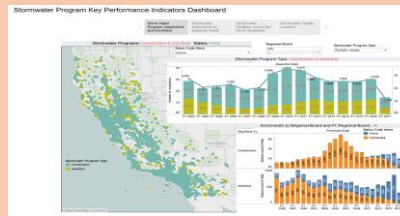
Other Open Data

Federal

CA DWR

APIs / Web
Services

S4 - Turning Data into Information



<https://public.tableau.com/profile/rafael.maestu#!/vizhome/StormwaterProgramKeyPerformanceIndicatorsUpdate1/Story1>

Storm Water Dashboard for Managers, Program Managers and Staff using Tableau



- » [CalEPA](#)
- » [State and Regional Water Boards' Map](#)
- » [Board Priorities](#)
- » [Laws/Regulations](#)
- » [Make a Payment](#)
- » [Plans/Policies](#)
- » [Programs](#)
- » [Decisions Pending and Opportunities for Public Participation](#)



[Home](#) » [About Us](#) » [Performance Report 1516](#)

The California Water Boards' Annual Performance Report - Fiscal Year 2015-16

This eighth annual Performance Report provides a mechanism to measure and evaluate both what we do and how the environment is responding to our actions, and is part of our overall effort toward developing as performance-based organizations. The Water Boards regulate more than 40,000 dischargers, and our core regulatory workload achievements for the fiscal year included review, update, or issuance of more than 700 individual permits and conducting more than 7,800 inspections.

The report presents numerous performance measures for specific outputs and outcomes that are currently tracked through Water Board data systems. These performance measures are organized under key functional categories of Water Board work and can be explored through the tabs below.

Percentage of Targets Achieved in FY 2015-16



» Previous Year Performance Reports: [FY 2008-09](#) | [FY 2009-10](#) | [FY 2010-11](#) | [FY 2011-12](#) | [FY 2012-13](#) | [FY 2013-14](#) | [FY 2014-15](#)

FEEDBACK

WHAT WE DO AND HOW WE ARE DOING FY 2015-16



Plan & Assess

[Regulate](#)

[Clean Up](#)

[Enforce](#)

[Fund](#)

