Automating Access to Password Protected ERDDAP Datasets

Madison Richardson[3] madison.richardson@noaa.gov
Dale Robinson[2] dale.robinson@noaa.gov
Cara Wilson[1] cara.wilson@noaa.gov
Sunny Hospital[2] sun.bak-hospital@noaa.gov

ERDDAP Showcase
January 2026

[1]Southwest Fisheries Science Center, NOAA
[2]UCSC Institute of Marine Sciences | Fisheries Collaborative Program | Cooperative Institute for Marine, Earth, and Atmospheric Systems
[3]Miami TSPi

What We'll Cover

- Why Access Control Matters in ERDDAP

- The CoastWatch slide archive problem

- Using ERDDAP for secure dataset workflows

- Automating login and data access with Selenium WebDriver

- From query to PowerPoint: the end-to-end workflow

Access Control in ERDDAP Workflows

Typical ERDDAP Usage

- Publicly distributed datasets
- Environmental and oceanographic data
- Designed for open discovery and reuse
- Anonymous access (no login required)

When Access Must Be Restricted

- Some datasets contain:
  - Personally identifiable information (PII)
  - Proprietary data
- These datasets require login via web interface

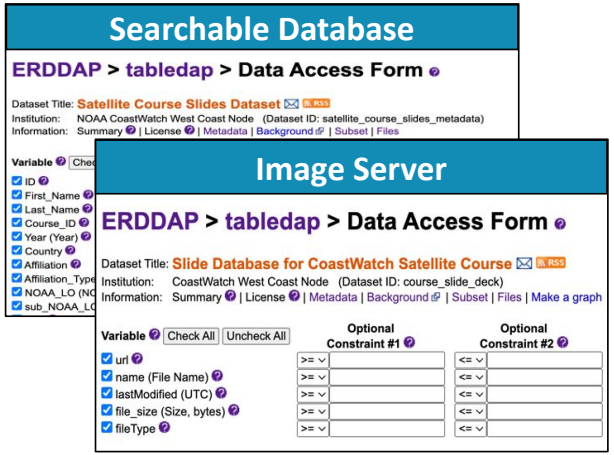
Manual login works for interactive use, but presents a problem for automated workflows.

What We Showcase in This Project

1. **Integrating Authentication into an Automated Workflow**

   - Incorporate ERDDAP login into an automated workflow
   - Authentication handled once, securely with Selenium WebDriver

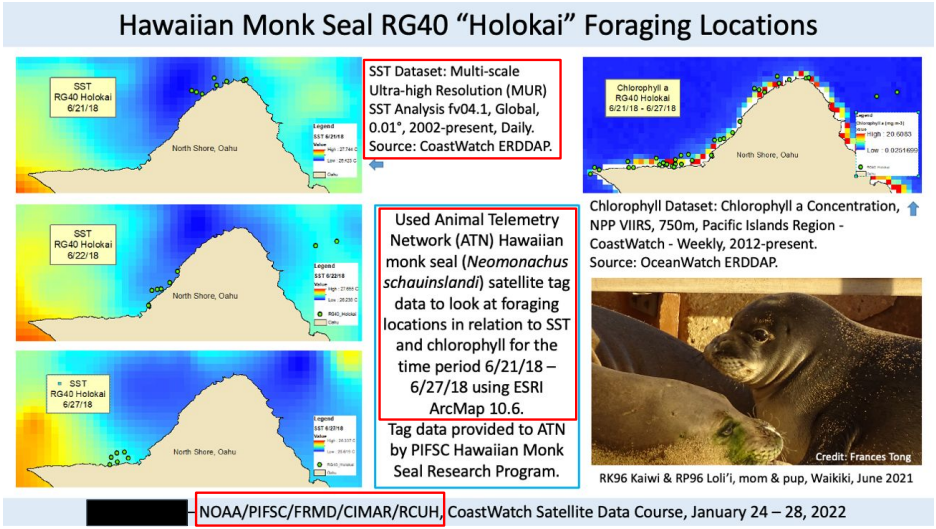
2. **A Non-Traditional Use of ERDDAP:**
   - As a searchable database
   - As an image server



CoastWatch Satellite Training Course Overview

- CoastWatch offers satellite data training workshops each year
- Participants submit a project summary slide
- Slides capture satellite use and participant/project metadata
- Slides contain PII → access is restricted
- **Goal**: Allow authorized users to generate presentations from an automated workflow

Example Participant Project Slide



Over 250 slides have been collected since 2013 (and growing!)

