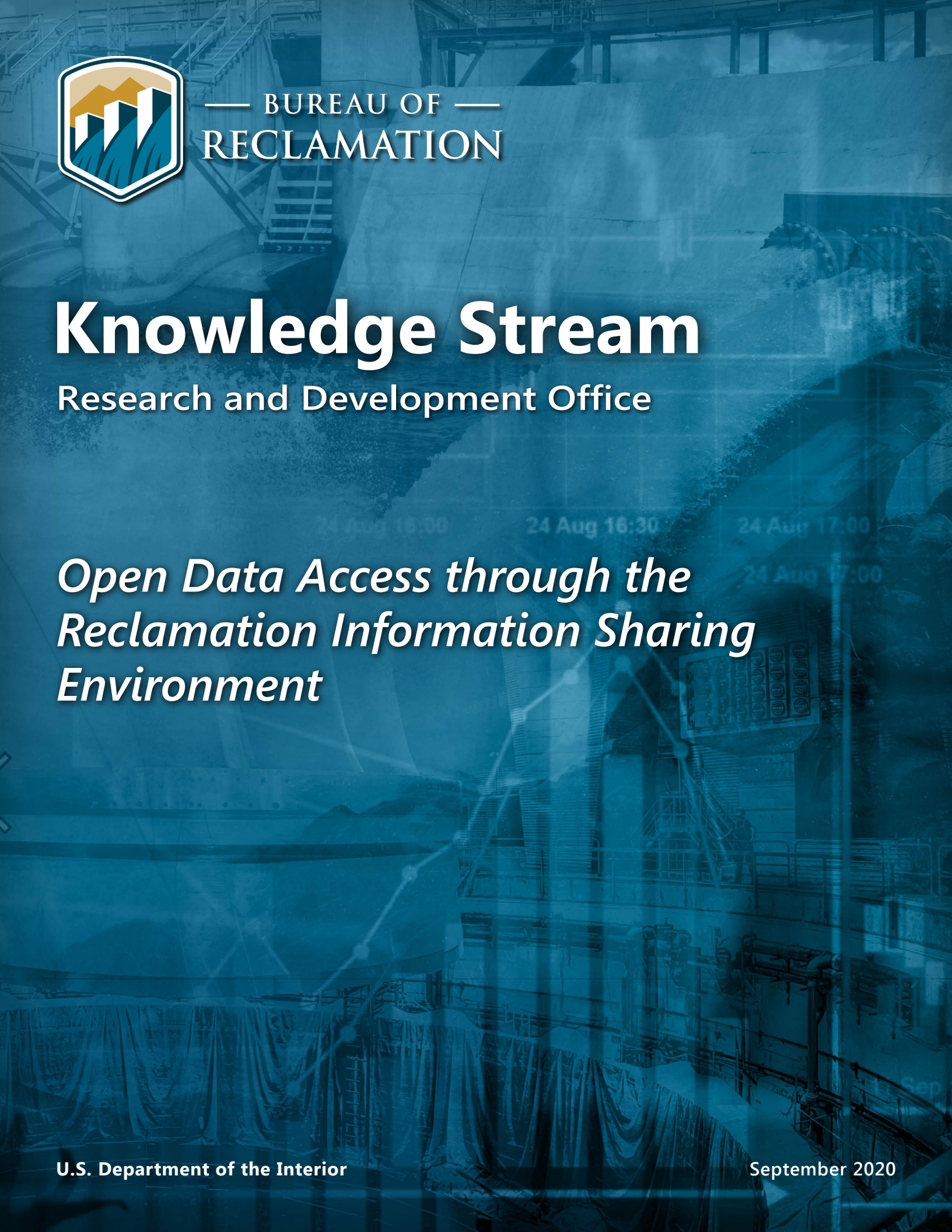


BUREAU OF RECLAMATION

Knowledge Stream

Research and Development Office

Open Data Access through the Reclamation Information Sharing Environment

U.S. Department of the Interior

September 2020

Research and Development Office (R&D) Contacts

Program Manager
Levi Brekke
lbrekke@usbr.gov

Science and Technology Program Administrator
John Whitler
jwhitler@usbr.gov

Desalination and Water Purification Research Program Administrator & Advanced Water Treatment Coordinator
Yuliana Porras-Mendoza
yporrasmendoza@usbr.gov

Hydropower and Water Infrastructure Research Coordinator
Erin Foraker
eforaker@usbr.gov

Water Availability Research Coordinator
Ken Nowak
knowak@usbr.gov

Technology Transfer Specialist
Vacant
research@usbr.gov

Prize Competitions Program Administrator
Jennifer Beardsley
jbeardsley@usbr.gov

Open Water Data Coordinator
Allison Odell
aodell@usbr.gov

Programs Analyst
Katie Hill
khill@usbr.gov

Budget Analyst
Rosann Velnich
rvelnich@usbr.gov

Administrative Assistant
Pat Loetel
ploetel@usbr.gov

Message from R&D

Welcome to this issue of the Knowledge Stream. In this issue we spotlight Reclamation's new platform – the Reclamation Information Sharing Environment (RISE) – for modernized sharing of mission data that benefits the public, the private sector, and the U.S. economy. RISE includes troves of data generated by Reclamation as it carries out its water and power mission, including data on water operations, water quality, infrastructure, hydropower, and environmental resources.

Development of RISE responds to federal policy and legislative mandates to make federal data assets more "open" – that is, more easily found, shared and used by public and private sectors. Through RISE, Reclamation programs and offices can publish their datasets in open formats, enabling them to be accessed via a map interface, catalog search, time series data query, or web-service. Data types include observed, modeled or analyzed data; time series or geospatial data; and miscellaneous data such as reports and media files. RISE also centralizes Reclamation's data publication, making it faster and easier for users to find and obtain the data they need. RISE is supported by the R&D Open Water Data program through partnerships with Reclamation regions, programs, and enterprise information management technology services.

Whether you're a data creator, data user, or just curious about what RISE has to offer, this issue of the Knowledge Stream has something for you. It will introduce you to RISE and give you insights into how you can use it. Please visit RISE and take a look around (https://data.usbr.gov). If you have questions or feedback on the RISE site, you may reach out to the RISE team at either data@usbr.gov or via the RISE contact form (https://data.usbr.gov/contact).

As always, we appreciate you reading about innovation funded by Reclamation's R&D programs. Please enjoy this issue of the Knowledge Stream and offer us any feedback for improving our strategies to transfer solutions to users.

About the *Knowledge Stream*

The *Knowledge Stream*, published by the Bureau of Reclamation's Research and Development Office, is a quarterly magazine bringing mission-critical news about the agency's innovations in the following:

- Science and Technology Program
- Desalination and Water Research Purification Program
- Prize Competitions
- Technology Transfer
- Open Water Data...and more.

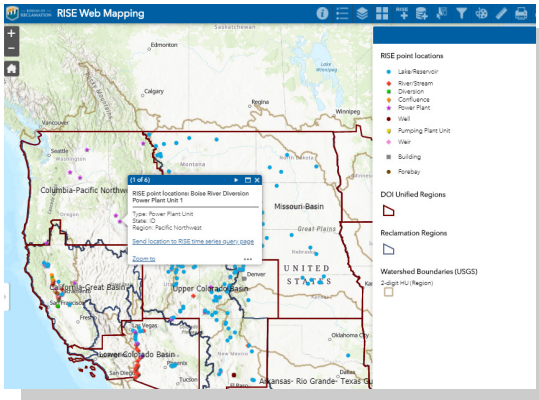
September 2020 Content Lead

Contact
Allison Odell
Open Water Data Coordinator
aodell@usbr.gov

More Information
https://data.usbr.gov
data@usbr.gov

For more information on articles within this issue, please contact listed author or Ms. Odell.

Contents



Covers: FRONT: Picture collage BACK: Picture collage

Community Needs

Unlocking Access to Reclamation's Mission Data

By Levi Brekke, Program Manager
lbrekke@usbr.gov

When federal agencies publish their mission data, they benefit citizens, businesses, and the U.S. economy. Federal agencies generate large amounts of publishable data as they carry out their missions. Reclamation, for example, generates much data describing water delivery, hydropower generation, environmental resources, and other conditions of interest. While some of these data are published, large amounts are managed with a focus on data archiving rather than a focus on publishing.

