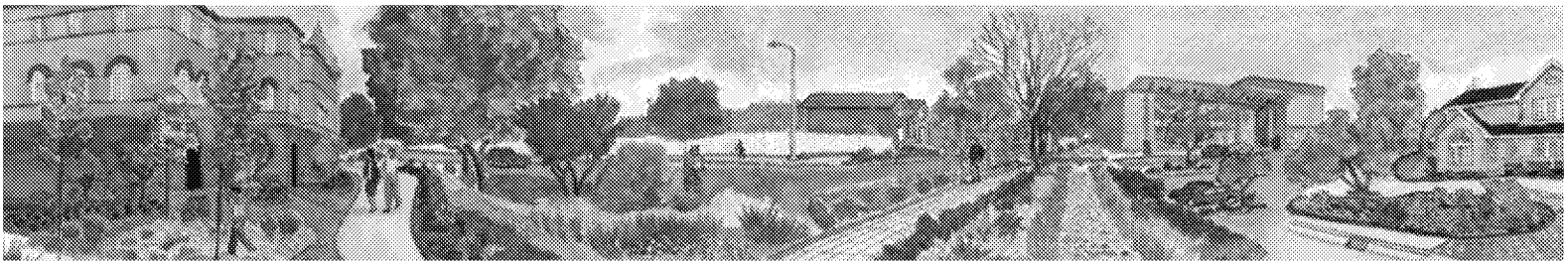

Strategy to Optimize Resource Management of Storm Water

Staff Report with Recommendations for Incorporating Open Data Concepts and Collaborative Activities for the Storm Water Program

Project:
Create Storm Water Program
Data and Information as Open Data

December 15, 2017

STORMS
STORM WATER STRATEGY
CALIFORNIA WATER BOARDS

DIVISION OF WATER QUALITY

STATE WATER RESOURCES CONTROL BOARD

ED_002551_00000338-00001

Introduction

One of the four goals identified in the State Water Resources Control Board's (State Water Board's) Strategy to Optimize Resource Management of Storm Water (STORMS) is to implement efficient and effective regulatory programs. The Project to Create Storm Water Program Data and Information "Open Data" (Storm Water Open Data Project) is identified in STORMS to help achieve this goal. Open data is defined generally as machine readable, fresh and well-defined datasets that can be used, modified, and shared by parties interested in acquiring information and gaining knowledge for any purpose.

The objective of the Storm Water Open Data Project is to increase the flow (amount, quality, and accessibility) of data and information useful for storm water management entered into and flowing out of the State Water Boards' existing data collection systems. The Storm Water Open Data Project Core Implementation Team[1] (Project Team) published datasets obtained from the Storm Water Multiple Application and Report Tracking System (SMARTS) on an open data platform. The Project Team also conducted public outreach to promote the use of the published datasets.

This Staff Report provides recommendations to the Water Board Division of Water Quality (DWQ) and Office of Information Management and Analysis (OIMA) to incorporate open data concepts and collaborative activities for the Storm Water Program at the State Water Resources Control Board and nine Regional Water Quality Quality Control Boards (collectively, Water Boards). These recommendations include: (1) creating a Storm Water Data Analysis Group; (2) refreshing the published datasets; (3) identifying additional potential datasets; (4) creating visualizations; and (5) conducting education and outreach.

Background

In 2015, the Government Operations Agency (GovOps) launched a Pilot Open Data Portal (Pilot Portal) to centralize datasets from various state agencies to allow citizens to easily access and use. The Pilot Portal was home to datasets from many state agencies in machine-readable format. Also, the Pilot Portal provided tools to help understand the data and tell visual stories across agencies. In 2015, GovOps conducted the Green Gov Challenge to seek innovative solutions that will help California improve its sustainable operations and engage with the public by leveraging the datasets available on the Pilot Portal and other public datasets available elsewhere.[2]

In 2016, the Water Boards participated in GovOps' open data effort spearheaded by the State Water Board Office of Information Management and Analysis (OIMA). The Water Boards published datasets to the Portal from three databases: California Environmental Data Exchange Network (CEDEN), SMARTS, and California Integrated Water Quality System (CIWQS). The Project Team published five datasets from SMARTS including information pertaining to the statewide Construction General Permit and Industrial

---

[1] The Storm Water Open Data Project Core Implementation Team consists of staff representing Division of Water Quality Storm Water Planning and Storm Water Permitting, Office of Information Management and Analysis, and the Santa Ana Regional Water Quality Control Board.