The Participant & Slide Archive: Then vs. Now

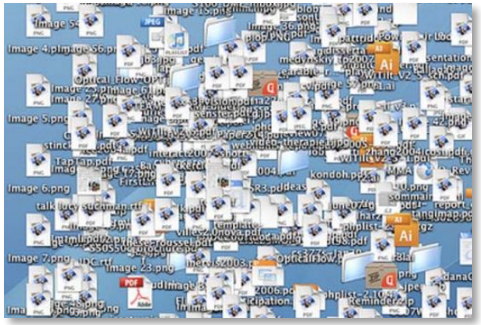
Then

- Slides scattered across folders and drives
- Inconsistently named folders and files
- Manual slide discovery (and what does Cara remember?)

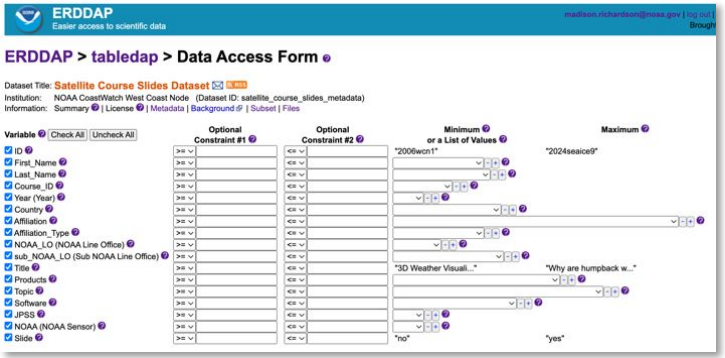
Now

- All slides and metadata hosted in ERDDAP
- Fully searchable by affiliation, product, topics, etc.
- Information on participants found instantly

Before



After



Server-Side Configuration of Datasets for Secure Access

ERDDAP is configured to use Google-based authentication for user logins.

Inside setup.xml:

```
<authentication>google</authentication>

<googleClientID>
  {token}.apps.googleusercontent.com
</googleClientID>
```

Results:

- ERDDAP adds a “Log in” link to its web interface
- Users authenticate using NOAA-authorized Google accounts

Who Can Access the Datasets?

Once authenticated, users are granted access to datasets based on assigned roles.

Inside datasets.xml:

```
<user username="john.doe@noaa.gov"
      roles="coastSlideUser" />

<accessibleTo>coastSlideUser</accessibleTo>
```

Results:

- Only users with this role can query metadata
- Only authorized users can retrieve slide images

Where the ERDDAP Data Comes From

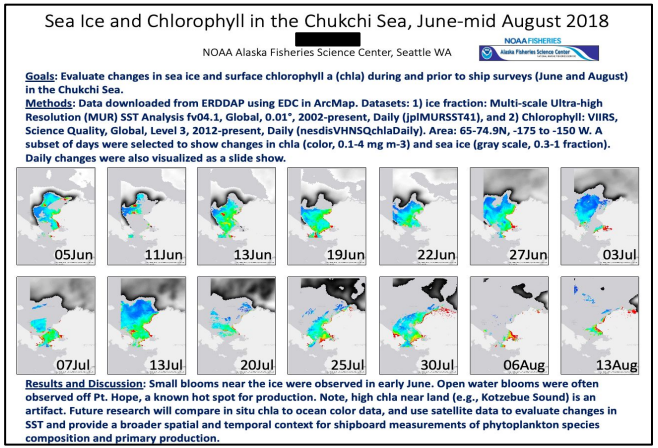
Participant metadata (CSV)

- Table with satellite course participants' information such as:
  - Course, Year, Affiliation, Products, SlideID, etc.