Easing public and private sector access to federal data is critical to maximizing benefits gained from generating those data resources.

In 2013, the U.S. Office of Management and Budget (OMB) issued a new Open Data Policy, recognizing that "making information resources accessible, discoverable, and usable by the public can help fuel entrepreneurship, innovation, and scientific discovery - all of which improve Americans' lives and contribute significantly to job creation."

In 2019, The Open Public, Electronic, and Necessary (OPEN) Government Data Act was signed into law, requiring federal agencies to publish most government data in open formats.

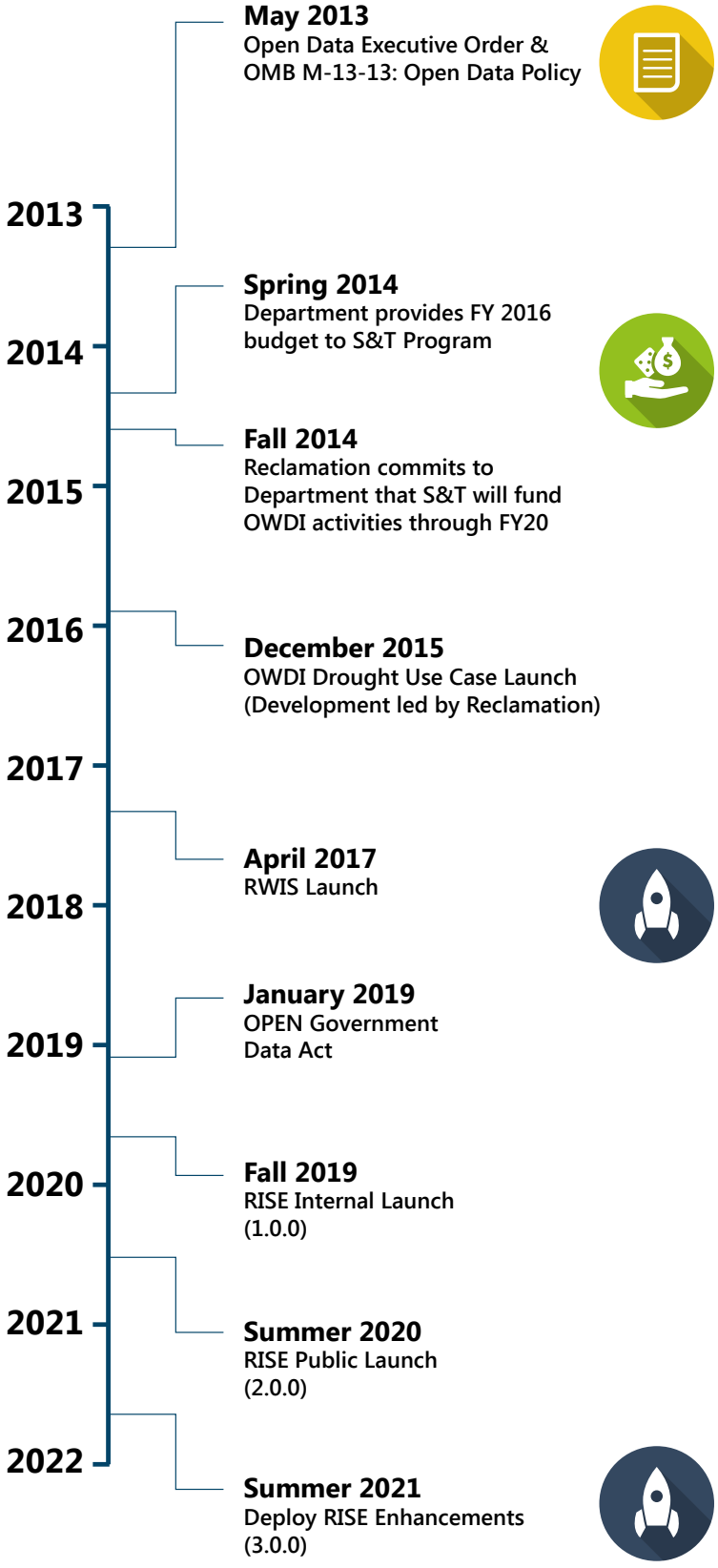
To implement the Open Data Policy and the OPEN Government Data Act, federal agencies are challenged to migrate legacy data publishing systems to those featuring open access ideals: data stored in common and machine-readable data formats, data served through modern web portals under open license, and equipping users with data discovery tools that inform how data fit their needs.

Reclamation's open water data journey began with two events in 2014. The first was a decision by the Department of the Interior (DOI) to include "Open Water Data Initiative" (OWDI) budget in Reclamation's Fiscal Year 2016 budget request via the Science and Technology (S&T) Program; OMB later approved and increased this request. The second was a commitment by Reclamation to have the S&T Program fund OWDI activities through FY 2020, namely to support planning, development and operation of new systems that would provide a common platform for Reclamation programs to publish data and a common web portal to ease public and private sector access to those data.

Early Reclamation efforts focused on formulating an Open Data Team to assess needs, partnering with U.S. Geological Survey to demonstrate an open data use case on drought in the Colorado River Basin, and developing Reclamation's first open data system, the Reclamation Water Information System (RWIS), which serves time-series water data from Reclamation's reservoirs across the West. Positive reactions to these efforts underscore the opportunity for Reclamation to invest in open data systems that provide broad benefits.

In FY 2017, Reclamation pivoted efforts to a large-scale solution, initiating development of the Reclamation Information Sharing Environment (RISE). RISE scope includes all types of publishable data that Reclamation generates in its water and power mission. Subjects include reservoir conditions, water delivery, hydropower generation, invasive species, threatened or endangered species, habitat conditions, water quality, etc. Types includes observations, analyses, time series data, geographic information, reports, and others. Data may relate to water and power management, operations, and field activities, as well as developments in science and technology.

RISE development has been a bureauwide collaboration, involving contributions from Region data administrators, Reclamation's Information Resources Office (software programmers, database developers, and cybersecurity professionals), Public Affairs, and the Science Advisor organization. Moving forward, the RISE team will seek feedback from RISE data providers and consumers to inform expansion of RISE data stores and features. The team also plans to implement regular technology refresh to sustain and advance RISE open data publishing capabilities.



Key Perspectives

Voices of RISE

Reclamation staff members who collect, produce, and manage data are the key to RISE's success. They also know firsthand what the value of RISE is to Reclamation, so we asked a few of them to share their thoughts on why they are using RISE and what they are excited about as people begin to access their data through RISE.

Lower Colorado River Multi-Species Conservation Program (LCR MSCP)

Travis Sizemore Physical Scientist / GIS Specialist
David Gundlach Physical Scientist / GIS Specialist
Jimmy Knowles Adaptive Management Group Manager

The Lower Colorado River Multi-Species Conservation Program (LCR MSCP) was created to balance the use of Colorado River water resources with the conservation of native species and their habitats. LCR MSCP's scientists routinely collect biological data on over 20 plant and animal species along 400 miles of the LCR, from Lake Mead to Mexico. Nearly all of LCR MSCP's data includes spatial and biological attributes, such as species abundance, location, apparent health, behavior, and interactions. LCR MSCP's plant and wildlife biologists use the data for species-specific studies, long-term monitoring, and decision-making in cooperation with LCR MSCP's engineers and field staff who help design and maintain conservation areas and raise and release native fish.

Why do you or your office/group want to publish data in RISE?

The LCR MSCP wants to publish our data in RISE because we support the OPEN Government Data Act's presumption that data should be published as machine-readable data in an open format, and under open licenses. We strongly support Reclamation's responsibility in complying with this legislation. We feel that it is important that our stakeholders and the general public have access to the data that they helped to pay for. LCR MSCP's goal is to share, to the fullest legal extent possible, all the data that we collect.

How do you expect RISE users (both internal and external) to use your data?

Programs with data-driven, adaptive management components can use RISE data to conduct analyses and make comparisons within and across regions that might otherwise not be possible.

External users of RISE may use the standardized information found there to discover how our data collection efforts complement and/or supplement their own research.

What feature(s) of RISE are most useful for interacting with your data? What are you most excited to do with your data in RISE?

As data scientists, we often search the web for new data sources and understand the challenge of finding high-quality data that has been properly validated and given appropriate metadata. Too often data are uploaded to servers without the proper quality control and are made difficult to locate or access to external users.