[2] "Green Gov Challenge." Green Gov Challenge. California Department of General Services, 2016. Web. 27 Mar. 2017.

ED_002551_00000338-00002

General Permit: Notices of Intent (permit coverage applications), inspections, permit violations, and enforcement actions. In addition to publishing datasets from SMARTS, the Project Team participated in a Water Boards themed data fair in March 2016 and the Data Innovation Challenge: The Water Data Dive in April 2016, both hosted by OIMA. These two events kicked off the Water Boards' open data initiative and focused on data from CEDEN, SMARTS, and CIWQS. Entries included applications, visualizations and other tools to understand the data better. Using SMARTS data uploaded by the Project Team, CloudCompli, Inc. together with Orange County Public Works (OCPW) won first place for their "Aggregating Municipal and State Open Data for Water Quality Investigations" submission. Their submission showed how the published data on violations, combined with OCPW's water quality monitoring data could be used to focus their inspections efforts, leading to cost efficiency and greater protection of water resources.

In September 2016, GovOps transitioned from the Pilot Portal into the new California Open Data Portal, transferring over the published datasets from SMARTS, CIWQS, and CEDEN. The California Open Data Portal (www.data.ca.gov) currently warehouses various datasets from many agencies.

In addition to the pilot projects, OIMA is in the process of developing a statewide Data Management Strategy to establish a vision, principles, and values for the State Water Board and Regional Water Boards. The Data Management Strategy will support data-driven management activities and decision making, and transparency to the public and stakeholder community.

To continue the STORMS open data effort, and align with the Data Management Strategy, the Project Team discussed various options moving forward. Successfully incorporating open data concepts and collaborative activities for the Storm Water Program depends on ensuring that the published datasets are useful, continuing to make data available on an open data platform, enabling stakeholders to actively interact with published datasets, and encouraging stakeholder engagement.
The Project Team developed the following five recommendations to State Water Board Executive Management.

Recommendation 1 - Create a Storm Water Data Analysis Group

The Project Team recommends creating the Storm Water Data Analysis Group, with representation from DWQ and OIMA. The Storm Water Data Analysis Group would be responsible for creating and implementing a work plan to address the recommendations in this Staff Report. Work plan tasks would include refreshing published datasets, identifying new datasets, creating visualizations, and conducting education and outreach. It may require approximately 1 to 2 hours to manually refresh a dataset, 1 to 2 hours to create a dashboard, and 2 to 4 hours to prepare and conduct each education and outreach effort. The amount of resource necessary to identify new datasets may vary depending on data format and accessibility.

Recommendation 2 - Refresh Published Datasets

Datasets published as open data are not automatically or continuously updated. The Project Team recommends a frequent refresh of the datasets published on the open data platform. This is necessary

ED_002551_00000338-00003

to ensure that users have access to current information allowing them to make informed decisions on storm water management issues. For example, when an entity like OCPW develops an application that integrates the published datasets to help identify potential water quality degradation, access to current information is imperative. The Project Team recommends two ways to refresh the published datasets – automatic and manual.

Refreshing the datasets automatically involves using an Extract, Transform, and Load (ETL) software. ETL refers to the process of extracting data from a source, transforming it to a desired format, and loading it to a target data system. The ETL software enables updating datasets more efficiently and frequently, which provides the opportunity to make near real-time data from SMARTS available on an open data platform. The Project Team recommends that State Water Board Executive Management consider investment in ETL or equivalent software. Until that time, the Project Team recommends manual refresh of the published datasets.