ID	First Name	Last Name	CourseID	Year
2006wcn1	John	Doe	2006wcn	2006
2006wcn2	Jane	Doe	2006wcn	2006

Participant Slide PNGs

- Over 250 individual slide PNGs, each given an ID:
  - [Year][Course][Number].png



How the ERDDAP Dataset XMLs Were Set Up

Participant Metadata Dataset

Searchable table of participant metadata

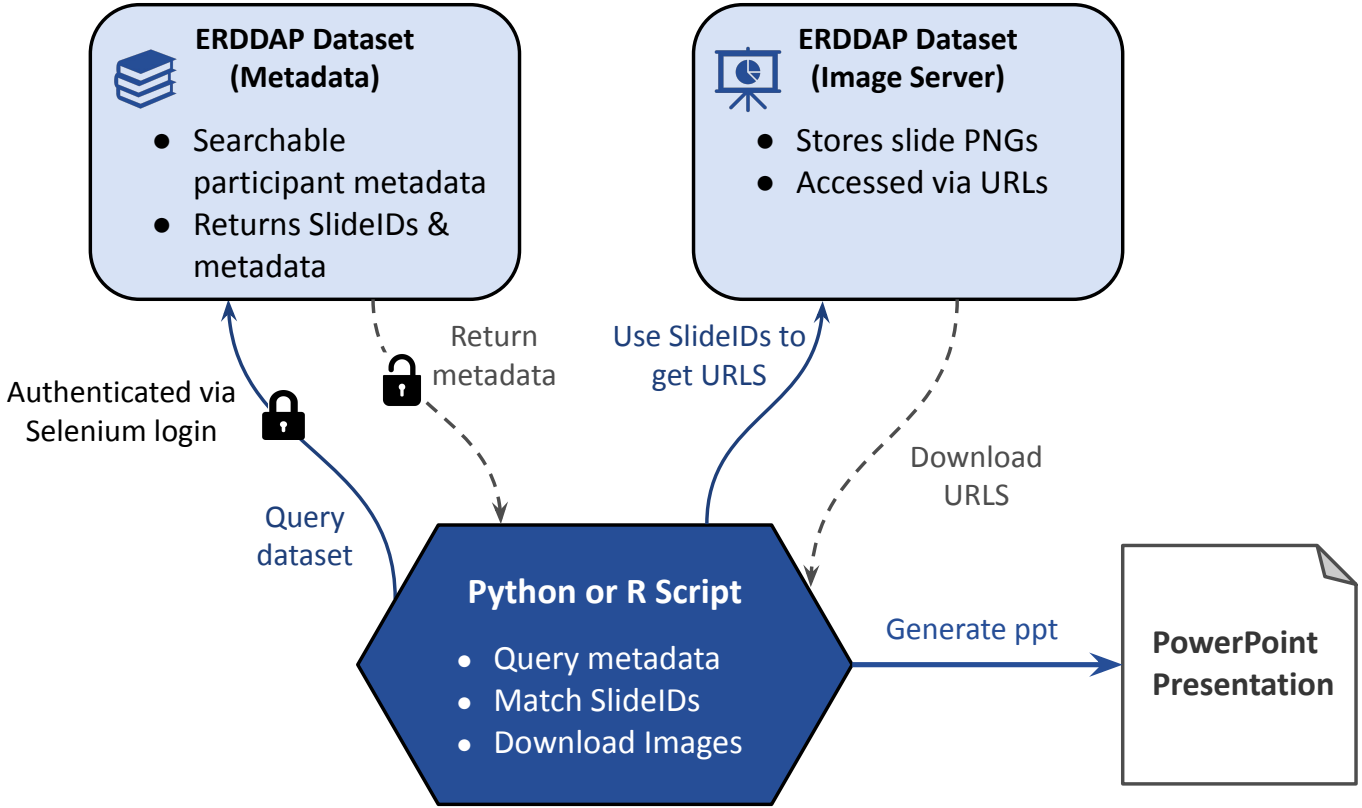
- **ID:** satellite_course_slides_metadata
- **Built with:** EDDTableFromAsciiFiles
- **Source:** Local CSV file
- **Columns:** Name, Affiliation, Course ID, Products, Year, SlideID, etc.
- **Access:** Limited to coastSlideUser

Slide Image Dataset

Linked archive of slide images

- **ID:** course_slide_deck
- **Built with:** EDDTableFromFileNames
- **Source:** Folder of PNG slides
- **Auto-indexed fields:** file name, URL, size, type, and modified date
- **Access:** Limited to coastSlideUser