Elf Owl at the Bill Williams National Wildlife Refuge

Relict Leopard frog at Red Rock Spring, NV

We see RISE as a mechanism to ensure data quality and make it available in an easily accessible and user-friendly portal. This will help us effectively share our data with the broader scientific community and any member of the public who seeks to learn about LCR MSCP's resources.

Brackish Groundwater National Desalination Research Facility (BGNDRF)

Randy Shaw Research Facility Manager
Zach Stoll Research Scientist

The Brackish Groundwater National Desalination Research Facility (BGNDRF) collects various datasets that inform potential clients and researchers about the quality of waters available for testing. These waters include four on-site well waters and evaporation ponds (all time series). Collecting data in this manner allows BGNDRF staff and researchers to see fluctuations in well water quality, and informs decisions on how to best operate and maintain the evaporation ponds.

Why do you or your office/group want to publish data in RISE?

BGNDRF has historically provided this data on their website, however, typically researchers and potential clients were only aware of the data because they directed them to the website.

How do you expect RISE users (both internal and external) to use your data?

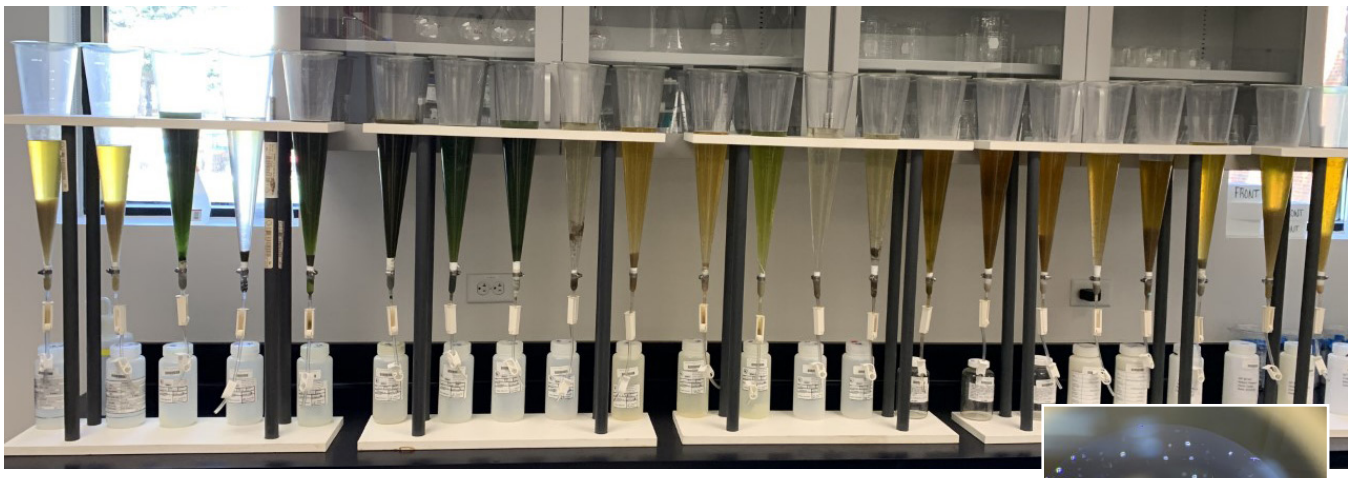
We anticipate that external RISE users will consist of our clients, who are usually consulting firms, University researchers/professors, and entrepreneurs. Historically, they have used the BGNDRF's well water and evaporation pond data as inputs to their modeling, which ultimately help inform them how to design a novel treatment system, or what the operational parameters should be for a system. Other possibilities we could envision are utilizing the time-series data to understand how our brackish groundwater chemistry changes over time, and/or by location (if compared to other brackish groundwater well data).

What feature(s) of RISE are most useful for interacting with your data? What are you most excited to do with your data in RISE?

Being able to select certain species is most exciting to us, because most of the time, their clients are only interested in a handful of chemical constituents. We're most excited to be contributing all this data in a centralized database that the public has access to.

"By publishing our data in a Reclamation-wide data base, we hope to obtain more visibility and the opportunity to let a broader audience know that we exist!" – Zach Stoll

BGNDRF's state-of-the-art facility including lab space, indoor and outdoor testing sites, brackish water wells, and office space, supports research for a variety of constituents.



Invasive mussel early detection water samples are settled in Imhoff cones for more efficient analysis. Only the bottom 15 mL of the settled sample needs to be analyzed under a microscope after the mussel larvae fall to the bottom of the cone.

Technical Service Center (TSC) Ecological Research Laboratory (EcoLab)

Yale Passamaneck Biologist
Rheannan (Annie) Quattlebaum Biologist

The Technical Service Center's Ecological Research Laboratory (EcoLab) is publishing data related to the early detection and monitoring of invasive dreissenid mussels (quagga and zebra mussels). Since 2008 the EcoLab has analyzed over 23,000 samples from across Reclamation's regions for the presence of mussels by looking for veligers, the free-swimming larval stage of the mussel. Early detection data has also been collected by Reclamation's Columbia-Pacific Northwest Soil and Water Quality Regional Laboratory. Data on sample collection will be published in RISE. Data on counts of veligers from waterbodies with established populations will also be published.

Microscopic quagga mussel larvae (veligers) glow against a dark background when viewed under a microscope using cross-polarized light. This particular sample had approximately 649,000 veligers in it.

Why do you or your office/group want to publish data in RISE?

Reclamation's monitoring has generated one of the largest available datasets on monitoring for invasive species, in terms of both its temporal and geographic breadth. Making these data available will provide an important resource for understanding mussel invasions and helping to prevent their further spread.

How do you expect RISE users (both internal and external) to use your data?

We expect that the publication of mussel data will be used in a wide variety of applications, including studies of mussel population dynamics and gap analyses to identify where monitoring efforts may need to be intensified in the future. We are hoping that these data will inspire analyses we have not yet considered and will benefit efforts to combat these harmful species.

What feature(s) of RISE are most useful for interacting with your data? What are you most excited to do with your data in RISE?

Pairing mussel monitoring data with the mapping capabilities of RISE will provide a powerful tool for visualizing the dataset.

California-Great Basin's Environmental Monitoring and Assessment Branch

Melanie Lowe Physical Scientist
Lynn Razzini Physical Scientist
Laurel Dodgen Physical Scientist

The California-Great Basin Environmental Monitoring and Assessment Branch primarily collects water quality data, and to a lesser extent sediment and biota, for many analytes and sites across the California-Great Basin region. The data is generally used to characterize and assess site conditions as they relate to environmental quality goals for designated uses.

Why do you or your office/group want to publish data in RISE? Consider how publishing the data in RISE will positively impact your work and how it will benefit internal or external users of your data.

Publishing in RISE supports Open Data initiatives and makes our extensive monitoring data available to other users to aid their management planning and decision-making.

How do you expect RISE users (both internal and external) to use your data?

We hope that the data will support study planning, decision-making, and guiding research questions about environmental health.

What feature(s) of RISE are most useful for interacting with your data? What are you most excited to do with your data in RISE?

We are excited to use RISE's visual display tools.

Surface water collection at Sliger Mine.

Soil core collection at a former airfield near the Delta Mendota Canal.

Power Resources Office

Clark Bishop

Program Analyst

Reclamation's hydropower program collects data to satisfy a number of program needs - including operations, maintenance, and investment decision making, performance assessments, stakeholder outreach, and compliance with mandatory reporting requirements.

The Power Resources Office oversees the Power Operations and Maintenance Tracking System ("POMTS"), Reclamation's internal power operations and maintenance database, populated by regional power personnel. The POMTS database stores, in part, Reclamation powerplant capacity and generation values, unit starts and run times, outage and maintenance records, and associated water data (e.g. elevations and releases) – all reported on a monthly timestep in accordance with Facilities Instructions Standards and Techniques (FIST) Volume 1-2 Conduct of Power Operations available on the public-facing FIST **website: https://www.usbr.gov/power/data/fist_pub.html.**

Why do you or your office/group want to publish data in RISE?

Non-sensitive POMTS data elements – including plant and unit capacity and generation values will be published to RISE, complementing existing datasets published to the Reclamation Power Program **website: https://www.usbr.gov/power/data/data.html**

POMTS data elements to be published to RISE are in demand by internal and external data consumers and making these elements available – in formats conducive to analysis, alongside complementary domain data (e.g. water data) will improve the data consumer experience and alleviate burdens on Reclamation Offices responding to power data requests.

This would not have been possible without support from the Reclamation hydropower community, the Research and Development Office, and the Denver Application Service Group.

How do you expect RISE users (both internal and external) to use your data?

Given past requests – I would expect RISE users (both internal and external) to use power data to inform a number of tasks, including, but not limited to, broader power system assessments, hydropower resource and grid support analyses, economic analysis, performance evaluations, etc. With that said, I'm sure users will find new and novel uses for published power data.

Glen Canyon Dam, AZ

What feature(s) of RISE are most useful for interacting with your data? What are you most excited to do with your data in RISE?

Specific data needs would determine what feature is the most helpful in interacting with the data. I'm most excited to see what others do with published data and what findings can be gleaned – especially when integrating other domain data.

Reclamation staff perform avian monitoring on the Rio Grande.

TSC Fisheries and Wildlife Resources Group
Rebecca Siegle Natural Resources Specialist

The TSC Fisheries and Wildlife Resources Group collects data related to aquatic and terrestrial wildlife communities in support of understanding and reducing impacts of water development facilities and operations. Southwestern Willow Flycatchers and Western Yellow-billed Cuckoos are federally listed species of concern, both of which nest along the Rio Grande. Therefore, projects must comply with the Endangered Species Act in order to protect these species from significant impacts caused by project activities. All avian data were collected along the Rio Grande from Los Lunas, NM to the U.S./Mexico border and are presented as geospatial. Territory data includes the location and number of territories by avian survey site for the previous five years. Avian habitat suitability data are identified using vegetation maps, updated every 4 to 5 years. Suitability was determined by analyzing avian territory sites to identify habitat types most often occupied by flycatchers and cuckoos as well as other factors, including nest success and population size. These data are used to inform management decisions by locating current breeding sites as well as potential breeding sites. Timing and location of construction projects are designed to cause the least disruption at breeding sites.

Why do you or your office/group want to publish data in RISE?

RISE will provide direct access to data for users, bypassing the need for our group to send data based on individual requests.

How do you expect RISE users (both internal and external) to use your data?

Data are often requested by entities managing projects along the Rio Grande, as well as by experts who study these species and are interested in population and habitat data to provide a better understanding of the species' status.

What feature(s) of RISE are most useful for interacting with your data? What are you most excited to do with your data in RISE?

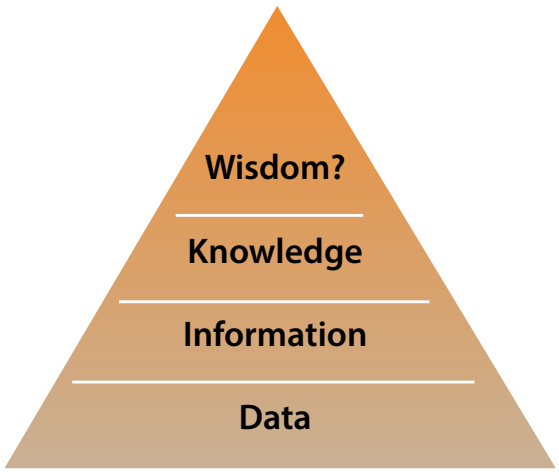
The ability to provide data on a spatial scale that will not require us to send numerous files in various formats and will be readily and easily available to those requesting this information.

What is Open Data and Why is it Important

By Jim Nagode, Data Resource Manager
jbnagode@usbr.gov

Back in 2009, the Open Government Directive launched all Federal agencies into the world of increased transparency through a digital data strategy. I remember being at a conference in 2010 when I learned that the National Science Foundation had just created a policy requiring that all funded research proposals were to include a data management plan and if awarded, provide access to their datasets. In a few short years, the Office of Management and Budget (OMB) adopted this approach and issued the Open Data Policy—Managing Information as an Asset, stating that "making information resources accessible, discoverable, and usable by the public can help fuel entrepreneurship, innovation, and scientific discovery—all of which improve Americans' lives and contribute significantly to job creation". While Reclamation's Science and Technology program was able to leverage these incremental policies of open data to stand up its first 'open' publishing platform for enterprise water data in RWIS, there was no law that compelled Federal agencies to create systems and applications to support open data.

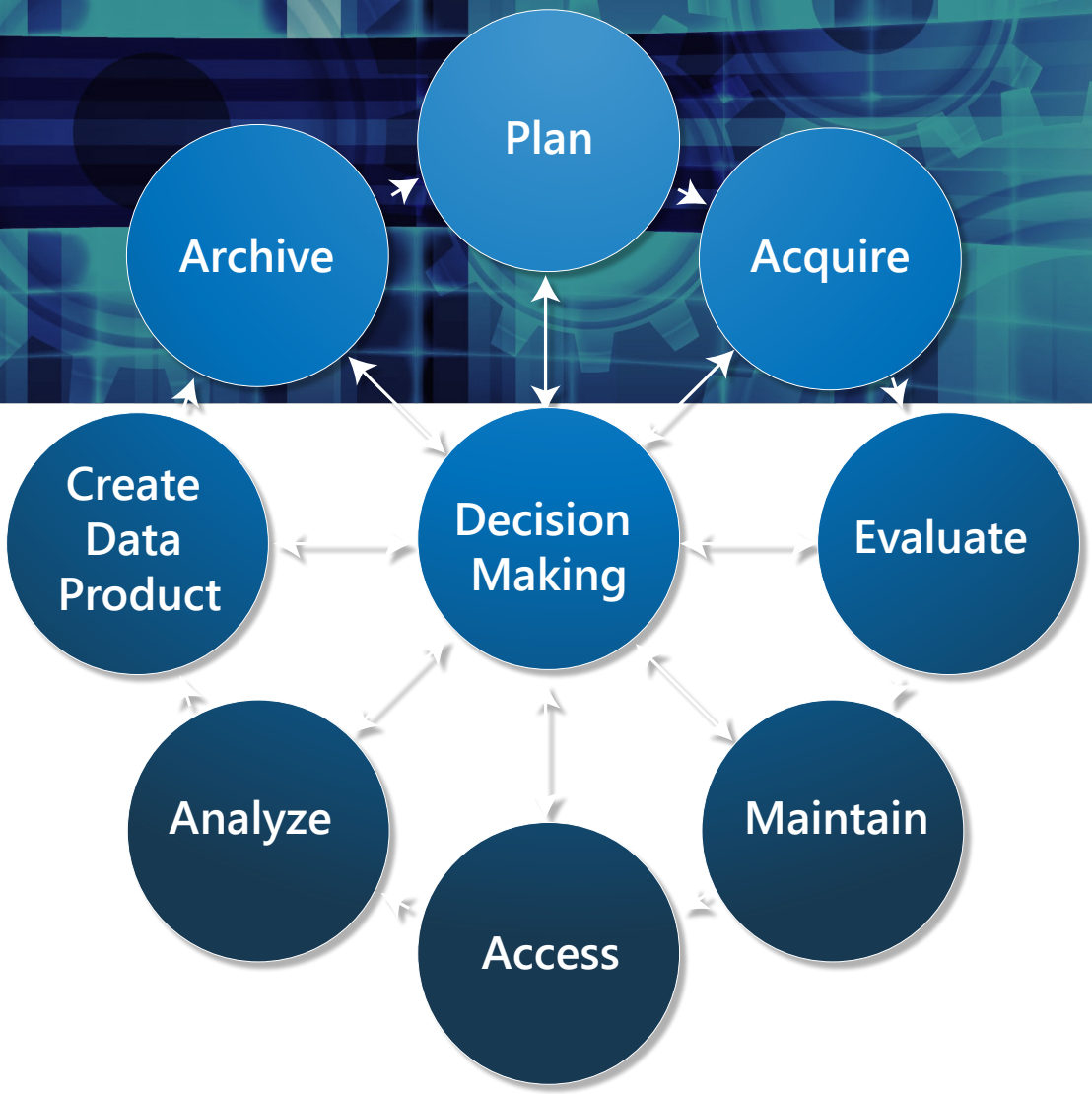
In 2018, Congress passed the Foundations for Evidence-Based Policymaking Act, which, and under Title II it created the OPEN Government Data Act. This codified many elements of the preceding Open Data Policy and was signed into law in January 2019. The law requires that government data assets be made available by Federal agencies (subject to confidentiality, security, and privacy restrictions) and free to the public in standardized, open formats. Then in 2020, OMB created the Federal Data Strategy, which guides agencies to implement open data with what is called the FAIR methodology. Findable, Accessible, Interoperable, and Reusable.



What Open Data Accomplishes

So now we are calling open data FAIR data. But why is it important, what are the benefits, and how should Reclamation move forward with this legal precedent?

Here's what open data can do: Data that is richly described, easy to discover, procedurally accessible, and



interoperable with other computerized systems and machine-readable is the best possible asset for informed decision-making. Senior leaders make decisions every day and will do so with or without all available data; but data—linked to from system to system, expertly analyzed with novel applications, and presented visually to a decision-maker's fingertips in a timely manner—would certainly help them to make the best decision possible. The presence of linked, interoperable data can be used to create information dashboards, and interactive visualizations that may allow Reclamation to reduce the number and effects of data calls, and perhaps one day eliminate the need for annual reports.

Finally, at the level of the scientist, engineer, or program lead, FAIR and open data will expand the knowledge base. Experts in their field will be able to better communicate with colleagues and their ever-widening data-centric community of practice. Problems can be solved faster, and more accurately through the ability to use multiple, interoperable data assets, that are well described in metadata.

The Reclamation Information Sharing Environment (RISE) is well positioned to support our bureau in achieving the goals of the Federal Data Strategy, and to work toward meeting the requirements of the OPEN Government Data Act. In addition to this, RISE will continue to have a central role in building data communities, contributing to standards development, and the development of tools and applications for data access and usage.

Data Options to Meet Your Needs

RISE offers numerous data types and themes to meet the needs of Reclamation users wanting to publish data and internal and external data consumers

Themes describe the topical areas of data in RISE:

- Water: Water resources, including water operations
- Water Quality: Chemical, physical, biological, and radiological characteristics of water
- Biological: Animals, plants, or other living organisms
- Environmental: Characteristics of the natural environment
- Infrastructure and Assets: Physical and organizational structures and facilities
- Hydropower: Generation of hydroelectricity

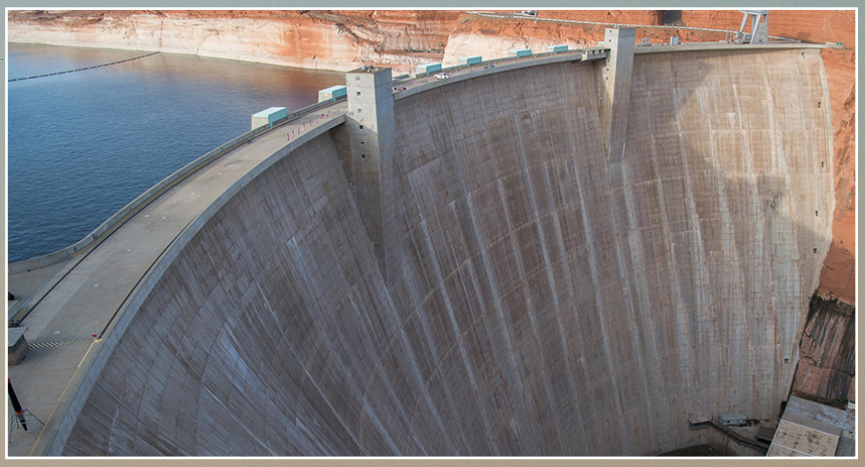
Data types include:

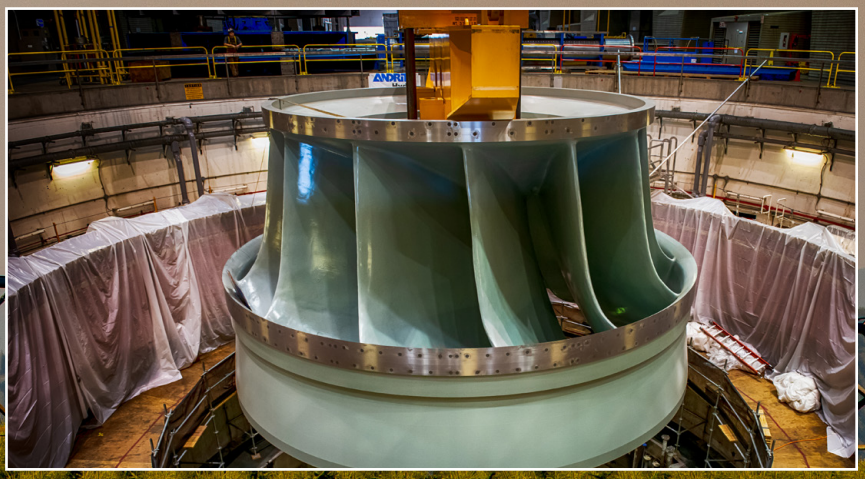
- Time Series: Data that consists of a series of values and corresponding time stamps and can be plotted in a line graph. Time series data may be observed or modeled, and time stamps may occur at regular intervals or be irregularly spaced. Time series datasets are characterized by their location and the parameter that the value represents.

- Geospatial: Data that have a spatial structure and can be displayed on a map. A geospatial layer can include one or more features and its associated attributes. The features may represent locations, objects, phenomena, observations, or model results.

- Other: A flexible category of data that can include a wide variety of files uploaded to RISE and available for download. Filetypes include documents/reports, spreadsheets, images and videos, model files, zipped archives, and more.

https://www.usbr.gov/research/ks.html





Data at Your Fingertips: RISE Features for Data Access

The RISE portal is a one-stop-shop with multiple ways for users to access data:

Catalog Search https://data.usbr.gov/catalog

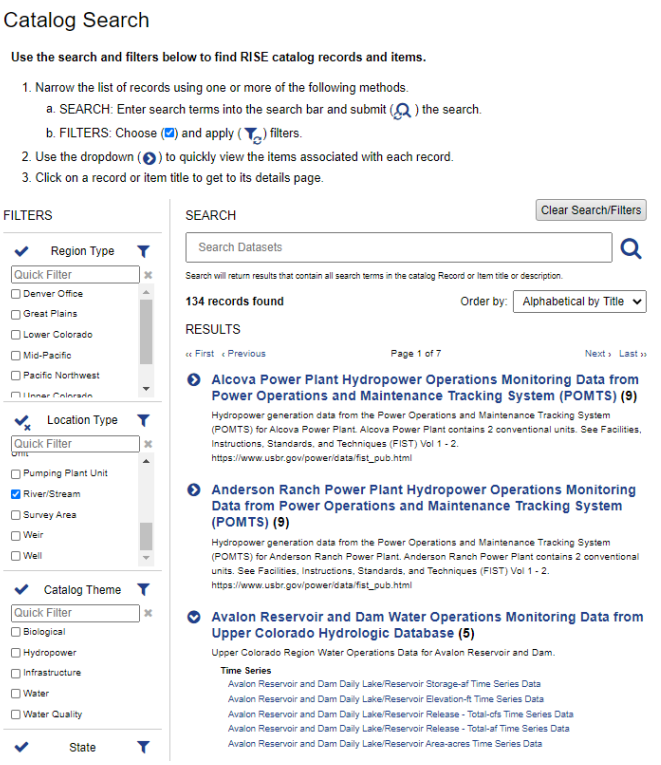
The catalog lists all the datasets that are available in RISE. It is a great place to start looking for data or browse around to see what is available. The catalog is organized into "Records" and "Items". Records are groupings of data that pertain to a particular location and come from the same source. Items are specific datasets within a record, and may contain time series, geospatial, or other data types. When you first enter the catalog, you will see a list of every record in RISE. Use the filters and search bar to narrow the list of records, then click on a record to see information about it or use the dropdown arrow to see the items within that record. Clicking on an item will bring you to the item details page, where you can see more information about the item and link to the map, query tool, or data download

Map Interface https://data.usbr.gov/rise-map/

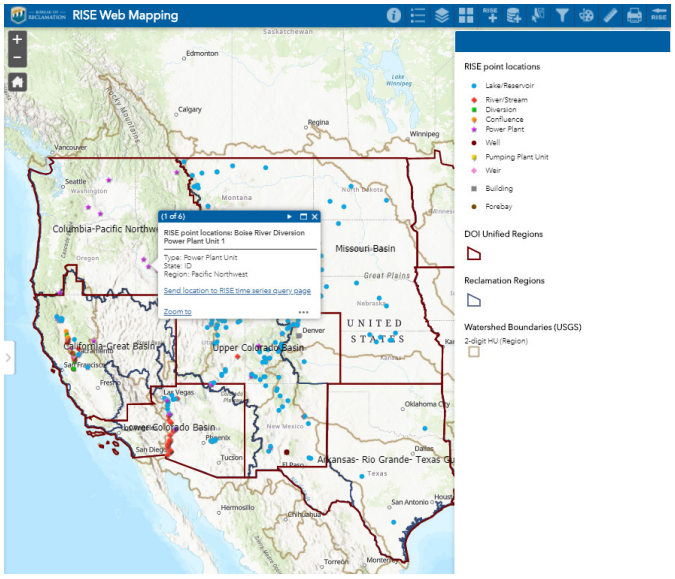
The map interface can be used in two ways: (1) to display the locations of all RISE datasets and (2) to interact directly with RISE geospatial datasets.

When you first enter the map, the interface displays a view that is useful for finding locations of RISE datasets. RISE has three types of locations: point locations, line locations, and area locations. The initial view shows various RISE point locations along with some reference layers. You can turn on additional layers for RISE line and area locations if desired. Widgets allow you to filter the locations to find ones you are interested in and select one or more locations. Clicking on a location brings up a popup with additional options, such as sending the location to the time series query page if it has time series data associated with it.

To interact with a geospatial dataset, you can add it to the map using the 'Add RISE Geospatial Layers' widget. Once it has been added, you will see the points, lines, and/or polygons associated with the layer and be able to click on the features of the layer to see data attributes or view an attribute table. Widgets can be used to further interact with the data, including measuring distances and areas, adding annotations, and printing the map.



Screenshot of RISE catalog search results showing a filter option selected and one of the catalog records expanded to show its underlying items.



Screenshot of the RISE map showing the RISE point locations layer with one point selected to show the popup with a link to send the location to the RISE query page.

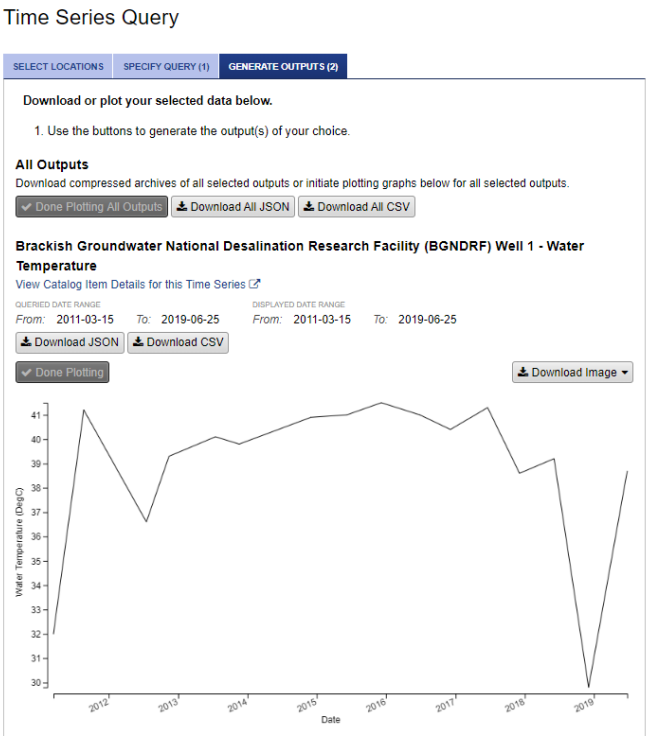
Time Series Data Query

https://data.usbr.gov/time-series/search?v=1

The time series data query tool allows you to build customized requests for time series data. Using the tool, you can select and retrieve time series data for specific locations, parameters, and time periods of interest. If the data you want comes from a model, the query tool also allows you to select which model runs to query. To perform a query, begin by selecting the location(s) associated with the data. Then, for each available time series parameter (and model run, if applicable) that you want to retrieve, input a date range. The output screen lets you quickly see the data in an on-screen plot and then easily download and use the data in your own analysis outside of RISE in machine-readable CSV or JSON files.

API https://data.usbr.gov/rise/api

The RISE Application Programming Interface (API) enables automated retrieval of information from the RISE catalog and time series databases. After setting up an API call, you can easily incorporate RISE data into interactive apps and visualizations or perform recurring analyses without having to manually retrieve data each time. Queries are formatted as URLs. The RISE API page shows which fields are available and can help you build API calls.



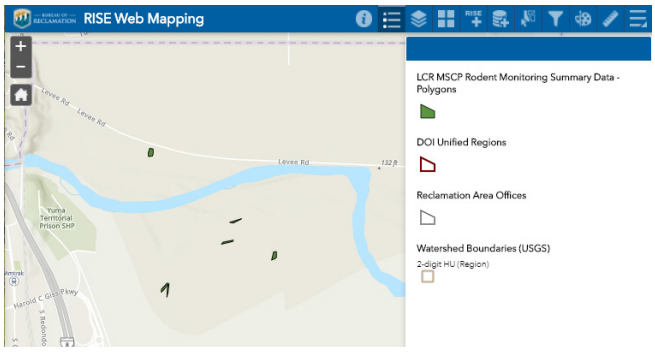
Screenshot of the RISE time series query output page showing a line graph for a timeseries of water temperature from Well 1 at the Brackish Groundwater National Desalination Research Facility.

Test Your Skills!

Send your answer to data@usbr.gov to find out if you are correct and/or let us know about any issues you encounter with the RISE website.

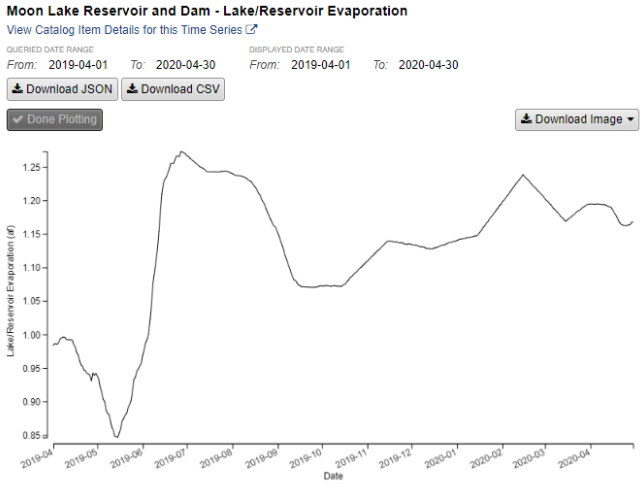
Which subspecies of cotton rat has been observed near Yuma, Arizona according to the Lower Colorado River Multi-Species Conservation Program rodent monitoring data?

Hint: Add the LCR MSCP Rodent Monitoring Summary Data to the map interface and check the features near Yuma for the subspecies described here.



How much water evaporated from Moon Lake Reservoir on April 1st, 2020?

Hint: Use the query tool to query daily reservoir evaporation for Moon Lake Reservoir for a date range that includes April 1, 2020. Download a csv file to see the value on April 1st.



Go Live! How to Publish Your Data in RISE

Any Reclamation mission-related data can be published in RISE. In general, our goal is to publish data that serves at least one of the following purposes:

- Describes past, current, or future conditions of infrastructure and resources managed by Reclamation
- Increases transparency in decision-making
- Enables replication of analytical or modeled results
- Supports stakeholder engagement with Reclamation's mission
- Supports safe and effective management and use of resources
- Encourages enhanced understanding of resources managed by Reclamation
- Enables private-sector innovation

There are four steps for publishing data in RISE:

1. PLAN: Plan for data publication by identifying which data to publish and how it will be structured in the RISE catalog.

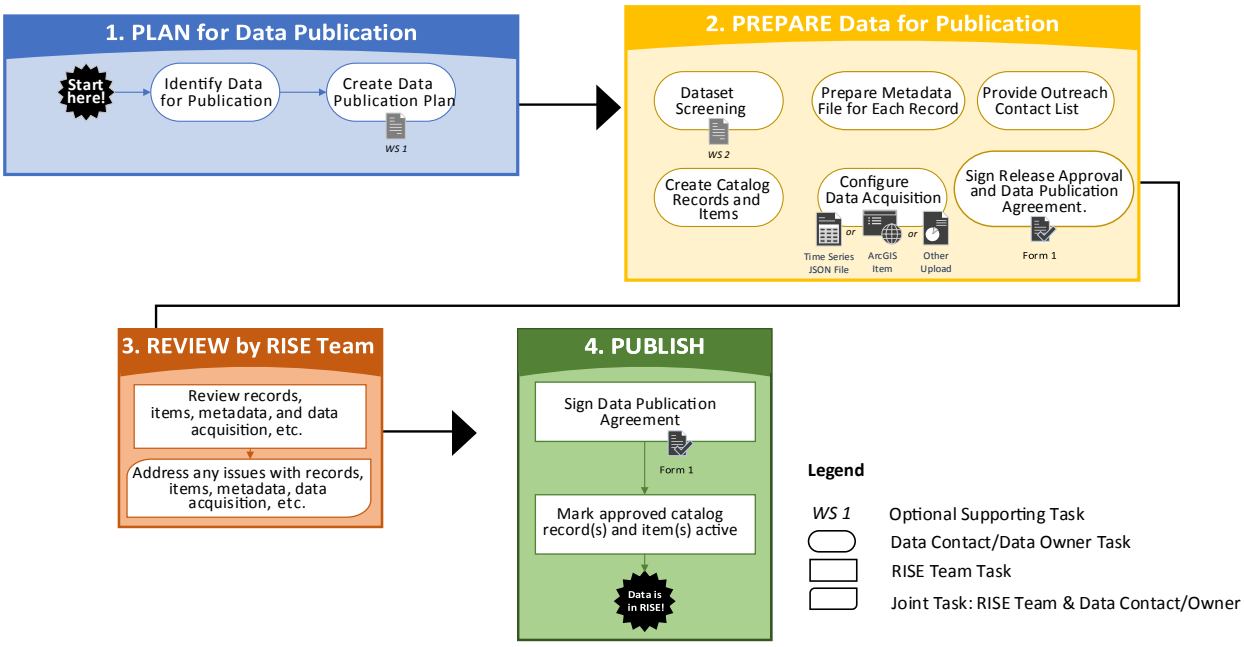
2. PREPARE: Prepare data for publication by defining RISE catalog records, items, and metadata and uploading the data or initiating data ingestion into the RISE database. This step also includes performing dataset screening to identify and mitigate any concerns about the suitability of the data for publication, such as security or privacy risks and signing a release approval document.

3. REVIEW: The RISE team reviews the configuration of the data for publication

4. PUBLISH: Data publication begins

The RISE team works with anyone who wants to publish data to help them complete these steps. To make the process easier, there are tools available to use, including worksheets for planning and performing screening, and a Data Administration User Interface, which provides a form-based interface for setting up catalog records and items. Depending on the complexity of the dataset, publication can take anywhere from a couple hours (a small, simple dataset managed by an individual) to a few weeks (a large and complex dataset that requires coordination among staff to complete the publication steps). The RISE team will be by your side every step of the way to guide and assist.

Get Started! Reach out to the RISE team to let us know you'd like to publish data. We'll get you set up with all the tools and information you need and guide you through the process.



Leveraging Data in RISE: Custom Visualizations

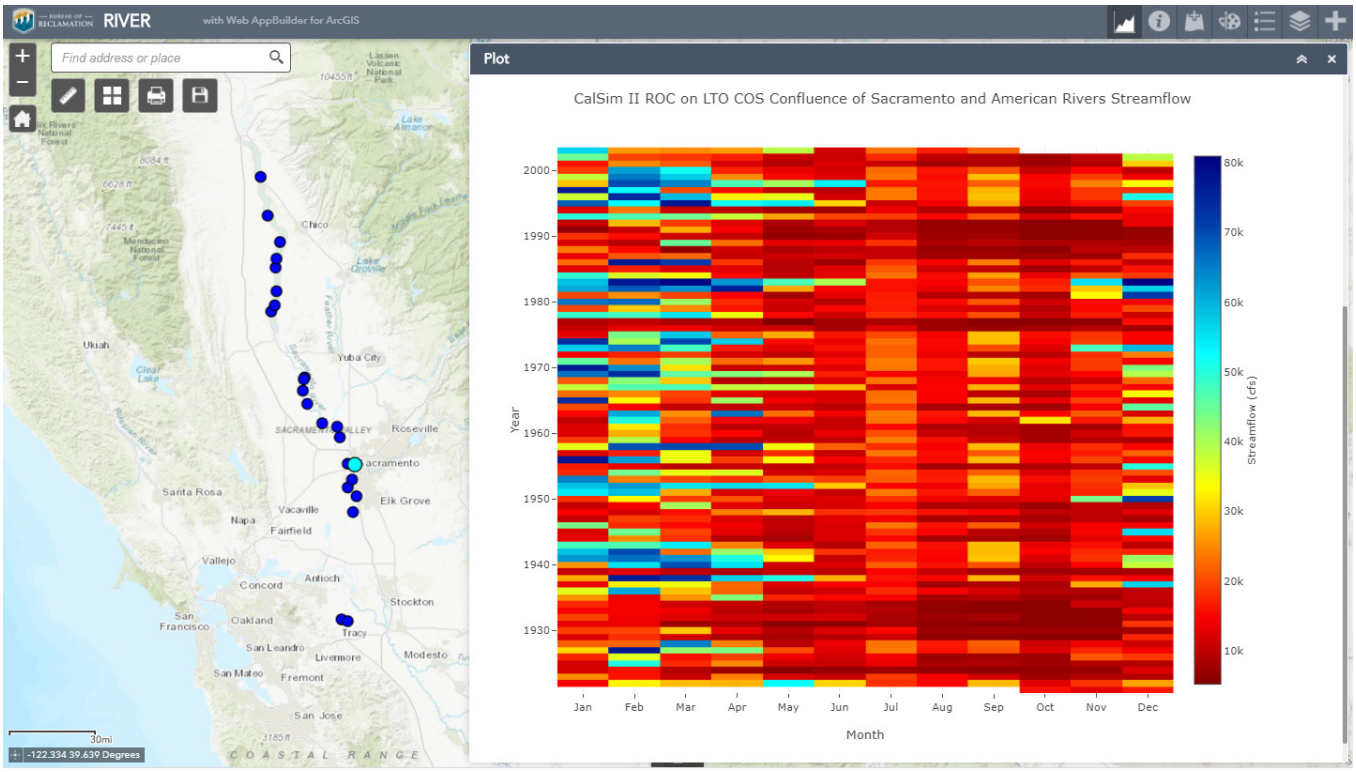
Publishing data in RISE promotes their utility by increasing discoverability and accessibility. This is further enhanced by effective visualizations that contextualize the data for specific uses. RISE offers the ability to generate static plots of data published as time series, which facilitates cursory data exploration and orientation. However, data and their use can be complex, often requiring nuanced presentation, integration of multiple data sets, and dynamic functionality.

Recognizing this, RISE includes server resources available for hosting custom visualizations that leverage data published in RISE. This visualization server aims to maximize the utility of RISE for Reclamation data stewards and programs by allowing them to focus on how to present their data, rather than procuring the IT resources to do so. Currently, the server is configured to support applications developed in Shiny, an R package that makes it easy to build interactive web applications using the R programming language, and Javascript. Other Linux compatible environments can likely also be supported in response to demand. Governance and guidance for hosting applications on the RISE visualization server is currently being finalized.

The first use of this visualization resource will be the Reclamation Interactive Visualization and Exploration Resource (RIVER; S&T project #19221 https://www.usbr.gov/research/projects/detail.cfm?id=19221). RIVER is a new web application to visualize water modeling results, using data hosted in RISE. The goal of RIVER is for anyone who uses modeled data – including managers, stakeholders, and modelers – to be able to view the data in a dynamic, interactive, format, without requiring any specialized software knowledge.

Users can navigate through modeled time series data stored in RISE through RIVER's user interface and load data into the application. They can then view the location of each time series on a map and create various interactive plots and charts, including bar charts, box plots, scatter plots, raster plots, and tables. They can also download both the charts and the underlying data, and navigate back to RISE to view associated metadata.

