

INAUGURAL CALIFORNIA TRASH DATA DIVE NOVEMBER 16, 2018



2018 California Trash Data Dive

BACKGROUND

In recent years there has been an increasing interest in trash in California's streets, waterways and in the ocean, as land-based sources contribute large amounts of trash through runoff to the marine system. The California State Water Resources Control Board and many Regional Water Quality Control Boards have passed policies limiting the amounts of trash in the environment. These policies include Total Maximum Daily Loads (TMDLs), which limit the amounts of trash in specific waterbodies, and the "Trash Amendments", which call for municipalities to install either full-capture systems or employ the use of alternative trash capture methods and provide monitoring to ensure similar results as with full capture systems. Given these policies there are multiple needs that need to be addressed to adequately tell the story of trash in California and to determine whether or not policies have been effective.



Many studies of trash in various California environments have been done; however, historically most have been done as one-off studies and many have been only local in nature. A few larger statewide and regional studies, such as the California Coastal Cleanup Day and Southern California Bight Regional Epibenthic and Riverine Studies, have been done and can provide some insight into trash both spatially and temporally, but all three represent different habitats. The challenge becomes using the available datasets together to derive a meaningful story.

The goal of this workshop was to bring together the available data, and data scientists to work alongside trash experts, municipalities, and policy writers to discover how and where trash is created, transported, and deposited. The desired outcome was a series of data visualizations which tells a statewide trash story that can direct monitoring design, implementation, and reporting.

The specific objectives included:

1. Gaining better insight into the trash life cycle (e.g. Where is it coming from? How much is in our waterways? Are BMPs working? Where are the hotspots? How much is making its way into the ocean?)
2. Obtaining a better understanding of data standards, schema, and collection methodologies necessary to use monitoring data effectively
3. Collaborating across multiple disciplines and sectors to better understand the issue as a whole

THE EVENT

On November 16, 2018, the State Water Board's Office of Information Management and Analysis held the first ever 2018 Trash Data Dive (TDD). The TDD was convened at the San Francisco Estuary Institute, located in Richmond, California. It was well-attended by over 80 attendants and included participants from all over California to discuss the opportunities to address entrenched problems with respect to trash-related information. Representatives from Regional and

State Water Boards gathered together with those from other state agencies, non-profits, municipalities, private consultants, and the interested public (see Appendix I for a full list of participants) to develop new concepts and visualizations using available trash data as collated by the event organizers (see Appendix II for a list of available data sets).



Ambassadors

Ambassadors were selected prior to the TDD to represent knowledge about trash from a variety of different aspects. These included government agencies that serve both regulatory and regulated purposes, as well as research organizations and NGOs. Each ambassador was asked to give a 3-minute elevator pitch to describe what they do and what questions they were interested in having answered. In addition, ambassadors were polled before the TDD to develop a list of challenge questions (Appendix 3) for the participants to ponder going in to the event.

List of Ambassadors and sectors they represent:

- **State Government**
 - Greg Gearheart (State Water Resources Control Board)
 - Walter Yu (Caltrans)
 - April Roeseler (California Department of Public Health)
- **Local Government**
 - MS4
 - Gary Conley (2ND Nature)
 - Lisa Moretti (City of Sacramento)
 - Local Municipality
 - Shelly Walther (Sanitation Districts of Los Angeles County)
- **Research Organizations**
 - Tony Hale (San Francisco Estuary Institute)
 - Shelly Moore (Southern California Coastal Water Research Project)
- **NGOs**
 - Carolyn Box (5Gyres)
 - Jeff Kirschner (Litterati)
- **Federal**
 - Sherry Lippiatt (NOAA Marine Debris Program)

Speed Dating

Challenge questions were used to identify six areas of interest based on different sectors and included Marine Debris, Public, California Tobacco Control Program, Machine Learning, Municipal, and Regulation.

Posters were created to describe each sector and identify the pertinent questions.

Ambassadors were assigned a posters to stand next to and answer TDD participant questions in a “speed dating” round. This gave the TDD participants the opportunity to interact with each other and determine which questions were most relevant to themselves and develop groups to identify the data necessary and the desired visualizations. Once groups were formed they split out into different areas of the building to begin work on formulating ideas.