Secure Automated ERDDAP Workflow



Example: NMFS Participants who used JPSS Data

File type: (more information)

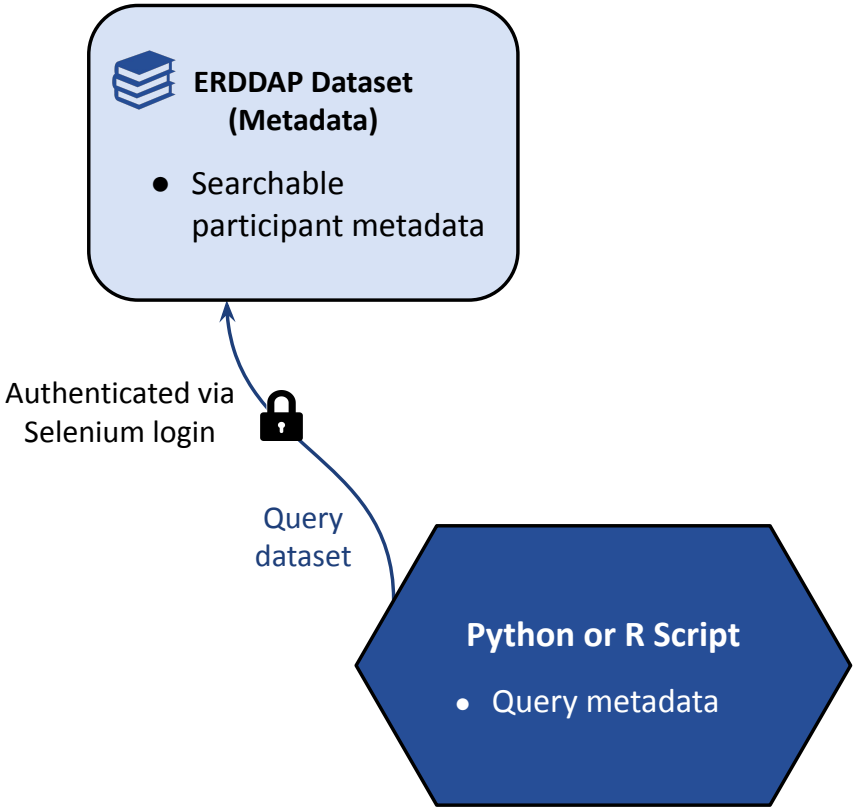
.csv - Download a ISO-8859-1 comma-separated text table (line 1: names; line 2: units; ISO 8601 times).

Just generate the URL: https://coastwatch.pfeg.noaa.gov/wcn/erddap/tabledap/satellite_cour

(Documentation / Bypass this form)

Submit (Please be patient. It may take a while to get the data.)