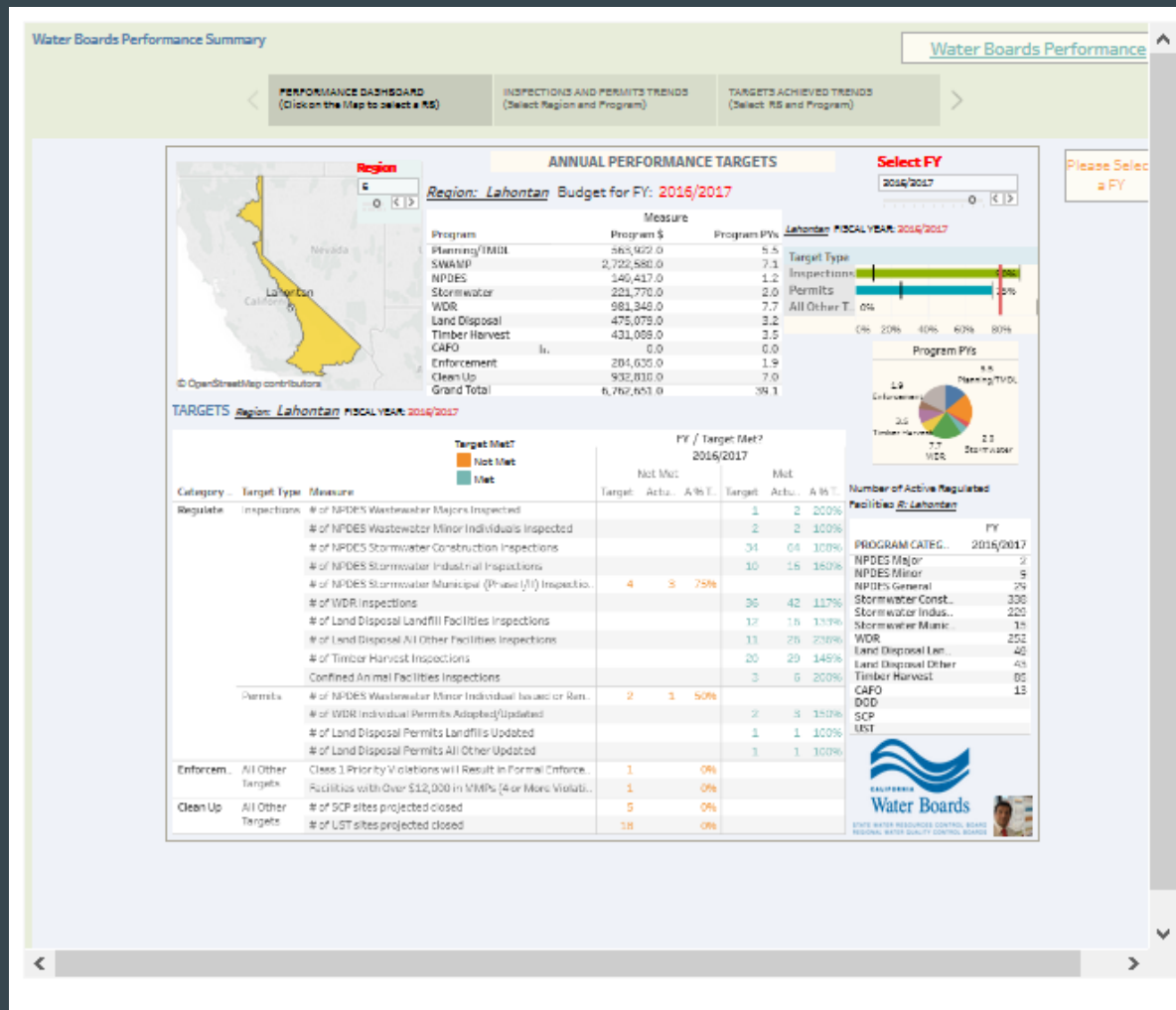
[Allocate](#)

[Targets](#)

The State and Regional Water Boards adopt plans and policies to carry out federal and State water quality

This dashboard showing inspection and permit trends for 9 Regional Boards for about 5 major programs could replace almost 50 individual report cards currently posted as individual HTML pages.

Additionally, the tool encourages more interaction and allows users to see more granular data. All due to open data techniques.



First Annual “Water Quality Status Report” - Series of Data Stories, Piloting of Tool for Outcomes

2017 Water Quality Status Report



Setting Flow Targets to Support Biological Integrity in Southern California Streams

Among the range of approaches available for setting flow targets that support biological integrity, a recently completed project in southern California utilized the Ecological Limits of Hydrologic Alteration (ELOHA) framework to assess the effect of flow alteration on the condition of benthic macroinvertebrate (BMI) communities across the region. The framework establishes recommended flow targets using a process that includes estimation of flow alteration and development of flow-ecology relationships based on the response of biological communities to changes in flow.

Biological Data Provide Ecological Context for Hydrologic Data

Data Indicators

Establishing flow targets based on ecology requires large data sets of both biological and hydrological conditions in order to derive relationships that are applicable to streams across a broad range of conditions. SWAMP's large bioassessment dataset (20 years) and tools made establishing flow targets possible. The dataset, collected by



R package to use web services to access data in CEDEN

The screenshot shows a web browser window with multiple tabs. The active tab is the GitHub repository page for `daltare/cedenTools`. The page header includes navigation links like Features, Business, Explore, Marketplace, and Pricing, along with a search bar and login/signup options. Below the header, the repository name `daltare/cedenTools` is displayed with statistics: 2 watches, 1 star, and 0 forks. The 'Code' tab is selected, showing a list of files and commits. A banner for 'Join GitHub today' is visible. The commit history table shows the following entries:

Commit Message	Author	Time
<code>daltare</code> fixed typo in README.md	a24398d	6 days ago
<code>R</code> updated the Roxygen2 documentation to roughly match the README.md file		8 days ago
<code>man</code> updated the Roxygen2 documentation to roughly match the README.md file		8 days ago
<code>.Rbuildignore</code> Initial commit		8 days ago
<code>.gitignore</code> Initial commit		8 days ago
<code>CEDEN Web Services - External Web ...</code> added the users guide		7 days ago
<code>DESCRIPTION</code> Initial commit		8 days ago
<code>NAMESPACE</code> Initial commit		8 days ago
<code>README.md</code> fixed typo in README.md		6 days ago
<code>cedenTools.Rproj</code> updated the Roxygen2 documentation to roughly match the README.md file		8 days ago

We have proof of concept results using computer vision (a form of artificial intelligence and machine learning) to recognize trash shapes in images, which can be captured via street sweepers, refuse trucks or other means.

OIMA's Trash Tracker: A.I. on a street sweeper

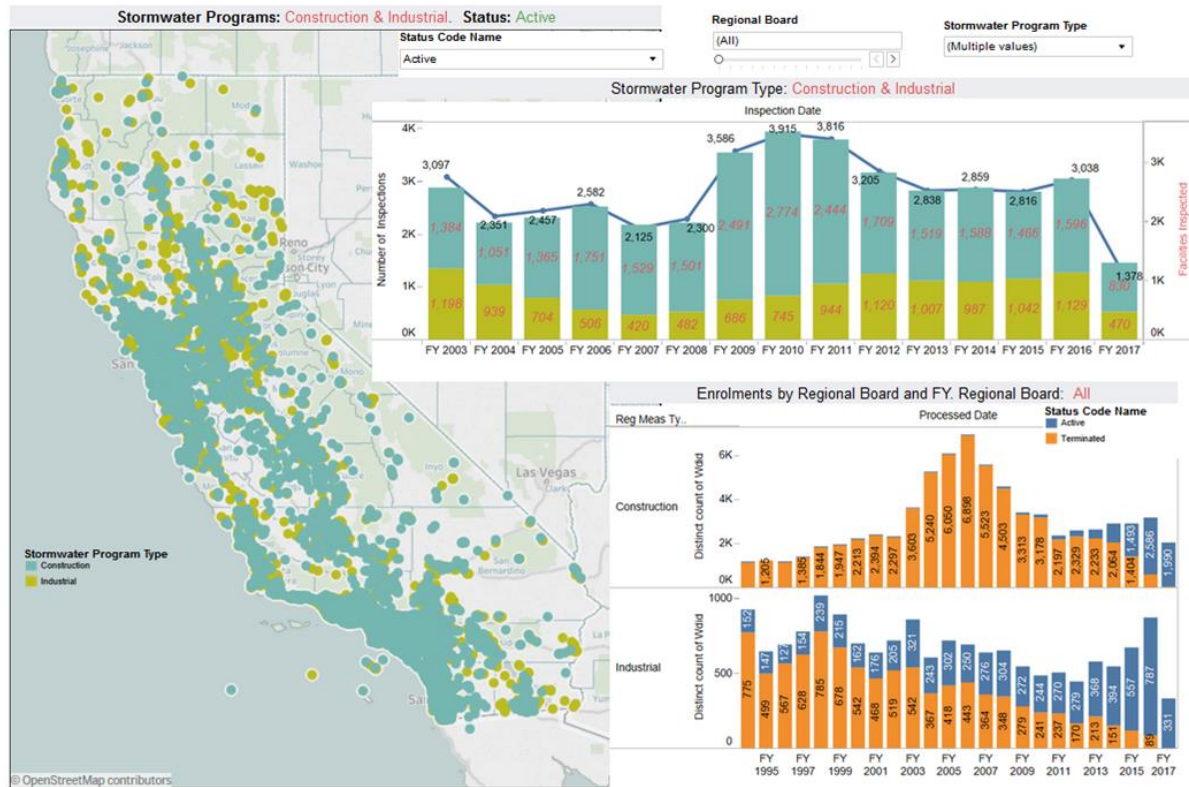


Storm Water Program: Inspections and

Stormwater Inspections by Inspector Name

Stormwater Facilities: Active but Never Inspected

Stormwater Facility Location



Stormwater Enforcement Tool

Filters:

Select Water Board Region:

9

Select Standard:

CTR

Select Monitoring Period:

2016 - 2017

Select Facility WDIDs (Optional):

All WDIDs

Select WQI Score Range:

0

100

Select a CES Parameter:

Pollution Burden

Filter by Score of Selected CES Parameter:

0

100

Filter for proximity to a 303d listed water body (ft):

Enter a distance in feet

☐ Show 303d proximity buffer

☐ Show excluded points

☐ Show parameters included in WQI score for each facility

WQI Scores:



WQI Scores - Tabular Data:

Column visibility Download Data

Search:

WDID	Monitoring.Period	Standard.Type	Exceedances	Sum.Excursion	Total.Samples	F1	NSE	F
9 30I005260	2016 - 2017	CTR	0	0	2	0	0	
9 30I013802	2016 - 2017	CTR	3	1.31957364341085	11	27.2727272727273	0.119961240310078	10.711195708
9 30I013854	2016 - 2017	CTR	3	3.55	3	100	1.18333333333333	54.198473282
9 33I000856	2016 - 2017	CTR	2	1120	2	100	560	99.821746880
9 33I023411	2016 - 2017	CTR	2	1.35833333333333	6	33.3333333333333	0.226388888888889	18.459796149
9 37I000277	2016 - 2017	CTR	49	236.98	112	43.75	2.11589285714286	67.906470284