Hacking Session

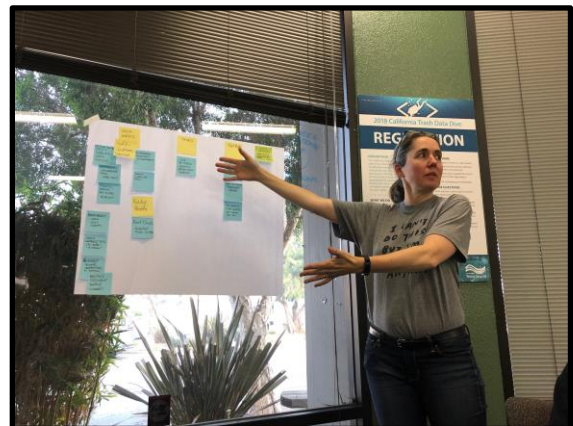
Following the presentations by the ambassadors and the speed dating round, participants broke out into groups and were given 4 hours to work together to develop their ideas and collaborate on visualizations and products from the trash data relative to their interests.

GROUP IDEAS

Participants detailed concepts spanning the gamut, from identifying plastics in multiple environments, to tobacco-pollution-focused problem-solving, to ideas for addressing trash among the homeless and overburdened communities, to a “power portal” designed to gather and distribute heterogeneous data types. The ideas generated and the opportunity for networking among like-minded individuals across many disciplines made this event valuable for the project team and participants. Below is a more detailed description of each groups project ideas and concepts.

Persistence of Plastics in the Marine Environment off of Southern California

While we have some idea of how much plastic is in our waters and how long it takes to decay, there has been little research into the residency time of plastic in each part of watersheds. Our question is how long does it take for a plastic cup on the beach to be converted to microplastics in the ocean and then to a chemical contaminant within the ocean sediments? While no model exists, much data has been collected in Southern California that could be used to potentially answer that question. The proposal here is to enlist a research graduate student and/or data scientist to put together known datasets relative to macro- and microplastics collected as part of beach, trawl, benthic, rocky reef and water column sampling efforts as well as using bathymetry, current and even polymer data. The researcher would line up these data sets and map and model for peer-reviewed published article.



Toolbox of Solutions to Enhance Data Collection by Citizen Monitors

There is a need to collect data while at the same time successfully engage citizen monitors in cleaning up the trash in the environment. The difficulty is in collecting consistent high-quality data at the same time as meeting the citizen monitoring expectations of participating not as data collectors but as helpers of the environment. The more complex the data collection card, you lose the a larger percentage of the crowd that will collect the data. Hence, the more simplistic the data card the higher the return rate on data. There is a need to create a simple toolbox of solutions to collect the most robust and applicable data to turn into meaningful policy recommendations. These toolbox solutions should be easy enough to use by NGOs, small

citizen groups, and even neighborhood beautification groups and should include a variety of data cards that can be used by different levels of users.

Quantifying Tobacco Waste in Litter

One major component of trash is cigarette butts and tobacco products which are a form of non-biodegradable litter. In addition, this litter contains tobacco which may lead to health impacts when incorporated into discharged runoff. This project recommended analyzing and characterizing trash data from the Bay Area Stormwater Management Agencies Association (BASMAA) street sweeping and curb inlet screen evaluation study and the CalEnviroScreen dataset provided by the CA Environmental Protection Agency (CalEPA) to determine how much tobacco waste there was and the potential for its effects on water quality.



Analysis by BASMAA showed in one area there was an estimated 404 gallons of waste, of which 10 gallons were of cigarette trash. This made for approximately 18,000 cigarettes total. The idea is to replicate and use this data in coordination with socioeconomic data to identify hot spots for tobacco waste and then associate those with health issues. The group also wants to investigate the effectiveness of cigarette receptacles. Initial results have shown that cigarette receptacles may actually lead to more cigarette litter in the environment. The group also proposed to use the data sets along with additional attributes and machine learning exercises to develop different algorithms and make predictions and identify areas of concern. Walter Yu from Caltrans has done some additional work to demonstrate this, which can be found on [github](https://github.com).

Addressing Agricultural Trash

Trash derived from agricultural practices has been a neglected topic. From creek clean-up surveys it is abundantly clear that farming practices (e.g. tilling over plastic row covers) can be a large contributor to plastic trash in local waterways. We sought to estimate the amount of agricultural trash by using cover of certain crops which are most likely to have tilled crop covers (e.g. strawberries, melons) and tracking their proximity to waterways. In the future, efforts will be made to identify agricultural trash from river surveys as well as identify factors that contribute to increased agricultural trash (storms, inability to dispose of inexpensively, wind, etc.) Additionally, efforts should be made to educate non-native communities on best practices.

The Effect of Homelessness on Track 2 MS4 Programs

Research shows the lower the socio-economic status of a community, the higher the trash production rate. Trash produced by the homeless creates policy decision points and financial hardships to those involved in prevention and cleanup. For example,

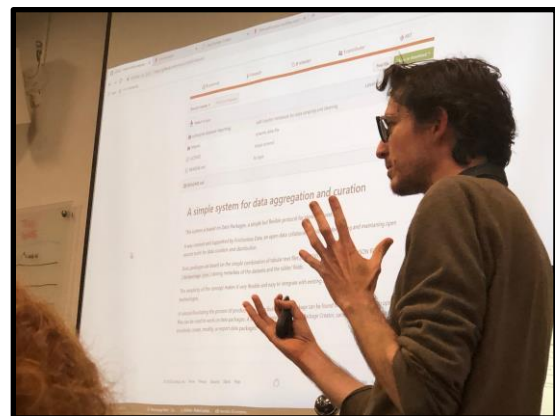
Sacramento Park and Recreation has a budget of \$5M for abatements and has conducted 4,080 abatements with 1,371 tons of trash during the first 10 months of 2018. Additionally, the city of San Jose has a \$1.5M homeless response budget, of which \$800K goes to specifically to a trash contractor. An additional \$1M dollars are spent on additional staff, such as park staff, police, etc. The question for this group is what is the distribution of funds needed to address the homelessness trash problems. Questions to be addressed: Who are the major offenders? What are the policy considerations? Is it cheaper to provide affordable housing? Where does responsibility start or end re: MS4 stormwater vs MS4 hardscape?

Municipal Options for Trash and Homelessness

Goal was to determine the solutions that are available and how they can be compared to choose the best solutions for any given region based on cost and amount of trash reduction. The group researched what programs are out there and what cities were using them. The goal became to build a dataset that allowed for comparisons and to analyze the data based on cost, number of homeless, available housing, time scale for the homeless to receive housing, amount of trash removed and to determine what is most effective. The primary question of the group was what are the most effective programs for eliminating trash from homeless encampments? The answer is not to eliminate homeless encampments but to look at demographics such as what ties a homeless encampment to a region, where are the most encampments, population density, household income, etc. The needs included an accessible platform for municipalities to access the data and data visualizations.

Power Portal

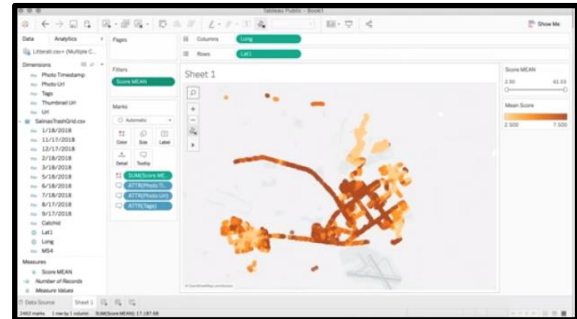
Focus is on plastic and big data through creating a Power Portal. There is lots of data from a variety of sources with different formats, different inputs and different needs. The Power Portal needs to have common formats, the input mechanisms must be simple, and it needs to be a one-stop shop where everything is easy to find and use. It also needs some form of governance, i.e. what is considered clean data, who should have access and what types of data validation should be involved. The group looked at the data available to the TDD participants and wanted to get an idea of how it looked spatially and counted and categorized it as litter vs waste management to create a heat map. Next step is to look at permitted vs non-permitted, non-point vs point (i.e. homeless) sources. Group in the end wanted to focus on outputs because most want different things, and has



different needs and objectives, all of which demonstrates the need for the outputs to be flexible, customizable and most of all insightful. Goal is to get the data to create action that creates positive change no matter your user level. Portal would include the ability to look at hotspots, trends, and predictive analytics giving us the ability to measure the impact of policies and practices that have been put into place. A simple system for data aggregation and curation was initiated and is available through [github](#).

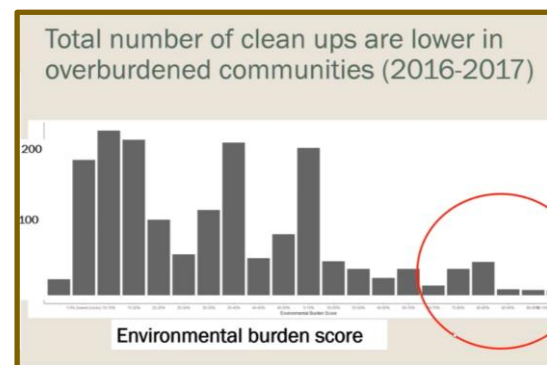
Defining Urban Trash Risk

We envisioned a holistic way to combine multiple data sources for determining the risk of urban trash impacts to local receiving waters. We wanted to make a tool that allowed trash data to be cross-walkable to each other so that although this data was collected in different ways, it could be combined to give an overview of trash in an area to help cities target priorities and track progress. Trash impact risk is defined as the magnitude of the potential impact and the likelihood of that impact occurring. Goal was to create a TrashMASH application to mash all the trash data together and reduce the risk of trash impact. Also, gives us the knowledge to update sampling design moving forward with new information to allow the reallocation of resources. Data sources include visual surveys, imagery, crowd-sourced trash/cleanup data, public works, and information collected by the application. Final product is a framework that allows for better decision making.



Do Overburdened Communities have a Disproportionate Litter Issues?

There are sensitive subpopulations that are disproportionately impacted by environmental pollution, and in California there are efforts underway to protect communities regardless, of race, color, national origin or income (environmental justice). Some of the anticipated stressors relative to litter are the high cost of waste disposal, availability of trash receptacles, lack of educational resources, and the prevalence of packaged food. Questions for this group included: 1. Are there fewer clean ups in overburdened communities? And 2. What are the litter-related costs associated to overburdened communities? Data sources to answer these questions included CalEnviroScreen, CalRecycle, Ocean Conservancy Coastal Clean-Up Day, Contra Costa County Trash Assessment Hotspot, and NRDC: Waste in our Water data. Data that identified overburdened communities in California was used to plot the number of cleanups versus and environmental burden score, showing that overburdened communities had fewer cleanups.



BREAKOUT SESSIONS

Nerd out with Drew

This session was led by Andrew Hill of the State Water Resources Control Board. The goal of his session was to build out a model that could handle all types of trash data. For example, if you were out in the field and encountered a piece of trash (a discrete trash entity), four types of information would need to be recorded: 1. Surveyor Information; 2. Attributes; 3. Geospatial/Timestamp; and 4. Status. Under each one of the four categories, additional information can be collected. Under Surveyor Information, for instance, you could collect the surveyors Organization, Contact Information, and Program Information. Under Status information such as was the trash item collected?, what condition was it in, etc. Information from other programs could also be included, such as the C.O.M.B. (category, object, material, and brand) classification system used by Litterati. The beauty of this object-oriented data structure is that it makes it easy to incorporate existing data sets, using programming languages such as Python and R, into a single, analyzable data set. The following image is the data model developed by participants in the 2018 Trash Data Dive on the day of the event.