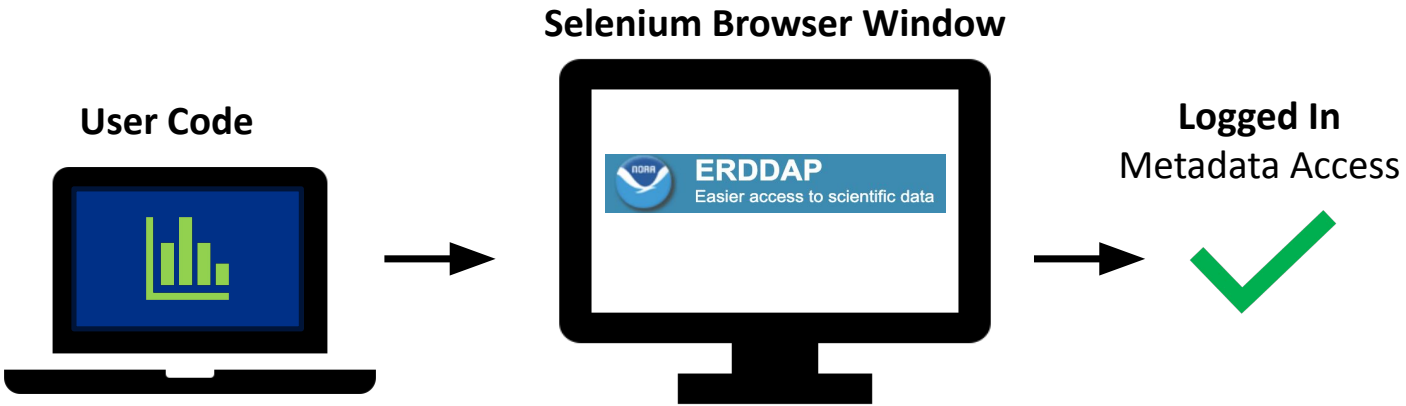
Accessing Password Protected Datasets



- Script initiates workflow
- Establishes an authenticated ERDDAP session via Selenium login
- Constructs an ERDDAP tabledap query URL
- Applies user-defined search criteria
  - Ex: ID, Year, Affiliation, etc.

Selenium WebDriver: Automating Access to ERDDAP

- **Selenium WebDriver** is an open-source software tool that automates browser actions
- In the tutorials, it's used to log into ERDDAP using your NOAA Google account and retrieve cookies for secure access to the password protected datasets



Double Login

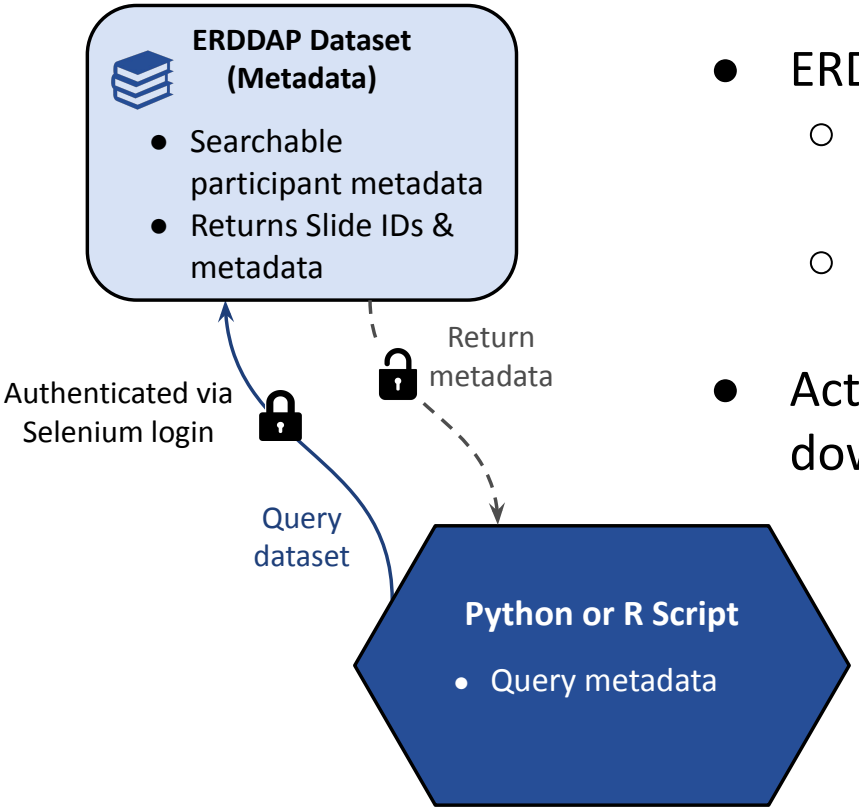
ERDDAP > log in

This ERDDAP is configured to let you log in with your Google-managed email address (including @noaa.gov). This allows ERDDAP to verify your identity (name and email address) and access your profile image, but does not give ERDDAP access to your Google account information, emails, Google Drive, or any other private information. Anyone with a Google email account can log in but the administrator of this ERDDAP determines who gets access to additional datasets.