Recommendation 3 - Identify Additional Potential Datasets

The Project Team recommends that the Storm Water Data Analysis Group identify additional data appropriate to make available on an open data platform by identifying potential datasets for inclusion in SMARTS. An example of a potential dataset is the Phase I Municipal Separate Storm Sewer System (MS4) permits. While Phase II MS4 permit data is entered into SMARTS, at this time Phase I MS4 permit data are not. Each Phase I MS4 permit requires permittees to collect and report to the Water Board on their activities. Currently, there is no standard reporting or data collection mechanism for these MS4 permits to promote consistency. With no consistent datasets, consolidating and making data available on an open data platform can be challenging. The State Water Board should develop minimum standard reporting requirements to allow for consistent data collection, accessibility, and analysis on a site-specific, watershed, and statewide level. Language supporting this mechanism for reporting would need to be adopted by the Regional Water Boards as the Phase I MS4 permits are issued/reissued.

The regulation of storm water overlaps with other State and Regional Water Boards' regulatory planning and permitting programs. To allow full water quality evaluations for storm water discharges based on open data and information from other programs, the State Water Board should also consider similar or compatible standard reporting and data collection mechanisms for other overlapping water quality programs including, but not limited to:

- Total Maximum Daily Loads & Clean Water Act Section 303(d) Surface Water Assessments
- Clean Water Act Section 401 Certifications
- Waste Discharge Requirements (WDR) Permitting

Recommendation 4 – Create Visualizations

The Project Team recommends utilizing the open data platform and available software tools to create informational dashboards for internal and external use. Open data coupled with dynamic and

ED_002551_00000338-00004

interactive data visualization tools help program managers immediately assess performance and set targets and goals leading to an efficient and effective Storm Water Program.

Each published dataset from SMARTS on an open data platform contains thousands of records. Data visualization enables users to analyze thousands of records into manageable pieces of visual information such as maps, charts, and graphs. The conventional way of using a spreadsheet application such as Excel to analyze and visualize data can be laborious, especially if it involves multiple rows or columns of information and integrating multiple datasets. Tools such as Tableau, R Shiny, and QlikView make data visualization easier using drag and drop functionalities. Currently, the Water Boards have a set number of licenses for these tools that can be increased with support from State Water Board Executive Management. With these tools, a user can easily consolidate different pieces of visual information necessary for making decisions on a single screen by creating a dashboard. Adding filters and drop down menus enables users to dynamically interact with the data. The following is an example of a dashboard created using the publishable datasets from SMARTS.