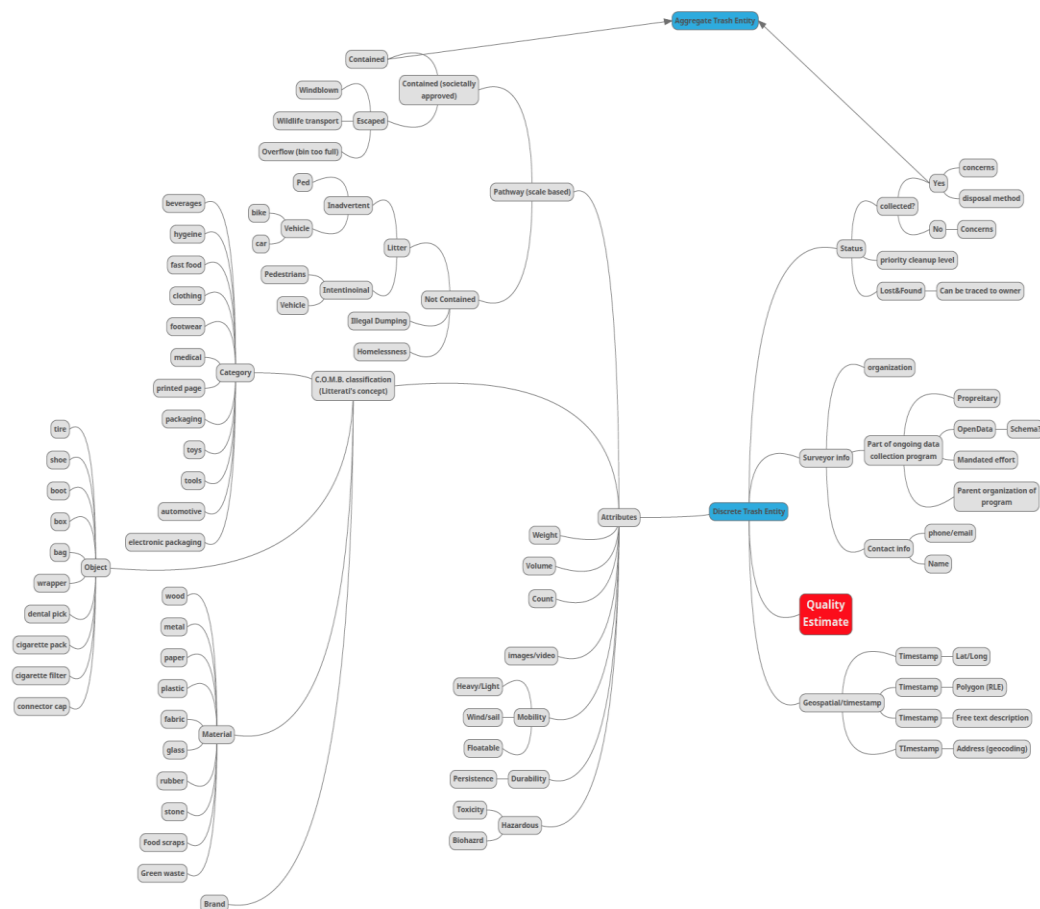
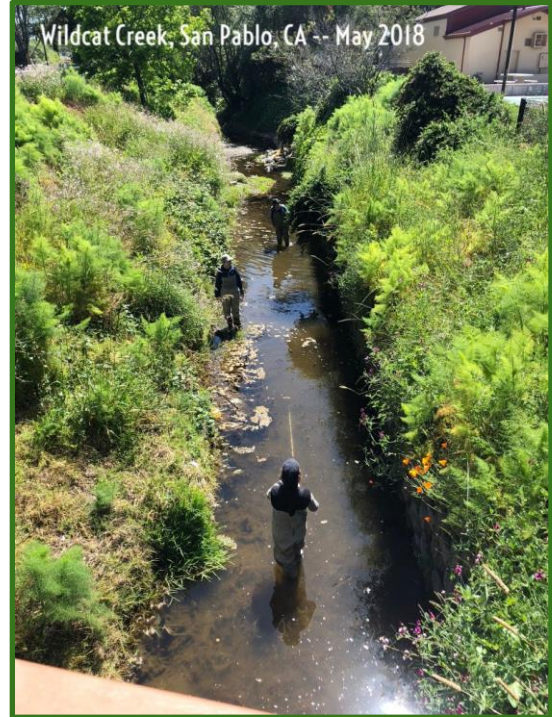


Diagram of the data model for trash. To view and download this image in pdf format click on the image and link.