You are logged in as **madison.richardson@noaa.gov** . (log out)

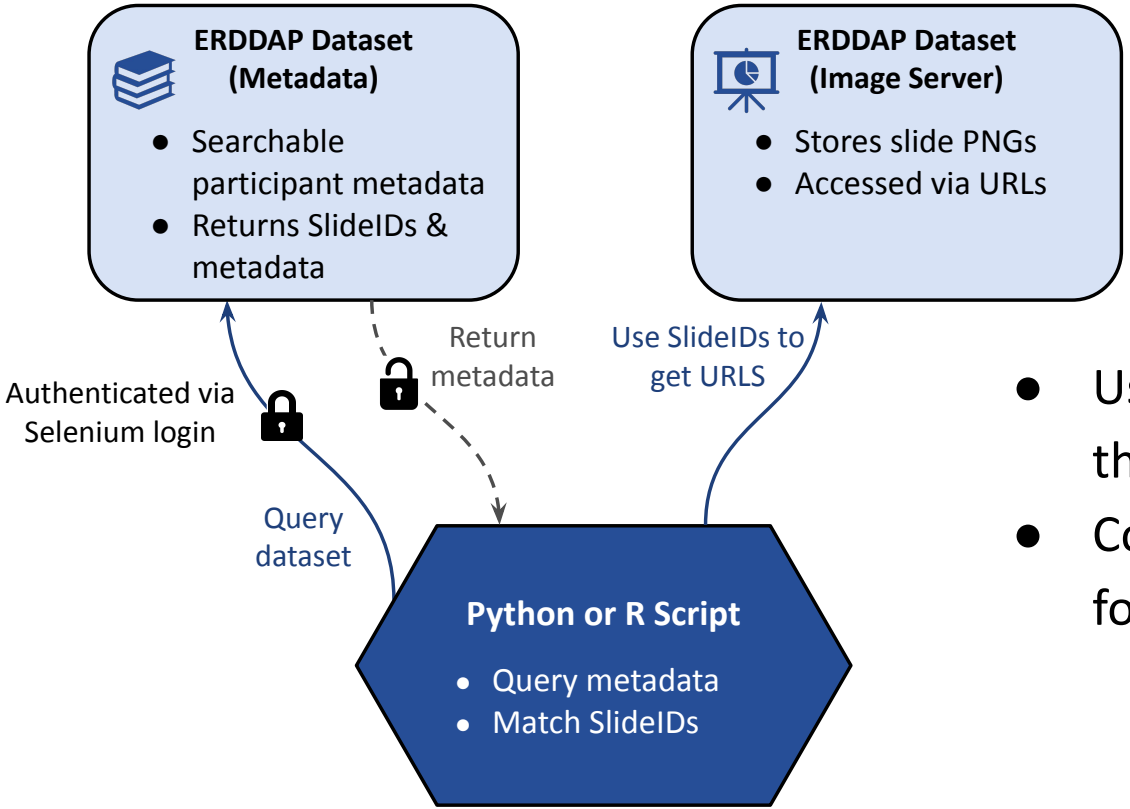
Don't use your browser's `Back` button. Use the "ERDDAP" link above, then use other links to go to ERDDAP pages you are interested in. If a cached web page says you aren't logged in, reload the page.