RIVER will be hosted on the RISE visualization server, with public launch tentatively planned for October 2020. New features will continue to be added throughout FY 2021.



A screenshot of RIVER, showing a raster plot. The locations of all loaded time series are displayed on a map, with the location of the selected time series highlighted.

What's Next for RISE

The current version of RISE is only the beginning. We have lots more datasets and features planned for the future!

Upcoming Features

• User Accounts – A new component of the RISE website that will let users (both internal and public) create an account through Login.gov to save searchers, queries and preferences

• Improved Data Administration User Interface – New features include the ability to sign on automatically with your active directory credentials, upload data publication documentation, save draft submissions, and automatically create metadata files.

• Model run member plotting in the query tool

• Additional filtering options for catalog search

Upcoming Datasets

• CGB and CPN water operations data

• Invasive mussels data

• Colorado River sediment and water quality data

• Water quality data from LCB, UCB, and CGB

• S&T project data

We want to hear from you about what you want to see in RISE! Send us an email at data@usbr.gov to suggest enhancements to the website or tell us what data you want.

Featured Face: Allison Odell

Allison Odell is Reclamation's Open Water Data Coordinator. Allison joined the Research and Development Office in March 2020, after working on open data projects funded through the S&T Program for more than five years. In her role as the Open Water Data Coordinator, Allison is responsible for overseeing the Reclamation Information Sharing Environment (RISE) system, coordinating with Reclamation staff on open data-related projects, and developing partnerships with other Federal agencies and outside organizations on open data efforts.

Prior to joining R&D, Allison was a Civil Engineer (Hydrologic) in Reclamation's Lower Colorado Basin Region Engineering Services Office and, before that, in the Lahontan Basin Area Office. At ESO, she managed water resources planning and hydrology projects including the San Diego Basin Study. At LBAO, her work focused on operations of the Newlands Project, including monthly operations modeling, canal flow monitoring, and irrigated acreage analysis. Allison also served as LCB's Value Program Coordinator starting in 2018 and facilitated numerous Value Engineering studies.

Allison studied Environmental Engineering at Cornell University, receiving her BS in 2010. She then went on to study Water Resources Engineering at Oregon State University, graduating with her MS in 2013. Her Masters thesis focused on the potential need for modifications to reservoir operations of a flood control reservoir in Oregon due to climate change.

Outside of work, Allison enjoys running, hiking, camping, traveling, and cooking. She and her husband Chris recently welcomed their first child, daughter Everly, in June 2020. They also have two dogs who love sharing adventures or curling up on the couch for a movie. Allison and her family are in the process of relocating to Denver from Henderson, NV. They look forward to exploring Colorado's landscapes, and will welcome a little snow after a hot Vegas summer.

Also Happening in Data

Streamflow Forecast Rodeo Prize Competition

Streamflow forecasts are integral to water management. Higher skill forecasts better equip water managers to operate facilities for high flows, mitigate impacts of drought, and achieve other improved outcomes like hydropower generation. Reclamation aims to spur innovation on streamflow forecasting by running the Streamflow Forecast Rodeo. This year-long prize competition starts October 1, 2020.

Prize competitions are a form of crowdsourcing. A specific question or requirement is posed and almost anyone can participate by submitting solutions that address or meet the requirements. The solutions that best meet requirements are awarded prizes. This paradigm rewards performance and leverages the ingenuity of the public to tackle challenging topics in the areas of infrastructure, water availability, and environment.

The Streamflow Forecast Rodeo competition seeks to improve the skill of short-term streamflow forecasts (10 days) via a year-long competition. Participants will develop and implement their methods for locations across the western United States with the goal of outperforming state-of-practice streamflow forecasts. With this approach, Reclamation aims to spur innovation using data science communities and Artificial Intelligence (AI)/Machine Learning (ML) methods toward enhancing streamflow forecasts.

Prior to the start of the year-long portion of the competition, teams were provided the opportunity to participate in a "pre-season" to build and refine their forecasting systems. This pre-season helps generate interest in the real-time competition and better positions teams to compete in the year-long competition. $500,000 in prizes are available between the "pre-season" and year-long portions of the competition.

Reclamation has partnered with the Centre for Energy Advancement through Technological Innovation International's Hydropower Operations and Planning Interest Group (CEATI HOPIG), NASA Tournament Lab, and Topcoder on this competition. Partnering with CEATI HOPIG includes a companion project that is evaluating existing streamflow forecasting technologies at locations across North America, including the locations specific to this prize competition. The CEATI effort will provide the benchmark forecasts that prize competition participations will be evaluated against by RTI International. Other contributors include Department of Energy's Water Power Technologies Office and the Tennessee Valley Authority.

Prize Competition Forecast Locations

Rio Puerco near Bernardo, NM

James River above Arrowwood Lake near Kensal, ND

Trinity River by Coffee Creek near Trinity Center, CA

North Fork Red River near Carter, OK

Big Thompson River above Lake Estes, CO

North Fork American River at North Fork Dam, CA

Green River near La Barge, WY

Prairie Dog Creek above Keith Sebilius Lake, KS

Taylor River above Taylor Park, CO

North Fork Shoshone near Wapiti, WY

Reclamation Data Council

The Reclamation Data Council provides programmatic and data-specific subject matter expertise to support decision making and coordinates Reclamation enterprise data management activities. The Data Council develops policy and guidance and promotes data stewardship and best practices to efficiently and effectively support the Bureau of Reclamation mission through its programs, systems, and projects. Reclamation develops and disseminates data through numerous programs and projects in fulfilment of its mission and initiatives that continue to increase in volume and scope. The goals of the Data Council are as follows: to develop and promote data governance, to reduce potential redundancy of activity; to increase efficiency in the use of information and dissemination systems to support customers and stakeholders; and to maintain compliance with the Department of the Interior, Office of Management and Budget (OMB), and statutory requirements, including the Foundations of Evidence Based Policy Making Act. The Data Council will coordinate on making data findable, accessible, interoperable and reusable both internally and externally.

The Reclamation Data Council works under the authority of the Chief Engineer and is co-chaired by Kenneth Richard, karichard@usbr.gov and Jim Nagode, jbnagode@usbr.gov

Open Water Data Beyond Reclamation

by Allison Odell

Reclamation is just one of many agencies and organizations working towards making water and related data more accessible. Below, we highlight one effort to create an "Internet of Water" by building a framework for integrating water data. Others doing work in this space include federal agencies (USGS, EPA), states including California and New Mexico that have passed open data laws, the Consortium of Universities for the Advancement of Hydrologic Science (CUAHSI), and the Western States Water Council. Reclamation is coordinating with these and others to make sure our efforts align with the practices of the broader open data community.

Internet of Water

by Peter Colohan
Executive Director, Internet of Water, Nicholas Institute for Environmental Policy Solutions, peter.colohan@duke.edu

The Internet of Water (IoW) project was established in 2018 at Duke University's Nicholas Institute for Environmental Policy Solutions on the premise that efforts to improve water management outcomes in the United States routinely stumble on the problem of inaccessible water data and information. Decision makers around the country often have difficulty discovering, accessing or integrating the data required to proactively plan for and manage an emerging water crisis or address routine, high-value questions for water stakeholders.

"The data needed to answer these questions often exist, although collected by multiple agencies across different scales of government and non-government organizations for different purposes. Since data are scattered across multiple platforms with different standards, much of it cannot be re-used beyond the primary purpose for which it was collected ..." Internet of Water: Sharing and Integrating Water Data for Sustainability, A Report from the Aspen Institute Dialogue Series on Water Data, 2017, p. VII

To address these issues and empower a 21st-century approach to water data management, the IoW seeks to advance the transformation and modernization of water data infrastructure in the United States. The IoW team at Duke works in close partnership with the Consortium of Universities for the Advancement of Hydrologic Services Inc.(CUAHSI), the Water Data Exchange (WaDE) of the Western States Water Council, the U.S. Geological Survey (USGS), the U.S Environmental Protection Agency (EPA), and a growing list of state agencies and local utilities in California, New Mexico, North Carolina and Texas. Together, the IoW community assists state and local agencies with low-cost technologies for sharing and integrating water data implemented throughout a broad network and demonstrates the advantage of those technologies for improving water outcomes in communities at risk.

Knowledge Stream

Research and Development Office