Trash Method Evaluation Project

With the passage of the Trash Amendments in 2015 by the State Water Board, many programs across the state of California have sought to develop trash monitoring programs that would assess how much trash is entering receiving waters. Considering the many different ways to monitor trash, these programs have adopted various methods to address their primary management questions. As we look at the "big picture" of trash in California's creeks, streams, and wetlands, this variability in monitoring methods makes it very challenging to report, with any certainty, the amounts of trash and whether conditions are getting better or worse.

The California Ocean Protection Council (OPC), in close partnership with the State Water Resources Control Board, recognized the importance of standard methods for trash monitoring. The OPC seeks to accelerate the adoption of trash monitoring across the state by funding the Southern California Coastal Water Research Project (SCCWRP) and the San Francisco Estuary Institute (SFEI) to collaboratively develop and implement the California Trash Monitoring Methods Project. The project's goal is to develop a library of trash monitoring methods with known levels of precision, accuracy, and cross-comparability of their results, which will be linked to management questions. This session was presented by the project partners to inform and update stakeholders on the goals and current status of this project.



OUTCOMES

People's Choice Award

At the end of the day the TDD participants voted on the top idea for the day.

And the winner was...the **Power Portal Team!!!** Congratulations to the team for a great idea!

California Monitoring Council Trash Monitoring Workgroup

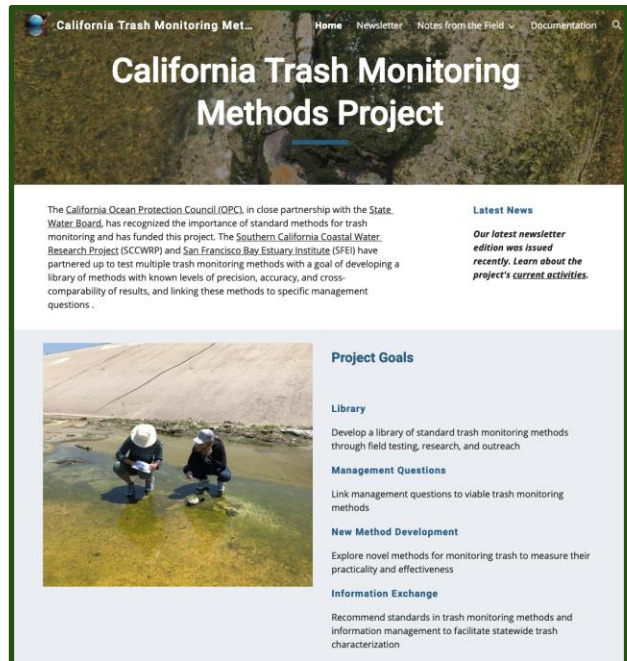
Given the level of participation by a large number of stakeholders and the great ideas that stemmed from the TDD, the State Water Resources Control Board created the Trash Monitoring Workgroup as part of the California Water Quality Monitoring Council. The goals of this workgroup are to use the momentum created during the TDD to promote and encourage the completion of the suggested projects that resulted from putting Trash Experts, Data Scientists and Interested Parties all in the same room. To sign up for the email list for this group go to the

[email signup page](#), choose General Interests and under Monitoring Council Workgroups & My Water Quality Portals choose the California Trash Monitoring Workgroup.

Trash Monitoring Website

The Ocean Protection Council and State Water Resources Control Board have supported a project to evaluate and develop methods for trash monitoring. As part of this project a website has been created to document progress of the project and monitoring efforts that are happening in California. A section of this website has been set aside to report the findings of the Trash Data Dive and will be used to update stakeholders on all things relative to monitoring. You can also sign up through the website for a newsletter of updates and happenings.