https://daltare.shinyapps.io/Stormwater_Enforcement_Tool/

CA Water Data Profile #1 - Lean Water Conservation Data

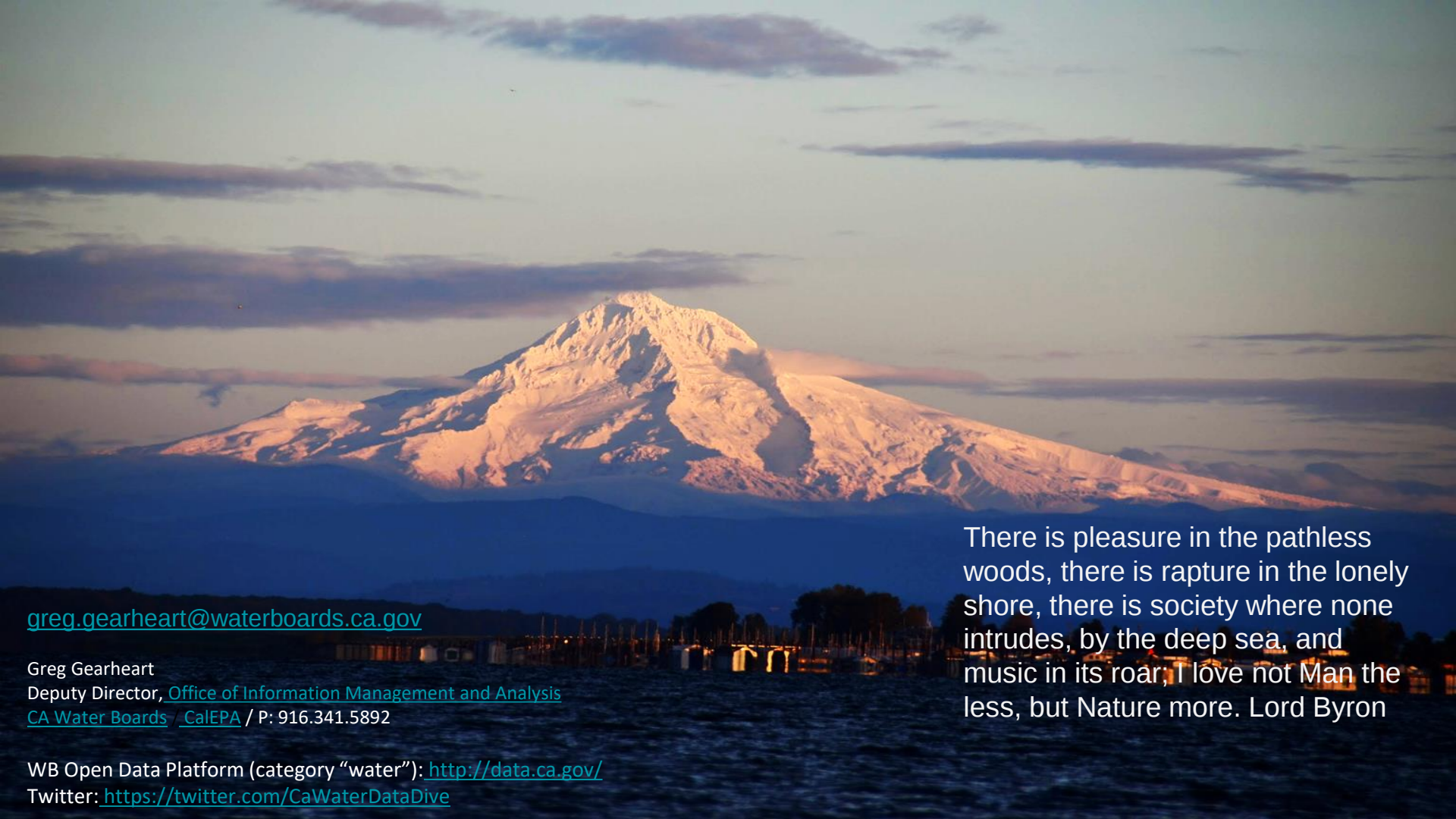
- Urban Water Supply Data - used to motivate Californians to conserve water (see Stanford <https://news.stanford.edu/2017/10/25/media-attention-drought-produced-water-savings/>)
- Extremely lean resources assigned to program
- Data driven messaging from top CA water leaders
- Data is now published (automatically) via data.ca.gov

CA Water Data Profile #2 - Not Lean Regulatory Data

- The National Pollutant Discharge Elimination System permit system regulates over 20,000 facilities in CA
- Two data systems manage the program - CIWQS (traditional) and SMARTS (stormwater)
- Self reporting data gets loaded regularly
- Over 200 staff are employed to run the program and use the data
- Millions of dollars spent to build and maintain the systems

CA Water Data Story #3 - The Ghost of IoW Future

- Legacy data systems with vital data & no open access
- Many more key datasets on desktop systems or unsustainable data platforms
- Multi year process to build an open water information architecture (OWIA) compliant platform (aka the Field of Dreams)
- Multiple agencies need to coordinate
- Ad hoc and operational decisions will suffer from lack of data
- External stressors uncertain (climate, human, etc.)



greg.gearheart@waterboards.ca.gov

Greg Gearheart
Deputy Director, [Office of Information Management and Analysis](#)
[CA Water Boards](#) / [CalEPA](#) / P: 916.341.5892

WB Open Data Platform (category “water”): <http://data.ca.gov/>
Twitter: <https://twitter.com/CaWaterDataDive>

There is pleasure in the pathless
woods, there is rapture in the lonely
shore, there is society where none
intrudes, by the deep sea, and
music in its roar; I love not Man the
less, but Nature more. Lord Byron