Retrieving Participant Metadata



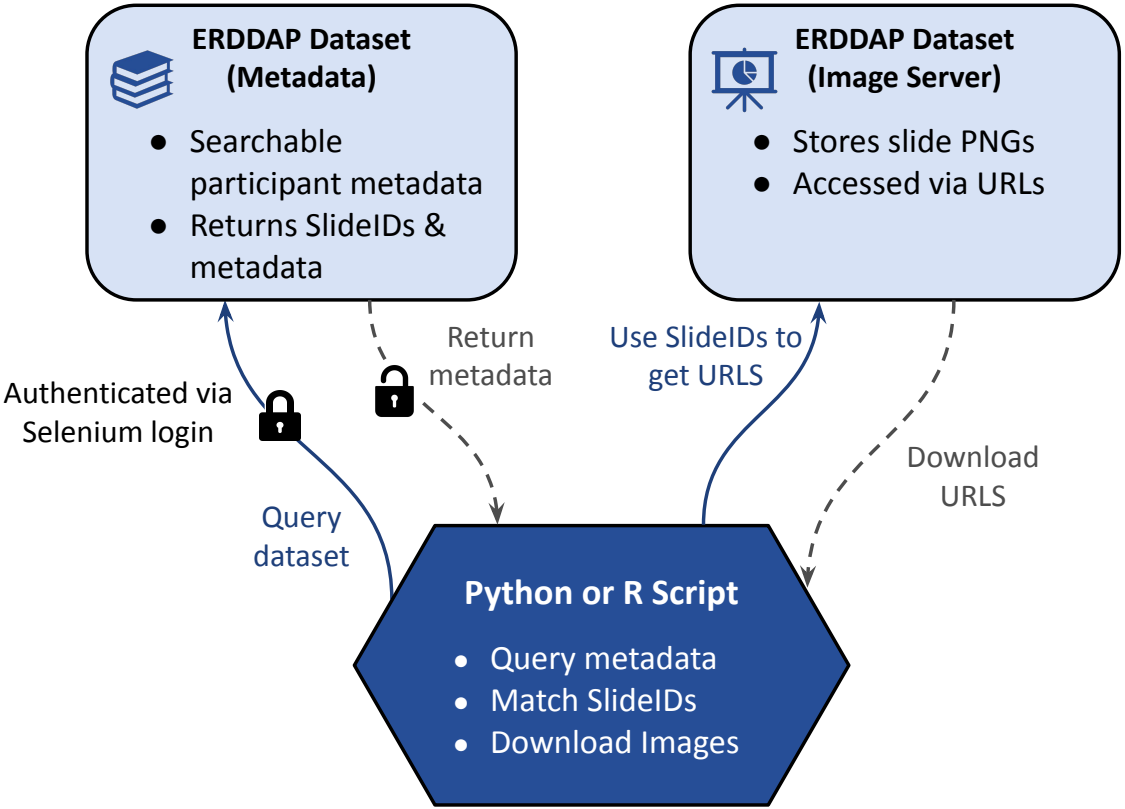
- ERDDAP returns:
  - Returns a CSV with metadata that matches user-defined criteria
  - Slide IDs used in the next step of the workflow
- Acts as the index for which slides to download

Query Participant Slide PNGs



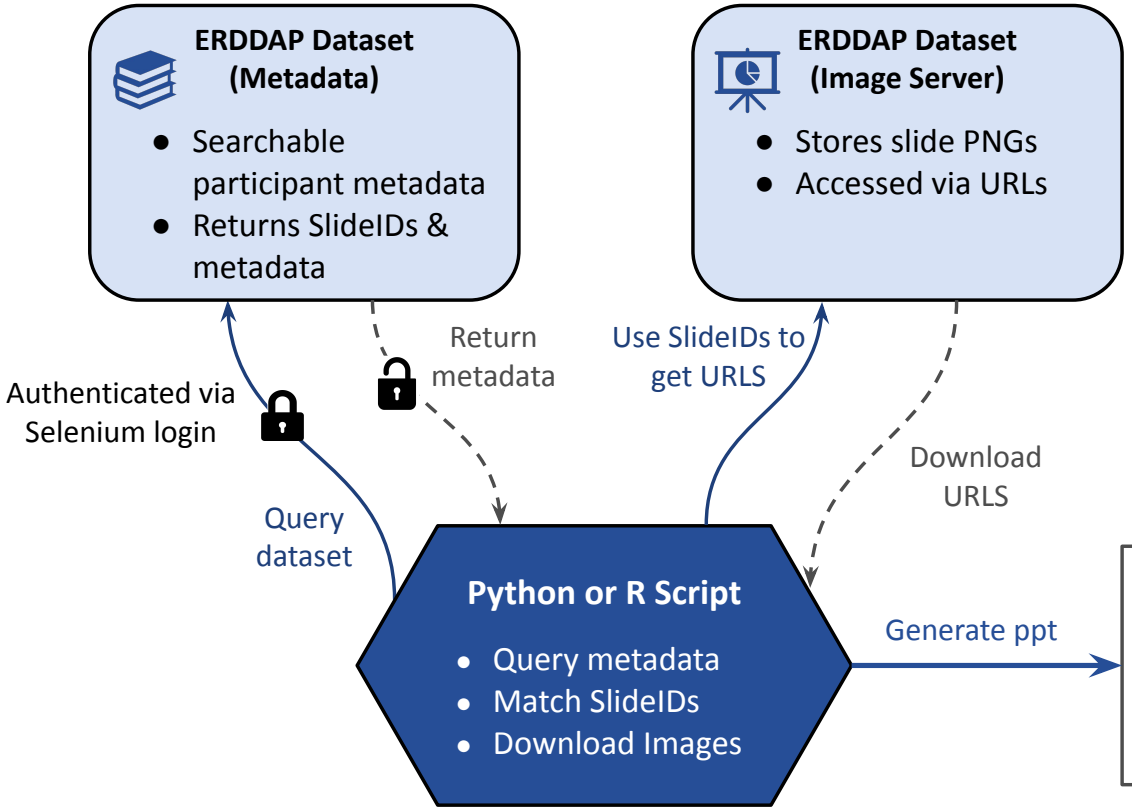
- Uses Slide IDs returned from the metadata query
- Constructs ERDDAP file URLs for each matching slide

Download Participant Slide PNGs



- Accesses slide PNGs through the authenticated ERDDAP session
- Downloads the PNGs to be included in PowerPoint

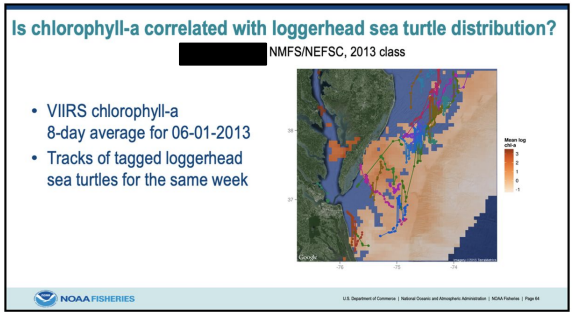
Generate Custom PowerPoint



- Uses the downloaded slide PNGs and associated metadata
- Automatically creates PowerPoint slides
- Populates slide titles and annotations using metadata fields

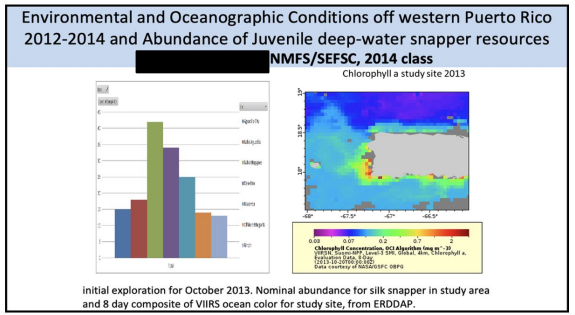
End Product: From Query to Presentation

Topic: Telemetry, Animal Distribution



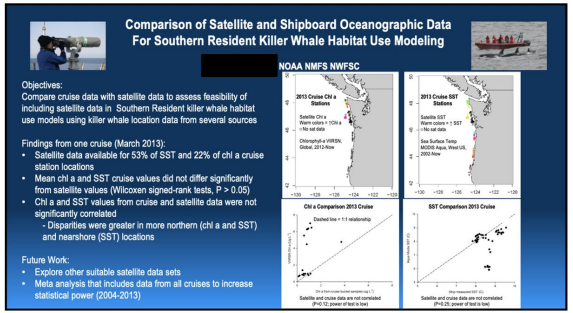
Products: Ocean Color

Topic: Fish Habitat



Products: Ocean Color

Topic: Animal Distribution



Products: Ocean Color, SST

Each slide combines metadata and participant images into a ready-to-use presentation.

Strengths and Limitations of This ERDDAP Approach

Strengths

- Leverages existing ERDDAP infrastructure
- Secure: uses Selenium + NOAA Google login, no stored passwords
- Demonstrates ERDDAP's versatility

Limitations

- Requires browser-based authentication with Selenium WebDriver due to PII-protected data
- Login step still requires user interaction
- Sensitive to ERDDAP structure changes or dataset renaming

Why This Matters

Improve Access and Efficiency

- Replaces manual slide hunting with a fast, repeatable, secure workflow

Enhance Course Impact

- Living archive of participant work
- Real-world satellite applications over time

Promotes Secure Reproducible Practices

- Uses external identity-based authentication (e.g., Google, ORCID), no credentials in code
- Reusable pattern for other ERDDAP workflows

Acknowledgements

West Coast Node & PolarWatch

- Dale Robinson
- Cara Wilson
- Sunny Hospital

CW Central

- Veronica Lance
- Michael Soracco

SWFSC

- Roy Mendelssohn

JPSS PGRR Program

Thank you for funding the slide collection project!

All instructors, participants, and hosts of the satellite course trainings!

Questions?