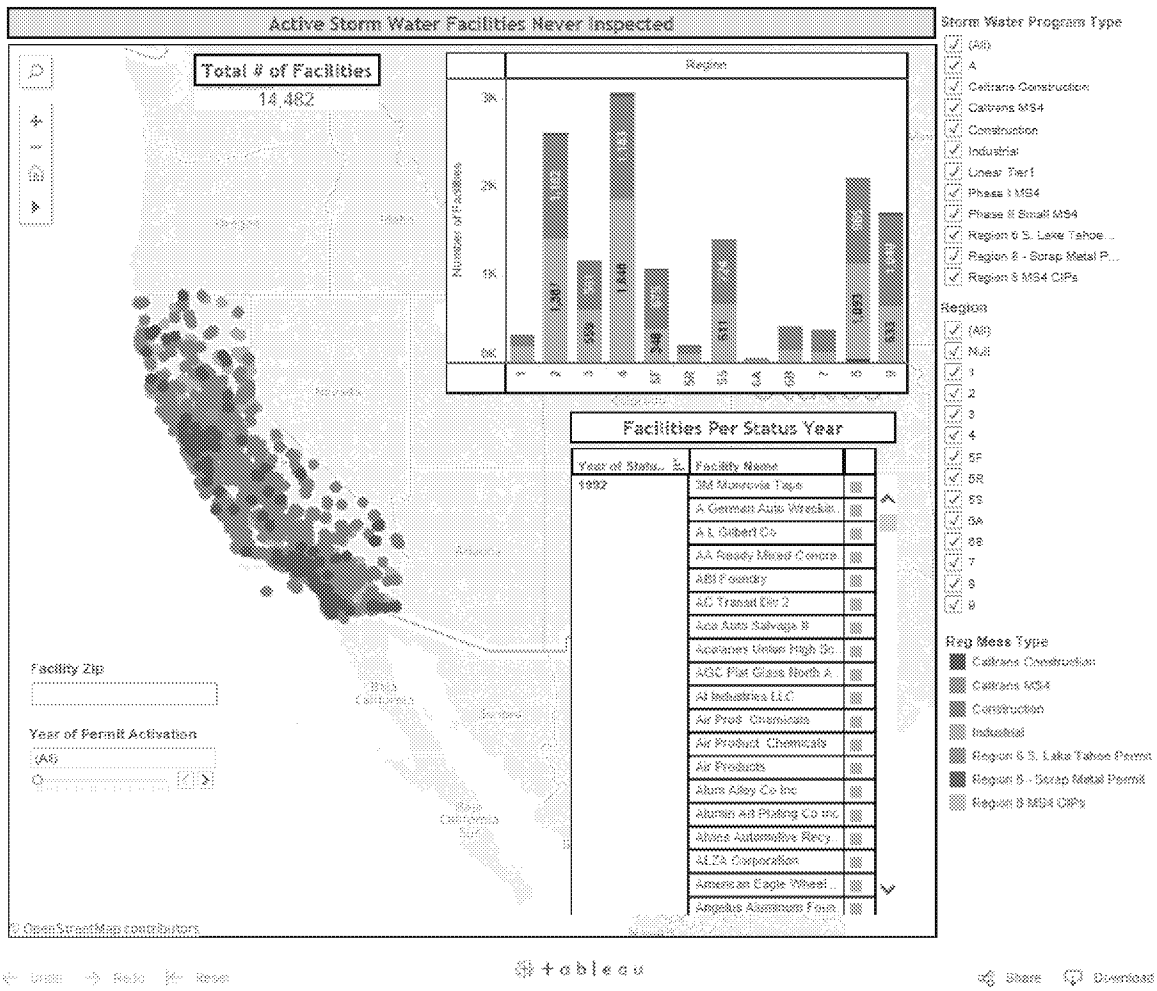

Figure 1: Active Storm Water Discharge Sites Never Inspected

Figure 1 aggregates two datasets including thousands of records on inspections and Notices of Intent (NOI) from SMARTS to provide data users a management tool for uninspected active storm water discharge sites between 1992 and 2017. The bar graph on this dashboard shows the number of uninspected discharge sites per region for each permit type and underneath it is a list of all the uninspected discharge sites sorted by year starting in 1992. The map shows the location of uninspected facilities in California. Using the interactive map as a filter along with the text box, slider, and other interactive components of this dashboard, a regulator can drill down to a specific zip code, year of permit activation, permit type, region, and site location.

A regulator can use this dashboard as a tool to quickly prioritize inspection efforts. For example, Figure 2 shows a map of uninspected sites, a list of uninspected sites sorted by year of permit activation, and the number of uninspected sites by permit type (i.e. Construction and Industrial) that are within a zip code. Duplicating this effort using MS Excel coupled with Geographical Information System (GIS) could take at least three to four times the amount of time it takes using Tableau.

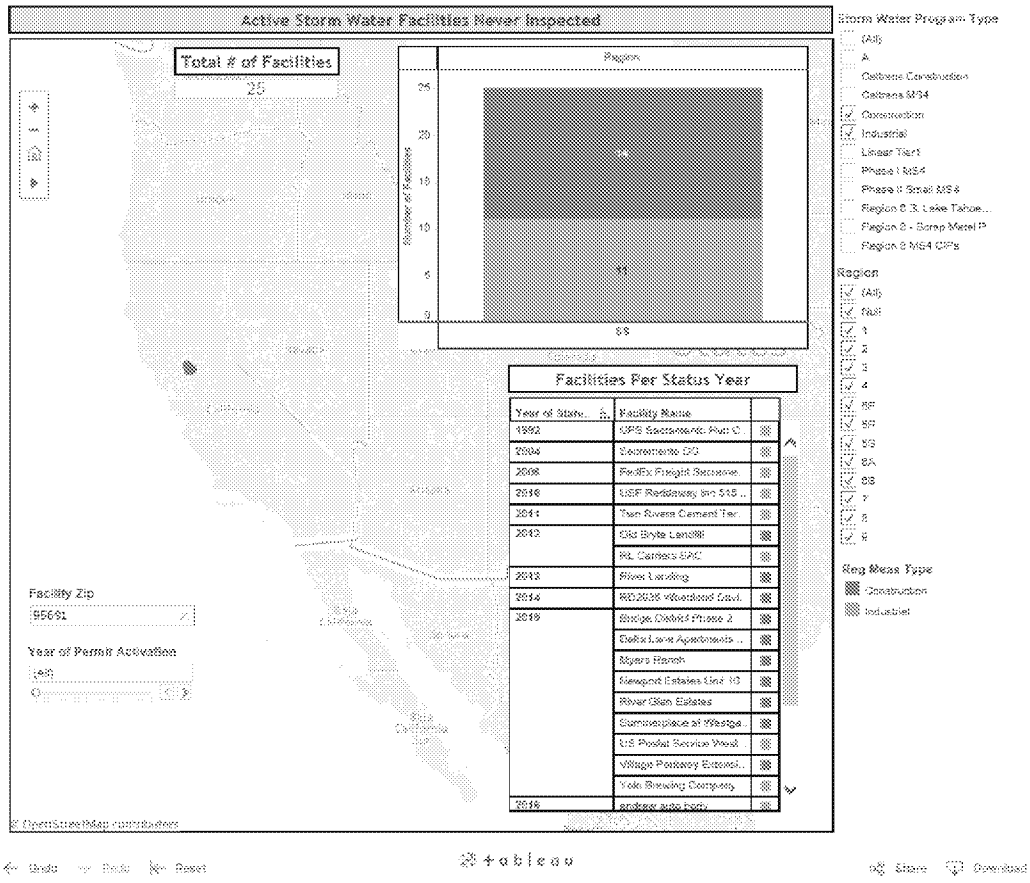

Figure 2. Active Storm Water Discharge Sites Never Inspected by Zip Code

Recommendation 5 - Conduct Education and Outreach

The Project Team recommends conducting education and outreach within the Water Boards and with stakeholders to promote data literacy and create a feedback loop (Figure 3) for improvements. Water Board program specific meetings called Round Tables and internal ad-hoc meetings can be useful for internal outreach and feedback. External outreach can be achieved by leveraging the STORMS Core Implementation Committee.[3]

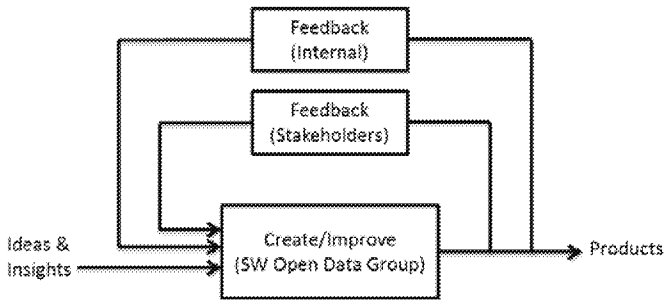

Figure 3. Feedback Loop

Data literacy is also key to fully benefit from open data. Awareness of the data available and how to use it will help users extract meaningful information, analyze the data, and interpret data visualizations to make informed decisions quickly. Internal and external feedback through communication is integral to the success of the open data effort by allowing data users the opportunity to identify their needs, and for the Storm Water Data Analysis Group to identify updates to address those needs. Feedback will provide insight on how to improve the open data effort.

