2018 NASCIO State IT Recognition Awards Submission

California Natural Resources Open Data Ecosystem

Category: Data Management, Analytics & Visualization

State: California

Contact:
Tim Garza
Agency Chief Information Officer
California Natural Resources Agency
(916) 653-8364
tim.garza@resources.ca.gov

Project Initiation Start Date: February 2017
Project Launch Date: December 2017

Executive Summary

While California is the 5th largest economy in the world, it is the state's natural resources that are its most treasured assets. California's natural resources are a key to its healthy economy and the California way of life that its citizens have. The state's natural resources consist of water and wetlands, ocean and coastal ecosystems, oil and gas, fish and wildlife, cultural lands, forestry and parks, natural energy, and rivers and mountains.

It is the mission of the California Natural Resources Agency, working with stakeholders, to manage, protect, and restore the state's natural, historical and cultural resources for current and future generations. The California Natural Resources Agency consists of thirty-six different organizations and 22,000 employees, who work with local, federal, public, and private entities and stakeholders to achieve the mission through use of creative and innovative solutions and approaches based on technology, science, engineering, and collaboration.

The ability to effectively manage, restore, and protect California's natural resources has become very challenging with the rise of climate change, increased population, droughts, wildfires, and other factors. The average global temperature has increased at the fastest rate in recorded history and experts predict the trend is accelerating, all but one of the 16 hottest years in NASA's 134-year record have occurred since 2000. California's population has expanded to nearly forty million. California has experienced excessive periods of droughts over the last seven years. The number of wildfires have increased and in 2017 reached a record high.

To help meet the challenges, California State leaders, led by Governor Edmund G. Brown Jr., have passed various legislation and issued Executive Orders to address the state's carbon footprint, management of surface and ground water, protection and restoration of fish and wildlife habitats and species, use of renewable energy, and use of protected lands.

The intent of these legislative and executive actions is to reduce California's carbon footprint, decrease non-renewable energy usage and greenhouse gas emissions, better manage water usage, stop ocean and coastal erosions, and protect forests, wetlands, and wildlife and fish habitat from risk.

As the Governor has called on other leaders (public and private) to join in the fight against climate change and other factors that threaten California's valued natural resources, a key enabling element was required, data and information openness and transparency. The efforts to gather hundreds of researchers and scientists around the call to action, required translation of environmental and scientific data and findings from disparate fields into an Open Data Ecosystem.

Recognizing the need for, and the benefits and value of a data ecosystem for environmental information and directional use, the California Natural Resources Agency (CNRA), in partnership with various stakeholders, developed and launched an Open Data Ecosystem that unifies the concepts of "useful, usable, and used" data and information for all parties related to California's natural resources.

Project Narrative

Concept

Problem:

The ability to effectively manage, restore, and protect California's natural resources has become very challenging with the rise of climate change, increased population, droughts, and wildfires. The average global temperature has increased at the fastest rate in recorded history and experts see the trend is accelerating. California's population has expanded to nearly forty million. The state has experienced excessive periods of droughts over the last seven years. The number of wildfires have increased and in 2017 reached a record high.

The California Natural Resources Agency prepared reports on California's Climate Adaptation Strategy, and the Agency also produced three Climate Change Assessments based on peer reviewed science. Those reports detail the existing and expected impacts of climate change, drought, population growth, and other factors in California. Some key impacts, but not all, include:

- **Sea level rise, coastal flooding and coastal erosion**
  85% of California's population live and work in coastal counties. Sea level along California's coasts has risen nearly 8 inches and is projected to rise by as much as 20 to 55 inches by the end of the century. Major sea level rises could increase risk of flooding by 2100, and threaten $100 billion in property and infrastructure. Coastal erosion could have a significant impact on California's ocean-dependent economy, which is estimated to be $46 billion per year. As sea levels rise, saltwater contamination of the state's delta and levee systems will increase. Saltwater contamination will threaten wildlife and the source of drinking water for millions of Californians and harm farmland in low areas.

- **Losses to the Sierra snowpack and water supply**
  The Sierra Nevada snowpack functions as the most important natural reservoir of water in California. Under current conditions, the snowpack is created in fall and winter and slowly releases about 15 million acre-feet of water in the spring and summer, when California needs it most. California's dams and water storage facilities are built to handle the snow melt as it happened in the past. Higher temperatures are now causing the snowpack to melt earlier and all at once. Earlier and larger releases of water could overwhelm California's water storage facilities, creating risk of floods and water shortages.