Please visit: trashmonitoring.org



APPENDICES

APPENDIX 1 - LIST OF THE 2018 TRASH DATA DIVE ATTENDEES

First Name	Last Name	Participant Type
David	Altare	Organizer
Alex	Aruj	Interested Party
Simona	Balan	Interested Party
Ramona	Beville	Interested Party
Ludovico	Bianchi	Data Scientist
Dale	Bowyer	Trash Expert
Carolynn	Box	Ambassador
Trixie	Bozin	Trash Expert
Nancy	Carr	Trash Expert
Nia	carrico-Diener	Trash Expert
Cortney	Ceccato	Interested Party
Tami	Church	Data Scientist
Kristine	Corneillie	Trash Expert
Sienna	Courter	Trash Expert

First Name	Last Name	Participant Type
Emily	Coven	Trash Expert
Win	Cowger	Data Scientist
Emma	Davidson	Data Scientist
Chrise	de Tournay	Trash Expert
Yanyi	Djamba	Data Scientist
Dagmara	E. Saini	Interested Party
Christine	Flowers	Trash Expert
Emilie	Fons	Data Scientist
Matthew	Freese	Interested Party
Alyssa	Gagnon	Trash Expert
Greg	Gearheart	Ambassador
kathy	grant	Interested Party
Nicole	Hack	Organizer
Tonia	Hagaman	Trash Expert
Tony	Hale	Ambassador
Liz	Hendrix	Trash Expert

First Name	Last Name	Participant Type
Orestis	Herodotou	Data Scientist
Andrew	Hill	Organizer
Trent	Hodges	Trash Expert
Peter	Haupt	Interested Party
Jaimie	Huynh	Interested Party
Ian	Kelmartin	Interested Party
Annalisa	Kihara	Interested Party
Jeff	Kirschner	Ambassador
Leila	Kordestani	Data Scientist
Pam	Krone	Data Scientist
Rich	Kwong	Trash Expert
Amelia	Labbe	Interested Party
Gwen	Lattin	Data Scientist
Sherry	Lippiatt	Ambassador
Chris	Lopez	Data Scientist
Laura-Lee	Love	Interested Party

First Name	Last Name	Participant Type
Emily	Lucic	Trash Expert
Kendra	Mann	Trash Expert
Nick	Martorano	Organizer
Meredith	McCarthy	Data Scientist
Lara	Meeker	Interested Party
Terra	Miller-Cassman	Trash Expert
Daphne	Molin	Interested Party
Shelly	Moore	Ambassador
Lisa	Moretti	Ambassador
Amanda	Orozco	Trash Expert
Michaela	Palmer	Data Scientist
Erika	Pinsker	Data Scientist
Emily	Pomeroy	Interested Party
April	Roeseler	Ambassador
Aiden	Sanchez	Trash Expert
Steven	Sander	Trash Expert

First Name	Last Name	Participant Type
Eben	Schwartz	Trash Expert
Scott	Silberman	Interested Party
Darin	Simmons	Data Scientist
Caleb	Smith	Interested Party
Jamie	Smith	Data Scientist
Jenny	Stephenson	Trash Expert
Lena	Tran	Data Scientist
Kirk	Van Rooyan	Interested Party
Nina	Venuti	Trash Expert
Ted	Von Bitner	Trash Expert
Tam	Vuong	Data Scientist
Shelly	Walther	Ambassador
Alisha	Wenzel	Data Scientist
Paula	White	Trash Expert
Elisabeth	Wilkinson	Trash Expert
Kate	Wing	Organizer

First Name	Last Name	Participant Type
Holly	Wyer	Ambassador
Walter	Yu	Ambassador
Leslie	Zellers	Trash Expert
Nick	Zigler	Trash Expert

APPENDIX 2 - LIST OF DATA AVAILABLE TO THE 2018 TRASH DATA DIVE DATA PARTICIPANTS

Dataset Name
Street Sweeper Data
Coastal Clean-up day
Annual solid waste
List of Landfills
NPDES
Marine Debris Items
SB Channelkeeper stream team
BASMAA
Adventure Scientist Microplastic project
Landfill locations
Coastal Cleanup Day Trash summary
Entangled animals
LA trash receptacles
LA clean streets
Southern California Bight Trawl Debris Data

Southern California Bight Riverine Trash Data
San Diego Coastkeeper Clean up data
Tuna Canyon trash images
Seal Beach trash images
Bay area trash images
Impairment of Waters in San Francisco Bay Region
Dataset Name
Ventura Stream Team Calleguas Creek Watershed
San Mateo Trash Assessment
SF Rapid Trash Assessment
Redemption and Recycle rates
Trash in Stormwater
Lottery ticket vendors
School locations
Underwater trash
Cost of waste management
Disposal export by county
Total trash created by city

Total trash disposed in a city
Alternative Daily Cover by material
Alternative Intermediate Cover by material
Beneficial Reuse
Fishing line recycling stations
Solid Waste Facility violations
Listing of Solid Waste Facilities and Disposal Sites Under Enforcement Orders
CA Communities Environmental Health Screening
Microplastics in the ocean
Manzanita Canyon trash over time
Alternative Daily Cover by facility
Alternative Intermediate Cover by facility
Dataset Name
Alternative Daily Cover by jurisdiction
Alternative Intermediate Cover by jurisdiction
Jurisdiction Transformation
Other beneficial reuse
Trash used as biomass (burned)

Transfer Station totals received and sent
Trash transfers
Transfer station allocations
Jurisdiction review status
Residential and Commercial waste by type
Solid waste facilities
Sacramento Clean-up locations
Traffic Data

APPENDIX 3 - LIST OF PRE-EVENT CHALLENGE QUESTIONS POSED BY THE EVENT AMBASSADORS

Challenge Questions
How do trash generation rates correlate with traffic count?
When applying machine learning to trash imagery, is it most effective to analyze by quantity, quality, or categories of trash?
What is the uncertainty associated with trash visual assessments for quantifying trash loading ?
How can machine learning be used to reliably quantify trash volumes and/or types in images?
Can we design an app to maximize public engagement that also captures useful data for track trash mitigation effectiveness?
Characterizing differences between upstream and coastal debris data - how does the assemblage of trash / litter / debris change within the watershed, and do we see a clear signal of ocean-sourced debris in the coastal data?
How do estimates of type and quantity of trash in local and regional trawl surveys contribute to knowledge of sources and loadings to Santa Monica Bay?
How can trawl trash data be combined with other datasets to estimate microplastics generated via trash in the ocean?
Do socioeconomic factors affect trash generate rates?
In Sacramento, If the vegetation is green in the summer, more than likely there is a hired landscaper that is also picking up trash. Would there be a way to look at satellite imagery and get correlations between green vs brown roadside edges and trash scores to use to exclude these areas from our priority land uses?
Do trash cans at bus stops alleviate litter issues or increase trash issues? Is there a difference between open trash cans and the compactors?
What are the costs per unit trash removal or acre treated using street sweeping, drain inlet devices, hydrodynamic separators, litter pick-up crews vs in-stream clean ups?
Are there any factors in rivers/streams that can predict the level of trash?

How much trash is in California?

Challenge Questions

What is the spatial distribution of trash?

What are the most prevalent trash types? Has this changed over time?

Create a comprehensive litter dataset to identify the most common item types according to volume, weight, flux, material, product, source, brand, and other units of importance.

Analyze the effectiveness of Trash Total Maximum Daily Loads (TMDLs) on reducing the flow of trash from stormwater systems to the ocean.

Analyze the effectiveness of full capture systems versus other best management practices as removing trash from waterways.

What trash removal credit could be claimed through a straw ban?

Are bans/tmdls effective (e.g. bag, styrofoam)?

Is there a difference in trash amounts in rivers/streams in Track 1 vs Track 2 areas?

Analyze the impact of the single-use plastic carryout bag ban on reducing disposable bag use, preventing ocean litter, and reducing government costs.

How do we identify trash mitigation actions that provide the greatest trash reduction benefits?

How do local, regional, and statewide policies, like bag bans, straw bans, and extended producer responsibility affect the amount and composition of debris found on California's beaches?

What sources of data are available to track the quantity and type of Tobacco Product Waste (TPW) ?

Where are TPW hot spots?

What are different group perceptions and actions on TPW? • Parks and Recreation • Environmental • Public • Tobacco Users