Conclusion

The Project Team supports implementation of the five recommendations identified above through resources from Water Board DWQ and OIMA. The recommendations will improve the flow of data and information useful for storm water management supporting the STORMS' goal to implement an effective and efficient regulatory program. In addition, the recommendations align with OIMA's Data Management Strategy that includes a Data Flow Model (Figure 4) that supports management activities, decision making, and transparency to the public and stakeholder community.

[3] The STORMS Core Implementation Committee consists of participants representing one or more of the nine sectors identified in STORMS Section 5 - Implementation Resources. The Core Implementation Committee is a coordinating and advisory body that provides a forum for information exchange, consideration of STORMS project challenges & improvements, and development of collective feedback and recommendations to Water Board Staff and Management.

ED_002551_00000338-00007

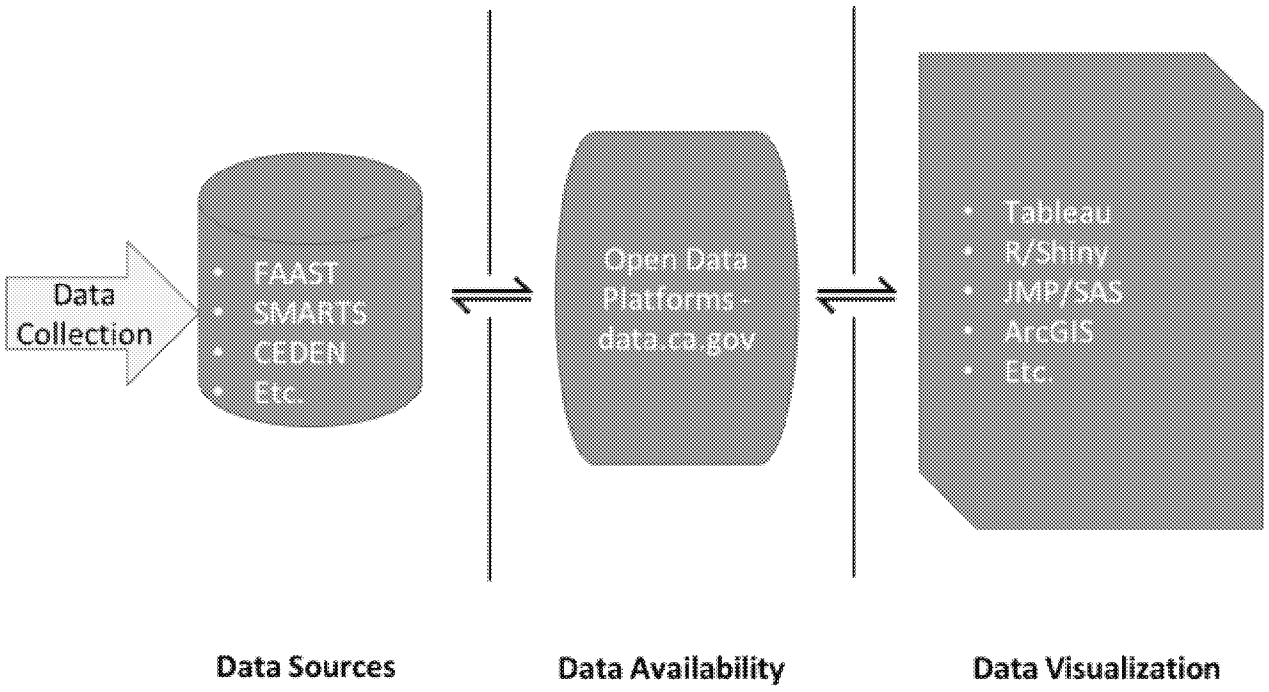

Figure 4. OIMA's Simplistic Water Board Data Flow Model