- **Forestry and higher risk of fires**
  Forest and rangelands cover over 80% of California's 100 million acres. Climate change and drought affect tree survival and growth, reducing these lands' productivity and changing their habitats. Climate change makes forests more vulnerable to fires by increasing temperatures and making forests and brush drier. Wildfire occurrence statewide could increase several fold by the end of the century, increasing fire suppression and emergency response costs and damage to property.

- **Damage to agriculture**
  Global warming can cause drought, higher temperatures, saltwater contamination through rising sea levels, flooding, and increased risk of pests. These changes pose a very serious threat to California's agricultural industry, which generates billion in revenue and which is responsible for more than half of all US domestic fruits and vegetables. Because California feeds not only its own residents, but the entire U.S. and other countries as well, production declines could lead to food shortages and higher prices.

- **Habitat destruction and loss of ecosystems**
  California is one of the most biologically diverse regions of the world, with the highest number of unique plant and animal species of all 50 states and the greatest number of endangered species. Climate change will adversely affect plant and wildlife habitats and the ability of the State's varied ecosystems to support clean water, wildlife, fish, timber and other goods and services important for our well-being.

- **Energy usage and Public health impacts**
  Increased demand for electricity. Higher temperatures and more heat waves will drive up demand for electricity. As people turn up their air conditioners, increased electricity use will be greatest in southern California and the Central Valley, and may be as high as 60% above present demand by the end of the century. Hotter temperatures lead to more smog, which can damage lungs, and increases childhood asthma, respiratory and heart disease and death. Certain segments of the population are at greater risk.

A problem area that the California Natural Resources Agency had in addressing California's natural resources issues was the ability to effectively collect, share, analyze, visualize, and publish large and disparate volumes of data in a useful and usable manner in collaboration with stakeholders, researchers, local/state/federal agencies, and other public and private entities.

The goal was to create an open data culture, enhancing data sharing and data use across the various stakeholders related to the utilization of data and analytics in addressing California's natural resources challenges.

To meet the goal, CNRA needed to implement and maintain an Open Data Ecosystem that will be utilized to enhance data collection, data transparency and access, and which would provide large amounts of high-value datasets. The Open Ecosystem could be utilized across a multitude of stakeholders that are committed to maximizing the use and value of the data/information to drive progress toward achieving the goals of California natural resources initiatives, and strategically and collaboratively address the issues facing the state's natural resources environment.

Solution:
Implement and maintain a “New Style” Open Data Ecosystem. The new CNRA Open Data Ecosystem has both Open Data and Analytics capabilities, capacity, and interoperability.

The California Natural Resources Agency used an Agile Methodology approach and utilized a combination of open source and best practice technologies to architect and implement the CNRA Open Data Ecosystem environment. The CNRA Open Data Ecosystem consist of interoperable components that provide both enterprise industry level open data platform and big data analytical platform capabilities and features.

The Open Data Ecosystem is a comprehensive solution that is flexible, scalable, and addresses the complex type of datasets and business functionality required by CNRA programs and their constituencies, stakeholders, partners, and the public. The CNRA Open Data Ecosystem launched officially in December 2017 and has over 1,500 data/information sets currently and is growing.

The CNRA Open Data Ecosystem is an interoperable environment that supports the entire data information value chain including: collection; analytics; visualization; and publishing. See Appendix [Diagram A] on page 8 for a high-level view of the CNRA Open Data Ecosystem environment, and below for a list of some key features:

- Built on CKAN and Hadoop open-source platforms
- Strong API capabilities and customization for special requirements
- Addresses large and small data volume capacity
- Data transformation and data extractor capabilities
- Flexible data capabilities: spatial; time-series; tabular; documents; images; maps; etc.
- Federation and data harvesting capabilities
- Powerful analytical and visualization capabilities
- Data import and export capabilities
- Rich search capabilities and rich set of metadata for each dataset and data dictionary
- Machine-readable format and role based security
- Provides big data analytical capabilities and capacity

In addition, the Open Data Ecosystem environment has strong and powerful big data analytical capabilities built on a Hadoop open source platform. CNRA program areas and defined collaboration partners can utilize this platform to perform complex analytical analysis and modeling of large and disparate datasets. The results could be published and utilized to make natural resources directional decisions.

The CNRA Open Data Ecosystem environment enhances data sharing across the Agency’s program areas and changes the way they interface with and provide information to constituencies, stakeholders, partners, and the public. The Open Data Ecosystem features data, showcases, analysis, and information visualizations related to Water, Wildlife, Conservation, Energy, Protected Lands, Oceans, and much more.

Making data open and accessible has a significant value to the management and decision-making processes of California's natural resources. The following benefits were identified as beneficial for the implementation of the CNRA Open Data Ecosystem:

- Improve Information Transparency
- Improve Data Accessibility
- Make Data Useful to Everyone (all interested parties)
- Improve External and Internal Data Contributions
- Improve Data and Information Collaboration
- Provide Optimization of Open and Analytical Data features and technical capabilities
- Help Transform the Natural Resources Community to an Open Data Sharing Culture

Significance

The California Natural Resources Open Data Ecosystem is significant because it is an invaluable resource for all parties and stakeholders interested in California's natural resources. It provides organizations, citizens, stakeholders, and interested parties better access to holistic data and allows them to use it in a manner that fits their needs. It helps local, federal, and state agencies make natural resources-related decisions by giving them access to data that was previously locked in silos.

The CNRA Open Data Ecosystem was developed to provide information to California citizens, agencies, and interested stakeholders and parties in a transparent and useful manner. It supports the CNRA organizations programs' missions by providing an environment to collect, share, analyze, visualize, and publish data that can be utilized effectively.

Recognizing the significant benefits and value from an active open data ecosystem and program, California launched into a new era of utilizing a multitude of datasets to set direction, analyze impacts, make adjustments, and communicate with affected and interested parties related to California's valued natural resources.

By utilizing data on the Open Data ecosystem and engaging stakeholder participation and contributions, all parties are able to assist in addressing California's natural resources challenges. All stakeholders are able to judge the affect and progress of California's efforts to manage, restore, and protect its valued natural resources.

The success of the CNRA Open Data Ecosystem environment was brought about as a result of business and information technology alignment that created enabling capabilities and collaboration for the state's natural resources stakeholders. In a broader perspective and in alignment to the California State Information Technology Strategic Plan and the California Natural Resources Agency Strategic Direction, the Open Data Ecosystem functions as a statewide resource to do the following:

- **Improve Government:** Open data and analytics improves government, by increasing transparency, and enhancing services and direction.
- **Empower Citizens:** Opening access to government data gives citizens and stakeholders the ability to make informed decisions and impact change.

- **Create Opportunity:** Open data creates new opportunities for interested parties and organizations by fostering innovation.
- **Solve Public Problems:** Open data plays an increasingly important role in solving large problems, primarily by allowing citizens and policymakers access to new forms of data-driven assessment of the problems at hand. It also enables data-driven engagement, producing more targeted interventions and enhanced collaboration.

Impact

The CNRA Open Data Ecosystem is an invaluable resource for all parties and stakeholders interested in California natural resources related data and information. It gives program areas, agencies, citizens, stakeholders, and interested parties better access to holistic data and allows them to view and use the information in a manner that fit their needs.

The CNRA Open Data Ecosystem is a collaborative data environment that will continue to grow and progress over time. The more data that is collected, analyzed, and published the more valuable the ecosystem, as a resource, will be.

The use of the new Open Data Ecosystem environment is open to all CNRA organizational programs and all California's natural resources stakeholders or interested parties. The California Natural Resources Agency plans on continuing to work with departments, commissions, boards, conservancies, local agencies, researchers, and others to add much more data, visualizations, and features to it. The Open Data Ecosystem already features data, showcases, analytical results, and information visualizations related to water, fish and wildlife, conservation, energy, protected lands, and oceans.

The Open Data Ecosystem facilitates government improvement by assisting officials in critical decision-making based on data, transparency, innovation and stakeholders' engagement.

With the launch of the CNRA Open Data Ecosystem, the California Natural Resources Agency has already seen improvements in data collection, data analysis, data access, data sharing and collaboration, and data driven decision-making processes.

Links: California Natural Resources Agency Open Data - https://data.cnra.ca.gov

Appendix

[Diagram A]: